Statistics 341

Fall 2008 - Homework Assignment #5 Solutions

Due Wednesday, October 29

This assignment is worth a total of 79 points.

1. (24 pts) Suppose a continuous random variable Y possesses the density function

$$f(y) = \begin{cases} cy & 0 \le y \le 2 \\ 0 & \text{elsewhere} \end{cases}$$

   (a) (3 pts) Find the value of c that makes $f(y)$ a probability density function.
   For $f(y)$ to be a probability density function, the integral of the function over the real line must be equal to 1.

$$\begin{aligned} \int_{-\infty}^{\infty} f(y)dy &= 1 \\ \int_{-\infty}^{0} 0dy + \int_{0}^{2} cydy + \int_{2}^{\infty} 0dy &= 1 \\ 0 + \int_{0}^{2} cydy + 0 &= 1 \\ \frac{c}{2}y^2|_0^2 &= 1 \\ \frac{c}{2}(4-0) &= 1 \\ 2c &= 1 \\ c &= 1/2 \end{aligned}$$

   (b) (5 pts) Find $F(y)$.
   The p.d.f. of Y divides the real line into three regions: $y < 0$, $0 \le y \le 2$ and $y > 2$.
   For the region $y < 0$:

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

   For the region $0 \le y \le 2$:

$$\begin{aligned} F(y) &= \int_{-\infty}^{y} f(t)dt \\ &= \int_{-\infty}^{0} 0dt + \int_{0}^{y} \frac{t}{2}dt \end{aligned}$$

$$\begin{aligned} &= 0 + \frac{1}{4}t^2|_0^y \\ &= \frac{1}{4}(y^2 - 0) \\ &= \frac{y^2}{4} \end{aligned}$$

For the region $y > 2$:

$$\begin{aligned} F(y) &= \int_{-\infty}^{y} f(t)dt \\ &= \int_{-\infty}^{0} 0dt + \int_{0}^{2} \frac{t}{2}dt + \int_{2}^{y} 0dt \\ &= 0 + \frac{1}{4}t^2|_0^2 + 0 \\ &= \frac{1}{4}(2^2 - 0) \\ &= 1 \end{aligned}$$

So the distribution function $F(y)$ is

$$F(y) = \begin{cases} 0 & y < 0 \\ \frac{y^2}{4} & 0 \le y \le 2 \\ 1 & y > 2 \end{cases}$$

(c) (3 pts) Find $P(1 \le Y \le 2)$ using $F(y)$.

$$\begin{aligned} P(1 \le Y \le 2) &= P(Y \le 2) - P(Y \le 1) \\ &= F(2) - F(1) \\ &= \frac{2^2}{4} - \frac{1^2}{4} \\ &= 1 - \frac{1}{4} \\ &= \frac{3}{4} \end{aligned}$$

(d) (3 pts) Find $P(0.5 \le Y \le 1.5)$ using $f(y)$.

$$\begin{aligned} P(0.5 \le Y \le 1.5) &= \int_{0.5}^{1.5} \frac{y}{2}dy \\ &= \frac{1}{4}y^2|_{0.5}^{1.5} \\ &= \frac{1}{4}(1.5^2 - 0.5^2) \\ &= 0.5 \end{aligned}$$

(e) (4 pts) Find $E(Y)$.

The only interval of the real line where the p.d.f. is non-zero is between 0 and 2. So the expected value is

$$E(Y) = \int_0^2 yf(y)dy = \int_0^2 y\frac{y}{2}dy = \int_0^2 \frac{y^2}{2} = \frac{1}{6}y^3|_0^2 = \frac{1}{6}(2^3 - 0^3) = \frac{8}{6} = \frac{4}{3} = 1.3333$$

(f) (6 pts) Find $V(Y)$.

