7 Bedingte Erwartungswerte und Bedingte Verteilungen

Sei $(\Omega, \mathcal{A}, P)$ ein W'Raum, $(\Omega', \mathcal{A}')$ ein Messraum, $Y : \Omega \to \Omega'$ sei $(\mathcal{A}, \mathcal{A}')$-messbar und nehme die Werte $y_1, \dots, y_n \in \Omega'$ an. $Y^{-1}(y_k) = \{\omega \in \Omega \mid Y(\omega) = y_k\} =: A_k \Rightarrow \Omega = A_1 + \dots + A_n$ und $\sigma(Y) = \{\sum_{k \in I} A_k \mid I \subset \{1, \dots, n\}\}$.

Definition *Sei $X : \Omega \to \mathbb{R}$ eine ZV mit $E|X| < \infty$. Dann ist der bedingte Erwartungswert von X unter der Bedingung $Y = y_k$ defniniert durch:*

$$E[X|Y = y_k] := \frac{1}{P(A_k)} \int_{A_k} X dP, \quad k = 1, \dots, n$$

Falls X diskret mit $x_1, \dots, x_m$:

$$\begin{aligned} E[X|Y = y_k] &= \frac{1}{P(Y = y_k)} \sum_{j=1}^{m} x_j \cdot P(X = x_j, Y = y_k) \\ &= \sum_{j=1}^{m} x_j \cdot P(X = x_j | Y = y_k) \end{aligned}$$

Definition *Der **bedingte Erwartungswert von** X **gegeben** Y ist $E[X|Y] : \Omega \to \mathbb{R}$ mit*

$$E[X|Y](\omega) := \sum_{k=1}^{n} E[X|Y = y_k] \cdot \mathbf{1}_{[Y = y_k]}(\omega)$$

Bemerkung a) Offenbar ist $E[X|Y]$ $(\sigma(Y), \mathfrak{B})$-messbar.

b) Sei $Z := E[X|Y]$. Dann gilt

$$\begin{aligned} \int_{A_k} Z \mathrm{d}P &= \int_{\Omega} \mathbf{1}_{A_k} Z \mathrm{d}P \\ &= E[X, Y = y_k] \cdot P(A_k) \\ &= \int_{A_k} X \mathrm{d}P \end{aligned}$$

Wegen der Struktur von $\sigma(Y)$ folgt auch

$$\int_A Z \mathrm{d}P = \int_A X \mathrm{d}P \quad \forall A \in \sigma(Y)$$

c) $E[X|Y] = g(Y)$ mit

$$g(y) = \sum_{k=1}^{n} E[X|Y = y_k] \cdot \mathbf{1}_{\{y_k\}}(y)$$

d) Offenbar hängt die Definition von $E[X|Y]$ nur davom ab, auf welchen Mengen A_k Y die verschiedenen Werte annimmt, nicht aber welche Werte das genau sind.
Deshalb schreibt man auch:

$$E[X|Y] = E[X|\sigma(Y)]$$

Beispiel 7.1 Sei $([0,1), \mathfrak{B}_{[0,1)}, \underbrace{\lambda_{[0,1)}}_{=:P}), X(\omega) = \omega$

- Hier fehlt ein Bild -

$$A_k = \left[\frac{k-1}{n}, \frac{k}{n}\right), k = 1, \ldots, n, \quad \mathfrak{F} := \left\{\sum_{k \in I} A_k | I \subset \{1, \ldots, n\}\right\}$$

$$\begin{aligned} E[X, A_k] &= \frac{1}{P(A_k)} \int_{A_k} \omega P(\mathrm{d}\omega) \\ &= n \int_{\frac{k-1}{n}}^{\frac{k}{n}} \omega \mathrm{d}\omega \\ &= \frac{1}{2} \frac{2k-1}{n} \end{aligned}$$

$E[X, \mathfrak{F}]$ ist also eine „Approximation" oder „Vergröberung" von X. Bezüglich einer beliebigen Sub-σ-Algebra $\mathfrak{F} \subset \mathcal{A}$ wird der bedingte Erwartungswert wie folgt definiert:

