EXERCISE 2: There are two bushes. The length (in cm) of the leaves of bush A is a RV whose PDF is:

$$f(x)=\begin{cases} -\frac{k}{256}\cdot x^2+\frac{k}{32}\cdot x & 0\le x\le 8 \\ 0 & otherwise \end{cases}$$

The length of the leaves of bush B is a RV whose PDF is:

$$g(x)=\begin{cases} \frac{1}{10} & 0\le x\le 10 \\ 0 & otherwise \end{cases}$$

1) Find k
2) Compute the mean and variance of $f(\ )$ and $g(\ )$

Assume now that we pluck a leaf from one of the bushes at random.

3) What is the probability that it is from bush B, given that it is more than 4cm long?

Assume now that we take a sample of 10 leaves from another bush (which is neither A nor B). The length of the leaves is: 5.4, 6.3, 5.8, 7.2, 4.9, 9.2, 7.0, 7.3, 6.9, 10.4

4) Compute the 95% confidence interval for the sample mean
5) Assume that every leaf is
   a. 5cm longer than before.
   b. Double as long as before

   What about the sample mean and the confidence interval?

ESERCIZIO 2: Ci sono due siepi. La lunghezza in cm delle foglie della siepe A è una VA la cui PDF è:

$$f(x) = \begin{cases} -\frac{k}{256} \cdot x^2 + \frac{k}{32} \cdot x & 0 \leq x \leq 8 \\ 0 & otherwise \end{cases}$$

La lunghezza delle foglie della siepe B è una VA la cui PDF è:

$$g(x) = \begin{cases} \frac{1}{10} & 0 \leq x \leq 10 \\ 0 & otherwise \end{cases}$$

1) Calcolare k
2) Calcolare media e varianza di $f(\ )$ e $g(\ )$

Si assuma adesso che venga presa una foglia a caso da una delle due siepi.

3) Qual è la probabilità che sia della siepe B, posto che è più lunga di 4 cm?

Si assuma adesso di estrarre un campione di 10 foglie da una terza siepe. La lunghezza delle foglie è: 5.4, 6.3, 5.8, 7.2, 4.9, 9.2, 7.0, 7.3, 6.9, 10.4

4) Si calcoli l'intervallo di confidenza al 95% per la media campionaria.
5) Si assuma che ogni foglia sia:
   a) 5cm più lunga rispetto a prima.
   b) Lunga il doppio rispetto a prima

Cosa succeed alla media campionaria ed all'intervallo di confidenza

Solution

1) *k* can be computed based on the normalization condition:

$$\int_0^8 \left(-\frac{k}{256}\cdot x^2+\frac{k}{32}\cdot x\right)dx=$$

$$\frac{k}{32}\cdot\int_0^8 \left(-\frac{1}{8}\cdot x^2+x\right)dx=$$

$$\frac{k}{64}\cdot\left[-\frac{1}{12}\cdot x^3+x^2\right]_0^8=$$

$$\frac{k}{64}\cdot\left[-\frac{512}{12}+64\right]=$$

$$\frac{k}{64}\cdot\left[\frac{-128+192}{3}\right]=$$

$$\frac{k}{3}$$

Since, by normalization, the integral must be equal to 1, it is $k=3$

2) $g(\ )$ is a uniform RV, hence its mean is $\mu_g=\frac{1}{b-a}=5$, and its variance is $\sigma^2=\frac{(b-a)^2}{12}=\frac{25}{3}$.

On the other hand, the mean and variance of $f(\ )$ are computed by solving these integrals:

$$\mu_f=\int_0^8 x\cdot 3\cdot\left(-\frac{1}{256}\cdot x^2+\frac{1}{32}\cdot x\right)dx=$$

$$=\frac{3}{32}\cdot\int_0^8\left(-\frac{1}{8}\cdot x^3+x^2\right)dx$$

$$=\frac{3}{32}\cdot\left[-\frac{1}{32}\cdot x^4+\frac{1}{3}\cdot x^3\right]_0^8$$

$$=\frac{\cancel{3}}{32}\cdot\left[\frac{-3\cdot 4096+32\cdot 512}{\cancel{3}\cdot 32}\right]$$

$$=\frac{1}{32}\cdot\frac{4096}{32}=4$$

$$\overline{X_f^2}=\int_0^8 x^2\cdot 3\cdot\left(-\frac{1}{256}\cdot x^2+\frac{1}{32}\cdot x\right)dx=$$

$$=\frac{3}{32}\cdot\int_0^8\left(-\frac{1}{8}\cdot x^4+x^3\right)dx$$

$$=\frac{3}{32}\cdot\left[-\frac{1}{40}\cdot x^5+\frac{1}{4}\cdot x^4\right]_0^8$$

$$=\frac{3}{128}\cdot\left[\frac{-32768+40960}{10}\right]$$

$$=\frac{3}{128}\cdot\frac{8192}{10}=3\cdot\frac{64}{10}=\frac{96}{5}$$

Hence, $\sigma_f^2 = \frac{96}{5} - (4)^2 = \frac{96 - 16 \cdot 5}{5} = \frac{16}{5}$

3) Call L the event "the leaf is more than 4 cm". By Bayes' Theorem, we have:

$$P(B|L) = \frac{P(L|B) \cdot P(B)}{P(L|B) \cdot P(B) + P(L|A) \cdot P(A)}.$$

However, $P(B) = P(A) = 1/2$, since the choice is "at random". Furthermore, $P(L|B) = (10-4)/10 = 0.6$, and:

$$\begin{aligned}
P(L|A) &= \frac{3}{32} \cdot \int_4^8 \left(-\frac{1}{8} \cdot x^2 + x \right) dx \\
&= \frac{3}{64} \cdot \left[-\frac{1}{12} \cdot x^3 + x^2 \right]_4^8 \\
&= \frac{3}{64} \cdot \left[\left(-\frac{512}{12} + 64 \right) - \left(-\frac{64}{12} + 16 \right) \right] \\
&= \frac{3}{64} \cdot \left[48 - \frac{112}{3} \right] \\
&= \frac{3}{64} \cdot \frac{32}{3} = \frac{1}{2}
\end{aligned}$$

Hence:

$$P(B|L) = \frac{P(L|B)}{P(L|B) + P(L|A)} = \frac{0.6}{0.6 + 0.5} = 0.545$$

4) The sample mean is

$$\bar{X} = \frac{1}{n} \cdot \sum_i x_i = 7.04.$$

The sample variance is:

$$S^2 = \frac{1}{n-1} \cdot \sum_i \left(x_i - \bar{X} \right)^2 = \frac{25.424}{9} = 2.824.$$

From the tabulated Student-s t function, I need $t_{\alpha/2,n-1} = t_{0.025,9} = 2.262$. The semi-width of the confidence interval centered around $\bar{X}$ is therefore:

$$w = \frac{S}{\sqrt{n}} \cdot t_{\alpha/2,n-1} = \sqrt{\frac{2.824}{9}} \cdot 2.262 = 1.27.$$

The confidence interval is thus $I = [5.77; 8.31]$

5) if each leaf is 5 cm longer, the mean will be 5cm larger and the variance will stay the same, hence the confidence interval will have the same width, i.e. $I_a = I + 5 = [10.77; 13.31]$. If, on the other hand, the length is doubled, the mean will be doubled as well, and the standard deviation will be double as much. Hence, we will have $I_b = 2I = [11.54; 16.62]$.