Tartu Ülikool
Loodus- ja täppisteaduste valdkond
Matemaatika ja statistika instituut

Kelli Kukk

Kindlustuskahjude modelleerimine Balti riikides

Matemaatilise statistika eriala
Bakalaureusetöö (9 EAP)

Juhendaja: vanemteadur Tõnu Kollo

Tartu 2020

Kindlustuskahjude modelleerimine Balti riikides

Bakalaureusetöö

Kelli Kukk

Käesoleva bakalaureusetöö eesmärgiks on uurida kindlustusettevõtte kaskokahjude jaotust kolmes Balti riigis ja tuua välja võimalikud erinevused. Esmalt antakse ülevaade töös kasutatavatest jaotustest. Praktilises osas kirjeldatakse andmete struktuuri ning koostatakse ja võrreldakse mudeleid. Töö lõpus tuuakse välja kokkuvõtlikud järeldused.

Märksõnad: Gammajaotus, log-normaalne jaotus, Pareto jaotus, Weibulli jaotus, Waldi jaotus, modelleerimine

CERCS teaduseriala: P160 Statistika, operatsioonanalüüs, programmeerimine, finants- ja kindlustusmatemaatika

Modelling non-life insurance claims in Baltic countries

Bachelor's thesis

Kelli Kukk

The aim of this thesis is to study the distribution of Motor Own Damage claims of an insurance company in Baltic countries and discover possible differences. Firstly, the used distributions are introduced. In the practical part data structures are described and models are composed and compared. Finally some conclusions are made.

Keywords: Gamma distribution, lognormal distribution, Pareto distribution, Weibull distribution, Wald distribution, modelling

CERCS research specialisation: P160 Statistics, operation research, programming, actuarial mathematics

Sisukord

Sissejuhatus

Bakalaureusetöö eesmärgiks on uurida kindlustusettevõtte kaskokahjude jaotust kolmes Balti riigis ja tuua välja erinevused.

Esimeses peatükis anname ülevaate töös kasutavatest jaotustest. Jaotuste lõikes toome välja tihedusfunktsiooni, momendid ja jaotuse kuju iseloomustavad kujuparameetrid. Lisaks hindame parameetreid momentide või suurima tõepära meetodil.

Teises peatükis kirjeldame andmete struktuuri tervikuna ja kolme riigi lõikes. Iga riigi andmete puhul toome välja empiirilised arvkarakteristikud, kvantiilid ja histogrammi.

Kolmandas peatükis koostame ja võrdleme mudeleid. Iga riigi andmete puhul toome välja sobitatud jaotuste parameetrid ja Kolmogorov-Smirnovi teststatistiku väärtused. Lisaks kujutame joonistel empiirilisele jaotusele sobitatud jaotusi iga riigi korral ning võrdleme riikide lõikes ja riigiti saadud jaotusi.

Bakalaureusetöö analüüsi teostame ja joonised koostame statistikatarkvaras R.

1 Ülevaade kasutatud jaotustest

Järgnev peatükk põhineb raamatul (Forbes jt, 2011), kui ei ole viidatud teisiti.

Kindlustuskahjud on positiivsed reaalarvud ja seetõttu vaatame jaotusi, mis on defineeritud positiivsel reaaltelje osal. Vaadeldavate jaotuste parameetrite hindamisel lähtume valimist mahuga n: $x_1, x_2, \ldots, x_n$, valimi keskmisega $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ ja valimi dispersiooniga $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$.

Kuna töö käigus selgus, et Weibulli jaotuse ühe parameetri hinnang ei olnud raamatus Forbes jt (2011) korrektne, siis esitame töös kõigi kasutatud jaotuste parameetrite hinnangud koos tuletuskäiguga.

1.1 Gammajaotus

Gammajaotus sisaldab erijuhtudena χ^2 jaotust, eksponentjaotust ja Erlangi jaotust. Gammajaotus on kaheparameetriline jaotus, mille tõenäosustiheduse kuju varieerub sõltuvalt parameetrite väärtustest ning gammajaotuse parameetreid on lihtne hinnata momentide meetodil. Gammajaotuse tähistame $\Gamma(b, c)$, kus skaalaparameeter b ja kujuparameeter c on positiivsed reaalarvud, $b, c > 0$.
Gammajaotuse tihedusfunktsioon on järgmine (Forbes jt. 2011, lk 109):

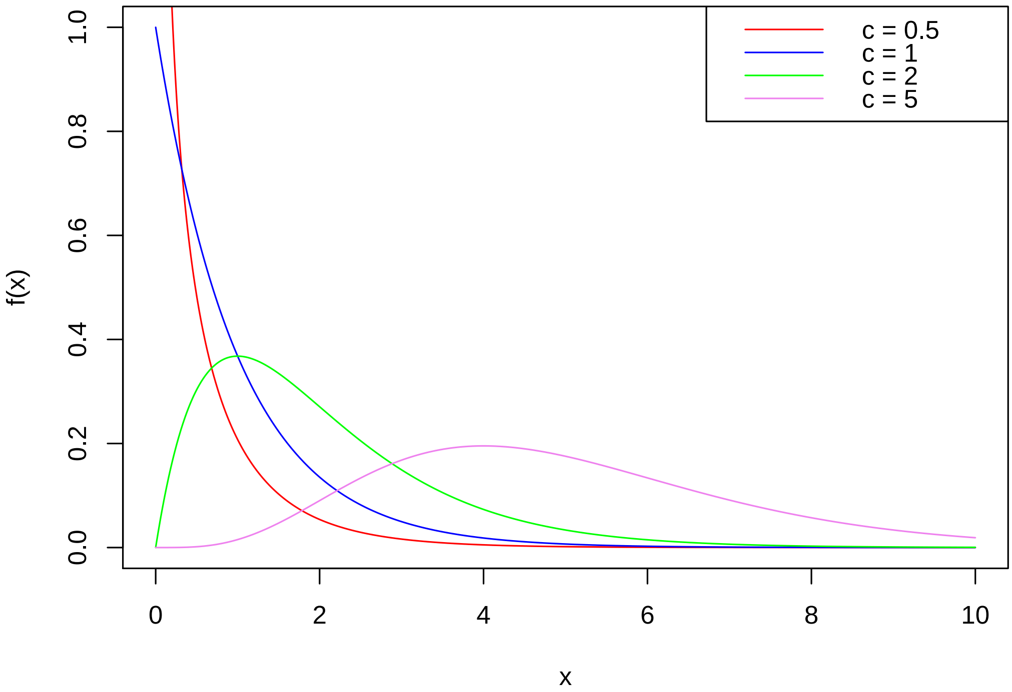
$$f(x) = \left(\frac{x}{b}\right)^{c-1} \frac{\exp\left(\frac{-x}{b}\right)}{b\Gamma(c)}, \quad x \geq 0,$$

kus gammafunktsioon $\Gamma(c) = \int_0^\infty \exp{(-u)} u^{c-1} du$.

On olemas ka alternatiivne parametrisatsioon, kus b on asendatud parameetriga λ, $\lambda = \frac{1}{b}$.

Gammajaotusega juhusliku suuruse X arvkarakteristikud:

- keskväärtus: $EX = bc$,
- dispersioon: $DX = b^2c$,
- k-s moment: $m_k = \frac{b^k\Gamma(c+k)}{\Gamma(c)}$,
- asümmeetriakordaja: $\beta_1 = \frac{E(X-EX)^3}{(DX)^{\frac{3}{2}}} = \frac{2}{\sqrt{c}}$,
- järsakuskordaja: $\beta_2 = \frac{E(X-EX)^4}{DX^2} = 3 + \frac{6}{c}$.



Joonis 1: Gammajaotuse tihedusfunktsiooni kuju sõltuvalt kujuparameetrist c ($b = 1$).

1.1.1 Parameetrite hindamine

Parameertite hindamisel lähtume valimist ning kasutame momentide meetodit. Momentide meetodi korral lähtume võrdustest

$$EX = \bar{x};$$

$$DX = s^2.$$

Saame järgmised seosed:

$$\bar{x} = bc;$$
$$s^2 = b^2 c = b\bar{x}.$$

Eelnevate seoste põhjal saame avaldada b ja c hinnangud.

$$\widehat{b} = \frac{s^2}{\bar{x}};$$
$$\widehat{c} = \frac{\bar{x}^2}{s^2}.$$

1.2 Log-normaalne jaotus

Log-normaalse jaotusega juhuslik suurus omandab suuri väärtusi suurema tõenäosusega kui gammajaotus. Log-normaalse jaotuse tõenäosustihedus on kujult asümmeetriline ning positiivselt kaldu. Log-normaalse jaotuse tähistame $L(m, \sigma)$, kus skaalaparameeter m ja kujuparameeter σ on positiivsed reaalarvud, $m > 0$, $\sigma > 0$. Ühtlasi on skaalaparameeter m log-normaalse jaotuse mediaan.

Log-normaalse jaotuse tihedusfunktsioon on järgmine (Forbes jt, 2011, lk 131-132):

$$f(x) = \frac{1}{x\sigma(2\pi)^{\frac{1}{2}}} \exp\left(\frac{-(\ln\left(\frac{x}{m}\right))^2}{2\sigma^2}\right).$$

Log-normaalsel jaotusel on olemas ka alternatiivne parametrisatsioon, kus skaalaparameeter m on asendatud parameetriga μ ($\mu > 0$). Parameetrite vahel kehtivad seosed:

$$m = \exp(\mu),$$
$$\mu = \ln(m).$$

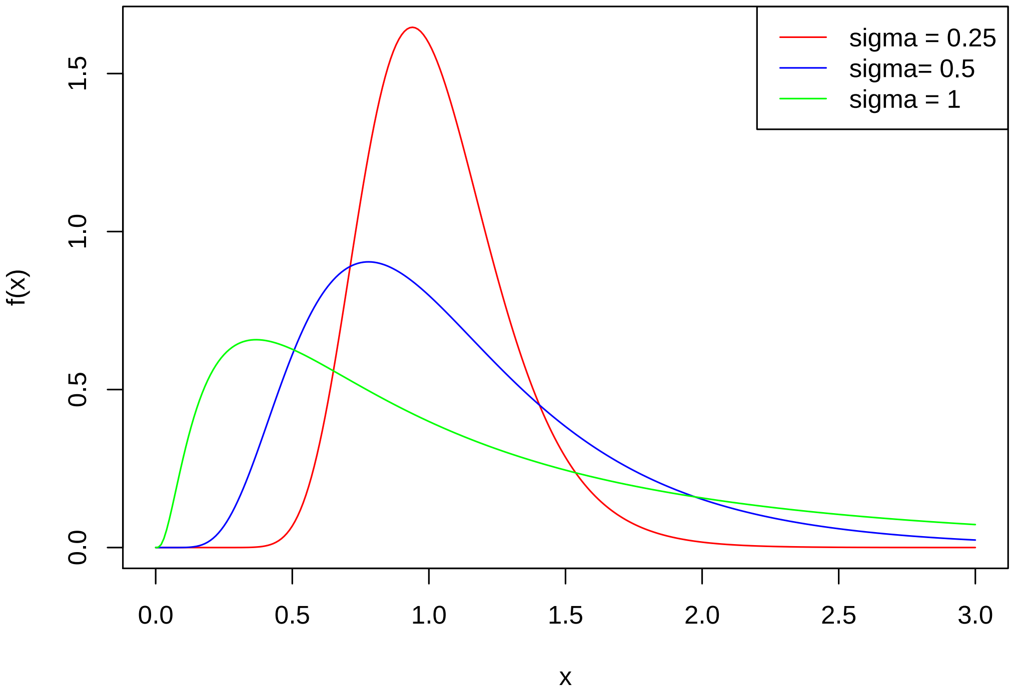
Alternatiivse parametrisatsiooni korral on log-normaalse jaotuse tihedusfunktsioon järgmine:

$$f(x) = \frac{1}{x\sigma(2\pi)^{\frac{1}{2}}} \exp\left(\frac{-(\ln x - \mu)^2}{2\sigma^2}\right).$$

Edaspidiseks teeme asenduse $\omega = \exp(\sigma^2)$.

Log-normaalse jaotusega juhusliku suuruse X arvkarakteristikud:

- keskväärtus: $EX = m\exp(\frac{1}{2}\sigma^2)$,
- dispersioon: $DX = m^2\omega(\omega - 1)$,
- k-s moment: $m_k = m^k \exp\left(\frac{1}{2}k^2\sigma^2\right) = \exp(k\mu + \frac{1}{2}k^2\sigma^2)$,
- asümmeetriakordaja: $\beta_1 = (\omega + 2)(\omega - 1)^{\frac{1}{2}}$,
- järsakuskordaja: $\beta_2 = \omega^4 + 2\omega^3 + 3\omega^2 - 3$.



Joonis 2: Log-normaalse jaotuse tihedusfunktsiooni kuju sõltuvalt parameetrist σ $(m = 0)$.

