Combinations of variables, especially so-called linear combinations, are an important area of the understanding and use of statistical methods. With a linear combination of variables we mean sums or differences between random variables. All other types of combinations are then called non-linear. One reason for this division is that e.g. expected values and standard deviation of the result variable (Y) easily can be calculated for linear combinations and because of this type being very common. Examples:

Linear combinations: $Y_1 = X_1 + X_2$ $\qquad$ $Y_2 = a \cdot X_1 + b \cdot X_2 - X_3$

Non-linear combinations: $Y_3 = \sqrt{X_1^2 + X_2^2}$ $\qquad$ $\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}$

See the first page of the routine for information of expected value and standard deviation for the variable Y. Note that we do not express anything regarding the distribution of Y. However, this can be derived with mathematical tools if the distributions are known for the single terms.

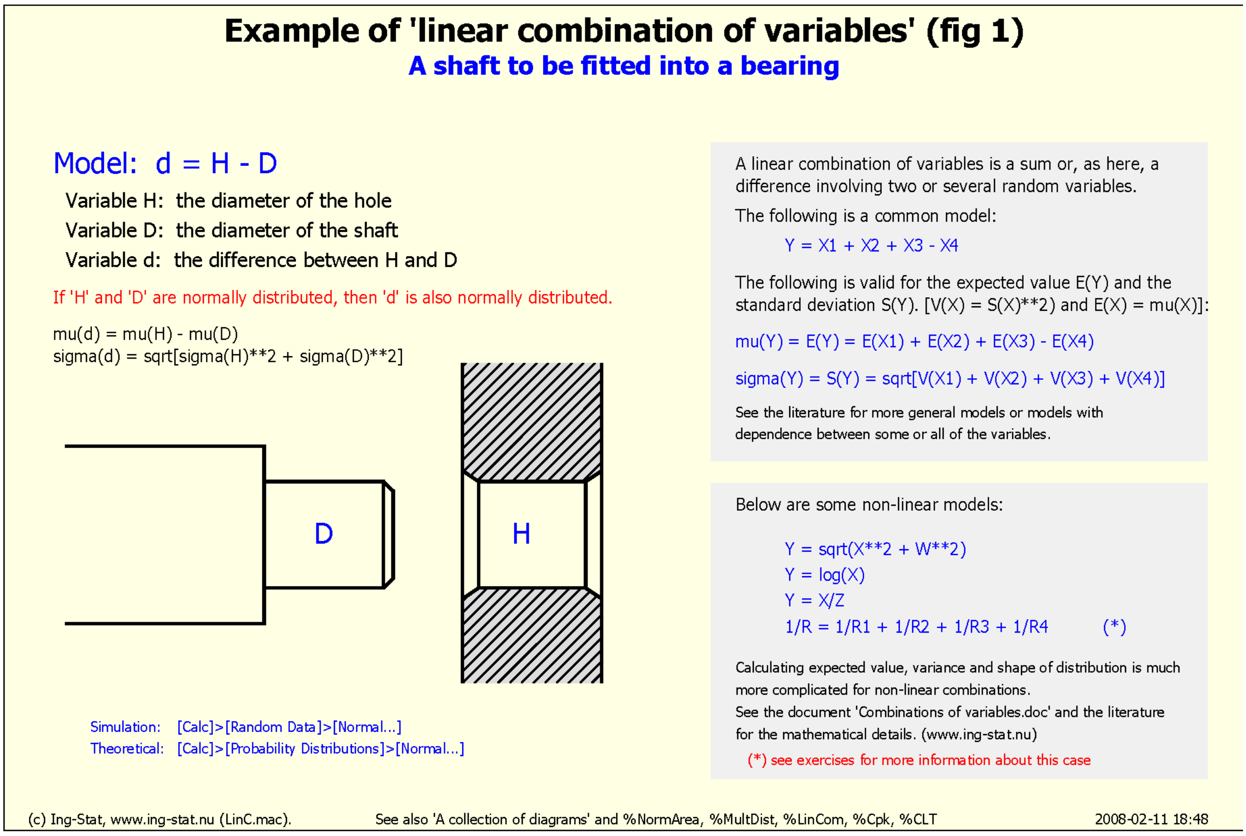
In the linear combinations above there are a fixed number of terms. We can also imagine examples where the number of terms varies. The total inflow of daily cash in a shop consists of a random number of terms (customers) that buy for a random sum of money. Such situations are not treated in this routine.

Macros and menus
%LinCom, %LinModel, %CLT, %Die

[Calc]>[Random Data]>[Normal…]
[Calc]>[Probability Distributions]>[Normal…]
[Graph]>[Histogram…]

Data on the printout

There is no data on the printout



The idea with combinations of variables is rather simple. We can regard it as a formula even if we within statistics use the expression model. It is rather easy to e.g. transfer a drawing to a mathematical formula or a model. Suppose that we build e.g. a high pressure pump assembled from a number of identical segments. The total length is then a simple sum of a number of lengths.

Suppose that we build an electronic circuit. Even in this case we can, if we know the area properly, calculate a result given the variables of the circuit. It is possible that the relationship is more or less complicated and need more or less mathematics. It is easy to see that many technical and administrative processes contain a number of combinations of variables.

The basic principle is that if there is variation in the variables this variation will transfer to the resulting variable. (It is possible to think that there is a so-called negative correlation between variables decreasing the total variation. This fact is used when trimming different types of electronic circuits.) The main three questions are thus expected value, standard deviation and type of distribution.

The model: $Y = a \cdot X_1 + b \cdot X_2 + c \cdot X_3 + d \cdot X_4$

Mean and std: $\mu_Y = a \cdot \mu_{X_1} + b \cdot \mu_{X_2} + c \cdot \mu_{X_3} + d \cdot \mu_{X_4}$ $\qquad \sigma_Y = \sqrt{a^2 \cdot \sigma^2_{X_1} + b^2 \cdot \sigma^2_{X_2} + c^2 \cdot \sigma^2_{X_3} + d^2 \cdot \sigma^2_{X_4}}$

Data on the printout

Expected value (μ)	=	Expected value
Standard dev. (σ)	=	Standard deviation
Average (x-bar)	=	Average value calculated from the data
Standard dev. (s)	=	Standard dev calculated from the data
Maximum value	=	Maximum value in the data
Minimum value	=	Minimum value in the data
Number of values (n)	=	Number of values in the simulation

Macros and menus
%LinCom, %LinModel, %CLT, %Die

[Calc]>[Random Data]>[Normal…]
[Calc]>[Probability Distributions]>[Normal…]
[Graph]>[Histogram…]