Традиционна теория на вероятностите

Леда Д. Минкова

10 януари 2015 г.

Съдържание

Глава 1

Измеримо пространство от събития

1.1 Пространство от елементарни изходи

Много житейски ситуации зависят от шанса на дадено събитие да се сбъдне или не. От друга страна, сбъдването или не на дадено събитие зависи от редица условия. Тези условия се наричат експеримент. Резултатите от експеримента се наричат **елементарни събития**. В общия случай, изходът от експеримента не може да се предскаже със сигурност. Можем само да знаем всички възможни изходи при провеждане на експеримента.

Определение 1.1 *Множеството* Ω *от всички елементарни събития се нарича* **основно пространство**.

Пример 1.1 *При хвърляне на монета има два възможни изхода. Основното пространство* $\Omega = \{\omega_1, \omega_2\}$ *съдържа елементарните събития*

ω_1 = "Герб" и ω_2 = "Лице". При провеждане на опита се реализира точно едно от елементарните събития.

Пример 1.2 *При хвърляне на зар има шест възможни изхода. В този случай*

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

Основното пространство Ω се нарича *крайно*, ако съдържа краен брой елементи. Множеството е *изброимо*, ако съдържа изброимо много елементи, т.е. всички елементи могат да се поставят в съответствие на целите положителни числа.

Пример 1.3 *При хвърляне на монета до първа поява на лице, основното пространство съдържа изброимо много елементарни събития*

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_n, \dots\},$$

където $\omega_n = \{$пада се лице при $n -$ тия опит$\}$.

Ако при хвърляне на зар се интересуваме от това, ще се паднат ли четен брой точки, то това е събитието $A = \{2, 4, 6\}$, състоящо се от три елементарни събития. Събитието A е подмножество на Ω.

Определение 1.2 *Всяко подмножество на Ω се нарича* **събитие.**

Събитията се означават с главни латински букви $A, B, C, \dots$.

1.2 Действия със събития

Обичайните операции върху множества и техните свойства се пренасят без изменение върху събитията.

Ако елементарното събитие ω се сбъдва и $\omega \in A$, то се сбъдва и събитието A. Казваме, че елементарния изход ω е благоприятен за сбъдването на A.

Обединение (или **сума**) на две събития A и B се нарича събитието $A \cup B$, което се сбъдва когато се сбъдне поне едно от събитията A или B.

Сечение (или **произведение**) на събитията A и B се нарича събитието $A \cap B$, което се сбъдва когато се сбъднат едновременно A и B.

Събитието $\{A$ не се сбъдва$\}$ се нарича **допълнение** на A, означава се $\overline{A}$ и съдържа елементарните събития от Ω, които не се съдържат в A.

Определенията за сума и произведение на две събития се обобщават за произволно, крайно или безкрайно множество от събития $\{A_k\}$: $\bigcup\limits_k A_k$ означава, че се сбъдва поне едно от събитията, а $\bigcap\limits_k A_k$ означава, че се сбъдват всичките събития едновременно.

Празното множество $\emptyset$ се нарича **невъзможно събитие**, което при условията на експеримента никога не се сбъдва. Например, при хвърляне на зар да се паднат седем точки. Ω се нарича **достоверно събитие,** което винаги се сбъдва. Например, при хвърляне на два зара, събитието $A = \{$сумата от точките е не по-малка от 2$\}$.

Събитията A и B се наричат **несъвместими**, ако $A \cap B = \emptyset$. В този случай $A \cup B$, може да се замени с $A + B$.

Ако за събитията A и B е в сила $A \subset B$, т.е. при сбъдването на A се сбъдва и B, казваме, че A **влече след себе си** B, или B **следва** от A.

Ако се сбъдват едновременно $A \subset B$ и $B \subset A$, то това означава, че при всеки експеримент събитията A и B едновременно настъпват или не настъпват. В този случай събитията A и B се наричат **еквивалентни** и това се изразява с равенство, $A = B$.

Събитието, при което A се сбъдва, а B не се сбъдва се нарича **разлика** на A и B и се означава $A \backslash B$.

1.3 Редици от събития.

Нека $A_1, A_2, \dots$ е редица от подмножества на Ω. Ако $A_1 \subset A_2 \subset \dots$ и $\bigcup_{n=1}^{\infty} A_n = A$ то редицата $\{A_n\}$ се нарича растяща редица от множества с граница множеството A, и се означава $\lim_n \uparrow A_n = \bigcup_n A_n$, $(A_n \uparrow A)$.

Ако $A_1 \supset A_2 \supset \dots$ и $\bigcap_{n=1}^{\infty} A_n = A$ то редицата $\{A_n\}$ е намаляваща, има граница A, и се означава $\lim_n \downarrow A_n = \bigcap A_n$, $(A_n \downarrow A)$.

За събитията са в сила законите на де Морган

$$\overline{\bigcup_n A_n} = \bigcap_n \overline{A_n}, \quad \overline{\bigcap_n A_n} = \bigcup_n \overline{A_n}.$$

Задача 1.1 *Да се покаже, че ако $A_n \uparrow A$ то $\overline{A}_n \downarrow \overline{A}$ и ако $A_n \downarrow A$ то $\overline{A}_n \uparrow \overline{A}$.*

Въвеждаме означенията

$$A^* = \limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k \quad \text{и} \quad A_* = \liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k. \tag{1.1}$$

Събитието A^* се сбъдва тогава и само тогава когато за всяко n съществува $k \geq n$, такова че A_k се сбъдва. С други думи, $\omega \in A^*$ тогава и само тогава когато $\omega \in A_n$ за безбройно много n. Аналогично $\omega \in A_*$ тогава и само тогава когато съществува n, такова че $\omega \in A_k$ за всяко $k \geq n$.

Събитието A^* се нарича **горна граница** на редицата $\{A_n\}$, а A_* - **долна граница.** Терминологията е аналогична на тази при редиците от реални числа. Ясно е, че $A_* \subset A^*$.

Задача 1.2 *Да се покаже, че*

$$\overline{\limsup_{n\to\infty} A_n} = \liminf_{n\to\infty} \overline{A_n}$$

$$\overline{\liminf_{n\to\infty} A_n} = \limsup_{n\to\infty} \overline{A_n}$$

$$\liminf_{n\to\infty} A_n \subset \limsup_{n\to\infty} A_n$$

Ако $A_n \uparrow A$ *или* $A_n \downarrow A$ *то*

$$\liminf_{n\to\infty} A_n = \limsup_{n\to\infty} A_n = A.$$

В общия случай

Определение 1.3 *Ако* $A_* = A^*$, *редицата* $\{A_n\}$ *се нарича сходяща и при* $n \to \infty$ *има граница* $\lim A_n = \liminf A_n = \limsup A_n$.

1.4 $\sigma-$ алгебра от събития

Нека заедно с основното пространство Ω да разгледаме непразното множество $\mathcal{F}$, от подмножества на Ω.

Определение 1.4 *Множеството* $\mathcal{F}$ *се нарича* **алгебра**, *ако:*

1. $\Omega \in \mathcal{F}$;
2. *Ако* $A \in \mathcal{F}$, *то* $\overline{A} \in \mathcal{F}$;
3. *Ако* $A_1, A_2, \dots A_n \in \mathcal{F}$, *то* $\bigcup_{i=1}^{n} A_i \in \mathcal{F}$.

От свойствата 1. и 2. следва, че $\emptyset \in \mathcal{A}$. Свойство 3. означава, че алгебрата е затворена относно крайно обединение. От второто свойство и законите на де Морган следва, че алгебрата е затворена и относно крайно сечение.

