

大同大學 89 學年度研究所招生入學考試試題

考試科目 統計學

所別 資訊經營研究所 第 1/3 頁

註：本次考試 不可 參考自己的書籍及筆記 不可 查字典 可 使用計算器

Reference Data:

$z_{0.025} = 1.96, z_{0.05} = 1.645,$
 $t_{0.025,18} = 2.101, t_{0.05,18} = 1.734,$
 $t_{0.025,19} = 2.093, t_{0.05,19} = 1.729,$
 $t_{0.025,20} = 2.086, t_{0.05,20} = 1.725,$
 $t_{0.025,21} = 2.080, t_{0.05,21} = 1.721,$
 $t_{0.025,22} = 2.074, t_{0.05,22} = 1.717.$
 $\chi^2_{0.025,18} = 31.526, \chi^2_{0.05,18} = 28.869, \chi^2_{0.975,18} = 8.231, \chi^2_{0.95,18} = 9.390,$
 $\chi^2_{0.025,19} = 32.852, \chi^2_{0.05,19} = 30.143, \chi^2_{0.975,19} = 8.906, \chi^2_{0.95,19} = 10.117,$
 $\chi^2_{0.025,20} = 34.170, \chi^2_{0.05,20} = 31.410, \chi^2_{0.975,20} = 9.591, \chi^2_{0.95,20} = 10.851,$
 $\chi^2_{0.025,21} = 35.478, \chi^2_{0.05,21} = 32.670, \chi^2_{0.975,21} = 10.283, \chi^2_{0.95,21} = 11.591,$
 $\chi^2_{0.025,22} = 36.781, \chi^2_{0.05,22} = 33.924, \chi^2_{0.975,22} = 10.982, \chi^2_{0.95,22} = 12.338.$

Multiple Choice:

In the following, there are 20 problems. Each of them has only one correct choice and scores 5 points. Making a wrong choice on a problem, you will get -2 points for that problem. The minimum score of this test is zero.

Answer problem 1 and 2 using the following problem setting:
Let X be a random variable with pdf

$$f(x) = k(-8 - x^2 + 6x + 9)$$

$$= k(-x^2 + 6x - 9)$$

$$= k(-x^2 + 6x - 9)$$

$$= k(-x^2 + 6x - 9)$$

$$f(x) = \begin{cases} k[1 - (x-3)^2] & 2 \leq x \leq 4 \\ 0 & \text{otherwise} \end{cases}$$

$$= k(1 - (x-3)^2)$$

$$= k(1 - (x^2 - 6x + 9))$$

$$= k(-x^2 + 6x - 8)$$

- The value of $k = ?$ (a) 0.25; (b) 0.5; (c) 0.75; (d) 1.
- The value of $E[X] = ?$ (a) 2.5; (b) 3; (c) 3.5; (d) 3.6.

Answer problem 3 to 5 using the following problem setting:

Two components of a minicomputer have the following joint pdf for their useful lifetimes X and Y :

$$f(x,y) = \begin{cases} xe^{-x(1+y)} & x \geq 0 \text{ and } y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

- What is the probability that the lifetime X of the first components exceeds 3? (a) e^{-3} ; (b) $\frac{e^{-3}}{2}$; (c) $\frac{e^{-3}}{6}$; (d) $\frac{1}{3}$.
- What is the probability that the lifetime Y of the second components exceeds 3? (a) $\frac{1}{2}$; (b) $\frac{1}{3}$; (c) $\frac{1}{4}$; (d) e^{-4} .
- What is the probability that the lifetime of at least one component exceeds 3? (a) $\frac{1}{3}$; (b) $\frac{1}{4} + e^{-3} - \frac{1}{4}e^{-12}$; (c) $\frac{3}{4} - e^{-3} + \frac{1}{4}e^{-12}$; (d) $\frac{1}{4} + e^{-2} - \frac{1}{4}e^{-12}$.

$$f(x) = \int_0^{\infty} x e^{-x-y} dy$$

$$= \int_0^{\infty} x e^{-x} e^{-y} dy$$

$$= x e^{-x} \int_0^{\infty} e^{-y} dy$$

$$= x e^{-x} [-e^{-y}]_0^{\infty}$$

$$= x e^{-x} (0 - (-1))$$

$$= x e^{-x}$$

$$f(x) = \int_0^{\infty} x e^{-x(1+y)} dy$$

$$= x \int_0^{\infty} e^{-x-y} dy$$

$$= x [-e^{-x-y}]_0^{\infty}$$

$$= x (0 - (-e^{-x}))$$

$$= x e^{-x}$$

$$f(x) = \int_0^{\infty} x e^{-x(1+y)} dy$$

$$= x \int_0^{\infty} e^{-x-y} dy$$

$$= x [-e^{-x-y}]_0^{\infty}$$

$$= x (0 - (-e^{-x}))$$

$$= x e^{-x}$$

大同大學 89 學年度研究所招生入學考試試題

考試科目 **統計學**

所別 **資訊經營研究所** 第 $\frac{2}{3}$ 頁

註：本次考試 不可 參考自己的書籍及筆記 不可 查字典 可 使用計算器

$\bar{x} = 1.8134$
 $s^2 = 36.9 = \frac{\sum x_i^2 - n\bar{x}^2}{n-1}$
 $36.9 = \frac{291 - 10\bar{x}^2}{9}$

