ulm university universität
uulm

Dr. Markus Kunze
Dipl.-Math. Stefan Roth

WS 2013/14
11.02.2014

Stochastik für WiWi - Klausurvorbereitung

Gesetz der totalen Wahrscheinlichkeit und Satz von Bayes

$(\Omega, \mathcal{F}, \mathbb{P})$ Wahrscheinlichkeitsraum, $E_1, \ldots, E_n \in \mathcal{F}$ Ereignisse mit $E_i \cap E_j = \emptyset$ für alle $i \neq j$ und $\Omega = E_1 \cup \cdots \cup E_n$. Dann gilt für jedes beliebige Ereignis $E \in \mathcal{F}$

1.)

$$\mathbb{P}(E) = \sum_{j=1}^{n} \mathbb{P}(E|E_j) \cdot \mathbb{P}(E_j)$$

2.) Falls $\mathbb{P}(E) > 0$ ist, so gilt außerdem

$$\mathbb{P}(E_i|E) = \frac{\mathbb{P}(E|E_i) \cdot \mathbb{P}(E_i)}{\mathbb{P}(E)} = \frac{\mathbb{P}(E|E_i) \cdot \mathbb{P}(E_i)}{\sum\limits_{j=1}^{n} \mathbb{P}(E|E_j) \cdot \mathbb{P}(E_j)}, \quad i = 1, \ldots, n$$

Beispiel

Eine statistische Auswertung der SWU hat ergeben, dass im Mittel 10 % aller Fahrgäste im öffentlichen Nahverkehr schwarz fahren. 70 % der Schwarzfahrer haben keine Fahrkarte, die restlichen 30 % haben entweder gefälschte oder illegale Fahrscheine. Von den ehrlichen Fahrgästen haben im Schnitt 5 % ihre Fahrkarte vergessen.

(a) Mit welcher Wahrscheinlichkeit kann ein kontrollierter Fahrgast keine Fahrkarte vorzeigen?

(b) Mit welcher Wahrscheinlichkeit ist ein kontrollierter Fahrgast, der keine Fahrkarte vorzeigen kann ein Schwarzfahrer?

Lösung

$E = \{\text{ehrlicher Fahrgast}\}$, $S = \{\text{Schwarzfahrer}\}$, $K = \{\text{Fahrgast kann keinen Fahrschein vorzeigen}\}$

(a) $\mathbb{P}(K) = \mathbb{P}(K|E)\mathbb{P}(E) + \mathbb{P}(K|S)\mathbb{P}(S) = 0.05 \cdot 0.9 + 0.7 \cdot 0.1 = 0.115$

(b) $\mathbb{P}(S|K) = \frac{\mathbb{P}(K|S)\mathbb{P}(S)}{\mathbb{P}(K)} = \frac{0.7 \cdot 0.1}{0.115} \approx 0.61$

Tschebyscheff Ungleichung

Sei $X : \Omega \to \mathbb{R}$ eine Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ mit $\mathbb{E}X^2 < \infty$. Dann gilt für jedes $\varepsilon > 0$

$$\mathbb{P}(|X - \mathbb{E}X| \geq \varepsilon) \leq \frac{\mathsf{Var}(X)}{\varepsilon^2}$$

Varianz, Kovarianz und Korrelationskoeffizient

$X, Y : \Omega \to \mathbb{R}$ Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ mit $\mathbb{E}X^2, \mathbb{E}Y^2 < \infty$. Wichtige Formeln:

1.) $\mathsf{Var}(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2$

2.) $\mathsf{Cov}(X,Y) = \mathbb{E}XY - \mathbb{E}X\mathbb{E}Y$

3.) $\rho(X,Y) = \frac{\mathsf{Cov}(X,Y)}{\sqrt{\mathsf{Var}(X)\cdot\mathsf{Var}(Y)}}$

$X_1, X_2, Y_1, Y_2 : \Omega \to \mathbb{R}$ Zufallsvariablen auf $(\Omega, \mathcal{F}, \mathbb{P})$ mit $\mathbb{E}X_1^2, \mathbb{E}X_2^2, \mathbb{E}Y_1^2, \mathbb{E}Y_2^2 < \infty$ und $\alpha, \beta \in \mathbb{R}$. Rechenregeln für Varianz und Kovarianz:

