QF609 Risk Analysis
Lecture Notes 4

Dr. Tony Wong

Adjunct Professor of MQF
Head of Market Risk Model Validation, Maybank

tonywong@smu.edu.sg

Overview

Main topics:

- Recap: Duration Model
- Review of Basic Statistics

Recap: Duration Model

Duration offers an alternative approach for measuring interest rate risks:

- **Macaulay duration** of an asset or liability is the weighted-average time-to-maturity of its cash flows using the present value (PV) of the cash flows as weights:

$$
\begin{aligned}
D &= \frac{\sum_{i=1}^{n} CF_i \cdot DF(T_i) \cdot T_i}{\sum_{i=1}^{n} CF_i \cdot DF(T_i)} = \frac{\sum_{i=1}^{n} PV_i \cdot T_i}{PV} = \sum_{i=1}^{n} w_i \cdot T_i, \\
DF(T_i) &= (1+R)^{-T_i}, \quad w_i = \frac{PV_i}{PV},
\end{aligned}
$$

  where all cash flows $\{CF_i\}$ are assumed to have the same sign.

- **Modified duration** is defined as:

$$
MD = -\frac{\frac{dPV}{dR}}{PV} = \frac{D}{1+R}.
$$

  As a result, $-MD \cdot 0.01$ gives the percentage PV change for $+100$bp shift in the rate R.

- **Dollar duration** is defined as:

$$
DD = -MD \cdot PV
$$

  As a result, $-DD \cdot 0.01$ gives the dollar PV change for $+100$bp shift in the rate R.

Recap: Duration Model

- The overall durations D_A and D_L of assets and liabilities can be calculated from individual asset/liability durations per below:

$$D_A = \omega_A^1 \cdot D_A^1 + \cdots + \omega_A^m \cdot D_A^m, \quad \sum_{i=1}^{m} \omega_A^i = 1,$$

$$D_L = \omega_L^1 \cdot D_L^1 + \cdots + \omega_L^n \cdot D_L^n, \quad \sum_{i=1}^{n} \omega_L^i = 1,$$

  where ω_A^i (resp. ω_L^i) is a weight given by the ratio between the individual asset (resp. liability) PV and the total asset (resp. liability) PV.

- For a given rate change δy, the change in equity value $E = A - L$ is given by:

$$\Delta E = -[D_A - D_L \cdot K] \cdot A \cdot \frac{\Delta y}{1+y}$$

  where $K = \frac{L}{A}$ is the leverage ratio of the bank or FI and $D_A - D_L \cdot K$ is called the **leverage-adjusted duration gap (D-Gap)** .

- To immunize equity value $E = A - L$ against interest rate change, FI can adjust either D_A, D_L, A, or L (or combination of them) to have a zero **leverage-adjusted duration gap (D-Gap)**.

Limitations of Duration-Based Immunization

- Immunizing the entire balance sheet can be time consuming and costly
  - derivative instruments can be used to hedge positions instead of rebalancing.
- Immunization is a dynamic process which involves high transaction fees
  - in practice, only approximately dynamically immunized at discrete intervals
- Large interest rate change effects are not accurately captured
  - higher order sensitivity can be included to improve accuracy
- It cannot be directly applied to transcations with cash flows and/or cash flows timing depending on interest rates.
- The basic version of duration analysis assumes a flat yield curve and parallel rate shift.

Recap: Duration Model

Price approximation using both duration and convexity:

$$\begin{aligned}\frac{\Delta P}{P} &= -MD \cdot \Delta y + \frac{1}{2} \cdot Convexity \cdot (\Delta y)^2 \\ Convexity &= \frac{1}{P} \cdot \frac{d^2 P}{dy^2} = -\frac{dMD}{dy} + MD^2\end{aligned}$$

where MD is the modified duration and the last equality follows from the modified duration formula:

$$MD = -\frac{1}{P} \cdot \frac{dP}{dy}.$$

Random Variables and Probability Distributions

A **Random Variable (RV)** is a variable whose values are stochastic. This means that there is uncertainty about the value that the variable will realize.

Discrete RV

This is a RV which can take on only a countable number of distinct values.

- For every discrete RV X, there is a **probability mass function (PMF)** $p(x)$ that gives the probability that X takes on a particular value x_i from the set of its possible values $E = \{x_1, x_2, \cdots\}$, i.e.:

$$p(x_i) \;\; =_{def} \;\; Probability\,[X = x_i] \;\geq\; 0, \qquad \sum_{x_i \in E} p(x_i) \;=\; 1.$$

  Note that the set E can be either finite or countably infinite (e.g. the set of all positive integers).

