The following is a review of the Quantitative Analysis principles designed to address the learning objectives set forth by GARP®. This topic is also covered in:

Volatility

Topic 28

Exam Focus

Traditionally, volatility has been synonymous with risk. Thus, the accurate estimation of volatility is crucial to understanding potential risk exposure. This topic pertains to methods that employ historical data when generating estimates of volatility. Simplistic models tend to generate estimates assuming volatility remains constant over short time periods. Conversely, complex models account for variations over time. For the exam, be able to estimate volatility using both the exponentially weighted moving average (EWMA) and the generalized autoregressive conditional heteroskedasticity [GARCH(1,1)] models.

Volatility, Variance, and Implied Volatility

LO 28.1: Define and distinguish between volatility, variance rate, and implied volatility.

The **volatility** of a variable, σ, is represented as the standard deviation of that variable's continuously compounded return. With option pricing, volatility is typically expressed as the standard deviation of return over a one-year period. This differs from risk management, where volatility is typically expressed as the standard deviation of return over a one-day period.

The traditional measure of volatility first requires a measure of change in asset value from period to period. The calculation of a continuously compounded return over successive days is as follows:

$$u_i = \ln\left(\frac{S_i}{S_{i-1}}\right)$$

where:
S_i = asset price at time i

This is similar to the proportional change in an asset, which is calculated as follows:

$$u_i = \frac{S_i - S_{i-1}}{S_{i-1}}$$

From a risk management perspective, the daily volatility of an asset usually refers to the standard deviation of the daily proportional change in asset value.

By assuming daily returns are independent with the same level of variation, daily volatility can be extended over a number of days, T, by multiplying the standard deviation of the return by the square root of T. This is known as the *square root of time rule*. For example, if the daily volatility is 1.5%, the standard deviation of the return (compounded continuously) over a 10-day period would be computed as $1.5\% \times \sqrt{10} = 4.74\%$. Note that when converting daily volatility to annual volatility, the usual practice is to use the square root of 252 days, which is the number of business days in a year, as opposed to the number of calendar days in a year.

Risk managers may also compute a variable's **variance rate**, which is simply the square of volatility (i.e., standard deviation squared: σ^2). In contrast to volatility, which increases with the square root of time, the variance of an asset's return will increase in a linear fashion over time. For example, if the daily volatility is 1.5%, the variance rate is $1.5\%^2 = 0.0225\%$. Thus, over a 10-day period, the variance will be 0.225% (i.e., $0.0225\% \times 10$).

In addition to variance and standard deviation, which are computed using historical data, risk managers may also derive implied volatilities. The **implied volatility** of an option is computed from an option pricing model, such as the Black-Scholes-Merton (BSM) model. The volatility of an asset is not directly observed in the BSM model, so we compute implied volatility as the volatility level that will result when equating an option's market price to its model price.

Professor's Note: Computing option prices using the BSM model will be demonstrated in Book 4.

The most widely used index for publishing implied volatility is the Chicago Board Options Exchange (CBOE) Volatility Index (ticker symbol: VIX). The VIX demonstrates implied volatility on a wide variety of 30-day calls and puts on the S&P 500 Index. Note that trading in futures and options on the VIX is a bet on volatility only. Since its inception, the VIX has mainly traded between 10 and 20 (which corresponds to volatility of 10%–20% on the S&P 500 Index options), but it reached a peak of close to 80 in October 2008, after the collapse of Lehman Brothers. The VIX is often referred to as the fear index by market participants because it reflects current market uncertainties.

The Power Law

LO 28.2: Describe the power law.

It is typically assumed that the change in asset prices is normally distributed. This makes it convenient to apply standard deviation when determining confidence intervals for an asset's price. For example, by assuming an asset price of \$50 and a volatility of 4.47%, we can compute a one-standard-deviation move as $50 \times 0.0447 = 2.24$. With this information, we can define the 95% confidence interval as $50 \pm 1.96 \times 2.24$.

In practice, however, the distribution of asset price changes is more likely to exhibit fatter tails than the normal distribution. Thus, heavy-tailed distributions can be used to better capture the possibility of extreme price movements (e.g., a five-standard-deviation move). An alternative approach to assuming a normal distribution is to apply the power law.

The power law states that when *X* is large, the value of a variable *V* has the following property:

$$P(V > X) = K \times X^{-\alpha}$$

where:
V = the variable
X = large value of *V*
K and α = constants

Example: The power law

Assume that data on asset price changes determines the constants in the power law equation to be the following: K = 10 and α = 5. **Calculate** the probability that this variable will be greater than a value of 3 and a value of 5.

Answer:

$$P(V > 3) = 10 \times 3^{-5} = 0.0412 \text{ or } 4.12\%$$
$$P(V > 5) = 10 \times 5^{-5} = 0.0032 \text{ or } 0.32\%$$

By taking the logarithm of both sides in the power law equation, we can perform regression analysis to determine the power law constants, *K* and α:

$$\ln[P(V > X)] = \ln(K) - \alpha \ln(X)$$

In this case, the dependent variable, ln[P(V > X)], can be plotted against the independent variable, ln(X). Furthermore, if we assume that *X* represents the number of standard deviations that a given variable will change, we can determine the probability that *V* will exceed a certain number of standard deviations. For example, if regression analysis indicates that K = 8 and α = 5, the probability that the variable will exceed four standard deviations will be equal to 8×4^{-5} = 0.0078 or 0.78%. The power law suggests that extreme movements have a very low probability of occurring, but this probability is still higher than what is indicated by the normal distribution.

Estimating Volatility

LO 28.3: Explain how various weighting schemes can be used in estimating volatility.

By collecting continuously compounded return data, u_i, over a number of days (as shown in LO 28.1), we can compute the mean return of the individual returns as follows:

$$\bar{u} = \frac{1}{m}\sum_{i=1}^{m} u_{n-i}$$

where:
m = number of observations leading up to the present period

If we assume that the mean return is zero, which would be true when the mean is small compared to the variability, we obtain the maximum likelihood estimator of variance:

$$\sigma_n^2 = \frac{1}{m}\sum_{i=1}^{m} u_{n-i}^2$$

In simplest terms, historical data is used to generate returns in an asset-pricing series. This historical return information is then used to generate a volatility parameter, which can be used to infer expected realizations of risk. However, the straightforward approaches just presented weight each observation equally in that more distant past returns have the same influence on estimated volatility as observations that are more recent. If the goal is to estimate the current level of volatility, we may want to weight recent data more heavily. There are various weighting schemes, which can all essentially be represented as:

$$\sigma_n^2 = \sum_{i=1}^{m} \alpha_i u_{n-i}^2$$

where:
α_i = weight on the return i days ago

The weights (the α's) must sum to one, and if the objective is to generate a greater influence on recent observations, then the α's will decline in value for older observations.

One extension to this weighting scheme is to assume a long-run variance level in addition to the weighted squared return observations. The most frequently used model is an **autoregressive conditional heteroskedasticity model**, ARCH(m), which can be represented by:

$$\sigma_n^2 = \gamma V_L + \sum_{i=1}^{m} \alpha_i u_{n-i}^2 \text{ with } \gamma + \sum \alpha_i = 1 \text{ so that}$$

$$\sigma_n^2 = \omega + \sum_{i=1}^{m} \alpha_i u_{n-i}^2$$

where:
$\omega = \gamma V_L$ (long-run variance weighted by the parameter γ)

Therefore, the volatility estimate is a function of a long-run variance level and a series of squared return observations, whose influence declines the older the observation is in the time series of the data.

