國立成功大學107學年度碩士班招生考試試題

系　所：企業管理學系

考試科目：統計學

考試日期：0206，節次：3

※ 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。

I. (28 points, 4 points each) Choose the most appropriate answer.

1. Let X be a Poisson random variable with E(X)=μ, median $\tilde{\mu}$ and mode μ^*. Then
   (a) $\mu < \tilde{\mu} < \mu^*$
   (b) $\mu^* < \tilde{\mu} < \mu$
   (c) $\tilde{\mu} < \mu < \mu^*$
   (d) $\mu < \mu^* < \tilde{\mu}$
   (e) None of the above.
2. Let (X_i, Y_i), i=1,2,…,n, be n pairs of random vectors, the Pearson sample correlation coefficient r is found to be 0.88, then
   (a) Strong linear association exists for X and Y.
   (b) The available information is not sufficient to claim the existence of linear association for X and Y.
   (c) If r is found to be 0 or near 0, then we can conclude that no relationship exists for X and Y.
   (d) True for (a), (c).
   (e) True for all the above (a), (b), (c), (d).
3. Let p_1 and p_2 be the proportions for some characteristic in populations 1 and 2. Random samples with size n_1 and n_2, respectively, are drawn from the two populations and found the sample proportions are $\hat{p}_1, \hat{p}_2$. We are interested in testing $H_0: p_1 = p_2$.
   (a) The test statistic t should be taken to be
   $$t = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$
   (b) The test statistic
   $$t = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}},$$
   where $\hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1+n_2}$, is better than the one in (a).
   (c) The test statistic in (b) is also good for the test $H_0: p_1 - p_2 = d_0$ vs $H_a: p_1 - p_2 \neq d_0$, where d_0 is some known value.
   (d) The test statistic in (a) is also good if $n_1 + n_2$ is large enough.
   (e) The above (a),(b),(d) are correct.

※ 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。

4. Two independent random samples are drawn from $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ with sizes $n_1 = 10,\ n_2 = 16$, respectively. It is found that $\overline{x_1} = 8,\ \overline{x_2} = 5,\ s_1^2 = 3.5,\ s_2^2 = 1$.

   (a) The random variable $\frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2}$ can be used to construct confidence interval for $\frac{\sigma_1^2}{\sigma_2^2}$ by using F random variable with 9 and 15 degrees of freedom.

   (b) Based on the results in (a), the test $H_0: \mu_1 = \mu_2$ would be concluded if $\alpha = 0.05$.

   (c) Based on the results in (a), to test $H_0: \mu_1 = \mu_2$, the test statistic to be used is a *t* with 26 degrees of freedom.

   (d) True for all above (a),(b),(c) and (d).

   (e) None of the above.

5. Consider the paired data: $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ and we want to compare the means of X and Y, μ_x, μ_y. Suppose we have $\bar{x}, \bar{y}, S_x^2, S_y^2$ and *r*, the sample correlation coefficient, where S_x^2 and S_y^2 are the unbiased estimators for σ_x^2, σ_y^2, and r is positive.

   (a) To test $H_0: \mu_x = \mu_y$, the test statistic $t = \frac{(\bar{x}-\bar{y})}{\sqrt{\frac{S_x^2}{n}+\frac{S_y^2}{n}}}$ is a good choice.

   (b) Let $d_i = x_i - y_i$, the test statistic $t_d = \frac{(\bar{x}-\bar{y})}{S_d\sqrt{\frac{1}{n}}}$ is a better choice than the *t* in (a) because $S_d^2 < S_x^2 + S_y^2$, where S_d is the sample standard deviation of d_i.

   (c) Since $S_d^2 = S_x^2 + S_y^2 - 2S_{xy}$, S_{xy} has to be given so that t_d in (b) can be computed, where S_{xy} is the sample covariance of X and Y.

   (d) True for the above (b) and (c).

   (e) None of the above.

6. Let $X_1, X_2, \ldots, X_n$, n≥4, be i.i.d. sample from some population with finite variance σ^2. Which of the following estimators is unbiased for σ^2 and has the smallest variance? ($\bar{X} = \sum_{i=1}^{n} X_i / n,\ \bar{X}_1 = \frac{\sum_{i=1}^{n_1} X_i}{n_1},\ \bar{X}_2 = \frac{\sum_{i=n_1+1}^{n} X_i}{n_2},\ n_1 + n_2 = n,\ n_1,\ n_2 \geq 2$)

※ 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。

(a) $X_1^2 - X_2X_3$
(b) $S^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2/(n-1)$
(c) $(X_1 - X_2)^2/2$
(d) $\hat{\sigma}^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2/n$
(e) $[\sum_{i=1}^{n_1}(X_i - \bar{X}_1)^2 + \sum_{i=n_1+1}^{n}(X_i - \bar{X}_2)^2]/(n_1 + n_2 - 2)$.

