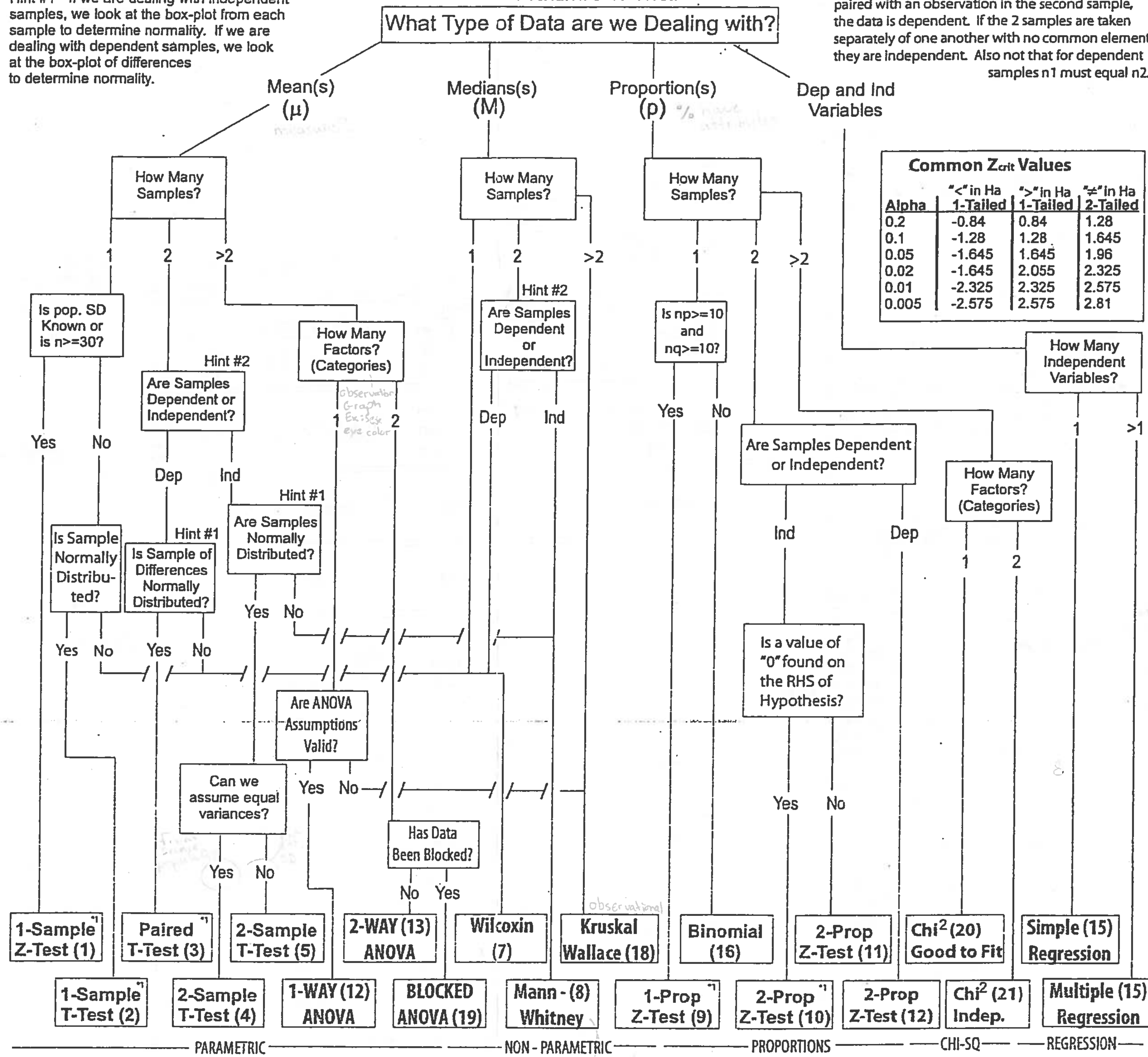
www.examreviews.ca

Hint #1 - If we are dealing with independent samples, we look at the box-plot from each sample to determine normality. If we are dealing with dependent samples, we look at the box-plot of differences to determine normality.

Hint #2 - If each piece of data from the first sample is paired with an observation in the second sample, the data is dependent. If the 2 samples are taken separately of one another with no common element, they are independent. Also not that for dependent samples n1 must equal n2.



