Chapter 3.

3.1. Distribution of a Random variable

A random variable is a function of an outcome,

$$X = f(\omega).$$

In other words, it is a quantity that depends on chance. collection of all the probabilities related to X is the distribution of X.

Probability mass function (pmf): $p(x) = P\{X = x\}$

Cumulative distribution function (cdf):

$$F(x) = P\{X \le x\} = \sum_{y \le x} P(y)$$

The set of possible values of X is called the support of the distribution F.

Types of Random variables:

i). discrete random variable.

ii) continuous random variable

iii). mixed random variable.

3.2. Distribution of a Random vector.

If X and Y are random variables, then the pair (X, Y) is a random vector.

Its distribution is called the joint distribution of X and Y.

Individual distributions of X and Y are called the marginal distributions.

Addition Rule $\begin{cases} P_X(x) = P\{X = x\} = \sum_y P_{(X,Y)}(x, y) \\ P_Y(y) = P\{Y = y\} = \sum_x P_{(X,Y)}(x, y). \end{cases}$

In general, the joint distribution $P_{(X,Y)}$ cannot be computed from marginal distributions because they carry no information about interrelations between random variables.

Random variable X and Y are independent if

$$P_{(X,Y)}(x, y) = P_X(x) \cdot P_Y(y)$$

for all values of x and y.

3.3. Expectation and variance.

Expectation or expected value of a random variable X is its mean, the average value. (a weighted average)

$$\mu = E(X) = \sum_x x \cdot P(x)$$

Properties of expectations:

For any random variables X and Y, and any non-random numbers a, b, c. we have:

i) $E(aX + bY + c) = aE(X) + bE(Y) + c$

ii) $E(X + Y) = E(X) + E(Y)$

$E(aX) = a \cdot E(X)$

$E(c) = c$

iii) For independent X and Y, $E(XY) = E(X) \cdot E(Y)$.

Variance :

variance of a random variable is defined as the expected squared deviation from the mean.

$$\sigma^2 = Var(X) = E(X - EX)^2 = \sum_{X} (X-\mu)^2 \, P(X)$$

Expectation

Standard deviation :

$$\sigma = Std(X) = \sqrt{Var(X)}$$

Covariance $\sigma_{xy} = cov(X, Y)$ is defined as

$$Cov(X, Y) = E\{(X - EX)(Y - EY)\}$$
$$= E(XY) - E(X)E(Y).$$

It summarizes interrelation of two random variables.

i) if $Cov(X, Y) > 0$, X and Y are positively correlated.
ii) if $cov(X, Y) < 0$, X and Y are negatively correlated.
iii) if $Cov(X, Y) = 0$, X and Y are uncorrelated.

Correlation coefficient between variables X and Y is defined

as $\rho = \dfrac{Cov(X, Y)}{(Std\, X)(Std\, Y)}$ $\quad (-1 \le \rho \le 1)$.

where $|\rho| = 1$ is possible only when all values of X and Y lie on a straight line.

$\rho \to 1$ indicate strong positive correlation.
$\rho \to -1$ indicate strong negative correlation.
$\rho \to 0$ indicate weak correlation or no correlation

Properties of variances and covariances:

$$Var(aX+bY+c) = a^2 Var(X) + b^2 Var(Y) + 2ab\,Cov(X,Y)$$

$$Cov(aX+bY,\ cZ+dW)$$
$$= ac\,Cov(X,Z) + ad\,Cov(X,W) + bc\,Cov(Y,Z) + bd\,Cov(Y,$$

$$Cov(X,Y) = Cov(Y,X)$$

$$\rho(X,Y) = \rho(Y,X)$$

In particular,

$$Var(aX+b) = a^2 Var(X)$$

$$Cov(aX+b,\ cY+d) = ac\,Cov(X,Y)$$

$$\rho(aX+b,\ cY+d) = \rho(X,Y).$$

For independent X and Y,

$$Cov(X,Y) = 0$$

$$Var(X+Y) = Var(X) + Var(Y)$$

Chebyshev's inequality:

expectation

$$P\{|X-u| > \varepsilon\} \le \left(\frac{\delta}{\varepsilon}\right)^2$$

any positive

any distribution.

