

CSE 107
Probability and Statistics for Engineers
Expected Value Rules

Discrete Case

The simplest version of the Expected Value Rule occurs when we compose a discrete random variable with a function of one variable.

Theorem 1

Let $X: \Omega \rightarrow \mathbb{R}$ be a discrete random variable, and $g: \mathbb{R} \rightarrow \mathbb{R}$ be a function. Then

$$E[g(X)] = \sum_{x \in \text{range}(X)} g(x)p_X(x)$$

where $p_X(x)$ denotes the PMF of X .

Proof:

Let $Y = g \circ X = g(X)$ and observe that for any $y \in \mathbb{R}$,

$$\begin{aligned} \{Y = y\} &= \{\omega \in \Omega \mid Y(\omega) = y\} \\ &= \{\omega \in \Omega \mid g(X(\omega)) = y\} \\ &= \bigcup_{\{x \mid g(x)=y\}} \{\omega \in \Omega \mid X(\omega) = x\} \\ &= \bigcup_{\{x \mid g(x)=y\}} \{X = x\}. \end{aligned}$$

The last expression is a disjoint union, so by axiom (iii),

$$\begin{aligned} p_Y(y) &= P(Y = y) \\ &= \sum_{\{x \mid g(x)=y\}} P(X = x) \\ &= \sum_{\{x \mid g(x)=y\}} p_X(x). \end{aligned}$$

Therefore

$$\begin{aligned} E[g(X)] &= E[Y] \\ &= \sum_{y \in \text{range}(Y)} y \cdot p_Y(y) \end{aligned}$$

$$\begin{aligned}
&= \sum_y y \cdot \left(\sum_{\{x|g(x)=y\}} p_X(x) \right) \\
&= \sum_y \sum_{\{x|g(x)=y\}} y \cdot p_X(x) \\
&= \sum_y \sum_{\{x|g(x)=y\}} g(x) p_X(x) \\
&= \sum_x g(x) p_X(x).
\end{aligned}$$

■

To establish the two-variable version, we use the *total expectation theorem for discrete random variables*. This says that if events $A_1, A_2, \dots, A_n$ form a partition of the sample space Ω , and Z is a discrete random variable, then

$$E[Z] = \sum_{i=1}^n P(A_i) E[Z|A_i]$$

This formula remains valid even if the sequence of events forming the partition is infinite, making the above expression an infinite series.

Theorem 2

Let X and Y be discrete random variables, $g: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ be a function of two variables, and $Z = g(X, Y)$. Then

$$E[Z] = E[g(X, Y)] = \sum_x \sum_y g(x, y) p_{X,Y}(x, y)$$

where the sum is over all $x \in \text{range}(X)$ and $y \in \text{range}(Y)$.

Proof:

Observe that the events $\{Y = y\}$ for all $y \in \text{range}(Y)$, form a partition of Ω . The total expectation theorem gives

$$\begin{aligned}
E[Z] &= \sum_{y \in \text{range}(Y)} P(Y = y) E[Z|Y = y] \\
&= \sum_y p_Y(y) E[g(X, Y)|Y = y]
\end{aligned}$$

$$\begin{aligned}
&= \sum_y p_Y(y) E[g(X, y) | Y = y] \\
&= \sum_y p_Y(y) \sum_x g(x, y) p_{X|Y}(x|y) \\
&= \sum_y \sum_x g(x, y) p_Y(y) p_{X|Y}(x|y) \\
&= \sum_x \sum_y g(x, y) p_{X,Y}(x, y)
\end{aligned}$$

as required. Line 4 above follows since for each fixed $y \in \text{range}(Y)$, $g(x, y)$ can be considered a function of the single variable x , and from the 1-variable, conditional expected value rule. ■

Continuous Case

For a continuous random variable $X: \Omega \rightarrow \mathbb{R}$, we first prove the following preliminary result.

Theorem 3

Let $X: \Omega \rightarrow \mathbb{R}$ be a continuous random variable. Then

$$E[X] = \int_0^{\infty} P(X > x) dx - \int_0^{\infty} P(X < -x) dx$$

Proof:

Consider the first term.

$$\begin{aligned}
\int_0^{\infty} P(X > x) dx &= \int_0^{\infty} \left(\int_x^{\infty} f_X(t) dt \right) dx \\
&= \int_0^{\infty} \left(\int_0^t f_X(t) dx \right) dt \\
&= \int_0^{\infty} f_X(t) \left(\int_0^t dx \right) dt \\
&= \int_0^{\infty} t f_X(t) dt
\end{aligned}$$

The second line follows by integrating over the infinite triangle $\{(x, t) \mid 0 \leq x \leq t\}$ in two ways. Similarly, the second term becomes

$$\begin{aligned}
 \int_0^\infty P(X < -x) dx &= \int_0^\infty \left(\int_{-\infty}^{-x} f_X(t) dt \right) dx \\
 &= \int_{-\infty}^0 \left(\int_0^{-t} f_X(t) dx \right) dt \\
 &= \int_{-\infty}^0 f_X(t) \left(\int_0^{-t} dx \right) dt \\
 &= \int_{-\infty}^0 (-t) f_X(t) dt
 \end{aligned}$$

