Expected Value

Reading

- Malik Magdon-Ismail. Discrete Mathematics and Computing.
  - Chapter 19

Overview

- Expected value approximates the sample average.
- Mathematical Expectation
- Examples
  - Sum of dice.
  - Bernoulli.
  - Uniform.
  - Binomial.
  - Waiting time.
- Conditional Expectation
- Law of Total Expectation

Sample Average: Toss Two Coins Many Times

Rensselaer

- Suppose I have 2 coins. Suppose I toss them 1 time each
  - What is the probability that I get 1 H?
  - Let X denote the number of heads

ω	HH	HT	TH	TT
$\mathbb{P}[\omega]$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
$X(\omega)$	2	1	1	0

x	0	1	2
$P_X[x]$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

- Repeat the experiment $n = 24$ times and suppose I get the following outcomes:

HH	TH	HT	HH	HH	TH	TT	TT	HH	TT	HT	HT	HH	HT	TT	HT	TT	HT	HT	TH	HH	TH	TT	TH
2	1	1	2	2	1	0	0	2	0	1	1	2	1	0	1	0	1	1	1	2	1	0	1

  - What is the average value of X?

$$\frac{2+1+1+2+2+1+0+0+2+0+1+1+2+1+0+1+0+1+1+1+2+1+0+1}{24} = \frac{24}{24} = 1$$

Sample Average: Toss Two Coins Many Times

Rensselaer

- Suppose I have 2 coins. Suppose I toss them 1 time each
  - What is the probability that I get 1 H?
  - Let X denote the number of heads

ω	HH	HT	TH	TT
$\mathbb{P}[\omega]$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
$X(\omega)$	2	1	1	0

x	0	1	2
$P_X[x]$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

- Repeat the experiment $n = 24$ times and suppose I get the following outcomes:

HH	TH	HT	HH	HH	TH	TT	TT	HH	TT	HT	HT	HH	HT	TT	HT	TT	HT	HT	TH	HH	TH	TT	TH
2	1	1	2	2	1	0	0	2	0	1	1	2	1	0	1	0	1	1	1	2	1	0	1

- To make it easier to count, let's reorder outcomes

TT	TT	TT	TT	TT	TT	HT	HT	HT	HT	HT	HT	HT	TH	TH	TH	TH	TH	HH	HH	HH	HH	HH	HH
		$n_0 = 6$									$n_1 = 12$									$n_2 = 6$			
0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2

  - Can you rewrite the average by grouping similar terms?

$$\frac{6 \times 0 + 12 \times 1 + 6 \times 2}{24} = \frac{24}{24}$$

Mathematical Expectation of a Random Variable X

- Let's go back to our 2 coins

TT	TT	TT	TT	TT	TT	HT	HT	HT	HT	HT	HT	HT	TH	TH	TH	TH	TH	HH	HH	HH	HH	HH	HH
		$n_0 = 6$									$n_1 = 12$									$n_2 = 6$			
0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2

- The general formula for the average value of X is:

$$\frac{n_0 \times 0 + n_1 \times 1 + n_2 \times 2}{n} = \frac{n_0}{n} \times 0 + \frac{n_1}{n} \times 1 + \frac{n_2}{n} \times 2$$

  – What is $\frac{n_0}{n}$ roughly?
    - Approximately $P_X(0)$
    - Law of Large Numbers: as n gets large, $\frac{n_i}{n}$ will converge to $P_X(i)$

- So,

$$\frac{n_0 \times 0 + n_1 \times 1 + n_2 \times 2}{n} \approx P_X(0) \times 0 + P_X(1) \times 1 + P_X(2) \times 2$$

$$= \sum_{x \in X(\Omega)} x P_X(x)$$

- For two coins, the expected value is $0 \times \frac{1}{4} + 1 \times \frac{1}{2} + 2 \times \frac{1}{4} = 1$

Mathematical Expectation of a Random Variable X, cont'd

- *Definition [Expected Value]*. The expected value of a random variable X is defined as the sum of all values x that X can take, weighted by their probabilities, $P_X(x)$:

$$\mathbb{E}[X] = \sum_{x \in X(\Omega)} x \cdot P_X(x)$$

- **Synonyms:** Expectation; Expected Value; Mean; Average.
- **Important Exercise.** Show that $\mathbb{E}[X] = \sum_{\omega \in \Omega} X(\omega) \cdot \mathbb{P}[\omega]$

Expected Number of Heads from 3 Coin Tosses is 1.5!

