R04 Common Probability Distributions

Version 1.0

1. Introduction and Discrete Random Variables

In nearly all investment decisions, we work with random variables. To make probability statements about a random variable, we need to understand its probability distribution. A probability distribution specifies the probabilities of all possible outcomes of a random variable. In this reading, we will look at the following seven probability distributions:

- Uniform
- Binomial
- Normal
- Lognormal
- Student's t
- Chi-square
- F-distribution

1.1 Discrete Random Variables

A random variable is a variable whose outcome cannot be predicted. There are two basic types of random variables:

- A **discrete random variable** is one where we can list all the possible outcomes. For example, the stocks traded on the New York Stock Exchange are quoted in ticks of $0.01. Quoted stock price is therefore a discrete random variable.
- A **continuous random variable** is where the number of points between the lower and upper bounds are infinite. Rate of return is an example of a continuous random variable.

Probability distribution specifies the probabilities of all the possible outcomes for a random variable.

- Probability function specifies the probability that a random variable takes on a specific value. It is expressed as: $p(x) = P(X = x)$. The capital X represents the random variable and lowercase x represents a specific value that the random variable may take.
- Probability density function denoted by $f(x)$ is used for continuous random variables.

A **cumulative distribution function** defines the probability that a random variable, X, takes on a value equal to or less than a specific value, x. It is denoted as: $F(x) = P(X \leq x)$. It represents the sum of the probabilities for the outcomes up to and including a specified outcome.

2. Discrete and Continuous Uniform Distribution

The Discrete Uniform Distribution

The discrete uniform distribution has a finite number of specified outcomes and each outcome is equally likely. Consider a roll of a dice. The outcome is a random variable and it can take a value of 1 to 6. The probability that the random variable takes on any of these values is the same for all outcomes. With six outcomes, p(x) = 1/6 for all values of X (X = 1, 2, 3, 4, 5, 6). The table below summarizes the two views of this random variable – the probability function and the cumulative distribution function.

X = x	**Probability Function p(x) = P(X = x)**	**Cumulative Distribution Function F(x) = P(X ≤ x)**
1	1/6	1/6
2	1/6	2/6
3	1/6	3/6
4	1/6	4/6
5	1/6	5/6
6	1/6	6/6

Example

Using the table above, find the following probabilities:

1. F(4)
2. P(3 ≤ X < 6)
3. F(9)

Solution to 1:

To find F(4), we must find the cumulative probability of P(X ≤ 4) using the cumulative distribution function (third column). From the table, we can see that P(X ≤ 4) = 4/6 = 2/3. Therefore, the probability is 2/3.

Solution to 2:

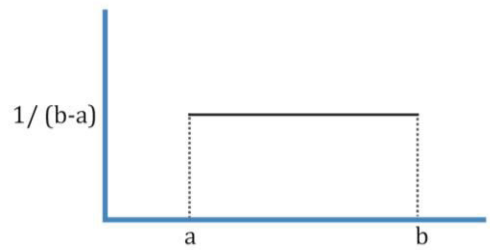
To find P (3 ≤ X < 6), we need to find the sum of three probabilities: p(3), p(4), p(5). This is 1/6 + 1/6 + 1/6 = 3/6 = 1/2.

Solution to 3:

To find F(9), we must find the cumulative probability of P(X ≤ 9). This includes all possible outcomes; hence the probability is 1.

2.1. Continuous Uniform Distribution

The continuous uniform distribution is defined over a range from a lower limit 'a' to an upper limit 'b'. These limits serve as the parameters of the distribution.



The probability that the random variable will take a value between x_1 and x_2, where x_1 and x_2 both lie within the range is given by:

$$P(x_1 \leq X \leq x_2) = \frac{x_2 - x_1}{b - a}$$

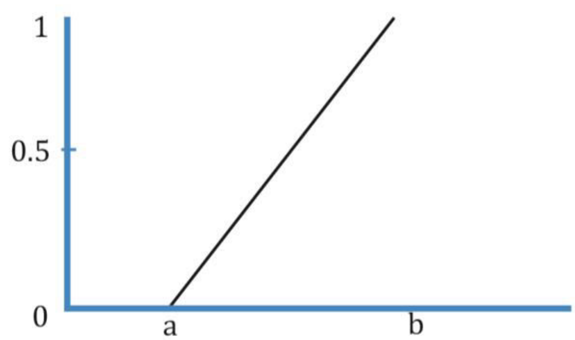
Example

X is a uniformly distributed continuous random variable between 10 and 20. Calculate the probability that X will fall between 12 and 18.

