

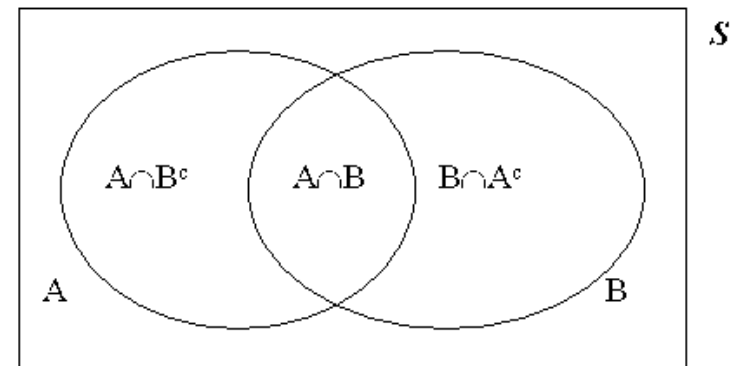
Basic Rules of Probability and Combinatorial Principles

Basic Set Operations

- $A \cup B = \{x \mid x \in A \text{ or } x \in B \text{ or both}\}; A \cap B = \{x \mid x \in A \text{ and } x \in B\}$
- A' or $A^c = \text{complement of } A = \{x \mid x \notin A\}; (A \cup B)^c = A^c \cap B^c; (A \cap B)^c = A^c \cup B^c$
- $A \subset B = B$ contains all the sample points in event A. It calls A is a subevent of B.
- If $A \subset B$, then $A \cap B = A$ and $A \cup B = B$

Basic Probability

- For a sample space S , $\Pr[S] = 1$; $\Pr[\emptyset] = 0$ where $\emptyset$ is the null or empty set
- For an event A in the sample space S, $1 \geq \Pr[A] \geq 0$
- For any event A, $\Pr[A] = 1 - \Pr[A^c]$
- If A and B are two mutually exclusive events, (no sample points in common $A \cap B = \emptyset$), then $\Pr[A \cap B] = 0$
- **General Formula :** $\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B]$
- $\Pr[A \cup B \cup C] = \Pr[A] + \Pr[B] + \Pr[C] - \Pr[A \cap B] - \Pr[A \cap C] - \Pr[B \cap C] + \Pr[A \cap B \cap C]$
- If $A \subset B$ then $\Pr[A] \leq \Pr[B]$
- $\Pr[A] = \Pr[A \cap B] + \Pr[A \cap B^c]$



Conditional Probability

- For two events A and B, the **conditional probability** of A given B has occurred is:

$$\Pr[A | B] = \frac{\Pr[A \cap B]}{\Pr[B]}$$

- If $A \subset B$, then $\Pr[A | B] = \frac{\Pr[A]}{\Pr[B]}$ and $\Pr[B | A] = 1$
- $\Pr[A^c | B] = 1 - \Pr[A | B]$ and if $A_1 \cap A_2 = \emptyset$ then $\Pr[A_1 \cup A_2 | B] = \Pr[A_1 | B] + \Pr[A_2 | B]$
i.e. $\Pr[\cdot | B]$ is a probability on S .

Baye's Theorem

Let $A_1, A_2, \dots, A_n$ be a collection of n mutually exclusive and exhaustive events with $\Pr[A_i] > 0$ for $i=1, \dots, n$.
Then, for any other event B,

$$\Pr[B] = \sum_{i=1}^n \Pr[B | A_i] \Pr[A_i]$$

$$\Pr[A_k | B] = \frac{\Pr[A_k \cap B]}{\Pr[B]} = \frac{\Pr[B | A_k] \Pr[A_k]}{\sum_{i=1}^n \Pr[B | A_i] \Pr[A_i]} \quad k = 1, \dots, n$$

Independence

Two events A and B are **independent** if and only if $\Pr[B | A] = \Pr[B]$

if and only if $\Pr[A | B] = \Pr[A]$

if and only if $\Pr[A \cap B] = \Pr[A] \cdot \Pr[B]$

Combinatorial Principles

Fundamental Principle of Counting: (also known as the multiplication rule for counting) If a task can be performed in n_1 ways, and for each of these a second task can be performed in n_2 ways, and for each of the latter a third task can be performed in n_3 ways, ... , and for each of the latter a k th task can be performed in n_k ways, then the entire sequence of k tasks can be performed in

$$n_1 \cdot n_2 \cdot n_3 \cdot \dots \cdot n_k \text{ ways.}$$

Permutations

An ordered sequence of k objects taken from a set of n distinct objects is called a permutation of the objects of size k . P_k^n denotes the number of size k permutations that can be constructed from the n objects.

