Лекция 14

ОСНОВНІ ПОНЯТТЯ ТЕОРІЇ ЙМОВІРНОСТЕЙ

1. σ-АЛГЕБРА

Нехай Ω — довільна множина. Система $\mathscr{F}$ її підмножин називається *σ-алгеброю*, якщо виконуються наступні три умови:

(σ_1) $\Omega \in \mathscr{F}$;
(σ_2) якщо $A \in \mathscr{F}$, то $\overline{A} \in \mathscr{F}$;
(σ_3) якщо $A_n \in \mathscr{F}$ для будь-якого $n \geq 1$, то $\bigcup_n A_n \in \mathscr{F}$.

1.1. Тривіальна σ-алгебра. Нехай $\mathscr{F}$ складається лише з двох елементів: $\mathscr{F} = \{\varnothing, \Omega\}$. Тоді $\mathscr{F}$ є σ-алгеброю, яка називається *тривіальною.* Всі три властивості (σ_1)–(σ_3) є очевидними в цьому випадку.

1.2. Найпотужніша σ-алгебра. Нехай $\mathscr{F}$ складається з усіх підмножин Ω включно з Ω та $\varnothing$. Зрозуміло, що така сукупність є σ-алгеброю, яка називається *найпотужнішою.* Найпотужніша σ-алгебра позначається 2^Ω.

Задача 1. *Як можна пояснити позначення 2^Ω для найпотужнішої σ-алгебри?*

Приклад 1. Нехай множина Ω складається з двох елементів, $\Omega = \{\omega_1, \omega_2\}$. В цьому випадку існують тільки дві σ-алгебри, тривіальна та найпотужніша. Крім них жодної іншої побудувати неможливо.

Задача 2. *Чому для $\Omega = \{\omega_1, \omega_2\}$ можна побудувати тільки дві σ-алгебри?*

Приклад 2. Нехай множина Ω складається з трьох різних елементів, тобто $\Omega = \{\omega_1, \omega_2, \omega_3\}$. Будемо позначати через A_i підмножину, яка складається з одного елемента ω_i, а через A_{ij} — підмножину, яка складається з двох елементів ω_i та ω_j. В цьому випадку, крім тривіальної та найпотужнішої σ-алгебр, можна побудувати і інші, наприклад $\mathscr{F} = \{\varnothing, \Omega, A_1, A_{23}\}$.

Задача 3. *Скільки різних σ-алгебр можна побудувати у випадку $\Omega = \{\omega_1, \omega_2, \omega_3\}$?*

Приклад 3. Нехай Ω — це інтервал $[0,1]$. Розглянемо сукупність інтервалів $\mathscr{A} = \{[a,b),\ 0 \leq a < b \leq 1\}$. Ця сукупність не є σ-алгеброю, оскільки $\Omega \notin \mathscr{A}$. Розширемо $\mathscr{A}$ до сукупності $\mathscr{A}_1 = \{\langle a,b\rangle,\ 0 \leq a < b \leq 1\}$, яка включає всі можливі інтервали, тобто $\langle a,b\rangle = [a,b), [a,b], (a,b), (a,b]$. Ця сукупність містить Ω, але все рівно не є σ-алгеброю, оскільки для $A = [0.25, 0.75)$

$$\overline{A} = [0, 0.25) \cup [0.75, 1) \notin \mathscr{A}_1.$$

Сукупність $\mathscr{A}_1$ можна розширити до $\mathscr{A}_2$, додавши підмножини вигляду $\langle a,b\rangle \cup \langle c,d\rangle$. Але і сукупність $\mathscr{A}_2$ не є σ-алгеброю.

[0] Printed from the file [14-random_variables.tex] on 9.6.2016

Задача 4. *Чому $\mathscr{A}_2$ не є σ-алгеброю?*

Процедуру розширення можна продовжити, додавши до $\mathscr{A}_2$ підмножини вигляду $\langle a,b\rangle \cup \langle c,d\rangle \cup \langle e,f\rangle$. Позначимо отриману сукупність через $\mathscr{A}_3$. Але і $\mathscr{A}_3$ не є σ-алгеброю. Більше того, при будь-якому n ми можемо означити сукупність $\mathscr{A}_n$, додавши до $\mathscr{A}_{n-1}$ підмножини вигляду $\langle a_1,b_1\rangle \cup \cdots \cup \langle a_n,b_n\rangle$. Жодна з сукупностей $\mathscr{A}_n$ не буде σ-алгеброю.

Задача 5. *Чому $\mathscr{A}_n$ не є σ-алгеброю для будь-якого $n \geq 1$?*

Більше того, об'єднання всіх $\mathscr{A}_n$ не є σ-алгеброю!

Задача 6. *Чому $\bigcup_{n\geq 1} \mathscr{A}_n$ не є σ-алгеброю?*

1.3. Борелівська σ-алгебра. Чи існує хоча б одна σ-алгебра, яка містить $\mathscr{A}_1$? Так, існує. Такою є, наприклад, найпотужніша σ-алгебра 2^Ω. А чи існує менш потужна σ-алгебра, ніж 2^Ω, яка тим не менш містить $\mathscr{A}_1$? І на це питання відповідь є ствердною. Ми опишемо спосіб побудови найменш потужної σ-алгебри, яка містить $\mathscr{A}_1$. Для цього ми скористаємось наступним твердженням.

Теорема 1. *Нехай $\mathscr{F}_\theta$ є σ-алгеброю для будь-якого θ, де θ — це деякий параметр (абсолютно довільної природи). Тоді перетин цих σ-алгебр*

$$\mathscr{F} \stackrel{\text{def}}{=} \bigcap_\theta \mathscr{F}_\theta$$

також є σ-алгеброю.

Доведення. Оскільки $\Omega \in \mathscr{F}_\theta$ для будь-якого θ, то $\Omega \in \mathscr{F}$, що доводить властивість (σ_1).

Нехай тепер $A \in \mathscr{F}$. Це означає, що $A \in \mathscr{F}_\theta$ для будь-якого θ. Звідси випливає, що $\overline{A} \in \mathscr{F}_\theta$ для будь-якого θ, тобто $\overline{A} \in \mathscr{F}$, що доводить властивість (σ_2).

Нехай нарешті $A_n \in \mathscr{F}$, $n \geq 1$. Це означає, що $A_n \in \mathscr{F}_\theta$ для будь-якого θ, звідки отримуємо $\bigcup_n A_n \in \mathscr{F}_\theta$ для будь-якого θ. Зрозуміло, що це означає $\bigcup_n A_n \in \mathscr{F}$, що доводить властивість (σ_3). $\square$

Означення 1. Нехай $\mathscr{A}$ — це сукупність інтервалів, означена у прикладі 3. *Борелевською σ-алгеброю $\mathscr{B}$* на інтервалі $[0,1]$ називається найменш потужна σ-алгебра, яка містить $\mathscr{A}$.

Борелівська σ-алгебра $\mathscr{B}$ будується наступним чином. Розглянемо всілякі σ-алгебри $\mathscr{F}_\theta$, які містять $\mathscr{A}$. Для зручності кожній такій σ-алгебрі ми дали індекс θ, якій не впливає на міркування. Тоді

$$\mathscr{B} = \bigcap_\theta \mathscr{F}_\theta. \tag{1}$$

Зауваження 1. В перетині у правій частині (1) приймає участь принаймні одна σ-алгебра (найпотужніша), тому результат цієї операції визначений.

Зауваження 2. Результат перетину $\mathscr{B}$ у рівності (1) є σ-алгеброю на підставі теореми 1.

Зауваження 3. Будь-який інтервал $[a,b)$ є елементом $\mathscr{B}$.

Задача 7. *Довести, що* $\mathscr{A}_n \in \mathscr{B}$ *для будь-якого* $n \geq 1$*, де сукупність* $\mathscr{A}_n$ *визначено у прикладі* 3.

Задача 8. *Довести, що* $\mathscr{A}_n \neq \mathscr{B}$ *для будь-якого* $n \geq 1$*, де сукупність* $\mathscr{A}_n$ *визначено у прикладі* 3.

Задача 9. *Довести, що* $\bigcup_n \mathscr{A}_n \in \mathscr{B}$*, де сукупність* $\mathscr{A}_n$ *визначено у прикладі* 3.

Задача 10. *Довести, що* $\bigcup_n \mathscr{A}_n \neq \mathscr{B}$*, де сукупність* $\mathscr{A}_n$ *визначено у прикладі* 3.

Зауваження 4. Борелівську σ-алгебру можна визначити на довільному інтервалі, а також на всій числовій прямій. Конструкція є абсолютно ідентичною до розглянутої для інтервалу $[0,1]$.

Теорема 2. *Якщо позначити через* $\mathscr{B}_{[0,1]}$ *борелівську* σ*-алгебру на* $[0,1]$*, а через* $\mathscr{B}_{\mathbf{R}}$ *борелівську* σ*-алгебру на* $\mathbf{R}$*, то* $\mathscr{B}_{[0,1]} \subseteq \mathscr{B}_{\mathbf{R}}$.

Теорема 3. *Нехай* $\mathscr{B}_{\mathbf{R}}$ *— це борелівська* σ*-алгебра на* $\mathbf{R}$*, а* $\mathscr{B}_{[0,1]}$ *— це борелівська* σ*-алгебра на* $[0,1]$*. Якщо* $B \in \mathscr{B}_{\mathbf{R}}$ *та* $B \subseteq [0,1]$*, то* $B \in \mathscr{B}_{[0,1]}$.

Задача 11. *Довести теорему* 3.

2. Ймовірнісна міра

Нехай Ω — довільна множина, а $\mathscr{F}$ — σ-алгебра її підмножин. Функція $\mathsf{P}(A)$, означена для $A \in \mathscr{F}$, називається ***ймовірнісною мірою***, якщо

- (P_1) $\mathsf{P}(A) \geq 0$ для будь-якого $A \in \mathscr{F}$;
- (P_2) $\mathsf{P}(\Omega) = 1$;
- (P_3) якщо $A_n \in \mathscr{F}$ для будь-якого $n \geq 1$ та $A_i \cap A_j = \varnothing$, $i \neq j$, то

$$\mathsf{P}\left(\bigcup_{n\geq 1} A_n\right) = \sum_{n=1}^{\infty} \mathsf{P}(A_n).$$

Приклад 4. Якщо $\mathscr{F}$ — це тривіальна σ-алгебра, то ймовірнісна міра визначена лише для $A = \varnothing$ або $A = \Omega$, причому $\mathsf{P}(\Omega) = 1$ за умовою (P_2). Щоб визначити значення $\mathsf{P}(\varnothing)$, припустимо, що $\mathsf{P}(\varnothing) = r > 0$. Тоді для послідовності $A_1 = A_2 = \cdots = \varnothing$ маємо $A_i \cap A_j = \varnothing$, $i \neq j$, й тому на підставі (P_3)

$$r = \mathsf{P}(\varnothing) = \mathsf{P}\left(\bigcup_{n\geq 1} A_n\right) = \sum_{n=1}^{\infty} \mathsf{P}(A_n) = \sum_{n=1}^{\infty} r.$$

Ряд у правій частині розбігається на підставі необхідної ознаки збіжності й не може дорівнювати значенню у лівій частині. Таким чином, єдиним значенням для $\mathsf{P}(\varnothing)$ є 0.

Приклад 5. Нехай $\Omega = \{\omega_1, \omega_2\}$, а $\mathscr{F} = 2^{\Omega}$. Тоді ймовірнісна міра має такі значення: $\mathsf{P}(\Omega) = 1$, $\mathsf{P}(\varnothing) = 0$. Значення для $\mathsf{P}(\omega_1)$ та $\mathsf{P}(\omega_2)$ не можуть бути довільними на підставі умови (P_3): якщо $\mathsf{P}(\omega_1) = p$, $0 \leq p \leq 1$, то $\mathsf{P}(\omega_2) = 1 - p$.

Задача 12. *Довести, що якщо* $\mathsf{P}(\omega_1) = p$*,* $0 \leq p \leq 1$*, для* $\Omega = \{\omega_1, \omega_2\}$*, то* $\mathsf{P}(\omega_2) = 1 - p$.

3. Ймовірнісний простір

Нехай Ω — довільна множина, $\mathscr{F}$ — σ-алгебра її підмножин, P — ймовірнісна міра на $\mathscr{F}$. Трійка $(\Omega, \mathscr{F}, \mathsf{P})$ називається *ймовірнісним простором.* Ми називаємо Ω *простором елементарних подій.* Елементи Ω ми називаємо *елементарними випадковими подіями.* Елементи σ-алгебри $\mathscr{F}$ ми називаємо *випадковими подіями.*

4. Випадкові величини

Нехай Ω — простір елементарних подій. Відображення X,

$$X : \Omega \to \mathbf{R},$$

це правило, за яким кожній елементарній події $\omega \in \Omega$ ставиться у відповідність дійсне число $X(\omega)$.

4.1. Відображення Бернуллі. Нехай $\Omega = \{\omega_0, \omega_1\}$. *Відображення Бернуллі* визначається правилом

$$X(\omega_0) = 0, \qquad X(\omega_1) = 1.$$

Іноді відображення Бернуллі зручно означати трохи іншим чином. Нехай Ω — довільна множина, A — її підмножина. Тоді

$$X(\omega) = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \notin A, \end{cases} \tag{2}$$

називається відображенням Бернуллі.

4.2. Відображення Радемахера. Нехай $\Omega = \{\omega_1, \omega_2\}$. *Відображення Радемахера* визначається правилом

$$X(\omega_1) = -1, \qquad X(\omega_2) = 1.$$

Зауваження 5. Відображення Бернуллі та Радемахера є різними, оскільки у них різна множина значень.

4.3. Ідентичне відображення. Нехай $\Omega = [0,1]$. Відображення $X(\omega) = \omega$ називається *ідентичним.* Таке ж правило можна застосувати і для іншого простору елементарних подій, наприклад для $\Omega = \mathbf{R}$.

Щоб розрізняти випадки $\Omega = [0,1]$ та $\Omega = \mathbf{R}$ будемо писати $X_{[0,1]}$ та $X_{\mathbf{R}}$ для ідентичного відображення. Відображення $X_{[0,1]}$ та $X_{\mathbf{R}}$ є різними не тільки тому, що різними є множини їх значень, але й тому, що різними є простори елементарних подій, на яких вони визначені.

Розглянемо відображення $Y(\omega) = \omega \cdot 1\!\!1_{[0,1]}(\omega)$ для $\Omega = \mathbf{R}$. Тоді область значень Y та $X_{[0,1]}$ є однаковими, але різними є простори елементарних подій, на яких вони визначені. До речі,Y визначено за правилом (2) для $A = [0,1]$.

4.4. Вимірні відображення. Нехай $\mathscr{F}$ — це σ-алгебра підмножин Ω, а $\mathscr{B}$ — це борелівська σ-алгебра в $\mathbf{R}$. Відображення $X : \Omega \to \mathbf{R}$ називається *вимірним*, якщо

$$X^{-1}(B) \in \mathscr{F} \qquad \text{для будь-якого } B \in \mathscr{B}.$$

Тут $X^{-1}(B)$ — це *прообраз* множини B, тобто

$$X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\}.$$

Приклад 6. Нехай $\Omega = \{\omega_0, \omega_1\}$, $\mathscr{F} = 2^\Omega$, X — відображення Бернуллі. Як і раніше, через A_i будемо позначати підмножину, яка складається з однієї елементарної події ω_i. Оскільки

$$X^{-1}(B) = \begin{cases} \Omega, & \text{якщо } 0, 1 \in B; \\ \varnothing, & \text{якщо } 0, 1 \notin B; \\ A_0, & \text{якщо } 0 \in B,\ 1 \notin B; \\ A_1, & \text{якщо } 0 \notin B,\ 1 \in B, \end{cases} \tag{3}$$

то відображення Бернуллі є вимірним.

Задача 13. *Довести, що відображення Радемахера є вимірним.*

Приклад 7. Нехай $\Omega = [0, 1]$, $\mathscr{F}$ — борелівська σ-алгебра на інтервалі $[0, 1]$, X — ідентичне відображення. Нехай B — борелівська множина в $\mathbf{R}$. Зрозуміло, що $X^{-1}(B) = B \cap [0, 1]$. Оскільки інтервал $[0, 1]$ є борелівською множиною в $\mathbf{R}$, то $B \cap [0, 1]$ також є борелівською множиною в $\mathbf{R}$. Але $B \cap [0, 1] \subseteq [0, 1]$ і тому $B \cap [0, 1]$ є борелівською множиною на $[0, 1]$. Це означає, що ідентичне відображення для $\Omega = [0, 1]$ є вимірним.

Задача 14. *Нехай $X_{\mathbf{R}}$ — це ідентичне відображення для $\Omega = \mathbf{R}$. Нехай Y — це відображення, означене в §4.3. Довести, що $X_{\mathbf{R}}$ та Y є вимірними відображеннями.*

Означення 2. Нехай $(\Omega, \mathscr{F}, \mathsf{P})$ — це ймовірнісний простір. Кожне вимірне відображення $X : \Omega \to \mathbf{R}$ називається *випадковою величиною.*

5. Розподіл випадкових величин

Приклад 8. Нехай $\Omega = \{\omega_0, \omega_1\}$, $\mathscr{F} = 2^\Omega$, X — випадкова величина Бернуллі. Нехай P — ймовірнісна міра на $(\Omega, \mathscr{F})$. Таблиця

$$\begin{pmatrix} X(\omega_0) & X(\omega_1) \\ \mathsf{P}(\omega_0) & \mathsf{P}(\omega_1) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ p & 1-p \end{pmatrix} \tag{4}$$

називається *розподілом (випадкової величини) Бернуллі.*

Приклад 9. Нехай $\Omega = \{\omega_0, \omega_1\}$, $\mathscr{F} = 2^\Omega$, X — відображення Радемахера. Нехай P — ймовірнісна міра на $(\Omega, \mathscr{F})$. Таблиця

$$\begin{pmatrix} X(\omega_0) & X(\omega_1) \\ \mathsf{P}(\omega_0) & \mathsf{P}(\omega_1) \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ p & 1-p \end{pmatrix} \tag{5}$$

називається *розподілом (випадкової величини) Радемахера.*

5.1. Симетрична випадкова величина Бернуллі. Випадкова величина Бернуллі називається *симетричною*, якщо її розподілом є

$$\begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

5.2. Симетрична випадкова величина Радемахера. Випадкова величина Радемахера називається *симетричною*, якщо її розподілом є

$$\begin{pmatrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Випадкові величини Бернуллі та Радемахера є представниками широкого класу дискретних випадкових величин.

Означення 3. Нехай $(\Omega, \mathscr{F}, \mathsf{P})$ — це ймовірнісний простір, X — випадкова величина. Якщо множина її значень є скінченою, то X називається *дискретною випадковою величиною.*

Означення 4. Нехай $(\Omega, \mathscr{F}, \mathsf{P})$ — це ймовірнісний простір, X — дискретна випадкова величина. Позначимо множину її значень через $\{x_1, \dots, x_m\}$ та покладемо $p_k = \mathsf{P}(X = x_k)$. Таблиця

$$\begin{pmatrix} x_1 & x_2 & \dots & x_m \\ p_1 & p_2 & \dots & p_m \end{pmatrix} \tag{6}$$

називається *розподілом дискретної випадкової величини* X.

Зауваження 6. Порівняймо загальне означення 4 розподілу випадкової величини та конкретне означення розподілу Бернуллі:

Означення Бернуллі	Загальне означення
$\begin{pmatrix} X(\omega_0) & X(\omega_1) \\ \mathsf{P}(\omega_0) & \mathsf{P}(\omega_1) \end{pmatrix}$	$\begin{pmatrix} 0 & 1 \\ \mathsf{P}(X = 0) & \mathsf{P}(X = 1) \end{pmatrix}$

Після заміни у лівій таблиці $X(\omega_0)$ на 0, а $X(\omega_1)$ на 1 перші рядки обох таблиць стають ідентичними. Другі рядки також не відрізняються, оскільки $\{\omega \colon X(\omega) = 0\} = A_0$ та $\{\omega \colon X(\omega) = 1\} = A_1$ (нагадаймо, що A_i це підмножина в Ω, яка складається з єдиної елементарної події ω_i). Таким чином, обидва означення є еквівалентними. Проте в більш складних випадках різниця між означеннями може стати суттєвою.

Приклад 10. Нехай $\Omega = [0, 1]$, $\mathscr{F} = \mathscr{B}_{[0,1]}$. Означимо відображення $X : \Omega \to \mathbf{R}$ за правилом

$$X(\omega) = \begin{cases} 0, & \text{якщо } 0 \le \omega < \frac{1}{2}, \\ 1, & \text{якщо } \frac{1}{2} \le \omega \le 1. \end{cases} \tag{7}$$

Задача 15. *Довести, що відображення X, яке означено правилом* (7), *є випадковою величиною.*

Множиною значень відображення (7) є $\{0,1\}$ і тому згідно означенню 4 її розподілом є

$$\begin{pmatrix} 0 & 1 \\ \mathsf{P}(X=0) & \mathsf{P}(X=1) \end{pmatrix}.$$

Якщо, наприклад, ймовірнісною мірою P є міра Лебега на інтервалі $[0,1]$, то цей розподіл стає розподілом симетричної випадкової величини Бернуллі, але $\mathsf{P}(X=0) \neq \mathsf{P}(\omega_0)$ для жодної $\omega_0 \in [0,1]$ і тому розподіл не можна записати у вигляді (4).

5.3. Розподіл ідентичної випадкової величини. Розподіл випадкової величини залежить від ймовірнісної міри P. Наприклад, розподіли (4) та (5) залежать від параметра p, який визначається мірою P. Тому, коли ми говоримо про розподіл, необхідно також згадувати відповідну ймовірнісну міру.

Розглянемо ідентичну випадкову величину X для $\Omega = [0,1]$, $\mathscr{F} = \mathscr{B}_{[0,1]}$. Її множиною значень є інтервал $[0,1]$, причому при кожному $x \in [0,1]$ існує тільки одна $\omega \in \Omega$, для якої $X(\omega) = x$. Якщо при цьому ймовірнісною мірою є міра Лебега, то

$$\mathsf{P}(X = x) = 0 \qquad \text{для будь-якого } x \in [0,1].$$

Спробувавши знайти розподіл такої випадкової величини, ми стікнемось з двома проблемами. Перша полягає у тому, що множина значень випадкової величини є незліченою і таблицю (6) не можна записати у тому ж вигляді: нам прийдеться використовувати позначення, притаманні відображенням. Більш серьозною проблемою є те, що другий рядок таблиці (6) буде містити тільки нулі і тому такий "розподіл" взагалі не має жодного смислу.

З іншого боку, нетривіальні значення ймовірності можна отримати не для окремих значень, а для їх сукупностей:

$$\mathsf{P}(X \in B), \qquad B \in \mathscr{B}_{\mathbf{R}}, \tag{8}$$

а саме

$$\mathsf{P}(X \in B) = \mathrm{Leb}\,(B \cap [0,1]), \qquad B \in \mathscr{B}_{\mathbf{R}}, \tag{9}$$

де $\mathrm{Leb}\,(\cdot)$ — це міра Лебега на інтервалі $[0,1]$. Наприклад,

$$\mathsf{P}(X \in [0.25, 0.75]) = \frac{1}{2}.$$

Означення 5. Нехай $(\Omega, \mathscr{F}, \mathsf{P})$ — ймовірнісний простір, X — довільна випадкова величина. Набір ймовірностей (8) називається *розподілом випадкової величини* X.

