Vorlesung 11a

Bedingte Verteilung,
bedingte Wahrscheinlichkeiten

Bisher legten wir das Hauptaugenmerk auf den

Aufbau der gemeinsamen Verteilung von X_1 und X_2

aus der Verteilung ρ von X_1

und Übergangswahrscheinlichkeiten $P(a_1, .)$:

$$\mathbf{P}(X_1 = a_1, X_2 = a_2) := \rho(a_1)P(a_1, a_2)$$

Jetzt:

Zerlegung der gemeinsamen Verteilung von X_1 und X_2

in die Verteilung von X_1

und die *bedingte Verteilung* von X_2 gegeben X_1

$$\mathbf{P}(X_1 = a_1, X_2 = a_2) = \mathbf{P}(X_1 = a_1) \cdot \quad ?$$

Jetzt:

Zerlegung der gemeinsamen Verteilung von X_1 und X_2

in die Verteilung von X_1

und die *bedingte Verteilung* von X_2 gegeben X_1

$$\mathbf{P}(X_1 = a_1, X_2 = a_2) = \mathbf{P}(X_1 = a_1)\frac{\mathbf{P}(X_1 = a_1, X_2 = a_2)}{\mathbf{P}(X_1 = a_1)}$$

Sei X_1 eine diskrete Zufallsvariable mit Zielbereich S_1

und X_2 eine Zufallsvariable mit Zielbereich S_2.

Dann ist die

bedingte Wahrscheinlichkeit von $\{X_2 \in A_2\}$*,*

gegeben $\{X_1 = a_1\}$

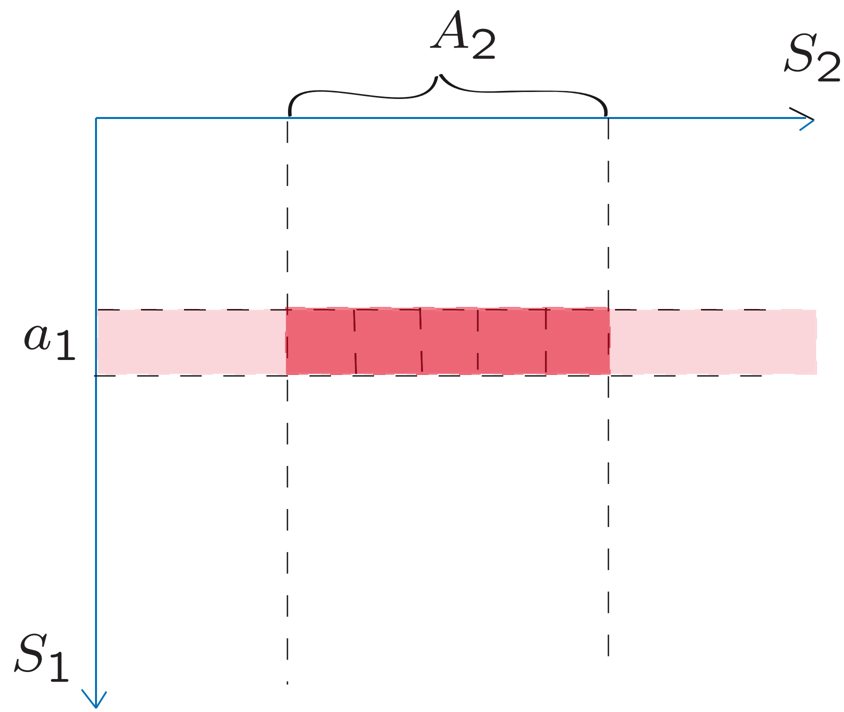
definiert als

$$\mathbf{P}(X_2 \in A_2 \,|\, X_1 = a_1) := \frac{\mathbf{P}(X_1 = a_1, X_2 \in A_2)}{\mathbf{P}(X_1 = a_1)} \,.$$

