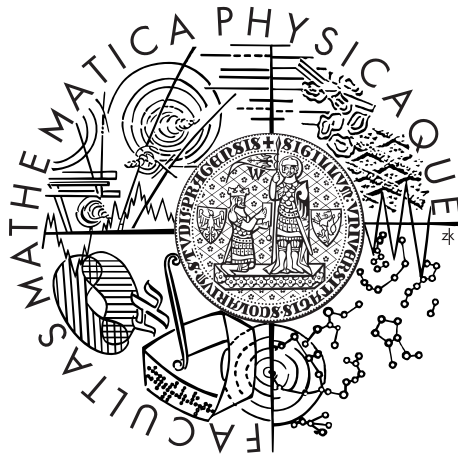


Univerzita Karlova v Praze  
Matematicko-fyzikální fakulta

## BAKALÁŘSKÁ PRÁCE



Eva Kielkowská

## Bodové procesy odvozené od Poissonova procesu

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: prof. RNDr. Viktor Beneš, DrSc.

Studijní program: Matematika

Studijní obor: Finanční matematika

Praha 2014

Ráda bych poděkovala svému vedoucímu, profesorovi Benešovi, za cenné rady a odborný dohled nad prací. Děkuji také za trpělivost a ochotu, kterou mi během konzultací v průběhu roku věnoval.

Prohlašuji, že jsem tuto bakalářskou práci vypracoval(a) samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

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 v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V ..... dne .....

Název práce: Bodové procesy odvozené od Poissonova procesu

Autor: Eva Kielkowská

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: prof. RNDr. Viktor Beneš, DrSc., Katedra pravděpodobnosti a matematické statistiky

Abstrakt: V práci jsou studovány konečné bodové procesy s hustotou vzhledem k Poissonovu procesu. Zabýváme se speciálními funkcionaly bodových procesů, které nazýváme  $U$ -statistiky. Střední hodnotu těchto funkcionalů dostaneme jako součin středních hodnot funkcionalů Poissonova procesu. S pomocí diferencí a jádra  $U$ -statistiky odvodíme obecný vzorec pro střední hodnotu tohoto funkcionalu, ve kterém se vyskytuje podmíněná intenzita bodového procesu. Výpočet středních hodnot vybraných  $U$ -statistik pro Straussův proces ukazuje použití zmíněného vzorce. Demonstrováno je také ověření předpokladů z vlastností Poissonova procesu. V závěru se uvádí výsledky numerických výpočtů středních hodnot na základě simulací Straussova procesu.

Klíčová slova: hustota vzhledem k Poissonovu procesu, podmíněná intenzita, modely a charakteristiky bodových procesů

Title: Point processes derived from the Poisson process

Author: Eva Kielkowská

Department: Department of Probability and Mathematical Statistics

Supervisor: prof. RNDr. Viktor Beneš, DrSc., Department of Probability and Mathematical Statistics

Abstract: Finite point processes having a density with respect to the Poisson process are investigated. We are interested in special functionals called  $U$ -statistics. The mean value of such a functional is obtained as a product of mean values of functionals of the Poisson point process. Using difference operators and kernels of an  $U$ -statistic, we can derive a general formula for the mean value of the functional, in which the conditional intensity of a point process is used. The calculation of mean values of selected  $U$ -statistics of the Strauss process shows the application of the formula. The assumptions of the formula are verified via characteristics of the Poisson process. At the end, the results of numerical calculations of mean values based on simulations of the Strauss process are presented.

Keywords: density w.r.t. the Poisson proces, conditional intensity, models and characteristics of point processes

# Obsah

Úvod	2
<b>1 Teoretické základy</b>	<b>3</b>
1.1 Bodové procesy . . . . .	3
1.2 Bodové procesy s hustotou . . . . .	4
1.3 Rozklad funkcionálu Poissonova procesu . . . . .	6
1.4 U-statistika . . . . .	7
<b>2 Teoretické výsledky pro bodové procesy s hustotou</b>	<b>12</b>
2.1 Střední hodnota funkcionálu . . . . .	12
2.2 Vztahy pro Straussův proces . . . . .	14
<b>3 Numerické výsledky</b>	<b>16</b>
3.1 Simulace Straussova procesu . . . . .	16
3.2 Aproximace vzorce pro střední počet bodů . . . . .	16
3.3 Výsledky numerických výpočtů . . . . .	17
3.4 Aproximace vzorce pro vybranou $U$ -statistiku druhého řádu . . . . .	19
3.5 Výsledky numerických výpočtů II . . . . .	20
<b>Literatura</b>	<b>21</b>
<b>Seznam obrázků</b>	<b>22</b>
<b>Seznam tabulek</b>	<b>23</b>

# Úvod

Tématem mé bakalářské práce jsou bodové procesy v  $\mathbb{R}^d$  odvozené od Poissonova procesu. Jsou to matematické modely pro náhodný výskyt událostí v prostoru. V textu se čtenář seznámí s konečnými bodovými procesy s hustotou a jejich podmíněnou intenzitou. Dále je zaveden pojem  $U$ -statistiky jako funkcionálu bodového procesu, jehož střední hodnota je potom hlavní zkoumanou vlastností bodových procesů v této práci.

Při psaní práce jsem pracovala s několika časopiseckými prameny, z jejichž poznatků je sestavena teoretická část. Dále jsem spočítala vlastní příklady a pro simulované modely jsem provedla numerickou studii k ověření některých vztahů.

V první kapitole jsou sepsány potřebné teoretické základy. Definují se zde základní pojmy a jsou uvedena tvrzení, která jsou dále využívána ve výpočtech středních hodnot. Kapitola zakončují dva příklady na výpočet střední hodnoty vybrané  $U$ -statistiky. Druhá kapitola se věnuje výhradně střední hodnotě funkcionálu bodového procesu a uvádí obecnou větu, podle níž je možné tuto hodnotu přímo vypočítat. Výsledky příkladů z první kapitoly jsou zde upraveny pro Straussův proces. Třetí kapitola přináší numerické ověření platnosti věty o střední hodnotě funkcionálu. Popisuje postup simulace bodového procesu, zde konkrétně Straussova procesu, aproximaci vzorce pro střední hodnotu a přináší srovnání numerických výsledků s odhadem skutečných hodnot.

Studium uvedených vlastností bodových procesů je zajímavé nejen z čistě matematického hlediska, ale může najít i široké uplatnění. V praxi mohou totiž bodové procesy simulovat v průmyslu například přítomnost příměsí v látce nebo v pojišťovnictví výskyt škodních událostí v určité oblasti.

# Kapitola 1

## Teoretické základy

### 1.1 Bodové procesy

Bud'  $(\mathbb{R}^d, \mathcal{B}^d)$   $d$ -rozměrný Eukleidovský prostor s Borelovskou  $\sigma$ -algebrou. Označme symbolem  $\nu_d$  Lebesgueovu míru.

**Definice 1.** Čítací míra na  $\mathbb{R}^d$  je míra s hodnotami v  $\{0, 1, 2, \dots\} \cup \{+\infty\}$ , která je konečná na každé kompaktní množině.

Nechť  $\mathbf{N}$  je množina všech čítacích měr na  $\mathbb{R}^d$ . Označme

$$E_{B,k} = \{N \in \mathbf{N} : N(B) = k\}$$

pro kompaktní množinu  $B \in \mathcal{B}^d$ , kde  $k$  je nezáporné celé číslo.

**Definice 2.** Nechť  $\mathcal{N}$  je  $\sigma$ -algebra podmnožin  $\mathbf{N}$  generovaná všemi množinami typu  $E_{B,k}$ . Prostor  $(\mathbf{N}, \mathcal{N})$  nazveme prostorem realizací bodového procesu na  $\mathbb{R}^d$ .

