

PÉCSI TUDOMÁNYEGYETEM
KÖZGAZDASÁGTUDOMÁNYI KAR

GAZDÁLKODÁSTANI DOKTORI ISKOLA

Uzsoki Máté

A piaci kockázat mérése neurális hálózattal és
statisztikai modellekkel

DOKTORI ÉRTEKEZÉS

Témavezető: Dr. Bugár Gyöngyi, PhD, habil

Pécs, 2021

Köszönetnyilvánítás

A kutatást az Innovációs és Technológiai Minisztérium Felsőoktatási Intézményi Kiválósági Programja finanszírozta, a Pécsi Tudományegyetem 4. - A hazai vállalatok szerepének növelése a nemzet újraiparosításában - tématerületi programja keretében.

Köszönettel tartozom Dr. Bugár Gyöngyinek, Dr. Kovács Balázsnak és Dr. Székely Balázsnak a szakmai irányításért és az értékes észrevételekért. Köszönettel tartozom családomnak, támogatásuk és biztatásuk nélkül ez a disszertáció nem készülhetett volna el.

Tartalomjegyzék

1	Témafelvetés és célkitűzés	1
1.1	<i>A téma jelentősége.....</i>	<i>1</i>
1.2	<i>A disszertáció hipotézisei</i>	<i>3</i>
1.3	<i>A disszertáció szerkezete</i>	<i>3</i>
2	A kockázati mutatókkal kapcsolatos szakirodalom	7
2.1	<i>A kockázat definíciója.....</i>	<i>7</i>
2.2	<i>A pénzügyi kockázatkezelés kezdetei</i>	<i>7</i>
2.3	<i>Miért szükséges a kockázat mérése és kezelése?.....</i>	<i>9</i>
2.4	<i>Kockázatok típusai.....</i>	<i>10</i>
2.5	<i>A koherens kockázati mutatók kritériumrendszere</i>	<i>10</i>
2.5.1	<i>Monotonitás</i>	<i>11</i>
2.5.2	<i>Transzlációinvariancia</i>	<i>11</i>
2.5.3	<i>Szubadditivitás</i>	<i>11</i>
2.5.4	<i>Pozitív homogenitás.....</i>	<i>11</i>
2.6	<i>A kockázati mutatók megfelelése a koherencia-kritériumoknak</i>	<i>12</i>
2.7	<i>Variancia.....</i>	<i>12</i>
2.8	<i>Maximális visszaesés (Maximal Drawdown, MDD)</i>	<i>13</i>
2.9	<i>Kockázatotott érték (VaR)</i>	<i>15</i>
2.10	<i>Várható többletveszteség (ES)</i>	<i>17</i>
2.11	<i>A várható többletveszteség visszatesztelhetősége</i>	<i>17</i>
2.12	<i>Historikus és hipotetikus terheléses tesztek.....</i>	<i>18</i>
3	A kockázatotott érték és a várható többletveszteség mutatók megjelenése a szabályozásban és a gyakorlati alkalmazásban	20
3.1	<i>A bank és a biztosítási szektor szabályozása.....</i>	<i>20</i>
3.1.1	<i>A Bázeli Bankfelügyeleti Bizottság és a Bázeli Keretrendszer</i>	<i>20</i>
3.1.2	<i>Solvency II, az EU szabályozása a biztosítási szektorra</i>	<i>22</i>
3.1.3	<i>Swiss Solvency Test (SST)</i>	<i>22</i>

3.2	<i>OTC és tőzsdei letétszámítás</i>	23
3.2.1	Az ES-mutató használata a SEC által felügyelt elszámolóházaknál	23
3.2.2	Gyakorlati tapasztalatok: ES-alapú letétszámítás a Nodal Clear elszámolóházánál	25
3.3	<i>A Risk.net kockázatkezelési felmérése a Covid első hulláma után</i>	26
4	A statisztikai modellek bemutatása	33
4.1	<i>Várható többletveszteség parametrikus modelljei</i>	33
4.2	<i>A historikus szimuláció (HS)</i>	33
4.3	<i>Kiemelt eloszlásalapú modellek</i>	34
4.3.1	Arcsine	34
4.3.2	Burr (burr7)	35
4.3.3	Chen-eloszlás (chen)	36
4.3.4	Exponential extension (expext)	37
4.3.5	Exponential logarithmic (explog)	38
4.3.6	Exponential Poisson eloszlás (exppois)	39
4.3.7	Generalized power Weibull eloszlás (genpowerweibull).....	40
4.3.8	Kumaraswamy-eloszlás (kum)	41
4.3.9	Kumaraswamy–Burr-eloszlás (kumburr7)	42
4.3.10	Linear failure rate (LFR)	43
4.3.11	Loglog-eloszlás (loglog)	44
4.3.12	Lomax-eloszlás (lomax)	46
4.3.13	Topp–Leone-eloszlás (TL2)	47
4.3.14	Triangular-eloszlás (triangular).....	48
4.3.15	Two-sided power eloszlás (TSP)	49
4.4	<i>További eloszlásalapú modellek</i>	51
4.5	<i>Az idősorok transzformációja</i>	58
4.6	<i>GARCH (feltételes modellek)</i>	62
4.6.1	Az ARCH-folyamat	62
4.6.2	Standard GARCH-modell (SGarch)	62
4.6.3	Exponential GARCH (Egarch)	63
4.6.4	GjrGarch.....	64
4.7	<i>Quantile regression</i>	64
4.7.1	A kvantilisregresszió bemutatása	64

4.7.2	A kvantilisregresszió módszertana.....	65
4.7.3	A kvantilisregresszió bemutatása egy példán keresztül	66
4.7.4	A kvantilisregresszió, mint lineáris programozási modell.....	67
4.7.5	A várható többletveszteség számítása kvantilisregresszió segítségével	70
4.8	<i>A felhasznált R-könyvtárak.....</i>	<i>71</i>
5	A neurális hálózaton alapuló modellek	72
5.1	<i>A neurális hálózatokkal kapcsolatos szakirodalom.....</i>	<i>72</i>
5.1.1	Történeti áttekintés – a neurális hálók fejlődésének fő mérföldkövei.....	72
5.1.2	A gépi tanulás kategorizálása a tanulás módja szerint	78
5.1.3	Kategóriák a terjesztés iránya alapján	80
5.2	<i>A neurális hálózat működése.....</i>	<i>81</i>
5.3	<i>A neurális hálózat tanítása</i>	<i>83</i>
5.4	<i>Legmeredekebb lejtő algoritmus (gradient descent).....</i>	<i>83</i>
5.5	<i>Backprop-algoritmus.....</i>	<i>85</i>
5.6	<i>A backprop-algoritmus bemutatása egy példán keresztül.....</i>	<i>87</i>
5.7	<i>A legmeredekebblejtő- eljárás (gradient descent - GD) fő típusai.....</i>	<i>91</i>
5.7.1	Kötegelt GD (batch gradient descent).....	92
5.7.2	Sztocasztikus GD (stochastic gradient descent)	92
5.7.3	Miniköteg-GD (mini-batch gradient descent)	92
5.8	<i>A GD-eljárásokkal kapcsolatos problémák.....</i>	<i>93</i>
5.9	<i>GD optimalizálási algoritmusok.....</i>	<i>94</i>
5.9.1	A momentum-GD optimalizálási algoritmus	94
5.9.2	Nesterov-féle gyorsított gradiens (Nesterov accelerated gradient, NAG)	94
5.9.3	Adagrad (adaptive gradient algorithm).....	95
5.9.4	Adadelta	97
5.9.5	RMSprop (Root Mean Squared Propagation).....	98
5.9.6	Adam (Adaptive Moment Estimation)	98
5.9.7	AdaMax.....	99
5.9.8	Nadam	99
5.10	<i>Overfitting & Dropout.....</i>	<i>100</i>
5.11	<i>Aktivációsfüggvények</i>	<i>102</i>

5.12	<i>Implementációs részletek: TensorFlow & Keras</i>	104
5.13	<i>ANN-AS-modell</i>	105
5.14	<i>Modellezési hibák & korábbi iterációk</i>	107
5.14.1	Neuralnet és a threshold-hiperparaméter	108
5.14.2	Az eredmények aggregációja és a vizuális megjelenítés	108
5.14.3	Neuralnet vagy Keras/TensorFlow	109
5.14.4	Veszteség- vagy hibafüggvény: négyzetes vagy abszolút hiba	109
5.14.5	Hibadefiníció: ESR és ESR-MA	110
6	Backtesting-módszerek	111
6.1	<i>A backtesting eljárások célja</i>	111
6.2	<i>A backtesting és a pénzügyi szabályozás aktuális kérdései</i>	112
6.3	<i>Elicitabilitás</i>	112
6.4	<i>A backtesting definíciója</i>	113
6.5	<i>A különböző backtesting-eljárások bemutatása</i>	114
6.5.1	VaR-hibasám-ellenőrzés – coverage test	114
6.5.2	VaR-függetlenségiteszt	115
6.5.3	Várható többletveszteség – coverage test	117
6.6	<i>Modellvalidáció: szignifikanciateszt</i>	119
6.7	<i>A backtesting-eljárás ereje</i>	120
6.8	<i>Az ellenőrzések kiegészítése és kiterjesztése</i>	123
6.8.1	Komplex termékek, portfóliók ellenőrzése	123
6.9	<i>Ex-ante becslés, ex-post ellenőrzés</i>	124
7	Empirikus elemzés	125
7.1	<i>Adatok</i>	125
7.2	<i>Az adatok statisztikai tulajdonságai</i>	129
7.2.1	A normalitás ellenőrzése a Shapiro–Wilk- és a Kolmogorov–Smirnov-tesztekkel	129
7.2.2	A normalitás ellenőrzése vizuális módszerekkel	131
7.2.3	A ferdeség, a csúcosság és a stacionaritás ellenőrzése	135
7.3	<i>A várható többletveszteség becslési modelljeinek empirikus összehasonlítása</i>	138

7.4	<i>Az ANN-AS-modell érzékenységvizsgálata</i>	144
7.4.1	Egyenlő tanítási és tartási időszakok (TAI = TRI).....	144
7.4.2	Eltérő tanítási és tartási időszakok (TAI > TRI).....	146
7.4.3	Aktivációs függvények.....	147
7.4.4	Különböző optimalizációs algoritmusok eredményei.....	149
7.4.5	Különböző dropout százalékok eredményei	150
7.4.6	Alternatív alapmodell.....	150
7.4.7	További paraméterek	151
7.5	<i>Összefoglaló</i>	152
7.5.1	Az induló munkahipotézisek igazolása.....	153
7.5.2	További lehetséges kutatási irányok.....	153
8	Felhasznált irodalom	i

Táblázatok jegyzéke

1. táblázat: A SEC által felügyelt elszámolóházak, ahol ES-mutatót használnak	24
2. táblázat: További eloszlások (A-Be).....	51
3. táblázat: További eloszlások (Bu – Ex)	52
4. táblázat: További eloszlások (Ex – G).....	53
5. táblázat: További eloszlások (K)	54
6. táblázat: További eloszlások (L-Lo)	55
7. táblázat: További eloszlások (Lo-Ma)	56
8. táblázat: További eloszlások (Mr-S).....	57
9. táblázat: További eloszlások (T-X).....	58
10. táblázat: A location paraméter kiszámításának módjai modellenként	59
11. táblázat: A scale paraméter kiszámításának módjai modellenként.....	60
12. táblázat: A modellimplementációk szempontjából lényeges R könyvtárak	71
13. táblázat: Gyakran használt értékelőfüggvények	113
14. táblázat: A hibák elfogadható száma 250 napos időszakra.....	115
15. táblázat: A backtesting-eljárás ereje	121
16. táblázat: Az elemzésben résztvevő részvények (A-K).....	126
17. táblázat: Az elemzésben résztvevő részvények (Ko-Z).....	127
18. táblázat: A részvények szórása és átlagos hozama	128
19. táblázat: Az idősorok normalitástesztje	130
20. táblázat: A részvényhozamok ferdesége	136
21. táblázat: A részvényhozamok csúcsossága.....	137
22. táblázat: $ Z $ -értékek összesítése modellenként	139
23. táblázat: Az arcsine- és az ANN-AS-modell összehasonlítása t- és z-próbával.....	143
24. táblázat: Az azonos hosszúságú tanítási és tartási időszakok hatása	145
25. táblázat: Az eltérő hosszúságú tanítási és tartási időszakok hatása	147
26. táblázat: Különböző aktivációs függvények eredményei	148
27. táblázat: Az optimalizációs algoritmusok hatása	149
28. táblázat: A dropout százalék hiperparaméter hatása.....	150
29. táblázat: A normális eloszlásra alapozó neurális háló teljesítménye	150

Ábrák jegyzéke

1. ábra: A Markowitz-féle hatékony portfóliók.....	13
2. ábra: A Maximum Drowdown mutató számítása (példa).....	15
3. ábra: Kockázatotott érték (VaR) és a várható többletveszteség (ES)	16
4. ábra: A várható többletveszteség és a kvantilis függvény kapcsolata	17
5. ábra: A megkérdezett befektetők által használt kockázati mutatók.....	26
6. ábra: A megkérdezett vállalatok által bevezetni tervezett új módszertanok.....	27
7. ábra: A megkérdezett vállalatok által megszüntetni tervezett módszertanok.....	28
8. ábra: A visszatekintési időszak hossza a felmérésben részt vevő pénzügyi vállalatoknál.....	29
9. ábra: A hipotetikus terheléses tesztek hasznosságának megítélése a Risk.net felmérésében résztvevő vállalatok között	30
10. ábra: A VaR-mérőszám által lefedett időszak	31
11. ábra: A VaR-mérőszám konfidenciaszintje	32
20. ábra: Az arcsine-eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye standard paraméterekkel.....	35
21. ábra: A Burr-eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye standard paraméterekkel.....	36
22. ábra: A Chen-eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye standard paraméterekkel	37
23. ábra: Az expext-eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye standard paraméterekkel.....	38
24. ábra: Az explog-sűrűségfüggvény ($p=0,5$)	39
25. ábra: Exppois-sűrűségfüggvény ($b = 1$ és $\lambda= 1,2,3$)	40
26. ábra: Az $F(x)$ sűrűségfüggvény (bal oldalon) és tipikus meghibásodási ráták (jobb oldalon).....	41
27. ábra: A Kumaraswamy-eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye $a=2$ és $b=2$ paraméterekkel	42
28. ábra A Kumaraswamy–Burr-eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye standard paraméterekkel	43
29. ábra: Az LFR-eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye standard paraméterekkel.....	44

30. ábra: Az $F(x)$ sűrűségfüggvény $\lambda = 2$ esetére.....	45
31. ábra: Meghibásodási ráta $\lambda = 2$ esetére	45
32. ábra: A Lomax-eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye standard paraméterekkel.....	47
33. ábra A Topp–Leone-eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye $b = 2$ alakparaméterrel.....	47
34. ábra: A Triangular-eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye standard paraméterekkel.....	49
35. ábra: A Two-sided power eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye $a=5$ alakparaméterrel.....	50
36. ábra: A parametrikus modellek működésének folyamatábrája	61
37. ábra: Apple Inc (AAPL) napi részvényára és a kvantilisek.....	66
38. ábra: A QR-modell.....	70
12. ábra: McCulloch és Pitts egyik neurális háló-modellje	72
13. ábra: A Rochester-féle szimulált neuron és kapcsolatai	73
14. ábra: Lineárisan szeparálható halmazok	74
15. ábra: Nemlineárisan szeparálható halmazok.....	75
16. ábra: Az Adaline-modell felépítése	76
17. ábra: Raszterizált betű.....	78
18. ábra: A megerősítéses tanulás modellje.....	80
19. ábra: Visszacsatolásos neurális hálózat	81
39. ábra: A többszintű perceptron.....	82
40. ábra: Függvény és gradiense vektormezőként ábrázolva az alsó síkon.....	84
41. ábra: A csomópontok kimenetének számítása.....	86
42. ábra: Példaként bemutatott egyszerű neurális háló	88
43. ábra: A paraméterek összefüggési hierarchiája	89
44. ábra: Nyeregfelület és a legmeredekebb lejtő algoritmus.....	93
45. ábra: A momentum- és a NAG-algoritmus közötti különbség	95
46. ábra: A Dropout-eljárás bemutatása	101
47. ábra: Sigmoid-aktivációsfüggvény	102
48. ábra: Tanh-aktivációsfüggvény.....	103

49. ábra: A ReLU-aktivációsfüggvény	103
50. ábra: TensorFlow-számításigráf-példa	104
51. ábra: A példában használt t-eloszlások sűrűségfüggvényei.....	121
52. ábra: A szimulált Z-értékek sűrűségfüggvényei	122
53. ábra: A szimulált Z-értékek kumulatív eloszlásfüggvényei	122
54. ábra: A backtesting eljárás ereje a kumulatív eloszlásfüggvények alapján	123
55. ábra: A becslési, az ES és az ellenőrzési időszak hosszának kapcsolata.....	124
56. ábra: Részvényárak alakulása 1988–2017	125
57. ábra: A részvények teljesítménye átlaghozam és szórás tekintetében	129
58. ábra: A WEC-idősor normalitásának vizuális ellenőrzése (sűrűségfüggvény).....	131
59. ábra: A WEC-idősor normalitásának vizuális ellenőrzése (kvantilisek)	132
60. ábra: A LUV-idősor normalitásának vizuális ellenőrzése (sűrűségfüggvény)	132
61. ábra: A LUV idősor normalitásának vizuális ellenőrzése (kvantilisek)	133
62. ábra A COO-idősor normalitásának vizuális ellenőrzése (sűrűségfüggvény)	134
63. ábra: A COO-idősor normalitásának vizuális ellenőrzése (kvantilisek).....	134
64. ábra: Az Arcsine és az ANN-modell $ Z $ -értékeinek összehasonlítása részvényenként	140
65. ábra: Az Arcsine és az ANN-modell $ Z $ -értékeinek összehasonlítása éves bontásban.	141
66. ábra: Az Arcsine és az ANN-modell Z-értékeinek összehasonlítása éves bontásban ..	141
67. ábra: Az Arcsine és az ANN-modell Z-értékeinek összehasonlítása a dotcom-válság alatt havi bontásban	142
68. ábra: Az Arcsine és az ANN-modell Z-értékeinek összehasonlítása a 2007-2008-as válság alatt havi bontásban	142
69. ábra: Egymást átfedő hathónapos tanítási és tartási időszakok	144
70. ábra: Az azonos hosszúságú tanítási és tartási időszakok hatása.....	145
71. ábra: Egymást átfedő eltérő hosszúságú tanítási és tartási időszakok	146
72. ábra: Különböző aktivációs függvények eredményei	149
73. ábra: A normális eloszlásra alapozó neurális háló teljesítményének összehasonlítása	151

A piaci kockázat mérése neurális hálózattal és statisztikai modellekkel

Absztrakt

Jelölt: *Uzsoki Máté*

Témavezető: *Dr. Bugár Gyöngyi, PhD, habil.*

Az értekezés célja egyrészt egy olyan átfogó empirikus elemzés elvégzése és bemutatása, ami nagy adathalmazon hasonlítja össze a várható többletveszteség több, mint 50 becslési modelljének teljesítményét. A szakirodalomban számos becslési technika megtalálható, de eddig hiányzott ezek egységes keretrendszerben történő széles körű összehasonlítása. Az empirikus elemzés elvégzéséhez a modelleket 100 S&P500 indexben résztvevő részvény 30 éves idősorán számítom ki. A várható többletveszteség visszatesztelhetőségével, így a modellrangsorolási probléma feloldásával az ilyen jellegű összehasonlítás elvégzése aktuálissá vált. Az értekezésben bemutatott modell összehasonlítás során olyan visszatesztelési eljárást alkalmazok, ami nem igényli a kockázati mérőszám elicálhatóságát és lehetővé teszi az eredmények összevonását, így a modellek rangsorolását is. Az értekezés eredményei között bemutatom, hogy a klasszikus valószínűségi-eloszlások használatánál lényegesen jobb eredmény érhető el más, kevésbé ismert eloszlások alkalmazásával. Az értekezés második, hasonlóan fontos célja annak vizsgálata, hogy adaptálhatók-e előjelzési modellként az egyre több területen sikeresen alkalmazott neurális hálózatok a várható többletveszteség becslésére. Ennek vizsgálatához egy új megközelítést dolgozok ki, és ez alapján konstruálok egy neurális hálózatot. Az így létrejött és az értekezésben bemutatott többszintű neurális háló a kiválasztott statisztikai modell eredményét tovább javítva, az összes, több mint 50 vizsgált modellenél pontosabban, azaz szignifikánsan kisebb hibával jelzi előre a várható többletveszteséget.

Kulcsszavak: Kockázati modellek, várható többletveszteség, neurális hálózatok, pénzügyi modellezés

1 Témafelvetés és célkitűzés

1.1 A téma jelentősége

A megfelelő pénzügyi kockázatkezelés jelentőségét számos pénzügyi katasztrófa mellett a 2007-es pénzügyi válság is bizonyítja. A kockázatkezelés egyrészt szabályozói szempontból fontos. A kockázatkezeléssel a szabályozó célja, hogy a pénzügyi szektor ellenálló legyen a különböző sokkhatásokkal szemben. Elengedhetetlen, hogy a szereplők kellő mértékű sokkelnyelő képességgel bírjanak ahhoz, hogy akár nagyobb vállalatok bukása esetén se sérüljön a pénzügyi piac olyan mértékben, hogy az a teljes gazdaság teljesítményét kritikus mértékben visszavesse (McNeil, et al., 2015). A 2007-es válság megmutatta, hogy a pénzügyi szereplők hibás kockázatmérése és a kockázatkezelés nem megfelelő mértéke a teljes gazdaság működését visszafoghatja, ezzel súlyos károkat okozva a társadalom számos szereplőjének. Egyértelmű, hogy a pénzügyi területen megjelenő új szabályozások – mint például a Bázeli III. keretrendszer – fontos célkitűzése, hogy választ adjanak a 2007-es válság tanulságaira, így megelőzhetőek a 2007-es válsághoz hasonló krízisek kialakulását a jövőben.

A kockázatkezelés központi eleme a kockázat megfelelő mérése. Ez alapvetően két részből áll: a megfelelő kockázati mérőszám kiválasztásából, valamint a mérőszám megfelelő becslési eljárásának, azaz modelljének kidolgozásából. Kockázati mérőszámként a várható többletveszteség (expected shortfall) mutató használata preferált a szakirodalom alapján. A várható többletveszteség legfőbb előnye, hogy a gazdasági szereplők számára jól értelmezhetően adja meg a kockázat mértékét ellentétben a szórással vagy az expectilisek használatával. A kockázati mérőszám másik előnye, hogy koherens kockázati mutató (Artzner, et al., 1999), ami azt biztosítja, hogy megfelel azoknak az elvárható kritériumoknak, amelyek azt biztosítják, hogy nem lesznek olyan esetek, amikor a mutató bizonyos kockázati faktorokat nem vesz figyelembe. A koherencia tulajdonságával a várható többletveszteség maga mögé utasítja a korábban iparági standardként elterjedt kockázatot érték mutatót (value at risk, VaR). A VaR szintén jól értelmezhetően adta meg a kockázat mértékét, de a koherencia hiánya miatt előfordulhatott olyan eset, amikor két különálló pozíció kockázatának összege kisebb volt, mint az összevont pozíció kockázata. Ez az eset alapvetően mond ellent a kockázatméréssel kapcsolatos intuitív elvárásainknak, és arra a tényre vezethető vissza,

hogy a VaR mutató szempontjából „vakfolt” az eloszlás legnagyobb veszteséggel járó széle. Ezekből a hiányosságokból következik a Bazel III. keretrendszer által előírt mérőszám váltás, ahol a kockázatosított érték mutató helyét a várható többletveszteség veszi át (Bugár & Ratting, 2016).

A pénzügyi kockázatkezelés a vállalatok szempontjából is nagy jelentőséggel bír. Az értekezésben bemutatott várható többletveszteség (expected shortfall) mutatót többek között a tőzsdei és OTC elszámolóházak (SEC, 2017a) (SEC, 2017b) (SEC, 2019) (NodalClear, 2021) is használják a derivatív pozíciók letéti követelményeinek meghatározásához. Egy elszámolóház fő tevékenysége, hogy a partnerek közötti kockázatot átvállalja, majd a letéti előírások segítségével minimálisra csökkenti azokat. Belátható, hogy egy elszámolóház esetében a módszertanilag megfelelő kockázatkezelés és a kockázatok pontos becslése komoly üzleti előnyt eredményez. A várható többletveszteség mutató erre kiváló lehetőséget biztosít. Jól bizonyítja a téma aktualitását, és azt, hogy a várható többletveszteség módszertant a szakmai közösség is elismeri, hogy a Nodal Exchange amerikai egyesült államokbeli energiatőzsde elszámolását végző Nodal Clear vállalat megnyerte a 2017-es év innovációja díjat a várható többletveszteségalapú letétszámítási módszertan bevezetésével (Risk.net, 2017).

A kockázatkezelés és a kockázati mutatók jelentőségén túl szintén fontos kérdés az, hogy milyen modell alkalmazásával lehet értéküket hatékonyan és pontosan becsülni, hiszen könnyen belátható, hogy a várható többletveszteség használata önmagában nem elég, a hatékony kockázatkezeléshez a megfelelő becslési modell és annak folyamatos visszatesztelése is szükséges. A szakirodalomban számos javaslatot találunk a várható többletveszteség különböző modelljeire. A becslésekre használhatunk többek között historikus szimulációt (Righi & Ceretta, 2015), különböző valószínűségi eloszlásokon alapuló modelleket (Nadarajah, et al., 2015), kvantilis regressziót (Koenker, 2005) vagy GARCH-folyamatokat (Engle, 1982) (Righi & Ceretta, 2015).

A várható többletveszteség (expected shortfall, ES) elterjedésének fontos mérföldköve, hogy Acerbi és Székely (2017) az ES-modellek visszatesztelésére olyan eljárást tett közzé, ami a modellek validitásának ellenőrzését és rangsorolását is lehetővé teszi. Ez azért bír nagy jelentőséggel, mert a várható többletveszteség mutatóval szemben a leggyakrabban felmerülő kritika, hogy az ezen az eljárásen alapuló modellek nehezen

ellenőrizhetőek (Gneiting, 2011). Erre nyújt megoldást az Acerbi-Székely-féle backtesting eljárás. A backtesting eljárás segítségével megállapítható, hogy a rendelkezésre álló modellek közül melyik a legalkalmasabb a mutató pontos előjelzésére.

Szintén releváns kérdés, hogy a korábban leírt modellek mellett milyen új modellezési megközelítés lehet még sikeres. Ezen a területen érdekes lehet a neurális hálók módszertana, illetve annak vizsgálata, hogy az hogyan és milyen eredménnyel adaptálható a várható többletveszteség becslésének problematikájára. Az utóbbi években számos tanulmány bizonyítja, hogy a neurális hálókra alapuló modellek kimagasló eredménnyel alkalmazhatók több tudományterületen is (LeCun, et al., 2015), többek között a genetika (Leung, et al., 2014), a gyógyszerkutató (Ma, et al., 2015), valamint a beszédfelismerés (Hinton & et al., 2012) területén.

Az elért eredmények ellenére az idősorok előrejelzése már korábban is a nehezen megfogható alkalmazási területek közé tartozott. Yang és Wu (2006) is a terület tíz kihívást jelentő problémája között sorolja fel az idősorok elemzését, elsősorban az adatokat szennyező magas zajszint miatt. Långkvist és szerzőtársai (2014) kiemelik, hogy a jelenleg elért eredmények többnyire az időben statikus adatok feldolgozásában születtek, de az idősorok elemzése is egyre nagyobb figyelmet kap. A kihívások ellenére a gépi tanuláson alapuló megoldások képesek arra, hogy a pénzügyi idősorok nem-lineáris struktúráját modellezzék.

Az irodalomban a neurális hálók módszertanát pénzügyi idősorok területén többnyire a hozam (Galeshchuk, 2016) (Badics, 2014) (Fischer & Krauss, 2018), kisebb mértékben a volatilitás (Lahmiri, 2017) (Kristjanpoller, et al., 2014) előrejelzésére alkalmazzák. Ezek közül a pénzügyi befektetések kockázatának mérésére Markowitz (1952) munkája alapján adódik a variancia használata, de a várható többletveszteség becsléséről is található ilyen jellegű munka (Uzsoki, 2020).

1.2 A disszertáció hipotézisei

Az értekezés a következő munkahipotézisekkel dolgozik:

H1: Érdemes a várható többletveszteség modellezésben az eloszlásfüggvények széles körét tesztelni, mert létezhetnek olyan kevésbé ismert eloszlások, amelyek a klasszikus modellekhez képest lényegesen jobb becsléseket eredményeznek.

H2: Lehetséges olyan neurális hálózat létrehozása, amely alkalmas a várható többletveszteség kockázati mutató előrejelzésére és a többi jól teljesítő modellel összemérhető pontosságú becslést ad.

H3: Neurális hálók segítségével létrehozható olyan modell, amely pontosabb várható többletveszteség-becslést ad, mint kizárólag a statisztikai modellek használata.

H4: A neurális hálózat becslésének pontossága javítható oly módon, hogy a neurális hálózat újratanítását gyakrabban végezzük el.

1.3 A disszertáció szerkezete

Az első fejezet a téma jelentőségét mutatja be egyrészt a kockázatkezelés, a várható többletveszteség (expected shortfall) koherens kockázati mérőszám, a 2007-es válság, a Bázeli III. banki keretrendszer és az elszámolóházak szempontjából. Ezt követően a neurális hálózatok eredményeinek áttekintése következik, azzal a felvetéssel, hogy neurális hálók adaptálhatóak a várható többletveszteség becslésére. Az első fejezet második szakasza az értekezés kiinduló hipotéziseit tartalmazza, valamint a disszertáció szerkezetét.

A második fejezet a pénzügyi kockázatméréssel kapcsolatos szakirodalmat tekinti át, kezdve a kockázat pénzügyi definícióival, majd rátérve a kockázatkezelés történeti kezdeteire, a kockázatmérés szükségességére és a kockázatok típusaira. Ezt követően a fejezet áttekinti a koherens kockázati mutatók kritériumait, valamint a gyakorlati szempontból releváns variancia és a maximális visszaesés mutatókat. Ezt követően a fejezet rátér a kockázatosított érték mutató bemutatására, ennek hátrányaira, és bemutatja a kockázatosított érték hiányosságaira válaszul bevezetett várható többletveszteség kockázati mérőszámot.

A harmadik fejezet a várható többletveszteség és a kockázatosított érték mutatók szerepét mutatja be a szabályozásban és a kockázatkezelés gyakorlatában, ezzel is alátámasztva a téma aktualitását. A fejezet első szakasza a Bázeli Bankfelügyeleti Bizottság keretrendszerét (BCBS, 2021) mutatja be, valamint az Európai Unió Solvency II biztosítási szabályozását (Európai Parlament, 2009), végül pedig a svájci Swiss Solvency Test biztosítási szabályozást (Federal Office of Private Insurance, 2006). Ezt követően a harmadik fejezet leírja az amerikai egyesült államokbeli SEC által felügyelt

OTC és tőzsdei elszámolóházak ES-mutatóhoz (Rockafellar & Uryasev, 2002) kapcsolódó kockázatmérési gyakorlatát (SEC, 2017a) (SEC, 2017b) (SEC, 2019). A fejezet a Risk.net befektetői kockázatkezelési felmérése releváns eredményeinek bemutatásával zárul (Risk.net, 2020).

A negyedik fejezetben a várható többletveszteség kockázati mutató becslésére alkalmazott statisztikai modellek részletes leírása következik. Elsőként a historikus szimuláció került bemutatásra, majd a különböző valószínűségi eloszlásokon alapuló modellek következnek két részre bontva. Az értekezés elsőként részletesebben azokat a kiemelt eloszlásokat taglalja, amelyek később az eredmények alapján az összesített rangsor felső negyedébe tartoznak. Ezt követi a jelen kontextusban rosszabb teljesítményű eloszlások kompakt, táblázatba foglalt leírása a számítások reprodukálhatósága végett. A fejezet a GARCH- (Engle, 1982) és a kvantilis regresszió (Koenker & Bassett, 1978) alapú modellek leírásával zárul.

Az ötödik fejezet a neurális hálózatokkal kapcsolatos elméleti háttérrel tekint át, valamint a szerző által kifejlesztett új, neurális hálókra alapuló modellt mutatja be. A fejezet elsőként bemutatja a neurális hálózatok kialakulásának főbb mérföldköveit a kezdeti, tanulásra képtelen MCP-neuronmodelltől (McCulloch & Pitts, 1943), a terület nagy áttöréseinek ismertetésén keresztül a jelenleg is használt modern megoldásokig. Az ötödik fejezet ezt követően a modern neurális hálók főbb működési mechanizmusait, illetve a neurális hálók kategorizálásához alkalmazott szempontrendszert írja le. Az ötödik fejezet részletesen leírja a neurális hálók módszertanának központi elemét képező tanítási folyamatot, valamint a modellben hiperparaméterként megjelenő optimalizációs algoritmusok működését. Ezt követően a fejezet bemutatja a túltanulás megakadályozására alkalmazott dropout, azaz a csomópontok véletlenszerű kihagyásának módszertanát, a szintén hiperparaméterként megadott aktivációs függvényeket, valamint a neurális hálók implementációjánál alkalmazott programkönyvtárakat. A fejezet végül a konkrét neurálisháló-modell leírásával zárul.

A hatodik fejezet a modellek összehasonlítására használt backtesting módszertant ismerteti. Az eljárás az értekezés szempontjából különösen nagy jelentőséggel bír, mivel ez teszi lehetővé a modellek rangsorolását, valamint az eredmények aggregációját. A fejezet bemutatja az elicitabilitás fogalmát. Ez a tulajdonság egyes tanulmányok szerint a

várható többletvesztés modellek nehézkes visszatesztelhetőségét okozza. Ezt követően a fejezet ismerteti a különböző backtesting eljárásokat, azok kialakulását egészen a VaR-alapú (Jorion, 2007) visszatesztelési módszerektől kezdve. Ezek után az értekezésben használt, az elicítabilitás tulajdonságának hiánya ellenére a modellek összehasonlítására és validálására alkalmas visszatesztelési eljárás kerül bemutatásra. A fejezet a visszatesztelési eljárások különböző kiterjesztési lehetőségeinek áttekintésével zárul.

A hetedik fejezet az empirikus elemzést, a neurális háló modell-hiperparamétereinek érzékenységvizsgálatát, valamint az értekezés eredményeit és következtetéseit foglalja össze. A fejezet elsőként az adatokat és azok statisztikai tulajdonságait írja le: az adatok normalitás-ellenőrzésének eredményét, az eloszlások ferdesége és csúcossága tesztelésének, valamint az idősorok stacionaritás-ellenőrzésének eredményét. Ezután a korábban bemutatott várható többletvesztésmodellek összesített rangsora következik, melyet a két legjobb modell eredményeinek részletes elemzése követ. A fejezet részletesen leírja a hiperparaméterek megválasztásának hatását a modell eredményességére. A fejezetben vizsgált hiperparaméterek: a belső és a kimeneti aktivációs függvények, a becslési és a tesztelési időszak hossza, az optimalizációs eljárás megválasztása és a dropout módszertan során használt csomópontelhagyás valószínűsége. A fejezet az értekezés eredményeivel és következtetéseivel zárul.

2 A kockázati mutatókkal kapcsolatos szakirodalom

2.1 A kockázat definíciója

A kockázat kifejezést számos különböző kontextusban használjuk, így egyetlen definícióval nem foglalható össze. A lenti definíciók elsősorban a fogalom pénzügyi, gazdasági aspektusaira koncentrálnak.

Révai nagy lexikona (1911, p. 775) a következő definíciót adja:

„A vállalkozásnak az az eleme, mely a vállalkozónak cselekvő hatáskörén kívül esik, amelynek eredményes vagy káros volta rajta kívül álló eshetőségektől függ.”

McNeil és társai a banki és biztosítási szektorra fókuszálva a következő két definíciót használják (McNeil, et al., 2015, p. 1):

„Egy esemény vagy cselekedet, ami hátráltatja a szervezetet céljai elérésében és stratégiájának megvalósításában.”

„A veszteség, vagy az elvárttól alacsonyabb hozam előfordulásának mérhető valószínűsége.”

A pénzügyi és gazdasági definíciók közös eleme a bizonytalanság és a kedvezőtlen esemény bekövetkezésének lehetősége (Bugár, 2015). Ez utóbbi elem fontosságát jól mutatja, hogy intuitívan a negatív kimenetekre gondolunk kockázatként és a vártnál pozitívabb eredményekre nem. Szintén nem nevezhetünk kockázatosnak egy olyan helyzetet, ahol teljesen biztos egy bizonyos negatív kimenetel megvalósulása.

2.2 A pénzügyi kockázatkezelés kezdetei

A pénzügyi kockázatkezelés kezdetben különféle derivatívák alkalmazásával történt. Az egyik legkorábbi kapcsolódó írásos emlék i.e. 18. századból származik. Hammurabi kőtábláinak 48. bekezdése egy opciós szerződésnek értelmezhető konstrukcióról szól, mely lehetőséget ad arra, hogy az adós a termés elmaradása esetén ne fizessen kamatot a hitelező részére (Dunbar, 2001).

„Ha valaki tartozik és egy vihar letarolja a gabonát, vagy megghiúsul az aratás, vagy a gabona nem nő a szárazság miatt, akkor abban az évben nem kell gabonát fizetnie a hitelezőjének.”

/Hammurabi törvényei/¹

Egy 17. századból származó későbbi forrás már az opciók használatát konkrétan megemlíti. Joseph de la Vega (De La Vega, 1688) holland kereskedő *Confusion de Confusiones* című könyve az amszterdami tőzsde működését írja le.

„Az 'opsies' magyarázataként: a prémium megfizetésével az egyik fél értéket ad át azért, hogy megvédje a készletét vagy profitot szerezzen. Az ember jótékony körülmények között úgy használja, mint a vitorlát egy örömteli utazáson, de horgonyként szolgál a viharban.

/De La Vega/²

De La Vega leírásának érdekessége, hogy az opciókra horgonyként hivatkozik, tehát itt az opció a kockázatkezelés eszköze és nem a spekulációé. Annak ellenére, hogy a derivatívokon alapuló kockázatkezelés nagy múltra tekint vissza, a kockázatkezelés tudományos módszertana jóval később, a 20. században alakult ki. Az 1950-es évek előtt a befektetések értékelése tipikusan a hozamuk alapján történt (Dunbar, 2001).

Markowitz (1952) úttörő munkájának eredménye, hogy a befektetések értékelésében a kockázat egy új szempontként jelent meg. Markowitz a portfóliók hatékonyságát a hozam-kockázat síkon határozta meg, ahol a kockázati mérőszám a hozamidősorok varianciája volt. Markowitz munkáját követően számos kockázatkezelési módszertan látott napilágot, többek között a CAPM-modell (Sharpe, 1964), vagy a Sharpe-mutató (Sharpe, 1966), de a kockázatkezelés területe azóta is folyamatosan fejlődik.

¹ (Dunbar, 2001, p. 23), fordítás a szerzőtől

² (McNeil, et al., 2015, p. 5), fordítás a szerzőtől

2.3 *Miért szükséges a kockázat mérése és kezelése?*

A kockázattal kapcsolatos természetes igény a kockázat mérése, mellyel a szereplők célja az, hogy a kockázatot bizonyos keretek között tartsák, kezeljék. Az ilyen jellegű kockázatkerülésnek két alapvető oka van. Egyrészt előfordulhat, hogy a gazdasági szereplő próbálja elkerülni a működését hátráltató negatív eseményeket, így természetéből adódóan kockázatkerülő. A másik alapvető ok, hogy külső szereplő írja elő a kockázatkerülő magatartást, például a szervezetre vonatkozó törvényi szabályozás kialakításával (Csóka, 2003). A kockázatkerülés, illetve a kockázatkezelés szükségességének a két alapvető oka számos további elvárásból és szempontból eredeztethető. A banki és biztosítási szektorra fókuszálva a következő szereplők szempontjait emelhetjük ki (McNeil, et al., 2015):

- A társadalom szempontjából lényeges elvárás, hogy a bankok és a biztosítók megbízhatóan működjenek, és a szolgáltatásaikat folyamatosan tudják nyújtani az ügyfelek felé. Egy bank vagy egy biztosító esetleges összeomlása az ügyfeleknek is komoly károkat okoz, hiszen ezek a pénzügyi vállalatok az alapvető gazdasági infrastruktúra részét képezik.
- Szabályozói és politikai szempontból mindenképpen kerülendő, hogy az egyes bankok vagy biztosítók bukása a piacok erős összekötöttsége miatt áterjedjen más szervezetekre vagy akár a teljes pénzügyi piacra, ezáltal a 2008-as válsághoz hasonló gazdasági recessziót idézve elő.
- A részvényesek elsődleges igénye, hogy a vállalat értéke növekedjen. Széles körben elterjedt nézet, hogy ehhez nagyban hozzájárul a kockázatok megfelelő kezelése. A vállalatok ezért gyakran olyan helyzetekben is komoly energiát fordítanak a kockázatok kezelésére, amikor ezt külön szabályozás nem írja elő számukra. McNeil és szerzőtársai (2015) szerint a kockázatkezelés többek között csökkenti az adóköltiséget a bevételek ingadozásának csökkentésével, könnyebbé és olcsóbbá teszi a tőke bevonását, csökkenti a csőd valószínűségét és a külső finanszírozás költségét.

2.4 Kockázatok típusai

Egy vállalat működése során a kockázatnak számos forrása lehet. A stabil működés biztosításának érdekében szükséges összes jelentős kockázatforrás, kockázattípus azonosítása, valamint a kockázatok kezelése. A pénzügyi kockázatok főbb típusai a következők (McNeil, et al., 2015):

- Piaci kockázat: egy pénzügyi pozíció vagy portfólió értékváltozásának kockázata, amelyet az azt felépítő komponensek árváltozásai okoznak, például a részvények vagy kötvények árváltozása. Az értekezés során tárgyalt kockázat legtöbb esetben ebbe a kategóriába esik.
- Hitelkockázat: a bankok azon kockázata, hogy az adós fizetéképtelenné válik és ezért nem tudja visszafizetni a hitelt és a kamatot lejáratkor.
- Működési kockázat: olyan veszteségek előfordulásának kockázata, amelyeket a működés során a helytelenül működő belső folyamatok vagy rendszerek, hibákat elkövető munkatársak vagy negatív külső események okoznak.
- Piaci likviditási kockázat: annak veszélye, hogy egy adott befektetést nem lehet elég gyorsan eladni vagy megvenni annak érdekében, hogy megakadályozható vagy minimalizálható legyen a veszteség.

2.5 A koherens kockázati mutatók kritériumrendszere

A koherens kockázati mutatók tulajdonságait Artzner és társai (1999) definiálták. Céljuk az volt, hogy a kockázati mutatók kiválasztását egy logikus elvárásokra épített axiómarendszer alapozza meg. A követelményrendszer axiómáinak eredménye, hogy a koherens kockázati mutató alkalmas számos gyakorlati kockázatkezelési cél ellátására, többek között banki tőkekövetelmények megadására, elszámolóházak létmeghatározására vagy a befektetési vállalatok kockázati limitjeinek kiszámítására. Artzner és társai (1999) a következő négy elvárt tulajdonságot, azaz axiómát fogalmazták meg a kockázati mutatókkal szemben.

2.5.1 Monotonitás

A monotonitási axióma kimondja, hogy ha egy pozíció minden esetben magasabb veszteségekhez vezet, mint egy másik, akkor az magasabb tőkekövetelménnyel jár.

$$\text{Ha } X \leq Y, \text{ akkor } \rho(X) \leq \rho(Y) \quad (1)$$

A fentiekben X és Y véletlen változók, amelyek az általuk jelölt pozíciók végső nettó értékét mutatják meg a jövő minden lehetséges állapotában, pénzösszegként megadva. ρ egy kockázati mutató, $\rho(X)$ pedig az X pozíció ρ mutatóval mért kockázata. $X \leq Y$ azt jelöli, hogy $X(\omega) \leq Y(\omega)$ a jövő minden lehetséges állapotában (ω). $\rho(X)$ a kockázatot pénzösszegben fejezi ki.

2.5.2 Transzlációinvariancia

A transzlációinvariancia axiómája azt jelenti, hogy amennyiben egy kockázatmentes befektetéssel egészítjük ki a pozíciót, úgy a kiegészített pozíció tőkekövetelménye ugyanakkora összeggel csökken.

$$\rho(X + \alpha \cdot r) = \rho(X) - \alpha \quad (2)$$

A fentiekben r a kockázatmentes befektetés hozamrátája, α pedig a kockázatmentes befektetés mértéke, pénzösszegként megadva.

2.5.3 Szubadditivitás

A szubadditivitás axiómája kimondja, hogy a diverzifikáció sosem növelheti a pozíció kockázatát. Megfordítva: két pozíció összevonása nem eredményezhet nagyobb tőkekövetelményt, mint a korábbi két külön pozíció tőkekövetelményeinek összege.

$$\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2) \quad (3)$$

2.5.4 Pozitív homogenitás

A pozitív homogenitás axiómája azt jelenti, hogy egy pozíció méretének növelése a tőkekövetelmény egyenesen arányos növekedéséhez vezet.

$$\rho(\lambda X) = \lambda \rho(X), \lambda \geq 0 \quad (4)$$

A fentiekben λ a pozíció növelésének arányát adja meg.

2.6 A kockázati mutatók megfelelése a koherencia-kritériumoknak

Azokat a kockázati mutatókat, amelyek a fenti axiómák mindegyikét teljesítik koherens kockázati mutatóknak hívjuk. A következő szakaszban bemutatott gyakorlati szempontból releváns kockázati mutatók közül egyedül a várható többletveszteség (expected shortfall) koherens mutató. A variancia és a maximális visszaesés relatív kockázati mutatók, azaz nem függenek a befektetés mértékétől, így nem teljesíthetik a pozitív homogenitás követelményét, ami azt írja elő, hogy a kockázat egyenesen arányosan növekedjen a befektetés méretének növekedésével. A kockázatotott érték (VaR) a szubadditivitás axiómáját nem teljesíti.

2.7 Variancia

A variancia az első, kockázatmérésre alkalmas mutató, amivel Markowitz (1952) korszakalkotó munkájában bevezette a portfóliók új értékelési szempontjaként a kockázat fogalmát a korábban is használt várható hozam mellé. Markowitz munkáját a tudományos kockázatkezelési módszertan kialakulásának első mérföldkövének tekinthetjük, ezt megelőzően a kockázat kezelésére csak szerződéses megoldások álltak a vállalkozások rendelkezésére.

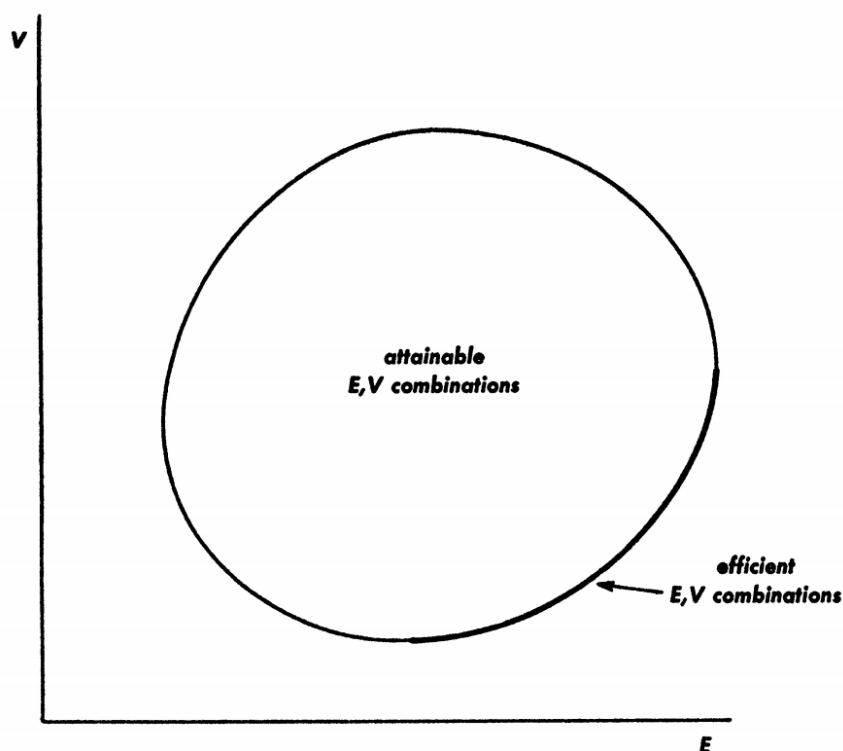
A kockázat és hozam kettős szempontrendszer alapján Markowitz definiálta a hatékony portfóliók fogalmát. Hatékony portfóliónak azokat a befektetéseket nevezhetjük, amelyeknél nincs magasabb hozamot (E) ígérő befektetés adott kockázati szinten (V). A hatékony portfóliók az 1. ábrán a körív vastagabb, jobb alsó részén találhatóak. Markowitz diszkrét esetben a következő képlettel adja meg a várható értéket és a varianciát:

$$E = p_1 y_1 + p_2 y_2 + \dots + p_N y_N \quad (5)$$

$$V = p_1 (y_1 - E)^2 + p_2 (y_2 - E)^2 + \dots + p_N (y_N - E)^2 \quad (6)$$

A fentiekben p_i az i . kimenet előfordulási valószínűsége, y_i az i . kimenethez tartozó hozam, N a lehetséges kimenetek száma, E a várható érték és V pedig a variancia.

1. ábra: A Markowitz-féle hatékony portfóliók



Forrás: (Markowitz, 1952, p. 82)

2.8 Maximális visszaesés (Maximal Drawdown, MDD)

A maximális visszaesés (maximal drawdown) kockázati mutatót széles körben használják a befektetési szektorban. A Risk.net felmérésében (Risk.net, 2020) résztvevő befektetési vállalatok több, mint 30%-a ezt a mutatót is alkalmazza a kockázatkezelési módszertanában. Az MDD kockázati mutató egy portfólió értékének egy adott időszakban elért legnagyobb relatív értékviszaesését adja meg, azaz a portfólió értékének maximuma és az azt követő minimuma közötti relatív különbséget. A maximális visszaesés formális definíciója a következő (Chekhlov, et al., 2005):

Legyen $x(t_k)$ egy olyan portfólió súlyvektora a t_k időpontban, ami egy kockázatmentes és m kockázatos eszközből áll.

$$x(t_k) = (x_0(t_k), x_1(t_k), x_2(t_k), \dots, x_m(t_k)) \quad (7)$$

A súlyok összege mindig 1, azaz meg kell felelniük a következő követelménynek:

$$\sum_{i=0}^m x_i(t_k) = 1 \quad (8)$$

Az i . eszköz hozamrátáját t_k időpillanatban $r_i(t_k)$ jelöli, és a következő képlet definiálja, ahol $p_i(t_k)$ az i . eszköz árfolyama t_k időpillanatban:

$$r_i(t_k) = \frac{p_i(t_k)}{p_i(t_{k-1})} - 1 \quad (9)$$

Ez alapján a portfólió hozamrátáját ($r_k^{(p)}$) a következő definíció adja meg:

$$r_k^{(p)}(x(t_k)) = r(t_k) \cdot x(t_k) = \sum_{i=0}^m r_i(t_k)x_i(t_k) \quad (10)$$

A következő definícióban $w_k(x(t_k))$ a portfólió hozamrátáinak összege (uncompounded cumulative portfolio rate or return) a k . időpontig, amit a későbbiekben az egyszerűség kedvéért w_k jelöl.

$$w_k(x(t_k)) = \begin{cases} 0, & \text{ha } k = 0 \\ \sum_{l=1}^k r_l^{(p)}(x(t_l)), & \text{ha } k = \overline{1, N} \end{cases} \quad (11)$$

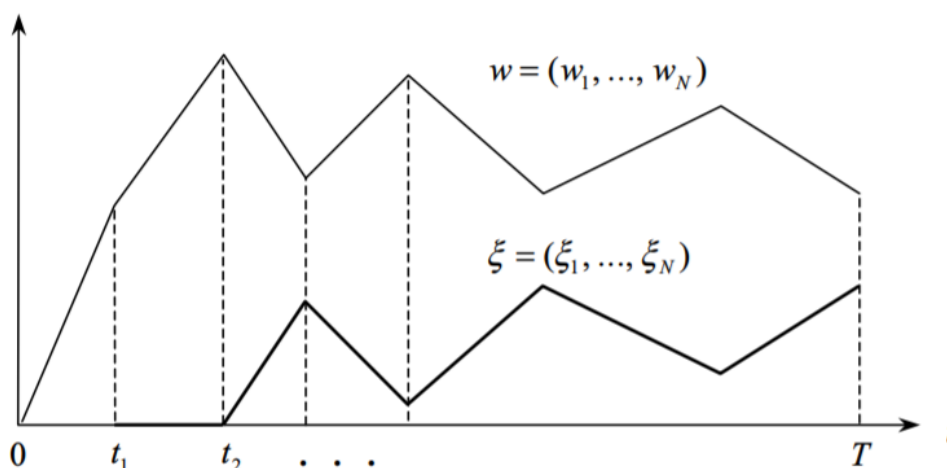
Fontos kiemelni, hogy a hozamráták kumulálása nem a „kamatos kamat” elve mentén történik, hanem összeadással. Ezek alapján minden időpontra kiszámítható az azt megelőző maximumtól számított csökkenés abszolút értéke (absolute drawdown, AD).

$$AD(w) = \xi = (\xi_1, \dots, \xi_N), \quad \xi_k = \max_{0 \leq j \leq k} \{w_j\} - w_k \quad (11)$$

Ahol w egy N elemű vektor. $k = \overline{1, N}$ felhasználásával tehát $w = (w_1, \dots, w_N)$. $AD(w)$ és ξ N elemű vektorok. A maximális visszaesés az abszolút visszaesések maximuma.

$$MDD(w) = \max_{1 \leq k \leq N} \{\xi_k\} \quad (12)$$

2. ábra: A Maximum Drowdown mutató számítása (példa)



Forrás: (Chekhlov, et al., 2005)

A maximális visszaesés mutató alapvetően relatív kockázati mutató, azaz a pozíció mérete nem befolyásolja az MDD mértékét. A fenti definícióban ez ott érhető tetten, ahol a portfólió kumulatív hozamát (w_k) definiáljuk a hozamok $r_t^{(p)}$ függvényében. A hozam a pozíció értékmozgásának relatív értéke, így a számítás végén megkapott mérőszám is relatív lesz. A relatív mérőszám-tulajdonság abból a szempontból releváns, hogy a koherencia axiómák közül nem összeegyeztethető a pozitív homogenitás axiómájával, így megállapíthatjuk, hogy a maximális visszaesés kockázati mutató nem koherens.

2.9 Kockáztatott érték (VaR)

A kockáztatott érték (value-at-risk, VaR) a mai napig az egyik leggyakrabban használt kockázati mutató, Risk.net felmérése szerint a befektetéssel foglalkozó vállalatok több, mint 70%-a ezt a mutatót is használja a kockázatkezelési módszertanában (Risk.net, 2020). A kockáztatott érték mutató bevezetése Jorion (2007) nevéhez köthető. Az Európai Unió Solvency II keretrendszere (Európai Parlament, 2009) jelenleg is a VaR használatát írja elő, valamint bizonyos helyzetekben a Bázeli III keretrendszer (BCBS, 2021) is használja. A VaR abból a felismerésből született, hogy a derivatívok és más mérlegen kívüli termékek kockázatát is mérni kell. A VaR mutató a JPMorgan RiskMetrics módszertanának részeként lett széles körben ismert (McNeil, et al., 2015), nagy előnye, hogy a kockázatot a varianciával ellentétben a befektetés pénznemében, azaz

egyszerűen értelmezhetően adja meg. A kockázatos érték mutató ezen utóbbi tulajdonsága miatt megfelel a pozitív homogenitás koherenciakritériumnak.