1.2.1 Parameetrite hindamine

Log-normaalse jaotuse parameetreid hindame suurima tõepära meetodiga. Esiteks paneme kirja tõepärafunktsiooni:

$$L(x_1, \ldots, x_n; \mu, \sigma) = \prod_{i=1}^{n} f(x_i; \mu, \sigma) =$$
$$= \prod_{i=1}^{n} \frac{1}{x_i(\sigma^2 2\pi)^{\frac{1}{2}}} \exp\left(\frac{-(\ln(x_i) - \mu)^2}{2\sigma^2}\right).$$

Järgmisena leiame logaritmilise tõepärafunktsiooni.

$$
\begin{aligned}
l(x_1,\ldots,x_n;\mu,\sigma) = & \\
&= \ln L(x_1,\ldots,x_n;\mu,\sigma) = \\
&= \sum_{i=1}^{n} \ln\left(\frac{1}{x_i(\sigma^2 2\pi)^{\frac{1}{2}}}\exp\left[\frac{-(\ln(x_i)-\mu)^2}{2\sigma^2}\right]\right) = \\
&= \sum_{i=1}^{n}\left[\ln 1 - \ln(x_i) - \frac{1}{2}\ln\sigma^2 - \frac{1}{2}\ln 2\pi - \frac{(\ln(x_i)-\mu)^2}{2\sigma^2}\right] = \\
&= -\sum_{i=1}^{n}\ln(x_i) - \sum_{i=1}^{n}\frac{1}{2}\ln\sigma^2 - \sum_{i=1}^{n}\frac{1}{2}\ln 2\pi - \sum_{i=1}^{n}\frac{(\ln(x_i)-\mu)^2}{2\sigma^2} = \\
&= -\sum_{i=1}^{n}\ln(x_i) - \frac{n}{2}\ln\sigma^2 - 0 - \frac{\sum_{i=1}^{n}\ln(x_i)^2 - 2\ln(x_i)\mu + \mu^2}{2\sigma^2} = \\
&= -\sum_{i=1}^{n}\ln(x_i) - \frac{n}{2}\ln\sigma^2 - \frac{\sum_{i=1}^{n}\ln(x_i)^2}{2\sigma^2} + \frac{\sum_{i=1}^{n}\ln(x_i)\mu}{\sigma^2} - \frac{n\mu^2}{2\sigma^2}
\end{aligned}
$$

Hinnangute leidmiseks peame lahendama võrrandisüsteemi:

$$
\begin{cases}
\frac{\partial l}{\partial \mu} & = 0 \\
\frac{\partial l}{\partial \sigma} & = 0.
\end{cases}
$$

Nüüd leiame parameetri μ hinnangu $\widehat{\mu}$, selleks võtame logaritmilisest tõepärafunktsioonist osatuletise μ järgi:

$$
\begin{aligned}
\frac{\partial l}{\partial \mu} = 0 \Rightarrow \frac{\sum_{i=1}^{n}\ln(x_i)}{\sigma^2} - \frac{2n\mu}{2\sigma^2} &= 0; \\
\Rightarrow \frac{n\mu}{\sigma^2} &= \frac{\sum_{i=1}^{n}\ln(x_i)}{\sigma^2}; \\
\Rightarrow n\mu &= \sum_{i=1}^{n}\ln(x_i); \\
\Rightarrow \mu &= \frac{1}{n}\sum_{i=1}^{n}\ln(x_i).
\end{aligned}
$$

Oleme saanud μ hinnangu $\widehat{\mu}$:

$$\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n} \ln\left(x_i\right)$$

Leiame $\widehat{\sigma}$, selleks võtame logaritmilisest tõepärafunktsioonist osatuletise σ järgi:

$$\begin{aligned}
\frac{\partial l}{\partial \sigma} = 0 &\Rightarrow -\frac{2n}{2\sigma} - \frac{(-2)\sum_{i=1}^{n}\ln\left(x_i\right)^2}{2\sigma^3} + \frac{(-2)\sum_{i=1}^{n}\ln\left(x_i\right)\mu}{\sigma^3} - \frac{(-2)n\mu^2}{2\sigma^3} = 0;\\
&\Rightarrow -\frac{n}{\sigma} + \frac{\sum_{i=1}^{n}\ln(x_i)^2 - 2\sum_{i=1}^{n}\ln(x_i)\mu + \sum_{i=1}^{n}\mu^2}{\sigma^3} = 0;\\
&\Rightarrow \frac{\sum_{i=1}^{n}(\ln(x_i)-\mu)^2}{\sigma^3} = \frac{n}{\sigma};\\
&\Rightarrow n\sigma^3 = \sigma\sum_{i=1}^{n}(\ln(x_i)-\mu)^2;\\
&\Rightarrow \sigma^2 = \frac{\sum_{i=1}^{n}(\ln\left(x_i\right)-\mu)^2}{n};\\
&\Rightarrow \sigma = \sqrt{\frac{\sum_{i=1}^{n}(\ln\left(x_i\right)-\mu)^2}{n}}.
\end{aligned}$$

Asendame μ tema hinnanguga $\widehat{\mu}$ ja saame σ hinnangu:

$$\widehat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n}(\ln\left(x_i\right)-\widehat{\mu})^2}{n}}.$$

Oleme saanud, et skaalaparameetri μ hinnang on kujul:

$$\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n}\ln(x_i)$$

ning kujuparameetri σ hinnang on järgmine:

$$\widehat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n}(\ln\left(x_i\right)-\widehat{\mu})^2}{n}}.$$

Alternatiivse skaalaparameetri m ehk log-normaalse jaotuse mediaani hinnang on kujul:

$$\widehat{m} = \exp\left(\widehat{\mu}\right).$$

1.3 Pareto jaotus

Pareto jaotus on pidev jaotus, mis on määratud reaaltelje positiivsel osal. Pareto jaotuse tähistame $Pa(a,c)$, kus kohaparameeter a ja kujuparameeter c on positiivsed reaalarvud ($a > 0,\ c > 0$). Pareto jaotus on raskema sabaga kui gammajaotus.

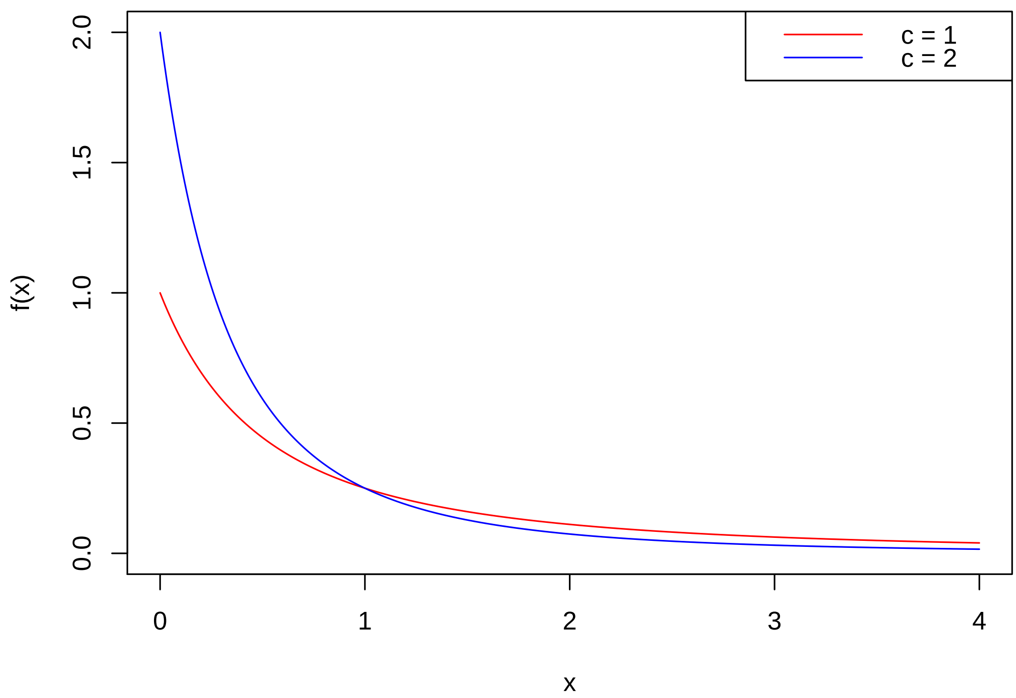
Pareto jaotuse tihedusfunktioon on järgmine (Forbes jt, 2011, lk 149):

$$f(x) = \frac{ca^c}{x^{c+1}}, \quad x > 0.$$

Pareto jaotusega juhusliku suuruse X arvkarakteristikud:

- keskväärtus: $EX = \frac{ca}{(c-1)}$, kus $c > 1$,
- dispersioon: $DX = \frac{ca^2}{[(c-1)^2(c-2)]}$, kus $c > 2$,
- k-s moment: $m_k = \frac{ca^k}{c-k}, \quad c > k$,
- asümmeetriakordaja $\beta_1 = 2\frac{c+1}{c-3}\sqrt{\frac{c-2}{c}}, \quad c > 3$,
- järsakuskordaja $\beta_2 = \frac{3(c-2)(3c^2+c+2)}{c(c-3)(c-4)}, \quad c > 4$.

Asümmeetria- ja järsakuskordaja valemid on raamatust Johnson jt, (1994), lk 577.



Joonis 3: Pareto jaotuse tihedusfunktsiooni kuju sõltuvalt kujuparameetrist c ($a = 1$).

1.3.1 Parameetrite hindamine

Pareto jaotuse parameetreid hindame suurima tõepära meetodiga. Esmalt paneme kirja tõepärafunktsiooni:

$$L(x_1, \ldots, x_n; c, a) = \prod_{i=1}^{n} f(x_i; c, a) = \prod_{i=1}^{n} \frac{ca^c}{x_i^{c+1}}.$$

Logaritmilise tõepärafunktsiooni saame kujul.

$$
\begin{aligned}
l(x_1, \ldots, x_n; c, a) = \ln L(x_1, \ldots, x_n; c, a) &= \\
&= \sum_{i=1}^{n} \ln\left(\frac{ca^c}{x_i^{c+1}}\right) = \\
&= \sum_{i=1}^{n} (\ln c + c\ln a - (c+1)\ln(x_i)) = \\
&= \sum_{i=1}^{n} \ln c + \sum_{i=1}^{n} c\ln a - \sum_{i=1}^{n}(c+1)\ln(x_i) = \\
&= n\ln c + nc\ln a - (c+1)\sum_{i=1}^{n} \ln(x_i).
\end{aligned}
$$

Hinnangute leidmiseks peame lahendama võrrandisüsteemi:

$$
\begin{cases}
\frac{\partial l}{\partial c} & = 0, \\
\frac{\partial l}{\partial a} & = 0.
\end{cases}
$$

Leiame esmalt osatuletise c järgi:

$$
\begin{aligned}
\frac{\partial l}{\partial c} = 0 &\Rightarrow \frac{n}{c} + n\ln a - \sum_{i=1}^{n} \ln(x_i) = 0; \\
&\Rightarrow \frac{n}{c} = -n\ln a + \sum_{i=1}^{n} \ln(x_i); \\
&\Rightarrow \frac{1}{c} = \frac{1}{n}\sum_{i=1}^{n} \ln(x_i) - \frac{n}{n}\ln a; \\
&\Rightarrow \frac{1}{c} = \frac{1}{n}\sum_{i=1}^{n} (\ln(x_i) - \ln a); \\
&\Rightarrow \frac{1}{c} = \frac{1}{n}\sum_{i=1}^{n} \ln\left(\frac{x_i}{a}\right); \\
&\Rightarrow c = \frac{n}{\sum_{i=1}^{n} \ln\left(\frac{x_i}{a}\right)}.
\end{aligned}
$$

Võtame osatuletise a järgi:

$$\frac{\partial l}{\partial a} = n\frac{1}{a}.$$

On näha, et saadud avaldist ei saa võrdsustada nulliga. Seega ei saa osatuletise abil a hinnangut leida. Siiski teame, et $x_i \geq a$. Seega a ei saa olla suurem kui valimi väikseim väärtus, sellest saame, et $\widehat{a} = \min\{x_i\}$, $i = 1, \ldots, n$.

Oleme saanud parameetri a hinnangu järgmisel kujul:

$$\widehat{a} = \min\{x_1, \ldots, x_n\}.$$

Asendame selle c avaldisse ja saame parameetri c hinnangu kujul:

$$\widehat{c} = \frac{1}{\left(\frac{1}{n}\right)\sum_{i=1}^{n} \ln\left(\frac{x_i}{\widehat{a}}\right)}.$$

1.4 Weibulli jaotus

Weibulli jaotust saab rakendada positiivsete väärtustega andmetel. Weibulli jaotust tähistame $W(\eta, \beta)$, kus skaalaparameeter η ja kujuparameeter β on positiivsed ($\eta, \beta > 0$).

Weibulli jaotuse tihedusfunktsioon on järgmine (Forbes jt, 2011, lk 193):

$$f(x) = \left(\frac{\beta x^{\beta-1}}{\eta^{\beta}}\right)\exp\left[-\left(\frac{x}{\eta}\right)^{\beta}\right].$$

Weibulli jaotusega juhusliku suuruse X arvkarakteristikud:

- keskväärtus: $EX = \eta\Gamma\left[\frac{\beta+1}{\beta}\right]$,
- dispersioon: $DX = \eta^2\left(\Gamma\left[\frac{\beta+2}{\beta}\right] - \left(\Gamma\left[\frac{\beta+1}{\beta}\right]\right)^2\right)$,
- k-s moment: $m_k = \eta^k\Gamma\left[\frac{\beta+k}{\beta}\right]$.

