Empirisches Mittel: Funktion von Messwerten

$$\bar{\mu}_n(x_1,\dots,x_n) = \frac{1}{n}\sum_{i=1}^{n} x_i \quad , x_i \in \mathbb{R}$$

als Funktion von Zufallsvariablen:

$$\bar{\mu}_n(X_1,\dots,X_n) = \frac{1}{n}\sum_{i=1}^{n} X_i$$

entspricht dem empirischen Mittel von Zufallsvariablen, vgl. z.B. Gesetz der großen Zahlen.

d.h. $\bar{\mu}_n(X_1,\dots,X_n)$ ist eine Zufallsvariable.

aus dem Gesetz der großen Zahlen wissen wir: Falls die X_i unabh. und identisch verteilt sind, dann konvergiert

$$\bar{\mu}_n(X_1,\dots,X_n) \xrightarrow{n\to\infty} \mathbb{E}[X_1]$$

Statistisches Modell:

Mögliche Parameter:

$$\vartheta \in \Theta \quad , \quad \underbrace{\Theta \subseteq \mathbb{R}}_{=\text{ Menge der möglichen Parameter}}$$

für jeden Wert $\vartheta \in \Theta$ existiert ein Wahrscheinlichkeitsmaß $\mathbb{P}_\vartheta$.

Die Familie $(\Omega, (\mathbb{P}_\vartheta)_{\vartheta\in\Theta})$ heißt statistisches Modell.

Bsp. Gegeben Daten $x_1,\dots,x_n$ von denen man annehmen kann, die zugrundeliegenden ZV seien geometrisch verteilt.

Der Parameter p ist unbekannt.

Dann ist $\Theta = [0,1]$ die Menge aller möglichen Parameter.

Mit Hilfe einer Schätzfunktion kann einer dieser Werte berechnet werden, welcher der Schätzer ist.

Möglicher Schätzer: $X \sim Geo(p)$, $\mathbb{E}[X] = \frac{1}{p}$

$\bar{\mu}_n(x_1,\dots,x_n) = \frac{1}{n}\sum_{i=1}^{n} x_i$ ist ein (möglicher) Schätzer für $\mathbb{E}[X]$, somit kann

$$\frac{1}{\frac{1}{n}\sum_{i=1}^{n} x_i} = \frac{n}{\sum_{i=1}^{n} x_i}$$

als Schätzer für p verwendet werden.

Eigenschaften von Schätzern

1) empirisches Mittel:

$$\bar{\mu}_n(x) = \frac{1}{n}\sum_{i=1}^{n} x_i$$

ist konsistent wg. Gesetz der großen Zahlen

erwartungstreu:

$$\mathbb{E}[\bar{\mu}_n(X_1,\dots,X_n)] = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[X_i] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[X_1] = \mathbb{E}[X_1]$$

die X_i identisch vert.

empirisches Mittel ist also erwartungstreu.

Empirische Varianz:

$$\bar{\sigma}_n^2(x_1,\dots,x_n) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{\mu}_n(x))^2$$

$$\mathbb{E}\left[\bar{\sigma}_n^2(X_1,\dots,X_n)\right] = \mathbb{E}\left[\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \frac{1}{n}\sum_{j=1}^{n}X_j\right)^2\right]$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\mathbb{E}\left[X_i^2 - \frac{2}{n}X_i\sum_{j=1}^{n}X_j + \frac{1}{n^2}\left(\sum_{j=1}^{n}X_j\right)^2\right]$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\Bigg(\underbrace{\mathbb{E}[X_i^2]}_{=\mathbb{E}[X_1^2]} - \frac{2}{n}\sum_{j=1}^{n}\mathbb{E}[X_iX_j] + \frac{1}{n^2}\sum_{j,k}\mathbb{E}[X_jX_k]\Bigg)$$

$$\mathbb{E}[X_iX_j] \overset{\text{unabh.}}{\underset{\text{falls } j\neq i}{=}} \mathbb{E}[X_i]\cdot\mathbb{E}[X_j] \overset{\text{id. vert.}}{=} \mathbb{E}[X_1]^2$$

$$\underset{i=j}{=} \mathbb{E}[X_i^2] = \mathbb{E}[X_1^2]$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\Bigg(\mathbb{E}[X_1^2] - \frac{2}{n}(n-1)\mathbb{E}[X_1]^2 - \frac{2}{n}\mathbb{E}[X_1^2] + \frac{1}{n^2}\left(n(n-1)\mathbb{E}[X_1]^2 + n\,\mathbb{E}[X_1^2]\right)\Bigg)$$

$= \dots$ (Termumformungen)

$= \dots = \mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2$

$= \underline{\underline{\mathbb{V}(X_1)}}$

empirische Varianz ist ein erwartungstreuer Schätzer für die Varianz.

