

## Hypothesis testing: comparing proportions

- Prof. Giuseppe Verlato
- Unit of Epidemiology & Medical Statistics,  
Dept. of Diagnostics & Public Health,  
University of Verona

### Mendel's experiment:

Mendel bred together smooth yellow peas (dominant traits) and wrinkled green peas (recessive traits), and further inbred the 1<sup>o</sup> generation of hybrids.

	Yellow	green	
Smooth	315	108	423
Wrinkled	101	32	133
	416	140	556

How many peas are **EXPECTED** in each cell under the hypothesis of statistical independence?

$$\begin{aligned}\text{Expected in the 1st cell} &= p(\text{smooth} \cap \text{yellow}) * N = \\ &= p(\text{smooth}) * p(\text{yellow}) * N = (423/556) * (416/556) * 556 = \\ &= 423 * 416 / 556 = 316.5\end{aligned}$$

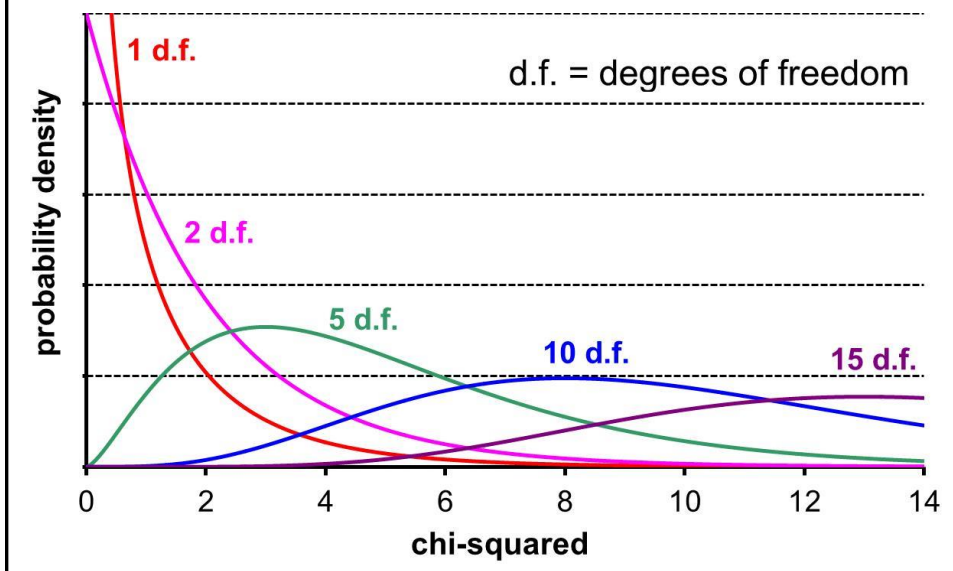
$$\text{Number of Expected} = \frac{(\text{row units}) * (\text{column units})}{\text{overall units}}$$

<b>OBSERVED</b>	yellow	green	
smooth	315	108	423
<b>wrinkled</b>	101	32	<b>133</b>
	416	140	556

<b>EXPECTED</b>	yellow	green	
smooth	316.5	106.5	423
<b>wrinkled</b>	99.5	33.5	<b>133</b>
	416	140	556

**At first sight** the hypothesis of statistical independence between surface characteristic (smooth / wrinkled) and color (yellow/green) seems to be valid: genetic traits are independently assorted.

The chi-square ( $\chi^2$ ) test, based on the chi-square distribution, allows to verify the hypothesis of statistical independence **IN A SCIENTIFIC WAY**.



## Chi-squared test for independence

$$\chi^2 = \sum (\text{observed} - \text{expected})^2 / \text{expected}$$

$$\begin{cases} H_0: \text{the two variables are statistically independent} \\ H_1: \text{the two variables are statistically dependent} \end{cases}$$

Significance level = 5%

$$\begin{aligned} \text{Degrees of freedom} &= (n^\circ \text{ rows} - 1) * (n^\circ \text{ columns} - 1) = \\ &= (2-1)*(2-1) = 1*1 = 1 \end{aligned}$$

$$\text{Critical threshold} = \chi^2_{1, 0.05} = 3.84$$

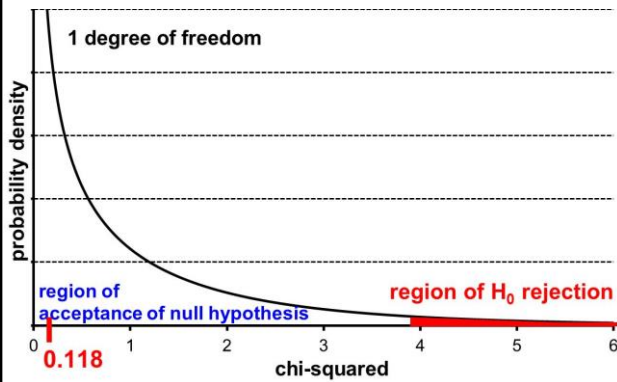
$$\begin{aligned} \chi^2 &= \frac{(315-316.5)^2}{316.5} + \frac{(108-106.5)^2}{106.5} + \frac{(101-99.5)^2}{99.5} + \frac{(32-33.5)^2}{33.5} \\ &= 0.007 + 0.021 + 0.023 + 0.067 = 0.118 \end{aligned}$$

observed  $\chi^2 <$  critical threshold  
 0.118                      3.84



**H<sub>0</sub> is accepted.**

Traits “surface characteristic” and “color” are independently assorted (**3rd Mendel’s law**).



