

# Konzistencia hĺbky funkcií II

Stanislav Nagy

Katedra pravděpodobnosti a matematické statistiky  
Univerzita Karlova v Praze

KU Leuven  
Dept. of Mathematics

Robust 2014

This work was supported by Research Foundation – Flanders (FWO-Vlaanderen)

1 Konzistencia hĺbky funkcií I

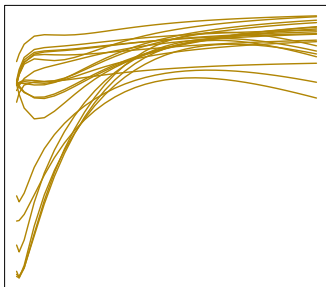
2 Integrálne hĺbky

3 Infimálne hĺbky

# Funkcionálne dáta

$X \sim P \in \mathcal{P}(C([0, 1]))$  a náhodný výber  $X_1, \dots, X_n$  z  $P$ . Uvažujme hĺbku funkcií náhodného výberu voči  $P$  (resp.  $P_n$ )

$$D: C([0, 1]) \times \mathcal{P}(C([0, 1])) \rightarrow [0, 1].$$



# Hĺbky pre funkcie

Tri základné typy hĺbok pre funkcie:

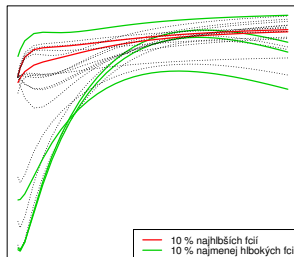
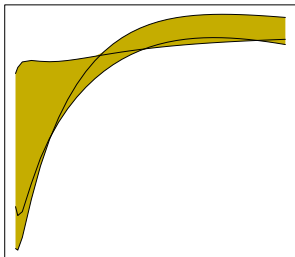
- **Pásové hĺbky** (López-Pintado, Romo (2009))
- **Integrálne hĺbky** (Fraiman, Muniz (2001))
- **Infimálne hĺbky** (Mosler, Polyakova (2013))

# Pásové hĺbky

López-Pintado, Romo (2009): pre  $J = 2, 3, \dots$

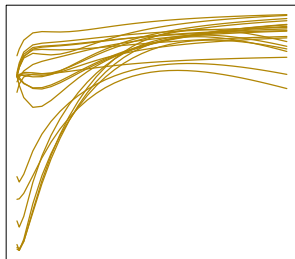
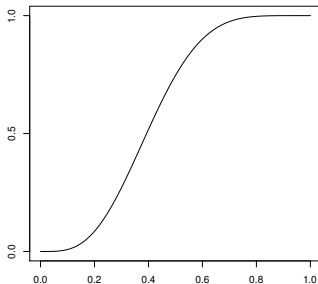
$$BD^{(J)}(x; P) = \frac{1}{J-1} \sum_{j=2}^J P(G(x) \subset B(X_1, X_2, \dots, X_j)),$$

kde  $G(x)$  je **graf funkcie**  $x$  a  $B(x_1, x_2, \dots, x_j)$  je **pás funkcií**



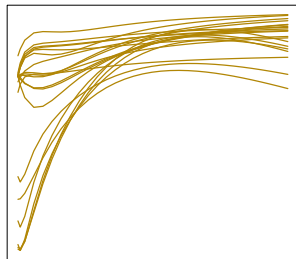
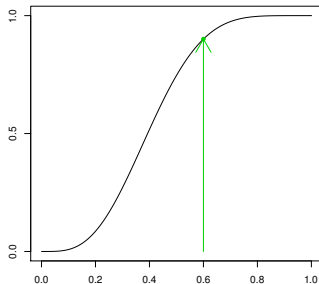
# Integrálne hĺbky

$$D_1(u; Q) = 1 - |0.5 - F_Q(u)|$$



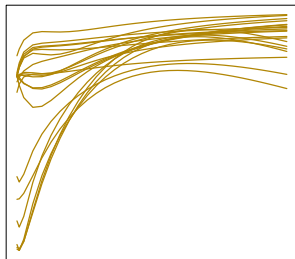
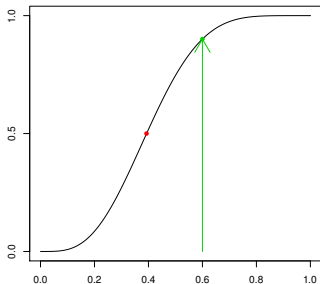
# Integrálne hĺbky

