IC105: Probability and Statistics **2021-22-M**

Lecture 8: Expectation (Expected Value) of a Random Variable

Instructor: Dr. Kuldeep Kumar Kataria *Scribe:*

8.1. Probability Distribution of a Function of Continuous Random Variable

Let X be a continuous r.v. with d.f. F, p.d.f. $f(\cdot)$ and support $S = \{x \in \mathbb{R} : F(x+h) - F(x-h) = \int_{x-h}^{x+h} f(t)\mathrm{d}t > 0,\ \forall\ h > 0\}$. For convenience assume that $S = [a, b]$ and $\{x \in \mathbb{R} : f(x) > 0\} = (a, b)$, for some $-\infty \leq a < b \leq \infty$ (with the convention that $[-\infty, b] \equiv (-\infty, b)$, $\forall\ b \in \mathbb{R}$, $[a, \infty] \equiv (a, \infty)$, $\forall\ a \in \mathbb{R}$ and $[-\infty, \infty] \equiv (-\infty, \infty)$).

Let $h : \mathbb{R} \to \mathbb{R}$ be a function such that h is strictly monotone and differentiable function on S. Then $Z = h(X)$ is a r.v. with d.f. $G(z) = P(Z \leq z) = P(h(X) \leq z)$, $z \in \mathbb{R}$.

For any sets $A \subseteq \mathbb{R}$ and $B \subseteq \mathbb{R}$, define $h(A) = \{h(x) : x \in A\}$ and $h^{-1}(B) = \{x \in \mathbb{R} : h(x) \in B\}$. Clearly $P(X \in (a, b)) = 1$ and therefore $P(h(X) \in h((a, b)) = 1)$. Consider the following cases:

Case I: $h(\cdot)$ **is strictly increasing on** S

We have $P(h(a) < Z < h(b)) = 1$. Therefore, for $z < h(a)$, $P(Z \leq z) = 0$ and for $z \geq h(b)$, $P(Z \leq z) = 1$. For $h(a) < z < h(b)$,

$$G(z) = P(h(X) \leq z) = P(X \leq h^{-1}(z)) = \int_{-\infty}^{h^{-1}(z)} f(t)\mathrm{d}t = \int_{a}^{h^{-1}(z)} f(t)\mathrm{d}t = \int_{h(a)}^{z} f(h^{-1}(y)) \left|\frac{\mathrm{d}}{\mathrm{d}y} h^{-1}(y)\right| \mathrm{d}y.$$

Thus,

$$G(z) = \begin{cases} 0, & \text{if } z < h(a), \\ \displaystyle\int_{h(a)}^{z} f(h^{-1}(y)) \left|\frac{\mathrm{d}}{\mathrm{d}y} h^{-1}(y)\right| \mathrm{d}y, & \text{if } h(a) \leq z < h(b), \\ 1, & \text{if } z \geq h(b). \end{cases}$$

Since f is continuous on (a, b) it follows that $G(z)$ is differentiable everywhere except possibly at $z = h(a)$ and $z = h(b)$. Moreover,

$$G'(z) = \begin{cases} f(h^{-1}(z)) \left|\frac{\mathrm{d}}{\mathrm{d}z} h^{-1}(z)\right|, & \text{if } h(a) < z < h(b), \\ 0, & \text{otherwise}, \end{cases}$$

and

$$\int_{-\infty}^{\infty} G'(z)\mathrm{d}z = \int_{h(a)}^{h(b)} f(h^{-1}(z)) \left|\frac{\mathrm{d}}{\mathrm{d}z} h^{-1}(z)\right| \mathrm{d}z = \int_{a}^{b} f(t)\mathrm{d}t = 1.$$

It follows that Z is a continuous r.v. with p.d.f.

$$g(z) = \begin{cases} f(h^{-1}(z)) \left|\frac{\mathrm{d}}{\mathrm{d}z} h^{-1}(z)\right|, & \text{if } h(a) < z < h(b), \\ 0, & \text{otherwise} \end{cases}$$

and support $T = [h(a), h(b)]$.

