STAT 641 Fall 2012
Solutions for ASSIGNMENT 9

(P1) (8 pts.)

(S1) Matched pairs - both algorithms are applied to the same problems

(S2) Independent samples - There are two independent random samples, one from inner city schools and the other from suburban schools.

(S3) Matched pairs - Each of the 250 were observed under both stimuli.

(S4) Matched pairs - The two viruses were applied to the same leaf, one half to Virus 1 and the other half to Virus 2.

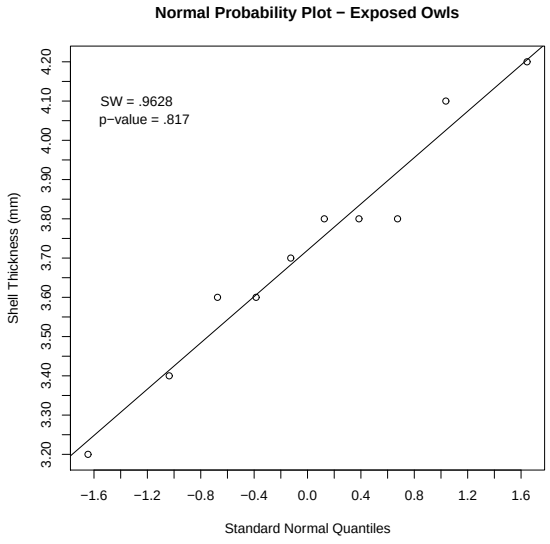
(P2) (12 pts.)

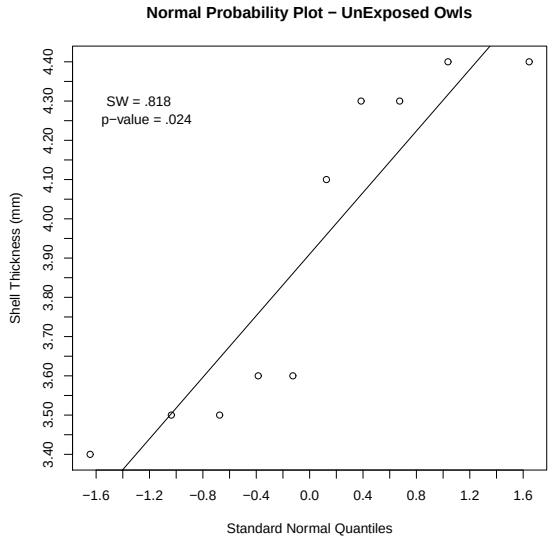
(A) Test $H_o : \mu_{exp} \geq \mu_{unexp}$ vs $H_o : \mu_{exp} < \mu_{unexp}$. A separate variance t-test will be implemented because the p-value from the BFL test is .0265 (see part (E) for details) implies there is sufficient evidence to conclude that there is a difference in the two population's variances. From the data, the $C = \frac{(.2974)^2/10}{(.4228)^2/10} = .4948$ and $df = \frac{(.4948+1)^2(10-1)(10-1)}{(.4948)^2(10-1)+(10-1)} = 16.154$

$$t = \frac{3.72 - 3.91}{\sqrt{(.2974)^2/10 + (.4228)^2/10}} = -1.1623 > -1.745 = -t_{.05,16.154} \Rightarrow$$

$$\text{p-value } = P[t_{16.154} \leq -1.1623] = pt(-1.1623, 16.154) = .131$$

with 95% Upper Bound $(-\infty,\ .095)$ $\Rightarrow$ Fail to reject H_o and conclude that there is not significant evidence that the average shell thickness for eggs from PCB exposed owls is less than for unexposed owls. For the t-test to be valid both the exposed and unexposed shell thickness populations would need to have normal distributions. From the exposed data we have that the Shapiro-Wilk's test has p-value=.817 and the unexposed data has p-value =.024 . Thus, we would conclude that the data indicates that the exposed shell thickness have a normal distribution but the unexposed shell thickness do not have a normal distribution. This is confirmed by the following normal reference distribution plots:





Therefore, the p-value for the t-test may not be valid.

(B) Because the distribution for the Unexposed shell thicknesses was not a normal distribution, the following calculations are not very reliable:

Let $\Delta = \frac{\mu_{exp}-\mu_{unexp}}{\sigma\sqrt{1/10+1/10}} = \frac{k\sigma}{\sigma\sqrt{2/10}} = k\sqrt{5}$, where k= 0, -.5, -1, -1.5, -2

$$\begin{aligned} P[\text{Type II error at } \mu_{exp}-\mu_{unexp}=k\sigma] &= P[\text{Fail to Reject } H_o \text{ at } \mu_{exp}-\mu_{unexp}=k\sigma] \\ &= P[t_{18,\Delta} > -1.734] \\ &= 1 - pt(-1.734, 18, \Delta] \end{aligned}$$

$\mu_{exp}-\mu_{unexp}$	0	$-.5\sigma$	-1.0σ	-1.5σ	-2.0σ
Δ	0	$-.5\sqrt{5}$	$-1.0\sqrt{5}$	$-1.5\sqrt{5}$	$-2.0\sqrt{5}$
P[Type II Error]	0	.7152	.3064	.0571	.0040

