# Probability 2 : Conditional Probability

Anand Systla

Masters in Financial Engineering Bootcamp
UCLA Anderson

August 11, 2022

# Today's Agenda

- Discuss the previous class take home questions
- Random Variables
- PMF and PDF
- Expected Value
- Bernoulli and Binomial Distributions
- Poisson Distribution
- Uniform Distribution
- Exponential Distribution

# Simpson's Paradox

- Classic example of how conditioning and aggregation affect things. Prof. Joe Blitzstein calls conditioning the "soul of statistics" and that all probabilities are conditional.

- Whenever we have data, we want to think about how it was sampled. We need to be careful what we are conditioning on.

# Random Variables

- Formally, a Random Variable (henceforth RV) is a function that maps events to the real line. A RV $X$ is a map $X : \Omega \to \mathbb{R}$
- The randomness comes from the underlying experiment involved. Variable comes from taking many possible outcomes of an experiment
- Example 1: $\Omega = \{H, T\}$. RV can be defined as $X : \Omega \to \mathbb{R}$ i.e. $X : \{H, T\} \to \{0, 1\}$. So $X = 1$ if we get Heads, and $X = 0$ if Tails
- Example 2: Say we toss a coin 3 times. $X = \{\text{No. of H}\}$. Sketching our RV

# Probability Mass Function (PMF)

- The PMF gives us a nice way to summarize a RV $X$. It gives us the probability (mass) of the RV $X$ taking a specific value

- More formally, Say $\{a_1, a_2, ..., a_n\}$ are $n$-discrete points of a RV $X$. Then the PMF $P(.)$ gives us the probability

$$P(X = a_j) = p_j \in [0,1] \qquad \text{and} \qquad \sum_j^n p_j = 1$$

- How would the PMF for $X = \{\text{No. of H}\}$ for our 3 coin toss example look like?

# Probability Density Function (PDF)

- Used for continuous distributions. The PMF gives us the probability of specific values of RV at $X = x$. The PDF gives us the probability over an interval. So $f(x)$ is a valid PDF for a RV $X$ iff

$$\int_{-\infty}^{\infty} f(x)dx = 1 \qquad \text{and} \qquad f(x) \geq 0$$

$$P(a < X < b) = \int_{a}^{b} f(x)dx$$

# Warning 1 : Interpreting PDFs as Probabilities

- PDFs are **not** probabilities. The PDF $f(x)$ can be greater than 1, but has to integrate to 1. A simple example involving the uniform distribution

# Warning 2 : Prob. of a Continuous RV at One Point

- For a continuous RV, $P(X = x) = 0$. If $x = 1$ for example, although the PDF at $f(x = 1)$ exists and is non-zero, the probability is zero.

- How can we compute probabilities in that case?

# Cumulative Distribution Function (CDF)

- CDF is another way to summarize our RV. As the name suggests, it is a cumulative sum of the PMF/PDF functions, i.e. we're interested in $X \leq x$

$$\text{Discrete RV:} \quad F(x) = \sum_{x_k \leq x} P(X = x_k)$$

$$\text{Continuous RV:} \quad F(x) = \int_{-\infty}^{x} f(x)dx$$

- How would the CDF look for our 3 coin toss example?
- Properties of the CDF function

1. $\frac{\partial F(x)}{\partial x} = f(x)$
2. $P(X = x) = F(x) - F(x^{-})$
3. $P(x < X \leq y) = F(y) - F(x)$
4. $P(X > x) = 1 - P(X \leq x) = 1 - F(x)$
5. $F(-\infty) = 0$ and $F(\infty) = 1$

# Expectation

- Say $X = \{2, 2, 2, 3, 3, 4\}$. What is the average of this RV?
- We can write down the expectation as

$$\text{Discrete RV}: \quad E[X] = \sum_{j=1}^{n} x_j \cdot P(X = x_j)$$

$$\text{Continuous RV}: \qquad E[X] = \int_{-\infty}^{\infty} x \cdot f(x) dx$$

- Properties of the expectation operator
  1. $E[X + Y] = E[X] + E[Y]$
  2. $E[cX] = c \cdot E[X]$
  3. $\sum E[X] = E\left[\sum X\right]$

# Bernoulli Trial and Binomial Distribution

- A Bernoulli trial is an experiment with only two outcomes - $\{$Success, Failure$\}$. The probability of success is denoted as $p$ and RV $X \sim Bern(p)$

$$\text{X} = \begin{cases} 1, \text{with probability } p \\ 0, \text{with probability } 1-p \end{cases}$$