**Definice 3** (Bodový proces). Nechť  $(\Omega, \mathcal{A}, \mathbb{P})$  je pravděpodobnostní prostor a  $(\mathbf{N}, \mathcal{N})$  je prostor realizací. Potom měřitelné zobrazení  $\mathbf{X} : \Omega \rightarrow \mathbf{N}$  je bodový proces, kde  $N_{\mathbf{X}}$  je čítací míra přiřazená tomuto procesu.

Nadále budeme předpokládat, že se jedná o jednoduchý bodový proces, tedy, že platí

$$N_{\mathbf{X}}(\{x\}) \leq 1 \quad \forall x \in \mathbb{R}^d$$

s pravděpodobností 1.

**Definice 4** (Rozdělení bodového procesu). Rozdělení bodového procesu  $\mathbf{X}$  je pravděpodobnostní míra  $\mathbf{P}_{\mathbf{X}}$  na prostoru realizací  $(\mathbf{N}, \mathcal{N})$  definovaná jako

$$\mathbf{P}_{\mathbf{X}}(A) = \mathbb{P}(N_{\mathbf{X}} \in A), \quad A \in \mathcal{N}.$$

**Definice 5.** Nechť  $\mathbf{X}$  je bodový proces. Potom vzorec

$$\lambda(B) = \mathbb{E}[N_{\mathbf{X}}(B)], \quad B \in \mathcal{B}^d,$$

definuje míru na  $\mathbb{R}^d$ , kterou nazýváme míra intenzity procesu  $\mathbf{X}$  (za předpokladu, že  $\lambda(B) < \infty$  pro všechny kompaktní množiny  $B$ ).

**Definice 6.** Předpokládejme, že míra intenzity  $\lambda$  bodového procesu  $\mathbf{X}$  na  $\mathbb{R}^d$  splňuje

$$\lambda(B) = \int_B \beta(u) du$$

pro nějakou nezápornou funkci  $\beta$ . Potom funkci  $\beta$  nazýváme intenzitou procesu  $\mathbf{X}$ .

**Věta 1** (Campbellův vzorec). (Baddeley, 2007, Věta 2.2). Nechť  $\mathbf{X}$  je bodový proces na  $\mathbb{R}^d$  s mírou intenzity  $\lambda$  a nechť  $f: \mathbb{R}^d \rightarrow \mathbb{R}$  je měřitelná funkce. Potom náhodná veličina

$$Z = \sum_{x \in \mathbf{X}} f(x)$$

je náhodná veličina se střední hodnotou

$$\mathbb{E}\left[\sum_{x \in \mathbf{X}} f(x)\right] = \int_{\mathbb{R}^d} f(x) \lambda(dx).$$

Tedy pro bodový proces  $\mathbf{X}$  na  $\mathbb{R}^d$  s intenzitou  $\beta$  dostáváme

$$\mathbb{E}\left[\sum_{x \in \mathbf{X}} f(x)\right] = \int_{\mathbb{R}^d} f(x) \beta(x) dx.$$

**Definice 7** (Poissonův proces). Poissonův proces  $\eta$  s konstantní intenzitou  $\beta > 0$  je bodový proces na  $\mathbb{R}^d$  takový, že:

- pro každou kompaktní množinu  $B$ , má náhodná veličina  $N_\eta(B)$  Poissonovo rozdělení se střední hodnotou  $\beta \nu_d(B)$ ;
- jestliže  $B_1, \dots, B_m \in \mathcal{B}^d$  jsou po dvou disjunktní, potom  $N_\eta(B_1), \dots, N_\eta(B_m)$  jsou nezávislé náhodné veličiny.

## 1.2 Bodové procesy s hustotou

Analogicky jako v sekci (1.1) můžeme definovat bodový proces na množině  $S \in \mathcal{B}^d$ , kde  $\nu_d(S) > 0$ .

**Definice 8.** Bodový proces  $\mathbf{X}$  na  $S \subset \mathbb{R}^d$ , pro který platí, že  $N_{\mathbf{X}}(S) < \infty$  skoro jistě, se nazývá konečný bodový proces.

Realizace konečného bodového procesu  $\mathbf{X}$  náleží do prostoru

$$\mathbf{N}^f = \{N_{\mathbf{X}} \in \mathbf{N} : N_{\mathbf{X}}(S) < \infty\},$$

což je prostor konečných čítacích měr na  $S$ . Tento prostor můžeme rozložit na podprostory podle celkového počtu bodů:

$$\mathbf{N}^f = \mathbf{N}_0 \cup \mathbf{N}_1 \cup \mathbf{N}_2 \cup \dots,$$

kde  $\mathbf{N}_k = \{N_{\mathbf{X}} \in \mathbf{N} : N_{\mathbf{X}}(S) = k\}$  pro každé  $k = 1, 2, \dots$  je množina všech čítacích měr s celkovou mírou  $k$ , tedy množina všech konfigurací  $k$  bodů. Toto můžeme vyjádřit explicitněji jako prostor uspořádaných  $k$ -tic

$$S^{!k} = \{(x_1, x_2, \dots, x_k) : x_i \in S, x_i \neq x_j \text{ pro všechna } i \neq j\}.$$



Definujme nyní zobrazení  $I_k : S^{lk} \rightarrow \mathbf{N}_k$  předpisem

$$I_k(x_1, \dots, x_k) = \delta_{x_1} + \dots + \delta_{x_k}, \quad (1.1)$$

kde  $\delta$  označuje Diracovu míru. Tímto získáváme

$$\mathbf{N}_k \equiv S^{lk} / \sim,$$

kde  $\sim$  je relace ekvivalence vzhledem k permutacím, máme tedy:

$$(x_1, \dots, x_k) \sim (y_1, \dots, y_k) \Leftrightarrow \{x_1, \dots, x_k\} = \{y_1, \dots, y_k\}.$$

Realizace bodového procesu je čítací míra nebo duálně podle (1.1) konečná množina bodů.

Označme dále symbolem  $\mathbf{P}_\eta$  rozdělení Poissonova procesu s mírou intenzity  $\lambda$ , nechť  $0 < \lambda(S) < +\infty$ .

**Definice 9** (Hustota bodového procesu). *Nechť  $p : \mathbf{N}^f \rightarrow \mathbb{R}_+$  je měřitelná funkce splňující*

$$\int_{\mathbf{N}} p(\mathbf{x}) \mathbf{P}_\eta(d\mathbf{x}) = 1.$$

Definujme

$$\mathbf{P}(A) = \int_A p(\mathbf{x}) \mathbf{P}_\eta(d\mathbf{x})$$

pro každé  $A \in \mathcal{N}$ . Potom  $\mathbf{P}$  je rozdělení bodového procesu. Funkci  $p$  nazýváme hustotou bodového procesu s rozdělením  $\mathbf{P}$  (vzhledem k Poissonovu procesu).

**Lemma 2.** (Baddeley, 2007, Lemma 4.1)

*Pro bodový proces  $\mathbf{X}$  s hustotou  $q$  máme*

$$\begin{aligned} \mathbb{P}(\mathbf{X} \in A) &= \\ &= e^{-\lambda(S)} \sum_{n=0}^{\infty} \frac{1}{n!} \int_S \dots \int_S \mathbf{1}\{I_n(x_1, \dots, x_n) \in A\} q(I_n(x_1, \dots, x_n)) \lambda(dx_1) \dots \lambda(dx_n) \end{aligned}$$

pro každou událost  $A \in \mathcal{N}$  a platí, že

$$\begin{aligned} \mathbb{E}[g(\mathbf{X})] &= \\ &= e^{-\lambda(S)} \sum_{n=0}^{\infty} \frac{1}{n!} \int_S \dots \int_S g(I_n(x_1, \dots, x_n)) q(I_n(x_1, \dots, x_n)) \lambda(dx_1) \dots \lambda(dx_n) \end{aligned} \quad (1.2)$$

pro každou integrovatelnou funkci  $g : \mathbf{N}^f \rightarrow \mathbb{R}_+$ .

