Due Thursday, April 11, beginning of class

Assigned Problems

1. **(5.6.1)** Suppose X has density function $f(x) = c(3-|x|)$ when $-3 < x < 3$. What value of c makes this a density function?

2. **(5.6.4)** Suppose X has density function $x^2/9$ for $0 < x < 3$ and 0 otherwise. Find (a) EX, (b) EX^2, and (c) $\text{var}X$.

3. **(5.6.8)** Let $F(x) = 3x - 2x^2$ for $0 \le x \le 1$, $F(x) = 0$ for $x \le 0$, and $F(x) = 1$ for $x \ge 1$. Is F a distribution function? If so, find its density function.

4. **(5.6.9)** Suppose X has density function $f(x) = x/2$ for $0 < x < 2$ and 0 otherwise. Find (a) the distribution function, (b) $P(X < 1)$, (c) $P(X > 3/2)$, and (d) the median.

5. **(5.6.18)** Suppose X has density function $f(x)$ for $a \le x \le b$ and $Y = cX + d$, where $c > 0$. Find the density function of Y.

6. **(5.6.24)** Suppose X is uniform on $(0, \pi/2)$ and $Y = \sin X$. Find the density function of Y. The answer is called the *arcsine law* because the distribution function contains the arcsine function.

7. **(5.6.28, modified)** Suppose X and Y have joint density $f(x, y) = 6xy^2$ for $0 < x$, $y < 1$.

   (a) Consider $P(X \le x, Y \le y) = \displaystyle\int_0^x \int_0^y 6vu^2 \, dudv$. Does this describe the joint distribution function of X and Y? If it is, explain why, and if it isn't, determine the correct distribution function.

   (b) What is $P(X + Y < 1)$?

8. **(5.6.29, modified)** Suppose X and Y have joint density $f(x, y) = 2$ for $0 < y < x < 1$. Find $P(X - Y > c)$, for some c.

9. **(5.6.34)** A pair of random variables X and Y takes values between 0 and 1 and has $P(X \le x, Y \le y) = x^3y^2$ when $0 \le x$, $y \le 1$. Find the joint density function.

10. Download the Matlab script RVfromData.m and the data file data.mat from Canvas. This script has been divided into sections; you can either run the whole script by pressing 'Run', or individual sections by pressing 'Run Section'.

    (a) Run the first section. This will load the data into Matlab and plot the histogram of the data. You can look at the data as a table by double-clicking on the name 'data' in the Workspace. The histogram resembles a density function. How is the histogram similar to or different from a density function?

    (b) In the next section, we divide the values of the histogram by their sum to normalize them. How is *this* set of numbers similar to or different from a density function?

    (c) Run the third section to create an approximation to the cumulative distribution of the normalized data. Why do we use the sum of the normalized values of the histogram rather than normalized data values?

    (d) In the next section, we draw a random number X from this distribution. Look up the documentation for the 'find' function. What is it doing?

(e) Alter the last section to include a 'for' loop that draws 100 random numbers from this distribution. Store each random number as an entry in an empty vector, created using `X=NaN(size(Y)))`. Check your work by visualizing your results with a histogram. Why is the shape of this histogram what you would expect? What result are we using? (*Turn in your code and a detailed sketch or printout of any plots.*)

(f) (**Bonus 5 points**) Alter this process to draw 1,000 random numbers, where each X has density $f(x) = x/2$ for $0 < x < 2$ (Problem 3). Use this to estimate the density for $Z = X^2$. Is this similar to the density you would expect for Z? Why or why not? (*Turn in your code and a detailed sketch or printout of any plots.*)