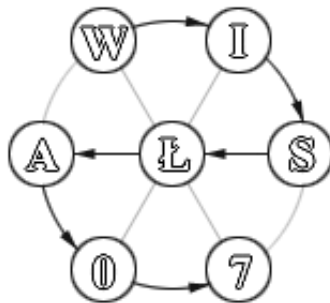


XXXIII Konferencja
„Statystyka Matematyczna Wisła 2007”



3-7 grudnia 2007

Optimalny wybór pary statystyk porządkowych do konstrukcji obszaru ufności dla parametrów położenia i skali

Magdalena Alama-Bućko

Uniwersytet Technologiczno-Przyrodniczy, Bydgoszcz
e-mail: mbucko@utp.edu.pl

Oleksandr Zaihraiev

Uniwersytet Mikołaja Kopernika, Toruń
e-mail: alzaig@mat.uni.torun.pl

Niech $x = (x_1, \dots, x_n)$ będzie próbą z rozkładu P_θ , $\theta = (\theta_1, \theta_2) \in \mathbb{R} \times \mathbb{R}_+$, gdzie θ_1 jest parametrem położenia, a θ_2 parametrem skali. Dla ustalonego poziomu ufności $\alpha \in (0, 1)$ rozważamy problem konstrukcji mocnego obszaru ufności dla (θ_1, θ_2) o minimalnym ryzyku

$$R(\theta, B) = E_\theta \lambda_2(B(x)),$$

gdzie λ_2 oznacza miarę Lebesgue'a na płaszczyźnie.

Metoda konstrukcji takich obszarów ufności (nazywamy je optymalnymi) jest znana (patrz np. [1]). Optymalny obszar ufności ma postać

$$B_A(x) = (t_1(x), t_2(x)) + t_2(x)A,$$

gdzie para statystyk $t_1(x)$ i $t_2(x)$ spełnia warunki:

$$t_1(bx + a\mathbf{1}_n) = bt_1(x) + a, \quad t_2(bx + a\mathbf{1}_n) = bt_2(x) \quad \forall a \in \mathbb{R}, b \in \mathbb{R}_+,$$

a zbiór borelowski A jest określony równością

$$A = \{u \in \mathbb{R}^2 : p(u) \geq z_\alpha\},$$

przy czym z_α wyznacza się z warunku $\int_A p(u)du = \alpha$. Przez $p(u)$ oznaczamy tutaj gęstość wektora losowego $(-t_1(x)/t_2(x), 1/t_2(x) - 1)$ dla próby x z rozkładu $P_{(0,1)}$. Zakładamy dodatkowo, że $\int_{\{u: p(u)=z\}} p(u)du = 0$ dla wszystkich $z > 0$.

Niech $t_1(x)$ i $t_2(x)$ będą kombinacjami liniowymi statystyk porządkowych $x_{k:n}$ i $x_{m:n}$, gdzie $k < m$. Celem referatu jest badanie asymptotycznego zachowania obszaru ufności $B_A(x)$, gdy $n \rightarrow \infty$, a $k = k_n$ i $m = m_n$ są takie, że $k/n \rightarrow p, m/n \rightarrow q, p < q$, oraz rozwiązanie problemu optymalnego wyboru pary statystyk porządkowych $x_{k:n}$ i $x_{m:n}$.

Literatura

[1] Czarnowska, A. i Nagaev, A.V., *Confidence regions of minimal area for the scale-location parameter and their applications*, *Applicationes Mathematicae* 2001, 28, 2, 125-142

Test pierwiastka jednostkowego dla procesów z długą pamięcią

Krystyna Ambroch

Wydział Fizyki Technicznej i Matematyki Stosowanej
Politechnika Gdańska, Gdańsk
e-mail: ambroch@mif.pg.gda.pl

Analiza szeregów czasowych często wymaga usunięcia pewnego rodzaju niestacjonarności. By wykryć jej występowanie, stosuje się testy pierwiastka jednostkowego. Najbardziej znanym wśród nich jest test Dickeya-Fullera, który został opracowany również dla procesów z długą pamięcią.

Celem referatu jest zaprezentowanie testu pierwiastka jednostkowego zakresów, z rozszerzeniem na procesy z długą pamięcią oraz porównanie jego możliwości z możliwościami testu DF w warunkach różnego rodzaju zakłóceń.

Literatura

- [1] Aparicio,F.,Escribano,A.,and Sipols,A.E., *Range unit-root (RUR) tests: robust against nonlinearities, error distributions, structural breaks and outliers*, Journal of Time Series Analysis, 2006
- [2] Dolado,J.J., Gonzalo,J.,and Mayoral,L., *A fractional Dickey-Fuller test for unit roots*, Econometrica, 2002
- [3] Sowell,F., *The fractional unit root distribution*, Econometrica, 1990

Model autopsji koherentnego systemu

Karol Andrzejczak

Instytut Matematyki
Politechnika Poznańska
e-mail: kandrzej@mat.put.poznan.pl

W teorii niezawodności szczególną rolę odgrywa klasa binarnych systemów koherentnych (patrz [Barlow i Proschan]). Do tej klasy należą m.in. systemy szeregowy, równoległy i "k z n". Zakładamy, że czasy zdatności elementów systemu koherentnego są absolutnie ciągłymi zmiennymi losowymi. System traci zdatność w chwili utraty zdatności przez pewien jego element. O tym, który element jest bezpośrednią przyczyną utraty zdatności systemu decyduje struktura niezawodnościowa. Rozważamy statystyczny model, w którym m replik systemu o znanej strukturze niezawodnościowej poddanych jest obserwacji aż do utraty ich zdatności. Uzyskane dane składają się ze zbioru niezdatnych elementów i czasów zdatności systemu. Czasy zdatności elementów są nieznanne. Tak opisany model jest modelem autopsji (the autopsy model) (patrz [Antoine i in.], [Nowik]). W modelu tym przedmiotem badania są dwa statystyczne problemy. Pierwszy problem to estymacja rozkładu czasu zdatności systemu, a drugi problem to estymacja czasów zdatności poszczególnych elementów systemu. W referacie podjęta zostanie próba odpowiedzi na pytanie: Czy w ogóle, bądź w jakich przypadkach, rozkłady czasów zdatności elementów systemu mogą być estymowane w modelu autopsji ?

Literatura

- [1] Antoine, R., Doss, H., Hollander, M., *On identifiability in the autopsy model of reliability theory*, J. Appl. Prob. 30, 913-930, 1993
- [2] Barlow, R.E., Proschan, F., *Theory of Reliability and Life Testing*, Holt, Rinehart and Winston Inc., New York, 1975
- [3] Nowik, S., *Identifiability problems in coherent systems*, J. Appl. Prob. 28, 862-872, 1990

Nowy przedział ufności dla prawdopodobieństwa sukcesu rozkładu dwumianowego

Jerzy Baran

Instytut Matematyki i Informatyki
Politechnika Wrocławska
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław
e-mail: Jerzy.Baran@pwr.wroc.pl

W pracy przedstawiono nową modyfikację przedziału ufności Walda dla prawdopodobieństwa sukcesu rozkładu dwumianowego. Zaproponowany przedział ufności posiada pewne pożądane własności. Pokazano, że skonstruowany przedział ufności, gdy nieznan parametr przyjmuje wartości brzegowe, odznacza się większym prawdopodobieństwem pokrycia w porównaniu z innymi rozważanymi w literaturze przedziałami ufności (Wald, Agresti-Coull, Jeffreys). Otrzymane wyniki zostały zilustrowane symulacjami. Ponadto zaproponowano pewną nową miarę służącą do porównywania przedziałów ufności.

Literatura

- [1] Agresti, Alan and Coull, Brent A., *Approximate is better than "exact" for interval estimation of binomial proportions*, The American Statistician, 52, 2, 119–126, 1998
- [2] Braess, Dietrich and Sauer, Thomas, *Bernstein polynomials and learning theory*, Journal of Approximation Theory, 128, 2, 187–206, 2004
- [3] Brown, Lawrence D. and Cai, T. Tony and DasGupta, Anirban, *Interval estimation for a binomial proportion*, Statistical Science. A Review Journal of the Institute of Mathematical Statistics, 16, 2, 101–133, 2001
- [4] Brown, Lawrence D. and Cai, T. Tony and DasGupta, Anirban, *Confidence intervals for a binomial proportion and asymptotic expansions*, The Annals of Statistics, 30, 1, 160–201, 2002
- [5] Brown, Lawrence D. and Cai, T. Tony and DasGupta, Anirban, *Interval estimation in exponential families*, Statistica Sinica, 13, 1, 19–49, 2003
- [6] Clopper, C. J. and Pearson, E. S., *The use of confidence of fiducial limits illustrated in the case of the binomial*, Biometrics, 26, 404–413, 1934
- [7] Santner, T. J., *Teaching large-sample binomial confidence intervals*, Teaching Statistics, 20, 20–23, 1998

Odporna estymacja parametru skali oparta na spacjach

Jarosław Bartoszewicz

Instytut Matematyczny, Uniwersytet Wrocławski
e-mail: jarbar@math.uni.wroc.pl

Paweł Błażej

Instytut Matematyczny, Uniwersytet Wrocławski
e-mail: Pawel.Blazej@math.uni.wroc.pl

Komunikat jest kontynuacją referatu Błażeja (2007). Rozważane jest pojęcie odporności statystyk w sensie Zielińskiego (1977) w sytuacji, gdy zakłócenia modelu statystycznego są generowane przez rozkłady wazone. Dla pewnych modeli wyznaczono jednostajnie najodporniejsze względem obciążenia estymatory parametru skali w klasie liniowych funkcji spacji, obszerniejszej niż klasa liniowych funkcji statystyk pozycyjnych, jednakże kosztem mocniejszych założeń. Otrzymane rezultaty są rozszerzeniami wyników Bartoszewicza (1986).

Literatura

- [1] Bartoszewicz, J., *Bias-robust estimation of scale parameter*, Prob. and Math. Statist. vol. 7, Fasc. 2, 103-113, 1986
- [2] Błażej, P., *Odporna estymacja parametru skali oparta na statystykach pozycyjnych*, XXXIII Konferencja ze Statystyki Matematycznej w Wiśle, 2007
- [3] Zieliński, R., *Robustness: a quantitative approach*, Bull. Acad. Polon. Sci., Ser. Sci. Math. Astronom. Phys. 25, 1281-1286, 1977

Propozycja systemu wag dla indeksów agregatowych

Jacek Białek

Wydział Ekonomiczno-Socjologiczny
Katedra Metod Statystycznych
Uniwersytet Łódzki
e-mail: jbialek@onet.eu

W pracy zaproponowano konstrukcję agregatowego indeksu wydajności pracy. Proponowany system wag wynika z teoretycznych rozważań nad sytuacją, gdy liczba obserwacji pochodzących od któregoś z analizowanych przedsiębiorstw jest niewystarczająca.

W pierwszej części pracy rozważania dotyczą grupy N - przedsiębiorstw i dwóch okresów ich funkcjonowania. Podajemy konstrukcję indeksu dla porównania tych okresów z punktu widzenia wydajności pracy. Następnie rozważamy przypadek, gdy chcemy zmierzyć przeciętną, jednookresową dynamikę wydajności pracy posiadając dane pochodzące z $T > 2$ okresów. Konstruujemy nowy indeks stanowiący ogólniejszą wersję poprzedniego indeksu.

