Assignment 3
Probability
Math 363
March 1, 2012

1. A lake contains 160 lake trout. 18 are captured, tagged, and released. A certain time later, 8 of the 160 are captured.
   (a) What is the probability that exactly two of them are tagged? State clearly what assumptions you are using.
   (b) Find the probability mass function for the number $x$ that are tagged. Check that the sum is 1.
   (c) Sketch the cumulative distribution function.
   (d) Find the mean and variance of the number tagged fish in the second capture.

2. Colorblindness is a sex linked trait on the X chromosome. Its frequency is 5% in the population.
   (a) A male is color blind if he received the allele for colorblindness. What fraction of the male population is colorblind?
   (b) A female is color blind if she received two alleles for colorblindness. Assuming that the probability of receiving the allele is 5% independently from each of her parents, what fraction of the female population is colorblind?
   (c) Assuming that the population is made up 1/2 men and 1/2 women, find the fraction of the population that is color blind.
   (d) If a person chosen at random is color blind, find the probability that the person is female.
   (e) If six randomly chosen individuals are color blind, find the probability one or fewer are female.

3. Let $X$ be a discrete random variable with probability mass function

   \[
   
   \begin{array}{c|ccccccc}
   x & 1 & 2 & 3 & 4 & 5 & 6 \\
   f_X(x) & c & 2c & 3c & 3c & 2c & c \\
   \end{array}
   \]

   (a) Find $c$.
   (b) Find $P\{X > 3\}$
   (c) Draw the cumulative distribution for $X$.
   (d) Find $EX$, and Var($X$).
   (e) Simulate 1000 random variables with this mass function and find their sample mean and variance.
   Compare your answer to your answer in part (d)
4. Consider a random variable $X$, a mixture of two exponential random variables, the distribution is

$$F_X(x) = 1 - \frac{1}{2}(e^{-x} + e^{-3x}), \quad x > 0,$$

and $F_X(x) = 0, \quad x \leq 0$.

(a) Give a plot of $F_X$ and explain using this plot why $F_X$ is a valid distribution function.

(b) On this plot show the 4 quintiles $X$.

(c) Find the probabilities $P\{X \leq 1\}, P\{X \leq 3\}$, and $P\{1 < X \leq 3\}$.

(d) Find the probability density for this distribution function.

(e) Sketch the distribution and density function indicating $P\{1 < X \leq 3\}$.

5. The Pareto random variable with parameters $\alpha > 0$ and $\beta > 0$ has probability density function

$$f_X(x) = \frac{\beta \alpha^\beta}{x^{\beta+1}}, \quad \alpha < x < \infty.$$

(a) Verify that $f_X$ is a density function.

(b) Find $P\{X > 2\alpha\}$.

(c) Find the mean and variance of $X$. What restriction do you have on $\beta$ in computing the variance?

(d) Use the probability transform to simulate 1000 Pareto random variables with $\alpha = 4$ and $\beta = 4$ and find their sample mean and variance.

6. In this problem, we shall use R to calculate probabilities and quantiles for random variables.

(a) For $X$ a negative binomial with $n = 4$ and $p = 1/3$, find $P\{X = x\}$ for $x = 0, 1, \ldots, 12$.

(b) For $X$ a gamma random variable with $\alpha = 5$ and $\beta = 3$, find $P\{X \leq x\}$ for $x = 0, 1, 2, 3$. Indicate these values on a plot of the cumulative distribution function of $X$.

(c) For $Z$ a standard normal, find values for $z$ so that $P\{Z \leq z\} = 0.05, 0.25, 0.50, 0.75, 0.95$ Indicate these values on a plot of the cumulative distribution function.

(d) Compare the probability mass functions of a $Bin(350, 1/100)$, and $Bin(3500, 1/1000)$ and $Pois(3.5)$ random variable by displaying them in a data.frame.

7. In this problem, we shall use R to simulate with random variables.

(a) Simulate 1000 independent beta random variables with $\alpha = 3$ and $\beta = 5$. Find the mean and variance of this sample and compare it to the actual values.

(b) Use rhyper to simulate 1000 times the drawing 10 marbles out of an urn containing 16 black and 12 white marbles. Compute the mean and standard deviation of the number of white marbles and compare it to the distributional answers.

(c) Repeat the exercise above using the appropriate binomial random variable to simulate sampling with replacement. Compute the mean and standard deviation of the number of white marbles and compare it to the distributional answers.

(d) For a random variable $X$ having the standard logistic distribution, the distribution is

$$F_X(x) = \frac{1}{1 + \exp(-x)} \quad x \in \mathbb{R}.$$

Use the probability transform to create 100 samples of the logistic random. Display the empirical cumulative distribution function and compare it to the actual distribution.