TD604, Tutorial 2

1. Load the scilab object *table* from the webpage (at `sohoni/TD604/sundry/table`). This is a 157 X 2 matrix of the heights of husband-wife pairs (in feet). For this data set, (i) what is the mean husband height μ_M and the mean wife height μ_F? (ii) plot the best fit line and verify that the line passes through the two means.

   - Divide the data set into to parts, (i) MF, where the men are taller than spouses, and (ii) FM, where the women are taller than their spouses. Compute means and variances of these sets.

   - Divide the men into two halves M_1 and M_2, men below and men above the mean. Similarly, separate the data set into two sets F_1 and F_2. For each of these sets, compute the mean of the husbands and wives.

2. Agricultural work is poorly paying and is thought to be done by the lower castes. From your district data, design an experiment to test this hypothesis.

3. Given the following data, compute my hand, the best-fit line, i.e., $y = ax + b$ and the best fit quadratic $y = ax^2 + bx + c$. Compute the R^2 for both fits.

x	-2	-1	0	1	2
y	1	1	0	0	1

4. Use `reglin` to compute the best fit linear and quadratic model for *fraction-under-6* vs. *literate-fraction* for your taluka. Compare the R^2. Do this for another pair of attributes (say employment vs. literacy) for your taluka.

5. List the outcome set for 5 coin tosses. If the probability of heads is 0.7, compute the probabilities of i heads, for $i = 0, \ldots, 5$. Now use an appropriate scilab command (it exists) to compute these probabilities for $n = 5, 10, 15$. Plot these probabilities.

6. A bag has 5 red balls and 3 black balls. A ball is randomly drawn from this bag and thrown away. Now, a second ball is drawn from the bag. What is the probability that this ball is red?

7. Construct a 1000×2 matrix iA of uniform random variables. Construct next $B(i) = \max(A(i, 1), A(i, 2))$. Thus B is the maximum of each row. Plot a histogram for B and try and explain the shape.

8. Let D_1 and D_2 be two dice. We roll these dice and compute the outcome as $d_1 - d_2$, i.e., the subtraction of the outcome of D_1 from D_2. What are the outcomes for this new random variable and what are the probabilities?