Introduction to Computational Finance and Financial Econometrics
Probability Theory Review: Part 1

Eric Zivot

June 26, 2012

1 Univariate Random Variables

Defnition: A random variable (rv) X is a variable that can take on a given set of values, called the sample space S_X, where the likelihood of the values in S_X is determined by the variable's probability distribution function (pdf).

Examples

- X = price of microsoft stock next month. $S_X = \{\mathbb{R} : 0 < X \leq M\}$

- X = simple return on a one month investment. $S_X = \{\mathbb{R} : -1 \leq X < M\}$

- $X = 1$ if stock price goes up; $X = 0$ if stock price goes down. $S_X = \{0, 1\}$

1.1 Discrete Random Variables

Definition: A discrete rv X is one that can take on a finite number of n different values $x_1, \cdots, x_n$

Definition: The pdf of a discrete rv X, $p(x)$, is a function such that $p(x) = \Pr(X = x)$. The pdf must satisfy

1. $p(x) \geqslant 0$ for all $x \in S_X$; $p(x) = 0$ for all $x \notin S_X$

2. $\sum_{x \in S_X} p(x) = 1$

3. $p(x) \leqslant 1$ for all $x \in S_X$

State of Economy	S_X = Sample Space	$p(x) = \Pr(X = x)$
Depression	-0.30	0.05
Recession	0.0	0.20
Normal	0.10	0.50
Mild Boom	0.20	0.20
Major Boom	0.50	0.05

Table 1: Discrete Distribution for Annual Return

Example: Probability Distribution for Annual Return on Microsoft

Example: Bernouli Distribution

Consider two mutually exclusive events generically called "success" and "failure".

Let $X = 1$ if success occurs and let $X = 0$ if failure occurs.

Let $\Pr(X = 1) = \pi$, where $0 < \pi < 1$, denote the probability of success. Then $\Pr(X = 0) = 1 - \pi$ is the probability of failure. A mathematical model describing this distribution is

$$p(x) = \Pr(X = x) = \pi^x(1-\pi)^{1-x},\ x = 0, 1.$$

When $x = 0$, $p(0) = \pi^0(1-\pi)^{1-0} = 1 - \pi$ and when $x = 1, p(1) = \pi^1(1-\pi)^{1-1} = \pi$.

1.2 Continuous Random Variables

Definition: A continuous rv X is one that can take on any real value

Definition: The pdf of a continuous rv X is a nonnegative function $f(x)$ such that for any interval A on the real line

$$\Pr(X \in A) = \int_A f(x)dx$$

$\Pr(X \in A) =$ "Area under probability curve over the interval A".

The pdf $f(x)$ must satisfy

1. $f(x) \geqslant 0$; $\int_{-\infty}^{\infty} f(x)dx = 1$

Example: Uniform distribution over $[a, b]$

Let $X \backsim U[a, b]$, where "$\backsim$" means "is distributed as". Then

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

Properties:

$f(x) \geq 0$, provided $b > a$, and

$$\int_{-\infty}^{\infty} f(x)dx = \int_a^b \frac{1}{b-a}dx = \frac{1}{b-a}\int_a^b dx$$
$$= \frac{1}{b-a}[x]_a^b = \frac{b-a}{b-a} = 1$$

1.3 The Cumulative Distribution Function (CDF)

Definition The CDF, F, of a rv X is $F(x) = \Pr(X \leq x)$ and

- If $x_1 < x_2$, then $F(x_1) \leq F(x_2)$
- $F(-\infty) = 0$ and $F(\infty) = 1$
- $\Pr(X \geq x) = 1 - F(x)$
- $\Pr(x_1 < X \leq x_2) = F(x_2) - F(x_1)$
- $\frac{d}{dx}F(x) = f(x)$ if X is a continuous rv.

