Vorlesung 10b

Bedingte Varianz

und

Weiteres zur Zweistufigkeit

$$(X_2 = X_1 + Z)$$

Zur Erinnerung:

Der bedingte Erwartungswert

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

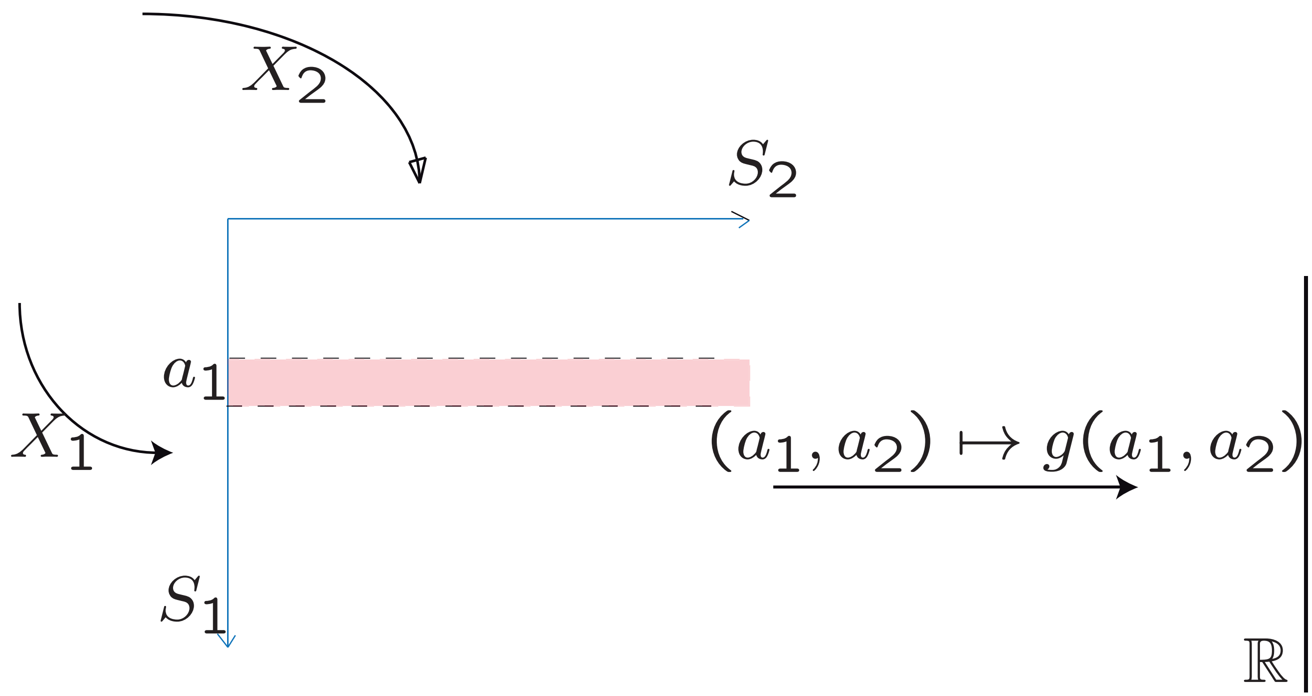
ist eine Zahl.

Er wird gebildet mit der Übergangsverteilung $P(a_1, .)$,

also mit den Wahrscheinlichkeitsgewichten,

die die Zeile $P(a_1, .)$ der Matrix P bilden:

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



X_2
S_2
a_1
X_1
$(a_1, a_2) \mapsto g(a_1, a_2)$
S_1
$\mathbb{R}$

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

ist eine Zufallsvariable.

Wir nennen $\mathbf{E}_{X_1}\big[g(X_1, X_2)\big]$ die

bedingte Erwartung von $g(X_1, X_2)$ gegeben X_1.

