81/12B

Statistical Formulae

Hypothesis testing

Sam. mean & variance	$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $s^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2\right]$
Accept region & CI	(all area in α refer to lower area)
$H_1: \ \mu > \mu_0$	$\bar{x} \in (-\infty, \mu_0 + t_{n-1,1-\alpha}\frac{s}{\sqrt{n}})$ or $\mu_0 \in (\bar{x} - t_{n-1,1-\alpha}\frac{s}{\sqrt{n}}, \infty)$
$H_1: \ \mu < \mu_0$	$\bar{x} \in (\mu_0 - t_{n-1,1-\alpha}\frac{s}{\sqrt{n}}, \infty)$ or $\mu_0 \in (-\infty, \bar{x} + t_{n-1,1-\alpha}\frac{s}{\sqrt{n}})$
$H_1: \ \mu \neq \mu_0$	$\bar{x} \in (\mu_0 - t_{n-1,1-\frac{\alpha}{2}}\frac{s}{\sqrt{n}}, \mu_0 + t_{n-1,1-\frac{\alpha}{2}}\frac{s}{\sqrt{n}})$ or $\mu_0 \in (\bar{x} - t_{n-1,1-\frac{\alpha}{2}}\frac{s}{\sqrt{n}}, \bar{x} + t_{n-1,1-\frac{\alpha}{2}}\frac{s}{\sqrt{n}})$
Rejection region	
$H_1: \ \mu > \mu_0$	$\bar{x} \geq \mu_0 + t_{n-1,1-\alpha}\frac{s}{\sqrt{n}}$
$H_1: \ \mu < \mu_0$	$\bar{x} \leq \mu_0 - t_{n-1,1-\alpha}\frac{s}{\sqrt{n}}$
$H_1: \ \mu \neq \mu_0$	$\bar{x} \leq \mu_0 - t_{n-1,1-\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$ or $\bar{x} \geq \mu_0 + t_{n-1,1-\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$
Note:	Above applies if σ^2 is unknown. If σ^2 is known, replace s by σ and t_{n-1} by z.
Type I error	$\alpha = \Pr(\text{reject } H_0 \mid H_0 \text{ is true.}) = \Pr(z_0 \geq z_\alpha \mid \mu = \mu_0)$
Type II error	$\beta_\alpha(\mu_1) = \Pr(\text{accept } H_0 \mid H_0 \text{ is false and } \mu = \mu_1) = \Pr(z_0 < z_\alpha \mid \mu = \mu_1 \geq \mu_0)$
Power	$\text{Power}(\mu_1) = \Pr(\text{reject } H_0 \mid H_0 \text{ is false and } \mu = \mu_1) = \Pr(z_0 \geq z_\alpha \mid \mu = \mu_1 \geq \mu_0)$
Note:	Above applies only if σ^2 is known and $H_1: \ \mu > \mu_0$.

Random variable

Binomial rv:	If $X \sim \mathcal{B}(n,p)$, $P(X=x) = \binom{n}{x} p^x(1-p)^{n-x}$, $E(X) = np$ & $Var(X) = np(1-p)$.
Normal rv:	If $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$, $P(X_i < x) = P(Z < \frac{x-\mu_i}{\sigma_i})$ & $aX_1 + bX_2 \sim \mathcal{N}(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2)$.
Central limit t.:	If $X_i \sim$ any $\mathcal{D}(\mu, \sigma^2)$ indept., $\overline{X} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$ & $\sum_{i=1}^{n} X_i \sim \mathcal{N}(n\mu, n\sigma^2)$ if $n > 25$.

