Formula Card for Johnson & Kuby, ELEMENTARY STATISTICS, Eleventh Edition

Sample mean:

$\bar{x} = \frac{\sum x}{n}$ **(2.1)**

Depth of sample median:

$d(\tilde{x}) = (n + 1)/2$ **(2.2)**

Range: $H - L$ **(2.4)**

Sample variance:

$s^2 = \frac{\sum(x - \bar{x})^2}{n - 1}$ **(2.5)**

or

$s^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n - 1}$ **(2.9)**

Sample standard deviation:

$s = \sqrt{s^2}$ **(2.6)**

Chebyshev's theorem: at least $1 - (1/k^2)$ **(p. 99)**

Sum of squares of x:

$SS(x) = \sum x^2 - ((\sum x)^2/n)$ **(2.8)**

Sum of squares of y:

$SS(y) = \sum y^2 - ((\sum y)^2/n)$ **(3.3)**

Sum of squares of xy:

$SS(xy) = \sum xy - ((\sum x \cdot \sum y)/n)$ **(3.4)**

Pearson's correlation coefficient:

$r = SS(xy)/\sqrt{SS(x) \cdot SS(y)}$ **(3.2)**

Equation for line of best fit: $\hat{y} = b_o + b_1 x$ **(p. 146)**

Slope for line of best fit: $b_1 = SS(xy)/SS(x)$ **(3.6)**

y-intercept for line of best fit:

$b_0 = [\sum y - (b_1 \cdot \sum x)]/n$ **(3.7)**

Empirical (observed) probability:

$P'(A) = n(A)/n$ **(4.1)**

Theoretical probability for equally likely sample space:

$P(A) = n(A)/n(S)$ **(4.2)**

Complement rule:

$P(\text{not } A) = P(\overline{A}) = 1 - P(A)$ **(4.3)**

General addition rule:

$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$ **(4.4)**

General multiplication rule:

$P(A \text{ and } B) = P(A) \cdot P(B \mid A)$ **(4.5)**

Special addition rule for mutually exclusive events:

$P(A \text{ or } B \text{ or} \ldots \text{or } E) = P(A) + P(B) + \cdots + P(E)$ **(4.6)**

Special multiplication rule for independent events:

$P(A \text{ and } B \text{ and} \ldots \text{and } E) = P(A) \cdot P(B) \cdot \cdots \cdot P(E)$ **(4.7)**

Mean of discrete random variable:

$\mu = \sum[xP(x)]$ **(5.1)**

Variance of discrete random variable:

$\sigma^2 = \sum[x^2 P(x)] - \{\sum[xP(x)]\}^2$ **(5.3a)**

Standard deviation of discrete random variable:

$\sigma = \sqrt{\sigma^2}$ **(5.4)**

Factorial: $n! = (n)(n-1)(n-2) \cdot \cdots \cdot 2 \cdot 1$ **(p. 248)**

Binomial coefficient:

$\binom{n}{x} = \frac{n!}{x!(n-x)!}$ **(5.6)**

Binomial probability function:

$P(x) = \binom{n}{x} \cdot p^x \cdot q^{n-x}, x = 0, 1, 2, \ldots, n$ **(5.5)**

Mean of binomial random variable: $\mu = np$ **(5.7)**

Standard deviation, binomial random variable:

$\sigma = \sqrt{npq}$ **(5.8)**

Standard score: $z = (x - \mu)/\sigma$ **(6.3)**

Standard score for $\bar{x}$: $z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$ **(7.2)**

Confidence interval for mean, μ (σ known):

$\bar{x} \pm z(\alpha/2) \cdot (\sigma/\sqrt{n})$ **(8.1)**

Sample size for $1 - \alpha$ confidence estimate for μ:

$n = [z(\alpha/2) \cdot \sigma/E]^2$ **(8.3)**

Calculated test statistic for H_o: $\mu = \mu_0$ (σ known):