Definition *Sei X eine Zufallsvariable mit $E|X| < \infty$ und $\mathfrak{F} \subset \mathcal{A}$ eine Sub-σ-Algebra von $\mathcal{A}$. Dann heißt $Z : \Omega \to \mathbb{R}$* ***eine Version des bedingten Erwartungswertes*** *$E[X|\mathfrak{F}]$* ***von*** *X unter $\mathfrak{F}$, wenn gilt*

(i) *Z ist $\mathfrak{F}$-messbar*

(ii) $\int_A Z\,dP = \int_A X\,dP \quad \forall A \in \mathfrak{F}$

Satz 7.1
Der bedingte Erwartungswert existiert und ist bis auf Nullmengen eindeutig.

Beweis Sei $X \geq 0$. Durch

$$Q(A) := \int_A X(\omega) P(\mathrm{d}\omega) \quad \forall A \in \mathfrak{F}$$

wird ein Maß auf $(\Omega, \mathfrak{F})$ definiert (Satz 2.7).
Sei $P_{\mathfrak{F}}$ die Einschränkung von P auf $\mathfrak{F}$. Offenbar $Q \ll P_{\mathfrak{F}}$. Satz von Radon-Nikodym $\Longrightarrow$ Q besitzt eine Dichte Z bzgl. $P_{\mathfrak{F}}$ und Z ist nach Definition $\mathfrak{F}$-messbar.
Falls X beliebig: $X = X^+ - X^-$
P-f.s. Eindeutigkeit: Seien $Z, \tilde{Z}$ Versionen von $E[X, \mathfrak{F}]$.

$$\Longrightarrow \int_A (Z - \tilde{Z}) \mathrm{d}P = 0 \quad \forall A \in \mathfrak{F}$$

Wegen $\{Z > \tilde{Z}\} \in \mathfrak{F}, \{Z < \tilde{Z}\} \in \mathfrak{F}$ folgt:

$$E|Z - \tilde{Z}| = \int_{\{Z>\tilde{Z}\}} (Z - \tilde{Z}) \mathrm{d}P - \int_{\{Z<\tilde{Z}\}} (Z - \tilde{Z}) \mathrm{d}P = 0$$

$\implies Z = \tilde{Z}$ P-f.s. ■

Bemerkung Der bedingte Erwartungswert ist also eigentlich die Äquivalenzklasse

$$E[X|\mathfrak{F}] = \left\{ Z \in L^1(\Omega, \mathfrak{F}, P) | \int_A Z \mathrm{d}P = \int_A X \mathrm{d}P \ \forall A \in \mathfrak{F} \right\}$$

Ein Element davon nennt man „Version". Oft wird $E[X|\mathfrak{F}]$ mit einer Version identifiziert.

Definition *Sei $A \in \mathfrak{F}$. Eine Version von $E[\mathbf{1}_A|\mathfrak{F}]$ bezeichnet man als **Version der bedingten Wahrscheinlichkeit** $P(A|\mathfrak{F})$.*

Bemerkung Es gilt für $B \in \mathfrak{F}$:

$$\int_B P(A|\mathfrak{F}) \mathrm{d}P \overset{(ii)}{=} \int_B \mathbf{1}_A \mathrm{d}P = P(A \cap B)$$

Satz 7.2
Sei $X \in L^2(\Omega, \mathcal{A}, P)$ mit $||X||^2 = EX^2$. Dann gilt:

$$||X - E[X|\mathfrak{F}]||^2 = \inf \left\{ ||X - Y||^2 | Y \in L^2(\Omega, \mathfrak{F}, P) \right\}$$

Beweis siehe Henze Stochastik II, S.214 ■

Satz 7.3 (Rechenregeln für bedingte Erwartungswerte)
Es seien $X, Y \in L^1(\Omega, \mathcal{A}, P), \mathfrak{F}, \mathfrak{F}_1, \mathfrak{F}_2$ Sub-σ-Algebren von $\mathcal{A}$. Dann gilt:

a) $E[aX + bY|\mathfrak{F}] = aE[X|\mathfrak{F}] + bE[Y|\mathfrak{F}]$ *P-f.s.* $a, b \in \mathbb{R}$

