

Exercise 6 – Linear fit and correlation coefficient

1. Use **cftool**, find chi-square of the fit.

```
x=[ 2 4 6 8 10 12 14 16]
y=[5.3 14.4 20.7 30.1 35.0 41.3 52.7 55.7]
```

Similarly, below are measurements of the temperature of water in a lab, against time in minutes. Find out if the slope is statistically significant

```
Time=[0 1 2 3 4 5]
T=[98.51 98.50 98.50 98.49 98.52 98.49]
```

2. Caveats in calculating r^2 . Plot the following data and calculate the correlation coefficient (e.g. using cftool)
 - a) `x1=1:10;`
`y1=x1.^2` % parabolic dependence
 - b) `x2=random('Normal',2,1,1,5)` % five random numbers from Normal (2,1)
`y2=random('Normal',2,1,1,5)` % another random set
`x2(6)=20` % the last (sixth) point is far on the right
`y2(6)=20`

Check if the spurious correlations could be caught using chi-square or by a t-test for the correlation coefficient:

$$t = r \sqrt{\frac{N-2}{1-r^2}}.$$

Remember that the t-test in this case tests the null hypothesis that the data are *not* correlated.