第一部分：社會統計 (50%)

單選題：每題兩分　※ 注意：請於試卷內之「選擇題作答區」依序作答。

_____ **1.** For two events A and B, P(A ∩ B) =
A. P(A) × P(A|B)
B. P(B) × P(B|A)
C. P(A) × P(B|A) = P(B) × (A|B)
D. P(A) × P(A|B) = P(B) × (B|A)
E. None of the above

_____ **2.** Suppose that Z is a standard normal random variable. What is the probability that Z lies between 0 and -2?
A. 0.95
B. 0.475
C. 0.34
D. 0.4772
E. None of the above

_____ **3.** Of the following, the most accurate statement concerning "sampling error" is:
A. This type of error cannot be totally prevented even when one is using the optimum sample design.
B. This type of error can be controlled by being accurate in calculations.
C. This type of error can be controlled by being accurate in measurements.
D. This type of error is the result of poor sample design.
E. All of the above

_____ **4.** When simple random sample is used,
A. Costs of sampling will be minimized.
B. Some members of the population will still have greater probabilities of being included in the sample than other members of the population.
C. There is still possibility of sampling error.
D. There is no possibility of chance error because random selection overcomes such error.
E. All of the above

_____ **5.** Which of the following statements is true about Chebychev's Theorem (CT) and the Empirical Rule (ER)?
A. CT is a more conservative rule than ER, since it assumes nothing about the underlying distribution of the measurements from which we are sampling.
B. ER is a more conservative rule than CT, since it assumes nothing about the underlying distribution of the measurements from which we are sampling.
C. CT assumes that the distribution from which we are sampling is bell-shaped.
D. A larger percentage of the measurements are expected to fall within the mean plus or minus two standard deviations when CT is used than when ER is used.
E. None of the above

見背面

_____ **6.** Suppose that A and B are independent events, with P(A)=0.2 and P(B)=0.7. Which is the probability that neither A nor B will occur?
A. 0.14
B. 0.24
C. 0.90
D. 0.10
E. None of the above

_____ **7.** Which of the following is correct statement concerning the Central Limit Theorem (CLT)?
A. The CLT states that the sample mean, $\bar{x}$, is always equal to μ.
B. The CLT states that for large samples, $\bar{x}$, is always equal to μ.
C. The CLT states that for large sample, the sampling distribution of population mean is approximately normal.
D. The CLT states that for large sample, the sampling distribution of sample mean is approximately normal.
E. Both B and C are correct.

_____ **8.** Suppose the statistic T is an unbiased estimator of θ and a sample of size n is used to compute T. Which of the following is NECESSARILY true?
A. The variance of the sampling distribution of T is $\sigma/\sqrt{n}$.
B. $T=\theta$
C. $E(\theta) = T$
D. $E(T) = \theta$
E. Variance (T) < variance (θ)

_____ **9.** For a given level of confidence, if the sample size is decreased,
A. The probability that the interval will not include the parameter increases.
B. The difference between the upper limit of the interval and the lower limit of the interval remains the same.
C. The difference between the upper limit of the interval and the lower limit of the interval increases.
D. The difference between the upper limit of the interval and the lower limit of the interval decreases.

_____ **10.** If a researcher is using a 95% level of confidence in calculating a confidence interval,
A. 95% of the time the computed interval will include the sample mean.
B. 5% of the time such intervals will not include the population value.
C. In the long run, 95% of all sample means will fall within the interval.
D. 95% of the time the interval will not include the population value.

_____ **11.** One characteristic of any t-distribution is
A. It is skewed to the right.
B. As n increases, the t-distribution has less and less resemblance to a normal curve.
C. It has a single parameter, which is degree of freedom.
D. It has $\mu = 0$ and $\sigma = 1$.
E. None of the above

接次頁

_____ **12.** To test H_0: $\sigma^2 = \sigma_0^2$ vs. H_1: $\sigma^2 \neq \sigma_0^2$ the test statistic is
A. s^2/σ_0^2
B. $s^2/(\mathrm{n}-1)\sigma_0^2$
C. $\mathrm{n}s^2/\sigma_0^2$
D. $(\mathrm{n}-1)s^2/\sigma_0^2$

_____ **13.** The ratio $(s_1^2/s_2^2)/(\sigma_1^2/\sigma_2^2)$ is
A. t-distributed
B. z-distributed
C. χ^2-distributed
D. F-distributed

_____ **14.** Of the following, the one that is NOT a property of multinomial experiment is
A. the experiment consists of n identical trials.
B. the outcome of each trial falls into one of the k classes, or cells.
C. The trials are independent
D. The probability that the outcome of a single trial will fall in a particular cell varies from trial to trial.

_____ **15.** For a Chi-square test of a contingency table,
A. The expected cell counts should be 10 or more.
B. The degrees of freedom will always equal k-1.
C. All cell counts should be greater than or equal to 5.
D. All expected cell counts should be greater than or equal to 5.

