Key Formulas

Relative Frequency

$$\frac{Frequency\ of\ the\ class}{n} \quad (2.1)$$

Approximate class Width

$$\frac{Largest\ data\ value - Smallest\ data\ value}{Number\ of\ classes} \quad (2.2)$$

Key Formulas

Sample Mean

$$\bar{x} = \frac{\sum x_i}{n} \quad (3.1)$$

Population Mean

$$\mu = \frac{\sum x_i}{N} \quad (3.2)$$

Weighted Mean

$$\bar{x} = \frac{\sum w_i x_i}{\sum w_i} \quad (3.3)$$

Geometry Mean

$$\bar{x}_g = \sqrt[n]{(x_1)(x_2)...(x_n)} = [(x_1)(x_2)...(x_n)]^{1/n} \quad (3.4)$$

Location of the pth Percentile

$$L_p = \frac{p}{100}(n+1) \quad (3.5)$$

Interquartile Range

$$IQR = Q_3 - Q_1 \quad (3.6)$$

Population Variance

$$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N} \quad (3.7)$$

Sample Variance

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} \quad (3.8)$$

Standard Deviation

Sample standard deviation = $s = \sqrt{s^2}$ (3.9)

Population standard deviation = $\sigma = \sqrt{\sigma^2}$ (3.10)

Coefficient of Variation

$$\left(\frac{\text{Standard deviation}}{\text{Mean}} \times 100\right)\% \quad (3.11)$$

$z-$ *Score*

$$z_i = \frac{x_i - \bar{x}}{s} \qquad (3.12)$$

Sample Covariance

$$s_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1} \qquad (3.13)$$

Population Covariance

$$\sigma_{xy} = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{N} \qquad (3.14)$$

Person Product Moment Correlation Coefficient: Sample Data

$$r_{xy} = \frac{S_{xy}}{S_x S_y} \qquad (3.15)$$

Person Product Moment Correlation Coefficient: Population Data

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \qquad (3.16)$$

Key Formulas

Counting Rule for Combinations

$$C_n^N = \binom{N}{n} = \frac{N!}{n!(N-n)!} \qquad (4.1)$$

Counting Rule for Permutations

$$P_n^N = n!\binom{N}{n} = \frac{N!}{(N-n)!} \qquad (4.2)$$

Counting Probability Using the Complement

$$P(A) = 1 - P(A^c) \qquad (4.5)$$

Addition Law

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \qquad (4.6)$$

Conditional probability

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \qquad (4.7)$$

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} \qquad (4.8)$$

Multiplication Law

$$P(A \cap B) = P(B)P(A \mid B) \qquad (4.11)$$

$$P(A \cap B) = P(A)P(B \mid A) \qquad (4.12)$$

Multiplication Law for Independent Events

$$P(A \cap B) = P(A)P(B) \qquad (4.13)$$

Bayes Theorem

$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2) + ... + P(A_n)P(B \mid A_n)} \qquad (4.19)$$