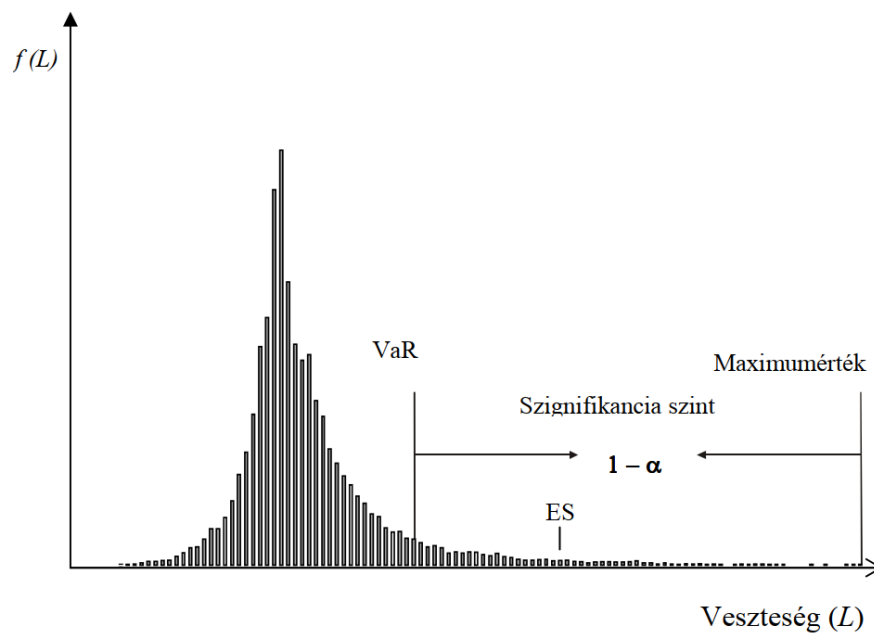
A kockázatos érték azt adja meg, hogy egy adott időszak alatt, adott konfidenciaszint mellett, mi az a maximális veszteség amire számíthatunk. Például a kockázatos érték 99 százalékos konfidenciaszint mellett megadja azt a veszteséget, aminél nagyobb veszteséget csak a kereskedési napok 1%-ában várhatunk. Ebből a definícióból a mutató legnagyobb hátránya is világosan látszik: a mutató nem mond semmit a konfidenciaszintet meghaladó leginkább veszteséges napokról. A VaR-mutatót legegyszerűbben a konfidenciaszintnek megfelelő kvantilisként definiálhatjuk:

$$VaR_u = q_u(F_X) \quad (13)$$

A fentiekben q_u a kvantilis függvény és F_X a hozam eloszlása.

A VaR az ES elterjedését megelőző időszakban iparági standard volt, de használatával kapcsolatban számos hiányosság merült fel. A VaR-mutató nem koherens (Artzner, et al., 1999), és ez többek között ahhoz vezet, hogy bizonyos esetekben a portfóliódiverzifikáció a mért kockázat növekedéséhez vezet, ami ellentmond annak az alapvető, logikus elvárásnak, hogy a diverzifikáció csökkentse a kockázatot.

3. ábra: Kockázatos érték (VaR) és a várható többletveszteség (ES)



Forrás: (Bugár, 2015, p. 77)

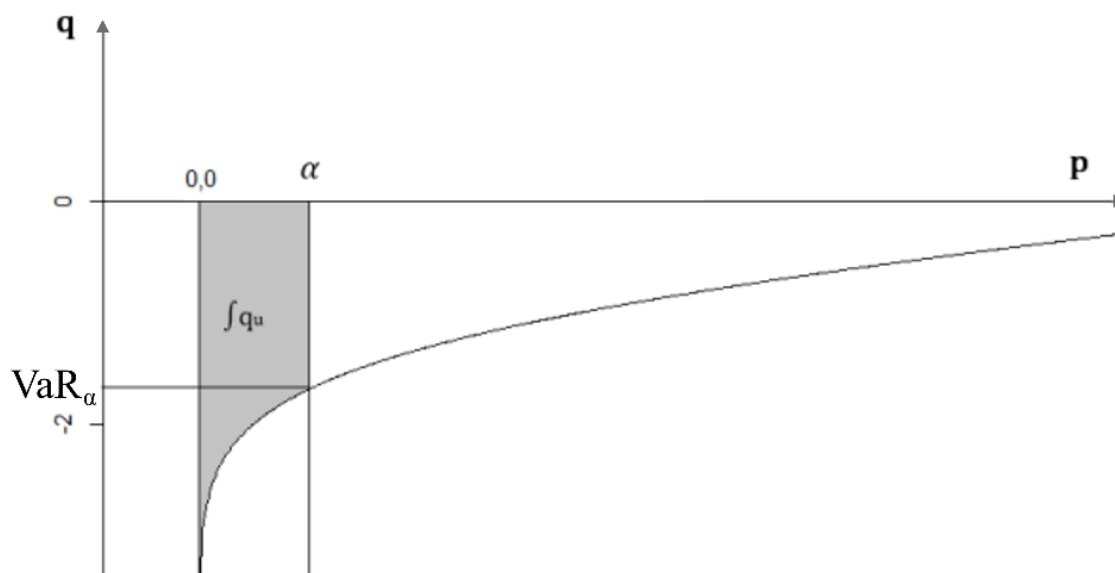
2.10 Várható többletveszteség (ES)

A várható többletveszteség (expected shortfall, ES, conditional value at risk) egy olyan kockázati mutató, ami a VaR-küszöbértéknél nagyobb veszteségek átlagos értékét adja meg (Rockafellar & Uryasev, 2002). A mutató az elmúlt időszakban folyamatosan veszi át elődje, a kockázatosított érték helyét. Az ES egy koherens kockázati mérőszám, így használata számos előnnyel bír a VaR-mutatóval szemben. Folytonos hozam eloszlásra az ES-mutatót a következő definícióval kapjuk meg:

$$ES_{\alpha} = -\frac{1}{\alpha} \int_0^{\alpha} VaR_u(X) du = -\frac{1}{\alpha} \int_0^{\alpha} q_u(F_X) du \quad (14)$$

A fentiekben X a portfólió hozamát jelöli, és $\alpha \in (0,1)$ a konfidenciaszintet.

4. ábra: A várható többletveszteség és a kvantilis függvény kapcsolata



forrás: (saját szerkesztés)

2.11 A várható többletveszteség visszatesztelhetősége

A várható többletveszteséggel kapcsolatos legnagyobb kritika az elicabilitás fogalmához köthető. A várható többletveszteség mutatóról Gneiting (2011) bebizonyította, hogy nem rendelkezik az elicabilitás tulajdonsággal. Több tanulmány jelent meg az irodalomban, azzal az állásponttal, hogy emiatt a várható többletveszteség mutató nem megfelelően ellenőrizhető (Ziegel, 2016) (Embrechts & Hofert, 2014) .

Acerbi és Székely (2017) kifejtik, hogy az elicítabilitás hiányából nem következik, hogy az ES-mutatót ne lehetne visszatesztelni. A szerzőpáros bemutat egy olyan mérőszámot, ami mind modellvalidáció, mind modellszelekció szempontjából megfelelően tudja értékelni az ES-alapú modelleket. A visszatesztelés témakörének részletes kifejtését a 7. fejezet tartalmazza.

2.12 Historikus és hipotetikus terheléses tesztek

A korábban bemutatott kockázati mérőszámok becslésének leginkább kézenfekvő és több szempontból leginkább megfelelő módja a múltbeli adatok használata. Ezek egyfajta biztonságot adnak és hitelességet az ellenőrzések eredményeinek, de bizonyos eseteket nem tudnak kezelni. A historikus adatok használata például helytelen eredményt adhat, ha nincs megfelelő mennyiségű adat egy adott eszközhöz. Egy újonnan megjelent részvény idősoraiból például hiányozhatnak azok az extrém időszakokat leíró adatok, amelyekből a különösen veszteséges, válság alatti időszakok becsléseit lehetne elkészíteni. Ezen kívül az irodalomban további érvek is megjelennek a terheléses tesztek elvégzése mellett. Berkowitz (2000) és Christoffersen (2012, p. 314) a következő esetekben javasolja a terheléses tesztek használatát:

- Ha szükséges olyan sokkok szimulálása, amelyek előre láthatóan gyakrabban fognak megtörténni, mint ahogy azt a korábbi történeti adatok sugallják.
- Olyan sokkokat szeretnénk szimulálni, amelyek még nem történtek meg, de megtörténhetnének. Ilyen esetre lehet példa egy tőzsdepiaci válság szimulációja.
- Ha szimulálni szeretnénk, hogy mi történik, ha a jelenlegi statisztikai minták megváltoznak, például szeretnénk szimulálni, hogy milyen hatással lenne, ha a veszteségek csoportosan, egy időben realizálódnának.
- Ha olyan strukturális változások hatását szeretnénk megvizsgálni, amelyekről elképzelhetőnek tartjuk, hogy a jövőben megtörténhetnek. Erre példa lehet a thaiföldi THB deviza árfolyamrögzítésének 1997-es eltörlése.

A fenti esetekre külön-külön tesztek készülhetnek, ahol az elemzők a scenáriókhöz megfelelő idősorokat generálnak és ezen futtatják az ellenőrzési módszereket, megvizsgálva azok eredményét. Bizonyos esetekben az is szükséges lehet, hogy ezek a scenáriók a becslési modellekbe is beépüljenek, oly módon, hogy a meglévő valós adatokkal együtt, azokat kiegészítve befolyásolják a kockázati mutatók értékeit. A

fenti külön-külön elkészített tesztek esetében szintén problémát jelent, hogy a kockázatkezelők nehezen tudják objektívan és egységesen kezelni az egymásnak akár ellentmondó eredményeket. Elképzelhető, hogy bizonyos scenáriók túl nagy súlyt kapnak a modellek vagy a portfóliók kialakításakor, míg más scenáriók fontosak lennének, de nem kapnak akkora figyelmet. Ezekre a problémákra, illetve a becslési szakasz módosítására javasolja Berkowitz (2000) az olyan egységes eloszlások létrehozását, melyek a különböző scenáriókat a hozzájuk tartozó valószínűségi súly használatával kapcsolja össze.

$$x \sim \begin{cases} f(\cdot), 1 - \sum \alpha_i \text{valószínűséggel} \\ f_{stress-1}(\cdot), \alpha_1 \text{valószínűséggel} \\ \vdots \\ f_{stress-m}(\cdot), \alpha_m \text{valószínűséggel} \end{cases} \quad (15)$$

A fenti összetett eloszlásban a $f(\cdot)$ jelöli az idősorokból adódó eloszlást, míg a többi az m darab terheléses teszt különböző scenáriói alapján generált eloszlásokat. Az i . scenárió előfordulási esélyét az α_i valószínűség adja meg. Az adatok szimulálásához elegendő egy egyenletes eloszlás generálása, ahol a 0 és 1 közötti intervallumot a scenáriók valószínűségei alapján kell felosztani, majd a megfelelő aleloszlásból kell elemet generálni, ha a hozzá tartozó sávba esik az első generált érték.

3 A kockázatosított érték és a várható többletveszteség mutatók megjelenése a szabályozásban és a gyakorlati alkalmazásban

3.1 A bank és a biztosítási szektor szabályozása

3.1.1 A Bázeli Bankfelügyeleti Bizottság és a Bázeli Keretrendszer

A Bázeli Bankfelügyeleti Bizottságot 1974-ben a G10 központi bankjainak vezetői hozták létre. A bizottság közvetlen szabályozási hatáskörrel nem rendelkezik az egyes nemzetek felett, ezért ajánlások útján egységes normákat hoz létre a hitelintézetek szabályozásával kapcsolatban. Az egyes nemzeti hatóságok saját hatáskörben a nemzeti sajátosságaikhoz igazítva építik az ajánlásokat a saját jogrendjükbe (Széles, et al., 2016).

Az első bázeli keretrendszer 1988-ban jelent meg „Tőkefelmérés és tőkeszabványok nemzetközi konvergenciája” címmel és fontos lépés volt a minimum tőke nemzetközi standardizálása felé (Bugár & Rattig, 2016). A Bazel I keretrendszer elsődlegesen a hitelkockázatra fókuszált, amit akkoriban a bankok kockázatának legjelentősebb forrásának tekintettek. Visszatekintve a Bazel I megközelítése túlzottan leegyszerűsítő volt, és a keretrendszerből hiányzott a kellő differenciáltság. Példa erre, hogy hitelkockázat mérésénél a szereplőket három egyszerű kategóriára osztották: államok, szabályozott bankok és más szereplők. A tőkekövetelmény egy vállalati hitelező esetében ötször akkora volt, mint egy OECD bank esetében, de a vállalati hitelezők között a keretrendszer egyáltalán nem tett különbséget, sem hitelminősítés sem más szempont alapján. Szintén hátrány volt, hogy az ellenkező irányú short és long pozíciók nettósítását a szabályozás nem engedte (McNeil, et al., 2015). 1996-ban a Bázeli Bankfelügyeleti Bizottság egy fontos kiegészítést fűzött hozzá a Bazel I keretrendszerhez. A módosítás a piaci kockázat kezelésére egy standardizált modellt írt elő, ugyanakkor lehetővé tette, hogy a nagyobb bankok egy belső, saját fejlesztésű VaR-alapú modellt használjanak.

A Bazel II keretrendszer 2004-ben került elfogadásra. A keretrendszer a szabályozás három pillérét határozta meg: az első a szükséges tőke kiszámításának módszerét szabályozta elsősorban a hitelkockázatra fókuszálva. A módosított rendszer lehetővé tette a bankok számára, hogy kifinomultabb módszereket alkalmazzanak a

hitelportfóliójukkal kapcsolatos kockázat értékelésére. A bankok az előírt eljárásokon kívül választhattak saját, belső vagy külső hitelminősítési modelleket is. A Bázeli II keretrendszer megalkotói arra számítottak, hogy a kifinomultabb kockázati modellek a hitelkockázattal kapcsolatos tőkekövetelmények csökkenését eredményezik majd. A bizottság új típusként bevezette a működési kockázatot annak érdekében, hogy az összességében szükséges tőkekövetelmények ne csökkenjenek (McNeil, et al., 2015). A Bázeli II keretrendszer második pillére a modellezési folyamat felügyeletét érintette, míg a harmadik az információ közzétételét szabályozta. A keretrendszer bevezetésének hatásaként a magyarországi piaci szereplők a kockázatkezelési osztály szerepének növekedését, a belső folyamatok átrendeződését, az adatnyilvántartási rendszerek szükségszerű fejlesztését és a tőkekövetelmények növekedését említik (Sallay & Szigel, 2014). A Bázeli II a tőkemegfelelési mutató számítására a kockázatosított érték (VaR) mutató használatát javasolta.

A keretrendszert több kritika is érte (McNeil, et al., 2015) (Repullo & Suarez, 2008) (Behn, et al., 2016). Egyrészt a keretrendszer a kockázatosabb időszakokban – azaz jellemzően gazdasági válságok alatt – nagyobb tőkemegfelelési rátát írt elő ezzel tovább csökkentve a piacok ilyen időszakokban amúgy is alacsony likviditását. A Bázeli II keretrendszer másik hátránya az volt, hogy lehetővé tette a szabályozási arbitrázst. A keretrendszer meghagyta annak lehetőségét, hogy a CDO-k (collateralized debt obligation) és más hozzájuk hasonló termékek segítségével a bankok a kockázatot átmozgassák a magasabb tőkekövetelménnyel járó banki könyvből az alacsonyabb tőkekövetelménnyel járó kereskedési könyvbe.

A Bázeli III keretrendszer szükségességét a 2007-es pénzügyi válság mutatta meg, mely mind a közvélemény, mind a döntéshozók számára azt bizonyította, hogy a Bázeli II keretrendszer nem ad kellő stabilitást a pénzügyi szektornak (McNeil, et al., 2015). Az új keretrendszer több területen módosította a korábbi javaslatokat. A Bázeli III a banki tőke mennyiségi és minőségi növelését írja elő a lényeges mutatók definícióinak felülvizsgálatával, így a bankok veszteségkezelési kapacitása nő. További módosítás, hogy a Bázeli III lehetővé teszi a tőkekövetelményt befolyásoló mutatók kontraproduktív kiigazítását, ezzel a Bázeli II azon hibáját korrigálva, hogy a tőkekövetelmények rendszere gazdasági válságok alatt tovább csökkenti a piacok likviditását. A Bázeli III másik újítása, hogy bevezeti az áttételi arányt és annak korlátozását ezáltal megakadályozza a túlzott

tőkeáttételt (McNeil, et al., 2015). A Bázeli III szintén jelentős módosítása az, hogy a VaR kockázati mutató helyett a várható többletveszteség (ES) mutató használatát javasolja a kereskedési könyv kockázatának mérésére.

3.1.2 Solvency II, az EU szabályozása a biztosítási szektorra

Az Európai Unió az 1970-es években adta ki első irányelveit a biztosítási szektorban előírt szavatoló tőke mértékének meghatározásával kapcsolatban (Az Európai Közösségek Tanácsa, 1973) (Az Európai Közösségek Tanácsa, 1979). A szavatoló tőke egy puffer, ami olyan események ellen védi a vállalatot, mint a káresemények előre nem látott megnövekedése vagy a kivételesen nagy befektetési veszteségek. A szabályozást a különböző európai országok eltérően alkalmazták, így megszületett az igény egy harmonizált szabályozásra. A Solvency I keretrendszer 2002-ben került elfogadásra (Az Európai Parlament és a Tanács, 2002a) (Az Európai Parlament és a Tanács, 2002b), és egy viszonylag egyszerű rendszer volt. A vállalatoknak egy minimum tőke alapot kellett fenntartaniuk, kiegészítve a biztosítási állományok összegének megadott százalékkal. Az elvárt kiegészítő tőkekövetelmény számításában két fő kategóriát különböztetett meg: az élet- és a nem-életbiztosításokét. A Solvency I rendszer előnye, hogy könnyen érthető és egyszerűen ellenőrizhető volt. A rendszer hátránya az volt, hogy csak a volument vette figyelembe és a tényleges kockázatot nem. A Solvency II keretrendszer 2009-ben került elfogadásra (Európai Parlament, 2009). Célja a szabályozás további harmonizálása, a biztosítottak védelme és a pénzügyi piac stabilitásának növelése. A Solvency II keretrendszer kockázati mérőszámként a VaR mutató használatát írja elő (McNeil, et al., 2015).

3.1.3 Swiss Solvency Test (SST)

Az SST a biztosítási szektor 2011 óta hatályos svájci szabályrendszere. A rendszer alapja az eszközök és kötelezettségek piaccal konzisztens értékelése. Az SST a piaci, hitel és biztosítási kockázatokat vizsgálja. A szabályrendszer tőkekövetelmény-számítása nagy vonalakban a következő összefüggések mentén történik (Federal Office of Private Insurance, 2006):

- MVM (market value margin): A kötelezettségek piaccal konzisztens értékének és a hozzájuk rendelhető kifizetések legjobb becslése közötti különbség.

- RBC (risk-bearing capital): Az RBC az eszközök piaccal konzisztens értéke és a kötelezettségek diszkontált legjobb becslése közötti különbség.
- TC (target capital): A vállalatnak a TC összegének megfelelő tőkekövetelménynek kell megfelelnie. A TC összege az egy éves diszkontált RBC, valamint az aktuális RBC különbségének várható többletvesztesége kiegészítve az MVM értékével.

Az SST tehát az EU Solvency II szabályozásával szemben nem a kockázatot érték (VaR) mutatót, hanem a várható többletveszteség (ES) kockázati mutatót használja a tőkekövetelmény kiszámításához.

3.2 OTC és tőzsdei letétszámítás

A tőzsdei kereskedés elszámolása egy központi partneren, az elszámolóházon keresztül történik. Mind az eladó, mind a vevő az elszámolóházzal áll kapcsolatban, így a kereskedő felek partnerkockázata minimálisra csökken, hiszen azt az elszámolóház veszi át. Az elszámolóház partnerkockázatának csökkentése érdekében a partnereitől letét megfizetését követeli meg. A letét összege az ügylet kockázatának függvénye, valamint napi szinten változik a napi elszámolás eredményeként. Jelentős fejlemény, hogy a 2007-es válságot követően az amerikai Dodd-Frank Act és az Európai Unió derivatív piacon hatályos EMIR (European Market Infrastructure Regulation) szabályozása előírja, hogy a tőzsdén kívüli (over-the-counter, OTC) ügyletek esetében is kötelező egy központi szerződő fél (central counterparty, CCP) közbeiktatása az elszámolási folyamatban. Ahhoz, hogy a központi szerződő fél bevezetése valóban stabilitáshoz vezessen, az elszámolóháznak megfelelő kockázatkezelési eljárásokat kell alkalmazniuk (Berlinger, et al., 2016).

3.2.1 Az ES-mutató használata a SEC által felügyelt elszámolóházaknál

A 2010-es, amerikai egyesült államokbeli Dodd-Frank „Clearing Supervision Act” alapján az elszámolóházaknak a felügyeleten keresztül közzé kell tenniük a módszertani változtatásaikat. A közzétett anyagok jól mutatják, hogy mely szervezetek hogyan használják a várható többletveszteség mutatót (ld. 1. táblázat).

1. táblázat: A SEC által felügyelt elszámolóházak, ahol ES-mutatót használnak

Szervezet	Paraméterek	Alkalmazási terület
<i>Options Clearing Corporation (OCC)</i> (OCC, 2021).	99,99% konfidenciaszint, 2 napos időszak	Az OCC a letétszámítási módszertanában (STANS) Monte-Carlo-szimulációk segítségével a nettó pozíciók várható többletveszteségét számítják ki, majd az így kapott értéket kiegészítik a terheléses tesztből kapott tétellel (SEC, 2019).
Banque Centrale de Compensation (LCH SA)	99,7% konfidenciaszint, 5 napos időintervallum	CDS (credit default swap), és CDS-opciók letéti követelményeinek meghatározása az ES-mutató alapján történik, azzal a kiegészítéssel, hogy az EMIR előírásai alapján a pozíciók összevonásából adódó letétszökkenést mértéke maximum 80% lehet (SEC, 2017a).
<i>ICE Clear Europe</i>	95%, 5 napos intervallum	A piaci volatilitás (procyclicality) mérésére alkalmazott keretrendszer részeként a kezdő letét összegének százalékos változásaira számítják ki az ES-mutatót egy 250 napos mozgó ablakban. Amennyiben a szervezet a fenti eljárás segítségével megnövekedett volatilitást érzékel, úgy elindítja a helyzet felülvizsgálatát. (SEC, 2017b)
<i>Nodal Clear</i>	99,7%	A kezdő letét kiszámításához a Nodal Clear a portfólió várható többletveszteségét veszi alapul. A számításokhoz az elmúlt 1-2 év adatait veszik alapul kiegészítve olyan stresszidőszakokkal, amelyeket a kockázati csoport választ ki. (NodalClear, 2021)

Forrás: (saját szerkesztés a megjelölt források alapján)

3.2.2 Gyakorlati tapasztalatok: ES-alapú letétszámítás a Nodal Clear elszámolóháznál

A Nodal Exchange amerikai egyesült államokbeli árutőzsde elsődleges fókusza az energiaszektor. A Nodal tőzsde 2009-es indulásakor más tőzsdékhez hasonlóan egy külső elszámolóház, az LCH Clearnet szolgáltatásait használta. Később, vertikális integráció keretében a Nodal megalapította saját tulajdonú Nodal Clear elszámolóházát. A Nodal Clear működését az amerikai egyesült államokbeli Határidős Árutőzsdei Kereskedelmet Felügyelő Hatóság (CFTC) 2015-ben engedélyezte derivatív elszámolóházként (Nodal, 2015). A Nodal Clear elszámolóházat 2017-ben fogadta el az Európai Értékpapírpiaci Felügyeleti Hatóság (ESMA) harmadik országbeli központi szerződő félként (Nodal, 2017), majd a Nodal Exchange ugyanebben az évben az European Energy Exchange (EEX) 100%-os tulajdonába került (Nodal, 2017).

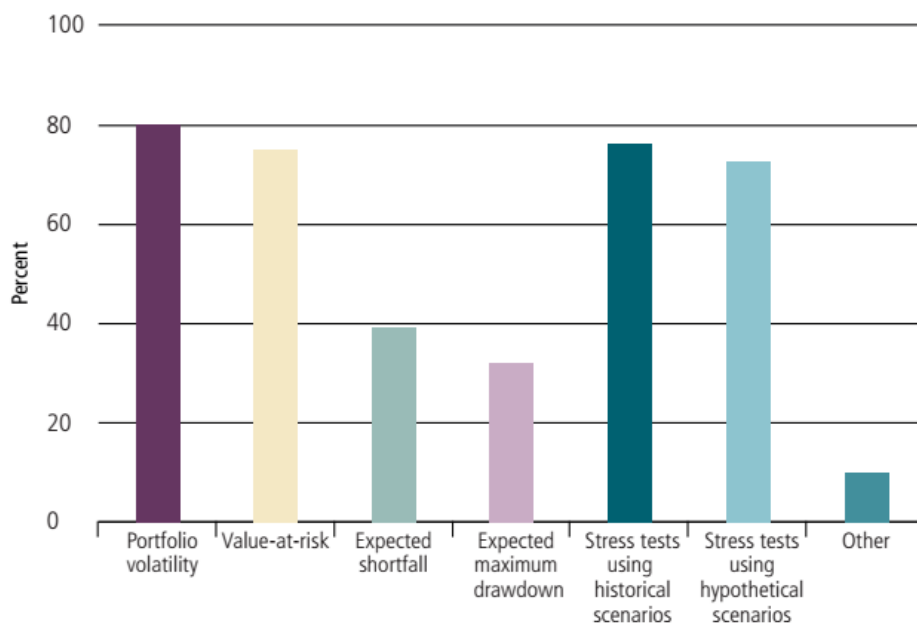
2015-ben a Nodal a korábbi elszámolóház, az LCH, VaR-alapú letétszámítási rendszerének átvételét tervezte, de végül az ES módszertani előnyeit figyelembe véve az utóbbi mutató használata mellett döntött. Ennek elsődleges indoka az volt, hogy a VaR-mutatóval szemben az ES a vastag szélű veszteség eloszlások kezelésére is képes, és a Nodal elszámolóháznál arra számítottak, hogy az általuk lefedett piacon előfordulhatnak olyan extrém esetek, amelyeknek fontos a letéti modellekben is megjelenniük.

Az ES bevezetését követően a kereskedelmi partnerek a kezdeti letétek növekedését tapasztalták, a modell tehát konzervatívabb, mint a régi VaR-alapú rendszer volt. Az ES-alapú modell hatékonyságával a Nodal elégedett volt, a Risk.net portálnak adott 2017-es nyilatkozatában a vállalat vezetője legnagyobb pozitívumként azt emeli ki, hogy az ES-modell kockázatmérése nagyon pontos, így az elszámolóház megfelelő helyen megfelelő mértékű letéti követelményt számít ki (Risk.net, 2017). A vállalat vezérigazgatója a Nodal piaci részesedésének növekedését is az ES használatának tudja be, annak kiemelésével, hogy a szervezet elsődleges feladata a kockázat kezelése. A Nodal piaci részesedése az amerikai energiaszektoron belül 2016-ban egy év alatt 21%-ról 26,4%-ra nőtt. A Nodal módszertanát a szakmai közösség is elismerte, 2017-ben a Nodal az ES-alapú modell bevezetéséért megnyerte a Risk Awards – Clearing house innovation of the year díjat (Risk.net, 2017).

3.3 A Risk.net kockázatkezelési felmérése a Covid első hulláma után

A 2020-ban a Risk.net portál a befektetési piac 158 vásárlóoldali vállalatának bevonásával kérdőíves felmérést végzett a kockázatkezelési szokásokról, illetve a koronavírus-járványt követő piaci folyamatok hatásairól (Risk.net, 2020). A befektetési piac vásárló-oldali vállalatai azok a szereplők, amelyek a piacokon vásárlóként lépnek fel, azaz eszközközkezelők, biztosítók, nyugdíjalapok, befektetési alapok és más intézményi befektetők.

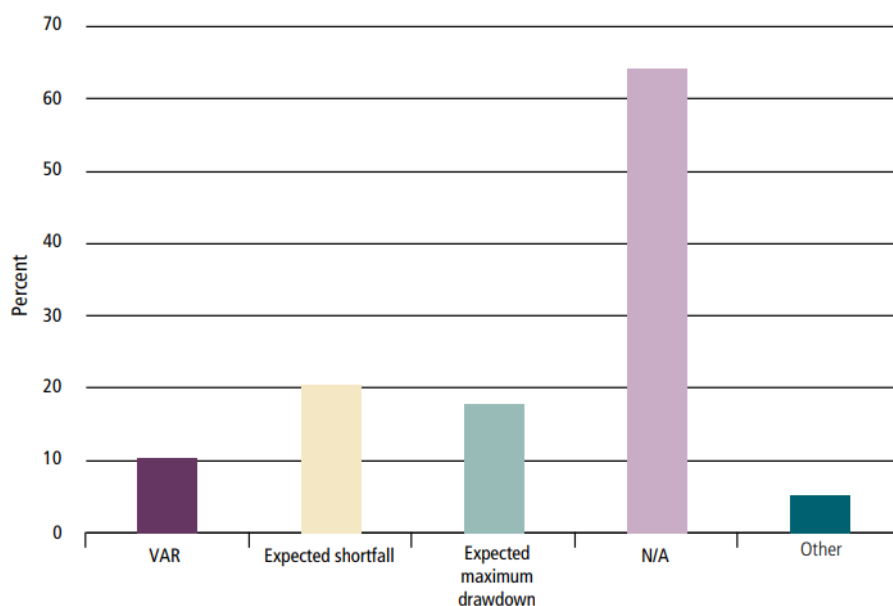
5. ábra: A megkérdezett befektetők által használt kockázati mutatók



Forrás: (Risk.net, 2020)

A megkérdezett vállalatok 58%-a az európai, közel-keleti és afrikai régióba, 23%-a az észak és dél amerikai régióba, 18% pedig az ázsiai/pacifikus és ausztráliai régióba tartozik. A felmérésben résztvevő vállalatok 53%-a 500 vagy több alkalmazottat foglalkoztat, 18%-a 200-499 alkalmazottat, 13%-a 100-199 alkalmazottat, 9%-a 50-99 alkalmazottat, 6%-a 10-49 alkalmazottat, a maradék 2% pedig kevesebb, mint 10 alkalmazottat foglalkoztat. A felmérésből az is kiderül, hogy a megkérdezett vállalatok 85%-a korlátozza a portfólió menedzserek által vállalható kockázatot.

6. ábra: A megkérdezett vállalatok által bevezetni tervezett új módszertanok



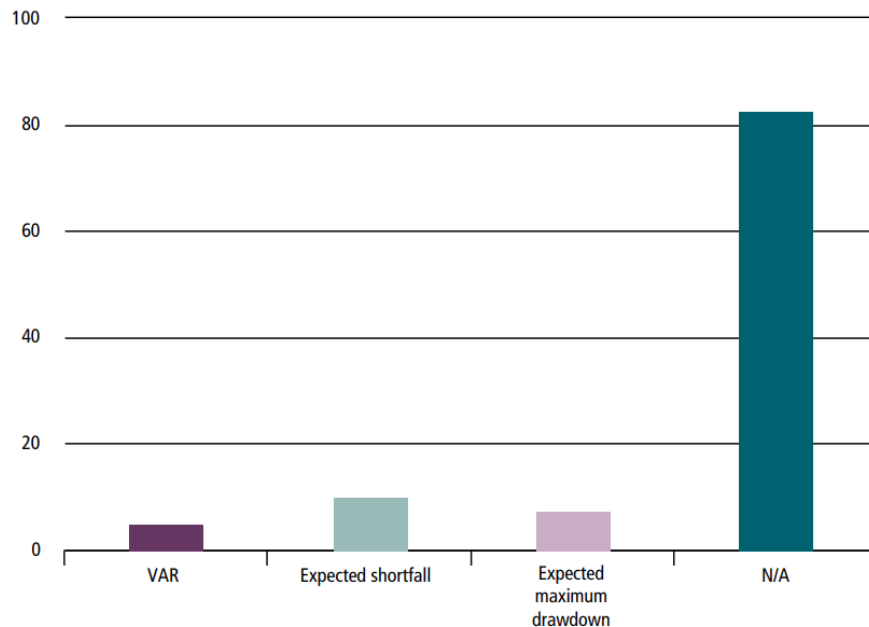
Forrás: (Risk.net, 2020)

A 6. ábra megmutatja, hogy azok a vállalatok, amelyek a járványt követő időszakban új módszertan bevezetése mellett döntöttek többnyire a várható többletveszteség (expected shortfall) mutatót választották. Ez részben annak tudható be, hogy a kockázatos érték (VaR) alapú módszertant számos kritika érte. A felmérés külön kitért arra, hogy a megkérdezett vállalatok számára mennyire volt hasznos a meglévő VaR-modelljük a koronavírus-járványt követő piaci nehézségek kezelésében. Az eredményekből kiderül, hogy a vállalatok 14%-a tartotta a VaR-modelljét nagyon hasznosnak, 53%-a részben hasznosnak és 33%-a egyáltalán nem tartotta a VaR-modelljét hasznosnak. A Risk.net-felmérés kapcsán megkérdezett szereplők nyilatkozataiból is kiderül, hogy a Covid-krízis hatására sok ipari szereplő elégedetlen a VaR-alapú modelljével. Ez egyik szakértő külön kiemeli, hogy a VaR-alapú metrikáik februárban még semmit sem mutattak, csak márciusban, a krízis közepén kaptak extrém magas kockázatot mutató értékeket (Mannix & St. Clair, 2020).

A Risk.net-felmérés érdekessége, hogy azt is megmutatja, hogy a krízis nyomán nem csak a VaR-alapú modellekkel, hanem más modellekkel kapcsolatban is elégedetlenek voltak a befektetési vállalatok, sőt a megszüntetni tervezett módszertanok között a várható többletveszteség volt a leggyakoribb. Véleményem szerint ez annak

tudható be, hogy a várható többletveszteség valóban megoldja a VaR azon problémáját, hogy a VaR elhanyagolja a legrosszabb 1-5%-át az eloszlásnak, de a várható többletveszteség hiába veszi figyelembe a teljes eloszlás szélet, ha nincs becslési időszakban komoly pénzügyi krízis.

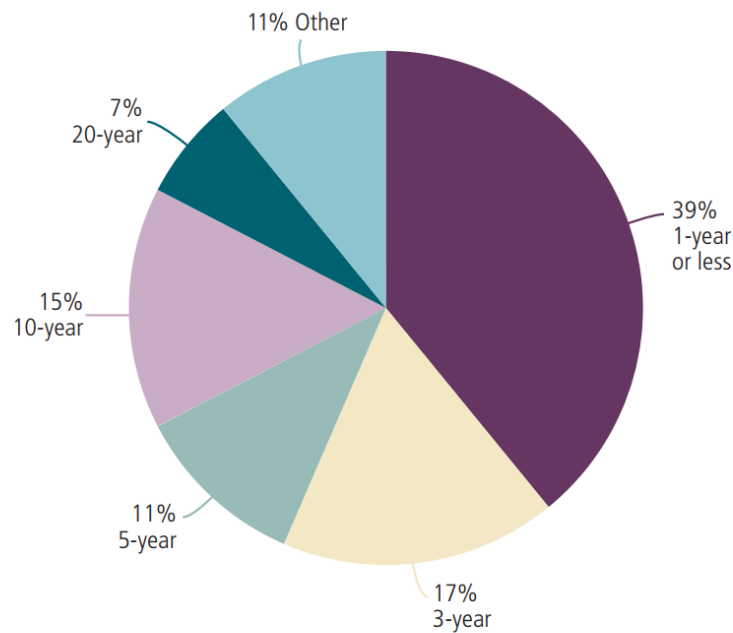
7. ábra: A megkérdezett vállalatok által megszüntetni tervezett módszertanok



Forrás: (Risk.net, 2020)

A következő ábrából kiderül, hogy a megkérdezett vállalatok 39%-a kevesebb, mint egyéves visszatekintési, más néven becslési időszakot használ a modelljében. A visszatekintési időszak az az intervallum, ami alapján a modellparamétereket becsüljük. Az ES-mutató esetén tehát ez azt jelenti, hogy ha a visszatekintési időszak egy év, akkor az elmúlt egy év legrosszabb napi hozamainak átlaga fogja megadni a kockázat mértékét. Ezzel az a probléma, hogy ha az elmúlt egy év éppen egy stabil gazdasági időszak volt, akkor a legrosszabb napok sem tudják megmutatni azt, hogy mi fog egy hirtelen bekövetkező válság alatt történni és ennek megfelelően a valós veszteségek is jelentősen meg fogják haladni a becsült kockázatot. A 2008-as pénzügyi válság volt az utolsó, a koronavírus-járvány időszakához mérhető negatív pénzügyi időszak, amely 12 évvel korábban volt. A Risk.net felméréséből kiderül, hogy a vállalatok 82%-a ennél rövidebb visszatekintési időszakot használt a koronavírus-időszakot megelőzően, így világos, hogy a historikus adatok alapján a modellek nem dolgoztak kellően veszteséges idősorokkal.

8. ábra: A visszatekintési időszak hossza a felmérésben részt vevő pénzügyi vállalatoknál

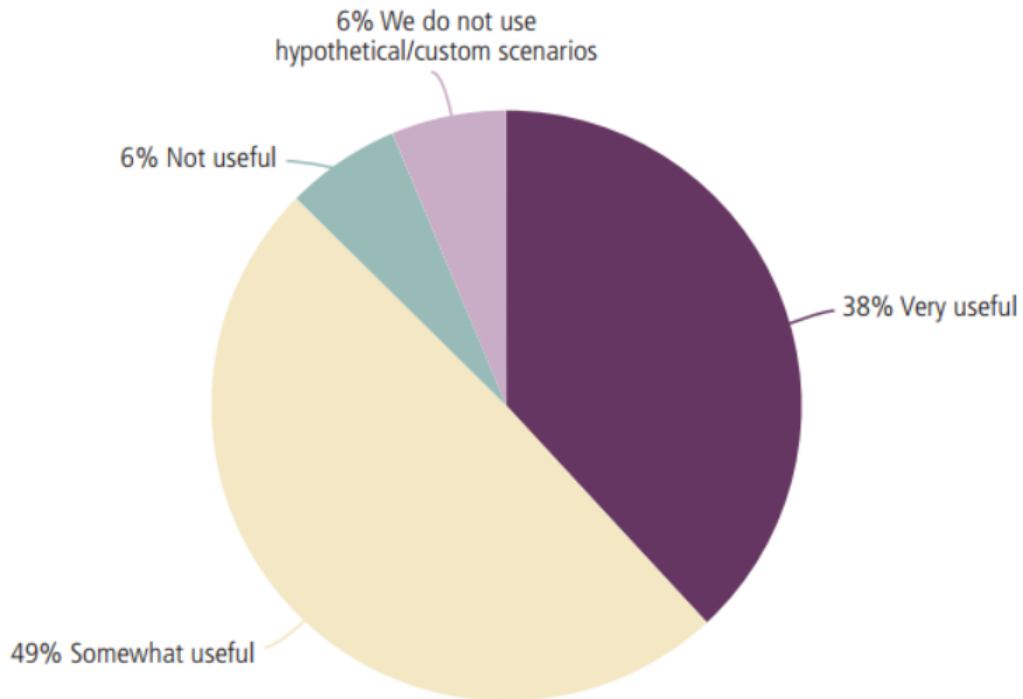


Forrás: (Risk.net, 2020)

Szintén érdekes kérdés, hogy miért pont a várható többletveszteség modellje a leggyakrabban megszüntetett modell. Véleményem szerint ez abból a téves elvárásból adódhat, hogy a VaR kockázati mérőszámot alkalmazó vállalatok gyakran a VaR-mérőszámnak tulajdonítják a modelljeik azon hiányosságait, amelyek valójában nem a VaR-mutató hiányosságai, hanem a becslési eljárásé. Ilyenkor pedig önmagában az ES bevezetése nem eredményezi azt a minőségi javulást az előrejelzésekben, amiket esetleg elvárnak az új modelltől. Jó példa a mérőszámon túlmutató becslési paraméterekre a fent leírt visszatekintési időszak megválasztása, de ugyanígy a becsléshez használt valószínűségi eloszlás vagy más alternatív eljárás is jelentős hatással van az előrejelzés pontosságára.

A visszatekintési időszakkal kapcsolatos problémára ad megoldást a terheléses tesztek használata. A Risk.net (2020) felmérése alapján a befektetők több, mint 80%-a használ historikus, és ugyanennyi százalékuk használ hipotetikus terheléses teszteket. A felmérésből sajnos nem derül ki, hogy a két módszert használók között pontosan mekkora az átfedés, de adódik, hogy a vállalatok legalább 60%-a mindkét módszert használja.

9. ábra: A hipotetikus terheléses tesztek hasznosságának megítélése a Risk.net felmérésében résztvevő vállalatok között



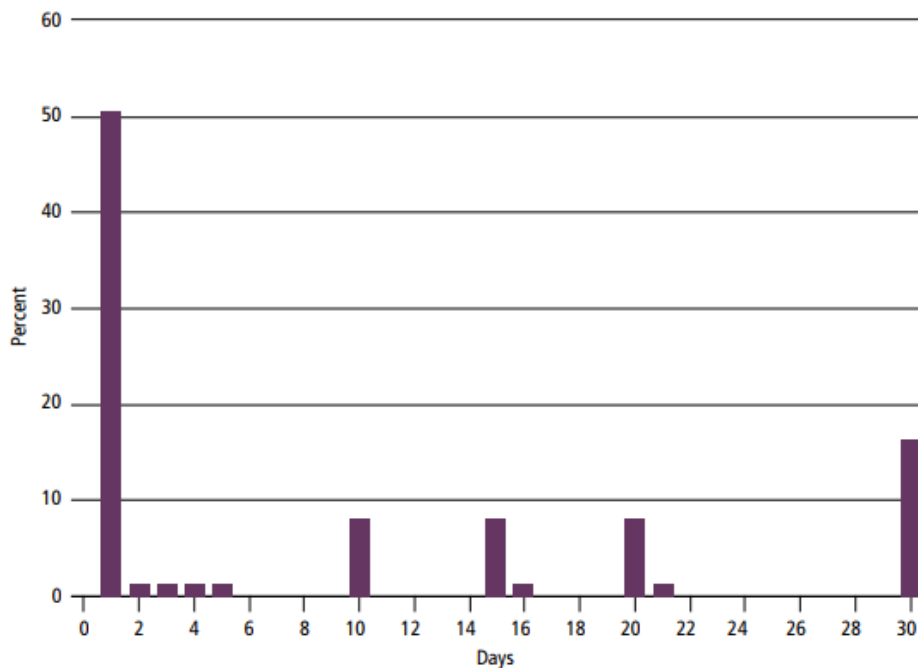
Forrás: (Risk.net, 2020)

A felmérésből kiderül, hogy a historikus terheléses tesztek kevésbé voltak hasznosak a Covid-időszak kezelésében, mint a hipotetikus eseteket vizsgáló tesztek. Amíg a hipotetikus teszteket 38% nagyon, 49% valamennyire és 6% nem találta hasznosnak, addig a historikus eljárást 32% találta nagyon, 48% valamennyire és 14% nem találta hasznosnak. A felmérésből az is kiderül, hogy a hipotetikus forgatókönyvek forrása többnyire a kockázat és a portfóliómenedzserek szakértői véleménye, kisebb részben pedig valamilyen forgatókönyv-generáló szoftver. A forgatókönyveket többnyire évente (32%), negyedévente (30%) vagy havonta (27%) frissítik, és a portfóliók terheléses tesztjeit évente (18%), negyedévente (22%), havonta (29%) vagy naponta (21%) készítik el.

Modellezés szempontjából szintén érdekes látni, hogy a befektetők milyen modellparamétereket használnak a kockázat becslésére. Az értekezés szempontjából hasznosabb lenne a várható többletveszteség modellekre vonatkozó adatokat látni, de sajnos a Risk.net felmérése erre nem terjedt ki. A 9. ábra tanulsága, hogy a leggyakoribb

az egy napos VaR-mérőszám használata, amit a vállalatok több, mint fele használ. A második leggyakoribb választás a 30 napos VaR-időszak használata.

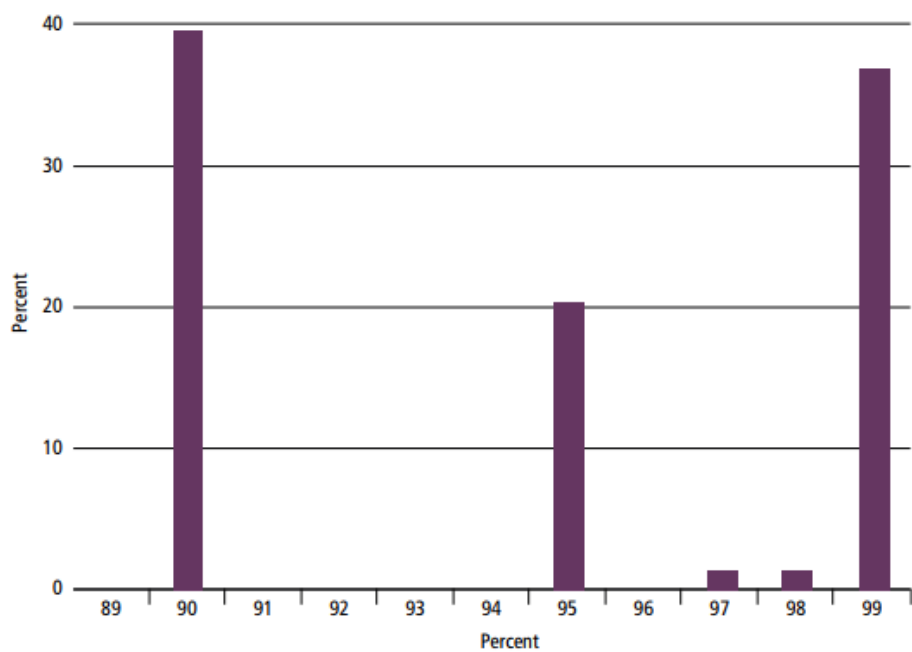
10. ábra: A VaR-mérőszám által lefedett időszak



Forrás: (Risk.net, 2020)

A konfidenciaszintet illetően meglepő, hogy a leggyakoribb választás a 90%. Ezek a modellek a veszteségek legrosszabb 10%-át teljesen figyelmen kívül hagyják. Ennek az a következménye, hogy a VaR értéke nem változik annak ellenére, hogy egy kockázatos időszakban a feldolgozott idősorokban egyre több extrém veszteség jelenik meg, mert ezek az extrém veszteségek a figyelmen kívül hagyott legrosszabb 10%-ra fognak esni. A VaR értéke csak akkor fog változni, amikor az extrém veszteségek olyan gyakorisággal jelentkeznek, hogy azok már „túlsordulnak” az eloszlás szélső 10%-án. Ezek a modellek tehát biztosan később jelzik a veszélyes időszakokat, mint azok, amelyek prudens konfidenciaszinteket használnak. A felmérés sajnos arra nem tér ki, hogy milyen arányban voltak a 90%-os konfidenciaszintet alkalmazó vállalatok között azok, akik aztán a kockázatkezelő modelljüket nem találták hasznosnak a válságkezelésben. A konfidenciaszintet illetően a második és harmadik leggyakoribb választás a 99%, valamint a 95%.

11. ábra: A VaR-mérőszám konfidenciaszintje



Forrás: (Risk.net, 2020)

4 A statisztikai modellek bemutatása

4.1 Várható többletveszteség parametrikus modelljei

A várható többletveszteség becslésére a szakirodalom számos különböző modell használatát veti fel. A mutató becsléséhez használható az egyszerűsége miatt a gyakorlatban leginkább népszerű historikus szimuláció, amelyet többek között Taylor (2008), Righi és Ceretta (2015) is leír. Ez a modell az empirikus eloszlásra épít. A pénzügyi kockázatmérés területén vélhetőleg a legismertebb megközelítés, hogy Markowitz (1952) munkája alapján normális eloszlást feltételezve a kockázat mértékét a variancia határozza meg.

A várható többletveszteség becslésére Broda és Paoella (2011) több eloszlásfüggvény használatát vizsgálja, többek között a Student-féle t - és a ferde t -eloszlást (Azzalini & Capitano, 2003). Taylor (2008) egy kvantilis-expektilis regresszió (Koenker & Bassett, 1978) alapuló modellt mutat be. Ennek a modellnek az előnye, hogy a jelen tanulmányban leírt neurálháló-modellhez hasonlóan közvetlenül az ES-mutatót becsli és nem eloszlásfüggvényre épít.

Az értekezésben a fentiekén túl számos további modell kerül összehasonlításra. Elsőként a legegyszerűbb historikus szimuláció kerül bemutatásra, majd ezt követően az eloszlásalapú, majd a Garch-alapú és végül pedig a kvantilisregresszió-alapú modellek. Az eloszlásalapú modellek esetében az eloszlásfüggvények két csoportban kerülnek bemutatásra: a kiemelt modellek közé az eredményük alapján a legjobb 25%-hoz tartozó modellek kerültek, a további modellek szakaszba pedig a fennmaradó 75% eloszlás-alapú modelljei.

4.2 A historikus szimuláció (HS)

A HS-becslés³ a legegyszerűbb és egyben a banki kockázatkezelési gyakorlatban népszerű eljárás (Taylor, 2008) (Righi & Ceretta, 2015). A módszertan semmilyen feltételezéssel nem él a hozamok eloszlásával kapcsolatban, egyedül arra épít, hogy a hozamértékek függetlenek és azonos eloszlást követnek. A kockázati mutató ΔT

³ A bemutatott eljárás nem az ES torzítatlan becslése. Bemutatását és későbbi használatát az indokolja, hogy Righi és Ceretta tanulmányában (2015, p. 15) a legnépszerűbb ES-becslési eljárásaként kerül bemutatásra, így a modell viszonyítási alapként szolgál.

megfigyelésből álló és T időpillanatig tartó időszakra vonatkozó becslését egyszerűen az azt megelőző N megfigyelés alapján számítjuk ki.

$$VaR_T^\alpha = E^{-1}(\alpha) \quad (16)$$

A fentiekben E az $\{X\}_{T-N-\Delta T}^{T-\Delta T}$ hozamok empirikus eloszlásfüggvénye, $E^{-1}(\alpha)$ pedig az empirikus hozamok α -kvantilise⁴.

$$ES_T^\alpha = (N\alpha)^{-1} \sum \left(\{X\}_{T-N-\Delta T}^{T-\Delta T} \times I_{\{X\}_{T-N-\Delta T}^{T-\Delta T} < VaR_T^\alpha} \right) \quad (17)$$

I_p egy indikátorfüggvény, ami az 1 értéket veszi fel, ha p értéke igaz és a 0 értéket, ha hamis, a $\{X\}_{T-N-\Delta T}^{T-\Delta T}$ pedig egy N elemű intervallum, amely segítségével a $T - \Delta T$ -től kezdődő és T -ig tartó időszakra vonatkozó ES-érték becslése készül.

4.3 Kiemelt eloszlásalapú modellek

Az eloszlásalapú megközelítés első lépése a feltételezett modell paramétereinek becslése. Ez a historikus szimulációhoz hasonlóan a becslési időszak hozamainak felhasználásával történik, azaz a $T - \Delta T - N$ és a $T - \Delta T$ közötti becslési időszakra vonatkozó eloszlásparaméterek kiszámításával. Az előrejelzés alapját az így kiszámított paraméterekkel rendelkező eloszlás adja.

Az eloszlás alapú modellek ismertetésénél az eloszlások angol nyelvű megnevezése került feltüntetésre. Ennek oka az, hogy a lent leírt eloszlások között vannak olyanok is, amelyek nem ismertek széles körben, így az angol nyelvű megnevezésük alapján könnyebben beazonosíthatóak. Annak érdekében, hogy az angol és magyar nyelvű megnevezések ne keveredjenek, egységesen az angol elnevezések kerültek feltüntetésre.

4.3.1 Arcsine

Az arcsine a béta-eloszlás egy speciális esete: $X \sim Beta\left(\frac{1}{2}, \frac{1}{2}\right)$, elsőként Erdős és Kac (1947) munkájában jelent meg. Kiterjesztett változatának sűrűségfüggvénye a következő:

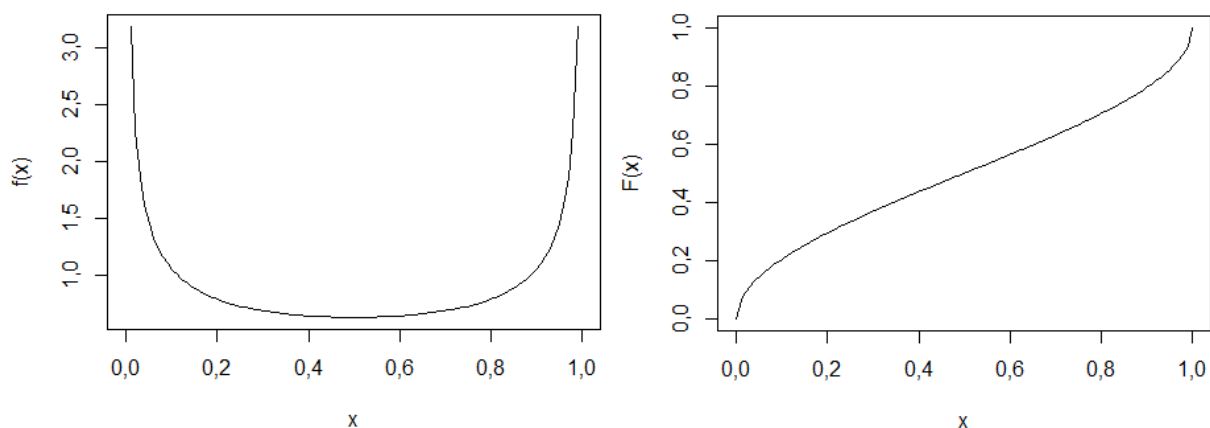
$$f(x) = \frac{1}{\pi x \sqrt{ab-x}} \quad (18)$$

⁴ Részletesen ld.: Righi és Ceretta (2015, p. 19):

Az arcsine kumulatív eloszlásfüggvényét a következő képlet definiálja:

$$F(x) = \frac{2}{\pi} \arcsin \left(\sqrt{\frac{x-a}{b-a}} \right) \quad (19)$$

12. ábra: Az arcsine-eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye standard paraméterekkel



Forrás: (saját szerkesztés)

Az eloszlás kockázatos értéke és várható többletvesztése a következő (Nadarajah, et al., 2015):

$$VaR_p(X) = a + (b - a) \sin^2 \left(\frac{\pi p}{2} \right) \quad (20)$$

$$ES_p(X) = a + \frac{b-a}{p} \int_0^p \sin^2 \left(\frac{\pi v}{2} \right) dv \quad (21)$$

A fentiekben $a < x < b$, $x > 0$, $0 < p < 1$ és $-\infty < a < b < \infty$, illetve a az első, b pedig a második helyparaméter (Nadarajah, et al., 2015).

4.3.2 Burr (burr7)

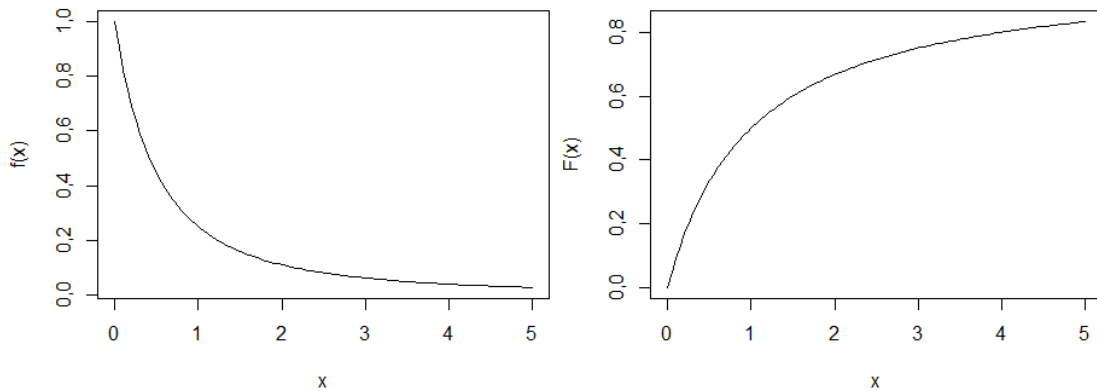
Az eloszláscsaládot Burr vezette be (1942), amellet érvelve, hogy számos esetben előnyösebb közvetlenül a kumulatív eloszlásfüggvényt meghatározni és használni,

mintsem azt később a sűrűségfüggvényből integrálással levezetni. A Burr-eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye a következő (Rodriguez, 1977):

$$f(x) = \frac{kcx^{c-1}}{(1+x^c)^{k+1}} \quad (22)$$

$$F(x) = 1 - (1 + x^c)^{-k} \quad (23)$$

13. ábra: A Burr-eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye standard paraméterekkel



Forrás: (saját szimuláció)

Az eloszlás kockázatos értékét és várható többletveszteségét a következő képletek adják meg (Nadarajah, et al., 2015):

$$VaR_p = [(1 - p)^{-1/k} - 1]^{1/c} \quad (24)$$

$$ES_p(X) = \frac{1}{p} \int_0^p [(1 - v)^{-1/k} - 1]^{1/c} dv \quad (25)$$

A fentiekben $x > 0$, $0 < p < 1$, $c > 0$ az első, $k > 0$ pedig a második alakparaméter (Nadarajah, et al., 2015).

