# Bickel–Rosenblatt test

Audris Ločmelis, Jānis Valeinis

University of Latvia

28.05.2011.

# A classical Bickel–Rosenblatt test

- Let $X_1, \ldots, X_n$ be i.i.d. random variables with a continuous probability density function $f$.
- Consider a simple hypothesis $H_0 : f = f_0$ with a significance level $\alpha$ and completely specified $f_0$.
- Given the kernel density estimate

$$f_n(x) = \frac{1}{nh_n} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h_n}\right),$$

where $h_n = h(n)$, a test statistic can be defined.

**The classical Bickel–Rosenblatt test statistic [Bickel and Rosenblatt(1973)]**

$$\hat{T}_n^{br} = nh_n \int [f_n(x) - f_0(x)]^2 a(x) dx.$$

# A smoothed modification

To avoid bias problems a smoothed version of $f_0$, namely

$$(K_{h_n} * f_0)(\cdot) = \int h_n^{-1} K\left(\frac{\cdot - z}{h_n}\right) f_0(z) dz,$$

where $*$ is a convolution operator, is employed. And $a(x) \equiv 1$ is used as the arbitrary weight function, which leads to a modification of the Bickel–Rosenblatt test statistic

$$T_n = nh_n^{d/2} \int [f_n(x) - (K_{h_n} * f_0)(x)]^2 \, dx,$$

and for composite hypothesis

$$T_{n,\hat{\theta}} = nh_n^{d/2} \int \left[f_n(x) - (K_{h_n} * f_{\hat{\theta}})(x)\right]^2 dx.$$

# Graphical representation

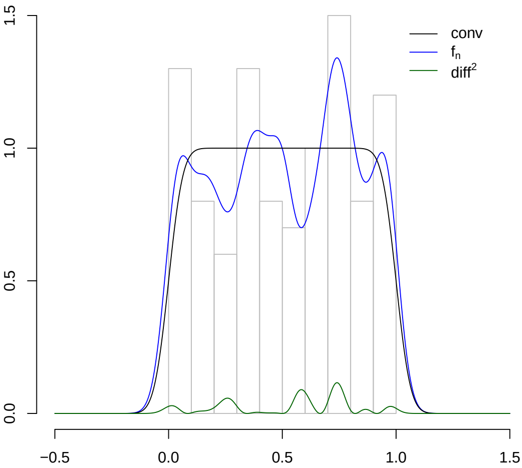



Figure: Convolution, $f_n$ and the squared error.

# An absolutely regular weakly dependent process

Let $(X_t)_{t \in \mathbb{Z}}$, $X_t \in \mathbb{R}$ be a strictly stationary process on a probability space $(\Omega, \mathcal{F}, P)$. For any two $\sigma$-fields $\mathcal{A}$ and $\mathcal{B} \subset \mathcal{F}$ define the following measure of dependence

$$\beta(\mathcal{A}, \mathcal{B}) := \sup \frac{1}{2} \sum_{i=1}^{I} \sum_{j=1}^{J} |P(A_i \cap B_j) - P(A_i)P(B_j)|,$$

such that $A_i \in \mathcal{A} \; \forall i$ and $B_j \in \mathcal{B} \; \forall j$, where $\forall i, j \; A_i, B_j \subset \Omega$. Define $\mathcal{F}_J^L := \sigma(X_k, J \leq k \leq L)$, when $-\infty \leq J \leq L \leq \infty$.

**Definition**

$(X_t)_{t \in \mathbb{Z}}$ is called absolutely regular or $\beta$-mixing if

$$\beta(n) = \sup_{J \in \mathbb{Z}} \beta(\mathcal{F}_{-\infty}^{J}, \mathcal{F}_{J+n}^{\infty}) \to 0, \text{ when } n \to \infty.$$

# Asymptotic distribution of $T_n$

**Theorem ([Neumann and Paparoditis(2000)])**

*If certain assumptions are fulfilled, then under $H_0$,*

$$(T_n - \mu) \xrightarrow{d} N(0, \sigma^2),$$

**where $\mu$ and $\sigma^2$ are**

$$\mu = h_n^{-d/2} \int K^2(u) du,$$

$$\sigma^2 = 2 \int f_0^2(x) dx \times \int \left[ \int K(u) K(u+v) du \right]^2 dv.$$

# Motivation

$(+)$
The test statistic can be used for:

- simple as well as composite hypothesis,
- independent and dependent identically distributed data without modification.

