

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 the outcome of roll of fair octahedral die.

- (a) Describe the probability distribution of T , that is, list the outcomes and the corresponding probabilities.
- (b) Determine the expected value and variance of T .

2. Expectation and Variance of a Continuous Random Variable (T)

Let X be a continuous random variable with the density function

$$f_X(x) = \begin{cases} x+1 & \text{if } -1 \leq x < 0 \\ -x+1 & \text{if } 0 \leq x < 1 \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Compute the expectation and variance of X .

3. Linearity of the Expectation Operator (T)

Show that the expectation operator is linear; that is, for functions $f, g : \mathbb{R} \rightarrow \mathbb{R}$, applied on the random variable X , and any scalars $\alpha, \beta \in \mathbb{R}$,

$$\mathbb{E}[\alpha f(X) + \beta g(X)] = \alpha \mathbb{E}[f(X)] + \beta \mathbb{E}[g(X)]. \quad (2)$$

Consider the cases where

- (a) X is a discrete random variable taking values $a_1, a_2, \dots \in \mathbb{R}$,
- (b) X is a continuous random variable taking values on the real axis.

4. Transforming a Random Variable (T)

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

- (a) Determine the distribution function F_X .
- (b) Let $Y = \sqrt{X}$. Determine the distribution function F_Y .
- (c) Determine the probability density of Y .

5. 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?

6. Flight Overbooking (R)

To maximise the seats occupied during flights, the airlines has the customs to overbook them. Assume that the total number of seats on a flight is 150 and the number of people showing up at the airport is a random variable $X \in 1, 2, \dots, M$, where all the outcomes are equally probable, and M is the number of bookings made. Assume that each passenger onboard means 500 EUR cash inflow for the airline whereas each refused passenger implies 1000 EUR penalty to the airline. Operating the plane costs 40000 EUR. For how many bookings would you advice the airline to take?

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