Expectation of a discrete random variable

Definition: Suppose X is a discrete random variable and its probability function (p.f.) is f. Then the expectation of X, denoted by EX, is defined as

$$EX = \sum_x x f(x). \quad (1)$$

Remark: When X has infinite many possible values, then EX is a sum of an infinite series. There is a technical condition in order to well define the sum of an infinite series. We only talk about EX if

$$\sum_x |x| f(x) < \infty. \quad (2)$$

If

$$\sum_x |x| f(x) = \infty \quad (3)$$

we say EX does not exist. In this class, we do not need to worry about this technical condition.

Examples

Ex1: Let A be some event and $X = I(A\ occurs)$. That is, $X = 1$ if A occurs and $X = 0$ if not. Let's compute EX. By the definition of EX,

$$EX = 1 \cdot f(1) + 0 \cdot f(0) = f(1)$$

We know $f(1) = P(X = 1) = P(A)$. Thus $EX = P(A)$.

Ex2: Roll a die. Let X be the number we see. What is EX? By the definition of EX,

$$\begin{aligned} EX &= 1 \cdot f(1) + 2 \cdot f(2) + 3 \cdot f(3) \\ &\quad + 4 \cdot f(4) + 5 \cdot f(5) + 6 \cdot f(6) \end{aligned}$$

We can assume an equally-likely probability model, $f(i) = \frac{1}{6}$, $1 \leq i \leq 6$.

$$EX = \sum_{i=1}^{6} i\frac{1}{6} = 3.5.$$

Ex3

Game of roulette. You bet on 38 numbers: $00, 0, 1, 2, \ldots, 36$. You win 1 dollar if the number is odd and lose 1 dollar if the number is even. Let X be the money you win.

$X = 1$ or -1.

$$f(1) = P(X = 1) = \frac{18}{38}$$

$$f(-1) = P(X = -1) = \frac{20}{38}$$

$$EX = 1 \cdot f(1) + (-1) \cdot f(-1) = \frac{18}{38} - \frac{20}{38} = -\frac{1}{19}$$

You are expected to loss 5 cents per game.

Expectation of a continuous random variable

Definition: Suppose X is a continuous random variable and its probability density function (p.d.f.) is f. Then the expectation of X, denoted by EX, is defined as

$$EX = \int_{-\infty}^{\infty} xf(x)dx. \quad (1)$$

Remark: We only talk about EX if

$$\int_{-\infty}^{\infty} |x|f(x)dx < \infty. \quad (2)$$

If

$$\int_{-\infty}^{\infty} |x|f(x) = \infty \quad (3)$$

we say EX does not exist. In this class, we do not need to worry about this technical condition.

Examples

Ex1: Let $X \sim Unif(0,1)$.

$$EX = \int_0^1 x \cdot 1 dx = \frac{1}{2}x^2|_0^1 = \frac{1}{2}.$$

Ex2: Let $X \sim N(0,1)$

$$EX = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 0$$

The integral is zero because $g(x) = x\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ is an odd function, that is,

$$g(-x) + g(x) = 0.$$

Note that

$$\int_{-\infty}^{\infty} g(x)dx = \int_0^{\infty} [g(x) + g(-x)]dx.$$

Thus if $g(x) + g(-x) = 0$ then $\int_{-\infty}^{\infty} g(x)dx = 0$.

Ex3

Let X be a continuous random variable with pdf $f(x) = e^{-x}I(x > 0)$.

$$\begin{aligned} EX &= \int_0^\infty x \cdot e^{-x} dx \\ \text{integration by part} &= (-e^{-x})x|_0^\infty - \int_0^\infty (-e^{-x}) dx \\ &= \int_0^\infty e^{-x} dx \\ &= (-e^{-x})|_0^\infty = 1. \end{aligned}$$

Remark: We have discussed the expectation of discrete and continuous random variables. Often we call EX the mean of X or the mean of the distribution of X. For example, the mean of $N(0,1)$ is 0 and the mean of $Unif(0,1)$ is $1/2$.

Expectation of $r(X)$

Suppose X is a random variable and $r(t)$ is a function, e.g., $r(t) = t^2$. Note that $r(X)$ is another random variable. We want to compute the expectation of $r(X)$.

