

Text Exercise Set 27

NAME:

27-1 Three different methods, labeled A, B, and C, to purify a particular drug are available to a manufacturer. A 0.01 significance level is chosen for a hypothesis test to see if there is any evidence of a difference among these three methods in the mean amount of impurities left after purification. The data consisting of independent random samples is displayed in the table on the right.

Impurity Levels (ppm) for Three Methods						
Method A:	104	112	108	114	110	112
Method B:	109	104	105	112		
Method C:	106	102	110	108	104	

- (a) Explain how the data for this hypothesis test is appropriate for a one-way ANOVA.

- (b) Complete the four steps of the hypothesis test below. As part of the second step, complete the construction of the ANOVA table below, where you should find that $SSB = 45$, $SSE = 145$, and Fisher's f statistic is $f_{2,12} = 1.86$.

Step 1 H_0 :
 H_1 :
 $\alpha =$

Step 2

Step 3

Step 4

<i>Source</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>f</i>	<i>p-value</i>
Error					
Total					

27-1 - continued

- (c) In the list below, circle the best graphical display for this data and say why.
- (i) multiple pie charts (ii) scatter plot (iii) multiple box plots
- (d) Considering the results of the hypothesis test, decide which of the Type I or Type II errors is possible, and describe this error.
- (e) Decide whether H_0 would have been rejected or would not have been rejected with each of the following significance levels: (i) $\alpha = 0.05$, (ii) $\alpha = 0.10$.
- (f) What would the presence of one or more outliers in the data suggest about using the f statistic?
- (g) In each table, alter the data in this exercise so that the stated condition is true.

Data where SSB is equal to zero, but SSE is not equal to zero.

Impurity Levels (ppm) for Three Methods

Method A:

Method B:

Method C:

Data where SSE is equal to zero, but SSB is not equal to zero.

Impurity Levels (ppm) for Three Methods

Method A:

Method B:

Method C:

27-2 A 0.01 significance level is chosen for a hypothesis test to see if there is any evidence of a difference among five varieties of oranges in the mean sugar content. The data consisting of independent random samples is displayed in the table on the right.

Sugar Content for Five Varieties of Oranges (hundredths of an ounce)	
Variety A:	41 36 28
Variety B:	29 43 39
Variety C:	33 43 41
Variety D:	40 32 36
Variety E:	34 38 27

- (a) Explain how the data for this hypothesis test is appropriate for a one-way ANOVA.
- (b) Complete the four steps of the hypothesis test below. As part of the second step, complete the construction of the ANOVA table below, where you should find that $SSB = 60$, $SSE = 340$, and Fisher's f statistic is $f_{4, 10} = 0.44$.

Step 1 H_0 :
 H_1 :
 $\alpha =$

Step 2

Step 3

Step 4

<i>Source</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>f</i>	<i>P-value</i>
Error					
Total					

27-2 - continued

- (c) In the list below, circle the best graphical display for this data and say why.
- (i) multiple pie charts (ii) scatter plot (iii) multiple box plots
- (d) Considering the results of the hypothesis test, decide which of the Type I or Type II errors is possible, and describe this error.
- (e) Decide whether H_0 would have been rejected or would not have been rejected with each of the following significance levels: (i) $\alpha = 0.05$, (ii) $\alpha = 0.10$.
- (f) What would the presence of one or more outliers in the data suggest about using the f statistic?
- (g) In each table, alter the data in this exercise so that the stated condition is true.

Data where SSB is equal to zero, but SSE is not equal to zero.

**Sugar Content for Five Varieties
of Oranges (hundredths of an ounce)**

Variety A:
Variety B:
Variety C:
Variety D:
Variety E:

Data where SSE is equal to zero, but SSB is not equal to zero.

**Sugar Content for Five Varieties
of Oranges (hundredths of an ounce)**

Variety A:
Variety B:
Variety C:
Variety D:
Variety E:

27-3 A 0.05 significance level is chosen to see if there is any evidence of an increase in mean milk sales after a series of half-minute commercials promoting milk consumption are shown on TV regularly for one week. Unit weekly dairy sales are recorded at 18 stores before and after the promotion, and the results are displayed in the table on the right.