1.) $\mathsf{Cov}(\alpha X, \beta Y) = \alpha\beta\mathsf{Cov}(X,Y)$, insbesondere $\mathsf{Var}(\alpha X) = \mathsf{Cov}(\alpha X, \alpha X) = \alpha^2\mathsf{Cov}(X,X) = \alpha^2\mathsf{Var}(X)$

2.) $\mathsf{Cov}(X_1 + X_2, Y_1 + Y_2) = \mathsf{Cov}(X_1, Y_1) + \mathsf{Cov}(X_1, Y_2) + \mathsf{Cov}(X_2, Y_1) + \mathsf{Cov}(X_2, Y_2)$

3.) $\mathsf{Var}(X+Y) = \mathsf{Var}(X) + 2\mathsf{Cov}(X,Y) + \mathsf{Var}(Y)$

Falls X und Y unkorreliert sind, d.h. $\mathsf{Cov}(X,Y) = 0$, so gilt außerdem

4.) $\mathsf{Var}(X+Y) = \mathsf{Var}(X) + \mathsf{Var}(Y)$

Wichtig:

$$X \text{ und } Y \text{unabhängig} \;\Rightarrow\; X \text{ und } Y \text{unkorreliert}$$

aber im Allgemeinen **nicht** umgekehrt!

Blatt 5, Aufgabe 2

Es sollen 100.000 € in Aktien investiert werden. Dabei stehen 2 Aktien zur Auswahl: Aktie 1 kostet 80 €, der erwartete Kurs in einem Jahr beträgt 90 €, bei einer Standardabweichung von 2 €. Aktie 2 kostet 120 €, der erwartete Kurs in einem Jahr beträgt 150 €, bei einer Standardabweichung von 10 €. Das Geld soll mit minimalem Risiko investiert werden. Als Risikomaß verwenden wir die Varianz, d.h. wir wollen die 100.000 € so investieren, dass der Wert unseres Portfolios in einem Jahr minimale Varianz hat. Wie ist das Geld zu investieren, wenn

(a) die Kurse in einem Jahr unabhängig sind?

(b) die Kurse in einem Jahr einen Korrelationskoeffizienten von $-0,4$ besitzen?

Zusätzlich wird stets davon ausgegangen, dass Leerverkäufe nicht erlaubt sind und dass die 100.000 € restlos investiert werden.

Lösung

Es bezeichne X den Kurs von Aktie 1 nach einem Jahr und Y den Kurs von Aktie 2 nach einem Jahr. Gesucht sind $\alpha, \beta \geq 0$, so, dass

$$\begin{cases} \alpha \cdot 80 + \beta \cdot 120 = 100000 \\ \mathsf{Var}(\alpha X + \beta Y) \to \min \end{cases}$$

Eine Lösung bekommt man, indem man die erste Bedingung nach α auflöst und danach in die zweite Bedingung einsetzt. Danach verwendet man die obigen Eigenschaften der Kovarianz um die zu minimierende Funktion konkret zu bestimmen. Das Minimum findet man dann wie üblich.

Blatt 6, Aufgabe 1

Die Zufallsvariable X sei gleichverteilt auf der Menge $\{-2, -1, 0, 1, 2\}$. Sei $Y = X^2$.

(a) Bestimme $\mathsf{Cov}(X,Y)$.

(b) Sind X und Y unabhängig? (Die Antwort ist zu begründen!)

Lösung

(a) Es gilt:
$$\mathsf{Cov}(X,Y) = \mathsf{Cov}(X,X^2) = \mathbb{E}X^3 - \mathbb{E}X\mathbb{E}X^2$$
Da die Verteilung von X symmetrisch ist, gilt $\mathbb{E}X = \mathbb{E}X^3 = 0$, also
$$\mathsf{Cov}(X,Y) = 0$$
Also sind X und Y unkorreliert.

