Conditional Expectation and Independence

$\mathbb{E}_{\mathcal{G}}[X]$ is the "best guess" at X using only information in $\mathcal{G}$. What if $\mathcal{G}$ has no information about X?

Prop: Let $X : (\Omega, \mathcal{F}) \to (S, \mathcal{B})$ be a random variable, and let $\mathcal{G} \subseteq \mathcal{F}$ be a sub-σ-field.

If $\sigma(X), \mathcal{G}$ are independent, and $f : S \to \mathbb{R}$ is s.t. $f(X) \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, then

$$\mathbb{E}_{\mathcal{G}}[f(X)] = \mathbb{E}[f(X)] \quad \text{a.s.}$$

Conversely, if $\nearrow$ holds for all $f \in B(S, \mathcal{B})$, then $\sigma(X), \mathcal{G}$ are independent.

Pf. ($\Longrightarrow$) Let $Y \in B(\Omega, \mathcal{G})$. Then $\mathbb{E}[f(X)Y] = \underbrace{\mathbb{E}[f(X)]}_{a}\mathbb{E}[Y] = \mathbb{E}[aY]$

$\Rightarrow \mathbb{E}_{\mathcal{G}}[X] = a = \mathbb{E}[f(X)]$ a.s.

($\Longleftarrow$) $\mathbb{E}[f(X)Y] = \mathbb{E}[\mathbb{E}_{\mathcal{G}}[f(X)]Y] = \mathbb{E}[f(X)]\mathbb{E}[Y]$.

$\Rightarrow$ $\sigma(X), \mathcal{G}$ are independent. ///

Conditioning on a Random Variable / Vector

If $X : (\Omega, \mathcal{F}) \to (S, \mathcal{B})$ (think $S = \mathbb{R}^d$),
and $Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, we denote

$$\mathbb{E}_{\sigma(X)}[Y] = \mathbb{E}[Y \mid \sigma(X)] =: \mathbb{E}[Y \mid X]$$

This is in $L^1(\Omega, \sigma(X), \mathbb{P})$. In particular, it is $\sigma(X)/\mathcal{B}(\mathbb{R})$-measurable.
By the Doob-Dynkin representation,
there is a $\mathcal{B}/\mathcal{B}(\mathbb{R})$-measurable function $f_Y : S \to \mathbb{R}$
s.t. $\mathbb{E}[Y \mid X] = f_Y(X)$.

Notation: $f_Y(s) =: \mathbb{E}[Y \mid X = s]$

$$\left(= \frac{\mathbb{E}[Y \mathbb{1}_{\{X = s\}}]}{\mathbb{P}(X = s)} \right)$$

Equivalently: $f_Y : S \to \mathbb{R}$ is characterized by

$$\mathbb{E}[Y \cdot h(X)] = \mathbb{E}[f_Y(X) h(X)] \quad \forall h \in B(S, \mathcal{B})$$

$\hookrightarrow h = \mathbb{1}_{\{s\}} \; \therefore \; \mathbb{E}[Y \mathbb{1}_{\{X=s\}}] = \mathbb{E}[f_Y(s) \mathbb{1}_{\{X=s\}}] = f_Y(s) \mathbb{P}(X = s)$

If X, Y are independent, "Y is constant wrt X".
We can make this precise as follows.

Prop: Let $X : (\Omega, \mathcal{F}) \to (S, \mathcal{B})$, $Y : (\Omega, \mathcal{F}) \to (T, \mathcal{C})$ be random variables. Let $\mathbb{P}$ be a probability measure on $(\Omega, \mathcal{F})$. If X, Y are independent, and $f \in \mathbb{B}(S \times T, \mathcal{B} \otimes \mathcal{C})$, then

$$\mathbb{E}[f(X,Y) \mid X = x] = \mathbb{E}[f(x,Y)] = \int_T f(x,y)\mu_Y(dy).$$

I.e. $\mathbb{E}[f(X,Y) \mid X] = \mathbb{E}[f(x,Y)]\big|_{x=X}$

Pf. Since X, Y are independent, $\mu_{(X,Y)} = \mu_X \otimes \mu_Y$. Thus, if $h \in B(S, \mathcal{B})$,

$$\mathbb{E}[f(X,Y)h(X)] = \iint_{S \times T} f(x,y)h(x)\,\mu_X \otimes \mu_Y(dx\,dy)$$

$$= \int_S \mu_X(dx)\, h(x) \underbrace{\int_T \mu_Y(dy)\, f(x,y)}_{\mathbb{E}[f(x,Y)] = f_Y(x)} = \int_S f_Y(x) h(x) \mu_X(dx) = \mathbb{E}[f_Y(X) h(X)].$$

///

Eg. Suppose (X,Y) has a joint density $\rho = \rho_{X,Y}$.

We want to identify $\mathbb{E}[f(X,Y)\,|\,X] = g(X)$

This means we want, $\forall\, h \in \mathbb{B}(\mathbb{R}, \mathcal{B}(\mathbb{R}))$,

$$\mathbb{E}[\,\overbrace{\mathbb{E}[f(X,Y)\,|\,X]}\,h(X)] = \mathbb{E}[f(X,Y)\,h(X)]$$

$$\mathbb{E}[\,g(X)\,h(X)] \qquad \qquad \iint f(x,y)\,h(x)\,\rho(x,y)\,dx\,dy.$$

$$\int g(x)\,h(x)\,\mu_X(dx)$$

Note: since (X,Y) has a joint density, X has a marginal density

$$\mathbb{P}(X \in A) = \mathbb{P}((X,Y) \in A \times \mathbb{R}) = \iint_{A \times \mathbb{R}} \rho(x,y)\,dx\,dy = \int_A \left(\int_{\mathbb{R}} \rho(x,y)\,dy\right) dx$$

I.e. $\rho_X(x) = \int_{\mathbb{R}} \rho(x,y)\,dy$

$$\int h(x)\,g(x)\,\rho_X(x)\,dx \qquad \qquad \int h(x) \left(\int dy\, f(x,y)\,\rho(x,y)\right) dx.$$

Prop. Let (X,Y) have density $\rho = \rho_{X,Y}$.
Let $\rho_X(x) = \int \rho_{X,Y}(x,y)\,dy$ be the marginal density of X. Define

$$\rho_{Y|X}(y|x) := \mathbb{1}_{0<\rho_X<\infty} \frac{\rho_{X,Y}(x,y)}{\rho_X(x)}$$

Then for $f \in \mathcal{B}(\mathbb{R}^2)$, $\qquad E[f(X,Y) \mid X=x] \approx$

$$\mathbb{E}[f(X,Y) \mid X] = g(X) \quad \text{where } g(x) = \int f(x,y)\,\rho_{Y|X}(y|x)\,dy.$$

Pf. We saw on the last slide, $\uparrow$ holds provided g satisfies

$$g(x)\,\rho_X(x) = \int f(x,y)\,\rho_{X,Y}(x,y)\,dy \quad \text{for } [\lambda] \text{ a.e. } x \in \mathbb{R}.$$

$\uparrow$

We defined g to make this true if $\rho_X(x) > 0$.

Claim: If $\rho_X(x) = 0$ then $\left|\int f(x,y)\,\rho_{X,Y}(x,y)\,dy\right| = 0$.

$$\leq \sup_y |f(x,y)| \underbrace{\int |\rho_{X,Y}(x,y)|\,dy}_{\rho_X(x)}.$$

///