$(-)$
No procedure for selecting the bandwidth $h_n$.

# Simulation study

We define by $f_u$ the probability density function of the uniform $U[0,1]$ distribution

$$f_u = F'_u, \quad F_u = U[0,1].$$

For the process $(X_t)_{t \in \mathbb{Z}}$ we test the single hypothesis

$$H_0 : f = f_u \ \text{ versus } \ H_1 : f \neq f_u.$$

Suppose that a random variable $X$ has a continuous cumulative distribution function $F_X$, then

$$F_X(X) = Y \sim U[0,1].$$

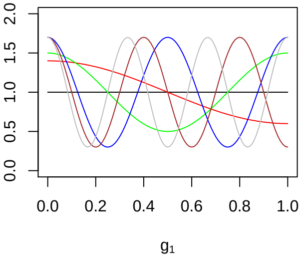
# Alternatives close to $U[0,1]$

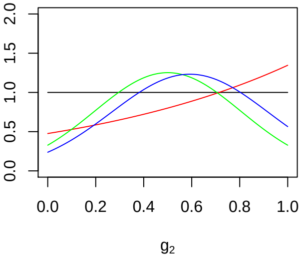
**[Kallenberg and Ledwina(1995)] uses**

$$g_1(x) = 1 + \rho \cos(j\pi x),$$

$$g_2(x) = \exp\left(\sum_{j=1}^{k} \theta_j \pi_j(x) - \psi_k(\theta)\right),$$

with $\{\pi_j\}$ the orthonormal Legendre polynomials on $[0,1]$,
$\psi_k(\theta) = \log \int_0^1 \exp(\theta \circ \phi(x))dx$, $\theta \in \mathbb{R}^k$.



# Simulated power for dependent ($AR(1)$, $\theta = -0.3$) data

Table: $T_n$ percentage rejections of the true $H_0$ at 5% significance level with $n = 20, 50, 100, 500, 1000$ for AR(1) case with $\phi = -0.3$ made with 10,000 replications; $h = h_0 n^{-1/4}$; kernel $U(0,1)$.

| $n$ | $h_0$: 0.005 | 0.01 | 0.02 | 0.03 | 0.05 | 0.1 | 0.15 | 0.2 | 0.25 | 0.3 |
|---|---|---|---|---|---|---|---|---|---|---|
| 20 | 6.53 | 6.10 | 5.42 | **4.68** | 4.26 | 2.97 | 2.39 | 1.86 | 1.48 | 1.08 |
| 50 | 6.26 | 5.97 | **5.31** | **5.31** | 4.59 | 3.59 | 2.91 | 2.48 | 2.07 | 1.63 |
| 100 | 6.02 | 5.77 | **4.98** | 4.82 | 4.44 | 3.69 | 3.40 | 3.01 | 2.56 | 2.33 |
| 500 | 5.91 | 5.94 | 6.00 | **5.21** | **5.20** | 4.34 | 3.91 | 3.53 | 3.13 | 2.85 |
| 1000 | 5.95 | 6.05 | 5.99 | 4.49 | 4.29 | 4.52 | 4.88 | 3.66 | 3.42 | 3.36 |

Table: *AR(1) case with* $\phi = -0.3$. Simulated power for alternatives $g_1$ and $g_2$ with $n = 50$ and 10,000 replications; $h = h_0 n^{-1/4}$, kernel $U(0,1)$.

| $\rho$ | $j$ | $h_0$: 0.005 | 0.01 | 0.02 | 0.03 | 0.05 | 0.1 | 0.15 | 0.2 | 0.25 | 0.3 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.4 | 1 | 6.1 | 9.2 | 10.8 | 12.8 | 15.2 | 22.5 | 28.6 | 34.4 | 37.3 | 41.5 |
| 0.5 | 2 | 8.9 | 12.8 | 16.6 | 19.2 | 26.3 | 38.9 | 46.0 | 50.7 | 54.6 | 57.5 |
| 0.7 | 4 | 15.3 | 21.9 | 33.8 | 40.5 | 53.0 | 70.0 | 75.7 | 77.9 | 75.4 | 70.4 |
| 0.7 | 5 | 16.3 | 22.6 | 33.6 | 39.7 | 53.6 | 67.9 | 70.8 | 68.1 | 58.3 | 42.2 |
| 0.7 | 6 | 16.1 | 22.0 | 32.0 | 38.9 | 52.2 | 64.2 | 64.8 | 55.3 | 39.0 | 22.2 |
| $\theta$ | | | | | | | | | | | |
| (0, 3) | | 25.1 | 25.3 | 25.2 | 24.5 | 27.6 | 34.6 | 40.1 | 45.3 | 49.8 | 53.7 |
| (0,-0. 4) | | 7.8 | 11.1 | 15.4 | 17.1 | 23.2 | 33.4 | 38.4 | 42.5 | 44.8 | 47.0 |
| (0.25,-0. 35) | | 9.3 | 11.7 | 15.5 | 17.4 | 23.7 | 33.8 | 40.2 | 45.7 | 48.5 | 51.6 |

