

THE UNIVERSITY OF SYDNEY
STAT2012 STATISTICAL TESTS

Semester 2	Tutorial Week 12	2015
------------	-------------------------	------

Summary of week 11

- *Chi-square* test on proportions:

If y_1, y_2, \dots, y_k are the observed frequencies of k classes, the test statistic for the hypothesis on the k class proportions:

$$H_0 : p_1 = p_{10}, p_2 = p_{20}, \dots, p_k = p_{k0} \quad H_1 : \text{Not all equalities hold}$$

is

$$\chi_0^2 = \sum_{i=1}^k \frac{(y_i - np_{i0})^2}{np_{i0}} = \sum_{i=1}^k \frac{y_i^2}{np_{i0}} - n \sim \chi_{k-1}^2$$

Note that the χ_{k-1}^2 approximation is good only when the expected frequencies $E_i = np_i \geq 5$.

Tutorial Questions

1. Portable personal computers, sometimes called laptops, represent a fast-growing segment of the PC market. According to the Market Intelligence Research Company, the use of portable laptop computers can be classified into the following user segments.

User Segment	1988 percentages	Current survey frequency
Business-Professional	69%	102
Government	21%	32
Education	7%	12
Home	3%	4
Total	100%	150

Do the data provide sufficient evidence to indicate that the figures obtained in the current survey agree with the percentages in 1988?

2. A study of patients with insulin-dependent diabetes was conducted to investigate the effects of cigarette smoking on renal and retinal complications. Before examining the results of the study, a researcher expects that the proportions of four different subgroups are as follow:

Subgroup	Proportion
Nonsmokers	0.50
Current Smokers	0.20
Tobacco Chewers	0.10
Ex-smokers	0.20

Of 100 randomly selected patients, there are 44 nonsmokers, 24 current smokers, 13 tobacco chewers and 19 ex-smokers. Should the researcher revise his estimates? Use 0.01 level of significance.

3. A scientist believes that the gender of a child is a Bernoulli random variable with probability = 0.5 for a boy and 0.5 for a girl. To help test her belief she randomly samples 100 families with 4 children and records the number of boys as below:

Number of boys	0	1	2	3	4
Number of families	10	18	35	28	9

Can the scientist infer at the 5% significance level that the number of boys in families with 4 children is a binomial random variable.

Extra Practice Problems

1. To deal with the water pollution problem, three proposals are suggested: 1. Remove the industrial plant, 2. Relocate the industrial plant to the river mouth and 3. Build a sewage plant. Thirty government officials are interviewed and their opinions are given below. Test, at the 5% level of significance, the null hypothesis that there is no preference among the three proposals. To facilitate the working, you may complete the following table:

Proposal	Observed O_i	Expected E_i	$(O_i - E_i)$	$\frac{(O_i - E_i)^2}{E_i}$
Remove the plant	6			
Relocate the plant to river mouth	9			
Build a sewage plant	15			
Total	30			$\chi_0^2 =$