1. A lake contains 400 lake trout. 48 are captured, tagged, and released. A certain time later, 10 of the
400 are captured.

(a) What is the probability that exactly three 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) Simulate the mark and capture process 1000 times and compare the simulation data to the mass
function in part (b).

(d) What is the probability that a fish is tagged? Use this to find the mean number of tagged fish in
the second capture.

2. Duchenne muscular dystrophy is a sex linked recessive trait on the X chromosome. About 1 out of
every 3,600 X chromosomes has the variant allele for the protein dystrophin that causes Duchenne
muscular dystrophy.

(a) A male inherits Duchenne muscular dystrophy if he receives the variant allele. What fraction of
the male population have hemophilia?

(b) A female has Duchenne muscular dystrophy if she receives the variant allele from both of her
parents. Assuming that the probability of receiving the allele is 1/3600 independently from each
of her parents, what fraction of the female population has Duchenne muscular dystrophy?

(c) Assuming that the population is made up 1/2 men and 1/2 women, find the fraction of the
population that have hemophilia.

(d) If a person chosen at random is a Duchenne muscular dystrophy, find the probability that the
person is female.

3. Let \( X \) be a discrete random variable with probability mass function

\[
\begin{array}{c|cccccc}
 x & 1 & 2 & 3 & 4 & 5 & 6 \\
 f_X(x) & c & 2c & 0 & c & 2c & 0 \\
\end{array}
\]

(a) Find \( c \).

(b) Find \( P\{X > 4\} \)
(c) Draw the cumulative distribution for $X$.
(d) Find $EX$, and $\text{Var}(X)$.
(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

$$F_X(x) = \begin{cases} 
0, & \text{for } x \leq 0 \\
1 - \frac{1}{2}(e^{-x} + e^{-3x}), & \text{for } x > 0.
\end{cases}$$

(a) Give a plot of $F_X$ and explain using this plot why $F_X$ is a valid distribution function.
(b) Find the probabilities $P\{X \geq 2\}, P\{X \leq 4\}$, and $P\{2 < X \leq 4\}$.
(c) Find the probability density for this distribution function.
(d) Sketch the distribution and density function indicating $P\{2 < X \leq 4\}$.

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 $\alpha = 1$ 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 $n = 3$ and $p = 3/4$, find $P\{X = x\}$ for $x = 0, 1, \ldots, 10$.
(b) For $X$ a gamma random variable with $\alpha = 4$ and $\beta = 2$, 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 $\text{Bin}(50, 1/50)$, and $\text{Bin}(500, 1/500)$ and $\text{Pois}(1)$ 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 $\alpha = 2$ and $\beta = 7$. 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 10 black and 15 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) Compare the means and standard deviations for parts (b) and (d) and explain in words if they differ and why.

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

$$F_X(x) = \frac{1}{1 + e^{-x}}$$

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.