Review 6

Use the traditional method to test the given hypothesis. Assume that the samples are independent and that they have been randomly selected

1) A researcher finds that of 1,000 people who said that they attend a religious service at least once a week, 31 stopped to help a person with car trouble. Of 1,200 people interviewed who had not attended a religious service at least once a month, 22 stopped to help a person with car trouble. At the 0.05 significance level, test the claim that the two proportions are equal. 1) ______________

Construct the indicated confidence interval for the difference between the two population means. Assume that the two samples are independent and that they have been randomly selected.

2) Independent samples from two different populations yield the following data. $\overline{x}_1 = 383$, $\overline{x}_2 = 448$, $s_1 = 44$, $s_2 = 18$. The sample size is 177 for both samples. Find the 95 percent confidence interval for $\mu_1 - \mu_2$. 2) _______

A) $-79 < \mu_1 - \mu_2 < -51$
B) $-72 < \mu_1 - \mu_2 < -58$
C) $-71 < \mu_1 - \mu_2 < -59$
D) $-66 < \mu_1 - \mu_2 < -64$

Test the indicated claim about the means of two populations. Assume that the two samples are independent and that they have been randomly selected.

3) A researcher wishes to determine whether people with high blood pressure can reduce their blood pressure by following a particular diet. Use the sample data below to test the claim that the treatment population mean μ_1 is smaller than the control population mean μ_2. Test the claim using a significance level of 0.01. 3) ______________

Treatment Group	Control Group
$n_1 = 85$	$n_2 = 75$
$\overline{x}_1 = 189.1$	$\overline{x}_2 = 203.7$
$s_1 = 38.7$	$s_2 = 39.2$

Construct a confidence interval for μ_d, the mean of the differences d for the population of paired data. Assume that the population of paired differences is normally distributed.

4) A coach uses a new technique in training middle distance runners. The times for 9 different athletes to run 800 meters before and after this training are shown below. 4) _______

Athlete	A	B	C	D	E	F	G	H	I
Time before training (seconds)	115.2	120.9	108.0	112.4	107.5	119.1	121.3	110.8	122.3
Time after training (seconds)	116.0	119.1	105.1	111.9	109.1	115.2	118.5	110.7	120.9

Construct a 99% confidence interval for the mean difference of the "before" minus "after" times.

A) $-0.82 < \mu_d < 3.26$
B) $-0.76 < \mu_d < 3.20$
C) $-0.54 < \mu_d < 2.98$
D) $-0.85 < \mu_d < 3.29$

Use the traditional method of hypothesis testing to test the given claim about the means of two populations. Assume that two dependent samples have been randomly selected from normally distributed populations.

5) The table below shows the weights of seven subjects before and after following a particular diet for two months. 5) ______________

Subject	A	B	C	D	E	F	G
Before	157	160	156	183	190	157	161
After	150	151	154	188	176	159	149

Using a 0.01 level of significance, test the claim that the diet is effective in reducing weight.

Test the indicated claim about the variances or standard deviations of two populations. Assume that the populations are normally distributed. Assume that the two samples are independent and that they have been randomly selected.

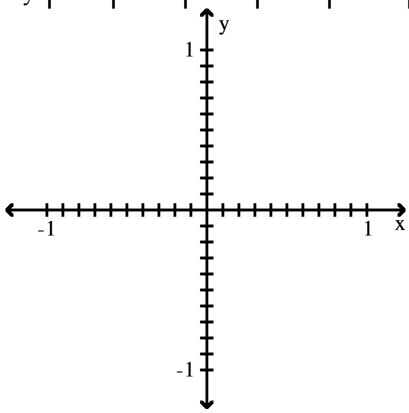
6) Test the claim that populations A and B have different variances. Use a significance level of 0.10. 6) ______________

Sample A	Sample B
$n = 28$	$n = 41$
$\bar{x}_1 = 19.2$	$\bar{x}_2 = 23.7$
$s = 4.7$	$s = 5.89$

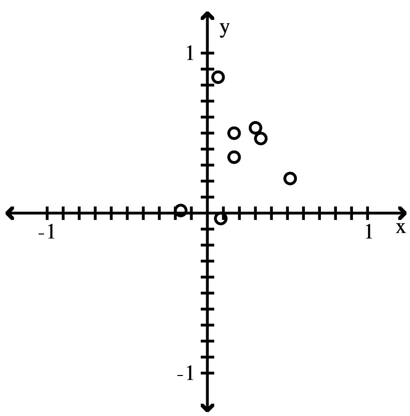
.**Construct a scatter diagram for the given data.**

