注意：考試開始鈴響前，不得翻閱試題，

並不得書寫、畫記、作答。

國立清華大學 108 學年度碩士班考試入學試題

系所班組別：**統計學研究所**

考試科目(代碼)：**統計學(0203)**

—作答注意事項—

1. 請核對答案卷（卡）上之准考證號、科目名稱是否正確。
2. 作答中如有發現試題印刷不清，得舉手請監試人員處理，但不得要求解釋題意。
3. 考生限在答案卷上標記「⬇由此開始作答」區內作答，且不可書寫姓名、准考證號或與作答無關之其他文字或符號。
4. 答案卷用盡不得要求加頁。
5. 答案卷可用任何書寫工具作答，惟為方便閱卷辨識，請儘量使用藍色或黑色書寫；答案卡限用 2B 鉛筆畫記；如畫記不清（含未依範例畫記）致光學閱讀機無法辨識答案者，其後果一律由考生自行負責。
6. 其他應考規則、違規處理及扣分方式，請自行詳閱准考證明上「**國立清華大學試場規則及違規處理辦法**」，無法因本試題封面作答注意事項中未列明而稱未知悉。

國立清華大學 108 學年度碩士班考試入學試題

系所班組別：統計學研究所（0502）

考試科目（代碼）：統計學（0203）

共 4 頁，第 1 頁　　*請在【答案卷】作答

***考生請注意：

- 共有二十五題，每題均為四選一之單選題。只需將題號與答案（填入括號內）直接寫入答案卷，不必附上計算過程（因不計分），作答型式如下所示：
  1. (　)
  2. (　)
  ⋮
  25. (　)
- 每題四分，若不會之題目請勿隨意猜測（括號中留空白），回答錯者每題倒扣 1.5 分。
- 總分數 ＝ 答對題數 ×4 － 答錯題數 ×1.5，若總分數為負數則以零分計算。

Suppose that $X_1, X_2, \cdots, X_n$ are a random sample from $N(\mu, \sigma^2)$, where $-\infty < \mu < \infty$ and $0 < \sigma^2 < \infty$ are both unknown parameters. Let $Pr(X_n \leq \theta) = 0.8, \bar{X} = \sum_{i=1}^n X_i/n, S^2 = \sum_{i=1}^n (X_i - \bar{X})^2/(n-1)$, and $\Phi(.)$ be the c.d.f. (cumulative distribution function) of $N(0,1)$.

1. What is the maximum likelihood estimator of θ?
(A) $\bar{X} + S\Phi^{-1}(0.8)$ (B) $\bar{X} + \sqrt{\frac{n-1}{n}}S\Phi^{-1}(0.8)$ (C) $\bar{X} + \sqrt{\frac{n}{n-1}}S\Phi^{-1}(0.8)$ (D) none of above

2. What is the uniformly minimum variance unbiased estimator of θ? (Here $\Gamma(.)$ is the gamma function)
(A) $\bar{X} + \frac{\Gamma((n-1)/2)}{\Gamma(n/2)}S\Phi^{-1}(0.8)$ (B) $\bar{X} + \sqrt{\frac{n-1}{2}}\frac{\Gamma(n/2)}{\Gamma((n-1)/2)}S\Phi^{-1}(0.8)$
(C) $\bar{X} + \sqrt{\frac{n-1}{2}}\frac{\Gamma((n-1)/2)}{\Gamma(n/2)}S\Phi^{-1}(0.8)$ (D) none of above

Suppose that a single random variable X has the p.d.f. (probability density function) $f(x;\theta) = 2(1-\theta)x + \theta, 0 \leq x \leq 1$, and $f(x;\theta) = 0$ otherwise, where the value of θ is unknown, but $0 \leq \theta \leq 2$. Based on this "single" random variable X, the following hypotheses are to be tested: $H_0 : \theta = 2$ versus $H_1 : \theta = 0$.