Case II: $h(\cdot)$ is strictly decreasing on S

Here $P(h(b) < h(X) < h(a)) = 1$ and $G(z) = P(h(X) \leq z)$, $z \in \mathbb{R}$. Clearly, for $z < h(b)$, $G(z) = 0$ and for $z \geq h(a)$, $G(z) = 1$. For $h(b) < z < h(a)$,

$$G(z) = P(X \geq h^{-1}(z)) = \int_{h^{-1}(z)}^{\infty} f(t)\mathrm{d}t = \int_{h^{-1}(z)}^{b} f(t)\mathrm{d}t = \int_{h(b)}^{z} f(h^{-1}(y)) \left|\frac{\mathrm{d}}{\mathrm{d}y} h^{-1}(y)\right| \mathrm{d}y.$$

Thus,

$$G(z) = \begin{cases} 0, & \text{if } z < h(b), \\ \int_{h(b)}^{z} f(h^{-1}(y)) \left|\frac{\mathrm{d}}{\mathrm{d}y} h^{-1}(y)\right| \mathrm{d}y, & \text{if } h(b) \leq z < h(a), \\ 1, & \text{if } z \geq h(a). \end{cases}$$

Since f is continuous on (a, b), it follows that $G(\cdot)$ is differentiable everywhere except possibly at $h(a)$ and $h(b)$. Moreover,

$$G'(z) = \begin{cases} f(h^{-1}(z)) \left|\frac{\mathrm{d}}{\mathrm{d}z} h^{-1}(z)\right|, & \text{if } h(b) < z < h(a), \\ 0, & \text{otherwise} \end{cases}$$

and

$$\int_{-\infty}^{\infty} G'(z)\mathrm{d}z = \int_{h(b)}^{h(a)} f(h^{-1}(z)) \left|\frac{\mathrm{d}}{\mathrm{d}z} h^{-1}(z)\right| \mathrm{d}z = \int_{a}^{b} f(t)\mathrm{d}t = 1.$$

Consequently, Z is a continuous r.v. with p.d.f.

$$g(z) = \begin{cases} f(h^{-1}(z)) \left|\frac{\mathrm{d}}{\mathrm{d}z} h^{-1}(z)\right|, & \text{if } h(b) < z < h(a), \\ 0, & \text{otherwise} \end{cases}$$

and support $T = [h(b), h(a)]$.

Combining Case I and Case II, we get the following result:

Theorem 8.1. *Let X be a continuous r.v. with p.d.f. $f(\cdot)$ and support $S = [a, b]$ for some $-\infty \leq a < b \leq \infty$. Suppose that $\{x \in \mathbb{R} : f(x) > 0\} = (a, b)$ and that f is continuous on (a, b). Let $h : \mathbb{R} \to \mathbb{R}$ be a function that is differentiable and strictly monotone on (a, b). Then, $Z = h(X)$ is a continuous r.v. with p.d.f*

$$\begin{cases} f(h^{-1}(z)) \left|\frac{\mathrm{d}}{\mathrm{d}z} h^{-1}(z)\right|, & \textit{if } z \in h((a, b)), \\ 0, & \textit{otherwise}, \end{cases}$$

and support $T = [\min\{h(a), h(b)\}, \max\{h(a), h(b)\}]$.

The following theorem is a generalization of the above result and can be proved on similar lines.

Theorem 8.2. *Let X be a continuous r.v. with p.d.f. $f(\cdot)$ and support $S = \bigcup_{i \in \Lambda} [a_i, b_i]$, where Λ is a countable set and $[a_i, b_i]$'s are disjoint intervals. Suppose that $\{x \in \mathbb{R} : f(x) > 0\} = \bigcup_{i \in \Lambda} (a_i, b_i)$ and that f is continuous in each (a_i, b_i), $i \in \Lambda$. Let $h : \mathbb{R} \to \mathbb{R}$ be a function that is differentiable and strictly monotone in each (a_i, b_i), $i \in \Lambda$ (h may be monotonic increasing in some (a_i, b_i) and monotonic decreasing in some (a_i, b_i)). Let $h_i^{-1}(\cdot)$ be the inverse function of h_i on (a_i, b_i), $i \in \Lambda$. Then, $Z = h(X)$ is a continuous r.v. with p.d.f.*

$$g(z) = \sum_{j \in \Lambda} f(h_j^{-1}(z)) \left|\frac{\mathrm{d}}{\mathrm{d}z} h_j^{-1}(z)\right| I_{h_j((a_j, b_j))}(z), \quad \textit{where} \quad I_{h_j((a_j, b_j))}(z) = \begin{cases} 1, & z \in h_j((a_j, b_j)), \\ 0, & \textit{otherwise}. \end{cases}.$$

Remark 8.3. *Theorem 8.1 and Theorem 8.2 hold even in situations where the function h is differentiable everywhere except possibly at a finite number of points in S.*

Example 8.4. *Let X be a r.v. with p.d.f.*

$$f(x) = \begin{cases} 3x^2, & 0 < x < 1, \\ 0, & \textit{otherwise}. \end{cases}$$

Find the p.d.f. and d.f. of $Y = 1/X^2$. What is the support of d.f. of Y.

