Вероятности 2

ДИМИТЪР Л. ВЪНДЕВ

Записки, 1999

Съдържание

Увод

Тези записки са основани на материал, набран (и прочетен от Д.Въндев) в различни курсове. Лекциите са предназначени за студенти специализиращи "Вероятности и статистика" в Факултет по Математика и Информатика. Курсът е непълен и предстои да бъде допълнен за да стане стандартен в програмата на специалността.

Включени са доказателства на основните теореми на Каратеодори, Бохнер, Леви. Част от текста е написана от Стилян Стоев.

Д.Въндев

Глава 1

Теорема за продължението

Основната цел при изграждане на вероятностно пространство е минимизирането на множеството събития, върху което трябва да проверяваме адитивността и непрекъснатостта на вероятността. В тази лекция думата мярка се използва само като синоним на вероятност.

Тук ще докажем знаменитата теорема за продължението в два варианта.

1.1 Въведение и примери

Определение 1.1. *Наричаме едно множество $\mathcal{A}$ от подмножества на множеството X булова полу-алгебра, ако удовлетворява следните условия:*

1. $X \in \mathcal{A}$;
2. *Ако* $A, B \in \mathcal{A}$, *то* $A \cap B \in \mathcal{A}$;
3. $\forall A \in \mathcal{A}\ \exists n < \infty$ *и* $B_i \in \mathcal{A}$ *такива, че* $X \setminus A = \sum_{i=1}^{n} B_i$.

Теорема 1.1. *[Теорема на Каратеодори] Ако вероятността* P*, зададена върху буловата полу-алгебра* $\mathcal{A}$*, е* σ*-адитивна на* $\mathcal{A}$*, то тя е продължима еднозначно върху* $\sigma(\mathcal{A})$.

Пример 1.1. *Множеството от подинтервали* $\mathcal{A} = \{[a, b)\, 0 \leq a \leq b < 1\}$ *на единичния интервал* $[0, 1)$ *образува полуалгебра.*

Теорема 1.2. *Мярката на Лебег, породена от дължината им е определена еднозначно на Бореловите множества.*

Доказателство: Достатъчно е да покажем, че при разделянето на интервал на подинтервали имаме адитивност на дължината. Така можем да използваме теоремата на Каратеодори. □

Пример 1.2. *Множеството от подинтервали* $\mathcal{A} = \{[a,b) \ -\infty \leq a \leq b < \infty\}$ *на* R^1 *образува полуалгебра. Нека* $F(x)$ *е функция на разпределение. Да определим вероятността* $\mathsf{P}([a,b)) = F(b) - F(a)$.

Покажете, че $(R^1, \ \mathcal{B} \ , \mathsf{P})$ е вероятностно пространство ($\mathcal{B}$ - Борелова σ-алгебра).

Пример 1.3. *Нека разгледаме единичния квадрат* $[0,1) \times [0,1)$. *Множеството* $\mathcal{A}$ *от правоъгълници* $[a,b) \times [c,d)$ *в него образува полуалгебра* $(a < b, c < d)$. *Мярката* μ *задаваме* $\mu(A) = (b-a)(d-c)$.

Покажете, че μ е σ-адитивна на $\mathcal{A}$. Значи $(R^2, \ \mathcal{B} \ , \mu)$ е вероятностно пространство ($\mathcal{B}$ - Борелова σ-алгебра).

1.2 Спомагателни резултати

Тук за упражнение ще докажем една лема, която има самостоятелно значение, а и вече доказахме в курса по теория на вероятностите.

Лема 1.1. *Нека вероятността* P *е зададена върху буловата алгебра* $\mathfrak{F}$ *и е адитивна. Тогава, необходимо и достатъчно условие да е непрекъсната в* $\emptyset$, *е тя да е* σ*-адитивна върху* $\mathfrak{F}$.

Доказателство: *Необходимост.* Нека P е непрекъсната. Нека $A = \sum_{n=1}^{\infty} A_n$ и $A, A_n \in \mathfrak{F}$. Първо да отбележим, че за $\forall n$ множеството $B_n = \sum_{i=n+1}^{\infty} A_i \in \mathfrak{F}$ защото. От адитивността на P за $\forall n$ следва равенството

$$\mathsf{P}(A) = \mathsf{P}(A_1) + \mathsf{P}(A_2) + \cdots + \mathsf{P}(A_n) + \mathsf{P}(\sum_{i=n+1}^{\infty} A_i). \tag{1.1}$$

От непрекъснатостта следва, че $\mathsf{P}(\sum_{i=n+1}^{\infty} A_i) \to 0$. Следователно, P е σ-адитивна върху $\mathfrak{F}$.

Достатъчност. Нека P е σ-адитивна. Нека $B_{n+1} \subset B_n \in \mathfrak{F}$ и $\bigcap B_n = \emptyset$.Да положим $A_n = B_n \setminus B_{n+1}$. Тъй като $\forall n, B_n = \sum_{i=n}^{\infty} A_i$ и P е σ-адитивна върху $\mathfrak{F}$, получаваме, че

$$\mathsf{P}(B_n) = \sum_{i=n}^{\infty} \mathsf{P}(A_i) \longrightarrow 0$$

като остатък на сходящ се ред. Значи P е непрекъсната. $\square$

Тази лема не е верна за полуалгебра. Защо?

Пример 1.4. *Нека разгледаме единичния квадрат* $[0,1) \times [0,1)$ *с отстранен диагонал* $I = \{(x,y) : \quad x = y\}$. *Множеството* $\mathcal{A}$ *от правоъгълници* $[a,b) \times [c,d)$ *в него образува полуалгебра. Мярката* μ *на* $\mathcal{A}$ *задаваме като дължина на частта от диагонала отрязан от правоъгълника.*

Покажете, че

1. μ е непрекъсната в $\varnothing$ на $\mathcal{A}$, но не е σ-адитивна;

2. μ не е непрекъсната в $\varnothing$ на породената от $\mathcal{A}$ е Борелова алгебра;

3. $\sigma(\mathcal{A})$ е бореловата алгебра $\mathcal{B}$ на $X = ([0,1) \times [0,1)) \setminus I$;

4. μ не е σ- адитивна на $\mathcal{B}$.

1.3 Конструктивно доказателство

Това доказателство е конструктивно и не се нуждае от единственост - конструираната вероятност е единствена по-построение. Предхождат го някои леми.

Лема 1.2. *Всяка вероятност определена на полуалгебра $\mathcal{A}$ се продължава еднозначно на породената от нея булова алгебра $\mathcal{B}$. Ако вероятността е σ-адитивна на $\mathcal{A}$, то тя е σ-адитивна на $\mathcal{B}$.*

Доказателство: 1. Нека $\mathcal{B}$ е породената от $\mathcal{A}$ булова алгебра. Съгласно условие 3. допълнението в X на всяко $A \in \mathcal{A}$ е обединение на краен брой елементи на $\mathcal{A}$ и е $\in \mathcal{B}$. Това значи, че всяко множество $B \in \mathcal{B}$ е крайна сума: $B = \sum_{i=1}^{n} A_i$ на непресичащи се елементи от $\mathcal{A}$.

Нека е зададена адитивна вероятност на $\mathcal{A}$. Значи определянето на вероятността върху $\mathcal{B}$ по формулите $\mathsf{P}(X \setminus B) = 1 - \mathsf{P}(B)$ и $\mathsf{P}(B) = \sum_{i=1}^{n} \mathsf{P}(A_i)$ е коректно и получената вероятност остава адитивна.

2. Нека вероятността на $\mathcal{A}$ е σ-адитивна .

$$\begin{array}{rcl} B_0 = & \sum_{i=1}^{\infty} B_i, & \\ B_i = & \sum_{k=1}^{n_i} A_i^k, & i = 0, 1, \dots \\ A_0^k = & \sum_{i=1}^{\infty} A_0^k \cap A_i^k, & k = 0, 1, \dots, n_0 \\ \mathsf{P}(A_0^k) = & \sum_{i=1}^{\infty} \mathsf{P}(A_0^k \cap A_i^k), & k = 0, 1, \dots, n_0 \\ \mathsf{P}(B_0) = & \sum_{k=1}^{n_0} \sum_{i=1}^{\infty} \mathsf{P}(A_0^k \cap A_i^k). & \square \end{array}$$

Следната конструкция е изключително удобна за теория на мярката.

Определение 1.2. *Наричаме едно множество $\mathcal{M}$ от подмножества на множеството X монотонен клас, ако*
1. заедно с всяка намаляваща редица $B_n \supset B_{n+1}$ от елементи на $\mathcal{M}$ съдържа и тяхното сечение $B = \cap B_n$.

2. заедно с всяка растяща редица $B_n \subset B_{n+1}$ от елементи на $\mathcal{M}$ съдържа и тяхното обединение $B = \cup B_n$.

Лема 1.3. *Монотонен клас породен от булова алгебра $\mathcal{B}$ е σ-алгебра ($\mathcal{M}$ ($\mathcal{B}$) $=$ $\sigma($ $\mathcal{B}$ $))$.*

Доказателство: Ясно е, че $\mathcal{M}(\mathcal{B}) \subset \sigma(\mathcal{B})$ защото $\sigma(\mathcal{B})$ съдържа всички множества от този вид. Нека покажем, че $\mathcal{M}(\mathcal{B})$ е σ-алгебра. Наистина,
1. $\varnothing, X, \overline{B} \in \mathcal{B}$.
2. $\cap B_n \in \mathcal{M}(\mathcal{B})$ защото $A_n = \cap_{i=1}^n B_i \in \mathcal{B}$ и $\cap A_n = \cap B_n$.
Значи $\mathcal{M}(\mathcal{B})$ е σ-алгебра и $\mathcal{M}(\mathcal{B}) \supset \sigma(\mathcal{B})$. $\square$

Лема 1.4. *Ако вероятността на буловата алгебра $\mathcal{B}$ е непрекъсната в $\varnothing$, то тя се продължава еднозначно на монотонния клас $\mathcal{M}(\mathcal{B})$ породен от $\mathcal{B}$.*

Доказателство: Нека $B = \cap B_n, B_n \supset B_{n+1}$. Ще определим $\mathsf{P}(B) = \lim \mathsf{P}(B_n)$. Границата винаги съществува, защото редицата от вероятности е монотонна. Нека $B = \cap A_n, A_n \supset A_{n+1}$ е друга такава редица. Нека означим с $C_n = A_n \cap B_n, D_n = A_n \cup B_n$. Имаме $B = \cap C_n = \cap D_n$ и неравенството: $\mathsf{P}(C_n) \leq \mathsf{P}(A_n), \mathsf{P}(B_n) \leq \mathsf{P}(D_n)$.Ако допуснем, че двете редици $\mathsf{P}(A_n), \mathsf{P}(B_n)$ са сходящи към различни граници, значи $\exists \varepsilon > 0$, такова че $|\mathsf{P}(C_n) - \mathsf{P}(D_n)| > \varepsilon$.Да фиксираме n така,че $|\mathsf{P}(C_n) - \lim \mathsf{P}(C_n)| < \varepsilon/2$. Да разгледаме сега редицата $E_m = \overline{C_n} \cap D_m$. Имаме $E_m \in \mathcal{B}$, $E_m \supset E_{m+1}$ и $\cap E_m = \overline{C_n} \cap B = \varnothing$. Значи съгласно предположението $\mathsf{P}(E_m) \to 0$. Получаваме противоречивото неравенство:

$$\varepsilon \leq \mathsf{P}(D_m) - \mathsf{P}(C_m) = \mathsf{P}(D_m) - \mathsf{P}(C_n) + \mathsf{P}(C_n) - \mathsf{P}(C_m) \leq \mathsf{P}(E_m) + \varepsilon/2.$$

Аналогично доказваме твърдението за множества от вида $B = \cup B_n, B_n \subset B_{n+1}$. $\square$

Остана да проверим адитивността, което е тривиално, и да сумираме получените резултати в твърдението на теоремата на Каратеодори.

1.4 Пълно доказателство

Второто доказателство дава възможност за "по-просторно" продължение - получават се повече измерими множества. То е и по-популярно в учебниците по теория на вероятностите и мярката (Обретенов 1974),(Партасарати 1983),(Обретенов и Хаджиев 1983).

Тук ще дадем независимо от горното доказателство на отслабения вариант на Теорема 1.1.

