2. Chi^2-tests for equality of proportions

Introduction: Two Samples

Consider comparing the sample proportions p_1 and p_2 in independent random samples of size n_1 and n_2 out of two populations which have a certain characteristic with respective probabilities π_1 and π_2.

We know from STAT.1030 that the relevant test statistic for equality of proportions is:

$$z = \frac{p_1 - p_2}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim N(0,1) \text{ under } H_0\colon \pi_1 = \pi_2,$$

where $p = \dfrac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$ denotes the combined proportion of both samples. Recall that this z-statistic is based upon the normal approximation of the binomial distribution and requires therefore sufficiently large sample sizes ($n \gtrsim 30$).

The same test may be equivalently formulated as a χ^2 independence test as follows.

Consider the contingency table below:

	success	failure	sum
pop.1	$n_1 p_1$	$n_1(1-p_1)$	$n_1 = f_{1\bullet}$
pop.2	$n_2 p_2$	$n_2(1-p_2)$	$n_2 = f_{2\bullet}$
sum:	$np = f_{\bullet 1}$	$n(1-p) = f_{\bullet 2}$	n

Since p_1 and p_2 are unbiased estimators of their population counterparts π_1 and π_2, H_0: $\pi_1 = \pi_2$ is equivalent to independence of the success and population variables. Applying the formula for two-way tables,

$$\chi^2 = \frac{n(f_{11}f_{22} - f_{12}f_{21})^2}{f_{1\bullet}f_{2\bullet}f_{\bullet 1}f_{\bullet 2}},$$

yields after some manipulation:

$$\chi^2 = \frac{(p_1 - p_2)^2}{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \sim \chi^2(1) \text{ under } H_0.$$

It should not surprise us that this is just the square of our familiar z-statistic, since we know from STAT.1030 that the square of a $N(0,1)$-distributed random variable is $\chi^2(1)$-distributed.

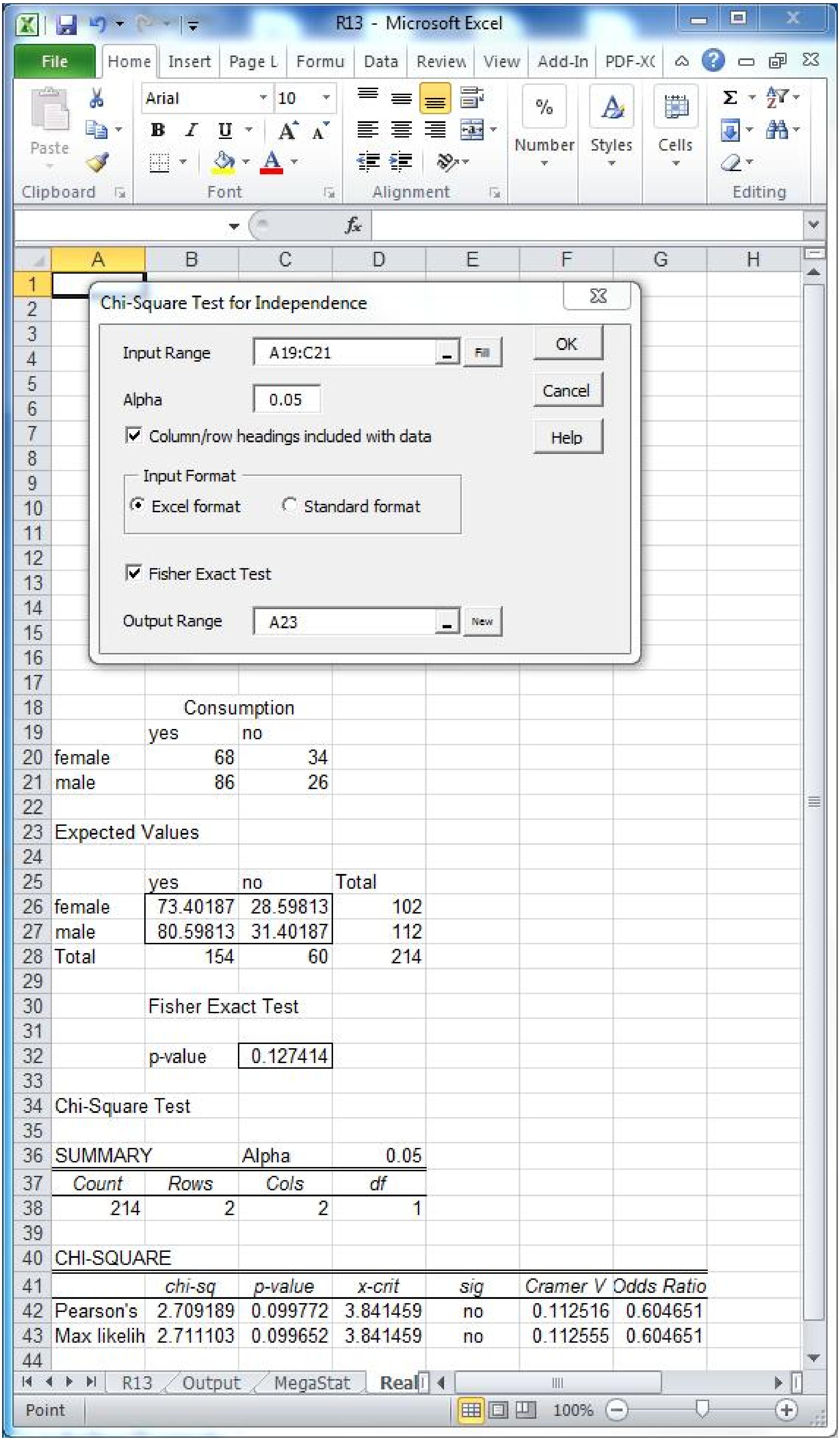
Example. Consider the following table about consumption of alcohol for men and women:

Consumption:	yes	no
female (pop.1)	68	34
male (pop.2)	86	26

As is evident from the output on the next slide, we may not reject $H_0: \pi_1 = \pi_2$ against the two-sided alternative $H_1: \pi_1 \neq \pi_2$ ($p = 0.0998$), but we may reject H_0 against the one-sided alternative $H_1: \pi_2 > \pi_1$ ($p = 0.0499$).

Using Fishers exact test, the p-value of $H_0: \pi_1 = \pi_2$ against the two-sided alternative $H_1: \pi_1 \neq \pi_2$ is $p = 0.1274$, leading to the same conclusion as above. But against the one-sided alternative $H_1: \pi_2 > \pi_1$ it is $p = 0.0676$, which means that we cannot even reject H_0 in a one-sided test at $\alpha = 5\%$.

If we have access to software, we prefer using the output from Fishers exact test. It is however too hard to calculate by hand, and in larger tables even software might fail to calculate it.


R13 - Microsoft Excel
File
Home
Insert
Page L
Formu
Data
Review
View
Add-In
PDF-X(
Paste
Clipboard
Arial
10
Font
Alignment
Number
Styles
Cells
Editing
Chi-Square Test for Independence
Input Range
A19:C21
Fill
Alpha
0.05
Column/row headings included with data
Input Format
Excel format
Standard format
Fisher Exact Test
Output Range
A23
New
OK
Cancel
Help
Consumption
yes
no
female
68
34
male
86
26
Expected Values
yes
no
Total
female
73.40187
28.59813
102
male
80.59813
31.40187
112
Total
154
60
214
Fisher Exact Test
p-value
0.127414
Chi-Square Test
SUMMARY
Alpha
0.05
Count
Rows
Cols
df
214
2
2
1
CHI-SQUARE
chi-sq
p-value
x-crit
sig
Cramer V
Odds Ratio
Pearson's
2.709189
0.099772
3.841459
no
0.112516
0.604651
Max likelih
2.711103
0.099652
3.841459
no
0.112555
0.604651
R13
Output
MegaStat
Real
Point
100%

General test for homogeneity of proportions

The χ^2-test for equality of proportions may be generalized to more than two independent samples as follows. Assemble the frequencies of success np_i and failure $n_i(1-p_i)$ for the respective populations i in a ($2\times c$) contingeny table, where n_i and p_i denote the size and proportion of success in population i, and c denotes the number of populations.