b) $E[E[X|Y]] = EX$

c) $X \leq Y \implies E[X|\mathfrak{F}] \leq E[Y|\mathfrak{F}]$ *P-f.s.*

d) *Für* $\mathfrak{F}_1 \subset \mathfrak{F}_2$ *gilt* $E[E[X|\mathfrak{F}_2]|\mathfrak{F}_1] = E[X|\mathfrak{F}_1]$
Für $\mathfrak{F}_1 \supset \mathfrak{F}_2$ *gilt* $E[E[X|\mathfrak{F}_2]|\mathfrak{F}_1] = E[X|\mathfrak{F}_2]$

e) *Falls Y $\mathfrak{F}$-messbar und $EXY < \infty$ gilt:*

$$E[XY|\mathfrak{F}] = YE[X|\mathfrak{F}]$$

f) *Falls X von $\mathfrak{F}$ unabhängig ist (d.h. falls die X und $\mathbf{1}_A$ $\forall A \in \mathfrak{F}$ unabhängig sind), dann gilt:*

$$E[X|\mathfrak{F}] = EX$$

Bemerkung Aus Satz 7.3 bekommt man:

1. $X \equiv c \in \mathbb{R} \overset{\text{f)}}{\Rightarrow} E[c|\mathfrak{F}] = c$

2. $\mathfrak{F} = \{\emptyset, \Omega\} \overset{\text{f)}}{\Rightarrow} E[X|\mathfrak{F}] = EX$

3. X $\mathfrak{F}$-messbar $\overset{\text{e)}}{\Rightarrow} E[X|\mathfrak{F}] = X$

4. $X \geq 0 \overset{\text{c)}}{\Rightarrow} E[X|\mathfrak{F}] \geq 0$ P-f.s.

Beweis von Satz 7.3:

a)

$$\begin{aligned}
\int_A E[aX + bY|\mathfrak{F}]\mathrm{d}P &= \int_A aX + bY \mathrm{d}P \\
&\overset{\text{Linearität}}{=} a\int_A X\mathrm{d}P + b\int_A Y\mathrm{d}P \\
&= a\int_A E[X|\mathfrak{F}]\mathrm{d}P + b\int_A E[Y|\mathfrak{F}]\mathrm{d}P \\
&= \int_A (aE[X|\mathfrak{F}] + bE[Y|\mathfrak{F}])\,\mathrm{d}P \quad \forall A \in \mathfrak{F}
\end{aligned}$$

$\Longrightarrow$ Behauptung, da $aE[X|\mathfrak{F}] + bE[Y|\mathfrak{F}]$ $\mathfrak{F}$-messbar und Radon-Nikodym-Dichte P-f.s. eindeutig.

b)

$$E[E[X|\mathfrak{F}]] = \int_\Omega E[X|\mathfrak{F}]\mathrm{d}P = \int_\Omega X\mathrm{d}P = EX$$

c)

$$\begin{aligned}
A &:= \{\omega \in \Omega\,|\, E[X|\mathfrak{F}](\omega) > E[Y|\mathfrak{F}](\omega)\} \in \mathfrak{F} \\
&= \bigcup_{n\in\mathbb{N}} \underbrace{\left\{\omega \in \Omega \,\middle|\, E[X|\mathfrak{F}](\omega) > E[Y|\mathfrak{F}](\omega) + \frac{1}{n}\right\}}_{A_n}
\end{aligned}$$

Annahme: $P(A) > 0 \Longrightarrow \exists n \in \mathbb{N}$ mit $P(A_n) > 0$

$$\begin{aligned}
\Longrightarrow 0 &\leq \int_{A_n} (Y - X)\mathrm{d}P \\
&= \int_{A_n} E[Y|\mathfrak{F}]\mathrm{d}P - \int_{A_n} E[X|\mathfrak{F}]\mathrm{d}P \\
&= \int_{A_n} (E[Y|\mathfrak{F}] - E[X|\mathfrak{F}])\,\mathrm{d}P \\
&\leq -\frac{1}{n} \cdot P(A_n) \\
&< 0 \quad \text{Widerspruch!}
\end{aligned}$$

d) Z.z. Für $\mathfrak{F}_1 \subset \mathfrak{F}_2$ gilt: $E[E[X|\mathfrak{F}_2]|\mathfrak{F}_1] = E[X|\mathfrak{F}_1]$. Sei $A \in \mathfrak{F}_1 \implies A \in \mathfrak{F}_2$ und

$$\int_A E[X|\mathfrak{F}_1]\mathrm{d}P = \int_A X\mathrm{d}P = \int_A E[X|\mathfrak{F}_2]\mathrm{d}P = \int_A E[E[X|\mathfrak{F}_2]|\mathfrak{F}_1]\mathrm{d}P$$

