Expected Values

1. Find the mean and variance of the following PDFs:

   (a) $f_X(x) = \frac{5}{8}(1 - x^4)$, $-1 < x \leq 1$.
   Ans: The mean of X is

$$\begin{aligned} E[X] &= \frac{5}{8}\int_{-1}^{1} x - x^5 dx \\ &= \frac{5}{8}\left[\frac{x^2}{2} - \frac{x^6}{6}\right]_{-1}^{1} \\ &= 0. \end{aligned}$$

   (Actually, since $f_X(-x) = f_X(x)$, we know from symmetry that $E[X] = 0$.)
   The second moment of X is

$$\begin{aligned} E[X^2] &= \frac{5}{8}\int_{-1}^{1} x^2 - x^6 dx \\ &= \frac{5}{8}\left[\frac{x^3}{3} - \frac{x^7}{7}\right]_{-1}^{1} \\ &= \frac{5}{8}\cdot 2\left(\frac{1}{3} - \frac{1}{7}\right) \\ &= \frac{5}{21}. \end{aligned}$$

   Therefore, $\sigma_X^2 = \frac{5}{21} - 0 = \frac{5}{21}$.

   (b) $f_Y(y) = 6y(1 - y)$, $0 \leq y \leq 1$.
   Ans: This PDF is a quadratic function with $f_Y(0) = f_Y(1) = 0$, and hence is symmetric about $y = 0.5$. Therefore, $E[Y] = 0.5$. The second moment of Y is

$$\begin{aligned} E[Y^2] &= 6\int_{0}^{1} x^3 - x^4 dx && (1) \\ &= 6\left[\frac{x^4}{4} - \frac{x^5}{5}\right]_{0}^{1} && (2) \\ &= \frac{3}{10}. && (3) \end{aligned}$$

   Therefore, $\sigma_Y^2 = \frac{3}{10} - \frac{1}{4} = \frac{1}{20}$.

   (c) $f_Z(z) = 0.5\delta(z + 2) + 0.25[u(z) - u(z - 2)]$.

Ans: The mean value of Z is

$$\begin{aligned} E[Z] &= \int_{-\infty}^{\infty} t f_Z(t) dt & (4) \\ &= 0.5 \int_{-2-\epsilon}^{-2+\epsilon} t\delta(t+2)dt + 0.25 \int_0^2 tdt & (5) \\ &= -1 + 0.25 \times 2 = -0.5. & (6) \end{aligned}$$

The second moment is

$$\begin{aligned} E[Z^2] &= 0.5 \int_{-2-\epsilon}^{-2+\epsilon} t^2\delta(t+2)dt + 0.25 \int_0^2 t^2 dt & (7) \\ &= 2 + 0.25 \times \frac{8}{3} = 2\frac{2}{3} & (8) \end{aligned}$$

Therefore, the variance of Z is

$$\sigma_Z^2 = 2\frac{2}{3} - (-0.5)^2 = 2\frac{5}{12}. \qquad (9)$$

2. Let the function $g(x)$ be defined as

$$g(x) = \begin{cases} -a & x \le -a \\ x & -a < x \le a \\ a & x > a \end{cases}$$

(a) Sketch $g(x)$.
Ans: It is what is called a "clipping" function – any input with magnitude exceeding a will be turned into $\pm a$ (depending on the sign of the input). An input with magnitude smaller than a remains unaltered.

(b) Let $Y = g(X)$, where X is a random variable with the PDF

$$f_X(x) = \frac{1}{2} e^{-|x|}.$$

Find $E[Y]$ and $\text{var}(Y)$.
Ans: The mean of Y is given by

$$\begin{aligned} E[g(X)] &= \int_{-\infty}^{\infty} g(x) f_X(x) dx \\ &= 0 \end{aligned}$$

because $g(x)$ is an odd function and $f_X(x)$ is even. Therefore $g(x)f_X(x)$ is odd, which means that the total area under it is zero.

The second moment of Y is given by

$$
\begin{aligned}
E[Y^2] &= \int_{-\infty}^{\infty} g^2(x) f_X(x)dx \\
&= \int_{-\infty}^{-a} a^2 f_X(x)dx + \int_{-a}^{a} x^2 f_X(x)dx + \int_{a}^{\infty} a^2 f_X(x)dx \\
&= 2a^2 \int_{a}^{\infty} f_X(x)dx + \int_{-a}^{a} x^2 f_X(x)dx \qquad (10)
\end{aligned}
$$

where the last line comes from noting that $f_X(x)$ is even and hence $\int_{-\infty}^{-a} f_X(x)dx = \int_a^{\infty} f_X(x)dx$.

