MATEMATICKO-FYZIKÁLNÍ FAKULTA
Univerzita Karlova

BAKALÁŘSKÁ PRÁCE

Zdeněk Pustějovský

Glivenkova-Cantelliho věta a její zobecnění

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: doc. Ing. Marek Omelka, Ph.D
Studijní program: Matematika
Studijní obor: Obecná matematika

Praha 2021

Prohlašuji, že jsem tuto bakalářskou práci vypracoval(a) samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů. Tato práce nebyla využita k získání jiného nebo stejného titulu.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V .... dne ....

Podpis autora

Zde bych rád upřímně poděkoval doc. Ing. Marku Omelkovi, Ph.D za vstřícnost, ochotu a cenné rady, které mi věnoval po celou dobu zpracování této práce.

Název práce: Glivenkova-Cantelliho věta a její zobecnění

Autor: Zdeněk Pustějovský

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: doc. Ing. Marek Omelka, Ph.D, Katedra pravděpodobnosti a matematické statistiky

Abstrakt: V této práci se zabýváme Glivenkovou-Cantelliho větou a jejím zobecněním. Nejprve dokážeme klasickou verzi této věty s empirickou distribuční funkcí a jako její důsledek ukážeme stejnoměrnou konvergenci výběrových kvantilů ke skutečným. Dále zadefinujeme pojem bracketing numbers a dokážeme zobecněnou Glivenkovu-Cantelliho větu pro třídy funkcí s konečným bracketing number. Následně ukážeme, jak ze zobecněné verze plyne ta klasická nejen pro reálné náhodné veličiny, ale i pro náhodné vektory. Nakonec uvádíme některé příklady Glivenkových-Cantelliho tříd funkcí. V celé práci klademe důraz také na ukázky aplikací dokázaných vět.

Klíčová slova: empirická distribuční funkce, Glivenkova-Cantelliho věta, bracketing numbers, konvergence skoro jistě

Title: Glivenko-Cantelli theorem and its generalization

Author: Zdeněk Pustějovský

Department: Department of probability and mathematical statistics

Supervisor: doc. Ing. Marek Omelka, Ph.D, Department of probability and mathematical statistics

Abstract: In this thesis we look at Glivenko-Cantelli theorem and its generalization. Firstly we prove the classical version of this theorem with empirical ditribution function and as its corollary we show uniform convergence of sample quantiles to the actual ones. Next we give definition of the bracketing numbers and prove generalized version of Glivenko-Cantelli theorem for function classes with finite bracketing number. Then we show how from the generalized version of the theorem follows the classical one not only for real random variables, but also for random vectors. Lastly we give examples of some Glivenko-Cantelli classes of functions. Throughout the work we also have applications of proven theorems in mind.

Keywords: empirical distribution function, Glivenko-Cantelli theorem, bracketing numbers, almost sure convergence

Obsah

Úvod

Pokud chceme z náhodného výběru $X_1, .., X_n$ reálných náhodných veličin, o jehož rozdělení nic nevíme, odhadnout jeho distribuční funkci F, dobrou volbou je tzv. empirická distribuční funkce

$$F_n(t) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}[X_i \leq t].$$

Bodovou konvergenci F_n k F zaručuje silný zákon velkých čísel. Stejnoměrnou konvergenci F_n k F na $\mathbb{R}$ dokázali Glivenko (1933) pro F spojitou a Cantelli (1933) pro F obecnou, odtud tedy věta nese svůj název.

Uvažujme teď nějakou třídu měřitelných reálných funkcí $\mathcal{F}$ a náhodný výběr $X_1, ..., X_n$ z rozdělení P a pro $f \in \mathcal{F}$ označme $Pf = \mathbb{E}f(X_1), \mathbb{P}_n f = \frac{1}{n}\sum_{i=1}^{n} f(X_i)$. Ze silného zakona velkých čísel pro nezávislé, stejně rozdělené náhodné veličiny pro libovolnou $f \in \mathcal{F}$ dostaneme $\mathbb{P}_n f \xrightarrow[n\to\infty]{s.j.} Pf$. Řekneme, že $\mathcal{F}$ je Glivenkova-Cantelliho, pokud $\mathbb{P}_n$ konverguje k P stejnoměrně na $\mathcal{F}$, tedy

$$\sup_{f\in\mathcal{F}} |\mathbb{P}_n f - Pf| \xrightarrow[n\to\infty]{s.j.*} 0,$$

kde $s.j.*$ je určité zobecnění konvergence skoro jistě (viz Definice 1). O $\mathcal{F}$ totiž nic nevíme a většinou supremum přes $\mathcal{F}$ nebude měřitelné, tedy standardní konvergence $s.j.$ nedává smysl. Zobecněná Glivenkova-Cantelliho věta říká, že $\mathcal{F}$ je Glivenkova-Cantelliho třída, pokud je $\mathcal{F}$ v nějakém smyslu malá. Tato velikost lze měřit pomocí tzv. bracketing numbers.

Cílem této práce je podrobně dokázat obě verze Glivenkových-Cantelliho vět, protože v literatuře bývají jejich důkazy napsány velmi stručně, také uvést jejich aplikace a některé příklady tříd funkcí, které jsou Glivenkovy-Cantelliho.

První kapitola se věnuje klasické verzi věty. Nejprve dokážeme pomocná tvrzení, která zaručí, že tvrzení věty dává smysl, nebo je budeme potřebovat v důkazu samotné věty. Ve druhé půlce kapitoly uvedeme jako důležitý důsledek větu o stejnoměrné konvergenci výběrových α-kvantilů ke skutečným pro α na vhodné množině $I \subseteq [0,1]$ a odvodíme tvar I pro některá rozdělení.

Druhá kapitola začíná definicí G-C tříd funkcí, konvergence $s.j.*$ a bracketing number. Následně vyslovíme a dokážeme zobecněnou G-C větu a ukážeme, jak z ní plyne její klasická verze i pro náhodné vektory. Dále ukážeme nějaké aplikace této věty, např. stejnoměrnou konvergenci empirické charakteristické funkce ke skutečné na libovolné omezené množině $A \subset \mathbb{R}$. Nakonec uvedeme některé příklady G-C funkcí, zejména funkce spojité v parametru na kompaktní množině a monotónní funkce.

1. Klasická Glivenkova-Cantelliho věta

1.1 Empirická distribuční funkce

V této kapitole podrobně dokážeme Glivenkovu-Cantelliho větu a odvodíme z ní stejnoměrnou konzistenci výběrových kvantilů. Situace je následující. Máme $X_i : \Omega \to \mathbb{R}, i \in \{1, ..., n\}$, vzájemně nezávislé, stejně rozdělené náhodné veličiny s neznámou distribuční funkcí F a tuto funkci chceme odhadnout. Pokud o rozdělení $X_1, X_2, ..., X_n$ nic nevíme, jako přirozená volba pro odhad se jeví tzv. *empirická distribuční funkce*

$$F_n(t) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}[X_i \leq t].$$

Tento odhad je nestranný, protože $\forall t \in \mathbb{R}$:

$$\mathbb{E}[F_n(t)] = \frac{1}{n}\sum_{i=1}^{n} E[\mathbb{1}\{X_i \leq t\}] = F(t)$$

z linearity střední hodnoty a faktu $E[\mathbb{1}\{X_i \leq t\}] = \mathsf{P}(X_i \leq t) = F(t)$. Dále máme $\forall t \in \mathbb{R} : E\,|\mathbb{1}\{[X_i \leq t]\}| < \infty$ a tedy ze silného zákona velkých čísel pro nezávislé, stejně rozdělené náhodné veličiny (SZVČ), viz Věta 4.8 v Dupač a Hušková (2005), dostaneme pro pevné $t \in \mathbb{R}$

$$F_n(t) \xrightarrow[n\to\infty]{s.j.} F(t)$$

a tedy je i konzistentní. Glivenkova-Cantelliho věta tyto výsledky dále posiluje a říká, že pokud se na F_n díváme jako na funkci proměnné t, tak F_n konverguje skoro jistě k F na $\mathbb{R}$ dokonce stejnoměrně, tedy $\sup_{t\in\mathbb{R}} |F_n(t) - F(t)| \xrightarrow[n\to\infty]{s.j.} 0$.

Poznámka. Často budeme pro funkci F a bod $t \in \mathbb{R}$ v práci používat značení

$$F(t-) = \lim_{x\to t-} F(x).$$

V důkazu Glivenkovy-Cantelliho věty budeme potřebovat následující lemma.

Lemma 1. *Nechť F je distribuční funkce a $\varepsilon > 0$. Pak existuje konečné dělení $-\infty = t_0 < t_1 < ... < t_m = \infty$ takové, že pro všechna $k \in \{1, ..., m\}$ platí*

$$F(t_k-) - F(t_{k-1}) \leq \varepsilon. \tag{1.1}$$

Důkaz. Máme $t_0 = -\infty$ a další body dělení konstruujeme následovně:

$$t_k = \sup_{x\in\mathbb{R}}\{x \mid F(x) - F(t_{k-1}) \leq \varepsilon\}, \qquad k = 1, 2, ...$$

Ukážeme, že platí (1.1). Protože $\varepsilon > 0$ a F je neklesající, je $t_k > t_{k-1}$ a z definice t_k tedy dostaneme, že $\forall \delta \in (0, t_k - t_{k-1})$:

$$F(t_k - \delta) - F(t_{k-1}) \leq \varepsilon$$

a tedy (1.1) platí. Zároveň ale platí, že

$$F(t_k) - F(t_{k-1}) \geq \varepsilon$$

a to ukážeme sporem. Předpokládejme, že $F(t_k) - F(t_{k-1}) < \varepsilon$. Pak ale ze spojitosti F zprava existuje $\delta > 0$ takové, že $F(t_k + \delta) - F(t_{k-1}) < \varepsilon$ a to je spor s definicí t_k. Naše dělení musí tedy být konečné, protože kdyby nebylo, tak pro $n \in \mathbb{N}, n > 1/\varepsilon$ bychom měli $F(t_n) > 1$ a F by nebyla distribuční funkce.

□

Poznámka. Všimněme si, že Lemma 1 nám dá v (1.1) i ostrou nerovnost. Pro $\varepsilon > 0$ ho stačí použít na $0 < \varepsilon' < \varepsilon$.

1.2 Důkaz Glivenkovy-Cantelliho věty

Věta 2 (Glivenkova-Cantelliho)**.** *Nechť pro $n \in \mathbb{N}$ jsou $X_1, X_2, ..., X_n$ vzájemně nezávislé, stejně rozdělené náhodné veličiny s distribuční funkcí F. Potom*

$$F_n(t) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}[X_i \leq t]$$

stejnoměrně konverguje k F na $\mathbb{R}$ skoro jistě, tedy

$$\sup_{t \in \mathbb{R}} |F_n(t) - F(t)| \xrightarrow[n \to \infty]{s.j.} 0.$$

Poznámka. Ve funkci $|F_n(t) - F(t)|$ závisí na ω pouze člen F_n. Pokud budeme dále pracovat s vybranými ω, píšeme $|F_n(t, \omega) - F(t)|$.