$\left[\mathbb{E}X = \frac{1}{5}(-2-1+0+1+2) = 0 \text{ und } \mathbb{E}X^3 = \frac{1}{5}(-8-1+0+1+8) = 0\right]$

(b) X und Y sind aber nicht unabhängig, denn z.B.
$$\mathbb{P}(X=1, Y=1) = \mathbb{P}(X=1, X^2=1) = \mathbb{P}(X=1) = \frac{1}{5}$$
und
$$\mathbb{P}(X=1)\mathbb{P}(Y=1) = \mathbb{P}(X=1)\mathbb{P}(X^2=1) = \mathbb{P}(X=1)\mathbb{P}(X\in\{-1,1\}) = \frac{1}{5}\cdot\frac{2}{5} = \frac{2}{25}$$
also
$$\mathbb{P}(X=1,Y=1) \neq \mathbb{P}(X=1)\mathbb{P}(Y=1)$$

Schätzen von Parametern

$\Theta \subseteq \mathbb{R}^n$ Parameterraum, $\{F_\theta;\ \theta\in\Theta\}$ parametrische Familie von Verteilungen, $X_1,\dots,X_n$ Zufallsstichprobe zur Verteilung F_θ, $\theta = (\theta_1,\dots,\theta_n) \in \Theta$.

Momentenschätzer

Ziel ist es, $g(\theta)$ zu schätzen, wobei $g:\Theta\to\mathbb{R}^m$ eine Funktion des Parametervektors θ ist. Ist
$$g(\theta) = f(m_1(\theta),\dots,m_l(\theta))$$
eine Funktion $f:\mathbb{R}^l\to\mathbb{R}^m$ gewisser Momente von F_θ, so wählt man den Schätzer $\widehat{g(\theta)}$ für $g(\theta)$ als
$$\widehat{g(\theta)} = f(\hat{m}_1,\dots,\hat{m}_l),$$
wobei $\hat{m}_k = \frac{1}{n}\sum_{i=1}^n X_i^k$.

Beispiel

Sei $X_1,\dots,X_n$ eine Zufallsstichprobe zur $\mathsf{LN}_{\mu,\sigma^2}$-Verteilung. Bestimme mit Hilfe der Momentenmethode einen Schätzer für den Parametervektor (μ,σ^2).

Lösung

$\hat{m}_1 = \frac{1}{n}\sum\limits_{j=1}^n x_j$, $\hat{m}_2 = \frac{1}{n}\sum\limits_{j=1}^n x_j^2$, $\mathbb{E}X_1 = e^{\mu+\frac{\sigma^2}{2}}$, $\mathbb{E}X_1^2 = e^{(2\mu+\sigma^2)\sigma^2}$. Betrachte folgendes Gleichungssystem
$$\begin{aligned}\hat{m}_1 &= \mathbb{E}X_1\\ \hat{m}_2 &= \mathbb{E}X_1^2\end{aligned}$$
d.h.
$$\begin{aligned}&\hat{m}_1 = e^{\mu+\frac{\sigma^2}{2}} \Leftrightarrow \mu + \frac{\sigma^2}{2} = \log(\hat{m}_1) \Leftrightarrow \mu = \log(\hat{m}_1) - \frac{\sigma^2}{2}\\ &\hat{m}_2 = e^{(2\mu+\sigma^2)\sigma^2} \Leftrightarrow (2\mu+\sigma^2)\sigma^2 = \log(\hat{m}_2) \Leftrightarrow 2\log(\hat{m}_1)\sigma^2 = \log(\hat{m}_2) \Leftrightarrow \sigma^2 = \frac{\log(\hat{m}_2)}{2\log(\hat{m}_1)}\end{aligned}$$

Daraus ergibt sich der Momentenschätzer

$$(\hat{\mu}, \hat{\sigma}^2) = (\log(\hat{m}_1) - \frac{\log(\hat{m}_2)}{4\log(\hat{m}_1)}, \frac{\log(\hat{m}_2)}{2\log(\hat{m}_1)})$$

