Testing

Menu:	QCExpert	Testing

The group *Testing* consists of three modules: *Power and Sample Size*, *Tests* and *Contingency Table*. These modules are described below.

Power and Sample size

Menu:	QCExpert	Testing	Power and Sample Size

Modules in the group *Power and Sample Size* compute power of a test, required sample size an minimal difference of parameters that can be detected by the test. The tests support normal and binomial distribution. Inputs are significance level (or type I error) α, type of the test (one-sided or two-sided and theoretical (expected, specified) distribution parameter value. This parameter is the mean value for normal distribution, or probability for binomial distribution. Further, it is necessary to specify two of the following three numbers: Sample size, expected sample statistic and the power of the test 1 – β (β is the type II error). This module does not use any data from the data editor.

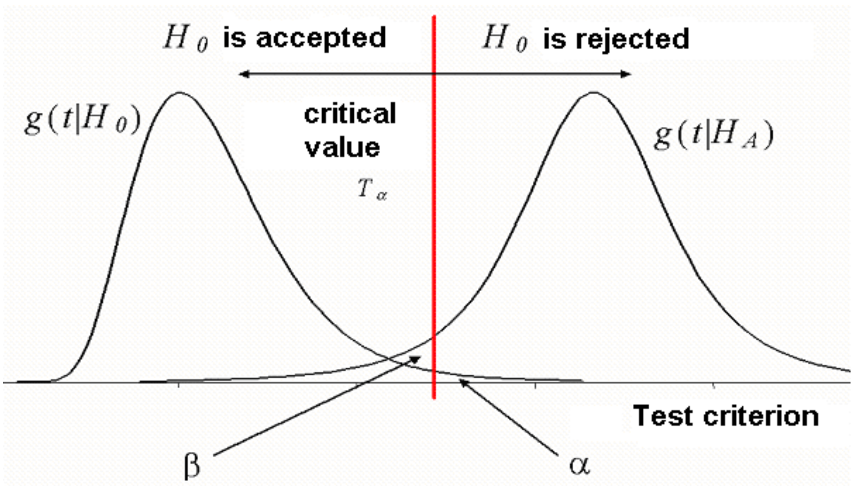

Fig. 1 Probabilities of type I (α) and type II (β) errors

Fig. 1 illustrates the types 1 and II type errors. H_0 denotes simple hypothesis to be tested, or zero hypothesis, for example equality of mean and some given number. H_A denotes alternative hypothesis, not(H_0). The result of the test is rejection or acceptance of H_0 based on comparing test criterion T calculated from sample statistics with critical value for this test T_α.

Example: For testing equality of arithmetic average x and a given value μ we use the Student t-test, where the test criterion is $T = |x - \mu|/s$ and the critical value $T_\alpha = t_{1-\alpha/2}(N-1)$ is 100(1–α)% quantile of the Student distribution. The two curves on the above figure illustrate the densities of distribution of the test criterion value for the cases when H_0 holds ($g(t|H_0)$) and when H_0 does not hold ($g(t|H_A)$). Two types of a mistake may happen:

The type I error, when we mistakenly reject H_0, despite the fact that H_0 is true. This will happen, when we happen to select the data from population (or measure items from a box) that all have untypically high or low value compared to other data. This will lead to too high value of T, which consequently, compared with T_α yields refusing H_0. This situation is called the type I error and is illustrated on Fig. 2. Its probability is α and can be specified by user. Usually, we set α = 0.05, or 5%.

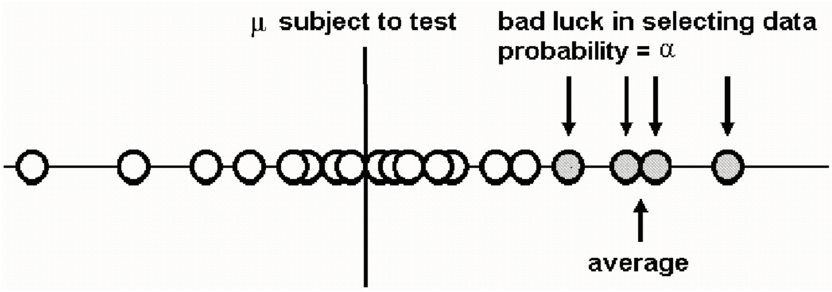

Fig. 2 Type I error, H_0 is rejected based on 4 unlucky measurements, though in fact H_0 holds.

Similarly, we can think of type II error, when we accept H_0, though it does not hold, see Fig. 3. Probability (or risk) of this situation is β. Obviously, number of data N, α, β, and difference between real and estimated parameter Δx are interdependent. When we want for example to have low both α and β, we have to take more data. When there is big Δx, we need less data. When we have available only small data set and expect small Δx, we will obtain lower „reliability“ of the test in term of high α and β, etc. All methods of Power and Sample Size have both one-sided and two-sided option. One-sided option means, that we are testing only „bigger“ or only „less“, and we don’t take into account the other possibility. By two-sided test we do not distinguish between „bigger“ or „less“. One-sided option tests always $x > \mu$ in one-sample normal tests, or $x_2 > x_1$ in two-sample normal and $P_A > P_0$ in one-sample binomial proportion tests or $P_2 > P_1$ in two-sample binomial proportion tests.

