MTH219 Summary Sheet

Tags

Created by Ho Han Sheng

- Sample mean

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{X_1 + X_2 + ... + X_n}{n}$$

- Population mean

$$\mu = \frac{\sum_{i=1}^{N} X_i}{N} = \frac{X_1 + X_2 + ... + X_N}{N}$$

- Median:

$$x_{(\frac{1}{2}(n+1))}$$

$Range = X_{largest} - X_{smallest}$

- Sample Variance

$$s^2 = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n-1}$$

- Population Variance

$$\sigma^2 = \frac{\sum_{i=1}^{N} (X_i - \mu)^2}{N}$$

- Sample Standard Deviation

$$s = \sqrt{\frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n-1}}$$

▼ Population Standard Deviation

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}}$$

Lower sample quartile (25th percentile) / first quartile (Q_L) --- Q1

$$Q_L = x_{(\frac{1}{4}(n+1))}$$

Upper sample quartile (75th percentile) / third quartile (Q_U) --- Q3

$$Q_U = x_{(\frac{3}{4}(n+1))}$$

The Interquartile range (iqr) is the difference between (Q_U) & (Q_L)

▼ Coefficient of variation (cv)

$$cv = s/\overline{X}$$

Sample skewness of a data sample $X_1, X_2, ..., X_n$ is given by

$$\frac{n}{(n-1)(n-2)} \sum_{i=1}^{n} (\frac{X_i - \overline{X}}{s})^3$$

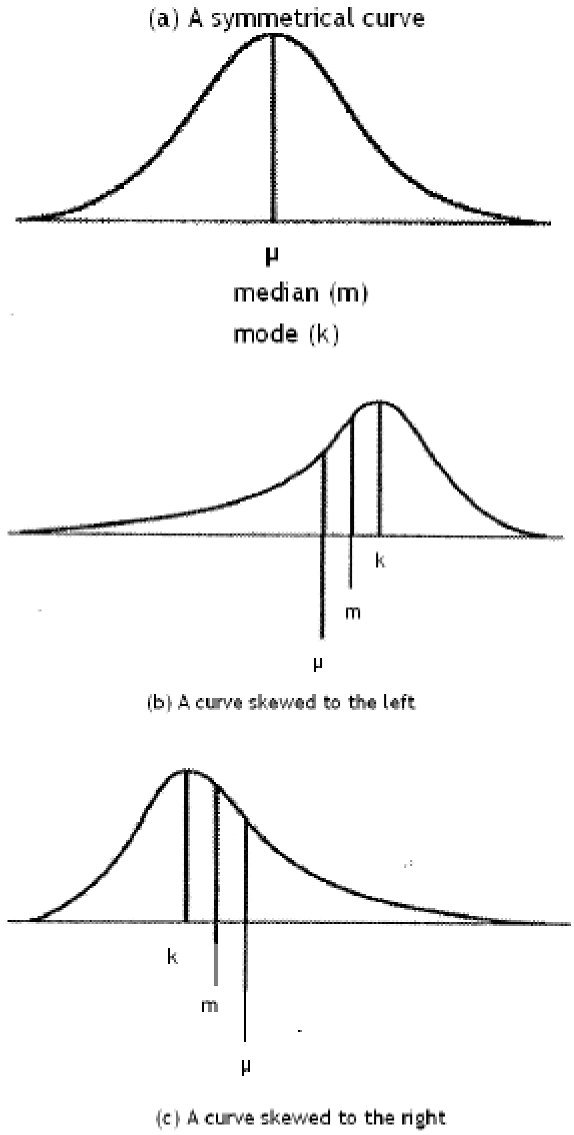

Figure 1.13 Relationship between mean, median and mode

(Source: Ho, Xie & Goh, 2011, p.69)

If an event A is impossible, that means it never happens

$$P(A) = 0$$

If an event A always happens, then

$$P(A) = 1$$

For any event A, the complement event is denoted A^c

$$P(A^c) = 1 - P(A)$$

$P(A \cap B)$ is the probability that both events A and B occur simultaneously

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

Consists of the outcomes that are contained within at least one of the events A and B

- Mutually exclusive

  Happens when events A and B have no outcomes in common

  A and B cannot happen simultaneously

  $A \cap B = \emptyset$

  $P(A \cap B) = 0$

  Hence,

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

- Independent

  Events A and B are independent if the occurrence of one event does not alter the probability of the other event occurring

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

  ^ Probability of A given B

  If A and B are independent, then

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

  Probability of both independent events occurring simultaneously is the product of both probabilities

The conditional probability of an event B given an event A

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

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A^C)P(A^C) + P(B|A)P(A)}$$

From,

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

we can get

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

Yes. The complement rule holds for conditional probabilities.