The Exponentially Weighted Moving Average Model

LO 28.4: Apply the exponentially weighted moving average (EWMA) model to estimate volatility.

LO 28.8: Explain the weights in the EWMA and GARCH(1,1) models.

The **exponentially weighted moving average** (EWMA) model is a specific case of the general weighting model presented in the previous section. The main difference is that the weights are assumed to decline exponentially back through time. This assumption results in a specific relationship for variance in the model:

$$\sigma_n^2 = \lambda\sigma_{n-1}^2 + (1-\lambda)u_{n-1}^2$$

where:
λ = weight on previous volatility estimate (λ between zero and one)

The simplest interpretation of the EWMA model is that the day-*n* volatility estimate is calculated as a function of the volatility calculated as of day $n - 1$ and the most recent squared return. Depending on the weighting term λ, which ranges between zero and one, the previous volatility and most recent squared returns will have differential impacts. High values of λ will minimize the effect of daily percentage returns, whereas low values of λ will tend to increase the effect of daily percentage returns on the current volatility estimate.

Example: EWMA model

The decay factor in an exponentially weighted moving average model is estimated to be 0.94 for daily data. Daily volatility is estimated to be 1%, and today's stock market return is 2%. What is the new estimate of volatility using the EWMA model?

Answer:

$$\sigma_n^2 = 0.94 \times 0.01^2 + (1-0.94) \times 0.02^2 = 0.000118$$

$$\sigma_n = \sqrt{0.000118} = 1.086\%$$

One benefit of the EWMA is that it requires few data points. Specifically, all we need to calculate the variance is the current estimate of the variance and the most recent squared return. The current estimate of variance will then feed into the next period's estimate, as will this period's squared return. Technically, the only "new" piece of information for the volatility calculation will be that attributed to the squared return.

The GARCH(1,1) Model

LO 28.5: Describe the generalized autoregressive conditional heteroskedasticity (GARCH (p,q)) model for estimating volatility and its properties.

LO 28.6: Calculate volatility using the GARCH(1,1) model.

One of the most popular methods of estimating volatility is the **generalized autoregressive conditional heteroskedastic** (GARCH)(1,1) model. A GARCH(1,1) model not only incorporates the most recent estimates of variance and squared return, but also a variable that accounts for a long-run average level of variance.

Professor's Note: In the GARCH(p,q) notation, the p stands for the number of lagged terms on historical returns squared, and the q stands for the number of lagged terms on historical volatility.

The best way to describe a GARCH(1,1) model is to take a look at the formula representing its determination of variance, which can be shown as:

$$\sigma_n^2 = \omega + \alpha u_{n-1}^2 + \beta\sigma_{n-1}^2$$

where:

α = weighting on the previous period's return

β = weighting on the previous volatility estimate

ω = weighted long-run variance $= \gamma V_L$

V_L = long-run average variance $= \dfrac{\omega}{1-\alpha-\beta}$

$\alpha + \beta + \gamma = 1$

$\alpha + \beta < 1$ for stability so that γ is not negative

The EWMA is nothing other than a special case of a GARCH(1,1) volatility process, with $\omega = 0$, $\alpha = 1 - \lambda$, and $\beta = \lambda$. Similar to the EWMA model, β represents the exponential decay rate of information. The GARCH(1,1) model adds to the information generated by the EWMA model in that it also assigns a weighting to the average long-run variance estimate. An additional characteristic of a GARCH(1,1) estimate is the implicit assumption that variance tends to revert to a long-term average level. Recognition of a mean-reverting characteristic in volatility is an important feature when pricing derivative securities such as options.

Example: GARCH(1,1) model

The parameters of a generalized autoregressive conditional heteroskedastic (GARCH)(1,1) model are ω = 0.000003, α = 0.04, and β = 0.92. If daily volatility is estimated to be 1%, and today's stock market return is 2%, what is the new estimate of volatility using the GARCH(1,1) model, and what is the implied long-run volatility level?

Answer:

$$\sigma_n^2 = 0.000003 + 0.04 \times 0.02^2 + 0.92 \times 0.01^2 = 0.000111$$

$$\sigma_n = \sqrt{0.000111} = 1.054\%$$

$$\text{long-run average variance} = \frac{\omega}{(1-\alpha-\beta)} = \frac{0.000003}{(1-0.04-0.92)} = 0.000075$$

$$\bar{\sigma} = \sqrt{0.000075} = 0.866\% = \text{long-run volatility}$$

Mean Reversion

LO 28.7: Explain mean reversion and how it is captured in the GARCH(1,1) model.

Empirical data indicates that volatility exhibits a mean-reverting characteristic. Given that stylized fact, a GARCH model tends to display a better theoretical justification than the EWMA model. The method for estimating the GARCH parameters (or weights), however, often generates outcomes that are not consistent with the model's assumptions. Specifically, the sum of the weights of α and β are sometimes greater than one, which causes instability in the volatility estimation. In this case, the analyst must resort to using an EWMA model.

The sum of α + β is called the **persistence**, and if the model is to be stationary over time (with reversion to the mean), the sum must be less than one. The persistence describes the rate at which the volatility will revert to its long-term value following a large movement. The higher the persistence (given that it is less than one), the longer it will take to revert to the mean following a shock or large movement. A persistence of one means that there is no reversion, and with each change in volatility, a new level is attained.

Estimation and Performance of GARCH Models

As was previously mentioned, one way to estimate volatility (e.g., variance) is to use a **maximum likelihood estimator**. Maximum likelihood estimators select values of model parameters that maximize the likelihood that the observed data will occur in a sample. Any variable of interest can be estimated via the maximum likelihood method, which requires formulating an expression or function for the underlying probability distribution of the data and then searching for the parameters that maximize the value generated by the expression.

One important consideration relates to which distribution is chosen when calculating probability. The most popular is the normal distribution, but normally distributed data are not often found in financial markets.

GARCH models are estimated using maximum likelihood techniques. The estimation process begins with a guess of the model's parameters. Then a calculation of the likelihood function based on those parameter estimates is made. The parameters are then slightly adjusted until the likelihood function fails to increase, at which time the estimation process assumes it has maximized the function and stops. The values of the parameters at the point of maximum value in the likelihood function are then used to estimate GARCH model volatility.

LO 28.9: Explain how GARCH models perform in volatility forecasting.

LO 28.10: Describe the volatility term structure and the impact of volatility changes.

One of the useful features of GARCH models is that they do a very good job at modeling volatility clustering when periods of high volatility tend to be followed by other periods of high volatility and periods of low volatility tend to be followed by subsequent periods of low volatility. Thus, there is autocorrelation in u_i^2. If GARCH models do a good job of explaining volatility changes, there should be very little autocorrelation in u_i^2 / σ_i^2. GARCH models appear to do a very good job of explaining volatility.

The question then arises, if GARCH models do a good job at explaining past volatility, how well do they forecast future volatility? The simple answer to this question is that GARCH models do a fine job at forecasting volatility from a volatility term structure perspective (e.g., estimates of volatility given time to expiration for options). Even though the actual volatility term structure figures are somewhat different from those forecasted by GARCH models, GARCH-generated volatility data does an excellent job in predicting how the volatility term structure responds to changes in volatility. This modeling tool is quite frequently used by financial institutions when estimating exposure to various option positions.

Key Concepts

LO 28.1

The volatility of a variable is the standard deviation of that variable's continuously compounded return. The variance rate of a variable is the square of its standard deviation. Variance and standard deviation are computed using historical data. Risk managers may also compute implied volatility, which is the volatility that forces a model price (i.e., option pricing model) to equal the market price.

