

## A PORTFÓLIÓMÉRET HATÁSA A BIZTOSÍTÁSI DÍJRA ÉS A KOCKÁZATI MÉRTÉKEKRE REPÜLŐGÉP-SZERENCSETLENSÉGEK ADATAINAK ELEMZÉSE

Papp Márk Ádám (A szerző a Budapesti Corvinus Egyetem és az ELTE közös Biztosítási és pénzügyi matematika mesterszakán 2018-ban végzett aktuáriusként. Ez a cikk az év legjobb szakdolgozata címet elnyert szakdolgozat rövidített változata.) pappmark821@gmail.com

### ÖSSZEFOGLALÓ

A biztosítások működése veszélyközösségek létrehozásán alapul, mert ezáltal a károk alakulásában lévő bizonytalanság kiszámíthatóbbá válik. A biztosítási portfólió mérete tehát elsődleges a biztosító szempontjából, így azt kívántam megvizsgálni, hogy ez miképp hat a kárkifizetés eloszlására, illetve a díjakra.

Bár tanulmányom középpontjában egy ennek megfelelő módszertan áll, ezt egy olyan adatbázison akartam tesztelni, amelynek a vizsgálata önmagában is hasznos és érdekes eredményeket adhat. Ennek megfelelően repülőgép-szerencsetlenségek adatainak felhasználásával mutatom be azt, hogyan vizsgálható egy biztosítási portfólió méretének hatása a biztosítás kockázati mértékeire, illetve a díjakra.

Az eredmények mind a légi közlekedés biztonságosságát (mivel átlagosan 10.000.000 útra 6 olyan jut, ahol valaki életét veszti), mind a portfólióméret hatását tekintve (nagyobb méret jelentősen alacsonyabb díjakat tesz lehetővé) megfeleltek az előzetes várakozásaimnak.

### SUMMARY

Operation of the insurances are based on risk pooling so that the uncertainty of the total claims becomes more predictable. This means that the size of the insurance portfolio is fundamental for the insurer. In this article my aim was to investigate its effect on the distribution of the total claims and on the insurance premiums.

Although the main goal in my study was to show a method, I also wanted to test this method on a database, from which I can also get useful and interesting results. According to that I have decided to carry out the impact analysis of the portfolio size on the risk measures and premiums based on plane crash data.

Results have met my expectations concerning both the safety of air traffic (in average there are only 6 flights out of 10.000.000 ending with fatal accidents) and the impact of portfolio size (bigger portfolio allows lower premiums).

**Kulcsszavak:** Monte-Carlo szimuláció, fontosságot kifejező mintavételezés (Importance Sampling), portfólióméret, légi közlekedés, biztosítási díj

**Keywords:** Monte-Carlo simulation, Importance Sampling, portfolio size, air traffic, insurance premium

**JEL:** C20, C53, C63, G22, R41

**DOI:** 10.18530/BK.2018.4.28

<http://dx.doi.org/1018530/BK.2018.4.28>

### Bevezetés

Napjaink talán legbiztonságosabb közlekedési módja a repülés (IATA Economics (2018)). Évente alig pár balesetről hallani – különösen igaz ez a fejlett országok között közlekedő nagyobb utasszállító repülőkre –, és ezekenél is csak igen kis százalékban történik végzetes baleset. Az összes utas- és teherszállítót figyelembe véve is csak kevesebb, mint minden egymilliomodik repülés járt halálos balesettel 2000 óta. Ráadásul a trendeket figyelve az is egyértelmű, hogy a repülésbiztonság folyamatosan fejlődik, és az idő előrehaladtával még ez az arány is rohamosan csökken.

Habár az előző bekezdésben leírtak alapján akár azt is mondhatnánk, hogy biztosítási szempontból elhanyagolható terület a repülőutak biztosítása, mégis fontos téma. Ennek okai abban keresendők, hogy évente több mint 30 millió utas- és teherszállító repül, ezzel behálózva a bolygó minden pontját, valamint, hogy a szektor hozzájárulása hatalmas a globális GDP-hez, és erre káros hatása lehetne annak, ha az emberek bizalma megrendülne ebben a közlekedési formában.

Sok elméleti modell esetén ennek a kérdésnek a megválaszolása nem lenne nehéz feladat tisztán valószínűségszámítási és matematikai eszközökkel, de mint a későbbiekben bemutatom majd, az általam vizsgált modell keretei között az összetett kockázat (a teljes portfólió kárkifizetése) nem vezethető vissza egy hagyományos eloszlásra, így nem határozhatóak meg annak alapján az egyes mutatói. Ilyen esetekben gyakran használt módszer valamilyen mikroszimulációs eljárás alkalmazása, mint például a Monte-Carlo szimuláció.

A számítások során egyaránt használtam az R szoftvert – ez segített az adatok elemzésében és a statisztikai modellek illesztésében –, illetve az Excel programot. A szimulációkat az Excel (és a VBA) segítségével hoztam létre, hogy minél jobban átlátható legyen az eljárás és a modell szerkezete.

Végző célom alátámasztani, hogy nagyobb biztosítási portfólió a biztosító számára biztonságosabb pozíciót eredményez, ami abban is jelentkezik, hogy az egyes szerződések díja azonos paraméterek esetén akár alacsonyabb is lehet. További célom annak megállapítása, hogy – bizonyos elméleti feltételezések mellett – mekkora a minimálisan szükséges portfólióméret, hogy ne legyen túlárazott a biztosítás.

## Adatok

Bár egy kezdeti modell létrehozásához elég lehetne az az információ is, hogy összesen hány repülőjárat volt egy vizsgált időszakban, és melyiknél hány haláleset történt (mint a bevezetőben is írtam, ez a gépek több mint 99,9999 %-ában nullát jelent), a céloom egy ennél szofisztikáltabb elemzés elkészítése volt. Mivel ezt alkalmas statisztikai modellek illesztésével akartam elérni, elengedhetetlen volt a repülőjáratok néhány további paraméterének megismerése. Természetesen adódott a lehetőség, hogy egy kategorikus változókat tartalmazó GLM-et (általánosított lineáris modellt) használjak, hogy így a gépeket homogén alcsoportokra osztva határozzam meg a kárkifizetés nagyságát befolyásoló valószínűségi változókat (például utasok száma és halottak száma).

Több lehetőség is felmerült a felhasználandó paraméterek körében. Először is különbségek lehetnek a légitársaságok között. Valószínű továbbá, hogy különböző géptípusok is különböző (paraméterű) eloszlás szerinti számokat produkálnak. Sajnos ezzel a két változóval kapcsolatban felmerültek problémák is, ugyanis nem sikerült ilyen mélységű adatokat találnom a repülőgép-szerencsétlenségekhez.

Az általánosított lineáris modellben a lehetséges magyarázó változók körét a felhasznált adatbázis határozta meg. Mivel jellemzően pontos adatokat találhattam a gépek útvonalára (kiindulási hely és végcél), valamint a gépek működésének típusára (utasszállító, teherszállító, katonai, mentő, városnéző stb.), ezért végül ezeket a tényezőket megtartottam a megfelelő adatbázis kiválasztásánál.

### Az adatok közül felhasználtam a gép utasainak számát, a halottak számát, a gép rendeltetését, illetve a kiindulási és a célállomását.

A kárnagyság mint valószínűségi változó mellett hasonlóan fontos (vagy talán az alacsony valószínűségek miatt még fontosabb) a kárgyakoriság. Egy szerződésre vetítve ez egy Bernoulli-eloszlású valószínűségi változót jelent, hiszen vagy bekövetkezik halálos baleset egy járatnál, vagy nem. Ennek is lehetne a paraméterét egy modellel becsülni különböző változók alapján, de ehelyett itt azt a megoldást választottam, hogy minden alcsoporthoz (ezekről később részletesen írok) egyszerűen számítottam egy bekövetkezési valószínűséget.

Az interneten számos, repülőgép-baleseteket tartalmazó adatbázist lehet találni. Ezek egy része ingyenesen nem érhető el, de a maradék lehetőség is kellően megbízhatónak bizonyult. Végül a választásom a [www.planecrashinfo.com](http://www.planecrashinfo.com) adatbázisára esett. Ennek az adatbázisnak az előnye, hogy teljes körű, és ad néhány extra információt a gépekre vonatkozóan, melyek jól használhatók a későbbi általánosított lineáris modellekben. Az adatok közül felhasználtam a gép utasainak számát, a halottak számát

(fedélzeten életüket veszítők és földi áldozatok külön-külön), a gép rendeltetését, illetve a kiindulási és a célállomását.

