CS 530: Geometric and Probabilistic Methods in Computer Science
Homework 1 (Fall '10)

1. The joint p.m.f. of two discrete random variables X and Y is given below:

$$\begin{array}{cc} & Y \\ X & \begin{bmatrix} 1/72 & 1/36 & 5/72 \\ 1/36 & 1/18 & 5/36 \\ 1/12 & 1/6 & 5/12 \end{bmatrix} \end{array}.$$

   Determine whether X and Y are statistically independent.

2. The joint p.m.f. of two discrete random variables X and Y is given below:

$$\begin{array}{cc} & Y \\ X & \begin{bmatrix} 0 & 5/36 & 1/3 \\ 1/12 & 1/9 & 1/18 \\ 1/36 & 1/4 & 0 \end{bmatrix} \end{array}.$$

   Compute and tabulate:

   (a) Marginal p.m.f., $p_X(x_i)$.

   (b) Marginal p.m.f., $p_Y(y_j)$.

   (c) Conditional p.m.f., $p_{X|Y}(x_i \,|\, y_j)$.

   (d) Conditional p.m.f., $p_{Y|X}(y_j \,|\, x_i)$.

3. Three boxes of identical appearance contain two coins each. In the first box, both are gold, in the second box both are silver, and in the third box one is silver and one is gold. Suppose that a box is selected at random, and further, that a coin in that box is selected at random. If the coin proves to be gold, what is the probability that the other coin in the box is also gold? [Hint: There are two discrete random variables: 1) the box, X, with outcomes, $\{x_{gg}, x_{ss}, x_{sg}\}$ and; 2) whether the first coin selected from the box is gold or silver, Y, with outcomes $\{y_g, y_s\}$. You need to compute $P_{X|Y}(x_{gg} \mid y_g)$. Note that $P_{X|Y}(x_{gg} \mid y_g) \neq \frac{1}{2}$.]

4. Let X be a discrete random variable with the following p.m.f.:

$$p_X(x_i) = \begin{cases} c\, x_i^2 & x_i \in \{1,2,3,4\} \\ 0 & \text{otherwise.} \end{cases}$$

- Find c.
- Find the expected value and variance of X.

5. The continuous random variable X has p.d.f.:

$$f_X(x) = \begin{cases} c\,x^2 & \text{if } 1 \le x \le 4 \\ 0 & \text{otherwise.} \end{cases}$$

   - Find c.
   - Find the expected value and variance of X.
   - Find the probability that X exceeds 2.0.

6. A p.d.f. is defined as follows:

$$f_X(x) = \begin{cases} 4x/50 & \text{if } 0 \le x \le 5 \\ 0 & \text{otherwise.} \end{cases}$$

   Find the value of $F_X(3)$, i.e., the value of the c.d.f. at 3.

7. A pattern recognition system has been developed for classifying objects into six categories. It has been correct on 80 percent of the test data. What is the probability that exactly four of the next five samples will be classified correctly?

8. Let X and Y be continuous random variables where

$$f_X(x) = \begin{cases} \frac{2}{\sqrt{2\pi}} e^{-x^2/2} & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

   and let $Y = \ln X$. Derive an expression for f_Y.

9. A company wants to hire three new employees. Eight people apply for the job. Three of the applicants are male and five are female. The company chooses which applicants it will hire at random. Let M be the number of males hired. Compute $p_M(m)$ for all m.

10. Prove that

$$1 - (1-p)^n = \sum_{k=1}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) p^k (1-p)^{n-k}.$$

11. Two fair dice are rolled, one of which is red and the other is green. Let $W = R + G$ be the sum of the values of the two dice. Compute and plot $p_W(w)$.

12. Write a MATLAB function which computes the $2D$ joint histogram, G_{XY}, of a pair of images, X and Y, of equal size. A joint histogram is a table of size 256×256 where the (i, j)-th element is a count of how frequently a pixel in image X has value i and the pixel at the same location in image Y has value j. Test your program on the red and green color components of the *cactus* image found on the class homepage. Display the joint histogram, G_{XY}, as a grey level image. [Note: Use the online documentation to find out how to generate hardcopies of images in MATLAB (or Octave).]

13. Write a MATLAB function which, given a joint histogram, G_{XY}, returns the marginal histograms, G_X and G_Y. Using the *cactus* image, verify that the marginal histograms you compute are the same as those computed using the $1D$ histogram function found on the class homepage. [Note: Use the online documentation to find out how to plot and generate hardcopies of histograms in MATLAB (or Octave).]