Задача 16. *Нехай X — дискретна випадкова величина. Виразити розподіл* (8) *через розподіл* (6).

5.4. Міра, породжена випадковою величиною. Формулою (8) означено функцію множин $\mathsf{P}_X(B) = \mathsf{P}(X \in B)$, визначену для борелівських підмножин $B \in \mathscr{B}_{\mathbf{R}}$. Фактично $\mathsf{P}_X(B)$ дорівнює значенню мірі P прообраза $X^{-1}(B)$ підмножини B:

$$\mathsf{P}_X(B) = \mathsf{P}(X^{-1}(B)), \qquad B \in \mathscr{B}_{\mathbf{R}}.$$

Теорема 4. *Нехай* $(\Omega, \mathscr{F}, \mathsf{P})$ *— ймовірнісний простір,* X *— випадкова величина. Тоді* $\mathsf{P}_X(B)$ *є ймовірнісною мірою на* $(\mathbf{R}, \mathscr{B}_{\mathbf{R}})$.

Задача 17. *Довести теорему* 4.

Приклад 11. Нехай $\Omega = \{\omega_0, \omega_1\}$, $\mathscr{F} = 2^\Omega$, X — випадкова величина Бернуллі. За формулою (3) всі значення міри, породженої випадковою величиною X, записано в наступній таблиці

(10)

B	$0,1 \notin B$	$0 \in B,\ 1 \notin B$	$0 \notin B,\ 1 \in B$	$0,1 \in B$
$\mathsf{P}_X(B)$	0	p	$1-p$	1

Приклад 12. Нехай $\Omega = [0,1]$, $\mathscr{F} = \mathscr{B}_{[0,1]}$, X — ідентична випадкова величина. Міра, породжена випадковою величиною X, визначається формулою (9). Міра $\mathsf{P}_X(B)$ відрізняється від міри Лебега на інтервалі $[0,1]$ тим, що перша визначена для всіх $B \in \mathscr{F}_{\mathbf{R}}$, а не тільки для $B \in \mathscr{F}_{[0,1]}$. Міра $\mathsf{P}_X(B)$ відрізняється також від міри Лебега на $\mathbf{R}$ своїми знвченнями. Наприклад, якщо $B = [0,2]$, то $\mathsf{P}_X(B) = 1$, але $\mathrm{Leb}\,(B) = 2$.

6. Функція розподілу випадкової величини

Розподіл довільної випадкової величини включає значення ймовірності для інтервалів $B = (-\infty, x)$ для довільних $x \in \mathbf{R}$.

Означення 6. *Функцією розподілу* випадкової величини X називається

$$F_X(x) = \mathsf{P}(X \in (-\infty, x)) = \mathsf{P}(X < x), \qquad x \in \mathbf{R}.$$

Приклад 13. Нехай $\Omega = \{\omega_0, \omega_1\}$, $\mathscr{F} = 2^\Omega$, X — випадкова величина Бернуллі. За формулою (10)

$$F_X(x) = \begin{cases} 0, & x < 0, \\ p, & 0 \le x < 1, \\ 1, & x \ge 1. \end{cases}$$

Задача 18. *Нехай* $\Omega = [0,1]$, $\mathscr{F} = \mathscr{B}_{[0,1]}$, X *— ідентична випадкова величина. Побудквати функцію розподілу* $F_X(x)$.

6.1. Зв'язок з розподілом. Звичайно, інтервали $(-\infty, x)$ складають лише незначну частину всіх борелівських множин, але розподіл випадкової величини однозначно визначається її функцією розподілу.

Теорема 5. *Розподіл довільної випадкової величини однозначно визначається її функцією розподілу.*

6.2. Властивості функції розподілу. Функція розподілу F довільної випадкової величини має такі властивості:

(F_1) F не спадає;
(F_2) $\lim_{x\to-\infty} F(x) = 0$;
(F_3) $\lim_{x\to\infty} F(x) = 1$;
(F_4) F є неперервною зліва.

Задача 19. *Довести властивості* (F_1)–(F_4).

Фактично теорему 5 можна подати у наступному вигляді.

Теорема 6. *Нехай функція F має властивості* (F_1)–(F_4). *Тоді на* $(\mathbf{R}, \mathscr{B}_\mathbf{R})$ *можна побудувати ймовірнісну міру μ так, що*

$$\mu((-\infty, x)) = F(x), \qquad x \in \mathbf{R}.$$

7. Розподіли Лебега–Стільтьеса ідентичної випадкової величини

Виявляється, що чотири властивості (F_1)–(F_4) є характеристичними для функцій розподілу.

Теорема 7. *Нехай деяка функція F задовольняє умови* (F_1)–(F_4). *Тоді F є функцією розподілу.*

Доведення. За теоремою 6 функція F генерує на $(\mathbf{R}, \mathscr{B}_\mathbf{R})$ ймовірнісну міру μ. Нехай $\Omega = \mathbf{R}$, $\mathscr{F} = \mathscr{B}_\mathbf{R}$, $\mathsf{P} = \mu$, X — ідентична випадкова величина. Тоді для $B_x = (-\infty, x)$

$$F_X(x) = \mathsf{P}(X < x) = \mathsf{P}_X(B_x) = \mu(B_x) = F(x), \qquad x \in \mathbf{R}.$$

Це і означає, що F є функцією розподілу випадкової величини (ми навіть визначили цю випадкову величину). $\square$

Зауваження 7. З доведення теореми 7 видно, що ідентична випадкова величина X має особливу властивість: саме X використовується для побудови розподілів, а зрештою і самих випадкових величин.

7.1. Абсолютно неперервні функції розподілу. Значення розподілу $\mathsf{P}_X(B)$, який породжує функція F з властивостями (F_1)–(F_4), обчислюються за допомогою *інтеграла Лебега–Стільтьеса*

$$\mu(B) = \int_B dF, \qquad B \in \mathscr{B}_\mathbf{R}.$$

Тому вони називаються *розподілами Лебега–Стільтьеса.* Серед таких розподілів особливе значення мають такі, для яких

$$\int_B dF = \int_B f(u)\,du, \qquad B \in \mathscr{B}_\mathbf{R},$$

для деякої невідємної функції f. Такі розподіли називаються *абсолютно неперевними*, а сама функція f називається *щільністю розподілу.* Випадкові величини, розподіли яких є абсолютно неперервними, також називаються *абсолютно неперервними випадковими величинами.*

Задача 20. *Навести приклад не абсолютно неперервного розподілу.*

8. Гауссівський розподіл

Особливу роль в теорії ймовірностей грає *гауссівський* (або *нормальний*) розподіл. Цей абсолютно неперервний розподіл має щільність

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \qquad x \in \mathbf{R}.$$

Відомою є формула (*інтеграл Пуассона*)

$$\int_{-\infty}^{\infty} \varphi(x)\,dx = 1. \tag{11}$$

З парності функції φ випливає, що

$$\int_{-\infty}^{\infty} x\varphi(x)\,dx = 0. \tag{12}$$

Нескладно також отримати, що

$$\begin{aligned}\int_{-\infty}^{\infty} x^2\varphi(x)\,dx &= 2\int_0^{\infty} x^2\varphi(x)\,dx \\ &= \left[\begin{smallmatrix}\text{Частинами:}\\ u=x,\ dv=x\varphi(x)\,dx\\ du=dx,\ v=-\varphi(x)\end{smallmatrix}\right] = 2\left(-x\varphi(x)\Big|_0^{\infty} + \int_0^{\infty}\varphi(x)\,dx\right) \\ &= 2\int_0^{\infty}\varphi(x)\,dx,\end{aligned}$$

тобто

$$\int_{-\infty}^{\infty} x^2\varphi(x)\,dx = 1. \tag{13}$$

8.1. Симетричні випадкові величини. Гауссівський розподіл має властивість *симетричності*. Нехай X — гауссівська випадкова величина, $Y = -X$. Тоді

$$\begin{aligned}\mathsf{P}_Y(x) = \mathsf{P}(Y < x) = \mathsf{P}(X > -x) = \int_{-x}^{\infty}\varphi(u)\,du = \int_{-\infty}^{x}\varphi(-u)\,du \\ = \int_{-\infty}^{x}\varphi(u)\,du = \mathsf{P}_X(x).\end{aligned}$$

Це означає, що X та $-X$ мають однаковий розподіл. Іншими словами, X має симетричний розподіл.

8.2. Гауссівський розподіл з параметрами. Нехай $a, \sigma^2 \in \mathbf{R}$, $\sigma^2 \neq 0$. Гауссівський розподіл з параметрами a та σ^2 визначається щільністю

$$\varphi_{a,\sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-a)^2/2\sigma^2}, \qquad x \in \mathbf{R}.$$

Заміною змінної інтегрування нескладно довести, що

$$\int_{-\infty}^{\infty}\varphi_{a,\sigma^2}(x)\,dx = 1, \tag{14}$$

$$\int_{-\infty}^{\infty} x\varphi_{a,\sigma^2}(x)\,dx = a, \tag{15}$$

$$\int_{-\infty}^{\infty}(x-a)^2\varphi_{a,\sigma^2}(x)\,dx = \sigma^2. \tag{16}$$

Задача 21. *Довести справедливість формул* (14)–(16).

† Всего в тексте было 0 вопросов

Лекция 15

СЛУЧАЙНЫЕ ВЕКТОРЫ

1. Функция распределения случайного вектора

Пусть $n > 1$ и $X_1, \ldots, X_n$ — случайные величины. Вектор-столбец $\mathbf{X}$, $\mathbf{X}' = (X_1, \ldots, X_n)$, называется *случайным вектором.* Тут $\mathbf{X}'$ означает транспонирование вектора $\mathbf{X}$.

Функцией распределения вектора $\mathbf{X}$ называется

$$F_{\mathbf{X}}(\mathbf{x}) = \mathsf{P}(X_1 < x_1, \ldots, X_n < x_n), \qquad \mathbf{x} \stackrel{\text{def}}{=} (x_1, \ldots, x_n)' \in \mathbf{R}^n. \tag{1}$$

Функция $F_{\mathbf{X}}(\mathbf{x})$ определена для любых аргументов $\mathbf{x} \in \mathbf{R}^n$ поскольку

$$\{\omega\colon X_1 < x_1, \ldots, X_n < x_n\} = \bigcap_{k=1}^{n} \{\omega\colon X_k < x_k\}. \tag{2}$$

Задача 1. *Почему правая часть* (2) *принадлежит σ-алгебре $\mathscr{F}$?*

Вектор $\mathbf{X}$ называется *абсолютно непрерывным*, если

$$F_{\mathbf{X}}(\mathbf{x}) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_{\mathbf{X}}(t_1, \ldots, t_n)\, dt_1 \ldots dt_n, \qquad \mathbf{x} \in \mathbf{R}^n, \tag{3}$$

для некоторой функции $f_{\mathbf{X}}(\mathbf{x}) \ge 0$, которая в этом случае называется *плотностью вектора* $\mathbf{X}$.

Теорема 1. *Пусть $\mathbf{X}$ — абсолютно непрерывный вектор, а B — борелевское подмножество в $\mathbf{R}^n$. Тогда*

$$\mathsf{P}(\mathbf{X} \in B) = \int_B f_{\mathbf{X}}(t_1, \ldots, t_n)\, dt_1 \ldots dt_n.$$

Заметим, что равенство (3) является частным случаем теоремы 1 при

$$B = (-\infty, x_1) \times \cdots \times (-\infty, x_n).$$

1.1. Независимые координаты. *Если случайные величины $X_1, \ldots, X_n$ независимы в совокупности и каждая является абсолютно непрерывной случайной величиной, то вектор $\mathbf{X}$ также является абсолютно непрерывным с плотностью $f_{\mathbf{X}}(x_1, \ldots, x_n) = f_{X_1}(x_1) \ldots f_{X_n}(x_n)$.* Действительно, в этом случае функция распределения (1) преобразуется таким образом

$$\begin{aligned} F_{\mathbf{X}}(\mathbf{x}) = F_{X_1}(x_1) \cdot \ldots \cdot F_{X_n}(x_n) &= \int_{-\infty}^{x_1} f_{X_1}(t_1)\, dt_1 \cdot \ldots \cdot \int_{-\infty}^{x_n} f_{X_n}(t_n)\, dt_n \\ &= \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_{X_1}(t_1) \ldots f_{X_n}(t_n)\, dt_1 \ldots dt_n \end{aligned}$$

и поэтому равенство (3) выполнено для $f_{\mathbf{X}}(x_1, \ldots, x_n) = f_{X_1}(x_1) \ldots f_{X_n}(x_n)$.

---

[0]Printed from the file [15-random_vectors.tex] on 9.6.2016

2. Линейные преобразования

Пусть $\mathbf{A}$ — некоторая $n \times n$ матрица. Если $\det(\mathbf{A}) \neq 0$, то существует обратная матрица $\mathbf{A}^{-1}$, причем $\det(\mathbf{A}^{-1}) = \frac{1}{\det(\mathbf{A})}$. Кроме того, система уравнений $\mathbf{Ax} = \mathbf{y}$ имеет единственное решение при любом $\mathbf{y}$.

Рассмотрим линейное преобразование $\mathbf{Y} \stackrel{\text{def}}{=} \mathbf{AX}$ вектора $\mathbf{X}$. Поскольку координаты вектора $\mathbf{Y}$ представляют собой конечные линейные комбинации координат вектора $\mathbf{X}$, то они являются случайными величинами. Следовательно, $\mathbf{Y}$ является случайным вектором.

Задача 2. *Пусть $\mathbf{A}$ — $n \times n$ матрица, а $\mathbf{X}$ — n-мерный вектор. Доказать, что координаты вектора $\mathbf{Y} = \mathbf{AX}$ являются случайными величинами.*

Теорема 2. *Пусть $\det(\mathbf{A}) \neq 0$. Если вектор $\mathbf{X}$ является абсолютно непрерывным, то и вектор $\mathbf{Y} = \mathbf{AX}$ является абсолютно непрерывным. При этом*

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{|\det(\mathbf{A})|} f_{\mathbf{X}}(\mathbf{A}^{-1}\mathbf{y}). \tag{4}$$

Задача 3. *Доказать теорему 2 для $n = 1$.*

Доказательство. Пусть B — некоторое подмножество в $\mathbf{R}^n$. Прообраз множества B при отображении $\mathbf{A}$ обозначим $\mathbf{A}^{-1}B \stackrel{\text{def}}{=} \{\mathbf{x} : \mathbf{Ax} \in B\}$. Если $\mathbf{y} = \mathbf{Ax}$, то $\mathbf{x} \in \mathbf{A}^{-1}B \iff \mathbf{y} \in B$. Поэтому $\{\omega\colon \mathbf{Y} \in B\} = \big\{\omega\colon \mathbf{X} \in \mathbf{A}^{-1}B\big\}$. Обозначим $g(\mathbf{s}) \stackrel{\text{def}}{=} \mathbf{A}^{-1}\mathbf{s}$, а якобиан преобразования g — через $\mathbf{J}$. Тогда $\mathbf{J} = \mathbf{A}^{-1}$. Действительно, $g : \mathbf{R}^n \to \mathbf{R}^n$, то есть $g(\mathbf{s}) = (g_1(\mathbf{s}), \ldots, g_n(\mathbf{s}))'$. Если $\mathbf{A}^{-1} = \big(a_{ij}\big)_{i,j=1}^n$ и $\mathbf{s} = (s_1, \ldots, s_n)'$, то

$$g_i(\mathbf{s}) = \sum_{l=1}^n a_{il} s_l, \qquad 1 \le i \le n.$$

По определению якобиана $\mathbf{J} = \left(\frac{\partial g_i(\mathbf{s})}{\partial s_k}\right)_{i,k=1}^n$, то есть

$$\frac{\partial g_i(\mathbf{s})}{\partial s_k} = \sum_{k=1}^n a_{il} \frac{\partial s_l}{\partial s_k} = a_{ik}, \qquad 1 \le i, k \le n.$$

Это и означает, что якобиан $\mathbf{J}$ совпадает с обратной матрицей $\mathbf{A}^{-1}$. В частности, $\det(\mathbf{J}) = \frac{1}{\det(\mathbf{A})}$. Поэтому для любого измеримого множества $B \in \mathbf{R}^n$

$$\begin{aligned} \mathsf{P}(\mathbf{Y} \in B) = \mathsf{P}(\mathbf{X} \in \mathbf{A}^{-1}B) &= \int_{\mathbf{A}^{-1}B} f_{\mathbf{X}}(\mathbf{t})\, d\mathbf{t} = \left[\begin{smallmatrix}\text{Замена:} \\ \mathbf{t}=\mathbf{A}^{-1}\mathbf{s}\end{smallmatrix}\right] \\ &= \int_B f_{\mathbf{X}}(\mathbf{A}^{-1}\mathbf{s}) |\det(\mathbf{J})|\, d\mathbf{s} = \int_B \frac{1}{|\det(\mathbf{A})|} f_{\mathbf{X}}(\mathbf{A}^{-1}\mathbf{s})\, d\mathbf{s} \end{aligned} \tag{5}$$

по формуле замены переменных в кратном интеграле. Если в (5) выбрать $B = (-\infty, y_1) \times \cdots \times (-\infty, y_n)$, то получим формулу (3), откуда и вытекает (4). $\square$

Задача 4. *Доказать, что при замене $\mathbf{t} = \mathbf{A}^{-1}\mathbf{s}$ областью интегрирования становится B вместо $\mathbf{A}^{-1}B$.*

2.1. Линейные преобразования двух гауссовских случайных величин. *Если X_1 и X_2 независимые стандартные гауссовские случайные величины, то $Y_1 \stackrel{\text{def}}{=} X_1+X_2$ и $Y_2 \stackrel{\text{def}}{=} X_1-X_2$ также независимые $\mathcal{N}(0,2)$ случайные величины.* Действительно, $X_1 = \frac{Y_1+Y_2}{2}$ и $X_2 = \frac{Y_1-Y_2}{2}$, то есть $\mathbf{Y} = \mathbf{A}\mathbf{X}$, где

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \qquad \det(\mathbf{A}) = -2, \qquad \mathbf{A}^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}.$$

Имеем $f_{X_1}(x) = f_{X_2}(x) = \varphi(x)$ и $f_{\mathbf{X}}(x_1,x_2) = f_{X_1}(x_1)f_{X_2}(x_2)$ по критерию независимости абсолютно непрерывных векторов. В силу теоремы 2 при $n=2$

$$\begin{aligned} f_{\mathbf{Y}}(\mathbf{y}) &= \frac{1}{2} f_{\mathbf{X}}\left(\frac{y_1+y_2}{2}, \frac{y_1-y_2}{2}\right) = \frac{1}{2}\varphi\left(\frac{y_1+y_2}{2}\right)\varphi\left(\frac{y_1-y_2}{2}\right) \\ &= \frac{1}{2}\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{8}(y_1+y_2)^2}\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{8}(y_1-y_2)^2} \\ &= \frac{1}{\sqrt{2\pi\sigma^2}}e^{-y_1^2/2\sigma^2}\cdot\frac{1}{\sqrt{2\pi\sigma^2}}e^{-y_2^2/2\sigma^2} = \varphi_{0,\sigma^2}(y_1)\varphi_{0,\sigma^2}(y_2) \end{aligned}$$

для $\sigma^2 = 2$. Так как $f_{Y_1}(y_1) = \int_{-\infty}^{\infty} f_{\mathbf{Y}}(y_1,y_2)\,dy_2 = \varphi_{0,\sigma^2}(y_1)$ и аналогично $f_{Y_2}(y_2) = \varphi_{0,\sigma^2}(y_2)$, то из последнего равенства вытекает, что $f_{\mathbf{Y}}(\mathbf{y}) = f_{Y_1}(y_1)f_{Y_2}(y_2)$. Значит Y_1 и Y_2 независимы по критерию независимости абсолютно непрерывных случайных величин.

Теорема 3. *Пусть $X_1,\dots,X_n$ — независимые гауссовские случайные величины, а $c_1,\dots,c_n$ — произвольные действительные числа, $|c_1|+\dots+|c_n|>0$. Тогда линейная комбинация $c_1X_1+\dots+c_nX_n$ является гауссовской случайной величиной.*

Доказательство. Сначала рассмотрим случай $n=2$. Пусть $\mathsf{E}[X_i] = \mu_i$ и $\mathsf{var}[X_i] = \sigma_i^2$, $i=1,2$. Тогда

$$Z_1 = \frac{X_1-\mu_1}{\sqrt{\sigma_1^2}} \quad \text{и} \quad Z_2 = \frac{X_2-\mu_2}{\sqrt{\sigma_2^2}}$$

являются независимыми стандартными гауссовскими случайными величинами. Поэтому их линейная комбинация

$$c_1\sqrt{\sigma_1^2}\cdot Z_1 + c_2\sqrt{\sigma_2^2}\cdot Z_2 = c_1X_1 + c_2X_2 - \mu, \qquad \mu = c_1\mu_1 + c_2\mu_2,$$

является гауссовской случайной величиной. Отсюда и вытекает, что $c_1X_1 + c_2X_2$ также является гауссовской случайной величиной.

Задача 5. *Пусть X — гауссовская случайная величина, а $\mu \in \mathbf{R}$. Доказать, что $X+\mu$ также гауссовская случайная величина.*

Случай общего $n \geq 2$ вытекает из частного случая $n=2$ по индукции. ① □

---

[1] С помощью метода математической индукции, доказать теорему 3 для произвольного $n \geq 2$.

2.2. Формула свертки. *Пусть X_1 и X_2 независимые абсолютно непрерывные случайные величины. Тогда X_1+X_2 абсолютно непрерівная случайная величина, причем*

$$f_{X_1+X_2}(x) = \int_{-\infty}^{\infty} f_{X_1}(y) f_{X_2}(x-y)\,dy. \tag{6}$$

Положим $Y_1 = X_1 + X_2$ и $Y_2 = X_1$. Тогда $X_1 = Y_2$, $X_2 = Y_1 - Y_2$, то есть $\mathbf{Y} = \mathbf{A}\mathbf{X}$, где

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \det(\mathbf{A}) = -1, \qquad \mathbf{A}^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}.$$

В силу (4)

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{|\det(\mathbf{A})|} f_{\mathbf{X}}\left(\mathbf{A}^{-1}\mathbf{y}\right) = f_{\mathbf{X}}(y_2, y_1 - y_2). \tag{7}$$

Согласно критерию независимости абсолютно непрерывных случайных величин $f_{\mathbf{X}}(\mathbf{x}) = f_{X_1}(x_1) f_{X_2}(x_2)$ для любых $\mathbf{x} \stackrel{\text{def}}{=} (x_1, x_2) \in \mathbf{R}^2$. Отсюда и из (7) вытекает (6).

