國立臺北大學 109 學年度碩士班一般入學考試試題

系（所）組別：都市計劃研究所甲組、乙組

科　　目：統計學

☑可 ☐不可使用計算機

一、Multiple choice (單選題)(每題 2 分)

1. What is the data type of the drainage area (集水區面積; unit: square meters)?
   A.Count B.Interval C.Nominal D.Ordinal E.Ratio

2. What is the data type of the residential location (住宅區位) such as metropolitan area, suburban and urban?
   A.Count B.Interval C.Nominal D.Ordinal E.Ratio

3. What is the data type of the preference of living in a condo measured in a five level Likert scale (Like very much, Like, Fair, Dislike, Very dislike)?
   A.Count B.Interval C.Nominal D.Ordinal E.Ratio

4. What is the data type of the temperature (unit: °C)?
   A.Count B.Interval C.Nominal D.Ordinal E.Ratio

5. Which is the most appropriate statistical chart for displaying data for drainage area (unit: square meters)?
   A.Bar plot B.Line plot C.Histogram D.Pie chart E.Scatter plot

6. Which is the most appropriate statistical chart to display the data for the residential location versus the type of the residential location?
   A.Bar chart B.Box plot C.Histogram D.Scatter plot E.Stem and leaf plot

7. Which is the most appropriate statistical chart to display the data for the residential location versus the house value (unit: thousand dollars)?
   A.Bar chart B.Box plot C.Histogram D.Scatter plot E.Stem and leaf plot

8. Which is the most appropriate statistical chart to display the data for house value (unit: thousand dollars) and the household income (unit: thousand dollars)?
   A.Bar chart B.Box plot C.Histogram D.Scatter plot E.Stem and leaf plot

9. Which is the most appropriate statistic to summary the data for the residential location?
   A.Interquartile B.Mean C.Median D.Range E.None of the above

10. Which of the following is the least appropriate statistical statistic to summary the central tendency of the data for household income (unit: thousand dollars)?
    A.Mean B.Median C.Midrange D.Mode E.None of the above

11. Which is the lease appropriate statistic to summary the variability of the data for household income (unit: thousand dollars)?
    A.Correlation B.Interquartile C.Standard deviation D.Range E.Variance

12. Which of the following is the most appropriate statistic to summary the association between the data for house value (unit: thousand dollars) and household income (unit: thousand dollars)?
    A.Correlation B.Interquartile C.Mean D.Relative frequency E.Skewness

13. Which of the following is the most appropriate statistic to examine the shape of the distribution for the data of the house value (unit: thousand dollars)?
    A.Correlation B.Interquartile C.Mean D.Relative frequency E.Skewness

14. Which of the following data can be summarized by the central tendency?
    A.Residential location B.The commuting distance from home to school C.The method of commute
    D.Type of districts E.None of the above

15. A survey for the method of commute to school for 200 selected students commuters in a particular city is conducted. The data is summarized as follows:

	Method of Commutes		
Sex	Metro	Bus	Walk
Female	60	20	40
Male	40	30	10

Determine the probability of the students who commute to school using metro or bus.

A.0.1 B.0.2 C.0.25 D.0.5 E.0.75

16. (Continued Question 15.) Given the respondents who were female, Determine the probability of the students who walked to school?

A.0.05 B.0.2 C.0.25 D.1/3 E.0.5

17. What of the following properties are not used to describe event A and event B are independent?

A.$P[A|B] = P[A]$ B.$P[B|A] = P[B]$ C.$P[A \cap B] = 0$ D.$P[A \cap B] = P[A]P[B]$

E.$P[A \cup B] = P[A] + P[B](1 - P[A])$

18. In a particular city, 20% of the people subscribe to the morning newspaper, 30% subscribe to the evening newspaper, and 10% subscribe to both. Determine the probability that an individual from this city subscribe at least one of the two newspaper?

A.0.1 B.0.2 C.0.3 D.0.4 E.2/3

19. (Continued Question 18.) What is the probability that the evening subscribers do not subscribe to the morning newspaper?

A.0.1 B.0.2 C.0.3 D.0.4 E.2/3

20. (Continued Question 18.) Given that a person subscribes to the evening newspaper, what is the probability that this person do not subscribe to the morning newspaper?

A.0.1 B.0.2 C.0.3 D.0.4 E.2/3

21. What of the following state is incorrect?

A.A random variable has to have numerical values.

B.The probability has to be greater than 0 but not equal to 0.

C.The probability has to be smaller than or equal to 1.

D.The probability of all possible values of a random variable has to equal 1.

E.The probability of mutually exclusive events equals 0.

22. Suppose there are three tax returns. An IRS auditor tries to audit them. Let R denote the event of refund and N denote the event of non-refund (owe). Define the random variable X to equal the number of tax refund. What of the following statement is correct?

