Vorlesung 9a

Zweistufige Zufallsexperimente

Teil 2

Zur Erinnerung: Wir denken uns das zufällige Paar

$$X = (X_1, X_2)$$

auf zweistufige Weise zustande gekommen:

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

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

nennen wir den

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

gegeben $\{X_1 = a_1\}$.

$$\begin{aligned}
& \mathbf{E}[h(X_1, X_2)] \\
&= \sum_{a_1 \in S_1} \sum_{a_2 \in S_2} h(a_1, a_2) \mathbf{P}(X_1 = a_1, X_2 = a_2) \\
&= \sum_{a_1 \in S_1} \sum_{a_2 \in S_2} h(a_1, a_2) \mathbf{P}(X_1 = a_1) \mathbf{P}_{a_1}(X_2 = a_2) \\
&= \sum_{a_1 \in S_1} \mathbf{P}(X_1 = a_1) \mathbf{E}_{a_1}[h(X_1, X_2)] \\
&= \sum_{a_1 \in S_1} \mathbf{P}(X_1 = a_1) \mathbf{E}_{a_1}[h(a_1, X_2)] \\
&= \mathbf{E}[\mathbf{E}_{X_1}[h(X_1, X_2)]].
\end{aligned}$$

“Zerlegung des Erwartungswertes nach der ersten Stufe”

Merke: Der bedingte Erwartungswert

$$\mathbf{E}_{a_1} [h(X_1, X_2)]$$

ist eine Zahl.

$$\mathbf{E}_{X_1} [h(X_1, X_2)]$$

ist eine **Zufallsvariable**.

Wir sprechen von der

bedingten Erwartung von $h(X_1, X_2)$ gegeben X_1 .

**Die bedingte Erwartung als
beste Prognose im quadratischen Mittel:**

Satz:

Sei X_2 reellwertige Zufallsvariable mit $\mathbf{E}[X_2^2] < \infty$.

Dann minimiert die bedingte Erwartung $\mathbf{E}_{X_1}[X_2]$
unter allen reellwertigen Zufallsvariablen der Form $h(X_1)$
den erwarteten quadratischen Abstand

$$\mathbf{E}[(X_2 - h(X_1))^2].$$

Beweis:

Wir zerlegen $\mathbf{E}[(X_2 - h(X_1))^2]$ nach X_1 :

$$\begin{aligned}\mathbf{E}[(X_2 - h(X_1))^2] &= \mathbf{E}\left[\mathbf{E}_{X_1}[(X_2 - h(X_1))^2]\right] \\ &= \sum_{a_1} \mathbf{P}(X_1 = a_1) \mathbf{E}_{a_1}[(X_2 - h(a_1))^2]\end{aligned}$$

Wir wissen schon (aus Vorlesung 8a):

$$\begin{aligned}\mathbf{E}_{a_1}[(X_2 - h(a_1))^2] \text{ wird minimal f\u00fcr} \\ h(a_1) := \mathbf{E}_{a_1}[X_2]. \quad \square\end{aligned}$$

Definieren wir die
bedingte Varianz von X_2 , gegeben $\{X_1 = a_1\}$:

$$\mathbf{Var}_{a_1}[X_2] := \mathbf{E}_{a_1}[(X_2 - \mathbf{E}_{a_1}[X_2])^2]$$

Dann haben wir die Zerlegung

$$\mathbf{E}_{a_1}[(X_2 - h(a_1))^2] = \mathbf{Var}_{a_1}[X_2] + \left(\mathbf{E}_{a_1}[X_2] - h(a_1)\right)^2$$

Denn was der Varianz recht ist:

$$\mathbf{E}[(Z - c)^2] = \mathbf{Var}[Z] + (\mathbf{E}Z - c)^2$$

ist der bedingten Varianz billig!

$$\mathbf{E}_{a_1}[(X_2 - h(a_1))^2] = \mathbf{Var}_{a_1}[X_2] + \left(\mathbf{E}_{a_1}[X_2] - h(a_1)\right)^2$$

Ersetzen wir a_1 durch die Zufallsvariable X_1

und bilden wir den Erwartungswert,

dann bekommen wir die Zerlegung

$$\begin{aligned} & \mathbf{E}[(X_2 - h(X_1))^2] \\ &= \mathbf{E}[\mathbf{Var}_{X_1}[X_2]] + \mathbf{E}\left[\left(\mathbf{E}_{X_1}[X_2] - h(X_1)\right)^2\right]. \end{aligned}$$

Wählen wir speziell $h(X_1) := \mathbf{E}[X_2]$, dann ergibt sich

$$\mathbf{Var}[X_2] = \mathbf{E}[\mathbf{Var}_{X_1}[X_2]] + \mathbf{Var}[\mathbf{E}_{X_1}[X_2]]$$

“Zerlegung der Varianz nach der ersten Stufe”

Beispiel: Zufällige Anzahl unabhängiger Summanden.

$$Y := \sum_{i=1}^N Z_i$$

mit $Z_1, Z_2, \dots$ unabh., ident. vert. und unabhängig von N .

