Stochastik

Barth, Friedrich

München, [20]03

18. 5. Das Stichprobenmittel

urn:nbn:de:hbz:466:1-83580

1. Die Schätzgröße H_n hat als Stichprobenfunktion die Gestalt

$$H_n = H_n(X_1, X_2, \ldots, X_n) = \frac{1}{n} \sum_{i=1}^{n} X_i .$$

Sie ist also ein spezielles arithmetisches Mittel, nämlich das der sämtlich nach $\mathrm{B}(1; p)$ verteilten Zufallsgrößen X_i, die den Erwartungswert $\mathscr{E} X_i = p$ und Varianz $\operatorname{Var} X_i = pq$ besitzen, wie in **14.5.** gezeigt wurde. Damit erhalten wir nach Satz 212.1

$$\mathscr{E} H_n = p .$$

Da wir p schätzen, ist also H_n erwartungstreu bezüglich p.
2. Zum Nachweis der Konsistenz von H_n schätzen wir den zu untersuchenden Grenzwert mit Hilfe der *Bienaymé-Tschebyschow*-Ungleichung ab. Für die in dieser Ungleichung auftretende $\operatorname{Var} H_n$ gilt nach Satz 212.2: $\operatorname{Var} H_n = \frac{pq}{n}$. Damit erhalten wir

$$0 \leqq \lim_{n \to \infty} P(|H_n - p| \geqq \varepsilon) \leqq \lim_{n \to \infty} \frac{pq}{n\varepsilon^2} = 0 ,$$

was zu zeigen war.
Ohne Beweis teilen wir mit, daß H_n auch effizient ist.

18.5. Das Stichprobenmittel

Eine Zufallsgröße X besitze den Erwartungswert μ. Zu seiner Schätzung zieht man aus X eine Stichprobe $(X_1 | X_2 | \ldots | X_n)$ und bildet als Stichprobenfunktion das arithmetische Mittel $\bar{X}_n = \bar{X}_n(X_1, X_2, \ldots, X_n) = \frac{1}{n} \sum_{i=1}^{n} X_i$, das in diesem Zusammenhang auch »Stichprobenmittel« heißt. Da die X_i die gleiche Verteilung wie X haben, lassen sich wieder die Sätze 212.1 und 212.2 anwenden, und wir erhalten $\mathscr{E} \bar{X}_n = \mu$ und

$$0 \leqq \lim_{n \to \infty} P(|\bar{X}_n - \mu| \geqq \varepsilon) \leqq \lim_{n \to \infty} \frac{\operatorname{Var} \bar{X}_n}{\varepsilon^2} = \lim_{n \to \infty} \frac{\operatorname{Var} X}{n\varepsilon^2} = 0 .$$

Ohne Beweis teilen wir mit, daß das Stichprobenmittel auch effizient ist, und halten fest

> **Satz 380.1:** Ist $(X_1 | X_2 | \ldots | X_n)$ eine Stichprobe aus der Zufallsgröße X, dann ist das **Stichprobenmittel** $\bar{X}_n := \frac{1}{n} \sum_{i=1}^{n} X_i$ eine erwartungstreue, konsistente und effiziente Schätzgröße für den Erwartungswert μ der Verteilung von X.

Satz 380.1 beinhaltet auch die Erkenntnisse aus **18.4.**; denn die relative Häufigkeit H_n ist ein spezielles Stichprobenmittel und schätzt den Erwartungswert p der nach $\mathrm{B}(1; p)$ verteilten Zufallsgröße X.

Auf Grund der bei den Beweisen verwendeten Sätze 212.1 und 212.2 gilt die Erwartungstreue von $\bar{X}_n$ auch für Stichproben, bei denen die X_i nicht unabhängig sind, und die Konsistenz von $\bar{X}_n$ für Stichproben, bei denen die X_i nur paarweise unabhängig sind.

18.6. Die Stichprobenvarianz

Wir wollen nun eine Schätzgröße für die Varianz σ^2 einer Zufallsgröße X ausfindig machen. Wenn möglich, soll sie erwartungstreu sein. Dazu ziehen wir aus X eine Stichprobe $(X_1|X_2|\ldots|X_n)$. Die X_i sind Kopien von X; also ist $\mathscr{E}X_i = \mathscr{E}X = \mu$ und $\sigma^2 = \operatorname{Var}X = \operatorname{Var}X_i = \mathscr{E}[(X_i-\mu)^2]$.
Da wir in **18.5.** unbekannte Erwartungswerte durch das Stichprobenmittel geschätzt haben, ist es naheliegend, als Schätzfunktion für σ^2 das Stichprobenmittel $V := \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$ der Zufallsgrößen $(X_i-\mu)^2$ zu verwenden. Wie in Aufgabe 384/**4** gezeigt werden soll, ist V eine erwartungstreue Schätzgröße für σ^2. Darüber hinaus ist sie konsistent und effizient.
In den meisten Fällen ist aber auch μ nicht bekannt. Wir schätzen dann μ gemäß **18.5.** durch das Stichprobenmittel $\bar{X}_n$, das wir an Stelle von μ in den Ausdruck für V einführen. Wir erhalten somit als Schätzfunktion für σ^2 die Stichprobenfunktion

$$\tilde{S}^2 := \frac{1}{n}[(X_1-\bar{X}_n)^2+\ldots+(X_n-\bar{X}_n)^2] = \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X}_n)^2.$$

Zur Vereinfachung der Schreibweise läßt man üblicherweise beim Stichprobenmittel $\bar{X}_n$ den Index n weg und schreibt

$$\tilde{S}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2.$$

Um zu prüfen, ob $\tilde{S}^2$ eine erwartungstreue Schätzgröße ist, formen wir die Summe um. Wir gehen dabei genauso vor wie bei der Herleitung des Verschiebungssatzes 207.1:

$$\begin{aligned}
\tilde{S}^2 &= \frac{1}{n}\sum_{i=1}^{n}[(X_i-\mu)-(\bar{X}-\mu)]^2 = \\
&= \frac{1}{n}\sum_{i=1}^{n}[(X_i-\mu)^2+(\bar{X}-\mu)^2-2(\bar{X}-\mu)(X_i-\mu)] = \\
&= \frac{1}{n}\left[\sum_{i=1}^{n}(X_i-\mu)^2+\sum_{i=1}^{n}(\bar{X}-\mu)^2-2(\bar{X}-\mu)\sum_{i=1}^{n}(X_i-\mu)\right] = \\
&= \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2+\frac{1}{n}\cdot n(\bar{X}-\mu)^2-\frac{2}{n}\cdot(\bar{X}-\mu)(n\bar{X}-n\mu) = \\
&= \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2+(\bar{X}-\mu)^2-2(\bar{X}-\mu)^2 = \\
&= \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2-(\bar{X}-\mu)^2.
\end{aligned}$$