Oszacowania średniej uogólnionych
statystyk porządkowych
z rozkładów o malejącej uogólnionej funkcji hazardowej

Mariusz Bieniek

Instytut Matematyki
Uniwersytet Marii Curie Skłodowskiej, Lublin
e-mail: mbieniek@hektor.umcs.lublin.pl

Dla ustalonej dystrybuanty W możemy określić rodzinę rozkładów

$$\mathcal{F}_W = \{F : F \succ_c W\},$$

gdzie \succ_c oznacza wypukły porządek stochastyczny dystrybuant. W przypadku, gdy W jest rozkładem jednostajnym lub wykładniczym, otrzymujemy w ten sposób znane rodziny rozkładów o malejącej gęstości (DD) lub malejącej funkcji hazardowej (DFR). Analogicznie, jeżeli W_α oznacza dystrybuantę uogólnionego rozkładu Pareto z parametrem $\alpha \in \mathbb{R}$ (zob. [1]), to rodzinę \mathcal{F}_{W_α} oznaczamy symbolem $\text{DGFR}(\alpha)$, oraz jeżeli $F \in \text{DGFR}(\alpha)$, to mówimy, że rozkład F ma malejącą uogólnioną funkcję hazardową.

Niech teraz $\alpha \in \mathbb{R}$ będzie ustaloną liczbą oraz niech $X_*^{(r)}$, $r \geq 1$, oznacza r -tą uogólnioną statystykę porządkową z rozkładu $F \in \text{DGFR}(\alpha)$ ze średnią μ i wariancją σ^2 . W referacie pokażemy zastosowanie metody projekcji do wyznaczenia oszacowań na iloraz

$$\frac{EX_*^{(r)} - \mu}{\sigma},$$

optymalnych w klasie $\text{DGFR}(\alpha)$. Uzyskane rezultaty są rozszerzeniem wyników przedstawionych w pracach [2] i [3].

Literatura

- [1] Pickands, J., *Statistical inference using extreme order statistics*, Ann. Statist. 3 (1975), 119–131
- [2] Bieniek, M., *Projection bounds on expectations of generalized order statistics from DFR and DFRA families*, Statistics 40 (2006), 339–352
- [3] Bieniek, M., *Projection bounds on expectations of generalized order statistics from DD and DDA families*, J. Statist. Plann. Inference (praca przyjęta do druku)

Odporna estymacja parametru skali oparta na statystykach pozycyjnych

Paweł Błażej

Instytut Matematyczny, Uniwersytet Wrocławski
e-mail: Pawel.Blazej@math.uni.wroc.pl

Rozważane jest pojęcie odporności statystyk w sensie Zielińskiego (1977) w sytuacji, gdy zakłócenia modelu statystycznego są generowane przez rozkłady ważone. Dla pewnych modeli wyznaczono jednostajnie najodporniejsze względem obciążenia estymatory parametru skali w klasie liniowych funkcji statystyk pozycyjnych. Otrzymane rezultaty są rozszerzeniami wyników Zielińskiego (1983) i Bartoszewicza (1986).

Literatura

- [1] Bartoszewicz, J., *Bias-robust estimation of scale parameter*, Prob. and Math. Statist. vol. 7, Fasc. 2, 103-113, 1986
- [2] Błażej, P., *Robust estimation of the scale and weighted distributions*, Applicationes Mathematicae. XXXIV, 29-38, 2007
- [3] Zieliński, R., *Robustness: a quantitative approach*, Bull. Acad. Polon. Sci., Ser. Sci. Math. Astronom. Phys. 25, 1281-1286, 1977
- [4] Zielinski, R., *A most bias-robust linear estimate of the scale parameter of the exponential distribution*, Applicationes Mathematicae XVIII., 73-77, 1983

Metody Monte Carlo w energetyce

Jacek Bojarski

Wydział Matematyki, Informatyki i Ekonometrii
Uniwersytet Zielonogórski
e-mail: j.bojarski@wmie.uz.zgora.pl

Każdy odbiorca energii wprowadza zakłócenia do systemu energetycznego. By zachować jak najlepsze parametry dostarczanego prądu, zgodnie z obowiązującymi normami, stosowane są *międzysystemowe sterowniki rozptywu mocy*. Koszt takich systemów, w głównej mierze, zależy od liczby i charakterystyk odbiorców energii co z kolei przekłada się wymaganą moc sterowników. Symulacje Monte Carlo pozwoliły na wskazanie pewnych optymalnych rozwiązań, które mogą obniżyć koszty sterowników.

Literatura

- [1] G. Benysek, *Improvement in the quality of delivery of electrical energy using power electronics systems*, London : Springer-Verlag, 2007
- [2] R. Strzelecki, G. Benysek, Z. Fedyczak, J. Bojarski, *Interline power flow controller - probabilistic approach*, 2002 IEEE 33rd Annual IEEE Power Electronics Specialists Conference - PESC '02 : Conference proceedings. Cairns, Australia, 2002 .- Piscataway : IEEE Service Center, 2002 .- Vol. 2
- [3] R. Strzelecki, J. Bojarski, G. Benysek, *Probabilistic Method for Parallel Filter's Power Selection in Interline Power Flow Controller*, Probabilistic Methods Applied to Power Systems - PMAFS 2002 : 7th International Conference. Naples, Włochy, 2002 .- [B.m.], 2002 .- Vol. 2
- [4] R. Strzelecki, G. Benysek, J. Bojarski, *Międzysystemowy sterownik przepływu mocy*, Sterowanie w Energoelektronice i Napędzie Elektrycznym - SENE 2001 : V Krajowa Konferencja Naukowa. Arturówek, Polska, 2001 .- Łódź : Instytut Automatyki Politechniki Łódzkiej, 2001 .- T. 2

O pewnych twierdzeniach o punktach stałych w przestrzeniach metrycznych statystycznie

Czesław Bylka

Instytut Matematyki
Politechnika Poznańska
e-mail: bylka@math.put.poznan.pl

Twórcą przestrzeni metrycznych statystycznie jest K. Menger, który główne idee tych przestrzeni przedstawił w pracy „Statistical Metrics” w 1942 roku. W ogólności, pomysł Mengera [7] polega na zastąpieniu deterministycznej metryki Frecheta $d(p, q)$, którą można interpretować jako odległość punktu p od punktu q , odpowiednią metryką statystyczną $F(p, q; x)$, którą między innymi można interpretować jako prawdopodobieństwo zdarzenia, że odległość punktu p od punktu q jest mniejsza niż x . Więcej informacji dotyczących przestrzeni metrycznych statystycznie można znaleźć w następujących pracach: [5,7 i 10]. Pojęcie kontrakcji i jej punktu stałego w przestrzeniach metrycznych statystycznie zainicjowali V. M. Sehgal i A. T. Bharucha-Reid [9]. Bogata różnorodność twierdzeń o punktach stałych kontrakcji określonych w przestrzeniach metrycznych lub w przestrzeniach Banacha, oraz wielu ich zastosowań, inspiruje wielu autorów do badania punktów stałych kontrakcji określonych w przestrzeniach metrycznych statystycznie; C. Bylka [1], S.S. Chang [2] oraz T.L. Hicks [4]. W prezentowanej pracy autor podaje twierdzenie typu Dedeica - Sarpa [3] o punktach stałych kontrakcji określonej w przestrzeni metrycznej, a później rozważa jego statystyczną wersję. Ponadto w przedstawionej pracy autor zajmuje się statystycznymi wersjami twierdzeń Nadlera [8] o punktach stałych pewnych ciągów kontrakcji oraz formułuje dwa twierdzenia Kirka Srinivasana i Veeramani [6] w przestrzeniach metrycznych statystycznie.

Literatura

- [1] Bylka C., *Fixed point theorems of Matkowski on probabilistic metric spaces*, *Demon. Math.*, Vol. XXIX, No. 1, (1996), 159-164
- [2] Chang S.S., *On some fixed point theorems in probabilistic metric space and applications*, *Wahrscheinlichkeitshorie Verw. Gebiete* 63, (1983), 463-473
- [3] Dedeic R., Sarpa N., *A common fixed point theorem for three mappings on Menger spaces*, *Math.Japonica* 34, No 6, (1989), 919-923
- [4] Hicks T.L., *Fixed point theory in probabilistic metric spaces*, *Review of Research Faculty of Science, Univ. of Novi Sad* 19, (1983), 63-72
- [5] Istratescu V. T., *Probabilistic metric space. An introduction*, Ed. Technica Bucharest, (1974)
- [6] Kirk W.A., Srinivasan P.S., Veeramani P., *Fixed points for mappings satisfying cyclical contractive conditions*, *Fixed Point Theory*, Vol.14, No.1, (2003), 79-89
- [7] Menger K., *Statistical metrics*, *Proc. Nat. Acad. Sci., USA*, 28, (1942), 535-537
- [8] Nadler Sam B., Jr., *Sequences of contractions and fixed points*, *Pac.Jour.of*

Math. Vol.27,No.3, (1968), 579-585

[9] Sehgal V.M., Bharucha-Reid A.T., *Fixed points of contraction mappings on probabilistic metric spaces*, Math. System Theory 6, (1972), 97-102

[10] Schweizer B., Sklar A., *Statistical Metric Spaces*, Pac. J. Math. 10, (1960), 313-334

Bootstrapowe obszary predykcyjne dla wielowymiarowych modeli szeregów czasowych

Grzegorz Chłapiński

Wydział Matematyki i Informatyki
Politechnika Wrocławska
e-mail: grzegorz.chlapinski@pwr.wroc.pl

Roman Różański

Wydział Matematyki i Informatyki
Politechnika Wrocławska
e-mail: roman.rozanski@pwr.wroc.pl

Przedstawimy metodę konstrukcji bootstrapowych obszarów predykcyjnych dla wielowymiarowych modeli stacjonarnych szeregów czasowych. W replikacji danych, stosowany jest schemat bootstrapu sitowego. Obszary ufności konstruowane są z wykorzystaniem bootstrapu hybrydowego i bootstrapu studentyzowanego. Rezultaty dotyczące zgodności obu metod zostaną sformułowane. Działanie wspomnianych metod będą zilustrowane na przykładach danych symulacyjnych.

Literatura

- [1] Brockwell, P.J. and Davis, R.A., *Time Series: Theory and Methods (second edition)*, Springer-Verlag New York, Inc. 1991
- [2] Bühlmann, P., *Sieve bootstrap for time series*, Bernoulli 1997, 3(2), 123-148
- [3] Lütkepohl, H., *Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin 1991
- [4] Różański, R. i Zagdański, A., *On the consistency of sieve bootstrap prediction intervals for stationary time series*, *Discussiones Mathematicae. Probability and Statistics* 2004, Vol. 24, 5-40
- [5] Zagdański, A., *On the construction and properties of bootstrap-t prediction intervals for stationary time series*, *Probability and Mathematical Statistics* 2005, 25, 133-153

Estymacja modeli ze zmiennymi ukrytymi za pomocą metody SML

Antoni Leon Dawidowicz

Institut Matematyki,
Uniwersytet Jagielloński, Kraków
e-mail: Antoni.Leon.Dawidowicz@im.uj.edu.pl

Katarzyna Brzozowska-Rup

Wydział Zarządzania i Modelowania Komputerowego,
Politechnika Świętokrzyska, Kielce
e-mail:brzozows@poczta.fm

W zagadnieniach teorii szeregów czasowych często pojawia się problem polegający na tym, że rozkład obserwowanej zmiennej losowej jest de facto rozkładem warunkowym. Fakt istnienia zmiennych ukrytych generuje pewne trudności w estymacji parametrów, a co za tym idzie również w prognozowaniu. Korzystanie ze standardowej metody największej wiarygodności nie jest w tej sytuacji możliwe. Wobec tego faktu zaproponowano zastosowanie zmodyfikowanej metody największej wiarygodności uwzględniającej metody Monte Carlo.