Järgmisena leiame Weibulli jaotuse asümmeetriakordaja ja järsakuskordaja avaldised. Esmalt tuletame asümmeetriakordaja β_1.

$$
\begin{aligned}
\beta_1 &= \frac{E(X-EX)^3}{(DX)^{\frac{3}{2}}} = \\
&= \frac{E(X^3 - 3X^2EX + 3X(EX)^2 - (EX)^3)}{(DX)^{\frac{3}{2}}} = \\
&= \frac{E(X^3) - 3EXE(X^2) + 2(EX)^3}{(DX)^{\frac{3}{2}}} = \\
&= \frac{m_3(X) - 3EXm_2(X) + 2(EX)^3}{(DX)^{\frac{3}{2}}}.
\end{aligned}
$$

Gammafunktsiooni kaudu avaldub asümmeetriakordaja järgmiselt:

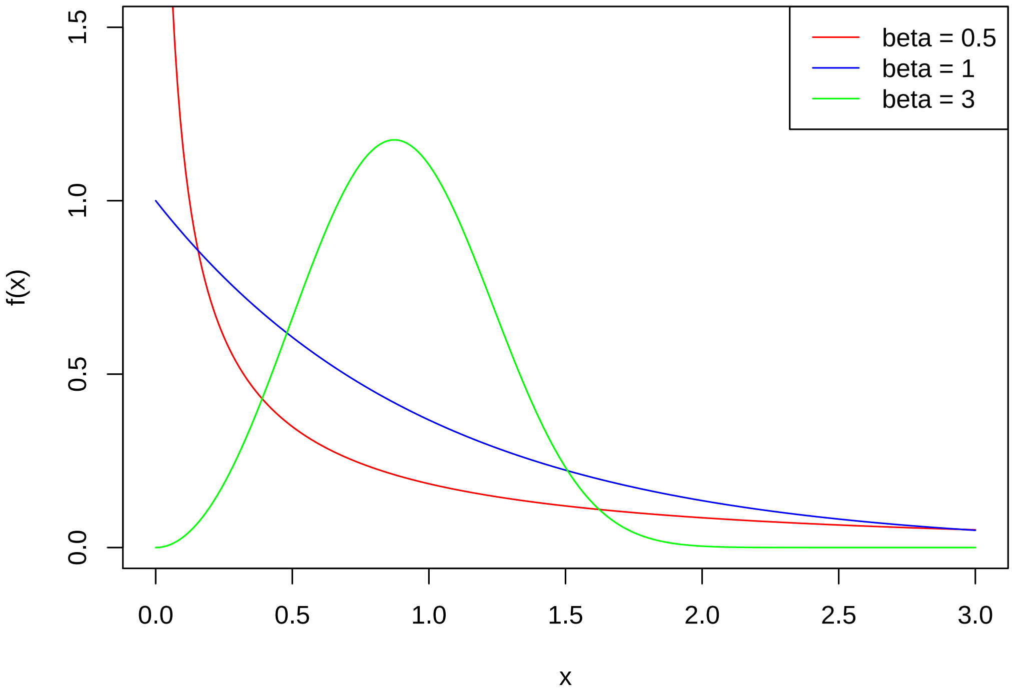
$$
\beta_1 = \frac{\Gamma\left[\frac{\beta+3}{\beta}\right] - 3\Gamma\left[\frac{\beta+1}{\beta}\right]\Gamma\left[\frac{\beta+2}{\beta}\right] + 2\left(\Gamma\left[\frac{\beta+1}{\beta}\right]\right)^3}{\left(\Gamma\left[\frac{\beta+2}{\beta}\right] - \left(\Gamma\left[\frac{\beta+1}{\beta}\right]\right)^2\right)^{\frac{3}{3}}}.
$$

Järsakuskordaja β_2 jaoks saame esituse momentide kaudu:

$$\begin{aligned}\beta_2 &= \frac{E(X-EX)^4}{(DX)^2} = \\ &= \frac{E(X^4 - 4X^3EX + 6X^2(EX)^2 - 4X(EX)^3 + (EX)^4)}{(DX)^2} = \\ &= \frac{E(X^4) - 4E(X^3)EX + 6E(X^2)(EX)^2 - 4E(X)(EX)^3 + (EX)^4)}{(DX)^2} = \\ &= \frac{E(X^4) - 4E(X^3)EX + 6E(X^2)(EX)^2 - 3(EX)^4)}{(DX)^2} = \\ &= \frac{m_4(X) - 4m_3(X)EX + 6m_2(X)(EX)^2 - 3(EX)^4}{(DX)^2}.\end{aligned}$$

Gammafunktsiooni kaudu avaldub järsakuskordaja kujul:

$$\beta_2 = \frac{\Gamma\left[\frac{\beta+4}{\beta}\right] - 4\Gamma\left[\frac{\beta+3}{\beta}\right]\Gamma\left[\frac{\beta+1}{\beta}\right] + 6\Gamma\left[\frac{\beta+2}{\beta}\right]\left(\Gamma\left[\frac{\beta+1}{\beta}\right]\right)^2 - 3\left(\Gamma\left[\frac{\beta+1}{\beta}\right]\right)^4}{\left(\Gamma\left[\frac{\beta+2}{\beta}\right]\right)^2 - 2\Gamma\left[\frac{\beta+2}{\beta}\right]\left(\Gamma\left[\frac{\beta+1}{\beta}\right]\right)^2 + \left(\Gamma\left[\frac{\beta+1}{\beta}\right]\right)^4}.$$



Joonis 4: Weibulli jaotuse tihedusfunktsiooni kuju sõltuvalt kujuparameetrist β ($\eta = 1$).

1.4.1 Parameetrite hindamine

Weibulli jaotuse parameetreid hindame suurima tõepära meetodiga. Paneme esmalt kirja tõepärafunktsiooni:

$$
\begin{aligned}
L(x_1,\dots,x_n;\beta,\eta) &= \prod_{i=1}^{n} f(x_i;\beta,\eta) = \\
&= \prod_{i=1}^{n} \frac{\beta}{\eta^{\beta}} x_i^{\beta-1} \exp\left[-\left(\frac{x_i}{\eta}\right)^{\beta}\right].
\end{aligned}
$$

Järgmisena leiame logaritmilise tõepärafunktsiooni :

$$
\begin{aligned}
l(x_1,\dots,x_n;\beta,\eta) &= \ln L(x_1,\dots,x_n;\beta,\eta) = \\
&= \sum_{i=1}^{n} \ln\left(\frac{\beta}{\eta^{\beta}} x_i^{\beta-1} \exp\left[-\left(\frac{x_i}{\eta}\right)^{\beta}\right]\right) = \\
&= \sum_{i=1}^{n} \left[\ln\beta - \beta\ln\eta + (\beta-1)\ln(x_i) - \left(\frac{x_i}{\eta}\right)^{\beta}\right] = \\
&= \sum_{i=1}^{n} \ln\beta - \sum_{i=1}^{n} \beta\ln\eta + \sum_{i=1}^{n}(\beta-1)\ln(x_i) - \sum_{i=1}^{n}\left(\frac{x_i}{\eta}\right)^{\beta} = \\
&= n\ln\beta - n\beta\ln\eta + (\beta-1)\sum_{i=1}^{n}\ln(x_i) - \sum_{i=1}^{n}\left(\frac{x_i}{\eta}\right)^{\beta}.
\end{aligned}
$$

Hinnangute leidmiseks peame lahendama võrrandisüsteemi:

$$
\begin{cases}
\frac{\partial l}{\partial \eta} &= 0, \\
\frac{\partial l}{\partial \beta} &= 0.
\end{cases}
$$

Leiame esmalt osatuletise η järgi:

$$\begin{aligned}
\frac{\partial l}{\partial \eta} = 0 \Rightarrow 0 - \frac{n\beta}{\eta} + 0 + \beta \sum_{i=1}^{n} \frac{x_i^{\beta}}{\eta^{\beta+1}} = 0;\\
\Rightarrow -\frac{n\beta}{\eta} + \frac{\beta}{\eta} \sum_{i=1}^{n} \frac{x_i^{\beta}}{\eta^{\beta}} = 0;\\
\Rightarrow -1 + \frac{1}{n} \sum_{i=1}^{n} \frac{x_i^{\beta}}{\eta^{\beta}} = 0;\\
\Rightarrow 1 = \frac{1}{n\eta^{\beta}} \sum_{i=1}^{n} x_i^{\beta};\\
\Rightarrow \eta^{\beta} = \frac{1}{n} \sum_{i=1}^{n} x_i^{\beta};\\
\Rightarrow \eta = \sqrt[\beta]{\frac{1}{n} \sum_{i=1}^{n} x_i^{\beta}}.
\end{aligned}$$

Tähistame

$$\bar{x}_{\beta} = \frac{1}{n} \sum_{i=1}^{n} x_i^{\beta},$$

siis

$$\eta = \sqrt[\beta]{\bar{x}_{\beta}}.$$

Leiame osatuletise β järgi:

$$\begin{aligned}
\frac{\partial l}{\partial \beta} = 0 &\Rightarrow \frac{n}{\beta} - n\ln\eta + \sum_{i=1}^{n}\ln(x_i) - \sum_{i=1}^{n}\ln\left(\frac{x_i}{\eta}\right)^{\beta}\left(\frac{x_i}{\eta}\right) = 0;\\
&\Rightarrow \frac{n}{\beta} - n\ln\eta + \sum_{i=1}^{n}\ln(x_i) - \frac{1}{\eta^{\beta}}\sum_{i=1}^{n}\ln\left(\frac{x_i}{\eta}\right)x_i^{\beta} = 0;\\
&\Rightarrow \frac{n}{\beta} - n\ln\eta + \sum_{i=1}^{n}\ln(x_i) - \frac{1}{\eta^{\beta}}\sum_{i=1}^{n}(\ln x_i - \ln\eta)\,x_i^{\beta} = 0;\\
&\Rightarrow \frac{n}{\beta} - n\ln\eta + \sum_{i=1}^{n}\ln(x_i) - \frac{1}{\eta^{\beta}}\sum_{i=1}^{n}\left(\ln(x_i)x_i^{\beta} - \ln(\eta)x_i^{\beta}\right) = 0;\\
&\Rightarrow \frac{n}{\beta} - n\ln\eta + \sum_{i=1}^{n}\ln(x_i) - \frac{1}{\eta^{\beta}}\sum_{i=1}^{n}\ln(x_i)x_i^{\beta} + \frac{1}{\eta^{\beta}}\sum_{i=1}^{n}\ln(\eta)x_i^{\beta} = 0;\\
&\Rightarrow \frac{1}{\beta} = \ln\eta - \frac{\sum_{i=1}^{n}\ln(x_i)}{n} + \frac{1}{n\eta^{\beta}}\sum_{i=1}^{n}\ln(x_i)x_i^{\beta} - \frac{1}{n\eta^{\beta}}\ln(\eta)\sum_{i=1}^{n}x_i^{\beta};\\
&\Rightarrow \frac{1}{\beta} = -\frac{\sum_{i=1}^{n}\ln(x_i)}{n} + \frac{1}{n\eta^{\beta}}\sum_{i=1}^{n}\ln(x_i)x_i^{\beta};\\
&\Rightarrow \beta = \frac{n}{\left(\frac{1}{\eta^{\beta}}\right)\sum_{i=1}^{n}x_i^{\beta}\ln(x_i) - \sum_{i=1}^{n}\ln(x_i)}.
\end{aligned}$$

Kujuparameetri β ja skaalaparameetri η hinnangud on võrrandisüsteemi lahendid:

$$\begin{cases} \widehat{\eta} = \sqrt[\widehat{\beta}]{\bar{x}_{\beta}} \\ \widehat{\beta} = \dfrac{n}{\left(\frac{1}{\widehat{\eta}^{\widehat{\beta}}}\right)\sum_{i=1}^{n}x_i^{\widehat{\beta}}\ln(x_i) - \sum_{i=1}^{n}\ln(x_i)}. \end{cases}$$

Märkus: Raamatus (Forbes jt, 2011, lk 196) on kujuparameetri β hinnangu esituses viga.

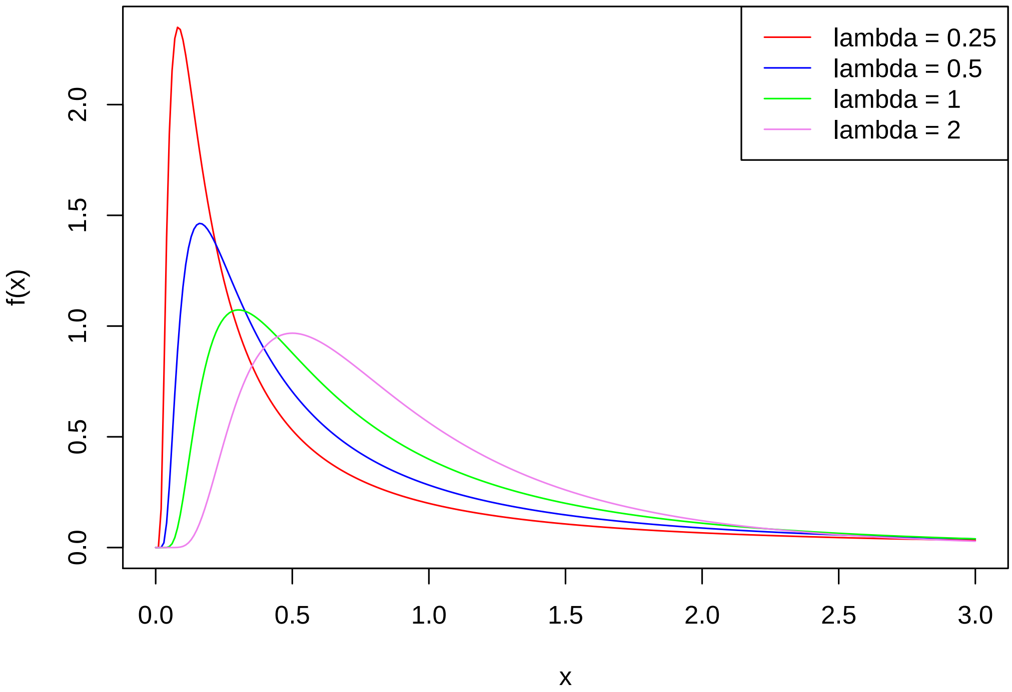
1.5 Gaussi pöördjaotus (Waldi jaotus)

Waldi jaotust tähistame $I(\mu, \lambda)$, kus kohaparameeter μ ja skaalaparameeter λ on positiivsed ($\mu, \lambda > 0$).
Waldi jaotuse tihedusfunktsioon on järgmine (Forbes jt, 2011, lk 120-121):

$$f(x) = \left(\frac{\lambda}{2\pi x^3}\right)^{\frac{1}{2}} \exp\left(\frac{-\lambda(x-\mu)^2}{2\mu^2 x}\right).$$

Waldi jaotusega juhusliku suuruse X arvkarakteristikud:

- keskväärtus: $EX = \mu$,
- dispersioon: $DX = \frac{\mu^3}{\lambda}$,
- k-s moment: $m_k = \mu^k \sum_{i=0}^{k-1} \frac{(k-1+i)!}{i!(k-1-i)!} \left(\frac{\mu}{2\lambda}\right)^i, \quad k \geq 2,$
- asümmeetriakordaja: $\beta_1 = 3\left(\frac{\mu}{\lambda}\right)^{\frac{1}{2}}$,
- järsakuskordaja: $\beta_2 = 3 + \frac{15\mu}{\lambda}$.



Joonis 5: Waldi jaotuse tihedusfunktsiooni kuju sõltuvalt kujuparameetrist λ ($\mu = 1$).

