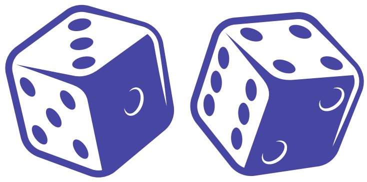
Probability theory



Random Experiment

An experiment that can result in different outcomes, even though it is repeated in the same manner every time, is called a random experiment.

Sample Space

The set of all possible outcomes of a random experiment is called the sample space of the experiment. The sample space is denoted as S.

Discrete

A sample space is discrete if it consists of a finite or countable infinite set of outcomes.

Continuous

A sample space is continuous if it contains an interval (either finite or infinite) of real numbers.

Event

An event is a subset of the sample space of a random experiment.

* We can use basic set operations to form other events of interest.

Union

The union of two events is the event that consists of all outcomes that are contained in either of the two events.

$E_1 \cup E_2$

Intersection

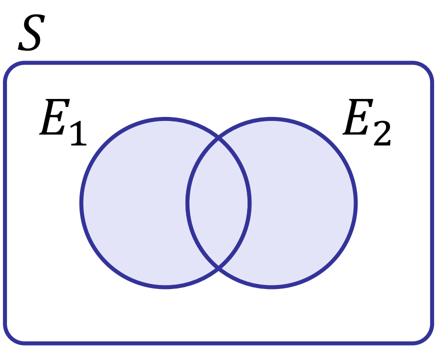
The intersection of two events is the event that consists of all outcomes that are contained in both of the two events.

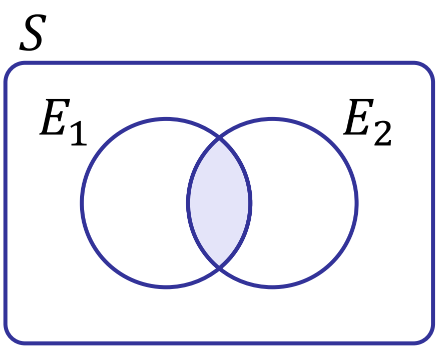
$E_1 \cap E_2$

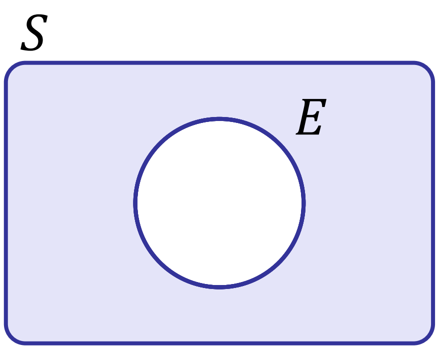
Complement

The complement of an event in a sample space is the set of outcomes in the sample space that are not in the event.

E' or E^C





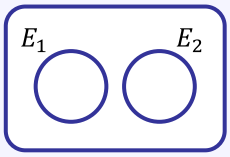


Mutually Exclusive Events

Two events, denoted as E_1 and E_2, such that

$$E_1 \cap E_2 = \emptyset$$

are said to be mutually exclusive or disjoint.



DeMorgan's laws

$$(A \cap B)^C = A^C \cup B^C \qquad (A \cup B)^C = A^C \cap B^C$$

* Counts of the numbers of outcomes in sample spaces and various events used to analyze random experiments are referred to as counting techniques.

Fundamental Counting Principle

Assume an operation can be described as a sequence of k steps, and

- the number of ways of completing step 1 is n_1, and
- the number of ways of completing step 2 is n_2 for each way of completing step 1, and
- the number of ways of completing step 3 is n_3 for each way of completing step 2, and so forth.