Common Z_{crit} Values

Alpha	"<" in Ha 1-Tailed	">" in Ha 1-Tailed	"≠" in Ha 2-Tailed
0.2	-0.84	0.84	1.28
0.1	-1.28	1.28	1.645
0.05	-1.645	1.645	1.96
0.02	-1.645	2.055	2.325
0.01	-2.325	2.325	2.575
0.005	-2.575	2.575	2.81

16

Binomial

Ho:p = 0.3 Ho:p = 0.4
Ha:p < 0.3 Ha:p ≠ 0.4

$P(X=x) = nCr^x p^x q^{n-x}$

n=# of trials (sample size)
p=probability of a success
q=probability of a failure
q=1-p
x=# of successes in "n" trials
Reject Ho if $p < \alpha$

Required Sample Size (always round up)

-proportion $n=\frac{Z_{\alpha/2}^2 \hat{p}\hat{q}}{E^2}$

*$\hat{p}$, $\hat{q}$ and E are always entered as decimals
*If $\hat{p}$ and $\hat{q}$ are not given use values of 0.5

-means (σ known) $n=\frac{Z_{\alpha/2}^2 \sigma^2}{E^2}$

-means (σ unknown) $n=\frac{T_{\alpha/2}^2 s^2}{E^2}$

20

Chi Sq-Goodness to Fit

Ho: Data follows described dist.
Ha: Data follows some other dist.

$Chi^2_{stat}=\sum_{i=1}^{k}\frac{(o_i-e_i)^2}{e_i}$ o=observed e=expected df=r-1

Reject Ho if $|Chi^2_{stat}| > |Chi^2_{crit}|$ or $p < \alpha$.

Step 1 - Calculate the total of observed counts
Step 2 - Derive expected values as described in the question
Follow Step 3,4,5 in box below.

21

Chi Sq-Test of independence

Ho: Factor A is independent of Factor B
Ha: Factor A is not independent of Factor B

$Chi^2_{stat}=\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(o_{ij}-e_{ij})^2}{e_{ij}}$ o=observed e=expected df=(r-1)(c-1)

$e_{ij}=\frac{(i \text{ row total})(j \text{ column total})}{\text{Total sample size}}$

Reject Ho if $|Chi^2_{stat}| > |Chi^2_{crit}|$ or $p < \alpha$.

Step 1 - Calculate all row totals, col totals and grand total
Step 2 - Calculate expected value for each cell
Step 3 - Calculate Chi² for each cell
Step 4 - Add the Chi² from each cell to get your Chi^2_{stat}
Step 5 - Finish Hypothesis Test.

[1]FPCF - If a sample(s) is more than 5% of the population we must apply the FPCF. Therefore, if n/N>0.05, we multiply the denominator by SQRT((N-n)/(N-1)).

→ Expected cell condition (Expected count > 5) no casual link cuz observational study.

