

Basics on Probability

Coin Flips

- You flip a coin
 - Head with probability 0.5
- You flip 100 coins
 - How many heads would you expect

Coin Flips cont.

- You flip a coin
 - Head with probability p
 - Binary random variable
 - Bernoulli trial with success probability p
- You flip k coins
 - How many heads would you expect
 - Number of heads X : discrete random variable
 - Binomial distribution with parameters k and p

Discrete Random Variables

- Random variables (RVs) which may take on only a **countable** number of **distinct** values
 - E.g., the total number of heads X you get if you flip 100 coins
- X is a RV with arity k if it can take on exactly one value out of $\{x_1, \dots, x_k\}$
 - E.g., the possible values that X can take are 0, 1, 2, ..., 100

Probability of Discrete RV

- Probability mass function (pmf): $P(X = x_i)$
- Easy facts about pmf
 - $\sum_i P(X = x_i) = 1$
 - $P(X = x_i \cap X = x_j) = 0$ if $i \neq j$
 - $P(X = x_i \cup X = x_j) = P(X = x_i) + P(X = x_j)$ if $i \neq j$
 - $P(X = x_1 \cup X = x_2 \cup \dots \cup X = x_k) = 1$

Common Distributions

- Uniform $X \sim U[1, \dots, N]$
 - X takes values 1, 2, ..., N
 - $P(X = i) = 1/N$
 - E.g., picking balls of different colors from a box
- Binomial $X \sim \text{Bin}(n, p)$
 - X takes values 0, 1, ..., n
 - $P(X = i) = \binom{n}{i} p^i (1 - p)^{n-i}$
 - E.g., coin flips

Coin Flips of Two Persons

- Your friend and you both flip coins
 - Head with probability 0.5
 - You flip 50 times; your friend flip 100 times
 - How many heads will both of you get

Joint Distribution

- Given two discrete RVs X and Y , their **joint distribution** is the distribution of X and Y together

– E.g., $P(\text{You get 21 heads AND you friend get 70 heads})$

- $$\sum_x \sum_y P(X = x \cap Y = y) = 1$$

– E.g.,

$$\sum_{i=0}^{50} \sum_{j=0}^{100} P(\text{You get } i \text{ heads AND your friend get } j \text{ heads}) = 1$$

Conditional Probability

- $P(X = x | Y = y)$ is the probability of $X = x$, given the occurrence of $Y = y$
 - E.g., you get 0 heads, given that your friend gets 61 heads
- $$P(X = x | Y = y) = \frac{P(X = x \cap Y = y)}{P(Y = y)}$$

Law of Total Probability

- Given two discrete RVs X and Y , which take values in $\{x_1, \dots, x_m\}$ and $\{y_1, \dots, y_n\}$, we have

$$\begin{aligned} P(X = x_i) &= \sum_j P(X = x_i \cap Y = y_j) \\ &= \sum_j P(X = x_i | Y = y_j) P(Y = y_j) \end{aligned}$$

Marginalization

Marginal Probability

Joint Probability

$$P(X = x_i) = \sum_j P(X = x_i \cap Y = y_j)$$

$$= \sum_j P(X = x_i | Y = y_j) P(Y = y_j)$$

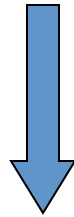
Conditional Probability

Marginal Probability

Bayes Rule

- X and Y are discrete RVs...

$$P(X = x | Y = y) = \frac{P(X = x \cap Y = y)}{P(Y = y)}$$



$$P(X = x_i | Y = y_j) = \frac{P(Y = y_j | X = x_i) P(X = x_i)}{\sum_k P(Y = y_j | X = x_k) P(X = x_k)}$$

Independent RVs

- Intuition: X and Y are independent means that $X = x$ **neither** makes it **more or less** probable that $Y = y$
- Definition: X and Y are independent iff
$$P(X = x \cap Y = y) = P(X = x)P(Y = y)$$

More on Independence

- $P(X = x \cap Y = y) = P(X = x)P(Y = y)$

A diagram consisting of a horizontal line with two arrows pointing downwards from its center. The left arrow points to the first conditional probability formula, and the right arrow points to the second conditional probability formula.

$$P(X = x|Y = y) = P(X = x) \quad P(Y = y|X = x) = P(Y = y)$$

- **E.g.**, no matter how many heads you get, your friend will not be affected, and vice versa

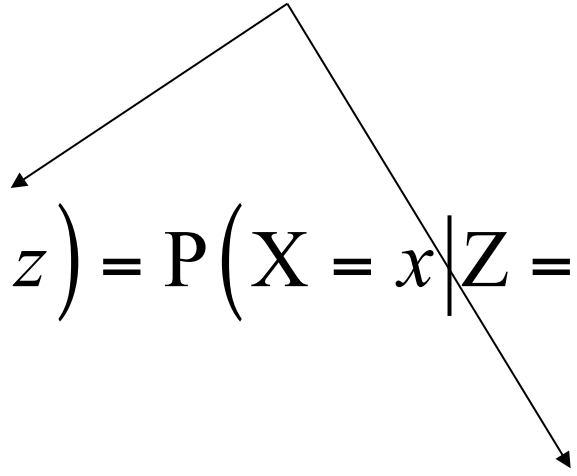
Conditionally Independent RVs

- Intuition: X and Y are conditionally independent given Z means that once Z is **known**, the value of X does not add any **additional** information about Y
- Definition: X and Y are conditionally independent given Z iff

$$P(X = x \cap Y = y | Z = z) = P(X = x | Z = z)P(Y = y | Z = z)$$

More on Conditional Independence

$$P(X = x \cap Y = y | Z = z) = P(X = x | Z = z) P(Y = y | Z = z)$$


$$P(X = x | Y = y, Z = z) = P(X = x | Z = z)$$

$$P(Y = y | X = x, Z = z) = P(Y = y | Z = z)$$

Probability of Events

- X denotes an event that could possibly happen
 - E.g., X = “you will fail in this course”
- $P(X)$ denotes the **likelihood** that X happens, or X = true
 - E.g., what’s the probability that you will fail in this course?
- Ω denotes the entire event set
 - $\Omega = \{X, \bar{X}\}$

The Axioms of Probabilities

- $0 \leq P(X) \leq 1$
- $P(\Omega) = 1$
- $P(X_1 \cup X_2 \cup \dots) = \sum_i P(X_i)$, where X_i are disjoint events
- Useful rules
 - $P(X_1 \cup X_2) = P(X_1) + P(X_2) - P(X_1 \cap X_2)$
 - $P(\bar{X}) = 1 - P(X)$

Interpreting the Axioms