Again the only interval of the real line where the p.d.f. is non-zero is between 0 and 2. So $E(Y^2)$ is

$$E(Y^2) = \int_0^2 y^2 f(y)dy = \int_0^2 y^2\frac{y}{2}dy = \int_0^2 \frac{y^3}{2} = \frac{1}{8}y^4|_0^2 = \frac{1}{8}(2^4 - 0^3) = 2$$

The variance is then

$$V(Y) = E(Y^2) - (E(Y))^2 = 2 - \left(\frac{4}{3}\right)^2 = 2 - \frac{16}{9} = \frac{2}{9} = 0.2222$$

2. (26 pts) A supplier of kerosene has a 150-gallon tank that is filled at the beginning of each week. His weekly demand shows a relative frequency behavior that increases steadily up to 100 gallons and then levels off between 100 and 150 gallons. If Y denotes weekly demand in hundreds of gallons, the relative frequency of demand can be modeled by

$$f(y) = \begin{cases} y & 0 \leq y \leq 1 \\ 1 & 1 < y \leq 1.5 \\ 0 & \text{elsewhere} \end{cases}$$

(a) (8 pts) Find $F(y)$.

The p.d.f. divides the real line into four regions: $y < 0$, $0 \leq y \leq 1$, $1 < y \leq 1.5$ and $y > 1.5$.

For the region $y < 0$:

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

For the region $0 \leq y \leq 1$:

$$\begin{aligned} F(y) &= \int_{-\infty}^{y} f(t)dt \\ &= \int_{-\infty}^{0} 0dt + \int_0^y tdt \\ &= 0 + \frac{1}{2}t^2|_0^y \\ &= \frac{1}{2}(y^2 - 0) \\ &= \frac{y^2}{2} \end{aligned}$$

For the region $1 < y \leq 1.5$:

$$\begin{aligned} F(y) &= \int_{-\infty}^{y} f(t)dt \\ &= \int_{-\infty}^{0} 0dt + \int_{0}^{1} tdt + \int_{1}^{y} 1dt \\ &= 0 + \frac{1}{2}t^2|_0^1 + t|_1^y \\ &= \frac{1}{2}(1^2 - 0) + (y - 1) \\ &= y - \frac{1}{2} \end{aligned}$$

For the region $y > 1.5$:

$$\begin{aligned} F(y) &= \int_{-\infty}^{y} f(t)dt \\ &= \int_{-\infty}^{0} 0dt + \int_{0}^{1} tdt + \int_{1}^{1.5} 1dt + \int_{1.5}^{y} 0dt \\ &= 0 + \frac{1}{2}t^2|_0^1 + t|_1^{1.5} \\ &= \frac{1}{2}(1^2 - 0) + (1.5 - 1) \\ &= 0.5 + 0.5 \\ &= 1 \end{aligned}$$

The distribution function $F(y)$ is then

$$F(y) = \begin{cases} 0 & y < 0 \\ \frac{y^2}{2} & 0 \leq y \leq 1 \\ y - 0.5 & 1 < y \leq 1.5 \\ 1 & y > 1.5 \end{cases}$$

(b) (3 pts) Find $P(0 \leq Y \leq 0.5)$.

Since you have just calculated the distribution function, I would use it to calculate this probability.

$$\begin{aligned} P(0 \leq Y \leq 0.5) &= P(Y \leq 0.5) - P(Y \leq 0) \\ &= F(0.5) - F(0) \\ &= \frac{0.5^2}{2} - \frac{0^2}{2} \\ &= 0.125 \end{aligned}$$

(c) (3 pts) Find $P(0.5 \le Y \le 1.2)$.

Again, since you have just calculated the distribution function, I would use it to calculate this probability.

$$\begin{aligned} P(0.5 \le Y \le 1.2) &= P(Y \le 1.2) - P(Y \le 0.5) \\ &= F(1.2) - F(0.5) \\ &= (1.2 - 0.5) - \frac{0.5^2}{2} \\ &= 0.7 - 0.125 \\ &= 0.575 \end{aligned}$$

(d) (5 pts) Find $E(Y)$.

