Practice questions

1. Let X be an Exponential(λ) random variable. Find the PDF of the random variables (a) $Y = X^2$ and (b) $Z = e^{-\lambda X}$.

   Solution:

   (a) For $y \geq 0$, the CDF of Y is $F_Y(y) = \mathrm{P}(X^2 \leq y) = \mathrm{P}(X \leq \sqrt{y}) = 1 - e^{-\lambda\sqrt{y}}$. The PDF is the derivative of the CDF which is

   $$f_Y(y) = \begin{cases} \dfrac{\lambda}{2\sqrt{y}} e^{-\lambda\sqrt{y}} & \text{if } y > 0 \\ 0 & \text{otherwise} \end{cases}$$

   (b) Since X only takes nonnegative values, Z will take values between 0 and 1. For $0 < z \leq 1$, the CDF of Z is

   $$F_Z(z) = \mathrm{P}(e^{-\lambda X} \leq z) = \mathrm{P}\Big(X \geq -\frac{\log z}{\lambda}\Big) = e^{-\lambda(-\log z/\lambda)} = z.$$

   Its derivative is

   $$f_Z(z) = \begin{cases} 1 & \text{if } 0 < z \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

   In words, Z is a Uniform$(0, 1)$ random variable.

2. Raindrops hit the ground at a rate of 1 per second. An observatory has a raindrop sensing equipment. A signal is received by the computer with a maximum delay of 1 second after sensing a raindrop, with all delays equally likely. Find

   (a) The joint PDF of the time T of the first raindrop and the time S of the signal reception.

   (b) The marginal PDF of S.

   (c) The conditional PDF of T given S.

   Solution:

   (a) S is a Uniform$(T, T+1)$ random variable, where T is an Exponential(1) random variable. The joint PDF is $f_{S,T}(s, t) = f_T(t) f_{S|T}(s|t)$. We are given that

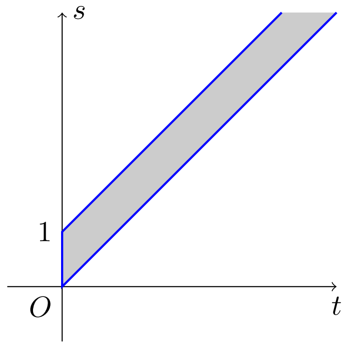
   $$f_T(t) = \begin{cases} e^{-t} & \text{if } t \geq 0 \\ 0 & \text{otherwise} \end{cases}, \qquad f_{S|T}(s|t) = \begin{cases} 1 & \text{if } t \leq s \leq t + 1 \\ 0 & \text{otherwise} \end{cases}$$

   Therefore $f_{S,T}(s, t) = f_T(t) f_{S|T}(s|t) = \begin{cases} e^{-t} & \text{if } t \geq 0,\ t \leq s \leq t + 1 \\ 0 & \text{otherwise} \end{cases}$

   Alternative Solution: We can write $S = T + D$ where the *delay* D is a Uniform$(0, 1)$ random variable that is independent of T. S and T take values s and t whenever T and D take values t and $s - t$, respectively, so by independence

   $$f_{S,T}(s, t) = f_T(t) f_D(s - t) = \begin{cases} e^{-t} & \text{if } t \geq 0,\ 0 \leq s - t \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

(b) The marginal PDF of S is $f_S(s) = \int_0^{+\infty} f_{S,T}(s,t)dt$. Let R be the region greyed out in the following plot in which $f_{S,T}(s,t)$ takes nonzero value e^{-t}.



Thus the integrand equals e^{-t} with bounds 0 to s when s is between 0 and 1, and s to $s+1$ when s is larger than 1. This gives

$$f_S(s) = \int_0^s e^{-t}dt = 1 - e^{-s}, \qquad \text{when } 0 \leq s \leq 1,$$
$$f_S(s) = \int_{s-1}^s e^{-t}dx = e^{-s}(e-1), \qquad \text{when } s > 1.$$

Alternative Solution: As T and D are independent we can calculate the PDF of S using the convolution formula:

$$\text{If } s \geq 1: \qquad f_S(s) = \int_{-\infty}^{+\infty} f_D(x) f_T(s-x)dx = \int_0^1 e^{-(s-x)}dx = e^{-s}(e-1),$$
$$\text{If } 0 \leq s \leq 1: \qquad f_S(s) = \int_{-\infty}^{+\infty} f_D(x) f_T(s-x)dx = \int_0^s e^{-(s-x)}dx = 1 - e^{-s}.$$

(c) The conditional PDF is given by

$$f_{T|S}(t|s) = \frac{f_{S,T}(s,t)}{f_S(s)} = \begin{cases} \frac{e^{-t}}{1-e^{-s}} & \text{if } t \geq 0,\ 0 < s \leq 1 \\ \frac{e^{-t}}{e^{-s}(e-1)} & \text{if } t \geq 0,\ s \geq 1 \\ 0 & \text{otherwise} \end{cases}$$

3. The body temperatures of a healthy person and an infected person are Normal(36.8, 0.5) and Normal(37.8, 1.0) random variables, respectively. About 1% of the population is infected. What is the conditional probability that I am infected given that my temperature is t? For which values of t am I more likely to be infected than not?