Теорема 1.3. *[Теорема на Каратеодори] Ако една адитивна вероятност* P*, зададена върху буловата алгебра $\mathcal{A}$, е непрекъсната в $\varnothing$, то тя е продължима еднозначно върху $\sigma(\mathcal{A})$.*

Първо ще въведем основно в теория на мярката и интеграла понятие.

Определение 1.3. *Нека определим върху всички подмножества на X функцията горна мярка:*

$$\mu(A) = \inf\{\mathsf{P}(B) : B \in \mathfrak{F}, A \subset B\}.$$

Горната мярка μ притежава следните важни и почти очевидни свойства:

1. монотонност - $\mu(B) \geq \mu(A), \forall B \supset A$;
2. изчислимост - $\forall \varepsilon > 0, \exists B \in \mathfrak{F} : A \subset B, \mu(B) - \mu(A) < \varepsilon$;
3. полуадитивност (крайна или изброима)- $\mu(\bigcup_i A_i) \leq \sum_i \mu(A_i)$;
4. на $\mathfrak{F}$ горната мярка μ съвпада с P и (следователно) е σ-адитивна (и непрекъсната в $\varnothing$) на $\mathfrak{F}$. Да отбележим, че това свойство не е длъжно да бъде изпълнено извън $\mathfrak{F}$.

Определение 1.4. *Да означим с $\mathfrak{R}$ класа от подмножества $B \subset X$, за които е изпълнено равенството*

$$1 = \mu(B) + \mu(\overline{B}).$$

Ясно е, че $\mathfrak{R} \supset \mathfrak{F}$ и $\mu(B) = \mathsf{P}(B)$, когато $B \in \mathfrak{F}$.

Лема 1.5. *За да принадлежи множеството B на $\mathfrak{R}$ е необходимо и достатъчно да бъде "апроксимируемо", т.е. да съществуват редиците $\{B_n^+\}, \{B_n^-\} \in \mathfrak{F}$ такива, че:*

$$B_n^- \subset B_{n+1}^- \subset B \subset B_{n+1}^+ \subset B_n^+, \qquad \mu(B_n^+) - \mu(B_n^-) = \mathsf{P}(B_n^+ \setminus B_n^-) \to 0.$$

Доказателство: Необходимост. За всяко B по определение (1.3) съществува редица $B_n \in \mathfrak{F}$ такава, че $B_n \supset B$ и $\mu(B) = \lim_n \mathsf{P}(B_n)$. Да означим с $B_n^+ = \cap_{i=1}^n B_i$. Тогава $\mu(B) = \lim_n \mathsf{P}(B_n^+)$. Същото е валидно и за $\overline{B}$ - $\exists B_n \in \mathfrak{F} : \mu(\overline{B}) = \lim_n \mathsf{P}(B_n), B_n \supset \overline{B}$. Нека означим с $B_n^- = \overline{\cap_{i=1}^n B_i}$. Така получаваме, че $B_n^- \subset B \subset B_n^+$. Тъй като $B \in \mathfrak{R}$, то $\mu(B) = 1 - \mu(\overline{B})$ и $\mathsf{P}(B_n^-) \uparrow \mu(B)$, $\mathsf{P}(B_n^+) \downarrow \mu(B)$.

Достатъчността е очевидна:

$$1 \leq \mu(B) + \mu(\overline{B}) \leq \mathsf{P}(B_n^+) + 1 - \mathsf{P}(B_n^-) \to 1.\square$$

Доказателство: [Теорема 1.3]
1.Да покажем, че $\mathfrak{R}$ е алгебра. $\mathfrak{R}$ очевидно съдържа $\Omega, \varnothing$, както и допълнението на всяко множество $B \in \mathfrak{R}$. Нека $A, B \in \mathfrak{R}$. Да означим с $A_n^+, B_n^+, A_n^-, B_n^-$ някои апроксимиращи редици на двете множества. Имаме очевидното включване: $A_n^- \cap B_n^- \subset A \cap B \subset A_n^+ \cap B_n^+$.Тогава:

$$\mathsf{P}(A_n^+ \cap B_n^+) - \mathsf{P}(A_n^- \cap B_n^-) =$$

$$\mathsf{P}(A_n^+ \cap B_n^+) - \mathsf{P}(A_n^- \cap B_n^+) + \mathsf{P}(A_n^- \cap B_n^+) - \mathsf{P}(A_n^- \cap B_n^-) =$$

$$\mathsf{P}((A_n^+ \setminus A_n^-) \cap B_n^+) + \mathsf{P}((B_n^+ \setminus B_n^-) \cap A_n^-) \leq$$

$$\mathsf{P}(A_n^+ \setminus A_n^-) + \mathsf{P}(B_n^+ \setminus B_n^-) \to 0$$

Така множеството $A \cap B$ е "апроксимируемо" и съгласно лема 1.5 $A \cap B \in \mathfrak{R}$, а следователно, и множествата $A \cup B, A + B, A \setminus B \in \mathfrak{R}$. $\mathfrak{R}$ е булова алгебра.От същите апроксимационни сметки следва, че μ е адитивна функция на $\mathfrak{R}$:

$$\mathsf{P}(A_n^-) + \mathsf{P}(B_n^-) \leq \mu(A + B) \leq \mu(A) + \mu(B) \leq \mathsf{P}(A_n^+) + \mathsf{P}(B_n^+).$$

2. Да покажем сега, че $\mathfrak{R}$ е и σ-алгебра. Нека $\{B_n \in \mathfrak{R}, B_n \supset B_{n+1}\}$ е намаляваща редица. Да означим с $A_n = B_n \setminus B_{n+1}$. Да означим с $B = \cap B_n$. От монотонността следва, че съществува граница на намаляващата редица $\mu(B_n)$ и $\mu(B) \leq \lim_n \mu(B_n)$. От полуадитивността следва, че $\mu(B) + \mu(\overline{B}) \geq 1$. От друга страна, от доказаната вече адитивност получаваме (за всяко n):

$$\sum_{i=1}^{n} \mu(A_i) = \mu(B_1) - \mu(B_{n+1}), \text{ т.е. } \sum_{i=1}^{\infty} \mu(A_i) = \mu(B_1) - \lim_n \mu(B_n).$$

Тъй като $\overline{B} = \overline{B_1} + \sum_{n=1}^{\infty} A_n$ ще получим: $\mu(\overline{B}) \leq 1 - \mu(B_1) + \sum_{n=1}^{\infty} \mu(A_n) = 1 - \lim_n \mu(B_n)$.Като съберем двете неравенства получаваме: $\mu(B) + \mu(\overline{B}) \leq 1$. Значи $B \in \mathfrak{R}$. 3. За да покажем, че μ е σ-адитивна на $\mathfrak{R}$, ще покажем, че е непрекъсната в $\varnothing$ и ще използуваме лема 1.1. Нека $\{B_n \in \mathfrak{R}, B_{n+1} \subset B_n\}$ е намаляваща редица такава, че $\cap B_n = \varnothing$. Както и в предната част получаваме, че $\mu(B) \leq \lim_n \mu(B_n^+) = \lim_n \mu(B_n^-)$. Но $\cap B_n^- \subset \cap B_n = \varnothing$. Тъй като $B_n^- \in \mathfrak{F}$ и $\mu = \mathsf{P}$, която е непрекъсната в $\varnothing$, получаваме $\mu(B) \leq \lim_n \mathsf{P}(B_n^-) = 0$.

Тъй като $\mathfrak{F} \subset \mathfrak{R}$, то и $\sigma(\mathfrak{F}) \subset \mathfrak{R}$. Единствеността на така построената мярка върху $\sigma(\mathfrak{F})$ следва от нейната изчислимост от P. Лесно се получава и, че $\mathfrak{R}$ е пълна, т.е. съдържа подмножествата на всички множества с нулева горна мярка. □

1.5 Следствия и коментарии

Като гледаме двете доказателства възниква законния въпрос: доколко $\mathfrak{R}$ е по-голяма от $\sigma(\mathcal{A})$? Нека разгледаме множеството $\mathbf{N}$ от подмножества $B \subset X$ с нулева горна мярка: $\mu(B) = 0$.

Теорема 1.4. *За всяко* $B \in \mathfrak{R} \quad \exists A \in \sigma(\mathcal{A}) : \mu(B \Delta A) = 0$.

Покажете го.

Теоремата на Каратеодори има фундаментално значение за анализа и от там за Теория на вероятностите. Видяхме как се изгражда мярката на Лебег върху единичния интервал [0,1).В курса по елементарната теория на вероятностите построихме вероятност съответствуваща на произволна ненамаляваща и непрекъсната отляво ограничена функция и мярка на декартово произведение от вероятностни пространства. Тук ще приведем още един пример, който показва редица възможности за разширяване на концепцията, заложена в теоремата на Каратеодори.

Пример 1.5 (Ц.Игнатов). *Да разгледаме в* $[0, 1) \times [0, 1)$ *семейството* $\mathcal{A}$, *състоящо се от множества от вида:* $A = [a, b) \times [0, c), 0 \leq a \leq b < 1, 0 \leq c < 1$. *Да определим върху него мярка по формулата* $\mu(A) = c * (b - a))$.

1. Семейството $\mathcal{A}$ не е съвсем полуалгебра: $A \cap B \in \mathcal{A}$, но $\overline{A} \notin \mathcal{A}$. μ е σ-адитивна на $\mathcal{A}$.
2. Нека сега разгледаме $\mathcal{B}$ - крайните обединения на непресичащи се множества от $\mathcal{A}$. То също не е съвсем булова алгебра- $\overline{B} \notin \mathcal{B}$, но μ коректно се определя на $\mathcal{B}$ и е σ-адитивна на $\mathcal{A}$.
3. Нека $\mathcal{M}$ е монотонният клас породен от $\mathcal{B}$. Той също не е булова алгебра, но μ коректно се разширява на $\mathcal{M}$, поради непрекъснатостта си в $\varnothing \in \mathcal{B}$.Покажете, че на класа $\mathcal{M}$ принадлежат подграфиките $F(f) = \{(x, y) : 0 \leq y < f(x)\}$ на всички непрекъснати и ограничени неотрицателни функции f, дефинирани на $[0, 1)$. Т.е. μ е интеграл на Риман: $\mu(F(f)) = \int_0^1 f(x)dx$. Интегрируеми в този смисъл са и функции с изброим брой точки на прекъсване (но от първи род). Този интеграл по линейност лесно се продължава и на функции с произволен знак.

Глава 2

Теория на мярката

Тук са дадени две прости теореми от теория на мярката, чието място е по-скоро в курса по анализ. Това са теоремите за разложението на Жордан-Хан и теоремата за производната на Радон-Никодим. Тук техните доказателства са дадени в този ред, защото втората теорема се доказва с помощта на първата. Това е и общоприетия път на доказване в рамките на класическия математически анализ. В заключение ще приведем една илюстрация на това как тези теореми следват леко от общите теореми на функционалния анализ.

2.1 Определения

Определение 2.1. *Реалната функция μ, определена върху елементите на буловата σ-алгебра $\mathfrak{A}$ и приемаща стойности в $R_1 \cup \{\infty\}$, се нарича мярка, ако удовлетворява условията:*

1. *неотрицателност:* $\mu(A) \geq 0, \forall A \in \mathfrak{A}$;
2. *адитивност:* $\mu(A_1 + A_2 + \dots) = \mu(A_1) + \mu(A_2) + \dots$;

Да напомним, че за множества операцията $+$ означава обединение на непресичащи се множества. По сравнение с вероятността в това определение е изпусната аксиомата за нормираност. По тази причина мярка може да бъде и функция приемаща само стойности 0 и ∞. За да се избегне тази неприятност се въвежда следното определение.

Определение 2.2. *Казваме, че една мярка μ, зададена на измеримото пространство $(X, \mathfrak{A})$ е σ-крайна, ако съществува изброимо разделяне на пространството $\gamma = \{H_1, H_2, \dots\}$, такова че $0 < \mu(H_i) < \infty$.*

По-нататък ще разглеждаме само σ-крайни мерки.

Определение 2.3. *Казваме, че мярката μ се доминира от мярката ν (или че мярката μ е абсолютно непрекъсната по отношение на мярката ν), когато $\forall A \in \mathfrak{A}$, за което $\nu(A) = 0$, е изпълнено $\mu(A) = 0$.*

2.2 Теорема на Жордан-Хан

За да се превърне множеството от мерки (зададени върху една буловата σ-алгебра) в линейно пространство, е достатъчно да разглеждаме линейни комбинации от вероятностни мерки. Тях наричаме знакопроменливи. За това е достатъчно в определение **??** да се откажем от условието за неотрицателност. Следната теорема ни осигурява коректността на такова определение.

