國立中正大學
113學年度碩士班招生考試
試 題

[第2節]

科目名稱	統計學
系所組別	財務金融學系

—作答注意事項—

※作答前請先核對「試題」、「試卷」與「准考證」之系所組別、科目名稱是否相符。

1. 預備鈴響時即可入場，但至考試開始鈴響前，不得翻閱試題，並不得書寫、畫記、作答。
2. 考試開始鈴響時，即可開始作答；考試結束鈴響畢，應即停止作答。
3. 入場後於考試開始40分鐘內不得離場。
4. 全部答題均須在試卷（答案卷）作答區內完成。
5. 試卷作答限用藍色或黑色筆（含鉛筆）書寫。
6. 試題須隨試卷繳還。

國立中正大學113學年度碩士班招生考試試題

科目名稱：統計學　　本科目共 7 頁　第 1 頁

系所組別：財務金融學系

第一大題：單選題（15題，每題3分，佔總分之45%）

1. Which of the following statements is correct?
   A) A Type I error occurs if we reject the null hypothesis when it is actually true.
   B) A Type I error occurs if we accept the null hypothesis when it is actually false.
   C) The power of test is the probability that we do not make either a Type I or a Type II error.
   D) A and C
   E) B and C

2. The higher the p-value, the greater the evidence that
   A) the alternative hypothesis is true.
   B) the null hypothesis is true.
   C) the significance level is high.
   D) the lower the Type I error may occur.
   E) B and D

3. We want to know the one of the two mutual fund managers indeed has better performance than the other. Their average rates of returns on investment are $\bar{x}_1 = 8$ and $\bar{x}_2 = 7.2$. Using μ_1 and μ_2 to denote the population mean, how should the null and alternative hypothesis be specified?
   A) $H_0: \mu_1 = \mu_2$ vs. $H_1: \mu_1 \neq \mu_2$
   B) $H_0: \mu_1 = \mu_2$ vs. $H_1: \mu_1 > \mu_2$
   C) $H_0: \mu_1 = \mu_2$ vs. $H_1: \mu_1 < \mu_2$
   D) B and C will conclude the same result.
   E) All specification will lead to the same result.

4. In a multiple regression, define SST as the total sum of squares and SSE as the error sum of squares. If SST = m×SSE and m > 1, what happens to the coefficient of determination if m decreases given other factors unchanged?
   A) Unchanged.
   B) Increases.
   C) Decreases.
   D) Need more information, such as the number of independent variables and the sample size, to answer.
   E) None of the above.

5. In stepwise regression analysis, the test statistic used to determine whether or not an additional variable should be included in a regression is
   A) t statistics if the sample size is smaller than 30.
   B) Z statistic if the sample size is smaller than 120.
   C) F statistic.

D) χ^2 statistic.

E) Nonparametric methods are required.

6. The lifespan, in years, of a certain computer is exponentially distributed. The probability that its lifespan exceeds four years is 0.3. Let $f(x)$ represent the density function of the computer's lifespan, in years, for $x > 0$, and zero elsewhere. Determine which of the following is an expression for $f(x)$.

A) $1-(0.3)^{-x/4}$

B) $1-(0.7)^{x/4}$

C) $-\frac{\ln 0.7}{4}(0.7)^{x/4}$

D) $-\frac{\ln 0.3}{4}(0.3)^{x/4}$

E) None of the above

7. For which of the exponential, normal, and continuous uniform distributions does doubling the mean also double the median?

A) All three

B) All but the normal

C) All but the uniform

D) All but the exponential

E) Fewer than two

8. In estimating a population mean, which of the following 95% confidence intervals has the smallest width given the same population deviation? (n = sample size)

A) n = 40

B) n = 120

C) n = 200

D) All the same

E) Need more information

9. Why is heteroscedasticity problematic?

A) The least-squares estimation procedure places more weight on observations that have small errors and variances.

B) The standard deviation of the regression coefficient will be less efficient.

