[6.13] Find coefficients of determination for all regression equations

The 89/92+ can fit 10 regression equations, but do not find the coeffcient of determination r^2 for all the equations. However, r^2 is defined for any regression equation, and these two functions calculate it:

```
r2coef(lx,ly)
func
©Find coefficient of determination r^2, no adjustment for DOF
©lx is list of x-data points
©ly is list of y-data points
©24 nov 99/dburkett@infinet.com

1-sum((regeq(lx)-ly)^2)/sum((ly-mean(ly))^2)

Endfunc


r2coefdf(lx,ly)
func
©Find coefficient of determination r^2, adjusted for DOF
©lx is list of x-data points
©ly is list of y-data points
©24 nov 99/dburkett@infinet.com
local n
dim(lx)→n

1-((n-1)*sum((regeq(lx)-ly)^2))/((n-dim(regcoef)-1)*sum((ly-mean(ly))^2))

Endfunc
```

Note that neither of these functions will account for weighting of the data points.

Both functions should be run in Approx mode. Both take the list of x-data and y-data values in the lists *lx* and *ly*. Be sure to run the regression command (*CubicReg*, *ExpReg*, etc) before using these functions, because they depend on the system variables *regeq* and *regcoef* being up-to-date.

regeq is a function that is updated by the regression command. It is the regression equation itself, and I use it in both routines to calculate needed values. *regcoef* is the list of regression equation coefficients, and I use it in *r2coefdf()* to find the degrees of freedom.

The coefficient of determination, without adjustment for degrees of freedom, is defined as

$$r2 = 1 - \frac{SSE}{SSM}$$

SSE is the sum of the squares of the error, defined as

$$SSE = \sum_{i=1}^{n} (\hat{y}_i - y_i)^2$$

SSM is the sum of squares about the mean, defined as

$$SSM = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

Also,

n is the number of data points
y_i are the y-data values
$\hat{y}_i = f(x_i)$ are the estimated y-values, that is, the regression equation evaluated at each x-data value

$\bar{y}$ is the mean of the y-data values

It is a simple matter to adjust r^2 for degrees of freedom:

$$r2 = 1 - \frac{(n-1)\cdot SSE}{(DOF-1)\cdot SSM}$$

Here, DOF is the degrees of freedom, defined as

$$DOF = n - m$$

where *m* is the number of coefficients in the equation. For the simple linear equation y = ax + b, m = 2, because we have two coefficients *a* and *b*.

It is appropriate to use r^2 without DOF adjustment when the data to be fit has no experimental errors. For example, this is the case if you are fitting a curve to data generated by some other mathematical function. However, the DOF adjusted r^2 should be used when the y-data values include experimental error.