Ještě před samotným důkazem věty ukážeme, že její tvrzení dává smysl. Konvergence skoro jistě je totiž definovaná pouze pro měřitelné funkce. Pro libovolné $t \in \mathbb{R}$ je náhodná veličina $Z_n(\omega) = |F_n(t, \omega) - F(t)|$ měřitelná, protože $F_n(t, \omega)$ je součet měřitelných $\mathbb{1}[X_i(\omega) \leq t]$ a $F(t)$ je konstanta. Z teorie míry víme, že supremum přes spočetnou množinu měřitelných funkcí je měřitelné. V tvrzení věty ale máme supremum přes nespočetnou $\mathbb{R}$, které nutně měřitelné být nemusí. Toto vyjasní následující lemma.

Lemma 3. *Nechť F a F_n jsou jako výše a $\omega \in \Omega$ je libovolné. Potom platí*

$$\sup_{t \in \mathbb{R}} |F_n(t, \omega) - F(t)| = \sup_{t \in \mathbb{Q}} |F_n(t, \omega) - F(t)| .$$

Důkaz. Množina $\{|F_n(t, \omega) - F(t)| \mid t \in \mathbb{R}\}$ je zřejmě neprázdná a shora omezená a tedy supremum na levé straně rovnosti existuje. Označme

$$S = \sup_{t \in \mathbb{R}} |F_n(t, \omega) - F(t)| ,$$

pak z definice suprema máme

$$\forall t \in \mathbb{R} : |F_n(t, \omega) - F(t)| \leq S \tag{1.2}$$

a také

$$\forall s < S \; \exists t \in \mathbb{R} : |F_n(t,\omega) - F(t)| > s. \tag{1.3}$$

Ukážeme, že tyto podmínky platí i pro $t \in \mathbb{Q}$ a tím bude důkaz hotov. Podmínka (1.2) platí zřejmě pro každé $t \in \mathbb{Q}$ z toho, že $\mathbb{Q} \subset \mathbb{R}$. Nechť $s \in \mathbb{R}, s < S$, potom z (1.3) existuje $t \in \mathbb{R}$ takové, že $|F_n(t,\omega) - F(t)| > s$. Jenže ze spojitosti $|F_n(t,\omega) - F(t)|$ zprava a z toho, že $\mathbb{Q}$ je hustá v $\mathbb{R}$, existuje $q \in \mathbb{Q}, q > t$ takové, že i $|F_n(q,\omega) - F(q)| > s$.

□

Poznámka. Ukázali jsme, že funkce $Y_n(\omega) = \sup_{t\in\mathbb{R}} |F_n(t,\omega) - F(t)|$ je v každém ω rovna náhodné veličině $\sup_{t\in\mathbb{Q}} |F_n(t,\omega) - F(t)|$, tedy i Y_n je náhodná veličina.

Důkaz. [Věty 2] Nechť $\varepsilon > 0$. Z lemmatu 1 existuje konečné dělení $-\infty = t_0 < t_1 < ... < t_m = \infty$ takové, že $F(t_k-) - F(t_{k-1}) \leq \varepsilon$. Protože F a F_n jsou neklesající na $\mathbb{R}$, pro libovolné $k \in \{1, ..., m\}$ platí $\forall t \in [t_{k-1}, t_k)$:

$$F(t_{k-1}) \leq F(t) \leq F(t_k-),$$

$$F_n(t_{k-1}) \leq F_n(t) \leq F_n(t_k-),$$

z čehož plyne

$$F_n(t) - F(t) \leq F_n(t_k-) - F(t_{k-1}).$$

Dále můžeme psát

$$F_n(t) - F(t) \leq F_n(t_k-) - F(t_k-) + F(t_k-) - F(t_{k-1}).$$

Teď využijeme toho, že

$$F(t_k-) - F(t_{k-1}) \leq \varepsilon$$

a dostaneme

$$F_n(t) - F(t) \leq F_n(t_k-) - F(t_k-) + \varepsilon.$$

Konečně z vlastností suprema (neporuší nám nerovnost) a funkce max můžeme nerovnost rozšířit na celé $\mathbb{R}$

$$\sup_{t\in\mathbb{R}}(F_n(t) - F(t)) \leq \max_{k\in\{1,...,m\}}(F_n(t_k-) - F(t_k-)) + \varepsilon. \tag{1.4}$$

Analogicky z nerovnosti

$$F_n(t) - F(t) \geq F_n(t_{k-1}) - F(t_k-)$$

odvodíme, že

$$\inf_{t\in\mathbb{R}}(F_n(t) - F(t)) \geq \min_{k\in\{1,...,m\}}(F_n(t_{k-1}) - F(t_{k-1})) - \varepsilon. \tag{1.5}$$

Protože náhodné veličiny $X_1, ..., X_n$ jsou vzájemně nezávislé a stejně rozdělené, jsou i $\mathbb{1}[X_1 \leq t], ..., \mathbb{1}[X_n \leq t]$, $t \in \mathbb{R}$ vzájemně nezávislé a stejně rozdělené. Označme $Y_i = \mathbb{1}[X_i \leq t]$, potom $\overline{Y_n} = F_n(t)$ a také $\mathbb{E}|Y_1| = \mathbb{E}\mathbb{1}[X_1 \leq t] = P(X_1 \leq t) = F(t)$ a tedy ze SZVČ dostáváme, že $F_n(t) \xrightarrow[n\to\infty]{s.j.} F(t)$ a podobně také $F_n(t-) \xrightarrow[n\to\infty]{s.j.} F(t-)$. Tedy pro libovolné $k \in \{1, ..., m\}$ máme z definice konvergence skoro jistě, že na domovském pravděpodobnostním prostoru $(\Omega, \mathcal{A}, \mathsf{P})$ náhodných veličin $X_1, ..., X_n$, je

$$\mathsf{P}\left(\{\omega \in \Omega : \lim_{n\to\infty} |F_n(t_k,\omega) - F(t_k)| = 0\}\right) = 1$$

a tedy zřejmě i

$$\mathsf{P}\left(\{\omega \in \Omega : \lim_{n\to\infty} \max_{k\in\{1,\dots,m\}} |F_n(t_k-,\omega) - F(t_k-)| = 0\}\right) = 1,$$

$$\mathsf{P}\left(\{\omega \in \Omega : \lim_{n\to\infty} \left|\min_{k\in\{1,\dots,m\}} (F_n(t_{k-1},\omega) - F(t_{k-1}))\right| = 0\}\right) = 1.$$

Jinými slovy existuje množina $A_1 \subset \Omega, \mathsf{P}(A_1) = 0$ taková, že pro $\omega \in \Omega \setminus A_1$ platí

$$\max_{k\in\{1,\dots,m\}} |F_n(t_k-,\omega) - F(t_k-)| \xrightarrow[n\to\infty]{} 0$$

a množina $A_2 \subset \Omega, \mathsf{P}(A_2) = 0$ taková, že pro $\omega \in \Omega \setminus A_2$ platí

$$\left|\min_{k\in\{1,\dots,m\}} (F_n(t_{k-1},\omega) - F(t_{k-1}))\right| \xrightarrow[n\to\infty]{} 0.$$

Z (1.4) a (1.5) a faktu

$$\sup_{t\in\mathbb{R}} |F_n(t) - F(t)| = \max\left\{\sup_{t\in\mathbb{R}}(F_n(t) - F(t)), -\inf_{t\in\mathbb{R}}(F_n(t) - F(t))\right\}$$

dostaneme pro $\omega \in \Omega \setminus (A_1 \cup A_2)$

$$\limsup_{n\to\infty} \sup_{t\in\mathbb{R}} |F_n(t,\omega) - F(t)| \le \varepsilon \tag{1.6}$$

a protože $\mathsf{P}(A_1 \cup A_2) \le \mathsf{P}(A_1) + \mathsf{P}(A_2) = 0$, platí nerovnost (1.6) skoro jistě. Zatím jsme dokázali, že

$$\forall \varepsilon > 0 : \mathsf{P}\left(\limsup_{n\to\infty} \sup_{t\in\mathbb{R}} |F_n(t) - F(t)| \le \varepsilon\right) = 1. \tag{1.7}$$

Dále postupujeme podobně jako v první půlce důkazu věty 4.4 v knize Dupač a Hušková (2005). Označme $Y_n(\omega) = \sup_{t\in\mathbb{R}} |F_n(t,\omega) - F(t)|$, a také

$$C = \{\omega \in \Omega : \limsup_{n\to\infty} Y_n(\omega) = 0\}.$$

K dokončení důkazu stačí ukázat, že $\mathsf{P}(C^{\mathsf{c}}) = 0$. Z nezápornosti Y_n máme

$$C^{\mathsf{c}} = \{\omega \in \Omega : \limsup_{n\to\infty} Y_n(\omega) > 0\} = \bigcup_{k=1}^{\infty} \{\omega \in \Omega : \limsup_{n\to\infty} Y_n(\omega) > \tfrac{1}{k}\}$$

a tedy

$$\mathsf{P}(C^{\mathsf{c}}) \le \sum_{k=1}^{\infty} \mathsf{P}\left(\limsup_{n\to\infty} Y_n > \tfrac{1}{k}\right).$$

Jenže z (1.7) plyne, že každý člen sumy na pravé straně předchozí nerovnosti je nulový a odtud $\mathsf{P}(C^{\mathsf{c}}) = 0$, což ukončuje důkaz

□

1.3 Stejnoměrná konvergence výběrových kvantilů

Jedním důležitým důsledkem Věty 2 je stejnoměrná konvergence výběrových kvantilů na vhodné podmnožině intervalu $(0,1)$. Nechť X je náhodná veličina, která má distribuční funkci F. Pro $\alpha \in (0,1)$ definujeme $\alpha-$kvantil rozdělení X jako

$$u_X(\alpha) = \inf\{x \in \mathbb{R} \mid F(x) \geq \alpha\}.$$

Obdobně pro nahodný výběr $X_1, ..., X_n$ definujeme výběrový $\alpha-$kvantil jako

$$u_n(\alpha) = \inf\{x \in \mathbb{R} \mid F_n(x) \geq \alpha\}.$$

Poznámka. Pro $u_n(\alpha)$ platí

$$u_n(\alpha) = X_{(k_\alpha)},$$

kde $X_{(k_\alpha)}$ je $k_\alpha-$tá pořádková statistika a

$$k_\alpha = \begin{cases} n\alpha, & (n\alpha) \in \mathbb{N}, \\ \lfloor n\alpha \rfloor + 1, & (n\alpha) \notin \mathbb{N}. \end{cases} \tag{1.8}$$

Teď uvedeme větu, která zaručuje bodovou konvergenci $u_n(\alpha)$ k $u_X(\alpha)$ skoro jistě.