Определение 1.5 *Ако свойство 3. е в сила за изброимо множество от съботия, т.е.*

4. От $A_n \in \mathcal{F}$, *при* $n = 1, 2, \dots$ *следва* $\bigcup\limits_{n} A_n \in \mathcal{F}$,

множеството $\mathcal{F}$ *се нарича* $\sigma-$**алгебра**.

Така, $\sigma-$ алгебрата е е затворена относно изброимо много операции обединение и сечение.

Пример 1.4 *Най - малката $\sigma-$ алгебра, свързана с Ω е множеството $\mathcal{F}_0 = \{\emptyset, \Omega\}$. Нарича се тривиална $\sigma-$***алгебра**.

Пример 1.5 *Ако A е подмножество на Ω, то $\mathcal{F} = \{\emptyset, A, \overline{A}, \Omega\}$ е $\sigma-$алгебра, породена от A.*

Пример 1.6 *Нека $\mathcal{F}(\Omega)$ означава множеството от всички подмножества на Ω. Ако $\#(\Omega) = n$, ($\#(\Omega)$ означава броя на елементите в Ω), то $\#(\mathcal{F}(\Omega)) = 2^n$. За всяка алгебра $\mathcal{F}$ от подмножества на Ω е в сила $\mathcal{F}_0 \subset \mathcal{F} \subset \mathcal{F}(\Omega)$.*

Определение 1.6 *Множеството $\mathfrak{M}$ от подмножества на Ω се нарича монотонно (монотонен клас), ако е затворено относно* $\lim \uparrow$ *и* $\lim \downarrow$, *т.е*

$$A_n \in \mathcal{F},\ n \geq 1,\ A_n \uparrow\ \Longrightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$$

и

$$A_n \in \mathcal{F},\ n \geq 1,\ A_n \downarrow\ \Longrightarrow \bigcap_{n=1}^{\infty} A_n \in \mathcal{F}.$$

Теорема 1.1 *В сила са следните твърдения:*

а) Всяка $\sigma-$ алгебра е алгебра;

б) Една алгебра е $\sigma-$ алгебра тогава и само тогава когато е монотонен клас;

в) Всяка $\sigma-$ алгебра е монотонен клас;

г) Сечение на $\sigma-$ алгебри (изброимо или неизброимо) е отново $\sigma-$ алгебра.

Доказателство.

б) Ясно е, че всяка $\sigma-$ алгебра е монотонен клас. Трябва да се покаже, че затвореният относно крайно обединение клас $\mathfrak{M}$ от подмножества на Ω е затворен относно изброимо обединение тогава и само тогава когато е затворен относно $\lim\uparrow$. За това е достатъчно да се отбележи, че ако $\{A_n,\ n \geq 1\}$ е произволна редица от подмножества на Ω, то

$$\bigcup_n A_n = \lim_m \uparrow (\bigcup_{n\leq m} A_n).$$

□

Нека $\mathcal{U}$ е множество от подмножества на Ω, т.е. $\mathcal{U} \subseteq \mathcal{F}(\Omega)$. $\mathcal{F}$ се нарича минимална σ- алгебра, съдържаща $\mathcal{U}$, ако за всяка $\sigma-$ алгебра $\mathcal{G}$, за която $\mathcal{U} \subseteq \mathcal{G}$, следва, че $\mathcal{F} \subseteq \mathcal{G}$.

Определение 1.7 *Минималната $\sigma-$ алгебра, съдържаща дадено множество от подмножества се нарича $\sigma-$ алгебра, породена от това множество и се означава $\mathcal{F} = \sigma(\mathcal{U})$. Аналогично, минималният монотонен клас, съдържащ $\mathcal{U}$ се нарича монотонен клас породен от $\mathcal{U}$ и се означава $\mathfrak{M}(\mathcal{U})$.*

Теорема 1.2 *Нека $\mathcal{U}$ е алгебра. Тогава*

$$\mathfrak{M}(\mathcal{U})) = \sigma(\mathcal{U}).$$

Най-съществените примери са Бореловите $\sigma-$ алгебри.

Нека сега $\Omega = \mathbb{R}^1$, където $\mathbb{R}^1$ е реалната права и $\mathcal{U}_1$ е множеството от всички отворени интервали. σ- алгебрата $\mathcal{B}(\mathcal{U}_1)$ породена от $\mathcal{U}_1$ се нарича Борелова $\sigma-$ алгебра. Елементите на $\mathcal{B}(\mathcal{U}_1)$ се наричат Борелови множества. Аналогично се дефинира Борелова $\sigma-$ алгебра $\mathcal{B}(\mathbb{R}^n)$ върху $n-$ мерното Евклидово пространство

Определение 1.8 *Двойката $(\Omega, \mathcal{F})$, където $\mathcal{F}$ е алгебра върху Ω се нарича измеримо пространство.*

Елементите на $\mathcal{F}$ се наричат **случайни събития**. Ако $\mathcal{F}$ не съвпада с $\mathcal{F}(\Omega)$, то съществуват елементарни събития, които не са случайни събития.

1.5 Задачи

Задача 1.3 *1) Нека $\mathcal{U}$ е множеството от интервали от вида $(a,b] = \{x : 0 < x \le b\}, -\infty \le a < b < \infty$. Да се покаже, че $\mathcal{U}$ е алгебра, но не е $\sigma-$ алгебра.*

2) Нека $\mathcal{C}$ е множеството от затворени множества на $\mathbb{R}$. Да се покаже, че $\mathcal{B}(\mathcal{U}) = \sigma(\mathcal{C})$.

Глава 2

Вероятностно пространство

2.1 Вероятност

Аксиоми на вероятността На всяко събитие A се съпоставя число $P(A)$, такова че

A1. $P(A) \geq 0$ (неотрицателност);

A2. $P(\Omega) = 1$ (нормираност);

A3. $P(A \cup B) = P(A) + P(B), \quad A \cap B = \emptyset$.

Свойства на вероятността:

От аксиомите следва, че

1. $P(\overline{A}) = 1 - P(A)$;

2. $P(\emptyset) = 0$;

3. $P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$ за непресичащи се елементи A_i на $\mathcal{F}$ (**крайна адитивност**);

4. За произволни събития A и B, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (**формула за събиране на вероятности**).

5. $P(A \cup B) + P(A \cap B) = P(A) + P(B)$;

6. $P(A) \leq P(B), \quad$ ако $\quad A \subset B$;

Определение 2.1 $(\Omega, \mathcal{F}, P)$, *където* Ω *е крайно множество,* $\mathcal{F}$ *е алгебра,* P *е крайно адитивна вероятност се нарича крайно адитивно вероятностно пространство.*

Нека $(\Omega, \mathcal{F})$ е измеримо пространство, такова, че $\Omega = \{\omega_1, \omega_2, \dots \omega_n\}$ е крайно множество, а $\mathcal{F} = \mathcal{B}(\Omega)$ е множеството от всички подмножества на Ω. Нека на всяко елементарно събитие $\omega_k \in \Omega$ е съпоставено неотрицателно число $p(\omega_k)$, така, че

$$\sum_{k=1}^{n} p(\omega_k) = 1. \tag{2.1}$$

Определение 2.2 *Числото* $p(\omega_k)$ *се нарича вероятност на елементарното събитие* ω_k.

Определение 2.3 *Вероятност на събитието* $A \subseteq \Omega$ *се нарича числото*

$$P(A) = \sum_{\omega_k \in A} p(\omega_k), \tag{2.2}$$

където сумирането е по всички елементарни събития, принадлежащи на A.