Solution: The support of F is $[0,1]$ and $\{x \in \mathbb{R} : f(x) > 0\} = (0,1)$. Moreover, f is continuous on $(0,1)$ and $h(x) = 1/x^2$ is differentiable and strictly monotone on $(0,1)$.

$h((0,1)) = (1,\infty)$. Now

$$y = \frac{1}{x^2} \implies x = \frac{1}{\sqrt{y}}, \ \ i.e., \ \ h^{-1}(y) = \frac{1}{\sqrt{y}} \implies \frac{\mathrm{d}}{\mathrm{d}y} h^{-1}(y) = -\frac{1}{2y\sqrt{y}}, \ \ y \in (1,\infty).$$

Thus, $Y = 1/X^2$ is continuous r.v. with p.d.f $g(y)$ given by

$$\begin{aligned} g(y) &= f(h^{-1}(y)) \left| \frac{\mathrm{d}}{\mathrm{d}y} h^{-1}(y) \right| I_{h((0,1))}(y) \\ &= f(h^{-1}(y)) \left| \frac{\mathrm{d}}{\mathrm{d}y} h^{-1}(y) \right| I_{(1,\infty)}(y) \\ &= \begin{cases} \dfrac{3}{y} \cdot \dfrac{1}{2y\sqrt{y}}, & \text{if } y > 1, \\ 0, & \text{otherwise}, \end{cases} = \begin{cases} \dfrac{3}{2y^2\sqrt{y}}, & \text{if } y > 1, \\ 0, & \text{otherwise}. \end{cases} \end{aligned}$$

The d.f. of Y is

$$G(y) = \int_{-\infty}^{y} g(t)\mathrm{d}t = \begin{cases} 0, & \text{if } y < 1, \\ \displaystyle\int_1^y \frac{3}{2t^2\sqrt{t}}\mathrm{d}t, & \text{if } y > 1, \end{cases} = \begin{cases} 0, & \text{if } y < 1, \\ 1 - \dfrac{1}{y^{3/2}}, & \text{if } y > 1. \end{cases}$$

Clearly the support of G is $[1,\infty)$.

Example 8.5. *Let X be r.v. with p.d.f.*

$$f(x) = \begin{cases} |x|/2, & \textit{if } -1 < x < 1, \\ x/3, & \textit{if } 1 \le x \le 2, \\ 0, & \textit{otherwise} \end{cases}$$

and let $Y = X^2$.

(a) Find the p.d.f of Y directly and hence find the d.f. of Y.

(b) Find the d.f. of Y and hence find the p.d.f. of Y.

(c) Find the support of d.f. of Y.

Solution: (a) The support of F is $S = [-1,2]$ and we may take $S = [-1,0] \cup [0,2]$, $\{x \in \mathbb{R} : f(x) > 0\} = (-1,0) \cup (0,2)$. The p.d.f f is continuous on $(-1,0) \cup (0,1) \cup (1,2)$, $h(x) = x^2$ is differentiable on $(-1,0) \cup (0,2)$, $h(\cdot)$ is strictly decreasing on $(-1,0)$ and strictly increasing on $(0,2)$.

$h(x) = x^2$ is strictly decreasing on $S_1 = (-1, 0)$ with inverse function $h_1^{-1}(y) = -\sqrt{y}$, $y \in (0, 1)$, $h(S_1) = (0, 1)$.
$h(x) = x^2$ is strictly decreasing on $S_2 = (0, 2)$ with inverse function $h_2^{-1}(y) = \sqrt{y}$, $y \in (0, 4)$, $h(S_2) = (0, 4)$.

Thus, $Y = X^2$ is a continuous r.v. with p.d.f.

$$\begin{aligned}
g(y) &= f(h_1^{-1}(y)) \left| \frac{\mathrm{d}}{\mathrm{d}y} h_1^{-1}(y) \right| I_{(0,1)}(y) + f(h_2^{-1}(y)) \left| \frac{\mathrm{d}}{\mathrm{d}y} h_2^{-1}(y) \right| I_{(0,4)}(y) \\
&= f(-\sqrt{y}) \left| \frac{-1}{2\sqrt{y}} \right| I_{(0,1)}(y) + f(\sqrt{y}) \left| \frac{1}{2\sqrt{y}} \right| I_{(0,4)}(y) \\
&= \frac{1}{2\sqrt{y}} \left[f(-\sqrt{y}) I_{(0,1)}(y) + f(\sqrt{y}) I_{(0,4)}(y) \right] \\
&= \begin{cases} \frac{1}{2}, & \text{if } 0 < y < 1, \\ \frac{1}{6}, & \text{if } 1 < y < 4, \\ 0, & \text{otherwise.} \end{cases}
\end{aligned}$$