Věta 4. *Nechť $\alpha \in (0,1)$ a $u_X(\alpha)$ je jediné číslo x splňující*

$$F(x-) \leq \alpha \leq F(x). \tag{1.9}$$

Potom platí

$$u_n(\alpha) \xrightarrow[n\to\infty]{s.j.} u_X(\alpha)$$

Důkaz. Viz str. 75 v knize Serfling (2002).

□

Poznámka. Předpoklad (1.9) lze ekvivalentně nahradit tím, že nesmí nastat zároveň $F(u_X(\alpha)) = \alpha$ a F je konstantní na pravém okolí $u_X(\alpha)$. My ukážeme, že $u_X(\alpha)$ není jednoznačné řešení (1.9) právě tehdy, když obě podmínky zmíněné výše platí. Na jednu stranu, pokud $F(u_X(\alpha)) = \alpha$ a F je konstantní na pravém okolí $u_X(\alpha)$, pak zřejmě $u_X(\alpha)$ splňuje (1.9) a navíc existuje $\delta > 0$ takové, že $F(u_X(\alpha) + \delta) = \alpha$, tedy i bod $u_X(\alpha) + \delta$ splňuje (1.9). Na druhou stranu, nechť k $u_X(\alpha)$ existuje $x_1 > u_X(\alpha)$ takové, že

$$F(x_1-) \leq \alpha \leq F(x_1)$$

Kdyby F nebyla konstantní na pravém okolí $u_X(\alpha)$, potom by na něm byla ostře rostoucí, tedy by platilo

$$\alpha \leq F(u_X(\alpha)) < F(x_1-) \leq \alpha,$$

což je spor. Kdyby $F(u_X(\alpha)) \neq \alpha$, potom by muselo být $F(u_X(\alpha)) > \alpha$ a odtud

$$\alpha < F(u_X(\alpha)) \leq F(x_1-) \leq \alpha,$$

což je opět spor.

Bodovou konvergenci z Věty 4 ještě zesílíme v následující větě, která mluví o stejnoměrné silné konzistenci $u_n(\alpha)$. Budeme ale muset o něco také zesílit předpoklady a stejnoměrnou konvergenci dokážeme na $I \subseteq [0,1]$, kde je F stejnoměrně rostoucí.

Věta 5. *Nechť $X_1, ..., X_n$ je náhodný výběr z rozdělení s distribuční funkcí F a $I \subseteq [0,1]$ je taková množina, že $u_X(\alpha), \alpha \in I \cap (0,1)$ je jediné řešení* (1.9) *a pro každé $\delta > 0$ platí*

$$\inf_{\alpha \in I}\{F(u_X(\alpha)) - F(u_X(\alpha) - \delta)\} > 0. \tag{1.10}$$

Potom

$$\sup_{\alpha \in I} |u_n(\alpha) - u_X(\alpha)| \xrightarrow[n\to\infty]{s.j.} 0.$$

Důkaz. Potřebujeme ukázat, že pro každé $\varepsilon > 0$ je

$$\mathsf{P}\left(\limsup_{n\to\infty} \sup_{\alpha \in I} |u_n(\alpha) - u_X(\alpha)| > \varepsilon\right) = 0 \tag{1.11}$$

a potom stačí použít stejný postup jako v závěru důkazu Věty 2. Nechť $\varepsilon > 0$. Nejprve ukážeme, že

$$\mathsf{P}\left(\limsup_{n\to\infty} \sup_{\alpha \in I} (u_n(\alpha) - u_X(\alpha)) > \varepsilon\right) = 0.$$

Je totiž

$$\begin{aligned}\mathsf{P}\left(\sup_{\alpha \in I}(u_n(\alpha) - u_X(\alpha)) > \varepsilon\right) &= \mathsf{P}\left(\sup_{\alpha \in I}(X_{(k_\alpha)} - u_X(\alpha)) > \varepsilon\right) \\ &= \mathsf{P}\left(\sup_{\alpha \in I}(\textstyle\sum_{i=1}^n \mathbb{1}\{X_i \le u_X(\alpha) + \varepsilon\} - k_\alpha) < 0\right) \\ &= \mathsf{P}\left(\sup_{\alpha \in I}(F_n(u_X(\alpha) + \varepsilon) - \tfrac{k_\alpha}{n}) < 0\right) \end{aligned} \tag{1.12}$$

Z Věty 2 víme, že $\sup_{t \in \mathbb{R}} |F_n(t) - F(t)| \xrightarrow[n\to\infty]{s.j.} 0$ a tedy pro každé $\varepsilon > 0$ je i $\sup_{\alpha \in I}(F_n(u_X(\alpha)+\varepsilon) - F(u_X(\alpha)+\varepsilon)) \xrightarrow[n\to\infty]{s.j.} 0$. Navíc pro libovolné $\alpha \in I$ a $\varepsilon > 0$ je

$$F(u_X(\alpha) + \varepsilon) - \tfrac{k_\alpha}{n} \xrightarrow[n\to\infty]{} F(u_X(\alpha) + \varepsilon) - \alpha \ge F(u_X(\alpha) + \varepsilon) - F(u_X(\alpha)) > 0$$

z předpokladu (1.9). Tedy

$$\limsup_{n\to\infty} \sup_{\alpha \in I} (F(u_X(\alpha) + \varepsilon) - \tfrac{k_\alpha}{n}) \ge 0.$$

A protože můžeme psát

$$\sup_{\alpha \in I}(F_n(u_X(\alpha)+\varepsilon) - \tfrac{k_\alpha}{n}) = \sup_{\alpha \in I}(F_n(u_X(\alpha)+\varepsilon) - F(u_X(\alpha)+\varepsilon) + F(u_X(\alpha)+\varepsilon) - \tfrac{k_\alpha}{n}),$$

dostaneme, že $\limsup_{n\to\infty} \sup_{\alpha\in I}(F_n(u_X(\alpha)+\varepsilon) - \frac{k_\alpha}{n}) \geq 0$ skoro jistě a tedy z (1.12) dostaneme

$$\mathsf{P}\left(\limsup_{n\to\infty} \sup_{\alpha\in I}(u_n(\alpha) - u_X(\alpha)) > \varepsilon\right) = 0. \tag{1.13}$$

Teď ukážeme, že

$$\mathsf{P}\left(\limsup_{n\to\infty} \sup_{\alpha\in I}(u_n(\alpha) - u_X(\alpha)) < -\varepsilon\right) = 0.$$

Máme totiž

$$\begin{aligned}
\mathsf{P}\left(\sup_{\alpha\in I}(u_n(\alpha) - u_X(\alpha)) < -\varepsilon\right) &= \mathsf{P}\left(\sup_{\alpha\in I}(X_{(k_\alpha)} - u_X(\alpha)) < -\varepsilon\right)\\
&= \mathsf{P}\left(\sup_{\alpha\in I}(\textstyle\sum_{i=1}^n \mathbb{1}\{X_i < u_X(\alpha)-\varepsilon\} - k_\alpha) \geq 0\right)\\
&\leq \mathsf{P}\left(\sup_{\alpha\in I}(F_n(u_X(\alpha)-\varepsilon) - \frac{k_\alpha}{n}) \geq 0\right),
\end{aligned}$$

kde poslední nerovnost plyne z toho, že

$$\frac{1}{n}\sum_{i=1}^n \mathbb{1}\{X_i < u_X(\alpha)-\varepsilon\} \leq F_n(u_X(\alpha)-\varepsilon).$$

Z Věty 2 opět dostaneme, že $\sup_{\alpha\in I}(F_n(u_X(\alpha)-\varepsilon) - F(u_X(\alpha)-\varepsilon)) \xrightarrow[n\to\infty]{s.j.} 0$. Dále z (1.8) máme

$$\sup_{\alpha\in I}\left|\frac{k_\alpha}{n} - \alpha\right| \leq \sup_{\alpha\in I}\left|\frac{n\alpha+1}{n} - \alpha\right| = \frac{1}{n} \xrightarrow[n\to\infty]{} 0,$$

odtud konvergence $\frac{k_\alpha}{n}$ k α je stejnoměrná na I a tedy i $F(u_X(\alpha)-\varepsilon) - \frac{k_\alpha}{n}$ konverguje stejnoměrně k $F(u_X(\alpha)-\varepsilon) - \alpha$ na I. Odtud a z předpokladu (1.10) dostaneme

$$\begin{aligned}
\limsup_{n\to\infty} \sup_{\alpha\in I}(F(u_X(\alpha)-\varepsilon) - \frac{k_\alpha}{n}) &= \sup_{\alpha\in I}(F(u_X(\alpha)-\varepsilon) - \alpha)\\
&\leq \sup_{\alpha\in I}(F(u_X(\alpha)-\varepsilon) - F(u_X(\alpha)))\\
&= -\inf_{\alpha\in I}(F(u_X(\alpha)) - F(u_X(\alpha)-\varepsilon))\\
&< 0.
\end{aligned}$$

Odtud plyne, že $\limsup_{n\to\infty} \sup_{\alpha\in I}(F_n(u_X(\alpha)-\varepsilon) - \frac{k_\alpha}{n}) < 0$ skoro jistě a tedy

$$\mathsf{P}\left(\limsup_{n\to\infty} \sup_{\alpha\in I}(u_n(\alpha) - u_X(\alpha)) < -\varepsilon\right) = 0. \tag{1.14}$$

Kombinací (1.13) a (1.14) dostáváme (1.11) a dále stačí postupovat analogicky se situací v (1.7) ve Větě 2.

□

Na konec kapitoly uvedeme několik příkladů, jak může vypadat množina I z Věty 5 pro vybraná rozdělení. Abychom dostali konkrétní tvar I, uvažujeme rozdělení se známými parametry, analogicky lze ale I získat i pro rozdělení s parametry neznámými.

Příklad. Nechť $X_1, ..., X_n$ je náhodný výběr z rozdělení $X \sim N(0,1)$. X má tedy distribuční funkci

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-t^2}{2}} dt,$$

která je rostoucí a spojitá na $\mathbb{R}$. Pro každé $\alpha \in (0,1)$ je tedy $u_X(\alpha)$ jediným řešením (1.9). I ale nemůže být interval $(0,1)$, protože pro $\delta > 0$ je

$$\inf_{\alpha \in (0,1)} (F(u_X(\alpha)) - F(u_X(\alpha) - \delta)) = \lim_{x \to \pm\infty} (F(x) - F(x - \delta)) = 0.$$

Na jakémkoliv uzavřeném podintervalu $(0,1)$ už ale (1.10) platí, I tedy může být libovolně velký uzavřený podinterval $(0,1)$.