Identity v Lemmatu 2 můžeme také přepsat jako

$$\begin{aligned} \mathbb{P}(\mathbf{X} \in A) &= \mathbb{E}[q(\eta) \mathbf{1}_A(\eta)] \\ \mathbb{E}[g(\mathbf{X})] &= \mathbb{E}[g(\eta) q(\eta)], \end{aligned} \quad (1.3)$$

kde  $\eta$  je Poissonův proces s mírou intenzity  $\lambda$ . Náhodnou veličinu  $g(\mathbf{X})$  nazýváme také funkcional bodového procesu.

**Definice 10** (Podmíněná intenzita). *Nechť  $p$  je hustota konečného bodového procesu  $\mathbf{X}$  na omezené množině  $S$  na  $\mathbb{R}^d$ . Předpokládejme, že*

$$p(\mathbf{x}) > 0 \Rightarrow p(\mathbf{y}) > 0 \quad \text{pro každé } \mathbf{x} \subset \mathbf{y},$$

kde  $\mathbf{x}, \mathbf{y} \in \mathbb{N}^f$ . Podmíněná intenzita procesu  $\mathbf{X}$  je rovna

$$\lambda^*(u, \mathbf{x}) = \frac{p(\mathbf{x} \cup \{u\})}{p(\mathbf{x})},$$

pro  $\mathbf{x} \in \mathbb{N}^f, u \in S$ .

**Definice 11.** *Podmíněnou intenzitu  $n$ -tého řádu definujeme pro  $n \geq 1$  jako*

$$\lambda_n^*(u_1, \dots, u_n, \mathbf{x}) = \frac{p(\mathbf{x} \cup \{u_1, \dots, u_n\})}{p(\mathbf{x})},$$

pro  $\mathbf{x} \in \mathbb{N}^f, u_1, \dots, u_n \in S$ . Dále klademe  $\lambda_0^* \equiv 1$ .

**Příklad 1.** (Straussův proces) *Pro parametry  $\beta > 0, 0 \leq \gamma \leq 1$  a  $r > 0$ , máme hustotu Straussova procesu*

$$p(\mathbf{x}) = \alpha \beta^{n(\mathbf{x})} \gamma^{s(\mathbf{x})},$$

kde  $n(\mathbf{x})$  je počet bodů  $\mathbf{x}$  a

$$s(\mathbf{x}) = \sum_{x, y \in \mathbf{x}} 1_{[\|x-y\| \leq r]}$$

je počet dvojic bodů, které jsou od sebe vzdáleny nejvýše o vzdálenost  $r$ .  $\alpha$  je normalizující konstanta, kterou neumíme explicitně vyjádřit. Pro podmíněnou intenzitu získáváme vztah

$$\lambda^*(u, \mathbf{x}) = \beta \gamma^{t(u, \mathbf{x})},$$

kde

$$t(u, \mathbf{x}) = \sum_{y \in \mathbf{x}} 1_{[\|u-y\| \leq r]}.$$

### 1.3 Rozklad funkcionálu Poissonova procesu

**Definice 12.** *Pro každou měřitelnou funkci  $F : \mathbb{N}^f \rightarrow \mathbb{R}$  je funkce  $D_y F$  na  $\mathbb{N}^f$  definována jako*

$$D_y F = F(\mathbf{x} + \delta_y) - F(\mathbf{x}), \quad \mathbf{x} \in \mathbb{N}^f,$$

kde  $\delta_y$  je Diracova míra v bodě  $y \in S$ . Funkci  $D_y F$  nazveme diferencí prvního řádu.

Iterováním této funkce můžeme definovat i difference vyšších řádů, a to jako:

$$D_{y_1, \dots, y_n}^n = D_{y_1}^1 D_{y_2, \dots, y_n}^{n-1} F, \quad y_1, \dots, y_n \in S,$$

kde  $D^1$  je difference prvního řádu. Dále budeme definovat pojem jádra funkcionálu bodového procesu.  $\eta$  je Poissonův proces s mírou intenzity  $\lambda$ .

**Definice 13.** *Bud'  $\mathbf{X}$  konečný bodový proces;  $F : \mathbf{N}^f \rightarrow \mathbb{R}$ . Měřitelnou funkci  $T_n^{\mathbf{X}}F$  na  $S^n$  definujeme předpisem*

$$T_n^{\mathbf{X}}F(y_1, \dots, y_n) = \mathbb{E}D_{y_1, \dots, y_n}^n F(\mathbf{X})$$

*(pokud střední hodnota existuje) a nazveme ji jádrem funkcionálu bodového procesu. V případě, že se bude jednat o Poissonův proces, budeme místo  $T_n^{\mathbf{X}}F$  psát zkráceně  $T_n F$ .*

Definujeme  $T_0^{\mathbf{X}}F = \mathbb{E}F(\mathbf{X})$ ,  $\|\cdot\|_n$  a  $\langle \cdot, \cdot \rangle_n$  značí normu resp. skalární součin v prostoru funkcí  $L^2(S^n, \mathcal{B}(S)^n, \lambda^n)$ , tj.

$$\langle f, g \rangle_n = \int_{S^n} f(x)g(x)\lambda^n(dx).$$

Dále značíme

$$L^2(\mathbf{P}_\eta) = \{F : \mathbf{N}^f \rightarrow \mathbb{R}; \mathbb{E}F(\eta)^2 < \infty\}.$$

**Věta 3.** *(Last a Penrose, 2011, Věta 1.1) Nechť  $F \in L^2(\mathbf{P}_\eta)$ . Potom*

$$\mathbb{E}F(\eta)^2 = (\mathbb{E}F(\eta))^2 + \sum_{n=1}^{\infty} \frac{1}{n!} \|T_n F\|_n^2.$$

*Pokud je také  $G \in L^2(\mathbf{P}_\eta)$ , potom*

$$\mathbb{E}F(\eta)G(\eta) = (\mathbb{E}F(\eta))(\mathbb{E}G(\eta)) + \sum_{n=1}^{\infty} \frac{1}{n!} \langle T_n F, T_n G \rangle_n. \quad (1.4)$$

Vzorec (1.4) budeme aplikovat v situaci z (1.3), kdy pomocí střední hodnoty součinu funkcionálů Poissonových procesů dostaneme střední hodnotu funkcionálu bodového procesu s hustotou. Omezíme se na třídu funkcionálů nazývaných  $U$ -statistiky.

## 1.4 U-statistika

Bud'  $L_s^2(S^k)$  prostor funkcí  $f : S^k \rightarrow \mathbb{R}$  symetrických ve všech proměnných s  $\int_{S^k} f^2 d\lambda^k < +\infty$ . Tyto funkce jsou tedy invariantní vzhledem k permutacím argumentů.

**Definice 14** ( $U$ -statistika). *Náhodná veličina*

$$F(\mathbf{X}) = \sum_{(x_1, \dots, x_k) \in \mathbf{X}_{\neq}^k} f(x_1, \dots, x_k) \quad (1.5)$$

*pro  $f \in L_s^2(S^k)$  se nazývá  $U$ -statistika řádu  $k$  bodového procesu  $\mathbf{X}$ .*

Symbol  $\mathbf{X}_{\neq}^k$  označuje množinu všech  $k$ -tic bodů procesu  $\mathbf{X}$ , které jsou navzájem různé.

