Conditional distributions. The continuous case.

Suppose Y is a continous random vector.

For any event E we let

$$F_{E,Y}(y) = P(E \cap \{Y \le y\}),\ y \in \mathbf{R}.$$

We let

$$f_{E,Y} = F'_{E,Y}$$

with the understanding that the domain of $f_{E,Y}$ is the set of $y \in \mathbf{R}$ such that $F'_{E,Y}$ exists at y. It follows from real variable theory that $f_{E,Y}$ is locally integrable. Since

$$\sum_{i=1}^{m} F_{E,Y}(y_i) - F_{E,Y}(y_{i-1}) == P(E \cap \{y_0 < Y \le y_m\}) \le P(y_0 < Y \le y_m) = \int_{y_0}^{y_m} f_Y(y)\,dy$$

whenever $-\infty < y_0 \le y_1 \le \cdots \le y_m < \infty$ it follows from real variable theory that

$$F_{E,Y}(y) = \int_{-\infty}^{y} f_{E,Y}(w)\,dw,\ y \in \mathbf{R}. \tag{1}$$

Note that if $y \in \mathbf{R}$ then $P(E|Y = y)$ is undefined because $P(Y = y) = 0$ since Y is continous. Let us *define*

$$P(E|Y = y) = \begin{cases} \lim_{h\downarrow 0} P(E|y < Y \le y + h) & \text{if } f_Y(y) > 0 \text{ and this limit exists,} \\ 0 & \text{if } f_Y(y) = 0. \end{cases}$$

Suppose $y \in \mathbf{R}$ and $f_Y(y) > 0$ and $f_{E,Y}(y) = F'_{E,Y}(y)$ exists. Then for $h > 0$ we have

$$\begin{aligned} P(E|y < Y \le y + h) &= \frac{P(E \cap \{y < Y \le y + h\})}{P(y < Y \le y + h)} \\ &= \frac{F_{E,Y}(y + h) - F_{E,Y}(y)}{h} \frac{h}{P(Y \le y + h) - P(Y \le y)} \\ &\to \frac{f_{E,Y}(y)}{f_Y(y)}. \end{aligned}$$

Thus,

$$P(E|Y = y) = \begin{cases} \frac{f_{E,Y}(y)}{f_Y(y)} & \text{if } f_Y(y) > 0 \text{ and } f_{E,Y}(y) \text{ is defined,} \\ 0 & \text{if } f_Y(y) = 0. \end{cases} \tag{2}$$

It follows from (1) and (2) that

$$P(E \cap \{Y \le y\}) = \int_{\infty}^{y} P(E|Y = y) f_Y(w)\,dw \quad \text{whenever } y \in \mathbf{R}. \tag{3}$$

We shall not make this precise, but $P(\cdot|Y = y)$ behaves like a probability for essentially all $y \in \mathbf{R}$.

Suppose X is a random variable such that (X, Y) is a continuous random vector.
For $(x, y) \in \mathbf{R}^2$ we let

$$F_{X|Y}(x|y) = P(\{X \le x\}|Y = y)$$

and we let

$$f_{X|Y}(x|y) = \frac{d}{dx} F_{X|Y}(x|y).$$

We have

$$f_{\{X\le x\},Y}(y) = \lim_{h\downarrow 0}\frac{P(X\le x, y<Y\le y+h)}{h}$$
$$= \lim_{h\downarrow 0}\frac{1}{h}\int_y^{y+h}\left(\int_{-\infty}^x f_{X,Y}(v,w)\,dv\right)dw$$
$$= \int_{-\infty}^x f_{X,Y}(v,y)\,dv$$

provided

(4) $$\mathbf{R}\ni y'\mapsto \int_{-\infty}^{y'}\left(\int_{-\infty}^x f_{X,Y}(v,w)\,dv\right)dw \quad \text{is continuous at } y.$$

Thus, by (2),

(5) $$F_{X|Y}(x|y) = \begin{cases} \frac{\int_{-\infty}^x f_{X,Y}(v,y)\,dv}{f_Y(y)} & \text{if } f_Y(y)>0 \text{ and (4) holds,}\\ 0 & \text{if } f_Y(y)=0\end{cases}$$

If, in addition to (4), $\mathbf{R}\ni v\mapsto f_{X,Y}(v,y)$ is continuous at x we find that

(6) $$f_{X|Y}(x|y) = \frac{d}{dx}F_{X|Y}(x|y) = \begin{cases}\frac{f_{X,Y}(x,y)}{f_Y(y)} & \text{if } f_Y(y)>0,\\ 0 & \text{if } f_Y(y)=0.\end{cases}$$

It follows that

$$F_{X|Y}(x|y) = \int_{-\infty}^x f_{X|Y}(w|y)\,dw \quad \text{whenever (4) holds.}$$

It seems natural to *define*

$$E(X|Y=y) = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\,dx \quad \text{whenever (4) holds;}$$

had we made precise our previously made assertion that $E(\cdot|Y=y)$ behave like a probability this would be a theorem and not a definition. It is true that

$$E(X|Y=y) = \lim_{h\downarrow 0} E(X|y<Y\le y+h);$$

we leave the proof as an exercise for the reader.

We let

$$E(X|Y),$$

the **conditional expectation of X given Y**, be the random variable such that

$$E(X|Y) = \begin{cases} E(X|Y=y) & \text{if } Y=y \text{ and } f_Y(y)>0,\\ 0 & \text{else.}\end{cases}$$

We have the fundamental formula

(7) $$E(E(X|Y)) = E(X)$$

which follows because

$$E(E(X|Y)) = \int_{-\infty}^{\infty} E(X|Y=y) f_Y(y)\,dy$$
$$= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} x f_{X|Y}(x|y)\,dx\right) f_Y(y)\,dy$$
$$= \int\int_{\mathbf{R}^2} x f_{X,Y}(x,y)\,dxdy$$
$$= E(X).$$