Příklad. Náhodná veličina s rovnoměrným rozdělením na $(0,1)$ má distribuční funkci

$$F(x) = \begin{cases} 0, & x < 0, \\ x, & x \in [0,1], \\ 1, & x > 1, \end{cases}$$

která je rostoucí na $[0,1]$ a $u_X(\alpha)$ je jediným řešením (1.9) pro všechna $\alpha \in (0,1)$. Pro $\delta > 0$ máme

$$\inf_{\alpha \in [0,1]} \{F(u_X(\alpha)) - F(u_X(\alpha) - \delta)\} = \lim_{x \to -\infty} (F(x) - F(x - \delta)) = 0.$$

Pokud ostře vyřadíme nějaké pravé okolí nuly, podmínka (1.10) už platí a I tedy může být uzavřený interval tvaru $[\eta, 1]$, kde $\eta \in (0,1)$. I můžeme ještě zvětšit na interval $(0,1]$. Pro $\alpha \in (0,1]$ je $u_X(\alpha) = \alpha$, tedy

$$\sup_{\alpha \in (0,1]} |u_n(\alpha) - \alpha| = \max \left\{ \sup_{\alpha \in (0,\eta)} |u_n(\alpha) - \alpha|, \sup_{\alpha \in [\eta,1]} |u_n(\alpha) - \alpha| \right\}.$$

Druhý člen z maxima výše jde skoro jistě k nule z Věty 5. Teď stačí ukázat, že

$$\mathsf{P}\left(\limsup_{n \to \infty} \sup_{\alpha \in (0,\eta)} |u_n(\alpha) - \alpha| > 3\eta \right) = 0. \tag{1.15}$$

Z monotonie $u_n(\alpha)$ potom

$$\sup_{\alpha \in (0,\eta)} |u_n(\alpha) - \alpha| \leq \sup_{\alpha \in (0,\eta)} |u_n(\alpha)| + \sup_{\alpha \in (0,\eta)} |\alpha| = u_n(\eta) + \eta.$$

Věta 4 nám dá, že

$$u_n(\eta) \xrightarrow[n \to \infty]{s.j.} u_X(\eta) = \eta,$$

tedy

$$\limsup_{n \to \infty} \sup_{\alpha \in (0,\eta)} |u_n(\alpha) - \alpha| \leq 2\eta < 3\eta$$

skoro jistě a 1.15 platí. Celkem

$$\sup_{\alpha \in (0,1]} |u_n(\alpha) - \alpha| \xrightarrow[n \to \infty]{s.j.} 0.$$

Příklad. Pro náhodný výběr z Poissonova rozdělení s parametrem 1 máme distribuční funkci

$$F(x) = e^{-1} \sum_{k=0}^{\lfloor x \rfloor} \frac{1}{k!}.$$

Nechť $\alpha \in [0,1]$. Snadno vidíme, že $\forall \delta > 0 : F(u_X(\alpha)) - F(u_X(\alpha) - \delta) > 0$ nastává právě tehdy, když $\forall \delta > 0 : \lfloor u_X(\alpha) \rfloor > \lfloor u_X(\alpha) - \delta \rfloor$ a $\alpha > 0$, což je ekvivalentní tomu, že $u_X(\alpha) \in \mathbb{N}_0$ a to nastává pro každé $\alpha \in (0,1)$. Jenže pro každé $\delta > 0$ je

$$\inf_{\alpha \in (0,1)} (F(u_X(\alpha)) - F(u_X(\alpha) - \delta)) = \lim_{n \to \infty} (F(n) - F(n - \delta)) = 0,$$

musíme se tedy omezit na intervaly $(0, 1-\eta]$, kde $\eta \in (0,1)$. Navíc aby $u_X(\alpha)$ bylo jediné řešení (1.9) z poznámky za Větou 4, nesmí nastat $F(u_X(\alpha)) = \alpha$ současně s tím, že F je konstantní na pravém okolí $u_X(\alpha)$. Toto nastává právě pro taková $\alpha \in (0,1)$, že $\forall K \in \mathbb{N} : \alpha \neq e^{-1} \sum_{k=1}^{K} \frac{1}{k!}$. Obě podmínky tedy platí zaroveň na množině $I = (0, 1-\eta] \setminus \{e^{-1} \sum_{k=1}^{K} \frac{1}{k!} \mid K \in \mathbb{N}\}$, kde $\eta \in (0,1)$ je libovolné.

2. Zobecnění pomocí bracketing numbers

V následující kapitole ještě zobecníme výsledky z kapitoly předešlé. Teď budeme uvažovat náhodný výběr $X_1,...,X_n$ na obecném měřitelném prostoru $(\mathcal{X},\mathcal{C})$ s rozdělením P, tedy mírou definovanou jako $P(B) = \mathsf{P}(X_1 \in B)$, kde $B \in \mathcal{C}$. Pro $x \in \mathcal{X}$ a $A \in \mathcal{C}$ označme tzv. Diracovu míru jako

$$\delta_x(A) = \mathbb{1}[x \in A].$$

Místo empirické distribuční funkce budeme v této kapitole uvažovat tzv. empirické rozdělení

$$\mathbb{P}_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i}.$$

Dále pro měřitelnou funkci $f : \mathcal{X} \to \mathbb{R}$ budeme střední hodnotu f podle $\mathbb{P}_n$ a P značit jako

$$\mathbb{P}_n f = \int_{\mathcal{X}} f(x) d\mathbb{P}_n(x) = \frac{1}{n}\sum_{i=1}^{n} f(X_i)$$

a

$$Pf = \int_{\mathcal{X}} f(x) dP(x) = \mathbb{E} f(X_1).$$

Pro f měřitelnou jsou $f(X_1),..., f(X_n)$ vzájemně nazávislé a stejně rozdělené, $\mathbb{E}[f(X_1)] = Pf$ a tedy pokud $P|f| < \infty$, ze SZVČ máme

$$\mathbb{P}_n f \xrightarrow[n\to\infty]{s.j.} Pf.$$

Teď si zadefinujeme speciální případ zobecněné konvergence skoro jistě, kterou budeme potřebovat pro vyslovení zobecněné Glivenkovy-Cantelliho věty. Tato definice nebude vyžadovat měřitelnost jako ta klasická. Dále totiž budeme pracovat se $\sup_{f\in\mathcal{F}} |\mathbb{P}_n f - Pf|$, kde $\mathcal{F}$ je nějaká třída funkcí, přičemž o této třídě obecně nevíme nic a nemůžeme tedy s měřitelností počítat.

Definice 1. *Nechť $X_n : \Omega \to \mathbb{R}, n \in \mathbb{N}$ je posloupnost reálných funkcí. Potom řekneme, že tato posloupnost konverguje k 0 skoro jistě z vnějšku (outer almost surely), píšeme*

$$X_n \xrightarrow[n\to\infty]{s.j.*} 0,$$

pokud pro každé $n \in \mathbb{N}$ existuje náhodná veličina $\Delta_n : (\Omega,\mathcal{A}, \mathsf{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ (tedy Δ_n je borelovsky meřitelná) taková, že $\forall\omega \in \Omega : |X_n(\omega)| \le \Delta_n(\omega)$ a $\Delta_n \xrightarrow[n\to\infty]{s.j.} 0$.

Všimněme si, že pokud jsou X_n měřitelné, dostaneme klasickou $\xrightarrow[n\to\infty]{s.j.}$. Definice se dá zobecnit na X_n s hodnotami v obecném metrickém prostoru s limitou jinou než 0 a je k nalezení například na str. 258 ve van der Vaart (2000).

Definice 2. *Třída měřitelných funkcí $\mathcal{F}$ se nazývá Glivenkova-Cantelliho (G-C), pokud*

$$\sup_{f\in\mathcal{F}} |\mathbb{P}_n f - Pf| \xrightarrow[n\to\infty]{s.j.*} 0.$$

Zde vidíme, proč byla definice 1 nutná. Třída $\mathcal{F}$ může být velmi obecná a proto vůbec není jasné, jak je to s měřitelností suprema z definice 2. Naše zobecněná konvergence měřitelnost nevyžaduje. Pokud je $\mathcal{F}$ konečná, je zřejmě G-C. Zobecněná verze věty 2 říká, že postačující podmínka pro to, aby $\mathcal{F}$ byla G-C, je, že bude v jistém smyslu malá. Velikost $\mathcal{F}$ se dá změřit pomocí tzv. bracketing numbers, které si teď definujeme.

Definice 3. *Pro f měřitelnou a $r > 0$ definujeme $L_r(P)$-normu jako*

$$\|f\|_{r,P} = (P\,|f|^r)^{\frac{1}{r}}.$$

Definice 4. *Nechť u, l jsou měřitelné funkce s konečnou $L_r(P)$-normou, pak bracket $[l, u]$ definujeme jako*

$$[l, u] = \{f \in \mathcal{F} \mid \forall x \in \mathcal{X} : l(x) \leq f(x) \leq u(x)\}$$

Pro $\varepsilon > 0$ definujeme ε-bracket vzhledem k $L_r(P)$-normě jako

$$[l, u]_\varepsilon = \{f \in \mathcal{F} \mid \forall x \in \mathcal{X} : l(x) \leq f(x) \leq u(x), P(u - l)^r < \varepsilon^r\}$$

a konečně bracketing number $N_{[\,]}(\varepsilon, \mathcal{F}, L_r(P))$ je minimální počet ε-bracketů, které pokryjí celou $\mathcal{F}$.

Poznámka. Funkce u, l z předchozí definice samy nemusí být prvky $\mathcal{F}$.

Věta 6 (zobecněná Glivenkova-Cantelliho)**.** *Nechť $\mathcal{F}$ je třída měřitelných funkcí $f : (\mathcal{X}, \mathcal{A}) \to \mathbb{R}$ takových, že $\forall \varepsilon > 0 : N_{[\,]}(\varepsilon, \mathcal{F}, L_1(P)) < \infty$, pak je tato třída Glivenkova-Cantelliho.*

Důkaz. Nechť $\varepsilon > 0$ a označme $N_\varepsilon = N_{[\,]}(\varepsilon, \mathcal{F}, L_1(P))$. Z předpokladu věty je $\forall f \in \mathcal{F}\ \exists i \in \{1,..., N_\varepsilon\} : f \in [l_i, u_i]_\varepsilon$. Z faktu $f \leq u_i$ a monotonie $\mathbb{P}_n$ máme $\mathbb{P}_n f \leq \mathbb{P}_n u_i$ a z definice $\varepsilon-$bracketu $Pu_i - Pf \leq Pu_i - Pl_i < \varepsilon$ pak pro libovolnou $f \in [l_i, u_i]_\varepsilon$ máme

$$\mathbb{P}_n f - Pf = \mathbb{P}_n f - Pu_i + Pu_i - Pf < \mathbb{P}_n u_i - Pu_i + \varepsilon$$

a odtud z vlastností suprema také

$$\sup_{f \in \mathcal{F}} (\mathbb{P}_n f - Pf) \leq \max_{i \in \{1,...,N_\varepsilon\}} (\mathbb{P}_n u_i - Pu_i) + \varepsilon.$$

Odhad zdola dostaneme podobně jako ten shora

$$\mathbb{P}_n f - Pf = \mathbb{P}_n f - Pl_i + Pl_i - Pf > \mathbb{P}_n l_i - Pl_i - \varepsilon$$

a odtud

$$\inf_{f \in \mathcal{F}} (\mathbb{P}_n f - Pf) \geq \min_{i \in \{1,...,N_\varepsilon\}} (\mathbb{P}_n l_i - Pl_i) - \varepsilon.$$

