Statistics for Data Science
Unit 4 Part 2 Homework: Continuous Random Variables

Brad Andersen
W203 Section 4

February 6, 2019

1. **Processing Pasta**

   A certain manufacturing process creates pieces of pasta that vary by length. Suppose that the length of a particular piece, L, is a continuous random variable with the following probability density function.

$$f(l) = \begin{cases} 0, & l \leq 0 \\ l/2, & 0 < l \leq 2 \\ 0, & 2 < l \end{cases}$$

   (a) **Write down a complete expression for the cumulative probability function of L.**

   Identifying the antiderivative of the function for range (0,l]:

$$\begin{aligned} F_L(l)dl &= \int_0^l f(m)dm \\ &= \int_0^l \frac{m}{2} dm \\ &= \int_0^l \frac{m^2}{4} \end{aligned}$$

   Calculating for the interval (0,l]:

$$\begin{aligned} F_L(l)dl &= \left[\frac{m^2}{4}\right]_0^l \\ &= \left[\frac{l^2}{4} - \frac{0^2}{4}\right] \\ &= \frac{l^2}{4} \end{aligned}$$

$$\text{For } l \in (0, 2] \; F_L(l)dl = \frac{l^2}{4}$$

Applying this equation to complete the cumulative probability function:

$$F_L(l) = \begin{cases} 0, & l \leq 0 \\ \dfrac{l^2}{4}, & 0 < l \leq 2 \\ 1, & 2 < l \end{cases}$$

(b) **Using the definition of expectation for a continuous random variable, compute the expected length of the pasta, $E(L)$.**

$$\begin{aligned} E(L) &= \int_{-\infty}^{\infty} l \cdot f(l) dl \\ &= \int_0^2 l \cdot \frac{l}{2} dl \\ &= \frac{1}{2} \int_0^2 l^2 dl \\ &= \frac{1}{2} \left(\frac{l^3}{3} \right) \Bigg|_{l=0}^{l=2} \\ &= \frac{1}{2} \left(\frac{8}{3} \right) \\ &= \underline{\underline{\frac{4}{3}}} \end{aligned}$$

2. **The Warranty is Worth It**

Suppose the life span of a particular (shoddy) server is a continuous random variable, T, with a uniform probability distribution between 0 and 1 year. The server comes with a contract that guarantees you money if the server lasts less than 1 year. In particular, if the server lasts t years, the manufacturer will pay you $g(t) = \$100(1-t)^{1/2}$. Let $X = g(T)$ be the random variable representing the payout from the contract.

Compute the expected payout from the contract, $E(X) = E(g(T))$.

Given that T is a continuous random variable with uniform distribution on the interval [0 years, 1 year], we can represent the probability density function as follows:

$$f(t; 0, 1) = \begin{cases} \dfrac{1}{1-0}, & 0 \leq t \leq 1 \\ 0, & otherwise \end{cases}$$

Validating uniformity, the expected value of T can be calculated as follows:

$$
\begin{aligned}
E(T) &= \int_{-\infty}^{\infty} t \cdot f(t)dt \\
&= \int_{0}^{1} t \cdot f(t)dt \\
&= \int_{0}^{1} t \cdot 1 \cdot dt \\
&= \int_{0}^{1} \frac{t^2}{2}
\end{aligned}
$$

Calculating for the interval [0, 1]:

$$
\begin{aligned}
E(T) &= \frac{1^2}{2} - \frac{0^2}{2} \\
&= \frac{1}{2}
\end{aligned}
$$

The expected value of T is 0.5 years.

Given that $g(T)$ represents the function of a random variable, it itself is a random variable. The probability density function of $g(t)$ is as follows:

$$
g(t) = \begin{cases} 0, & t < 0 \\ \$100(1-t)^{1/2}, & 0 \leq t \leq 1 \\ 0, & 1 < t \end{cases}
$$

The expected value of $g(t)$ can be calculated as follows:

$$
\begin{aligned}
E(g(t)) &= \int_{-\infty}^{\infty} g(t) \cdot f(t)dt \\
&= \int_{0}^{1} (\$100(1-t)^{1/2}) \cdot 1 \cdot dt \\
&= \$100 \int_{0}^{1} (1-t)^{1/2}dt \\
&= \$100 \int_{0}^{1} -\frac{2}{3}(1-t)^{3/2} \\
&= \$100 \int_{0}^{1} -\frac{2(1-t)^{3/2}}{3} \\
&= \$100\left(-\frac{2(1-1)^{3/2}}{3}\right) - \$100\left(-\frac{2(1-0)^{3/2}}{3}\right) \\
&= \$100\left(0\right) - \$100\left(-\frac{2}{3}\right) \\
&= \frac{\$200}{3}
\end{aligned}
$$

The expected payout from the contract is \$66.67 rounded to the nearest cent.

3. **(Lecture)#Fail**

Suppose the length of Paul Laskowski's lecture in minutes is a continuous random variable C, with pmf $f(t) = e^{-t}$ for $t > 0$. This is an example of an exponential random variable, and it has some special properties. For example, suppose you have already sat through t minutes of the lecture, and are interested in whether the lecture is about to end immediately. In statistics, this can be represented by something called the *hazard rate*:

$$h(t) = \frac{f(t)}{1 - F(t)}$$

To understand the hazard rate, think of the numerator as the probability the lecture ends between time t and time $t + dt$. The denominator is just the probability the lecture does not end before time t. So you can think of the fraction as the conditional probability that the lecture ends between t and $t + dt$ given that it did not end before t.

Compute the hazard rate for C.

The probability density function for $f(t)$ can be represented as follows:

$$f(t) = \begin{cases} e^{-t}, & 0 < t \\ 0, & otherwise \end{cases}$$

Computing $F(t)$:

$$\begin{aligned} F(t) &= \int_{-\infty}^{\infty} f(t)dt \\ &= \int_{-\infty}^{\infty} (e^{-t})dt \\ &= \int_{0}^{\infty} -e^{-t} \text{ where } t > 0 \end{aligned}$$

Substituting values, the hazard rate for C can be calculated as follows:

$$\begin{aligned} h(t) &= \frac{e^{-t}}{1 - (-e^{-t})} \text{ where } t > 0 \\ &= \frac{e^{-t}}{1 + e^{-t}} \text{ where } t > 0 \end{aligned}$$