Tail event probability for different distribution functions: A comparison

Bachelor's Thesis submitted

to

Prof. Dr. Wolfgang K. Härdle

and

Prof. Dr. Christoph Breunig

Humboldt-Universität zu Berlin
School of Business and Economics
Ladislaus von Bortkiewicz Chair of Statistics

by

Caspar Moser
(565110)

Bachelor of Science in Economics

Paris, November , 2017

Abstract

Aim of this paper is to discuss and compare tail risk for a small variety of continuous distributions. Necessary mathematical foundations from measure theory and probability theory as well as the mathematic definition of tail events are introduced. With these at hand the probability of the tail events, the tail risk is calculated for each distribution function. We find that the tail risk of uniform distribution is zero, that the tail risk of a normal distributions is well below one percent and that the exponential distribution and the Laplace distribution both have tail risk below two percent. We evaluate our findings and contextualize them.

Contents

1 Introduction

The term tail risk refers to extreme outcomes of uncertain experiments, which may or may not be unlikely. Generally, when confronted with an event with uncertain outcomes, people tend to expect the realization to be in a certain range according to their past experience with similar events. For example, when going to a football match, you wouldn't expect to see ten goals. It is possible and it has happened before. But it is so unlikely, that people tend to ignore the possibility.

Similarly when investing in a well thought out portfolio with a wide range of assets you would not expect it to lose all of its value. The probability of such events are thought of being near zero. But they might not be. When estimating the risks of a financial portfolio, it is often assumed that the prices follow a normal distribution. Classic portfolio theorys, like Markowitz (1991) and Black and Scholes (1973) rely on this assumption. The tail risk of a normal distribution however is very small. Assuming market outcomes to be normally distributed therefore might understate the actual tail risk, because extreme events occur more often than traditionally expected.

Recent history has brought up a large number of economics crisis that, ergo, lead traders to rethink the assumption of normality. So-called fat-tailed distributions might be an adequate replacement according to Mandelbrot (1963). The difference is basically, that classic distributions like the normal or exponential distribution only have finite variances, which limits the probability of tail events. Fat-tailed distributions like the Cauchy distributions have diverging moments.

This works focuses on distributions with finite moments. Its aim is to introduce the mathematical foundations of tail risk as well as to study and compare tail risk for four different "classic" distributions, i.e. the continuous uniform distribution, the exponential distribution, the Laplace distribution and the normal distribution.

The term tail event may not be confused with the term associated with Kolmogorov's zero-one law, as in Stroock (2010).

2 Mathematical foundations

2.1 Basic terms of measure theory

In order to be able to define Tail events mathematically, we first need to introduce some basic terms of measure theory. Any uncertain situation that we try to modelize statistically has a set of different possible outcomes. Let Ω be such a set. Furthermore, let $\mathcal{F}$ be a collection of subsets of Ω. (This is important. In many cases we are interested not in the probability of one specific outcome, but in the probability, that the actual outcome will be part of a group of possible outcomes. We will come back to this later.) $\mathcal{F}$ is called a σ-Algebra, if the following properties hold:

a. $\Omega \in \mathcal{F}$,

b. $\forall A \in \mathcal{F} : A^c \in \mathcal{F}$,

c. $\forall A_n \in \mathcal{F}, \forall n \geq 1 : \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$

The introduction of the σ-Algebra will allow us to determine for any of its elements the probability, that the actual outcome will be within that subset of Ω. We do so using probability measures. The function $\mathrm{P} : \mathcal{F} \to [0,1]$ is called a probability measure on Ω, if the following properties hold:

a. $\mathrm{P}(\Omega) = 1$,

b. $\forall A \in \mathcal{F}, \mathrm{P}(A) \geq 0$,

c. For any disjoint sequence of subsets $(A_n)_{n\geq 1}$ of $\mathcal{F}$:

$$\mathrm{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathrm{P}(A_n)$$

The triplet $(\Omega, \mathcal{F}, \mathrm{P})$ is called probability space. So far it does not allow us to actually calcutlate the probability of any element $A \in \mathcal{F}$ because we have not modelized the outcomes yet. In order to so, we need to quantify them. *For the upcoming sections, let* $(\Omega, \mathcal{F}, \mathrm{P})$ *be a probability space.*

2.2 Basic terms of probability theory

In order to quantify the outcomes of experiments, we introduce random variables. Random variables are functions, which assign to any single outcome of an experiment a real or complex number. Here we restrict ourselves on real random variables.

$X : \Omega \to \mathbb{R}$ is called real random variable, if X is measurable, i.e.

$$\forall B \in \mathcal{B}(\mathbb{R}) : X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{F}$$

where $\mathcal{B}(\mathbb{R})$ is the σ-algebra generated by the borel sets of $\mathbb{R}$ (any open subset of $\mathbb{R}$). Hence, random variables allow us to switch from spaces in the form of $(\Omega, \mathcal{F})$ to spaces in the form of $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. We now define the probability measure that is associated to a random variable, its probability distribution.

Let $X : \Omega \to \mathbb{R}$ be a random variable. The probability distribution of X is defined as

$$\mathrm{P}_X \stackrel{\text{def}}{=} \mathrm{P} \circ X^{-1}.$$

It can easily be proven that the function

$$\begin{aligned} \mathrm{P}_X : \mathcal{B}(\mathbb{R}) &\longrightarrow [0,1] \\ B &\longmapsto \mathrm{P}\{X^{-1}(B)\} \end{aligned}$$

is a probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. From now on we will write the expression $\mathrm{P}\{X^{-1}(B)\}$ as $\mathrm{P}(X \in B)$. Furthermore, if P_X is the probability measure of X, we will say X follows P_X and write $X \sim \mathrm{P}_X$.

