Solution: Probability and Statistics (FALL 2023) Max Marks :100

Display 3 decimal places and draw neat and clean diagram where necessary

Question 01: **[5+5=10]**

a)

A stem-and-leaf plot is shown below.

Stem	Leaf	Frequency
0	22233457	8
1	023558	6
2	035	3
3	03	2
4	057	3
5	0569	4
6	0005	4

Five number summary

$Min = 0.2$

$Q_1 = 0.7$

$Q_2 = 2.15$

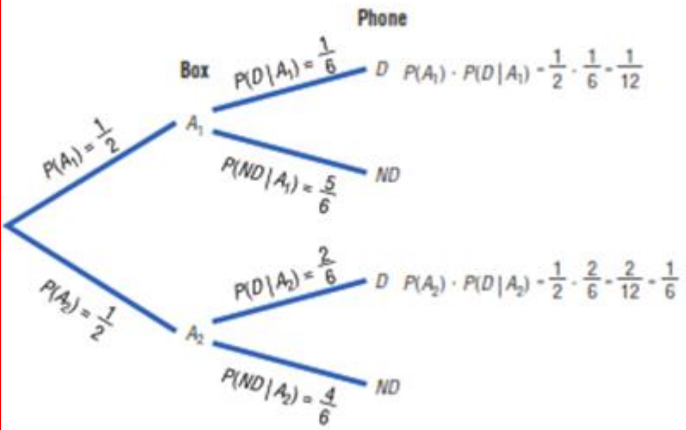
$Q_3 = 5$

$Max = 6.5$

$sample\ Mean = 2.797$

$Sample\ SD = 2.227$

b)



$$P(A_2|D) = \frac{P(A_2)\cdot P(D|A_2)}{P(A_1)\cdot P(D|A_1) + P(A_2)\cdot P(D|A_2)}$$

$$= \frac{\frac{1}{2}\cdot\frac{2}{6}}{\frac{1}{2}\cdot\frac{1}{6}+\frac{1}{2}\cdot\frac{2}{6}} = \frac{\frac{1}{6}}{\frac{1}{12}+\frac{2}{12}} = \frac{\frac{1}{6}}{\frac{3}{12}}$$

$$= \frac{1}{6} \div \frac{3}{12} = \frac{1}{\cancel{6}_{1}}\cdot\frac{\cancel{12}^{2}}{3} = \frac{2}{3}$$

Question 02: [8+3+6+3=20]

a)

By definition of the marginal density. for $0 < x < 2$,

$$g(x) = \int_{-\infty}^{\infty} f(x,y)\, dy = \int_0^1 \frac{x(1+3y^2)}{4} dy$$
$$= \left(\frac{xy}{4} + \frac{xy^3}{4}\right)\Big|_{y=0}^{y=1} = \frac{x}{2},$$

and for $0 < y < 1$,

$$h(y) = \int_{-\infty}^{\infty} f(x,y)\, dx = \int_0^2 \frac{x(1+3y^2)}{4} dx$$
$$= \left(\frac{x^2}{8} + \frac{3x^2y^2}{8}\right)\Big|_{x=0}^{x=2} = \frac{1+3y^2}{2}.$$

$$\sigma_{XY} = E(XY) - \mu_X\mu_Y.$$

$$E(x) = \int_0^2 \frac{x^2}{2} dx = \frac{4}{3}\ , \qquad E(y) = \int_0^1 \frac{y+3y^3}{2} dy = \frac{5}{8}\ ,$$

$$E(xy) = \int_0^1 \int_0^2 xy\, f(x,y)\, dxdy = \frac{5}{6}\ ,$$

$$covarriance\ of\ xy\ variable \qquad \sigma_{xy} = \frac{5}{6} - \frac{4}{3}\left(\frac{5}{8}\right) = 0$$

Therefore, the correlation coefficient between X and Y is

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y}$$

$$\boxed{\rho = 0}$$

b)

Using the geometric distribution

(a) $P(X = 3) = g(3; 0.7) = (0.7)(0.3)^2 = 0.0630.$

(b) $P(X < 4) = \sum_{x=1}^{3} g(x; 0.7) = \sum_{x=1}^{3} (0.7)(0.3)^{x-1} = 0.9730.$

c)

(a) $z = (10.075 - 10.000)/0.03 = 2.5$; $P(X > 10.075) = P(Z > 2.5) = 0.0062$.

