Math 3501 - Probability and Statistics I

2.2 - Mathematical Expectation

Mathematical Expectation

$\underbrace{0 \cdot f(0)}_{=0}$

Definition

Let X be a random variable of the discrete type with space S and pmf $f(x)$, and let u be a real-valued function of a single real variable.

The mathematical expectation or expected value of $u(X)$, denoted $E[u(X)]$, is given by

$$\boxed{E[u(X)] = \sum_{x \in S} u(x) f(x) \,,}$$

provided the sum is absolutely convergent.

Interpretation: $E[u(X)]$ may be regarded has a weighted mean of $u(x)$, $x \in S$, where the weights are the probabilities $f(x) = P(X = x)$, $x \in S$.

$u(x) = x$

Special case: $E(X)$ is often denoted by the Greek letter μ and called the mean of X or of its distribution:

$$\mu = E(X) = \sum_{x \in S} x f(x) \,.$$

Example

Let the random variable X have the pmf

$$\underset{x}{f}(x) = \frac{1}{3}\,, \quad x \in \overbrace{\{-1, \underset{=}{0}, 1\}}^{S}\,.$$

pmf of X

Evaluate both $E[X]$ and $E[X^2]$.

$$E[X] = \sum_{x \in S} x \cdot f(x) = \sum_{x=-1}^{1} x \cdot f(x) = (-1) \cdot f(-1) + 0 \cdot f(0) + 1 \cdot f(1)$$

$$= (-1) \cdot \frac{1}{3} + \underbrace{0 \cdot \frac{1}{3}}_{0} + 1 \cdot \frac{1}{3} = 0$$

$$E[X^2] = \sum_{x \in S} x^2 \cdot f(x) = (-1)^2 \cdot f(-1) + 0^2 \cdot f(0) + 1^2 \cdot f(1)$$

$$= (-1)^2 \cdot \frac{1}{3} + 0^2 \cdot \frac{1}{3} + 1^2 \cdot \frac{1}{3} = \frac{2}{3}$$

$$E[X^3] = \sum_{x \in S} x^3 \cdot \underbrace{f(x)} = (-1)^3 \cdot \underline{f(-1)} + 0^3 \cdot \underline{f(0)} + 1^3 \cdot f(1) = 0$$

Alternative computation for $\underbrace{E[X^2]}$

Define a new r.v. $Y = X^2 \longrightarrow$ pmf of Y : $f_Y(y) = \begin{cases} \frac{1}{3} & \text{if } y=0 \\ \frac{2}{3} & \text{if } y=1 \end{cases}$

$E[X^2] = E[Y] = ?$

Since X takes values $-1, 0, 1$ then $\boxed{Y = X^2}$ takes only two values : $0, 1$

The space of Y is then $S_Y = \{0, 1\}$ and $f_Y(0) = P\{Y=0\} = P\{X^2=0\} = P\{X=0\} = \frac{1}{3}$

and $f_Y(1) = P\{Y=1\} = P\{X^2=1\} = P\{X=1 \text{ or } X=-1\}$

$= \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$

$$E[X^2] = E[Y] = \sum_{y \in S_Y} y \cdot f_Y(y) = 0 \cdot f_Y(0) + 1 \cdot f_Y(1)$$

$$= 0 \cdot \frac{1}{3} + 1 \cdot \frac{2}{3} = \frac{2}{3}$$

Properties of mathematical expectation

Theorem

When it exists, the mathematical expectation satisfies the following properties:

(a) If c is a constant, then $\boxed{E(c) = c}$

real number

$$E[c] = \sum_{x \in S} c \cdot f(x) = c \underbrace{\sum_{x \in S} f(x)}_{1} = c \cdot 1 = c$$