$$\mathbb{E}[X] = \sum_{x \in X(\Omega)} x \cdot P_X(x)$$

- Recall the outcome space:

ω	HHH	HHT	HTH	HTT	THH	THT	TTH	TTT
$\mathbb{P}[\omega]$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
$X(\omega)$	3	2	2	1	2	1	1	0

x	0	1	2	3
$P_X(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

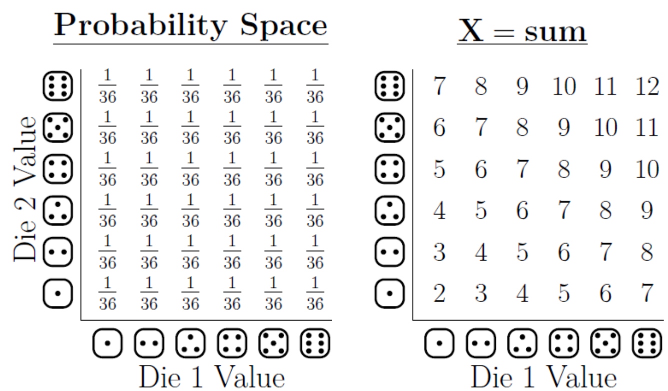
- So, what is $\mathbb{E}[X]$?

$$\mathbb{E}[X] = 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8}$$
$$= \frac{12}{8} = 1.5$$

- What does this mean?!? CAN'T GET HALF A HEAD!
  - It's just an average. The average may not correspond to a possible outcome
- **Exercise.** Let X be the value of a fair die roll. Show that $\mathbb{E}[X] = 3.5$

Expected Sum of Two Dice

- Consider two fair dice, X_1 and X_2. Let $X = X_1 + X_2$



- What is the probability of each outcome for X?

x	2	3	4	5	6	7	8	9	10	11	12
$P_X(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

- What is $\mathbb{E}[X]$?

$$\mathbb{E}[X] = \sum_x x \cdot P_X(x) =$$

$$= \frac{1}{36}(2 \times 1 + 3 \times 2 + 4 \times 3 + 5 \times 4 + 6 \times 5 + 7 \times 6 +$$
$$+8 \times 5 + 9 \times 4 + 10 \times 3 + 11 \times 2 + 12 \times 1) = 7$$

Expected Sum of Two Dice, cont'd

- Hm,

$$\mathbb{E}[X] = 7$$

- But $\mathbb{E}[X_1] = \mathbb{E}[X_2] = 3.5$
- So $\mathbb{E}[X] = \mathbb{E}[X_1] + \mathbb{E}[X_2]$
- Interesting…

Expected Value of a Bernoulli Random Variable

- A Bernoulli random variable X takes a value in $\{0,1\}$

x	0	1
$P_X(x)$	$1-p$	p

- The expected value is

$$\mathbb{E}[X] = 0 \times (1-p) + 1 \times p = p$$

- A Bernoulli random variable with success probability p has expected value p
  - Duh…
  - Again, the mean does not correspond to an actual outcome

Expected Value of a Uniform Random Variable

Rensselaer

- A uniform random variable X takes a value in $\{1, \ldots, n\}$

x	1	2	…	n
$P_X(x)$	$\frac{1}{n}$	$\frac{1}{n}$	$\frac{1}{n}$	$\frac{1}{n}$

- The expected value is

$$\begin{aligned}\mathbb{E}[X] &= 1 \times \frac{1}{n} + 2 \times \frac{1}{n} + \cdots + n \times \frac{1}{n} \\ &= \frac{1}{n}(1 + 2 + \cdots + n) \\ &= \frac{1}{n} \times \frac{1}{2} n(n+1) = \frac{n+1}{2}\end{aligned}$$

- A uniform random variable on $[1, n]$ has expected value $= \frac{1}{2}(n+1)$
  - Notation: $[1, n]$ means $\{1,2, \ldots, n\}$

What is the Expected Number of Heads in n Coin Tosses?