За произволна σ-крайна мярка μ нека означим $\mu^+(A) = \max(\mu(A), 0)$, $\mu^-(A) = \max(-\mu(A), 0)$.

Теорема 2.1 (Жордан-Хан). *Нека μ е знакопроменлива мярка върху $\langle \Omega, \mathcal{F} \rangle$, тогава са изпълнени следните твърдения:*

1. *μ^+ и μ^- са положителни мерки;*
2. *μ^- е крайна и $\mu = \mu^+ - \mu^-$;*
3. *съществува положителен носител D_+ на μ*
   $\mu^+(A) = \mu(A \cap D_+)$ и $\mu^-(A) = \mu(A \cap \overline{D_+})$

Доказателство: (По (Обретенов и Хаджиев 1983))

Свойствата $\sigma-$адитивност и положителност на μ^+ и μ^- са очевидни.

Нека $\mathcal{B} = \{A \in \mathcal{F} : \mu(B) \leq 0 \forall B \subseteq A\}$ от определението за μ^+ имаме, че $\mathcal{B} = \{A \in \mathcal{F} : \mu^+(A) = 0\}$.

Лесно се вижда, че ако $B_n \in \mathcal{B}$, то $\cup B_n \in \mathcal{B}$. Наистина нека $B \subseteq \cup B_n$

$$\mu(B) = \mu(\cup B_n \cap B) = \mu(\sum_{n=1}^{\infty} C_n), \text{ където}$$

$C_1 = B_1 \cap B$, $C_2 = (B_2 \cap B) \backslash C_1$, и т.н. $C_n = (B_n \cap B) \backslash (C_{n-1} \cup \cdots \cup C_1)$. Очевидно $C_n \subseteq B_n$ и следователно $\mu(C_n) \leq 0$, откъдето

$$\mu(B) = \sum_{n=1}^{\infty} \mu(C_n) \leq 0.$$

Нека да разгледаме $B_n \in \mathcal{B}$ - минимизираща редица за

$$m = \inf_{B \in \mathcal{B}} \mu(B), \text{ тогава } \mu(\cup B_n) = m < \infty,$$

наистина $C = \cup B_n \in \mathcal{B}$, тогава $C \backslash B_n \in \mathcal{B}$, но $\mu(C \backslash B_n) \leq 0$ и $\mu(B_n) \leq 0$, откъдето

$$m \leq \mu(C) = \mu(C \backslash B_n) + \mu(B_n) \leq \mu(B_n) \underset{n \to \infty}{\longrightarrow} m$$

получаваме, че $m = \mu(C) < \infty$.

Да разгледаме множеството $D_+ = \bar{C} = \overline{\cup B_n}$. Ще покажем, че то е положителен носител на μ. По-точно ще покажем, че за всяко $A \subseteq D_+$ $\mu(A) \geq 0$.

1. Нека $A \subseteq D_+$, ако $\alpha = \mu(A) < 0$, то $\mu^+(A) > 0$. Наистина да допуснем противното, тогава $\mu(B) \leq 0$ за всяко $B \subseteq A$ и следователно $A \in \mathcal{B}$, но тогава $m = \inf_{B\in \mathcal{B}} \mu(B) > \mu(C + A) = m + \alpha$, което е противоречие тъй като $C + A \in \mathcal{B}$.

2. Да допуснем, че съществува $A \subseteq D_+$ с $\mu(A) < 0$ и $d = \mu^+(A) < \infty$. Според **1.** $\mu^+(A) > 0$. Според определението за μ^+ избираме $E_1 \subset A$ и $\mu^+(A)/2 < \mu(E_1)$, $E_2 \subset A\backslash E_1$ и $\mu^+(A\backslash E_1)/2 < \mu(E_2)$ и т.н. $E_{n+1} \subset A\backslash(E_1 + \cdots + E_n)$ с $\mu^+(A\backslash(E_1 + \cdots + E_n)) < \mu(E_{n+1})$. Да отбележим, че μ^+ е $\sigma-$адитивна и положителна и тогава по индукция

$$\mu^+(A\backslash(E_1 + \cdots + E_n)) < d/2^n, \text{ тъй като}$$

$$\mu^+(A\backslash(E_1 + \cdots + E_n)) - \mu(E_{n+1}) = \mu^+(A\backslash(E_1 + \cdots + E_n + E_{n+1})) <$$

$$< \mu^+(A\backslash(E_1 + \cdots + E_n))(1 - 1/2) < d/2^{n+1}.$$

Следователно, ако $E = \sum E_n$, то $\mu^+(E) = \mu^+(A)$, наистина

$$\mu^+(A\backslash E) = \lim_{n\to\infty} \mu^+(A\backslash(E_1 + \cdots + E_n)) = 0, \text{ и}$$

$$\mu^+(A) = \mu^+(A\backslash E) + \mu^+(E).$$

Имаме, че $0 < \mu(E) = \mu(E_1) + \cdots + \mu(E_n) + \cdots$, следователно $\mu(A\backslash E) < 0$, т.к. $\mu(A) < 0$. Да допуснем, че $\exists B \subset A\backslash E$, така че $\mu(B) > 0$ тогава $\mu^+(A) \geq \mu^+(B + E) > \mu^+(E) = \mu^+(A)$ и стигаме до противоречие.

Следователно за всяко $B \subseteq A\backslash E$ $\mu(B) < 0$, но тогава $A\backslash E \in \mathcal{B}$ и това води до противоречие тъй като

$$m = \inf_{B\in \mathcal{B}} \mu(B) < \mu(C + A\backslash E) = m + \mu(A\backslash E).$$

Това е в противоречие с допускането, че съществува $A \subseteq D_+$ с $\mu(A) < 0$ и $\mu^+(A) < \infty$. За да се уверим, че за всяко $A \subseteq D_+$ $\mu(A) \geq 0$, което би завършило доказателството, че D_+ е положителен носител на μ, остава да разгледаме:

3. за всяко $A \subseteq D_+$ с $\mu(A) < 0$ имаме $\mu^+(A) = \infty$. Този случай изглежда 'абсурден', но за пълнота трябва да се разгледа. Съществува $G_1 \subset A$ с $\mu(G_1) > 1$, по индукция, тъй като $\mu(A\backslash G_1) < 0$ и $\mu^+(A\backslash G_1) = \infty$ съществува $G_2 \subset A\backslash G_1$ с $\mu(G_2) > 1$, и т.н. Окончателно за $G = G_1 + G_2 + \cdots$ $\mu(G) = \infty$, но

$$\mu(A) = \mu(A\backslash G) + \mu(G) < 0, \text{ откъдето } \mu(A\backslash G) = -\infty,$$

което е в противоречие, т.к. $\mu : \mathcal{F} \longrightarrow (-\infty, \infty]$. Това показва, че този случай е невъзможен. $\square$

2.3 Теорема на Радон-Никодим

Теорема 2.2 (Радон-Никодим). *Нека в измеримото пространство $\langle\Omega,\mathcal{F}\rangle$ са дадени $\sigma-$адитивните мерки μ и ν, като μ е абсолютно непрекъсната относно ν. Тогава съществува единствена (с точност до ν – нулево множество) измерима функция $f:\Omega\longrightarrow\Omega$, такава че*

$$\mu(A)=\int\limits_A f(\omega)\nu(d\omega),\ \textit{за всяко измеримо } A.$$

Функцията f се нарича производна на Радон-Никодим на μ относно ν, $f=d\mu/d\nu$.

Доказателство:

I. Нека μ и ν са крайни и положителни и нека тогава за определеност да считаме, че са вероятностни мерки.

Разглеждаме съвкупността от измерими функции

$$\mathcal{M}=\{f>0:I_f(A)=\int\limits_A fd\nu\le\mu(A)\text{ за всяко } A\in\mathcal{F}\}.$$

Ще покажем, че $\mathcal{M}$ е непразна с помощта на Теоремата на Жордан-Хан. Нека $\lambda>0$ означаваме $\mu_\lambda(A)=\mu(A)-\lambda\nu(A)$. Нека $A\in\mathcal{F}$ и $\mu(A)>0$, $\mu_\lambda(A)$ е непрекъсната функция на λ при фиксирано A и $\mu_0(A)=\mu(A)>0$, то съществува $\lambda=\lambda(A)$, така че $\mu_{\lambda(A)}(A)>0$. Според теоремата на Жордан-Хан нека $D(\lambda)$ е положителният носител на $\mu_\lambda(\cdot)$, според горните разсъждения $\mu_\lambda(D(\lambda))=\mu^+_\lambda(\Omega)$ е строго положително число за достатъчно малки, но положителни, λ. Тогава разглеждаме

$$f_\lambda(\omega)=\lambda 1_{D(\lambda)}(\omega),$$

от горното следва, че $f_\lambda,\lambda>0$ са нетривиални (т.е. не п.с. константни) функции. Ясно е, че за всяко $\lambda>0$ $f_\lambda\in\mathcal{M}$ следователно $\varnothing\neq\mathcal{M}$.

Ако $g_n\in\mathcal{M}$, то очевидно $\max\{g_1,g_2,\dots,g_n,\dots\}\in\mathcal{M}$, тогава съвкупността

$$\mathcal{M}_0=\{f=\max_{\lambda\in\Delta}f_\lambda:\Delta\subset(0,\infty)\ \Delta-\text{изброимо }\}$$

се съдържа в $\mathcal{M}$.

Да разгледаме минимизираща редица $\{g_n,n\in\mathbf{N}\}$ за супремума:

$$1\ge M=\sup_{f\in\mathcal{M}}\int\limits_\Omega fd\nu,\ \text{тогава за}$$

$g=\max_{n\in\mathbf{N}}g_n$ имаме $I_g(\Omega)=\int_\Omega gd\nu=M$. Т.е. този супремумс се достига в g. Ще покажем, че g е единствена $\nu-$почти сигурно. Нека за $h\in\mathcal{M}$ $I_h(\Omega)=M$, но тогава ако $I_h(A) > I_g(A)$, то $f = \max\{g, h\} \in \mathcal{M}$, но $I_f(\Omega) \geq I_g(\overline{A}) + I_h(A) > M$, което е противоречие с определението на M.

Нека тогава g_μ е единствената функция от $\mathcal{M}$, за която $M = I_{g_\mu}(\Omega)$. От определението на $\mathcal{M}$ имаме, че $\mu_1 = \mu - I_{g_\mu} \geq 0$. Да допуснем, че съществува $A \in \mathcal{F}$, такова че $\mu_1(A) > 0$, тогава тъй като и $\mu_1 \prec \nu$ ще приложим същите разсъждения за мерките $\frac{1}{\mu_1(A)}\mu_1(A \cap \cdot)$ и $\nu(\cdot)$ и ще намерим нетривиална функция h, за която $0 \leq I_h(\cdot) \leq \mu_1(\cdot)$. Но това води до противоречие тъй като очевидно $g_\mu + h \in \mathcal{M}$, и тогава

$$M \geq I_{g_\mu + h}(\Omega) > I_{g_\mu}(\Omega) = M.$$

Следователно $\mu_1 = \mu - I_{g_\mu} = 0$, което означава, че за всяко $A \in \mathcal{F}$

$$\mu(A) = \int\limits_A g_\mu(\omega) d\nu(\omega).$$

□

2.4 "Модерни" доказателства

Теорема 2.3. *Множеството от крайни мерки $\mathcal{M}$ върху $\langle \Omega, \mathcal{F} \rangle$ е Банахово пространство и решетка относно конуса на неотрицателните мерки.*

Доказателство: $\mathcal{M}$ е локално изпъкнало и затворено относно силната норма: $||\mu|| = \sup_{A \in \mathcal{F}} |\mu(A)|$. □

Да разгледаме преобразованието $F : \mathcal{M} \longrightarrow \mathcal{M}$

$$\nu(A) = \int_A f(x) d\mu(x),$$

където $0 < f < C$ е ограничена измерима функция.

Това преобразование е непрекъснат линеен оператор на $\mathcal{M}$ с норма $< C$.

