國立成功大學 108 學年度碩士班招生考試試題

系　　所：經濟學系

考試科目：統計學

考試日期：0223，節次：1

※ 考生請注意：本試題不可使用計算機。　請於答案卷(卡)作答，於本試題紙上作答者，不予計分。

Part I: Multiple Choice [40pts, 2pt each]

Instruction: Identify one choice that best completes the statement or answers the question. Answers gain no point if it is not understandable.

1. The numerical value of the variance
a. is always larger than the numerical value of the standard deviation
b. is always smaller than the numerical value of the standard deviation
c. is negative if the mean is negative
d. can be larger or smaller than the numerical value of the standard deviation

2. Which of the following is **not** a measure of dispersion?
a. mode
b. standard deviation
c. range
d. interquartile range

3. If $P(A) = 0.62$, $P(B) = 0.47$, and $P(A \cup B) = 0.88$; then $P(A \cap B) =$
a. 0.2914
b. 1.9700
c. 0.6700
d. 0.2100

4. Which of the following is a required condition for a discrete probability function?
a. $\Sigma f(x) = 0$
b. $f(x) \geq 1$ for all values of x
c. $f(x) \leq 0$
d. $\Sigma f(x) = 1$

5. Twenty percent of the students in a class of 100 are planning to go to graduate school. The standard deviation of this binomial distribution is
a. 20
b. 16
c. 4
d. 2

6. For a continuous random variable x, the probability density function f(x) represents
a. the probability at a given value of x
b. the area under the curve at x
c. the area under the curve to the right of x
d. the height of the function at x

7. The uniform probability distribution is used with
a. a continuous random variable
b. a discrete random variable
c. a normally distributed random variable
d. any random variable, as long as it is not nominal

8. Since the sample size is always smaller than the size of the population, the sample mean
a. must always be smaller than the population mean
b. must be larger than the population mean
c. must be equal to the population mean
d. can be smaller, larger, or equal to the population mean

9. As the sample size increases, the
a. standard deviation of the population decreases
b. population mean increases
c. standard error of the mean decreases
d. standard error of the mean increases

10. To compute an interval estimate for the difference between the means of two populations, the t distribution
a. is restricted to small sample situations
b. is not restricted to small sample situations
c. can be applied when the populations have equal means
d. None of these alternatives is correct.

11. If two independent large samples are taken from two populations, the sampling distribution of the difference between the two sample means
a. can be approximated by a Poisson distribution
b. will have a variance of one
c. can be approximated by a normal distribution
d. will have a mean of one

12. The sampling distribution of the ratio of two independent sample variances taken from normal populations with equal variances is
a. an F distribution
b. a chi-square distribution
c. a t distribution
d. a normal distribution

Suppose X = the number of cars owned by a family in the U.S. The probability distribution of X is shown in the table below.

X	0	1	2	3
Probability	0.56	0.23	0.12	0.09

13. What is the chance that a family owns more than one car?
a. 0.23
b. 0.21
c. 0.44
d. None of these choices.

14. The Student t distribution:
a. is symmetrical.
b. approaches the normal distribution as the degrees of freedom increase.
c. has more area in the tails than the standard normal distribution does.
d. All of these choices are true.

15. An estimator is said to be consistent if:
a. the difference between the estimator and the population parameter grows smaller as the sample size grows larger.
b. it is an unbiased estimator.
c. the variance of the estimator is zero.
d. the difference between the estimator and the population parameter stays the same as the sample size grows larger.

16. The problem with relying on a point estimate of a population parameter is that:
a. it is virtually certain to be wrong.
b. it doesn't have the capacity to reflect the effects of larger sample sizes.
c. it doesn't tell us how close or far the point estimate might be from the parameter.
d. All of these choices are true.

17. A confidence interval is defined as:
a. a point estimate plus or minus a specific confidence level.
b. a lower and upper confidence limit associated with a specific level of confidence.
c. an interval that has a 95% probability of containing the population parameter.
d. a lower and upper confidence limit that has a 95% probability of containing the population parameter.

18. If a hypothesis is not rejected at the 0.10 level of significance, it:
a. must be rejected at the 0.05 level.
b. may be rejected at the 0.05 level.
c. will not be rejected at the 0.05 level.
d. must be rejected at the 0.025 level.

19. The degrees of freedom for the test statistic for μ when σ is unknown is:
a. 1
b. n
c. $n-1$
d. None of these choices.

20. The statistic $(n-1)s^2 / \sigma^2$ has a chi-squared distribution with $n-1$ degrees of freedom if:
a. the sample has a Student t-distribution with degrees of freedom equal to $n-1$.
b. the sample is normally distributed with variance equal to s^2.
c. the population is normally distributed with variance equal to σ^2.
d. All of these choices are true.

Part II: Short Answer [60pts]
Instruction: Answers gain no point if there is no explanation, calculation, or if it is not understandable.

1. [4pts, 2pts each] Provide **short and concise** answers to the following questions. **Briefly explain** your logic.

(a) If the mean is less than the mode for a data set, what can you conclude about the data's distribution?

(b) Your score of statistics exam states that you scored in the 95th percentile. What does this mean?

