

Formula Sheet

$$s^2 = \frac{\sum_{n-1} (X - \bar{X})^2}{n-1} = \frac{\sum_{n-1} X^2 - \left[(\sum X)^2 / n \right]}{n-1}, \text{MAD} = \frac{\sum_n |X - \bar{X}|}{n}$$

$P(A|B) = P(A \text{ and } B)/P(B)$. If A is independent of B $\Leftrightarrow P(A \text{ and } B) = P(A) \cdot P(B)$.

$${}_N P_n = \frac{N!}{(N-n)!}, {}_N C_n = \binom{N}{n} = \frac{N!}{n!(N-n)!}, P\{X=x\} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}, \mu = np,$$

$$\sigma^2 = np(1-p). Z = \frac{X - \mu}{\sigma}. \mu_{\bar{X}} = \mu, \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}, \text{when sampling is without}$$

$$\text{replacement sample; } \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}, \text{when sampling is with replacement. } Z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}.$$

$$\bar{X} \pm Z_{1-(\alpha/2)} \frac{\sigma}{\sqrt{n}}; \bar{X} \pm Z_{1-(\alpha/2)} \frac{s}{\sqrt{n}}; \bar{X} \pm t_{1-(\alpha/2)} \frac{s}{\sqrt{n}}, t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}}, \text{df} = n-1. \text{Warning the}$$

formula Power = $P\left(Z > Z_{1-(\alpha/2)} - \frac{|\mu_0 - \mu_1|}{\sigma/\sqrt{n}}\right)$ can give wrong answer when $|\mu_0 - \mu_1|$ is small

relative to σ/\sqrt{n} , to compute the power of a two-sided test for μ . The power of a one-sided test for $\mu = P\left(Z > Z_{1-\alpha} - \frac{|\mu_0 - \mu_1|}{\sigma/\sqrt{n}}\right)$. The sample size required to ensure a specific level of power in

a two-sided test is $n = \left(\frac{Z_{1-(\alpha/2)} + Z_{1-\beta}}{ES} \right)^2$, where $ES = \frac{|\mu_0 - \mu_1|}{\sigma}$, this formula can give wrong

answer when $|\mu_0 - \mu_1|$ is small relative to σ/\sqrt{n} . The sample size required to ensure a specific

$$\text{level of power in a one-sided test is } n = \left(\frac{Z_{1-\alpha} + Z_{1-\beta}}{ES} \right)^2. Z = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}},$$

$$\bar{X}_1 - \bar{X}_2 \pm Z_{1-(\alpha/2)} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}. F = \frac{s_1^2}{s_2^2}, \text{critical values: lower } 1/F_{1-(\alpha/2)}(\text{df}_2, \text{df}_1), \text{upper}$$

$$F_{1-(\alpha/2)}(\text{df}_1, \text{df}_2), \text{df}_1 = n_1 - 1, \text{df}_2 = n_2 - 1. Z = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

$$\bar{X}_1 - \bar{X}_2 \pm Z_{1-(\alpha/2)} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, \quad Z = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}, \quad \bar{X}_1 - \bar{X}_2 \pm Z_{1-(\alpha/2)} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$t = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, \quad \bar{X}_1 - \bar{X}_2 \pm t_{1-(\alpha/2)} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, \quad S_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}},$$

$$df = n_1 + n_2 - 2. \quad t = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}, \quad \bar{X}_1 - \bar{X}_2 \pm t_{1-(\alpha/2)} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}},$$

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right)^2}{\frac{(s_1^2/n_1)^2}{n_1-1} + \frac{(s_2^2/n_2)^2}{n_2-1}}. \quad \bar{X}_d \pm t_{1-(\alpha/2)} \frac{s_d}{\sqrt{n}}, \quad t = \frac{\bar{X}_d - \mu_d}{\frac{s_d}{\sqrt{n}}}, \quad df = n-1.$$

$$ES = \frac{|(\mu_1 - \mu_2)_{H_1}|}{\sigma} = \frac{|\mu_1 - \mu_2|}{\sigma}. \quad \text{Warning the formula Power} = P\left(Z > Z_{1-(\alpha/2)} - \frac{|\mu_1 - \mu_2|}{\sqrt{2\sigma^2/n}}\right)$$

can give wrong answer when $|\mu_1 - \mu_2|$ is small relative to $\sqrt{2\sigma^2/n}$, to compute the power of a two-sided test for $\mu_1 - \mu_2$. The sample size required to ensure a specific level of power in a two-

sided test is $n = 2 \left(\frac{Z_{1-(\alpha/2)} + Z_{1-\beta}}{ES} \right)^2$, this formula can give wrong answer when $|\mu_1 - \mu_2|$ is

small relative to $\sqrt{2\sigma^2/n}$. The sample size required to ensure a specific level of power in a one-

$$\text{sided test is } n = 2 \left(\frac{Z_{1-\alpha} + Z_{1-\beta}}{ES} \right)^2. \quad \hat{p} \pm Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \quad Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}};$$

$$(\hat{p}_1 - \hat{p}_2) \pm Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}; \quad Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p}) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}, \quad \hat{p} = \frac{x_1 + x_2}{n_1 + n_2}.$$

$$\chi^2 = \sum \frac{(O - E)^2}{E}, \quad df = k-1, \quad \text{for independence test} \quad df = (R-1)(C-1). \quad ES = |p_1 - p_2|,$$

$$\bar{p} = \frac{p_1 + p_2}{2}, \quad \bar{q} = 1 - \bar{p}, \quad n_i = \left(\frac{\sqrt{\bar{p}\bar{q}} Z_{1-(\alpha/2)} + \sqrt{p_1 q_1 + p_2 q_2} Z_{1-\beta}}{ES} \right)^2.$$