Теорема 2.4. *Линейното пространство* $\mathbf{R}$ *породено от всички такива оператори образува комутативна Банахова алгебра.*

Доказателство: Достатъчно е да проверим комутативността. □

Остава да зафиксираме мярката μ и разгледаме образа ѝ $\mathcal{M}(\mu)$ с всички оператори от $\mathbf{R}$.

Теорема 2.5 (Радон-Никодим). *Множеството $\mathcal{M}(\mu)$ е затворено линейно подпространство на $\mathcal{M}$ и се състои от доминираните от μ мерки.*

Да отбележим, че мерките в $\mathcal{M}(\mu)$ са крайни и знакопроменливи. За да докажем пълната теорема, трябва само да преобразуваме σ-крайните мерки в еквивалентни крайни, което е тривиално.

Доказателство: □

Теорема 2.6 (Жордан-Хан).

Доказателство: □

Глава 3

Теорема на Колмогоров

3.1 Безкрайномерни разпределения

Даден е случайният процес $\{\xi_t, t \in T\}$, за $T \subseteq \mathbf{R}$. Ще построим едно измеримо пространство, т.е. множество от елементарни събития със $\sigma-$алгебра от негови подмножества.

Нека $\tilde{\Omega} = \mathbf{R}^T = \{\omega : T \to \mathbf{R}\}$ е множеството от всички функции $f : T \to \mathbf{R}$.

Определение 3.1. *Нека $B \in \mathcal{B}(\mathbf{R}^n)$ е произволно Борелово подмножество на $\mathbf{R}^n$. Казваме, че множеството*

$$C(t_1, \ldots, t_n; B) = \{\omega \in \mathbf{R}^T : (\omega(t_1), \ldots, \omega(t_n)) \in B\}$$

е цилиндрично с основа B.

Лесно се вижда, че:

1. Допълнението на цилиндрично множество е цилиндрично;
2. Крайно сечение на цилиндрични множества е цилиндрично;
3. Крайно обединение на цилиндрични множества е цилиндрично.

Следователно класът на цилиндричните множества в $\tilde{\Omega}$ образува булова алгебра. Тогава определяме буловата $\sigma-$алгебра $\mathcal{B}(\mathbf{R}^T)$ като минималната $\sigma-$ алгебра, съдържаща всички цилиндрични множества.

Така е определено едно измеримо пространство $\left\langle \mathbf{R}^T, \mathcal{B}(\mathbf{R}^T) \right\rangle$, в което (както по нататък ще видим) благодарение на теоремата на Колмогоров може да се дефинира разпределение P_ξ. За това разпределение е изпълнено:

$$\mathsf{P}_\xi(C) = \mathsf{P}\{(\xi_{t_1}, \ldots, \xi_{t_n}) \in B\},$$

за всяко цилиндрично множество $C = C(t_1, \ldots, t_n; B), B \in \mathcal{B}(\mathbf{R}^n)$.

3.2 Съгласувани семейства от разпределения

Определение 3.2. *За случайният процес $\{\xi_t, t \in T\}$ разпределенията*

$$F_{t_1,\ldots,t_n}(x_1,\ldots,x_n) = \mathsf{P}\{\xi_{t_1} \leq x_1, \cdots, \xi_{t_n} \leq x_n\},$$

се наричат крайномерни разпределения на процеса.

За крайномерните разпределения са очевидни следните свойства:

$$F_{t_1,\ldots,t_n}(x_1,\ldots,x_n) = F_{t_{\sigma(1)},\ldots,t_{\sigma(n)}}(x_{\sigma(1)},\ldots,x_{\sigma(n)}), \tag{3.1}$$

където σ е произволна пермутация на $\{1,\ldots,n\}$.

$$F_{t_1,\ldots,t_n,t_{n+1}}(x_1,\ldots,x_n,\infty) = F_{t_1,\ldots,t_n}(x_1,\ldots,x_n) \tag{3.2}$$

Определение 3.3. *Произволно семейство от многомерни разпределения $\{F_{t_1,\ldots,t_n}(x_1,\ldots,x_n), t_i \in T\}$, което удовлетворява условията (3.1) и (3.2) се нарича съгласувано, а условията - условия за съгласуваност.*

Теорема 3.1 (Колмогоров). *Нека е дадено съгласуваното семейство от разпределения $\{F_{t_1,\ldots,t_n}(x_1,\ldots,x_n), t_i \in T\}$, където $T \subseteq \mathbf{R}$. Тогава в измеримото пространство $\left\langle \mathbf{R}^T, \mathcal{B}(\mathbf{R}^T) \right\rangle$ съществува единствена вероятностна мярка P_F, за която:*

$$\mathsf{P}_F(C(t_1,\ldots,t_n;B)) = F_{t_1,\ldots,t_n}(B), \text{ за произволно}$$

цилиндрично множество $C = C(t_1,\ldots,t_n;B)$ с основа B.

Това означава, че действието на едно ”безкрайномерно” вероятностно разпределение в $\left\langle \mathbf{R}^T, \mathcal{B}(\mathbf{R}^T) \right\rangle$ се определя от действието му върху ”крайномерните” цилиндрични множества.

Когато имаме даден случаен процес, то естествено възниква семейството от крайномерни разпределения. Тези разпределения, според горната теорема пораждат единствено вероятностно разпределение P_ξ, в $\left\langle \mathbf{R}^T, \mathcal{B}(\mathbf{R}^T) \right\rangle$, за което

$$\mathsf{P}_\xi(\{\omega \in \mathbf{R}^T : (\omega(t_1),\ldots,\omega(t_n)) \in B\}) = \mathsf{P}\{(\xi_{t_1},\ldots,\xi_{t_n}) \in B\}, B \in \mathcal{B}(\mathbf{R}^n).$$

Разпределението P_ξ се интерпретира като мярка върху траекториите на процеса $\{\xi_t, t \in T\}$, които са елементи на $\tilde{\Omega} = \mathbf{R}^T$. Естествено е да поискаме P_ξ върху цилиндричните множества да се определя чрез крайномерните разпределения на процеса, а теоремата на Колмогоров показва, че това е достатъчно за да се определи мярката P_ξ върху всички $\mathcal{B}(\mathbf{R}^T)-$измерими множества от траектории.

В този смисъл P_ξ представлява разпределението на процеса $\{\xi_t, t \in T\}$ в пространството от траектории $\left\langle \mathbf{R}^T, \mathcal{B}(\mathbf{R}^T) \right\rangle$.

3.3 Доказателство

Доказателство: Да означим с X множеството от всички функции f определени на R.

1. За всеки фиксиран набор t да означим с $\mathcal{B}_t$ буловата алгебра от подмножества на X, породена от "интервалите" $I = I(t, x, y) = \{f : x_i \le f(t_i) < y_i, i = 1, 2, \dots, n\}$ при произволни набори $x < y$ със същия като набора t размер. Върху $\mathcal{B}_t$ с помощта на разпределенията F_t определяме вероятностна мярка P_t. Това определение е коректно, поради условието за симетричност (3.1).

2. Между наборите и алгебрите съществува естественото съотношение:
   - ако $t \subset s$, то $\mathcal{B}_t \subset \mathcal{B}_s$.
   Сега условието за съгласуваност (3.2) ни позволява да се убедим и в съгласуваността на определените по-горе вероятности:
   - ако $B \in \mathcal{B}_t$, то $P_t(B) = P_s(B)$, за всяко $t \subset s$.

3. Да означим с $\mathcal{B}_0 = \cup \mathcal{B}_t$ за всички крайни набори t. Това, разбира се е булова алгебра и вероятността $P(B) = P_t(B)$ върху нея е правилно определена. Да разгледаме сега минималната σ-алгебра $\mathcal{B}$ съдържаща $\mathcal{B}_0$. За да използуваме теоремата на Каратеодори за продължението, трябва да покажем σ-адитивността на P върху $\mathcal{B}_0$. Това е най-трудната част на теоремата.

Достатъчно е да покажем σ-адитивността на P за интервали. Трябва да покажем, че ако

$$I = I_1 + I_2 + \dots, \quad I_i \cap I_j = \varnothing \text{ то } P(I) = P(I_1) + P(I_2) + \dots.$$

Да допуснем противното, т.е. $\exists h > 0 : \forall r : P(I) - \sum_{i=1}^{r} P(I_i) > h$. Да наречем за простота основа на интервала $I \in \mathcal{B}_t$ множеството t. Ясно е, че можем да приемем без да намаляваме общността, че основите на интервалите I_n нарастват - $t_n = \{\tau_1, \tau_2, \dots, \tau_{k_n}\}$, $t \subset t_1 \subset t_2 \subset \dots$ и $k_n \to \infty$. Да означим с $J_n = I - \sum_{i=1}^{n} I_i \in \mathcal{B}_t$. Съгласно допускането, имаме: $P(J_n) > h > 0$. Тъй като вероятностите P_{t_n} са Борелови, ако ги разгледаме върху крайномерните пространства R^{k_n}, съществуват компактни множества $B_n \subset J_n$ такива, че $P(B_n \setminus J_n) < \varepsilon 2^{-n}$. Нека $A_n = \cap_{i=1}^{n} B_i$. Имаме очевидното съотношение

$$P(J_n \setminus A_n) = P(J_n \overline{A_n}) = P(\cup_{i=1}^{n} J_n \overline{B_i}) < \varepsilon$$
$$P(A_n) \ge P(J_n) - \varepsilon \ge h - \varepsilon > 0$$

Тъй като $A_n \supset A_{n+1}$ сечението им е непразно за всяко фиксирано τ_k. Така вече лесно можем да построим функция в $f \in \cap A_n \subset I$. Така стигаме до противоречие с предположението - $I = I_1 + I_2 + \dots$.

За да завършим доказателството остава да използуваме теоремата на Каратеодори за продължението. Така получаваме вероятностното пространство $(X, \mathcal{B}, P)$. Определяме на него случайния процес $\xi(x, \tau) = x(\tau)$. Лесно се проверява коректността на това определение. $\square$

Глава 4

Проектори

4.1 Определения

В тази секция ще разгледаме проекторите - най-простите и навярно най-използувани в математиката преобразования.

Определение 4.1. *Нека P е преобразование на множеството X в себе си. Казваме, че P е проектор, ако $P^2 = P$. Понякога това свойство се нарича идемпотентност.*

От това определение следват следните свойства:

- Множеството X се разпада на две непресичащи се подмножества: $X = PX + \overline{PX}$;
- $PX \neq \varnothing$;
- ако $\overline{PX} = \varnothing$, т.е. $(PX = X)$, то $P = I$;
- траекториите на P съдържат най-много два различни елемента - единият е винаги в PX, другият - в $\overline{PX}$;

Теорема 4.1. *Ако P и Q са комутиращи проектори ($PQ = QP$), то*

1. *PQ е проектор;*
2. *$(PQ)X = QX \bigcap PX$.*

Доказателство: 1.$(PQ)(PQ) = PQ^2P = PQP = P^2Q = PQ$.
2. Следователно, $\forall x \in X$ е изпълнено: $QPx \in QX, PQx \in Px$. Т.е. $PQX \subset PX \bigcap QX$. Нека сега $p \in PX \bigcap QX$. Тогава $Pp = p, Qp = p$. Но тогава $PQp = p$, значи $p \in PQX$. Значи $QX \bigcap PX \subset PQX$. $\square$

4.2 Проектори в линейно пространство

Теорема 4.2. *Нека сега X е векторно пространство (над полето на реалните числа) и P е линеен оператор ($P(ax) = aPx, P(x+y) = Px + Py$) и проектор ($P^2 = P$). Тогава:*

1. *образът PX е линейно подпространство;*
2. *ядрото $N_P = \{x : Px = 0\}$ е линейно подпространство;*
3. *$Q = I - P$ е проектор, $QP = PQ = 0$, $QX = N_P$;*
4. *за всеки комутиращ с P проектор Q имаме, че $P - QP$,$Q - QP$ са проектори.*

Доказателство: 1. Нека $x, y \in PX$. Тогава $P(ax + by) = aPx + bPy = ax + by$.
2. Нека $x, y \in N_P$. Тогава $P(ax + by) = aPx + bPy = 0$.
3. Нека $Qx = x - Px = (I - P)x$,$P(I - P) = P - P^2 = 0$. Тогава $Q^2 = (I - P)(I - P) = I - 2 * P + P = I - P$.
4. Да докажем,че $Q - PQ$ е проектор. Ако $QP = PQ$, то
$(I - P)Q = Q - PQ = Q - QP = Q(I - P)$,
$(I - P)Q(I - P)Q = (I - P)Q(I - P) = (I - P)(I - P)Q = (I - P)Q$. $\square$

4.3 Ортогонални проектори

Нека сега X е Хилбертово пространство (пак над полето на реалните числа) и с (x, y) сме означили скаларното произведение в X.