Literatura

- [1] Danielson, J. , *Stochastic volatility in asset prices Estimation with simulated maximum likelihood*, Journal of Econometrics 64(1994),375-400
- [2] Danielson and J., Richard, J.F. , *Accelerated Gaussian Importance Sampler with Application to Dynamics Latent Variable Model* , Journal of Applied Econometrics Vol. 8(1993),S153-S173
- [3] Liesenfeld, R. and Jung R.C. , *Stochastic volatility models: conditional normal versus heavy-tailed distribution* , Journal of Applied Econometrics Vol. 15(2000),137-160

Jednoczesne przedziały ufności dla funkcji hazardu z wykorzystaniem metody bootstrap

Anna Dudek

Wydział Matematyki Stosowanej
Akademia Górniczo-Hutnicza, Kraków
e-mail: aedudek@agh.edu.pl

Rozważymy okresowy model multiplikatywnej intensywności. Przedstawimy zastosowanie zaproponowanej przez nas metody bootstrap dla niestacjonarnych procesów punktowych do konstrukcji wersji estymatora histogram owego największej wiarygodności dla funkcji hazardu. Pokażemy zgodność zaproponowanej metody. Zaprezentujemy jednoczesne przedziały ufności dla okresowej funkcji hazardu uzyskane przy wykorzystaniu zaproponowanej metody bootstrap. Przedstawimy wyniki przeprowadzonych symulacji mających na celu porównanie rzeczywistych prawdopodobieństw pokrycia uzyskanych obszarów ufności z nominalnym. Rezultaty teoretyczne zostaną zilustrowane zastosowaniem do danych rzeczywistych.

Literatura

- [1] Andersen, P.K., Borgan, O., Gill, R.D. i Keiding, N., *Statistical Models Based on Counting Processes*, Springer-Verlag, 1993.
- [2] Cowling, A., Hall, P. i Phillips, M.J., *Bootstrap confidence regions for the intensity of a Poisson point process*, Am. Stat. Ass. 91(436), 1516-1524, 1996.
- [3] Leśkow, J., *Histogram maximum likelihood estimator of a periodic function in the multiplicative intensity model*, Statistics and Decisions 6, 79-88, 1988.

Problemy jednoczesnego testowania
wielu hipotez statystycznych i ich
zastosowanie w analizie mikromacierzy DNA

Konrad Furmańczyk

Katedra Zastosowań Matematyki SGGW

e-mail: konfur@wp.pl

W referacie zostaną przedstawione procedury testowania wielu hipotez takie jak: cut off, step down i step up. Omówione zostaną również podstawowe miary błędów: FWER, k-FWER, FDP, FDR jakie można popełnić stosując powyższe procedury. Rozpatrzemy przypadki zależnych i niezależnych testowań wielu hipotez wraz z ich zastosowaniem do analizy mikromacierzy DNA.

Optymalne oszacowania L -statystyk w modelu ciągłym

Agnieszka Goroncy

Wydział Matematyki i Informatyki
Uniwersytet Mikołaja Kopernika, Toruń
e-mail: gemini@mat.uni.torun.pl

Rozważamy ciągi niezależnych zmiennych losowych o tym samym rozkładzie z dopuszczalnie ciągłą dystrybucją F , dla których wyznaczamy optymalne górne ujemne oszacowania wartości oczekiwanych liniowych kombinacji statystyk pozycyjnych. Oszacowania te wyrażamy w jednostkach skali generowanych przez centralne momenty absolutne różnych rzędów. Podajemy również rozkłady, dla których oszacowania są osiągalne. Pokazujemy, że oszacowania są zerowe za wyjątkiem tych, które są wyrażone w terminach średniego absolutnego odchylenia od średniej. Podajemy również ogólne wyniki dotyczące dodatnich dolnych oszacowań dla ważnych przykładów L -statystyk, takich jak maksimum z próby, rozstęp z próby oraz średnie różnice Giniego.

Metody konstrukcji jąder

Tomasz Górecki

Wydział Matematyki i Informatyki
Uniwersytet im. Adama Mickiewicza, Poznań
e-mail: drizzt@amu.edu.pl

Oryginalnym powodem wprowadzenia jąder w analizie statystycznej było poszukiwanie przestrzeni (a co za tym idzie przekształcenia), w której nieliniowe funkcje mogą być rozpatrywane jako funkcje liniowe. Taki efekt uzyskuje się właśnie poprzez zastosowanie jąder. Oczywiście okazuje się, że nie każda funkcja może służyć za jądro. W pracy opisane są podstawowe typy jąder, omówione są metody jak ze znanych jąder uzyskać inne, bardziej skomplikowane. Okazuje się dodatkowo, że w ogólności sprawdzenie czy dana funkcja jest jądrem może okazać się skomplikowane. Wskazana jest metoda przekształcenia danych na sferę, gdzie znane są stosunkowo proste metody sprawdzania czy dana funkcja jest jądrem.

Literatura

- [1] Burges, C. J. C., *Geometry and invariance in kernel based methods*, MIT Press, 1999
- [2] Herbrich, R., *Learning Kernel Classifiers*, MIT Press, 2002
- [3] Schölkopf, B., Smola, A. J., *Learning with Kernels*, MIT Press, 2002
- [4] Schoenberg, I. J., *Positive definite functions on spheres*, Duke Mathematical Journal, 1942

O estymatorach największej wiarogodności macierzy kowariancji i podprzestrzeniach kwadratowych

Mariusz Grządziel

Katedra Matematyki
Uniwersytet Przyrodniczy we Wrocławiu
e-mail: mg@ar.wroc.pl

Rozważany jest problem estymacji macierzy kowariancji (jednej lub kilku) w modelach liniowych przy użyciu metody największej wiarogodności lub metody największej wiarogodności z ograniczeniami (REML). Dla szerokiej klasy modeli funkcja wiarogodności, po odpowiedniej zamianie zmiennych, jest wypukła. Klasa tych modeli zdefiniowana jest w terminach podprzestrzeni kwadratowych. W pracy została przedstawiona metoda pozwalająca na rozstrzygnięcie, czy dany model należy do tej klasy, bazująca na wynikach z pracy [1].

Literatura

[1] Grządziel M., *Quadratic subspaces and construction of Bayes invariant quadratic estimators of variance components in mixed linear models*, Praca przyjęta do druku w Statistical Papers, 2007

O asymptotycznych rozkładach dwóch testów wielowymiarowej normalności opartych na kurtozie

Zofia Hanusz

Katedra Zastosowań Matematyki
Akademia Rolnicza
Akademicka 13, 20-950 Lublin
e-mail: zofia.hanusz@ar.lublin.pl

Joanna Tarasińska

Katedra Zastosowań Matematyki
Akademia Rolnicza
Akademicka 13, 20-950 Lublin
e-mail: joanna.tarasinska@ar.lublin.pl

Analiza statystyczna wielowymiarowych obserwacji opiera się często na założeniu, że dane mają wielowymiarowy rozkład normalny. W literaturze istnieje wiele testów sprawdzających to założenie. W praktyce, najłatwiejsze do zastosowania są te z nich, których funkcje testowe mają znane (asymptotyczne) rozkłady i w związku z tym nie wymagają specjalnych tablic lub wyznaczania symulacyjnych wartości krytycznych. Dość często jednak stosuje się asymptotyczne rozkłady do zbyt małych prób.

W pracy rozważa się dwa testy wielowymiarowej normalności oparte na kurtozie. Jednym z nich jest test Mardii ([1]), występujący niemal we wszystkich pracach porównujących moce testów wielowymiarowej normalności. Drugi z rozważanych testów to test Srivastavy ([2]), oparty na średniej arytmetycznej jednowymiarowych współczynników kurtozy dla składowych głównych.

Dla obu rozważanych testów często używa się ich asymptotycznych rozkładów normalnych z "grubo" oszacowanymi parametrami. Tymczasem, nawet dla 50-elementowej próby, rzeczywiste rozkłady obu statystyk są od normalności bardzo odległe. Wnioskowanie można znacznie poprawić biorąc dokładniejsze oszacowania parametrów asymptotycznych rozkładów normalnych.

Literatura

[1] Mardia, K.V., *Measures of multivariate skewness and kurtosis with applications*, Biometrika 57 (3), pp.519-530, 1970

[2] Srivastava, M.S., *A measures of skewness and kurtosis and a graphical method for assessing multivariate normality*, Statistics and Probability Letters 2, pp.263-267, 1984

Uwagi o testach wielowymiarowej normalności Srivastavy i Hui

Zofia Hanusz

Katedra Zastosowań Matematyki
Akademia Rolnicza
Akademicka 13, 20-950 Lublin
e-mail: zofia.hanusz@ar.lublin.pl

Joanna Tarasińska

Katedra Zastosowań Matematyki
Akademia Rolnicza
Akademicka 13, 20-950 Lublin
e-mail: joanna.tarasinska@ar.lublin.pl

Celem pracy jest wskazanie na pewną niedokładność w testach wielowymiarowej normalności zaproponowanych przez Srivastavę i Hui. Procedury te są uogólnieniem testu Shapiro-Wilka o jednowymiarowej normalności. Przeprowadzono badania symulacyjne wskazujące na to, że na rozkłady funkcji testowych istotny wpływ ma estymator macierzy kowariancji. Uzyskane wyniki zilustrowano przykładami.

Literatura

[1] Srivastava, M.S. and Hui, T.K., *On Assessing Multivariate Normality Based on Shapiro-Wilk Statistic*, Statistics and Probability Letters 5, pp. 15-18, 1987

Jak optymalne są adaptacyjne testy wynikowe dla testowania jednostajności?

Tadeusz Inglot

Instytut Matematyki i Informatyki
Politechnika Wroclawska

oraz

Zakład Statystyki Matematycznej
Polskiej Akademii Nauk

e-mail: Tadeusz.Inglot@pwr.wroc.pl

Alicja Janic

Instytut Matematyki i Informatyki
Politechnika Wroclawska

e-mail: Alicja.Janic@pwr.wroc.pl

W pracy skonstruowano nową klasę adaptacyjnych testów jednostajności, które są lepsze od dotychczas istniejących w sensie mocy średniej. Za pomocą symulacji wykazano, że osiągają średnią moc bliską optymalnemu testowi Bayesowskiemu. Pokazano również, że strata mocy optymalnego testu Bayesowskiego względem testu Neymana-Pearsona w środkowym zakresie funkcji mocy wyraża się poprzez entropię rozkładu a priori.

Estymacja minimaksowa prawdopodobieństwa
sukcesu przy funkcji straty
LINEX

Alicja Jokiel-Rokita

Instytut Matematyki i Informatyki, Politechnika Wrocławska
e-mail: Alicja.Jokiel-Rokita@pwr.wroc.pl

Ryszard Magiera

Instytut Matematyki i Informatyki, Politechnika Wrocławska
e-mail: Ryszard.Magiera@pwr.wroc.pl

Rozważa się problem minimaksowej estymacji parametru rozkładu dwumianowego przy funkcji straty LINEX. Pokazano, że w rozpatrywanym problemie nie istnieje estymator wyrównujący i decyzja optymalna może być określona poprzez wyznaczenie estymatora bayesowskiego względem rozkładu najmniej korzystnego o skończonym nośniku. W tej sytuacji wartości estymatora minimaksowego i rozkład najmniej korzystny mogą być wyznaczone jedynie wykorzystując metody numeryczne. Opisano zachowanie się estymatorów minimaksowych i odpowiadających im rozkładów najmniej korzystnych w zależności od parametrów funkcji straty. Pokazano, że w rozpatrywanym problemie klasa estymatorów niemalejących jest istotnie zupełna. Otrzymane rezultaty mogą być wykorzystane w problemie nieparametrycznym przy wyznaczaniu minimaksowych estymatorów dystrybuanty i funkcji przeżycia.

