S115_LE4_002

Statistics 115
Sample Fourth Long Examination

TGCapistrano

I. FILL IN THE BLANKS.

1. In a 4x3 table, the number of cells with expected frequencies less than 5 but greater than 1 must not be more than _________ for the chi-square test to be valid.
2. The population linear correlation coefficient of X and Y, denoted by ρ, is defined as ρ = ________.
3. The alternative hypothesis in a test for homogeneity is ________________.
4. If Y_i is the value of the response variable of the ith element and X_i is the value of the explanatory variable of the ith element then the simple linear regression model is defined as (a) Y_i = _______________ that satisfies the following conditions about the error term ε_i: (b) _____________________.
5. The individual deviations of the observations y_i from the predicted values using the estimated regression equation are called ___________.
6. In simple linear regression analysis based on a sample of size 112, if the SSE=220 then the estimate for the variance of ε_i is equal to __________.
7. The basic principles of experimental design are (a) ___________, (b) ________, and (c) _____________.
8. The extraneous variations in the values of the response variable which tend to mask the true effects of the treatments due to the inherent variations among experimental units and lack of uniformity in the physical conduct of the experiment is called ____________.
9. The component of a time series that describes the long-term movement in the time series and usually modeled by a smooth curve is called ___________.
10. The smoothed value at time t+1 using single exponential smoothing can be expressed as

$$S_{t+1} = \alpha Y_t + (1-\alpha)\alpha Y_{t-1} + (a)\_\_\_\_\_ Y_{t-2} + (b)\_\_\_\_\_ Y_{t-3} + \ldots$$

II. TRUE OR FALSE. Write 'True' if the statement is always true; otherwise, write 'False'.

1. If ρ is close to -1 then the linear relationship between X and Y is strong.
2. If ρ is equal to 1 then either X causes Y or Y causes X.
3. If ρ is equal to 0 then there is no linear relationship between X and Y.
4. The Pearson product moment correlation coefficient between X and Y, denoted by r, can be any real number.
5. If r is equal to 1 then all of the sample points (X_i, Y_i) will fall on a line whose slope is positive.
6. If r is equal to 0 then all of the sample points (X_i, Y_i) will fall on a line whose slope is 0.
7. In the simple linear regression model, the mean of the response variable depends on the value of the explanatory variable.
8. The criterion used to derive the estimates for the regression coefficients using the method of least squares is to minimize the sum of the squares of the error terms □is.
9. If the estimated regression equation is $\hat{Y} = 5 - 3X$ then for every unit increase in X the mean of Y is estimated to decrease by 3 units.

10. The length of the (1-α)100% confidence interval estimate for the regression coefficient β_1 is longer when the MSE is smaller .
11. The cyclical component of a time series recurs in varying length and magnitude.
12. The seasonal component of a time series recurs in varying length and magnitude.
13. The smaller the value of T in the single moving average then the more the moving averages will follow the pattern of the data set.
14. If the weight α used in single exponential smoothing is close to 1 then the greater the smoothing effect.
15. The exponential decrease in the weights of the past observations is faster if the weight α used in single exponential smoothing is closer to 1.

III. COMPUTATIONS.

1. The accompanying paired data give X, Narco Medical's advertising expense during a period (in hundreds of dollars), and Y, the amount of the company's sales (in thousands of dollars) for the period.

X	8	9	7	6	5	1
Y	15	11	10	11	8	5

a) Find the estimated regression equation with X as the explanatory variable and Y as the response variable.
b) Construct the scatterplot. Superimpose the graph of the regression line derived in (a).
c) Find the predicted value of Y when X=5 using the estimated regression equation in (a).
d) What is the critical region when we test Ho: $\beta_1 = 0$ vs Ha: $\beta_1 \neq 0$ at α=0.05?
e) Compute for the value of the test statistic to test Ho: $\beta_1 = 0$ vs Ha: $\beta_1 \neq 0$.
f) What percentage of total variation in the values of Y in the sample can be accounted for or explained by its linear relationship with the values of X?
g) Compute for the estimate for ρ .

2. A random sample of buyers of cars from a large automobile agency was taken to determine if the type of automobile purchased and sex of buyer are related. The collected data from the sample were summarized as follows:

Sex	Type of Car			
	Subcompact	Compact	Midsize	Full Size
Male	90	82	68	76
Female	15	12	27	30

a) State Ho and Ha.
b) Write the formula of the test statistic to be used.
c) State the decision rule at 0.05 level of significance.
d) Compute for the value of the test statistic.
e) Is there sufficient evidence at 0.05 level of significance to conclude that the type of automobile purchased and sex of buyer are related?
f) Compute for Cramer's V.

3. The following data show the number of mergers that took place in an industry over a 10-year period.

Year	Mergers
1990	23
1991	32
1992	32
1993	42
1994	64
1995	47
1996	96
1997	125
1998	140
1999	160

Compute for the smoothed value for 1993 using the specified method.

a) single moving average with T=4

b) single moving average with T=2

c) single exponential smoothing with α =0.25 and initial value=Y1

d) single exponential smoothing with α =0.75 and initial value=Y1

4. Five different baking temperatures (A, B, C, D, E) are being tested to bake a ready-mix cake. The experiment will be performed on five days (1st day of the months Dec, Jan, Feb, March, and April) and five timeslots for each day (8:00, 10:00, 12:00, 2:00, 4:00.) Indicate the baking temperature to be used for each time slot of each month using the Latin square design by filling-up the following table:

Month	Timeslot				
	8:00	10:00	12:00	2:00	4:00
Dec					
Jan					
Feb					
Mar					
Apr					

Use the following random numbers and the 5x5 latin square below:

Row 1: 26 Row 2: 58 Row 3: 14 Row 4: 76 Row 5: 8

Col 1: 94 Col 2: 74 Col 3: 88 Col 4: 25 Col 5: 68

Formulas:

$$S_{b_1} = \sqrt{\frac{MSE}{\sum_{i=1}^{n} X_i^2 - \frac{\left(\sum_{i=1}^{n} X_i\right)^2}{n}}}$$

Latin square:

A	B	C	D	E
B	A	E	C	D
C	D	A	E	B
D	E	B	A	C
E	C	D	B	A