Some Probability Rules

The probability of complement of A

$P(A') = 1 - P(A)$

If A and B are independent events,

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

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

If A and B are mutually exclusive events,

$P(A \cap B) = 0 \; since \; (A \cap B) = \emptyset$

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

The conditional probability of A for given B is

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Bayes' Rule

If the events $B_1, B_2, \ldots, B_k$ constitute a partition of the sample space *S* such that $P(B_i) \neq 0$ for *i* = 1, 2, . . . , *k*, then for any event *A* in *S* such that $P(A) \neq 0$,

$$P(B_r|A) = \frac{P(B_r \cap A)}{P(A)} = \frac{P(B_r)P(A\,|B_r)}{\sum_{i=1}^{k} P(B_i \cap A)}$$

where

$$P(A) = \sum_{i=1}^{k} P(B_i \cap A) = P(B_1 \cap A) + P(B_2 \cap A) + \ldots + P(B_k \cap A)$$
$$= P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + \cdots + P(B_k)P(A|B_k)$$

Discrete Probability Distributions

The set of ordered pairs (*x*, *f*(*x*)) is a **probability function**, **probability mass function**, or **probability distribution** of the discrete random variable *X* if, for each possible outcome *x*,

1. $f(x) \geq 0$
2. $\sum_x f(x) = 1$
3. $P(X = a) = f(a)$

The **cumulative distribution function** *F*(*x*) of a discrete random variable *X* with probability distribution *f*(*x*) is

$F(a) = P(X \leq a) = \sum_{t \leq a} f(t)$ for $-\infty < a < \infty$

Mathematical Expectation

Let *X* be a discrete random variable with probability distribution *f*(*x*). The expected value of the *g*(*X*) is

$$\mu_{g(x)} = E[g(x)] = \sum_x g(x)f(x)$$

Let *X* be a discrete random variable with probability distribution *f*(*x*). The expected value of the *X* is

$$\mu = E[X] = \sum_x xf(x)$$

Let *X* be a discrete random variable with probability distribution *f*(*x*) and mean μ. The variance of *X* is

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

Binomial Distribution

Probability of x successes in n trials with P(Success)=p

$X \sim b(x;n,p)$, $\quad P(X=x) = b(x;n,p) = \binom{n}{x} p^x q^{n-x}$ for $x = 0,1,\ldots,n$

The mean and standard deviation of the binomial distribution $b(x;n,p)$ are $\mu = np$ and $\sigma = \sqrt{npq}$

Approximation of Binomial Distribution by a Poisson Distribution

if n is large and p is close to 0, the Poisson dist. can be used, with $\mu = np$, to approximate binomial dist.

Hypergeometric Distribution

Probability of x successes in a random sample of size n selected from N items of which k are labeled success and $N - k$ labeled failure, is

$X \sim h(x;N,n,k)$, $\quad P(X=x) = h(x;N,n,k) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$ for $max\{0, n-(N-k)\} \leq x \leq min\{n,k\}$

Approximation of Hypergeometric Distribution by a Binomial Distribution

if n is small compared to N, the nature of the N items changes very little in each draw. So a binomial distribution can be used to approximate the hypergeometric distribution when n is small compared to N. In fact, as a rule of thumb, the approximation is good when $n/N \leq 0.05$.

Negative Binomial Distribution

Probability of the k^{th} success occurs at x^{th} trial is

$X \sim b^*(x;k,p)$, $\quad P(X=x) = b^*(x;k,p) = \binom{x-1}{k-1} p^k q^{x-k}$ for $x = k, k+1, k+2, \ldots$

Geometric Distribution

Probability of the *first* success occurs at x^{th} trial is

$X \sim g(x,p)$, $\quad P(X=x) = g(x,p) \quad = pq^{x-1}$ for $x = 1, 2, \ldots$

The mean and standard deviation of the geometric distribution $g(x,p)$ are $\mu = \frac{1}{p}$ and $\sigma = \frac{\sqrt{1-p}}{p}$

Poisson Distribution

The probability distribution of the Poisson random variable X, representing the number of outcomes occurring in a given time interval or specified region

$X \sim p(x;\lambda)$, $\quad P(X=x) = p(x;\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}$ for $x = 0, 1, 2, \ldots$

where λ is the average number of outcomes per unit time, distance, area, or volume.