1.5.1 Parameetrite hindamine

Waldi jaotuse parameetreid hindame suurima tõepära meetodiga. Esmalt paneme kirja tõepärafunktsiooni.

$$\begin{aligned} L(x_1, \ldots, x_n; \mu, \lambda) &= \prod_{i=1}^{n} f(x_i; \mu, \lambda) = \\ &= \prod_{i=1}^{n} \left(\frac{\lambda}{2\pi x_i^3} \right)^{\frac{1}{2}} \exp\left(\frac{-\lambda (x_i - \mu)^2}{2\mu^2 x_i} \right). \end{aligned}$$

Järgmisena leiame logaritmilise tõepärafunktsiooni:

$$
\begin{aligned}
l(x_1,\dots,x_n;\mu,\lambda) = \ln L(x_1,\dots,x_n;\mu,\lambda) &= \\
&= \sum_{i=1}^{n} \ln\left[\left(\frac{\lambda}{2\pi x_i^3}\right)^{\frac{1}{2}} \exp\left(\frac{-\lambda(x_i-\mu)^2}{2\mu^2 x_i}\right)\right] = \\
&= \sum_{i=1}^{n}\left[\frac{1}{2}\left(\ln\lambda - \ln 2\pi - 3\ln(x_i)\right) - \frac{\lambda(x_i-\mu)^2}{2\mu^2 x_i}\right] = \\
&= \sum_{i=1}^{n}\frac{1}{2}\ln\lambda - \sum_{i=1}^{n}\frac{1}{2}\ln 2\pi - \sum_{i=1}^{n}\frac{3}{2}\ln(x_i) - \sum_{i=1}^{n}\frac{-\lambda(x_i-\mu)^2}{2\mu^2 x_i} = \\
&= \frac{n}{2}\ln\lambda - \frac{n}{2}\ln 2\pi - \frac{3}{2}\sum_{i=1}^{n}\ln(x_i) - \sum_{i=1}^{n}\frac{-\lambda(x_i-\mu)^2}{2\mu^2 x_i}.
\end{aligned}
$$

Esiteks võtame osatuletise μ järgi ja võrdsustame nulliga:

$$
\begin{aligned}
\frac{\partial l}{\partial \mu} &= \sum_{i=1}^{n}\frac{-2\lambda(x_i-\mu)}{4\mu x_i} = 0; \\
&\Rightarrow \frac{\sum_{i=1}^{n} -\lambda(x_i-\mu)}{\sum_{i=1}^{n} 2\mu x_i} = 0; \\
&\Rightarrow \frac{\sum_{i=1}^{n} -\lambda x_i + \sum_{i=1}^{n}\lambda\mu}{\sum_{i=1}^{n} 2\mu x_i} = 0; \\
&\Rightarrow \frac{-\lambda\sum_{i=1}^{n} x_i + n\lambda\mu}{2\mu\sum_{i=1}^{n} x_i} = 0; \\
&\Rightarrow \frac{\lambda\sum_{i=1}^{n} x_i}{2\mu\sum_{i=1}^{n} x_i} = \frac{n\lambda\mu}{2\mu\sum_{i=1}^{n} x_i}; \\
&\Rightarrow 2\mu^2 n\lambda\sum_{i=1}^{n} x_i = 2\mu\lambda\left(\sum_{i=1}^{n} x_i\right)^2.
\end{aligned}
$$

Saame μ hinnangu:

$$
\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}.
$$

Järgmisena leiame osatuletise λ järgi:

$$\begin{aligned}
\frac{\partial l}{\partial \lambda} &= \frac{n}{2\widehat{\lambda}} - \frac{1}{2\widehat{\mu}^2}\sum_{i=1}^{n}\frac{(x_i-\widehat{\mu})^2}{x_i} = 0;\\
&\Rightarrow \frac{n}{2\widehat{\lambda}} - \frac{1}{2\widehat{\mu}^2}\sum_{i=1}^{n}(x_i - 2\widehat{\mu} + \frac{\widehat{\mu}^2}{x_i}) = 0;\\
&\Rightarrow \frac{n}{2\widehat{\lambda}} - \frac{1}{2\widehat{\mu}^2}\left(\sum_{i=1}^{n}x_i - n2\widehat{\mu} + \widehat{\mu}^2\sum_{i=1}^{n}\frac{1}{x_i}\right) = 0;\\
&\Rightarrow \frac{n}{2\widehat{\lambda}} - \frac{1}{2\bar{x}^2}\left(n\bar{x} - n2\bar{x} + \bar{x}^2\sum_{i=1}^{n}\frac{1}{x_i}\right) = 0;\\
&\Rightarrow \frac{n}{2\widehat{\lambda}} - \frac{1}{2\bar{x}^2}\left(\bar{x}^2\sum_{i=1}^{n}\frac{1}{x_i} - \bar{x}\right) = 0;\\
&\Rightarrow \frac{n}{2\widehat{\lambda}} - \frac{\bar{x}^2\sum_{i=1}^{n}\frac{1}{x_i}}{2\bar{x}^2} + \frac{\bar{x}}{2\bar{x}^2} = 0;\\
&\Rightarrow \frac{n}{2\widehat{\lambda}} = \frac{1}{2}\sum_{i=1}^{n}\frac{1}{x_i} - \frac{1}{2\bar{x}} = 0;\\
&\Rightarrow \frac{1}{\widehat{\lambda}} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i} - \frac{1}{n\bar{x}} = 0;\\
&\Rightarrow \bar{\lambda} = \frac{n}{\sum_{i=1}^{n}\frac{1}{x_i} - \frac{1}{\bar{x}}}.
\end{aligned}$$

Oleme saanud, et kohaparameetri μ hinnang on valimi keskmine

$$\widehat{\mu} = \bar{x}$$

ning skaalaparameetri λ hinnang on kujul:

$$\widehat{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i^{-1} - (\bar{x})^{-1}}.$$

2 Andmete kirjeldus

Töös kasutame If P&C Insurance AS kaskokindlustuse kahjuandmeid. Andmestikus on Eesti, Läti ja Leedu andmed ning andmestikku on kaasatud sõidu- ja pakiautod. Andmestik sisaldab kahjusid, mis on toimunud aastatel 2008-2019, aga on kindlustusse kahjuna teatatud aastatel 2010-2019. Andmestikus on 153 343 rida ning iga rida vastab ühele kahjule. Andmestikus on seitse tunnust, nendeks on riik, kus kahju toimus, teatamise kuu, toimumise kuupäev, kahju klassifikatsioon, hüvitis eurodes (omavastutus on maha arvestatud), kliendi tüüp ja sõiduki vanus täisaastates kahju toimumise ajal. Tunnus *kahju klassifikatsioon* näitab, mis liiki kahjuga on tegemist, näiteks: klaasikahju, vargus. Kahjude klassifikatsioonid pole riikide lõikes ühtsed. Tunnus *kliendi tüüp* kirjeldab, kas kindlustusvõtja on juriidiline isik või eraklient. Tunnus *hüvitis* on kindlustuse poolt makstud või potentsiaalselt makstav hüvitis.

Töö eesmärgiks on uurida kahjude hüvitiste jaotusi riikide lõikes, seetõttu koostame tunnuse *riik* põhjal algandmestikust kolm alamandmestikku ja töös kasutame tunnust *hüvitis*. Kolmest alamandmestikust eemaldame hüvitiste suurused, mis on väiksemad kui 1 euro. Lisaks eemaldame autoabi kahjud, sest need käituvad teistest kahjudest erinevalt.

2.1 Eesti

Alamandmestikus, kus on ainult Eestis toimunud kahjud on 50 844 rida. Eesti andmetabel sisaldab kahjusid, mis on toimunud aastatel $2008 - 2017$ ning, mis on kahjuna teatatud aastatel $2010 - 2019$. Tabelis 1 toome välja Eesti kaskokahjude hüvitiste empiirilise jaotuse arvkarakteristikud. Näeme, et kõige väiksem suurus on $1,16$ eurot ja suurim hüvitis on 56 501,48 eurot. Eesti andmete keskmine hüvitis on $958,59$ eurot ja mediaan $459,81$ eurot ehk mediaan on keskmisest ligikaudu kaks korda väiksem. Sellele viitab ka positiivne asümmeetriakordaja, mille väärtuseks on $9,07$. Järsakuskordaja väärtuseks on $162,86$. Standardhälve on $1701,47$ eurot, mis on suur, seega kahju hüvitise väärtuste hajuvus on suur.

Tabel 1: Eesti kaskokahjude hüvitiste empiirilised karakteristikud.

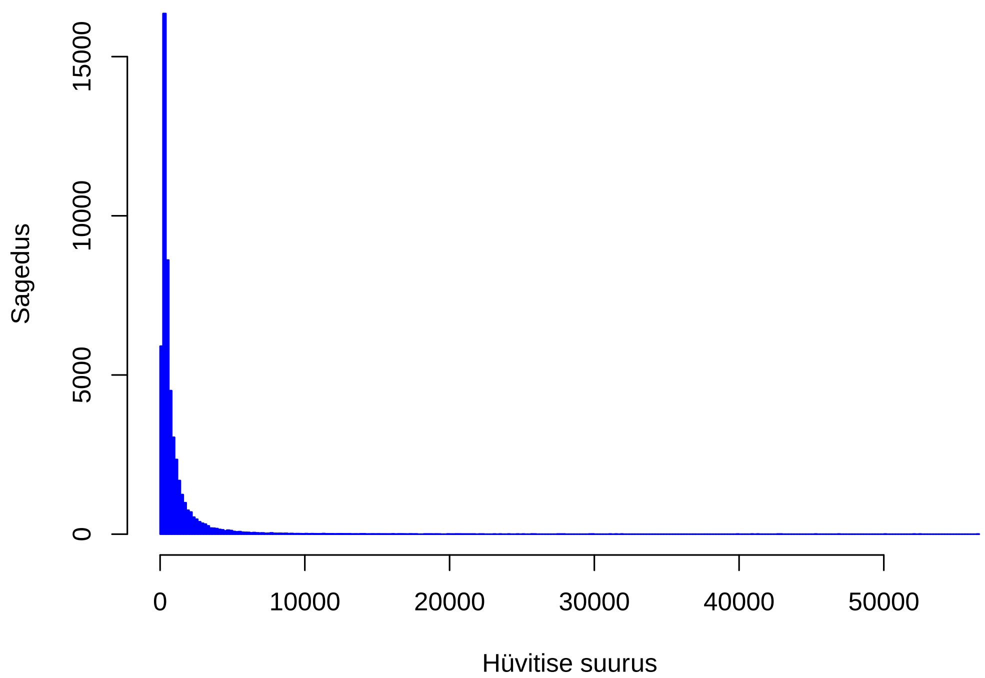
	Miinimum	Maksimum	Keskmine	Mediaan	Standardhälve	Järsakuskordaja	Asümmeetriakordaja
Kahju maksumus	1,16	56 501,48	958,59	459,81	1701,47	162,86	9,07

Tabelis 2 toome välja Eesti kaskokahjude hüvitiste kvantiilid. Kvantiilide tabelist näeme, et ühe protsendi kvantiili hüvitise väärtuseks on 50 eurot, see tähendab, et antud väärtusest, esineb üks protsent väiksemaid hüvitisi. Üheksakümne protsendi kvantiili hüvitise väärtuseks on $2088,98$ eurot ehk antud väärtusest esineb kümme protsenti suuremaid hüvitisi. Üheksakümne viie protsendi kvantiili hüvitise väärtuseks on $3299,11$ eurot ehk antud väärtusest esineb viis protsenti suuremaid hüvitisi. Üheksakümne üheksa protsendi kvantiili hüvitise väärtuseks on $7811,97$ eurot, seega antud väärtusest esineb üks protsent suuremaid hüvitisi.

Tabel 2: Eesti kaskokahjude hüvitiste kvantiilid.

	Kvantiilid					
	1%	5%	10%	90%	95%	99%
Kahju maksumus	50	140,68	190,21	2088,98	3299,11	7811,97

Hüvitiste kogusumma Eesti kaskokahjude puhul on 48 738 369 eurot ja 1% kvantiilist väiksemate hüvitiste kogusumma on 17 459,04 eurot. 1% kvantiilist väiksemate hüvitiste kogusumma moodustab $0,036\%$ kõikide hüvitiste kogusummast. Joonisel 6 kujutame Eesti kaskokahjude hüvitiste histogrammi. Jooniselt näeme, et suurem osa hüvitistest on väikesed ning suuremaid hüvitisi esineb harva. Suurte kahjude esinemine venitab jaotuse saba pikaks.



Joonis 6: Eesti kaskokahjude hüvitiste histogramm.

2.2 Läti

Läti kaskokahjude andmestikus on 30 087 kahjut. Läti andmestikus esinevad kahjud, mis on toimunud aastatel 2009-2019 ning, mis on kahjuna teatatud aastatel 2010-2019. Tabelis 3 toome välja Läti kaskokahjude hüvitiste empiirilise jaotuse arvkarakteristikuid. Näeme, et väikseim suurus on 1 euro ja suurim hüvitis on 171 359,8 eurot. Läti kahjude keskmine hüvitis on $1018,7$ eurot ja mediaan on $435,9$ eurot ehk keskväärtus on mediaanist ligikaudu kaks korda suurem. Sellele tulemusele viitab ka positiivne ning küllaltki suure väärtusega asümmeetriakordaja, milleks on $18,13$. Järsakuskordaja väärtuseks on $815,74$. Standardhälbe väärtuseks on $2511,4$ eurot, mis on suurem kui Eesti andmete standardhälve.

Tabel 3: Läti kaskokahjude hüvitiste empiirilised karakteristikud.