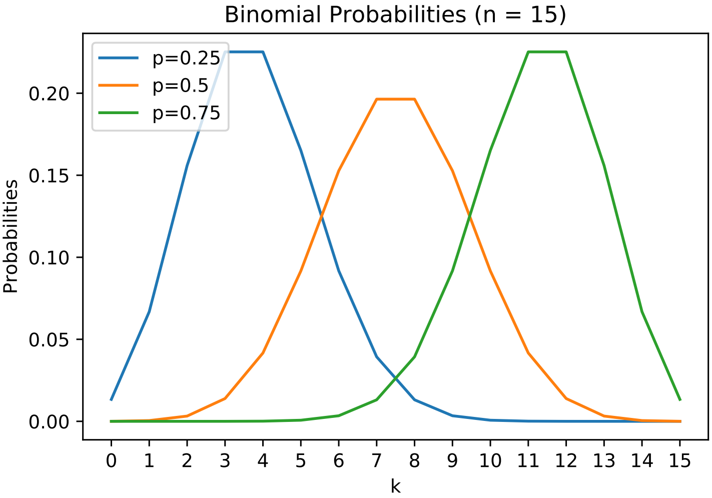
$$E[X] = 0 \cdot (1-p) + 1 \cdot p = p$$

- When we combine $n-$independent and identically distributed (i.i.d) Bernoulli trials with probability of success $p$, we get a Binomial distribution. RV $X \sim Bin(n, p)$

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

$$E[X] = \sum_{k=0}^{n} P(X = k) \cdot k = np$$

# Binomial Distribution Example



# Poisson Distribution

- Closely linked to the Binomial distribution. Poisson distribution can be derived by from counting events that occur in a given time interval if the events are occurring independently in the time interval
- Common examples - number of clicks on a website in an hour, number of buses arriving at a bus stop in a day, number of patients visiting a doctor in any 'time interval', etc.

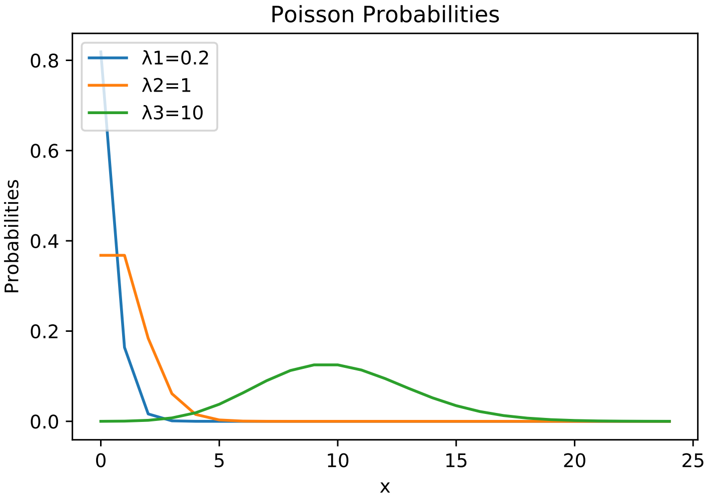
$$X \sim Pois(\lambda)$$

$$p(x) = e^{-\lambda}\frac{\lambda^x}{x!}, \qquad x = 0, 1, 2, ...$$

$$E[X] = \sum_{x=0}^{\infty} x \cdot e^{-\lambda}\frac{\lambda^x}{x!} = \lambda$$

- Key idea is that period of time is 1 and $\lambda$ can be thought of as the intensity per unit time. Also, the time interval between the arrival of two Poisson events is exponentially distributed

# Poisson Distribution Example



# Binomial To Poisson

- Say we are interested in number of people visiting a website. Lets call the RV $X = \{\text{No. of visits per hour}\}$. Prior data tells us that we see about 20 visits per hour.
- If $X$ was Binomial, $E[X] = np = 20$. We can think of $n = 60$ for example, as conducting 60 independent bernoulli trials each minute
- If $X$ was poisson, $E[X] = \lambda = 20$
- Going from Bernoulli to Poisson as $n \to \infty$

# Take Home Questions

1. Show that sum of two independent binomial distributions is also binomial, i.e. $X \sim \text{Bin}(n, p), Y \sim \text{Bin}(m, p)$ then $X + Y \sim \text{Bin}(n + m, p)$. *Hint*: You might want to use the Vandermonde's identity, which states $\sum_{i=0}^{k} {}^{n}C_i \; {}^{m}C_{k-i} = {}^{n+m}C_k$.
2. Say $X \sim \text{Pois}(3)$ and $Y \sim \text{Pois}(2)$ are two independent poisson distributions. Compute the probability that $X + Y = 5$. Additionally, compute the probability that $X = 4, Y = 1$ given $X + Y = 5$. *Hint*: Sum of two poisson distributions is also a poisson distribution. For the second part, independence assumption is key!
3. The arrival of buses at the bus stop can be modeled as a poisson process. Say buses arrive at a rate of 5 per hour ($\lambda = 5$). You just arrived at the bus stop. What is the probability that the next buss will arrive at the stop in the next 10 minutes. Additionally, find the probability that there will be a bus arriving in the next 5 minutes if there was no bus in the last 10 minutes.
4. Suppose you are given the PMF function of a random variable $X$ such that $P(X = -2) = P(X = 2) = 1/8, P(X = -1) = P(X = 1) = 1/8, P(X = 0) = 1/2$. Lets define $Y = X^2$. Compute the PMF and the CDF of the random variable $Y$.
5. Show that the Poisson PMF is a valid PMF. *Hint*: Recall, that to show validity of a PMF, we need the probabilities to sum to 1.