(b) If c is a constant and u is a function, then

$$E[c \cdot u(X)] = c \cdot E[u(X)]$$

$$E[c \cdot u(x)] = \sum_{x \in S} c \cdot u(x) \cdot f(x) = c \sum_{x \in S} u(x) \cdot f(x) = c\, E[u(x)]$$

mathematical expectation is LINEAR

(c) If c_1 and c_2 are constants and u_1 and u_2 are functions, then

$$E[c_1 u_1(X) + c_2 u_2(X)] = c_1 E[u_1(X)] + c_2 E[u_2(X)]$$

$$E[u_1(x) + u_2(x)] = \sum_{x \in S} (u_1(x) + u_2(x)) \cdot f(x) = \sum_{x \in S} (u_1(x) \cdot f(x) + u_2(x) \cdot f(x)) = \sum_{x \in S} u_1(x) \cdot f(x) + \sum_{x \in S} u_2(x) \cdot f(x) = E[u_1(x)] + E[u_2(x)]$$

Note: Property (c) can be extended to more than two terms

$$E\left[\sum_{i=1}^{k} c_i u_i(X)\right] = \sum_{i=1}^{k} c_i E[u_i(X)]$$

Example

Let the random variable X have the pmf

$$f(x) = \frac{x}{10}, \quad x \in \overbrace{\{1,2,3,4\}}^{S}. \rightarrow f(x) = \begin{cases} 1/10 & \text{if } x = 1 \\ 2/10 & \text{if } x = 2 \\ 3/10 & \text{if } x = 3 \\ 4/10 & \text{if } x = 4 \\ 0 & \text{otherwise} \end{cases}$$

Find $E[X(5-X)]$.

Two ways to evaluate:

(1) $E[X(5-X)] \underset{\text{definition}}{=} \sum_{x \in S} x(5-x) \cdot f(x) = \dots \quad \dots = 5$

(2) $E[X(5-X)] = E[5X - X^2] \underset{\text{linearity}}{=} 5\underbrace{\overbrace{E[X]}^{3}}_{?} - \underbrace{\overbrace{E[X^2]}^{10}}_{?} = 5 \cdot 3 - 10 = 5$

↑ simpler approach

$E[X] = \sum_{x \in S} x \cdot f(x) = 1 \cdot f(1) + 2 \cdot f(2) + 3 \cdot f(3) + 4 \cdot f(4) = 1 \cdot \frac{1}{10} + 2 \cdot \frac{2}{10} + 3 \cdot \frac{3}{10} + \frac{4 \cdot 4}{10} = 3$

$E[X^2] = \sum_{x \in S} x^2 \cdot f(x) = 1^2 \cdot f(1) + 2^2 \cdot f(2) + 3^2 \cdot f(3) + 4^2 \cdot f(4) = 1^2 \cdot \frac{1}{10} + 2^2 \cdot \frac{2}{10} + 3^2 \cdot \frac{3}{10} + 4^2 \cdot \frac{4}{10}$

$= 10$

Example

Let $u(x) = (x-b)^2$, where b is not a function of X, and suppose that the mathematical expectation $E\left[(X-b)^2\right]$ exists.

Find the value of b for which $E\left[(X-b)^2\right]$ is minimum.

constant

$E[(X-b)^2]$

think of this as a function of b.

$$g(b) = E[(x-b)^2] = E[x^2 - 2bX + b^2] = E[x^2] - 2bE[x] + \underbrace{E[b^2]}_{b^2}$$

linearity

$$g(b) = E[x^2] - 2b\,E[x] + b^2$$

← quadratic polynomial on b

numbers

derivative of g w.r.t b

$g'(b) = -2E[x] + 2b$ and so $g'(b) = 0$ yields $-2E[x] + 2b = 0 \Rightarrow \boxed{b = E[x]}$

$g''(b) = 2 > 0 \Rightarrow$ critical number is indeed a minimum!