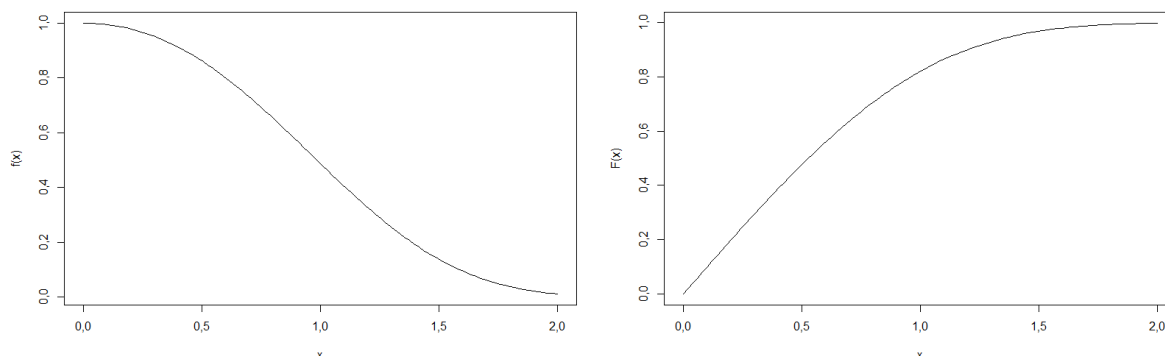
4.3.3 Chen-eloszlás (chen)

Az eloszlás egy fürdőkádgörbét ad a meghibásodási rátára. Az eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye a következő (Chen, 2000):

$$f(x) = \lambda \beta x^{\beta-1} \exp(x^\beta) \exp[\lambda - \lambda \exp(x^\beta)] \quad (26)$$

$$F(x) = 1 - \exp[\lambda - \lambda \exp(x^\beta)] \quad (27)$$

14. ábra: A Chen-eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye standard paraméterekkel



forrás: (saját szimuláció)

Az eloszlás kockázatos értéke és várható többletvesztése a következő (Nadarajah, et al., 2015):

$$VaR_p(X) = \left\{ \log \left[1 - \frac{\log(1-p)}{\lambda} \right] \right\}^{1/\beta} \quad (28)$$

$$ES_p(X) = \frac{1}{p} \int_0^p \left\{ \log \left[1 - \frac{\log(1-v)}{\lambda} \right] \right\}^{1/\beta} dv \quad (29)$$

A fentiekben $x > 0$, $0 < p < 1$, $\beta > 0$ az alakparaméter, $\lambda > 0$ pedig a skálaparaméter (Nadarajah, et al., 2015).

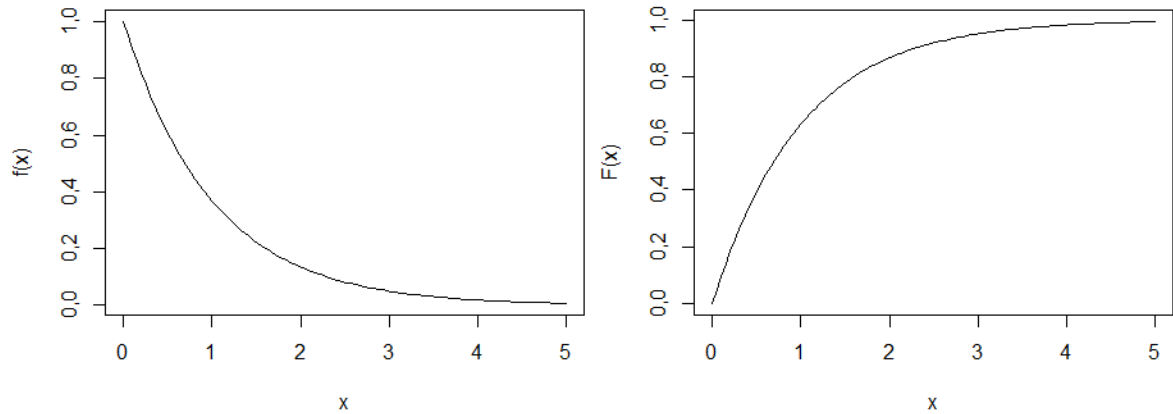
4.3.4 Exponential extension (expept)

Nadarajah és Haghghi (2011) az exponenciális eloszlás általánosításaként mutatja be az expept-eloszlást. Az eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye a következő (Nadarajah & Haghghi, 2011):

$$f(x) = a\lambda(1 + \lambda x)^{a-1} \exp[1 - (1 + \lambda x)^a] \quad (30)$$

$$F(x) = 1 - \exp[1 - (1 + \lambda x)^a] \quad (31)$$

15. ábra: Az expext-eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye standard paraméterekkel



Forrás: (saját szimuláció)

Az eloszlás kockázatos értéke és várható többletvesztése a következő (Nadarajah, et al., 2015):

$$VaR_p(X) = \frac{[1 - \log(1-p)]^{1/a} - 1}{\lambda} \quad (32)$$

$$ES_p(X) = -\frac{1}{\lambda} + \frac{1}{\lambda p} \int_0^p [1 - \log(1-v)]^{1/a} dv \quad (33)$$

A fentiekben $x > 0$, $0 < p < 1$, $a > 0$ az alakparaméter, $\lambda > 0$ pedig a skálaparaméter (Nadarajah, et al., 2015).

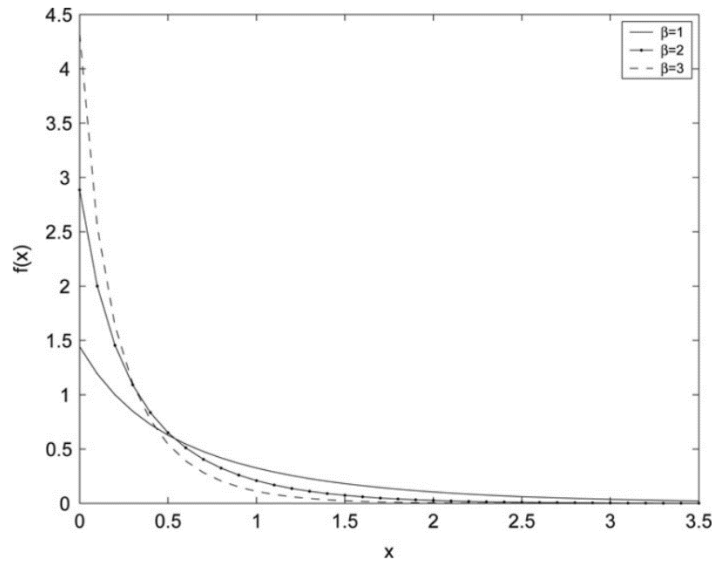
4.3.5 Exponential logarithmic (explog)

Az explog egy kétparaméteres élettartameloszlás, csökkenő meghibásodási rátával (Tahmasbi & Rezaei, 2008). Az eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye a következő:

$$f(x) = -\frac{\beta(1-a) \exp(-\beta x)}{\log a [1 - (1-a) \exp(-\beta x)]} \quad (34)$$

$$F(x) = 1 - \frac{\log[1 - (1-a) \exp(-\beta x)]}{\log a} \quad (35)$$

16. ábra: Az explog-sűrűségfüggvény (p=0,5)



Forrás: (Tahmasbi & Rezaei, 2008, p. 3890)

Az eloszlás kockázatosított értékét és várható többletveszteségét a következő függvények adják meg: (Nadarajah, et al., 2015):

$$VaR_p(X) = -\frac{1}{\beta} \log \left[\frac{1-a^{1-p}}{1-a} \right] \quad (36)$$

$$ES_p(X) = -\frac{1}{\beta p} \int_0^p \log \left[\frac{1-a^{1-v}}{1-a} \right] dv \quad (37)$$

A fentiekben $x > 0$, $0 < p < 1$, $0 < a < 1$ az első, $\beta > 0$ pedig a második skálaparaméter (Nadarajah, et al., 2015).

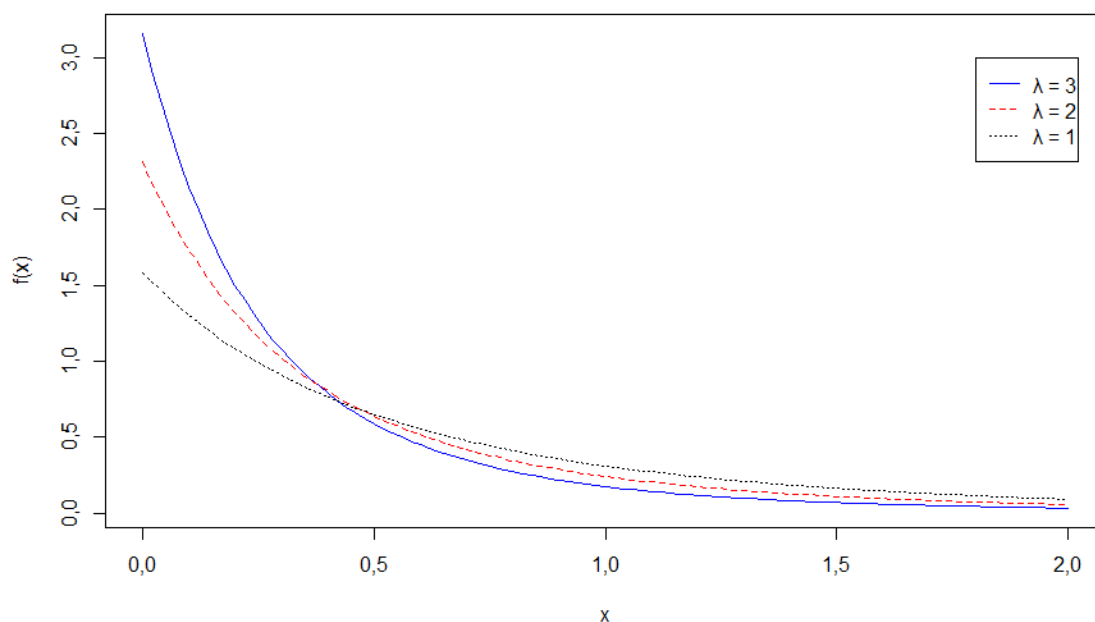
4.3.6 Exponential Poisson eloszlás (exppois)

Az exppois egy kétváltozós élettartam-eloszlás, amelyet Kus definiált (2007). Az eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye a következő:

$$f(x) = \frac{b\lambda \exp[-bx-\lambda+\lambda \exp(-bx)]}{1-\exp(-\lambda)} \quad (38)$$

$$F(x) = \frac{1-\exp[-\lambda+\lambda \exp(-bx)]}{1-\exp(-\lambda)} \quad (39)$$

17. ábra: Exppois-sűrűségfüggvény ($b = 1$ és $\lambda = 1, 2, 3$)



Forrás: (saját szimuláció)

Az eloszlás kockázatosított értékét és várható többletveszteségét a következő függvények adják meg: (Nadarajah, et al., 2015):

$$VaR_p(X) = -\frac{1}{b} \log \left\{ \frac{1}{\lambda} \log[1 - p + p \exp(-\lambda)] + 1 \right\} \quad (40)$$

$$ES_p(X) = -\frac{1}{bp} \int_0^p \log \left\{ \frac{1}{\lambda} \log[1 - v + v \exp(-\lambda)] + 1 \right\} dv \quad (41)$$

A fentiekben $x > 0$, $0 < p < 1$, $b > 0$ az első, $\lambda > 0$ pedig a második skálaparaméter (Nadarajah, et al., 2015).

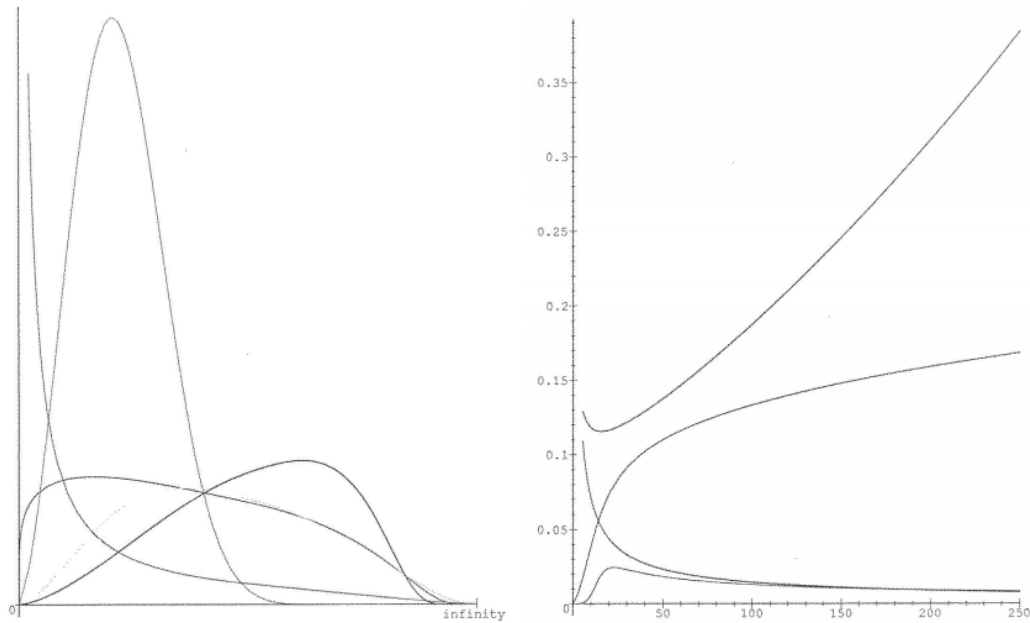
4.3.7 Generalized power Weibull eloszlás (genpowerweibull)

Az eloszlás a Weibull-eloszláscsalád egyik kiterjesztése (Nikulin & Haghighi, 2006). Az eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye a következő:

$$f(x) = a\theta x^{a-1} [1 + x^a]^{\theta-1} \exp\{1 - [1 + x^a]^\theta\} \quad (42)$$

$$F(x) = 1 - \exp\{1 - [1 + x^a]^\theta\} \quad (43)$$

18. ábra: Az $F(x)$ sűrűségfüggvény (bal oldalon) és tipikus meghibásodási ráták (jobb oldalon)



Forrás: (Nikulin & Haghghi, 2006, p. 1337)

Az eloszlás kockázatosított értékét és várható többletvesztését a következő függvények adják meg: (Nadarajah, et al., 2015):

$$VaR_p(X) = \{[1 - \log(1 - p)]^{1/\theta} - 1\}^{1/a} \quad (44)$$

$$ES_p(X) = \frac{1}{p} \int_0^p \{[1 - \log(1 - v)]^{1/\theta} - 1\}^{1/a} dv \quad (45)$$

A fentiekben $x > 0$, $0 < p < 1$, $a > 0$ az első, $\theta > 0$ pedig a második alakparaméter (Nadarajah, et al., 2015).

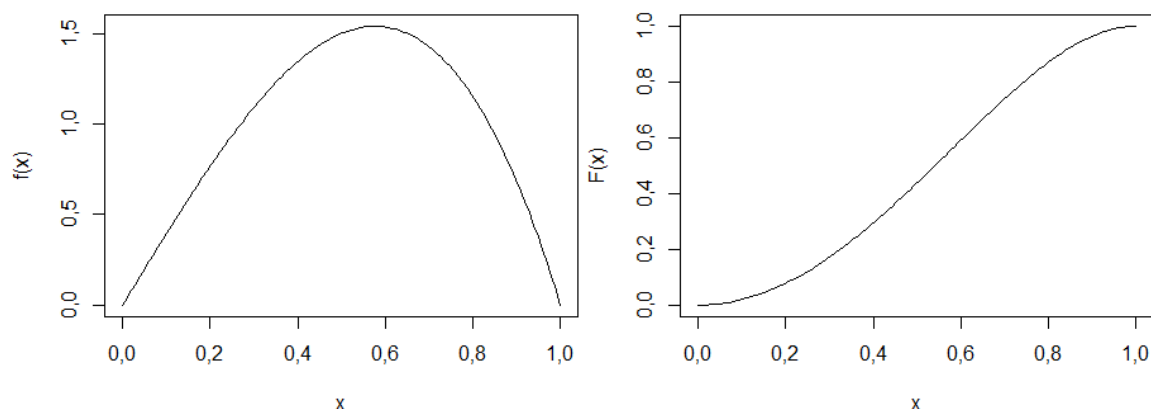
4.3.8 Kumaraswamy-eloszlás (kum)

Kumaraswamy (1980) hidrológiai valószínűségi változók leírására vezette be ezt az eloszlást, mivel a normál, log-normál, béta és az empirikus eloszlásokkal nem tudott olyan változókat megfelelően modellezni, mint a napi csapadékmennyiség. Az eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye a következő:

$$f(x) = abx^{a-1}(1 - x^a)^{b-1} \quad (46)$$

$$F(x) = 1 - (1 - x^a)^b \quad (47)$$

19. ábra: A Kumaraswamy-eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye a=2 és b=2 paraméterekkel



Forrás: (saját szerkesztés)

Az eloszlás kockázatosított értékét és várható többletveszteségét a következő függvények adják meg (Nadarajah, et al., 2015):

$$VaR_p(X) = [1 - (1 - p)^{1/b}]^{1/a} \quad (48)$$

$$ES_p(X) = \frac{1}{p} \int_0^p [1 - (1 - v)^{1/b}]^{1/a} dv \quad (49)$$

A fentiekben $0 < x < 1$, $0 < p < 1$, $a > 0$ az első, $b > 0$ pedig a második alakparaméter (Nadarajah, et al., 2015).

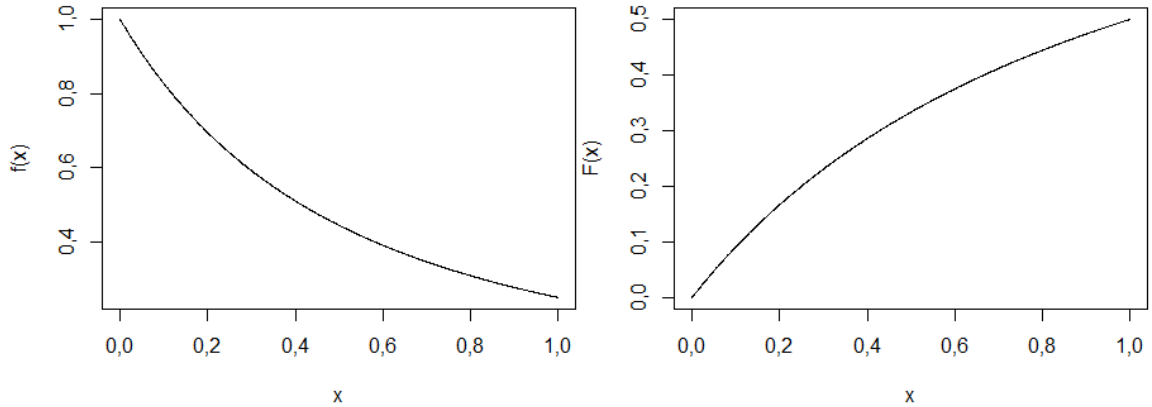
4.3.9 Kumaraswamy–Burr-eloszlás (kumburr7)

Az eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye a következő (Paranaiba, et al., 2013):

$$f(x) = \frac{abkcx^{c-1}}{(1+x^c)^{k+1}} [1 - (1 + x^c)^{-k}]^{a-1} \{1 - [1 - (1 + x^c)^{-k}]^a\}^{b-1} \quad (50)$$

$$F(x) = 1 - \{1 - [1 - (1 + x^c)^{-k}]^a\}^b \quad (51)$$

20. ábra A Kumaraswamy–Burr-eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye standard paraméterekkel



Forrás: (saját szimuláció)

Az eloszlás kockázatosított értékét és várható többletveszteségét a következő függvények adják meg: (Nadarajah, et al., 2015):

$$VaR_p(X) = \left[\left\{ 1 - \left[1 - (1 - p)^{1/b} \right]^{1/a} \right\}^{-1/k} - 1 \right]^{1/c} \quad (52)$$

$$ES_p(X) = \frac{1}{p} \int_0^p \left[\left\{ 1 - \left[1 - (1 - v)^{1/b} \right]^{1/a} \right\}^{-1/k} - 1 \right]^{1/c} dv \quad (53)$$

A fentiekben $0 < x$, $0 < p < 1$, $a > 0$ az első, $b > 0$ a második, $c > 0$ a harmadik, $k > 0$ pedig a negyedik alakparaméter (Nadarajah, et al., 2015).

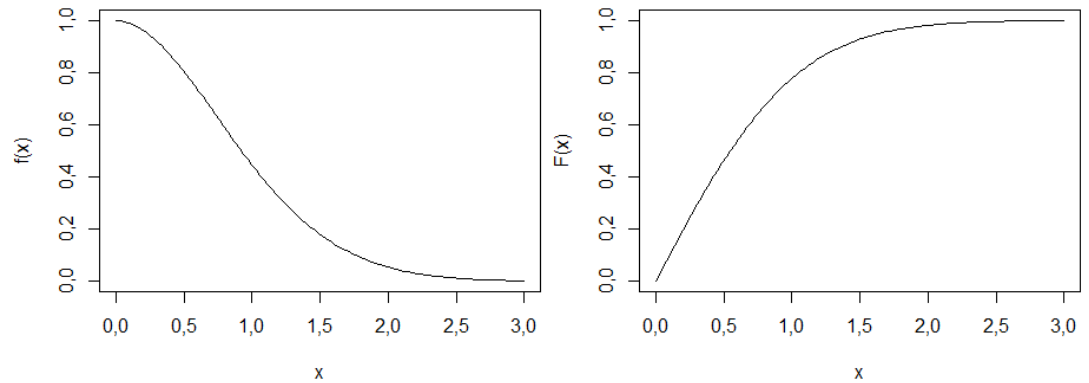
4.3.10 Linear failure rate (LFR)

Az LFR-eloszlást Bain vezette be az amerikai légierő által szponzorált kutatásában, ahol a motorok meghibásodásának modellezését vizsgálta (Bain, 1974). Az LFR-eloszlás az élettartam-elemzésre leggyakrabban használt Rayleigh- és az exponenciális eloszlásokhoz köthető. Az exponenciális eloszlás csak konstans meghibásodási arányt enged meg, míg a Rayleigh-eloszlás csak növekvő meghibásodási arányt. Az LFR-eloszlás ezen eloszlások általánosításával a csökkenő meghibásodási arányt is megengedi (Sarhan & Kundu, 2009).

$$f(x) = (a + bx) \exp(-ax - bx^2/2) \quad (54)$$

$$F(x) = 1 - \exp(-ax - bx^2/2) \quad (55)$$

21. ábra: Az LFR-eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye standard paraméterekkel



Forrás: (saját szimuláció)

Az eloszlás kockázatos értéke és várható többletvesztesége a következő (Nadarajah, et al., 2015):

$$VaR_p(X) = \frac{-a + \sqrt{a^2 - 2b \log(1-p)}}{b} \quad (56)$$

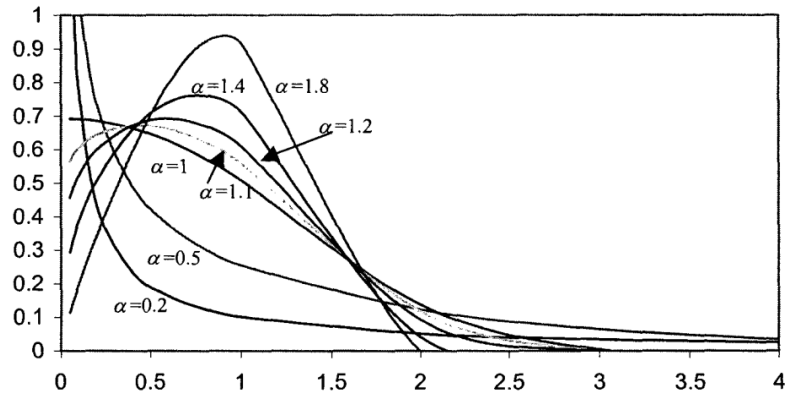
$$ES_p(X) = -\frac{a}{b} + \frac{1}{bp} \int_0^p \sqrt{a^2 - 2b \log(1-v)} dv \quad (57)$$

A fentiekben $x > 0$, $0 < p < 1$, $a > 0$ az első, $b > 0$ pedig a második skálaparaméter (Nadarajah, et al., 2015).

4.3.11 Loglog-eloszlás (loglog)

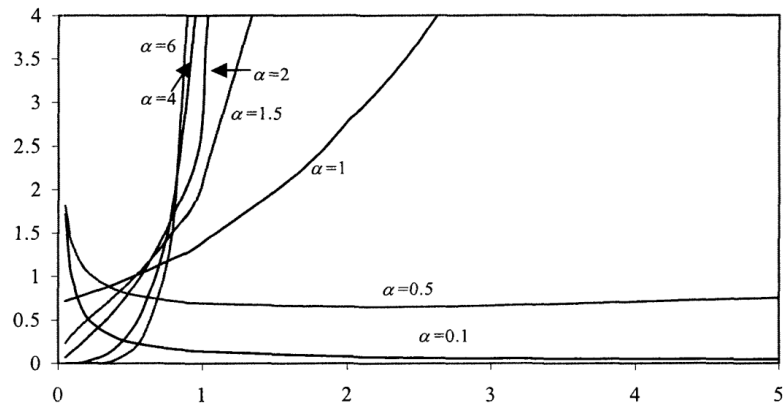
A Loglog-eloszlásfüggvény a megbízhatósággal kapcsolatos műszaki területeken az életciklus és a meghibásodási ráta modellezésére gyakran használt fürdőkádgörbe egyik változata. A görbe alakját az adja, hogy kezdetben a meghibásodás esélye a gyártási hibák miatt relatív magas, majd ezek kiküszöbölését követően drasztikusan visszaesik, majd lassan, folyamatosan növekszik a használati idő növekedésével. Pham (2002) tanulmányában egy új, úgynevezett „Vtub-alakú” kétparaméteres élettartam-eloszlásfüggvény használatát javasolja, amit Loglog-eloszlásnak nevez és helikopteralkatrészek meghibásodási idősorain tesztel.

22. ábra: Az $F(x)$ sűrűségfüggvény $\lambda = 2$ esetére



Forrás: (Pham, 2002, p. 3)

23. ábra: Meghibásodási ráta $\lambda = 2$ esetére



Forrás: (Pham, 2002, p. 4)

Az eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye a következő:

$$f(x) = \alpha \log(\lambda) x^{\alpha-1} \lambda x^\alpha e^{(1-\lambda x^\alpha)} \quad (58)$$

$$F(x) = 1 - e^{(1-\lambda x^\alpha)} \quad (59)$$

A Loglog-eloszlás kockázatos értéke és várható többletvesztesége a következő (Nadarajah, et al., 2015):

$$VaR_p(X) = \left\{ \frac{\log[1-\log(1-p)]}{\log \lambda} \right\}^{1/\lambda} \quad (60)$$

$$ES_p(X) = \frac{1}{p(\log \lambda)^{1/\lambda}} \int_0^p \{\log[1 - \log(1 - v)]\}^{1/\lambda} dv \quad (61)$$

A fentiekben $x > 0$, $0 < p < 1$, $\alpha > 0$ az első, $\lambda > 1$ pedig a második skálaparaméter (Nadarajah, et al., 2015).

4.3.12 Lomax-eloszlás (lomax)

Lomax (1954) az általa javasolt eloszlással vállalatok megszűnését modellezte. Az eloszlás sűrűségfüggvénye:

$$f(x) = \frac{a}{\lambda} \left(1 + \frac{x}{\lambda}\right)^{-a-1} \quad (62)$$

Az eloszlás kumulatív eloszlásfüggvénye:

$$F(x) = 1 - \left(1 + \frac{x}{\lambda}\right)^{-a} \quad (63)$$

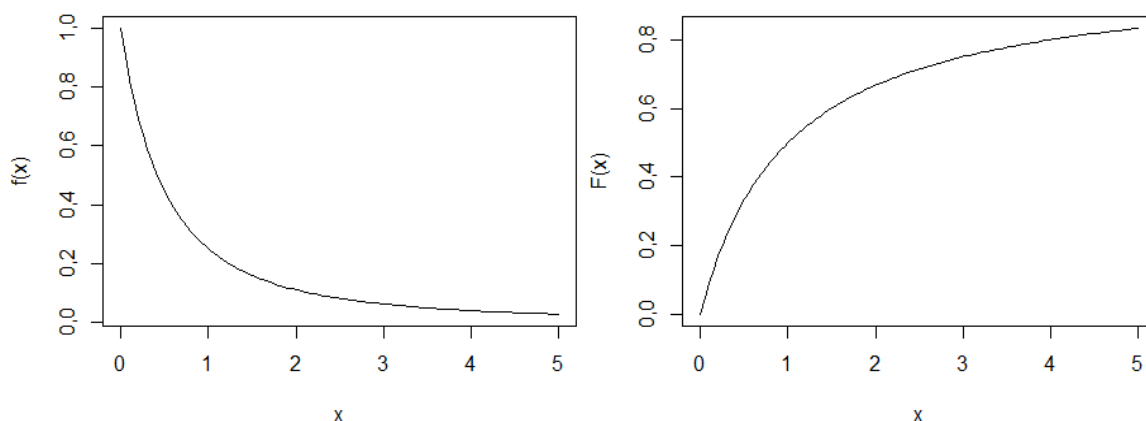
Az eloszlás kockázatosított értékét és várható többletveszteségét a következő képletek adják meg (Nadarajah, et al., 2015):

$$VaR_p(X) = \lambda \left[(1-p)^{-1/a} - 1 \right] \quad (64)$$

$$ES_p(X) = -\lambda + \frac{\lambda - \lambda(1-p)^{1-1/a}}{p - p/a} \quad (65)$$

A fentiekben $x > 0$, $0 < p < 1$, $a > 0$ az alakparaméter, $\lambda > 0$ pedig skálaparaméter (Nadarajah, et al., 2015).

24. ábra: A Lomax-eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye standard paraméterekkel



Forrás: (saját szimuláció)

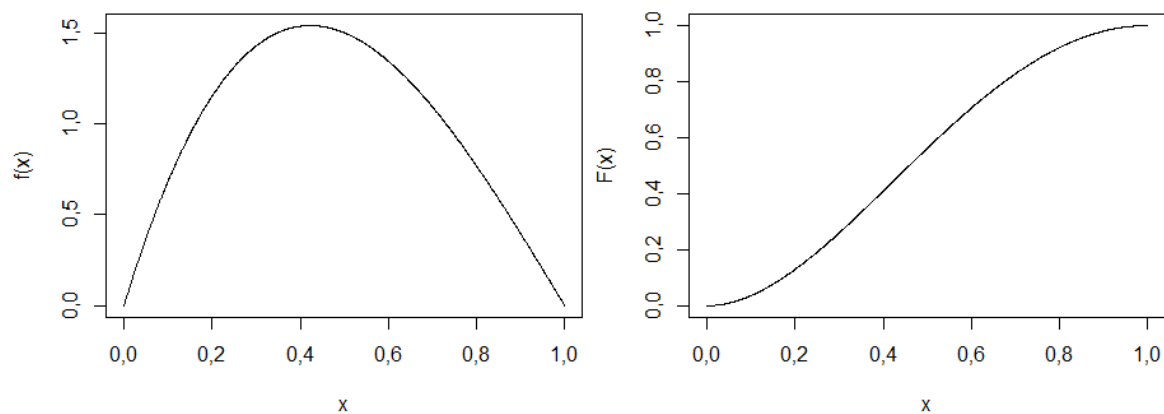
4.3.13 Topp–Leone-eloszlás (TL2)

Az eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye Topp és Leone alapján (1955):

$$f(x) = 2b[x(2 - x)]^{b-1}(1 - x) \quad (66)$$

$$F(x) = [x(2 - x)]^b \quad (67)$$

25. ábra A Topp–Leone-eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye $b = 2$ alakparaméterrel



Forrás: (saját szimuláció)

Az eloszlás kockázatosított értékét és várható többletveszteségét a következő képletek adják meg (Nadarajah, et al., 2015):

$$VaR_p(X) = 1 - \sqrt{1 - p^{1/b}} \quad (68)$$

$$ES_p(X) = 1 - \frac{b}{p} B_{p^{1/b}}\left(b, \frac{3}{2}\right) \quad (69)$$

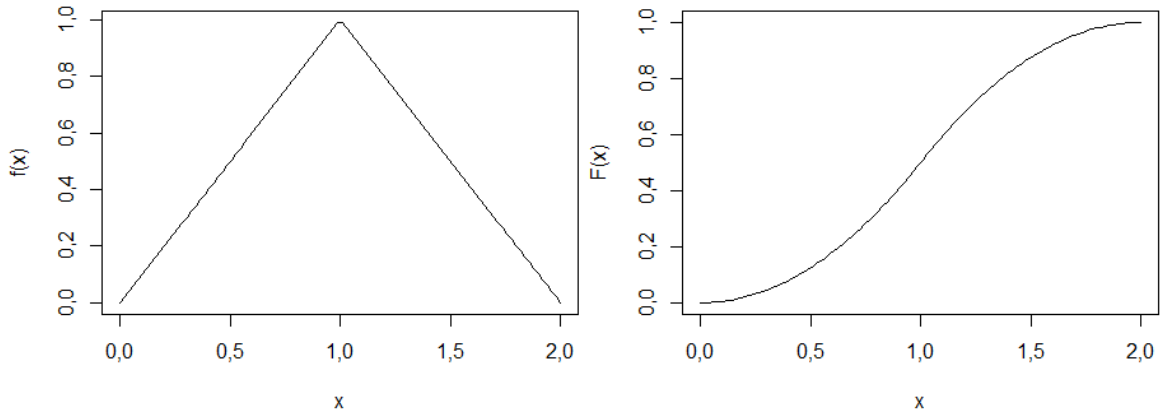
A fentiekben $x > 0$, $0 < p < 1$, $b > 0$ pedig az alakparaméter (Nadarajah, et al., 2015).

4.3.14 Triangular-eloszlás (triangular)

$$f(x) = \begin{cases} 0, & \text{ha } x < a \\ \frac{2(x-a)}{(b-a)(c-a)}, & \text{ha } a \leq x \leq c \\ \frac{2(b-x)}{(b-a)(b-c)}, & \text{ha } c < x \leq b \\ 0, & \text{ha } b < x \end{cases} \quad (70)$$

$$F(x) = \begin{cases} 0, & \text{ha } x < a \\ \frac{(x-a)^2}{(b-a)(c-a)}, & \text{ha } a \leq x \leq c \\ 1 - \frac{(b-x)^2}{(b-a)(b-c)}, & \text{ha } c < x \leq b \\ 1, & \text{ha } b < x \end{cases} \quad (71)$$

26. ábra: A Triangular-eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye standard paraméterekkel



Forrás: (saját szimuláció)

Az eloszlás kockázatosított értékét és várható többletveszteségét a következő képletek adják meg (Nadarajah, et al., 2015):

$$VaR_p(X) = \begin{cases} a + \sqrt{p(b-a)(c-a)}, & \text{ha } 0 < p < \frac{c-a}{b-a} \\ b - \sqrt{(1-p)(b-a)(b-c)}, & \text{ha } \frac{c-a}{b-a} \leq p < 1 \end{cases} \quad (72)$$

$$ES_p(X) = \begin{cases} a + \frac{2}{3}\sqrt{p(b-a)(c-a)}, & \text{ha } 0 < p < \frac{c-a}{b-a} \\ b + \frac{a-c}{p} + \frac{2(2c-a-b)}{3p} + 2\sqrt{(b-a)(b-c)}\frac{(1-p)^{3/2}}{3p}, & \text{ha } \frac{c-a}{b-a} \leq p < 1 \end{cases} \quad (73)$$

A fentiekben $a \leq x \leq b$, $0 < p < 1$, a az első, c a második, b pedig a harmadik helyparaméter, illetve $-\infty < a < c < b < \infty$ (Nadarajah, et al., 2015).

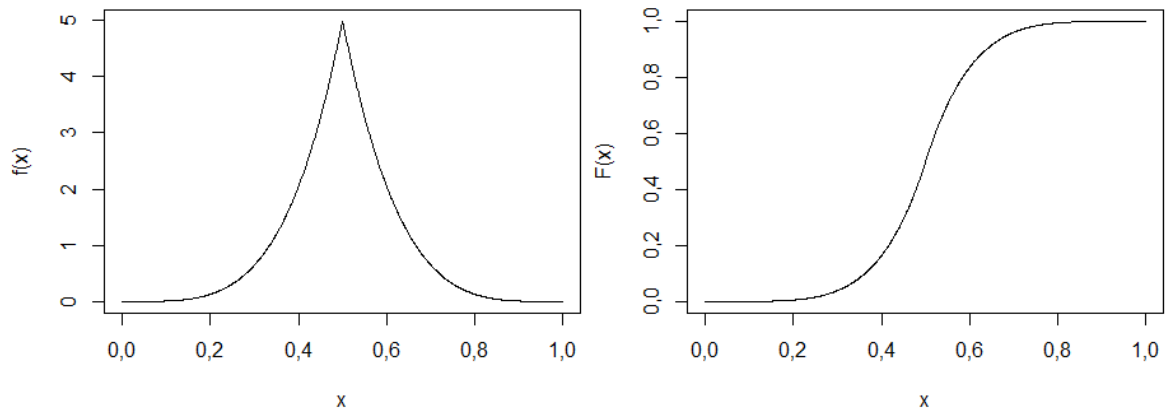
4.3.15 Two-sided power eloszlás (TSP)

A TSP-eloszláscsalád használatát van Dorp és Kotz kifejezetten a pénzügyi területen javasolják (van Dorp & Kotz, 2002). Az eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye a következő:

$$f(x) = \begin{cases} a \left(\frac{x}{\theta}\right)^{a-1} & \text{ha } 0 < x \leq \theta \\ a \left(\frac{1-x}{1-\theta}\right)^{a-1} & \text{ha } \theta < x < 1 \end{cases} \quad (74)$$

$$F(x) = \begin{cases} \theta \left(\frac{x}{\theta}\right)^a & \text{ha } 0 < x \leq \theta \\ 1 - (1-\theta) \left(\frac{1-x}{1-\theta}\right)^a & \text{ha } \theta < x < 1 \end{cases} \quad (75)$$

27. ábra: A Two-sided power eloszlás sűrűségfüggvénye és kumulatív eloszlásfüggvénye $a=5$ alakparaméterrel



Forrás: (saját szimuláció)

Az TSP-eloszlás kockázatos értéke és várható többletvesztesége a következő (Nadarajah, et al., 2015):

$$VaR_p(X) = \begin{cases} \theta \left(\frac{p}{\theta}\right)^{1/a} & \text{ha } 0 < p \leq \theta \\ 1 - (1-\theta) \left(\frac{1-p}{1-\theta}\right)^{1/a} & \text{ha } \theta < p < 1 \end{cases} \quad (76)$$

$$ES_p(X) = \begin{cases} \frac{a\theta}{a+1} \left(\frac{p}{\theta}\right)^{1/a} & \text{ha } 0 < p \leq \theta \\ 1 - \frac{\theta}{p} + \frac{a(2\theta-1)}{(a+1)p} + \frac{a(1-\theta)^2}{(a+1)p} \left(\frac{1-p}{1-\theta}\right)^{1+1/a} & \text{ha } \theta < p < 1 \end{cases} \quad (77)$$

A fentiekben $0 < x < 1$, $0 < p < 1$, $a > 0$ az alakparaméter és $-\infty < \theta < \infty$ pedig a helyparaméter (Nadarajah, et al., 2015).

4.4 További eloszlásalapú modellek

2. táblázat: További eloszlások (A-Be)

Az eloszlás angol megnevezése, rövidítése és sorszáma	Az eloszlás sűrűségfüggvénye
Asymmetric exponential power (aep) (78) (Zhu & Zinde-Walsh, 2009)	$f(x) = \begin{cases} \frac{\alpha}{\alpha^*} K(q_1) \exp\left[-\frac{1}{q_1} \left \frac{x}{2\alpha^*}\right ^{q_1}\right], & \text{ha } x \leq 0 \\ \frac{1-\alpha}{1-\alpha^*} K(q_2) \exp\left[-\frac{1}{q_2} \left \frac{x}{2-2\alpha^*}\right ^{q_2}\right], & \text{ha } x > 0 \end{cases}$
Generalized asymmetric Student's t (ast) (79) (Zhu és Galbraith (2010))	$f(x) = \begin{cases} \frac{\alpha}{\alpha^*} K(v_1) \left[1 + \frac{1}{v_1} \left(\frac{x}{2\alpha^*}\right)^2\right]^{-\frac{v_1+1}{2}}, & \text{ha } x \leq 0 \\ \frac{1-\alpha}{1-\alpha^*} K(v_2) \left[1 + \frac{1}{v_2} \left(\frac{x}{2(1-\alpha^*)}\right)^2\right]^{-\frac{v_2+1}{2}}, & \text{ha } x > 0 \end{cases}$
Asymmetric Laplace (asylaplace) (80) (Hinkley & Revankar, 1977)	$f_{\theta,\kappa,\tau}(x) = \frac{\sqrt{2}}{\tau} \frac{\kappa}{1+\kappa^2} \begin{cases} \exp\left(-\frac{\sqrt{2}\kappa}{\tau} x-\theta \right) & , \text{ha } x \geq \theta \\ \exp\left(-\frac{\sqrt{2}}{\tau\kappa} x-\theta \right) & , \text{ha } x < \theta \end{cases}$
Beta (betadist) (81)	$f(x) = \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)}$

Forrás: saját szerkesztés, (Nadarajah, et al., 2015) alapján

3. táblázat: További eloszlások (Bu – Ex)

Az eloszlás angol megnevezése, rövidítése és sorszáma	Az eloszlás sűrűségfüggvénye
Burr (burr) (82) (Burr, 1942)	$f(x) = \frac{ba^b}{x^{b+1}} [1 + (x/a)^{-b}]^{-2}$
Cauchy (Cauchy) (82)	$f(x) = \frac{1}{\pi} \frac{\sigma}{(x - \mu)^2 + \sigma^2}$
Compound Laplace gamma (clg) (83)	$f(x) = \frac{ab}{2} \{1 + b x - \theta \}^{-(a+1)}$
Complementary beta (compbeta) (84) (Jones, 2002)	$f(x) = B(a, b) \{I_x^{-1}(a, b)\}^{1-a} \{1 - I_x^{-1}(a, b)\}^{1-b}$
Exponentiated exponential (expexp) (85) (Gupta & Kundu, 1999) (Gupta & Kundu, 2001)	$f(x) = a\lambda \exp(-\lambda x) [1 - \exp(-\lambda x)]^{a-1}$

Forrás: saját szerkesztés, (Nadarajah, et al., 2015) alapján

4. táblázat: További eloszlások (Ex – G)

Az eloszlás angol megnevezése, rövidítése és sorszáma	Az eloszlás sűrűségfüggvénye
Exponentiated logistic distribution (86) (explogis)	$f(x) = (a/b) \exp(-x/b) [1 + \exp(-x/b)]^{-a-1}$
F (f) (87) (Pintér & Rappai, 2007)	$f(x) = \frac{1}{B\left(\frac{d_1}{2}, \frac{d_2}{2}\right)} \left(\frac{d_1}{d_2}\right)^{\frac{d_1}{2}} x^{\frac{d_1}{2}-1} \left(1 + \frac{d_1}{d_2} x\right)^{-\frac{d_1+d_2}{2}}$
Fréchet (frechet) (88) (Fréchet, 1927)	$f(x) = \frac{\alpha \sigma^\alpha}{x^{\alpha+1}} \exp\left\{-\left(\frac{\sigma}{x}\right)^\alpha\right\}$
Gamma (Gamma) (89)	$f(x) = \frac{b^a x^{a-1} \exp(-bx)}{\Gamma(a)}$
Gumbel (gumbel) (90) (Gumbel, 1954)	$f(x) = \frac{1}{\sigma} \exp\left(\frac{\mu-x}{\sigma}\right) \exp\left[-\exp\left(\frac{\mu-x}{\sigma}\right)\right]$

Forrás: saját szerkesztés, (Nadarajah, et al., 2015) alapján

5. táblázat: További eloszlások (K)

Az eloszlás angol megnevezése, rövidítése és sorszám	Az eloszlás sűrűségfüggvénye
Kumaraswamy exponential (kumexp) (91) (Cordeiro & de Castro, 2011)	$f(x) = ab\lambda \exp(-\lambda x) [1 - \exp(-\lambda x)]^{a-1} \{1 - [1 - \exp(-\lambda x)]^a\}^{b-1}$
Kumaraswamy Gumbel (kumgumbel) (92) (Cordeiro, et al., 2012a)	$f(x) = \frac{ab}{\sigma} \exp\left(\frac{\mu - x}{\sigma}\right) \exp\left[-a \exp\left(\frac{\mu - x}{\sigma}\right)\right] \left\{1 - \exp\left[-a \exp\left(\frac{\mu - x}{\sigma}\right)\right]\right\}^{b-1}$
Kumaraswamy half normal eloszlás (kumhalfnorm) (93) (Cordeiro, et al., 2012b)	$f(x) = \frac{2ab}{\sigma} \phi\left(\frac{x}{\sigma}\right) \left[2\Phi\left(\frac{x}{\sigma}\right) - 1\right]^{a-1} \left\{1 - \left[2\Phi\left(\frac{x}{\sigma}\right) - 1\right]^a\right\}^{b-1}$
Kumaraswamy normal eloszlás (kumnorm) (94) (Cordeiro & de Castro, 2011)	$f(x) = \frac{ab}{\sigma} \phi\left(\frac{x - \mu}{\sigma}\right) \Phi^{a-1}\left(\frac{x - \mu}{\sigma}\right) \left[1 - \Phi^a\left(\frac{x - \mu}{\sigma}\right)\right]^{b-1}$

Forrás: saját szerkesztés, (Nadarajah, et al., 2015) alapján

6. táblázat: További eloszlások (L-Lo)

Az eloszlás angol megnevezése, rövidítése és sorszáma	Az eloszlás sűrűségfüggvénye
Laplace (laplace) (95)	$f(x) = \frac{1}{2\sigma} \exp\left(-\frac{ x - \mu }{\sigma}\right)$
Libby-Novick beta (LNbeta) (96) (Libby & Novick, 1982)	$f(x) = \frac{\lambda^a x^{a-1} (1-x)^{b-1}}{B(a,b)[1 - (1-\lambda)x]^{a+b}}$
Log-Cauchy (logcauchy) (97)	$f(x) = \frac{1}{x\pi} \frac{\sigma}{(\log x - \mu)^2 + \sigma^2}$
Log Gamma (loggamma) (98) (Consul & Jain, 1971)	$f(x) = \frac{a^r x^{a-1} (-\log x)^{r-1}}{\Gamma(r)}$
Logistic exponential (logisexp) (99) (Lan & Leemis, 2008)	$f(x) = \frac{a\lambda \exp(\lambda x) [\exp(\lambda x) - 1]^{a-1}}{\{1 + [\exp(\lambda x) - 1]^a\}^2}$

Forrás: saját szerkesztés, (Nadarajah, et al., 2015) alapján

7. táblázat: További eloszlások (Lo-Ma)

Az eloszlás angol megnevezése, rövidítése és sorszáma	Az eloszlás sűrűségfüggvénye
Logistic Rayleigh (logisrayleigh) (100) (Lan & Leemis, 2008)	$f(x) = a\lambda x \exp(\lambda x^2/2) [\exp(\lambda x^2/2) - 1]^{a-1} \{1 + [\exp(\lambda x^2/2) - 1]^a\}^{-2}$
Logistic (logistic) (101)	$f(x) = \frac{1}{\sigma} \exp\left(-\frac{x-\mu}{\sigma}\right) \left[1 + \exp\left(-\frac{x-\mu}{\sigma}\right)\right]^{-2}$
Lognormal (lognorm) (102)	$f(x) = \frac{1}{\sigma x} \phi\left(\frac{\log x - \mu}{\sigma}\right)$
McGill Laplace eloszlás (Mlaplace) (103) (McGill, 1962)	$f(x) = \begin{cases} \frac{1}{2\psi} \exp\left(\frac{x-\theta}{\psi}\right), & \text{ha } x \leq \theta \\ \frac{1}{2\phi} \exp\left(\frac{\theta-x}{\phi}\right), & \text{ha } x > \theta \end{cases}$
Marshall-Olkin Weibull eloszlás (moweibull) (104) (Marshall & Olkin, 1997)	$f(x) = b\lambda^b x^{b-1} \exp[(\lambda x)^b] \{\exp[(\lambda x)^b] - 1 + a\}^{-2}$

Forrás: saját szerkesztés, (Nadarajah, et al., 2015) alapján

8. táblázat: További eloszlások (Mr-S)

Az eloszlás angol megnevezése, rövidítése és sorszáma	Az eloszlás sűrűségfüggvénye
McDonald-Richards beta (MRbeta) (105) (McDonald & Richards, 1987a) (McDonald & Richards, 1987b)	$f(x) = \frac{x^{ar-1}(bq^r - x^r)^{b-1}}{(bq^r)^{a+b-1}B(a, b)}$
Poiraud-Casanova-Thomas-Agnan Laplace (PCTAlaplace) (106) (Poiraud-Casanova & Thomas-Agnan, 2000)	$f(x) = \begin{cases} a(1-a) \exp\{(1-a)(x-\theta)\}, & \text{ha } x \leq \theta \\ a(1-a) \exp\{a(\theta-x)\}, & \text{ha } x > \theta \end{cases}$
Perks (perks) (107) (Perks, 1932)	$f(x) = \frac{a \exp(bx) [1+a]}{[1+a \exp(bx)]^2}$
Reflected gamma (rgamma) (108) (Borgi, 1965)	$f(x) = \frac{1}{2\phi\Gamma(a)} \left \frac{x-\theta}{\phi} \right ^{a-1} \exp\left\{-\left \frac{x-\theta}{\phi} \right \right\}$
Stacy (stacygamma) (109) (Stacy, 1962)	$f(x) = \frac{cx^{c\gamma-1} \exp[-(x/\theta)^c]}{\theta^{c\gamma}\Gamma(\gamma)}$

Forrás: saját szerkesztés, (Nadarajah, et al., 2015) alapján

9. táblázat: További eloszlások (T-X)

Az eloszlás angol megnevezése, rövidítése és sorszáma	Az eloszlás sűrűségfüggvénye
Students't (t4) (110)	$f_X(x) = \frac{n^{-1/2}}{\sigma B(n/2, 1/2)} \left(1 + \frac{(x - \mu)^2}{n} \right)^{-(n+1)/2}$
Weibull (weibull) (111) (Weibull, 1951)	$f(x) = \frac{\alpha x^{\alpha-1}}{\sigma^\alpha} \exp\left\{-\left(\frac{x}{\sigma}\right)^\alpha\right\}$
Xie (xie) (112) (Xie, et al., 2002)	$f(x) = \lambda b \left(\frac{x}{a}\right)^{0-1} \exp[(x/a)^b] \exp(\lambda a) \exp\{-\lambda a \exp[(x/a)^b]\}$

Forrás: saját szerkesztés, (Nadarajah, et al., 2015) alapján

4.5 Az idősorok transzformációja

Az eloszlásfüggvényen alapú modellek gyakran az idősorok transzformációját teszik szükségesé, azon egyszerű oknál fogva, hogy bizonyos eloszlások csak bizonyos tartományokban értelmezettek. Ezekben az esetekben az eloszlás paraméterbecslésére használt maximum likelihood algoritmus nem konvergál, ha az illesztéshez használt idősor részben vagy egészben a megengedett tartományon kívül esik. A probléma megkerülhető, ha a bemeneti adatokat az aktuális modell által értelmezett tartományra vetítjük, majd a modell által becsült várható többletveszteség mutató értékét a korábbi művelet inverzével állítjuk vissza.

$$z = \frac{x-l}{s} \quad (113)$$

A fentiekben az x az eredeti idősor, a z a transzformált idősor, az l (location) és az s (scale) értékei pedig a transzformációhoz szükséges paraméterek.

$$ES_x = s \cdot ES_z + l \quad (114)$$

ES_x a becslési modell végeredménye, ES_z pedig a transzformált idősor alapján megkapott várható többletvesztés-becslés. Az empirikus elemzés során az idősorok transzformációja minden becslési időszakra külön-külön történik.

10. táblázat: A location paraméter kiszámításának módjai modellenként

A location paraméter (l) kiszámítása	Modellek	Transzformált idősor (z_i)	Sorszám
$l = \min(x)$	betadist, f, frechet, gamma, weibull	$z_i \geq 0$	(115)
$l = \min(x) - \varepsilon_1$ $\varepsilon_1 = 0,0001$	burr, burr7, chen, compbeta, expexp, expext, explog, exppois, genpowerweibull, kum, kumburr7, kumexp, kumhalfnorm, lfr, LNbeta, loggamma, logisexp, logisrayleigh, logcahcy, loglog, lognorm, lomax, moweibull, MRbeta, stacygamma, tsp, tl2, xie	$z_i > \varepsilon_1 > 0$	(116)
$l = \min(x) - \varepsilon_2$ $\varepsilon_2 = 0,01$	invgamma	$z_i > \varepsilon_2 > 0$	(117)
$l = 0$	Az összes, fent nem megjelölt eloszlás modellje	$\infty > z_i > -\infty$	(118)

Forrás: (saját szerkesztés)

A legegyszerűbb az az eset, amikor $z_i \geq 0$ tartományra transzformáljuk az idősort. A 10. táblázatból látszik, hogy számos esetben nem elég a nemnegatív tartományra transzformálni az idősort, az is fontos, hogy ne legyen olyan időszak, amikor a transzformált hozam pontosan nulla, mert ezekben az időszakokban az eloszlások olyan szélsőséges értékeket adnak, amelyeket a maximum likelihood algoritmus nem tud

kezelni. Ezekben az esetekben az idősorokat egy további ($\varepsilon_1, \varepsilon_2$) korrekciós tényezővel kell eltolni pozitív irányban. Hasonló okokból válik szükségessé a scale paraméter használata. Bizonyos eloszlások nem értelmezhetőek az 1 feletti tartományban, ilyenkor az idősor skálázása jelent megoldást. A location paraméterhez hasonlóan itt is nehezen kezelhető az értelmezési tartomány szélső értéke, így az ε_3 paraméter segítségével a transzformáció valamivel kisebb érték felhasználásával történik.