The p.d.f. of Y is non-zero only in the interval between 0 and 1.5. The value of the p.d.f. changes at 1. So the expected value is

$$\begin{aligned} E(Y) &= \int_0^{1.5} y f(y) dy \\ &= \int_0^1 yy dy + \int_1^{1.5} y 1 dy \\ &= \int_0^1 y^2 dy + \int_1^{1.5} y dy \\ &= \frac{1}{3} y^3 |_0^1 + \frac{1}{2} y^2 |_1^{1.5} \\ &= \frac{1}{3}(1^3 - 0^3) + \frac{1}{2}(1.5^2 - 1^2) \\ &= \frac{1}{3} + \frac{5}{8} \\ &= \frac{23}{24} = 0.9583 \end{aligned}$$

(e) (7 pts) Find $V(Y)$.

Again, the p.d.f. of Y is non-zero only in the interval between 0 and 1.5. The value of the p.d.f. changes at 1. So $E(Y^2)$ is

$$\begin{aligned} E(Y^2) &= \int_0^{1.5} y^2 f(y) dy \\ &= \int_0^1 y^2 y dy + \int_1^{1.5} y^2 1 dy \\ &= \int_0^1 y^3 dy + \int_1^{1.5} y^2 dy \\ &= \frac{1}{4} y^4 |_0^1 + \frac{1}{3} y^3 |_1^{1.5} \\ &= \frac{1}{4}(1^4 - 0^4) + \frac{1}{3}(1.5^3 - 1^3) \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{4} + \frac{1}{3}\left(\frac{19}{8}\right) \\
&= \frac{1}{4} + \frac{19}{24} \\
&= \frac{25}{24} = 1.0417
\end{aligned}
$$

$$
\begin{aligned}
V(Y) &= E(Y^2) - (E(Y))^2 \\
&= \frac{25}{24} - \left(\frac{23}{24}\right)^2 \\
&= \frac{71}{24^2} = 0.1233
\end{aligned}
$$

3. (15 pts) A gas station operates two pumps, each of which can pump up to 10,000 gallons of gas in a month. The total amount of gas pumped at the station in a month is a random variable Y (measured in 10,000 gallons) with a probability density function given by

$$
f(y) = \begin{cases} y & 0 < y < 1 \\ 2 - y & 1 \le y < 2 \\ 0 & \text{elsewhere} \end{cases}
$$

(a) (3 pts) Find the probability that the station will pump between 8000 and 12,000 gallons in a particular month.

This probability is the same as $P(0.8 \le Y \le 1.2)$. Since we don't have the distribution function, I would calculate this probability using the p.d.f.

$$
\begin{aligned}
P(0.8 \le Y \le 1.2) &= \int_{0.8}^{1.2} f(y)dy \\
&= \int_{0.8}^{1} ydy + \int_{1}^{1.2} (2-y)dy \\
&= \frac{1}{2}y^2|_{0.8}^{1} + 2y|_{1}^{1.2} - \frac{1}{2}y^2|_{1}^{1.2} \\
&= \frac{1}{2}(1^2 - 0.8^2) + 2(1.2-1) - \frac{1}{2}(1.2^2 - 1^2) \\
&= 0.18 + 0.4 - 0.22 \\
&= 0.36
\end{aligned}
$$

(b) (5 pts) Find the expected number of gallons the station will pump in a particular month.

The range where the p.d.f. is non-zero is between 0 and 2. So the expected value is

$$
\begin{aligned}
E(Y) &= \int_{0}^{2} yf(y)dy \\
&= \int_{0}^{1} yydy + \int_{1}^{2} y(2-y)dy
\end{aligned}
$$

$$
\begin{aligned}
&= \int_0^1 y^2 dy + \int_1^2 2y - y^2 dy \\
&= \frac{1}{3} y^3|_0^1 + y^2|_1^2 - \frac{1}{3} y^3|_1^2 \\
&= \frac{1}{3}(1^3 - 0^3) + (2^2 - 1^2) - \frac{1}{3}(2^3 - 1^3) \\
&= \frac{1}{3} + 3 - \frac{7}{3} \\
&= 1
\end{aligned}
$$

Note: this expectation makes sense geometrically, since the p.d.f. is symmetric about 1.