Example: Uniform distribution over $[0,1]$

$$X \backsim U\,[0,1]$$

$$f(x) = \begin{cases} \frac{1}{1-0} = 1 & \text{for } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Then

$$\begin{aligned} F(x) = \Pr(X \le x) &= \int_0^x dz \\ &= [z]_0^x = x \end{aligned}$$

and, for example,

$$\begin{aligned} \Pr(0 \le X \le 0.5) &= F(0.5) - F(0) \\ &= 0.5 - 0 = 0.5 \end{aligned}$$

Note

$$\frac{d}{dx}F(x) = 1 = f(x)$$

Remark:

For a continuous rv

$$\mathsf{Pr}(X \leq x) = \mathsf{Pr}(X < x)$$
$$\mathsf{Pr}(X = x) = 0$$

1.4 Quantiles of a Distribution

X is a rv with continuous CDF $F_X(x) = \Pr(X \leq x)$

Definition: The $\alpha * 100\%$ quantile of F_X for $\alpha \in [0, 1]$ is the value q_α such that

$$F_X(q_\alpha) = \Pr(X \leq q_\alpha) = \alpha$$

The area under the probability curve to the left of q_α is α. If the inverse CDF F_X^{-1} exists then

$$q_\alpha = F_X^{-1}(\alpha)$$

Note: F_X^{-1} is sometimes called the "quantile" function.

Example:

$$
\begin{aligned}
1\% \text{ quantile} &= q_{.01} \\
5\% \text{ quantile} &= q_{.05} \\
50\% \text{ quantile} &= q_{.5} = \text{ median}
\end{aligned}
$$

Example: Quantile function of uniform distn on [0,1]

$$F_X(x) = x \Rightarrow q_\alpha = \alpha$$
$$q_{.01} = 0.01$$
$$q_{.5} = 0.5$$

1.5 The Standard Normal Distribution

Let X be a rv such that $X \backsim N(0,1)$. Then

$$f(x) = \phi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right), \quad -\infty \leq x \leq \infty$$

$$\Phi(x) = \Pr(X \leq x) = \int_{-\infty}^{x} \phi(z)dz$$

Shape Characteristics

- Centered at zero

- Symmetric about zero (same shape to left and right of zero)

$$\text{Pr}(-1 \leq x \leq 1) = \Phi(1) - \Phi(-1) = 0.67$$
$$\text{Pr}(-2 \leq x \leq 2) = \Phi(2) - \Phi(-2) = 0.95$$
$$\text{Pr}(-3 \leq x \leq 3) = \Phi(3) - \Phi(-3) = 0.99$$

Finding Areas under the Normal Curve

- $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = 1$, via change of variables formula in calculus

- $\Pr(a < X < b) = \int_{a}^{b} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = \Phi(b) - \Phi(a)$, cannot be computed analytically!

- Special numerical algorithms are used to calculate $\Phi(z)$

Excel functions

1. `NORMSDIST` computes $\Pr(X \leq z) = \Phi(z)$ or $p(z) = \phi(z)$

2. `NORMSINV` computes the quantile $z_\alpha = \Phi^{-1}(\alpha)$

R functions

1. `pnorm` computes $\Pr(X \leq z) = \Phi(z)$

2. `qnorm` computes the quantile $z_\alpha = \Phi^{-1}(\alpha)$

3. `dnorm` computes the density $\phi(z)$

Some Tricks for Computing Area under Normal Curve

$N(0, 1)$ is symmeric about 0; total area $=1$

$$\text{Pr}(X \leq z) = 1 - \text{Pr}(X \geq z)$$
$$\text{Pr}(X \geq z) = \text{Pr}(X \leq -z)$$
$$\text{Pr}(X \geq 0) = \text{Pr}(X \leq 0) = 0.5$$

Example In Excel use

$$\begin{aligned}\Pr(-1 \le X \le 2) &= \Pr(X \le 2) - \Pr(X \le -1) \\ &= \texttt{NORMSDIST(2)} - \texttt{NORMSDIST(-1)} \\ &= 0.97725 - 0.15866 = 0.81860\end{aligned}$$

In R use

$$\texttt{pnorm(2)} - \texttt{pnorm(-1)} = 0.81860$$

The 1%, 2.5%, 5% quantiles are

$$\begin{aligned}\text{Excel:} z_{.01} &= \Phi^{-1}(0.01) = \texttt{NORMSINV(0.01)} = -2.33 \\ \text{R} &: \texttt{qnorm(0.01)} = -2.33 \\ \text{Excel:} z_{.025} &= \Phi^{-1}(0.025) = \texttt{NORMSINV(0.025)} = -1.96 \\ \text{R} &: \texttt{qnorm(0.025)} = -1.96 \\ \text{Excel:} z_{.05} &= \Phi^{-1}(.05) = \texttt{NORMSINV(.05)} = -1.645 \\ \text{R} &: \texttt{qnorm(0.05)} = -1.645\end{aligned}$$

