Dr. Katharina Best
Sommersemester 2011
30. Juni 2011

Reader – Teil 10: Resampling-Methoden

WiMa-Praktikum

In der multivariaten Analyse ist die Angabe von Konfidenzintervallen oder Hypothesentests für den Vektor der Erwartungswerte sowie die Kovarianzmatrix zu sehr vielen Verfahren wichtig. Die klassischen Verfahren setzen dabei die multivariate Normalverteilung voraus, da ansonsten die asymptotischen Verteilung sehr kompliziert und oft unberechenbar werden.

Resampling-Methoden ermöglichen es, asymptotische Aussagen unter nicht sehr restriktiven Bedingungen zu machen.[1] Allen gemeinsam ist, dass sie nicht alleingestellt zu sehen sind, sondern in Verbindung mit anderen Verfahren angewandt werden.

1 Bootstrap

Der Bootstrap bietet die Möglichkeit, die Verteilung einer Teststatistik mit einfachen Mitteln zu erhalten. Die Grundidee, ausgehend von unabhängigen und identisch Stichprobenelementen, dabei ist, die empirische Verteilungsfunktion der Stichprobe als Ausgangspunkt der Untersuchung zu nehmen. Dazu betrachte $X = (X_1, \dots, X_n)$ die Zufallsstichprobe und $x = (x_1, \dots, x_n)$ ihre Realisation. Sei $\theta = \theta(F)$ der zu schätzende Parameter der Verteilungsfunktion F und $\hat{\theta} = s(X)$ der Schätzer für θ.

Bootstrap-Stichprobe Wir gehen jetzt von der erhaltenen Realisation $x = (x_1, \dots, x_n)$ aus und bezeichnen mit $\hat{F}$ ihre empirische Funktion. Im Anschluss ziehen wir n mal mit Zurücklegen aus $(x_1, \dots, x_n)$ und bekomme $x^* = (x_1^*, \dots, x_n^*)$ die Bootstrap-Stichprobe. Diese besteht aus Elementen von x, wobei einige mehrmals, andere vielleicht gar nicht vorkommen. Zu dieser können wir $\hat{\theta}^* = s(x^*)$ berechnen, *Bootstrap-Kopie* von $\hat{\theta}$ genannt.

Bootstrap-Algorithmus Der Bootstrap-Algorithmus wiederholt das Ziehen der Bootsrap-Stichprobe und die Berechnung der Bootstrap-Kopie. Im Anschluss werden diese analysiert:

1. Erhalte m unabhängige Bootstrap-Stichproben $x^{*1}, \dots, x^{*m}$, jeweils der Größe n durch Ziehen mit Zurücklegen aus den Elementen der Stichprobe x.

[1] Die hier exemplarisch vorgestellten Methoden sind historisch in umgekehrter Reihenfolge.

2. Berechne $\hat{\theta}^{*i}$ für alle Bootstrap-Stichproben, $1 \leq i \leq n$.

3. Schätze die interessierende Eigenschaft von $\hat{\theta}$ aus $\hat{\theta}^{*1}, \ldots, \hat{\theta}^{*m}$.

Beispiel: Varianzschätzung Interessieren wir uns beispielsweise für die Varianz des Schätzers, so berechnen wir die empirische Standardabweichung der Bootstrap-Kopien des Schätzers mit Hilfe von $\overline{\hat{\theta}^*} = \frac{1}{m} \sum_{i=1}^{m} \hat{\theta}^{*i}$ und

$$\widehat{\sigma}(\theta)_m := \left[\frac{1}{m-1} \sum_{i=1}^{m} \left(\hat{\theta}^{*i} - \overline{\hat{\theta}^*} \right) \right]^{1/2}. \tag{1}$$

Es kann gezeigt werden, dass

$$\lim_{m \to \infty} \widehat{\sigma}(\theta)_m = \widehat{\sigma}(\hat{\theta}), \tag{2}$$

wir also mit steigendem m gegen die Standardabweichung konvergieren. Daraus ergibt sich die Frage nach einer sinnvollen Anzahl von Bootstrap-Kopien. Der Rechenaufwand ist linear und erfahrungsgemäß sind $50 - 100$ Wiederholungen zur Berechnung der Standardabweichung bzw. der Varianz genug.

Beispiel: Konfidenzintervall Wollen wir aus den Bootstrap-Kopien ein Konfidenzintervall für den Schätzer angeben, so werden wesentlich mehr Wiederholungen benötigt.

2 Jackknife

Beim *Jackknife* handelt es sich um ein kontrolliertes Resampling aus der Stichprobe x. Haben wir die Stichprobe der Länge n vorliegen, so werden wir genau n Stichproben der Länge $n-1$ untersuchen und zwar

$$x^{*i} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \tag{3}$$

für alle $1 \leq i \leq n$, d. h. genau die i-te Beobachtung wird weggelassen. Aus diesen werden dann wieder $\hat{\theta}^{*i}$ berechnet.

Schätzung der Bias Um die Bias eines Schätzers zu evaluieren betrachtet

$$\overline{\hat{\theta}^*} = \frac{1}{n} \sum_{i=1}^{n} \hat{\theta}^{*i} \tag{4}$$

$$(\widehat{\text{Bias}}\, \theta)_{jack} = (n-1) \left(\overline{\hat{\theta}^*} - \hat{\theta}, \right) \tag{5}$$

wobei $\hat{\theta}$ der gewünschte Schätzer aus der Originalstichprobe ist. Es kann gezeigt werden, dass die Wahl des Vorfaktors in der Größenordnung n, auch *Inflationsfaktor* genannt, diese Schätzung mit Bootstrap vergleichbar macht. Der Grund liegt in der großen Ähnlichkeit, jeder einzelnen Jackknife-Stichprobe zu der Originalstichprobe

Gerade bei kleinen Stichproben ist die Jackknife-Methode effizienter zu rechnen. Verwendung findet die Methode vor allem zur Schätzung der Bias und der Varianz eines Schätzers. Sie ist oft leichter und schneller zu rechnen als Bootstrap, erlaubt jedoch nur verschiedene Varianzschätzungen anstatt einer Schätzung der Verteilung.

3 Permutationstest

Permutationstests gehen auf R. A. Fisher zurück und waren mehr eine theoretische Argumentationslinie als ein praktisches Vorgehen in der damaligen Zeit. Den ersten Permutationstest, den exakten Test nach Fisher, haben wir schon kennengelernt. Die Hauptanwendung ist vor allem die Betrachtung der Nullhypothese der Gleichheit der Erwartungswerte. Haben wir die Stichprobe $x = (x_1, \ldots, x_n)$ gegeben und untersuchen wir die Teststatistik T, so berechnen wir den p-Wert des Permutationstest als

$$p = \sum_{z:T(z)>T(x)} \mathbb{P}(z), \tag{6}$$

wobei wir hier von einem einseitige Test ausgehen. Für einen zweiseitigen Test verwenden wir die Teststatistik $\tilde{T} = |T|$.

Letztendlich werden alle oder zufällige Permutation (daher der Name) der möglichen Ausgänge unter der Nullhypothese angesehen, das sind die Stichproben x^* und ihre Wahrscheinlichkeiten berechnet. Danach wird sortiert.

Der Unterschied zum Bootstrap ist, dass wir hier Ziehen mit Zurücklegen haben. Die Anzahl der Stichproben m kann durch die Parameter (Gesamtanzahl, Aufteilungen,. . .) vorgegeben sein, wie beim exakten Test nach Fisher.