Minimaksowa predykcja dystrybuanty empirycznej

Alicja Jokił-Rokita

Instytut Matematyki i Informatyki, Politechnika Wrocławska
e-mail: Alicja.Jokił-Rokita@pwr.wroc.pl

Ryszard Magiera

Instytut Matematyki i Informatyki, Politechnika Wrocławska
e-mail: Ryszard.Magiera@pwr.wroc.pl

Niech $\mathbf{X} = (X_1, \dots, X_n)$ będzie próbą z rozkładu o nieznanym dystrybuancie $F : \mathbb{R} \rightarrow [0, 1]$, o której niczego się nie zakłada. Rozpatruje się problem predykcji dystrybuanty empirycznej $\hat{F}_m(t; \mathbf{Y})$ z przyszłej próby $\mathbf{Y} = (Y_1, \dots, Y_n)$ z rozkładu o dystrybuancie F i niezależnej od wektora \mathbf{X} . Do rozwiązania problemu stosuje się podejście teoriodecyzyjne. Prezentuje się metodę wyznaczania minimaxowego predyktora $d(t; \mathbf{Y})$ dystrybuanty empirycznej $\hat{F}_m(t; \mathbf{Y})$ przy funkcji straty

$$\mathcal{L}(\hat{F}_m(t; \mathbf{Y}), d(t; \mathbf{X})) = \int_{\mathbb{R}} [\hat{F}_m(t; \mathbf{Y}) - d(t; \mathbf{X})]^2 h(F(t)) dW(t),$$

gdzie $h(\cdot)$ jest funkcją ciągłą i dodatnią, W jest niezerową miarą skończoną na $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, gdzie $\mathcal{B}_{\mathbb{R}}$ oznacza σ -ciało zbiorów borelowskich na \mathbb{R} . Metoda rozwiązania polega na sprowadzeniu problemu predykcji dystrybuanty empirycznej \hat{F}_m do problemu estymacji dystrybuanty F przy odpowiedniej funkcji straty, a następnie sprowadzeniu nieparametrycznego problemu estymacji dystrybuanty do parametrycznego problemu estymacji prawdopodobieństwa sukcesu rozkładu dwumianowego. Główny rezultat pozostaje prawdziwy, jeżeli klasę dystrybuant ograniczy się do klasy dystrybuant o nośniku nieujemnym.

O szeregu Grincevičiusa

Agnieszka Klama

Wydział Matematyki, Informatyki i Ekonometrii
Uniwersytet Zielonogórski
e-mail: A.Klama@wmie.uz.zgora.pl

Bardzo interesujący okazuje się warunek konieczny i wystarczający na to, aby szereg zmiennych losowych postaci

$$\eta_1 + \eta_2 \xi_1 + \eta_3 \xi_1 \xi_2 + \cdots + \eta_n \xi_1 \cdots \xi_{n-1} + \cdots$$

był zbieżny p.n. Powyższy szereg zastosowano w znajdowaniu niestałych rozwiązań $\varphi: X \rightarrow \mathbb{R}$ równania funkcyjnego

$$\varphi(x) = \int_{\Omega} \varphi(L(\omega)x + M(\omega)) P(d\omega),$$

gdzie $L: \Omega \rightarrow (0, +\infty)$ i $M: \Omega \rightarrow X$ są mierzalnymi funkcjami na przestrzeni probabilistycznej (Ω, \mathcal{A}, P) , a X jest rzeczywistą i ośrodkową przestrzenią unormowaną.

Estymacja parametrów procesu Markowa modelującego aktywność egzoptydaz

Bogusław Kluge

Instytut Informatyki
Uniwersytet Warszawski
e-mail: bogklug@mimuw.edu.pl

Anna Gambin

Instytut Informatyki
Uniwersytet Warszawski
e-mail: aniag@mimuw.edu.pl

Wojciech Niemiro

Wydział Matematyki i Informatyki
Uniwersytet Mikołaja Kopernika, Toruń
e-mail: wniem@mat.uni.torun.pl

Egzoptydazy to enzymy odcinające pojedyncze aminokwasy z końców peptydów. Ich działanie można pośrednio zaobserwować w danych pochodzących z analizy surowicy krwi spektrometrem masowym. Niedawno [1] zasugerowano, że użyteczność tych danych w różnicowaniu pacjentów z chorobą nowotworową wynika właśnie ze zmiany aktywności egzoptydaz w stanie patologicznym. Zaprezentujemy probabilistyczny model aktywności egzoptydaz bazujący na procesie Markowa z czasem ciągłym. Jest to w istocie rodzaj wielowymiarowego procesu urodzin i śmierci [2]. Sformułujemy twierdzenie o postaci rozkładu stacjonarnego dla tego procesu. Zaproponujemy procedurę estymacji parametrów modelu opartą o algorytm Metropolisa-Hastingsa.

Literatura

- [1] J. Villanueva et. al., *Differential Exoprotease Activities Confer Tumor-Specific Serum Peptidome Patterns*, Journal of Clinical Investigation, 116:271–284, 2006
- [2] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, North-Holland, Amsterdam, 1992

O pewnej metodzie odróżniania od siebie dwóch modeli regresji

Józef Kolonko

Katedra Statystyki, Akademia Ekonomiczna w Katowicach
ul. Bogucicka 14, 40-226 Katowice
e-mail: wywial@ae.katowice.pl

Janusz Wywiał

Katedra Statystyki, Akademia Ekonomiczna w Katowicach
ul. Bogucicka 14, 40-226 Katowice
e-mail: kolonko@ae.katowice.pl

W różnych zagadnieniach praktycznych mamy do czynienia ze zbiorami danych dotyczących wzajemnego położenia względem siebie dwóch funkcji regresji. Rozważamy tutaj przypadek, że każda z tych funkcji (opisująca inną zmienną) zależy od tej samej zmiennej objaśniającej. Przykładem takiego problemu może być położenie względem siebie funkcji popytu i podaży na pewne dobro zależnych od jego ceny. Ekonomistów interesuje oczywiście cena równowagi będąca współrzędną punktu przecięcia obu funkcji. W niniejszej pracy jest rozważany prostszy przypadek punktu przecięcia dwóch liniowych funkcji regresji. Przy klasycznych założeniach dotyczących danych, na podstawie których są szacowane parametry obu linii regresji jest analizowany rozkład prawdopodobieństwa współrzędnej punktu ich przecięcia. Rezultaty wykorzystano do testowania hipotezy głoszącej, że obie linie regresji pokrywają się.

Literatura

- [1] Antoniewicz R. Metoda najmniejszych kwadratów dla zależności niejawnych i jej zastosowania w ekonomii, *Prace naukowe Akademii Ekonomicznej we Wrocławiu*, 1988, nr 445
- [2] Gerstenkorn T., Śródka. T., *Kombinatoryka i rachunek prawdopodobieństwa*, PWN, Warszawa, 1974
- [3] Johnson N.L., S. Kotz, *Continuous Univariate Distributions-1*, Hughton Mifflin Company, Boston 1970
- [4] Hinkley D.V., *On the Ratio of two Correlated Normal Random Variables*, *Biometrika*, 1969, vol. 56, nr 3, s. 635 - 639
- [5] Fieller E.C., *The Distribution of the Index in the Normal Bivariate Population*, *Biometrika*, 1932, vol. 24, s. 428 - 440

Analiza dyskryminacyjna wielowymiarowych powtarzanych pomiarów

Mirosław Krzyśko

Wydział Matematyki i Informatyki
Uniwersytet im. Adama Mickiewicza, Poznań
e-mail: mkrzysko@amu.edu.pl

Michał Skorzybut

Wydział Matematyki i Informatyki
Uniwersytet im. Adama Mickiewicza, Poznań
e-mail: skorzmic@amu.edu.pl

W referacie przedstawiona zostanie procedura klasyfikacyjna bazująca na danych pochodzących z wielowymiarowych, powtarzanych w czasie, pomiarów na obiektach pochodzących z wielu wyróżnionych populacji. Rozpatrzone zostaną różne postacie struktury macierzy kowariancyjnych i problemy estymacji parametrów w takich modelach.

Resampling w dziedzinie częstotliwości dla m -zależnych prawie okresowo skorelowanych szeregów czasowych

Łukasz Lenart

Zakład Ekonometrii
Wyższa Szkoła Biznesu - NLU, Nowy Sącz
e-mail: llenart@wsb-nlu.edu.pl

Modelle dla szeregów czasowych okresowo lub prawie okresowo skorelowanych znalazły zastosowanie w analizie sygnałów telekomunikacyjnych, meteorologii, ekonometrii oraz wielu innych dziedzinach (patrz Gardner, Napolitano oraz Paura (2006)). W dziedzinie częstotliwości problemem identyfikacji szeregów czasowych o strukturze okresowo skorelowanej zajmowali się m.in. Hurd oraz Gerr (1991). Z kolei, w dziedzinie czasu metody identyfikacji danych okresowo skorelowanych zostały zaprezentowane (z zastosowaniem metody subsampling) w Lenart, Leśkow oraz Synowiecki (2007). Celem wystąpienia jest prezentacja nowych testów graficznych dotyczących testowania stacjonarności oraz okresowości opartych na dziedzinie częstotliwości oraz metodzie bootstrap i subsampling. Przedstawione zostaną rezultaty dotyczące zgodności subsamplingowych przedziałów ufności dla wartości bezwzględnej gęstości spektralnej oraz koherencji w klasie m - zależnych prawie okresowo skorelowanych szeregów czasowych. Rezultaty te prowadzą do teoretycznego uzasadnienia stosowania wyżej wspomnianych testów graficznych.

Literatura

- [1] Gardner, W. A., Napolitano, A. and Paura, L., *Cyclostationarity: Half a century of research*, Signal Process, **86**, 639–697, 2006
- [2] Hurd, H., Gerr N. L., *Graphical Methods for determining the presence of periodic correlation*, J. Time Ser. Anal., **12** (4), 337–350, 1991
- [3] Lenart, Ł., Leśkow, J. and Synowiecki, R., *Subsampling in estimation of autocovariance for PC time series*, submitted, 2007

Częściowo zrównoważone układy blokowe z zagnieżdżonymi wierszami i kolumnami

Agnieszka Łacka, Maria Kozłowska

Katedra Metod Matematycznych i Statystycznych
Akademia Rolnicza im. Augusta Cieszkowskiego w Poznaniu
e-mail: aga@riders.pl; markoz@au.poznan.pl

Celem podjętych badań była analiza wyników doświadczeń dwuczynnikowych z pojedynczą kontrolą, obejmująca estymację funkcji parametrów obiektowych w układzie blokowym z zagnieżdżonymi wierszami i kolumnami (krótko NRC; Singh, Day 1979). Rozważane układy należą do klasy układów o ortogonalnej strukturze blokowej (Nelder 1965) i posiadają właściwość ogólnego zrównoważenia (Houtman, Speed 1983). Badania dotyczyły planowania doświadczeń zakładanych w częściowo zrównoważonych układach NRC zwanych układami grup podzielnymi (GD NRC), z dwiema lub trzema klasami partnerów. Przedstawione zostały konstrukcje takich układów. W rozważaniach uwzględniono wyniki między innymi z prac Bagchi (2004), Caliński i Kageyama (2000), Kozłowska (2001), Sinha i Kageyama (2006).