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

Combinations

An unordered subset of k objects taken from a set of n distinct objects is called a combination. $\binom{n}{k}$ denotes the number of combinations of size k that can be formed from n objects.

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

$$\binom{n}{n} = 1 \quad \binom{n}{0} = 1 \quad \binom{n}{1} = \binom{n}{n-1} = n \quad \binom{n}{k} = \binom{n}{n-k}$$

Random Variables and Probability Distributions, Expectation and Other Distributional Parameters

Random Variables Probability Distributions and Independence

- Given an experiment with sample space S , a random variable X is any rule that associates a number with each outcome in S .
- The cumulative distribution function (cdf) of a random variable written $F(x)$ is defined for every number x by: $F(x) = \Pr[X \leq x]$
- Two random variables X and Y are independent if

$$\Pr[X \in A, Y \in B] = \Pr[X \in A] \Pr[Y \in B] \text{ for every } A \text{ and } B.$$

Discrete Random Variables

- A set is described as discrete if its elements can be listed in sequence or if it consists of a finite number of elements.
- The probability distribution or probability mass function of a discrete random variable written $p(x)$ is defined for every number by: $p(x) = \Pr[X = x]$
- The probability mass function $p(x)$ must satisfy: $0 \leq p(x) \leq 1$ and $\sum_x p(x) = 1$.
- The cumulative distribution function (cdf) of a discrete random variable with pmf $p(x)$ written $F(x)$ is defined by:

$$F(x) = \sum_{k \leq x} \Pr[X = k] \text{ for every number } x$$

- Two random discrete variables X and Y are independent if $\Pr[X=x, Y=y] = \Pr[X=x] \Pr[Y=y]$ for every x and y .

Continuous Random Variables

- A random variable X is continuous if its set of possible values is an entire interval of numbers.
- The probability distribution or probability density function (pdf) of a continuous random variable written $f(x)$ is :

$$\Pr[a \leq X \leq b] = \int_a^b f(x) dx$$

- The probability density function $f(x)$ must satisfy: $f(x) \geq 0$ and $\int_{-\infty}^{\infty} f(x) dx = 1$.
- The cumulative distribution function of a continuous random variable written $F(x)$ is: $F(x) = \Pr[X \leq x] = \int_{-\infty}^x f(y) dy$
- We have $\Pr[a \leq x \leq b] = F(b) - F(a)$ and $F'(x) = f(x)$

Expectation, Variance, Standard deviation, Covariance

- The **expected value** or **mean** value of a discrete random variable written $E(X)$ or μ_x is:

$$E(X) = \sum_x x \Pr[X = x] \text{ if } X \text{ is discrete and } E(X) = \int_{-\infty}^{\infty} x f(x) dx \text{ if } X \text{ is continuous.}$$

- The **variance** of a discrete random variable written $Var(X)$ or σ_x^2 is: $Var(X) = E[(X - E(X))^2] = E[X^2] - [E(X)]^2$

$$Var(X) = \sum_x (x - \mu_x)^2 \Pr[X = x] \text{ if } X \text{ is discrete and } Var(X) = \int_{-\infty}^{\infty} (x - \mu_x)^2 f(x) dx \text{ if } X \text{ is continuous.}$$

- $\sigma_x = \text{standard deviation} = \sqrt{Var(X)}$
- $Cov(X, Y) = E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y) = Cov(Y, X)$
If X, Y are independent, then $Cov(X, Y) = 0$

Rules of expected value:

- For any constant a , $E(aX) = aE(X)$
- For any random variables X, Y , $E(X + Y) = E(X) + E(Y)$

Rules of variance:

- For any constants a and b , $Var(aX + b) = a^2[Var(X)]$
- For any random variables X, Y , $Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)$
If X, Y are independent, then $Var(X + Y) = Var(X) + Var(Y)$

Expected value of a function of a random variable

$$E[h(X)] = \sum_x h(x) \Pr[X = x] \text{ if } X \text{ is discrete and } E[h(X)] = \int_{-\infty}^{\infty} h(x) f(x) dx \text{ if } X \text{ is continuous.}$$

Moment generating function

The moment generating function of a discrete random variable written $M_X(t)$ is: $M_X(t) = E(e^{tx})$

- **Expectation, variance, and moments via mgf's:** $M_X(0) = 1$,

$$M_X'(0) = E(X), \quad M_X''(0) = E(X^2), \quad M_X'''(0) = E(X^3), \text{ etc.} \quad \text{Var}(X) = M_X''(0) - M_X'(0)^2.$$

- **Mgf of a multiple of a r.v.:** If X has mgf $M_X(t)$, and $Y = cX$ with c a constant, then the mgf of Y is

$$M_Y(t) = E(e^{tY}) = E(e^{tcX}) = M_X(tc).$$

- **Mgf of a sum of independent r.v.'s X and Y :** If X and Y are independent, then $X + Y$ has mgf

$$M_{X+Y}(t) = M_X(t)M_Y(t).$$

(An analogous formula holds for sums of more than two independent r.v.'s.)

