What is Risk Management
Risk management is the process of understanding and managing, to an acceptable level, the risk of loss to a company's earnings and balance sheet.

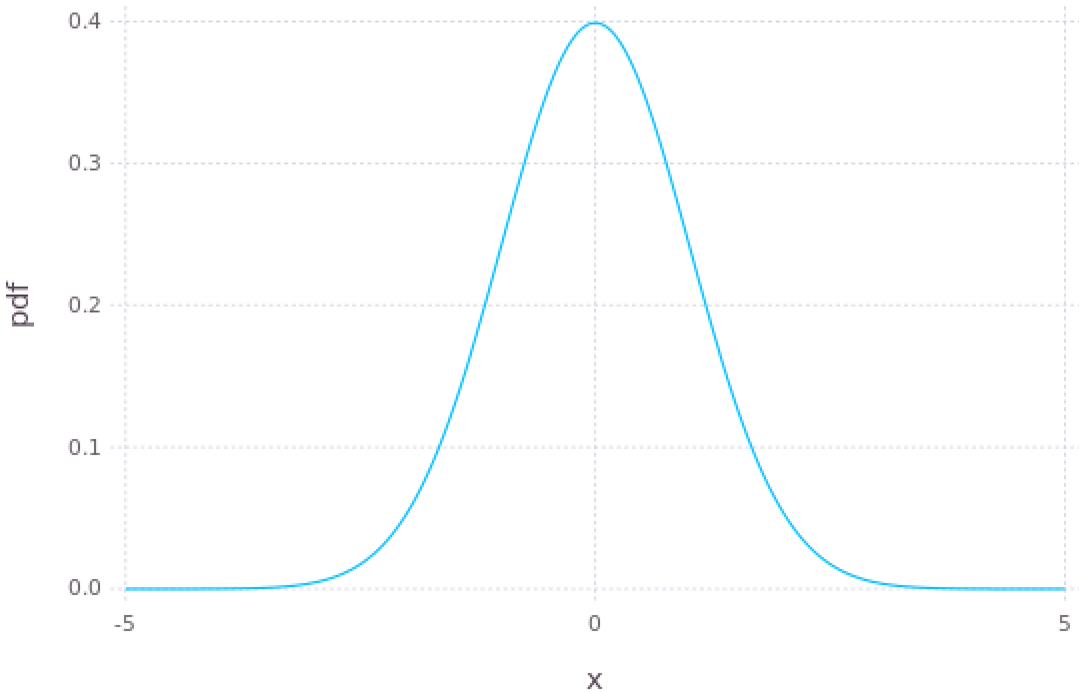
As Wikipedia puts it:
Risk management is the identification, evaluation, and prioritization of risks followed by coordinated and economical application of resources to minimize, monitor, and control the probability or impact of unfortunate events or to maximize the realization of opportunities.

We are concerned with the distribution of outcomes. We are NOT only interested in the expected, or forecasted, outcome.

Type of Risk
- Market Risk
- Credit Risk
- Liquidity Risk
- Operational Risk
- Business Risk

Univariate Statistics
To manage risks, we need to understand the distribution of outcomes.

Probability Density Function, PDF(x), measures the relative likelihood of a the value a random variable (x).



Cumulative Density Function, CDF(x), measures the probability a value of random variable X will be less than or equal to some value x

$$F_X(x) = \mathrm{P}(X \leq x)$$

$F_X(x)$ is the CDF(x) function. $f_X(x)$ is the PDF function.

