MA 6.101
Probability and Statistics

Tejas Bodas

Assistant Professor, IIIT Hyderabad

Conditioning with random variables

Today's class

- Conditioning X on an event $A \in \mathcal{F}$.
- Conditional Expectation $E[X|A]$.
- Conditioning X with disjoint partitions $\{A_i\}$ of Ω.
- Conditioning X on an event $\{X \in A\} \in \mathcal{F}'$
- Conditioning X on another random variable Y.
- Conditional expectation $E[X|Y = y]$.

A new running example

- Pick 2 integers from $\{1, 2, 3\}$ without replacement.
- $\Omega = \{(1, 2), (1, 3), (2, 1), (2, 3), (3, 1), (3, 2)\}$
- $\mathbb{P}\{\omega\} = \frac{1}{6}$ for all $\omega \in \Omega$.
- Denote them by random variables X and Y.
- For $\omega = (1, 3)$ $X(\omega) = 1$ and $Y(\omega) = 3$.
- Write down their joint PMF $p_{X,Y}(x, y)$.
- Write down their marginal PMFs p_X and p_Y?
- What is $E[X], E[Y]$ and $E[XY]$?

Remember Conditional probability?

- Given/If dice rolls odd, what is the probability that the outcome is 1?
- Given/If $\bar{\omega} \in [0, 0.5]$ what is the probability that $\bar{\omega} \in [0, 0.25]$?
- The conditional probability of event B given event A is defined as $\mathbb{P}(B/A) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}$ when $\mathbb{P}(A) > 0$.

Conditioning on an event A

- Consider a discrete r.v. X with pmf $p_X(x)$. Suppose an event A has happened where $A \in \mathcal{F}$.

- Consider event $\{\omega \in \Omega : X(\omega) = x\}$. We will use shorthand $\{X = x\}$.

- What is $\mathbb{P}(X = x|A)$? $\mathbb{P}(X = x|A) = \frac{\mathbb{P}(\{X=x\}\cap A)}{\mathbb{P}(A)}$.

$$p_{X|A}(x) := \frac{\mathbb{P}(\{X=x\}\cap A)}{\mathbb{P}(A)}.$$

- $p_{X|A}(x)$ denotes the conditional PMF of X under event A.

- In the running example say A is the event that the first number is odd and second is even. $A = \{(1,2),(3,2)\}$. Compute $p_{X|A}(\cdot)$.

- How do we know that it is consistent, i.e., $\sum_x p_{X|A}(x) = 1$?

Consistency of conditional PMF

$$\sum_x p_{X|A}(x) = 1.$$

Proof:

- $\sum_x p_{X|A}(x) = \sum_x \frac{\mathbb{P}(\{X=x\}\cap A)}{\mathbb{P}(A)}$
- $\{\omega \in \Omega : X(\omega) = x\}$ are disjoint sets for different x.
- From theorem of total probability, this implies that $\{X = x\} \cap A$ are disjoint sets for all x.
- $\sum_x p_{X|A}(x) = \frac{\mathbb{P}(\bigcup_x \{\{X=x\}\cap A)\}}{\mathbb{P}(A)} = \frac{\mathbb{P}(A)}{\mathbb{P}(A)} = 1.$ $\square$

Another Example

- Lets X denote the outcome of a dice.
- Let A denote the event that the roll is odd.
- What is $p_{X|A}(x)$?
- Given that event A has happened, what is the average value of the dice, i.e., $E[X|A]$?

$$E[X/A] = \sum_x x p_{X|A}(x).$$

Using LOTUS,

$$E[g(X)/A] = \sum_x g(x) p_{X|A}(x).$$

Today's class

- Conditioning X on an event $A \in \mathcal{F}$.
- Conditional Expectation $E[X|A]$.
- Conditioning X with disjoint partitions $\{A_i\}$ of Ω.
- Conditioning X on an event $\{X \in A\} \in \mathcal{F}'$
- Conditioning X on another random variable Y.
- Conditional expectation $E[X|Y = y]$.

Conditioning with disjoint partitions

- Now let $\{A_i, i = 1, 2, \ldots, n\}$ be a disjoint partition of Ω.
- Prove the following using law of total probability

$$p_X(x) = \sum_{i=1}^n \mathbb{P}(A_i) p_{X|A_i}(x)$$

Proof:

- $\sum_{i=1}^n \mathbb{P}(A_i) \frac{\mathbb{P}(\{X=x\} \cap A_i)}{\mathbb{P}(A_i)} = \sum_{i=1}^n \mathbb{P}(\{X = x\} \cap A_i) = \mathbb{P}(\{X = x\})$. □
- The last equality follows from the law of total probability.
- An important consequence is the following.

$$E[X] = \sum_{i=1}^n \mathbb{P}(A_i) E[X|A_i]$$

Today's class

- Conditioning X on an event $A \in \mathcal{F}$.
- Conditional Expectation $E[X|A]$.
- Conditioning X with disjoint partitions $\{A_i\}$ of Ω.
- Conditioning X on an event $\{X \in A\} \in \mathcal{F}'$
- Conditioning X on another random variable Y.
- Conditional expectation $E[X|Y = y]$.

Conditioning on event $X \in A$

- Consider a discrete r.v. X with pmf $p_X(x)$. Suppose an event $X \in A$ has happened where $A \in \mathcal{F}'$.
- $X \in A = \{\omega \in \Omega : X(\omega) \in A\}$ and $\mathbb{P}\{X \in A\} = \sum_{x \in A} p_X(x)$.
- We will use the same notation $p_{X|A}(x) := \frac{\mathbb{P}(\{X=x\} \cap \{X \in A\})}{\mathbb{P}(X \in A)}$.
- If $x \notin A$, we have $p_{X|A}(x) = 0$.
- Otherwise (when $x \in A$,), we have $p_{X|A}(x) = \frac{p_X(x)}{\mathbb{P}(X \in A)}$.
- Running example: Suppose we are given $X \in A$ where $A = \{2, 3\}$. What is $p_{X|A}(x)$?

Revisiting Geometric random variable

- Let N be a geometric random variable with parameter p.
- Its pmf is $p_N(k) = (1-p)^{k-1}p$.
- Suppose we are given the event $A := N > n$. $P(A) = (1-p)^n$.
- What is $p_{N|A}(k)$?
- For $k > n$, $p_{N|A}(k) = \frac{P\{(N>n) \cap N=k\}}{P(N>n)} = (1-p)^{k-1-n}p$. For $k \leq n$, we have $p_{N|A}(k) = 0$.

Memoryless property of Geometric random variable

- What is $P(N > n + m | N > n)$?
- $P(N > n + m | N > n) = \frac{P(N > n+m)}{P(N > n)} = (1 - p)^m = P(N > m)$.
- If N denotes number of tosses till you first get a head, and having already tossed more than n times, the probability of having to toss more than $n + m$ is same as starting the experiment (forgetting that you have already tossed more than n times) fresh and having to toss more than m times.
- How much you have tossed till now has no bearing on how much you will be required to toss.

$$\boxed{P(N > n + m | N > n) = P(N > m) \text{ (Memoryless property).}}$$

HW: Find $E[N|A]$ where event $A = \{N > n\}$ and $n > 0$.