Suppose X is a discrete RV and its probability function is f, then

$$E[r(X)] = \sum_x r(x)f(x).$$

Suppose X is a continuous RV and its pdf is $f(x)$, then

$$E[r(X)] = \int_{-\infty}^{\infty} r(x) \cdot f(x)dx.$$

Ex1. Let $X \sim Unif(0, 1)$. Compute $E[X^2]$.

$$E[X^2] = \int_0^1 x^2 \cdot 1dx = \frac{1}{3}x^3|_0^1 = \frac{1}{3}.$$

Ex2. Let X take values at $1, -1, 0$ and $f(1) = f(-1) = 1/4$ and $f(0) = 1/2$. Compute $E[X^2]$.

$$\begin{aligned} E[X^2] &= 1^2 f(1) + (-1)^1 f(-1) + 0^2 f(0) \\ &= 1/4 + 1/4 + 0 = 1/2. \end{aligned}$$

Ex3. Let $X \sim N(0, 1)$. Compute $E[X^2]$.

$$\begin{aligned} E[X^2] &= \int_{-\infty}^{\infty} x^2 \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ &= \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi}} d(-e^{-\frac{x^2}{2}}) \\ &\quad \text{use integral by part} \\ &= (-\frac{x}{\sqrt{2\pi}} e^{-\frac{x^2}{2}})|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} -e^{-\frac{x^2}{2}} d(\frac{x}{\sqrt{2\pi}}) \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ &\quad \text{integral of the pdf of standard normal} \\ &= 1. \end{aligned}$$

Variance

Let X be a random variable. Suppose EX exists and we write $EX = \mu$. The Variance of X is defined as

$$Var(X) = E[(X - \mu)^2].$$

Often σ^2 is used to denote the variance of X.

Remark: We only talk about the variance of X when $E[(X - \mu)^2]$ is well defined.

Ex1: Let $X \sim Unif(0,1)$. Then $\mu = EX = \frac{1}{2}$.

$$\begin{aligned} Var(X) &= E[(X - \frac{1}{2})^2] \\ &= \int_0^1 (x - \frac{1}{2})^2 dx \\ &= \frac{1}{3}(x - \frac{1}{2})^3|_0^1 = \frac{1}{12}. \end{aligned}$$

Ex2: Let $X \sim N(0,1)$. Then $\mu = EX = 0$.

$$\begin{aligned} Var(X) &= E[(X-0)^2] \\ &= E[X^2] \\ &= 1. \end{aligned}$$

In the notation $N(0,1)$, 0 is the mean of $N(0,1)$ and 1 is the variance of $N(0,1)$.

A useful identity

$$Var(X) = E[X^2] - (EX)^2.$$

Explanation: Suppose X is a continuous r.v. with pdf $f(x)$.

$$\begin{aligned} E[(X-\mu)^2] &= \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx \\ &= \int_{-\infty}^{\infty} (x^2 - 2\mu x + \mu^2) f(x)dx \\ &= E[X^2] - 2\mu \int_{-\infty}^{\infty} f(x)dx \\ &\quad +\mu^2 \int_{-\infty}^{\infty} f(x)dx \\ &= E[X^2] - 2\mu \cdot \mu + \mu^2 \\ &= E[X^2] - \mu^2 \end{aligned}$$

Expectation of functions of multivariate random variables

Definition: Suppose that the joint pdf of $X_1, \ldots, X_n$ is $f(x_1, \ldots, x_n)$. Then the expectation of $Y = r(X_1, \ldots, X_n)$ can be computed by

$$E[r(X_1, \ldots, X_n)]$$
$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} r(x_1, \ldots, x_n) f(x_1, \ldots, x_n) dx_1 \cdots dx_n$$

For example, if $r(X_1, \ldots, X_n) = \frac{X_1 + \cdots + X_n}{n}$, then

$$E[\frac{X_1 + \cdots + X_n}{n}]$$
$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{x_1 + \cdots + x_n}{n} f(x_1, \ldots, x_n) dx_1 \cdots dx_n$$

Or if $r(X_1, \ldots, X_n) = X_1 \cdots X_n$, then

$$E[X_1 \cdots X_n]$$
$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1 \cdots x_n f(x_1, \ldots, x_n) dx_1 \cdots dx_n$$

