Chapter 3 Modeling Loss Severity

A short course authored by the Actuarial Community

3.1 Basic Distributional Quantities

Overview

In this section, we learn how to define some basic distributional quantities:

- 3.1.1 moments,
- 3.1.2 percentiles, and
- 3.1.3 generating functions.

Moments – Raw Moments

- Let X be a continuous random variable with probability density function (*pdf*) $f_X(x)$ and distribution function $F_X(x)$.
- The k-th raw moment of X, denoted by μ'_k, is the expected value of the k-th power of X, provided it exists.
- The first raw moment μ'_1 is the mean of X usually denoted by μ.
- The formula for μ'_k is given as

$$\mu'_k = \mathrm{E}\left(X^k\right) = \int_0^\infty x^k f_X(x)\,dx.$$

Moments – Central Moments

- The k-th central moment of X, denoted by μ_k, is the expected value of the k-th power of the deviation of X from its mean μ.
- The formula for μ_k is given as

$$\mu_k = \mathrm{E}\left[(X-\mu)^k\right] = \int_0^\infty (x-\mu)^k f_X(x)\, dx.$$

- The second central moment μ_2 defines the variance of X, denoted by σ^2. The square root of the variance is the standard deviation σ.

Moments – Skewness and Kurtosis

- The ratio of the third central moment to the cube of the standard deviation (μ_3/σ^3) defines the coefficient of skewness which is a measure of symmetry.
- A positive coefficient of skewness indicates that the distribution is skewed to the right (positively skewed).
- The ratio of the fourth central moment to the fourth power of the standard deviation (μ_4/σ^4) defines the coefficient of kurtosis.
- The normal distribution has a coefficient of kurtosis of 3.
- Distributions with a coefficient of kurtosis greater than 3 have heavier tails than the normal, whereas distributions with a coefficient of kurtosis less than 3 have lighter tails and are flatter.

Quantiles

- When the distribution of X is continuous, for a given fraction $0 \leq p \leq 1$ the corresponding quantile is the solution of the equation
$$F_X(\pi_p) = p.$$
- For example, the middle point of the distribution, $\pi_{0.5}$, is the median.
- A percentile is a type of quantile; a $100p$ percentile is the number such that $100 \times p$ percent of the data is below it.

Skewness, Mean, and Median

- The relationship between mean and median under different skewness[1].

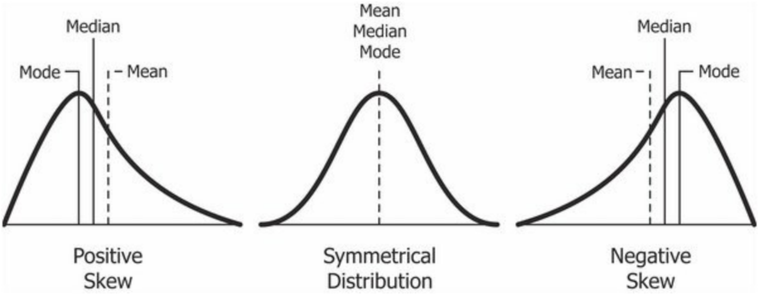


Figure 1: Different Skewness

[1] By Diva Jain - https://codeburst.io/2-important-statistics-terms-you-need-to-know-in-data-science-skewness-and-kurtosis-388fef94eeaa, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=84219892

Moment Generating Function

- The moment generating function (mgf), denoted by $M_X(t)$ uniquely characterizes the distribution of X.
- The moment generating function is given by

$$M_X(t) = \mathrm{E}\left(e^{tX}\right) = \int_0^\infty e^{\mathrm{tx}} f_X\left(x\right) dx$$

  for all t for which the expected value exists.
- The *mgf* is a real function whose k-th derivative at zero is equal to the k-th raw moment of X. In symbols, this is

$$\left.\frac{d^k}{dt^k} M_X(t)\right|_{t=0} = \mathrm{E}\left(X^k\right).$$

Probability Generating Function

- One can also use the moment generating function to compute the probability generating function

$$P_X(z) = \mathrm{E}\left(z^X\right) = M_X\left(\log z\right).$$

- The probability generating function is more useful for discrete random variables.

Exercise for 3.1.1

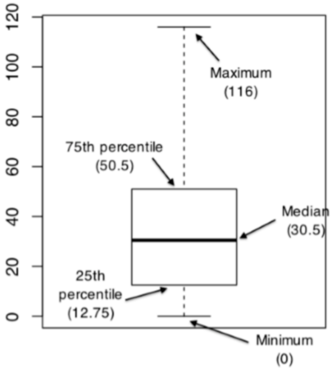
- Data: Anscombe's quartet, `anscombe`
- Variables: `y1`, `y2`, `y3`, and `y4`
- R functions: `summary`, `mean`, `var`, and `sd`

Exercise for 3.1.2

- Data: Wisconsin Property Fund data, `Insample_nz`
- Variables: It includes claim values `y` as well as the claim year `Year`. For this exercise we will work with the natural logarithm of the claim values which are in the `lny` variable. We have filtered the data to exclude zero claims.
- R functions: `boxplot` and `quantile`

Boxplot

Below is an illustration of a boxplot[2].



[2]Boxplots. (2020, August 11). Retrieved June 15, 2021, from https://stats.libretexts.org/@go/page/3976 “Boxplots” by Danielle Navarro, LibreTexts is licensed under CC BY-SA .

Exercise for 3.1.3

- Moment generating function for a gamma distribution
- R functions: `parse`, `class`, `eval`, `D`

Review

In this section, we learnt how to define some basic distributional quantities:

- moments,
- percentiles, and
- generating functions.

Chapter 3 Modeling Loss Severity

A short course authored by the Actuarial Community

3.2 Continuous Distributions for Modeling Loss Severity

Overview

In this section, you learn how to define and apply four fundamental severity distributions:

- gamma,
- Pareto,
- Weibull, and
- generalized beta distribution of the second kind.

