

# Vorlesung 9b

## Bedingte Verteilungen und bedingte Wahrscheinlichkeiten

Voriges Mal:

**Aufbau** der gemeinsamen Verteilung von  $X_1$  und  $X_2$

aus der Verteilung  $\rho$  von  $X_1$

und Übergangswahrscheinlichkeiten  $P(a_1, \cdot)$ :

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

Heute:

**Zerlegung** der gemeinsamen Verteilung von  $X_1$  und  $X_2$

in die Verteilung von  $X_1$

und die *bedingte Verteilung* von  $X_2$  gegeben  $X_1$

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

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

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

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

Definieren wir Übergangswahrscheinlichkeiten durch

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

dann bekommen wir die

aus der vorigen Vorlesung vertraute Formel

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

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  $Y + Z = b$ , ist

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

Ein instruktiver Spezialfall:

$Y$  und  $Z$  seien unabhängig und  $\text{Geom}(p)$ -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  $\{1, \dots, b - 1\}$

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  $\{1, \dots, b - 1\}$ .

## 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

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

und sprechen von der

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

## Beispiel: Exponentialverteilte Summanden

$Y$  und  $Z$  seien unabhängig und  $\text{Exp}(1)$ -verteilt.

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

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

$$e^{-y} e^{-(b-y)} dy db = e^{-b} dy db$$

Die Dichte von  $Y + Z$  ist (man erinnere sich!)  $b e^{-b} db$

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

## Münzwurf mit zufälliger Erfolgswahrscheinlichkeit

$$\mathbf{P}(U \in du, X_n = k) = du \binom{n}{k} u^k (1 - u)^{n-k}$$

$$\mathbf{P}(X_n = k) = \frac{1}{n + 1}$$

$$\frac{\mathbf{P}(U \in du, X_n = k)}{\mathbf{P}(X_n = k)} = du \binom{n}{k} u^k (1 - u)^{n-k} / \frac{1}{n + 1}$$

$$\mathbf{P}(U \in du | X_n = k) = \frac{(n+1)!}{k!(n-k)!} u^k (1 - u)^{n-k} du$$

Alte Regeln (für zweistufige Experimente) im neuen Gewand:

*Formel für die totale Wahrscheinlichkeit:*

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

*Bedingter Erwartungswert von  $h(X_1, X_2)$ ,  
gegeben  $\{X_1 = a_1\}$ :*

$$\begin{aligned} & \mathbf{E}[h(X_1, X_2) \mid X_1 = a_1] \\ := & \sum_{a_2 \in S_2} h(a_1, a_2) \mathbf{P}(X_2 = a_2 \mid X_1 = a_1) \end{aligned}$$

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

$$\begin{aligned} \mathbf{E}[h(X_1, X_2) \mid X_1] &:= e(X_1), \\ &\text{mit} \\ e(a_1) &:= \mathbf{E}[h(X_1, X_2) \mid X_1 = a_1]. \end{aligned}$$



### Definition.

Seien  $E_1, E_2$  Ereignisse. Dann ist die  
*bedingte Wahrscheinlichkeit von  $E_2$ , gegeben  $E_1$ ,*  
definiert als

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

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

*Gedächtnislosigkeit der geometrischen Verteilung:*

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

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  $\lambda$  gilt für  $r, s > 0$

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

Zweistufigkeit - Spieß umgedreht:

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

*Formel von Bayes*

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

$$\mathbf{P}(E \mid E') = \frac{\mathbf{P}(E' \mid E)\mathbf{P}(E)}{\mathbf{P}(E' \mid E)\mathbf{P}(E) + \mathbf{P}(E' \mid E^c)\mathbf{P}(E^c)}$$

Beispiel: Reihenuntersuchungen:

Eine kranke Person wird in 100% der Fälle positiv getestet,  
eine gesunde Person in 1%.

