Chapter 2: Random Variables

Example: You toss a coin 3 times

- $\Omega = \{\text{HHH}, \ldots, \text{TTT}\}$
- $|\Omega| = 8$
- A random variable is formally a function $X : \Omega \to \mathbb{R}$
- Say X is number of heads in the experiment
  - $X(HTH) = 2$

A pdf summarizes all statistical info there is to about a random variable

- $p_X(x) = P(X = x)$ where X is random and x is fixed

Another function that summarizes everything about a random variable X is the cdf

- $F_X(x) = P(X \leq x)$ where X is random and x is fixed
- Note: all cdfs are non-decreasing
- $lim_{x \to -\infty} F_X(x) = 0$
- $lim_{x \to \infty} F_X(x) = 1$

Independent Random Variables

- Let X and Y be random variables taking possible values $x_1, x_2, ...$ and $y_1, y_2, ...$
- We say X, Y are independent if $P(\{X = x_i\} \cap \{X = x_j\}) = P(\{X = x_i\})P(\{X = x_j\})$ for all x_i, y_j
- If X, Y, Z are random variables, we say they are independent if $P(\{X = x_i\} \cap \{X = x_j\} \cap \{Z = z_k\}) = P(\{X = x_i\})P(\{X = x_j\})P(\{Z = z_k\})$ for all x_i, y_j, z_k
- This applies for any number of random variables

Zoo of random variables

$X \sim Bernoulli(p)$

- $p \in [0, 1]$ fixed parameter
- X represents the outcome of a coin toss
  - $X = 1$ if toss is H
  - $X = 0$ if toss is T
- $P(X = 1) = p, p_X(1) = p$
- $P(X = 0) = 1 - p, p_X(0) = 1 - p$
- $p_X(x) = p^x(1 - p)^{1-x}, x \in [0, 1]$

$Y \sim Binomial(n, p)$

- two parameters $n \geq 1, n \in \mathbb{Z}, p \in [0, 1]$
- Y represents the story number of heads among n independent tosses of a coin with head probability p
- Y takes values in $\{0, 1, \ldots, \text{n}\}$
- Derive pmf $p_Y(k) = P(Y = k)$
- Special case: $n = 3, k = 2$
  - $\{Y = 2\} = A_1 \cup A_2 \cup A_3$
  - $A_1 = \{\text{hht}\}$
  - $A_2 = \{\text{hth}\}$
  - $A_3 = \{\text{thh}\}$
- By 3rd axiom: $P(Y = 2) = P(A_1) + P(A_2) + P(A_3)$
  - $A_{11} = \{\text{h on first toss}\}$
  - $A_{12} = \{\text{h on second toss}\}$

- $A_{13} = \{$t on third toss$\}$
  - $A_1 = A_{11} \cap A_{12} \cap A_{13}$
- Independent: $P(A_1) = P(A_{11})P(A_{12})P(A_{13})$
- Likewise: $P(A_2) = P(A_3) = p^2(1-p)$
- So: $P(Y = 2) = p_Y(2) = 3p^2(1-p)$
- General Case: $p_Y(k) = \binom{n}{k}p^k(1-p)^{n-k}$
  - Note: $\binom{n}{k}$ is number of ways getting k heads among n tosses

$Z \sim Geometric(g)$

- Story: Number of attempts of tossing a coin with head probability p needed to observe the first head
  - All tosses independent, $p \in [0,1]$
- Z is number of attempts until success
- The possible values of Z are 1, 2, $\ldots$, ∞
- $p_Z(k) = P(Z = k)$ when k has heads and tails for 1 to $k-1$

$W \sim NegativeBinomial(r, p)$

- $NegativeBinomial(1, p) = Geometric(p)$
- $r \geq 1, r \in \mathbb{Z}, p \in [0,1]$
- Number of attempts of tossing a coin with head p needed to observe r heads (all tosses independent)
- Possible values of W are $r, r+1, \ldots$
- $p_W(k) = P(W = k)$ where the last toss is a head and the tosses from 1 to $k-1$ must have $r-1$ heads
- $\{W = k\} = \{r-1$ heads among first $k-1$ tosses$\} \cap \{$head on k-th attempt$\}$
- $P(W = k) = \mathrm{P}(\{r-1$ heads among first $k-1$ tosses$\})\mathrm{P}(\{$head on k-th attempt$\})$
  - $P(W = k) = \binom{k-1}{r-1}p^{r-1}(1-p)^{(k-1)-(r-1)}p$
  - $p_W(k) = \binom{k-1}{r-1}p^r(1-p)^{k-r}$

$X \sim Hypergeometric(n, m, r)$

- Jar with n marbles, r of them are red, $n-r$ are blue, we draw m marbles from jar without replacement
  - All draws are equally likely
- X represents number of red marbles among the drawn ones
- $p_X(k) = P(X = k)$
  - Think about a tray, k red ones and $m-k$ blue ones, m spaces
  - $\binom{r}{k}$ ways of filling tray of red marbles
  - $\binom{n-r}{m-k}$ ways of filling tray of blue marbles
- $p_X(k) = \frac{\binom{r}{k}\binom{n-r}{m-k}}{\binom{n}{m}}$
- X ranges from 0, 1, $min(r, m)$