

STATISTICS FOR BUSINESS DECISIONS

PAST QUESTIONS

QUIZ2

1. The assumption of one way ANOVA include all the following ‘All populations are normally distributed, variance are equal, observations are independent’ except data are normal’.

A. True B. False

Solution

True (refer to page 171)

Use the following information to answer questions 2 and 3

Source of variation	SS	df	MS	F -Ratio
Between samples	22.00	3	7.33	y
Within samples	x	28	7.10	
Totals	220.88	31		

2. From the ANOVA results given above, find the value of x
 A. 189.88 B. 165.88 C. 198.88 D. 188.89

Solution

$$22+X = 220.88 \Rightarrow x = 220.88 - 22 = 198.88. \text{ Ans.C}$$

3. Compute the value of y in the table given above.
 A. 1.333 B. 3.03 C. 2.03 D. 1.03

Solution

$$F = \frac{MSB}{MSW} = y = \frac{7.33}{7.10} = 1.0323943662 = 1.03. \text{ Ans D.}$$

Use the following information to answer questions 4 and 5.

Source	size		Total
	small	large	
White	276	3	279
Black	473	66	539
Total	749	69	818

4. Find the expected count e11 from the data given above.
 A. 245.60 B. 256.78 C. 255.47 D. 256.01

Solution

Visit archornews.com for great content in mathematics.

$$e_{ij} = \frac{(i^{\text{th}} \text{ row total})(j^{\text{th}} \text{ column total})}{\text{total sample size}}$$

$$e_{11} = \frac{276 \times 749}{818} = 255.47 \text{ .Ans.C}$$

5. Using the data above, compute the statistic X^2 .

A. 29.00 B. 28.93 C. 29.01 D. 29.68

Solution

$$e_{12} = \frac{279 \times 69}{818} = 23.53 e_{21} = \frac{539 \times 749}{818} = 493.53 e_{22} = \frac{539 \times 69}{818} = 45.47$$

$$X^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(o_{ij} - e_{ij})^2}{e_{ij}}$$

$$X^2 = \frac{(o_{11} - e_{11})^2}{e_{11}} + \frac{(o_{12} - e_{12})^2}{e_{12}} + \frac{(o_{21} - e_{21})^2}{e_{21}} + \frac{(o_{22} - e_{22})^2}{e_{22}}$$

$$X^2 = \frac{(276 - 255.47)^2}{255.47} + \frac{(3 - 23.53)^2}{23.53} + \frac{(473 - 493.53)^2}{493.53} + \frac{(66 - 45.47)^2}{45.47} = 29.68. \text{ Ans.D}$$

6. A radio station publicizes that its proportion of local listening audience is greater than or equal to 39%. Find the mathematical expression of this hypothesis.

A. $H_0: p = 0.39$, $H_A: p \neq 0.39$

B. $H_0: P \geq 0.39$, $H_A: p < 0.39$

C. $H_0: \mu \geq 0.39$, $H_A: \mu < 0.39$

D. $H_0: \mu = 0.39$, $H_A: \mu \neq 0.39$

Solution

$H_0: p \geq 0.39$

$H_A: p < 0.39. \text{ Ans.B}$

7. Results of analysis of students preference level in respect to programmes of study gave us p-value = 0.0287. Hint: $\alpha = 0.05$. In conclusion. We fail to reject the null hypothesis.

A. True B. false

Solution

If P-value $< \alpha$, we reject the null hypothesis.

If the p-value $\geq \alpha$, we failed to reject the null hypothesis.

The p-value of 0.0287 is less than($<$) than $\alpha = 0.05$. We reject the null hypothesis. **True**

Use the following information to answer questions 8 to 10.

Group A	GroupB
$n_1 = 15$	$N_2 = 15$
$X_1 = 2119.7$	$X_2 = 1777.7$
$S_1 = 709.7$	$S_2 = 593.9$

8. Compute the t statistic for the two groups

A. 0.4313 B. 2.3100 C. 1.4313 D. 1.1233

Solution

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (u_1 - u_2)}{sp \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

but $sp = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} = \sqrt{\frac{(15-1)709.7^2 + (15-1)1777.7^2}{15+15-2}} =$

$$\sqrt{\frac{7051437.26 + 44243042.06}{28}} = \sqrt{1831945.69} = 654.37$$

$$t = \frac{(2119.7 - 1777.7) - (0 - 0)}{654.37 \sqrt{\frac{1}{15} + \frac{1}{15}}} = \frac{342}{238.97865476} = 1.4313. \text{ Ans.C}$$

9. Compute the population standard deviation s_p for the two groups.
 A. 754.37 B. 600.37 C. 654.37 D. 854.37

Solution

$$S_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} = \sqrt{\frac{(15-1)709.7^2 + (15-1)593.9^2}{15+15-2}} =$$

$$\sqrt{\frac{7051437.26 + 4938040.94}{28}} = \sqrt{428195.65} = 654.37. \text{ Ans. C}$$

10. Given the critical value of 2.0484, what is the conclusion?
 A. The two groups has the same mean value B. The two groups do not have the same mean value
 C. The two groups do not exist D. None of the above.

Solution

Hypothesis

$$H_0: u_1 = u_2 \text{ or } H_0: u_1 - u_2 = 0$$

$$H_A: u_1 \neq u_2 \text{ or } H_A: u_1 - u_2 \neq 0$$

Decision rule: Reject if test statistic is greater than critical region

OR: fail to reject if test statistic is less than critical value.