Assumptions	**Confidence Interval for one sample** (for two-, upper- and lower-sided H_1 resp.)
normal; σ^2 known	For μ: $(\bar{x} - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \ \bar{x} + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}})$, $(\bar{x} - z_{1-\alpha}\frac{\sigma}{\sqrt{n}}, \ \infty)$; $(-\infty, \ \bar{x} + z_{1-\alpha}\frac{\sigma}{\sqrt{n}})$
normal; σ^2 unknown	For μ: $(\bar{x} - t_{1-\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}}, \ \bar{x} + t_{1-\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}})$, $(\bar{x} - t_{1-\alpha,n-1}\frac{s}{\sqrt{n}}, \ \infty)$; $(-\infty, \ \bar{x} + t_{1-\alpha,n-1}\frac{s}{\sqrt{n}})$
normal	For σ^2: $(\frac{(n-1)s^2}{\chi^2_{n-1,1-\alpha/2}}, \ \frac{(n-1)s^2}{\chi^2_{n-1,\alpha/2}})$, $(\frac{(n-1)s^2}{\chi^2_{n-1,1-\alpha}}, \ \infty)$, $(0, \ \frac{(n-1)s^2}{\chi^2_{n-1,\alpha}})$

Assumptions	**Confidence Interval for two samples** (for two-, upper- and lower-sided H_1 resp.)
normal; σ_1^2, σ_2^2 known	For $\mu_1 - \mu_2$: $(\bar{x}_1 - \bar{x}_2 - z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}, \ \bar{x}_1 - \bar{x}_2 + z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}})$, $(\bar{x}_1 - \bar{x}_2 - z_{1-\alpha}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}, \ \infty)$; $(-\infty, \ \bar{x}_1 - \bar{x}_2 + z_{1-\alpha}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}})$
normal; σ_1^2, σ_2^2 unknown	For $\mu_1 - \mu_2$: $(\bar{x}_1 - \bar{x}_2 - t_{n_1+n_2-2,1-\frac{\alpha}{2}} s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, \ \bar{x}_1 - \bar{x}_2 + t_{n_1+n_2-2,1-\frac{\alpha}{2}} s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}})$, $(\bar{x}_1 - \bar{x}_2 - t_{n_1+n_2-2,1-\alpha} s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, \ \infty)$; $(-\infty, \ \bar{x}_1 - \bar{x}_2 + t_{n_1+n_2,1-\alpha} s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}})$
normal	For $\frac{\sigma_1^2}{\sigma_2^2}$ $(\sigma_1^2 > \sigma_2^2)$: $(\frac{s_1^2}{s_2^2}F^{-1}_{n_1-1,n_2-1,1-\frac{\alpha}{2}}, \ \frac{s_1^2}{s_2^2}F^{-1}_{n_1-1,n_2-1,\frac{\alpha}{2}})$, $(\frac{s_1^2}{s_2^2}F^{-1}_{n_1-1,n_2-1,1-\alpha}, \ \infty)$