Következő lépésben az adatok rendszerezését végeztem el, besoroltam minden gépet a tevékenysége, illetve földrajzi adatok alapján. Tevékenységüket tekintve a gépek a következő kategóriákba kerülhettek: utasszállító, kis utasszállító (19 főt nem meghaladó kapacitás esetén), teherszállító, illetve voltak olyan további kategóriák, mint helikopter, katonai, mentő vagy városnéző. Ez utóbbiakra azért nem fordítottam nagyobb figyelmet, mert egyrészt lényegesen kisebb volt az arányuk jellegükből adódóan, másrészt a teljes járatszámra nem találtam megfelelő adatot, így a kárgyakoriságra nem lett volna megfelelő modellem, ami nagyban torzította volna az eredményeket. Az utasszállítókon belüli szétválasztás mindenképp hasznos (ezt a GLM illesztésnél külön vizsgálom), de meghatározása önkényesen történt, illetve egyes esetekben kérdéses a besorolás, ha a géptípus alapján ez nem derült ki – ellenben ezek az esetek épp a határon lehetnek, így szignifikáns hibát semmiképp sem okozhattak. A földrajzi adatok tekintetében több lehetőség is felmerült. Először is lehetne a távolságot használni, mint magyarázó változót a káresemény bekövetkezésének valószínűségéhez, de ehhez megint csak teljes állomány (az összes járat) adataira lett volna szükség. Kárnagysághoz alkalmazhattam volna külön kategorikus változóként a kiindulási és a célországot, de így nagyon elaprózódhattak volna a csoportok. Így végül a járat kiindulási kontinensét tartottam meg (az esetek több mint 90%-ában ez megegyezett a célállomás kontinensével), illetve két kontinens között közlekedő járat esetén egy külön dummy változót (interkontinentális) hoztam létre. Annak alapján, hogy az összes repülésre milyen megbontású adatokat kaptam, végül a következő módon alakultak a csoportok:

- Típus: kis utasszállító, utasszállító, teherszállító
- Kontinens: Afrika, Ázsia és Ausztrália, Európa, Észak-Amerika, Latin-Amerika
- Interkontinentális

A kontinensek esetén érdemes megjegyezni, hogy nem biztos, hogy önmagában elegendő információt ad, tekintve, hogy lehet rövid út két kontinens között, és hosszabb egy kontinensen belül, de ez országszintű változó esetén is megeshet (például az olyan nagy kiterjedésűek miatt, mint Oroszország). Így összességében ezt találtam a leginkább alkalmas megoldásnak, és az egész adatbázist nézve nem is jelentkezik a leírt probléma olyan nagy mértékben. Az ezekben a csoportokba nem illő, illetve a hibásnak tűnő vagy hiányos adatok törlése után 657 elem maradt az adatbázisomban. Ez a 2018. február 15-i állapot szerint készült a 2000-2018-as adatok alapján. (Korábbi adatokat nem akartam bevonnai, mert az iparban a biztonság terén egyértelmű a fejlődés, így érdemes csak viszonylag friss adatokat felhasználni – kevesebb évvel pedig a modell megbízhatósága sérül a kevés adat miatt.) Az adatok végső megoszlását az ismertetett kategóriák szerint az 1. táblázat tartalmazza.

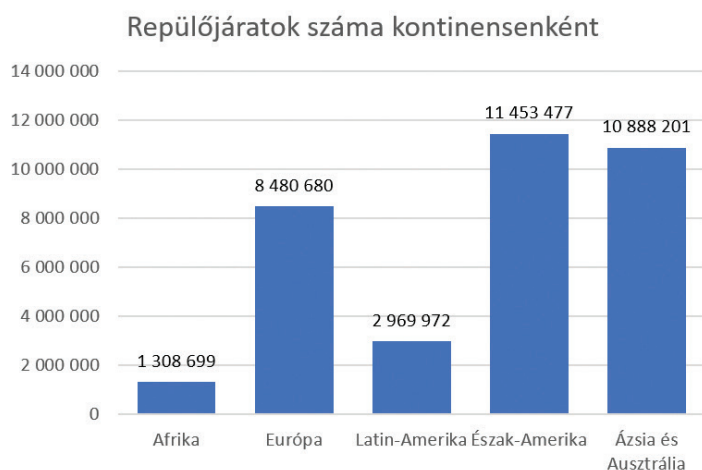
1. táblázat: Halálos áldozattal járó repülőgép-balesetek megoszlása (2000-2018) - zárójelben az ebből interkontinentálisok száma

	Afrika	Ausztr.	Ázsia	Észak-A.	Európa	Latin-A.	Összes
Teher	42 (3)	3 (0)	31 (5)	39 (3)	20 (5)	21 (2)	156 (18)
Kis utas	37 (0)	19 (0)	32 (5)	91 (5)	26 (2)	79 (2)	284 (9)
Utas	43 (3)	4 (0)	75 (12)	20 (5)	41 (8)	34 (3)	217 (31)
Összes	122 (6)	26 (0)	138 (17)	150 (13)	87 (15)	134 (7)	657 (58)

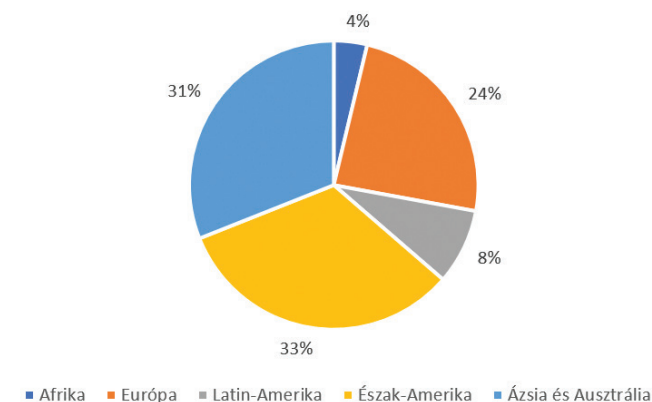
Forrás: saját gyűjtés

A káresemény bekövetkezésének valószínűségét tapasztalati relatív gyakorisággal akartam becsülni, így szükségem volt az összes repülőútra vonatkozó adatokra is. Ehhez az olyan jelentéseket használtam fel, amelyek kontinensekre lebontva tartalmazták az adott évekhez tartozó repülőjáratok számát (ATAG, 2014; ATAG, 2016). Ezek a számok csak az utas- és teherszállító gépeket tartalmazzák, ezért az egyéb rendeltetésű járatokat kiszűrtem az adatbázisomból. Az összes járat megoszlása kontinensek szerint az 1. ábráról olvasható le.

1. ábra: Kontinens szintű adatok a repülőjáratok számára (2012. és 2014. évek átlaga)



Repülőjáratok megoszlása kontinensek szerint



Forrás: saját gyűjtés

Néhány megjegyzést fontos tenni az adatbázissal kapcsolatban. Először is bizonyos kockázatokat elhanyagoltam, ugyanis az előfordulási gyakoriságuk alacsony volt, még az eleve ritka, halálessel járó baleseteken belül is. Ilyen szempontból fontos, hogy míg a valóságban a terrorcselekmény következtében történő elhalálozás ki lehet zárva a biztosító fizetési kötelezettségében, én ezeket – az információ esetleges hiánya miatt – nem kezeltem külön. Hasonló módon elhanyagoltam azokat az eseteket is, amikor két repülő ütközött a kifutópályán vagy a levegőben. Ugyan ez önmagában jelentősen módosíthatja a halottak számát vizsgáló valószínűségi változót, de egyrészt nagyon ritka volt az ilyen esemény, másrészt a baleset időpontját folytonos valószínűségi változóként tekintve az egybeesés valószínűsége nulla lesz.

Figyelemmel kell lenni továbbá arra is, ami a repülőgépek biztosításának talán legfontosabb tulajdonsága, hogy a katasztrófák bekövetkezési valószínűsége nagyon alacsony (kevesebb, mint egy a millióhoz), sőt, ez hosszú távon még mindig szignifikánsan javul. Ennek azért van jelentősége, mert egy-egy kiugró év észrevehetően módosíthatja az eredményeket (például 20 baleset helyett 22 már azonnal 10%-os valószínűségemelkedést eredményezhet). Ez természetesen nem azt jelenti, hogy a felhasznált módszertan ne lenne megfelelő, és a végső eredmények és következtetések haszontalanok lennének, azonban mindig tisztában kell lenni azzal, hogy mindez érzékeny lehet az inputként használt adatokra.

Mindezeket az adatokat felhasználva elő tudtam állítani egy olyan modellt, amely a járatok eloszlásában jól tükrözi a valóságot, illetve elégséges információt ad a károk bekövetkezéséről és nagyságát illetően azok szimulációjához.

## Modell

A modellillesztés előtt szükséges volt az adatbázis tisztítása. Ez részben azt jelentette, hogy bizonyos adatsorokat hiányzó információk miatt töröltem, illetve egyes kategóriákat összevontam. Külön figyelemmel kellett lenni a kiugró adatokra, hiszen egy-egy ilyen képes lehet jelentősen torzítani, és ez akár hibás eredményekhez is vezethetett volna.

Természetesen, mint minden modellbecslésnél, itt is fontos kérdés volt, hogy milyen adatokat használjak fel. Annak érdekében, hogy az eredmények hasznosak legyenek a későbbiekben, az összes áldozat számát, azaz a fedélzeten utazó és a földi áldozatok összegét használtam. A modell illesztése során egy további módosítást kellett eszközölni az adatbázisban. Ennek oka a 2001. szeptember 11-én történt katasztrófa. Ebben az esetben a földi áldozatok száma olyan kiugróan magas volt, hogy az már szignifikánsan torzította az eredményeket. Ennek érdekében próbálkoztam a földi áldozatok számára vonatkozó adatsor transzformálásával (logaritmus, gyök), de végül ennek az egy adatnak az elhagyása segített a megfelelő modell létrehozásában. Elvi megalapozása is van természetesen a változtatásnak, hiszen a legtöbb biztosításban a terrorcselekmény következtében bekövetkező káresemény a biztosító mentesülését eredményezheti a kifizetés alól. Bár több hasonló baleset is lehet, de azokat (ugyanúgy, mint itt a gépen utazók közül elhunytakat) nem szűrtem ki, az alacsony előfordulási gyakoriság és a baleset esetlegesen hiányzó vagy nem egyértelmű oka miatt.

