CS 70 Discrete Mathematics and Probability Theory
Fall 2025 Sabin, Hug

DIS 09B

Random Variables Intro

Note 18

Random Variable: A random variable X is a function from $\Omega \to \mathbb{R}$, mapping the possible outcomes to real numbers. Note that this function itself is not random; the *outcomes* are random. We define

$$\mathbb{P}[X = k] = \mathbb{P}[\{\omega \in \Omega : X(\omega) = k\}].$$

Distribution of a random variable: the set of all $(k, \mathbb{P}[X = k])$, describing the probability of attaining each value of the random variable.

Bernoulli Distribution: $X \sim \text{Bernoulli}(p)$; X represents the outcome of a biased coin flip. X is oftentimes also called an *indicator random variable* of an event with probability p. The distribution is described by the following:

$$\mathbb{P}[X = k] = \begin{cases} p & \text{if } k = 1 \\ 1 - p & \text{if } k = 0 \end{cases}$$

Binomial Distribution: $X \sim \text{Binomial}(n, p)$; X represents the number of successes in n independent trials, where p is the probability of success in each trial.

Geometric Distribution: $X \sim \text{Geometric}(p)$; X represents the number of independent trials until the first success (including the success), where p is the probability of success in each trial.

Poisson Distribution: $X \sim \text{Poisson}(\lambda)$; X represents the number of occurrences of an event in one unit of time, if on average there are λ occurrences in one unit of time. The distribution is described by the following:

$$\mathbb{P}[X = k] = \frac{\lambda^k}{k!} e^{-\lambda}$$

Further, if $X \sim \text{Poisson}(\lambda_x)$ and $Y \sim \text{Poisson}(\lambda_y)$ are independent, then $X + Y \sim \text{Poisson}(\lambda_x + \lambda_y)$.

1 Head Count

Note 18

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) What is $\mathbb{P}[X = k]$, for some $0 \leq k \leq 20$? Express your answer in terms of k. (Do not just copy down a formula—re-derive it yourself!)

(b) What is the name of the distribution of X, and what are its parameters?

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

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

(e) Now consider a second coin also with $\mathbb{P}[\text{Heads}] = 2/5$. Suppose you flip this second coin 30 times, and define Y to be the number of heads. What is the distribution of the *total* number of heads among these two coins, i.e. what is the distribution of $X + Y$?

2 Head Count II

Note 18

Consider a coin with $\mathbb{P}[\text{Heads}] = 3/4$. Suppose you flip the coin until you see heads for the first time, and define X to be the number of times you flipped the coin.

(a) What is $\mathbb{P}[X = k]$, for some $k \geq 1$? Express your answer in terms of k. (Do not just copy down a formula—re-derive it yourself!)

(b) What is the name of the distribution of X, and what are its parameters?

(c) What is $\mathbb{P}[X > k]$, for some $k \geq 0$? (You should not have any summations.)

(d) What is $\mathbb{P}[X < k]$, for some $k \geq 1$? (You should not have any summations.)

(e) What is $\mathbb{P}[X > k \mid X > m]$, for some $k \geq m \geq 0$? Show that your answer is equal to $\mathbb{P}[X > k - m]$. Why do we call this the memoryless property?

(f) Suppose $X \sim \text{Geometric}(p)$ and $Y \sim \text{Geometric}(q)$ are independent. Find the distribution of $\min(X,Y)$ and justify your answer.

Hint: consider flipping two coins (with $\mathbb{P}[\text{Heads}] = p$ and $\mathbb{P}[\text{Heads}] = q$ respectively) simultaneously.

3 Shuttles and Taxis at Airport

Note 18

In front of terminal 3 at San Francisco Airport is a pickup area where shuttles and taxis arrive according to a Poisson distribution. The shuttles arrive at a rate $\lambda_1 = 1/20$ (i.e. 1 shuttle per 20 minutes) and the taxis arrive at a rate $\lambda_2 = 1/10$ (i.e. 1 taxi per 10 minutes) starting at 00:00. The shuttles and the taxis arrive independently.

(a) Write in terms of a Poisson distribution, the distribution of taxis between 00:00 and 00:01:

(b) What is the distribution of the following:

(i) The number of taxis that arrive between times 00:00 and 00:20?

(ii) The number of shuttles that arrive between times 00:00 and 00:20?

(iii) The total number of pickup vehicles that arrive between times 00:00 and 00:20?

(c) What is the probability that exactly 1 shuttle and 3 taxis arrive between times 00:00 and 00:20?

(d) Given that exactly 1 pickup vehicle arrived between times 00:00 and 00:20, what is the conditional probability that this vehicle was a taxi?

(e) Suppose you reach the pickup area at 00:20. You learn that you missed 3 taxis and 1 shuttle in those 20 minutes. What is the probability that you need to wait for more than 10 mins until either a shuttle or a taxi arrives?