One and two sample parametric test

H_a	Test statistics	RR	p-value	Assumption
One sample: $H_0:\ \mu = \mu_0$ or $H_0:\ \sigma^2 = \sigma_0^2$				
$\mu > \mu_0$ $\mu < \mu_0$ $\mu \neq \mu_0$	$z_0 = \frac{\overline{y} - \mu_0}{\sigma/\sqrt{n}}$ or $\hat{\sigma} = s$	$z_0 \geq z_{1-\alpha}$ $z_0 \leq -z_{1-\alpha}$ $\lvert z_0\rvert \geq z_{1-\alpha/2}$	$P(Z \geq z_0)$ $P(Z \leq z_0)$ $2(Z \geq \lvert z_0\rvert)$	normality, σ^2 known; or large n, σ^2 unknown
$\mu > \mu_0$ $\mu < \mu_0$ $\mu \neq \mu_0$	$t_0 = \frac{\overline{y} - \mu_0}{s/\sqrt{n}}$	$t_0 \geq t_{1-\alpha,n-1}$ $t_0 \leq -t_{1-\alpha,n-1}$ $\lvert t_0\rvert \geq t_{1-\alpha/2,n-1}$	$P(t_{n-1} \geq t_0)$ $P(t_{n-1} \leq t_0)$ $2P(t_{n-1} \geq \lvert t_0\rvert)$	n small; normality σ^2 unknown
$\sigma^2 > \sigma_0^2$ $\sigma^2 < \sigma_0^2$ $\sigma^2 \neq \sigma_0^2$	$\chi_0^2 = \frac{(n-1)s^2}{\sigma_0^2}$	$\chi_0^2 \geq \chi^2_{1-\alpha,n-1}$ $\chi_0^2 \leq \chi^2_{\alpha,n-1}$ $\chi_0^2 \geq \chi^2_{1-\alpha/2,n-1}$ or $\chi_0^2 \leq \chi^2_{\alpha/2,n-1}$	$P(\chi^2_{n-1} \geq \chi_0^2)$ $P(\chi^2_{n-1} \leq \chi_0^2)$ $2\min[P(\chi^2_{n-1} \geq \chi_0^2),$ $P(\chi^2_{n-1} \leq \chi_0^2)]$	n large or normality; σ^2 unknown
Two samples: $H_0:\ \mu_1 - \mu_2 = d_0$ or $H_0:\ \sigma_1^2 = \sigma_2^2$				
$\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 \neq d_0$	$z_0 = \frac{(\overline{y}_1 - \overline{y}_2) - d_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	$z_0 \geq z_{1-\alpha}$ $z_0 \leq -z_{1-\alpha}$ $\lvert z_0\rvert \geq z_{1-\alpha/2}$	$P(Z \geq z_0)$ $P(Z \leq z_0)$ $2(1 - \Phi(\lvert z_0\rvert))$	normality σ_1^2, σ_2^2 known; independence
