Stat 403 - Solution to Assignment 6

Turned in Tuesday, October 31, 2000

1. Two machines are used to measure a vibration property of a rubber product. The data below are the vibrational force (kg) of samples from each machine. The items entering each machine are randomly selected from the production process that produces items that are similar in their vibrational characteristics. Any differences apparent in the data are thus assumed to be due to the machines.

   (a) Machine 1 tends to have lower vibrational forces than machine 2. The median for machine 1 is around 22 while the median for machine 2 is around 40. The spread of both machines is about the same, although machine 2 has one unusually large value which is affecting it's spread considerably. Without that one value, machine 2 would be less variable than machine 1. The shape of the distribution for machine 1 is skewed towards higher values. The distribution for machine 2 is fairly symmetric especially if you discount the one unusually large value.

   (b) For the two-sample t-test the value of the test statistic is $t = -2.95$ with a P-value of 0.0055. Since this P-value is so small, we should reject the null hypothesis of no difference in means. There is a significant difference in mean vibrational forces for the two machines. The estimated difference in mean vibrational force is from 4.5 to 24.1 kg less for machine 1.

   Note: This analysis assumes equal variances for the two distributions.

   (c) For the ANOVA the value of the test statistic is $F = 8.71$ with a P-value of 0.005. Since this P-value is so small, we should reject the null hypothesis of no difference in means. There is a significant difference in mean vibrational force for the two machines.

   (d) Both procedures reach the same conclusions. Both procedures have the same P-values (0.005). Both procedures have the same estimate of the population standard deviation (15.1). The value of the test statistic for the ANOVA (F = 8.71) is equal to the square of the t statistic ($t^2 = [-2.95]^2 = 8.71$).

   (e) Wilcoxon-Mann-Whitney test.

Machine 1						Machine 2					
	R		R		R		R		R		R
10.6	1	20.2	9	34.8	19	19.5	8	39.4	23	50.1	31
11.3	2	22.1	10	36.0	20	25.6	13	40.3	24	51.3	32
14.2	3	23.0	11	43.3	26	27.0	14	42.7	25	51.6	33
17.0	4	23.2	12	44.3	27	33.5	17	45.2	28	52.2	34
17.8	5	27.3	15	53.9	35	37.6	21	48.3	29	55.5	36
18.6	6	31.7	16	59.1	37	38.4	22	48.9	30	86.2	39
19.3	7	34.0	18	63.5	38						

$W_1 = 321 \Rightarrow U_1 = 321 - \frac{21(22)}{2} = 321 - 231 = 90$

$W_2 = 459 \Rightarrow U_2 = 459 - \frac{18(19)}{2} = 459 - 171 = 288$

mean of U $= \frac{mn}{2} = \frac{21(18)}{2} = 189$

std dev of U $= \sqrt{\frac{mn(m+n+1)}{12}} = \sqrt{\frac{21(18)(40)}{12}} = \sqrt{1260} = 35.5$

P-value $= 2Pr(U \geq 288) \doteq 2Pr\left(Z \geq \frac{288-0.5-189}{35.5}\right) = 2Pr(Z \geq 2.77) = 2(.0028) =$ .0056

Since the P-value is so small we should reject the hypothesis of no difference in distributions. The two machines have significantly different distributions for vibrational force.

Note: If you do not use the correction for continuity then

P-value $= 2Pr(U \geq 288) \doteq 2Pr\left(Z \geq \frac{288-189}{35.5}\right) = 2Pr(Z \geq 2.79) = 2(.0026) =$ .0052

(f) Kruskal-Wallis test.

$R_1 = 321 \quad R_2 = 459$

$$\begin{aligned} H &= \frac{12}{39(40)}\left[\frac{321^2}{21} + \frac{459^2}{18}\right] - 3(40) \\ &= \frac{12}{39(40)}[4906.71 + 11704.5] - 120 \\ &= 127.78 - 120 = 7.78 \end{aligned}$$

The P-value is between 0.001 and 0.01 (0.005 from Minitab). Since the P-value is so small we should reject the hypothesis of no difference in distributions. The two machines have significantly different distributions for vibrational force.

(g) Both procedures reach the same conclusion. The P-values for both procedures are approximately the same (they are exactly the same if you do not use the continuity correction for the Mann-Whitney-Wilcoxon). The value of the Kruskal-Wallis test statistic (H = 7.78) is the square of the Mann-Whitney-Wilcoxon z value using no continuity correction ($z^2 = 2.77^2 = 7.78$).