Ezen a ponton érdemes tenni egy rövid kitérőt. Biztosítási kockázatok elemzésének egy gyakori eszköze az extrémérték elmélet, illetve azon belül az általánosított Pareto-eloszlás illesztése. Ennek a lényege abban áll, hogy bizonyos feltételek teljesülése esetén meghatározható egy kockázat farokeloszlása, és ezáltal közvetlenül nyerhető információk a kiugró adatokat illetően. Mivel a halálos kimenetelű repülőgép-szerencsétlenségek elég ritka események, és értelmezhetőek az összes út azon részhalmazaként, ahol az áldozatok száma 0 fölött volt, ezért ezt fel lehetne fogni úgy, mint egy minden utat figyelembe vevő eloszlás farokeloszlását. Az általánosított Pareto-eloszlás tulajdonságait, illetve alkalmazását itt nem részletezem, csak a legfontosabb eredményeket ismertetem röviden.

Az adatokból megállapítható, hogy az áldozatok számának eloszlása korlátos. Ez kézenfekvő megfontolások alapján egyezik a várakozásaimmal, hiszen a halottak számát korlátozza a gépen utazók száma maximumának (ami a legnagyobb utasszállító kapacitás, tehát nyilvánvalóan véges) és a földi áldozatok száma maximumának (ami megint csak véges) összege. Itt érdemes megjegyezni egy külön vizsgálat eredményét. A 2001. 09. 11-i katasztrófa földi áldozatainak számát elhagytam bár, de külön megvizsgáltam, hogy megtartása mit eredményezne. Abban az esetben, amikor az eredeti adatokkal dolgoztam, a halottak számának eloszlása nemkorlátosnak adódott, ami az előbb leírt logikának ellentmond, így nem megfelelő a modellezésre.

## Kárgyakoriságok vizsgálata

Összetett kockázatok modellezésénél jellemzően két részre szokás bontani a problémát. Ennek fényében külön kerül modellezésre a károk nagysága és a károk gyakorisága. Ebben a tanulmányban én is ezt a megközelítést fogom alkalmazni, így először a halálos kimenetelű balesetek előfordulási gyakoriságát fogom elemezni, majd a repülőgép-szerencsétlenségekben elhunytak számának eloszlását.

A halálos kimenetelű balesetek bekövetkezése egy egyszerű Bernoulli valószínűségi változóval jellemezhető, ami 0 értéket vesz fel, ha nem volt halálos áldozat, és 1-et, ha volt. Mivel egy Bernoulli-eloszlásnak egyetlen paramétere van (annak a valószínűsége, hogy a valószínűségi változó értéke 1 lesz), csak ezt kellett becsülnöm. Ehhez a tapasztalati gyakoriságokat használtam fel, azaz a valószínűségszámítás egyik legelemibb eljárása alapján vettem a „jó esetek számát” osztva az „összes esettel”. Mivel az összes repülőútra volt adatom kontinensenként (Észak- és Latin-Amerika szétválasztva, viszont Ausztrália és Ázsia összevonva), így eszerint megbontottam az adatokat. A típusok arányára nem találtam ilyen jellegű adatot, de az Amerikai Egyesült Államokra vonatkozóan volt adat a teherszállítók arányára az összes járaton belül, így ezt használtam globálisan is. A további adatok (kis és nagy utasszállító, illetve interkontinentális járat) tekintetében nem tudtam tovább bontani a sokaságot, így ezeken a szinteken nincs külön bekövetkezési valószínűség.

Mivel a korábbi évek adatai alapján egyértelműen kijelenthető, hogy az ilyen balesetek valószínűsége folyamatosan csökken, ezért mindenképpen érdemes csak viszonylag friss adatokat használni. A járatok teljes számára vonatkozóan 2012-es és 2014-es adatokat találtam (ATAG, 2014; ATAG, 2016), így ezek átlagát használtam fel. Hogy ezzel konzisztens legyek, de a véletlen fluktuációkat is lehetőleg csökkentsem, ezért a szerencsétlenségek számára egy hosszabb idősort használtam, így a 2010–2017 közötti évek átlagát vettem alapul. A tehergépek arányára 2012 és 2017 közötti időszakra havi szintű adatokat találtam, és ezek átlagát vetítettem az összes gépre (2012 és 2014 átlaga), hogy a teherszállítók számát megkapjam. Mivel a teherszállítók aránya a vizsgált időszakban stabilan 7 százalék körül mozgott (Bureau of Transportation Statistics), így a teherszállítók szerencsétlenségeire is a 2010–2017-es adatok átlagát használtam.

A kapott eredmények a várakozásoknak megfelelően alakultak, és a következő megfigyelések tehetőek:

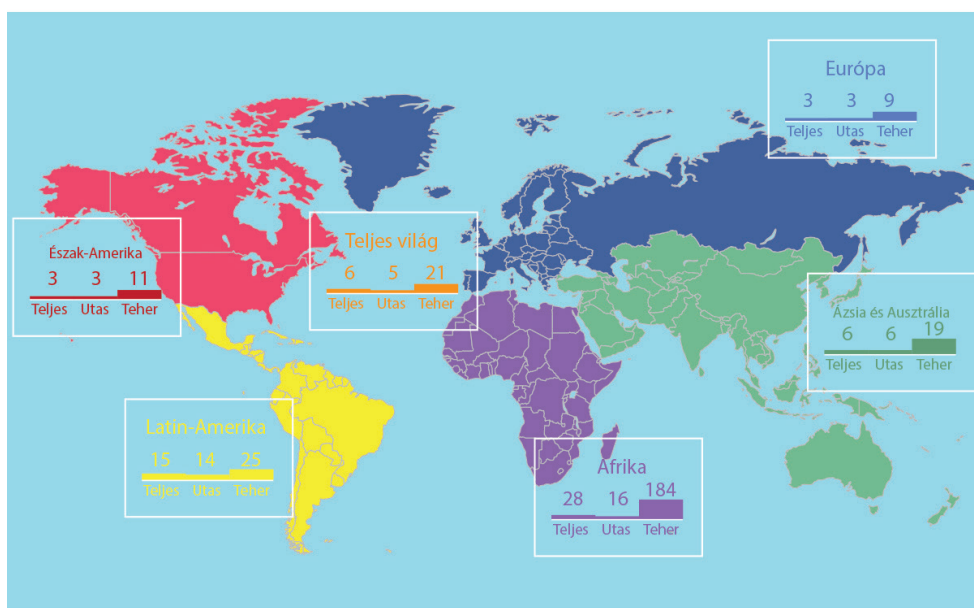
- Az átlagos valószínűsége annak, hogy egy utas- vagy teherszállító olyan balesetet szenved a repülés során (vagy a kifutópályán), melynek halálos áldozata lesz: 0,000061 százalék, vagyis kevesebb, mint minden 1.600.000-edik repülésnél következik be várhatóan.
- A teherszállítók esetében ez az arány (1:485000) lényegesen nagyobb, nagyjából négyszerese az utasszállítókénak (1:1985000).
- A fejlettebb országokat magában foglaló Észak-Amerika, illetve Európa esetén lényegesen nagyobb a biztonság, mint Afrika és Latin-Amerika esetén. Az Ausztráliát és Ázsiát magában foglaló kategória a két típus között nagyjából középen helyezkedik el.



- A legkisebb valószínűsége egy ilyen balesetnek egy Észak-Amerikából induló utasszállító esetében van, ez 0,000026 százalék. Ez azt jelenti, hogy várhatóan csak minden 3.880.000-edik úton lesz halálos áldozattal járó baleset.
- A legrosszabb biztonság az Afrikából induló teherszállítóknál jelentkezik, méghozzá 0,00184 százalékos valószínűséggel. Ugyan még ez is nagyon alacsony szám, de arányaiban már rossznak számít, hogy nagyjából minden 55.000 repülésből lesz egy, amely halálos áldozatot követel.

A könnyebb átláthatóság érdekében a teljes eredményt a 2. ábrán mutatom be. Ez tartalmazza kontinens és típus szerint (utasszállító, teherszállító, illetve összevontan) a 10 millió repülőjáratra jutó várható halálos szerencsétlenségek számát.

2. ábra: 10.000.000 útra jutó halálos kimenetelű balesetek várható száma típus és kontinens szerint



Forrás: saját gyűjtés

### Kárnagyságok vizsgálata

A kárnagyságot befolyásoló tényezők vizsgálatához az aktuáriusi munka egy kedvelt eszközkészletét, az általánosított lineáris modelleket (GLM) használtam. A GLM lényege a hagyományos lineáris modellekhez képest abban áll, hogy annak bizonyos megkötéseit el lehet hagyni. Ezek közül a legfontosabbak azok, ahol nemcsak normális

eloszlást lehet feltételezni a magyarázott változónál, hanem bármilyen más eloszlást az exponenciális családból, illetve, hogy nem szükségszerűen lineáris a kapcsolat az eredményváltozó várható értéke és a magyarázó változók között (Kaas, 2008).

Szemléletesen azt lehet mondani, hogy a GLM lényege, hogy – adott feltevés mellett a valószínűségi változó eloszlására – az eloszlás várható értékének valamilyen transzformáltjára illesztünk egy hagyományos lineáris modellt. Leggyakrabban ez a transzformált a logaritmus, így a modell a függő változó szintjén gyakorlatilag multiplikatív lesz. Több paraméterrel rendelkező eloszlás esetén a kettő között a modell valamilyen (lineáris, négyzetes...) kapcsolatot tételez fel. Ezek a tulajdonságok alkalmassá teszik arra, hogy a vizsgált változók eloszlását modellezem.

