Conditional expectation toy example

Márton Balázs*

October 3, 2017

We start with Kolmogorov's Theorem on conditional expectations.

Theorem 1 (Thm 9.2 in Williams) *Let X be a random variable on the probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*, with $\mathbb{E}|X| < \infty$. Let $\mathcal{G}$ be a sub σ-algebra. Then there exists a random variable V such that*

(a) V is $\mathcal{G}$-measurable,

(b) $\mathbb{E}|V| < \infty$,

(c) $\mathbb{E}(V\,;\,G) = \mathbb{E}(X\,;\,G)$ *for any $G \in \mathcal{G}$.*

This V is unique up to zero-measure sets and is called a version of the conditional expectation $\mathbb{E}(X \,|\, \mathcal{G})$.

Our toy example will be the following. Let $\Omega = \{1, 2, \dots, 12\}$, $\mathcal{F} = \mathcal{P}(\Omega)$, and $\mathbb{P}$ be the uniform measure on the finite set Ω. Elementary outcomes in Ω will be denoted by ω. Define the random variables

$$Y := \left\lceil \frac{\omega}{4} \right\rceil = \begin{cases} 1, & \text{if } \omega = 1, 2, 3, 4, \\ 2, & \text{if } \omega = 5, 6, 7, 8, \\ 3, & \text{if } \omega = 9, 10, 11, 12, \end{cases} \qquad X := \left\lceil \frac{\omega}{2} \right\rceil = \begin{cases} 1, & \text{if } \omega = 1, 2, \\ 2, & \text{if } \omega = 3, 4, \\ 3, & \text{if } \omega = 5, 6, \\ 4, & \text{if } \omega = 7, 8, \\ 5, & \text{if } \omega = 9, 10, \\ 6, & \text{if } \omega = 11, 12. \end{cases}$$

The σ-algebra generated by Y is

$$\begin{aligned} \mathcal{G} := \sigma(Y) &:= \sigma\big(Y^{-1}\big(\mathcal{B}(\mathbb{R})\big)\big) = \sigma\big(\{1, 2, 3, 4\}, \{5, 6, 7, 8\}, \{9, 10, 11, 12\}\big) \\ &= \big\{\emptyset, \{1, 2, 3, 4\}, \{5, 6, 7, 8\}, \{9, 10, 11, 12\}, \\ &\qquad \{1, 2, 3, 4, 5, 6, 7, 8\}, \{1, 2, 3, 4, 9, 10, 11, 12\}, \{5, 6, 7, 8, 9, 10, 11, 12\}, \Omega\big\}. \end{aligned}$$

Similarly, the σ-algebra generated by X is

$$\mathcal{H} := \sigma(X) := \sigma\big(X^{-1}\big(\mathcal{B}(\mathbb{R})\big)\big) = \sigma\big(\{1, 2\}, \{3, 4\}, \{5, 6\}, \{7, 8\}, \{9, 10\}, \{11, 12\}\big).$$

We see that $\mathcal{G} \subset \mathcal{H} \subset \mathcal{F}$. The σ-algebra $\mathcal{G}$ is coarser (contains less information), while $\mathcal{H}$ is finer (more information). We also see that

- Y is $\mathcal{G}$-measurable (by definition).
- Y is $\mathcal{H}$-measurable (due to $\mathcal{G} \subset \mathcal{H}$).
- X is $\mathcal{H}$-measurable (by definition).
- X is *not* $\mathcal{G}$-measurable (e.g., $X^{-1}\{1\} = \{1, 2\} \notin \mathcal{G}$).

Next, we find the conditional expectation $\mathbb{E}(X \,|\, \mathcal{G})$ based on the definition above. As $\mathcal{G} = \sigma(Y)$, an equivalent notation for this is $\mathbb{E}(X \,|\, \mathcal{G}) = \mathbb{E}(X \,|\, Y)$. Due to $|\Omega| = 12 < \infty$, finite mean of $V = \mathbb{E}(X \,|\, \mathcal{G})$ is not an issue. We look for a $\mathcal{G}$-measurable random variable V with $\mathbb{E}(V\,;\,G) = \mathbb{E}(X\,;\,G)$ for any $G \in \mathcal{G}$.