- The **cumulative distribution function (CDF)** of X, denoted by $F(x)$, is defined by:

$$F(x) \quad =_{def} \quad Probability\,[X \le x] \;=\; \sum_{\{x_i | x_i \le x\}} p(x_i)$$

- The **m-th moment** of X is defined by:

$$E[X^m] \quad =_{def} \quad \sum_{x_i \in E} p(x_i) \cdot x_i^m$$

The **mean** of X is simply the first moment of X while its **variance** is defined by:

$$V[X] \quad =_{def} \quad E[X^2] - (E[X])^2 \;=\; \sum_{x_i \in E} p(x_i) \cdot x_i^2 - \left(\sum_{x_i \in E} p(x_i) \cdot x_i \right)^2$$

Example:

Let X denote the number that shows up from rolling a fair dice. We have:

- $E = 1, 2, ..., 6$
- The PMF of X is given by $p(1) = p(2) = \cdots = p(6) = \frac{1}{6}$
- The CDF of X a step function:

$$F(x) = \begin{cases} 0 & x < 1 \\ 1/6 & x < 2 \\ 1/3 & x < 3 \\ 1/2 & x < 4 \\ 2/3 & x < 5 \\ 5/6 & x < 6 \\ 1 & x \geq 6 \end{cases}$$

- $E[X] = 3.5$ and $V[X] = \frac{91}{6} - 3.5^2 = 2.9167$

Some Well-known Discrete Distributions

- **Binomial(n,p)**: X denotes the number of successes in n indepedent trials with a probablity of success p in each trial. In the particular case $n = 1$, the RV is also known as a **Bernoulli(p) RV**. In this case, we have:

$$p(k) \;=\; C_k^n \cdot p^k \cdot (1-p)^{n-k}, \quad C_k^n \;=\; \frac{n!}{k! \cdot (n-k)!}$$

for $k \in E \;=\; \{0, 1, \cdots, n\}$. The mean and variance are respectively given by:

$$E[X] \;=\; n \cdot p, \quad V[X] \;=\; n \cdot p \cdot (1-p)$$

- **Geometric(p)**: X denotes the number of independent trials until the 1st success where each trial has a probablity of success p.

$$p(k) \;=\; (1-p)^{k-1} \cdot p$$

for $k \in E \;=\; \{0, 1, \cdots\}$. Note that in this case E is the set of all positive integers. The mean and variance are respectively given by:

$$E[X] \;=\; \frac{1}{p}, \quad V[X] \;=\; \frac{1-p}{p^2}$$

- **NegativeBinomial(n,p)**: Let X denote number of independent trials until the n-th success where each trial has a probablity of success p.

$$p(k) \;=\; C_{k-1}^{n-1} \cdot p^{k-1} \cdot (1-p)^{n-k+1} \cdot p$$

for $k \in E \;=\; \{0, 1, \cdots\}$. Note that in this case E is the set of all positive integers. The mean and variance are respectively given by:

$$E[X] \;=\; \frac{n}{p}, \quad V[X] \;=\; \frac{n(1-p)}{p^2}$$

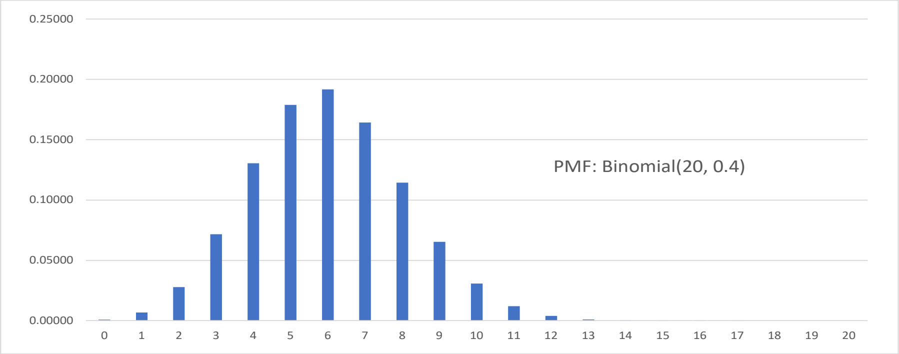
- **Poisson(λ)**: This is used to describe the probability distribution of the number of independent occurrences of an event over a given time period where λ can be regarded as the number of event arrivals per unit of time (e.g. per year). Let X_T denote the number of occurrences over the time interval $[0, T]$. Under this distribution, the PMF of X_T is given by