$$\begin{aligned} \Pr(B) &= \Pr((A \cap B) \cup (A' \cap B)) && \text{by total probability law} \\ &= \Pr(A \cap B) + \Pr(A' \cap B) && \text{because of mutual exclusion} \\ \implies \Pr(A \cap B) &= \Pr(B) - \Pr(A' \cap B) && \text{by rearangement} \\ \therefore \Pr(A \mid B) &= 1 - \Pr(A' \mid B) && \text{by division by } \Pr(B) \end{aligned}$$

▼ Discrete random variables

$E(X) = \mu$

$$E(X) = \sum_{i=1}^{\infty} x_i p(x_i) = \mu$$

$$E(X) = X_1 P(X_1) + X_2 P(X_2) + ... + X_n P(X_n)$$

$V(X) = \sigma^2$

$$V(X) = E[(X - \mu^2)] = \sum_{i=1}^{n} (x - \mu)^2 \cdot P(x)$$

$$= (X_1 - \mu)^2 \cdot P(X_1)$$

Alternatively,

$$V(X) = \sum_{i=1}^{n} X_i^2 \ \cdot \ P(X_i) - \mu^2$$

$$= \{[X_1^2 \cdot P(X_1)] + [X_2^2 \cdot P(X_2)] + ... + [X_n^2 \cdot P(X_n)]\} - \mu^2$$

- Continuous random variables

$$\mu = E(X) = \int_x x f(x)\, dx$$

Where the integral is taken over all values in the range of X

$$\sigma^2 = V(X) = E[(X-\mu)^2] = \int\limits_x (x-\mu)^2 f(x)\, dx$$

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

For continuous random variables only,

$P(X \geq 10) = 1 - P(X < 10) = 1 - P(X \leq 10)$

- Bernoulli distribution

If X is a Bernoulli r.v. with parameter p then,

$P(X = 1) = p$

$P(X = 0) = 1 - p$

$X \sim$ Bernoulli (p)

$\mu = E(X) = p$

$\sigma^2 = V(X) = p(1-p)$

- Binomial distribution

The total number of successes out of the n <u>independent Bernoulli trials</u>

$X \sim B(n, p)$

If $X \sim B(n, p)$ then

$$P(X = r) = {}^nC_r p^r (1-p)^{n-r}$$

for $r = 0, 1, 2, ..., n$

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

$$n! = n \cdot (n-1) \cdot (n-2) \cdot \ ... \ \cdot (3) \cdot (2) \cdot (1)$$

$\mu = E(X) = np$

$\sigma^2 = V(X) = np(1-p)$

▼ Conditions for binomial experiment

- There are n “trials” where n is determined in advance and is not a random value.
- Two possible outcomes on each trial, called “success” and “failure” and denoted S and F.
- Outcomes are independent from one trial to the next.
- Probability of a “success”, denoted by p, remains the same from one trial to the next.
- Probability of “failure” is 1 – p.

▼ Geometric Distribution

The number of trials up to and including the first success in a sequence of independent Bernoulli trials is a random variable N

$N \sim G(p)$

With a probability mass function (pmf) given by:

$$P(N) = P(N=n) = q^{n-1}p$$

For $n = 1, 2, ...,$

*n starts from 1 as it does not make sense if attempt 0 is successful

Where $q = 1 - p$,

$$q^{n-1}p = p(1-p)^{n-1}$$

▼ Cumulative distribution function (cdf)

$$F(N) = P(N \leq n) = 1 - q^n$$

- Mean

$$E(N) = 1/p$$

- Variance

$$V(N) = (1 - p)/p^2$$

If the parameter p of the geometric model is high, e.g. in the range from 0.7 - 0.9, this implies that it is quite unlikely you have to wait long for the first successful trial

Since p represents the probability of success, when the value p is much lower, e.g. in the range from 0.1- 0.3, we would expect the wait to be longer for the first success to occur

- Uniform Distribution

A discrete uniform r.v. if X can assume a set of values with the same probability for each value.

$$P(X = x) = 1/n$$

$x = 1, 2, ..., n$
$E(X) = \frac{n+1}{2}$
$V(X) = \frac{1}{12}(n^2 - 1)$

- Continuous Uniform Random Variable

  Common for when we have no *a priori* knowledge favouring the distribution of outcomes except for the end points

  Outcomes are constrained to lie within a known interval and all outcomes are equally likely to occur

  i.e. we do not know when a business call will come, but it must come, say, between 10am and 2pm.

