國立臺北大學112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、True and False（是非題）questions. [State the reason for your answer. 3% for each，30%]

1. A distribution table is required when using the p value of a test statistic to make a decision.
2. A random variable cannot be negative.
3. Assuming a random variable X has a normal distribution with mean μ and standard deviation σ, then $P[X \geq \mu] = 0.5$.
4. Consider two events A and B and suppose $P[B] > 0$, then $P[A|B] > P[A]$.
5. Consider two events A and B, then we have $P[A \text{ or } B] \leq P[A] + P[B]$.
6. Events A and B are mutually exclusive which imply they are independent.
7. The probability of an event A has to be less than 1.
8. The shape of the binomial distribution is symmetry when the probability of success equals 0.5.
9. The significance level of a test statistic has to be given in advance when performing the hypothesis test problem.
10. When the correlation of coefficient of random variables X and Y equals zero, we have X and Y are independent.

二、Multiple choice（單選題）[2% for each，70%]

1. What is the data type of the floor area (unit: in square foot)?
   A. Count
   B. Interval
   C. Nominal
   D. Ordinal
   E. Ratio

2. What is the data type of the type of transportation to work such as walk, bus, metro or car?
   A. Count
   B. Interval
   C. Nominal
   D. Ordinal
   E. Ratio

3. What is the data type for the preference of a certain type of coffee 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. Which of the following is the most appropriate statistic to summarize the type of transportation to work?
   A. Interquartile
   B. Mean
   C. Median
   D. Range
   E. None of the above

5. Which of the following is the least appropriate statistical statistic to summary the central tendency of the data for the floor area (unit: square foot)?
   A. Mean
   B. Median
   C. Midpoint
   D. Midrange
   E. Mode

6. Which of the following is the lease appropriate statistic to summary the variability of the data for the floor area (unit: square foot)?
   A. Correlation
   B. Interquartile
   C. Range
   D. Standard deviation
   E. Variance

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

8. Which of the following is the most appropriate statistic to examine the shape of the distribution for the floor area (unit: square foot)?
   A. Correlation
   B. Interquartile
   C. Mean
   D. Relative frequency
   E. Skewness

9. A survey is distributed to understand the commuting methods to NTPU. 270 students were responded and the data were summarized as follows:

Level of students	Method of Commutes		
	Metro	Bus	Walk
freshman	70	20	65
Sophomore	50	40	25

Determine the probability of the students who commute to school using metro or bus.
   A. 1/9
   B. 2/9
   C. 1/3
   D. 4/9
   E. 2/3

10. (Continued Question 9.) Given the respondents who were freshman, determine the probability of the students who walked to school?
   A. 1/6
   B. 13/31
   C. 1/3
   D. 5/18
   E. 13/18

11. What of the following property is not appropriate 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. $P[A \cup B] = P[A] + P[B](1 - P[A])$

12. A doctor is concerned about the relationship between blood pressure and sugar level. Among her patients, she classifies blood pressures as high, normal, or low and sugar level as normal and abnormal. She finds that (a) 15% have high blood pressure; (b) 20% have low blood pressure; (c) 18% have an abnormal sugar level; (d) of those with an abnormal sugar level, 35% have high blood pressure; and (e) of those with normal blood pressure, 12% have an abnormal sugar level. What percentage of her patients have a normal sugar level and low blood pressure?
    A. 0.161
    B. 0.176
    C. 0.184
    D. 0.920
    E. 0.039

13. (Continued Question 12.) Given patients with high blood pressure, what percentage of her patients have an abnormal sugar level?
    A. 0.063
    B. 0.12
    C. 0.18
    D. 0.35
    E. 0.42

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

X	0	1	2	3	4
Probability	0.3	?	0.1	0.3	0.25

What is the missing value?
    A. 0
    B. 0.05
    C. 0.1
    D. 0.15
    E. 0.2

15. Owing to the extravagant housing prices, the government would like to estimate the rate of the vacant house in the Taipei metropolitan area. From the past records, the rate of the vacant house is 10%. The government randomly selects 400 houses and assigns an agent to visit the house. Let X denote the number of non- vacant houses. What of the following condition is not necessary for X to be a binomial random variable?
    A. A selected 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 uncorrelated.
    E. Trials are independent.

