Chap. IV – χ^2 TEST (« chi square »)

This test allows to estimate whether or not the distribution function of a random variable follows a known distribution (e.g. binomial, normal, … distribution).

The observed distribution is based on sample (of the unknown population). We divide the sample values into defined classes, and determine the observed frequencies n_{obs} for each class.

The test invokes the comparison of the observed frequencies n_{obs} with the expected frequencies n_{th}, which are calculated using the theoretical distribution that we intend to test.

The variable χ^2 is defined as follows:

This variable χ^2 is a random variable, its numerical value depends on the selected sample.
The variable χ^2 is described by a Pearson distribution ("χ^2 distribition") with ν degrees of freedom:

For a given number of degrees of freedom, the Pearson distribution ("χ^2 distribition") gives the probability to obtain a χ^2 value smaller than the value calculated based on the sample.

We reject the proposed hypothesis if the calculated χ^2 value is larger than the critical value obtained from the inverse χ^2 distribution.

Generally speaking, employing the χ^2 test requires that all classes have a non-zero frequency and that at least 80% of the classes have a frequency of at least 5 (Cochran criterium, 1954).

<u>Example :</u> We would like to test whether the probability of having a boy is 1/2.

A survey among 320 families of 5 children has given the following results:

Number of boys	0	1	2	3	4	5
Obs. frequencies	8	40	88	110	56	18

Hypothesis to test:

Calculation of the expected frequencies (based on the hypothesis):

Number of boys	0	1	2	3	4	5
Exp. frequencies						

Question: Hypothesis confirmed by expected frequencies ?

Number of boys	0	1	2	3	4	5
Diff. $n_{th} - n_{obs}$						

Calculation of χ^2 :

Is the obtained χ^2 value compatible with the hypothesis?