Solution:

$$P(12 \leq X \leq 18) = \frac{18 - 12}{20 - 10} = 0.6$$

The cumulative distribution function for a continuous random variable is shown below:



Example

A commodity analyst predicts that the price per ounce of gold three years from now will be between \$1,500 and \$1,700. Assume gold prices follow a continuous uniform distribution. What is the probability that the price will be less than \$1,600 three years from now?

Solution:

$F(1{,}600) = \frac{1{,}600 - 1{,}500}{1{,}700 - 1{,}500} = 50\%$ The probability that gold price will be less than \$1,600 per ounce three years from now is 50%.

3. The Binomial Distribution

The building block of the binomial distribution is the **Bernoulli random variable**. A

Bernoulli trial is one where there are only two possible outcomes: success or failure. Flipping a coin is an example of a Bernoulli trial – you either get heads or tails, but nothing else. This can be expressed as:

$P(Y = 1) = p$
$P(Y = 0) = 1 - p$

where:
p = probability that the trial is a success

In a **binomial distribution**, the random variable, X, is the number of successes in a given number of Bernoulli trials. Continuing with the coin example, say we flip the coin 10 times and we define success as 'Heads'. Clearly with 10 flips we can get 0 to 10 successes.

The probability distribution of a binomial random variable for the probability of "x" success in "n" trials is calculated using the following formula:

$$P(x) = P(X = x) = ({}_{n}C_{x})\, p^{x}\, (1 - p)^{n-x}$$

where:
p = the probability of success on each trial
x = number of successes
n = number of trials

Instructor's Note:

Two important points help illustrate the intuition behind the formula:

- The successes can be in any order. That is why we use the combination function and not the permutation function.
- The events are independent. That is why we simply multiply the probability of each event.

Example
If we flip a fair coin (p = 0.5) ten times (n = 10), what is the probability of seven successes?

Solution:
$P(7) = P(X = 7) = {}_{10}C_{7}\, 0.5^{7}\, 0.5^{3} = 0.117$

Mean and variance of a binomial variable

The mean and variance of a binomial variable can be calculated as:

Random Variable	**Mean**	**Variance**
Bernoulli, B (1, p)	p	p (1 - p)
Binomial, B (n, p)	np	np (1 - p)

For our coin-flip example, the mean value of the binomial random variable is np = 10 x 0.5 = 5. The intuition: if we perform the binomial experiment several times, where each experiment refers to 10 coin-flips, on average we will have 5 successes. The actual number of successes will be distributed equally on either side of the mean value. The random value for every trial moves closer to the expected value as the number of trials grows.

Example

Over the last 10 years, Abro corporation's EPS increased year over year six times and decreased year over year four times. You decide to model the number of EPS increases for the next decade as a binomial random variable.

1. If success is defined as an increase in the annual EPS, determine the probability of success.
2. What is the probability that EPS will increase in exactly 5 of the next 10 years?
3. Calculate the expected number of yearly EPS increases during the next 10 years.
4. Calculate the variance and standard deviation of the number of yearly EPS increases during the next 10 years.

