Introduction to probability with MATLAB, Spring 2014
University of Helsinki, Department of Mathematics and Statistics
Exercise set 8

1. Let A and B be two events such that $0 < P(B) < 1$, i.e. both B and its complement are possible. **Prove:** If any two of the three probabilities $P(A)$, $P(A \mid B)$ and $P(A \mid B^c)$ are equal, then the third one is also equal.

Hint: Start from the law of total probability. Assume that two of the probabilities are equal, and solve the third one.

2. X and Y are independent, and each has uniform distribution over the unit interval $(0, 1)$. Let $Z = X + Y$. It is known that Z has the following density function:

$$f(z) = \begin{cases} z, & \text{if } 0 < z \leq 1, \\ 2 - z, & \text{if } 1 < z < 2, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Calculate $E(Z)$ by using the linearity of expectation.

(b) Calculate $E(Z)$ from the density function.

(c) Calculate $P(0.5 < Z < 1.5)$.

(d) Find the cumulative distribution function of Z.

3. The random variable X has the *Laplace distribution* (or *two-sided exponential distribution*), meaning that it has density

$$f(x) = \frac{1}{2}\exp(-|X|)$$

everywhere on the real line.

(a) Check by integration that this is indeed a density function.

(b) Calculate $E(Z)$.

(c) Calculate $P(|X| > 1)$.

(d) Try to think of a way to simulate (generate) random numbers from this distribution. Once you have a method, simulate a large number of them, plot a histogram (`hist`), and see if it has the same shape as the density function as defined above.

4. The random variable θ has uniform distribution over $(0, 2\pi)$. Two other random variables are then defined as $X = \cos\theta$ and $Y = \sin\theta$. In this exercise we study their distribution experimentally.

(a) Simulate 100 values of θ, compute the corresponding values of X and Y, and plot them as points in the 2D plane, for example `plot(X,Y,'.')`. Where do the points seem to be and how are they distributed there?

Hint: The functions `cos` and `sin` can be applied to vectors as well.

(b) Simulate 1000 (or more) values of θ, then plot a histogram of the values of X. Does X seem to have a uniform distribution? Why / why not?

5. Previous problem continued. Now we will solve the distribution of X exactly.

(a) Let x be any real number within the interval $(-1, 1)$. Which values of θ fulfill the condition that $X \leq x$? (Hint: Think about the graph of the cosine function.)

(b) What is the probability that $X \leq x$? (Hint: What is the probability that θ is such that the condition holds?) In other words, find the cdf of X.

(c) Solve the density function of X. Hint: the derivative of arccosine is

$$\frac{d}{dx} \arccos x = -\left(\frac{1}{\sqrt{1-x^2}}\right).$$

(d) Plot the density function and compare to the histogram.

6. Previous continued. It was possible but not quite easy to solve the exact distribution of $X = \cos\theta$. Suppose, however, that we only need to know $E(X)$. Calculate it by using Theorem 6.11 (G&S page 270).

7. In this problem we try to study an empirical histogram (which is random by its nature) theoretically.

Let $X_1, \ldots, X_{99}$ be independent and identically distributed, each having the uniform distribution over $(0, 1)$. Suppose that a very coarse histogram is drawn with only three bars:

- First bar has height N_1, which is the count of the values X_i that are within $(0, 1/3]$.
- Second bar has height N_2, which is the count of the values X_i that are within $(1/3, 2/3]$.
- Third bar has height N_3, which is the count of the values X_i that are within $(2/3, 1)$.

What is the distribution of N_1, and what is its expected value $\mu = E(X)$? What is the probability that $N_1 = \mu$? What about N_2 and N_3?

8. Previous continued. What is the probability that the histogram is completely "uniform", in the sense that $N_1 = N_2 = N_3$? Hint: Multinomial distribution.

9. Are the random variables N_1, N_2, N_3 independent?

10. (Monte Carlo integration) There is a strange shape S in the plane. We will try to find its area a.

From the auxiliary functions (link at course page) you will find `strangeshape.m`, which tests for a given point (x, y) whether it is inside shape S or not. If inside, the call `strangeshape(x,y)` returns 1, otherwise 0. (x and y may also be vectors.)

We know that the shape S is somewhere inside the square $(-1, 1)^2$, that is, the square that has side 2 and center at the origin. Let b be the area of this square.

Simulate $n = 1\,000\,000$ points (x, y) randomly within the square. Count how many of them are inside shape S, and denote this by k.

Since each point in the square has probability a/b of being inside the shape, we expect that $k/n \approx a/b$. Compute now the approximation $a \approx bk/n$.

11. Of the points you generated, plot the inside points in red and outside points in blue. Hint: If `F` is a vector of zeros and ones, then `x(F)` picks only those elements of `x` where `F` is 1. For changing the color see `help plot`. You may also want to use `hold on`.