$$D_1(u; Q) = 1 - |0.5 - F_Q(u)|$$



# Integrálne hĺbky

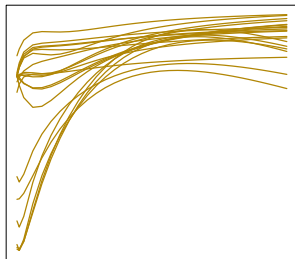
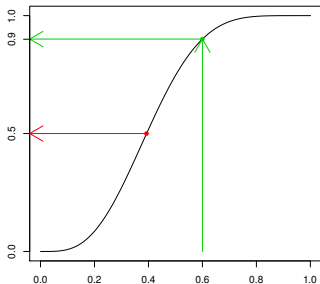
$$D_1(u; Q) = 1 - |0.5 - F_Q(u)|$$





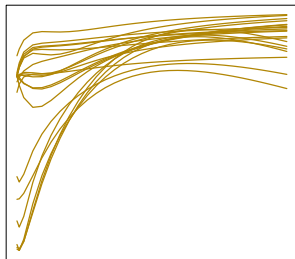
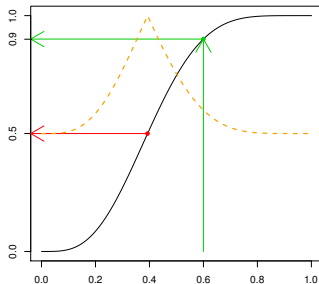
# Integrálne hĺbky

$$D_1(u; Q) = 1 - |0.5 - F_Q(u)|$$



# Integrálne hĺbky

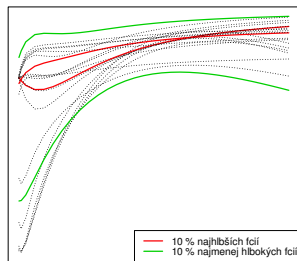
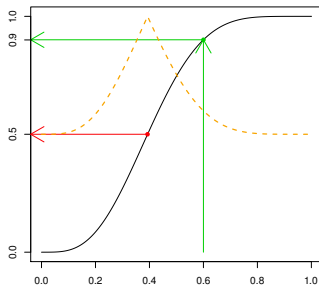
$$D_1(u; Q) = 1 - |0.5 - F_Q(u)|$$



# Integrálne hĺbky

Frailman, Muniz (2001):

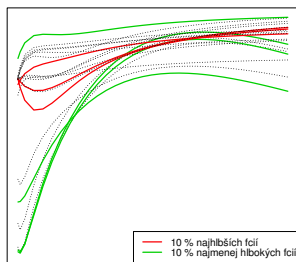
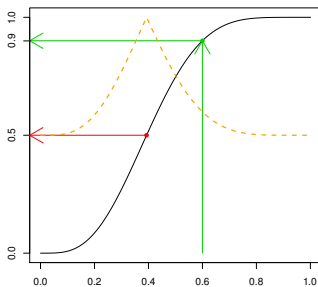
$$D_{FM}(x; P) = \int_0^1 D_1(x(t); P_t) dt.$$



# Infimálne hĺbky

Mosler, Polyakova (2013):

$$D_{MP}(x; P) = \inf_{t \in [0,1]} D_1(x(t); P_t).$$



# Slabá a silná konzistencia

Hĺbka  $D$  je na množine  $S \subset C([0, 1])$  **konzistentná**

- **slabo rovnomerne** ak

$$\sup_{x \in S} |D(x; P_n) - D(x; P)| \xrightarrow[n \rightarrow \infty]{P} 0,$$

- **slabo univerzálne** ak

$$\sup_{x \in S} |D(x; P_n) - D(x; P)| \xrightarrow[n \rightarrow \infty]{P} 0 \text{ pre každé } P \in \mathcal{P}(C([0, 1])).$$