$\implies$ Behauptung, da Radon-Nikodym-Dichte eindeutig.
Für $\mathfrak{F}_1 \supset \mathfrak{F}_2$ ähnlich.

e) Mit algebraischer Induktion:

- Sei $Y = \mathbf{1}_B, B \in \mathfrak{F}$ und $A \in \mathfrak{F}$ beliebig.

$$\int_A Y \cdot E[X|\mathfrak{F}]\mathrm{d}P = \int_{A\cap B} E[X|\mathfrak{F}]\mathrm{d}P = \int_{A\cap B} X\mathrm{d}P = \int_A YX\mathrm{d}P$$

  Außerdem ist $Y \cdot E[X|\mathfrak{F}]$ $\mathfrak{F}$-messbar $\implies$ Behauptung, da Radon-Nikodym-Dichte P-f.s. eindeutig.

- Linearität des Integrals + Teil a) $\implies$ Aussage für $Y \in \mathcal{E}.Y \geq 0$: Bedingte Version des Satzes von der monotonen Konvergenz ($\to$ Übung).

- Dann $Y = Y^+ - Y^-$

f)

$$\begin{aligned}
\int_A E[X|\mathfrak{F}]\mathrm{d}P &= \int_A X\mathrm{d}P \\
&= \int_\Omega \mathbf{1}_A X\mathrm{d}P \\
&\overset{\text{unabh.}}{=} \int \mathbf{1}_A \mathrm{d}P \cdot \underbrace{\int X\mathrm{d}P}_{=EX} \\
&= \int_A EX\mathrm{d}P
\end{aligned}$$

$\implies$ Behauptung, da EX $\mathfrak{F}$-messbar. ∎

Satz 7.4 (Faktorisierungssatz)
Es seien $(\Omega, \mathcal{A}), (\Omega', \mathcal{A}')$ Messräume und $Y : \Omega \to \Omega'$ ein Zufallsgröße. Ist $X : \Omega \to \mathbb{R}$ eine $(\sigma(Y), \mathfrak{B})$-messbare Zufallsvariable. Dann gibt es eine $\mathfrak{B}$-messbare Funktion $g : \Omega' \to \mathbb{R}$ mit

$$X = g \circ Y.$$

Beweis Algebraische Induktion:

(i) Sei $X = \sum_{j=1}^n a_j \mathbf{1}_{A_j} \in \mathcal{E}$ mit $a_j \geq 0, A_j \in \sigma(Y)$.
$\implies$ $A_j = Y^{-1}(A'_j), A'_j \in \mathcal{A}'$. Wähle $g = \sum_{j=1}^n a_j \mathbf{1}_{A'_j}$
$\implies$ $X = g \circ Y$
$\implies$ Behauptung

(ii) Sei $X \geq 0$ und $(\sigma(Y), \mathfrak{B})$-messbar. $\implies \exists (X_n) \subset \mathcal{E}, 0 \leq X_n \uparrow X$ und wegen (i) $\exists (\mathcal{A}', \mathfrak{B})$-messbare Funktion g_n mit $X_n = g_n \circ Y, n \in \mathbb{N}$.

$$\implies X = \sup_{n\in\mathbb{N}} X_n = \sup_{n\in\mathbb{N}} (g_n \circ Y) = (\sup_{n\in\mathbb{N}} g_n) \circ Y$$

Wähle also $g = \sup_{n\in\mathbb{N}} g_n$

(iii) $X = X^+ - X^- \overset{\text{(ii)}}{\implies} X = g_1 \circ Y - g_2 \circ Y$. Wähle $g = g_1 - g_2$. ■

Bemerkung Statt $E[X|\sigma(Y)]$ schreiben wir auch $E[X|Y]$ und wegen Satz 7.4 $\exists g : \Omega' \to \mathbb{R}$ $(\mathcal{A}', \mathfrak{B})$-messbar mit $E[X|Y] = g \circ Y$ P-f.s.. Die Funktion g ist P^Y-f.s. eindeutig.