(C) Using the normal based procedure with an estimated common variance of $\hat{\sigma}^2 = \frac{(10-1)(.2974)^2+(10-1)(.4228)^2}{20-1} = (.3655)^2$, even though it is of questionable validity due to the conclusion of unequal variance and lack of normality,

$$m = \left(\frac{3+1}{3}\right)\frac{(\hat{\sigma})^2(Z_\alpha+Z_\beta)^2}{(\delta)^2} = \left(\frac{3+1}{3}\right)\frac{(.3655)^2(1.645+.84)^2}{(.3)^2} = 12.22 \Rightarrow m = 13 \text{ and } n = 3m = 39.$$

(D) Using the Wilcoxon Rank Sum test (even though the distributions of the exposed and unexposed egg thicknesses are not in the same family of distributions), we obtain the following ranks:

Exposed	3.6	3.2	3.8	3.6	4.1	3.8	4.2	3.4	3.7	3.8
Ranks	7.5	1.0	12.0	7.5	14.5	12.0	16.0	2.5	10.0	12.0
UnExposed:	4.3	4.4	3.6	3.5	4.4	3.5	3.4	3.6	4.1	4.3
Ranks	17.5	19.5	7.5	4.5	19.5	4.5	2.5	7.5	14.5	17.5

The sum of the ranks for the Exposed owls is $W_1 = 95$, with

$p - value = pwilcox(W_1 - (10)(11)/2, 10, 10) = .2406$ which would imply that there is not sufficient evidence to conclude that the egg thickness for the exposed owls is shifted to the left of the egg thicknesses for the Unexposed owls.

Using the R-function: **wilcox.test(ex,uex,alternative="l",paired=FALSE)** we obtain

```
        Wilcoxon rank sum test with continuity correction
data:  ex and uex
W = 40, p-value = 0.2347
alternative hypothesis: true location shift is less than 0
Warning message:
In wilcox.test.default(ex, uex, alternative = "l", paired = FALSE) :
  cannot compute exact p-value with ties
```

(E) There is not a valid test for testing $H_o : \sigma_{exp} \leq \sigma_{unexp}$ versus $H_1 : \sigma_{exp} > \sigma_{unexp}$ because the distribution of the shell thicknesses from the Unexposed owls have a non-normal distribution. We can test $H_o : \sigma_{exp} = \sigma_{unexp}$ versus $H_1 : \sigma_{exp} \neq \sigma_{unexp}$ using the Brown-Forsythe-Levene test (SAS) which yields a p-value of .0265 which would indicate there is sufficient evidence to conclude there is a difference in the variability of the two populations of egg thicknesses.

(F) Based on the lack of normality in the data for the unexposed egg thicknesses, the t-test would not appear to be a valid procedure, especially considering the small sample sizes. Therefore, I would be more confident in using the results from the Wilcoxon Rank Sum test, although, the conditions for using this test are not valid either. However, the Wilcoxon Rank Sum test is much more robust against deviations from its conditions than is the t-test.

(P3) (8 pts.) Let $\tilde{\mu}_1$ and $\tilde{\mu}_2$ be the medians of Vitamin B and Placebo groups, respectively.

Test $H_0 : \tilde{\mu}_1 \leq \tilde{\mu}_2$ vs $H_1 : \tilde{\mu}_1 > \tilde{\mu}_2$.

Let $D = VitB - Placebo$. The values of D and the ranks of their magnitudes $|D_i|$ are given below:

Pair	1	2	3	4	5	6	7	8	9	10	11	12
$Y = X_1 - X_2$	6	8	9	5	-7	5	3	3	-12	3	0	1
Rank	7	9	10	5.5	8	5.5	3	3	11	3	*	1

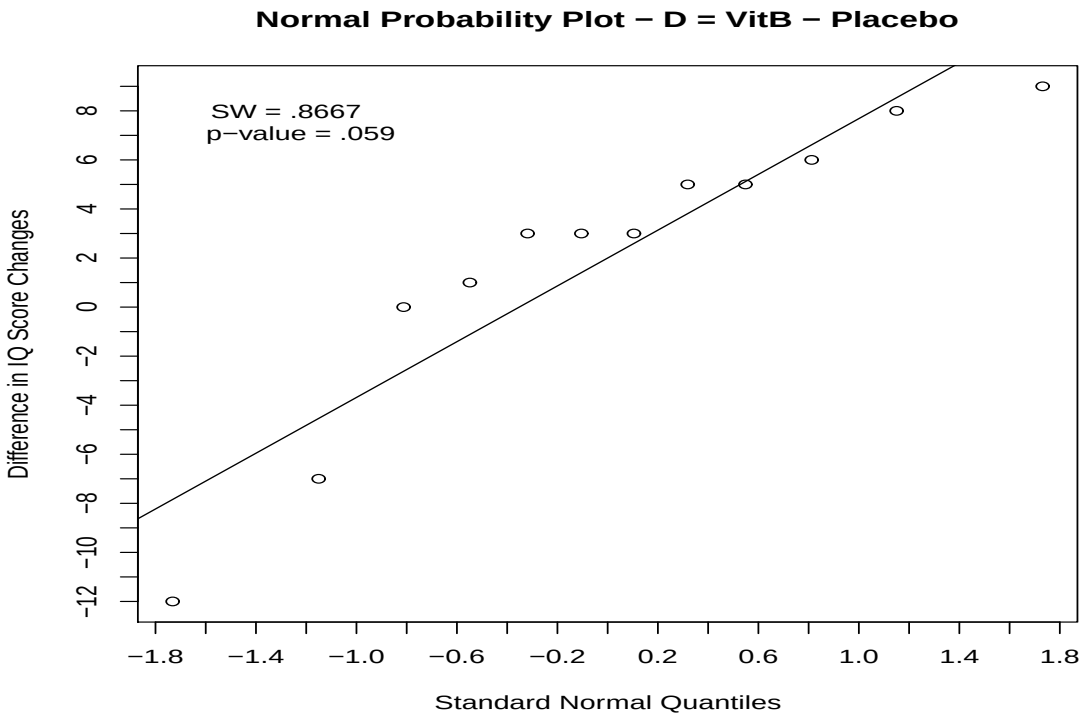
- **t-test:** $t = \frac{\bar{D}}{S_D/\sqrt{12}} = \frac{2}{6.0603/\sqrt{12}} = 1.143$ with $p-value = P[t_{12} \geq 1.143] = 1 - pt(1.143, 12) = .138$. Therefore, fail to reject H_o and conclude there is insufficient evidence that the median change in IQ score for the Vitamin B group is larger than the median change in IQ score for the Placebo group.

