1 Optimalitätseigenschaft der bedingten Erwartung

Satz 1.1. *Es seien X, Y Zufallsvariablen mit $Y \in L^2$ und $\mathbb{P}(X = x) > 0$ für alle $x \in X(\Omega)$. Setze*

$$\mathcal{F}_X := \{Z : \Omega \to \mathbb{R} \mid \exists h : X(\Omega) \to \mathbb{R} : \, Z = h(X)\}.$$

Dann nimmt die Funktion

$$\varphi : \mathcal{F}_X \to \mathbb{R}, \quad \varphi(Z) = \mathbb{E}[(Y - Z)^2]$$

ihr Minimum bei $\mathbb{E}[Y \mid X]$ an.

Beachte, dass $\varphi(Z)$ nicht die Komposition einer Funktion φ mit Z bezeichnet und daher keine Zufallsvariable ist. Vielmehr ist $\varphi(Z)$ ein Wert, der der gesamten Zufallsvariablen (Funktion) Z zugeordnet wird. Wir benötigen zuerst eine wichtige Eigenschaft bedingter Erwartungen.

Lemma 1.2. *Es seien X, Z Zufallsvariablen und $h : X(\Omega) \to \mathbb{R}$ eine Funktion mit $Z, h(X) \in L^1$ und $\mathbb{P}(X = x) > 0$ für alle $x \in X(\Omega)$. Dann gilt*

$$\mathbb{E}[Z \bullet h(X) \mid X] = h(X) \, \mathbb{E}[Z \mid X].$$

Beweis. Wir zeigen die entsprechende Identität für bedingte Erwartungswerte. Für $x \in X(\Omega)$ gilt

$$\begin{aligned}
\mathbb{E}[Z \bullet h(X) \mid X = x] &= \sum_{y \in (Z \bullet h(X))(\Omega)} y \, \mathbb{P}(Z \bullet h(X) = y \mid X = x) \\
&= \sum_{z \in Z(\Omega)} \sum_{\eta \in h(X(\Omega))} z\eta \, \mathbb{P}(Z = z, \, h(X) = \eta \mid X = x) \\
&= \sum_{z \in Z(\Omega)} \sum_{\eta \in h(X(\Omega))} z\eta \, \mathbb{P}(Z = z, \, h(x) = \eta \mid X = x) \\
&= \sum_{z \in Z(\Omega)} z h(x) \, \mathbb{P}(Z = z \mid X = x) \\
&= h(x) \, \mathbb{E}[Z \mid X = x].
\end{aligned}$$

□

Beweis des Satzes. Für $Z \in \mathcal{F}_X$ mit $Z = h(X)$ gilt

$$\begin{aligned}
&\varphi(Z) = \mathbb{E}[((Y - \mathbb{E}[Y \mid X]) - (Z - \mathbb{E}[Y \mid X]))^2] \\
&= \varphi(\mathbb{E}[Y \mid X]) - 2\,\mathbb{E}[(Y - \mathbb{E}[Y \mid X])(Z - \mathbb{E}[Y \mid X])] + \mathbb{E}[(Z - \mathbb{E}[Y \mid X])^2] \\
&= \varphi(\mathbb{E}[Y \mid X]) - 2\,\mathbb{E}[\mathbb{E}[(Y - \mathbb{E}[Y \mid X])(h(X) - \mathbb{E}[Y \mid X]) \mid X]] + \mathbb{E}[(Z - \mathbb{E}[Y \mid X])^2] \\
&= \varphi(\mathbb{E}[Y \mid X]) - 2\,\mathbb{E}[\mathbb{E}[Y - \mathbb{E}[Y \mid X] \mid X](h(X) - \mathbb{E}[Y \mid X])] + \mathbb{E}[(Z - \mathbb{E}[Y \mid X])^2] \\
&= \varphi(\mathbb{E}[Y \mid X]) - 2\,\mathbb{E}[(\mathbb{E}[Y \mid X] - \mathbb{E}[Y \mid X])(h(X) - \mathbb{E}[Y \mid X])] + \mathbb{E}[(Z - \mathbb{E}[Y \mid X])^2] \\
&= \varphi(\mathbb{E}[Y \mid X]) + \mathbb{E}[(Z - \mathbb{E}[Y \mid X])^2].
\end{aligned}$$

Es wurden nacheinander benutzt: die Definition von φ, die Linearität des Erwartungswertes, die Eigenschaft $\mathbb{E}[\mathbb{E}[Z \mid X]] = \mathbb{E}[Z]$, das obige Lemma und die Linearität der bedingten Erwartung. Weil der letzte Term stets nicht-negativ ist, ergibt sich die Behauptung. □

Markus Reiß

Vorlesung
Einführung in die Statistik
Sommersemester 2005

Lösungsansatz zur Zusatzaufgabe in 4.4

Wir bestimmen die Verteilung der Anzahl A von gekauften Bildchen, bis jeder Spieler doppelt vertreten ist. Das Ereignis $\{A \leqslant m\}$ tritt also ein, falls unter den m Bildchen jeder Spieler mindestens zweifach vorkommt. Wie groß ist die Wahrscheinlichkeit für dieses Ereignis?