Definition *Ist $E[X|Y] = g \circ Y$ wie oben, so heißt $E[X|Y = y] = g(y)$ (ein)* ***bedingter Erwartungswert von X unter der Bedingung $Y = y$.***

Satz 7.5
Für alle $A' \in \mathcal{A}'$ gilt:

$$\int_{A'} E[X|Y = y] P^Y(dy) = \int_{Y^{-1}(A')} X \, dP$$

Beweis

$$\int_{A'} E[X|Y = y] P^Y(\mathrm{d}y) = \int_{A'} g \mathrm{d}P^Y \overset{\text{Sa. 2.4}}{=} \int_{Y^{-1}(A')} g \circ Y \mathrm{d}P = \int_{Y^{-1}(A')} X \mathrm{d}P.$$

Bemerkung Für $A \in \mathcal{A}$ heißt $P(A|Y = y) := E[\mathbf{1}_A|Y = y]$ (eine) **bedingte Wahrscheinlichkeit von A unter der Bedingung** $Y = y$. Bedingte Wahrscheinlichkeiten treten oft bei gekoppelten Zufallsexperimenten auf. Die folgende Sichtweise ist konstruktiver:

Definition *Es seien $(\Omega_1, \mathcal{A}_1), (\Omega_2, \mathcal{A}_2)$ messbare Räume. Eine Abbildung $Q : \Omega_1 \times \mathcal{A}_2 \to [0, 1]$ mit*

(i) $\omega_1 \mapsto Q(\omega_1, A_2)$ ist $\mathcal{A}_1$-messbar $\forall A_2 \in \mathcal{A}$.

(ii) $A_2 \mapsto Q(\omega_1, A_2)$ ist ein Wahrscheinlichkeitsmaß auf $(\Omega_2, \mathcal{A}_2)$ $\forall \omega_1 \in \Omega_1$

nennt man ***Übergangskern*** *oder* ***Kern*** *von $(\Omega_1, \mathcal{A}_1)$ nach $(\Omega_2, \mathcal{A}_2)$.*

Satz 7.6
Es seien $(\Omega_1, \mathcal{A}_1, P_1)$ ein Wahrscheinlichkeitsraum, $(\Omega_2, \mathcal{A}_2)$ ein Messraum und Q ein Übergangskern von $(\Omega_1, \mathcal{A}_1)$ nach $(\Omega_2, \mathcal{A}_2)$. Dann wird durch

$$P(A) := \int_{\Omega_1} \left(\int_{\Omega_2} \mathbf{1}_A(\omega_1, \omega_2) Q(\omega_1, d\omega_2) \right) P_1(d\omega_1)$$

ein Wahrscheinlichkeitsmaß $P =: P_1 \otimes Q$ auf $\mathcal{A}_1 \otimes \mathcal{A}_2$ definiert. P heißt ***Koppelung*** *und ist das einzige Wahrscheinlichkeitsmaß auf $\mathcal{A}_1 \otimes \mathcal{A}_2$ mit der Eigenschaft*