A.The outcome of this experiment is RRR, NRR, NNR, NNN. B.The possible outcome of X is 0, 1, 2, 3.

C.P[RRR]=1/4 D.P[NRR]=2/4 E.P[NNR]=1/4

23. The probability distribution of random variable, X, is defined as follows:

X	0	1	2	3	4
Probability	?	0.3	0.1	0.3	0.3

What is the missing value?

A.0 B.0.1 C.0.2 D.0.3 E.0.4

24. Airline overbooking is a common practice. At a recent survey, 5% of customers making the reservation in the China Air do not show up for the flight, 10% of customers making the reservation in Eva Air do not show up for the flight, 20% of customers making the reservation in Eagle Air do not show up for the flight. The market shares for China Air, Eva Air and Eagle Air are 35%, 40% and 25%. Given a customer who made the reservation but didn't show up for the flight, what is the probability that this customer booked Eagle Air?

A.0.05 B.0.2 C.0.4651 D.0.5714 E.0.8

25. Owing to the housing shortage, the government would like to know the rate of the vacant house in the Taipei metropolitan area. From the past records, the rate of the vacant house is 20%. The government randomly selects 200 houses and assigns an agent to visit the house. Let X denote the number of vacant houses. What of the following condition is not necessary for X to be a binomial random variable?
A.A house is considered as a trial. B.Each trial has only two outcomes such as "vacant" and "non-vacant".
C.Each trial has the same probability of the vacant house. D.Trials are independent. E.Trials are uncorrelated.

26. (Continued Question 25.) Assume X is a binomial random variable. What of the following statement is incorrect?
A.The probability that all of the house are non-vacant is 0.2^{200}. B.The mean number of vacant houses equals 40.
C.The mean number of non-vacant houses equals 160. D.The standard deviation of X equals 5.66.
E.The variance of X equals 32.

27. What of the following statement is incorrect for a normal random variable X?
A.The mean and median of X are equal. B.The mean and mode of X are equal.
C.The normal curve is symmetry about mean. D.The probability above the mean and below the mean has to be equal.
E.The standard deviation of X has to equal 1.

28. A survey describes the distribution of total sleep time (in hours per day) among college students as approximately normal with mean $\mu = 7$ and standard deviation $\sigma = 1$. What is the 1.5 percentile of the distribution of total sleep time among all college students?
A.3.55 B.4.83 C.5 D.6.12 E.9

29. (Continued Question 28.) If the survey asks this question to a random sample of n =100 college students, what is the probability that $\frac{41}{6} \leq \bar{x} \leq \frac{43}{6}$ (or in words, that the sample mean differs from the population mean by at most 10 minutes)?
A.0.1025 B.0.2186 C.0.5468 D.0.9050 E.0.9876

30. What of the following condition is not necessary for the Central Limit Theorem?
A.Each observation in the sample has equal mean. B.Each observation in the sample has equal variance.
C.Each observation in the sample has to have a normal distribution. D.Observations in the sample are independent.
E.The sample size of the sample has to be large.

31. What of the following condition is not necessary for a random sample?
A.Each observation has equal mean. B.Each observation has equal variance.
C.Each observation has the same distribution. D.Observations are independent.
E.The sample size of this random sample has to be large

32. What of the following properties is not true for the sample mean?
A.The distribution of the sample mean is approximately normal when the sample size is large.
B.The mean of the sample mean equals the population mean.
C.The observations used to compute the sample mean are independent.
D.The population mean can be estimated by it.
E.The variance of the sample mean has to equal the population variance.

33. As the sample size increases, the
A.Population mean increases. B.Sample mean increases. C.Standard deviation of the population decreases.
D.Standard error of the sample mean decreases. E.Standard error of the sample mean increases.

34. The Tourism Bureau of Taiwan would like to estimate the mean amount of money spent by a tourist to within $1000 with 95% confidence. If the amount of money spent by tourists is considered to be normally distributed with a standard deviation of $2550, what sample size would be necessary for the Chamber of Commerce to meet its objective in estimating the mean spending amount?
A.9 B.16 C.25 D.44 E.62

35. What of the following properties is not true for the *t* distribution?
A.As the degrees of freedom increase, the t distribution is nearly identical to the standard normal distribution.
B.The mean of the t distribution equals 0.
C.The shape of t distribution depends on the degrees of freedom.
D.The t distribution has thinner tails than the standard normal distribution.
E.The t distribution is symmetric about 0.

36. What of the following properties is correct for the confidence interval?
A.The confidence interval has the same length regardless of the random samples.
B.The length of confidence interval tends to be shorter as the confidence level gets larger.
C.The length of confidence interval tends to be shorter as the mean gets larger.
D.The length of confidence interval tends to be shorter as the sample size increases.
E.The length of confidence interval tends to be shorter as the variance gets larger, the length of confidence interval tends to be shorter.