Treatment or Comparison group	Outcome		
	1	0	
1	a	b	a+b
0	c	d	c+d
	a+c	b+d	N

$$\hat{RD} \pm Z_{1-\alpha/2} \sqrt{\left(\frac{\hat{p}_0(1-\hat{p}_0)}{n_0} \right) + \left(\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} \right)}, \quad exp\left(\ln(\hat{RR}) \pm Z_{1-\alpha/2} \sqrt{\frac{(d/c)}{n_0} + \frac{(b/a)}{n_1}} \right),$$

$$\exp\left(\ln(\hat{OR}) \pm Z_{1-\alpha/2} \sqrt{\left(\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d} \right)} \right). \quad \chi^2 = \frac{(ad-bc)^2(N)}{(a+b)(c+d)(a+c)(b+d)}.$$

$$\chi^2_{CMH} = \frac{\left(\sum \frac{(ad-bc)}{N} \right)^2}{\sum \frac{(a+b)(c+d)(a+c)(b+d)}{(N-1)N^2}}, \quad \hat{RR}_{CMH} = \frac{\sum \frac{a(c+d)}{N}}{\sum \frac{c(a+b)}{N}}.$$

$$n_i = \left(\frac{\sqrt{2pq}Z_{1-(\alpha/2)} + \sqrt{p_1q_1 + p_0q_0}Z_{1-\beta}}{\hat{RD}} \right)^2.$$

$$s_w^2 = \sum_{j=1}^k \frac{s_j^2}{k} = \frac{s_1^2 + s_2^2 + \dots + s_k^2}{k}, \quad s_{\bar{X}_j}^2 = \sum_{j=1}^k \frac{(\bar{X}_j - \bar{X})^2}{k-1}, \quad s_b^2 = ns_{\bar{X}_j}^2, \quad F = \frac{s_b^2}{s_w^2}.$$

Source of Variation	Sums of Squares (SS)	Degrees of Freedom (df)	Mean Squares (MS)	F
Between	$SS_b = \sum n_j (\bar{X}_{..j} - \bar{X}_{..})^2$	k-1	$s_b^2 = MS_b = \frac{SS_b}{k-1}$	$F = \frac{MS_b}{MS_w}$
Within	$SS_w = \sum \sum (X_{ij} - \bar{X}_{..j})^2$	N-k	$s_w^2 = MS_w = \frac{SS_w}{N-k}$	
Total	$SS_{total} = \sum \sum (X_{ij} - \bar{X}_{..})^2$	N-1		

$$\eta^2 = \frac{SS_b}{SS_{\text{total}}} . F = \frac{(\bar{X}_i - \bar{X}_j)^2}{MSE(\frac{1}{n_i} + \frac{1}{n_j})} \text{ Reject Ho if } F \geq (k-1) F_{k-1, N-k} . \text{ Do not reject Ho if } F < (k-1) F_{k-1, N-k} . q_k = \frac{(\bar{X}_1 - \bar{X}_{k'})}{\sqrt{\frac{s_w^2}{n}}} = \frac{(\bar{X}_1 - \bar{X}_{k'})}{\sqrt{\frac{MS_{\text{error}}}{n}}} \text{ Reject Ho if } q_k \geq q_\alpha(k, df_{\text{error}}) \text{ Do not reject Ho if } q_k < q_\alpha(k, df_{\text{error}}).$$

Source of Variation	Sums of Squares (SS)	Degrees of Freedom (df)	Mean Squares (MS)	F
Between Subjects	$SS_{\text{subj}} = \sum k (\bar{X}_{s.} - \bar{X}_{..})^2$	n-1		
Between Treatments	$SS_b = \sum n (\bar{X}_{.j} - \bar{X}_{..})^2$	k-1	$s_b^2 = MS_b = \frac{SS_b}{k-1}$	$F = \frac{MS_b}{MS_w}$
Within	$SS_w = SS_{\text{total}} - SS_{\text{subj}} - SS_b$	(n-1)(k-1)	$s_w^2 = MS_w = \frac{SS_w}{(n-1)(k-1)}$	
Total	$SS_{\text{total}} = \sum \sum (X_{s.j} - \bar{X}_{..})^2$	nk-1		

$$r = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}}, Cov(X, Y) = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{n-1}. t = r \sqrt{\frac{n-2}{1-r^2}}, df = n-2.$$

$$\hat{\beta}_1 = r \sqrt{\frac{Var(Y)}{Var(X)}}, \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}. R^2 = \frac{SS_{\text{model}}}{SS_{\text{total}}} = \frac{SS_{\text{regression}}}{SS_{\text{total}}}.$$

$$\ln \left\{ \frac{p}{(1-p)} \right\} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon.$$

$$Z = \frac{T - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}, T = \text{sum of the positive ranks. } Z = \frac{S - \frac{n_1(n_1+n_2+1)}{2}}{\sqrt{\frac{n_1 n_2 (n_1+n_2+1)}{12}}}, S = \text{sum}$$

of the ranks corresponding to treatment 1.