$$P(A_1 \times A_2) = \int_{A_1} Q(\omega_1, A_2) P_1(d\omega_1) \quad (*)$$

Beweis

1. Ähnlich wie in §3 zeigt man: für $f : \Omega_1 \times \Omega_2 \to \mathbb{R}_+$, f $(\mathcal{A}_1 \otimes \mathcal{A}_2)$-messbar ist $\omega_1 \mapsto \int_{\Omega_2} f(\omega_1, \omega_2) Q(\omega_1, \mathrm{d}\omega_2)$ $\mathcal{A}_1$-messbar.
2. Für $A = A_1 \times A_2$ ist $\mathbf{1}_A(\omega_1, \omega_2) = \mathbf{1}_{A_1}(\omega_1)\mathbf{1}_{A_2}(\omega_2) \quad \Longrightarrow (*)$.
3. $P(\Omega_1 \times \Omega_2) = 1$ wegen (*). $P \geq 0$ ist klar.
$$\begin{aligned} P\left(\sum_{n=1}^{\infty} A_n\right) &= \int_{\Omega_1} \left(\int_{\Omega_2} \underbrace{\mathbf{1}_{\sum_{n=1}^{\infty}} (\omega_1, \omega_2)}_{=\sum_{n=1}^{\infty} \mathbf{1}_{A_n}(\omega_1,\omega_2)} Q(\omega_1, \mathrm{d}\omega_2) \right) P_1(\mathrm{d}\omega_1) \\ &= \sum_{n=1}^{\infty} \left(\int_{\Omega_1} \left(\int_{\Omega_2} \mathbf{1}_{A_n}(\omega_1, \omega_2) \, Q(\omega_1, \mathrm{d}\omega_2) \right) P_1(\mathrm{d}\omega_1) \right) \\ &= \sum_{n=1}^{\infty} P(A_n). \end{aligned}$$
4. Eindeutigkeitssatz für Maße. ∎

Satz 7.7 *Es seien $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum, $(\Omega_1, \mathcal{A}_1)$ ein messbarer Raum, $Y : \Omega \to \Omega_1$ $(\mathcal{A}, \mathcal{A}_1)$-messbar und X ein d-dimensionaler Zufallsvektor. Dann existiert ein Kern Q von $(\Omega_1, \mathcal{A}_1)$ nach $(\mathbb{R}^d, \mathfrak{B}^d)$ derart, dass*

$$P^{X,Y} = P^Y \otimes Q.$$

Q ist eine Version der bedingten Verteilung von X unter Y. Schreibweise:

$$Q(y, \cdot) = P^X(\cdot|Y = y).$$

Beweis - ohne Beweis - ∎

Bemerkung Für $A \in \mathcal{A}, B \in \mathfrak{B}^d$ gilt:

$$P(X \in B, Y \in A) = \int_A Q(y, B) \, P^Y \mathrm{d}y = \int_A P^X(B|Y = y) \, P^Y(\mathrm{d}y)$$

Satz 7.8
Es seien μ und ν σ-endliche Maße auf $\mathcal{A}_1$ bzw. $\mathfrak{B}^d$. $P^{(Y,X)}$ besitze eine Dichte f bezüglich $\mu \otimes \nu$. Es sei $f_Y(y) := \int_{\mathbb{R}^d} f(x, y)\nu(dx)$ die (Rand-)Dichte von P^Y bzgl. μ. Weiterhin sei

$$f(x|y) := \frac{f(x, y)}{f_Y(y)} \quad \text{und} \quad \frac{0}{0} := 0.$$

So wird durch

$$P^X(B|Y = y) := \int_B f(x|y)\nu(dx) \quad \forall B \in \mathfrak{B}^d,\ y \in \Omega_1$$

eine bedingte Verteilung von X unter der Bedingung $Y = y$ definiert.
*$f(\cdot|y)$ heißt **bedingte ν-Dichte von X unter der Bedingung** $Y = y$.*

Beweis
$y \mapsto \int_B f(x|y)\nu(\mathrm{d}x)$ ist messbar $\forall B \in \mathfrak{B}^d$ (Satz von Tonelli),
$B \mapsto \int_B f(x|y)\nu(\mathrm{d}x)$ ist ein Wahrscheinlichkeitsmaß $\forall y \in \Omega_1$.
Für $A \in \mathcal{A}_1, B \in \mathfrak{B}^d$ gilt:

$$\begin{aligned}
P^{(Y,X)}(A \times B) &= \int_{A\times B} f\mathrm{d}(\mu \otimes \nu) \\
&= \int_A \left(\int_B f(x,y)\,\nu(\mathrm{d}x)\right)\mu(\mathrm{d}y) \\
&= \int_A \left(\int_B f(x|y)\,\nu(\mathrm{d}x)\right) f_Y(y)\,\mu(\mathrm{d}y) \\
&\overset{!}{=} \int_A P^X(B|Y=y)\,\underbrace{P^Y(\mathrm{d}y)}_{=f_Y(y)\mu(\mathrm{d}y)}
\end{aligned}$$

■

Satz 7.9
Es seien $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum, $Y : \Omega \to \mathbb{R}^d$ ein Zufallsvektor und X eine Zufallsvariable mit $E|X| < \infty$. Dann ist

$$h(y) := \int_{\mathbb{R}} xP^X(dx|Y=y)$$

ein bedingter Erwartungswert von X unter der Bedingung $Y = y$.