Key Formulas

Discrete Uniform Probability Function

$$f(x) = 1/n \qquad (5.3)$$

Expected Value of a Discrete Random Variable

$$E(x) = \mu = \sum xf(x) \qquad (5.4)$$

Variance of a Discrete Random Variable

$$Var(x) = \sigma^2 = \sum (x - \mu)^2 f(x) \qquad (5.5)$$

Covariance of Random Variables x and y

$$\sigma_{xy} = [Var(x+y) - Var(x) - Var(y)]/2 \qquad (5.6)$$

Correlation between Random Variables x and y

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \qquad (5.7)$$

Expected Value of a Linear Combination of Variables x and y

$$E(ax + by) = aE(x) + bE(y) \qquad (5.8)$$

Variance of a Linear Combination of Two Random Variables

$$Var(ax + by) = a^2 Var(x) + b^2 Var(y) + 2ab\sigma_{xy} \qquad (5.9)$$

where σ_{xy} is the covariance of x and y

Number of Experimental Outcomes Providing Exactly Successes in Trials

$$\binom{n}{x} = \frac{n!}{x!(n-x)!} \qquad (5.10)$$

Binomial probability Function

$$f(x) = \binom{n}{x} p^x (1-p)^{(n-x)} \qquad (5.12)$$

Expected Value for the Binomial Distribution

$$E(x) = \mu = np \qquad (5.13)$$

Variance for the Binomial Distribution

$$Var(x) = \sigma^2 = np(1-p) \qquad (5.14)$$

Poisson probability Function

$$f(x) = \frac{\mu^x e^{-\mu}}{x!} \qquad (5.15)$$

Hypergeometric Probability Function

$$f(x)=\frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} \qquad (5.16)$$

Expected Value for the Hypergeometric Probability Function

$$E(x)=\mu=n\left(\frac{r}{N}\right) \qquad (5.17)$$

Variance for the Hypergeometric Distribution

$$Var(x)=\sigma^2=n\left(\frac{r}{N}\right)\left(1-\frac{r}{N}\right)\left(\frac{N-n}{N-1}\right) \qquad (5.18)$$

Key Formulas

Uniform probability Density Function

$$f(x)=\begin{cases}\dfrac{1}{b-a} & for a \le x \le b\\ 0 & \text{elsewhere}\end{cases} \qquad (6.1)$$

Normal probability Density Function

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad (6.2)$$

Converting to the standard Normal Random variable

$$z=\frac{x-\mu}{\sigma} \qquad (6.3)$$

Exponential Probability Density Function

$$f(x)=\frac{1}{\mu}e^{-\frac{x}{\mu}} \quad \text{for} \quad x \ge 0 \qquad (6.4)$$

Exponential Distribution: Cumulative Probabilities

$$P(x \le x_0)=1-e^{-\frac{x_0}{\mu}} \qquad (6.5)$$

Key Formulas

Expected Value of $\bar{x}$

$$E(\bar{x}) = \mu \qquad (7.1)$$

Standard Deviation of $\bar{x}$ (Standard Error)

Finite Population

$$\sigma_{\bar{x}} = \sqrt{\frac{N-n}{N-1}}\left(\frac{\sigma}{\sqrt{n}}\right) \qquad (7.2)$$

Infinite Population

$$\sigma_{\bar{x}} = \left(\frac{\sigma}{\sqrt{n}}\right)$$

Expected Value of $\bar{p}$

$$E(\bar{p}) = p \qquad (7.4)$$

Standard Deviation of $\bar{p}$ (Standard Error)

Finite Population

$$\sigma_{\bar{p}} = \sqrt{\frac{N-n}{N-1}}\sqrt{\frac{p(1-p)}{n}} \qquad (7.5)$$

Infinite Population

$$\sigma_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}}$$

Key Formulas

Interval Estimate of a Population Mean: σ Known

$$\bar{x} \pm z_{a/2}\frac{\sigma}{\sqrt{n}} \qquad (8.1)$$

Interval Estimate of a Population Mean: σ UnKnown

$$\bar{x} \pm t_{a/2}\frac{S}{\sqrt{n}} \qquad (8.2)$$

Sample Size for an Interval Estimate of a Population Mean

$$n = \frac{(z_{a/2})^2\sigma^2}{E^2} \qquad (8.3)$$

Interval Estimate of a Population Proportion

$$\bar{p} \pm z_{a/2}\sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \qquad (8.6)$$

Sample size for an Interval Estimate of a Population Proportion

$$n = \frac{(z_{a/2})^2 p^*(1-p^*)}{E^2} \qquad (8.7)$$

Key Formulas

Test Statistic for Hypothesis Tests About a Population Mean: σ Known

$$z = \frac{\overline{x} - \mu_0}{\sigma / \sqrt{n}} \qquad (9.1)$$

Test Statistic for Hypothesis Tests About a Population Mean: σ unKnown

$$t = \frac{\overline{x} - \mu_0}{s / \sqrt{n}} \qquad (9.2)$$

Test Statistic for Hypothesis Tests About a Population Proportion

$$z = \frac{\overline{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \qquad (9.4)$$

Key Formulas

Point Estimator of the Difference Between Two Population Means

$$\overline{x_1} - \overline{x_2} \qquad (10.1)$$

Standard Error of $\overline{x_1} - \overline{x_2}$

$$\sigma_{\overline{x_1} - \overline{x_2}} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \qquad (10.2)$$

Interval Estimate of the Difference Between Two Population Means: σ_1 and σ_2 Known

$$\overline{x_1} - \overline{x_2} \pm z_{a/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \qquad (10.4)$$

