Freie Universität Berlin

Wahrscheinlichkeit & Statistik
Wintersemester 2019/20

Prof. Matthias Beck
Hauke Sprink
Institut für Mathematik
AG Diskrete Geometrie
Arnimallee 2

15. Übung Musterlösung

Aufgabe 15.1
Es sei $(X_1, X_2, \ldots, X_n)$ eine Stichprobe eines Bernoulliexperiments mit Parameter p; insbesondere ist $X_1 + X_2 + \cdots + X_n$ binomialverteilt.

a. Zeigen Sie, daß $\overline{X} := \frac{1}{n}(X_1 + X_2 + \cdots + X_n)$ eine erwartungstreue Schätzfunktion für p ist.

b. Zeigen Sie, daß $V(\overline{X}) = \frac{p(1-p)}{n}$ ist.

c. Berechnen Sie $\mathbb{E}\left(\frac{\overline{X}(1-\overline{X})}{n}\right)$ und schließen Sie hieraus, daß $\frac{\overline{X}(1-\overline{X})}{n}$ nicht eine erwartungstreue Schätzfunktion für $V(\overline{X})$ ist.

Lösung:

a.

$$\begin{aligned}\mathbb{E}(\overline{X}) &= \mathbb{E}\left(\frac{1}{n}(X_1 + X_2 + \cdots + X_n)\right) = \frac{1}{n}\left(\mathbb{E}(X_1) + \mathbb{E}(X_2) + \cdots + \mathbb{E}(X_n)\right) \\ &= \frac{np}{n} = p\,.\end{aligned}$$

b.

$$\begin{aligned}V(\overline{X}) &= V\left(\frac{1}{n}(X_1 + X_2 + \cdots + X_n)\right) = \frac{1}{n^2}V(X_1 + X_2 + \cdots + X_n) \\ &= \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}\,.\end{aligned}$$

c. Mit

$$\underbrace{V(\overline{X})}_{\frac{p(1-p)}{n}} = \mathbb{E}(\overline{X}^2) - \underbrace{\mathbb{E}(\overline{X})^2}_{p^2} \implies \mathbb{E}(\overline{X}^2) = \frac{p(1-p)+np^2}{n}$$

berechnen wir

$$\begin{aligned}\mathbb{E}\left(\frac{\overline{X}(1-\overline{X})}{n}\right) &= \frac{1}{n}\left(\mathbb{E}(\overline{X}) - \mathbb{E}(\overline{X}^2)\right) = \frac{1}{n}\left(p - \frac{p(1-p)+np^2}{n}\right) \\ &= \frac{p}{n}\left(\frac{n-(1-p)-np}{n}\right) = \frac{p(1-p)}{n}\left(\frac{n-1}{n}\right) \neq V(\overline{X})\,.\end{aligned}$$

Aufgabe 15.2
Wir betrachten eine geometrische Verteilung mit (unbekanntem) Parameter p. Berechnen Sie eine Maximum-likelihood Schätzfunktion für p, falls die Stichprobe $(X_1, X_2, \ldots, X_n)$ vorliegt. Interpretieren Sie Ihre Antwort.

Lösung: Wir definieren die Maximum-likelihood Funktion

$$f(p) \;=\; \prod_{i=1}^{n}(1-p)^{X_i-1}p \;=\; p^n\,(1-p)^{\sum_{i=1}^n X_i - n}.$$

Diese wird maximal, falls

$$\begin{aligned} f'(p)=0 \iff & np^{n-1}(1-p)^{\sum_{i=1}^n X_i-n} = p^n\left(\sum_{i=1}^n X_i - n\right)(1-p)^{\sum_{i=1}^n X_i-n-1} \\ \iff & n(1-p) = p\left(\sum_{i=1}^n X_i - n\right) \\ \iff & p = \frac{n}{\sum_{i=1}^n X_i}\,. \end{aligned}$$

Der Schätzwert für p ist das Verhätnis von Gesamtanzahl der Erfolge zu Gesamtanzahl der Versuche.

