Resampling

in »Statistische Methoden in der Physik«

Referent: Alex Ortner

Studenten-Seminar Sommersemester 2007

Gliederung

1. Resampling Prinzip
   - Einleitung
   - Resampling Methoden
2. Bootstrap Methoden
   - Bootstrap: Punktschätzung
   - Bootstrap: Konfidenzintervalle
   - Bootstrap: Tests
3. Permutationstest
4. Ausblick und weiterführende Fragen

Einleitung

intuitv

Was ist Resampling?

Prinzip: Stichprobenziehung aus Stichprobe

Warum macht man so etwas?

Zweck: Simulation der Variabilität der Daten

Wenn man für viele solcher Stichproben die gewünschte Statistik berechnet, erhält man die (simulierte) Verteilung der Statistik

$\Downarrow$

Schätzung von Standardabweichung, Varianz, Verzerrung usw. der Statistik

Einleitung
formal

Problematik:

F unbekannt $\Rightarrow t(F)$ unbekannt

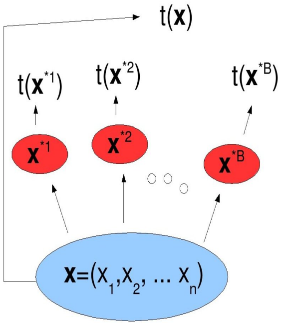
Einzige Möglichkeit:

Schätzung über endliche Stichprobe

1. Schätzung der Verteilung $\hat{F}$
2. Schätzung der Statistiken $\hat{t}(F) = t(\hat{F})$

Wie gut ist Schätzung?, Wie ist Verteilung von $\hat{t}$? $\Rightarrow$Resampling

1. Ziehen von B Resamples aus Stichprobe: $\mathbf{x}^{*1}$, $\mathbf{x}^{*2}$, ... $\mathbf{x}^{*B}$
2. Testgröße $\hat{t^{*i}}$aus Samples bestimmen
3. Vergleichen und bewerten



Resampling Methoden

Permutationstest

$$\mathbf{x}^* = (\mathbf{y}, \mathbf{z})^* = ((y_1, z_2), (y_2, z_3), \ldots, (y_n, z_{n-1}))$$

Jackknife

$$\mathbf{x}^* = (x_1, x_2, \ldots x_{i-1}, x_{i+1}, \ldots x_n)$$

Bootstrap

$$\mathbf{x}^* = (x_1^*, x_2^*, \ldots, x_n^*) = \underbrace{(x_4, x_4, x_1, \ldots, x_3)}_{n-\,Werte}$$

Subsampling

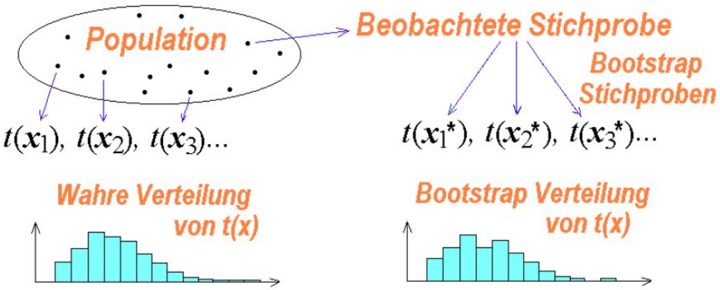
$$\mathbf{x}^* = (x_1^*, x_2^*, \ldots, x_m^*) = \underbrace{(x_4, x_4, x_1, \ldots, x_3)}_{m-\,Werte \qquad m<n}$$

Bootstrap: Punktschätzung

Die Grundidee

Der Name: Boostrap=Stiefelschlaufe

Aus Märchen »der Lügenbaron von Münchausen«

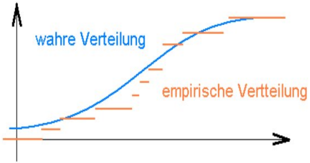


Bootstrap: Punktschätzung

Der »Plug-in« Schätzer

- $\mathbf{x}$ Stichprobe aus F verteilter Menge
- Prüfgröße $t(F)$ gesucht
- $\hat{F}_n$ empirische Verteilung von $\mathbf{x}$