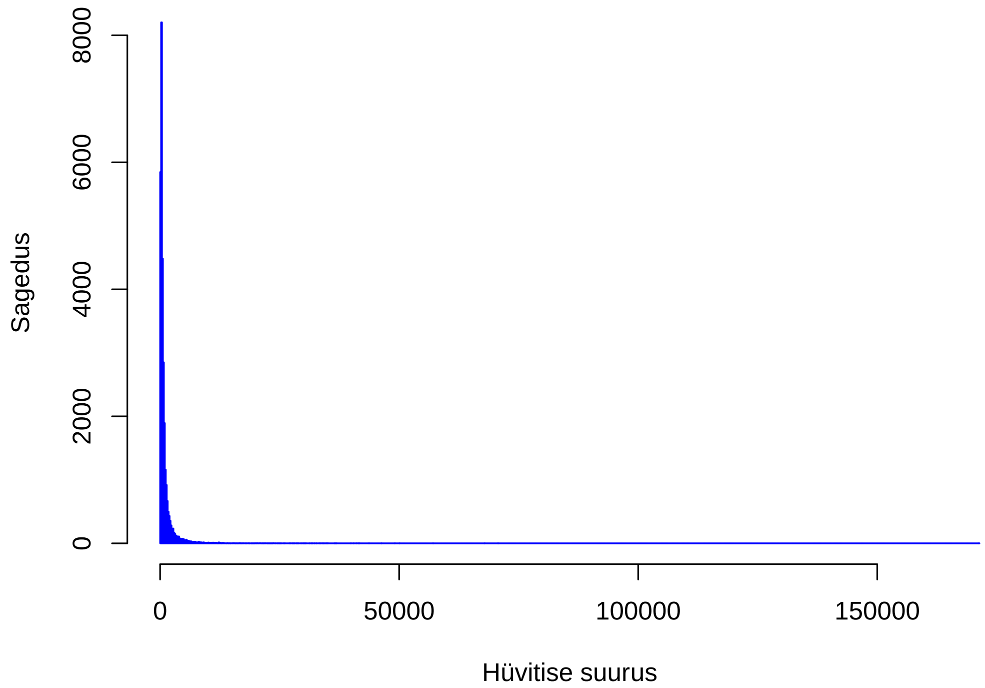
	Miinimum	Maksimum	Keskmine	Mediaan	Standardhälve	Järsakuskordaja	Asümmeetriakordaja
Kahju maksumus	1	171 359,8	1018,7	435,9	2511,4	815,74	18,13

Tabelis 4 esitame Läti kaskokahjude hüvitiste kvantiilid. Kvantiilide tabelist näeme, et ühe protsendi kvantiili hüvitise väärtuseks on 40 eurot. Üheksakümne protsendi kvantiili hüvitise väärtuseks on $2063,17$ eurot. Üheksakümne viie protsendi kvantiili hüvitise väärtuseks on $3550,18$ eurot. Üheksakümne üheksa protsendi kvantiili hüvitise väärtuseks on 10 139,44 eurot.

Tabel 4: Läti kaskokahjude hüvitiste kvantiilid.

	Kvantiilid					
	1%	5%	10%	90%	95%	99%
Kahju maksumus	40	91,05	139	2063,17	3550,18	10 139,44

Hüvitiste kogusumma on 30 648 883 eurot ja 1% kvantiilist väiksemate hüvitiste kogusumma on 6962 eurot. 1% kvantiilist väiksemate hüvitiste kogusumma moodustab $0,023\%$ kõikide hüvitiste kogusummast. Joonisel 7 kujutame Läti kaskokahjude hüvitiste histogrammi. Jooniselt näeme, et suurem osa hüvitistest on väikesed ning suuremaid kahjude hüvitisi esineb harva. Suurte kahjude esinemine venitab jaotuse saba pikaks. Eriti suure kahju esinemine toob kaasa märgatavalt suuremad asümmeetria- ja järsakuskordaja väärtused võrreldes Eesti vastavate suurustega.



Joonis 7: Läti kaskokahjude hüvitiste histogramm.

2.3 Leedu

Leedu kaskokahjude andmestikus on 52 498 kahjut. Leedu andmete hulgas on kahjud, mille toimumise aasta jääb vahemikku 2008-2019 ning nende kahjude teatamise aasta jääb vahemikku 2010-2019. Tabelis 5 toome välja Leedu kaskokahjude hüvitiste empiirilise jaotuse arvkarakteristikud. Tabelist näeme, et Leedu andmete väikseim väärtus on $1,07$ eurot ja suurim hüvitis on 72 762 eurot. Keskmine hüvitis on $734,76$ eurot ja mediaan on $326,22$ eurot. Keskmine on mediaanist ligikaudu kaks korda suurem. Standardhälve Leedu kahjude hüvitiste puhul on $1639,06$ eurot, mis on suurem nii Eesti kui ka Läti andmetest. Suur standardhälve väärtus viitab, et kahjude hüvitiste hajuvus on suur. Leedu kahjude järsakuskordaja väärtus on $281,28$ ja asümmetriakordaja on $12,01$, mis viitavad sellele, et jaotusel on pikk parempoolne saba ja mediaan on keskmisest väiksem.

Tabel 5: Leedu kaskokahjude hüvitiste empiirilised karakteristikud.

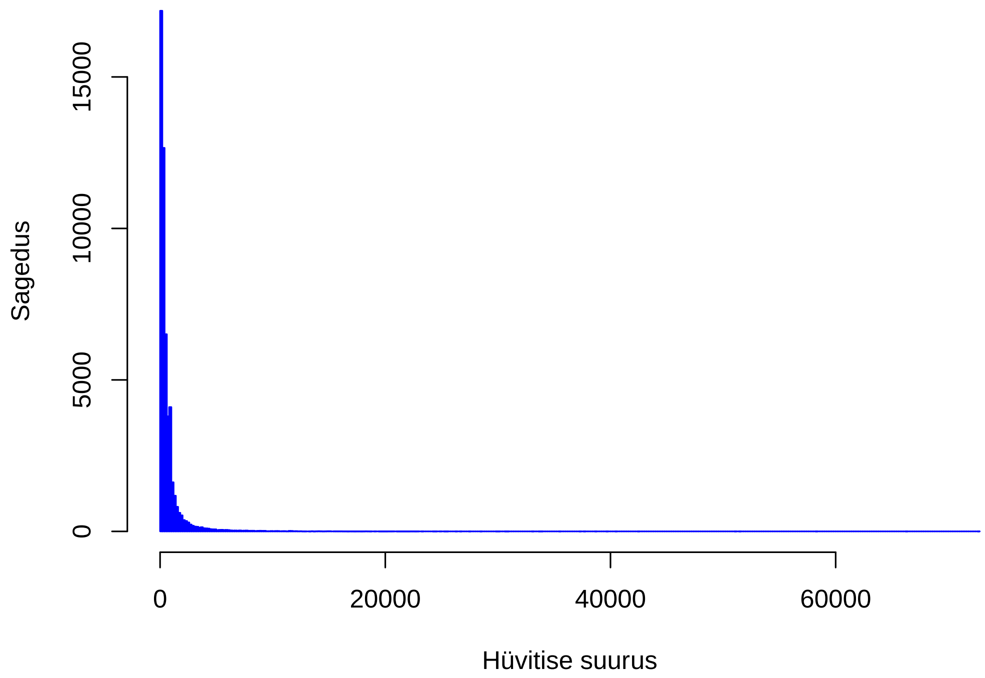
	Miinimum	Maksimum	Keskmine	Mediaan	Standardhälve	Järsakuskordaja	Asümmeetriakordaja
Kahju maksumus	1,07	72 762	734,76	326,22	1639,06	281,28	12,01

Tabelis 6 toome välja Leedu kaskokahjude hüvitiste kvantiilid. Kvantiilide tabelist näeme, et ühe protsendi kvantiili hüvitise väärtuseks on $20,62$ eurot. Üheksakümne protsendi kvantiili hüvitise väärtuseks on $1443,17$ eurot. Üheksakümne viie protsendi kvantiili hüvitise väärtuseks on $2493,68$ eurot. Üheksakümne üheksa protsendi kvantiili hüvitise väärtuseks on $7257,75$ eurot.

Tabel 6: Leedu kaskokahjude hüvitiste kvantiilid.

	Kvantiilid					
	1%	5%	10%	90%	95%	99%
Kahju maksumus	20,62	54,74	88,03	1443,17	2493,68	7257,75

Leedu kaskokahjude hüvitiste kogusumma on 38 573 390 eurot ja 1% kvantiilist väiksemate hüvitiste kogusumma on $6540,78$ eurot. 1% kvantiilist väiksemate hüvitiste kogusumma moodustab $0,017\%$ kõikide hüvitiste kogusummast. Joonisel 8 kujutame Leedu kaskokahjude hüvitiste histogrammi. Jooniselt näeme, et suurem osa hüvitistest on väikesed ning suuremaid kahjude hüvitisi esineb harvemini.



Joonis 8: Leedu kaskokahjude hüvitiste histogramm.

2.4 Võrdlus

Tabelis 7 toome välja kõigi kolme riigi karakteristikud. Tabelist näeme, et Läti kaskokahjude puhul on palju suurem maksimaalne hüvitise väärtus, kui Eesti ja Leedu kaskokahjude puhul. Lisaks näeme, et keskmised ja mediaanid on riikide lõikes üpriski erinevad. Suurim keskmine on Lätil ja väikseim Leedul. Mediaan on Leedu kahjude hüvitiste puhul väiksem kui Eesti ja Läti kaskokahjude hüvitiste puhul. Lisaks näeme, et standardhälbed ja asümmeetriakordajad on kõigi riigi andmete puhul erinevad. Leedu andmete puhul on standardhälve kõige väiksem ja Läti andmete puhul on standardhälve kõige suurem. Läti asümmeetriakordaja ja järsakuskordaja on palju suuremad kui Eestil ja Leedul tänu üksikutele väga suurtele kahjuhüvitistele.

Tabel 7: Hüvitiste empiirilised karakteristikud Balti riikides.

Kahju maksumus	Miinimum	Maksimum	Keskmine	Mediaan	Standardhälve	Järsakuskordaja	Asümmeetriakordaja
Eesti	1,16	56 501,48	958,59	459,81	1701,47	162,86	9,07
Läti	1	171 359,8	1018,7	435,9	2511,4	815,74	18,13
Leedu	1,07	72 762	734,76	326,22	1639,06	281,28	12,01

3 Jaotuste sobitamine ja võrdlus

Modelleerimisel kasutame gammajaotust, log-normaalset, Pareto, Weibulli ja Waldi jaotust. Töös toome välja sobitatud jaotuste parameetrite hinnangud, Kolmogorov-Smirnovi teststatistiku D_n väärtused iga jaotuse puhul, kahjude empiirilise jaotuse hüvitiste väärtuste korral, mis on vahemikus $(0, 10\,000)$ koos sobitatud jaotuste tihedusfunktsiooniga ja kahjude empiirilise jaotuse kõigi hüvitiste väärtuste korral koos kõige paremini sobinud tihedusfunktsiooniga. Modelleerimisel vaatame riikide lõikes nii kõiki kahjusid kui ka juhtu, kus 1% kvantiilist väiksemad hüvitised on välja jäetud. Nagu eespool nägime, moodustab 1% kvantiilist väiksemate hüvitiste kogusumma väikese osa üldisest hüvitiste kogusummast ja firma jaoks ei ole see märkimisväärne suurus.

Kolmogorov-Smirnovi teststatistik D_n avaldub kujul (Kollo, 2004, lk 29):

$$D_n = max(D_n^+, D_n^-)$$

kus

- $D_n^+ = \sup\limits_{-\infty<x<\infty} (F_n(x) - F_X(x))$,
- $D_n^- = \sup\limits_{-\infty<x<\infty} (F_X(x) - F_n(x))$,
- $F_X(x)$ on lähendava jaotuse jaotusfunktsioon,
- $F_n(x) = \frac{\#(x_i|x_i \leq x)}{n}$ on empiiriline jaotusfunktsioon.

3.1 Eesti

Esmalt modelleerime Eesti kaskokahjude hüvitisi. Tabelis 8 toome välja sobitatud jaotuste parameetrite hinnangud.

Tabel 8: Sobitatud jaotuste parameetrite hinnangud Eesti andmetel.

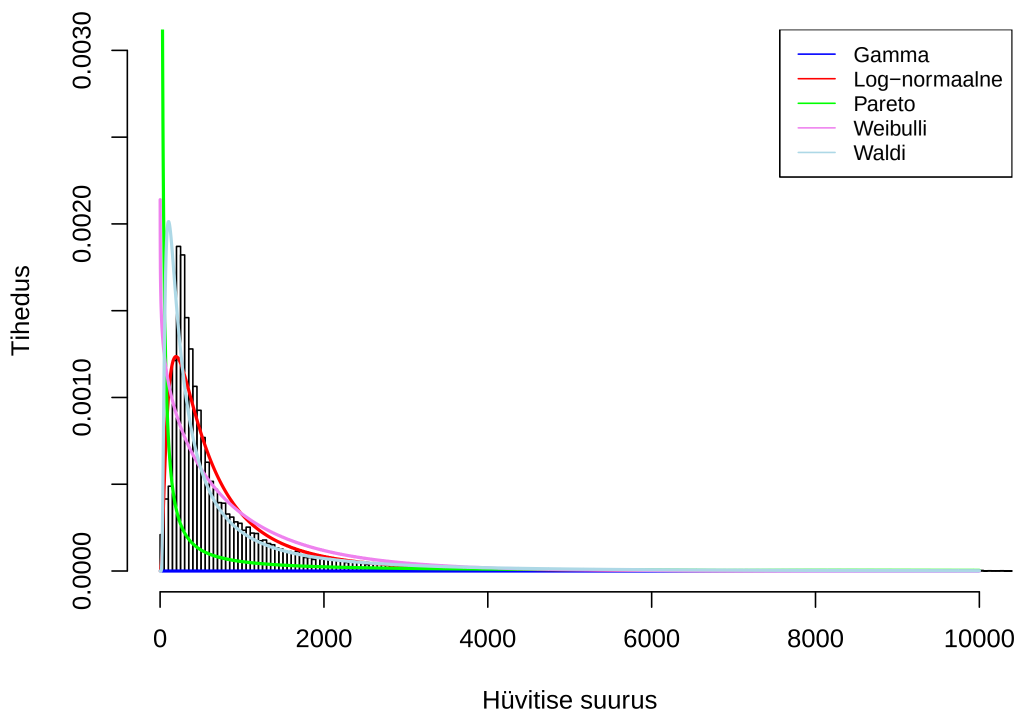
Jaotus	Parameetrite hinnangud	
Gamma	$b = 3020,06$	$c = 0,32$
Log-normaalne	$\mu = 6,28$	$\sigma = 1$
Pareto	$a = 1,16$	$c = 0,16$
Weibulli	$\eta = 895,93$	$\beta = 0,90$
Waldi	$\mu = 958,59$	$\lambda = 313,11$

Tabelis 9 toome välja Kolmogorov-Smirnovi kooskõlatesti teststatistiku D_n väärtuse. Kõige väiksem teststatistiku väärtus on log-normaalsel jaotusel, $D_n = 0,06638$. Suuruselt teine teststatistiku väärtus on Weibulli jaotusel, $D_n = 0,13046$. Suuruselt kolmanda teststatistiku väärtusega on Waldi jaotus, $D_n = 0,1707$. Suuruselt järgmine teststatistiku väärtus on Pareto jaotusel, $D_n = 0,48969$. Suurim teststatistiku väärtus on gammajaotusel, $D_n = 1$. Näeme, et gammajaotus on sobimatu andmete kirjeldamisel, seetõttu järgmisi andmeid me gammajaotusega ei sobita.