LO 28.2

The power law is an alternative approach to using probabilities from a normal distribution. It states that when X is large, the value of a variable V has the following property, where K and α are constants:

$$P(V > X) = K \times X^{-\alpha}$$

LO 28.3

Historical price data is used to generate return estimates, which are then used to estimate volatility. Traditional volatility estimation methods weight past information equally across time. Weighting schemes can be used to weight recent information more heavily than distant data.

LO 28.4

The EWMA model generates volatility estimates based on weightings of the last estimate of volatility and the latest current price change information. The objective is to account for previous volatility estimates, as well as to account for the latest return information.

$$\sigma_n^2 = \lambda\sigma_{n-1}^2 + (1-\lambda)u_{n-1}^2$$

where:
λ = weight on previous volatility estimate (λ between zero and one)

LO 28.5

One of the most popular methods of estimating volatility is the generalized autoregressive conditional heteroskedastic (GARCH)(p,q) model. In a GARCH(p,q) model, the *p* stands for the number of lagged terms on historical returns squared, and the *q* stands for the number of lagged terms on historical volatility.

LO 28.6

GARCH(1,1) models not only incorporate the most recent estimates of volatility and return, but also incorporate a long-run average level of variance.

$$\sigma_n^2 = \omega + \alpha u_{n-1}^2 + \beta\sigma_{n-1}^2$$

where:

α = weighting on the previous period's return
β = weighting on the previous volatility estimate
ω = weighted long-run variance $= \gamma V_L$
V_L = long-run average variance $= \dfrac{\omega}{1-\alpha-\beta}$
$\alpha + \beta + \gamma = 1$
$\alpha + \beta < 1$ for stability so that γ is not negative

GARCH(1,1) estimates of volatility have a better theoretical justification than the EWMA model. In the event that model parameter estimates indicate instability, however, EWMA volatility estimates may be used.

LO 28.7

In a GARCH(1,1) model, the sum of $\alpha + \beta$ is called the persistence. The persistence describes the rate at which the volatility will revert to its long-term value. A persistence equal to one means there is no mean reversion.

LO 28.8

The EWMA is nothing other than a special case of a GARCH(1,1) volatility process, with $\omega = 0$, $\alpha = 1 - \lambda$, and $\beta = \lambda$. Similar to the EWMA model, β in the GARCH(1,1) equation represents the exponential decay rate of information. The GARCH(1,1) model adds to the information generated by the EWMA model in that it also assigns a weighting to the average long-run variance estimate.

LO 28.9

GARCH models do a very good job at modeling volatility clustering when periods of high volatility tend to be followed by other periods of high volatility and periods of low volatility tend to be followed by subsequent periods of low volatility.

LO 28.10

When forecasting future volatility, GARCH-generated volatility data does an excellent job in predicting the volatility term structure (i.e., differing volatilities for options given differing maturities). This modeling tool is quite frequently used by financial institutions when estimating exposure to various option positions.

Concept Checkers

1. An analyst is attempting to compute a confidence interval for asset Z prices. Assume a daily volatility of 1% and a current asset price of 100. What is the 95% confidence interval for asset price at the end of five days, assuming price changes are normally distributed?
   A. 100 ± 1.96.
   B. 100 ± 2.24.
   C. 100 ± 4.39.
   D. 100 ± 9.80.

2. The parameters of a generalized autoregressive conditional heteroskedastic (GARCH)(1,1) model are $\omega = 0.00003$, $\alpha = 0.04$, and $\beta = 0.92$. If daily volatility is estimated to be 1.5%, and today's stock market return is 0.8%, what is the new estimate of the standard deviation?
   A. 1.68%.
   B. 1.55%.
   C. 1.45%.
   D. 2.74%.

3. The λ of an exponentially weighted moving average (EWMA) model is estimated to be 0.9. Daily standard deviation is estimated to be 1.5%, and today's stock market return is 0.8%. What is the new estimate of the standard deviation?
   A. 1.68%.
   B. 1.55%.
   C. 1.45%.
   D. 2.74%.

4. The parameters of a GARCH(1,1) model are $\omega = 0.00003$, $\alpha = 0.04$, and $\beta = 0.92$. These figures imply a long-run daily standard deviation of:
   A. 1.68%.
   B. 1.55%.
   C. 1.45%.
   D. 2.74%.

5. GARCH(1,1) models can only be used to estimate volatility in the case where:
   A. $\alpha + \beta > 0$.
   B. $\alpha + \beta < 1$.
   C. $\alpha > \beta$.
   D. $\alpha < \beta$.

【梦轩考资网www.mxkaozi.com】QQ106454842 专业提供CFA FRM全程高清视频+讲义

Concept Checker Answers

1. **C** First, convert daily volatility to weekly volatility using the square root to time: $1\% \times \sqrt{5} = 2.24\%$. Next, compute the one-standard-deviation move: $100 \times 0.0224 = 2.24$. Finally, derive the confidence interval: $100 \pm 1.96 \times 2.24 = 100 \pm 4.39$.

2. **B** $\sigma_n^2 = 0.00003 + (0.008)^2 \times 0.04 + (0.015)^2 \times 0.92 = 0.00023956$

   $$\sigma_n = \sqrt{0.00023956} = 0.0155 = 1.55\%$$

3. **C** $\sigma_n^2 = 0.9 \times (0.015)^2 + (1 - 0.9) \times (0.008)^2 = 0.0002089$

   $$\sigma_n = \sqrt{0.0002089} = 0.0145 = 1.45\%$$

4. **D** The long-run variance rate can be estimated by dividing the ω of a GARCH(1,1) model by $1 - \alpha - \beta$. This yields $0.00003 / (1 - 0.04 - 0.92) = 0.00075$; long-run standard deviation = $\sqrt{0.00075} = 0.0274 = 2.74\%$.

5. **B** Stable GARCH(1,1) models require $\alpha + \beta < 1$; otherwise the model is unstable.

The following is a review of the Quantitative Analysis principles designed to address the learning objectives set forth by GARP®. This topic is also covered in:

Correlations and Copulas

Topic 29

Exam Focus

This topic examines correlation and covariance calculations and how covariance is used in exponentially weighted moving average (EWMA) and generalized autoregressive conditional heteroskedasticity (GARCH) models. The later part of this topic defines copulas and distinguishes between several different types of copulas. For the exam, be able to calculate covariance using EWMA and GARCH(1,1) models. Also, understand how copulas are used to estimate correlations between variables. Finally, be able to explain how marginal distributions are mapped to known distributions to form copulas.

Correlation and Covariance

LO 29.1: Define correlation and covariance and differentiate between correlation and dependence.

Correlation and covariance refer to the co-movements of assets over time and measure the strength between the linear relationships of two variables. Correlation and covariance essentially measure the same relationship; however, correlation is standardized so the value is always between –1 and 1. This standardized measure is more convenient in risk analysis applications than covariance, which can have values between –∞ and ∞. **Correlation** is mathematically determined by dividing the covariance between two random variables, cov(X,Y), by the product of their standard deviations, $\sigma_X\sigma_Y$.

$$\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sigma_X\sigma_Y}$$

Multiplying each side of this equation by $\sigma_X\sigma_Y$ provides the formula for calculating **covariance:**

$$\text{cov}(X,Y) = \rho_{X,Y} \times \sigma_X\sigma_Y$$

In practice, it is necessary to first calculate the covariance between two random variables using the following equation and then solve for the standardized correlation.

$$\text{cov}(X,Y) = E\left[(X - E(X)) \times (Y - E(Y))\right] = E(X,Y) - E(X) \times E(Y)$$

In this covariance equation, E(X) and E(Y) are the means or expected values of random variables *X* and *Y*, respectively. E(X,Y) is the expected value of the product of random variables *X* and *Y*.