$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$
 $= \frac{\sum x_i^2 - n\bar{x}^2}{n-1}$
 $36.9 = \frac{291 - 10\bar{x}^2}{9}$

- Let A and B be two independent events, and let A' and B' denote their complements, respectively. Which of the following statements is false: (a) $P(A|B) = P(A)$; (b) $P(A|B') = P(A)$; (c) $P(A \cup B) = P(A) + P(B)$; (d) $P(A \cap B') = P(A)(1 - P(B))$.
- The probability of a man hitting a target is $\frac{1}{3}$. At least how many times must he fire so that the probability of hitting the target at least once is more than 90%: (a) 4; (b) 5; (c) 6; (d) 7.
- Let x_1, x_2, \dots, x_{10} be a sample from a population. Suppose that $\sum_{i=1}^{10} x_i^2 = 297$ and sample variance $s^2 = 36.9$. Then, the sample mean \bar{x} satisfies: (a) $48 < \bar{x} \leq 50$; (b) $50 < \bar{x} \leq 52$; (c) $52 < \bar{x} \leq 54$; (d) $54 < \bar{x} \leq 56$.
- Let $X \sim N(\mu_1, \sigma^2)$, $Y \sim N(\mu_2, 2\sigma^2)$ be two independent and normally distributed random variables. Which of the following statements is false: (a) $Var[X + Y] = Var[X - Y]$; (b) $X - Y \sim N(\mu_1 - \mu_2, 3\sigma^2)$; (c) $Var[2X + Y] = 3Var[Y]$; (d) $Var[2X] = Var[Y]$.
- z_{α} denotes the z critical value of an $N(0, 1)$ population. Suppose that $0 < \alpha_1 < \alpha_2 < \alpha_3 < 0.5$ and $\alpha_2 - \alpha_1 = \alpha_3 - \alpha_2$. Then, (a) $z_{\alpha_1} - z_{\alpha_2} = z_{\alpha_2} - z_{\alpha_3}$; (b) $z_{\alpha_1} - z_{\alpha_2} > z_{\alpha_2} - z_{\alpha_3}$; (c) $z_{\alpha_1} - z_{\alpha_2} < z_{\alpha_2} - z_{\alpha_3}$; (d) all of the above statements are incorrect.

Answer problem 11 and 12 using the following problem setting:
 Let a population have pdf $f(x) = \frac{1}{k^2}x, 0 \leq x \leq \sqrt{2}k$, and let x_1, x_2, \dots, x_n be a sample obtained from the population.

- Let \hat{k}_1 be the MLE of parameter k . Then $\hat{k}_1 = ?$ (a) $\max(x_i)$; (b) $\frac{1}{2} \max(x_i)$; (c) $\sqrt{2} \max(x_i)$; (d) $\frac{\sqrt{2}}{2} \max(x_i)$.
- Let \hat{k}_2 be the estimator of parameter k obtained using the method of moments. Then $\hat{k}_2 = ?$ (a) $\frac{3\sqrt{2}}{4} \bar{x}$; (b) $\frac{4\sqrt{2}}{3} \bar{x}$; (c) $\frac{\sqrt{3}}{2} \bar{x}$; (d) $\frac{2\sqrt{3}}{3} \bar{x}$.

Answer problem 13 to 15 using the following problem setting:
 There are two populations $X \sim N(\mu_X, \sigma^2)$ and $Y \sim N(\mu_Y, \sigma^2)$. The following table lists some sample statistics calculated from the random samples of the two populations:

	Sample Size	Sample Mean	Sample Variance
Population X	11	540	441
Population Y	11	554	225

- The 95% confidence interval of σ^2 : (a) (194.908, 694.401); (b) (197.108, 680.055); (c) (199.179, 667.092); (d) (331.04, 334.96);
- The 95% confidence interval of $\mu_Y - \mu_X$ is: (a) (-2.138, 30.138); (b) (-2.185, 30.185); (c) (-2.231, 30.231); (d) (1.251, 26.749).

$P(A|B) = P(A)$
 $P(A|B') = P(A)$
 $P(A \cup B) = P(A) + P(B)$
 $P(A \cap B') = P(A)(1 - P(B))$
 $P(X \geq 1) = 1 - P(X=0) = 1 - \frac{1}{2} \times \frac{1}{3} = \frac{5}{6}$
 $P(X \geq 2) = 1 - P(X=0) - P(X=1) = 1 - \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$
 $P(X \geq 3) = 1 - P(X=0) - P(X=1) - P(X=2) = 1 - \frac{1}{2} - \frac{1}{3} - \frac{1}{6} = 0$
 $\frac{1 - P(X=0)}{6} = \frac{1}{6}$
 $\frac{1 - P(X=0) - P(X=1)}{6} = \frac{1}{6}$
 $\frac{1 - P(X=0) - P(X=1) - P(X=2)}{6} = \frac{1}{6}$
 $\frac{1 - P(X=0) - P(X=1) - P(X=2) - P(X=3)}{6} = \frac{1}{6}$