$$p(k; \lambda, T) \;=\; \frac{(\lambda \cdot T)^k e^{-\lambda \cdot T}}{k!}$$

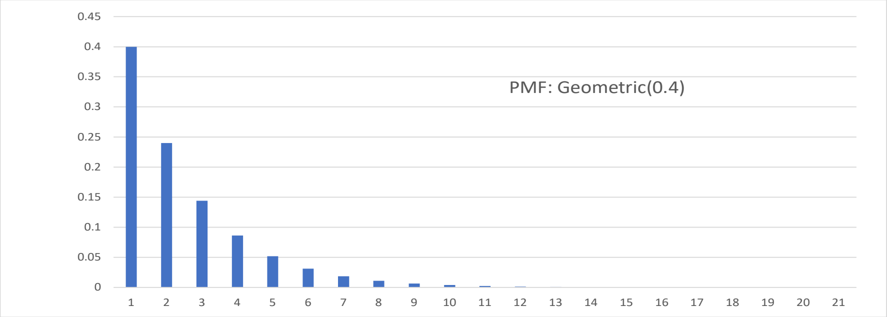
for $k \in E \;=\; \{0, 1, \cdots\}$. The mean and variance are respectively given by:

$$E[X_T] \;=\; \lambda \cdot T, \quad V[X_T] \;=\; \lambda \cdot T$$

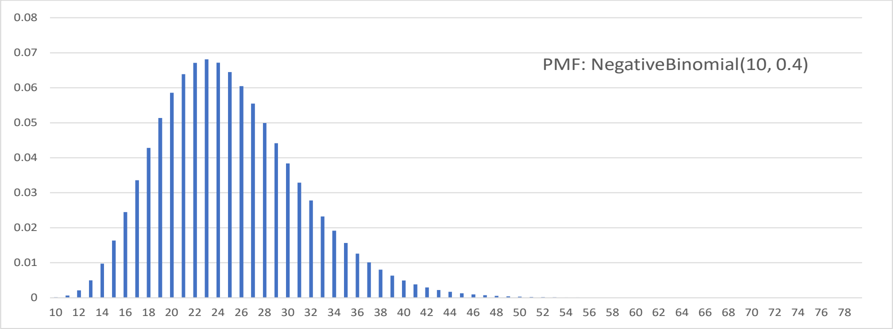
Review of Basic Statistics



Review of Basic Statistics



Review of Basic Statistics



Example:

a. Consider an experiment of rolling a fair dice indefinitely.
    - Let X denote the number of rolls under one obtains a number greater than 2 for the first time. What is the distribution, mean, and variance of X?
    - Let X denote the number of rolls under one obtains a number greater than 2 for the 5th time. What is the distribution, mean, and variance of X?

b. A final exam has 10 MCQs, each with 5 possible answers and only one is correct. A student plans to guess the answer to each question randomly. Let X denote the number of questions he answers correctly.
    - What are the expectation and variance of X?
    - What is the probability that all of his answers are wrong?
    - What is the probability that he has two questions answered correctly?

Continuous RV

For most practical cases, it suffices to consider a continuous RV X being one that

- takes on values from an interval rather than a countable set
- assigns a positive probability to intervals of values only via a **non-negative probability density function (PDF)** $f(x)$ that satisfies:

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)dx &= 1 \\ Probability\,[X \in (l,u)] &= \int_{l}^{u} f(x)dx \;\; for \;\; any \;\; interval \;\; [l,u]\end{aligned}$$

  As a result, it is clear that

$$Probability\,[X = x] = 0 \;\; for \;\; all \;\; x.$$

- The **cumulative distribution function (CDF)** $F(x)$ is given by:

$$F(x) \;\; =_{def} \;\; Probability\,[X \leq x] \; = \int_{-\infty}^{x} f(u)du$$

Review of Basic Statistic

- As a result, if we know the CDF $F(x)$, we can obtain the PDF via

$$f(x) = \frac{dF(x)}{dx}$$

- The **m-th moment** of X is given by:

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

The **mean** of X is simply the first moment of X while its **variance** is given by:

$$V[X] =_{def} E[X^2] - (E[X])^2 = \int_{-\infty}^{\infty} x^2 f(x)dx - \left(\int_{-\infty}^{\infty} xf(x)dx\right)^2$$

Example:

Let X denote a continuous RV with its PDF given by:

$$f(x) = \begin{cases} \kappa \cdot x^3 & 0 \leq x \leq 1 \\ 0 & otherwise \end{cases}$$

where κ is some positive constant.

- How do we know the value of κ? This can be determined using the fact that $f(x)$ needs to integrate to 1 over the real line, i.e.:

$$\int_{-\infty}^{\infty} \kappa \cdot x^3 dx = \int_0^1 \kappa \cdot x^3 dx = 1$$

This gives $\kappa = 4$.

