

考試科目	統計學	所(組)別	統計學系	考試時間	11月22日 星期六 10:00-11:40
------	-----	-------	------	------	---------------------------

1. Suppose 5% of birds are infected by avian flu. Within this 5% of infected birds, 60% are infected by type A avian flu and 40% are infected by type B avian flu. A researcher at the Academia Sinica has developed a test for these two types of avian flu. 90% of birds with type A avian flu will have positive reaction to the test. 80% of birds with type B avian flu will have positive reaction to the test. 92% of birds without avian flu will have negative reaction to the test. What is the probability that a bird is infected by type A avian flu when the test is positive. (10%)

2. An elevator carries nine people in a load. The weights of people vary according to many factors but may be described by a normal distribution with mean 160 pounds. Over many loads, each of nine people, there are about 0.3% of loads will exceed the safe load limit of 1800 pounds. What is the standard deviation of this normal distribution? (10%)

1. (15 points) The following is the number of goals scored during each game of a baseball team in the past season:
3, 5, 2, 2, 1, 4, 3, 3, 2, 1, 8, 2, 4, 1, 1, 3, 2, 5, 6.
- (a) (3 points) Construct a frequency plot of the data.
 - (b) (2 points) Is the distribution symmetric, right-skewed, or left-skewed?
 - (c) (3 points) Give the mean, median, and mode of the data.
 - (d) (5 points) Construct a boxplot. (Indicate the five-number summary on the plot.)
 - (e) (2 points) About 25% of the time, the team scored at least x goals. What is the value of x ?
2. (5 points) During the decade of the 1990s, and into the 2000s, the population in city A increased from 258,295 in 1990 to 534,847 in 2005. What is the average annual increase over the 15-year period?

考試科目	統計學	所(組)別	統計學系	考試時間	11月22日 星期六 10:00-11:40
------	-----	-------	------	------	---------------------------

三.

A large sample of loafs of bread was taken to count the number of raisins contained in each loaf. The average number of raisins per loaf in the sample is 40 with a sample standard deviation of 12.

- (a) The 95% confidence interval for the mean number of raisins based on the sample is between 36.08 and 43.92. What is the sample size?
- (b) How large a sample would be required so that the margin of error is 3 raisins at 98% confidence?

(4%, 4%)

First Bank wants to know if the mean amount of cash withdrawn per transaction from the ATM is greater than \$5000. A sample of 36 cash withdrawn transactions will be taken. Assume $\sigma = \$900$, and a 0.05 level of significance will be used in the hypothesis testing.

- (a) What is the probability of making a type II error if the true mean is \$5100?
- (b) Now ignore part (a), if the sample average is \$5300, what is the p-value? What is your conclusion?

(6%, 6%)

四.

The following data represent the distribution of ages for a sample of observed purchasers of major league baseball game tickets. Use the chi-square goodness-of-fit test to determine whether this distribution is the normal distribution. Assume that $\alpha = 0.05$.

(20%)

Age of purchaser	Frequency
10-under 20	16
20-under 30	44
30-under 40	61
40-under 50	56
50-under 60	35
60-under 70	19

考試科目	統計學	所(組)別	統計學系	考試時間	11月22日 星期六 10:00-11:40
------	-----	-------	------	------	---------------------------

- 五. The following data are from 10 ski resorts in New England during a period of normal snow conditions. Here,
 Y = total # visitor-days, in thousands of days
 X_1 = miles of intermediate trails
 X_2 = lift capacity, skiers/hour, in thousands of skiers

X_1	X_2	Y	
10.5	2.200	19.929	
2.5	1.000	5.839	$\sum X_{1i} = 101.4$
13.1	3.250	23.696	$\sum X_{1i}^2 = 1419.78$
4.0	1.475	9.881	$\sum X_{2i} = 26.45$
14.7	3.800	30.011	$\sum X_{2i}^2 = 90.21125$
3.6	1.200	7.241	$\sum Y_i = 202.459$
7.1	1.900	11.634	$\sum Y_i^2 = 5741.9988$
22.5	5.575	45.684	
17.0	4.200	36.476	$\sum X_{1i}X_{2i} = 356.4225$
6.4	1.850	12.068	$\sum X_{1i}Y_i = 2848.8415$

- (a) Obtain the least squares estimates for the parameters in the model (3%)
 $Y_i = \beta_0 + \beta_1 X_{1i} + \varepsilon_i$, where ε_i is a random error, and construct an analysis of variance table.

	df	SS
Model		
Error		
Total		1643.0341

- (b) Assuming the model in (a) with normal random errors, conduct a test of (3%)
 $H_0 : \beta_1 = 0$ vs. $H_1 : \beta_1 \neq 0$ at level of significance 0.05.

考試科目	統計學	所(組)別	統計學系	考試時間	11月22日 星期六 10:00-11:40
------	-----	-------	------	------	---------------------------

五. (c) Below are some results from a least squares fit to the model . (4%)

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i \quad \hat{\beta}_0 = -1.8068, \hat{\beta}_1 = 1.1310, \hat{\beta}_2 = 4.0015,$$

Estimated standard error of $\hat{\beta}_1 = 0.6158$.

Complete the analysis of variance table:

Source	df	SS	MS	F
Model		1623.7302		
Error		19.3039		
Total		1643.0341		

(d) Write down the null and alternative hypotheses tested by the F-ratio in the ANOVA in (c). What is the result of this F-test (conducted at level 0.05)? (4%)

(e) For the model (c), conduct tests of the hypotheses (6%)

(i) $H_0 : \beta_1 = 0$ vs. $H_1 : \beta_1 \neq 0$

(ii) $H_0 : \beta_2 = 0$ vs. $H_1 : \beta_2 \neq 0$