RECAP

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

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

For disjoint partitions $A_i \in \mathcal{F}$, $p_X(x) = \sum_{i=1}^n \mathbb{P}(A_i) p_{X|A_i}(x)$

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

For $A \in \mathcal{F}'$, $p_{X|A}(x) = \frac{p_X(x)}{\mathbb{P}(X\in A)}$ if $x \in A$ and 0 otherwise.

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 ?

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. $P(A) = \sum_i P(A|B_i)P(B_i)$.

Conditional expectation $E[X|Y=y]$

It is easy to guess that

$$E[X|Y=y] := \sum_x x p_{X|Y}(x|y)$$
$$E[Y|X=x] := \sum_y y p_{Y|X}(y|x)$$

Can you write $E[X]$ in terms of $E[X|Y=y]$?

$$E[X] = \sum_y p_Y(y) E[X|Y=y]$$

Proof:
$$\begin{aligned}\sum_y p_Y(y)E[X|Y=y] &= \sum_y p_Y(y)\sum_x x p_{X|Y}(x|y)\\ &= \sum_x \sum_y x p_{X,Y}(x,y)\\ &= \sum_x x p_X(x)\\ &= E[X]\end{aligned}$$

Summary

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

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

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

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

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

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

$$E[X|Y=y] := \sum_x x p_{X|Y}(x|y)$$

$$E[X] = \sum_y p_Y(y) E[X|Y=y]$$

How about all this for continuous X & Y?

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

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

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

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

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

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

$$E[X|Y=y] := \sum_x x p_{X|Y}(x|y)$$

$$E[X] = \sum_y p_Y(y) E[X|Y=y]$$

How about all this for continuous X & Y?

$$\int_{x \in B} f_{X|A}(x)dx = \mathbb{P}(X \in B|A).$$

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

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

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

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

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

$$E[X|Y=y] := \sum_x x p_{X|Y}(x|y)$$

$$E[X] = \sum_y p_Y(y) E[X|Y=y]$$

How about all this for continuous X & Y?

$$\int_{x\in B} f_{X|A}(x) = \mathbb{P}(X \in B|A).$$

$$E[X/A] = \int_{-\infty}^{\infty} x f_{X|A}(x)dx$$

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

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

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

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

$$E[X|Y=y] := \sum_x x p_{X|Y}(x|y)$$

$$E[X] = \sum_y p_Y(y)E[X|Y=y]$$

How about all this for continuous X & Y?

$$\int_{x \in B} f_{X|A}(x) = \mathbb{P}(X \in B|A).$$

$$E[X/A] = \int_{-\infty}^{\infty} x f_{X|A}(x) dx$$

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

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

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

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

$$E[X|Y=y] := \sum_x x p_{X|Y}(x|y)$$

$$E[X] = \sum_y p_Y(y) E[X|Y=y]$$

How about all this for continuous X & Y?

$$\int_{x\in B} f_{X|A}(x) = \mathbb{P}(X \in B|A).$$

$$E[X/A] = \int_{-\infty}^{\infty} x f_{X|A}(x)dx$$

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

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

$$f_{X,Y}(x,y) = f_{X|Y}(x|y) f_Y(y)$$

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

$$E[X|Y=y] := \sum_x x p_{X|Y}(x|y)$$

$$E[X] = \sum_y p_Y(y) E[X|Y=y]$$

How about all this for continuous X & Y?

$$\int_{x \in B} f_{X|A}(x) = \mathbb{P}(X \in B|A).$$

$$E[X/A] = \int_{-\infty}^{\infty} x f_{X|A}(x) dx$$

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

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

$$f_{X,Y}(x,y) = f_{X|Y}(x|y) f_Y(y)$$

$$f_X(x) = \int_y f X|Y(x|y) f_Y(y) dy$$

$$E[X|Y=y] = \int_x x f_{X|Y}(x|y) dx$$

$$E[X] = \int_y E[X|Y=y] f_Y(y) dy$$

Conditional expectation $E[X|Y]$

Recall that

$$E[X|Y = y] := \sum_x x p_{X|Y}(x|y)$$

- Is $E[X|Y = y]$ a constant? Is it a function of y?
- $E[X|Y = y]$ is a function of y.
- Now consider $E[X|Y]$. Is it still a function of y ?
- $E[X|Y]$ is a function of Y, say $g(Y)$.
- When Y takes the value y, (this happens with probability $p_Y(y)$) $E[X|Y]$ takea the value $E[X|Y = y]$.
- What is the expectation of $E[X|Y]$?

Conditional expectation $E[X|Y]$

- $g(Y) = E[X|Y]$.
- What is $E[g(Y)] = E[E[X|Y]]$?
- $E[g(Y)] = \sum_y g(y)p_Y(y) = \sum_y E[X|Y = y]p_Y(y)$.
- This implies $E[g(Y)] = E[E[X|Y]] = E[X]$. This is the law of iterated expectation.

$$\boxed{E[E[X|Y]] = E[X]}$$

Conditional expectation $E[X|Y]$ – Example

- Consider $Y = \begin{cases} \lambda_1 \text{ with prob } p \\ \lambda_2 \text{ with prob } 1-p \text{ .} \end{cases}$
- Now consider an exponential random variable X with a random parameter Y.
- What is $E[X]$?
- $E[X] = E[E[X|Y]] = \sum_y E[X|Y=y]p_Y(y)$
- We have $X \sim Exp(\lambda_1)$ with probability p when $Y = \lambda_1$.
- Similarly $X \sim Exp(\lambda_2)$ with probability $1-p$ when $Y = \lambda_2$.
- $E[X|Y=\lambda_i] = \frac{1}{\lambda_i}$
- $E[X] = \frac{p}{\lambda_1} + \frac{1-p}{\lambda_2}$.

Conditional expectation $E[X|Y]$ – Example 2

- Consider $Y = X_1 + X_2 + \ldots X_N$ where N is a positive integer valued r.v. with PMF $p_N(\cdot)$ and $X_i's$ are independent and identically distributed (i.i.d) with mean $E[X]$.
- What is $E[Y]$? Use $E[Y] = E[E[Y|N]]$.
- What is $E[Y|N = n]$?
- $E[Y|N = n] = E[X_1 + X_2 + \ldots X_n] = nE[X]$.
- This implies $E[Y|N] = NE[X]$.
- Now $E[Y] = E[E[Y|N]] = E[NE[X]] = E[X]E[N]$.
- What is $Var(Y)$? (section 4.5)