Gamma Distribution (1)

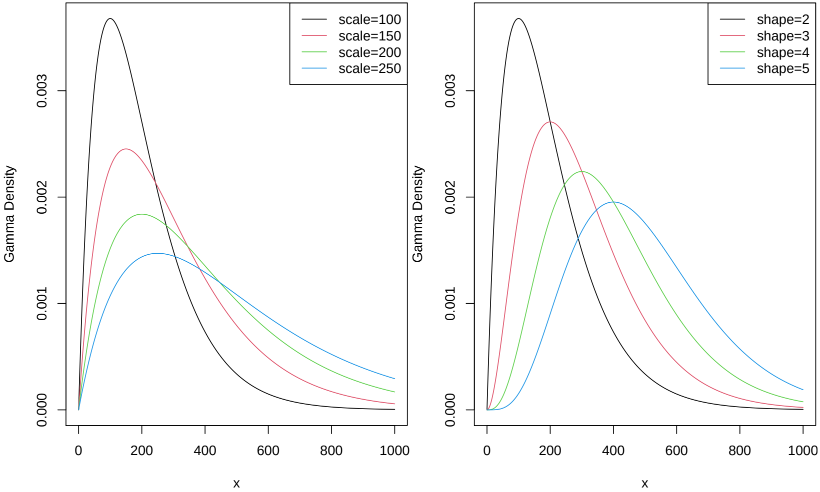
- The continuous variable X is said to have the gamma distribution with shape parameter α and scale parameter θ if its probability density function is given by

$$f_X(x) = \frac{(x/\theta)^{\alpha}}{x\,\Gamma(\alpha)} \exp(-x/\theta) \quad \text{for } x > 0.$$

  Note that $\alpha > 0,\ \theta > 0$.
- When $\alpha = 1$ the gamma reduces to an exponential distribution.
- When $\alpha = \frac{n}{2}$ and $\theta = 2$ the gamma reduces to a chi-square distribution with n degrees of freedom.

Gamma Distribution (2)

Gamma Densities. The left-hand panel is with shape=2 and varying scale. The right-hand panel is with scale=100 and varying shape.



Gamma Distribution (3)

- The distribution function of the gamma model is the *incomplete gamma function*, denoted by $\Gamma\left(\alpha; \frac{x}{\theta}\right)$, and defined as

$$F_X(x) = \Gamma\left(\alpha; \frac{x}{\theta}\right) = \frac{1}{\Gamma(\alpha)} \int_0^{x/\theta} t^{\alpha-1} e^{-t}\, dt,$$

with $\alpha > 0,\ \theta > 0$. For an integer α, it can be written as $\Gamma\left(\alpha; \frac{x}{\theta}\right) = 1 - e^{-x/\theta} \sum_{k=0}^{\alpha-1} \frac{(x/\theta)^k}{k!}$.

Gamma Distribution (4)

- The k-th raw moment of the gamma distributed random variable for any positive k is given by

$$\mathrm{E}\left(X^k\right) = \theta^k \frac{\Gamma\left(\alpha + k\right)}{\Gamma\left(\alpha\right)}.$$

- The mean and variance are given by $\mathrm{E}\left(X\right) = \alpha\theta$ and $\mathrm{Var}\left(X\right) = \alpha\theta^2$, respectively.
- Since all moments exist for any positive k, the gamma distribution is considered a light tailed distribution.

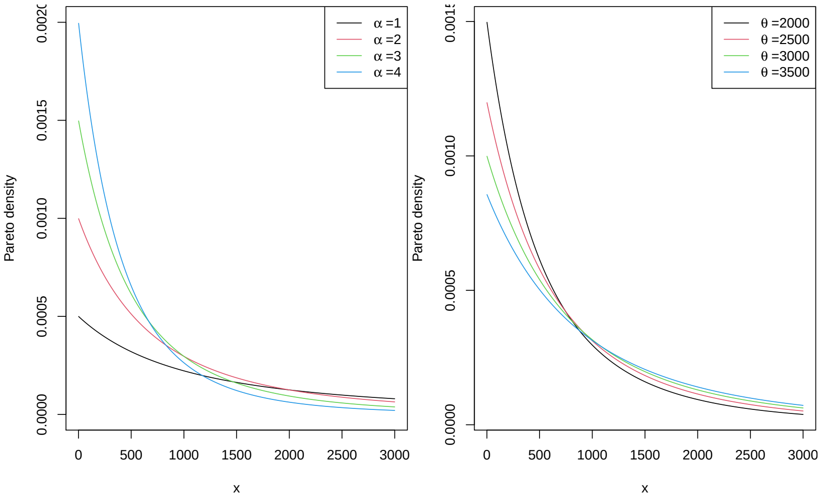
Pareto Distribution (1)

- A positively skewed and heavy-tailed distribution
- For extreme insurance claims, the tail of the severity distribution (losses in excess of a threshold) can be modeled using a Generalized Pareto distribution.
- The continuous variable X is said to have the (two parameter) Pareto distribution with shape parameter α and scale parameter θ if its pdf is given by

$$f_X(x) = \frac{\alpha\theta^{\alpha}}{(x+\theta)^{\alpha+1}} \quad x > 0,\ \alpha > 0,\ \theta > 0.$$

Pareto Distribution (2)

Pareto Densities. The left-hand panel is with scale=2000 and varying shape. The right-hand panel is with shape=3 and varying scale.



Pareto Distribution (3)

- The distribution function of the Pareto distribution is given by

$$F_X(x) = 1 - \left(\frac{\theta}{x+\theta}\right)^{\alpha} \quad x > 0,\ \alpha > 0,\ \theta > 0.$$

- The k-th raw moment of the Pareto distributed random variable exists, if and only if, $\alpha > k$. If k is a positive integer then

$$\mathrm{E}\left(X^k\right) = \frac{\theta^k\ k!}{(\alpha-1)\cdots(\alpha-k)} \quad \alpha > k.$$

- The mean and variance are given by

$$\mathrm{E}(X) = \frac{\theta}{\alpha-1} \quad \text{for } \alpha > 1$$

and

$$\mathrm{Var}(X) = \frac{\alpha\theta^2}{(\alpha-1)^2(\alpha-2)} \quad \text{for } \alpha > 2,$$

respectively.

Weibull Distribution (1)

- The continuous variable X is said to have the Weibull distribution with shape parameter α and scale parameter θ if its *pdf* is given by

$$f_X(x) = \frac{\alpha}{\theta}\left(\frac{x}{\theta}\right)^{\alpha-1} \exp\left(-\left(\frac{x}{\theta}\right)^{\alpha}\right) \quad x > 0,\ \alpha > 0,\ \theta > 0.$$

- The shape parameter α describes the shape of the hazard function of the Weibull distribution. The hazard function is a decreasing function when $\alpha < 1$ (heavy tailed distribution), constant when $\alpha = 1$ and increasing when $\alpha > 1$ (light tailed distribution).
- The distribution function of the Weibull distribution is given by

$$F_X(x) = 1 - \exp\left(-\left(\frac{x}{\theta}\right)^{\alpha}\right) \quad x > 0,\ \alpha > 0,\ \theta > 0.$$

Weibull Distribution (2)

- The k-th raw moment of the Weibull distributed random variable is given by

$$\mathrm{E}\left(X^k\right) = \theta^k \, \Gamma\left(1 + \frac{k}{\alpha}\right).$$

- The mean and variance are given by

$$\mathrm{E}\left(X\right) = \theta \, \Gamma\left(1 + \frac{1}{\alpha}\right)$$

and

$$\mathrm{Var}(X) = \theta^2 \left(\Gamma\left(1 + \frac{2}{\alpha}\right) - \left[\Gamma\left(1 + \frac{1}{\alpha}\right)\right]^2\right),$$

respectively.