11. táblázat: A scale paraméter kiszámításának módjai modellenként

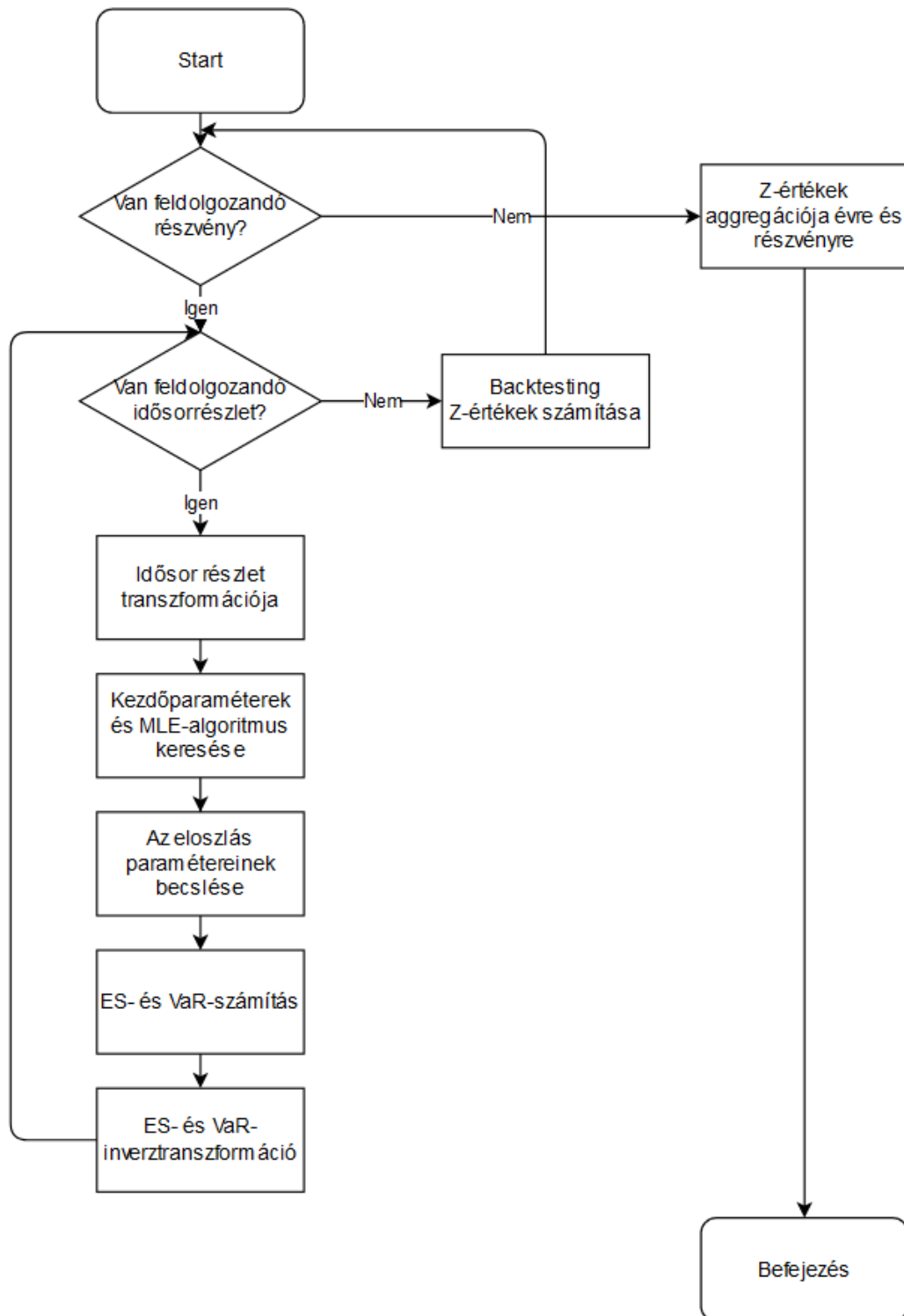
A scale paraméter (l) kiszámítása	Modellek	Transzformált idősor (z)	Sorszám
$s = \begin{cases} \max(x)/\varepsilon_3, & \text{ha } \max(x) > \varepsilon_3 \\ 1, & \text{ha } \max(x) \leq \varepsilon_3 \end{cases}$ $\varepsilon_3 = 0,99$	arcsine, betadist, compbeta, tsp, TL2, kum, LNbeta, loggamma, MRbeta, xie	$1 > \varepsilon_3 \geq z_i$	(119)
$s = 1$	Az összes, fent nem megjelölt eloszlás modellje.	$\infty > z_i > -\infty$	(120)

Forrás: (saját szerkesztés)

A location és a scale paraméter együttes használatával az ezt igénylő modellek olyan transzformált idősorokon futtathatóak, amelyek ráhagyással ($\varepsilon_1, \varepsilon_2$ és ε_3 függvényében) 0 és 1 közé esnek, azaz:

$$1 > \varepsilon_3 \geq z_i > \varepsilon_1 > 0 \quad (120)$$

28. ábra: A parametrikus modellek működésének folyamatábrája



Forrás: (saját szerkesztés)

4.6 GARCH (feltételes modellek)

4.6.1 Az ARCH-folyamat

Az Engle (1982) által bevezetett ARCH⁵-folyamatok megjelenése óta megkülönböztethetünk feltételes és feltétel nélküli varianciamodelleket. A feltételes modellek abban különböznek a feltétel nélküliektől, hogy megengedik, hogy a variancia az előző időszakok hibáinak függvényében változzon.

Az ARCH-folyamat definíciója a következő:

Legyen $(Z_t)_{t \in \mathbb{Z}}$ FAE(0,1), azaz független és azonos eloszlású fehér zaj. Az $(X_t)_{t \in \mathbb{Z}}$ akkor ARCH(p)-folyamat, ha a következő egyenletek írják le:

$$X_t = \sigma_t Z_t \quad (121)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 \quad (122)$$

A fentiekben $\alpha_0 > 0$ és $\alpha_i \geq 0, i = 1, \dots, p$.

Az ARCH-modell autoregresszív, mivel X_t egyértelműen függ X_{t-1} értékétől és bír a feltételes heteroszkedaszticitás tulajdonságával, mivel a feltételes variancia folyamatosan változik. A modell varianciacsoportokat generál, azaz az alacsony varianciájú „nyugodt” és a magas varianciájú „változékony” időszakok váltják egymást (Nagy & Balogh, 2013). Ennek oka az, hogy ha $|X_{t-1}|, \dots, |X_{t-p}|$ értékei közül egy vagy több magas, akkor az X_t értékét is egy nagy szórású eloszlásból fogjuk megkapni (McNeil, et al., 2015).

4.6.2 Standard GARCH-modell (SGarch)

A GARCH⁶-folyamatok általánosított ARCH-folyamatok abban az értelemben, hogy a variancia függhet korábbi időszakok értékeitől és varianciájától is (Bollerslev, 1986).

⁵ autoregressive conditional heteroskedasticity, azaz autoregresszív feltételes heteroszkedaszticitás

⁶ generalized autoregressive conditional heteroskedasticity, azaz általános autoregresszív feltételes heteroszkedaszticitás

A GARCH-folyamatot a következő definíció írja le:

Legyen $(Z_t)_{t \in \mathbb{Z}}$ FAE(0,1). Az $(X_t)_{t \in \mathbb{Z}}$ akkor GARCH(p,q)-folyamat, ha a következő egyenletek írják le:

$$X_t = \sigma_t Z_t \quad (123)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 \quad (124)$$

ahol $\alpha_0 > 0$ és $\alpha_i \geq 0, i = 1, \dots, p, \beta_j \geq 0, j = 1, \dots, q$.

A modell p és q paraméterei azt adják meg, hogy hány korábbi X (p) és variancia (q) értéket vesz figyelembe a modell.

4.6.3 Exponential GARCH (Egarch)

Nelson (1991) a GARCH-modellben azt kifogásolja, hogy a GARCH három fő probléma miatt nem ideális eszköz árazási esetekben. Az első, hogy a hozamok és a jövőbeni volatilitás negatív korrelációt mutat, ezt pedig a GARCH-modellek nem engedik. Nelson második problémaként a GARCH-paraméter korlátait említi, amely a feltételes variancia folyamatdinamikájának megtörését eredményezheti. A harmadik probléma az, hogy nehéz értelmezni, hogy a sokkok mennyire hosszú távon hatnak a feltételes varianciára.

$$\ln(\sigma_t^2) = \alpha_t + \sum_{k=1}^{\infty} \beta_k g(z_{t-k}), \quad \beta_1 \equiv 1 \quad (125)$$

$$g(z_t) \equiv \theta z_t + \gamma[|z_t| - E|z_t|] \quad (126)$$

A fenti képletből jól látható, hogy a γ paraméter szabályozza, hogy a z_t eltérése a várható értékétől mekkora növekedést eredményez a varianciában. A θ paraméter azt teszi lehetővé, hogy a modell aszimmetrikus legyen, azaz máshogy reagáljon a negatív eltérésekre, mint a pozitívokra (Hamilton, 1994).

4.6.4 GjrGarch

Glosten és szerzőtársai (1993) olyan GARCH-modell használatát javasolják, amely megengedi, hogy a hozamok pozitív és negatív elmozdulása különböző hatással legyen a feltételes volatilitásra. A GjrGarch-modellt a következő összefüggés definiálja:

$$\sigma_t^2 = K + \delta\sigma_{t-1}^2 + \alpha\epsilon_{t-1}^2 + \phi\epsilon_{t-1}^2 I_{t-1} \quad (127)$$

A fentiekben $I_{t-1} = 0$, ha $\epsilon_{t-1} \geq 0$ és $I_{t-1} = 1$, ha $\epsilon_{t-1} < 0$

4.7 Quantile regression

4.7.1 A kvantilisregresszió bemutatása

A kvantilisregresszió alapját képező elképzelések a 18. századi Boscovich (Stigler, 1984) és a 19. századi Edgeworth (1888) munkájáig vezethetőek vissza. Boscovich a Föld ellipszoid alakját kutatta, és számításait az abszolút hiba minimalizálásával végezte el oly módon, hogy az általa alkalmazott eljárás a súlyozott medián számításának tekinthető. Egy évszázaddal később Edgeworth folytatta Boscovich munkáját, hangsúlyozva, hogy az átlag használata csak akkor ad optimális eredményt, ha a megfigyelések ugyanabból a normális eloszlásból származnak. Abban az esetben viszont amikor a megfigyelések különböző eloszlásból – például különböző szórású eloszlásokból – származnak a medián használata könnyen jobb eredményt adhat az átlagnál. Edgeworth az abszolút hibát és a mediánt a lineáris regresszióba illesztve használta, ami a kvantilisregresszió alapját képezi. Az Edgeworth által javasolt becslési eljárás meglehetősen nehezen megfogható volt, ami nagyban akadályozta a módszer elterjedését. A kvantilisregresszió számítása csak jóval később, a lineáris programozási eljárás használatával vált egyszerűvé és hatékonyvá (Koenker, 2005).

A kvantilisregresszió használata alapvetően két problémára nyújt megoldást. Egyrészt az átlag használatát a mediánnal helyettesíti, mivel a medián számos helyzetben alkalmasabb eredményt ad (Koenker, 2005). A második problémát az jelenti, hogy az egyetlen középpértéket megadó általános regressziós eljárások nem adnak teljes képet. Ezt a hiányosságot Mosteller és Tukey (1977) a következőképpen fogalmazzák meg könyvükben:

“A regressziós görbe összefoglalja az x változóhoz tartozó eloszlások átlagait. Ezen túlmenően kiszámíthatnánk az eloszlások különböző százalékpontjaihoz tartozó regressziós görbéket azért, hogy teljesebb képet kapjunk a halmazról. Többnyire ez nem történik meg, így a regresszió gyakran meglehetősen hiányos képet ad. Ahogy az átlag nem ad teljes képet az eloszlásról, úgy a regressziós görbe sem ad teljes képet az eloszlások halmazáról”

Koenker (2005) könyvében kiemeli, hogy ahhoz hasonlóan, ahogy teljesebbé teszi egy eloszlás megértését az, ha az átlagon túl a szórását, ferdeségét vagy a csúcosságát is figyelembe vesszük, az egyetlen középértéken alapuló regresszió kiegészítése is szükséges. A kvantilisregresszió az eloszlások leírását úgy teszi teljesebbé, hogy a középérték megadásán túl a különböző kvantilisekhez tartozó összefüggéseket is megmutatja. A kvantilisregresszió – angol nyelven quantile regression (QR) – kifejezést Koenker és Bassett (1978) használta először. Koenker és Bassett tanulmánya a statisztikai eljárások egy új osztályát vezeti be.

4.7.2 A kvantilisregresszió módszertana

Egy Y valószínűségi változó kvantilisfüggvénye a változó eloszlásfüggvényének inverze. Ennek megfelelően a feltételes kvantilisfüggvény a feltételes eloszlásfüggvény inverze (Xiao, et al., 2015).

$$Q_Y(\tau|X) = F_Y^{-1}(\tau|X) = \inf\{y: F_Y(y|X) \geq \tau\} \quad (128)$$

A fentiekben $F_Y(y|X) = P(Y \leq y|X)$.

A kvantilisregresszió hasonló eljárás a lineáris regresszióhoz alkalmazott legkisebb négyzetek módszeréhez, azzal a különbséggel, hogy ez utóbbi a függő változó átlagát határozza meg. Ezzel szemben a QR-eljárás az abszolút eltérések minimalizálásával a medián, vagy más tetszőleges kvantilis meghatározására alkalmazható. Koenker és Bassett (1978) a módszer bemutatása mellett kiemelik, hogy a QR használata jobb eredményeket ad a legkisebb négyzetek módszerénél olyan lineáris modellek esetén, ahol a hibák nem normális eloszlásúak.

A θ -hoz és (y_1, \dots, y_n) mintához tartozó regresszió kvantilise a következő (Koenker & D'Orey, 1987)⁷:

$$R_\theta(\beta) = \min_{\beta} \sum_{i=1}^n \rho_\theta(y_i - x_i\beta) \quad (129)$$

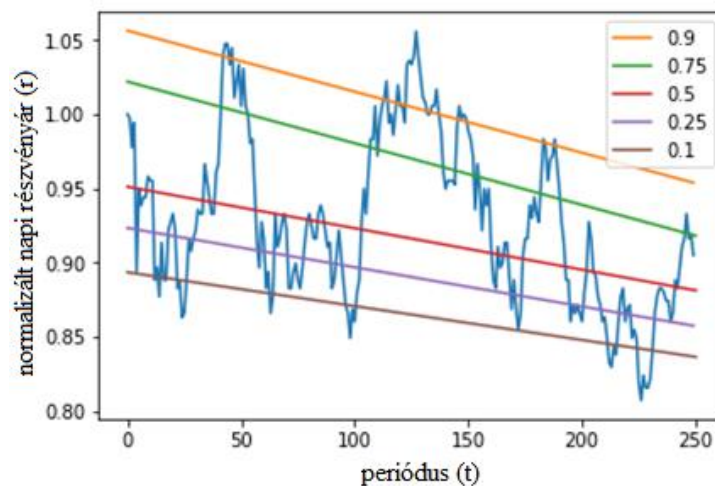
$$\rho_\theta(u) = \begin{cases} \theta u & u \geq 0 \\ (1 - \theta)u & u < 0 \end{cases} \quad (130)$$

$$\theta \in (0,1)$$

4.7.3 A kvantilisregresszió bemutatása egy példán keresztül

A 29. ábra az Apple normalizált napi részvényárának (r) alakulását, illetve öt különböző kvantilis értékét mutatja az 1988-as évre. A mediánál ($\theta = 0,5$) a megfigyelések 50%-a az egyenes felett, a maradék 50%-a pedig az egyenes alatt található. A számítás során a modell egy konstans elemet használ. A függő változó a normalizált részvényár, a független változó pedig a periódus (t).

29. ábra: Apple Inc (AAPL) napi részvényára és a kvantilisek



Forrás: (saját számítás)

⁷ Az R_θ képletében nincs szükség konstans tagra. A kvantilisregresszió módszertanának áttekintéséhez ld. (Koenker, 2005, p. 10).

4.7.4 A kvantilisregresszió, mint lineáris programozási modell

A regresszió kvantilis definíciójában megadott minimalizációs probléma megoldható egy lineáris programozási modell segítségével. A lineáris programozási feladat egységes, kanonikus alakja a következő (Komáromi, 2002):

$$\begin{aligned} \max_x \quad & \mathbf{c}\mathbf{x} \\ \text{Az} \quad & \mathbf{Ax} = \mathbf{b} \\ \mathbf{x} \quad & \geq 0 \end{aligned} \tag{131}$$

A egy $m \times n$ méretű mátrix, \mathbf{b} egy m komponensű vektor és \mathbf{c} pedig egy n komponensű vektor. A lineáris programozási feladat általános formájában tartalmazhat egyenlőtlenségi jeleket, negatív értékű változókat és minimalizálási célfüggvényt, de az ilyen típusú problémák mind átalakíthatók a kanonikus formára.

Koener (2005) a kvantilis regressziós problémát a lenti lineáris programozási feladattal fejezi ki.

$$\min_{(\boldsymbol{\beta}, \mathbf{u}, \mathbf{v}) \in \mathbb{R}^p \times \mathbb{R}_+^{2n}} \{ \theta \mathbf{1}_n^T \mathbf{u} + (1 - \theta) \mathbf{1}_n^T \mathbf{v} \} \tag{132}$$

$$X\boldsymbol{\beta} + \mathbf{u} - \mathbf{v} = \mathbf{y}$$

A fentiekben θ az előre megválasztott kvantilis, \mathbf{u} a hiba nemnegatív része, \mathbf{v} a hiba negatív része, $\mathbf{1}_n^T$ egy n elemű csupa egyesekből álló sorvektor, X a független változók $n \times p$ méretű mátrixa, $\boldsymbol{\beta}$ pedig a független változók együtthatóit tartalmazó vektor. A célfüggvényből látható, hogy a modell az abszolút hiba minimalizálásával adja meg a független változók együtthatóit ($\boldsymbol{\beta}$). A θ kvantilis megválasztása pedig azt befolyásolja, hogy pozitív hibákat és a negatív hibákat milyen súllyal veszi figyelembe az optimalizáció. $\theta = 0,5$ esetén mindkét hiba azonos súllyal bír, így az optimalizáció eredménye a medián. Ettől eltérő θ mellett az optimalizáció eredménye egy olyan hipersík lesz, amelyet a megfigyelések hozzávetőlegesen $1-\theta$ arányban fognak meghaladni és θ arányban lesznek a hipersík alatt.

Tegyük fel, hogy rendelkezésre áll n darab megfigyelés, amely alapján elvégezhető a lineáris regressziós modell illesztése, azaz a β -együtthatók kiszámítása. A megfigyelések alapján rendelkezésre áll a független változók megfigyeléseit tartalmazó mátrix (X) és a hozzájuk tartozó függő változók megfigyeléseit tartalmazó vektor (y):

$$X := \begin{bmatrix} \mathbf{x}_1^T \\ \vdots \\ \mathbf{x}_n^T \end{bmatrix} \quad (133)$$

$$\mathbf{y} := \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \quad (134)$$

A fentiekben \mathbf{x}_i^T a független változók i . megfigyelését tartalmazó vektor y_i pedig az i . periódusban megfigyelt függő változó értéke. Amennyiben a modell egy konstans tagot is tartalmaz, úgy a konstans tagot is fel kell venni a megfigyelések közé, például az X mátrix első oszlopába. Ebben az esetben $x_{i,1} = 1 \quad \forall i$.

$$X := \begin{bmatrix} 1 & x_{1,2} & \cdots & x_{1,k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n,2} & \cdots & x_{n,k} \end{bmatrix} \quad (135)$$

Ez alapján a lineáris regressziós modell illesztéséhez szükséges A mátrixot, illetve a \mathbf{b} és \mathbf{c} vektorokat a következő módon kaphatjuk meg (Hybel, 2018).

$$A := [X, -X, I_n, -I_n] \quad (136)$$

ahol I_n egy $n \times n$ méretű egységmátrix.

$$A := \begin{bmatrix} 1 & x_{1,2} & \cdots & x_{1,k} & -1 & -x_{1,2} & \cdots & -x_{1,k} & 1 & 0 & 0 & -1 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 1 & 0 & 0 & -1 & 0 \\ 1 & x_{n,2} & \cdots & x_{n,k} & -1 & -x_{n,2} & \cdots & -x_{n,k} & 0 & 0 & 1 & 0 & 0 & -1 \end{bmatrix} \quad (137)$$

Az A mátrix sorai a feltételek baloldali részét képezik. Az A mátrix n sorból áll, azaz a modellillesztéshez használt minden egyes megfigyelés egy külön feltételként jelenik meg a lineáris programozási modellben. Az A mátrix oszlopai a lineáris modell független változói. Az X és $-X$ mátrixok, azaz az A mátrix első $2k$ oszlopa a modellillesztés szempontjából lényeges változókat tartalmazzák, ezek a változók együtthatói fogják

megadni a lineáris regresszió eredményét. Az I_n és $-I_n$ mátrixok, azaz az A mátrix utolsó $2n$ oszlopa a pozitív és a negatív hibának megfelelő változókat tartalmazzák.

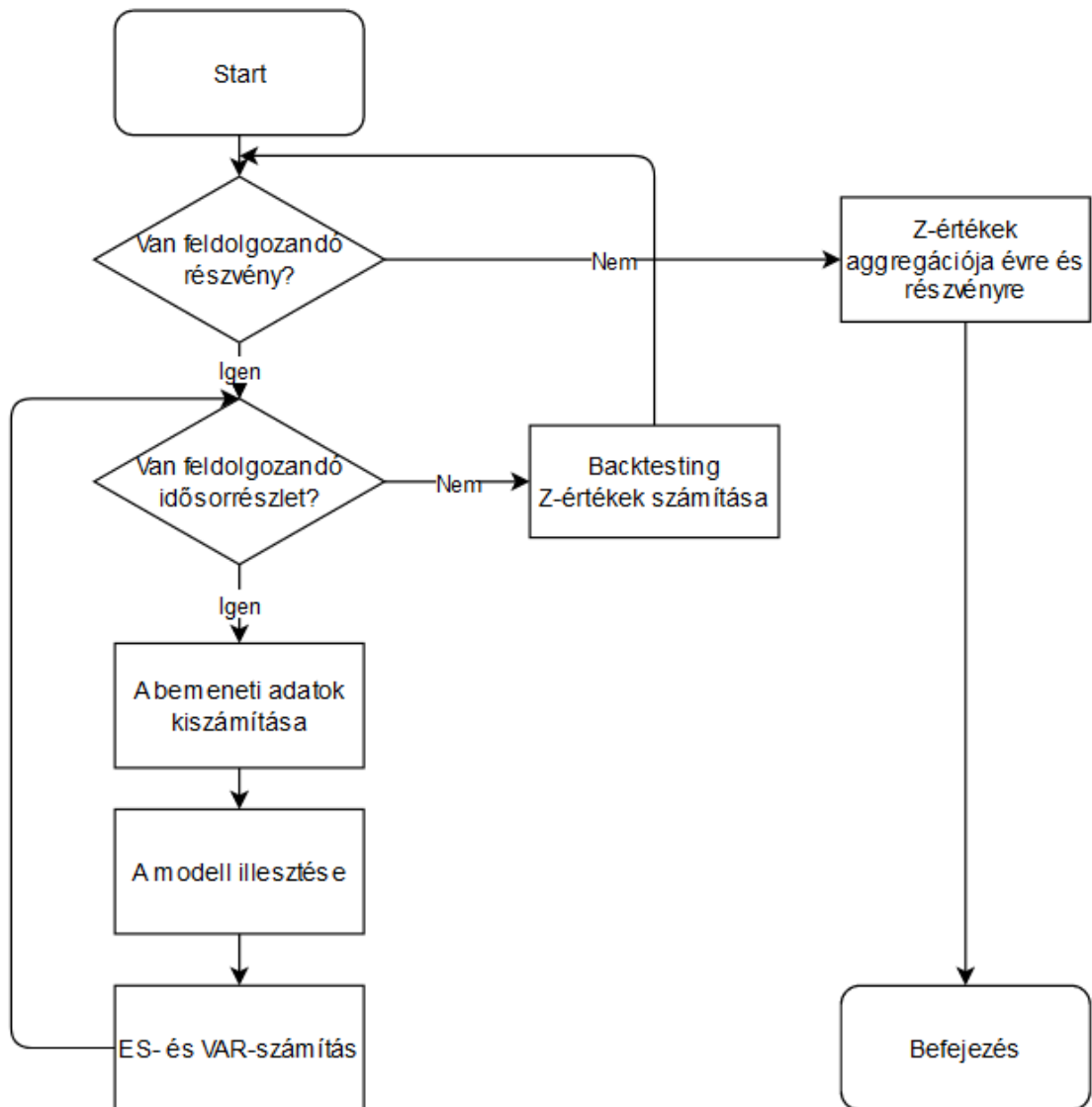
$$b := y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \quad (138)$$

A b mátrix a feltételek jobboldali részét képezik, azért kerülnek ide a megfigyelt függő változók értékei.

$$c := \begin{bmatrix} 0_{2k} \\ \theta 1_n \\ (1 - \theta) 1_n \end{bmatrix} = \begin{bmatrix} 0_1 \\ \vdots \\ 0_{2k} \\ \theta_1 \\ \vdots \\ \theta_n \\ (1 - \theta)_1 \\ \vdots \\ (1 - \theta)_n \end{bmatrix} \quad (139)$$

A c vektor az A mátrix minden oszlopához (a modell minden változójához) tartalmaz egy súlyt, amivel az a célfüggvény értékébe számít. A modell célja, hogy a releváns, első $2k$ változóhoz olyan együtthatókat találjon, amikor a hibák – azaz az utolsó $2n$ változó – minimálisak. Ez alapján a célfüggvényben az első $2k$ érték súlya 0 , míg az ezt követő, a pozitív hibákként értelmezett n érték súlya θ és az utolsó, a negatív hibáknak megfeleltethető n érték súlya pedig $(1 - \theta)$.

30. ábra: A QR-modell



Forrás: (saját szerkesztés)

4.7.5 A várható többletveszteség számítása kvantilisregresszió segítségével

Bayer és Dimitriadis (2019b) egy olyan regressziós keretrendszert mutat be, amely segítségével a VaR és az ES mutatók egyszerre modellezhetőek. A regressziós eljárás Koenker (2005) munkáján alapul, és a két mutató párosának szigorúan konzisztens veszteségfüggvényeit használja. A veszteségfüggvények használata lehetővé teszi az együttes regressziós paraméterek M-becslését. A VaR és az ES mutató együttes modellezése elengedhetetlen, mivel az ES önálló becslése nem megvalósítható (Bayer &

Dimitriadis, 2019b). Ez abból adódik, hogy nem létezik olyan megfelelő veszteségfüggvény, amit az ES minimalizál és ami az M-becslés célfüggvényeként szolgálhatna (Gneiting, 2011). Fissler és Ziegel egy olyan veszteségfüggvényt mutat be, ami a két mutató párosára épül. Bayer és Dimitriadis (2019b) erre az eredményre alapozva mutatja be az ES és a VaR együttes becslésére alkalmas regressziós modelljét. A szerzőpáros az eljárást közzétette R-könyvtárként.

4.8 A felhasznált R-könyvtárak

A fenti modellek implementációja az R programozási nyelv használatával történt. A több, mint 20 éves múltra visszatekintő R programozási nyelv a matematikai statisztika területén rendkívül népszerű és több, mint 7500 standard könyvtárat tartalmaz. Használata valamivel bonyolultabb, mint az olyan menüalapú programoké, mint az SPSS. Ennek köszönhetően viszont az R lényegesen rugalmasabb, mint egy standard statisztikai program és számos komplex művelet elvégezhető néhány függvényhívással (Pröhle & Zempléni, 2016). A lényegesebb felhasznált R-komponenseket és azok felhasználási területét a következő táblázat tartalmazza.

12. táblázat: A modellimplementációk szempontjából lényeges R könyvtárak

R könyvtár	Hivatkozás	Felhasználási terület
rugarch	(Ghalanos, 2020)	Garch-alapú modellek
cvar	(Boshnakov, 2019)	t4-modell
VaRES	(Nadarajah, et al., 2015)	Eloszlásalapú modellek
esreg	(Bayer & Dimitriadis, 2019b)	QR-modell
stats4	(R-core, 2019)	Maximum likelihood paraméterbecslés

Forrás: (saját szerkesztés)

5 A neurális hálózaton alapuló modellek

5.1 A neurális hálózatokkal kapcsolatos szakirodalom

5.1.1 Történeti áttekintés – a neurális háló fejlődésének fő mérföldkövei

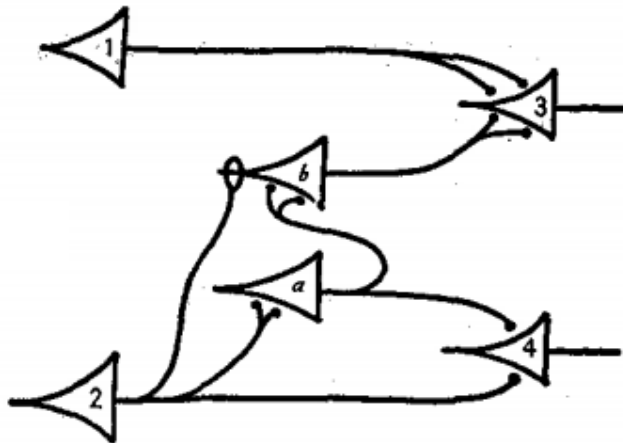
5.1.1.1 Lineáris regresszió

A neurális háló első modelljei lényegében a lineáris regresszió különböző változatai voltak. A lineáris regresszió egy olyan neurális hálónak tekinthető, amelynek mélysége mindössze egy szint (Schmidhuber, 2015).

5.1.1.2 A McCulloch–Pitts-neuron (MCP-neuron)

A neurális háló történetének első lépése McCulloch és Pitts (1943) publikációja, amelyben a szerzőpáros a később McCulloch–Pitts-modellként hivatkozott neurális hálót és a hozzá kapcsolódó matematikai definíciókat írja le. A hálózat rugalmasságát a neuronokat összekötő élek változó súlyai adják, de a rendszer fontos korlátja volt, hogy nem volt képes a tanulásra. A MCP-neuronok minden szinapszisukon beérkező impulzust figyelembe vesznek és hatásukat összegzik. A szerzők a modellt egy egyszerű elektromos áramkörként írták le.

31. ábra: McCulloch és Pitts egyik neurálisháló-modellje



Forrás: (McCulloch & Pitts, 1943, p. 130)

A McCulloch–Pitts-neuron összegzi a bemenetek súlyozott értékeit, és ha az így kapott érték meghaladja az előre megadott határértéket, akkor a kimenetén 1 értéket ad, különben pedig nullát. Az n bemenetű McCulloch–Pitts-neuron kimenetét tehát a következő összefüggés adja meg:

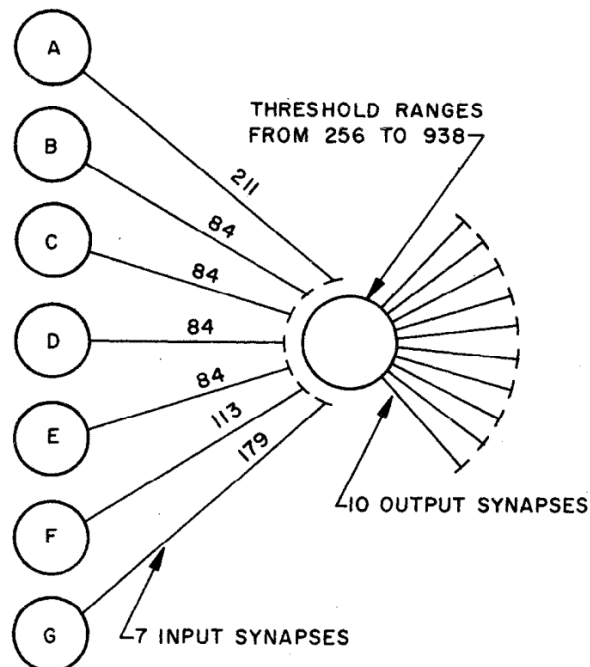
$$y = \begin{cases} 1, & \text{ha } \sum_{i=1}^n x_i > b \\ 0, & \text{ha } \sum_{i=1}^n x_i \leq b \end{cases} \quad (140)$$

A McCulloch–Pitts-neuron egyik legnagyobb problémája, hogy az inputok és outputok binárisak, azaz csak 0 vagy 1 értéket vehetnek fel. Szintén a modell egyszerűségéből adódott, hogy az XOR, azaz a „kizáró vagy” logikai kapcsolatot MCP-neuronok segítségével nem sikerült szimulálni (Dr. Aradi, et al., 2014).

5.1.1.3 A Hebb-tanulás

A fejlődés következő lépéseként Hebb (1949) bevezette azt a rendkívül fontos koncepciót, hogy a neuronok közötti kapcsolat erősödik minden alkalommal, ha az egyik neuron a másikkal egyszerre aktiválódik. Ez a mechanizmus ad lehetőséget a tanulás és a memória működésére. A jelenségre „Hebb-tanulásként” vagy „Hebb-szabályként” hivatkozik az irodalom (Altrichter, et al., 2006). A Hebb-tanulás a felügyelet nélküli tanulás kategóriájába tartozik, és előre vetíti az MCP-neuron továbbfejlesztésének lehetőségét.

32. ábra: A Rochester-féle szimulált neuron és kapcsolatai



Forrás: (Rochester, et al., 1956, p. 91)

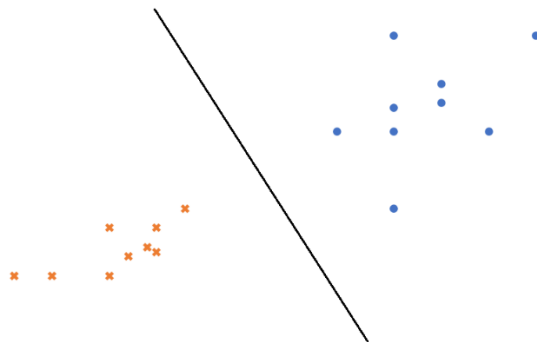
5.1.1.4 Az első digitálisan megvalósított neurális hálózatok

Farley és Clark (1954) használt először digitális gépeket egy Hebb-hálózat szimulálására. A létrehozott információs rendszer igen kezdetleges volt, de képes volt önállóan eljutni egy kezdeti véletlenszerű állapotból a céljának megfelelő szervezett állapotba. 1955-ben McCarthy, Minsky, Rochester és Shannon felhívást adott közre, amelyben kéthónapos kutatási projektet javasoltak a mesterséges intelligencia témájában a hanoveri Dartmouth College-ban. Megközelítésük szerint egy gépnek képesnek kell lennie az intelligencia szimulálására, amennyiben annak minden lényeges aspektusát kellő pontossággal leírjuk (McCarthy, et al., 1955). Ezt követően Rochester és az IBM kutatócsoportja (Rochester, et al., 1956) egy IBM 704 típusú számítógépen valósították meg a Hebb-hálózat digitális változatát.

5.1.1.5 A Rosenblatt-perceptron

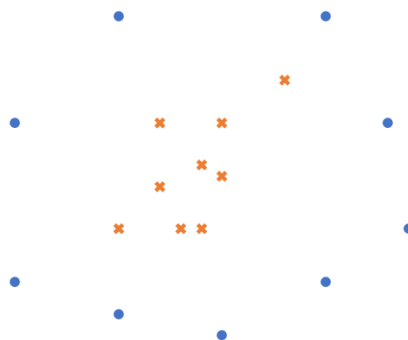
Egy mai napig használt modell a Rosenblatt-perceptron (Rosenblatt, 1958). Az egyszerű perceptronnak is hívott modell egy olyan neurális háló, amely képes arra, hogy egy lineárisan szeparálható bemeneti halmazzt két csoportba kategorizáljon.

33. ábra: Lineárisan szeparálható halmazok



Forrás: (saját szerkesztés)

34. ábra: Nemlineárisan szeparálható halmazok



Forrás: (saját szerkesztés)

Rosenblatt kísérletében egy gép képes volt a kezdeti tanítási szakaszt követően megkülönböztetni a jobb oldalon megjelölt kártyákat a bal oldalon megjelölt kártyáktól. A Rosenblatt-perceptron az MCP-neuron több megkötését is feloldotta, például a bemenetek már nem csak bináris értékeket vehettek fel, és különböző súlyokkal rendelkezhetnek. Az n bemenetű Rosenblatt-perceptron kimenetét a következő összefüggés adja meg:

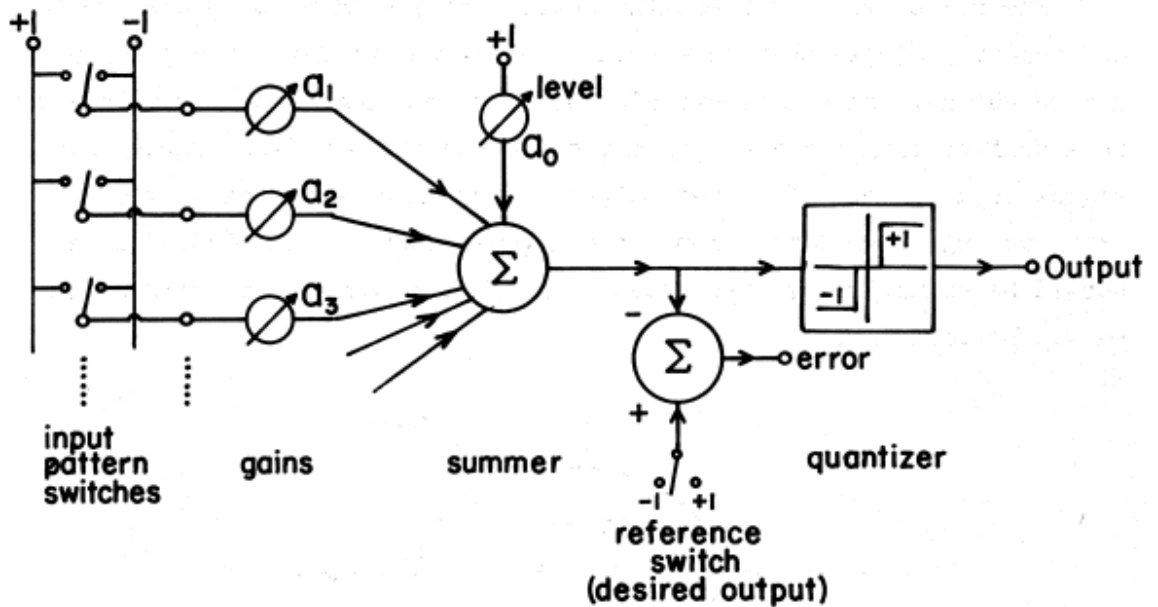
$$y = \begin{cases} 1, & \text{ha } \sum_{i=1}^n w_i x_i > b \\ 0, & \text{ha } \sum_{i=1}^n w_i x_i \leq b \end{cases} \quad (141)$$

Annak ellenére, hogy a Rosenblatt-perceptron lényegesen rugalmasabban használható, mint az MCP-neuron még ez a modell sem képes az XOR logikai művelet szimulálására.

5.1.1.6 Adaline (Adaptive Linear Neuron / Adaptive Linear Element)

Az Adaline-modellt Widrow és Hoff (Widrow & Hoff, 1960) javasolta egy olyan neurális hálóként, amely hatékonyabban adaptálódik és tanul a korábbiaknál. Az Adaline-modell nagyban hasonlít a perceptronhoz, a két modell különbsége a hiba számításában és a súlyok korrigálásában rejlik. A perceptron a bináris output eredményéből számítja a hibát és ezzel korrigálja a súlyokat. Az Adaline-modell az aktivációs függvény (ld. *quantizer* a 35. ábrán) alkalmazása előtt kapott folytonos eredményt használja a súlyok korrigálására.

35. ábra: Az Adaline-modell felépítése



Forrás: (Widrow & Hoff, 1960, p. 6)

5.1.1.7 Az egyszintűperceptron-modell korlátainak felismerése

A mesterséges intelligencia témakörét övező kezdeti lelkesedést nagyban visszafogta Minsky és Papert *Perceptrons* című könyve (1969), amely matematikai elemzések segítségével levezeti, hogy az akkor ismert modellek milyen elméleti korlátokkal rendelkeznek. Az egyik legjelentősebb probléma az volt, hogy az akkoriban használt egyrétegű hálózatok nem tudták megvalósítani az XOR logikai függvényt. A szerzőpáros által bemutatott másik legnagyobb probléma, a modellek számítási igénye volt. Ahhoz, hogy a modelleket valós problémák megoldására lehessen alkalmazni nagyságrendekkel nagyobb számítási kapacitás lett volna szükséges, mint ami a könyv megjelenésekor rendelkezésre állt. Ez többek között azért okozott komoly problémát, mert a technológiával szemben komoly gyakorlati elvárások is voltak, többek között már a 50-es években automatikus fordítógépek megvalósítását várták el a neurális hálóktól, ami a fenti korlátozások miatt csak évtizedekkel később tudott valamennyire megvalósulni (Hutchins, 1995). A szerzőpáros könyvének megjelenését követően a mesterséges intelligencia kutatási területének finanszírozása drasztikusan visszaesett az addig élenjáró USA területén, ezzel a következő időszakban a fejlődés is lassult.

5.1.1.8 *A hatékony hibavisszaterjesztés (backpropagation, backprop, BP)*

Az irodalomban számos egymásnak ellentmondó információ található arról, hogy a jelenlegi fejlődési hullámot elindító hatékony hibavisszaterjesztési algoritmust, másnéven az automatikus differenciálás visszafejtő módszerét ki írta le először. Számos forrás Werbos (1974) Ph.D. disszertációját említi a neurális hálókkal kapcsolatos áttörő eredményként, annak ellenére, hogy a munka egyáltalán nem említi a neurális hálókat, és a szerző a bemutatott algoritmust nem alkalmazza a neurális hálók kontextusában. Szintén fontos kiemelni, hogy a hibavisszaterjesztés (backpropagation) fogalma már korábban is ismert volt. Többek között Rosenblatt is leírt ilyen néven egy eljárást, de az nem volt alkalmas komplex, többszintű neurális hálók tanítására. A jelenleg használatos backpropagation algoritmust először egy finn egyetemi hallgató, Linnainmaa (1970) írta le a finn nyelvű szakdolgozatához csatolt FORTRAN kódban (Schmidhuber, 2015). Linnainmaa Werboshoz hasonlóan eredeti munkájában nem említette a neurális hálókat.

Werbos (1974) Ph.D. disszertációjában politikai előjelzéseket készített statisztikai modellekkel. A modellezés során problémásnak találta a paraméterek illesztését a valós adatokhoz, illetve az eredményként létrejött modellek becslései pontatlannak bizonyultak. A probléma feloldásaként Linnainmaa munkájától függetlenül új algoritmusokat dolgozott ki, amelyek segítségével a modellek illesztése hatékonyabbá vált. Annak ellenére, hogy Werbos a disszertációban a maximum likelihood módszernél hatékonyabbnak bizonyuló modellillesztést ARMA és regressziós modelleken mutatta be, a módszertan később a neurális hálók fejlődésében fontos mérföldkövé vált.

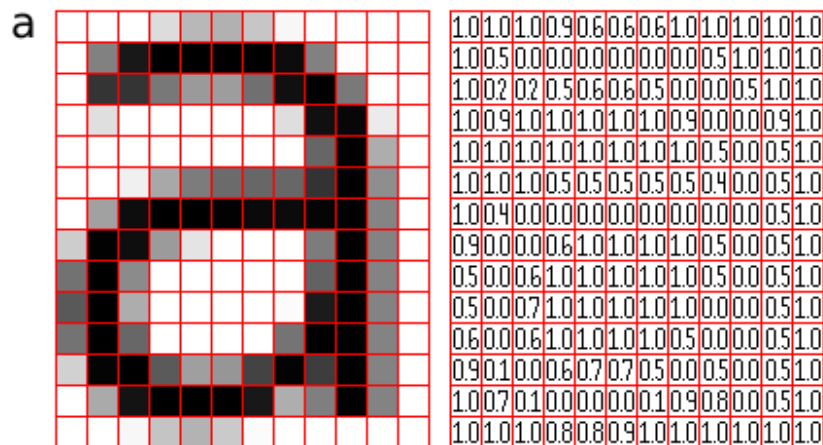
A hatékony hibavisszaterjesztési algoritmus új lendületet adott a mesterséges intelligencia kutatásának. A backprop algoritmus lényege, hogy kiszámítja, hogy a hálózat egyes súlyainak változtatása milyen mértékben csökkenti a költségfüggvényt, így kiszámítható a költségfüggvényt leginkább csökkentő súlykorrekció. A backprop algoritmus egyik nagy előnye, hogy többszintű hálózatok tanítására is alkalmazható, ezzel megnyitva az utat a mély tanuláson alapuló modellek alkalmazása előtt.

5.1.1.9 *Konvolúciós neurális hálózatok (Convolutional neural networks, CNN, convnet)*

Wiesel és Hubel (1959) megfigyelték, hogy a macskák látókérgét egyszerű és komplex sejtek építik fel. A komplex sejtek nagyobb térbeli invarianciát mutattak, mint az egyszerű sejtek. Kutatási eredményeik később különböző többszintű neurálisháló-architektúra inspirációjaként szolgáltak. Az első ilyen modell Fukushima

(1979) nevéhez köthető, aki az általa megalkotott neurális hálót neocognitronnak nevezte. A modell egyike volt az első valóban többszintű neurális hálóknak és elsősorban képek feldolgozására alkalmas. A feldolgozás során a feldolgozandó kép pixeleinek tulajdonságai külön-külön bemenetet képeznek. Az első rétegekben a konvolúciós neurális hálózat csak egyszerűbb formákat ismer fel, mint például egyeneseket és görbéket, a későbbi szinteken pedig komplexebb objektumokat, mint például, hogy milyen betű látható egy képen.

36. ábra: Raszterizált betű



Forrás: (Kolås, 2005)

5.1.1.10 A Hopfield-hálózat

A Hopfield-hálózat egy visszacsatolt neurális háló (recurrent neural network, RNN) (Hopfield, 1982). A hálózat asszociatív, tartalom szerint címezhető memóriaként működik és az emberi emlékezet működésére is modellként szolgál. Hopfield a neurális háló gyakorlati alkalmazását szorgalmazta, munkája eredményeként a neurális háló területe világszerte újra felkapott kutatási területté vált (Macukow, 2016).

5.1.2 A gépi tanulás kategorizálása a tanulás módja szerint

5.1.2.1 Kezdeti, tanulásra képtelen modellek

A neurális háló legkezdetlegesebb változatai még nem voltak képesek a tanulásra. Az egyetlen ismert modell ebben a kategóriában az MCP-hálózat (McCulloch & Pitts, 1943).

5.1.2.2 *Felügyelt tanulás (supervised learning – SL)*

Jelenleg a gyakorlati felhasználás során a legnagyobb figyelmet a felügyelt tanuláson alapuló megoldások kapják. Az SL-módszer lényege, hogy a neurális háló felcímkezett adatokon tanul, azaz egy olyan adathalmazon, amely a bemenethez tartozó elvárt kimeneteket is tartalmazza. A neurális háló a tanítási fázisban megkeresi azokat a súlyokat, amikkel a tanítási adathalmazon a legkisebb hibát éri el. A fenti modellek közül ebbe a csoportba tartozik a Rosenblatt-féle perceptron (1958) és az Adaline-modell (Widrow & Hoff, 1960).

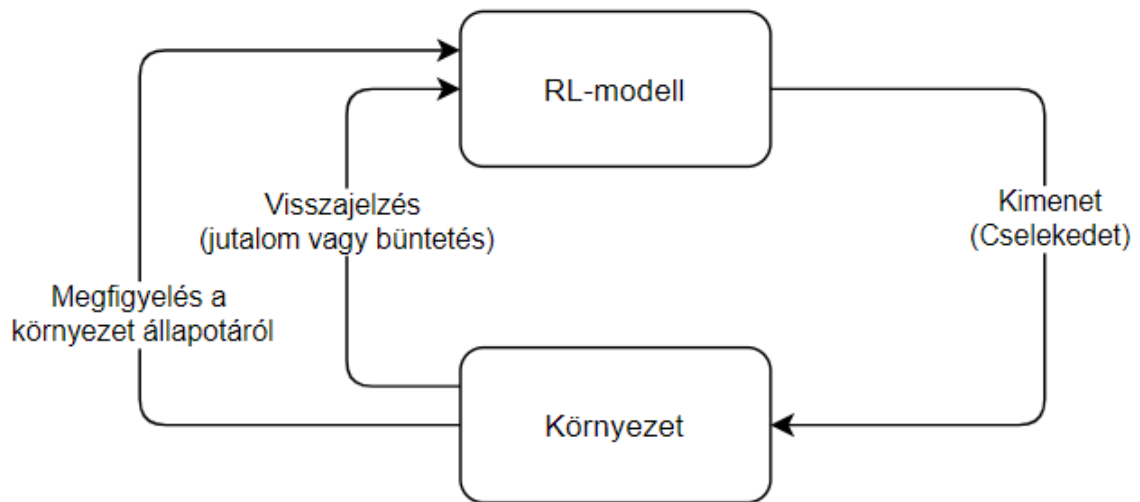
5.1.2.3 *Felügyelet nélküli tanulás (unsupervised learning – UL)*

A felügyelet nélküli tanulás során a modellek új mintákat keresnek az adatokban. Az eljárás során az adatok nincsenek felcímkezve. A leggyakoribb feladattípus a klaszterezés, azaz az adatok csoportokra bontása. A felügyelet nélküli tanulás egyik érdekessége, hogy mind a megerősítéses, mind a felügyelt tanulást elősegítheti oly módon, hogy a bejövő nyers adatokat kompaktabb, megfelelőbb formára kódolja például a videó vagy szövegfelismerés esetén (Schmidhuber, 2015). A korábbi modellek közül ebbe a csoportba tartozik a Hebb-féle neurális háló (1949).

5.1.2.4 *Megerősítéses tanulás (reinforcement learning – RL)*

A megerősítéses tanulás lényege, hogy a modell olyan kimenetek kiszámítására törekszik, amelyek egy bizonyos kumulatív célfüggvény értékét maximalizálják. Az RL-modell által megoldott problémák tipikusan felírhatóak egy dinamikus programozási feladatként is, azaz egy célfüggvény és különböző megkötések segítségével. Az RL és a dinamikus programozás lényeges különbsége, hogy az RL modell működéséhez nem szükséges a célfüggvény pontos ismerete. Az RL modell esetében az is elégséges, ha a környezetből érkező pozitív vagy negatív visszajelzésekből tanulhat. A megerősítéses tanulási feladatoknál az is problémát jelent, hogy az egyes kimenetek és az azokra érkező visszajelzések között ismeretlen hosszúságú eltolódás lehet (Schmidhuber, 2015).

37. ábra: A megerősítéses tanulás modellje



Forrás: (saját szerkesztés)

5.1.3 Kategóriák a terjesztés irányja alapján

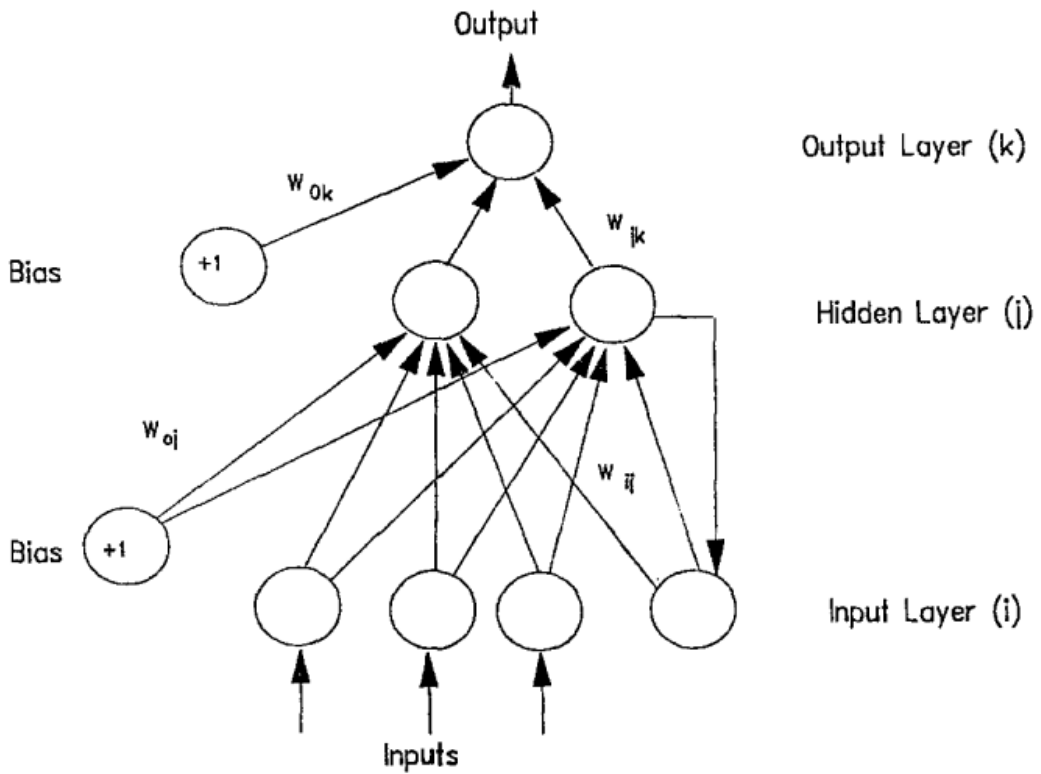
5.1.3.1 Előreterjesztéses neurális hálózatok (Feedforward neural network ,FNN)

Az egyszerűbb FNN, azaz előreterjesztéses neurális hálózatok legfontosabb jellemzője, hogy az adat továbbítás iránya mindig a bemenetek felől a kimenetek felé történik és nincs semmilyen visszacsatolás. Az FNN hálózatok alapvetően akkor alkalmazhatóak, ha a feldolgozandó adatokban nincs autokorreláció, vagy más olyan összefüggés, amely alapján egy adott kimenetet a korábbi becslések értékei befolyásolnának.

5.1.3.2 Visszacsatolásos neurális hálózatok (Recurrent neural network ,RNN)

A visszacsatolásos neurális hálózatok az FNN hálókkal szemben képesek a visszacsatolások kezelésére, ezért alkalmasak olyan összefüggések kezelésére is, amikor az adatok időfüggőek vagy autokorrelációt mutatnak (Dematos, et al., 1996).

38. ábra: ViSSzacSatoLásos neurális hálózat



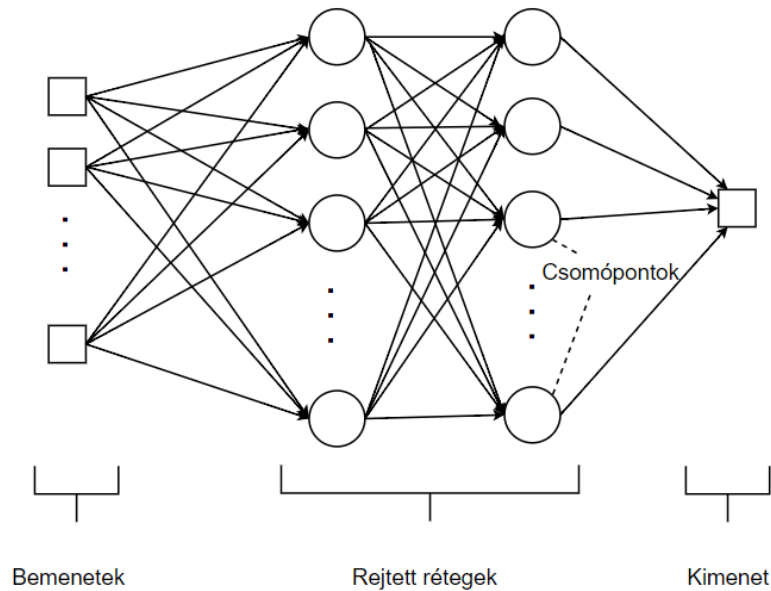
Forrás: (Dematos, et al., 1996, p. 62)

5.2 A neurális hálózat működése

A neuron elektromos jelek összegyűjtését, feldolgozását és továbbítását végző agysejt. A mesterséges neurális háló (ANN) alapkoncepciója, hogy az agy információfeldolgozó kapacitását adó neuronhálózatot mesterséges, matematikai modellen alapuló neuronok segítségével leképezze. A neurális háló csomópontok halmaza, amelyeket adaptív súlyú kapcsolatok kötnek össze. Minden csomópont összegzi a bemeneteire érkező jeleket, majd az aktivációs függvény segítségével transzformálja az így kapott értéket és továbbítja azt a csomópont kimeneteként. Ez a kimenet a megfelelő súlyokkal módosítva más csomópontok bemeneteként szolgál (Russell & Norvig, 2005).

A neurális hálók a fentieknek megfelelően csomópontokból és az azokat összekötő irányított kapcsolatokból épülnek fel (Russell & Norvig, 2005, p. 845). A csomópontok bemenetét a bejövő kapcsolatok értékei és az azokhoz kapcsolódó súlyok határozzák meg.

39. ábra: A többszintű perceptron



Forrás: (saját szerkesztés, Russel és Norvig (2005) alapján)

$$in_i = \sum_{j=0}^n W_{j,i} a_j \quad (142)$$

A fenti képletben in_i az i . csomópont bemeneteiből kapott érték, $W_{j,i}$ a j . csomópontból az i . csomópontba érkező irányított kapcsolat súlya, a_j a j . csomópont kimenete, n pedig az i . csomópont bemenetére kötött csomópontok száma.

A tanítási fázisban a rendszer felcímkézett adatokat dolgoz fel, azaz olyan adathalmazt kap, amely a bemenetek mellett az elvárt kimeneteket is tartalmazza. Kezdetben a csomópontokat összekötő háló $W_{j,i}$ súlyai véletlenszerű értéket kapnak, majd a tanulási adatok alapján a súlyok lépésenként korrigálódnak, oly módon, hogy amikor a neurális háló rossz becslést ad, akkor a hibás becslésért felelős csomópontok súlyai csökkennek, míg a helyes értékhez hozzájáruló csomópontok súlyait növeli a visszacsatoló algoritmus. A tanulási fázis egy ismétlődő, iteratív folyamat, mely során a modell a visszacsatolási algoritmus segítségével a fent leírt módon lépésenként a csomópontok közti kapcsolatok súlyát próbálja úgy módosítani, hogy a modell becslési hibája csökkenjen (Russell & Norvig, 2005).

