STAT/MATH 415 HW#8

Nov 27, 2017

EXERCISES

9.3.1 Let μ_1, μ_2, μ_3 be, respectively, the means of three normal distributions with a common but unknown variance σ^2. In order to test, at $\alpha = 0.05$ significance level, the hypothesis $H_0 : \mu_1 = \mu_2 = \mu_3$ against all possible alternative hypotheses, we take a random sample of size 4 from each distribution. Determine whether we accept or reject H_0 if the observed values are as follows:

x_1 :	5	9	6	8
x_2 :	11	13	10	12
x_3 :	10	6	9	9

Answer: Here $m = 3,\ n = 12$ and $\bar{x}_{1\cdot} = 7,\ \bar{x}_{2\cdot} = 11.5,\ \bar{x}_{3\cdot} = 8.5,\ \bar{x}_{\cdot\cdot} = 9$

$$SS(TO) = \sum_{i=1}^{3}\sum_{j=1}^{4}(x_{ij} - \bar{x}_{\cdot\cdot})^2 = 66 \qquad SS(T) = \sum_{i=1}^{3} 4(\bar{x}_{i\cdot} - \bar{x}_{\cdot\cdot})^2 = 42 \qquad SS(E) = \sum_{i=1}^{3}\sum_{j=1}^{4}(x_{ij} - \bar{x}_{i\cdot})^2 = 24$$

The ANOVA table is

Source	d.f.	SS	MS	F
Treatment	2	42	21	7.875
Error	9	24	8/3	
Total	11	66	6	

$F_{0.05}(2, 9) = 4.26$, so the calculated F = 7.875 is in critical region $[4.26, \infty)$, we should reject H_0.

9.3.15 Ledolter and Hogg report that an operator of a feedlot wants to compare the effectiveness of three different cattle feed supplements. He selects a random sample of 15 one-year-old heifers from his lot of over 1000 and divides them into three groups at random. Each group gets a different feed supplement. Upon noting that one heifer in group A was lost due to an accident, the operator records the gains in weight (in pounds) over a six-month period as follows:

Group A:	500	650	530	680	
Group B:	700	620	780	830	860
Group C:	500	520	400	580	410

a) Test whether there are differences in the mean weight gains due to the three different feed supplements.

Answer: Let μ_1, μ_2, μ_3 denote the mean weight gains for three supplements respectively, then $H_0 : \mu_1 = \mu_2 = \mu_3$.

$$\bar{x}_{1\cdot} = 590 \qquad \bar{x}_{2\cdot} = 758 \qquad \bar{x}_{3\cdot} = 482 \qquad \bar{x}_{\cdot\cdot} = 604.286 \qquad n_1 = 4 \qquad n_2 = 5 \qquad n_3 = 5$$

$$SS(TO) = \sum_{i=1}^{3}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_{\cdot\cdot})^2 = 278171.429 \qquad SS(T) = \sum_{i=1}^{3} n_i(\bar{x}_{i\cdot} - \bar{x}_{\cdot\cdot})^2 = 193011.429 \qquad SS(E) = \sum_{i=1}^{3}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_{i\cdot})^2 = 85160$$

Source	d.f.	SS	MS	F
Treatment	2	193011.429	96505.714	12.466
Error	11	85160	7741.818	
Total	13	278171.429	21397.802	

As shown in the ANOVA table, the calculated F is 12.466, greater than $F_{0.05}(2, 11) = 3.98$, therefore we should reject H_0 and conclude the differences in mean weight gains are due to different feed supplements.

8.4.3 Let X equal the weight (in grams) of a Hershey's grape-flavored Jolly Rancher. Denote the median of X by m. We shall test $H_0 : m = 5.900$ against $H_1 : m > 5.900$. A random sample of size $n = 25$ yielded the following ordered data:

5.625	5.665	5.697	5.837	5.863	5.870	5.878	5.884	5.908	5.967	6.019	6.020	6.029
6.032	6.037	6.045	6.049	6.050	6.079	6.116	6.159	6.186	6.199	6.307	6.387	

a) Use the sign test to test the hypothesis.

Answer: Let $Y_i = sign(X_i - m),\ Y = \sum Y_i = 17,\ Y \sim binom(25, p),\ \hat{p} = 0.68$. Under $H_0,\ p = 0.5$

$$Z = \frac{0.68 - 0.5}{\sqrt{0.5 \times 0.5/25}} = \frac{0.18}{0.1} = 1.8$$

Assume the significance level $\alpha = 0.05$, then critical region is $Z > Z_{0.05} = 1.645$, thus we should reject H_0.

b) Use the Wilcoxon test statistic to test the hypothesis.