The Generalized Beta Distribution of the Second Kind (1)

- A four-parameter, very flexible, distribution
- The continuous variable X is said to have the *GB2* distribution with parameters σ, θ, α_1 and α_2 if its *pdf* is given by

$$f_X(x) = \frac{(x/\theta)^{\alpha_2/\sigma}}{x\sigma\ \mathrm{B}(\alpha_1, \alpha_2)\left[1 + (x/\theta)^{1/\sigma}\right]^{\alpha_1+\alpha_2}} \quad \text{for } x > 0,$$

$\sigma, \theta, \alpha_1, \alpha_2 > 0$, and where the beta function $\mathrm{B}(\alpha_1, \alpha_2)$ is defined as

$$\mathrm{B}(\alpha_1, \alpha_2) = \int_0^1 t^{\alpha_1 - 1}(1-t)^{\alpha_2 - 1}\, dt.$$

The Generalized Beta Distribution of the Second Kind (2)

- The *GB2* provides a model for heavy as well as light tailed data.
- It includes the exponential, gamma, Weibull, Burr, Lomax, F, chi-square, Rayleigh, lognormal and log-logistic as special or limiting cases.
- Suppose that G_1 and G_2 are independent random variables where G_i has a gamma distribution with shape parameter α_i and scale parameter 1. Then, the random variable $X = \theta \left(\frac{G_1}{G_2}\right)^{\sigma}$ has a *GB2* distribution.
- When the moments exist, one can show that the k-th raw moment of the *GB2* distributed random variable is given by

$$\mathrm{E}\left(X^k\right) = \frac{\theta^k \ \mathrm{B}\left(\alpha_1 + k\sigma, \alpha_2 - k\sigma\right)}{\mathrm{B}\left(\alpha_1, \alpha_2\right)}, \quad k > 0.$$

Exercise 3.2.1

- Data: Wisconsin Property Fund. We have read the data and created a vector of the log of non-zero claim values. The name of that vector is `wisc_prop`.
- Use the `hist` function to plot a histogram.
- Use the `str` function to review the elements of the list returned by the histogram function.
- R functions: `hist` and `str`.

Exercise 3.2.2

- Data: same as Exercise 3.2.1
- Aim: develop an intuition as to whether a gamma distribution is likely to fit this data reasonably well.
- Plot the histogram and specify the probability argument so as to plot probability densities; calculate shape and scale parameters using the method of moments.
- R functions: `hist`, `dgamma`, and `curve`.

Exercise 3.2.3

- Aim: write your own function to find the density (the pdf) of a random variable that follows a Pareto distribution.
- R function: `function`.

Review

In this section, you learnt how to define and apply four fundamental severity distributions:

- gamma,
- Pareto,
- Weibull, and
- generalized beta distribution of the second kind.

Modeling Loss Severity

A short course authored by the Actuarial Community

11 May, 2021

Methods of Creating New Distributions I

Creating Severity Distributions

- In this section, we consider two main approaches to creating new distributions.
- Approach 1: We consider distributions that are created by transforming the random variable of a distribution:
  - Multiplication by a constant ($Y = cX$)
  - Raising to a power ($Y = X^{\tau}$)
  - Exponentiation ($Y = e^{X}$)
- Approach 2 (next video): We consider ways of combining distributions to form a distribution of interest:
  - Mixing
  - Splicing

Multiplication by a Constant

- Multiplying a random variable by a positive constant
  - Think of X as this year's losses and assume that we have an 8% inflation rate. We can model next year's losses as $Y = 1.08X$
  - Can readily go from dollars to thousands of dollars ($c = 1/1000$) or from dollars to Euros
- More generally, let $Y = cX$ and use

$$
\begin{aligned}
F_Y(y) &= \Pr(Y \leq y) = \Pr\left(X \leq \frac{y}{c}\right) = F_X\left(\frac{y}{c}\right) \\
f_Y(y) &= \frac{1}{c}\, f_X\left(\frac{y}{c}\right)
\end{aligned}
$$

Scale Distributions

- In a scale distribution, the transformed variable $Y = cX$ has a distribution from the same family as the random variable X
- Many loss distributions are scale distributions
- Typically, one uses θ as the scale parameter

If X comes from a distribution with parameter θ, then $Y = cX$ has the same distribution with scale parameter $\theta^* = c\theta$

- Gamma distribution is an example of a scale distribution

Raising to a Power

- Consider $Y = X^\tau$. We examine three cases:
- Case 1: $\tau > 0$ (**transformed**). Consider the transformation $Y = X^\tau$, then we have

$$\begin{aligned} F_Y(y) &= \Pr(Y \leq y) = \Pr\left(X^\tau \leq y\right) = \Pr\left(X \leq y^{1/\tau}\right) = F_X\left(y^{1/\tau}\right) \\ f_Y(y) &= \frac{1}{\tau} y^{(1/\tau)-1}\, f_X\left(y^{1/\tau}\right) \end{aligned}$$

- Case 2: $\tau < 0$ (**inverse transformed**). We have

$$\begin{aligned} F_Y(y) &= \Pr(Y \leq y) = \Pr\left(X^\tau \leq y\right) = \Pr\left(X \geq y^{1/\tau}\right) = 1 - F_X\left(y^{1/\tau}\right) \\ f_Y(y) &= \left|\frac{1}{\tau}\right| y^{(1/\tau)-1}\, f_X\left(y^{1/\tau}\right) \end{aligned}$$

- Case 3: $\tau = -1$ (**inverse**). This is a special case of when $\tau < 0$, Where $Y = 1/X$.

