考試科目	統計學	所(組)別	統計學系	考試時間	102 年 11 月 3 日 星期日 10：00－11：40

Part 1

1. (10%)
Listed below are scores (from smallest to largest) of 29 students for midterm calculus exam.

47 47 52 58 60 62 63 66 68 73 73 73 75 76 77 77 79 79 80 82 83 83 85 86 86 87 87 90 91

(a) (4%) Determine whether the data is left-skewed, symmetric, or right-skewed. (Briefly explain why.)
(b) (6%) Determine the first and the ninth deciles.

2. (10%)
We are interested in the population in Taiwan.

(a) (5%) The population was estimated to be about 22301000 in 2000, and 23162000 in 2010. What is the average annual growth rate from 2000 to 2010?
(b) (5%) The average annual population growth rate is about 1.8% from 1920 to 1925, and about 2.8% from 1925 to 1930. Suppose that the population in 1930 was about 4593000. What was the population in 1920?

(背面還有試題)

考試科目	統計學	所(組)別	統計學系	考試時間	102 年 11 月 3 日 星期日 10：00－11：40

Part 2

1. (12%) An Olympic archer is able to hit the bull's-eye 80% of the time. Assume each shot is independent of the others.
   (a) If she keeps shooting arrows until she hits the bull's-eye, how long do you expect it will take?
   (b) What is the probability that her first bull's-eye comes on the fourth or fifth arrow.
   (c) Suppose she shoots 10 arrows. What is the probability that she hits the bull's-eye more than she misses?
   (d) The archer will be shooting 200 arrows in a large competition. Would you be surprised if she made only 140 bull's-eyes or less? Explain.

2. (8%) In a large city school system with 20 elementary schools, the school board is considering the adoption of a new policy that would require elementary students to pass a test in order to be promoted to the next grade. The PTA wants to find out whether parents agree with this plan. Listed below are some of the ideas proposed for gathering data. For each, indicate what kind of sampling strategy is involved and what (if any) biased might result.
   (a) Put a big advertisement in the newspaper asking people to log their opinions on the PTA Web site.
   (b) Randomly select one of the elementary schools and contact every parent by phone.
   (c) Send a survey home with every student, and ask parents to fill it out and return it the next day.
   (d) Randomly select 20 parents from each elementary school. Send them a survey, and follow up with a phone call if they do not return the survey within a week.

TABLE OF RANDOM DIGITS

Row										
1	96299	07196	98642	20639	23185	56282	69929	14125	38872	94168
2	71622	35940	81807	59225	18192	08710	80777	84395	69563	86280
3	03272	41230	81739	74797	70406	18564	69273	72532	78340	36699
4	46376	58596	14365	63685	56555	42974	72944	96463	63533	24152
5	47352	42853	42903	97504	56655	70355	88606	61406	38757	70657
6	20064	04266	74017	79319	70170	96572	08523	56025	89077	57678
7	73184	95907	05179	51002	83374	52297	07769	99792	78365	93487
8	72753	36216	07230	35793	71907	65571	66784	25548	91861	15725
9	03939	30763	06138	80062	02537	23561	93136	61260	77935	93159
10	75998	37203	07959	38264	78120	77525	86481	54986	33042	70648

(下頁還有試題)

考試科目	統計學	所(組)別	統計學系	考試時間	102 年 11 月 3 日 星期日 10：00－11：40

part 3 (20%)

1. The management of the Lion baseball team decided to sell only low alcohol beer in their ballpark to help combat rowdy fan conduct. They claimed that more than 40% of the fans would approve of this decision. Let p equal the proportion of Lion fans on opening day who approved of this decision. We shall test the null hypothesis $H_0: p = 0.4$ against alternative hypothesis $H_1: p > 0.40$.
   (a) Define a critical region that has an $\alpha = 0.05$ significant level. (4%)
   (b) If out of a random sample of $n = 1278$ fans, $y = 550$ said that they approved of this new policy; what is your conclusion? (6%)

2. Consider the butterfat production (in ponds) for a cow during a 305-day milk production period following the birth of a calf. Let X and Y equal the butterfat production for such cows on a farm in Wisconsin and a farm in Michigan. Twelve observations of X are:

   649 657 714 877 975 468
   567 849 721 791 874 405

   Sixteen observations of Y are:

   699 891 632 815 589 764 524 727
   597 868 652 978 479 733 549 790

   (a) Assuming that X is $N(\mu_x, \sigma^2)$ and Y is $N(\mu_Y, \sigma^2)$, find a 95%confidence interval for $\mu_X - \mu_Y$. (4%)
   (b) Construct a box-and whisker diagrams for these two sets of data on the same graph. (4%)
   (c) Does these seems to be a significant difference in butterfat production for cows on these two farms? (2%)