2. [6pts, 2pts each] The following is a partial list of statistical methods that students learned from their statistics class:

mean	percentile	box plot
median	coefficient of variation	confidence interval for a mean
mode	scatter plot	confidence interval for a proportion
standard deviation	histogram	one sample hypothesis test of means
z-score	pareto chart	one sample hypothesis test of proportions
		two sample hypothesis test of proportions

Decide which method is most applicable in the following situations. **Briefly explain** your logic.

(a) A researcher wants to estimate the mean weight loss when on the *Super Fit Diet*. The researcher takes a random sample of 300 people's weight loss who used the diet.

(b) The Taiwan Bureau of Statistics wants to test the claim that the unemployment rate is above 4 percent using a random sample of 1,000 people.

(c) A researcher would like to estimate the mean weight of elephant.

3. [10pts, 2pts each] The following questions regard hypothesis testing in general. Provide **short and concise** answers to the following questions. **Briefly explain** your logic.

(a) If you are using a hypothesis test to make a decision where the effect of a Type I error may negatively affect human health, should you increase or decrease α?

(b) What variable represents the actual Type I error for a study?

(c) Two studies were conducted, study A had a power of 0.6 and study B had a power of 0.8. Which study would be more likely to support a true alternative hypothesis?

(d) A researcher takes a sample, conducts a hypothesis test, and fails to reject the null hypothesis since the p-value was not small enough. The researcher concludes that "the sample data supports that the mean height of men is equal to 171.4 centimeters." What is wrong with this conclusion?

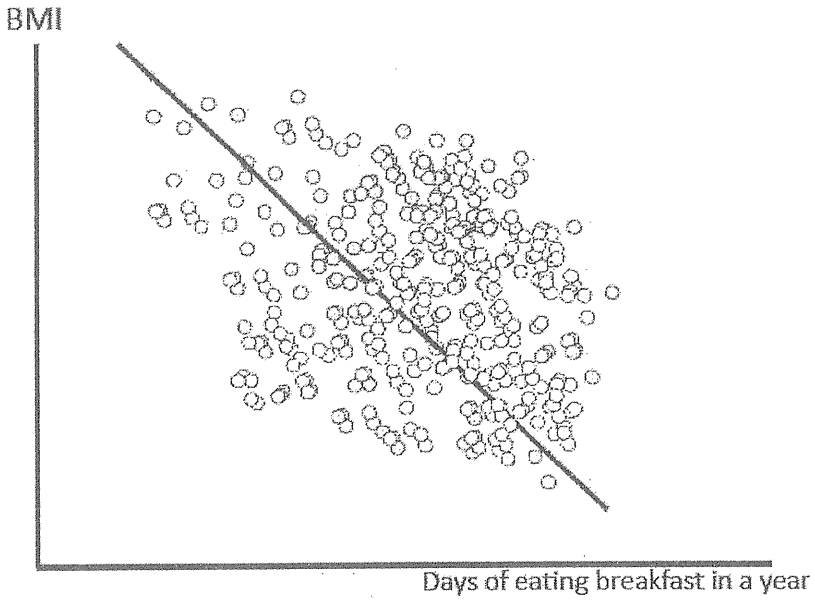
(e) For the one-sample mean hypothesis test, what is the distribution of the test statistic if standard deviation (σ) is unknown? (Give the specific name.)

4. [20pts, 10pts each] A researcher wants to investigate the effect of eating breakfast on obesity. Let x be the number of days an individual eats breakfast in a year. Let y be the body mass index (BMI). He collects the following data:

	Individual 1	Individual 2	Individual 3
x	50	150	250
y	60	40	20

(a) For regression model $y = \alpha + \beta x + \varepsilon$, find estimated $\hat{\alpha}$ and $\hat{\beta}$.

(b) This researcher collects more data and plot the following graph (each dot represents an individual):



After seeing this graph, he claims that "individuals who regularly ate breakfast tended to gain less weight and had a lower body mass index than breakfast skippers." The researcher concludes that "eating breakfast can cause people to be less obese". Do you agree? Why or why not? Explain your logic in detail.

國立成功大學 108 學年度碩士班招生考試試題

系　　所：經濟學系

考試科目：統計學

考試日期：0223，節次：1

5. [10pts] We know that the normal distribution $N(\mu, \sigma^2)$ has pdf

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

The standard normal distribution $N(0, 1)$ has mean 0, variance 1, and pdf $\phi(z)$ given by setting $\mu = 0$ and $\sigma = 1$ above. $\Phi(z)$ is the cdf and it does not have a nice formula. Suppose $Z \sim N(0, 1)$ and $X = aZ + b$.

(a) [2pts] Compute the mean μ and variance σ^2 of X.

(b) [8pts] Express the cdf $F_X(x)$ of X in terms of Φ and then use the chain rule to find the pdf $f_X(x)$ of X.

6. [10pts] The relation between degrees Celsius and degrees Fahrenheit is:

$$\text{degrees Fahrenheit} = \text{degrees Celsius} \times \frac{9}{5} + 32$$

Let X and Y be the daily high temperature in degrees Fahrenheit for the summer in Taipei and Tainan. Let U and G be the daily high temperatures in degrees Celsius for the summer in Taipei and Tainan. Suppose that Cov(X, Y) = 4 and Corr(X, Y) = 0.8 (Corr(X, Y) = correlation between X and Y). Find Cov(U, G) and Corr(U, G).