CONCLUSION: $E[(X-b)^2]$ is minimum when $b = E[X]$

The minimum value is then

Variance of X $\leftarrow$ $\left\{ \begin{array}{l} E[(X - E[X])^2] \text{ also denoted as} \\ \\ E[(X-\mu)^2] \leftarrow \end{array} \right\}$

two outcomes → success → prob p
→ failure → prob $1-p$

Example

An experiment has probability of success $p \in (0,1)$ and probability of failure $q = 1 - p$. This experiment is repeated independently until the first success occurs, say this happens on trial X.

Determine:

a) • the space of X. —— $S = \{1, 2, 3, 4, \dots\} = \mathbb{N}$ ← infinite countable set

1st success observed on 1st toss

1st success on 2nd toss

1st success on 3rd toss

b) • the pmf of X. $f(x) = (1-p)^{x-1} \cdot p, \; x \in S$

c) • the mean of X. $E[X] = \frac{1}{p}$

b) Let us find $f(x) = P(X = x)$ for each $x = 1, 2, 3, 4, 5, \dots$

if $x = 1$ $\quad f(1) = P(X=1) = p$

$x = 2$ $\quad f(2) = P(X=2) = (1-p) \cdot p$

1st success on 2nd trial

$\underline{F} \quad \underline{S}$

$x = 3$ $\qquad f(3) = P(X = 3) = (1-p)^2 \, p$

1st success observed on 3rd trial $\qquad \underline{F} \ \underline{F} \ \underline{S}$

For any $x \in \mathbb{N}$

$$f(x) = P(X = x) = (1-p)^{x-1} \cdot p$$

1st success observed on trial # x

$\underline{F} \ \underline{F} \ \underline{F} \ \ldots \ \underline{F} \ \underline{S}$: $x-1$ F's, only 1 S

$x-1$ failures followed by one success

CONCLUSION: The pmf of X is $\boxed{f(x) = (1-p)^{x-1} \cdot p, \quad \underbrace{x = 1, 2, 3, 4, \ldots}_{x \in \mathbb{N}}}$

X is said to have a geometric distribution with parameter p

Let us check that $f(x)$ is actually a pmf. We have two conditions to check:

① $f(x) > 0$ for all $x \in S$ ✓

② $\sum_{x \in S} f(x) = 1$ ✓

Recall that $f(x) = (1-p)^{x-1} \cdot p \quad , \quad x = 1, 2, 3, \ldots$

Since $p \in (0,1)$, then $(1-p) \in (0,1)$ and so is $(1-p)^{x-1}$ for each $x \in \mathbb{N}$

Then $f(x) = \underbrace{(1-p)^{x-1}}_{>0} \cdot \underbrace{p}_{>0} > 0$ for all $x \in \mathbb{N}$ $\Rightarrow$ ① holds

To check condition ② we need to compute $\sum_{x \in S} f(x)$, that is,

$$\sum_{x=1}^{\infty} (1-p)^{x-1} \cdot p = p \sum_{x=1}^{\infty} (1-p)^{x-1} = ?????$$

$\boxed{\text{Review (Calculus 2):}}$

$\underbrace{\sum_{x=0}^{\infty} r^x}_{1+r+r^2+r^3+r^4+\cdots}$ is called a geometric series with ratio r.

We know that the geometric series converges if $|r|<1$ and diverges otherwise.

When $|r|<1$, the sum of the series is

$$\boxed{\sum_{x=0}^{\infty} r^x = \frac{1}{1-r}} \quad \leftarrow$$

we will use this often

What if we have a series of the form

$$\sum_{x=0}^{\infty} a r^{x}$$ converges provided $|r| < 1$

$$\underbrace{a + ar + ar^2 + ar^3 + ar^4}_{\text{geometric series with first term } \underline{\underline{a}} \text{ and ratio } r} + \cdots$$

If $|r| < 1$, the sum of the series is

$$\sum_{x=0}^{\infty} a r^{x} = a \sum_{x=0}^{\infty} r^{x} = a \cdot \frac{1}{1-r} = \frac{a}{1-r}$$