Again, the second line follows by integrating over the infinite triangle $\{(x, t) \mid 0 \leq x \leq -t\}$ in two ways. Therefore

$$\begin{aligned}
 E[X] &= \int_{-\infty}^\infty t f_X(t) dt \\
 &= \int_0^\infty t f_X(t) dt + \int_{-\infty}^0 t f_X(t) dt \\
 &= \int_0^\infty t f_X(t) dt - \int_{-\infty}^0 (-t) f_X(t) dt \\
 &= \int_0^\infty P(X > x) dx - \int_0^\infty P(X < -x) dx
 \end{aligned}$$

as claimed. ■

Theorem 4

Let $X: \Omega \rightarrow \mathbb{R}$ be a continuous random variable, and $g: \mathbb{R} \rightarrow \mathbb{R}$ be a function. Then

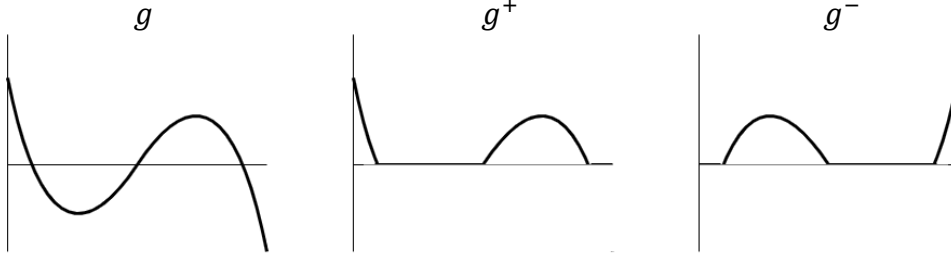
$$E[g(X)] = \int_{-\infty}^\infty g(x) f_X(x) dx$$

Proof:

First define the two non-negative functions $g^+(x) = \max(g(x), 0)$ and $g^-(x) = \max(-g(x), 0)$ called the *positive part* and the *negative part* of $g(x)$, respectively. Then

$$g(x) = g^+(x) - g^-(x)$$

For instance



Now let $U = g(X)$. Then Theorem 3 says

$$E[U] = \int_0^{\infty} P(U > u) du - \int_0^{\infty} P(U < -u) du$$

i.e.

$$E[g(X)] = \int_0^{\infty} P(g(X) > u) du - \int_0^{\infty} P(g(X) < -u) du$$

The first term is

$$\begin{aligned} \int_0^{\infty} P(g(X) > u) du &= \int_0^{\infty} \int_{\{x \mid g(x) > u\}} f_X(x) dx du \\ &= \int_{-\infty}^{\infty} \int_{\{u \mid 0 \leq u < g(x)\}} f_X(x) du dx \\ &= \int_{-\infty}^{\infty} \left(\int_0^{g(x)} du \right) f_X(x) dx \\ &= \int_{-\infty}^{\infty} g^+(x) f_X(x) dx \end{aligned}$$

The second term is

$$\begin{aligned}
\int_0^\infty P(g(X) < -u) du &= \int_0^\infty \int_{\{x \mid g(x) < -u\}} f_X(x) dx du \\
&= \int_{-\infty}^\infty \int_{\{u \mid 0 \leq u < -g(x)\}} f_X(x) du dx \\
&= \int_{-\infty}^\infty \left(\int_0^{-g(x)} du \right) f_X(x) dx \\
&= \int_{-\infty}^\infty g^-(x) f_X(x) dx
\end{aligned}$$

Thus

$$\begin{aligned}
E[g(X)] &= \int_{-\infty}^\infty g^+(x) f_X(x) dx - \int_{-\infty}^\infty g^-(x) f_X(x) dx \\
&= \int_{-\infty}^\infty (g^+(x) - g^-(x)) f_X(x) dx \\
&= \int_{-\infty}^\infty g(x) f_X(x) dx
\end{aligned}$$

as claimed. ■

For the two variable version, we use the *total expectation theorem for continuous random variables*. If Y and Z are continuous random variables, then $\{Y = y \mid y \in \text{range}(Y)\}$ forms a partition of Ω , and

$$E[Z] = \int_{-\infty}^\infty f_Y(y) E[Z|Y = y] dy$$

Theorem 5

Let X and Y be continuous random variables, $g: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ be a function of two variables, and let Z be the composition $Z = g(X, Y)$. Then

$$E[Z] = E[g(X, Y)] = \iint_{\mathbb{R}^2} g(x, y) f_{X,Y}(x, y) dx dy$$

Proof:

By the total expectation theorem above we have

$$\begin{aligned} E[Z] &= \int_{-\infty}^{\infty} f_Y(y) E[Z|Y = y] dy \\ &= \int_{-\infty}^{\infty} f_Y(y) E[g(X, Y)|Y = y] dy \\ &= \int_{-\infty}^{\infty} f_Y(y) E[g(X, y)|Y = y] dy \\ &= \int_{-\infty}^{\infty} f_Y(y) \left(\int_{-\infty}^{\infty} g(x, y) f_{X|Y}(x|y) dx \right) dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_Y(y) f_{X|Y}(x|y) dx dy \\ &= \iint_{\mathbb{R}^2} g(x, y) f_{X,Y}(x, y) dx dy \end{aligned}$$

Line 4 above follows from the 1-variable, conditional expected value rule, and because for fixed $y \in \mathbb{R}$, $g(x, y)$ can be considered a function of the single variable x . ■