- **Signed Rank Test:** Let W_+ be sum of the ranks of positive differences.
  Note that the sample size is reduced to $n^* = 12 - 1$ due to one of the pairs having D=0.
  $W_+ = 47$ and $p - value = P[W_+ \geq 47] > P[W_+ \geq 48] = 0.103$ from Table A.10 in Textbook
  $p - value = P[W_+ \geq 47] = psignrank(46, 11, FALSE) = 0.120$ using R-function
  An approximate p–value is obtained using the Central Limit Theorem for W_+:
  $Z = \frac{W_+ - 0.5 - 33}{\sqrt{(11)(23)/2}} \sim N(0,1)$ approximately for large n

  $$p-value = P[W_+ \geq 47] \approx P\left[Z \geq \frac{47 - 0.5 - 33}{\sqrt{(11)(23)/2}}\right] = 1 - pnorm\left(\frac{47 - 0.5 - 33}{\sqrt{(11)(23)/2}}\right) = 1 - pnorm(1.20) = 0.115$$

  The following R code can be used to obtain the same results as above:

```
x = c(14,26,2,4,-5,14,3,-1,1,6,3,4)
y = c(8,18,-7,-1,2,9,0,-4,13,3,3,3)

t.test(x,y,alternative=c("greater"),paired=TRUE)

wilcox.test(x,y,alternative=c("greater"),paired=TRUE)

OUTPUT From R:
        Paired t-test

data:  V and P
t = 1.1432, df = 11, p-value = 0.1386
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -1.141826       Inf
sample estimates:
mean of the differences 2


OUTPUT From R:
       Wilcoxon signed rank test with continuity correction

data:  x and y
V = 47, p-value = 0.1144
alternative hypothesis: true location shift is greater than 0
```



- The Wilcoxon Signed Rank test provides a more reliable p-value than the paired t-test because the t-test requires that the differences have a normal distribution. From the normal reference distribution plot of the 12 differences, it would appear that a normal distribution is not a very good fit and the Shapiro-Wilk's test has p-value .059.

(P4) (10 pts.) Let p_1 be the probability that a Normal patient will have Low Excretions and p_2 be the probability that a Diabetic patient will have Low Excretions. We want to test $H_0 : p_1 = p_2$ vs $H_1 : p_1 \neq p_2$.

(A) The $E'_{ij}s$ would be given by

$\hat{E}_{11} = (14)(12)/24 = 7,\ \hat{E}_{21} = (14)(12)/24 = 7,\ \hat{E}_{12} = (10)(12)/24 = 5,\ \hat{E}_{22} = (10)(12)/24 = 5$

We could use the Chi-square test because $E_{ij} \geq 5$ for all (i, j).

$$\chi^2 = \sum_{i=1}^{2}\sum_{j=1}^{2}(O_{ij} - \hat{E}_{ij})^2/\hat{E}_{ij} = \frac{(10-7)^2}{7} + \frac{(4-7)^2}{7} + \frac{(2-5)^2}{5} + \frac{(8-5)^2}{5} = 6.171$$

With the two of the four expected values at 5, the Fisher Exact test will also be calculated:

$$p(x_0) = \frac{\binom{12}{x_0}\binom{12}{14-x_0}}{\binom{24}{14}} \quad \Rightarrow \quad p(10) = \frac{\binom{12}{10}\binom{12}{14-10}}{\binom{24}{14}} = dhyper(10, 12, 12, 14) = 0.0167.$$

(B) For the Pearson Chi-squared Test: $p - value = P[\chi^2_1 > 6.171] = 1 = pchisq(6.171, 1) = .013$

For the Fisher Exact Test: p-value=$\sum_{x\in\Omega_0} p(x) = 0.0361$,

where $\Omega_0 = \{x : p(x) \leq p(x_0)\}$ and $p(x) = \frac{\binom{12}{x}\binom{12}{14-x}}{\binom{24}{14}} = dhyper(x, 12, 12, 14)$ for $2 \leq x \leq 12$.

x	2	3	4	5	6	7	8	9	10	11	12
p(x)	0.00003	0.00135	0.01666	0.08884	0.23321	0.31983	0.23321	0.08884	0.01666	0.00135	0.00003

Because of the small values of expected counts and the discrepancies between the two p-values, we will use the values from the Fisher Exact Test:

p-value $= .00003 + .00135 + .01666 + .01666 + .00135 + .00003 = 0.03608 < 0.05$,

Therefore, we reject H_0 and conclude that there is significant evidence (p-value=.036) of a difference in the urinary thromboglobulin excretion between normal and diabetic persons.

