Clex 4: Continuous Random Variables - Solutions

Problem 1: Corridor Aérobique

Morin-Heights is the starting point for Corridor Aérobique which during the winter months is a 58 km long cross-country skiing trail. Yvan & Co. ski on this trail frequently. The distance covered during these excursions is well-modelled by a random variable with a probability density function

$$f(x) = \begin{cases} (x-12)^2/1152 & 12 \leq x < 24 \\ 9/(x-12)^2 & 24 \leq x \leq 48 \\ 0 & \text{otherwise} \end{cases}$$

a) Find the cumulative distribution function of the distance covered.

The CDF is defined as $F(x) = \int_{-\infty}^{x} f(t)dt$.

- For $x < 12$: $F(x) = 0$.
- For $12 \leq x < 24$:

$$F(x) = \int_{12}^{x} \frac{(t-12)^2}{1152} dt = \left[\frac{(t-12)^3}{3456}\right]_{12}^{x} = \frac{(x-12)^3}{3456}$$

- For $24 \leq x \leq 48$: First, note that $F(24) = \frac{(24-12)^3}{3456} = \frac{1728}{3456} = 0.5$.

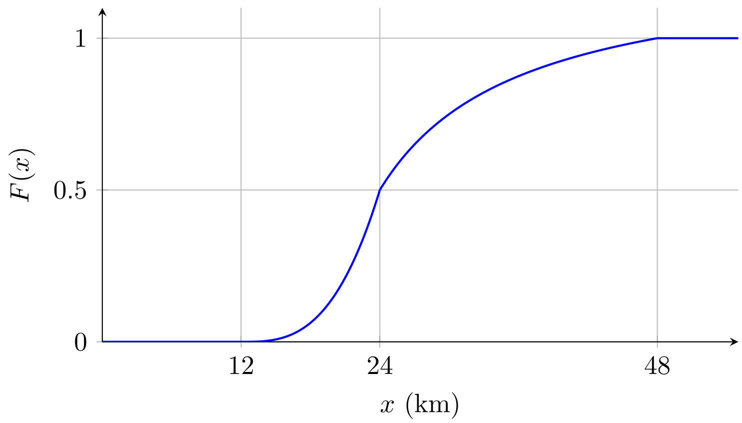
$$F(x) = 0.5 + \int_{24}^{x} \frac{9}{(t-12)^2} dt = 0.5 + 9\left[-\frac{1}{t-12}\right]_{24}^{x}$$

$$F(x) = 0.5 + 9\left(\frac{1}{12} - \frac{1}{x-12}\right) = 0.5 + 0.75 - \frac{9}{x-12} = 1.25 - \frac{9}{x-12}$$

- For $x > 48$: $F(x) = 1$.

$$F(x) = \begin{cases} 0 & x < 12 \\ \frac{(x-12)^3}{3456} & 12 \leq x < 24 \\ 1.25 - \frac{9}{x-12} & 24 \leq x \leq 48 \\ 1 & x > 48 \end{cases}$$

b) Sketch the graph of the cumulative distribution function.



c) Find the ninety-eighth percentile, P_{98}, of the distance covered.

We solve $F(x) = 0.98$. Since $0.98 > 0.5$, we use the second piece of the CDF:

$$1.25 - \frac{9}{x-12} = 0.98 \implies 0.27 = \frac{9}{x-12} \implies x = 45.33 \text{ km}$$

d) Determine the average distance covered.

$E[X] = \int_{12}^{24} x \frac{(x-12)^2}{1152} dx + \int_{24}^{48} x \frac{9}{(x-12)^2} dx$. Using substitution $u = x - 12$:

- Integral 1: $\frac{1}{1152} \int_0^{12} (u+12)u^2 du = \frac{1}{1152} [\frac{u^4}{4} + 4u^3]_0^{12} = \frac{5184+6912}{1152} = 10.5$
- Integral 2: $9 \int_{12}^{36} \frac{u+12}{u^2} du = 9[\ln u - \frac{12}{u}]_{12}^{36} = 9(\ln 3 + \frac{2}{3}) = 9 \ln 3 + 6 \approx 15.89$

$E[X] = 10.5 + 15.89 = 26.39$ km.

Problem 2: Rates of calls and exponential waiting

A physicians office receives an average of six calls per hour. We define the random variable T as the time (in minutes) between consecutive calls. Since the average rate is $\lambda = 6$ calls/hour, the rate per minute is $\lambda = 0.1$ calls/min.

For an exponential distribution, the probability that the waiting time T is less than or equal to t is given by the cumulative distribution function:

$$F(t) = P(T \leq t) = 1 - e^{-\lambda t}$$

Conversely, the probability that we wait *longer* than t minutes is $P(T > t) = e^{-\lambda t}$.

1. **What is the probability that the office will receive no calls in the next 10 minutes?** Receiving "no calls" in 10 minutes is equivalent to the event that the waiting time for the next call is greater than 10 minutes ($T > 10$).

$$P(T > 10) = e^{-0.1(10)} = e^{-1} \approx 0.3679$$

Note: This is identical to the Poisson result for $k = 0$ arrivals over a 10-minute interval.

2. **What is the probability that there will be at least one phone call within the next half hour?** Half an hour is $t = 30$ minutes. "At least one call" is the event that the waiting time for the first call is less than or equal to 30 minutes ($T \leq 30$).

$$P(T \leq 30) = F(30) = 1 - e^{-0.1(30)} = 1 - e^{-3} \approx 0.9502$$

3. **If no phones calls were received at the office in the last 20 minutes, what is the probability that one will be received in the next 8 minutes?** The exponential distribution is characterized by the **memoryless property**: $P(T > s + t \mid T > s) = P(T > t)$. The fact that 20 minutes have passed without a call ($s = 20$) does not change the distribution of the remaining time until the next call. Thus, we simply find the probability of a call within the next 8 minutes:

$$P(T \leq 8) = 1 - e^{-0.1(8)} = 1 - e^{-0.8} \approx 0.5507$$

4. **How many minutes would the office have to wait in order to get at least one phone call with 99% probability?** We seek the time t such that the probability of at least one arrival (i.e., the waiting time is $\leq t$) is 0.99.

$$1 - e^{-0.1t} = 0.99, \quad 0.01 = e^{-0.1t}$$

Taking the natural logarithm of both sides:

$$t = \frac{\ln(0.01)}{-0.1} \approx \frac{-4.605}{-0.1} = 46.05 \text{ minutes}$$

Problem 3: Standardization

Show that if X is a normal random variable with mean μ and st.dev σ then $Z = (X - \mu)/\sigma$ is standard normal.

1. **Write the cdf $F_Z(z)$ in terms of the cdf of the random variable X.**

$$F_Z(z) = P(Z \leq z) = P\left(\frac{X - \mu}{\sigma} \leq z\right) = P(X \leq \sigma z + \mu) = F_X(\sigma z + \mu)$$

2. **Change appropriately the variable of integration.** The PDF of X is $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$. Thus:

$$F_Z(z) = \int_{-\infty}^{\sigma z + \mu} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx$$

Let $u = \frac{x-\mu}{\sigma}$, then $dx = \sigma du$. When $x = \sigma z + \mu$, $u = z$.

$$F_Z(z) = \int_{-\infty}^{z} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}u^2} (\sigma du) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$$

This is the integral defining the CDF of the standard normal distribution $N(0, 1)$.