So far we have defined certain terms that allow us to describe our random experiment mathematically. But we are not yet able to calculate the probabilities of its outcomes. For simple experiments, such as the toss of a coin or a dice, we know well there probability distributions (assuming we have a fair coin or dice). But normally, we have to calculate the probabilities. Therefore, we characterize the distribution by a function, the density function.

Let ν be a measure on $\mathbb{R}$. $f : \mathbb{R} \to \mathbb{R}$ is called density function of X with respect to ν if

a.
$$\int_{\mathbb{R}} f(x) d\nu(x) = 1$$

b.
$$\forall A \in \mathcal{B}(\mathbb{R}) \ \mathrm{P}_X(A) = \int_A f(x) d\nu(x)$$

We can identify unambiguously a probability distribution by its density function.

Now we have all the neccessary tools that we need to calculate the probabilities of an experiment. Before we can introduce the notion of Tail events there is one statistic we have to define, the expectation.

Let $\varphi : \mathcal{B}(\mathbb{R}) \to \mathcal{B}(\mathbb{R})$ be measurable and X a random variable. The expectation of $\varphi(X)$ is defined as
$$\mathsf{E}[\varphi(X)] \stackrel{\text{def}}{=} \int_{\Omega} \varphi(X)(\omega) d\mathrm{P}(\omega).$$
The definition alone does not yet enable us to calculate it for any random variable. Again we have to find the link from $(\Omega, \mathcal{F}, \mathrm{P})$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathrm{P}_X)$.

Suppose that X has a density function f with respect to ν. Using theorem 7.1 from **Gut (2012)**, $\mathsf{E}[X]$ can be rewritten as
$$\int_{\Omega} X(\omega) d\mathrm{P}(\omega) = \int_{\mathbb{R}} x f(x) d\nu(x).$$
We call $\mathsf{E}[X]$ the mean of X. The mean can be losely interpreted as the value that splits the distribution of its random variable in half. More precisely it measures the center of gravity of a distribution.

The variance of X is defined as
$$\mathsf{Var}(X) \stackrel{\text{def}}{=} \mathsf{E}[(X - \mathsf{E}[X])^2]$$
It can be rewritten as
$$\mathsf{E}[(X - \mathsf{E}[X])^2] = \mathsf{E}[X^2] - \mathsf{E}[X]^2.$$

Proof
Using the linearity of the integral,
$$\begin{aligned} \mathsf{E}[(X - \mathsf{E}[X])^2] &= \mathsf{E}[X^2 - 2X\,\mathsf{E}[X] + \mathsf{E}[X]^2] \\ &= \mathsf{E}[X^2] - 2\,\mathsf{E}[X]^2 + \mathsf{E}[X]^2 \\ &= \mathsf{E}[X^2] - \mathsf{E}[X]^2. \end{aligned}$$

The variance measures the dispersion of a random variable, that is to say it measures how far on average a random varibale moves from its expected value, the mean. For the definition of tail events we will need its square root, called standard deviation. We note
$$\mathsf{sd} \stackrel{\text{def}}{=} \sqrt{\mathsf{Var}(X)}.$$

3 Tail event probability

We call tail risk the risk of a random variable being realised more than three times the standard deviation away from the mean. These relatively rare events are called tail events. Formally, we can calculate the probabiliy of tail events for different distributions as the probability of an associated random variable being in the tail region.

Let X be random variable on $(\Omega, \mathcal{F}, \mathrm{P})$. We call tail region of X the set

$$TR_X \stackrel{\text{def}}{=}]-\infty,\ \mathsf{E}[X] - 3 \cdot \mathsf{sd}(X)\,[\ \cup\ ]\ \mathsf{E}[X] + 3 \cdot \mathsf{sd}(X),\ +\infty\,[\ .$$

We call $T1_X \stackrel{\text{def}}{=} \mathsf{E}[X] - 3\,\mathsf{sd}(X)$ lower boundary and $T2_X \stackrel{\text{def}}{=} \mathsf{E}[X] + 3\,\mathsf{sd}(X)$ upper boundary of the tail region. We call tail event probability or tail risk the probability

$$\boxed{p_X^{TR} \stackrel{\text{def}}{=} \mathrm{P}(X \in TR_X).}$$

Now we have the neccesary tools to calculate tail risk for any probability distribution. In a next step, we will do so for some common distributions and discuss our findings. The procedure will be more or less the same for all dristibutions we will approach. Using the density function, we will calculate the mean and the variance. With these at hand, we will procede to calculate the probability mass of the tail region.

3.1 Continuous distributions

We call X continuous if it has a density with respect to the Lebesgue measure λ. In that case we call P_X continuous distribution.

3.1.1 Uniform distribution (continuous)

Let X be a random variable. X follows a continuous uniform distribution with parameters $a, b \in \mathbb{R},\ a < b$ (i.e. $\mathrm{P}_X = \mathcal{U}[a, b]$), if its density function $f : \mathbb{R} \to \mathbb{R}$ is given by

$$f(t) = \frac{1}{b-a}\mathbf{1}_{[a,b]}(t).$$

Let us calculate the mean of X.

$$
\begin{aligned}
\mathsf{E}[X] &= \int_{\mathbb{R}} t f(t) d\lambda(t) \\
&= \int_{\mathbb{R}} \frac{t}{b-a} \mathbf{1}_{[a,b]}(t) d\lambda(t) \\
&= \frac{1}{b-a} \int_a^b t d\lambda(t) \\
&= \frac{1}{b-a} \left(\frac{b^2}{2} - \frac{a^2}{2} \right) \\
&= \frac{1}{b-a} \frac{(b-a)(b+a)}{2} \\
&= \frac{a+b}{2}.
\end{aligned}
$$

To find the standard deviation of X, we will calculate $\mathsf{E}[X^2]$.