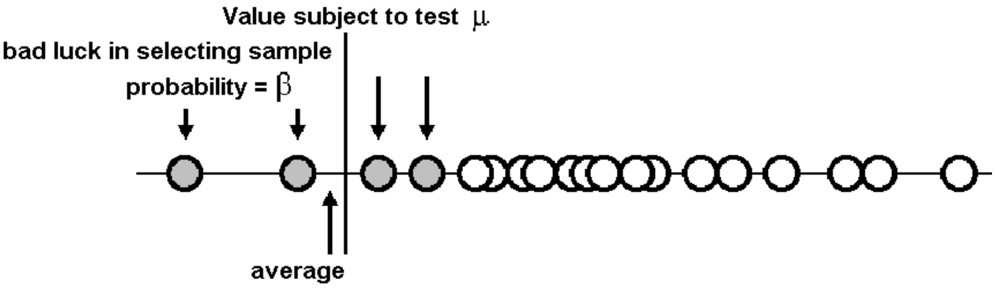

Fig. 3 The type II error, H_0 is falsely accepted based of 4 measurements, although the true mean is not equal to the value subject to test.

The module Power and Sample Size can answer three types of questions:

- What would be the least sample size to prove the given difference between a hypothesized statistic (sample average or proportion) and a given number (or between two statistics) at a given risks α, β;
- What is the least difference difference between a hypothesized statistic (sample average or proportion) and a given number (or between two statistics) that could be proved by the test at a given sample size (or sizes) and at a given risks α, β;
- What would be the power 1 – β of a test that will prove a given difference between hypothesized statistic (sample average or proportion) and a given number (or between two statistics) at a given sample size (or sizes) and at a given risk of the type I error α.

Normal distribution, 1 sample

Menu:	QCExpert	Testing	Power and Sample Size	Normal distribution 1 sample

This module calculates parameters for testing arithmetic average of one normally distributed sample.

Parameters

In the dialog window (Fig. 4) specify the significance level (here also called type I error probability), the given number to be compared with average and the expected standard deviation σ.

Select one- or two-sided option. Then you must specify two of the three fields: *Sample size*, *Sample average*, *Power*. At the field you want to calculate, click the corresponding button. The last calculated value will be marked in red. After calculation, the dialog window will not close, nor there is any output to the protocol. You may specify next parameters and make another calculations. Clicking *Output to protocol* will produce the output to the protocol window, the dialog window *Power and Sample Size* is closed by clicking on *Close*. The *Close* button itself does not produce an output to protocol.

The power of the test may be calculated from N, μ, X, σ, α according to

$$1-\beta=\Phi\left(\frac{\sqrt{N}\,(\mu-X)}{\sigma}-Z_{1-\alpha/2}\right)+\Phi\left(\frac{\sqrt{N}\,(X-\mu)}{\sigma}-Z_{1-\alpha/2}\right)$$

The minimal sample size is given by

$$N=\left\{\left(\sigma\left(Z_{1-\alpha/2}+Z_{1-\beta}\right)\right)\Big/\left|\mu-X\right|\right\}^2,$$

where Z_α is the α-quantile of normal distribution and Φ is the distribution function (or cumulative density) of normal distribution. The unknown X (sample average) is calculated iteratively.

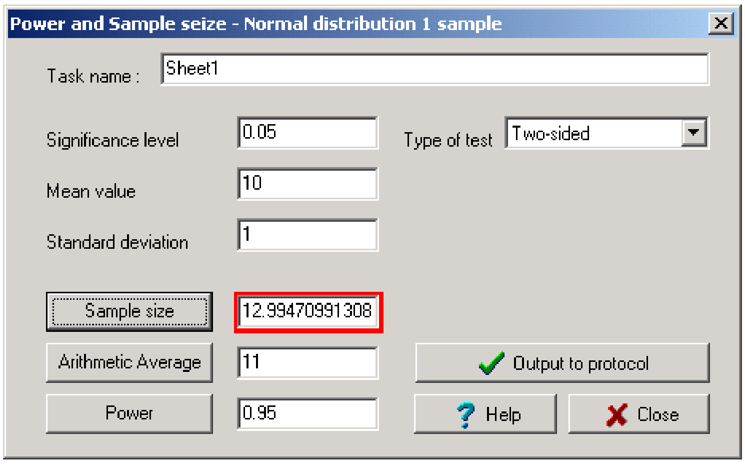

Fig. 4 Dialog window for Power and Sample size, normal distribution, 1 sample

Protocol

Power and sample size Normal dist., 1 sample	
Computation of sample size	Specifies which of the parameters was calculated
Task name	Task name from the dialog window
Significance level	Significance level α
Mean value M	Specified constant μ
Expected mean X	Specified or calculated arithmetic average.
Type of test	Specified mode: one-sided of two-sided
Null hypothesis H_0	X = M
Alternative hypothesis H_A	X <> M in case of two-sided test, X > M in case of one-sided test. It is always assumed that $x > \mu$.
Standard deviation	Specified assumed standard deviation of data
Sample size	Specified or calculated sample size, non-integer value must be always rounded to the higher integer value.

Rounded sample size	Calculated sample size rounded to the nearest higher integer.
Power of test	Specified or calculated power of the test.

Normal distribution, 2 samples

Menu:	QCExpert	Testing	Power and Sample Size	Normal distribution 2 samples

This module calculates parameters for testing arithmetic averages of two normally distributed samples.

Parameters

In the dialog window (Fig. 5) specify the significance level (here also called type I error probability), the average of the first sample and the expected standard deviations of the first and second sample. If the Sample size is to be computed, you can also specify the ratio of the sample sizes of the first and second sample N_2/N_1. Select one- or two-sided option. Then you must specify two of the three fields: *Sample sizes* (two fields), *Second sample average*, *Power*. At the field you want to calculate, click the corresponding button. The last calculated value will be marked in red. After calculation, the dialog window will not close, nor there is any output to the protocol. You may specify next parameters and make another calculations. Clicking *Output to protocol* will produce the output to the protocol window, the dialog window *Power and Sample Size* is closed by clicking on *Close*. The *Close* button itself does not produce an output to protocol.