2.2 Класическа схема

Да разгледаме частния случай, когато основното пространство Ω има краен брой елементарни събития, т.е. $\Omega = \{\omega_1, \omega_2, \dots \omega_n\}$ и на всяко елементарно събитие е съпоставено едно и също неотрицателно число $p = p(\omega_i),\ i = 1, 2, \dots, n$. Тогава, съгласно (2.2)

$$\sum_{i=1}^{n} p(\omega_i) = np = 1,$$

от където $p = \frac{1}{n}$, т.е. всяко елементарно събитие има вероятност равна на $\frac{1}{n} = \frac{1}{\#(\Omega)}$.

Тогава от (2.2) се получава

$$P(A) = \frac{\#(A)}{\#(\Omega)} = \frac{m}{n}. \tag{2.3}$$

Вероятността в този случай се нарича **класическа вероятност.** Накратко, класическата дефиниция на вероятност се прилага в случаите, когато при провеждане на опита има **краен брой, равновъзможни изходи**. Във формула (2.3), в знаменателя се съдържа броят на всички възможни изходи при провеждане на опита ($n = \#(\Omega)$). В числителя ($m = \#(A)$) се съдържа броят на изходите, благоприятстващи сбъдването на събитието A. Класическата дефиниция на вероятност е приложима, например, при задачите, свързани със зарове, карти, лотарии, спорт тото.

2.3 Условна вероятност

Нека разгледаме следната задача: От колода от 52 карти се избира една и тя се оказва червена. Каква е вероятността избраната карта да е дама? Ако означим събитието $A = \{$избраната карта е дама$\}$ и $B = \{$избраната карта е червена$\}$, то търсената вероятност може да се означи $P(A|B)$ и се нарича условна вероятност на A.

След като знаем, че избраната карта е червена, вече не се намираме в основното пространство Ω, което съдържа 52 елементарни събития. Новото пространство съдържа 26 елементарни събития, колкото са червените карти. Така $P(A|B) = \frac{2}{26} = \frac{1}{13}$.

Съществува формула, с която не се налага излизане от основното пространство Ω. Ако A и B са събития за които $P(A) > 0$ и $P(B) > 0$, то вероятността за съвместното им сбъдване се определя от

$$P(AB) = P(B)P(A|B) = P(A)P(B|A). \tag{2.4}$$

От (2.4) следва

$$P(A|B) = \frac{P(AB)}{P(B)}, \tag{2.5}$$

което се нарича вероятност на събитието A при условие че е настъпило събитието B, или просто **условна вероятност** на A. (2.4) се нарича формула за умножение на вероятности за две събития. Общата формула се доказва по индукция.

Определение 2.4 (Формула за умножение на вероятности) *Вероятността за съвместното сбъдване на събитията $A_1, A_2, \dots A_n$ се изразява с формулата*

$$P(A_1 A_2 \dots A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1, A_2) \dots P(A_n|A_1, A_2, \dots A_{n-1}).$$

2.4 Независимост

В пример видяхме, че в общия случай $P(A)$ и $P(A|B)$ са различни, т.е. настъпването на B може да промени вероятността за настъпване на A.

Определение 2.5 (Независимост) *Събитието A се нарича независимо от B, ако*

$$P(A|B) = P(A). \tag{2.6}$$

За независими събития A и B с $P(A) > 0$ и $P(B) > 0$ (2.6) приема вида

$$P(AB) = P(A)P(B). \tag{2.7}$$

Може да се покаже, че и от (2.7) следва (2.6). И двете равенства характеризират свойството независимост на две събития.

При независимост на повече от две събития се налагат някои уточнения.

Определение 2.6 (Независимост на $n > 2$ събития) *Събитията $A_1, A_2, \dots A_n$ са съвместно независими или просто независими, ако за* *всяко* $k = 2, 3, \ldots n$ *е в сила*

$$P(A_{i_1} \ldots A_{i_k}) = \prod_{j=1}^{k} P(A_{i_j}), \tag{2.8}$$

където $i_1, \ldots i_k$ *е произволно подмножество на* $1, 2, \ldots n$.

Ако (2.8) е в сила само при $k = 2$, то събитията се наричат **независими две по две**. Ясно е, че от независимост по двойки не следва съвместна независимост.

Определение 2.7 (Условна независимост) *Две събития* A *и* B *с* $P(A) > 0$ *и* $P(B) > 0$ *са независими при условие събитието* C*, ако*

$$P(AB|C) = P(A|C)P(B|C).$$

Ясно е, че от условна независимост не следва независимост. Може A и B да са независими при условие C, но да не са независими.

2.5 Формула за събиране на вероятности

За произволни случайни събития A и B е в сила следната *формула за събиране на вероятностите:*

$$P(A \cup B) = P(A) + P(B) - P(AB).$$

Когато A и B са несъвместими ($A \cap B = \emptyset$), формулата съвпада с Аксиома **А3**.

В общия случай, може да се намери вероятността за сбъдване на поне едно от събитията $A_1, \ldots, A_n,\ n \geq 2$. Формулата за събиране на вероятности има вида:

$$\begin{aligned} P(\textstyle\bigcup_{i=1}^n A_i) &= \textstyle\sum_{i=1}^n P(A_i) - \sum_{i=1}^{n-1} \sum_{j=i+1}^n P(A_i A_j) \\ &+ \textstyle\sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^n P(A_i A_j A_k) - \ldots + (-1)^{n-1} P(A_1 \ldots A_n). \end{aligned}$$

2.6 Формула за пълната вероятност и формула на Бейс

Да предположим, че събитието A се сбъдва заедно с едно от събитията $H_1, H_2, \dots H_n$, такива че $H_i \cap H_j = \emptyset,\ i \neq j,\ \ i, j = 1, 2, \dots n$ и $\bigcup_{i=1}^n H_i = \Omega$. Събития с това свойство образуват **пълна група от събития.** Събитията $H_1, H_2, \dots H_n$, се наричат хипотези. Вероятността на събитието A се определя по формулата

$$P(A) = \sum_{i=1}^{n} P(H_i)P(A|H_i) \tag{2.9}$$

и се нарича **формула за пълната вероятност.** Доказателството следва от това, че $AH_1, AH_2, \dots AH_n$ са несъвместими събития и $A = AH_1 + AH_2 + \dots + AH_n$. По правилото за събиране на вероятности намираме $P(A) = \sum_{i=1}^n P(AH_i)$. За всеки член на тази сума се прилага формулата за умножение на вероятности $P(AH_i) = P(H_i)P(A|H_i)$ и се получава (2.9).

Нека при същата схема е даден резултата от сбъдването на A. При това условие вероятностите на хипотезите се определят по формулата

$$P(H_j|A) = \frac{P(H_j)P(A|H_j)}{\sum_{i=1}^n P(H_i)P(A|H_i)},\ j = 1, 2, \dots n, \tag{2.10}$$

която се нарича **формула на Бейс.**

Доказателството следва от формулата за условна вероятност и формулата за умножение на вероятности.

2.7 Вероятност върху изброимо пространство

В този параграф се предполага, че пространството $\Omega = \{\omega_1, \omega_2, \dots\}$ съдържа изброимо много елментарни събития. В този случай крайната адитивност на случайните събития не е достатъчна. Измеримото пространство $(\Omega, \mathcal{F})$ е снабдено със $\sigma-$ алгебра $\mathcal{F}$, която съдържа всички подмножества на Ω.

Следващата теорема характеризира вероятността, определена на изброимо пространство.

