Dirk Metzler SS 2021

# STATISTICS FOR EES — EXERCISE SHEET 3

**1.** You have two dice that look exactly the same, but one is fair and the other one is loaded and gives a six in 1/3 of the cases (and other results with a probability of 2/15 each). You forgot which of the two dice is which, so you roll each dice five times. The first dice shows a six only once and with the second dice you obtain two sixes. Given this result, calculate the probability that the first dice is fair.

**2.** The R script sem_cll.R contains commands to perform the following steps:

(a) Simulates a population of 1,000,000 values of some variable $x$.

(b) Compute the mean $\mu$ and the standard deviation $\sigma$ of all values $x$.

(c) Draw 1,000 samples of size $n = 10$ from the population of values $x$.

(d) Compute the sample mean $\overline{x}$ and the standard deviation $s$ for each of the 1,000 samples.

(e) Determine for what fraction of the samples the interval between $\overline{x} - s/\sqrt{n}$ and $\overline{x} + s/\sqrt{n}$ contained the population mean $\mu$.

(f) Visually compare the distribution of sample means to the normal distribution with mean $\mu$ and standard deviation $\sigma/\sqrt{n}$.

At the end of the R script another population of values $y$ is simulated. Perform the steps listed above also for the population $y$ with various values for the sample size $n$. For which $n$ is the normal distribution a good approximation for the distribution of sample means?

**3.** Let $X$ be a random variable that is uniformly distributed on $[0, 1]$ and be $f : [0, 1] \to \mathbb{R}$ defined by $f(x) = 10 \cdot (x - 0.25)^2 - 0.2$. Compute $\Pr(f(X) \in [0, 1])$.

**4.** A plant of a certain species reaches on average a height of 123 cm after four months. The height, however, also depends in the amount of sun ($S$) and the amount of rain ($R$) that the plant will get during this time. $S$ and $R$ (which are measured in certain units) vary from year to year, and their standard deviations are $\sigma_S = 3.6$ and $\sigma_R = 4.7$. Furthermore, sunny years tend to be less rainy, which can be quantified by $\text{Cor}(R, S) = -0.1$. Each unit of sun increases the plant height (after four months) by 1.2 cm, and each unit of rain increases the height by 4.2 cm. Additional genetical and environmental factors, which are not correlated to $R$ and $S$, adds a random increase or decrease $F$ to plant height, with $\sigma_F = 5.7$. Calculate the total standard deviation of plant height after four months.

**5.** Let $X$ and $Y$ be random variables with values in $\{1, 2, 3\}$ and $\{0, 1\}$ and

$$\begin{aligned} \Pr(X = 1, Y = 0) &= \frac{1}{3} \qquad \Pr(X = 1, Y = 1) = 0 \\ \Pr(X = 2, Y = 0) &= \frac{1}{4} = \Pr(X = 2, Y = 1) \\ \Pr(X = 3, Y = 0) &= \frac{1}{12} = \Pr(X = 3, Y = 1). \end{aligned}$$

Compute

(a) $\Pr(Y = 0)$ and $\Pr(Y = 1)$

(b) $\mathbb{E}X$ and $\mathbb{E}Y$

(c) $\mathbb{E}(X^2)$ and $\mathbb{E}(Y^2)$

(d) $\mathrm{Var}(X)$ and $\mathrm{Var}(Y)$

(e) $\mathrm{Cov}(X, Y)$

(f) $\mathrm{Cor}(X, Y)$

**6.** (This exercise is a bit more advanced.) Let $X$ be a random variable that is uniformly distributed on $[0, 1]$.

(a) Calculate the probability that an $n \in \{1, 2, 3, \ldots\}$ exists such that

$$X \in \left[\frac{2}{3^{n+1}}, \frac{1}{3^n}\right],$$

that is, expressed less formally, the probability that $X$ takes a value in the union set of the infinitely many intervals

$$\ldots \cup \left[\frac{2}{243}, \frac{1}{81}\right] \cup \left[\frac{2}{81}, \frac{1}{27}\right] \cup \left[\frac{2}{27}, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right].$$

(Hint: in the end you may come to a point where the geometric series formula can be applied.)

(b) Let

$$\begin{aligned}
\mathbb{Q}_+ &= \left\{\frac{1}{1}, \frac{2}{1}, \frac{1}{2}, \frac{3}{1}, \frac{2}{2}, \frac{1}{3}, \frac{4}{1}, \frac{3}{2}, \frac{2}{3}, \frac{1}{4}, \frac{5}{1}, \frac{4}{2}, \frac{3}{3}, \frac{2}{4}, \frac{1}{5}, \ldots\right\} \\
&= \left\{\frac{1}{1}, \frac{2}{1}, \frac{1}{2}, \frac{3}{1}, \quad \frac{1}{3}, \frac{4}{1}, \frac{3}{2}, \frac{2}{3}, \frac{1}{4}, \frac{5}{1}, \qquad \frac{1}{5}, \ldots\right\} \\
&= \{1, 2, 1/2, 3, 1/3, 4, 3/2, 2/3, 1/4, 5, 1/5, \ldots\}
\end{aligned}$$

be the set of all positive rational numbers, that is, the numbers that can be represented as a fraction of two positive integers. Calculate $\Pr(X \in \mathbb{Q}_+)$.