Test statistic = 1.4313

Critical value = 2.0484

Since $1.4313 < 2.0484$, we **fail to reject**, meaning the two groups don't have the same mean value. **Ans.B**

Use the following to answer questions 11 to 13.

Given the following information $\bar{x} = 40.78$, $\sigma = 12.6$ and a total population $n=100$.

11. Compute the confidence interval estimate for 99%.
 A. (33.54 to 44.02) B. (36.54 to 44.02) C. (37.54 to 42.02) D. None of the above.

Solution

$$\text{Confidence interval} = \bar{x} \pm z \frac{\sigma}{\sqrt{n}} = 40.78 + 2.575 \left(\frac{12.6}{\sqrt{100}} \right) =$$

$$40.78 \pm 3.24 = 40.78 - 3.24 \text{ or } 40.78 + 3.24 =$$

37.54 to 44.02. refer to page 124 of text book. **Ans.D**

12. Compute the confidence interval estimate for 90%.

- A.(36.71,42.02) B.(39.71, 45.85) C.(38.71, 42.85) D.None of the above.

Solution

$$\text{Confidence interval} = \bar{x} \pm z \frac{\sigma}{\sqrt{n}} = 40.78 + 1.26 \left(\frac{12.6}{\sqrt{100}} \right) = 40.78 \pm 2.07 = \\ 40.78 - 2.07 \text{ or } 40.78 + 2.07 = 38.71 \text{ to } 42.85. \text{ Ans.C}$$

13. Find the standard error
A. 0.36 B. 2.16 C. 1.26 D.0.40

Solution

$$\text{Standard error} = \frac{\sigma}{\sqrt{n}} = \left(\frac{12.6}{\sqrt{100}} \right) = 1.26. \text{ Ans.C}$$

14. A random sample of 100 footballers shows that 25 are left-footed. Given that z=1.96, find the 95% confidence interval for the true proportion of the left-footers.
A.(0.320,0.335) B.(0.165,0.335) C.(0.160,0.450) D.(0.600,0.332)

$$\text{Confidence interval} = \bar{p} \pm z \sqrt{\frac{\bar{p}(1-\bar{p})}{n}} = \frac{25}{100} \pm 1.96 \sqrt{\frac{0.25(1-0.25)}{100}} = \\ 0.25 \pm 0.085 = 0.25 - 0.085 \text{ or } 0.25 + 0.085 =$$

0.165 to 0.335. **Ans.B.** refer to page 130 of the text.

$$\text{Note a proportion}(\bar{p}) = \frac{x}{n} = \frac{25}{100} = 0.25$$

15. A consumer analyst reports that mean life of a certain type of automobile battery is not 50 months. Find the mathematical expression of the hypothesis.
A. Ho: $\mu = 50$, HA: $\mu \neq 50$ B. Ho: $p = 50$, HA: $p \neq 50$
C. Ho: $\mu \leq 50$, HA: $\mu > 50$ D. Ho: $p \leq 50$, HA: $p > 50$

Solution

Ans. A. Ho: $\mu = 50$, HA: $\mu \neq 50$. this is a two tail test. Note the key word NOT. Refer to page 133 of the text.

16. The sampling error is a computed measure of difference between a statistic and a parameter.
A. True B. False.

Solution

True. Refer to page 115 of the text.

17. If the population mean is $\mu= 98.6$ degrees and the computed measure of difference in mean is 0.6 degrees, then the sample mean is 99.5 degrees.
A. True B. False

Solution

$99.5 - 98.6 \neq 0.6$. **False.** Refer to page 115 of the text.

18. When the population is irregularly shaped, the required sample size which is defined to be sufficiently large is $n \geq 20$.
A. True B. False

Solution

False. It is $n \geq 30$. Refer to page 119 of the text.

19. A group of students were asked whether they are able to use Minitab to analyse data. Out of a total of 200, 90 answered “Yes” and the rest answered “No”. Find the proportion of students that answered “No”.
 A. 10% B. 55% C. 45% D. None of the above.

Solution

Note a proportion(\bar{p}) = $\frac{x}{n} = \frac{(200-90)}{200} = \frac{110}{200} = 0.55 = 55\%$. **Ans.B**

20. Population parameter that is estimated from sample statistic is known as
 A. Confidence level estimate B. Point estimate C. Confidence interval estimate D. None of the above.

Solution

Ans.B. refer to page 123 of the text.

END OF SEMESTER EXAMINATION FOR 2018/2019

1. Two-dimensional chart showing time on the horizontal axis and the variable of interest on the vertical axis is
 A. Pie chart B. Bar Chart C. Line chart D. None of the above.

Solution

Ans.C

Use the information to answer questions 2 to 4.

Given that x numbers are 568, 570, 575, 578, 584, and the y numbers are 573, 574, 575, 577, 578, and they are the inner diameter of two grades of tires.

index	x	X ²	Y	Y ²
1	568	322624	573	328329
2	570	324900	574	329476
3	575	330625	575	330625
4	578	334084	577	332929
5	584	341056	578	334084
Totals of columns	2875	1653289	2877	1655443

2. Find the mean of grade X.
 A. 575.00 B. 577.00 C. 579.00 D. 557.00

Solution

$$\text{Mean} = \frac{\sum x}{n} = \frac{568+570+575+578+584}{5} = 575. \text{ Ans.A}$$

3. Find the variance of grade x.
 A. 40.00 B. 41.00 C. 43.00 D. 39.20

Solution

$$\text{Variance} = \frac{\sum (xi - \mu)^2}{n-1} = \frac{(568-575)^2 + (570-575)^2 + (575-575)^2 + (578-575)^2 + (584-575)^2}{4} = 41$$