3. What is the critical region for which $\alpha + 2\beta$ is a minimum, where α and β are the respective probabilities of Type I and Type II errors?
(A) $x > 1/3$ (B) $x > 1/2$ (C) $x > 1/4$ (D) none of above

4. What is the minimum value of $\alpha + 2\beta$ for the critical region you found in Problem 3?
(A) 4/9 (B) 5/9 (C) 6/9 (D) none of above

5. Suppose that for testing H_0 against H_1, the significance level α is given, $0 < \alpha < 1$. What is the best critical region for which the probability of Type II error is minimized?
(A) $x \geq 1/(k+1)$, where $k > 0$ satisfies $k^2/(k+1)^2 = \alpha$ (B) $x \leq 1/k$, where $k > 0$ satisfies $k^2/(k+1)^2 = \alpha$ (C) $x > 1/k$, where $k > 0$ satisfies $k^2/(k+1)^2 = \alpha$ (D) none of above

Suppose that "four observations are taken at random from a normal distribution for which the mean μ, $-\infty < \mu < \infty$, is unknown and the variance is 1. Suppose also that the following hypotheses are to be tested: $H_0 : \mu \geq 10$ versus $H_1 : \mu < 10$.

6. Define $\bar{X}$ to be the sample mean of these four observations. What is the best critical region for the uniformly most powerful test for testing H_0 against H_1 with the size (maximum significance level) $\alpha = 0.1$? (hint: $\Phi(1.282) = 0.9$, where $\Phi(.)$ is the c.d.f. of $N(0,1)$)
(A) $\bar{X} \leq 10.641$ (B) $\bar{X} \leq 8.359$ (C) $\bar{X} \leq 9.359$ (D) none of above

7. What is the power of this uniformly most powerful test at $\mu = 9$?
(A) $\Phi(-1.282)$ (B) $\Phi(0.718)$ (C) $\Phi(1.282)$ (D) none of above

8. What is the probability of not rejecting H_0 if $\mu = 11$?
(A) $\Phi(3.282)$ (B) $\Phi(-3.282)$ (C) $2\Phi(3.282) - 1$ (D) none of above

Suppose that "one observation" X is drawn from a distribution with the following p.d.f.: $f(x|\theta) = 1/\theta$ for $0 < x < \theta$ and $f(x|\theta) = 0$ otherwise. Also, suppose that the prior p.d.f. of θ is $\xi(\theta) = \theta e^{-\theta}$ for $\theta > 0$ and $\xi(\theta) = 0$ otherwise.

9. What is the posterior distribution of θ given X?
(A) $e^{-(x+\theta)}, 0 < x < \theta$ (B) $e^{\theta-x}, 0 < x < \theta$ (C) $e^{x-\theta}, 0 < x < \theta$ (D) none of above

10. What is the Bayes estimator of θ under squared error loss?
(A) $1 - X$ (B) $1 + X$ (C) X (D) none of above

11. What is the Bayes estimator of θ under absolute error loss?
(A) $\ln 2 + X$ (B) $\ln 2 - X$ (C) X (D) none of above

Let $X_1, X_2, \cdots, X_n$ be a random sample from $N(\mu, 1)$, where $-\infty < \mu < \infty$ is the unknown parameter.

12. What is the Fisher information of this random sample of size n?
(A) n (B) $n - 1$ (C) $n + 1$ (D) none of above

13. What is the Rao-Cramer Lower Bound for all unbiased estimators of μ^2?
(A) $2\mu^2/n$ (B) $4\mu^2/n$ (C) u^2/n (D) none of above

14. What is the uniformly minimum variance unbiased estimator of μ^2?
(A) $\bar{X}^2 + 1$ (B) $\bar{X}^2 - 1$ (C) $\bar{X}^2 + 1/n$ (D) none of above

Let the random variable X_n have a Poisson distribution with expectation $n > 0$.

15. What is the limiting distribution of $2X_n/\sqrt{n} - 2\sqrt{n} + 2$ as $n \to \infty$?
(A) $N(0,4)$ (B) $N(2,2)$ (C) $N(2,4)$ (D) none of above

Let X_1, X_2, X_3 be a random sample of size 3 from Gamma (α, β) (with p.d.f. $f(x;\alpha,\beta) = x^{\alpha-1}e^{-x/\beta}/\Gamma(\alpha)\beta^{\alpha}$), where $\alpha = 3$ and $\beta > 0$ is the unknown parameter.