Other parameters of the distribution

- For $0 \leq p \leq 1$, the $(100p)^{th}$ percentile of a continuous random variable written $\eta(p)$ is:

$$p = F[\eta(p)] = \int_{-\infty}^{\eta(p)} f(y) dy$$

- The median of a continuous distribution written $\tilde{\mu}$ satisfies $0.5 = F(\tilde{\mu}) = \text{the } 50^{th} \text{ percentile of the distribution.}$
- The mode of a distribution is any point m at which the probability mass function $p(x)$ or the density function $f(x)$ is maximized.

- The skewness of a distribution is:
$$\frac{E[(X - \mu)^3]}{\sigma^3}$$

Discrete Distributions

1. Uniform A distribution of N points, $1, 2, \dots, N$, where N is an integer.

$$p(x) = \frac{1}{N} \quad E(X) = \frac{N+1}{2} \quad Var(X) = \frac{N^2-1}{12} \quad M_X(t) = \frac{e^t(e^{Nt}-1)}{N(e^t-1)}$$

2. Binomial An experiment consisting of a fixed number of n trials. Each trial is identical and results in one of two possible outcomes, denoted success (S) or failure (F). The trials are independent and the probability of a success is denoted p . The binomial random variable X is equal to the number of successes in n trials.

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, 2, \dots, n, \quad E(X) = np \quad Var(X) = np(1-p) \quad M_X(t) = (1-p + pe^t)^n$$

3. Hypergeometric Closely related to binomial, but it is known that the population (N) contains M successes. X is the number of successes in a random sample of size n drawn from a population of M successes.

$$p(x) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}} \quad E(X) = n \frac{M}{N} \quad Var(X) = n \frac{M}{N} \left(1 - \frac{M}{N}\right) \frac{N-n}{N-1}$$

4. Pascal Related to binomial. An experiment continues until a total of r successes have been observed (r is a positive integer). X is equal to the number of trials until the r^{th} success is completed.

$$p(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \quad x = r, r+1, r+2, \dots \quad E(X) = \frac{r}{p} \quad Var(X) = \frac{r(1-p)}{p^2} \quad M_X(t) = \left[\frac{pe^t}{1-(1-p)e^t} \right]^r$$

5. Geometric Special case of negative binomial when $r = 1$. So, want to perform the experiment until a success occurs. X is equal to the number of trials until the first success is completed.

$$p(x) = p(1-p)^{x-1}, \quad x = 1, 2, \dots \quad E(X) = \frac{1}{p} \quad Var(X) = \frac{1-p}{p^2} \quad M_X(t) = \frac{pe^t}{1-(1-p)e^t}$$

6. Poisson Not based on any simple experiment. Most often used to count the number of a certain type of event that occurs in a period of time.

$$p(x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, 2, \dots \quad E(X) = \lambda \quad Var(X) = \lambda \quad M_X(t) = e^{\lambda(e^t-1)}$$

for some $\lambda > 0$ where λ is frequently some rate per unit of time.

Continuous Distributions

1. **Uniform** X is distributed uniformly over (a, b) .

$$f(x) = \frac{1}{b-a}, \quad a < x < b \quad E(X) = \frac{b+a}{2} \quad Var(X) = \frac{(b-a)^2}{12} \quad M_X(t) = \frac{e^{bt} - e^{at}}{(b-a)t}$$

2. **Normal** Distribution given that is the mean and is the variance of the distribution.

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad E(X) = \mu \quad Var(X) = \sigma^2 \quad M_X(t) = \exp\left[\mu t + \frac{\sigma^2 t^2}{2}\right]$$

When $\mu = 0$ and $\sigma = 1$, the normal distribution is called a **standard normal** distribution and the standard normal random variable is denoted Z .

