MÜNCHENER STOCHASTIK-TAGE 1998 Sekt.-Nr.:2

Asymptotische Statistik u.a. Nichtparametrische und Semiparametrische Verfahren

Sektionsleiter: M. Nussbaum, Berlin

Drees, Holger
Mathematisches Institut, Universität zu Köln

Asymptotische Minimax-Schranken in der Extremwerttheorie

Hall und Welsh (1984) und Drees (1997) haben für gewisse nicht-parametrische Familien von Verteilungen die optimale Rate bestimmt, mit der eine beliebige Folge von Schätzern für den Extremwertindex von i.i.d. Zufallsvariablen gleichmässig gegen den wahren Wert konvergieren kann. In diesem Vortrag werden lokale Approximationen von geeigneten Teilexperimenten hergeleitet, die es ermoeglichen, untere Schranken für den asymptotischen Minimax-Fehler einer beliebigen Schätzerfolge zu bestimmen.

Dümbgen, Lutz
Medizinische Universität zu Lübeck

Optimale nichtparametrische Tests von qualitativen Annahmen

Angenommen man beobachtet einen stochastischen Prozess Y auf dem Einheitsintervall, wobei $dY = f_o + n^{-1/2}dW$ mit einer unbekannten Funktion f_o, gegebenem Skalenparameter $n \geq 1$ ("Stichprobenumfang") und einer Brownschen Bewegung W. Es werden zwei Klassen von Tests von qualitativen Hypothesen über f_o wie Monotonie oder Konvexität vorgeschlagen. Diese Tests sind asymptotisch optimal und adaptiv bezüglich zweier Kriterien. Modifikationen für nichtparametrische Regression werden kurz beschrieben. Als Nebenprodukt erhalten wir eine Verallgemeinerung des Lévyschen Stetigkeitsmoduls der Brownschen Bewegung. Diese hat potentielle Anwendungen auf simultane Konfidenzintervalle in der nichtparametrischen Kurvenschätzung. (gemeinsames Projekt mit Vladimir G. Spokoiny, WIAS Berlin).

Dyckerhoff, Rainer
Seminar für Wirtschafts- und Sozialstatistik, Universität zu Köln

Multivariate rank tests based on different notions of data depth

Data depth is a concept that measures the "depth" of a point y in a given cloud of data points $x_1, x_2, \ldots, x_n$ in $\mathbb{R}^d$. Any sensible notion of data depth should be affine-invariant, which means that the data depth is independent of the chosen coordinate system. Many such notions of data depth have been proposed by, e.g., Mahalanobis (1936), Tukey (1975), Liu (1990), Singh (1991), Koshevoy and Mosler (1997).
The concept of data depth can be used to construct multivariate rank tests. The problem in generalizing rank tests such as the Wilcoxon test to the multivariate case is the absence of a natural linear order in d-space. Thus, there is no natural ranking of points. If the assignment of ranks is made with respect to the depth of the points, distribution free test statistics can be obtained.
We describe such a rank test which can be employed to detect scale differences between two distributions. We show that under the null hypothesis the distribution of the test statistic is the same as the distribution of the test statistic of the usual Wilcoxon test, which makes it easy to calculate critical values for this test.

While the null distribution of the test statistic is the same for all depths, the power of the test strongly depends on which depth is used to assign the ranks. We simulate the power of these rank tests for the Mahalanobis depth, the halfspace depth and the zonoid depth on different alternatives. Finally, the power functions are compared with the power function of a parametric test for the normal distribution. The comparisons show that the rank tests are nearly as good as the parametric test when the underlying distribution is multivariate normal, while under deviations from the assumption of a normal distribution the rank tests are in general much better than the parametric test.

