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:

- $\sum_{i=1}^n \mathbb{P}(A_i) \frac{\mathbb{P}(\{X=x\}\cap A_i)}{\mathbb{P}(A_i)} = \sum_{i=1}^n \mathbb{P}(\{X = x\} \cap A_i) = \mathbb{P}(\{X = x\}).$ $\square$
- 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 $k \le n$, we have $p_{N|A}(k) = 0$.

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 $E[X|Y = y]$.
- Law of iterated expectation $E[X|Y]$
- Bayes rule revisited
- Sums of random variables.

Conditioning X on random variable Y

- Consider a discrete r.v's X and Y with joint pmfs $p_{XY}(x,y)$ and with marginal pmf $p_X(x)$ and $p_Y(y)$.
- Suppose an event $A : \{Y = y\}$ has happened and we are interested in the probability that $X = x$ given $Y = y$.
- This conditional pmf is denoted by $p_{X|Y}(x|y)$.
- In fact, $p_{X|Y}(x|y) := \frac{P(X=x, Y=y)}{P(Y=y)} = \frac{p_{X,Y}(x,y)}{p_Y(y)}$.

$$\boxed{p_{X,Y}(x,y) = p_{X|Y}(x|y)p_Y(y)}$$

- This is essentially same as $P(A \cap B) = P(A|B)P(B)$
- Is $p_{X|Y}(x|y)$ consistent?
- $\sum_x p_{X|Y}(x|y) = \sum_x \frac{p_{X,Y}(x,y)}{p_Y(y)} = 1.$

What if X and Y are independent ?

- When do we say that X and Y are independent ? When $p_{X,Y}(x,y) = p_X(x)p_Y(y)$.
- We also know that

$$\boxed{p_{X,Y}(x,y) = p_{X|Y}(x|y)p_Y(y)}$$

- This implies that $p_{X|Y}(x|y) = p_X(x)$.
- NOTE: Independence implies $E[XY] = E[X]E[Y]$.

Independent random variables are uncorrelated ($Cov(X,Y) = 0$). But Uncorrelated random variables need not be independent!! (See Example 4.13 in Bertsekas)

Conditioning X on random variable Y

$$p_{X,Y}(x,y) = p_{X|Y}(x|y)p_Y(y)$$

- Now summing on both sides over y, we have

$$p_X(x) = \sum_y p_{X|Y}(x|y)p_Y(y)$$

- Similarly from $p_{X,Y}(x,y) = p_{Y|X}(y|x)p_X(x)$, summing on both sides over x, we have

$$p_Y(y) = \sum_x p_{Y|X}(y|x)p_X(x)$$

- Notice similarity to the law of total probability. $P(A) = \sum_i P(A|B_i)P(B_i)$.