Dále protože $X_1, X_2, ...$ jsou nezávislé a stejně rozdělené, jsou i $\forall i \in \{1,..., N_\varepsilon\}$ $l_i(X_1), l_i(X_2), ...$ nezávislé a stejně rozdělené a stejně tak pro u_i. Taky $\mathbb{E}\,|l_i(X_1)| \leq$

$E\,|u_i(X_1)| < \infty$ z definice bracketu a SZVČ nám dá pro každé i: $\mathbb{P}_n l_i - P l_i \xrightarrow[n\to\infty]{s.j.} 0$ a $\mathbb{P}_n u_i - P u_i \xrightarrow[n\to\infty]{s.j.} 0$. Celkem tedy

$$\begin{aligned}\min_{i\in\{1,\dots,N_\varepsilon\}}(\mathbb{P}_n l_i - P l_i) - \varepsilon &\le \inf_{f\in\mathcal{F}}(\mathbb{P}_n f - Pf)\\ &\le \sup_{f\in\mathcal{F}}(\mathbb{P}_n f - Pf)\\ &\le \max_{i\in\{1,\dots,N_\varepsilon\}}(\mathbb{P}_n u_i - P u_i) + \varepsilon.\end{aligned}$$

Odtud můžeme psát

$$\sup_{f\in\mathcal{F}}|\mathbb{P}_n f - Pf| \le \max_{i\in\{1,\dots,N_\varepsilon\}}\{|\mathbb{P}_n l_i - P l_i|\,, |\mathbb{P}_n u_i - P u_i|\} + \varepsilon.$$

Označme pravou stranu předchozí nerovnosti $\Delta_n(\varepsilon)$ a všimněme si, že je měřitelná. Dále položme

$$\Delta_n = \inf_{k\in\mathbb{N}} \Delta_n(\tfrac{1}{k}). \tag{2.1}$$

Potom Δ_n je také meřitelná, protože se jedná o infimum přes spočetně mnoho měřitelných funkcí, a dále platí

$$\sup_{f\in\mathcal{F}}|\mathbb{P}_n f - Pf| \le \Delta_n.$$

Pro zobecněnou konvergenci skoro jistě už stačí jen ukázat, že $\Delta_n \xrightarrow{s.j.} 0$. Položme $C = \{\limsup \Delta_n = 0\}$ a ukážeme, že $\mathsf{P}(C^{\mathsf{c}}) = 0$. Máme

$$C^{\mathsf{c}} = \{\limsup_{n\to\infty} \Delta_n > 0\} = \bigcup_{j=1}^{\infty}\{\limsup_{n\to\infty}\Delta_n > \tfrac{1}{j}\}$$

a ze $\sigma-$aditivity P je

$$\mathsf{P}(C^{\mathsf{c}}) \le \sum_{j=1}^{\infty}\mathsf{P}\left(\limsup_{n\to\infty}\Delta_n > \tfrac{1}{j}\right)$$

a ukážeme, že všechny členy sumy na pravé sraně jsou nulové. Z definice (2.1) je totiž $\forall j \in \mathbb{N} : \Delta_n \le \Delta_n(\frac{1}{2j})$ a tedy

$$\mathsf{P}\left(\limsup_{n\to\infty}\Delta_n > \tfrac{1}{j}\right) \le \mathsf{P}\left(\limsup_{n\to\infty}\Delta_n(\tfrac{1}{2j}) > \tfrac{1}{j}\right) = 0,$$

protože $\Delta_n(\frac{1}{2j}) \xrightarrow[n\to\infty]{s.j.} \frac{1}{2j}$ ze silného zákona velkých čísel pro nezávislé, stejně rozdělené náhodné veličiny.

$\square$

Poznámka. Tvrzení Věty 6 nelze obrátit. Jednoduchým protipříkladem je třída

$$\mathcal{F} = \{x \mapsto x + a \mid a \in \mathbb{R}\}.$$

Pro každé $\varepsilon > 0$ je zřejmě $N_{[\,]}(\varepsilon, \mathcal{F}, L_1(P)) = \infty$, ale pro $f_a \in \mathcal{F}$ máme

$$\mathbb{P}_n f_a = \frac{1}{n}\sum_{i=1}^{n} X_i + a$$

a také
$$Pf_a = \mathbb{E}X_1 + a.$$
Ze SZVČ dostáváme
$$\sup_{a\in\mathbb{R}} |\mathbb{P}_n f_a - Pf_a| = \left|\tfrac{1}{n}\textstyle\sum_{i=1}^n X_i - \mathbb{E}X_1\right| \xrightarrow[n\to\infty]{s.j.} 0.$$

Teď ukážeme, jak z Věty 6 plyne Věta 2. Pokud položíme
$$\mathcal{F} = \{x \mapsto \mathbb{1}_{(-\infty,t]}(x) \mid t \in \mathbb{R}\},$$
potom
$$\mathbb{P}_n \mathbb{1}_{(-\infty,t]} = \frac{1}{n}\sum_{i=1}^n \mathbb{1}[X_i \le t] = F_n(t)$$
a také
$$P\mathbb{1}_{(-\infty,t]} = \int_{\mathbb{R}} \mathbb{1}_{(-\infty,t]}(x)\, dP(x) = \mathsf{P}(X_i \le t) = F(t).$$
Tedy v tomto případe můžeme psát
$$\sup_{f\in\mathcal{F}} |\mathbb{P}_n f - Pf| = \sup_{t\in\mathbb{R}} |F_n(t) - F(t)|\,.$$

Teď už jen stačí ukázat, že $N_{[\,]}(\varepsilon, \mathcal{F}, L_1(P))$ je konečné pro každé $\varepsilon > 0$ a zobecněná Glivenkova-Cantelliho věta nám dá klasickou.

Nechť tedy $\varepsilon > 0$ je libovolné. Z lemmatu 1 najdeme konečné dělení splňující $F(t_i-) - F(t_{i-1}) < \varepsilon$ a $\varepsilon-$brackety konstruujeme jako
$$[l_i, u_i] = [\mathbb{1}_{(-\infty,t_{i-1}]}, \mathbb{1}_{(-\infty,t_i)}].$$
Potom pro každé $t \in [t_{i-1}, t_i)$ je
$$l_i(x) \le \mathbb{1}_{(-\infty,t]}(x) \le u_i(x) \quad \forall x \in \mathbb{R}$$
a navíc
$$P(u_i - l_i) = F(t_i-) - F(t_{i-1}) < \varepsilon.$$
Celkem dostaneme $\sup_{t\in\mathbb{R}} |F_n(t) - F(t)| \xrightarrow[n\to\infty]{s.j.*} 0$ a protože pro každé $n \in \mathbb{N}$ jsou náhodné veličiny $\sup_{t\in\mathbb{R}} |F_n(t) - F(t)|$ měřitelné z Lemmatu 3, je i
$$\sup_{t\in\mathbb{R}} |F_n(t) - F(t)| \xrightarrow[n\to\infty]{s.j.} 0.$$

2.1 Zobecnění klasické G-C věty do $\mathbb{R}^d$

Pomocí Věty 6 můžeme tvrzení klasické Glivenkovy-Cantelliho věty pro náhodné veličiny zobecnit i na náhodné vektory. Nechť $\mathbf{X_1}, ..., \mathbf{X_n}$ jsou náhodné vektory s hodnotami v $\mathbb{R}^d, d \in \mathbb{N}$, rozdělením P a distribuční funkcí F. Pro $\mathbf{t} \in \mathbb{R}^d$ definujme empirickou distribuční funkci jako
$$F_n(\mathbf{t}) = \frac{1}{n}\sum_{i=1}^n \mathbb{1}[\mathbf{X_i} \le \mathbf{t}],$$

kde pro $d > 1$ výraz $\mathbf{X_i} \leq \mathbf{t}$ značí, že j-tá složka vektoru $\mathbf{X_i}$ není větší, než j-tá složka vektoru $\mathbf{t}$ pro všechna $j \in \{1,...,d\}$. Potom platí

$$\sup_{\mathbf{t}\in\mathbb{R}^d} |F_n(\mathbf{t}) - F(\mathbf{t})| \xrightarrow[n\to\infty]{s.j.} \mathbf{0}.$$

Důkaz. Položme

$$\mathcal{F}_d = \{\mathbf{x} \mapsto \mathbb{1}_{(-\infty,\mathbf{t}]}(\mathbf{x}) \mid \mathbf{t} \in \mathbb{R}^d\}.$$

Potom dostaneme

$$\mathbb{P}_n \mathbb{1}_{(-\infty,\mathbf{t}]} = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}[\mathbf{X_i} \leq \mathbf{t}] = F_n(\mathbf{t})$$

a také

$$P\mathbb{1}_{(-\infty,\mathbf{t}]} = \int_{\mathbb{R}^d} \mathbb{1}_{(-\infty,\mathbf{t}]}(\mathbf{x})\, dP(\mathbf{x}) = \mathsf{P}(\mathbf{X_i} \leq \mathbf{t}) = F(\mathbf{t}).$$

Teď stačí ukázat, že $\forall\varepsilon > 0 : N_{[\,]}(\varepsilon, \mathcal{F}_d, L_1(P)) < \infty$ a použít Větu 6. Toto dokážeme indukcí podle d. Pro $d = 1$ toto platí z diskuse za důkazem Věty 6. Předpokládejme tedy, že bracketing number $\mathcal{F}_{d-1}$ je konečné pro všechna $\varepsilon > 0$. Všimněme si, že pro $\mathbf{v} \in \mathbb{R}^{d-1}, s \in \mathbb{R}$, kde $\mathbf{t} = (\mathbf{v}, s)^T$, můžeme psát

$$\mathbb{1}_{(-\infty,\mathbf{t}]} = \mathbb{1}_{(-\infty,\mathbf{v}]}\mathbb{1}_{(-\infty,s]}.$$