$$
\begin{aligned}
\mathsf{E}[X^2] &= \int_{\mathbb{R}} t^2 f(t) d\lambda(t) \\
&= \frac{1}{b-a} \int_a^b t^2 d\lambda(t) \\
&= \frac{1}{b-a} \left(\frac{b^3}{3} - \frac{a^3}{3} \right) \\
&= \frac{1}{b-a} \frac{(b-a)(b^2+ab+a^2)}{3} \\
&= \frac{a^2+ab+b^2}{3}.
\end{aligned}
$$

It follows

$$
\begin{aligned}
\mathsf{sd}(X) &= \sqrt{\mathsf{E}[X^2] - \mathsf{E}[X]^2} \\
&= \sqrt{\frac{a^2+ab+b^2}{3} - \left(\frac{a+b}{2} \right)^2} \\
&= \frac{b-a}{2\sqrt{3}}
\end{aligned}
$$

.

The tail region starts on the left side at

$$
T1_X = \mathsf{E}[X] - 3\,\mathsf{sd}(X) = \frac{a+b}{2} - 3\frac{b-a}{2\sqrt{3}}
$$

and on the right side at

$$
T2_X = \mathsf{E}[X] + 3\,\mathsf{sd}(X) = \frac{a+b}{2} + 3\frac{b-a}{2\sqrt{3}}.
$$

We will proove that $TR_X \bigcap [a,b] = \varnothing$. Suppose that

$$\begin{aligned}
& T1_X > a \\
\iff & \frac{a+b}{2} - 3\frac{b-a}{2\sqrt{3}} > a \\
\iff & \frac{a(1+\sqrt{3}) + b(1-\sqrt{3})}{2} > a \\
\iff & b(1-\sqrt{3}) > 2a - a(1+\sqrt{3}) \\
\iff & b(1-\sqrt{3}) > a(1-\sqrt{3}) \\
\iff & b < a
\end{aligned}$$

which disagrees with the condition that $a < b$. Consequently, we have $T1_X < a$. Analogously, we can show that $b < T2_X$.

As a result $TR_X \bigcap [a,b] = \big( ]-\infty, T1_X[\bigcap]T2_X, +\infty[\big) \bigcap [a,b] = \varnothing$. It obviously follows that $p_X^{TR} = 0$. Formally we calculate

$$\begin{aligned}
p_X^{TR} &= \int_{TR_X} \frac{1}{b-a} \mathbf{I}_{[a,b]}(t) d\lambda(t) \\
&= \frac{1}{b-a} \int_{\mathbb{R}} \mathbf{I}_{TR_X}(t) \, \mathbf{I}_{[a,b]}(t) d\lambda(t) \\
&= \frac{1}{b-a} \int_{\mathbb{R}} 0 d\lambda(t) \\
&= 0
\end{aligned}$$

In Figure 1, which shows the bell curve of a uniform distribution on the intervall $[a,b] = [0,1]$ along with the boundaries of the tail region, we see that no probability mass is in the tail region.

3.1.2 Exponential distribution

Let X be a random variable. $\mathrm{P}_X = \mathcal{E}(\theta)$, $(\theta > 0)$ if its density funcion is in the form of

$$f(t) = \theta \; \exp(-\theta t) \mathbf{1}_{]0,+\infty[}(t)$$

The mean is

$$\begin{aligned}
\mathsf{E}[X] &= \int_{\mathbb{R}} t f(t) d\lambda(t) \\
&= \int_{\mathbb{R}_+} \theta t \exp(-\theta t) d\lambda(t) \\
&= \theta \left(\left[-\frac{1}{\theta} t \exp(-\theta t) \right]_0^{+\infty} - \int_{\mathbb{R}_+} -\frac{1}{\theta} \exp(-\theta t) d\lambda(t) \right) \\
&= \int_{\mathbb{R}_+} \exp(-\theta t) d\lambda(t) \\
&= -\frac{1}{\theta} \left[\exp(-\theta t) \right]_0^{+\infty} \\
&= \frac{1}{\theta}
\end{aligned}$$