N.B. chi-squared test is always one-sided.



**Computing DEGREES OF FREEDOM of a CHI-SQUARE test**

?	?	423				416-111
?	111	133	133-111	22	111	133
140	416	556		140	416	556

			140-22
423-305	118	305	423
	22	111	133
	140	416	556

**In a 2\*2 table  
 DEGREES OF FREEDOM = 1**

**Computing DEGREES OF FREEDOM  
of a CHI-SQUARE test**

?	?	?	463
?	72	56	193
100	140	416	656

	$100-65$	$140-72$	$416-56$	
$463-68-360$	35	68	360	463
$193-72-56$	65	72	56	193
	100	140	416	656

**In a 2\*3 table  
DEGREES OF FREEDOM = 2**

**Computing DEGREES OF  
FREEDOM of a CHI-SQUARE test**

?	?	423
?	111	133
140	416	556

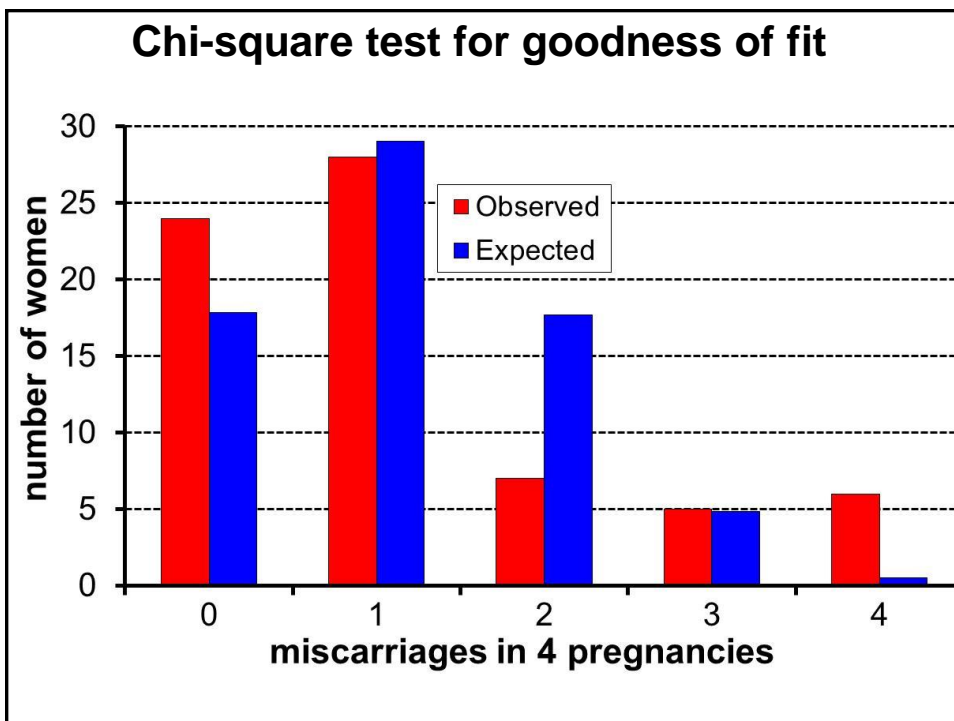
**degrees of freedom = 1**

?	?	?	463
?	72	56	193
100	140	416	656

**degrees of freedom = 2**

**Is there unifying formula ? YES  
DEGREES OF FREEDOM =  
(n ROWS - 1) \* (n COLUMNS -1)**

- The chi-square test is based on the normal approximation.
- The chi-square test is suited when there are at least 5 expected per cell in a 2\*2 table.
- Otherwise, the Fisher's exact test must be used instead.



### Chi-square test for goodness-of-fit

$$\chi^2 = \sum (\text{observed} - \text{expected})^2 / \text{expected}$$


- $\{ H_0: \text{observed data follow a binomial distribution}$
- $\{ H_1: \text{observed data DO NOT follow a binomial distribution}$

Level of significance = 5%

Degrees of freedom =  $n^\circ \text{ cells} - n^\circ \text{ parameters} = 5 - 2 = 3$

Critical threshold =  $\chi^2_{3, 0.05} = 7.81$

$$\begin{aligned} \chi^2 &= \frac{(24-17.9)^2}{17.9} + \frac{(28-29.1)^2}{29.1} + \frac{(7-17.7)^2}{17.7} + \frac{(5-4.8)^2}{4.8} + \frac{(6-0.5)^2}{0.5} \\ &= 2.12 + 0.04 + 6.48 + 0.01 + 61.96 = 70.60 \end{aligned}$$

Observed  $\chi^2 >$  critical threshold   $H_0$  is rejected

70.60                      7.81