Science Maths 3

MA4704 Lecture 7B

Kevin O'Brien

Kevin.obrien@ul.ie

Dept. of Mathematics & Statistics,
University *of* Limerick

Autumn Semester 2016

The data for these tests are contingency tables showing the relationship between 2 qualitative variables. For example, suppose we have the following information regarding hair and eye colour

	Red hair	Blonde hair	Dark hair	$\sum$
Blue eyes	30	90	30	150
Brown eyes	20	60	70	150
$\sum$	50	150	100	300

i.e. 30 people have red hair and brown eyes.

Chi Square Test For Independence

Test for independence

Let $O_{i,j}$ be the entry in the i-th row and j-th column of the contingency table. We wish to choose between the hypotheses

$\mathbf{H_0}$: hair colour and eye colour are independent.

$\mathbf{H_1}$: hair and eye colour are dependent.

Chi Square Test For Independence

Row and column sums

- The number of people in the sample with blue eyes is the sum of the entries in the first row (150).
- The number of people in the sample with brown eyes is the sum of the entries in the second row (150).
- The sum of all the entries is the number of individuals in the sample (300).

Chi Square Test For Independence

- If the traits are independent, then the probability that an individual has a given hair colour and given eye colour is the product of the two corresponding probabilities e.g.

$$\text{P(blond hair, blue eyes)} = \text{P(blond hair)} \times \text{P(blue eyes)}$$

- In order to test whether two traits are independent, we need to calculate what we would expect to observe if the traits were independent.
- The following calculations allow us to calculate what we expect to see under the null hypothesis of independence.