Mathematical Statistics 1

Chapter 1. Introduction to Probability

1.8. Expectation of a Random Variables—Proofs of Theorems

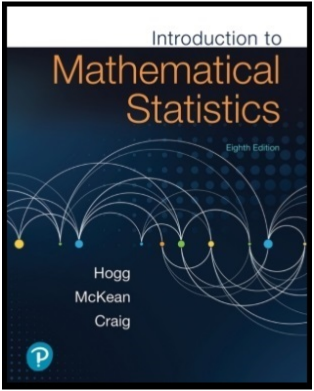

Table of contents

Theorem 1.8.1

Theorem 1.8.1. Let X be a random variable and let $Y = g(X)$ for some function g.

(a) Suppose X is continuous with probability density function $f_X(x)$. If

$$\int_{-\infty}^{\infty} |g(x)| f_X(x)\, dx < \infty,$$

then the expectation of Y exists and is $E(Y) = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx$.

(b) Suppose X is a discrete random variable with probability mass function $p_X(x)$. Suppose the support of X is denoted by $\mathcal{S}_X$. If $\sum_{x \in \mathcal{S}_X} |g(x)| p_X(x) < \infty$, then the expectation of Y exists and it is given by $E(Y) = \sum_{x \in \mathcal{S}_X} g(x) p_X(x)$.

Proof. The text states (page 62): "The proof of the continuous case requires some advanced results in analysis…"

Theorem 1.8.1

Theorem 1.8.1. Let X be a random variable and let $Y = g(X)$ for some function g.

(a) Suppose X is continuous with probability density function $f_X(x)$. If
$$\int_{-\infty}^{\infty} |g(x)| f_X(x)\, dx < \infty,$$
then the expectation of Y exists and is $E(Y) = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx$.

(b) Suppose X is a discrete random variable with probability mass function $p_X(x)$. Suppose the support of X is denoted by $\mathcal{S}_X$. If $\sum_{x \in \mathcal{S}_X} |g(x)| p_X(x) < \infty$, then the expectation of Y exists and it is given by $E(Y) = \sum_{x \in \mathcal{S}_X} g(x) p_X(x)$.

Proof. The text states (page 62): "The proof of the continuous case requires some advanced results in analysis..."

Theorem 1.8.1 (continued 1)

Proof (continued). This is done in Theorem 4.10.2 of my online notes for a class on Measure Theory Based Probability (not a formal ETSU class) on 4.10. Expectation; the necessary background is Real Analysis 1 and 2 (MATH 5210/5220)...at least it doesn't require functional analysis!

Since $\sum_{x \in \mathcal{S}_X} |g(x)| p_X(x)$ converges, it follows from the Rearrangement Theorem for Absolutely Convergent Series then

$$\sum_{x \in \mathcal{S}_X} |g(x) p_X(x) = \sum_{y \in \mathcal{S}_Y} \sum_{\{x \in \mathcal{S}_X | g(x) = y\}} |g(x)| p_X(x)$$

$$\text{since } \{x \in \mathbb{R} \mid x \in \mathcal{S}_X\} = \{x \in \mathbb{R} \mid s \in \mathcal{S}_X, g(x) \neq 0\}$$

$$\cup \{x \in \mathbb{R} \mid x \in \mathcal{S}_X, g(x) = 0\}$$

$$= \sum_{y \in \mathcal{S}_Y} \sum_{\{x \in \mathcal{S}_X X | g(x) = y\}} |y| p_X(x)$$

$$= \sum_{y \in \mathcal{S}_Y} |y| \sum_{\{x \in \mathcal{S}_X | g(x) = y\}} p_X(x)$$

Theorem 1.8.1 (continued 1)

Proof (continued). This is done in Theorem 4.10.2 of my online notes for a class on Measure Theory Based Probability (not a formal ETSU class) on 4.10. Expectation; the necessary background is Real Analysis 1 and 2 (MATH 5210/5220)...at least it doesn't require functional analysis! Since $\sum_{x\in\mathcal{S}_X}|g(x)|p_X(x)$ converges, it follows from the Rearrangement Theorem for Absolutely Convergent Series then

$$\sum_{x\in\mathcal{S}_X}|g(x)p_X(x) = \sum_{y\in\mathcal{S}_Y}\sum_{\{x\in\mathcal{S}_X|g(x)=y\}}|g(x)|p_X(x)$$

$$\text{since } \{x\in\mathbb{R}\mid x\in\mathcal{S}_X\} = \{x\in\mathbb{R}\mid s\in\mathcal{S}_X, g(x)\neq 0\}$$

$$\cup\{x\in\mathbb{R}\mid x\in\mathcal{S}_X, g(x)=0\}$$

$$= \sum_{y\in\mathcal{S}_Y}\sum_{\{x\in\mathcal{S}_X X|g(x)=y\}}|y|p_X(x)$$

$$= \sum_{y\in\mathcal{S}_Y}|y|\sum_{\{x\in\mathcal{S}_X|g(x)=y\}}p_X(x)$$

Theorem 1.8.1 (continued 2)

