

PART I. Multiple Choice Questions, 3 points each. You do not need to show the detailed steps, 無需列出計算過程。

Use the following to answer 1-2

A random sample of 160 commercial customers of PayMor Lumber revealed that 32 had paid their accounts within a month of billing.

1. The 90% confidence interval for the true proportion of customers who pay within a month would be
A) 0.148 to 0.252. B) 0.138 to 0.262. C) 0.144 to 0.256. D) 0.153 to 0.247.
2. How many additional customers would have to be sampled to estimate the proportion of all customers of PayMor Lumber who had paid their accounts within a month of billing to within $\pm 3\%$ with 99% reliability?
A) 523 B) 683 C) 1024 D) 1184
3. The true proportion of accounts payable with some kind of error is .20 for some manufacturing company. If an auditor randomly samples 225 accounts receivable, what is the probability that 39 or fewer will contain errors?
A) 0.0544 B) 0.2097 C) 0.1587 D) 0.1788
4. In establishing warranties on laptops, the manufacturer wants to set the limits so that few will need repair at the manufacturer's expense. On the other hand, the warranty period must be long enough to make the purchase attractive to the buyer. For a new laptop the mean number of months until repairs are needed is 36 months with a standard deviation of 3 months. Suppose the number of months until repairs follows the normal distribution. Where should the warranty limits be set so that only 1.5 percent of the laptops need repairs at the manufacturer's expense?
A) 32.78 B) 29.52 C) 32.88 D) 29.49

Use the following to answer 5-6

The lengths of brook trout caught in a certain stream have a mean of 14 inches and a standard deviation of 3 inches. Assume the lengths of brook trout follow the normal distribution.

5. What proportion of brook trout caught will be between 12 and 18 inches in length?
A) .6568 B) .6826 C) .2486 D) .4082
6. What lower limit should the State Game Commission set on length, if it is desired that 80% of the catch may be kept by fisherpersons?
A) 12.80 inches B) 12.00 inches C) 11.48 inches D) 9.22 inches
7. If tubing averages 16 defects per 100 meters, what is the probability of finding exactly 2 defects in a randomly-chosen 20-meter piece of tubing?
A) 0.2584 B) 0.2087 C) 0.4174 D) 0.5168
8. A clinic employs 9 physicians. Five of the physicians are female. Four patients arrive at once. Assuming the doctors are assigned randomly to patients, what is the probability that all of the assigned physicians are female?
A) 0.0533 B) 0.0295 C) 0.0808 D) 0.0397
9. A cell phone salesperson has kept records on the customers who visited the store. 40% of the customers who visited the store were female. Furthermore, the data show that 35% of the females who visited his store purchased a cell phone, while 20% of the males who visited his store purchased a cell phone. The salesperson has just informed us that a cell phone was purchased. What is the probability that customer was male?
A) 0.20 B) 0.35 C) 0.46 D) 0.40

10. A professor gave an exam intended to separate those who knew the material from those who didn't. Amazingly, half the students got 0 and the others got 100. There were N students, which are to be treated as a population. What is the standard deviation?
A) 2500 B) 500 C) 250 D) 50

Use the following to answer 11-12

Capital One Bank, like most other larger banks, found that using automatic teller machines (ATMs) reduces the cost of routine bank transactions. Capital One installed an ATM in the corporate offices of Fun Toy Company. The ATM is for the exclusive use of Fun's 500 employees. After several months of operation, a sample of 100 employees revealed the following use of the ATM machine by Fun employees in a month.

Number of times ATM used	0	1	2	3	4	5
Frequency	10	40	20	10	10	10

11. What is the 90 percent confidence interval for this estimate of the proportion of employees who use the ATM at least twice in a month?
A) 0.3552 to 0.6448 B) 0.4175 to 0.5825
C) 0.4261 to 0.5739 D) 0.3383 to 0.6617
12. Can Capital One be sure that more than 40 percent of the employees of Fun Toy Company will use the ATM at least twice in a month?
A) Yes B) No. C) Cannot determine from information given.
13. A researcher is interested in knowing whether the distributions of the STAT test score between Class A and Class B are the same (homogeneous). He obtains the following results :

Score	Over 70	Below 70	Sum
Class A	30	10	40
Class B	40	20	60

The p value of this test is ?

- A) $p > 10\%$ B) $5\% < p < 10\%$ C) $2.5\% < p < 5\%$ D) $1\% < p < 2.5\%$ E) $p < 1\%$
14. You are given that the probability that a salesman can successfully sale a product to a customer is 0.4. The salesman visits two customers per day. The following table gives his sale result during the past 100 days.

