Random variables and expectations

Daniele Bianchi[1]

whitesphd.com

[1]School of Economics and Finance
Queen Mary, University of London

Summary

This week we cover the basics of probability theory as it is applied to both discrete and continuous random variables. The material covered this week serves as a foundation for the rest of the module. We start by discussing the properties of univariate random variables. The attention then turns to the discussion of multivariate random variables and their difference with respect to univariate random variables. Then we will focus on describing the concept of expectations operator and moments.

Contents

Univariate random variables

What is a random variable (RV)

Definition (Probability space)

A probability space is denoted using the tuple $(\Omega, \mathcal{F}, \mathrm{Pr})$ where Ω is the sample space, $\mathcal{F}$ is the event space and Pr is the probability set function which has a given domain $\omega \in \mathcal{F}$.

Definition (Random variable)

Let $(\Omega, \mathcal{F}, \mathrm{Pr})$ be a probability space. IF $X : \Omega \to \mathbb{R}$ is a real-valued function have its domain elements of Ω, then X is called a random variable (RV).

- A RV is essentially a function that takes a value $\omega \in \Omega$ as an input and returns a value $x \in \mathbb{R}$ as an output, where $\mathbb{R}$ is the symbol the set of real numbers.
- We are going to cover two main forms of RVs: discrete and continuous.

Discrete random variables

Definition (Discrete random variable)

A RV is called **discrete** if its range, i.e., the values for which it is defined, consists of countable (possibly infinite) number of elements.

- Example 1: A RV that takes on values in $\{0, 1\}$ is known as a **Bernoulli** RV.
  - Bernoulli random variables are often used to model "success" vs "failure", where failure can be defined, for e.g., as a negative return, a recession and/or a corporate default.
- Example 2: A RV that takes values in $\{0, 1, 2, 3, \ldots\}$ is known as a **Poisson** RV.
  - Poisson random variables are often used to model the number of occurrences of an event in an interval, e.g., how many Covid cases in a month.

Discrete random variables

Definition (Probability mass function)

The probability mass function f for a discrete random variable X is defined as $f(x) = \Pr(x)$ for all $x \in R(X)$, and $f(x) = 0$ for all $x \notin R(X)$ where $R(X)$ is the range of X.

Example 1: The probability mass function of a Bernoulli RV takes the form

$$f(x, p) = p^x (1 - p)^{1-x},$$

where $p \in [0, 1]$ is the probability of success.

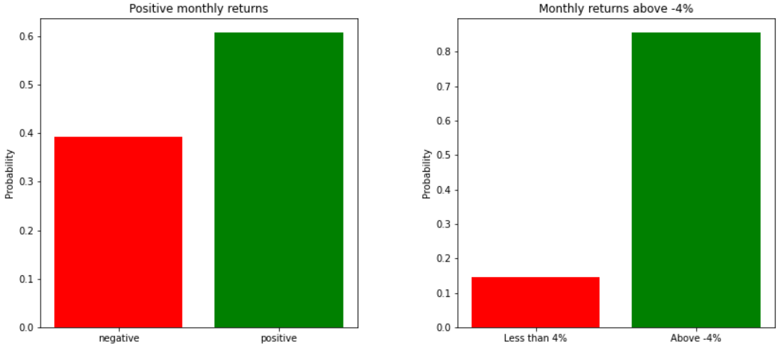
Example 2: The probability mass function of a Poisson random variable is

$$f(x, \lambda) = \frac{\lambda^x}{x!} \exp(-\lambda),$$

where $\lambda \in [0, \infty)$ determines the intensity of arrival.

"Success" in the stock market

Figure 1: Market returns as a Bernoulli



These two charts show examples of Bernoulli random variables using returns on the NYSE/AMEX stock market. The left panel shows the "success" as a positive return, whereas the right panel shows the "success" as a return above 4% on a monthly basis.

Continuous random variables

Definition (Continuous random variable)

A RV is called **continuous** if its range is uncountably infinite and there exists a non-negative-valued function $f(x)$ defined over $x \in (-\infty, \infty)$ such that for any event $B \subset R(X)$, $\Pr(B) = \int_{x \in B} f(x)dx$ and $f(x) = 0$ for all $x \notin R(X)$ where $R(X)$ is the range of X.

- The Probability Mass Function (PMF) of a discrete random variable is replaced with a Probability Density Function (PDF) for continuous random variables.
- This change in naming reflects that the probability of a single point of a continuous random variable is 0, although the probability of observing a value within a small interval in $R(X)$ is not.