Proof (continued).

$$\begin{aligned}\sum_{x \in \mathcal{S}_X} |g(x) p_X(x) &= \sum_{y \in \mathcal{S}_Y} |y| \sum_{\{x \in \mathcal{S}_X | g(x) = y\}} p_X(x) \\ &= \sum_{y \in \mathcal{S}_Y} |y| p_Y(y)\end{aligned}$$

where the last equality holds since

$$p_Y(y) = P(Y = y) = P(g(X) = y) = P(\{x \in X \mid g(x) = y\})$$
$$= p_X(\{x \in X \mid g(x) = y\}) = p_X(\{x \in X \mid g(x) = y\})$$
$$= \sum_{\{x \in \mathcal{S}_X | g(x) = y\}} p_X(x).$$

Now $\sum_{x \in \mathcal{S}_X} g(x) p_X(x)$, and so $\sum_{y \in \mathcal{S}_Y} y p_Y(y)$, converge absolutely by hypothesis (and hence converge by "The Absolute Convergence Test" mentioned above).

Theorem 1.8.1 (continued 2)

Proof (continued).

$$\sum_{x\in\mathcal{S}_X} |g(x)p_X(x) = \sum_{y\in\mathcal{S}_Y} |y| \sum_{\{x\in\mathcal{S}_X|g(x)=y\}} p_X(x) = \sum_{y\in\mathcal{S}_Y} |y|p_Y(y)$$

where the last equality holds since

$$p_Y(y) = P(Y=y) = P(g(X)=y) = P(\{x\in X \mid g(x)=y\})$$

$$= p_X(\{x\in X \mid g(x)=y\}) = p_X(\{x\in X \mid g(x)=y\})$$

$$= \sum_{\{x\in\mathcal{S}_X|g(x)=y\}} p_X(x).$$

Now $\sum_{x\in\mathcal{S}_X} g(x)p_X(x)$, and so $\sum_{y\in\mathcal{S}_Y} yp_Y(y)$, converge absolutely by hypothesis (and hence converge by "The Absolute Convergence Test" mentioned above).

Theorem 1.8.1 (continued 3)

Proof (continued). So we similarly have

$$E(Y) = \sum_{y \in \mathcal{S}_Y} y p_Y(y) = \sum_{y \in \mathcal{S}_Y} y \sum_{y \in \mathcal{S}_Y} y \sum_{\{x \in \mathcal{S}_X \mid g(x) = y\}} p_X(x)$$

$$= \sum_{y \in \mathcal{S}_Y} \sum_{\{x \in \mathcal{S}_X \mid g(x) = y\}} g(x) p_X(x) = \sum_{x \in \mathcal{S}_X} g(x) p_X(x),$$

as claimed. □

Exercise 1.8.9

Exercise 1.8.9. Let $f(x) = 2x$, $0 < x < 1$, zero elsewhere, be the pdf of X.

(a) Compute $E(1/X)$.

(b) Find the cdf and the pdf of $Y = 1/X$.

(c) Compute $E(Y)$ directly from the pdf of Y.

Solution. **(a)** We take $g(x) = 1/x$ and apply Theorem 1.8.1(a). First, notice that

$$\int_{-\infty}^{\infty} |g(x)| f_X(x)\, dx = \int_{-\infty}^{\infty} \frac{1}{|x|} f(x)\, dx = \int_{-\infty}^{0} 0\, dx + \int_{0}^{1} \left(\frac{1}{x} 2x\right) dx + \int_{1}^{\infty} 0\, dx$$

$= \int_0^1 \left(2\frac{x}{x}\right) dx = \lim_{a \to 0^+} \int_a^1 \left(2\frac{x}{x}\right) dx$ since we have an improper integral

Exercise 1.8.9

Exercise 1.8.9. Let $f(x) = 2x$, $0 < x < 1$, zero elsewhere, be the pdf of X.

(a) Compute $E(1/X)$.