The minimal sample sizes N_1 and N_2 are given by

$$N_1 = \left(\sigma_1^2 + \frac{\sigma_2^2}{k}\right)\left\{\frac{\left(Z_{1-\alpha/2} + Z_{1-\beta}\right)}{\left|X_2 - X_1\right|}\right\}^2,$$

$$N_2 = kN_1$$

where k is ratio N_2/N_1, Z_α is α-quantile of normal distribution.

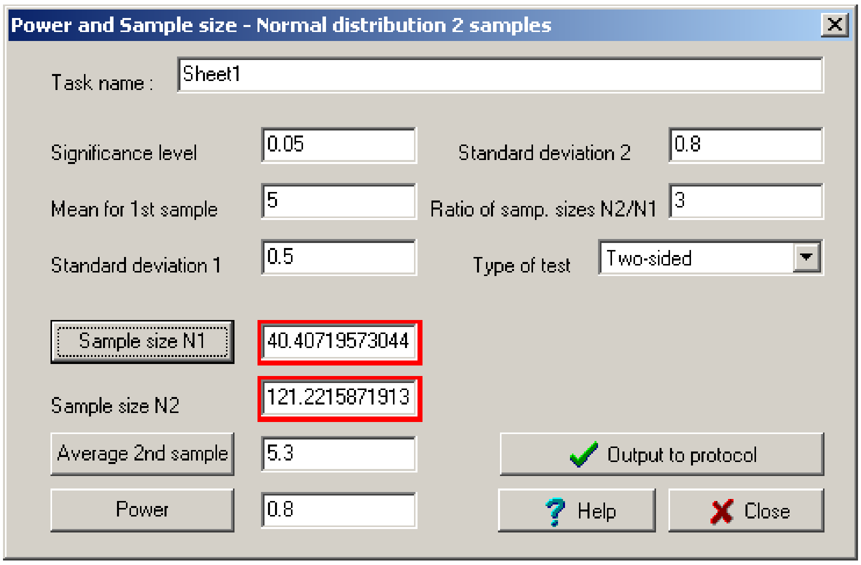

Fig. 5 Dialog window Power and Sample size – Normal distribution 2 samples

Note

Though the total number of measurements $N_1 + N_2$ is minimal when $N_1 = N_2$, or $N_2/N_1 = 1$, it may be profitable to force one of the two samples to be smaller, say N_1, by specifying $k > 1$, for example when the first sample is much more expensive or difficult to measure, even at a price of significant increase of N_2.

Protocol

Power and sample size Normal dist., 2 samples	
Computation of sample size	Specifies which of the parameters was calculated
Task name	Task name from the dialog window
Significance level	significance level α
Mean for 1st sample	Specified arithmetic average of the 1st sample
Mean for 2nd sample	Specified or calculated arithmetic average of the 2nd sample
Type of test	Specified mode: one-sided of two-sided
Null hypothesis H_0	$X_1 = X_2$
Alternative hypothesis H_A	$X_1 \neq X_2$ in case of two-sided test, $X_1 < X_2$ in case of one-sided test. It is always assumed that $X_1 < X_2$.
Standard deviation 1	Specified assumed standard deviation of the first sample
Standard deviation 2	Specified assumed standard deviation of the second sample
Ratio of samp. sizes N2/N1	Specified ratio of sizes of the second and first sample. This ratio is used only when calculating *Sample size*.
Sample size N1 Sample size N2	Specified or calculated sample sizes, non-integer values must be always rounded to the higher integer value.
Rounded sample sizes	Calculated sample sizes rounded to the nearest higher integer.
Power of test	Specified or calculated power of the test.

Binomial distribution, 1 sample

Menu:	QCExpert	Testing	Power and Sample Size	Binomial distribution 1 sample

This module calculates parameters for testing equality of probability of occurrence and a given constant (here called *Tested ratio*) based on the hypothesized ratio of observed trials and occurrences (here called *Ratio*).

Parameters

In the dialog window (Fig. 6) specify the significance level (here also called probability of the type I error), the given number P_0 to be compared with hypothesized ratio and the expected ratio of successes and trials P_1. Select one- or two-sided option. Then you must specify two of the three fields: *Sample size N*, *Ratio*, *Power*. Then, at the field you want to calculate, click the coresponding button. The last calculated value will be marked in red. After calculation, the dialog window will not close, nor there is any output to the protocol. You may specify next parameters and make another calculations. Clicking *Output to protocol* will produce the output to the protocol window, the dialog window *Power and Sample Size* is closed by clicking on *Close*. The *Close* button itself does not produce an output to protocol.

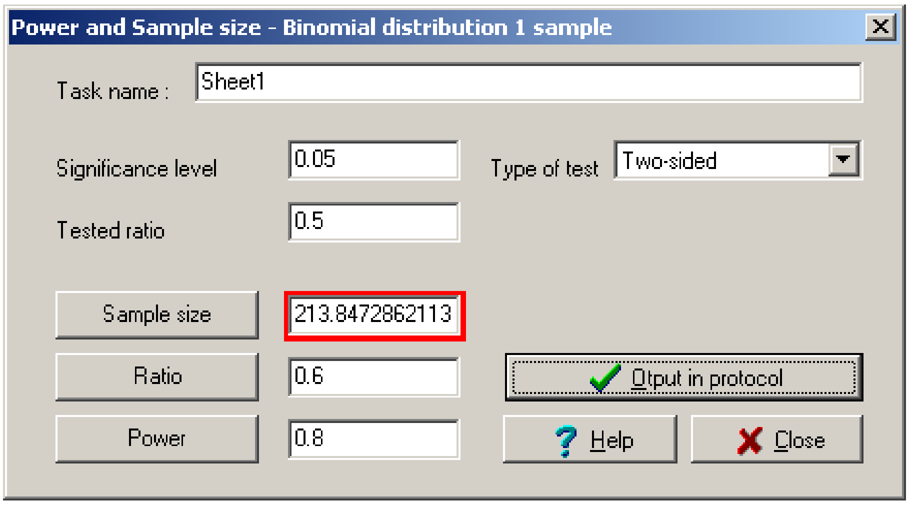

Fig. 6 Dialog window Power an sample size – Binomial distribution 1 sample

The sample size N is given by

$$N = \left\{ \frac{\sqrt{P_0(1-P_0)}Z_{1-\alpha/2} + \sqrt{P_0(1-P_0)}Z_{1-\beta}}{|P-P_0|} \right\}^2 + \frac{2}{|P-P_0|}$$

where Z_α is α-quantile of the normal distribution. Normal approximation is used, which is precise enough for $NP(1-P) > 5$. The second term is the correction for continuous approximation.

