Lesson 2: Continuous Random Variables, the Continuous Uniform Distribution, the Normal Distribution and the Lognormal Distribution

A **continuous uniform distribution** is described by an upper limit, *a*, and a lower limit, *b*.

- These limits serve as the parameters of the distribution.
- The probability of any outcome or range of outcomes outside these limits is 0.
- Individual outcomes also have a probability of 0.
- The distribution is often denoted as U(*a*,*b*).

The probability that the random variable will take a value that falls between x_1 and x_2 that both lie within the range, *a* to *b*, is the proportion of the total area taken up by the range, x_1 to x_2.

Example

X is a uniformly distributed continuous random variable that cannot take values lower than 2 or greater than 10. Calculate the probability that X will fall between 5 and 7.

The **normal distribution** is a continuous distribution that plays a central role in quantitative analysis.

Properties of the Normal Distribution

It is completely described by its mean (μ) and variance (σ^2).

- The distribution is stated as $X \sim N(\mu,\sigma^2)$.
- The random variable X follows the normal distribution with mean, μ and variance, σ^2.

The distribution has a skewness of 0, which means that it is symmetric about its mean.

- $P(X \leq \text{mean}) = P(X \geq \text{mean}) = 0.5$.
- The mean, median and mode are the same.

Kurtosis equals 3 and excess kurtosis equals 0

A linear combination of normally distributed random variables is also normally distributed.

The probability of the random variable lying in ranges further away from the mean gets smaller and smaller, but never goes to zero.

- The tail on either side extends to infinity.

UNIVARIATE VERSUS MULTIVARIATE DISTRIBUTIONS

Univariate distributions describe the distribution of a single random variable.

Multivariate distributions specify probabilities associated with a group of random variables taking into account the interrelationships that may exist between them.
A multivariate normal distribution for the return on a portfolio with n stocks is completely defined by the following three parameters:

- The mean returns on all the n individual stocks
- The variances of returns of all n individual stocks
- The return correlations between each possible pair of stocks.

Note: The need to determine correlations is what differentiates multivariate normal distributions from univariate normal distributions.

A **confidence interval** represents the range of values within which a certain population parameter is expected to lie a specified percentage of the time.

A 95% confidence interval between 4 and 8 implies that:

- We can be 95% confident that the relevant parameter lies between 4 and 8.
- The probability that the parameter lies between 4 and 8 equals 0.95.
- There is a 5% chance that the parameter does not lie between 4 and 8.

Example

The average annual return on a stock is 15% and the standard deviation of returns equals 5%. Given that the stock's returns are distributed normally, calculate the 90% confidence interval for the return in any given year.

The **standard normal distribution curve** is a normal distribution that has been standardized so that it has a mean of zero and a standard deviation of 1.

Z-Score

- Standardized value.
- Calculated by subtracting the mean of the population from the observed value of the random variable, and dividing the result by the standard deviation.
- z = (observed value - population mean)/standard deviation = $(x - \mu)/\sigma$.

Essentially, the z-value represents the number of standard deviations away from the population mean, a given observation lies.

Example

Assume that the monthly salaries of the employees of Nippon Enterprises are normally distributed with a mean of $1,500 and standard deviation of $500. What are the z-scores for salaries of $1,200 and $3,000?

Calculating Probabilities Using Z-Scores

Calculating Probabilities Using Z-Scores

Example

The points, X, scored by students in a class on their final exam are normally distributed with a mean of 60 and standard deviation of 15. What is the probability that the points scored by a given student will be:

a. less than 80
b. more than 70
c. less than 40
d. between 40 and 70.

Shortfall risk refers to the probability that a portfolio's value or return, $E(R_P)$, will fall below a particular target value or return (R_T) over a given period.

Roy's safety-first criterion states that an optimal portfolio minimizes the probability that the actual portfolio return, R_P, will fall below the target return, R_T.

- The *higher* the value of the SF Ratio, the better the risk-return tradeoff the portfolio offers, given the investor's threshold level.
- A portfolio with a *higher* SF Ratio also has a *lower* probability of attaining returns lower than the threshold level.

Example

An investor has $1,000 to invest. His minimum acceptable portfolio value at the end of the year is $1,050. He is considering two portfolios, A and B. Portfolio A has expected returns of 8% and a standard deviation of 10%, while Portfolio B has expected returns of 12% and a standard deviation of 15%. Using Roy's safety-first criterion, select the optimal portfolio, and calculate the probability that the portfolio's return will fall short of its target.

A random variable, Y, follows the **lognormal distribution** if its natural logarithm (ln Y) is normally distributed. Three important features differentiate the lognormal distribution from the normal distribution:

1. It is bounded by zero on the lower end.
2. The upper end of its range is unbounded.
3. It is skewed to the right (positively skewed).

The holding period return on any asset can range between -100% and + ∞.

Instead of using the return on the investment as the random variable, if we use the ratio of the ending value of the investment to its beginning value, the distribution will still be skewed, but with a lower bound of zero, and no upper bound.

Finally, if we were to take the distribution of the natural logarithm of V_t/V_0 we would find that the distribution is unbounded at both ends.

Conclusion: While the distribution of the variable (V_t/V_0) follows the lognormal distribution (lower bond = 0; no upper bound), the distribution of the natural logarithm of (V_t/V_0) is normally distributed with a lower bound of $-\infty$ and upper bound of $+\infty$.

The **discretely compounded rate of return** is based on discrete or defined compounding periods, such as 12 months or 6 months.

- As the compounding periods get *shorter*, the effective annual rate (EAR) *rises.*

The **continuously compounded rate of return** is based on continuous compounding.

- The effective annual rate with continuous compounding is given as:

Calculating the continuously compounded stated annual rate of return given a holding period return

Example

A stock that was purchased for $65 one year ago was sold for $86 today. What was the continuously compounded return on this investment?

Relationship Between HPY and the Continuously Compounded Rate

Example

An investment of $1,000 appreciates to a value of $1,250 in 2 years. Calculate the continuously compounded return on this investment.