7) 7) _______

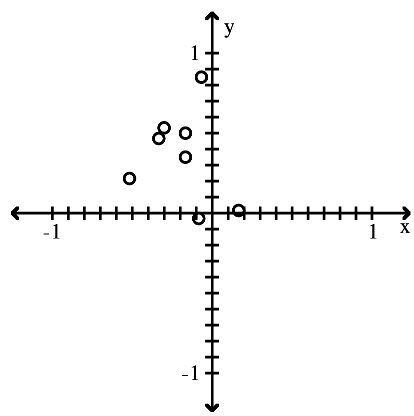
x	0.17	0.07	0.16	0.34	–0.16	0.3	0.51	0.08
y	0.5	0.85	0.35	0.46	0.01	0.53	0.22	–0.03



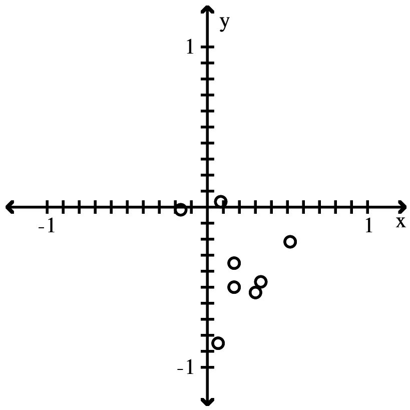
A)



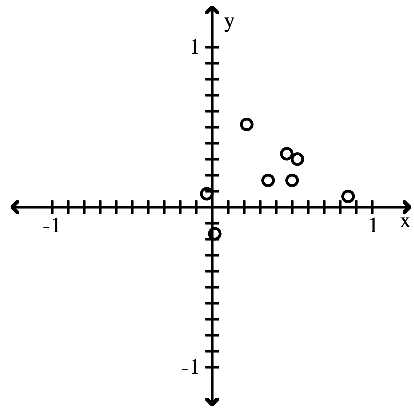
B)



C)



D)



Find the value of the linear correlation coefficient r.

8)

x	62	53	64	52	52	54	58
y	158	176	151	164	164	174	162

A) –0.775 B) 0.754 C) 0 D) –0.081

8) _______

Use the given data to find the equation of the regression line. Round the final values to three significant digits, if necessary.

9)

x	6	8	20	28	36
y	2	4	13	20	30

A) $\hat{y} = -2.79 + 0.950x$ B) $\hat{y} = -3.79 + 0.897x$

C) $\hat{y} = -3.79 + 0.801x$ D) $\hat{y} = -2.79 + 0.897x$

9) _______

Solve the problem.

10) For the data below, determine the value of the linear correlation coefficient r between y and x^2. Test whether the correlation is significant. Use a significance level of 0.01.

10) _______

x	1.2	2.7	4.4	6.6	9.5
y	1.6	4.7	9.9	24.5	39.0

A) 0.985 B) 0.913 C) 0.873 D) 0.990

Find the total variation for the paired data.

11) The equation of the regression line for the paired data below is $\hat{y} = 6.18286 + 4.33937x$. Find the explained variation, the unexplained variation, the total variation, and the coefficient of determination.

11) _______

x	9	7	2	3	4	22	17
y	43	35	16	21	23	102	81

A) 6,531.37 B) 6,693.27 C) 13.479 D) 6,544.86

Construct the indicated prediction interval for an individual y.

12) The equation of the regression line for the paired data below is $\hat{y} = 6.1829 + 4.3394x$ and the standard error of estimate is $s_e = 1.6419$. Find the 99% prediction interval of y for $x = 12$. 12) ______

x	9	7	2	3	4	22	17
y	43	35	16	21	23	102	81

A) $51.1 < y < 65.4$
B) $59.9 < y < 64.9$
C) $56.4 < y < 68.5$
D) $53.6 < y < 71.3$