In der Matrix der gemeinsamen Verteilungsgewichte

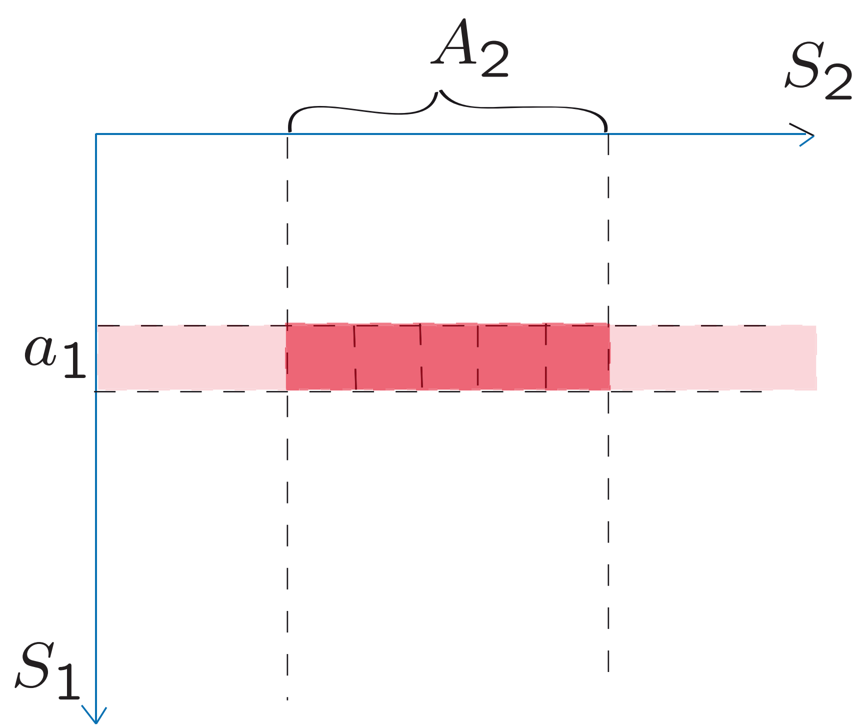
$$\mu(a_1, a_2) = \mathbf{P}(X_1 = a_1, X_2 = a_2)$$

ist $\mathbf{P}(X_2 \in A_2 \,|\, X_1 = a_1)$ das relative Gewicht von A_2 bezogen auf das Gesamtgewicht der Zeile a_1



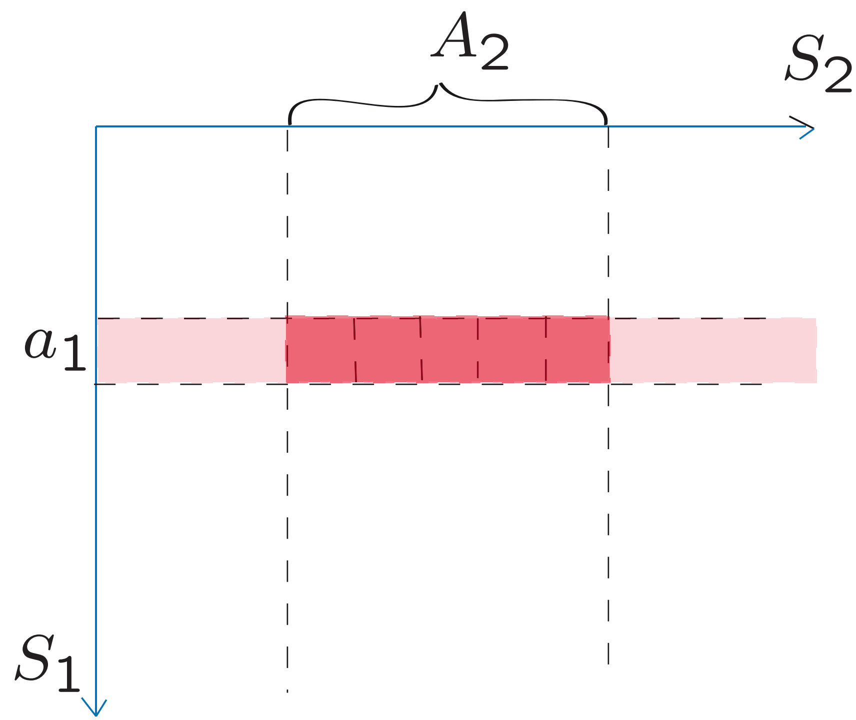
Die Verteilung $\mathbf{P}(X_2 \in \cdot \,|\, X_1 = a_1)$

heißt die *bedingte Verteilung von* X_2*, gegeben* $\{X_1 = a_1\}$.



Definieren wir Übergangswahrscheinlichkeiten durch

$$\mathbf{P}_{a_1}(X_2 \in A_2) := \mathbf{P}(X_2 \in A_2 \,|\, X_1 = a_1)$$



Definieren wir Übergangswahrscheinlichkeiten durch

$$\mathbf{P}_{a_1}(X_2 \in A_2) := \mathbf{P}(X_2 \in A_2 \,|\, X_1 = a_1)$$

dann bekommen wir die

aus den vorigen Vorlesungen vertraute Formel

$$\mathbf{P}(X_1 = a_1, X_2 \in A_2) = \mathbf{P}(X_1 = a_1)\,\mathbf{P}_{a_1}(X_2 \in A_2).$$

Beispiel:

Es seien Y und Z unabhängige $\mathbb{Z}$-wertige Zufallsvariable

$$X_1 := Y,\ X_2 := Y + Z$$

$$\mathbf{P}(Y = a, Y + Z = b) = \mathbf{P}(Y = a, Z = b - a)$$

$$= \mathbf{P}(Y = a)\,\mathbf{P}(Z = b - a)$$

Also:

$$\mathbf{P}(Y + Z = b \,|\, Y = a) = \mathbf{P}(a + Z = b)$$

Die bedingte Verteilung von $Y + Z$, gegeben $\{Y = a\}$,
ist gleich der Verteilung von $a + Z$.