The second term in (10) needs to be simplified as follows:

$$
\begin{aligned}
\int_{-a}^{a} x^2 f_X(x)dx &= \int_{-a}^{0} x^2 \frac{1}{2} e^{x} dx + \int_{0}^{a} x^2 \frac{1}{2} e^{-x} dx \qquad (11) \\
&= \int_{0}^{a} x^2 e^{-x} dx. \qquad (12)
\end{aligned}
$$

Evaluation of the integral above requires integration by parts twice. Skipping the details, we eventually obtain

$$
\int_{0}^{a} x^2 e^{-x} dx = 1 - e^{-a}(a^2 + a + 1). \qquad (13)
$$

The first term in (10) is straightforward to simplify:

$$
2a^2 \int_{a}^{\infty} f_X(x)dx = 2a^2 \int_{a}^{\infty} \frac{1}{2} e^{-x} dx = a^2 e^{-a} \qquad (14)
$$

Therefore, $E[Y^2] = 1 - e^{-a}(a+1) = \sigma_Y^2$ since $E[Y] = 0$.

3. Let a and b be constants such that $a < b$, and let X have the PDF

$$
f_X(x) = \lambda e^{-\lambda(x-a)}, \quad x > a, \lambda > 0.
$$

Find the conditional PDF of X given $\{X \leq b\}$, and hence find $E[X|X \leq b]$.

Ans: We have $P[X \leq b] = \int_a^b \lambda e^{-\lambda(x-a)} dx = 1 - e^{-\lambda(b-a)}$. Therefore,

$$
\begin{aligned}
f_X(x|X \leq b) &= \frac{f_X(x)}{P[X \leq b]}, \quad x \leq b \qquad (15) \\
&= \frac{\lambda \exp[-\lambda(x-a)]}{1 - \exp[-\lambda(b-a)]}, \quad a < x \leq b. \qquad (16)
\end{aligned}
$$

To find $E[X|X \leq b]$, we need to evaluate (using integration by parts with $u = \lambda x$ and $dv = e^{-\lambda(x-a)}dx$)

$$
\begin{aligned}
\int_{a}^{b} \lambda x e^{-\lambda(x-a)} dx &= x e^{-\lambda(x-a)} \Big|_{b}^{a} + \int_{a}^{b} e^{-\lambda(x-a)} dx \qquad (17) \\
&= a + \frac{1}{\lambda} - \left(b + \frac{1}{\lambda}\right) e^{-\lambda(b-a)}. \qquad (18)
\end{aligned}
$$

Finally,

$$
\begin{aligned}
E[X|X \le b] &= \frac{\int_a^b \lambda x e^{-\lambda(x-a)} dx}{1 - e^{-\lambda(b-a)}} && (19) \\
&= \frac{a - b e^{-\lambda(b-a)}}{1 - e^{-\lambda(b-a)}} + \frac{1}{\lambda} && (20)
\end{aligned}
$$

4. If the first and second moments of X are $E[X] = 1$ and $E[X^2] = 2$, find

   (a) $E[(X-2)^2]$;
   Ans: Expanding the square gives us

$$
\begin{aligned}
E[(X-2)^2] &= E[X^2 - 4X + 4] && (21) \\
&= E[X^2] - 4E[X] + 4 && (22) \\
&= 2 - 4 + 4 = 2. && (23)
\end{aligned}
$$

   (b) the value of a that minimizes $E[(aX-1)^2]$;
   Ans: Let $g(a) = E[(aX-1)^2]$, then

$$
\begin{aligned}
g(a) &= E[a^2X^2 - 2aX + 1] && (24) \\
&= a^2E[X^2] - 2aE[X] + 1 && (25) \\
&= 2a^2 - 2a + 1 && (26) \\
&= 2[(a-0.5)^2 + 0.25]. && (27)
\end{aligned}
$$

   Therefore, $g(a)$ is minimized at $a = 0.5$.

