Unit 4 - Part 2 Homework

W203 Statistics for Data Science

Student Name: Youzhi Wu

Section Number: 05

1. Processing Pasta

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

From the PDF we know that $\forall l \leq 0 : F_L(l) = 0$, because $\forall l \leq 0 : f_L(l) = 0$.

$$\begin{aligned}\forall l \in (0, 2] : F_L(l) &= \int_0^l f(l)dl \\ &= \int_0^l \frac{l}{2}dl \\ &= \left[\frac{l^2}{4}\right]_0^l \\ &= \frac{l^2}{4}\end{aligned}$$

From the PDF we know that $\forall l > 2 : F_L(l)$ will not increase, because $\forall l > 2 : f_L(l) = 0$.
Combining the three intervals, we get the following cumulative probability function of L:

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

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

$$
\begin{aligned}
E(L) &= \int_{-\infty}^{\infty} l \cdot f_L(l) dl \\
&= \int_{-\infty}^{0} l \cdot 0 dl + \int_{0}^{2} l \cdot \frac{l}{2} dl + \int_{2}^{\infty} l \cdot 0 dl \\
&= 0 + \left[\frac{l^3}{6}\right]_0^2 + 0 \\
&= \frac{4}{3}
\end{aligned}
$$

2. The Warranty is Worth It

Because continuous random variable T has a uniform probability distribution between 0 and 1 year, its probability density function can be written as follows:

$$
f_T(t) = \begin{cases} 0, & \text{for } t < 0 \\ 1, & \text{for } 0 \geq t < 1 \\ 0, & \text{for } 1 \geq t \end{cases}
$$

$$
\begin{aligned}
E(X) &= E(g(T)) \\
&= \int_{-\infty}^{\infty} g_T(t) \cdot f_T(t) dt \\
&= \int_{-\infty}^{0} g_T(t) \cdot 0 dt + \int_{0}^{1} g_T(t) \cdot 1 dt + \int_{1}^{\infty} g_T(t) \cdot 0 dt \\
&= 0 + \int_{0}^{1} 100(1-t)^{1/2} \cdot 1 dt + 0 \\
&= \left[-100 \cdot \frac{2}{3} \cdot (1-t)^{3/2}\right]_0^1 \\
&= \frac{200}{3}
\end{aligned}
$$

3. (Lecture)#Fail

$$
\begin{aligned}
\forall t > 0 : h(t) &= \frac{f(t)}{1 - F(t)} \\
&= \frac{f(t)}{1 - \int_{-\infty}^{t} f(t)dt} \\
&= \frac{f(t)}{1 - \int_{0}^{t} f(t)dt} \\
&= \frac{e^{-t}}{1 - \int_{0}^{t} e^{-t}dt} \\
&= \frac{e^{-t}}{1 - [-e^{-t}]_{0}^{t}} \\
&= \frac{e^{-t}}{1 - (-e^{-t} + 1)} \\
&= 1
\end{aligned}
$$

4. Optional Advanced Exercise: Characterizing a Function of a Random Variable

Because $Y = h(X)$ and that h is an invertible function, we have that $X = h^{-1}(Y)$

$$g(y) = \frac{d}{dy}[P(Y \leq y)]$$
$$= \frac{d}{dy}[P(h(X) \leq y)]$$

Since h is an invertible function, it is either monotonically increasing or monotonically decreasing. (1) If it is monotonically increasing, the above equation can be written as follows:

$$g(y) = \frac{d}{dy}\left[P(X \leq h^{-1}(y))\right]$$
$$= \frac{d}{dy}\int_{-\infty}^{h^{-1}(y)} f(x)dx$$
$$= \frac{d}{dy}\left[F(h^{-1}(y)) - F(-\infty)\right] \leftarrow \text{F as the CDF of X.}$$
$$= f(h^{-1}(y))\frac{d}{dy}h^{-1}(y) \leftarrow F(-\infty) \text{ is a constant, the derivative of it is zero.}$$

(2) If it is monotonically decreasing, the above equation can be written as follows:

$$g(y) = \frac{d}{dy}\left[P(X \geq h^{-1}(y))\right]$$
$$= \frac{d}{dy}\int_{h^{-1}(y)}^{\infty} f(x)dx$$
$$= \frac{d}{dy}\left[F(\infty) - F(h^{-1}(y))\right]$$
$$= -f(h^{-1}(y))\frac{d}{dy}h^{-1}(y)$$

Since $f(h^{-1}(y))$ must be non-negative, combining the two scenarios, the equation can be written as follows:

$$g(y) = f(h^{-1}(y)) \cdot \left|\frac{dh^{-1}(y)}{dy}\right|$$