A visszacsatolási algoritmus hatására a feldolgozott adatmennyiség növekedésével a modell többnyire képes egyre jobb eredményeket elérni az előre meghatározott feladat elvégzésében. A felhasznált szintek számának növelésével az ANN

tetszőleges függvényt le tud képezni, ezért számos felhasználási terület mellett a várható többletveszteségmutató-számításra is alkalmas (Hornik, 1989).

5.3 A neurális hálózat tanítása

Egy FNN-hálózatban minden csomópont össze van kötve az előző réteg összes csomópontjával és egy konstans taggal. Ebből következik, hogy a hálózatban a súlyok száma (n_w), amennyiben a rétegekben csomópontok száma azonos (Russell & Norvig, 2005):

$$n_w = n_i n_h + (d - 1) n_h^2 + n_h n_o + n_b \quad (143)$$

, ahol n_i az inputértékek darabszáma, n_h a rejtett rétegek csomópontjainak száma, n_o a kimeneti csomópontok száma, d a rejtett rétegek száma, n_b pedig a konstans tényezőkhöz kapcsolódó súlyok száma.

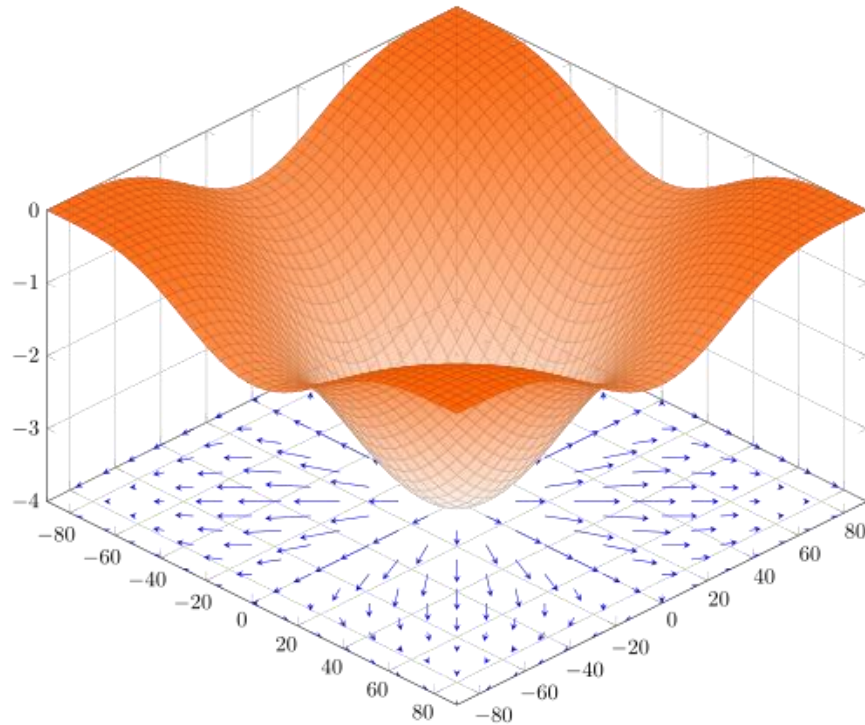
$$n_b = d n_h + n_o \quad (144)$$

Konstans tényezők a rejtett rétegek és a kimeneti réteg csomópontjaihoz kapcsolódnak, a bemeneti csomópontokhoz nem. A hálózat tanításának első lépése az, hogy a csomópontokat összekötő súlyok véletlenszerű értéket kapnak. Ezt követően a rendszer feldolgozza a tanításra használt, felcímkezett adathalmaz első elemét, de a véletlenszerű súlyok miatt még rossz választ ad (Russell & Norvig, 2005). A modell megfelelő működése esetén a tanulási folyamat eredményeként a hálózat egyre jobb válaszokat ad. A tanítási folyamat nem ér véget a tanításra használt adathalmaz első feldolgozása után, mivel az adathalmaz többször is feldolgozásra kerülhet. Ez ismétlések számát az epochszám adja meg. Epochszám alatt tehát a neurális háló tanítási fázisának ismétlésszámát értjük, azaz azt, hogy hányszor fogja a háló feldolgozni a felcímkezett tanítási idősort.

5.4 Legmeredekebb lejtő algoritmus (gradient descent)

Mivel felcímkezett adatokat használunk, a bemenetekhez tartozó elvárt válaszok is rendelkezésre állnak, így egy költségfüggvény segítségével kiszámítjuk az elvárt és a becsült eredmények különbségét. Költségfüggvényként többek között az átlagos abszolút hibát vagy az átlagos négyzetes hibát használhatjuk.

40. ábra: Függvény⁸ és gradiense vektormezőként ábrázolva az alsó síkon



Forrás: (Thoma, 2008)

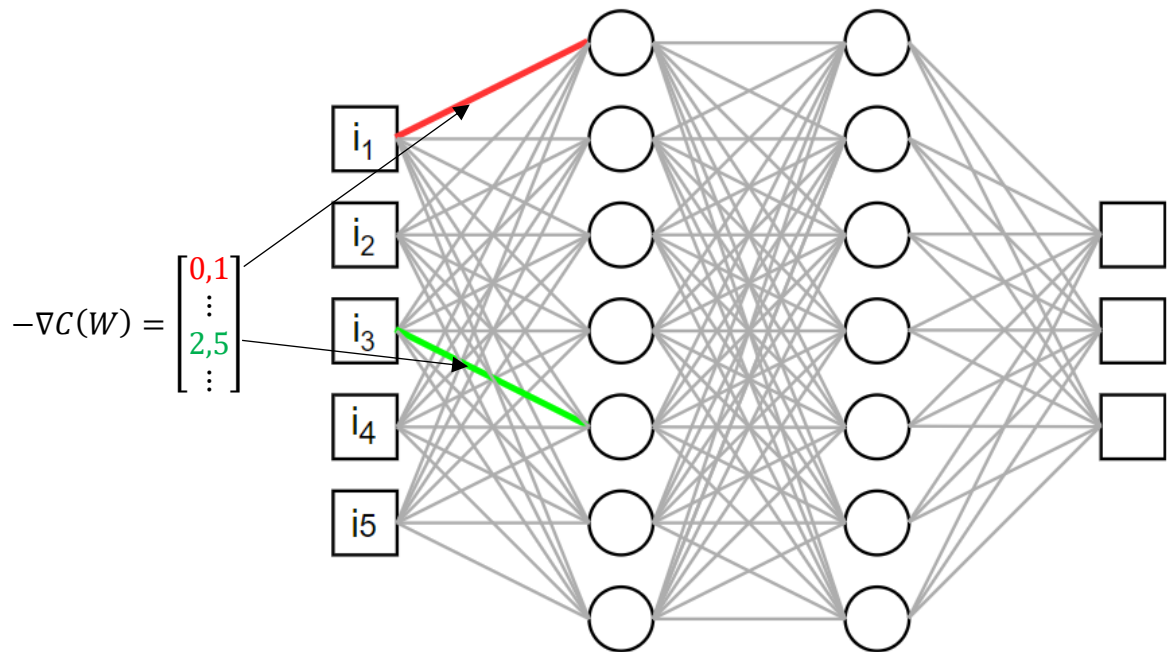
A költségfüggvény gradiense $\nabla C(W)$ az aktuális w -súlyvektorhoz tartozó pontban megadja, hogy melyik irányban növekszik legjobban a költségfüggvény értéke, illetve az eredményként kapott vektor hossza a növekedés mértékével azonos. Mivel a célunk a költségfüggvény csökkentése, így a súlyvektort a kapott vektor negatívjával kell korrigálnunk:

$$W_i = W_{i-1} - \nabla C(W_{i-1}) \quad (145)$$

Könnyen belátható, hogy az eljárás alkalmas a költségfüggvény lokális minimumainak megtalálására (Russell & Norvig, 2005) (Sanderson, 2018). A $\nabla C(W)$ értékei azt is megmutatják, hogy a neurális háló mely súlyai bírnak nagy jelentőséggel az adott tanítási adat fényében, hiszen minél nagyobb súlymódosítást ír elő a $\nabla C(W)$ adott súlyhoz tartozó értéke, annál nagyobb lesz az adott kapcsolat jelentősége a későbbi eredmények szempontjából. Egy egyszerű példával élve, ha $-\nabla C(W)$ vektor egyik értéke 0,1 míg egy másik értéke 2,5, akkor az utóbbi súlyhoz kapcsolódó input (i_3) 25-ször

⁸ $f(x, y) = -(\cos^2(x) + \cos^2(y))^2$

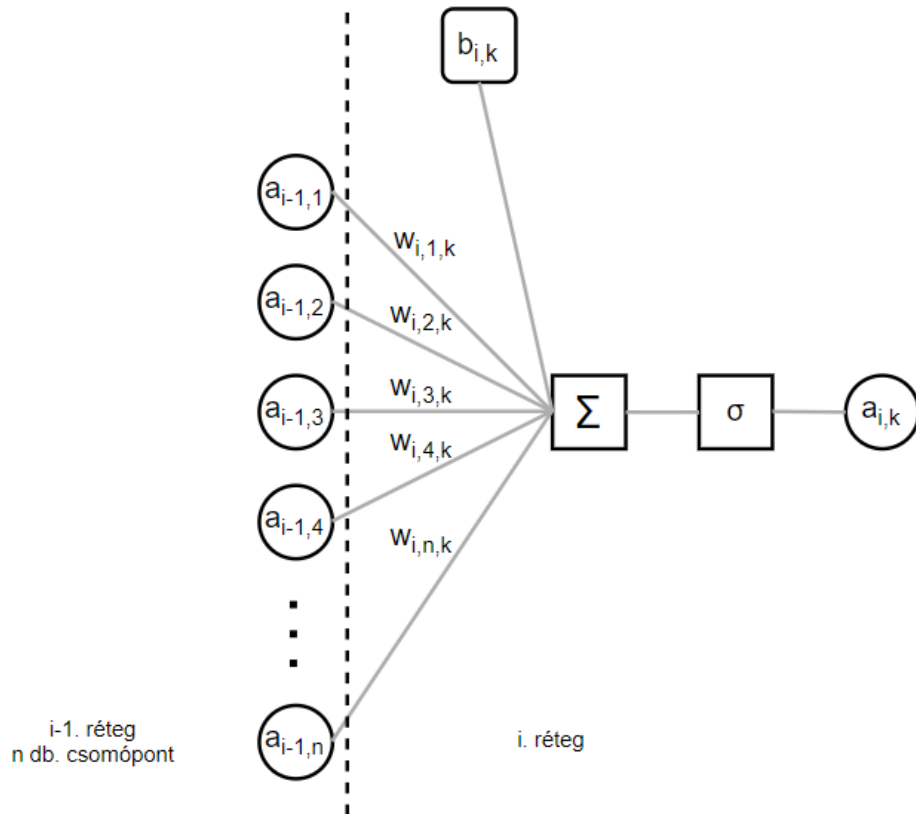
akkora hatással lesz a végeredményre, mint az első (i_1), azt feltételezve, hogy az inputokat azonos tartományra transzformáltuk.



5.5 Backprop-algoritmus

A tanítási fázis lényege tehát a $C(W)$ -költségfüggvény értékének csökkentése, amit úgy érhetünk el, hogy egy adott kimeneti csomópont értékét közelítjük az elvárt értékhez.

41. ábra: A csomópontok kimenetének számítása



Forrás: (saját szerkesztés)

Az i . (rejtett vagy kimeneti) réteg k . csomópontjának értékét a következő képlet adja meg:

$$a_{i,k} = \sigma_i(z_{i,k}) \quad (146)$$

$$z_{i,k} = w_{i,1,k}a_{i-1,1} + w_{i,2,k}a_{i-1,2} + \dots + w_{i,n,k}a_{i-1,n} + b_{i,k} \quad (147)$$

A fentiekben:

n : az $(i-1)$. réteg csomópontjainak száma

$a_{i,k}$: az i . réteg k . csomópontjának, vagy a kimenet k . csomópontjának aktivációs értéke

σ_i : az i . réteg aktivációs függvénye

$w_{i,j,k}$: $a_{i-1,j}$ és $a_{i,k}$ közötti kapcsolat súlya

$b_{i,k}$: az i . réteg k . csomópontjához tartozó konstans tag

A fentiek alapján könnyen belátható, hogy amennyiben $a_{i,k}$ értéke nem megfelelő, például kisebb az elvárttól, úgy a következő három módszer egyikével lehet ezt korrigálni:

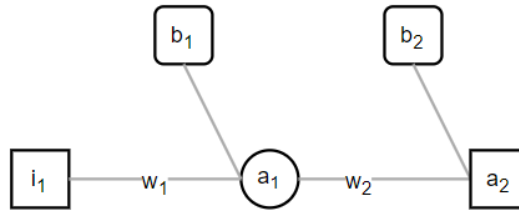
- a b konstans növelésével
- a $w_{i,j,k}$ súlyok növelésével
- az előző réteg, azaz az $a_{i,j}$ csomópontok értékeinek növelésével

A backprop-algoritmus a kimeneti csomópontoktól indul és meghatározza, hogy ezek mindegyike mennyivel tér el az elvárttól. Ezt követően az algoritmus a hálózatban visszafelé haladva a gradiens segítségével meghatározza, hogy az adott csomópont korrekciós értékét milyen mértékben kell módosítani. Mivel több kimeneti csomópont is lehet, ezért a szükséges paramétermódosítások mindegyikre kiszámításra kerülnek, és ezek hatása összeadódik. A kimeneti réteg feldolgozásának befejeztével az ezt megelőző réteg csomópontjaira is megkapjuk, hogy mi lett volna az elvárt értékük, így a fenti eljárás újra megismételhető az egyel korábbi rétegben, addig amíg a kimeneten megfigyelt hiba a teljes hálózatban visszaterjesztésre kerül (Russell & Norvig, 2005).

5.6 A backprop-algoritmus bemutatása egy példán keresztül

A példában egy egyszerű neurális háló kerül bemutatásra. A neurális háló egy bemenettel (i_1) és egy rejtett réteggel rendelkezik, és egyetlen csomópontja (a_1), valamint egyetlen kimenete (a_2) van. A konstans tagoknál a b_1 és b_2 jelöléseket használom.

42. ábra: Példaként bemutatott egyszerű neurális háló



Forrás: (saját szerkesztés)

A cél a ∇C –gradiens kiszámítása, aminek a fenti példa esetén a következő négy eleme lesz:

$$\nabla C_{\theta} = \begin{bmatrix} \frac{dC}{dw_1} \\ \frac{dC}{db_1} \\ \frac{dC}{dw_2} \\ \frac{dC}{db_2} \end{bmatrix} \quad (148)$$

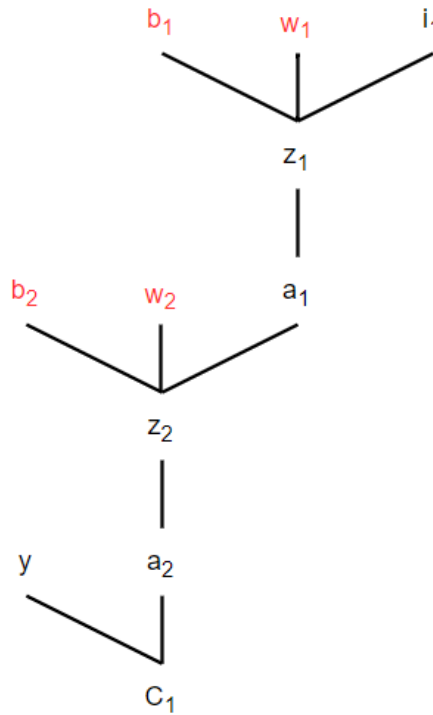
A fentiekben θ a modellparaméterek vektora. A vektor négy eleme megfelel annak a négy paraméternek, amit a neurális háló tanítása során közvetlenül módosítani lehet. Ez a két súlyt és a két konstans tagot jelenti. A 43. ábra bemutatja, hogy az egyes paraméterek hogyan befolyásolják egymást közvetlenül. Például az a_1 érték csak a z_1 értéktől függ közvetlenül, de a z_1 a b_1 , w_1 és i_1 paraméterektől is. Az indirekt függőségek fastruktúrát követe épülnek egymásra, indirekt módon a paraméterek az összes felettük található paramétertől függenek.

Fontos kiemelni, hogy a backprop-algoritmus alapvetően visszafelé halad, azaz a 43. ábra estében C_1 -től kezdve felfelé haladunk a fastruktúrán. Eszerint a logika alapján az utolsó réteghez tartozó b_2 és w_2 paraméterek kerülnek elsőként kiszámításra.

$$C_1 = |a_2 - y| \quad (149)$$

,ahol y az elvárt kimenet, C_1 pedig az 1. tanítási adathoz tartozó költség.

43. ábra: A paraméterek összefüggési hierarchiája



Forrás: saját szerkesztés, (Sanderson, 2018) alapján

Az első kiszámított paraméter azt fogja megadni, hogy b_2 módosítása milyen hatással lesz a költségfüggvényre. A $\frac{dC_1}{db_2}$ elem kiszámítása a láncszabály segítségével:

$$\frac{dC_1}{db_2} = \frac{dz_2}{db_2} \frac{da_2}{dz_2} \frac{dC_1}{da_2} \quad (150)$$

$$z_2 = w_2 a_1 + b_2 \quad (151)$$

$$\frac{dz_2}{db_2} = 1 \quad (152)$$

$$a_2 = \sigma(z_2) \quad (153)$$

$$\frac{da_2}{dz_2} = \sigma'(z_2) \quad (154)$$

$$\frac{dC}{da_2} = \frac{a_2 - y}{|a_2 - y|} \quad (155)$$

$$\frac{dC_1}{db_2} = \sigma'(z_2) \frac{a_2 - y}{|a_2 - y|} \quad (156)$$

Az $\frac{dC_1}{dw_2}$ elem kiszámítása az előbbihez hasonlóan megadja a költségfüggvény érzékenységét w_2 változásaira.

$$\frac{dC_1}{dw_2} = \frac{dz_2}{dw_2} \frac{da_2}{dz_2} \frac{dC_1}{da_2} \quad (157)$$

$$\frac{dz_2}{dw_2} = a_1 \quad (158)$$

$$\frac{dC_1}{dw_2} = a_1 \sigma'(z_2) \frac{a_2 - y}{|a_2 - y|} \quad (159)$$

A 43. ábrán jól látható, hogy az a_1 a harmadik paraméter, amely a C_1 költségfüggvényt ezen a szinten befolyásolja. Ez a paraméter eltér az előző kettőtől abban a tekintetben, hogy ennek értékét közvetlenül nem lehet módosítani, csak az őt befolyásoló paraméterek változtatásával. Ettől függetlenül szükséges kiszámítani, hogy milyen mértékben kellene az a_1 értékének módosulnia.

$$\frac{dC_1}{da_1} = \frac{dz_2}{da_1} \frac{da_2}{dz_2} \frac{dC_1}{da_2} \quad (160)$$

$$\frac{dz_2}{da_1} = w_2 \quad (161)$$

A többi tag nem változott, tehát az eredmény a következő:

$$\frac{dC_1}{da_1} = w_2 \sigma'(z_2) \frac{a_2 - y}{|a_2 - y|} \quad (162)$$

A fenti eredmény felhasználásával a neurális hálózat becslési hibáját visszaterjesztjük a korábbi szintre. Intuitíven a kezdeti $(a_2 - y)$ különbséghez (azaz a becslés és az elvárt érték különbségéhez) hasonlóan most is rendelkezésre áll egy elvárt és egy valós érték az a_1 esetében. Ezt a logikát követve tetszőleges mélyégben elvégezhető a hiba visszaterjesztése. Az előző szinthez hasonlóan itt is az a_1 értékét meghatározó három paraméter releváns azzal a különbséggel, hogy az i_1 input nem változtatható, és nem is függ más paraméterektől, így annak kiszámítása felesleges.

$$\frac{dC_1}{db_1} = \frac{dz_1}{db_1} \frac{da_1}{dz_1} \frac{dC_1}{da_1} \quad (163)$$

$$z_1 = w_1 i_1 + b_1 \quad (164)$$

$$\frac{dz_1}{db_1} = 1 \quad (165)$$

$$a_1 = \sigma(z_1) \quad (166)$$

$$\frac{da_1}{dz_1} = \sigma'(z_1) \quad (167)$$

$$\frac{dC_1}{db_1} = \sigma'(z_1)w_2\sigma'(z_2)\frac{a_2 - y}{|a_2 - y|} \quad (168)$$

Hasonlóan:

$$\frac{dC_1}{dw_1} = \frac{dz_1}{dw_1} \frac{da_1}{dz_1} \frac{dC_1}{da_1} \quad (169)$$

$$\frac{dC_1}{dw_1} = i_1 \sigma'(z_1)w_2 \sigma'(z_2) \frac{a_2 - y}{|a_2 - y|} \quad (170)$$

Az így megkapott eredmények megadják a szükséges gradiens értékeit egyetlen tanítási adatra. Mivel a tanítás folyamán a teljes tanítási adathalmazzal többször elvégezzük a fenti számítást, az eredmények aggregálása is szükségessé válik. Ezt egy egyszerű átlagolással végezzük el.

$$\frac{dC}{dw_1} = \frac{1}{r} \sum_{x=1}^r \frac{dC_x}{dw_1} \quad (171)$$

$$\frac{dC}{dw_2} = \frac{1}{r} \sum_{x=1}^r \frac{dC_x}{dw_2} \quad (172)$$

$$\frac{dC}{db_1} = \frac{1}{r} \sum_{x=1}^r \frac{dC_x}{db_1} \quad (173)$$

$$\frac{dC}{db_2} = \frac{1}{r} \sum_{x=1}^r \frac{dC_x}{db_2} \quad (174)$$

, ahol r a tanítási adathalmaz rekordszáma.

$$\theta_{t+1} = \theta_t - \eta \nabla C(\theta_t) \quad (175)$$

A fenti képletben θ a modellparamétereket, t az epoch sorszámát és η a tanulási rátát jelöli.

5.7 A legmeredekebbeljő- eljárás (gradient descent - GD) fő típusai

Az eljárás típusai elsősorban abban különböznek, hogy a súlyok és konstansok módosítása előtt mennyi adatot dolgoznak fel. Minél több adatot dolgoz fel egy adott eljárástípus, annál pontosabb lesz a kiszámított gradiens, de egyúttal az egyetlen paraméterfrissítésre jutó számítási idő is megnő (Ruder, 2017).

5.7.1 Kötegelt GD (batch gradient descent)

Ez az eljárás a teljes tanítási adathalmazt feldolgozza minden paraméterfrissítés előtt. A frissítések között relatív hosszú idő telik el, de a frissítések mindig a lehető legmeredekebb irányban csökkentik a költségfüggvényt. Az eljárás determinisztikus, azaz, ha az eljárást ugyanazonokon az adatokon többször futtatjuk, úgy az eredmény is mindig ugyanaz lesz. Az eljárás konvex felületen a globális minimumhoz konvergál, de ha a felület nem konvex úgy egy lokális minimumpontra megragadhat. Az eljárás másik hátránya, hogy egyszerűen nem tud feldolgozni olyan nagy adathalmazokat, amelyek egyszerre nem férnek be a feldolgozást végző számítógép vagy szerver memóriájába (Ruder, 2017). A módszer használata szintén előnytelen azokban az esetekben, amikor a modellünket online módon, folyamatosan új adatokkal szeretnénk frissíteni.

5.7.2 Sztochasztikus GD (stochastic gradient descent)

Ez az eljárás minden egyes tanítási adat után frissíti a paramétereket, ezért gyors és az online tanítást is lehetővé teszi. A gyakori paraméterfrissítések miatt a költségfüggvény értéke erősen ingadozik. Az eljárás sztochasztikus jellege miatt lehetővé teszi, hogy az algoritmus kiszabaduljon egy lokális minimumból és esetleg egy jobb eredményhez ugorjon tovább. Az eljárás hátránya, hogy nehezen érkezik meg a minimumpontokhoz, mivel gyakran túllő azokon. A túllövés problémája a kötegelt GD esetében is előfordulhat, de a sztochasztikus GD esetében ez a jelenség sokkal gyakrabban fordul elő, így komolyabban akadályozza a konvergenciát. A problémára az jelent megoldást, ha az algoritmus tanulási rátáját folyamatosan csökkentjük a tanítási folyamat előrehaladtával. A tanulási ráta egy olyan paraméter, ami a legmeredekebblejtő-algoritmusoknál a lépések méretét befolyásolja. Az előbbi megoldás azt eredményezi, hogy az algoritmus egyre kisebb lépéseket tesz, azaz egyre kisebb mértékben módosítja a súly- és konstanstag-paramétereket (Ruder, 2017).

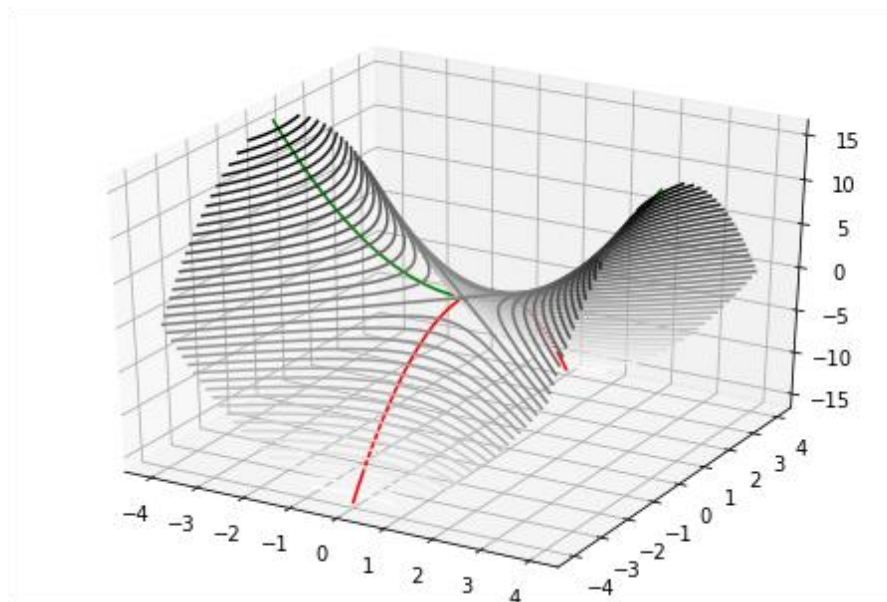
5.7.3 Miniköteg-GD (mini-batch gradient descent)

Ez az eljárás a két korábbi előnyeit ötvözi. Lényege, hogy a paraméterek frissítése nem minden elem után, és nem a teljes adathalmaz feldolgozása után történik, hanem köztes megoldásként egy megadott kötegméret elérése után. Az eljárás előnye az, hogy a paraméterfrissítések által okozott költségfüggvény-variáncia csökken, így az algoritmus stabilabban konvergál a minimumköltség felé.

5.8 A GD-eljárásokkal kapcsolatos problémák

A tanulási ráta egy olyan hiperparaméter, amely a legmélyebb leszállás algoritmusoknál a lépések méretét befolyásolja. A GD-algoritmusok működését jelentősen befolyásolja a tanulási ráta, viszont megfelelő megválasztása egyáltalán nem triviális feladat. Ha a tanulási ráta túl kicsi, akkor a GD-algoritmus rendkívül lassan fog a minimum felé konvergálni, ha pedig túl nagy, akkor a GD-algoritmus nagy költségfüggvény-fluktuációt fog produkálni, és gyakran túl fog lőni a minimumértéken. Ezen kívül problémát jelenthet, ha bizonyos esetekben a tanulási rátát szükséges a tanulási folyamat során változtatni, például csökkenteni kell a tanulási folyamat végén. Szintén problémát jelenthet, ha a neurális háló jellegzetességeiből adódóan nem kívánatos, hogy minden paramétert azonos tanulási rátával módosítsunk (Ruder, 2017).

44. ábra: Nyeregfelület és a legmeredekebb lejtő algoritmus



Forrás: (Li-Bland, 2018)

Egy másik probléma, hogy amennyiben a költségfüggvény által képzett felület nem konvex, úgy nincs arra garancia, hogy az egyszerű GD algoritmusok nem egy globális minimumot találnak meg, hanem egy lokálisat. Ez különösen nagy problémát jelent olyan helyzetekben, ahol a költségfüggvény ún. nyeregfelületet képez, azaz az egyik dimenzió irányába csökken, míg a másik dimenzió irányába növekszik (Dauphin, et al., 2014).

5.9 GD optimalizálási algoritmusok

5.9.1 A momentum-GD optimalizálási algoritmus

A Rumelhart és szerzőtársai (1986) által először leírt momentumot alkalmazó optimalizálási algoritmus arra a problémára ad megoldást, amikor a minimum környékén az egyik dimenzió irányába a felület lényegesen jobban görbül, mint a másik irányba. Ilyenkor a súlyfrissítések után a költségfüggvény értéke a meredek irányba erős oszcilláló mozgást mutat, míg az alacsony meredekségű irányban csak nagyon lassan halad (Qian, 1999). A momentum optimalizálási eljárás a korábbi lépés mértékét egy paraméterrel csökkentve hozzáadja az aktuális lépéshez. Az eljárás eredménye, hogy az egymást követő ellentétes irányú mozgások kioltják egymást, míg az azonos irányba történő lépések összeadódnak, ezzel egyfajta lendületet adva az algoritmus haladásának.

$$v_t = \gamma v_{t-1} + \eta \nabla C(\theta_t) \quad (176)$$

$$\theta_{t+1} = \theta_t - v_t \quad (177)$$

A fentiekben γ a momentumparaméter, v_t pedig a t . időszak paramétermódosításra használt vektora.

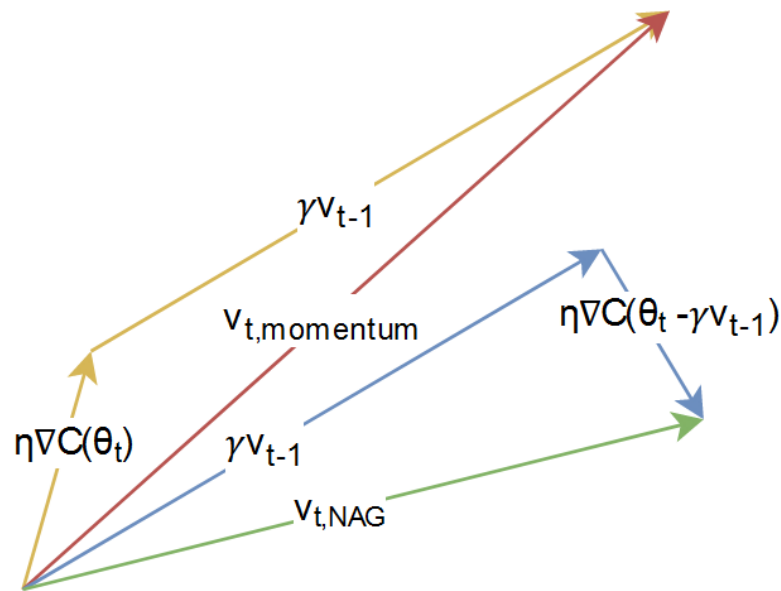
5.9.2 Nesterov-féle gyorsított gradiens (Nesterov accelerated gradient, NAG)

A momentum optimalizálási algoritmus problémája az, hogy a gradiensvektor értékét még azelőtt kiszámolja, hogy az előző lépésből átemelt, lendületet adó módosító tétel (γv_{t-1}) hatása érvényesülne. Ennek következménye, hogy az algoritmus könnyen túllőhet a minimumpontra. A NAG-eljárás (Nesterov, 1983) lényege, hogy a gradiensvektort a lendülettag hozzáadása után számoljuk ki.

$$v_t = \gamma v_{t-1} + \eta \nabla C(\theta_t - \gamma v_{t-1}) \quad (178)$$

$$\theta_{t+1} = \theta_t - v_t \quad (179)$$

45. ábra: A momentum- és a NAG-algoritmus közötti különbség



Forrás: (saját szerkesztés)

A 45. ábra bemutatja a momentumalgoritmus lépéseit (sárga vektorok) és eredményét (piros vektor), valamint a NAG algoritmus lépéseit (kék vektorok) és eredményét (zöld vektor).

5.9.3 Adagrad (adaptive gradient algorithm)

A tanulási ráta hiperparaméter megfelelő megválasztása bizonyos esetekben nem lehetséges, mivel az egyik dimenzió szempontjából optimális érték egy másik dimenziónál lehetséges, hogy nem megfelelő. Az ilyen helyzetekre megoldást jelenthetne, ha a modellekben minden dimenzióra külön tanulási rátát tudnánk megadni, de nagy méretű neurális hálók esetében ezek manuális megadása rendkívül időigényes lenne. Erre jelent megoldást az Adagrad-eljárás (Duchi, et al., 2011), amely a tanulási rátákat minden paraméterre automatikusan beállítja annak függvényében, hogy az adott paraméter milyen mértékben volt korábban módosítva. Gylberth (2018) levezetését követve az Adagrad a lenti eljárás szerint korrigálja a tanulási rátát:

Legyen $g_t^{(i)}$ a t . időponthoz tartozó gradiensvektor i . paraméterhez tartozó eleme.

$$g_t^{(i)} = \nabla C(\theta_t^i) \quad (180)$$

g_t a teljes, m elemű gradiens vektor:

$$g_t = \begin{bmatrix} g_t^{(1)} \\ g_t^{(2)} \\ \vdots \\ g_t^{(m)} \end{bmatrix} \quad (181)$$

A G_t -mátrix a g_t -gradiensvektorok keresztszorzatának összege a t . időpontig.

$$G_t = \sum_{\tau=1}^t g_{\tau} g_{\tau}^T = \sum_{\tau=1}^t \begin{bmatrix} g_{\tau}^{(1)} \\ g_{\tau}^{(2)} \\ \vdots \\ g_{\tau}^{(m)} \end{bmatrix} \begin{bmatrix} g_{\tau}^{(1)} & g_{\tau}^{(2)} & \dots & g_{\tau}^{(m)} \end{bmatrix} = \sum_{\tau=1}^t \begin{bmatrix} g_{\tau}^{(1)} g_{\tau}^{(1)} & \dots & g_{\tau}^{(1)} g_{\tau}^{(m)} \\ \vdots & \ddots & \vdots \\ g_{\tau}^{(m)} g_{\tau}^{(1)} & \dots & g_{\tau}^{(m)} g_{\tau}^{(m)} \end{bmatrix} \quad (182)$$

Belátható, hogy a G_t főátlójában található értékek a következő képlettel is megkaphatók:

$$G_t^{(i,i)} = \sum_{\tau=1}^t (g_{\tau}^{(i)})^2 \quad (183)$$

Ezt felhasználva a paraméterek frissítése a következő képlettel történik:

$$\theta_{t+1} = \theta_t - \eta (\varepsilon I + \text{diag}(G_t))^{-1/2} \cdot g_t \quad (184)$$

$\text{diag}(G_t)$ a G_t -mátrix főátlójából képzett diagonális mátrix, I az egységmátrix ε egy kis (10^{-8} nagyságrendű) korrekciós tényező, ami a nullával való osztás elkerüléséhez szükséges.

$$\theta_{t+1} = \theta_t - \eta \left(\begin{bmatrix} \varepsilon & 0 & \dots & 0 \\ 0 & \varepsilon & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \varepsilon \end{bmatrix} + \begin{bmatrix} G_t^{(1,1)} & 0 & \dots & 0 \\ 0 & G_t^{(2,2)} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & G_t^{(m,m)} \end{bmatrix} \right)^{-1/2} \cdot g_t \quad (185)$$

$$\theta_{t+1} = \theta_t - \begin{bmatrix} \frac{\eta}{\sqrt{\varepsilon + G_t^{(1,1)}}} g_t^{(1)} \\ \frac{\eta}{\sqrt{\varepsilon + G_t^{(2,2)}}} g_t^{(2)} \\ \vdots \\ \frac{\eta}{\sqrt{\varepsilon + G_t^{(m,m)}}} g_t^{(m)} \end{bmatrix} \quad (186)$$

A fenti eredményből látszik, hogy az Adagrad-algoritmus nem egy fix η tanulási rátával szorozza a gradiensvektort, hanem elemenként automatikusan külön korrigálja

azt. G_t a korábbi időszakok gradienseinek négyzetes összege, azaz $G_t^{(1,1)}$ annál nagyobb lesz, minél nagyobb mértékben módosult az 1. paraméter a korábbi tanítási időszakok alatt. Tehát minél nagyobb módosításokon esik át egy paraméter, annál nagyobb számmal osztjuk a rá vonatkozó tanulási rátát, így az idő előrehaladtával egyre nehezebben fog az adott paraméter változni. Az Adagrad-optimalizálásjeljárás hibája, hogy a tanulási folyamat során az osztóban folyamatosan nő a korábbi gradiensek értéke, így egy idő után a tanulási ráta olyan kicsire csökken, hogy a neurális háló már nem lesz képes új összefüggések megtanulására.

5.9.4 Adadelta

A módszert Zeiler (2012) vezette be az Adagrad-algoritmus továbbfejlesztéseként. Az Adagrad hibája az, hogy a tanulási ráta minden időszakban monoton csökken, ezért a frissítések egyre kisebbek lesznek. Az Adadelta alapötlete, hogy a gradiensek négyzetes összegét csak egy w hosszú időablakban gyűjti és nem a teljes időszakon. Az alapötlet problémája az, hogy kiszámításához w darab négyzetes gradiens értéket kell tárolni, ami túlságosan nagy erőforrást igényel. Ennek feloldására az Adadelta a négyzetes gradiensek w hosszú sorozata helyett a négyzetes gradiensek exponenciális mozgóátlagát használja:

$$E[G_t^2]_t = \rho E[G_{t-1}^2]_{t-1} + (1 - \rho)g_t^2 \quad (187)$$

A fenti képletben ρ a csökkenési állandó, amihez tipikusan 0,9 körüli értéket használunk.

$$RMS[G_t]_t = \sqrt{E[G_t^2]_t + \varepsilon} \quad (188)$$

Az RMS (root mean squared)-korrekcióstényező hátránya, hogy nem ugyanabban a mértékegységben van, mint a paraméterek, így közvetlenül nem használható. Az $RMS[G_t]_t$ értékét a következő módon skálázzuk:

$$E[V_t^2]_t = \rho E[V_{t-1}^2]_{t-1} + (1 - \rho)v_t^2 \quad (189)$$

$$E_0 = 0 \quad (190)$$

$$RMS[V_t]_t = \sqrt{E[V_t^2]_t + \varepsilon} \quad (191)$$

$$v_t = -\frac{RMS[V_{t-1}]_{t-1}}{RMS[G_t]_t} g_t \quad (192)$$

$$\theta_{t+1} = \theta_t + v_t \quad (193)$$

A fentiekben v_t a súlyváltoztatási vektor. Az ε tag használatát az indokolja a tanulási ráta szerepét átvevő $RMS[V_t]_t$ képletében, hogy RMS_0 nullától eltérő értéket kapjon mivel így $RMS_0 = \sqrt{\varepsilon}$.

5.9.5 RMSprop (Root Mean Squared Propagation)

Az eljárás használatát először Geoffrey Hinton javasolta egy online kurzus keretében (Hinton, n.d.). Az RMSprop szintén az Adagrad csökkenőtanulási ráta-problémájára keres megoldást, és nagyban hasonlít az Adadelta-eljáráshoz, azzal a különbséggel, hogy itt az utolsó, skálázási lépés kimarad.

$$E[V_t^2]_t = \rho E[V_{t-1}^2]_{t-1} + (1 - \rho)v_t^2 \quad (194)$$

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{E[G_t^2]_t + \varepsilon}} g_t \quad (195)$$

5.9.6 Adam (Adaptive Moment Estimation)

Az Adam-optimalizációs eljárás (Kingma & Ba, 2014) szintén különböző tanulási rátákat számol az egyes paraméterekre. Az eljárás az előzőkhöz hasonlóan a gradiensek négyzeteinek csökkenőátlagát használja, de emellett a gradiensek csökkenőátlagára is épít, így lényegében az RMSprop és a momentum ötvözeteként viselkedik.

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (196)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (197)$$

Az m_t az első momentum, a v_t pedig a második momentum, az eljárás neve ezek használatából ered. Átlagolás első lépéseként mindkét változó nulla értékről indul, ezért az átlagok a zérus felé hibáznak.

$$m_0 = 0 \quad (198)$$

$$v_0 = 0 \quad (199)$$

Ennek korrigálására az eljárás szerzői a következő korrekciót javasolják:

$$\widehat{m}_t = \frac{m_t}{1-\beta_1^t} \quad (200)$$

$$\widehat{v}_t = \frac{v_t}{1-\beta_2^t} \quad (201)$$

A frissített paraméterek vektorát a következő képlettel kapjuk meg:

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\widehat{v}_t + \epsilon}} \widehat{m}_t \quad (202)$$

5.9.7 AdaMax

Az AdaMax (Kingma & Ba, 2014) az Adam-eljárás általánosított változata. Az Adam-eljárásnál használt v_t képlet a következőképpen általánosítható:

$$v_t = \beta_2^p v_{t-1} + (1 - \beta_2^p) |g_t|^p \quad (203)$$

Az Adam esetében a $p = 2$, az AdaMax esetében pedig $p = \infty$.

$$\begin{aligned} u_t &= \lim_{p \rightarrow \infty} (\beta_2^p v_{t-1} + (1 - \beta_2^p) |g_t|^p)^{1/p} \\ &= \max(\beta_2 \cdot u_{t-1}, |g_t|) \end{aligned} \quad (204)$$

A fenti u_t értéke az Adam-eljárásnál használt v_t értékétől eltérően nem hibázik a nulla felé, így az Adamnál használt korrekció az u_t esetében nem szükséges. Ez alapján a paraméterek frissítése a következő képlettel történik:

$$\theta_{t+1} = \theta_t - \frac{\eta}{u_t} \widehat{m}_t \quad (205)$$

5.9.8 Nadam

Ahogy az Adam-eljárás a momentum és a RMSprop ötvözetete, úgy a Nadam a momentum továbbfejlesztett változata, a NAG és az RMSprop egységbe foglalt optimalizációs eljárása. Dozat (2016) a Nadam-eljárás levezetésének első lépéseként a NAG-algoritmus egy hatékonyabb változatát javasolja.

Ez eredeti NAG-algoritmus:

$$g_t = \nabla_{\theta_t} J(\theta_t - \gamma m_{t-1}) \quad (206)$$

$$m_t = \gamma m_{t-1} + \eta g_t \quad (207)$$

$$\theta_{t+1} = \theta_t - m_t \quad (208)$$

$$\theta_{t+1} = \theta_t - (\gamma m_{t-1} + \eta g_t) \quad (209)$$

Dozat módosított NAG-megvalósítása:

$$g_t = \nabla_{\theta_t} J(\theta_t) \quad (210)$$

$$m_t = \gamma m_{t-1} + \eta g_t \quad (211)$$

$$\theta_{t+1} = \theta_t - (\gamma m_t + \eta g_t) \quad (212)$$

A módosított NAG-eljárás előnye, hogy a $-\gamma m_{t-1}$ tagot csak egyszer alkalmazza és nem két helyen, mint az eredeti megoldásban. Fontos észrevenni, hogy a módosított NAG-eljárásnál a paraméterfrissítéshez az aktuális időszak momentumértékét (m_t) használjuk, és nem az előző időszakét, mint az eredeti algoritmus szerint. Ugyanezt a logikát követve módosíthatjuk az Adam-eljárást is. Az Adam paraméterfrissítési képlete $\frac{m_{t-1}}{1-\beta_1^t}$ kiemelésével:

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t + \epsilon}} \left(\beta_1 \frac{m_{t-1}}{1-\beta_1^t} + \frac{(1-\beta_1)g_t}{1-\beta_1^t} \right) \quad (213)$$

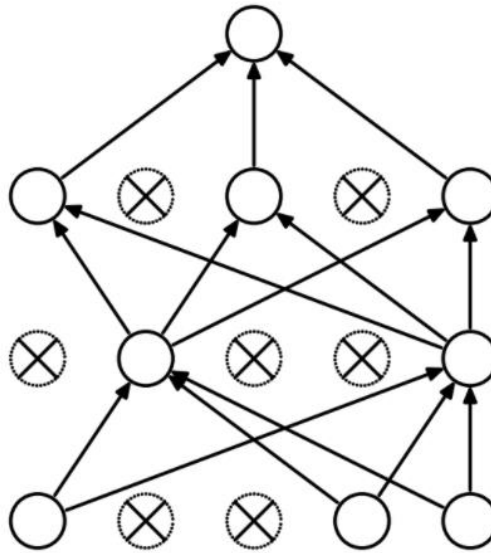
A fenti logika szerint a $\frac{m_{t-1}}{1-\beta_1^t}$ tagot helyettesítjük az aktuális időszak momentumértékével ($\frac{m_t}{1-\beta_1^{t+1}}$), azaz:

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t + \epsilon}} \left(\beta_1 \frac{m_t}{1-\beta_1^{t+1}} + \frac{(1-\beta_1)g_t}{1-\beta_1^t} \right) \quad (214)$$

5.10 Overfitting & Dropout

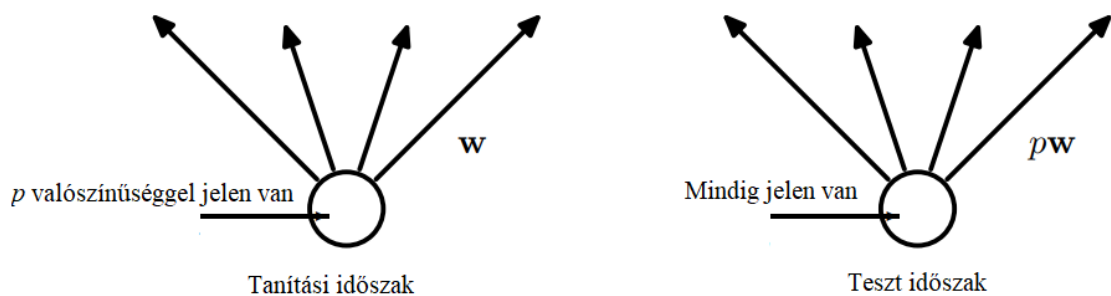
Az overfitting (túltanulás) jelensége az egyik, a neurális hálókkal kapcsolatban felmerülő probléma. Az overfitting többnyire kisebb minták esetén merül fel, és azt jelenti, hogy a modell olyan szabályokat tanul meg, amik csak az adott mintára érvényesek, de a teljes populációra nem. Overfitting esetén a modell jól teljesít a tanítási adatokon, de kifejezetten rosszul a tesztidőszakban. Srivastava és szerzőtársai (2014) az overfitting kezelésére a Dropout-eljárás használatát javasolják, ami a csomópontok véletlenszerű kihagyását jelenti.

46. ábra: A Dropout-eljárás bemutatása



Forrás: (Srivastava, et al., 2014)

A neurális háló különböző szintjeire akár eltérő dropout-hiperparamétert (p) is definiálhatunk. A p -hiperparaméter jelentése az, hogy az adott szint egy csomópontja milyen valószínűséggel nem kerül kihagyásra az adott tanítási szakaszban. Egy csomópont kihagyása esetén a kimenő és bejövő élek is kihagyásra kerülnek. Szintén fontos kiemelni, hogy az eljárás az összes rétegben alkalmazható, azaz a bemenetek, a rejtett rétegek és a kimenetek szintjén is. Mivel a csomópontok működését befolyásolja az, hogy $1-p$ valószínűséggel kihagyjuk őket a tanítási szakaszban, a tesztelési időszakban a súlyukat korrigálni kell, így egy $1-p$ valószínűséggel kihagyott csomópontból induló él súlya a tesztelési időszakban w helyett pw lesz.



Forrás: (Srivastava, et al., 2014) alapján

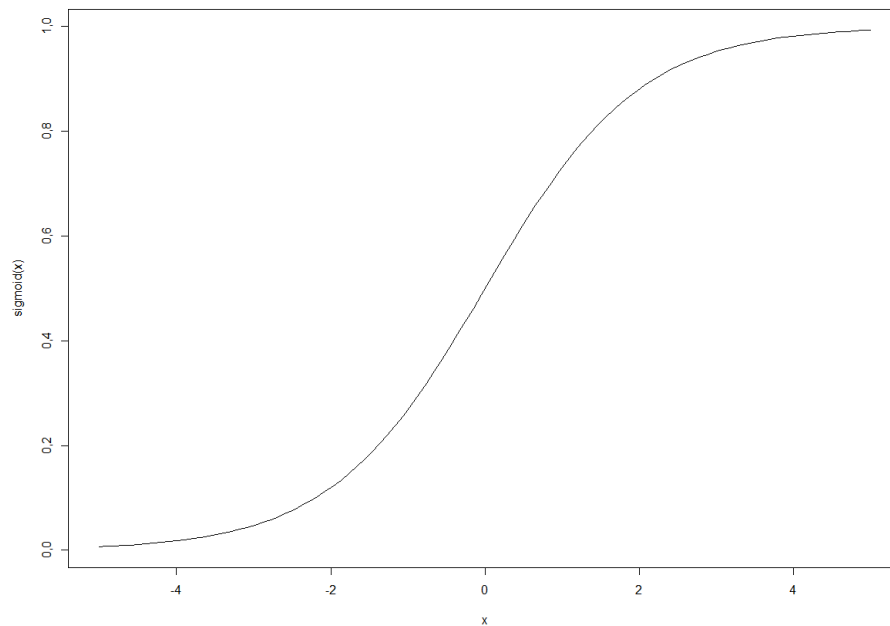
Fontos szempont a modellalkotásnál, hogy a Dropout-eljárás használata többnyire kétszeres epochszám használatát követeli meg ahhoz, hogy a modell megfelelően konvergáljon (Budhiraja, 2016).

5.11 Aktivációsfüggvények

A kimenetek értékét a bemenetektől kapott értékek és az aktivációs függvény határozza meg. A jelen értekezésben a szigmoid-, a tanh- és a Relu-aktivációsfüggvények teljesítménye került tesztelésre az érzékenységvizsgálat során. A szigmoid függvény definíciója a következő (Hagan, et al., 2014):

$$\text{sigmoid}(x) = \frac{1}{1+e^{-x}} \quad (215)$$

47. ábra: Sigmoid-aktivációsfüggvény

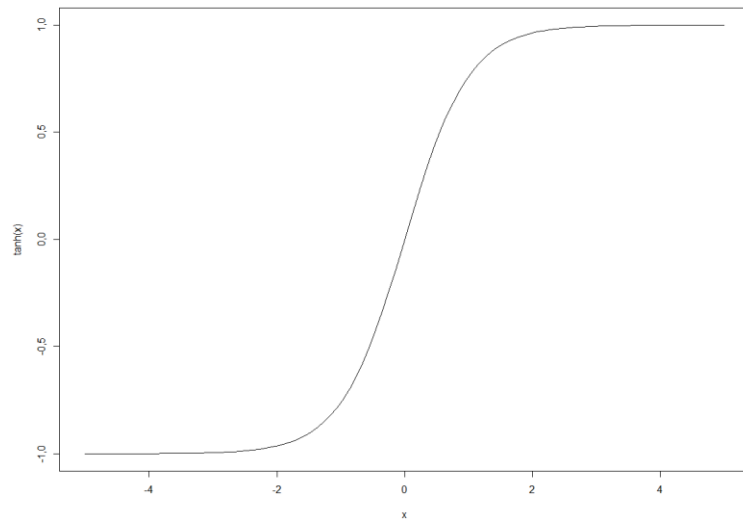


forrás: (saját szerkesztés)

A tanh-aktivációsfüggvény definíciója a következő (Hagan, et al., 2014):

$$\text{tanh}(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (216)$$

48. ábra: Tanh-aktivációsfüggvény

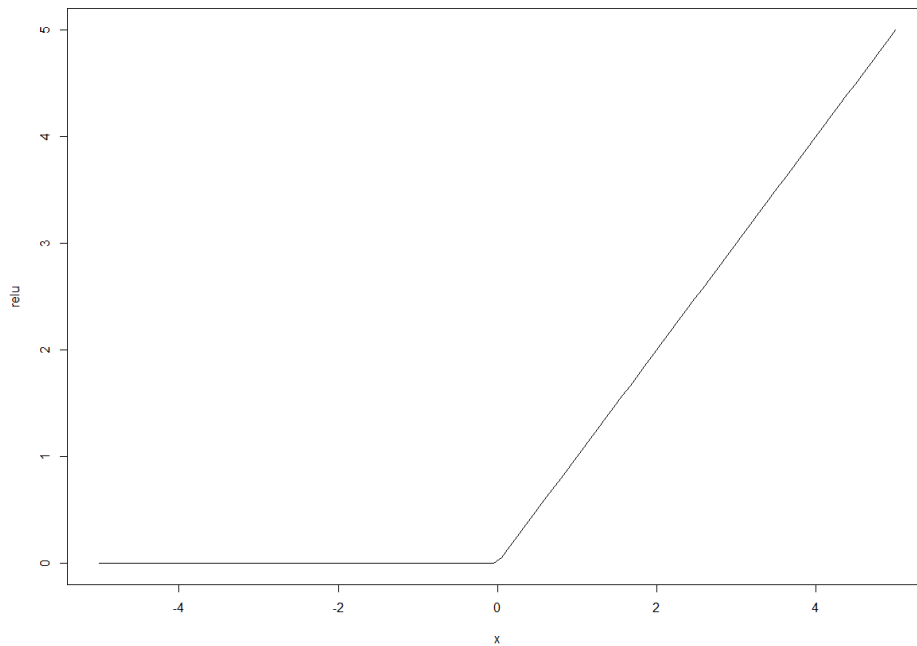


forrás: (saját szerkesztés)

A ReLU (rectified linear activation function)-aktivációsfüggvény egyszerűen a bemenet 0 feletti részét adja vissza. A ReLU-aktivációsfüggvény definíciója (IRMA, 2019):

$$ReLU(x) = \max(0, x) \quad (217)$$

49. ábra: A ReLU-aktivációsfüggvény

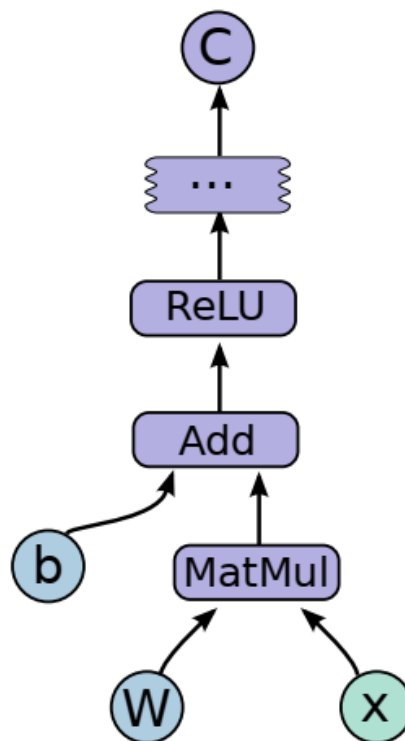


forrás: (saját szerkesztés)

5.12 Implementációs részletek: TensorFlow & Keras

A TensorFlow a Google nyílt forráskódú fejlesztése, amely a 2011-ben indult Google Brain projekt részeként jött létre. A TensorFlow célja, hogy rugalmas és egységes interfészt adjon a gépi tanulással kapcsolatos algoritmusoknak és futtatásuknak, és kiszolgálja mind a kutatók mind a Google termékeinek igényeit. A TensorFlow lehetővé teszi, hogy ezek a modellek a hardvereszközök széles skáláján futtathatóak legyenek, többek között mobiltelefonokon, egyetlen számítógép processzorán és videokártyáján, vagy nagy elosztott rendszereken (Abadi, et al., 2015).

50. ábra: TensorFlow-számításgráf-példa



Forrás: (Abadi, et al., 2015, p. 3)

A TensorFlow-számítások irányított gráfokként írhatóak le, ahol a csomópontok a műveleteknek felelnek meg. A tenzorok n -dimenziós adatobjektumok, amelyek ezen a gráfon haladnak végig a bemenetektől a kimenetekig, miközben a csomópontok különféle

műveleteket végeznek rajtuk. A különböző műveletek az egyes támogatott eszközökön külön-külön, az eszköz hardverkövetelményeinek megfelelően kerültek megvalósításra ezzel garantálva a rendszer platformfüggetlen kialakítását.