- The CDF of X is given by:

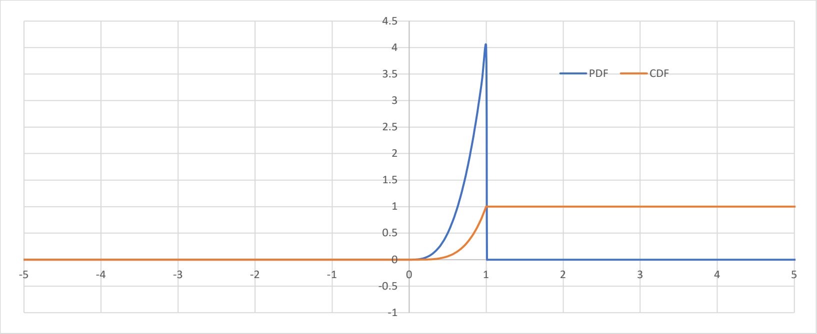
$$F(x) = \int_{-\infty}^{x} f(u)du = \begin{cases} 0 & x < 0 \\ \int_0^x 4u^3 du = x^4 & 0 \leq x \leq 1 \\ 1 & x > 1 \end{cases}$$

- The mean and variance of X are given by:

$$E[X] = \int_0^1 u \cdot 4u^3 du = \frac{4}{5}$$

$$V[X] = \int_0^1 u^2 \cdot 4u^3 du - \left(\frac{4}{5}\right)^2 = \frac{2}{3} - \frac{16}{25} = 0.0267$$

- The PDF and CDF of X is plotted below:



Some Well-known Continuous Distributions

- **Uniform(a,b)**: a Uniform(a,b) distribution is defined by the following PDF

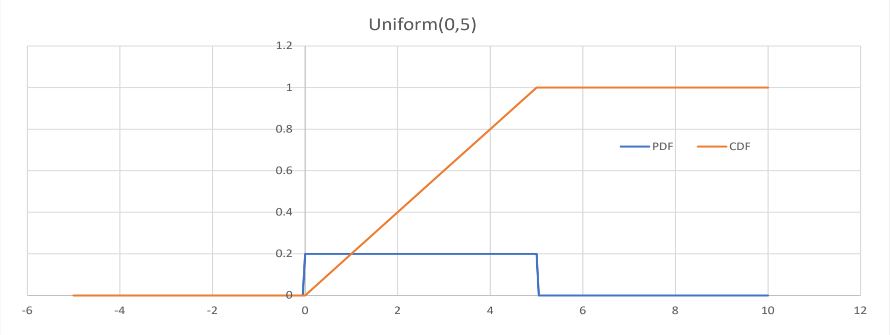
$$f(x) = \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & otherwise \end{cases}$$

In this case, the CDF, mean, variance are given by:

$$F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & x \in [a, b] \\ 1 & x > b \end{cases}$$

$$E[X] = \frac{a+b}{2}, \quad V[X] = \frac{(b-a)^2}{12}$$

Among the class of uniform distribution, Uniform(0,1) is of particular importance as it is used for simulating all kinds of distributions (we shall touch this topic later). Below is a plot of the PDF and CDF of a sample uniform distribution.


Uniform(0,5)
1.2
1
0.8
0.6
0.4
0.2
0
-6
-4
-2
0
2
4
6
8
10
12
PDF
CDF

- **Exp(λ)**: an Exp(λ) distribution is defined by the following PDF:

$$f(x) = \begin{cases} xe^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

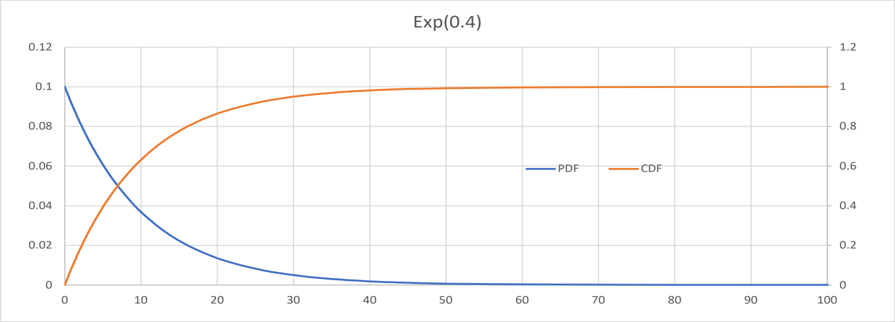
In this case, the CDF, mean, variance are given by:

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x} & x \geq 0 \end{cases}$$

$$E[X] = \frac{1}{\lambda}, \quad V[X] = \frac{1}{\lambda^2}$$

The Exp(λ) RV can be seen as the time until the first event occurrence within the context of a Poisson(λ) variable (**Hint**: *consider the probability of having zero event occurrence within the time interval [0,T] for a Poisson RV*)

Review of Basic Statistics



- **Standard Normal RV**: a standard normal distribution is defined by the following PDF:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in (-\infty, \infty)$$

In this case, the CDF, mean, and variance of X are given by:

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du = \frac{1}{2}\left[1 + \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2} du\right] \\ E[X] &= 0, \ \ V[X] = 1. \end{aligned}$$

While there is no closed-form expression available for analytical evaluation of $F(x)$, the integral $\int_0^x e^{-u^2} du$ can be numerically computed with a high accuracy. There is a standized table for looking up values of $F(x)$ at standard grids.