Exponential to get an Inverse Exponential

- Suppose that X has an exponential distribution with parameter θ^* and consider $Y = 1/X$
- Cdf of Y is

$$F_Y(y) = \Pr(Y \leq y) = \Pr(\frac{1}{X} \leq y) = \Pr(X \geq \frac{1}{y}) = \exp\left(-\frac{1}{y\theta^*}\right).$$

- Define a new parameter $\theta = \frac{1}{\theta^*}$. With this notation,

$$F_Y(y) = \Pr(Y \leq y) = \exp\left(-\frac{\theta}{y}\right).$$

- This is an inverse exponential distribution with parameter θ

Exponential to get a Weibull

- Start with $X \sim$ exponential distribution with parameter 1. Define transformed random variable with positive parameters τ and θ:

$$Y = \theta X^{1/\tau}$$

- This has distribution

$$\begin{aligned} F_Y(y) &= \Pr(Y \le y) \\ &= \Pr(X^{1/\tau} \le \frac{y}{\theta}) = \Pr(X \le \left(\frac{y}{\theta}\right)^{\tau}) \\ &= 1 - \exp\left(-\left(\frac{y}{\theta}\right)^{\tau}\right), \end{aligned}$$

known as a Weibull distribution

Exponentiation/ Natural Log

- Another type of transformation involves exponentiating a random variable so that $Y = e^X$
- Develop the distribution of the new random variable through the cdf

$$F_Y(y) = \Pr(Y \leq y) = \Pr(e^X \leq y) = \Pr(X \leq \ln y) = F_X(\ln y)$$

and the pdf

$$f_Y(y) = \frac{1}{y} f_X(\ln y).$$

- If $X \sim$ normal, then $Y = e^X \sim$ a lognormal distribution

Exponentiation/ Natural Log

- Another type of transformation involves exponentiating a random variable so that $Y = e^X$
- Develop the distribution of the new random variable through the cdf

$$F_Y(y) = \Pr(Y \leq y) = \Pr(e^X \leq y) = \Pr(X \leq \ln y) = F_X(\ln y)$$

and the pdf

$$f_Y(y) = \frac{1}{y} f_X(\ln y).$$

- If $X \sim$ normal, then $Y = e^X \sim$ a lognormal distribution
- Another commonly used transformation involves taking the natural logarithm a random variable so that $Y = \ln(X)$
- The pdf and cdf of the new variable Y is given by:

$$F_Y(y) = F_X(e^y) \quad \text{and} \quad f_Y(y) = e^y f_X(e^y) = x f_X(x).$$

Log Transformation: R code

```
#Assign values for alpha and theta for Gamma dist.
alpha<-0.8
theta<-10000

#Find the pdf for the claim severities (x)
x<-seq(5,300000)
fx<-dgamma(x,shape=alpha,scale=theta)

#Find the transformed pdf (Y) and plot
# Y = log(X)
# F_Y(y) = F_X(e^y)
# f_Y(y) = e^y f_X(e^y) = x*f_X(x)
x<-seq(5,300000)
y<-log(x)
fy<-fx*x
plot(y,fy,type="l",lty=1,xlab="log claim severity",
     ylab="density",ylim=c(0,0.35))
```

Review

In this section, you learned how to:

- Provide foundations for creating new distributions: multiplying by a constant, raising to a power, and exponentiation.
- Understand connections among the distributions

Modeling Loss Severity

A short course authored by the Actuarial Community

13 June, 2021

Methods of Creating New Distributions II

Creating Severity Distributions

- Approach 1 (previous video): We discussed distributions that are created by transforming the random variable of a distribution:
  - Multiplication by a constant ($Y = cX$)
  - Raising to a power ($Y = X^{\tau}$)
  - Exponentiation ($Y = e^{X}$)
- Approach 2: We consider ways of combining distributions to form a distribution of interest:
  - Mixing
  - Splicing

Discrete Mixture Severity Distributions

- *Definition.* Let $X_1, \ldots, X_k$ be random variables and define

$$Y = \begin{cases} X_1 & \textit{with probability } \alpha_1 \\ \vdots & \vdots \\ X_k & \textit{with probability } \alpha_k \end{cases}$$

Here, $\alpha_j > 0$ and $\alpha_1 + \cdots + \alpha_k = 1$. Then, Y is a k-point mixture random variable

Cdf is

$$F_Y(y) = \alpha_1 F_{X_1}(y) + \cdots + \alpha_k F_{X_k}(y)$$

with mean

$$E(Y) = \alpha_1 E(X_1) + \cdots + \alpha_k E(X_k).$$

Example: Actuarial Exam Question

Question

A collection of insurance policies consists of two types. 25% of policies are Type 1 and 75% of policies are Type 2. For a policy of Type 1, the loss amount per year follows an exponential distribution with mean 200, and for a policy of Type 2, the loss amount per year follows a Pareto distribution with parameters $\alpha = 3$ and $\theta = 200$. For a policy chosen at random from the entire collection of both types of policies, find the probability that the annual loss will be less than 100, and find the average loss.

Example: Actuarial Exam Question

Solution:
The two types of losses are the random variables X_1 and X_2. X_1 has an exponential distribution with mean 200, so

$$F_{X_1}(100) = 1 - e^{-\frac{x}{\theta}} = 1 - e^{-\frac{100}{200}} = 0.393$$

X_2 has a Pareto distribution with parameters $\alpha = 3$ and $\theta = 200$, so

$$F_{X_2}(100) = 1 - \left(\frac{\theta}{x+\theta}\right)^{\alpha} = 1 - \left(\frac{200}{100+200}\right)^{3} = 0.704$$

Hence,

$$F_X(100) = (0.25 \times 0.393) + (0.75 \times 0.704) = 0.626.$$

Further, the average loss is given by:

$$E(X) = 0.25E(X_1) + 0.75E(X_2) = (0.25 \times 200) + (0.75 \times 100) = 125.$$

R Code to Produce a Mixture Distribution

```
# Create a vector of alphas and thetas for the two
# subpopulations and plot the pdfs
alpha<-c(0.8, 0.5) ; theta<-c(15000, 30000)
plot(log(x),dgamma(x,shape=alpha[1],scale=theta[1])*x,
      type="l",xlab="log claim severity",ylab="density",
      ylim=c(0,0.35))
lines(log(x), dgamma(x,shape=alpha[i],scale=theta[i])*x)
# Create a vector of weights and create an object
# called pdfmix for the mixture pdf on log scale
weight<-c(0.25,0.75)
x<-seq(5,300000)
pdfmix<-0
for(i in 1:2){
  pdfmix<-dgamma(x,shape=alpha[i],
      rate=1/theta[i])*x*weight[i]+pdfmix}
#Superimpose a plot of the mixture pdf
lines(log(x), pdfmix,col="red",lwd=2)
```

Continuous Mixtures for Severity

- Infinite number of subgroups within a population

Each subgroup has $\mathrm{F}(\cdot|\theta)$ (e.g., exponential) but with a parameter θ that accounts for population differences