1.6 Shape Characteristics of pdfs

- Expected Value or Mean - Center of Mass

- Variance and Standard Deviation - Spread about mean

- Skewness - Symmetry about mean

- Kurtosis - Tail thickness

Expected Value - Discrete rv

$$E[X] = \mu_X = \sum_{x \in S_X} x \cdot p(x)$$
$$= \sum_{x \in S_X} x \cdot \Pr(X = x)$$

$E[X] =$ probability weighted average of possible values of X

Expected Value - Continuous rv

$$E[X] = \mu_X = \int_{-\infty}^{\infty} x \cdot f(x) dx$$

Note: In continuous case, $\sum_{x \in S_X}$ becames $\int_{-\infty}^{\infty}$

Expected value of discrete random variable

Using the discrete distribution for the return on Microsoft stock in Table 1, the expected return is

$$
\begin{aligned}
E[X] &= (-0.3)\cdot(0.05) + (0.0)\cdot(0.20) + (0.1)\cdot(0.5) \\
&\quad + (0.2)\cdot(0.2) + (0.5)\cdot(0.05) \\
&= 0.10.
\end{aligned}
$$

Example: $X \backsim U[1,2]$

$$E[X] = \int_1^2 x dx = \left[\frac{x^2}{2}\right]_1^2$$
$$= \frac{1}{2}[4-1] = \frac{3}{2}$$

Example: $X \backsim N(0,1)$

$$\mu_X = E[X] = \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = 0$$

Expectation of a Function of X

Definition: Let $g(X)$ be some function of the rv X. Then

$$E[g(X)] = \sum_{x \in S_X} g(x) \cdot p(x) \text{ Discrete case}$$

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) \cdot f(x) dx \text{ Continuous case}$$

Variance and Standard Deviation

$$
\begin{aligned}
g(X) &= (X - E[X])^2 = (X - \mu_X)^2 \\
\mathsf{Var}(X) &= \sigma_X^2 = E[(X - \mu_X)^2] = E[X^2] - \mu_X^2 \\
\mathsf{SD}(X) &= \sigma_X = \sqrt{\mathsf{Var}(X)}
\end{aligned}
$$

Note: Var(X) is in squared units of X, and SD(X) is in the same units as X. Therefore, SD(X) is easier to interpret.

Computation of Var(X) **and** SD(X)

$$
\begin{aligned}
\sigma_X^2 &= E[(X-\mu_X)^2] \\
&= \sum_{x \in S_X} (x-\mu_X)^2 \cdot p(x) \text{ if } X \text{ is a discrete rv} \\
&= \int_{-\infty}^{\infty} (x-\mu_X)^2 \cdot f(x)dx \text{ if } X \text{ is a continuous rv} \\
\sigma_X &= \sqrt{\sigma_X^2}
\end{aligned}
$$

Remark: For "bell-shaped" data, σ_X measures the size of the typical deviation from the mean value μ_X.

Example: Variance and standard deviation for a discrete random variable

Using the discrete distribution for the return on Microsoft stock in Table 1 and the result that $\mu_X = 0.1$, we have

$$\begin{aligned}
\mathrm{Var}(X) &= (-0.3 - 0.1)^2 \cdot (0.05) + (0.0 - 0.1)^2 \cdot (0.20) \\
&\quad + (0.1 - 0.1)^2 \cdot (0.5) + (0.2 - 0.1)^2 \cdot (0.2) \\
&\quad + (0.5 - 0.1)^2 \cdot (0.05) \\
&= 0.020 \\
\mathrm{SD}(X) &= \sigma_X = \sqrt{0.020} = 0.141.
\end{aligned}$$

Given that the distribution is fairly bell-shaped we can say that typical values deviate from the mean value of 0.10 by about 0.141

$$\mu \pm \sigma = -0.10 \pm 0.141 = [-0.041,\ 0.241]$$

Example: $X \backsim N(0,1)$.

$$\mu_X = \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = 0$$

$$\sigma_X^2 = \int_{-\infty}^{\infty} (x-0)^2 \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = 1$$

$$\sigma_X = \sqrt{1} = 1$$

$\Rightarrow$ size of typical deviation from $\mu_X = 0$ is $\sigma_X = 1$

The General Normal Distribution

$$
\begin{aligned}
X &\backsim N(\mu_X,\ \sigma_X^2) \\
f(x) &= \frac{1}{\sqrt{2\pi\sigma_X^2}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right) dx,\ \ -\infty \le x \le \infty \\
E[X] &= \mu_X = \text{ mean value} \\
\text{Var}(X) &= \sigma_X^2 = \text{ variance} \\
\text{SD}(X) &= \sigma_X = \text{ standard deviation}
\end{aligned}
$$