Определение 4.2. *Линеен самоспрегнат ($(Px, y) = (x, Py)$) проектор се нарича ортогонален.*

Теорема 4.3. *За да бъде линеен проектор P в Хилбертовото пространство X самоспрегнат е необходимо и достатъчно PX и $(I - P)X = N_P$ да са ортогонални линейни подпространства: $(Px, (I - P)y) = 0, \forall x, y \in X$;*

Доказателство: Необходимост. Нека P е самоспрегнат. Тогава $(Px, (I - P)y) = (x, P(I - P)y) = 0$, Значи $PX \perp N_P$.
Достатъчност. Нека $(Px, (I - P)y) = 0, \forall x, y \in X$. Нека $H = P^*$ е спрегнат оператор на P (т.е.$(Px, y) = (x, Hy)$ и $(Hx, y) = (x, Py)$). Нека $y \in X$. Можем да го запишем във вида $y = (I - P)y + Py = y_1 + y_2$, където $y_1 =\in N_P, \quad y_2 \in PX$. Имаме:

$$
\begin{aligned}
((P - H)x, y) = ((P - H)x, y_1) + ((P - H)x, y_2) = \\
(Px, y_1) - (Hx, y_1) + (Px, y_2) - (Hx, y_2) = \\
(Px, y_1) - (x, Py_1) + (Px, y_2) - (x, Py_2) = \\
(Px, (I - P)y) - 0 - (Px, Py) - (x, Py) = \\
(Px, (I - P)y) + ((P - I)x, Py) = 0.
\end{aligned}
$$

Следователно, $||P - H||^2 = sup_{||y|| \leq 1}((P - H)y, (P - H)y) = 0$. $\square$

Пример 4.1. *Не ортогонален проектор.*

Да разгледаме матрицата

$$Z = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

като линеен оператор в $X = R^2$. Очевидно е, че $Z^2 = Z$. Но

$$ZX = \begin{pmatrix} x \\ 0 \end{pmatrix}, \quad N_P = \begin{pmatrix} x \\ -x \end{pmatrix}$$

и тези две едномерни подпространства не са ортогонални.

Теорема 4.4. *Нека Z е линейно подпространство на X. За всяко $x \in X$ да означим с $H : X \to Z$ преобразованието*

$$x \longrightarrow \underset{z \in Z}{argmin} ||z - x||.$$

H е ортогонален проектор.

Доказателство: Очевидно преобразованието H е проектор ($H^2 = H, H(\lambda x) = \lambda H(x)$). Да допуснем, че е изпълнено неравенството $\delta = (Hx_0, (I - H)y_0) \neq 0$ за някакви $x_0, y_0 \in X$. Да запишем $x \overset{def}{=} Hx_0, y \overset{def}{=} (I - H)y_0, \delta = (x, y)$. Нека $z = x + y$, т.е. $Hz = x, (I - H)z = y$. Имаме

$$||z - \lambda x||^2 = ||z||^2 - 2\lambda(\delta + ||x||^2) + \lambda^2 ||x||^2.$$

Следователно минимум на тази квадратична функция на λ ще получим при

$$\lambda_0 = \frac{\delta + ||x||^2}{||x||^2}, \quad Hz = \lambda_0 x.$$

Следователно, $\lambda_0 = 1, \delta = 0$. $\square$

Теорема 4.5. *Нека е зададена редица H_n от ортогонални проектори в Хилбертовото пространство X, такава че подпространствата им монотонно нарастват: $H_n(X) \subset H_{n+1}(X)$. Тогава съществува единствен ортогонален проектор H, който е граница на тази редица в слаб смисъл (на норма на стойностите).*

Доказателство: Нека означим $Z_0 = 0$, $Z_n = H_n(X)$ и $Z = \cup Z_n$. Очевидно Z е линейно подпространство на X. Нека H е ортогоналния проектор върху него. За всяко $x \in X$ имаме

$$Hx = \sum_{n=0}^{\infty} (H_{n+1} - H_n)x$$

Тъй като $(H_n x, (H_{n+1} - H_n)x) = 0$, имаме:

$$||Hx||^2 = \sum_{n=0}^{\infty} ||(H_{n+1} - H_n)x||^2 < \infty$$

Следователно,

$$||(H - H_n)x||^2 = \sum_{k=n+1}^{\infty} ||(H_{k+1} - H_k)x||^2 \to 0. \quad \forall x \in X.\square$$

Глава 5

Измерими разделяния

В тази лекция ще се опитаме да представим накратко основните резултати от теорията на мярката, които ще ни потрябват нататък. Ще разгледаме и условни вероятностни пространства.

5.1 Изоморфизми и класификация

За начало ще разглеждаме само вероятностни пространства със σ-адитивна вероятност и разделяща точките и попълнена σ-алгебра.

Определение 5.1. *Една точка в такова вероятностно пространство се нарича атом, ако има положителна вероятност.*

Теорема 5.1. *Всяко вероятностно пространство има не повече от изброим брой атоми.*

Доказателство: Броят атоми с вероятности по-големи от $1/n$ е не повече от $n-1$. Следователно те последователно могат да се отделят и номерират. $\Box$

Да напомним понятието изоморфизъм в теорията на мярката. Ще започнем със следното определение:

Определение 5.2. *Изображението*

$$T : B' \longrightarrow B'', \qquad B' \in \mathfrak{B}', \quad B'' \in \mathfrak{B}''$$

се нарича хомоморфизъм на вероятностни пространства $(\Omega', \mathfrak{B}', \mathsf{P}')$ *и* $(\Omega'', \mathfrak{B}'', \mathsf{P}'')$*, ако*

1. *T е определено п.с., т.е. върху достоверно събитие в Ω' (т.е. $\mathsf{P}(B') = 1$);*

2. *T е измеримо: $T^{-1}(A'') \in \mathfrak{B}'$ за $\forall A'' \in \mathfrak{B}''$;*

3. *T запазва вероятността: $\mathsf{P}'(T^{-1}(A'')) = \mathsf{P}''(A'')$ за $\forall A'' \in \mathfrak{B}''$;*

(Тук $T^{-1}(B)$ означава пълен праобраз, т.е. $T^{-1}(B) = \{\omega : T(\omega) \in B\}$.)

Определение 5.3. *Изображението*

$$T : (\Omega', \mathfrak{B}', \mathsf{P}') \longrightarrow (\Omega'', \mathfrak{B}'', \mathsf{P}'')$$

се нарича изоморфизъм на вероятностни пространства, ако е хомоморфизъм и T е обратимо п.с. в Ω''.

Така ако две вероятностни пространства са изоморфни, стават изоморфни и сл.в. зададени върху двете пространства. При този изоморфизъм очевидно стойностите им върху подмножества с мярка 0 нямат значение. Така вместо с конкретни функции ние се занимаваме с класове на еквивалентност.

Определение 5.4. *Да означим $\prod(\lambda)$ вероятностното пространство $(X, \mathfrak{A}, \mathsf{P})$:*

- *X е произведение от единични интервали в количество равно на кардиналността на числото λ.*
- *$\mathfrak{A}$ е породена от цилиндрически множества с крайни борелови основи;*
- *P - произведение от лебегови мерки.*

Да напомним определението на цилиндрично множество в декартово произведение от множества $\prod_i X_i$. Това е множество B, за което съществува крайно n, че то може да бъде представено във вида:

$$B = A^n \times X_{n+1} \times X_{n+2} \times \dots, \qquad A^n \subset \prod_{i=1}^{n} X_i.$$

Пример 5.1. *Множеството M от всички функции определени на интервала $[0,1]$ със стойности в същия интервал.*

Множеството M е декартово произведение на континуум интервали $[0,1]$. Само-то то има мощност по-голяма от континуум. Цилиндрично множество $C \subset M$ е, например, множеството:

$$C(t_1.t_2, \dots, t_n, x_1.x_2, \dots, x_n) = \{f : f(t_1) = x_1, f(t_2) = x_2, \dots, f(t_2) = x_2\}$$

за всеки два фиксирани набора $t_1.t_2, \dots, t_n$ и $x_1.x_2, \dots, x_n$ от елементи на $[0,1]$.

Теорема 5.2. *(Теорема на Дороти Махарам) Всяко вероятностно пространство без атоми е изоморфно на следната сума от несъвместими събития:*

$$p_0[0,1] + \sum_{i=1}^{\infty} p_i \prod(\lambda_i), \quad \sum_{i=0}^{\infty} p_i = 1, \quad 0 \le p_i, \tag{5.1}$$

λ_i - различни и по-големи от изброимост.

Нека сега видим какво става с изброимо породените вероятностни пространства.

Теорема 5.3. *Всяко такова вероятностно пространство без атоми е изоморфно на интервала* $[0, 1]$*, снабден с мярката на Лебег върху бореловата* σ*-алгебра.*

Доказателство: Просто следствие от теоремата на Махарам. □

Пример 5.2. *Схема на Бернули е редица от независими еднакво разпределени случайни величини* $\{\xi_i, i = 1, 2, \dots\}$*, всяка от които приема две стойности: 1 и 0 с вероятности (съответно)* p *и* $q = 1 - p$*.*

Съгласно теоремата на Каратеодори този набор от случайни величини поражда вероятностна мярка в пространството от безкрайни двоични редици. Да напомним това построение. Нека означим това пространство с $\mathfrak{E}$ и неговите елементи с $\varepsilon = \{\varepsilon_1, \varepsilon_2, \dots\}$. За всяко крайно n наборът от първите n случайни величини е векторна случайна величина, която определя в основното вероятностно пространство набор от събития:

$$W_{\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n} = \{\omega : \{\xi_i = \varepsilon_i\}, i = 1, 2, ...n\} = \bigcap_{i=1,2,...n} \{\xi_i = \varepsilon_i\}.$$

Определената от тях булева подалгебра W^n се състои от крайни обединения на непресичащи се множества от този тип. От друга страна тези булеви алгебри се влагат в пространството $\mathfrak{E}$, като определението на съответните множества се продължава цилиндрично — неупоменатите координати са свободни от ограничение. Там алгебрите образуват нарастваща редица. Върху множествата от тази алгебра вероятността се определя просто:

$$\mathsf{P}(W_{\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n}) = p^k q^{n-k}, \quad k = \sum_{k=1}^{n} \varepsilon_i. \tag{5.2}$$

Минималната σ-алгебра, която съдържа W ще означим с $\sigma(W)$. Това че вероятността е непрекъсната върху W следва от нейното определение — всяко ограничение върху нарастващ брой индекси поражда намаляваща към нула редица от вероятности. Така че от теоремата на Каратеодори следва съществуването и единствеността на вероятностната мярка P на $\sigma(W)$ и получаваме вероятностното пространство $(\mathfrak{E}, \sigma(W), \mathsf{P})$.

Съгласно теорема 5.3 това вероятностно пространство е изброимо породено, няма атоми и следователно е изоморфно на интервала $[0, 1]$ с мярката на Лебег. В лекциите по вероятности ние построихме един конкретен изоморфизъм.

5.2 Разделяния и алгебри

Да напомним някои основни свойства на булевите подалгебри. С всяко крайно разделяне (пълна група от събития) γ на пространството Ω е свързана еднозначно определена булова подалгебра $\mathfrak{A}_\gamma \subset \mathfrak{A}$. Ако разделянето γ се състои от N непразни множества, то буловата подалгебра е крайна и съдържа точно 2^N елемента.

Изобщо казано е верна следната

Теорема 5.4. *Съществува взаимно - еднозначно съответствие между крайните булеви алгебри и крайните разделяния на пространството Ω.*

Доказателство: 1. Да съпоставим на всяко разделяне булова алгебра. Нека $\gamma = \{H_1, H_2, \dots H_n\}$. Нека означим с N множеството $\{1, 2, \dots, n\}$ и нека $I \subset N$. Да разгледаме семейството от множества от вида: $A_I = \sum_{i \in I} H_i$. Те образуват булева подалгебра на $\mathfrak{A}$. Наистина, $A_\varnothing = \varnothing$, $A_N = \Omega$, $A_{\overline{I}} = \overline{A_I}$, $A_{I \cup J} = A_I \cup A_J$. Да означим тази подалгебра $\mathfrak{A}_\gamma$.

