The averaging property is the major workhorse for conditioning: $\mathbb{E}_{\mathcal{G}}[X] \in L^1(\Omega, \mathcal{G}, \mathbb{P})$ and

$$\mathbb{E}[\mathbb{E}_{\mathcal{G}}[X]Y] = \mathbb{E}[XY] \quad \forall X \in L^1(\Omega, \mathcal{F}, \mathbb{P}) \quad Y \in B(\Omega, \mathcal{G})$$

Eg. $\mathbb{E}_{\mathcal{F}}[X] = X$ b/c $\mathbb{E}[ZY] = \mathbb{E}[XY] \quad \forall Y \in B(\Omega, \mathcal{F})$

$\uparrow$ $\mathcal{F}$-meas.

$$\Downarrow$$

$$\mathbb{E}[(Z-X)Y] = 0 \qquad Y = \operatorname{sgn}(X-Z)\mathbb{1}_{|X-Z| \le n}$$

$$\therefore 0 = \mathbb{E}[|Z-X|\mathbb{1}_{|X-Z| \le n}] \xrightarrow[DCT]{} \mathbb{E}[|Z-X|]$$

$\therefore Z = X$ a.s.

Eg. If $\mathcal{G} = \{\emptyset, \Omega\}$, $L^0(\Omega, \mathcal{G}) = \{\text{const. fns}\}$ $\quad Y^{-1}\{t\} = \emptyset$ or Ω

$\mathbb{E}_{\mathcal{G}}[X]$ is $\mathcal{G}$-measurable, $\therefore \mathbb{E}_{\mathcal{G}}[X] = \mathbb{E}[X]$ a.s.

Eg. $\mathbb{E}[\mathbb{E}_{\mathcal{G}}[X]] = \mathbb{E}[\mathbb{E}_{\mathcal{G}}[X] \cdot 1] = \mathbb{E}[X \cdot 1] = \mathbb{E}[X]$

$$\langle P_k(X), 1 \rangle = \langle X, P_k(1) \rangle = \langle X, 1 \rangle$$

Lemma: $Z = \mathbb{E}_{\mathcal{G}}[X]$ iff $Z \in L^1(\Omega, \mathcal{G}, \mathbb{P})$, and

$$\mathbb{E}[Z : B] = \mathbb{E}[X : B] \quad \forall B \in \mathcal{G}.$$

Pf. ($\Rightarrow$) $\mathbb{1}_B \in B(\Omega, \mathcal{G})$ $\quad \mathbb{E}[Z \mathbb{1}_B] = \mathbb{E}[X \mathbb{1}_B]$

($\Leftarrow$) Dynkin's Multiplicative Systems theorem. ///

Eg. $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2, \mathbb{P}_1 \otimes \mathbb{P}_2)$

$$\mathcal{G} = \{ A \times \Omega_2 : A \in \mathcal{F}_1 \} \quad (\subseteq \mathcal{F}_1 \otimes \mathcal{F}_2) \qquad \phi \times \Omega_2 = \phi.$$

For $X \in L^1(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2, \mathbb{P}_1 \otimes \mathbb{P}_2)$, what is $\mathbb{E}_{\mathcal{G}}[X] = Z$?

$$\mathbb{E}[Z \mathbb{1}_{A \times \Omega_2}] = \mathbb{E}[X \mathbb{1}_{A \times \Omega_2}] \quad \forall A \in \mathcal{F}_1$$

$$\int_{\Omega_2} \left(\int_A Z'(\omega_1)\, \mathbb{P}_1(d\omega_1) \right) \mathbb{P}_2(d\omega_2) = \int_A \left(\int_{\Omega_2} X(\omega_1, \omega_2)\, \mathbb{P}_2(d\omega_2) \right) \mathbb{P}_1(d\omega_1)$$

$$= \int_A Z'\, d\mathbb{P}_1$$

Z is $\mathcal{G}$-meas.

$Z^{-1}\{t\} \in \mathcal{G} = \mathcal{F}_1 \times \Omega_2$

I.e. $Z(\omega_1, \omega_2) = t$.

$\Rightarrow Z(\omega_1, \omega_2') = t$ (〃)

$Z(\omega_1, \omega_2) = Z'(\omega_1)$

$$\therefore \quad Z(\omega_1, \omega_2) = \int_{\Omega_2} X(\omega_1, \omega_2)\, \mathbb{P}_2(d\omega_2)$$

Remark: "Conditional" is a terrible name for this thing!

$\mathbb{E}[X \mid \mathcal{G}]$ ← not "putting conditions" on X

The bigger $\mathcal{G}$ is, the less constrained $\mathbb{E}[X \mid \mathcal{G}]$ is!

Every time you here "conditioned on", say to yourself "projected on" and everything will make much more sense.

The last example showed $\mathbb{E}_{\mathcal{G}}[\,\cdot\,]$ can mean partial integration: integrating out some (but not necessarily all) variables. This is the sense in which it still makes sense to call it "expectation".
(A much better name would be "partial expectation, given $\mathcal{G}$".)

→ This motivates the fact that every integral inequality / convergence theorem has a "conditional" version.