$MST = \frac{SST}{k-1}$ $MSE = \frac{SSE}{n-K}$

*see bottom of page for 1 tailed

#	1-Tailed Hyp	2-Tailed Hyp	Assumptions	Formulas	Decision (2-tailed)	CI (2-Sided)	CI (1-Sided Upper)	CI (1-Sided Lower)
1	**1 Sample Z-Test** Ho:$\mu=5$ Ha:$\mu>5$	Ho:$\mu=3$ Ha:$\mu\neq3$	-Random Sample -Sample mean is normally distributed	$Z_{stat}=\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$	Reject Ho if $\lvert Z_{stat}\rvert > \lvert Z_{crit}\rvert$ or $p<\alpha$	$\bar{x} \pm Z_{\alpha/2}\,\sigma/\sqrt{n}$	if ">" in Ha UB="infinite" LB=$\bar{x} - Z_\alpha\,\sigma/\sqrt{n}$	if "<" in Ha UB=$\bar{x} + Z_\alpha\,\sigma/\sqrt{n}$ LB="0" or "-infinite"
2	**1 Sample T-Test** Ho:$\mu=6$ Ha:$\mu>6$	Ho:$\mu=8$ Ha:$\mu\neq8$	-Random Sample -Sample mean is normally distributed	$T_{stat}=\frac{\bar{x}-\mu}{s/\sqrt{n}}$ df = n -1	Reject Ho if $\lvert T_{stat}\rvert > \lvert T_{crit}\rvert$ or $p<\alpha$	$\bar{x} \pm T_{\alpha/2}\,s/\sqrt{n}$	if ">" in Ha UB="infinite" LB=$\bar{x} - T_\alpha\,s/\sqrt{n}$	if "<" in Ha UB=$\bar{x} + T_\alpha\,s/\sqrt{n}$ LB="0" or "-infinite"
3	**Paired T-Test** Ho:$\mu_d=0$ Ha:$\mu_d<0$	Ho:$\mu_d=0$ Ha:$\mu_d\neq0$	-Dependent Samples -Sample mean of differences is normally distributed	$T_{stat}=\frac{\bar{d}-\mu_d}{s_d/\sqrt{n}}$ df = n -1	Reject Ho if $\lvert T_{stat}\rvert > \lvert T_{crit}\rvert$ or $p<\alpha$	$\bar{d} \pm T_{\alpha/2}\,s_d/\sqrt{n}$	if ">" in Ha UB="infinite" LB=$\bar{d} - T_\alpha\,s_d/\sqrt{n}$	if "<" in Ha UB=$\bar{d} + T_\alpha\,s_d/\sqrt{n}$ LB="-infinite"
4	**2 Sample T-Test** Ho:$\mu_1-\mu_2=0$ Ha:$\mu_1-\mu_2<0$	Ho:$\mu_1-\mu_2=0$ Ha:$\mu_1-\mu_2\neq0$	-Independent Samples -Population Variances are Equal -Sample means are both normally distributed	$T_{stat}=\frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{s_p\sqrt{1/n_1+1/n_2}}$ $s_p=\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$ df $=n_1+n_2-2$	Reject Ho if $\lvert T_{stat}\rvert > \lvert T_{crit}\rvert$ or $p<\alpha$ Sufficient evidence	$(\bar{x}_1-\bar{x}_2) \pm T_{\alpha/2}^{*}\, s_p\sqrt{1/n_1+1/n_2}$	if ">" in Ha UB="infinite" LB=$(\bar{x}_1-\bar{x}_2)-T_\alpha^{*}\, s_p\sqrt{1/n_1+1/n_2}$	if "<" in Ha UB=$(\bar{x}_1-\bar{x}_2)+T_\alpha^{*}\, s_p\sqrt{1/n_1+1/n_2}$ LB="-infinite"
5	**2 Sample T-Test** Ho:$\mu_1-\mu_2=0$ Ha:$\mu_1-\mu_2<0$	Ho:$\mu_1-\mu_2=0$ Ha:$\mu_1-\mu_2\neq0$	-Independent Samples -Population Variances are NOT Equal -Sample means are both normally distributed	$T_{stat}=\frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{s_1^2/n_1+s_2^2/n_2}}$ df $=\frac{(s_1^2/n_1+s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1-1}+\frac{(s_2^2/n_2)^2}{n_2-1}}$	Reject Ho if $\lvert T_{stat}\rvert > \lvert T_{crit}\rvert$ or $p<\alpha$	$(\bar{x}_1-\bar{x}_2) \pm T_{\alpha/2}^{*}\sqrt{s_1^2/n_1+s_2^2/n_2}$	if ">" in Ha UB="infinite" LB=$(\bar{x}_1-\bar{x}_2)-T_\alpha^{*}\sqrt{s_1^2/n_1+s_2^2/n_2}$	if "<" in Ha UB=$(\bar{x}_1-\bar{x}_2)+T_\alpha^{*}\sqrt{s_1^2/n_1+s_2^2/n_2}$ LB="-infinite"
7	**Wilcoxin** Ho:$M_d=0$ Ha:$M_d<0$	Ho:$M_d=0$ Ha:$M_d\neq0$	Paired -Dependent Samples -Sample of Diff is NOT Normally Distributed	p-value = from Minitab output T^-(negative difference) T^+(add ranks positive) Tstat is smaller	Reject Ho if $p<\alpha$	$\alpha=0.05$ n=6 w=? (critical value) Minitab	Minitab	Minitab