Shape Characteristics

- Centered at μ_X

- Symmetric about μ_X

$$
\begin{aligned}
\Pr(\mu_X - \sigma_X \leq X \leq \mu_X + \sigma_X) &= 0.67 \\
\Pr(\mu_X - 2 \cdot \sigma_X \leq X \leq \mu_X + 2 \cdot \sigma_X) &= 0.95 \\
\Pr(\mu_X - 3 \cdot \sigma_X \leq X \leq \mu_X + 3 \cdot \sigma_X) &= 0.99
\end{aligned}
$$

- Quantiles of the general normal distribution:

$$q_\alpha = \mu_X + \sigma_X \cdot \Phi^{-1}(\alpha) = \mu_X + \sigma_X \cdot z_\alpha$$

Remarks:

- $X \backsim N(0,1)$: Standard Normal $\Longrightarrow \mu_X = 0$ and $\sigma_X^2 = 1$

- The pdf of the general Normal is completely determined by values of μ_X and σ_X^2

Finding Areas under General Normal Curve

Excel Functions

- `NORMDIST`(x, μ_X, σ_X,`cumulative`). If `cumulative` = `true`: $\Pr(X \leq x)$ is computed; If `cumulative` = `false`, $f(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2}$ is computed

- `NORMINV`(α, μ_x, σ_x) computes $q_\alpha = \mu_X + \sigma_X z_\alpha$

R Runctions

- simulate data: `rnorm(n, mean, sd)`
- compute CDF: `pnorm(q, mean, sd)`
- compute quantiles: `qnorm(p, mean, sd)`
- compute density: `dnorm(x, mean, sd)`

Standard Deviation as a Measure of Risk

$$
\begin{aligned}
R_A &= \text{monthly return on asset A} \\
R_B &= \text{monthly return on assetB} \\
R_A &\backsim N(\mu_A, \sigma_A^2), R_B \backsim N(\mu_B, \sigma_B^2)
\end{aligned}
$$

where

$$
\begin{aligned}
\mu_A &= E[R_A] = \text{ expected monthly return on asset A} \\
\sigma_A &= \text{SD}(R_A) \\
&= \text{ std. deviation of monthly return on asset A}
\end{aligned}
$$

Typically, if

$$\mu_A > \mu_B$$

then

$$\sigma_A > \sigma_B$$

Example: Why the normal distribution may not be appropriate for simple returns

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \text{simple return}$$

$$\text{Assume } R_t \sim N(0.05, (0.50)^2)$$

Note: $P_t \geq 0 \implies R_t \geq -1$. However, based on the assumed normal distribution

$$\Pr(R_t < -1) = \texttt{NORMDIST(-1,0.05,0.50,TRUE)} = 0.018$$

This implies that there is a 1.8% chance that the asset price will be negative. This is why the normal distribution may not be appropriate for simple returns.

Example: The normal distribution is more appropriate for cc returns

$$r_t = \ln(1 + R_t) = \text{cc return}$$
$$R_t = e^{r_t} - 1 = \text{ simple return}$$
$$\text{Assume } r_t \sim N(0.05, (0.50)^2)$$

Unlike R_t, r_t can take on values less than -1. For example,

$$r_t = -2 \implies R_t = e^{-2} - 1 = -0.865$$
$$\Pr(r_t < -2) = \Pr(R_t < -0.865)$$
$$= \texttt{NORMDIST(-2,0.05,0.50,TRUE)} = 0.00002$$

The Log-Normal Distribution

$$X \sim N(\mu_X, \sigma_X^2), \quad -\infty < X < \infty$$
$$Y = \exp(X) \sim \text{lognormal}(\mu_X, \sigma_X^2),\ 0 < Y < \infty$$
$$E[Y] = \mu_Y = \exp(\mu_X + \sigma_X^2/2)$$
$$\text{Var}(Y) = \sigma_Y^2 = \exp(2\mu_X + \sigma_X^2)(\exp(\sigma_X^2) - 1)$$

