國立中正大學
114學年度碩士班招生考試
試 題

[第4節]

科目名稱	統計學
系所組別	財務金融學系

**PART I. Multiple Choice**（單選題 25 題，每題 3 分）(75%)

1. Which of the following statement about panel and pooled cross-sectional data is correct?

    A. A panel data set consists of data on different cross-sectional units over a given period of time while a pooled data set consists of data on the same cross-sectional units over a given period of time.

    B. A panel data set consists of data on the same cross-sectional units over a given period of time while a pooled data set consists of data on different cross-sectional units over a given period of time.

    C. A panel data consists of data on a single variable measured at a given point in time while a pooled data set consists of data on the same cross-sectional units over a given period of time.

    D. A panel data set consists of data on a single variable measured at a given point intime while a pooled data set consists of data on more than one variable at a given point in time.

2. Order the following histograms from least to most variability.
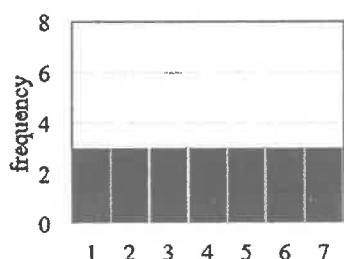
(i)　　　　　　　　　　(ii)　　　　　　　　　　(iii)
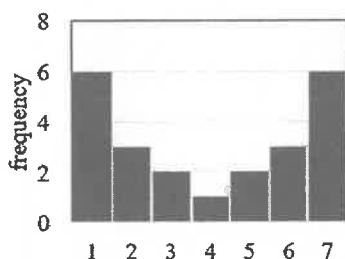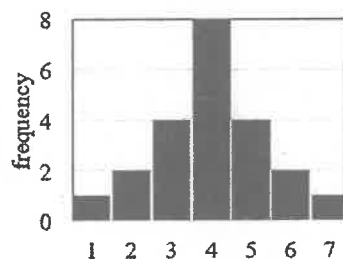
A. (i), (ii), (iii)　　　B. (iii), (i), (ii)　　　C. (ii), (iii), (i)

D. (ii), (i), (iii)　　　E. none of the above.

Use the histogram to answer Questions 3 and 4:

Length of time on hold in minutes
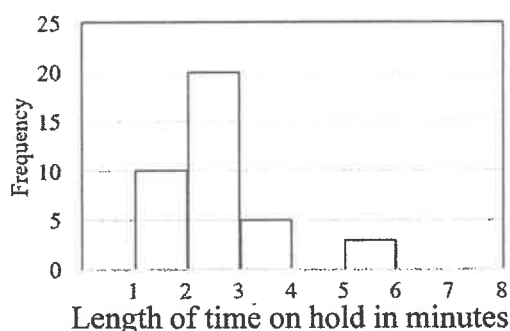
3. The histogram above displays the distribution of the length of time on hold, for a collection of customers, calling a repair call center. Select the true statement.

    (i) The distribution is right-skewed.

    (ii) The distribution is left-skewed.

    (iii) Most callers waited on hold for less than three minutes.

    (iv) Most callers waited on hold for at least three minutes.

    (v) The distribution shows that the data was highly variable with some callers waiting on hold as many as 20 minutes.

   A. (i), (iii)　　　　　　　　B. (i), (iv), (v)　　　　　　　C. (ii), (iii)

   D. (i), (iii), (v)　　　　　　E. none of the above

4. What do you expect for the values of the mean, median, and the mode?

   A. mean = median　　　　　　　　　　　B. mean > median

   C. mean < median　　　　　　　　　　　D. mean = mode

5. In many financial studies, researchers replaced the Normal distribution with Student's t distribution to capture the dynamics of stock returns because the t distribution

   A. is asymmetric and is negatively skewed.　　B. is asymmetric and is positively skewed.

   C. has fat tails.　　　　　　　　　　　　　D. has thin tails.

   E. is a sampling distribution while normal distribution is a population distribution.

6. A political scientist is studying voters in California. It is appropriate for him to use a mean to describe

   A. the age of a typical voter.　　　　　　　B. the party affiliation of a typical voter.

   C. the sex of a typical voter.　　　　　　　D. the county of residence of a typical voter.

7. If population A has a larger standard deviation than population B, which of the following is NOT true?

   A. Population B has a smaller variance than population A.

   B. The mean of a sample of 20 from population A has a larger standard deviation than the mean of a sample of 20 from population B.

   C. A typical observation from population A will be farther from the mean of population A than a typical observation from B will be from the mean of population B.

   D. The mean of a sample from population A will on average be larger than the mean of a sample from population B.

8. At a private university with 20,000 students, admission administrators find that the enrollment for the current year consist of 64% in-state students and 36% out-of-state students. If a random sample of 400 is taken, what is the probability that the sample mean out-of-state students is between 33% and 41%?