- Assume the random variable Θ has pdf $f_\Theta(\theta)$
- Cdf:

$$\begin{aligned}
\mathrm{F}_X(x) = \Pr(X \leq x) &= E_\Theta[\Pr(X \leq x|\Theta)] \\
&= \int \Pr(X \leq x|\theta) f_\Theta(\theta) d\theta \\
&= \int \mathrm{F}_{X|\Theta}(x|\theta) f_\Theta(\theta) d\theta
\end{aligned}$$

- Pdf:

$$f_X(x) = \int f_{X|\Theta}(x|\theta) f_\Theta(\theta) d\theta$$

Special Case: Gamma Mixtures of Exponentials

- Suppose $X|\Theta \sim \text{exponential}(\frac{1}{\Theta})$:

$$f_{X|\Theta}(x|\theta) = \theta e^{-\theta x}$$

- Suppose $\Theta \sim \text{gamma}(\alpha, \beta)$

$$f_{\Theta}(\theta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}}\theta^{\alpha-1}e^{-\theta/\beta}$$

- Pdf of X is

$$\begin{aligned} f_X(x) &= \int f_{X|\Theta}(x|\theta) f_{\Theta}(\theta) d\theta \\ &= \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \int_0^{\infty} \theta^{\alpha} e^{-\theta(x+1/\beta)} d\theta = \frac{\alpha\beta}{(1+x\beta)^{\alpha+1}} \end{aligned}$$

- This is a Pareto distribution with parameters α and $\theta = 1/\beta$

Mixture Expectations

- Law of iterated expectation:

$$E(X) = E_{\Theta}[\mathrm{E}(X|\Theta)]$$

- This is easily extended to kth moment:

$$E(X^k) = E_{\Theta}[E(X^k|\Theta)]$$

Law of total variance:

$$Var(X) = E_{\Theta}[Var(X|\Theta)] + Var_{\Theta}[E(X|\Theta)]$$

Splicing

- Join (splice) together different probability density functions to form a pdf over support of a random variable

$$f_X(x) = \begin{cases} \alpha_1 f_1(x) & c_0 < x < c_1 \\ \alpha_2 f_2(x) & c_1 < x < c_2 \\ \vdots & \vdots \\ \alpha_k f_k(x) & c_{k-1} < x < c_k \end{cases}$$

$\alpha_1 + \alpha_2 \cdots + \alpha_k = 1$

Each f_j is a pdf, so that $\int_{c_{j-1}}^{c_j} f_j(x)dx = 1$

c_j's are typically known

Review

In this section, you learned how to:

- Provide foundations for creating new distributions: Mixing and Splicing
- Understand connections among the distributions.

Modeling Loss Severity

A short course authored by the Actuarial Community

13 June, 2021

Coverage Modifications: Deductibles

Risk Retention Framework

- Now consider the following framework:
  - Policyholder or insured suffers a loss of amount X
  - Under an insurance contract, the insurer is obligated to cover a portion of X, denoted as Y
  - Y represents the insurer's claim payment
- In this section:
  - We introduce standard mechanisms that insurers use to reduce, or mitigate, their risk, including deductibles, policy limits and coinsurance.
  - We will also introduce reinsurance, a mechanism of insurance for insurance companies.
  - We examine how the distribution of the insurer's obligations depends on these mechanisms.
- In this video, we focus on impacts of deductibles.

Ordinary Deductible

- Under **Ordinary Deductible (d)**, we first define the payment per loss variable Y^L, given by:

  $$Y^L = (X-d)_+ = \begin{cases} 0 & X \le d \\ X-d & X > d. \end{cases}$$

  Notation "$(\cdot)_+$" means "Take the positive part of".

- Insurance only pays amounts in excess of the deductible d. If a loss is less than the deductible, the insurer does not observe the loss.
  - Random variable Y^P is the claim that an insurer observes
  - "P" superscript indicates that the retained loss is on a per payment basis

- Therefore, the payment per payment variable Y^P, given by:

  $$Y^P = \begin{cases} \text{undefined/not observed} & X \le d \\ X-d & X > d \end{cases}$$

Distribution of Payment Per Loss Variable Y^L

- The distribution of Y^L is a hybrid combination of discrete and continuous components:
  - The discrete part of the distribution is concentrated at $Y = 0$ when $(X \leq d)$
  - the continuous part is spread over the interval $Y > 0$ when $(X > d)$
- Using the transformation $Y^L = X - d$ for the continuous part of the distribution, we can find the *pdf* of Y^L given by:

$$f_{Y^L}(y) = \begin{cases} F_X(d) & y = 0 \\ f_X(y+d) & y > 0. \end{cases}$$

- The distribution functions of Y^L can be found directly using the *pdf* of X as follows

$$F_{Y^L}(y) = \begin{cases} F_X(d) & y = 0 \\ F_X(y+d) & y > 0. \end{cases}$$

Distribution of Payment Per Payment Variable Y^P

- We can see that the payment per payment variable is the payment per loss variable $Y^P = Y^L$ conditional on the loss exceeding the deductible $(X > d)$, i.e.,$Y^P = (X - d)|X > d$.
- Hence, the *pdf* of Y^P is given by:

$$f_{Y^P}(y) = \frac{f_X(y+d)}{1 - F_X(d)}, \qquad \text{for } y > 0.$$

- The distribution functions of Y^P can be found directly using the *pdf* of X as follows

$$F_{Y^P}(y) = \frac{F_X(y+d) - F_X(d)}{1 - F_X(d)}, \qquad \text{for } y > 0$$

Raw Moments of Y^L and Y^P

- The raw moments of Y^L and Y^P can be found directly using the *pdf* of X as follows:

$$E\left[\left(Y^L\right)^k\right] = \int_d^\infty (x-d)^k f_X(x)dx, \quad \text{and}$$

$$E\left[\left(Y^P\right)^k\right] = \frac{\int_d^\infty (x-d)^k f_X(x)dx}{1-F_X(d)} = \frac{E\left[\left(Y^L\right)^k\right]}{1-F_X(d)}.$$

- For $k=1$, we can use the survival function to calculate $E\left(Y^L\right)$ as:

$$E\left(Y^L\right) = \int_d^\infty [1-F_X(x)]dx$$

- For $k=1$, the expectation of Y^P is known as the mean excess loss $e_X(d)$. Thus:

$$E\left(Y^P\right) = e_X(d) = E(X-d|X>d) = \frac{E\left(Y^L\right)}{1-F_X(d)} = \frac{\int_d^\infty [1-F_X(x)]dx}{1-F_X(d)}.$$

