# Topic 8

# The Expected Value

## Definition and Properties

# Outline

# Definition

For a data set taking numerical values $x_1, x_2, \ldots, x_n$, the sample mean of a real-valued function $h$ of the data is

$$\overline{h(x)} = \sum_x h(x)p(x),$$

where $p(x)$ is the proportion of observations taking the value $x$.

For a finite sample space $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$ and a probability $P$ on $\Omega$, we can define the expectation or the expected value of a random variable $X$ by an analogous average,$g(X)$ by an analogous average,

$$EX = \sum_{j=1}^{N} X(\omega_j)P\{\omega_j\}.$$

$$Eg(X) = \sum_{j=1}^{N} g(X(\omega_j))P\{\omega_j\}.$$

# Definition

Roll one die. Then $\Omega = \{1, 2, 3, 4, 5, 6\}$. Let $X$ be the value on the die. So, $X(\omega) = \omega$. If the die is fair, then the probability model has $P\{\omega\} = 1/6$ for each outcome $\omega$. An example of an unfair dice would be the probability with $P\{1\} = P\{2\} = P\{3\} = 1/12$ and $P\{4\} = P\{5\} = P\{6\} = 1/3$.

| $\omega$ | $X(\omega)$ | $P\{\omega\}$ | $X(\omega)P\{\omega\}$ | $P\{\omega\}$ | $X(\omega)P\{\omega\}$ |
|---|---|---|---|---|---|
| 1 | 1 | 1/6 | 1/6 | 1/12 | 1/12 |
| 2 | 2 | 1/6 | 2/6 | 1/12 | 2/12 |
| 3 | 3 | 1/6 | 3/6 | 1/12 | 3/12 |
| 4 | 4 | 1/6 | 4/6 | 3/12 | 12/12 |
| 5 | 5 | 1/6 | 5/6 | 3/12 | 15/12 |
| 6 | 6 | 1/6 | 6/6 | 3/12 | 18/12 |
| sum | | 1 | $EX = 21/6 = 7/2$ | 1 | $EX = 51/12 = 17/4$ |

# Properties

Two properties of expectation are immediate from the formula

$$EX = \sum_{\omega} X(\omega)P\{\omega\}.$$

1. If $X(\omega) \geq 0$ for every outcome $\omega \in \Omega$, then every term in the sum is nonnegative and consequently their sum $EX \geq 0$.
2. Let $X_1$ and $X_2$ be two random variables and $c_1$, $c_2$ be two real numbers, then by using $g(x_1, x_2) = c_1 x_1 + c_2 x_2$ and the distributive property, we find out that

$$E[c_1 X_1 + c_2 X_2] = c_1 EX_1 + c_2 EX_2.$$

Taking these two properties, we say that the operation of taking an expectation

$$X \mapsto EX$$

is a positive linear functional.

# Discrete Random Variables

Example. Toss a coin three times and $X$ denote the number of heads.

| A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|
| $\omega$ | $X(\omega)$ | $x$ | $P\{\omega\}$ | $P\{X=x\}$ | $X(\omega)P\{\omega\}$ | $xP\{X=x\}$ |
| $HHH$ | 3 | 3 | $P\{HHH\}$ | $P\{X=3\}$ | $X(HHH)P\{HHH\}$ | $3P\{X=3\}$ |
| $HHT$ | 2 | | $P\{HHT\}$ | | $X(HHT)P\{HHT\}$ | |
| $HTH$ | 2 | 2 | $P\{HTH\}$ | $P\{X=2\}$ | $X(HTH)P\{HTH\}$ | $2P\{X=2\}$ |
| $THH$ | 2 | | $P\{THH\}$ | | $X(THH)P\{THH\}$ | |
| $HTT$ | 1 | | $P\{HTT\}$ | | $X(HHT)P\{HHT\}$ | |
| $TTH$ | 1 | 1 | $P\{THT\}$ | $P\{X=1\}$ | $X(HTH)P\{HTH\}$ | $1P\{X=1\}$ |
| $THT$ | 1 | | $P\{TTH\}$ | | $X(THH)P\{THH\}$ | |
| $TTT$ | 0 | 0 | $P\{TTT\}$ | $P\{X=0\}$ | $X(TTT)P\{TTT\}$ | $0P\{X=0\}$ |

$$EX = 0 \cdot P\{X=0\} + 1 \cdot P\{X=1\} + 2 \cdot P\{X=2\} + 3 \cdot P\{X=3\}.$$