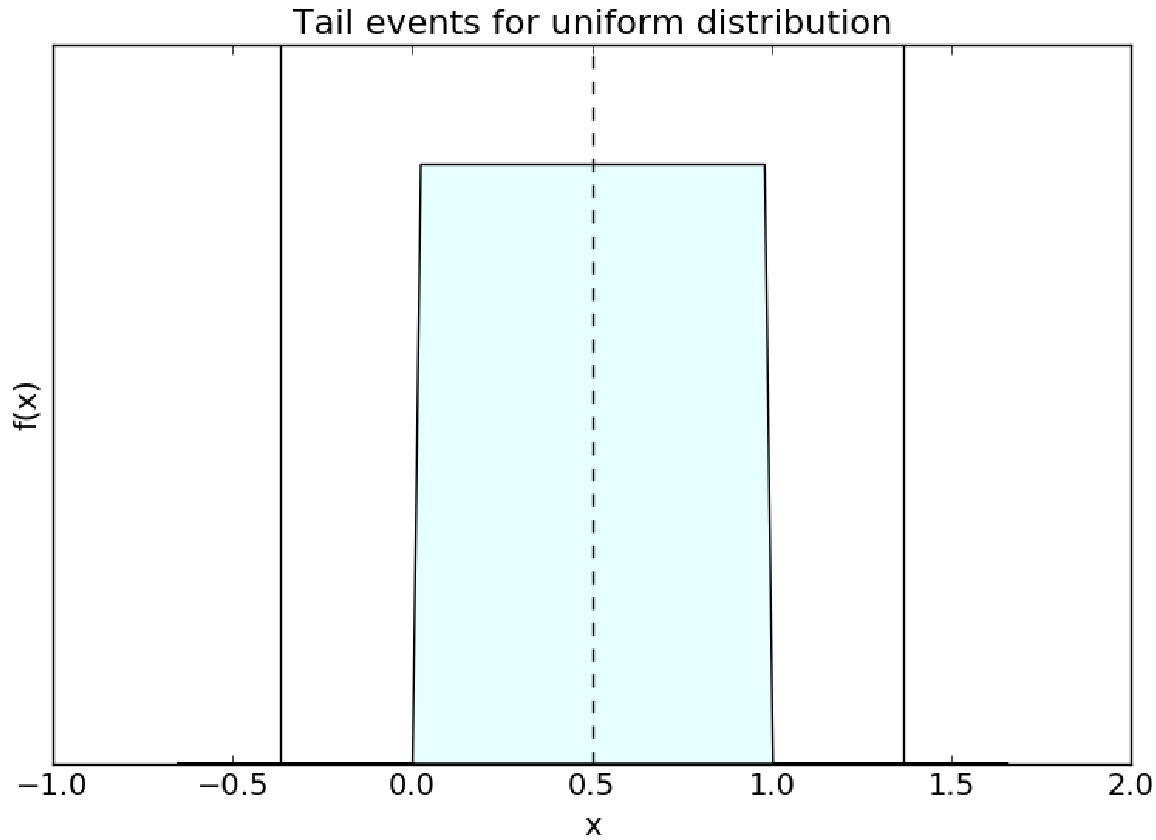

Figure 1: Bell curve of a uniform distribution $P_X = \mathcal{U}[0,1]$ with mean as dashed line, tail boundaries as solid lines [*TailEventGraphs*]

The second moment is

$$\begin{aligned}
\mathsf{E}[X^2] &= \int_{\mathbb{R}} t^2 f(t) d\lambda(t) \\
&= \theta \int_{\mathbb{R}_+} t^2 e^{-\theta t} d\lambda(t) \\
&= \theta \left(\left[-\frac{1}{\theta} t^2 \exp(-\theta t) \right]_0^{+\infty} - \int_{\mathbb{R}_+} -\frac{2t}{\theta} \exp(-\theta t) d\lambda(t) \right) \\
&= 2 \int_{\mathbb{R}_+} t \exp(-\theta t) d\lambda(t) \\
&= \frac{2}{\theta} \int_{\mathbb{R}_+} \theta t \exp(-\theta t) d\lambda(t) \\
&= \frac{2}{\theta^2}
\end{aligned}$$

The standard deviation is

$$\begin{aligned}
\mathsf{sd}(X) &= \sqrt{\mathsf{E}[X^2] - (\mathsf{E}[X])^2} \\
&= \sqrt{\frac{2}{\theta^2} - \left(\frac{1}{\theta}\right)^2} \\
&= \frac{1}{\theta}
\end{aligned}$$

and the lower (resp. upper) tail is $T1_X = -\frac{2}{\theta}$ (resp. $T2_X = \frac{4}{\theta}$).

When calculating p_X^{TR}, we will see that it is non-zero, but that the lower part of the tail region has no mass, so that the probability of the variable being in that lower part of TR_X is zero. In many practical cases one is more interested in the probability of the lower tail.

$$\begin{aligned}
p_X^{TR} &= \int_{TR_X} f(t) d\lambda(t) \\
&= \int_{TR_X} \theta \exp(-\theta t) \mathbf{1}_{\mathbb{R}_+}(t) d\lambda(t) \\
&= \theta \int_{TR_X \bigcap \mathbb{R}_+} \exp(-\theta t) d\lambda(t) \\
&= \theta \int_{]\frac{4}{\theta}, +\infty[} \exp(-\theta t) d\lambda(t) \\
&= [-\exp(-\theta t)]_{\frac{4}{\theta}}^{+\infty} \\
&= \exp(-4) \\
&\approx 0.0183
\end{aligned}$$

Figure 2 shows the bell curve of an exponential distribution with parameter $\theta = 1$ along with the tail region. Note that, as we have correctly calculated, there is no probability mass in the lower part of the tail region.

3.1.3 Laplace distribution

The density function of a Laplace distributed random variable, in contrast to the normal distribution that we will treat hereafter, expresses the absolute value of the distance to the mean, rather than the squared distance.

Let X be a random variable, so that $\mathrm{P}_X = \mathcal{L}(\mu, b)$, where $\mu \in \mathbb{R}$ and $b > 0$. Then,

$$\forall t \in \mathbb{R},\ f(t) = \frac{1}{2b} \exp\left(-\frac{|t-\mu|}{b}\right).$$