Bei der Untersuchung von zwei Zufallsvariablen X_1, X_2

kann man immer

zu einer 2**-stufigen Betrachtungsweise** übergehen.

Man kann dabei wählen,

ob man X_1 in die erste Stufe aufnimmt oder in die zweite.

Beispiel:

Es seien Y und Z unabhängige $\mathbb{Z}$-wertige Zufallsvariable und

$X_1 := Y$, $X_2 := Y + Z$.

Wir haben gesehen:

Die bedingte Verteilung von $Y + Z$, gegeben $\{Y = a\}$,

ist die Verteilung von $a + Z$.

Was ergibt sich für die bedingte Verteilung von Y,

gegeben $\{Y + Z = b\}$?

“Wie war der erste Schritt?”

Die bedingte Verteilung von Y, gegeben $Y + Z = b$,

hat die Gewichte

$$\mathbf{P}(Y = a \,|\, Y + Z = b) = \frac{\mathbf{P}(Y = a)\,\mathbf{P}(Z = b - a)}{\mathbf{P}(Y + Z = b)}\,.$$

Ein instruktiver Spezialfall:

Y und Z seien unabhängig und Geom(p)-verteilt. Dann ist

$$\mathbf{P}(Y=a)\,\mathbf{P}(Z=b-a) = \qquad\qquad\qquad\qquad\qquad\qquad .$$

Ein instruktiver Spezialfall:

Y und Z seien unabhängig und Geom(p)-verteilt. Dann ist

$$\mathbf{P}(Y = a)\,\mathbf{P}(Z = b - a) = q^{a-1}p\,q^{(b-a)-1}p$$

Ein instruktiver Spezialfall:

Y und Z seien unabhängig und Geom(p)-verteilt. Dann ist

$$\mathbf{P}(Y = a)\,\mathbf{P}(Z = b - a) = q^{a-1}p\,q^{(b-a)-1}p = p^2 q^{b-2}\ .$$

Dieses hängt nicht von a ab.

Also ist die bedingte Verteilung

die uniforme auf $\{1, \ldots, b-1\}$

Das ist auch ohne Rechnung plausibel:

Gegeben, dass in einem p-Münzwurf

der zweite Erfolg beim b-ten Versuch kommt,

ist der Zeitpunkt des ersten Erfolges

uniform verteilt auf $\{1, \dots, b-1\}$.

Beispiel: Erfolgreiche Würfe beim Münzwurf

Bei einem 10-maligen p-Münzwurf sei

K die Anzahl der Erfolge,

und $G \subset \{1, \ldots, 10\}$ die zufällige Menge der Zeiten,

zu denen die Erfolge eintreten.

Wie ist die bedingte Verteilung von G, gegeben $\{K = 4\}$?

Für jede 4-elementige Teilmenge a von $\{1, \ldots, 10\}$ ist

$$\mathbf{P}(K = 4, G = a) = \mathbf{P}(G = a) = p^4(1-p)^6.$$

Das hängt nicht von a ab, also ist

die bedinge Verteilung von G uniform.

Bedingte Dichten.

Ist $f(a_1, a_2)\, da_1\, da_2$ gemeinsame Dichte von X_1 und X_2

und $f_1(a_1)\, da_1$ Dichte von X_1,

dann setzen wir

$$\mathbf{P}(X_2 \in da_2 \,|\, X_1 = a_1) := \frac{f(a_1, a_2)}{f_1(a_1)}\, da_2$$

und sprechen von der

bedingten Dichte von X_2, gegeben $\{X_1 = a_1\}$.

Beispiel: Exponentialverteilte Summanden

Y und Z seien unabhängig und Exp(1)-verteilt.

Was ist die bedingte Dichte von Y, gegeben $\{Y + Z = b\}$?