The following SAS output confirms our calculations:

```
The SAS System                                                          1

The FREQ Procedure

Table of DISEASE by EXCRETION

DISEASE         EXCRETION

Frequency      |
Expected       |
Cell Chi-Square|
Percent        |
Row Pct        |
Col Pct        |H       |L       |  Total
---------------------------------------
D              |      8 |      4 |     12
               |      5 |      7 |
               |    1.8 | 1.2857 |
               |  33.33 |  16.67 |  50.00
               |  66.67 |  33.33 |
               |  80.00 |  28.57 |
---------------------------------------
N              |      2 |     10 |     12
               |      5 |      7 |
               |    1.8 | 1.2857 |
               |   8.33 |  41.67 |  50.00
               |  16.67 |  83.33 |
               |  20.00 |  71.43 |
---------------------------------------
Total               10       14       24
                 41.67    58.33   100.00


Statistics for Table of DISEASE by EXCRETION

Statistic                     DF       Value      Prob
------------------------------------------------------------
Chi-Square                     1      6.1714    0.0130
Likelihood Ratio Chi-Square    1      6.5115    0.0107
Continuity Adj. Chi-Square     1      4.2857    0.0384
Mantel-Haenszel Chi-Square     1      5.9143    0.0150
Phi Coefficient                       0.5071
Contingency Coefficient               0.4523
Cramer's V                            0.5071


     Pearson Chi-Square Test
----------------------------------
Chi-Square                   6.1714
DF                                1
Asymptotic Pr >  ChiSq       0.0130
Exact      Pr >= ChiSq       0.0361
```

(P5) 10 pts.

(A) $H_o : p_{1\cdot} = p_{\cdot 1}$ vs $H_a : p_{1\cdot} \neq p_{\cdot 1}$,

where $p_{1\cdot}, p_{\cdot 1}$ are the probabilities that drug 1 and drug 2 remained anesthetized, respectively.

Because the data consists of the two responses from the same individual, the McNemar's test statistic is appropriate.

The Pearson chi-square or Fisher Exact test would be inappropriate due to the correlation in the two responses.

(B) $p - value = 2min(P[B \leq n_{12}], P[B \geq n_{12}]) = 2min(P[B \leq 13], P[B \geq 13]) = 2min(.9979, .0106) = .0212$,

where B has a $Bin(m, .5) = Bin(3 + 13, .5) = Bin(16, .5)$ distribution.

There is significant evidence (p-value=.0212) of a difference in the two drugs' probabilities that the membrane remained anesthetized.

(P6) 10 pts. Let p_1 = proportion of offspring that are Tall, cut-leaf,

p_2 = proportion of offspring that are Dwarf, cut-leaf,

p_3 = proportion of offspring that are Tall, potato-leaf,

p_4 = proportion of offspring that are Dwarf, potato-leaf.

Test the hypotheses

$H_0 : p_1 = \frac{9}{16}, \;\; p_2 = \frac{3}{16}, \;\; p_3 = \frac{3}{16}, \;\; p_4 = \frac{1}{16}$ vs H_1 : at least one of the p_i's differs from its specified value.

This is a multinomial distribution so we can use the chi-square goodness of fit test from Handout 9 to test the hypotheses as we did for testing the fit of a discrete distribution.

$\widehat{E}_1 = (1611)\left(\frac{9}{16}\right) = 906.1875, \;\; \widehat{E}_2 = \widehat{E}_3 = (1611)\left(\frac{3}{16}\right) = 302.0625, \;\; \widehat{E}_4 = (1611)\left(\frac{1}{16}\right) = 100.6875.$

i. T.S: $\chi^2 = \sum_{i=1}^{4} \frac{(O_i - \widehat{E}_i)^2}{\widehat{E}_i} = \frac{(926-906.1875)^2}{906.1875} + \frac{(293-302.0625)^2}{302.0625} + \frac{(288-302.0625)^2}{302.0625} + \frac{(104-100.6875)^2}{100.6875} = 1.4687$ with df=4-1=3

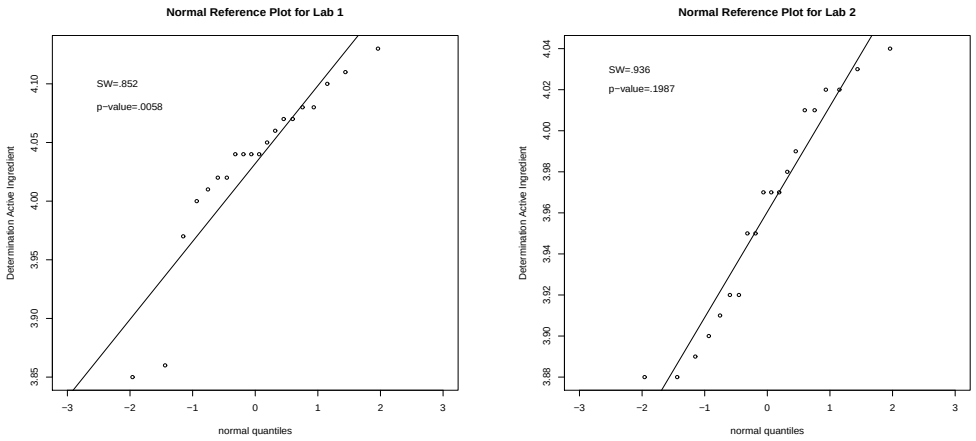
ii. p-value=$Pr(\chi^2_3 \geq 1.4687) = 1 - pchisq(1.4687, 3) = 0.6895 > 0.05 \quad \Rightarrow$ fail to reject H_0.