Solution to 1:
There are only two possible outcomes: increase in the EPS and no increase in the EPS.
Probability of success: p = 6/10 = 0.6
Probability of failure: 1 – p = 1 – 0.6 = 0.4

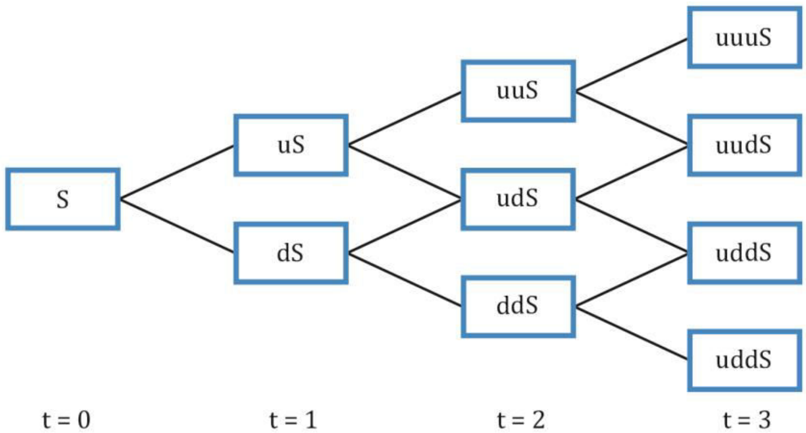
Solution to 2:
Using the binomial model:
$P(X = 5) = ({}_{n}C_{x})\ p^{x}\ q^{n-x}$
$P\ (X = 5) = ({}_{10}C_{5})\ 0.6^{5}\ 0.4^{5}$
$P\ (X = 5) = 252 \times 0.07776 \times 0.01024 = 20.06\%$

Solution to 3:
Expected number of yearly EPS increases: $E(x) = np = 10 \times 0.6 = 6$

Solution to 4:
Variance = $np\ (1 - p) = 6 \times 0.4 = 2.4$
Standard Deviation = $\sqrt{2.4}$ = 1.549
The variance of the distribution is calculated as n p (1 – p) = 10 x 0.5 x (1 – 0.5) = 2.5

Binomial tree

A binomial tree can be used to model stock price movements. Refer to the tree diagram below. 'S' represents the initial stock price. 'u' represents an up move and 'd' represents a down move. The nodes show each possible value of the stock after 1, 2 and 3 time periods.



The expected stock price after each period is equal to the sum of possible stock prices at the end of the period multiplied by their respective probabilities.

Example

Consider an initial stock price of \$100. In one time period, the stock can either rise by a factor of 1.1 or go down by a factor of 1/1.1. In any given time period, the probability of an up move is 0.6 and the probability of a down move is 0.4. After two periods, what are the possible stock prices and their respective probabilities? What is the expected stock price?

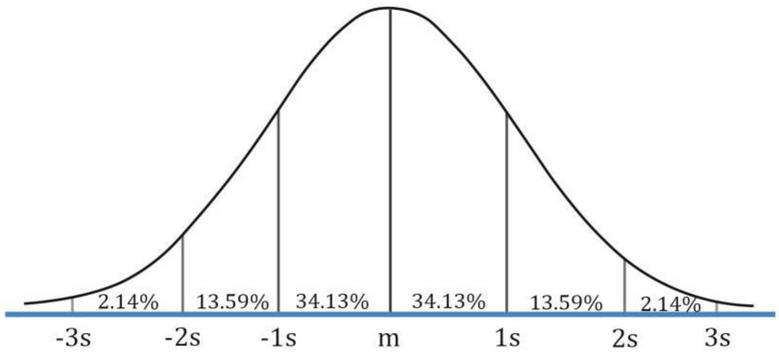
Solution:

uuS = 1.1 x 1.1 x 100 = 121 with probability 0.6 x 0.6 = 0.36
udS = 1.1 x 1/1.1 x 100 = 100 with probability 0.6 x 0.4 = 0.24
duS = 1/1.1 x 1.1 x 100 = 100 with probability 0.4 x 0.6 = 0.24
ddS = 1/1.1 x 1/1.1 x 100 = 82.64 with probability 0.4 x 0.4 = 0.16
Expected stock price = 121 x 0.36 + 100 x 0.24 + 100 x 0.24 + 82.64 x 0.16 = \$104.78

ounce three years from now is 50%.