Example: log-normal distribution for simple returns

$$r_t \sim N(0.05, (0.50)^2)$$
$$1 + R_t \sim \text{lognormal}(0.05, (0.50)^2)$$
$$\mu_{1+R} = \exp(0.05 + (0.5)^2/2) = 1.191$$
$$\sigma_{1+R}^2 = \exp(2(0.05) + (0.5)^2)(\exp(0.5^2) - 1) = 0.563$$

R Runctions

- simulate data: `rlnorm(n, mean, sd)`
- compute CDF: `plnorm(q, mean, sd)`
- compute quantiles: `qlnorm(p, mean, sd)`
- compute density: `dlnorm(y, mean, sd)`

Skewness - Measure of symmetry

$$
\begin{aligned}
g(X) &= ((X - \mu_X)/\sigma_X)^3 \\
\text{Skew}(X) &= E\left[\left(\frac{X - \mu_X}{\sigma_X}\right)^3\right] \\
&= \sum_{x \in S_X} \left(\frac{x - \mu_X}{\sigma_X}\right)^3 p(x) \text{ if } X \text{ is discrete} \\
&= \int_{-\infty}^{\infty} \left(\frac{x - \mu_X}{\sigma_X}\right)^3 f(x) dx \text{ if } X \text{ is continuous}
\end{aligned}
$$

Intuition

- If X has a symmetric distribution about μ_X then $\mathsf{Skew}(X) = 0$

- $\mathsf{Skew}(X) > 0 \Longrightarrow$ pdf has long right tail, and median $<$ mean

- $\mathsf{Skew}(X) < 0 \Longrightarrow$ pdf has long left tail, and median $>$ mean

Example: Using the discrete distribution for the return on Microsoft stock in Table 1, the results that $\mu_X = 0.1$ and $\sigma_X = 0.141$, we have

$$
\begin{aligned}
\text{skew}(X) &= [(-0.3 - 0.1)^3 \cdot (0.05) + (0.0 - 0.1)^3 \cdot (0.20) \\
&+ (0.1 - 0.1)^3 \cdot (0.5) + (0.2 - 0.1)^3 \cdot (0.2) \\
&+ (0.5 - 0.1)^3 \cdot (0.05)]/(0.141)^3 \\
&= 0.0
\end{aligned}
$$

Example: $X \backsim N(\mu_X, \sigma_X^2)$. Then

$$\text{Skew}(X) = \int_{-\infty}^{\infty} \left(\frac{x - \mu_X}{\sigma_X} \right)^3 \frac{1}{\sqrt{2\pi\sigma_X^2}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2 \right) dx = 0$$

Example: $Y \backsim \text{lognormal}(\mu_X, \sigma_X^2)$. Then

$$\text{Skew}(Y) = \left(\exp(\sigma_X^2) + 2 \right) \sqrt{\exp(\sigma_X^2) - 1} > 0$$

Kurtosis - Measure of tail thickness

$$
\begin{aligned}
g(X) &= ((X - \mu_X)/\sigma_X)^4 \\
\text{Kurt}(X) &= E\left[\left(\frac{X - \mu_X}{\sigma_X}\right)^4\right] \\
&= \sum_{x \in S_X} \left(\frac{x - \mu_X}{\sigma_X}\right)^4 p(x) \text{ if } X \text{ is discrete} \\
&= \int_{-\infty}^{\infty} \left(\frac{x - \mu_X}{\sigma_X}\right)^4 f(x)dx \text{ if } X \text{ is continuous}
\end{aligned}
$$

Intuition

- Values of x far from μ_X get blown up resulting in large values of kurtosis
- Two extreme cases: fat tails (large kurtosis); thin tails (small kurtosis)

Example: Kurtosis for a discrete random variable

Using the discrete distribution for the return on Microsoft stock in Table 1, the results that $\mu_X = 0.1$ and $\sigma_X = 0.141$, we have

$$\begin{aligned}
\text{Kurt}(X) &= [(-0.3 - 0.1)^4 \cdot (0.05) + (0.0 - 0.1)^4 \cdot (0.20) \\
&+ (0.1 - 0.1)^4 \cdot (0.5) + (0.2 - 0.1)^4 \cdot (0.2) \\
&+ (0.5 - 0.1)^4 \cdot (0.05)]/(0.141)^4 \\
&= 6.5
\end{aligned}$$

Example: $X \backsim N(\mu_X, \sigma_X^2)$

$$\text{Kurt}(X) = \int_{-\infty}^{\infty} \left(\frac{x - \mu_X}{\sigma_X} \right)^4 \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-\frac{1}{2}(\frac{x-\mu_X}{\sigma_X})^2} dx = 3$$

Definition: Excess kurtosis = Kurt$(X) - 3$ = kurtosis value in excess of kurtosis of normal distribution.