We conclude that there is not significant evidence (p-value=.6895) in the data that the tomato plants deviated from the current theory.

(P7) 20 pts.

(A) Lab 1: plotted points are not close to the line and p-value=0.0058 from Shapiro-Wilk which implies that the normal distribution provides a poor fit to the data.

Lab 2: plotted points are close to the line and p-value=0.1987 from Shapiro-Wilk which implies that the normal distribution provides a good fit to the data



(B) $H_0 : \sigma_1^2 = \sigma_2^2$ vs $H_1 : \sigma_1^2 \neq \sigma_2^2$.

From the data we have $\hat{\sigma}_1^2 = .005133; \quad \hat{\sigma}_2^2 = .002752$

From part (A) we have Lab 1 data has a non-normal distribution and therefore use BFL - test.

The BFL test yields p-value=0.799 and hence we would fail to reject H_o.

Thus, we conclude that there is not significant evidence that the two labs have different levels of variability.

- Note that the positive correlation for the data from Lab 2 found in part C, would somewhat invalidate this conclusion.

(C) Test the hypotheses: H_0 : Data not correlated vs H_1: Data is correlated

i. Because the data for Lab 1 was determined to be non-normally distributed, a Runs test will be used.

Runs test for Lab 1:
T.S: $r = \#$ of runs $= 9$.
$n_1 = 13,\ n_2 = 7 \quad \Rightarrow \quad$ From Table in Handout 13: $r_L = 5,\ r_U = 15$.
Because $r_L < r < r_U$, fail to reject H_0.

ii. Because there is a strong indication that the data is normally distributed, the von Neumann test will be used to evaluate correlation.

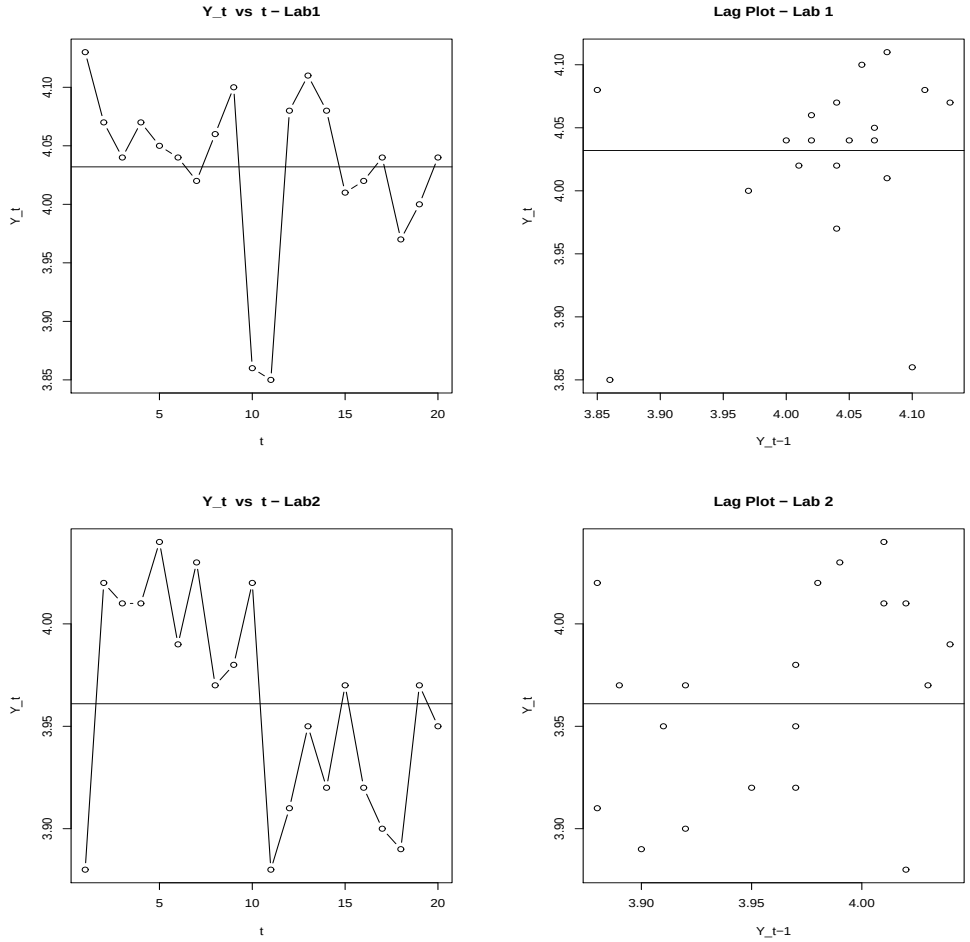
Reject H_o : Data is not correlated if $Q < Q_{P,.05} = 1.368$. From the data $\hat{\rho} = .3126$ and $Q = 1.314 < 1.368$. Conclude that there is sufficient evidence of correlation in the 20 observations from Lab 2.

Alternatively: Runs test for Lab 2:
$r = \#$ of runs $= 7$.
$n_1 = 11,\ n_2 = 9 \quad \Rightarrow \quad$ From Table in Handout 13: $r_L = 6,\ r_U = 16$.
Because $r_L < r < r_U$, fail to reject H_0.

When the data is from a normal distribution, the von Neumann test is more powerful that the Runs test and hence less likely to commit a Type II error.

Thus we conclude that there is not significant evidence that the daily determinations within Lab 1 are correlated but there is positive correlation in the Lab 2 data.