Or

$$\text{Variance} = \frac{\sum x^2 - n\bar{x}^2}{n-1} = \frac{1653289 - 5(575)^2}{5-1} = \frac{164}{4} = 41$$

4. Compute the standard deviation of y.
 A. 2.07 B. 2.33 C. 3.00 D. 2.87

Solution

$$\text{Mean of } y = \frac{\sum y}{n} = \frac{573+574+575+577+578}{5} = \frac{2877}{5} = 575.4$$

$$\text{Standard deviation} = \sqrt{\frac{\sum y^2 - n\bar{y}^2}{n-1}} = \sqrt{\frac{1655443 - 5(574.5)^2}{4}} = \sqrt{4.3} = 2.07. \text{ Refer to page 68 of the text.}$$

Or

$$\sqrt{\left(\frac{(573-575.4)^2 + (574-575.4)^2 + (575-575.4)^2 + (577-575.4)^2 + (578-575.4)^2}{4}\right)} = 2.07. \text{ Ans.A}$$

5. Suppose you are given a 3 question multiple –choice test. Each question has 4 responses and only one is correct. Suppose you want to find the probability that you can just guess the answers and get 2 questions right. This is typical of a Experiment.

- A. Hypergeometric B. geometric C. Binomial D. Poisson

Solution

Ans.C. Refer to page 94 of the text.

6. A Scale is a measurement, in which numbers serve as tags or labels only to identify or classify an object.
 A. Ordinal B. Nominal C. ratio D. None of the above.

Solution

Ans.B. Page 13.

Use the following information to answer questions 7 to 9.

The bivariate joint distribution are often presented as Two –way contingency table with marginal frequencies.

	Dance	Sports	TV	Total
Men	2	10	8	20
Women	16	6	8	30
Total	18	16	16	50

7. From the table above, compute the joint total relative frequency of women.
 A. 0.40 B. 0.65 C. 0.60 D. 0.40

Solution

	Dance	Sports	TV	Total
Men	2/50 =0.04	10/50=0.2	8/50=0.16	20/50=0.4
Women	16 /50 =0.32	6/50=0.12	8/50=0.16	30/50=0.6
Total	18/50=0.36	16/50=0.32	16/50=0.32	50/50= 1

Total for women = 30

Joint relative frequency of total women = $30/50 = 0.6$. **Ans.C**

8. Compute the joint total relative frequency of dancers given the table table above.
A. 0.36 B. 0.32 C. 0.60 D.0.40

Solution

Total frequency of dancers = 18. Relative total frequency of dancers = $18/50 = 0.36$.

Ans.A

9. Find the joint relative frequency of men in TV
A. 0.32 B.0.12 C.0.16 D. 0.60

Solution

Total relative frequency of men in TV = $8/50 = 0.16$. **Ans.C**

10. A group of students were asked whether they are able to use Minitab to analyze data. Out of a total of 200, 90 answered “Yes” and the rest answered “No”. Find the proportion of students that answered “No”.
A. 10% B. 55% C. 45% D.None of the above.

Solution

Proportion of no = $110/200 = 0.55$ or 55%. **Ans.B**

11. A random sample of 100 footballers shows that 35 are left-handed. Find the 95% confidence interval for the true proportion of the left-handers, given that z-value for 95% confidence level is 1.96.
A.(0.403,0.207) B.(0.043,0.257) C.(0.443,0.257) D.(0.043,0.257) D.(0.443,0.057)

Solution

$$\text{Confidence interval of proportion} = \bar{p} \pm z \sqrt{\frac{\bar{p}(1-\bar{p})}{n}} = 0.35 \pm 1.96 \sqrt{\frac{0.35(1-0.35)}{100}} = \\ 0.35 - 0.093 \text{ or } 0.35 + 0.093 \text{ is } (0.443 \text{ to } 0.257)$$

12. Which one of the following tools is used to measure multi-collinearity in related variables.
A.Variance inflation factor B. Adjusted R-square C. Coefficeint of variation D. None of the baove.

Solution

Ans.A. Refer to page 219 of the text

13. If the population mean is $\mu = 98.6$ degrees and the computed measure of differences in mean is 0.6 degrees, then the sampl mean is 99.5 degrees.
A. True B. False

Solution

False. Because $99.5 - 98.6 \neq 0.6$

14. Highly correlated independent variables included in a regression model can never adversely affect the regression results.
A. True B.False

Solution

False. They can adversely affect the regression results. Refer to page 219 of the text.

Visit archornews.com for great content in mathematics.

15. It is known that 10% of Parliamentarians voted for a particular policy implementation. There are only 273 parliamentarians, so the researcher can count what every single one of them voted. This statistical measure is statistic.

A. True B. False

Solution

True, because a statistic is a measure computed from the population parameter. Page 115 of the text.

Use the following information to answer questions 16 to 18.

Given the information $\bar{x} = 40.78$, $\sigma = 12.6$ and a total population $n=100$.

16. Compute the confidence interval estimate for 99%.
A.(37.54 to 44.02) B.(36.54 to 44,02) C. (37.54 to 42.02) D.None of the above.

Solution

$$\text{Confidence interval} = \bar{x} \pm z \frac{\sigma}{\sqrt{n}} = 40.78 + 2.575 \left(\frac{12.6}{\sqrt{100}} \right) = 40.78 \pm 3.24 =$$

$40.78 - 3.24$ or $40.78 + 3.24 = 37.54$ to 44.02 . refer to page124 of text book.