Der Ausdruck

$$\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{\mu}_n(x))^2$$

ist ebenfalls ein Schätzer für die Varianz, aber nicht erwartungstreu.

beide sind konsistent.

Vorsicht: nicht jeder erwartungstreue Schätzer ist ein "guter" Schätzer:

Bsp. $(x_1,\dots,x_n)$

$$\vartheta(x_1,\dots,x_n) = x_1$$

das ist ein erwartungstreuer Schätzer für den Erwartungswert

$$\mathbb{E}[\vartheta(X_1,\dots,X_n)] = \mathbb{E}[X_1]$$

Für jedes mögliche $\vartheta \in \Theta$ kann ich

$$\boxed{\mathbb{P}_\vartheta(X_1 = x_1,\dots,X_n = x_n)}$$

berechnen, wobei $X_1,\dots,X_n$ Zufallsvariablen mit Verteilung definiert durch ϑ sind, und die $x_1,\dots,x_n$ die Messwerte.

Def. p_ϑ sei eine diskrete Verteilung, also $p_\vartheta(x) = \mathbb{P}_\vartheta(X = x)$

Bsp. $(x_1,\dots,x_n)$ seien Realisierungen von unabh. $Poi(\lambda)$ Zufallsvariablen, $\lambda > 0$

$\vartheta = \lambda$, $\quad \Theta =]0,\infty[$

$$p_\lambda(k) = \frac{e^{-\lambda}\lambda^k}{k!} = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x\in\mathbb{N}_0$$

Möchte

$$L((x_1,\dots,x_n);\underset{\vartheta}{\lambda}) = p_\lambda(x_1)\cdot\ldots\cdot p_\lambda(x_n)$$

$$= \frac{e^{-\lambda}\lambda^{x_1}}{x_1!}\cdot\ldots\cdot\frac{e^{-\lambda}\lambda^{x_n}}{x_n!}$$

$$= \boxed{e^{-n\cdot\lambda}\cdot\frac{\lambda^{x_1+x_2+\dots+x_n}}{x_1!\cdot\ldots\cdot x_n!}}$$

diese Funktion L möchte man – als Funktion von λ – maximieren,

also z.B. $L((x_1,\dots,x_n),\lambda)$ nach λ ableiten und 0 setzen

Schwierigkeit: Ableitung nach λ ist etwas rechenaufwändig

Trick:

$$\ell((x_1,\dots,x_n);\lambda) = \log L((x_1,\dots,x_n),\lambda)$$

ℓ hat sein Maximum an derselben Stelle wie L, und ist leichter abzuleiten

$$\ell((x_1,\dots,x_n),\lambda) = \log\left(e^{-n\lambda}\cdot\frac{\lambda^{x_1+\dots+x_n}}{x_1!\cdot\ldots\cdot x_n!}\right)$$

$$= -n\cdot\lambda\underbrace{\log(e)}_{=1} + (x_1+\dots+x_n)\cdot\log\lambda \underbrace{- \log(x_1!\cdot\ldots\cdot x_n!)}_{\text{verschwindet beim Ableiten nach }\lambda}$$

Ableiten nach λ:

$$\frac{d\ell}{d\lambda} = -n + (x_1+\dots+x_n)\frac{1}{\lambda}$$

null setzen:

$$0 = -n + (x_1+\dots+x_n)\cdot\frac{1}{\lambda}$$

auflösen nach λ:

$$\underline{\underline{\lambda = \frac{x_1+\dots+x_n}{n}}} \quad \Big| \text{ Maximum likelihood-Schätzer}$$