Problem Set 2
Due April 15, 2014 **BENG 100 S14**

1. This is another Matlab program now related to the lecture on April 8, 2014. Consider a sequential experiment involving two rolls of a 4-sided die. For the first die roll, the probabilities are as follows:

$$1 : 0.4 \quad 2 : 0.2 \quad 3 : 0.3 \quad 4 : 0.1$$

   The second roll takes on the same parity as the first roll with probability 0.6 and takes on parity different than the first with probaility 0.4. Define A_i to be the event that four-sided die takes the value i and B_k to be the event that the second value takes the value k.

   - create a column vector called **prior** where the ith entry corresponds to $\mathbb{P}(A_i)$.
   - create a matrix called **likelihood** where rows correspond to i and columns correspond to k with $L(i,k)$ corresponding to $\mathbb{P}(B_k|A_i)$.
   - write a function **totalProbabilityTheorem** with inputs being **prior** vector and **likelihood** matrix that spits out a vector **pb** whose kth entry is $\mathbb{P}(B_k)$.
   - write a function called **Bayes** with inputs being index k, **prior** vector, and **likelihood** matrix. Build this function so that it performs Bayes rule: the output of the function is a vector **pagivenb** whose ith entry is $\mathbb{P}(A_i|B_k)$.
   - use these functions to calculate $\mathbb{P}(A_1)$, $\mathbb{P}(B_3|A_4)$, $\mathbb{P}(A_1 \cap B4)$, and $\mathbb{P}(B_3)$, and $\mathbb{P}(A_2|B_2)$.

2. Use the axioms of probability and the law of conditional probability to show that if $A_1, A_2, \ldots, A_m$ form a partition of the sample space Ω, then for any event B:

$$\sum_{i=1}^{m} \mathbb{P}(A_i|B) = 1.$$

3. Show that if events A and B are statistically independent, then A^c and B are also statistically independent, and likewise for the pairs (A, B^c) and (A^c, B^c).

4. (a) Prove De Morgan's laws from the lecture notes. (*remember: to show that $A = B$, you must show that $A \subset B$ and $B \subset A$. This means that you must reason that if $x \in A$ then $x \in B$, and analogously if $x \in B$ then $x \in A$*).

   (b) Consider events A_1, A_2, and A_3 that are statistically independent with

$$\mathbb{P}(A_1) = 0.8 \quad \mathbb{P}(A_2) = 0.35 \quad \mathbb{P}(A_3) = 0.75$$

   Calculate $\mathbb{P}(A_1 \cup A_2 \cup A_3)$. *Hint: use part (a) and the previous problem. You will not receive full credit unless you justify every step.*

5. The a priori likelihood that a drug maker has a defect during manufacturing is 0.1. If it does not have a defect, then the testing by the FDA provides risk levels as follows: 0 with probability 0.7, 1 with probability 0.2, and 2 with probability 0.1. If the drug does have a defect during manufacturing, then the FDA's testing procedure provides risk levels at: 1 with probability 0.3, 2 with probability 0.5, and 3 with probability 0.2.

   (a) Draw a sequential description of the sample space.

   (b) Calculate the probability that there is a drug defect and the risk level equals 2.

   (c) Calculate the probability the risk level equals 2.

   (d) Calculate the probability there is a defect given the readout equals 1.

   (e) Calculate the probability that the risk level 1 or 2.

   (f) Calculate the probability that there is a drug defect given the readout equals 1 or 2.

6. In this problem we seek to prove the LOTUS theorem. Consider a discrete random variable X and a function g whose range is the integers.

   (a) Define $Y = g(X)$. Show that

   $$\mathbb{P}(Y = m) = \sum_{k:g(k)=m} P_X(k)$$

   (b) For any m and any $m' \neq m$, show that if $g(k) = m$ then $g(k) \neq m'$ (*hint: this is not rocket science - give a simple example*).

   (c) Argue that

   $$\mathbb{E}[Y] = \sum_m m P_y(m)$$

   can be expressed as

   $$\mathbb{E}[Y] = \sum_k g(k) P_X(k)$$

   by arguing that exhausing all k in the latter is the same as exhausting all m in the former. Use a simple example with $g(k) = k^2$ where X can take on values 1, 2, or 3 with nonzero probability to describe the result.