UNIVERSITY OF THE PUNJAB

Fifth Semester – 2019

Examination: B.S. 4 Years Program

Roll No. in Fig. ....

Roll No. in Words. ....

....

PAPER: Statistical Computer Packages
Course Code: STAT-309 Part–I (Compulsory)

MAX. TIME: 15 Min.
MAX. MARKS: 10

....
Signature of Supdt.:

Attempt this Paper on this Question Sheet only.
Please encircle the correct option. Division of marks is given in front of each question.
This Paper will be collected back after expiry of time limit mentioned above.

Q.1. Encircle the right answer, cutting and overwriting is not allowed. (1x10=10)

i. SPSS data file extension is:

a) .sps b) .sav c) .spo d) .xls

ii. To find the Mode for a variable, which set of instructions should you give SPSS?

a) Analyze → Descriptive Statistics → Explore → Statistics → Descriptives → ok.
b) Graphs → Descriptive Statistics → Frequencies → Statistics → ok.
c) Analyze → Descriptive Statistics → Descriptives→ Options→ ok
d) None of the Above

iii. Types of Blood Group is an example of:

a) Nominal Data b) Ordinal Data c) Interval Data d) Ratio Data

iv. In Testing of Hypothesis problems, One may reject the null hypothesis, when:

a) $p > \alpha$ b) $p < \alpha$ c) $p \leq \alpha$ d) $p \geq \alpha$

v. In an experiment to determine if antibiotics increase the final dressed weight of cattle, the following were measured on each animal in the study.

sex, initial weight, weight gain, grade of meat.

where grade is recorded as (A, B, or C). The scales of measurement of these variables are:

a) Nominal, ratio, interval, nominal
b) Nominal, ratio, ratio, nominal
c) Nominal, ratio, ratio, ordinal
d) Ordinal, ratio, ratio, ordinal

vi. The appropriate graph for a continuous nature of data is:

a) Histogram b) Simple Bar Chart c) Component Bar Chart d) Pie Chart

vii. In SPSS v 20 or above, one can define a variable name of maximum....characters?

a) 8 b) 16 c) 32 d) 64

viii. The first character of variable name defined in SPSS should be.

a) Numeric b) Alphabet c) Special Character d) None of these

ix. Shoe-size is an example of .....:

a) Nominal Data b) Ordinal Data c) Interval Data d) Ratio Data

x. Time on a Clock is an example of .....:

a) Nominal Data b) Ordinal Data c) Interval Data d) Ratio Data

UNIVERSITY OF THE PUNJAB

Fifth Semester – 2019

Examination: B.S. 4 Years Program

Roll No. ....

PAPER: Statistical Computer Packages
Course Code: STAT-309 Part – II

MAX. TIME: 2 Hrs. 45 Min.
MAX. MARKS: 50

ATTEMPT THIS (SUBJECTIVE) ON THE SEPARATE ANSWER SHEET PROVIDED

Q.2. Give short answers. **(4 x 5 = 20)**

a) State the Rules for defining a variables Name in SPSS.
b) Nominal, Ordinal, Interval and Ratio Scale of Measurement
c) Use of Compute Command in SPSS
d) Assumptions of Parametric Tests.

SUBJECTIVE

Note: You are required to justify your answers in this section by stating hypothesis, statistics, reference of graphs etc. on your answer sheet.

Q3: The following is the SPSS output from a data containing the Horsepower of 156 cars.

			Statistic	Std. Error
Horsepower	Mean		185.95*	4.540
	95% Confidence Interval for Mean	Lower Bound	176.98	
		Upper Bound	194.92	
	5% Trimmed Mean		183.04	
	Median		177.50	
	Variance		3214.926	
	Std. Deviation		56.700	
	Minimum		55	
	Maximum		450	
	Range		395	
	Interquartile Range		67	
	Skewness		1.001	.194
	Kurtosis		2.407	.386

Runs Test

	Horsepower
Test Value[a]	178
Cases < Test Value	78
Cases >= Test Value	78
Total Cases	156
Number of Runs	62
Z	-2.731
Asymp. Sig. (2-tailed)	.006

a. Median

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Horsepower	.077	156	.026	.949	156	.000

	N	Mean	Std. Deviation	Std. Error Mean
Horsepower	156	185.95	56.700	4.540

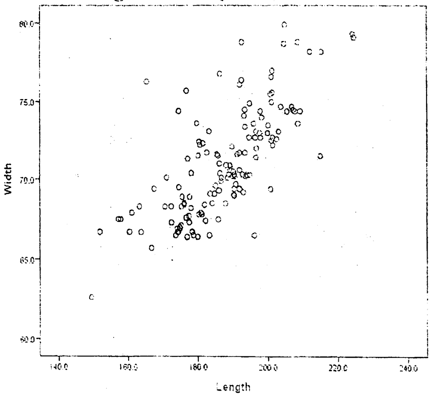
One-Sample Test

	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Horsepower	-3.095	155	.002	-14.051	-23.02	-5.08

Study the above output carefully answer the following questions: (2+2+2)

a) Test the normality of the data.
b) Test the randomness of the data.
c) Can we say that the average horsepower of the cars is 200?