**Lemma 4.** (Reitzner a Schulte, 2013, Lemma 3.3) Necht'  $F \in L^2(\mathbf{P}_\eta)$  je  $U$ -statistika řádu  $k$ . Potom pro diferenci  $F$  máme

$$D_{y_1}F(\mathbf{X}) = k \sum_{(x_1, \dots, x_{k-1}) \in \mathbf{X}_{\neq}^{k-1}} f(y_1, \dots, x_{k-1}).$$

*Důkaz.* Podle Definice 12 máme pro  $U$ -statistiku:

$$\begin{aligned} D_{y_1}F(\mathbf{X}) &= \sum_{(x_1, \dots, x_k) \in (\mathbf{X} \cup \{y_1\})_{\neq}^k} f(x_1, \dots, x_k) - \sum_{(x_1, \dots, x_k) \in \mathbf{X}_{\neq}^k} f(x_1, \dots, x_k) = \\ &= \sum_{(x_1, \dots, x_{k-1}) \in \mathbf{X}_{\neq}^{k-1}} (f(y_1, x_1, \dots, x_{k-1}) + \dots + f(x_1, \dots, x_{k-1}, y_1)) = \\ &= k \sum_{(x_1, \dots, x_{k-1}) \in \mathbf{X}_{\neq}^{k-1}} f(y_1, x_1, \dots, x_{k-1}). \end{aligned}$$

□

Při výpočtu diferencí vyššího řádu opakujeme postup z důkazu Lemmatu 4 a po  $k$  krocích již difference nezávisí na bodovém procesu. To znamená, že od řádu  $k + 1$  jsou difference rovny nule. Tak dostáváme pro jádra:

**Lemma 5.** (Reitzner a Schulte, 2013, Lemma 3.5) Bud'  $F$   $U$ -statistika řádu  $k$  (1.5), potom

$$T_j F(y_1, \dots, y_j) = \begin{cases} \frac{k!}{(k-j)!} \int_{S^{k-j}} f(y_1, \dots, y_j, x_1, \dots, x_{k-j}) \lambda(d(x_1, \dots, x_{k-j})), & \text{pro } j \leq k, \\ 0, & \text{pro } j > k. \end{cases}$$

Důsledkem Lemmatu 5 je, že pro  $U$ -statistiky jsou řady ve Větě 3 konečné, což umožní výpočet (1.3). Ukažme si nyní tento výpočet střední hodnoty pro vybranou  $U$ -statistiku prvního a druhého řádu.

**Příklad 2.** (Výpočet střední hodnoty  $U$ -statistiky prvního řádu) Bud'  $X$  konečný bodový proces na  $S$  s hustotou  $p$  vzhledem k Poissonovu procesu  $\eta$  s intenzitou  $\lambda$ . Mějme  $C \subset S$  omezenou množinu, necht'  $f(x) = 1_{[x \in C]}$ . Uvažme  $U$ -statistiku 1.řádu

$$F(\mathbf{X}) = \sum_{x \in \mathbf{X}} 1_{[x \in C]} = \mathbf{X}(C).$$

Máme

$$D_y F(\eta) = F(\eta + \delta_y) - F(\eta) = \sum_{x \in \eta + \delta_y} 1_{[x \in C]} - \sum_{x \in \eta} 1_{[x \in C]} = 1_{[y \in C]}$$

a dále

$$D_{y_1, y_2}^2 F(\eta) = 0$$

a tedy

$$T_1 F(y) = \mathbb{E} 1_{[y \in C]} = 1_{[y \in C]}.$$

Nyní můžeme počítat střední hodnotu  $U$ -statistiky  $\mathbf{X}(C)$ , tedy střední hodnotu počtu bodů procesu. Víme, že

$$\mathbb{E}p(\eta) = \int_N p(\mathbf{x})d\mathbf{P}_\eta(\mathbf{x}) = \int_N d\mathbf{P}_\mathbf{X}(\mathbf{x}) = 1.$$

Dále také

$$\mathbb{E}F(\eta) = \int 1_{[y \in C]}d\lambda(y) = \lambda(C),$$

podle Campbellova vzorce (Věta 1). Vypočteme ještě jádro:

$$D_y p(y) = p(\mathbf{x} \cup y) - p(\mathbf{x})$$

a tedy

$$\begin{aligned} T_1^\eta p(y) &= \mathbb{E}D_y(p(\eta)) = \int D_y(p(\mathbf{x}))d\mathbf{P}_\eta(\mathbf{x}) = \int \frac{p(\mathbf{x} \cup y) - p(\mathbf{x})}{p(\mathbf{x})} p(\mathbf{x})d\mathbf{P}_\eta(\mathbf{x}) = \\ &= \int (\lambda^*(y; \mathbf{x}) - 1)d\mathbf{P}_\mathbf{X}(\mathbf{x}) = \mathbb{E}\lambda^*(y; \mathbf{X}) - 1. \end{aligned}$$

Nyní použijeme vzorec (1.4). Předpokládejme, že  $p \in L^2(\mathbf{P}_\eta)$  a ověříme, že  $F \in L^2(\mathbf{P}_\eta)$ . K ověření využijeme vzorec (1.2) s hustotou  $q \equiv 1$ . Dostaneme tedy:

$$\mathbb{E}g(\eta) = e^{-\lambda(S)} \sum_{n=0}^{\infty} \frac{1}{n!} \int_S \dots \int_S g(I_n(x_1, \dots, x_n)) \lambda(dx_1) \dots \lambda(dx_n) \quad (1.6)$$

Potom

$$\mathbb{E}F^2(\eta) \leq \text{const.} \sum_{n=0}^{\infty} \frac{1}{n!} n^2 \lambda(S)^n < +\infty.$$

Předpoklady jsou splněny, a tak dostáváme

$$\begin{aligned} \mathbb{E}F(\mathbf{X}) &= \mathbb{E}[F(\eta)p(\eta)] = \\ &= \mathbb{E}F(\eta)\mathbb{E}p(\eta) + \langle T_1 F, T_1 p \rangle_1 = \\ &= \mathbb{E}F(\eta) + \int_S T_1 F(y)T_1 p(y)d\lambda(y) = \\ &= \lambda(C) + \int_C (\mathbb{E}\lambda^*(y; \mathbf{X}) - 1)d\lambda(y) = \\ &= \int_C \mathbb{E}\lambda^*(y; \mathbf{X})d\lambda(y). \end{aligned} \quad (1.7)$$

**Příklad 3.** (Výpočet střední hodnoty  $U$ -statistiky druhého řádu) Mějme funkci

$$f(x, y) = 1_{[\|x-y\| \leq r]}, \quad x, y \in S, r > 0.$$

Potom máme  $U$ -statistiku druhého řádu

$$F(\mathbf{X}) = \sum_{(x, y) \in \mathbf{X}_\neq^2} 1_{[\|x-y\| \leq r]}.$$

Nejprve vypočteme diferenci prvního a druhého řádu funkcionálu  $F$

$$D_z F(\eta) = F(\eta + \delta_z) - F(\eta) = 2 \sum_{x \in \eta} 1_{[\|x-z\| \leq r]},$$

$$D_{z,w} F(\eta) = 2 \sum_{x \in \eta + \delta_w} 1_{[\|x-z\| \leq r]} - 2 \sum_{x \in \eta} 1_{[\|x-z\| \leq r]} = 2 \cdot 1_{[\|w-z\| \leq r]}.$$

