الجمهورية العربية المتحدة

معهد التخطيط القومي — نسخة واحدة فقط

Memo. No. 59o.

Note On Statistical Methods

By

Dr. Moharram W. Mahmoud

B.Sc., M.Sc., M.A., Ph.D.

Operations Research Centre.

10/8/1965.

القاهرة

٣ شارع محمد مظهر - بالزمالك

Contents

This note has been prepared by the author for a course in Statistical methods which is especially designed for those of mathematical background as engineers and scientists.

It was given as lectures for

(i) the group of operations research of the 4th long term training course 1965.

&(ii) the group who attended the short term training period (Jan. - Feb. 1965) at the Operations Research Centre.

The author takes this opportunity to acknowledge the encouragement of Dr Salah Hamid; the general director of the Operations Research Centre.

He would also like to thank Mrs. Nadia Amer and Miss Ellen Zaki for the great care they took in writing this note.

Moharram W. Mahmoud

(1)

(I) MOMENTS

(1-1) <u>Moments around the origin</u>:-

$$m_r'(x) = \frac{1}{n} \sum x_i^r \qquad \text{(ungrouped data)}$$

$$= \frac{1}{n} \sum f_i \, x_i^r \qquad \text{(grouped data)}$$

For grouped data x_i denote the mid points of the intervals and f_i the corresponding frequencies

(1-2) <u>Moments around the mean</u>:-

$$m_r(x) = \frac{1}{n} \sum (x_i - \bar{x})^r \qquad \text{(ungrouped data)}$$

$$= \frac{1}{n} \sum f_i (x_i - \bar{x})^r \qquad \text{(grouped data)}$$

<u>Remarks</u> (i) $m_o'(x) = m_o(x) = 1$

(ii) $m_1'(x) = \bar{x}$ = the mean

(iii) $m_1(x) = o$

(iV) $m_2(x) = s^2$ = the variance.

(1.3) <u>Expressing $m_r(x)$ in terms of $m_r'(x)$</u>

$$m_r(x) = \sum_{t=o}^{r} (-1)^t \binom{r}{t} (m_1')^t \, m_{r-t}'$$

In particular

$$m_2 = m_2' - m_1'^2$$

$$m_3 = m_3' - 3m_2' \, m_1' + 2m_1'^3$$

$$m_4 = m_4' - 4m_3' \, m_1' + 6m_2' \, m_1'^2 - 3m_1'^4$$

Remark In the case of frequency distributions with equidistant intervals we replace the mid-points x_i by d_i where

$$d_i = \frac{x_i - a}{\ell}$$

where "a" is an arbitrary origin

"ℓ" is the length of the interval

Then

$$m_r(x) = \ell^r \, m_r(d)$$

(1.4) Measures of skewness

$$\beta_1 = m_3^2 \,/\, m_2^3$$

$$\gamma_1 = \sqrt{\beta_1}$$

Measure of Kurtosis

$$\beta_2 = m_4/m_2^2$$

$$\gamma_2 = \beta_2 - 3$$

(1.5) Numerical Example (illustrative

Intervals	f_i	d_i	d_i^2	d_i^3	d_i^4
10-20	10	-1	1	-1	1
20-30	18	0	0	0	0
30-40	14	1	1	1	1
40-50	8	2	4	8	16
	50	20	56	72	152
$m_r(d)$		$\sum f_i d_i$	$\sum f_i d_i^2$	$\sum f_i d_i^3$	$\sum f_i d_i^4$
		0.4	1.02	1.44	3.04

$m_2(d) = 1.02 - (0.4)^2 = 0.86$

$m_3(d) = 1.44 - 3(1.02)(0.4)+2(0.4)^3 = 0.344$

$m_4(d) = 3.04 - 4(1.44)(0.4)+6(1.02)(0.4)^2-3(0.4)^4 = 1.6394$

$m_1(x) = \ell\, m_1'(d) + a = 10(0.4) + 25 = 29$

$m_2(x) = S^2 = 10^2\ (0.86) = 86$

$m_3(x) = 10^3\ (0.344) = 344$

$m_4(x) = 10^4\ (1.6394) = 16394$

$\beta_1 = (344)^2/(86)^3 = 0.1860$

$\gamma_1 = \sqrt{0.1860} = 0.43$

& $\beta_2 = 16394/7396 = 2.22$

$\therefore \gamma_2 = \beta_2 - 3 = 2.22-3 = -0.78$

<u>Exercise</u> given the following frequency table, calculate the mean, the variance and β_1 & β_2

Intervals	f	Intervals	f
55 - 75	3	155 - 175	209
75 - 95	21	175 - 195	81
95 -115	78	195 - 215	21
115 -135	182	215 - 235	5
135 -155	305		905

(II) FUNDAMENTAL CONCEPTS OF PROBABILITY

(2-1) The theory of probability is a branch of applied mathematics dealing with the effects of chance. If we throw a die upon a board we are certain that one of the six faces will turn up, but whether a particular face will show, depends on what we call chance. Also, if equal numbers of white and black balls are put in an urn and we draw one of them blindly, we are certain that its colour will be either white or black, but whether it will be black, that depends on chance.

The word event which we are going to use frequently is used to signify an observation satisfying some specified conditions.

Two events are said to be "equally likely" if after taking into consideration all relevent evidence one of them cannot be expected in preference to the other. e.g. in the case of the urn with equal number of white and black balls, if we draw a ball, it is equally likely to be either white or black.

In the field of statistical analysis there would seem to be two definitions:

(i) Mathematical theory of arrangements which is as old as gambling & playing cards. The probability (p) of an event is the ratio of the no. of ways in which the event may happen divided by the total no. of ways in which the event may or may or may not happen. This is under the condition that all the events are equally likely. e.g. in the case of an unbiased coin, the probability that the head appears uppermost is $p=\frac{1}{2}$. Also in throwing a die the probability a particular face will show is $p=1/6$.