4. Normal Distribution

The normal distribution is the most extensively used probability distribution in quantitative work. A normal distribution is symmetrical and bell-shaped as shown in the graph below:



In this figure 'm' stands for mean, 1s means one standard deviation, 2s means two standard deviations, and so on. We can make the following probability statements for a normal distribution:

- Approximately 68% of all observations fall in the interval m ± 1s.
- Approximately 95% of all observations fall in the interval m ± 2s.
- Approximately 99% of all observations fall in the interval m ± 3s.

The intervals indicated above are easy to remember but are only approximate for the stated probabilities. More precise intervals (confidence intervals) are:

- 90% of all observations are in the interval m ± 1.65s.
- 95% of all observations are in the interval m ± 1.96s.
- 99% of all observations are in the interval m ± 2.58s.

The characteristics of a normal distribution are as follows:

- The normal distribution is completely described by two parameters – its mean, μ, and variance, σ^2. We indicate this as $X \sim N(\mu, \sigma^2)$.
- The normal distribution has a skewness of 0 (it is symmetric) and a kurtosis (measure of peakedness) of 3. Due to the symmetry, the mean, median and mode are all equal for a normal random variable.
- A linear combination of two or more normal random variables is also normally distributed.

Standard normal distribution

The normal distribution with mean (μ) = 0 and standard deviation (σ) = 1 is called the standard normal distribution.

The formula for standardizing a random variable X is:

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

where:
μ is the population mean.

σ is the population standard deviation.

The Z-table is used to find the probability that X will be less than or equal to a given value. Suppose we have a normal random variable, X, with μ = 10 and σ = 2. If the value of X is 11, we standardize X with Z = (11 – 10)/2 = 0.5.

The probability that we will observe a value less than 11 for X ~ N (10, 2) is exactly the same as the probability that we will observe a value less than 0.5 for Z ~ N (0, 1).

We can answer probability questions about X using standardized values and probability tables for Z. A snapshot of a table showing cumulative probabilities for a standard normal distribution is shown below.

x or *z*	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.00	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.10	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.20	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.30	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.40	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.50	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224

To find the probability that a standard normal variable is less than or equal to 0.5, for example, locate the row that contains 0.50, look at the 0 column, and find the entry 0.6915. Thus, P(Z ≤ 0.5) = 0.6915 or 69.15%.

The probability that a standard normal variable is less than or equal to 0 is 0.5000. This is true by definition because the mean of a standard normal distribution is 0. The table above validates this fact.

For a non-negative number x, we can use N(x) directly from the table. For a negative number –x, N(-x) = 1.0 – N(x). Essentially, we are using the fact that the normal distribution is symmetric around the mean.

Example

A portfolio has a mean return of 15% and a standard deviation of return of 20% per year. What is the probability that the portfolio return would be below 18%? We are given the following information from the z-table: P(Z < 0.15) = 0.5596, P(Z > 0.15) = 0.4404, P(Z < 0.18) = 0.5714, P(Z > 0.18) = 0.4286.

Solution:

$$\mathrm{P}\left(\mathrm{Z} < \frac{\mathrm{X} - \mu}{\sigma}\right) = \mathrm{P}\left(\mathrm{Z} < \frac{0.18 - 0.15}{0.20}\right) = \mathrm{P}(\mathrm{Z} < 0.15) = 0.5596$$

Univariate v/s multivariate distribution

Univariate distribution

A univariate distribution describes the probability distribution of a single random variable. For example, the distribution of expected return of one stock from a portfolio.

Multivariate distribution

A multivariate distribution describes the probability distribution for a group of related random variables. For example, the distribution of expected return of a portfolio with multiple stocks.

A multivariate normal distribution for the returns on n stocks is completely defined by the following three sets of parameters:

- Mean returns on the individual stocks (n means in total).
- Variances of the individual stocks (n variances in total).
- Pairwise return correlations between the stocks (n (n - 1)/2 distinct correlations in total).

5. Applications of the Normal Distribution

Shortfall risk

Shortfall risk is the risk that portfolio's return will fall below a specified minimum level of return over a given period of time.

