

EIN KRITERIUM FÜR DIE ANWENDBARKEIT EINES NÄHERUNGSVERFAHRENS BEI DER BESTIMMUNG VON KONFIDENZINTERVALLEN

von StD G. Bung, Gymnasium zum Altenforst, 5210 Troisdorf

Bei der Bestimmung von Konfidenzintervallen für unbekannte Wahrscheinlichkeiten (analog für Erwartungswerte) sind quadratische Ungleichungen der Art $|r-p| \leq k \cdot \frac{\sigma}{n} \Leftrightarrow (r-p)^2 \leq k^2 \cdot \frac{p(1-p)}{n}$ zu lösen, wobei r für $\frac{X}{n}$ steht. STRICK gibt in [1] eine den Rechenaufwand erheblich verringernde Näherungsmethode an, bei der σ durch $\sqrt{r(1-r)}$ und damit die rechte Seite der letzten Ungleichung durch $k^2 \cdot \frac{r(1-r)}{n}$ ersetzt wird. Als Kriterium für die Brauchbarkeit des Näherungsverfahrens wird die Bedingung $0,3 < p < 0,7$ genannt.

Dieses Kriterium läßt sich erheblich verschärfen, wenn man den Stichprobenumfang n berücksichtigt. Offensichtlich darf $k^2 \cdot \frac{p(1-p)}{n}$ durch $k^2 \cdot \frac{r(1-r)}{n}$ ersetzt werden, wenn der Unterschied

$$|k^2 \cdot \frac{p(1-p)}{n} - k^2 \cdot \frac{r(1-r)}{n}|$$

bezüglich der halben Intervallbreite vernachlässigbar klein gegenüber dem Schätzwert r für das Intervallzentrum ist, d. h. wenn $|p(1-p) - r(1-r)| \ll \frac{r \cdot n}{k^2}$ gilt.

Für $|p(1-p) - r(1-r)|$ gewinnt man nun leicht mehrere Abschätzungen:

- (1) $|p(1-p) - r(1-r)| \leq |p(1-p)| + |r(1-r)| \leq 0,25 + 0,25 = 0,5$
- (2) $|p(1-p) - r(1-r)| = |p^2 - r^2 - r + p| = |p-r + (r-p)(r+p)| = |p-r| \cdot |1+(r+p)| \leq |p-r|$

Diskutiert man $f(p,r) = |p(1-p) - r(1-r)|$ als Funktion zweier Variablen mit Mitteln der Differentialrechnung, so läßt sich weiter zeigen:

(3) $f(p,r) \leq 0,25$

Stochastik in der Schule, Band 4, Heft 1 (1984)

Schon Abschätzung (1) liefert ein sehr brauchbares Kriterium:

$$|p(1-p) - r(1-r)| \ll \frac{r \cdot n}{k^2} \ll \frac{1}{2} \ll \frac{r \cdot n}{k^2} \Leftrightarrow \frac{k^2}{2} \ll r \cdot n$$

Es muß also nur überprüft werden, ob $\frac{k^2}{2}$ gegenüber dem Produkt $r \cdot n$ vernachlässigt werden kann. Dies ist z. B. für $n = 1000$, $r = 0,2$, $k = 2$ der Fall, obwohl $p \notin]0,3; 0,7[$.

Literatur

- [1] STRICK, H. K.: Einführung in die beurteilende Statistik. - Hannover: Schroedel, 1980