Assignment 2 due Sep 29

1. A probability distribution function is described by

$$p(x) = \sqrt{\frac{c}{2\pi}} \frac{e^{-c/2x}}{x^{3/2}},$$

where c=10 is a parameter. (By the way, this is one form of a so-called Lévy distribution. On the course webpage I posted a reference to a *Nature* paper that discusses its applications in describing foraging behaviors of many animals. An important property of this distribution is that it has a “fat tail”, i.e. power law (non-exponential) decay for large x. Thus it can describe “unlikely but very large events”, such as when albatrosses fly 1000s of km to find a new feeding area.)

a) Find the most probable value of *x*.
b) Plot the c.d.f. and verify if this p.d.f. can indeed describe probability.
c) Find the mean and variance of this distribution. Do your results make sense? (you may need to read more about Levy distributions).
d) If this distribution describes the frequency with which albatrosses fly *x* kilometers to find a new foraging area, what is the probability that a given flight will be over 100 km?

2. Take the height of men data (second column in body_men.mat file on the class website). Fit a normal distribution, e.g. by using ’dfittool’. Perform a chi-squared test and conclude whether the distribution can be considered normal.

3. The New York Times opinion poll on September 14, 2010 found that 40% of the respondents would vote for a Republican candidate, whereas 38% would vote for a Democratic candidate, with the rest giving some other answer. The report specified that a total of 990 people were polled. Assuming, for the purpose of this exercise, that only the people who gave a definite answer would vote, and also that the only two options available to them are either Republican or Democrat,
   a) find the uncertainty in the reported poll numbers.
   b) find the probability that in reality (or, in the limit of infinite sample size) the actual number for Democrats was higher than the number for Republicans.

4. The initial activity N_0 and the mean life τ of a radioactive source are known with uncertainties of 1% each. The activity follows the exponential distribution $N = N_0\exp(-t/\tau)$. The uncertainty in the initial activity N_0 dominates for small times t; the uncertainty in the mean life τ dominates for large t ($t >> \tau$) (Why?). For what value of t/τ do the uncertainties in N_0 and τ contribute equally to the uncertainty in N? What is the uncertainty in N at that point?