Задача 6. *Пусть $X_1, \dots, X_n$ — независимые в совокупности абсолютно непрерывные случайные величины. Тогда их сумма $S = X_1 + \dots + X_n$ также абсолютно непрерывная случайная величина, причем*

$$f_S(x) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{X_1}(y_1) \dots f_{X_{n-1}}(y_{n-1}) f_{X_n}(x - y_1 - \dots - y_{n-1})\, dy_1 \dots dy_{n-1}.$$

3. Нелинейные преобразования

Пусть по-прежнему $\mathbf{X}$ — абсолютно непрерывный n-мерный случайный вектор. Носитель его плотности обозначим $\operatorname{supp} f_{\mathbf{X}} \stackrel{\text{def}}{=} S$ ($\subseteq \mathbf{R}^n$). Пусть, далее, $g : \mathbf{R}^n \to \mathbf{R}^n$ — биективное измеримое отображение S в некоторое множество $T \subseteq \mathbf{R}^n$. Тогда $\mathbf{Y} \stackrel{\text{def}}{=} g(\mathbf{X})$ называется *преобразованием вектора* $\mathbf{X}$. Поскольку g — биекция между S и T, то любому $\mathbf{y} \in T$ отвечает единственный $\mathbf{x} \in S$, для которого $\mathbf{y} = g(\mathbf{x})$. Это соответствие между T и S, называемое обратным к g, обозначим $h : T \to S$. Отображение h также является биективным. Соответствия g и h можно записать более подробно таким образом:

$$\mathbf{Y} = g(\mathbf{X}) \Longleftrightarrow \begin{cases} Y_1 = g_1(X_1, \dots, X_n), \\ \quad\vdots \\ Y_n = g_n(X_1, \dots, X_n), \end{cases} \qquad \mathbf{X} = h(\mathbf{Y}) \Longleftrightarrow \begin{cases} X_1 = h_1(Y_1, \dots, Y_n), \\ \quad\vdots \\ X_n = h_n(Y_1, \dots, Y_n). \end{cases}$$

Теорема 4. *Предположим, что функции $h_1, \dots, h_n$ являются непрерывно дифференцируемыми в области T. Обозначим через $\mathbf{J}$ матрицу, составленную из производных $\frac{\partial h_i}{\partial y_j}$, $1 \le i, j \le n$. Тогда*

$$f_{\mathbf{Y}}(\mathbf{y}) = \begin{cases} f_{\mathbf{X}}(h(\mathbf{y})) \cdot |\det(\mathbf{J})|, & \mathbf{y} \in T, \\ 0, & \textit{в других случаях.} \end{cases} \tag{8}$$

Доказательство теоремы 4 полностью аналогично доказательству теоремы 2. Заметим, что $|\det(\mathbf{J})|$ — это якобиан для замены переменных в доказательстве (5).

4. Двумерное гауссовское распределение

Пусть $-1 < \rho < 1$ и

$$(9)\qquad f(x,y) = f_\rho(x,y) \stackrel{\text{def}}{=} \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}(x^2-2\rho xy+y^2)}, \qquad (x,y)\in\mathbf{R}^2.$$

Если $\rho = 0$, то

$$(10)\qquad f_0(x,y) = \varphi(x)\varphi(y),$$

где $\varphi = \varphi_{0,1}$ — стандартная гауссовская плотность Следовательно в этом случае f является плотностью некоторого случайного вектора.

Задача 7. *Доказать, что* (10) *является плотностью некоторого случайного вектора.*

Напомним, что φ_{a,σ^2} — это гауссовская плотность с параметрами $a\in\mathbf{R}$ и $\sigma^2>0$ (см. §8.2, лекция 14):

$$\varphi_{a,\sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}}, \qquad x\in\mathbf{R}.$$

Покажем, что функция (9) является плотностью не только для $\rho = 0$, но и для любого $-1<\rho<1$. Так как $f_\rho(x,y)\ge 0$, то достаточно доказать, что

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_\rho(x,y)\,dx\,dy = 1.$$

Поскольку $x^2-2\rho xy+y^2 = (x-\rho y)^2+(1-\rho^2)y^2$, то

$$(11)\qquad f_\rho(x,y) = \frac{1}{\sqrt{2\pi(1-\rho^2)}} e^{-\frac{1}{2(1-\rho^2)}(x-\rho y)^2}\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac12 y^2} = \varphi_{\rho y,1-\rho^2}(x)\varphi_{0,1}(y).$$

Задача 8. *Доказать, что* $f_\rho(x,y) = \varphi_{\rho x,1-\rho^2}(y)\varphi_{0,1}(x)$.

Так как $\varphi_{\rho y,1-\rho^2}(\cdot)$ — плотность гауссовского распределения при любом y и любом $-1<\rho<1$, то

$$\int_{-\infty}^{\infty}\varphi_{\rho y,1-\rho^2}(x)\,dx = 1$$

для любого y и поэтому

$$(12)\qquad \int_{-\infty}^{\infty} f_\rho(x,y)\,dx = \varphi_{0,1}(y), \quad\text{откуда}\quad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_\rho(x,y)\,dx\,dy = 1.$$

Определение 1. Случайный вектор (X_1,X_2) с плотностью (9) называется *двумерным стандартным гауссовским вектором.* Функция (9) называется *стандартной двумерной гауссовской плотностью.*

4.1. Маргинальные распределения гауссовского вектора. *Маргинальные распределения гауссовского вектора с плотностью* (9) *являются гауссовскими* $\mathcal{N}(0,1)$ *распределениями. Это означает, что координаты стандартного гауссовского вектора являются стандартными гауссовскими случайными величинами.*

Чтобы найти маргинальное распределение X_2, необходимо проинтегрировать плотность (9) по x. Это уже сделано в (12). Следовательно плотностью X_2 является $\varphi_{0,1}$. Так как функция (9) симметрична относительно x и y, такой же результат справедлив и для X_1.

Задача 9. *Являются ли координаты стандартного гауссовского вектора независимыми случайными величинами?*

4.2. Линейная комбинация координат имеет гауссовское распределение. *Пусть* $\mathbf{X}$ *— стандартный двумерный гауссовский вектор,* $c_1, c_2 \in \mathbf{R}$, $|c_1| + |c_2| \neq 0$. *Тогда случайная величина* $c_1X_1 + c_2X_2$ *имеет гауссовское распределение с параметрами* 0 *и* $c_1^2 + 2c_1c_2\rho + c_2^2$.

Для доказательства необходимо воспользоваться методом доказательства теоремы из §2.2. В случае $c_1 = c_2 = 1$ можно воспользоваться формулой (7).

Задача 10. *Пусть двумерный стандартный гауссовский вектор. Пусть* $c_1, c_2 \in \mathbf{R}$, $|c_1|+|c_2| \neq 0$. *Доказать, что* $c_1X_1+c_2X_2$ *имеет гауссовское распределение с параметрами* 0 *и* $2(1+\rho)$, *то есть* $c_1X_1 + c_2X_2 \in \mathcal{N}(0, 2(1+\rho))$.

4.3. Смысл параметра ρ. Найдем $\mathsf{E}\,[X_1X_2]$:

$$\begin{aligned}\mathsf{E}\,[X_1X_2] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_\rho(x,y)\,dx\,dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\varphi_{\rho y,1-\rho^2}(x)\varphi_{0,1}(y)\,dx\,dy\\ &= \int_{-\infty}^{\infty} y\varphi_{0,1}(y)\left[\int_{-\infty}^{\infty} x\varphi_{\rho y,1-\rho^2}(x)\,dx\right]dy = \rho\int_{-\infty}^{\infty} y^2\varphi_{0,1}(y) = \rho,\end{aligned}$$

так как интеграл в квадратных скобках равен ρy, а последний интеграл равен 1. Таким образом $\rho = \mathsf{E}\,[X_1X_2]$, то есть ρ — это ковариация $\mathsf{cov}\,[X_1, X_2]$ (так как $\mathsf{E}\,[X_1] = \mathsf{E}\,[X_2] = 0$).

4.4. Критерий факторизации гауссовской плотности. *Если* $f_\rho(x,y) = \varphi_{a_1,\sigma_1^2}(x)\varphi_{a_2,\sigma_2^2}(y)$ *для любых* $x, y \in \mathbf{R}$ *и для каких-либо* a_1, a_2 *и* $\sigma_1^2 \neq 0$, $\sigma_2^2 \neq 0$, *то* $\rho = 0$ *и обязательно* $a_1 = a_2 = 0$, $\sigma_1^2 = \sigma_2^2 = 1$.

Доказательство. Из §4.3 и условия факторизации плотности f_ρ вытекает , что $\rho = a_1a_2$. Из §4.1 вытекает, что $\mathsf{E}\left[X_1^2\right] = 1$. ② С другой стороны,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2 f_\rho(x,y)\,dx\,dy = \int_{-\infty}^{\infty} x^2\varphi_{a_1,\sigma_1^2}(x)\,dx\cdot\int_{-\infty}^{\infty}\varphi_{a_2,\sigma_2^2}(y)\,dy = (\sigma_1^2 + a_1^2)\cdot 1,$$

то есть $\sigma_1^2 + a_1^2 = 1$. Аналогично получаем $\sigma_2^2 + a_2^2 = 1$. Таким же образом вычисляем $\mathsf{E}\left[X_1^2X_2\right]$:

$$\begin{aligned}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2yf_\rho(x,y)\,dx\,dy &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2y\varphi_{\rho y,1-\rho^2}(x)\varphi_{0,1}(y)\,dx\,dy\\ &= \int_{-\infty}^{\infty} y\varphi_{0,1}(y)\left[\int_{-\infty}^{\infty} x^2\varphi_{\rho y,1-\rho^2}(x)\,dx\right]dy\\ &= \int_{-\infty}^{\infty}(1-\rho^2+\rho^2y^2)y\varphi_{0,1}(y)\,dy = 0,\end{aligned}$$

так как подинтегральная функция нечетна. С другой стороны,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2yf_\rho(x,y)\,dx\,dy = \int_{-\infty}^{\infty} x^2\varphi_{a_1,\sigma_1^2}(x)\,dx\cdot\int_{-\infty}^{\infty} y\varphi_{a_2,\sigma_2^2}(y)\,dy = (\sigma_1^2+a_1^2)a_2,$$

откуда $a_2 = 0$. Аналогично $a_1 = 0$. Следовательно $\rho = 0$ и $\sigma_1^2 = \sigma_2^2 = 1$. □

Замечание 1. Учитывая (10) и теорему 4.4, заключаем, что если $(X_1, X_2)'$ — это стандартный двумерный гауссовский вектор, то случайные величины X_1 и X_2 независимы тогда и только тогда, когда $\rho = 0$.

[2] Почему $\rho = a_1a_2$ в случае факторизации плотности стандартного гауссовского вектора?

Задача 11. *Доказать, что если* $(X_1, X_2)'$ *— это двумерный стандартный гауссовский вектор, то случайные величины* X_1 *и* X_2 *независимы тогда и только тогда, когда* $\rho = 0$.

Теорема 5. *Пусть* f_ρ *— плотность двумерного стандартного гауссовского вектора* $\mathbf{X}$, *заданная равенством* (9). *Тогда* $\mathsf{cov}\,[X_1, X_2] = \rho$ *и любая линейная комбинация* $c_1X_1 + c_2X_2$ *координат вектора,* $|c_1| + c_2| \neq 0$, *является гауссовской* $\mathcal{N}\left(0, c_1^2 + 2c_1c_2\rho + c_2^2\right)$ *случайной величиной.*

Поэтому $X_1 \in \mathcal{N}(0,1)$ *и* $X_2 \in \mathcal{N}(0,1)$. *В частности,*

$$\begin{aligned} \mathsf{E}\,[X_1] &= \mathsf{E}\,[X_2] = 0, \\ \mathsf{var}\,[X_1] &= \mathsf{var}\,[X_2] = 1. \end{aligned}$$

Наконец, координаты X_1 *и* X_2 *вектора* $\mathbf{X}$ *являются независимыми случайными величинами тогда и только тогда, когда* $\rho = 0$.

[†]Всего в тексте было 2 вопросов

Лекция 16

ГАУССОВСКИЕ СЛУЧАЙНЫЕ ВЕКТОРЫ

1. Двумерные гауссовские случайные векторы

В лекции 15 мы рассматривали двумерную стандартную гауссовскую плотность. Случайный вектор, который имеет такую плотность, мы называли стандартным гауссовским вектором.

Как и в случае случайных величин (см. §8.2, лекция 14), мы определим общие гауссовские векторы, но при этом будем пользоваться другим определением. Позднее мы докажем, что новое определение эквивалентно определению из раздела 15.4.

Определение 1. Случайный вектор $\mathbf{X} = (X_1, X_2)'$ называется *гауссовским случайным вектором*, если линейная комбинация $c_1X_1 + c_2X_2$ является гауссовской случайной величиной при любых c_1 и c_2, $|c_1| + |c_2| \neq 0$.

Напомним, что таким свойством обладают стандартные гауссовские случайные векторы (см. §4.2, лекция 15).

1.1. Простейшие свойства. Непосредственно из определения 1 вытекают такие свойства гауссовских векторов:

(i) гауссовские векторы существуют: например, если X_1 и X_2 независимые стандартные гауссовские случайные величины, то вектор $\mathbf{X} = (X_1, X_2)'$ является гауссовским;

(ii) если $\mathbf{X}$ — гауссовский случайный вектор, то X_1 и X_2 являются гауссовскими случайными величинами; (не обязательно независимыми!)

(iii) если $\mathbf{X}$ — гауссовский случайный вектор, то $X_1 + X_2$ является гауссовской случайной величиной.

Доказательство. (i) Сначала докажем, что $Y \stackrel{\text{def}}{=} cX \in \mathcal{N}\big(0, c^2\sigma^2\big)$ для любого $c \neq 0$, если $X \in \mathcal{N}\big(0, \sigma^2\big)$. Действительно, из теоремы 15.2 получаем $f_Y(y) = \frac{1}{|c|} f_X(y/c)$ или $f_Y(y) = \frac{1}{\sqrt{2\pi c^2\sigma^2}} e^{-(y/c)^2/2\sigma^2} = \varphi_{0,c^2\sigma^2}(y)$.

Теперь покажем, что $cX_1 + X_2 \in \mathcal{N}\big(0, 1 + c^2\big)$ для любого $c \in \mathbf{R}$. Это утверждение тривиально для $c = 0$. Докажем его и для $c \neq 0$. Так как $cX_1 \in \mathcal{N}\big(0, c^2\big)$, то из формулы свертки (15.6) получаем

$$f_{cX_1+X_2}(x) = \frac{1}{2\pi|c|} \int_{-\infty}^{\infty} e^{-v^2/2c^2} \cdot e^{-(x-v)^2/2}\, dv.$$

Положим $\sigma^2 = \frac{c^2}{1+c^2}$. Тогда

$$\frac{v^2}{c^2} + (x-v)^2 - \frac{x^2}{1+c^2} = \frac{v^2}{\sigma^2} - 2xv + \sigma^2 x^2 = \Big(\frac{v}{\sigma} - \sigma x\Big)^2 = \frac{1}{\sigma^2}(v - \sigma^2 x)^2.$$

[0]Printed from the file `[16-Gauss-vectors.tex]` on 9.6.2016

Поэтому

$$\begin{aligned} f_{cX_1+X_2}(x) &= \frac{1}{2\pi|c|}\int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{v^2}{\sigma^2}+(x-v)^2\pm\frac{x^2}{1+c^2}\right)}\,dv \\ &= \varphi_{0,1+c^2}(x)\cdot\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}(v-\sigma^2x)^2}\,dv \\ &= \varphi_{0,1+c^2}(x)\int_{-\infty}^{\infty}\varphi_{\sigma^2x,\sigma^2}(v)\,dv = \varphi_{0,1+c^2}(x), \end{aligned}$$

что и требовалось доказать.

Теперь легко закончить доказательство утверждения (i). Пусть, например, $c_2 \neq 0$. Тогда $c_1X_1 + c_2X_2 = c_2(cX_1 + X_2)$, где $c = c_1/c_2$. Поскольку $cX_1 + X_2 \in \mathcal{N}\left(0, 1+c^2\right)$, то $c_2(cX_1 + X_2) \in \mathcal{N}\left(0, c_2^2(1+c^2)\right) = \mathcal{N}\left(0, c_1^2 + c_2^2\right)$ по доказанному выше. Значит $c_1X_1 + c_2X_2$ является гауссовской случайной величиной с параметрами 0 и $c_1^2 + c_2^2$.

(ii) Достаточно выбрать $c_1 = 1, c_2 = 0$ или $c_1 = 0, c_2 = 1$ в определении 1.

(iii) Достаточно выбрать $c_1 = c_2 = 1$ в определении 1. $\square$

1.2. Вектор негауссовский, а координаты гауссовские. Если координаты вектора являются гауссовскими, то сам вектор не обязательно является гауссовским (свойство (ii) обратить нельзя). Действительно, пусть $X_1 \in \mathcal{N}(0,1)$, а β — симметричная случайная величина Бернулли, то есть $\mathsf{P}(\beta = -1) = \mathsf{P}(\beta = +1) = \frac{1}{2}$. Предположим, что β не зависит от X_1 и положим $X_2 = \beta X_1$. Так как

$$\begin{aligned} \{X_2 < v\} &= \{X_2 < v, \beta = -1\} \cup \{X_2 < v, \beta = +1\} \\ &= \{-X_1 < v, \beta = -1\} \cup \{X_1 < v, \beta = +1\}, \end{aligned}$$

то в силу независимости

$$\mathsf{P}(X_2 < v) = \frac{1}{2}\mathsf{P}(-X_1 < v) + \frac{1}{2}\mathsf{P}(X_1 < v) = \mathsf{P}\mathsf{P}(X_1 < v)$$

в силу симметричности гауссовского распределения (см. §8.1, лекция 14). С другой стороны,

$$\mathsf{P}(X_1 + X_2 = 0) = \mathsf{P}(\beta = -1) = \frac{1}{2},$$

откуда мы заключаем, что $X_1 + X_2$ негауссовская случайная величина. [1] В силу свойства (iii) это означает, что вектор $\mathbf{X}$ не гауссовский.

1.3. Ковариационная матрица. Поскольку координаты двумерного гауссовского вектора являются гауссовскими случайными величинами, то они имеют математические ожидания и дисперсии. Пусть $\boldsymbol{\mu}$ — вектор-столбец математических ожиданий вектора $\mathbf{X}$: $\boldsymbol{\mu} = \begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}$, где $\mu_1 = \mathsf{E}[X_1]$ и $\mu_2 = \mathsf{E}[X_2]$. Положим $\sigma_1^2 = \mathsf{var}[X_1]$ и $\sigma_2^2 = \mathsf{var}[X_2]$, а ковариацию между X_1 и X_2 обозначим через $\rho = \mathsf{E}[(X_1 - \mu_1)(X_2 - \mu_2)]$. Матрица

$$\boldsymbol{\Lambda} \stackrel{\text{def}}{=} \begin{pmatrix}\mathsf{cov}[X_1, X_1] & \mathsf{cov}[X_1, X_2]\\ \mathsf{cov}[X_2, X_1] & \mathsf{cov}[X_2, X_2]\end{pmatrix} \stackrel{\text{(a)}}{=} \begin{pmatrix}\sigma_1^2 & \rho\\ \rho & \sigma_2^2\end{pmatrix}$$

[1]почему?

называется *ковариационной матрицей* вектора $\mathbf{X}$. ②[2]

Пусть $Q(z_1, z_2) \stackrel{\text{def}}{=} \mathbf{z}'\boldsymbol{\Lambda}\mathbf{z}$, $\mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$. Такое выражение называется *квадратичной формой* от $\mathbf{z}$. Квадратичную форму можно записать в явном виде:

$$\begin{aligned} Q(z_1, z_2) &= \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}' \begin{pmatrix} \sigma_1^2 & \rho \\ \rho & \sigma_2^2 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} \sigma_1^2 z_1 + z_2\rho \\ \rho z_1 + \sigma_2^2 z_2 \end{pmatrix}' \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \\ &= \sigma_1^2 z_1^2 + 2\rho z_1 z_2 + \sigma_2^2 z_2^2. \end{aligned}$$

Утверждение 1. *Квадратичная форма Q является неотрицательно определенной, то есть $Q(z_1, z_2) \geq 0$ для любых z_1 и z_2.*

Замечание 1. Матрица $\boldsymbol{\Lambda}$, для которой $Q(\mathbf{z}) > 0$ для всех $\mathbf{z}$ называется *положительно определенной.*

Доказательство. Мы рассмотрим случай $\sigma_2^2 \neq 0$. Утверждение очевидно, если $z_1 = 0$ и $z_2 = 0$. Если же, например, $z_1 \neq 0$, то

$$Q(z_1, z_2) = z_1^2 \left(\sigma_1^2 + 2\rho t + \sigma_2^2 t^2\right), \qquad t = z_2/z_1.$$

Квадратный трехчлен в скобках представим в виде

$$\sigma_1^2 + 2\rho t + \sigma_2^2 t^2 = \left(\sigma_2 t + \frac{\rho}{\sigma_2}\right)^2 + \sigma_1^2 - \frac{\rho^2}{\sigma_2^2}. \tag{1}$$

Теперь видно, что $Q(z_1, z_2) \geq 0$, так как $\rho^2 \leq \sigma_1^2 \sigma_2^2$ в силу неравенства Коши–Буняковского. □

1.4. Неравенство Коши–Буняковского. Последний шаг в доказательстве утверждения 1 опирается на неравенство

$$\left(\mathsf{E}\left[\xi\eta\right]\right)^2 \leq \mathsf{E}\left[\xi^2\right] \cdot \mathsf{E}\left[\eta^2\right], \tag{2}$$

которое верно для любых случайных величин ξ и η с конечными вторыми моментами.

Для доказательства неравенства (2) заметим, что $(s\xi+\eta)^2 \geq 0$ для любого s. Поэтому $\mathsf{E}\left[(s\xi+\eta)^2\right] \geq 0$. Значит $\mathsf{E}\left[\xi^2\right] s^2 + 2\mathsf{E}\left[\xi\eta\right] s + \mathsf{E}\left[\eta^2\right] \geq 0$ для любого s. Следовательно, дискриминант неположителен: $(\mathsf{E}\left[\xi\eta\right])^2 - \mathsf{E}\left[\xi^2\right] \mathsf{E}\left[\eta^2\right] \leq 0$, что и требовалось доказать.

Задача* 1. *Как доказать утверждение 1 при $\sigma_2^2 = 0$?*

Задача* 2. *Может ли выполняться равенство $Q(z_1, z_2) = 0$ и для каких $\mathbf{z}$?*

[2]объяснить (a)

2. Многомерный гауссовский случайный вектор

Определение 2. Случайный вектор $\mathbf{X}$ размерности n называется *n-мерным гауссовским случайным вектором*, если случайная величина $c_1X_1 + \cdots + c_nX_n$ является гауссовской при любых $c_1, \ldots, c_n$: $|c_1| + \cdots + |c_n| \neq 0$.

Как и для двумерных векторов справедливы такие свойства:

(G_1) гауссовские векторы существуют для любой размерности n: например, вектор $\mathbf{X}$, составленный из независимых стандартных гауссовских случайных величин $X_1, \ldots, X_n$, является гауссовским;

(G_2) любая из случайных величин $X_1, \ldots, X_n$ является гауссовской, если $\mathbf{X}$ — гауссовский случайный вектор;

(G_3) если $\mathbf{X}$ — гауссовский случайный вектор, то $X_1 + \cdots + X_n$ является гауссовской случайной величиной;

(G_4) любое маргинальное распределение вектора $\mathbf{X}$ является гауссовским.

Напомним, что маргинальным распределением случайного вектора $\mathbf{X} = (X_1, \ldots, X_n)'$ называется распределение его подвектора $(X_{i_1}, \ldots, X_{i_m})'$, $m < n$.