Example: Actuarial Exam Question

Question: For an insurance, losses have the density function

$$f_X(x) = \begin{cases} 0.02x & 0 < x < 10 \\ 0 & \text{elsewhere} \end{cases}$$

The insurance has an ordinary deductible of 4 per loss. Y^P is the claim payment per payment random variable. Calculate $E[Y^P]$

Solution: We define Y^P as follows

$$Y^P = \begin{cases} \text{undefined} & X \leq 4 \\ X - 4 & X > 4 \end{cases}$$

So, $E(Y^P) = \frac{\int_4^{10}(x-4)0.02x dx}{1-F_X(4)} = \frac{2.88}{0.84} = 3.43$

Loss Elimination Ratio (LER)

- Consider an ordinary deductible, cost (amount of payment) per loss event.
- The Loss elimination ratio at deductible d is the fraction of the losses have been eliminated by introducing the deductible and given by:

$$LER = \frac{E(X) - E(Y^L)}{E(X)}$$

Franchise Deductible

- A little less common type of policy deductible is the franchise deductible.
- The payment per loss and payment per payment variables are defined as
$$Y^L = \begin{cases} 0 & X \leq d \\ X & X > d. \end{cases}$$
and
$$Y^P = \begin{cases} \text{Undefined} & X \leq d \\ X & X > d. \end{cases}$$

Review

Random Variable	**Expectation**
Excess loss random variable	$e_X(d) = \mathrm{E}(X - d \mid X > d)$
$Y^P = X - d$ if $X > d$	mean excess loss function
left truncated and shifted variable	mean residual life function
	complete expectation of life
	$e_X^k(d) = E\left[(X-d)^k \mid X > d\right]$
$Y^L = (X-d)_+ = \begin{cases} 0 & X \leq d \\ X - d & X > d \end{cases}$	$E\,(X-d)_+ = e(d)S(d)$
left-censored and shifted variable	$E\,(X-d)_+^k = e^k(d)S(d)$

For nonnegative, continuous random variables,

$$E\left(Y^L\right) = E(X-d)_+ = \int_d^\infty S\,(x)\,dx$$

Modeling Loss Severity

A short course authored by the Actuarial Community

10 May, 2021

Coverage Modifications: Limits

Policy Limit

- Under policy limits, the insurer is responsible for covering the actual loss X up to the limit of its coverage u.
- If the loss exceeds the policy limit, the difference $X - u$ has to be paid by the policyholder.
- The amount paid by the insurer Y is known as the *limited loss variable*, denoted by $X \wedge u$, and expressed as:

$$Y = X \wedge u = \begin{cases} X & X \leq u \\ u & X > u. \end{cases}$$

Notation "$\wedge$" means "take the minimum of"

- The distinction between Y^L and Y^P is not needed under limited policy as the insurer will always make a payment.

Distribution of the Limited Loss Variable

- The distribution of Y is a hybrid combination of discrete and continuous components:
  - The discrete part of the distribution is concentrated at $Y = u$ when $(X > u)$
  - the continuous part is spread over the interval $Y < u$ when $(X \leq u)$
- The *pdf* of Y is given by:

$$f_Y(y) = \begin{cases} f_X(y) & 0 < y < u \\ 1 - F_X(u) & y = u. \end{cases}$$

- Accordingly, the distribution function of Y is given by:

$$F_Y(y) = \begin{cases} F_X(x) & 0 < y < u \\ 1 & y \geq u. \end{cases}$$

Raw Moments of Limited Loss Variable

- The raw moments of Y can be found directly using the *pdf* of X as follows:

$$E\left(Y^k\right) = E\left[(X \wedge u)^k\right] \quad \begin{aligned} &= \textstyle\int_0^u x^k f_X(x)dx + \int_u^\infty u^k f_X(x)dx \\ &= \textstyle\int_0^u x^k f_X(x)dx + u^k\left[1 - F_X(u)\right]. \end{aligned}$$

- An alternative expression using the survival function is:

$$E\left(Y^k\right) = E\left[(X \wedge u)^k\right] = \int_0^u kx^{k-1}\left[1 - F_X(x)\right]dx.$$

- For $k = 1$, this is

$$E\left(Y\right) = E(X \wedge u) = \int_0^u \left[1 - F_X(x)\right]dx.$$

Limited Expected Value: Pareto Policy Limit

- Expected value of limited loss variable $(X \wedge u)$ is

$$E(Y) = E(X \wedge u) = \int_0^u (1 - F(x))\, dx = \int_0^u S(x) dx.$$

Pareto Policy Limit. Recall

$$1 - F(x) = S(x) = \Pr(X > x) = \left(\frac{\theta}{x + \theta} \right)^{\alpha}$$

with mean $\mathrm{E}\,(X) = \frac{\theta}{\alpha - 1}$. Thus, the limited expected value is

$$\begin{aligned} E(X \wedge u) &= \theta^{\alpha} \int_0^u (x + \theta)^{-\alpha} dx = \theta^{\alpha}\, \frac{(x + \theta)^{-\alpha+1}}{-\alpha + 1} \Bigg|_0^u \\ &= \theta^{\alpha} \left(\frac{\theta^{-\alpha+1} - (u + \theta)^{-\alpha+1}}{\alpha - 1} \right) \\ &= \frac{\theta}{\alpha - 1} \left\{ 1 - \left(\frac{\theta}{u + \theta} \right)^{\alpha - 1} \right\}. \end{aligned}$$

Expected Payments Under Deductible Revisited

- Using the definitions of $(X-d)_+$ and $X \wedge d$, we see that:

$$X = (X-d)_+ + X \wedge d \qquad \text{then}$$

$$E(X) = E(X-d)_+ + E(X \wedge d)$$

Pareto distribution Example. Recall

$$E(X) = \frac{\theta}{\alpha-1} \quad \text{and} \quad E(X \wedge d) = \frac{\theta}{\alpha-1}\left\{1-\left(\frac{\theta}{d+\theta}\right)^{\alpha-1}\right\}.$$

Thus,

$$\begin{aligned} E(X-d)_+ &= E(X) - E(X \wedge d) \\ &= \frac{\theta}{\alpha-1} - \frac{\theta}{\alpha-1}\left\{1-\left(\frac{\theta}{d+\theta}\right)^{\alpha-1}\right\} \\ &= \frac{\theta}{\alpha-1}\left\{\left(\frac{\theta}{d+\theta}\right)^{\alpha-1}\right\} \end{aligned}$$

Loss Elimination Ratio (LER) Revisited

- Recall:

$$X = (X - d)_+ + X \wedge d \qquad \text{and}$$

,

$$LER = \frac{E(X) - E(X - d)_+}{E(X)}$$

- Then the Loss elimination ratio at deductible d can be expressed as

$$\begin{aligned} LER &= \frac{\mathrm{E}\,(X \wedge d)}{E(X)} \\ &= \frac{\text{limited exp value}}{\text{exp value}} \end{aligned}$$

Example: Actuarial Exam Question

Question: Under a group insurance policy, an insurer agrees to pay 100% of the medical bills incurred during the year by employees of a small company, up to a maximum total of one million dollars. The total amount of bills incurred, X, has pdf

$$f_X(x) = \begin{cases} \frac{x(4-x)}{9} & 0<x<3 \\ 0 & \text{elsewhere} \end{cases}$$

where x is in millions. Calculate the total amount, in millions of dollars, the insurer would expect to pay under this policy.