## Slabá a silná konzistencia

Hĺbka  $D$  je na množine  $S \subset C([0, 1])$  **konzistentná**

- **rovnomerne** ak

$$\sup_{x \in S} |D(x; P_n) - D(x; P)| \xrightarrow[n \rightarrow \infty]{\text{s.i.}} 0,$$

- **univerzálne** ak

$$\sup_{x \in S} |D(x; P_n) - D(x; P)| \xrightarrow[n \rightarrow \infty]{\text{s.i.}} 0 \text{ pre každé } P \in \mathcal{P}(C([0, 1])).$$

# Konzistencia integrálnych hĺbok

Fraiman, Muniz (2001), Theorem 3.1:

**Tvrdenie:**

*Nech pre  $X \sim P$  platia nasledujúce podmienky:*

- *existuje  $L > 0$  tak, že trajektórie procesu  $X$  sú skoro iste  $L$ -Lipschitzovské,*
- *existuje  $C > 0$  tak, že pre každú  $u \in C([0, 1])$   $L$ -Lipschitzovskú a každé  $\varepsilon > 0$  platí*

$$E[\lambda(\{t \in [0, 1]: X(t) \in [u(t); u(t) + C\varepsilon]\})] < \varepsilon/2.$$

*Potom  $D_{FM}$  je rovnomerne konzistentná voči  $P$  na množine  $L$ -Lipschitzovských funkcií.*

# Konzistencia integrálnych hĺbok: Dôkaz

**Dôkaz:** Fraiman a Munizová dokazujú rovnomernú konzistenciu

$$\sup_{x \in \text{Lip}_L} \left| \int_0^1 F_{n,t}(x(t)) - F_t(x(t)) dt \right| \xrightarrow[n \rightarrow \infty]{\text{s.i.}} 0,$$

z čoho **okamžite plynie**

$$\sup_{x \in \text{Lip}_L} |D_{FM}(x; P_n) - D_{FM}(x; P)| \xrightarrow[n \rightarrow \infty]{\text{s.i.}} 0.$$





# Dokazovanie konzistencie

Opakovane budeme používať pre  $u \in \mathbb{R}$  a  $Q \in \mathcal{P}(\mathbb{R})$  vzťah

$$\begin{aligned}
 |D_1(u; Q) - D_1(u; Q_n)| &= |1 - |0.5 - F_Q(u)| - (1 - |0.5 - F_{Q_n}(u)|)| \\
 &\leq |F_Q(u) - F_{Q_n}(u)|.
 \end{aligned}
 \tag{*}$$

# Silná konzistencia hĺbky

Rozpíšme pomocou (\*)

$$\begin{aligned} \sup_{x \in \text{Lip}_L} |D_{FM}(x; P) - D_{FM}(x; P_n)| &= \sup_{x \in \text{Lip}_L} \left| \int_0^1 D_1(x(t); P_t) - D_1(x(t); P_{n,t}) dt \right| \\ &\leq \sup_{x \in \text{Lip}_L} \int_0^1 |D_1(x(t); P_t) - D_1(x(t); P_{n,t})| dt \\ &\leq \sup_{x \in \text{Lip}_L} \int_0^1 |F_{n,t}(x(t)) - F_t(x(t))| dt, \end{aligned}$$

# Silná konzistencia hĺbky

Rozpíšme pomocou (\*)

$$\begin{aligned} \sup_{x \in \text{Lip}_L} |D_{FM}(x; P) - D_{FM}(x; P_n)| &= \sup_{x \in \text{Lip}_L} \left| \int_0^1 D_1(x(t); P_t) - D_1(x(t); P_{n,t}) dt \right| \\ &\leq \sup_{x \in \text{Lip}_L} \int_0^1 |D_1(x(t); P_t) - D_1(x(t); P_{n,t})| dt \\ &\leq \sup_{x \in \text{Lip}_L} \int_0^1 |F_{n,t}(x(t)) - F_t(x(t))| dt, \end{aligned}$$

my ale máme iba

$$\sup_{x \in \text{Lip}_L} \left| \int_0^1 F_{n,t}(x(t)) - F_t(x(t)) dt \right| \xrightarrow[n \rightarrow \infty]{\text{s.i.}} 0.$$

# Slabá konzistencia hĺbky

Podobne vďaka (\*) a Fubiniho vete (merateľnosť Nagy et al. (2013))

$$\mathbb{E} \left[ \sup_{x \in \mathcal{C}([0,1])} |D_{FM}(x; P) - D_{FM}(x; P_n)| \right] \leq \int_0^1 \mathbb{E} \left[ \sup_{u \in \mathbb{R}} |F_{n,t}(u) - F_t(u)| \right] dt,$$

integrand konverguje pre každé  $t$  skoro iste k 0 (Glivenko-Cantelli) a  $D_{FM}$  je **slabo univerzálne konzistentná**.