   A. 0.8757　　　　　　B. 0.5852　　　　　　C. 0.4312　　　　　　D. 0.6213

9. A company prices its hurricane insurance using the following assumptions: i. In any calendar year, there can be at most one hurricane. ii. In any calendar year, the probability of a hurricane is 0.1. iii. The number of hurricanes in any calendar year is independent of the number of hurricanes in any other calendar year. Calculate the probability that there are fewer than 3 hurricanes in a 10-year period.

   A. 0.9872　　　　　　B. 0.9298　　　　　　C. 0.0702　　　　　　D. 0.0128

10. The time to travel from A to B through city center (road R1) is normally distributed with a mean of 20 minutes and a standard deviation of 5 minutes. The time to travel from A to B through a new ring road (road R2) is normally distributed with a mean of 15 minutes and a standard deviation of 8 minutes. You have 17 minutes to travel from A to B on an important appointment. What is your correct decision to choose between R1 and R2?

A. R2, since the probability of reaching the destination over 17 minutes is higher for R1 (59.87%) compared to R2 (27.43%).

B. R1, since the probability of reaching the destination over 17 minutes is higher for R2 (59.87%) compared to R1 (27.43%).

C. R1, since the probability of reaching the destination in 17 minutes is higher for R1 (59.87%) compared to R2 (27.43%).

D. R2, since the probability of reaching the destination in 17 minutes is higher for R2 (59.87%) compared to R1 (27.43%).

Question 11~12

11. You are conducting a one-sided test of the null hypothesis that the population mean is 532 versus the alternative that the population mean is less than 532. If the sample mean is 529 and the p-value is 0.03, which of the following statements is true?

A. There is a 0.03 probability that the population mean is smaller than 529.

B. There is a 0.03 probability that the population mean is smaller than 532.

C. The probability of observing a sample mean smaller than 529 as the population mean is 532 is 0.03.

D. If the significance level is 0.05, you will accept the null hypothesis.

12. Using the same data as in question 11 but now conducting a two-sided test of the null hypothesis that the population mean is 532, which of the follower(s) is/are correct

(i)   If the significance level is 5%, you will accept the null hypothesis.

(ii) If the significance level is 5%, you will reject the null hypothesis.

(iii) If the significance level is 10%, you will accept the null hypothesis.

(iv) If the significance level is 10%, you will reject the null hypothesis.

A. (i), (iv)   B. (ii), (iii)   C. (i), (iii)

D. (ii), (iv)   E. none of the above

13. A chi-square test involves a set of counts called "expected counts." What are the expected counts?

A. Hypothetical counts that would occur if the alternative hypothesis were true.

B. Hypothetical counts that would occur on average if the null hypothesis were true.

C. The actual counts that did occur in the observed data.

D. The long-run counts that would be expected if the observed counts are representative

14. A statistically significant relationship between two categorical variables is illustrated in the sample as one that

A. is small enough that it is likely to have occurred in the observed sample even if there is no relationship in the population.

B. is small enough that it is unlikely to have occurred in the observed sample if there is no relationship in the population.

C. is large enough that it is likely to have occurred in the observed sample even if there is no relationship in the population.

D. is large enough that it is unlikely to have occurred in the observed sample if there is no relationship in the population.

15. If a relationship between two variables is called statistically significant, it means the investigators think the variables are
   A. not related in the population represented by the sample.
   B. not related due to sampling errors.
   C. related in the population represented by the sample.
   D. related in the sample due to chance alone.

16. A sample of 150 new cell phones produced by HTC found that 12 had cosmetic flaws. A 90% confidence interval for the proportion of all new HTC phones with cosmetic flaws is 0.044 to 0.116. Which statement below provides the correct interpretation of this confidence interval?
   A. There is a 90% chance that the proportion of new phones that have cosmetic flaws is between 0.044 and 0.116.
   B. There is at least a 4.4% chance that a new phone will have a cosmetic flaw.
   C. A sample of 150 phones will have no more than 11.6% with cosmetic flaws.
   D. If you selected a very large number of samples and constructed a confidence interval for each, 90% of these intervals would include the proportion of all new phones with cosmetic flaws.

Question 17~18

17. A research firm conducted a survey to determine the mean amount smokers spend on cigarettes during a day. A sample of 100 smokers revealed that the sample mean is $5 and sample standard deviation is $2. Assume that the sample was drawn from a normal population. The upper limit of the 95% confidence interval for the population mean is most closed to
   A. 5.1　　　　　B. 5.6　　　　　C. 6.2　　　　　D. 5.4

18. When a 99% confidence interval is calculated instead of a 95% confidence interval with the sample size being the same, the maximum error of estimate will be
   A. larger　　　　B. smaller　　　　C. the same　　　　D. it cannot be determined

19. Which of the following statement is true?
   A. A variable has a causal effect on another variable if both variables increase or decrease simultaneously.
   B. The notion of 'ceteris paribus' plays an important role in causal analysis.
   C. Difficulty in inferring causality disappears when studying data at fairly high levels of aggregation.
   D. The problem of inferring causality arises if experimental data is used for analysis.