   (c) $\text{var}(X-1)$.
   Ans: Let $Y = X - 1$. Then $E[Y^2] = E[X^2] - 2E[X] + 1 = 1$ and $E[Y] = E[X] - 1 = 0$, thus $\text{var}(X-1) = 1$.
   In fact, $\text{var}(aX + b) = a^2\text{var}(X)$ for any a, b.

Important Random Variables

1. An average of 50 pedestrians pass a given point on a sidewalk per hour in the day, and the average drops to 10 in the night. Assuming that the number of pedestrians per hour is Poisson, find the probability density function of the time between successive pedestrians passing this point.

   Ans: Let T be the time interval between successive pedestrians in hours. In the day, T is exponentially distributed with $\lambda = 50$ per hour; in the night, T is exponentially distributed with $\lambda = 10$ per hour. In essence,

$$
\begin{aligned}
f_T(t|\text{“day”}) &= 50e^{-50t}, \quad t > 0, \\
f_T(t|\text{“night”}) &= 10e^{-10t}, \quad t > 0.
\end{aligned}
$$

A randomly chosen inter-pedestrian interval has an equal chance of occurring in the day and in the night, therefore we have

$$\begin{aligned} f_T(t) &= 0.5 f_T(t|\text{``day''}) + 0.5 f_T(t|\text{``night''}) \\ &= 25e^{-50t} + 5e^{-10t}, \quad t > 0. \end{aligned}$$

2. The r-th percentile, $\pi(r)$, of a random variable X is defined by $P[X \leq \pi(r)] = r/100$.

   (a) Find the 90th, 95th, and 99th percentiles of the exponential random variable with parameter λ.

   Ans: The CDF of the exponential r.v. is $F_X(x) = 1 - e^{-\lambda x}$, $x > 0$. To find the 90-th percentile, we solve for x in

$$\begin{aligned} F_X(x) &= 0.9 && (28) \\ \Rightarrow \quad e^{-\lambda x} &= 0.1 && (29) \\ \Rightarrow \quad x &= \frac{1}{\lambda} \ln 10 = \pi(90). && (30) \end{aligned}$$

   Similarly, we find $\pi(95)$ and $\pi(99)$ by changing the RHS of the first line above to 0.95 and 0.99, to obtain

$$\pi(95) = \frac{1}{\lambda} \ln 20, \quad \pi(99) = \frac{1}{\lambda} \ln 100. \tag{31}$$

   (b) Repeat part (a) for the Gaussian random variable with parameters $\mu = 0$ and σ^2.

   Ans: If $X \sim \mathcal{N}(0, \sigma^2)$, then $P[X > x] = Q\left(\frac{x}{\sigma}\right)$. Therefore,

$$\begin{aligned} P[X > \pi(90)] &= Q\left(\frac{\pi(90)}{\sigma}\right) = 0.1 && (32) \\ \Rightarrow \quad \pi(90) &= \sigma Q^{-1}(0.1) && (33) \\ &= 1.2815\sigma && (34) \end{aligned}$$

   where the final result was obtained using Table 4.3 in the textbook. Similarly,

$$\begin{aligned} Q(\pi(95)/\sigma) &= 0.05 && (35) \\ \Rightarrow \quad \pi(95) &= \sigma Q^{-1}(0.05) && (36) \\ &= 1.65\sigma, && (37) \\ Q(\pi(99)/\sigma) &= 0.01 && (38) \\ \Rightarrow \quad \pi(99) &= \sigma Q^{-1}(0.01) && (39) \\ &= 2.3263\sigma. && (40) \end{aligned}$$

   Since $Q^{-1}(0.05)$ is not tabulated in Table 4.3, and from Table 4.2, $Q(1.6) > 0.05 > Q(1.7)$, we estimate $Q^{-1}(0.05)$ as 1.65.

3. Let X be an exponential random variable with parameter λ.

   (a) For $d > 0$ and k a non-negative integer, find $P[kd < X \leq (k+1)d]$.
   Ans: The CDF of X is $F_X(x) = 1 - e^{-\lambda x}$, $x > 0$, thus

   $$P[kd < X \leq (k+1)d] = e^{-\lambda kd} - e^{-\lambda(k+1)d} = e^{-\lambda kd}(1 - e^{-\lambda d}).$$

