

Vorlesung 8b

Bedingte Erwartung und bedingte Varianz

Teil 3:

Erwartete quadratische Prognosefehler
und Optimalität der bedingten Erwartung

(Buch S. 90)

Wir erinnern uns an die “hilfreiche Formel für die Varianz”:

$$\mathbf{E}[Y^2] = \mathbf{Var}[Y] + (\mathbf{E}[Y])^2.$$

Verschieben um eine Konstante ändert die Varianz nicht:

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

Was der Varianz recht ist, ist der bedingten Varianz billig:

Für alle $a_1 \in S_1$ und beliebiges $h(a_1) \in \mathbb{R}$ ist

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

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

Anstelle von a_1 können wir die Zufallsvariable X_1 einsetzen:

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

Bilde in dieser Gleichheit den Erwartungswert und verwende links dessen Zerlegung nach X_1 .

(also die Formel auf Folie 8 in Teil 1, hier mit

$$g(a_1, a_2) := (a_2 - h(a_1))^2:$$

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

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

Diese Formel können wir so lesen:

Der **erwartete quadratische Prognosefehler**
bei Verwendung der Prognose $h(X_1)$ ist die Summe
aus dem **Erwartungswert der bedingten Varianz**
und dem “**erwarteten quadratischen Bias**”,
der dadurch entsteht, dass man $h(X_1)$ anstelle von
 $\mathbf{E}_{X_1}[X_2]$ als Prognose von X_2 verwendet.

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 wissen schon:

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

Der Satz über die Positivität des Erwartungswerts impliziert:

Der **letzte Term rechts** wird minimal (nämlich 0)

genau dann, wenn

$$\mathbf{P}\left(h(X_1) = \mathbf{E}_{X_1}[X_2]\right) = 1.$$

Äquivalent dazu:

$$h(a_1) = \mathbf{E}_{a_1}[X_2] \quad \text{für alle } a_1 \text{ mit } \mathbf{P}(X_1 = a_1) > 0. \quad \square$$

Fazit:

1. 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.

2. 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.