Ferger, Dietmar
TU Dresden, Institut für Mathematische Stochastik

Optimale Tests für das allgemeine Zweistichprobenproblem

Es seien $X_1, \ldots, X_m$ und $Y_1, \ldots, Y_n$ unabhängige Zufallsvariablen mit Werten in einem meßbaren Raum $(\mathcal{X}, \mathcal{F})$ und gemeinsamen unbekannten Verteilungen $Q_X = \mathcal{L}(X_i)$, $1 \leq i \leq m$, bzw. $Q_Y = \mathcal{L}(Y_j)$, $1 \leq j \leq m$. Das Zweistichprobenproblem besteht darin, die Hypothese $H_0 : Q_X = Q_Y$ gegen die Alternative $H_1 : Q_X \neq Q_Y$ zu testen. Wir motivieren für diesen Zweck einen Test, der als Prüfgröße eine verallgemeinerte U-Statistik mit antisymmetrischem Kern ψ verwendet. Eine ausführliche asymptotische Untersuchung des Tests umfaat insbesondere das Verhalten unter lokalen Alternativen und die Bestimmung der zugehörigen lokalen Güte. Diese läßt sich in einem sehr allgemeinen Rahmen bzgl. des Kerns ψ maximieren. Die Maximalstelle ψ_{opt} ist explizit bekannt und von überraschend einfacher Gestalt. Es zeigt sich, daß beispielsweise für reellwertige Daten ($\mathcal{X} = \mathbb{R}$) sowohl im einfachen Lokations- als auch im einfachen Skalenmodell die lokale Güte unseres Tests zum optimalen Kern ψ_{opt} übereinstimmt mit der des optimalen Rangtests. Da jedoch an den Stichprobenraum $\mathcal{X}$ keinerlei einschränkende Bedingungen gestellt werden, läßt sich unser Verfahren - im Gegensatz zu z. B. Rangverfahren - nicht nur auf univariate ($\mathcal{X} = \mathbb{R}$) sondern auch auf multivariate ($\mathcal{X} = \mathbb{R}^d$) oder auf funktionale Daten ($\mathcal{X} = L_2[0, 1]$) anwenden. Wir demonstrieren dies anhand vieler Beispiele.

Gaffke, Norbert
Otto-von-Guericke-Universität Magdeburg

Testen der Unkonfundiertheit von multivariat normalverteilten Variablen

Betrachtet werden die linearen Regressionsmodelle

$$\begin{aligned} \mathrm{E}\big(\, Y \mid X = x, W = w \,\big) &= \beta_0 + \beta_1' x + \beta_2' w \ , \\ \mathrm{E}\big(\, Y \mid X = x \,\big) &= \gamma_0 + \gamma_1' x \ , \end{aligned}$$

wobei Y eine reelle Zufallsvariable und X und W mehrdimensionale Zufallsvariablen sind, die gemeinsam einer multivariaten Normalverteilung genügen. Man sagt, daß X *unkonfundiert* mit W bezüglich Y ist, wenn $\beta_1 = \gamma_1$. Im statistischen Modell mit i.i.d. Beobachtungen (Y_i, X_i, W_i), $i = 1, \ldots, n$, und unbekanntem Erwartungswert und unbekannter (positiv definiter) Kovarianzmatrix stellt die Nullhypothese

$$H_0 \ : \ \beta_1 = \gamma_1$$

eine nicht-lineare Hypothese über die Kovarianzmatrix dar. Die Teststatistik des *Wald Tests* ist bekanntlich unter der Nullhypothese asymptotisch (für $n \to \infty$) χ^2 verteilt, sofern die Jacobi-Matrix der Funktion $\beta_1 - \gamma_1$ in allen Punkten der Nullhypothese vollen Rang hat. Es stellt sich aber heraus, daß 'singuläre Punkte' der Nullhypothese existieren, d.h. solche mit *nicht* vollem Rang der Jacobi-Matrix. Die asymptotische Verteilung der Wald-Statistik in singulären Punkten wird untersucht und insbesondere die Frage, ob der Wald-Test asymptotisch konservativ ist.