$\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 \neq d_0$	$t_0 = \frac{(\overline{y}_1 - \overline{y}_2) - d_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ $df = \frac{\left[s_1^2/n_1 + s_2^2/n_2\right]^2}{\frac{(s_1^2/n_1)^2}{n_1-1} + \frac{(s_2^2/n_2)^2}{n_2-1}}$	$t_0 \geq t_{1-\alpha,df}$ $t_0 \leq -t_{1-\alpha}$ $\lvert t_0\rvert \geq t_{1-\alpha/2}$	$P(t_{df} \geq t_0)$ $P(t_{df} \leq t_0)$ $2P(t_{df} \geq \lvert t_0\rvert)$	n_1, n_2 large; normality, σ_1^2, σ_2^2 unknown; independence
$\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 \neq d_0$	$t_0 = \frac{(\overline{y}_1 - \overline{y}_2) - d_0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ $s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$	$t_0 \geq t_{1-\alpha,n_1+n_2-2}$ $t_0 \leq -t_{1-\alpha,n_1+n_2-2}$ $\lvert t_0\rvert \geq t_{1-\frac{\alpha}{2},n_1+n_2-2}$	$P(t_{n_1+n_2-2} \geq t_0)$ $P(t_{n_1+n_2-2} \leq t_0)$ $2P(t_{n_1+n_2-2} \geq \lvert t_0\rvert)$	n_1, n_2 small; normality; independence; σ_1^2, σ_2^2 unknown but equal
$\mu_d > d_0$ $\mu_d < d_0$ $\mu_d \neq d_0$	$t_0 = \frac{(\overline{y}_1 - \overline{y}_2) - d_0}{s_d/\sqrt{n}}$	$t_0 \geq t_{1-\alpha,n-1}$ $t_0 \leq -t_{1-\alpha,n-1}$ $\lvert t_0\rvert \geq t_{1-\alpha/2,n-1}$	$P(t_{n-1} \geq t_0)$ $P(t_{n-1} \leq t_0)$ $2P(t_{n-1} \geq \lvert t_0\rvert)$	n small; n pairs and dependence; normality for d_i; σ_d^2 unknown
$\sigma_1^2 > \sigma_2^2$ $\sigma_1^2 \neq \sigma_2^2$	$f_0 = \frac{s_1^2}{s_2^2}$ $(s_1^2 \geq s_2^2)$	$f_0 \geq F_{1-\alpha,n_1-1,n_2-1}$ $f_0 \geq F_{1-\frac{\alpha}{2},n_1-1,n_2-1}$ or $f_0 \leq F_{\frac{\alpha}{2},n_1-1,n_2-1}$, i.e. $f_0^{-1} \geq F_{1-\frac{\alpha}{2},n_2-1,n_1-1}$	$P(F_{n_1-1,n_2-1} \geq f_0)$ $2\min[P(F_{n_1-1,n_2-1} \geq f_0),$ $P(F_{n_2-1,n_1-1} \geq \frac{1}{f_0})]$	normality σ_1^2, σ_2^2 unknown

