Random Variables

DEFINITION 3.1

> A **random variable** is a function of an outcome,
>
> $$X = f(\omega).$$
>
> In other words, it is a quantity that depends on chance.

If X is a discrete variable,

DEFINITION 3.2

> Collection of all the probabilities related to X is the **distribution** of X. The function
>
> $$P(x) = \boldsymbol{P}\{X = x\}$$
>
> is the **probability mass function**, or **pmf**. The **cumulative distribution function**, or **cdf** is defined as
>
> $$F(x) = \boldsymbol{P}\{X \le x\} = \sum_{y \le x} \boldsymbol{P}(y). \tag{3.1}$$
>
> The set of possible values of X is called the **support** of the distribution F.

For continuous distributions we have a probability density function rather than a pmf

Probability density function

$$f(x) = F'(x)$$

$$\boldsymbol{P}\{a < X < b\} = \int_a^b f(x)dx$$

Distribution	**Discrete**	**Continuous**
Definition	$P(x) = P\{X = x\}$ (pmf)	$f(x) = F'(x)$ (pdf)
Computing probabilities	$\boldsymbol{P}\{X \in A\} = \sum_{x \in A} P(x)$	$\boldsymbol{P}\{X \in A\} = \int_A f(x)dx$
Cumulative distribution function	$F(x) = \boldsymbol{P}\{X \le x\} = \sum_{y \le x} P(y)$	$F(x) = \boldsymbol{P}\{X \le x\} = \int_{-\infty}^{x} f(y)dy$
Total probability	$\sum_x P(x) = 1$	$\int_{-\infty}^{\infty} f(x)dx = 1$

Joint Distributions

DEFINITION 3.3

If X and Y are random variables, then the pair (X, Y) is a **random vector**. Its distribution is called the **joint distribution** of X and Y. Individual distributions of X and Y are then called the **marginal distributions**.

Joint Distributions for Continuous Variables

DEFINITION 4.2

For a vector of random variables, the **joint cumulative distribution function** is defined as

$$F_{(X,Y)}(x, y) = \boldsymbol{P}\{X \le x \ \cap Y \le y\}.$$

The **joint density** is the *mixed derivative* of the joint cdf,

$$f_{(X,Y)}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{(X,Y)}(x, y).$$

The following table shows how to calculate marginal distributions, check for independence and how to calculate probabilities of random vectors using joint probability distributions

Distribution	Discrete	Continuous
Marginal distributions	$P(x) = \sum_y P(x, y)$ $P(y) = \sum_x P(x, y)$	$f(x) = \int f(x, y)dy$ $f(y) = \int f(x, y)dx$
Independence	$P(x, y) = P(x)P(y)$	$f(x, y) = f(x)f(y)$
Computing probabilities	$\boldsymbol{P}\{(X, Y) \in A\}$ $= \sum\sum_{(x,y)\in A} P(x, y)$	$\boldsymbol{P}\{(X, Y) \in A\}$ $= \iint_{(x,y)\in A} f(x, y)\, dx\, dy$

Properties (or Moments) of a Distribution

Discrete	Continuous
$\mathbf{E}(X) = \sum_x xP(x)$	$\mathbf{E}(X) = \int xf(x)dx$
$\mathrm{Var}(X) = \mathbf{E}(X-\mu)^2 = \sum_x (x-\mu)^2 P(x) = \sum_x x^2P(x) - \mu^2$	$\mathrm{Var}(X) = \mathbf{E}(X-\mu)^2 = \int (x-\mu)^2 f(x)dx = \int x^2 f(x)dx - \mu^2$
$\mathrm{Cov}(X,Y) = \mathbf{E}(X-\mu_X)(Y-\mu_Y) = \sum_x \sum_y (x-\mu_X)(y-\mu_Y)P(x,y) = \sum_x \sum_y (xy)P(x,y) - \mu_x\mu_y$	$\mathrm{Cov}(X,Y) = \mathbf{E}(X-\mu_X)(Y-\mu_Y) = \iint (x-\mu_X)(y-\mu_Y)f(x,y)\,dx\,dy = \iint (xy)f(x,y)\,dx\,dy - \mu_x\mu_y$

Std(X) is square root of Var(X)

DEFINITION 3.9

Correlation coefficient between variables X and Y is defined as

$$\rho = \frac{\mathrm{Cov}(X,Y)}{(\mathrm{Std}X)(\mathrm{Std}Y)}$$

Properties of expectations

$$\mathbf{E}(aX + bY + c) = a\,\mathbf{E}(X) + b\,\mathbf{E}(Y) + c$$

In particular,

$$\begin{aligned} \mathbf{E}(X+Y) &= \mathbf{E}(X) + \mathbf{E}(Y) \\ \mathbf{E}(aX) &= a\,\mathbf{E}(X) \\ \mathbf{E}(c) &= c \end{aligned}$$

For **independent** X and Y,

$$\mathbf{E}(XY) = \mathbf{E}(X)\,\mathbf{E}(Y)$$

Properties of variances and covariances

$$\mathrm{Var}(aX + bY + c) = a^2\,\mathrm{Var}(X) + b^2\,\mathrm{Var}(Y) + 2ab\,\mathrm{Cov}(X,Y)$$

$$\begin{aligned}&\mathrm{Cov}(aX + bY, cZ + dW)\\&\quad = ac\,\mathrm{Cov}(X,Z) + ad\,\mathrm{Cov}(X,W) + bc\,\mathrm{Cov}(Y,Z) + bd\,\mathrm{Cov}(Y,W)\end{aligned}$$

$$\begin{aligned}\mathrm{Cov}(X,Y) &= \mathrm{Cov}(Y,X)\\ \rho(X,Y) &= \rho(Y,X)\end{aligned}$$

In particular,

$$\begin{aligned}\mathrm{Var}(aX + b) &= a^2\,\mathrm{Var}(X)\\ \mathrm{Cov}(aX + b, cY + d) &= ac\,\mathrm{Cov}(X,Y)\\ \rho(aX + b, cY + d) &= \rho(X,Y)\end{aligned}$$

For independent X and Y,

$$\begin{aligned}\mathrm{Cov}(X,Y) &= 0\\ \mathrm{Var}(X + Y) &= \mathrm{Var}(X) + \mathrm{Var}(Y)\end{aligned}$$

Only a large variance will allow a Variable to vary greatly from the expected value. This is shown by Chebyshev's Inequality

Chebyshev's inequality

$$\boldsymbol{P}\left\{|X - \mu| > \varepsilon\right\} \le \left(\frac{\sigma}{\varepsilon}\right)^2$$

for any distribution with expectation μ and variance σ^2 and for any positive ε.

This shows that a Variable with a high variance has bigger risk of varying from the expected amount by a large value.