Prof. Vincent Tassion
Laurin Köhler-Schindler

Wahrscheinlichkeit & Statistik

Quiz 9

Onlineabgabe vor Beginn der Übungsstunde:
Montag (02.05.2022) um 16:15 Uhr oder Dienstag (03.05.2022), um 14:15 Uhr

Dieser Quiz beschäftigt sich mit Schätzern und deren Eigenschaften am Beispiel der geometrischen Verteilung.

Weitere Informationen und Instruktionen zur Abgabe unter
https://metaphor.ethz.ch/x/2022/fs/401-0614-00L/

---

1. Eddy schlägt Sophie ein Spiel vor: Er wirft eine Münze solange, bis zum ersten Mal Kopf fällt. Angenommen beim k-ten Wurf fällt das erste Mal Kopf. Wenn $k \geq 3$, bekommt er von Sophie k CHF. Wenn $k \in \{1, 2\}$, bekommt Sophie von ihm k CHF.
Was ist Sophies erwarteter Gewinn, wenn die Münze gezinkt ist (Kopf mit Wahrscheinlichkeit p)?
Tipp: Der Erwartungswert einer Geom(p)-verteilten Zufallsvariable ist $1/p$.

(a) $\frac{2p^2-1}{p}$

Leider nicht.

√ (b) $\frac{6p^2-4p^3-1}{p}$

Richtig!

(c) $\frac{3p^2-2p^3-1}{p}$

Leider nicht.

Sei K die Zufallsvariable, die die Nummer des ersten Wurfs, der Kopf zeigt, bezeichnet. Dann ist Sophies erwarteter Gewinn

$$\sum_{k=1}^{2} k \cdot \mathbb{P}[K = k] - \sum_{k=3}^{\infty} k \cdot \mathbb{P}[K = k] = 2(p + 2p(1-p)) - \underbrace{\sum_{k=1}^{\infty} k \cdot \mathbb{P}[K = k]}_{=1/p}$$

$$= 2(3p - 2p^2) - 1/p = \frac{6p^2 - 4p^3 - 1}{p},$$

wobei wir verwendet haben, dass der Erwartungswert einer Geom(p)-verteilten Zufallvariable gleich $1/p$ ist.

2. Wir betrachten weiterhin das gleiche Spiel. Sophie weiss nicht, ob die Münze gezinkt ist, und bittet Eddy, die Münze *vorab* einige Male zu werfen, bis insgesamt 5 Mal Kopf gefallen ist. Sie beobachtet die Ergebnisse der Münzwürfe

$$0, 1, 0, 0, 1, 1, 0, 0, 0, 1, 0, 1$$

und notiert sich die Daten

$$2, 3, 1, 4, 2$$

also wie lange es jeweils bis zum nächsten Mal Kopf gedauert hat. Sophie betrachtet die Daten als realisierte Werte von Zufallsvariablen $X_1, \dots, X_5$. Welcher Parameterraum Θ und welche Modellfamilie $(\mathbb{P}_\theta)_{\theta\in\Theta}$ eignen sich in diesem Fall?

√ (a) Parameterraum $\Theta = [0, 1]$ und Wahrscheinlichkeitsmasse $(\mathbb{P}_\theta)_{\theta\in\Theta}$, sodass $X_1, \dots, X_5$ unter $\mathbb{P}_\theta$ unabhängig, identisch verteilt sind mit $X_i \sim \mathrm{Geom}(\theta)$.

Richtig! Diese Modellfamilie eigent sich gut.

(b) Parameterraum $\Theta \in (0, +\infty)$ und Wahrscheinlichkeitsmasse $(\mathbb{P}_\theta)_{\theta\in\Theta}$, sodass $X_1, \dots, X_5$ unter $\mathbb{P}_\theta$ unabhängig, identisch verteilt sind mit $X_i \sim \mathrm{Geom}(\theta)$.