Continuous random variables

Definition (Probability density function)

A function $f : \mathbb{R} \to \mathbb{R}$ is a member of the class of continuous density functions is an only if $f(x) \geq 0$ for all $x \in (-\infty, \infty)$ and $\int_{-\infty}^{\infty} f(x)dx = 1$.

Two essential properties:

- The function is non-negative (it is a probability at the end)
- The function needs to integrate, i.e., sum up, to 1.

Example: Perhaps the most used continuous RV is the Normal, which has a PDF of the form

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),$$

with $N\left(\mu, \sigma^2\right)$ used as short-hand notation.

Continuous random variables

A closely related function to the PDF is the cumulative distribution function (CDF), which returns the total probability of observing a value of the RV *less* than its input.

Definition (Cumulative distribution function)

The cumulative distribution function (CDF) for a random variable X is defined as $F(c) = Pr\,(x \leq c)$ for all $c \in (-\infty, \infty)$.

When X is a discrete random variable, the CDF is

$$F\,(X) = \sum_{s \leq x} f(s), \qquad \text{for } x \in (-\infty, \infty),$$

- Example: The CDF of a Bernoulli is

$$F\,(x, p) = \begin{bmatrix} 0 & \text{if} & x < 0 \\ p & \text{if} & 0 \leq x < 1 \\ 1 & \text{if} & x \geq 1 \end{bmatrix}$$

Continuous random variables

When X is a continuous random variable, the CDF is

$$F(X) = \int_{-\infty}^{x} f(s)ds \qquad \text{for } x \in (-\infty, \infty),$$

- Example: The CDF of a Normal $N\left(\mu, \sigma^2\right)$ is given by

$$F(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} \exp\left(-\frac{(s-\mu)^2}{2\sigma^2}\right) ds,$$

The relationship between the PDF and the CDF is rather simple. In fact, the density function of X can be defined as $f(x) = \partial F(x)/\partial x$ whenever $f(x)$ is continuous and $f(x) = 0$ elsewhere.

Quantile functions

Before defining quantile functions, it is necessary to define a quantile.

Definition (Quantile)

Any number q satisfying $\Pr(x \leq q) = \alpha$ and $\Pr(x \geq q) = 1 - \alpha$ is known as the α-quantile of X and is denoted q_α.

Notice:

- The quantile function is closely related to the CDF.
- The quantile function is the inverse (function) of the CDF.

N.B.: Strictly speaking, a **quantile** is just the point on the CDF where the total probability that a random variable is smaller than q is α and the probability that a random variable takes a larger value than q is $1 - \alpha$.

Quantile functions

Definition (Quantile function)

Let X be a continuous RV with CDF $F(X)$. The quantile function for X is defined as $G(\alpha) = q$ where $\Pr(x \leq q)$ and $\Pr(x > q) = 1 - \alpha$

Notice that:

- When $F(X)$ is one-to-one, i.e., X is strictly continuous, then $G(\alpha) = F^{-1}(\alpha)$.
- When the PDF does not contain any region with zero probability, the quantile function is simply the inverse of the CDF.
- The quantile function plays an important role in simulating RVs; for instance, is $u \sim U(0,1)$, then $x = F^{-1}(u)$ is distributed as F.
  - E.g., if $u \sim U(0,1)$ and $F^{-1}(\alpha)$ is the quantile function of a standard Normal, i.e., $N(0,1)$, then $x = F^{-1}(u)$ follows a $N(0,1)$.

Common distributions

Bernoulli distribution

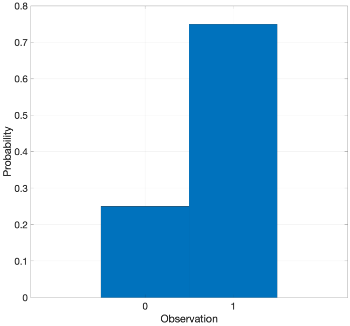
Bernoulli: A Bernoulli RV is a discrete random variable which takes one of two values, 0 or 1. The Bernoulli is often used to model success vs failure, loosely defined.