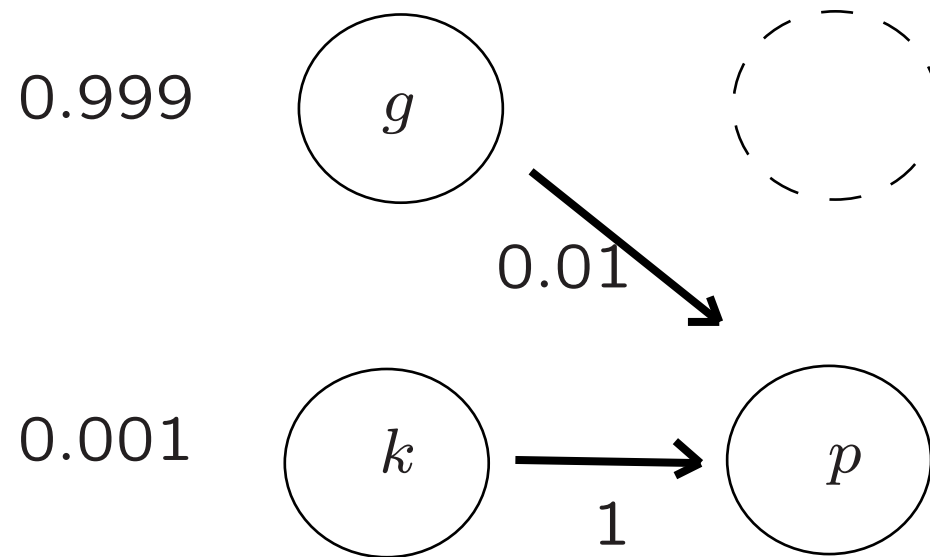
Wie groß ist die Wahrscheinlichkeit, dass eine positiv  
getestete Person wirklich krank ist?

Der Prozentsatz der kranken Personen sei 0.1% .

$$P(E) = 0.001, \quad P(E' | E) = 1, \quad P(E' | E^c) = 0.01.$$

$$\text{Bayes: } P(E | E') = 0.091.$$

Die Wahrscheinlichkeit, bei positivem Testbefund wirklich krank zu sein,  
liegt also bei nur 9%!



$$\frac{\mathbf{P}(X_1 = k, X_2 = p)}{\mathbf{P}(X_2 = p)} = \frac{0.001 \cdot 1}{0.999 \cdot 0.01 + 0.001 \cdot 1}$$

$$\approx \frac{1}{11}$$

Zur Erinnerung:

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

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

ist der Erwartungswert unter der bedingten Verteilung.

Im diskreten Fall

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

und im Fall von Dichten

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

Beispiel:

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

Die *Runs* in  $(z_1, \dots, z_n)$  sind  
die Maximalserien aus nur Nullen oder nur Einsen.

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

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

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



Es gilt in jedem Fall:

$$R = 1 + \sum_{i=1}^9 I_{\{Z_i \neq Z_{i+1}\}}.$$

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

$$\mathbf{P}(Z_i \neq Z_{i+1}) = 2 \cdot \frac{4}{10} \cdot \frac{6}{9},$$

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

Beispiel:

$Z_1, Z_2, \dots$  sei ein Münzwurf mit uniform auf  $[0, 1]$  verteiltem zufälligem Erfolgsparameter  $U$ ,

$X_n$  sei die Anzahl der Erfolge in den ersten  $n$  Versuchen.

Gefragt ist nach  $\mathbf{E}[U \mid X_n = k]$ .

Wir wissen schon:

Die bedingte Dichte von  $U$  gegeben  $\{X_n = k\}$  ist

$$\mathbf{P}_k(U \in du) = \frac{1}{\frac{1}{n+1}} \binom{n}{k} u^k (1-u)^{n-k} du$$

$$\begin{aligned}
\mathbf{E}[U \mid X_n = k] &= \int \frac{1}{\frac{1}{n+1}} \binom{n}{k} u^{k+1} (1-u)^{n-k} du \\
&= \frac{k+1}{n+2} \int \frac{1}{\frac{1}{n+2}} \binom{n+1}{k+1} u^{k+1} (1-u)^{(n+1)-(k+1)} du \\
&= \frac{k+1}{n+2}.
\end{aligned}$$

Man nennt dies auch den

**Bayes-Schätzer für die Erfolgswahrscheinlichkeit**

(bei a priori uniform verteilter Erfolgswahrscheinlichkeit).

Hier ist -außer Konkurrenz - noch ein probabilistisches Argument für

$$\mathbf{E}[U \mid X_n = k] = \frac{k+1}{n+2}, \text{ ohne analytische Rechnung.}$$

Wie in Vorlesung 9a seien

$U, U_1, \dots$  unabhängig und uniform auf  $[0, 1]$ .

Die aufsteigend geordneten  $U, \dots, U_{n+1}$  nennen wir  $V_0, \dots, V_{n+1}$ .

Die Verteilung von  $U$  bedingt unter  $X_n = k$  stimmt überein mit der Verteilung von  $V_k$  (siehe Vorlesung 9a).

Sei  $W_i := V_i - V_{i-1}$ ,  $i = 1, \dots, n+2$ , mit  $V_{-1} := 0$ ,  $V_{n+2} := 1$ .

Die  $W_i$  sind identisch verteilt mit  $\sum_{i=0}^{n+1} \mathbf{E}[W_i] = 1$ . Also folgt

$$\mathbf{E}[V_k] = \sum_{i=0}^k \mathbf{E}[W_i] = \frac{k+1}{n+2}. \quad \square$$