2. Обратно, нека $\mathfrak{B} \subset \mathfrak{A}$ е крайна (състояща се от краен брой множества) булева подалгебра. Ще наречем атом на $\mathfrak{B}$ всяко множество от вида:

$$A(\omega) = \cap_{B \in \mathfrak{B}} B^{(\omega)}, \qquad B^{(\omega)} = \begin{cases} B, & \omega \in B \\ \overline{B}, & \omega \in \overline{B} \end{cases}$$

Така на всяко $\omega \in \Omega$ съпоставяме единствено множество $A(\omega) \in \mathfrak{B}$, което го съдържа - значи множеството от атоми образува разделяне. Взаимната еднозначност на съответствието предоставяме на читателя. $\square$

Това съответствие поражда интересна връзка между операциите определени върху подалгебри и разделяния. На тривиалната алгебра $\{\varnothing, \Omega\}$ отговаря тривиалното разделяне $\nu = \{\Omega\}$.

Определение 5.5. *Казваме, че δ е по-ситно (или по фино) от γ($\gamma \le \delta$) ако $\mathfrak{A}_\gamma \subset \mathfrak{A}_\delta$.*

Това означава, че елементите на разделянето γ са по-груби, те се състоят от по няколко елемента на разделянето δ. И двете множества – това на булевите алгебри и това на разделянията – са частично наредени. Но тъй като в тях са определени и операциите max и min, те са и решетки. В частност,

$$max(\mathfrak{A}_\delta, \mathfrak{A}_\gamma) = \sigma(\mathfrak{A}_\delta, \mathfrak{A}_\gamma) = \mathfrak{A}_{max(\delta,\gamma)}.$$

Разделянето $max(\delta, \gamma)$ се състои от всевъзможни сечения на елементи от двете разделяния.

$$min(\mathfrak{A}_\delta, \mathfrak{A}_\gamma) = \mathfrak{A}_\delta \bigcap \mathfrak{A}_\gamma = \mathfrak{A}_{min(\delta,\gamma)}.$$

Разделянето $min(\delta, \gamma)$ се състои от събития, които могат да се представят като обединения на елементи, както от едното, така и от другото разделяне.

Благодарение на това двете структури - на булови алгебри и разделяния стават изоморфни.

Теорема 5.5. *За всяко изброимо породено разделяне γ съществува редица от крайни разделяния $\gamma_n < \gamma_{n+1}$, такава че $\gamma = \sup_n \gamma_n$.*

Доказателство: Нека $\mathfrak{N} = \{A_1, A_2, \dots\}$ са събитията пораждащи γ. Да означим с γ_n разделянето и, съответно, $\mathfrak{A}_n$ - буловата му алгебра, породени от първите n събития. Да означим с $\mathfrak{A} = \sigma(\mathfrak{N})$. Очевидно, $\mathfrak{A} = \sigma(\cup \mathfrak{A}_n)$. Тъй като елементите на γ са сечения на намаляващи елементи на γ_n и за него това е изпълнено. $\square$

До тук работихме само с измеримото пространство $(\Omega, \mathfrak{A})$ без фиксирана вероятност на него.

5.3 Пространства на Лебег - Рохлин

Нека сега разгледаме $(\Omega, \mathfrak{A}, \mathsf{P})$ след въвеждането на фиксирана вероятност.

Определение 5.6. *Нека σ-алгебрата $\mathfrak{A}$ е породена от някакво изброимо семейство множества, което разделя точките във вероятностното пространство Ω. Такива пространства наричаме пространства на Лебег - Рохлин (ЛР-пространства).*

Сега всички σ-подалгебри се попълват с подмножествата на събития с нулева вероятност и се разглеждат само класовете еквивалентни събития.

Определение 5.7. *Нека γ и δ са разделяния, а $\mathfrak{A}_\gamma$ и $\mathfrak{A}_\delta$ – съответните им σ-алгебри. Казваме, че $\gamma < \delta$, когато $\forall A \in \mathfrak{A}_\gamma$ $\exists B \in \mathfrak{A}_\delta$ такова, че $\mathsf{P}(A\Delta B) = 0$. Тук $A\Delta B = A\overline{B} + B\overline{A}$.*

От тук нататък ще разглеждаме само класове еквивалентни алгебри и съответните им разделяния. Теорема 5.4 може лесно да се обобщи.

Теорема 5.6. *Съществува взаимно - еднозначно съответствие между изброимо породените σ-подалгебри на $\mathfrak{A}$ и изброимо породени разделяния на Ω.*

Доказателство: Директно следствие от теорема 8.1. □

Винаги ще предполагаме, че са изпълнени условията за изброима породеност на $\mathfrak{A}$ и всички разглеждани подалгебри. Във пространството R^n, например, това е изпълнено, защото бореловата алгебра е изброимо породена. Така получаваме ключовата за статистиката теорема:

Теорема 5.7. *Нека ξ е векторна сл.в. и P_ξ е нейното разпределение в R^n, определено на бореловата σ-алгебра $\mathfrak{B}$. Тогава тройката $(R^n, \mathfrak{B}, \mathsf{P}_\xi)$ е ЛР-пространство.*

Определение 5.8. *Ще казваме, че γ е измеримо разделяне във вероятностното пространство $(\Omega, \mathfrak{A}, \mathsf{P})$, ако съществуват изброим брой $A_i, i = 1, 2, \ldots$ пораждащи го събития.*

Пример 5.3. *Разделянето на точки ε в интервала $[0, 1]$ се поражда например от двоично - рационалните интервали, които са изброимо много.*

5.4 Условни вероятностни пространства

Нека γ е измеримо разделяне в ЛР-пространството $(\Omega, \mathfrak{A}, \mathsf{P})$. Да означим $\mathfrak{A}_\gamma$ съответната му алгебра и с P_γ ограничението на вероятноста P върху нея. Да разгледаме разделянето γ като множество от елементите си C. Тогава преобразованието $T : \omega \to C$, съпоставящо на ω елемента C, който го съдържа става измеримо относно породената от $T(A_i)$ σ-алгебра $\mathfrak{A}_\gamma$. Пренасяме вероятността тривиално: $\mathsf{P}(T(A_i)) = \mathsf{P}(A_i)$. Така тройката $(\gamma, \mathfrak{A}_\gamma, \mathsf{P}_\gamma)$ става ЛР-пространство, а преобразованието T се превръща в хомоморфизъм. В литературата по теория на мерките $(\gamma, \mathfrak{A}_\gamma, \mathsf{P}_\gamma)$ се нарича фактор-пространство.

Теорема 5.8. *Нека* γ *е измеримо разделяне в ЛР-пространството* $(\Omega, \mathfrak{A}, \mathsf{P})$*. Тогава*

1. *Върху п.в.(относно* P_γ*) елемент* C *е определено ЛР-пространство* $(C, \mathfrak{A}_C, \mathsf{P}_C)$*.*
2. *Изображението* $C(A) : A \to C \cap A$ *от* $\mathfrak{A}$ *в* $\mathfrak{A}_C$ *е добре дефинирано за п.в.* C*.*
3. *Функцията* $\mathsf{P}_C(C(A))$ *е измерима в* $(\gamma, \mathfrak{A}_\gamma, \mathsf{P}_\gamma)$ $\forall A \in \mathfrak{A}$*.*
4. *Изпълнена е следната формула на (пълната вероятност):*

$$\mathsf{P}(A) = \int_{C\in\gamma} \mathsf{P}_C(C(A))d\,\mathsf{P}_\gamma(C) \tag{5.3}$$

или формалното, но еквивалентно:

$$d\,\mathsf{P}(C,t) = d\,\mathsf{P}_C(t)d\,\mathsf{P}_\gamma(C) \tag{5.4}$$

Доказателство: Тази теорема няма да доказваме. Тя следва от теорията на Рохлин за пространствата на Лебег, изложена подробно в (Г.Самородницкий 1985). $\square$

Глава 6

Условни математически очаквания

В тази лекция ще напомним определенията на условните математически очаквания (у.м.о.) и изведем необходимите ни свойства.

Интензивно ще използуваме понятието ортогонален проектор в хилбертово пространство.

6.1 Пространства случайни величини

Да разгледаме вероятностното пространство $(\Omega, \mathfrak{A}, \mathsf{P})$ и линейното пространство на сл.в. определени на него.

Математическо очакване на $(\Omega, \mathfrak{A}, \mathsf{P})$ може да се определи абстрактно. Първо го определяме за прости (с краен брой стойности) сл.в. Всяка неотрицателна сл.в. ξ^+ може да се представи като граница на монотонно нарастваща редица от прости сл.в. $\xi^+ = \lim_n \xi_n =\uparrow \xi_n$. Следователно, винаги съществува границата $\mathbf{E}\,\xi = \lim \mathbf{E}\,\xi_n$, възможно равна на ∞. Всяка сл.в. ξ може да се представи като сума на две неотрицателни сл.в. - $\xi = \xi^+ - \xi^-$. Така, ако и двете м.о. са крайни, можем спокойно да определим $\mathbf{E}\,\xi = \mathbf{E}\,\xi^+ - \mathbf{E}\,\xi^-$.

Класовете еквивалентни с вероятност 1 сл.в. (к.е.сл.в.), зададени на дадено фиксирано вероятностно пространство $(\Omega, \mathfrak{A}, \mathsf{P})$ образуват линейно пространство. Нека го означим с $\mathfrak{M} = \mathfrak{M}(\Omega, \mathfrak{A}, \mathsf{P})$. Да разгледаме подмножествата $\mathbf{L}^r \subset \mathfrak{M}, (1 \le r \le \infty)$ от сл.в. с краен r-ти абсолютен момент. Да напомним известното неравенство за моментите

$$(\mathbf{E}\,|\xi|^r)^{1/r} \le (\mathbf{E}\,|\xi|^k)^{1/k}, \quad \text{при всички } r < k. \tag{6.1}$$

Тъй като в $\mathbf{L}^r \subset \mathfrak{M}$ е в сила същото отношение на еквивалентност, к.е. сл.в. от съответните множества образуват пълни линейни нормирани пространства, ако за норма във всяко от тях служи: $||\xi||_r = (\mathbf{E}\,|\xi|^r)^{1/r}$. При това от неравенството (6.1) следва включването:

$$\mathbf{L}^\infty \subset \mathbf{L}^r \subset \mathbf{L}^p \subset \mathbf{L}^1 \subset \mathfrak{M}, \quad 1 < p < r < \infty.$$

Особен интерес представлява пространството $\mathbf{L}^2$ от сл.в. с крайна дисперсия. То е хилбертово и скаларното произведение в него е $(\xi, \eta) = \mathbf{E}\,\xi\eta$. Така центрираните сл.в., ако са независими, са ортогонални в $\mathbf{L}^2$. Ортогоналните сл.в., разбира се, може и да не са независими.

Когато вероятностното пространство е ЛР-пространство, $\mathbf{L}^2$ е сепарабелно, т.е. има изброим базис. В тези лекции ние винаги ще предполагаме, че това е изпълнено.

6.2 Условни м.о.

Нека $\mathfrak{F} \subset \mathfrak{A}$ и разделянето, което ѝ съответствува е γ. Ще предполагаме, че γ (и следователно $\mathfrak{F}$) е изброимо породено. Да разгледаме подмножеството $\mathbf{L}^2_\gamma$ на $\mathbf{L}^2$ от сл.в.ξ, такива, че техните лебегови множества $\{\xi < x\} \in \mathfrak{F}$. Те очевидно образуват линейно подпространство на $\mathbf{L}^2$. Проекторът върху това подпространство да означим с $\mathbf{E}_\gamma$. Ще го наричаме условно м.о. при условие разделянето γ, или σ-алгебрата, която му съответствува.

Теорема 6.1. *Пресмятането на условното м.о. може да се извърши по следния начин:*

$$E_\gamma(\xi) = \int_{\omega \in C} \xi(\omega) d\,\mathsf{P}_C(\omega), \quad \omega \in C \tag{6.2}$$

Доказателство: 1. Достатъчно е да проверим за крайни разделяния γ_n, че горната формула дава проектор. Нека $\gamma_n = \{H_1, H_2, \ldots, H_n\}$ е крайно разделяне. Тогава условните вероятностни пространства върху всеки елемент на разделянето са определени еднозначно и формула 6.2 действително дава като резултат елемент от $\mathbf{L}^2_\gamma$.

