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

Pf. $(\Rightarrow)$ Let $Y \in B(\Omega, \mathcal{G})$. Then $\mathbb{E}[f(X)Y]$

$\downarrow$

$(\Leftarrow)$ $\mathbb{E}[f(X)Y]$

Conditioning on a Random Variable / Vector

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

$$E_{\sigma(X)}[Y] = E[Y \mid \sigma(X)]$$

This is in $L^1(\Omega, \sigma(X), P)$. 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. $E[Y \mid X] = f_Y(X)$.

Notation: $f_Y(s) =:$

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

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

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

$$\mathbb{E}[f(X,Y) \mid X = x] = \mathbb{E}[f(x,Y)]$$

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 $h \in \mathbb{B}(S, \mathcal{B})$,

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

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

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

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

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

Note: since (X,Y) has a joint density, X has a 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. $\rho_X(x) = \int_{\mathbb{R}} \rho(x,y)\, dy$

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

$$\rho_{Y|X}(y|x) := \frac{\rho_{X,Y}(x,y)}{\rho_X(x)}$$

Then for $f \in \mathbb{B}(\mathbb{R}^2)$,

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

Pf. We saw on the last side, $\uparrow$ 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