Exact L_2-small ball probabilities for Durbin's processes

Yulia Petrova

Saint-Petersburg State University
& Chebyshev laboratory

6 May 2019

Outline: small deviations for Durbin's processes

1. Introduction:
   - Basic notion: small deviation probability
   - Problem statement
   - Durbin processes
2. Historical review
   - Relation to the spectral problem
   - Critical and noncritical perturbations
3. Main results
   - Durbin process for testing exponentiality
   - Kac-Kiefer-Wolfowitz processes
   - Durbin process for Gumbel distribution
4. Main ingredient of proof
   - Slowly varying functions
   - Asymptotics of integrals

Basic notion: small deviation probability

$X(t)$, $t \in (0,1)$, — Gaussian process, $\mathbb{E}X(t) \equiv 0$, $G_X(s,t) = \mathbb{E}X(s)X(t)$.

Definition

To find the asymptotics of small deviation probability of the process $X(t)$ in L_2-norm means to find the asymptotics:

$$\mathbb{P}\left(\|X\|_2 < \varepsilon\right) = \mathbb{P}\left(\int_0^1 (X(t))^2\, dt < \varepsilon^2\right), \qquad \varepsilon \to 0 \tag{1}$$

$$\mathbb{P}\left(\|W\|_2 < \varepsilon\right) \sim \frac{4}{\sqrt{\pi}}\, \varepsilon\, \exp\bigl(-\frac{1}{8}\, \varepsilon^{-2}\bigr)$$

«Typical» answer:

$$\mathbb{P}\left(\|X\|_2 < \varepsilon\right) \sim D \cdot \varepsilon^C \cdot \exp\bigl(-B\varepsilon^{-A}\bigr)$$

A, B — *Logarithmic* asymptotics; A, B, C, D — *Exact* asymptotics

Problem statement (general setting)

$X_0(t)$ — Gaussian process:

- $\mathbb{E}X_0(t) \equiv 0$
- $G_0(s,t) = \mathbb{E}X_0(s)X_0(t)$

$\mathbb{P}\left(\|X_0\|_2 < \varepsilon\right)$ is known

$X(t)$ — finite-dimensional perturbation of rank m of the process $X_0(t)$:

- $\mathbb{E}X(t) \equiv 0$
- $G(s,t) = \mathbb{E}X(s)X(t)$

$$\boxed{G(s,t) = G_0(s,t) + \vec{\psi}^{\,T}(s) \cdot D \cdot \vec{\psi}(t)}$$

Parameters of the perturbation:

- $\vec{\psi}(t) = (\psi_1(t), \dots, \psi_m(t))^T$
- $D \in M_{m\times m}$ — symmetric matrix (w.l.o.g.)

Question:

What is the relation between $\mathbb{P}\left(\|X_0\|_2 < \varepsilon\right)$ and $\mathbb{P}\left(\|X\|_2 < \varepsilon\right)$?

Problem statement (Durbin processes)

- important for statistics
- appear as limiting ones when building goodness-of-fit tests of ω^2-type when parameters of the distribution are estimated from the sample

Given a sample $x_1, \dots, x_n \sim F(x, \theta)$.
$\theta = (\theta_1, \dots, \theta_m)$ — parameters of the distribution.

parameters *known*	parameters *not known*
($\theta = \theta^0$ fix) $\Downarrow$ limiting process — Brownian bridge $B(t)$	(evaluated from a sample) $\Downarrow$ limiting process — m-dimensional perturbation of $B(t)$

Problem:

Find exact L_2-small ball asymptotics for Durbin processes

Kac-Kiefer-Wolfowitz processes (KKW)

Important example: test for normality, $x_1, \dots, x_n \sim F(x, \theta)$

$$f(x,\theta) = \frac{1}{\beta\sqrt{2\pi}} \exp\left(-\frac{1}{2}\Big(\frac{x-\alpha}{\beta}\Big)^2\right); \qquad F(x,\theta) = \int_{-\infty}^{x} f(y,\theta)\,dy$$

- $\widehat{\alpha}$ estimated, $\beta = 1$:

$$G_1(s,t) = G_B(s,t) - \psi_1(s)\psi_1(t), \qquad \psi_1(t) = \varphi(\Phi^{-1}(t))$$

- $\alpha = 0$, $\widehat{\beta}$ estimated:

$$G_2(s,t) = G_B(s,t) - \psi_2(s)\psi_2(t), \qquad \psi_2(t) = \psi_1(t) \cdot \frac{\Phi^{-1}(t)}{\sqrt{2}}$$

- $\widehat{\alpha}$, $\widehat{\beta}$ estimated:

$$G_3(s,t) = G_B(s,t) - \psi_1(s)\psi_1(t) - \psi_2(s)\psi_2(t)$$

Rmk: Here φ and Φ is the density and DF with standard parameters $\alpha = 0, \beta = 1$.