### Utassok száma

Mivel az utassok, illetve az áldozatok számáról nem állt rendelkezésemre adat minden repülőútra vonatkozóan, csupán azokra, amikor volt halálos áldozat, ezért a becsléshez csak a halálos kimenetelű utak adatbázisában szereplő adatokat használtam. Ezáltal egy feltételes eloszlást tudtam csak vizsgálni (arra a feltételre nézve, hogy volt halálos áldozat), de feltehetően ez nem okoz problémát az eredmények szempontjából.

A GLM modellekben a magyarázó változók mellett a vizsgált valószínűségi változó eloszlását illetően is szükség van valamilyen feltételezésre. Az utassok számára a Poisson-eloszlás jó választásnak látszik, hiszen csak nemnegatív értékeket vehet fel, és a hisztogramok is megfelelnek ennek. Mivel a különböző géptípusok (és akár kiindulási kontinensek is) más-más várható értékűek, így különbözik a paraméterük, érdemes ezeket is kipróbálni magyarázó változóként.

A modellillesztésre az *R glm* függvényét alkalmaztam. Ez a függvény a megadott paraméterek segítségével (magyarázó változók, interakciók és eloszlás) elvégzi a maximum likelihood becslést.

Az így kapott modell: konstans nélkül; a repülő tevékenységi köre, kiindulási kontinense és interkontinentális volta alapján becsült paramétereket, egy olyan lineáris modellre, melynek magyarázott változója az utassok számát jellemző Poisson-eloszlás paramétere. Mivel a magyarázó változók mind kategorikus változók, azaz kategóriákat jelölnek, gyakorlatilag a hozzájuk rendelt dummy (0;1) változókkal kell szorozni a megadott paramétereket, majd ezeket összegezni. Végül mivel a Poisson-eloszláshoz alapbeállításként a logaritmus link-függvény tartozik, ezért a kapott eredményt az exponenciális függvény hatványaként kell felhasználni a Poisson-paraméter becsült értékének meghatározásához. A modell paraméteres felírását ebben az esetben a következőképpen adhatjuk meg:

$$\lambda_{\text{utass}} = e^n = e^{K\beta_{KU} + U\beta_U + T\beta_T + A\beta_{AF} + EA\beta_{EA} + EU\beta_{EU} + LA\beta_{LA} + n\beta_{nIC}} \quad (1)$$

ahol  $KU$ =kis utas,  $U$ =utas,  $T$ =teher,  $AF$ =Afrika,  $EA$ =Észak-Amerika,  $EU$ =Európa,  $LA$ =Latin-Amerika,  $AA$ = Ázsia és Ausztrália és  $nIC$ = nem interkontinentális a megfelelő dummy változók, melyek értéke 1, ha az adott repülő abba a kategóriába tartozott, különben 0.

A modellszelekció fontos lépés a GLM alkalmazásánál, így teszteltem több különböző modellt is. A tesztelés során egyaránt néztem szűkebb modelleket (csak interkontinentális és/vagy típus vagy kontinens elhagyása) és bővebbet is (típus és kontinens interakciójának bevonása), továbbá külön vizsgáltam azt a modellt is, amikor a két utasszállító kategóriát egyként kezeltem. Végül a korábban bemutatott modell bizonyult a legjobbnak, így a későbbi szimulációkhoz is ennek az eredményeit használtam.

#### Halálos áldozatok száma az utasok között

Az utasok közül életüket veszítők számának modellezésére kézenfekvő választás egy binomiális modell. Ennek oka, hogy csak nemnegatív egész szám lehet, amelyet felülről természetesen korlátoz a gépen utazók száma.

A modellszelekciót itt is az előző részben leírt elvek alapján hajtottam végre. Bár ennek alapján egy olyan modell bizonyult a legjobbnak, amelyben a magyarázó változók között az utasok száma is szerepelt, ezt végül elutasítottam, tekintettel arra, hogy ez a függő változó oszlopainak lineáris kombinációjaként előállítható. Végeredményben így azt a modellt választottam, amelyben a rögzített számú utas mellett a binomiális eloszlás  $p$  paraméterének becsléséhez a lineáris prediktorban a kontinens és az interkontinentális változók szerepeltek.

A glm függvény alapbeállítás szerinti link függvénye binomiális eloszlás illesztése esetén a logit függvény, azaz a becsült modell:

$$p = \frac{\exp(\eta)}{1 + \exp(\eta)} = \frac{\exp(AA\beta_{AA} + AF\beta_{AF} + EA\beta_{EA} + EU\beta_{EU} + LA\beta_{LA} + nIC\beta_{nIC})}{1 + \exp(AA\beta_{AA} + AF\beta_{AF} + EA\beta_{EA} + EU\beta_{EU} + LA\beta_{LA} + nIC\beta_{nIC})} \quad (2)$$

#### Földi áldozatok száma

A földi halálos áldozatok számának modellezéséhez az előző kettőtől eltérő modellt kellett alkalmazni. Ennek az eloszlásnak a tulajdonságai hasonlóak az utasszám eloszlásának jellemzőihez, de az, hogy a halálos kimenetelű balesetek jelentős részénél (közel 95%) nem volt földi áldozat, megnehezíti egy egyszerű Poisson-eloszlás alkalmazását.

Ezek alapján megállapíthatjuk, hogy olyan eloszlásra van szükség, amely pozitív súlyt helyez a 0-ra, mintha diszkrét lenne, és folytonos eloszlású a pozitív értékeken. Szerencsére ilyen eloszlásra gyakran van szükségük az aktuáriusoknak, így létezik rá megoldás. Ezt az eloszlást Tweedie-eloszlásnak nevezzük, amely gyakorlatilag az összetett Poisson-eloszlás egy speciális esete. A Tweedie-eloszlás gyakorlatilag több eloszlást foglal magában, bizonyos paraméterek mellett kiadja a normális, a

gamma- vagy a Poisson-eloszlást. Jelen helyzetben a lehetőségek közül az összetett Poisson-eloszlásra van szükség. Egy  $Y$  valószínűségi változónak akkor ez az eloszlása, ha felírható

$$Y = \sum_{i=1}^N X_i \quad (3)$$

formában, ahol  $N$  Poisson-eloszlású, az  $X_i$ -k pedig független, gamma-eloszlású valószínűségi változók azonos paraméterekkel.

A Tweedie-eloszlás illesztése valamivel bonyolultabb egy GLM keretein belül, mint például a Poisson-paraméter becslése, ugyanis szükség van egy extra paraméterre, amely a variancia és a várható érték közötti kapcsolatot írja le.<sup>1</sup>

A Tweedie-eloszlással rendelkező valószínűségi változóra illesztett GLM értelmezése lényegesen bonyolultabb, mint az előző alfejezetekben látott példákban volt. Ennek ellenére ezt is meg lehet tenni, mégpedig a következőképpen (Generate Tweedie random deviates): egy Tweedie valószínűségi változó  $1 < p < 2$  esetén nem más, mint  $N$  darab független azonos, gamma-eloszlású valószínűségi változó ( $X$ ) összege, ahol  $N$  Poisson-eloszlású. Az eloszlások paraméterezése:

$$N \sim \text{Poisson}\left(\frac{\mu^{2-p}}{(2-p)\phi}\right) \quad (4)$$

$$X \sim \text{Gamma}\left(\frac{2-p}{p-1}; \phi(p-1)\mu^{p-1}\right) \quad (5)$$

ahol  $\mu$  a lineáris prediktor, és a link függvény alapján adódik a becsült paraméterek segítségével:

$$\mu = e^\eta = e^{KU\beta_{KU} + U\beta_U + T\beta_T} \quad (6)$$

Ezekre a képletekre a későbbiekben, a szimulációkban is szükség lesz.

#### Szimuláció

A legtöbb esetben a biztosító számára nemcsak a kárgyakoriság vagy kárnagyság a fontos kérdés, hanem a teljes portfólió összesített kára, azaz gyakorlatilag az egyedi károk összege. Ez a valószínűségi változó úgy adódik, hogy az ismeretlen számú, ismeretlen nagyságú kárt összegezzük. Mivel az összes kár ilyen módon matematikailag könnyen előállítható, és a mögötte lévő változók egyes tulajdonságai ismertek, van esély az eloszlásának meghatározására. Sok esetben, amikor a kárgyakoriság és a kárnagyság eloszlása nem túl bonyolult, ez elméleti úton is megtehető. Ezt a módszert hívhatjuk az összetett kockázat modelljének, hiszen gyakorlatilag egy összetett valószínűségi változó eloszlását kívánjuk meghatározni.

Amennyiben ismert a teljes kárkifizetés eloszlása, könnyen alkalmazható rá bármilyen szimulációs eljárás, hiszen csak ennek az eloszlásnak megfelelő véletlen értékeket kell generálni. A szimuláció azért hasznos akár ismert eloszlás esetén is, mert nem minden eloszlásnál lehet könnyedén, csupán az eloszlásfüggvény és a paraméterek alapján meghatározni olyan kockázati mértékeket, mint a VaR vagy a TVaR.

Matematikailag ugyan nem túl bonyolult levezetni az összesített kárkifizetés sűrűségfüggvényét, de ez nem karakterizál semmilyen ismert eloszlást sem. Így – noha numerikusan kiszámolhatóak lennének a sűrűségfüggvény értékei – az általam vizsgálni kívánt kockázati mértékek meghatározására nem alkalmas ez a megközelítés. Ennek a problémának a kiküszöbölése érdekében használtam az úgynevezett Monte-Carlo eljárást.