Solution: Define the total amount paid by the insurer as

$$Y = X \wedge 1 = \begin{cases} X & X \leq 1 \\ 1 & X > 1. \end{cases}$$

So, $E(Y) = E(X \wedge 1) = \int_0^1 \frac{x^2(4-x)}{9} dx + 1 \cdot \int_1^3 \frac{x(4-x)}{9} dx = 0.935.$

R Code to Calculate Limited Expected Value

```r
set.seed(12345)
mu <- 12; sigma <- 1.5
claims <- rlnorm(n = 100000, meanlog = mu, sdlog = sigma)
## Functions to calculate unlimited and limited expected
## values
mlnorm <- function(meanlog, sdlog) exp(meanlog+sdlog^2/2)
levlnorm <- function(limit, meanlog, sdlog){
  t1 <- mlnorm(meanlog = meanlog, sdlog = sdlog) *
    pnorm(q = (log(limit)-meanlog-sdlog^2)/sdlog)
  t2 <- limit*(1- pnorm((log(limit)- meanlog)/sdlog))
  t1 + t2
}
## Unlimited
mean(claims)
mlnorm(meanlog = mu, sdlog = sigma)
## Claim amounts retained under a $1 million deductible
mean(pmin(1e6, claims))
levlnorm(limit = 1e6, meanlog = mu, sdlog = sigma)
```

Review

Random Variable	Expectation
limited loss variable $\min(X, u) = X \wedge u = \begin{cases} X & X \leq u \\ u & X > u \end{cases}$ right censored	$E(X \wedge u)$ – limited expected value

Note that $(X - u)_+ + (X \wedge u) = X$. Thus, $E(X - u)_+ + E(X \wedge u) = E(X)$

For nonnegative, continuous random variables,

$$E(X \wedge u) = \int_0^u S(x)\, dx$$

Chapter 3 Modeling Loss Severity

A short course authored by the Actuarial Community

3.5 Maximum Likelihood Estimation

Overview

In this section, you learn how to:

- Define a likelihood for a sample of observations from a continuous distribution
- Define the maximum likelihood estimator for a random sample of observations from a continuous distribution
- Estimate parametric distributions based on grouped, censored, and truncated data

Maximum Likelihood Estimators for Complete Data

- A random sample $X_1, \ldots, X_n$ from a distribution with distribution function F_X. x_i is the observation.
- $\boldsymbol{\theta}$ denotes the set of parameters for F_X.
- With the independence assumption of complete (individual) and continuous data, we define the **likelihood** to be

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} f(x_i).$$

- Each individual observation is recorded, and its contribution to the likelihood function is the density at that value.
- The maximum likelihood estimator is that value of the parameters in $\boldsymbol{\theta}$ that maximize $L(\boldsymbol{\theta})$.
- The **logarithmic likelihood** is denoted as

$$l(\boldsymbol{\theta}) = \log L(\boldsymbol{\theta}) = \sum_{i=1}^{n} \log f(x_i).$$

Maximum Likelihood Estimators using Modified Data

- Maximum Likelihood Estimators for Grouped Data
- Maximum Likelihood Estimators for Censored Data
- Maximum Likelihood Estimators for Truncated Data

Maximum Likelihood Estimators for Grouped Data

- The observations are only available in grouped form, and the contribution of each observation to the likelihood function is the probability of falling in a specific group (interval).
- Let n_j represent the number of observations in the interval $(c_{j-1}, c_j]$. The grouped data likelihood function is thus given by

$$L(\theta) = \prod_{j=1}^{k} \left[F_X(c_j|\theta) - F_X(c_{j-1}|\theta)\right]^{n_j},$$

where c_0 is the smallest possible observation (often set to zero) and c_k is the largest possible observation (often set to infinity).

Maximum Likelihood Estimators for Censored Data

- Only partial information is available; all that may be known is that the observation exceeds a specific value.
- The contribution of the censored observation to the likelihood function is the probability of the random variable exceeding this specific limit.
- The likelihood function for censored data is then given by

$$L(\theta) = \left[\prod_{i=1}^{r} f_X(x_i) \right] [S_X(u)]^m ,$$

where r is the number of known loss amounts below the limit u and m is the number of loss amounts larger than the limit u.

Maximum Likelihood Estimators for Truncated Data

- If the values of X are truncated at d, then it should be noted that we would not have been aware of the existence of these values had they not exceeded d.
- The contribution to the likelihood function of an observation x truncated at d will be a conditional probability and the $f_X(x)$ will be replaced by $\frac{f_X(x)}{S_X(d)}$.
- The likelihood function for truncated data is then given by

$$L(\theta) = \prod_{i=1}^{k} \frac{f_X(x_i)}{S_X(d)},$$

where k is the number of loss amounts larger than the deductible d.

Exercise 3.5.1

- Data: Extract claim severity data from the Wisconsin Property Fund data.
- Aim: (1) Write a function to calculate the negative log-likelihood for the claim severity data, assuming that claim severity is Pareto distributed. (2) Find the maximum likelihood estimates and plot a contour plot of the likelihood.
- R functions: `dpareto`, `optim`, `contour`, and `points`.

Exercise 3.5.2

- Data: same as Exercise 3.5.2
- Aim: calculate 95% confidence intervals for the parameter estimates that we found in exercise 3.5.1.
- Instructions: use the Hessian matrix (which is a matrix of the second-order partial derivatives) to find the variance of the estimators. Specifically, the standard errors of the estimators are equal to the square root of the diagonal elements of the inverse of the Hessian matrix.
- R functions: `solve`, `diag`, and `qnorm`.