16. (Continued Question 15.) Assume X is a binomial random variable. What of the following is a correct statement?
    A. The mean number of vacant houses equals 40.
    B. The mean number of non-vacant houses equals 15.
    C. The probability that all of the house are non-vacant is $1 - 0.9^{400}$.
    D. The standard deviation of X equals 36.
    E. The variance of X equals 6.

17. Regarding a normal random variable X, what of the following statement is incorrect?
    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 mode.
    D. The probability below the mean equals 1/2.
    E. The probability that X exceeds 1.96 equals 0.025.

18. The "fill" problem is important in many industries, such as those making cereal, toothpaste, beer, and so on. If an industry claims that it is selling 20 ounces of its product in a container, it must have a mean greater than 20 ounces, or else the FDA will crack down, although the FDA will allow a very small percentage of the containers to have less than 20 ounces. Assume the content X of a container follows a normal distribution with mean μ=20.1 and variance σ^2=0.25. What is P[X<20]?
    A. 0.023
    B. 0.484
    C. 0.5
    D. 0.516
    E. 0.977

19. Which of the following is not a necessary condition for the well-known Central Limit Theorem?
    A. All observations in the sample are independent.
    B. Each observation in the sample has equal mean.
    C. Each observation in the sample has equal variance.
    D. Each observation in the sample has to have a normal distribution.
    E. The sample size of the sample has to be large.

20. Which of the following is not a necessary condition for a random sample?
    A. All observations are correlated.
    B. All observations are drawn from the same population.
    C. Each observation has equal mean.
    D. Each observation has equal variance.
    E. Each observation has a common distribution.

21. What of the following is not the property for the sample mean?
    A. The mean of the sample mean equals the population mean.
    B. The observations used to derive the sample mean have to be uncorrelated.
    C. The population mean can be estimated by it.
    D. The standard deviation of the sample mean is a function of the sample size.
    E. When the sample size is large, the distribution of the sample mean is approximately normal.

22. Which of the following is a correct statement when the sample size increases?
    A. The length of the confidence interval of the population mean increases.
    B. The population mean increases.
    C. The sample mean increases.
    D. The standard deviation of the population decreases.
    E. The standard deviation of the sample mean decreases.

23. The Tourism Bureau of Taiwan would like to estimate the mean amount of money spent by a tourist to within $1000 with 90% confidence. If the amount of money spent by tourists is considered to be normally distributed with a standard deviation of $2500, what sample size would be necessary for the Chamber of Commerce to meet its objective in estimating the mean spending amount?
    A. 5
    B. 7
    C. 17
    D. 25
    E. 44

24. What of the following is a correct property for the confidence interval?
    A. The length of the confidence interval is not related to the size of the random sample.
    B. The length of confidence interval is shorter as the mean gets larger.
    C. The length of confidence interval is shorter as the sample size increases.
    D. The length of confidence interval is shorter as the standard deviation gets larger.
    E. The length of confidence interval is shorter when the confidence level gets larger.