The homogeneity of the populations
(with respect to the success variable)

H_0: $\pi_1 = \pi_2 = \cdots = \pi_c$ versus
H_1: Not all $\pi_i,\ i = 1, \ldots, c$ are equal

is then assessed by a χ^2 independence test of the population and success variables.

Example.

An insurance company wants to test whether the proportion of people who submit claims for automobile accidents is about the same for the three age groups 25 and under, over 25 and under 50, and 50 and over:

H_0: The age goups are homogeneous wrt. claims,
H_1: The age goups are not homogeneous wrt. claims.

The (2×3) contingency table for this example is:

	age< 25	25 ≤age< 50	50 ≤age	Total
Claim	40	35	60	135
No claim	60	65	40	165
Total	100	100	100	300

The expected frequencies are:

	age< 25	25 ≤age< 50	50 ≤age	Total
Claim	45	45	45	135
No claim	55	55	55	165
Total	100	100	100	300

The χ^2 statistic is:

$$\chi^2 = \frac{(40-45)^2}{45} + \frac{(35-45)^2}{45} + \frac{(60-45)^2}{45}$$
$$\frac{(60-55)^2}{55} + \frac{(66-55)^2}{55} + \frac{(40-55)^2}{55} = 14.14$$

and the degrees of freedom are (2-1)(3-1)=2. Looking up in a table or calling CHIDIST(14.14;2) in Excel establishes that the p-value in this case is less than 0.1%, so we reject homogeneity with respect to claims in favour of inhomogeneity at all conventional significance levels.

The median test

The hypotheses for the median test are

H_0: The c populations have the same median,
H_1: Not all c populations have the same median.

This is in fact a special case of the homogeneity test for independent samples which we just discussed, with the indicator: value $\gtrless$ median as the success variable. Median means here grand median, which is the median of all observations regardless of the population.

Example:

An economist wants to test the null hypothesis that the median family income in three rural areas are approximately equal. Random samples of family incomes (in $1000/year) in three regions (A/B/C) are given below:

A	22	29	36	40	35	50	38	25	62	16
B	31	37	26	25	20	43	27	41	57	32
C	28	42	21	47	18	23	51	16	30	48

Example: (continued.)
Arranging all observations in order reveals that the grand median is 31.5. The observed and expected counts for each region are displayed below:

	Region A	Region B	Region C	Total
≤median	4	5	6	15
(expected)	(5)	(5)	(5)	
>median	6	5	4	15
(expected)	(5)	(5)	(5)	
Total	10	10	10	30

Note that here: $e_{ij} = \frac{f_{i\bullet} f_{\bullet j}}{n} = \frac{n/2 \cdot n_i}{n} = \frac{n_i}{2}.$

The χ^2 statistic is:

$$\chi^2 = \frac{(4-5)^2}{5} + \frac{(5-5)^2}{5} + \frac{(6-5)^2}{5}$$
$$\frac{(6-5)^2}{5} + \frac{(5-5)^2}{5} + \frac{(4-5)^2}{5} = 0.8$$

and the degrees of freedom are (2-1)(3-1)=2. Comparing this value with critical points $\chi^2_\alpha(2)$ of the χ^2-distribution with 2 degrees of freedom, we conclude that there is no evidence to reject the null hypothesis. The p-value of the test is CHIDIST(0.8;2)=0.67.

The median test in Excel

You find the median test under the name Moods Test in the Nonparametric Tests tool of the Real Statistics toolbox.