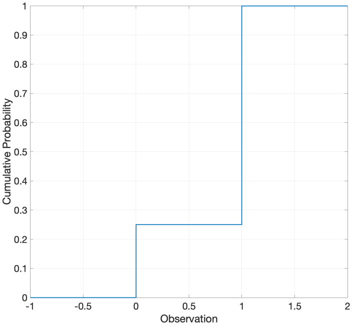
- Parameters $p \in [0, 1]$
- Support $x \in \{0, 1\}$
- Probability mass function

$$f(x, p) = p^x (1 - p)^{1-x} \qquad p \geq 0,$$

- Moments:
  - Mean p
  - Variance $p(1 - p)$

Figure 2: Bernoulli PDF and CDF





This figure plots PDF (left panel) and the CDF (right panel) of a Bernoulli distribution with $p = 0.75$.

Poisson distribution

Poisson: A Poisson RV is a discrete random variable taking values in $\{0, 1, \ldots\}$. Poisson random variables are often used to model counts of events during some interval, e.g., the number of trades within a day.

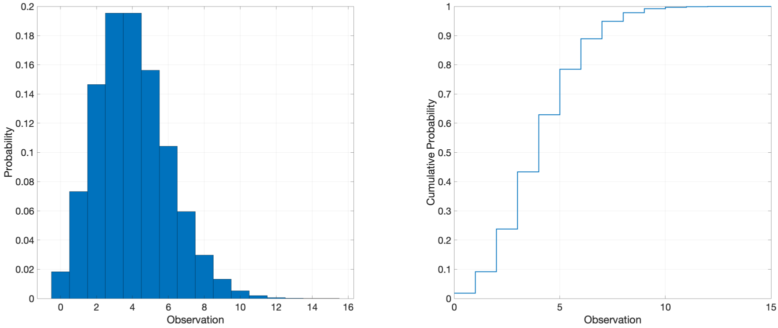
- Parameters $\lambda \geq 0$
- Support $x \in \{0, 1, \ldots\}$
- Probability mass function

$$f(x, \lambda) = \frac{\lambda^x}{x!} \exp(-\lambda),$$

- Moments:
  - Mean λ
  - Variance λ

Poisson distribution

Figure 3: Poisson PDF and CDF



This figure plots PDF (left panel) and the CDF (right panel) of a Poisson distribution with $\lambda = 4$.

Normal (Gaussian) distribution

Normal (Gaussian): The normal is the most used distribution in financial economics. It is the familiar "bell-shaped" distribution and it is heavily used in hypothesis testing and in modeling asset returns.

- Parameters $\mu \in (-\infty, \infty), \sigma^2 \geq 0$
- Support $x \in (-\infty, \infty)$
- Probability mass function

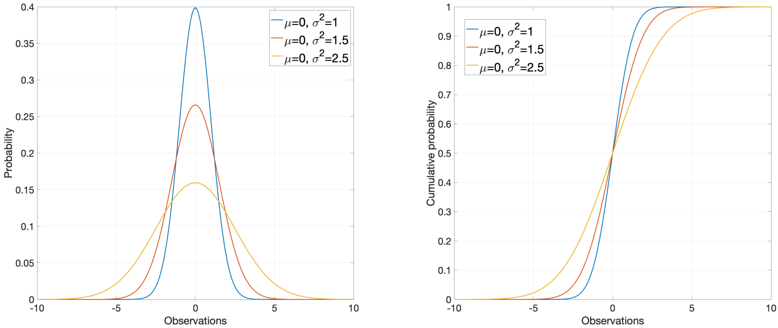
$$f(x, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),$$

- Moments:
  - Mean μ
  - Variance σ^2
  - Skewness 0
  - Kurtosis 3

N.B., A normally distributed random variable with $\mu = 0$ and $\sigma^2 = 1$ is known as a **standard normal**.

Normal (Gaussian) distribution

Figure 4: Normal PDF and CDF



This figure plots PDF (left panel) and the CDF (right panel) of a Normal distribution with $\mu = 0$ and $\sigma^2 = 1, 1.5, 2.5$.

Student's t distribution

Student's t: The Student's t RVs are also commonly encountered in hypothesis testing and are closely related to standard normals. Student's t are similar to normals except that they are heavy tailed, although as $\nu \to \infty$ a Student's t converges to a standard normal.