# Slabá konzistencia hĺbky

Podobne vďaka (\*) a Fubiniho vete (merateľnosť Nagy et al. (2013))

$$\mathbb{E} \left[ \sup_{x \in \mathcal{C}([0,1])} |D_{FM}(x; P) - D_{FM}(x; P_n)| \right] \leq \int_0^1 \mathbb{E} \left[ \sup_{u \in \mathbb{R}} |F_{n,t}(u) - F_t(u)| \right] dt,$$

integrand konverguje pre každé  $t$  skoro iste k 0 (Glivenko-Cantelli) a  $D_{FM}$  je **slabo univerzálne konzistentná**.

Platí niečo podobné aj **v silnom zmysle?**

# Konvergencia náhodných procesov

Majme postupnosť náhodných procesov

$$X_n = \sup_{u \in \mathbb{R}} |F_{n,t}(u) - F_t(u)| : \Omega \times [0, 1] \rightarrow \mathbb{R}.$$

Vieme, že  $X_n(\cdot, t) \xrightarrow[n \rightarrow \infty]{P_t\text{-s.i.}} 0$  pre každé  $t \in [0, 1]$ , t.j.

$$X_n(\omega, t) \xrightarrow[n \rightarrow \infty]{} 0 \text{ pre každé } \omega \notin N_t, P(N_t) = 0.$$

# Konvergencia náhodných procesov

Majme postupnosť náhodných procesov

$$X_n = \sup_{u \in \mathbb{R}} |F_{n,t}(u) - F_t(u)| : \Omega \times [0, 1] \rightarrow \mathbb{R}.$$

Vieme, že  $X_n(\cdot, t) \xrightarrow[n \rightarrow \infty]{P\text{-s.i.}} 0$  pre každé  $t \in [0, 1]$ , t.j.

$$X_n(\omega, t) \xrightarrow[n \rightarrow \infty]{} 0 \text{ pre každé } \omega \notin N_t, P(N_t) = 0.$$

Znamená to  $X_n(\cdot, t) \xrightarrow[n \rightarrow \infty]{P\text{-s.i.}} 0$  pre každé  $t \in [0, 1]$ ? T.j. platí

$$X_n(\omega, t) \xrightarrow[n \rightarrow \infty]{} 0 \text{ pre každé } t \in [0, 1] \text{ a } \omega \notin N, P(N) = 0?$$

# Konvergencia náhodných procesov

Majme postupnosť náhodných procesov

$$X_n = \sup_{u \in \mathbb{R}} |F_{n,t}(u) - F_t(u)| : \Omega \times [0, 1] \rightarrow \mathbb{R}.$$

Vieme, že  $X_n(\cdot, t) \xrightarrow[n \rightarrow \infty]{P_t\text{-s.i.}} 0$  pre každé  $t \in [0, 1]$ , t.j.

$$X_n(\omega, t) \xrightarrow[n \rightarrow \infty]{} 0 \text{ pre každé } \omega \notin N_t, P(N_t) = 0.$$

Znamená to  $X_n(\cdot, t) \xrightarrow[n \rightarrow \infty]{P\text{-s.i.}} 0$  pre každé  $t \in [0, 1]$ ? T.j. platí

$$X_n(\omega, t) \xrightarrow[n \rightarrow \infty]{} 0 \text{ pre každé } t \in [0, 1] \text{ a } \omega \notin N, P(N) = 0?$$

Zrejme  $N \subset \bigcup_{t \in [0,1]} N_t$ , z toho ale stále **nevieme ukázať**  $P(N) = 0$ !