2. За произволно разделяне γ можем да построим редицата $\gamma_n \to \gamma$ монотонно нарастваща и да се възползуваме от сходимостта в $\mathbf{L}^2$ дадена от теорема 4.5. □

В частност, ако означим с ε разделянето, което съответствува на $\mathfrak{A}$ и $\nu = \{\varnothing, \Omega\}$, то $\mathbf{E}_\varepsilon = I$, $\mathbf{E}_\nu = \mathbf{E}$. Условното м.о. на всяка сл.в. ξ при условие крайното разделяне γ е проста сл.в. с константни стойности върху елементите му.

Теорема 6.2. *Нека $\gamma < \delta$ са прости и $\xi \in \mathbf{L}^2_\gamma$, $\eta \in \mathbf{L}^2_\delta$, $\xi * \eta, \theta \in \mathbf{L}^2$. Изпълнени са следните свойства:*

1. *проекторите комутират:* $\mathbf{E}_\delta \mathbf{E}_\gamma \theta = \mathbf{E}_\gamma \mathbf{E}_\delta \theta = \mathbf{E}_\gamma \theta$.

2. $\mathbf{E}_\gamma \xi\eta = \xi \mathbf{E}_\gamma \eta$. $\mathbf{D}_\gamma \xi\eta = \xi^2 \mathbf{D}_\gamma \eta$.

Доказателство: Достатъчно е да ги проверим с елементарото определение на у.м.о.□

Да отбележим, че докато първото твърдение следва от елементарните свойства на проекторите, то второто няма аналог в Хилбертовите пространства - там изобщо казано не е определено произведение на елементи. Това показва удобството от паралелното използуване на двата подхода.

Теорема 6.3. *Условното м.о. е ортогонален проектор в хилбертовото пространство $\mathbf{L}^2$.*

Доказателство: Ще използуваме очевидното свойство: $\mathbf{E} = \mathbf{E}\,\mathbf{E}_\gamma$ и че всяка γ-измерима сл.в. за $\mathbf{E}_\gamma$ е равносилна на константа.

$$(\mathbf{E}_\gamma\xi, \eta) = \mathbf{E}\,\mathbf{E}_\gamma(\mathbf{E}_\gamma\xi)\eta = \mathbf{E}\,(\mathbf{E}_\gamma\eta)(\mathbf{E}_\gamma\xi) = \mathbf{E}\,(\mathbf{E}_\gamma\xi)(\mathbf{E}_\gamma\eta) = \mathbf{E}\,\xi(\mathbf{E}_\gamma\eta) = (\xi, \mathbf{E}_\gamma\eta).$$

С други думи проекторът $\mathbf{E}_\gamma$ е самоспрегнат - значи ортогонален. □

Така от двете представяния на у.м.о. можем да извлечем всички необходими свойства. Ще се опитаме да ги подредим в следните теореми.

Теорема 6.4. *Нека γ е измеримо разделяне и $0 \le \xi \le \eta \in \mathbf{L}^2$. Изпълнени са следните свойства:*

1. *(монотонност)* $0 \le \mathbf{E}_\gamma\xi \le \mathbf{E}_\gamma\eta$.
2. *(неравенство на Чебишов)* $\mathbf{E}_\gamma I_{\{\xi>\varepsilon\}} \le \mathbf{E}_\gamma\xi/\varepsilon$.
3. *(неравенство за моментите, когато съществуват)*

$$(\mathbf{E}_\gamma|\xi|^r)^{1/r} \le (\mathbf{E}_\gamma|\xi|^k)^{1/k}, \quad \forall 0 < r < k. \tag{6.3}$$

Доказателство: Достатъчно е да го проверим за прости сл.в. и да използуваме представянето (6.2). □

Определение 6.1. *Да означим с $\mathbf{D}_\gamma\xi = \mathbf{E}_\gamma(\xi - \mathbf{E}_\gamma\xi)^2$ условната дисперсия на ξ.*

Теорема 6.5. *Нека $\gamma \perp\!\!\!\perp \delta$ и $\xi \in \mathbf{L}^2_\gamma$, $\eta \in \mathbf{L}^2_\delta$, $\theta \in \mathbf{L}^2$. Изпълнени са следните свойства:*

1. $\mathbf{E}_\gamma\eta = \mathbf{E}\,\eta$ *и* $\mathbf{E}_\delta\xi = \mathbf{E}\,\xi$.
2. *сл.в.* $\mathbf{E}_\gamma\theta$ *и* $\mathbf{E}_\delta\theta$ *са независими;*
3. *проекторите комутират:* $\mathbf{E}_\delta\mathbf{E}_\gamma\theta = \mathbf{E}_\gamma\mathbf{E}_\delta\theta = \mathbf{E}\,\theta$.
4. $\mathbf{E}_\gamma\theta - \mathbf{E}\,\theta$ *и* $\mathbf{E}_\delta\theta - \mathbf{E}\,\theta$ *са ортогонални в* $\mathbf{L}^2$, *т.е. с нулева корелация.*
5. $\mathbf{D}\theta \ge \mathbf{D}(\mathbf{E}_\gamma\theta) + \mathbf{D}(\mathbf{E}_\delta\theta)$*;*

Доказателство: Достатъчно е да го проверим за крайни разделяния. □

Теорема 6.6. *Нека γ_n е нарастваща редица от измерими разделяния и $\gamma = \sup_n \gamma_n$. Тогава за всяка сл.в. $\xi \in \mathbf{L}^2$ редицата $\mathbf{E}_{\gamma_n}\xi$ е сходяща по норма в $\mathbf{L}^2$ към сл.в. $\mathbf{E}_\gamma\xi$.*

Доказателство: Директно следствие от теореми 4.5 и 6.3. □

Да отбележим, че тази теорема е най-слабия аналог на така наречената мартингална теорема защото редицата $\xi_n = \mathbf{E}_{\gamma_n}\xi$ е мартингал.

6.3 Коментарии

Накратко изложената теория на у.м.о. показва колко удобен апарат са те. Нищо чудно, че много автори предпочитат аксиоматиката им пред тази на теория на мярката за целите на образованието по вероятности (виж.(Хог 1978)). В крайна сметка и двата подхода си имат своите преимущества и недостатъци.

Глава 7

Изпъкнали множества

В тази лекция

- ще напомним определението на изпъкнало множество,
- ще покажем един вариант на теоремата на Шоке

7.1 Изпъкнали множества

Тук ще приведем няколко теореми от функционалния анализ (Рудин 1975).

Определение 7.1. *Казваме, че множеството C в линейното пространство X е изпъкнало, ако* $\forall x, y \in C, 0 < \alpha < 1 : \quad \alpha x + (1 - \alpha) y \in C$.

Определение 7.2. *За всяко множеството $A \subset X$ определяме изпъкнала обвивка $conv(A)$, като най-малкото изпъкнало множество $\supset A$.*

Определението е коректно, понеже сечението на произволно семейство изпъкнали множества е изпъкнало. Нека множеството C в линейното пространство X е изпъкнало.

Определение 7.3. *Казваме, че точката z е крайна в C, ако не съществуват $x, y \in C, x \neq y, 0 < \alpha < 1$ такива, че $z = \alpha x + (1 - \alpha) y$.*

Нека C е изпъкнал компакт в X. Да означим с $\partial(C)$ множеството от крайните точки на C.

7.2 Теорема на Крейн-Милман

Теорема 7.1 (Теорема на Крейн-Милман). *Ако X е локално изпъкнало, то $C \subset cl(conv(\partial(C)))$ (C е подмножество на затворената изпъкнала обвивка на $\partial(C)$.)*

С други думи на всяка точка $x \in C$ може да бъде съпоставена мярка μ на $\partial(C)$ такава, че $x = \int_{\partial(C)} y d\mu(y)$.

Пример 7.1. *Единичният интервал* $[0, 1]$ *притежава две крайни точки:* $\{0\}, \{1\}$. *Всяка точка* x *от него еднозначно се представя в вида* $(1 - \alpha)\{0\} + \alpha\{1\}$.

Пример 7.2. *Крайните точки* $\partial(C)$ *на единичното кълбо* $C = \{||x|| \leq 1\} \subset R^p$ *са точките от единичната сфера* $S = \{||x|| = 1\}$. *Всяка точка* $z \in C$ *по много различни начини може да се представи в вида* $(1 - \alpha)x + \alpha y, x, y \in \partial(C)$.

Разбира се, от теорема 7.1 не следва че представянето $x = \int_{\partial(C)} y d\mu(y)$ е единствено, нито пък следва, че $x \in C, \forall \mu$. С втората неприятност се справяме лесно като въведем допълнително изискване за топологията в X.

Определение 7.4. *Казваме, че линейното пространство* X *е пространство на Фреше, ако е локално изпъкнало и топологията в* X *е породена от пълна инвариантна метрика.*

Пространствата на Фреше са удобни, защото затворената изпъкналата обвивка на компакт е компакт. Така получаваме, че за изпъкналия компакт C имаме равенство: $C = cl(conv(\partial(C))) = conv(\partial(C))$.

Пример 7.3. *Да разгледаме множеството* M *от неотрицателни мерки върху измеримото пространство* X.

Крайните точки $\partial(M)$ на множеството $M = \{\mu(X) \leq 1\}$ са атомарните вероятностни мерки $\delta_x, x \in X\}$ и нулевата мярка. Всяка точка $\mu \in M$ по единствен начин може да се представи в вида $\mu = \int_{\partial(M)} \delta_y d\mu(y)$. Линейното пространство (от знакопроменливи мерки) породено от M става пространство на Фреше, ако вземем метриката на Леви (или друга метризираща слабата топология).

Хубавото на понятието крайна точка е неговото чисто геометрично определение.

Лема 7.1. *Нека* X, Y *са векторни пространства, нека* T *е обратим линеен оператор от* X *на* Y. *Тогава* $\partial(TC) = T\partial(C)$.

Доказателство: Трябва само да проверим определение 7.3 на крайна точка в Y. Поради линейността $Tz = \alpha Tx + (1 - \alpha)Ty$. Поради обратимостта, ако $x \neq y$,то $Tx \neq Tz$. Следователно "вътрешните" точки на C преминават във вътрешни на Tx, а крайните – в крайни. $\square$

7.3 Теорема на Шоке

Следната теорема ни дава общите условия за еднозначност на интегралното представяне.

Определение 7.5. *Казваме, че компактът C е симплекс на Шоке ако, когато вложим множеството $C \times \{1\}$ в линейното пространство $X \times R^1$ и образуваме положителният конус $\mathbf{P} = \{\alpha x, x \in C \times \{1\}, \alpha \geq 0\}$, множеството от разлики $\mathbf{R} = \{x - y, x, y \in P\}$ става решетка (относно частичната наредба породена от P), т.е. съществуват единствени* $\min$, $\max$ *за всеки два елемента на* $\mathbf{R}$.

Пример 7.4. *В крайномерното пространство R^p симплексът на Шоке съвпада с обикновения симплекс.*

Крайните му точки $\partial(C)$ са точно $p+1$ (ако не лежи в подпространство). Всяка точка $z \in C$ по единствен начини може да се представи в вида на претеглена линейна комбинация: $z = \sum_{y \in \partial(C)} w(y)y$, $\sum_{y \in \partial(C)} w(y) = 1, w(y) \geq 0$.

Пример 7.5. *Да разгледаме множеството $\mathbf{M}$ на всички вероятностни мерки върху Бореловата σ-алгебра на единичния интервал $(0, 1)$. Множеството е очевидно изпъкнало.*

За да ги направим симплекс на Шоке нека го вложим в линейното пространство на всички крайни знакопроменливи мерки. Това множество е решетка по отношение на конуса на положителните мерки $\lambda\mu, \mu \in \mathbf{M}, \lambda > 0$.

Проверете аксиомите за $sup(\mu, \nu)$ като използувате теоремата на Жордан-Хан 2.1.

Следователно, $\mathbf{M}$ е симплекс на Шоке.