*University of Bristol

An efficient choice for G is $\{1, 2, 3, 4\}$. As V is $\mathcal{G}$-measurable, and $\mathcal{G}$ has no set that distinguishes between these four outcomes, we find that $V(\omega)$ is the same for $\omega = 1, 2, 3, 4$. The above expectations turn into

$$\begin{aligned}
V(1)\mathbb{P}\{1\} + V(2)\mathbb{P}\{2\} + V(3)\mathbb{P}\{3\} + V(4)\mathbb{P}\{4\} &= X(1)\mathbb{P}\{1\} + X(2)\mathbb{P}\{2\} + X(3)\mathbb{P}\{3\} + X(4)\mathbb{P}\{4\} \\
V(1)\mathbb{P}\{1\} + V(1)\mathbb{P}\{2\} + V(1)\mathbb{P}\{3\} + V(1)\mathbb{P}\{4\} &= X(1)\mathbb{P}\{1\} + X(2)\mathbb{P}\{2\} + X(3)\mathbb{P}\{3\} + X(4)\mathbb{P}\{4\} \\
V(1) = V(2) = V(3) = V(4) &= \frac{1\cdot\frac{1}{12} + 1\cdot\frac{1}{12} + 2\cdot\frac{1}{12} + 2\cdot\frac{1}{12}}{\frac{1}{12}+\frac{1}{12}+\frac{1}{12}+\frac{1}{12}} = 1.5.
\end{aligned}$$

Similarly, with the respective choices $G = \{5, 6, 7, 8\}$ and $G = \{9, 10, 11, 12\}$,

$$\begin{aligned}
V(5) = V(6) = V(7) = V(8) &= \frac{3\cdot\frac{1}{12} + 3\cdot\frac{1}{12} + 4\cdot\frac{1}{12} + 4\cdot\frac{1}{12}}{\frac{1}{12}+\frac{1}{12}+\frac{1}{12}+\frac{1}{12}} = 3.5, \\
V(9) = V(10) = V(11) = V(12) &= \frac{5\cdot\frac{1}{12} + 5\cdot\frac{1}{12} + 6\cdot\frac{1}{12} + 6\cdot\frac{1}{12}}{\frac{1}{12}+\frac{1}{12}+\frac{1}{12}+\frac{1}{12}} = 5.5.
\end{aligned}$$

Hence the conditional expectation is the random variable

$$\mathbb{E}(X \,|\, \mathcal{G})(\omega) = V(\omega) = \begin{cases} 1.5, & \text{if } \omega = 1, 2, 3, 4, \\ 3.5, & \text{if } \omega = 5, 6, 7, 8, \\ 5.5, & \text{if } \omega = 9, 10, 11, 12, \end{cases}$$

being just the average of X over the smallest nontrivial respective units in $\mathcal{G}$.

In a similar way one can check

$$\mathbb{E}(Y \,|\, \mathcal{G})(\omega) = \left\{\begin{matrix} 1, & \text{if } \omega = 1, 2, 3, 4, \\ 2, & \text{if } \omega = 5, 6, 7, 8, \\ 3, & \text{if } \omega = 9, 10, 11, 12 \end{matrix}\right\} = Y(\omega),$$

and indeed it is always the case that $\mathbb{E}(Y \,|\, Y) = Y$ almost everywhere (a.e.).

Further examples are $\mathbb{E}\big(X \,|\, \{\emptyset, \Omega\}\big)$, where the random variable V we are looking for is measurable w.r.t. the trivial σ-algebra $\{\emptyset, \Omega\}$, in other words is a constant. Picking $G = \emptyset$ gives $\mathbb{E}(V\,;\,\emptyset) = 0 = \mathbb{E}(X\,;\,\emptyset)$, which is not very informative. The choice $G = \Omega$ on the other hand fixes the value of the constant V:

$$\begin{aligned}
\mathbb{E}(V\,;\,\Omega) &= \mathbb{E}(X\,;\,\Omega) \\
\mathbb{E}V &= \mathbb{E}X \\
V &= \mathbb{E}X,
\end{aligned}$$

that is, $\mathbb{E}\big(X \,|\, \{\emptyset, \Omega\}\big) = \mathbb{E}X$. This is again true a.e. in general, conditioning on the trivial σ-algebra always produces a full expectation.

If, on the other hand, one conditions on the full σ-algebra $\mathcal{F}$ that has all information that can be available in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, then every event $G \in \mathcal{F}$ can be substituted, and the very detailed ones completely fix the conditional expectation. In our example we can e.g., take $\{7\}$ to obtain

$$\begin{aligned}
\mathbb{E}(V\,;\,\{7\}) &= \mathbb{E}(X\,;\,\{7\}), \\
V(7)\cdot\mathbb{P}\{7\} &= X(7)\cdot\mathbb{P}\{7\}, \\
V(7) &= X(7) = 4.
\end{aligned}$$

Similarly, for any $\omega \in \Omega$ one has $V(\omega) = X(\omega)$, which leads us to $\mathbb{E}(X \,|\, \mathcal{F}) = V = X$. This is again a.e. true for general probability spaces: conditioning on the full information does not do any averaging and gives back the random variable instead.