Rensselaer

- What is the distribution of the number of heads in n coin tosses?
  - A binomial distribution

$$P_X(k) = B(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}$$

x	0	1	…	k	…	n
$P_X(x)$	$\binom{n}{0} p^0 (1-p)^n$	$\binom{n}{1} p^1 (1-p)^{n-1}$	…	$\binom{n}{k} p^k (1-p)^{n-k}$	…	$\binom{n}{n} p^n (1-p)^0$

- What is the expected number of heads?

$$\mathbb{E}[X] = 0 \times \binom{n}{0} p^0 q^n + 1 \times \binom{n}{1} p^1 q^{n-1} + \cdots + n \times \binom{n}{n} p^n q^0$$

  - where $q = 1 - p$

- Hm, this looks very similar to the standard binomial theorem!

$$(p+q)^n = \binom{n}{0} p^0 q^n + \binom{n}{1} p^1 q^{n-1} + \cdots + \binom{n}{n} p^n q^0$$

What is the Expected Number of Heads in n Coin Tosses?, cont'd

Rensselaer

- What is the distribution of the number of heads in n coin tosses?
  - A binomial distribution

$$P_X(k) = B(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}$$

- What is the expected number of heads?

$$\mathbb{E}[X] = 0 \times \binom{n}{0} p^0 q^n + 1 \times \binom{n}{1} p^1 q^{n-1} + \cdots + n \times \binom{n}{n} p^n q^0$$

  - where $q = 1 - p$

- Hm, this looks very similar to the standard binomial theorem!

$$(p+q)^n = \binom{n}{0} p^0 q^n + \binom{n}{1} p^1 q^{n-1} + \cdots + \binom{n}{n} p^n q^0$$

- Taking the derivative of both sides w.r.t. p

$$n(p+q)^{n-1} = 1 \times \binom{n}{1} p^0 q^{n-1} + 2 \times \binom{n}{2} p^1 q^{n-2} + \cdots + n \times \binom{n}{n} p^{n-1} q^0$$

- Finally, multiplying both sides by p

$$pn(p+q)^{n-1} = 1 \times \binom{n}{1} p^1 q^{n-1} + 2 \times \binom{n}{2} p^2 q^{n-2} + \cdots + n \times \binom{n}{n} p^n q^0$$

- RHS is = $\mathbb{E}[X]$ (why?)

$$\mathbb{E}[X] = 0 \times \binom{n}{0} p^0 q^n + RHS$$

What is the Expected Number of Heads in n Coin Tosses?, cont'd

- What is the distribution of the number of heads in n coin tosses?
  - A binomial distribution

$$P_X(k) = B(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}$$

- What is the expected number of heads?

$$\mathbb{E}[X] = 0 \times \binom{n}{0} p^0 q^n + 1 \times \binom{n}{1} p^1 q^{n-1} + \cdots + n \times \binom{n}{n} p^n q^0$$

  - where $q = 1 - p$
- Hm, this looks very similar to the standard binomial theorem!

$$(p+q)^n = \binom{n}{0} p^0 q^n + \binom{n}{1} p^1 q^{n-1} + \cdots + \binom{n}{n} p^n q^0$$

- Taking the derivative of both sides w.r.t. p

$$n(p+q)^{n-1} = 1 \times \binom{n}{1} p^0 q^{n-1} + 2 \times \binom{n}{2} p^1 q^{n-2} + \cdots + n \times \binom{n}{n} p^{n-1} q^0$$

- Finally, multiplying both sides by p

$$pn(p+q)^{n-1} = 1 \times \binom{n}{1} p^1 q^{n-1} + 2 \times \binom{n}{2} p^2 q^{n-2} + \cdots + n \times \binom{n}{n} p^n q^0$$

- But $p + q = 1$ by definition; RHS is = $\mathbb{E}[X]$

$$\mathbb{E}[X] = np$$

Binomial Distribution Example

- **Example.** Answer randomly 15 multiple choice questions with 5 choices ($p = \frac{1}{5}$)
  - What do you expect to get?

$$15 \times \frac{1}{5} = 3 \text{ correct answers}$$

Expected Waiting Time to Success

Rensselaer

- Recall our success-waiting problem
  - We derived the **Exponential Waiting Time Distribution** as: $P_X(t) = \beta(1-p)^t$
  - where $\beta = \frac{p}{1-p}$

x	1	2	3	…	k
$P_X(x)$	$\beta(1-p)$	$\beta(1-p)^2$	$\beta(1-p)^3$	…	$\beta(1-p)^k$

- Note that $\mathbb{E}[X] = \beta\big(1 \times (1-p) + 2 \times (1-p)^2 + \cdots + k \times (1-p)^k + \cdots\big)$
- We'll use a similar derivation to the binomial distribution
- Recall the geometric series:

$$\frac{1}{1-a} = 1 + a + a^2 + a^3 + \cdots$$

- Taking the derivative w.r.t. a:

$$\frac{1}{(1-a)^2} = 1 + 2a + 3a^2 + 4a^3 + \cdots$$

- Multiplying by a:

$$\frac{a}{(1-a)^2} = 1a + 2a^2 + 3a^3 + 4a^4 + \cdots$$

Expected Waiting Time to Success

- Recall our success-waiting problem
  - We derived the **Exponential Waiting Time Distribution** as: $P_X(t) = \beta(1-p)^t$
  - where $\beta = \frac{p}{1-p}$

x	1	2	3	…	k
$P_X(x)$	$\beta(1-p)$	$\beta(1-p)^2$	$\beta(1-p)^3$	…	$\beta(1-p)^k$

- Note that $\mathbb{E}[X] = \beta\big(1 \times (1-p) + 2 \times (1-p)^2 + \cdots + k \times (1-p)^k + \cdots\big)$
- We'll use a similar derivation to the binomial distribution

$$\frac{a}{(1-a)^2} = 1a + 2a^2 + 3a^3 + 4a^4 + \cdots$$

- So finally,

$$\mathbb{E}[X] = \beta \times \frac{1-p}{p^2} = \frac{1}{p}$$

- Expected waiting time is $\frac{1}{p}$

Conditional Expectation

- New information changes a probability. Hence, the expected value also changes.
- Average annual precipitation in the US is 30.28 inches
  - Source: https://www.ncei.noaa.gov/access/monitoring/monthly-report/national/202013

$$\mathbb{E}[Precipitation] = 30.28$$

- Average annual precipitation in the northeast is 43.61 inches

$$\mathbb{E}[Precipitation|Northeast] = 43.61$$

Source: https://gisgeography.com/us-precipitation-map/

Conditional Expectation Definition

- **Conditional Expected Value $\mathbb{E}[X|A]$:**

$$\mathbb{E}[X|A] = \sum_{x \in X(\Omega)} x \cdot \mathbb{P}[X = x|A] = \sum_{x \in X(\Omega)} x \cdot P_X(x|A)$$

Law of Total Expectation

Rensselaer

- Case by case analysis for expectation (similar to the Law of Total Probability).

- **Law of Total Expectation**:

$$\mathbb{E}[X] = \mathbb{E}[X|A] \times \mathbb{P}[A] + \mathbb{E}[X|\bar{A}] \times \mathbb{P}[\bar{A}]$$

- For example,

$$\begin{aligned}\mathbb{E}[Precipitation] &= \mathbb{E}[Precipitation|Northeast] \times \mathbb{P}[Northeast] + \\ &\quad + \mathbb{E}[Precipitation|\overline{Northeast}] \times \mathbb{P}[\overline{Northeast}] \\ &\approx 43.61 \times 0.05 + 29.58 \times 0.95\end{aligned}$$

- Of course, can sub-divide $\bar{A}$ (i.e., $\overline{N\text{ortheast}}$) into more regions for more refined information

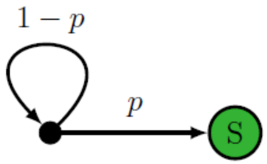
Example

- Suppose a jar has 9 fair coins and 1 two-headed coin.
- Choose a random coin and flip it 10 times.
- Let X be the number of heads you see. What is $\mathbb{E}[X]$?
- Use the law of total expectation
  - What are the events?
  - Coin is fair vs. coin is 2-headed

$$\mathbb{E}[X] = \mathbb{E}[X|fair] \times \mathbb{P}[fair] + \mathbb{E}[X|2 - headed] \times \mathbb{P}[2 - headed]$$
$$= 5 \times \frac{9}{10} + 10 \times \frac{1}{10} = 5.5$$

Expected Waiting Time from Law of Total Expectation

- Go back to the expected waiting time example, where success comes w.p. p
- Let X be the waiting time. There are two cases:
  - First trial is a succeeds (S) with probability p, i.e., $X = 1$
  - First trial is a fails (F) with probability $1 - p$, i.e., "restart"



- Using the law of total expectation:

$$\begin{aligned}\mathbb{E}[X] &= \mathbb{E}[X|S] \times \mathbb{P}[S] + \mathbb{E}[X|F] \times \mathbb{P}[F] \\ &= 1 \times p + (1 + \mathbb{E}[X]) \times (1 - p)\end{aligned}$$

- Solving for $\mathbb{E}[X]$:

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

  - which we already knew
  - the law of total expectation gives a neat way to calculate this expectation!
- **Practice.** Exercise 6.5 and 6.8.