Задача 3. *Доказать, что случайный вектор размерности n имеет $2^n - 2$ маргинальных распределений.*

Задача 4. *Доказать свойства (G_1)–(G_3).*

Доказательство свойства (G_4). Ограничимся случаем маргинального распределения случайного подвектора, составленного из первых $m < n$ координат вектора $\mathbf{X}$. Докажем, что оно является m-мерным гауссовским. Действительно, пусть $c_1, \ldots, c_m$ — произвольные константы, $|c_1| + \cdots + |c_m| \neq 0$. Тогда $c_1X_1 + \cdots + c_mX_m$ является гауссовской случайной величиной, так как $c_1X_1 + \cdots + c_mX_m = c_1X_1 + \cdots + c_mX_m + 0 \cdot X_{m+1} + \cdots + 0 \cdot X_n$. □

2.1. Векторные обозначения. Векторная нотация более экономна при изучении гауссовских векторов. Например, $c_1X_1 + \cdots + c_nX_n = \mathbf{c}'\mathbf{X}$. В дальнейшем мы используем векторные обозначения и матричные операции. Векторы понимаются как вектор-столбцы. Если матрица имеет m строк и n столбцов, то говорим о $m \times n$-матрице. Транспонирование матрицы $\mathbf{B}$ обозначаем $\mathbf{B}'$. В частности, если $\mathbf{X}$ — это n-вектор (то есть $n \times 1$-матрица [3]), то $\mathbf{XX}'$ — это $n \times n$-матрица. [4]

Если все элементы b_{ij} матрицы $\mathbf{B}$ являются случайными величинами, то мы говорим о *случайной матрице*. Если b_{ij} — это элементы матрицы $\mathbf{B}$ и все они имеют математические ожидания, то через $\mathsf{E}[\mathbf{B}]$ мы обозначаем матрицу, составленную из $\mathsf{E}[b_{ij}]$. В частности, если $\mathbf{X}$ — это случайный n-вектор (случайная $n \times 1$-матрица), то $\mathsf{E}[\mathbf{X}]$ — это вектор составленный из $\mathsf{E}[X_i]$.

Лемма 1. *Пусть $\mathbf{X}$ — это случайный n-вектор, а $\mathbf{b}$ — неслучайный n-вектор. Тогда $\mathsf{E}[\mathbf{X} + \mathbf{b}] = \mathsf{E}[\mathbf{X}] + \mathbf{b}$.*

Доказательство. Вытекает из линейности математического ожидания случайных величин. [5] □

---

[3] почему не $1 \times n$ матрица?

[4] проверить!

[5] доказать!

Лемма 2. *Пусть* $\mathbf{A}$ *— случайная матрица, а* $\mathbf{B}$ *и* $\mathbf{C}$ *— неслучайные матрицы, согласованных с* $\mathbf{A}$ *размерностей. Тогда* $\mathsf{E}[\mathbf{BA}] = \mathbf{B}\cdot\mathsf{E}[\mathbf{A}]$ *и* $\mathsf{E}[\mathbf{AC}] = \mathsf{E}[\mathbf{A}]\cdot\mathbf{C}$*. В частности, если* $\mathbf{X}$ *— случайный вектор,* $\mathbf{B}$ *— неслучайная матрица и* $\boldsymbol{\mu} = \mathsf{E}[\mathbf{X}]$*, то* $\mathsf{E}[\mathbf{BX}] = \mathbf{B}\boldsymbol{\mu}$*.*

Доказательство. Обозначим элементы (i, j) матриц $\mathbf{BA}$ и $\mathbf{B}\cdot\mathsf{E}[\mathbf{A}]$ через ξ_{ij} и e_{ij} соответственно. Тогда в силу линейности математического ожидания для случайных величин

$$\xi_{ij} = \sum_\nu b_{i\nu}a_{\nu j}, \quad e_{ij} = \sum_\nu b_{i\nu}\mathsf{E}[a_{\nu j}], \quad \text{откуда} \quad \mathsf{E}[\xi_{ij}] = e_{ij}.$$

Это и доказывает первое утверждение леммы 2. $\square$

Задача 5. *Доказать второе утверждение леммы 2.*

2.2. Ковариационная матрица. Пусть $\mathbf{X}$ — случайный n-вектор. Предположим, что существуют вторые моменты у всех случайных величин X_i (то есть у всех координат вектора $\mathbf{X}$).

Определение 3. Матрица $\mathsf{Cov}[\mathbf{X}]$, составленная из ковариаций $\mathsf{cov}[X_i, X_j]$, $1 \le i, j \le n$, называется *ковариационной матрицей* вектора $\mathbf{X}$.

Лемма 3. *Пусть* $\mathbf{X}$ *— случайный* n*-вектор. Обозначим* $\mathsf{E}[\mathbf{X}] = \boldsymbol{\mu}$ *и пусть* $\boldsymbol{\Lambda}$ *— ковариационная матрица вектора* $\mathbf{X}$*. Тогда* $\boldsymbol{\Lambda} = \mathsf{E}[(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})']$*. В частности,* $\boldsymbol{\Lambda}$ *является симметричной матрицей.*

Напомним, что свойство симметричности матрицы $\boldsymbol{\Lambda} = \left(\lambda_{ij}\right)_{i,j=1}^n$ означает, что $\lambda_{ij} = \lambda_{ji}$.

Доказательство. Поскольку

$$(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})' = \begin{pmatrix} X_1-\mu_1 \\ X_2-\mu_2 \\ \dots \\ X_n-\mu_n \end{pmatrix} (\, X_1-\mu_1, X_2-\mu_2, \dots, X_n-\mu_n \,)$$

$$= \begin{pmatrix} (X_1-\mu_1)^2 & (X_1-\mu_1)(X_2-\mu_2) & \dots & (X_1-\mu_1)(X_n-\mu_n) \\ (X_2-\mu_2)(X_1-\mu_1) & (X_2-\mu_2)^2 & \dots & (X_2-\mu_2)(X_n-\mu_n) \\ \dots & \dots & \dots & \dots \\ (X_n-\mu_n)(X_1-\mu_1) & (X_n-\mu_n)(X_2-\mu_2) & \dots & (X_n-\mu_n)^2 \end{pmatrix},$$

то $\mathsf{Cov}[\mathbf{X}] = \mathsf{E}[(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})']$. Симметричность ковариационной матрицы вытекает из ее определения. $\square$

Теорема 1. *Ковариационная матрица является неотрицательно определенной, то есть* $Q(\mathbf{z}) \overset{\text{def}}{=} \mathbf{z}'\boldsymbol{\Lambda}\mathbf{z} \ge 0$ *для любого неслучайного вектора* $\mathbf{z}$*.*

Доказательство. Из лемм 2 и 3 вытекает, что

$$\mathbf{z}'\boldsymbol{\Lambda}\mathbf{z} = \mathbf{z}'\mathsf{E}[(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})']\,\mathbf{z} = \mathsf{E}[\mathbf{z}'(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})'\mathbf{z}]\,.$$

Для любых векторов $\mathbf{a}$ и $\mathbf{b}$ выполнено $(\mathbf{a}'\mathbf{b})' = \mathbf{b}'\mathbf{a}$. ⑥ Поэтому

$$\mathbf{z}'\boldsymbol{\Lambda}\mathbf{z} = \mathsf{E}[\boldsymbol{\xi}'\boldsymbol{\xi}]\,, \qquad \boldsymbol{\xi} \overset{\text{def}}{=} (\mathbf{X}-\boldsymbol{\mu})'\mathbf{z}.$$

[6]проверить!

Поскольку $\boldsymbol{\xi}'\boldsymbol{\xi} = \xi_1^2 + \cdots + \xi_n^2$, то $\mathbf{z}'\mathbf{\Lambda}\mathbf{z} \geq 0$, что и требовалось доказать. □

Замечание 2. Утверждение 1 (частный случай теоремы 1 для $n = 2$) было доказано с помощью неравенства Коши–Буняковского (см. §1.4). В свою очередь, неравенство Коши–Буняковского вытекает из теоремы 1. Действительно, пусть $n = 2$. Тогда, как доказано в теореме 1, $Q(z_1, z_2) \geq 0$ для любых z_1, z_2. Квадратичная форма $Q(z_1, z_2)$ имеет вид (1). Подставляя $t = -\frac{\rho}{\sigma_2^2}$ в (1), видим, что $\sigma_1^2 - \frac{\rho^2}{\sigma_2^2} \geq 0$. А это и есть неравенство Коши–Буняковского (2).

Следствие 1. *Пусть* $\mathbf{X}$ *— случайный* n*-вектор,* $\mathbf{b}$ *— неслучайный* n*-вектор. Тогда* $\mathsf{Cov}\,[\mathbf{X} + \mathbf{b}] = \mathsf{Cov}\,[\mathbf{X}]$.

Доказательство. Обозначим $\mathbf{Y} \stackrel{\text{def}}{=} \mathbf{X} + \mathbf{b}$. Тогда матрица $\mathsf{Cov}\,[\mathbf{X} + \mathbf{b}]$ составлена из чисел $\mathsf{cov}\,[Y_i, Y_j]$, а матрица $\mathsf{Cov}\,[\mathbf{X}]$ — из $\mathsf{cov}\,[X_i, X_j]$. Так как $\mathsf{cov}\,[Y_i, Y_j] = \mathsf{cov}\,[X_i, X_j]$, ⑦ то следствие доказано. □

Лемма 4. *Пусть* $\mathbf{X}$ *— случайный* n*-вектор,* $\mathbf{B}$ *— неслучайная* $m \times n$*-матрица. Обозначим* $\boldsymbol{\mu} = \mathsf{E}\,[\mathbf{X}]$, $\mathbf{\Lambda} = \mathsf{Cov}\,[\mathbf{X}]$. *Тогда* $\mathbf{Y} = \mathbf{B}\mathbf{X}$ *— случайный* m*-вектор, для которого* $\mathsf{E}\,[\mathbf{Y}] = \mathbf{B}\boldsymbol{\mu}$, $\mathsf{Cov}\,[\mathbf{Y}] = \mathbf{B}\mathbf{\Lambda}\mathbf{B}'$.

Доказательство. Утверждение о математическом ожидании вытекает из леммы 2. Она же вместе с леммой 3 позволяет доказать и утверждение о ковариационной матрице:

$$\mathsf{Cov}\,[\mathbf{Y}] = \mathsf{E}\,[(\mathbf{Y} - \mathsf{E}\,[\mathbf{Y}])(\mathbf{Y} - \mathsf{E}\,[\mathbf{Y}])'] = \mathsf{E}\,[(\mathbf{B}\mathbf{X} - \mathbf{B}\boldsymbol{\mu})(\mathbf{B}\mathbf{X} - \mathbf{B}\boldsymbol{\mu})']$$
$$\stackrel{(c_1)}{=} \mathsf{E}\,[(\mathbf{B}(\mathbf{X} - \boldsymbol{\mu}))(\mathbf{X} - \boldsymbol{\mu})'\mathbf{B}'] \stackrel{(c_2)}{=} \mathbf{B}\mathbf{\Lambda}\mathbf{B}'. \qquad \Box$$

⑧

2.3. Линейные преобразования гауссовских векторов. Если гауссовский вектор $\mathbf{X}$ имеет вектор средних $\boldsymbol{\mu}$ и матрицу ковариаций $\mathbf{\Lambda}$, то мы пишем $\mathbf{X} \in \mathcal{N}\,(\boldsymbol{\mu}, \mathbf{\Lambda})$. Иногда мы также пишем $\mathbf{X} \in \mathcal{N}_n\,(\boldsymbol{\mu}, \mathbf{\Lambda})$, чтобы отметить размерность векторов $\mathbf{X}$, $\boldsymbol{\mu}$ и матрицы $\mathbf{\Lambda}$.

Заметим, что математические ожидания и ковариации координат гауссовского вектора существуют в силу свойства (G_2).

Теорема 2. *Пусть* $\mathbf{B}$ *матрица размерности* $m \times n$, $\mathbf{b}$ *— вектор размерности* m. *Если случайный* n*-вектор* $\mathbf{X}$ *является гауссовским, причем* $\mathbf{X} \in \mathcal{N}_n\,(\boldsymbol{\mu}, \mathbf{\Lambda})$, *то случайный* m*-вектор* $\mathbf{Y} = \mathbf{B}\mathbf{X} + \mathbf{b}$ *также является гауссовским, причем* $\mathbf{Y} \in \mathcal{N}_m\,(\mathbf{B}\boldsymbol{\mu} + \mathbf{b}, \mathbf{B}\mathbf{\Lambda}\mathbf{B}')$.

⑨

Доказательство. Сначала докажем, что $\mathbf{Y}$ является гауссовским вектором. Для любого вектора $\mathbf{c}'$ имеем $\mathbf{c}'\mathbf{Y} = \mathbf{c}'\mathbf{B}\mathbf{X} + \mathbf{c}'\mathbf{b} = \mathbf{u}'\mathbf{X} + v$, где $\mathbf{u} = \mathbf{B}'\mathbf{c}$ и $v = \mathbf{c}'\mathbf{b}$. ⑩ Случайная величина $\mathbf{u}'\mathbf{X}$ является гауссовской по определению 2, а v — действительной константой. Поэтому и $\mathbf{c}'\mathbf{Y}$ является гауссовской случайной величиной. Утверждение о математическом ожидании вытекает из лемм 1

[7]доказать!
[8]почему выполнены (c_1) и (c_2)?
[9]Почему вектор $\mathbf{Y}$ имеет размерность m?
[10]почему?

и 2. Следствие 1 и лемма 4 доказывают утверждение о ковариационной матрице. □

† Всего в тексте было 10 вопросов

Лекция 17

ХАРАКТЕРИСТИЧЕСКАЯ ФУНКЦИЯ ГАУССОВСКОГО СЛУЧАЙНОГО ВЕКТОРА

1. Характеристическая функция случайного вектора

Определение 1. Пусть $\mathbf{X}$ — случайный n-вектор. *Характеристической функцией случайного вектора* $\mathbf{X}$ называется

$$h(\mathbf{t}) = h_{\mathbf{X}}(\mathbf{t}) = \mathsf{E}\left[e^{i\mathbf{t}'\mathbf{X}}\right] = \mathsf{E}\left[e^{i(t_1X_1+\cdots+t_nX_n)}\right], \qquad \mathbf{t} \in \mathbf{R}^n. \tag{1}$$

Пример 1. Пусть координаты вектора $\mathbf{X}$ являются независимыми случайными величинами. Поскольку преобразования его координат $e^{it_1X_1}, \ldots, e^{it_nX_n}$ являются независимыми для любых фиксированных действительных чисел $t_1, \ldots, t_n$, то

$$\mathsf{E}\left[e^{i(t_1X_1+\cdots+t_nX_n)}\right] = \mathsf{E}\left[e^{it_1X_1}\right] \ldots \mathsf{E}\left[e^{it_nX_n}\right]$$

и поэтому

$$h_{\mathbf{X}}(\mathbf{t}) = h_{X_1}(t_1) \ldots h_{X_n}(t_n).$$

Таким образом, характеристическая функция случайного вектора с независимыми координатами равна произведению характеристических функций его координат.

Вычисление характеристической функции вектора с зависимыми координатами часто представляет сложную задачу, но ее можно решить для гауссовских случайных векторов. Сначала мы напомним вид характеристической функции гауссовской случайной величины.

1.1. Характеристическая функция гауссовской случайной величины. Согласно определению характеристической функции, $h_\xi(t) = \mathsf{E}\left[e^{it\xi}\right]$ для любой случайной величины ξ.

Лемма 1. *Пусть ξ — гауссовская случайная величина с параметрами a и σ^2. Тогда ее характеристическая функция $h(t)$ равна*

$$h(t) = e^{iat - t^2\sigma^2/2}. \tag{2}$$

Доказательство. Пусть сначала $a = 0$ и $\sigma^2 = 1$. Согласно определению

$$h(t) = \int_{-\infty}^{\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx. \tag{3}$$

Дифференцируя обе части этого равенства по t, получаем

$$h'(t) = \frac{i}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{itx} e^{-x^2/2}\, dx.$$

[0] Printed from the file [17-Gauss-cf.tex] on 9.6.2016

Дифференцирование под знаком интеграла разрешено, поскольку интеграл $\int_{-\infty}^{\infty} xe^{itx}e^{-x^2/2}\,dx$ сходится равномерно (см. Фихтенгольц, т. II, глава 520, §712, стр. теорема 3). [1] Проинтегрируем теперь по частям интеграл для производной:

$$\int_{-\infty}^{\infty} xe^{itx}e^{-x^2/2}\,dx = -\int_{-\infty}^{\infty} e^{itx}\,d\left(e^{-x^2/2}\right)$$
$$= -\,e^{itx}e^{-x^2/2}\Big|_{-\infty}^{\infty} + it\int_{-\infty}^{\infty} e^{itx}e^{-x^2/2}\,dx,$$

откуда $h'(t) = -th(t)$. [2] Решением этого дифференциального уравнение является $h(t) = Ce^{-t^2/2}$ при любой константе C. Поскольку $h(0) = 1$, то $C = 1$ и $h(t) = e^{-t^2/2}$. [3]

Возвращаясь к общему случаю, рассмотрим случайные величины $\xi \in \mathcal{N}(0,1)$ и $\eta \stackrel{\text{def}}{=} \sigma\xi + a$. Тогда $\eta \in \mathcal{N}\left(a,\sigma^2\right)$ и $h_\eta(t) = e^{iat}h_\xi(\sigma t)$. С учетом уже рассмотренного случая $\mathcal{N}(0,1)$ это и влечет (2). $\square$

1.2. Х. ф. гауссовского случайного вектора. Логарифм характеристической функции гауссовской случайной величины является полиномом второй степени. Аналогичное свойство верно и для гауссовских векторов. Сначала мы проверим это свойство для векторов с независимыми координатами.

Пример 2. Пусть $\mathbf{X}$ — это гауссовский вектор с независимыми координатами. Согласно свойству (G_2) (лекция 16, стр. 125) его координаты являются гауссовскими случайными величинами. Обозначим их параметры через $\mu_1, \sigma_1^2, \ldots, \mu_n, \sigma_n^2$, то есть $X_k \in \mathcal{N}\left(\mu_k, \sigma_k^2\right)$, $1 \le k \le n$. Тогда

$$h_{\mathbf{X}}(\mathbf{t}) = h_{X_1}(t_1)\ldots h_{X_n}(t_n) = e^{i\mu_1 - \frac{\sigma_1^2}{2}t_1^2}\cdot\ldots\cdot e^{i\mu_n - \frac{\sigma_n^2}{2}t_n^2}.$$

Если обозначить $\boldsymbol{\mu}' = (\mu_1, \ldots, \mu_n)$ и

$$\boldsymbol{\Lambda} = \begin{pmatrix} \sigma_1^2 & 0 & 0 & \ldots & 0 & 0 \\ 0 & \sigma_2^2 & 0 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 & \sigma_n^2 \end{pmatrix},$$

то получим $h_{\mathbf{X}}(\mathbf{t}) = e^{i\mathbf{t}'\boldsymbol{\mu} - \frac{1}{2}\mathbf{t}'\boldsymbol{\Lambda}\mathbf{t}}$. [4]

Теорема 1. *Если* $\mathbf{X} \in \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Lambda})$, *то* $h_{\mathbf{X}}(\mathbf{t}) = e^{i\mathbf{t}'\boldsymbol{\mu} - \frac{1}{2}\mathbf{t}'\boldsymbol{\Lambda}\mathbf{t}}$.

Доказательство. Предположим, что $\mathbf{X} \in \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Lambda})$. Зафиксируем действительные числа $t_1, \ldots, t_n$, для которых $|t_1| + \cdots + |t_n| > 0$. В соответствии с определением 16.2 случайная величина $Z \stackrel{\text{def}}{=} \mathbf{t}'\mathbf{X}$ в равенстве (1) является

[1] Убедиться, что дифференцирование под знаком интеграла в (3) допустимо.

[2] Объяснить равенство $e^{itx}e^{-x^2/2}\Big|_{-\infty}^{\infty} = 0$.

[3] почему?

[4] Проверить формулу $h_{\mathbf{X}}(\mathbf{t}) = e^{i\mathbf{t}'\boldsymbol{\mu} - \frac{1}{2}\mathbf{t}'\boldsymbol{\Lambda}\mathbf{t}}$ для гауссовских векторов с независимыми координатами.

гауссовской. В силу теоремы 16.2 с $\mathbf{B} = \mathbf{t}'$ и $\mathbf{b} = \mathbf{0}$ параметры этой случайной величины равны $\mathsf{E}[Z] = \mathbf{t}'\boldsymbol{\mu} \stackrel{\mathrm{def}}{=} a$ и $\mathsf{var}[Z] = \mathbf{t}'\boldsymbol{\Lambda}\mathbf{t} \stackrel{\mathrm{def}}{=} \sigma^2$. Поскольку характеристическая функция гауссовской случайной величины Z равна $h_Z(u) \stackrel{\mathrm{def}}{=} \mathsf{E}\left[e^{iuZ}\right] = e^{iua-\frac{1}{2}u^2\sigma^2}$ (см. лемму 1) и $h_{\mathbf{X}}(\mathbf{t}) = h_Z(1)$, то теорема доказана. $\square$

Теперь мы покажем, что только гауссовские векторы имеют характеристические функции вида $e^{i\mathbf{t}'\boldsymbol{\mu}-\frac{1}{2}\mathbf{t}'\boldsymbol{\Lambda}\mathbf{t}}$. Для этого напомним некоторые полезные сведения из алгебры.

2. Диагонализация неотрицательно определенной матрицы

Квадратная $n \times n$ матрица $\mathbf{A} = (a_{ij})$ называется *симметричной*, если $\mathbf{A}' = \mathbf{A}$. Это условие можно записать иначе: $a_{ij} = a_{ji}$ для любых $1 \le i, j \le n$.

Матрица $\mathbf{C}$ называется *ортогональной*, если $\mathbf{C}'\mathbf{C} = \mathbf{I}$ или, иными словами, если $\mathbf{C}' = \mathbf{C}^{-1}$. Тут $\mathbf{I}$ — единичная матрица.

Для нас важным является такой факт:

Утверждение 1. *Если $\mathbf{A}$ — симметричная матрица, то существует диагональная матрица $\mathbf{D}$ и ортогональная матрица $\mathbf{C}$, для которых $\mathbf{A} = \mathbf{C}\mathbf{D}\mathbf{C}'$, причем диагональ матрицы $\mathbf{D}$ составлена из собственных чисел матрицы $\mathbf{A}$.*

Замечание 1. Напомним, что *собственным числом матрицы* $\mathbf{A}$ называется любое (комплексное) число λ, для которого система уравнений $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ имеет нетривиальное решение. Вектор $\mathbf{x}$, который отвечает собственному числу λ, называется *собственным вектором матрицы* $\mathbf{A}$. Отметим, что если $\mathbf{x}$ — собственный вектор матрицы $\mathbf{A}$, то $c\mathbf{x}$ также ее собственный вектор каким бы ни было действительное число c. ⑤

Если систему, которая определяет собственные числа, переписать в виде $(\mathbf{A}-\lambda\mathbf{I})\mathbf{x} = 0$, то легко найти условие существования нетривиального решения, а именно $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$. ⑥ Правая часть последнего равенства является полиномом степени n от переменной λ. ⑦ По основной теореме алгебры это, в частности, означает, что $n \times n$ матрица $\mathbf{A}$ имеет ровно n собственных чисел $\lambda_1, \ldots, \lambda_n$ (возможно кратных или комплексных). ⑧

Лемма 2. *Пусть $\mathbf{A}$ — симметричная вещественная матрица, а λ — некоторое собственное число матрицы $\mathbf{A}$. Тогда $\lambda \in \mathbf{R}$. Иными словами, любое собственное число симметричной неотрицательно определенной матрицы является действительным.*

Доказательство. По определению собственного числа, матричное уравнение $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ имеет нетривиальное решение $\mathbf{x}$; само число λ является решением уравнения $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$. Это алгебраическое уравнение степени n, поэтому если λ является его корнем, то комплексно сопряженное число $\mu = \overline{\lambda}$ также его корень. Следовательно $\mathbf{A}\mathbf{y} = \mu\mathbf{y}$ для некоторого вектора $\mathbf{y}$. Больше того, $A\overline{\mathbf{x}} = \overline{\mathbf{A}\mathbf{x}} = \overline{\lambda\mathbf{x}} = \mu\overline{\mathbf{x}}$. ⑨ Иными словами, $\overline{\mathbf{x}}$ является собственным вектором

[5] почему?