Protocol

Power and sample size Binomial dist., 1 sample	
Computation of sample size	Specifies which of the parameters was calculated
Task name	Task name from the dialog window
Significance level	Significance level α
Ratio to be tested, P_0	Specified value to be compared with *Ratio* P_A., $0<P_0<1$.
Expected ratio P_A	Specified constant value $0<P_A<1$.
Type of test	Specified mode: one-sided of two-sided
Null hypothesis H_0	$P_0 = P_A$
Alternative hypothesis H_A	$P_0 \neq P_A$ in case of two-sided test, $P_0 < P_A$ in case of one-sided test. It is always assumed that $P_0 < P_A$.
Sample size	Specified or calculated sample size, non-integer value must be always rounded to the higher integer value.
Rounded sample size	Calculated sample size rounded to the nearest higher integer.
Power of test	Specified or calculated power of the test.

Binomial distribution, 2 samples

Menu:	QCExpert	Testing	Power and Sample Size	Binomial distribution 2 samples

This module calculates parameters for testing equality of probability of occurrence in two experiments (here called *Tested ratio*) based on the hypothesized ratio of observed trials and occurrences for both experiments (here called *Ratio*).

Parameters

In the dialog window (Fig. 7) specify the significance level (here also called probability of the type I error), hypothesized ratios of successes and trials in the first and second experiment $P_1 = X_1/N_1$, $P_2 = X_2/N_2$. Select one- or two-sided option. Then you must specify two of the three fields: *Sample size N*, *Ratio X2/N2*, *Power*. Then, at the field you want to calculate, click the corresponding button. The last calculated value will be marked in red. After calculation, the dialog window will not close, nor there is any output to the protocol. You may specify next parameters and make another calculations. Clicking *Output to protocol* will produce the output to the protocol window, the dialog window *Power and Sample Size* is closed by clicking on *Close*. The *Close* button itself does not produce an output to protocol.

Sample sizes N_1 and N_2 are given by

$$N_1 = \left\{ \frac{\sqrt{P_1(1-P_1)+\frac{P_2(1-P_2)}{k}}\,Z_{1-\beta} + \sqrt{\bar{P}(1-\bar{P})+1\frac{1}{k}}\,Z_{1-\alpha/2}}{|P_2-P_1|} \right\}^2 + \frac{k+1}{k|P_2-P_1|},$$

where Z_α is α-quantile of the normal distribution. Normal approximation is used, which is precise enough for $NP(1-P) > 5$. The second term is the correction for continuous approximation.

$\bar{P} = \frac{P_1 + kP_2}{1+k}$; k is the user-specified specified ratio, $k = N_2/N_1$, so that $N_2 = k\,N_1$.

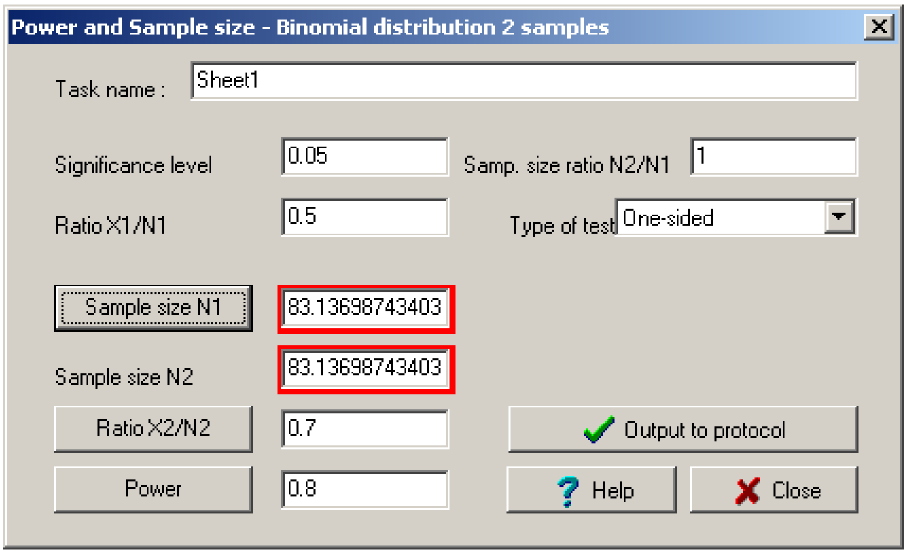

Fig. 7 Dialog window Power and Sample size – Binomial distribution 2 samples

Protocol

Power and sample size Binomial dist., 1 sample	
Computation of sample size	Specifies which of the parameters was calculated
Task name	Task name from the dialog window
Significance level	Significance level α
Expected ratio P_1	Specified value of the ratio of successes in the first experiment.