The problem is related to:

Stochastic processes

$X(t)$, $t \in (0,1)$, —

- Gaussian processes
- $\mathbb{E}X(t) \equiv 0$
- $G(s,t) = \mathbb{E}X(s)X(t)$.

Spectral theory

$\mathbb{G} : L_2[0,1] \to \mathrm{Im}(\mathbb{G})$

- integral operator

$$(\mathbb{G}u)(s) = \int_0^1 G(s,t)u(t)\,dt$$

- trace operator: $\sum \mu_k < \infty$

Small ball probabilities

$$\mathbb{P}\left(\|X\|_2 < \varepsilon\right), \quad \varepsilon \to 0$$

Asymptotics of eigenvalues μ_k

Find «good» approximation to μ_k

Hilbert structure $\Longrightarrow$ spectral problem

Karhunen–Loève expansion (KL-expansion):

(due to K. Karhunen'1947, M. Loève'1948)

$$X(t) \stackrel{d}{=} \sum_{k=1}^{\infty} \sqrt{\mu_k}\, u_k(t)\, \xi_k$$

- ξ_k, $k \in \mathbb{N}$, — iid standard normal r.v.
- $u_k(t)$, μ_k — orthonormal eigenfunctions and positive eigenvalues of the covariance operator $\mathbb{G}_X$:

$$\mu_k u_k = \mathbb{G}_X u_k \qquad \Longleftrightarrow \qquad \mu_k\, u_k(t) = \int_0^1 G_X(s,t)\, u_k(s)\, ds.$$

The small deviation problem ($\varepsilon \to 0$):

$$\mathbb{P}(\|X\|_2 < \varepsilon) = \mathbb{P}\left(\sum_{k=1}^{\infty} \mu_k \xi_k^2 < \varepsilon^2\right).$$

Main idea: all information about the process is contained in the spectrum.

What is already known?

1974 — G. Sytaya: implicit solution of the problem in terms of Laplace transform of $\sum \mu_k \xi_k^2$

from 1974 — V. M. Zolotarev, J. Hoffmann-Jorgensen , L. Shepp, R. Dudley, Il. A. Ibragimov, M. A. Lifshits, . . . :
simplification under different assumptions

1998 — T. Dunker, M. A. Lifshits, W. Linde (DLL):
Rather simple formulas for

$$\mathbb{P}\left(\sum \mu_k \xi_k^2 < \varepsilon^2\right) \qquad \text{when}$$

- μ_k — decreasing, logarifmically convex
- $\mu_k = k^{-d}, \quad d > 0,$ — polynomial decreasing
- $\mu_k = A^{-k}, \quad A > 0,$ — exponential decreasing

Useful fact: Wenbo Li principle

Let $\widehat{\mu}_k \approx \mu_k$ be some approximation.
Question: How the following small deviation probabilities are related

$$\mathbb{P}\left(\sum \mu_k \xi_k^2 < \varepsilon^2\right) \text{ and } \mathbb{P}\left(\sum \widehat{\mu}_k \xi_k^2 < \varepsilon^2\right)?$$

Theorem (The Wenbo Li principle 1992, Gao et al. 2003)

Let μ_k, $\widehat{\mu}_k$ — two summable sequences. If

$$\prod_{k=1}^{\infty} \frac{\widehat{\mu}_k}{\mu_k} < \infty, \tag{2}$$

then as $\varepsilon \to 0$

$$\mathbb{P}\left(\sum_{k=1}^{\infty} \mu_k \xi_k^2 < \varepsilon^2\right) \sim \mathbb{P}\left(\sum_{k=1}^{\infty} \widehat{\mu}_k \xi_k^2 < \varepsilon^2\right) \cdot \left(\prod \frac{\widehat{\mu}_k}{\mu_k}\right)^{1/2}$$