Nyní můžeme vypočítat jádra procesu

$$T_1 F(z) = \mathbb{E} \left( 2 \sum_{x \in \eta} 1_{[\|x-z\| \leq r]} \right) = 2 \int_S 1_{[\|x-z\| \leq r]} d\lambda(x) = 2\lambda(\{x \in S; \|x-z\| \leq r\}),$$

$$T_2 F(z,w) = 2 \cdot 1_{[\|w-z\| \leq r]}.$$

Spočtíme ještě difference a jádra pro hustotu Poissonova procesu.

$$D_u p(\mathbf{x}) = p(\mathbf{x} + \delta_u) - p(\mathbf{x})$$

Ze vzorce pro podmíněnou intenzitu máme

$$p(\mathbf{x} + \delta_u) = p(\mathbf{x}) \lambda^*(u, \mathbf{x}),$$

dosazením do původního vzorce dostáváme

$$D_u p(\mathbf{x}) = p(\mathbf{x}) (\lambda^*(u, \mathbf{x}) - 1).$$

Diference druhého řádu:

$$\begin{aligned} D_{u,v}^2 p(\mathbf{x}) &= D_u(D_v p(\mathbf{x})) = D_u(p(\mathbf{x}) (\lambda^*(v, \mathbf{x}) - 1)) = \\ &= p(\mathbf{x} \cup u) (\lambda^*(v, \mathbf{x} \cup u) - 1) - p(\mathbf{x}) (\lambda^*(v, \mathbf{x}) - 1) = \\ &= p(\mathbf{x} \cup u) \left( \frac{p(\mathbf{x} \cup u \cup v)}{p(\mathbf{x} \cup u)} - 1 \right) - p(\mathbf{x}) \left( \frac{p(\mathbf{x} \cup v)}{p(\mathbf{x})} - 1 \right) = \\ &= p(\mathbf{x} \cup u \cup v) - p(\mathbf{x} \cup u) - p(\mathbf{x} \cup v) + p(\mathbf{x}) \end{aligned}$$

Vynásobíme-li obě strany rovnice  $\frac{1}{p(\mathbf{x})}$ , získáváme

$$D_{u,v}^2 p(\mathbf{x}) = (\lambda_2^*(\mathbf{x}; u, v) - \lambda^*(\mathbf{x}, u) - \lambda^*(\mathbf{x}, v) + 1) p(\mathbf{x}).$$

Z diferencí vypočteme jádra procesu

$$\begin{aligned} T_1^\eta p(u) &= \mathbb{E} \lambda^*(y; \mathbf{X}) - 1, \\ T_2^\eta p(u, v) &= \mathbb{E} D_{u,v}^2 p(\eta) = \int D_{u,v}^2(p(\mathbf{x})) d\mathbf{P}_\eta(\mathbf{x}) = \\ &= \int \frac{p(\mathbf{x} \cup \{u, v\}) - p(\mathbf{x} \cup u) - p(\mathbf{x} \cup v) + p(\mathbf{x})}{p(\mathbf{x})} p(\mathbf{x}) d\mathbf{P}_\eta(\mathbf{x}) = \\ &= \int_N (\lambda_2^*(\mathbf{x}; u, v) - \lambda^*(\mathbf{x}, u) - \lambda^*(\mathbf{x}, v) + 1) d\mathbf{P}_\mu(\mathbf{x}) = \\ &= \mathbb{E} (\lambda_2^*(\mathbf{X}; u, v) - \lambda^*(\mathbf{X}, u) - \lambda^*(\mathbf{X}, v)) + 1. \end{aligned}$$

Vypočteme ještě  $\mathbb{E}F(\eta)$  podle Campbellova vzorce (Věta 1) .

$$\begin{aligned}\mathbb{E}F(\eta) &= \iint 1_{[\|x-y\|\leq r]} \lambda(dx)\lambda(dy) = \\ &= \iint_{\{\|x-y\|\leq r\}} \lambda(dx)\lambda(dy) = \lambda^2(\{(x,y); \|x-y\|\leq r\})\end{aligned}$$

Před dosazením do vzorce pro střední hodnotu  $U$ -statistiky, ověříme předpoklady. Opět předpokládáme, že  $p \in L^2(\mathbf{P}_\eta)$  a ověříme, že  $F \in L^2(\mathbf{P}_\eta)$ . Použitím vzorce (1.6) dostáváme

$$\mathbb{E}F^2(\eta) \leq \text{const.} \sum_{n=0}^{\infty} \frac{1}{n!} n^4 \lambda(S)^n < +\infty.$$

Můžeme dále počítat:

$$\begin{aligned}\mathbb{E}F(\mathbf{X}) &= \mathbb{E}F(\eta) + \langle T_1 F, T_1 p \rangle_1 + \frac{1}{2} \langle T_2 F, T_2 p \rangle_2 = \\ &= \lambda^2(\{(x,y); \|x-y\|\leq r\}) + \\ &+ 2 \int \lambda(\{y \in S; \|y-z\|\leq r\}) (\mathbb{E}\lambda^*(\mathbf{X}, z) - 1) \lambda(dz) + \\ &+ \iint 1_{\|w-z\|\leq r} [\mathbb{E}(\lambda_2^*(\mathbf{X}; u, v) - \lambda^*(\mathbf{X}, u) - \lambda^*(\mathbf{X}, v)) + 1] \lambda(dw)\lambda(dz)\end{aligned}$$

Vzmemme-li nyní členy s  $\lambda^2$  získáváme:

$$\begin{aligned}\lambda^2(\{(x,y); \|x-y\|\leq r\}) - 2 \int \lambda(\{y \in S; \|x-z\|\leq r\}) \lambda(dz) + \\ + \iint 1_{\|w-z\|\leq r} \lambda(dw)\lambda(dz) = 0.\end{aligned}$$

Nyní vezmeme členy s  $\lambda^*$ :

$$\begin{aligned}2 \iint 1_{\|y-z\|\leq r} \mathbb{E}\lambda^*(z, \mathbf{X}) \lambda(dz)\lambda(dy) - \iint 1_{\|w-z\|\leq r} \mathbb{E}\lambda^*(w, \mathbf{X}) \lambda(dz)\lambda(dw) - \\ - \iint 1_{\|y-z\|\leq r} \mathbb{E}\lambda^*(z, \mathbf{X}) \lambda(dz)\lambda(dw) = 0.\end{aligned}$$

Tudíž z původního vzorce nám zbývá:

$$\mathbb{E}F(\mathbf{X}) = \iint 1_{\|w-z\|\leq r} \mathbb{E}(\lambda_2^*(w, z; \mathbf{X})) \lambda(dz)\lambda(dw). \quad (1.8)$$

Cílem dalšího výkladu je odvodit obecný vzorec pro střední hodnotu funkcionálu bodového procesu s hustotou, pro něž uvedené příklady budou speciální případy.

# Kapitola 2

## Teoretické výsledky pro bodové procesy s hustotou

### 2.1 Střední hodnota funkcionálu

Bud'  $\mathbf{X}$  konečný bodový proces s hustotou  $p$  vzhledem k Poissonovu procesu  $\eta$  s mírou intenzity  $\lambda$ . Hlavním tvrzením je Věta 10, která udává vztah pro střední hodnotu funkcionálu  $F(\mathbf{X})$ .