#### Monte-Carlo szimuláció

Az aggregált káreloszlás vizsgálatára a legelterjedtebb módszer a Monte-Carlo szimuláció alkalmazása (Glasserman, 2004). A módszer lényege abban áll, hogy az ismert eloszlások alapján generálunk valószínűségi változókat, egymás után többször. Így gyakorlatilag a véletlen folyamatokra kapunk több realizációt, amelyekből utána már könnyedén kalkulálhatjuk a vizsgált változó bizonyos tulajdonságait. A legkézenfekvőbb példa az, amikor az aggregált kár várható értékét akarjuk ezzel meghatározni. Ebben az esetben létrehozunk kellően sok scenáriót, melyekben ugyanazzal a módszerrel egyenként szimuláljuk az aggregált kárt, majd az így kapott eredményeket átlagoljuk. Amennyiben például az eloszlás valamely kvantilisére (azaz Value-at-Risk-jére) vagyunk kíváncsiak, akkor a szimulált értékek megfelelő (tapasztalati) kvantilise lesz a becslés. A módszer előnye, hogy könnyen kivitelezhető és jól interpretálható. Fontos tulajdonsága, hogy a scenáriók számának növelésével csökken a becslés hibája, így megfelelő számítási kapacitás mellett pontos eredményeket kaphatunk.

### A scenáriók számának növelésével csökken a becslés hibája.

Megvizsgálva az összesített kárkifizetés eloszlását, azt az eredményt kaptam, hogy ezt a valószínűségi változót nem tudom könnyen közvetlenül szimulálni. Ennek következtében azt a megoldást választottam, hogy a kárkifizetés minden összetevőjét külön szimuláltam minden repülőjárat esetén.

A járatok egyes típusok szerinti megoszlását a tapasztalati arányok alapján állítottam be. A kiindulási kontinensek szerinti megoszlásra volt adatom, így ezt közvetlenül tudtam alkalmazni. Ezekon kívül volt még adatom külső adatforrásból a teherszállítók összes járaton belüli arányára. A kis- és nagy utasszállítók arányára nem volt külön adatom, így csak a saját gyűjtésem (melyben csak a halálos kimenetelű balesettel járó utak voltak) belüli arányokat tudtam használni. Annak érdekében, hogy ez konzisztens

legyen a teherszállítók arányával, a 2012–2017 közötti időszak átlagát vettem. Az interkontinentális járatok arányára ismét csak a saját adatbázisomat tudtam használni, itt az összes (2000–2017 közötti időszak) megfigyelésen belüli arány alapján modelleztem.

A halálos kimenetelű baleset bekövetkezését jellemző változó csupán két értéket vehet fel:  $1$ , ha a repülőjárat során volt halálos áldozat, és  $0$  különben. A különböző specifikációk esetén az  $1$  érték előfordulási valószínűsége különbözik. Más-más valószínűsége van a halálos szerencsétlenségnek az egyes kontinenseken, illetve teher- vagy utasszállító esetén (itt nincs különbség a kis és nagy utasszállító között). Minden járat esetén külön szimuláltam egy standard egyenletes eloszlású valószínűségi változót, majd ezt hasonlítottam a megfelelő valószínűséghez (Vékás, 2012). Amennyiben a valószínűségnél kisebb számot kaptam, akkor az érték egy lett, azaz valamilyen okból kifolyólag a járaton vagy a gép miatt a földön legalább egy személy az életét veszítette.

Mivel az előző fejezetben sikerült mindhárom, a kárnagyságot befolyásoló változónak (utasok, halottak és földi halottak száma) meghatározni az eloszlását, ezért csak az ott kapott eredményeket kellett felhasználni. Első lépésként az utasok számát szimuláltam: a Poisson-eloszlás inverz eloszlásfüggvényébe tettem a standard egyenletes eloszlásból szimulált valószínűségi változót, így kapva Poisson-eloszlásút ( $P(F^{-1}(U) \leq x) = P(U \leq F(x))$ ) (Vékás, 2012)). A Poisson-eloszlás paramétere minden esetben a lineáris prediktor és a link függvény alapján számolható, így azt közvetlenül befolyásolták a járat paraméterei.

A fedélzeten utazók közül életüket veszítők száma binomiális eloszlást követ, feltéve, hogy az utasok száma ismert. Mivel az első lépésben ezt határoztam meg, így ez adott volt, tehát ténylegesen lehetett a binomiális eloszlásból szimulálni. Hasonlóan, mint az előbb, itt is az inverz eloszlásfüggvény segítségével hoztam létre a kívánt eloszlású változókat. Az eloszlás  $p$  paramétere a GLM becsült együtthatói és a logit link függvény segítségével jött ki.

A halálos áldozatok számához szükséges adatok közül a földi áldozatok számának szimulálása volt a legösszetettebb feladat. A (4)-(6) képletek segítségével ez is kivitelezhető feladat az Excelben. Az inverz függvények segítségével szimuláltam Poisson- és gamma-eloszlású valószínűségi változókat. Majd ezeket adtam össze megfelelően.

Mivel azzal a feltételezéssel éltem, hogy kárkifizetés csak a halálos áldozatok után jár, esetükben pedig egységesen azonos összeg, így az aggregált kár meghatározásához gyakorlatilag csak az áldozatok számát (ami a gépen és a földön életüket veszítők összegeként adódott) kellett egy biztosítási összeggel szorozni. A 3. ábrán az látható, hogy miként néztek ki a szimulációk az Excelben (a fontosságot kifejező mintavételezéshez (Importance Sampling) és a súlyhoz tartozó oszlopok jelentését és szerepét később ismertetem).

3. ábra: Egy realizáció a Monte-Carlo szimulációban, két szerződés mellett

#	Típus	Kontinens	Interkont.	Véletlen	Imp. Sampl.	Halálos áld.	Utaz	Halott	Földi halott	Össz halott	Kifizetés	Súly
$\Sigma$	2			0,33	0,44	0,0000	20	14	0	0,00000	0,15	1,1445
1	Kis utas	EA	nem IC	0,05571638	0,00000001	1	20	14	0	14	140 000	0,00000107
2	Utas	AA	nem IC	0,59821611	0,88004893	0				0		2,28898131

Forrás: saját gyűjtés

Első megközelítésben megvizsgáltam, hogy miként alakulhat a repülőgép-katasztrófákban elhunytak biztosítása. Pontosabban azt állítottam a középpontba, hogy milyen díjak adódnak különböző díjelvek (várható érték, kvantilis vagy szórás elv alapján). Ehhez alapvetően arra volt szükségem, hogy a kárkifizetés eloszlását megismerjem, és meghatározzam ennek várható értékét, szórását és VaR értékét. Ezekon kívül, minthogy sok tulajdonságát tekintve a TVaR alkalmasabb kockázati mérték lehet a VaR-nál, ezt is szerettem volna meghatározni.

Ezen mutatók értékei mellett egy másik kérdésre is kerestem a választ. A szimuláció során egyenként hoztam létre a repülőjáratokat minden tulajdonságukkal együtt, viszont egyelőre azt nem részleteztem, hogy hány ilyen elem szerepelt a szimulációban. Egyes statisztikák, mint az összesített kárkifizetés várható értéke (egy járatra jutó esetben) várhatóan nem függ a portfólió nagyságától, azaz az egy szimulációs lépésben összegzett járatok számától, de ez nem mindenre igaz. Például a VaR egyik fontos tulajdonsága, hogy adott esetben lehet akár szubadditív vagy szuperadditív is. Éppen ezek miatt mindenképp vizsgálni akartam a portfólióméret növelésének hatását a kockázati mértékekre, illetve a díjelvekre. Ez azért fontos kérdés, mert például egy olyan esetben, amikor a nagyobb portfólió csökkenti az egy szerződésre jutó szórást, akkor egy szórásalapú díjelv mellett a nagyobb méret csökkentheti a szerződés díját, ami mind a biztosító, mind a biztosítottak számára előnyös lehet.

*A hagyományos Monte-Carlo módszer gyengesége és a fontosságot kifejező mintavételezés (Importance Sampling) szerepe*

Ugyan a Monte-Carlo eljárás nagyon hasznos eszköz, de a hatékonyságával kapcsolatban bizonyos esetekben felmerülhetnek kételyek. A leggyakoribb probléma az, hogy az általa nyerhető becslés nem elég pontos, illetve ahhoz, hogy elég pontos legyen az eredmény, nagyon sok scenáriót kell létrehozni. Ennek oka, hogy a becsléseknek gyakorta nagy a varianciája kisebb szimulációs méret esetében. Ennek kiküszöbölésére valamilyen varianciacsökkentő technikát érdemes alkalmazni.

A hagyományos Monte-Carlo módszer önmagában nem elég az általam vizsgált esetben. Ennek oka, hogy azoknak az eseteknek, amelyek befolyásolják a kárkifizetéseket, nagyon alacsony a bekövetkezési valószínűsége. Ha (jelentősen felülbecsülve) azt vesszük, hogy a halálos kimenetelű baleset valószínűsége 0,000001, akkor például

még egy 10.000 elemű portfólió esetén, 1.000 szimuláció mellett is csak várhatóan 10 ilyen következik be. Ez alapvetően nem lenne probléma, ám mivel könnyen előfordulhat, hogy az adott szimulációban csak 9 vagy éppen 12 ilyen lesz, nagy lesz a becslés varianciája.