Теорема 2.1 *Нека $p_k = p(\omega_k), \quad k = 1, 2, \dots$ е редица от реални числа, определени на Ω. Тогава $p_k \geq 0$ и $\sum_{k=1}^{\infty} p_k = 1$ е необходимо и достатъчно условие за съществуване на единствена вероятност P, за която $P(\{\omega_k\}) = p_k$.*

Доказателство. Нека приемем че $P(\{\omega_k\}) = p_k, \quad k = 1, 2, \dots$. Тогава $p_k \geq 0$ и

$$1 = P(\Omega) = P\left(\bigcup_{k=1}^{\infty}\{\omega_k\}\right) = \sum_{k=1}^{\infty} P(\{\omega_k\}) = \sum_{k=1}^{\infty} p_k.$$

Обратно, ако за числата $p_k \geq 0$ е в сила $\sum_{k=1}^{\infty} p_k = 1$, то може да се дефинира вероятност P, такава че за всяко $A \in \mathcal{F}$

$$P(A) = \sum_{\omega \in A} p(\omega),$$

при условие че празната сума е равна на нула. Тогава $P(\emptyset) = 0$ и $P(\Omega) = \sum_{k=1}^{\infty} p_k = 1$.

$\square$

В този случай може да се даде следното определение за вероятност.

Определение 2.8 *Вероятност на $(\Omega, \mathcal{F})$ се нарича всяка числова функция P, дефинирана върху $\mathcal{F}$, за която са в сила аксиомите:*

А1. $P(A) \geq 0$ *за всяко* $A \in \mathcal{F}$ **(неотрицателност)**;

А2. $P(\Omega) = 1$ **(нормираност)**;

А3*. *Ако събитията $\{A_n\}$ са две по две несъвместими (т.е. $A_i A_j = \emptyset, i \neq j$) и $\sum_{n=1}^{\infty} A_n \in \mathcal{F}$ то*

$$P(\sum_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n) \quad (\sigma - \textbf{адитивност}). \tag{2.11}$$

Може да се провери, че за така определената вероятност са в сила свойствата **1** - **6** на вероятност върху крайно пространство.

Следната теорема съдържа свойства, които са еквивалентни на хипотезата за $\sigma-$ адитивност.

Теорема 2.2 *Предполагаме, че вероятностната мярка P удовлетворява аксиомите* **A1** *и* **A2**. *Да се покаже, че* **A3*** *е еквивалентна на всяко от следните твърдения:*

а) Ако $A_n \in \mathcal{F}$ и $A_n \downarrow \emptyset$, то $P(A_n) \downarrow 0$;

б) Ако $A_n \in \mathcal{F}$ и $A_n \downarrow A$, то $P(A_n) \downarrow P(A)$.

в) Ако $A_n \in \mathcal{F}$ и $A_n \uparrow \Omega$, то $P(A_n) \uparrow 1$;

г) Ако $A_n \in \mathcal{F}$ и $A_n \uparrow A$, то $P(A_n) \uparrow P(A)$.

Доказателство. От това, че $P(A) + P(\overline{A}) = 1$ следва, че а) $\Longleftrightarrow$ в) и б) $\Longleftrightarrow$ г). Очевидно е, че от г) следва в).

Нека сега е всила в) и нека $A_n \in \mathcal{F}$ и $A_n \uparrow A$. Означаваме с $B_n = A_n \cup \overline{A},\ n = 1, 2, \ldots$, множествата, които покриват Ω. Тъй като $A_n \cap \overline{A} = \emptyset$, то $P(B_n) = P(A_n) + P(\overline{A})$ и при $n \to \infty \quad P(B_n) \to 1$, от където $P(A_n) = P(B_n) - P(\overline{A}) = P(B_n) - 1 + P(A) \to P(A)$.

Сега е в сила аксиомата **A3*** и за $A_n \in \mathcal{F} \quad A_n \uparrow A$.

Означаваме $B_k = A_k \setminus A_{k-1},\ k = 2, 3, \ldots$. Тогава $A = A_1 \cup B_2 \cup B_3 \cup \ldots$, където $A_1, B_2, B_3, \ldots$ са две по две несъвместими и

$$\begin{aligned} P(A) &= P(A_1) + P(B_2) + P(B_3) + \ldots \\ &= P(A_1) + \lim_{n\to\infty} \textstyle\sum_{k=2}^{n} P(B_k) \\ &= P(A_1) + \lim_{n\to\infty} \textstyle\sum_{k=2}^{n} [P(A_k) - P(A_{k-1}] = \lim_{n\to\infty} P(A_n). \end{aligned}$$

Обратното твърдение се получава аналогично.

□

Свойството (2.11) е в сила когато събитията $\{A_n\}$ са две по две несъвместими. При произволна редица от събития е в сила следната теорема.

Теорема 2.3 *За всяка редица от събития* $\{A_n\}$*, такава, че* $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$*, е в сила*

$$P(\bigcup_{n=1}^{\infty}) \leq \sum_{n=1}^{\infty} P(A_n) \quad (\sigma - \textit{полуадитивност});$$

Доказателство.Означаваме $B_n = \Omega - \sum_{k=1}^{n-1} A_k$. От това, че $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} A_n B_n$ следва, че $A_n B_n$ са несъвместими и

$$P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n B_n) \leq \sum_{n=1}^{\infty} P(A_n).$$

□

2.8 Закон за нулата и единицата на Борел-Кантели

Теорема 2.4 *Нека* $\{A_n, \ \ n = 1, 2, \ldots\}$ *е редица от независими събития на* $(\Omega, \mathcal{F}, P)$ *и нека* $A = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$ *(А съдържа* ω*, които принадлежат на безбройно много* A_k*.) Тогава* $P(A) = 1$ *или 0 в зависимост от това*

$$\sum_{k=1}^{\infty} P(A_k) = \infty \quad \textit{или} \quad \sum_{k=1}^{\infty} P(A_k) < \infty.$$

Доказателство.

$$P(A) = \lim_{n \to \infty} P(\bigcup_{k=n}^{\infty} A_k) \leq \lim_{n \to \infty} \sum_{k=n}^{\infty} P(A_k) = 0.$$

Следователно
$$\sum_{k=1}^{\infty} P(A_k) < \infty \quad \longrightarrow \quad P(A) = 0$$
дори за независими събития.

Нека $\sum_{k=1}^{\infty} P(A_k) = \infty$. Тогава

$$\begin{aligned} P(A) &= \lim_{n\to\infty} P(\bigcup_{k=n}^{\infty} A_k) = \lim_{n\to\infty} P(\Omega - \bigcap_{k=n}^{\infty} \overline{A}_k) \\ &= 1 - \lim_{n\to\infty} P(\bigcap_{k=n}^{\infty} \overline{A}_k) = 1 - \lim_{n\to\infty} \lim_{m\to\infty} P(\bigcap_{k=n}^{m} \overline{A}_k) \\ &= 1 - \lim_{n\to\infty} \prod_{k=n}^{\infty} (1 - P(A_k)). \end{aligned}$$

От неравенството $\ln(1-x) \leq -x$ следва
$$\prod_{k=n}^{\infty} (1 - P(A_k)) \leq e^{-\sum_{k=n}^{\infty} P(A_k)}.$$

Следователно
$$\prod_{k=n}^{\infty} (1 - P(A_k)) \leq e^{-\infty} = 0.$$

□

2.9 Вероятностни мерки

Да припомним Определение 1.8, съгласно която $(\Omega, \mathcal{F})$, където Ω е пространството от елементарни събития и $\mathcal{F}$ е $\sigma-$ алгебра се нарича измеримо пространство. Видяхме, че когато Ω е крайно или изброимо, конструирането на вероятностна мярка е лесно. Когато Ω е неизброимо, тази техника е неприложима. В този случай $P(\omega) = 0$ за всяко ω и множеството от стойности $P(\omega)$, $\omega \in \Omega$ в общия случай не може да характеризира вероятността. Трудностите се виждат от следния пример. Нека $\Omega = [0,1]$ и да въведем естествената вероятностна мярка $P((a,b]) = b - a$ върху интервалите от вида $(a, b]$, $0 \leq a \leq b \leq 1$. Ако искаме да разширим P, по единствен начин върху всички подмножества на $[0, 1]$ и да са в сила аксиомите от Определение 2.8, това е невъзможно. Съвкупността от тези множества е твърде голяма. Оказва се, че това може да се направи върху по-малка съвкупност от множества, т.е. върху $\sigma-$ алгебрата, съдържаща интервалите от вида $(a, b]$. Това се съдържа в следната теорема.

