Section 3

Probability Distributions

Introduction

A probability distribution is similar to the frequency distribution of a quantitative population because both provide a long-run frequency for outcomes.

In other words, a probability distribution is listing of all the possible values that a random variable can take along with their probabilities.

In topic 2, we said a **Random Variable** is a quantity resulting from a random experiment that, by chance, can assume different values.

There are probability distributions for **discrete** and **continuous** random variables.

Probability Distributions for Discrete Random Variables

A discrete random variable is a variable that takes on distinct values for example number of male children in families. Let X be a discrete random variable, the probability distribution of X describes how the probabilities are distributed over the possible values of X. The probability that a r.v X will assume the value x_i is denoted by $P(X = x_i)$.

Properties of Discrete Probability Distributions

Any probability distribution for a discrete random variable(r.v) X has the properties.

- $0 \leq P(X = x_i) \leq 1$ for i=1,2,...,k.
- $\sum_{i=1}^{k} P(X = x_i) = 1$

The functional probability distribution for discrete random variable is often reffered to as probability mass function. More often, the probability distribution for discrete varible is given in tabular form.

Example Suppose a team palys 3 games and has the folllowing probabilities of winning.

$$P(winning\ 0\ games) = 0.65$$
$$P(winning\ 1\ game) = 0.15$$
$$P(winning\ 2\ games) = 0.1$$
$$P(winning\ 3\ games) = 0.1$$

The probability distribution associated with each outcome is given in table form as follows:

X	0	1	2	3
$P(X = x_i)$	0.65	0.15	0.1	0.1

Cumulative Probability Distribution

It is the probability that a r.v X takes a value less than or equal to x, $P(X \le x)$.

Example The cumulative probability distribution for the previous example is

X	0	1	2	3
$P(X \le x_i)$	0.65	0.80	0.90	1

Expectation and Variance of a Discrete Random Variable *X*

The mean value or expected value of a r.v in many trials denoted as $E(X)$ for a discrete r.v is given by

$$E(X) = \sum_{i=1}^{k} x_i P(X = x_i)$$

Variance explains how outcomes of a discrete variable vary and it is given by

$$Var(X) = \sum_{i=1}^{k} (x_i - E(X))^2 P(X - x_i) = E(X^2) - E^2(X)$$

where $E(X^2) = \sum_{i=1}^{k} x_i^2 P(X = x_i)$ and $E^2(X) = [E(X)]^2$.

Example Referring to the first example find $E(X)$ and $Var(X)$.

$$
\begin{aligned}
E(X) &= \sum_{i=1}^{k} x_i P(X = x_i) \\
&= 0 * 0.65 + 1 * 0.15 + 2 * 0.1 + 3 * 0.1 \\
&= 0.65
\end{aligned}
$$

$$
\begin{aligned}
Var(X) = \sum_{i=1}^{k} (x_i - E(X))^2 P(X - x_i) &= E(X^2) - E^2(X) \\
E(X^2) &= \sum_{i=1}^{k} x_i^2 P(X = x_i) \\
= 0^2 * 0.65 + 1^2 * 0.15 + 2^2 * 0.1 + 3^2 * 0.1 &= 1.45 \\
Var(X) &= 1.45 - 0.65^2 \\
&= 1.0275
\end{aligned}
$$

Special Discrete Probability Distributions

These are discrete probability distributions used to model count data

1. Discrete Uniform Distribution

2. Binomial Distribution

3. Poisson Distribution

Discrete Uniform Distribution

It is a variable that takes on integer values within a given interval with equal probabilities. A discrete random variable X is said to follow a Uniform distribution with parameters a and b, written $X \sim U(a, b)$, if it has probability mass function,

$$P(X = x_i) = \frac{1}{b - a + 1}, \; x_i = a, a + 1, a + 2, ..., b - 1, b.$$

Expectation and variance of a Uniform r.v is given by,

$$E(X) = \frac{a + b}{2}$$

$$Var(X) = \frac{(b - a + 1)^2 - 1}{12}$$

Binomial Distribution

One of the ***most widely known*** of all discrete probability distributions is the binomial distribution.

Several characteristics underlie the use of the binomial distribution.

Characteristics of the Binomial Distribution

1. The experiment consists of n identical trials.
2. Each trial has only one of the two possible mutually exclusive outcomes, success or a failure.
3. The probability of each outcome does not change from trial to trial, and
4. The trials are independent, thus we must sample with replacement.

If $X = X_1+X_2+...+X_n$ where X_i are independent and identically distributed (iid) Bernoulli r.v, then the discrete random variable X is said to follow a Binomial distribution with parameters n and p, written $X \sim Bin(n,p)$, if it has probability mass function,

$$P(X = x_i) =^n C_x p^x (1-p)^{n-x}, \ x = 0, 1, 2, ..., n.$$

Expectation and variance of a Binomial r.v is given by

$$E(X) = \sum_{i=1}^{k} x_i P(X = x_i) = np$$

$$Var(X) = \sum_{i=1}^{k} (x_i - E(X))^2 P(X = x_i) = npq.$$

where $q = 1 - p$.

Poisson Distribution

Typically, a Poisson random variable is a count of the number of events that occur in a certain time interval or spatial area. For example, the number of cars passing a fixed point in a 5 minute interval, or the number of calls received by a switchboard during a given period of time.

A discrete random variable X is said to follow a Poisson distribution with parameter λ, written $X \sim Po(\lambda)$, if it has probability mass function,

$$P(X = x_i) = \frac{e^{-\lambda}\lambda^x}{x!}, \ x_i = 0, 1, 2, \ldots.$$

Expectation and variance of a Poisson r.v is given by:

$$E(X) = \lambda$$

$$Var(X) = \lambda$$

Under certain circumstances the Poisson distribution can be used as an approximation to the Binomial distribution.

