Stat 403 - Solution to Assignment 3

Turned in Tuesday, September 26, 2000

1. Use the Wilcoxon signed rank test to analyze the data in problem 3.22 of the text.

Pair	1	2	3	4	5	6	7	8	9	10
M.I.	4.2	1.3	1.0	3.2	5.1	6.2	5.6	2.3	2.8	1.4
Control	3.5	1.1	1.1	0.8	2.3	2.8	4.3	1.1	2.2	1.0
Diff	0.7	0.2	-0.1	2.4	2.8	3.4	1.3	1.2	0.6	0.4
Rank	5	2	1	8	9	10	7	6	4	3

H: $\eta_d = 0$
A: $\eta_d > 0$

$T_+ = 54 \quad T_- = 1$

Table G

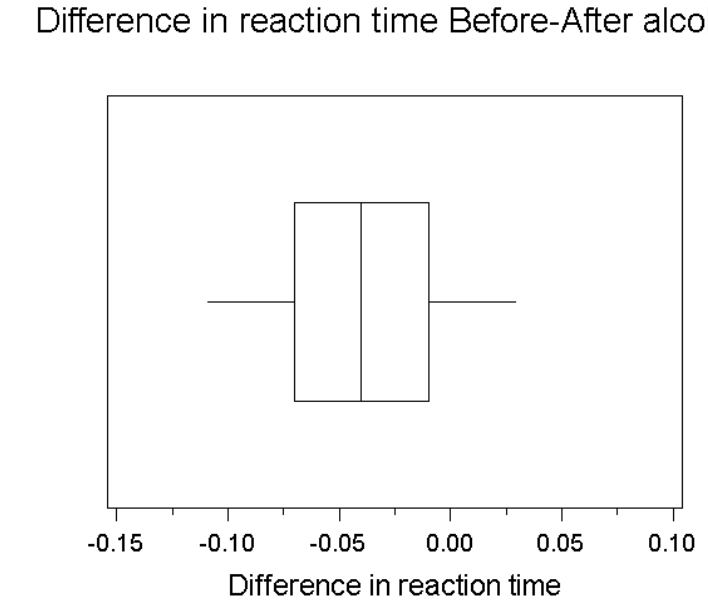
	Left		Right
n	T	P	T
10	1	0.002	54

The P-value is 0.002. Since the P-value is so small we reject H. These data support the study in *The Lancet* that suggests that acute myocardial infarction (M.I.) patents consume more coffee than the other patients.

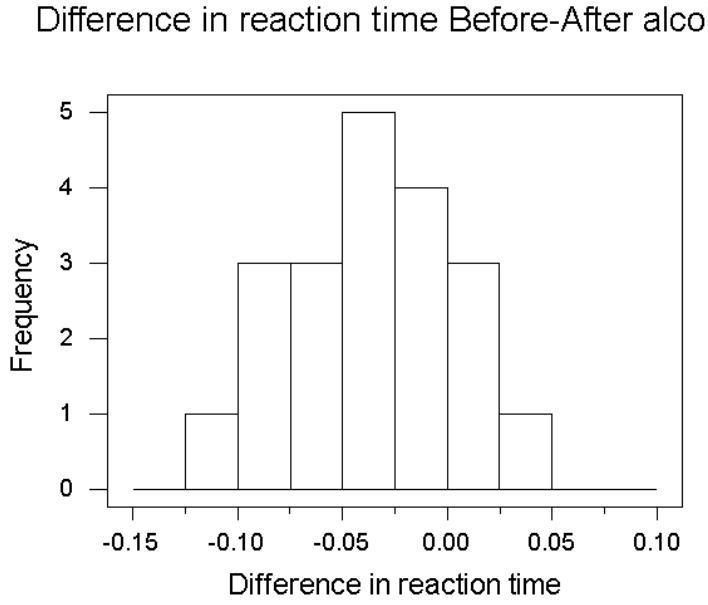
2. A random sample of 20 people who drove automobiles was selected to see if alcohol affected reaction time. Each driver's reaction time was measured in a laboratory before and after drinking a specified amount of beverage containing alcohol. The reaction times in seconds are given below. We are interested in the change in reaction time, Before – After.