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

If X has a normal distribution, it can be standardized by : $Z = \frac{X-\mu}{\sigma}$

3 **Gamma** For $\alpha > 0$, the gamma function $\Gamma(\alpha)$ is defined by: $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$.

$$\Gamma(\alpha+1) = \alpha \Gamma(\alpha) \quad \text{and} \quad \Gamma(n) = (n-1)! \quad \text{for } n \text{ positive integer.}$$

A continuous random variable X is said to have a gamma distribution with parameters $\alpha, \beta > 0$ if:

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad x > 0 \quad E(X) = \frac{\alpha}{\beta} \quad Var(X) = \frac{\alpha}{\beta^2} \quad M_X(t) = \left(\frac{\beta}{\beta-t}\right)^\alpha \quad \text{for } t < \beta$$

4. **Exponential** The exponential distribution is a special case of the gamma distribution with $\alpha = 1$ and $\beta = \lambda$

$$f(x) = \lambda e^{-\lambda x}, \quad x > 0 \quad E(X) = \frac{1}{\lambda} \quad Var(X) = \frac{1}{\lambda^2} \quad M_X(t) = \frac{\lambda}{\lambda-t} \quad \text{for } t < \lambda$$

Joint, Marginal and Conditional Distributions

- Let X and Y be two discrete random variables defined on the sample space S of an experiment. The joint probability mass function $p_{X,Y}(x, y)$, is defined for each (x,y) pair by:

$$p_{X,Y}(x, y) = p(x, y) = \Pr[X = x, Y = y]$$

and $p(x, y)$ must satisfy:

$$0 \leq p(x, y) \leq 1 \text{ and } \sum_x \sum_y p(x, y) = 1.$$

- Let X and Y be two continuous random variables. Then, the joint probability density function $f_{X,Y}(x, y) = f(x, y)$ must satisfy:

$$f(x, y) \geq 0 \quad \text{and} \quad \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1.$$

Marginal Distribution

- The marginal probability mass function for the discrete random variables X and Y , denoted $p_X(x)$ and $p_Y(y)$ are given by:

$$p_X(x) = \sum_y p(x, y) \quad \text{and} \quad p_Y(y) = \sum_x p(x, y)$$

- In the continuous case, they are given by:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy \quad \text{and} \quad f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx$$

Expected Value

Let X and Y be two jointly distributed random variables. The **expected value of a function** h of (X, Y) denoted $E[h(X, Y)]$ is:

$$E[h(X, Y)] = \sum_x \sum_y h(x, y) p(x, y) \quad \text{If } X \text{ and } Y \text{ are discrete}$$

$$E[h(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) f(x, y) dx dy \quad \text{If } X \text{ and } Y \text{ are continuous}$$

Moment Generating Function

The moment generating function for two jointly distributed random variables X and Y , denoted $M_{X,Y}(t_1, t_2)$ is:

$$M_{X,Y}(t_1, t_2) = E[e^{t_1 X + t_2 Y}]$$

Variance, Covariance, Correlation

- The covariance between two random variables X and Y is :

$$Cov(X, Y) = E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y)$$

$$Cov(X, Y) = \sum_x \sum_y xy p(x, y) - \mu_X \mu_Y \text{ in the discrete case} \quad Cov(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f(x, y) dx dy - \mu_X \mu_Y \text{ in the continuous case}$$

- If X and Y are independent, **then** $E[X \cdot Y] = E[X]E[Y]$ $Cov(X, Y) = 0$
- $Var[X + Y] = Var[X] + Var[Y] + 2 Cov[X, Y]$
- The correlation coefficient of X and Y , denoted by $corr(X, Y)$ or $\rho_{X,Y}$ is defined by:

$$\rho_{X,Y} = \frac{Cov(X, Y)}{\sigma_X \sigma_Y}$$

If a and c are either both positive or both negative, $Corr(aX + b, cY + d) = Corr(X, Y)$

For any two random variables X and Y , $-1 \leq Corr(X, Y) \leq 1$.

Independence

Two random variables X and Y are said to be independent if and only if one of the following holds:

(i) for every pair of x and y values:

$$p(x, y) = p_X(x) p_Y(y) \text{ if } X \text{ is discrete and } f(x, y) = f_X(x) f_Y(y) \text{ if } X \text{ is continuous.}$$

(ii) for every pair of functions g and h

$$E[g(X) h(Y)] = E[g(X)] E[h(Y)]$$

(iii) for every pair of t_1 and t_2 values:

$$M_{X,Y}(t_1, t_2) = M_X(t_1) M_Y(t_2)$$

Conditional Distributions in the continuous case

Let X and Y be two random variables with joint pdf $f(x, y)$ and marginal X pdf $f_X(x)$. For any x -value for which $f_X(x) > 0$, the conditional probability density function of Y given that $X = x$ is:

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

Conditional expectation and variance

For two jointly distributed random variables X and Y :

$$E[X | Y = y] = \int_{-\infty}^{\infty} x f_{X|Y=y}(x) dx = g(y)$$

$$E[X | Y] = g(Y)$$

$$Var[X | Y] = E[X^2 | Y] - E[X | Y]^2$$

$$E[X] = E[E[X | Y]]$$

$$Var[X] = E[Var[X | Y]] + Var[E[X | Y]]$$