Literatura

- [1] Bagchi S., *Construction of group divisible designs and rectangular designs from resolvable and almost resolvable balanced incomplete block designs*, Journal of Statistical Planning and Inference 119 pp. 401-410, 2004
- [2] Caliński T., Kageyama S., *Block designs: A randomization approach. Volume II: Design*, Springer. New York, Berlin, Heidelberg, Hong Kong, London, Milan, Paris, Tokyo, 2000
- [3] Houtman A.M., Speed T.P., *Balance in designed experiments with orthogonal block structure*, Ann. Stat. 11, pp. 1069-1085, 1983
- [4] Kozłowska M., *Planowanie doświadczeń z zakresu ochrony roślin w układach blokowych z zagnieżdżonymi wierszami i kolumnami*, Roczn. AR Pozn. 313, 2001
- [5] Singh M., Day A., *Block designs with nested rows and columns*, Biometrika 66, 2, pp.321-326, 1979
- [6] Sinha K., Kageyama S., *Constructions of group divisible designs and related combinatorial structures*, Journal of Statistical Planning and Inference 136, pp. 3281-3286, 2006
- [7] Nelder J.A., *The analysis of randomised experiments with orthogonal block structure. 1. Block structure and the null analysis of variance. 2. Treatment structure and general analysis of variance*, Proc. R. Soc. Ser. A, 283, pp. 147-178, 1965

Sekwencyjna, asymptotyczna ($\varepsilon - \alpha$)–aprosymacja metodą MCMC

Krzysztof Łatuszyński

Szkoła Główna Handlowa, Instytut Ekonometrii
e-mail: latuch@gmail.com

W wielu problemach, w których stosuje się algorytmy MCMC, brak jest jakichkolwiek teoretycznych oszacowań na tempo zbieżności tych procedur, lub też istniejące oszacowania są rażąco słabe i nie nadają się do stosowania w praktyce. Jednym ze sposobów radzenia sobie w tej sytuacji, jest estymacja odchylenia standardowego uzyskiwanych estymatorów w trakcie działania procedury i zatrzymywanie algorytmu w momencie $T(\varepsilon, \alpha)$, gdy szerokość przedziału ufności $\text{Int}[T(\varepsilon, \alpha)]$ na poziomie $1 - \alpha$ dla poszukiwanej wartości $E_{\pi}g$ spada poniżej pewnej z góry ustalonej wartości 2ε . Przy odpowiednich założeniach taka sekwencyjna procedura prowadzi do asymptotycznej ($\varepsilon - \alpha$)–aprosymacji, gdzie asymptotyka jest względem ε , a dokładnie

$$P(E_{\pi}g \in \text{Int}[T(\varepsilon, \alpha)]) \rightarrow 1 - \alpha, \quad \text{gdy } \varepsilon \rightarrow 0. \quad (1)$$

Naszkuje tę metodologię w oparciu o pracę [2] i podamy osłabione założenia uzyskane w [1], przy których zachodzi (1).

Literatura

- [1] Bednorz, W., Łatuszyński, K., *A few Remarks on "Fixed-Width Output Analysis for Markov Chain Monte Carlo" by Jones et al.*, to appear in Journal of the American Statistical Association.
- [2] Jones, G. L., Haran, M., Caffo, B. S., Neath, R., *Fixed-Width Output Analysis for Markov Chain Monte Carlo*, Journal of the American Statistical Association, (2006) 101, 1537-1547.

Aproksymacje najmocniejszych testów niezmienniczych wielowymiarowej normalności

Piotr Majerski

Wydział Matematyki Stosowanej
Akademia Górniczo-Hutnicza, Kraków
e-mail: majerski@agh.edu.pl

Rozważamy problem testowania wielowymiarowej normalności przeciwko alternatywie niezmienniczej ze względu na pewną podgrupę transformacji liniowych. Ogólna postać najmocniejszego testu niezmienniczego dla tego problemu została wyprowadzona w [1]. Wyznaczanie jawnej postaci testu w konkretnych przypadkach jest zazwyczaj jednak niemożliwe, ze względu na skomplikowane całki, za pomocą których jest wyrażony test. Dlatego stosowalność tego testu jest ograniczona do najprostszych alternatyw i do przypadku dwuwymiarowych obserwacji. Używając metody Laplace'a rozwijania całek wyprowadzamy asymptotyczne przybliżenia dla statystyk testowych najmocniejszych testów niezmienniczych. Szczegółowo omawiany jest przypadek alternatywy zawierającej zwykły dwuwymiarowy rozkład jednostajny na kwadracie. Pokazujemy również, że tak otrzymane przybliżenie można jeszcze bardziej uprościć poprzez dalsze rozwinięcie - w ten sposób otrzymując statystykę testową testu ilorazu wiarygodności. Jako pośredni wniosek otrzymujemy więc, że w rozważanym problemie test ilorazu wiarygodności jest (przy dostatecznie dużej ilości obserwacji) przybliżeniem najmocniejszego testu niezmienniczego. Symulacje Monte Carlo pokazują, że to przybliżenie jest zaskakująco dobre nawet dla małych próbek.

Literatura

[1] Szkutnik Z., (1988), *Most powerful invariant tests for binormality*, Annals of Statistics, 16, 292-301

Estymacja parametrów rozkładu
Rayleigha i logistycznego w
terminach k -tych wartości rekordowych

Iwona Malinowska

Wydział Elektrotechniki i Informatyki
Politechnika Lubelska, Lublin
e-mail: i.malinowska@pollub.pl

Dominik Szynal

Instytut Matematyki
Uniwersytet Marii Curie-Skłodowskiej, Lublin
e-mail: szynal@golem.umcs.lublin.pl

Podamy - w terminach k -tych wartości rekordowych, estymatory klasyczne (ML, MVU) i Bayesowskie parametrów rozkładu Rayleigha i logistycznego. W estymacji Bayesowskiej korzystamy z symetrycznych (np. kwadratowej) i niesymetrycznych (np. LINEX) funkcji straty. Porównania symulacyjne wartości estymatorów i ich błędów średniokwadratowych są również analizowane.

Liniowa dostateczność i dopuszczalność w podzielonym modelu liniowym

Augustyn Markiewicz

Akademia Rolnicza im. Augusta Cieszkowskiego
ul. Wojska Polskiego 28, 60-637, Poznań
e-mail: amark@au.poznan.pl

W pracy rozważamy estymację liniowych funkcji estymowalnych wektora $\mathbf{X}\boldsymbol{\beta}$ w podzielonym modelu liniowym $\{\mathbf{y}, \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma}, \sigma^2\mathbf{V}\}$. Wyprowadzamy w szczególności charakterystykę estymatorów liniowych, które są liniowo dostateczne i dopuszczalne dla liniowych funkcji estymowalnych wektora $\mathbf{X}\boldsymbol{\beta}$. Następnie charakteryzujemy te estymatory liniowo dostateczne i dopuszczalne, które w zredukowanym modelu liniowym $\{(\mathbf{Z}^\perp)'\mathbf{y}, (\mathbf{Z}^\perp)'\mathbf{X}\boldsymbol{\beta}, \sigma^2(\mathbf{Z}^\perp)'\mathbf{V}\mathbf{Z}^\perp\}$ zachowują swe własności.

O różnych zasadach w problemie estymacji metodą największej wiarogodności

Andrzej Michalski

Katedra Matematyki Uniwersytet Przyrodniczy we Wrocławiu
e-mail: max@ozi.ar.wroc.pl

Autor przedstawia różne aspekty wnioskowania statystycznego przy zastosowaniu metody największej wiarogodności w estymacji parametrów dla modeli liniowych. Klasyczna zasada wiarogodności dotyczyła parametrów stałych (tzw. Fishers's likelihood), natomiast ogólne modele liniowe (GLM) w swojej strukturze zawierają nieobserwowalne zmienne losowe i dla efektywnej estymacji, zarówno parametrów stałych jak i losowych, wymagane jest wprowadzenie ogólniejszej koncepcji wiarogodności. Wielu statystyków np.: Birnbaum (1962); Bayarri, DeGroot i Kadane (1988); Bjornstad(1996) usiłowało odpowiedzieć na kluczowe pytanie: *czy są teoretyczne podstawy dla wyboru szczególnej postaci ogólnej wiarogodności?* Jedną z ważniejszych własności, którą powinna posiadać zasada wiarogodności jest niezmienniczość ze względu na pewną grupę przekształceń (patrz np. praca [2]). W tych rozważaniach autor chciałby przede wszystkim zaprezentować dwie zasady wiarogodności tzw. **extended likelihood** (por.[1]) i **h-likelihood** w zastosowaniu do hierarchicznych uogólnionych modeli liniowych (HGLM)(por. [3])

Literatura

- [1] Bjornstad, J.F., *On the generalization of the likelihood function and likelihood principle*, Journal of the American Statistical Association, 91, 791-806, 1996
- [2] Gnot S., Michalski A., Urbańska - Motyka A., *On some properties of ML and REML estimators in mixed normal models with two variance components*, Discussiones Mathematicae - Probability and Statistics 24, 109-126, 2004
- [3] Lee, Y. and Nelder, J.A., *Hierarchical generalized linear model (with discussion)*, Journal of the Royal Statistical Society B, 58, 619-656, 1996

Empiryczne porównania mocy testów dla rozkładu wykładniczego i Rayleigha

Anna Nosalewicz

Instytut Matematyki UMCS, Lublin
e-mail: anna.nosalewicz@umcs.lublin.pl

Dominik Szynal

Instytut Matematyki UMCS, Lublin
e-mail: szynal@golem.umcs.lublin.pl

Waldemar Wołyński

Wydział Matematyki i Informatyki
Uniwersytet im. A. Mickiewicza, Poznań
e-mail: wołynski@amu.edu.pl

Moce testów zgodności dla rozkładu wykładniczego i Rayleigha, skonstruowanych na bazie warunków charakteryzujących rozkłady ciągłe w terminach wartości oczekiwanych dwóch funkcji wartości rekordowych ([6]–[9]), porównuje się z mocami testów rekomendowanych w literaturze ([1]–[5]).