Challenge 1: Can you show the second equality below?

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du = \frac{1}{2}\left[1 + \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2} du\right]$$

Challenge 2: Since a PDF should be integrated to 1 over the the entire domain, we should have

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{u^2}{2}} du = 1$$

Can you show the above is indeed integrated to 1? **Hint**: all you need is to work out the value of $\int_{-\infty}^{\infty} e^{-u^2} du$. This is a populuar interview question.

- **Normal(μ, σ) RV**: a Normal(μ, σ) RV X can be defined via the following mapping:

$$X = \mu + \sigma \cdot Z,$$

where Z is standard normal RV which we've discussed already. Let $F(x)$ and $\Phi(z)$ denote the CDF of X and Z, respectively. We have

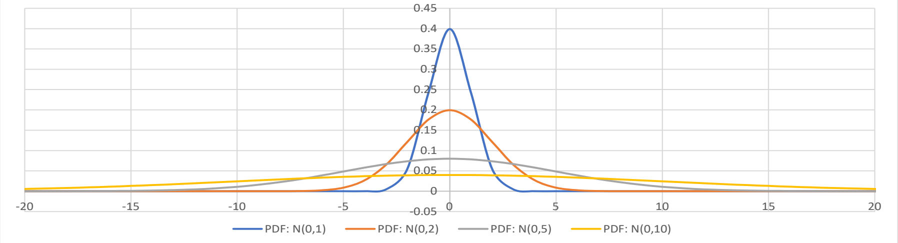
$$F(x) = Probability(X \leq x) = Probability\left(Z \leq \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right).$$

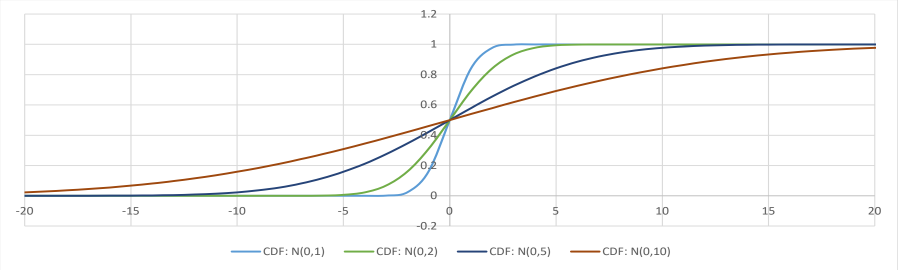
Hence, the CDF of X can be effectively calcualted from that of a standard normal distribution. How about the PDF? Let $f(x)$ and $\phi(z)$ denote the PDF of X and Z, respectively. We have:

$$\begin{aligned} f(x) &= \frac{dF(x)}{dx} = \frac{d\left[\Phi\left(\frac{x-\mu}{\sigma}\right)\right]}{dx} = \phi\left(\frac{x-\mu}{\sigma}\right) \cdot \frac{1}{\sigma} \\ &= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\cdot\left(\frac{x-\mu}{\sigma}\right)^2} \end{aligned}$$

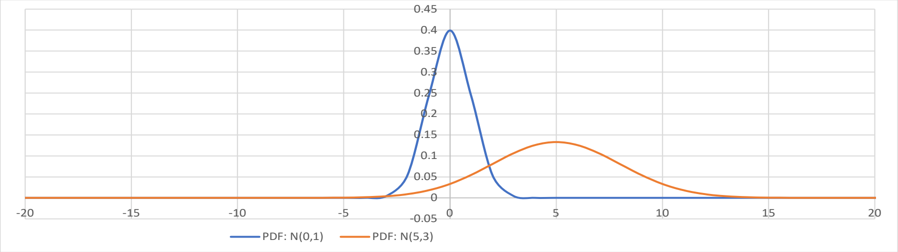
Clearly, the mean and variance of X are μ and σ^2, respectively.