[6] Почему, если $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$, то система $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ имеет нетривиальное решение?

[7] объяснить!

[8] доказать!

[9] Объяснить равенство $\overline{\mathbf{A}\mathbf{x}} = \mathbf{A}\overline{\mathbf{x}}$ для вещественной матрицы $\mathbf{A}$.

матрицы $\mathbf{A}$, отвечающим собственному значению μ. Именно этот вектор мы и выбираем в качестве $\mathbf{y}$. Итак,

$$\overline{\mathbf{x}}'\mathbf{A}\mathbf{x} = \lambda\overline{\mathbf{x}}'\mathbf{x} \quad \text{и} \quad \mathbf{x}'\mathbf{A}\overline{\mathbf{x}} = \mu\mathbf{x}'\overline{\mathbf{x}}.$$

В силу симметричности матрицы $\mathbf{A}$ имеем $\mathbf{A} = \mathbf{A}'$ и значит

$$\overline{\mathbf{x}}'\mathbf{A}\mathbf{x} = \mathbf{x}'\mathbf{A}\overline{\mathbf{x}},$$

так как

$$\overline{\mathbf{x}}'\mathbf{A}\mathbf{x} = \sum_{i=1}^{n}\sum_{j=1}^{n}\overline{x}_i a_{ij} x_j = \sum_{i=1}^{n}\sum_{j=1}^{n}\overline{x}_i a_{ji} x_j = \mathbf{x}'\mathbf{A}\overline{\mathbf{x}}.$$

Отсюда вытекает, что

$$\lambda\overline{\mathbf{x}}'\mathbf{x} = \mu\mathbf{x}'\overline{\mathbf{x}} = \mu\left(\overline{\mathbf{x}}'\mathbf{x}\right)' = \mu\overline{\mathbf{x}}'\mathbf{x},$$

так как $\overline{\mathbf{x}}'\mathbf{x}$ — действительное число. Таким образом, $(\lambda-\mu)\overline{\mathbf{x}}'\mathbf{x} = 0$. Так как вектор $\mathbf{x}$ невырожденный и $\overline{\mathbf{x}} = \overline{\mathbf{x}}$, то $\overline{\mathbf{x}}'\mathbf{x} > 0$ и поэтому $\lambda = \mu = \overline{\lambda}$, то есть λ — действительное число. $\square$

Лемма 3. *Пусть* $\mathbf{A}$ *— симметричная вещественная матрица,* $\lambda \neq \mu$ *два ее собственных числа,* $\mathbf{x}$ *— собственный вектор для* λ *и* $\mathbf{y}$ *— собственный вектор для* μ. *Тогда* $\mathbf{x}'\mathbf{y} = 0$. *Иными словами, собственные векторы, отвечающие разным собственным числам симметричной вещественной матрицы, являются ортогональными.*

Доказательство. Имеем $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ и $\mathbf{A}\mathbf{y} = \mu\mathbf{y}$. Поэтому $\mathbf{y}'\mathbf{A}\mathbf{x} = \lambda\mathbf{y}'\mathbf{x}$ и $\mathbf{x}'\mathbf{A}\mathbf{y} = \mu\mathbf{x}'\mathbf{y}$. Поскольку $\mathbf{A}$ является симметричной матрицей, то

$$\mathbf{x}'\mathbf{A}\mathbf{y} = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i a_{ij} y_j = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i a_{ji} y_j = \mathbf{y}'\mathbf{A}'\mathbf{x} = \mathbf{y}'\mathbf{A}\mathbf{x}.$$

Поэтому $\mu\mathbf{x}'\mathbf{y} = \lambda\mathbf{y}'\mathbf{x}$. Так как $\mathbf{x}'\mathbf{y}$ является действительным числом, то $\mathbf{y}'\mathbf{x} = (\mathbf{x}'\mathbf{y})' = \mathbf{x}'\mathbf{y}$ и поэтому $(\lambda-\mu)\mathbf{x}'\mathbf{y} = 0$. Это и означает, что $\mathbf{x}'\mathbf{y} = 0$. $\square$

Доказательство утверждения 1. Мы рассмотрим только случай, когда все собственные числа матрицы $\mathbf{A}$ разные. Итак, пусть $\lambda_1, \dots, \lambda_n$ — собственные числа матрицы $\mathbf{A}$, а $\mathbf{x}_1, \dots, \mathbf{x}_n$ — соответствующие собственные векторы. Эти векторы выберем такими, чтобы $\mathbf{x}_i'\mathbf{x}_i = 1$ для всех $1 \le i \le n$. [10] Рассмотрим матрицу $\mathbf{T}$, столбцами которой являются векторы $\mathbf{x}_1, \dots, \mathbf{x}_n$. Тогда строками матрицы $\mathbf{T}'$ являются векторы $\mathbf{x}_1', \dots, \mathbf{x}_n'$. Элементами матрицы $\mathbf{T}'\mathbf{T}$ являются числа $\mathbf{x}_i'\mathbf{x}_j$. В силу леммы 3 имеем $\mathbf{x}_i'\mathbf{x}_j = 0$, $i \neq j$, то есть $\mathbf{T}'\mathbf{T} = \mathbf{I}$. Это означает, что $\mathbf{T}$ — ортогональная матрица.

Заметим теперь, что матрица $\mathbf{AT}$ составлена из столбцов $\mathbf{A}\mathbf{x}_1, \dots, \mathbf{A}\mathbf{x}_n$, то есть из столбцов $\lambda_1\mathbf{x}_1, \dots, \lambda_n\mathbf{x}_n$. Поэтому

$$\mathbf{T}'\mathbf{A}\mathbf{T} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \stackrel{\text{def}}{=} \mathbf{D}.$$

Домножим это матричное равенство слева на $\mathbf{T}$, а справа — на $\mathbf{T}'$ и воспользуемся ортогональностью матрицы $\mathbf{T}$:

$$\mathbf{T}\mathbf{D}\mathbf{T}' = \mathbf{T}(\mathbf{T}'\mathbf{A}\mathbf{T})\mathbf{T}' = (\mathbf{T}\mathbf{T}')\mathbf{A}(\mathbf{T}\mathbf{T}') = \mathbf{A}. \qquad \square$$

[10] почему так можно выбрать собственные векторы?

3. Квадратный корень из матрицы

Матрица $\mathbf{B}$ называется *квадратным корнем* матрицы $\mathbf{A}$, если $\mathbf{B}^2 = \mathbf{A}$. Квадратный корень матрицы $\mathbf{A}$ обозначается $\mathbf{A}^{1/2}$. Квадратный корень существует не для всех матриц. Например, матрица $\mathbf{A} = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$ не имеет квадратного корня. ⑪

Утверждение 2. *Если* $\mathbf{A}$ *— неотрицательно определенная симметричная матрица, то квадратный корень* $\mathbf{A}^{1/2}$ *существует.*

Доказательство. Действительно, так как $\mathbf{A}$ симметричная матрица, то $\mathbf{A} = \mathbf{C}\mathbf{D}\mathbf{C}'$ на основании утверждения 1, причем $\mathbf{C}$ — ортогональная, а $\mathbf{D}$ — диагональная. Так как $\mathbf{A}$ неотрицательно определена, то $\mathbf{x}'\mathbf{A}\mathbf{x} \geq 0$ для любого собственного вектора $\mathbf{x}$. Поэтому $\lambda\mathbf{x}'\mathbf{x} \geq 0$ для любого собственного числа λ и любого собственного вектора $\mathbf{x}$, отвечающего этому числу. Отсюда вытекает, что $\lambda \geq 0$, то есть любое собственное значение симметричной неотрицательно определенной матрицы является неотрицательным числом.

В частности, квадратный корень $\mathbf{D}^{1/2}$ существует: им будет диагональная матрица с элементами $\sqrt{d_i}$ на главной диагонали, ⑫ где d_i — это диагональные элементы матрицы $\mathbf{D}$ (которые, напомним, совпадают с собственными числами матрицы $\mathbf{A}$). Положим $\widetilde{\mathbf{A}} = \mathbf{C}\mathbf{D}^{1/2}\mathbf{C}'$. Тогда в силу ассоциативности умножения матриц и ортогональности $\mathbf{C}$

$$\widetilde{\mathbf{A}}^2 = \widetilde{\mathbf{A}}\widetilde{\mathbf{A}} = \mathbf{C}\mathbf{D}^{1/2}\mathbf{C}'\mathbf{C}\mathbf{D}^{1/2}\mathbf{C}' = \mathbf{C}\mathbf{D}^{1/2}\mathbf{D}^{1/2}\mathbf{C}' = \mathbf{C}\mathbf{D}\mathbf{C}' = \mathbf{A}.$$

Значит $\widetilde{\mathbf{A}}^2 = \mathbf{A}$ и $\mathbf{A}^{1/2}$ действительно существует. □

4. Второе определение гауссовского случайного вектора

Определение 2. Случайный вектор $\mathbf{X}$ называется гауссовским, если его характеристическая функция равна $h_{\mathbf{X}}(\mathbf{t}) = e^{i\mathbf{t}'\boldsymbol{\mu} - \frac{1}{2}\mathbf{t}'\boldsymbol{\Lambda}\mathbf{t}}$, где $\boldsymbol{\mu} \in \mathbf{R}^n$, а $\boldsymbol{\Lambda}$ – симметричная неотрицательно определенная матрица.

Мы покажем, что определения 2 и 16.2 эквивалентны. Согласно теореме 1, любой гауссовский вектор в смысле определения 16.2 является гауссовским в смысле определения 2. Поэтому остается доказать обратное утверждение.

Сначала мы покажем, что $h^*(\mathbf{t}) = e^{i\mathbf{t}'\boldsymbol{\mu} - \frac{1}{2}\mathbf{t}'\boldsymbol{\Lambda}\mathbf{t}}$ является характеристической функцией какими бы ни были вектор $\boldsymbol{\mu}$ и неотрицательно определенная симметричная матрица $\boldsymbol{\Lambda}$.

Лемма 4. *Пусть* $\boldsymbol{\mu}$ *— неслучайный вектор, а* $\boldsymbol{\Lambda}$ *— неотрицательно определенная симметричная матрица. Тогда* $h^*(\mathbf{t}) = e^{i\mathbf{t}'\boldsymbol{\mu} - \frac{1}{2}\mathbf{t}'\boldsymbol{\Lambda}\mathbf{t}}$ *является характеристической функцией некоторого случайного вектора* $\mathbf{X}$, *у которого* $\mathsf{E}[\mathbf{X}] = \boldsymbol{\mu}$ *и* $\mathsf{Cov}[\mathbf{X}] = \boldsymbol{\Lambda}$.

Доказательство. Пусть $\mathbf{Y}$ — это n-вектор, координаты $Y_1, \ldots, Y_n$ которого являются независимыми $\mathcal{N}(0,1)$ (то есть, стандартными гауссовскими) случайными величинами. Определим другой вектор по формуле

$$\mathbf{X} = \boldsymbol{\Lambda}^{1/2}\mathbf{Y} + \boldsymbol{\mu}. \tag{4}$$

[11] доказать!
[12] проверить!

Напомним, что квадратный корень $\mathbf{\Lambda}^{1/2}$ существует согласно утверждению 2. Представим матрицу $\mathbf{\Lambda}$ в виде $\mathbf{\Lambda} = \mathbf{CDC}'$, где $\mathbf{C}$ — симметричная матрица, а $\mathbf{D}$ — диагональная. Поскольку $\mathsf{E}[\mathbf{Y}] = \mathbf{0}$ и $\mathsf{Cov}[\mathbf{Y}] = \mathbf{I}$, [13] то из лемм 16.2 и 16.4 вытекает, что

$$\mathsf{E}[\mathbf{X}] = \boldsymbol{\mu}, \qquad \mathsf{Cov}[\mathbf{X}] = \mathbf{\Lambda}, \tag{5}$$

поскольку $\left(\mathbf{\Lambda}^{1/2}\right)' = \left(\mathbf{CD}^{1/2}\mathbf{C}'\right)' = (\mathbf{C}')'\left(\mathbf{D}^{1/2}\right)'\mathbf{C}' = \mathbf{CD}^{1/2}\mathbf{C}' = \mathbf{\Lambda}^{1/2}$, так как $\mathbf{D}^{1/2}$ — диагональная матрица (см. доказательство утверждения 2). [14] Поскольку любая линейная комбинация независимых гауссовских случайных величин является гауссовской, то $\mathbf{Y}$ является гауссовским вектором (см. теорему 15.3) в смысле определения 16.2. Поэтому из теоремы 16.2 вытекает, что и вектор $\mathbf{X}$ является гауссовским, причем $\mathbf{X} \in \mathcal{N}(\boldsymbol{\mu}, \mathbf{\Lambda})$. Следовательно $h_{\mathbf{X}}(\mathbf{t}) = h^*(\mathbf{t})$ по теореме 1. $\square$

Заметим, что в лемме 4 не доказывается, что любой вектор $\mathbf{X}$, который имеет характеристическую функцию h^*, обязательно является гауссовским в смысле определения 16.2. К доказательству этого результата мы и приступаем.

Лемма 5. *Пусть $\boldsymbol{\mu}$ — вектор, а $\mathbf{\Lambda}$ неотрицательно определенная симметричная матрица. Если характеристической функцией некоторого случайного вектора $\mathbf{X}$ является $h_{\mathbf{X}}(\mathbf{t}) = e^{i\mathbf{t}'\boldsymbol{\mu} - \frac{1}{2}\mathbf{t}'\mathbf{\Lambda}\mathbf{t}}$, то $\mathbf{X} \in \mathcal{N}(\boldsymbol{\mu}, \mathbf{\Lambda})$ в смысле определения* 16.2.

Доказательство. Пусть $\mathbf{c} \in \mathbf{R}^n$. Найдем характеристическую функцию случайной величины $\mathbf{c}'\mathbf{X}$ (линейной комбинации координат вектора $\mathbf{X}$):

$$h_{\mathbf{c}'\mathbf{X}}(u) = \mathsf{E}\left[e^{iu\mathbf{c}'\mathbf{X}}\right] = h_{\mathbf{X}}(u\mathbf{c}) = e^{i(u\mathbf{c})'\boldsymbol{\mu} - \frac{1}{2}(u\mathbf{c})'\mathbf{\Lambda}(u\mathbf{c})} = e^{iua - \frac{1}{2}u^2\sigma^2}, \tag{6}$$

где $a = \mathbf{c}'\boldsymbol{\mu}$, $\sigma^2 = \mathbf{c}'\mathbf{\Lambda}\mathbf{c}$. В силу взаимно однозначного соответствия характеристических функций и функций распределения имеем $\mathbf{c}'\mathbf{X} \in \mathcal{N}\left(a, \sigma^2\right)$. Так как вектор $\mathbf{c}$ произвольный, то $\mathbf{X}$ — гауссовский вектор в смысле определения 16.2. Чтобы найти его параметры, воспользуемся леммой 16.2: так как $\mathsf{E}[\mathbf{c}'\mathbf{X}] = a = \mathbf{c}'\boldsymbol{\mu}$ в силу (6) и $\mathsf{E}[\mathbf{c}'\mathbf{X}] = \mathbf{c}'\mathsf{E}[\mathbf{X}]$ в силу леммы 16.2, то $\mathbf{c}'(\mathsf{E}[\mathbf{X}] - \boldsymbol{\mu}) = 0$. В силу произвольности вектора $\mathbf{c}$ получаем $\mathsf{E}[\mathbf{X}] = \boldsymbol{\mu}$. [15]

Для нахождения ковариационной матрицы используем теорему 16.2: так как $\mathsf{Cov}[\mathbf{c}'\mathbf{X}\mathbf{c}] = \mathbf{c}'\,\mathsf{Cov}[\mathbf{X}]\,\mathbf{c}$ в силу теоремы 16.2 и $\mathsf{Cov}[\mathbf{c}'\mathbf{X}\mathbf{c}] = \mathbf{c}'\mathbf{\Lambda}\mathbf{c}$ в силу (6), то $\mathbf{c}'(\mathsf{Cov}[\mathbf{X}] - \mathbf{\Lambda})\mathbf{c} = \mathbf{O}$ (тут $\mathbf{O}$ — нулевая матрица). В силу произвольности вектора $\mathbf{c}$ получаем $\mathsf{Cov}[\mathbf{X}] = \mathbf{\Lambda}$. [16] $\square$

4.1. Эквивалентность двух определений. Теперь мы легко доказываем эквивалентность двух определений гауссовского вектора.

Теорема 2. *Случайный вектор $\mathbf{X}$ является гауссовским в смысле определения* 16.2 *тогда и только тогда, когда его характеристическая функция имеет вид $h_{\mathbf{X}}(\mathbf{t}) = e^{i\mathbf{t}'\boldsymbol{\mu} - \frac{1}{2}\mathbf{t}'\mathbf{\Lambda}\mathbf{t}}$, где $\boldsymbol{\mu} \in \mathbf{R}^n$, а $\mathbf{\Lambda}$ – симметричная неотрицательно определенная матрица.*

---

[13] показать!

[14] Почему $\left(\mathbf{D}^{1/2}\right)' = \mathbf{D}^{1/2}$?

[15] объяснить!

[16] объяснить!

Доказательство. Действительно, если $\mathbf{X}$ — гауссовский вектор в смысле определения 16.2, то из теоремы 1 вытекает, что его характеристическая функция имеет необходимый вид. Матрица $\mathbf{\Lambda}$ неотрицательно определена в силу теоремы 16.1. Ее симметричность вытекает из леммы 16.3.

Обратное утверждение теоремы 2 вытекает из леммы 5. $\square$

4.2. Построение гауссовского вектора из простейшей формы. Как видно из доказательства леммы 4, общий $\mathcal{N}(\boldsymbol{\mu}, \mathbf{\Lambda})$ гауссовский вектор $\mathbf{X}$ можно построить из $\mathcal{N}(\mathbf{0}, \mathbf{I})$ вектора $\mathbf{Y}$ с помощью линейного преобразования (4), причем координаты $\mathbf{Y}$ являются независимыми гауссовскими случайными величинами.

Следствие 1. *Пусть $X_1, \ldots, X_n$ независимые стандартные гауссовские случайные величины. Тогда вектор $\mathbf{X}$ является $\mathcal{N}(\mathbf{0}, \mathbf{I})$-вектором.*

Доказательство. Так как $\mathsf{E}\left[e^{i(t_1 X_1 + \cdots + t_n X_n)}\right] = e^{-\frac{1}{2}(t_1^2 + \cdots + t_n^2)} = e^{-\frac{1}{2}\mathbf{t}'\mathbf{I}\mathbf{t}}$ в силу независимости, то следствие 1 вытекает из теоремы 2. $\square$

[†]Всего в тексте было 16 вопросов

Лекция 18

ПЛОТНОСТЬ ГАУССОВСКОГО СЛУЧАЙНОГО ВЕКТОРА

1. Характеристическая функция двумерного гауссовского случайного вектора

Напомним, что двумерный вектор $\mathbf{X}' = (X_1, X_2)$ называют (см. определение 15.1) *стандартным гауссовским*, если для некоторого $-1 < \rho < 1$ его плотность $f_\rho(x,y)$ имеет вид

$$f_\rho(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}(x^2-2\rho xy+y^2)}, \qquad (x,y) \in \mathbf{R}^2. \tag{1}$$

Гауссовские векторы в определениях 16.1 и 17.2 определялись иначе, а именно с помощью свойства гауссовости любой невырожденной линейной комбинации его координат или с помощью характеристической функции вектора. В §4.2, лекция 15, мы отметили, что вектор с плотностью (1) является гауссовским в смысле определения 16.1. Сейчас мы найдем его характеристическую функцию и увидим, что она именно такая, какой является характеристическая функция гауссовского вектора в смысле определения 17.2. Это и неудивительно, так как определения 16.1 и 17.2 эквивалентны (см. §4.1, лекция 17).

Поскольку $x^2 - 2\rho xy + y^2 = (x-\rho y)^2 + (1-\rho^2)y^2$, то

$$f_\rho(x,y) = \frac{1}{\sqrt{2\pi(1-\rho^2)}} e^{-\frac{1}{2(1-\rho^2)}(x-\rho y)^2} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} = \varphi_{\rho y, 1-\rho^2}(x)\varphi_{0,1}(y). \tag{2}$$

Аналогично $f_\rho(x,y) = \varphi_{\rho x, 1-\rho^2}(y)\varphi_{0,1}(x)$.

Применим дважды формулу для характеристической функции гауссовской случайной величины (лемма 17.1): для $\mathbf{t}' = (t_1, t_2)$ имеем

$$\begin{aligned}
h_{\mathbf{X}}(\mathbf{t}) &= \mathsf{E}\left[e^{i(t_1X_1+t_2X_2)}\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i(t_1x+t_2y)} f_\rho(x,y)\,dx\,dy \qquad ① \\
&= \int_{-\infty}^{\infty} e^{it_2y}\varphi_{0,1}(y) \underbrace{\left[\int_{-\infty}^{\infty} e^{it_1x}\varphi_{\rho y,1-\rho^2}(x)\,dx\right]}_{\text{х.ф. } \mathcal{N}(\rho y, 1-\rho^2) \text{ с.в. для аргумента } t_1} dy \\
&= \int_{-\infty}^{\infty} e^{it_2y}\varphi_{0,1}(y) e^{it_1\rho y - \frac{1}{2}t_1^2(1-\rho^2)}\,dy \\
&= e^{-\frac{1}{2}t_1^2(1-\rho^2)} \underbrace{\int_{-\infty}^{\infty} e^{iy(t_2+t_1\rho)}\varphi_{0,1}(y)\,dy}_{\text{х.ф. } \mathcal{N}(0,1) \text{ с.в. для аргумента } t_2+t_1\rho} = e^{-\frac{1}{2}t_1^2(1-\rho^2)} \cdot e^{-\frac{1}{2}(t_2+t_1\rho)^2}.
\end{aligned}$$

[0]Printed from the file [18-Gauss-density.tex] on 9.6.2016

Поэтому $h_{\mathbf{X}}(\mathbf{t}) = e^{-\frac{1}{2}(t_1^2+2\rho t_1 t_2+t_2^2)} = e^{-\frac{1}{2}\mathbf{t}'\mathbf{\Lambda}\mathbf{t}}$ для $\mathbf{\Lambda} = \left(\begin{smallmatrix} 1 & \rho \\ \rho & 1 \end{smallmatrix}\right)$. ② Так как $\mathbf{\Lambda}$ неотрицательно определенная матрица, ③ то $\mathbf{X} \in \mathcal{N}_2(\mathbf{0}, \mathbf{\Lambda})$ в смысле определения 17.2.

Таким образом, определение стандартного двумерного гауссовского вектора с помощью плотности (1) эквивалентно определениям 16.1 и 17.2 согласно теореме 17.2.

2. Плотность распределения гауссовского вектора

Рассмотрим теперь общий случай, понимая гауссовский вектор в смысле определений 16.1 и 17.2. Пусть $\boldsymbol{\mu}$ — неслучайный вектор, а $\mathbf{\Lambda}$ — неотрицательно определенная симметричная матрица. При вычислении плотности $\mathcal{N}(\boldsymbol{\mu}, \mathbf{\Lambda})$ вектора $\mathbf{X}$ необходимо различать два случая: $\det(\mathbf{\Lambda}) = 0$ и $\det(\mathbf{\Lambda}) > 0$.