The total number of ways of completing the operation is

$$n_1 \times n_2 \times \cdots \times n_k$$

Permutations

The number of permutations, ordered sequences, of n different elements is $n!$ where

$$n! = n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1$$

The number of permutations of subsets of r elements selected from a set of n different elements is

$$P_r^n = n \times (n-1) \times (n-2) \times \cdots \times (n-r+1) = \frac{n!}{(n-r)!}$$

Permutations of Similar Objects

The number of permutations of $n = n_1 + n_2 + \cdots + n_r$ objects of which n_1 are of one type, n_2 are of a second type,…, and n_r are of an rth type is

$$\frac{n!}{n_1!\, n_2!\, n_3! \ldots n_r!}$$

Combinations

The number of combinations, subsets of r elements that can be selected from a set of n elements, is denoted as C_r^n or $\binom{n}{r}$ and

$$C_r^n = \binom{n}{r} = \frac{n!}{r!\,(n-r)!}$$

Probability

Probability is used to quantify the likelihood, or chance, that an outcome of a random experiment will occur.

- The probability of an outcome can be interpreted as our subjective probability, or degree of belief, that the outcome will occur.
- Another interpretation of probability is based on the conceptual model of repeated replications of a random experiment or the long-run relative frequency of occurrence.

Equally Likely Outcomes

Whenever a sample space consists of N possible outcomes that are equally likely, the probability of each outcome is $1/N$.

Probability of an Event

For a discrete sample space, the probability of an event E, denoted as $P(E)$, equals the sum of the probabilities of the outcomes in E.

Axioms of Probability

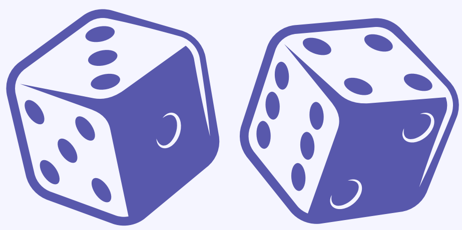
Probability is a number assigned to each member of a collection of events from a random experiment that satisfies the following:

If S is the sample space and E is any event in a random experiment,

(1) $P(E) \geq 0$

(2) $P(S) = 1$

(3) For two mutually exclusive events
$$P(E_1 \cup E_2) = P(E_1) + P(E_2)$$



Probability of a Union

The probability of event A or B is interpreted as $P(A \cup B)$ and the following general addition rule applies.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Conditional Probability

The probability of an event B under the knowledge that the outcome will be in event A is denoted as $P(B|A)$ and this is called the conditional probability of B given A.

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \quad \text{for} \quad P(A) > 0$$

Multiplication Rule

$$P(A \cap B) = P(B|A)P(A) = P(A|B)P(B)$$

Total Probability Rule

$$P(B) = P(B \cap A) + P(B \cap A^C)$$
$$= P(B|A)P(A) + P(B|A^C)P(A^C)$$

Independence

Two events are independent if any one of the following equivalent statements is true:

(1) $P(A|B) = P(A)$

(2) $P(B|A) = P(B)$

(3) $P(A \cap B) = P(A)P(B)$

Bayes' Theorem

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad \text{for} \quad P(B) > 0$$

Random Variable

A random variable, denoted by an uppercase letter, is a function that assigns a real number to each oucome in a sample space of a random experiment. Its measured values are denoted by lowercase letters.

Discrete

A discrete random variable is a random variable with a finite (or countably infinite) range

Continuous

The range of a continuous random variable is an interval (either finite or infinite) of the reals.

Probability Mass Function

For a discrete random variable X with possible values $x_1, x_2, \dots, x_n$, a probability mass function is a function such that

(1) $f(x_i) \geq 0$

(2) $\sum_{i=1}^{n} f(x_i) = 1$

(3) $f(x_i) = P(X = x_i)$

Cumulative Distribution Function

The cumulative distribution function of a discrete random variable X is a is a function such that

(1) $F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$

(2) $0 \leq F(x) \leq 1$

(3) If $x \leq y$, then $F(x) \leq F(y)$

Mean, Variance, and Standard Deviation

The mean or expected value, variance, and standard deviation of the discrete random variable X are defined respectively as

$$\mu = E(X) = \sum_{x} x f(x)$$

$$\sigma^2 = V(X) = E\left((X-\mu)^2\right) = \sum_{x} x^2 f(x) - \mu^2$$

$$\sigma = \sqrt{\sigma^2}$$

Bernoulli trial

A random experiment with two possible outcomes, success or failure.