Tabel 9: Kooskõlatesti väärtused Eesti kaskokahjude hüvitise suuruse korral.

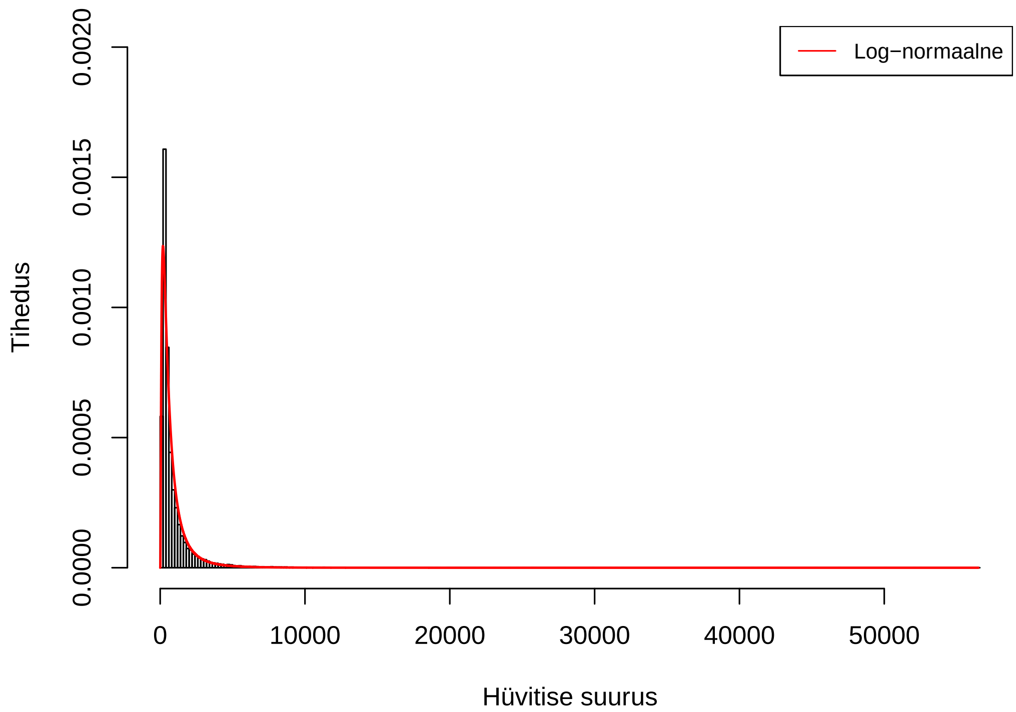
Jaotus	D_n
Gamma	1
Log-normaalne	0,06638
Pareto	0,48969
Weibulli	0,13046
Waldi	0,1707

Joonisel 9 kujutame Eesti kaskokahjude hüvitiste suuruseid, mis on väiksemad kui 10 000 eurot. Uurime 10 000 eurost väiksemaid hüvitisi, sest siis on jooniselt võimalik näha sobitatud jaotuste kujusid. Kui joonisele kanda ka suuremaid väärtusi, siis pole võimalik ühisel joonisel sobitatud jaotusi eristada. Lisaks kujutame joonisel empiirilisele jaotusele sobitatud jaotuste tihedusfunktsioonid. Jooniselt näeme, et visuaalselt kõige täpsemalt kirjeldab hüvitise suuruse empiirilist jaotust log-normaalne jaotus. Kolmogorov-Smirnovi teststatistiku D_n väärtus viitab ka sellele, et log-normaalne jaotus sobitub kõige paremini Eesti koguandmetele.



Joonis 9: Eesti kaskokahjude hüvitiste (kuni 10 000 eurot) empiiriline jaotus ja sobitatud jaotuste tihedusfunktsioonid.

Eelnevalt leidsime, et log-normaalne jaotus sobitub kõige paremini. Koostame joonise (joonis 10), kus kujutame Eesti kaskokahjude hüvitiste suuruste kõik väärtused koos log-normaalse jaotuse tihedusfunktsiooniga.



Joonis 10: Eesti kaskokahjude hüvitiste empiiriline jaotus ja log-normaalse jaotuse tihedusfunktsioon.

Teame, et Eesti kahjud järgivad log-normaalset jaotust parameetritega $\mu = 6,28$ ja $\sigma = 1$. Arvutame statistikatarkvara R abil tõenäosuse, et firmal tuleb välja maksta hüvitis, mille väärtus on suurem kui 20 000 eurot.

$$X \sim (\mu = 6,28; \sigma = 1)$$
$$P(X > 20\,000) = 0,000145$$

Tõenäosus, et firmal tuleb välja maksta hüvitis, mille väärtus on suurem kui 20 000 eurot on $0,000145$.

Uurime kas mudel on optimistlik või konservatiivne. Eesti koguandmestikus on 50 844 kahju ja hüvitisi, mis on suuremad kui 20 000 eurot on 42 tükki. Mudel prognoosib, et 50 844 kahju korral tekib $50\,844 \cdot 0,000145 = 7,37$ ehk 7 kahju, mille hüvitis on suurem, kui 20 000 eurot. Kuna Eesti andmete puhul on aga 42 kahju, mille hüvitise väärtus on suurem kui 20 000 eurot, siis antud mudel on liiga optimistlik.

3.1.1 Eesti kahjud, kui 1% kvantiilist väiksemad hüvitised on eemaldatud

Modelleerime Eesti kaskokahjude hüvitisi, kui 1% kvantiilist väiksemad hüvitised on välja jäetud. 1% kvantiilist väiksemate hüvitiste kogusumma on 17 459,04 eurot. Tabelis 10 toome välja sobitatud jaotuste parameetrite hinnangud.

Tabel 10: Sobitatud jaotuste parameetrite hinnangud Eesti andmetel, kui 1% kvantiilist väiksemad hüvitised on eemaldatud.

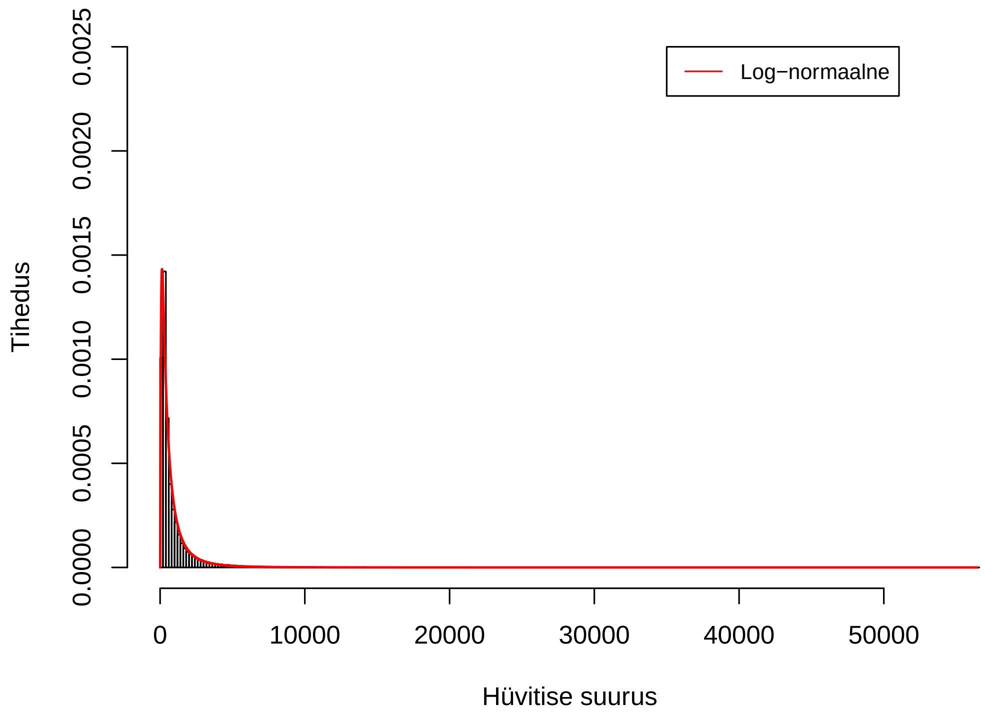
Jaotus	Parameetrite hinnangud	
Log-normaalne	$\mu = 6,14$	$\sigma = 1,13$
Pareto	$a = 0,34$	$c = 0,14$
Weibulli	$\eta = 822,21$	$\beta = 0,85$
Waldi	$\mu = 918,28$	$\lambda = 208,65$

Tabelis 11 toome välja Kolmogorov-Smirnovi kooskõlatesti teststatistiku D_n väärtused. Kõige väiksem Kolmogorov-Smirnovi teststatistiku väärtus on log-normaalsel jaotusel, $D_n = 0,04772$. Suuruselt järgmine väärtus on Weibulli jaotusel, $D_n = 0,11618$. Suuruselt kolmas väärtus on Waldi jaotusel, $D_n = 0,19048$. Suuruselt neljas väärtus on Pareto jaotusel, $D_n = 0,49956$.

Tabel 11: Kooskõlatesti väärtused Eesti andmetel, kui 1% kvantiilist väiksemad kahjud on eemaldatud.

Jaotus	D_n
Log-normaalne	0,04772
Pareto	0,49956
Weibulli	0,11618
Waldi	0,19048

Joonisel 11 kujutame Eesti kaskokahjude hüvitisi, kui andmetest on eemaldatud 1% kvantiilist väiksemad hüvitised, koos log-normaalse jaotuse tihedusfunktsiooniga.



Joonis 11: Eesti kaskokahjude hüvitiste empiiriline jaotus, kui 1% kvantiilist väiksemad hüvitised on eemaldatud, ja log-normaalse jaotuse tihedusfunktsioon.

3.2 Läti

Modelleerime Läti kaskokahjude hüvitisi. Tabelis 12 toome välja sobitatud jaotuste parameetrite hinnangud.

Tabel 12: Sobitatud jaotuste parameetrite hinnangud Läti andmetel.

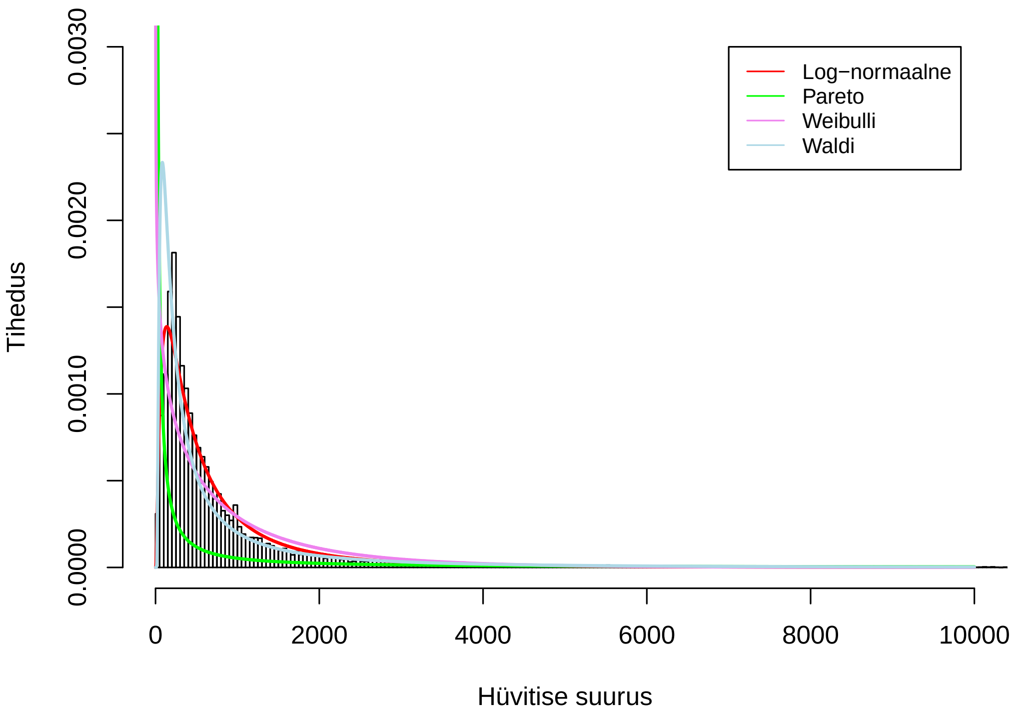
Jaotus	Parameetrite hinnangud	
Log-normaalne	$\mu = 6,17$	$\sigma = 1,12$
Pareto	$a = 1$	$c = 0,16$
Weibulli	$\eta = 856,3$	$\beta = 0,80$
Waldi	$\mu = 1018,68$	$\lambda = 251,23$

Tabelis 13 toome välja Kolmogorov-Smirnovi kooskõlatesti teststatistiku D_n väärtused Läti andmete põhjal. Kõige väiksem teststatistiku väärtus on log-normaalse jaotuse puhul, $D_n = 0,04052$, suuruselt järgmine väärtus on Weibulli jaotusel, $D_n = 0,11068$. Suuruselt järgmine teststatistik on Waldi jaotusel, $D_n = 0,14232$. Suurim teststatistiku väärtus on Pareto jaotusel, $D_n = 0,46877$. Teststatistiku põhjal sobib Läti andmetele kõige paremini log-normaalne jaotus.

Tabel 13: Kooskõlatesti väärtused Läti kaskokahjude hüvitise suuruse korral.