«good» approx. $\widehat{\mu}_k$ for μ_k	+	Wenbo Li principle	+	DLL theorem	=	small ball probability

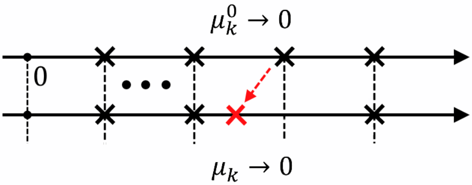
One-dimensional perturbation: first observation

$$G_X(s,t) = G_0(s,t) + D\psi(s)\psi(t), \qquad D \in \mathbb{R}$$

- $D = 0$ — unperturbed operator

The simplest case: $\psi(t)$ — eigenfunction of the integral operator $\mathbb{G}_0$

What happens if we change D?



Decrease $D \downarrow$

Asymptotically
$\mu_k^0 = \mu_k,\ k \to \infty$

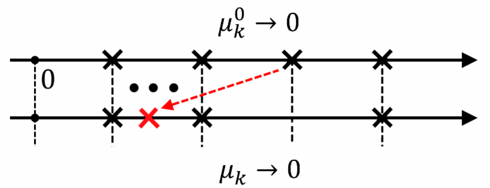
One-dimensional perturbation: first observation

$$G_X(s,t) = G_0(s,t) + D\psi(s)\psi(t), \qquad D \in \mathbb{R}$$

- $D = 0$ — unperturbed operator

The simplest case: $\psi(t)$ — eigenfunction of the integral operator $\mathbb{G}_0$

What happens if we change D?



Decrease $D \downarrow$

Asymptotically
$\mu_k^0 = \mu_k,\ k \to \infty$

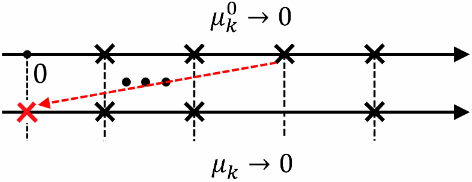
One-dimensional perturbation: first observation

$$G_X(s,t) = G_0(s,t) + D\psi(s)\psi(t), \qquad D \in \mathbb{R}$$

- $D = 0$ — unperturbed operator

The simplest case: $\psi(t)$ — eigenfunction of the integral operator $\mathbb{G}_0$

What happens if we change D?



D — critical

Asymptotically
$\mu_k^0 = \mu_{k-1}$, $k \to \infty$

Similar effect can be observed in a more general situation (when $\psi(t)$ is not necessarily the eigenfunction)

One-dimensional perturbation (A.Nazarov'2009)

Let $\quad Q := \langle \mathbb{G}_0^{-1}\psi, \psi\rangle < \infty \quad \Leftrightarrow \quad \psi \in \mathrm{Im}(\mathbb{G}_0^{1/2}).$
Exists critical value $D_{crit} = -1/Q$ such that:

Non critical

If $D > D_{crit} = -1/Q$,
then as $\varepsilon \to 0$

$$\mathbb{P}\left(\|X\|_2 < \varepsilon\right) \sim \frac{\mathbb{P}\left(\|X_0\|_2 < \varepsilon\right)}{|1 + QD|}.$$

Critical

If $D = D_{crit}$, $\boxed{\psi \in \mathrm{Im}(\mathbb{G}_0)}$,
then as $\varepsilon \to 0$

$$\mathbb{P}\left(\|X\|_2 < \varepsilon\right) \sim \frac{\sqrt{Q}}{\|\varphi\|_2} \cdot \sqrt{\frac{2}{\pi}} \cdot$$
$$\cdot \int_0^{\varepsilon^2} \frac{d}{dt}\mathbb{P}(\|X_0\|_2 < t) \cdot \frac{dt}{\sqrt{\varepsilon^2 - t^2}}$$

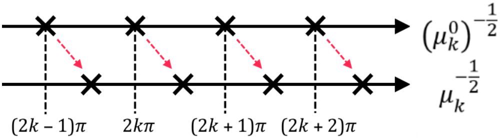
- In critical case there is an extra assumption $\psi \in \mathrm{Im}(\mathbb{G}_0)$
- also true for finite-dimensional perturbations (Yu. Petrova'2018)