Safety first ratio

Safety first ratio is used to measure shortfall risk. It is calculated as:

$$\text{SF Ratio} = \frac{[R_P - R_L]}{\sigma_P}$$

where:
R_p = Expected portfolio return
R_L = Threshold level
σp = Standard deviation of portfolio returns

The portfolio with the highest SF-Ratio is preferred, as it has the lowest probability of falling below the target return.

Example

An investor is considering two portfolios A and B. Portfolio A has an expected return of 10% and a standard deviation of 2%. Portfolio B has an expected return of 15% and a standard deviation of 10%. The minimum acceptable return for the investor is 8%. According to Roy's safety first criteria, which portfolio should the investor select?

Solution:

$$SF_A = \frac{10 - 8}{2} = 1$$

$$SF_B = \frac{15 - 8}{10} = 0.7$$

Since A has a higher safety first ratio, the investor should select portfolio A.

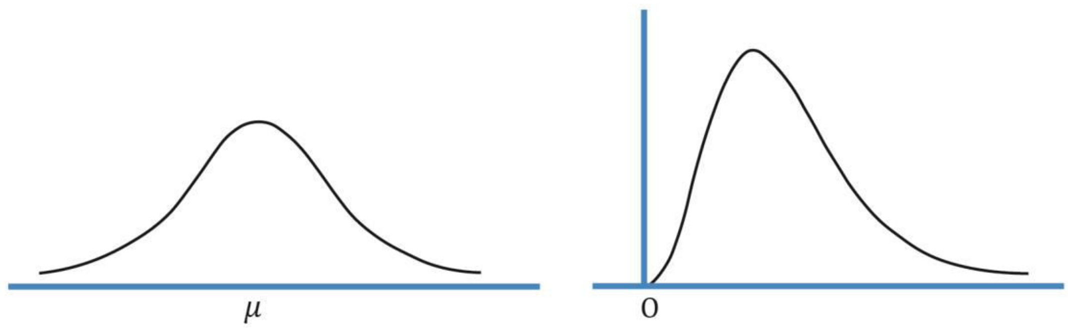
Roy's safety first criteria

It states that an optimal portfolio minimizes the probability that the actual portfolio return will fall below the target return.

6. Lognormal Distribution and Continuous Compounding

6.1 The Lognormal Distribution

If x is a random variable that is normally distributed, then to create a lognormal distribution of x we take e^x and plot the values on a graph.



Normal Distribution Lognormal Distribution

The properties of a lognormal distribution are:

- It cannot be negative.
- The upper end of its range extends to infinity.
- It is positively skewed.

Instructor's Note: A normal distribution is more suitable as a model for returns than for asset prices, because returns generally vary about the mean, with a high probability of returns being close to the mean.
However, asset prices do not vary equally about a mean price, since the probability of extreme changes in price decreases as the price approaches zero. This means asset prices will not form a symmetrical graph like that of the normal distribution. Instead, asset prices follow a lognormal distribution, which is skewed to the right and cannot be negative.

6.2 Continuously Compounded Rates of Return

Discretely compounded rates of returns have defined compounding periods such as quarterly, monthly etc. As we decrease the length of the compounding period, the effective annual rate rises. For continuous compounding, the EAR is given by:

$EAR = e^r - 1$

If we are given the holding period return over any time period, we can calculate the equivalent continuously compounded rate of return for that period as:

r = ln (HPR +1)

Example

If the holding period return of a stock was 10% for a period of one year. What is the equivalent continuously compounded rate of return for the year?