A Keras a TensorFlow csomagra épülő könyvtár, ami a neurális hálók kezelését teszi egyszerűbbé (Ketkar, 2017). A könyvtár korábbi változatai több számításgráf-technológiát is támogattak, de a legújabb verzióban már csak a TensorFlow megengedett. A Keras- és a TensorFlow-csomagok eredetileg csak Python környezetben voltak elérhetőek, de újabb fejlesztések következtében már az R programnyelvből is használhatók.

5.13 ANN-AS-modell

Az ANN-AS-modell lényege, hogy az arcsine nevű parametrikus statisztikai modell aktuális időszakban várható hibáját becsüli meg az előző időszak hibája (ε_t) alapján, majd az arcsine-modell becslését korrigálja ezzel az értékkel. A hiba számításának alapja az Acerbi–Székely (2019) szerzőpáros által leírt megvalósult várható többletveszteség (realized expected shortfall, $ESRMA_\alpha$), azzal a különbséggel, hogy a megvalósult ES a jelen modellben használt $ESR_{\alpha,t}$ értékek átlaga.

$$ESR_{\alpha,t} = v_t + \frac{1}{\alpha}(x_t + v_t)_- \quad (218)$$

$$ESRMA_\alpha \equiv \frac{1}{T} \sum_{t=1}^T [ESR_{\alpha,t}] \quad (219)$$

$$\varepsilon_t = ESR_{\alpha,t} - ES_{arcsine,t} \quad (222)$$

$$ES_{ANN-AS,t} = ES_{arcsine,t} + \varepsilon_{estimated,t} \quad (223)$$

A fentiekben α a konfidenciaszint (97,5%), x_t a t. időszak hozama, ami veszteségek esetén negatív értéket vesz fel, v_t a kockázatotott érték, ami egy pozitív szám, $\varepsilon_{estimated,t}$ pedig a neurálisháló által becsült hiba a t. időszakban. $ESR_{\alpha,t}$ egy olyan érték, amely megegyezik a VaR t. periódushoz tartozó értékével akkor, ha az adott napi veszteség a VaR értéke alatt marad, de a konfidenciaszinttel arányos mértékben növekszik akkor, ha az adott napi veszteség meghaladja a VaR értékét. Az $(x_t + v_t)_-$ egy olyan tag, aminek értéke nulla, ha $x_t + v_t$ pozitív és $-(x_t + v_t)$, ha $x_t + v_t$ értéke

negatív. $x_t + v_t$ akkor lesz negatív, ha az aktuális napi hozam (x_t) nagyobb veszteséget mutat, mint a t . időszakra becsült kockázatos érték, azaz egy ún. VaR-hiba (violation) következik be. A képletből világos, hogy amennyiben nincs violation, úgy az ESR értéke a kockázatos értékkel egyezik. Ha a teljes T periódusból álló időszak alatt nincs egyetlen a VaR értékét meghaladó veszteség sem, akkor az $ESRMA_\alpha$ értéke megegyezik a VaR értékek átlagával. Ezen az értéken fog minden violation annak mértékével arányosan emelni.

$$\varepsilon_{estimated,t} = f(\varepsilon_{t-1}) \quad (224)$$

Az ANN-AS-neurálsháló egy bemenettel és egy kimenettel dolgozik. Két rejtett rétegben 4-4 neuron található. Az inputréteg és a rejtett rétegek a *tanh*-aktivációsfüggvényt alkalmazzák, míg a kimenetnél a *sigmoid* aktivációs függvény használata adta a legjobb eredményt.

A modell nem alkalmazza a Dropout-eljárást. A hibafüggvény az átlagos abszolút hiba (mae), a legmeredekebb lejtő algoritmus optimalizációját pedig a Nadam-eljárás végzi. A kötegméret 10, az epochok száma pedig 200. A modell tanítása az arcsine-modell hibaidősorával történik, olyan adathalmaz képzésével, amiben az előző időszak hibája párban van az aktuális időszak hibájával.

$$\varepsilon = \begin{bmatrix} \varepsilon_1 & \varepsilon_2 \\ \varepsilon_2 & \varepsilon_3 \\ \vdots & \vdots \\ \varepsilon_{T-1} & \varepsilon_T \end{bmatrix} \quad (225)$$

A neurális háló egzakt leírását a következő kódrészlet adja meg:

```

model <- keras_model_sequential()
model %>%
  layer_dense(units = 1, activation = 'tanh', input_shape = 1) %>%
  layer_dense(units = 4, activation = 'tanh') %>%
  layer_dense(units = 4, activation = 'tanh') %>%
  layer_dense(units = 1, activation = 'sigmoid')

model
%>%
compile(loss = 'mae', optimizer = optimizer_nadam(), metrics = c('mae') )

model %>% fit(
  X_train, Y_train,
  epochs = 200, batch_size = 10, verbose = 0,
  validation_split = 0.1
)
Y_predict = model %>% predict ( X_test )

```

A tanítás során az egyes részvények idősorai felhasználásra kerülnek az arcsine-modell becslésére, amely alapján az $ESR_{\alpha,t}$ tag, majd végül az ε_t hibatag kerül kiszámítására. Minden részvény minden tanítási időszakra külön neurális hálót képez, tehát sem a részvények, sem az időszakok között nem keverednek a tanítási idősorok.

A modellépítés alatt számos, a fent bemutatottnál minden szempontból lényegesen komplexebb modell tesztelése is megtörtént, de azok eredményei nem voltak kielégítőek, a jól teljesítő valószínűségeloszlásokon alapuló modellek teljesítményétől elmaradtak. Az eredmények alapján a fent leírt egyszerű megközelítés a mutatkozik a legalkalmasabbnak az ES becslésére.

5.14 Modellezési hibák & korábbi iterációk

Jelen értekezésben egy olyan neurális hálón alapuló modell került bemutatásra, ami szignifikánsan jobb eredményt ad a korábban bemutatott modelleknél. A modell számos korábbi, kevésbé jól vagy kifejezetten rosszul teljesítő változat átdolgozásának eredményeként született, melyek közül a lényegesebb modellek bemutatásra kerülnek a 9. fejezetben. A következő szakaszban leírt sikertelen próbálkozások azért kerülnek bemutatásra, mert hasznosak lehetnek más, hasonló területen végzett modellezési folyamat során. Véleményem szerint a buktatók többsége más területen is felmerülhet, így releváns lehet.

5.14.1 Neuralnet és a threshold-hiperparaméter

A neuralnet-csomag egy R-alapú neurálisháló-implementáció (Fritsch & Guenther, 2019). A csomag előnye, hogy teljes mértékben R-alapokon került megvalósításra, így egyszerűen telepíthető és viszonylag könnyen felépíthető belőle egy neurális hálón alapuló modell. Az értekezésben ES-becslésre használt kezdeti modellek is ezt a csomagot használták. A modell optimalizálása során a legnagyobb áttörést a „threshold” hiperparaméter beállítása jelentette. A threshold-hiperparaméter adja meg az algoritmus számára, hogy melyik az a hibaérték, amelynél a tanítási folyamatot be lehet fejezni, mivel már a neurális háló kellően pontos eredményt ad a tanítási adatokon. Az ES-modell lényeges javulásához vezetett ennek a threshold paraméternek a pontos beállítása. Túlságosan kicsi threshold-érték megadása ahhoz vezetett, hogy az algoritmus hibát dobott és nem futott le, vélhetőleg azért, mert az optimalizálás során nem volt képes olyan megoldást találni, ami az alacsony threshold-értéknek megfelelő mértékű hibát ért volna el a tanítási adatokon. Ettől függetlenül az optimálisnak bizonyuló threshold értéke (0,005) lényegesen kisebb volt a csomag alapértelmezett értékénél (0,01), azaz annál az értéknél, amelyet az algoritmus a hiperparaméter explicit megadása nélkül használ.

5.14.2 Az eredmények aggregációja és a vizuális megjelenítés

Viszonylag triviális hiba, de teljesen félrevezető következtetésekhez vezet, ha az eredmények nem megfelelően kerülnek összevonásra. Az értekezésben az ES-modell becslésének pontosságát az Acerbi és Székely (Acerbi & Szekely, 2017) által javasolt Z-értékkel mérem, aminek egyik tulajdonsága, hogy pozitív és negatív értéket is felvehet. Ebből adódik, hogy amennyiben az összes részvény összes Z-értékét egyszerűen összeadom, úgy az alulbecsült időszakok korrigálják a felülbecsült időszakokat, így ilyen esetben egészen rossz becslések is kis összesített Z értéket adnak. Az ilyen jellegű triviális hibák kiszűrésének legegyszerűbb módja az eredmények vizualizációja, ahol egyből feltűnik, ha egy olyan modell került az első helyre, ami éves vagy részvényenkénti bontásban láthatóan távolabb van az optimálistól, mint más modellek. A modellek Z-értékeinek jelenlegi összegzése részvényenkénti éves bontásban történik, majd ezek abszolút értéke kerül összeadásra.

5.14.3 Neuralnet vagy Keras/TensorFlow

A neuralnet csomag hátránya, hogy bizonyos hiperparaméterek rögzítve vannak benne, vagy azok csak korlátozottan módosíthatóak. Ilyen hiperparaméter például a hibafüggvény (err.fct). A neuralnet-csomag dokumentációja alapján jelenleg a hibafüggvény vagy a négyzetes hibák összege (sse), vagy a keresztentropia (ce) lehet. A neuralnet másik megkötése, hogy az aktivációs függvények nem módosíthatók szintenként, illetve nem alkalmazhatóak olyan eljárások, mint például a túltanulás ellen a csomópontok véletlenszerű kihagyásával dolgozó Dropout. A probléma feloldása egy szélesebb körben, nagyobb paraméterezhetőséggel bíró csomag használata, ilyen például a Google által közzétett Keras/Tensorflow.

5.14.4 Veszteség- vagy hibafüggvény: négyzetes vagy abszolút hiba

A tanítási folyamat során a neurális háló egy célfüggvény segítségével értékeli, hogy egy adott becslés mennyire tér el az elvárt értéktől. A hálózat súlyait a hibavisszaterjesztés során ennek a különbségnek az értékéből kiindulva frissítjük. A két leggyakrabban használt veszteségfüggvény az átlagos négyzetes hiba (MSE) és az átlagos abszolút hiba (MAE). A várható többletvesztés becslésére használt neurálisháló-modell korábbi változatai a négyzeteshiba-függvényt (MSE) használták, de ez kifejezetten rossz eredményeket produkált. A neurális hálókkal kapcsolatos irodalom viszonylag kis hangsúlyt fektet a veszteségfüggvény megválasztására, a döntés így nem feltétlenül tűnik különösebben lényegesnek. Az értekezésemben bemutatott modellnek viszont van egy lényeges sajátossága. A legtöbb neurális háló esetében azért nem fontos a hibafüggvény megválasztása, mert elegendő arra figyelni, hogy a tesztelési szakaszban majd ugyanez a hibafüggvény szerint értékeljék a modellt. A jelen értekezésben bemutatott várható többletvesztést becsülő modell esetében a tanítási szakaszban a hibát a MAE- vagy az MSE-függvények adták meg. A tesztelési szakaszban viszont a modell értékelését nem ezek használatával, hanem a szakirodalomban javasolt backtesting eljárással végzem el, amire azért van szükség, hogy az eredményeket megfelelően lehessen aggregálni. A modell kezdeti MSE-alapú változata tehát összességében rossz eredményeket adott. A hibakeresés során kiderült, hogy a modell néhány részvény esetében jól becsült, míg más esetekben rosszul, ráadásul a becslési hibák konzisztensen az egyik oldalra estek, azaz egy adott részvény esetén vagy folyamatosan túlbecsülték az ES értékét vagy konzisztensen alábecsülték azt. A hibát az okozta, hogy a backtesting-eljárásban

hibafüggvényként használt Z-érték nem használ négyzetes tagokat, azaz Z értéke alapvetően abszolút távolságon alapszik. A négyzetes hibát alkalmazó modell azokban az esetekben adott jó eredményt, amikor a kiválasztott részvény eloszlása viszonylag szimmetrikus volt, mivel így a nagy súlyt kapó extrém hibák mindkét irányban megjelentek és a tanítási folyamat során a neurális háló nem hajlott el sem az alábecslés sem a felülbecslés irányába. A legtöbb részvény esetében a hozamok eloszlása ferdeséget mutatott vagy a veszteségek, vagy a nyereségek irányába, így az extrém hibák költségének csökkentése érdekében a neurális háló olyan súlyokat tanult meg, amelyekkel túlkorrigálta az adott eredményt a hozameloszlás ferdeségének megfelelően. A probléma feloldása a tanulási fázisban az átlagosabszolúthiba-függvény használata volt.

5.14.5 Hibadefiníció: ESR és ESR-MA

Az ANN-AS-neurálisháló számításánál komoly előrelépést jelentett az arcsine-modell hibaidősorának elsőre viszonylag lényegtelennek tűnő változtatása. A modell első változatában a hibaidősor az ESRMA-képlet szerint került kiszámításra, amely a backtesting-eljárás szerint kiszámított valós várható többletveszteség értéke (Acerbi & Székely, 2019). Az ESRMA-mutató, ahogy az a korábbi képletekből is kiderül, az ESR értékének az adott ellenőrzési időszakra vonatkozó átlaga. A korábban leírtak alapján az ESR egy olyan érték, amely megegyezik a VaR értékével akkor, ha az adott napi veszteség a VaR értéke alatt marad, de a konfidenciaszinttel arányos mértékben növekszik akkor, ha az adott napi veszteség meghaladja a VaR értékét. A modell kezdeti változata tehát a következő képletek alapján működött:

$$\varepsilon_i = ESRMA_{\alpha,i} - ES_{arcsine,i} \quad (226)$$

$$ES_{ANN-AS,i} = ES_{arcsine,i} + \varepsilon_{estimated,i} \quad (227)$$

A neurálisháló-modell inputadatsora a jelenlegi változathoz hasonlóan az előző időszak hibája (ε_{i-1}) volt, az elvárt output adatsora pedig a modell fenti képlettel leírt aktuális hibája (ε_i). A modell kifejezetten rossz eredményeket adott, feltehetőleg annak köszönhetően, hogy az ESR-MA képletében használt átlagolás a hiba idősorban található információ mértékét csökkentette, hiszen az ESR-MA mutató az átlagolás következtében elveszti azt az információt, hogy a VaR violation az előző időszakban, vagy azt megelőzően fordult elő. A neurális háló teljesítménye drasztikusan javult azzal, hogy a hibafüggvény módosításra került, és nem az ESR-MA-mutatót, hanem egyszerűen az ESR-részeredményt használta a hiba meghatározásához.

6 Backtesting-módszerek

6.1 A backtesting eljárások célja

A számos rendelkezésre álló becslési modell közötti sorrend felállítására lehetőséget adnak olyan szempontrendszerek, melyek a kockázati mérőszámokkal szemben általános követelményeket állítanak, mint például Artzner, et al. (1999) koherenciakritériumai, vagy Giorgi (2005) munkája. Ezeknek a követelményeknek egy lényeges tulajdonsága, hogy a mutatókról a matematika eszközkészletével bizonyítható, hogy az egyes tulajdonságoknak megfelelnek vagy sem. Ezen túlmenően viszont továbbra is szükséges marad a fennmaradó, és az általános követelményeknek megfelelő modellek közötti sorrend felállítása, és annak meghatározása, hogy egy adott kockázati modell bemeneti paraméterei valóban megfelelően kerültek-e meghatározásra.

A 2007-es pénzügyi válság megmutatta, hogy a gyakorlat során használt modellek gyakran alulbecsülték a kockázatot (Sheedy, 2008). A nagyobb stabilitás elérésének érdekében fontos tehát a következő kérdések megválaszolása: Egy koherens kockázati mutató mely becslési eljárásai adják a legpontosabb értéket a kockázatra nézve? A becslési minta egy évre vagy hosszabb időszakra szóljon? Hogyan validálhatók az elkészült modellek? Ezekre a kérdésekre adhatnak választ a modellek visszaellenőrzésére használt backtesting-módszerek.

A backtesting-módszerek lehetőséget adnak a modellek validálására, és a modellek közötti választásra. A megfelelő backtesting-eljárásoknak minél kevesebb feltételezéssel kell élniük az idősorok eloszlásaira nézve annak érdekében, hogy a helytelen feltételezések kockázata kiküszöbölhető legyen. A backtesting célja, hogy a modell megbízhatóságát igazolja, valamint az, hogy a rendszeres ellenőrzés eredményeképpen feltárható legyen, ha bizonyos időszakokban – például egy válság során – a modell nem megfelelő előjelzéseket ad. Ezen túlmenően fontos kérdés, hogy milyen tulajdonságok kerülnek ellenőrzésre a backtesting-lépés során.

6.2 A backtesting és a pénzügyi szabályozás aktuális kérdései

A pénzügyi kockázatkezelés területén az elmúlt években a backtesting témaköre kimagasló figyelmet kapott. Ez annak köszönhető, hogy a Bázeli Bizottság a várható többletveszteség mutató használatát javasolta a banki tőke megfelelési mutató számításához a korábbi kockázatosított érték mutató helyett. A várható többletveszteség mutató használatával kapcsolatban a legkomolyabb kritika az, hogy a mutató nem rendelkezik az elicítabilitás tulajdonságával (Gneiting, 2011), így több kutató véleménye szerint nem létezik hozzá tartozó „természetes” vagy könnyen alkalmazható backtesting-eljárás (Bellini, et al., 2014) (Ziegel, 2016) és ezért a mutató használata nem ajánlott (Béli & Váradi, 2017). Ennek következménye, hogy a Bázeli Bizottság a tőke megfelelés ellenőrzésére továbbra is a VaR-mutató használatát javasolja, valamint a IAIS (International Association of Insurance Supervisors) többek között az elicítabilitás hiánya miatt javasolja a VaR-mutató használatát az ES helyett (IAIS, 2014). Ezzel ellentétes álláspontot képviselve több tanulmány mutat be olyan backtesting-eljárást, amely az elicítabilitás hiánya mellett is alkalmazható (Acerbi & Szekely, 2014) (Acerbi & Szekely, 2017) (Kratz, et al., 2018). A kockázati mérőszámok ellenőrzése azért komplex feladat, mert az összehasonlítás alapját képező realizált hozamok/veszteségek nem a teljes valós eloszlást mutatják meg, hanem csak egy konkrét véletlen „sorsolás” eredményét.

6.3 Elicitabilitás

Az elicítabilitás fogalmát Gneiting (2011) definálta. Az elicítabilitás egy olyan tulajdonság, amely a pontbecslések értékelésében játszik szerepet, így fontos a backtesting-eljárások jellemzésében.

Definíció 2.1: Egy \mathbf{y} statisztika \mathcal{F} -elicítálható, ha létezik olyan $S_{\mathbf{y}}(\boldsymbol{\psi}, X)$ értékelőfüggvény, ahol az \mathbf{y} -statisztika a következőképpen fejezhető ki:

$$\mathbf{y}(F) = \underset{\boldsymbol{\psi}}{\operatorname{argmin}} \mathbb{E}_F[S_{\mathbf{y}}(\boldsymbol{\psi}, X)], \quad \forall F \in \mathcal{F}, \quad (228)$$

A fenti képletben y a statisztika, X az F valós eloszlásfüggvényt követő valós valószínűségi változó, ψ pedig az y statisztikai adat becslése. A fenti definíció úgy értelmezhető, hogy egy statisztika akkor bír az elicítabilitás tulajdonságával, ha létezik hozzá olyan $S_y(\psi, X)$ hibafüggvény⁹, amelynek minimumhelye a tökéletes $\psi = y(F)$ előrejelzésnél található (Acerbi & Székely, 2017). Acerbi és Székely (2017) a következő gyakran használt értékelőfüggvényeket emeli ki példaként:

13. táblázat: Gyakran használt értékelőfüggvények

y	$S_y(x, y)$
μ (átlag)	$(x - y)^2$
$q_{1/2}$ (medián)	$ x - y $
q_α (α -kvantilis)	$ (x - y)/y $
e_α (α -expektilis)	$S(x, y) = (x - y)/x $

Forrás: saját szerkesztés, (Acerbi & Székely, 2017) alapján

6.4 A backtesting definíciója

Az elicítabilitás és egy adott kockázati mutatóhoz tartozó megfelelő backtesting-eljárás létezése közötti kapcsolatról szóló vita feloldását megnehezíti, hogy annak ellenére, hogy a pénzügyi kockázatkezelés területén évtizedek óta használnak backtesting-eljárásokat, nincs pontos definíció arra, hogy mit nevezhetünk backtesting-eljárásnak. Ilyen definíció hiányában viszont az sem egyértelműen megállapítható, hogy egy adott kockázati mutató ellenőrizhető-e. Acerbi és Székely (2017) a fenti probléma feloldása érdekében a backtesting definiálására tesz kísérletet.

⁹ *scoring function.*

Acerbi és Székely (2017) a backtesting intuitív definíciójaként a következő leírást javasolja:

„The term backtesting [...] refers to the practice of validating ex-ante model predictions of some statistics against observed ex-post realizations of the random variable.”

„A backtesting kifejezés egy statisztikai adat modelljéből származó ex-ante becsléseknek a valószínűségi változó ex-post megvalósulásai alapján történő ellenőrzésére vonatkozik.”¹⁰

6.5 A különböző backtesting-eljárások bemutatása

6.5.1 VaR-hibaszám-ellenőrzés – coverage test

A kockázatos érték mutató esetében adódik egy magától értetődő ellenőrzési pont, hiszen a VaR jelentése, hogy a VaR-összegnél nem lesz nagyobb veszteség az esetek konfidenciaszintnek megfelelő százalékában. Ez az állítás könnyen visszaellenőrizhető a becslések és a valós adatok ismeretében.

Jorion (2007, pp. 143-147) ennek megfelelően első lépésként meghatározza a hibák arányát a teljes minta elemszámához viszonyítva, ahol a hibának az minősül, ha az adott időszak – tipikusan egy nap – alatt a veszteség meghaladja a becsült VaR-értéket.

$$I_t = (L_t > VaR_t) \quad (229)$$

$$x = \sum_{t=1}^T I_t \quad (230)$$

$$\text{Hibaarány} = \frac{x}{T} \quad (231)$$

A fentiekben x a hibák száma, és T pedig a minta elemszáma. A következő lépés annak meghatározása, hogy a hibaarány mekkora eltérése után kerül a modell elvetésre. Erre a következő képlet alkalmas (Jorion, 2007, p. 144):

¹⁰ Fordítás a szerzőtől

$$z = \frac{x-pT}{\sqrt{p(1-p)T}} \sim N(0,1) \quad (232)$$

A fenti képletben p a VaR feletti veszteség valószínűsége, azaz 99% -os VaR mellett 1%. A bankok számára rendkívül költséges a második típusú hiba. Második típusú hiba akkor fordul elő, amikor olyan modell kerül elfogadásra, ami nem megfelelő. Ennek elkerülésére Christoffersen (2012, p. 303) 10%-os szignifikanciaszint használatát javasolja a modellek ellenőrzésére. Ezzel összhangban és a fenti képlet használatával könnyen megkaphatjuk a hibák elfogadható számát 250 napos időszakra.

14. táblázat: A hibák elfogadható száma 250 napos időszakra

VaR-szint	A hibák elfogadható száma
99%	< 5
97,5%	$3 < x < 9$
95%	$8 < x < 17$

Forrás: (saját szerkesztés)

6.5.2 VaR-függetlenségiteszt

A hibaszámok ellenőrzése fontos információt ad a modellről, de a hibaszámok ellenőrzése önmagában nem elegendő a modell értékelésére. Olyan eset is elképzelhető, amikor a megengedett hibák számát a modell nem lépi túl, de a hibák nagy része egy időszak alatt, vagy csoportosan realizálódik. Ilyenkor elképzelhető, hogy a modell bizonyos időszakokat nem tud megfelelően kezelni, ráadásul a mutatót használó szervezet stabilitásának szempontjából is rendkívül nagy kockázatot jelent a veszteségek ilyen jellegű sorozata. Megállapítható, hogy egy ideális modell hibái egymástól függetlenül jelentkeznek, ezzel növelve az alkalmazó pénzügyi szervezet stabilitását. A gyakorlatban többek között függetlenségi hibát mutathat például a historical simulation módszerrel kialakított modell, akkor, ha az idősor egy részidőszakra magasabb varianciát mutat, mint máskor (Christoffersen, 2012, p. 304).

Függetlenségi hibáról akkor beszélhetünk, hogyha egy adott időszakban a hiba előfordulásának valószínűsége nagyobb abban az esetben, ha az előző időszakban szintén hiba volt, mint abban az esetben, amikor az előző időszakban nem volt hiba.

Christoffersen (1998) (2012) elsőrendű Markov-folyamatot használ a függetlenségi teszt elvégzéséhez.

A Markov-folyamatot a következő átmenetmátrix definiálja:

$$\Pi_1 = \begin{bmatrix} 1 - \pi_{01} & \pi_{01} \\ 1 - \pi_{11} & \pi_{11} \end{bmatrix} \quad (233)$$

Az átmenetmátrix valószínűségei a következők:

$$\pi_{11} = P(I_{t+1} = 1 \mid I_t = 1) \quad (234)$$

$$\pi_{01} = P(I_{t+1} = 1 \mid I_t = 0) \quad (235)$$

π_{11} azt adja meg, hogy mekkora valószínűsége van annak, hogy egy VaR-hibát (violation) egy újabb hiba követ. π_{01} azt adja meg, hogy mekkora valószínűséggel követ VaR-hiba egy olyan napon, amikor nem volt VaR-hiba. Az elsőrendű Markov-folyamat használatának következménye, hogy azt feltételezzük, hogy a VaR-hiba előfordulási valószínűségét csak az előző időszak befolyásolja és az azt megelőző időszakok nem. Az elsőrendű Markov-folyamat likelihood-függvénye a következő:

$$L(\Pi_1) = (1 - \pi_{01})^{T_{00}} \pi_{01}^{T_{01}} (1 - \pi_{11})^{T_{10}} \pi_{11}^{T_{11}} \quad (236)$$

T_1 a hibák száma, T_0 a nem hibás periódusok száma. T_{ij} pedig azoknak a megfigyeléseknek a száma, ahol T_j megfigyelés T_i megfigyelést követ. T_{01} például azoknak az eseteknek a száma, ahol hiba nélküli periódust követően hibás periódus következett. L a likelihood-függvényt jelöli. A π_{01} és a π_{11} valószínűségeket a konkrét megfigyelések alapján is kiszámíthatjuk:

$$\widehat{\pi}_{01} = \frac{T_{01}}{T_{00} + T_{01}} \quad (237)$$

$$\widehat{\pi}_{11} = \frac{T_{11}}{T_{10} + T_{11}} \quad (238)$$

A valószínűségek behelyettesítésével megkapjuk a következő likelihood függvényt:

$$L(\widehat{\Pi}_1) = \left(1 - \frac{T_{01}}{T_{00} + T_{01}}\right)^{T_{00}} \left(\frac{T_{01}}{T_{00} + T_{01}}\right)^{T_{01}} \left(1 - \frac{T_{11}}{T_{10} + T_{11}}\right)^{T_{10}} \left(\frac{T_{11}}{T_{10} + T_{11}}\right)^{T_{11}} \quad (239)$$

A megfigyelt átmenetmátrix tehát a következő:

$$\hat{\Pi}_1 = \begin{bmatrix} 1 - \hat{\pi}_{01} & \hat{\pi}_{01} \\ 1 - \hat{\pi}_{11} & \hat{\pi}_{11} \end{bmatrix} \quad (240)$$

Ha azt feltételezzük, hogy a VaR-hibák függetlenek, azaz előfordulásuk független attól, hogy az előző időszak eredménye mi volt, akkor a következő Markov átmenetmátrixot kapjuk:

$$\hat{\Pi} = \begin{bmatrix} 1 - \hat{\pi} & \hat{\pi} \\ 1 - \hat{\pi} & \hat{\pi} \end{bmatrix} \quad (241)$$

A fentiek alapján a függetlenségi hipotézis ellenőrzése a következő próba segítségével végezhető el:

$$LR_{\text{ind}} = -2 \ln \left[\frac{L(\hat{\Pi})}{L(\hat{\Pi}_1)} \right] \sim \chi_1^2 \quad (242)$$

6.5.3 Várható többletveszteség – coverage test

Acerbi és Székely (2014) (2017) bemutat több különböző módszert arra, hogy a várható veszteség mutató teljesítménye hogyan ellenőrizhető a kockázatos értékhez hasonlóan. Acerbi és Székely (2014) szerint a várható veszteség esetében csak a hibaszám/hibaméret ellenőrzést (coverage test) szükséges elvégezni, a függetlenségi tesztet a piacon elterjedt gyakorlatnak megfelelően elegendő vizuális vizsgálattal ellenőrizni.

A következő szakaszban több ellenőrzési módszer kerül bemutatásra. Az első ellenőrzés akkor alkalmazható, ha a kockázatos érték ellenőrzése már megtörtént. A második módszer a kockázatos értéktől függetlenül alkalmazható, a harmadik pedig a második továbbfejlesztett változata. Mindegyik esetben az elvárás a modellel kapcsolatban a következő:

$$E \left(\frac{X_t}{ES_{\alpha,t}} + 1 \mid X_t + VaR_{\alpha,t} < 0 \right) = 0 \quad (243)$$

A fenti képletben X_t a t időszak hozamváltozója, $ES_{\alpha,t}$ az α konfidenciaszinthez tartozó, t időszakban várható veszteség. Mivel X_t változó esetében a veszteség negatív,

$ES_{\alpha,t}$ és $VaR_{\alpha,t}$ esetében pedig pozitív előjelű, így $\frac{X_t}{ES_{\alpha,t}} = -1$, ha X_t pontosan az $ES_{\alpha,t}$ -nek megfelelő értéket vesz fel. Ebből következik, hogy a lenti egyenlőtlenség teljesülése esetén a modell alábecsüli a kockázatot:

$$E \left(\frac{X_t}{ES_{\alpha,t}} + 1 \mid X_t + VaR_{\alpha,t} < 0 \right) < 0 \quad (243)$$

A további elemzéshez definiáljuk a következő függvényt:

$$I_t = (X_t + VaR_{\alpha,t} < 0) \quad (244)$$

Ahol I_t értéke egy lesz, ha a feltétel teljesül és nulla, ha nem.

6.5.3.1 VaR-alapú ellenőrzés

A modell ellenőrzéséhez használt próbastatisztika lényegében a várható veszteségtől való átlagos eltérés, definíciója a következő (Acerbi & Szekely, 2014):

$$Z_1(\vec{X}) = \frac{\sum_{t=1}^T \frac{X_t I_t}{ES_{\alpha,t}}}{N_T} + 1 \quad (245)$$

$N_T = \sum_{t=1}^T I_t$, $N_T > 0$ és $\vec{X} = \{X_t\}$ független, de nem azonos eloszlású véletlen változók halmaza. Minél kisebb a $Z_1(\vec{X})$ értéke, annál rosszabb becslést adott a várható veszteség modell, annál inkább alulbecsülte a kockázatot. A próba nullhipotézise az, hogy a becsült eloszlásszél megegyezik a valós eloszlásszéllel. Az alternatív hipotézis pedig az, hogy a valós várható veszteség nagyobb bizonyos időszakokban, mint a modell által megadott várható veszteség. Az első ellenőrzési módnál az alternatív hipotézis is azt feltételezi, hogy a becsült VaR-értékek megfelelnek a valós értékeknek. Ebből is adódik, hogy ez az ellenőrzési módszer valóban csak a VaR előzetes, külön ellenőrzése után használható.

6.5.3.2 Közvetlenül számított várható többletveszteség-ellenőrzés

A közvetlen ellenőrzés elvégzésére a következő próbastatisztika alkalmas (Acerbi & Szekely, 2014):

$$Z_2(\vec{X}) = \sum_{t=1}^T \frac{X_t I_t}{T \alpha ES_{\alpha,t}} + 1 \quad (246)$$

A próba null hipotézise továbbra is az, hogy a becült eloszlásszél megegyezik a valós eloszlásszéllel. Az alternatív hipotézis annyiban változott a korábbihoz képest, hogy megengedi, hogy a VaR-érték a becült modellnél kisebb legyen, mint a valós kockázatos érték. Ezen kívül ez a második teszt nem követeli meg, hogy az X_t hozamok függetlenek legyenek egymástól.

6.5.3.3 *A közvetlenül számított várhatóöbbitveszteség-ellenőrzés továbbfejlesztett változata*
Acerbi és Székely (2017) a Z_2 teszt következő továbbfejlesztett változatát javasolja:

$$Z_{ES_\alpha}(e, v, x) = \alpha(e - v) + (x + v)(x + v < 0) \quad (247)$$

A fentiekben $v = VaR_\alpha$ és $e = ES_\alpha$. A Z_{ES_α} -backtestfüggvény előnye, hogy sokkal kevésbé érzékeny a VaR becslésének pontosságára, mint a Z_2 mutató volt. A Z_{ES_α} -függvény alkalmas a modellek validációjára és összehasonlítására is. Az értekezésben a modellek összehasonlítása Z_{ES_α} -függvény segítségével történik.

6.6 Modellvalidáció: szignifikanciateszt

A próbastatisztika értékéből a szignifikanciateszthez szükséges p-érték számításához Acerbi és Székely (2014), valamint Christoffersen (2012) a következő Monte-Carlo-szimuláción alapuló módszert használják:

1. A korábbi t periódusra becült eloszlás P_t felhasználásával X_t^i értékeket generálnak, ahol $i = 1, \dots, M$ és M egy tetszőleges nagy szám.
2. Kiszámítják a Z^i -értéket, ahol $Z^i = Z(\bar{X}^i)$.
3. Kiszámítják a p-értéket, ahol $p = \sum_{i=1}^M (Z^i < Z(\vec{x})) / M$ és $Z(\vec{x})$ a p-értékhez tartozó konkrét megvalósult Z -érték.

Az előre meghatározott ϕ szignifikancia szint mellett a modell akkor kerül elfogadásra, ha p kisebb, mint ϕ . A fenti módszer előnye, hogy kisebb minta mellett is alkalmazható, illetve független a Z -értékek eloszlásától, így minden ellenőrzési módszerre egységesen alkalmazható.

6.7 A backtesting-eljárás ereje

Egy backtesting-eljárás ereje alatt azt értjük, hogy az adott eljárás mennyire képes a hibás becsléseket produkáló modellek megkülönböztetésére a helyes, azaz adott szignifikancia szint mellett elfogadható modellektől. Acerbi és Székely (2014) egy szimuláción alapuló eljárást javasol a backtesting-módszer erejének mérésére. Minden $t = 1, \dots, T$ banki napon X_t a bank hozamát reprezentáló valószínűségi változó, amely az F_t valós eloszlást követi, és a P_t becslési eloszlás alapján becsülhető. A becslési modell tesztelésekor a H_0 –nullhipotézis az, hogy a becslési modell megfelelő.

$$H_0: P_t^{[\alpha]} = F_t^{[\alpha]}, \forall t \quad (248)$$

$$H_1: ES_{\alpha,t}^F \geq ES_{\alpha,t}^P, \forall t,$$

$$\text{és } \exists t: ES_{\alpha,t}^F > ES_{\alpha,t}^P \quad (249)$$

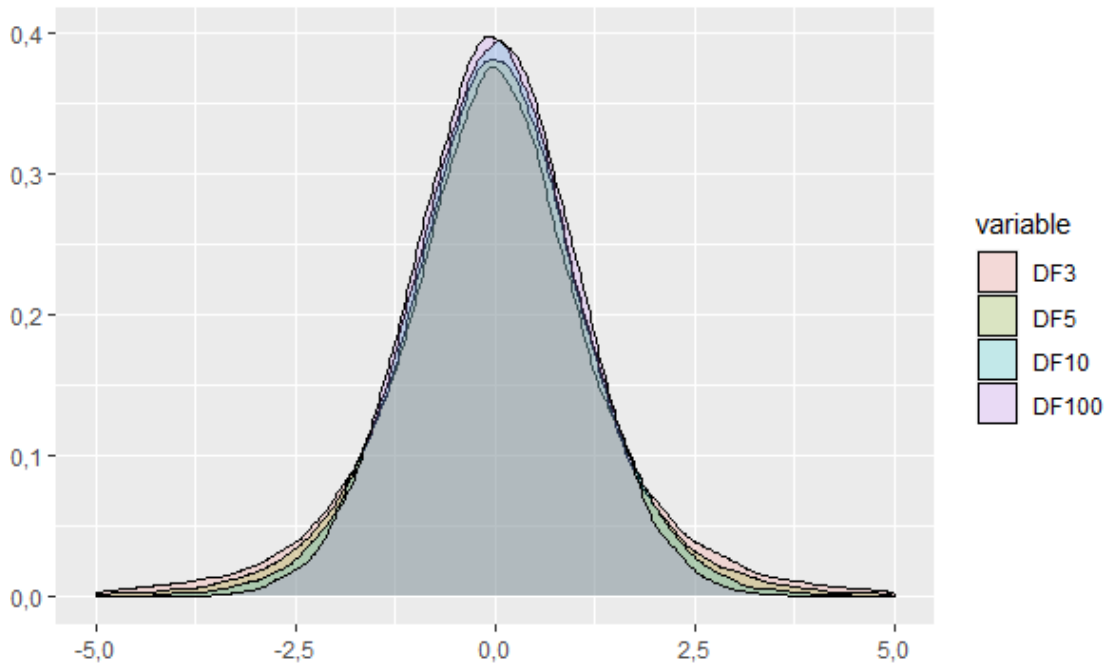
A backtesting eljárás erejének mérését valós idősor használata helyett egy valószínűségi F és egy becsült P hozameloszlás meghatározása mellett történik, így kiküszöbölhető az a probléma, hogy a valós eloszlás nem ismert. A számítás lépései a következők:

1. A P eloszlás felhasználásával X_P^i értékeket generálunk, ahol $i = 1, \dots, M$ és M egy tetszőleges nagy szám.
2. Kiszámítjuk a Z_P^i értékeit: $Z_P^i = Z(\vec{X}_P^i)$
3. Az F eloszlás felhasználásával X_F^i értékeket generálunk, ahol $i = 1, \dots, M$ és M egy tetszőleges nagy szám.
4. Kiszámítjuk a Z_F^i értékeit: $Z_F^i = Z(\vec{X}_F^i)$
5. Az M elemből álló Z_P^i vektor α szignifikancia szinthez tartozó kvantilisének ($Z_{P,\alpha}$) kiszámítása: $\alpha = P(Z_P^i < Z_{P,\alpha})$
6. $power = \sum_{i=1}^M (Z_{F,i} < Z_{P,\alpha}) / M$

Az eljárás bemutatásához Acerbi és Székely (2014) megközelítéséhez hasonlóan különböző szabadságfokú Student-féle t-eloszlásokat használok, azzal a

különbséggel, hogy a lenti példában már a későbbi tanulmányukban megjelent Z_{ES_α} -backtesting-függvény (Acerbi & Szekely, 2017) erejét számítom ki.

51. ábra: A példában használt t-eloszlások sűrűségfüggvényei



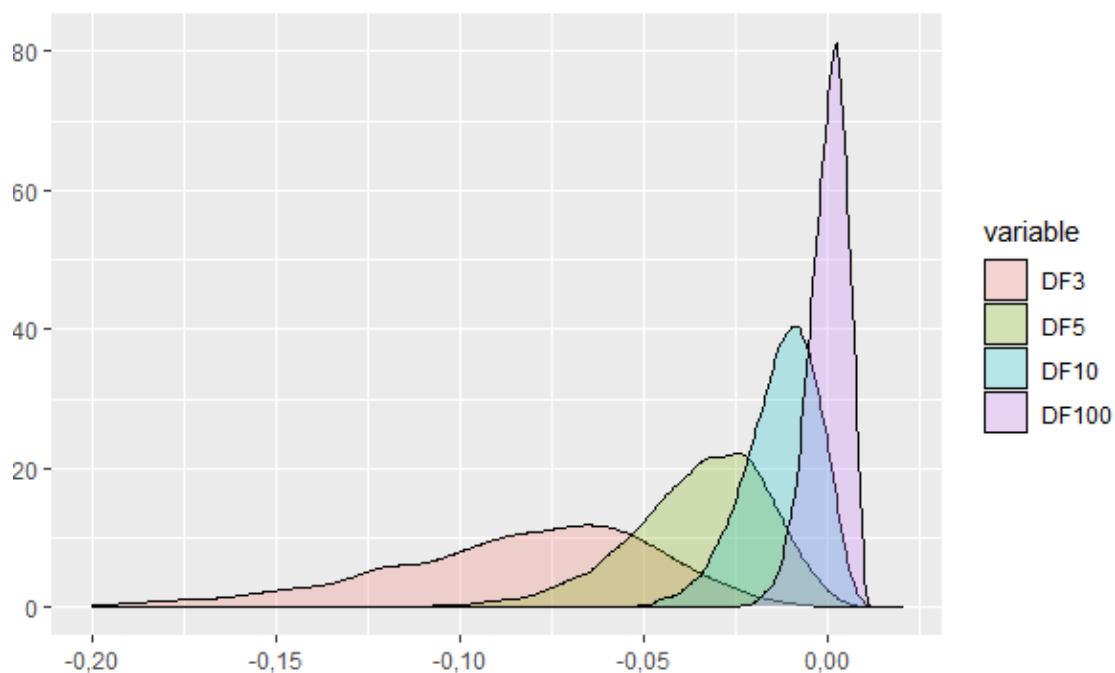
Forrás: (saját szimuláció)

15. táblázat: A backtesting-eljárás ereje

F_t	P_t	A Z-mutató ereje
t-szabadságfok: 3	t-szabadságfok: 100	98,81%
t-szabadságfok: 5	t-szabadságfok: 100	93,74%
t-szabadságfok: 10	t-szabadságfok: 100	56,3%
t-szabadságfok: 100	t-szabadságfok: 100	5%

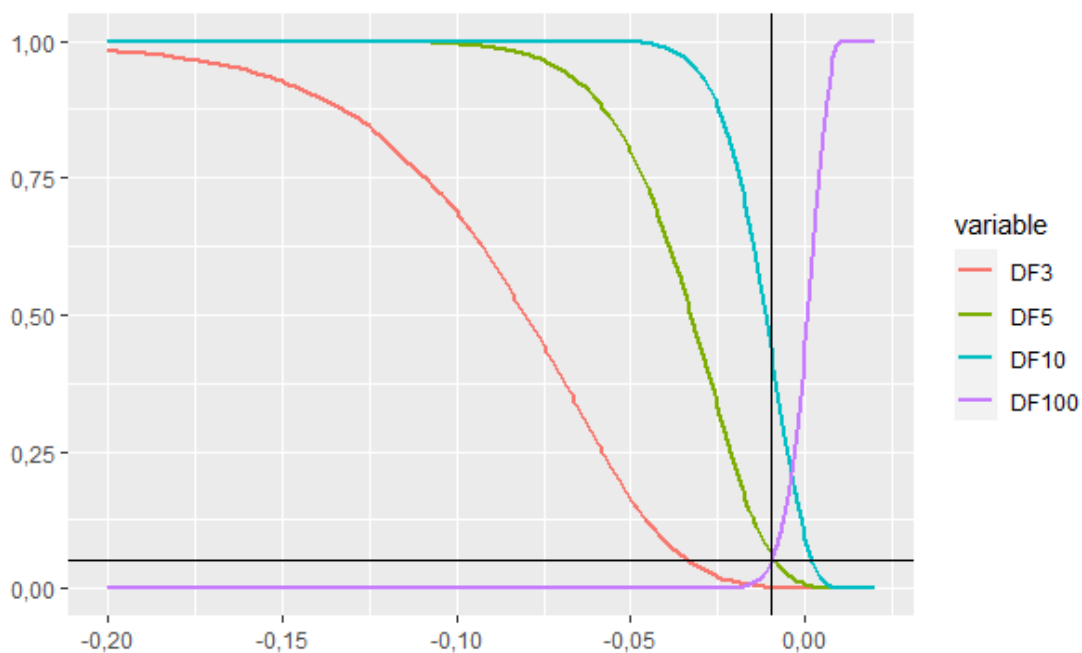
Forrás: (saját számítások)

52. ábra: A szimulált Z-értékek sűrűségfüggvényei



Forrás: (saját számítások)

53. ábra: A szimulált Z-értékek kumulatív eloszlásfüggvényei

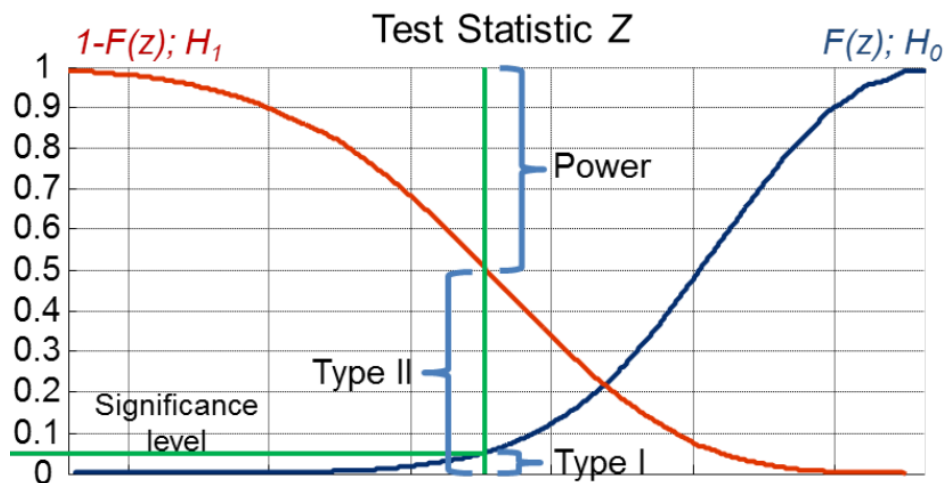


Forrás: (saját számítások)

Az 53. ábra a 3, 5 és 10-es szabadságfokhoz tartozó túlélési függvényeket, azaz $1-F(z)$ értékeit mutatja. A 100-as szabadságfokhoz tartozó függvény egyszerűen a Z-

értékhez tartozó kumulatív eloszlásfüggvény. A vízszintes egyenes az 5%-os szignifikancia szintet jelöli, a függőleges egyenes pedig azt a pontot, ahol az 5%-os egyenes metszi a 100-as szabadság fokú t eloszláshoz tartozó Z-értékek kumulált eloszlásfüggvényét.

54. ábra: A backtesting eljárás ereje a kumulatív eloszlásfüggvények alapján



Forrás: (Acerbi & Szekely, 2014, p. 21)

6.8 Az ellenőrzések kiegészítése és kiterjesztése

6.8.1 Komplex termékek, portfóliók ellenőrzése

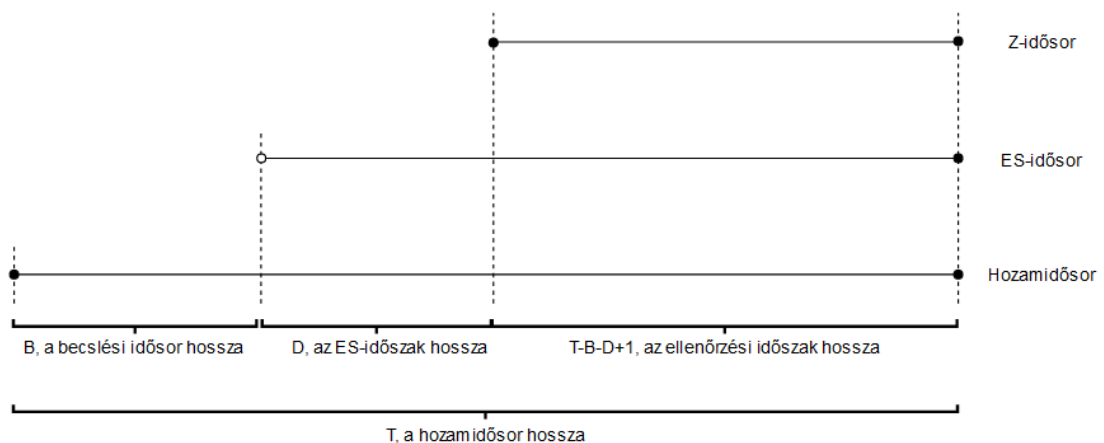
Komplex termékek vagy portfóliók ellenőrzéséhez a múltbeli alapadatokból – például részvényadatokból – elkészítésre kerül az összetett terméket, vagy portfóliót leíró idősor. Egy portfólió esetében ez egyszerűen abból áll, hogy az egyes alkotóelemek periodikus adatai Monte-Carlo-szimuláció során generálásra kerülnek, majd a részvények adatai a portfóliónak megfelelően súlyozásra kerülnek. Ezt követően ebből az idősorból már az egyszerű eszközökre alkalmazott módszerek segítségével elvégezhető az ellenőrzés. Amennyiben az elemzés szempontjából a portfóliókialakítási stratégia teljesítménye is érdekes, úgy a fenti módszer a portfólióoptimalizálási lépés periódusonként történő ismétlésével is elvégezhető Bugár és Uzsoki (2005).

6.9 Ex-ante becslés, ex-post ellenőrzés

A modellek működése a folyamatos ellenőrzések mellett könnyen visszaellenőrizhető a múltbeli adatok segítségével. Ehhez minél hosszabb idősor szükséges, lehetőleg olyan, amely különösen veszteséges, válságos időszakokat is magában foglal. Az ellenőrzések elvégzéséhez az idősor első B darab eleme a modell paramétereinek becsléséhez szükséges, majd a modell eredményei összehasonlításra kerülnek a becslési időszakot követő valós eredményekkel. A lehető legtöbb ellenőrzési pont úgy kapható, ha a becslési időszakot követően egy-egy periódussal eltolva megismétlődik az ellenőrzés. T-hosszú minta, B-hosszú becslési és D-hosszú ES-időszak esetén $T-B-D+1$ darab ellenőrzési pontot kapunk, azaz ez lesz a Z-idősor hossza. Az időszakok hosszának kapcsolatát az 55. ábra mutatja be. Egy ES-becslés kiszámítása mindig egy B-hosszú idősor alapján történik, és egy Z-érték kiszámítása pedig egy D-hosszú ES-becslésekből álló idősor alapján történik. A B és a D paraméterek értéke nem változik az ablakok eltolása közben, tehát a becslési időszak és az ES-időszak mindig ugyanolyan hosszú.

$$T \geq B + D \quad (250)$$

55. ábra: A becslési, az ES és az ellenőrzési időszak hosszának kapcsolata



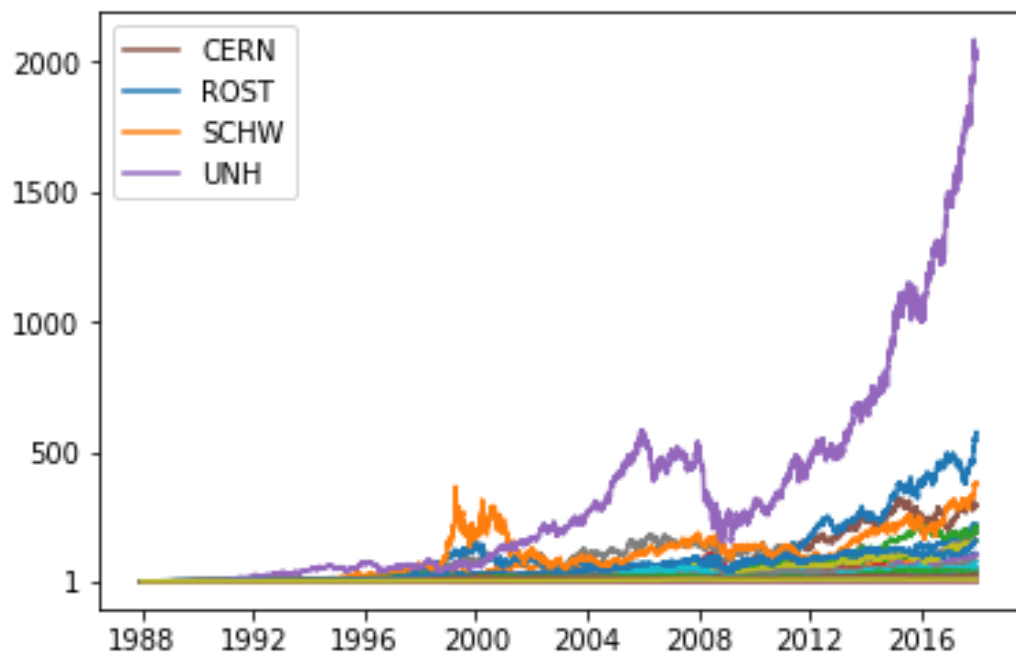
Ha a számítási kapacitás szűkössége, vagy más ok miatt a becslési ablakok egyperiódusos eltolása nem megfelelő, úgy ahelyett K periódussal kell eltolni a becslési ablakot, így szükség esetén az is elérhető, hogy a becslési időszakok ne legyenek átfedésben egymással.

7 Empirikus elemzés

7.1 Adatok

Az empirikus elemzéshez a tanulmány a Standard & Poor's S&P 500 indexben résztvevő részvények közül 100, legalább 30 évre visszamenően elérhető adatokkal rendelkező értékpapír napi korrigált záróárfolyamait használja. Az adatok forrása a finance.yahoo.com adatbázisa. A felhasznált idősorok minden részvény esetében a 1988.01.01-től 2017.12.31-ig terjedő, 30 éves időszakot fedik le.