$$\hat{F}_n(x) = \sum_{x_i < x} \frac{1}{n}$$



«plug-in principle«

$t(\hat{F}_n)$ ist die »Plug-in«-Schätzung von $t(F)$

Bootstrap: Punktschätzung

Die Bootstrap Verteilung

Schätzen der Verteilung von $t(F)$

Real World	**Bootstrap World**
$\mathbf{x}$ Stichprobe aus F	$\mathbf{x}^*$ Stichprobe aus $\hat{F}_n$
$\hat{F}_n$ empirische Verteilung basierend auf $\mathbf{x}$	$\hat{F}_n^*$ empirische Verteilung basierend auf $\mathbf{x}^*$
$\hat{t} = t(\hat{F}_n)$ beobachteter Wert der Prüfgröße	$\hat{t}^* = t(\hat{F}^*{}_n)$ boostrap Wert der Prüfgröße
$\hat{Q}(t)$ zu schätzende Verteilung der Prüfgröße	$\hat{Q}(\hat{t}^*)$bootstrap Verteilung der Prüfgröße

$\Longrightarrow$ Schätzung $Q(t)$ durch empirische Verteilung von $\hat{t}^* = (\hat{t}^{*1}, \hat{t}^{*2}, \ldots, \hat{t}^{*B})$

Bootstrap: Punktschätzung
Bootstrap Schätzung

Standartfehler des Schätzers $\hat{t}$

Mit Standardabweichung der Bootstrapverteilung

$$\overline{\hat{t}(B)}^* = \frac{1}{B}\sum_{b=1}^{B}\hat{t}(b)^*$$

$$se(\hat{t}) = \sqrt{var(\hat{t})} = \left\{\frac{1}{B-1}\sum(\hat{t}(b)^* - \overline{\hat{t}(B)}^*)^2\right\}^{\frac{1}{2}}$$

Beispiel: Erwartungswert $E(F)$ mit Mittelwert $\overline{x}$ geschätzt

$$\hat{t} = \overline{x} \quad \Rightarrow se(\overline{x}) = \left\{\tfrac{1}{B-1}\sum(\overline{x}(b)^* - \overline{\overline{x}(B)}^*)^2\right\}^{\frac{1}{2}},$$

Bootstrap: Punktschätzung
Komplizierte Datenstrukturen

Bis jetzt betrachtet: »One-Sample« Problem

$F \longrightarrow \mathbf{x} = (x_1, x_2, \ldots, x_n)$ x_i beliebig dimensionale Zufallsvariablen

Beliebig viele Samples

$P = (F, G, \ldots) \longrightarrow \mathbf{x} = (\mathbf{y}, \mathbf{z}, \ldots)$
Wobei $F \longrightarrow \mathbf{y} = (\mathrm{y}_1, \mathrm{y}_2, \ldots, \mathrm{y}_\mathrm{m})$ unabhängig von $G \longrightarrow \mathbf{z} = (z_1, z_2, \ldots, z_n)$

Beispiel: »Two-Sample« Problem

Prüfgröße $t = E(F) - E(G)$
Schätzung über $\hat{t} = \hat{\mu}_y - \hat{\mu}_z = \bar{z} - \bar{y}$
Schätzung $se(\hat{t})$ mit Bootstrap, wobei $\mathbf{x}^* = (\mathbf{y}^*, \mathbf{z}^*) = (y_{i1}, y_{i2}, \ldots, y_{im}, z_{j1}, z_{j2}, \ldots, z_{jn})$

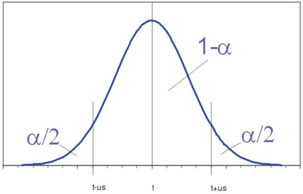
Bootstrap: Konfidenzintervalle

Konfidenzintervall

Konfidenzintervall=Vertrauensbereich

Bereich um geschätzten Parameter $\hat{t}$ in dem t mit Wahrscheinlichkeit $p = 1 - \alpha$ liegt

- $Q(\hat{t})$ asymptotisch normalverteilt
- $1 - \alpha = P(t_- < t < t_+) = \int_{t_-}^{t_+} \hat{q}(t)dt \Longrightarrow t_-, t_+$
- klas. Konfidenzintervall zum Niveau $1 - \alpha$
  $[\hat{t} - u(1 - \frac{\alpha}{2})\sqrt{var(\hat{t})}\ ,\ \hat{t} + u(1 - \frac{\alpha}{2})\sqrt{var(\hat{t})}]$

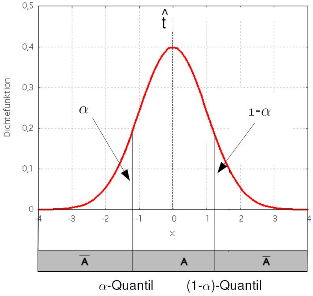


Exkurs:
Quantil und Perzentil

Quantil, Perzentil

Ein α-Quantil ist ein Wert, der die Messwerte so aufteilt, dass der α Anteil der Daten kleiner ist und der $1-\alpha$ Anteil der Daten größer

