Random Variables

$X : \Omega \longrightarrow \mathbb{R}$

- function from sample space to $\mathbb{R}$

$P(X=k) = P(\{\omega \in \Omega : X(\omega) = k\})$

ex. flip a fair coin

$\Omega = \{H, T\}$

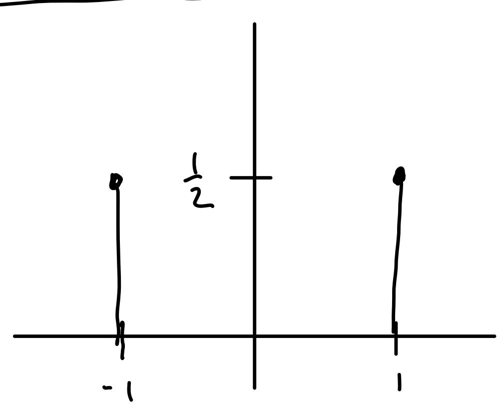
let X be the random variable s.t.

$X(H) = 1, \quad X(T) = -1$

$P(X=1) = P(\{\omega \in \Omega : X(\omega) = 1\}) = P(H) = \frac{1}{2}$

$P(X=-1) = P(\{\omega \in \Omega : X(\omega) = -1\}) = P(T) = \frac{1}{2}$

Distribution of X



Binomial Distribution

- counts the number of successes in a sequence of n independent trials where each trial has success probability p.

$X \sim Bin(n, p)$

$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$

1 Head Count

Consider a coin with $\mathbb{P}(\text{Heads}) = 2/5$. Suppose you flip the coin 20 times, and define X to be the number of heads.

(a) Name the distribution of X and what its parameters are.

20 independent trials, each have $2/5$ probability of success

$X \sim \text{Bin}(20, 2/5)$

(b) What is $\mathbb{P}(X = 7)$?

$P(X=7) = \binom{20}{7}\left(\frac{2}{5}\right)^7\left(\frac{3}{5}\right)^{13}$

(c) What is $\mathbb{P}(X \geq 1)$? Hint: You should be able to do this without a summation.

$P(X \geq 1) = 1 - P(X<1) = 1 - P(X=0) = 1 - \left(\frac{3}{5}\right)^{20}$

(d) What is $\mathbb{P}(12 \leq X \leq 14)$?

$P(12 \leq X \leq 14) = P(X=12) + P(X=13) + P(X=14)$

$= \binom{20}{12}\left(\frac{2}{5}\right)^{12}\left(\frac{3}{5}\right)^8 + \binom{20}{13}\left(\frac{2}{5}\right)^{13}\left(\frac{3}{5}\right)^7$

$+ \binom{20}{14}\left(\frac{2}{5}\right)^{14}\left(\frac{3}{5}\right)^6$

Expectation: "weighted average of values r.v. can take on"

$$E[X] = \sum_{k \in range(X)} k P(X=k)$$

Indicator: indicates success or failure of each trial in a sequence of trials

$$X_i = \begin{cases} 1 & w.p.\ p \\ 0 & w.p.\ 1-p \end{cases}$$

ex. 3 coin flips: X_1, X_2, X_3
indicate if ith flip heads

* indicators are Bernoulli r.v.s

2 How Many Queens?

You shuffle a standard 52-card deck, before drawing the first three cards from the top of the pile. Let X denote the number of queens you draw.

(a) What is $\mathbb{P}(X=0)$, $\mathbb{P}(X=1)$, $\mathbb{P}(X=2)$ and $\mathbb{P}(X=3)$?

$$P(X=0) = \frac{\binom{48}{3}}{\binom{52}{3}} = \frac{4324}{5525}$$

$$P(X=1) = \frac{\binom{4}{1}\binom{48}{2}}{\binom{52}{3}} = \frac{1128}{5525}$$

$$P(X=2) = \frac{\binom{4}{2}\binom{48}{1}}{\binom{52}{3}} = \frac{72}{5525}$$

$$P(X=3) = \frac{\binom{4}{3}}{\binom{52}{3}} = \frac{1}{5525}$$

(b) What do your answers you computed in part a add up to?

$$P(X=0)+P(X=1)+P(X=2)+P(X=3) = \frac{4324+1128+72+1}{5525} = 1$$

(c) Compute $\mathbb{E}(X)$ from the definition of expectation.

$$E[X] = \sum_{k=0}^{3} k\,P(X=k) = 0\cdot\frac{4324}{5525} + 1\cdot\frac{1128}{5525} + 2\cdot\frac{72}{5525} + 3\cdot\frac{1}{5525} = \frac{3}{13}$$

(d) Let X_i be an indicator random variable that equals 1 if the ith card a is queen and 0 otherwise. Are the X_i indicators independent?

