Exercises chapter 4

1. Make a short program to generate figure 4.1 on page 103. Try out a number of sets of propositions with alternative error frequencies, i.e, A="a% defective", B="b% defective" and C = "c% defective" with some suitable values for a, b, and c. Interpret your plots.

2. Consider the two hypotheses

   - $A =$ "we have $f_A = 1/6$ defective widgets."
   - $B =$ "we have at least f_B defective widgets".

   Note that B can be considered as a logical sum of a large number of propositions $B = B_1 + B_2 + ...$ that are mutually exclusive. The proposition B_i is then "the fraction f of bad widgets lies in the interval $(f_i, f_i + df)$".

   For the two cases in (a) and (b) below, plot the evidences $e(A|DX)$ and $e(B|DX)$ (as functions of m) where D stands for "m widgets were tested and all were found defective". Use the priors

   $$P(A|X) = \frac{10}{11} \qquad P(B|X) = \frac{1}{11} \tag{1}$$

   (a) $f_B = 1/4$

   (b) $f_B = 1/2$. Note that we here have, to some small extent, overlapping propositions. Do you need to make adjustments in your derived formulas for this?

3. Consider again the widget scenario. Derive an expression for, and write a function for calculating the probability $P(m|n, f \in [f_1, f_2], X)$ where m is shorthand for "m widgets were found defective" and n is shorthand for "n widgets were drawn". Assume a uniform prior distribution for f in the interval. Use this function to calculate the evidences for the three hypotheses

   - $A =$ "$f \in [0, f_a]$"
   - $B =$ "$f \in (f_a, f_b]$"
   - $C =$ "$f \in (f_b, 1]$"

   Choose, for instance, $f_a = 0.1$ and $f_b = 0.3$. Then choose some f_{true} to be the true value of f. Preferably, choose $f_{true} = 1/N_{true}$ with N_{true} being some integer. For this f you will on the average see one bad widget every N_{true}'th draw. Let D stand for "$n = m \times N_{true}$ widgets were drawn, m were found defective." Plot the evidences $e(A|DX)$, $e(B|DX)$, and $e(C|DX)$, as functions of m. Specify what priors $P(A|X)$, $P(B|X)$, and $P(C|X)$ you have used.

   Hint 1: You may find the matlab functions `betainc` and `beta` useful when writing the code.
   Hint 2: To check that your code is correct, you can choose a very small interval $[f_1, f_1 + \Delta f]$ and compare $P(m|n, f \in [f_1, f_1 + \Delta f], X)$ with $P(m|n, f = f_1, , X)$. These probabilities should be approximately equal.
   Hint 3: You can find a numerically stable implementation of the binomial coefficient on the course home page. Although the coefficients will cancel when you calculate the evidences you might want to have a correct implementation of the function calculating the probability $P(m|n, f \in [f_1, f_2], X)$ anyway.