56. ábra: Részvényárak alakulása 1988–2017



Forrás: (saját szerkesztés, finance.yahoo.com alapján)

16. táblázat: Az elemzésben résztvevő részvények (A-K)

AAPL	Apple Inc.	CERN	Cerner
ABT	Abbott Laboratories	CHD	Church & Dwight
ADBE	Adobe Inc.	CI	CIGNA Corp.
ADSK	Autodesk Inc.	CNP	CenterPoint Energy
AET	Aetna Inc. (CVS felvásárolta 2018-ban)	COO	The Cooper Companies
AFL	Aflac Inc.	CVX	Chevron Corp.
AJG	Arthur J. Gallagher & Co.	DIS	The Walt Disney Company
ALK	Alaska Air Group Inc	DOV	Dover Corporation
AMD	Advanced Micro Devices Inc	DVN	Devon Energy
AME	AMETEK Inc.	EFX	Equifax Inc.
APD	Air Products & Chemicals Inc	EMR	Emerson Electric Company
AVY	Avery Dennison Corp	EQT	EQT Corporation
AXP	American Express Co	EXC	Exelon Corp.
BA	Boeing Company	FDX	FedEx Corporation
BAC	Bank of America Corp	FOX	Fox Corporation
BAX	Baxter International Inc.	FRT	Federal Realty Investment Trust
BBT	BB&T Corp. (Nevet változtatott 2019-ben: Truist Financial Corp/ TFC)	GE	General Electric Company
BBY	Best Buy Co. Inc.	GIS	General Mills, Inc. (GIS)
BDX	Becton Dickinson	HAL	Halliburton Company
BLL	Ball Corp	HRB	H&R Block, Inc.
CA	CA Inc. (Broadcom Inc. felvásárolta 2018- ban)	IBM	International Business Machines Corporation
CAH	Cardinal Health Inc.	IP	International Paper Company
CAT	Caterpillar Inc.	JNJ	Johnson & Johnson
CB	Chubb Corp	JPM	JPMorgan Chase & Co.
CCL	Carnival Corp.	K	Kellogg Company

Forrás: (saját szerkesztés)

17. táblázat: Az elemzésben résztvevő részvények (Ko-Z)

KO	The Coca-Cola Company	STI	SunTrust Banks (BB&T felvásárolta 2019-ben)
L	Loews Corporation	STT	State Street Corporation
LMT	Lockheed Martin Corp.	T	AT&T Inc.
LNT	Alliant Energy Corp	TGT	Target Corporation
LUV	Southwest Airlines	TIF	Tiffany & Co.
MAS	Masco Corp.	TROW	T. Rowe Price Group
MCD	McDonald's Corp.	TRV	The Travelers Companies
MMM	3M Company	TXN	Texas Instruments Incorporated
MO	Altria Group Inc	TXT	Textron Inc.
MRK	Merck & Co.	UNH	UnitedHealth Group Incorporated
MSFT	Microsoft Corporation	USB	U.S. Bancorp
NEM	Newmont Corporation	UTX	United Technologies (2020-ban nevet változtatott: Raytheon Technologies /RTX)
NKE	NIKE, Inc.	VAR	Varian Medical Systems, Inc.
OKE	ONEOK, Inc.	VFC	V.F. Corporation
PEG	Public Service Enterprise Group Incorporated	VLO	Valero Energy Corporation
PFE	Pfizer Inc.	VNO	Vornado Realty Trust
PG	The Procter & Gamble Company	VZ	Verizon Communications Inc.
PSA	Public Storage	WBA	Walgreens Boots Alliance, Inc.
RJF	Raymond James Financial, Inc.	WEC	WEC Energy Group
ROK	Rockwell Automation, Inc.	WMB	The Williams Companies, Inc.
ROST	Ross Stores, Inc.	WMT	Walmart Inc.
SCHW	The Charles Schwab Corporation	WY	Weyerhaeuser Company
SEE	Sealed Air Corporation	XEL	Xcel Energy Inc.
SO	The Southern Company	XOM	Exxon Mobil Corporation
SPGI	S&P Global Inc.	ZION	Zions Bancorp

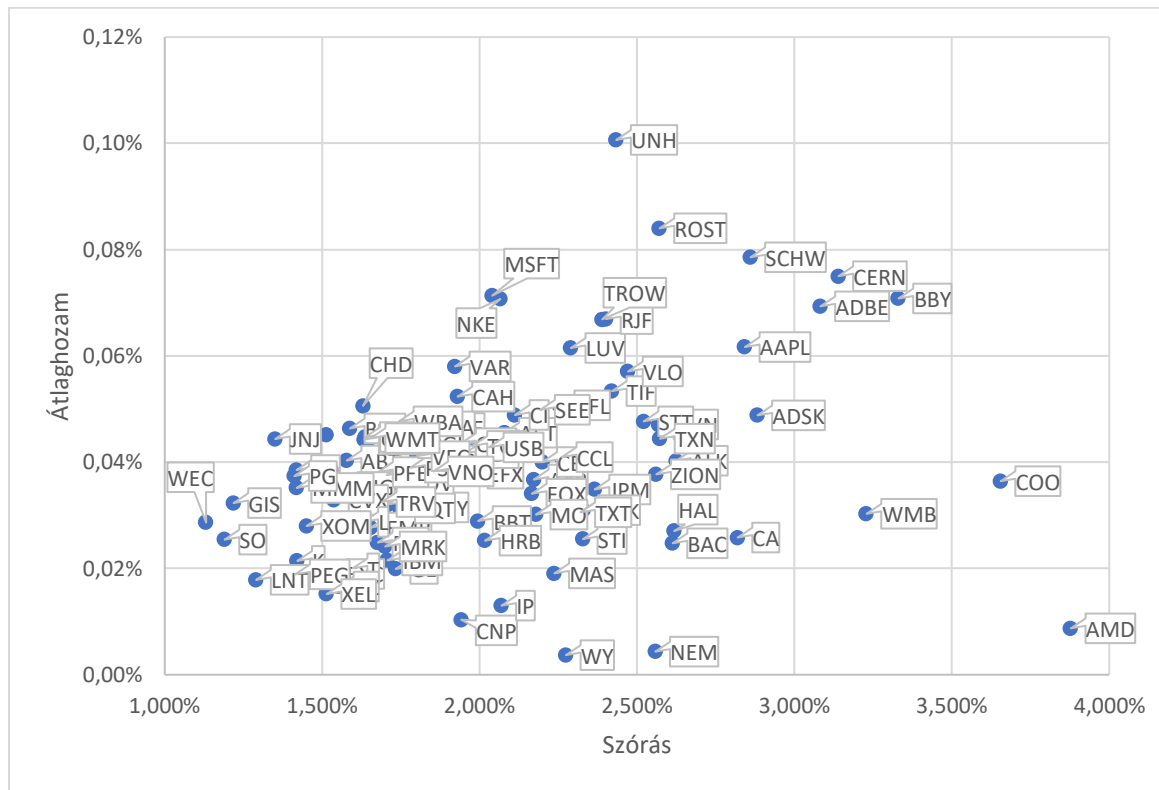
Forrás: (saját szerkesztés)

18. táblázat: A részvények szórása és átlagos hozama

Részvény	Szórás	Átlag-hozam	Részvény	Szórás	Átlag-hozam	Részvény	Szórás	Átlag-hozam
AAPL	2,843%	0,062%	DVN	2,570%	0,047%	PG	1,411%	0,037%
ABT	1,578%	0,040%	EFX	1,979%	0,038%	PSA	1,780%	0,039%
ADBE	3,083%	0,069%	EMR	1,655%	0,028%	RJF	2,402%	0,067%
ADSK	2,883%	0,049%	EQT	1,748%	0,031%	ROK	2,330%	0,031%
AET	2,079%	0,046%	EXC	1,531%	0,019%	ROST	2,572%	0,084%
AFL	2,255%	0,051%	FDX	1,954%	0,042%	SCHW	2,860%	0,079%
AJG	1,566%	0,035%	FOX	2,166%	0,034%	SEE	2,193%	0,050%
ALK	2,624%	0,040%	FRT	1,677%	0,025%	SO	1,190%	0,025%
AMD	3,878%	0,009%	GE	1,734%	0,020%	SPGI	1,758%	0,044%
AME	1,821%	0,046%	GIS	1,219%	0,032%	STI	2,328%	0,026%
APD	1,726%	0,038%	HAL	2,618%	0,027%	STT	2,521%	0,048%
AVY	1,792%	0,031%	HRB	2,017%	0,025%	T	1,597%	0,020%
AXP	2,172%	0,037%	IBM	1,708%	0,022%	TGT	1,978%	0,043%
BA	1,852%	0,047%	IP	2,070%	0,013%	TIF	2,420%	0,053%
BAC	2,613%	0,025%	JNJ	1,351%	0,044%	TROW	2,389%	0,067%
BAX	1,708%	0,032%	JPM	2,365%	0,035%	TRV	1,690%	0,032%
BBT	1,995%	0,029%	K	1,421%	0,021%	TXN	2,573%	0,044%
BBY	3,331%	0,071%	KO	1,419%	0,039%	TXT	2,320%	0,031%
BDX	1,588%	0,046%	L	1,631%	0,029%	UNH	2,434%	0,101%
BLL	1,704%	0,038%	LMT	1,634%	0,044%	USB	2,029%	0,043%
CA	2,820%	0,026%	LNT	1,290%	0,018%	UTX	1,637%	0,045%
CAH	1,930%	0,052%	LUV	2,290%	0,061%	VAR	1,922%	0,058%
CAT	1,966%	0,040%	MAS	2,238%	0,019%	VFC	1,801%	0,041%
CB	2,199%	0,040%	MCD	1,513%	0,045%	VLO	2,472%	0,057%
CCL	2,261%	0,041%	MMM	1,419%	0,035%	VNO	1,853%	0,038%
CERN	3,140%	0,075%	MO	2,179%	0,030%	VZ	1,553%	0,017%
CHD	1,630%	0,051%	MRK	1,700%	0,024%	WBA	1,744%	0,048%
CI	2,112%	0,049%	MSFT	2,040%	0,071%	WEC	1,131%	0,029%
CNP	1,942%	0,010%	NEM	2,558%	0,004%	WMB	3,229%	0,030%
COO	3,655%	0,036%	NKE	2,066%	0,071%	WMT	1,654%	0,045%
CVX	1,538%	0,033%	OKE	1,942%	0,044%	WY	2,274%	0,004%
DIS	1,838%	0,040%	PEG	1,414%	0,019%	XEL	1,514%	0,015%
DOV	1,728%	0,036%	PFE	1,678%	0,038%	XOM	1,451%	0,028%
						ZION	2,561%	0,038%

Forrás: (saját számítás)

57. ábra: A részvények teljesítménye átlaghozam és szórás tekintetében



Forrás: (saját számítás)

7.2 Az adatok statisztikai tulajdonságai

Az ES-mutató használatát többek között az is indokolja, hogy a gyakorlatban a veszteségek nem normális eloszlást követnek, valamint a veszteségeloszlások vastagszélűek (Rockafellar & Uryasev, 2002). Annak ellenére, hogy az ES-mutató alapvetően jobban tudja kezelni a gyakorlatban előforduló eloszlásokat, fontos megvizsgálni az idősorok eloszlásait, mert az adatok statisztikai tulajdonságai hatással lehetnek a különböző modellek teljesítményére.

7.2.1 A normalitás ellenőrzése a Shapiro–Wilk- és a Kolmogorov–Smirnov-tesztekkel

A loghozamok normalitását két külön teszttel ellenőriztem: a Shapiro–Wilk (1965) és a Kolmogorov–Smirnov-próbával (Pintér & Rappai, 2007). 5%-os szignifikanciaszint mellett egyik teszt sem adott pozitív eredményt egyik részvényre sem, azaz a tesztek alapján a részvények loghozamai nem követnek normális eloszlást. Az ellenőrzött minta jelen esetben nagy elemszámúnak minősül. A fenti tesztek jellemzője,

hogy nagy elemszám mellett az eloszlások legkisebb eltérése is azt az eredményt adja, hogy az empirikus eloszlás nem követi a normális eloszlást. Ennek feloldása végett javasolt az eloszlások vizuális vizsgálata is, melyre egyrészt a sűrűségfüggvények összehasonlítása szolgál, másrészt pedig az úgynevezett QQ-ábra vizsgálata.

19. táblázat: Az idősorok normalitástesztje

Részvény	SW p	KS p	Részvény	SW p	KS p	Részvény	SW p	KS p
AAPL	0	2.09E-37	EFX	0	1.47E-40	RJF	0	1.09E-44
ABT	3.36E-44	1.49E-17	EMR	9.81E-45	2.59E-21	ROK	0	3.03E-56
ADBE	0	1.17E-40	EQT	0	1.04E-27	ROST	0	1.75E-39
ADSK	0	1.27E-26	EXC	0	6.88E-25	SCHW	4.20E-45	2.37E-31
AET	0	1.67E-42	FDX	1.26E-44	1.16E-27	SEE	0	7.34E-61
AFL	0	6.18E-70	FOX	0	1.48E-35	SO	2.46E-42	6.23E-24
AJG	0	3.85E-50	FRT	0	1.31E-59	SPGI	0	1.35E-41
ALK	0	5.68E-27	GE	0	6.91E-39	STI	0	7.60E-78
AMD	0	2.94E-34	GIS	3.67E-42	6.72E-20	STT	0	2.57E-80
AME	0	1.07E-31	HAL	0	4.76E-24	T	0	2.18E-28
APD	7.01E-45	8.95E-25	HRB	0	2.14E-31	TGT	0	1.65E-28
AVY	0	4.17E-32	IBM	0	3.31E-34	TIF	0	4.17E-36
AXP	0	4.24E-41	IP	0	8.65E-32	TROW	0	6.75E-48
BA	0	6.50E-23	JNJ	7.01E-45	7.33E-27	TRV	0	9.11E-49
BAC	0	1.51E-84	JPM	0	4.20E-53	TXN	1.56E-42	3.90E-27
BAX	0	8.10E-25	K	0	1.32E-30	TXT	0	6.43E-66
BBT	0	4.78E-50	KO	0	3.49E-25	UNH	0	2.65E-42
BBY	0	1.48E-48	L	0	3.13E-53	USB	0	6.91E-62
BDX	0	1.79E-34	LMT	0	8.26E-34	UTX	0	2.20E-28
BLL	0	2.14E-31	LNT	0	1.15E-37	VAR	0	2.62E-32
CA	0	6.77E-60	LUV	3.68E-42	5.31E-17	VFC	0	1.03E-34
CAH	0	4.48E-39	MAS	0	1.14E-34	VLO	0	6.17E-24
CAT	1.51E-43	1.97E-22	MCD	7.85E-44	2.58E-21	VNO	0	5.29E-87
CB	0	1.03E-78	MMM	0	2.34E-29	VZ	1.26E-44	4.52E-24
CCL	0	5.66E-38	MO	0	1.32E-120	WBA	2.94E-44	7.47E-20
CERN	0	1.95E-53	MRK	0	6.72E-26	WEC	7.24E-40	1.03E-23
CHD	0	4.86E-34	MSFT	0	1.12E-26	WMB	0	1.52E-117
CI	0	2.89E-55	NEM	2.31E-43	1.13E-18	WMT	2.80E-44	9.82E-29
CNP	0	5.63E-89	NKE	0	1.50E-38	WY	0	7.63E-43
COO	0	2.02E-127	OKE	0	1.01E-38	XEL	0	3.75E-75
CVX	0	1.92E-17	PEG	0	2.40E-31	XOM	0	3.44E-19
DIS	0	8.86E-27	PFE	1.37E-40	7.19E-21	ZION	0	1.49E-72
DOV	5.32E-44	1.78E-22	PG	0	5.04E-32			
DVN	0	4.07E-30	PSA	0	6.12E-45			

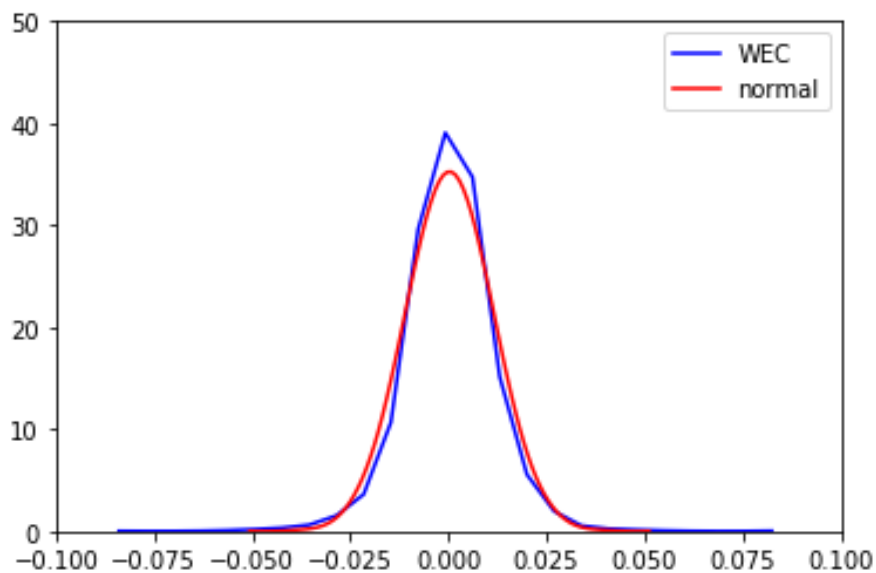
Forrás: (saját számítások)

A 19. táblázat tartalmazza a Shapiro–Wilk- és a Kolmogorov–Smirnov-tesztek p-értékeit, ahol az 5% feletti értékek jelentenék azt, hogy az adott részvény hozamai normális eloszlást követnek. A vizuális ellenőrzésre a Shapiro–Wilk-teszt szerint kapott legmagasabb p-értékkel rendelkező részvényt (WEC) választottam ki, a legmagasabb KS szerinti p-értékű (LUV), valamint a mindkét módszer szerinti legalacsonyabb p-értékű (COO) részvényt.

7.2.2 A normalitás ellenőrzése vizuális módszerekkel

A sűrűségfüggvényekhez az adott részvény idősorait, valamint 100 000, az adott részvény hozamának átlagával és szórásával parametrizált normális eloszlásból szimulált hozamadatot használtam.

58. ábra: A WEC-idősor normalitásának vizuális ellenőrzése (sűrűségfüggvény)



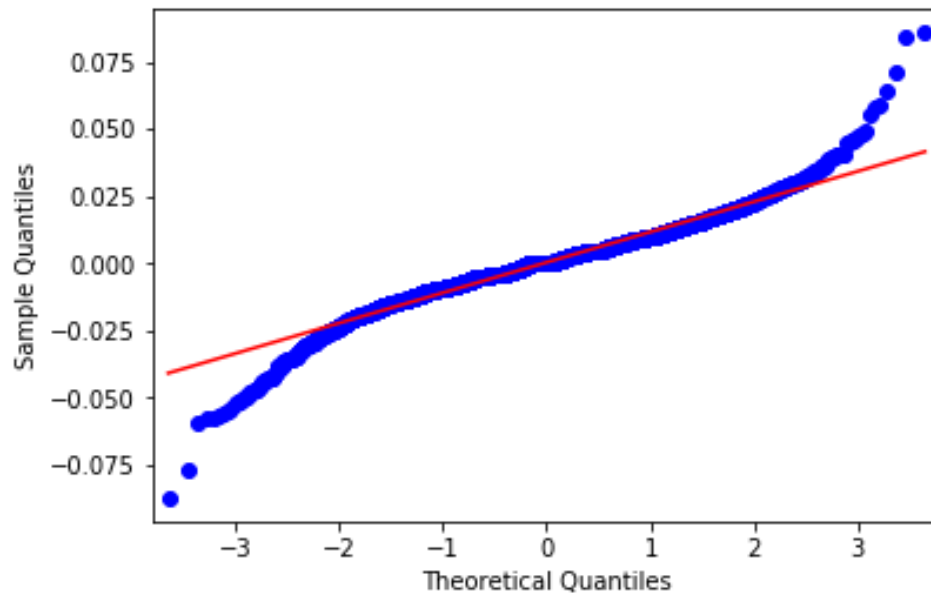
Forrás: (saját szerkesztés)

Az 58. ábra a WEC részvény sűrűségfüggvényét mutatja, amely a Shapiro-Wilk teszt alapján a vizsgált 100 részvény közül a legközelebb áll ahhoz, hogy normális eloszlásnak tekinthető legyen. Az ábra alapján a WEC részvény hozamának empirikus eloszlása relatív jól követi a normális eloszlást, de látható, hogy az empirikus eloszlás csúcsosabb.

Az 59. ábra a WEC részvény QQ-ábráját mutatja, amely a részvény empirikus eloszlásának kvantiliseit (y-tengely) veti össze a normális eloszlás kvantiliseivel (x-

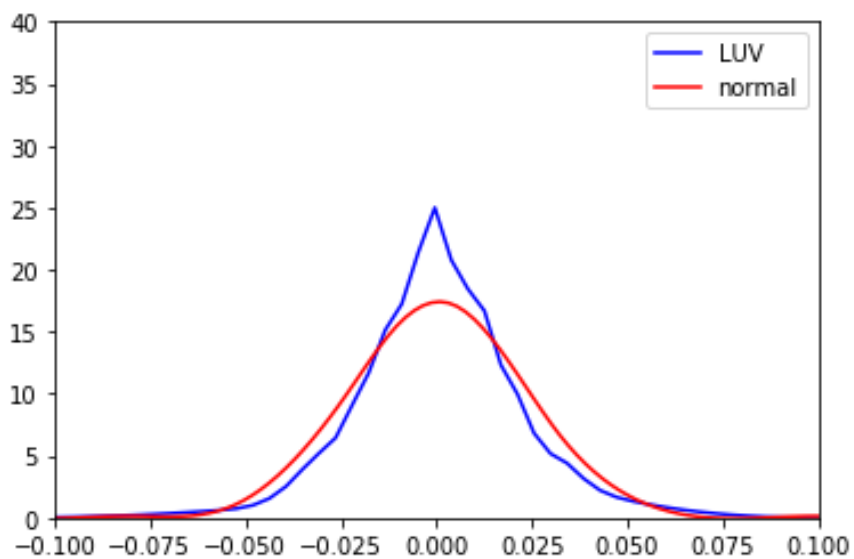
tengely). A piros egyenes azt adja meg, ha a vizsgált részvény normális eloszlást követne, akkor a kék adatpontok hol helyezkednének el. Ez alapján megállapítható, hogy a legnagyobb eltérések a WEC részvény empirikus eloszlásának széleinél figyelhető meg.

59. ábra: A WEC-idősor normalitásának vizuális ellenőrzése (kvantilisek)



Forrás: (saját számítás)

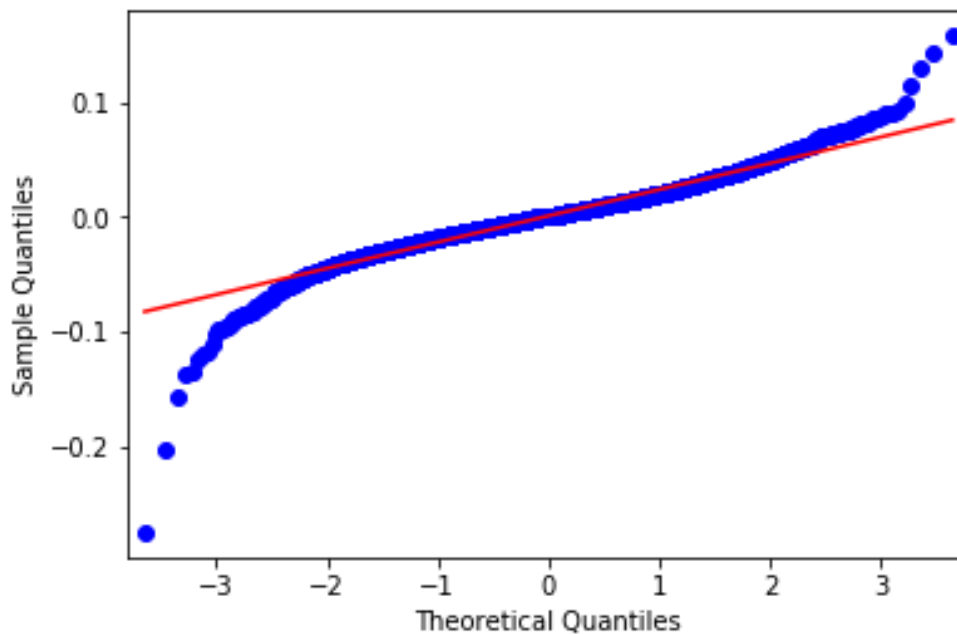
60. ábra: A LUV-idősor normalitásának vizuális ellenőrzése (sűrűségfüggvény)



Forrás: (saját szerkesztés)

A LUV részvény a Komolgorov–Smirnov-teszt alapján áll legközelebb ahhoz, hogy normális eloszlásúnak tekinthető legyen. Az 60. ábra a LUV sűrűségfüggvényét mutatja be. Az ábrán lényeges eltérés látható a normális eloszlás és a LUV empirikus eloszlásának sűrűségfüggvénye között. A korábban vizsgált WEC részvényhez hasonlóan a LUV esetében is nagyobb csúcsosság, illetve a QQ ábrán (61. ábra) az eloszlás széleinek eltérése figyelhető meg.

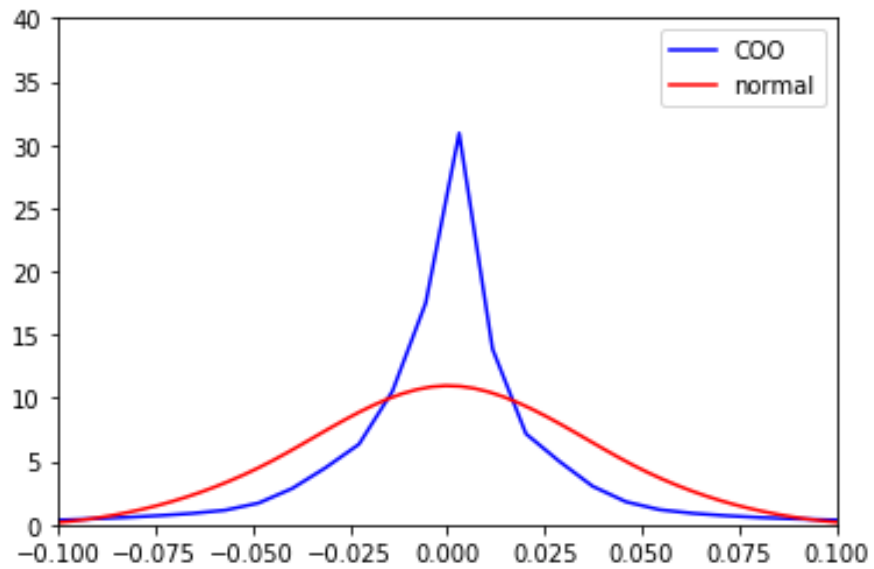
61. ábra: A LUV idősor normalitásának vizuális ellenőrzése (kvantilisek)



Forrás: (saját számítás)

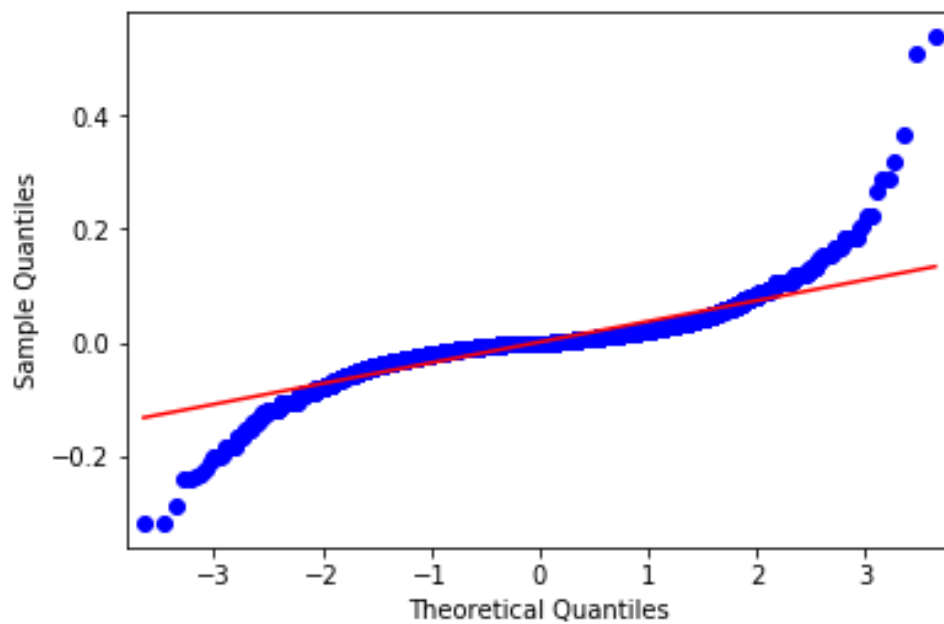
A COO részvény mind a Komolgorov–Smirnov-, mind a Shapiro–Wilk-teszt alapján a legalacsonyabb p-értéket kapta, így a tesztek alapján a COO empirikus eloszlása áll leginkább távol a normális eloszlástól. Ezt a vizuális tesztek is megerősítik, lényegesen nagyobb eltérést mutatva, mint a korábbi két részvény esetében. Az eltérések jellege viszont megegyezik a korábbi két vizsgált részvényével: a sűrűségfüggvényből a csúcsosság, a QQ-ábra alapján pedig a szélek eltérése állapítható meg. A sűrűségfüggvényekről egy további jellegzetesség is leolvasható: az empirikus eloszlások szimmetrikusnak tűnnek.

62. ábra A COO-idősor normalitásának vizuális ellenőrzése (sűrűségfüggvény)



Forrás: (saját számítás)

63. ábra: A COO-idősor normalitásának vizuális ellenőrzése (kvantilisek)



Forrás: (saját számítás)

7.2.3 A ferdeség, a csúcosság és a stacionaritás ellenőrzése

A ferdeség az eloszlás aszimmetriájának vizsgálatára alkalmas mérőszám, amely az aszimmetriát a normális eloszláshoz méri (Pintér & Rappai, 2007):

$$S = \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\sigma} \right)^3 \quad (251)$$

A fenti képletben \bar{x} az eloszlás számtani átlaga, x_i az idősor i . eleme, σ pedig a szórás. Jobb oldali aszimmetria esetén S értéke pozitív, tökéletesen szimmetrikus eloszlás esetén 0, míg bal oldali aszimmetria esetén negatív. A 20. táblázatból jól látható, hogy a részvény idősorok eloszlásai közül 72 esetben bal oldali aszimmetriát figyelhetünk meg, egyetlen esetben az eloszlás közel szimmetrikus és 27 esetben pedig jobb oldali aszimmetria jellemző rá.

Az eloszlás alakvizsgálatának másik fontos mérőszáma a csúcosság, amely alatt az eloszlást jellemző görbe meredekségét értjük. A csúcosság mérőszámai is a normális eloszláshoz viszonyítva adják meg a vizsgált eloszlás relatív lapultságát. A csúcsos eloszlást az átlag körüli tömörülés jellemzi (Pintér & Rappai, 2007):

$$K = \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{\sigma} \right)^4 \quad (252)$$

A fenti képletben \bar{x} az eloszlás számtani átlaga, x_i az idősor i . eleme, σ pedig a szórás. Normális eloszlás esetén a K -mutató értéke 3. Ennél kisebb K -érték lapult, háromnál nagyobb K -érték pedig csúcsos eloszlást jelent. A részvények idősoraihoz tartozó csúcosságértékeket a 21. táblázat tartalmazza. A táblázatból kiderült, hogy az idősorok eloszlásaira minden esetben jellemző a csúcosság. A legkevésbé csúcsos eloszlással bíró részvény a PFE, míg a leginkább csúcsos eloszlás az MO részvényhez tartozik.

20. táblázat: A részvényhozamok ferdesége

részvény	ferdeség	részvény	ferdeség	részvény	ferdeség	részvény	ferdeség
MO	-22,751	TXT	-0,719	GIS	-0,113	CVX	0,030
WY	-10,153	CCL	-0,651	MMM	-0,113	T	0,037
ROK	-9,338	BDX	-0,558	FRT	-0,106	WMT	0,055
CB	-9,078	L	-0,503	CAT	-0,105	FDX	0,061
STT	-5,662	HRB	-0,471	WBA	-0,086	BBT	0,075
XEL	-4,598	AVY	-0,467	MSFT	-0,082	SO	0,098
EFX	-3,074	AET	-0,447	NKE	-0,078	SPGI	0,103
WMB	-3,006	ADBE	-0,443	EQT	-0,073	K	0,122
CI	-2,535	ADSK	-0,420	LMT	-0,070	VZ	0,127
PG	-2,339	STI	-0,394	MCD	-0,063	TXN	0,130
CNP	-2,289	AMD	-0,372	DOV	-0,056	TIF	0,137
AAPL	-2,189	BA	-0,340	DIS	-0,053	BLL	0,148
SEE	-1,839	LUV	-0,331	IP	-0,046	RJF	0,170
CA	-1,831	BAC	-0,315	EMR	-0,042	TROW	0,197
BAX	-1,487	ABT	-0,309	TGT	-0,040	USB	0,204
CAH	-1,277	AME	-0,285	IBM	-0,039	JPM	0,211
HAL	-1,219	EXC	-0,160	PEG	-0,037	VNO	0,246
BBY	-1,172	WEC	-0,159	PSA	-0,031	ROST	0,254
LNT	-1,040	MAS	-0,157	VFC	-0,029	TRV	0,281
MRK	-1,006	ZION	-0,157	GE	-0,005	NEM	0,317
AFL	-0,945	CHD	-0,155	AXP	-0,004	VAR	0,320
CERN	-0,812	JNJ	-0,147	APD	-0,002	SCHW	0,348
UTX	-0,808	PFE	-0,143	KO	0,000	FOX	0,460
VLO	-0,758	AJG	-0,127	OKE	0,010	COO	0,896
UNH	-0,723	ALK	-0,120	XOM	0,021	DVN	2,059

Forrás: (saját számítás)

21. táblázat: A részvényhozamok csúcossága

részvény	csúcosság	részvény	csúcosság	részvény	csúcosság	részvény	csúcosság
PFE	6,728	EMR	9,195	RJF	12,796	COO	25,118
WEC	6,989	LUV	9,350	SPGI	13,347	CAH	25,320
TXN	7,021	ABT	9,415	ALK	13,564	BAX	25,521
WMT	7,161	ADSK	9,500	FOX	13,707	STI	27,680
SCHW	7,432	ROST	9,840	AMD	14,068	CERN	29,810
CAT	7,517	BA	9,924	JPM	14,245	BAC	30,521
FDX	7,523	CHD	9,937	ADBE	14,363	HAL	33,962
SO	7,744	TIF	9,937	VLO	16,591	LNT	35,059
APD	7,801	NKE	10,184	AET	16,936	TXT	38,373
NEM	7,997	JNJ	10,310	TRV	18,053	AFL	38,446
MCD	8,002	AVY	10,360	PSA	18,249	CA	42,974
BLL	8,093	VAR	10,546	CCL	19,128	DVN	53,624
MMM	8,105	IBM	10,768	BBT	19,291	PG	63,142
VZ	8,205	DIS	10,913	USB	19,617	AAPL	66,924
GIS	8,208	EXC	10,913	ZION	20,719	CI	71,995
DOV	8,277	LMT	10,935	BDX	20,832	EFX	81,911
K	8,314	AXP	11,069	UNH	21,995	SEE	98,324
T	8,547	AME	11,186	L	22,263	WMB	153,789
VFC	8,847	GE	11,360	UTX	22,279	XEL	157,444
KO	8,896	XOM	11,461	OKE	22,701	CNP	169,917
MSFT	8,910	PEG	11,485	MRK	23,451	STT	223,575
MAS	9,016	CVX	11,500	AJG	23,682	ROK	377,083
WBA	9,034	IP	11,519	VNO	23,821	CB	389,606
TROW	9,049	HRB	12,112	FRT	24,206	WY	433,005
TGT	9,128	EQT	12,620	BBY	24,456	MO	1235,142

Forrás: (saját számítás)

A kiterjesztett Dickey–Fuller (ADF)-teszt alapján 5%-os szignifikancia szint mellett az összes részvény logaritmikushozam-idősora stacionárius. A p-érték minden esetben 1% alatt van 1-120 lag paraméter mellett.

7.3 A várható többletveszteség becslési modelljeinek empirikus

összehasonlítása

Az értekezésben a modellek összehasonlítása a 6.5.3.3 fejezetben leírt Z_{ES_α} -függvényen (Acerbi & Szekely, 2017) alapul. A Z_{ES_α} -mutató kiválasztása előtt megvizsgáltam a Z_1 – és a Z_2 -mutatót is. Z_2 -mutató hátránya, hogy a Z_{ES_α} -mutatóval ellentétben jelentős érzékenységet mutat a VaR-becslésre. Ezen túlmenően a backtesting-mutatók közötti különbség a tesztek különböző erejében is megjelenik (Acerbi & Szekely, 2014) (Acerbi & Szekely, 2017). Mindhárom mutatót az Acerbi–Szekely szerzőpáros definiálta, tanulmányukban a Z_{ES_α} -mutató használatát javasolják (Acerbi & Szekely, 2017). A Z_{ES_α} -értékre a későbbiekben az egyszerűség kedvéért Z-értékként hivatkozom. A számításokhoz használt ES-konfidenciaszint értéke 97,5%.

A 22. táblázat a Z-értékek összesítését tartalmazza a modellek teljesítménye szerinti sorrendben. Az összesítés alapját a modellek részvényenkénti, éves bontásban kiszámított Z-értéke adta, majd ezek abszolút értéke került összesítésre. Így elkerülhető, hogy az egyik részvény alábecsült ES-értékét egy másik részvény túlbecsült ES-értéke kompenzálja, vagy ugyanígy későbbi évek pozitív és negatív eltérései tűnjenek el átlagolás közben.

$$|Z| = \sum_{y=1}^Y \sum_{s=1}^S \left| \sum_{p=1}^{P_y} Z_{s,y,p,ES_\alpha} \right| \quad (253)$$

A fenti képletben Y az évek száma, S a részvények száma, P_y az időszakok száma az y. évben, Z_{s,y,p,ES_α} pedig az s. részvény, y. évének p. időszakához tartozó Z-érték. Minden egyes modell Z-értéke 100 részvény 29 éves Z-értékének, azaz 2900 éves Z-érték összesítésének eredménye. A 65. ábrán a két legjobb modell teljesítménye látható éves bontásban. Az ábrán jól látható, hogy mindkét modell rosszabb teljesítményt mutatott a 2001-es dotcom-válság és a 2007–2008-as pénzügyi válság idején.

22. táblázat: |Z|-értékek összesítése modellenként

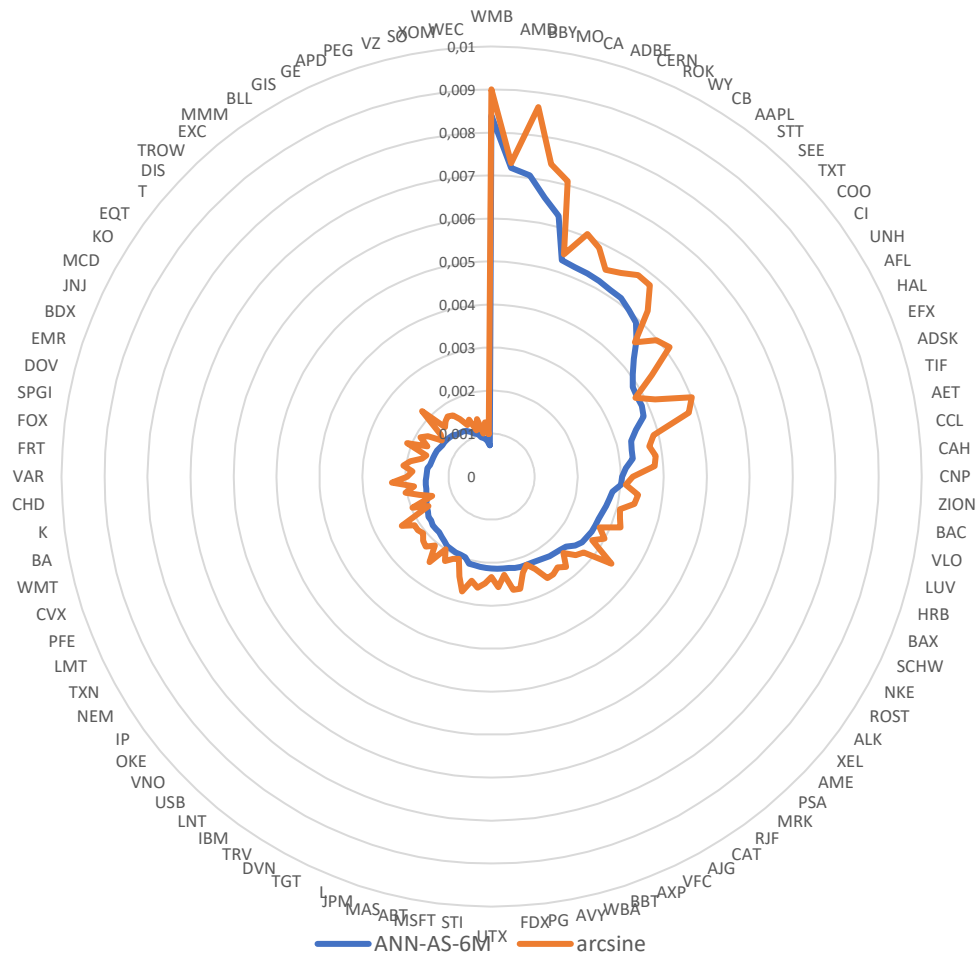
S.	Modell	Z	S.	Modell	Z	S.	Modell	Z
1	ANN-AS	0,260	22	logisrayleigh	0,721	43	kumgumbel	0,997
2	arcsine	0,302	23	Weibull	0,758	44	gjrgarch-arma-norm	1,000
3	TL2	0,310	24	loggamma	0,764	45	gumbel	1,003
4	tsp	0,313	25	stacygamma	0,769	46	egarch-arma-norm	1,019
5	exppois	0,314	26	kumexp	0,800	47	Gamma	1,153
6	lomax	0,314	27	kumnormal	0,824	48	logisexp	1,197
7	expext	0,317	28	LNbeta	0,826	49	betadist	1,271
8	explog	0,318	29	asylaplace	0,837	50	egarch-arma-sstd	1,310
9	loglog	0,318	30	sgarch-arma-std	0,885	51	burr	1,326
10	lfr	0,397	31	gjrgarch-arma-std	0,888	52	clg	2,253
11	genpowerweibull	0,488	32	explogis	0,889	53	aep	3,145
12	kum	0,502	33	kumhalfnorm	0,898	54	f	3,482
13	chen	0,503	34	normal	0,928	55	compbeta	3,680
14	burr7	0,507	35	logistic	0,933	56	MIaplace	11,724
15	triangular	0,509	36	egarch-arma-std	0,934	57	perks	19,626
16	kumburr7	0,559	37	t4	0,971	58	PCTAlaplace	1,45E+02
17	lognorm	0,614	38	QR	0,976	59	laplace	4,69E+02
18	rgamma	0,647	39	gjrgarch-arma	0,983	60	logcauchy	5,57E+02
19	xie	0,666	40	sgarch-arma-sstd	0,983	61	Cauchy	6,52E+03
20	frechet	0,666	41	sgarch-arma-norm	0,989	62	MRbeta	1,11E+27
21	expexp	0,691	42	historical	0,996	63	moweibull	1,76E+58
						64	ast	1,20E+123

Forrás: (saját számítások)

A 22. táblázat értékeiből kiderül, hogy a bemutatott idősorokon a legjobb modell az ANN-AS-neurálisháló volt. Az eloszlásalapú modellek közül az arcsine adta a legjobb eredményt. A Garch- és a kvantilisregresszió-alapú modellek közepes eredményt mutattak. A legjobb Garch alapú modell az sgarch-arma-std modell volt, ami a 30. legjobb eredményt érte el, a QR pedig a 38. helyen végzett. A széles körben elterjedt normális és t-eloszláson alapú, valamint a historikus szimuláció modellek a 34., 37. és a 42. helyet érték el. Fontos kiemelni, hogy a normális eloszláson alapuló modell hibája az arcsine modell hibájának 307%-a, azaz több, mint háromszorosa. A t4 modell hibája 330% az arcsine-modellhez viszonyítva. Ez alapján elmondható, hogy az eloszlásalapú modellek

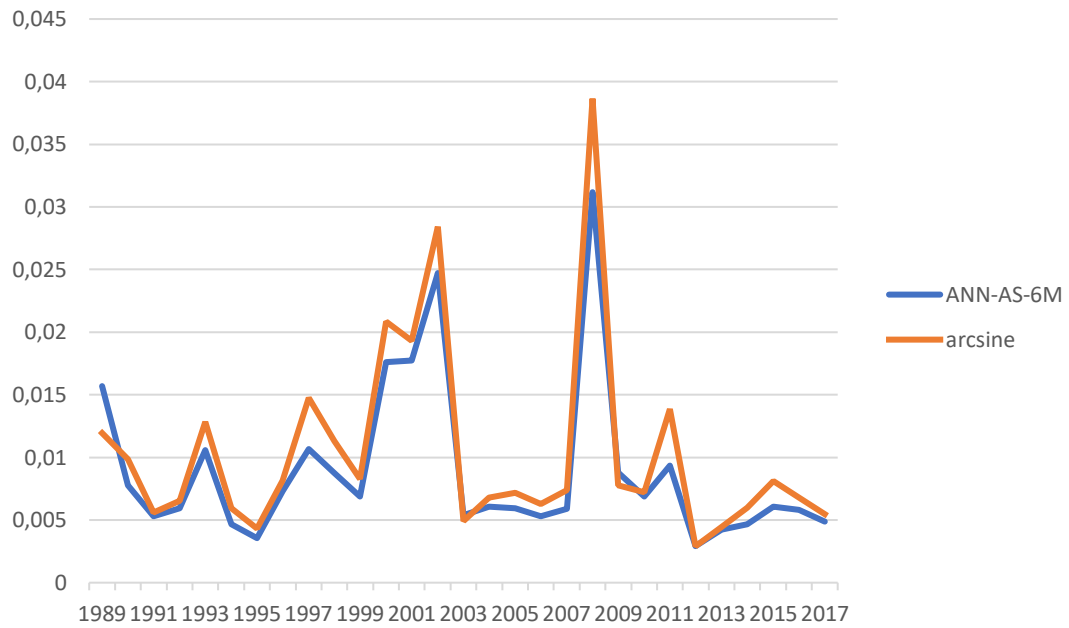
mellett maradvány is nagyon komoly előrelépést lehet elérni megfelelő eloszlásfüggvény megválasztásával.

64. ábra: Az Arcsine és az ANN-modell |Z|-értékeinek összehasonlítása részvényenként



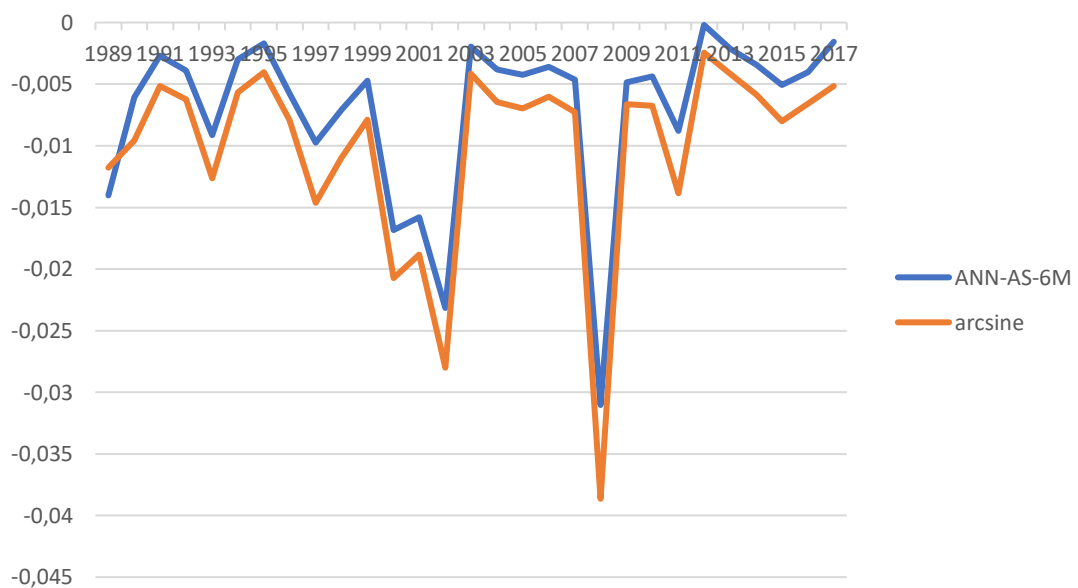
Forrás: (saját számítások)

65. ábra: Az Arcsine és az ANN-modell $|Z|$ -értékeinek összehasonlítása éves bontásban



Forrás: (saját számítások)

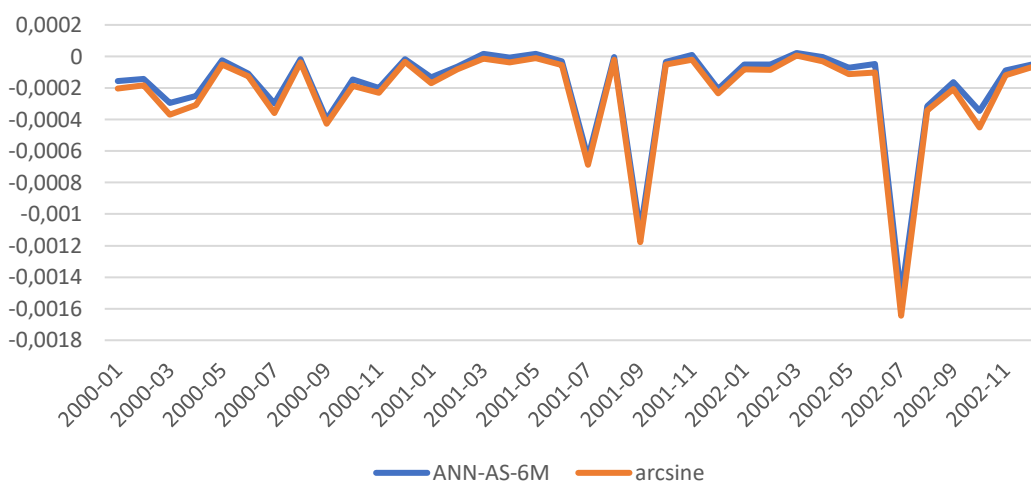
66. ábra: Az Arcsine és az ANN-modell Z-értékeinek összehasonlítása éves bontásban



Forrás: (saját számítások)

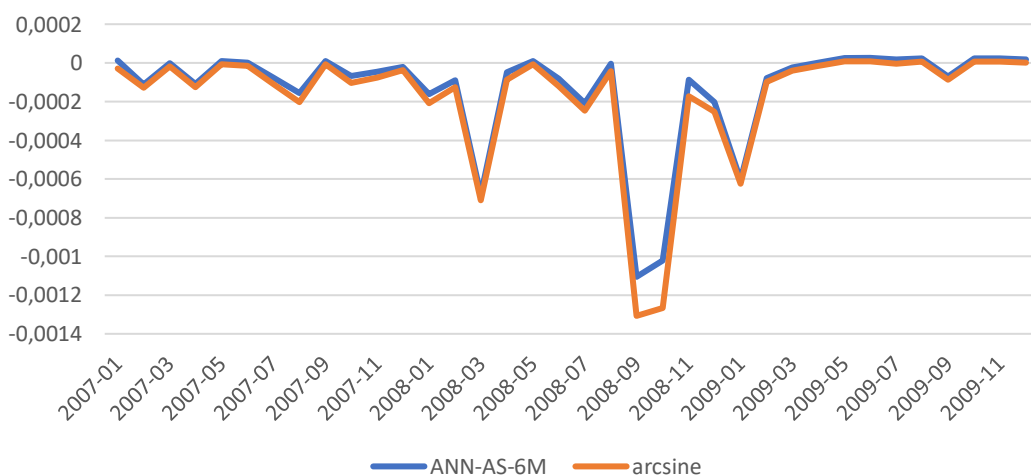
A legjobb eloszlásalapú modell, az arcsine 16%-kal nagyobb hibát mutat, mint az ANN-AS-modell. A 65. ábra mutatja, hogy az ANN-modell nem csak összesítésben nyújt jobb teljesítményt, hanem külön-külön minden évben és minden részvényre is. A 66. ábra alapján megállapíthatjuk, hogy mindkét modell a negatív Z-értékek felé hibázik, de az ANN-AS-modell hibája kisebb.

67. ábra: Az Arcsine és az ANN-modell Z-értékeinek összehasonlítása a dotcom-válság alatt havi bontásban



Forrás: (saját számítások)

68. ábra: Az Arcsine és az ANN-modell Z-értékeinek összehasonlítása a 2007-2008-as válság alatt havi bontásban



Forrás: (saját számítások)

A 67. ábra és a 68. ábra két válságos időszakban, a 2001-es dotcom-válság és a 2007–2008-as pénzügyi válság alatt mutatja be a két modell teljesítményét. Az ábrák a Z-érték előjelét megtartva mutatják be a modellek teljesítményét havi bontásban. A két ábra tanulsága alapján látszik, hogy a havi bontásban is mindkét modell a negatív Z-értékek irányába hibázik, azaz a modellek alábecsülik a kockázatot. Megfigyelhetjük, hogy a 2001-es dotcom-válság nagyobb kilengéseket eredményezett a modellek teljesítményében, mint a 2007-2008-as válság.

23. táblázat: Az arcsine- és az ANN-AS-modell összehasonlítása t- és z-próbával

	Próba	H_0	H_1	p-érték	Δ
1.	t-próba – Z	$\Delta = 0$	$\Delta \neq 0$	3,488e-06	$\Delta = \overline{Z}_{ANN} - \overline{Z}_{AS}$
2.	t-próba – Z	$\Delta \leq 0$	$\Delta > 0$	1,744e-06	$\Delta = \overline{Z}_{ANN} - \overline{Z}_{AS}$
3.	páros t-próba – Z	$\Delta = 0$	$\Delta \neq 0$	2,2e-16	$\Delta = \frac{1}{T} \sum_{t=1}^T Z_{ANN,t} - Z_{AS,t}$
4.	t-próba – Z	$\Delta = 0$	$\Delta \neq 0$	0,02308	$\Delta = \overline{ Z }_{ANN} - \overline{ Z }_{AS}$
5.	t-próba – Z	$\Delta \geq 0$	$\Delta < 0$	0,01154	$\Delta = \overline{ Z }_{ANN} - \overline{ Z }_{AS}$
6.	páros t-próba – Z	$\Delta = 0$	$\Delta \neq 0$	2,2e-16	$\Delta = \frac{1}{T} \sum_{t=1}^T Z _{ANN,t} - Z _{AS,t}$
7.	z-próba	$\Delta \leq 0$	$\Delta > 0$	1,706e-06	$\Delta = \overline{Z}_{ANN} - \overline{Z}_{AS}$

Forrás: (saját számítások)

A teljes időszakot vizsgálva a két modell közötti különbség 5%-os szignifikancia mellett szignifikáns, ahogy ezt a 23. táblázat is mutatja. A táblázatban \overline{Z}_{ANN} az ANN-modell, \overline{Z}_{AS} pedig az arcsine-modell Z-értékeinek átlagát, $\overline{|Z|}_{ANN}$ az ANN-modell, $\overline{|Z|}_{AS}$ pedig az arcsine-modell Z-értékeiből számított abszolútértékek átlagát jelöli. A táblázatban a T a Z-értékek idősorának hosszát, H_0 az alap, H_1 pedig az alternatív hipotézist jelöli. Az első próba azt ellenőrzi, hogy a két eloszlás átlagának különbsége nem nulla, a második azt, hogy az arcsine-modell Z-értékeinek átlaga kisebb, mint az ANN-AS-modellé. A 3., páros próba az ellenőrzi, hogy két idősor páros különbségeinek átlaga eltér-e a nullától. A 4-6. próbák ugyanezt ellenőrzik, csak abszolút Z-értékeket használnak a különbséggel, hogy az abszolút érték miatt az 5. próba azt ellenőrzi, hogy az ANN-AS-modell átlagos Z-értéke kisebb-e, mint arcsine-modellé.

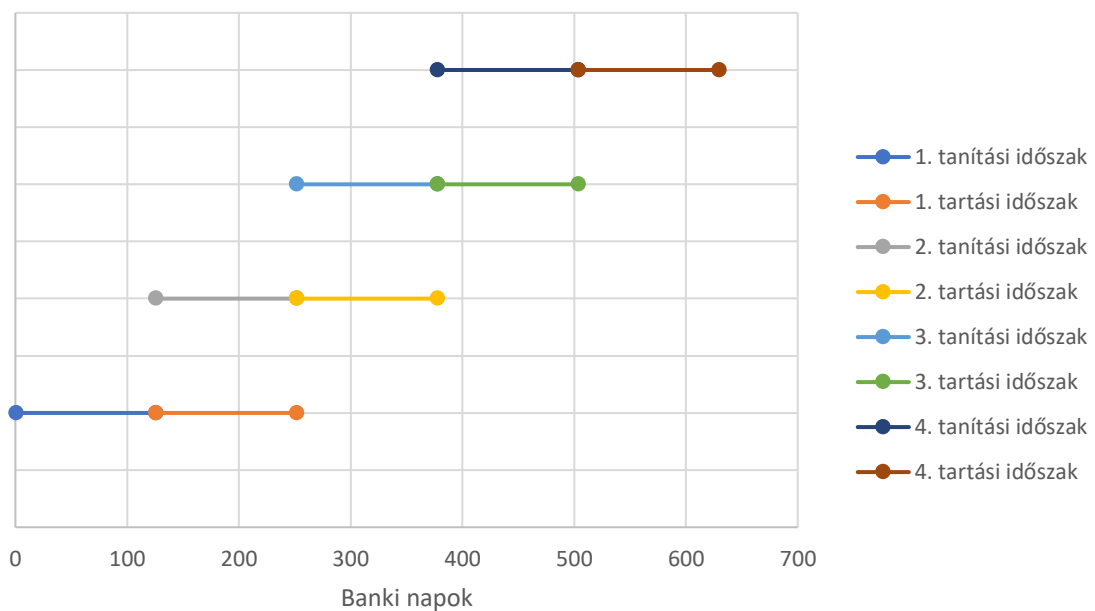
7.4 Az ANN-AS-modell érzékenységvizsgálata

Az ANN-modellek definiálásában néhány paraméter finomhangolása lényeges javulást érhet el a modell pontosságában. Az első két ilyen paraméter a tanítási időszak (TAI) és a tartási időszak (TRI) hossza. A bemutatott modellvariációknál minden esetben az epochok száma 200. Nincs külön kritérium definiálva a tanulás leállítására, így az a tanítási idősor végén, az epochok számával megadott ismétlésszám után fog leállni.

7.4.1 Egyenlő tanítási és tartási időszakok (TAI = TRI)

A tanítási időszak hossza azt befolyásolja, hogy hány banki nap, azaz hány adat alapján tanul az adott neurális háló. A neurális háló tanítása újra és újra megtörténik minden részvényfeldolgozás és minden tartási időszak végén. A 69. ábra ezeket az egymást átfedő időszakokat mutatja be egyenlő, hathónapos tanítási és hathónapos tartási idősorozatokra.