$z\bigstar = (\bar{x} - \mu_0)/(\sigma/\sqrt{n})$ **(8.4)**

Confidence interval estimate for mean, μ (σ unknown):

$\bar{x} \pm t(\text{df}, \alpha/2) \cdot (s/\sqrt{n})$ with $\text{df} = n - 1$ **(9.1)**

Calculated test statistic for H_o: $\mu = \mu_0$ (σ unknown):

$t\bigstar = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$ with $\text{df} = n - 1$ **(9.2)**

Confidence interval estimate for proportion, p:

$p' \pm z(\alpha/2) \cdot \sqrt{(p'q')/n}, \quad p' = x/n$ **(9.6)**

Sample Size for 1-α Confidence Interval of p

$n = \frac{[z(\alpha/2)]^2 p^* q^*}{E^2}$ (9.8)

Calculated test statistic for H_o: $p = p_0$:

$z\bigstar = (p' - p_0)/\sqrt{(p_0 q_0/n)}, \quad p' = x/n$ **(9.9)**

Calculated test statistic for H_o: $\sigma^2 = \sigma_0^2$ or $\sigma = \sigma_0$:

$\chi^2\bigstar = (n-1)s^2/\sigma_0^2, \quad \text{df} = n - 1$ **(9.10)**

Mean difference between two dependentsamples:

Paired difference: $d = x_1 - x_2$ **(10.1)**

Confidence interval for mean difference, μ_d:

$\bar{d} \pm t(\text{df}, \alpha/2) \cdot s_d/\sqrt{n}$ with $\text{df} = n - 1$ **(10.2)**

Sample mean of paired differences:

$\bar{d} = \sum d/n$ **(10.3)**

Sample standard deviation of paired differences:

$$s_d = \sqrt{\frac{\sum^{d^2} - \frac{(\sum d)^2}{n}}{n-1}} \quad \textbf{(10.4)}$$

Calculated test statistic for H_o: $\mu_d = \mu_0$:

$t\bigstar = (\bar{d} - \mu_0)/(s_d/\sqrt{n})$, df $= n - 1$ **(10.5)**

Difference between means of two independent samples:

Degrees of freedom:

df = smaller of $(n_1 - 1)$ or $(n_2 - 1)$ **(p. 496)**

Confidence interval estimate for $\mu_1 - \mu_2$:

$(\bar{x}_1 - \bar{x}_2) \pm t(\text{df}, \alpha/2)\sqrt{(s_1^2/n_1) + (s_2^2/n_2)}$ **(10.8)**

Calculated test statistic for H_o: $\mu_1 - \mu_2 = (\mu_1 - \mu_2)_0$:

$t\bigstar = [(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)_0]/\sqrt{(s_1^2/n_1) + (s_2^2/n_2)}$ **(10.9)**

Difference between proportions of two independent samples:

Confidence interval for $p_1 - p_2$:

$$(p'_1 - p'_2) \pm z(\alpha/2) \cdot \sqrt{\frac{p'_1 q'_1}{n_1} + \frac{p'_2 q'_2}{n_2}} \quad \textbf{(10.11)}$$

Pooled observed probability:

$p'_p = (x_1 + x_2)/(n_1 + n_2)$ **(10.13)**

$q'_p = 1 - p_p{}'$ **(10.14)**

Calculated test statistic for H_o: $p_1 - p_2 = 0$:

$$z\bigstar = \frac{p'_1 - p'_2}{\sqrt{(p'_p)(q'_p)\left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right]}} \quad \textbf{(10.15)}$$

Ratio of variances between two independent samples:

Calculated test statistic for H_o: $\sigma_1^2/\sigma_2^2 = 1$:

$F\bigstar = s_1^2/s_2^2$ **(10.16)**

Calculated test statistic for enumerative data:

$\chi^2\bigstar = \Sigma[(O - E)^2/E]$ **(11.1)**

Multinomial experiment:

Degrees of freedom: df $= k - 1$ **(11.2)**

Expected frequency: $E = n \cdot p$ **(11.3)**

Test for independence or Test of homogeneity:

