國立政治大學 112 學年度第 2 學期 Midterm Exam 考試命題紙

Subject： Statistics (II) 開課班別： 統計學整合開課 Teacher: Han-Ming Wu

Date：21 Mar. (Thur) 15:00-16:00	※Allowed:「O」, Prohibited:「×」				1. 需加發計算紙或答案紙請在試題內封袋備註。 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：
Pages: 3，Copies: 65	Calculator	Textbook	Class notes	3C product	
Scope: §ch9-11	O	×	×	×	

Notes:

(1) Fill in the student ID number and name on the answer sheet。

(2) Answer all questions in English (部份寫中文可)。

(3) Answer each question in the order it appears。

(4) It is recommended to use a dark ballpoint pen。(鉛筆可)

(5) The calculation process (for parts **IV** and **V**) is required (calculate to 4 decimal places)。

(6) Return both the answer sheet and the question sheet.

(7) The total score is 120.

(8) Please transcribe the following oath onto the answer sheet. (請複寫下列宣誓詞至答案卷上)。

0. Please write in either Chinese or English. (10 points will be deducted if not written.)(不寫扣 10 分)

" 本人姓名 恪遵各項考試規則，若如違反，願受校方最嚴厲處罰，謹誓。"

"I (your name here) will strictly adhere to all examination rules. If I break this oath, I am willing to accept the most severe punishment imposed by the school. Solemnly sworn."

(I) Multiple choice (20%, 10% each); select one correct answer.

1. For a lower tail test, the p-value is the probability of obtaining a value for the test statistic at least as
   (A) small as that provided by the sample.
   (B) large as that provided by the sample.
   (C) small as that provided by the population.
   (D) large as that provided by the population.

2. To avoid the problem of not having access to tables of the F distribution when a one-tailed test is required and with F values given for the lower tail, let the
   (A) smaller sample variance be the numerator of the test statistic.
   (B) larger sample variance be the numerator of the test statistic.
   (C) sample variance from the population with the smaller hypothesized variance be the numerator of the test statistic.
   (D) sample variance from the population with the larger hypothesized variance be the numerator of the test statistic.

考試日期：21 Mar. (Thur) 15:00-16:00	※准帶項目打「O」，否則打「×」					1. 需加發計算紙或答案紙請在試題內封袋備註。 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3頁，印刷份數: 65 份	計算機	課本	筆記	字典	手機平板筆電	
備註：注意事項要看!! (Scope: §ch9-11)	O	×	×	×	×	

(II) Fill-in-the-blank (Correct spelling should be used.) (20%, 10% each)

3. In the matched sample design the two production methods are tested under similar conditions (i.e., with the same workers); hence this design often leads to a smaller _____ than the independent sample design. The primary reason is that in a matched sample design, _____ between workers is eliminated because the same workers are used for both production methods.

4. Whenever independent simple random samples of sizes n_1 and n_2 are selected from two _____ populations with _____, the sampling distribution of s_1^2/s_2^2 is an F distribution with n_1-1 degrees of freedom for the numerator and n_2-1 degrees of freedom for the denominator; where s_1^2 (s_2^2) is the sample variance for the random sample of n_1 (n_2) items from population 1 (2).

(III) Short answer (20%, 4% each blank)

5. What are the *Practical Advices* for making inferences about the difference between two population means when σ_1 and σ_2 are unknown? (Filling in the blank would be sufficient.)

   Solution: There are four practical advices provided in the textbook as follows.

   (a) The interval estimation and hypothesis testing procedures are robust and can be used with _____.

   (b) In most applications, equal or nearly equal sample sizes such that the total sample size _____ can be expected to provide very good results even if the populations are not normal.

   (c) Larger sample sizes are recommended if the distributions of the populations are _____ or contain _____.

   (d) Smaller sample sizes should only be used if the analyst is satisfied that the distributions of the populations are at least _____.

(IV) Calculation (40%, 20% each)

6. **Golf Scores.** A sample of 20 tournament scores from LPGA (the women's professional golf tour) events showed a standard deviation of 2.4623 strokes, and a sample of 30 tournament scores from PGA (the men's professional golf tour) events showed a standard deviation of 2.2118. Conduct a hypothesis test for equal population variances to determine if there is any statistically significant difference in the variability of golf scores for male and female professional golfers. Use $\alpha = 0.10$.

   (a) Formulate hypotheses.

   (b) What are the value of test statistic and the rejection rule under H_0?

   (c) Make a decision using the p-value approach. What is your conclusion?

考試日期：21 Mar. (Thur) 15:00-16:00	※准帶項目打「O」, 否則打「×」					1. 需加發計算紙或答案紙請在試題內封袋備註。 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3頁，印刷份數: 65 份	計算機	課本	筆記	字典	手機平板筆電	
備註：注意事項要看!! (Scope: §ch9-11)	O	×	×	×	×	

7. **PGA Tour Scores.** Scores in the first and fourth (final) rounds for a sample of 20 golfers who competed in PGA tournaments are shown in the following table. Suppose you would like to determine if the mean score for the first round of a PGA Tour event is significantly different than the mean score for the fourth and final round. Use $\alpha = 0.10$.

Player	First Round	Final Round	Player	First Round	Final Round
Michael Letzig	70	72	Aron Price	72	72
Scott Verplank	71	72	Charles Howell	72	70
D. A. Points	70	75	Jason Dufner	70	73
Jerry Kelly	72	71	Mike Weir	70	77
Soren Hansen	70	69	Carl Pettersson	68	70
D. J. Trahan	67	67	Bo Van Pelt	68	65
Bubba Watson	71	67	Ernie Els	71	70
Reteif Goosen	68	75	Cameron Beckman	70	68
Jeff Klauk	67	73	Nick Watney	69	68
Kenny Perry	70	69	Tommy Armour III	67	71

(a) Test for a statistically significantly difference between the population means for first- and fourth-round scores. What is the hypothesis? What is the (range of) p-value? What is your conclusion?

(b) What is the point estimate of the difference between the two population means? For which round is the population mean score lower?

(c) What is the margin of error for a 90% confidence interval (CI) estimate for the difference between the population means? Could this CI have been used to test the hypothesis in (a)? Explain.

(V) Bonus (20%)

8. Let the random samples $X_1, X_2, \cdots, X_{n_1} \sim N(\mu_x, \sigma_x^2)$ and $Y_1, Y_2, \cdots, Y_{n_2} \sim N(\mu_y, \sigma_y^2)$, where σ_x and σ_y are known. Derive the interval estimate of the difierence between two population means:

$$(\bar{X} - \bar{Y}) \pm z_{\alpha/2}\sqrt{\frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n_2}},$$

where $\bar{X}, \bar{Y}$ are sample means and $1-\alpha$ is the confidence coefficient. (導出上述公式))

機率表

Area in Upper Tail	0.20	0.10	0.05	0.025	0.01	0.005
t value ($df = 19$)	0.861	1.328	1.729	2.093	2.539	2.861
F Value ($df_1 = 19, df_2 = 29$)	1.404	1.685	1.958	2.231	2.599	2.885
F Value ($df_1 = 29, df_2 = 19$)	1.452	1.763	2.077	2.402	2.855	3.221

<The blank pages at the back can be used as scratch paper. (後面空白頁可當計算紙)>