8	**Mann-Whitney** Ho:$M_1-M_2=0$ Ha:$M_1-M_2<0$	Ho:$M_1-M_2=0$ Ha:$M_1-M_2\neq0$	-Independent Samples -One or both of the samples is NOT normally distributed	p-value = from Minitab output Tnew (Add ranks) $E(T_i)=\frac{n_i(n_1+n_2+1)}{2}$	Reject Ho if $p<\alpha$ $Var(T_i)=\frac{n_1 n_2(n_1+n_2+1)}{12}$	Minitab $SD(T_i)=\sqrt{Var(T_i)}$	Minitab $Z=\frac{T_{new}-E(T_i)}{SD(T_i)}$	Minitab P-Value = Z(≥ 72.34) = 0.5-(2.34) p-Value = 0.237
9	**1 Proportion** Ho:p=0.3 Ha:p<0.3	Ho:p=0.4 Ha:p≠0.4	-np ≥10 -nq ≥ 10 Sample Proportion normally distributed	$Z_{stat}=\frac{\hat{p}-p}{\sqrt{pq/n}}$ $\hat{p}=x/n$ $\hat{q}=1-\hat{p}$	Reject Ho if $\lvert Z_{stat}\rvert > \lvert Z_{crit}\rvert$ or $p<\alpha$	$\hat{p} \pm Z_{\alpha/2}\sqrt{\hat{p}\hat{q}/n}$ $n=\left(\frac{Z_{\alpha/2}}{ME}\right)^2 pq$	if ">" in Ha UB="1" LB=$\hat{p} - Z_\alpha\sqrt{\hat{p}\hat{q}/n}$	if "<" in Ha UB=$\hat{p} + Z_\alpha\sqrt{\hat{p}\hat{q}/n}$ LB="0"
10	**2 Proportions** Ho:$p_1-p_2=0.0$ Ha:$p_1-p_2<0.0$ (Independent Pooled)	Ho:$p_1-p_2=0.0$ Ha:$p_1-p_2\neq0.0$	Sample Proportions normally distributed	$Z_{stat}=\frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\bar{p}(1-\bar{p})(1/n_1+1/n_2)}}$ $\bar{p}=(n_1\hat{p}_1+n_2\hat{p}_2)/(n_1+n_2)$ $\hat{p}_1=x_1/n_1$ $\hat{p}_2=x_2/n_2$	Reject Ho if $\lvert Z_{stat}\rvert > \lvert Z_{crit}\rvert$ or $p<\alpha$	$(\hat{p}_1-\hat{p}_2) \pm Z_{\alpha/2}^{*}\sqrt{\hat{p}_1\hat{q}_1/n_1+\hat{p}_2\hat{q}_2/n_2}$	if ">" in Ha UB="1" LB=$(\hat{p}_1-\hat{p}_2)-Z_\alpha^{*}\sqrt{\hat{p}_1\hat{q}_1/n_1+\hat{p}_2\hat{q}_2/n_2}$	if "<" in Ha UB=$(\hat{p}_1-\hat{p}_2)+Z_\alpha^{*}\sqrt{\hat{p}_1\hat{q}_1/n_1+\hat{p}_2\hat{q}_2/n_2}$ LB="-1"
11	**2 Proportions** Ho:$p_1-p_2=0.1$ Ha:$p_1-p_2<0.1$ (Independent)	Ho:$p_1-p_2=0.1$ Ha:$p_1-p_2\neq0.1$	Sample Proportions normally distributed	$Z_{stat}=\frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\hat{p}_1\hat{q}_1/n_1+\hat{p}_2\hat{q}_2/n_2}}$ $\hat{p}_1=x_1/n_1$ $\hat{p}_2=x_2/n_2$	Reject Ho if $\lvert Z_{stat}\rvert > \lvert Z_{crit}\rvert$ or $p<\alpha$	$(\hat{p}_1-\hat{p}_2) \pm Z_{\alpha/2}^{*}\sqrt{\hat{p}_1\hat{q}_1/n_1+\hat{p}_2\hat{q}_2/n_2}$	if ">" in Ha UB="1" LB=$(\hat{p}_1-\hat{p}_2)-Z_\alpha^{*}\sqrt{\hat{p}_1\hat{q}_1/n_1+\hat{p}_2\hat{q}_2/n_2}$	if "<" in Ha UB=$(\hat{p}_1-\hat{p}_2)+Z_\alpha^{*}\sqrt{\hat{p}_1\hat{q}_1/n_1+\hat{p}_2\hat{q}_2/n_2}$ LB="-1"
12	**2 Proportions** Ho:$p_1-p_2=0.0$ Ha:$p_1-p_2<0.0$ (Dependent, $\hat{p}_1$ and $\hat{p}_2$ from same sample)	Ho:$p_1-p_2=0.0$ Ha:$p_1-p_2\neq0.0$	Sample Proportions normally distributed	$Z_{stat}=\frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\hat{p}_1\hat{q}_1/n+\hat{p}_2\hat{q}_2/n+2\hat{p}_{12}\hat{p}/n}}$ $\hat{p}_1=x_1/n$ $\hat{p}_2=x_2/n$	Reject Ho if $\lvert Z_{stat}\rvert > \lvert Z_{crit}\rvert$ or $p<\alpha$	$(\hat{p}_1-\hat{p}_2) \pm Z_{\alpha/2}^{*}\sqrt{\hat{p}_1\hat{q}_1/n+\hat{p}_2\hat{q}_2/n+2\hat{p}\hat{p}_{12}/n}$	if ">" in Ha UB="1" LB="do only subtraction" *Remember use 1 sided Z	if "<" in Ha UB="Do only addition" LB="-1"

