國立臺灣師範大學 114 學年度碩士班招生考試試題

科目：**統計學**　　適用系所：**管理研究所**

注意：1.本試題共 3 頁，請依序在答案卷上作答，並標明題號，不必抄題；必須列示計算過程，不可只有答案。
2.答案必須寫在指定作答區內，否則依規定扣分。(**註：可以使用電子計算機**)

1. Bag A contains 8 yellow flower seeds, 8 white flower seeds, and 8 purple flower seeds. Bag B contains 6 yellow flower seeds, 6 white flower seeds, and 12 purple flower seeds. One bag is randomly selected. Please answer the following questions:
   (1) If three seeds are randomly selected from the chosen bag, and all of them grow into purple flowers, calculate the probability that the selected bag is Bag B. ***(15 points)***
   (2) If three seeds are randomly selected from the chosen bag, and one grows into a yellow flower, one into a white flower, and one into a purple flower, calculate the probability that the selected bag is Bag A. ***(15 points)***
2. A finished product in a production line can be classified as good, second, or defective with probabilities of 0.6, 0.3, and 0.1, respectively. Suppose 15 items are randomly selected from the production line. Let X represent the number of good items, Y represent the number of second items, and $15-X-Y$ represent the number of defective items. Answer the following questions:
   (1)What is the joint probability mass function (*p.m.f.*) of X and Y? ***(5 points)***
   (2)Calculate the probability $P(X=10, Y=4)$. ***(5 points)***
   (3)What is the marginal *p.m.f.* of X? ***(5 points)***
   (4)Compute the expectation and variance of X. ***(5 points)***
3. Starbucks has 4 stores in the Da'an District, each with some customers who filled out satisfaction surveys. Higher scores indicate greater satisfaction. The data is organized in the table below.

	Store A	Store B	Store C	Store D	Total
	94	75	70	68	
	90	68	73	70	
	85	77	76	72	
	80	83	78	65	
		88	80	74	
			68	65	
			65		
Sample size n=	4	5	7	6	22
Mean M=	87.25	78.20	72.86	69.00	75.64
Variance s^2=	36.92	58.70	30.14	13.60	70.72
Sum of squares SS=	110.75	234.80	180.86	68.00	1485.09

(1) Present the summary table of the analysis of variance (ANOVA) to display the results of the analysis of variance. ***(10 points)***

(2) Write down the null and alternative hypotheses, and at $\alpha=0.05$, test whether there is a significant difference in customer satisfaction among the four stores. (Note: $F_{(0.025, 3,18)}=3.954$; $F_{(0.025, 3,21)}=3.819$; $F_{(0.05, 3,18)}=3.160$; $F_{(0.05, 3,21)}=3.072$) ***(10 points)***

4. A market researcher collected data from a sample of 10 families to explore the relationship between household expenditure, family size, and socioeconomic status. After conducting an initial analysis, the sample descriptive statistics are summarized in the table below:

	mean	Standard deviation	*SS and SP			Covariances and variances		
			X_1	X_2	Y	X_1	X_2	Y
X_1	5	1.826	30			3.333		
X_2	5	2.582	27	60		3.000	6.667	
Y	30	10.687	167	182	1028	18.556	20.222	114.222

*SS: sum of products, SP: sum of cross products

(1) The market researcher aims to construct a multiple regression model to predict household expenditure (Y) using family size (X_1) and socioeconomic status (X_2) as independent variables. Using the sample statistics provided in the table above, please calculate the sample regression coefficients: b_0 (intercept), b_1, and b_2 (slopes for X_1 and X_2) to support the researcher's analysis. ***(10 points)***

(2) Given that the residual sum of squares ($\sum_{=1}^{10} e^2 = 70.218$), please calculate the model's explanatory power, including the coefficient of determination ($\boldsymbol{R^2}$) and the adjusted $\boldsymbol{R^2}$. ***(10 points)***

(3) Please compute the overall F-statistic for the model to evaluate its significance. Based on the results, perform a hypothesis test to determine whether the model as a whole is statistically significant. You should write the null and alternative hypotheses before conducting the hypothesis test. (Note: $F_{(0.025, 2,7)}=6.542$; $F_{(0.025, 2,9)}=5.715$; $F_{(0.05, 2,7)}=4.737$; $F_{(0.05, 2,9)}=4.256$). ***(10 points)***

Hint: The following three equations are based on the least squares approach and can be used to calculate b_0, b_1, and b_2.

$$b_0 = \bar{y} - b_1\bar{x}_1 - b_2\bar{x}_2$$

$$\sum(x_1 - \bar{x}_1)(y - \bar{y}) = b_1\sum(x_1 - \bar{x}_1)^2 + b_2\sum(x_1 - \bar{x}_1)(x_2 - \bar{x}_2)$$

$$\sum(x_2 - \bar{x}_2)(y - \bar{y}) = b_1\sum(x_1 - \bar{x}_1)(x_2 - \bar{x}_2) + b_2\sum(x_2 - \bar{x}_2)^2$$