Теорема 2.5 *Да предположим, че $\sigma-$ алгебрата $\mathcal{F}$ е породена от алгебрата $\mathcal{F}_0$. Тогава всяка вероятностна мярка, дефинирана на алгебрата $\mathcal{F}_0$ притежава единствено продължение върху $\sigma-$ алгебрата $\mathcal{F}$.*

Продължението на вероятностната мярка се означава отново с P и тройката $(\Omega, \mathcal{F}, P)$ се нарича вероятностно пространство.

Глава 3

Случайни величини

Нека $(\Omega, \mathcal{F})$ е произволно измеримо пространство и $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ е реалната права с алгебра от борелови множества $\mathcal{B}(\mathbb{R})$.

Определение 3.1 *Реалната функция $X = X(\omega)$, дефинирана на $(\Omega, \mathcal{F})$, се нарича $\mathcal{F}$ -* **измерима функция** *или* **случайна величина***, ако за всяко $B \in \mathcal{B}(\mathbb{R})$*

$$X^{-1}(B) = \{\omega : X(\omega) \in B\} \in \mathcal{F}. \tag{3.1}$$

В случая, когато $\Omega = \mathbb{R}$ и $\mathcal{F} = \mathcal{B}$, функцията $X(\omega)$ се нарича *борелова.*

Същността на (3.1) се състои в това, че се разглеждат функции $X(\omega)$, при които праобразите на бореловите множества $X^{-1}(B)$ са събития.

Теорема 3.1 *Условието (3.1) е еквивалентно на*

$$\{\omega : X(\omega) \leq x\} \in \mathcal{F} \text{ за всяко } x \in R. \tag{3.2}$$

Случайните величини се означават с главни латински букви X, Y, Z или с гръцки букви ξ, η, ζ. Лесно се проверява, че множеството

$$\mathcal{F}_X = \{A : A = X^{-1}(B),\ B \in \mathbb{R}\}$$

е σ - алгебра.

Определение 3.2 *$\mathcal{F}_X$ се нарича σ - алгебра, породена от X. Означава се $\mathcal{F}_X = \sigma(X)$.*

3.1 Дискретни случайни величини

Нека $(\Omega, \mathcal{F}, P)$ е вероятностно пространство, при което Ω е изброимо и $\sigma-$ алгебрата $\mathcal{F}$ съдържа всички подмножества на Ω.

Определение 3.3 *Функцията $X = X(\omega)$, дефинирана на $(\Omega, \mathcal{F})$, със стойности в $\mathbb{R}$, такава че*

1) образът на Ω е изброимо подмножество на $\mathbb{R}$ и

2) за всяко $x \in \mathbb{R}$, $\{\omega \in \Omega : X(\omega) = x\} \in \mathcal{F}$, се нарича **дискретна случайна величина.**

Второто условие в дефиницията е еквивалентно на

$$X^{-1}(x) = \{\omega : X(\omega) = x\} \in \mathcal{F}. \tag{3.3}$$

Със всяка случайна величина X се определя вероятност

$$P^X(B) = P(\omega : X(\omega) \in B) = P(X^{-1}(B)) = P(X \in B), \tag{3.4}$$

която се нарича **вероятностно разпределение** на случайната величина X. Тази формула дефинира вероятностната мярка P. Тъй като множеството от стойности е най - много изброимо, то тази вероятност е напълно определена от числата

$$p^X(k) = P(X = k) = \sum_{\{\omega : X(\omega) = k\}} p(\omega).$$

Множеството от стойности $\{p^X(k)\}$ също се нарича вероятностно разпределение на X. Ясно е, че $P^X(A) = \sum_{\{k \in A\}} p^X(k)$.

Нека $A_1, A_2, \dots A_n$ е едно разлагане на Ω.

Определение 3.4 *Случайна величина, която приема краен брой стойности $x_1, x_2, \ldots, x_n$ и има вида*

$$X = \sum_{k=1}^{n} x_k I_{A_k}$$

се нарича елементарна.

В този случай

$$p_k = P(X = x_k) = P(A_k), \quad \sum_{k=1}^{n} p_k = 1.$$

Когато в (3.4), $B = (-\infty, x]$, се получава функцията $F_X(x) = P(X \leq x)$, определена върху цялата реална права, която се нарича функция на разпределение.

3.2 Функция на разпределение

Функцията на разпределение се дефинира за всяка случайна величина.

Нека $(\Omega, \mathcal{F}, P)$ е вероятностно пространство и X е случайна величина, дефинирана върху него.

Определение 3.5 *За всяко $x \in \mathbb{R}^1$, функцията*

$$F_X(x) = P(\{\omega : X(\omega) \leq x\}) = P(X \leq x), \tag{3.5}$$

се нарича **функция на разпределение** *(ф.р.) на случайната величина X.*

Когато няма опасност от недоразумение функцията на разпределение се записва $F(x)$.

С помощта на $F(x)$ може да се изрази следната вероятност

$$P(a < X \leq b) = F(b) - F(a).$$

Свойства на функцията на разпределение.

Непосредствено от дефиницията на $F(x)$ следва, че всяка ф.р. притежава следните свойства:

1. $0 \leq F(x) \leq 1$ за всяко x;

2. За $x_1 < x_2$ е в сила $F(x_1) \leq F(x_2)$(**монотонно растяща**);

3. Ако редицата $x_n \downarrow x_0$, то $\lim_{n\to\infty} F(x_n) = F(x_0)$ (**непрекъсната отдясно**);

4. $\lim_{x\to-\infty} F(x) = 0, \quad \lim_{x\to\infty} F(x) = 1.$

Непрекъснатост отдясно означава още, че за всяко $x \in \mathbb{R}^1$ и $\varepsilon > 0$

$$\lim_{\varepsilon\to 0}[F(x+\varepsilon) - F(x)] = 0,$$

което записваме кратко $F(x+0) = F(x)$.

Нека сега редицата $x_n \uparrow x_0$. Тъй като съответната редица $F(x_n)$ е монотонна и ограничена, тя е сходяща, т.е. $\lim_{n\to\infty} F(x_n) = F(x-0)$. В общия случай $p_x = F(x) - F(x-0)$. Ако $F(x)$ е непрекъсната в точката x, то $p_x = 0$.

Точките, в които $p_x > 0$ се наричат *точки на скок* (точки на прекъсване). В тези точки графиката на функцията $F(x)$ има скок с големина p_x.