# of sales in a day	0	1	2
Observed days	46	38	16

To test if his sale result follows a binomial distribution, the p value of this test is:

- A) $p > 10\%$ B) $5\% < p < 10\%$ C) $2.5\% < p < 5\%$ D) $1\% < p < 2.5\%$ E) $p < 1\%$
15. Suppose a continuous random variable X has $p.d.f.$ $f(x) = \frac{1}{4}e^{-x/4}$, use Chebyshev inequality to find the lower bound for $P(-28 < X < 36)$ is given by:
A) $3/4$. B) $7/8$ C) $1/16$ D) $15/16$ E) $63/64$
16. We are interested in testing whether the means of different populations are equal using analysis of variance, which of the following statements about ANOVA is **most incorrect**?
A) The expectation of mean square error is an unbiased estimator of population error variance when the null hypothesis is not true
B) The expectation of mean square error is an unbiased estimator of population error variance if the null hypothesis is true

- C) The expectation of mean square treatment is an unbiased estimator of population error variance when the null hypothesis is not true
- D) The expectation of mean square treatment is an unbiased estimator of population error variance when the null hypothesis is true
17. Which of the following statements about R^2 is **most incorrect** in a multiple regression model?
- A) Because R^2 is interpreted assuming the model is correct, it provides no direct procedure for comparing alternative specification.
- B) Increasing the number of independent variables in the regression can never lower R^2 and is likely to raise it.
- C) When the model is constrained to have zero intercept, the ratio of regression sum of square to total sum of square (i.e., R^2) need not lie within $[0,1]$
- D) Subtracting the value of one of the independent variables from both sides of the regression equation can change R^2 while leaving the residual unaffected.
- E) The adjusted R^2 used to evaluate the performance for models with different independent variables can not be negative.
18. A financial manager ran a regression to predict the return of its stock price, in which $X_2 =$ the net profit of current year. $X_3 = 1$ if interest rate is increasing, and $=0$ otherwise. $X_4 = 1$ if interest rate is decreasing, and $=0$ otherwise. $Y =$ the return of its stock price. The data he used is based on past 42 years, and he fits two regressions:
- (I). $Y = \beta_1 + \beta_2 X_2 + \beta_3 X_2^2 + \beta_4 X_3 + \beta_5 X_4 + \varepsilon$ was fit and its $R^2 = 0.940$
- (II). $Y = \gamma_1 + \gamma_2 X_2 + \gamma_3 X_2^2 + \varepsilon^*$ was fit and its $R^2 = 0.915$.
- Determine the value of the F statistic used to test whether the interest rate effect is significant
- A) 5.4 B) 7.3 C) 7.7 D) 7.9 E) 8.3
19. The following model are fitted to 30 observations:
- Model (I): $Y = \beta_1 + \beta_2 X_2 + \varepsilon$
- Model (II): $Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \varepsilon$
- You are also given: $\sum(Y - \bar{Y})^2 = 160$, $\sum(X_2 - \bar{X}_2)^2 = 10$. In addition, for model (I), its $\hat{\beta}_2 = -2$, and for model (II), its $R^2 = 0.70$. Determine the value of the F statistic used to test whether β_3 and β_4 are jointly equal to zero
- A) Less than 15 B) At least 15, but less than 18 C) At least 18, but less than 21
- D) At least 21, but less than 24 E) At least 24
20. You are given the following two regressions, each based on a different population of data.
- Model A: $Y_i = \alpha_1 + \alpha_2 X_{2i} + \alpha_3 X_{3i} + \varepsilon_i$ where $i = 1, 2, \dots, 30$
- Model B: $Y_j = \beta_1 + \beta_2 X_{2j} + \beta_3 X_{3j} + \varepsilon_j$ where $j = 1, 2, \dots, 50$
- You assume that the variances of the two models are equal and pool the data into one model:
- Model C: $Y_k = \gamma_1 + \gamma_2 X_{2k} + \gamma_3 X_{3k} + \varepsilon_k$ where $k = 1, 2, \dots, 80$
- You also calculate the associated R^2 and error sum of square for each model. Which of the following is the F statistic for testing the hypothesis that Model A is identical (homogeneous) to Model B?

$$\begin{aligned}
 \text{A) } F_{3,74} &= \frac{(R^2_C - R^2_A - R^2_B)/3}{(R^2_A + R^2_B)/74} & \text{B) } F_{6,77} &= \frac{(R^2_C - R^2_A - R^2_B)/6}{(R^2_A + R^2_B)/77} \\
 \text{C) } F_{6,74} &= \frac{(ESS_C - ESS_A - ESS_B)/6}{(ESS_A + ESS_B)/74} & \text{D) } F_{6,77} &= \frac{(ESS_C - ESS_A - ESS_B)/6}{(ESS_A + ESS_B)/77} \\
 \text{E) } F_{3,74} &= \frac{(ESS_C - ESS_A - ESS_B)/3}{(ESS_A + ESS_B)/74}
 \end{aligned}$$

21. For the model $Y_i = \alpha + \beta X_i + \varepsilon_i$, where $i = 1, 2, 3, \dots, 10$, you are given that

(i). $X_i = \begin{cases} 1, & \text{if the } i\text{th individual belongs to the specified group} \\ 0, & \text{otherwise} \end{cases}$

(ii) 40% of the individuals belong to the specified group.