37. Which is the least appropriate statistical test to evaluate the association between the data for the residential location and house value (unit: thousand dollars)?
A.Chi-square test　B.*F* test　C.Mann-Whitney test　D.One-way ANOVA　E.Two-independent sample *T* test

38. What of the following assumption is not required when building a regression model?
A.The mean value of each error term equals to 0.　B.The observations are independent.
C.The observations have the common variance.　D.The response variable has a common normal distribution.
E.The response variable has to be linear associated with the explanatory variable.

39. A real-estate company would like to evaluate whether the performance in terms of the number of houses sold by its agents for year 2019 is better than 2020. He randomly selects some agents and compare their performance in 2019 and 2020. Which is the most appropriate statistical test to answer this question?
A.Chi-square test　B.*F* test　C.One-way ANOVA　D.Paired T test　E.Two-independent sample *T* test

40. Consider the problem of time use for housework for female and male. The time is reported by the average minutes per day spent on cooking and washing up. The following provides some summary statistics:

	N	Sample mean	Sample standard deviation
Women	733	37	16
Men	1219	23	32
95% CI for difference: (11.9, 16.1)			

Which of the following statement is the most appropriate statement?
A.The average time use for women is significantly different than that for men.
B.The average time use for women is significantly higher than that for men.
C.The average time use for women is significantly lower than that for men.
D.The average time use for women is similar to that for men.
E.The information is not sufficient to make any inference.

41. Suppose a potential customer would like to evaluate where there exist difference in house price (unit: thousand dollars) between urban and rural areas. Let μ_1 and μ_2 denote the mean of the house price for urban and rural areas. Also, let $\bar{x}_1$ and $\bar{x}_2$ denote the sample mean of the house price for urban and rural areas. Which of the following statement is correct?
A.$H_0: \mu_1 \geq \mu_2$ versus $H_a: \mu_1 < \mu_2$.　B.$H_0: \mu_1 = \mu_2$ versus $H_a: \mu_1 \neq \mu_2$.　C.$H_0: \mu_1 < \mu_2$ versus $H_a: \mu_1 \geq \mu_2$.
D.$H_0: \bar{x}_1 = \bar{x}_2$ versus $H_a: \bar{x}_1 \neq \bar{x}_2$.　E.$H_0: \bar{x}_1 \leq \bar{x}_2$ versus $H_a: \bar{x}_1 > \bar{x}_2$.

42. Which is the correct statement of the null hypothesis to evaluate the association between the method of commute and gender?
A.H_0: the method of commute and gender are not significantly associated.
B.H_0: the method of commute and gender are dependent.
C.H_0: the method of commute and gender are independent.
D.H_0: the method of commute and gender are significantly associated.
E.H_0: the method of commute and gender are uncorrelated.

43. Which is the most appropriate test to evaluate the association between the method of commute and gender?
A.Chi-square test B.F test C.Mann-Whitney test D.Paired T test E.Two-independent sample T test

44. Assume the significant level is 0.05. When H_0 is true in a given test, what is the most range for the p-value?
A.Equals to 0.05. B.Greater than or equal to 0.95. C.Greater than or equal to 0.05.
D.Less than 0.05. E.Less than 0.95.

45. Which of the following statement is correct?
A.A 95% confidence interval for μ =population mean IQ is (96, 110). So, in the test of $H_0: \mu = 100$ versus $H_a: \mu \neq 100$, the p-value is greater than 0.05.
B.For a fixed significance level α, the probability of a Type II error increases when the same size increases.
C.If we reject $H_0: \mu = 0$ in a study about change in weight on a new diet using a significance level of $\alpha = 0.05$, then we also reject it using $\alpha = 0.01$.
D.P[Type II error]=1-P[Type I error]
E.The p value is defined as the probability that H_0 is true.

46. An industrial plant claims to discharge no more than 1000 gallons of wastewater per hour, on the average, into a neighboring lake. An environmental action group took some random samples from the lake over a period of time to analyze the true mean discharge of wastewater per hour. Consider a one-sided hypotheses and use a significance level of $\alpha = 0.05$. What of the following statement is correct?
A.If the plant is exceeding the limit, there is only a 5% chance that we will conclude that they are not exceeding the limit.
B.If the plant is not exceeding the limit, there is only 5% chance that we will conclude they are exceeding the limit.
C.If we reject H_0, the probability that it is actually true is 0.05.
D.The probability that the sample mean equals exactly the observed value would equal 0.05 if H_0 was true.
E.All of the above.