Grama , Ion
Institute of Mathematics, Chişinău, Moldova

Asymptotic equivalence for some nonparametric regression models with random design

Assume that X is a set of real numbers and the parameter set Θ is a real interval. Let $\{p(x,\theta),\theta\in\Theta\}$ be a family of densities on the measurable space $(X,\mathcal{X},\mu)$ endowed with a σ-finite measure μ. We consider a sequence of two dimensional independent identically distributed observations $(\xi_1,X_1),...,(\xi_n,X_n)$, where $\xi_1,...,\xi_n$ is an independent sequence of covariates with a common density $\rho(t)$, $t\in[0,1]$, and $X_1,...,X_n$ is the corresponding sequence of responses in X, such that the conditional μ-density of X_i given $\xi_i=t$ is $p(x,f(t))$, $t\in[0,1]$, for $i=1,...,n$. The unknown function $f(t):[0,1]\to\Theta$ is assumed to be in the Hölder smoothness class with exponent greater than $1/2$. We show that, under some natural regularity conditions, this model is asymptotically equivalent in the sense of the Le Cam deficiency pseudodistance to the Gaussian model, where we observe $Y_i=\Gamma(f(t_i))+\varepsilon_i$, $i=1,...,n$, with deterministic design points $t_i=\int_0^{i/n}\rho(s)ds$, $i=1,...,n$, chosen according to the density of the covariates $\xi_1,...,\xi_n$, and with independent identically distributed standard normal random variables $\varepsilon_1,...,\varepsilon_n$. The function $\Gamma(\theta)$ is related to Fisher's information $I(\theta)$ corresponding to the density $p(x,\theta)$ in the following way: $\Gamma'(\theta)=I(\theta)^{1/2}$. Our approximation is in the global sense. We give also a local variant of the asymptotic equivalence result, which involves only the Fisher information. Applications to the nonparametric linear regression model and generalized linear regression model with random design will be considered.

Hesse, Christian
Mathematisches Institut A, Universität Stuttgart

Dichteschätzung aus fehlerbehafteten Beobachtungen

Wir betrachten das Problem der Dichteschätzung aus unabhängigen und identisch verteilten Beobachtungen, die mit additiven Messfehlern kontaminiert sind. Unter milden Voraussetzungen an die zu schätzende Dichte wird ein adaptives Verfahren entwickelt, das sich an die unbekannte Glattheit der Dichtefunktion anpasst und asymptotisch optimal ist für ISE und MISE.

Jähnisch , Michael
Weierstrass Institut Berlin

Asymptotische Äquivalenz für ein Modell mit unabhängigen, nicht identisch verteilten Daten

Wir betrachten das folgende nichtparametrische Modell:

$$Y_1,...,Y_n\in[0,1]$$

seien unabhängige Daten mit den Dichten:

$$Y_i\sim h(i/n,.)\qquad h\in\Sigma$$

Dabei ist Σ enthalten in einer Hölderklasse von Funktionen h, mit $h\geq\epsilon>0$ für ein festes $\epsilon>0$. Das Ziel ist es, hierfür ein begleitendes Experiment zu finden, d.h. eine Folge von Gaußschen Experimenten für die der Le Camsche Δ-Abstand gegen 0 konvergiert.
Dabei wird Gebrauch gemacht von einer funktionalen Version der starken Approximation des empirischen Folgenprozesses (d.h. des empirischen Prozesses aufgefasst als Prozess von 2 Parametern) durch den Kiefer Prozess.

Kovac, Arne
Universität-GH Essen, University of Bristol

Vernachlässigen der Annahmen von Wavelet-Thresholding

Wir betrachten das Problem der nicht-parametrischen Regression. Gegeben sind Beobachtungen $y_1, \ldots, y_n$ zu Zeitpunkten $t_1, \ldots, t_n$, die modelliert werden als