Unit Dairy Sales Before and After a Promotion		
<u>Store No.</u>	<u>Before</u>	<u>After</u>
1	124	136
2	107	102
3	82	89
4	114	128
5	940	1080
6	75	85
7	105	105
8	94	95
9	865	985
10	620	820
11	80	75
12	760	725
13	330	350
14	110	115
15	125	119
16	400	425
17	400	450
18	175	215

- (a) Verify that, when subtracting the sales before the promotion from the sales after the promotion, the paired t test statistic is $t_{17} = +2.325$.
- (b) Complete the four steps of the hypothesis test by completing the following:

Step 1 H_0 :
 H_1 :
 $\alpha =$

Step 2

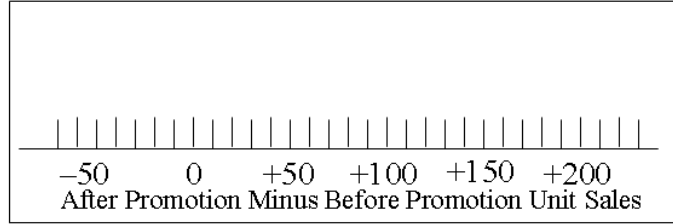
Step 3

Step 4

- (c) Considering the results of the hypothesis test, explain why a confidence for the mean increase in milk sales after the promotion would be of interest; then, verify that $n = 18$, $\bar{d} = 32.9444$, and $s_d = 60.1053$, and find and interpret a 95% confidence interval for the mean increase.

27-3 - continued

- (d) Complete the construction of the box plot on the right, and comment on whether the paired t statistic appears to be appropriate.



- (e) Considering the results of the hypothesis test, decide which of the Type I or Type II errors is possible, and describe this error.
- (f) Decide whether H_0 would have been rejected or would not have been rejected with each of the following significance levels: (i) $\alpha = 0.01$, (ii) $\alpha = 0.10$.

27-4 The variables "Expenses" and "Sales" in the CHAIN DATA, displayed as Data Set 23-1 at the end of Unit 23, are being used to see if there is any evidence that the mean sales exceeds the mean expenses among restaurants in a chain known as McDoogles. A 0.05 significance level is chosen for a hypothesis test, and the data (millions of dollars) are displayed for convenience as follows:

Expenses	1.0	1.2	2.8	1.9	0.3	1.5	3.4	1.0	1.6	0.9	1.6	1.9	0.6	1.0	2.5	1.0	1.9
Sales	0.1	1.8	4.0	6.1	5.3	4.0	7.4	3.6	2.2	7.5	4.6	8.0	3.3	2.5	2.6	8.1	1.7
Expenses	1.4	0.7	1.0	1.3	2.4	1.2	1.7	1.4	1.0	1.2	2.3	2.2	0.8				
Sales	6.7	3.8	5.2	7.8	5.1	7.8	7.7	2.8	4.9	8.0	2.0	2.5	5.1				

- (a) Verify that, when subtracting expenses from sales, the paired t test statistic is $t_{29} = +7.266$.
- (b) Complete the four steps of the hypothesis test by completing the following:

Step 1 H_0 :
 H_1 :
 $\alpha =$

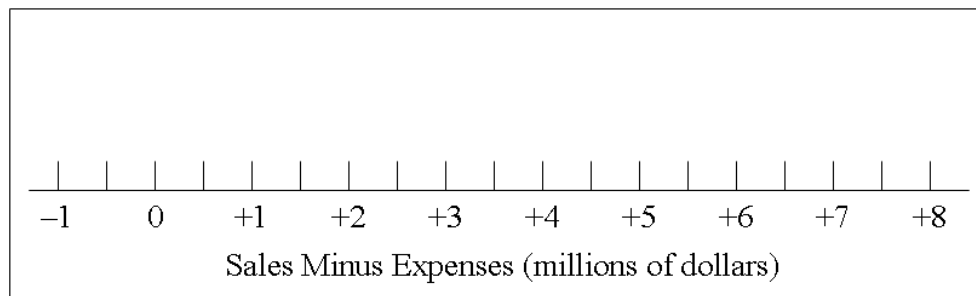
Step 2

Step 3

Step 4

27-4 - continued

- (c) Considering the results of the hypothesis test, explain why a confidence for the mean difference by which sales exceeds expenses would be of interest; then, verify that $n = 30$, $\bar{d} = 3.250$, and $s_d = 2.450$, and find and interpret a 95% confidence interval for the mean increase.



- (d) Complete the construction of the box plot above, and comment on whether the paired t statistic appears to be appropriate.
- (e) Considering the results of the hypothesis test, decide which of the Type I or Type II errors is possible, and describe this error.
- (f) Decide whether H_0 would have been rejected or would not have been rejected with each of the following significance levels: (i) $\alpha = 0.01$, (ii) $\alpha = 0.10$.

[This page is intentionally left blank]