Expected ratio P_2	Specified value of the ratio of successes in the second experiment.
Type of test	Specified mode: one-sided of two-sided
Null hypothesis H_0	$P_1 = P_2$
Alternative hypothesis H_A	$P_1 \neq P_2$ in case of two-sided test, $P_1 < P_2$ P_A in case of one-sided test. It is always assumed that $P_1 < P_2$.
Ratio of sample sizes N2/N1	Specified ratio of sizes of the second and first sample. This ratio is used only when calculating *Sample size*.
Sample size N1 Sample size N2	Specified or calculated sample sizes, non-integer values must be always rounded to the higher integer value.
Rounded sample sizes	Calculated sample sizes rounded to the nearest higher integer.
Power of test	Specified or calculated power of the test.

Tests

The group *Tests* performs statistical testing for one-sample and two-sample binomial and normal distribution, for multinomial distribution and for contingency tables. Testing is based on actual experimental data, or on known statistics like average or standard deviation.

Binomial test, 1 sample

Menu:	QCExpert	Testing	Tests	Binomial test 1 sample

This module tests the hypothesis H_0, whether the observed number of occurrences X in N tested trials is in accordance with a given constant probability P of occurrence of X in one (any) trial. Standard Chi-square test is employed. Assuming that the true (usually unknown) probability P_A of occurrence is equal to the given P (the null hypothesis H_0), it would hold $P = X/N$ for $N \to \infty$. It is good to keep in mind, that not rejecting H_0 does not necessarily mean that H_0 is true. Often it only means that the number of trials is not sufficient to reject H_0.

Parameters

This module does not use any data from the data spreadsheet. All information needed for the calculation are specified in the dialog window, see Fig. 8. You may modify the task name and the significance level (default value is $\alpha = 0.05$). Than you must specify the given probability P to be tested, number of trials N and number of occurrences X. After clicking the RunTest button the test is performed and results are displayed in the fields *Chi2 statistic*, *Critical value*, *p-value*. If the statistic is bigger than the critical value, the possible equality between hypothesized P and true unknown probability of success estimated by X/N is rejected. The field *Conclusion* will show verbal conclusion of the test: „*Equality of ratios is REJECTED or ACCEPTED*“. For very low number of occurrences and/or low P, $XP < 5$, the test is less reliable.

Clicking *Output to protocol* will produce a record of the last performed test in the Protocol window, while the dialog window still remains open. Clicking *Close* will close the dialog window. χ^2 or Chi-square test is used in this module. The χ^2 test statistic Z

$$Z = \frac{\left(X - NP\right)^2}{NP\left(1 - P\right)}$$

has asymptotically distribution $\chi^2_{(1)}$. This statistic is compared to the quantile $U = \chi^2_{(1)}(1-\alpha)$. If $Z > U$ then H_0 is rejected.

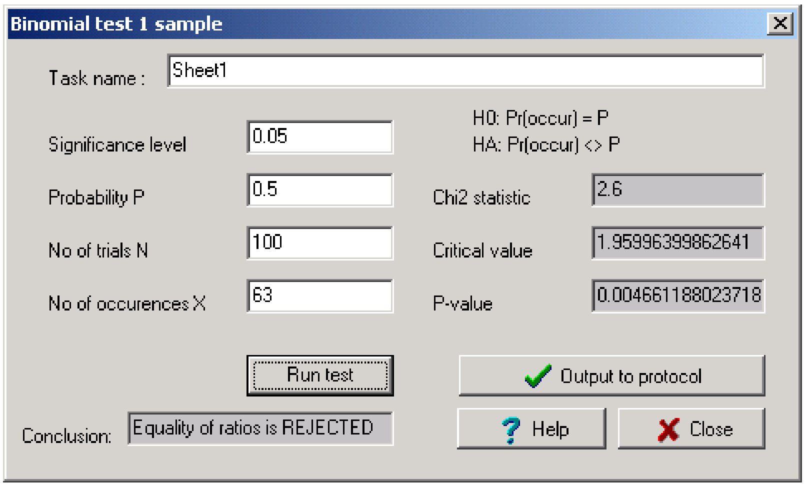

Fig. 8 Dialog window for Binomial test, 1 sample

Protocol

Binomial test for equal ratio, 1 sample	
Task name	Task name from the dialog window
Overall sample size	Number of trials
Number of occurrences	Number of occurrences
Sample probability X/N	Calculated ratio X/N.
Probability to be tested	Given probability P to be tested
Significance level	Significance level α
Statistic Z	Calculated χ^2 statistic
Critical value U	Quantile of the χ^2 distribution
p-value	Biggest significance value at which H_0 would be rejected
Conclusion	Verbal test conclusion

Binomial test, N samples

Menu:	QCExpert	Testing	Tests	Binomial test N samples

This module generalizes the previous test. It tests simultaneously K hypotheses based on K binomial experiments, if observed numbers of occurrences X_i in N_i trials correspond to the hypothesized probabilities of this occurrences P_i. Null hypothesis H_0 is defined as H_0: $P_i = P_{Ai}$ for $i = 1, .. K$. Chi-square test is employed again. Assuming that all true (unknown)probabilities P_{Ai} of occurrence of A are equal to P_i, than for $N_i \to \infty$ the probabilities P_i would be equal to X_i/N_i.

Data and parameters

This module expects data in two or three columns in the data editor. One column must contain numbers of trials for each experiment, second column must contain the numbers of successes, or occurrences of A_i, the third column may contain the probabilities P_i. The third column is not required, you may set all P_i ($i = 1 .. K$) to empirical average value $P_i = \Sigma X_i/\Sigma N_i$. An example of input data is in the following table for $K = 4$.

Trials	*Success*	*Probability*
200	22	0.1
200	46	0.25
100	56	0.5
250	103	0.4

In the dialog window, see Fig. 9, you may modify the task name and the significance level (default value is $\alpha = 0.05$). Then select the columns with N_i, X_i and P_i respectively. If the field *Use empirical probability* is checked, the program will use equal values $P_i = \Sigma X_i/\Sigma N_i$ and will ignore the third column, if any.