(背面還有試題)

考試科目	統計學	所(組)別	統計學系	考試時間	102 年 11 月 3 日 星期日 10：00－11：40

Part 4

1. (20%)

(a) The error terms e_i's of the regression equation $Y_i = \beta_0 + \beta_1 X_i + e_i$ are usually assumed to follow normal distribution $N(0, \sigma^2)$. State the reasons why the conditions normality and constant variance are required. (Or, you may state what would go wrong if they are not valid.)

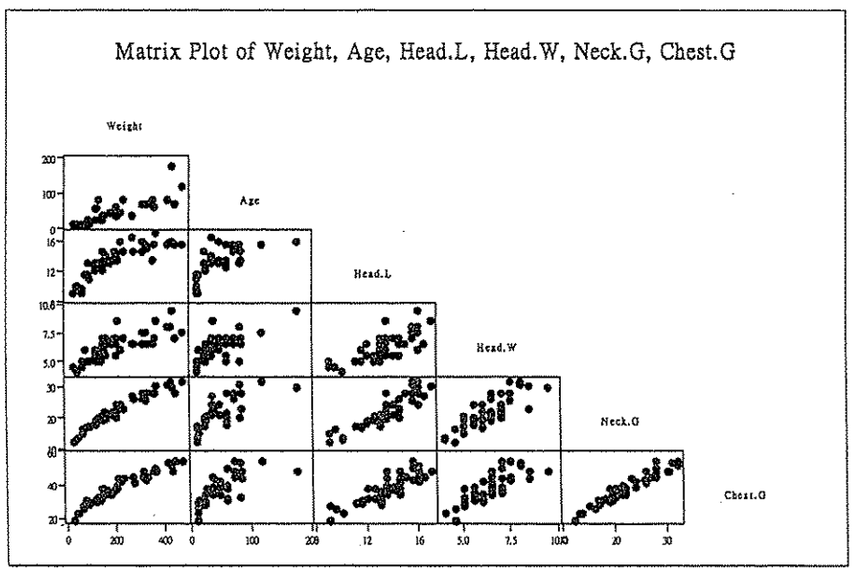
(b) We are interested in using the multiple regression, together with the independent variables Age, Head.L, Head.W, Neck.G, and Chest.G, to determine the weight of polar bears. (Data Source: Minitab) Judging from the following matrix plot and output of regression analysis, comment on the regression equation. (For example, is this regression equation acceptable or what should be done to improve the fitting?)

```
The regression equation is
Weight = - 191 + 0.704 Age - 8.77 Head.L - 4.61 Head.W + 11.7 Neck.G
         + 6.62 Chest.G

35 cases used, 15 cases contain missing values

Predictor      Coef  SE Coef      T      P     VIF
Constant   -191.13    38.67  -4.94  0.000
Age         0.7038   0.2033   3.46  0.002   2.695
Head.L      -8.773    4.742  -1.85  0.075   4.684
Head.W      -4.614    6.365  -0.72  0.474   3.162
Neck.G      11.711    3.591   3.26  0.003  18.114
Chest.G      6.619    1.763   3.75  0.001  14.063

S = 25.9146   R-Sq = 96.3%   R-Sq(adj) = 95.7%
```



(下頁還有試題)

考試科目	統計學	所(組)別	統計學系	考試時間	102 年 11 月 3 日 星期日 10：00－11：40

Part 5

1. (4pts). Please determine whether each of the following statements is true or false.

   (a) Parametric tests are preferred over their nonparametric counterparts, if the assumptions can be met.

   (b) The sign test can be used to see whether a median is greater than a specific value.

2. (2pts). Which of the following tests can be used to test whether two population means are differ when samples are dependent and the normality assumption cannot be met.

   (A) paired Student's t-test (B) Wilcoxon rank sum test

   (C) Wilcoxon signed-rank test (D) ANOVA F-test

3. (14pts). Listed below is information regarding the sale of a furniture retailer with store in Taipei for three different months.

	Type of Furniture		
	Sofa	Recliner	Loveseat
June	14	11	17
July	7	13	10
August	11	8	9

Based on these data, is there sufficient evidence at $\alpha = 0.05$ to conclude that a relationship exists between month and type of furniture sold ? Please include the test hypotheses, test statistics, decision rule, and your conclusion. Show your work.

備考：試題隨卷繳交