Nonparametric test

H_a	**Test statistics**	**RR**	**p-value**	**Assumption**
One sample or match-pair sample:			$H_0:\ p=p_0$ or $H_0:\ \mu=\mu_0$	
$p>p_0$ $p<p_0$ $p\neq p_0$	Binomial test: $x_0=\#$success	NA	$P(B_{n,p_0}\ge x_0)=\sum_{i=x_0}^{n}\binom{n}{i}p_0^i(1-p_0)^{n-i}$ $P(B_{n,p_0}\le x_0)=\sum_{i=0}^{x_0}\binom{n}{i}p_0^i(1-p_0)^{n-i}$ $2\min[P(B_{n,p_0}\ge x_0),$ $P(B_{n,p_0}\le x_0)]$	$x_0\sim B(n,p_0)$ $x_0\sim N(np_0,np_0(1-p_0))$ if n large
$\mu>\mu_0$ $\mu<\mu_0$ $\mu\neq\mu_0$	Sign test: $x_0=\#(d_i>0)$ $d_i=y_i-\mu_0$ Discard $d_i=0$ & revise n	NA	$P(B_{n,\frac{1}{2}}\ge x_0)=\frac{1}{2^n}\sum_{i=x_0}^{n}\binom{n}{i}$ $P(B_{n,\frac{1}{2}}\le x_0)=\frac{1}{2^n}\sum_{i=0}^{x_0}\binom{n}{i}$ $2\min[P(B_{n,\frac{1}{2}}\ge x_0),P(B_{n,\frac{1}{2}}\le x_0)]$	Symmetric dist. for y_i $x_0\sim B(n,\frac{1}{2})$ $x_0\sim N(\frac{n}{2},\frac{n}{4})$ if n large
$\mu>\mu_0$ $\mu<\mu_0$ $\mu\neq\mu_0$	Wilcoxon sign rank test $w_0=\min(w_0^+,w_0^-)$ $w_0^+=\sum_{i,d_i>0}r_i,\ \ w_0^-=\sum_{i,d_i<0}r_i$ Discard $d_i=0$ & revise n	NA	$P(W_n\le\frac{n(n+1)}{2}-w_0^+)$ or $P(Z\ge\frac{w_0^+-E(w^+)}{\sqrt{V(w^+)}})$ if ties $P(W_n\le w_0^+)$ or $P(Z\le\frac{w_0^+-E(w^+)}{\sqrt{V(w^+)}})$ $2P(W_n\le w_0)$ or $2P(Z\ge\frac{\lvert w_0^+-E(w^+)\rvert}{\sqrt{V(w^+)}})$	Symmetric dist. for y_i $w^+\sim N(E(w^+),V(w^+))$ $E(w^+)=\frac{n(n+1)}{4}$ $V(w^+)=\frac{1}{4}\sum_{i,d_i\neq0}r_i^2$ if ties
Two samples:			$H_0:\ \mu_1=\mu_2$	
$\mu_1>\mu_2$ $\mu_1<\mu_2$ $\mu_1\neq\mu_2$	Wilcoxon rank-sum test $w_0=\sum_{i=1}^{n_1}r_{y_1,i}$ $r_{y_1,i}$: rank for y_{1i}, $n_1\le n_2$	NA	$P(W_{n_1,n_2}\le n_1(N+1)-w_0)$ or $P(Z\ge\frac{w_0-E(w)}{\sqrt{V(w)}})$ if ties $P(W_{n_1,n_2}\le w_0)$ or $P(Z\le\frac{w_0-E(w)}{\sqrt{V(w)}})$ $2P(W_{n_1,n_2}\le w_0)$ if $w_0\le E(w)$ or $2P(W_{n_1,n_2}\le n_1(N+1)-w_0)$ or $2P(Z\ge\frac{\lvert w_0-E(w)\rvert}{\sqrt{V(w)}})$	Same dist. for each group $w\sim N(E(w),V(w))$ $E(w)=\frac{n_1}{2}(N+1)$ $V(w)=\dfrac{n_1n_2}{N(N-1)}\left[\sum_{i=1}^{N}r_i^2-\frac{N(N+1)^2}{4}\right]$ if $n_1,n_2\ge10$ or with ties
Three or more samples:			$H_0:\ \mu_1=\cdots=\mu_c$	
At least 1 equality not hold	Kruskal-Wallis test $k_0=(N-1)\times\dfrac{\sum_{j=1}^{c}n_j\bar{r}_{\cdot j}^2-N\bar{r}^2}{\sum_{j=1}^{c}\sum_{i=1}^{n_j}r_{ij}^2-N\bar{r}^2}$ r_{ij}: rank over all samples	$k_0>\chi^2_{1-\alpha,c-1}$	$P(\chi^2_{c-1}\ge k_0)$	Same dist. for each group $k_0=\dfrac{12}{N(N+1)}\sum_{j=1}^{c}n_j\bar{r}_{\cdot j}^2$ $-3(N+1)$ if no ties
Two way data:			$H_0:\ \beta_1=\cdots=\beta_c$	
At least 1 equality not hold	Friedman test $q_0=(rc-r)\times\dfrac{r\sum_{j=1}^{c}\bar{r}_{\cdot j}^2-rc\bar{r}^2}{\sum_{i=1}^{r}\sum_{j=1}^{c}r_{ij}^2-rc\bar{r}^2}$ r_{ij}: rank over row i	$q_0>\chi^2_{1-\alpha,c-1}$	$P(\chi^2_{c-1}\ge q_0)$	$q_0=\dfrac{12r}{c(c+1)}\sum_{j=1}^{c}\bar{r}_{\cdot j}^2-$ $-3r(c+1)$ if no ties

