Homework Assignment 9: Applied Probabilistic Models

Expected Value and Variance

5273

1 Exercises

Exercises solved in this work are provided on the book Grinstead and Snell [1].

1.1 Expected Value of Discrete Random Variables

For this exercises the expected value and the variance of discrete random variables are discussed.

Exercise 1 page 247

As the deck consists of cards between 2 through 10, it will have 36 cards. Let X be the number of the selected card. The player will win a dollar if the number of the card is odd and loses one dollar if the number is even, so the expected value of his winnings will be:

$$E(X) = -1\left(\frac{4}{36}\right)+1\left(\frac{4}{36}\right)-1\left(\frac{4}{36}\right)+1\left(\frac{4}{36}\right)-1\left(\frac{4}{36}\right)+1\left(\frac{4}{36}\right)-1\left(\frac{4}{36}\right)+1\left(\frac{4}{36}\right)-1\left(\frac{4}{36}\right) = -\frac{1}{9}.$$

Exercise 15 page 249

In this exercise, the game stops whenever it is one dollar profit or run out of gold balls. For one gold ball, a player wins one dollar and loses one dollar if a silver ball is drawn as the box contains two gold balls and three silver balls and let X be the results of the draws until the game is finished. There are seven possible outcomes, which are detailed below with its corresponding probability. Therefore, the expected value is:

$$E(X) = 1\left(\frac{2}{5}\right) + 1\left(\frac{1}{10}\right) + 0\left(\frac{1}{10}\right) + 0\left(\frac{1}{10}\right) - 1\left(\frac{1}{10}\right) - 1\left(\frac{1}{10}\right) - 1\left(\frac{1}{10}\right) = \frac{2}{10} = \frac{1}{5}.$$

Because $E(X) > 0$ it can be said this is a favorable game.

Exercise 18 page 249

Six similar keys are given, and let X be the number of tried keys before the success of opening the door.

$$E(X) = 0\left(\frac{1}{6}\right) + 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + 3\left(\frac{1}{6}\right) + 4\left(\frac{1}{6}\right) + 5\left(\frac{1}{6}\right) = \frac{5}{2}.$$

Table 1: Possible outcomes for Exercise 15 page 249.

Outcome	Probability	Profit
G	$\left(\frac{2}{5}\right)$	1
SGG	$\left(\frac{3}{5}\right)\left(\frac{1}{2}\right)\left(\frac{1}{3}\right)$	1
SGSG	$\left(\frac{3}{5}\right)\left(\frac{1}{2}\right)\left(\frac{2}{3}\right)\left(\frac{1}{2}\right)$	0
SSGG	$\left(\frac{3}{5}\right)\left(\frac{1}{2}\right)\left(\frac{2}{3}\right)\left(\frac{1}{2}\right)$	0
SSSGG	$\left(\frac{3}{5}\right)\left(\frac{1}{2}\right)\left(\frac{1}{3}\right)(1)(1)$	-1
SGSSG	$\left(\frac{3}{5}\right)\left(\frac{1}{2}\right)\left(\frac{2}{3}\right)\left(\frac{1}{2}\right)$	-1
SSGSG	$\left(\frac{3}{5}\right)\left(\frac{1}{2}\right)\left(\frac{2}{3}\right)\left(\frac{1}{2}\right)$	-1

Exercise 19 page 249

For every correct answer, the student gets three points and for every incorrect one loses one point. The problem has four possible answers. Let X be the result if the answer is correct or incorrect. Therefore, the probability of choosing the correct answer just guessing is 0.25, whereas the probability of choosing the incorrect one is 0.75. The expected value is:

$$E(X) = 3(0.25) - 1(0.75) = 0.$$

Exercise 31 page 254

(a) For a pooled sample of k people, if the test is positive, it means that each person has a positive result on the test independently with probability p. Therefore, the probability that each person has a negative result is $1 - p$. The probability that all of the k subjects have negative result is $(1 - p)^k$. Consequently:

$$\begin{aligned} P(\text{sample is positive}) &= 1 - P(\text{all } k \text{ subjects have negative results}), \\ &= 1 - (1 - p)^k. \end{aligned}$$

(b) Let X be the number of blood tests necessary under the plan (2). There are $\dfrac{N}{k}$ groups of k individuals. For each of these groups, if someone is positive, a $k + 1$ tests are needed, and otherwise, only one test. The expected value for the number of tests for one group is:

$$(k+1)P(Positive) + 1 * P(Negative) = (k+1)(1 - (1-p)^k) + 1(1-p)^k,$$

For the whole group of N subjects is:

$$\begin{aligned} E(X) &= \frac{N}{k}\left[(k+1)(1-(1-p)^k) + 1(1-p)^k\right] \\ &= \frac{N}{k}\left[(k+1) - (k+1)(1-p)^k + (1-p)^k\right] \\ &= \frac{N}{k}\left[k + 1 - k(1-p)^k - (1-p)^k + (1-p)^k\right] \\ &= \frac{N}{k}\left[k + 1 - k(1-p)^k\right]. \end{aligned}$$