Variables
E=margin of error
μ = pop. mean
$\bar{x}$ = sample mean
σ = pop. SD
s = sample SD
σ^2 = pop. variance
s^2 = sample variance
μ_d = pop. mean of differences
$\bar{d}$ = sample mean of differences
s_d = sample SD of differences
s_p = pooled SD
n = sample size
M_d = pop. median of differences
M = pop. median
p = pop. proportion
$\hat{p}$ = sample proportion
$\bar{p}$ = pooled proportion

Testing Steps
1) State Assumptions
2) Write Hypothesis
3) Calculate STAT Value
4) Calculate CRIT Value
5) Make Decision
6) Write Conclusion

18 **Kruskal-Wallace**
Ho:$M_1 = M_2 = M_3 ... = M_n$
Ha:not all M equal
-Get p-value from minitab and compare to alpha.
Reject Ho p < α

Correlation
-Measure the relationship between 2 variables
-Always between 1 and -1
-The closer to 1 or -1 the stronger the relationship

Ho:ρ=0 Ha:ρ≠0
Ho:ρ=0 Ha:ρ>0
$T_{stat}=\frac{r}{\sqrt{\frac{1-r^2}{n-1}}}$ df = n-2
- "r" is calculated by minitab
- Reject Ho if $\lvert T_{stat}\rvert > \lvert T_{crit}\rvert$ *see 1-tailed rules

***If 1 tailed and you have a "<" sign in Ha, reject Ho if $Z_{stat} < Z_{crit}$ or $T_{stat} < T_{crit}$. If 1 tailed and you have a ">" sign in Ha, reject Ho if $Z_{stat} > Z_{crit}$ or $T_{stat} > T_{crit}$**

---

12 **1-WAY ANOVA**
Assumptions
-Samples Normal
-Equal Variances
-Observations Ind

Ho:$\mu_1 = \mu_2 = \mu_3 = ... \mu_n$
Ha:not all means equal
***Review 1 WAY ANOVA table in the book for detailed calculations.

$F_{stat}=MS_{fac}/MS_{err}$
Reject Ho if $\lvert F_{stat}\rvert > \lvert F_{crit}\rvert$ or p < α.
Both F_{stat} and p come from minitab.

$df_f = k-1$
$df_e = N-k$
$df_t = N-1$
k=# rows

13 **2-WAY ANOVA**
Assumptions
-See 1-Way ANOVA
**Review 2 WAY ANOVA table in the book for detailed calculations.
3 Possible Tests

Test #1-Interaction
Ho: Factors A and B do not Interact
Ha:Factors A and B Interact
$F_{stat}=MS_{int}/MS_{err}$

Test #2-Factor A Means
Ho:$\mu_{1a} = \mu_{2a} = ... \mu_{na}$
Ha:not all means equal
$F_{stat}=MS_{facA}/MS_{err}$

Test #3-Factor B Means
Ho:$\mu_{1b} = \mu_{2b} = ... \mu_{nb}$
Ha:not all means equal
$F_{stat}=MS_{facB}/MS_{err}$

For all 3 tests Reject Ho if $\lvert F_{stat}\rvert > \lvert F_{crit}\rvert$ or p < α.
F-table
There will be 3 F-values and/or p-values in minitab.
$df_{facA} = r-1$ $df_{facB} = c-1$ $df_{int} = (r-1)*(c-1)$ $df_{err} = r*c*(n-1)$ $df_{tot} = N-1$

19 **ANOVA-BLOCKED**
Assumptions
-See 1-Way ANOVA
**Review BLOCKED ANOVA table in the book for detailed calculations.
2 Possible Tests

Test #1-Was Blocking Needed?
Ho:$\mu_{b1} = \mu_{b2} = ... \mu_{bn}$
Ha:not all block means equal
$F_{stat}=MS_{block}/MS_{err}$

Test #2-Factor Means
Ho:$\mu_{a1} = \mu_{a2} = ... \mu_{an}$
Ha:not all factor means equal
$F_{stat}=MS_{facA}/MS_{err}$

***If in test #1 we DO NOT reject Ho we must run a 1-WAY ANOVA to complete test #2, otherwise, you can proceed with test #2 using the output.