20. Which of the following is a nonlinear regression model?
   A. $y = \beta_0 + \beta_1 x^{1/2} + \varepsilon$　　　　B. $\log y = \beta_0 + \beta_1 \log x + \varepsilon$　　　　C. $y = 1/(\beta_0 + \beta_1 x) + \varepsilon$
   D. all of the above　　　　E. none of the above

21. In the regression of y on x, the error term exhibits heteroskedasticity if
    A. It has a constant variance.　　　　B. Var($y|x$) is a function of x.
    C. x is a function y while at the same time y is a function of x.
    D. x is correlated with the error term.　　　　E. all of the above except A.

22. Which of the following statement is true?
    A. The term "linear" in a multiple linear regression model means that the equation is linear in independent variables.
    B. The key assumption for the general multiple regression model is that all factors in the unobserved error term be correlated with the explanatory variables.
    C. The coefficient of determination decreases when an independent variable is added to a multiple regression model.
    D. An explanatory variable is said to be exogeneous if it is correlated with the error term.
    E. A larger error variance makes it difficult to estimate the partial effect of any of the independent variables on the dependent variable.

23. If $\hat{\beta}_j$, an unbiased estimator of β_j, is consistent, then the distribution of $\hat{\beta}_j$ _____ as the sample size grows.
    A. becomes more and more loosely distributed around β_j
    B. becomes more and more tightly distributed around β_j
    C. tends toward a standard normal distribution
    D. remains unaffected
    E. collapses to a single value of zero

24. A change in the unit of measurement of the dependent variable in a model does NOT lead to a change in
    A. The slope coefficient
    B. The standard error of the regression
    C. The standard error of the slope coefficient
    D. The goodness-of-fit of the regression
    E. The confidence interval of the regression

25. The following simple model is used to determine the annual savings of an individual on the basis of his annual income and education:
$$Saving = \beta_0 + \beta_1 Edu + \beta_2 Inc + \varepsilon$$
where the variable "*Edu*" takes a value of 1 if the person is a collage graduate and 0 otherwise, and the variable "*Inc*" measures the income of the individual. To interpret β_1, the benchmark group in this model is
    A. The group of college graduates.
    B. The group of non-college graduates.
    C. The group of individuals with high income.
    D. The group of individuals with low income.

PART II. Essay Questions（問答計算題，共 3 題，需詳列計算過程）(25%)

1. An auto insurance company is implementing a new bonus system. In each month, if a policyholder does not have an accident, he or she will receive a $5 cash-back bonus from the insurer. Among the 1000 policyholders of the auto insurance company, 400 are classified as low-risk drivers and 600 are classified as high-risk drivers. In each month, the probability of zero accidents for high-risk drivers is 0.8 and the probabilities for low-risk drivers is 0.9. Calculate the expected bonus payment from the insurer to the 1000 policyholders in one year. (5%)

2. The table below shows the ratings given by 1200 students in an introductory statistics course. There were two instructors for the course.

| Instructor | Rating | | Total |
	High Quality	Not High Quality	
Peter	490	210	700
Jane	280	220	500
Total	770	430	1200

(1) Explain which instructor has the higher percentage of high-quality teaching rating? (2%)

(2) For further investigation, another question on the evaluation was given by asking whether the students considered themselves as being good at math or not. The tables below show the ratings for each instructor for these two groups of students.

Good at Math

| Instructor | Rating | | Total |
	High Quality	Not High Quality	
Peter	480	120	600
Jane	180	20	200
Total	660	140	800

Not Good at Math

| Instructor | Rating | | Total |
	High Quality	Not High Quality	
Peter	10	90	100
Jane	100	200	300
Total	110	290	400

Among those students who consider themselves good at math, which instructor has the higher percentage of high-quality teaching ratings? Among those students who consider themselves not good at math, which instructor has the higher percentage of high-quality teaching ratings? (4%)

(3) Provide an explanation for the apparent contradiction between (1) and (2). (4%)

3. A man purchases a life insurance policy on his 40th birthday. The policy will pay 5,000 only if he dies before 50th birthday and will pay 0 otherwise. The length of lifetime, in years from birth, of a male born the same year as the insured has the cumulative distribution function as

$$F(t) = \begin{cases} 0, & \text{for } t \le 0 \\ 1 - \exp\left(\dfrac{1-1.1^t}{1000}\right), & \text{for } t > 0 \end{cases}$$

Calculate the expected payment to the man under this policy. (10%)

STANDARD NORMAL DISTRIBUTION: Table Values Represent AREA to the LEFT of the Z score.