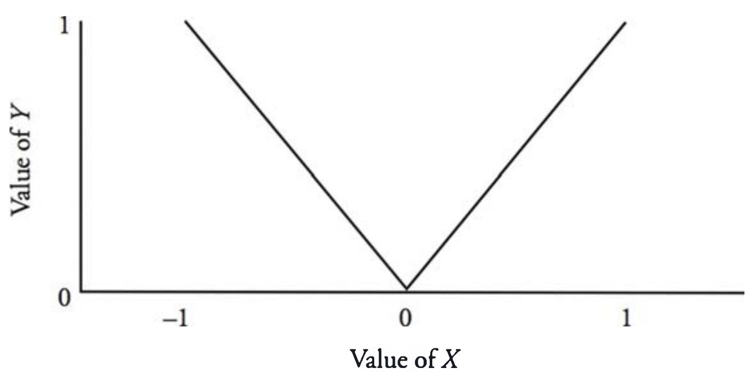
Variables are defined as independent variables if the knowledge of one variable does not impact the probability distribution for another variable. In other words, the conditional probability of V_2 given information regarding the probability distribution of V_1 is equal to the unconditional probability of V_2 as expressed in the following equation:

$$P(V_2 \mid V_1 = x) = P(V_2)$$

A correlation of zero between two variables ***does not*** imply that there is no dependence between the two variables. It simply implies that there is no linear relationship between the two variables, but the value of one variable can still have a nonlinear relationship with the other variable.

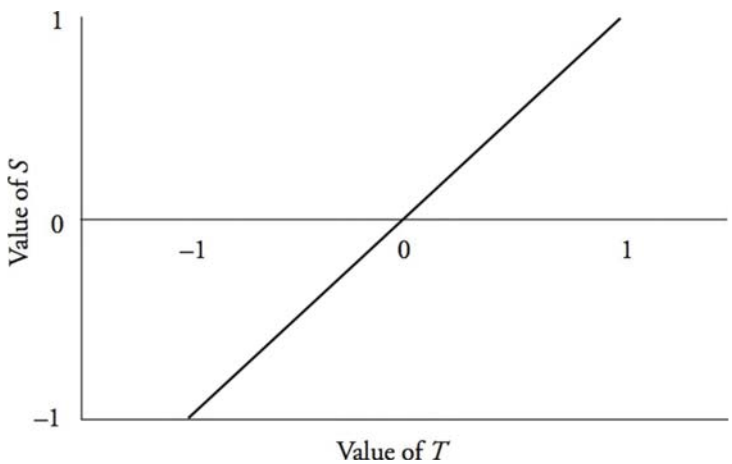
As an example, suppose variable *X* has three expected values of –1, 0, and 1 with an equal probability of occurrence, and variable *Y* has a value of 1 when variable *X* has a value of either –1 or 1. When variable *X* has a value of 0, then variable *Y* has a value of 0. This V-shaped relationship is illustrated in Figure 1.

Figure 1: Relationship between *X* and *Y*



Also suppose that variables *S* and *T* are perfectly positively correlated and that variable *S* has three expected values of –1, 0, and 1 with an equal probability of occurrence. When variable *S* has a value of –1, 0, or 1, then variable *T* has a value of –1, 0, and 1, respectively. This relationship is illustrated in Figure 2.

Figure 2: Relationship between S and T



With the above information, we can now determine the correlation coefficient and dependency of these two pairs of variables. In this example, the coefficient of correlation between variables X and Y is zero, and the coefficient of correlation between variables S and T is one.

The coefficient of correlation is a statistical measure of linear dependency. If we know the value of X, it will change our expectations of the value or probability distribution of Y. Likewise, if we know the value of Y, it will change our expectations of the probability distribution of X. Clearly, there is a dependency between X and Y, as well as a dependency between S and T. A practical example of the V-shaped dependency in Figure 1 is with respect to financial derivatives that may have more value with large market movements in either direction.

Covariance Using EWMA and GARCH Models

LO 29.2: Calculate covariance using the EWMA and GARCH(1,1) models.

EWMA Model

Covariance is a statistical measure that is calculated over historical time periods. Conventional wisdom suggests that more recent observations should carry more weight because they more accurately reflect the current market environment. The following equation calculates a new covariance on day n using an **exponentially weighted moving average (EWMA) model.** This model is designed to vary the weight given to more recent observations (by adjusting λ).

$$\text{cov}_n = \lambda\,\text{cov}_{n-1} + (1-\lambda)X_{n-1}Y_{n-1}$$

where:
λ = the weight for the most recent covariance on day $n - 1$
X_{n-1} = the percentage change for variable X on day $n - 1$
Y_{n-1} = the percentage change for variable Y on day $n - 1$

Example: Calculating covariance using the EWMA model

Assume an analyst uses the EWMA model with $\lambda = 0.90$ to update correlation and covariance rates. The correlation estimate for two variables *X* and *Y* on day *n* –1 is 0.7. In addition, the estimated standard deviations on day *n* – 1 for variables *X* and *Y* are 1.5% and 2%, respectively. Also, the percentage change on day *n* – 1 for variables *X* and *Y* are 2% and 1%, respectively. What is the updated estimate of the covariance rate and correlation between *X* and *Y* on day *n*?

Answer:

The estimated covariance rate between variables *X* and *Y* on day *n* – 1 can be calculated as:

$$\text{cov}(X,Y) = \rho_{X,Y} \times \sigma_X \sigma_Y = 0.7 \times 0.015 \times 0.02 = 0.00021$$

With this value, the EWMA model can update the covariance rate for day *n*.

$$\text{cov}_n = 0.9 \times 0.00021 + 0.1 \times 0.02 \times 0.01 = 0.000189 + 0.00002 = 0.000209$$

Note that the covariance of an asset with itself is equal to the variance of the asset ($\text{cov}(X,X) = \sigma_X^2$). Thus, the EWMA equation can also be used to estimate the new variances for variables *X* and *Y*. The modified equation for updating the variance of *X* becomes:

$$\sigma_{X,n}^2 = \lambda \sigma_{X,n-1}^2 + (1-\lambda) X_{n-1}^2$$

$$\sigma_{X,n}^2 = 0.9 \times 0.015^2 + 0.1 \times 0.02^2 = 0.0002025 + 0.00004 = 0.0002425$$

Similarly, the updated variance for variable *Y* is calculated as follows:

$$\sigma_{Y,n}^2 = 0.9 \times 0.02^2 + 0.1 \times 0.01^2 = 0.00036 + 0.00001 = 0.00037$$

The new standard deviation estimates for *X* and *Y* are found by taking the square root of their respective variances. The new volatility measure of *X* is:

$$\sigma_{X,n} = \sqrt{0.0002425} = 0.0155724$$

The new volatility measure of *Y* is:

$$\sigma_{Y,n} = \sqrt{0.00037} = 0.0192354$$

Therefore, the new correlation on day n can be found by dividing the updated covariance (cov_n) by the updated standard deviations for X and Y:

$$\frac{0.000209}{0.0155724 \times 0.0192354} = 0.6977$$

GARCH(1,1) Model

An alternative method for updating the covariance rate for two variables X and Y uses the **generalized autoregressive conditional heteroskedasticity (GARCH) model**. The GARCH(1,1) model for updating covariance rates is defined as follows:

$$cov_n = \omega + \alpha X_{n-1} Y_{n-1} + \beta cov_{n-1}$$

GARCH(1,1) applies a weight of α to the most recent observation on covariance ($X_{n-1}Y_{n-1}$) and a weight of β to the most recent covariance estimate (cov_{n-1}). In addition, a weight of ω is given to the long-term average covariance rate.