$$y_i = 3Df(t_i) + \varepsilon_i.$$

Hierbei ist f die Funktion, die geschätzt werden soll, und $\varepsilon_1, \ldots, \varepsilon_n$ Rauschen. Ein noch relativ neues Verfahren ist Wavelet-Thresholding, das zu Beginn der 90er Jahre von Donoho und Johnstone vorgeschlagen wurde (siehe [1]) und zur Zeit Gegenstand vieler Untersuchungen ist.
Ein entscheidender Nachteil von Wavelet-Thresholding besteht darin, daß eine Reihe von Annahmen getroffen werden, die in der Praxis oft nicht erfüllt sind. Insbesondere müssen die Anzahl der Zeitpunkte eine Zweierpotenz und die Zeitpunkte selbst äquidistant sein. Ferner wird angenommen, daß das Rauschen normalverteilt und unkorreliert oder zumindestens stationär (siehe zum Beispiel [2]) ist.
Als Ergebnis einer gemeinsamen Arbeit mit Bernard Silverman werden Möglichkeiten präsentiert, wie man diese Restriktionen umgehen kann. Die nicht notwendigerweise äquidistanten Daten werden dazu durch lineare Interpolation auf ein äquidistantes Gitter transformiert, wobei ein neuer Datensatz entsteht, der im allgemeinen eine nicht-stationäre Kovarianzmatrix besitzt. Zentraler Bestandteil unseres Verfahrens ist ein $O(n)$-Algorithmus, mit dem Varianzen von Wavelet-Koeffizienten bestimmt werden können, wenn die Kovarianzmatrix der Daten bekannt ist.
Dieser Algorithmus wird dazu benutzt, um für jeden Wavelet-Koeffizienten einen geeigneten Threshold zu berechnen. Durch Invertieren der Wavelet-Transformation wird dann wie üblich das Signal f an den Gitterpunkten geschätzt. Daten, bei denen die Varianz nicht konstant ist, können ebenso behandelt werden wie korrelierte Daten, bei denen die Korrelationsmatrix bekannt ist. Schließlich untersuchen wir noch Anwendungen auf Datensätze mit Ausreißern und verallgemeinerte lineare Modelle.

Literatur

[1] Donoho, D. L., Johnstone, I. M., Kerkyacharian, G., and Picard, D. (1995).

[2] Johnstone, I. M. and Silverman, B. W. (1997).

Löwendick , Martin
Universität - GHS Essen / SFB 475

Ein schnelles Verfahren zur Berechnung von kubischen Splines mit vorgegebenen Schranken und Monotonie-Eigenschaften

Wir betrachten das Modell

$$y_i = f(x_i) + \varepsilon_i,$$

mit unabhängig identisch verteilten Rauschen ε_i mit $\mathtt{MED}(\varepsilon_i) = 0$. In [1] wird ein Verfahren angegeben, welches zu einem gegebenen Datensatz $\{(x_1, y_1), \ldots, (x_n, y_n)\}$ Extrema und Monotonie-Eigenschaften ermittelt, die eine adäquate Regressionsfunktion $\tilde{f}$ erfüllen sollte. Das Verfahren liefert Schranken $l_i \leq f(x_i) \leq u_i$, die minimale Anzahl von nötigen lokalen Extremwerten sowie Approximationsbereiche für die Extremstellen. In [2] wird ein iteratives Verfahren angegeben, welches unter diesen Nebenbedingungen die glatteste Regressionsfunktion f bestimmt.

Man kann Konvergenz nachweisen, jedoch ist die Konvergenzgeschwindigkeit in verschiedenen Situationen unbefriedigend klein. In bestimmten Fällen sind mehrere tausend Iterationen nötig, um der Lösung nahe zu kommen. Typischerweise sind die Veränder ungen pro Iteration sehr klein, so daß auch ein vernünftiges Abbruchkriterium ein Problem darstellt.
Weil man weiß, daß zwischen zwei aktiven Nebenbedingungen die Lösung des Problems ein kubischer Spline ist, können wir nun ein erheblich schnelleres Verfahren angeben. Die neue Hybridmethode basiert auf iterativem Hinzufügen und Wegnehmen von Knoten und anschließender Kontrolle auf korrekte Wahl der Knoten mit einem Schritt des Verfahrens aus [2]. Beispiele künstlicher und auch realer Datensätze verdeutlichen den Geschwindigkeitsgewinn. Ein weiterer Vorteil der neuen Methode ist, daß in vielen Fällen nicht nur eine Approximation, sondern die glatteste Lösung selber das Ergebnis ist. Es gibt Beispiele, in denen das Verfahren nach [2] selbst nach mehreren tausend Iterat ionen noch sichtbar von der exakten Lösung nach der neuen Methode abweicht.