Ans.A

17. Compute the confidence interval estimate for 90%.
A.(36.71,42.05) B.(39.71,45.85) C.(38.71, 42.85) D. None of the above

Solution

$$\text{Confidence interval} = \bar{x} \pm z \frac{\sigma}{\sqrt{n}} = 40.78 + 1.26 \left(\frac{12.6}{\sqrt{100}} \right) = 40.78 \pm 2.07 =$$

$40.78 - 2.07$ or $40.78 + 2.07 = 38.71$ to 42.85 . **Ans.C**

18. Find the standard error
A. 0.36 B. 2.16 C.1.26 D.0.40

Solution

$$\text{Standard error} = \frac{\sigma}{\sqrt{n}} = \left(\frac{12.6}{\sqrt{100}} \right) = 1.26. \text{ **Ans.C**}$$

19. Given that accidents occur with Poisson distribution at an average of 4 per week. What is the probability that at least two weeks will elapse between accidents?
A. 3.4×10^{-3} B. 2.4×10^{-4} C. 3.4×10^{-4} D. 3.4×10^{-2}

Solution

$$P(x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!} = (\lambda = 4, x \geq 2)$$

20. A data expressing a certain quantity, amount or range is..... Usually, there is measurement units associated with the data.
A. Qualitative data B. Quantitative data C. Cross-sectional data D. Categorical data.

Solution

Ans.B. Refer to page 11 of the text.

THEORY

- 1.a) A doctor believes that the number of births in Ghana on each day of the week is equal. A simple random sample of 700 births from a recent year is selected, and the results are below. At a significance level of 0.01, is there enough evidence to support the doctor's claim?

Day	Total number of Births Count	
	Observe O _i	Expected e _i
Sunday	4,502	6,461.54
Monday	6,623	8,615.38
Tuesday	8,308	8,615.38
Wednesday	10,420	8,615.38
Thursday	11,032	8,615.38
Friday	10,754	8,615.38
Saturday	4,361	6,461.54

Solution

H₀: The customer demand distribution is evenly spread through the weekdays and is 25% lower on the weekend

H₁: The customer demand follows some other distribution

Day	Total customer count		O _i -e _i	(o _i -e _i) ²	$\frac{(o_i - e_i)^2}{e_i}$
	Observed(oi)	Expected (ei)			
Sunday	4502	6,461.54	-1,959.54	3839797.012	594.254
Monday	6,623	8,615.38	-1,992.38	3969578.064	460.755
Tuesday	8,308	8,615.38	-307.38	94482.4644	10.967
Wednesday	10,420	8,615.38	1,804.62	3256653.344	378.005
Thursday	11,032	8,615.38	2,416.62	5840052.224	677.864
Friday	10,754	8,615.38	2,138.62	4573695.504	530.876
saturday	4,361	6,461.54	-2,100.54	4412268.2925	682.851
Total	56,000	56000			3,3335.6

Test statistic = $X^2 = \sum_{i=1}^k \frac{(o_i - e_i)^2}{e_i} = \sum_{i=1}^7 \frac{(o_i - e_i)^2}{e_i} = 3,3335.6$ as seen in the table.

Test statistic = $X^2 = 3,3335.6$, Critical value or table value $X^2_{0.05(6)} = 12.5916$

Decision rule: reject null hypothesis if X^2 calculated is greater than X^2 table

Since $X^2 = 3,3335.6 >$ than $X^2_{0.05(6)} = 12.5916$, we reject the null hypothesis and conclude that the distribution is not the same as previously indicated.

- b) 279 Males and 539 Females UCC students were asked about their drink preferences in terms of size of the bottle. Among the Males, 276 preferred small and 3 preferred large size. Of the Females, 473 preferred small and 66 preferred large size. The summary is given as follows.

Sex	Size	
	small	large
Male	276	3
Female	473	66
	749	818

Required

Find if sex is related to drink size preference, using 5% significance level.

Solution

H_0 : Sex is independent of size

H_1 : Sex is not independent of size

$$e_{11} = \frac{276 \times 749}{818} = 255.47$$

$$e_{12} = \frac{279 \times 69}{818} = 23.53 \quad e_{21} = \frac{539 \times 749}{818} = 493.53$$

$$e_{22} = \frac{539 \times 69}{818} = 45.47.$$

$$X^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(o_{ij} - e_{ij})^2}{e_{ij}}$$

$$X^2 = \frac{(o_{11} - e_{11})^2}{e_{11}} + \frac{(o_{12} - e_{12})^2}{e_{12}} + \frac{(o_{21} - e_{21})^2}{e_{21}} + \frac{(o_{22} - e_{22})^2}{e_{22}}$$

$$X^2 = \frac{(276 - 255.47)^2}{255.47} + \frac{(3 - 23.53)^2}{23.53} + \frac{(473 - 493.53)^2}{493.53} + \frac{(66 - 45.47)^2}{45.47} = 29.68. \text{ Ans.D}$$

2. a) A critical components of a handled hair dryer is the motor-heater unit. The company has recently created a new motor-heater unit with fewer parts than the current unit. Suppose the following information is given where new unit: $n_1=250$, $x_1=55$ and old unit: $n_2=250$, $x_2=75$. Determine the motor-heater which is more reliable using samples of old and new units.