Теорема 3.2 *Функцията на разпределение $F(x)$ може да има не повече от изброимо много скокове.*

Доказателство. Функцията на разпределение може да има не повече от 1 скок, по-голям от $\frac{1}{2}$; скоковете, при които $\frac{1}{4} < p_x \leq \frac{1}{2}$, са не повече от 3; при $\frac{1}{8} < p_x \leq \frac{1}{4}$ те са не повече от 7. В общия случай функцията на разпределение $F(x)$ може да има не повече от $2^n - 1$ такива скока, че $\frac{1}{2^n} < p_x \leq \frac{1}{2^{n-1}},\ n = 1, 2, \ldots$

Ясно е, че всички скокове могат да бъдат номерирани, т.е. те са не повече от изброимо много.

□

Показахме, че на всяка случайна величина се съпоставя една функция на разпределение $F(x)$, която е дрфинирана за всяко реално x. Може да се случи различни случайни величини да имат една и съща ф.р.

Пример: Нека Ω е съвкупността от всички числа в интервала $[0, 1]$ и $\mathcal{F}$ е σ - алгебрата, породена от всички подинтервали на $[0, 1]$. Вероятността P, дефинирана върху $\mathcal{F}$ е мярката (дължината) на съответния елемент от $\mathcal{F}$. Дефинираме случайните величини $X(\omega) = \omega$ и $Y(\omega) = 1 - \omega$ за $\omega \in [0, 1]$. Лесно се вижда, че $P(X \leq x) = P(Y \leq x) = x$ за $0 \leq x \leq 1$. Следователно X и Y имат една и съща ф.р.

$$F(x) = \begin{cases} 0, & x < 0 \\ x, & 0 \leq x < 1 \\ 1, & x \geq 1. \end{cases}$$

Случайна величина с тази функция на разпределение се нарича равномерно разпределена в интервала $(0, 1)$. Означава се $X \in \mathcal{U}(0, 1)$.

Този пример показва, че ф.р. определя цял клас от случайни величини. От вероятностна гледна точка, случайни величини с една и съща ф.р. се предполагат неразличими.

Нека X е една дискретна сл.в. и $x_1, x_2, \ldots, x_n, \ldots$ са стойностите ѝ, подредени по големина $x_1 < x_2 < \ldots < x_n \ldots$. Да означим с

$$p_n = P(\omega : X = x_n)$$

вероятността, с която случайната величина приема съответните стойности. Ф.р. на тази случайна величина приема вида

$$F(x) = P(X \leq x) = \sum_{x_k \leq x} P(X = x_k) = \sum_{x_k \leq x} p_k. \tag{3.6}$$

От (3.6) се проверява, че

$$\lim_{x \uparrow x_n} F(x) = F(x_n - 0) = p_1 + p_2 + \ldots + p_{n-1},$$

и

$$\lim_{x \downarrow x_n} F(x) = F(x_n) = p_1 + p_2 + \ldots + p_n,$$

от където, големината на скока на $F(x)$ в точката x_n е $p_n = F(x_n) - F(x_n - 0)$.

Така, ф.р. на една дискретна сл.в. е напълно определена, ако са известни нейните стойности и съответните вероятности. Това може да се представи във вид на таблица и се нарича **дискретно разпределение.**

X	x_1	$x_2 \dots$	$x_n \dots$
P	p_1	$p_2 \dots$	$p_n \dots$

Тази таблица, където $p_n \geq 0$ и $\sum_{n=1}^{\infty} p_n = 1$, се нарича още *ред на разпределение* на случайната величина.

3.3 Непрекъснати случайни величини

Друг основен клас случайни величини са тези, при които съществува неотрицателна функция $f(t)$, такава, че за всяко $x \in \mathbb{R}^1$

$$F(x) = \int_{-\infty}^{x} f(t)dt. \tag{3.7}$$

Такива случайни величини се наричат **непрекъснати**, а функцията $f(t)$ - плътност на разпределението (**плътност**). От (3.7) се вижда, че функцията $F(x)$ е абсолютно непрекъсната и $F'(x) = f(x)$.

Плътността притежава следните основни свойства

$$f(t) \geq 0, \quad \int_{-\infty}^{\infty} f(t)dt = 1. \tag{3.8}$$

Лесно се вижда и верността на следната формула

$$P(x_1 < X \leq x_2) = \int_{x_1}^{x_2} f(t)dt.$$

3.4 Функции на разпределение и вероятностни мерки в борелово пространство

Видяхме, че всяка функция на разпределение притежава свойствата 1. - 4. Оказва се, че тези свойства са характеризационни, т.е. определят еднозначно функцията на разпределение.

Следната теорема установява връзката между вероятностите в измеримо борелово пространство $(\mathbb{R}^1, \mathcal{B}_1)$ и функциите $F(x)$ със свойствата 1. - 4.

Означаваме $I_x = (-\infty, x]$.

Теорема 3.3 (за съответствие). *Съотношението $P(I_x) = F(x)$ установява взаимно еднозначно съответствие между функциите на разпределение и вероятностите P в $(\mathbb{R}^1, \mathcal{B}_1)$.*

Следствие 3.1 *Всяка реална функция $F(x),\ x \in \mathbb{R}^1$ със свойствата 1. - 4. е функция на разпределение на някаква случайна величина, дефинирана на подходящо вероятностно пространство.*

Доказателство. Нека P е вероятността в $(\mathbb{R}^1, \mathcal{B}_1)$ съгласно теоремата за съответствие. Във вероятностното пространство $(\mathbb{R}^1, \mathcal{B}_1, P)$ разглеждаме сл. в. $X(\omega) = \omega,\ \omega \in R^1$. Тогава както се вижда от следните равенства

$$P(\omega : X(\omega) \leq x) = P(\omega \leq x) = P(I_x) = F(x),$$

$F(x)$ е функция на разпределение на X.

$\square$

3.5 Свойства на случайните величини

1. Ако X_1 и X_2 са случайни величини и $a \in \mathbb{R}$, то $X_1 + X_2,\ aX_i,\ X_1X_2$ и $\frac{X_1}{X_2}$ са също случайни величини.

В множеството от случайни величини приемаме естествената наредба: $X_1 \leq X_2$ ако $X_1(\omega) \leq X_2(\omega)$ за всяко $\omega \in \Omega$.

2. Ако X_1 и X_2 са случайни величини, то $\max(X_1, X_2)$ и $\min(X_1, X_2)$ са случайни величини. Твърдението следва от това, че за всяко $x \in \mathbb{R}$ са в сила твърденията

$$\{\max(X_1, X_2) \leq x\} = \{X_1 \leq x\} \bigcap \{X_2 \leq x\} \in \mathcal{F}$$

и

$$\{\min(X_1, X_2) > x\} = \{X_1 > x\} \bigcap \{X_2 > x\} \in \mathcal{F}.$$

3. Ако $\{X_n\}$ е редица от случайни величини, то $\sup X_n$ и $\inf X_n$ също са случайни величини. Следва от това, че за всяко $x \in \mathbb{R}$

$$\{\sup_{n\geq 1} X_n \leq x\} = \bigcap_{n=1}^{\infty} \{X_n \leq x\} \in \mathcal{F}$$

и

$$\{\inf_{n\geq 1} X_n \leq x\} = \bigcup_{n=1}^{\infty} \{X_n \leq x\} \in \mathcal{F}.$$

От последните равенства следва още, че

$$\limsup X_n = \inf_{n\geq 1} (\sup_{m\geq n} X_m) \quad \text{и} \quad \liminf X_n = \sup_{n\geq 1} (\inf_{m\geq n} X_m)$$

са също случайни величини.