47. Let p_F and p_M denote the probability of possessing house properties for female and male in Taiwan. Suppose that a 95% confidence interval for the difference $p_F - p_M$ equals (0.02, 0.09). What of the following statement is correct?
A.Since the confidence interval does not contain 0, it is impossible that $p_F = p_M$
B.We are 95% confident that the probability of possessing house properties is between 0.02 and 0.09.
C.We are 95% confident that the probability of possessing house properties for female is between 0.02 and 0.09 larger than that for male.
D.We are 95% confident that a minority of Taiwan residents possesses house properties.
E.We can conclude that the population proportions may be equal.

48. Which of the following statement is when the one-way ANOVA provides relatively more evidence that $H_0: \mu_1 = \mu_2 = \cdots = \mu_g$ is rejected, where μ_i are the population mean for the ith group, $i = 1, \cdots, g$?
A.The smaller the between-groups variation. B.The smaller the within-groups variation.
C.The larger the between-groups variation. D.The larger the within-group variation. E.None of the above.

49. Which of the following statement about coefficient of determination R^2 is correct?
A.As the number of explanatory variables increases, R^2 may increase.
B.It is an index that can be used to measure any kind of associations between the response variable and the explanatory variable.
C.Its value must fall in 0 and 1.
D.The correlation of coefficient must equal to R^2.
E.When $R^2 = 1$, the slope of the linear relation between the response variable and the explanatory variable has to be positive.

50. What of the following assumption is not required when using the one-way ANOVA?
A.The mean value of each error term equals to 0. B.The number of observations have to the same for each population.
C.The observations are independent. D.The observations from each population follow a common normal distribution.
E.The variance for observations for each population have to be the same.

國立臺北大學109學年度碩士班一般入學考試試題

系（所）組別：都市計劃研究所甲組、乙組

科　　目：統計學

☑可 ☐不可使用計算機

Table 1: Area under the standard normal curve

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.00	0.500	0.504	0.508	0.512	0.516	0.519	0.523	0.527	0.531	0.535
0.10	0.539	0.543	0.547	0.551	0.555	0.559	0.563	0.567	0.571	0.575
0.20	0.579	0.583	0.587	0.591	0.594	0.598	0.602	0.606	0.610	0.614
0.30	0.617	0.621	0.625	0.629	0.633	0.636	0.640	0.644	0.648	0.651
0.40	0.655	0.659	0.662	0.666	0.670	0.673	0.677	0.680	0.684	0.687
0.50	0.691	0.695	0.698	0.701	0.705	0.708	0.712	0.715	0.719	0.722
0.60	0.725	0.729	0.732	0.735	0.738	0.742	0.745	0.748	0.751	0.754
0.70	0.758	0.761	0.764	0.767	0.770	0.773	0.776	0.779	0.782	0.785
0.80	0.788	0.791	0.793	0.796	0.799	0.802	0.805	0.807	0.810	0.813
0.90	0.815	0.818	0.821	0.823	0.826	0.828	0.831	0.834	0.836	0.838
1.00	0.841	0.843	0.846	0.848	0.850	0.853	0.855	0.857	0.859	0.862
1.10	0.864	0.866	0.868	0.870	0.872	0.874	0.877	0.879	0.881	0.883
1.20	0.884	0.886	0.888	0.890	0.892	0.894	0.896	0.898	0.899	0.901
1.30	0.903	0.904	0.906	0.908	0.909	0.911	0.913	0.914	0.916	0.917
1.40	0.919	0.920	0.922	0.923	0.925	0.926	0.927	0.929	0.930	0.931
1.50	0.933	0.934	0.935	0.937	0.938	0.939	0.940	0.941	0.942	0.944
1.60	0.945	0.946	0.947	0.948	0.949	0.950	0.951	0.952	0.953	0.954
1.70	0.955	0.956	0.957	0.958	0.959	0.959	0.960	0.961	0.962	0.963
1.80	0.964	0.964	0.965	0.966	0.967	0.967	0.968	0.969	0.969	0.970
1.90	0.971	0.971	0.972	0.973	0.973	0.974	0.975	0.975	0.976	0.976
2.00	0.977	0.977	0.978	0.978	0.979	0.979	0.980	0.980	0.981	0.981
2.10	0.982	0.982	0.983	0.983	0.983	0.984	0.984	0.985	0.985	0.985
2.20	0.986	0.986	0.986	0.987	0.987	0.987	0.988	0.988	0.988	0.989
2.30	0.989	0.989	0.989	0.990	0.990	0.990	0.990	0.991	0.991	0.991
2.40	0.991	0.992	0.992	0.992	0.992	0.992	0.993	0.993	0.993	0.993
2.50	0.993	0.994	0.994	0.994	0.994	0.994	0.994	0.994	0.995	0.995
2.60	0.995	0.995	0.995	0.995	0.995	0.996	0.996	0.996	0.996	0.996
2.70	0.996	0.996	0.996	0.996	0.996	0.997	0.997	0.997	0.997	0.997
2.80	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.998	0.998
2.90	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998
3.00	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.999	0.999

試題隨卷繳交