Literatur

[1] Davies, P.L. (1994): *Data Features* , Statistica Neerlandica 49, Nr. 2, Seiten 185-245

[2] Metzner, L. (1997): *Facettierte nichtparametrische Regression* , Doktorarbeit, Uni GHS Essen.

Marohn , Frank
Katholische Universität Eichstätt

Über Grenzexperimente von trankierten Modellen

Trankierte Modelle treten beispielsweise in der Extremwertstatistik und in der Regressionsanalyse auf. Trankierung bedeutet, daß man nur solche Beobachtungen betrachtet, die in eine gewisse Menge D fallen. Wir untersuchen für den unabhängigen Fall die Frage, welche Grenzexperimente auftreten knnen. Bekannt ist, daß unter gewissen Annahmen lokale asymptotische Normalität (LAN) gilt, d.h. die Limesexperimente sind Gaua–Shifts (Falk (1997), Marohn (1997)). Es zeigt sich, daa aber auch Poisson–Experimente auftreten knnen. Wir untersuchen, welche Invarianzeigenschaften (Translationsinvarianz, Skaleninvarianz) diese besitzen. Dabei verwenden wir die von Strasser (1985) entwickelte Theorie lokalisierter Produktexperimente.
Literatur:

Falk, M. (1997): Local asymptotic normality of truncated empirical processes, Ann. Statist., erscheint.
Marohn, F. (1997): Local asymptotic normality of truncation models, Preprint Katholische Universität Eichstätt.
Strasser, H. (1985): Scale invariance of statistical experiments, Probab. Math. Statist. 5, 1-20.

Munk, Axel
Fakultät für Mathematik, Ruhr Universität Bochum

Comparsion of regression curves

A new test is proposed for the comparison of several regression curves $f_1, \cdots, f_k$. We prove an asymptotic normal law under fixed alternatives which can be applied for power calculations, for constructing confidence regions and for testing precise hypotheses of a weighted L^2 distance between these curves. In particular, the problem of nonequal sample sizes is treated which is related to a peculiar formula of the area between two step functions. These results are extended in various directions such as the optimal allocation of the sample sizes or to a random predictor.

Spokoiny , Vladimir **Hauptvortrag**
Humboldt-Universität Berlin

On estimating a dynamic function of stochastic system with averaging

We consider a two-scaled diffusion system, when drift and diffusion parameters of a "slow" component are contaminated by an unobservable " fast" one. The goal is to estimate the dynamic function which is defined by averaging the drift coefficient of the "slow" component w.r.t. the stationary distribution of the "fast" one. For estimation we use a locally linear smoother with a data-driven choice of bandwidth. A procedure proposed is fully adaptive and nearly optimal up to a *log log* factor.

Steland, Ansgar
Europa-Universität Frankfurt (Oder)

Analyse ordinaler Strukturen - Der Rangmodellierungsansatz (RMA)

Die Interpretation insbesondere ökonomischer Daten erfolgt häufig ordinal; man denke nur an die laufend veröffentlichten Rankinglisten der Volkswirtschaften hinsichtlich ihrer Wirtschaftskraft. Somit erscheint es sinnvoll, semiparametrische Regressionsmodelle zur ordinalen Datenanalyse zu studieren.
Seien $(Y, \mathbf{x}), (Y_1, \mathbf{x}_1), \ldots, (Y_n, \mathbf{x}_n) \overset{i.i.d.}{\sim} F_{(Y,\mathbf{X})}(y, \mathbf{x})$. Ist $H(y) = P(Y \leq y)$, so heißt $H(Y_i)$ *asymptotische Rangtransformation.* Ein nichtparametrischer Schätzer ist durch $\widehat{H}(Y_i) = R_i/n$ gegeben, wobei $\widehat{H}$ die empirische Verteilungsfunktion der $\{Y_i\}$ bezeichnet und R_i der Rang von Y_i unter $Y_1, \ldots, Y_n$ ist. Um Rankings zu modellieren liegt es also nahe, ein Regressionsmodell der Form