The following graphs display the plot of the data along with a lag plot. In the lag plot for Lab 1 there is no distinct pattern but a somewhat positive trend appears in the plot for Lab 2.



(D) From part (A) we have that Lab 1 data is non-normal distributed and n_1 is relatively small. So, use Wilcoxon Rank sum test.

i. Let $\tilde{\mu}_1$ and $\tilde{\mu}_2$ be the location parameters of the distributions for Lab 1 and Lab 2, respectively. Test: $H_0 : \tilde{\mu}_1 = \tilde{\mu}_2$ vs $H_1 : \tilde{\mu}_1 \neq \tilde{\mu}_2$.

ii. T.S: Let W_1 and W_2 be sums of ranks from Lab 1 data and Lab 2 data, respectively, $W_1 = 541.5$ and $W_2 = 278.5$.

iii. p-value=$2Pr(W_1 \geq w_{\max}) = 2Pr(W_1 \geq 541.5) = 2(1 - pwilcox(541 - 210, 20, 20) = .00020 < \alpha = 0.05 \quad \Rightarrow$ reject H_0.

We conclude there is significant evidence that the two labs have the different average determinations.

- Note that the positive correlation for the data from Lab 2 would somewhat invalidate this conclusion.

(E) Because the data from Lab 1 appears non-normal and the sample size, $n_1 = 20$ is small, confidence intervals will be placed on the medians instead of the means:

From Table VII.3 on page 32 in Handout 11, we have that with k=6, a 95.9% C.I. on the median determination is $(X_{(6)},\ X_{(15)})$

For Lab 1, a 95.9% C.I. on the median determination is (4.02, 4.07)

For Lab 2, there was strong evidence (p-value $\approx$ .2) that the data was from a normal distribution with positive correlation, $\hat{\rho} = .3126$, therefore an approximate 95% C.I. on the median determination is given by

$$\bar{Y} \pm t_{.025,19}\widehat{SE}(\bar{Y}) = \bar{Y} \pm (2.093)\frac{S}{\sqrt{n}}\sqrt{\frac{1+\hat{\rho}}{1-\hat{\rho}}} = 3.9605 \pm (2.093)\frac{.05246}{\sqrt{20}}\sqrt{\frac{1+.3126}{1-.3126}} = 3.9605 \pm .0339 = (3.93,\ 3.99)$$

The distribution-free 95.9% C.I. on the median determination for Lab 2 is $(X_{(6)},\ X_{(15)}) = (3.92,\ 4.01)$

(P8) 10 pts. Let p_{ij} be the probability that a randomly selected child has the ith level of Tonsil Size and jth level of Carrier Status.

Test for Independence between Tonsil Size and Carrier Status:

$H_0 : p_{ij} = p_{i\cdot}p_{\cdot j}$ for all pairs (i, j) vs $H_1 : p_{ij} \neq p_{i\cdot}p_{\cdot j}$ for some (i, j)

$$\widehat{E}_{11} = \frac{516)(72)}{1398} = 26.58, \quad \widehat{E}_{12} = \frac{516)(1326)}{1398} = 489.42, \quad \widehat{E}_{21} = \frac{589)(72)}{1398} = 30.33,$$
$$\widehat{E}_{22} = \frac{589)(1326)}{1398} = 558.67, \quad \widehat{E}_{31} = \frac{293)(72)}{1398} = 15.09, \quad \widehat{E}_{32} = \frac{293)(1326)}{1398} = 277.91$$

$\chi^2 = \sum_{i=1}^{3}\sum_{j=1}^{2}\frac{(O_{ij}-\widehat{E}_{ij})^2}{\widehat{E}_{ij}} = 7.885$ with $df = (3-1)(2-1) = 2 \quad \Rightarrow$

p-value=$Pr(\chi^2_2 \geq 7.885) = 1 - pchisq(7.885, 2) = 0.0194 < 0.05 \quad \Rightarrow$ reject H_0.

We thus conclude that there is significant evidence (p-value=.0194) that Tonsil Size and Carrier Status are associated.

The following SAS output will confirm the above calculations:

```
    Table of TONSIL by CARRIER

TONSIL          CARRIER

Frequency      |
Expected       |
Cell Chi-Square|
Percent        |
Row Pct        |
Col Pct        |C       |NC      |  Total
-----------------------------------------------
L              |     29 |    560 |    589
               | 30.335 | 558.67 |
               | 0.0587 | 0.0032 |
               |   2.07 |  40.06 |  42.13
               |   4.92 |  95.08 |
               |  40.28 |  42.23 |
-----------------------------------------------
N              |     19 |    497 |    516
               | 26.575 | 489.42 |
               | 2.1592 | 0.1172 |
               |   1.36 |  35.55 |  36.91
               |   3.68 |  96.32 |
               |  26.39 |  37.48 |
-----------------------------------------------
VL             |     24 |    269 |    293
               |  15.09 | 277.91 |
               | 5.2608 | 0.2857 |
               |   1.72 |  19.24 |  20.96
               |   8.19 |  91.81 |
               |  33.33 |  20.29 |
-----------------------------------------------
Total          |   72   |  1326  |   1398
               |  5.15  | 94.85  | 100.00


Statistics for Table of TONSIL by CARRIER

Statistic                     DF       Value      Prob
------------------------------------------------------------
Chi-Square                     2      7.8848    0.0194
Likelihood Ratio Chi-Square    2      7.3209    0.0257
Sample Size = 1398
```

(P9) 12 pts.