Jedes der m Bildchen zeigt mit Wahrscheinlichkeit $1/n$ einen der n Spieler, und diese Ereignisse sind unabhängig. Daher ist die Wahrscheinlichkeit, dass jeweils k_i-mal Spieler i, $i = 1, \dots, n$, gekauft wurde, gerade durch die Multinomialverteilung mit Erfolgswahrscheinlichkeiten $1/n$ gegeben:

$$\binom{m}{k_1 \; \cdots \; k_n} (1/n)^{k_1} \cdots (1/n)^{k_n} = \binom{m}{k_1 \; \cdots \; k_n} n^{-m}.$$

Es ergibt sich folglich durch Summation

$$\mathbb{P}(A \leqslant m) = n^{-m} \sum_{\substack{k_1 \geqslant 2, \dots, k_n \geqslant 2 \\ k_1 + \cdots + k_n = m}} \binom{m}{k_1 \; \cdots \; k_n},$$

sofern $m \geqslant 2n$, während natürlich $\mathbb{P}(A \leqslant m) = 0$ für $m < 2n$ gilt. Als Erwartunswert erhalten wir durch geschickte Umformung und Benutzung des Gegenereignisses

$$\begin{aligned}
\mathbb{E}[A] &= \sum_{k=0}^{\infty} k \, \mathbb{P}(A = k) \\
&= \sum_{m=0}^{\infty} \mathbb{P}(A \geqslant m + 1) \\
&= \sum_{m=0}^{\infty} \Big(1 - n^{-m} \sum_{\substack{k_1 \geqslant 2, \dots, k_n \geqslant 2 \\ k_1 + \cdots + k_n = m}} \binom{m}{k_1 \; \cdots \; k_n} \Big) \\
&= \sum_{m=0}^{\infty} n^{-m} \sum_{\substack{k_1 \geqslant 0, \dots, k_n \geqslant 0 \\ \exists i \colon k_i \leqslant 1 \\ k_1 + \cdots + k_n = m}} \binom{m}{k_1 \; \cdots \; k_n}.
\end{aligned}$$

Leider sehe ich nicht, wie diese Reihen vereinfacht werden können. Für kleine n ist jedoch alles per Hand berechenbar. Beispielsweise gilt im Fall $n = 2$:

$$\mathbb{E}[A] = 4 + \sum_{m=4}^{\infty} 2^{-m} \sum_{k_1 \in \{0, 1, m-1, m\}} \binom{m}{k_1} = 4 + \sum_{m=4}^{\infty} 2^{-m+1}(1 + m) = 5,5.$$

Wenn nur ein Fan die 2 Spieler sammelt, so kauft er im Schnitt 3 Bildchen, durch Tauschen fällt sein erwarteter Kaufanteil immerhin auf $5,5/2 = 2,75$ Bildchen.

Markus Reiß

Vorlesung
Einführung in die Statistik
Sommersemester 2005

Lösung der Aufgabe 6.4

Annahme: $\varphi : \{0, \ldots, n\} \to [0,1]$ ist gleichmäßig bester Test zum Niveau α von $H_0 : p = 1/2$ gegen $H_1 : p \neq 1/2$.
Dann gilt also $\mathbb{E}_{1/2}[\varphi] \leqslant 0{,}05$ sowie $\mathbb{E}_p[\varphi] \geqslant \mathbb{E}_p[\psi]$ für alle $p \neq 1/2$ und alle Tests ψ zum Niveau α. Setze

$$\varphi_-(x) := \varphi(n-x) \text{ sowie } \varphi_s(x) := \frac{\varphi(x) + \varphi_-(x)}{2}, \quad x \in \{0, \ldots, n\}.$$

Dann ist φ_- ein Test zum Niveau α, da wegen der Symmetrie von $\mathrm{Bin}_{n,1/2}$ gilt $\mathbb{E}_{1/2}[\varphi_-] = \mathbb{E}_{1/2}[\varphi(n - \bullet)]) = \mathbb{E}_{1/2}[\varphi]$. Aus $\mathbb{E}_{1/2}[\varphi_s] = \frac{1}{2}(\mathbb{E}_{1/2}[\varphi] + \mathbb{E}_{1/2}[\varphi_-])$ folgt, dass auch φ_s Niveau α besitzt. Weil φ gleichmäßig bester Test ist, folgt $\mathbb{E}_p[\varphi] \geqslant \mathbb{E}_p[\varphi_-]$ für alle $p \neq 1/2$. Gäbe es nun ein $p \neq 1/2$ mit $\mathbb{E}_p[\varphi] > \mathbb{E}_p[\varphi_-]$, so wäre aus Symmetriegründen $\mathbb{E}_{1-p}[\varphi_-] > \mathbb{E}_{1-p}[\varphi]$, was wegen der Optimalität von φ nicht sein kann. Also gilt $\mathbb{E}_p[\varphi_-] = \mathbb{E}_p[\varphi] = \mathbb{E}_p[\varphi_s]$ für alle $p \in [0,1]$, und φ_s ist ebenfalls ein gleichmäßig bester Test, der zusätzlich eine symmetrische Entscheidungsregel aufweist.

Wähle nun ein $p_1 \in (1/2, 1)$ und bezeichne mit φ_{NP} den gleichmäßig besten Neyman-Pearson-Test zum exakten Niveau α von $H_0 : p = 1/2$ gegen $H_1 : p = p_1$. Wegen des monotonen Likelihood-Quotienten erhalten wir

$$\varphi_{NP}(x) = \begin{cases} 1, & x \geqslant x_\alpha + 1 \\ 0, & x \leqslant x_\alpha - 1 \\ \gamma_\alpha, & x = x_\alpha \end{cases}$$

mit x_α, γ_α geeignet. Weil φ_{NP} nicht symmetrisch ist und exaktes Niveau α besitzt, ist die Menge $A := \{x : \varphi_{NP}(x) > \varphi_s(x)\}$ nicht leer. Folgt man dem Beweis von Teil (i) im Neyman-Pearson-Lemma, so kann man daher auf die strikte Ungleichung $\mathbb{E}_{p_1}[\varphi_{NP}] > \mathbb{E}_{p_1}[\varphi_s]$ schließen – im Widerspruch zur behaupteten Optimalität von φ_s bzw. φ.