- Parameters $\nu > 1$
- Support $x \in (-\infty, \infty)$,
- Probability mass function

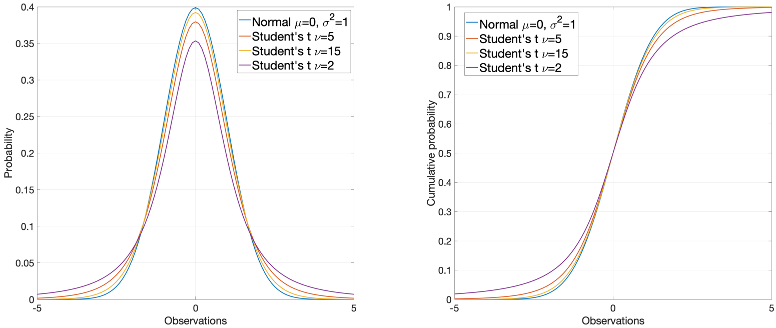
$$f(x, \mu, \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi\Gamma\left(\frac{\nu}{2}\right)}}\left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

- Moments:
  - Mean 0, $\nu > 1$.
  - Median 0
  - Variance $\frac{\nu}{\nu-2}, \nu > 2$.
  - Skewness 0, $\nu > 3$.
  - Kurtosis $3\frac{(\nu-2)}{\nu-4}$, $\nu > 4$.

N.B., When $\nu = 1$ a Student's t is known as a Cauchy RV.

Student's t distribution

Figure 5: Student's t PDF and CDF



This figure plots PDF (left panel) and the CDF (right panel) of a Student's t distribution with $\nu = 2, 5, 15$.

Expectations and moments

Expectations

The expectation is the value, on average, of a random variable (or a function of a random variable).

Definition (Expectation of a discrete RV)

The expectation of a discrete RV, defined $E[X] = \sum_{x \in R(X)} x f(x)$, exists if and only if $\sum_{x \in R(X)} |x| f(x) < \infty$

- When the range of X is finite then the expectation always exists.
- When the range is infinite, the PDF must be sufficiently small for large values in order for the expectation to exist.

Definition (Expectation of a continuous RV)

The expectation of a discrete RV, defined $E[X] = \int_{-\infty}^{\infty} x f(x) dx$, exists if and only if $\sum_{-\infty}^{\infty} |x| f(x) dx < \infty$

Expectations

The expectations operator has a number of simple and useful properties:

- If c is a constant, then $E[c] = c$.
- If c is a constant, then $E[cX] = cE[X]$.
- The expectation of the sum is the sum of the expectations,

$$E\left[\sum_{i=1}^{k} g_i\left(X\right)\right] = \sum_{i=1}^{k} E\left[g_i\left(X\right)\right],$$

- If a is a constant, then $E[a + X] = a + E[X]$.
- $E\left[f\left(X\right)\right] = f\left(E\left[X\right]\right)$ when $f\left(X\right)$ is affine, i.e., $f\left(x\right) = a + bx$.
- $E\left[X^p\right] \neq E\left[X\right]^p$ except when $p = 1$.

Moments

Moments are expectations of particular functions of a random variable, typically $g(x) = x^s$ for $s = 1, 2, \ldots$.

Definition (Noncentral moment)

The r^{th} noncentral moment of a continuous random variable X is

$$\mu_r^{nc} = E\left[X^r\right] = \int_{-\infty}^{\infty} x^r f(x)dx, \qquad \text{for } r = 1, 2, \ldots,$$

N.B., The 1st non-central moment of a random variable X is the mean μ

Definition (Central moment)

The r^{th} central moment of a RV X is defined

$$\mu_r = E\left[(X-\mu)^r\right] = \int_{-\infty}^{\infty} (x-\mu)^r f(x)dx, \qquad \text{for } r = 2, 3, \ldots,$$

Variance

The second central moment of a RV X, $E\left[(X-\mu)^2\right]$ is called the variance and is denoted σ^2, or $V[X]$. The variance operator has a number of useful properties:

- If c is a constant, then $V[c]=0$.
- If c is a constant, then $V[cX]=c^2V[X]$.
- If a is a constant, then $V[a+X]=V[X]$.
- The variance of the sum is the sum of the variances plus twice all of the covariances.

$$V\left[\sum_{i=1}^{n}X_i\right]=\sum_{i=1}^{n}V[X_i]+2\sum_{j=1}^{n}\sum_{k=j+1}^{n}Cov[X_j,X_k],$$

- The square root of the variance, known as standard deviation (or volatility) is typically more useful.

