Chapter 3

Discrete random variables

In probability and statistics, a random variable, is described informally as a variable whose values depend on outcomes of a random phenomenon. A discrete random variable X is a pair of values x_i and probabilities $p_i \equiv P(X = x_i)$, for all $i = 1, 2, ...n$, where n is finite integer. We call ***probability distribution*** the list of probabilities $\{p_i\}_{i=1}^n$ associated to outcomes $\{x_i\}_{i=1}^n$

Linear combination Suppose that X and Y are two discrete random variables with the same probability distribution $\{p_i\}_{i=1}^n$. Let a and b be real numbers. The random variable $aX + bY$ is called a linear combination of X and Y with coefficients a and b.

3.1. Probability distributions

Joint probability function Let X and Y be two random variables. A joint probability function $P(x, y)$ is used to express the probability that X takes the specific value x_i and simultaneously Y takes the value y, i.e. $P(x, y) = P(X = x, Y = y)$. By definition of probability, it must be that $\sum_x \sum_y P(x, y) = 1$.

Cumulative probability function Given a random variable X with probability distribution $P(x)$, the cumulative distribution function of x, $F(x)$ is defined as follows as the probability that X is less than or equal to x, i.e.

$$F(x) = P(X \leq x)$$

By construction, $0 \leq F(x) \leq 1$.
Consider now a pair of random variable X and Y with joint probability function $P(x, y)$. Analogously, the cumulative distribution function for this pair of random variables is defined in terms of their joint probability distribution:

$$F(x, y) = P(X \leq x, Y \leq y)$$

Marginal probability function Consider a pair of random variable X and Y with joint probability function $P(x, y)$. Their marginal probabilities are defined as

$$P(x) = \sum_y P(X \leq x, Y \leq y)$$

and

$$P(y) = \sum_x P(X \leq x, Y \leq y)$$

The marginal probability of X (Y) gives the probability of x (y) irrespective of Y (X).

3.2. Moments

Expected value The expected value of a discrete random variable is equal to

$$E[X] = \sum_{i=1}^{n} P(X = x_i)x_i$$

Notice that for any random variables, X and Y, with the same probability distribution $\{p_i\}_{i=1}^n$, and any numbers, a and b one has

$$E[aX + bY] = aE[X] + bE[Y]$$

Proof $E[aX + bY] = \sum_{i=1}^n (ax_i + by_i)p_i = \sum_{i=1}^n ax_i p_i + \sum_{i=1}^n by_i p_i = a\sum_{i=1}^n x_i p_i + b\sum_{i=1}^n y_i p_i = aE[X] + bE[Y]$.

If a random variable X is such that there exit an event x_i with $p(X = x_i) = 1$ and $p(X = x_j) = 0$ for all $j \neq i$, then $E(X) = x_i$.

Notice finally that in general, if $g(X)$ is some function of the random variable X then $E[g(X)]$ and $g[E(x)]$ are not necessarily equal.

Variance The variance of a discrete random variable is given by:

$$VAR[X] = \sum_{i=1}^{n} P(X = x_i)(x_i - E[X])^2$$

The standard deviation is the positive square root of the variance.

Covariance Let X and Y be two discrete random variables, the covariance between X and Y is given by:

$$COV[X,Y] = \sum_{i=1}^{n}\sum_{j=1}^{n} P(X = x_i, Y = y_j)(x_i - E[X])(y_i - E[Y])$$

This expression is equivalent to:

$$COV[X,Y] = \sum_{i=1}^{n}\sum_{j=1}^{n} P(X = x_i, Y = y_j)(x_i y_i - x_i E[Y] - E[X]y_i + E[X]E[Y]) =$$
$$\sum_{i=1}^{n}\sum_{j=1}^{n} P(x_i, y_j)x_i y_i - \sum_{i=1}^{n}\sum_{j=1}^{n} P(x_i, y_j)x_i E[Y] - \sum_{i=1}^{n}\sum_{j=1}^{n} P(x_i, y_j)E[X]y_i + \sum_{i=1}^{n}\sum_{j=1}^{n} P(x_i, y_j)E[X]E[Y] =$$
$$\sum_{i=1}^{n}\sum_{j=1}^{n} P(x_i, y_j)x_i y_i - E[Y]\sum_{i=1}^{n}\sum_{j=1}^{n} P(x_i, y_j)x_i - E[X]\sum_{i=1}^{n}\sum_{j=1}^{n} P(x_i, y_j)y_i + E[X]E[Y]\sum_{i=1}^{n}\sum_{j=1}^{n} P(x_i, y_j) =$$
$$\sum_{i=1}^{n}\sum_{j=1}^{n} P(x_i, y_j)x_i y_i - E[Y]\sum_{i=1}^{n} P(x_i)x_i - E[X]\sum_{j=1}^{n} P(y_j)y_i + E[X]E[Y] =$$
$$\sum_{i=1}^{n}\sum_{j=1}^{n} P(x_i, y_j)x_i y_i - E[X]E[Y]$$

Notice that for any random variables, X and Y, and for a constant b, it is true that:

- $VAR[X + Y] = VAR[X] + VAR[Y] + 2COV[X,Y]$
- $VAR[bX] = b^2 VAR[X]$
- $VAR[b] = 0$

Independence We say that X and Y are independent if the events independence condition is satisfied for every pair of their values, i.e. if

$$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j)$$

for all i and for all j.