Review

In this section, you learnt how to:

- Define a likelihood for a sample of observations from a continuous distribution
- Define the maximum likelihood estimator for a random sample of observations from a continuous distribution
- Estimate parametric distributions based on grouped, censored, and truncated data

Modeling Loss Severity

A short course authored by the Actuarial Community

12 June, 2021

Coverage Modifications: Coinsurance and Reinsurance

Coinsurance

- Under the coinsurance, the retained loss by the insurer is a percentage α.
- The percentage α is often referred to as the coinsurance factor.
- Define $Y = \alpha X$. Typically, $0 < \alpha < 1$ and X denote the loss incurred.
- Combining three special cases of coverage modifications (deductible, limit, coinsurance) results in the payment per loss variable Y^L defined as:

$$Y^L = \begin{cases} 0 & X \leq d, \\ \alpha(X - d) & d < X \leq u, \\ \alpha(u - d) & X > u. \end{cases}$$

- Think about these as parameters in a contract between a policyholder and an insurer and so represent modifications of the underlying contract.

Raw Moments of Y^L

- Using $Y^L = \alpha\left[(X \wedge u) - (X \wedge d)\right]$

$$
\begin{aligned}
E\left(Y^L\right) &= \alpha\left[E(X \wedge u) - E(X \wedge d)\right] \\
&= \alpha \int_0^u \left[1 - F_X(x)\right] dx - \alpha \int_0^d \left[1 - F_X(x)\right] dx.
\end{aligned}
$$

- The k-th raw moment of Y^L is given by:

$$
E\left[\left(Y^L\right)^k\right] = \int_d^u \left[\alpha(x-d)\right]^k f_X(x) dx + \left[\alpha(u-d)\right]^k \left[1 - F_X(u)\right]
$$

Inflation

- To handle inflation, a *growth factor* $(1+r)$ may be applied to X resulting in an inflated random variable $(1+r)X$.
- The resulting per loss variable can be written as

$$Y^L = \begin{cases} 0 & X \le \frac{d}{1+r}, \\ \alpha[(1+r)X - d] & \frac{d}{1+r} < X \le \frac{u}{1+r}, \\ \alpha(u-d) & X > \frac{u}{1+r}. \end{cases}$$

- The first and second moments of Y^L can be expressed as:

$$E(Y^L) = \alpha(1+r)\left[E\left(X \wedge \tfrac{u}{1+r}\right) - E\left(X \wedge \tfrac{d}{1+r}\right)\right]$$

and

$$\begin{aligned} E\left[(Y^L)^2\right] &= \alpha^2(1+r)^2 \left\{ E\left[\left(X \wedge \frac{u}{1+r}\right)^2\right] - E\left[\left(X \wedge \frac{d}{1+r}\right)^2\right] \right. \\ &\quad \left. -2\left(\frac{d}{1+r}\right)\left[E\left(X \wedge \frac{u}{1+r}\right) - E\left(X \wedge \frac{d}{1+r}\right)\right]\right\} \end{aligned}$$

Example: Actuarial Exam Question

Question: The ground up loss random variable for a health insurance policy in 2006 is modeled with X, a random variable with an exponential distribution having mean 1000. An insurance policy pays the loss above an ordinary deductible of 100, with a maximum annual payment of 500. The ground up loss random variable is expected to be 5% larger in 2007, but the insurance in 2007 has the same deductible and maximum payment as in 2006. Find the percentage increase in the expected cost per payment from 2006 to 2007.

Solution: We define the amount per loss Y^L in both years as

$$Y_{2006}^L = \begin{cases} 0 & X \leq 100, \\ X - 100 & 100 < X \leq 600, \\ 500 & X > 600. \end{cases}$$

Example: Actuarial Exam Question

Solution Cont'd:

$$Y^L_{2007} = \begin{cases} 0 & X \leq 95.24, \\ 1.05X - 100 & 95.24 < X \leq 571.43, \\ 500 & X > 571.43,. \end{cases}$$

$$E(Y^P_{2006}) = \frac{E(Y^L_{2006})}{S_X(d)} \quad = \frac{E(X \wedge 600) - E(X \wedge 100)}{S_X(100)} = \frac{1000\left(1 - e^{-\frac{600}{1000}}\right) - 1000\left(1 - e^{-\frac{100}{1000}}\right)}{e^{-\frac{100}{1000}}} = 393.469$$

$$E(Y^P_{2007}) = \frac{E(Y^L_{2007})}{S_X(d)} \quad = \frac{1.05[E(X \wedge 571.43) - E(X \wedge 95.24)]}{S_X(95.24)} = \frac{1.05\left[1000\left(1 - e^{-\frac{571.43}{1000}}\right) - 1000\left(1 - e^{-\frac{95.24}{1000}}\right)\right]}{e^{-\frac{95.24}{1000}}} = 397.797$$

Then, $\dfrac{E(Y^P_{2007})}{E(Y^P_{2006})} - 1 = \dfrac{397.797}{393.469} - 1 = 0.011 = 1.1\%$

R Code to Calculate Expected Value of the Claims Under Deductible and Policy Limit

```
set.seed(12345)
mu <- 12; sigma <- 1.5
claims <- rlnorm(n = 100000, meanlog = mu, sdlog = sigma)

# Number of claims resulting in reimbursement
sum(claims > 1e6) # in the data
(1 - plnorm(q = 1e6, meanlog = mu, sdlog = sigma))
* length(claims) # lorgnormal model

# Average Value from claims in the data
insured <- claims[claims > 1e6]; mean(pmin(5e6, insured))

# Expected value of the claims based on lognormal model
1e6 + (levlnorm(limit = 5e6, meanlog = mu, sdlog = sigma) -
  levlnorm(limit = 1e6, meanlog = mu, sdlog = sigma)) /
  (1 - plnorm(q = 1e6, meanlog = mu, sdlog = sigma))
```

Reinsurance

- Reinsurance is a contractual arrangement under which an insurer transfers part of the underlying insured risk by securing coverage from another insurer (referred to as a reinsurer) in return for a reinsurance premium.
- The common form of reinsurance contracts is excess of loss coverage:
  - the primary insurer must make all required payments to the insured until the primary insurer's total payments reach a fixed reinsurance deducible.
  - the reinsurer is then only responsible for paying losses above the reinsurance deductible
- Benefits of reinsurance:
  - risk management tool for insurers
  - allows the primary insurer to benefit from underwriting skills, expertise and proficient complex claim file handling of the larger reinsurance companies.

Review

In this video, you learned how to:

- We evaluate the impacts of coinsurance and inflation on insurer's costs.
- We examine how the distribution of the insurer's obligations depends on these mechanisms.
- We will also introduce reinsurance, a mechanism of insurance for insurance companies.