16. What is the distribution of $2(X_1 + X_2 + X_3)/\beta$? Here the notation χ^2_p means a chi-square distribution with p degrees of freedom.
(A) χ^2_9 (B) χ^2_{18} (C) χ^2_6 (D) none of above

17. Based on your solution to Problem 16, which one below is a 95% equal-tail confidence interval for β? Here $\chi^2_{p,\gamma}$ is a positive number such that $Pr(\chi^2_p \le \chi^2_{p,\gamma}) = \gamma$.
(A) $[2(X_1+X_2+X_3)/\chi^2_{18,0.975}, 2(X_1+X_2+X_3)/\chi^2_{18,0.025}]$ (B) $[\chi^2_{18,0.975}/2(X_1+X_2+X_3), \chi^2_{18,0.025}/2(X_1+X_2+X_3)]$ (C) $[2(X_1+X_2+X_3)/\chi^2_{9,0.975}, 2(X_1+X_2+X_3)/\chi^2_{9,0.025}]$ (D) none of above

Following a presidential debate, people were asked how they might vote in the forthcoming election. Based on the data below, is there any association between one's gender and choice of presidential candidate? To test the hypotheses, you might need the information that $Pr(\chi^2_1 \le 3.841) = Pr(\chi^2_2 \le 5.991) = 0.95$ and $Pr(\chi^2_1 \le 2.706) = Pr(\chi^2_2 \le 4.605) = 0.9$.

	Gender	
Candidate Preference	Male	Female
Candidate A	150	130
Candidate B	100	120

18. Given the significance level $\alpha = 0.05$, what is your conclusion for the test?
(A) no association (B) rejection of no association (C) the observed value of test statistic is between 3.841 and 5.991 (D) the observed value of test statistic > 5.991

19. Given the significance level $\alpha = 0.1$, what is your conclusion for the test?

(A) no association (B) rejection of no association (C) the observed value of test statistic < 2.706 (D) the observed value of test statistic > 3.841

Three plants, A_1, A_2, and A_3 produce respectively, 10, 50, and 40 percent of a company's output. Although plant A_1 is a small plant, its manager believes in high quality and only 1 percent of its products are defective. The other two, A_2 and A_3, are worse and produce items that are 3 and 4 percent defective, respectively. All products are sent to a central warehouse.

20. If one item is selected at random from the warehouse, what is the probability that this item is defective?
(A) 0.016 (B) 0.017 (C) 0.031 (D) none of above

21. If one item is selected at random from the warehouse and observed to be defective, what is the probability that this item comes from plant A_1?
(A) 1/31 (B) 1/32 (C) 1/16 (D) none of above

After fitting a simple linear regression model $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, where ϵ_i are i.i.d. from $N(0, \sigma^2)$, we have the information that $\sum_{i=1}^{n}(y_i - \bar{y})^2 = 240, R^2 = 0.8$, and $n = 52$.

22. What is the estimate of the standard deviation of ϵ_i?
(A) $\sqrt{48/50}$ (B) $\sqrt{48/51}$ (C) $\sqrt{48/52}$ (D) none of above

23. If we have more information that $\hat{\beta}_1 = 2$, what is the sample correlation coefficient between x and y?
(A) $4/5$ (B) $\sqrt{2}/5$ (C) $2/\sqrt{5}$ (D) none of above

24. Two numbers are selected at random from the interval (0,1). If these values are uniformly and independently distributed, compute the probability that the three resulting line segments, by cutting the interval at the numbers, can form a triangle.
(A) 1/2 (b) 1/3 (C) 1/4 (D) none of above

25. Let X_1, X_2, X_3 be a random sample from a distribution of the continuous type having p.d.f. $f(x) = 2x, 0 < x < 1$, zero elsewhere. What is the probability that the smallest of these X_i exceeds the median of this distribution?
(A) 1/2 (B) 1/4 (C) 1/6 (D) none of above