Ez a gyakorlatban ennél is jelentősebb problémát okoz, tekintettel arra, hogy épp a portfólió méretének hatását szeretném vizsgálni, amihez kisebb méretű veszélyközösségek modellezése is szükség van. Amennyiben például egy csupán 10 járatot tartalmazó portfóliót vizsgálok, már ahhoz is több mint 100.000 scenárió kell, hogy várhatóan legalább egy ilyen szerencsétlenség belekerüljön. Így nyilván az is előfordulhat, hogy a várható érték becslése 0-nak adódik, pedig nyilvánvalóan pozitív kell, hogy legyen. Természetesen a problémát elvileg kezelni lehet a szimulációk számának növelésével, de a hosszabb futási idő miatt ez nem hatékony megoldás. Ahhoz, hogy a variancia jelentősen csökkenjen, akár több nagyságrenddel is szükséges lehet növelni a lépésszámot.

Szerencsére az ilyen esetek kezelésére van kidolgozott módszertan, amelyet könnyen lehet egy Monte-Carlo szimulációba implementálni. A módszer, melyet én is használtam, a fontosságot kifejező mintavételezés (Importance Sampling). Ennek a lényege abban áll, hogy a realizációk generálásához használt valószínűségi mértéket kicseréljük egy olyanra, amelyben a számunkra fontos esetek gyakrabban fordulnak elő.

Tulajdonképpen ennek az eljárásnak a segítségével el lehet érni azt, hogy a szimulációban a halálos balesetek előfordulási gyakorisága megnövekedjen, és ezáltal több adatból kalkuláljuk az olyan mutatókat, mint a várható érték és a szórás. Természetesen az érdekes esetek megnövelt súlya miatt az eredményeket korrigálni kell, hogy a végső becslésekben már az eredeti eloszlásnak megfelelő értékek adódjanak.

Alapesetben az  $E(h(x))=\alpha$  becslése a Monte-Carlo szimuláció segítségével a következőképpen adható meg:

$$\hat{\alpha} = \frac{1}{N} \sum_{i=1}^N h(X_i) \quad (7)$$

ahol  $N$  a szimulációk száma,  $h(X_i)$  pedig az  $i$ -edik szimulációban kapott  $h(X)$  érték. Következő lépésben vegyünk egy másik eloszlást,  $g(X)$  sűrűségfüggvénnyel, ami abszolút folytonos az eredeti eloszlásra nézve (tehát  $f(X)>0 \Rightarrow g(X)>0$ ). Ebben az esetben a megfelelő tartományon integrálva igaz lesz a következő összefüggés:

$$E(h(X)) = \int h(x) \frac{f(x)}{g(x)} g(x) dx = E_Q \left( h(X) \frac{f(X)}{g(X)} \right) \quad (8)$$

Tehát a keresett eredmény felírható az új mérték ( $Q$ ) szerinti várható értéként. Ennek következtében kapjuk a következő képletet:

$$\hat{\alpha}_Q = \frac{1}{N} \sum_{i=1}^N h(X_i) \frac{f(X_i)}{g(X_i)} \quad (9)$$



Ebben az egyenletben az  $X_i$ -k már nem az eredeti, hanem az új mérték szerinti eloszlásból származnak. Jól látható, hogy gyakorlatilag az új eloszlásból generálunk  $X$  értékeket, majd ezeknek a realizációknak a súlyozott átlagát vesszük. A használt  $f(X_i)/g(X_i)$  súlyokat szokás likelihood aránynak vagy Radon-Nikodym deriválnak is hívni. Szemléletesen azt mondhatnánk (bár ez matematikailag nem túl precíz), hogy azt vizsgálják, hogy a realizálódott  $X$  érték hányszor gyakrabban következne be az eredeti mérték esetén, mint az új mérték esetén. Az Importance Sampling jó beállítása esetén ez egynél kisebb a vizsgálat szempontjából fontosabb  $X$ -eknél, tehát azokat gyakrabban kapjuk meg, mint alapesetben.

Bár a tanulmányom központi kérdése a repülőgépjáratokból álló portfólió aggregált kárkifizetéséhez kapcsolódott, tekintettel arra, hogy ennek eloszlása nem volt könnyen felírható, erre nem tudtam az Importance Sampling módszert alkalmazni. Ehelyett azt a megoldást választottam, hogy azt a mértéket cseréltem le, amelyiken a halálos kimenetelű szerencsétlenség bekövetkezése alapult. Eredetileg ez egy egyenes eloszlású valószínűségi változó lett volna a  $[0,1]$  intervallumon, amit egy béta-eloszlással helyettesítettem, ennek hatására az igazán kicsi (egymilliomod nagyságrendű) értékek előfordulási gyakorisága jelentősen megnőtt.

Az összesített kárkifizetés eloszlása tulajdonképpen egy többdimenziós eloszlás, amelyet minden járat kárkifizetése befolyásol. Az ehhez tartozó együttes sűrűségfüggvény meghaladja ennek a tanulmánynak a kereteit, de erre tulajdonképpen nincs is szükség. Ami valójában szükséges, az az eredeti együttes sűrűségfüggvény és az Importance Samplinghez tartozó együttes sűrűségfüggvény hányadosa. Ez viszont – kihasználva az egyedi szerződések egymástól való függetlenségét – felírható a peremsűrűségfüggvények hányadosainak (egyenletes osztva a bétával) szorzataként.

Az így kapott súlyok segítségével már elméleti síkon alkalmazható lenne az Importance Samplinget használó Monte-Carlo eljárás. A megfelelő módszerek segítségével elméletileg ez adna egy torzítatlan és alacsony varianciájú becslést a várható értékre, a szórásra, a VaR-ra és a TVaR-ra. Ezzel szemben, amikor elvégeztem a szimulációt, azt tapasztaltam, hogy az egyes járatokhoz tartozó likelihood arányoknak a várható értéke ugyan 1 volt, de a szorzatuk nagyon alacsony lett. Különösen igaz volt ez nagyobb portfóliók esetén, ahol numerikusan gyakran 0-vá is vált.

Ennek a szituációnak az oka alapvetően abban keresendő, hogy a halálos áldozattal járó balesetek nagyon alacsony bekövetkezését kompenzáló, nagyon erős torzítást kellett alkalmaznom. Továbbá megmutatható (Glasserman, 2004), hogy a nagyobb portfólióméret esetén a  $h(X_i)$  értékekhez tartozó hányadosok tartanak a 0-hoz (bár várható értékük ettől még 1, tehát nem elhanyagolható valószínűséggel vesznek fel extrém magas értékeket).

Mivel ezen a módon nem lehetett elvégezni a modellezést, némiképp máshogy próbáltam az eredményeket megkapni a várható értéket és a szórást illetően. A VaR és TVaR Importance Sampling módszerrel történő meghatározása viszont nem sikerült, így azokra a későbbiekben csak a hagyományos Monte-Carlo szimuláció által létrehozott eredményeket tudom felhasználni.

A repülőgépjáratokból álló portfólió várható teljes (haláleseti) kárkifizetésének meghatározásához a várható érték linearitását használtam fel. Ez azt jelenti, hogy ténylegesen

az aggregált kárkifizetés ( $S$ ) várható értékét venném, ezt komponenseire bontom, azaz külön veszem a várható értékét minden egyes járat (összesen  $P$  darab) kárkifizetésének ( $X_i$ ), majd ezeket összegezem. Az Importance Sampling pedig nem az aggregált kár szintjén, hanem az egyes szerződések szintjén kerül alkalmazásra.

Ugyanezt az elvet alkalmazva sikerült meghatározni a második momentum becslését, amivel már számolható a variancia becslése is. Ezeknek a becsléseknek ugyan nem túl összetett feladat a levezetése, de ezt most nem részletezem, az eredményeket viszont a következő képletekben megadom ( $N$  a szimulációk,  $P$  a portfólió elemeinek száma).

$$\overline{E(S)} = \frac{1}{N} \sum_{j=1}^N \sum_{i=1}^P X_i \frac{f(X_i)}{g(X_i)} \quad (10)$$

$$\overline{E(S^2)} = \frac{1}{N} \sum_{k=1}^N \left( \sum_{i=1}^P X_i^2 \frac{f(X_i)}{g(X_i)} + \sum_{i=1}^P \sum_{j=1, i \neq j}^P X_i X_j \frac{f(X_i)f(X_j)}{g(X_i)g(X_j)} \right) \quad (11)$$

## Eredmények

A tanulmányomban a szimuláció során az aggregált kárösszeg eloszlásának néhány tulajdonságára koncentráltam. Ezek a várható érték, a szórás, a VaR, illetve a TVaR voltak. A várható érték és a szórás számításánál az Importance Samplinget is felhasználtam. A kettő között annyi különbség volt, hogy a szórás becsléséhez szükséges kettős szummázás nagyban növelte a futási időt, így azt csak egy külön – kevesebb scenárióval dolgozó – szimulációban vizsgáltam. A farkeloszlást jellemző kockázati mértékek becslésére csak a hagyományos Monte-Carlo szimuláció eredményei álltak rendelkezésemre.

A portfólióméret növekedésének vizsgálatára a járatok számát 10 és 20.000 között változtattam, növekvő lépésközökkel. Mint majd a későbbi szakaszok ábrái tanúsítják, ezek segítségével elég jól ki lehetett rajzolni az egyes mutatók alakulását. A VaR esetében külön tekintettel kell lenni arra a tényre, hogy kisebb portfólióméret esetén az értéke nem vizsgálható. Ahhoz, hogy minél inkább hasznos mutató legyen, érdemes igen nagy konfidenciaszintet választani. Éppen ezért én a biztosításban bevett 99,5 százalékos, illetve a 99,9 százalékos Value-at-Riskkel és Tail-Value-at-Riskkel dolgoztam.