Proof: $\delta^2 = \sum_{\text{all } x} (x-u)^2 p(x) \ge \sum_{\substack{\text{only } x \text{ with} \\ |x-u|>\varepsilon}} (x-u)^2 p(x) \ge \sum_{x:|x-u|>\varepsilon} \varepsilon^2 p(x)$

$$= \varepsilon^2 \sum_{x:|x-u|>\varepsilon} p(x)$$

$$= \varepsilon^2 P\{|X-u| > \varepsilon\}$$

Therefore, $P\{|X-u| > \varepsilon\} \le \left(\frac{\delta}{\varepsilon}\right)^2$

3.4. Families of discrete distributions.

Bernoulli distribution : (with two possible values , 0 and 1). $\to$ for one trio

If $P(1) = p$, then $P(0) = 1-p = q$.

$$E(X) = \sum_x p(x) = 0 \cdot (1-p) + 1 \cdot p = p$$

$$\begin{aligned} Var(X) &= \sum_x (x-p)^2 p(x) = (0-p)^2 \cdot (1-p) + (1-p)^2 p \\ &= p(1-p) \\ &= p \cdot q \end{aligned}$$

Binomial distribution : (binary outcomes : 0 or 1)

mass function :

$$p(x) = P\{X=x\} = \binom{n}{x} p^x \cdot q^{n-x}, \quad x = 0, 1, \cdots, n.$$

$\downarrow$

the probability of exactly x successes in n trials.

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

To compute a pmf instead of a cdf. ($F(x) = P(X \le x)$)

cumulative distribution function.

$$p(x) = F(x) - F(x-1)$$

$$\begin{cases} E(X) = np. \\ Var(X) = npq. \end{cases}$$

$$F(x) = P\{X \le x\} = \sum_{w=0}^{x} C(n,w) \cdot p^w \cdot q^{n-w}$$

※ For several values of p, n, x, the value of $F(x)$ is in the Binomial distribution table.

(P412)

Geometric distribution: the number of Bernouli trials needed to get the first success.

Geometric probability mass function:

$$p(x) = P\{\text{the 1st success occurs on the } x\text{-th trial}\}$$
$$= (1-p)^{x-1} \cdot p \qquad x = 1, 2, \cdots$$

Expectation $E(x) = \mu = \frac{1}{p}$.

variance $\delta^2 = \frac{(1-p)}{p^2}$

Negative Binomial distribution

probability mass function

$$p(x) = P\{\text{the } x\text{th trial results in the } k\text{th success}\}$$
$$= P\left\{\begin{array}{l}(k-1) \text{ successes in the first } (x-1) \text{ trials,} \\ \text{and the last trial is a success}\end{array}\right\}$$
$$= \binom{x-1}{k-1} (1-p)^{x-k} \cdot p^k$$

$E(x) = \frac{k}{p}$;

$var(x) = \frac{k(1-p)}{p^2}$.

Poisson distribution: the number of rare events occuring within a fixed period of time has poisson distribution.

λ = frequency, average number of events.

$$P(X) = e^{-\lambda} \cdot \frac{\lambda^x}{x!}, \quad x = 0, 1, 2, \cdots$$

$$\begin{cases} E(X) = \lambda. \\ Var(X) = \lambda. \end{cases}$$

Cumulative distribution function $F(x)$:

i). $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to +\infty} F(x) = 1$.

ii). $F(x) = P(X \le x) = \sum_{x_i \le x} P(X = x_i) = \sum_{x_i \le x} P(x_i)$

iii) $F(b) - F(a) = P(a < X \le b) = \int_a^b f(x)\, dx$

Poisson approximation to Binomial:

Binomial $(n, p) \approx$ Poisson (λ),

where $n \ge 30$, $p \le 0.05$, $np = \lambda$

P415

For moderate p $(0.05 \le p \le 0.95)$, the poisson approximation may not be accurate.