Замечание 1. Третий вариант $\det(\mathbf{\Lambda}) < 0$ невозможен, так как $\mathbf{\Lambda} = \mathbf{CDC}'$ в силу утверждения 17.1, причем $\mathbf{C}$ — ортогональная матрица, а $\mathbf{D}$ — диагональная матрица с неотрицательными диагональными элементами. Так как по определению $\mathbf{CC}' = \mathbf{I}$, то $\det(\mathbf{C})\det(\mathbf{C}') = 1$. Поэтому $\det(\mathbf{\Lambda}) = \det(\mathbf{D}) \geq 0$. ④

2.1. Случай $\det(\mathbf{\Lambda}) = 0$**.** Для простоты ограничимся стандартными двумерными векторами. В этом случае плотность задается формулой (1). Согласно §4.1, лекция 15, координаты вектора являются стандартными гауссовскими случайными величинами, поэтому

$$\mathsf{E}[X_1] = \mathsf{E}[X_2] = 0 \qquad \text{и} \qquad \mathsf{E}\left[X_1^2\right] = \mathsf{E}\left[X_2^2\right] = 1.$$

Ковариация же между X_1 и X_2 вычислена в §4.3, лекция 15:

$$\mathsf{cov}[X_1, X_2] = \rho.$$

Таким образом ковариационной матрицей вектора $\mathbf{X}$ является $\mathsf{Cov}[\mathbf{X}] = \mathbf{\Lambda}$, где $\mathbf{\Lambda} = \left(\begin{smallmatrix} 1 & \rho \\ \rho & 1 \end{smallmatrix}\right)$. Так как $\det(\mathbf{\Lambda}) = 1 - \rho^2$, то $\det(\mathbf{\Lambda}) = 0$ тогда и только тогда, когда $\rho^2 = 1$. Ограничение $|\rho| < 1$, сделанное выше, исключает этот случай.

А что бы случилось, если бы $\rho^2 = 1$? В этом случае, как мы видели, $\mathsf{cov}[X_1, X_2] = \mathsf{E}[X_1 X_2] = \pm 1$. Если, например, $\rho = 1$, то $\mathsf{E}\left[(X_1 - X_2)^2\right] = 0$. Если же $\rho = -1$, то $\mathsf{E}\left[(X_1 + X_2)^2\right] = 0$. ⑤

Утверждение 1. *Если $|\rho| = 1$, то $\mathsf{P}(X_1 = \rho X_2) = 1$.*

Задача 1. *Доказать утверждение 1.* Указание. Для доказательства использовать неравенство Чебышева.

Утверждение 1 показывает, что $X_1 = \rho X_2$ с вероятностью единица. Иными словами, $\mathsf{P}(\mathbf{X} \in L) = 1$, где $L = \{(x, y) : x = \rho y\}$.

[1]Вспомнить формулу для математического ожидания функции от случайного вектора.
[2]Проверить, что $h_{\mathbf{X}}(\mathbf{t}) = e^{-\frac{1}{2}\mathbf{t}'\mathbf{\Lambda}\mathbf{t}}$ для $\mathbf{\Lambda} = \left(\begin{smallmatrix} 1 & \rho \\ \rho & 1 \end{smallmatrix}\right)$.
[3]Доказать, что $\mathbf{\Lambda} = \left(\begin{smallmatrix} 1 & \rho \\ \rho & 1 \end{smallmatrix}\right)$ является положительно определенной матрицей.
[4]Почему $\det(\mathbf{D}) \geq 0$?
[5]Почему $\mathsf{E}\left[(X_1 - \rho X_2)^2\right] = 0$, если $|\rho| = 1$?

Отсюда уже несложно получить, что вектор $\mathbf{X} \in \mathbf{R}^2$ плотности не имеет, если $\det(\mathbf{\Lambda}) = 0$. Действительно, если бы существовала плотность $f_{\mathbf{X}}$, то распределение вектора имело бы следующее свойство: для любых $x, y \in \mathbf{R}$

$$1 = \mathsf{P}(\mathbf{X} \in L) = \int_L f_{\mathbf{X}}(u, v)\, du\, dv.$$

Сдругой стороны, правая часть равна нулю, так как L имеет нулевую меру Лебега. Полученное противоречие доказывает, что вектор не имеет плотности, если $\det(\mathbf{\Lambda}) = 0$.

Аналогичная ситуация наблюдается и в общем случае.

Теорема 1. *Пусть $n > 1$ и $\mathbf{X} \in \mathcal{N}_n(\boldsymbol{\mu}, \mathbf{\Lambda})$. Если $\det(\mathbf{\Lambda}) = 0$, то координаты вектора $\mathbf{X}$ линейно зависимы с вероятностью единица. Иными словами, существует гиперплоскость L для которой $\mathsf{P}(\mathbf{X} \in L) = 1$. Кроме того, вектор плотности не имеет.*

Задача* 2. *Доказать теорему 1 в общем случае.*

Замечание 2. Пусть $\mathbf{X}$ — двумерный стандартный гауссовский вектор с параметром ρ. Если $\rho = 0$, то $f(x, y) = \varphi_{0,1}(x)\varphi_{0,1}(y)$. Согласно критерию независимости абсолютно непрерывных случайных величин координаты вектора являются независимыми случайными величинами. Поэтому, если вектор $\mathbf{X}$ гауссовский, то X_1 и X_2 некоррелированы тогда и только тогда, когда X_1 и X_2 независимы. Это свойство выполнено в очень редких случаях (гауссовский один из них).

2.2. Случай $\det(\mathbf{\Lambda}) > 0$**.** В этом случае мы докажем, что плотность гауссовского вектора существует и выражается через параметры вектора.

Теорема 2. *Пусть $n \geq 1$ и $\mathbf{X} \in \mathcal{N}_n(\boldsymbol{\mu}, \mathbf{\Lambda})$ в смысле определений* 16.1 *и* 17.2. *Если $\det(\mathbf{\Lambda}) > 0$, то вектор $\mathbf{X}$ имеет плотность $f_{\mathbf{X}}$, причем*

$$f_{\mathbf{X}}(\mathbf{x}) = \left(\frac{1}{2\pi}\right)^{n/2} \frac{1}{\sqrt{\det(\mathbf{\Lambda})}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\mathbf{\Lambda}^{-1}(\mathbf{x}-\boldsymbol{\mu})}. \tag{3}$$

Замечание 3. Если $n = 2$, то формула (3) совпадает с (1) в случае $\boldsymbol{\mu} = \mathbf{0}$ и $\mathbf{\Lambda} = \left(\begin{smallmatrix} 1 & \rho \\ \rho & 1 \end{smallmatrix}\right)$, то есть для двумерного стандартного гауссовского вектора с параметром ρ.

Задача 3. *Доказать, что* (3) *совпадает с* (1) *в случае $\boldsymbol{\mu} = \mathbf{0}$, $\mathbf{\Lambda} = \left(\begin{smallmatrix} 1 & \rho \\ \rho & 1 \end{smallmatrix}\right)$.* Указание. Для нахождения обратной матрицы выразить $\mathbf{\Lambda}^{-1}$ через алгебраические дополнения $\mathbf{\Lambda}$ и $\det(\mathbf{\Lambda})$.

Сначала докажем следующий вспомогательный результат.

Лемма 1. *Если $\mathbf{\Lambda}$ — ковариаационная матрица некоторого вектора $\mathbf{X}$, причем $\det(\mathbf{\Lambda}) > 0$, то матрицы $\mathbf{\Lambda}^{1/2}$, $\left(\mathbf{\Lambda}^{1/2}\right)^{-1}$, $\mathbf{\Lambda}^{-1}$, $\left(\mathbf{\Lambda}^{-1}\right)^{1/2}$ существуют и являются симметричными.*

Доказательство леммы 1. Согласно утверждению 17.1 справедливо представление $\mathbf{\Lambda} = \mathbf{CDC}'$, где $\mathbf{C}$ — ортогональная матрица, а $\mathbf{D}$ — диагональная. Докажем, что $\mathbf{\Lambda}^{1/2} = \mathbf{CD}^{1/2}\mathbf{C}'$:

$$\begin{aligned} \mathbf{\Lambda}^{1/2}\mathbf{\Lambda}^{1/2} &= \mathbf{CD}^{1/2}\mathbf{C}'\mathbf{CD}^{1/2}\mathbf{C}' = \mathbf{CD}^{1/2}\mathbf{ED}^{1/2}\mathbf{C}' \\ &= \mathbf{CD}^{1/2}\mathbf{D}^{1/2}\mathbf{C}' = \mathbf{CDC}' = \mathbf{\Lambda}. \end{aligned}$$

Задача 4. *Доказать, что* $\mathbf{AEB} = \mathbf{AB}$ *для любых матриц* $\mathbf{A}$ *и* $\mathbf{B}$.

Так как $\mathbf{D}^{1/2}$ — диагональная матрица, то

$$\left(\mathbf{\Lambda}^{1/2}\right)' = \left(\mathbf{C}\mathbf{D}^{1/2}\mathbf{C}'\right)' = \mathbf{C}\left(\mathbf{D}^{1/2}\right)'\mathbf{C}' = \mathbf{C}\mathbf{D}^{1/2}\mathbf{C}' = \mathbf{\Lambda}^{1/2}.$$

⑥ Это означает, что $\mathbf{\Lambda}^{1/2}$ является симметричной матрицей.

Докажем теперь симметричность матрицы $\left(\mathbf{\Lambda}^{1/2}\right)^{-1}$. Пользуясь ортогональностью матрицы $\mathbf{C}$, убеждаемся, что $\left(\mathbf{\Lambda}^{1/2}\right)^{-1} = \mathbf{C}\left(\mathbf{D}^{1/2}\right)^{-1}\mathbf{C}'$. ⑦ Поскольку $\left(\mathbf{D}^{1/2}\right)^{-1}$ — диагональная матрица, то

$$\left(\left(\mathbf{\Lambda}^{1/2}\right)^{-1}\right)' = \mathbf{C}\left(\mathbf{D}^{1/2}\right)^{-1}\mathbf{C}' = \left(\mathbf{\Lambda}^{1/2}\right)^{-1},$$

значит матрица $\left(\mathbf{\Lambda}^{1/2}\right)^{-1}$ симметрична.

Утверждение леммы для других матриц доказывается аналогично. □

Задача 5. *Доказать утверждение леммы* 1 *для* $\mathbf{\Lambda}^{-1}$ *и* $\left(\mathbf{\Lambda}^{-1}\right)^{1/2}$.

Доказательство теоремы 2. Рассмотрим вектор $\mathbf{Y}$, координатами которого служат независимые стандартные гауссовские случайные величины. Положим $\mathbf{X} = \mathbf{\Lambda}^{1/2}\mathbf{Y} + \boldsymbol{\mu}$. Тогда $\mathbf{X} \in \mathcal{N}(\boldsymbol{\mu}, \mathbf{\Lambda})$, как в доказательстве леммы 17.4. ⑧ Поскольку координаты $\mathbf{Y}$ независимы, то его плотность равна произведению плотностей координат:

$$f_{\mathbf{Y}}(\mathbf{t}) = \prod_{k=1}^{n} f_{Y_k}(t_k) = \left(\frac{1}{2\pi}\right)^{n/2} \prod_{k=1}^{n} e^{-\frac{1}{2}t_k^2} = \left(\frac{1}{2\pi}\right)^{n/2} e^{-\frac{1}{2}\mathbf{t}'\mathbf{t}}, \qquad \mathbf{t} \in \mathbf{R}^n. \tag{4}$$

Пусть теперь $x_1, \dots, x_n \in \mathbf{R}$. Обозначим через $\Pi = (-\infty, x_1) \times \cdots \times (-\infty, x_n)$ "прямоугольник" в $\mathbf{R}^n$ с "верхней правой" вершиной $\mathbf{x}$. Рассмотрим отображение

$$\mathbf{\Lambda}^{-1/2}(\bullet - \boldsymbol{\mu}) : \mathbf{R}^n \to \mathbf{R}^n,$$

где символом $\bullet$ обозначено аргумент этого отображения. Образ прямоугольника Π при этом отображении обозначим через

$$\mathbf{\Sigma}_1 = \left\{\mathbf{u}\colon \ \text{существует } \mathbf{w} \in \Pi, \text{ для которого } \mathbf{u} = \Lambda^{-1/2}(\mathbf{w} - \boldsymbol{\mu})\right\}.$$

Тут $\mathbf{\Lambda}^{-1/2} = \left(\Lambda^{1/2}\right)^{-1}$ — обратная матрица к $\mathbf{\Lambda}^{1/2}$. Так как $\mathbf{Y} = \mathbf{\Lambda}^{-1/2}(\mathbf{X} - \boldsymbol{\mu})$, то $\mathbf{X} \in \mathbf{\Pi} \iff \mathbf{Y} \in \mathbf{\Sigma}_1$, ⑨ то

$$\mathsf{P}(\mathbf{X} \in \mathbf{\Pi}) = \mathsf{P}(\mathbf{\Lambda}^{1/2}\mathbf{Y} + \boldsymbol{\mu} \in \mathbf{\Pi}) = \mathsf{P}(\mathbf{Y} \in \mathbf{\Sigma}_1) = \int_{\mathbf{\Sigma}_1} f_{\mathbf{Y}}(\mathbf{u})\, d\mathbf{u}. \tag{5}$$

---

[6] Почему $\left(\mathbf{D}^{1/2}\right)' = \mathbf{D}^{1/2}$?

[7] Убедиться, что $\left(\mathbf{\Lambda}^{1/2}\right)^{-1} = \mathbf{C}\left(\mathbf{D}^{1/2}\right)^{-1}\mathbf{C}'$.

[8] Доказать, что $\mathbf{X} \in \mathcal{N}(\boldsymbol{\mu}, \mathbf{\Lambda})$, если $\mathbf{X} = \mathbf{\Lambda}^{1/2}\mathbf{Y} + \boldsymbol{\mu}$ и вектор $\mathbf{Y}$ составлен из независимых стандартных гауссовских случайных величин.

[9] Почему $\mathbf{\Lambda}^{1/2}\mathbf{Y} + \boldsymbol{\mu} \in \mathbf{\Pi} \iff \mathbf{Y} \in \mathbf{\Sigma}_1$?

Сделав в интеграле замену $\mathbf{v} = \mathbf{\Lambda}^{1/2}\mathbf{u} + \boldsymbol{\mu}$ или $\mathbf{u} = \mathbf{\Lambda}^{-1/2}(\mathbf{v} - \boldsymbol{\mu})$, получаем

$$(6) \qquad \mathsf{P}(\mathbf{X} \in \mathbf{\Pi}) = \int_{\mathbf{\Sigma}_1} f_{\mathbf{Y}}(\mathbf{u})\,d\mathbf{u} = \int_{\mathbf{\Sigma}_2} f_{\mathbf{Y}}(\mathbf{\Lambda}^{-1/2}(\mathbf{v} - \boldsymbol{\mu}))|\mathbf{J}|\,d\mathbf{v},$$

где $\mathbf{\Sigma}_2$ — это образ области $\mathbf{\Sigma}_1$ при отображении $\mathbf{u} \to \mathbf{v} = \mathbf{\Lambda}^{1/2}\mathbf{u} + \boldsymbol{\mu}$, а $\mathbf{J}$ — якобиан этого отображения. Формула замены переменных, которую мы применили, справедлива только в случае взаимно однозначных отображений, для которых элементы якобиана $\mathbf{J}$ непрерывны и $|\mathbf{J}| \neq 0$ ни для одного аргумента. Проверим все эти условия.

Обратное отображение задается формулой $\mathbf{v} \to \mathbf{u} = \mathbf{\Lambda}^{-1/2}(\mathbf{v} - \boldsymbol{\mu})$. Если обозначить через λ_{ij} элементы матрицы $\mathbf{\Lambda}^{-1/2}$ и $\mathbf{u}' = (u_1, \ldots, u_n)$, $\mathbf{v}' = (v_1, \ldots, v_n)$, то $u_i = \sum_j \lambda_{ij}(v_j - \mu_j)$, ⑩ откуда $\frac{\partial u_i}{\partial v_j} = \lambda_{ij}$, то есть $\mathbf{J} = \mathbf{\Lambda}^{-1/2}$ и поэтому

$$|\mathbf{J}| = \det(\mathbf{\Lambda}^{-1/2}) = \frac{1}{\sqrt{\det(\mathbf{\Lambda})}} \neq 0.$$

Таким образом, элементы якобиана являются константами (и поэтому непрерывны) и $|\mathbf{J}| \neq 0$. Последнее условие также гарантирует, что каждая из систем линейных уравнений

$$\mathbf{v} = \mathbf{\Lambda}^{1/2}\mathbf{u} + \boldsymbol{\mu} \qquad (\text{относительно неизвестного } \mathbf{u}),$$
$$\mathbf{u} = \mathbf{\Lambda}^{-1/2}(\mathbf{v} - \boldsymbol{\mu}) \qquad (\text{относительно неизвестного } \mathbf{v})$$

имеет единственное решение, что и доказывает взаимную однозначность отображения. ⑪

В формуле (6) осталось определить $\mathbf{\Sigma}_2$. Множество $\mathbf{\Sigma}_1$ состоит из таких векторов $\mathbf{u}$, для которых существует $\mathbf{w} \in \mathbf{\Pi}$, при котором $\mathbf{u} = \mathbf{\Lambda}^{-1/2}(\mathbf{w} - \boldsymbol{\mu})$. Замена переменных в интеграле (6) переводит $\mathbf{u}$ в

$$\mathbf{\Lambda}^{1/2}\mathbf{u} + \boldsymbol{\mu} = \mathbf{\Lambda}^{1/2}(\mathbf{\Lambda}^{-1/2}(\mathbf{w} - \boldsymbol{\mu})) + \boldsymbol{\mu} = \mathbf{w},$$

то есть в элемент $\mathbf{\Pi}$. Выше мы уже установили, что это преобразование взаимно однозначное. Это доказывает, что $\mathbf{\Sigma}_2 = \Pi$. ⑫ Окончательно получаем из (4) и (5)

$$\begin{aligned}
\mathsf{P}(\mathbf{X} \in \mathbf{\Pi}) &= \left(\frac{1}{2\pi}\right)^{n/2} \frac{1}{\sqrt{\det(\mathbf{\Lambda})}} \int_{\mathbf{\Pi}} e^{-\frac{1}{2}\left(\mathbf{\Lambda}^{-1/2}(\mathbf{v}-\boldsymbol{\mu})\right)'(\mathbf{\Lambda}^{-1/2}(\mathbf{v}-\boldsymbol{\mu}))}\,d\mathbf{v} \\
&= \left(\frac{1}{2\pi}\right)^{n/2} \frac{1}{\sqrt{\det(\mathbf{\Lambda})}} \int_{\mathbf{\Pi}} e^{-\frac{1}{2}(\mathbf{v}-\boldsymbol{\mu})'(\mathbf{\Lambda}^{-1/2})'\mathbf{\Lambda}^{-1/2}(\mathbf{v}-\boldsymbol{\mu})}\,d\mathbf{v}.
\end{aligned}$$

В силу леммы 1 матрица $\mathbf{\Lambda}^{-1/2}$ симметрична и поэтому $\mathsf{P}(\mathbf{X} \in \Pi) = \int_\Pi f_{\mathbf{X}}(\mathbf{v})\,d\mathbf{v}$ для

$$f_{\mathbf{X}}(\mathbf{v}) = \left(\frac{1}{2\pi}\right)^{n/2} \frac{1}{\sqrt{\det(\mathbf{\Lambda})}} e^{-\frac{1}{2}(\mathbf{v}-\boldsymbol{\mu})'(\mathbf{\Lambda}^{-1/2})'\mathbf{\Lambda}^{-1/2}(\mathbf{v}-\boldsymbol{\mu})}.$$

то есть теорема 2 доказана. ⑬ □

---

[10] Почему $u_i = \sum_j \lambda_{ij}(v_j - \mu_j)$?

[11] Почему каждая из систем $\mathbf{v} = \mathbf{\Lambda}^{1/2}\mathbf{u} + \boldsymbol{\mu}$ и $\mathbf{u} = \mathbf{\Lambda}^{-1/2}(\mathbf{v} - \boldsymbol{\mu})$ имеет единственное решение?

[12] Почему $\mathbf{\Sigma}_2 = \Pi$?

[13] Как закончить доказательство теоремы 2?

3. Еще одно эквивалентное определение гауссовского вектора

Еще одно определение гауссовского вектора можно предложить с помощью его плотности.

Определение 1. Случайный вектор называется гауссовским, если его плотность $f_{\mathbf{X}}$ задана равенством (3), в котором $\mathbf{\Lambda}$ — неотрицательно определенная симметричная матрица, $\det(\mathbf{\Lambda}) > 0$, а $\boldsymbol{\mu}$ — неслучайный вектор.

Теорема 3. *В несингулярном случае* (*если* $\det(\mathbf{\Lambda}) > 0$) *определение* 1 *эквивалентно определениям* 16.1 *и* 17.2.

Доказательство. В силу теоремы 2 достаточно доказать, что из определения 1 вытекает, что характеристическая функция $h_{\mathbf{X}}(\mathbf{t})$ вектора $\mathbf{X}$ равна $e^{i\mathbf{t}'\boldsymbol{\mu}-\frac{1}{2}\mathbf{t}'\mathbf{\Lambda}\mathbf{t}}$. ⑭ Из доказательства теоремы 2 мы знаем, что $\mathbf{X} = \mathbf{\Lambda}^{1/2}\mathbf{Y}+\boldsymbol{\mu}$, где вектор $\mathbf{Y}$ составлен из независимых стандартных случайных величин. В силу независимости координат вектора $\mathbf{Y}$ характеристическая функция вектора $\mathbf{Y}$ равна произведению маргинальных характеристических функций: $h_{\mathbf{Y}}(\mathbf{t}) = e^{-\frac{1}{2}(t_1^2+\cdots+t_n^2)} = e^{-\frac{1}{2}\mathbf{t}'\mathbf{t}}$. Поэтому для $\mathbf{s} = \mathbf{\Lambda}^{1/2}\mathbf{t}$ или $\mathbf{s}' = \mathbf{t}'\mathbf{\Lambda}^{1/2}$ (см. лемму 1) ⑮ имеем

$$\begin{aligned} h_{\mathbf{X}}(\mathbf{t}) &= \mathsf{E}\left[e^{i\mathbf{t}'\mathbf{X}}\right] = \mathsf{E}\left[e^{i\mathbf{t}'(\mathbf{\Lambda}^{1/2}\mathbf{Y}+\boldsymbol{\mu})}\right] = e^{i\mathbf{t}'\boldsymbol{\mu}}\mathsf{E}\left[e^{i\mathbf{s}'\mathbf{Y}}\right] = e^{i\mathbf{t}'\boldsymbol{\mu}}h_{\mathbf{Y}}(\mathbf{s}) \\ &= e^{i\mathbf{t}'\boldsymbol{\mu}}e^{-\frac{1}{2}\mathbf{s}'\mathbf{s}} = e^{i\mathbf{t}'\boldsymbol{\mu}-\frac{1}{2}\mathbf{t}'\mathbf{\Lambda}^{1/2}\mathbf{\Lambda}^{1/2}\mathbf{t}} = e^{i\mathbf{t}'\boldsymbol{\mu}-\frac{1}{2}\mathbf{t}'\mathbf{\Lambda}\mathbf{t}}. \end{aligned}$$

Это и заканчивает доказательство теоремы 3. □

[14] Почему для доказательства теоремы 3 достаточно показать, что характеристическая функция вектора $\mathbf{X}$ равна $e^{i\mathbf{t}'\boldsymbol{\mu}-\frac{1}{2}\mathbf{t}'\mathbf{\Lambda}\mathbf{t}}$?

