I .(20%) The mean number of hours during a work week for the population of workers was reported to be 44.75 hours. Suppose we take a current sample of workers to see whether the mean number of hours over a work week has changed from the previously reported 44.75 hours.

    a. (5%) State the hypotheses that will help us determine whether a change occurred in the mean length of work week? [Notations have to be defined properly.]

    b. (10%) Suppose a current sample of 121 workers provided a sample mean of 43.95 hours. What is the p-value if the population standard deviation is 5.48 hours?

    c. (5%) What is your conclusion assuming the significance level is 0.05?

II .(30%) Consider a study involving the graduate management admission test (GMAT), a standardized test used by graduate schools of business to evaluate an applicant's ability to pursue a graduate program in that field. Scores on the GMAT range from 200 to 800, with higher scores implying higher aptitude. In an attempt to improve student's performance on the GMAT exam, a major Texas university is considering offering the following three GMAT preparation programs.

    1. A three-hour review session covering the types of questions generally asked on GMAT.

    2. A one-day program covering relevant exam material, along with the taking and grading of a sample exam.

    3. An intensive 10-week course involving the identification of each student's weaknesses and the setting up of individualized programs for improvement.

The following table summaries the GMAT data collected:

		College		
		Business (1)	Engineering (2)	Arts and Science (3)
Preparation program	1	500, 580	540, 460	480, 400
	2	460, 540	560, 620	420, 480
	3	560, 600	600, 580	480, 410

    a. (10%) Write down the ANOVA model that corresponding to the following ANOVA and state your finding based on the significance level 0.05. [Notations have to be defined properly.]
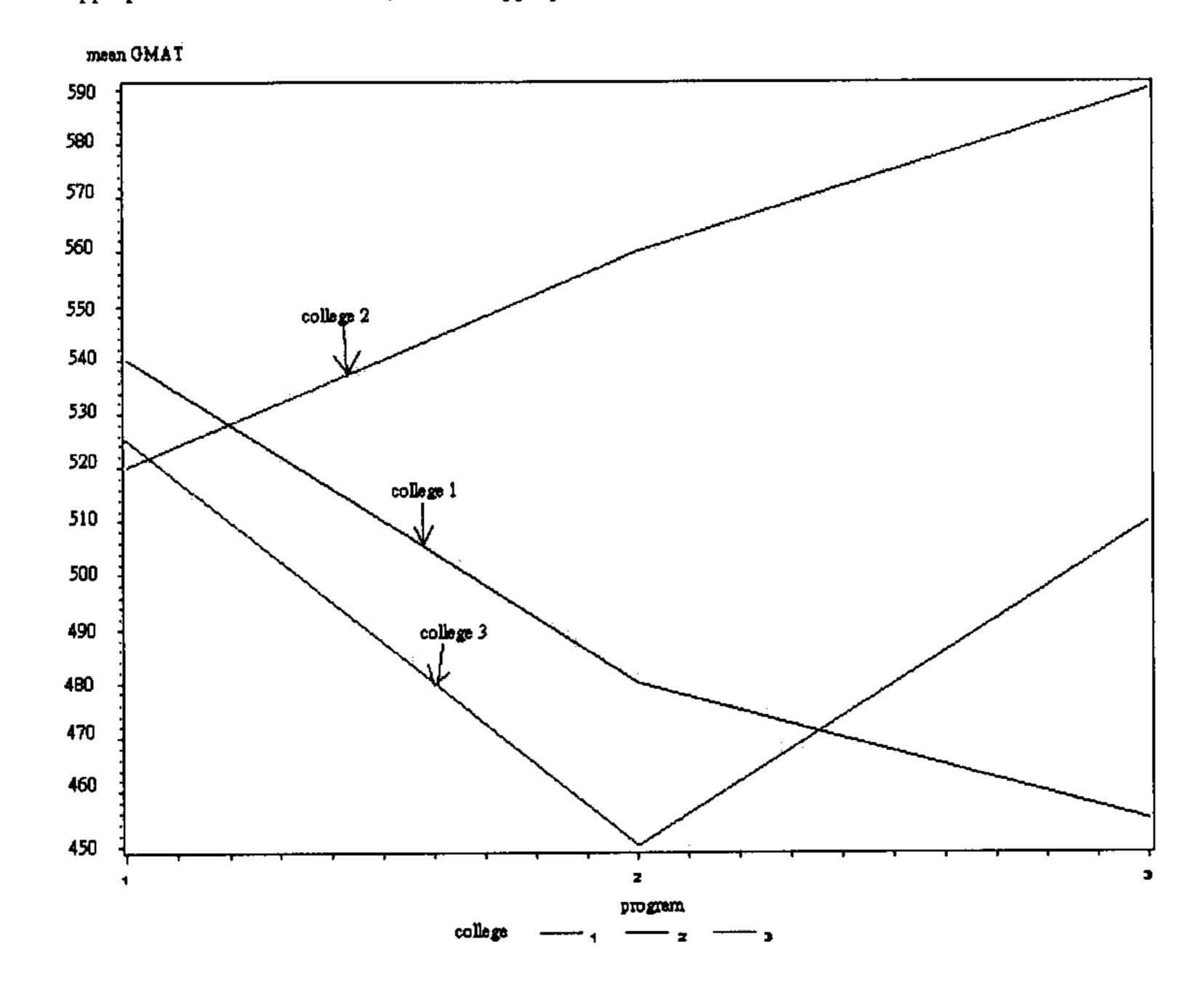
試題隨卷繳交　　　　　　　　　接背面