Maximum-Likelihood-Schätzer

Die Likelihood-Funktion $L : \mathbb{R}^n \times \Theta \to \mathbb{R}$ von F_θ ist definiert durch

$$L(x_1, \dots, x_n; \theta) = \begin{cases} \prod\limits_{k=1}^{n} \mathbb{P}(X_1 = x_k) & ; \; F_\theta \text{ diskret} \\ \prod\limits_{k=1}^{n} f(x_k) & ; \; F_\theta \text{ absolutstetig mit Dichte } f \end{cases}$$

Der Maximum-Likelihood-Schätzer $\hat{\theta}$ für θ ist derjenige Parametervektor, der die Likelihood-Funktion für jeden Datenvektor $(x_1, \dots, x_n)$ maximiert. Oft ist es einfacher die Log-Likelihood-Funktion $\mathrm{logL}(x_1, \dots, x_n; \theta)$ definiert durch

$$\mathrm{logL}(x_1, \dots, x_n; \theta) = \log(L(x_1, \dots, x_n; \theta))$$

zu betrachten.

Beispiel

Sei $X_1, \dots, X_n$ eine Zufallsstichprobe zur $\mathsf{LN}_{\mu,\sigma^2}$-Verteilung. Bestimme mit Hilfe der Maximum-Likelihood-Methode einen Schätzer für den Parametervektor (μ, σ^2), wobei die Dichte f der $\mathsf{LN}_{\mu,\sigma^2}$-Verteilung gegeben ist durch

$$f(x) = \frac{1}{\sigma x \sqrt{2\pi}} e^{-\frac{(\log(x)-\mu)^2}{2\sigma^2}}, \quad x > 0$$

Lösung

$$L(x_1, \dots, x_n; \mu, \sigma^2) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \cdot \frac{1}{x_1 \cdots x_n} e^{-\sum\limits_{j=1}^{n} \frac{(\log(x_j)-\mu)^2}{2\sigma^2}}$$

$$\mathrm{logL}(x_1, \dots, x_n; \mu, \sigma^2) = -n\log(\sqrt{2\pi\sigma^2}) - \sum_{j=1}^{n} \log(x_j) - \sum_{j=1}^{n} \frac{(\log(x_j) - \mu)^2}{2\sigma^2}$$

$$\frac{\partial}{\partial\mu}\mathrm{logL} = \sum_{j=1}^{n} \frac{\log(x_j) - \mu}{\sigma^2} = 0 \Leftrightarrow \mu = \frac{1}{n}\sum_{j=1}^{n} \log(x_j)$$

$$\frac{\partial}{\partial\sigma^2}\mathrm{logL} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{j=1}^{n}(\log(x_j) - \mu)^2 = 0 \Leftrightarrow \sigma^2 = \frac{1}{n}\sum_{j=1}^{n}(\log(x_j) - \mu)^2$$

Der ML-Schätzer für (μ, σ^2) ist also gegeben durch

$$(\hat{\mu}, \hat{\sigma}^2) = (\frac{1}{n}\sum_{j=1}^{n} \log(X_j), \frac{1}{n}\sum_{j=1}^{n}(\log(X_j) - \frac{1}{n}\sum_{j=1}^{n} \log(X_j))^2)$$

Güteeigenschaften von Schätzern

Sei $T_n(X_1, \dots, X_n)$ ein Schätzer für $g(\theta)$.

1.) Erwartungstreue: $\mathbb{E}T_n(X_1, \dots, X_n) = g(\theta)$.

2.) Asymptotische Erwartungstreue: $\lim\limits_{n\to\infty} \mathbb{E}T_n(X_1, \dots, X_n) = g(\theta)$.

3.) Starke Konsistenz: $\lim_{n\to\infty} T_n(X_1, \ldots, X_n) = g(\theta)$.

4.) Schwache Konsistenz: $\lim_{n\to\infty} \mathbb{P}(|T_n(X_1, \ldots, X_n) - g(\theta)| > \varepsilon) = 0$, für jedes $\varepsilon > 0$.

Wichtig: 1.) $\Rightarrow$ 2.) und 3.) $\Rightarrow$ 4.) Im Allgemeinen aber **nicht** andersherum!