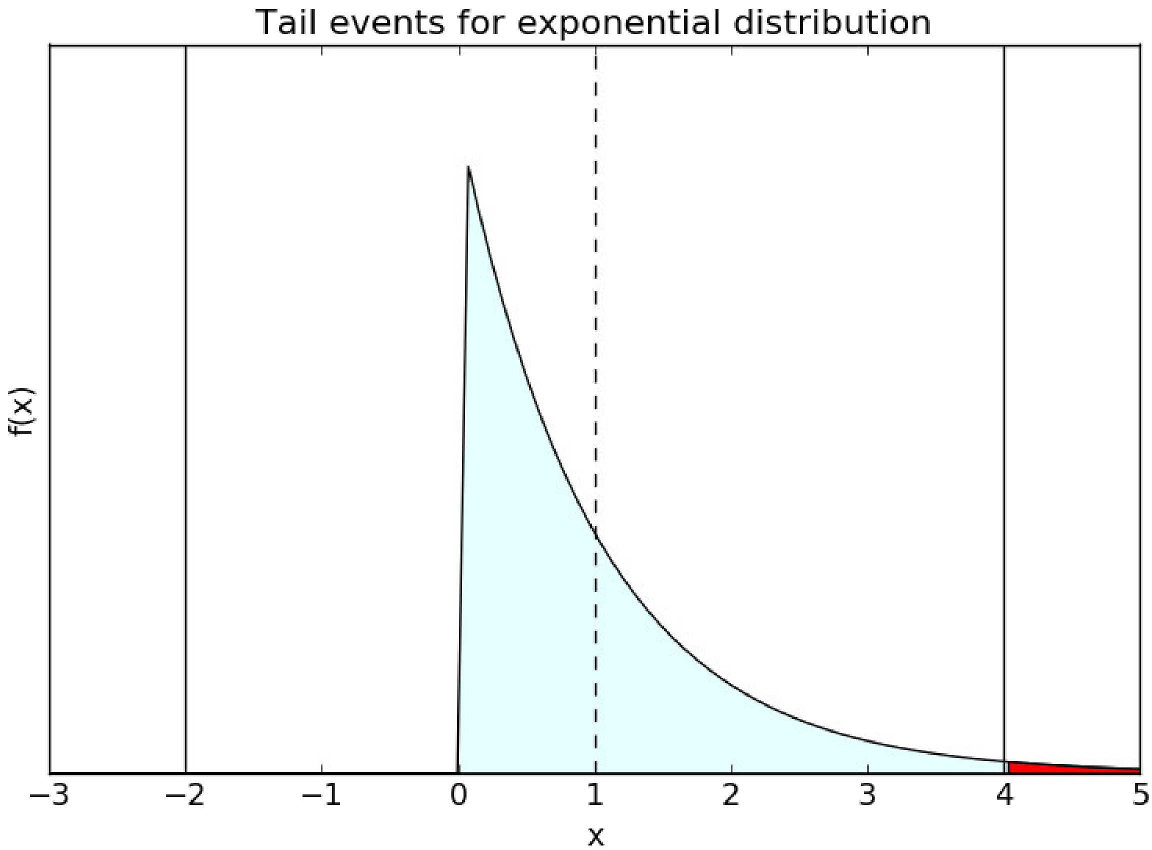

Figure 2: Bell curve of an exponential distribution $P_X = \mathcal{E}(1)$ with mean as dashed line, tail boundaries as solid lines and tail region in red [TailEventGraphs]

We'll show that $\mathsf{E}[X] = \mu$ and that $\mathsf{sd}(X) = \sqrt{2}b$.

$$
\begin{aligned}
\mathsf{E}[X] &= \int_{\mathbb{R}} t f(t) d\lambda(t) \\
&= \frac{1}{2b} \int_{\mathbb{R}} t \exp\left(-\frac{|t-\mu|}{b}\right) d\lambda(t) \\
&= \frac{1}{2b} \left\{ \int_{-\infty}^{\mu} t \exp\left(-\frac{\mu - t}{b}\right) d\lambda(t) + \int_{\mu}^{+\infty} t \exp\left(-\frac{t-\mu}{b}\right) d\lambda(t) \right\} \\
&= \frac{1}{2b} \left\{ b \left[t \exp\left(\frac{t-\mu}{b}\right) \right]_{-\infty}^{\mu} - b \int_{-\infty}^{\mu} \exp\left(\frac{t-\mu}{b}\right) d\lambda(t) \right\} \\
&\qquad + \frac{1}{2b} \left\{ (-b) \left[t \exp\left(\frac{\mu - t}{b}\right) \right]_{\mu}^{+\infty} - (-b) \int_{\mu}^{+\infty} \exp\left(\frac{\mu - t}{b}\right) d\lambda(t) \right\} \\
&= \frac{1}{2} \left\{ \mu - b \left[\exp\left(\frac{t-\mu}{b}\right) \right]_{-\infty}^{\mu} \right\} + \frac{1}{2} \left\{ \mu - b \left[\exp\left(\frac{\mu - t}{b}\right) \right]_{\mu}^{+\infty} \right\} \\
&= \frac{1}{2} (\mu - b + \mu + b) \\
&= \mu
\end{aligned}
$$

$$
\begin{aligned}
\mathsf{E}[X^2] &= \int_{\mathbb{R}} t^2 f(t) d\lambda(t) \\
&= \frac{1}{2b} \int_{\mathbb{R}} t^2 \exp\left(-\frac{|t-\mu|}{b}\right) d\lambda(t) \\
&= \frac{1}{2b} \left\{ \int_{-\infty}^{\mu} t^2 \exp\left(-\frac{\mu - t}{b}\right) d\lambda(t) + \int_{\mu}^{+\infty} t^2 \exp\left(-\frac{t-\mu}{b}\right) d\lambda(t) \right\} \\
&= \frac{1}{2b} \left\{ b \left[t^2 \exp\left(\frac{t-\mu}{b}\right)\right]_{-\infty}^{\mu} - 2b \int_{-\infty}^{\mu} t \exp\left(\frac{t-\mu}{b}\right) d\lambda(t) \right\} \\
&\quad + \frac{1}{2b} \left\{ (-b) \left[t^2 \exp\left(\frac{\mu - t}{b}\right)\right]_{\mu}^{+\infty} - 2(-b) \int_{\mu}^{+\infty} t \exp\left(\frac{\mu - t}{b}\right) d\lambda(t) \right\} \\
&= \frac{1}{2} \left\{ \left[t^2 \exp\left(\frac{t-\mu}{b}\right)\right]_{-\infty}^{\mu} - \left[t^2 \exp\left(\frac{\mu - t}{b}\right)\right]_{\mu}^{+\infty} \right\} \\
&\quad - \int_{-\infty}^{\mu} t \exp\left(\frac{t-\mu}{b}\right) d\lambda(t) + \int_{\mu}^{+\infty} t \exp\left(\frac{\mu - t}{b}\right) d\lambda(t) \\
&= \mu^2 - b(\mu - b) + b(\mu + b) \\
&= \mu^2 + 2b^2.
\end{aligned}
$$