One way ANOVA table for completely randomized design

Source	df	SS	MS	F
Groups	$c-1$	$SST=\sum_{j=1}^{c} n_j\overline{y}_j^2 - n\overline{y}^2$	$MST=\frac{SST}{c-1}$	$F=\frac{MST}{MSR}$
Residuals	$n-c$	$SSR=\sum_{j=1}^{c}(n_j-1)s_j^2 = SST_o - SST$	$MSR=\frac{SSR}{n-c}=s^2$	
Total	$n-1$	$SST_o=\sum_{j=1}^{c}\sum_{i=1}^{n_j} y_{ij}^2 - n\overline{y}^2$		

$\overline{y}_j=\frac{1}{n_j}\sum_{i=1}^{n_j} y_{ij}$, $n=\sum_{j=1}^{c} n_j$ and $\overline{y}=\frac{1}{n}\sum_{j=1}^{c}\sum_{i=1}^{n_j} y_{ij}$.

Two way ANOVA table for randomized block design

Source	df	SS	MS	F
Treatments	$c-1$	$SST=r\sum_{j=1}^{c}\overline{y}_{\cdot j}^2 - n\overline{y}^2$	$MST=\frac{SST}{c-1}$	$F_t=\frac{MST}{MSR}$
Blocks	$r-1$	$SSB=c\sum_{i=1}^{r}\overline{y}_{i\cdot}^2 - n\overline{y}^2$	$MSB=\frac{SSB}{r-1}$	$F_b=\frac{MSB}{MSR}$
Residuals	$(r-1)(c-1)$	$SSR=SST_o-SST-SSB$	$MSR=\frac{SSR}{(r-1)(c-1)}$	
Total	$n-1$	$SST_o=\sum_{i=1}^{r}\sum_{j=1}^{c} y_{ij}^2 - n\overline{y}^2$		

$\overline{y}_{i\cdot}=\frac{1}{c}\sum_{j=1}^{c} y_{ij}$, $\overline{y}_{\cdot j}=\frac{1}{r}\sum_{i=1}^{r} y_{ij}$, $n=rc$ and $\overline{y}=\frac{1}{n}\sum_{i=1}^{r}\sum_{j=1}^{c} y_{ij}$.

Two way ANOVA table for factorial design

Source	df	SS	MS	F
Treatments	$c-1$	$SST=rm\sum_{j=1}^{c}\overline{y}_{\cdot j\cdot}^2 - n\overline{y}^2$	$MST=\frac{SST}{c-1}$	$F_t=\frac{MST}{MSR}$
Blocks	$r-1$	$SSB=cm\sum_{i=1}^{r}\overline{y}_{i\cdot\cdot}^2 - n\overline{y}^2$	$MSB=\frac{SSB}{r-1}$	$F_b=\frac{MSB}{MSR}$
Interactions	$(r-1)(c-1)$	$SSI=SST_o-SST-SSB-SSR$	$MSI=\frac{SSI}{(r-1)(c-1)}$	$F_i=\frac{MSI}{MSR}$
Residuals	$rc(m-1)$	$SSR=(m-1)\sum_{i=1}^{r}\sum_{j=1}^{c} s_{ij}^2$	$MSR=\frac{SSR}{rc(m-1)}$	
Total	$n-1$	$SST_o=\sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{m} y_{ijk}^2 - n\overline{y}^2$		

$\overline{y}_{i\cdot\cdot}=\frac{1}{cm}\sum_{j=1}^{c}\sum_{k=1}^{m} y_{ijk}$, $\overline{y}_{\cdot j\cdot}=\frac{1}{rm}\sum_{j=1}^{r}\sum_{k=1}^{m} y_{ijk}$, $n=rcm$, $\overline{y}=\frac{1}{n}\sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{m} y_{ijk}$ and $s_{ij}^2=\frac{1}{m-1}(\sum_{k=1}^{m} y_{ijk}^2 - m\overline{y}_{ij\cdot}^2)$.

ANOVA table for regression

Source	df	SS	MS	F
Regression	1	$SST=S_{xy}^2/S_{xx}$	$MST=SST$	$F=\frac{MST}{MSR}$
Residuals	$n-2$	$SSR=S_{yy}-S_{xy}^2/S_{xx}$	$MSR=\frac{SSR}{n-2}=s^2$	
Total	$n-1$	$SST_o=S_{yy}$		