Test Statistic for Hypothesis Tests About $\mu_1 - \mu_2$: σ_1 and σ_2 Known

$$z = \frac{(\overline{x_1} - \overline{x_2}) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \qquad (10.5)$$

Interval Estimate of the Difference Between Two Population Means: σ_1 and σ_2 UnKnown

$$\overline{x}_1 - \overline{x}_2 \pm t_{a/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} \qquad (10.6)$$

Degree of Freedom: t Distribution with two Independent Random Samples

$$df = \frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1-1}\left(\frac{s_1^2}{n_1}\right)^2+\frac{1}{n_2-1}\left(\frac{s_2^2}{n_2}\right)^2} \qquad (10.7)$$

Test statistic for Hypothesis Tests About $\mu_1 - \mu_2$: σ_1 and σ_2 UnKnown

$$t = \frac{(\overline{x}_1 - \overline{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} \qquad (10.8)$$

Test Statistic for Hypothesis Tests Involving Matched Samples

$$t = \frac{\overline{d} - \mu_d}{S_d / \sqrt{n}} \qquad (10.9)$$

Point Estimator of the Difference Between Two Population Proportions

$$\overline{p}_1 - \overline{p}_2 \qquad (10.10)$$

Standard Error of $\overline{p}_1 - \overline{p}_2$

$$\sigma_{\overline{p}_1 - \overline{p}_2} = \sqrt{\frac{p_1(1-p_1)}{n_1}+\frac{p_2(1-p_2)}{n_2}} \qquad (10.11)$$

Interval Estimate of the Difference Between Two Population Means: σ_1 and σ_2 Known

$$\overline{p}_1 - \overline{p}_2 \pm z_{a/2}\sqrt{\frac{\overline{p}_1(1-\overline{p}_1)}{n_1}+\frac{\overline{p}_2(1-\overline{p}_2)}{n_2}} \qquad (10.13)$$

Standard Error of $\overline{p}_1 - \overline{p}_2$ When $p_1 = p_2 = p$

$$\sigma_{\overline{p}_1 - \overline{p}_2} = \sqrt{p(1-p)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} \qquad (10.14)$$

Pooled Estimator of p when $p_1 = p_2 = p$

$$\overline{p} = \frac{n_1\overline{p}_1 + n_2\overline{p}_2}{n_1+n_2} \qquad (10.15)$$

Test Statistic for Hypothesis Tests About $p_1 - p_2$

$$z = \frac{\left(\overline{p}_1 - \overline{p}_2\right)}{\sqrt{\overline{p}(1-\overline{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \qquad (10.16)$$

Key Formulas

Interval Estimate of a Population Variance

$$\frac{(n-1)S^2}{\chi^2_{\alpha/2}} \le \sigma^2 \le \frac{(n-1)S^2}{\chi^2_{(1-\alpha/2)}} \qquad (11.7)$$

Test Statistic for Hypothesis Tests About a Population Variance

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} \qquad (11.8)$$

Test Statistic for Hypothesis Tests About Population Variances with $\sigma_1^2 = \sigma_2^2$

$$F = \frac{s_1^2}{S_2^2} \qquad (11.10)$$

Key Formulas

Test Statistic for the Goodness of Fit Test

$$\chi^2 = \sum_{i=1}^{k} \frac{(f_i - e_i)^2}{e_i} \qquad (12.1)$$

Expected Frequencies: Test of Independence

$$e_{ij} = \frac{(\text{Row } i \text{ Total})(\text{Column } j \text{ Total})}{\text{Sample Size}} \qquad (12.2)$$

Chi-Square Test Statistic for Test of Independence and Test for Equality of three or More Population Proportions

$$\chi^2 = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} \qquad (12.3 \text{ and } 12.5)$$

Expected Frequencies Test for Equality of Three or More Population Proportions

$$e_{ij} = \frac{(\text{Row } i \text{ Total})(\text{Column } j \text{ Total})}{\text{Sum of Sample Sizes}} \qquad (12.4)$$

Critical Values for the Marascuilo Pairwise Comparison Procedure

$$CV_{ij} = \sqrt{\chi^2_\alpha} \sqrt{\frac{\overline{p_i}(1-\overline{p_i})}{n_i} + \frac{\overline{p_j}(1-\overline{p_j})}{n_j}} \qquad (12.6)$$