# Konzistentná verzia procesu

Pre každý proces  $X_n$  definujeme  $\tilde{X}_n$  ako

$$\tilde{X}_n(\omega, t) = \begin{cases} X_n(\omega, t) & \text{pre } \omega \notin N_t, \\ 0 & \text{pre } \omega \in N_t. \end{cases}$$

Potom  $P(X_n(\cdot, t) = \tilde{X}_n(\cdot, t)) = 1$  pre každé  $t \in [0, 1]$ , teda  $\tilde{X}_n$  je **verziou procesu**  $X_n$  (Gikhman, Skorokhod [2]) takou, že

$$\tilde{X}_n(\omega, t) \xrightarrow[n \rightarrow \infty]{} 0 \text{ pre každé } \omega \in \Omega \text{ a } t \in [0, 1].$$

Zrejme teda

$$\int_0^1 \tilde{X}_n(\omega, t) dt \xrightarrow[n \rightarrow \infty]{} 0 \text{ pre každé } \omega \in \Omega.$$

# Stochastická ekvivalencia verzíí

Teraz môžeme pre každú P-merateľnú množinu  $B \subset \Omega$  písať

$$\begin{aligned} \int_B \int_0^1 X_n(\omega, t) dt dP(\omega) &= \int_0^1 \int_B X_n(\omega, t) dP(\omega) dt \\ &= \int_0^1 \int_B \tilde{X}_n(\omega, t) dP(\omega) dt \\ &= \int_B \int_0^1 \tilde{X}_n(\omega, t) dt dP(\omega), \end{aligned}$$

teda

$$\int_0^1 X_n(., t) dt = E \left[ \int_0^1 X_n(., t) dt \middle| \mathcal{F} \right] = \int_0^1 \tilde{X}_n(., t) dt$$

kde posledná rovnosť platí P-skoro iste.

## Stochastická ekvivalencia verzíí

Teraz môžeme pre každú P-merateľnú množinu  $B \subset \Omega$  písať

$$\begin{aligned} \int_B \int_0^1 X_n(\omega, t) dt dP(\omega) &= \int_0^1 \int_B X_n(\omega, t) dP(\omega) dt \\ &= \int_0^1 \int_B \tilde{X}_n(\omega, t) dP(\omega) dt \\ &= \int_B \int_0^1 \tilde{X}_n(\omega, t) dt dP(\omega), \end{aligned}$$

teda

$$\int_0^1 X_n(\cdot, t) dt = E \left[ \int_0^1 X_n(\cdot, t) dt \middle| \mathcal{F} \right] = \int_0^1 \tilde{X}_n(\cdot, t) dt$$

kde posledná rovnosť platí **P-skoro iste**. Dostávame

$$\int_0^1 X_n(\omega, t) dt \xrightarrow{n \rightarrow \infty} 0 \text{ pre P-skoro všetky } \omega \in \Omega.$$

## Konzistencia integrálnych hĺbok 2

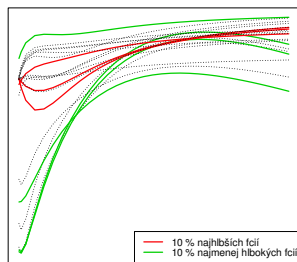
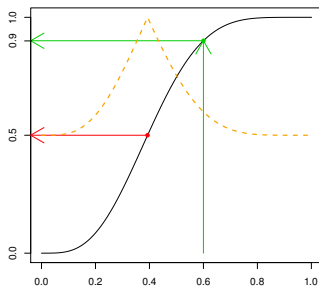
Tvrdenie:

$D_{FM}$  je univerzálne konzistentná hĺbka na celom priestore  $C([0, 1])$ .