_____ **16.** See the ANOVA table shown below. What is the total number of observations?
A. 7
B. 9
C. 11
D. 13
E. None of the above

Source	d.f.	SS	MS	F
Treatments	2	?	?	2
Error	?	?	25	
Total	?	325		

_____ **17.** If we regress y on x by the least-squares technique and the correlation between x and y, $r_{xy} = 0.7$, which of the following is most correct?
A. The percentage of squared deviations in the y values that are "explained" by the x values is 30%.
B. The percentage of the squared deviations in the y values that are "explained" by the x values is 70%
C. The percentage of squared deviations in the y values that are not "explained" by the x values is 9%.
D. The percentage of squared deviations in the y values that are not "explained" by the x values is 51%.

見背面

_____ **18.** The chi-square goodness-of-fit test
A. is a lower-tail test
B. is an upper-tail test
C. is a two-tail test
D. all of the above are correct

_____ **19.** A disadvantage of running multiple t-test in the comparison of several population means is
A. the probability of incorrectly rejecting at least one of the hypotheses increases as the number of t-tests increases.
B. α may have to be set as high as .10 for valid testing.
C. sample sizes may have to be very large.
D. null hypothesis may not be rejected for any of the t-tests.

_____ **20.** There are three basic assumptions for the F-test in an ANOVA to be valid, and those assumptions are
A. normality, large sample sizes, and equal group variances.
B. normality, independence, and equal variance.
C. equal variance, large sample sizes, and independence.
D. large sample sizes, normality, and equal variance
E. None of the above

_____ **21.** A straight line fitted by least-squares techniques has all of the following characteristics except
A. it gives the best fit to the data in the sense that it makes the sum of squared deviations from the line less than any other straight line
B. the larger the population from which the sample is drawn, the better the fit.
C. The deviations of the actual values from the predicted values sum to 0.
D. When the data represent a sample from a larger population, the least-squares line is the "best" estimate of the population regression line.
E. None of the above

_____ **22.** In a simple regression model $y_i = b_0 + b_1 x_i + \varepsilon_i$, the key assumption that allows one to construct confidence intervals for the slope is
A. y_i is independent of x_i.
B. the x_i are normally distributed
C. the relationship between y_i and x_i is linear
D. the ε_i are normally distributed

_____ **23.** The following regression equation has been estimated:

$$y_i = 6.4 + 3.26x_i - 3.87z_i + \varepsilon_i$$
$$\quad\quad (0.48) \quad (2.21) \quad (1.11)$$

The numbers in parentheses are estimated standard error and the sample size was 20, which of the following interpretations is NOT entirely correct?
A. x_i has a positive effect on y_i.
B. y_i decrease as z_i increases.
C. the relationship between y_i and x_i is linear
D. the model assumes ε_i are normally distributed
E. None of the above

接次頁

_____ **24.** When one uses a dummy variable in a regression model,

A. The number of dummy variables entered for the qualitative variable is always one less than the number of levels of the qualitative variable.
B. Only one dummy variable needs to be entered into the model
C. The number of dummy variables entered for the qualitative variable equals the number of values the qualitative variable can assume.
D. One should not enter a second qualitative variable because there will be too many dummy variables in the model.

_____ **25.** If $\hat{y} = 2 + 3.1x - 1.5x^2$,

A. as x increases, $\hat{y}$ increases to a maximum and then decreases.
B. $\hat{y}$ decreases as x increases.
C. as x increases, $\hat{y}$ decreases to a minimum and then increase.
D. as x decrease, $\hat{y}$ increases to a maximum and then decreases.
E. $\hat{y}$ increases as x increases.

第二部分：社會研究方法 (50%) 請於試卷內之「非選擇題作答區」作答，並應註明作答之題號。

1. 名詞解釋 (一題 5 分)

1) 分層抽樣 (stratified sampling)

2) 知情同意 (informed consent)

3) 效度 (Validity)

4) 橫斷面研究 (Cross-sectional study)

2. 研究設計 (一題 15 分，共 30 分)

少子化是台灣近年遇到的社會問題，以總人口數來說，2019 年達到 2360 萬多的歷史高峰後，2020 年開始已連續兩年人口減少，2021 年間，人口減少了約 18.6 萬人。2021 年的新生兒僅有 15.3 萬人，是史上新低，平均每位婦女僅生育 1.07 位新生兒，為世界最低。面對少子化，各方提出許多不同論點解釋其成因，有說是高房價、育兒成本、薪資停滯等因素，有人說家務性別不平等，有人提到婦女教育程度提高，還有人說是價值觀改變所造成。

理想的政策，需要有嚴謹的研究證據來支持，以下請你提出一個解釋少子化成因的假設，結合量化與質化的研究設計，來驗證該成因的影響。

1) 請設計一個量化研究，詳述你需要的數據，包含測量方法、主要變項(自變項、依變項和控制變項)，並說明用何種統計方法來分析，除此之外，也請設想一些可能的推論限制。

2) 除了量化之外，也請設計一個質化研究，請簡單說明你採取的研究策略(如什麼樣的田野工作、與誰訪談等等)，你的設計可以幫助因果推論嗎？為什麼？

試題隨卷繳回