Professor's Note: Recall that the EWMA is a special case of GARCH(1,1), where $\omega = 0$, $\alpha = 1 - \lambda$, and $\beta = \lambda$.

An alternative form for writing the GARCH(1,1) model is shown as follows:

$$cov_n = \gamma V_L + \alpha X_{n-1} Y_{n-1} + \beta cov_{n-1}$$

where:
γ = weight assigned to the long-term variance, V_L

In this equation, the three weights must equal 100% ($\gamma + \alpha + \beta = 1$). If α and β are known, the weight for the long-term variance, γ, can be determined as $1 - \alpha - \beta$. Therefore, the long-term average covariance rate must equal: $\omega / (1 - \alpha - \beta)$.

Example: Calculating covariance using the GARCH(1,1) model

Assume an analyst uses daily data to estimate a GARCH(1,1) model as follows:

$$cov_n = 0.000002 + 0.14 X_{n-1} Y_{n-1} + 0.76 cov_{n-1}$$

This implies $\alpha = 0.14$, $\beta = 0.76$, and $\omega = 0.000002$. The analyst also determines that the estimate of covariance on day $n - 1$ is 0.000324 and the most recent returns on X and Y are both 0.02. What is the updated estimate of covariance?

Answer:

The updated estimate of covariance on day n is 0.0304%, which is calculated as:

$$\begin{aligned}\text{cov}_n &= 0.000002 + \left(0.14 \times 0.02^2\right) + (0.76 \times 0.000324) \\ &= 0.000002 + 0.000056 + 0.000246 = 0.000304\end{aligned}$$

Evaluating Consistency for Covariances

LO 29.3: Apply the consistency condition to covariance.

A **variance-covariance matrix** can be constructed using the calculated estimates of variance and covariance rates for a set of variables. The diagonal of the matrix represents the variance rates where $i = j$. The covariance rates are all other elements of the matrix where $i \neq j$.

A matrix is known as *positive-semidefinite* if it is internally consistent. The following expression defines the necessary condition for an N × N variance-covariance matrix, Ω, to be internally consistent for all N × 1 vectors ω, where ω^T is the transpose of vector ω:

$$\omega^T \Omega \omega \geq 0$$

Variance and covariance rates are calculated using the same EWMA or GARCH model parameters to ensure that a positive-semidefinite model is constructed. For example, if a EWMA model uses $\lambda = 0.95$ for estimating variances, the same EWMA and λ should be used to estimate covariance rates.

When small changes are made to a small positive-semidefinite matrix such as a 3 × 3 matrix, the matrix will most likely remain positive-semidefinite. However, small changes to a large positive-semidefinite matrix such as 1,000 × 1,000 will most likely cause the matrix to no longer be positive-semidefinite.

An example of a variance-covariance matrix that is not internally consistent is shown as follows:

$$\begin{pmatrix} 1 & 0 & 0.8 \\ 0 & 1 & 0.8 \\ 0.8 & 0.8 & 1 \end{pmatrix}$$

Notice that the variances (i.e., diagonal of the matrix) are all equal to one. Therefore, the correlation for each pair of variables must equal the covariance for each pair of variables. This is true because the standard deviations are all equal to one. Thus, correlation is calculated as the covariance divided by one.

Also, notice that there is no correlation between the first and second variables. However, there is a strong correlation between the first and third variables as well as the second and third variables. This is very unusual to have one pair with no correlation while the other two pairs have high correlations. If we transpose a vector such that $\omega^T = (1, 1, -1)$, we would find that this variance-covariance matrix is not internally consistent since $\omega^T \Omega \omega \geq 0$ is not satisfied.

Another method for testing for consistency is to evaluate the following expression:

$$\rho_{12}^2 + \rho_{13}^2 + \rho_{23}^2 - 2\rho_{12}\rho_{13}\rho_{23} \leq 1$$

We can substitute data from the above variance-covariance matrix into this expression because all covariances are also correlation coefficients. When computing the formula, we would determine that the left side of the expression is actually greater than the right side, indicating that the matrix is not internally consistent.

$$0^2 + 0.8^2 + 0.8^2 - 2 \times 0 \times 0.8 \times 0.8 = 1.28$$
$$1.28 > 1$$

Generating Samples

LO 29.4: Describe the procedure of generating samples from a bivariate normal distribution.

Suppose there is a bivariate normal distribution with two variables, *X* and *Y*. Variable *X* is known and the value of variable *Y* is conditional on the value of variable *X*. If variables *X* and *Y* have a bivariate normal distribution, then the expected value of variable *Y* is normally distributed with a mean of:

$$\mu_Y + \rho_{XY} \times \sigma_Y \times \frac{X - \mu_X}{\sigma_X}$$

and a standard deviation of:

$$\sigma_Y \sqrt{1 - \rho_{XY}^2}$$

The means, μ_X and μ_Y, of variables *X* and *Y* are both unconditional means. The standard deviations of variables *X* and *Y* are both unconditional standard deviations. Also note that the expected value of *Y* is linearly dependent on the conditional value of *X*.

The following procedure is used to generate two sample sets of variables from a bivariate normal distribution.

Step 1: Independent samples Z_X and Z_Y are obtained from a univariate standardized normal distribution. Microsoft Excel® and other software programming languages have routines for sampling random observations from a normal distribution. For example, this is done in Excel with the formula = NORMSINV(RAND()).

Step 2: Samples ε_X and ε_Y are then generated. The first sample of *X* variables is the same as the random sample from a univariate standardized normal distribution, $\varepsilon_X = Z_X$.

Step 3: The conditional sample of *Y* variables is determined as follows:

$$\varepsilon_Y = \rho_{XY} Z_X + Z_Y \sqrt{1-\rho_{XY}^2}$$

where:
ρ_{XY} = correlation between variables *X* and *Y* in the bivariate normal distribution

Factor Models

LO 29.5: Describe properties of correlations between normally distributed variables when using a one-factor model.

A **factor model** can be used to define correlations between normally distributed variables. The following equation is a one-factor model where each U_i has a component dependent on one common factor (F) in addition to another component (Z_i) that is uncorrelated with other variables.

$$U_i = \alpha_i F + \sqrt{1-\alpha_i^2} Z_i$$

Between normally distributed variables, one-factor models are structured as follows:

- Every U_i has a standard normal distribution (mean = 0, standard deviation = 1).
- The constant α_i is between –1 and 1.
- F and Z_i have standard normal distributions and are uncorrelated with each other.
- Every Z_i is uncorrelated with each other.
- All correlations between U_i and U_j result from their dependence on a common factor, F.

There are two major advantages of the structure of one-factor models. First, the covariance matrix for a one-factor model is positive-semidefinite. Second, the number of correlations between variables is greatly reduced. Without assuming a one-factor model, the correlations of each variable must be computed. If there are N variables, this would require $[N \times (N - 1)] / 2$ calculations. However, the one-factor model only requires N estimates for correlations, where each of the N variables is correlated with one factor, F. The most well-known one factor model in finance is the *capital asset pricing model* (CAPM). Under the CAPM, each asset return has a systematic component (measured by beta) that is correlated with the market portfolio return. Each asset return also has a nonsystematic (or idiosyncratic) component that is independent of the return on other stocks and the market.