Both the mean and the variance of the Poisson distribution $p(x;\lambda)$ are λ.

Continuous Probability Distributions

The function *f*(*x*) is a **probability density function** (pdf) for the continuous random variable *X*, defined over the set of real numbers, if

4. $f(x) \geq 0$
5. $\int_{-\infty}^{\infty} f(x)dx = 1$
6. $P(a < X < b) = \int_{a}^{b} f(x)dx$

The **cumulative distribution function** *F*(*x*) of a continuous rv *X* with probability density function *f*(*x*) is

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

Mathematical Expectation

Let *X* be a continuous random variable with probability distribution *f*(*x*).

The expected value of the $g(X)$ is $\mu_{g(x)} = E[g(x)] = \int_{-\infty}^{\infty} g(x)f(x)dx$

The expected value of the X is $\mu = E[X] = \int_{-\infty}^{\infty} xf(x)dx$

The variance of X is $\sigma^2 = E[X^2] - \mu^2$

The Joint Density Function

- The function *f*(*x, y*) is a joint density function of the continuous random variables *X* and *Y* if

1. $f(x, y) \geq 0$ for all (x, y)
2. $\iint_{-\infty}^{\infty} f(x, y)dA = 1$
3. $P[(X, Y) \in R] = \iint_{\boldsymbol{R}} f(x, y)dA$ for any region R in the xy-plane.

- The **marginal distributions** of *X* alone and of *Y* alone are

$$g(x) = \int_{-\infty}^{\infty} f(x, y)dy \text{ and } h(y) = \int_{-\infty}^{\infty} f(x, y)dx \text{ for the continuous case.}$$

- Let *X* and *Y* be two random variables, discrete or continuous. The **conditional distribution** of the random variable *Y* given that *X* = *x* is

$$f(x|y) = \frac{f(x, y)}{h(y)}, \qquad g(x) > 0$$

- Similarly, the conditional distribution of *X* given that *Y* = *y* is

$$f(y|x) = \frac{f(x, y)}{g(x)}, \quad h(y) > 0$$

- The random variables *X* and *Y* are said to be **statistically independent** if and only if
  f(*x, y*) = *g*(*x*)*h*(*y*) for all (*x, y*) within their range.

Mathematical Expectation

Let *X* and *Y* be random variables with joint probability distribution *f*(*x, y*).

- The mean, or expected value, of the random variable *g*(*X, Y*) is

$$\mu_{g(x,y)} = E[g(x,y)] = \iint_{-\infty}^{\infty} g(x,y)f(x,y)dA$$

- The covariance of *X* and *Y* is

$$\sigma_{xy} = E[XY] - \mu_x\mu_y$$

where $\mu_x = E[X] = \int_{-\infty}^{\infty} xg(x)dx$ and $\mu_y = E[Y] = \int_{-\infty}^{\infty} yh(y)dy$

- The correlation coefficient of X and Y is $\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x\sigma_y}$

Uniform Distribution

- The density function of the continuous uniform random variable *X* on the interval [*A, B*] is

$$f(x;A,B) = \begin{cases} \frac{1}{B-A}, & A \le x \le B \\ 0, & elsewhere \end{cases}$$

- The mean and variance of the uniform distribution are

$$\mu = \frac{A+B}{2}\ , and\ \sigma^2 = \frac{(B-A)^2}{12}$$

Standard Normal distribution

we are able to transform all the observations of any normal random variable *X* into a new set of observations of a normal random variable *Z* with mean 0 and variance 1. This can be done by means of the transformation

$$Z = \frac{X-\mu}{\sigma}$$

$$So, P(x_1 < X < x_2) = P(z_1 < Z < z_2)$$

Normal Approximation to Binomial Distribution

If *X* is a binomial random variable with mean $\mu = np$ and variance $\sigma = \sqrt{npq}$, then the limiting form of the distribution of

$$Z = \frac{Y - np}{\sqrt{npq}}$$

as $n \to \infty$, is the standard normal distribution *n*(*z*; 0*,* 1), where Y is the upper or lower real limit of X.
The approximation will be good if *np* and *n*(1−*p*) are greater than or equal to 5.