(ii) The frequency theory: If in a series of n independent trials which are absolutely identical, the event E is found to occur in m trials, then the probability of E is $\frac{m}{n}$.

This gives us a way to estimate probabilities from experimental results in a simple way.

As n increases $\frac{m}{n}$ tends to p, i.e. $p = \lim_{n \to \infty} \frac{m}{n}$.

(2-2) <u>Definition</u> (1) <u>Fundamental probability set</u> (F.P.S.). is that set of individuals or units from which the probability is calculated.

In the case of die, the F.P.S. given by the mathematical theory of arrangements would be 6. If the die is biassed in some way and it is necessary to estimate a probability from the observations, then the F.P.S. would be the total number of throws of a die.

<u>Definition (2) Mutually exclusive</u>: Two events E_1, E_2 are said to be mutually exclusive if no element of the F.P.S. may possess both E_1, E_2. In other words the two events do not occur together.

<u>Remarks</u> (i) $Pr\{E_1+E_2\}$ means the probability of E_1 or E_2.

(ii) $Pr\{E_1E_2\}$ means the probability of E_1 & E_2.

(2.3) <u>Basic theorems</u>:-

In the following theorems we are going to assume that the fundamental probability set N, consists of

n_1	elements possessing	E_1
n_2	" "	E_2
n_{12}	" "	E_1 & E_2
n_o	" "	$\overline{E}_1$ & $\overline{E}_2$

(where $\overline{E}_1$ means the event E_1 does not occur)

In other words $N = n_1+n_2+n_{12}+n_o$

Theorem (1) If E_1, E_2 are mutually exclusive and the only possible events, then

$$\Pr\left\{E_1\right\} + \Pr\left\{E_2\right\} = 1$$

Proof:- Since the two events are mutually exclusive then $n_{12} = o$. Also the two events are the only possible then $n_o = o$. Therefore

$$n_1 + n_2 = N$$

$$\therefore \frac{n_1}{N} + \frac{n_2}{N} = 1$$

$$\therefore \Pr\left\{E_1\right\} + \Pr\left\{E_2\right\} = 1$$

Theorem (2) $\Pr\left\{E_1 + E_2\right\} = \Pr\left\{E_1\right\} + \Pr\left\{E_2\right\} - \Pr\left\{E_1E_2\right\}$

$$\Pr\left\{E_1\right\} = \frac{n_1+n_{12}}{N} \quad , \quad \Pr\left\{E_2\right\} = \frac{n_2 + n_{12}}{N}$$

$$\Pr\left\{E_1E_2\right\} = \frac{n_{12}}{N}$$

$$\therefore \Pr\left\{E_1\right\} + \Pr\left\{E_2\right\} - \Pr\left\{E_1E_2\right\} = \frac{n_1+n_{12}}{N} + \frac{n_2+n_{12}}{N} - \frac{n_{12}}{N}$$

$$= \frac{n_1 + n_2 + n_{12}}{N}$$

$$= \Pr\left\{E_1 + E_2\right\}$$

Cor. If E_1,E_2 are mutually exclusive them

$$\Pr\left\{E_1 + E_2\right\} = \Pr\left\{E_1\right\} + P_r\left\{E_2\right\}$$

In general for k mutually exclusive properties we have

$$\Pr\left\{\sum_{i=1}^{k} E_i\right\} = \sum_{i=1}^{k}\left\{\Pr\left(E_i\right)\right\}$$

<u>Definition</u> Conditional probability of an event E_2 given event E_1 is the probability of E_2 referred to the F.P.S. for E_1 and it is written $P_r\{E_2 | E_1\}$

<u>Theorem (3)</u> $$Pr\{E_1E_2\} = Pr\{E_1\} \cdot Pr\{E_2 | E_1\}$$
$$= Pr\{E_2\} \, Pr\{E_1 | E_2\}$$

Proof:- $$Pr\{E_1\} = \frac{n_1 + n_{12}}{N}$$
$$Pr\{E_2\} = \frac{n_2 + n_{12}}{N}$$
$$Pr\{E_2 | E_1\} = \frac{n_{12}}{n_1 + n_{12}}$$
$$Pr\{E_1 | E_2\} = \frac{n_{12}}{n_2 + n_{12}}$$

$$\therefore Pr\{E_1\} \cdot Pr\{E_2 | E_1\} = \frac{n_1 + n_{12}}{N} \times \frac{n_{12}}{n_1 + n_{12}} = \frac{n_{12}}{N}$$
$$Pr\{E_2\} \, Pr\{E_1 | E_2\} = \frac{n_2 + n_{12}}{N} \times \frac{n_{12}}{n_2 + n_{12}} = \frac{n_{12}}{N}$$

$$\therefore Pr\{E_1\} \, Pr\{E_2 | E_1\} = Pr\{E_2\} \, Pr\{E_1 | E_2\}$$
$$= \frac{n_{12}}{N} = Pr\{E_1E_2\}$$

(2-4) <u>Independence:</u> E_1 is independent of E_2 if

$$Pr\{E_1\} = Pr\{E_1 | E_2\}.$$

Consequently this implies (theorem 3) that E_2 is independent of E_1

Cor. If E_1, E_2 are mutually independent then

$$Pr\left\{E_1 E_2\right\} = Pr\left\{E_1\right\} Pr\left\{E_2\right\}$$

In general

$$Pr\left\{\prod_{i=1}^{k} E_i\right\} = \prod_{i=1}^{k} Pr\left\{E_i\right\}$$