- Excess kurtosis $(X) > 0 \Rightarrow X$ has fatter tails than normal distribution
- Excess kurtosis $(X) < 0 \Rightarrow X$ has thinner tails than normal distribution

The Student's-t Distribution

A distribution similar to the standard normal distribution but with fatter tails, and hence larger kurtosis, is the Student's t distribution. If X has a Student's t distribution with degrees of freedom parameter v, denoted $X \sim t_v$, then its pdf has the form

$$f(x) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{v\pi}\Gamma\left(\frac{v}{2}\right)}\left(1+\frac{x^2}{v}\right)^{-\left(\frac{v+1}{2}\right)}, \quad -\infty < x < \infty, \ v > 0.$$

where $\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt$ denotes the gamma function.

It can be shown that

$$
\begin{aligned}
E[X] &= 0,\ v > 1 \\
\mathrm{var}(X) &= \frac{v}{v-2},\ v > 2, \\
\mathrm{skew}(X) &= 0,\ v > 3, \\
\mathrm{kurt}(X) &= \frac{6}{v-4} + 3,\ v > 4.
\end{aligned}
$$

The parameter v controls the scale and tail thickness of distribution. If v is close to four, then the kurtosis is large and the tails are thick. If $v < 4$, then $\mathrm{kurt}(X) = \infty$. As $v \to \infty$ the Student's t pdf approaches that of a standard normal random variable and $\mathrm{kurt}(X) = 3$.

R Runctions

- simulate data: `rt(n, df)`

- compute CDF: `pt(q, df)`

- compute quantiles: `qt(p, df)`

- compute density: `dt(x, df)`

Here df is the degrees of freedom parameter v.

1.7 Linear Functions of a Random Variable

Let X be a discrete or continuous rv with $\mu_X = E[X]$, and $\sigma_X^2 = \mathsf{Var}(X)$. Define a new rv Y to be a linear function of X :

$$Y = g(X) = a \cdot X + b$$

a and b are known constants

Then

$$\begin{aligned}
\mu_Y &= E[Y] = E[a \cdot X + b] \\
&= a \cdot E[X] + b = a \cdot \mu_X + b \\
\sigma_Y^2 &= \mathsf{Var}(Y) = \mathsf{Var}(a \cdot X + b) \\
&= a^2 \cdot \mathsf{Var}(X) \\
&= a^2 \cdot \sigma_X^2 \\
\sigma_Y &= a \cdot \sigma_X
\end{aligned}$$

Linear Function of a Normal rv

Let $X \backsim N(\mu_X, \sigma_X^2)$ and define $Y = a \cdot X + b$. Then

$$Y \sim N(\mu_Y, \sigma_Y^2)$$

with

$$\begin{aligned}\mu_Y &= a \cdot \mu_X + b \\ \sigma_Y^2 &= a^2 \cdot \sigma_X^2\end{aligned}$$

Remarks

- Proof of result relies on change-of-variables formula for determining pdf of a function of a rv

- Result may or may not hold for random variables whose distributions are not normal

Example - Standardizing a Normal rv

Let $X \sim N(\mu_X, \sigma_X^2)$. The standardized rv Z is created using

$$
\begin{aligned}
Z &= \frac{X - \mu_X}{\sigma_X} = \frac{1}{\sigma_X} \cdot X - \frac{\mu_X}{\sigma_X} \\
&= a \cdot X + b \\
a &= \frac{1}{\sigma_X},\ b = -\frac{\mu_X}{\sigma_X}
\end{aligned}
$$

Properties of Z

$$
\begin{aligned}
E[Z] &= \frac{1}{\sigma_X} E[X] - \frac{\mu_X}{\sigma_X} \\
&= \frac{1}{\sigma_X} \cdot \mu_X - \frac{\mu_X}{\sigma_X} = 0 \\
\mathrm{Var}(Z) &= \left(\frac{1}{\sigma_X}\right)^2 \cdot \mathrm{Var}(X) \\
&= \left(\frac{1}{\sigma_X}\right)^2 \cdot \sigma_X^2 = 1 \\
Z &\sim N(0, 1)
\end{aligned}
$$