$$\mu := \mathbf{E}[Z_1], \quad \sigma^2 := \mathbf{Var}[Z_1]$$

$$\mathbf{E}_n[Y] = n\mu, \quad \mathbf{Var}_n[Y] = n\sigma^2.$$

$$\mathbf{E}[Y] = \mathbf{E}[\mathbf{E}_N Y] = \mathbf{E}[N\mu] = \mathbf{E}[N] \cdot \mu.$$

$$\begin{aligned} \mathbf{Var}[Y] &= \mathbf{E}[\mathbf{Var}_N[Y]] + \mathbf{Var}[\mathbf{E}_N[Y]] \\ &= \mathbf{E}[N] \cdot \sigma^2 + \mathbf{Var}[N] \cdot \mu^2. \end{aligned}$$

Beispiel: Addieren von unabhängigen ZV'en
– zweistufig aufgefasst (vgl. Buch S. 92)

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

$$\begin{aligned} & \mathbf{P}(X_1 = a_1, X_2 = a_2) \\ &= \mathbf{P}(Y = a_1, a_1 + Z = a_2) \\ &= \mathbf{P}(X_1 = a_1) \mathbf{P}(a_1 + Z = a_2) \end{aligned}$$

Dies führt zu den Übergangswahrscheinlichkeiten

$$P(a_1, a_2) := \mathbf{P}(Z = a_2 - a_1)$$

$$\mathbf{P}(Y + Z = b) = \sum_y \mathbf{P}(Y = y) \mathbf{P}(Z = b - y)$$

Beispiel (vgl. Buch S. 35)

Y, Z unabhängig und $\text{Geom}(p)$ -verteilt

$$\begin{aligned} \mathbf{P}(Y + Z = b) &= \sum_{y=1}^{b-1} pq^{y-1} pq^{b-y-1} \\ &= (b-1)p^2q^{b-2}, \quad b = 2, 3, \dots \end{aligned}$$

Die *negative Binomialverteilung* mit Parametern $2, p$
ist die Verteilung der Anzahl der Versuche
in einem p -Münzwurf bis einschließlich zum zweiten Erfolg.

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

Haben Y und Z die Dichten $f(y) dy$ und $g(z) dz$,
so hat man analog

$$\begin{aligned}\mathbf{P}(Y \in da, Y + Z \in db) &= \mathbf{P}(Y \in da, a + Z \in db) \\ &= \mathbf{P}(Y \in da) \mathbf{P}(a + Z \in db)\end{aligned}$$

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

Haben Y und Z die Dichten $f(y) dy$ und $g(z) dz$,
so hat man analog

$$\begin{aligned}\mathbf{P}(Y \in da, Y + Z \in db) &= \mathbf{P}(Y \in da, a + Z \in db) \\ &= \mathbf{P}(Y \in da) \mathbf{P}(a + Z \in db) \\ &= f(a) g(b - a) da db\end{aligned}$$

$$\mathbf{P}(Y \in da, Y + Z \in db) = f(a) g(b - a) da db$$

Die Formel von der totalen Wahrscheinlichkeit lautet hier

$$\mathbf{P}(Y + Z \in db) = \left(\int f(a) g(b - a) da \right) db .$$

Beispiel: Für Y und Z unabhängig und $\text{Exp}(1)$ -verteilt ist

$$\begin{aligned} \mathbf{P}(Y + Z \in db) &= \left(\int_0^b e^{-a} e^{-(b-a)} da \right) db \\ &= b e^{-b} db, \quad b \geq 0. \end{aligned}$$

(Dichte der Gamma(2)-Verteilung)

Beispiel: Münzwurf mit zufälliger Erfolgsw'keit.

Erste Stufe: Zufällige Erfolgswahrscheinlichkeit U .

Zweite Stufe: Gegeben $U = u$ führe u -Münzwurf durch.

Dadurch entsteht eine zufällige 01-Folge $Z_1, Z_2, \dots$

Ist U uniform verteilt auf dem Intervall $[0, 1]$,

dann ergibt sich für $X_n := Z_1 + \dots + Z_n$

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

$$\mathbf{P}(X_n = k) = \binom{n}{k} \int_0^1 u^k (1 - u)^{n-k} du, \quad k = 0, 1, \dots, n.$$

Eine Darstellung des Münzwurfs mit zufälliger,
uniform verteilter Erfolgswahrscheinlichkeit:

Seien $U, U_1, U_2, \dots$ unabhängig und uniform auf $[0, 1]$.

U gibt die zufällige Erfolgswahrscheinlichkeit, und

$$Z_1 := I_{\{U_1 < U\}}, \quad Z_2 := I_{\{U_2 < U\}}, \quad \dots \quad (1)$$

stellt einen Münzwurf mit zufälligem Parameter U dar.

$$\begin{aligned} & \{X_n = k\} = \{Z_1 + \dots + Z_n = k\} \\ & = \left\{ \text{von den } U_1, \dots, U_n \text{ sind genau } k \text{ Stück kleiner als } U \right\}. \end{aligned}$$

Aus Symmetriegründen hat jedes der Ereignisse $E_k := \{ \text{von den } U_1, \dots, U_n \text{ sind genau } k \text{ Stück kleiner als } U \}$ die Wahrscheinlichkeit $\frac{1}{n+1}$ (denn die Anordnung der $U, U_1, U_2, \dots$ ist rein zufällig).

Also folgt:

$$\mathbf{P}(X_n = k) = \frac{1}{n+1}, \quad k = 0, 1, \dots, n.$$

Fazit:

Die Anzahl der Erfolge im n -fachen Münzwurf mit zufälliger auf $[0, 1]$ uniform verteilter Erfolgswahrscheinlichkeit ist uniform verteilt auf $\{0, 1, \dots, n\}$.