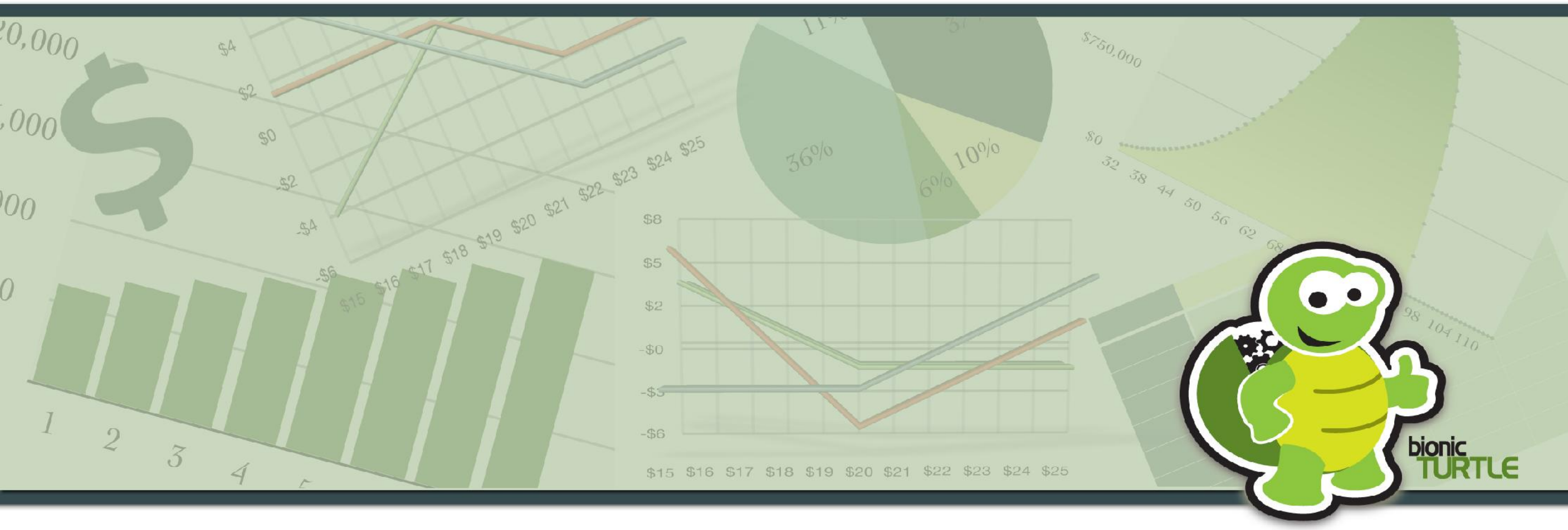


P1.T2. Quantitative Analysis

Chapter 4: Multivariate Random Variables

Bionic Turtle FRM Study Notes

By David Harper, CFA FRM CIPM
www.bionicturtle.com

Chapter 4: Multivariate Random Variables

Chapter 4: Multivariate Random Variables

- **Explain how a probability matrix can be used to express a probability mass function.**
- **Compute the marginal and conditional distributions of a discrete bivariate random variable.**
- **Explain how the expectation of a function is computed for a bivariate discrete random variable.**
- **Define covariance and explain what it measures.**
- **Explain the relationship between the covariance and correlation of two random variables, and how these are related to the independence of the two variables.**
- **Explain the effects of applying linear transformations on the covariance and correlation between two random variables.**
- **Compute the variance of a weighted sum of two random variables.**
- **Compute the conditional expectation of a component of a bivariate random variable.**
- **Describe the features of an iid sequence of random variables.**
- **Explain how the iid property is helpful in computing the mean and variance of a sum of iid random variables.**

Key concepts

- The interior cells of a **probability matrix** are joint probabilities that must sum to 100.0%. The exterior cells (which sum an interior row or column) are **unconditional (aka, marginal)** probabilities.
- Be able to manipulate the fundamental probability relationship: conditional equals joint divided multiplied by marginal; i.e., P(B|A) = P(B ∩ A) × P(A)
- Be very comfortable with the relationship between covariance and correlation: ρ(X,Y) = Cov(X,Y)/[σ(X) × σ(Y)]
- Correlation is a measure of linear relationship. Zero correlation does not imply independence, but independence does imply zero correlation.
- Memorize basic variance properties (https://en.wikipedia.org/wiki/Variance#Properties), including, for example: Variance(a + bX) = b^2 × Variance(X)
- Variance of two-asset portfolio return is given by $\boldsymbol{w^2\sigma_1^2 + w^2\sigma_2^2 + 2w(1-w)\sigma_{12}}$
- The **independent and identically distributed (iid)** property is an extremely common assumption, but it is often implicit. For example, when we use the square root rule (SRR) to scale volatility over time, we assume iid returns.

Explain how a probability matrix can be used to express a probability mass function.