Nechť $\varepsilon > 0$. Z indukčního předpokladu existuje konečně mnoho $\frac{\varepsilon}{2}-$bracketů $[l_i, u_i], i \in \{1, ..., N_{[\,]}(\frac{\varepsilon}{2}, \mathcal{F}_{d-1}, L_1(P))\}$. Dále z lemmatu 1 najdeme konečné dělení

$$-\infty = s_0 < ... < s_m = \infty$$

takové, že $\forall k \in \{1,...,m\}$ platí

$$G(s_k-) - G(s_{k-1}) < \frac{\varepsilon}{2},$$

kde G je distribuční funkce $d-$té složky vektoru $\mathbf{X_1}$. Položme brackety

$$[L_{ij},U_{ij}] = [l_i\mathbb{1}_{(-\infty,s_{j-1}]}, u_i\mathbb{1}_{(-\infty,s_j)}],$$

kde $j \in \{1,...,m\}, i \in \{1, ..., N_{[\,]}(\varepsilon, \mathcal{F}_{d-1}, L_1(P))\}$. Z konstrukce $[L_{ij},U_{ij}]$ vidíme, že u_i (a také l_i) je součin nějakých indikátorů a tedy platí $|u_i| \leq 1$ ($|l_i| \leq 1$). Potom je bracketů $[L_{ij},U_{ij}]$ konečně mnoho a pro každé $\mathbf{t} \in \mathbb{R}^d$ existují indexy i,j takové, že

$$L_{ij} \leq \mathbb{1}_{(-\infty,\mathbf{t}]} \leq U_{ij}$$

a navíc

$$\begin{aligned}
P(U_{ij} - L_{ij}) &= P(u_i\mathbb{1}_{(-\infty,s_j)} - l_i\mathbb{1}_{(-\infty,s_{j-1}]}) \\
&= P(u_i\mathbb{1}_{(-\infty,s_j)} - u_i\mathbb{1}_{(-\infty,s_{j-1}]} + u_i\mathbb{1}_{(-\infty,s_{j-1}]} - l_i\mathbb{1}_{(-\infty,s_{j-1}]}) \\
&= P(u_i\mathbb{1}_{(-\infty,s_j)} - u_i\mathbb{1}_{(-\infty,s_{j-1}]}) + P(u_i\mathbb{1}_{(-\infty,s_{j-1}]} - l_i\mathbb{1}_{(-\infty,s_{j-1}]}) \\
&< \varepsilon,
\end{aligned}$$

protože

$$\begin{aligned}P(u_i \mathbb{1}_{(-\infty,s_j)} - u_i \mathbb{1}_{(-\infty,s_{j-1}]}) &\leq P(\mathbb{1}_{(-\infty,s_j)} - \mathbb{1}_{(-\infty,s_{j-1}]}) \\ &= G(s_j-) - G(s_{j-1}) \\ &< \frac{\varepsilon}{2}\end{aligned}$$

a také

$$P(u_i \mathbb{1}_{(-\infty,s_{j-1}]} - l_i \mathbb{1}_{(-\infty,s_{j-1}]}) \leq P(u_i - l_i) < \frac{\varepsilon}{2}.$$

Našli jsme tedy konečný počet $\varepsilon-$bracketů, které pokryjí celou $\mathcal{F}_d$ a tedy také $N_{[\,]}(\varepsilon, \mathcal{F}_d, L_1(P))\}$ musí být konečné. Navíc všechny funkce v $\mathcal{F}_d$ jsou měřitelné, tedy z Věty 6 dostáváme, že

$$\sup_{f \in \mathcal{F}_d} |\mathbb{P}_n f - Pf| = \sup_{t \in \mathbb{R}^d} |F_n(t) - F(t)| \xrightarrow[n \to \infty]{s.j.*} 0.$$

Díky analogickému argumentu jako v důkazu Lemmatu 3 můžeme psát

$$\sup_{t \in \mathbb{R}^d} |F_n(t) - F(t)| = \sup_{t \in \mathbb{Q}^d} |F_n(t) - F(t)|$$

a ze spočetnosti $\mathbb{Q}^d$ je náhodná veličina $\sup_{t \in \mathbb{R}^d} |F_n(t) - F(t)|$ měřitelná pro každé $n \in \mathbb{N}$ a odtud plyne i klasická konvergence skoro jistě.

$\square$

2.2 Některé příklady G-C tříd funkcí

V této sekci uvedeme několik příkladů tříd funkcí, jejichž bracketing number je konečné pro každé $\varepsilon > 0$ a tedy jsou z věty 6 Glivenkovy-Cantelliho. Na začátek uvedeme definici kompaktních metrických prostorů, kterou využijeme v prvním příkladu.

Definice 5. *Nechť (Θ, ϱ) metrický prostor, pak řekneme, že Θ je kompaktní, pokud pro každý systém $\mathcal{G}$ otevřených množin takový, že $\Theta = \bigcup_{G \in \mathcal{G}} G$ existuje konečný podsystém $\mathcal{G}^\star \subset \mathcal{G}$, pro který platí $\Theta = \bigcup_{G \in \mathcal{G}^\star} G$.*

2.2.1 Spojité funkce v parametru na kompaktu

Jako první uvedeme Příklad 19.8 z knihy van der Vaart (2000). Uvažujme třídu měřitelných funkcí

$$\mathcal{F} = \{f_\theta : \mathcal{X} \mapsto \mathbb{R} \mid \theta \in \Theta\}.$$

Pokud je funkce $\theta \mapsto f_\theta(x)$ spojitá v θ pro každé $x \in \mathcal{C}$ a Θ je podle nějaké metriky kompaktní množina, potom je $\mathcal{F}$ Glivenkova-Cantelliho třída.

Věta 7. *Nechť $\mathcal{F} = \{f_\theta : \mathcal{X} \mapsto \mathbb{R} \mid \theta \in \Theta\}$ je třída měřitelných funkcí, které mají společnou integrabilní majorantu H. Nechť dále pro nějakou metriku ϱ je (Θ, ϱ) kompaktní metrický prostor. Pokud ještě navíc je funkce $\theta \mapsto f_\theta(x)$ spojitá pro každé $x \in \mathcal{X}$, potom pro libovolnou pravděpodobostní míru P a $\varepsilon > 0$ je $N_{[\,]}(\varepsilon, \mathcal{F}, L_1(P)) < \infty$.*

Důkaz. Pro otevřenou kouli B v (Θ, ϱ) označme

$$f_B(x) = \inf_{\theta \in B} f_\theta(x), \; f^B(x) = \sup_{\theta \in B} f_\theta(x)$$

a definujme brackety na B jako $[f_B, f^B]$. Z kompaktnosti plyne separabilita, tedy Θ určitě obsahuje hustou spočetnou množinu. Z předpokladu spojitosti f_θ v parametru a použitím analogického argumentu z důkazu Lemmatu 3 dostaneme, že pro libovolnou spočetnou hustou množinu $D \subset B$ lze psát f_B (resp. f^B) jako infimum (resp. supremum) z f_θ přes $\theta \in D$ a odtud jsou f_B a f^B měřitelné. Dále pro libovolný bod $\theta \in \Theta$ a posloupnost $\delta_n \searrow 0$ označme $B_n = B(\theta, \delta_n)$ otevřenou kouli se středem θ a poloměrem δ_n. Potom

$$f^{B_n}(x) - f_{B_n}(x) \xrightarrow[n \to \infty]{} f_\theta(x) - f_\theta(x) = 0$$

ze spojitosti f_θ v parametru θ pro každé $x \in \mathcal{X}$ (jde tedy o bodovou konvergenci). Navíc z existence integrovatelné majoranty H zde platí Lebesgueova věta, viz. Věta 11.32 v Rudin (1964), a tedy také

$$P(f^{B_n} - f_{B_n}) \xrightarrow[n \to \infty]{} P(f_\theta - f_\theta) = 0. \tag{2.2}$$

Nechť $\varepsilon > 0$. Pak z (2.2) pro každé $\theta \in \Theta$ existuje otevřená koule B_θ taková, že $P(f^{B_\theta} - f_{B_\theta}) < \varepsilon$. Zřejmě platí, že

$$\Theta = \bigcup_{\theta \in \Theta} B_\theta.$$

Tedy z Definice 5 existuje konečná $\Theta^\star \subset \Theta$, taková, že $\bigcup_{\theta\in\Theta^\star} B_\theta = \Theta$. Tedy množina $\varepsilon-$bracketů $[f_{B_\theta}, f^{B_\theta}], \theta \in \Theta^\star$ je konečná, pokryje celé $\mathcal{F}$ a Věta je dokázána.

□

Příklad. Jako aplikaci předchozí věty si ukážeme stejnoměrnou konvergenci a-tých (absolutních) empirických momentů ke skutečným. Mějme náhodný výběr X_1,..., X_n z rozdělení X, kde $X > 0$ skoro jistě, $\mathbb{E}X^a < \infty$ pro každé $a \in A$ (a může být i zaporné). Větu 7 lze totiž použít na třídu funkcí z $\mathbb{R}^+$ do $\mathbb{R}$

$$\mathcal{F} = \{x \mapsto x^a \mid a \in A \subset \mathbb{R}, A \text{ je omezená a uzavřená}\}.$$

Pro $f_a : x \mapsto x^a, a \in A$ potom máme

$$\mathbb{P}_n f_a = \frac{1}{n}\sum_{i=1}^{n} X_i^a$$

a také

$$Pf_a = \mathbb{E}X^a.$$

Protože omezená a uzavřená množina v $\mathbb{R}$ je kompaktní a funkce $a \mapsto x^a$ je zřejmě spojitá pro každé $x > 0$, Věta 6 nám tedy dá

$$\sup_{f_a\in\mathcal{F}} |\mathbb{P}_n f_a - Pf_a| = \sup_{a\in\mathbb{R}} \left|\tfrac{1}{n}\textstyle\sum_{i=1}^n X_i^a - \mathbb{E}X^a\right| \xrightarrow[n\to\infty]{s.j.*} 0.$$

Ze spojitosti f_a můžeme $\sup_{a\in\mathbb{R}} \left|\frac{1}{n}\sum_{i=1}^n X_i^a - \mathbb{E}X^a\right|$ z předchozí rovnosti psát přes $a \in \mathbb{Q}$, odtud plyne měřitelnost a tedy i klasická konvergence skoro jistě.

Příklad. Také můžeme uvažovat třídu

$$\mathcal{F} = \{x \mapsto s^x \mid s \in [-1, 1]\},$$

kde $x \in \mathbb{N}_0$. Interval $[-1, 1]$ je kompaktní a funkce $s \mapsto s^x$ je spojitá pro každé $x \in \mathbb{N}_0$. Potom pro diskrétní náhodný výběr $X_1, ..., X_n$ z rozdělení X, kde $\mathsf{P}(X \in \mathbb{N}_0) = 1$, a pro $f_s \in \mathcal{F}$ máme

$$Pf_s = \mathbb{E}s^X,$$

což je vytvořující funkce náhodné veličiny X, jejímž empirickým odhadem je

$$\mathbb{P}_n f_s = \frac{1}{n}\sum_{i=1}^{n} s^{X_i}.$$

Všechny funkce v $\mathcal{F}$ jsou zřejmě omezené jedničkou a tu můžeme vzít jako integrabilní majorantu. Ověřili jsme předpoklady Věty 7, tedy platí i Věta 6 a odtud

$$\sup_{s\in[-1,1]} \left|\tfrac{1}{n}\textstyle\sum_{i=1}^n s^{X_i} - \mathbb{E}s^X\right| \xrightarrow[n\to\infty]{s.j.*} 0.$$

Ze spojitosti argumentu suprema výše toto supremum můžeme psát přes $s \in \mathbb{Q} \cap [-1,1]$, tedy platí i klasická konvergence skoro jistě.