[15] Показать, что из $\mathbf{s} = \mathbf{\Lambda}^{1/2}\mathbf{t}$ вытекает $\mathbf{s}' = \mathbf{t}'\mathbf{\Lambda}^{1/2}$.

[†] Всего в тексте было 15 вопросов

Лекция 19

НЕЗАВИСИМЫЕ ГАУССОВСКИЕ ВЕКТОРЫ

1. Независимость случайных векторов: общий случай

Определение 1. Случайные m-вектор $\mathbf{X}$ и n-вектор $\mathbf{Y}$ называются *независимыми*, если для любых множеств $A \in \mathscr{B}_m$ и $B \in \mathscr{B}_n$

$$\mathsf{P}(\mathbf{X} \in A, \mathbf{Y} \in B) = \mathsf{P}(\mathbf{X} \in A)\mathsf{P}(\mathbf{Y} \in B). \tag{1}$$

Тут $\mathscr{B}_m$ — это σ-алгебра борелевских множеств в $\mathbf{R}^m$, а $\mathscr{B}_n$ — в $\mathbf{R}^n$.

Чтобы доказать независимость векторов, необходимо проверить формулу (1) для всех возможных борелевских множеств A и B, в частности для

$$A = A(\mathbf{x}) \stackrel{\text{def}}{=} (-\infty, x_1) \times \dots \times (-\infty, x_m), \tag{2}$$

$$A = A(\boldsymbol{\alpha}, \boldsymbol{\beta}) \stackrel{\text{def}}{=} [\alpha_1, \beta_1) \times \dots \times [\alpha_m, \beta_m) \tag{3}$$

при любых $\mathbf{x} \in \mathbf{R}^m$ и $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbf{R}^m$, $\alpha_1 < \beta_1, \dots, \alpha_m < \beta_m$, а также для

$$B = B(\mathbf{y}) \stackrel{\text{def}}{=} (-\infty, y_1) \times \dots \times (-\infty, y_n), \tag{4}$$

$$B = B(\boldsymbol{\gamma}, \boldsymbol{\delta}) \stackrel{\text{def}}{=} [\gamma_1, \delta_1) \times \dots [\gamma_n, \delta_n)$$

при любых $\mathbf{y} \in \mathbf{R}^n$ и $\boldsymbol{\gamma}, \boldsymbol{\delta} \in \mathbf{R}^n$, $\gamma_1 < \delta_1, \dots, \gamma_n < \delta_n$.

Замечание 1. Семейство множеств (2), как и (3), не исчерпывает всю σ-алгебру борелевских множеств $\mathscr{B}_m$. ①

1.1. Эквивалентные определения независимости. Классы множеств $\mathscr{B}_m$ и $\mathscr{B}_n$, для которых надо проверять равенство (1) при доказательстве независимости случайных векторов, можно существенно сузить.

Теорема 1. *Векторы $\mathbf{X}$ и $\mathbf{Y}$ являются независимыми тогда и только тогда, когда выполнено одно из следующих условий:*

(5) *равенство* (1) *выполняется для любых $A = A(\mathbf{x})$, $\mathbf{x} \in \mathbf{R}^m$, и $B = B(\mathbf{y})$, $\mathbf{y} \in \mathbf{R}^n$;*

(6) *равенство* (1) *выполняется для любых $A = A(\boldsymbol{\alpha}, \boldsymbol{\beta})$, $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbf{R}^m$, и $B = B(\boldsymbol{\gamma}, \boldsymbol{\delta})$, $\boldsymbol{\gamma}, \boldsymbol{\delta} \in \mathbf{R}^n$;*

(7) *равенство* (1) *выполняется для любых $A = A(\mathbf{x})$, $\mathbf{x} \in \mathbf{R}^m$, и $B = B(-\boldsymbol{\gamma}, \boldsymbol{\gamma})$, $\boldsymbol{\gamma} \in \mathbf{R}^n$.*

Тут мы обозначаем через $-\boldsymbol{\gamma}$ вектор $(-\gamma_1, \dots, -\gamma_n)'$ для $\boldsymbol{\gamma}' = (\gamma_1, \dots, \gamma_n)$.

Задача* 1. *Доказать, что определение 1 эквивалентно каждому из условий, сформулированных в теореме 1.* **Указание**. Использовать единственность продолжения меры.

[0] Printed from the file `[19a-independence.tex]` on 9.6.2016

[1] Привести пример борелевского множества, которое нельзя представить ни в виде (2), ни в виде (3).

2. Независимость случайных векторов: гауссовский случай

2.1. Критерий независимости в терминах плотностей. Для абсолютно непрерывных случайных векторов критерий независимости можно записать с помощью их плотностей. Этот критерий в гауссовском случае имеет следующий вид.

Теорема 2. *Пусть* $\mathbf{X}$ *и* $\mathbf{Y}$ *— два случайных вектора. Обозначим*

$$\mathbf{Z} \stackrel{\text{def}}{=} (\mathbf{X}', \mathbf{Y}')' = (X_1, \ldots, X_m, Y_1, \ldots, Y_n)'$$

и предположим, что $\mathbf{Z}$ *является несингулярным гауссовским* $(m+n)$*-вектором. Случайные векторы* $\mathbf{X}$ *и* $\mathbf{Y}$ *независимы тогда и только тогда, когда*

$$f_{\mathbf{Z}}(\mathbf{x}, \mathbf{y}) = f_{\mathbf{X}}(\mathbf{x}) f_{\mathbf{Y}}(\mathbf{y}) \qquad \textit{для любых } \mathbf{x} \in \mathbf{R}^m \textit{ и } \mathbf{y} \in \mathbf{R}^n. \qquad ②$$

Задача* **2.** *Доказать теорему* 2. Указание. Применить метод, который использовался для доказательства аналогичного критерия для абсолютно непрерывных случайных величин.

2.2. Пример, когда $X - Y$ и $X + Y$ независимы. Пусть X и Y — независимые случайные величины. В общем случае трудно ожидать, что $X - Y$ и $X + Y$ являются независимыми. Но если X и Y — гауссовские случайные величины, то это свойство выполнено.

Утверждение 1. *Пусть* X *и* Y *независимые стандартные гауссовские случайные величины. Тогда* $X - Y$ *и* $X + Y$ *независимы.*

Доказательство. В силу следствия 17.1 вектор $(X, Y)'$ является гауссовским. Поэтому линейная комбинация $c_1 X + c_2 Y$ является гауссовской случайной величиной для любых $c_1, c_2 \in \mathbf{R}$, если $|c_1| + |c_2| > 0$. Следовательно,

$$c_1(X - Y) + c_2(X + Y) = (c_1 + c_2)X + (c_1 - c_2)Y$$

является гауссовской случайной величиной для любых $c_1, c_2 \in \mathbf{R}$, если $|c_1| + |c_2| > 0$. Значит вектор $(X - Y, X + Y)'$ также является гауссовским (согласно определению 16.2). ③ Поскольку $\mathsf{E}[X - Y] = \mathsf{E}[X + Y] = 0$, то

$$\mathsf{cov}[X - Y, X + Y] = \mathsf{E}\left[X^2 - Y^2\right] = 0,$$

то есть случайные величины $X - Y$ и $X + Y$ некоррелированы. Ясно, что $\mathsf{E}\left[(X - Y)^2\right] = \mathsf{E}\left[(X + Y)^2\right] = 2$ и $\det(\mathbf{\Lambda}) = 4$, где $\mathbf{\Lambda}$ – это ковариационная матрица вектора $(X - Y, X + Y)'$:

$$\mathbf{\Lambda} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \qquad \mathbf{\Lambda}^{-1} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}, \qquad \det(\mathbf{\Lambda}^{-1}) = \frac{1}{4}.$$

Значит для $\mathbf{u} = (u_1, u_2)'$

$$\mathbf{u}' \mathbf{\Lambda}^{-1} \mathbf{u} = \frac{1}{2}(u_1^2 + u_2^2) \qquad ④$$

---

[2] Почему существуют плотности $f_{\mathbf{X}}$ и $f_{\mathbf{Y}}$?

[3] Такой вывод можно сделать только при условии, что из $|c_1| + |c_2| > 0$ вытекает, что $|c_1 + c_2| + |c_1 - c_2| > 0$. Почему это действительно так?

и поэтому

$$\begin{aligned} f_{X-Y,X+Y}(u_1,u_2) &= \frac{1}{2\pi}\cdot\frac{1}{\sqrt{\det(\mathbf{\Lambda})}}\cdot e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\mathbf{\Lambda}^{-1}(\mathbf{x}-\boldsymbol{\mu})} \\ &= \frac{1}{2\pi}\cdot\frac{1}{2}\cdot e^{-\frac{1}{2}\mathbf{u}'\mathbf{\Lambda}^{-1}\mathbf{u}} = \frac{1}{\sqrt{2\pi\cdot 2}}e^{-\frac{1}{2\cdot 2}u_1^2}\cdot\frac{1}{\sqrt{2\pi\cdot 2}}e^{-\frac{1}{2\cdot 2}u_2^2} \\ &= \varphi_{0,2}(u_1)\varphi_{0,2}(u_2) = f_{X-Y}(u_1)f_{X+Y}(u_2) \end{aligned}$$

(см. теорему 18.2). ⑤ Таким образом, из теоремы 2 вытекает, что $X-Y$ и $X+Y$ независимы. ☐

Задача 3. *Доказать утверждение* 1 *для общего случая независимых случайных величин* $X\in\mathcal{N}(a_1,\sigma_1^2)$ *и* $Y\in\mathcal{N}(a_2,\sigma_2^2)$.

2.3. Пример гауссовских некоррелированных, но зависимых случайных величин. При доказательстве независимости гауссовских случайных векторов с помощью плотностей необходимо обязательно проверять, что их совместное распределение является несингулярным. Если же распределение сингулярно, то из некоррелированности не обязательно вытекает независимость.

Пример 1. Рассмотрим случайные величины $X_1\in\mathcal{N}(0,1)$ и $X_2\in\mathcal{N}(0,1)$, построенные в лекции 16, §1.2, стр. 124, то есть $X_2=\beta X_1$, а $\beta\in\mathrm{Rad}\left(\frac{1}{2}\right)$ не зависит от X_1. Тогда $\mathsf{E}[X_1]=\mathsf{E}[X_2]=0$ и

$$\mathsf{cov}[X_1,X_2]=\mathsf{E}[X_1X_2]=\mathsf{E}\left[X_1^2\beta\right]=\mathsf{E}\left[X_1^2\right]\mathsf{E}[\beta]=0,$$

то есть X_1 и X_2 некоррелированы.

С другой стороны, $X_1^2=X_2^2$ и поэтому X_1 и X_2 зависимы. Если бы это было не так, то X_1^2 и X_2^2 также были бы независимы. Тогда бы $\mathsf{E}\left[X_1^2X_2^2\right]=\mathsf{E}\left[X_1^2\right]\mathsf{E}\left[X_2^2\right]=1$. С другой стороны,

$$\begin{aligned} \mathsf{E}\left[X_1^2X_2^2\right]=\mathsf{E}\left[X_1^4\right] &= \int_{-\infty}^{\infty}x^4\varphi(x)\,dx \\ &= \left[\begin{smallmatrix}\text{По частям:}\\ u=x^3,\ du=3x^2\\ dv=x\varphi(x)\,dx,\ v=-\varphi(x)\end{smallmatrix}\right] = 3\int_{-\infty}^{\infty}x^2\varphi(x)\,dx=3. \end{aligned}$$

⑥ Значит случайные величины X_1 и X_2 зависимы.

3. Критерий независимости координат гауссовского вектора

Теорему 2 можно расширить следующим образом.

Теорема 3. *Пусть* $X_1,\ldots,X_n$ *— случайные величины, для которых* $\mathbf{X}=(X_,\ldots,X_n)'$ *— несингулярный гауссовский вектор. В этом случае случайные величины* $X_1,\ldots,X_n$ *являются независимыми в совокупности тогда и только тогда, когда*

$$f_{\mathbf{X}}(\mathbf{x})=f_{X_1}(x_1)\ldots f_{X_n}(x_n)$$

для любых $x_1,\ldots,x_n\in\mathbf{R}$, *где* $\mathbf{x}=(x_1,\ldots,x_n)$.

---

[4] Доказать, что $\mathbf{u}'\mathbf{\Lambda}^{-1}\mathbf{u}=\frac{1}{2}(u_1^2+u_2^2)$.

[5] Почему $\varphi_{0,2}$ является плотностью как случайной величины $X-Y$, так и $X+Y$?

[6] Почему $uv\Big|_{-\infty}^{\infty}=0$?

Задача 4. *Доказать теорему 3.* Указание. Использовать теорему 2 и индукцию.

Из этого утверждения вытекает, что некоррелированность в гауссовском случае эквивалентна независимости.

Теорема 4. *Пусть* $\mathbf{X}$ — *несингулярный гауссовский вектор,* $\mathbf{X} \in \mathcal{N}_n(\boldsymbol{\mu}, \boldsymbol{\Lambda})$, $\det(\boldsymbol{\Lambda}) > 0$. *Если его координаты попарно некоррелированы, то они независимы в совокупности.*

Следствие 1. *Пусть* X_1 *и* X_2 *некоррелированные гауссовские случайные величины. Тогда* X_1 *и* X_2 *независимые случайные величины.*

Задача 5. *Доказать следствие 1.* Указание. Обратите внимание, что несингулярность вектора $(X_1, X_2)'$ не упоминается. Почему?

Доказательство теоремы 4. Если координаты вектора попарно некоррелированы, то его ковариационная матрица $\boldsymbol{\Lambda}$ диагональна, причем на ее диагонале стоят дисперсии $\sigma_1^2, \ldots, \sigma_n^2$. Значит диагональной будет и обратная матрица $\boldsymbol{\Lambda}^{-1}$, причем на ее диагонале стоят числа $1/\sigma_1^2, \ldots, 1/\sigma_n^2$. Поскольку в этом случае для любого $\mathbf{x} \in \mathbf{R}^n$

$$(\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Lambda}^{-1} (\mathbf{x} - \boldsymbol{\mu}) = \sum_{k=1}^{n} \frac{(x_k - \mu_k)^2}{\sigma_k^2}, \qquad \det(\boldsymbol{\Lambda}) = \sigma_1^2 \ldots \sigma_n^2, \qquad ⑦$$

то плотность вероятности вектора $\mathbf{X}$, полученная в теореме 18.2, распадается в произведение маргинальных плотностей. Согласно теореме 2 это и означает, что случайные величины независимы в совокупности. $\square$

4. Критерий независимости гауссовских случайных векторов

Теорему 4 можно обобщить на случай независимых гауссовский случайных векторов.

Теорема 5. *Пусть* $\mathbf{X} = (X_1, \ldots, X_m)'$ *и* $\mathbf{Y} = (Y_1, \ldots, Y_n)'$ *случайные векторы размерности* m *и* n *соответственно. Предположим, что* $\mathbf{Z} \stackrel{\text{def}}{=} (\mathbf{X}', \mathbf{Y}')'$ — *несингулярный гауссовский* $(m+n)$*-вектор. Случайные векторы* $\mathbf{X}$ *и* $\mathbf{Y}$ *независимы тогда и только тогда, когда* $\mathsf{cov}[X_i, Y_j] = 0$ *для любых* $1 \le i \le m$ *и* $1 \le j \le n$.

Доказательство. Мы рассмотрим только случай $m = 1$, то есть когда $\mathbf{X} = X$ — случайная величина. В этом случае ковариационная матрица вектора $\mathbf{Z}$ равна $\mathbf{C} \stackrel{\text{def}}{=} \begin{pmatrix} \sigma^2 & 0 \ldots 0 \\ 0 & \\ \ldots & \boldsymbol{\Lambda} \\ 0 & \end{pmatrix}$, где $\sigma^2 = \mathsf{var}[X]$, а $\boldsymbol{\Lambda}$ — это ковариационная матрица вектора $\mathbf{Y}$. Несложно убедиться, что $\det(\mathbf{C}) = \sigma^2 \det(\boldsymbol{\Lambda})$ ⑧ и

$$\mathbf{C}^{-1} = \begin{pmatrix} \sigma^{-2} & 0 \quad \ldots \quad 0 \\ 0 & \\ \ldots & \boldsymbol{\Lambda}^{-1} \\ 0 & \end{pmatrix}. \qquad ⑨$$

[7] Вычислить $(\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Lambda}^{-1} (\mathbf{x} - \boldsymbol{\mu})$ и $\det(\boldsymbol{\Lambda})$.

[8] Убедиться, что $\det(\mathbf{C}) = \sigma^2 \det(\boldsymbol{\Lambda})$

Если положить $\mathbf{z}' = (x, \mathbf{y}')$, $x \in \mathbf{R}$, $\mathbf{y} \in \mathbf{R}^n$, то легко видеть, что

$$\mathbf{z}'\mathbf{C}^{-1}\mathbf{z} = \frac{x^2}{\sigma^2} + \mathbf{y}'\mathbf{\Lambda}^{-1}\mathbf{y}. \qquad ⑩$$

Поэтому из теоремы 18.2 вытекает, что плотность вектора $\mathbf{Z}$ имеет вид

$$f_{\mathbf{Z}}(\mathbf{z}) = \frac{1}{(2\pi)^{(n+1)/2}} \frac{1}{\sqrt{\det(\mathbf{C})}} e^{-\frac{1}{2}\mathbf{z}'\mathbf{C}^{-1}\mathbf{z}} = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \cdot \frac{1}{\sqrt{\det(\mathbf{\Lambda})}} e^{-\frac{1}{2}\mathbf{y}'\mathbf{\Lambda}^{-1}\mathbf{y}}.$$

На основании теоремы 2 это и означает, что X и $\mathbf{Y}$ независимы. $\square$

Задача 6. *Доказать теорему 5 для $m > 1$.*

5. Независимость выборочного среднего и выборочной дисперсии

Пусть случайные величины $X_1, \ldots, X_n$ независимы и одинаково распределены. Положим

$$\overline{X}_n = \frac{1}{n}\sum_{k=1}^n X_k \qquad \text{и} \qquad \overline{\sigma}_n^2 = \frac{1}{n}\sum_{k=1}^n \left(X_k - \overline{X}_n\right)^2.$$

Эти величины называются *выборочным средним* и *выборочной дисперсией* соответственно. Они применяются в статистике для оценки неизвестных математического ожидания и дисперсии.

Случай $X_k \in \mathcal{N}\left(a, \sigma^2\right)$, $1 \le k \le n$, считается важнейшим в статистике. В этом случае многие гипотезы о параметрах проверяются с помощью *t-критерия Стьюдента*, который основан на следующем результате.

Теорема 6. *Если $X_1, \ldots, X_n$ независимые гауссовские случайные величины, то выборочные среднее и дисперсия, $\overline{X}_n$ и $\overline{\sigma}_n^2$, независимы.*

Доказательство. Мы рассмотрим только случай $a = 0$, $\sigma^2 = 1$, а доказательство разобьем на несколько частей.

Лемма 1. *Если $X_1, \ldots, X_n$ независимы и $X_k \in \mathcal{N}(0, 1)$, то $(n+1)$-вектор $\mathbf{Y} = (\overline{X}_n, X_1 - \overline{X}_n, \ldots, X_n - \overline{X}_n)'$ является гауссовским.*

Доказательство леммы 1. Прежде всего, отметим, что вектор $(X_1, \ldots, X_n)'$ является гауссовским (следстие 17.1). Кроме того, для любого $\mathbf{c} \in \mathbf{R}^{n+1}$

$$\mathbf{c}'\mathbf{Y} = c_1\overline{X}_n + c_2(X_1 - \overline{X}_n) + \cdots + c_{n+1}(X_n - \overline{X}_n) = d_1X_1 + \cdots + d_nX_n,$$

где $d_k = c_{k+1} + \frac{1}{n}(c_1 - c_2 - \cdots - c_{n+1})$, $1 \le k \le n$. Поэтому $\mathbf{c}'\mathbf{Y}$ — гауссовская случайная величина. Следовательно $\mathbf{Y}$ — гауссовский вектор (определение 16.2). $\square$

---

[9]Убедиться, что $\mathbf{C}^{-1}$ записывается именно в таком виде.

[10]Доказать, что $\mathbf{z}'\mathbf{C}^{-1}\mathbf{z} = \frac{x^2}{\sigma^2} + \mathbf{y}'\mathbf{\Lambda}^{-1}\mathbf{y}$.

Лемма 2. *Пусть $X_k \in \mathcal{N}(0,1)$, $1 \le k \le n$, — независимые в совокупности случайные величины. Тогда $\mathbf{Y} = (X_1 - \overline{X}_n, ..., X_n - \overline{X}_n)'$ является несингулярным гауссовским вектором.*

Доказательство леммы 2. Гауссовость вытекает из леммы 1. ⑪ Для доказательства несингулярности заметим, что для $i \ne j$

$$\mathsf{E}\left[(X_i - \overline{X}_n)(X_j - \overline{X}_n)\right] = \mathsf{E}\left[X_i X_j\right] - \mathsf{E}\left[X_i \overline{X}_n\right] - \mathsf{E}\left[X_j \overline{X}_n\right] + \mathsf{E}\left[\overline{X}_n^2\right] = -\frac{1}{n},$$

так как для любого $1 \le k \le n$

$$\mathsf{E}\left[X_k \overline{X}_n\right] = \frac{1}{n}, \qquad \mathsf{E}\left[\overline{X}_n^2\right] = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathsf{E}\left[X_i X_j\right] = \frac{1}{n}.$$

Задача 7. *Доказать, что если случайные величины $X_n \in \mathcal{N}(0,1)$, $1 \le k \le n$, независимы в совокупности, то*

$$\mathsf{E}\left[X_k \overline{X}_n\right] = \frac{1}{n}, \qquad \mathsf{E}\left[\overline{X}_n^2\right] = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathsf{E}\left[X_i X_j\right] = \frac{1}{n}.$$

Кроме того, для любого $1 \le k \le n$

$$\mathsf{E}\left[(X_k - \overline{X}_n)^2\right] = \mathsf{E}\left[X_k^2\right] - 2\mathsf{E}\left[X_k \overline{X}_n\right] + \mathsf{E}\left[\overline{X}_n^2\right] = 1 - \frac{1}{n}.$$

Таким образом, если обозначить $a = -\frac{1}{n}$, $b = 1 + a$, $c = b/a$, то

$$\mathsf{Cov}\left[\mathbf{Y}\right] = \begin{pmatrix} b & a & a & a & \dots & a & a \\ a & b & a & a & \dots & a & a \\ a & a & b & a & \dots & a & a \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a & a & a & a & \dots & a & b \end{pmatrix} = a \cdot \begin{pmatrix} c & 1 & 1 & 1 & \dots & 1 & 1 \\ 1 & c & 1 & 1 & \dots & 1 & 1 \\ 1 & 1 & c & 1 & \dots & 1 & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & 1 & \dots & 1 & c \end{pmatrix}.$$

Поскольку $\det(\mathsf{Cov}\left[\mathbf{Y}\right]) \ne 0$, то вектор $\mathbf{Y}$ является несингулярным. □

Лемма 3. *Пусть $n > 1$. Если случайные величины $X_1, \dots, X_n$ независимы в совокупности и $X_k \in \mathcal{N}(0,1)$, $1 \le k \le n$, то случайная величина $\overline{X}_n$ и случайный вектор $(X_1 - \overline{X}_n, \dots, X_n - \overline{X}_n)'$ независимы.*

Доказательство. Найдем $\mathsf{cov}\left[\overline{X}_n, X_k - \overline{X}_n\right]$, $1 \le k \le n$. Так как $\mathsf{E}\left[\overline{X}_n\right] = 0$, ⑫ то $\mathsf{E}\left[X_k - \overline{X}_n\right] = 0$. Зафиксируем $1 \le k \le n$. Поскольку

$$\mathsf{cov}\left[\overline{X}_n, X_k - \overline{X}_n\right] = \mathsf{E}\left[X_k \overline{X}_n\right] - \mathsf{E}\left[\overline{X}_n^2\right],$$

то $\mathsf{cov}\left[\overline{X}_n, X_k - \overline{X}_n\right] = 0$. Таким образом, $(\overline{X}_n, X_1 - \overline{X}_n, \dots, X_n - \overline{X}_n)'$ является несингулярным (см. лемму 2) ⑬ гауссовским $(n+1)$-вектором (см. лемму 1). Согласно теореме 5 случайная величина $\overline{X}_n$ и вектор $(X_1 - \overline{X}_n, \dots, X_n - \overline{X}_n)'$ независимы. □

[11] Как гауссовость в лемме 2 вытекает из леммы 1?
[12] Доказать, что $\mathsf{E}\left[\overline{X}_n\right] = 0$.
[13] Почему этот вектор несингулярен?