(b) Find the cdf and the pdf of $Y = 1/X$.

(c) Compute $E(Y)$ directly from the pdf of Y.

Solution. **(a)** We take $g(x) = 1/x$ and apply Theorem 1.8.1(a). First, notice that

$$\int_{-\infty}^{\infty} |g(x)| f_X(x)\,dx = \int_{-\infty}^{\infty} \frac{1}{|x|} f(x)\,dx = \int_{-\infty}^{0} 0\,dx + \int_{0}^{1} \left(\frac{1}{x} 2x\right)\,dx$$

$$+ \int_{1}^{\infty} 0\,dx$$

$= \int_0^1 \left(2\frac{x}{x}\right)\,dx = \lim_{a \to 0^+} \int_a^1 \left(2\frac{x}{x}\right)\,dx$ since we have an improper integral

Exercise 1.8.9 (continued 1)

Solution (continued).

$$\int_{-\infty}^{\infty} |g(x)| f_X(x)\, dx = \cdots = \lim_{a \to 0^+} \int_a^1 2\, dx = 2 < \infty,$$

as required. Next, $E(1/X) = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx = 2$, as above. □

(b) Since the support of X is $0 < x < 1$ then the support of $Y = 1/X$ is $1 < y < \infty$. The cdf of $Y = 1/X$ is

$$f_Y(y) = P(Y \leq y) = P(1/X \leq y) = P(X \geq 1/y)$$

$$= 1 - P(X < 1/y) = 1 - P(X \leq 1/y)$$

where the last equality holds since $P(X = 1/y) = 0$ because we have a continuous random variable.

Exercise 1.8.9 (continued 1)

Solution (continued).

$$\int_{-\infty}^{\infty} |g(x)| f_X(x)\, dx = \cdots = \lim_{a \to 0^+} \int_a^1 2\, dx = 2 < \infty,$$

as required. Next, $E(1/X) = \displaystyle\int_{-\infty}^{\infty} g(x) f_X(x)\, dx = 2$, as above. □

(b) Since the support of X is $0 < x < 1$ then the support of $Y = 1/X$ is $1 < y < \infty$. The cdf of $Y = 1/X$ is

$$f_Y(y) = P(Y \leq y) = P(1/X \leq y) = P(X \geq 1/y)$$

$$= 1 - P(X < 1/y) = 1 - P(X \leq 1/y)$$

where the last equality holds since $P(X = 1/y) = 0$ because we have a continuous random variable.

Exercise 1.8.9 (continued 2)

Solution (continued).
Now for $1 < y < \infty$ we have $0 < 1/y < 1$ and so

$$P(X \leq 1/y) = \int_{-\infty}^{1/y} f(x)\,dx = \int_{-\infty}^{0} 0\,dx + \int_{0}^{1/y} 2x\,dx = 0 + x^2\big|_0^{1/y} = 1/y^2.$$

Therefore the cdf of $Y = 1/X$ is $F_Y(y) = 1 - P(X \leq 1/y) = 1 - 1/y^2$ for $1 < y < \infty$. The pdf of Y is then

$$f_Y(y) = \frac{d}{dy}[F_Y(y)]\frac{d}{dy}[1 - 1/y^2] = 2/y^3,\ 1 < y < \infty.$$ □

(c) With pdf $f_y(y) = 2/y^3$ from part (b), we have the expectation

$$E(Y) = \int_1^\infty y f_Y(y)\,dy = \int_1^\infty y\left(\frac{2}{y^3}\right)dy = \int_1^\infty \frac{2}{y^2}\,dy$$

$$= \frac{-2}{y}\bigg|_1^\infty = \lim_{b\to\infty}\left(\frac{-2}{b} - \frac{-2}{1}\right) = 2,$$

in agreement with part (a). □

Exercise 1.8.9 (continued 2)

Solution (continued).
Now for $1 < y < \infty$ we have $0 < 1/y < 1$ and so

$$P(X \leq 1/y) = \int_{-\infty}^{1/y} f(x)\,dx = \int_{-\infty}^{0} 0\,dx + \int_{0}^{1/y} 2x\,dx = 0 + x^2\big|_0^{1/y} = 1/y^2.$$

Therefore the cdf of $Y = 1/X$ is $F_Y(y) = 1 - P(X \leq 1/y) = 1 - 1/y^2$ for $1 < y < \infty$. The pdf of Y is then

$$f_Y(y) = \frac{d}{dy}[F_Y(y)]\frac{d}{dy}[1 - 1/y^2] = 2/y^3,\ 1 < y < \infty.$$ □