(A) Table of Defendant's Race vs Death Penalty by ignoring the Victim's Race:

Defendant's Race	Death Penalty Yes	No	Row Total
White	19	141	160
Black	17	149	166
Column Total	36	290	326

Test $H_0 : p_1 = p_2$ vs $H_1 : p_1 \neq p_2$,

where p_1 and p_2 denote the proportions of White and Black convicts who received the death penalty, respectively.

Then $\hat{p}_1 = \frac{19}{160} = 0.119$, $\hat{p}_2 = \frac{17}{166} = 0.102$ and $\hat{p}_C = \frac{19+17}{160+166} = 0.110$.

Under H_0, the test statistic is

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_C(1-\hat{p}_C)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{0.119 - 0.102}{\sqrt{(0.11)(0.89)\left(\frac{1}{160}+\frac{1}{166}\right)}} = 0.4706$$

$p - value = 2P(Z \geq 0.4706) = 0.6379 > .05 = \alpha$. Thus fail to reject H_0 at level $\alpha = 0.05$ and conclude that there is not significant evidence that of a difference between the proportions of White and Black convicts who received the death penalty.

Note that you can use the chi-square test:

The $E'_{ij}s$ would be given by

$\hat{E}_{11} = (160)(36)/326 = 17.669$, $\hat{E}_{21} = (166)(36)/326 = 18.331$, $\hat{E}_{12} = (160)(290)/326 = 142.33$, $\hat{E}_{22} = (166)(290)/326 = 147.67$

We could use the Chi-square test because $E_{ij} \geq 5$ for all (i, j).

$$\chi^2 = \sum_{i=1}^{2}\sum_{j=1}^{2}(O_{ij} - \hat{E}_{ij})^2/\hat{E}_{ij} = \frac{(19-17.669)^2}{17.669} + \frac{(17-18.331)^2}{18.331} + \frac{(141-142.33)^2}{142.33} + \frac{(149-147.67)^2}{147.67} = 0.2214$$

$p - value = P[\chi^2_1 \geq .2214] = 1 - pchisq(.2214, 1) = .6379$

(B) (i) For White victims we have the following results:

Defendant's Race	Death Penalty Yes	No	Row Total
White	19	132	151
Black	11	52	63
Column Total	30	184	214

$\hat{p}_{1,W} = \frac{19}{151} = 0.126$, $\hat{p}_{2,W} = \frac{11}{63} = 0.174$ and $\hat{p}_{C,W} = \frac{19+11}{151+63} = 0.14$.

Under $H_0 : p_{1,W} = p_{2,W}$, the test statistic is

$$z_W = \frac{\hat{p}_{1,W} - \hat{p}_{2,W}}{\sqrt{\hat{p}_{C,W}(1-\hat{p}_{C,W})\left(\frac{1}{n_{1,W}}+\frac{1}{n_{2,W}}\right)}} = \frac{0.126 - 0.174}{\sqrt{(0.14)(0.86)\left(\frac{1}{151}+\frac{1}{63}\right)}} = -0.9370$$

$p - value = 2P(Z \geq 0.9370) = 0.3487$. Thus fail to reject H_0 at level $\alpha = 0.05$ and conclude that there is not significant evidence of a difference in the probability of receiving the death penalty between white and black defendant when the victim is white.

(ii) For Black victims we have the following results:

Defendant's Race	Death Penalty Yes	No	Row Total
White	0	9	9
Black	6	97	103
Column Total	6	106	112

$\hat{p}_{1,W} = \frac{0}{9} = 0$, $\hat{p}_{2,W} = \frac{6}{97} = 0.062$ and $\hat{p}_{C,W} = \frac{19+11}{151+63} = 0.14$.

The expected count in the (1,1) cell is $(9)(6)/112 = .48$ which is less than 1 hence you should not use the Z-test or chi-square test.

From Fisher's Exact test, $P[X = 0] = .5977 > P[X = x]$ for $2 \leq x \leq 6 \Rightarrow$

$p - value = \sum_{x=0}^{6} dhyper(x, 9, 103, 6) = 1$.

Thus fail to reject H_0 at level $\alpha = 0.05$ and conclude that there is not significant evidence of a difference in the probability of receiving the death penalty between white and black defendant when the victim is black.

(C) No. There is not significant evidence of a racial difference in the death penalty for white and black victims.