So we get

$$
\mathsf{sd}(X) = \sqrt{\mathsf{E}[X^2] - (\mathsf{E}[X])^2} = \sqrt{\mu^2 + 2b^2 - \mu^2} = \sqrt{2}b.
$$

The tail region is therefore,

$$
TR_X = \left[\mu - 3\sqrt{2}b, \mu + 3\sqrt{2}b\right]^C
$$

We calculate the tail risk

$$
\begin{aligned}
p_X^{TR} &= 1 - \mathrm{P}\left(X \in \left[\mu - 3\sqrt{2}b, \mu + 3\sqrt{2}b\right]\right) \\
&= 1 - \int_{\mu - 3\sqrt{2}b}^{\mu + 3\sqrt{2}b} \frac{1}{2b} \exp\left(-\frac{|t-\mu|}{b}\right) d\lambda(t)
\end{aligned}
$$

Suppose $u = t - \mu$. Then $du = dt$ and the integral changes to

$$
\begin{aligned}
p_X^{TR} &= 1 - \int_{-3\sqrt{2}b}^{3\sqrt{2}b} \frac{1}{2b} \exp\left(-\frac{|u|}{b}\right) d\lambda(t) \\
&= 1 - 2 \int_{0}^{3\sqrt{2}b} \frac{1}{2b} \exp\left(-\frac{u}{b}\right) d\lambda(t) \\
&= 1 + \left[\exp(-\frac{u}{b})\right]_0^{3\sqrt{2}b} \\
&= \exp\left(-3\sqrt{2}\right) \\
&\approx 0.0144.
\end{aligned}
$$

Figure 3 shows the bell curve of a Laplace distribution with parameters $(\mu, b) = (0, 1)$ along with the tail region.

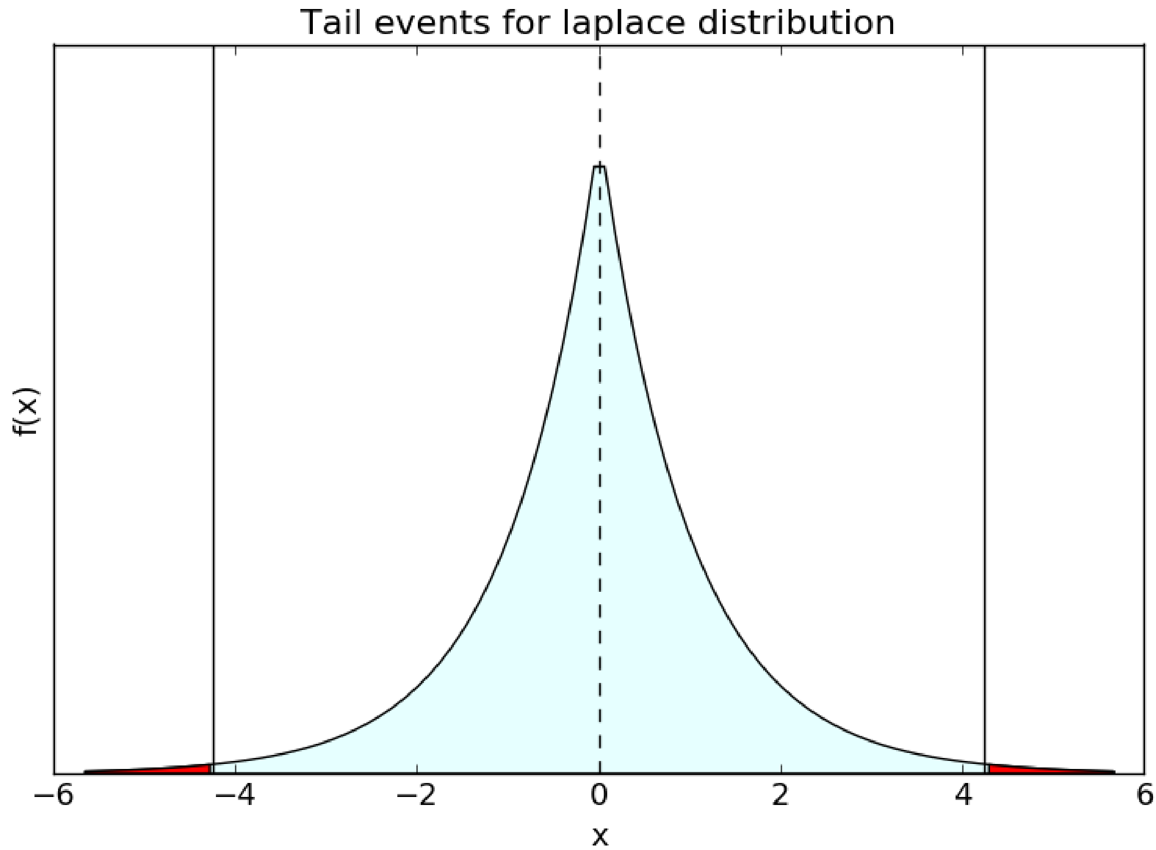

Figure 3: Bell curve of a Laplacel distribution $P_X = \mathcal{L}(0,1)$ with mean as dashed line, tail boundaries as solid lines and tail region in red [*TailEventGraphs*]

3.1.4 Normal distribution

Let X be a random variable, following a normal distribution with the tupel of parameters (μ, σ^2), i.e. $\mathrm{P}_X = \mathsf{N}(\mu, \sigma^2)$. Therefore, its density function g is given by

$$\forall t \in \mathbb{R}\ g(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right)$$

Suppose that $\sigma \neq 0$. We note $Z \overset{\text{def}}{=} \frac{X-\mu}{\sigma}$. Then $\mathrm{P}_Z = \mathsf{N}(0,1)$ with the density function f given by

$$\forall t \in \mathbb{R}\ f(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right)$$

.