25. Which is the least appropriate statistical test to evaluate the association between the data for the level of education for the head of household measured as less than high school (HS) and greater than HS and household income (unit: thousand dollars)?
    A. Chi-square test
    B. Mann-Whitney test
    C. One-way ANOVA
    D. Regression
    E. Two-independent sample *T* test

26. Which is the most appropriate statistical model or test to evaluate the association between the data for the floor area (unit: square foot) and household income (unit: thousand dollars)?
    A. Chi-square test
    B. Mann-Whitney test
    C. One-way ANOVA
    D. Regression
    E. Two-independent sample *T* test

27. The average time for a resident in Taiwan watches TV in Taiwan about 2 hours per day. The following provides some summary statistics for a sample collected from a survey in Japan:

	N	Sample mean	Sample standard deviation
Respondents	100	2.5	0.6
95% CI for average: (2.38, 2.62)			

    Which of the following statement is the most appropriate statement?
    A. The average time for a resident in Taiwan watching TV is significantly different than that for past research.
    B. The average time for a resident in Taiwan watching TV is significantly higher than that for past research.
    C. The average time for a resident in Taiwan watching TV is significantly lower than that for past research.
    D. The average time for a resident in Taiwan watching TV is not significantly different from that for the past research.
    E. The average time for a resident in Taiwan watching TV exactly equals to 2.5.

28. Suppose the city government would like to build a new metro route. They would like to know whether the local resident approves this policy. A survey is distributed to understand the residents' opinion. Let the probability of approving this policy be denoted as p. What is the least sample size that the city government needs to have the marginal error of E=0.06 with 95% confidence?
    A. 67
    B. 188
    C. 267
    D. 534
    E. 376

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

30. Which of the following is a correct statement about the null hypothesis that investigate the association between the method of commute and gender?
    A. H_0: the commuting method and gender are dependent.
    B. H_0: the commuting method and gender are independent.
    C. H_0: the commuting method and gender are not significantly associated.
    D. H_0: the commuting method and gender are significantly associated.
    E. H_0: the commuting method and gender are uncorrelated.

31. Which is the most appropriate test to evaluate the association between the commuting method (nominal scale) and gender?
    A. Chi-square test
    B. *F* test
    C. Mann-Whitney test
    D. Paired *T* test
    E. Two-independent sample *T* test

32. Let μ denote the population mean of IQ. Which of the following is a correct statement?
    A. For a fixed significance level α, the probability of a Type I error decreases when the sample size increases.
    B. 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 $\alpha = 0.1$.
    C. In the test of $H_0: \mu = 100$ versus $H_a: \mu \neq 100$, the *p*-value is greater than greater than 0.01. Then, a 99% confidence interval for μ should be (120, 130).
    D. *P*[Type II error]=1-*P*[Type I error]
    E. The *p* value is defined as the probability that H_0 is true.

33. 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 $\alpha = 0.05$. Which of the following statement is correct?
    A. If H_0 is rejected, the probability that H_0 is actually true is 0.05.
    B. If H_0 was true, the probability that the sample mean equals exactly the observed value would equal 0.05.
    C. If the plant exceeds the limit, there is only a 5% chance that we will conclude that they do not exceed the limit.
    D. If the plant does not exceed the limit, there is only 5% chance that we will conclude they exceed the limit.
    E. We have 5% confidence that this plant does not discharge more than 1000 gallons of wastewater.

34. Let p_F and p_M denote the probability of possessing a house for female and male in Taiwan. Suppose that a 95% confidence interval for the difference $p_M - p_F$ equals (-0.05, -0.01). What of the following statement is correct?
    A. The probability of possessing a house for male and female is not significantly different.
    B. The difference in the probability of possessing a house between male and female is between -0.05 and -0.01.
    C. We are 95% confident that the probability of possessing a house for female is significantly larger than that for male.
    D. We are 95% confident that a minority of Taiwan residents possesses a house.
    E. We can conclude that the population proportions for male and female are equal.

35. Let μ_i denote the population mean for the *i*th group, $i = 1, \cdots, g$. When the one-way ANOVA provides relatively more evidence that $H_0: \mu_1 = \mu_2 = \cdots = \mu_g$ is rejected, which of the following statement is appropriate?
    A. The between-groups variation is smaller.
    B. The between-groups variation is larger.
    C. The within-groups variation is smaller.
    D. The within-groups variation is larger.
    E. The total sum of square is unchanged.

國立臺北大學112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	$P[Z<z]$									
	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