Solution: Let A be the event that I am infected, and T be my body temperature. By the total probability theorem,

$$f_T(t) = \mathrm{P}(A) f_{T|A}(x) + \mathrm{P}(A^c) f_{T|A^c}(t),$$

where $T|A$ is a Normal(37.8, 101) random variable and $T|A^c$ is a Normal(36.8, 0.5) random variable. The (unconditional) PDF of X is

$$f_T(t) = \frac{0.01}{\sqrt{2\pi}} e^{-\frac{(t-37.8)^2}{2}} + \frac{0.99}{\sqrt{2\pi}(0.5)} e^{-\frac{(t-36.8)^2}{2(0.5)^2}} = \frac{0.01}{\sqrt{2\pi}} e^{-(t-37.8)^2/2} + \frac{1.98}{\sqrt{2\pi}} e^{-2(t-36.8)^2}.$$

By Bayes' rule, the conditional probability of A given T is

$$P(A|T=t) = \frac{P(A) f_{T|A}(t)}{f_T(t)} = \frac{0.01 e^{-(t-37.8)^2/2}}{0.01 e^{-(t-37.8)^2/2} + 1.98 e^{-2(t-36.8)^2}}$$

I am more likely to be infected than not when $\mathrm{P}(A) > \mathrm{P}(A^c)$, namely when $0.01e^{-(t-37.8)^2/2} > 1.98e^{-2(t-36.8)^2}$. Taking logarithms of both sides this is equivalent to a quadratic inequality in t. Solving this inequality, we obtain that $\mathrm{P}(A) > \mathrm{P}(A^c)$ holds when $t < t_-$ or $t > t_+$, where $t_- \approx 34.4742$ and $t_+ \approx 38.4591$.

4. Raindrops hit your head at a rate of 1 per second. What is the PDF of the time at which the second raindrop hits you? How about the third one? (**Hint:** convolution)

Solution:

The time before the second raindrop is $Y = X_1 + X_2$, where X_1 and X_2 are independent Exponential(1) random variables. We calculate the PDF of Y using the convolution formula:

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X_1}(x_1) f_{X_2}(y - x_1) dx_1 = \int_0^y e^{-x_1} e^{-y+x_1} dx_1 = ye^{-y}.$$

The third raindrop hits at time $Z = Y + X_3$, where X_3 is another independent Exponential(1) random variable. By the convolution formula again,

$$f_Z(z) = \int_{-\infty}^{\infty} f_Y(y) f_{X_3}(z - y) dy = \int_0^z ye^{-y} e^{-z+y} dy = \frac{z^2}{2} e^{-z}.$$

5. You draw 10 balls at random among 15 red and 5 blue balls. Let X be the number of red balls drawn.

   (a) What is the expected value of X?

   (b) Write $X = X_1 + X_2 + \cdots + X_{10}$, where X_i indicates if the i-th drawn ball is red. What is the variance of X_i?

   (c) What is the covariance of X_i and X_j $(i \neq j)$?

   (d) What is the variance of X?

Solution:

(a) Let $X = X_1 + X_2 + \cdots + X_{10}$, where X_i indicates if the i-th drawn ball is red. By linearity of expectation,

$$\mathrm{E}[X] = \mathrm{E}[X_1] + \cdots + \mathrm{E}[X_{10}] = 10 \cdot \frac{15}{20} = 7.5.$$

(b) $\mathrm{Var}[X_i] = \mathrm{E}[X_i^2] - \mathrm{E}[X_i]^2 = \mathrm{P}(X_i = 1) - \mathrm{P}(X_i = 1)^2 = \frac{3}{4} - (\frac{3}{4})^2 = \frac{3}{16}$.

(c) $\mathrm{Cov}[X_i, X_j] = \mathrm{E}[X_i X_j] - \mathrm{E}[X_i]\,\mathrm{E}[X_j] = \mathrm{P}(X_i = 1, X_j = 1) - \mathrm{P}(X_i = 1)\,\mathrm{P}(X_j = 1) = \frac{15}{20} \cdot \frac{14}{19} - (\frac{15}{20})^2 = -\frac{3}{304}$. The variables X_i and X_j are negatively correlated: Given that ball i is red, ball j is less likely to be red.

(d) The variance of X is the sum of the 10 variances from part (b) plus the $10 \cdot 9$ covariances from part (c), so $\mathrm{Var}[X] = 10 \cdot \frac{3}{16} + 10 \cdot 9 \cdot \frac{-3}{304} = 0.9868$.