$\mu - z \leq \bar{x} \leq \mu + z$
 $z = 1.645$
 $50 - 1.645 \times \sqrt{36.9} \leq \bar{x} \leq 50 + 1.645 \times \sqrt{36.9}$
 $50 - 10 \leq \bar{x} \leq 50 + 10$
 $40 \leq \bar{x} \leq 60$
 $50 - 10 \leq \bar{x} \leq 50 + 10$
 $40 \leq \bar{x} \leq 60$
 $50 - 10 \leq \bar{x} \leq 50 + 10$
 $40 \leq \bar{x} \leq 60$

$\mu_X = 16.131$
 $\sigma = 2.088 \times \sqrt{\frac{441}{11} + \frac{225}{11}}$

$291 - 10\bar{x}^2 = 332.1$
 $-10\bar{x}^2 = 33.1$
 $\bar{x} = 3.51$

$\sigma^2 = 10 \leq$
 $\frac{10 \times 441 + 10 \times 225}{11 + 11 - 2}$
 $= \frac{4410 + 2250}{20}$
 $= 333$

$\bar{y} - \bar{x} \pm z \leq \mu_Y - \mu_X \leq \bar{y} - \bar{x} \pm z$
 $554 - 540 \pm 1.96 \times \sqrt{\frac{225}{11} + \frac{441}{11}}$
 $14 \pm 1.96 \times \sqrt{\frac{666}{11}}$
 $14 \pm 1.96 \times \sqrt{60.545}$
 $14 \pm 1.96 \times 7.78$
 14 ± 15.25
 $-2.25 \leq \mu_Y - \mu_X \leq 29.25$

大同大學 89 學年度研究所招生入學考試試題

考試科目 統計學

所別 資訊經營研究所

第 3 / 3 頁

註：本次考試 參考自己的書籍及筆記 查字典 使用計算器

15. Using the statistics to perform hypothesis test on $H_0: \mu_X = \mu_Y$ versus $H_1: \mu_X \neq \mu_Y$ at significance level of $\alpha = 0.05$. Then, the conclusion will be: (a) to reject H_0 ; (b) to reject H_1 ; (c) to reject both; (d) the statistics are inadequate.

Answer problem 16 and 17 using the following problem setting:

To perform a single-factor ANOVA, suppose that we get the following partially filled table:

Source of Variation	df	Sum of Squares	Mean Square	f
Treatments	?	0.0608	?	?
Error	12	?	0.0308	?
Total	14	?		

16. Let m denote the number of populations (or treatments) involved in the ANOVA test, and let n denote the sum of sample sizes obtained from the m populations. Then, $m + n =$ (a) 14; (b) 15; (c) 16; (d) 18.
17. At significance level of $\alpha = 0.05$, we can conclude: (a) $\mu_1 = \mu_2 = \dots = \mu_m$; (b) there are some i, j , where $1 \leq i < j \leq m$, such that $\mu_i \neq \mu_j$; (c) $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_m^2$; (d) since the value of critical value $f_{0.05, \nu_1, \nu_2}$ is not available, no conclusion can be drawn.

Answer problem 18 to 20 using the following problem setting:

Given $n = 15$ data points (x_i, y_i) 's, $i=1, 2, \dots, 15$, the following quantifies were computed:

$S_{xx} = 50, S_{yy} = 25, S_{xy} = -30, \bar{x} = 1.3, \bar{y} = 27$
 $S_x = \sqrt{50} = 7.07, S_y = \sqrt{25} = 5$
 $r = \frac{S_{xy}}{S_x S_y} = \frac{-30}{7.07 \times 5} = -0.85$

18. To fitting the data points to a least squares line $y = \beta_1 + \beta_2 x + \epsilon$, $\hat{\beta}_1 =$ (a) $\frac{1}{2}$; (b) $-\frac{3}{5}$; (c) $-\frac{6}{5}$; (d) $\frac{270}{13}$.
19. The coefficient of determinations $r^2 = ?$ (a) 0.72; (b) 0.50; (c) 0.45; (d) 0.36.
20. In linear regression, we assume $\epsilon \sim N(0, \sigma^2)$. Using the above statistics, we get $\hat{\sigma} = ?$ (a) 1.52; (b) 1.62; (c) 1.72; (d) 1.82.

Handwritten calculations for problem 18:

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{-30}{50} = -\frac{3}{5}$$

Handwritten calculations for problem 19:

$$r^2 = \frac{S_{xy}^2}{S_{xx} S_{yy}} = \frac{(-30)^2}{50 \times 25} = \frac{900}{1250} = 0.72$$

Handwritten calculations for problem 20:

$$MSE = \frac{S_{yy} - \beta_2 S_{xy}}{n-2} = \frac{25 - (-\frac{3}{5})(-30)}{15-2} = \frac{25 - 18}{13} = \frac{7}{13}$$

$$\hat{\sigma} = \sqrt{MSE} = \sqrt{\frac{7}{13}} \approx 0.73$$