Note on Conditional Distributions

Bernd Pape
Department of Mathematics and Statistics
University of Vaasa
e-mail: bepa@uwasa.fi
url: www.uwasa.fi/∼bepa/

Consider coin-tossing with a fair coin, so head and tail are up with equal probability 1/2. Suppose you win 1$ each time head is up, and loose 1$ each time tail is up, and you toss the coin twice.

Collecting the possible outcomes of the experiment, we may then write for the sample space $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, where $\omega_1 = \{1, 1\}$, $\omega_2 = \{1, -1\}$, $\omega_3 = \{-1, 1\}$, and $\omega_4 = \{-1, -1\}$.

The collection of sets $\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}\}$ is called partition of the sample space, because $\{\omega_1\} \cup \{\omega_2\} \cup \{\omega_3\} \cup \{\omega_4\} = \Omega$ and $\{\omega_i\} \cap \{\omega_j\} = \emptyset$ for all $i \neq j$.

Let us now introduce the random variables X_1 and X_2 to denote the number of dollars earned in the first, respectively second, toss. The probability distributions of these two random variables are:

x_1	1	-1
$P(X_1 = x_1)$	1/2	1/2

,

x_2	1	-1
$P(X_2 = x_2)$	1/2	1/2

,

such that we obtain for the expected value of both random variables

$$\mathsf{E}(X_1) = \mathsf{E}(X_2) = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot (-1) = 0.$$

Let us now consider a new random variable $X := X_1 + X_2$ which represents the total amount of dollars made after the game. Obviously, X may take the values $1 + 1 = 2$, $1 - 1 = (-1) + 1 = 0$, and $-1 - 1 = -2$. Our goal is to find the (unconditional) probability distribution of X. As an intermediate step, we may write down the probability distributions of X conditional upon that the first toss was head (tail) up:

x	2	0
$P(X = x \mid X_1 = 1)$	1/2	1/2

,

x	0	-2
$P(X = x \mid X_1 = -1)$	1/2	1/2

,

which implies for the expected value of X conditional on the first toin coss

$$\mathsf{E}(X \mid X_1 = 1) = \frac{1}{2} \cdot 2 + \frac{1}{2} \cdot 0 = 1$$

and

$$\mathsf{E}(X|X_1 = -1) = \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot (-2) = -1.$$

We may summarize the two probability distribtutions of the total gain for head (tail) up in the first toss as the conditional distribution of X given X_1 (for short $X|X_1$):

$x\|X_1$	$X_1 + 1$	$X_1 - 1$
$P(X = x\|X_1)$	$1/2$	$1/2$

The conditional expectation of X given X_1 is then similar to the unconditional expectation:

$$\mathsf{E}(X|X_1) = \frac{1}{2}(X_1 + 1) + \frac{1}{2}(X_1 - 1) = X_1.$$

Note that the conditional expectation, unlike the unconditional expectation, is a random variable instead of just a number.

We may now proceed in our original project to find the unconditional probability distribution of X. Note for that purpose, that the events $\{X=2\}$, $\{X=0\}$ and $\{X=-2\}$ may be partitioned as follows:

$$\{X=2\}=\{1,1\}=\omega_1, \quad \{X=-2\}=\{-1,-1\}=\omega_4,$$

$$\text{and } \{X=0\}=\{1,-1\}\cup\{-1,1\}=\omega_2\cup\omega_3.$$

Now the law of total probability states that we may calculate the unconditional probability of an event ω as a weigthed sum of its conditional probabilities over any of its partions $\{\omega_1,\ldots,\omega_n\}$ as*

$$P(\omega)=\sum_{i=1}^{n}P(\omega|\omega_i)P(\omega_i).$$

This implies in our case:

$$\begin{aligned}
P(X=2)=&P(X=2|X_1=1)\cdot P(X_1=1)=\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4},\\
P(X=-2)=&P(X=-2|X_1=-1)\cdot P(X_1=-1)=\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4},\\
P(X=0)=&P(X=0|X_1=1)\cdot P(X_1=1)\\
&+P(X=0|X_1=-1)\cdot P(X_1=-1)\\
=&\frac{1}{2}\cdot\frac{1}{2}+\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{2}.
\end{aligned}$$

*for continious random variables the law of total probability reads $f(x)=\int_{-\infty}^{\infty}f(x|y)\,dy$.

The unconditional expectation of X becomes

$$\mathsf{E}(X) = \frac{1}{4} \cdot 2 + \frac{1}{2} \cdot 0 + \frac{1}{4} \cdot (-2) = 0.$$

Note that this coincides with the unconditional expectation of $\mathsf{E}(X|X_1)$:

$$\mathsf{E}[\mathsf{E}(X|X_1)] = \mathsf{E}(X_1) = 0 = \mathsf{E}(X).$$

Consider now n coin tosses with associated random variables $X_1, \ldots, X_n$ and the total gain again associated with X. It is then convenient to collect the outcomes of the first i tosses in the information set $\mathcal{F}_i = \{X_1, \ldots, X_i\}$. It turns out that also in this general case

$$\mathsf{E}[\mathsf{E}(X|\mathcal{F}_i)] = \mathsf{E}(X) \quad \text{for all } i = 1, \ldots, n,$$

known as the law of iterated expectations.