Solution:
r = ln (0.1 +1) = 0.0953 = 9.53%

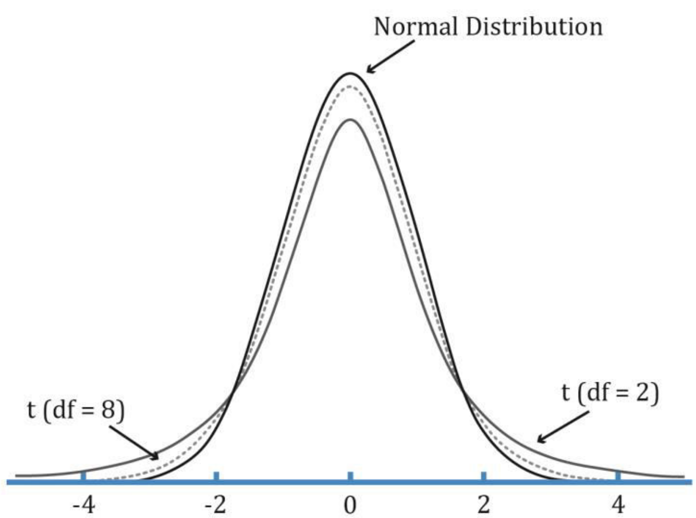
7. Student's t-, Chi-Square, And F-Distributions

The Student's t, chi-square, and F-distributions are mostly used to support statistical analyses, such as sampling, testing the statistical significance of estimated model parameters, or hypothesis testing.

7.1 Student's t-Distribution

The properties of a Student's t-distribution are:

- It is symmetrical, bell-shaped and similar to a normal distribution.
- It has a lower peak and fatter tails as compared to a normal distribution.
- It is defined by a single parameter, degrees of freedom (df) = n – 1.
- As the df increase the t-distribution approaches the standard normal distribution.

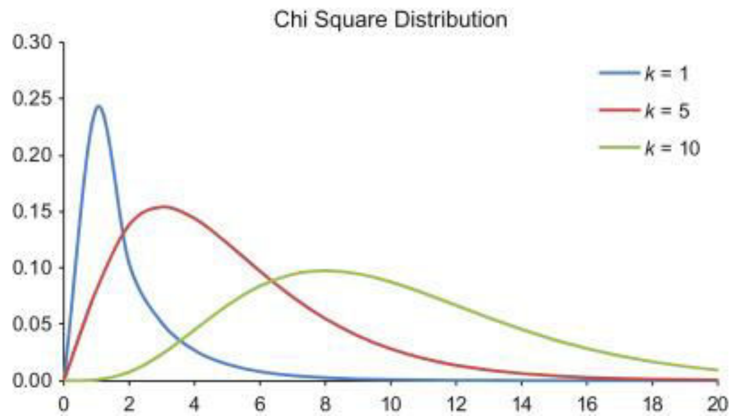


7.2 Chi-Square and F-Distribution

The properties of the Chi-square distribution are:

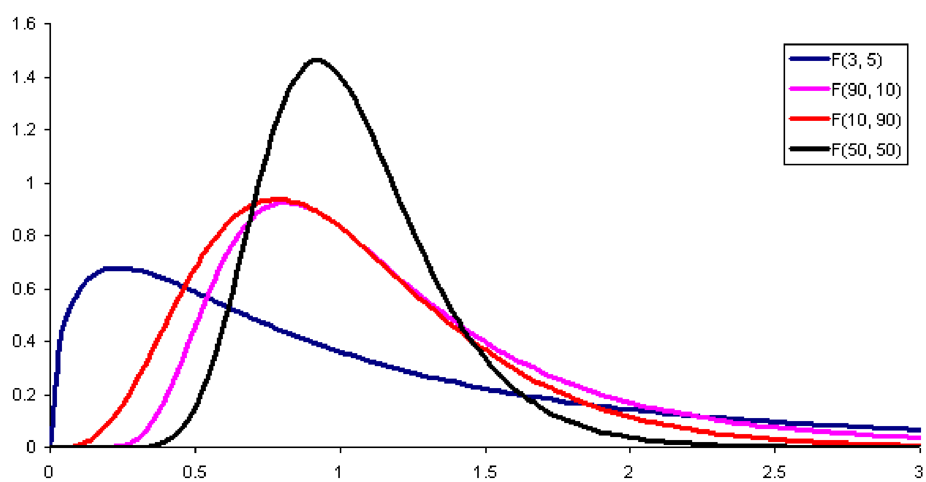
- It is asymmetrical and is defined by a single parameter, degrees of freedom (df) = n – 1

- With k degrees of freedom, the distribution is the sum of the squares of k independent standard normally distributed random variables. Therefore, the distribution does not take on negative values.
- A different distribution exists for each value of df, n-1.
- As the df increase the shape of the distribution becomes more similar to a bell curve.