Using the following SAS code and output, the CHM test can be conducted to test for an association between race of defendant and penalty after controlling for race of the victim.

```
OPTIONS PS=55 LS=75 NOCENTER NODATE;
DATA Dia;
INPUT DEF_RACE $ VIC_RACE $ DEATH $ CNT @@;
CARDS;
W W Y 19
W W N 132
W B Y 0
W B N 9
B W Y 11
B W N 52
B B Y 6
B B N 97

RUN;
PROC FREQ ORDER=DATA; TABLES DEF_RACE*DEATH/EXPECTED CELLCHI2 CHISQ;
WEIGHT CNT;
EXACT CHISQ;
RUN;
PROC SORT; BY VIC_RACE;
PROC FREQ ORDER=DATA; BY VIC_RACE;
TABLES DEF_RACE*DEATH/EXPECTED CELLCHI2 CHISQ;
WEIGHT CNT;
EXACT CHISQ;
RUN;
PROC FREQ ORDER=DATA;
TABLES VIC_RACE*DEF_RACE*DEATH/CMH;
WEIGHT CNT;
RUN;

SAS OUTPUT:
Table of DEF_RACE by DEATH
DEF_RACE        DEATH

Frequency      |
Expected       |
Cell Chi-Square|
Percent        |
Row Pct        |
Col Pct        |Y       |N       |  Total
-----------------------------------------
W              |     19 |    141 |    160
               | 17.669 | 142.33 |
               | 0.1003 | 0.0125 |
               |   5.83 |  43.25 |  49.08
               |  11.88 |  88.13 |
               |  52.78 |  48.62 |
-----------------------------------------
B              |     17 |    149 |    166
               | 18.331 | 147.67 |
               | 0.0967 |  0.012 |
               |   5.21 |  45.71 |  50.92
               |  10.24 |  89.76 |
               |  47.22 |  51.38 |
-----------------------------------------
Total          |    36  |   290  |     326
               |  11.04 |  88.96 |   100.00
```

```
Statistics for Table of DEF_RACE by DEATH
Statistic                     DF       Value      Prob
------------------------------------------------------------
Chi-Square                     1      0.2214    0.6379
Likelihood Ratio Chi-Square    1      0.2215    0.6379

       Fisher's Exact Test
----------------------------------------
Cell (1,1) Frequency (F)        19
Left-sided Pr <= F          0.7412
Right-sided Pr >= F         0.3843

Table Probability (P)       0.1255
Two-sided Pr <= P           0.7246

Sample Size = 326

For VICTIM RACE BLACK
The FREQ Procedure
Table of DEF_RACE by DEATH

DEF_RACE        DEATH

Frequency      |
Expected       |
Cell Chi-Square|
Percent        |
Row Pct        |
Col Pct        |N       |Y       |  Total
-----------------------------------------
W              |      9 |      0 |      9
               | 8.5179 | 0.4821 |
               | 0.0273 | 0.4821 |
               |   8.04 |   0.00 |   8.04
               | 100.00 |   0.00 |
               |   8.49 |   0.00 |
-----------------------------------------
B              |     97 |      6 |    103
               | 97.482 | 5.5179 |
               | 0.0024 | 0.0421 |
               |  86.61 |   5.36 |  91.96
               |  94.17 |   5.83 |
               |  91.51 | 100.00 |
-----------------------------------------
Total          |  106   |   6    |  112
               | 94.64  | 5.36   | 100.00


Statistics for Table of DEF_RACE by DEATH

Statistic                     DF       Value      Prob
------------------------------------------------------------
Chi-Square                     1      0.5539    0.4567
Likelihood Ratio Chi-Square    1      1.0344    0.3091

WARNING: 25% of the cells have expected counts less than 5.
         (Asymptotic) Chi-Square may not be a valid test.

       Fisher's Exact Test
----------------------------------------
Cell (1,1) Frequency (F)         9
Left-sided Pr <= F          1.0000
Right-sided Pr >= F         0.5977

Table Probability (P)       0.5977
Two-sided Pr <= P           1.0000

Sample Size = 112
```

For VICTIM RACE WHITE

The FREQ Procedure

Table of DEF_RACE by DEATH

DEF_RACE	DEATH		
Frequency Expected Cell Chi-Square Percent Row Pct Col Pct	Y	N	Total
W	19 21.168 0.2221 8.88 12.58 63.33	132 129.83 0.0362 61.68 87.42 71.74	151 70.56
B	11 8.8318 0.5323 5.14 17.46 36.67	52 54.168 0.0868 24.30 82.54 28.26	63 29.44
Total	30 14.02	184 85.98	214 100.00

Statistics for Table of DEF_RACE by DEATH

Statistic	DF	Value	Prob
Chi-Square	1	0.8774	0.3489
Likelihood Ratio Chi-Square	1	0.8475	0.3573

Fisher's Exact Test

Cell (1,1) Frequency (F)	19
Left-sided Pr <= F	0.2325
Right-sided Pr >= F	0.8744
Table Probability (P)	0.1070
Two-sided Pr <= P	0.3893

Sample Size = 214

```
Summary Statistics for DEFENDANT RACE by DEATH
Controlling for VICTIM RACE

  Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic    Alternative Hypothesis    DF       Value      Prob
------------------------------------------------------------------
    1        Nonzero Correlation        1      1.2097    0.2714
    2        Row Mean Scores Differ     1      1.2097    0.2714
    3        General Association        1      1.2097    0.2714


           Estimates of the Common Relative Risk (Row1/Row2)

Type of Study     Method                  Value     95% Confidence Limits
-------------------------------------------------------------------------
Case-Control      Mantel-Haenszel        1.5741       0.7096       3.4915
  (Odds Ratio)    Logit **               1.4545       0.6666       3.1735


     Breslow-Day Test for
Homogeneity of the Odds Ratios
--------------------------------
Chi-Square              0.3806
DF                           1
Pr > ChiSq              0.5373


Total Sample Size = 326
```

The Breslow-Day test fails to reject the hypothesis that the odds-ratios for the death penalty for Black and White Defendants are different for Black and White Victims. The estimate of the common odds-ratio is 1.57 with a 95% C.I. of (.71, 3.49). Because the C.I. contains 1 we can not reject the hypothesis of the odds of receiving the death penalty are different for Black and White Defendants. Note also that the CMH test has a p-value = .2714 from which we can conclude that after controlling for the Race of the Victim, there is not significant evidence of an association between the likelihood of receiving the death penalty and the race of the defendant.

(D) In this situation, the conclusions remain the same in all three analyses of the data.