# Discrete Random Variables

To find a similar formula for $Eg(X)$ for a discrete random variable $X$ ,

| A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|
| $\omega$ | $X(\omega)$ | $x$ | $P\{\omega\}$ | $P\{X = x\}$ | $g(X(\omega))P\{\omega\}$ | $g(x)P\{X = x\}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $\omega_k$<br>$\omega_{k+1}$<br>$\vdots$ | $X(\omega_k)$<br>$X(\omega_{k+1})$<br>$\vdots$ | $x_i$ | $P\{\omega_k\}$<br>$P\{\omega_{k+1}\}$<br>$\vdots$ | $P\{X = x_i\}$ | $g(X(\omega_k))P\{\omega_k\}$<br>$g(X(\omega_{k+1}))P\{\omega_{k+1}\}$<br>$\vdots$ | $g(x_i)P\{X = x_i\}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |

$$\begin{aligned} Eg(X) &= g(x_1)P\{X = x_1\} + \cdots + g(x_n)P\{X = x_n\} \\ &= g(x_1)f_X(x_1) + \cdots + g(x_n)f_X(x_n) \\ Eg(X) &= \sum_x g(x)f_X(x) \end{aligned}$$

# Discrete Random Variables

A similar formula holds if we have a vector of random variables $X = (X_1, X_2, \ldots, X_n)$,

$$f_X(x_1, x_2, \ldots, x_n) = P\{X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n\},$$

the joint probability mass function and $g$ a real-valued function of $x = (x_1, x_2, \ldots, x_n)$.

In the two dimensional case, this takes the form

$$Eg(X_1, X_2) = \sum_{x_1} \sum_{x_2} g(x_1, x_2) f_{X_1,X_2}(x_1, x_2).$$

We will return to this identity in computing the covariance of two random variables.

Exercise. Find $EX^2$ for both the fair dice and the unfair dice.

# Bernoulli Trials

Bernoulli trials are the simplest and among the most common models for an experimental procedure. Each trial has two possible outcomes, variously called,

heads-tails, yes-no, up-down, left-right, win-lose, female-male, green-blue, dominant-recessive, **success-failure**.

Random variables $X_1, X_2, \ldots, X_n$ are called a sequence of Bernoulli trials provided that:

- Each $X_i$ takes on two values, namely, 0 and 1. We call the value 1 a success and the value 0 a failure.
- Each trial has the same probability for success, i.e., $P\{X_i = 1\} = p$ for each $i$.
- The outcomes on each of the trials is independent.

# Bernoulli Trials

For each trial $i$, the expected value

$$EX_i = 0 \cdot P\{X_i = 0\} + 1 \cdot P\{X_i = 1\} = 0 \cdot (1-p) + 1 \cdot p = p$$

is the same as the success probability.

Let $S_n = X_1 + X_2 + \cdots + X_n$ be the total number of successes in $n$ Bernoulli trials.

Using the linearity of expectation, we see that

$$ES_n = E[X_1 + X_2 \cdots + X_n] = p + p + \cdots + p = np,$$

the expected number of successes in $n$ Bernoulli trials is $np$.

# Bernoulli Trials

To determine the probability mass function for $S_n$. Beginning with a concrete example, let $n = 8$, and the outcome

failure, success, failure, failure, success, failure, failure, success

Using the independence of the trials, we can compute the probability of this outcome:

$$(1-p) \times p \times (1-p) \times (1-p) \times p \times (1-p) \times (1-p) \times p = p^3(1-p)^5.$$

- Any of the $\binom{8}{3}$ sequences of 8 Bernoulli trials having 3 successes also has probability $p^3(1-p)^5$.
- Each of the outcomes are mutually exclusive.
- Their union is the event $\{S_8 = 3\}$.
- Consequently, by the axioms of probability,

$$P\{S_8 = 3\} = \binom{8}{3} p^3(1-p)^5.$$

# Bernoulli Trials

Replace 8 by $n$ and 3 by $x$ to see that any sequence of $n$ Bernoulli trials having $x$ successes has probability

$$p^x(1-p)^{n-x}.$$

In addition, we know that we have

$$\binom{n}{x}$$

mutually exclusive sequences of $n$ Bernoulli trials that have $x$ successes. Thus, we have the mass function

$$f_{S_n}(x) = P\{S_n = x\} = \binom{n}{x} p^x(1-p)^{n-x}, \quad x = 0, 1, \ldots, n.$$

The fact that the sum

$$\sum_{x=0}^{n} f_{S_n}(x) = \sum_{x=0}^{n} \binom{n}{x} p^x(1-p)^{n-x} = (p + (1-p))^n = 1^n = 1$$

follows from the binomial theorem. Thus, $S_n$ is called a binomial random variable

# Bernoulli Trials

Exercise. Let $S_3$ be the number of successes in 3 Bernoulli trials with success parameter $p = 2/3$.

1. Find the mass function $f_{S_3}(x)$, x=0,1,2,3.
2. Use $f_{S_3}$ to compute $ES_3$.
3. Sketch the cumulative distribution function $F_{S_3}$ for $S_3$.