# Infimálne hĺbky

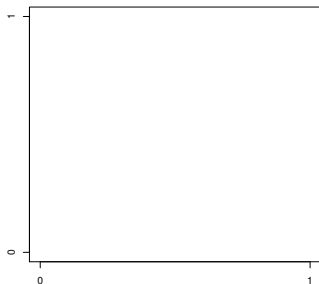
Mosler, Polyakova (2013):

$$D_{MP}(x; P) = \inf_{t \in [0,1]} D_1(x(t); P_t).$$



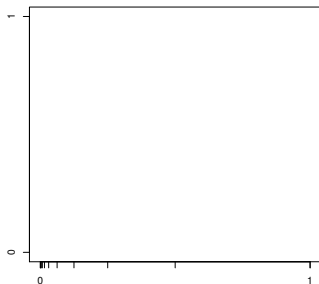
## Konzistencia infimálnych hĺbok: Príklad 1

Rozdeľme interval  $[0, 1]$  pomocou postupnosti  $u_j = 2^{-j}$ ,  $j = 0, 1, \dots$



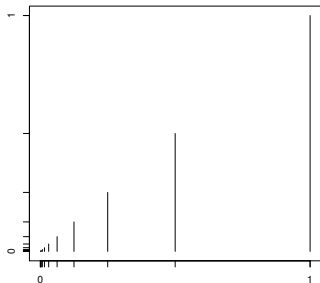
## Konzistencia infimálnych hĺbok: Príklad 1

Rozdeľme interval  $[0, 1]$  pomocou postupnosti  $u_j = 2^{-j}$ ,  $j = 0, 1, \dots$



# Konzistencia infimálnych hĺbok: Príklad 1

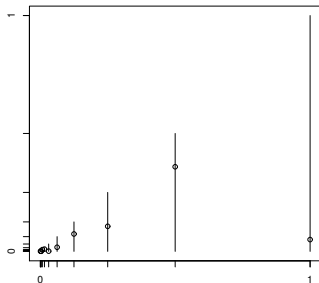
Definujme  $X(u_j) \sim \text{Unif}([0, 2^{-j}])$  **nezávisle od seba** pre každé  $j$  a  $X(0) = 0$ .





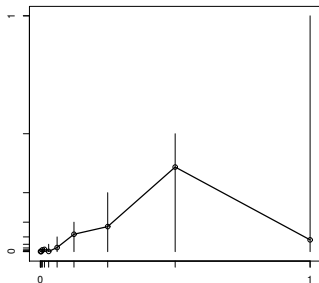
# Konzistencia infimálnych hĺbok: Príklad 1

Definujme  $X(u_j) \sim \text{Unif}([0, 2^{-j}])$  **nezávisle od seba** pre každé  $j$  a  $X(0) = 0$ .



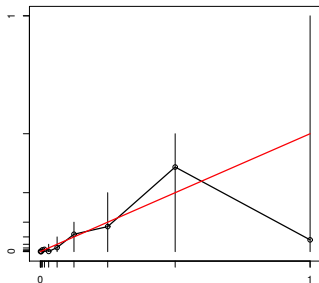
# Konzistencia infimálnych hĺbok: Príklad 1

Dodefinujme  $X$  **lineárne**,  $X \sim P \in \mathcal{P}(C([0, 1]))$ .



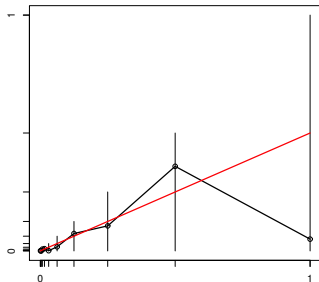
# Konzistencia infimálnych hĺbok: Príklad 1

Počítajme hĺbku funkcie  $x(t) = t/2$ .



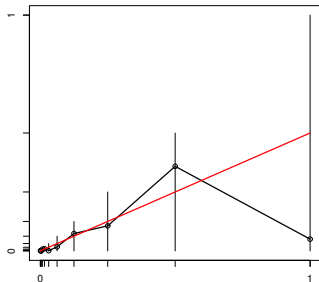
# Konzistencia infimálnych hĺbok: Príklad 1

$$\inf_{t \in (0,1]} 1 - |0.5 - F_t(x(t))| \geq 0.75$$



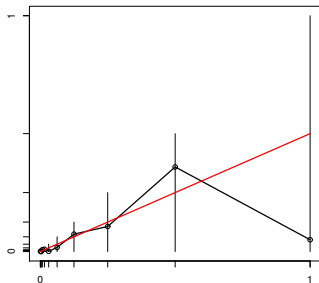
# Konzistencia infimálnych hĺbok: Príklad 1

$$\inf_{t \in (0,1]} 1 - |0.5 - F_{n,t}(x(t))| = 0.5$$



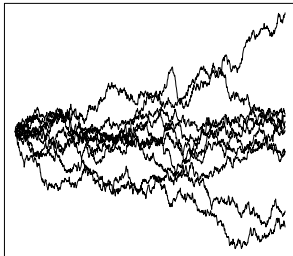
# Konzistencia infimálnych hĺbok: Príklad 1