Beweis Nach 7.5:

$$\int_B E[X|Y=y]\,P^Y(\mathrm{d}y) = \int_{Y^{-1}(B)} X\mathrm{d}P.$$

Für $B \in \mathfrak{B}^d$ und $T(Y,X) := X \cdot (\mathbf{1}_B \circ Y)$ gilt:

$$\begin{aligned}
\int_{Y^{-1}(B)} X\mathrm{d}P &= \int T(Y,X)\,\mathrm{d}P \\
&\overset{2.4}{=} \int T(y,x)\,P^{(Y,X)}(\mathrm{d}y,\mathrm{d}x) \\
&= \int x\mathbf{1}_B(y)\,P^{(Y,X)}(\mathrm{d}y,\mathrm{d}x) \\
&= \int_B \left(\int_{\mathbb{R}} xP^X(\mathrm{d}x|Y=y)\right) P^Y(\mathrm{d}y)
\end{aligned}$$

$\overset{7.5}{\Longrightarrow}$ Beh. ■

Beispiel 7.2
U und V seien unabhängig und $U(0,1)$-verteilt und entsprechen den zufälligen Seitenlängen eines Rechtecks. Es sei $X =$ Flächeninhalt des Rechtecks und $Y =$ Umfang des Rechtecks. Klar: X und Y sind nicht unabhängig.

Weiter ist $f_{U,V}(u,v) = \begin{cases} 1 & 0 < u < 1 \text{ und } 0 < v < 1 \\ 0 & \text{sonst} \end{cases}$ die gemeinsame Dichte von U und V.
$\Longrightarrow$ (Transformationssatz für Dichten) $f_{X,Y}(x,y) = \frac{2}{\sqrt{y^2-16x}}$ für $0 < x < 1$ und $4\sqrt{x} < y < 2+2x$; $f_X(x) = -\log x$ für $0 < x < 1$.
$\Longrightarrow$ $f(y|x) = -\frac{2}{\log x \sqrt{y^2-16x}}$ für $4\sqrt{x} < y < 2+2+x$.
$\Longrightarrow$ $E[Y|X=x] = \int y \cdot f(y|x) \mathrm{d}y = -\frac{4(1-x)}{\log x}$.

Beispiel 7.3 (Buffonsches Nadelproblem)
Wir werfen eine Nadel der Länge 1 zufällig auf einen unendlich langen Streifen der Breite 1. Wie groß ist die Wahrscheinlichkeit, dass die Nadel mindestens eine Wand des Korridors schneidet?
X = Abstand der Nadelmitte von der linken Wand
Y = Winkel der Nadel zum Lot
Annahme: $X \sim U(0,1)$, $Y \sim U(-\frac{\pi}{2}, \frac{\pi}{2})$ und X, Y unabhängig.
A = Nadel schneidet die Wand = $\{\omega \mid (X,Y)(\omega) \in B\}$ mit
$B = \{(x,y) \mid |y| < \frac{\pi}{2}, x \in [0, \frac{1}{2}\cos y] \cup [1 - \frac{1}{2}\cos y, 1]\}$
- hier fehlt eine Skizze -
Es ergibt sich:

$$\begin{aligned} P(A) &= P^{X,Y}(B) \\ &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^1 \mathbf{1}_B(x,y)\ P^X(\mathrm{d}x|Y=y)\ P^Y(\mathrm{d}y) \\ &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} P^X\left([0, \frac{\cos y}{2}] \cup [1 - \frac{\cos y}{2}, 1] \mid Y = y\right) P^Y(\mathrm{d}y) \\ &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos y \cdot \frac{1}{\pi} \mathrm{d}y \\ &= \frac{2}{\pi} \end{aligned}$$

So läßt sich zum Beispiel auch π näherungsweise bestimmen.