(c) With pdf $f_y(y) = 2/y^3$ from part (b), we have the expectation

$$E(Y) = \int_1^\infty y f_Y(y)\,dy = \int_1^\infty y\left(\frac{2}{y^3}\right)\,dy = \int_1^\infty \frac{2}{y^2}\,dy$$

$$= \frac{-2}{y}\bigg|_1^\infty = \lim_{b\to\infty}\left(\frac{-2}{b} - \frac{-2}{1}\right) = 2,$$

in agreement with part (a). □

Theorem 1.8.2

Theorem 1.8.2. Let $g_1(X)$ and $g_2(X)$ be functions of a random variable X. Suppose the expectations of $g_1(X)$ and $g_2(X)$ exist. Then for any constants k_1 and k_2 the expectation of $k_1 g_1(X) + k_2 g_2(X)$ exists and it is given by

$$E(k_1 g_1(X) + k_2 g_2(X)) = k_1 E(g_1(X)) + k_2 E(g_2(X)).$$

Proof. For the continuous case, we have by the Triangle Inequality on $\mathbb{R}$

$$\int_{-\infty}^{\infty} |f_1 g_1(x) + k_2 g_2(x)| f_X(f)\, dx \leq \int_{-\infty}^{\infty} (|k_1||g_1(x)| + |k_2||g_2(x)|) f_X(x)\, dx$$

$$= |k_1| \int_{-\infty}^{\infty} |g_1(x)| f_X(x)\, dx + |k_2| \int_{-\infty}^{\infty} |g_2(x)| f_X(x)\, dx < \infty$$

where the boundedness follows by the hypothesis that the expectations of $g_1(X)$ and $g_2(X)$ exist. Therefore the expectation of $k_1 g_1(X) + k_2 g_2(X)$ is defined.

Theorem 1.8.2

Theorem 1.8.2. Let $g_1(X)$ and $g_2(X)$ be functions of a random variable X. Suppose the expectations of $g_1(X)$ and $g_2(X)$ exist. Then for any constants k_1 and k_2 the expectation of $k_1 g_1(X) + k_2 g_2(X)$ exists and it is given by

$$E(k_1 g_1(X) + k_2 g_2(X)) = k_1 E(g_1(X)) + k_2 E(g_2(X)).$$

Proof. For the continuous case, we have by the Triangle Inequality on $\mathbb{R}$

$$\int_{-\infty}^{\infty} |f_1 g_1(x) + k_2 g_2(x)| f_X(f)\, dx \leq \int_{-\infty}^{\infty} (|k_1||g_1(x)| + |k_2||g_2(x)|) f_X(x)\, dx$$

$$= |k_1| \int_{-\infty}^{\infty} |g_1(x)| f_X(x)\, dx + |k_2| \int_{-\infty}^{\infty} |g_2(x)| f_X(x)\, dx < \infty$$

where the boundedness follows by the hypothesis that the expectations of $g_1(X)$ and $g_2(X)$ exist. Therefore the expectation of $k_1 g_1(X) + k_2 g_2(X)$ is defined.

Theorem 1.8.2 (continued 1)

Proof (continued). We now have, by the linearity of the integral,

$$E(k_1 g_1(X) + k_2 g_2(X)) = \int_{-\infty}^{\infty} (k_1 g_1(x) + k_2 g_2(x)) f_X(x)\,dx$$

$$= k_1 \int_{-\infty}^{\infty} g_1(x) f_X(x)\,dx + k_2 \int_{-\infty}^{\infty} g_2(x) f_X(x)\,dx$$

$$= k_1 E(g_1(X)) + k_2 E(g_2(X)),$$

as claimed.

For the discrete case, by the Triangle Inequality on $\mathbb{R}$ we have

$$\sum_{x \in \mathcal{S}_X} |k_1 g_1(x) + k_2 g_2(x)| f_X(x) \leq \sum_{x \in \mathcal{S}_X} (|k_1||g_1(x)| + |k_2||g_2(x)|) f_X(x)$$

$$= |k_1| \sum_{x \in \mathcal{S}_X} |g_1(x)| f_X(x) + |k_x| \sum_{x \in \mathcal{S}_X} |g_2(x)| f_X(x) < \infty$$

where the boundedness follows by the hypothesis that the expectations of $g_1(X)$ and $g_2(X)$ exist.