Skewness

Definition (Skewness)

The 3rd central moment, standardised by the 2nd central moment raised to the power 3/2,

$$\frac{\mu_3}{\left(\sigma^2\right)^{3/2}} = \frac{E\left[\left(X - E\left[X\right]\right)^3\right]}{E\left[\left(X - E\left[X\right]\right)^2\right]^{3/2}} = E\left[Z^3\right]$$

where $Z = \left(x - \mu\right)/\sigma$ the standardised version of x.

- The skewness is a general measure of **asymmetry**, and is 0 for symmetric distributions.

Kurtosis

Definition (Kurtosis)

The 4th central moment, standardised by the squared 2nd central moment,

$$\frac{\mu_4}{\left(\sigma^2\right)^2} = \frac{E\left[\left(X - E\left[X\right]\right)^4\right]}{E\left[\left(X - E\left[X\right]\right)^2\right]^2} = E\left[Z^4\right]$$

where $Z = \left(x - \mu\right)/\sigma$ the standardised version of x.

- The kurtosis measures the chance of observing a large value.
- It is often expressed as **excess kurtosis**, which is the $E\left[Z^4\right] - 3$, where 3 is the kurtosis of a normal random variable.
- RVs with positive excess kurtosis as often referred to as **heavy-tailed**.

Related measures

Definition (Median)

Any number m satisfying $\Pr(X \leq m) = 0.5$ and $\Pr(X \geq m) = 0.5$ is known as the median of X.

- The median measures the point where 50% of the distribution lies on either side.
- The median has few advantages over the mean, in particular it is not affected by outliers.

Definition (Interquartile range)

The value $q_{.75} - q_{.25}$ is known as the interquartile range, with q_i being the ith quantile.

Exercises

Exercises

Problem 1.1: Suppose

$$\begin{bmatrix} X \\ U \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_X \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_X^2 & 0 \\ 0 & \sigma_U^2 \end{bmatrix}\right),$$

and $Y = 2X + U$. What is $E(Y)$ and $V(Y)$?

Problem 1.2: Show that

$$Cov[aX + bY, cX + dY] = acV(X) + bdV(Y) + (ad + bc)Cov\left[X, Y\right],$$

Problem 1.3: Suppose $\{X_i\}$ is a sequence of random variables where $V\left[X_i\right] = \sigma^2$ for all i, $Cov\left[X_i, X_{i-1}\right] = \theta$ and $Cov\left[X_i, X_{i-j}\right] = 0$ for $j > 1$. What is $V\left[\overline{X}\right]$ where $\overline{X} = n^{-1}\sum_{i=1}^{n} X_i$?

Problem 1.4: Suppose $Y = X\beta + \epsilon$ where $X \sim N\left(\mu_X, \sigma_X^2\right), \epsilon \sim N\left(0, \sigma^2\right)$ and X and ϵ are independent. What is $\mathrm{Corr}\left[X, Y\right]$?

Problem 1.5: Prove that $E[a+bX] = a + bE[X]$.

Problem 1.6: Prove that $V[a+bX] = b^2 V[X]$.

Problem 1.7: Prove that $Cov[a+bX, c+dY] = bdCov[X,Y]$.

Problem 1.8: Prove that
$V[a+bX+cY] = b^2V[X] + c^2V[Y] + 2bcCov[X,Y]$.

Problem 1.9: Suppose $\{X_i\}$ is a sequence of RVs which are independent to each other and have the same variance σ^2. Show that

$$V\left[\overline{X}\right] = V\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sigma^2,$$

Problem 1.10: Suppose $F(X) = 1 - p^{x+1}$ for $x \in [0, 1, 2, \ldots]$ and $p \in (0,1)$. Find the Probability Mass Function.

Exercises

Problem 1.11: A fund manager tells you that her fund has non-linear as a function of the market and that his return is $r_{it} = 0.02 + 2r_{mt} - 0.5r_{mt}^2$ where r_{it} is the return on the fund and r_{mt} is the return on the market. She tells you her expectation of the market return this year is 10%, and that her fund will have an expected return of 22%. Can this be?

Problem 1.12: An investor can invest in stocks or bonds which have expected returns and covariances as

$$\mu = \begin{bmatrix} 0.10 \\ 0.03 \end{bmatrix}, \Sigma = \begin{bmatrix} 0.04 & -0.003 \\ -0.003 & 0.0009 \end{bmatrix},$$

where stocks are the first component.

- Suppose the investor has Â£1,000 to invest and splits the investment evenly. What is the expected return, standard deviation and Sharpe ratio (μ/σ) for the investment?