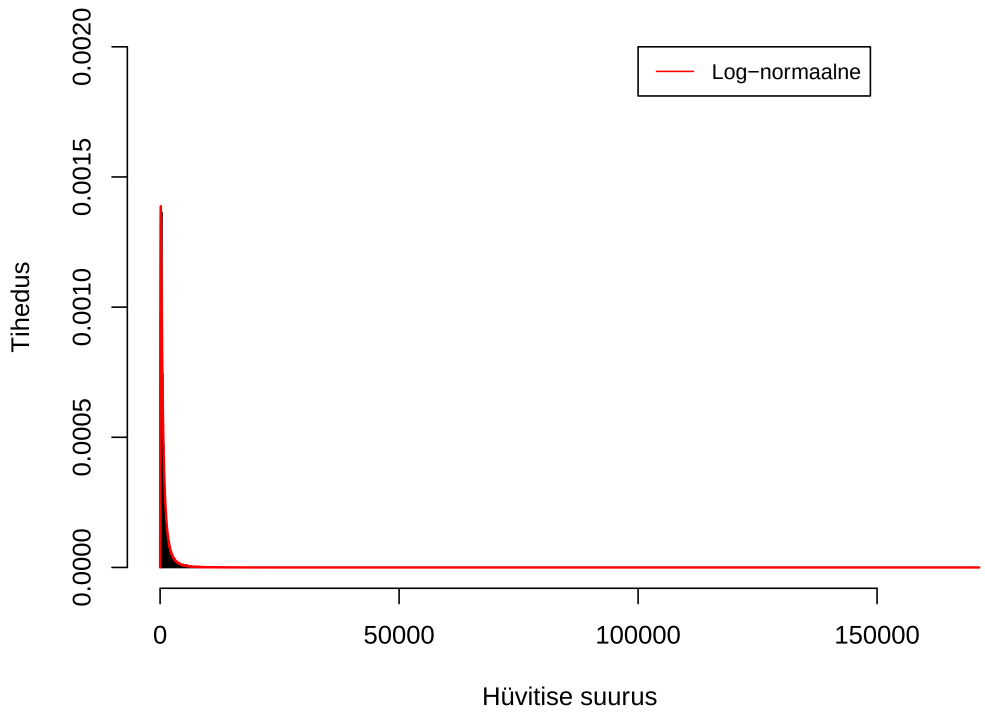
Jaotus	D_n
Log-normaalne	0,04052
Pareto	0,46877
Weibulli	0,11068
Waldi	0,14232

Joonisel 12 kujutame Läti kaskokahjude hüvitiste suuruseid, mis on väiksemad kui 10 000 eurot. Lisaks kujutame joonisel empiirilisele jaotusele sobitatud jaotuste tihedusfunktsioonid. Jooniselt 12 näeme, et visuaalselt kõige täpsemalt kirjeldab hüvitise suuruse empiirilist jaotust log-normaalne jaotus.



Joonis 12: Läti kaskokahjude hüvitiste (kuni 10 000 eurot) empiiriline jaotus ja sobitatud jaotuste tihedusfunktsioonid.

Kooskõlatesti teststatistiku ja visuaalsel hinnangul otsustame, et log-normaalne jaotus sobitub kõige paremini Läti andmetele. Koostame joonise (joonis 13), kus on kujutatud Läti kaskokahjude hüvitiste suuruste kõik väärtused koos log-normaalse jaotuse tihedusfunktsiooniga.



Joonis 13: Läti kaskokahjude hüvitiste empiiriline jaotus ja log-normaalse jaotuse tihedusfunktsioon.

Teame, et Läti kahjud järgivad log-normaalset jaotust parameetritega $\mu = 6,17$ ja $\sigma = 1,12$. Arvutame statistikatarkvara R abil tõenäosuse, et firmal tuleb välja maksta hüvitis, mille väärtus on suurem kui 20 000 eurot.

$$X \sim (\mu = 6,17; \sigma = 1,12)$$
$$P(X > 20\,000) = 0,0000944$$

Tõenäosus, et firmal tuleb välja maksta hüvitis, mille väärtus on suurem kui 20 000 eurot on $0,0000944$.

Uurime kas mudel on optimistlik või konservatiivne. Läti koguandmestikus on 30 087 kahju ja hüvitisi, mis on suuremad kui 20 000 eurot on 94 tükki. Mudel prognoosib, et 30 087 kahju korral tekib $30\,087 \cdot 0,0000944 = 2,84$ ehk 3 kahju, mille hüvitis on suurem, kui 20 000 eurot. Kuna Läti andmete korral esineb hoopis 94 kahju, mille hüvitise väärtus on suurem kui 20 000 eurot, siis antud mudel on liiga optimistlik.

3.2.1 Läti kahjud, kui 1% kvantiilist väiksemad hüvitised on eemaldatud.

Modelleerime Läti kahjusid juhul, kui 1% kvantiilist väiksemad hüvitised on välja jäetud. Tabelis 14 toome välja sobitatud jaotuste parameetrite hinnangud.

Tabel 14: Sobitatud jaotuste parameetrite hinnangud Läti andmetel, kui 1% kvantiilist väiksemad kahjud on eemaldatud

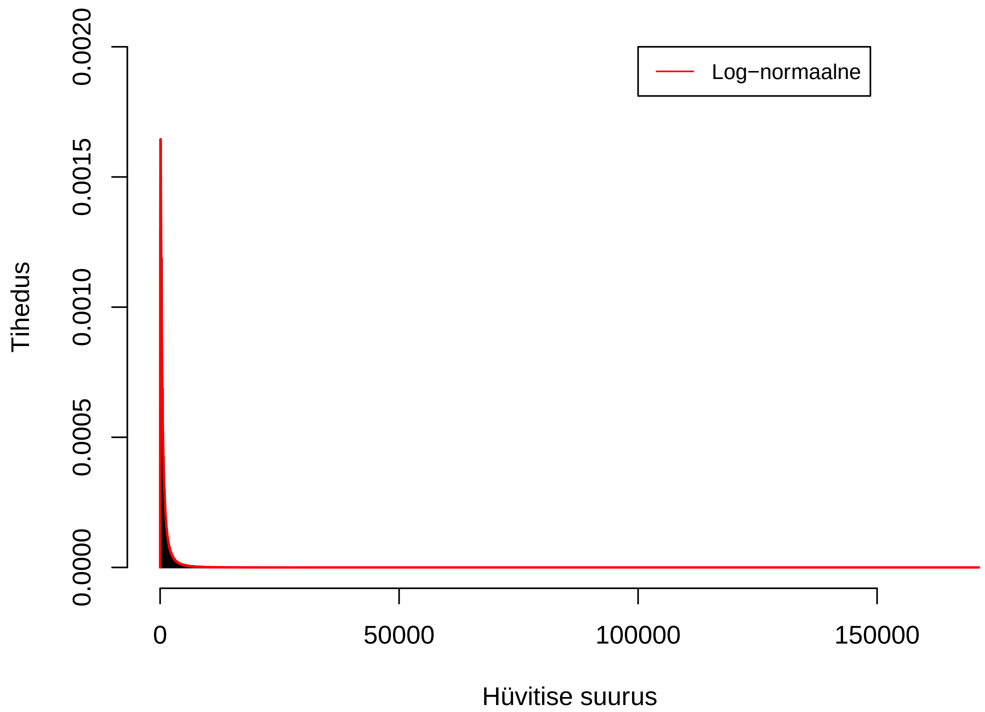
Jaotus	Parameetrite hinnangud	
Log-normaalne	$\mu = 6,05$	$\sigma = 1,25$
Pareto	$a = 0,28$	$c = 0,14$
Weibulli	$\eta = 791,51$	$\beta = 0,76$
Waldi	$\mu = 988,8$	$\lambda = 160,65$

Tabelis 15 toome välja Kolmogorov-Smirnovi kooskõlatesti teststatistiku D_n väärtused Läti andmete põhjal. Kõige väiksem teststatistiku väärtus on log-normaalse jaotuse puhul, $D_n = 0,04157$, suuruselt järgmine väärtus on Weibulli jaotusel, $D_n = 0,09498$. Suuruselt kolmas teststatistik on Waldi jaotusel, $D_n = 0,19249$. Kõige suurem teststatistiku väärtus on Pareto jaotusel, $D_n = 0,46841$. Teststatistiku põhjal sobib Läti andmetele, kui 1% kvantiilist väiksemad kahjud on välja jäetud, kõige paremini log-normaalne jaotus.

Tabel 15: Kooskõlatesti väärtused Läti kaskokahjude hüvitise suuruse korral, kui 1% kvantiilist väiksemad kahjud on eemaldatud.

Jaotus	D_n
Log-normaalne	0,04157
Pareto	0,46841
Weibulli	0,09498
Waldi	0,19249

Koostatame joonise (joonis 14), kus on kujutatud Läti kaskokahjude hüvitise suuruse kõik väärtused, kui 1% kvantiilist väiksemad kahjud on eemaldatud, koos log-normaalse jaotuse tihedusfunktsiooniga.



Joonis 14: Läti kaskokahjude hüvitiste empiiriline jaotus, kui 1% kvantiilist väiksemad kahjud on eemaldatud, ja log-normaalse jaotuse tihedusfunktsioon.

3.3 Leedu

Järgnevalt uurime Leedu kaskokahjude hüvitisi. Tabelis 16 toome välja sobitatud jaotuste parameetrite hinnangud.

Tabel 16: Sobitatud jaotuste parameetrite hinnangud Leedu andmetel.

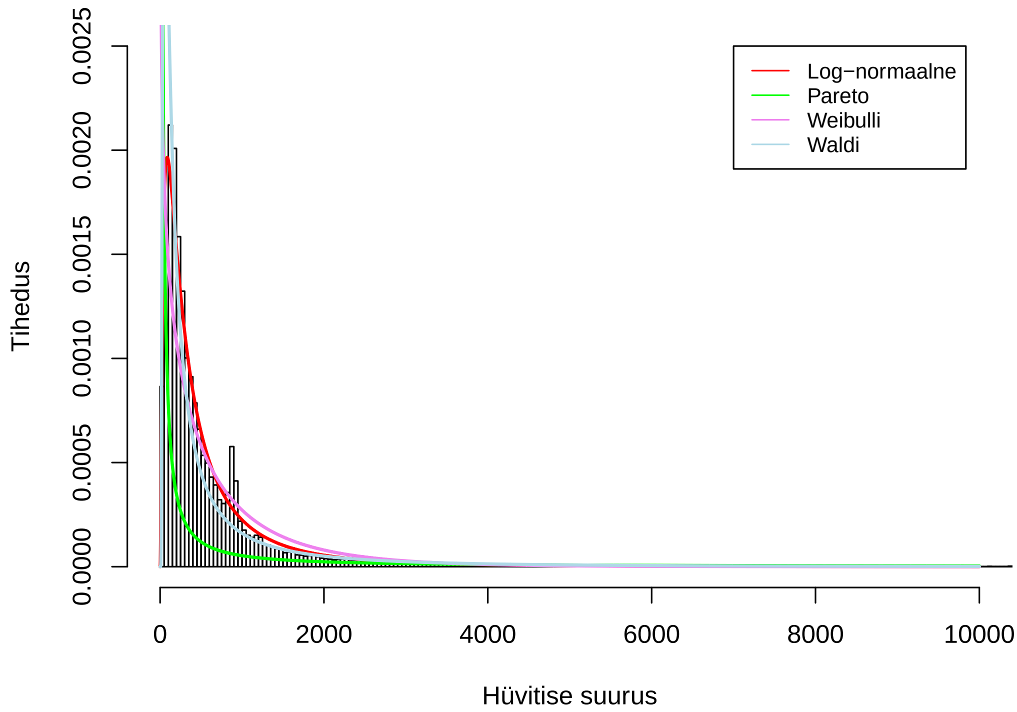
Jaotus	Parameetrite hinnangud	
Log-normaalne	$\mu = 5,84$	$\sigma = 1,17$
Pareto	$a = 1,07$	$c = 0,17$
Weibulli	$\eta = 623,12$	$\beta = 0,8$
Waldi	$\mu = 734,76$	$\lambda = 164,46$

Tabelis 17 toome välja Kolmogorov-Smirnovi kooskõlatesti teststatistiku D_n väärtused Leedu andmete põhjal. Kõige väiksem teststatistiku väärtus on log-normaalse jaotuse puhul, $D_n = 0,02617$, suuruselt järgmine väärtus on Weibulli jaotusel, $D_n = 0,09105$. Suuruselt järgmine teststatistiku väärtus on Waldi jaotusel, $D_n = 0,1406$. Suurim teststatistiku väärtus on Pareto jaotusel, $D_n = 0,44562$. Teststatistiku põhjal sobib Leedu andmetele kõige paremini log-normaalne jaotus, sest tal on väikseim teststatistiku väärtus.

Tabel 17: Kooskõlatesti väärtused Leedu kaskokahjude hüvitise suuruse korral.

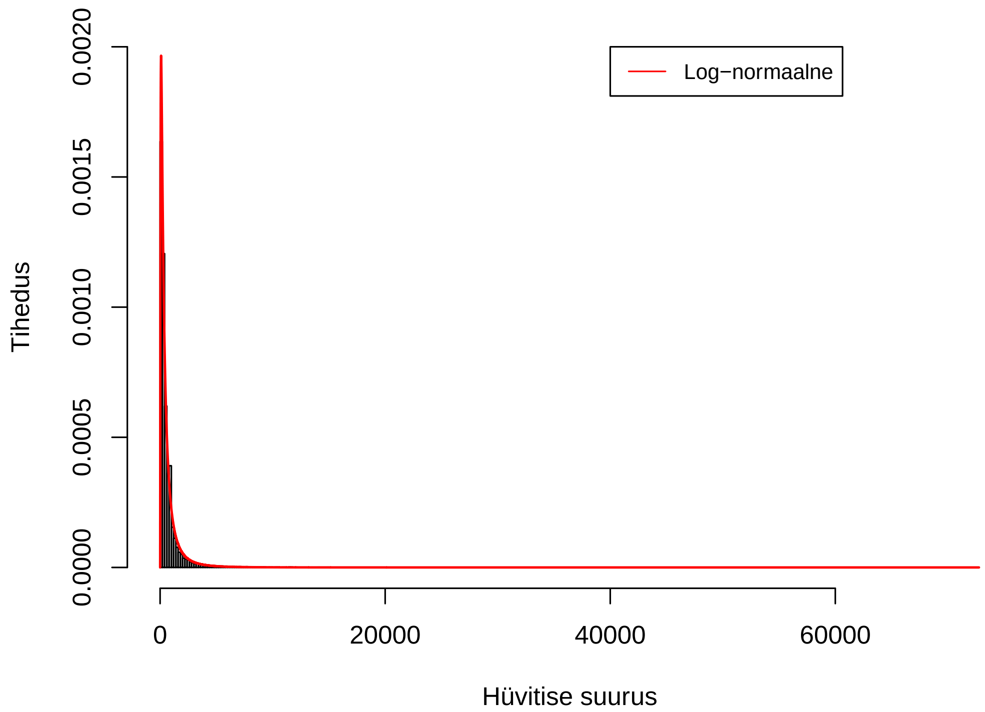
Jaotus	D_n
Log-normaalne	0,02617
Pareto	0,44562
Weibulli	0,09105
Waldi	0,1406

Joonisel 15 kujutame Leedu kaskokahjude hüvitiste suuruseid, mis on väiksemad kui 10 000 eurot. Lisaks kujutame joonisel empiirilisele jaotusele sobitatud jaotuste tihedusfunktsioonid. Jooniselt 15 näeme, et visuaalselt kõige täpsemalt kirjeldab hüvitise suuruse empiirilist jaotust log-normaalne jaotus.