Therefore, 0.62% of the rings have inside diameters exceeding 10.075 cm.

(b) $z_1 = (9.97 - 10)/0.03 = -1.0$, $z_2 = (10.03 - 10)/0.03 = 1.0$;
$P(9.97 < X < 10.03) = P(-1.0 < Z < 1.0) = 0.8413 - 0.1587 = 0.6826$.

d)

(a) $\mu = np = (1875)(0.004) = 7.5$, so $P(X < 5) = P(X \leq 4) \approx 0.1321$.

(b) $P(8 \leq X \leq 10) = P(X \leq 10) - P(X \leq 7)$

$\approx 0.8622 - 0.5246 = 0.3376$.

Question 03: [10+5+10=25]

a)

Before (X_1)	After (X_2)	$D = X_1 - X_2$
210	190	20
235	170	65
208	210	−2
190	188	2
172	173	−1
244	228	16
		$\Sigma D = 100$

Find the standard deviation of the differences.

$$s_D = \sqrt{\frac{n\Sigma D^2 - (\Sigma D)^2}{n(n-1)}} = 25.4$$

$$\overline{D} = \frac{\Sigma D}{n} = \frac{100}{6} = 16.7$$

$$\overline{D} - t_{\alpha/2}\frac{s_D}{\sqrt{n}} < \mu_D < \overline{D} + t_{\alpha/2}\frac{s_D}{\sqrt{n}}$$
$$16.7 - 2.015 \cdot \frac{25.4}{\sqrt{6}} < \mu_D < 16.7 + 2.015 \cdot \frac{25.4}{\sqrt{6}}$$
$$16.7 - 20.89 < \mu_D < 16.7 + 20.89$$
$$-4.19 < \mu_D < 37.59$$

t alpha /2 = 2.571 , [-9.98 , 43.316]

b)

$n_1 = 12, n_2 = 18, \bar{x}_1 = 84, \bar{x}_2 = 77, s_1 = 4, s_2 = 6$, and $s_p = 5.305$

$t_{0.005} = 2.763$ with 28 degrees of freedom.

$$(84 - 77) \pm (2.763)(5.305)\sqrt{1/12 + 1/18} = 7 \pm 5.46,$$

$$1.54 < \mu_1 - \mu_2 < 12.46.$$

c)

using calculator

$\overline{X}_1 = 11.8$ $\overline{X}_2 = 3.8$ $\overline{X}_3 = 7.6$

$s_1^2 = 5.7$ $s_2^2 = 10.2$ $s_3^2 = 10.3$

Source	Sum of squares	d.f.	Mean square	F
Between	160.13	2	80.07	9.17
Within (error)	104.80	12	8.73	
Total	264.93	14		

Since $k = 3$ and $N = 15$,

d.f.N. $= k - 1 = 3 - 1 = 2$

d.f.D. $= N - k = 15 - 3 = 12$

The critical value is 3.89, obtained from Table with $\alpha = 0.05$.

Make the decision. The decision is to reject the null hypothesis, since $9.17 > 3.89$.

Summarize the results. There is enough evidence to reject the claim and conclude that at least one mean is different from the others.

Question 04: [5+10=15]

a)

Solution

Step 1 State the hypotheses and identify the claim

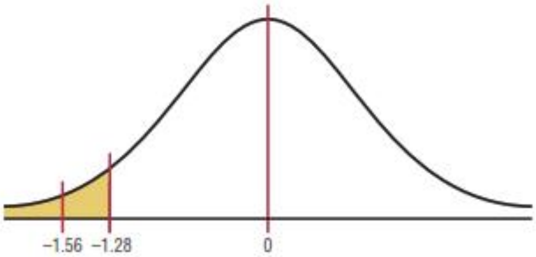
H_0: $\mu = \$80$ and H_1: $\mu < \$80$ (claim)

Step 2 Find the critical value. Since $\alpha = 0.10$ and the test is a left-tailed test, the critical value is -1.28.

Step 3 Compute the test value. Since the exercise gives raw data, it is necessary to find the mean of the data. Using the formulas in Chapter 3 or your calculator gives $\overline{X} = 75.0$ and $\sigma = 19.2$. Substitute in the formula

$$z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{75 - 80}{19.2/\sqrt{36}} = -1.56$$

Step 4 Make the decision. Since the test value, −1.56, falls in the critical region, the decision is to reject the null hypothesis.