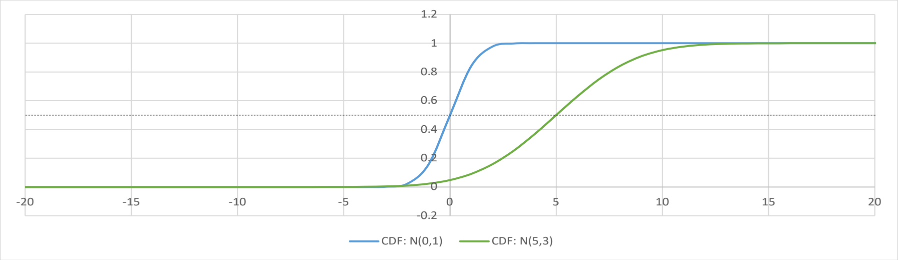
Review of Basic Statistics





Review of Basic Statistics





From the above, it is clear that the general normal(μ,σ) distribution is simply the result of: (1) shifting the mean of the standard normal from 0 to μ, and (2) rescaling the 'dispersion' of the distribution, as measured by the standard deviation. A more dispersed distribution means the underlying variable has a higher chance of varying over a wider range, hence in a sense is more 'risky'.

An Important Result:

Let X be a continuous random variable whereby its CDF $F(x)$ is strictly increasing. Then,

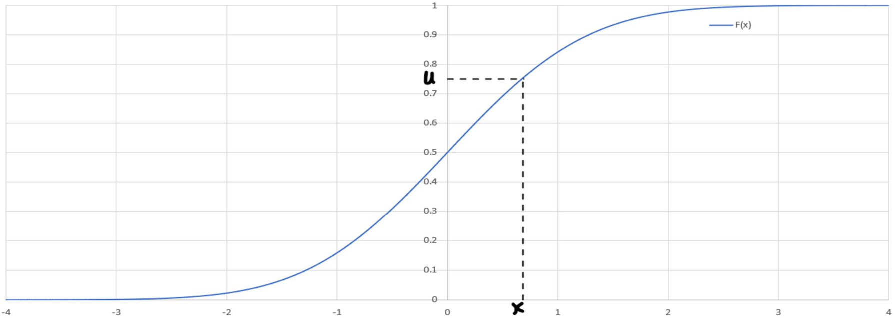
a. The RV $Y = F(X)$ has a Uniform(0,1) distribution
b. The RV $Z = F^{-1}(U)$ has the same distribution as X, where U is a Uniform(0,1) RV.

To prove a), we can work out the CDF/PDF of Y, $F_Y(u)$ and $f_Y(u)$ as follows:

$$
\begin{aligned}
F_Y(u) &= Probability(Y \leq u) = Probability(F(X) \leq u) \\
&= Probability(X \leq F^{-1}(u)) = F(F^{-1}(u)) = u \\
f_Y(u) &= \frac{dF_Y(u)}{du} = 1.
\end{aligned}
$$

You can apply a similar trick to prove b). Why is this result important?

Illustration to help previous question:



Example:

A RV X has a normal distribution with mean μ and standard deviation σ.

- Show that $Z = \frac{X-\mu}{\sigma}$ has a standard normal distribution assuming. that all we know is the *PDF* of X presented on previous slide.
- Assume $\mu = 5$, $\sigma = 8$, and a is some number such that $a \geq 5$. Show that

$$\mathit{Probability}(X > a) \quad = \quad \mathit{Probability}(X < 10 - a).$$

- Assume again $\mu = 5$ and $\sigma = 8$. What is $\mathit{Probability}(X^2 \geq 12)$? You can express your answer in terms of the standard normal CDF $\Phi(\cdot)$.

Multiple RVs

- Given two continuous RVs X_1 and X_2, their covariance and correlation is defined as:

$$\begin{aligned} COV(X_1, X_2) &= E[X_1 \cdot X_2] - E[X_1] \cdot E[X_2], \\ E[X_1 \cdot X_2] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2) x_1 x_2 dx_1 dx_2 \\ CORR(X_1, X_2) &= \frac{COV(X_1, X_2)}{\sqrt{V[X_1]} \cdot \sqrt{V[X_x]}} \end{aligned}$$

where $f(x_1, x_2)$ is the so-called joint PDF of $X1$ and X_2. We have:

$$V[X] = COV(X, X), \quad COV(X_1, X_2) = COV(X_2, X_1), \quad CORR(X_1, X_2) = CORR(X_2, X_1)$$

It can be proved that $-1 \leq CORR(X_1, X_2) \leq 1$. A more positive correlation generally means that the two RV vary more strongly together (in the same direction) while a more negative correlation generally means the opposive. A correlation of zero, while not necessary, generally means the two RVs are indepedent.