69. ábra: Egymást átfedő hathónapos tanítási és tartási idősorozatok



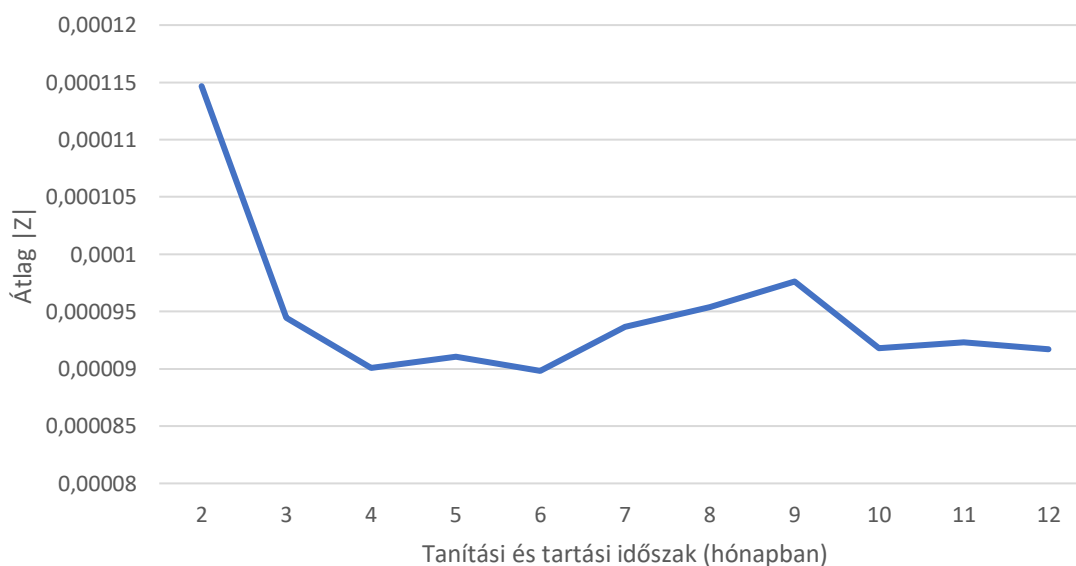
Forrás: (saját szerkesztés)

24. táblázat: Az azonos hosszúságú tanítási és tartási időszakok hatása

Tanítási és tartási időszak hónapban	Átlag Z	Max. Z	Min. Z
1	0,000257	0,006754	8,59E-07
2	0,000115	0,005005	1,37E-07
3	9,44E-05	0,007208	8,65E-08
4	9,01E-05	0,005116	4,6E-09
5	9,11E-05	0,005148	2E-08
6	8,98E-05	0,005154	1,68E-08
7	9,36E-05	0,00731	3,24E-09
8	9,54E-05	0,005158	5,72E-08
9	9,76E-05	0,005185	2,35E-08
10	9,18E-05	0,005285	1,39E-07
11	9,23E-05	0,007341	2,95E-08
12	9,17E-05	0,004794	2,07E-09

Forrás: (saját számítások)

70. ábra: Az azonos hosszúságú tanítási és tartási időszakok hatása



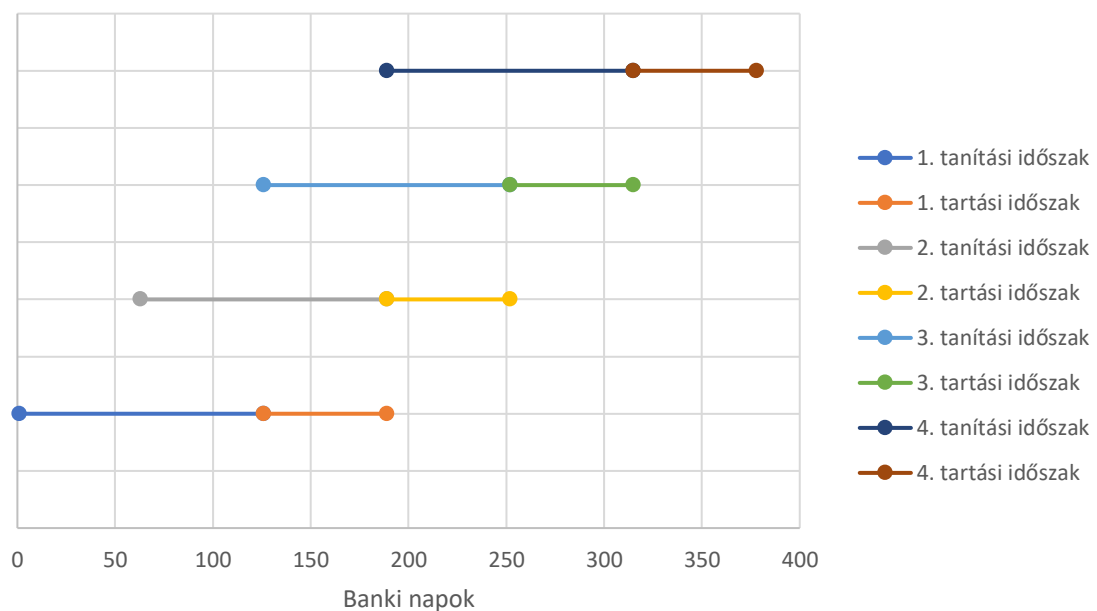
Forrás: (saját számítások)

A tanítási időszak hosszának megválasztásával az befolyásolható, hogy a felhasznált bementei adatok milyen hosszú ideig befolyásolják az aktuális modellt. Minél nagyobb a tanítási időszak hossza, annál tovább emlékszik a modell egy adott értékre. A

tartási időszak hossza gyakorlatilag azt adja meg, hogy milyen gyakran változnak meg a modellparaméterek. Minél rövidebb a tartási időszak, annál több idő szükséges az adott szimuláció futtatásához, mivel a rövidebb tartási időszak (TRI) azt eredményezi, hogy több modellt kell majd felhasználni a teljes 30 éves időszak lefedésére. Konkrétan az egy hónapos tartási időszakra paraméterezett modell futtatása 12-szer annyi idő, mint ugyanaz a modell 12 hónapos tartási időszakkal. A 24. táblázat és a 71. ábra azt mutatják be, hogy milyen hatással vannak a különböző tanítási és tartási időszakok a modell eredményességére. A paraméterérzékenység-vizsgálat alapján megállapítható, hogy a tesztelt idősoron a neurális hálók a 6 hónapos tartási időszak mellett eredményezik a legpontosabb becslést.

7.4.2 Eltérő tanítási és tartási időszakok (TAI > TRI)

71. ábra: Egymást átfedő eltérő hosszúságú tanítási és tartási időszakok



Forrás: (saját szerkesztés)

25. táblázat: Az eltérő hosszúságú tanítási és tartási időszakok hatása

Tartási időszak hónapokban	Tanítási időszak hónapokban	Átlag Z	Max. Z	Min. Z
1	6	9,17E-05	7,06E-03	2,98E-08
2	6	9,10E-05	5,07E-03	2,00E-08
3	6	9,14E-05	7,24E-03	3,32E-08
4	6	9,18E-05	5,23E-03	5,42E-08
5	6	9,20E-05	7,30E-03	6,89E-08
6	6	8,98E-05	5,15E-03	1,68E-08

Forrás: (saját számítás)

A 25. táblázat eredményeiből kiderül, hogy nem érdemes eltérő tanítási és tartási időszakot használni az ANN-AS-modell futtatásakor, mivel ez a módosítás minden esetben rosszabb Z-értéket eredményez, mint az, amikor a mindkét időszak hossza hat hónap.

7.4.3 Aktivációs függvények

A következő tesztben a különböző aktivációs függvények választásának a modellek eredményére gyakorolt hatását vizsgálom. A vizsgálathoz a tanh-, sigmoid- és a relu-aktivációs függvények kerültek kiválasztásra. A modellben minden szinthez különböző aktivációs függvényt lehet megadni. Az érzékenységvizsgálat során a bemeneti és a rejtett rétegek aktivációs függvénye mindig ugyanaz, a kimenet pedig a leírás alapján eltérhet. A 26. táblázatban az első oszlopban feltüntetett függvény párost adja meg. Az első aktivációs függvény az input- és a rejtett rétegeké a második pedig a kimenet. A TANH-SIGM sor értelmezése tehát az, hogy az input és rejtett rétegek a TANH-aktivációs függvényt használják, a kimenet pedig a SIGM-függvényt.

A 26. táblázat és a 72. ábra az aktivációs függvények Z-értékekre gyakorolt hatását írják le. A számítási eredményekből kiderül, hogy a legjobb eredményt az adja, ha a bemeneti és a rejtett réteg a TANH-, a kimeneti réteg pedig a SIGM-aktivációs függvényt használja. SIGM-aktivációs függvény használata abból a szempontból nem intuitív, hogy a SIGM értékkészlete $(0,1)$, míg az elvárt kimenet értékkészlete $[0, \infty)$. Ettől függetlenül a 26. táblázat eredményeiből megfigyelhetjük, hogy mégis az a modell teljesít legjobban, ahol a kimenet a SIGM-aktivációs függvényt használja. Ez akkor lehetséges,

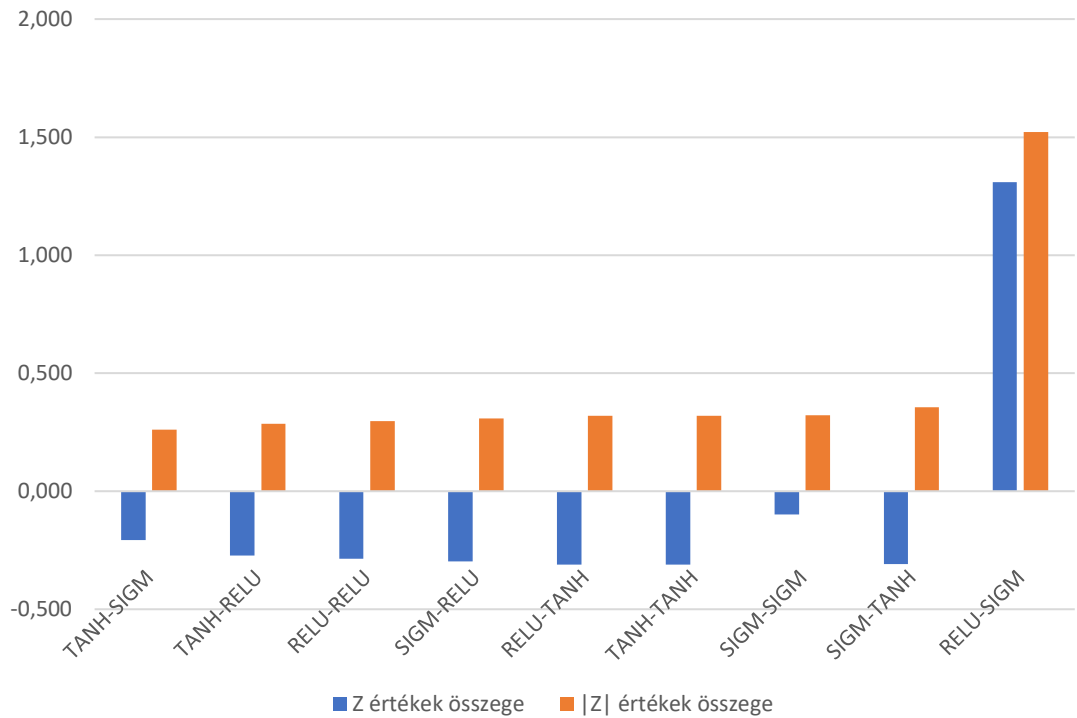
ha a tesztelési időszakban a modell elvárt kimenete többnyire a SIGM-aktivációs függvény értékkészletén belülre esik. A neurális hálók esetében elterjedt gyakorlat a bemenetek és a kimenetek normalizálása. Ennek célja pontosan az, hogy az így kapott idősorok a (0,1) tartományra essenek, ezzel elősegítve, hogy a korlátos értékkészletű aktivációs függvények használata mellett is megfelelő eredményeket kapjunk.

26. táblázat: Különböző aktivációs függvények eredményei

Aktivációs függvénypár	Z értékek összege	Z értékek összege
TANH-SIGM	-0,207	0,260
TANH-RELU	-0,272	0,285
RELU-RELU	-0,287	0,297
SIGM-RELU	-0,298	0,307
RELU-TANH	-0,312	0,319
TANH-TANH	-0,311	0,320
SIGM-SIGM	-0,099	0,323
SIGM-TANH	-0,308	0,355
RELU-SIGM	1,310	1,521

Forrás: (saját számítások)

72. ábra: Különböző aktivációs függvények eredményei



Forrás: (saját számítások)

7.4.4 Különböző optimalizációs algoritmusok eredményei

A legmélyebb leszállás eljárás működése különböző optimalizációs algoritmusokkal egészíthető ki. Az ANN-AS modell eredménye nagyban függ ettől a hiperparamétertől. A 27. táblázatban bemutatott eredmények alapján a legrosszabb optimalizációs algoritmus több, mint háromszor akkora $|Z|$ értéket eredményez, mint a legjobb optimalizációs algoritmus, a NADAM eljárás.

27. táblázat: Az optimalizációs algoritmusok hatása

Optimalizációs eljárás	$ Z $
NADAM	0,260
ADAMAX	0,268
ADAM	0,275
ADADELTA	0,277
SGDMOM	0,297
NAG	0,297
RMSPROP	0,300
ADAGRAD	0,787

Forrás: (saját számítások)

7.4.5 Különböző dropout százalékok eredményei

A dropout eljárás célja, hogy megakadályozza a modell túltanulását, úgy hogy bizonyos valószínűséggel elhagy egyes csomópontokat a tanulási fázisban. A dropout eljárás egyik bemenő paramétere az, hogy milyen valószínűséggel kerül elhagyásra egy adott csomópont. A 28. táblázatból látható, hogy a dropout eljárás használata a modell teljesítményének romlását eredményezi.

28. táblázat: A dropout százalék hiperparaméter hatása

Dropout százalék	Z
0	0,260
10%	0,268
20%	0,268
30%	0,275
40%	0,271

Forrás: (saját számítások)

7.4.6 Alternatív alapmodell

A neurális háló egy meglévő statisztikai modell eredményére épül. Adódik, hogy a statisztikai modell megválasztása is tekinthető a modell egyik hiperparaméterének. A 29. táblázat azt mutatja meg, hogy hogyan teljesít egy olyan neurális háló, amely a normális eloszlásra épít. Az eredmények alapján megállapíthatjuk, hogy a neurális háló használata egy sztenderd modell teljesítményét is javítja.

29. táblázat: A normális eloszlásra alapozó neurális háló teljesítménye

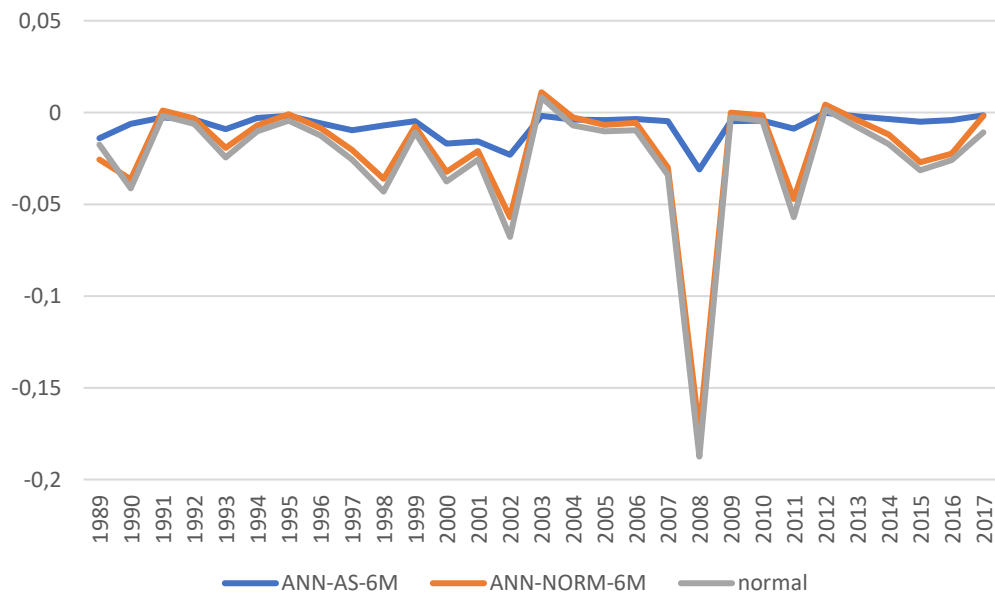
Modell	Z
ANN-AS-6M	0,260
ANN-NORM-6M	0,862
normal	0,947

Forrás: (saját számítások)

Ettől függetlenül az arcsine-alapú neurális háló teljesítménye lényegesen jobb. Ez arra enged következtetni, hogy a legjobb teljesítmény eléréséhez a megfelelő alapmodell

kiválasztása is elengedhetetlen. A 73. ábra a normális eloszlásra alapozó neurális háló teljesítményét a Z-értékek előjelének megtartásával mutatja be.

73. ábra: A normális eloszlásra alapozó neurális háló teljesítményének összehasonlítása



Forrás: (saját számítások)

7.4.7 További paraméterek

A modellalkotás során több további hiperparaméter alternatív értéke is tesztelésre került. Az egyik ilyen hiperparaméter az epochok száma, ahol a legjobb eredményt a jelenleg használt 200-as érték adja. Szintén felmerülhet az a kérdés, hogy a veszteségfüggvény különböző megválasztása tudná-e javítani az eredményt. A neurális hálót megvalósító szoftvercsomagban elérhető a MAE (átlagos abszolút hiba) és az MSE (átlagos négyzetes hiba) függvények, de lehetőség van további egyedi veszteségfüggvények definiálására is. Ezzel kapcsolatban egy fontos megkötés, hogy az egyedi veszteségfüggvény két bemenete csak az aktuális és az elvárt modellkimenet lehet. Ebből adódik, hogy a modellek értékeléséhez használt Z-függvény nem használható. A Z-érték kiszámításához a teljes ES-periódus hozamidősora, valamint a VaR- és az ES-beclés is szükséges. Az előbbi megkötés figyelembevételével elérhető veszteségfüggvények közül a MAE használata adta a legjobb eredményt.

7.5 Összefoglaló

Az értekezés elsődleges célja volt, hogy olyan várható-többletveszteség-becslési eljárásokat mutasson be, amelyek a közismert eljárásoknál lényegesen jobb teljesítményt adnak. Célját az értekezés egyrésztől azzal érte el, hogy több, mint 60 modell teljesítményét vizsgálta meg egységes keretrendszerben, 100 részvény 30 évet lefedő napihozam-adatbázisán. A modellek összehasonlítása az Acerbi–Székely-féle (2017) backtesting-eljárással történt, ami biztosította az eredmények aggregálhatóságát és lehetővé tette a modellek rangsorolását. Az összehasonlítás eredményeként bemutattam olyan valószínűségeloszlás-alapú modelleket, amelyek kimagasló teljesítményét a szakirodalom korábban nem tárgyalta, így feltehetőleg ezek kivételes becslési teljesítménye nem volt ismert. A legjobban teljesítő valószínűségeloszlás-alapú modellenél több, mint háromszor akkora becslési hibát vét a normális- és a t-eloszláson alapuló, vagy a historikus szimulációt alkalmazó modell. Az értekezés célját a modellek széleskörű összehasonlításán túl azzal érte el, hogy új megközelítésen alapuló modellt mutatott be. Az értekezésben olyan modellt konstruáltam és írtam le, amely a neurális hálók módszertanát alkalmazta arra, hogy a várható többletveszteség kockázati mutatót becsülje. Az így létrejött neurális hálón alapuló modell jobb becslést adott az összes többi modellenél, a második legjobb modellenél pedig szignifikánsan kisebb becslési hibát eredményezett.

Az értekezés áttekintette a kockázattal és a kockázati mutatókkal kapcsolatos szakirodalmat elsősorban a kockázatot érték (value-at-risk) és a várható többletveszteség (expected shortfall) mutatók szemszögéből. Értekezésemben külön részleteztem a két kockázati mutató megjelenését a szabályozásban és használatát a gyakorlatban, ezzel támasztva alá, hogy a téma nem csak tudományos szempontból, hanem többek között olyan gyakorlati szereplők számára is releváns, mint a bankok, tőzsdei elszámolóházak, biztosítók, nyugdíj- vagy befektetési alapok. Ezt követően értekezésemben bemutattam a neurális hálók szakirodalmi hátterét, majd az összehasonlításban szereplő valószínűségi eloszláson, Garch-folyamaton, kvantilis-regresszió, valamint neurális hálón alapuló modelleket. Az értekezés részletesen taglalta az összehasonlítás alapját képező idősorokat, azok statisztikai tulajdonságait, valamint a modell-összehasonlítás eredményeit. Értekezésemben külön kitértem a neurális háló

modellezési részleteire, valamint arra, hogy mely hiperparaméterek milyen hatással voltak a modell teljesítményére.

7.5.1 Az induló munkahipotézisek igazolása

H1: Érdeemes a várható többletveszteség modellezésben az eloszlásfüggvények széles körét tesztelni, mert létezhetnek olyan kevésbé ismert eloszlások, amelyek a klasszikus modellekhez képest lényegesen jobb becsléseket eredményeznek.

Az értekezésben bemutatott arcsine-eloszláson alapuló modellnél a közismert normális- és a t-eloszláson alapuló modellek több, mint háromszor akkora hibát vétettek, így kijelenthető, hogy fontos és eredményes volt a kevésbé ismert eloszlások teljesítményének tesztelése a várható többletveszteség modellezésének kontextusában.

H2: Lehetséges olyan neurális hálózat létrehozása, amely alkalmas a várható többletveszteség kockázati mutató előrejelzésére és a többi jól teljesítő modellel összemérhető pontosságú becslést ad.

A neurális hálók módszertana alkalmazható a várható többletveszteség kockázati mutató becslésére. A szerző által kifejlesztett és ismertetett megközelítés az arcsine-eloszláson alapuló modell becsléseit korrigálja a tanítási fázisban feldolgozott korábbi megfigyelések alapján.

H3: Neurális hálók segítségével létrehozható olyan modell, amely pontosabb várható többletveszteség-becslést ad, mint kizárólag a statisztikai modellek használata.

A szerző által kifejlesztett arcsine becslését korrigáló neurálisháló-modell szignifikánsan jobb eredményt adott, mint a teljes összehasonításban második helyet elérő arcsine-modell, így kijelenthetjük, hogy a neurális hálók módszertana nem csak adaptálható a várható többletveszteség becslési problémájára, hanem a módszertan alkalmazása a becslések további javítását teszi lehetővé.

H4: A neurális hálózat becslésének pontossága javítható oly módon, hogy a neurális hálózat újratanítását gyakrabban végezzük el.

Értekezésemben megvizsgáltam számos hiperparaméter hatását a neurális hálón alapuló modell teljesítményére. Az egyik hiperparaméter a tartási időszak hossza. A tartási időszak azt adja meg, hogy a neurális háló súlyait milyen hosszú ideig tartjuk meg. A tartási időszak végén a tanítási folyamat megismétlésével a neurális hálózat

csomópontjait összekötő élek új súlyokat kapnak, ezáltal a modell összefüggései is megváltoznak. Minél kisebb a tartási időszak hiperparaméter értéke, annál gyakrabban tanul meg új összefüggéseket a modell. Az eredményekből kiderül, hogy a tartási időszak csökkentése a 12 és a 6 hónap közötti intervallumon az eredmények javulásához vezet. A tartási időszak további csökkentése ezt követően a modell teljesítményének romlását vonja maga után. Az eredmények alapján kijelenthető, hogy a modellezés során érdemes a tartási időszak hatását viszonylag széles intervallumon tesztelni, majd kiválasztani az optimális értéket, mert a hiperparaméter csökkentése nem minden esetben eredményezi a modell javulását.

7.5.2 További lehetséges kutatási irányok

A várható többletveszteség mutató becslési modelljeivel kapcsolatban több továbblépési lehetőség is felmerül. Egyrésztől érdekesnek tartom a legsikeresebb eloszlásalapú modellek szűk körének vizsgálatát portfólió kontextusban különböző kopulák felhasználásával. Ez a kutatási irány az értekezésben bemutatott rendkívül széles modellpalettával nem megvalósítható, mivel a portfóliókkal kapcsolatban felmerülő szimulációk olyan jelentős számítási igénnyel járnak, hogy futtatásuk az elérhető hardverkapacitások mellett éveket vett volna igénybe.

A másik továbblépési lehetőség a neurális hálók továbbfejlesztése az inputidősorok bővítésével. Ez a potenciális kiegészítőadatok elérhetőségének függvényében valószínűleg a feldolgozott idősorok hosszának csökkenését vonja maga után, de könnyen elképzelhető, hogy vannak olyan releváns információkat tartalmazó idősorok, amelyek a neurális hálón alapuló modell becsléseit tovább javítják.

8 Felhasznált irodalom

Abadi, M. és mtsai., 2015. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. *arXiv preprint arXiv:1603.04467*.

Acerbi, C. & Szekely, B., 2014. Back-testing Expected Shortfall. *MSCI*.

Acerbi, C. & Szekely, B., 2017. General properties of backtestable statistics. *Working Paper*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2905109. kötet.

Acerbi, C. & Székely, B., 2019. The minimally biased backtest for ES. *Risk.net*, 29. kötet, pp. 1-6.

Altrichter, M. és mtsai., 2006. *Neurális hálózatok*. Budapest: Panem Könyvkiadó Kft..

Anon., 1911. *Révai nagy lexikona*. Budapest: Révai Testvérek Irodalmi Intézet Részvénytársaság.

Artzner, P., Delbaen, F., Eber, J.-M. & Heath, D., 1999. Coherent Measures of Risk. 9. kötet, pp. 203-228.

Az Európai Közösségek Tanácsa, 1973. *Az Európai Unió Hivatalos Lapja - 73/239/EEC*. [Online]

Available at: <https://eur-lex.europa.eu/legal-content/HU/TXT/PDF/?uri=CELEX:31973L0239&from=EN>
[Hozzáférés dátuma: 23 05 2021].

Az Európai Közösségek Tanácsa, 1979. *Az Európai Unió Hivatalos Lapja - 79/267/EEC*. [Online]

Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:31979L0267&from=en>
[Hozzáférés dátuma: 23 05 2021].

Az Európai Parlament és a Tanács, 2002a. *Az Európai Unió Hivatalos Lapja - 2002/12/EC*. [Online]

Available at: <https://eur-lex.europa.eu/legal->

[content/EN/TXT/PDF/?uri=CELEX:32002L0012&from=EN](https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32002L0012&from=EN)

[Hozzáférés dátuma: 23 05 2021].

Az Európai Parlament és a Tanács, 2002b. Az *Európai Unió Hivatalos Lapja* - 2002/13/EC.

[Online]

Available at: [https://eur-lex.europa.eu/legal-](https://eur-lex.europa.eu/legal-content/HU/TXT/PDF/?uri=CELEX:32002L0013&from=EN)

[content/HU/TXT/PDF/?uri=CELEX:32002L0013&from=EN](https://eur-lex.europa.eu/legal-content/HU/TXT/PDF/?uri=CELEX:32002L0013&from=EN)

[Hozzáférés dátuma: 23 05 2021].

Azzalini, A. & Capitano, A., 2003. Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t distribution. *Journal of the Royal Statistical Society*, 65. kötet, pp. 367-389.

Badics, M. C., 2014. Tőzsdei idősorok előrejelzése adatbányászati módszerekkel. *Hitelintézeti Szemle*, 13.(4.), pp. 207-227.

Bain, L. J., 1974. Analysis for the Linear Failure-Rate Life-Testing Distribution. *Technometrics*, 16(4), pp. 551-559.

Bayer, S. & Dimitriadis, T., 2019b. *Package 'esreg'*. [Online]

Available at: <https://cran.r-project.org/web/packages/esreg/esreg.pdf>

[Hozzáférés dátuma: 18 06 2021].

BCBS, 2021. *The Basel Framework*. [Online]

Available at: https://www.bis.org/basel_framework/index.htm?export=pdf

[Hozzáférés dátuma: 09 05 2021].

Behn, M., Haselmann, R. & Vig, V., 2016. *The limits of model-based regulation*, hely nélkül.: ECB Working Paper 1928.

Béli, M. & Váradí, K., 2017. Alapletét meghatározásának lehetséges módszertana. *Hitelintézeti Szemle*, 16.(2.), pp. 117-145.

Bellini, F., Klar, B., Müller, A. & Gianin, E. R., 2014. Generalized quantiles as risk measures. *Insurance: Mathematics and Economics*, Issue 54, pp. 41-48.

- Berkowitz, J., 2000. A coherent framework for stress testing. *Journal of Risk*, 2/2. kötet, pp. 5-15.
- Berlinger, E., Dömötör, B., Illés, F. & Váradi, K., 2016. A tőzsdei elszámolóházak vesztesége. *Közgazdasági Szemle*, LXIII. kötet, pp. 993-1010.
- Bernardi, M., 2013. Risk measures for skew normal mixtures. *Statistics and Probability Letters*, 83. kötet, pp. 1819-1824.
- Bollerslev, T., 1986. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31. kötet, pp. 307-327.
- Borgi, O., 1965. Sobre una distribución de frecuencias. *Trabajos Estadística*, 16. kötet, pp. 171-192.
- Boshnakov, G. N., 2019. *Package 'cvar'*. [Online]
Available at: <https://cran.r-project.org/web/packages/cvar/cvar.pdf>
[Hozzáférés dátuma: 16 02 2021].
- Broda, S. A. & Paolella, M. S., 2011. Expected shortfall for distributions in finance. In: *Statistical Tools for Finance and Insurance*. Berlin: Heidelberg, pp. 57-99.
- Budhiraja, A., 2016. *Dropout in (Deep) Machine learning*. [Online]
Available at: <https://medium.com/@amarbudhiraja/https-medium-com-amarbudhiraja-learning-less-to-learn-better-dropout-in-deep-machine-learning-74334da4bfc5>
[Hozzáférés dátuma: 12 01 2021].
- Bugár, G., 2015. *Piaci és hitelkockázat menedzsment*. Budapest: Akadémiai Kiadó.
- Bugár, G. & Rattig, A., 2016. A piaci kockázat számszerűsítésének változása a Bazel III szabályozásban. *Hitelintézeti Szemle*, 15. évf. 1. szám. kötet, p. 33–50.
- Bugár, G. & Uzsoki, M., 2005. Nemzetközi részvény befektetési lehetőségek Közép- és Kelet-Európa új Európai Unió tagállamainak szemszögéből. *Közgazdasági Szemle*, 52/6. kötet, pp. 576-598.
- Burr, I. W., 1942. Cumulative frequency functions. *The Annals of Mathematical Statistics*, 13(2), pp. 215-232.

- Chekhlov, A., Uryasev, S. & Zabarankin, M., 2005. Drawdown measure in portfolio optimization. *International Journal of Theoretical and Applied Finance*, 8(1), pp. 13-58.
- Chen, Z., 2000. A new two-parameter lifetime distribution with bathtub shape or increasing failure rate function. *Statistics & Probability Letters*, 49(2), pp. 155-161.
- Christoffersen, P. F., 1998. Evaluating Interval Forecasts. *International Economic Review*, 39/4. kötet, pp. 841-862.
- Christoffersen, P. F., 2012. *Elements of Financial Risk Management*. 2nd szerk. Oxford: Elsevier.
- Consul, P. & Jain, G., 1971. On the log-gamma distribution and its properties. *Statistische Hefte*, 12. kötet, pp. 100-106.
- Cordeiro, G. M. & de Castro, M., 2011. A new family of generalized distributions. *Journal of Statistical Computation and Simulation*, 81(7), pp. 883-898.
- Cordeiro, G., Narajajah, S. & Ortega, E., 2012a. The Kumaraswamy Gumbel Distribution. *Statistical Methods and Applications*, 168(21), p. 139.
- Cordeiro, G., Pescim, R. & Ortega, E., 2012b. The Kumaraswamy Generalized HalfNormal Distribution for Skewed Positive Data. *Journal of Data Science*, 10. kötet, p. 195-224.
- Csóka, P., 2003. Koherens kockázatmérés és tőkeallokáció. *Közgazdasági Szemle*, L. kötet, pp. 855-880.
- Dauphin, Y. és mtsai., 2014. Identifying and attacking the saddle point problem in high-dimensional non-convex optimization. *arXiv*, 1406.2572 . kötet, pp. 1-14.
- De La Vega, J., 1688. *Confusion De Confusiones*. Boston: Baker Library.
- Dematos, G. és mtsai., 1996. Feedforward versus recurrent neural networks for forecasting monthly japanese yen exchange rates. *Financial Engineering and the Japanese Markets*, 3. kötet, pp. 59-75.
- Dimitriadis, T. & Bayer, S., 2019. A joint quantile and expected shortfall regression framework. *Electornic Journal of Statistics*, 13. kötet, pp. 1823-1871.

- Dozat, T., 2016. *Incorporating Nesterov Momentum into Adam*. hely nélk., ICLR.
- Dr. Aradi, P., Gräff, J. & Dr. Lipovszki, G., 2014. *Számítógépes szimuláció*. hely nélk.:BME MOGI.
- Duchi, J., Hazan, E. & Singer, Y., 2011. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, 12. kötet, pp. 2121-2159.
- Dunbar, N., 2001. *Inventing Money*. 1st szerk. Chichester: Wiley.
- Edgeworth, F. Y., 1888. XXII. On a new method of reducing observations relating to several quantiles. *Philosophical Magazine Series*, 5. kötet, pp. 184-191.
- Embrechts, P. & Hofert, M., 2014. Statistics and Quantitative Risk Management for Banking and Insurance. *Annual Review of Statistics and Its Application*, 1. kötet, pp. 493-514.
- Emmer, S., Kratz, M. & Tasche, D., 2015. What is the best risk measure in practice? - A comparison of standard measures. *Journal of Risk*, 18(2), pp. 31-60.
- Engle, R. F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4), pp. 987-1007.
- Erdős, P. & Kac, M., 1947. On the number of positive sums of independent random variables. *Bulletin of the American Mathematical Society*, 53(10), pp. 1011-1020.
- Európai Parlament, 2009. *DIRECTIVE 2009/138/EC*. [Online]
Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32009L0138>
[Hozzáférés dátuma: 09 05 2021].
- Everitt, B. S. & Skrondal, A., 2010. *The Cambridge Dictionary of Statistics*. 4th szerk. Cambridge: Cambridge University Press.
- Farley, B. G. & Clark, W. A., 1954. Simulation of self-organizing systems by digital computer. *Transaction of the IRE Professional Group on Information Theory*, 4. kötet, pp. 76-84.

Federal Office of Private Insurance, 2006. *Technical document on the Swiss Solvency Test*.
[Online]

Available at:

https://www.finma.ch/FinmaArchiv/bpv/download/e/SST_techDok_061002_E_wo_Li_20070118.pdf

[Hozzáféres dátuma: 31 01 2021].

Fischer, T. & Krauss, C., 2018. Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, pp. 1-16.

Fissler, T. & Ziegel, J. F., 2016. Higher order elicibility and Osband's principle.. *Annals of Statistics*, 44(4), pp. 1680-1707.

Fréchet, M., 1927. Sur la loi de probabilité de l'écart maximum. *Annales de la Société Polonaise de Mathématique*, 6. kötet, pp. 93-116.

Fritsch, S. & Guenther, F., 2019. *Package 'neuralnet'*. [Online]

Available at: <https://cran.r-project.org/web/packages/neuralnet/neuralnet.pdf>

[Hozzáféres dátuma: 18 01 2021].

Fukushima, K., 1979. Neural network model for a mechanism of pattern recognition unaffected by shift in position-Neocognitron. *Transactions of the IECE*, 62(10), pp. 658-665.

Galeshchuk, S., 2016. Neural networks performance in exchange rate prediction. *Neurocomputing*, 172. kötet, pp. 446-452.

Ghalanos, A., 2020. *Package 'rugarch'*. [Online]

Available at: <https://cran.r-project.org/web/packages/rugarch/rugarch.pdf>

[Hozzáféres dátuma: 16 02 2021].

Giorgi, E., 2005. Reward-risk portfolio selection and stochastic dominance. *Journal of Banking & Finance*, 29. kötet, pp. 895-926.

Glosten, L. R., Jagannathan, R. & Runkle, D. E., 1993. On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *The Journal of Finance*, XLVIII(5), pp. 1779-1801.

- Gneiting, T., 2011. Making and Evaluating Point Forecasts. 106(494), pp. 746-762.
- Gumbel, E. J., 1954. The maxima of the mean largest value and of the range. *The Annals of Mathematical Statistics*, 25(1), pp. 76-84.
- Gupta, R. & Kundu, D., 1999. Generalized exponential distributions. *Australian & New Zealand Journal of Statistics*, 41(2), pp. 173-188.
- Gupta, R. & Kundu, D., 2001. Exponentiated exponential family: an alternative to gamma and Weibull distributions. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, 43(1), pp. 117-130.
- Gylberth, R., 2018. *An Introduction to AdaGrad*. [Online]
Available at: <https://medium.com/konvergen/an-introduction-to-adagrad-f130ae871827>
[Hozzáférés dátuma: 06 01 2021].
- Hagan, M. T., Demuth, H. B., Beale, M. H. & Jesús, O. D., 2014. *Neural Network Design (2nd Edition)*. hely nélk.: ismeretlen szerző
- Hamilton, J. D., 1994. *Time Series Analysis*. New Jersey: Princeton University Press.
- Hebb, D. O., 1949. *The Organization of Behavior*. New York: John Wiley & Sons.
- Hill, I. D., Hill, R. & Holder, R. L., 1976. Fitting Johnson Curves by Moments. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 25(2), pp. 180-189.
- Hinkley, D. V. & Revankar, N. S., 1977. Estimation of the Pareto law from underreported data. *Journal of Econometrics*, 5. kötet, pp. 1-11.
- Hinton, G. & et al., 2012. Deep neural networks for acoustic modeling in speech. *IEEE Signal Processing Magazine*, 29. kötet, pp. 82-97.
- Hinton, G., dátum nélk. *Lecture 6e rmsprop: Divide the gradient by a running average of its recent magnitude*. [Online]
Available at: http://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf
[Hozzáférés dátuma: 08 01 2021].

- Hopfield, J. J., 1982. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79(8), pp. 2554-2558.
- Hornik, K., 1989. Multilayer Feedforward Networks are Universal Approximators. *Neural Networks*, 2. kötet, pp. 359-366.
- Hutchins, W. J., 1995. Machine Translation: A Brief History. In: *Concise History of the Language Sciences*. hely nélk.:Pergamon, pp. 431-445.
- Hybel, J., 2018. *Formulating quantile regression as Linear Programming problem?*. [Online]
Available at: <https://stats.stackexchange.com/questions/384909/formulating-quantile-regression-as-linear-programming-problem/407478#407478>
[Hozzáférés dátuma: 15 05 2019].
- IAIS, 2014. *Risk-based Global Insurance Capital Standard, public consultation document*, hely nélk.: International Association of Insurance Supervisors.
- IRMA, 2019. *Deep Learning and Neural Networks: Concepts, Methodologies, Tools, and Applications*. Hershey: IGI Global.
- Johnson, N. L. & Tadikamalla, P., 1949. Systems of Frequency Curves Generated by Transformations of Logistic Variables. *Biometrika*, 36. kötet, pp. 149-176.
- Jones, M., 2002. The complementary beta distribution. *Journal of Statistical Planning and Inference*, 104(2), pp. 329-337.
- Jorion, P., 2007. *Value at Risk - The New Benchmark for Managing Financial Risk*. 3rd szerk. New York: McGraw-Hill.
- Ketkar, N., 2017. *Deep Learning with Python*. hely nélk.:ismeretlen szerző
- Kingma, D. P. & Ba, J., 2014. Adam: A Method for Stochastic Optimization. *arXiv*, 1412.6980. kötet.
- Koenker, R., 2005. *Quantile Regression*. Cambridge: Cambridge University Press.

- Koenker, R., 2021. *Package 'quantreg'*. [Online]
Available at: <https://cran.r-project.org/web/packages/quantreg/quantreg.pdf>
[Hozzáférés dátuma: 16 02 2021].
- Koenker, R. & Bassett, G., 1978. Regression Quantiles. *Econometrica*, 46(1), pp. 33-50.
- Koenker, R. W. & D'Orey, V., 1987. Computing Regression Quantiles. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 36(3), pp. 383-393.
- Kolås, Ø., 2005. *Image Processing with gluas*. [Online]
Available at: http://pippin.gimp.org/image_processing/chap_dir.html#id2549721
[Hozzáférés dátuma: 26 12 2020].
- Komáromi, É., 2002. *Operációkutatás 2.*. Budapest: BKÁE KTK.
- Kotosz, B. & Ferenci, T., 2010. Nemnormális, parametrizált eloszlású valószínűségi változók. *Statisztikai Szemle*, 88(7-8), pp. 803-832.
- Kotz, S., Kozubowski, T. J. & Podgórski, K., 2001. Asymmetric Laplace Distributions. In: *The Laplace Distribution and Generalizations*. Boston, MA: Birkhäuser.
- Kozubowski, T. & Podgórski, K., 2000. Asymmetric Laplace distributions. *Mathematical Science*, 25. kötet, pp. 37-46.
- Kratz, M., Lok, Y. H. & McNeil, J. A., 2018. Multinomial VaR backtests: A simple implicit approach to backtesting expected shortfall. *Journal of Banking and Finance*, 88. kötet, pp. 393-407.
- Kristjanpoller, W., Fadic, A. & Minutolo, M. C., 2014. Volatility forecast using hybrid Neural Network models. *Expert Systems with Applications*, 41. kötet, pp. 2437-2442.
- Kumaraswamy, P., 1980. A generalized probability density function for double-bounded random processes. *Journal of Hydrology*, 46(1), pp. 79-88.
- Kus, C., 2007. A new lifetime distribution. *Computational Statistics & Data Analysis*, 51(9), pp. 4497-4509.
- Lahmiri, S., 2017. Modeling and predicting historical volatility in exchange rate. *Physica*, pp. 387-395.

- Längkvist, M., Karlsson, L. & Loutfi, A., 2014. A review of unsupervised feature learning and deep learning A review of unsupervised feature learning and deep learning. *Pattern Recognition Letters*, 42. kötet, pp. 11-24.
- Lan, Y. & Leemis, . L. M., 2008. The logistic–exponential survival distribution. *Naval Research Logistics*, 55(3), pp. 252-264.
- LeCun, Y., Bengio, Y. & Hinton, G., 2015. Deep learning. *Nature*, 521. kötet, pp. 436-444.
- Leung, M., Xiong, H., Lee, L. & Frey, B., 2014. Deep learning of the tissue-regulated splicing code. *Bioinformatics*, 30. kötet, pp. 121-129.
- Libby, D. L. & Novick, M. R., 1982. Multivariate Generalized Beta Distributions with Applications to Utility Assessment. *Journal of Educational Statistics*, 7(4), pp. 271-294.
- Li-Bland, D., 2018. *Saddle Points and Stochastic Gradient Descent*. [Online]
Available at: <https://davidlibland.github.io/posts/2018-11-10-saddle-points-and-sdg.html>
[Hozzáférés dátuma: 05 01 2021].
- Linnainmaa, S., 1970. *The representation of the cumulative rounding error of an algorithm as a Taylor expansion of the local rounding errors*. Master's Thesis (in Finnish), hely nélk.: Univ. Helsinki.
- Lomax, K. S., 1954. Business Failures: Another Example of the Analysis of Failure Data. *Journal of American Statistical Association*, 49(268), pp. 847-852.
- Macukow, B., 2016. Neural Networks - State of Art, Brief History, Basic Models and Architecture. *15th IFIP Internation Conference on Computer Information Systems and Industrial Management*, pp. 3-14.
- Ma, J. és mtsai., 2015. Deep Neural Nets as a Method for Quantitative Structure–Activity Relationships. *Journal of Chemical Information and Modeling*, 55(2), pp. 263-274.
- Mannix, R. & St. Clair, B., 2020. *Why investors are stuck with flawed VAR models*. [Online]
Available at: <https://www.risk.net/investing/7659866/why-investors-are-stuck-with-flawed->

var-models

[Hozzáférés dátuma: 10 02 2021].

Markowitz, H. M., 1952. Portfolio Selection. *The Journal of Finance*, 7(1), pp. 77-91.

Marshall, A. & Olkin, I., 1997. : A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika*, 84. kötet, pp. 641-652.

McCarthy, J., Minsky, M. L., Rochester, N. & Shannon, C. E., 1955. *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, hely nélk.: ismeretlen szerző

McCulloch, W. S. & Pitts, W., 1943. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5. kötet, pp. 115-133.

McDonald, J. & Richards, D., 1987a. Hazard rates and generalized beta distribution. *IEEE Transactions on Reliability*, 36. kötet, pp. 1049-1074.

McDonald, J. & Richards, D., 1987b. Model selection: Some generalized distributions. *Communications in Statistics – Theory and Methods*, 16. kötet, p. 1049–1074.

McGill, W. J., 1962. Random fluctuations of response rate. *Psychometrika*, 27. kötet, pp. 3-17.

McNeil, A. J., Embrechts, P. & Frey, R., 2015. *Quantitative Risk Management*. 2nd szerk. New Jersey: Priceton University Press.

Minsky, M. L. & Papert, S. A., 1969. *Perceptrons : An introduction to computational geometry*. hely nélk.:MIT Press.

Miskolczi, P., 2016. A value at risk és az expected shortfall összehasonlítása történeti szimuláció segítségével. *Sigma*, XLVII.(3-4.), pp. 139-160.

Mosteller, F. & Tukey, W. J., 1977. *Data Analysis and Regression: A Second Course in Statistics*. hely nélk.:Addison-Wesley.

- Nadarajah, S., Chan, S. & Afuecheta, E., 2015. *Package 'VaRES'*. [Online]
Available at: <https://cran.r-project.org/web/packages/VaRES/VaRES.pdf>
[Hozzáférés dátuma: 31 7 2019].
- Nadarajah, S. & Haghghi, F., 2011. An extension of the exponential distribution. *Statistics: A Journal of Theoretical and Applied Statistics*, 45(6), pp. 543-558.
- Nagy, L. & Balogh, P., 2013. *Ökonometria*. Debrecen: Debreceni Egyetem.
- Nelson, D. B., 1991. Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*, 59(2), pp. 347-370.
- Nesterov, Y., 1983. A method for unconstrained convex minimization problem with the rate of convergence. *Doklady ANSSSR*, 269. kötet, pp. 543-547.
- Nikulin, M. & Haghghi, F., 2006. A chi-squared test for the generalized power weibull family for the head-and-neck cancer censored data. *Journal of Mathematical Sciences*, 133(3), pp. 1333-1341.
- Nodal, 2015. *CFTC Grants Nodal Clear Registration as a DCO*. [Online]
Available at: <https://www.nodalexchange.com/cftc-grants-nodal-clear-registration-as-a-dco/>
[Hozzáférés dátuma: 03 02 2021].
- Nodal, 2017. *EEX acquires 100% ownership of US Nodal Exchange*. [Online]
Available at: https://www.nodalexchange.com/wp-content/uploads/20170504_EEX_Nodal_Closing-Press-Release.pdf
[Hozzáférés dátuma: 03 02 2021].
- Nodal, 2017. *Nodal Clear recognized as third-country central counterparty by ESMA*. [Online]
Available at: <https://www.nodalexchange.com/wp-content/uploads/2017-03-31-Nodal-Clear-recognized-as-third-country-CCP-by-ESMA.pdf>
[Hozzáférés dátuma: 03 02 2021].
- NodalClear, 2021. *Margin Methodology*. [Online]
Available at: <https://www.nodalexchange.com/nodal-clear/risk-management/margin->

methodology/

[Hozzáférés dátuma: 02 02 2021].

OCC, 2021. *What Is OCC?*. [Online]

Available at: <https://www.theocc.com/Company-Information/What-Is-OCC>

[Hozzáférés dátuma: 02 02 2021].

Ou, S. & Yi, D., 2009. Robustness Analysis and Algorithm of Expected Shortfall Based on Extreme-Value Block Minimum Model. *Proceedings of the 2009 International Conference*, pp. 288-292.

Paranaíba, P., Ortega, E., Cordeiro, G. & Pascoa, M., 2013. The Kumaraswamy Burr XII Distribution: Theory and Practice. *Journal of Statistical Computation and Simulation*, 83. kötet, pp. 2117-2143.

Perks, W., 1932. On some experiments in the graduation of mortality statistics. *Journal of the Institute of Actuaries*, 63. kötet, pp. 12-40.

Pham, H., 2002. A Vtub-Shaped Hazard Rate Function with Applications to System Safety. *International Journal of Reliability and Applications*, 3(1), pp. 1-16.

Pintér, J. & Rappai, G., 2007. *Statisztika*. Pécs: Pécsi Tudományegyetem Közgazdaságtudományi Kar.

Poiraud-Casanova, S. & Thomas-Agnan, C., 2000. About monotone regression quantiles. *Statistics & Probability Letters*, 48(1), pp. 101-104.

Pröhle, T. & Zempléni, A., 2016. *Statistical Problem Solving in R*. hely nélk.: ismeretlen szerző

Qian, N., 1999. On the momentum term in gradient descent learning algorithms. *Neural Networks*, 12(1), pp. 145-151.

R-core, 2019. *stats4 v3.6.2*. [Online]

Available at: <https://www.rdocumentation.org/packages/stats4/versions/3.6.2>

[Hozzáférés dátuma: 16 02 2021].

- Repullo, R. & Suarez, J., 2008. *The Procyclical Effects of Basel II*. Washington, 9th Jacques Polak Annual Research Conference.
- Righi, M. B. & Ceretta, P. S., 2015. A comparison of Expected Shortfall estimation models. *Journal of Economics and Business*, 78. kötet, pp. 14-47.
- Risk.net, 2017. *Clearing house innovation of the year: Nodal Clear*. [Online]
Available at: <https://www.risk.net/awards/2480700/clearing-house-innovation-year-nodal-clear>
[Hozzáférés dátuma: 03 02 2021].
- Risk.net, 2020. *Before and after the Covid-19 storm: Risk.net's Buy-side Risk Management Survey 2020*. [Online]
Available at: <https://www.risk.net/media/download/1047451>
[Hozzáférés dátuma: 10 02 2021].
- Rochester, N., Holland, J. H., Haibt, L. H. & Duda, W. L., 1956. Test on a cell assembly theory of the action of the brain, using a large digital computer. *IEEE Transaction on Information Theory*, 2. kötet, pp. 80-93.
- Rockafellar, R. & Uryasev, S., 2002. Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, 26. kötet, pp. 1443-1471.
- Rodriguez, R. N., 1977. A guide to the Burr type XII distributions. *Biometrika*, 64(1), pp. 129-134.
- Rosenblatt, F., 1958. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), pp. 386-408.
- Ruder, S., 2017. An overview of gradient descent optimization algorithms. *arXiv*, Issue 1609.04747.
- Rumelhart, D. E., Hinton, G. E. & Williams, R. J., 1986. Learning representations by back-propagating errors. *Nature*, 323. kötet, pp. 533-536.
- Russell, S. J. & Norvig, P., 2005. *Mesterséges intelligencia - Modern megközelítésben*. 2. szerk. Budapest: Panem.

- Sallay, M. & Szigel, G., 2014. Elvek és gyakorlat: a magyar bankrendszer válasza az új tőkegyezmény kihívásaira. *Hitelintézeti Szemle*, 3(3-4), pp. 128-140.
- Sanderson, G., 2018. *Backpropagation calculus*. [Online]
Available at: <https://www.3blue1brown.com/>
[Hozzáférés dátuma: 04 01 2021].
- Sarhan, A. M. & Kundu, D., 2009. Generalized Linear Failure Rate Distribution. *Communications in Statistics - Theory and Methods*, 38(5), pp. 642-660.
- Schmidhuber, J., 2015. Deep learning in neural networks: An overview. *Neural Networks*, 61. kötet, pp. 85-117.
- SEC, 2017a. *Release No. 34-81399; File No. SR-LCH SA-2017-007*. [Online]
Available at: <https://www.sec.gov/rules/sro/lchsa/2017/34-81399.pdf>
[Hozzáférés dátuma: 02 020 2021].
- SEC, 2017b. *Release No. 34-82313; File No. SR-ICEEU-2017-013*). [Online]
Available at: <https://www.sec.gov/rules/sro/iceeu/2017/34-82313.pdf>
[Hozzáférés dátuma: 02 02 2021].
- SEC, 2019. *Release No. 34-85755; File No. SR-OCC-2019-004*. [Online]
Available at: <https://www.sec.gov/rules/sro/occ/2019/34-85755.pdf>
[Hozzáférés dátuma: 02 02 2021].
- Shapiro, S. & Wilk, M. B., 1965. An analysis of variance test for normality (complete samples). *Biometrika*, 52(3-4), pp. 591-611.
- Sharpe, W., 1966. Mutual Fund Performance. *Journal of Business*, 39. kötet, pp. 119-138.
- Sharpe, W. F., 1964. Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *Journal of Finance*, 19(3), pp. 425-442.
- Sheedy, E., 2008. Why VaR models fail and what can be done. *MAFC Research Papers*, 34. kötet.
- Simonato, J. G., 2011. The performance of Johnson distributions for computing value at risk and expected shortfall. *Journal of Derivatives*, 19. kötet, pp. 7-24.

- Srivastava, N. és mtsai., 2014. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, 15. kötet, pp. 1929-1958.
- Stacy, E., 1962. A Generalization of the Gamma Distribution. *Annals of Mathematical Statistics*, 33(3), pp. 1187-1192.
- Stigler, S. M., 1984. Boscovich, Simpson and a 1760 manuscript note on fitting a linear relation. *Biometrika*, 71(3), pp. 615-620.
- Széles, Z., Pataki, L. & Fóriánné Horváth, M., 2016. A bázeli tőkeszabályozás múltja, jelene és jövője. *Gazdaság & Társadalom*, 8(3), pp. 51-68.
- Tahmasbi, R. & Rezaei, S., 2008. A two-parameter lifetime distribution with decreasing failure rate. *Computational Statistics & Data Analysis*, 52(8), pp. 3889-3901.
- Taylor, J. W., 2008. Estimating Value at Risk and Expected Shortfall Using Expectiles. *Journal of Financial Economics*, 6. kötet, pp. 231-252.
- Thoma, M., 2008. *3d-gradient*. [Online]
Available at: <https://commons.wikimedia.org/wiki/File:3d-gradient-cos.svg#file>
[Hozzáférés dátuma: 31 12 2020].
- Topp, C. W. & Leone, F. C., 1955. A Family of J-Shaped Frequency Functions. *Journal of the American Statistical Association*, 50. kötet, pp. 209-219.
- Uzsoki, M., 2020. Neurális hálók a pénzügyi kockázatomérésben. *Sigma*, LI(3), pp. 287-299.
- van Dorp, J. R. & Kotz, S., 2002. The Standard Two-Sided Power Distribution and its Properties. *The American Statistician*, 56(2), pp. 90-99.
- Weibull, W., 1951. A Statistical Distribution Function of Wide Applicability. *Journal of Applied Mechanics*, 18. kötet, pp. 293-297.
- Werbos, P., 1974. *Beyond regression : new tools for prediction and analysis in the behavioral sciences (Ph. D. Thesis)*, hely nélk.: Harvard University.
- Widrow, B. & Hoff, M. E., 1960. *Adaptive Switching Circuits*, hely nélk.: Solid-State Electronics Laboratory.

- Wiesel, D. H. & Hubel, T. N., 1959. Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology*, 148. kötet, pp. 574-591.
- Xiao, Z., Guo, H. & Lam, M. S., 2015. Quantile Regression and Value at Risk. In: L. J. Lee CF., szerk. *Handbook of Financial Econometrics and Statistics*. New York: Springer.
- Xie, M., Tang, Y. & Goh, T. N., 2002. A modified Weibull extension with bathtub-shaped failure rate function. *Reliability Engineering System Safety*, 76(3), pp. 279-285.
- Yang, Q. & Wu, X., 2006. 10 challenging problems in data mining research. *International Journal of Intermination Technology & Decision Making*, 5(4), pp. 597-604.
- Yu, W., Yang, K., Wei, Y. & Lei, L., 2018. Measuring Value-at-Risk and Expected Shortfall of crude oil portfolio using extreme value theory and vine copula. *Physica A*.
- Zeiler, D. M., 2012. ADADELTA: An Adaptive Learning Rate Method. *arXiv: 1212.5701*.
- Zhu, D. & Galbraith, J. W., 2010. A generalized asymmetric Student-t distribution with application to financial econometrics. *Journal of Econometrics*, 157. kötet, pp. 297-305.
- Zhu, D. & Zinde-Walsh, V., 2009. Properties and estimation of asymmetric exponential power distribution. *Journal of Econometrics*, 148(1), pp. 86-99.
- Ziegel, J. F., 2016. Coherence and Elicitability. *Mathematical Finance*, 26(4), pp. 901-918.