1) H_0: $p_1 = p_2$. H_1: $p_1 \neq p_2$.
Test statistic: z = 1.93. Critical values: z = 1.96, –1.96.
Fail to reject the null hypothesis. There is not sufficient evidence to warrant rejection of the claim that the two proportions are equal.
2) B
3) H_0: $\mu_1 = \mu_2$. H_1: $\mu_1 < \mu_2$.
Test statistic t = –2.365. Critical value: t = –2.377.
Fail to reject the null hypothesis. There is not sufficient evidence to support the claim that the treatment population mean $\mu 1$ is smaller than the control population μ_2.
4) A
5) Test statistic t = 1.954. Critical value: t = 3.143.
Fail to reject H_0: $\mu_d = 0$. There is not sufficient evidence to support the claim that the diet is effective in reducing weight.
6) H_0: $\sigma^2{}_1 = \sigma^2{}_2$. H_1: $\sigma^2{}_1 \neq \sigma^2{}_2$.
Test statistic: F = 1.57. Critical value: F = 1.84.
Fail to reject the null hypothesis. There is not sufficient evidence to support the claim that populations A and B have different variances.
7) A
8) A
9) B
10) D
11) D
12) A

Problem 1

The data in the problem can be summarized as

$$n_1 = 1000, \quad \hat{p}_1 = \frac{31}{1000} = 0.0310,$$

$$n_2 = 1200, \quad \hat{p}_2 = \frac{22}{1200} \approx 0.0183,$$

$$\alpha = 0.05, \quad H_0: \quad p_1 = p_2, \quad H_1: \quad p_1 \neq p_2.$$

The alternative hypothesis indicates that we will use a two-tailed test.

The pooled sample proportion (p. 457)

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{31 + 22}{1000 + 1200} = \frac{53}{2200} \approx 0.0241.$$

The test statistic (p. 457)

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\bar{p}(1-\bar{p})}{n_1} + \frac{\bar{p}(1-\bar{p})}{n_2}}} =$$

$$= \frac{0.0310 - 0.0183}{\sqrt{\frac{0.0241(1-0.0241)}{1000} + \frac{0.0241(1-0.0241)}{1200}}} \approx 1.93$$

The critical value (from Table A1) $z_{\alpha/2} = z_{.0250} = 1.96$.

We fail to reject the null-hypothesis. There is not sufficient evidence to warrant rejection of the claim that the two proportions are equal.

Problem 2

We use the formula for confidence interval from page 470.

$$\left(\overline{x_1}-\overline{x_2}\right)-E<\left(\mu_1-\mu_2\right)<\left(\overline{x_1}-\overline{x_2}\right)+E$$

where the margin error *E* is given by the formula

$$E=t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}.$$

The number of degrees of freedom $df=\min(n_1-1,n_2-1)=176$**.**

Table A3 contains only values of *df* equal to 100 or 200. We take the closest value $df=200$**.**

From table A3 we find $t_{\alpha/2}=1.972$**. Therefore**

$$E=1.972\sqrt{\frac{44^2+18^2}{177}}\approx 7$$

Because $\overline{x_1}-\overline{x_2}=-65$ **we have the confidence interval as**

$$-72<\overline{x_1}-\overline{x_2}<-58.$$

The correct answer is "B".

Problem 3

The null hypothesis $H_0:\quad \mu_1=\mu_2$**.**

The alternative hypothesis $H_1:\quad \mu_1<\mu_2$**. The test is left-tailed.**

The test statistic

$$t=\frac{\overline{x_1}-\overline{x_2}}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_1^2}{n_1}}}=\frac{189.1-203.7}{\sqrt{\frac{38.7^2}{85}+\frac{39.2^2}{75}}}\approx -2.365$$

The critical value from Table A3 (area in one tail 0.01, *df* = 74) is 2.377 We fail to reject the null hypothesis. There is not sufficient evidence to support the claim that the treatment population mean is smaller than the control population mean.

Problem 4

The differences in pairs are

$$-0.8, 1.8, 2.9, 0.5, -1.6, 3.9, 2.8, 0.1, 1.4$$

The mean of differences is $\overline{d} \approx 1.222$**.**

The standard deviation of differences is $s_d \approx 1.826$

The margin error (page 485) $E = t_{\alpha/2} \frac{s_d}{\sqrt{n}} = 3.355 \frac{1.826}{3} \approx 2.04$**.**

The confidence interval

$$\overline{d} - E < \mu_d < \overline{d} + E$$

or

$$-0.82 < \mu_d < 3.26$$

The correct answer is "A".

Problem 5

The null hypothesis $H_0 : \mu_d = 0$**.**

The alternative hypothesis (the claim) $H_1 : \mu_d > 0$**. The test is right-tailed.**

The differences are 7, 9, 2, -5, 14, -2, 12.

