

Contingency Tables and Analysis of Variance

Math 130: Introductory Statistics

1

Goals of this section

Goal of this section

- Learn how to test for dependence/independence for categorical data.

Contingency table

2 Contingency table

Contingency table

- **Contingency table** (aka **two-way frequency table**) is a table which reports frequencies corresponding to two different sets of categories (variables).
- Example: mortality rates on the Titanic:

	Men	Women	Boys	Girls	Total
Survived	332	318	29	27	706
Died	1360	104	35	18	1517
Total	1692	422	64	45	2223

- Data in the table are always frequencies that fall into individual categories.
- Could use this table to test if two variables are independent.

3 Test for Independence

3.1 Chi-squared distribution

Test of independence

Test of independence

- **Null hypothesis:** there is no association between the row variable and the column variable.
- that is, the two variables are independent.
- **Alternative hypothesis:** The two variables are dependent.
- To test such a hypothesis, we'll need to use a new distribution
- **Chi-squared distribution:** a distribution skewed to the right whose support is always positive.
 - Only use one tailed test.
 - Values of χ^2 close to zero imply independence.
 - Large values of χ^2 occur when variables are dependent.

3.2 Test statistic

Test statistic

Test statistic

- Test statistic:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

- O : observed frequency in a cell from the contingency table.
- E : expected frequency assuming the row and column variable are independent.

$$E = \frac{(\text{row total})(\text{column total})}{\text{grand total}}$$

- Degrees of freedom = $(r - 1)(c - 1)$
 - r = number of rows.
 - c = number of columns.

4 Example

Example

Example

- Do helmet laws make a difference in the number of fatalities?

	No helmet law	Helmet Law
fatal accident	50	40
survived accident	180	200

- Compute Chi-squared statistic to test the hypothesis that motorcycle fatalities and helmet laws are independent.
- If they are independent, then helmet laws to not effect motorcycle fatalities. If they are dependent, then helmet laws do effect motorcycle fatalities.

5 Homework

Homework

Homework

- Section 11-3: problems 7 through 12, odds.