No, $P(X_2=1 \mid X_1=1) = \frac{3}{51} \neq P(X_2=1) = \frac{1}{13}$

Linearity of Expectation

For any n random variables on the same probability space:

$E[c_1X_1 + \dots + c_nX_n] = c_1E[X_1] + \dots + c_nE[X_n]$

$* X_1, \dots, X_n$ don't need to be independent!

3 Linearity

Solve each of the following problems using linearity of expectation. Explain your methods clearly.

(a) In an arcade, you play game A 10 times and game B 20 times. Each time you play game A, you win with probability $1/3$ (independently of the other times), and if you win you get 3 tickets (redeemable for prizes), and if you lose you get 0 tickets. Game B is similar, but you win with probability $1/5$, and if you win you get 4 tickets. What is the expected total number of tickets you receive?

expected # of tickets we receive $= 3 \cdot$ expected # of games won playing A $+ 4 \cdot$ expected # of games won playing B

let $A_i = \begin{cases} 1 & \text{if you win ith time playing A} \\ 0 & \text{o.w.} \end{cases}$

$B_i = \begin{cases} 1 & \text{if you win ith time playing B} \\ 0 & \text{o.w.} \end{cases}$

$E[A_i] = \cancel{0 \cdot P(A_i = 0)} + 1 \cdot P(A_i = 1)$ $\quad E[B_i] = \frac{1}{5}$

$= \frac{1}{3}$

let T_A = # of tickets won from playing A

T_B = " B

Want $E[T_A + T_B] = E[T_A] + E[T_B]$

$$E[T_A] = 3\underbrace{E\left[\sum_{i=1}^{10} A_i\right]}_{\text{expected \# of wins playing A}} = 3\sum_{i=1}^{10} E[A_i] = 3\sum_{i=1}^{10}\frac{1}{3} = 10$$

$$E[T_B] = 4E\left[\sum_{i=1}^{20} B_i\right] = 4\sum_{i=1}^{20} E[B_i] = 4\sum_{i=1}^{20}\frac{1}{5} = 16$$

$\Rightarrow E[T_A + T_B] = 10 + 16 = 26$

let $X_A \sim Bin(10, 1/3)$

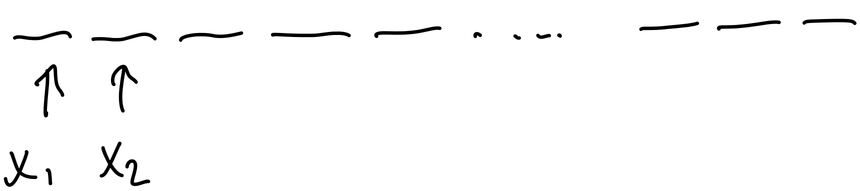
$X_B \sim Bin(20, 1/5)$

find $3E[X_A] + 4E[X_B]$

(b) A monkey types at a 26-letter keyboard with one key corresponding to each of the lower-case English letters. Each keystroke is chosen independently and uniformly at random from the 26 possibilities. If the monkey types 1 million letters, what is the expected number of times the sequence "book" appears?

let X = # of times "book" appears

$$X_i = \begin{cases} 1 & \text{if "book" appears starting at letter } i \\ 0 & \text{o.w.} \end{cases}$$



$$E[X_i] = P(X_i = 1) = \frac{1}{26^4}$$

There are $1{,}000{,}000 - 4 + 1 = 999{,}997$ places where book can start at that position

$$E[X] = E[X_1 + \dots + X_{999,997}]$$

$$= E[X_1] + \dots + E[X_{999,997}]$$

$$= \frac{999{,}997}{26^4}$$

$$\approx 2.19$$