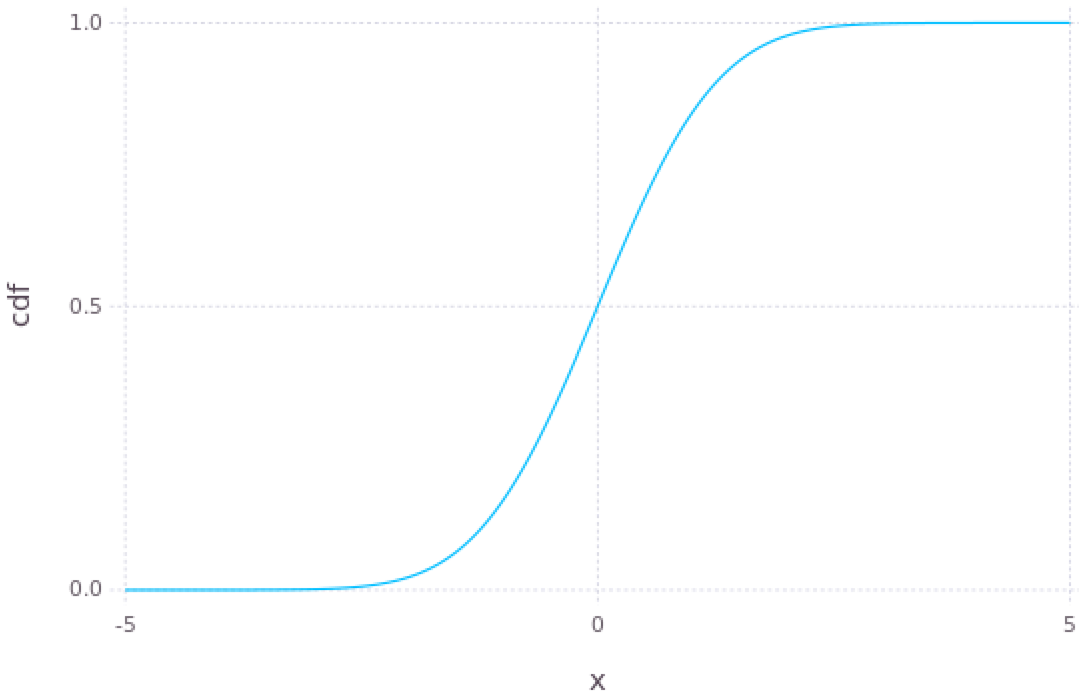
The CDF is the integral of the PDF over the range.

In the case of the above, Normal, distribution, the lower range is unbounded.

$$F_X(x) = \int_{-\infty}^{x} f_x(t)\,dt$$

If we are interested the the probability of a value being in an interval

$$P(a < x \leq b) = \int_a^b f_x(t)\,dt = \int_{-\infty}^b f_x(t)\,dt - \int_{-\infty}^a f_x(t)\,dt = F_X(b) - F_x(a)$$



Quantile Function is the inverse of the CDF. It is the value of x such that F(X) = p

$$F^{-1}(p)$$

This is useful in numerous ways, starting with hypothesis testing.

For example, given a distribution of X, what is the range of X such that 95% of random values fall inside the range?

$$(F^{-1}(0.025), F^{-1}(0.975))$$

Distribution moments describe the shape of the PDF. Generally

$$\mu_n = \int_{-\infty}^{\infty} (x - c)^n f(x)dx = E[(x - c)^n]$$

Where E is the expectation operator. Adjust the integral range for the support of the distribution.

In statistics when talking about the first moment, we set $c = 0$. The first moment is the mean, usually denoted μ.

The higher moments we set $c = \mu$.

The first 4 moments of a distribution are the most used, but do not, by themselves, define a distribution.

Moment:

1. The mean. The expected value of the distribution
2. Variance (σ^2). This measures the dispersion of the distribution around the mean.
3. Skewness. Measures the asymmetry of the distribution. Positive skew distributions skew to the right with the mean > median. Negative skew distributions skew to the left with mean < median.
4. Kurtosis. Measures the fattness of the tails.

Skewness and Kurtosis as generally reported are standardized values.

$$\hat{\mu}_3 = E[(\frac{X-\mu}{\sigma})^3] = \frac{u_3}{\sigma^3}$$

$$\hat{\mu}_4 = E[(\frac{X-\mu}{\sigma})^4] - 3 = \frac{u_4}{\sigma^4} - 3$$

We subtract 3 from the Kurtosis to report “Excess” Kurtosis. The kurtosis of the Normal Distribution is 3 – excess kurtosis is kurtosis in excess of the normal distribution.

