Given $(\Omega, \mathcal{F}, \mathbb{P})$ and $A \in \mathcal{F}$ with $\mathbb{P}(A) > 0$,
we have the conditioned probability measure

$$\mathbb{P}_A(B) = \mathbb{P}[B \mid A] = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}.$$

We can then take integrals of random variables $X \in L^1(\Omega, \mathcal{F}, \mathbb{P}_A)$

$$\mathbb{E}_A[X] = \mathbb{E}[X \mid A] = \int X \, d\mathbb{P}_A.$$

Lemma: $L^1(\mathbb{P}) = L^1(\mathbb{P}_A)$, and $\mathbb{E}_A[X] = \dfrac{\mathbb{E}[X : A]}{\mathbb{P}(A)}$.

Pf. If X is simple, $X = \sum_j \alpha_j \mathbb{1}_{B_j}$, then

$$\int X \, d\mathbb{P}_A =$$

If $\{A_n\}_{n=1}^{\infty}$ is a partition of Ω by events with $\mathbb{P}(A_n) > 0$, we can string the different $\mathbb{E}_{A_n}$ functionals together into a single function:

$$\mathbb{E}_{\{A_n\}_{n=1}^{\infty}}[X] =$$

I.e, $\omega \mapsto \begin{cases} \vdots \\ \vdots \end{cases}$

Labeling: instead of indexing this by the partition $\{A_n\}_{n=1}^{\infty}$, it is more convenient (and suggestive) to label it by $\mathcal{G} = \sigma(\{A_n\}_{n=1}^{\infty})$. This makes sense because:

Exercise: If $A_n \in \mathcal{F}$ with $\Omega = \bigsqcup_{n=1}^{\infty} A_n$, then

$$\sigma(\{A_n\}_{n=1}^{\infty}) = \left\{ \bigsqcup_{n \in \Lambda} A_n : \Lambda \subseteq \mathbb{N} \right\}$$

Thus, we can "recover" the partition from $\mathcal{G}$

Cor: If $\mathcal{G} = \sigma(\{A_n\}_{n=1}^{\infty})$ for some partition $\{A_n\}_{n=1}^{\infty}$, then $E_{\mathcal{G}}[X]$ is $\mathcal{G}$-measurable.

Pf. $E_{\mathcal{G}}[X] = \sum_{n=1}^{\infty} E_{A_n}[X]\, \mathbb{1}_{A_n}$ is a linear combination of the indicator functions $\mathbb{1}_{A_n}$ ///

Note: even in this special case where $\mathcal{G}$ is generated by a countable partition, it is not practically possible to recover the partition from $\mathcal{G}$ (unless $\mathcal{G}$, and $\therefore$ the partition, is finite)

(Motivating) Eg. Suppose $Y : \Omega \to S$ is a random variable whose state space S is countable, and $\mathbb{P}(Y=s) > 0$ for each $s \in S$. Then $\{ Y^{-1}(s) : s \in S \}$ forms a partition of Ω.

For $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$,

$$\mathbb{E}[X \mid Y] = \mathbb{E}[X \mid \sigma(Y)] = \sum_{s \in S} \mathbb{E}_{\{Y=s\}}[X] \mathbb{1}_{\{Y=s\}} = \sum_{s \in S} \frac{\mathbb{E}[X : Y=s]}{\mathbb{P}(Y=s)} \mathbb{1}_{\{Y=s\}}$$

$\mathbb{E}[X \mid Y]$ is $\sigma(Y)$-measurable

$\therefore$ by Doob-Dynkin, $\mathbb{E}[X \mid Y] = g(Y)$ for some $g : S \to \mathbb{R}$

There is a more "invariant" way to define conditional expectation, owing to the following.

Prop: Let $\{A_n\}_{n=1}^{\infty}$ be a partition of Ω by positive probability events, and set $\mathcal{G} = \sigma(\{A_n\}_{n=1}^{\infty})$. If $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ and $Y \in L^2(\Omega, \mathcal{G}, \mathbb{P})$, then

$$\mathbb{E}[XY] = \mathbb{E}[\mathbb{E}_{\mathcal{G}}[X]Y].$$

Pf. Step 1: the case that Y is simple. So $Y = \sum_j \beta_j \mathbb{1}_{B_j} \leftarrow B_j \in \mathcal{G}$

$$\therefore \quad \mathbb{E}[XY] = \sum_{k=1}^{\infty} \beta_k' \mathbb{E}[X \mathbb{1}_{A_k}]$$

$$\mathbb{E}_{\mathcal{G}}[X] = \sum_{j=1}^{\infty} \mathbb{E}_{A_j}[X] \mathbb{1}_{A_j}$$

So $\mathbb{E}[\mathbb{E}_{\mathcal{G}}[X]Y] =$

Step 2: If $Y \in L^2(\Omega, \mathcal{G}, \mathbb{P}) \subseteq L^1(\Omega, \mathcal{G}, \mathbb{P})$,

$\exists$ sequence of simple Y_n s.t.

$Y_n \to Y$ a.s. and $|Y_n| \leq |Y|$

$\therefore$ $XY_n \to XY$ a.s. and $|XY_n| \leq |XY| \in L^1$ b/c $XY \in L^2$

$\therefore$ By DCT, $\mathbb{E}[XY_n] \to \mathbb{E}[XY]$

$$\begin{array}{c} \mathbb{E}[XY_n] \\ \parallel \\ \mathbb{E}[\mathbb{E}_{\mathcal{G}}[X] Y_n] \end{array}$$

Note: $\| \mathbb{E}_{\mathcal{G}}[X] \|_{L^2}^2 = \mathbb{E}\left[\left(\sum_{j=1}^{n} \mathbb{E}_{A_j}[X] \mathbb{1}_{A_j}\right)^2\right]$