Going back to our computation

$$\sum_{x \in S} f(x) = \sum_{x=1}^{\infty} (1-p)^{x-1} \cdot p = p \underbrace{\boxed{\sum_{x=1}^{\infty} (1-p)^{x-1}}}_{1 + (1-p) + (1-p)^2 + (1-p)^3} = p \cdot \frac{1}{1-\underbrace{(1-p)}} = p \cdot \frac{1}{p} = 1$$

geometric series with ratio $1-p \in (0,1)$

ratio of the series

$\Rightarrow$ Condition (2) holds

c) The mean of X is

$$\mu = E[X] = \sum_{x \in S} x \cdot f(x) = \sum_{x=1}^{\infty} x \cdot (1-p)^{x-1} \cdot p = \sum_{x=1}^{\infty} x \cdot q^{x-1} \cdot p = ???$$

denote $q = 1-p$

Evaluation of sum:

$$\mu = \sum_{x=1}^{\infty} x \cdot q^{x-1} \cdot p = p + 2 \cdot q p + 3 \cdot q^2 \cdot p + 4 \cdot q^3 \cdot p + 5 q^4 \cdot p + \cdots$$

$$q\mu = qp + 2q^2 \cdot p + 3q^3 p + 4q^4 \cdot p + 5q^5 p$$

$$\mu - q\mu = p + qp + q^2 p + q^3 p + q^4 p + \cdots$$

$$= p \underbrace{\left(1 + q + q^2 + q^3 + q^4 + \cdots\right)}_{\frac{1}{1-q}} = p \cdot \frac{1}{1-q} = p \cdot \frac{1}{p} = 1$$

$q = 1 - p$

$\Rightarrow p = 1 - q$

Conclusion: $\mu - q\mu = 1 \Rightarrow$

$$\Rightarrow \mu(1-q) = 1 \Rightarrow \mu = \frac{1}{1-q} = \frac{1}{p}$$

$$\boxed{E[X] = \frac{1}{p}}$$

Math 3501 - Probability and Statistics I

2.3 - Special Mathematical Expectations

Mean

In what follows, suppose that X is a discrete random variable with space S and pmf $f(x)$.

Definition

The mean of the random variable X (or of its distribution) is

$$\boxed{\mu = E(X) = \sum_{x \in S} x f(x)\ .}$$

Notes:

1) the mean μ is also referred to as the *first moment of X about the origin*.

2) the first moment of X about the mean is always zero:

$$E[(X-\mu)] = E(X) - E(\mu) = \mu - \mu = 0\ .$$

linearity

Variance and standard deviation

Definition

The *variance of the random variable* X (or of its distribution), denoted $\mathrm{Var}(X)$, is the second moment of X about the mean:

$$\mathrm{Var}(X) = E[(X-\mu)^2] = \sum_{x \in S} (x-\mu)^2 f(x)$$

The positive square root of the variance is called the *standard deviation of* X and is denoted by the Greek letter σ (sigma)

$$\sigma = \sqrt{\mathrm{Var}(X)}\ .$$

Note: Variance may also be denoted by σ^2, since

$$\sigma^2 = \left(\sqrt{\mathrm{Var}(X)}\right)^2 = \mathrm{Var}(X)\ .$$

Property

Variance may be computed in another way:

$$\mathrm{Var}(X) = E\left(X^2\right) - \left(E(X)\right)^2.$$

Either use the definition to compute $\mathrm{Var}(x) = E\left[(X-\mu)^2\right]$ or

observe that:

$$\mathrm{Var}(x) = E\left[(x-\mu)^2\right] = E\left[X^2 - 2\mu X + \mu^2\right]$$

$$\overset{\text{linearity}}{=} E\left[x^2\right] - 2\mu \underbrace{E[x]}_{\mu} + \mu^2$$

$$= E\left[x^2\right] - 2\mu^2 + \mu^2 = E\left[x^2\right] - \mu^2$$

$$\boxed{\mathrm{Var}(x) = E\left[x^2\right] - \left(E[x]\right)^2}$$

← often easier than using the definition!