Week 5 exercises, statistics

1. Generate a 5 by 4 matrix of random numbers. Assuming that each column in the matrix contains measurements for that variable, compute the:

   - means of the variables
   - variances of the variables
   - covariance matrix of the variables
   - correlation matrix of the variables.

   Use the formula definition given in the lecture slides. **Do not** use the direct Matlab built-in functions (mean, var, cov, corr, corrcov).

2. Fit a straight line $y = \theta_0 + \theta_1 x$ to the data given below. Compute the:

   - covariance matrix of unknown parameters (hint: obtain an estimate of the measurement error using the repeated measurements)
   - the *t-values* of the unknown parameters
   - the R^2 value of the model.

x	0.0	1.0	1.0	2.0	1.8	3.0	4.0	5.2	6.5	8.0	10.0
y	5.00	5.04	5.12	5.28	5.48	5.72	6.00	6.32	6.68	7.08	7.52

3. For each of the data sets given below, perform the leave-one-out cross-validation to estimate the accuracy of a linear model ($y = ax + b$) and the accuracy of a quadratic model ($y = ax^2 + bx + c$). Use Q^2 to quantify the accuracy.

x	0.0	1.0	1.0	2.0	1.8	3.0	4.0	5.2	6.5	8.0	10.0
y	5.00	5.04	5.12	5.28	5.48	5.72	6.00	6.32	6.68	7.08	7.52

x	0.9	1.3	1.3	1.4	1.6	1.8	2.1	2.1	2.5	2.9	3.2	3.3	3.6	4.0	4.2	4.3
y	2.5	4.03	4.1	4.4	5.1	6.05	7.48	7.6	9.8	12.4	14.3	15.2	18.1	19.9	23.0	23.9

x	-3	-2	-1	0	1	2	3
y	7.5	3	0.5	1	3	6	14

4. Consider the simple nonlinear model which has been given before in a previous task, $y = \theta_1 x e^{(1+\theta_2 x^2)}$. Using xy.mat for calibration, compute the approximative covariance matrix of the parameters using the Jacobian of the model analytically.