



GUJARAT TECHNOLOGICAL UNIVERSITY

STATISTICAL FORMULA

for

- MBA
- MBA (Part Time)
- MBA (Integrated)
- MBA (International Business)



1. Descriptive Statistics:

Population mean (ungrouped)	$\mu = \frac{\sum x_i}{N}$
Sample mean (ungrouped)	$\bar{x} = \frac{\sum x_i}{n}$
Grouped Mean	$\mu_{\text{grouped}} = \frac{\sum f_i M_i}{N}$
Grouped Median	$\text{Median} = L + \frac{\frac{N}{2} - cf_p}{f_{\text{med}}}(W)$
Grouped Quartile	$Q_i = L + \frac{\frac{iN}{4} - cf_p}{f_{Q_i}}(W)$
Grouped Mode	$\text{Mode} = L + \frac{f_{\text{mode}} - f_1}{2f_{\text{mode}} - f_1 - f_2}(W)$
Range	Range = Highest Value – Lowest Value
Interquartile Range	$IQR = Q_3 - Q_1$
Mean Absolute Deviation	$MAD = \frac{\sum x_i - \mu }{N}$
Population standard deviation (ungrouped)	$\sigma = \sqrt{\sigma^2} \qquad \sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}}$ $\sigma = \sqrt{\frac{\sum x_i^2 - \frac{(\sum x_i)^2}{N}}{N}} \qquad \sigma = \sqrt{\frac{\sum x_i^2 - N\mu^2}{N}}$
Population variance (ungrouped)	$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N}$ $\sigma^2 = \frac{\sum x_i^2 - \frac{(\sum x_i)^2}{N}}{N}$



	$\sigma^2 = \frac{\sum x_i^2 - N\mu^2}{N}$
Population variance (grouped)	$\sigma^2 = \frac{\sum f_i (M_i - \mu)^2}{N} = \frac{\sum f_i M_i^2 - \frac{(\sum f_i M_i)^2}{N}}{N}$
Population standard deviation (grouped)	$\sigma = \sqrt{\frac{\sum f_i (M_i - \mu)^2}{N}} = \sqrt{\frac{\sum f_i M_i^2 - \frac{(\sum f_i M_i)^2}{N}}{N}}$
Sample variance	$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n-1} = \frac{\sum x_i^2 - n(\bar{x})^2}{n-1}$
Sample Standard deviation	$s = \sqrt{s^2} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n-1}}$ $= \sqrt{\frac{\sum x_i^2 - n(\bar{x})^2}{n-1}}$
Chebyshev's theorem	$1 - \frac{1}{k^2}$
z score	$z = \frac{x_i - \mu}{\frac{\sigma}{\sqrt{n}}}$
Coefficient of variation	$CV = \frac{\sigma}{\mu} \cdot 100$
Sample variance (grouped)	$s^2 = \frac{\sum (M_i - \bar{x})^2}{n-1} = \frac{\sum f_i M_i^2 - \frac{(\sum f_i M_i)^2}{N}}{n-1}$
Sample standard deviation (grouped)	$s = \sqrt{\frac{\sum (M_i - \bar{x})^2}{n-1}} = \sqrt{\frac{\sum f_i M_i^2 - \frac{(\sum f_i M_i)^2}{N}}{n-1}}$
Pearson coefficient of skewness	$S_k = \frac{3(\mu - M_i)}{\sigma}$

2. Probability:

Counting rule	Mn
sampling with replacement	N^n
sampling without replacement	${}^N C_n$



Combination formula	${}^N C_n = \binom{N}{n} = \frac{N!}{n! (N - n)!}$
General law of addition	$P(X \cup Y) = P(X) + P(Y) - P(X \cap Y)$
Special law of addition	$P(X \cup Y) = P(X) + P(Y)$
General law of multiplication	$P(X \cap Y) = P(X) \times P(Y X) = P(Y) \times P(X Y)$
Special law of multiplication	$P(X \cap Y) = P(X) \times P(Y)$
Law of conditional probability	$P(X Y) = \frac{P(X \cap Y)}{P(Y)} = \frac{P(X) \times P(Y X)}{P(Y)}$
Bayes Rule	$P(X_i Y) = \frac{P(X_i) \times P(Y X_i)}{P(X_1) \times P(Y X_1) + P(X_2) \times P(Y X_2) + \dots + P(X_n) \times P(Y X_n)}$

3. Probability Distribution:

Mean (expected) value of a discrete distribution	$\mu = E(x) = \sum [x \times P(x)]$
Variance of a discrete distribution	$\sigma^2 = \sum [(x - \mu)^2 \times P(x)]$
Standard deviation of a discrete distribution	$\sigma = \sqrt{\sum [(x - \mu)^2 \times P(x)]}$
Poisson formula	$P(x) = \frac{\lambda^x e^{-\lambda}}{x!}$
Binomial formula	$\binom{n}{x} \times p^x \times q^{n-x} = \frac{n!}{x! (n-x)!} \times p^x \times q^{n-x}$
Mean of binomial distribution	$\mu = n \times p$
Standard deviation of a binomial distribution	$\sigma = \sqrt{n \times p \times q}$
Mean and standard deviation of a uniform distribution	$\mu = \frac{a + b}{2} \quad \sigma = \frac{b - a}{\sqrt{12}}$
Probability distribution for uniform distribution	$P(x) = \frac{ x_1 - x_2 }{ b - a }$
z formula	$z = \frac{x - \mu}{\sigma}$