Key Formulas

Completely Randomized Design

Sample Mean for Treatment j

$$\bar{x}_j = \frac{\sum_{i=1}^{n_j} x_{ij}}{n_j} \qquad (13.1)$$

Sample Variance for Treatment j

$$s_j^2 = \frac{\sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2}{n_j - 1} \qquad (13.2)$$

Overall Sample Mean

$$\bar{\bar{x}} = \frac{\sum_{j=1}^{k} \sum_{i=1}^{n_j} x_{ij}}{n_T} \qquad (13.3)$$

$$n_T = n_1 + n_2 + \ldots + n_k \qquad (13.4)$$

Mean Square Due to Treatments

$$\text{MSTR} = \frac{\text{SSTR}}{k-1} \qquad (13.7)$$

Sum of Square Due to Treatments

$$\text{SSTR} = \sum_{j=1}^{k} n_j (\bar{x}_j - \bar{\bar{x}})^2 \qquad (13.8)$$

Mean Square Due to Error

$$\text{MSE} = \frac{\text{SSE}}{n_T - k} \qquad (13.10)$$

Sum of Square Due to Error

$$\text{SSE} = \sum_{j=1}^{k} (n_j - 1) s_j^2 \qquad (13.11)$$

Test Statistic for the Equality of k Population Means

$$F = \frac{\text{MSTR}}{\text{MSE}} \qquad (13.12)$$

Total Sum of Squares

$$\text{SST} = \sum_{j=1}^{k} \sum_{i=1}^{n_j} (x_{ij} - \bar{\bar{x}})^2 \qquad (13.13)$$

Partitioning of Sum of Squares

$$\text{SST} = \text{SSTR} + \text{SSE} \qquad (13.14)$$

Multiple Comparison Produres

Test Statistic for Fishes's LSD Procedure

$$t = \frac{\bar{x}_i - \bar{x}_j}{\sqrt{MSE\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}} \qquad (13.16)$$

Fisher's LSD

$$LSD = t_{\alpha/2}\sqrt{MSE\left(\frac{1}{n_i} + \frac{1}{n_j}\right)} \qquad (13.17)$$

Randomized Block Design

Total Sum of Squares

$$SST = \sum_{i=1}^{b}\sum_{j=1}^{k}(x_{ij} - \bar{\bar{x}})^2 \qquad (13.22)$$

Sum of Squares Due to Treatments

$$SSTR = b\sum_{j=1}^{k}(x_{\bullet j} - \bar{\bar{x}})^2 \qquad (13.23)$$

Sum of Squares Due to Blocks

$$SSBL = k\sum_{i=1}^{b}(x_{i\bullet} - \bar{\bar{x}})^2 \qquad (13.24)$$

Sum of Squares Due to Error

SSE= SST - SSTR - SSBL (13.25)

Factorial Experiment

Total Sum of Squares

$$SST = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}(x_{ijk} - \bar{\bar{x}})^2 \qquad (13.27)$$

Sum of Squares for Factor A

$$SSA = br\sum_{i=1}^{a}(x_{i\bullet} - \bar{\bar{x}})^2 \qquad (13.28)$$

Sum of Squares for Factor B

$$SSA = ar\sum_{j=1}^{b}(x_{\bullet j} - \bar{\bar{x}})^2 \qquad (13.29)$$

Sum of Squares for Interaction

$$SSAB = r\sum_{i=1}^{a}\sum_{j=1}^{b}(\overline{x_{ij}} - \overline{x_{i\bullet}} - \overline{x_{\bullet j}} + \bar{\bar{x}})^2 \qquad (13.30)$$

Sum of Squares for Error

SSE=SST - SSA - SSB - SSAB (13.31)