The properties of the F-distribution are:

- It is asymmetrical and defined by two parameters, degrees of freedom of the numerator (df_1) and degrees of freedom of the denominator (df_2).
- The distribution does not take on negative values.
- As both the numerator (df_1) and the denominator (df_2) degrees of freedom increase, the distribution becomes more similar to a bell curve.



The relationship between the chi-square and F-distributions is as follows: If χ_1^2 is one chi-square random variable with m degrees of freedom and χ_2^2 is another chi-square random variable with n degrees of freedom, then $F = (\chi_1^2/m)/(\chi_2^2/n)$ follows an F-distribution with m numerator and n denominator degrees of freedom.

8. Monte Carlo Simulation

Monte Carlo simulation is a computer simulation used to simulate possible security prices based on risk factors. As input, it uses randomly generated values for risk factors based on their assumed distributions. It processes this information as per the specified model and runs thousands of iterations. Then it gives the distribution of the expected value of the security as output.

Major applications include:

- financial planning
- developing VAR estimates
- valuing complex securities

Limitations include:

- It is fairly complex and will provide answers that are no better than the assumptions.
- Simulation is not an analytical method but a statistical one.

Summary

LO.a: Define a probability distribution and compare and contrast discrete and continuous random variables and their probability functions.

A random variable is a variable whose outcome cannot be predicted. A probability distribution lists all possible outcomes of a random variable along with their associated probabilities.

A discrete random variable is one for which the number of possible outcomes can be counted. It has measurable probabilities associated with each specific outcome.

A continuous random variable is one for which we cannot count the number of possible outcomes. Therefore, probabilities cannot be associated with specific outcomes, instead, it has to be assigned to a particular range.

LO.b: Calculate and interpret probabilities for a random variable, given its cumulative distribution function.

The probability that an outcome will be less than or equal to a specific value is represented by the area under the cumulative probability distribution to the left of that value.

LO.c: Describe the properties of a discrete uniform random variable, and calculate and interpret probabilities given the discrete uniform distribution function.

A discrete uniform random variable is one where the probability of all the possible outcomes is equal. For example, the roll of a dice.

Probabilities for a discrete uniform distribution: If the total number of outcomes is n, then the probability of each outcome = 1/n.

LO.d: Describe the properties of the continuous uniform distribution, and calculate and interpret probabilities given a continuous uniform distribution.

The continuous uniform distribution is defined over a range from a lower limit 'a' to an upper limit 'b'. The probability that the random variable will take a value between x_1 and x_2, where x_1 and x_2 both lie within the range is given by:

$$P(x_1 \leq X \leq x_2) = \frac{x_2 - x_1}{b - a}$$

LO.e: Describe the properties of a Bernoulli random variable and a binomial random variable, and calculate and interpret probabilities given the binomial distribution function.

A Bernoulli trial is an experiment that has only two possible outcomes: a success or a failure. For example, the toss of a coin.

If the experiment is carried out n times, the number of success (denoted by X) is called a Bernoulli random variable.

The distribution that X follows is known as the binomial distribution.

The probability distribution of a binomial random variable for the probability of x successes in n trials is calculated using the following formula:

$P(x) = P(X = x) = {}_{n}C_{x}\, p^{x} (1 - p)^{n-x}$

LO.f: Explain the key properties of the normal distribution.

A normal distribution is a bell-shaped curve, with two identical halves.

It is completely described by two parameters its mean (μ) and its variance (σ^2). This is stated as $X \sim N(\mu, \sigma^2)$.

It has a skewness of 0 and a kurtosis of 3.

A linear combination of two or more random variables is also normally distributed.

LO.g: Contrast a multivariate distribution and a univariate distribution, and explain the role of correlation in the multivariate normal distribution.

A univariate distribution describes a single random variable. An example is the expected return on a particular stock. A multivariate distribution specifies the probabilities for a group of related random variables. An example is the expected return on a portfolio of multiple stocks.