Leider nicht. Der Erfolgsparameter einer geometrische Verteilung liegt zwischen 0 und 1. Für Erfolgsparameter > 1 ist die Verteilung nicht definiert.

(c) Parameterraum $\Theta \in (0, +\infty)$ und Wahrscheinlichkeitsmasse $(\mathbb{P}_\theta)_{\theta\in\Theta}$, sodass $X_1, \dots, X_5$ unter $\mathbb{P}_\theta$ unabhängig, identisch verteilt sind mit $X_i \sim \mathrm{Exp}(\theta)$.

Leider nicht. Diese Modellfamilie eigenet sich weniger gut, da wir wissen, dass die Daten ganzzahlig sind, also in $\{1, 2, 3, \dots\}$ liegen. Zufallsvariablen, die exponentialverteilt sind, nehmen hingegen Werte in $\mathbb{R}_+$ an.

3. Sei $\Theta = \mathbb{N}$. Seien $X_1, \dots, X_n$ Zufallsvariablen. Wir betrachten die Modellfamilie $(\mathbb{P}_\theta)_{\theta\in\Theta}$, sodass $X_1, \dots, X_n$ unter $\mathbb{P}_\theta$ unabhängig, identisch verteilt sind mit $X_i \sim \mathrm{Bin}(\theta, 1/2)$. Welche der folgenden Schätzer sind erwartungstreu?
(Mehrere richtige Antworten möglich.)

(a) $T_1 = X_1$

Leider nicht. Wir wissen, dass $\mathrm{E}_\theta[X_1] = \theta \cdot \frac{1}{2}$, da $X_1 \sim \mathrm{Bin}(\theta, 1/2)$. Somit ist der Bias von T_1 gleich $\theta/2 - \theta = -\theta/2$ und T_1 ist nicht erwartungstreu.

√ (b) $T_2 = 2X_1$

Richtig! Wir wissen, dass $\mathrm{E}_\theta[X_1] = \theta \cdot \frac{1}{2}$, da $X_1 \sim \mathrm{Bin}(\theta, 1/2)$. Für T_2 erhält man den Bias $2 \cdot \theta/2 - \theta = 0$, der Schätzer T_2 ist also erwartungstreu.

(c) $T_3 = \frac{1}{n}\sum_{i=1}^n X_i$

Leider nicht. Der Bias von T_3 im Modell $\mathbb{P}_\theta$ ist $-\theta/2$ (mittels Linearität des Erwartungswerts), der Schätzer T_3 ist also nicht erwartungstreu.

√ (d) $T_4 = \frac{2}{n}\sum_{i=1}^n X_i$

Richtig! Der Bias von T_4 im Modell $\mathbb{P}_\theta$ ist 0 (mittels Linearität des Erwartungswerts), der Schätzer T_4 ist also erwartungstreu.

4. Wir betrachten erneut die Modellfamilie $(\mathbb{P}_\theta)_{\theta\in\Theta}$ aus Frage 3. Was ist der mittlere quadratische Schätzfehler $\mathrm{MSE}_\theta[T_1]$ des Schätzers $T_1 = X_1$ im Modell $\mathbb{P}_\theta$?

(a) θ

Leider nicht.

√ (b) $\frac{\theta+\theta^2}{4}$

Richtig!

(c) $\frac{\theta}{4n} + \frac{\theta^2}{4}$

Leider nicht.

(d) $\frac{\theta}{n}$

Leider nicht.

Wir berechnen

$$\mathrm{Var}_\theta[T_1] = \mathrm{Var}_\theta[X_1] = \theta/4,$$

wobei wir verwendet haben, dass die Varianz einer $\mathrm{Bin}(m,p)$-verteilten Zufallsvariable gleich $mp(1-p)$ ist. Analog zu Frage 3 erhält man, dass der Bias des Schätzers T_2 gleich $-\theta/2$ ist und somit erhalten wir

$$\mathrm{MSE}_\theta[T_2] = \theta/4 + (-\theta/2)^2 = \frac{\theta+\theta^2}{4}.$$

5. Wir betrachten erneut die Modellfamilie $(\mathbb{P}_\theta)_{\theta\in\Theta}$ aus Frage 3. Was ist der mittlere quadratische Schätzfehler $\mathrm{MSE}_\theta[T_2]$ des Schätzers $T_2 = 2X_1$ im Modell $\mathbb{P}_\theta$?