Copulas

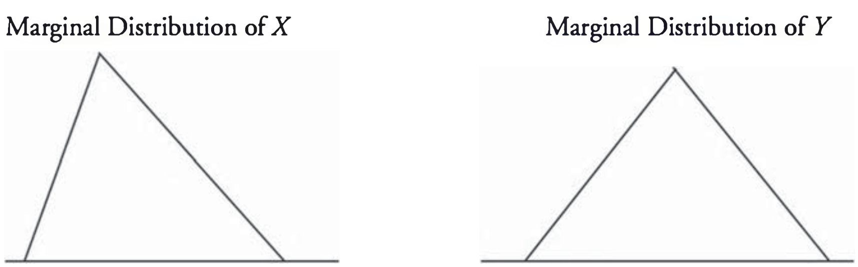
LO 29.6: Define copula and describe the key properties of copulas and copula correlation.

Suppose we have two **marginal distributions** of expected values for variables *X* and *Y*. The marginal distribution of variable *X* is its distribution with no knowledge of variable *Y*. The marginal distribution of variable Y is its distribution with no knowledge of variable X. If both distributions are normal, then we can assume the joint distribution of the variables is bivariate normal. However, if the marginal distributions are not normal, then a copula is necessary to define the correlation between these two variables.

A **copula** creates a joint probability distribution between two or more variables while maintaining their individual marginal distributions. This is accomplished by mapping the marginal distributions to a new known distribution. For example, a Gaussian copula (discussed in LO 29.8) maps the marginal distribution of each variable to the standard normal distribution, which, by definition, has a mean of zero and a standard deviation of one. The mapping of each variable to the new distribution is done based on percentiles.

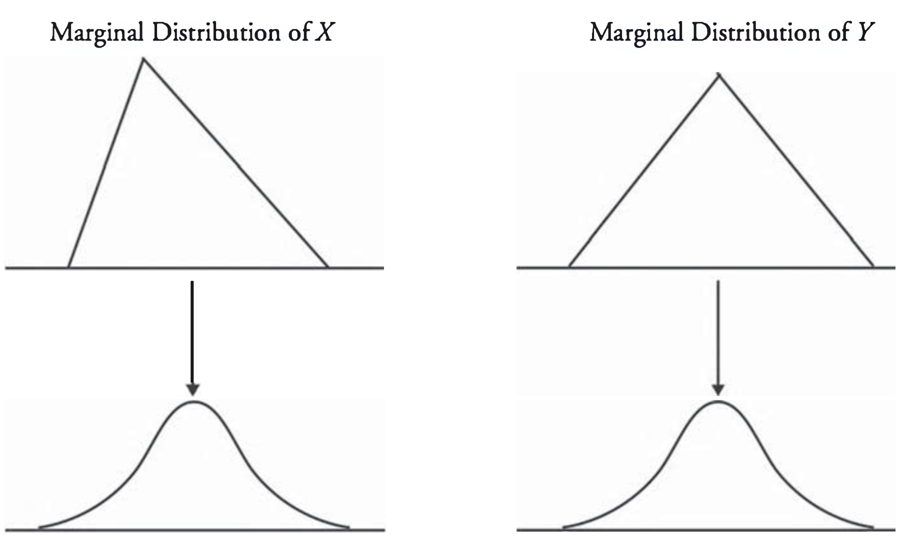
Suppose we have two triangular marginal distributions for two variables X and Y as illustrated in Figure 3.

Figure 3: Marginal Distributions



These two triangular marginal distributions for X and Y are preserved by mapping them to a known joint distribution. Figure 4 illustrates how a **copula correlation** is created.

Figure 4: Mapping Variables to Standard Normal Distributions



The key property of a copula correlation model is the *preservation of the original marginal distributions while defining a correlation between them*. A correlation copula is created by converting two distributions that may be unusual or have unique shapes and mapping them to known distributions with well-defined properties, such as the normal distribution. As mentioned, this is done by mapping on a percentile-to-percentile basis.

For example, the 5th percentile observation for the variable X marginal distribution is mapped to the 5th percentile point on the U_X standard normal distribution. The 5th percentile will have a value of –1.645. This is repeated for each observation on a percentile-to-percentile basis. The value that represents the 95th percentile of the X marginal distribution will have a value mapped to the 95th percentile of the U_X standard normal distribution and will have a value of +1.645. Likewise, every observation on the variable Y distribution is mapped to the corresponding percentile on the U_Y standard normal distribution. The new distribution is now a multivariate normal distribution.

Both U_X and U_Y are now normal distributions. If we make the assumption that the two distributions are joint bivariate normal distributions, then a correlation structure can be defined between the two variables. The triangular structures are not well-behaved structures. Therefore, it is difficult to define a relationship between the two variables. However, the normal distribution is a well-behaved distribution. Therefore, using a copula is a way to indirectly define a correlation structure between two variables when it is not possible to directly define correlation.

As mentioned, the correlation between U_X and U_Y is referred to as the copula correlation. The conditional mean of U_Y is linearly dependent on U_X, and the conditional standard deviation of U_Y is constant because the two distributions are bivariate normal.

For example, suppose the correlation between U_X and U_Y is 0.5. A partial table of the joint probability distribution between variables X and Y when the values of X and Y are 0.1, 0.2, and 0.3 is illustrated in Figure 5.

Figure 5: Partial Cumulative Joint Probability Distribution

	Variable Y		
Variable X	0.1	0.2	0.3
0.1	0.006	0.017	0.028
0.2	0.013	0.043	0.081
0.3	0.017	0.061	0.124

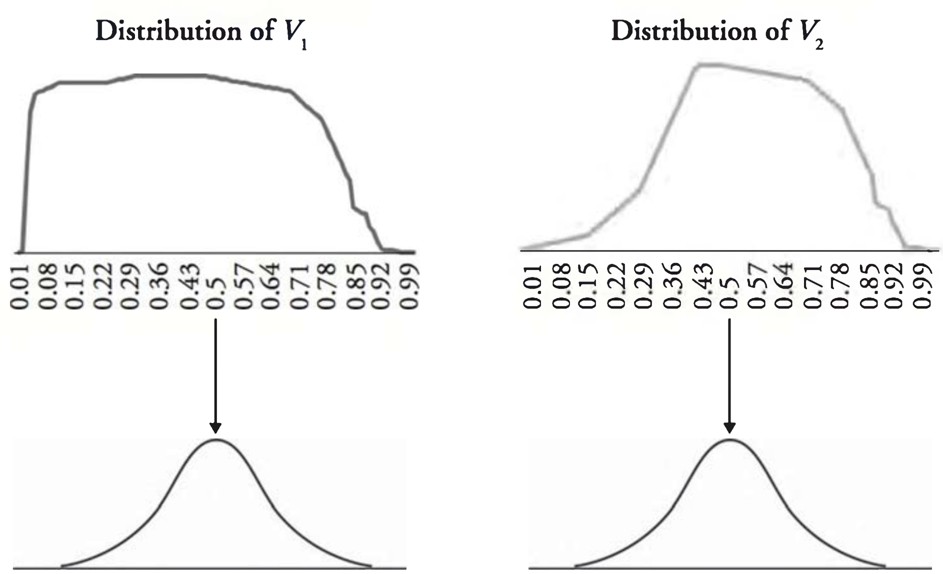
Now assume that the variable X under the original distribution had a value of 0.1 at the 5th percentile with a corresponding U_X value of –1.645. Also assume that the variable Y under the original distribution had a value of 0.1 with a corresponding value of –2.05. The joint probability that $U_X < -1.645$ and $U_Y < -2.05$ can be determined as 0.006 based on the row and column in Figure 5 that corresponds to a 0.1 value for both variables X and Y.