Proposition If X and Y are independent random variables, then $COV(X,Y) = 0$

Proof. Use the definition of covariance derived above.

$$COV[X,Y] = \sum_{i=1}^{n}\sum_{j=1}^{n} P(x_i, y_j)x_i y_i - E[X]E[Y]$$

If $P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j)$, then

$$COV[X,Y] = \sum_{i=1}^{n}\sum_{j=1}^{n} P(X = x_i)P(Y = y_j)x_i y_i - E[X]E[Y] =$$
$$\sum_{i=1}^{n} P(X = x_i)x_i \sum_{j=1}^{n} P(Y = y_j)y_i - E[X]E[Y] =$$
$$E[X]E[Y] - E[X]E[Y] = 0$$

3.3. Examples of Discrete Random Variables

3.3.1 Bernoulli distribution

A Bernoulli random variable X describes the outcome of an experiment with only two possible outcomes, i.e. success ($x = 1$) and failure ($x = 0$), and a probability of success $P(X = 1)$ equal to $0 \leq p \leq 1$.

Expected value The expected value of a Bernoulli random variable is equal to the probability of success, p, i.e.

$$E[X] = \sum_x p(x)x = (1-p)0 + p1 = p$$

Variance The variance of a Bernoulli random variable is equal to the probability of success times the probability of failure, $p(1-p)$

$$VAR[X] = \sum_x p(x)(x - E[X])^2 = (1-p)(0-p)^2 + p(1-p)^2 =$$
$$= (1-p)(p^2 - 2p) + p(1 + p^2 - 2p) = p^2 - p^3 + p + p^3 - 2p^2 = -p^2 + p =$$
$$p(1-p)$$

3.3.2 Binomial distribution

A Binomial random variable describes the outcome of a series of n independent Bernoulli trials. Two assumptions are needed to construct a Binomial variable as a combination of Binomial r.v.:

- the probability of success for each trial is constant
- trials are independent: the outcome of one observation does not affect the outcome of the other

The probability of x successes out of n trials is defined as follows

$$P(X = x) = C_x^n p^x (1-p)^{(n-x)}$$

where C_x^n is the number of combination of x elements out of n, i.e.

$$C_x^n = \frac{n!}{x!(n-x)!}$$

Expected value The expected value of a Binomial random variable with n trials is equal to n times the expected value of a Bernoulli r.v., i.e. $E[X] = np$.

Proof. Using the formula for the expected value, we get

$$E[X] = \sum_x x C_x^n p^x (1-p)^{(n-x)}$$

Notice that $xC_x^n = nC_{x-1}^{n-1}$. Therefore:

$$\begin{aligned} E[X] = \sum_x n C_{x-1}^{n-1} p^x (1-p)^{(n-x)} &= n \sum_x C_{x-1}^{n-1} p^x (1-p)^{(n-x)} \\ &= np \sum_x C_{x-1}^{n-1} p^{x-1} (1-p)^{(n-1)-(x-1))} = np \end{aligned}$$

since $\sum_x C_{x-1}^{n-1} p^{x-1} (1-p)^{(n-1)-(x-1))} = 1$.

Variance The variance of a Binomial random variable with n trials is equal to n times the variance of a Bernoulli r.v., i.e. $VAR[X] = np(1-p)$.

3.3.3 Poisson distribution

A Poisson random variable is a discrete r.v. that represents a number of successes occurring in a fixed interval of time or space if events occur with a known constant mean rate and independently of the time since the last event. The probability of x successes in a given length of time is defined as follows:

$$P(x) = \frac{e^{-\lambda} \lambda^x}{x!}$$

where $\lambda > 0$ denotes the average number of successes in a given subinterval of time.

Expected value The expected value of a Poisson r.v. is equal to $E[X] = \lambda$.

Variance The variance of a Poisson r.v. is equal to $VAR[X] = \lambda$.

3.4. Exercises

Exercise 1 X is the number of heads from two tosses of a coin, i.e.

$$X = 0 \quad \text{w/ prob} = 0.25$$
$$X = 1 \quad \text{w/ prob} = 0.50$$
$$X = 2 \quad \text{w/ prob} = 0.25$$

Calculate expected value and variance of X.

Exercise 2 Consider the joint probability distribution

$$P(X = 0, Y = 1) = 0.30$$
$$P(X = 1, Y = 1) = 0.25$$
$$P(X = 0, Y = 2) = 0.20$$
$$P(X = 1, Y = 2) = 0.25$$

where X represents the number of exams a student has in a day during final examinations and Y represents the number of snacks eaten by the student during the same day.

- Find the marginal probability distributions for X and Y .
- Calculate the expected value and variance of X and Y .
- Calculate the probability of having one exam conditional on eating two snacks.
- Calculate the covariance of X and Y.

Exercise 3 The probability that a student applying to a college will be admitted is 60%. Suppose a student applies to 6 colleges. What is the probability that exactly 2 colleges admit her?

Exercise 4 Smartphones are shipped from the manufacturer in packets of 12. The probability of a phone being faulty is 0.1 and such faults are independent.

- What is the probability that no more than 2 phones in a shipment are faulty?
- If a shop receives 6 shipments, what is the probability that at least one shipment will contain 3 or more faulty phones?
- Let Y denote the number of shipments containing 3 or more faulty phones. What is the probability that Y will exceed its mean by more than 2 standard deviations?