(a) $\frac{\theta}{4n} + \frac{\theta^2}{4}$

Leider nicht.

(b) $\frac{\theta+\theta^2}{4}$

Leider nicht.

(c) $\frac{\theta}{n}$

Leider nicht.

√ (d) θ

Richtig!

Wir berechnen

$$\mathrm{Var}_\theta[T_2] = 4\mathrm{Var}_\theta[X_1] = \theta,$$

wobei wir verwendet haben, dass die Varianz einer $\mathrm{Bin}(m,p)$-verteilten Zufallsvariable gleich $mp(1-p)$ ist. Analog zur vorherigen Frage erhält man, dass der Schätzer T_2 erwartungstreu ist, also ist der Bias von T_2 gleich 0 für jedes $\theta \in \Theta$. Somit erhalten wir

$$\mathrm{MSE}_\theta[T_2] = \theta.$$

6. Wir betrachten erneut die Modellfamilie $(\mathbb{P}_\theta)_{\theta\in\Theta}$ aus Frage 3. Was ist der mittlere quadratische Schätzfehler $\mathrm{MSE}_\theta[T_3]$ des Schätzers $T_3 = \frac{1}{n}\sum_{i=1}^n X_i$ im Modell $\mathbb{P}_\theta$?

(a) $\frac{\theta+\theta^2}{4}$

Leider nicht.

(b) θ

Leider nicht.

(c) $\frac{\theta}{n}$

Leider nicht.

√ (d) $\frac{\theta}{4n} + \frac{\theta^2}{4}$

Richtig!

Wir berechnen

$$\mathrm{Var}_\theta[T_3] = (1/n)^2 \sum_{i=1}^n \mathrm{Var}_\theta[X_i] = \frac{\theta}{4n},$$

wobei wir verwendet haben, dass die Varianz einer $\mathrm{Bin}(m,p)$-verteilten Zufallsvariable gleich $mp(1-p)$ ist. Analog zu Frage 3 erhält man, dass der Bias des Schätzers T_3 gleich $-\theta/2$ ist und somit erhalten wir

$$\mathrm{MSE}_\theta[T_3] = \frac{\theta}{4n} + (-\theta/2)^2 = \frac{\theta}{4n} + \frac{\theta^2}{4}.$$

7. Wir betrachten erneut die Modellfamilie $(\mathbb{P}_\theta)_{\theta\in\Theta}$ aus Frage 3. Was ist der mittlere quadratische Schätzfehler $\mathrm{MSE}_\theta[T_4]$ des Schätzers $T_4 = \frac{2}{n}\sum_{i=1}^n X_i$ im Modell $\mathbb{P}_\theta$?

√ (a) $\frac{\theta}{n}$

Richtig!

(b) $\frac{\theta+\theta^2}{4}$

Leider nicht.

(c) $\frac{\theta}{4n} + \frac{\theta^2}{4}$

Leider nicht.

(d) θ

Leider nicht.

Wir berechnen

$$\mathrm{Var}_\theta[T_3] = (2/n)^2 \sum_{i=1}^n \mathrm{Var}_\theta[X_i] = \frac{\theta}{n},$$

wobei wir verwendet haben, dass die Varianz einer $\mathrm{Bin}(m,p)$-verteilten Zufallsvariable gleich $mp(1-p)$ ist. Analog zur vorherigen Frage erhält man, dass der Schätzer T_4 erwartungstreu ist, also ist der Bias von T_2 gleich 0 für jedes $\theta \in \Theta$. Somit erhalten wir

$$\mathrm{MSE}_\theta[T_4] = \frac{\theta}{n}.$$