Literatura

- [1] A. Cabaña, E. Cabaña, *Goodness-of-fit to the exponential distribution, focused on Weibull alternatives*, Commun. Statist. – Simul. Comput., 34, 711–723 (2005)
- [2] P. Grzegorzewski, R. Wieczorkowski, *Entropy-based goodness-of-fit test for exponentiality*, Commun. Statist. – Theory Meth., 28 (5), 1183–1202, (1999)
- [3] N. Henze, S. G. Meintanis, *Recent and classical tests for exponentiality: a partial review with comparisons*, Metrika, 61, 29–45 (2005)
- [4] W. C. M. Kallenberg, T. Ledwina, *Data driven smooth tests for composite hypotheses: comparison of powers*, J. Statist. Comput. Simul., Vol. 59, 101–121 (1997)
- [5] S. Meintanis, G. Iliopoulos, *Tests of fit for the Rayleigh distribution based on the empirical Laplace transform*, Ann. Inst. Statist. Math. 55, no.1, 137–151 (2003)
- [6] K. W. Morris, D. Szynal, *Tests derived from characterizations in terms of moments of record values*, Appl. Math. 30, no. 1, 11–37 (2003)
- [7] K. W. Morris, D. Szynal, *Goodness-of-fit tests via characterizations*, International Journal of Pure and Applied Mathematics, 23, No. 4, 491–555 (2005)
- [8] K. W. Morris, D. Szynal, *Goodness-of-fit tests based on characterizations using record values and U -statistics*, J. Math. Sci. (submitted)
- [9] D. Szynal, *Goodness-of-fit tests via partitions of chi-squared test statistics derived from characterizing conditions*, International Journal of Pure and Applied Mathematics, 35, no. 1, 47–125 (2007)

Estymacja parametru kształtu w rozkładzie gamma

Agnieszka Podraza-Karakulska i Oleksandr Zaihraiev

Uniwersytet Mikołaja Kopernika, Toruń
e-mail: alzaig@mat.uni.torun.pl

W referacie rozważamy problem szacowania nieznanego parametru kształtu $\alpha > 0$ w rozkładzie gamma, gdy parametr skali $\sigma > 0$ też jest niezmany. Niech $x = (x_1, x_2, \dots, x_n)$ będzie próbą z rozkładu $\Gamma(\alpha, \sigma)$ o gęstości

$$p(u; \alpha, \sigma) = u^{\alpha-1} e^{-u/\sigma} / (\sigma^\alpha \Gamma(\alpha)), \quad u > 0.$$

Dla szacowania α najczęściej jest stosowany estymator największej wiarygodności α^* , który jak wiadomo (patrz np. [1]) jest pierwiastkiem równania

$$g(\alpha) = \ln \bar{x} - \sum_{j=1}^n \ln x_j / n,$$

gdzie $\bar{x} = \sum_{k=1}^n x_k / n$, $g(\alpha) = \ln \alpha - \Psi(\alpha)$ a $\Psi(\alpha) = (\ln \Gamma(\alpha))'$ jest tzw. psi (digamma) funkcją Eulera. Zauważmy, że estymator α^* jest niezmienniczy ze względu na parametr skali oraz zachodzi

$$E\left(\ln \bar{x} - \sum_{j=1}^n \ln x_j / n\right) = \Psi(n\alpha) - \Psi(\alpha) - \ln n = g(\alpha) - g(n\alpha).$$

Rozważmy nowy estymator α^{**} parametru α , który określamy jako pierwiastek równania

$$g(\alpha) - g(n\alpha) = \ln \bar{x} - \sum_{j=1}^n \ln x_j / n.$$

Estymator α^{**} można uzyskać stosując zasadę największej wiarygodności do gęstości miary określonej na σ -algebrze zbiorów niezmienniczych ze względu na parametr skali i generowanej przez rozkład wyjściowy względem standardowego rozkładu normalnego (patrz np. [2]).

Okazuje się, że nowy estymator jest lepszy od poprzedniego, bowiem zachodzi
Twierdzenie. *Jeżeli $x = (x_1, x_2, \dots, x_n)$ jest próbą z rozkładu $\Gamma(\alpha, \sigma)$, to*

$$E\alpha^* > E\alpha^{**} > \alpha, \quad \text{Var } \alpha^* > \text{Var } \alpha^{**}.$$

Uwaga. Z twierdzenia wynika, że $E(\alpha^* - \alpha)^2 > E(\alpha^{**} - \alpha)^2$.

W referacie porównujemy też asymptotykę błędów średniokwadratowych dwóch estymatorów, gdy $n \rightarrow \infty$.

Literatura

- [1] Barndorff-Nielsen, O., *Information and Exponential Families in Statistical Theory*, Wiley, New York, 1978
- [2] Nagaev, A. V., *Limit Theorems under Testing Hypotheses*, UMK, Toruń, 1996

Rozkład spektralny macierzy odległości euklidesowych

Piotr Pokarowski

Wydział Matematyki, Informatyki i Mechaniki
Uniwersytet Warszawski, Warszawa
e-mail: pokar@mimuw.edu.pl

Macierze odległości euklidesowych (EDM) są wykorzystywane m. in. w eksploracji danych do redukcji liczby cech w skalowaniu wielowymiarowym lub do redukcji liczby danych w klasteryzacji oraz w geometrii odległości do rekonstrukcji struktury biopolimerów na podstawie danych z jądrowego rezonansu magnetycznego. Zastosowania motywują prace nad strukturą macierzy EDM, a w szczególności nad ich rozkładem spektralnym.

W referowanej pracy wyprowadzam z prawa kosinusów *rozkład radialny* EDM, tj prosty rozkład postaci VDV^T , gdzie D jest macierzą diagonalną, ale niekoniecznie $V^T V = I$. Następnie definiuję *regularne* EDM (REDM), w których rozkład radialny jest równy rozkładowi spektralnemu. Równoważnie: macierz EDM jest REDM wtw, gdy wektor kwadratów odległości od centroidu punktów realizujących macierz jest ortogonalny do współrzędnych głównych tych punktów. Pokazałem, że klasa REDM zawiera znane wcześniej klasy macierzy EDM, dla których opisano rozkład spektralny.

Na podstawie symulacji Monte Carlo i bazy struktur białkowych uzasadniam nazwę 'regularne' EDM: w wygenerowanych macierzach rozkład radialny jest bliski rozkładowi spektralnemu, dlatego można przyjąć, że 'przeciętne' macierze EDM niewiele różnią się od macierzy regularnych. Symulacje prowadzą również do sformułowania dwóch prostych nierówności dla dominującej wartości własnej macierzy EDM. W wygenerowanych macierzach oszacowania te różniły się od policzonej numerycznie wartości własnej tylko o kilka procent.

Na zakończenie opowiem o wykorzystaniu rozkładu radialnego do predykcji struktury białka.

D- optymalna estymacja parametrów w modelu sprężynowego układu wagowego

Katarzyna Przybył

Wydział Matematyki i Informatyki
Uniwersytet im. Adama Mickiewicza, Poznań
e-mail: pkasik@amu.edu.pl

Referat prezentuje problem estymacji miar obiektów w modelu sprężynowego układu wagowego z wykorzystaniem kryterium D- optymalności. Celem jest podanie ograniczenia górnego wyznacznika macierzy informacji estymatorów wektora parametrów przy różnych założeniach dotyczących macierzy kowariancji błędów (dotychczas znane wyniki w literaturze dotyczą jedynie przypadku, gdy macierz kowariancji błędów jest macierzą jednostkową). Określone zostaną warunki dla macierzy modelu, po spełnieniu których układ jest regularny D- optymalny. Ponadto przedstawione będą D- optymalne układy sprężynowe wyznaczone na bazie istniejących układów regularnych.

Wstępne wyniki analizy „sample entropy” w badaniach sygnału elektrohisterograficznego

Dariusz Radomski

Instytut Radioelektroniki
Politechniki Warszawskiej
e-mail: D.Radomski@ire.pw.edu.pl

Sławomir Graczyk

Pracownia Promocji Matki i Dziecka
Uniwersytet Medyczny w Poznaniu

Antoni Grzanka

Instytut Systemów Elektronicznych
Politechniki Warszawskiej
e-mail: antoni.grzanka@ise.pw.edu.pl

Często obecne w procesach biologicznych nieliniowe zależności między zmiennymi stanu procesu powodują konieczność stosowania nieliniowej analizy szeregów czasowych [1]). W pracy skoncentrowano się na analizie sygnału opisującego aktywność elektryczną (EHG) macicy w trzech sytuacjach fizjologicznych: okres przedporodowy, okres porodu oraz okres połogu. Modele opisujące fizjologię skurczu mięśnia macicy wskazują na nieliniową dynamikę elektrycznej i mechanicznej fali skurczowej mięśnia macicy. Wnioskiem z tych modeli była hipoteza mówiąca o synchronizacji czynności elektrycznej mięśnia macicy podczas porodu. Aby zweryfikować tę hipotezę zastosowano uproszczoną wersję przybliżonej entropii, tzw sample entropy do oceny regularności sygnału EHG. Analizowano sygnał zarejestrowany od 61 kobiet w okresie poprzedzającym poród, w czasie porodu i połogu. Dla każdego przebiegu oszacowano wartość indeksu sample entropy. Przeprowadzona analiza wariancji wykazała, że wartość tego indeksu istotnie różnicuje stan przebiegu akcji porodowej.

Literatura

[1] Holger K.Schreiber T., *Nonlinear Time Series Ananlysis*, Wiley, 2004

Ograniczenia prawdopodobieństwa błędu w boostingu

Wojciech Rejchel

Wydział Matematyki i Informatyki
Uniwersytet Mikołaja Kopernika, Toruń
e-mail: iggypop@mat.uni.torun.pl

Chciałbym przedstawić problem klasyfikacji binarnej, w którym opierając się na niezależnych kopiach wektora losowego (X, Y) , chcemy wnioskować o $Y \in \{-1, 1\}$ na podstawie cechy $X \in \mathbb{R}^d$. Jednym z algorytmów klasyfikacyjnych jest boosting, w którym opierając się na, być może "słabych", klasyfikatorach bazowych, budujemy funkcję, która, jak wskazuje praktyka, celnie przewiduje nieznaną etykietę. Rozwój teorii procesów empirycznych pozwala również teoretycznie uzasadnić adekwatność algorytmu. W referacie przedstawimy ulepszone (patrz [1]) ograniczenie prawdopodobieństwa błędu przy założeniu, że wymiar Vapnika-Czervonenkisa rodziny klasyfikatorów bazowych jest skończony.

Literatura

[1] Schapire, R. E., Freund, Y., Bartlett P., Lee W. S. (1998), *Boosting the margin: a new explanation for the effectiveness of voting methods*, Ann. Statist. **26** 1651-1686.

Pewne własności binarnych kompletnych układów blokowych w modelu współoddziaływania

Rafał Różański

Instytut Zarządzania
Państwowa Wyższa Szkoła Zawodowa w Gorzowie Wlkp
e-mail: rozraf@poczta.onet.pl

W eksperymentach, w których obiekty mogą na siebie wzajemnie oddziaływać, często stosowany jest model współoddziaływania z efektami sąsiedztwa. Eksperymentami uniwersalnie optymalnymi w tym modelu są kołowe układy zrównoważone ze względu na sąsiedztwo (CNBD). Istnienie takich układów może wymagać jednak dużej liczby jednostek eksperymentalnych. Dlatego w sytuacji gdy CNBD nie istnieje, poszukuje się zazwyczaj układów optymalnych ze względu na wybrane kryteria. W pracy zostaną przedstawione wyniki dotyczące eksperymentów E-optymalnych w klasie układów kompletnych. Podana zostanie również charakterystyka układów spójnych, wśród których należy poszukiwać eksperymentów optymalnych.

Literatura

- [1] Druilhet, P., *Optimality of circular neighbour balanced designs*, J. Statist. Plann. Infer. 81, 141-152, 1999
- [2] Filipiak, K., Markiewicz, A., *Optimality of circular neighbor balanced designs under mixed effects model*, Statist. Probab. Lett. 61, 225-234, 2003
- [3] Filipiak, K., Konieczna, A., Różański, R., Wojtera-Tyrakowska, D., *On spectral norm minimizers indicating optimal designs*, 2007, Praca w recenzji
- [4] Filipiak, K., Różański, R., *E-optimal designs under an interference model*, Biometrical Letters 42, vol. 2,133-142, 2005

Oszacowania statystyk pozycyjnych niskiego rzędu z rozkładów o malejącej gęstości lub intensywności awarii

Tomasz Rychlik

Instytut Matematyczny PAN
Chopina 12, 87-100 Toruń
e-mail: trychlik@impan.gov.pl

Danielak (2003) wyznaczyła optymalne dodatnie górne oszacowania dla statystyk pozycyjnych rzędu wyższego niż mediana z próby, pochodzących z prób o malejącej gęstości i pokazała, że wartości oczekiwane statystyk niższego rzędu nie przekraczają średniej z populacji. Podobne rezultaty uzyskała dla węższej klasy rozkładów o malejącej intensywności awarii.

