Introduction to Mathematical Statistics

Definitions

- An EXPERIMENT is a process by which an observation is made.
- The SAMPLE SPACE, S, of an experiment is the set of possible observations (or possible outcomes) for the experiment.
- A SAMPLE POINT is an element of the sample space, i.e. a possible outcome of the experiment.
- An EVENT is a subset of the sample space, i.e. a collection of possible outcomes to the experiment. We say that an event A OCCURS if the outcome is an element of A.

Examples

1. Tossing a coin, sample space $S = \{h, t\}$.
2. Tossing a coin twice, sample space $= \{hh, ht, th, tt\}$.
3. The sample space for throwing a die is $\{1, 2, 3, 4, 5, 6\}$. $A = \{2, 4, 6\}$ is the event "the outcome is even" while $B = \{4, 5, 6\}$ is the event "greater than three".
4. Tossing a coin until a tail appears. The sample space is $S = \{t, ht, hht, hhht, \ldots\}$. Notice that this sample space is not finite.
5. Measuring the hours a computer disk runs. $S = [0, \infty)$. The event $A = [0, 168]$ represents "breakdown in the first week".

Random variables

A function which assigns a number to each possible outcome of an experiment is called a RANDOM VARIABLE. The set of all the values the random variable takes is called the RANGE or IMAGE of the random variable.

Definition

1. A random variable X is a function from a sample space S to $\mathbb{R}$,
2. A random variable X is said to be DISCRETE if the range of X, written as $X(S)$, is a finite or countable infinite set. Otherwise, the random variable is called CONTINUOUS.

Note An infinite set is said to be COUNTABLE if its elements can be written in a sequence. For example, $\mathbb{N} = \{1, 2, 3, \ldots\}$ is countable. An interval, e.g. $[0, 1] = \{x \in \mathbb{R} : 0 \leq x \leq 1\}$ is **not** countable.

Examples

1. Two fair dice are rolled. X is the total score. The sample space is $S = \{(1,1),(1,2),\ldots(6,6)\}$. The range of X is $X(S) = \{2,3,4,\ldots,11,12\}$ which is finite so X is a discrete random variable.
2. A coin tossed until heads appears has sample space $S = \{h, th, tth, \ldots\}$. Let X be the number of tosses required. Then $X : S \to \mathbb{R}$, and the range of X is $\mathbb{N}$ which is countable, so X is a discrete random variable.
3. The height X of a student selected at random is **not** a discrete random variable. The set of values X can take is not countable. (However, if we agree to round up to the nearest milimetre then it is a d.r.v.)

Probability function of a random variable

The PROBABILITY FUNCTION p_X for a discrete random variable X is a function that assigns probabilities to all distinct values that X can take, that is:

$$p_X(w) = P(X = w).$$

Example For Example 1 on previous slide, we have

$$p_X(4) = P(X = 4) = P(\{(1,3),(2,2),(3,1)\}) = \frac{3}{36} = \frac{1}{12}.$$

Of course, if we add up the probabilities of a discrete random variable X taking all values w in the range of the X, then we should get one:

If X is a discrete random variable and $X(S) = \{w_1, w_2, w_3, \ldots\}$ then

$$\sum_{i=1}^{\infty} p_X(w_i) = 1.$$

Independent random variables

Let X and Y be two discrete random variables on a sample space S. We say that X and Y are **independent** random variables if the events $\{X = x\}$ and $\{Y = y\}$ are independent events for any $x \in X(S)$, $y \in Y(S)$.

Exercise Consider the experiment of throwing two fair dice. Decide if the following pairs of random variables are independent.

(i) $X =$ score on first die, $Y =$ score on second die,

(ii) $X =$ total score, $Y =$ score on first die

Expectation of a random variable

Let X be a discrete random variable on a sample space S. The expected value of X, $E(X)$ is defined by

$$E(X) = \sum_{a \in X(S)} a \; p_X(a).$$

$E(X)$ is also called the EXPECTATION of X, or the MEAN of X and may be written as $\mu(X)$ or just μ.

