.

1. Conditional expectation.

See page 239 in the book.

1.1. **When Y is discrete.**

Suppose X and Y are random variables on the same probability space and Y is discrete.

Definition 1.1. We let

$$E(X|Y)=\sum_{y,\ P(Y=y)>0} E(X|Y=y)1_{\{Y=y\}}.$$

That is, $E(X|Y)$ is the random variable such that, for each $y\in\mathbb{R}$,

$$E(X|Y)=E(X|Y=y)\quad \text{on } \{Y=y\} \text{ if } P(Y=y)>0,$$

and

$$E(X|Y)=0\quad \text{on } \{Y=y\} \text{ if } P(Y=y)=0.$$

One calls $E(X|Y)$ the **conditional expectation of X given Y**. For any event F we set

$$E(F|Y)=E(1_F|Y).$$

Theorem 1.1. We have

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

Proof. We have

$$E(E(X|Y))=\sum_{y,\ P(Y=y)>0} E(X|Y=y)P(Y=y)=E(X)$$

□

Definition 1.2. Suppose X is discrete. We let

$$p_{X|Y}(x|y)=P(X=x|Y=y)\quad \text{whenver } x,y\in\mathbb{R} \text{ and } p_Y(y)=P(Y=y)>0.$$

Proposition 1.1. Suppose X is discrete. Then

$$p_{X|Y}(x|y)=\frac{p_{X,Y}(x,y)}{p_Y(y)}\quad \text{whenever } x,y\in\mathbb{R} \text{ and } p_Y(y)=P(Y=y)>0.$$

Moreover, $E(X|Y)$ is discrete and

$$E(X|Y=y)=\sum_x x p_{X|Y}(x|y).$$

Proof. This follows directly from the definitions. □

1.2. **When Y is continuous.** Suppose X and Y are random variables on the same probability space and Y is continuous.

Definition 1.3. For each $y \in \mathbb{R}$ we let

$$E(X|Y = y) \in \mathbb{R}$$

equal

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

provided this limit exists and

$$\lim_{h \downarrow} \frac{P(y - y < Y < y + h)}{2h} \quad \text{exists and is positive}$$

and we let it be zero otherwise. For an event F we let

$$P(F|Y = y) = E(1_F|Y = y).$$

We let

$$E(X|Y)$$

be the random variable such that, for each $y \in \mathbb{R}$,

$$E(X|Y) = E(X|Y = y) \quad \text{on } \{Y = y\}.$$

For an event F we let

$$P(F|Y) = E(1_F|Y).$$

Theorem 1.2. We have

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

Proof. Approximate X by a discrete random variable in much the same way we did when we defined the expectation of a continuous random variable. □

Definition 1.4. Suppose (X, Y) is continuous. we let

$$f_{X|Y}(x|y) = \lim_{h \downarrow 0} \frac{P(x - h < X < x + h|Y = y)}{2h}$$

whenever for $(x, y) \in \mathbb{R}^2$ for which this limit exists.

Proposition 1.2. Suppose (X, Y) is continuous . Then

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$$

whenever $(x, y) \in \mathbb{R}^2$, $f_Y(y) > 0$ and $f_{X,Y}$ is continuous at (x, y).

Proof. For any $(x, y) \in \mathbb{R}^2$

$$\begin{aligned}
&\frac{P(x - h < X < x + h|y - k < Y < y + k)}{2h} \\
&\qquad = \frac{\frac{P(x-h<X<x+h, y-k<Y<y+k)}{4hk}}{\frac{P(y-k<Y<y+k)}{2k}} \\
&\qquad \to \frac{f_{X,Y}(x, y)}{f_Y(y)}
\end{aligned}$$

as $(h, k) \to (0, 0)$. □

Proposition 1.3. Suppose (X, Y) is continuous. Then

$$E(X|Y = y) = \int_{-\infty}^{\infty} f_{X|Y}(x|y)dx$$

whenever $f_Y(y) > 0$.

Moreover $E(X|Y)$ is continuous and has pdf

$$y \ni \mathbb{R} \mapsto \begin{cases} \int_{-\infty}^{\infty} f_{X|Y}(x|y)\, dx & \text{if } f_Y(y) > 0, \\ 0 & \text{else.} \end{cases}$$

Finally,

$$E(X) = \int_{\{f_Y > 0\}} y \left(\int_{-\infty}^{\infty} f_{X|Y}(x|y)\, dx \right).$$

Proof. Turn the crank. □