- $\Phi(t_-) = P(t \leq t_-) = \alpha$
- $t \sim N(\mu, \sigma^2)\ \mu = \hat{t}$
- $\Rightarrow \frac{t-\hat{t}}{\sqrt{var(\hat{t})}} \sim N(0,1)$
- $\Phi(t_-) = \Phi(\frac{t_- - \hat{t}}{\sqrt{var(\hat{t})}}) = \alpha$
- mit $u = \Phi^{-1}$ folgt
- $t_- = \hat{t} + u(\alpha)\sqrt{var(\hat{t})}$



Bootstrap: Konfidenzintervalle

Probleme bei klas. Konfidenzintervallen

Problem:

- Häufig $var(\hat{t})$ nicht bekannt $\Longrightarrow$ Schätzung $\hat{var}(\hat{t})$
- $\hat{var}(\hat{t})$ schlechter Schätzer
- $Q(\hat{t})$ nicht Normalverteit (schlechte Approximation)

Lösung:

Konstruktion der Konfindenzintervalle für t mit Bootstrap-Verfahren

Bootstrap: Konfidenzintervalle

1. Standard-Bootstrap-Intervall

Anwendung:

$var(\hat{t})$ schlecht, $Q(\hat{t})$ annähernd normalverteilt

Methode

$var(\hat{t})$ wird aus simulierter Bootstrap-Verteilung von $\mathbf{t}^*$ geschätzt:
$\hat{var}(\hat{t}) = var(\hat{t}^*)$

Standart-Bootstrap-Intervall

$$I_{Standart} = [\hat{t} - u(1-\frac{\alpha}{2})\sqrt{var(\hat{t}^*)}\ ,\ \hat{t} + u(1-\frac{\alpha}{2})\sqrt{var(\hat{t}^*)}]$$

Bootstrap: Konfidenzintervalle

2. Bootstrap-t-Intervall

Anwendung:

$Q(\hat{t})$ nicht normalverteilt, $var(\hat{t})$ gut

Methode

Es werden die Quantile $u \sim \tau^*$ aus der Bootstrap Verteilung $\hat{Q}(\hat{t}^*)$ verwendet

Ansatz: $\frac{\hat{t}^* - \hat{t}}{\sqrt{var(\hat{t}^*)}}$

Bootstrap-t-Intervall

$$I_{Boot-t} = [\hat{t} - \tau^*(1 - \frac{\alpha}{2})\sqrt{var(\hat{t})}\ ,\ \hat{t} + \tau^*(1 - \frac{\alpha}{2})\sqrt{var(\hat{t})}]$$

Bootstrap: Konfidenzintervalle

3. Bootstrap-Perzentil-Intervall

Anwendung:

$Q(\hat{t})$ nicht normalverteilt, $var(\hat{t})$ schlecht

Methode

verwendet $\tau^*(\alpha)$ und $\tau^*(1-\alpha)$ Quantil der Bootstrapverteilung von t^*

Bootstrap-Perzentil-Intervall

$$I_{Perz} = [\tau^*(\alpha)\ ,\ \tau^*(1-\alpha)]$$

Bootstrap: Tests

Möglichkeiten Bootstrap für Hypothesentests zu verwenden

Intuitiv: «Schätzer basteln«

für Testgrößen t benötigte Schätzer mit Bootstrapverfahren schätzen
(z.B. T-Test, F-Test)

Annahme, Ablehnebereiche berechnen

Bestimmung von Konfidenzintervalle mit Bootstrap $\Longrightarrow$ Herleitung von Tests

Direkt mit Bootstrap testen

Bootstrap Samples aus einer Verteilung die in H_0 liegt

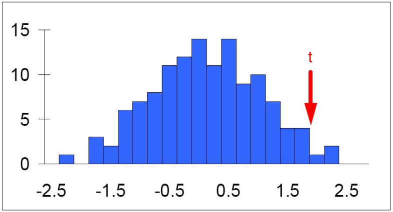
Bootstrap: Tests
logisches Prinzip

- Vergleich von zwei Verteilungen $F \to \mathbf{y}$, $G \to \mathbf{z}$
  - $H_0 : \quad F = G$
- benötigen Testgröße $t(\mathbf{x}) := \overline{y} - \overline{z}$
  $\mathbf{x} = (y_1, y_2, ..., y_n, z_1, z_2, ...z_m) = (x_1, x_2, ..., x_{n+m})$
- erzeugen $\mathbf{x}^* = (\underbrace{x_1^*, x_2^*, ..., x_n^*}_{\mathbf{y}^*}, \underbrace{x_{n+1}^*, x_{n+2}^*, ..., x_{n+m}^*}_{\mathbf{z}^*})$
- schätzen Verteilung $\hat{Q}(t)$ über $t(\mathbf{x}^{*b}) = \overline{\mathbf{y}}^* - \overline{\mathbf{z}}^*$

