Recall, conditional expectation is the \( \mathcal{G} \)-measurable function \( \mathbb{E}[X \mid \mathcal{G}] \) such that
\[
\int_A \mathbb{E}[X \mid \mathcal{G}] \, d\mathbb{P} = \int_A X \, d\mathbb{P}
\]
for every \( \mathcal{G} \)-measurable set \( A \).

\( \mathbb{E}[X \mid Y] \) is as above with \( \mathcal{G} \) replaced by \( \sigma(Y) \) and \( Y \) measurable.

This reduces to a function of \( Y \).

E.g., for a proof,

\[
P_{X \mid Y} = \frac{P_{X,Y}}{P_Y}
\]

\( \sigma(Y) \) is the \( Y \) marginal distribution, and \( \sum_j P_{X,Y}(x_i, y_j) = P_Y(y_j) \).

Claim
\[
\mathbb{E}\left[ I_{X_i}(X) \mid Y = y_j \right] w_i = P_{X\mid Y}(x_i \mid y_j)
\]

for \( w_i \in Y^{-1}(y_j) \).
This also works for functions with a pdf

\[ h_{x,y}(x,y) = \text{joint density function of } X \text{ and } Y \]

\[ h_y(y) = \int_{\mathbb{R}} h_{x,y}(x,y) \, dx \]

\[ h_{x|y}(x,y) = \frac{h_{x,y}(x,y)}{h_y(y)} \]

\[ E[f(x) \mid Y = y] = \int_{\mathbb{R}} f(x) \, h_{x|y}(x,y) \, dx \]

**Eq. 4**

Let \( X, Y \) be independent. Then

\[ E[\Phi(x,y) \mid Y = y] = E[\Phi(x,y)] \]

By suff to show

\[ E[\Phi(x,y) \mid Y = y] = \int_{\mathbb{R}^2} E_x[\Phi(x,y)] \, dP = \int_{\mathbb{R}^2} \Phi(x,y) \, dP \]

\[ Y \in (a) \]

\[ \Rightarrow \text{suff to show} \int_{\mathbb{R}^2} E_x[\Phi(x,y)] \, dP = \int_{\mathbb{R}^2} \Phi(x,y) \, dP \]

\[ \Rightarrow \int_{(a) \subset \mathbb{R}^2} \Phi(x,y) \, dP \]

\[ = \int_{(a) \subset \mathbb{R}^2} \Phi(x,y) \, d\mu_x \, d\mu_y \]

\[ = \int_{(a) \subset \mathbb{R}^2} \Phi(x,y) \, d\mu \]

\[ \Rightarrow \int_{\mathbb{R}^2} E_x[\Phi(x,y)] \, d\mu_y \]

\[ = \int_{\mathbb{R}^2} E_x[\Phi(x,y)] \, d\mu_y \]
The converse statement is false for non-independent \( X, Y \) e.g. 
\[
E[X] = E[Y] = 0 \\
\rho(X, Y) = x + y
\]
Then
\[
E[X + Y] = y \\
E[X + Y | Y = y] = 0
\]

\[
E[1_B(Y) E_X[\rho(X, Y)]] = \int E_X[\rho(X, Y)] dP = \text{LHS of (2.4)}
\]

Converse counterexample
\[
X = Y \\
\rho(X, Y) = x + y \\
E[X + Y | Y = y] = y
\]
But
\[
E[X + Y] = 1_y
\]

Consult notes pp 5-6 Ch 2 for properties of conditional inequality

\[
E[xy | G] \leq E[x^2 | G] E[y^2 | G] \\
\int E[xy | G] dP \leq \int E[x^2 | G] E[y^2 | G] dP
\]
\[
0 \Rightarrow \int E|xG|_p \leq \int X dP \quad \text{and} \quad p > 1
\]
\[
\int |E|xG|_p dP \leq \int E|X|^p |G| dP = \int (X^p dP = \|X\|_p^p)
\]
\[ E[E[X|G_2]] = E[X] \]

"smallest algebra wins"

\textbf{Thm 3} (Ch. 2 p. 6)

a) \[ \int E[X|G] \, dP = \int X \, dP = E[X] \]

b) \[ E\left[E[X|G_1] \mid G_2\right] = E\left[E[X|G_1] \mid G_1\right] = E[X|G_2] \]

\[ \text{P. a.e. in } G_1, \quad G_1 \subseteq G_2 \]

\[ E[X|G_1] \text{ is } G_1 \text{-measurable} \Rightarrow E\left[E[X|G_1] \mid G_2\right] = E[X|G_2] \]

\text{w.r.t.} \[ E\left[E[X|G_1] \mid G_2\right] = E[X|G_1] \]

\[ \text{b) } \text{p} \text{-measurable. \ Sufficient to show } \forall A \subseteq G_2 \]

\[ \int_A E\left(E[X|G_1] \mid G_2\right) \, dP = \int_A E[X|G_1] \, dP \]

\[ = \int_A E[X|G_2] \, dP \]

\[ \text {w.r.t. } A \subseteq G_1, A \subseteq G_2 \text{ so } \]

\[ \int_A X \, dP \]

\[ \text{if } Y \text{ already } G\text{-measurable, } \]

\[ \int E[X \mid G] \, dP = YE[X|G] \]

\[ \text{w.r.t. independence not needed.} \]
In $L^2(\Omega, \mathcal{G}, P)$, $E[X \mid \mathcal{G}]$ is orthogonal projection onto $L^2(\Omega, \mathcal{G}, P)$.

The remainder $X - E[X \mid \mathcal{G}]$ is orthogonal to $X$, is orthogonal

\[
E[X^2 - X E[X \mid \mathcal{G}]] = E[X^2] - E[X E[X \mid \mathcal{G}]]
\]

\[
= \int_{\Omega} x^2 - \int_{\Omega} x x = 0
\]

\[\text{c) } \text{Var}(X) = E[\text{Var}(X \mid \mathcal{G})] + \text{Var}(E[X \mid \mathcal{G}])
\]

\[\text{if } E[X] = \text{Var}(X) = 0 \implies \text{Var}(E[X \mid \mathcal{G}]) = 0
\]

\[\implies \text{Var}(X) = \|X\|^2
\]

**Regular Conditional Probabilities**

$X : \Omega \rightarrow \mathbb{R}$, $\mu_x(\cdot) = E[1_B(X) \mid \mathcal{G}]$

$X \in \mathcal{G} \implies$

$\mu_x(B \mid \mathcal{G})(\omega) = E[1_B(X) \mid \mathcal{G}](\omega)$

$\mu_x(B, \omega)$ is called a regular conditional probability if

1) $\mu_x(B, \omega)$ is a version of $E[1_B(X) \mid \mathcal{G}]$

2) $\forall \omega \mu_x(\cdot, \omega)$ is a probability measure on $\mathcal{G}$
Standard Borel spaces are those measurable spaces that are homeomorphic to a Borel space on $\mathbb{R}$, i.e., $\mathcal{E}$, $\mathcal{E}^{-1}$ measurable, like Borel spaces on $\mathbb{R}$ vector spaces.

For next time, prepare on Martingales.

A measurable space $(M, \Sigma)$ is a standard Borel space if we have a bijection $\varphi: M \rightarrow \mathbb{R}$ s.t. $\varphi, \varphi^{-1}$ both measurable.

Next time:

Martingales

Read & prepare to talk.