Solution

The null hypothesis : The new motor heater is better than the old one

Alternative hypothesis : The new one is no better than the old one

Mathematically

$$H_0: \bar{p}_1 - \bar{p}_2 \geq 0 \quad H_A: \bar{p}_1 - \bar{p}_2 < 0,$$

$$\alpha = 0.05, -z_{0.05} = -1.645$$

$$\bar{p}_1 = \frac{55}{250} = 0.22, \bar{p}_2 = \frac{75}{250} = 0.30$$

$$\text{But proportion for two populations} = \frac{n_1 \bar{p}_1 + n_2 \bar{p}_2}{n_1 + n_2} = \frac{x_1 + x_2}{n_1 + n_2} =$$

$$\frac{250(0.22) + 250(0.30)}{250+250} = \frac{55+75}{500} = 0.26$$

Since the two samples are with characteristics of interest, we use the Z-test statistic for difference between population and it is given below.

Visit archornews.com for great content in mathematics.

$$Z = \frac{(\bar{p}_1 - \bar{p}_2) - (\bar{p}_1 - \bar{p}_2)}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.22 - 0.3) - 0}{\sqrt{0.26(1-0.26)x\left(\frac{1}{250} + \frac{1}{250}\right)}} = -2.04$$

Decision rule: we reject the null hypothesis if Z-calculated is less than Z read from the table.

Z-calculated = -2.04, Z-table = -1.645

We therefore reject the null hypothesis and conclude that the new motor heater is no better than the old one.

- b) A professor has developed three different mid-term exams that are to be graded on a 1000 point scale. Before she uses the exams in a live class, she wants to determine if the tests will yield the same mean scores. To test this, a random sample of fourteen people is selected, categorized according to their level of understanding so that each student takes each test. From the results of the analysis given below.

Source of Variation	SS	df	MS	F-Ratio	F-critical
Between blocks	116,605.00	13	X2	0.9105	2.15
Between samples	X1	2	120,956.40	X3	3.4
Within Samples	256,123.90	26	9,850.90		
Total	614,641.61	41			

Required

- Find the missing values X1, X2 and X3.
- State the hypothesis of the problem.
- Draw your conclusion

Solution

- $SST = SSB + SSBL(X1) + SSW$
 $614,641.61 = 116,605 + X1 + 256,123.90$
 $X1 = 614,641.61 - 372,728.9 = \mathbf{24,912.71}$

$$X2 = MSB = \frac{SSB}{k-1} = \frac{116,605}{13} = 8,969.6$$

$$X3 = \frac{MSB}{MSW} = \frac{120,956.4}{9850.9} = 12.2787$$

- Main hypothesis

$$H_0: \mu_1 = \mu_2 = \mu_3$$

H_A : At least two of the populations are different

Secondary hypothesis

$$H_0: \mu_{b1} = \mu_{b2} = \mu_{b3} = \dots = \mu_{bb}$$

H_A : Not all block means are equal (blocking is effective).

- Decision rule: reject null hypothesis F-calculated more than F-table

F-calculated = 0.91, F-table = 2.15

Since $F=0.9105 < F_{0.05} = 2.15$; we fail to reject the null hypothesis and conclude that blocking was not effective

Also since $F= 12.2787 > F_{0.05} = 3.40$, we reject the null hypothesis and conclude that the tests have the same mean.

Note: we are using the **F-ratio** and the **F-critical** in the table for the conclusion.

The First F-value for the block effects and the second F-value to determine the equality of the means. Refer to pages 178 and 181 of the text.

- 3.a) In a certain test, 5 out of 20 students scored an 'A'. We chose three students at random out of the 20 students without replacement. Find the probability that all three are the ones who scored an 'A'.

Solution

This is a self-assessment question(3.1).

$$\text{Probability of first student with grade A} = \frac{5}{20}$$

Probability of second student with grade A = $\frac{4}{19}$ (because grade one students reduced to 4 and total number of students also reduces to 19 since the first student is not replaced).

$$\text{Probability of picking third with grade A} = \frac{3}{18}$$

Probability that all three students scored grade A = $\frac{5}{20} \times \frac{4}{19} \times \frac{3}{18} = \frac{1}{114}$. Refer to further explanations in the text in page 228.

- b) In the Arizona lottery called pick 3, a player pays GHS1 and then picks a three-digit number. If those three numbers are picked in right or specific order the person wins GH₵500.00

Required

- i. What is the expected value in this game?
- ii. Comment on the outcome of the results

Solution

Self assessment 3.3 question

i) Expected value = $\sum xP(x)$

If you pick the right numbers , you win Gh₵500, but in actual sense, you win GH₵499.00 since you have to pay GH₵1 to win.

Since there are ten numbers and we are to pick one . The probability of picking a right number = $\frac{1}{10}$, probability of picking a second right number = $1/10$ and that of the third is also = $1/10$.

$$\text{The probability of winning} = \frac{1}{10} \times \frac{1}{10} \times \frac{1}{10} = \frac{1}{1000} = 0.001$$

$$\text{The probability of losing} = 1 - \frac{1}{1000} = 0.999$$

Presenting the outcome in a table, we have

Win or lose	x	P(x)	x p(x)
Win	499	0.001	$499 \times 0.001 = 0.499$
lose	-1	0.999	$-1 \times 0.999 = -0.999$

$$\text{Expected value} = 0.499 + -0.999 = -0.50$$

- ii) The conclusion is that, in the long run, you will expect to lose \$0.50, since the expected value is not Zero(0). The game is not a fair game. It is fair when the expected value is zero. Refer to page 230 of the text.

4. a) Bailey has been playing golf on the weekends for the past three years. Recently, she started keeping track of her recorded scores. Her scores for June and July at her favorite 9-hole (par 36) golf course are provided below.

Data: 45 49 42 56 41 36 34 38 41 40 42 41 39 38 40 39 36 41

Required:

- I. Find the Mean, Median, and Mode from the above data.
- II. What does this information tell you about Bailey's golf game?
- III. Is this data skewed?