Aufgabe 15.3

Angenommen, eine Zufallsvariable ist normalverteilt mit unbekanntem Erwartungswert μ und bekannter Varianz σ^2. Wir möchten ein Konfidenzinterval für μ finden zu einem vorgegebenen Fehlerniveau a.

a. Sei $a = 1\%$. Wie groß muß eine Stichprobe mindestens sein, daß das Konfidenzintervall eine vorgegebene Länge l hat?

b. Welches Fehlerniveau a sollten wir für eine gegebene Stichprobenanzahl n und Konfidenzintervallänge l ansetzen?

Lösung:

a. Wir suchen ein Intervall der Länge l mit Mittelpunkt $\overline{X}$, sodass die Wahrscheinlichkeit, dass μ in diesem Interval liegt, 99 % ist, d.h. $\mathbb{P}(\mu \in [\,\overline{X} - \frac{l}{2}, \overline{X} + \frac{l}{2}]) = 0,99$.

$$X \sim \mathcal{N}(\mu,\sigma^2) \iff \overline{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right) \iff \underbrace{\left(\frac{(\overline{X}-\mu)\sqrt{n}}{\sigma}\right)}_{=:Y} \sim \mathcal{N}(0,1)\,.$$

Wir bestimmen jetzt r so, dass

$$\mathbb{P}(-r \le Y \le r) = 0,99 \iff \mathbb{P}(Y \le r) = 0,995\,.$$

Damit ist $r = 2,58$. Mit Wahrscheinlichkeit 0,99 ist also

$$\frac{(\overline{X}-\mu)\sqrt{n}}{\sigma} = Y \in [-r,r] \iff \mu \in \left[\overline{X} - \frac{\sigma r}{\sqrt{n}}, \overline{X} + \frac{\sigma r}{\sqrt{n}}\right].$$

Nun können wir durch Vergleich der Intervallgrenzen n bestimmen:

$$\frac{l}{2} = \frac{\sigma r}{\sqrt{n}} \iff n = \left(\frac{2\sigma r}{l}\right)^2.$$

b. Haben wir n und l gegeben, können wir mit der letzten Gleichung berechenen, dass $r = \frac{l\sqrt{n}}{2\sigma}$. Den Wert für a berechnen wir dann über $\mathbb{P}(Y \le r) = 1 - \frac{a}{2}$.

Aufgabe 15.4
Ihre Tante hat ein Käsegeschäft und verkauft u.a. 22-Pfund-Käselaibe. Sie hat gerade eine Lieferung bekommen und eine Stichprobe von 10 Laiben genommen, mit folgendem Ergebnis für die Gewichte (in Pfund):

$$21,50 \qquad 18,95 \qquad 18,55 \qquad 22,35 \qquad 22,90 \qquad 22,20 \qquad 19,40 \qquad 23,10 \qquad 22,84 \qquad 23,34\,.$$

Unter der Annahme, daß die Gewichte normalverteilt sind, berechnen Sie ein 95%-Konfidenzintervall für den Erwartungswert.

Lösung:
Wir bezeichnen mit X_i das Gewicht des i-ten Laibs. Wir schätzen

$$\overline{X} = \frac{1}{10}\sum_{i=1}^{10} X_i = 21,51$$

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{10}\left(X_i - \overline{X}\right)^2} \approx 1,84\,.$$

Nun können analog zu **Aufgabe 15.3** (gegeben a und n, gesucht l) vorgehen: Wir haben $r = 1,96$ so gewählt, dass

$$\mathbb{P}\left(\frac{(\overline{X}-\mu)\sqrt{n}}{s} \le r\right) = 0,975$$

Es ist nun

$$0,95 = \mathbb{P}\left(-r \le \frac{(\overline{X}-\mu)\sqrt{n}}{s} \le r\right) = \mathbb{P}\left(\overline{X} - \frac{rs}{\sqrt{n}} \le \mu \le \overline{X} + \frac{sr}{\sqrt{n}}\right).$$

Damit ist das 95%-Konfidenzintervall für den Erwartungswert $[20,37;\, 22,65]$.