Our final example is

$$\begin{aligned}
\mathcal{I} :&= \sigma\big(\{1, 5, 9\}, \{3, 7, 11\}\big) \\
&= \big\{\emptyset, \{1, 5, 9\}, \{3, 7, 11\}, \{2, 4, 6, 8, 10, 12\}, \{1, 3, 5, 7, 9, 11\}, \\
&\qquad \{1, 2, 4, 5, 6, 8, 9, 10, 12\}, \{2, 3, 4, 6, 7, 8, 10, 11, 12\}, \Omega\big\}.
\end{aligned}$$

We compute $V = \mathbb{E}(Y \,|\, \mathcal{I})$ as before. This is $\mathcal{I}$-measurable, hence constant on $\{1,\,5,\,9\}$, as well as on $\{3,\,7,\,11\}$ and on $\{2,\,4,\,6,\,8,\,10,\,12\}$. Substituting these as G (the rest in $\mathcal{I}$ will not provide additional help) in $\mathbb{E}(V\,;\,G) = \mathbb{E}(Y\,;\,G)$ results in

$$V(1) = V(5) = V(9) = \frac{Y(1)\mathbb{P}\{1\} + Y(5)\mathbb{P}\{5\} + Y(9)\mathbb{P}\{9\}}{\mathbb{P}\{1\} + \mathbb{P}\{5\} + \mathbb{P}\{9\}} = \frac{1\cdot\frac{1}{12} + 2\cdot\frac{1}{12} + 3\cdot\frac{1}{12}}{\frac{1}{12}+\frac{1}{12}+\frac{1}{12}} = 2,$$

$$V(3) = V(7) = V(11) = \frac{Y(3)\mathbb{P}\{3\} + Y(7)\mathbb{P}\{7\} + Y(11)\mathbb{P}\{11\}}{\mathbb{P}\{3\} + \mathbb{P}\{7\} + \mathbb{P}\{11\}} = \frac{1\cdot\frac{1}{12} + 2\cdot\frac{1}{12} + 3\cdot\frac{1}{12}}{\frac{1}{12}+\frac{1}{12}+\frac{1}{12}} = 2,$$

$$\begin{aligned} V(2) &= V(4) = V(6) \\ &= V(8) = V(10) = V(12) = \tfrac{Y(2)\mathbb{P}\{2\}+Y(4)\mathbb{P}\{4\}+Y(6)\mathbb{P}\{6\}+Y(8)\mathbb{P}\{8\}+Y(10)\mathbb{P}\{10\}+Y(12)\mathbb{P}\{12\}}{\mathbb{P}\{2\}+\mathbb{P}\{4\}+\mathbb{P}\{6\}+\mathbb{P}\{8\}+\mathbb{P}\{10\}+\mathbb{P}\{12\}} \\ &= \frac{1\cdot\frac{1}{12} + 1\cdot\frac{1}{12} + 2\cdot\frac{1}{12} + 2\cdot\frac{1}{12} + 3\cdot\frac{1}{12} + 3\cdot\frac{1}{12}}{\frac{1}{12}+\frac{1}{12}+\frac{1}{12}+\frac{1}{12}+\frac{1}{12}+\frac{1}{12}} = 2. \end{aligned}$$

We find that $\mathbb{E}(Y \,|\, \mathcal{I})$ is actually a constant, and in fact $= \mathbb{E}Y$.

We can repeat this calculation with any function $f : \mathbb{R} \to \mathbb{R}$ (in general this is chosen to be bounded and measurable) to find $\mathbb{E}\big(f(Y) \,|\, \mathcal{I}\big) = \mathbb{E}\big(f(Y)\big)$, a constant. This is when we say that the random variable Y is independent of the σ-algebra $\mathcal{I}$. Knowing which of the events $\{1,\,5,\,9\}$ and $\{3,\,7,\,11\}$ did or did not happen will not tell us any information about Y.

If $\mathcal{I}$ happens to be generated by yet another random variable Z, $\mathcal{I} = \sigma(Z)$, then the above is equivalent to variables Y and Z being independent.