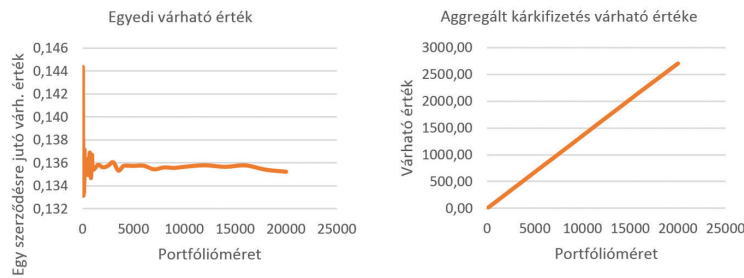
### A kárkifizetéshez kapcsolódó mutatók alakulása a portfólióméret függvényében

Elméleti megfontolások alapján könnyen belátható, hogy az aggregált kárkifizetés egy szerződésre jutó várható értéke nem függ a portfólió méretétől. Ennek oka a várható érték additivitása, valamint az, hogy az egyes szerződések független, azonos eloszlású valószínűségi változók.

Ebből következően nagyon pontosan meg lehet határozni, hogy egy repülőjárat esetén mennyi a kárnagyság várható értéke. Ez a szimulációs eredmények alapján 0,1356, azaz a halálos áldozatok száma szorozva a biztosítási összeggel. Ennél talán hasznosabb mutató az egy repülőút során vár-

ható halálos áldozatok száma, ami 0,00001356, azaz nagyságrendileg, átlagosan minden 75.000-edik járatra jut egy elhunyt személy. Nem véletlen, hogy a repülés az egyik legbiztonságosabb közlekedési forma. A várható érték alakulását a portfólióméret függvényében a 4. ábra mutatja.

4. ábra: Aggregált kárkifizetés egy szerződésre jutó, illetve teljes várható értéke



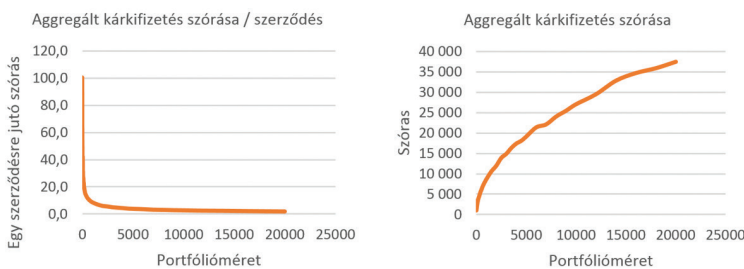
Forrás: saját szerkesztés

Az eredmények a szórás esetében is igazolják a hipotéziseket, miszerint fordítottan arányos a portfólióméret négyzetgyökével (hiszen független valószínűségi változókat összegzünk).

## Átlagosan minden 75 000-edik járatra jut egy elhunyt személy.

Ez a tulajdonság, hogy a veszélyközösség növelésével – állandó egy szerződésre jutó várható érték mellett – az egy szerződésre jutó szórás csökken, nagyon fontos eleme a biztosításoknak. Ennek köszönhetően egy olyan szervezet, amely több egyedi kockázatot fog össze, képes Pareto-javulást elérni egy biztosítás nélküli állapothoz képest. A szórás alakulását az 5. ábra szemlélteti.

5. ábra: Aggregált kárkifizetés egy szerződésre jutó, illetve teljes szórása



Forrás: saját szerkesztés

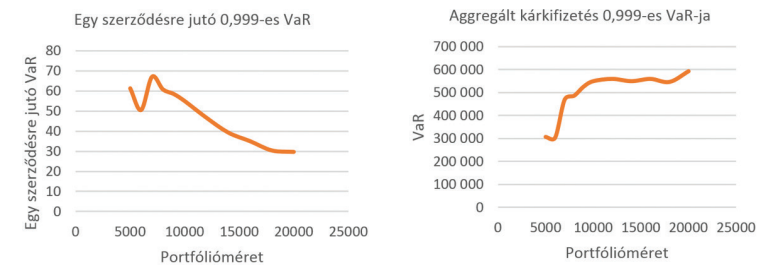
A VaR számításakor a legfontosabb kérdés, hogy adott esetben szubadditív vagy szuperadditív ez a kockázati mérték. A repülőgépjáratok szerencsétlenségei halálos áldozatainak biztosítása esetén

teljesül a szubadditivitás a 0,999-es Value-at-Riskre, ugyanis az összeg esetén kisebb a VaR, mint a VaR-ok összege.

## A veszélyközösség növelésével az egy szerződésre jutó szórás csökken.

Ez ekvivalens azzal, hogy az egy szerződésre jutó értéke a teljes kárösszeg VaR-jának csökken a portfólió növelésének hatására (6. ábra).

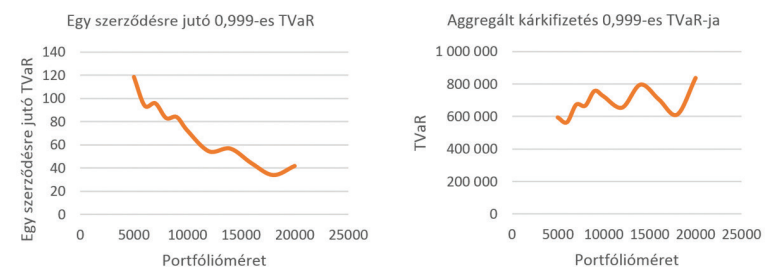
6. ábra: Aggregált kárkifizetés egy szerződésre jutó, illetve teljes Value-at-Riskje



Forrás: saját szerkesztés

Az eredmények a TVaR esetében is a várakozásoknak megfelelően alakultak (7. ábra), továbbá hasonló következtetések olvashatók le az ábrákról, mint a VaR-nál. További hasznos eredmény az, hogy míg a VaR 10-20.000-es portfólióméret esetén 550.000 körül alakult, addig a TVaR körülbelül 700.000-es értékeket mutat (ezek az értékek jelenthetnek dollárt, de alapvetően nincs mértékegységük). Ebből következően, amikor a veszteség a VaR-t meghaladó, akkor sem annak sokszorosa, várhatóan csak 25-30 százalékkal haladja meg. Ez a repülőgépjáratok szerencsétlenségei biztosításának szempontjából megnyugtató eredmény, hiszen egy kiugróan magas várható érték könnyen okozhatná a biztosító csődjét is. Ez, hogy látszólag nem túl vastagfarkú az eloszlás, egybecseng az általánosított Pareto-eloszlás segítségével kapott eredményekkel is.

7. ábra: Aggregált kárkifizetés egy szerződésre jutó, illetve teljes Tail-Value-at-Riskje



Forrás: saját szerkesztés

Biztosítási díjak alakulása a portfólióméret függvényében

A fogyasztók számára adott specifikációjú biztosítás (azonos biztosítási esemény, biztosítási összeg) esetén a legfontosabb kérdés az, hogy mennyi a biztosítás díja. Ebben az alfejezetben néhány nagyon egyszerű díjelv segítségével fogom vizsgálni a biztosítási díjak alakulását különböző portfólióméretetek esetében. A használt díjelveket a következő táblázat tartalmazza, a biztosítás díját pedig  $\pi(X)$  jelöli  $X$  kárösszeg mellett.

2. táblázat: Díjelvek neve és képlete

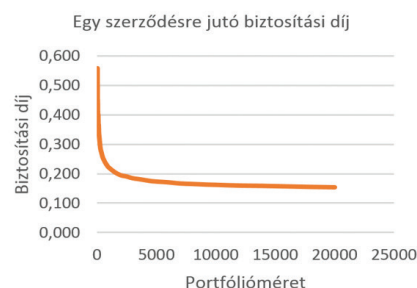
Várható érték elv	$\pi(X) = (1 + \alpha)E(X)$
Szórás elv	$\pi(X) = E(X) + \alpha D(X)$
Kvantilis elv	$\pi(X) = F^{-1}(\alpha) = VaR_{\alpha}$

Forrás: Kaas, 2008

Először vizsgáljuk meg a várható érték elv alapján számított díjat. Ebben az esetben gyakorlatilag azt veszi alapul a biztosító, hogy a kifizetések várható értékét kell, hogy fedezze a díj nettó része, illetve ezen felül van egy ezzel arányos díjrész (költségekre, nagyobb biztonsági szinthez és a profitra). Mivel a repülőgép-szerencsétlenségek áldozatainak fizetendő összeg várható értéke lineárisan aránylik a portfólió méretéhez, itt a mérethozadék konstans. Ennek következtében a biztosítás díja nem csökken a nagyobb portfólió esetén sem.

Mivel az egy szerződésre jutó szórás csökken a portfólióban lévő szerződések számának növelésével, ezért egy szórásalapú díjszabás esetén nagyobb portfólió mellett csökkenthető lehet a szerződés díja. Tételezzük fel, hogy egy biztosító az  $\alpha$  értékét 0,01-nek határozza meg. Továbbá egy szerződés várható kárkifizetése 0,1356, az aggregált összeg szórása pedig a szimulációban kapott eredmények alapján alakul. Az így kapott díjak alakulását a 8. ábra mutatja.