Moments can be infinite or undefined.

Estimation of moments from samples are messy.

The Mean

$$\hat{\mu}_1 = E[X] = \frac{1}{n}\sum_{i}^{n} x_i$$

Variance – in small samples, dividing by n leads to a biased estimator.

$$\hat{\mu}_2 = E[(X - \hat{\mu}_1)^2] = \frac{1}{n-1}\sum_{i}^{n}(x_i - \hat{\mu}_1)^2$$

As the sample size grows, the unbiased estimator above converges to the biased estimator.

Unbiased estimators for skew and kurtosis are more complicated. See https://modelingwithdata.org/pdfs/moments.pdf for mathematical derivations of estimators.

For reference:

Unnormalized skew

$$\hat{\mu}_3 = E[(X - \hat{\mu}_1)^3] = \frac{n}{(n-1)(n-2)}\sum_{i}^{n}(x_i - \hat{\mu}_1)^3$$

Normalize to the formula above by dividing by σ^3

Unnormalized kurtosis

$$\hat{\mu}_4 = E[(X - \hat{\mu}_1)^4] = \frac{n^2}{(n-1)^3(n^2-3n+3)}[\,[(n(n-1)^2 + (6n-9))]K_{\bar{x}}(x) - n(6n-9)\sigma_{\bar{x}}^4(x)$$

Where

$K_{\bar{x}}(x)$ is the biased estimator for kurtosis (see link for formula)

$\sigma_{\bar{x}}^4(x)$ is the square of the biased estimator for variance.

Statistical packages may or may not provide the bias corrected estimates.

In Julia, the 3rd and 4th moment estimators are presented as the 3rd and 4th central moment population estimators. The skew metric can be easily adjusted. The kurtosis estimator is more involved as it depends on the population kurtosis, the population variance, as well as the kurtosis of the mean estimator.

In general, the differences are minor.

Properties of Common Distributions

Normal or Gaussian Distribution is commonly used in statistics. Properties of the normal distribution make it tractable to many calculations. Further through the Central Limit Theorem, it is the limiting distribution of sample means.

Notation

$$X \sim N(\mu, \sigma^2)$$

Statistic	**Value**
Support	$x \in (-\infty, \infty)$
Mean	μ
Median	μ
Mode	μ
Variance	σ^2
Skewness	0
Excess Kurtosis	0

PDF

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

If we set μ=0 and σ=1, we call this the Standard Normal.

The CDF for the standard normal is:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2} dt$$

Standard normals are often notated Z

A property of the Normal Distribution is that any linear function of a normal variariable is also Normally distributed.

$$Y = A + BX$$

$$Y \sim N(A + B\mu,\ B^2\sigma^2)$$

Given that, we can transform any normal variable into a Standard normal

$$Z = \frac{(X-\mu)}{\sigma}$$

The CDF of any normal is then:

$$\Phi(\frac{(X-\mu)}{\sigma})$$

No analytic form exists for the value of the integral, therefore numerical approximations are used. The same applies for the quantile function.

A linear function of multiple random normal variables is also normally distributed. For a 2 variable case:

$$Z = Ax + By + c \ \ where\ x\ and\ y\ are\ random\ normals$$

$$Z \sim N(\mu_z, \sigma_z^2)$$

$$\mu_z = A\mu_x + B\mu_y + c$$

$$\sigma_z^2 = A^2\sigma_x^2 + B^2\sigma_y^2 + AB(cov_{x,y})$$

We cover the covariance and the general N variable case next week.