H_a	**Test statistic**	**RR**	**p-value**	**Assumptions**
One-way ANOVA table:		$H_0:\ \mu_1 = \cdots = \mu_c$		
At least 1 equality not hold	$f_0 = \dfrac{MST}{MSR}$	$f_0 > f_{1-\alpha,c-1,n-c}$	$P(f_{\alpha,c-1,n-c} \geq f_0)$	$y_{ij} \sim N(\mu_i, \sigma^2)$
Pairwise comparison:		$H_0:\ \mu_l = \mu_k$		
$\mu_l \neq \mu_k$	$t_{l,k} = \dfrac{\overline{y}_l - \overline{y}_k}{\sqrt{MSR}\sqrt{\frac{1}{n_l} + \frac{1}{n_k}}}$	$\lvert t_{l,k}\rvert > t_{1-\alpha/2,n-c}$	$2P(t_{n-c} \geq \lvert t_{l,k}\rvert)$	$y_{ij} \sim N(\mu_i, \sigma^2)$
Multiple comparison (Bonferroni):		$H_0:\ \mu_1 = \cdots = \mu_c$		
At least 1 equality not hold	$t_0 = \max\limits_{l,k} \dfrac{\overline{y}_l - \overline{y}_k}{\sqrt{MSR}\sqrt{\frac{1}{n_l} + \frac{1}{n_k}}}$	$\lvert t_0\rvert > t_{1-\alpha/r,n-c}$ $r = c(c-1)/2$	$2P(t_{n-c} \geq \lvert t_0\rvert)$	$y_{ij} \sim N(\mu_i, \sigma^2)$
Regression:		$H_0:\ \beta = \beta_0$ or $H_0:\ \rho = 0$		
$\beta > \beta_0$ $\beta < \beta_0$ $\beta \neq \beta_0$	$t_0 = \dfrac{\hat{\beta} - \beta_0}{s/\sqrt{S_{xx}}}$ $\overset{\beta_0=0}{=} \dfrac{\sqrt{n-2}S_{xy}}{\sqrt{S_{xx}S_{yy} - S_{xy}^2}}$ $s^2 = \frac{SSR}{n-2}$	$t_0 > t_{1-\alpha,n-2}$ $t_0 < -t_{1-\alpha,n-2}$ $\lvert t_0\rvert > t_{1-\alpha/2,n-2}$	$P(t_{n-2} \geq t_0)$ $P(t_{n-2} \leq t_0)$ $2P(t_{n-2} \geq \lvert t_0\rvert)$	$e_i \sim N(0, \sigma^2)$ $\mathrm{Cov}(e_i, e_j) = 0$
$\rho > 0$ $\rho < 0$ $\rho \neq 0$	$t_0 = \dfrac{r\sqrt{n-2}}{\sqrt{1-r^2}}$	$t_0 > t_{1-\alpha,n-2}$ $t_0 < -t_{1-\alpha,n-2}$ $\lvert t_0\rvert > t_{1-\alpha/2,n-2}$	$P(t_{n-2} \geq t_0)$ $P(t_{n-2} \leq t_0)$ $2P(t_{n-2} \geq \lvert t_0\rvert)$	$e_i \sim N(0, \sigma^2)$ $\mathrm{Cov}(e_i, e_j) = 0$
Chi-square goodness-of-fit test:		$H_0:\ p_1 = p_{10}, \ldots, p_c = p_{c0}$ where p_{j0} may be calculated from normal or other dist.		
At least 1 equality not hold	$\chi_0^2 = \sum\limits_i \dfrac{(y_i - np_{i0})^2}{np_{i0}}$	$\chi_0^2 > \chi^2_{1-\alpha,c-k-1}$ $k =$ no. of par.	$P(\chi^2_{c-k-1} \geq \chi_0^2)$	$E_i \geq 5$ for all i
Chi-square test for dependency:		H_0 : Factor i & factor j are independent		
Dependent	$\chi_0^2 = \sum\limits_{i,j} \dfrac{(y_{ij} - \frac{y_{i\cdot}y_{\cdot j}}{n})^2}{\frac{y_{i\cdot}y_{\cdot j}}{n}}$	$\chi_0^2 > \chi^2_{1-\alpha,(c-1)(r-1)}$	$P(\chi^2_{(c-1)(r-1)} \geq \chi_0^2)$	$E_{ij} \geq 5$ for all i, j