A bivariate random variable is a **joint probability** of two random variables. The probability mass function (pmf) generalizes the univariate pmf:

$$f_{x_1,x_2}(x_1, x_2) = \Pr\,(X_1 = x_1, X_2 = x_2)$$

An example of a bivariate distribution is the *trinomial distribution* (itself a member of the multinomial distribution family) which generalizes the binomial distribution. The trinomial distribution assumes (n) independent trials; instead two outcomes (as is the case in the binomial), each trial produces one of three outcomes.[1]

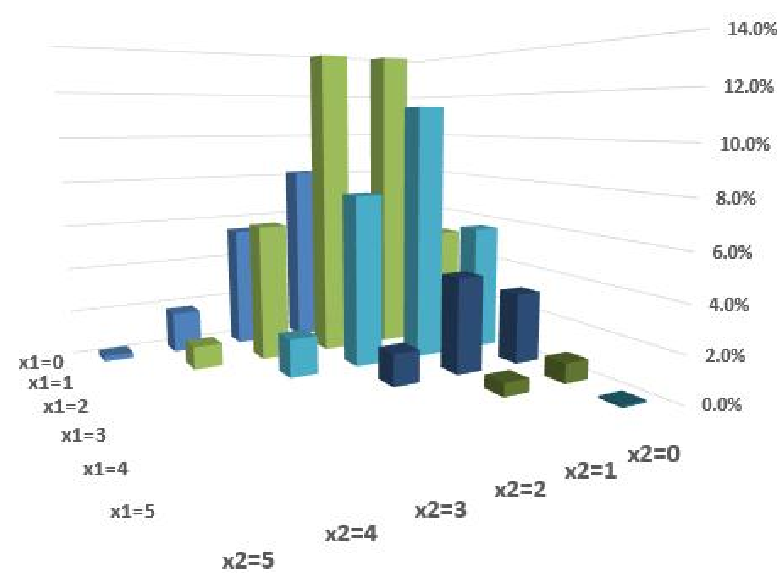
The trinomial PMF has three parameters: n, p_1 and p_2. As with the binomial, (n) is the number of trials; p_1 and p_2 are the probabilities, respectively, of observing the first and second outcomes. The third outcome is inferred: $p_3 = 1 - p_2 – p_1$

The PMF of a trinomial random variable is:

$$f_{x_1,x_2}(x_1, x_2) = \frac{n!}{x_1!\,x_2!\,(n - x_1 - x_2)!} p_1^{x_1} p_2^{x_2} (1 - p_1 - p_2)^{n - x_1 - x_2}$$

For example, when measuring the credit quality of a bond portfolio, each of (n) bonds can be classified as investment grade, speculative or unrated. In this case, X_1 is the count of the investment grade bonds and X_2 is the count of the high yield bonds.

The below figure shows the PMF of a trinomial with p_1 = 25%, p_2= 30%, p_3= 45%, and n=5. Notice that several combinations (in the chart) are not possible; e.g., x1 = 3 and x2 = 4.



[1] Example from GARP Chapter 4 (2020)

The CDF of a bivariate variable is a function returning the total probability that each component is less than or equal to a given value for that component, so that:

$$F_{X_1,X_2}(x_1,x_2) = \sum_{\substack{t_1 \epsilon R(X_1) \\ t_1 \le x_1}} \sum_{\substack{t_2 \epsilon R(X_2) \\ t_2 \le x_2}} f_{X_1,X_2}(t_1,t_2)$$

In this function, t_1 contains the values that X_1 may take as long as $t_1 \le X_1$, and t_2 is similarly defined only for X_2.

Probability matrix

The joint probability is the probability that the random variables (in this case, two random variables) take on certain values simultaneously. Thus, we refer to the probability of two events occurring together as their **joint probability**.

As Stock and Watson explain, “The joint probability distribution of two discrete random variables, say X and Y, is the probability that the random variables simultaneously take on certain values, say x and y. The probabilities of all possible (x, y) combinations sum to 1. The joint probability distribution can be written as the function P (X = x, Y = y).”[2]

When dealing with the joint probabilities of two variables, it is convenient to summarize the various probabilities in a **probability matrix** (a.k.a., probability table).

For example:[3] (Consider this same example for calculating joint, unconditional and conditional probabilities.) In Miller’s example, we assume a company that issues both **bonds** and **stock**. The bonds can either be downgraded, be upgraded, or have no change in rating. The stock can either outperform the market or underperform the market. This implies six different joint possibilities:

Probability Matrix: Interior cells are JOINT probabilities and sum to 100%. Exterior cells are unconditional probabilities

		Stock (S)		
		Out-perform (O)	Under-perform (U)	Uncond'l P[Bonds = i]
Bonds (B)	Upgrade	15.0%	5.0%	20.0%
	No Change	30.0%	25.0%	55.0%
	Downgrade	5.0%	20.0%	25.0%
Uncond'l P[Stock = i]		50.0%	50.0%	100.0%

[2] James Stock and Mark Watson, Introduction to Econometrics, 3rd Edition (Pearson Education, 2014)
[3] This example is borrowed from Michael Miller, Mathematics and Statistics for Financial Risk Management, 2nd Edition (Hoboken, NJ: John Wiley & Sons, 2013). But the exhibit was hand constructed by David Harper and available in our Learning Spreadsheets.