Types of Copulas

LO 29.8: Describe the Gaussian copula, Student's t-copula, multivariate copula, and one factor copula.

A **Gaussian copula** maps the marginal distribution of each variable to the standard normal distribution. The mapping of each variable to the new distribution is done based on percentiles. Figure 6 illustrates that V_1 and V_2 have unique marginal distributions. The observations of each distribution is mapped to the standard normal distribution on a percentile-to-percentile basis to create a Gaussian copula as follows:

Figure 6: Mapping Gaussian Copula to Standard Normal Distribution



Other types of copulas are created by mapping to other well-known distributions. The **Student's *t*-copula** is similar to the Gaussian copula. However, variables are mapped to distributions of U_1 and U_2 that have a bivariate Student's *t*-distribution rather than a normal distribution.

The following procedure is used to create a Student's *t*-copula assuming a bivariate Student's *t*-distribution with *f* degrees of freedom and correlation ρ.

Step 1: Obtain values of χ by sampling from the inverse chi-squared distribution with *f* degrees of freedom.

Step 2: Obtain values by sampling from a bivariate normal distribution with correlation ρ.

Step 3: Multiply $\sqrt{f/\chi}$ by the normally distributed samples.

A **multivariate copula** is used to define a correlation structure for more than two variables. Suppose the marginal distributions are known for *N* variables: $V_1, V_2, \ldots, V_N$. Distribution V_i for each *i* variable is mapped to a standard normal distribution, U_i. Thus, the correlation structure for all variables is now based on a multivariate normal distribution.

Factor copula models are often used to define the correlation structure in multivariate copula models. The nature of the dependence between the variables is impacted by the choice of the U_i distribution. The following equation defines a **one-factor copula** model where F and Z_i are standard normal distributions:

$$U_i = \alpha_i F + \sqrt{1 - \alpha_i^2} Z_i$$

The U_i distribution has a multivariate Student's t-distribution if Z_i and F are assumed to have a normal distribution and a Student's t-distribution, respectively. The choice of U_i determines the dependency of the U variables, which also defines the covariance copula for the V variables.

A practical example of how a one-factor copula model is used is in calculating the value at risk (VaR) for loan portfolios. A risk manager assumes a one-factor copula model maps the default probability distributions for different loans. The percentiles of the one-factor distribution are then used to determine the number of defaults for a large portfolio.

Tail Dependence

LO 29.7: Explain tail dependence.

There is greater **tail dependence** in a bivariate Student's t-distribution than a bivariate normal distribution. In other words, it is more common for two variables to have the same tail values at the same time using the bivariate Student's t-distribution. During a financial crisis or some other extreme market condition, it is common for assets to be highly correlated and exhibit large losses at the same time. This suggests that the Student's t-copula is better than a Gaussian copula in describing the correlation structure of assets that historically have extreme outliers in the distribution tails at the same time.

Key Concepts

LO 29.1

Correlation and covariance measure the strength between the linear relationship of two variables as follows:

$$\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y}$$

A correlation of zero between two variables does not imply that there is no dependence between the two variables.

LO 29.2

The formula for calculating a new covariance on day *n* using an exponentially weighted moving average (EWMA) model is:

$$\text{cov}_n = \lambda\,\text{cov}_{n-1} + (1-\lambda)X_{n-1}Y_{n-1}$$

GARCH(1,1) applies a weight of α to the most recent observation on covariance ($X_{n-1}Y_{n-1}$), a weight of β to the most recent covariance estimate (cov_{n-1}), and a weight of ω to the long-term average covariance rate as follows:

$$\text{cov}_n = \omega + \alpha X_{n-1}Y_{n-1} + \beta\,\text{cov}_{n-1}$$

LO 29.3

A matrix is positive-semidefinite if it is internally consistent. The following expression defines the necessary condition for an N × N variance-covariance matrix, Ω, to be internally consistent for all N × 1 vectors ω, where ω^T is the transpose of vector ω:

$$\omega^T \Omega \omega \geq 0$$

LO 29.4

Independent samples of two variables Z_X and Z_Y can be generated from a univariate standardized normal distribution. The conditional sample of *Y* variables for a bivariate normal distribution is then generated as:

$$\varepsilon_Y = \rho_{XY} Z_X + Z_Y \sqrt{1-\rho_{XY}^2}$$

LO 29.5

The covariance matrix for a one-factor model is positive-semidefinite. Also, the one-factor model only requires N estimates for correlations, where each of the N variables is correlated with one factor, F.

LO 29.6

A copula creates a joint probability distribution between two or more variables while maintaining their individual marginal distributions.

LO 29.7

The Student's t-copula is better than a Gaussian copula in describing the correlation structure of assets that historically have extreme outliers in tails at the same time.

LO 29.8

A Gaussian copula maps the marginal distribution of each variable to the standard normal distribution. The Student's t-copula maps variables to distributions of U_1 and U_2 that have a bivariate Student's t-distribution. The multivariate copula defines a correlation structure for three or more variables. The choice of U_i determines the dependency of the U variables in a one-factor copula model, which also defines the covariance copula for the V variables.

Concept Checkers

1. Suppose an analyst uses the EWMA model with $\lambda = 0.95$ to update correlation and covariance rates. The observed percentage change on day $n - 1$ for variables X and Y are 2.0% and 1.0%, respectively. The correlation estimate based on historical data for two variables X and Y on day $n - 1$ is 0.52. In addition, the estimated standard deviations on day $n - 1$ for variables X and Y are 1.4% and 1.8%, respectively. What is the new estimate of the correlation between X and Y on day n?
   A. 0.14.
   B. 0.42.
   C. 0.53.
   D. 0.68.

2. An equity analyst is concerned about satisfying the consistency condition for estimating new covariance rates. Which of the following procedures will most likely result in a positive-semidefinite matrix?
   A. The analyst uses an EWMA model with $\lambda = 0.95$ to update variances and a GARCH(1,1) model with $\lambda = 0.95$ to update the covariance rates for a 1,000 × 1,000 variance-covariance matrix.
   B. The analyst uses an EWMA model with $\lambda = 0.90$ to update variances and an EWMA model with $\lambda = 0.90$ to update the covariance rates for a 3 × 3 variance-covariance matrix.
   C. The analyst uses a GARCH(1,1) model with $\lambda = 0.95$ to update variances and a GARCH(1,1) model with $\lambda = 0.90$ to update the covariance rates for a 1,000 × 1,000 variance-covariance matrix.
   D. The analyst uses an EWMA model with $\lambda = 0.90$ to update variances and a GARCH(1,1) model with $\lambda = 0.90$ to update the covariance rates for a 3 × 3 variance-covariance matrix.

3. Suppose two samples, Z_X and Z_Y, are generated from a bivariate normal distribution. If variable Y is conditional on variable X, which of the following statements regarding these two samples is incorrect?
   A. The expected value of Y has a nonlinear relationship with all values of X.
   B. The mean and standard deviations for sample Z_X are unconditional.
   C. The value of variable Y is normally distributed.
   D. The conditional sample of Y variables is determined by:

$$\varepsilon_Y = \rho_{XY} Z_X + Z_Y \sqrt{1 - \rho_{XY}^2}\,.$$

4. Which of the following statements is most reflective of a characteristic of one-factor models between multivariate normally distributed variables? The one-factor model is shown as follows:

$$U_i = \alpha_i F + \sqrt{1-\alpha_i^2} Z_i$$

A. Each U_i has a component dependent on one common factor (F) in addition to another component (Z_i) that is uncorrelated with other variables.
B. F and Z_i must both have Student's t-distributions.
C. The covariance matrix for a one-factor model is not positive-semidefinite.
D. The number of calculations for estimating correlations is equal to $[N \times (N - 1)] / 2$.