Теорема 7.2 (Теорема на Шоке). *Необходимо и достатъчно условие интегралното представяне на Крейн-Милман на C да е единствено, е условието C да е симплекс на Шоке.*

Доказателство:

$\square$

Глава 8

Теорема на Бохнер

8.1 Характеристични функции

Определение 8.1. *Наричаме характеристична функция на вероятностната мярка μ в R^p функцията:*

$$f(t) = \int_{R^p} e^{it'x} d\mu(x). \tag{8.1}$$

Определение 8.2. *Казваме, че комплексната функция на R^p е положително определена (п.о.ф.), ако за всяко $n > 0$, за всеки набор от комплексни числа a_i и вектори t_i $(i = 1, 2, \dots, n)$ е изпълнено неравенството:*

$$\sum_{i=1}^{n} \sum_{k=1}^{m} a_i \bar{a}_k f(t_i - t_k) \geq 0. \tag{8.2}$$

Лема 8.1. *Всяка положително определена функция (п.о.ф.) има следните свойства:*

1. $f(-t) = \bar{f(t)}$;
2. $f(0) \geq |f(t)| \geq 0$;
3. $|f(t) - f(s)| \leq 2f(0).Re(f(0) - f(t-s))$.

Доказателство: (Виж (Круглов 1984)) Като положим в (8.2) $n = 1, a_1 = 1$ получаваме $f(0) \geq 0$. Като положим в (8.2) $n = 2, a_1 = 1, a_2 = c, t_1 = 0, t_2 = t$ получаваме

$$(1 + |c|^2)f(0) + cf(t) + \bar{c}f(-t) \geq 0. \tag{8.3}$$

1. Като положим в (8.3) отначало $c = 1$, а после $c = i$, получаваме, че $f(t) + f(-t)$ и $i(f(t) - f(-t))$ са реални. Това може да стане само ако $f(-t) = \bar{f(t)}$.
2. Като положим в (8.3) $c = |f(t)|/f(t)$, получаваме $f(0) \geq |f(t)|$.

3. Нека $f(t) \neq f(s)$. Да положим в (8.2) $n = 3$, $a_1 = 1, a_2 = c, a_3 = -c$, $t_1 = 0, t_2 = t, t_3 = s$. Получаваме

$$c = -\lambda \frac{|f(t) - f(s)|}{f(t) - f(s)},$$

λ - реално. Тогава (8.2) става

$$f(0)(1 + 2\lambda^2) + 2\lambda|f(t) - f(s)| - 2\lambda^2 Re(f(t - s)) \geq 0.$$

Това може да е изпълнено само ако дискриминантата е положителна. □

Така, ако положително определената функция е непрекъсната в нулата, то тя е равномерно непрекъсната на R^p.

8.2 Теорема на Бохнер

Теорема 8.1 (Теорема на Бохнер). *Положително определените непрекъснати в нулата функции, и такива, че $f(0) = 1$, се представят във вида (8.1), където μ е вероятностна мярка в R^p.*

Доказателство: Тук даваме само пътя на доказателството.

1. За да използуваме теоремата на Крейн-Милман се нуждаем да поместим множеството C от положително-определени функции в подходящо пространство. Нека това е множеството X от ограничени равномерно - непрекъснати комплексни функции зададени на R^p. Ще разширим C като добавим всички функции с $0 \leq f(0) \leq 1$.

2. Нека фиксираме $L < \infty$ и разгледаме в X нормата: $||f|| = \sup_{||t|| \leq L} |f(t)|$.

3. Лесно се проверява, че пространството е локално изпъкнало и C е изпъкнал компакт в тази норма. Положителният конус породен от C в X са п.о.ф., при които $f(0) \geq 0$.

4. Крайни точки на компакта C се оказват функциите: $f(t) \equiv 0$ и $f_x(t) = e^{it'x}, x \in R^p$. Трудното е, да се докаже, че други няма.

5. C е симплекс на Шоке и представянето е единствено. □

8.3 Мързеливо доказателство

Съществува и по-лесен път за доказателство на тази теорема. Нека се възползуваме от лема 7.1. Да означим с T преобразованието на Фурие:

$$f(t) \stackrel{def}{=} (TF)(t) = \int_{R^p} e^{it'x} dF(x)$$

от множеството на функциите с ограничена вариация $\mathcal{F}$ в множеството на непрекъснати функции. То е линейно и (съгласно формулата за обръщане) взаимно еднозначно.

Да разгледаме подмножеството $K \subset \mathcal{F}$, състоящо се от неотрицателните и ограничени от 1 мерки. Те образуват компакт в слабата топология на $\mathcal{F}$. Крайните точки на този компакт са атомарните мерки и нулевата мярка. Преобразованието T е взаимно еднозначно и ги прехвърля в положително определени функции такива, че $0 \leq f(0) \leq 1$.

Няколко заключителни думи.

Глава 9

Теорема на Леви

В тази тема ще разглеждаме сл.в. и разпределения в крайно мерното Евклидово пространство R^p.

Изложението следва (Вандев 1978) и (Линник и Островский 1972), Приложение 1.

9.1 Безгранично делими разпределения

Определение 9.1. *Казваме, че сл.в. има безгранично делимо разпределение, ако* $\forall n$ *може да се представи като сума на* n *независими еднакво разпределени сл.в.*

Определение 9.2. *Казваме,че характеристичната функция* $\psi(t)$ *е безгранично делима (б.д.х.ф.), ако* $\forall n$ *съществува х.ф.* $f_n(t)$ *такава, че* $\psi(t) = (f_n(t))^n$.

Еквивалентността на двете определения е очевидна.

Пример 9.1. *Атомарната мярка* δ_a *в точката* a *безгранично делима.*

Наистина, $\delta_a = \sum_{k=1}^{n} \delta_{a/n}$. Това формално равенство разбираме като равенство на съответните "сл."в.

Пример 9.2. *Всяка Гаусова сл.в. е безгранично делима.*

Характеристичната функция на многомерната гаусова сл.в. $N(\mu, \Sigma)$ е:

$$f(t, \mu, \Sigma) = \exp(it'\mu - \frac{1}{2}t'\Sigma t).$$

Следователно, нейната $f^{1/n}$ ще бъде също гаусова.

Теорема 9.1. *Основните свойства на б.д.х.ф. можем да сумираме така:*

1. $\forall t$ *б.д.х.ф.* $\varphi(t) \neq 0$;

2. *Произведение на краен брой б.д.х.ф. е б.д.х.ф.;*

3. $\forall \tau > 0$ $(\varphi(t))^\tau$ *е б.д.х.ф.;*

4. Ако $\psi(t)$ *е х.ф., то* $\exp(\psi(t) - 1)$ *е б.д.х.ф.*

Доказателство: 1.$|\varphi(t)|^2$ е също х.ф. $|\varphi(0)|^2 = 1$. Поради равномерната непрекъснатост: $\forall t |\varphi(t)|^{2/n} \longrightarrow 1$;
2. $\varphi(t)\psi(t) = (\varphi(t)^{(1/n)}\psi(t)^{(1/n)})^n$;
3. Ако $m/n \to \tau$, $\varphi(t)^{(m/n)} \to \varphi(t)^\tau$;
4. Следва от това $(1/n)\psi(t)$ е о.х.ф. □

Да означим с Q множеството от логаритми на б.д.х.ф. Те са добре определени, защото можем да изберем това листо, за което $\ln \varphi(0) = 0$.

Теорема 9.2. *Нека* $\varphi, \psi \in Q$. *Тогава:*

1. $\varphi(0) = \psi(0) = 0$*;*

2. $\varphi(-t) = \overline{\varphi(t)}$*;*

3. φ, ψ *са непрекъснати;*

4. ако $\alpha, \beta > 0$, *то* $\alpha\varphi + \beta\psi \in Q$ *(*Q *е конус);*

5. $\chi_a(t) = \varphi(t) - 1/2(\varphi(t-a) + \varphi(t+a))$ *е п.о.ф.*

Доказателство: 1.
□

9.2 Теорема на Леви

Тази теорема ще формулираме и докажем в едномерния случай. Доказателството, което предлагаме не върви в многомерния случай защото точката нула, съответстваща на дисперсията на нормалното разпределение, се разпада. .

Теорема 9.3 (Теорема на Леви). *Всяка б.д.х.ф.може да бъде представена във формата*

$$\varphi(u) = i\gamma_1 u + \int_{-\infty}^{\infty} (e^{itu} - 1 - itu\frac{2(1-\cos t)}{t^2})(\frac{t - \sin t}{t})^{-1} dF_1(t) \tag{9.1}$$

или във формата

$$\varphi(u) = i\gamma_2 u + \int_{-\infty}^{\infty} (e^{itu} - 1 - \frac{itu}{1+t^2})(\frac{1+t^2}{t^2}) dF_2(t) \tag{9.2}$$

Тук γ_i *са реални константи, а* $F_i(t)$ *- о.ф.р.*

Ще формулираме и докажем последователно няколко твърдения, от които лесно следват теореми 9.1 и 9.2. Те имат и самостоятелно значение.

Нека означим с $\widetilde{Q}$ подконуса на Q определено от условието

$$Im \int_0^1 \varphi(u)du = 0,$$

а с $\widetilde{K}$ изпъкналото подмножество на $\widetilde{Q}$ определено от условието:

$$-1 \le Re \int_0^1 \varphi(u)du \le 0.$$

Да означим с A преобразованието:

$$A\varphi(u) = \varphi(u) - \frac{1}{2}\int_{-1}^{1} \varphi(u+\alpha)d\alpha. \tag{9.3}$$

Лема 9.1. *Операторът A е линеен изоморфизъм на множеството $\widetilde{K}$ и множеството от о.х.ф. K, определено от условието $0 \le f(0) \le 1$.*

Доказателство: Достатъчна е проста проверка на линейността. След това трябва да се докаже, че интегралното уравнение $A\varphi(u) = 0$ допуска за решения в класа на непрекъснатите функции само линейни функции. Това става лесно с преобразование на Фурие. □

Върху множеството от о.х.ф. K съществува естествена топология, (породената от теоремата на Бохнер,например), която го прави компакт. С помощта на изоморфизма A тя се прехвърля в $\widetilde{K}$, който също става компакт.

Лема 9.2. *Крайни точки на компакта $\widetilde{K}$ са функциите:*

1. $\varphi(u) = -1$;
2. $\varphi_0(u) = -3u^2, \quad \gamma = 0$;
3. $\varphi_\gamma(u) = (e^{i\gamma u} - 1 - i\gamma u\frac{2(1-\cos\gamma)}{\gamma^2})(1 - \frac{\sin\gamma}{\gamma})^{-1}, \quad \gamma \neq 0.$

Доказателство: Проста проверка. Това, че това са всички следва от свойствата на линейния изоморфизъм.□

Лема 9.3. *За всяко $\varphi \in \widetilde{K}$ е изпълнена формулата 9.1, където константата γ_1 и функцията $F_1(t)$ се определят от условията:*

$$\gamma_1 = 2Im \int_0^1 \varphi(u)du, \tag{9.4}$$

$$F_1(b) - F_1(a) = \lim_{T\to\infty} \frac{1}{2\pi}\int_{-T}^{T} \frac{e^{ita} - e^{itb}}{it}(A\varphi)(t)dt. \tag{9.5}$$

Доказателство: □

Няколко заключителни думи.

Литература

Вандев, Д. (1978). Доказательство формулы Леви–Хинчина при помощи теорем Бохнера и Крейна–Милмана. *C.R.Acad.Bulgare Sci. 31*(4), 385–387. 9

Г.Самородницкий (1985). *Теория меры.* Киев: Вища школа. 5.4

Круглов, В. (1984). *Дополнительные главы теорий вероятностей.* Москва: Высшая школа. 8.1

Линник, Ю. В. и И. В. Островский (1972). *Разложения случайных величин и векторов.* Теория вероятностей и математическая статистика. Москва: Наука. 9

Обретенов, А. и Д. Хаджиев (1983). *Теория на вероятностите /Втора част/.* София: Унивеситетско издателство. 1.4, 2.2

Обретенов, А. (1974). *Теория на вероятностите.* София: Наука и изкуство. 1.4

Партасарати, К. (1983). *Введение в теорию вероятностей и теорию меры.* Москва: МИР. 1.4

Рудин, У. (1975). *Функциональный анализ.* Москва: МИР. 7.1

Хог, К. (1978). *Увод в теория на вероятностите.* София: Наука и искуство. 6.3