IE 121 – Spring 2003

Midterm exam 2

1. (20 points) In this problem, you need to select the correct answer.

   1. Which of the following distributions is asymmetric?
      a) χ^2 distribution
      b) normal distribution
      c) t distribution

   2. Suppose we want to test a goodness of fit to a normal distribution with mean equal to 12.5. The number of bins is chosen to be 8. What is the number of degrees of freedom that we have to use for the corresponding χ^2 test?
      a) 8
      b) 6
      c) 7
      d) 5

   3. The paired t-test has to be chosen over the pooled t-test if
      a) there is evidence for a strong positive correlation between pairs of observations
      b) the available sample size is small
      c) never – the pooled t-test is always better

   4. The null hypothesis is H_0: $\sigma^2 = \sigma_0^2$. Which test statistic has to be used for the hypothesis testing?
      a) z_0
      b) t_0
      c) χ_0^2

   5. Suppose that we have found that, with the given sample size $n_1 = n_2 = 300$, the power of the test of the hypothesis H_0: $p_1 = p_2$ was found to be 0.15. We are interested in the (common for the two populations) sample size required to to make the power of the test equal to 0.95. Which of the following can realistically be the right answer?
      a) 250
      b) 10000
      c) 305

   Answers: a-b-a-c-b

IE 121 – Spring 2003

Midterm exam 2

2. (30 points) The copper content in 1 g of an alloy is normally distributed. We want to conclusively demonstrate that the variance of the copper content is less than 20 $(\text{mg})^2$.

a) Formulate the appropriate null hypothesis and its alternative. What (rejection or not) should be considered a desirable outcome for our purpose?

b) The sample variance obtained from the sample of size 25 is equal to 15 $(\text{mg})^2$. What conclusion will we draw about our hypothesis at significance level of $\alpha = 0.05$?

c) Find the P-value (or its range) for the above test.

d) Determine a 95% confidence upper-confidence interval for σ^2.

Solution:

a) H_0: $\sigma^2 = 20$

H_1: $\sigma^2 < 20$

The desirable outcome is the rejection of H_0.

b) $\chi_0^2 = \frac{24 \cdot 15}{20} = 18 > 13.85 = \chi^2_{0.95,24}$ – we can't reject the null hypothesis.

c) From the χ^2 table, we get $0.1 < P < 0.5$

d) $\sigma^2 \leq \frac{24 \cdot 15}{13.85} = 25.99$

IE 121 – Spring 2003

Midterm exam 2

3. (25 points) Two different methods of advertising a certain product are compared. For this purpose two samples of 1000 customers each from the populations exposed to the first and the second method, respectively, are studied. It was found that, in the first sample, 150 people bought the product, and in the second, 180 people made the corresponding purchase.

a) Do these data indicate that there is a difference in the proportions of people who react positively to the advertising between the two populations? Use $\alpha = 0.05$.

b) Find the P-value for this test.

c) Suppose that the true proportions are equal to $p_1 = 0.14$ and $p_2 = 0.18$, respectively. Determine the sample size needed to detect this difference with probability of at least 0.9. Use $\alpha = 0.05$

Solution:

a) $\hat{p} = \frac{180+150}{1000+1000} = 0.165$

$z_0 = \frac{0.18-0.15}{\sqrt{0.165 \cdot 0.835\left(\frac{1}{1000}+\frac{1}{1000}\right)}} = 1.807 < 1.96 = z_{0.025}$ – so these data do not indicate the difference in proportions at 5% significance level.

b) Using the z-table we get: $P = 2(1 - 0.965) = 0.07$.

c) $n = \frac{\left(1.96\sqrt{(0.14+0.18)(0.86+0.82)/2}+1.28\sqrt{0.14 \cdot 0.86+0.18 \cdot 0.82}\right)^2}{(0.14-0.18)^2} = 1308$

IE 121 – Spring 2003

Midterm exam 2

4. (25 points) We want to conduct a goodness of fit test for the data in the table below. The null hypothesis states that these data came from a uniform distribution.

a) Estimate the parameters a and b of the uniform distribution from the data, using the MLE estimators: $\hat{a} = \min_i X_i$ and $\hat{b} = \max_i X_i$.

b) Choose the bins for the goodness of fit test of the above hypothesis so that the expected frequency in each bin is no less than 3 and the resulting number of degrees of freedom for the test is positive[1].

c) Test the above hypothesis at $\alpha = 0.05$ significance level.

i	1	2	3	4	5	6	7	8	9	10	11	12
x_i	2.7	2.5	4.6	1.0	5.0	4.2	4.5	4.4	2.3	2.4	3.2	3.8

Solution:

a) $\hat{a} = \min_i x_i = 1.0$

$\hat{b} = \max_i x_i = 5.0$

b) Since we have 12 observations, and we want the expected frequency in each bin to be no more than 3, we can have at most 4 bins, i.e. $k \leq 4$. On the other hand, since we have to estimate both parameters of the uniform distribution from the data, the number of degrees of freedom for the χ^2 test is $k - 2 - 1 = k - 3$. For this number to be positive, we need $k \geq 4$. Combining these two restrictions, we conclude that $k = 4$ necessarily. So the first bin will contain observations from 1.0 to 2.0, second – from 2.0 to 3.0, third – from 3.0 to 4.0, third – from 4.0 to 5.0. Thus, for the expected frequencies we get: $E_1 = E_2 = E_3 = E_4 = 3$.

c) From the table, the observed frequencies are: $O_1 = 1$, $O_2 = 4$, $O_3 = 2$, $O_4 = 5$.

$\chi_0^2 = \frac{1}{3}((1-3)^2 + (4-3)^2 + (2-3)^2 + (5-3)^2) = 3.33 < 3.84 = \chi^2_{0.05,1}$, therefore we can't reject the null hypothesis at 5% significance level.

[1] These two requirements taken together should leave you with a unique choice