W referacie pokażemy, że zerowe oszacowania odchylenia wartości oczekiwanych statystyk pozycyjnych odpowiednio niskich rzędów od średniej z populacji, mierzone w jednostkach skali generowanych przez absolutne centralne momenty populacji rzędu większego niż jeden, nie mogą być poprawione. Wyznamy też odpowiednie ściśle ujemne oszacowania w przypadku, gdy jednostką skali jest centralny absolutny moment rzędu jeden.

Poprawność pewnego modelu gamma-regresji

Jerzy Rydlewski

Wydział Matematyki Stosowanej
Akademia Górniczo Hutnicza, Kraków
e-mail: ry@agh.edu.pl

W zastosowaniach często wyznacza się estymatory największej wiarygodności w modelach, dla których brak matematycznych podstaw takiego postępowania. Podczas, gdy w klasycznym modelu regresji zakładamy, że zmienna losowa ma rozkład normalny, to w rozpatrywanym modelu regresji zakładamy, że ma ona rozkład gamma. Ponadto przyjmujemy, że parametr kształtu rozkładu gamma jest modelowany przez sumę pewnej liczby funkcji trygonometrycznych i funkcji liniowej. Celem niniejszej pracy jest pokazanie poprawności znajdowania estymatora metodą największej wiarygodności w opisywanej sytuacji, to znaczy, wykazanie, że istnieje dokładnie jedno maksimum funkcji wiarygodności dla rozpatrywanego modelu gamma-regresji.

Literatura

- [1] Clarke, P.S. i Smith, P.W.F., *On maximum likelihood estimation for log-linear models with non-ignorable non-response*, Statistics & Probability Letters 73 (2005) 441-448.

Optymalne wykrywanie jednorodnego segmentu początkowego w ciągu stochastycznym

mgr inż. Wojciech Sarnowski

Instytut Matematyki i Informatyki
Politechnika Wroclawska
e-mail: wojciech.sarnowski@pwr.wroc.pl

dr hab. Krzysztof Szajowski, prof. PWr

Instytut Matematyki i Informatyki
Politechnika Wroclawska
e-mail: Krzysztof.Szajowski@pwr.wroc.pl

Problem wyodrębnienia z sekwencji otrzymywanych na bieżąco danych informacji która pozwala zidentyfikować i tym samym możliwie szybko zareagować na określone zjawiska jest we współczesnym świecie sprawą o zasadniczym znaczeniu. Naukowcy pracują nad systemami wczesnego ostrzegania przed różnymi zagrożeniami: trzęsienia ziemi, usterki systemów technologicznych, załamania na rynkach finansowych. Matematyczny model dla takich zagadnień sformułowany został przez Shiryaeva [3]. Podejście to wykorzystuje metody teorii optymalnego stopowania. Modele rozregulowania ciągów stochastycznych mogą być interpretowane jako zagadnienia przełączania między procesami o różnych charakterystykach.

W pracy przedstawiona zostanie konstrukcja algorytmu wykrywania przełączenia między ciągami markowskimi (patrz [2], [4]). Przy znanych charakterystykach dwóch ciągów Markowa mamy do czynienia z obserwacją realizacji ustalonego z nich do pewnej losowej, nieobserwowalnej bezpośrednio chwili, po której obserwujemy realizację zgodnie z drugim z procesów. Analizowany model jest uogólnieniem znanego w literaturze modelu rozregulowania ciągów niezależnych zmiennych losowych. Konstruowany algorytm jest optymalną regułą zatrzymania obserwowanego ciągu losowego, taką która wskazuje z maksymalnym prawdopodobieństwem moment przełączenia ciągów przy zadanej z góry dokładności detekcji. Otrzymana optymalna reguła zatrzymania jest uogólnieniem wyników uzyskanych przez Bojdeckiego dla zadania rozregulowania niezależnych ciągów losowych analizowanego w pracy [1]. Podane zostały wartości progowe dla wybranych przykładów.

Literatura

- [1] T. Bojdecki., *Probability Maximizing Approach to Optimal Stopping and its Application to a Disorder Problem.*, Stochastics, vol. 3. pp. 61-71, 1979.
- [2] W. Sarnowski and K. Szajowski., *On line detection of a part of the sequence with unspecified distribution.*, Institute of Mathematics and Computer Science, Wrocław University of Technology, Technical Report XN I18/05/P-048, 2005.
- [3] A.N. Shiryaev., *The detection of spontaneous effects.*, Sov. Math, Dokl., 2:740-743, 1961. translation from Dokl. Akad. Nauk SSSR 138, 799-801, 1961.
- [4] K. Szajowski., *Optimal on-line detection of outside observation.*, J. Stat. Planning and Inference, 30:413-426, 1992.

Estymacja dystrybuanty rozkładu ciągłego na podstawie danych otrzymanych w chwilach losowych

Agnieszka Stępień-Baran

Instytut Matematyki i Informatyki
Politechnika Wrocławska
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław
e-mail: Agnieszka.Stepien@pwr.wroc.pl

Niech X_1, \dots, X_n będą niezależnymi zmiennymi losowymi o jednakowym rozkładzie o nieznannej dystrybuancie F . Zakładamy, że X_i jest obserwacją uzyskaną w chwili t_i , $i = 1, \dots, n$, gdzie t_1, \dots, t_n są statystykami pozycyjnymi dodatnich zmiennych losowych U_1, \dots, U_n , które są niezależne od X_1, \dots, X_n . Rozpatrujemy niezmienniczy problem estymacji dystrybuanty F względem grupy przekształceń ciągłych i ściśle rosnących. Zakłada się, że strata jaką ponosimy, gdy zakończymy obserwację w chwili t jest postaci

$$\mathcal{L}_t(F, d) = \mathcal{L}(F, d) + c_A k(t) + c(t),$$

gdzie $\mathcal{L}(F, d)$ oznacza stratę związaną z błędem estymacji, $c(t)$ wyraża koszt obserwacji prowadzonych do chwili t , $c_A \geq 0$ jest kosztem pobrania pojedynczej obserwacji, $k(t) = \sum_{i=1}^n I_{[0,t]}(U_i)$ oznacza liczbę obserwacji uzyskanych do chwili $t \geq 0$. Wyznaczono optymalne chwile zatrzymania minimalizujące oczekiwane ryzyko estymatora ekwiwariantnego przy wybranych niezmienniczych funkcjach straty.

Literatura

- [1] Aggarwal, O. P., *Some minimax invariant procedures for estimating a cumulative distribution function*, Ann. Math. Statist. 26, 450–462, 1955
- [2] Ferguson, T. S., *Mathematical Statistics: A Decision Theoretic Approach*, Academic, New York, 1967
- [3] Friedman, D., Gelman, A., Phadia, E., *Best invariant estimator of a distribution function under the Kolmogorov-Smirnov loss function*, Ann. Statist. 16, 1254–1261, 1988
- [4] Joki-Rokita, A., Stępień, A., *Sequential estimation of a location parameter from delayed observations*, Statistical Papers, 2006, w druku
- [5] Mohammadi, L., van Zwet, W. R., *Minimax invariant estimation of a continuous distribution function under entropy loss*, Metrika 56, 31–42, 2002
- [6] Ning, J., Xie, M., *Minimax invariant estimator of continuous distribution function under Linex loss*, Journal of Systems Science and Complexity 20, 119–126, 2007
- [7] Yu, Q., Chow, M.S., *Minimaxity of the empirical distribution function in invariant estimation*, Ann. Statist. 19, 935–951, 1991

Perfekcyjne modele liniowe i perfekcyjne funkcje parametryczne

Czesław Stępnia

Zakład Statystyki i Ekonometrii UMCS, Lublin
e-mail: stepniak@umcs.lublin.pl

Rozważamy klasę ogólnych modeli liniowych ze wspólną macierzą układu i z dowolnymi strukturami macierzy wariancji-kowariancji. Mówimy, że zadany model należący do tej klasy jest perfekcyjny dla zadanej funkcji parametrycznej, albo, że ta funkcja jest perfekcyjna w tym modelu, jeśli jest ona liniowo estymowalna z jednostajnie minimalną wariancją. Charakteryzujemy zarówno wszystkie perfekcyjne modele dla danej funkcji jak i wszystkie perfekcyjne funkcje w zadanym modelu. Otrzymane wyniki są ilustrowane na przykładach.

Analiza błędu pierwszego rodzaju oraz mocy testu dla danych symulowanych oraz różnych parametryzacji macierzy wystąpień

Tomasz Suchocki

Katedra Genetyki i Ogólnej Hodowli Zwierząt
Uniwersytet Przyrodniczy we Wrocławiu
e-mail: suchocki@gen.ar.wroc.pl

Joanna Szyda

Katedra Genetyki i Ogólnej Hodowli Zwierząt
Uniwersytet Przyrodniczy we Wrocławiu

Celem pracy jest zbadanie poziomu błędu pierwszego rodzaju oraz mocy dla testu opartego na ilorazie wiarygodności oraz testu Walda dla różnych parametryzacji macierzy wystąpień efektów genów. Analiza opiera się na danych symulowanych z uwzględnieniem różnych liczebności próby danych oraz różnych modelach genetycznej determinacji modelowanej cechy. Liniowy model mieszany zawiera losowy efekt addytywnie genetyczny zwierzęcia oraz stałe efekty pięciu genów reprezentujące ich addytywny oraz dominacyjny efekt genetyczny oraz epistazę – interakcję pomiędzy wybranymi genami. Efekty modelu estymowano przy wykorzystaniu metody największej wiarygodności przy założeniu znanych wariancji efektów addytywnie poligenicznego oraz efektu błędu.

Literatura

- [1] Cockerham, C. C., *An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present*, Genetics, 1954
- [2] Kao, C. H., and Z-B. Zeng, *Modeling Epistasis of Quantitative Trait Loci Using Cockerham's Model*, Genetics, 2002
- [3] Z-B. Zeng, Wang T., Zou W., *Modeling Quantitative Trait Loci and Interpretation of Models*, Genetics, 2005
- [4] Alvarez-Castro, J.M. and Carlborg, O., *A Unified Model for Functional and Statistical Epistasis and Its Application in Quantitative Trait Loci Analysis*, Genetics, 2007

Bootstrap dla danych okresowych

Rafał Synowiecki

Wydział Matematyki Stosowanej
Akademia Górniczo-Hutnicza, Kraków
e-mail: rsynowie@agh.edu.pl

Dane okresowe (sezonowe) pojawiają się w wielu dziedzinach, m.in. w klimatologii, ekonometrii, teorii sygnałów. Zwykle można modelować je jako szereg czasowy lub tablica zmiennych losowych o okresowej funkcji momentów pierwszego lub wyższych rzędów. Ponieważ modele te są niestacjonarne, wnioskowanie statystyczne może być trudne. Dlatego bardzo istotne jest rozwijanie procedur bootstrapowych, które umożliwiają efektywne aproksymowanie przedziałów ufności.

Dla danych okresowych można skonstruować kilka rodzajów procedur bootstrapowych:

- ruchome bloki (moving block bootstrap, MBB),
- sezonowe bloki (seasonal block bootstrap, SBB),
- okresowe bloki (periodic block bootstrap, PBB).

W referacie omówione zostaną warunki wystarczające zgodności tych metod dla średniej próbkowej przy założeniach dotyczących momentów i tempa α -mieszania. Trzy powyższe metody zostaną porównane symulacyjnie.