- For multiple RVs, the covariances and correlations between for any two of them can be summarized in the form of a variance-covariance matrix and a correlation matrix.
- Consider the RVs $X_1, \cdots, X_n$. Their variance-covariance matrix is given by:

$$
\begin{aligned}
\Gamma &= \begin{pmatrix} V(X_1) & COV(X_1,X_2) & COV(X_1,X_3) & \cdots & COV(X_1,X_n) \\ COV(X_2,X_1) & V(X_2) & COV(X_2,X_3) & \cdots & COV(X_2,X_n) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ COV(X_n,X_1) & COV(X_n,X_2) & \cdots & COV(X_n,X_{n-1}) & V(X_n) \end{pmatrix} \\
\Psi &= \begin{pmatrix} 1 & CORR(X_1,X_2) & CORR(X_1,X_3) & \cdots & CORR(X_1,X_n) \\ CORR(X_2,X_1) & 1 & CORR(X_2,X_3) & \cdots & CORR(X_2,X_n) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ CORR(X_n,X_1) & CORR(X_n,X_2) & \cdots & CORR(X_n,X_{n-1}) & 1 \end{pmatrix}
\end{aligned}
$$

- Clearly, the two matrices are symmetric.

Review of Basic Statistics

Below are some well-known and userful results on expectation, variance, and covariance:

- $E[X + C] = E[X] + C$
- $E[C \cdot X] = C \cdot E[X]$
- $E[a_1 \cdot X_1 + \cdots + a_n \cdot X_n] = a_1 \cdot E[X_1] + \cdots + a_N \cdot E[X_n]$
- $V[X + C] = V[X]$
- $V[C \cdot X] = C^2 \cdot V[X]$
- $V[a_1 \cdot X_1 + \cdots + a_n \cdot X_n] = \sum_{i=1}^{n} a_i^2 \cdot V[X_i] + \sum_{i \neq j} 2a_i a_j COV(X_i, X_j)$
- $\int_{-\infty}^{\infty} f(x_1, x_2) dx_2 = f_{X_1}(x_1)$ and $\int_{-\infty}^{x} \int_{-\infty}^{\infty} f(x_1, x_2) dx_2 = F_{X_1}(x)$

Let $X_1, \cdots, X_n$ be n RVs with variance-covariance matrix Γ and let $\vec{a} = (a_1, \cdots, a_n)$ be a row vector of n elements.

$$V[\sum_{i=1}^{n} a_i \cdot X_i] \;=\; \sum_{i=1}^{n} a_i^2 \cdot V[X_i] + \sum_{i \neq j} 2a_i a_j COV(X_i, X_j) \;=\; \vec{a} \cdot \Gamma \cdot \vec{a}^T$$

We know that $V[\sum_{i=1}^{n} a_i \cdot X_i] \geq 0$ for any choice of $\vec{a}$. What this means is that a variance-covariance matrix must be positive semi-definite.
How about correlation matrix?

Law of Large Number (LLN)

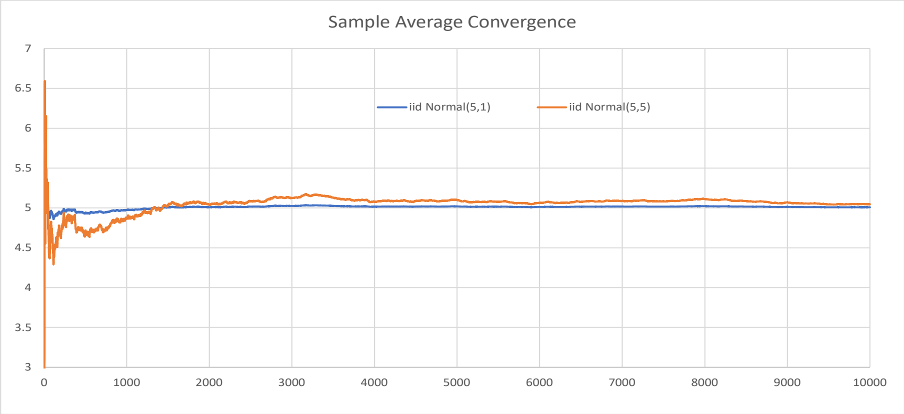
Let $\{X_1, \ldots, X_n\}$ be a sequence of indendepently identically distributed (iid) RVs with mean μ and finite variance σ^2. Then, their average

$$\bar{X}_n = \frac{X_1 + \cdots + X_n}{n}$$

converges (almost surely) to μ. The plot on the next slide illustrates the convergence under LLN.

Central Limit Theorem (CLT) Let $\{X_1, \ldots, X_n\}$ be a sequence of indendepently identically distributed (iid) RVs with mean μ and finite variance σ^2. Then, $\sqrt{n}(\bar{X}_n - \mu)$ coverge in distribution to $Normal(0, \sigma^2)$.