$$E[H(Y)|\mathbf{X} = \mathbf{x}] = g(\mathbf{x}'\beta_0)$$

zu betrachten, wobei g eine monotone Bijektion und β_0 der wahre unbekannte Parametervektor ist. Durch Wahl spezieller Linkfunktionen g können lineare Modelle in die Klasse der RMA-Modelle eingebettet werden. Somit bieten sich RMA-Modelle auch zur Robustifizierung des linearen Modells an. Zur Schätzung von β_0 wird ein M-Schätzer $\widehat{\beta}_n$ vorgeschlagen, der sich unter Regularitätsvoraussetzungen als asymptotisch normalverteilt erweist.
Bei der Analyse von Überlebenszeiten kann die Größe $R(Y) = 1 - H(Y)$ als *Risikomaß* eines Individuums mit Überlebenszeit Y interpretiert werden, so daß die Anwendung von RMA-Modellen sinnvoll erscheint.Setzt man zur Cox-Regression eine multiplikative Zerlegung des Risikos in Risikofaktoren,

$$E[R(Y)|\mathbf{X} = \mathbf{x}] = e^{\beta_0} \prod_{j=1}^{p} e^{\beta_j \cdot x_j}.$$

Vermöge der Resultate von Stute (1996) können die asymptotischen Ergebnisse auf den Fall rechtszensierter Daten ausgeweitet werden. Die Ergebnisse von Simulationsstudien zur Untersuchung der Güte der Verfahren werden vorgestellt.
Ferner soll eine erste Verallgemeinerung auf Modelle mit variierenden Koeffizienten vorgestellt werden. Hier ist $\beta_0 = \beta_0(t)$ eine Funktion der Zeit, die durch eine lokale Kernschätzung basierend auf *lokalen Rängen* geschätzt werden kann.

Stute, W. (1996). Distributional convergence under random censorship when covariables are present. *Scand. J. Statist. 23*, 461-471.

Zwanzig , Silvelyn
Universität Hamburg, Institut für Mathematische Stochastik

Orthogonalreihen-Schätzer im funktionellen Fehler-in-Variablen Modell

Das funktionelle Fehler-in-Variablen Modell kann durch zwei Modell/-gleichungen beschrieben werden. Die erste entspricht der des nicht/-parametrischen Regressions/-modells mit festen, nichtzufälligen Versuchspunkten. Die einzelnen Versuchspunkte (Variablen) sind jedoch nicht bekannt und können nur mit einem zufälligen Fehler beobachtet werden; dies wird in der zweiten Gleichung formuliert. Es wird allerdings angenommen, daß der asymptotische Versuchsplan bekannt ist. Die Versuchspunkte werden als Nebenparameter aufgefasst, deren Anzahl gleich dem Stichprobenumfang ist. Der interessierende "Parameter" ist die nichtparametrisierte Regressions/-funktion.
Für die Regressionsfunktion werden Orthogonalreihen-Schätzer vorgeschlagen und die Optimalität der Konvergenz-Raten wird diskutiert.
Die Orthogonalität bezieht sich auf den asymptotischen Versuchsplan. Im Unterschied zu den Orthogonalreihen-Schätzern in der nichtparametrischen Regressionstheorie werden hier die Werte der orthogonalen Basisfunktionen an den unbekannten Versuchspunkten erwartungstreu geschätzt. Die Konstruktion der erwartungstreuen Schätzer führt zu Integralgleichungen analog zum Dekonvolution - Problem.
Für verschiedene Verteilungen der Fehler in den Variablen lassen sich solche erwartungstreuen Schätzer angeben. Die Konsistenzraten hängen wesentlich von diesen Fehlerverteilungen ab. Im Falle der Fourierreihen-Schätzer sind die Konsistenzraten dieselben wie bei den Kernschätzern von Fan und Truong (1993) im strukturellen Modell. Für die Polynomialreihen-Schätzer ergeben sich andere, nicht vergleichbare Raten.