1.8 Value at Risk: Introduction

Consider a \$10,000 investment in Microsoft for 1 month. Assume

$$
\begin{aligned}
R &= \text{simple monthly return on Microsoft} \\
R &\sim N(0.05, (0.10)^2),\ \mu_R = 0.05,\ \sigma_R = 0.10
\end{aligned}
$$

Goal: Calculate how much we can lose with a specified probability α

Questions:

1. What is the probability distribution of end of month wealth, $W_1 = \$10,000 \cdot (1 + R)$?

2. What is $\Pr(W_1 < \$9,000)$?

3. What value of R produces $W_1 = \$9,000$?

4. What is the monthly value-at-risk (VaR) on the \$10,000 investment with 5% probability? That is, how much can we lose if $R \leq q_{.05}$?

Answers:

1. $W_1 = \$10,000 \cdot (1 + R)$ is a linear function of R, and R is a normally distributed rv. Therefore, W_1 is normally distributed with

$$
\begin{aligned}
E[W_1] &= \$10,000 \cdot (1 + E[R]) \\
&= \$10,000 \cdot (1 + 0.05) = \$10,500, \\
\mathsf{Var}(W_1) &= (\$10,000)^2 \mathsf{Var}(R) \\
&= (\$10,000)^2 (0.1)^2 = 1,000,000 \\
W_1 &\sim N(\$10,500, (\$1,000)^2)
\end{aligned}
$$

2. Using $W_1 \sim N(\$10,500, (\$1,000)^2)$

$$
\begin{aligned}
&\Pr(W_1 < \$9,000) \\
&= \texttt{NORMDIST(9000,10500,1000)} = 0.067
\end{aligned}
$$

3. To find R that produces $W_1 = \$9,000$ solve

$$R = \frac{\$9,000 - \$10,000}{\$10,000} = -0.10.$$

Notice that -0.10 is the 6.7% quantile of the distribution of R :

$$q_{.067} = \Pr(R < -0.10) = 0.067$$

4. Use $R \sim N(0.05, (0.10)^2)$ and solve for the the 5% quantile:

$$\begin{aligned} \Pr(R < q_{.05}^R) &= 0.05 \Rightarrow \\ q_{.05}^R &= \texttt{NORMINV(0.05, 0.05,0.10)} = -0.114. \end{aligned}$$

If $R = -11.4\%$ the loss in investment value is at least

$$\begin{aligned} \$10,000 \cdot (-0.114) &= -\$1,144 \\ &= 5\% \text{ VaR} \end{aligned}$$

In general, the $\alpha \times 100\%$ Value-at-Risk (VaR_α) for an initial investment of $\$W_0$ is computed as

$$\begin{aligned} \text{VaR}_\alpha &= \$W_0 \times q_\alpha \\ q_\alpha &= \alpha \times 100\% \text{ quantile of simple return distn} \end{aligned}$$

Remark:

Because VaR represents a loss, it is often reported as a positive number. For example, $-\$1,144$ represents a loss of $\$1,144$. So the VaR is reported as $\$1,144$.

VaR for Continuously Compounded Returns

$$r = \ln(1+R), \text{ cc monthly return}$$
$$R = e^r - 1, \text{ simple monthly return}$$

Assume

$$r \sim N(\mu_r, \sigma_r^2)$$
$$W_0 = \text{initial investment}$$

$100 \cdot \alpha\%$ VaR Computation

- Compute α quantile of Normal Distribution for r:

$$q_\alpha^r = \mu_r + \sigma_r z_\alpha$$

- Convert α quantile for r into α quantile for R:

$$q_\alpha^R = e^{q_\alpha^r} - 1$$

- Compute $100 \cdot \alpha\%$ VaR using q_α^R:

$$\mathrm{VaR}_\alpha = \$W_0 \cdot q_\alpha^R$$

Example: Compute 5% VaR assuming

$$r_t \sim N(0.05, (0.10)^2), W_0 = \$10,000$$

The 5% cc return quantile is

$$\begin{aligned} q^r_{.05} &= \mu_r + \sigma_r z_{.05} \\ &= 0.05 + (0.10)(-1.645) = -0.114 \end{aligned}$$

The 5% simple return quantile is

$$q^R_{.05} = e^{q^r_{.05}} - 1 = e^{-.114} - 1 = -0.108$$

The 5% VaR based on a $10,000 initial investment is

$$\text{VaR}_{.05} = \$10,000 \cdot (-0.108) = -\$1,077$$