1 Moment Statistics

Advantages

Mathematically interrelated and related to other moments.
All have the same assumptions.
They provide the only measures of skewness and kurtosis.
They provide sufficient information to reconstruct a frequency distribution function.

Assumption Interval or ratio data, univariate system

General Moment Equation

$$\pi_k = \frac{1}{N}\sum_{i=1}^{N}\Psi_i^k$$

Ψ = any variable having its origin at some point α.
α = real or arbitrary origin (operationally defined)
π_k = the k^{th} moment about α.

Moments about the origin $\alpha = 0$, X_i = raw score

First Moment (M_1) - Mean = $\bar{X}$

$$M_1 = \frac{1}{N}\sum_{i=1}^{N}X_i = \frac{1}{N}\sum_{i=1}^{N}(freq * X)$$

Second Moment (M_2)

$$M_2 = \frac{1}{N}\sum_{i=1}^{N}X_i^2 = \frac{1}{N}\sum_{i=1}^{N}(freq * X^2)$$

Third Moment (M_3)

$$M_3 = \frac{1}{N}\sum_{i=1}^{N}X_i^3 = \frac{1}{N}\sum_{i=1}^{N}(freq * X^3)$$

Fourth Moment (M_4)

$$M_4 = \frac{1}{N}\sum_{i=1}^{N}X_i^4 = \frac{1}{N}\sum_{i=1}^{N}(freq * X^4)$$

Moments about the mean $\alpha = \bar{X}$, $(x_1 = X_i - \bar{X})$ = raw score

First Moment (m_1)

$$m_1 = \frac{1}{N}\sum_{i=1}^{N} x_i = 0$$

Second Moment (m_2) - Variance

$$m_2 = \frac{1}{N}\sum_{i=1}^{N} x_i^2 = M_2 - M_1^2$$

Standard Deviation (s) = $\sqrt{m_2}$.

Third Moment (m_3)

$$m_3 = \frac{1}{N}\sum_{i=1}^{N} x_i^3 = M_3 - 3M_1M_2 + 2M_1^3$$

For a symmetrical distribution, m_3 =0. This function(m_3) is related to skewness, but it is influenced by the size of the unit of measure.

Fourth Moment (m_4)

$$m_4 = \frac{1}{N}\sum_{i=1}^{N} x_i^4 = M_4 - 4M_1M_3 + 6M_1^2M_2 - 3M_1^4$$

m_4 is directly related to kurtosis, but it influenced by the size of the metric unit.

Standardized Moments $\alpha = \bar{X}$, $(z_1 = \frac{x_i}{s})$ = standardized score

First Moment (a_1)

$$a_1 = \frac{1}{N}\sum_{i=1}^{N} z_i = 0$$

Second Moment (a_2)

$$a_2 = \frac{1}{N}\sum_{i=1}^{N} z_i^2 = 1$$

Third Moment (a_3) - Skewness

$$a_3 = \frac{1}{N}\sum_{i=1}^{N} z_i^3 = \frac{M_3 - 3M_1M_2 + 2M_1^3}{(\sqrt{M_2 - M_1^2})^3} = \frac{m_3}{s^3}$$

Fourth Moment (a_4) - Kurtosis = $(a_4$ -3.0)

$$a_4 = \frac{1}{N}\sum_{i=1}^{N} z_i^4 = \frac{M_4 - 4M_1M_3 + 6M_1^2M_2 - 3M_1^4}{(M_2 - M_1^2)^2} = \frac{m_4}{Variance^2}$$

Interpretation of Moment Statistics

Mean (M_1) - 1^{st} moment about the origin - central tendency measure.
Variance (m_2) - 2^{nd} moment about the mean - dispersion measure.
Skewness (a_3) - 3^{rd} standardize moment - skewness measure.

$a_3 = 0 \rightarrow$ symmetrical

$a_3 > 0 \rightarrow$ positively skewed

$a_3 < 0 \rightarrow$ negatively skewed

for a_3 between $\pm$ 0.2, the distribution can be assumed to be normal with respect to skewness.

Kurtosis (a_4-3.0) - (4^{th} standardized moment - 3) - kurtosis measure.

$a_4 - 3 = 0 \rightarrow$ same peakedness as normal curve.

$a_4 - 3 > 0 \rightarrow$ more peakedness than normal curve.

$a_4 - 3 < 0 \rightarrow$ flattter than normal curve.

for $a_4 - 3$ between $\pm$ 0.5, the curve can be considered normal with respect to kurtosis.

Example - # of years attending University of Houston

Years (X)	*Students* (F)	$F * X$	$F * X^2$	$F * X^3$	$F * X^4$
1	5	5	5	5	5
2	4	8	16	32	64
3	3	9	27	81	243
4	7	28	112	448	1792
5	1	5	25	125	625
6	1	6	36	216	1296

N=21

$$\begin{array}{llll}
\sum X = 61 & M_1 = 2.90476 & m_1 = 0.00000 & a_1 = 0.00000 \\
\sum X^2 = 221 & M_2 = 10.52381 & m_2 = 2.08617 & a_2 = 1.00000 \\
\sum X^3 = 907 & M_3 = 43.19048 & m_3 = 0.50167 & a_3 = 0.16649 \\
\sum X^4 = 4025 & M_4 = 191.66667 & m_4 = 9.02988 & a_4 = 2.07483
\end{array}$$

Mean = M_1 = 2.90476

Variance = m_2 = 2.08617

Standard Deviation = $\sqrt{m_2}$ = 1.4443572

Skewness = a_3 = 0.16649 $\rightarrow$ assumed normal

Kurtosis = $(a_4 - 3)$ = -0.92517 $\rightarrow$ flatter than a normal distribution curve.

2 Single Sample Tests

2.1 Single sample test of the mean ($\hat{\mu}$) where σ is known

Make Assumption

Level of measurement - interval

Model - random sampling, population nornally distributed, $\sigma =$ (some known value)

HO: $\mu =$ (some number)

Obtain a Sampling Distribution

Z distribution

$\mu =$ (some number)

$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}}$

Choose a Significance Level and Critical Region

$\alpha = 0.05$, 0.01 or 0.001

One or two tail test

H1: $\mu \neq$ (some number)

Compute a Test Statistics

$$Z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}}$$

Make a decision

Formula for computing the confidence interval

$$C((\bar{X} - Z_{\frac{\alpha}{2}}\sigma_{\bar{X}}) \leq \mu \leq (\bar{X} + Z_{\frac{\alpha}{2}}\sigma_{\bar{X}})) = 1 - \alpha$$