C) The estimation of the regression coefficient will be bias and inconsistent

D) B and C

E) A, B and C

10. If one of the relevant explanatory variables is omitted from a multiple regression, then the estimated regression coefficient will be affected.
    A) The phenomenon is generally called a specification bias problem.
    B) Increasing the sample size may help to eliminate the problem.
    C) Using transformations in regression models may help to solve the problem.
    D) B and C
    E) A, B and C

11. Under what condition should an interaction term be introduced in the multiple regression model?
    A) When two explanatory variables are highly correlated.
    B) When the correlation coefficient between two explanatory variables is small.
    C) When the relationship between one independent variable and the dependent variable changes for different values of another independent variable.
    D) When the partial effect between one independent variable and the dependent variable is increasing or decreasing, depending on the signs of the coefficients, as the value of the given independent variable increases.
    E) To make statistical analysis fancier and easier to estimate.

Questions 12 to 14. A study was conducted to compare the mean sulfur dioxide concentrations for three cities. Independent random samples were obtained from each city and the ANOVA table are provided below:

The ANOVA Table

Source	Sum of Squares	Degrees of Freedom	Mean Squares	F-statistic	p-value
Cities (Between)	24	2	12	4.49	0.064
Error (Within)	16	6	2.667		
Total	40	8			

12. Suppose the necessary assumptions hold. Using a 10% significance level, which of the following is the correct conclusion?
    A) It appears that the population variances of concentration levels are the same for the three city populations.
    B) It appears that the population mean concentration levels are the same for the three city populations.
    C) It appears that at least one of the population variances of concentration levels is different.
    D) It appears that at least one of the population mean concentration levels is different.
    E) It appears that the population mean concentration levels are all different.

13. How is the p-value of 0.064 in the above ANOVA table obtained?
    A) The area to the right of $F = 4.49$ under the F-distribution.
    B) The area to the left of $F = -4.49$ under the F-distribution.

C) The area to the left of F = 4.49 under the F-distribution.
D) The area left of F = −4.49 or the right of F = 4.49 under the F-distribution.
E) The area between F = −4.49 and F = 4.49 under the F-distribution.

14. Consider the following four statements. Clearly select all that are correct.
A) If this study were repeated many times, the null hypothesis would be true only 6.4% of the time.
B) If this study were repeated many times, we would see an F statistic as large as or larger than that observed in about 6.4% of the repetitions.
C) If this study were repeated many times and if the population means were all equal, we would see an F statistic as large as or larger than that observed in about 6.4% of the repetitions.
D) If this study were repeated many times and if the population means were all equal, we would reject the null hypothesis in about 10% of the repetitions.
E) C and D

15 A research study compares the risk of lung disease for men who smoke to the risk of lung disease for women who smoke. The sample relative risk is 1.03 and a 95% confidence interval for the relative risk is 0.5 to 1.9. Which of the following statements is most correct about the comparison of the two groups?
A) The population risks may be the same because the interval includes the value 1.03.
B) The population risks may be the same because the interval doesn't include the value 0.
C) The population risks might differ because the interval includes the value 1.03.
D) The population risks might differ because the interval doesn't include the value 0.
E) None of the above.

第二大題：填充題（8 題，每題 5 分，佔總分之 40%）
說明：不須列示計算過程，分數無法整除時請四捨五入至 2 位小數點

1. A blood test indicates the presence of a particular disease 95% of the time when the disease is actually present. The same test indicates the presence of the disease 0.5% of the time when the disease is not present. One percent of the population actually has the disease. Calculate the probability that a person has the disease given that the test indicates the presence of the disease.

2. For a simple regression analysis, the sample deviation of the error term is equal to 5, and the sample size is equal to 10. Also, the regression sum of squares is equal to 175. What is the adjusted determination of coefficient?

Question 3-4. The coefficient of determination of a 3-independent variable multiple regression is 0.8.

3. If the number of observations is 20, then what is the F statistic for the regression?

4. Will the null hypothesis be rejected under the 5% significance level and why?

5. S&P500 and DJIA have the following average prices (index) and standard deviation in 2023:

	Mean	Standard Deviation
S&P500	4,282	229
DJIA	34,107	1,158

Using an appropriate measure to find the market which has higher volatility.