Определение 3.6 *Редицата от случайни величини $\{X_n\}$ клони към случайната величина X, ако за всяко $\omega \in \Omega$ е в сила*

$$\limsup X_n(\omega) = \liminf X_n(\omega) = X(\omega).$$

Тази сходимост се нарича **поточкова сходимост.**

4. За всяка борелова функция $\varphi(x)$ и за всяка случайна величина X, функцията $Y = \varphi(X)$ е също случайна величина (т.е. борелова функция от случайна величина е случайна величина).

Доказателството следва от факта, че за всяко $B \in \mathcal{B}$, е в сила

$$\{\omega : Y(\omega) \in B\} = \{\omega : \varphi(X(\omega)) \in B\} = \{\omega : X(\omega) \in \varphi^{-1}(B)\} \in \mathcal{F},$$

тъй като $\varphi^{-1}(b) = \{x : \varphi(x) \in B\} \in \mathcal{B}$.

От това свойство веднага следва, че X^n, $X^+ = \max(X, 0)$, $X^- = \max(-X, 0)$, $|X| = X^+ + X^-$ също са случайни величини.

Теорема 3.4 *Класът на случайните величини, дефинирани в дадено измеримо пространство $(\Omega, \mathcal{F})$ съвпада с класа на измеримите функции, съдържащ елементарните случайни величини и затворен относно поточкова сходимост.*

Доказателство. Да припомним, че елементарните случайни величини приемат краен брой стойности (Определение 3.5). Нека $X = X(\omega)$ е случайна величина. Означаваме

$$X_n(\omega) = \sum_{k=1-n^2}^{n^2} \frac{k-1}{n} I_{A_k}(\omega) + nI_{B_n}(\omega) - nI_{C_n}(\omega), \tag{3.9}$$

където $A_k = \{\omega : \frac{k-1}{n} < X(\omega) \le \frac{k}{n}\}$, $B_n = \{\omega : X(\omega) > n\}$, $C_n = \{\omega : X(\omega) \le -n\}$.

$X_n(\omega)$ е проста случайна величина, тъй като приема краен брой стойности. Ще покажем, че за всяко $\omega \in \Omega$, $X_n(\omega) \to X(\omega)$ при $n \to \infty$, т.е. че за всяко $\omega \in \Omega$ и всяко $\varepsilon > 0$, съществува $N < \infty$, така, че при $n \ge N$ е в сила $|X_n(\omega) - X(\omega)| < \varepsilon$.

Нека $\omega \in \Omega$ и $\varepsilon > 0$ е произволно. Избираме N така, че $\frac{1}{N} < \varepsilon$. Тогава съществува k_0, такова, че $-N^2 + 1 \le k_0 \le N^2$ и $\frac{k_0-1}{N} < X(\omega) \le \frac{k_0}{N}$.

Съгласно (3.9) това означава, че $X_N(\omega) = \frac{k_0-1}{N}$ и следователно $X(\omega) - X_N(\omega) < \frac{1}{N} < \varepsilon$. Сега за всяко $n \ge N$ съществува k, така, че $1 - n^2 \le k \le n^2$ и $\frac{k-1}{n} \le X(\omega) \le \frac{k}{n}$. Съгласно (3.9), $X_n(\omega) = \frac{k-1}{n}$. Тогава $X(\omega) - X_n(\omega) < \frac{1}{n} < \varepsilon$.

Така показахме, че всяка случайна величина може да се представи като граница на редица от елементарни случайни величини.

□

От (3.9) следва, че ако $X \ge 0$, то $C_n = \emptyset$, $n = 1, 2, \ldots$. Следователно $X_n \uparrow X$, т.е. всяка неотрицателна случайна величина може да се представи като граница на монотонно растяща редица от неотрицателни прости случайни величини.

В общия случай от (3.9) следва, че $|X_n| \le |X|$.

3.6 Редица от независими опити

Много често в практиката се налага изучаването на случайни събития, които се повтарят многократно при едни и същи условия. Най-простата схема на повтарящи се събития е т.н. схема на Бернули. Това е безкрайна редица от независими опити, при които едно събитие се сбъдва с вероятност $p > 0$, или не се сбъдва с вероятност $q = 1 - p$. За улеснение, когато събитието се сбъдне, казваме че е настъпил успех. Ако събитието не се сбъдне, означава че е настъпил неуспех. Вероятността за успех при всеки опит е една и съща. Този тип опити е изследван за пръв път от швейцарския учен Якоб Бернули (1654 - 1705) и носи неговото име. Оказва се, че закономерностите, които се появяват в схемата на Бернули са основни в теория на вероятностите. Те дават насока за изучаване и на по-сложни схеми.

3.7 Често срещани дискретни разпределения

Разпределение на Бернули. Случайната величина X има разпределение на Бернули, ако приема две стойности 0 и 1, с вероятност

$$P(X = 1) = p, \qquad P(X = 0) = 1 - p = q.$$

Биномно разпределение. При постоянен състав на урната, се избират n топки с връщане. Случайната величина X е равна на броя на белите топки в извадката и приема стойности $0, 1, \dots n$. Броят на изходиte при експеримента е $\#(\Omega) = (M + N)^n$. Броят на благоприятните изходи за x бели топки в извадката е $\binom{n}{x} M^x N^{n-x}$. В този случай разпределението е

$$P(X = x) = \binom{n}{x} \left(\frac{M}{M+N}\right)^x \left(\frac{N}{M+N}\right)^{n-x}, \quad x = 0, 1, \dots n.$$

Приемаме $p = \frac{M}{M+N}$ за параметър на това разпределение. Тогава

$$P(X = k) = \binom{n}{x} p^k (1 - p)^{n-k}, \quad k = 0, 1, \dots n$$

и се нарича **биномно разпределение с параметри** $n \geq 1$ **и** $p \in (0,1)$, $(X \sim Bi(n,p))$. Биномно разпределената случайна величина може да се представи като сума на n бернулиеви сл. в. Нека $Y_1, \dots Y_n$ са бернулиеви случайни величини, т.е. за $k = 1, \dots n$

$$Y_k = \begin{cases} 1 \text{ (успех)}, & p \\ 0 \text{ (неуспех)}, & 1-p. \end{cases}$$

Тогава $X = Y_1 + \dots Y_n$ е равна на броя на успехите при n опита в схемата на Бернули.

Геометрично разпределение. Случайната величина X е геометрично разпределена с параметър p, $(X \sim Ge_0(p))$, ако

$$P(X = k) = (1-p)^k p,\ \ k = 0, 1, 2, \dots$$

Случайната величина X е равна на броя на неуспехите до първа поява на успех в редица от независими бернулиеви опити.

Отрицателно биномно разпределение. Случайната величина Z се нарича отрицателно биномно разпределена с параметри r и p, $(Z \sim NB(r,p))$, ако

$$P(X = k) = \binom{k+r-1}{k} p^r (1-p)^k,\ \ k = 0, 1, 2, \dots.$$

Случайната величина Z е равна на броя на неуспехите до появата на r успехи в редица от независими бернулиеви опити и може да се представи като сума от геометрично разпределени случайни величини: $Z = \sum_{k=1}^{r} X_k$, където $X_k \sim Ge_0(p)$.