   (b) Segment the positive real line into four equi-probable disjoint intervals.
   Ans: Basically, we need to find a, b, and c so that $P[0 < X \leq a] = P[a < X \leq b] = P[b < X \leq c] = P[c < X] = 0.25$. To find a, we solve

   $$\begin{aligned} 1 - e^{-\lambda a} &= 0.25 \\ \Rightarrow \quad a &= \frac{1}{\lambda} \ln \frac{4}{3}. \end{aligned} \tag{41}$$

   To find b, we solve

   $$1 - e^{-\lambda b} = 0.5 \tag{42}$$
   $$\Rightarrow \quad b = \frac{1}{\lambda} \ln 2. \tag{43}$$

   To find c, we solve

   $$e^{-\lambda c} = 0.25 \tag{44}$$
   $$\Rightarrow \quad c = \frac{1}{\lambda} \ln 4. \tag{45}$$

4. Let X be a Gaussian random variable with mean μ and variance σ^2.

   (a) Find $P[X \leq \mu]$.
   Ans: Since $f_X(x)$ has even symmetry about $x = m$, $P[X \leq m] = 0.5$.

   (b) Find $P[|X - \mu| < k\sigma]$, for $k = 1, 2, 3, 4$.
   Ans: Note that $\{|X - \mu| < k\sigma\}$ is equivalent to $\{-k\sigma + \mu < X < k\sigma + \mu\}$, and that $f_X(x)$ has even symmetry about $x = \mu$. (A sketch of the PDF is useful here.) Therefore,

   $$P[|X - \mu| < k\sigma] = 1 - 2P[X > k\sigma + \mu] = 1 - 2Q(k).$$

   From Table 4.2, we can then find that

   $$P[|X - \mu| < k\sigma] = \begin{cases} 0.682 & k = 1 \\ 0.9544 & k = 2 \\ 0.9973 & k = 3 \\ 0.9999 & k = 4 \end{cases}$$

   For larger values of k, the probability of X deviating from $E[X]$ by more than $k\sigma$ is basically zero.

(c) Find the value of k for which $P[X > \mu + k\sigma] = 10^{-j}$, for $j = 1, 2, 3, 4, 5, 6$.

Ans: Noting that $P[X > \mu + k\sigma] = Q(k)$, we want to basically solve $Q(k) = 10^{-j}$. From Table 4.3, we have

$$k = \begin{cases} 1.2815 & j = 1 \\ 2.3263 & j = 2 \\ 3.0902 & j = 3 \\ 3.7190 & j = 4 \\ 4.2649 & j = 5 \\ 4.7535 & j = 6. \end{cases}$$

In other words, the probability of X exceeding $E[X]$ by 4.75σ is only 1 in a million.

5. Two chips are being considered for use in a certain system. The lifetime of chip 1 is modelled by a Gaussian RV with mean 20,000 hours and standard deviation 5,000 hours. The lifetime of chip 2 is also Gaussian, with mean 22,000 hours and standard deviation 1,000 hours. Which chip is preferred if the target lifetime is 20,000 hours? What if the target is 24,000 hours?

Ans: Let the lifetimes of chip 1 and chip 2 be denoted by the random variables X_1 and X_2 respectively. We are told that $X_1 \sim \mathcal{N}(20000, 5000^2)$ and $X_2 \sim \mathcal{N}(22000, 1000^2)$. If the target lifetime of the chip is 20,000 hours, then we compare $P[X_1 > 20000]$ against $P[X_2 > 20000]$ and pick the one with the higher probability:

$$\begin{aligned} P[X_1 > 20000] &= Q(0) = 0.5 && (46) \\ P[X_2 > 20000] &= Q\left(-\frac{2000}{1000}\right) = 1 - Q(2) && (47) \\ &= 0.9772. && (48) \end{aligned}$$

Therefore we prefer Chip 2. If the target lifetime is 24,000 hours, we calculate

$$\begin{aligned} P[X_1 > 24000] &= Q\left(\frac{4000}{5000}\right) = Q(0.8) && (49) \\ &= 0.212 && (50) \\ P[X_2 > 24000] &= Q\left(\frac{2000}{1000}\right) = Q(2) && (51) \\ &= 0.0228. && (52) \end{aligned}$$

Thus now we prefer Chip 1. The reason is that Chip 1's lifetime has a much larger standard deviation than Chip 2's. Therefore, while on average Chip 2 will last longer than Chip 1, it is more likely for Chip 1 to last 4000 hours beyond its average than for Chip 2 to last 2000 hours beyond its average.