The d.f. of Y is

$$G(y) = P(X^2 \le y) = \int_{-\infty}^{y} g(t)\mathrm{d}t = \begin{cases} 0, & \text{if } y < 0, \\ \displaystyle\int_0^y \frac{\mathrm{d}t}{2}, & \text{if } 0 \le y < 1, \\ \displaystyle\int_0^1 \frac{\mathrm{d}t}{2} + \int_1^y \frac{\mathrm{d}t}{6}, & \text{if } 1 \le y < 4, \\ 1, & \text{if } y \ge 4. \end{cases} = \begin{cases} 0, & \text{if } y < 0, \\ \dfrac{y}{2}, & \text{if } 0 \le y < 1, \\ \dfrac{y+2}{6}, & \text{if } 1 \le y < 4, \\ 1, & \text{if } y \ge 4. \end{cases}$$

(b) The d.f. of Y is

$$G(y) = P(X^2 \le y) = \begin{cases} 0, & \text{if } y < 0, \\ P\{-\sqrt{y} \le X \le \sqrt{y}\}, & \text{if } y > 0. \end{cases}.$$

For $0 \le y < 1$,

$$G(y) = P\{-\sqrt{y} \le X \le \sqrt{y}\} = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{|x|}{2} \mathrm{d}x = \frac{y}{2}.$$

For $1 \le y < 4$ (so that $-2 < -\sqrt{y} \le -1$ and $1 \le \sqrt{y} \le 2$)

$$G(y) = P\{-\sqrt{y} \le X \le \sqrt{y}\} = \int_{-1}^{1} \frac{|x|}{2} \mathrm{d}x + \int_1^{\sqrt{y}} \frac{x}{3} \mathrm{d}x = \frac{y+2}{6}.$$

For $y \ge 4$, $G(y) = 1$. Therefore

$$G(y) = \begin{cases} 0, & \text{if } y < 0, \\ \dfrac{y}{2}, & \text{if } 0 \le y < 1, \\ \dfrac{y+2}{6}, & \text{if } 1 \le y < 4, \\ 1, & \text{if } y \ge 4. \end{cases}$$

Clearly G is differentiable everywhere except at finite number of points (0,1 and 4) and we may take

$$G'(y) = \begin{cases} 1/2, & \text{if } 0 < y < 1, \\ 1/6, & \text{if } 1 < y < 4, \\ 0, & \text{otherwise.} \end{cases}$$

Moreover, $\int_{-\infty}^{\infty} G'(y)\mathrm{d}y = \int_0^1 \frac{1}{2}\mathrm{d}y + \int_1^4 \frac{1}{6}\mathrm{d}y = 1$. Thus, Y is a continuous r.v. with p.d.f.

$$g(y) = \begin{cases} 1/2, & \text{if } 0 < y < 1, \\ 1/6, & \text{if } 1 < y < 4, \\ 0, & \text{otherwise.} \end{cases}$$

(c) The support of G is $[0, 4]$.

8.2. Expectation (or Mean) of Random Variables

Let X be a discrete r.v. with p.m.f. $f(\cdot)$ and support S. For any $x \in S$, $f(x)$ gives an idea about proportion of times we will observe the event $\{X = x\}$ if the experiment is repeated a large number of times. Thus $\sum_{x \in S} x f(x)$ represents the mean (or expected) value of r.v. X if the experiment is repeated a large number of times.

Similarly, if X is a continuous r.v. with p.d.f. $f(\cdot)$ then $\int_{-\infty}^{\infty} x f(x)\mathrm{d}x$ (provided the integral is finite) represents the mean (or expected) value of r.v. X.