Probability Distributions for Continuous Random Variables

A continuous random variable is one which takes an **infinite number** of possible values.

Continuous random variables are usually measurements.

Examples include *height*, *weight*, *the amount of sugar in an orange*, e.t.c.

Differences from a Discrete Probability Distribution

- The probability that a continuous random variable will assume a particular value is zero.
- As a result, a continuous probability distribution cannot be expressed in tabular form.
- Instead, an equation or formula is used to describe a continuous probability distribution.

Most often the equation describing a continuous probability distribution is called **probability density function** (*pdf*)

Properties of Continuous Probability Distributions

A function $f(x)$ defined in the interval, $-\infty \leq x \leq +\infty$ is said to be a continuous probability density function for random variable X if:

- $0 \leq \int_a^b f(x)dx \leq 1$
- $\int_{-\infty}^{\infty} f(x)dx = 1$

Cumulative probability distribution

It is a function giving the probability that the random variable X is less than or equal to x that is,

$$F(x) = \int_{-\infty}^{x} f(x)dx,\ where -\infty \leq x \leq \infty.$$

Special Continuous Probability Distributions

✓ ***These are specific distributions used to model continuous data.***

1. Continuous Uniform Distribution
2. Exponential Distribution
3. Normal Distribution

Continuous Uniform Distribution

This is also known as the rectangular distribution. A continuous uniform variable has a uniform probability function over an interval. The probability density function for a continuous uniform distribution on the interval $[a, b]$ is given by

$$f(x) = \frac{1}{b-a},\ a \leq x \leq b$$

Expectation and variance for a continuous uniform distribution are

$$E(X) = \frac{a+b}{2}$$

$$Var(X) = \frac{(b-a)^2}{12}$$

Exponential Distribution

This is the continuous counterpart of discrete Poisson distribution. It describes a process in which events occur continuously and independently at a constant average rate. A continuous random variable X, is said to follow a exponential distribution with parameter λ, written $X \sim Exp(\lambda)$ if it has probability density function,

$$f(x) = \lambda e^{-\lambda x},\ x \geq 0$$

Expectation and variance for a continuous uniform distribution are

$$E(X) = \frac{1}{\lambda}$$

$$Var(X) = \frac{1}{\lambda^2}$$

Normal Distribution

In many natural processes, random variation conforms to a particular probability distribution known as the normal distribution. Normal distributions are symmetrical with a single central peak at the mean (average) of the data. The shape of the curve is described as bell-shaped with the graph falling off evenly on either side of the mean. Fifty percent of the distribution lies to the left of the mean and fifty percent lies to the right of the mean.

A continuous random variable X, taking all real values in the range $(-\infty, \infty)$ is said to follow a Normal distribution with parameters μ and σ^2, written $X \sim N(\mu, \sigma^2)$ if it has probability density function,

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty \leq x \leq \infty.$$

Expectation and variance for a continuous uniform distribution are

$$E(X) = \mu$$

$$Var(X) = \sigma^2$$

Under certain conditions most of probability distributions (discrete and continuous) can be approximated by the Normal distribution.

The normal distribution is the most important of all the continuous distributions. There are two **parameters/characteristics/numbers** used to describe a normal distribution, and these are the mean denoted by $\boldsymbol{\mu}$ and variance denoted by $\boldsymbol{\sigma}^{\mathbf{2}}$

Determining A Probability For a Normal Random Variable

Standard Normal Curve is called the Z-Curve

$$\boldsymbol{Z \sim N(0, 1)}$$

Although you will probably never observe a random variable like $\boldsymbol{Z}$ in practice, it is a useful normal random variable. In fact, area under any normal curve can be determined by finding the curves pending area under the **standard normal curve**.

Area(s) Under Normal Curve

If $\boldsymbol{X}$ is a normal random variable, then

$$\boldsymbol{Z} = \frac{\boldsymbol{X}-\boldsymbol{\mu}}{\boldsymbol{\sigma}} \sim \boldsymbol{N}(\mathbf{0},\mathbf{1})$$

This procedure of subtracting $\boldsymbol{\mu}$ and diving by $\boldsymbol{\sigma}$ is referred to as **standardizing** the normal random variable $\boldsymbol{X}$. It allows us to determine probabilities for any random variable by first standardizing it and the use statistical tables.

Question

A manufacturer packs breakfast cereals in boxes which have normally distributed weights with a mean of 500g and a standard deviation of 4g. Determine the probability of a box of breakfast cereal weighing:

1. (a) Less than 495g, and

   (b) Between 495g and 505g

2. If there are 1000 boxes, what is the expected number of boxes that weigh more than 505g.

Solution

a) Let X be the random variable of packing breakfast cereals. $\mu = 500\ and\ \sigma = 4\quad X \sim N(500, 4^2)$

$$P(X < 495) = P\left(\frac{X-\mu}{\sigma} < \frac{495-\mu}{\sigma}\right)$$

$$= P\left(Z < \frac{495-500}{4}\right)$$

$$= P(Z < -1.25) = \Phi(-1.25) = \mathbf{0.1056}$$

cont

b)$P(495 < X < 505) = P\left(\frac{495-500}{4} < \frac{X-\mu}{\sigma} <\right.$

cont

$2.P(X > 505) = P\left(\frac{X-\mu}{\sigma} > \frac{505-\mu}{\sigma}\right) = P\left(Z > \frac{505-500}{4}\right) = P(Z > 1.25) = 1 - P(z \leq 1.25) = 1 - \Phi(1.25) = 1 - 0.8944 = \mathbf{0.1056}$

Thus, the expected number of boxes is $np = 1000 \times 0.1056 = 105.6 \approx \mathbf{106}\ \boldsymbol{boxes}$

THE END