Note $E(X)$ may not be an element of $X(S)$ (that is, $E(X)$ may not be a possible value of X). We think of $E(X)$ as the value X takes "on average".

Example

Let X be the score shown when a fair die is rolled.

$$\begin{aligned}
E(X) &= 1p_X(1) + 2p_X(2) + \cdots + 6p_X(6) \\
&= 1P(X = 1) + 2P(X = 2) + \cdots + 6P(X = 6) \\
&= 1(1/6) + 2(1/6) + 3(1/6) + 4(1/6) + 5(1/6) + 6(1/6) \\
&= \frac{21}{6} = 3.5
\end{aligned}$$

Properties of expectations

Let X and Y be discrete random variables and a, b real numbers. Then

1. $E(aX + b) = aE(X) + b$,
2. $E(X + Y) = E(X) + E(Y)$
3. If $a \leq X \leq b$ then $a \leq E(X) \leq b$,
4. If X and Y are independent random variables then $E(XY) = E(X)E(Y)$.

Example

Again let X be the score shown when a fair die is rolled. Let Y be the square of the score shown when a fair die is rolled.

$$\begin{aligned} E(Y) &= \sum_{w \in Y(S)} w p_Y(w) \\ &= 1p_Y(1) + 4p_Y(4) + \cdots + 36p_Y(36) \\ &= 1P(X=1) + 4P(X=2) + \cdots + 36P(X=6) \\ &= 1(1/6) + 4(1/6) + 9(1/6) + 16(1/6) + 25(1/6) + 36(1/6) \\ &= \frac{91}{6} \end{aligned} \tag{1}$$

The important point to notice here is that $\frac{91}{6} \neq (\frac{7}{2})^2$, so $E(Y) \neq E(X)^2$ even though $Y = X^2$. In other words

$$E(X^2) \neq E(X)^2$$

Variance and Standard Deviation

Let X be a random variable with expected value $\mu = E(X)$.

1. The VARIANCE of X, denoted $V(X)$ or σ^2 is defined by
$$\sigma^2 = E((X-\mu)^2).$$
2. The square root of the variance is called the STANDARD DEVIATION of X and is denoted by σ
$$\sigma = \sqrt{E([X-E(X)]^2)}.$$

Example The variance for the score shown when a fair die is thrown is

$$\sigma^2 = V(X) = \sum_{k=1}^{6}(k-3.5)^2\frac{1}{6} = \frac{35}{12} = 2.917.$$

The standard deviation is $\sigma = 1.7078$.

Properties of the variance

σ and $V(X)$ measure the "spread" of probability of X while $E(X)$ measures its "centre". A low value for σ means that the random variable X usually gives a value close to its average.

Let X and Y be random variables. Then

1. $V(X) = E(X^2) - (E(X))^2$,
2. $V(aX + b) = a^2 V(X)$ for $a, b \in \mathbb{R}$,
3. If X and Y are independent then $V(X + Y) = V(X) + V(Y)$.

Bernoulli Probability Function

A BERNOULLI TRIAL is an experiment with two possible outcomes, success or failure. That is, $S = \{s, f\}$.

Given *any* experiment, pick an event A and call it "success" and the complement of A "failure". (For example, tossing a coin and calling "heads" success.) Associated with a Bernoulli trial, we have a very particular random variable.

Definition Let $\{s, f\}$ be a Bernoulli trial. Let $P(s) = p$ and $P(f) = 1 - p$. We call the random variable X defined by

$$X(s) = 1 \qquad X(f) = 0$$

the BERNOULLI RANDOM VARIABLE WITH PARAMETER p.

It is common when dealing with Bernoulli random variables to write $q = 1 - p$.

Example

A coin is biased so the probability of scoring heads is p. Let X be the associated Bernoulli random variable. Calculate $E(X)$, $E(X^2)$ and $V(X)$.