Solution

i) Mean = $\frac{\sum x}{n} = \frac{45+49+42+56+41+36+34+38+41+40+42+41+39+38+40+39+36+41}{18} = 41$

Median =

Re-arranging

34, 36, 36, 38, 38, 39, 40, **40, 41, 41, 41, 41, 42, 42, 45, 49, 56**

Median = average of two middle numbers = $\frac{40+40}{2} = \frac{81}{2} = 40.5$

Mode = most common number = 41

- ii) It means Baileys mean score for the month June and July is 41 (only 5 above par). Over half of Baileys games had a score above 40.5 and half had a score below 40.5. her most common golf score is 41.
- iii) NO. The data is not skewed. The mean and median are almost identical, only half-stroke difference. refer to page 66 of the text.

- b) I. Convert the following data sets to stem and leaf.

23	58	62	62	63	65	67
71	71	72	80	82	82	82

- II. How is standard deviation better than range?

Solution

A self assessment 1.5Q2.

Stem	Leaf
2	3
3	
4	
5	8
6	2 2 3 5 7
7	1 1 2
8	0 2 2 2

QUIZ 1

Use the information to answer questions 1 and 2.

81,86,78,80,81,82,92,90,79,83,84,95,85,88,80,78,84,79,80,83,79,87,84,80.

1. Find the leaves associated with 7.
 A.88899 B.78999 C.88999 D.788999

Solution

78,79,78,7979: The

leaf
8,9,8,9,9

Ans.C.

2. Find the leaves associated with 9
 A.88999 B.022 C.026 D.025

Solution

90,92,95

Stem	leaf
9	0,2,5

Ans.D

3. Find the Lower Quartile (Q1) of the following data set: 3,5,7,8,9,11,15,16,20,21.
 A.7.5 B.8 C.7 D.8.5

Solution

Re-arrange the data. Pick the median (Middle number). Find the middle number of the first set.

First five numbers = 3,5,**7,8,9,11**.

Lower quartile = $\frac{7+8}{2} = 7.5$. **Ans.A**

4. Find the interquartile range (IQR) of the following list of numbers: 1,9,2,3,7,8,9,2.
 A.7.5 B.6.5 C.8.5 D.2.5

Solution

Re-arranging: **1,2,2,3,7,8,9,9**

$$Q_1 = \frac{2+2}{2} = 2, Q_2 = \frac{8+9}{2} = 8.5. \textbf{Ans.C}$$

5. The best measure of central tendency corresponding to the nominal data is A. mode B. median C. mean D. None of the above

Solution

Ans.A. refer to page 39 of the text

6. Find the sample mean for the following set of numbers:

12,13,14,16,17,40,43,55,56,67,78,78,79,80,81,90,99,101,102,304,306,400.
A.155 B.165 C.144 D.140.5

Solution

$$\text{Sample mean} = \frac{\sum x_i}{n} =$$

$$\frac{12+13+14+16+17+40+43+55+56+67+78+78+79+80+81+90+99+101+102+304+306+400}{26} =$$

$$\frac{3744}{26} = 144. \textbf{Ans.C.} \text{ Refer to page 42 of the text.}$$

7. Business statistics is a special area of statistics which are applied in the business setting. A. True B. False

Solution

True. Refer to page 1 of the text.

8. It is known that 10% of parliamentarians voted for a particular policy implementation. There are only 273 parliamentarians, so the researcher can count what every single one of them voted for. This statistical measure is statistic
A. True B. False

Solution

True.

9. 45% of Ghanaian residents in Germany report that they have been to at least one tourist Site at Cape Coast. It is very doubtful that anyone polled in excess of a million people for this data. Therefore a statistics is an appropriate measure for this data. A. True B. False

Solution

True

10. Which of the following is inherently categorical measurement scale?
A. Political affiliation B.Weight C.Volume D. Number of students in a class

Solution

Ans. D. refer page 11 of the text.

11. A/an----- scale is a measurement scale, in which numbers serve as “tags or “labels” only to identify or classify an object.

Visit archornews.com for great content in mathematics.

- A. Ordinal B. nominal C. ratio D. None of the above.

Solution

Ans.C. Refer to Page 13 of the text..

12. Compute the sample variance given that the scores 1,2,4, and 5 were sampled from a larger population.
A.3.33 B.3 C. 2.3 D.7

Solution

$$\bar{x} = \frac{1+2+4+5}{4} = 3.$$

$$\text{Sample variance} = \frac{\sum(x_i - \bar{x})^2}{n-1} = \frac{(1-3)^2 + (2-3)^2 + (4-3)^2 + (5-3)^2}{4-1} = \frac{10}{3} = 3.33. \text{ Ans.A}$$

13. Compute the coefficient of variation for an average price of GH₵50.00 and with standard deviation GH₵5.00 for 5 pineapples.
A.5% B. 10% C. 10.5% D. 12%

Solution

Coefficient of variation (CV) = $\frac{s}{\bar{x}} 100\% = \frac{5}{50} (100) = 10\%.$ **Ans.B** refer to page 55 of the text.