The mean $\overline{d} = \frac{7+9+2-5+14-2+12}{7} = \frac{37}{7} \approx 5.2857$**.**

The standard deviation $s_d \approx 7.0313$**.**

The test statistics $t = \frac{\overline{d} - \mu_d}{s_d} \sqrt{n} = \frac{5.2857 - 0}{7.0313} \sqrt{7} \approx 1.989$**.**

The critical value from Table A3 for the area in one tail 0.01 and 6 degrees of freedom is 3.143.

We fail to reject the null hypothesis. There is not sufficient evidence to support the claim that the diet is effective in reducing weight.

Problem 6

The null hypothesis $H_0 : \sigma_1^2 = \sigma_2^2$.

The alternative hypothesis $H_1 : \sigma_1^2 \neq \sigma_2^2$.

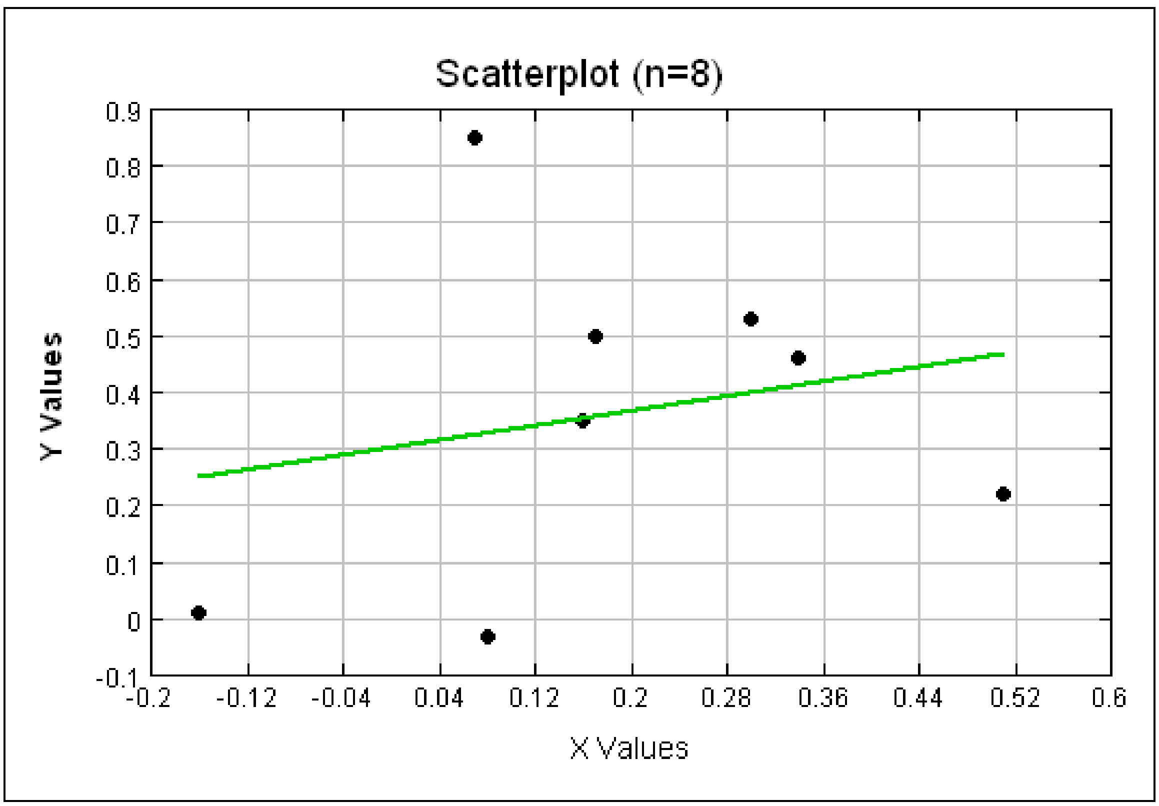
The test is two tailed.

The test statistics $F = \dfrac{s_1^2}{s_2^2} = \dfrac{5.89^2}{4.7^2} \approx 1.5705$

The critical value can be found in Table A5 on page 779. The area in one tail is 0.05, the numerator degrees of freedom is 40 and the denominator degrees of freedom is 27. The critical value is 1.8361.

We fail to reject the null hypothesis. There is not sufficient evidence to support the claim that populations A and B have different variances.

Problem 7



Above we see the scatterplot constructed by Statdisk.

The closest answer is "C".