$$X \sim \text{Bernoulli}(p) \qquad f(x) = \begin{cases} p, & x = 1 \\ 1 - p, & x = 0 \end{cases} \qquad \begin{array}{c} \mu = p \\ \sigma^2 = p(1-p) \end{array}$$

Binomial Distribution

A random experiment consists of n Bernoulli trials such that

(1) The trials are independent.
(2) Each trial results in only two possible outcomes.
(3) The probability p of a success in each trial remains constant.

The random variable X that equals the number of trials that result in a success is a binomial random variable with parameters p and $n = 1,2, \ldots$

$$X \sim \text{Binomial}(n, p) \qquad f(x) = \binom{n}{x} p^x (1-p)^{n-x} \quad x = 0,1, \ldots, n$$

$$\mu = np \qquad \sigma^2 = np(1-p)$$

Geometric Distribution

In a series of Bernoulli trials (independent trials with constant probability of a success), the random variable X that equals the number of trials until the first success is a geometric random variable with parameter p and

$$X \sim \text{Geometric}(p) \qquad f(x) = (1-p)^{x-1} p \qquad x = 1,2, \ldots$$

$$\mu = 1/p \qquad \sigma^2 = (1-p)/p^2$$

Poisson Distribution

The random variable X that equals the number of events in an interval of length T in a Poisson process, a random process that models the occurrence of events in time or space where λ is the average rate of events per unit time or space and the number of events in disjoint intervals are independent random variables, is a Poisson random variable with parameter $0 < \lambda$, and

$$X \sim \text{Poisson}(\lambda) \qquad f(x) = \frac{e^{-\lambda T} (\lambda T)^x}{x!} \qquad x = 0,1,2, \ldots$$

$$\mu = \lambda T \qquad \sigma^2 = \lambda T$$

Probability Density Function

For a continuous random variable X, a probability density function is a function such that

(1) $f(x) \geq 0$

(2) $\int_{-\infty}^{\infty} f(x)\,dx = 1$

(3) $P(a \leq X \leq b) = \int_{a}^{b} f(x)\,dx$

Cumulative Distribution Function

The cumulative distribution function of a continuous random variable X is

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)\,du \qquad f(x) = \frac{dF(x)}{dx}$$

Mean, Variance, and Standard Deviation

The mean or expected value, variance, and standard deviation of a continuous random variable X are defined respectively as

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x)\,dx$$

$$\sigma^2 = V(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx = E(X^2) - \mu^2$$

$$\sigma = \sqrt{\sigma^2}$$

Moment-Generating Function

The rth moment about the origin of the random variable X is

$$\mu_r' = E(X^r) = \begin{cases} \sum_{x} x^r f(x), & X \text{ discrete} \\ \int_{-\infty}^{\infty} x^r f(x)\,dx, & X \text{ continuous} \end{cases}$$

The moment-generating function of the random variable X is the expected value of e^{tX} and is denoted by $M_X(t)$. That is,

$$M_X(t) = E(e^{tX}) \qquad \mu_r' = \left.\frac{d^r M_X(t)}{dt^r}\right|_{t=0}$$

Continuous Uniform Distribution

A continuous random variable X with probability density function

$$f(x) = 1/(b-a), \quad a \leq x \leq b$$

is a continuous uniform random variable.

$$X \sim \text{Uniform}(a,b) \qquad \mu = \frac{(a+b)}{2} \qquad \sigma^2 = \frac{(b-a)^2}{12}$$