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.50000	.50399	.50798	.51197	.51595	.51994	.52392	.52790	.53188	.53586
0.1	.53983	.54380	.54776	.55172	.55567	.55962	.56356	.56749	.57142	.57535
0.2	.57926	.58317	.58706	.59095	.59483	.59871	.60257	.60642	.61026	.61409
0.3	.61791	.62172	.62552	.62930	.63307	.63683	.64058	.64431	.64803	.65173
0.4	.65542	.65910	.66276	.66640	.67003	.67364	.67724	.68082	.68439	.68793
0.5	.69146	.69497	.69847	.70194	.70540	.70884	.71226	.71566	.71904	.72240
0.6	.72575	.72907	.73237	.73565	.73891	.74215	.74537	.74857	.75175	.75490
0.7	.75804	.76115	.76424	.76730	.77035	.77337	.77637	.77935	.78230	.78524
0.8	.78814	.79103	.79389	.79673	.79955	.80234	.80511	.80785	.81057	.81327
0.9	.81594	.81859	.82121	.82381	.82639	.82894	.83147	.83398	.83646	.83891
1.0	.84134	.84375	.84614	.84849	.85083	.85314	.85543	.85769	.85993	.86214
1.1	.86433	.86650	.86864	.87076	.87286	.87493	.87698	.87900	.88100	.88298
1.2	.88493	.88686	.88877	.89065	.89251	.89435	.89617	.89796	.89973	.90147
1.3	.90320	.90490	.90658	.90824	.90988	.91149	.91309	.91466	.91621	.91774
1.4	.91924	.92073	.92220	.92364	.92507	.92647	.92785	.92922	.93056	.93189
1.5	.93319	.93448	.93574	.93699	.93822	.93943	.94062	.94179	.94295	.94408
1.6	.94520	.94630	.94738	.94845	.94950	.95053	.95154	.95254	.95352	.95449
1.7	.95543	.95637	.95728	.95818	.95907	.95994	.96080	.96164	.96246	.96327
1.8	.96407	.96485	.96562	.96638	.96712	.96784	.96856	.96926	.96995	.97062
1.9	.97128	.97193	.97257	.97320	.97381	.97441	.97500	.97558	.97615	.97670
2.0	.97725	.97778	.97831	.97882	.97932	.97982	.98030	.98077	.98124	.98169
2.1	.98214	.98257	.98300	.98341	.98382	.98422	.98461	.98500	.98537	.98574
2.2	.98610	.98645	.98679	.98713	.98745	.98778	.98809	.98840	.98870	.98899
2.3	.98928	.98956	.98983	.99010	.99036	.99061	.99086	.99111	.99134	.99158
2.4	.99180	.99202	.99224	.99245	.99266	.99286	.99305	.99324	.99343	.99361
2.5	.99379	.99396	.99413	.99430	.99446	.99461	.99477	.99492	.99506	.99520
2.6	.99534	.99547	.99560	.99573	.99585	.99598	.99609	.99621	.99632	.99643
2.7	.99653	.99664	.99674	.99683	.99693	.99702	.99711	.99720	.99728	.99736
2.8	.99744	.99752	.99760	.99767	.99774	.99781	.99788	.99795	.99801	.99807
2.9	.99813	.99819	.99825	.99831	.99836	.99841	.99846	.99851	.99856	.99861
3.0	.99865	.99869	.99874	.99878	.99882	.99886	.99889	.99893	.99896	.99900
3.1	.99903	.99906	.99910	.99913	.99916	.99918	.99921	.99924	.99926	.99929
3.2	.99931	.99934	.99936	.99938	.99940	.99942	.99944	.99946	.99948	.99950
3.3	.99952	.99953	.99955	.99957	.99958	.99960	.99961	.99962	.99964	.99965
3.4	.99966	.99968	.99969	.99970	.99971	.99972	.99973	.99974	.99975	.99976
3.5	.99977	.99978	.99978	.99979	.99980	.99981	.99981	.99982	.99983	.99983
3.6	.99984	.99985	.99985	.99986	.99986	.99987	.99987	.99988	.99988	.99989
3.7	.99989	.99990	.99990	.99990	.99991	.99991	.99992	.99992	.99992	.99992
3.8	.99993	.99993	.99993	.99994	.99994	.99994	.99994	.99995	.99995	.99995
3.9	.99995	.99995	.99996	.99996	.99996	.99996	.99996	.99996	.99997	.99997