Driver	Before	After	Diff	Rank	Driver	Before	After	Diff	Rank
1	0.68	0.73	-0.05	11.5	11	0.65	0.72	-0.07	14.5
2	0.62	0.64	-0.02	5	12	0.59	0.60	-0.01	2
3	0.68	0.66	0.02	5	13	0.81	0.78	0.03	7.5
4	0.82	0.92	-0.10	17.5	14	0.67	0.66	0.01	2
5	0.58	0.68	-0.10	17.5	15	0.65	0.68	-0.03	7.5
6	0.80	0.87	-0.07	14.5	16	0.76	0.77	-0.01	2
7	0.72	0.76	-0.04	9.5	17	0.61	0.72	-0.11	19
8	0.65	0.70	-0.05	11.5	18	0.86	0.86	0.00	—
9	0.84	0.88	-0.04	9.5	19	0.72	0.74	-0.02	5
10	0.73	0.79	-0.06	13	20	0.89	0.97	-0.08	16

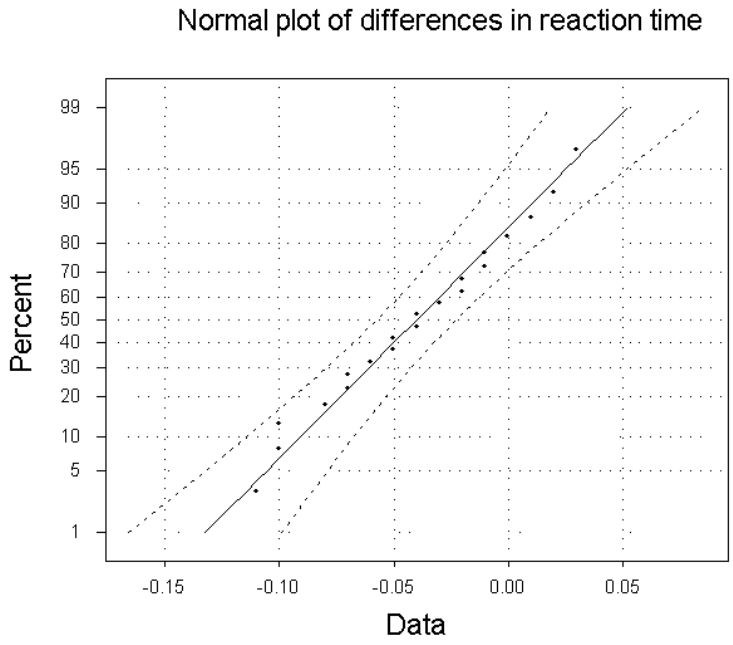
(a) The box plot of the differences looks fairly symmetric. This supports an assumption of symmetry.



The histogram looks fairly symmetric and mounded in the middle. This supports the assumption of normality.



The normal probability plot shows the data falling along a straight line, within reasonable bounds of natural variability. This supports the assumption of normality.



(b) Test of hypothesis for the mean difference.

H: $\mu_d = 0$
A: $\mu_d < 0$

$\overline{X}_d = \frac{-0.80}{20} = -0.04 \quad s_d = \sqrt{\frac{0.0314}{19}} = \sqrt{0.001653} = 0.04065$

$t = \frac{-0.04}{\left(\frac{0.04065}{\sqrt{30}}\right)} = \frac{-0.04}{0.00909} = -4.40$

P-value < 0.0005

Since the P-value is so small, we should reject H. There is evidence that the mean reaction time after consuming alcohol is longer than before.

(c) 95% Confidence interval on the mean difference.

df = 19 t*=2.093

$\overline{X}_d \pm t^* \left(\frac{s_d}{\sqrt{n}}\right)$

$-0.04 \pm 2.093 \left(\frac{0.04065}{\sqrt{20}}\right)$

-0.04 ± 0.019

-0.059 to -0.021

(d) Wilcoxon signed rank test for the median difference.