Example 1: Durbin process for testing exponentiality

$G_0(s,t) = \min(s,t) - st, \quad \psi(t) = t\ln(t)$, then

$$G(s,t) = G_0(s,t) - \psi(s)\psi(t), \text{ critical, not «good» perturbation.}$$

Writing down the equation on the eigenvalues and solving it directly, we get $(\mu_k^0)^{-1/2} = \pi k$ and:



Spectral asymptotics: $\quad \mu_k^{-1/2} = \pi k + \frac{\pi}{2} + O\left(\frac{1}{k}\right), \quad k \to \infty$

Small ball probability: $\quad \mathbb{P}\{\|X\|_2 < \varepsilon\} \sim \frac{4}{\pi^{3/2}}\,\varepsilon^{-1}\exp\Big(-\frac{1}{8\varepsilon^2}\Big), \quad \varepsilon \to 0.$

Important example: test for normality — no general theorem works

X_1	$\widehat{\alpha}$ estimated, $\beta = 1$	$\psi_1(t) = f(F^{-1}(t))$	critical, not «good»
X_2	$\alpha = 0$, $\widehat{\beta}$ estimated	$\psi_2(t) = \psi_1(t) \cdot \frac{F^{-1}(t)}{\sqrt{2}}$	critical, not «good»
X_3	$\widehat{\alpha}$, $\widehat{\beta}$ estimated:	$\psi_1(t), \psi_2(t)$	critical, not «good»

Theorem: Kac-Kiefer-Wolfowitz processes

Theorem (A.Nazarov, Yu.Petrova'2015)

$$X_1: \qquad \omega_{2k-1} = 2\pi k + \frac{\pi}{\ln(k)} + O\left(\frac{\ln(\ln(k))}{\ln^2(k)}\right), \qquad \omega_{2k} = 2\pi k.$$

$$\mathbb{P}\Big\{\|X_1\| < \varepsilon\Big\} \sim C \cdot \varepsilon^{-1} \cdot \ln^{\frac{1}{2}}\left(\frac{1}{\varepsilon}\right) \cdot \exp\Big(-\frac{1}{8\varepsilon^2}\Big)$$

$$X_2: \qquad \omega_{2k-1} = 2\pi k - \pi, \qquad \omega_{2k} = 2\pi k + \pi + O\left(\frac{1}{\ln^2(k)}\right)$$

$$\mathbb{P}\Big\{\|X_2\| < \varepsilon\Big\} \sim \frac{2\sqrt{2}}{\pi^{3/2}}\,\varepsilon^{-1}\exp\Big(-\frac{1}{8\varepsilon^2}\Big)$$

Example 3: general Durbin processes

Lemma (Yu. Petrova '2018)

All Durbin processes are critical.

However, perturbations are «often» not «good». We considered Durbin processes when testing for distributions with parameters $\theta = (\alpha, \beta)$:

- Laplace $$F(x,\theta) = \begin{cases} \frac{1}{2}\exp(\frac{x-\alpha}{\beta}), & x \leqslant \alpha; \\ 1 - \frac{1}{2}\exp(-\frac{x-\alpha}{\beta}), & x > \alpha. \end{cases}$$
- logistic $$F(x,\theta) = \left(1 + \exp(-\tfrac{x-\alpha}{\beta})\right)^{-1}.$$
- Gumbel $$F(x,\theta) = \exp\left(-\exp\left(-\tfrac{x-\alpha}{\beta}\right)\right).$$
- Gamma $$F(x,\theta) = \begin{cases} \int\limits_0^{x/\beta} \dfrac{y^{\alpha-1}e^{-y}}{\Gamma(\alpha)}\,dy, & x \geqslant 0; \\ 0, & x < 0. \end{cases}$$

Note: all perturbations, but X_1 for logistic dist, are «bad»

Example 3: Gumbel distribution

Theorem (Yu. Petrova '2017)