Wir haben gesehen:

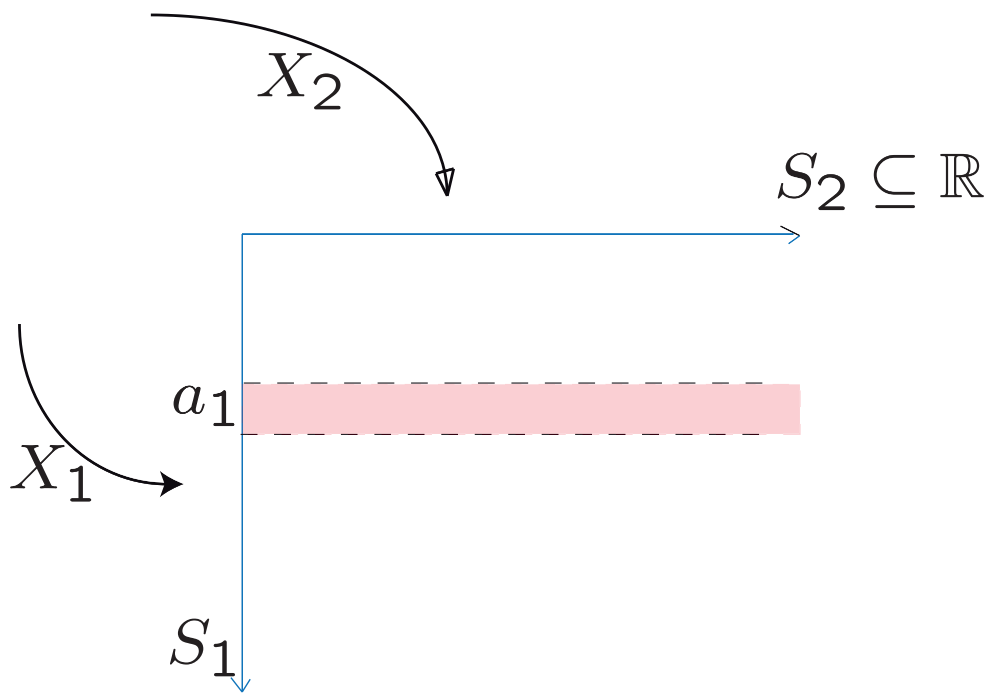
$$\mathbf{E}\Big[g(X_1, X_2)\Big] = \mathbf{E}\Big[\mathbf{E}_{X_1}[g(X_1, X_2)]\Big].$$

"Zerlegung des Erwartungswertes nach der ersten Stufe"

Der bedingte Erwartungswert als

beste Prognose im quadratischen Mittel:

Es sei jetzt $S_2 \subset \mathbb{R}$.



Der bedingte Erwartungswert als

beste Prognose im quadratischen Mittel:

Wir wissen schon (aus Vorlesung 9a):

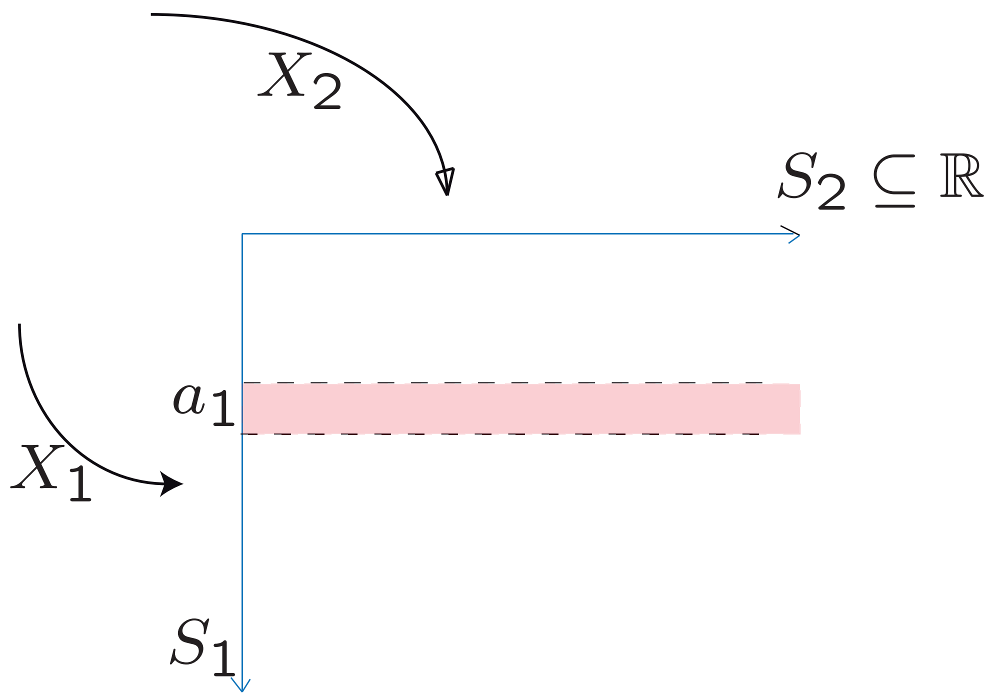
$\mathbf{E}_{a_1}\left[(X_2 - h(a_1))^2\right]$ wird minimal für

$$h(a_1) := \mathbf{E}_{a_1}[X_2].$$

In diesem Sinn ist also der bedingte Erwartungswert $\mathbf{E}_{a_1}[X_2]$ die (im Sinne des erwarteten quadratischen Abstandes) beste Prognose von X_2, gegeben $\{X_1 = a_1\}$