Exponential Distribution

The continuous random variable *X* has an **exponential distribution**, with parameter β, if its density function is given by

$$f(x;\beta) = \begin{cases} \frac{1}{\beta}e^{\frac{-x}{\beta}}, & x > 0 \\ 0, & elsewhere \end{cases}$$

where $\beta > 0$ and β denotes the average time.
The mean and standard deviation of the exponential distribution are $\mu = \beta\ and\ \sigma = \beta$.

The Moment Generating Function

$$M_t(x) = E[e^{xt}] = \sum_x e^{xt} f(x) \; for\ discrete\ random\ variables$$

$$M_t(x) = E[e^{xt}] = \int_{-\infty}^{\infty} g(x) f(x) dx \; for\ continuous\ random\ variables$$

$$n^{th}\ moment\ \mu_n = E[X^n] = \frac{d^n M_t(x)}{dt^n}|_{t=0}$$

The Joint Probability Distribution Function

- The function $f(x, y)$ is a joint Probability Distribution Function of the Discrete random variables X and Y if

  1. $f(x, y) \geq 0$ for all (x, y)
  2. $\sum_x \sum_y f(x, y) = 1$
  3. $P[(X, Y) \in R] = \sum_x \sum_y f(x, y)$ for any region R in the xy-plane.

- The **marginal distributions** of X alone and of Y alone are

  $g(x) = \sum_y f(x, y)$ and $h(y) = \sum_x f(x, y)$ for the discrete case.

- Let X and Y be two random variables, discrete or continuous. The **conditional distribution** of the random variable Y given that $X = x$ is

$$f(x|y) = \frac{f(x, y)}{h(y)}, \qquad g(x) > 0$$

- Similarly, the conditional distribution of X given that $Y = y$ is

$$f(y|x) = \frac{f(x, y)}{g(x)}, \quad h(y) > 0$$

- The random variables X and Y are said to be **statistically independent** if and only if

  $f(x, y) = g(x)h(y)$ for all (x, y) within their range.

The Hypotheses Testing

i. State the null and alternative hypotheses.
ii. Set the critical region
iii. Compute the test statistics
iv. Draw scientific or engineering conclusions.

H_0	Value of Test Statistic	H_1	Critical Region
$\mu = \mu_0$	$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$; σ known	$\mu < \mu_0$ $\mu > \mu_0$ $\mu \neq \mu_0$	$z < -z_\alpha$ $z > z_\alpha$ $z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$
$\mu = \mu_0$	$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$; $v = n - 1$, σ unknown	$\mu < \mu_0$ $\mu > \mu_0$ $\mu \neq \mu_0$	$t < -t_\alpha$ $t > t_\alpha$ $t < -t_{\alpha/2}$ or $t > t_{\alpha/2}$
$\mu_1 - \mu_2 = d_0$	$z = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$; σ_1 and σ_2 known	$\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 \neq d_0$	$z < -z_\alpha$ $z > z_\alpha$ $z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$
$\mu_1 - \mu_2 = d_0$	$t = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{s_p\sqrt{1/n_1 + 1/n_2}}$; $v = n_1 + n_2 - 2$, $\sigma_1 = \sigma_2$ but unknown, $s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$	$\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 \neq d_0$	$t < -t_\alpha$ $t > t_\alpha$ $t < -t_{\alpha/2}$ or $t > t_{\alpha/2}$

where s_p^2 is pooled variance.