**Infimálna hĺbka nie je slabo univerzálne konzistentná.**



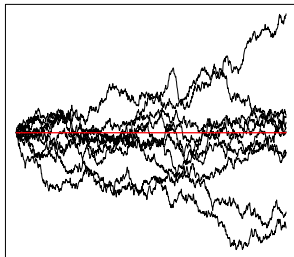
## Konzistencia infimálnych hĺbok: Príklad 2

Uvažujme náhodný výber **Wienerovych procesov** na  $[0, 1]$ .



## Konzistencia infimálnych hĺbok: Príklad 2

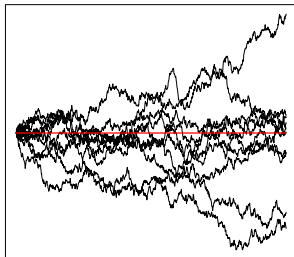
Počítajme hĺbku funkcie  $x(t) \equiv 0$ .





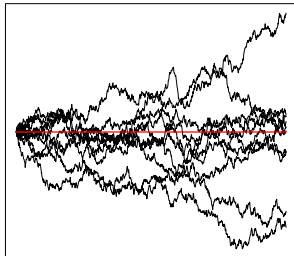
## Konzistencia infimálnych hĺbok: Príklad 2

$$\inf_{t \in (0,1]} 1 - |0.5 - F_t(x(t))| = 1$$



## Konzistencia infimálnych hĺbok: Príklad 2

$$\inf_{t \in (0,1]} 1 - |0.5 - F_{n,t}(x(t))| = 0.5$$



# Konzistencia infimálnej hĺbky

Použime opäť (\*):

$$\begin{aligned}
 & \sup_{x \in C([0,1])} \left| \inf_{t \in [0,1]} D_1(x(t); P_t) - \inf_{t \in [0,1]} D_1(x(t); P_{n,t}) \right| \\
 & \leq \sup_{x \in C([0,1])} \sup_{t \in [0,1]} |D_1(x(t); P_t) - D_1(x(t); P_{n,t})| \\
 & \leq \sup_{t \in [0,1], x \in \mathbb{R}} |F_{n,t}(x) - F_t(x)|
 \end{aligned}$$

# Konzistencia infimálnej hĺbky

Použime opäť (\*):

$$\begin{aligned} & \sup_{x \in C([0,1])} \left| \inf_{t \in [0,1]} D_1(x(t); P_t) - \inf_{t \in [0,1]} D_1(x(t); P_{n,t}) \right| \\ & \leq \sup_{x \in C([0,1])} \sup_{t \in [0,1]} |D_1(x(t); P_t) - D_1(x(t); P_{n,t})| \\ & \leq \sup_{t \in [0,1], x \in \mathbb{R}} |F_{n,t}(x) - F_t(x)| \end{aligned}$$

Potrebuje **rovnomerne obmedziť nespočetnú kolekciu ecdf procesov.**

## Rovnomerné Glivenkove-Cantelliho vety?

Kolekcia ecdf procesov je **rovnomerne GC** (napr. [6, Kapitola 2.8]), t.j.

$$\sup_{t \in [0,1]} P \left( \sup_{m \geq n} \sup_{x \in \mathbb{R}} |F_{m,t}(x) - F_t(x)| > \varepsilon \right) \xrightarrow{n \rightarrow \infty} 0 \text{ pre každé } \varepsilon > 0.$$

## Rovnomerné Glivenkove-Cantelliho vety?

Kolekcia ecdf procesov je **rovnomerne GC** (napr. [6, Kapitola 2.8]), t.j.

$$\sup_{t \in [0,1]} P \left( \sup_{m \geq n} \sup_{x \in \mathbb{R}} |F_{m,t}(x) - F_t(x)| > \varepsilon \right) \xrightarrow{n \rightarrow \infty} 0 \text{ pre každé } \varepsilon > 0.$$