Lognormal Distribution – is the transform of a random normal, X. The exponential of X is distributed Log Normal

$$x \sim N(\mu, \sigma^2)$$

$$Y = e^{(x)} : Y \sim ln(\mu, \sigma^2)$$

Statistic	**Value**
Support	$x \in (0, \infty)$
Mean	$e^{(\mu + \frac{\sigma^2}{2})}$
Median	e^{μ}
Mode	$e^{(\mu - \sigma^2)}$
Variance	$(e^{\sigma^2} - 1)e^{(2\mu + \sigma^2)}$
Skewness	$(e^{\sigma^2} + 2)\sqrt{e^{\sigma^2} - 1}$
Excess Kurtosis	$e^{4\sigma^2} + 2e^{3\sigma^2} + 3e^{2\sigma^2} - 6$

NOTE: The mean and variance of the distribution are not μ **and** σ^2

PDF

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{(-\frac{ln(x-\mu)^2}{2\sigma^2})}$$

CDF

Because we can transform the log normal distribution into a normal distribution with the natural logarithm function, we can use the standard normal CDF function, $\Phi(x)$

$$F(x) \ = \ \Phi(\frac{ln(x) - \mu}{\sigma})$$

Student's-t Distribution or t distribution is used for hypothesis testing. Specifically it is the distribution of the sample mean when the variance is unknown. As the number of observations increases, the Student's-t distribution converges to the normal distribution.

The t distribution has 3 parameters

1. ν - degrees of freedom
2. μ - location
3. σ - scale

Usually only the 1st parameter is used as the standardized T ($\mu = 0$ and $\sigma = 1$) is used for hypothesis testing. I will reference the standardized distribution as T and the generalized as t. Like the Normal, the standardized T can be transformed with the location and scale:

$$X \ = \ \mu + \ \sigma T$$

Notation:

$$X \sim t(\nu, \ (\mu, \ \sigma))$$

Statistic	**Value**
Support	$x \ \in \ (-\ \infty, \ \infty)$
Mean	μ if $\nu > \ 1$, otherwise undefined
Median	μ
Mode	μ
Variance	$\sigma^2 \frac{\nu}{\nu - 2}$ if $\nu > 2$, $\infty \ if \ 1 < \ \nu \le 2$, otherwise undefined
Skewness	0 if $\nu > 3$, otherwise undefined
Excess Kurtosis	$\frac{6}{\nu - 4}$, if $\nu > 4$, $\infty \ if \ 2 < \ \nu \le 4$, otherwise undefined

Note that σ (σ^2) is not the standard deviation (variance).

PDF

$$f(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sigma\sqrt{\pi\nu}}\left(1 + \frac{1}{\nu}\left(\frac{x-\mu}{\sigma}\right)^2\right)^{-\frac{\nu+1}{2}}$$

Notes

The t distribution has $\nu - 1$ finite moments.

$$\lim_{\nu \to \infty} t(\nu, \mu, \sigma) = N(\mu, \sigma^2)$$

The T distribution is defined as

$$T = Z\sqrt{\frac{\nu}{V}}$$

Where
1. $Z \sim N(0, 1)$ - the standard normal
2. $V \sim \chi^2(\nu)$ - Chi Square distribution with ν degrees of freedom (independently study the Chi Square)
3. $Z\ and\ V$ are independant

As a test statistic

$$T = \left(\overline{X} - \mu_0\right)\frac{\sqrt{n}}{S_n}$$

Where
1. $\overline{X}$ is the sample mean
2. S_n is the sample standard deviation
3. n is the number of samples
4. μ_0 is the hypothesis value
5. $T \sim T(n - 1)$ - Standardized T with $\nu = n - 1$

Recognize this is standardizing the value of $\overline{X}$ with a hypothesized mean and calculated variance.

Example: Test if the kurtosis calculated by your distribution package is biased.

Steps

1. Sample 100,000 standardized random normal values.
2. Calculate the kurtosis
3. Sample the kurtosis by repeating steps 1 and 2 100 times.
4. Calculate the mean kurtosis $\overline{k}$ and standard deviation S_k
5. Calculate the T statistic ($\mu_0 = 0$).
6. Use the CDF function to find the p-value of the absolute value of the statistic and subtract from 1. Multiply the value by 2 because this is a 2 sided test.
7. If the value is lower than your threshold (typically 5%), then you reject the hypothesis that the kurtosis function is unbiased.

Alternatively, use a T test function in your statistics package.

Run that a few times, how confident are you in the result? Try increasing the number of samples.

We know in Julia the kurtosis function should be biased. If it is failing to reject the null hypothesis, why and what can we change?