14. IQ scores in a large population have a bell-shape distribution with mean $\mu = 100$ and standard deviation $\sigma = 15.$ Find the standardized score (Z-score) for a person with an IQ of 121.
A. 1.5 B. 1.4 C. 1.47 D.0.9

Solution

$$\text{Z-score} = \frac{x-\mu}{\sigma} = \frac{121-100}{15} = \frac{21}{15} = 1.4. \text{ Ans.B. refer to page 59 of the text.}$$

15. Two cards have been drawn from the deck of 52 cards without replacing the first one back. Find the probability of getting first card king and second card queen.
A. 4/51 B. 4/52 C. 4/663 D. 4/664

Solution

Probability of drawing a first king(A) = 4/52

Probability of drawing first queen(B) = 4/51

Probability of drawing king and queen = $\frac{4}{52} \times \frac{4}{51} = \frac{4}{663}.$ **Ans.C.** refer to page 75 of the text.

16. Given that accident occur with a poison distribution at an average of 4 per week. What is the probability that at least two weeks will elapse between accidents? A. 3.4×10^{-3} B. 2.4×10^{-4} C. 3.4×10^{-4} D. 3.4×10^{-2} .

Solution

17. Which measurement scale is appropriate for the statement, "How many hours a day do you spend on a computer"?
A. Nominal scale B. Ratio scale C. Ordinal scale D. Interval scale

Solution

Ans.B. because the question ask respondents to respond in a measurable way.

Ratio scale questions have a true zero and often times will be presented in an ordinal way. Refer to page 225 of the text for further details.

Use the following information to answer questions 18 to 20.

Two way contingency table with marginal frequencies.

	Dance	Sports	TV	Total
Men	2	10	8	20
Women	16	6	8	30
Total	18	16	16	50

18. From the table above, compute the joint total relative frequency of female.
A. 0.040 B. 0.65 C. 0.60 D. 0.40

Solution

Total relative frequency for female = $30/50 = 0.6$. **Ans.C**

19. Compute the joint total relative frequency of Dances given in the table above.
A. 0.36 B. 0.32 C. 0.60 D. 0.40

Solution

Total relative frequency for dancers = $18/50 = 0.36$. **Ans.A**

20. Find the joint relative frequency of men in Tv.
A. 0.32 B. 0.12 C. 0.16 D. 0.60

Solution

Joint total relative frequency for men in TV = $8/50$
0.16. **Ans.C.** Refer to page 24 of the text.

END OF SEMESTER EXAMS

1. A two -dimensional chart showing time on the horizontal axis and the variable of interest on the vertical axis is
A.Pie Chart B. Bar Chart C. Line Chart D.None of the above.

Solution

Ans. C. refer to page 29 of the text.

Use the following information to answer questions 2 to 4.

Given that x numbers are 568,570,575,578,584, and the y numbers are 573,574,575,578, and they are the inner diamters of two grades of tires.

2. Find the mean of grade x
A.575.00 B. 577.00 C.579.00 D.557.00
3. Which of the following

SBU: 302:STATISTICS FOR BUSINESS DECISIONS: SECOND SEMESTER (FULL TIME STUDENTS ON CAMPUS)

Question One

- A. Define type I and II errors. How does the choice of a given level of type I error influence the level of Type II error? How is Type I error related to the level of significance?

Solution

Type I error is the probability of rejecting the null hypothesis when in actual case, it is true.

Type II is the probability of failing to reject a null hypothesis when in actual case, it is false. The level of significance is related to the type I error because it is the maximum allowable probability of committing a type I error. Refer to page 134 of the text.

- B. For the Table below, complete the spaces X,Y and Z.

	Null Hypothesis True	Null Hypothesis False
Reject Null Hypothesis	X	Correct decisions
Fail to reject Null Hypothesis	Y	Z

Solution

X = Type I error

Y= correct decision

Z= Type II error.

- C. There is a raging debate about the salaries of new graduates of economics and the business school. Some prospective applicants are keeping an eye on the future as the apply to these programmes at the University. The HR manager of the school of Business Incubator decided to test whether there is any different in the salaries

earned by the respective graduates of these disciplines upon graduation. She collected a sample of 15 Economics and 15 business school graduates for study. Her study revealed that the mean salary of economics graduates is 4.5 units of currency with a variance of 1 and the mean salary of business school graduates is 3.4 units of currency with a variance of 1. Assuming these are independent samples drawn from a normally distributed population and using $\alpha = 0.05$
 Required:

- i) State the null hypothesis
- ii) State the decision rule
- iii) Calculate the test statistic
- iv) Make a statistical decision
- v) State your conclusion based on your answers to (i), (ii) and (iii) above.

Solution

- i) $H_0: \mu_1 = \mu_2$ $H_A: \mu_1 \neq \mu_2$
- ii) Decision rule: reject Null hypothesis if t-calculated is greater than t-table or t-calculated is less than negative t-calculated.

$$\text{iii) } t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} = \sqrt{\frac{(15-1)1 + (15-1)1}{15+15-2}} = \sqrt{\frac{28}{28}} = 1$$

$$t = \frac{(4.5 - 3.4) - 0}{\sqrt{\frac{1}{15} + \frac{1}{15}}} = \frac{1.1}{0.365} = 3.014.$$

- iv) $t_{0.025} = 2.0484$, since it is two tail test, the significance level is divided into two $0.05/2 = 0.025$. Since $3.014 > 2.0484$. we reject the null hypothesis.
- v) The rejection of the null hypothesis indicates that there are differences in the salaries of graduates of the two disciplines.

Question Two

- A. Differentiate between time series and cross-sectional data.

Solution

Time series : a set of consecutive data values observed at successive points in time whiles cross-sectional data are those set of data values observed at a fixed point in time. Refer to page 12 of the text.

- B. The equation for the margin of error when estimating a population mean is given below. $\sigma + Z \frac{\sigma}{\sqrt{n}}$, where σ , Z and n connote usual notation.