We will proove that $\mathsf{E}[Z] = 0$.

$$\begin{aligned}\mathsf{E}[Z] &= \int_{\mathbb{R}} t f(t) d\lambda(t) \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \underbrace{t \exp\left(-\frac{t^2}{2}\right)}_{\overset{\text{def}}{=}\psi(t)} d\lambda(t)\end{aligned}$$

We know that ψ is odd, i.e.

$$\forall t \in \mathbb{R} \ \psi(t) = -\psi(-t).$$

Subsequently we get for any $M \in \mathbb{R}_+$

$$\int_{-M}^{M} \psi(t) d\lambda(t) = 0 \tag{1}$$

We conclude that

$$\begin{aligned}
\mathsf{E}[Z] &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \psi(t) d\lambda(t) \\
&= \frac{1}{\sqrt{2\pi}} \lim_{M \to +\infty} \int_{-M}^{M} \psi(t) d\lambda(t) \\
&\overset{(1)}{=} \frac{1}{\sqrt{2\pi}} \lim_{M \to +\infty} 0 \\
&= 0
\end{aligned}$$

We omit the proof that $\mathsf{sd}(Z) = 1$

The tail region of Z is simply $TR_Z = [-3, 3]^C$. Let us calculate the tail risk.

$$\begin{aligned}
p_Z^{TR} &= \mathrm{P}(Z \in [-3, 3]^C) \\
&= 1 - \mathrm{P}(Z \in [-3, 3]) \\
&= 1 - \int_{-3}^{3} f(t) d\lambda(t) \\
&= 1 - \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-3}^{3} \exp\left(-\frac{t^2}{2}\right) d\lambda(t)}_{I}
\end{aligned}$$

We cannot integrate I as it is. Instead, we will transform it and use the following numerical approximation of the error function which can be found in chapter 7.1 of Abramowitz and Stegun (1964).

$$\forall x \in \mathbb{R}_+ \ \operatorname{erf}(x) = 1 - \left(a_1 t + a_2 t^2 + a_3 t^3\right) exp(-x^2) + \epsilon(x), \ t = \frac{1}{1 + px} \tag{2}$$

We know that $|\epsilon(x)| \leq 2,5 \times 10^{-5}$ if we use

$$\begin{cases} p &= 0.47047 \\ a_1 &= 0.3480242 \\ a_2 &= -0.0958798 \\ a_3 &= 0.7478556 \end{cases} \tag{3}$$

$$\begin{aligned}
I &= \frac{1}{\sqrt{2\pi}} \int_{-3}^{3} \underbrace{\exp\left(-\frac{t^2}{2}\right)}_{\phi} d\lambda(t) \\
&\overset{\phi \text{ even}}{=} \frac{2}{\sqrt{2\pi}} \int_{0}^{3} \exp\left(-\frac{t^2}{2}\right) d\lambda(t) \\
&= \frac{2}{\sqrt{2\pi}} \int_{0}^{3} \exp\left\{-\left(\frac{t}{\sqrt{2}}\right)^2\right\} d\lambda(t)
\end{aligned}$$

We use the substitution

$$\frac{t}{\sqrt{2}} = x, \ \frac{1}{\sqrt{2}} d\lambda(t) = d\lambda(x)$$

and get

$$\begin{aligned} I &= \frac{2}{\sqrt{\pi}} \int_0^{\frac{3}{\sqrt{2}}} \exp\left(-x^2\right) d\lambda(x) \\ &= \operatorname{erf}\left(\frac{3}{\sqrt{2}}\right) \end{aligned}$$

(2) and (3) yield

$$\operatorname{erf}\left(\frac{3}{\sqrt{2}}\right) \approx 0.99729$$

We conclude that

$$\begin{aligned} p_Z^{TR} &= 1 - I \\ &= 1 - \operatorname{erf}(\frac{3}{\sqrt{2}}) \\ &\approx 1 - 0.99729 \\ &= 0.00271 \end{aligned}$$

This calculation is rather difficult and time consuming. Tables with the probabilities of the standard normal distribution can be found for a variety of values in most intoductary statistics books, for example Dixon and Massey Frank (1950). There the probability is given by $\mathrm{P}(Z \in [-3, 3]) = 0.9973$, which leads to a tail risk of $p_Z^{TR} = 0.0027$.
Figure 4 shows the bell curve of a standard normal distribution along with the tail region.

3.2 Results

The results of our analysis contain some intuitive information about the notion of tail events. First of all, the continuous uniform distribution has no tail risk at all. All of its probability mass is on the interval of its parameters. Any event outside of the interval is impossible. When we move three times the standard deviation from the mean, then we are outside of the interval. This becomes especially clear when looking at the standard deviation. We calculated it as the distance from one of the two parameters (the boundaries of the interval) to the mean divided by the square root of three. When multiplying it by square root of three, we already get the distance from the mean to either one of the parameters, yielding in probability mass of zero beyond that point.

Remarkable about the exponential distribution is that it is a one sided distribution. Therefore, it is no surprise that it does not have any probability mass in the left part of the tail region, as the bell curve has no left tail, as can be seen in Figure 2 .