For Durbin process $X(t)$ when testing for Gumbel distribution

$$G(s,t) = G_0(s,t) - \psi(t)\psi(s), \qquad \psi(t) = C\, t \ln(t) \cdot \ln(-\ln(t))$$

the asymptotics of corresponding eigenvalues is the following

$$\mu_k^{-1/2} = \pi k + \frac{\pi}{2} + (-1)^k \cdot 2\,\mathrm{arctg}\Big(\frac{1}{\ln(\ln(k)) + 1}\Big) - \frac{1}{\ln(k)\ln(\ln(k))} + \\ + O\Big(\frac{1}{\ln(k)(\ln(\ln(k)))^2}\Big).$$

And asymptotics of small ball probabilities

$$\mathbb{P}\Big\{\|X\| < \varepsilon\Big\} \sim C \cdot \ln^{-1}(\ln(\varepsilon^{-1})) \cdot \varepsilon^{-1} \cdot \exp\big(-\frac{1}{8\varepsilon^2}\big)$$

Spectral asymptotics $\longleftrightarrow$ small ball probabilities

Theorem (A. Nazarov, Yu. Petrova '2015)

If $\hat{\mu}_k = (\vartheta(k+\delta+F(k)))^{-2}$. *Then we have, as* $\varepsilon \to 0$,

$$\mathbb{P} \sim C \cdot \exp\left(\frac{1}{2} \cdot F_{-1}(\varepsilon^{-2})\right) \cdot \varepsilon^{-2\delta} \cdot \exp\left(-\left(\frac{\pi}{2\vartheta}\right)^2 \cdot \frac{\varepsilon^{-2}}{2}\right),$$

where $F(t)$ *is a slowly-varying function at infinity,* $F(t) \to 0$ *as* $t \to \infty$,

$$F_{-1}(t) = \int\limits_1^t \frac{F(x)}{x}\, dx.$$

Note: $\exp\left(\frac{1}{2} \cdot F_{-1}(t)\right)$ is also a slowly-varying function as $t \to \infty$.

Example 3: logistic, Gumbel distributions etc.

Theorem (Yu. Petrova '2017)

Small deviations probabilities for some Durbin processes:

LOG 1	$\frac{2\sqrt{15}}{\sqrt{\pi}} \cdot \varepsilon^{-2} \exp\left(-\frac{1}{8\varepsilon^2}\right)$
LOG 2	$\frac{4\sqrt{3+\pi^2}}{3\sqrt{2}\,\pi^{3/2}} \cdot \varepsilon^{-1} \exp\left(-\frac{1}{8\varepsilon^2}\right)$
LOG 3	$\frac{4\sqrt{15(3+\pi^2)}}{3\,\pi^{3/2}} \cdot \varepsilon^{-3} \exp\left(-\frac{1}{8\varepsilon^2}\right)$
GUM 1	$\frac{4}{\pi^{3/2}} \cdot \varepsilon^{-1} \exp\left(-\frac{1}{8\varepsilon^2}\right)$
GUM 2	$C \cdot \ln^{-1}(\ln(\varepsilon^{-1})) \cdot \varepsilon^{-1} \exp\left(-\frac{1}{8\varepsilon^2}\right)$
GUM 3	$C \cdot \exp\left(2\pi \ln^2(\ln(\varepsilon^{-1}))\right) \cdot \varepsilon^{-2} \exp\left(-\frac{1}{8\varepsilon^2}\right)$

Thank you for your attention!

Using standard methods we get the equation on the eigenvalues $\mu = \omega^{-2}$:

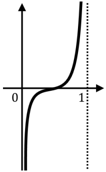
$$P(\mathcal{S}(\omega), \mathcal{C}(\omega), \mathcal{I}(\omega)) = 0, \qquad \omega \to \infty,$$

where

$$\mathcal{C}(\omega) = \int_0^{\frac{1}{2}} F_{st}^{-1}(t) \cos(\omega t)\, dt \qquad \mathcal{S}(\omega) = \int_0^{\frac{1}{2}} F_{st}^{-1}(t) \sin(\omega t)\, dt$$

$$\mathcal{I}(\omega) = \int_0^{\frac{1}{2}} \int_0^{\tau} F_{st}^{-1}(t) F_{st}^{-1}(\tau) \sin(\omega \tau) \cos(\omega t)\, dt\, d\tau$$

- $F_{st}(t)$ — standard normal DF
- $F_{st}^{-1}(t) \sim -\sqrt{-2\ln(t)}, \quad t \to 0,$
- $F_{st}^{-1}(t)$ has singularity at $t = 0$



Slowly varying functions = SVF

Definition

Function $V(t)$ is called SVF at infinity, if it doesn't change sign on some $[A, \infty)$, $A > 0$, and for any $\lambda > 0$

$$\lim_{t\to\infty} \frac{V(\lambda t)}{V(t)} = 1.$$

Function $V(t)$ is called SVF at zero, if $V(1/t)$ is SVF at infinity.
For example, $\ln^{\alpha}(t)$, $\alpha \in \mathbb{R}$.