For both tests Reject Ho if $\lvert F_{stat}\rvert > \lvert F_{crit}\rvert$ or p < α.
There will be 2 F-values and/or p-values in minitab.
$df_{block} = b-1$ $df_{fac} = c-1$ $df_{err} = (b-1)*(c-1)$ $df_{tot} = N-1$ b = # of rows

Bonferroni used to identify where the actual differences between means are.
Step 1 - Determine number of rows (r), number of columns (c) and number of observations in each samples being compared (n)
Step 2 - Calculate "J" (the number of possible comparisons) J=[(rc)*(rc-1)]/2 s=SQRT(MSE)
Step 3 - Calculate alpha/2(J) and the df = dfe
Step 4 - Get T-value from table at a 1 tailed significance level of alpha/2(J) and dfe (pick closest value)
Step 5 - Interval = (x_bar1 - x_bar2) +/- $t_{\alpha/2J}$ * s *SQRT(1/n1 + 1/n2)
critical difference = margin of error = $t_{\alpha/2J}$ * s *SQRT(1/n1 + 1/n2)
Rule - if the interval contains "0", there is no difference between the means being compared
OR - if |actual diff| < |critical diff|, there is no difference between the means
-Used for 1-WAY ANOVA and 2-WAY ANOVA, do not use for BLOCKED ANOVA
-For 2-WAY ANOVA there are 3 types of possible comparisons, row comparisons, column comparisons and cell comparisons, all steps will be exactly the same except step 1. The values of "r", "c" "n" are determined based on the type of comparison.

15 **Multiple Regression** or Simple Regression
Assumptions
-Error Terms Norm Dist.
-Constant Variance
-Error Terms Ind.
-Linear Model

Test #1-Significance of Model
Ho:$B_1=B_2=B_3...=B_n=0$
Ha:Atleast one $B_n\neq0$
$F_{stat}=MS_{reg}/MS_{err}$
Reject Ho if $\lvert F_{stat}\rvert > \lvert F_{crit}\rvert$ or p < α.
F-table

Test #2-Significance of Variable in Model
Ho:$B_n=0$
Ha:$B_n\neq0$
$T_{stat}=\frac{b_n - B_n}{s_{bn}}$ df = dfe
Reject Ho if $\lvert T_{stat}\rvert > \lvert T_{crit}\rvert$ or p < α
-All numbers come from minitab.
-This test can be repeated for each independent variable.
CI (coefficient) = b_n +/- $T_{\alpha/2}$*s_{bn}

CI = fit +/- $T_{\alpha/2}$ *SE_{fit} (interval for average of all observations) df_{reg}= k df_{err}= n-k-1
PI = fit +/- $T_{\alpha/2}$ *SQRT(SE_{fit}^2 + S^2) (interval for 1 observation) df_{tot}= n-1
-Regression Equation Given in Output as $\hat{y}=c+B_1x_1+B_2x_2+....+B_nx_n$
-Residual= actual value - value predicted by model by subbing independent variables into the regression equation.
-The best model will be the one with the highest R^2 or R^2_{adj}, the lowest S and the lowest number of ind. variables.

Variance Inflation Factor (VIF)
VIF = 1/(1-R^2)
Multicollinearity is a problem caused by highly correlated independent variables. It is a problem if VIF>10.

ANOVA and REGRESSION
R^2 = 1-(SS_{err}/SS_{tot})
R^2_{adj}= 1-(MS_{err}/MS_{tot})
s=SQRT(MSE) MS = SS/df
SSe = SUM[$(n_i-1)*s_i^2$]

Manual Simple Regression Calculations:
Slope = $b_1 = r * (s_x / s_y)$
Intercept = $b_0 = \bar{y} - b_1\bar{x}$
Regression equation:
$\hat{y} = b_0 + b_1\bar{x}$
r = coefficient of correlation
s_x = standard dev. of x
s_y = standard dev. of y
$\bar{x}$ = mean of x values
$\bar{y}$ = mean of y values

spread of residual increasing (linear may not be appropriate)