Joonis 15: Leedu kaskokahjude hüvitiste (kuni 10 000 eurot) empiiriline jaotus ja sobitatud jaotuste tihedusfunktsioonid.

Joonise 15 ja tabeli 17 põhjal otsustasime, et kõige paremini sobitub Leedu kaskokahjude hüvitise empiirilisele jaotusele log-normaalse jaotuse tihedusfunktsioon. Seega koostame joonise (joonis 16), kus on kujutatud Leedu kaskokahjude hüvitise suuruse kõik väärtused koos log-normaalse jaotuse tihedusfunktsiooniga.



Joonis 16: Leedu kaskokahjude hüvitiste empiiriline jaotus ja log-normaalse jaotuse tihedusfunktsioon.

Teame, et Leedu kahjud järgivad log-normaalset jaotust parameetritega $\mu = 5,84$ ja $\sigma = 1,17$. Arvutame statistikatarkvara R abil tõenäosuse, et firmal tuleb välja maksta hüvitis, mille väärtus on suurem kui 20 000 eurot.

$$X \sim (\mu = 5,84; \sigma = 1,17)$$
$$P(X > 20\,000) = 0,000024$$

Tõenäosus, et firmal tuleb välja maksta hüvitis, mille väärtus on suurem kui 20 000 eurot on $0,000024$.

Uurime kas mudel on optimistlik või konservatiivne. Leedu koguandmestik sisaldab 52 498 kahju ja hüvitisi, mis on suuremad kui 20 000 eurot on 51 tükki. Mudel prognoosib, et 52 498 kahju korral tekib $52\,498 \cdot 0,000024 = 1,26$ ehk 1 kahju, mille hüvitis on suurem, kui 20 000 eurot. Kuna Leedu andmete korral esineb hoopis 51 kahju, mille hüvitise väärtus on suurem kui 20 000 eurot, siis antud mudel on liiga optimistlik.

3.3.1 Leedu kahjud, kui 1% kvantiilist väiksemad hüvitised on eemaldatud

Uurime Leedu kaskokahjude hüvitisi juhul, kui 1% kvantiilist väiksemad hüvitised on eemaldatud. Tabelis 18 toome välja sobitatud jaotuste parameetrite hinnangud.

Tabel 18: Sobitatud jaotuste parameetrite hinnangud Leedu andmetel, kui 1% kvaniilist väiksemad hüvitised on eemaldatud.

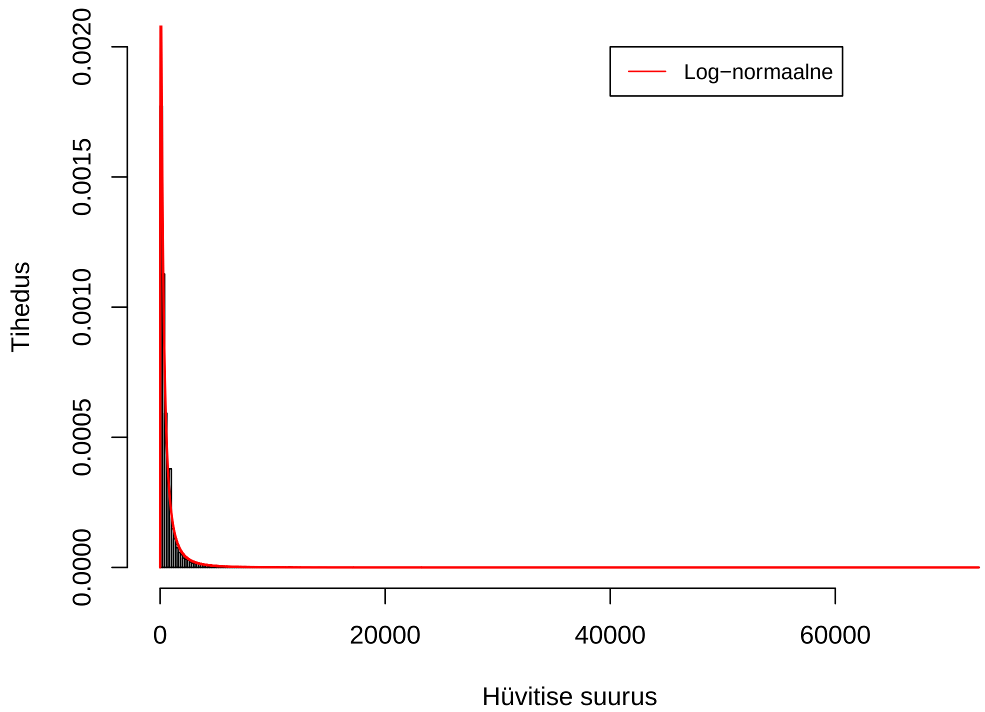
Jaotus	Parameetrite hinnangud	
Log-normaalne	$\mu = 5,76$	$\sigma = 1,26$
Pareto	$a = 0,64$	$c = 0,16$
Weibulli	$\eta = 593,54$	$\beta = 0,78$
Waldi	$\mu = 722,07$	$\lambda = 121,52$

Tabelis 19 toome välja Kolmogorov-Smirnovi kooskõlatesti teststatistiku D_n väärtused Leedu andmete põhjal. Kõige väiksem teststatistiku väärtus on log-normaalse jaotuse puhul, $D_n = 0,03081$, suuruselt järgmine väärtus on Weibulli jaotusel, $D_n = 0,08101$. Suuruselt järgmine teststatistik on Waldi jaotusel, $D_n = 0,17381$. Suurim teststatistiku väärtus on Pareto jaotusel, $D_n = 0,4447$. Teststatistiku põhjal sobib Leedu andmetele juhul, kui 1% kvantiilist väiksemad kahjud on välja jäetud, kõige rohkem log-normaalne jaotus.

Tabel 19: Kooskõlatesti väärtused Leedu kaskokahjude hüvitise suuruse korral, kui 1% kvantiilist väiksemad kahjud on eemaldatud.

Jaotus	D_n
Log-normaalne	0,03081
Pareto	0,4447
Weibulli	0,08101
Waldi	0,17381

Tabeli 19 põhjal otsustame, et kõige paremini sobitub Leedu kaskokahjude hüvitiste empiirilisele jaotusele log-normaalse jaotuse tihedusfunktsioon. Koostame joonise (joonis 17), kus on kujutatud Leedu kaskokahjude hüvitised, kui 1% kvantiilist väiksemad kahjud on eemaldatud, ja log-normaalne jaotus.



Joonis 17: Leedu kaskokahjude hüvitiste empiiriline jaotus (kui 1% kvantiilist väiksemad kahjud on eemaldatud) ja log-normaalse jaotuse tihedusfunktsioon.

3.4 Riikide vaheline võrdlus

Koostatud mudelite põhjal näeme, et log-normaalne jaotus sobis kõige paremini nii iga riigi koguandmetele kui ka juhule, kui 1% kvantiilist väiksemad hüvitised on välja jäetud.

Võrreldes Eesti koguandmeid ja andmeid, kui 1% kvantiilist väiksemad väärtused on eemaldatud, siis nende andmete põhjal modelleerides väga suuri erinevusi ei esine. Empiirilise jaotuse ja log-normaalse jaotuse tihedusfunktsiooni joonised on kujult sarnased. Eesti koguandmete puhul oli Kolmogorov-Smirnovi teststatistiku väärtus log-normaalse jaotuse puhul $D_n = 0,06638$. Eesti andmete puhul, kui 1% kvantiilidest väiksemad hüvitised on eemaldatud oli Kolmogorov-Smirnovi teststatistiku väärtus log-normaalse jaotuse puhul $D_n = 0,04772$. Sellest võib järeldada, et log-normaalne jaotus sobitub paremini andmetele, kui 1% kvantiilist väiksemad kahjud on eemaldatud, mis ei ole firma seisukohalt olulised.

Võrdleme Läti koguandmeid ja andmeid, kui 1% kvantiilist väiksemad väärtused on eemaldatud, siis nende andmete põhjal modelleerides väga suuri erinevusi ei esine. Empiirilise jaotuse ja log-normaalse jaotuse tihedusfunktsiooni joonised on sarnased. Läti koguandmete puhul oli Kolmogorov-Smirnovi teststatistiku väärtus log-normaalse jaotuse puhul $D_n = 0,040516$. Juhul, kui 1% kvantiilist väiksemad hüvitised on eemaldatud oli Kolmogorov-Smirnovi teststatistiku väärtus log-normaalse jaotuse puhul $D_n = 0,041566$. Sellest järeldame, et log-normaalne jaotus sobitub paremini Läti koguandmetele.

Kui võrrelda Leedu koguandmeid ja andmeid, kui 1% kvantiilist väiksemad väärtused on eemaldatud, siis nende andmete põhjal modelleerides suuri erinevusi ei esine. Empiirilise jaotuse ja log-normaalse jaotuse tihedusfunktsiooni joonised on kujult sarnased. Leedu koguandmete puhul oli Kolmogorov-Smirnovi teststatistiku väärtus log-normaalse jaotuse puhul $D_n = 0,02617$. Andmetel, kui 1% kvantiilist väiksemad hüvitised on eemaldatud, oli Kolmogorov-Smirnovi teststatistiku väärtus log-normaalse jaotuse puhul $D_n = 0,03081$. Leedu andmete puhul sobitub log-normaalne jaotus paremini koguandmetele.

Tabelis 20 toome iga riigi koguandmete puhul välja tõenäosused, et firmal tuleb välja maksta hüvitis, mille väärtus on suurem kui 20 000 eurot log-normaalse jaotuse korral. Näeme, et 20 000 eurost suuremate hüvitiste esinemise tõenäosus on suurim Eesti andmetel. Väikseim tõenäosus on Leedu andmete korral.

Tabel 20: Tõenäosus, et firmal tuleb välja maksta hüvitis, mille väärtus on suurem kui 20 000 eurot iga riigi kohta log-normaalse jaotuse korral.

Riik	$P(X > 20\,000$
Eesti	0,000145
Läti	0,000094
Leedu	0,000024

Kokkuvõte

Antud bakalaureusetöö eesmärgiks oli uurida kindlustusettevõtte kaskokahjude jaotust kolmes Balti riigis ja tuua välja võimalikud erinevused. Esimeses osas andsime ülevaate töös kasutavatest jaotustest, esitasime nende olulisemad karakteristikud ja tuletasime parameetrite hinnangud. Teises osa kirjeldasime kindlustusettevõttest saadud andmeid kolme riigi lõikes. Kolmandas peatükis lähendasime kolme riigi kahjuandmete empiirilisi jaotusi esimeses peatükis kirjeldatud tõenäosusjaotustega. Iga riigi puhul modelleerisime empiirilist jaotust nii koguandmete puhul kui ka juhul, kui 1% kvantiilist väiksemad kahjud olid andmetest eemaldatud. Jaotuste rakendamisel tõime töös välja sobitatud jaotuste parameetrid. Jaotuste sobivust võrdlesime Kolmogorov-Smirnovi testi abil. Lisaks kujutasime graafiliselt empiirilist jaotust sobitatud jaotuste tihedusfunktsioonidega ja ka kõige paremini sobinud log-normaalse jaotuse tihedusfunktsiooniga.

Modelleerimise käigus leidsime, et Balti riigi andmetele sobitub kõige paremini log-normaalne jaotus. Leidsime, et Eesti andmete puhul olid tulemused mõnevõrra paremad juhul, kui 1% kvantiilist väiksemad kahjud olid eemaldatud. Läti ja Leedu andmete puhul olid tulemused paremad koguandmete korral. Üldiselt võib öelda, et Balti riikide kahjud käituvad kõik sarnaselt, sest iga riigi andmeid kirjeldas kõige paremini log-normaalne jaotus. Tulemustest saime teada, et kui kasutada log-normaalset jaotust, siis suurte kahjude esinemise tõenäosus on suurim Eesti koguandmete puhul ja väikseim Leedu koguandmete puhul.

Töös uurisime, kuidas log-normaalne jaotus prognoosib üle 20 000 euro maksvaid hüvitisi. Selgus, et iga riigi korral oli mudel liiga optimistlik. Kuna log-normaalse jaotuse mudel oli liiga optimistlik ning firma võiks eelistada konservatiivsemat mudelit, siis uurisime, kuidas teised jaotused suuri kahjusid prognoosivad. Saime teada, et Waldi jaotuse kahekordsed prognoosid on ligikaudu kooskõlas seniste andmetega. Kui firma soovib väga konservatiivseid hinnanguid, siis peaks ta kasutama Pareto jaotust, kuid sellisel juhul on firma liiga ettevaatlik. Weibulli jaotuse prognoosid olid liiga optimistlikud.

Kasutatud kirjandus

Forbes, C., Evans, M., Hastings, N., Peacock, B. (2011). *Statistical Distributions.* Fourth edition. New York: John Wiley & Sons, Inc.

Johnson, N., Kotz, S., Balakrishnan, N. (1994). *Continuous Univariate Distributions.* Volume 1, 2nd Edition. New York: John Wiley & Sons, Inc.

Kollo, T. (2004). *Monte Carlo meetodid.* Tartu: Tartu Ülikooli Kirjastus.

Mina, Kelli Kukk,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose "Kindlustuskahjude modelleerimine Balti riikides", mille juhendaja on Tõnu Kollo, reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.

2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commonsi litsentsiga CC BY NC ND 3.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.

3. Olen teadlik, et punktis 1 ja 2 nimetatud õigused jäävad alles ka autorile.

4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Kelli Kukk
12.05.2019