Ex1

X and Y have joint pdf $f(x,y) = (x+y)I(0 < x < 1, 0 < y < 1)$. Compute $E[XY]$ and $E[X+Y]$.

$$
\begin{aligned}
E[XY] &= \int_0^1 \int_0^1 xy \cdot (x+y) dx dy \\
&= \int_0^1 dx [\int_0^1 dy xy \cdot (x+y)] \\
&= \int_0^1 dx [\frac{1}{2}x^2 + \frac{1}{3}x] \\
&= \frac{1}{3}.
\end{aligned}
$$

$$
\begin{aligned}
E[X+Y] &= \int_0^1 \int_0^1 (x+y) \cdot (x+y) dx dy \\
&= \int_0^1 dx [\int_0^1 dy (x+y)^2] \\
&= \int_0^1 dx [\frac{1}{3}((x+1)^3 - x^3)] \\
&= \frac{7}{6}.
\end{aligned}
$$

Ex2

X and Y have joint pdf $f(x,y) = \frac{1}{\pi} I(x^2 + y^2 \leq 1)$. Compute $E\sqrt{X^2 + Y^2}$.

$$E\sqrt{X^2+Y^2} \;=\; \int\int_{x^2+y^2\leq 1} \sqrt{x^2+y^2} \cdot \frac{1}{\pi} dxdy$$

Use the polar coordinates

$$x = r\cos(\theta) \quad y = r\sin(\theta)$$

$$dxdy = rdrd\theta$$

$$\begin{aligned} & \int\int_{x^2+y^2\leq 1} \sqrt{x^2+y^2} \cdot \frac{1}{\pi} dxdy \\ = \; & \int_0^1 \int_0^{2\pi} r \cdot \frac{1}{\pi} r dr d\theta \\ = \; & \int_0^1 2r^2 dr \\ = \; & \frac{2}{3}. \end{aligned}$$

so $E\sqrt{X^2+Y^2} = \frac{2}{3}$.

Properties of Expectations

Property 0. Let $X = c$ be a constant random variable, then $E(c) = c$ and $Var(c) = 0$.

Property 1. Let a, b be constants and $Y = aX + b$, then $EY = aEX + b$ and $Var(Y) = a^2 Var(X)$. Proof: Assume X has a pdf $f(x)$ then

$$\begin{aligned} EY &= E[aX + b] \\ &= \int_{-\infty}^{\infty} (ax + b) f(x) dx \\ &= a \cdot \int_{-\infty}^{\infty} x f(x) dx + b \cdot \int_{-\infty}^{\infty} f(x) dx \\ &= a \cdot EX + b \cdot 1 = aEX + b \end{aligned}$$

$$\begin{aligned} Var(Y) &= E[(Y - EY)^2] \\ &= E[(aX + b - (aEx + b))^2] \\ &= E[a^2 (X - EX)^2] \\ &= \int_{-\infty}^{\infty} a^2 (x - EX)^2 f(x) dx \\ &= a^2 \int_{-\infty}^{\infty} (x - EX)^2 f(x) dx \\ &= a^2 Var(X). \end{aligned}$$

Applications of Property 1

Ex1. Let $EX = 1$ and $Var(X) = 2$. If $Y = -2X + 3$. Find EY and $Var(Y)$.

We use property 1. $a = -2$ and $b = 3$.

$$EY = (-2)EX + 3 = -2 \cdot 1 + 3 = 1.$$

$$Var(Y) = (-2)^2 Var(X) = 4 \cdot 2 = 8.$$

Ex2. Let $X \sim N(0, 1)$. If $Y = \frac{1}{2}X + 1$. Find EY and $Var(Y)$.

We use property 1. $a = \frac{1}{2}$ and $b = 1$.

$$EY = \frac{1}{2}EX + 1 = \frac{1}{2} \cdot 0 + 1 = 1.$$

$$Var(Y) = (\frac{1}{2})^2 Var(X) = 1/4 \cdot 1 = \frac{1}{4}.$$

Property 2

$X_1, X_2, \ldots, X_n$ are n random variables and let $a_1, a_2, \ldots, a_n$ be n constants. b is another constant. Then

$$E[a_1X_1+\cdots+a_nX_n+b] = a_1EX_1+\cdots+a_nEX_n+b.$$

Some special cases:

$$E[X_1+X_2+\cdots+X_n] = EX_1+EX_2+\cdots+EX_n$$

where $a_1 = a_2 = \cdots = a_n = 1$ and $b = 0$.

$$E[\frac{X_1+X_2+\cdots+X_n}{n}] = \frac{EX_1+EX_2+\cdots+EX_n}{n}.$$

where $a_1 = a_2 = \cdots = a_n = \frac{1}{n}$ and $b = 0$.