Лемма 4. *Пусть случайная величина ξ и случайный вектор $\boldsymbol{\eta} = (\eta_1, \dots, \eta_n)'$ независимы. Обозначим $\boldsymbol{\zeta} = (\eta_1^2, \dots, \eta_n^2)'$. Тогда ξ и $\boldsymbol{\zeta}$ также независимы.*

Доказательство леммы 4. Пусть $A \in \mathscr{B}^1$ и $B \in \mathscr{B}^n$ произвольные борелевские множества. Согласно определению (1),

$$\mathsf{P}(\xi \in A, \boldsymbol{\eta} \in B) = \mathsf{P}(\xi \in A)\mathsf{P}(\boldsymbol{\eta} \in B).$$

В этом равенстве выберем $A = (-\infty, x)$, $x \in \mathbf{R}$, и $C = (-\infty, z_1) \times \dots \times (-\infty, z_n)$, $z_1, \dots, z_n \in \mathbf{R}$. Если $z_1 \ge 0, \dots, z_n \ge 0$, то обозначим $y_k = \sqrt{z_k}$, $1 \le k \le n$, и $B \overset{\text{def}}{=} (-y_1, y_1) \times \dots \times (-y_n, y_n)$. Тогда $\{\boldsymbol{\zeta} \in C\} = \{\boldsymbol{\eta} \in B\}$. ⑭ Поэтому

$$\mathsf{P}(\xi \in A, \boldsymbol{\zeta} \in C) = \mathsf{P}(\xi \in A, \boldsymbol{\eta} \in B) = \mathsf{P}(\xi \in A)\mathsf{P}(\boldsymbol{\eta} \in B) = \mathsf{P}(\xi \in A)\mathsf{P}(\boldsymbol{\zeta} \in C).$$

Если $z_k < 0$ для какого-нибудь k, то равенство

$$\mathsf{P}(\xi \in A, \boldsymbol{\zeta} \in C) = \mathsf{P}(\xi \in A)\mathsf{P}(\boldsymbol{\zeta} \in C)$$

все равно выполняется, так как $\{\boldsymbol{\zeta} \in C\} = \varnothing$. ⑮ В силу определения 1 и и теоремы 1 случайная величина ξ и вектор $\boldsymbol{\zeta}$ независимы. ⑯ □

Из лемм 3 и 4 вытекает, что

(8) $\overline{X}_n$ и вектор $((X_1 - \overline{X}_n)^2, \dots, (X_n - \overline{X}_n)^2)'$ независимы. ⑰

Лемма 5. *Пусть случайная величина ξ и случайный вектор $\boldsymbol{\eta} = (\eta_1, \dots, \eta_n)'$ независимы. Тогда независимы случайные величины ξ и $\eta_1 + \dots + \eta_n$.*

Доказательство леммы 5. Идею доказательства продемонстрируем для случая $n = 2$. Упорядочим рациональные числа по какому-либо правилу и пусть $\mathbf{Q} = \{q_1, q_2, \dots\}$. Обозначим $S = \eta_1 + \eta_2$. Тогда для любого $y \in \mathbf{R}$

$$\{S < y\} = \bigcup_{l=1}^{\infty} \{\eta_1 < q_l, \eta_2 < y - q_l\}. \tag{9}$$

Действительно, если элементарное событие ω принадлежит правой части (9), то для некоторого рационального числа $r = q_l$ выполнены два неравенства $\eta_1(\omega) < r$ и $\eta_2(\omega) < y - r$. ⑱ Отсюда вытекает $S(\omega) < y$, что и доказывает включение $\omega \in \{S < y\}$.

Наоборот, пусть $\omega \in \{S < y\}$, то есть $S(\omega) < y$. Если бы при этом для любого рационального числа $r > \eta_1(\omega)$ выполнялось $\eta_2(\omega) \ge y - r$, то мы бы перешли к пределу в последнем неравенстве при $r \downarrow \eta_1(\omega)$ и получили бы $\eta_2(\omega) \ge y - \eta_1(\omega)$,

[14] Почему $\{\boldsymbol{\zeta} \in C\} = \{\boldsymbol{\eta} \in B\}$?

[15] Почему $\{\boldsymbol{\zeta} \in C\} = \varnothing$, если $z_k < 0$ для хотя-бы одного k?

[16] Почему $\mathsf{P}(\xi \in A, \boldsymbol{\zeta} \in C) = \mathsf{P}(\xi \in A)\mathsf{P}(\boldsymbol{\zeta} \in C)$, если $z_k < 0$ для хотя-бы одного k?

[17] Почему из лемм 3 и 4 вытекает (8)?

[18] Почему $\eta_1(\omega) < r$ и $\eta_2(\omega) < y - r$ для некоторого рационального числа r, если элементарное событие ω принадлежит правой части (9)?

что несовместимо с условием $S(\omega) < y$. Следовательно для какого-то рационального $r > \eta_1(\omega)$ обязательно выполнено $\eta_2(\omega) < y - r$, откуда и вытекает принадлежность ω правой части (9), поскольку $\omega \in \{\eta_1 < q_l, \eta_2 < y - q_l\}$, если $q_l = r$.

Применяя теперь лемму о расщеплении, доказываем, что

$$\{\xi < x, S < y\} = \bigcup_{l=1}^{\infty} \{\xi < x, \eta_1 < q_l, \eta_2 < y - q_l\}.$$

В силу непрерывности вероятности

$$\mathsf{P}(\xi < x, S < y) = \lim_{L\to\infty} \mathsf{P}\Bigg(\bigcup_{l=1}^{L} \{\xi < x, \eta_1 < q_l, \eta_2 < y - q_l\}\Bigg), \quad ⑲$$

$$\mathsf{P}(S < y) = \lim_{L\to\infty} \mathsf{P}\Bigg(\bigcup_{l=1}^{L} \{\eta_1 < q_l, \eta_2 < y - q_l\}\Bigg). \quad ⑲$$

Докажем, что для любого $L \geq 1$

$$\mathsf{P}\Bigg(\bigcup_{l=1}^{L} \{\xi < x, \eta_1 < q_l, \eta_2 < y - q_l\}\Bigg) = \mathsf{P}(\xi < x)\mathsf{P}\Bigg(\bigcup_{l=1}^{L} \{\eta_1 < q_l, \eta_2 < y - q_l\}\Bigg). \tag{10}$$

Мы ограничимся доказательством равенства (10) только для случаев $L = 1$ и $L = 2$.

Случай $L = 1$. Будем пользоваться обозначениями $A(x)$ и $B(\mathbf{y})$, введенными формулами (2) и (4). Тогда при $r = q_l$

$$\{\xi < x\} = \{\xi \in A(x)\}, \qquad \{\eta_1 < r, \eta_2 < y - r\} = \{\boldsymbol{\eta} \in B(\mathbf{y})\},$$

где $\boldsymbol{\eta} = (\eta_1, \eta_2)'$ и $\mathbf{y} = (r, y - r)'$. Аналогично

$$\{\xi < x, \eta_1 < r, \eta_2 < y - r\} = \{\xi \in A(x), \boldsymbol{\eta} \in B(\mathbf{y})\}.$$

Поэтому $\mathsf{P}(\xi \in A(x), \boldsymbol{\eta} \in B(\mathbf{y})) = \mathsf{P}(\xi \in A(x))\mathsf{P}(\boldsymbol{\eta} \in B(\mathbf{y}))$ согласно свойству (1), что и требовалось доказать.

Случай $L = 2$. Положим $U = \{\xi < x\}$ и $V_{kl} = \{\eta_1 < q_k, \eta_2 < y - q_l\}$. Тогда события U и V_{kl} независимы для любых $k, l \geq 1$ согласно определению 1, поскольку $U = A(x)$ и $V_{kl} = B(q_k, y - q_l)$ (см. обозначение (2)). Кроме того

$$V_{11} \cap V_{22} = \{\eta_1 < \min\{q_k, q_l\}, \eta_2 < y - \max\{q_k, q_l\}\}.$$

Действительно, событие $V_{11} \cap V_{22}$ равносильно выполнению четырех неравенств $\eta_1 < q_k$, $\eta_1 < q_l$, $\eta_2 < y - q_k$ и $\eta_2 < y - q_l$, ⑳ которые эквивалентны совокупности двух неравенств

$$\eta_1 < \min\{q_1, q_2\} \quad \text{и} \quad \eta_2 < y - \max\{q_1, q_2\}. \quad ㉑$$

[19] Как это вытекает из непрерывности вероятности?

[20] Почему событие $V_{11} \cap V_{22}$ равносильно выполнению четырех неравенств $\eta_1 < q_k$, $\eta_1 < q_l$, $\eta_2 < y - q_k$ и $\eta_2 < y - q_l$?

Значит $V_{11} \cap V_{22} = V_{12}$, если $q_k < q_l$, или $V_{11} \cap V_{22} = V_{21}$, если $q_k > q_l$. В любом случае события U и $V_{11} \cap V_{22}$ независимы согласно результату для $L = 1$, доказанному выше.

Следовательно, при $q_k < q_l$

$$\begin{aligned}
\mathsf{P}\left(\bigcup_{l=1}^{2}\{\xi < x, \eta_1 < q_l, \eta_2 < y - q_l\}\right) &= \mathsf{P}([U \cap V_{11}] \cup [U \cap V_{22}]) \\
&= \mathsf{P}(U \cap V_{11}) + \mathsf{P}(U \cap V_{22}) - \mathsf{P}(U \cap V_{12}) \\
&= \mathsf{P}(U)\left[\mathsf{P}(V_{11}) + \mathsf{P}(V_{22}) - \mathsf{P}(V_{12})\right] \\
&= \mathsf{P}(U)\left[\mathsf{P}(V_{11}) + \mathsf{P}(V_{22}) - \mathsf{P}(V_{11} \cap V_{22})\right] \\
&= \mathsf{P}(U)\mathsf{P}(V_{11} \cup V_{22}).
\end{aligned}$$

[22] Отсюда и вытекает равенство (10) при $L = 2$. □

Задача 8. *Доказать* (10) *для любого* $L \geq 2$.

Задача⋆ 9. *Доказать лемму* 5 *для любого* n. <u>**Указание**</u>. Воспользоваться формулой включений-исключений для объединения событий.

Задача 10. *Доказать теорему* 6 *для произвольных* $a \in \mathbf{R}$ *и* $\sigma^2 > 0$. <u>**Указание**</u>. Вывести общий случай из уже доказанного.

Итак, в силу леммы 5 и свойства (8) случайные величины

$$\overline{X}_n \quad \text{и} \quad (X_1 - \overline{X}_n)^2 + \cdots + (X_n - \overline{X}_n)^2$$

независимы. Поэтому для любого $c \in \mathbf{R}$ независимы и случайные величины $\overline{X}_n$ и $c\sum_{k=1}^{n}(X_k - \overline{X}_n)^2$, [23] в частности, они независимы для $c = \frac{1}{n}$. □

Существует более общий результат, чем теорема 6.

5.1. Теорема Дели. *Пусть* $\mathbf{X} \in \mathcal{N}_n\left(\boldsymbol{\mu}, \sigma^2\mathbf{I}\right)$. *Предположим, что функция* $g : \mathbf{R}^n \to \mathbf{R}$ *обладает свойством*:

$$g(\mathbf{x}) = g(\mathbf{x} + a \cdot \mathbf{1})$$

для любых векторов $\mathbf{x} \in \mathbf{R}^n$ *и чисел* $a \in \mathbf{R}$, *где* $\mathbf{1} = (1, \ldots, 1)'$. *Тогда случайные величины* $\overline{X}_n$ *и* $g(\mathbf{X})$ *независимы.*

Замечание 2. Ясно, что теорема 6 вытекает из теоремы Дели для функции $g(\mathbf{x}) = \frac{1}{n}\sum_{k=1}^{n}(x_k - \overline{x}_n)^2$. [24]

5.2. Следствие для размаха выборки. Теорема Дели применима и для других функций g, например для $g(\mathbf{x}) = \max x_k - \min x_k$. Разность $X_{(n)} - X_{(1)}$ называется *размахом выборки.*

---

[21] Почему совокупность неравенств $\eta_1 < q_k$, $\eta_1 < q_l$, $\eta_2 < y - q_k$, $\eta_2 < y - q_l$ равносильна совокупности неравенств $\eta_1 < \min\{q_1, q_2\}$, $\eta_2 < y - \max\{q_1, q_2\}$?

[22] Какие изменения надо сделать в этой части доказательства при $q_k > q_l$?

[23] Почему независимы $\overline{X}_n$ и $c\sum_{k=1}^{n}(X_k - \overline{X}_n)^2$ для любого c, если независимы $\overline{X}_n$ и $(X_1 - \overline{X}_n)^2 + \cdots + (X_n - \overline{X}_n)^2$?

[24] Доказать, что теорема 6 вытекает из теоремы 5.1.

Теорема 7. *Пусть* $\mathbf{X} \in \mathcal{N}_n\left(\boldsymbol{\mu}, \sigma^2\mathbf{I}\right)$. *Положим* $X_{(1)} = \min_{1\le k\le n} X_k$ *и* $X_{(n)} = \max_{1\le k\le n} X_k$. *Тогда случайные величины* $\overline{X}_n$ *и* $X_{(n)} - X_{(1)}$ *независимы.*

Доказательство теоремы Дели. Пусть $\mathbf{X} = (X_1, \dots, X_n)'$, $S_n = X_1 + \dots + X_n$, $\mathbf{Y} = (S_n, X_2 - X_1, X_3 - X_1, \dots, X_n - X_1)'$. Для любого вектора $\mathbf{c} \in \mathbf{R}^n$ найдется другой вектор $\mathbf{b} \in \mathbf{R}^n$, при котором $\mathbf{c}'\mathbf{Y} = \mathbf{b}'\mathbf{X}$, [25] значит $\mathbf{Y}$ — гауссовский вектор. Поскольку $\mathsf{cov}\left[S_n, X_k - X_1\right] = 0$ для любого $1 \le k \le n$, [26] то S_n не зависит от $(X_2 - X_1, X_3 - X_1, \dots, X_n - X_1)'$ в силу теоремы 4.

По условию $g(x_1, \dots, x_n) = g(x_1 + a, \dots, x_n + a)$ для любых $x_1, \dots, x_n$ и a. [27] Выбрав $a = -x_1$, получаем $g(x_1, \dots, x_n) = g(0, x_2 - x_1, \dots, x_n - x_1)$. Как и при доказательстве лемм 4 и 5, показываем, что S_n не зависит от $h(X_2 - X_1, X_3 - X_1, \dots, X_n - X_1)$ для любой измеримой функции h.

Задача 11. *Пусть* $\mathbf{X}' = (X_1, \dots, X_n)$, $S_n = X_1 + \dots + X_n$. *Доказать, что* S_n *не зависит от* $h(X_2 - X_1, X_3 - X_1, \dots, X_n - X_1)$ *для любой измеримой функции* h.

Полагая

$$h(x_2 - x_1, x_3 - x_1, \dots, x_n - x_1) = g(0, x_2 - x_1, \dots, x_n - x_1),$$

доказываем, что S_n и $g(0, X_2 - X_1, \dots, X_n - X_1)$ независимы. Поэтому S_n и $g(X_1, X_2, \dots, X_n)$ также независимы. □

6. Ортогональные преобразования гауссовских векторов

Ортогональными мы называем преобразования вида $\mathbf{Y} = \mathbf{C}\mathbf{X}$, если $\mathbf{C}$ — ортогональная матрица. Пусть $\mathbf{X} \in \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Lambda})$.

6.1. Случай $\boldsymbol{\Lambda} = \mathbf{C}\mathbf{D}\mathbf{C}'$. Напомним, что существует ортогональная матрица $\mathbf{C}$, при которой $\boldsymbol{\Lambda} = \mathbf{C}\mathbf{D}\mathbf{C}'$, причем главная диагональ диагональной матрицы $\mathbf{D}$ составлена из собственных чисел матрицы $\boldsymbol{\Lambda}$ (утверждение 17.1). Оказывается, что координаты вектора $\mathbf{C}'\mathbf{X}$ независимы.

Теорема 8. *Пусть* $n > 1$ *и* $\mathbf{X} \in \mathcal{N}_n(\boldsymbol{\mu}, \boldsymbol{\Lambda})$, *причем* $\boldsymbol{\Lambda} = \mathbf{C}\mathbf{D}\mathbf{C}'$. *Положим* $\mathbf{Y} \stackrel{\text{def}}{=} \mathbf{C}'\mathbf{X}$. *Тогда* $\mathbf{Y} \in \mathcal{N}(\mathbf{C}'\boldsymbol{\mu}, \mathbf{D})$. *В частности, координаты* $Y_1, \dots, Y_n$ *вектора* $\mathbf{Y}$ *независимы, а их дисперсии равны собственным числам матрицы* $\boldsymbol{\Lambda}$.

Замечание 3. Некоторые из собственных чисел могут равняться 0. В этом случае соответствующие координаты вырождены (равны константе с вероятностью 1).

Доказательство. Из теоремы 16.2 для $\mathbf{B} = \mathbf{C}'$ и $\mathbf{b} = \mathbf{0}$ вытекает, что $\mathsf{E}\left[\mathbf{Y}\right] = \mathbf{C}'\boldsymbol{\mu}$, $\mathsf{Cov}\left[\mathbf{Y}\right] = \mathbf{C}'\boldsymbol{\Lambda}\mathbf{C} = \mathbf{D}$. Поэтому координаты вектора $\mathbf{Y}$ некоррелированы, так как $\mathbf{D}$ — диагональная матрица.

Теперь из теоремы 4 выводим, что координаты вектора $\mathbf{Y}$ независимы в совокупности. Наконец, главная диагональ ковариационной матрицы $\mathsf{Cov}\left[\mathbf{Y}\right]$ составлена из $\mathsf{var}\left[Y_1\right], \dots, \mathsf{var}\left[Y_n\right]$; с другой стороны, $\mathsf{Cov}\left[\mathbf{Y}\right] = \mathbf{D}$, то есть главная диагональ матрицы $\mathsf{Cov}\left[\mathbf{Y}\right]$ составлена из собственных чисел матрицы $\boldsymbol{\Lambda}$. □

Одним из следствий теоремы 8 является следующий результат.

[25] Почему для любого $\mathbf{c} \in \mathbf{R}^n$ найдется $\mathbf{b} \in \mathbf{R}^n$, при котором $\mathbf{c}'\mathbf{Y} = \mathbf{b}'\mathbf{X}$?
[26] Почему $\mathsf{cov}\left[S_n, X_k - X_1\right] = 0$ для любого $1 \le k \le n$?
[27] Почему $g(x_1, \dots, x_n) = g(x_1 + a, \dots, x_n + a)$ для любых $x_1, \dots, x_n$ и a?

Следствие 2. *Гауссовский вектор* $\mathbf{X}' = (X_1, \ldots, X_n)$ *имеет нулевой вектор математических ожиданий и единичную ковариационную матрицу* $\mathbf{I}$, *то есть* $\mathbf{X} \in \mathcal{N}_n(\mathbf{0}, \mathbf{I})$, *тогда и только тогда, когда* $X_1, \ldots, X_n$ *— независимые стандартные гауссовские случайные величины.* ㉘

6.2. Случай произвольной ортогональной матрицы C. Важным частным случаем теоремы 6 является $\mathsf{var}[X_1] = \cdots = \mathsf{var}[X_n]$. При таком предположении координаты вектора $\mathbf{CX}$ являются независимыми при любой ортогональной матрице $\mathbf{C}$, а не только для той, которая входит в разложение ковариационной матрицы вектора $\mathbf{X}$.

Теорема 9. *Пусть* $\mathbf{X} \in \mathcal{N}\left(\boldsymbol{\mu}, \sigma^2\mathbf{I}\right)$ *и* $\sigma^2 > 0$, *а ортогональная матрица* $\mathbf{C}$ *произвольна. Положим* $\mathbf{Y} = \mathbf{CX}$. *Тогда* $\mathbf{Y} \in \mathcal{N}\left(\mathbf{C}\boldsymbol{\mu}, \sigma^2\mathbf{I}\right)$. *В частности,* $Y_1, \ldots, Y_n$ *независимые случайные величины с дисперсией* σ^2.

Доказательство. Обозначим $\boldsymbol{\Lambda} = \mathsf{Cov}[\mathbf{X}]$. Тогда $\boldsymbol{\Lambda} = \sigma^2\mathbf{I}$. Из теоремы 16.2 для $\mathbf{B} = \mathbf{C}$ и $\mathbf{b} = \mathbf{0}$ вытекает, что $\mathsf{Cov}[\mathbf{Y}] = \mathbf{C}\boldsymbol{\Lambda}\mathbf{C}' = \sigma^2\mathbf{I} \stackrel{\text{def}}{=} \mathbf{D}$, то есть $\mathbf{CDC}' = \boldsymbol{\Lambda}$ и теорема 9 вытекает из теоремы 8. $\square$

6.3. Другое доказательство независимости $X-Y$ **и** $X+Y$**.** Дадим другое доказательство утверждения 1. Пусть $\mathbf{X} \in \mathcal{N}_2(\mathbf{0}, \mathbf{I})$. Матрица $\mathbf{C} \stackrel{\text{def}}{=} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$ является ортогональной. ㉙ Согласно теореме 9, компоненты вектора $\mathbf{Y} \stackrel{\text{def}}{=} \mathbf{CX}$ независимы. Легко видеть, что $\mathbf{CX} = \left(\frac{X_1+X_2}{\sqrt{2}}, \frac{X_1-X_2}{\sqrt{2}}\right)'$. ㉚ Значит $X_1 + X_2$ и $X_1 - X_2$ независимы. ㉛

[28] Доказать следствие 2.

[29] Проверить, что матрица $\mathbf{C}$ является ортогональной.

[30] Доказать, что $\mathbf{CX} = \left(\frac{X_1+X_2}{\sqrt{2}}, \frac{X_1-X_2}{\sqrt{2}}\right)'$.

[31] Как закончить доказательство независимости $X_1 + X_2$ и $X_1 - X_2$?

[†] Всего в тексте было 31 вопросов