8. ábra: Egy szerződés biztosítási díja szórás díjlevvel  $\alpha = 0,01$  mellett



Forrás: saját szerkesztés

Erről a 8. grafikonról néhány igen fontos következtetést vonhatunk le. Egyrészt nyilvánvaló, hogy a biztosítási díja egy repülőjáratnak jelentősen csökken a nagyobb portfóliók létrehozásával. Ebből az következik, hogy amennyiben a biztosítónak a szerződésenkénti költsége konstans (vagy csökken), akkor nagyobb portfólió Pareto-javulást eredményezhet, ugyanis meghatározható olyan díj, amellyel a szerződésenkénti díj csökken, az összes bevétel pedig nő.

Másik fontos tanulsága az eredményeknek, hogy bizonyos biztosított magatartást feltételezve van egy minimális szükséges portfólióméret, amely nélkül a termék nem eladható. Tételezzük fel azt, hogy a légitársaságok kellően pontosan ismerik a légikatasztrófák káralakulását. Ezen kívül van egy korlát arra vonatkozóan, hogy milyen díj mellett választják inkább a biztosítást a saját megtartás helyett. Tegyük fel továbbá az egyszerűség kedvéért, hogy ez azt jelenti, hogy legfeljebb a kárösszeg várható értékének kétszeresét hajlandók biztosítási díjként kifizetni.

**Meghatározható olyan díj, amellyel a szerződésenkénti díj csökken, az összes bevétel pedig nő.**

A 3. táblázat alapján egyértelműen megállapítható, hogy az előző bekezdésben leírt viselkedésű légitársaságok esetén csak akkor lehet a biztosítást eladni, ha a portfólióban legalább 400 repülőjárat lesz. Továbbá érdemes megfigyelni, hogy – főleg kis veszélyközösségek esetén – a portfólió növelésének nagy a hozadéka a díj csökkenésében, de még a nagyobb portfóliók esetén is lehet a díjat szignifikánsan csökkenteni.

3. táblázat: Biztosítási díj a várható érték százalékában különböző portfólió nagyságok mellett

Szerződés (db)	10	50	100	200	400	700	1.000	5.000	20.000
Biztosítási díj	840 %	375 %	282 %	243 %	200 %	176 %	164 %	128 %	114 %

Forrás: saját szerkesztés

A kvantilisalapú árazás vizsgálatához a 6. ábra első grafikonját kell megnézni. Ez ugyanis egy  $\alpha = 0,999$  melletti VaR díjlevvel meghatározott biztosítási díj egy szerződésre jutó részét mutatja a portfólióméret függvényében. A megállapítások hasonlóak, mint a szórás elv esetén, könnyen csökkenthető egy járat biztosításának a díja, ha nagyobb a portfólió. Ami érdekes a szórásalapú díjjal szemben, az az, hogy itt lényegesen nagyobb díjak adódnak. Egyrészt ez függ az árazási paraméterek megválasztásától, de van technikai oka is. Ez abban áll, hogy kellően nagy portfólió esetén ez esetek egy ezrelékében már várható, hogy bekövetkezik a biztosítási esemény, így a VaR érték gyakorlatilag egy ténylegesen realizált aggregált kárral egyezik meg. Ha csak egyetlen járaton is volt halálos áldozat ekkor, és ott se sok, akkor is lényegesen nagyobb a kárkifizetés, mint annak a várható értéke (vagy szórása). Éppen ezért fontos a VaR és TVaR a biztosítók számára, mert

magas biztonsági szinteknek felelnek meg (például itt csak minden 1000-edik portfólió esetén lesz az aggregált VaR-nál nagyobb az összes kárkifizetés).

Több további elemzési lehetőség is felmerülhet a tanulmányban vizsgált területek mellett. Egyrészt érdemes lehet pontosabb becslés készítése a VaR és TVaR értékeket illetően, illetve még nagyobb portfólióméretekre is annak érdekében, hogy jobb képet kapjunk arról, miként alakulnak ezek a kockázati mértékek a szerződések számának növelésével. Másik érdekes kérdés, hogy a biztosítási díjak, illetve a kockázati mértékek hogyan viselkednek, ha az eredeti portfólió helyett például egy kockázatosabbat veszünk. Ennek azért is lehet kiemelt jelentősége, mert az utóbbi évek trendjei alapján a repülőjáratoknak egyre nagyobb része közlekedik Afrikában és Ázsiában (IATA Economics, b), ahol a baleseteknek relatíve magas a kockázata. Ennek akár az is lehet a következménye, hogy a kárgyakoriság csökkenő trendje valamikor a következő évtizedekben megtorpan vagy akár meg is fordul, ami mindenképpen fontos tényező lehet a repülési iparágban.

### Összegzés

A cikk alapjául szolgáló szakdolgozat témájának igyekeztem olyan területet választani, amely az aktuáriusi tevékenység szempontjából jelentős, ugyanakkor a róla szóló ismereteim valamilyen szempontból hiányosak. Miután eldöntöttem, hogy módszertanként a mikroszimulációs modellezést szeretném alkalmazni, adódott az ötlet, hogy a veszélyközösségek nagyságának hatásait vizsgáljam. Ez a terület a közeljövőben egyébként is komoly érdeklődésre tarthat számot, hiszen a technológiák fejlődésével egyre inkább igény lehet kisebb (akár önszerveződő) biztosítási hálózatokra is. Hogy érdekesebbé tegyem az amúgy jelentős részben matematikai-statisztikai témát, a modellezéshez egy mindenki számára közérthető – és kiemelt jelentőségű – területet választottam: a légi közlekedési biztosításokat.

Elsősorban azt igyekeztem bemutatni, hogy miért lehet érdekes az alacsony kárgyakoriság következtében ennek a megközelítésnek az alkalmazása a repülőgép-balesetek vizsgálatában. Az összegyűjtött adatok alapszintű vizsgálata, illetve az extrémérték elmélet alkalmazása éppúgy hasznos eredményeket adott, mint a Monte-Carlo eljárással végzett szimulációk. Sikertelen megmutatni, hogy milyen alacsony a halálos kimenetelű baleset valószínűsége, illetve azt, hogy még egy ilyen esemény bekövetkezése esetén se túlzottan nagy az áldozatszám.

A repülőgép-szerencsétlenségek vizsgálata után, azok adatait felhasználva bevezettem egy általánosabb módszertant, amelynek a segítségével bármilyen biztosítási ágazatnál elemezni lehet a portfólió nagyságának hatását különböző kockázati mértékekre. Ezek eredményeként bemutattam, hogy a biztosítási díj hogyan változik, ha nő a portfólió mérete, illetve azt, hogy speciális feltételek mellett létezhet minimálisan szükséges portfólióméret is.

Ugyan több különböző kérdéssel foglalkoztam, mégis a terület szerteágazósága miatt rengeteg nyitott kérdés marad, amelyek vizsgálatára a területi korlátok nem adtak lehetőséget. Ennek ellenére úgy gondolom, hogy a két központi témát (a repülőjáratok biztonságossága, illetve a portfólió nagyságának hatása a kockázati mértékekre) sikerült teljeskörűen bemutatnom, miközben törekedtem arra, hogy a modell a lehető legjobban tükrözze a valóságot, és akár más területeken is lehessen alkalmazni.

Összességében elmondható, hogy a bemutatott módszer más biztosítási kérdések elemzésében is alkalmazható, továbbá a kapott eredmények a valós életben is hasznos jelentéssel bírhatnak.

## HIVATKOZÁSOK

<sup>1</sup>Ez az  $R$  glm függvényében egy input változó, de például a fishMod bővítményben van egy tglm függvény, ami ezt a  $p$ -t is becsüli a maximum likelihood eljárás segítségével. En ez utóbbi függvényt használtam, itt továbbá egy  $\Phi$  paraméter is adódik, amire a későbbiekben a szimulációhoz szükség is lesz.

## IRODALOMJEGYZÉK

- ATAG (2014), Aviation Benefits Beyond Borders  
 ATAG (2016), Aviation Benefits Beyond Borders  
 Bureau of Transportation Statistics. U.S Air Carrier Traffic Statistics. Online: <https://www.transtats.bts.gov/TRAFFIC/>.  
 Letöltés dátuma: 2018.02.27.  
 Generate Tweedie random deviates. Online: <https://stat.ethz.ch/R-manual/R-devel/library/mgcv/html/rTweedie.html>.  
 Letöltés dátuma: 2018.10.15.  
 IATA Economics' Chart of the Week. Aviation's centre of gravity to continue to shift eastward.  
 Online: <http://www.iata.org/publications/economics/Reports/chart-of-the-week/chart-of-the-week-22-Apr-2016.pdf>.  
 Letöltés dátuma: 2018.04.24.)  
 IATA Economics' Chart of the Week. Flying is by far the safest form of transport.  
 Online: <http://www.iata.org/publications/economics/Reports/chart-of-the-week/chart-of-the-week-23-feb-2018.pdf>.  
 Letöltés dátuma: 2018.04.24.  
 Paul Glasserman (2004): Monte Carlo Methods in Financial Engineering, New York: Springer, <https://doi.org/10.1007/978-0-387-21617-1>  
 plane crashinfo.com. Accident Database. Online: <http://www.plane crashinfo.com/database.htm>. Letöltés dátuma: 2018.02.15.  
 Rob Kaas et al. (2008): Modern Actuarial Risk Theory, Berlin: Springer. <https://doi.org/10.1007/978-3-540-70998-5>  
 Vékás Péter (2012): Összefüggő biztosítási kockázatok modellezése, Oktatási segédlet.