6. The skewness and kurtosis coefficients are equal to 0.18 and 4.51, respectively, for the rate of returns of TSE stock index in 2023. Compared to those in 2020 of −0.71 and 5.23, what do the numbers mean in terms of risk the investors are encountered? Explain in your own words.

7. A researcher intends to test for the racial discrimination in the mortgage loan market through the following regression

$$approve = \beta_0 + \beta_1 white + other\ \ factors$$

where *approve* is a binary variable to be explained which is equal to one if a mortgage loan to an individual was approved, and zero otherwise. The key explanatory variable is *white*, a dummy variable equal to one if the applicant was white, and zero otherwise. If the estimate of coefficient of *white* is equal to 0.12, and is significant under 5% level, how will you interpret the result? (i.e., what exactly this 0.12 means)

8. The latest voting intention poll shows the Conservatives on 38% of the vote to Labour's 46%. Will you claim that the Labour has better chance to win the election when the margin of error has been carefully considered given a sample size of 1000? Why or why not.

第三大題：問答題（3 題，每題 5 分，佔總分之 15%）

說明：務必列示計算過程，分數無法整除時請四捨五入至 2 位小數點

Let X and Y be the independent and the response variable, respectively. Four observations of (x, y) are (0.8, 5.0), (0.9, 5.3), (1.0, 5.5), (1.1, 5.4).

1. Regress Y on X and report the result.
2. What is the 95% confidence interval estimate for the mean response if the fifth observation of the independent variable is 1.2?
3. What is the 95% prediction interval for an individual response if the fifth observation of the independent variable is 1.2?

Critical Values of the *t* Distribution

		Significance Level				
1-Tailed:		.10	.05	.025	.01	.005
2-Tailed:		.20	.10	.05	.02	.01
Degrees of Freedom	1	3.078	6.314	12.706	31.821	63.657
	2	1.886	2.920	4.303	6.965	9.925
	3	1.638	2.353	3.182	4.541	5.841
	4	1.533	2.132	2.776	3.747	4.604
	5	1.476	2.015	2.571	3.365	4.032
	6	1.440	1.943	2.447	3.143	3.707
	7	1.415	1.895	2.365	2.998	3.499
	8	1.397	1.860	2.306	2.896	3.355
	9	1.383	1.833	2.262	2.821	3.250
	10	1.372	1.812	2.228	2.764	3.169
	11	1.363	1.796	2.201	2.718	3.106
	12	1.356	1.782	2.179	2.681	3.055
	13	1.350	1.771	2.160	2.650	3.012
	14	1.345	1.761	2.145	2.624	2.977
	15	1.341	1.753	2.131	2.602	2.947
	16	1.337	1.746	2.120	2.583	2.921
	17	1.333	1.740	2.110	2.567	2.898
	18	1.330	1.734	2.101	2.552	2.878
	19	1.328	1.729	2.093	2.539	2.861
	20	1.325	1.725	2.086	2.528	2.845
	21	1.323	1.721	2.080	2.518	2.831
	22	1.321	1.717	2.074	2.508	2.819
	23	1.319	1.714	2.069	2.500	2.807
	24	1.318	1.711	2.064	2.492	2.797
	25	1.316	1.708	2.060	2.485	2.787
	26	1.315	1.706	2.056	2.479	2.779
	27	1.314	1.703	2.052	2.473	2.771
	28	1.313	1.701	2.048	2.467	2.763
	29	1.311	1.699	2.045	2.462	2.756
	30	1.310	1.697	2.042	2.457	2.750
	40	1.303	1.684	2.021	2.423	2.704
	60	1.296	1.671	2.000	2.390	2.660
	90	1.291	1.662	1.987	2.368	2.632
	120	1.289	1.658	1.980	2.358	2.617
	∞	1.282	1.645	1.960	2.326	2.576

5% Critical Values of the F Distribution

Denominator Degrees of Freedom	Numerator Degrees of Freedom									
	1	2	3	4	5	6	7	8	9	10
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99
90	3.95	3.10	2.71	2.47	2.32	2.20	2.11	2.04	1.99	1.94
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83