7. One box contains 1 red ball and 4 white balls. Three persons are to draw one ball from the box in order. Let X_i, i=1.2.3, be the outcome for person i that draws the red ball.
   (a) Each person, no matter his order in drawing ball from the box, has the same probability in drawing the red ball, $E(X_i)=1/4$.
   (b) The variances of each X_i are equal, which is 3/16.
   (c) The X_i, i=1.2.3, are identically distributed random variables.
   (d) The probability of drawing a red ball of the second person depends on the outcome of the first person.
   (e) True for all the above.

II. (12 points, three points each) For the following statements, cycle T for true and F for false.

1. Let A, B be sets in sample space S, ∅ be empty set to S.
(a) T F Sets A and ∅ are mutually exclusive,
(b) T F If A and B both are not empty sets, then they cannot be independent and must be mutually exclusive.
2. T F The skewness of a Poisson distribution is always positive. It cannot be negative.
3. T F Let X be an exponential distribution with parameter μ. Then the skewness of X, like normal distribution, can be zero, positive, or negative.

III. (60 points) Fill in the blanks for each problem.

1. (8 points, 4 points each)
   (a) The maximal value for the variance of a Bernoulli random variable with probability of success is __(A)__.

※ 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。

(b) In a survey sampling, what is the smallest sample size __(B)__ required if the margin of error (suppose α is 0.05) in estimating the population proportion p being set to be less than 0.03? Assume $0.2 \le p \le 0.4$.

2. (10 points, two point each) A sample with size n=27 is obtained with $\hat{y}=1.2-0.8x$, SSE (sum of squares due to error)=150, SSR (sum of squares due to regression)=24. Then R^2(coefficient of determination)=__(C)__, correlation coefficient of Y_i and the predicted value $\hat{Y}_i$=__(D)__ , correlation coefficient of Y_i and X_i =__(E)__, t=__(F)__ with__(G)__ degrees of freedom.

3. (3 points, one point each) Let X be a random variable taking two values 1 and 0, with $P(X=1) = \frac{3}{8}$; Y be another random variable taking two values 10 and 20, with $P(Y=10) = \frac{1}{4}$. It is known that $P(X=1, Y=10) = \frac{3}{32}$. Fill in the blanks for the following table

		Y		
		10	20	
X	0	(H)	(I)	
	1	3/32	(J)	3/8
		1/4		

4. (12 points, one point each) Let $X_1, X_2, \ldots, X_n$ be independent, identical distributed Bernoulli random variables with probability of success E(X) = p.
   (a) If $Y=\sum_{i=1}^{n} X_i$, then Y follow__(K)__distribution with mean__(L)__ and variance__(M)__.
   (b) (continued) If sample size n large, but p small, n×p constant, then Y approximates to a__(N)__distribution with mean__(O)__, variance__(P)__. If n×p is not small, say n×p ≥10, the distribution can be, in turn, approximated by__(Q)__distribution with mean__(R)__, and variance__(S)__.

※ 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。

(c) For binomial, if p is not too extreme, say $0.2 \le p \le 0.8$, the probability distribution can be approximated by a___(T)___distribution with mean___(U)___and variance___(V)___.

5. (27 points, three points each) The following data are collected to examine the existence of treatment effect:

Treatment		
1	2	3
8	14	10
7	16	12
9	12	16
13	17	15

(a) Fill in the blanks for the following ANOVA table

	SS	df	MS	F
SSB	(W)			(Y)
SSE	(X)			

Suppose the treatment effect does exist, and we want to know which of the differences, $\mu_1 - \mu_2$, $\mu_1 - \mu_3$, $\mu_2 - \mu_3$, contribute to the rejection of $H_0: \mu_1 = \mu_2 = \mu_3$ (μ_i: mean of treatment *i*) at α=0.05.
Give, respectively, the 95% joint confidence intervals for $\mu_1 - \mu_2$:___(Z)___, $\mu_1 - \mu_3$:___(Z1)___, $\mu_2 - \mu_3$:___(Z3)___.

(b) Apply Tukey's method to obtain a joint 95% confidence intervals for $\mu_1 - \mu_2$:___(Z4)___, $\mu_1 - \mu_3$:___(Z5)___, $\mu_2 - \mu_3$:___(Z6)___.
It is known, in this case, that $q_{0.05}$(k,df)=3.95.

$F_{9,15}(0.975) = 0.265$, $F_{9,15}(0.95) = 0.327$, $F_{9,15}(0.05) = 2.59$,
$F_{9,15}(0.025) = 3.12$
$t_{12}(0.05) = 1.782$, $t_{12}(0.025) = 2.179$, $t_{12}(0.01) = 2.681$,
$t_9(0.05) = 1.833$, $t_9(0.025) = 2.262$, $t_9(0.01) = 2.821$.