(c) To minimize the expected value, it can be calculated the derivate of with respect to k and set equal to 0.

$$E(k) = \frac{N}{k}\left[k + 1 - k(1-p)^k\right],$$

$$\begin{aligned}\frac{\mathrm{d}E}{\mathrm{d}k} &= 0\\ \frac{-n\ln(1-p)(1-p)^k k^2 + N}{k^2} &= 0,\end{aligned}$$

If in the expression above, a sufficient small p is considered approximately 0, the equality would be:

$$\begin{aligned}\frac{\cancel{-n\ln(1-p)(1-p)^k k^2} + N}{k^2} &= 0\\ \frac{N}{k^2} &= 0,\end{aligned}$$

When substituting the value of k to $\dfrac{1}{\sqrt{p}}$:

$$\begin{aligned}\frac{N}{k^2} &= 0\\ \frac{N}{\left(\frac{1}{\sqrt{p}}\right)^2} &= 0\\ \frac{N}{\left(\frac{1}{p}\right)} &= 0\\ Np &= 0.\end{aligned}$$

At this point, if the value of p is sufficient small (approximately 0), the equality is fullfilled for a value of $k = \dfrac{1}{\sqrt{p}}$, where the expected number of test is minimum.

Exercise 1 page 263

If $\{S = -1, 0, 1\}$:

- The expected value would be:

$$E(X) = \frac{\sum(x)}{N} = -1\left(\frac{1}{3}\right) + 0\left(\frac{1}{3}\right) + 1\left(\frac{1}{3}\right) = 0.$$

- To calculate the variance it can be used Table 2. Then, the variance would be:

$$\sigma^2 = \frac{\sum(x-\bar{x})^2}{N} = \frac{(-1-0)^2 + (0-0)^2 + (1-0)^2}{3} = \frac{2}{3}.$$

- The standard deviation $\sigma = \sqrt{\sigma^2} = \sqrt{\frac{2}{3}} \approx 0.816.$

Table 2: Variance in Exercise 1 page 263.

x	$x-\bar{x}$	$(x-\bar{x})^2$
-1	-1	1
0	0	0
1	1	1
$\sum x_i = 0$		$\sum (x_i-\bar{x})^2 = 2$

1.2 Expected Value of Continuous Random Variables

This section corresponds to exercises of expected value and variance of continuous random variables.

Exercise 3 page 278

The expected value of a continuous random variable is given by

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx,$$

Therefore the expected lifetime of the light bulb would be:

$$\begin{aligned}
E(T) &= \int_0^{\infty} t(\lambda)^2 t e^{-\lambda t} dt \\
\lambda^2 \int t^2 e^{-\lambda t} dt &= \lambda^2 \left(\frac{-t^2 e^{-\lambda t}}{\lambda} - \int \frac{-2te^{-\lambda t}}{\lambda} dt \right) \\
&= \lambda^2 \left[\frac{-t^2 e^{-\lambda t}}{\lambda} + \frac{2}{\lambda} \left(-\frac{te^{-\lambda t}}{\lambda} - \int \frac{-e^{-\lambda t}}{\lambda} dt \right) \right] \\
&= \lambda^2 \left[\frac{-t^2 e^{-\lambda t}}{\lambda} + \frac{2}{\lambda} \left(-\frac{te^{-\lambda t}}{\lambda} + \frac{1}{\lambda} \left[\frac{-e^{-\lambda t}}{\lambda} \right] \right) \right] \\
&= \left[\frac{\left(-t^2\lambda^2 - 2t\lambda - 2\right) e^{-\lambda t}}{\lambda} \right]_0^{\infty} \\
&= \frac{2}{\lambda} \\
&= 40.
\end{aligned}$$

The variance would be:

$$\begin{aligned}
\lambda^2 \int t^3 e^{-\lambda t} dt &= \lambda^2 \left(\frac{-t^3 e^{-\lambda t}}{\lambda} - \int \frac{-3t^2 e^{-\lambda t}}{\lambda} dt \right) \\
&= \lambda^2 \left[\frac{-t^3 e^{-\lambda t}}{\lambda} + \frac{3}{\lambda} \left(-\frac{t^2 e^{-\lambda t}}{\lambda} - \frac{2te^{-\lambda t}}{\lambda^2} - \frac{2e^{\lambda t}}{\lambda^3} \right) \right] \\
&= \left[\frac{\left(-t^3\lambda^3 - 3t^2\lambda^2 - 6t\lambda - 6\right) e^{-\lambda t}}{\lambda^2} \right]_0^{\infty} \\
&= \frac{6}{\lambda^2} \\
&= 2400,
\end{aligned}$$

$$
\begin{aligned}
V(T) &= E(X^2 - E(X)^2) \\
&= 2400 - 1600 \\
&= 800.
\end{aligned}
$$