Required:

Indicate the effect on the margin of error resulting from an increase in each of the following items.

- i) Confidence level

- ii) Standard deviation
- iii) Sample size
- iv) Standard error

Solution

- i) If the confidence level is increased, the margin of error will increase also.
- ii) If the standard deviation is increased, margin of error will increase
- iii) If the sample size increased, the margin of error reduces.
- iv) If the standard error increases, margin of error also increases.

Refer to page 124 of the text.

- C. A human resource manager wishes to estimate the proportion of 3000 employees in her large company who have supplemental health insurance. What is the largest size sample she should select if she wants 95 percent confidence and a margin of error of ± 0.01 ?

Solution

$$\text{Sample size}(n) = \left(\frac{z\sigma}{e} \right)^2 =$$

$$\text{But } e = z \frac{\sigma}{\sqrt{n}} \Rightarrow 0.01 = 1.96 \frac{\sigma}{\sqrt{3000}} \Rightarrow \frac{0.01 \times \sqrt{3000}}{1.96} = \sigma \Rightarrow \sigma = 0.2795$$

$$\text{Sample size}(n) = \left(\frac{z\sigma}{e} \right)^2 = \left(\frac{1.96 \times 0.2795}{0.01} \right)^2 = \frac{0.3001}{0.001} = 300.10 \cong 300.$$

- D. Indicate the level of data measurement for each of the following variables.

- i) Type of business
- ii) Pieces of luggage
- iii) Sex(Male/Female)
- iv) Day of week

Solution

- i) Types of business-nominal
- ii) Pieces of luggage- ordinal
- iii) Sex(male/female) - Nominal
- iv) Day of week- ordinal

- E. In each of the variables in question (2D) above, identify the appropriate central tendency.

Solution

- v) Types of business-nominal-mode
- vi) Pieces of luggage- ordinal-median
- vii) Sex(male/female) - Nominal-mode
- viii) Day of week- ordinal-median

Question Three

- A. State one assumption Chi-Square distribution makes about the expected frequency cell and how it can be overcome.

Solution

- B. The ATM card Department of Awazi Bank knows from experience that 5% of its card holders whose cards have been called in for failure to pay their charges this month, 50 had some high school education 100 had completed high school. 190 had some University education and 160 had completed University.

Required:

Can we conclude that the distribution of cards holders who do not pay their chargers is different from all others? Use the 0.01 significance level.

Solution

- C. i) Explain the following terms and give an example each?
ii) Quantitative variables
iii) Discrete variables
v) Continuous variables

Solution

- i) quantitative variable –is a variable that is observed or measured in an experiment.
ii) discrete variable – is a variable that can assume only finite or infinite countable number of distinct values such as 0,1,2,3 etc.
iii) Continuous variable is one that takes all values in a given interval of numbers.

Question Four

The following regression output is the results of a multiple regression application in which we are interested in explaining the variation in academic performance of students based on six independent variables. Type of lecturer (1= flexible, 0 = not flexible), income level, relationship status (1=yes,0 = No), number of lectures attended. Hours of tutorial sessions and IQ. The following computer printout shows the final output. However, several values are omitted from the printout.

Summary Output

Regression statistics

Multiple R	0.8069
R square	0.6511
Adjusted R Sqaure	0.4900
Standard Error	4.0361
Observation	20

ANOVA

	df	SS	MS	F	F
Regression	6	3951745	C	E	0.0166
Residual	B	211.7755	D		
Total	19	606.9500			

	Coefficient	Standard Error	tStat	P-value	Lower 95%
Intercept	10.0878	6.3696	1.5837		
Type of lecturer	1.6803	2.1841	0.7693		
Income level	-0.7255	0.2201	- 3.2962		
Relationship status	-0.6914	2.3526	- 0.2939		
No. of lectures attended	1.1045	0.3252	3.3964		
Hours of Tutorial session –	0.3306	0.5809	- 0.5691		
IQ	0.6610	2.6397	0.2504		

Given this output

- Interpret the Adjusted R-square and explain why it's better than R-square in multiple regression
- Compute the following missing values B,C ,D and E.
- Compute the t-statistics for each of the independent variable.
- Identify which of the variables is/are significant in explaining the variation in the dependent variable and interpret the results. Test using 0.05 level of alpha.

Solution

- The adjusted R square is a measure of the percentage of explained variation in the dependent variable in a multiple regression model that takes into account the relationship between the sample size and the number of independent variables. It is expected to be less than the R-square as can be seen in the table.
It means there is 49% relationship between the sample size and the independent variables. The R-square increases with the addition of more independent variables. The Adjusted R square corrects this situation.

Visit **archornews.com** for great content in mathematics.

b) $B = 19 - 6 = 13$

$$C = \frac{ss}{df} = \frac{395.1745}{6} = 65.8624$$

$$D = \frac{211.7755}{13} = 16.2904$$

$$E = \frac{65.8624}{16.2904} = 4.0430$$

c) t-statistics calculations:

$$\frac{10.0878}{6.3696} = 1.5837, \frac{1.6803}{2.1841} = 0.7693, \frac{-0.7255}{0.2201} = -3.2962$$

$$\frac{-0.6914}{2.3526} = -0.2939, \frac{1.1045}{0.3252} = 3.3964$$

$$, \frac{-0.3306}{5809} = -0.5691, \frac{0.6610}{2.6397} = 0.2504$$

- d) To determine which variable is significant, we need to determine the probabilities of the variables and compare those probabilities with the significance level of 0.05. You can ask your tutors to guide you in calculating the probabilities.