Review of Basic Statistics



Sample Statistics

Following LLN, given a sample of IID observations $x_1, \cdots, x_n$ from a distribution with mean μ and variance σ^2, one can have sample estiamtes of the mean and variance of the distribution:

- Sample Mean:

$$\hat{\mu} = \frac{x_1 + \cdots + x_n}{n}$$

  This is an unbiased estimate of the true mean: $E[\hat{\mu}] = \mu$.

- Sample Variance:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} (x_i - \hat{\mu})^2}{n}$$

  We have:

Review of Basic Statistics

$$
\begin{aligned}
E[\hat{\sigma}^2] &= \frac{\sum_{i=1}^n (x_i - \hat{\mu})^2}{n} = \frac{1}{n}\sum_{i=1}^n E\left[x_i^2\right] - \hat{\mu}^2 \\
&= \frac{1}{n}\sum_{i=1}^n E\left[x_i^2\right] - \frac{1}{n^2} E\left[(x_1 + \cdots + x_n)^2\right] \\
&= \frac{1}{n}\sum_{i=1}^n (\sigma^2 + \mu^2) - \frac{1}{n^2}\sum_{i=1}^n E\left[x_i^2\right] - \frac{1}{n^2}\sum_{i \neq j} E\left[x_i \cdot x_j\right] \\
&= (\sigma^2 + \mu^2) - \frac{\sigma^2 + \mu^2}{n} - \frac{2 \cdot C_2^n \cdot \mu^2}{n^2} \\
&= \frac{n-1}{n}\sigma^2
\end{aligned}
$$

Hence, $\hat{\sigma}^2$ is an biased estimate of the true variance σ^2. But we can turn it into an unbiased estimate with a slight modification of its definition

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (x_i - \hat{\mu})^2}{n-1}$$

Do you see why the above is unbiased?

- Similarly, the sample skewness and kurtosis can be calculated:

$$\hat{\gamma} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{x_i - \hat{\mu}}{\hat{\sigma}}\right)^3, \quad \hat{\delta} \ = \ \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{x_i - \hat{\mu}}{\hat{\sigma}}\right)^4$$

Given iid observations $\{(x_i, y_i)\}_i = 1^n$ from two distributions (X and Y) with means μ_x and μ_y, and variances σ_x^2 and σ_y^2, respectively.

- Sample Covariance:

$$\widehat{COV} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \hat{\mu}_x)\cdot(y_i - \hat{\mu}_y)$$

- Sample Correlation:

$$\widehat{Corr} = \frac{\widehat{COV}}{\hat{\sigma}_x \cdot \hat{\sigma}_y}$$

In general, $\widehat{Corr}$ is a biased estiamte of the true correlation.

Monte Carlo (MC) Simulation

- Recall that simulation of any continuous RV X with a known CDF $F(\cdots)$ can be effectively reduced to simulating a $Uniform[0, 1]$ RV U:
  - $F^{-1}(U)$ follows the distribution of X
- In general, MC simulation tries to address the following problem:
  - Given some function $f(X_1, \cdots, X_n)$ of n RVs $X_1, \cdots, X_n$, we want to estimate the expected value $E[f(X_1, \cdots, X_n)]$ whereby the distributions of $f(\cdot)$ and/or the RVs random variables $X_1, \cdots, X_n$ may not be explicitly known.
- While distributions of $X_1, \cdots, X_n$ may not be explicitly known, suppose that we have some way to simulate independent samples of $X_1, \cdots, X_n$. Then, we know by LLN:

$$E[f(X_1, \cdots, X_n)] = \lim_{m\to\infty} \frac{\sum_{i=1}^m f(X_1^i, \cdots, X_n^i)}{m}$$

where for each i, $X_1^i, \cdots, X_n^i$ denote an independent sample of the RVs.

- In most practical applications, the problem typically boils down to simulating (potentially correlated) standard normal RVs, e.g. you may want to simulate independent samples of n correlated standard normal RVs with a known correlation matrix A. For this, the typical procedure is:
  - identify a decomposition of the correlation matrix A:

$$A = \Gamma \cdot \Gamma^T,$$

    where Γ^T denotes the transpose of the *factor loading matrix* Γ.
  - provided a factor loading matrix Γ is obtained, the problem reduced to simulating samples of n independent standard normal RVs which is straightforward.
  - There are different matrix decomposition methods to obtain a factor loading matrix Γ, i.e. Eigen decomposition, Cholesky decomposition etc..

Exercises (popular interview questions)

- How do you simulate two correlated standard normal RVs with a correlation ρ?
- Can you think of a way to estimate the value of π using Monte Carlo simulation? (*Hint: use the fact the area of circle is equal to* πr^2)