A multivariate normal distribution for the returns on n stocks is completely defined by three lists of parameters:

- The list of the mean returns on the individual securities (n means in total).
- The list of the securities' variances of return (n variances in total).
- The list of all the distinct pairwise return correlations (n (n - 1)/2 distinct correlations in total).

LO.h: Calculate the probability that a normally distributed random variable lies inside a given interval.

The property of normal distribution gives us intervals stated below:

- 90% of all observations are in the interval m ± 1.65s.
- 95% of all observations are in the interval m ± 1.96s.
- 99% of all observations are in the interval m ± 2.58s.

This implies that if a random variable is less than or equal to m ± 1.65s, then the probability of it lying inside the distribution is 90%.

LO.i. Explain how to standardize a random variable.

The normal distribution with μ = 0 and σ = 1 is called the standard normal distribution. The formula for standardizing a random variable X is:

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

LO.j: Calculate and interpret probabilities using the standard normal distribution.

The Z-table is used to find the probability that X will be less than or equal to a given value. Suppose we have a normal random variable, X, with $\mu = 10$ and $\sigma = 2$. If the value of X is 11, we standardize X with Z = (11 – 10)/2 = 0.5.

The probability that we will observe a value less than 11 for $X \sim N(10, 2)$ is exactly the same as the probability that we will observe a value less than 0.5 for $Z \sim N(0, 1)$.

LO.k: Define shortfall risk, calculate the safety-first ratio, and identify an optimal portfolio using Roy's safety-first criterion.

Shortfall risk is the risk that portfolio value will fall below some minimum acceptable level over some time horizon.

$$\text{SF Ratio} = \frac{[E(R_p) - R_L]}{\sigma_p}$$

To select the optimal portfolio according to Roy's criterion, we follow the following steps:

- Calculate each portfolio's SF-Ratio.
- Choose the portfolio with the highest SF-Ratio.

LO.l: Explain the relationship between normal and lognormal distributions and why the lognormal distribution is used to model asset prices.

If x is a random variable that is normally distributed, then to create a lognormal distribution of x we take e^x and plot the values on a graph.

A lognormal distribution is often used to model asset prices because the asset prices need to be positive, they cannot be negative.

LO.m: Calculate and interpret a continuously compounded rate of return, given a specific holding period return.

For continuous compounding, the EAR is given by: $EAR = e^r - 1$.

If we are given the holding period return over any time period, we can calculate the equivalent continuously compounded rate of return for that period as: $r = \ln(HPR + 1)$

LO.n: Describe the properties of the Student's t-distribution, and calculate and interpret its degrees of freedom.

The properties of a Student's t-distribution are:

- It is symmetrical, bell-shaped and similar to a normal distribution.
- It has a lower peak and fatter tails as compared to a normal distribution.
- It is defined by a single parameter, degrees of freedom (df) = n – 1.

- As the df increase the t-distribution approaches the standard normal distribution.

LO.o: Describe the properties of the chi-square distribution and the F-distribution, and calculate and interpret their degrees of freedom.

The properties of the Chi-square distribution are:

- It is asymmetrical and is defined by a single parameter, degrees of freedom (df) = n – 1
- With k degrees of freedom, the distribution is the sum of the squares of k independent standard normally distributed random variables. Therefore, the distribution does not take on negative values.
- A different distribution exists for each value of df, n-1.
- As the df increase the shape of the distribution becomes more similar to a bell curve.

The properties of the F-distribution are:

- It is asymmetrical and defined by two parameters, degrees of freedom of the numerator (df_1) and degrees of freedom of the denominator (df_2).
- The distribution does not take on negative values.
- As both the numerator (df_1) and the denominator (df_2) degrees of freedom increase, the distribution becomes more similar to a bell curve.

LO.p: Describe Monte Carlo simulation.

Monte Carlo simulation is a computer simulation that produces a distribution of possible security prices using randomly generated values for risk factors, based on their assumed distributions.

Limitations include:

- It is fairly complex and will provide answers that are no better than the assumptions.
- Simulation is not an analytical method but a statistical one.