The ANOVA Procedure

Dependent Variable: GMAT

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	51400.00000	12850.00000	5.38	0.0088
Error	13	31050.00000	2388.46154		
Corrected Total	17	82450.00000			

R-Square	Coeff Var	Root MSE	GMAT Mean
0.623408	9.489687	48.87189	515.0000

Source	DF	Anova SS	Mean Square	F Value	Pr > F
program	2	6100.00000	3050.00000	1.28	0.3117
college	2	45300.00000	22650.00000	9.48	0.0029

b. (5%) Using a significance level 0.05, state your finding based on a.

c. (5%) Interpret the meaning of R-square.

d. (10%) The following Figure lists the mean GMAT score for preparation and college. Is the model from a. appropriate for this data? If not, state the appropriate model. [Notations have to be defined properly.]

Ⅲ.(15%) Use the following to answer question 1-5

During the early part of the 2004 baseball season, many sports fans and baseball players noticed that the number of home runs being hit seemed to be unusually large. Below are the team by team statistics on home runs hit through Friday, June 4, 2004(from the Columbus Dispatch Sports Section, Sunday, June 6, 2004). Below are separate stemplots for the number of home runs by American and National League teams.

American League	National League
2		2	9
3	5	3	1
4	0 3 9	4	2 6 7 8 8
5	1 4 7 8 8	5	3 5 5 5
6	4 8 8	6	3 3 7
7	5 7	7	

(　　　)1. The median for the number of home runs for the American League teams is
　　a) lower than for the National League teams.
　　b) 45
　　c) 50
　　d) 57.5

(　　　)2. The median for the number of home runs for the National League teams is
　　a) higher than for the American League teams.
　　b) 46
　　c) 50.5
　　d) 55

(　　　)3. Which of the following is a correct statement ?
　　a) The American League plot is reasonably symmetric.
　　b) The National League plot is slightly skewed to the left.
　　c) The median number of home runs hit by American League teams was higher than by National League teams.
　　d) All of the above

(　　　)4. The mean for the number of home runs for the American League teams is _____

(　　　)5. The mean for the number of home runs for the National League teams is _____

Ⅳ.(18%)張森先生同時擲一骰子 (x)，丟一錢幣(y) 及抽一張撲克牌 (z)。骰子每點可得 6 元，正面錢幣可得 10 元，撲克牌則一點一元（J =11 點，Q = 12 點，K=13 點），此三種實驗相互獨立，隨機變數 x、y、z 同為機率均等分配，試求：
(1) 期望值 E(x), E(y), E(z)
(2) 變異數 V(x), V(y), V(z)
(3) 張森所得錢數為 W，E(W)及 V(W)為多少？

V.(17%) 計算下面各題之正確機率:

1. You are tossing a balanced die that has probability 1/6 of coming up 1 on each toss. Tosses are independent.

   a. The probability of a 1 on the first toss is 1/6. What is the probability that the first toss is not a 1 and the second toss is a 1? (3%)

   b. What is the probability that the second 1 occurs on the fifth toss? (5%)

2. The weight of the eggs produced by a certain breed of hen is normally distributed with mean 65 grams (g) and standard deviation 5g. Think of cartons of such eggs as random sample of size 12 from the population of all eggs. What is the probability that the weight of a carton fall between 750g and 825g? (9%)

常態分配表

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \quad \Phi(z) = P(Z \leq z) = \int_{-\infty}^{z}$$

試題隨卷繳交