Theorem 1.8.2 (continued 1)

Proof (continued). We now have, by the linearity of the integral,

$$E(k_1 g_1(X) + k_2 g_2(X)) = \int_{-\infty}^{\infty} (k_1 g_1(x) + k_2 g_2(x)) f_X(x)\, dx$$

$$= k_1 \int_{-\infty}^{\infty} g_1(x) f_X(x)\, dx + k_2 \int_{-\infty}^{\infty} g_2(x) f_X(x)\, dx$$

$$= k_1 E(g_1(X)) + k_2 E(g_2(X)),$$

as claimed.

For the discrete case, by the Triangle Inequality on $\mathbb{R}$ we have

$$\sum_{x \in \mathcal{S}_X} |k_1 g_1(x) + k_2 g_2(x)| f_X(x) \leq \sum_{x \in \mathcal{S}_X} (|k_1||g_1(x)| + |k_2||g_2(x)|) f_X(x)$$

$$= |k_1| \sum_{x \in \mathcal{S}_X} |g_1(x)| f_X(x) + |k_x| \sum_{x \in \mathcal{S}_X} |g_2(x)| f_X(x) < \infty$$

where the boundedness follows by the hypothesis that the expectations of $g_1(X)$ and $g_2(X)$ exist.

Theorem 1.8.2 (continued 2)

Theorem 1.8.2. Let $g_1(X)$ and $g_2(X)$ be functions of a random variable X. Suppose the expectations of $g_1(X)$ and $g_2(X)$ exist. Then for any constants k_1 and k_2 the expectation of $k_1g_1(X) + k_2g_2(X)$ exists and it is given by

$$E(k_1g_1(X) + k_2g_2(X)) = k_1E(g_1(X)) + k_2E(g_2(X)).$$

Proof (continued). Therefore the expectation of $k_1g_1(X) + k_2g_2(X)$ is defined. We now have by the absolute summability and the Rearrangement Theorem for Absolutely Convergent Series that

$$E(k_1g_1(x) + k_2g_2(x)) = \sum_{x \in \mathcal{S}_X} (k_1g_1(x) + k_sg_2(x))f_X(x)$$

$$= k_1 \sum_{x \in \mathcal{S}_X} g_1(x)f_X(x) + k_2 \sum_{x \in \mathcal{S}_X} g_2(x)f_X(x)$$

$$= k_1E(g_1(x)) + k_2E(g_2(X)),$$

as claimed. □

Exercise 1.8.7

Exercise 1.8.7. Let X have the pdf $f(x) = 3x^2$, $0 < x < 1$, zero elsewhere. Consider a random rectangle where sides are X and $(1 - X)$. Determine the expected value of the area of the rectangle.

Solution. We let $A = X - X^2$ be the random variable representing the area of the rectangle. By Theorem 1.8.2 we have
$E(A) = E(X - X^2) = E(X) - E(X^2)$ where

$$E(X) = \int_{-\infty}^{\infty} xf(x), dx = \int_0^1 x(3x^2)\,dx = \int_0^1 3x^3\,dx = \frac{3}{4}x^4\Big|_0^1 = \frac{3}{4}$$

and by Theorem 1.8.1(a)

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx = \int_0^1 x^2(3x^2)\,dx = \int_0^1 3x^4\,dx = \frac{3}{5}x^5\Big|_0^1 = \frac{3}{5}.$$

So the expected area is $E(A) = E(X) - E(X^2) = 3/4 - 3/5 = 3/20$. □

Exercise 1.8.7

Exercise 1.8.7. Let X have the pdf $f(x) = 3x^2$, $0 < x < 1$, zero elsewhere. Consider a random rectangle where sides are X and $(1 - X)$. Determine the expected value of the area of the rectangle.

Solution. We let $A = X - X^2$ be the random variable representing the area of the rectangle. By Theorem 1.8.2 we have
$E(A) = E(X - X^2) = E(X) - E(X^2)$ where

$$E(X) = \int_{-\infty}^{\infty} xf(x), dx = \int_0^1 x(3x^2)\,dx = \int_0^1 3x^3\,dx = \frac{3}{4}x^4\bigg|_0^1 = \frac{3}{4}$$

and by Theorem 1.8.1(a)

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx = \int_0^1 x^2(3x^2)\,dx = \int_0^1 3x^4\,dx = \frac{3}{5}x^5\bigg|_0^1 = \frac{3}{5}.$$

So the expected area is $E(A) = E(X) - E(X^2) = 3/4 - 3/5 = 3/20$. □