Příklad. Mějme náhodný výběr $X_1, ..., X_n$ z rozdělení X, $\mathbb{E}\,|X| < \infty$. Jedním z možných odhadů toho, jak moc se X vzdaluje od své střední hodnoty (variability) je $\sqrt{S_n^2}$, kde S_n^2 je výběrový rozptyl. Alternativou může být $\frac{1}{n}\sum_{i=1}^n \left|X_i - \overline{X}_n\right|$. Ukážeme, že

$$\frac{1}{n}\sum_{i=1}^n \left|X_i - \overline{X}_n\right| \xrightarrow[n\to\infty]{\mathsf{P}} \mathbb{E}\,|X - \mathbb{E}X|\,. \tag{2.3}$$

Uvažujme třídu

$$\mathcal{F} = \{x \mapsto |x-a| \mid a \in A \text{ kompaktní}, \mathbb{E}X \in int(A)\}.$$

Taková A zřejmě existuje, stačí volit napríklad $A = [\mathbb{E}X - 1, \mathbb{E}X + 1]$. $\mathcal{F}$ je z Věty 7 Glivenkova-Cantelliho a tedy

$$\sup_{a\in A}\left|\tfrac{1}{n}\textstyle\sum_{i=1}^n |X_i - a| - \mathbb{E}\,|X-a|\right| \xrightarrow[n\to\infty]{s.j.} 0. \tag{2.4}$$

Označme $M(a) = \mathbb{E}\,|X-a|$ a $M_n(a) = \frac{1}{n}\sum_{i=1}^n |X_i - a|$. Potom máme

$$\begin{aligned}\left|\tfrac{1}{n}\textstyle\sum_{i=1}^n \left|X_i - \overline{X}_n\right| - \mathbb{E}\,|X - \mathbb{E}X|\right| &\le \left|M_n(\overline{X}_n) - M(\overline{X}_n)\right| \\ &\quad + \left|M(\overline{X}_n) - M(\mathbb{E}X)\right|. \end{aligned} \tag{2.5}$$

Nejprve ukážeme, že první člen na pravé straně předchozí rovnosti jde k nule v pravděpodobnosti. Ze SZVČ je $\overline{X}_n \xrightarrow[n\to\infty]{s.j.} \mathbb{E}X$ a tedy také

$$\mathsf{P}\left(\overline{X}_n \in A\right) \xrightarrow[n\to\infty]{} 1. \tag{2.6}$$

Nechť $\varepsilon > 0$. Potom s pomocí (2.4) a (2.6) dostaneme

$$\begin{aligned}\mathsf{P}\left(\left|M_n(\overline{X}_n) - M(\overline{X}_n)\right| > \varepsilon\right) &\le \mathsf{P}\left(\left|M_n(\overline{X}_n) - M(\overline{X}_n)\right| > \varepsilon, \overline{X}_n \in A\right) \\ &\quad + \mathsf{P}\left(\overline{X}_n \notin A\right) \\ &\le \mathsf{P}\left(\sup_{a\in A}\left|\tfrac{1}{n}\textstyle\sum_{i=1}^n |X_i - a| - \mathbb{E}\,|X-a|\right| > \varepsilon\right) \\ &\quad + \mathsf{P}\left(\overline{X}_n \notin A\right) \\ &\xrightarrow[n\to\infty]{} 0. \end{aligned}$$

Abychom ukázali, že člen (2.5) konverguje k 0, stačí nám díky větě o spojité transformaci (Věta 2.3 ve van der Vaart (2000)) spojitost funkce M v bodě $a = \mathbb{E}X$. Nechť $\varepsilon > 0$. Potom pro každé $\delta \in (a-\varepsilon, a+\varepsilon)$ platí

$$\begin{aligned}|\mathbb{E}|X-\delta| - \mathbb{E}|X-a|| &\le |\mathbb{E}|X-a| + \mathbb{E}|a-\delta| - \mathbb{E}|X-a|| \\ &= |a-\delta| < \varepsilon.\end{aligned}$$

Odtud plyne (2.3).

Příklad. Další aplikací bude stejnoměrná konvergence empirické charakteristické funkce

$$\varphi_n(t) = \frac{1}{n}\sum_{j=1}^n e^{itX_j}$$

k pravé, kde $X_1, ..., X_n$ je náhodný výběr z rozdělení X, ke skutečné char. funkci $\varphi(t) = \mathbb{E}e^{itX}$. Položme

$$\mathcal{F}_1 = \{x \mapsto \cos(tx) \mid t \in T \subset \mathbb{R}, T \text{ je kompaktní}\}$$

a

$$\mathcal{F}_2 = \{x \mapsto \sin(tx) \mid t \in T \subset \mathbb{R}, T \text{ je kompaktní}\}.$$

Obě třídy jsou zřejmě omezeny jedničkou a spojité s parametru, tedy $\mathcal{F}_1$ i $\mathcal{F}_2$ jsou G-C z Vět 7 a 6. Dále pro $f_1^t \in \mathcal{F}_1, f_2^t \in \mathcal{F}_2$ máme $P(f_1^t + if_2^t) = \varphi(t)$ a $\mathbb{P}_n(f_1^t + if_2^t) = \varphi_n(t)$. Celkem

$$\begin{aligned}
\sup_{t\in T} \left|\mathbb{P}_n(f_1^t + if_2^t) - P(f_1^t + if_2^t)\right| &\leq \sup_{t\in T} \left|\mathbb{P}_n f_1^t - Pf_1^t\right| + \sup_{t\in T} \left|i\mathbb{P}_n f_2^t - iPf_2^t\right| \\
&= \sup_{t\in T} \left|\mathbb{P}_n f_1^t - Pf_1^t\right| + \sup_{t\in T} \left|\mathbb{P}_n f_2^t - Pf_2^t\right| \\
&\xrightarrow[n\to\infty]{s.j.*} 0.
\end{aligned}$$

S měřitelností opět není problém a klasická konvergence skoro jistě zde platí také.

Předpoklady Věty 7 můžeme lehce přeformulovat tak, že požadavek spojitosti zesílíme na Lipschitzovskost a požadavek kompaktnosti parametrického prostoru zeslabíme na jeho omezenost. Navíc pro bracketing number dostaneme konkrétní horní mez jako funkci ε. Podobnou verzi následující Věty lze nalézt jako Příklad 19.7 ve van der Vaart (2000).

Věta 8. *Nechť $\mathcal{F} = \{f_\theta : \mathcal{X} \mapsto \mathbb{R} \mid \theta \in \Theta\}$, kde $\Theta \subset \mathbb{R}^d$ je omezená. Pokud existuje měřitelná funkce $m : \mathcal{X} \mapsto \mathbb{R}$ taková, že pro každé $x \in \mathcal{X}$ a každé $\theta_1, \theta_2 \in \Theta$ platí*

$$|f_{\theta_1}(x) - f_{\theta_2}(x)| \leq m(x)\|\theta_1 - \theta_2\|, \tag{2.7}$$

kde $\|\cdot\|$ značí Eukleidovskou normu v $\mathbb{R}^d$. Pokud ještě navíc pro pravděpodobnostní míru P platí $Pm < \infty$, pak pro každé $\varepsilon > 0$ je

$$N_{[\,]}(\varepsilon, \mathcal{F}, L_1(P)) < \left(\frac{\sqrt{d}K\,diam(\Theta)}{\varepsilon}\right)^d,$$

kde $K = Pm$ a

$$diam(\Theta) = \sup_{x,y\in\Theta} \|x - y\|.$$

Důkaz. Nechť $\varepsilon > 0$. Θ je z předpokladu omezená, tedy v ní existuje konečná $\frac{\varepsilon}{2K}$-síť A. Brackety pak volíme jako $[f_{\theta_a} - \frac{m\varepsilon}{2K}, f_{\theta_a} + \frac{m\varepsilon}{2K}], a \in A$. Tyto brackety mají zřejmě velikost ε a pro všechna θ t. ž. $\|\theta - \theta_a\| \leq \frac{\varepsilon}{2K}$, pak z Lipschitzovské vlastnosti (2.7)

$$f_{\theta_a} - \frac{m\varepsilon}{2K} \leq f_\theta \leq f_{\theta_a} + \frac{m\varepsilon}{2K},$$

tedy ε-bracketů, které pokryjí $\mathcal{F}$ je potřeba stejně jako koulí o poloměru $\frac{\varepsilon}{2K}$, které pokryjí Θ. Prostor Θ jistě pokryjeme $\left(\frac{\sqrt{d}K\mathrm{diam}(\Theta)}{\varepsilon}\right)^d$ krychlemi o straně $\frac{\varepsilon}{\sqrt{d}K}$ a koule opsané těmto krychlím tedy také pokryjí Θ a mají poloměr $\frac{\sqrt{d}}{2}\frac{\varepsilon}{\sqrt{d}K} = \frac{\varepsilon}{2K}$. □

Pokud nás zajímá pouze, zda $N_{[\,]}(\varepsilon, \mathcal{F}, L_1(P)) < \infty$, nemá Věta 8 příliš velký význam, protože omezený parametrický prostor $\Theta \subset \mathbb{R}^d$ ke kompaktnosti stačí uzavřít. Z toho, že k hranici se dá dokonvergovat z vnitřku její množiny, má $\mathcal{F}$ Lipschitzovskou vlastnost (2.7) (a tedy je spojitá ve θ) i na uzávěru Θ. Pokud tedy pro $\mathcal{F}$ existuje integrabilní majoranta, dostáváme Větu 7. Aplikaci Věty 8 najdeme například v situaci, kdy nás zajímá řád konvergence v Definici 2. Přesněji, pokud chceme vědět, pro jaká $\{a_n\}_{n\in\mathbb{N}}, a_n \nearrow \infty$ ještě platí

$$a_n \sup_{f\in\mathcal{F}} |\mathbb{P}_n f - Pf| \xrightarrow[n\to\infty]{s.j.*} 0.$$

2.2.2 Monotónní funkce a funkce s omezenou variací

Ukazuje se, že i libovolná třída monotónních funkcí na intervalu $[0,1]$ je Glivenkova-Cantelliho.

Věta 9. *Nechť $\varepsilon > 0, r \geq 1$ P je pravděpodobnostní míra a*

$$\mathcal{F} = \{f : \mathbb{R} \mapsto [0,1] \mid f \text{ je monotónní}\},$$

potom existuje konstanta K závislá pouze na r taková, že platí

$$\log N_{[\,]}(\varepsilon, \mathcal{F}, L_r(P)) \leq K\left(\frac{1}{\varepsilon}\right)\log\left(\frac{1}{\varepsilon}\right).$$

Důkaz je proveden v knize van der Vaart a Wellner (1996), Věta 2.7.5. Je velmi technický a my jej v práci neuvádíme.