Step 5 Summarize the results. There is enough evidence to support the claim that the average cost of men's athletic shoes is less than $80.

b)

Solution

Step 1 State the hypotheses and identify the claim.

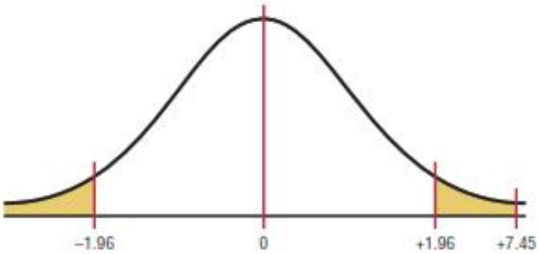
H_0: $\mu_1 = \mu_2$ and H_1: $\mu_1 \neq \mu_2$ (claim)

Step 2 Find the critical values. Since $\alpha = 0.05$, the critical values are +1.96 and −1.96.

Step 3 Compute the test value.

$$z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(88.42 - 80.61) - 0}{\sqrt{\frac{5.62^2}{50} + \frac{4.83^2}{50}}} = 7.45$$

Step 4 Make the decision. Reject the null hypothesis at $\alpha = 0.05$, since $7.45 > 1.96$.



Step 5 Summarize the results. There is enough evidence to support the claim that the means are not equal. Hence, there is a significant difference in the rates.

Question 05: [5+15+10=30]

a) Quadratic Modal by using calculator

$b_0 = 270.525\,, \quad b_1 = 7.949\,, \quad b_2 = -0.067$

$$\mu_{Y|35} = 270.525 + 7.949(35) - 0.067(35^2) = 466.665$$

b)

	y	x1	x2	x1y	x2y	x1x2	x1^2	x2^2
	193	1.6	851	308.8	164243	1361.6	2.56	724201
	172	22	1058	3784	181976	23276	484	1119364
	113	33	1357	3729	153341	44781	1089	1841449
	230	15.5	816	3565	187680	12648	240.25	665856
	91	43	1201	3913	109291	51643	1849	1442401
	125	40	1115	5000	139375	44600	1600	1243225
Total	924	155.1	6398	20299.8	935906	178309.6	5264.81	7036496

Normal equation of Multiple Regression:

$$\boldsymbol{y = a + bx_1 + cx_2}$$

$$\sum y = na + b\sum x_1 + c\sum x_2$$

$$\sum x_1 y = a\sum x_1 + b\sum x_1{}^2 + c\sum x_1 x_2$$

$$\sum x_2 y = a\sum x_2 + b\sum x_1 x_2 + c\sum x_2{}^2$$

Substitute values from table in normal equation and Form the linear equation and solve via calculator for multiple linear coefficient

$$\boldsymbol{b_0 = 350.99427\,, \quad b_1 = -1.27199\,, \quad b_2 = -0.15390}$$

$$\boxed{\hat{y} = 350.99427 - 1.27199x_1 - 0.15390x_2.}$$

$$\boxed{\hat{y} = 350.99427 - (1.27199)(20) - (0.15390)(1200) = 140.86930.}$$

$$\boldsymbol{r_{x_1 y} = -0.851}$$

$$\boldsymbol{r_{x_1 x_2} = 0.788}$$

$$\boldsymbol{r_{x_2 y} = -0.898}\ ,$$

$$\boldsymbol{R = 0.9277} \quad , \quad \boldsymbol{R^2 = 0.8606}$$

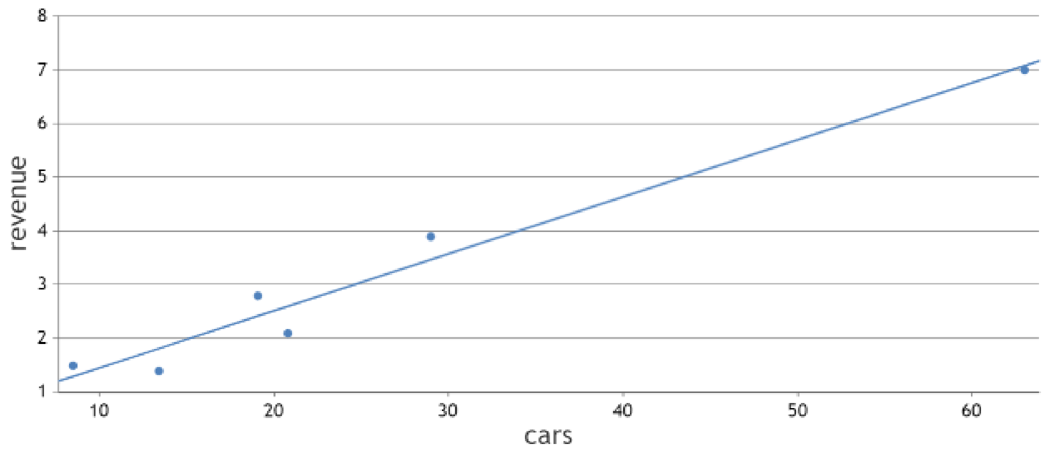
C)