After clicking the RunTest button the test is performed and results are displayed in the fields *Chi2 statistic*, *Critical value*, *p-value*. If the statistic is bigger than the critical value, the possible equality between all hypothesized P_i and true probabilities of success estimated by X_i/N_i is rejected. The field *Conclusion* will show verbal conclusion of the test: „*Equality of ratios is REJECTED or ACCEPTED*“. For very low number of occurrences and/or low *P*, $X_i P_i < 5$, the test is less reliable. Clicking *Output to protocol* will produce a record of the last performed test in the Protocol window, while the dialog window still remains open. Clicking *Close* will close the dialog window.

Standard χ^2 test is used based on statistic *C*, which has asymptotic distribution $\chi^2_{(K\text{-}1)}$.

$$C = \sum_{i=1}^{K} \frac{1}{P_i\left(1-P_i\right)}\left(X_i - N_i P_i\right)$$

This statistic is compared to the quantile $U = \chi^2_{(1)}\,(1-\alpha)$. If $C > U$ then H_0 is rejected.

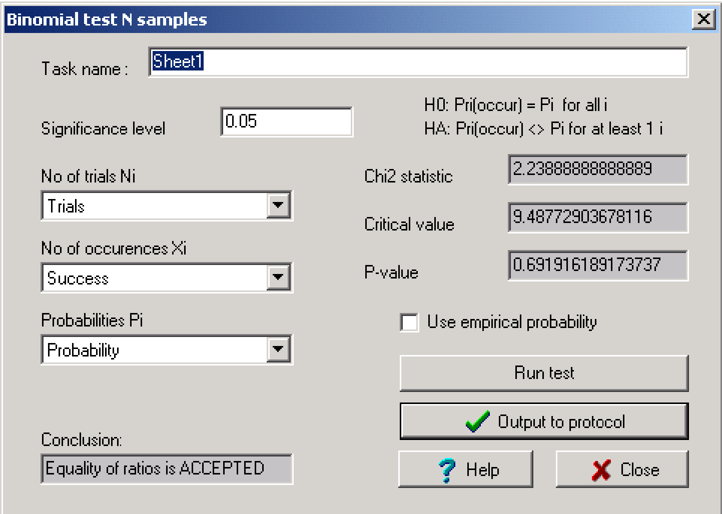

Fig. 9 Dialog window for Binomial test - N samples

Protocol

Binomial test for equal ratio, N samples	
Task name	Task name from the dialog window
Number of samples K	Number of tests.
Sample sizes Ni	Numbers of trials in each test
Number of occurrences Xi	Number of occurrences in each test
Theoretical occurrences Ni*Pi	Theoretical numbers of occurrences if H_0 were true
Actual ratios Xi/Ni	Observed number of occurrences
Ratio to be tested Pi	Given probabilities to be tested
Hypothesis H0	All true probabilities are equal to P_i
Hypothesis HA	Alternative to HA
Significance level	Significance level α
Degrees of freedom	Degrees of freedom
Statistic Chi2	Test statistic

Critical value	Maximal acceptable value of test statistic when H_0 holds
p-value	Biggest significance value at which H_0 would be rejected
Conclusion	Verbal conclusion of the test

Multinomial test

Menu:	QCExpert	Testing	Tests	Multinomial test

This module tests probabilities of multinomial distribution. It is used when the trials may have more than two exclusive outputs (events) with probabilities P_{Ai}, $i = 1, .. , K$, $K > 2$ and P_A are the true unknown probabilities of occurrence of the event A_i. If we perform N trials, we receive K frequencies, or numbers $X_1, X_2, ..., X_K$ of occurrences of events $A_1, A_2, ..., A_K$, $\Sigma X_i = N$. Here we test H_0: $P_{Ai} = P_i$ for all $i = 1, .. , K$ based on the observed $P_{Ai} = X_i/N$, whereby $\Sigma P_i = 1$ and $\Sigma P_{Ai} = 1$. Standard Chi-squared test is used. Assuming that all true probabilities P_{Ai} of the occurrences of A_i are equal to the given user-specified values P_i, the used statistic C has the distribution $\chi^2_{(K-1)}$.

$$C = \sum_{i=1}^{K} \frac{\left(X_i - NP_i\right)^2}{NP_i}$$

This statistic is then compared with critical quantile $\chi^2_{(K-1)}(1-\alpha)$. If C is bigger than critical quantile, we reject the H_0 hypothesis on the significance level α.

Data and parameters

This module expects data in two columns in the data editor. One column must contain numbers of occurrences X_i of each event A_i. Second column must contain the expected probabilities P_i, or occurrences of A_i, the third column may contain the probabilities P_i. If the third column is missing, you may set all P_i ($i = 1 .. K$) to empirical average value $P_i = \Sigma X_i/\Sigma N_i$. An example of input data is in the following table for $K = 4$. Note that the probabilities must sum to unity, $\Sigma P_i = 1$.

Occurrences	*Probability*
120	0.125
140	0.125
260	0.25
480	0.5

In the dialog window, see Fig. 10, you may modify the task name and the significance level (default value is $\alpha = 0.05$). Then select the columns with numbers of occurrences X_i and the probability values P_i respectively. After clicking the *Run Test* button the test is performed and results are displayed in the fields *Chi2 statistic*, *Critical value*, *p-value*. If the statistic is bigger than the critical value, the possible equality between all hypothesized P_i and true probabilities of the i^{th} event estimated by X_i/N is rejected. The field *Conclusion* will show verbal conclusion of the test: „*Equality of ratios is REJECTED or ACCEPTED*“. Clicking *Output to protocol* will produce a record of the last performed test in the Protocol window, while the dialog window still remains open. Clicking *Close* will close the dialog window.