Note that $X_1, \ldots, X_n$ can be arbitrary random variables.

Proof of property 2:

$$
\begin{aligned}
& E[a_1X_1 + \cdots + a_nX_n + b] \\
= & \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (a_1x_1 + \cdots + a_nx_n + b) f(x_1, \ldots, x_n) dx_1 \cdots dx_n \\
= & \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} a_1x_1 f(x_1, \ldots, x_n) dx_1 \cdots dx_n + \cdots \\
& + \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} a_nx_n f(x_1, \ldots, x_n) dx_1 \cdots dx_n \\
& + \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} b f(x_1, \ldots, x_n) dx_1 \cdots dx_n \\
= & a_1 \int_{-\infty}^{\infty} dx_1 x_1 [\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_n) dx_2 \cdots dx_n] + \cdots \\
& + a_n \int_{-\infty}^{\infty} dx_n x_n [\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_n) dx_1 \cdots dx_{n-1}] \\
& + b \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_n) dx_1 \cdots dx_n \\
= & a_1 \int_{-\infty}^{\infty} dx_1 x_1 [f_{X_1}(x_1)] + \cdots + a_n \int_{-\infty}^{\infty} dx_n x_n [f_{X_n}(x_n)] + b \cdot 1 \\
= & a_1 EX_1 + \cdots + a_n EX_n + b
\end{aligned}
$$

Applications of Property 2

Ex3. $EX_1 = 1, EX-2 = 2, EX_3 = -1$. Find 1. $E(\frac{X_1+X_2+X_3}{3})$ and 2. $E(-X_1 + 2X_2 - X_3 - 2)$.

$$E(\frac{X_1 + X_2 + X_3}{3}) = \frac{1}{3}EX_1 + \frac{1}{3}EX_2 + \frac{1}{3}EX_3$$
$$= \frac{1}{3}(1 + 2 - 1) = 2/3.$$

$$E(-X_1 + 2X_2 - X_3 - 2)$$
$$= (-1)EX_1 + 2EX_2 + (-1)EX_3 + (-2)$$
$$= (-1) \cdot 1 + 2 \cdot 2 + (-1) \cdot (-1) - 2 = 2.$$

Ex4. A team of 10 players entered a restaurant for lunch. They all wear the same team hat. The waiter collected their hats. When the 10 players left the waiter randomly returned them the hats. The waiter wondered how many players can actually get their own hats.

Let N be the number of players who got their own hats back. Since N is a random quantity, let's compute the average value of N, or the expectation of N.

For $1 \leq i \leq 10$, we define

$$X_i = \begin{cases} 1 & \text{if player i got his hat} \\ 0 & \text{Otherwise} \end{cases}$$

Then $N = X_1 + X_2 + \cdots + X_{10}$. By property 2 we have

$$EN = E[X_1 + X_2 + \ldots, X_{10}] = EX_1 + \cdots + EX_{10}.$$

For each EX_i, we know

$$EX_i = 1P(X_i = 1) + 0P(X_i = 0) = P(X_i = 1) = \frac{1}{10}$$

$$EN = \frac{1}{10} + \cdots + \frac{1}{10} = 1.$$

Property 3

If $X_1, X_2, \ldots, X_n$ are independent random variables, then

$$E[h_1(X_1)h_2(X_2)\cdots h_n(X_n)]$$
$$= E[h_1(X_1)]E[h_2(X_2)]\cdots E[h_n(X_n)].$$

In particular,

$$E[X_1X_2\cdots X_n] = EX_1 \cdot EX_2 \cdots EX_n.$$

Proof:

$$\begin{aligned} & E[h_1(X_1)h_2(X_2)\cdots h_n(X_n)] \\ = & \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} h_1(x_1)h_2(x_2)\cdots h_n(x_n)f(x_1,x_2,\ldots,x_n)dx_1\cdots dx_n \end{aligned}$$