2. A graduate in exercise physiology wants to compare four training methods suitable for marathon runners. Forty-nine individuals interested in training for a marathon are randomly divided into four groups of roughly equal sizes. Each group was then randomly assigned to a particular training method. At the end of three months, the participants are to be timed on a trial marathon run. For various reasons, not all participants complete the training course or trial marathon run. Below are times (hr:min) for the 40 participants grouped according to training methods 1, 2, 3 and 4.

1		2		3		4	
Time	Rank	Time	Rank	Time	Rank	Time	Rank
3:01	10	2:49	1	2:53	3	2:50	2
3:32	19	2:54	4	2:59	8	3:11	14
3:37	20	2:55	5	3:10	13	3:12	15
3:38	21	2:56	6	3:27	17	3:39	22
3:43	23	2:58	7	3:46	24	4:17	31
4:06	27	3:00	9	3:58	25	4:25	33
4:15	29	3:05	11	4:01	26	4:32	35
4:16	30	3:06	12	4:13	28	4:39	36
4:27	34	3:14	16	4:21	32	4:46	37
4:59	39	3:28	18			4:54	38
5:04	40						
$R_1 =$	292	$R_2 =$	89	$R_3 =$	176	$R_4 =$	263

(a) Kruskal-Wallis test.

$$H = \frac{12}{40(41)}\left[\frac{292^2}{11}+\frac{89^2}{10}+\frac{176^2}{9}+\frac{263^2}{10}\right]-3(41)$$

$$= \frac{12}{40(41)}[7751.27+792.1+3441.78+6916.9]-123$$

$$= 138.31-123 = 15.31$$

The P-value is between 0.001 and 0.01 (0.002 from Minitab). Since the P-value is so small we should reject the hypothesis of no difference in distributions. Some of the four training methods produce significantly different median times to run a marathon.

(b) Multiple comparisons.

Comparison	Difference in Average Ranks	Cut off	Significant?
1 & 2	17.65	13.16	Yes
1 & 3	6.99	13.54	No
1 & 4	0.02	13.16	No
2 & 3	10.66	13.84	No
2 & 4	17.40	13.47	Yes
3 & 4	6.74	13.84	No

Training method 2 is significantly different from both methods 1 and 4. Method 2 tends to have considerably lower times.

(c) Compute the 90 differences and construct a confidence interval based on these differences and Table H.

	261	253	241	238	226	207	190	179	173
169	92	84	72	69	57	38	21	10	4
174	87	79	67	64	52	33	16	5	−1
175	86	78	66	63	51	32	15	4	−2
176	85	77	65	62	50	31	14	3	−3
178	83	75	63	60	48	29	12	1	−5
180	81	73	61	58	46	27	10	−1	−7
185	76	68	56	53	41	22	5	−6	−12
186	75	67	55	52	40	21	4	−7	−13
194	67	59	47	44	32	13	−4	−15	−21
208	53	45	33	30	18	−1	−18	−29	−35

Table H

n	Left U	P	Right U
		m=9	
10	20	0.022	70
	21	0.027	69

A 95.6% confidence interval for the difference in median time goes from the 21st to the 70th ordered Δ_{ij}: (4, 66)

A 94.6% confidence interval for the difference in median time goes from the 22nd to the 69th ordered Δ_{ij}: (5, 65)

Training method 2 has a median time that is from about 5 to 65 minutes shorter than the median time for method 3.

(d) Mood's median test.

The combined sample median is 3:37.5.

	Training Method 1	2	3	4	
$\# < 3:37.5$	3	10	4	3	20
$\# > 3:37.5$	8	0	5	7	20

$$
\begin{aligned}
\chi^2 &= \sum \frac{(O-E)^2}{E} \\
&= \frac{2.5^2}{5.5} + \frac{5^2}{5} + \frac{0.5^2}{4.5} + \frac{2^2}{5} \\
&+ \frac{2.5^2}{5.5} + \frac{5^2}{5} + \frac{0.5^2}{4.5} + \frac{2^2}{5} \\
&= 13.98
\end{aligned}
$$

With three degrees of freedom, the P-value is between 0.001 and 0.01 (0.003 form Minitab). Since the P-value is so small we should reject the hypothesis of no difference in distributions. Some of the four training methods produce significantly different times to run the marathon.