Note: $F^{-1}(t)$ has the following properties:

- $V_0(t) := F^{-1}(t)$, $V_{n+1}(t) := tV_n'(t)$, $n \geq 0$, are SVF at zero.
- $F^{-1}\left(\frac{1}{2}\right) = 0$.

Note: for any SVF at zero: $tV'(t) = o(V(t))$ when $t \to 0$.
So $\forall n \geq 0 \quad V_{n+1}(t) = o(V_n(t))$.

Let

- $V_0(t)$ and $V_{n+1}(t) = t \cdot V_n'(t)$ be SVF at zero.
- $V_0(\frac{1}{2}) = 0$.

Theorem (A.Nazarov, Yu.Petrova'2015)

As $\omega \to \infty$:

$$\mathcal{C} = \int\limits_0^{\frac{1}{2}} V(t)\cos(\omega t)\,dt = \frac{1}{\omega}\sum_{k=1}^{N} c_k V_k\Big(\frac{1}{\omega}\Big) + R_N, \tag{3}$$

where

$$|R_N| \le C(V,N) \cdot \frac{\left|V_{N+1}(\frac{1}{\omega})\right|}{\omega}.$$

Example: $\int\limits_0^{1/2} \sqrt{-\ln(2t)}\cos(\omega t)\,dt = \frac{\pi}{2\ln^{1/2}(2\omega)} - \frac{\gamma\pi}{2\ln^{3/2}(2\omega)} + O\left(\frac{1}{\ln^{5/2}(\omega)}\right)$

Asymptotics of integrals

Theorem (A.Nazarov, Yu.Petrova'2015)

$$\int_0^{\frac{1}{2}}\int_0^{\tau} V(t)V(\tau)\sin(\omega\tau)\cos(\omega t)\,dt\,d\tau =$$

$$= \frac{1}{2\omega}\int_0^{\frac{1}{2}} V^2(t)\,dt + \sum_{n=2}^{N}\sum_{\substack{k+m=n\\ k,m\geq 1}} a_{k,m}\frac{V_k(\frac{1}{\omega})V_m(\frac{1}{\omega})}{\omega^2} + R_N,$$

where $|R_N| \leq C(V,N)\sum_{\substack{i+j=N+1\\ i,j\geq 1}} \frac{|V_i(\frac{1}{\omega})V_j(\frac{1}{\omega})|}{\omega^2}.$

Thank you again!

Durbin processes

Definition

A sample $x_1, \dots, x_n \sim F(x, \theta)$
$\theta = (\theta_1, \dots, \theta_m)$ — parameters of the distribution. Let's consider:

$$\hat{F}_n(t) = \frac{\{\text{number of } x_i \colon F(x_i, \hat{\theta}_n) \leqslant t\}}{n}, \quad \hat{\theta}_n \text{ — estimated from the data}$$

Then Durbin process $DP(t)$ is defined as limit

$$n^{1/2}\big[\hat{F}_n(t) - t\big] \xrightarrow{w} DP(t)$$

It is a Gaussian process with zero mean and covariance function:

$$G(s,t) = G_B(s,t) - \vec{\psi}^{\,T}(s)\, S^{-1}\, \vec{\psi}(t)$$

- $G_B(s,t) = \min(s,t) - st$
- $S_{ij} = \mathbb{E}\left(\frac{\partial}{\partial \theta_i} \ln(f(x,\theta)) \frac{\partial}{\partial \theta_j} \ln(f(x,\theta))\right)\Big|_{\theta=\theta_0}$ — Fisher information
- $\psi_j(t) = \left.\dfrac{\partial F}{\partial \theta_j}\right|_{\theta=\theta_0}, \ (j = 1 \dots m), \quad \theta_0$ — fixed vector of parameters