The Laplace distribution is sometimes referred to as double exponential distribution, as it seems to be mirroring the exponential distribution on $\mathbb{R}_-$. One might therefore expect its tail risk to be twice as large as the tail risk of an exponential distribution. Obviously that is not the case. The difference lies in the distribution function. As we have defined in the first

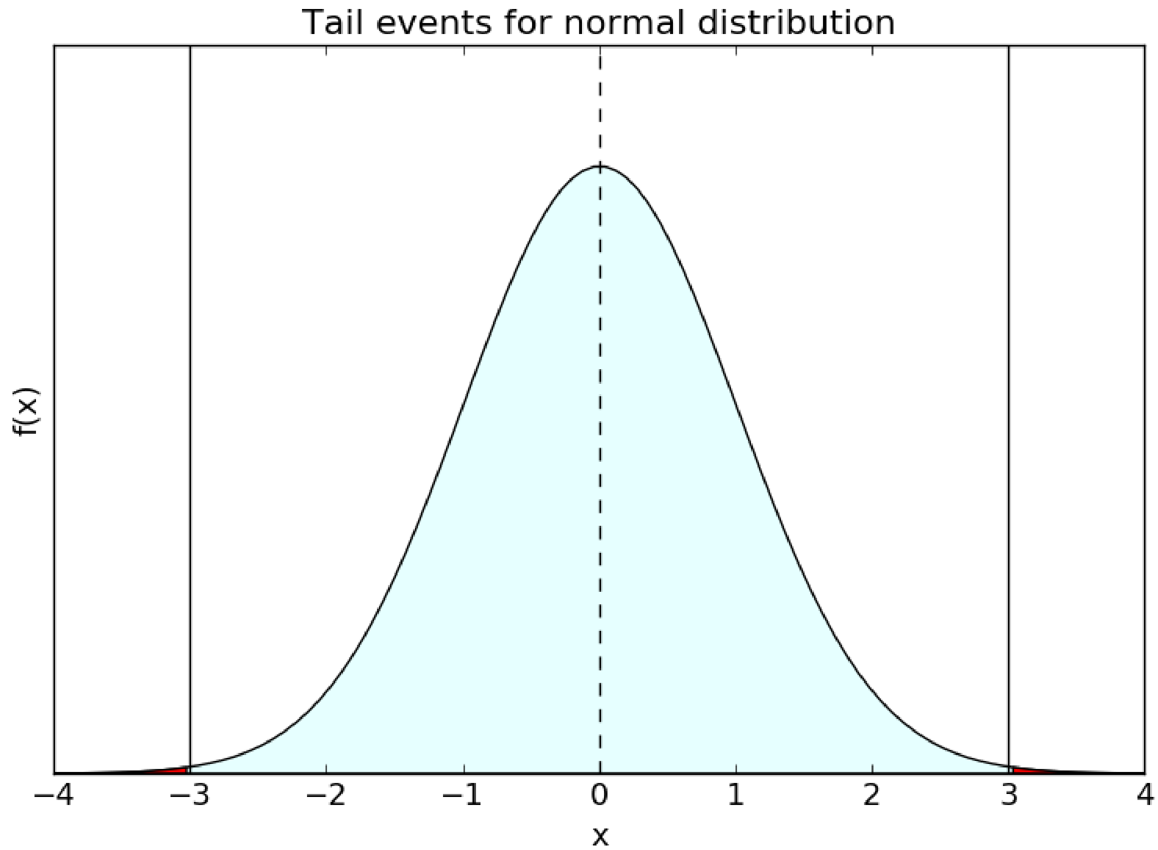

Figure 4: Bell curve of the standard normal distribution with mean as dashed line, tail boundaries as solid lines and tail region in red [*TailEventGraphs*]

chapter, integrating the distribution function of a real random variable over $\mathbb{R}$ generates one. In other words, the area under the bell curve of an exponential distribution is as big as the area under the bell curve of a Laplace distribution. Therefore, we have the factor 0.5 in the Laplace distribution function and as a result it has a larger standard deviation (which still depends on the parameter) than an exponential distribution. This yields smaller tail regions and thus smaller tail risk.

Finally the normal distribution differs from the Laplace distribution as its densit function expresses the squared distance from the mean instead of the absolute distance. This leads in fact to smaller tail regions, ergo less tail risk.

The following table summarizes our results.

Distribution	Parameters	Mean	Standard Deviation	Tail risk in % (≈)
$\mathcal{U}[a,b]$	$a, b \in \mathbb{R},\ a < b$	$\frac{a+b}{2}$	$\frac{b-a}{2\sqrt{3}}$	0
$\mathcal{E}(\theta)$	$\theta > 0$	$\frac{1}{\theta}$	$\frac{1}{\theta}$	1.83
$\mathcal{L}(\mu, b)$	$\mu \in \mathbb{R},\ b > 0$	μ	$\sqrt{2}b$	1.44
$\mathsf{N}(\mu, \sigma^2)$	$\mu \in \mathbb{R},\ \sigma^2 > 0$	μ	σ	0.27

Table 1: Table of different distributions with tail events [*Distributiontable*]

4 Conclusion

We have determined the tail regions and calculated its probability masses for four different basic continuous distributions. All distributions had a tail risk well below two percent.

Especially, the normal distribution which plays a significant role in financial modeling has a low tail risk of 0.27%. With regard to the rather recent frequency of financial crises occurring which can effect entire portfolios, it tends to underestimate the risk of a total devaluation of a portfolio. Alternative modeling using fat-tailed distribution should be considered.

References

Abramowitz, M. and I. A. Stegun (1964): *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, vol. 55, Courier Corporation.

Black, F. and M. Scholes (1973): "The pricing of options and corporate liabilities," *Journal of political economy*, 81, 637–654.

Dixon, W. J. and J. Massey Frank (1950): *Introduction To Statistical Analsis*, McGraw-Hill Book Company, Inc; New York.

Gut, A. (2012): *Probability: a graduate course*, vol. 75, Springer Science & Business Media.

Mandelbrot, B. (1963): "New methods in statistical economics," *Journal of political economy*, 71, 421–440.

Markowitz, H. M. (1991): "Foundations of portfolio theory," *The journal of finance*, 46, 469–477.

Stroock, D. W. (2010): *Probability theory: an analytic view*, Cambridge university press.