Key Formulas

Simple Linear Regression Model

$$y = \beta_0 + \beta_1 x + \varepsilon \qquad (14.1)$$

Simple Linear Regression Equation

$$E(y) = \beta_0 + \beta_1 x \qquad (14.2)$$

Estimated Simple Linear Regression Equation

$$\hat{y} = b_0 + b_1 x \qquad (14.3)$$

Least Squares Criterion

$$\min \sum (y_i - \hat{y}_i)^2 \qquad (14.5)$$

Slope and y-Intercept for the Estimated Regression Equation

$$b_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \qquad (14.6)$$

$$b_0 = \bar{y} - b_1 \bar{x} \qquad (14.7)$$

Sum of Squared Due to Error

$$\text{SSE} = \sum (y_i - \hat{y}_i)^2 \qquad (14.8)$$

Total Sum of Squares

$$\text{SST} = \sum (y_i - \bar{y})^2 \qquad (14.9)$$

Sum of Squares due to Regression

$$\text{SSR} = \sum (\hat{y}_i - \bar{y})^2 \qquad (14.10)$$

Relationship Among SST, SSR, and SSE

$$\text{SST} = \text{SSR} + \text{SSE} \qquad (14.11)$$

Coefficient of Determination

$$r^2 = \frac{SSR}{SST} \qquad (14.12)$$

Sample Correlation Coefficient

$$r_{xy} = (\text{sign of } b_1)\sqrt{\text{Coefficient of determination}}$$

$$r_{xy} = (\text{sign of } b_1)\sqrt{r^2} \qquad (14.13)$$

Mean Square Error (Estimate of σ^2)

$$s^2 = MSE = \frac{SSE}{n-2} \qquad (14.15)$$

Standard Error of the Estimate

$$s = \sqrt{MSE} = \sqrt{\frac{SSE}{n-2}} \qquad (14.16)$$

Standard Deviation of b_1

$$\sigma_{b_1} = \frac{\sigma}{\sqrt{\sum(x_i - \bar{x})^2}} \qquad (14.17)$$

Estimated Standard Deviation of b_1

$$s_{b_1} = \frac{s}{\sqrt{\sum(x_i - \bar{x})^2}} \qquad (14.18)$$

t Test Statistic

$$r = \frac{b_1}{s_{b_1}} \qquad (14.19)$$

Mean Square Regression

$$\text{MSR} = \frac{\text{SSR}}{\text{Number of independent variables}} \qquad (14.20)$$

F Test Statistic

$$F = \frac{MSR}{MSE} \qquad (14.21)$$

Estimated Standard Deviation of $\hat{y}^*$

$$s_{\hat{y}^*} = s = \sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{\sum(x_i - \bar{x})^2}} \qquad (14.23)$$

Confidence Interval for $E(y^*)$

$$\hat{y}^* \pm t_{\alpha/2} s_{\hat{y}_i^*} \qquad (14.24)$$

Estimated Standard Deviation of an Individual Value

$$s_{pred} = s\sqrt{1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{\sum(x_i - \bar{x})^2}} \qquad (14.26)$$

Prediction Interval for y^*

$$\hat{y}^* \pm t_{\alpha/2} s_{pred} \qquad (14.27)$$

Residual for Observation i

$$y_i - \hat{y}_i \qquad (14.28)$$

Standard Deviation of the ith Residual

$$s_{y_i - \hat{y}_i} = s\sqrt{1 - h_i} \qquad (14.30)$$

Standardized Residual for Observation i

$$\frac{y_i - \hat{y}_i}{s_{y_i - \hat{y}_i}} \qquad (14.32)$$

Leverage of observation i

$$h_i = \frac{1}{n} + \frac{(x_i - \bar{x})^2}{\sum(x_i - \bar{x})^2} \qquad (14.33)$$

Key Formulas

Multiple Regression Model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p + \varepsilon \qquad (15.1)$$

Multiple Regression Equation

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p \qquad (15.2)$$

Estimated Multiple Regression Equation

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + ... + b_p x_p \qquad (15.3)$$

Least Squares Criterion

$$\min \sum(y_i - \hat{y}_i)^2 \qquad (15.4)$$

Relationship Among SST, SSR, and SSE

SST = SSR + SSE (15.7)

Multiple Coefficient of Determination

$$R^2 = \frac{SSR}{SST} \qquad (15.8)$$

Adjusted Multiple Coefficient of Determination

$$R_a^2 = 1 - (1 - R^2)\frac{n-1}{n-p-1} \qquad (15.9)$$

Mean Square Due to Regression

$$\text{MSR} = \frac{\text{SSR}}{p} \qquad (15.12)$$

Mean Square Due to Error

$$\text{MSE} = \frac{\text{SSE}}{n-p-1} \qquad (15.13)$$

F Test Statistic

$$F = \frac{\text{MSR}}{\text{MSE}} \qquad (15.14)$$

t Test Statistic

$$t = \frac{b_i}{s_{b_i}} \qquad (15.15)$$