Problem 8

We use the formula 10-1 on page 520.

$$r = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{\sqrt{n\left(\sum x^2\right) - \left(\sum x\right)^2}\sqrt{n\left(\sum y^2\right) - \left(\sum y\right)^2}} =$$

$$= \frac{7 \cdot 64636 - 395 \cdot 1149}{\sqrt{7 \cdot 22437 - 395^2}\sqrt{7 \cdot 189053 - 1149^2}} \approx -0.775.$$

The correct answer is "B".

Problem 9

We use formula 10-2 on page 542 to compute the slope of the regression line.

$$b_1 = \frac{n\left(\sum xy\right) - \left(\sum x\right)\left(\sum y\right)}{n\left(\sum x^2\right) - \left(\sum x\right)^2} = \frac{5 \cdot 1944 - 98 \cdot 69}{5 \cdot 2580 - 98^2} \approx 0.897$$

To find the y-intercept of the regression line we use formula 10-3.

$$b_0 = \overline{y} - b_1\overline{x} = \tfrac{69 - 0.897 \cdot 98}{5} \approx -3.78$$

The correct answer is "C".

Problem 10

First we compute the correlation coefficient. Because we use values x^2 instead of x formula 10-1 takes the form

$$r = \frac{n\sum x^2 y - \left(\sum x^2\right)\left(\sum y\right)}{\sqrt{n\left(\sum x^4\right) - \left(\sum x^2\right)^2}\sqrt{n\left(\sum y^2\right) - \left(\sum y\right)^2}} =$$

$$= \frac{5 \cdot 4815.201 - 161.9 \cdot 79.7}{\sqrt{5 \cdot 10472.5634 - 161.9^2}\sqrt{5 \cdot 2243.91 - 79.7^2}} \approx .990.$$

The correct answer to the first part of the problem is "C".

Next we follow the procedure described on page 527.

The null-hypothesis $H_0: \rho = 0$ (There is no linear correlation).

The alternative hypothesis $H_1: \rho \neq 0$ (There is linear correlation).

The test statistics

$$t = \frac{r}{\sqrt{(1 - r^2)/(n-2)}} = \frac{0.990}{\sqrt{0.0199 / 3}} \approx 12.1554.$$

From table A-3 we see that the critical value for two-tailed test at a significance level 0.01 is 5.841. Because the test statistics is larger than

the critical value we reject the null-hypothesis. There is a significant correlation between values of x^2 and y.

Problem 11

We will use formula 10-4 on page 559. First we need to compute the predicted values according to the formula

$$\hat{y} = 6.18286 + 4.33937x$$

The results of computations are presented in the table below.

x	9	7	2	3	4	22	17
y	43	35	16	21	23	102	81
$\hat{y}$	45.23719	36.55845	14.8616	19.20097	23.54034	101.649	79.95215

The mean of y-values is $\overline{y} = 45.85714$**.**

The explained variation is $\sum(\hat{y} - \overline{y})^2 \approx 6531.36578$**.**

The unexplained variation is $\sum(y - \hat{y})^2 \approx 13.47940$**.**

The total variation is 6531.36578 + 13.47940 = 6544.84519. The closest answer is "D".

The coefficient of determination is

$$r^2 = \frac{\text{explained variation}}{\text{total variation}} = \frac{6531.36578}{6544.84519} \approx 0.99794.$$

Problem 12

We use the formulas for prediction interval on page 561.

$$\hat{y} - E < y < \hat{y} + E,$$

$$E = t_{\alpha/2} s_e \sqrt{1 + \frac{1}{n} + \frac{n(x_0 - \overline{x})^2}{n\left(\sum x^2\right) - \left(\sum x\right)^2}} \text{ , and}$$

$$s_e = \sqrt{\frac{\sum (y - \hat{y})^2}{n-2}}.$$

First we compute the standard error of estimate. From the previous problem we know that the unexplained variation $\sum(y - \hat{y})^2$ **is 13.47940, whence**

$$s_e = \sqrt{\frac{13.47940}{5}} \approx 1.64191$$

Next from Table A-3 keeping in mind that the number of degrees of freedom is $n-2=5$ **we find that** $t_{0.005}=4.032$**.**

Finally

$x_0=12, \quad \overline{x}=9.14286, \quad \sum x=64, \ \sum x^2=932, \text{ and } \hat{y}=6.1829+4.3394\cdot 12\approx 58.3$

Plugging in these numbers we get that the margin of error is

$$E=4.032\cdot 1.64191\cdot\sqrt{1+\frac{1}{7}+\frac{7(12-9.14286)^2}{7\cdot 932-64^2}}\approx 7.1$$

The prediction interval is 51.1<y<65.4. The correct answer is "B".