Příklad. Mějme opět náhodný výběr $X_1,...,X_n$ z rozdělení X s distribuční funkcí F. K odhadování F se někdy používá tzv. vyhlazovací metoda. Pokud pouze předpokládáme, že rozdělení X s hustotou f je spojité, můžeme místo empirické distribuční funkce uvažovat odhad

$$G_n(t) = \frac{1}{n}\sum_{i=1}^{n} K\left(\tfrac{t-X_i}{h_n}\right),$$

kde K je distribuční funkce nějaké spojité náhodné veličiny a $\{h_n\}_{n\in\mathbb{N}}$ je nezáporná monotónní posloupnost jdoucí k nule zprava. Namísto indikátorů u empirické distribuční funkce zde sčítáme spojité distribuční funkce posunuté o hodnoty X_i, každou s váhou $\frac{1}{n}$, a tedy i F_n je spojitá. Ukážeme, že

$$\sup_{t\in\mathbb{R}} |G_n(t) - F(t)| \xrightarrow[n\to\infty]{s.j.} 0.$$

Definujme

$$\mathcal{F} = \{x \mapsto K\left(\tfrac{t-x}{h}\right) \mid h > 0, t \in \mathbb{R}\}.$$

Tato třída je z Věty 9 G-C. Dále máme

$$\begin{aligned}\sup_{t\in\mathbb{R}}\left|\frac{1}{n}\sum_{i=1}^{n} K\left(\tfrac{t-X_i}{h_n}\right) - F(t)\right| &\leq \sup_{f\in\mathcal{F}}\left|\frac{1}{n}\sum_{i=1}^{n} K\left(\tfrac{t-X_i}{h}\right) - \mathbb{E}K\left(\tfrac{t-X}{h}\right)\right| \\ &\quad + \sup_{t\in\mathbb{R}}\left|\mathbb{E}K\left(\tfrac{t-X}{h_n}\right) - F(t)\right|.\end{aligned}$$

První člen na pravé straně předchozí nerovnosti jde skoro jistě k nule. Dále máme pro $t \in \mathbb{R}$

$$\mathbb{E}K\left(\tfrac{t-X}{h_n}\right) = \int_{-\infty}^{\infty} K\left(\tfrac{t-x}{h_n}\right) f(x)dx,$$

kde f je hustota rozdělení X. Vidíme, že posloupnost $\{K\left(\tfrac{t-x}{h_n}\right) f(x)\}_{n\in\mathbb{N}}$ má pro $x < t$ limitu $f(x)$ a pro $x > t$ má limitu 0. Také zřejmě pro každé $x \in \mathbb{R}$ platí, že $0 \leq K\left(\tfrac{t-x}{h_n}\right) f(x) \leq f(x)$ a tedy z Lebesgueovy věty

$$\mathbb{E}K\left(\tfrac{t-X}{h_n}\right) = \int_{-\infty}^{\infty} K\left(\tfrac{t-x}{h_n}\right) f(x)dx \xrightarrow[n\to\infty]{} \int_{-\infty}^{t} f(x)dx = F(t). \tag{2.8}$$

Tedy platí bodová konvergence. Z toho, že K je distribuční funkce a z linearity a monotonie střední hodnoty splňuje také $\mathbb{E}K\left(\tfrac{t-X}{h_n}\right)$ vlastnosti distribuční funkce. K dokázání stejnoměrné konvergence už stačí postupovat analogicky jako v první půlce důkazu Věty 2, kde místo F_n teď máme G_n. Pro $\varepsilon > 0$ najdeme z Lemmatu 1 konečné dělení splňující (1.1) a $\sup_{t\in\mathbb{R}} \left|\mathbb{E}K\left(\tfrac{t-X}{h_n}\right) - F(t)\right|$ odhadneme pomocí ε a členu, který jde k nule z (2.8). Tedy

$$\limsup_{n\to\infty} \sup_{t\in\mathbb{R}} \left|\mathbb{E}K\left(\tfrac{t-X}{h_n}\right) - F(t)\right| \leq \varepsilon.$$

Protože $\varepsilon > 0$ bylo libovolné, platí

$$\sup_{t\in\mathbb{R}} \left|\mathbb{E}K\left(\tfrac{t-X}{h_n}\right) - F(t)\right| \xrightarrow[n\to\infty]{} 0.$$

Celkem tedy

$$\sup_{t\in\mathbb{R}} \left|\frac{1}{n}\sum_{i=1}^{n} K\left(\tfrac{t-X_i}{h_n}\right) - F(t)\right| \xrightarrow[n\to\infty]{s.j.} 0.$$

Z Věty 9 lze celkem snadno ukázat, že i třída monotónních funkcí na obecném intervalu $[a, b]$ má pro každé $\varepsilon > 0$ konečné bracketing number.

Věta 10. *Třída monotónních funkcí $f : \mathbb{R} \mapsto [a, b], a < b$ má pro každé $\varepsilon > 0$ konečné bracketing number.*

Důkaz. Nechť $\varepsilon > 0$. Pro libovolnou f s hodnotami $[a, b]$ je funkce $\frac{f-a}{b-a}$ je monotónní s hodnotami na $[0, 1]$. Aplikací Věty 9 na třídu

$$\mathcal{F} = \{\tfrac{f-a}{b-a} : \mathbb{R} \mapsto [0, 1] \mid f \text{ je monotónní}\}$$

existuje konečně mnoho $\frac{\varepsilon}{b-a}$-bracketů $[l_i, u_i]$, které pokryjí celou $\mathcal{F}$, tedy pro každou $\frac{f-a}{b-a} \in \mathcal{F}$ exituje index i takový, že

$$l_i \leq \frac{f-a}{b-a} \leq u_i$$

a odtud

$$(b-a)l_i + a \leq f \leq (b-a)u_i + a,$$

tedy brackety $[(b-a)l_i + a, (b-a)u_i + a]$ pokryjí všechny monotónní f na $[a, b]$ a navíc

$$P((b-a)u_i + a - (b-a)l_i - a) = (b-a)P(u_i - l_i) < (b-a)\frac{\varepsilon}{b-a} = \varepsilon.$$

□

Teď budeme uvažovat třídu funkcí s omezenou variací na $\mathbb{R}$, viz definice níže, a ukážeme, že tato třída je Glivenkova-Cantelliho.

Definice 6. *Nechť $\alpha, \beta \in \mathbb{R}, \alpha < \beta$ a $f : [\alpha, \beta] \mapsto \mathbb{R}$ je funkce. Označme D množinu všech konečných dělení $a = x_0 < x_1 < ... < x_n = b$. Potom výrazu*

$$V(f, \alpha, \beta) = \sup_D \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|$$

říkáme totální variace funkce f na intervalu $[\alpha, \beta]$ a pokud $V(f, \alpha, \beta) < \infty$ pak f má tzv. omezenou variaci na $[\alpha, \beta]$.

Věta 11. *Pro každé $a, b, \alpha, \beta \in \mathbb{R}, a < b, \alpha < \beta, \varepsilon > 0$, $K > 0$ a třídu*

$$\mathcal{F}_K = \{f : [\alpha, \beta] \mapsto [a, b] \mid V(f, \alpha, \beta) \leq K\}$$

je $N_{[\,]}(\varepsilon, \mathcal{F}_K, L_1(P)) < \infty$.

Důkaz. Použijeme charakterizaci funkcí s omezenou variací, např. Věta 17. 8 v Laczkovich a Sós (2015), podle které můžeme každou takovou $f : [\alpha, \beta] \mapsto [a, b]$ psát jako $f = g - h$, kde g a h jsou neklesající, konkrétně $g(x) = V(f, \alpha, x)$ a $h(x) = V(f, \alpha, x) - f(x)$, kde $x \in [\alpha, \beta]$. Navíc g nabývá hodnot v intervalu $[0, K]$ a h v intervalu $[-b, K - a]$. Označme $\mathcal{G} = \{g : [\alpha, \beta] \mapsto [0, K] \mid g \text{ je neklesající}\}$ a $\mathcal{H} = \{h : [\alpha, \beta] \mapsto [-b, K - a] \mid h \text{ je neklesající}\}$ a nechť $\varepsilon > 0$. Z Věty 10 pro tyto třídy existuje konečně mnoho $\frac{\varepsilon}{2}$-bracketů $[l_i^g, u_i^g]_{\frac{\varepsilon}{2}}$ a $[l_j^h, u_j^h]_{\frac{\varepsilon}{2}}$. Pro každou $f \in \mathcal{F}_K, f = g - h$ tedy existují indexy i, j takové, že

$$l_i^g - u_j^h \leq f \leq u_i^g - l_j^h$$

a navíc $P(u_i^g - l_j^h - l_i^g + u_j^h) < \varepsilon$. Našli jsme tedy konečně pro libovolné $\varepsilon > 0$ konečné mnoho ε-bracketů, které pokryjí celou $\mathcal{F}_K$.

□

Poznámka. Podobnou úvahou by se dala předchozí Věta zobecnit pro funkce $f : \mathbb{R} \mapsto [a, b]$ s variací omezenou K.

Závěr

V práci jsme podrobně dokázali klasickou i zobecněnou verzi Glivenkovy-Cantelliho věty, uvedli jsme některé jejich aplikace a důsledky a několik netriviálních příkladů G-C tříd funkcí.

Obsah práce ale pokrývá jen malou část teorie empirických procesů. Dále by se daly studovat například tzv. covering numbers, které jsou v jistém smyslu podobné bracketing numbers a pomocí nichž lze nalézt jinou postačující podmínku, podle které je třída $\mathcal{F}$ Glivenkova-Cantelliho, viz např. Věta 24 v Pollard (1984). Zobecněné Glivenkově-Catelliho větě se někdy také říká stejnoměrný zákon velkých čísel.

Další partií teorie empirických procesů jsou Donskerovy věty, které zase mluví o konvergenci empirických procesů v distribuci a někdy se jim říká stejnoměrné centrální limitní věty. Více lze najít ve van der Vaart a Wellner (1996).

Seznam použité literatury

Cantelli, F. P. (1933). Sulla determinazione empirica delle leggi di probabilita. *Giorn. Inst. Ital. Attuari*, **4**, 421 – 424.

Dupač, V. a Hušková, M. (2005). *Pravěpodobnost a matematická statistika.* Nakladatelství Karolinum, Praha. ISBN 976-80-246-0009-3.

Glivenko, V. I. (1933). Sulla determinazione empirica delle leggi di probabilita. *Giorn. Inst. Ital. Attuari*, **4**, 92 – 99.

Laczkovich, M. a Sós, V. (2015). *Real Analysis.* Springer, New York. ISBN 978-1-4939-2766-1.

Pollard, D. (1984). *Convergence of stochastic processes.* Springer-Verlag, New York. ISBN 0-387-90990-7.

Rudin, W. (1964). *Principles of Mathematical Analysis.* McGraw-Hill, Inc., New York. ISBN 0-07-054235-X.

Serfling, R. J. (2002). *Approximation Theorems of Mathematical Statistics.* John Wiley and Sons, inc., New York. ISBN 0-471-21927-4.

van der Vaart, A. W. (2000). *Asymptotic Statistics.* Cambridge University press, New York. ISBN 978-0-521-78450-4.

van der Vaart, A. W. a Wellner, J. A. (1996). *Weak Convergence and Empirical Processes.* Springer-Verlag, New York. ISBN 0-387-94640-3.