Bootstrap: Tests
Prinzip

- Nullverteilung der Prüfgröße $t(\mathbf{x}^{*b}) = \bar{\mathbf{y}}^* - \bar{\mathbf{z}}^*$

- jedes t^* gleich wahrscheinlich
- $p(t^*) = \frac{1}{B}$



geschätzter Fehler 1. Art

$P(H_0 \textit{ abgelehnt} \mid \textit{obwohl } H_0 \textit{ richtig}) = P(t^* \geq t) = \frac{\mathbf{1}}{\mathbf{B}}\#\{\mathbf{t}^{*\mathbf{b}} \geq \mathbf{t}\} = \alpha$

Permutationstest
Prinzip

- **Prinzip wie bei Bootstraptest**
- Vergleich von zwei Verteilungen $F \rightarrow \mathbf{y}$, $G \rightarrow \mathbf{z}$
- $H_0 : \quad F = G$
- benötigen Testgröße $t(\mathbf{x}) := \overline{y} - \overline{z}$
  $\mathbf{x} = (y_1, y_2, ..., y_n, z_1, z_2, ...z_m) = (x_1, x_2, ..., x_{n+m})$
- andere Samples $\mathbf{x}^* = (\underbrace{y_2, y_3, ..., z_1}_{\mathbf{y}^*}, \underbrace{z_2, z_3, ..., y_1}_{\mathbf{z}^*})$

Fehler 1. Art

$P(t^* \geq t) = \frac{1}{\mathbf{B}} \#\{\mathbf{t}^{*\mathbf{b}} \geq \mathbf{t}\} = \alpha$

Permutationstest

Vergleich mit Bootstrap Test

Anzahl der Permutationen

- alle $(n+m)!$, ohne Doppelte $\frac{(n+m)!}{(n-m)!m!}$
- oftmals sehr groß $\Longrightarrow$Simulation mit zufälliger Auswahl

Bootstrap Test v.s. Permutationstest

Permutationstest:

- H_0: immer Gleichheit zweier Verteilungen
- bedingter Test
- exakt, wenn Austauschbarkeit besteht

Boostrap Test:

- H_0: allgemeiner (Mittelwert, Median)
- kein bedingter Test
- nicht exakt

Ausblick und weiterführende Fragen

Ausblick

- vom Sampling Prinzip zum Randomierungsprinzip (Monte-Carlo)
- parametrisches Bootstrap
- Block Bootstrap
- Regressions Modelle mit Bootstrap
- Konsistenz des Bootstraps
- Spezielle Boostrap/Permutationsmethoden für Spezielle Probleme
- Was kann man über Abhängige Verteilungen/Korrelationen mit Bootstrap aussagen?

Noch Fragen?

Danke

Ausblick und weiterführende Fragen

Ausblick

- vom Sampling Prinzip zum Randomierungsprinzip (Monte-Carlo)
- parametrisches Bootstrap
- Block Bootstrap
- Regressions Modelle mit Bootstrap
- Konsistenz des Bootstraps
- Spezielle Boostrap/Permutationsmethoden für Spezielle Probleme
- Was kann man über Abhängige Verteilungen/Korrelationen mit Bootstrap aussagen?

Noch Fragen?

Danke

Quellangaben

- B.Efron,R.Tibshirani (1993): An Introduction to the Bootstrap, Chapmann&Hall, London
- B.Efron (1985): The jackknife, the bootstrap and other resampling plans, Chapmann&Hall, London
- P.Good (2000): Permutation tests, Springer, Heidelberg
- P.Good (1999): Resampling methods, Birkhäuser, Boston
- B.Winkler (1998): Bootstrap-Methoden bei nichtparametrischer Regression, Dis. Uni München
- K.Bosch (1996): Großes Lehrbuch der Statistik, Oldenburg Verlag, München
- www.univet.hu/users/jreiczig/dortmund2005/ [Stand: 28.06.07]
- www.wikipedia.de [Stand: 28.06.07]