Literatura

- [1] Politis D., *Resampling time series with seasonal components*, in *Frontiers in Data Mining and Bioinformatics: Proceedings of the 33rd Symposium on the Interface of Computing Science and Statistics*, Orange County, California, June 13-17, 2001
- [2] Chan V., Lahiri S., Meer W., *Bootstrap for outdoor degradation*, *Techometrics*, 46, 2004
- [3] Leśkow J., Synowiecki R., *On bootstrapping periodic random arrays with increasing period*, submitted
- [4] Synowiecki R., *Consistency and application of moving block bootstrap for non-stationary time series with periodic and almost periodic structure*, *Bernoulli*, 13(4), 2007

Estymacja dopuszczalna liniowych estymatorów efektów stałych i losowych w pewnym modelu liniowym

Ewa Synówka-Bejenka

Wydział Matematyki, Informatyki i Ekonometrii
Uniwersytet Zielonogórski,
Zielona Góra
e-mail: e.synowka@wmie.uz.zgora.pl

Stefan Zontek

Wydział Matematyki, Informatyki i Ekonometrii
Uniwersytet Zielonogórski,
Zielona Góra,
e-mail: s.zontek@wmie.uz.zgora.pl

Rozważany jest problem jednoczesnej estymacji efektów stałych i losowych w modelu liniowym

$$Y = X\beta + Z_1u_1 + \dots + Z_ku_k + e.$$

Zakładając, że $\beta \in \mathcal{R}$ oraz $Z_iZ_i' = \sum_{j=0}^i h_{i+1,j+1}E_j$ dla $i = 0, \dots, k+1$, gdzie $h_{i+1,j+1} \geq 0$ a E_j są idempotentnymi i ortogonalnymi macierzami, badany jest wpływ wartości współczynników $h_{i+1,j+1}$ na postać liniowych estymatorów dopuszczalnych dla wektora

$$[\beta, (Z_1u_1)', \dots, (Z_ku_k)']'.$$

Literatura

- [1] LaMotte, L.R. (1982), *Admissibility in linear estimation*, Ann. Statist. 10.1, 245-255.
- [2] LaMotte, L.R. (1997), *On limits of uniquely best linear estimators*, Metrika 45, 197-211.
- [3] Synówka-Bejenka, E., Zontek, S., *A characterization of admissible linear estimators of fixed and random effects in linear models*, Metrika (praca przyjęta do druku).

Kostki Assouada i tempa zbieżności estymatorów

Zbigniew Szkutnik

Wydział Matematyki Stosowanej
AGH, Kraków
e-mail: szkutnik@agh.edu.pl

W referacie przedyskutowane będą niektóre problemy związane z wyznaczeniem dolnych ograniczeń ryzyka estymatorów z wykorzystaniem lematu Fano, nierówności van Treesa i lematu Assouada. To ostatnie podejście będzie rozważone dokładniej w zastosowaniu do poissonowskich problemów odwrotnych.

Nowe nieobciążone estymatory Zolotarewa
dla indeksu i parametru skali rozkładów
sub-gaussowskich sferycznie niezmienniczych

Piotr Szymański

Instytut Informatyki Stosowanej
Wyższa Szkoła Gospodarki, Bydgoszcz
oraz

Wydział Matematyki i Informatyki
Uniwersytet Mikołaja Kopernika, Toruń
e-mail: galant@mat.uni.torun.pl

W referacie przedstawione zostaną nowe estymatory służące do szacowania parametrów rozkładów sub-gaussowskich sferycznie niezmienniczych.

Definicja 1. Rozkład wektora losowego $\zeta = (\zeta_1, \dots, \zeta_d)$ nazywamy α -stabilnym sub-gaussowskim sferycznie niezmienniczym i oznaczamy $\zeta \sim S_{d,\alpha}(\sigma)$, jeżeli funkcja charakterystyczna $\hat{s}_{d,\alpha}(t)$ tego wektora ma postać

$$\hat{s}_{d,\alpha}(t) = \exp(-\sigma^\alpha |t|^\alpha), \quad \alpha \in (0, 2], \quad \sigma > 0, \quad t \in \mathbb{R}^d.$$

Zolotarev w swoich pracach [1,2,3,4] przedstawia estymatory dla indeksu α i parametru skali σ oparte na metodzie momentów dla zlogarytmowanych długości wektorów losowych. Są one bardzo proste ale charakteryzują się obciążeniem i znacznym ryzykiem. W referacie przedstawione zostaną nowe estymatory, bazujące na estymatorach Zolotarewa ale niezmiennicze i o mniejszym ryzyku.

Literatura

- [1] V. V. Uchaikin, V. M. Zolotarev, *Chance and Stability, Stable Distributions and Their Applications*, Netherlands, Utrecht, VSP Science, 1999
- [2] V. M. Zolotarev, *An Estimate of the parameters of multivariate stable distributions*, Theory Probab. Appl. 23 (1978), pp. 472-473
- [3] V. M. Zolotarev, *Statistical estimates of the parameters of stable laws*, Mathematical Statistics Banach Center Publications, 1980
- [4] V. M. Zolotarev, *One-dimensional Stable Distributions*, Volume 65 of Translations of mathematical monographs. American Mathematical Society. Translation from the original 1983 Russian edition, 1986

Macierze GIG i macierze Wisharta

Jacek Wesołowski

Wydział Matematyki i Nauk Informatycznych
Politechnika Warszawska
e-mail: wesolo@mini.pw.edu.pl

Po raz pierwszy na powiązania między macierzowymi rozkładami GIG i Wisharta zwrócił uwagę Butler w 1998 roku. Na lepsze zrozumienie tych związków znaczny wpływ miało odkrycie tzw. własności Matsumoto-Yora w 2000 roku, a szczególnie jej wersji macierzowej dla macierzy o różnych wymiarach. Referat będzie koncentrował się właśnie na tej własności i jej konsekwencjach. Dodatkowo przedstawiony będzie pewien model mieszkankowy, który wskazuje na inny aspekt powiązań między tymi rozkładami. Omawiane wyniki powstały w ciągu kilku ostatnich lat we współpracy m.in. z G. Letac (Tuluza), H. Massam (Toronto), V. Seshadri (Montreal).

Modele wielowymiarowych szeregów czasowych o zredukowanym rzędzie z zastosowaniami do analizy danych mikromacierzowych

Adam Zagdański

Instytut Matematyki i Informatyki
Politechnika Wrocławska
e-mail: Adam.Zagdanski@pwr.wroc.pl

Rafal Kustra

Department of Public Health Sciences
University of Toronto
Canada

Przedstawione zostanie nowe podejście do analizy wielowymiarowych szeregów czasowych, charakteryzujących się dużą liczbą silnie skorelowanych składowych, obserwowanych w krótkim horyzoncie czasowym. Dane tego typu pojawiają się m.in. jako wynik eksperymentów mikromacierzowych, pozwalających na monitorowanie aktywności genów w kolejnych odstępach czasowych (*time-course microarray experiments*).

Zaproponowana metoda analizy oparta jest na klasycznym wielowymiarowym modelu o zredukowanym rzędzie (*reduced-rank model*) zaproponowanym przez Andersona (1951), uogólnionym następnie na przypadek modeli wielowymiarowych szeregów czasowych (m.in: Box i Tiao (1977) oraz Velu i inni (1986)).

Zasadniczym rezultatem jest uogólnienie modelu o zredukowanym rzędzie na przypadek, kiedy liczba składowych jest znacznie większa od długości szeregu czasowego ($p \gg n$). W tym celu zastosowano odpowiednie metody regularyzacji dla estymacji macierzy kowariancji oraz zaproponowano nową metodę wyboru optymalnych parametrów modelu skonstruowaną na bazie schematu *cross-validation*.

W czasie referatu przedstawione zostaną także przykłady zastosowania omawianej metody do analizy danych mikromacierzowych.

Literatura

- [1] Anderson, T., *Estimating linear restrictions on regression coefficients for multivariate normal distributions*, Ann. Math. Statis. 22, 327–351, 1951
- [2] Box, G. and Tiao, G., *A canonical analysis of multiple time series*, Biometrika 64, 355–365, 1977
- [3] Velu, R., Reinsel, G. and Wichern, D., *Reduced rank models for multiple time series*, Biometrika 73, 105–118, 1986

Identyfikacja odstających podzbiorów w danych medycznych

Marta Zalewska

Zakład Profilaktyki Zagrożeń Środowiskowych,
Akademia Medyczna Warszawa

Antoni Grzanka

Instytut Systemów Elektronicznych
Politechnika Warszawska

Wojciech Niemirowicz

Wydział Matematyki i Informatyki
Uniwersytet Mikołaja Kopernika, Toruń
e-mail: wniem@mat.uni.torun.pl

Bolesław Samoliński

Zakład Profilaktyki Zagrożeń Środowiskowych,
Akademia Medyczna Warszawa

Umiejętność rozpoznawania nietypowych (odmiennych od reszty, np. obciążonych błędem) podzbiorów danych jest niezbędna w wielu różnych dyscyplinach nauki (np. w medycynie, ekonomii, socjologii). Szczęólnego znaczenia nabiera wówczas, gdy mamy do czynienia z dużymi zbiorami danych, a gotowe narzędzia statystyczne np. w postaci analizy dyskryminacji czy składowych głównych są niewystarczające do rozwiązania problemu. W pracy przedstawiono algorytm rozpoznawania nietypowych podzbiorów za pomocą nowego wskaźnika opartego na estymatorze prawdopodobieństwa błędnej klasyfikacji. Wykorzystano „niemal nieobciążony” estymator związany z metodą weryfikacji krzyżowej tzw. „leave-one-out” [1]. Zaproponowano dwie wersje tego wskaźnika. Ustalanie poziomu krytycznego powstałego w ten sposób testu jednorodności jest oparte na idei zapożyczony z metody „bootstrap”.

Literatura

[1] Lachenbruch, P.A., *An Almost Unbiased Method of Obtaining Confidence Intervals for the Probability of Misclassification in Discriminant Analysis*, *Biometrics*. 23, pp. 639-645, 1967

Nierówność Dvoretzkiego-Kiefera-Wolfowitza dla estymatorów jądrowych

Ryszard Zieliński

Instytut Matematyczny PAN
Zakład Statystyki Matematycznej
e-mail: R.Zielinski@impan.gov.pl

Standardowe estymatory jądrowe nie są zbieżne JEDNOSTAJNIE do rozkładu, z którego pochodzą obserwacje. W konsekwencji nie da się skonstruować nierówności typu nierówność Kiefera-Dvoretzkiego-Wolfowitza i odpowiedzieć na pytanie, ile potrzeba obserwacji aby osiągnąć zamierzona dokładność przybliżenia estymowanego rozkładu. Okazuje się, że można to uzyskać za pomocą estymatorów jądrowych z losową szerokością okna.

Wnioskowanie statystyczne w modelach liniowych

Roman Zmyślony

Wydział Matematyki, Informatyki i Ekonometrii
Uniwersytet Zielonogórski
e-mail: r.zmyslony@wmie.uz.zgora.pl

1. Wartość oczekiwana i operator kowariancji w przestrzeniach z iloczynem skalarnym.
2. Model liniowy i optymalna estymacja parametrów.
3. Modele liniowe normalne – postać operatora kowariancji dla estymacji wariancji.
4. Charakterystyka JNNE (BLUE) dla parametrów wartości oczekiwanej i (BQUE) dla operatora kowariancji.
5. Minimalne statystyki dostateczne oraz warunki konieczne i dostateczne zupełności.
6. Testowanie hipotez dla wariancji.
7. Jednolita teoria testowania dla parametrów modelu liniowego.
8. Przedziały ufności, zastosowanie w rozkładzie logarytmiczno-normalnym.