Exercise 5

Applied Statistics, IT University of Copenhagen

T=Theoretical exercise, R=R exercise

Preparation

- Read pages 33–44, 50–55, and 60–61 from Verzani (2014).

Problems

1. Octrahedral Die (T)

Let T be random variable characterising the outcome of rolling a fair, octahedral die.

(a) Write down the probability mass function of T.

(b) Determine the expected value and variance of T.

2. Expected Value and Variance (T)

Let $X \sim U(1, 2)$.

(a) Compute the expected value of X.

(b) Compute the variance of X.

(c) Let $f(x) = e^x$. Compute the expected value of $f(X)$.

3. Transforming a Random Variable (T)

Given is a random variable X with the probability density function f given by $f(x) = 4x - 4x^3$ for $0 \leq x \leq 1$, and $f(x) = 0$, otherwise.

(a) Determine the distribution function F_X.

(b) Let $Y = \sqrt{X}$. Determine the distribution function F_Y.

(c) Determine the probability density function of Y.

4. Accessing Data and Numeric Summaries (R)

(a) Take `Cars93` (MASS) data set. What is the type of the Cylinders variable? What does the summary command do for the Cylinders variable? Get the names of the cars having 8 cylinders. What is the mean horsepower of the cars having 8 cylinders, how about standard deviation? How about those for the cars having 6 cylinders? Is the result what you expect?

(b) For the `precip` data set, find the mean and standard deviation of the rain fall over cities. Find all the cities with the average annual rain fall exceeding 50 inches. Which cities are the dryest? Does this match your expectation?

(c) The `rivers` contains the lengths of the 141 major rivers in North America. Compare the mean and 25% trimmed mean on the data set. What does the result tell you? How big is the standard deviation?

5. Simulation (R)

Let X be a continuous random variable with the probability density function

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

a) Derive the distribution function of $f(x)$.

b) Describe the principle how you can simulate the random variables following the probability density $f(x)$ when you have an access to a random number generator of uniformly distributed random numbers.

c) Write a program that uses stochastic simulation to draw 100 independent samples from $f(x)$.

Verzani, John. 2014. *Using R for Introductory Statistics.* CRC Press.