Answer: The table shows the rank and the sign of $|X_i - m|$

X_i	$\lvert X_i - m\rvert$	R_i	S_i	X_i	$\lvert X_i - m\rvert$	R_i	S_i
5.625	0.275	21	-1	6.032	0.132	11	+1
5.665	0.235	19	-1	6.037	0.137	12	+1
5.697	0.203	17	-1	6.045	0.145	13	+1
5.837	0.063	6	-1	6.049	0.149	14	+1
5.863	0.037	5	-1	6.050	0.150	15	+1
5.870	0.030	4	-1	6.079	0.179	16	+1
5.878	0.022	3	-1	6.116	0.216	18	+1
5.884	0.016	2	-1	6.159	0.259	20	+1
5.908	0.008	1	+1	6.186	0.286	22	+1
5.967	0.067	7	+1	6.199	0.299	23	+1
6.019	0.119	8	+1	6.307	0.407	24	+1
6.020	0.120	9	+1	6.387	0.487	25	+1
6.029	0.129	10	+1				

$$W = \sum_{i=1}^{25} R_i S_i = 171 \qquad Z = \frac{W}{\sqrt{n(n+1)(2n+1)/6}} = \frac{171}{74.330} = 2.300$$

When $\alpha = 0.05$, then the critical region is $[1.645, +\infty)$, thus we should reject H_0.

c) Use a t test to test the hypothesis.

Answer: $\bar{X} = 5.996,\ s^2 = 0.0341$. Since $t_{0.05}(24) = 1.711$, T is in the critical region, so we should reject H_0

$$T = \frac{\bar{X} - m}{\sqrt{s^2/n}} = \frac{5.996 - 5.900}{0.0369} = 2.608$$

d) Write a short comparison of the three tests.

Answer: All three tests reject H_0, but Wilcoxon test statistic is larger than that in sign test, indicating that Wilcoxon test may have a greater power in detecting deviations. This increased power is probably due to the consideration of values of the differences in addition to the signs. For t test, we applied the method for mean parameter estimate to median estimate as an approximation.

8.4.7 Let X equal the weight in pounds of a "1-pound" bag of carrots. Let m equal the median weight of a population of these bags. Test the null hypothesis $H_0 : m = 1.14$ against the alternative hypothesis $H_1 : m > 1.14$.

a) With a sample size of $n = 14$, use Wilcoxon statistic to define a critical region. Use $\alpha \approx 0.10$.

Answer: $Z_{0.10} = 1.28$ so the critical region is $[1.28, +\infty)$

b) What would be your conclusion if the observed weights were

1.12 1.13 1.19 1.25 1.06 1.31 1.12 1.23 1.29 1.17 1.20 1.11 1.18 1.23

Answer: Here shows the absolute values, ranks and signs.

$\lvert X_i - m\rvert$	0.02	0.01	0.05	0.11	0.08	0.17	0.02	0.09	0.15	0.03	0.06	0.03	0.04	0.09
R_i	2.5	1	7	12	9	14	2.5	10.5	13	4.5	8	4.5	6	10.5
S_i	-1	-1	+1	+1	-1	+1	-1	+1	+1	+1	+1	-1	+1	+1

$$W = \sum_{i=1}^{14} R_i S_i = 66 \qquad Z = \frac{W}{\sqrt{n(n+1)(2n+1)/6}} = \frac{66}{31.859} = 2.07$$

Since $Z = 2.07$ is in the critical region, we should reject H_0.

c) What is the p-value of your test?

Answer: p-value is $P(Z > 2.07) = 0.0192$

8.4.15 With $\alpha = 0.05$, use the Wilcoxon statistic to test $H_0 : m_X = m_Y$ against a two-sided alternative. Use the following observations of X and Y, which have been ordered for your convenience:

x:	-2.3864	-2.2171	-1.9148	-1.9097	-1.4883	-1.2007	-1.1077
-0.3601	0.4325	1.0598	1.3035	1.5241	1.7133	1.7656	2.4912
y:	-1.7613	-0.9391	-0.7437	-0.5530	0.2469	0.0647	0.2031
0.3219	0.3579	0.6431	0.6557	0.6724	0.6762	0.9041	1.3571

Answer: Here shows the rank of combined values and their origins.

rank	-2.3864	-2.2171	-1.9148	-1.9097	-1.7613	-1.4883	-1.2007	-1.1077	-0.9391	-0.7437
origin	x	x	x	x	y	x	x	x	y	y
rank	-0.5530	-0.3601	-0.2469	0.0647	0.2031	0.3219	0.3579	0.4325	0.6431	0.6557
origin	y	x	y	y	y	y	y	x	y	y
rank	0.6724	0.6762	0.9041	1.0598	1.3035	1.3571	1.5241	1.7133	1.7656	2.4912
origin	y	y	y	x	x	y	x	x	x	x

Let W be the sum of ranks from Y, under H_0

$$E(W) = \frac{n_2(n_1+n_2+1)}{2} = \frac{15 \times 31}{2} = 232.5$$

$$Var(W) = \frac{n_1 n_2(n_1+n_2+1)}{12} = \frac{15 \times 15 \times 31}{12} = 581.25$$

$$W = 5+9+10+11+13+14+15+16+17+19+20+21+22+23+26 = 241$$

$$Z = \frac{W - E(W)}{\sqrt{Var(W)}} = \frac{241 - 232.5}{24.11} = 0.3526$$

The critical region is $|Z| \leq Z_{0.025} = 1.96$, so we do not reject H_0.