Die bedingte Erwartung als

beste Prognose im quadratischen Mittel:



Die bedingte Erwartung als beste Prognose im quadratischen Mittel:

Satz:

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

Unter allen reellwertigen Zufallsvariablen der Form $h(X_1)$

minimiert die bedingte Erwartung $\mathbf{E}_{X_1}[X_2]$

den erwarteten quadratischen Abstand

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

Beweis:

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

$$\mathbf{E}\big[(X_2 - h(X_1))^2\big] = \mathbf{E}\Big[\mathbf{E}_{X_1}\big[(X_2 - h(X_1))^2\big]\Big]$$
$$= \sum_{a_1} \mathbf{P}(X_1 = a_1)\mathbf{E}_{a_1}\big[(X_2 - h(a_1))^2\big]$$

Wir wissen schon (aus Vorlesung 9a):

$\mathbf{E}_{a_1}\big[(X_2 - h(a_1))^2\big]$ wird minimal für

$h(a_1) := \mathbf{E}_{a_1}[X_2]$. $\square$

Fazit:

Unter allen Zahlen $h(a_1)$

ist der bedingte Erwartungswert $\mathbf{E}_{a_1}[X_2]$

diejenige Zahl, für die $\mathbf{E}_{a_1}[(X_2 - h(a_1))^2]$ minimal wird.

Unter allen Zufallsvariablen der Form $h(X_1)$

ist die bedingte Erwartung $\mathbf{E}_{X_1}[X_2]$

diejenige, für die

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

minimal wird.

Definieren wir die

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

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

Die Zerlegung

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

überträgt sich

auf den *bedingten* Erwartungswert und die *bedingte Varianz*:

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

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

Ersetzen wir a_1 durch die Zufallsvariable X_1:

$$\mathbf{E}_{X_1}\Big[(X_2 - h(X_1))^2\Big] = \mathbf{Var}_{X_1}[X_2] + \Big(\mathbf{E}_{X_1}[X_2] - h(X_1)\Big)^2$$

und bilden den Erwartungswert, dann bekommen wir

$$\mathbf{E}\Big[(X_2 - h(X_1))^2\Big]$$
$$= \mathbf{E}\Big[\mathbf{Var}_{X_1}[X_2]\Big] + \mathbf{E}\Big[\Big(\mathbf{E}_{X_1}[X_2] - h(X_1)\Big)^2\Big].$$

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

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

“Zerlegung der Varianz nach der ersten Stufe”

Beispiel: Zufällige Anzahl unabhängiger Summanden.

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

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

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

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

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

$$\mathbf{Var}[Y] = \mathbf{E}\big[\mathbf{Var}_N[Y]\big] + \mathbf{Var}\big[\mathbf{E}_N[Y]\big]$$

$$= \mathbf{E}[N] \cdot \sigma^2 + \mathbf{Var}[N] \cdot \mu^2.$$

Weiteres zur Zweistufigkeit:

Beispiel: Addieren von unabhängigen ZV'en

– zweistufig aufgefasst (vgl. Buch S. 92)

$$X_2 = X_1 + Z$$

mit X_1, Z unabhängig

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

$$= \mathbf{P}(X_1 = a_1, Z = a_2 - a_1)$$

$$= \mathbf{P}(X_1 = a_1)\,\mathbf{P}(Z = a_2 - a_1)$$

Dies führt zu den Übergangswahrscheinlichkeiten

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

Die Formel von der totalen Wahrscheinlichkeit ergibt

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

$$\mathbf{P}(X_1 + Z = a_2) = \sum_{a_1} \mathbf{P}(X_1 = a_1)\, \mathbf{P}(Z = a_2 - a_1)$$

Y, Z unabhängig:

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

Beispiel

Y, Z unabhängig und Geom(p)-verteilt

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

Die *negative Binomialverteilung* mit Parametern 2, p

ist die Verteilung der Anzahl der Versuche

in einem p-Münzwurf bis einschließlich zum zweiten Erfolg.