Exercise 12 page 280

The variables X and Y are independent, and both are uniformly distributed on $[0,1]$. The expected value of both variables is given by:

$$
\begin{aligned}
E(X^Y) &= \int_0^1 \int_0^1 x^y f(x) f(y) dx dy \\
&= \int_0^1 \left[\frac{x^{y+1}}{y+1}\right]_0^1 dy \\
&= \int_0^1 \frac{1}{y+1} dy \\
&= [\ln(y+1)]_0^1 \\
&= \ln 2 \\
&\approx 0.6931.
\end{aligned}
$$

Results of the simulation are shown in Figure 1. It is performed in R software in its version 4.0.2 [3], and the code used is available on the GitHub repository of [2].

Exercise 28 page 284

The length of the needle is L (much bigger than 1). If it is dropped on a grid with a horizontal line, it will form an angle θ with intersections is $L\cos\theta$ and with a vertical line with intersections is $L\sin\theta$. The uniform probability density function of θ between 0 and $\frac{\pi}{2}$ is:

$$
\begin{cases}
\frac{2}{\pi} & \text{if } 0 \leq \theta \leq \frac{\pi}{2}, \\
0 & \text{elsewhere.}
\end{cases}
$$

Therefore the average number of lines crossed approximately is:

$$
\begin{aligned}
a &= \frac{2}{\pi} \int_0^{\frac{\pi}{2}} L\left(cos\theta + sin\theta\right] d\theta \\
&= \frac{2L}{\pi} \left[sin\theta - cos\theta\right]_0^{\frac{\pi}{2}} \\
&= \frac{4L}{\pi}.
\end{aligned}
$$

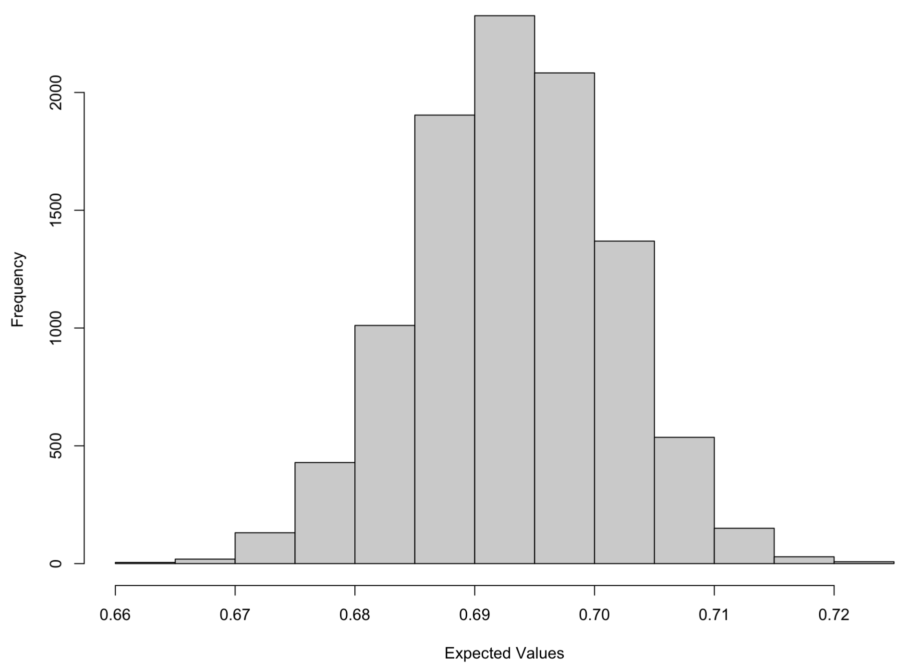

Figure 1: Histogram of the simulated expected values in Exercise 12 page 280

References

[1] Charles Miller Grinstead and James Laurie Snell. *Introduction to probability*. American Mathematical Soc., 2012.

[2] Oscar Alejandro Hernadez Lopez. Probability in R. `https://github.com/oscaralejandro1907/probability-in-R/blob/master/assignment1/t1.R`, 2020.

[3] The R Foundation. The R Project for Statistical Computing. `https://www.r-project.org/`, 2020.