H: $\eta_d = 0$
A: $\eta_d < 0$

$T_+ = 14.5 \quad T_- = 175.5$

Because Table G only goes up to n=15, we must use a normal approximation.

mean of $T_+ = \frac{n(n+1)}{4} = \frac{19(20)}{4} = 95$

std dev of $T_+ = \sqrt{\frac{n(n+1)(2n+1)}{24}} = \sqrt{\frac{19(20)(39)}{24}} = \sqrt{617.5} = 24.85$

corrected std dev of $T_+ = \sqrt{\frac{19(20)(39)}{24} - \frac{3(8)+3(8)+2(3)+2(3)+2(3)+2(3)+2(3)}{48}} = \sqrt{617.5 - \frac{78}{48}} = \sqrt{615.875} = 24.82$

$$\begin{aligned} P-value &= Pr(T \geq 175.5) \\ &\doteq Pr\left(Z \geq \frac{175.5 - 0.5 - \frac{19(20)}{4}}{\sqrt{\frac{19(20)(39)}{24}}}\right) \\ &= Pr\left(Z \geq \frac{175 - 95}{\sqrt{617.5}}\right) \\ &= Pr\left(Z \geq \frac{80}{24.85}\right) \\ &= Pr(Z \geq 3.22) = 0.0006 \end{aligned}$$

Since the P-value is so small, we should reject H. There is evidence that the median reaction time after consuming alcohol is longer than before.

(e) Confidence interval based on the Wilcoxon signed rank test.

For n = 20, mean of T = $\frac{20(21)}{4} = 105$, std dev of T = $\sqrt{\frac{20(21)(41)}{24}} = \sqrt{717.5} = 26.79$. The approximate value of k for 95% confidence is:

$$k = 105 - 0.5 - 1.96(26.79) = 52$$

Rounding down to k = 52, an approximate 95% confidence interval will extend from the 53^{rd} ordered Walsh average to the 158^{th} ordered Walsh average. See the next page for a table of the 210 Walsh averages. That is from –0.060 to –0.020. This does not cover zero and so a median difference of zero is not acceptable.

(f) Sign test for the median difference.

H: $\eta_d = 0$
A: $\eta_d > 0$

$S_+ = 3 \quad S_- = 16$

Table F	Left		Right
n	S	P	S
19	3	0.0022	16

Since the P-value is so small, we should reject H. There is evidence that the median reaction time after consuming alcohol is longer than before.

(g) Confidence interval for the median difference based on the sign test.

Table F	Left		Right
n	S	P	S
20	5	0.0207	15

A 95.86% Confidence interval for the median difference will extend from the 6^{th} ordered value to the 15^{th} ordered value. That is from –0.07 to –0.01. This does not cover zero and so a median difference of zero is not acceptable.

(h) All three procedures come to the same conclusion that there is a significant difference in reaction times before and after consumption of alcohol. The t-test has the smallest P-value, followed by the Wilcoxon signed rank test with the sign test having the largest (but still very small) P-value.

The 95% confidence intervals also show a pattern in their widths. The t-CI is the narrowest. The CI based on the Walsh averages is slightly wider. The CI based on the ordered values is the widest.

3. Use the Wilcoxon signed rank test to analyze the data in problem 3.19 of the text.

Student	1st	2nd	Diff	Rank	Student	1st	2nd	Diff	Rank
1	101	104	3	7.5	11	110	114	4	10
2	115	117	2	5	12	125	128	3	7.5
3	96	107	11	18	13	104	109	5	12
4	93	92	−1	2	14	110	111	1	2
5	80	90	10	17	15	104	104	0	—
6	102	102	0	—	16	106	104	−2	5
7	95	94	−1	2	17	100	107	7	14
8	125	131	6	13	18	107	115	8	15
9	130	132	2	5	19	122	118	−4	10
10	111	120	9	16	20	114	118	4	10

H: $\eta_d = 0$
A: $\eta_d > 0$

$T_+ = 152 \quad T_- = 19 \quad n = 18$

Because Table G only goes up to n=15, we must use a normal approximation.