**Lemma 6.** (*Last a Penrose, 2011, Vztah 2.4*) Bud'  $\mathbf{X}$  konečný bodový proces,  $F : \mathbb{N}^f \rightarrow \mathbb{R}$  měřitelná funkce. Potom

$$D_{y_1, \dots, y_n}^n F(\mathbf{X}) = \sum_{J \subset \{1, 2, \dots, n\}} (-1)^{n-|J|} F(\mathbf{X} + \sum_{j \in J} \delta_{y_j}),$$

kde  $|J|$  označuje počet prvků  $J$ .

**Lemma 7.** *Platí*

$$T_n p(y_1, \dots, y_n) = \sum_{J \subset \{1, \dots, n\}} (-1)^{n-|J|} \mathbb{E} \lambda_{|J|}^* (\{y_j, j \in J\}; \mathbf{X}).$$

*Důkaz.* Z Lemmatu 6 máme

$$T_n p(y_1, \dots, y_n) = \mathbb{E} D_{y_1, \dots, y_n}^n p(\eta) = \int \sum_{J \subset \{1, \dots, n\}} (-1)^{n-|J|} p(\mathbf{x} \cup \{y_j, j \in J\}) \frac{d\mathbf{P}_\eta(\mathbf{x})}{p(\mathbf{x})}.$$

□

Budeme potřebovat některé kombinatorické identity.

**Lemma 8.** (*Kaucký, 1975*) Pro přirozené číslo  $s$  platí

$$\sum_{r=0}^s (-1)^r \binom{s}{r} = 0. \quad (2.1)$$



*Důkaz.* Dokážeme indukcí. Pro  $s = 1$  je to identita  $1 - 1 = 0$ . Nechť tedy vzorec platí pro nějaké  $s \geq 1$ . Potom

$$\sum_{r=0}^{s+1} (-1)^r \binom{s+1}{r} = \sum_{r=1}^{s+1} (-1)^r \binom{s}{r-1} + \sum_{r=0}^s (-1)^r \binom{s}{r} = - \sum_{r=0}^s (-1)^r \binom{s}{r} = 0,$$

takže vzorec platí obecně. □

**Lemma 9.** (*Kaucký, 1975*) Pro všechna nezáporná celá čísla  $j$  a  $k > j$  platí:

$$\sum_{n=j}^k (-1)^n (n)_j \binom{k}{n} = 0,$$

kde

$$(n)_j = \begin{cases} n(n-1)(n-2) \cdots (n-j+1), & \text{pro } k \text{ přirozené,} \\ 1, & \text{pro } k = 0. \end{cases}$$

*Důkaz.*

$$\begin{aligned} \sum_{n=j}^k (-1)^n (n)_j \binom{k}{n} &= \\ &= \sum_{n=j}^k (-1)^n \frac{n!}{(n-j)!} \binom{k}{n} = \\ &= j! \sum_{n=j}^k (-1)^n \binom{n}{j} \binom{k}{n} = \\ &= (k)_j \sum_{n=j}^k (-1)^n \binom{k-j}{n-j} = \\ &= (k)_j (-1)^j \sum_{r=0}^{k-j} (-1)^r \binom{k-j}{r} = \\ &= 0, \end{aligned}$$

v posledním řádku se využívá Lemma 8. □

**Věta 10.** (*Beneš a Zikmundová, 2014, Věta 2*) Nechť  $F \in L^2(\mathbf{P}_\eta)$  je  $U$ -statistika řádu  $k$ , viz (1.5) a hustota  $p$  bodového procesu  $\mathbf{X}$  splňuje také  $p \in L^2(\mathbf{P}_\eta)$ , potom

$$\mathbb{E}F(\mathbf{X}) = \int_{B^k} f(x_1, \dots, x_k) \mathbb{E}[\lambda_k^*(x_1, \dots, x_k; \mathbf{X})] \lambda(d(x_1, \dots, x_k)). \quad (2.2)$$

*Důkaz.* Mějme  $\lambda_0^* \equiv 1$ ,  $T_0F = \mathbb{E}F(\eta)$ , třída  $C_j^n$  kombinací  $c = \{c_1, \dots, c_j\}$  z  $\{1, \dots, n\}$ . Dosadíme do vzorce (1.4) z Lemmat 6 a 7 a s použitím vztahu  $\mathbb{E}p(\eta) = 1$  získáváme

$$\begin{aligned}
\mathbb{E}F(\mathbf{X}) &= \mathbb{E}F(\eta) + \sum_{n=1}^k \frac{1}{n!} \langle T_n F, T_n p \rangle_n = \\
&= \sum_{n=0}^k \frac{1}{n!} \int_{B^n} \frac{k!}{(k-n)!} \int_{B^{k-n}} f(y_1, \dots, y_n, x_1, \dots, x_{k-n}) \lambda(d(x_1, \dots, x_{k-n})) \\
&\quad \cdot \sum_{j=0}^n (-1)^{n-j} \sum_{c \in C_j^n} \mathbb{E} \lambda_j^*(y_{c_1}, \dots, y_{c_j}; \mathbf{X}) \lambda(d(y_1, \dots, y_n)) = \\
&= \sum_{j=0}^k \sum_{n=j}^k (-1)^{n-j} \binom{k}{n} \int_{B^k} \sum_{c \in C_j^n} \mathbb{E} \lambda_j^*(y_{c_1}, \dots, y_{c_j}; \mathbf{X}) f(y_1, \dots, y_k) \lambda(d(y_1, \dots, y_k)) = \\
&= \sum_{j=0}^k \sum_{n=j}^k (-1)^{n-j} \binom{k}{n} \binom{n}{j} \int_{B^k} \mathbb{E} \lambda_j^*(y_1, \dots, y_j; \mathbf{X}) f(y_1, \dots, y_k) \lambda(d(y_1, \dots, y_k))
\end{aligned}$$

Nyní ukážeme, že pro  $0 \leq j \leq k-1$  je součet přes  $n$  roven nule a pro  $j = k$  dává výsledek. Z Lemmatu 9 víme, že pro  $k > j$  je

$$j! \sum_{n=j}^k (-1)^n \binom{n}{j} \binom{k}{n} = 0.$$

Pro  $j = k$  je tato suma rovna 1 a tedy

$$\mathbb{E}F(\mathbf{X}) = \int_{B^k} f(x_1, \dots, x_k) \mathbb{E}[\lambda_k^*(x_1, \dots, x_k; \mathbf{X})] \lambda(d(x_1, \dots, x_k)).$$

□

## 2.2 Vztahy pro Straussův proces

Ověříme nejprve předpoklady Věty 10 v případě Straussova procesu. Fakt, že studované  $U$ -statistiky patří do prostoru  $L^2(\mathbf{P}_\eta)$ , jsme ověřili už v příkladech 2 a 3. Zbývá tedy ověřit, že do tohoto prostoru patří i hustota Straussova procesu. Pro hustotu Straussova procesu máme

$$p^2(\mathbf{x}) = \alpha^2 \beta^{2n(\mathbf{x})} \gamma^{2s(\mathbf{x})},$$

kde

$$s(\mathbf{x}) = \sum_{x, y \in \mathbf{x}} 1_{[\|x-y\| \leq r]},$$

pro  $\mathbf{x} = (x_1, \dots, x_n)$  je  $n(\mathbf{x}) = n$ ,  $s(x_1, \dots, x_n) \leq \binom{n}{2}$ . Protože  $0 < \gamma \leq 1$ , tak s použitím (1.6) máme

$$\mathbb{E}p^2(\eta) \leq \text{const.} \sum_{n=0}^{\infty} \frac{1}{n!} \beta^{2n} \lambda(S)^n < +\infty.$$