Die gemeinsame Dichte von Y und $Y + Z$ ist

$$e^{-y}e^{-(b-y)}\,dy\,db = e^{-b}\,dy\,db,\quad 0 \le y \le b$$

Die Dichte von $Y + Z$ ist $\quad \left(\int_0^b dy\right)e^{-b}\,db \quad = be^{-b}db$

Also:

$$\mathbf{P}(Y \in dy \,|\, Y + Z = b) = \frac{1}{be^{-b}}e^{-b}\,dy = \frac{1}{b}\,dy, \quad 0 \le y \le b.$$

Bedingte Wahrscheinlichkeiten:

Definition.

Seien E_1, E_2 Ereignisse. Dann ist die

bedingte Wahrscheinlichkeit von E_2, gegeben E_1,

definiert als

$$\mathbf{P}(E_2 \,|\, E_1) := \frac{\mathbf{P}(E_2 \cap E_1)}{\mathbf{P}(E_1)} = \mathrm{P}(I_{E_2} = 1 \,|\, I_{E_1} = 1)$$

…die Wahrscheinlichkeit von E_2, wenn man schon weiß, dass E_1 eingetreten ist.

Eine vertraute Regel (für zweistufige Experimente)

im neuen Gewand:

Formel für die totale Wahrscheinlichkeit:

$$\mathbf{P}(X_2 \in A_2) = \sum_{a_1 \in S_1} \mathbf{P}(X_1 = a_1)\mathbf{P}\big(X_2 \in A_2 \mid X_1 = a_1\big)$$

Zweistufigkeit - Spieß umgedreht:

$$\mathbf{P}(X_1 = a_1 \,|\, X_2 = a_2) = \frac{\mathbf{P}(X_2{=}a_2 \,|\, X_1{=}a_1)\mathbf{P}(X_1{=}a_1)}{\mathbf{P}(X_2{=}a_2)}$$

Formel von Bayes:

$$\mathbf{P}(X_1 = a_1 \,|\, X_2 = a_2) = \frac{\mathbf{P}(X_2{=}a_2 \,|\, X_1{=}a_1)\mathbf{P}(X_1{=}a_1)}{\sum_{b \in S_1} \mathbf{P}(X_2{=}a_2 \,|\, X_1{=}b)\mathbf{P}(X_1{=}b)}$$

$$\mathbf{P}(E_1 \,|\, E_2) = \frac{\mathbf{P}(E_2 \,|\, E_1)\mathbf{P}(E_1)}{\mathbf{P}(E_2 \,|\, E_1)\mathbf{P}(E_1) + \mathbf{P}(E_2 \,|\, E_1^c)\mathbf{P}(E_1^c)}$$

Beispiel: Bei einer bestimmten Reihenuntersuchung wird eine kranke Person in 100% der Fälle positiv getestet, eine gesunde Person in 1%.

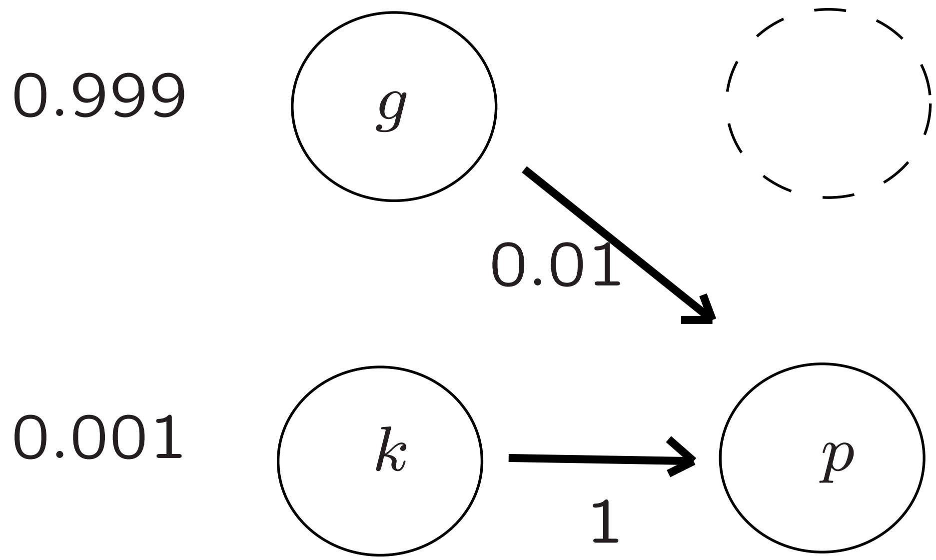
Wie groß ist die Wahrscheinlichkeit, dass eine positiv getestete Person wirklich krank ist?

Der Prozentsatz der kranken Personen sei 0.1% .