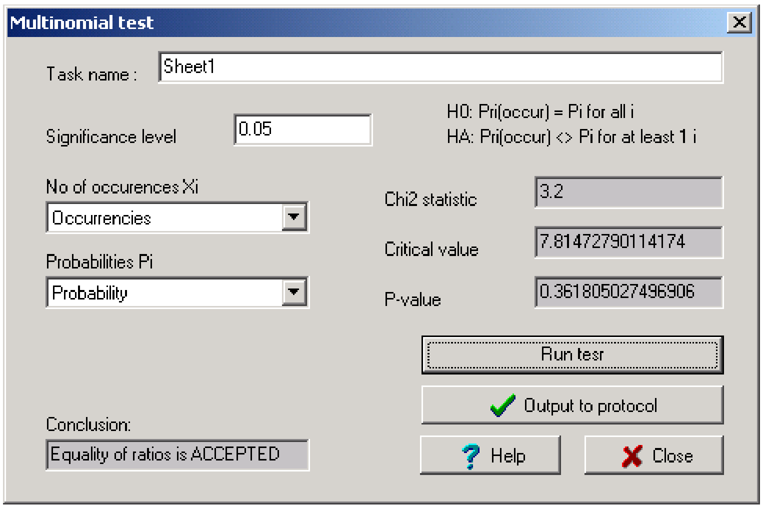

Fig. 10 Dialog window for multinomial test

Example

When throwing randomly a theoretical homogeneous matchbox with dimensions $a > b > c$ on a solid horizontal plane in an homogeneous conservative gravitational field (e.g. on a table in a common pub on the Earth), the matchbox may fall on either of the three sides: on the biggest one (side $A = ab$, event A) the smaller side, $B = ac$, event B, or on the smallest side, $C = bc$, event C. We want to carry out an experiment to support our hypothesis H_0 that the probabilities P_A, P_B, P_C of the positions A, B, C, see Fig. 11 are in the same ratio as the areas S_i of the sides divided by squared potential energy E_i of the respective position, $P_A : P_B : P_C = ab/c^2 : ac/b^2 : bc/a^2$, or $P_i \sim S_i/E_i^2$. The dimensions of the box are $a = 47$mm, $b = 35$mm and $c = 15$mm. Thus, the hypothesized probabilities would be $P_A = 0.89991$, $P_B = 0.07084$, $P_C = 0.02925$, since $P_A + P_B + P_C = 1$. In the experiment we received the position A in 1495 cases, position B in 115 cases and position C in 41 cases out of 1651 throws. We decided to carry out the test on the confidence level $\alpha = 0.05$. The data table will have the following form:

Events	Probabilities
1495	0.8999082056
115	0.0708382553
41	0.0292535391

Open the window *Multinomial test*. Leave the *Significance level* at 0.05. Select the first and second column in *No. of occurrences* and *Probabilities* and click on *Run test*. The conclusion reads *Equality of ratios is ACCEPTED*, which means that the experiment does not contradict our theory. (On the other hand, of course, by no means this confirms it.)

Fig. 11 The three matchbox positions

Protocol

Multinomial test for equal ratio	The module name

Task name	Task name from the dialog window
Number of classes K	Numbers of the classes *K*
Number of occurrences	Number of occurrences of each event
Theoretical occurrences N*Pi	Theoretical number of occurrences when H_0 holds
Actual ratios Ni/N	Observed ratios of the frequencies P_{Ri}
Ratio to be tested Pi	Given values P_i to be tested
Hypothesis H0	$P_{Ri} = P_i$
Hypothesis HA	$P_{Ri} <> P_i$
Significance level	Significance level α, usually α = 0.05
Degrees of freedom	Degrees of freedom η
Statistic Chi2	The χ^2 statistic calculated from data
Critical value	Theoretical quantile of the χ^2 distribution H_0
p-value	Calculated *p*-value of the test
Conclusion	Verbal conclusion of the test

Normal test, 1 sample

Menu:	QCExpert	Testing	Tests	Normal test 1 sample

This test is used to test equality between the mean value x_1 of normally distributed data and a given constant x_0. Null hypothesis is then H_0: $x_1 = x_0$ and alternative hypothesis H_A: $x_1 \neq x_0$ for two-sided test, or H_A: $x_1 > x_0$ for one-sided test. The test is based on the known arithmetic average x_1 and standard deviation *s*, that have been calculated from *n* measured samples.

We use the one-sample *t*-test, where the t-statistic T_1 is compared with the critical value *T*:

$$T_1 = \frac{x_1 - x_0}{s}\sqrt{n}; \quad T = t_{n-1}\left(1 - \alpha / 2\right)$$

$t_n(\alpha)$ denotes α-quantile of the Student distribution with *n* degrees of freedom. H_0 is rejected, if $|T_1| > T$. In one-sided mode of the test the critical value $T = t_{n-1}(1 - \alpha)$ is used.

Parameters

In the dialog window (Fig. 12) specify the significance level α, hypothesized mean value *X0*, measured average *X1*, standard deviation of the data *S* and sample size *N*. If only one-sided inequality is taken into account, one-sided alternative of the test should be selected in the field *Type of test*. After clicking *Run test* the values of *t*-statistic, critical value and *p*-value will be displayed in the respective fields. If the computed statistic is bigger than the critical value, then H_0 is rejected. The verbal conclusion has the form „*Difference between X0 and X1 is SIGNIFICANT*“ if H_0 is rejected or „*INSIGNIFICANT*“ if H_0 is accepted. Clicking *Output to protocol* will produce a record of the last performed test in the Protocol window, while the dialog window still remains open. Clicking *Close* will close the dialog window.