Q4. The following is the output of bivariate data of 156 cars width and length:



Correlations				
			Length	Width
Length	Pearson Correlation		1	.710
	Sig. (2-tailed)			.000
	N		156	156

Correlations				
			Length	Width
Spearman's rho	Length	Correlation Coefficient	1.000	.716**
		Sig. (2-tailed)	.	.000
		N	156	156

i) Can we say that the relationship between cars width and length is linear?

ii) Test the significance of the relationship between cars width and length by using an appropriate method? (2+3)

Q4: a) The following is the SPSS output of a data set measuring an association of consumer's satisfaction between color and taste of soft drinks.

			Taste				
			Very Satisfied	Satisfied	Dissatisfied	Very Dissatisfied	
Color	White	Count	108	68	60	47	283
		Expected Count	102.2	65.4	69.1	46.3	283.0
		% within Color	38.2%	24.0%	21.2%	16.6%	100.0%
		% within Taste	77.7%	76.4%	63.8%	74.6%	73.5%
		% of Total	28.1%	17.7%	15.6%	12.2%	73.5%
	Black	Count	10	4	18	4	36
		Expected Count	13.0	8.3	8.8	5.9	36.0
		% within Color	27.8%	11.1%	50.0%	11.1%	100.0%
		% within Taste	7.2%	4.5%	19.1%	6.3%	9.4%
		% of Total	2.6%	1.0%	4.7%	1.0%	9.4%
	Green	Count	21	17	16	12	66
		Expected Count	23.8	15.3	16.1	10.8	66.0
		% within Color	31.8%	25.8%	24.2%	18.2%	100.0%
		% within Taste	15.1%	19.1%	17.0%	19.0%	17.1%
		% of Total	5.5%	4.4%	4.2%	3.1%	17.1%
Total		Count	139	89	94	63	385
		Expected Count	139.0	89.0	94.0	63.0	385.0
		% within Color	36.1%	23.1%	24.4%	16.4%	100.0%
		% within Taste	100.0%	100.0%	100.0%	100.0%	100.0%
		% of Total	36.1%	23.1%	24.4%	16.4%	100.0%

		Value	Approx. Sig
Nominal by Nominal	Contingency Coefficient	.197	.017
N of Valid Cases		385	

i) State the null and alternative hypothesis for the association for satisfaction between Color and Taste of the soft drink.

ii) What percentage of consumers was very satisfied with the black color of the soft drink?

iii) What is the decision regarding null hypothesis? (1+1+4)

b) We find some depressed people and check that they are all equivalently depressed to begin with. Then we allocate each person randomly to one of three physical exercise groups: no exercise; 20 minutes of jogging per day; or 60 minutes of jogging per day. The following output of allocated groups is generated through SPSS

Levene Statistic	df1	df2	Sig.
1.708	2	21	.205

	Exercise Group
Chi-Square	7.290
df	2
Asymp. Sig.	.026

Kruskal Wallis Test

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1306.083	2	653.042	4.414	.025
Within Groups	3107.250	21	147.964		
Total	4413.333	23			

i) Test the homogeneity of exercise groups

ii) Can we conclude that there is no significant difference among the exercise groups (2+2)

Q5. The administrator of a new paralegal program at Seagate Technical College want to estimate the grade point average in the new program. He thought that high school GPA, the verbal score on the Scholastic Aptitude Test (SAT) and the mathematics score on the SAT would be good predictors of paralegal GPA. The following output is generated from the 9 students recorded data using SPSS:

Correlations

		High School GPA	SAT Verbal	SAT Math	Paralegal GPA
		1	.609	.636	.911
High School GPA	Sig. (2-tailed)		.082	.066	.001
	N	9	9	9	9
		.609	1	.599	.616
SAT Verbal	Sig. (2-tailed)	.082		.088	.077
	N	9	9	9	9
		.636	.599	1	.487
SAT Math	Sig. (2-tailed)	.066	.088		.183
	N	9	9	9	9
		.911	.616	.487	1
Paralegal GPA	Sig. (2-tailed)	.001	.077	.183	
	N	9	9	9	9

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.928[a]	.861	.778	.37513	1.974

a. Predictors: (Constant), High School GPA, SAT Verbal, SAT Math

b. Dependent Variable: Paralegal GPA