X_1 sei der Gesundheitszustand $(S_1 = \{g, k\})$,

X_2 der Testbefund $(S_2 = \{p, n\})$

(X_1, X_2) entsteht über ein zweistufiges Experiment:



$$\mathbf{P}(X_1 = k \,|\, X_2 = p) = \frac{\mathbf{P}(X_1 = k, X_2 = p)}{\mathbf{P}(X_2 = p)}$$

$$= \frac{0.001 \cdot 1}{0.999 \cdot 0.01 + 0.001 \cdot 1} \approx \frac{1}{11}$$

Bedingter Erwartungswert

In einem zweistufigen Experiment hatten wir

$$\mathbf{E}_{a_1}[h(X_1, X_2)] = \sum_{a_2 \in S_2} h(a_1, a_2)\mathbf{P}_{a_1}(X_2 = a_2)$$

Wegen

$$\mathbf{P}_{a_1}(X_2 = a_2) = \mathbf{P}\big(X_2 = a_2 \,|\, X_1 = a_1\big)$$

ist damit die folgende Definition konsistent:

Bedingter Erwartungswert von $h(X_1, X_2)$,

gegeben $\{X_1 = a_1\}$:

$$\mathbf{E}\big[h(X_1, X_2) \,|\, X_1 = a_1\big]$$

$$:= \sum_{a_2 \in S_2} h(a_1, a_2)\, \mathbf{P}\big(X_2 = a_2 \,|\, X_1 = a_1\big)$$

Bedingte Erwartung von $h(X_1, X_2)$, gegeben X_1:

$$\mathbf{E}\big[h(X_1, X_2) \,|\, X_1\big] := e(X_1),$$

mit

$$e(a_1) := \mathbf{E}\big[h(X_1, X_2) \mid X_1 = a_1\big].$$

Zum Merken:

Der **bedingte Erwartungswert** von Y, gegeben $X = x$

(Symbol : $\mathbf{E}[Y \,|\, X = x]$ oder $\mathbf{E}_x[Y]$)

ist *der Erwartungswert unter der bedingten Verteilung*.

Im diskreten Fall

$$\sum_y y\, \mathbf{P}(Y = y \,|\, X = x),$$

und im Fall von Dichten

$$\int y \, \frac{f(x,y)}{f_1(x)} \, dy.$$

Beispiel:

$Z_1, \ldots, Z_{10}$ sei ein p-Münzwurf der Länge 10, $K := \sum_{i=1}^{10} Z_i$.

Die *Runs* in $(z_1, \ldots, z_n)$ sind die

(in keinem größeren Block enthaltenen)

Blöcke aus nur Nullen oder nur Einsen.

Z. B. hat $(0, 1, 1, 0, 0, 1, 1, 0, 0, 0)$ fünf Runs:

$0, 11, 00, 11, 000.$

Sei R die Anzahl der Runs in $(Z_1, \ldots, Z_{10})$.

Gefragt ist nach $\mathbf{E}[R|K = 4]$.

$$
\begin{aligned}
R &= \sum_{i=1}^{n} I_{\{\text{beim } i\text{-ten Wurf beginnt ein Run}\}} \\
&= 1 + \sum_{i=1}^{9} I_{\{Z_i \neq Z_{i+1}\}}.
\end{aligned}
$$

Und die bedingte Verteilung von $(Z_1, \ldots, Z_{10})$ gegeben $K = 4$ entsteht so, dass man aus den Plätzen $1, \ldots, 10$ rein zufällig 4 auswählt, auf die man die 4 Einsen setzt.

$$
\mathbf{P}(Z_i \neq Z_{i+1} | K = 4) = 2 \cdot \frac{4}{10} \cdot \frac{6}{9} \quad \text{(warum?)},
$$

also ist der gesuchte Wert gleich $\quad 1 + 9 \cdot 2 \cdot \frac{4}{10} \cdot \frac{6}{9} = \frac{29}{5}.$

Gedächtnislosigkeit der geometrischen Verteilung:

T sei Geom(p)-verteilt. Dann gilt

$$\mathbf{P}(T > k + l \,|\, T > k) = q^{k+l}/q^k = q^l \; .$$

Die bedingte Verteilung von $T - k$, gegeben $\{T > k\}$,

ist somit gleich Geom(p).

Die Kenntnis, dass T einen Wert größer als k annimmt,

ändert also die Verteilung für die “restliche Wartezeit” nicht.

Gedächtnislosigkeit der Exponentialverteilung

Für exponentialverteiltes T zum Parameter λ gilt für $r, s > 0$

$$\mathbf{P}(T > r + s \,|\, T > r) = e^{-\lambda s} \,.$$

Die bedingte Verteilung von $T - s$, gegeben $\{T > s\}$,

ist somit gleich Exp(λ).