Do vztahů ze sekce 2.1 nyní můžeme dosadit podmíněnou intenzitu Straussova procesu. Připomeňme:

$$\lambda^*(u, \mathbf{x}) = \beta \gamma^{t(u, \mathbf{x})},$$

kde

$$t(u, \mathbf{x}) = \sum_{y \in \mathbf{x}} 1_{[\|u-y\| \leq r]}$$

a tedy, označíme-li Straussův proces písmenem  $\mu$ , pak po dosazení do vzorce (1.7), resp. (2.2) dostáváme vztah pro  $U$ -statistiku prvního řádu z příkladu 2:

$$\mathbb{E}\mu(C) = \beta \int_C \mathbb{E}\gamma^{t(u, \mu)} \lambda(du) = \beta \int_C \mathbb{E}\left[\prod_{y \in \mu} \gamma^{1_{[\|u-y\| \leq r]}}\right] \lambda(du). \quad (2.3)$$

Uvažme  $U$ -statistiku druhého řádu z příkladu 3 :

$$F(\mu) = \sum_{(x,y) \in \mu^2_{\neq}} 1_{[\|x-y\| \leq Q]}, \quad Q > 0.$$

Podmíněná intenzita druhého řádu Straussova procesu vypadá takto:

$$\lambda_2^*(\mathbf{x}; u, v) = \frac{p(\mathbf{x} \cup \{u, v\})}{p(\mathbf{x})} = \frac{\beta^{n(\mathbf{x} \cup \{u, v\})} \gamma^{s(\mathbf{x} \cup \{u, v\})}}{\beta^{n(\mathbf{x})} \gamma^{s(\mathbf{x})}} = \beta^2 \gamma^A,$$

kde

$$\begin{aligned} A &= \sum_{y, z \in \mathbf{x} \cup \{u, v\}} 1_{[\|y-z\| \leq r]} - \sum_{y, z \in \mathbf{x}_{\neq}} 1_{[\|y-z\| \leq r]} = \\ &= 1_{[\|u-v\| \leq r]} + \sum_{y \in \mathbf{x}} 1_{[\|y-u\| \leq r]} + \sum_{y \in \mathbf{x}} 1_{[\|y-v\| \leq r]} = \\ &= 1_{[\|u-v\| \leq r]} + t(u, \mathbf{x}) + t(v, \mathbf{x}). \end{aligned}$$

Nyní můžeme dosadit do vzorce (1.8) a dostaneme

$$\mathbb{E}F(\mu) = \iint_{\{\|w-z\| \leq Q\}} \beta^2 \gamma^{1_{[\|w-z\| \leq r]}} \mathbb{E}[\gamma^{t(w, \mu) + t(z, \mu)}] \lambda(dw) \lambda(dz). \quad (2.4)$$

# Kapitola 3

## Numerické výsledky

Cílem této části je ověřit identitu z věty 10 na základě simulační studie, kde střední hodnota je odhadována průměry přes realizace bodového procesu.

### 3.1 Simulace Straussova procesu

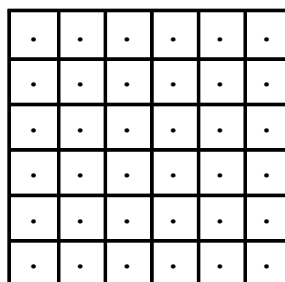
Pro získání dat k výpočtům bylo nejprve nutné simulovat bodový proces. Použili jsme simulaci Straussova procesu, který je implementován v softwaru R. Jednotlivé simulace jsme získali pomocí funkce  $rStrauss(beta, gamma, R, W)$  se čtyřmi parametry:

- beta -  $\beta > 0$ ,
- gamma - parametr interakce,  $0 \leq \gamma \leq 1$ ,
- $R$  - rozsah interakce,  $R > 0$ ,
- $W$  - okno, ve kterém se proces vygeneruje,  $S = W$ .

Ze získaných simulací jsme uložili souřadnice bodů jednotlivých realizací procesu, které jsme dále zpracovávali.

### 3.2 Aproximace vzorce pro střední počet bodů

Jako simulace procesu nám poslouží vygenerování  $n$  realizací procesu v jednotkovém čtverci  $S$ . Pro aproximaci integrálu využijeme mřížku vytvořenou v tomto čtverci, která vznikne rozdělením na  $k \times k$  stejných čtverců. Ve výpočtech byla použita mřížka s dělením  $6 \times 6$ , viz obr.3.1.



Obrázek 3.1: Mřížka jednotkového okna pro numerický výpočet s vyznačením středů buněk,  $k = 6$ .

Budeme nyní pracovat se vzorcem pro střední hodnotu  $U$ -statistiky prvního řádu, tedy vzorcem (2.3)

$$\mathbb{E}\mu(C) = \beta \int_C \mathbb{E} \left[ \prod_{y \in \mu} \gamma^{1_{[\|u-y\| \leq R]}} \right] \lambda(du).$$

Zvolme  $C = S$ . Integrand můžeme přepsat jako

$$\mathbb{E} \left[ \prod_{y \in \mu} \gamma^{1_{[\|u-y\| \leq R]}} \right] = \mathbb{E} [\gamma^{\#\{y \in \mu; 1_{[\|u-y\| \leq R]}\}}]. \quad (3.1)$$

Označme exponent u parametru  $\gamma$  na pravé straně (3.1) pro pevný bod mřížky  $y_{i,j}$  o souřadnicích  $i,j$  jako  $M_{y_{i,j}}$ . V  $l$ -té realizaci  $\mu_l$  Straussova procesu je hodnota  $M_{y_{i,j}}$  rovna číslu  $m_{l,y_{i,j}}$ . Střední hodnotu budeme odhadovat aritmetickým průměrem přes jednotlivé realizace.

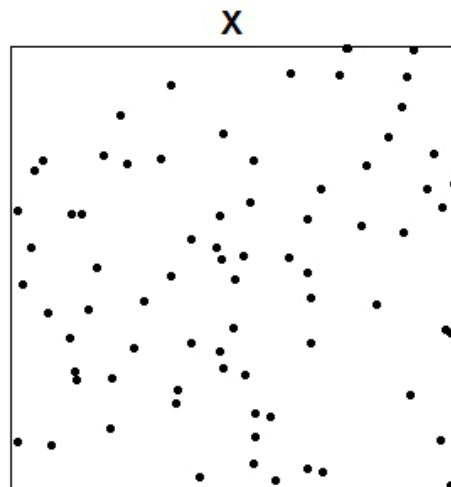
$$\mathbb{E} \gamma^{M_{y_{i,j}}} \approx \frac{1}{n} \sum_{l=1}^n \gamma^{m_{l,y_{i,j}}} := \widehat{\mathbb{E} \gamma^{M_{y_{i,j}}}}$$

Pro numerický výpočet bylo nutné ještě aproximovat integrál. Ten jsme nahradili součtem přes body mřížky vynásobené plochou, která jim v mřížce náleží. Vzorec pro odhad středního počtu bodů tedy vypadá:

$$\widehat{\mathbb{E} \mu(C)} = \beta \frac{1}{k^2} \sum_{i=1}^k \sum_{j=1}^k \widehat{\mathbb{E} \gamma^{M_{y_{i,j}}}}. \quad (3.2)$$

### 3.3 Výsledky numerických výpočtů

Podle výše uvedeného vzorce (3.2) byly provedeny výpočty v softwaru Mathematica (zdrojový kód přiložen na CD). Jako simulace procesu nám posloužilo 50 realizací Straussova procesu s parametry  $\beta = 100, \gamma = 0,5; R = 0,05$  na jednotkovém čtverci. Obrázek jedné z realizací:



Obrázek 3.2: Realizace  $\mathbf{X}$  Straussova procesu s parametry  $\beta = 100, \gamma = 0,5, R = 0,05$  v okně  $W = [0,1]^2$ .