# Bandwidth selection for nonparametric kernel tests

[Gao and Gijbels(2008)] consider a statistic $\hat{T}_n(h)$, similar to $T_n$, for regression fit and derive Edgeworth expansions for size and power functions:

$$\alpha_n(h) = P(\hat{T}_n(h) > l_\alpha | H_0) \quad \text{and}$$
$$\beta_n(h) = P(\hat{T}_n(h) > l_\alpha | H_1),$$

where $l_\alpha$ is a simulated critical value of $\hat{T}_n(h)$.

The Edgeworth expansions of $\alpha_n(h)$ and $\beta_n(h)$ are then used to choose a suitable bandwidth

$$\beta_n(h_{ew}) = \max_{h \in H_n(\alpha)} \beta_n(h),$$

with $H_n(\alpha) = \{h : \alpha - c_{\min} < \alpha_n(h) < \alpha + c_{\min}\}$ for a small $0 < c_{\min} < \alpha$.

- Gao an Gijbels used Edgeworth expansions for quadratic forms.
- [Bachmann and Dette(2005)] states that under $H_0$ $(T_n/nh)$ is a degenerate $U$–statistic.
- For i.i.d. random variables and fixed bandwidth [Tenreiro(2005)] states that statistic $I_n^2(h) = T_n/h$ is a $V$–statistic,

$$I_n^2(h) = \frac{1}{n}\sum_{i,j=1}^{n} Q_h(X_i, X_j),$$

$$Q_h(u,v) = \int k(x,u,h)k(x,v,h)dx \ \ldots$$

- [Fan and Linton(2003)] have derived Edgeworth expansions for a regression model specification test statistic, that is also a degenerate $U$–statistic.

$$\frac{T_n}{nh} - \frac{1}{nh}\int K^2(x)dx - \int [H_h * (f - f_0)]^2(x)dx$$
$$= U_n + \frac{2}{n}\sum_{i=1}^{n} Y_i + O_P\left(\frac{1}{n}\right),$$

where $Y_i = (K_h * g_h)(Z_i) - E[K_h * g_h(Z_i)]$ and

$$U_n = \frac{2}{n^2}\sum_{i<j} H_n(Z_i, Z_j)$$
$$= \frac{2}{n^2}\sum_{i<j}\int [K_h(x - Z_i) - K_h * f(x)]\,[K_h(x - Z_j) - K_h * f(x)]\,dx$$

and $U_n$ is a degenerate $U$–statistic.

# Bibliography I


D. Bachmann and H. Dette.
A note on the Bickel - Rosenblatt test in autoregressive time series.
*Stat. Probab. Lett.*, 74(3):221–234, 2005.

P. J. Bickel and M. Rosenblatt.
On some global measures ot the deviations of density function estimates.
*The Annals of Statistics*, 1(6):1071–1095, 1973.

Y. Fan and O. Linton.
Some higher-order theory for a consistent non-parametric model specification test.
*J. Stat. Plann. Inference*, 109(1-2):125–154, 2003.

J. Gao and I. Gijbels.
Bandwidth selection in nonparametric kernel testing.
*Journal of the American Statistical Association*, 103(484):1584–1594, 2008.

W. C. M. Kallenberg and T. Ledwina.
Consistency and monte carlo simulation of a data driven version of smooth goodness-of-fit tests.
*Annals of Statistics*, 23(5):1594–1608, 1995.

M. H. Neumann and E. Paparoditis.
On bootstrapping $L_2$-type statistics in density testing.
*Statistics & Probability Letters*, 50(2):137–147, 2000.

C. Tenreiro.
On the role played by the fixed bandwidth in the bickel-rosenblatt goodness-of-fit test.
*SORT*, 29(2):201–216, 2005.

Thank you!