

國立政治大學 110 學年度第 2 學期 Quiz(1) 考試命題紙

考試科目：Regression Analysis (I)

開課班別：統計二

命題教授：吳漢銘

考試日期：03 月 24 日 (四) 11:10-12:00

※准帶項目打「O」· 否則打「×」

1. 需加發計算紙或答案紙請備註。
2. 為環保節能減碳· 試題一律採雙面印刷· 如有特殊印製需求· 請註記：

本試題共 2 頁· 印刷份數：90 份

Calculator

Book Notes

Dictionary

Cell phone Laptop

備註：注意事項要看!! (§1~§2.7)

O

×

×

×

Note: (1) Fill in your name and student ID number ° (2) Answer all questions in English ° (3) Answer each question in the order it appears ° (4) Pencils are permitted ° (5) Hand in the question, the answer sheet, and the sketch paper ° (6) The calculation process is required.

1. (20%) Explain the following:

(a) What is the "Regression Analysis"?

(b) Let α be the level of the significance. What is the so-called "(1 - α)% Confidence Interval" for a parameter θ of the population.

(c) What is the "ANOVA table" for simple linear regression? What is it used for?

(d) For a simple linear regression based on the normal error model, why do we need to derive the sampling distribution of the β_1 ? (i.e., What is the sampling distribution of the β_1 used for?)

2. (20%) For the given sample observations $\{(X_i, Y_i), i = 1, \dots, n\}$, we assume a normal error regression model as $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$, where ϵ_i are independent normally distributed with mean 0 and variance σ^2 . Find the MLEs of the parameters β_0 and β_1 .

3. (20%) (Referring to the above question) Derive the sampling distribution of $\frac{b_1 - \beta_1}{s(b_1)}$, where b_1 is the MLE of β_1 and $s(b_1)$ is the sample standard deviation of b_1 .

4. (10%) Given a random sample of data, $\{(X_i, Y_i), i = 1, \dots, n\}$, and the level of the significance α , describe how to conduct the two-sided test concerning whether or not there is a linear association between X and Y for a normal error regression model. (State the null hypothesis, alternative hypothesis, test statistics (in terms of data), and decision rule.)

5. (20%) **Grade point average.** The director of admissions of a small college selected 120 students at random from the new freshman class in a study to determine whether a student's grade point average (GPA) at the end of the freshman year (Y) can be predicted from the ACT test score (X). The results of the study follow. Assume that a simple linear regression model is appropriate.

$i:$	1	2	3	...	118	119	120
$X_i:$	21	14	28	...	28	16	28
$Y_i:$	3.897	3.885	3.778	...	3.914	1.860	2.948

國立政治大學 110 學年度第 2 學期 Quiz(1) 考試命題紙

考試科目：Regression Analysis (I)

開課班別：統計二

命題教授：吳漢銘

考試日期：03 月 24 日 (四) 11:10-12:00

※准帶項目打「O」· 否則打「×」

1. 需加發計算紙或答案紙請備註。
2. 為環保節能減碳· 試題一律採雙面印刷· 如有特殊印製需求· 請註記：

本試題共 2 頁· 印刷份數：90 份

Calculator

Book
Notes

Dictionary

Cell phone
Laptop

備註：注意事項要看!! (§1~§2.7)

O

×

×

×

The regression analysis report conducted by R is given in Table 1.

- (a) (5%) Obtain a 95 percent confidence interval for β_1 . Interpret your confidence interval. Does it include zero? Why might the director of admissions be interested in whether the confidence interval includes zero? ($t_{0.025,120} = -1.97993$, $t_{0.05,120} = -1.657651$, $t_{0.025,119} = -1.9801$, $t_{0.05,119} = -1.657759$, $t_{0.025,118} = -1.980272$, $t_{0.05,118} = -1.65787$)
- (b) (5%) Test, using the test statistic t^* , whether or not a linear association exists between student's ACT score (X) and GPA at the end of the freshman year (Y). Use a level of significance of 0.05. State the alternatives, decision rule, and conclusion.
- (c) (5%) What is the P -value of your test in part (b)? How does it support the conclusion reached in part (b)?
- (d) (5%) The ANOVA table is shown in Table 2. How to you interpret ANOVA results?
6. (10%) In a simple linear regression, show that the total sum of squares can be partitioned into the error sum of squares and the regression sum of squares.

注意：1、考試求公平及公正· 請同學務必自律· 維護學校與學生之榮譽。

2、考試時不得有交談、窺視、夾帶、抄襲、傳遞、代考或其它作弊等舞弊行為· 考畢務必交卷· 不得攜卷出場· 違者依考場規則議處。

Table 1: Regression analysis for Grade point average data

```

Call:
lm(formula = GPA ~ ACT, data = ex2.4.data)

Residuals:
    Min       1Q   Median       3Q      Max
-2.74004 -0.33827  0.04062  0.44064  1.22737

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.11405    0.32089   6.588 1.3e-09 ***
ACT           0.03883    0.01277   3.040 0.00292 **
---
Signif. codes:  0  '***'  0.001  '**'  0.01  '*'  0.05  '.'  0.1  ''  1

Residual standard error: 0.6231 on 118 degrees of freedom
Multiple R-squared:  0.07262, Adjusted R-squared:  0.06476
F-statistic:  9.24 on 1 and 118 DF,  p-value: 0.002917

```

Table 2: Analysis of Variance Table for Grade point average data

```

Analysis of Variance Table

Response: GPA
      Df Sum Sq Mean Sq F value    Pr(>F)
ACT     1  3.588   3.5878   9.2402 0.002917 **
Residuals 118 45.818   0.3883
---
Signif. codes:  0  '***'  0.001  '**'  0.01  '*'  0.05  '.'  0.1  ''  1

```