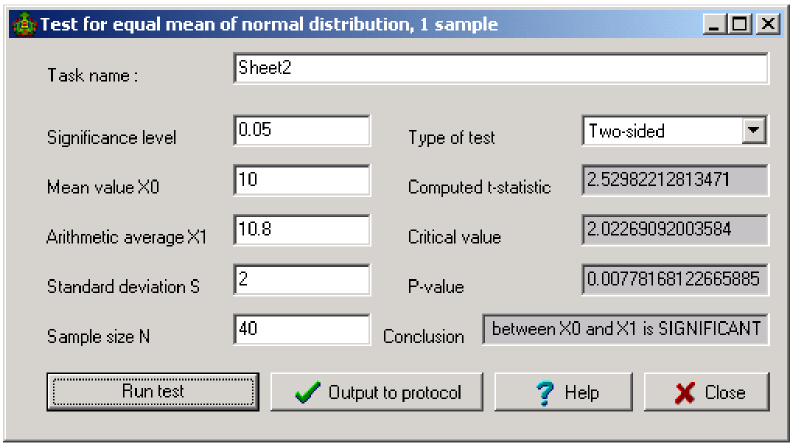

Fig. 12 Dialog window for normal test, 1 sample

Protocol

t-test one sample	Name of the module
Task name	Task name from the dialog window
Mean value X0	The given tested value
Sample average X1	Average of the data
Standard deviation S	Standard deviation of the data
Degrees of freedom	$n-1$
Computed t-statistic	The value of the sample *t*-statistic T_1
Critical value T	Critical quantile $t_{(n-1)}(1-\alpha)$ of the *t*-distribution
p-value	Computed *p*-value
Conclusion	Verbal conclusion of the test

Normal test, 2 samples

Menu:	QCExpert	Testing	Tests	Normal test 2 samples

This test is used to test equality between two mean values of normally distributed data. Null hypothesis is H_0: $\mu_1 = \mu_2$ and alternative hypothesis H_A: $\mu_1 \neq \mu_2$ for two-sided test, or H_A: $\mu_1 > \mu_2$ for one-sided test. The test is based on the known sample arithmetic averages x_1, x_2 and standard deviations s_1 and s_2 of the samples, that have been calculated from n_1 and n_2 measurements of the first and second sample respectively. The test is based only on the 4 sample statistics x_1, x_2, s_1 and s_2, does not use original measurements and assumes normality and not too different variances of the data. If the original data are available, the module *Two samples comparison*, is recommended to test the mean values.

We use the two-sample *t*-test, where the *t*-statistic T_1 is compared with the critical value T:

$$T_1 = \frac{|x_2 - x_1|}{\sqrt{(n_1-1)s_1^2 + (n_2-1)s_2^2}} \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}}; \quad T = t_{n_1+n_2-2}(1-\alpha/2)$$

$t_n(\alpha)$ denotes α-quantile of the Student distribution with n degrees of freedom. H_0 is rejected, if $|T_1| > T$. In one-sided mode of the test the critical value $T = t_{n-1}(1-\alpha)$ is used.

Parameters

In the dialog window (Fig. 13) specify the significance level α, average of the first and second sample *X1*, *X2*, standard deviations of the samples *S1 and S2* and sample sizes *N1*, *N2*. If only one-sided inequality is taken into account, one-sided alternative of the test should be selected in the field *Type of test*. After clicking *Run test* the values of *t*-statistic, critical value and *p*-value will be displayed in the respective fields. If the computed statistic is bigger than the critical value, then H_0 is rejected. The verbal conclusion has the form „*Difference between X1 and X2 is SIGNIFICANT*" if H_0 is rejected or „*INSIGNIFICANT*" if H_0 is accepted. Clicking *Output to protocol* will produce a record of the last performed test in the Protocol window, while the dialog window still remains open. Clicking *Close* will close the dialog window.

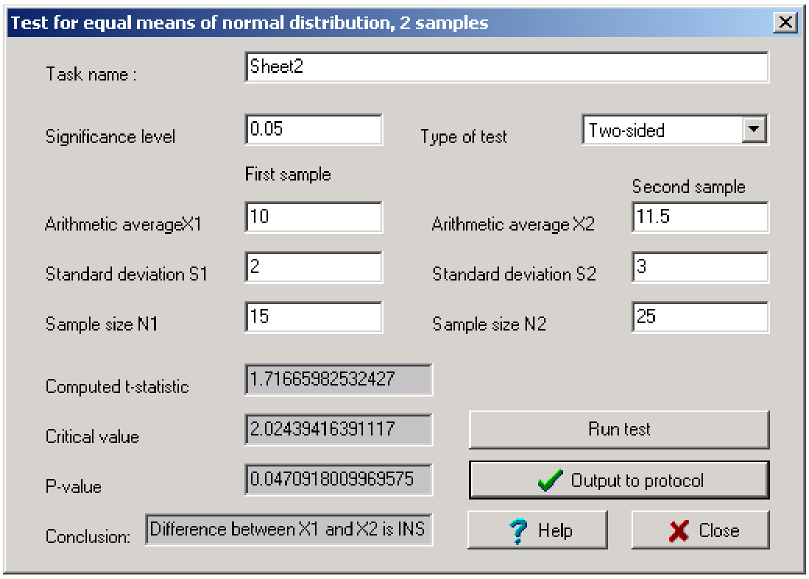

Fig. 13 Dialog window for normal test, 2 samples

Protocol

t-test two sample	Module name
Task name	Task name from the dialog window
Sample average X1	Average of the first data sample
Standard deviation S1	Standard deviation of the first data sample
Sample average X2	Average of the second data sample
Standard deviation S2	Standard deviation of the second data sample
Degrees of freedom	$n_1 + n_2 - 2$.
Computed t-statistic	The value of the sample *t*-statistic T_1
Critical value T	Critical quantile of the *t*-distribution on the significance level α
p-value	Computed *p*-value
Conclusion	Verbal conclusion of the test