My ale potrebujeme **oveľa silnejší výsledok**

$$P \left( \sup_{m \geq n} \sup_{t \in [0,1], x \in \mathbb{R}} |F_{m,t}(x) - F_t(x)| > \varepsilon \right) \xrightarrow{n \rightarrow \infty} 0 \text{ pre každé } \varepsilon > 0.$$

# Rovnomerné Glivenkove-Cantelliho vety?

Kolekcia ecdf procesov je **rovnomerne GC** (napr. [6, Kapitola 2.8]), t.j.

$$\sup_{t \in [0,1]} P \left( \sup_{m \geq n} \sup_{x \in \mathbb{R}} |F_{m,t}(x) - F_t(x)| > \varepsilon \right) \xrightarrow{n \rightarrow \infty} 0 \text{ pre každé } \varepsilon > 0.$$

My ale potrebujeme **oveľa silnejší výsledok**

$$P \left( \sup_{m \geq n} \sup_{t \in [0,1], x \in \mathbb{R}} |F_{m,t}(x) - F_t(x)| > \varepsilon \right) \xrightarrow{n \rightarrow \infty} 0 \text{ pre každé } \varepsilon > 0.$$

**Vieme toto dokázať?**

# Uniformná Glivenkova-Cantelliho vlastnosť ecdf procesu

Tvrdenie:

Nech  $X \sim P \in \mathcal{P}(C([0, 1]))$  je zmesou  $P^{(1)}$  a  $P^{(2)}$  takých, že:

- 1 Trajektórie  $X_1 \sim P^{(1)}$  skoro iste ležia v rovnako spojitej množine funkcií a existuje  $K > 0$  tak že pre všetky  $t \in [0, 1]$  je  $X_1(t) \sim P_t^{(1)} \in \mathcal{P}(\mathbb{R})$  abs. spojitě s hustotou obmedzenou  $K$ .
- 2  $X_2 \sim P^{(2)}$  je koncentrovaná v konečnorozmernom podpriestore  $C([0, 1])$ .

Potom

$$\sup_{t \in [0, 1], x \in \mathbb{R}} |F_{n,t}(x) - F_t(x)| \xrightarrow[n \rightarrow \infty]{\text{s.i.}} 0.$$



# Konzistencia infimálnych hĺbok

Tvrdenie:

*Ak  $P$  je rozdelenie ako v predchádzajúcom, potom je  $D_{MP}$  rovnomerne konzistentná hĺbka voči  $P$  na celom priestore  $C([0, 1])$ .*

# Záver

Ukázali sme, že:

- **integrálne hĺbky** sú **vždy silne univerzálne konzistentné**, ale
- **infimálne hĺbky** sú konzistentné **iba za istých predpokladov**.

# Záver

Ukázali sme, že:

- **integrálne hĺbky** sú **vždy silne univerzálne konzistentné**, ale
- **infimálne hĺbky** sú konzistentné **iba za istých predpokladov**.

Do budúcnosti:

- Charakterizácia funkcionálov **stochasticky ekvivalentných s každou verziou náh. procesu**.
- Funkcionálne **uniformné Glivenkove-Cantelliho vety**.

# Literatúra

-  Ricardo Fraiman and Graciela Muniz. Trimmed means for functional data. *Test*, 10(2):419–440, 2001.
-  I. I. Gikhman and A. V. Skorokhod. *Introduction to the theory of random processes*. Translated from the Russian by Scripta Technica, Inc. W. B. Saunders Co., Philadelphia, Pa., 1969.
-  Sara López-Pintado and Juan Romo. On the concept of depth for functional data. *J. Amer. Statist. Assoc.*, 104(486):718–734, 2009.
-  Karl Mosler and Yulia Polyakova. General notions of depth for functional data. *arXiv preprint arXiv:1208.1981*, 2013.
-  Stanislav Nagy, Marek Omelka, and Daniel Hlubinka. Integrated data depth for vector valued functions. *Submitted*, 2013.
-  Aad W. van der Vaart and Jon A. Wellner. *Weak convergence and empirical processes*. Springer Series in Statistics. Springer-Verlag, New York, 1996.