Definition 8.6. *(a) Let X be a discrete r.v. with p.m.f. $f(\cdot)$ and support S. We say that the expected value of X (or the mean of X, which we denote by $E(X)$) is finite and equals*

$$E(X) = \sum_{x \in S} x f(x), \quad \textit{provided} \quad \sum_{x \in S} |x| f(x) < \infty.$$

(b) Let X be a continuous r.v. with p.d.f. $f(\cdot)$ and support S. We say that the expected value of X (or the mean of X, which we denote by $E(X)$) is finite and equals

$$E(X) = \int_{-\infty}^{\infty} x f(x)\mathrm{d}x, \quad \textit{provided} \quad \int_{-\infty}^{\infty} |x| f(x)\mathrm{d}x < \infty.$$

Example 8.7. *(a) Let X be a discrete r.v. with p.m.f.*

$$f(x) = \begin{cases} \frac{1}{2^x}, & \textit{if } x \in \{1, 2, 3, \dots\}, \\ 0, & \textit{otherwise}. \end{cases}.$$

Show that $E(X)$ is finite. Find $E(X)$.

(b) Let X be a r.v. with p.m.f.

$$f(x) = \begin{cases} \frac{3}{\pi^2 x^2}, & \textit{if } x \in \{\pm 1, \pm 2, \dots\}, \\ 0, & \textit{otherwise}. \end{cases}.$$

Show that $E(X)$ is not finite.

(c) Let X be a continuous r.v. with p.d.f. $f(x) = \frac{e^{-|x|}}{2}$, $-\infty < x < \infty$. Show that $E(X)$ is finite. Find $E(X)$.

(d) Let X be a continuous r.v. with p.d.f. $f(x) = \frac{1}{\pi(1 + x^2)}$, $-\infty < x < \infty$. Show that $E(X)$ is not finite.

Solution: (a) The support of the distribution is $S = \{1, 2, \dots\}$. Also,

$$\sum_{x \in S} |x| f(x) = \sum_{n=1}^{\infty} \frac{n}{2^n} = \sum_{n=1}^{\infty} a_n \ \text{(say)},$$

where $a_n = \dfrac{n}{2^n} > 0, \forall\, n = 1, 2, \dots$ and

$$\frac{a_{n+1}}{a_n} = \frac{n+1}{2n} \to \frac{1}{2} < 1, \ \text{ as } \ n \to \infty.$$

Thus by the ratio test $\sum_{x \in S} |x| f(x) = \sum_{n=1}^{\infty} \dfrac{n}{2^n} < \infty$. It can be seen that $E(X) = 2$ (Exercise).

(b) Here the support of the distribution is $S = \{\pm 1, \pm 2, \dots\}$.

$$\sum_{x \in S_X} |x| f(x) = \frac{6}{\pi^2} \sum_{n=1}^{\infty} \frac{1}{n} = \infty \implies E(X) \text{ is not finite.}$$

(c) We have

$$\int_{-\infty}^{\infty} |x| f(x) \mathrm{d}x = \int_{-\infty}^{\infty} |x| \frac{e^{-|x|}}{2} \mathrm{d}x = \int_{0}^{\infty} x e^{-x} \mathrm{d}x = 1 < \infty \implies E(X) \ \text{ is finite}$$

and

$$E(X) = \int_{-\infty}^{\infty} x f(x) \mathrm{d}x = \int_{-\infty}^{\infty} x \frac{e^{-|x|}}{2} \mathrm{d}x = 0.$$

(d) We have

$$\int_{-\infty}^{\infty} |x| f(x) \mathrm{d}x = \int_{-\infty}^{\infty} |x| \frac{1}{\pi(1+x^2)} \mathrm{d}x = \frac{2}{\pi} \int_{0}^{\infty} \frac{x}{1+x^2} \mathrm{d}x = \infty \implies E(X) \ \text{ is not finite.}$$

Example 8.8 (St. Petersburg Paradox)**.** *To make some money a gambler plays a sequence of fair games with the following strategy:*

In the first bet he bet Rs. 1 million. If the first bet is lost he doubles his bet in the second game. He keeps on doubling his bet until he wins a game. If the gambler has not won by the mth trial he bets Rs. 2^m million in the $(m+1)$th game. If he wins in kth game then

$$\textit{Investment}{=}1 + 2 + 4 + \cdots + 2^{k-1} = 2^k - 1 \ \textit{ million rupee}, \ \ \textit{win}{=}2^k \ \textit{ million rupee}.$$

Total earning if he wins on the kth game= 1 million rupee.

The above scheme seems to be foolproof for earning Rs. 1 million rupee. By this logic all gamblers should be billionaries!

X *: the amount of money bet on the last game (the game he wins). Then*

$$P(X = 2^k) = \frac{1}{2^{k+1}}, \ \ k = 0, 1, 2, \dots, \quad E(X) = \sum_{k=0}^{\infty} \frac{2^k}{2^{k+1}} = \infty \ \ (E(X) \textit{ is not finite}).$$

This implies enormous amount of money would be required.