$n = 6, \Sigma x = 153.8, \Sigma y = 18.7, \Sigma xy = 682.77,$ and $\Sigma x^2 = 5859.26.$

$$a = \frac{(\Sigma y)(\Sigma x^2) - (\Sigma x)(\Sigma xy)}{n(\Sigma x^2) - (\Sigma x)^2} = \frac{(18.7)(5859.26) - (153.8)(682.77)}{(6)(5859.26) - (153.8)^2} = 0.396$$

$$b = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2} = \frac{6(682.77) - (153.8)(18.7)}{(6)(5859.26) - (153.8)^2} = 0.106$$

$$y' = 0.396 + 0.106x$$



Cars x (in 10,000s)	Revenue y (in billions)	xy	x^2	y^2
63.0	7.0	441.00	3969.00	49.00
29.0	3.9	113.10	841.00	15.21
20.8	2.1	43.68	432.64	4.41
19.1	2.8	53.48	364.81	7.84
13.4	1.4	18.76	179.56	1.96
8.5	1.5	12.75	72.25	2.25
$\Sigma x = 153.8$	$\Sigma y = 18.7$	$\Sigma xy = 682.77$	$\Sigma x^2 = 5859.26$	$\Sigma y^2 = 80.67$

$$r = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{\sqrt{[n(\Sigma x^2) - (\Sigma x)^2][n(\Sigma y^2) - (\Sigma y)^2]}}$$

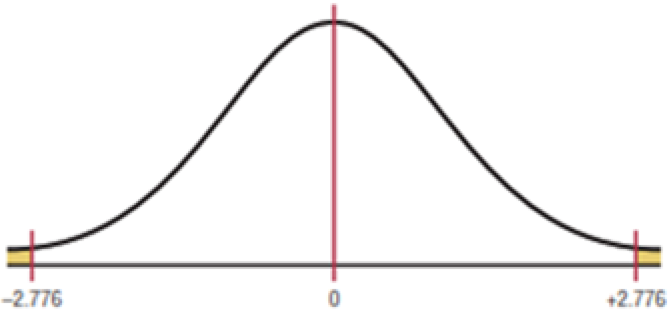
$$= \frac{(6)(682.77) - (153.8)(18.7)}{\sqrt{[(6)(5859.26) - (153.8)^2][(6)(80.67) - (18.7)^2]}} = 0.982$$

The correlation coefficient suggests a strong relationship between the number of cars a rental agency has and its annual revenue.

Step 1 State the hypotheses.

$H_0: \rho = 0$ and $H_1: \rho \neq 0$

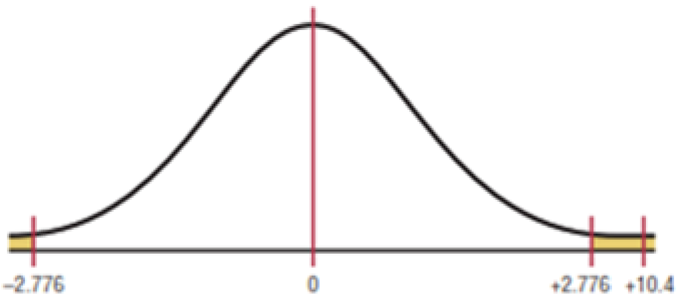
Step 2 Find the critical values. Since $\alpha = 0.05$ and there are $6 - 2 = 4$ degrees of freedom, the critical values obtained from Table are ± 2.776,



Step 3 Compute the test value.

$$t = r\sqrt{\frac{n-2}{1-r^2}} = 0.982\sqrt{\frac{6-2}{1-(0.982)^2}} = 10.4$$

Step 4 Make the decision. Reject the null hypothesis, since the test value falls in the critical region,



Step 5 Summarize the results. There is a significant relationship between the number of cars a rental agency owns and its annual income.

---------THE END ---------