mean of $T_+ = \frac{n(n+1)}{4} = \frac{18(19)}{4} = 85.5$

std dev of $T_+ = \sqrt{\frac{n(n+1)(2n+1)}{24}} = \sqrt{18(19)(39)24} = \sqrt{527.25} = 22.962$

corrected std dev of $T_+ = \sqrt{\frac{n(n+1)(2n+1)}{24} - \frac{\sum t_i(t_i^2-1)}{48}} = \sqrt{\frac{18(19)(37)}{24} - \frac{24+24+6+24}{48}}$
$= \sqrt{527.25 - \frac{78}{48}} = \sqrt{525.625} = 22.9265$

$$\begin{aligned} P-value &= Pr(T \geq 152) \\ &\doteq Pr\left(Z \geq \frac{152 - 0.5 - 85.5}{\sqrt{525.625}}\right) \\ &= Pr\left(Z \geq \frac{66}{22.93}\right) \\ &= Pr(Z \geq 2.88) = 0.0020 \end{aligned}$$

Since the P-value is so small, we should reject H. There is evidence that the median score after study is higher than before.

Walsh averages for part (e) of problem 2.

−0.110	−0.100	−0.100	−0.080	−0.070	−0.070	−0.060	−0.050	−0.050	−0.040
−0.110	−0.105	−0.105	−0.095	−0.090	−0.090	−0.085	−0.080	−0.080	−0.075
	−0.100	−0.100	−0.090	−0.085	−0.085	−0.080	−0.075	−0.075	−0.070
		−0.100	−0.090	−0.085	−0.085	−0.080	−0.075	−0.075	−0.070
			−0.080	−0.075	−0.075	−0.070	−0.065	−0.065	−0.060
				−0.070	−0.070	−0.065	−0.060	−0.060	−0.055
					−0.070	−0.065	−0.060	−0.060	−0.055
						−0.060	−0.055	−0.055	−0.050
							−0.050	−0.050	−0.045
								−0.050	−0.045
									−0.040

−0.040	−0.030	−0.020	−0.020	−0.010	−0.010	0.000	0.010	0.020	0.030
−0.075	−0.070	−0.065	−0.065	−0.060	−0.060	−0.055	−0.050	−0.045	−0.040
−0.070	−0.065	**−0.060**	−0.060	−0.055	−0.055	−0.050	−0.045	−0.040	−0.035
−0.070	−0.065	−0.060	−0.060	−0.055	−0.055	−0.050	−0.045	−0.040	−0.035
−0.060	−0.055	−0.050	−0.050	−0.045	−0.045	−0.040	−0.035	−0.030	−0.025
−0.055	−0.050	−0.045	−0.045	−0.040	−0.040	−0.035	−0.030	−0.025	−0.020
−0.055	−0.050	−0.045	−0.045	−0.040	−0.040	−0.035	−0.030	−0.025	−0.020
−0.050	−0.045	−0.040	−0.040	−0.035	−0.035	−0.030	−0.025	−0.020	−0.015
−0.045	−0.040	−0.035	−0.035	−0.030	−0.030	−0.025	−0.020	−0.015	−0.010
−0.045	−0.040	−0.035	−0.035	−0.030	−0.030	−0.025	−0.020	−0.015	−0.010
−0.040	−0.035	−0.030	−0.030	−0.025	−0.025	−0.020	−0.015	−0.010	−0.005
−0.040	−0.035	−0.030	−0.030	−0.025	−0.025	−0.020	−0.015	−0.010	−0.005
	−0.030	−0.025	−0.025	−0.020	**−0.020**	−0.015	−0.010	−0.005	0.000
		−0.020	−0.020	−0.015	−0.015	−0.010	−0.005	0.000	0.005
			−0.020	−0.015	−0.015	−0.010	−0.005	0.000	0.005
				−0.010	−0.010	−0.005	0.000	0.005	0.010
					−0.010	−0.005	0.000	0.005	0.010
						0.000	0.005	0.010	0.015
							0.010	0.015	0.020
								0.020	0.025
									0.030