Linear Regression

Regression line	$\hat{y} \;=\; \hat{\alpha} + \hat{\beta}\, x$
Slope	$\hat{\beta} \;=\; \dfrac{S_{xy}}{S_{xx}} = \dfrac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}, \quad \hat{\beta} \sim \mathcal{N}\left(\beta, \dfrac{\sigma^2}{S_{xx}}\right)$
y-intercept	$\hat{\alpha} \;=\; \bar{y} - \hat{\beta}\,\bar{x}, \quad \hat{\alpha} \sim \mathcal{N}\left(\alpha, \sigma^2\left[\dfrac{1}{n} + \dfrac{\bar{x}^2}{S_{xx}}\right]\right)$
Correlation coeff.	$r \;=\; \dfrac{S_{xy}}{\sqrt{S_{xx}\cdot S_{yy}}} = \dfrac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sqrt{\left[\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right]\left[\sum_{i=1}^{n} y_i^2 - n\bar{y}^2\right]}}$
Residual variance	$s^2 \;=\; \dfrac{SSR}{n-2}, \quad SSR = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = S_{yy} - \hat{\beta}S_{xy}$
Coeff. of determination	$R^2 \;=\; \dfrac{SST}{SST_o} = r^2, \quad SST = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 = \hat{\beta}S_{xy}, \quad SST_o = \sum_{i=1}^{n}(y_i - \bar{y})^2 = S_{yy}$
Estimated mean $E(\widehat{Y}_0)$	$E(\widehat{Y}_0) = \hat{\alpha} + \hat{\beta}x_0, \quad E(\widehat{Y}_0) \;\sim\; \mathcal{N}\left(\alpha + \beta\, x_0,\, \sigma^2\left[\dfrac{1}{n} + \dfrac{(x_0-\bar{x})^2}{S_{xx}}\right]\right)$
Predicted point $\widehat{Y}_0$	$\widehat{Y}_0 = \hat{\alpha} + \hat{\beta}x_0, \quad \widehat{Y}_0 \;\sim\; \mathcal{N}\left(\alpha + \beta\, x_0,\, \sigma^2\left[1 + \dfrac{1}{n} + \dfrac{(x_0-\bar{x})^2}{S_{xx}}\right]\right)$
Conf. interval for β	$\left(\hat{\beta} - t_{n-2,1-\frac{\alpha}{2}}\dfrac{s}{\sqrt{S_{xx}}},\; \hat{\beta} + t_{n-2,1-\frac{\alpha}{2}}\dfrac{s}{\sqrt{S_{xx}}}\right)$
Conf. interval for α	$\left(\hat{\alpha} - t_{n-2,1-\frac{\alpha}{2}}\; s\sqrt{\dfrac{1}{n} + \dfrac{\bar{x}^2}{S_{xx}}},\; \hat{\alpha} + t_{n-2,1-\frac{\alpha}{2}}\; s\sqrt{\dfrac{1}{n} + \dfrac{\bar{x}^2}{S_{xx}}}\right)$
Est. interval for $E(Y_0)$	$\left(\hat{\alpha} + \hat{\beta}x_0 - t_{n-2,1-\frac{\alpha}{2}}\; s\sqrt{\dfrac{1}{n} + \dfrac{(x_0-\bar{x})^2}{S_{xx}}},\; \hat{\alpha} + \hat{\beta}x_0 + t_{n-2,1-\frac{\alpha}{2}}\; s\sqrt{\dfrac{1}{n} + \dfrac{(x_0-\bar{x})^2}{S_{xx}}}\right)$
Pred. interval for Y_0	$\left(\hat{\alpha} + \hat{\beta}x_0 - t_{n-2,1-\frac{\alpha}{2}}\; s\sqrt{1 + \dfrac{1}{n} + \dfrac{(x_0-\bar{x})^2}{S_{xx}}},\; \hat{\alpha} + \hat{\beta}x_0 + t_{n-2,1-\frac{\alpha}{2}}\; s\sqrt{1 + \dfrac{1}{n} + \dfrac{(x_0-\bar{x})^2}{S_{xx}}}\right)$