Here $X(S) = \{0,1\}$ so

$$E(X) = 0P(X=0) + 1P(X=1) = 0(q) + 1(p) = p$$

and

$$E(X^2) = 0^2P(X=0) + 1^2P(X=1) = p$$

and

$$V(X) = E(X^2) - E(X)^2 = p - p^2 = p(1-p) = pq$$

Binomial probability function

1. An experiment consisting of n independent repeated Bernoulli trials each of which has probability p is called a BINOMIAL EXPERIMENT WITH n TRIALS AND PARAMETER p.
2. Let X be the total number of successes in a binomial experiment with n trials and parameter p. X is called the BINOMIAL RANDOM VARIABLE WITH PARAMETERS n AND p.

Example Suppose each player scores a penalty with probability 0.8. Then a penalty shootout is a binomial experiment with 5 trials and parameter 0.8. The number of goals scored is the binomial random variable with parameters 5 and 0.8.

Binomial formula

Recall the binomial formula:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

where the binomial coefficient

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

represents the number of ways to choose k elements out of a set of n.

Binomial random variables

If X is the binomial random variable with parameters n and p then the probability of k successes in the binomial experiment is

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

For example, the probability of scoring 4 out of 5 penalties, if the probability of scoring each one is 0.8, is $\binom{5}{4}(.8)^4(.2)^1 = .4096$.

Let X be the binomial random variable with parameters n and p. Then its expectation and variance are given by

$$E(X) = np \quad \text{and} \quad V(X) = npq.$$

Example

A component manufacturer knows that 2% of components produced are defective, but guarantees that there are no more than two defective items in a box of 20. What is the probability that a box satisfies the guarantee?

Let X be the number of defective items in a box. X is a binomial random variable with parameters $n = 20$ and $p = .02$. We have

$$p_X(k) = \binom{20}{k}(0.02)^k(0.98)^{20-k}$$
$$p_X(0) = \binom{20}{0}(0.02)^0(0.98)^{20} = 0.668$$
$$p_X(1) = \binom{20}{1}(0.02)^1(0.98)^{19} = 0.273$$
$$p_X(2) = \binom{20}{2}(0.02)^2(0.98)^{18} = 0.053$$

This means that $P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = 0.994$.

Continuous Random Variables

The DISTRIBUTION FUNCTION of a random variable X is defined by

$$F(x) := P(X \leq x) \qquad x \in \mathbb{R}.$$

Example Let X be a binomial random variable with parameters $n = 2$ and $p = 0.5$. The probability function for X is given by

$$P_X(x) = \binom{2}{x}(0.5)^x(0.5)^{2-x} \qquad (x = 0, 1, 2)$$

Hence $P_X(0) = \frac{1}{4}$, $P_X(1) = \frac{3}{4}$, $P_X(2) = \frac{1}{4}$ and

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{4} & 0 \leq x < 1 \\ \frac{3}{4} & 1 \leq x < 2 \\ 1 & x \geq 2 \end{cases}$$

The "step-like" appearance of the distribution function for the random variable in the above example is characteristic of discrete random variables.

Definition If X is a random variable whose distribution function is a continuous function on $\mathbb{R}$ then we call X a CONTINUOUS RANDOM VARIABLE.

Definition A function $f : \mathbb{R} \to [0, \infty)$ is said to be a PROBABILITY DENSITY FUNCTION (or pdf) for the random variable X if

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt.$$

Properties

1. If X is a continuous random variable then $P_X(a) = 0$ for all a.
2. $\int_{-\infty}^{\infty} f(t)\,dt = 1$ for any pdf f.
3. Let f be a pdf for the continuous random variable X. For $a \leq b$,

$$P(a \leq X \leq b) = \int_a^b f(t)\,dt.$$

Expectation

The EXPECTATION of X is defined by

$$E(X) = \mu = \int_{-\infty}^{\infty} t\, f(t)\, dt.$$

We also have $E(X^2) = \int_{-\infty}^{\infty} t^2 f(t)\, dt$.

As before, the VARIANCE is defined to be

$$V(X) = \sigma^2 = E([X - \mu]^2) = E(X^2) - E(X)^2$$

and the STANDARD DEVIATION is given by $\sigma = \sqrt{V(X)}$.