(iii). The least square estimate of β is 4

(iv). $\sum (Y_i - \hat{\alpha} - \hat{\beta} X_i)^2 = 92$

Calculate the t statistic for testing $H_0: \beta = 0$

- A) 0.9 B) 1.2 C) 1.5 D) 1.8 E) 2.1

22. You are given that the total sum of square for a regression line is $SST = 160$. If we include one explanatory variable X_1 , the associated regression sum of square is $SSR(X_1) = 124$. If we include only explanatory variable X_2 , the regression sum of square is $SSR(X_2) = 136$. If we include both X_1 and X_2 , the regression sum of square is $SSR(X_1, X_2) = 142$. What is the partial coefficient of determination of Y and X_2 , given X_1 , i.e., $r_{Y2.1}^2$

- A) $\frac{1}{2}$ B) $\frac{1}{3}$ C) $\frac{12}{160}$ D) $\frac{136}{160}$ E) None of the above.

23. The residuals for a linear regression based on time series data is given as.

t	1	2	3	4	5
$\hat{\varepsilon}_t$	-0.6	-0.7	2.3	0	-1

You are given that the associated Durbin-Watson statistic is 2.143. Find the lag-one serial correlation coefficient for the residuals, which is approximately

- A) Less than -0.2 B) At least -0.2, but less than -0.1 C) At least -0.1, but less than 0.0
 D) At least 0.0, but less than 0.1 E) At least 0.1

24. You fit the following model to 48 observations: $Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \varepsilon$

You are given the following ANOVA table:

Source of Variation	Degree of Freedom	Sum of Square
Regression	3	103.66
Error	44	69.20

Then the adjusted R^2 of this regression model is

- A) 0.57 B) 0.58 C) 0.59 D) 0.60 E) 0.61

PART II. You must show detailed steps to get the credits. (需列出計算過程才能獲得分數).

1. (8 points) Suppose that we want to test the null hypothesis that the mean of a normal population with $\sigma^2 = 1$ is μ_0 against the alternative hypothesis that it is μ_1 , where $\mu_1 > \mu_0$.
 - a. Find the value of k such that $\bar{x} > k$ provides a critical region of size $\alpha = 0.015$ for a random sample of size n .
 - b. Determine the minimum sample size needed to test the null hypothesis $\mu_0 = 10$ against the alternative hypothesis $\mu_1 = 10.5$ with $\beta \leq 0.017$.

2. (6 points) Suppose that independent random samples of size n from two normal populations with the known variances σ_1^2 and σ_2^2 are to be used to test the null hypothesis $\mu_1 - \mu_2 = \varepsilon$ against the alternative hypothesis $\mu_1 - \mu_2 = \varepsilon'$ and that the probabilities of type I and type II errors are to have the preassigned values α and β . Show that the required sample size is given by

$$n = \frac{(\sigma_1^2 + \sigma_2^2)(z_\alpha + z_\beta)^2}{(\varepsilon - \varepsilon')^2}$$

3. A bank likes to know whether the amount of loans granted to customers is associated with branch offices and customer groups (Group A: high-risk customers; Group B: low-risk customers). **Three replicates** are recorded in each branch-group treatment combination. The mean values of branch-group combinations are given as below (thousands) :

	Branch A	Branch B	Branch C	Row mean
Group A	80	60	70	70
Group B	100	70	70	80
Column mean	90	65	70	75

You are given that the total sum of square is equal to 3750.

- (a). (8 points) Calculate the two-way ANOVA table **without the interaction term**.

Source	SS	df	MS	F
Branch				
Group				
Error				
Total	3750			

- (b). (6 points) Calculate the two-way ANOVA table **with the interaction term**.

Source	SS	df	MS	F
Branch				
Group				
Interaction				
Error				
Total	3750			

國立中正大學 102 學年度碩士班招生考試試題

系所別：財務金融學系

科目：統計學

第 3 節

第 6 頁，共 6 頁

附表 1 z 分配表

$P(0 < Z < z) = \alpha$

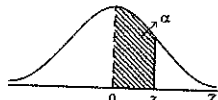


Table with columns for z (0.00 to 3.5) and rows for alpha values (0.00 to 0.99). It provides cumulative probabilities for the standard normal distribution.

附表 2 t 分配臨界值表

$P(t_{df} > t_{\alpha, df}) = \alpha$

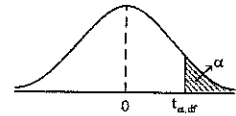


Table with columns for df (1 to 1000) and rows for alpha values (0.25 to 0.0005). It provides critical values for the t-distribution.

附表 3 卡方分配臨界值表

$P(\chi^2_{df} > \chi^2_{\alpha, df}) = \alpha$

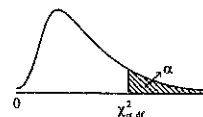


Table with columns for df (1 to 100) and rows for alpha values (0.995 to 0.005). It provides critical values for the chi-square distribution.