Degrees of freedom:

df $= (r - 1) \cdot (c - 1)$ **(11.4)**

Expected value: $E = (R \cdot C)/n$ **(11.5)**

Mathematical model:

$x_{c,k} = \mu + F_c + \epsilon_{k(c)}$ **(12.13)**

Total sum of squares:

$$\text{SS(total)} = \sum(x^2) - \frac{(\sum x)^2}{n} \quad \textbf{(12.2)}$$

Sum of squares due to factor:

$$\left[\left(\frac{C_1^2}{k_1}\right) + \left(\frac{C_2^2}{k_2}\right) + \left(\frac{C_3^2}{k_3}\right) + \cdots\right] - \left[\frac{(\sum x)^2}{\text{n}}\right] \quad \textbf{(12.3)}$$

Sum of squares due to error:

$\text{SS(error)} = \sum(x^2) - [(C_1^2/k_1) + (C_2^2/k_2) + (C_3^2/k_3) + \cdots]$ **(12.4)**

Degrees of freedom for total:

df(total) $= n - 1$ **(12.6)**

Degrees of freedom for factor:

df(factor) $= c - 1$ **(12.5)**

Degrees of freedom for error:

df(error) $= n - c$ **(12.7)**

Mean square for factor:

MS(factor) = SS(factor)/df(factor) **(12.10)**

Mean square for error:

MS(error) = SS(error)/df(error) **(12.11)**

Calculated test statistic for H_o: Mean value is same at all levels:

$F\bigstar$ = MS(factor)/MS(error) **(12.12)**

Covariance of x and y:

$\text{covar}(x, y) = \Sigma[(x - \bar{x})(y - \bar{y})]/(n - 1)$ **(13.1)**

Pearson's correlation coefficient:

$r = \text{covar}(x, y)/(s_x \cdot s_y)$ **(13.2)**

or

$r = \text{SS}(xy)/\sqrt{\text{SS}(x) \cdot \text{SS}(y)}$ **(3.2)** or **(13.3)**

Experimental error: $e = y - \hat{y}$ **(13.5)**

Estimated variance of error: $s_e^2 = \sum(y - \hat{y})^2/(n - 2)$ **(13.6)**

or

$$s_e^2 = \frac{(\sum y^2) - (b_0)(\sum y) - (b_1)(\sum xy)}{n - 2} \quad \textbf{(13.8)}$$

Standard deviation about the line of best fit:

$s_e = \sqrt{s_e^2}$ **(13.9)**

Estimate for variance of slope:

$$s_{b_1}^2 = \frac{s_e^2}{\text{SS}(x)} = \frac{s_e^2}{\sum x^2 - [(\sum x)^2/n]} \quad \textbf{(13.12)}$$

Confidence interval for β_1:

$b_1 \pm t(\text{df}, \alpha/2) \cdot s_{b1}$ **(13.14)**

Calculated test statistic for H_o: $\beta_1 = 0$:

$t\bigstar = (b_1 - \beta_1)/s_{b1}$ with df $= n - 2$ **(13.15)**

Confidence interval for mean value of y at x_0:

$$\hat{y} \pm t(n - 2, \alpha/2) \cdot s_e \cdot \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\text{SS}(x)}} \quad \textbf{(13.17)}$$

Prediction interval for y at x_0:

$$\hat{y} \pm t(n - 2, \alpha/2) \cdot s_e \cdot \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\text{SS}(x)}} \quad \textbf{(13.16)}$$

Mann–Whitney U test:

$U_a = n_a \cdot n_b + [(n_b) \cdot (n_b + 1)/2] - R_b$ **(14.3)**

$U_b = n_a \cdot n_b + [(n_a) \cdot (n_a + 1)/2] - R_a$ **(14.4)**

Spearman's rank correlation coefficient:

$$r_s = 1 - \left[\frac{6\sum d^2}{n(n^2 - 1)}\right] \quad \textbf{(14.11)}$$