5. Suppose a risk manager wishes to create a correlation copula to estimate the risk of loan defaults during a financial crisis. Which type of copula will most accurately measure tail risk?
A. Gaussian copula.
B. Student's t-copula.
C. Gaussian one-factor copula.
D. Standard normal copula.

Concept Checker Answers

1. C First, calculate the estimated covariance rate between variables X and Y on day $n - 1$ as:

$$\text{cov}(X,Y) = \rho_{X,Y} \times \sigma_X \sigma_Y = 0.52 \times 0.014 \times 0.018 = 0.00013$$

The EWMA model is then used to update the covariance rate for day n:

$$\text{cov}_n = 0.95 \times 0.00013 + 0.05 \times 0.02 \times 0.01 = 0.0001235 + 0.00001 = 0.0001335$$

The updated variance of X is:

$$\sigma^2_{X,n} = 0.95 \times 0.014^2 + 0.05 \times 0.02^2 = 0.0001862 + 0.00002 = 0.0002062$$

The new volatility measure of X is then:

$$\sigma^2_{X,n} = \sqrt{0.0002062} = 0.0143597$$

The updated variance for variable Y is:

$$\sigma^2_{Y,n} = 0.95 \times 0.018^2 + 0.05 \times 0.01^2 = 0.0003078 + 0.000005 = 0.0003128$$

The new volatility measure of Y is then:

$$\sigma^2_{Y,n} = \sqrt{0.0003128} = 0.01768615$$

The new correlation is found by dividing the new cov_n by the new standard deviations for X and Y as follows:

$$\frac{0.0001335}{0.0143597 \times 0.0176862} = 0.5257$$

2. B A matrix is positive-semidefinite if it is internally consistent. Variance and covariance rates must be calculated using the same EWMA or GARCH model and parameters to ensure that a positive-semidefinite model is constructed. For example, if an EWMA model is used with $\lambda = 0.90$ for estimating variances, the same EWMA model and λ should be used to estimate covariance rates.

3. A Both samples are normally distributed. The expected value of variable Y is normally distributed with a mean of:

$$\mu_Y + \rho_{XY} \times \sigma_Y \times \frac{X - \mu_X}{\sigma_X}$$

and a standard deviation of:

$$\sigma_Y \sqrt{1 - \rho^2_{XY}}$$

The expected value of Y is therefore linearly dependent on the conditional value of X.

4. A Each U_i has a component dependent on one common factor (F) in addition to another component (Z_i) that is uncorrelated with other variables. F and Z_i have standard normal distributions and are uncorrelated with each other. The covariance matrix for a one-factor model is positive-semidefinite and the one-factor model only requires N estimates for correlations, where each of the N variables is correlated with one factor, F.

5. B There is greater *tail dependence* in a bivariate Student's *t*-distribution than a bivariate normal distribution. This suggests that the Student's *t*-copula is better than a Gaussian copula in describing the correlation structure of assets that historically have extreme outliers in tails at the same time.

The following is a review of the Quantitative Analysis principles designed to address the learning objectives set forth by GARP®. This topic is also covered in:

Simulation Methods

Topic 30

Exam Focus

Simulation methods model uncertainty by generating random inputs that are assumed to follow an appropriate probability distribution. This topic discusses the basic steps for conducting a Monte Carlo simulation and compares this simulation method to the bootstrapping technique. For the exam, be able to explain ways to reduce Monte Carlo sampling error, including the use of antithetic and control variates. Also, understand the pseudo-random number generation method and the benefits of reusing sets of random number draws in Monte Carlo experiments. Finally, be able to describe the advantages and disadvantages of the bootstrapping technique in comparison to the traditional Monte Carlo approach.

Monte Carlo Simulation

LO 30.1: Describe the basic steps to conduct a Monte Carlo simulation.

Monte Carlo simulations are often used to model complex problems or to estimate variables when there are small sample sizes. A few practical finance applications of Monte Carlo simulations are: pricing exotic options, estimating the impact to financial markets of changes in macroeconomic variables, and examining capital requirements under stress-test scenarios.

There are four basic steps required to conduct a Monte Carlo simulation.

Step 1: Specify the data generating process (DGP)
Step 2: Estimate an unknown variable or parameter
Step 3: Save the estimate from step 2
Step 4: Go back to step 1 and repeat this process *N* times

The first step of conducting a simulation requires generating random inputs that are assumed to follow a specific probability distribution. The DGP could be a simple time series model or a more complex full structural model that requires multiple DGPs.

The second step of the simulation generates scenarios or trials based on randomly generated inputs drawn from a pre-specified probability distribution. The most common probability distribution used is the standard normal distribution. However, Student's *t* distribution is often used if the user believes it is a better fit for the data. A well-defined simulation model requires the generation of variables that follow appropriate probability distributions.

The last two steps in the simulation process allow for data analysis related to the properties of the probability distributions of the output variables. In other words, rather than making just one output estimate for a problem, the model generates a probability distribution of estimates. This provides the user with a better understanding of the range of possible outcomes. The quantity N in step four is the number of times the simulation is repeated. This is referred to as the number of replications or iterations and is typically 1,000 to 10,000 times depending on how costly it is to generate the sample size.

For example, suppose we are managing an investment portfolio and desire to estimate the ending capital in the portfolio in one year, C_1. The initial capital investment, C_0, is $100 invested in the Standard & Poor's 500 index (S&P 500). The return is a random variable that depends on how the market performs over the next year.

If we assume the return over the next year is equal to a historical mean return, we can calculate one point estimate of the ending capital based on the equation: $C_1 = C_0(1 + r)$. The return over the next period is a random variable, and a simulation model estimates multiple scenarios to represent future returns based on a probability distribution of possible outcomes. The output variable is an estimate of an ending amount of capital that is also a random variable. The simulation model allows us to visualize the output and analyze the probability distribution of the ending capital amounts generated by the model.

Reducing Monte Carlo Sampling Error

LO 30.2: Describe ways to reduce Monte Carlo sampling error.

The sampling variation for a Monte Carlo simulation is quantified as the standard error estimate. The standard error of the true expected value is computed as $s / \sqrt{N}$, where s is the standard deviation of the output variables and N is the number of scenarios or replications in the simulation. Based on this equation, it intuitively follows that in order to reduce the standard error estimate by a factor of 10, the analyst must increase N by a factor of 100. (Because the square root of 100 is 10, if we increase the sample size 100 times it will reduce the standard error estimate by dividing by 10.)

Suppose we continue the illustration from the previous example and run a simulation to estimate the ending capital amount for an initial investment portfolio of $100. The number of replications is initially 100 (i.e., N = 100), resulting in a mean ending capital of $110 and a standard deviation of $14.798. For this example, the standard error estimate is computed as $1.4798 (i.e., $14.798 / 10). Now, suppose we want to increase the accuracy by reducing the standard error estimate. How can we increase the accuracy of the simulation?

The accuracy of simulations depends on the standard deviation and the number of scenarios run. We cannot control the standard deviation, but we can control the number of replications. Assume we rerun the previous simulation with 400 replications that result in the same mean ending capital of $110, and the standard deviation remains at $14.798. The standard error estimate for the simulation with 400 replications is then $0.7399 (i.e., 14.798 / 20). With four times the number of scenarios (4 × N, or 400, in this example) the standard error estimate is cut in half to $0.7399. In other words, quadrupling the number of scenarios will improve the accuracy twofold.