4. Parametric tests:



z test for a single mean	$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$
z test for a population proportion	$z = \frac{\hat{p} - p}{\sqrt{\frac{p \times q}{n}}}$
Formula to test hypothesis about μ with a finite population	$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}}$
Formula for testing hypothesis about a population variance	$\chi^2 = \frac{(n-1)s^2}{\sigma^2}, df = n-1$
t test for a single mean	$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}, df = n-1$
z test for the difference in two independent sample means	$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$
t test for two independent sample means, and population variances unknown but assumed to be equal (assume also that the two populations are normally distributed)	
$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2(n_1-1) + s_2^2(n_2-1)}{n_1 + n_2 - 2}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, df = n_1 + n_2 - 2$	
t test for difference in two related samples (the differences are normally distributed in the population)	
$t = \frac{\bar{d} - D}{\frac{s_d}{\sqrt{n}}}; df = n-1, \quad \bar{d} = \frac{\sum d}{n}, s_d = \sqrt{\frac{\sum(d - \bar{d})^2}{n-1}} = \sqrt{\frac{\sum d^2 - \frac{(\sum d)^2}{n}}{n-1}}$	
z formula for testing the difference in population proportions	
$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{(\bar{p} \times \bar{q}) \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \text{ where } \bar{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}, \bar{q} = 1 - \bar{p}$	
F test for two population variances (assume the two populations are normally distributed)	
$F = \frac{s_1^2}{s_2^2} \quad df_{\text{numerator}} = v_1 = n_1 - 1, df_{\text{denominator}} = v_2 = n_2 - 1$	
Formula for determining the critical value for the lower-tail F	
$F_{1-\alpha, v_2, v_1} = \frac{1}{F_{\alpha, v_1, v_2}}$	



Pearson's Product moment correlation coefficient

$$r = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sqrt{\sum(x - \bar{x})^2 \sum(y - \bar{y})^2}} = \frac{\sum xy - \frac{(\sum x \sum y)}{n}}{\sqrt{\left[\sum x^2 - \frac{(\sum x)^2}{n}\right] \left[\sum y^2 - \frac{(\sum y)^2}{n}\right]}}$$

Equation of the simple regression line

$$\hat{y} = \beta_0 + \beta_1 x$$

Sum of Squares

$$SS_{xx} = \sum x^2 - \frac{(\sum x)^2}{n}, SS_{yy} = \sum y^2 - \frac{(\sum y)^2}{n}, SS_{xy} = \sum xy - \frac{\sum x \sum y}{n}$$

Slope of the regression line

$$b_1 = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sum(x - \bar{x})^2} = \frac{\sum xy - n\bar{x}\bar{y}}{\sum x^2 - n\bar{x}^2} = \frac{\sum xy - \frac{(\sum x)(\sum y)}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}}$$

y-intercept of the regression line

$$b_0 = \bar{y} - b_1 \bar{x} = \frac{\sum y}{n} - b_1 \frac{(\sum x)}{n}$$

Sum of squares of error

$$SSE = \sum(y - \hat{y})^2 = \sum y^2 - b_0 \sum y - b_1 \sum xy$$

Standard error of the estimate

$$s_e = \sqrt{\frac{SSE}{n-2}}$$

Coefficient of determination

$$r^2 = 1 - \frac{SSE}{SS_{yy}} = 1 - \frac{SSE}{\sum y^2 - \frac{(\sum y)^2}{n}}$$

Computation formula for r^2

$$r^2 = \frac{b_1^2 SS_{xx}}{SS_{yy}}$$

One Way ANOVA

$$SSC = \sum_{j=1}^c n_j (\bar{x}_j - \bar{x})^2, SSE = \sum_{i=1}^{n_j} \sum_{j=1}^c (x_{ij} - \bar{x}_j)^2, SST = \sum_{i=1}^{n_j} \sum_{j=1}^c (x_{ij} - \bar{x})^2$$

$$df_c = C - 1, df_E = N - C, df_T = N - 1$$

$$MSC = \frac{SSC}{df_c}, MSE = \frac{SSE}{df_E}, F = \frac{MSC}{MSE}$$

5. Non-parametric tests:



Chi-square goodness of fit	$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}, \text{ df} = k - 1 - c$
Chi-square test of independence	$\chi^2 = \sum \sum \frac{(f_o - f_e)^2}{f_e}, \text{ df} = (r - 1)(c - 1)$
Spearman's Rank correlation	$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$
Mann Whitney U test Small sample –	$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - W_1$ $U_2 = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - W_2$ $U' = n_1 \times n_2 - U$
Mann Whitney U test Large sample –	$\mu_U = \frac{n_1 \times n_2}{2}$ $\sigma_U = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$ $z = \frac{U - \mu_U}{\sigma_U}$
Wilcoxon matched-pair signed rank test	$\mu_T = \frac{(n)(n + 1)}{4}$ $\sigma_T = \sqrt{\frac{(n)(n + 1)(2n + 1)}{24}}$ $z = \frac{T - \mu_T}{\sigma_T}$
Kruskal-Wallis test	$K = \frac{12}{n(n + 1)} \left(\sum_{j=1}^c \frac{T_j^2}{n_j} \right) - 3(n + 1)$
Friedman test	$\chi_r^2 = \frac{12}{bc(c + 1)} \sum_{j=1}^c R_j^2 - 3b(c + 1)$