  X is a continuous uniform random variable on the interval (a, b) if and only if the probability function of X is given by

$$f(x) = \frac{1}{b - a}$$

Where $a < x < b$

$X \sim U(a, b)$

$$E(X) = \frac{a+b}{2}$$

$$V(X) = \frac{1}{12}(b-a)^2$$

- Poisson distribution

  Suppose the rate of occurence λ is expressed in number per unit time

  At a fixed time t, the total number of occurrence is given by $N(t)$

  Let $X = N(t)$

  The number of events X that occur in an interval of length t has a Poisson distribution

  $$P(N(t) = r) = P(X = r) = \frac{e^{-\lambda t}(\lambda t)^r}{r!}$$

  Where r = 0, 1, 2, ...,

  $X \sim Poisson(\lambda t)$

  $E(X) = \lambda t$

  $V(X) = \lambda t$

  - Conditions

    The number of outcomes occurring in a given time interval (for example) is independent of the number of outcomes in any other disjoint time interval

    Usually used for calculations involving density

    e.g. defects per wafer, failures per period, accidents per hour

  - Rules

    Consider 2 disjoint time intervals, which are independent

    X_1 is the number of occurrences in the 1st time interval

    X_2 is the number of occurrences in the 2nd time interval

    Such that

$X_1 \sim Poisson(\lambda_1 t)$
$X_2 \sim Poisson(\lambda_2 t)$

Then,

$$X_1 + X_2 \sim Poisson((\lambda_1 + \lambda_2)t)$$

▼ Poisson Approximation to Binomial Model

Distribution of X when

$X \sim B(n, p)$

can be well approximated by the Poisson distribution with mean

$\mu = np = \lambda t$

When n is large and parameter p is small

▼ Poisson's Approximation for Rare Events

- For binomial random variable, the approximation improves as n increases
- If n is large and p is small (e.g. $n \geq 50$ & $p \leq 0.05$), the approximation is good
- When p is small enough, the approximation is good even for quite small values of n

▼ Exponential distribution

The probability distribution of the time between events in a Poisson process

Random variable T follows an exponential distribution with parameter λ if the probability density function (pdf) of T is given as:

$$f(t) = \lambda e^{-\lambda t}$$

Where $t \geq 0$

▼ Cumulative distribution function

cdf of t can be determined as follows

$$F(t) = 1 - e^{-\lambda t}$$

- Mean

$$\mu = E(T) = \frac{1}{\lambda}$$

- Variance

$$\sigma^2 = V(T) = \frac{1}{\lambda^2}$$

Hence in the exponential model,

$\mu = \sigma$

- Relationship between Poisson and Exponential Models

  $\lambda \to \frac{1}{\lambda}$

- Normal distribution

  $X \sim N(\mu, \sigma^2)$

  - Symmetric distribution
  - Mean μ is at the center
  - Data set normally distributed has one mode (unimodal)
  - $f(x) > 0$
  - $P(a < x < b) = \int_a^b f(x)dx$

- General Rules (for any random variables)

  - Mean of sum of random variables is equal to the sum of mean of random variables

  $$E(X_1 + X_2 + ... + X_n) = E(X_1) + E(X_2) + ... + E(X_n)$$

  - If the random variables are independent, then the variance of their sum is equal to the sum of their variance

  $$V(X_1 + X_2 + ... + X_n) = V(X_1) + V(X_2) + ... + V(X_n)$$

  - If $Y = aX + b$, then

    $E(Y) = E(aX + b) = aE(X) + b$

$$V(Y) = V(aX + b) = a^2 V(X)$$

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

- A binomial random variable may be regarded as a sum of independent Bernoulli random variables each with the same parameter
- If X and Y are normally distributed, then $X - Y$ also follows a normal distribution
- The mean of the difference between two random variables is equal to the difference between their means:

$$E(X - Y) = E(X) - E(Y)$$

- If the random variables are independent, then the variance of their difference is equal to the sum of their variances.

$$V(X - Y) = V(X) + V(Y)$$

Common Functions	Function	Integral
Constant	$\int a\,dx$	$ax + C$
Variable	$\int x\,dx$	$x^2/2 + C$
Square	$\int x^2\,dx$	$x^3/3 + C$
Reciprocal	$\int (1/x)\,dx$	$\ln\lvert x\rvert + C$
Exponential	$\int e^x\,dx$	$e^x + C$
	$\int a^x\,dx$	$a^x/\ln(a) + C$
	$\int \ln(x)\,dx$	$x \ln(x) - x + C$
Trigonometry (x in radians)	$\int \cos(x)\,dx$	$\sin(x) + C$
	$\int \sin(x)\,dx$	$-\cos(x) + C$
	$\int \sec^2(x)\,dx$	$\tan(x) + C$
Rules	**Function**	**Integral**
Multiplication by constant	$\int cf(x)\,dx$	$c\int f(x)\,dx$
Power Rule ($n \neq -1$)	$\int x^n\,dx$	$\frac{x^{n+1}}{n+1} + C$
Sum Rule	$\int (f + g)\,dx$	$\int f\,dx + \int g\,dx$
Difference Rule	$\int (f - g)\,dx$	$\int f\,dx - \int g\,dx$
Integration by Parts	See Integration by Parts	
Substitution Rule	See Integration by Substitution	