Example

A builder estimates that the probability of an apartment block being built within x months is 0 for $x < 1$ and $1 - \frac{1}{x^3}$ for $x \geq 1$. Let X be the number of months the builder takes. Calculate

(a) $P(X \leq 2)$ and $P(3 \leq X \leq 4)$,

(b) the probability density function f_X for X,

(c) the expectation of X,

(d) the standard deviation of X.

The Normal Distribution

The normal distribution, also known as the Gaussian distribution, is the most important continuous distribution function as it is an excellent approximation to many distributions which occur in science and economics.

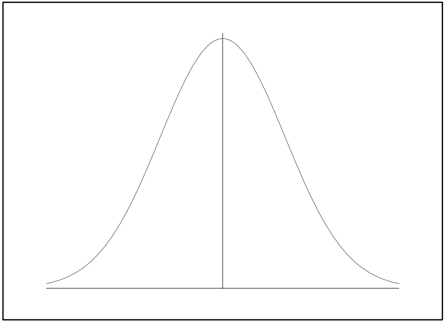
Definition The random variable X has a NORMAL DISTRIBUTION if it has a pdf of the form

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left[\frac{X-\mu}{\sigma}\right]^2\right) \tag{2}$$

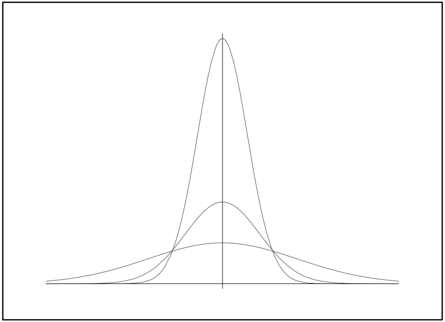
for all $x \in \mathbb{R}$ where $\mu \in \mathbb{R}$ and $\sigma > 0$ are fixed.

We will say that X has distribution $N(\mu, \sigma^2)$, or simply that X is $N(\mu, \sigma^2)$ if the above holds.

The graph of $f_X(x)$ is a "bell-shaped curve".



(a) $N(\mu,\sigma)$



(b) $N(0,\sigma)$ for different values of σ

The value of μ affects where along the x-axis the curve is centered. It is symmetric about μ and the value of $f_X(x)$ is negligible for values of x more than a distance of about 3σ away from μ. The "flatness" of the bell is controlled by σ but the area under the curve is always equal to 1.

Properties of the normal distribution

1. Since
$$\int_{-\infty}^{\infty} e^{-y^2}\,dy = \sqrt{\pi} \tag{3}$$
then $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$
2. If X has distribution $N(\mu, \sigma^2)$ then $E(X) = \mu$ and $V(X) = \sigma^2$.
3. If X has distribution $N(\mu, \sigma^2)$ and $Y = aX + b$ where a and b are constants then Y has distribution $N(a\mu + b, a^2\sigma^2)$.
4. If X has distribution $N(\mu, \sigma^2)$ then $Y = \frac{X-\mu}{\sigma}$ has distribution $N(0,1)$.

Definition We say X has STANDARD NORMAL DISTRIBUTION if X has distribution $N(0,1)$, that is, X has probability density function given by

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

New Cambridge Statistical Tables

The probability density function for the normal distribution can be calculated at any point on a standard scientific calculator. However, in practice it is the value of the probability distribution which is of interest, and so one must integrate the pdf. A table of values for the probability distribution of the random variable $N(0,1)$ is contained in the *New Cambridge Statistical Tables*. On page 34-35, a range of values of

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\,dt$$

is given, that is, a range of values of $\Phi(x) = P(N(0,1) \leq x)$.

Example

Suppose X is the random variable with distribution $N(0, 1)$. Find

1. $P(X < 3)$
2. $P(X < -3)$,
3. $P(-3 < X < 3)$,
4. $P(X > 1.26)$,
5. $P(X < 0.195)$.