By independence assumption,

$$f(x_1,x_2,\ldots,x_n) = f_{X_1}(x_1)f_{X_2}(x_2)\cdots f_{X_n}(x_n)$$

$$\begin{aligned} & E[h_1(X_1)h_2(X_2)\cdots h_n(X_n)] \\ = & \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} h_1(x_1)h_2(x_2)\cdots h_n(x_n)f_{X_1}(x_1)f_{X_2}(x_2)\cdots f_{X_n}(x_n)dx_1\cdots dx_n \\ = & (\int_{-\infty}^{\infty} h_1(x_1)f_{X_1}(x_1)dx_1)\cdots(\int_{-\infty}^{\infty} h_n(x_n)f_{X_n}(x_n)dx_n) \\ = & E[h_1(X_1)]E[h_2(X_2)]\cdots E[h_n(X_n)]. \end{aligned}$$

Applications of Property 3

Ex5. Let X_1, X_2, X_3 be independent $N(0,1)$. Find (a) $E[X_1^2 X_2^2 X_3^2]$ and (b) $E[X_1 X_2^2 X_3^3]$.

By the independence condition, we can use property 3 to get

$$\begin{aligned} & E[X_1^2 X_2^2 X_3^2] \\ = & E[X_1^2]E[X_2^2]E[X_3^2] \\ = & 1 \cdot 1 \cdot 1 = 1. \end{aligned}$$

$$\begin{aligned} & E[X_1 X_2^2 X_3^3] \\ = & E[X_1]E[X_2^2]E[X_3^3] \\ = & 0 \cdot 1 \cdot E[X_3^3] = 0. \end{aligned}$$

Ex6. If X and Y are independent, then

$$E[XY] = E[X]E[Y].$$

This is an often used equation.

Without independence the equation can be wrong. For example, let $Y = X$ and $X \sim N(0,1)$, then $E[XY] = E[X^2] = 1$ but $E[X]E[Y] = 0 \cdot 0 = 0$.

Ex7: Variance of the sum of independent random variables

If $X_1, X_2, \ldots, X_n$ are independent random variables, $a_1, a_2, \ldots, a_n$ are constants. Then

$$\begin{aligned} & Var(a_1X_1 + \cdots + a_nX_n) \\ = \; & a_1^2 Var(X_1) + \cdots + a_n^2 Var(X_n). \end{aligned}$$

Two special case:

$$\begin{aligned} & Var(X_1 + \cdots + X_n) \\ = \; & Var(X_1) + \cdots + Var(X_n), \end{aligned}$$

where $a_1 = a_2 = \cdots = a_n = 1$.

$$\begin{aligned} & Var(\frac{X_1 + \cdots + X_n}{n}) \\ = \; & \frac{Var(X_1) + \cdots + Var(X_n)}{n^2}, \end{aligned}$$

where $a_1 = a_2 = \cdots = a_n = \frac{1}{n}$.

Proof of Ex7: Variance of the sum of independent random variables

We show for any two independent RV X, Y, $Var(X+Y) = Var(X) + Var(Y)$.

$$
\begin{aligned}
& Var(X+Y) \\
= \; & E([(X+Y) - (\mu_X + \mu_Y)]^2) \\
= \; & E((X-\mu_X)^2 - 2(X-\mu_X)(Y-\mu_Y) + (Y-\mu_Y)^2) \\
= \; & E(X-\mu_X)^2 - 2E[(X-\mu_X)(Y-\mu_Y)] + E(Y-\mu_Y)^2 \\
= \; & Var(X) - 2E[(X-\mu_X)]E[(Y-\mu_Y)] + Var(Y) \\
& \text{because of independence} \\
= \; & Var(X) + Var(Y).
\end{aligned}
$$

For the general result, let $Z_1 = a_1X_1 + \cdots + a_{n-1}X_{n-1}$ and $Z_2 = a_nX_n$ then we know Z_1, Z_2 are independent and thus $Var(Z_1 + Z_2) = Var(Z_1) + Var(Z_2) = Var(a_1X_1 + \cdots + a_{n-1}X_{n-1}) + a_n^2 Var(X_n)$. Then by induction arguments we prove the general result.