Exponential Distribution

The random variable X that equals the distance between successive events from a Poisson process with mean number of events $0 < \lambda$ per unit interval is an exponential random variable with parameter λ. The probability density function of X is

$$X \sim \text{Exponential}(\lambda) \qquad f(x) = \lambda e^{-\lambda x} \quad 0 \leq x < \infty$$

$$\mu = \frac{1}{\lambda} \qquad \sigma^2 = \frac{1}{\lambda^2}$$

$$P(X < t_1 + t_2 | X > t_1) = P(X < t_2)$$ **Lack of Memory Property**

Normal Distribution

A random variable X with probability density function

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad -\infty < x < \infty$$

is a normal random variable with parameters $-\infty < \mu < \infty$ and $\sigma > 0$. Also,

$$X \sim N(\mu, \sigma^2) \qquad E(X) = \mu \qquad V(X) = \sigma^2$$

A standard normal random variable is defined as

$$Z \sim N(0,1) \qquad \mu = 0 \qquad \sigma^2 = 1$$

The cumulative distribution function of a standard normal random variable is a non-elementary antiderivative and is denoted as

$$\Phi(z) = P(Z \leq z)$$

We can standarize a normal random variable to calculate probabilities

$$Z = \frac{X-\mu}{\sigma} \qquad P(X \leq x) = P\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) = P(Z \leq z)$$

Joint Probability Mass Function

The bivariate probability distribution of two discrete random variables X and Y is a function that satisfies

(1) $f_{XY}(x,y) \geq 0$

(2) $\sum_x \sum_y f_{XY}(x,y) = 1$

(3) $f_{XY}(x,y) = P(X = x, Y = y) = P(X = x \text{ and } Y = y)$

Marginal Probability Mass Function

The marginal probability mass functions of X and Y are defined as

$$f(x) = \sum_y f(x,y) \qquad f(y) = \sum_x f(x,y)$$

Conditional Probability Mass Function

The conditional probability mass functions of X and Y are defined as

$$f_{Y|X}(y) = \frac{f_{XY}(x,y)}{f_X(x)} \qquad f_{X|Y}(x) = \frac{f_{XY}(x,y)}{f_Y(y)}$$

Expected Value of a Function

$$E\big(h(X,Y)\big) = \sum_x \sum_y h(x,y) f(x,y)$$

Variance, Covariance, and Correlation

The variance, covariance, and correlation of two discrete bivariate random variables X and Y are defined respectively as

$$V(X+Y) = V(X) + V(Y) + 2\,\mathrm{cov}(X,Y)$$

$$\mathrm{cov}(X,Y) = \sigma_{XY} = E[(X-\mu_X)(Y-\mu_Y)] = E(XY) - \mu_X \mu_Y$$

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}, \qquad -1 \leq \rho_{XY} \leq +1$$

Independence

$$f_{XY}(x,y) = f_X(x) f_Y(y) \qquad E(XY) = E(X)E(Y) \qquad \sigma_{XY} = \rho_{XY} = 0$$

Statistics

Statistics is the discipline that concerns the collection, organization, analysis, interpretation, and presentation of data.

Descriptive

Descriptive statistics deals with organizing and summarizing data using numerical summaries

Inferential

Inferential statistics deals with drawing conclusions about populations based on samples.

Measures of Location

Sample mean
affected by extreme values (outliers)

$$\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

Sample median
not affected by extreme values (outliers)

$$\tilde{x} = \begin{cases} x_{(n+1)}/2, & n \quad \text{odd} \\ \left(x_{n/2} + x_{n/2+1}\right)/2, & n \quad \text{even} \end{cases}$$

Quartiles are values that separate the data into four equal parts.

Percentiles are values that separate the data into 100 equal parts.

Measures of Variability

Sample range

$$r = \max(x_i) - \min(x_i)$$

Sample variance
where $n-1$ is called the degree of freedom

$$s^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1} = \frac{\left[\sum_{i=1}^{n} x_i^2\right] - n(\bar{x})^2}{n-1}$$

Sample standard deviation

$$s = \sqrt{s^2}$$

Sample covariance

$$\operatorname{cov}(x, y) = \sum_{i=1}^{n} \frac{(x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

Interquartile range is the difference between the third and first quartiles.