Porovnání numerických výpočtů střední hodnoty Straussova procesu pro  $k = 6$  s průměrným počtem bodů jednotlivých 50 realizací je vidět v tabulce 3.1, kde byla uvedená simulace opakována desetkrát. Při výpočtu přes 50 realizací byla průměrná absolutní odchylka odhadu střední hodnoty od průměrného počtu bodů realizací 1.28904.

Číslo pokusu	num. výpočet	prům. počet bodů	abs. odchylka
1	73,1944	74,64	1,44556
2	73,6701	74,3	0,629861
3	74,1285	73,72	0,408472
4	73,6563	75,22	1,56375
5	74,5243	73,62	0,904306
6	73,7431	73,98	0,236944
7	75,4583	72,52	2,93833
8	75,1736	73,7	1,47361
9	75,1667	73,86	1,30667
10	73,1771	75,16	1,98292
<b>průměr</b>	74,1892	74,072	1,28904

Tabulka 3.1: Porovnání numerických výpočtů pro 50 realizací.

Při výpočtu přes 100 realizací byla průměrná absolutní odchylka odhadu střední hodnoty od průměrného počtu bodů realizací 0.816285. Další údaje jsou uvedeny v následující tabulce:

Číslo pokusu	num. výpočet	prům. počet bodů	abs. odchylka
1	74,092	74,62	0,527986
2	73,6962	74,49	0,793819
3	74,9583	74,66	0,298333
4	74,5382	75,21	0,671806
5	73,5677	74,98	1,41229
6	73,474	74,39	0,916042
7	74,3906	75,79	1,39938
8	74,0521	73,96	0,0920833
9	73,2969	74,82	1,52312
10	74,092	74,62	0,527986
<b>průměr</b>	74,0158	74,754	0,816285

Tabulka 3.2: Porovnání numerických výpočtů pro 100 realizací.

### 3.4 Aproximace vzorce pro vybranou $U$ -statistiku druhého řádu

Mějme následující  $U$ -statistiku druhého řádu:

$$F(\mu) = \sum_{(x,y) \in \mu_{\neq}^2} 1_{[\|x-y\| \leq Q]}, \quad Q > 0.$$

Připomeňme si vzorec (2.4) pro střední hodnotu této  $U$ -statistiky:

$$\mathbb{E}F(\mu) = \iint_{\{\|w-z\| \leq Q\}} \beta^2 \gamma^{1_{[\|w-z\| \leq R]}} \mathbb{E}[\gamma^{t(w,\mu)+t(z,\mu)}] \lambda(dw) \lambda(dz).$$

Budeme postupovat analogicky jako při aproximaci vzorce pro střední hodnotu  $U$ -statistiky prvního řádu. Uvažujme nyní dvojice bodů  $z$  a  $w$  z  $S$ . Pomocná mřížka pro aproximaci integrálu má opět  $k \times k$  buněk,  $z$  a  $w$  jsou body této mřížky. Za simulaci Straussova procesu považujeme jeho  $n$  realizací v programu R. Volíme  $0 < R < Q$ . Nejprve upravíme exponenty parametru  $\gamma$ :

$$t(w,\mu) = \sum_{y \in \mu_l} 1_{[\|y-w\| \leq R]} = \#\{y \in \mu_l; 1_{[\|y-w\| \leq R]}\} := m_l$$

$$t(z,\mu) := n_l \quad \text{pro } l\text{-tou realizaci.}$$

Střední hodnotu odhadujeme opět průměrem přes realizace, a tak dostáváme

$$\mathbb{E}[\gamma^{t(w,\mu)} \gamma^{t(z,\mu)}] \approx \frac{1}{n} \sum_{l=1}^n \gamma^{m_l} \gamma^{n_l}.$$

Nakonec nahradíme ještě integrál součtem přes jednotlivé body mřížky a jim odpovídající plochy. Získáváme tak

$$\mathbb{E}F(\mu) = \frac{\beta^2}{nk^4} \sum_{j=1}^k \sum_{i=1}^k \sum_{o=1}^k \sum_{p=1}^k 1_{[\|z_{ij}-w_{op}\| \leq Q]} \gamma^{1_{[\|z_{ij}-w_{op}\| \leq R]}} \sum_{l=1}^n \gamma^{m_l} \gamma^{n_l}, \quad (3.3)$$

kde výraz v sumě započítáme, pokud je splněna podmínka  $(i \neq o) \vee (j \neq p)$ , tedy, jedná-li se o dva různé body mřížky.

### 3.5 Výsledky numerických výpočtů II

Podle vzorce (3.3) byl numericky vypočten střední počet dvojic bodů vzdálených od sebe o méně než  $Q$ . Pro výpočet jsme použili data ze 100 realizací Straussova procesu s parametry  $\beta = 100, \gamma = 0,5; R = 0,05$  na jednotkovém čtverci. Pomocnou mřížku jsme zjemnili na  $12 \times 12$  buněk. Zkoumanou vzdálenost  $Q$  jsme položili rovnu 0,25. Tabulka 3.3 ukazuje výsledky získané výpočtem podle vzorce (3.3) v porovnání s průměrným počtem dvojic bodů se zkoumanou vlastností vypočteným přímo z bodů realizací (v tabulce označeno jako *průměr z dat*).

Číslo pokusu	num. výpočet	průměr z dat	abs. odchylka
1	842,073	872,68	30,606900
2	846,07	845,82	0,249637
3	844,59	850,32	5,729820
4	845,948	864,52	18,572500
5	841,141	862,8	21,659000
6	848,282	839,32	8,962350
7	836,944	876,86	39,916200
8	853,018	829,04	23,977600
9	848,328	848,12	0,207712
10	856,459	846,36	10,099400
<b>průměr</b>	846,285	853,584	15,9981

Tabulka 3.3: Porovnání numerických výpočtů II.

Průměrná odchylka je 17.9566 dvojic bodů. Zlepšení výsledku by bylo možné zjemněním mřížky, která slouží k aproximaci integrálu. Pro ještě jemnější mřížku ale rychle roste doba trvání výpočtu.



# Literatura

BADDELEY, A. (2007). Spatial point processes and their applications. In *Stochastic Geometry*, pages 1–75. Springer Berlin Heidelberg. ISBN 978-3-540-38174-7.

BENEŠ, V. a ZIKMUNDOVÁ, M. (2014). Functionals of spatial point processes having a density with respect to the poisson process. *arXiv:1401.2783v3 [math.PR]* 14 Feb 2014.

KAUCKÝ, J. (1975). *Kombinatorické identity*. První vydání. Vydavatelstvo Slovenskej akadémie vied, Bratislava.

LAST, G. a PENROSE, M. (2011). Poisson process Fock space representation, chaos expansion and covariance inequalities. *Probability Theory and Related Fields*, **150**(3-4), 663–690.

REITZNER, M. a SCHULTE, M. (2013). Central limit theorems for U-statistics of poisson point processes. *The Annals of Probability*, **41**(6), 3879–3909.

# Seznam obrázků

3.1	Mřížka jednotkového okna pro numerický výpočet s vyznačením středů buněk, $k = 6$ . . . . .	16
3.2	Realizace <b>X</b> Straussova procesu s parametry $\beta = 100, \gamma = 0.5, R = 0.05$ v okně $W = [0,1]^2$ . . . . .	17

# Seznam tabulek

3.1	Porovnání numerických výpočtů pro 50 realizací. . . . .	18
3.2	Porovnání numerických výpočtů pro 100 realizací. . . . .	18
3.3	Porovnání numerických výpočtů II. . . . .	20