(c) (7 pts) Find the variance of the number of gallons the station will pump in a particular month.

Again, the range where the p.d.f. is non-zero is between 0 and 2. So $E(Y^2)$ and $V(Y)$ are

$$
\begin{aligned}
E(Y^2) &= \int_0^2 y^2 f(y) dy \\
&= \int_0^1 y^2 y dy + \int_1^2 y^2(2-y) dy \\
&= \int_0^1 y^3 dy + \int_1^2 2y^2 - y^3 dy \\
&= \frac{1}{4} y^4|_0^1 + \frac{2}{3} y^3|_1^2 - \frac{1}{4} y^4|_1^2 \\
&= \frac{1}{4}(1^4 - 0^4) + \frac{2}{3}(2^3 - 1^3) - \frac{1}{4}(2^4 - 1^4) \\
&= 0.25 + \frac{14}{3} - 3.75 \\
&= \frac{-7}{2} + \frac{14}{3} \\
&= \frac{7}{6} = 1.1667
\end{aligned}
$$

$$
\begin{aligned}
V(Y) &= E(Y^2) - (E(Y))^2 \\
&= \frac{7}{6} - 1^2 \\
&= \frac{1}{6} = 0.1667
\end{aligned}
$$

4. (14 pts) As a measure of intelligence, mice are timed when going through a maze to reach a reward of food. The time (in seconds) required for any mouse is a random variable Y with a density function given by

$$
f(y) = \begin{cases} \frac{b}{y^2} & y \geq b \\ 0 & \text{elsewhere} \end{cases}
$$

where b is the minimum possible time needed to traverse the maze.

(a) (5 pts) Show the $f(y)$ has the properties of a density function.

First, we must show that $f(y) \geq 0$ for all y. This is true since $b > 0$.

Second, we must show that $\int_{-\infty}^{\infty} f(y)dy = 1$. This is true since

$$\int_{-\infty}^{\infty} f(y)dy = \int_{-\infty}^{b} 0dy + \int_{b}^{\infty} \frac{b}{y^2}dy = -b\frac{1}{y}|_b^{\infty} = -b(\frac{1}{\infty} - \frac{1}{b}) = 0 + b\frac{1}{b} = 1$$

(b) (3 pts) Find $P(Y > b + c)$ for a positive constant c.

Since we don't have the distribution function, I would find this probability using the p.d.f.

$$\begin{aligned} P(Y > b + c) &= \int_{b+c}^{\infty} \frac{b}{y^2}dy \\ &= -b\frac{1}{y}|_{b+c}^{\infty} \\ &= -b\left(\frac{1}{\infty} - \frac{1}{b+c}\right) \\ &= \frac{b}{b+c} \end{aligned}$$

(c) (3 pts) Find the expected number of seconds required for a particular mouse to reach the reward.

The p.d.f. will be non-zero only when $y \geq b$. So the expected value is

$$\begin{aligned} E(Y) &= \int_{b}^{\infty} yf(y)dy \\ &= \int_{b}^{\infty} y\frac{b}{y^2}dy \\ &= \int_{b}^{\infty} \frac{b}{y}dy \\ &= b\ln y|_b^{\infty} \\ &= \infty \end{aligned}$$

The expected value for this random variable does not converge. Therefore, the expected value is undefined. While not common, random variables do exist where moments, such as the mean, are undefined.

(d) (3 pts) Find the variance of the number of seconds required for a particular mouse to reach the reward.

For any random variable with an undefined mean $E(Y)$, the second moment $E(Y^2)$ and all other higher moments will also be undefined. This random variable has an undefined variance.