3.8 Някои непрекъснати вероятностни разпределения

3.8.1 Експоненциално разпределение и отсъствие на памет

Случайната величина X се нарича експоненциално разпределена с параметър $\lambda > 0$, $(X \sim \exp(\lambda))$, ако

$$F(x) = 1 - e^{-\lambda x} \quad \text{и} \quad f(x) = \lambda e^{-\lambda x}, \ x \geq 0.$$

Ако $\lambda = 1$, случайната величина $X \sim \exp(1)$ се нарича стандартна експоненциално разпределена. Нека $X \sim \exp(1)$ и $Y = a + \frac{X}{\lambda}, \quad \lambda > 0, \ -\infty < a < \infty$. Тогава разпределението на Y се определя от функцията на разпределение и плътността

$$F(x) = 1 - e^{-\lambda(x-a)} \quad \text{и} \quad f(x) = \lambda e^{-\lambda(x-a)}, \ x \geq a$$

и се означава $Y \sim \exp(a, \lambda)$.

Предимството на експоненциалното разпределение се състои в наличието на свойството отсъствие на памет.

Случайната величина X притежава свойството **отсъствие на памет**, ако за всеки избор на $t, s \geq 0$, така че $P(X \geq t) > 0$

$$P(X \geq t + s | X \geq t) = P(X \geq s). \tag{3.10}$$

Ako X е продължителност на живот, то (6.5) изразява вероятността, че индивид, доживял възраст t ще доживее поне до възраст $s + t$ е равна на вероятността този индивид да надживее възраст s. Ако функцията на разпределение не е изродена в нулата, то условието (6.10) е еквивалентно на

$$\frac{P(X \geq t + s, X \geq t)}{P(X \geq t)} = P(X \geq s)$$

или

$$P(X \geq t + s) = P(X \geq s)P(X \geq t). \quad (3.11)$$

В равенство (6.11) може да се включи и случая на израждане. Решението се дава от следната теорема ([2]).

Теорема 3.5 *Сред функциите на разпределение съществуват само две решения на уравнение (6.11). $F(x)$ или е изродена в нулата или за дадена константа $\lambda > 0$, $\quad F(x) = 1 - e^{-\lambda x},\ x \geq 0$.*

Така показахме, че експоненциалното разпределение притежава свойството отсъствие на памет и то е единственото непрекъснато разпределение с това свойство.

Свойството отсъствие на памет при експоненциалното разпределение се илюстрира и при понятието хазартна функция.

Нека предположим, че продължителността на живот има експоненциално разпределение. Тогава, съгласно свойството отсъствие на памет следва, че разпределението на останалото време на живот на индивидите на възраст t е същото както при младите индивиди. Така $\lambda(t)$ трябва да е константа. Това се проверява така:

$$\lambda(t) = \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}} = \lambda.$$

Хазартната функция на експоненциалното разпределение е константа.

Нека X и Y са независими, експоненциално разпределени случайни величини с параметри, съответно λ и μ. Тогава разпределението на $Z = \min(X, Y)$ се определя от

$$\begin{aligned} P(Z \leq z) &= 1 - P(Z > z) = 1 - P(X > z, Y > z) = \\ &= 1 - P(X > z)P(Y > z) = 1 - e^{-\lambda z}e^{-\mu z} = 1 - e^{-(\lambda+\mu)z}. \end{aligned}$$

Така, $Z = \min(X, Y)$ е отново експоненциално разпределена с параметър $\lambda+\mu$. Повтаряйки това свойство, може да се покаже, че минималната от краен брой експоненциално разпределени случайни величини е отново експоненциална.

$$P(X \leq Y \quad | \min(X, Y) > z) = \frac{P(X \leq Y, Z > z)}{P(Z > z)} = \frac{P(X \leq Y, X > z, Y > z)}{P(X > z, Y > z)}$$

$$= \frac{P(X \leq Y, X > z)}{P(X > z)P(Y > z)} = \frac{\int_z^\infty \int_x^\infty \lambda e^{-\lambda x} \mu e^{-\mu z} dz dx}{e^{-\lambda z} e^{-\mu z}}$$

$$= \frac{\int_z^\infty \lambda e^{-\lambda x} e^{-\mu x} dx}{e^{-\lambda z} e^{-\mu z}} = \frac{\frac{\lambda}{\lambda+\mu} e^{-\lambda z} e^{-\mu z}}{e^{-\lambda z} e^{-\mu z}} = \frac{\lambda}{\lambda + \mu}.$$

Това означава, че вероятността, че минималната стойност ще се достигне от случайната величина X е пропорционална на параметъра на X и не зависи от $\min(X, Y)$.

3.8.2 Гама разпределение

Нека за $\alpha > 0$ означим $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$, която се нарича Гама функция. Свойствата на Гама функцията, от които се нуждаем са $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$, $\Gamma(n) = (n - 1)!$ за всяко цяло $n \geq 1$ и $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

С помощта на Гама функцията се определя Гама разпределена случайна величина ($X \sim \Gamma(\alpha, \beta)$) с параметри α и β, с плътност на разпределение

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad x > 0.$$

В частност, при $\alpha = 1$, $\Gamma(1, \beta) = exp(\beta)$.

3.8.3 Бета разпределение

Нека за $\alpha > 0$ и $\beta > 0$ означим $B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx$, която се нарича Бета функция. Най-същественото от свойствата на Бета функцията е

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}.$$

Случайната величина X се нарича Бета разпределена ($X \sim B(\alpha, \beta)$) с параметри α и β, ако се задава с плътност на разпределение

$$f(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad x \in (0, 1).$$

3.8.4 Разпределение на Вайбул (Weibull)

Случайната величина X има разпределение на Вайбул с параметри $\beta > 0$ и $\sigma > 0$,, ($X \sim W(\beta, \sigma)$) ако функцията на разпределение е

$$F(x) = 1 - e^{-(\frac{x}{\sigma})^\beta}, \quad x \geq 0.$$

Функцията на плътността е

$$f(x) = \frac{\beta}{\sigma} \left(\frac{x}{\sigma}\right)^{\beta-1} e^{-(\frac{x}{\sigma})^\beta}, \quad x \geq 0.$$

При $\beta = 1, \quad W(1, \sigma) = exp(\sigma)$.

3.8.5 Разпределение на Коши

Случайната величина X има разпределение на Коши с параметри $\mu \in (-\infty, \infty)$ и $\sigma > 0$, ($X \sim C(\mu, \sigma)$), ако функцията на плътността е

$$f(x) = \frac{1}{\pi\sigma} \frac{1}{1 + (\frac{x-\mu}{\sigma})^2}, \quad -\infty < x < \infty.$$

При $\mu = 0$ и $\sigma = 1$, разпределението се нарича стандартно разпределение на Коши и се задава с плътността

$$f(x) = \frac{1}{\pi} \frac{1}{1 + x^2}, \quad -\infty < x < \infty.$$

3.8.6 Разпределение на Парето

Случайната величина X има разпределение на Парето с параметри α и λ, ($X \sim Par(\alpha, \lambda)$), ако функцията на разпределение е

$$F(x) = 1 - \left(\frac{\lambda}{\lambda + x}\right)^{\alpha}, \quad x > 0.$$

Параметрите λ и α са неотрицателни.

Плътността на X се определя от

$$f(x) = \frac{\alpha\lambda^{\alpha}}{(\lambda + x)^{\alpha+1}}, \quad x > 0.$$

Често се налага използването на изместено разпределение на Парето с плътност

$$f_1(x) = f(x - \lambda) = \frac{\alpha\lambda^{\alpha}}{x^{\alpha+1}}, \quad x > \lambda.$$

3.8.7 Лог - нормално разпределение

Случайната величина X има лог - нормално разпределение, ако $Y = \log X$ има нормално разпределение. Ако $Y \sim N(\mu, \sigma)$, то плътността на X е

$$f_X(x) = f_Y(\log x)\frac{1}{x} = \frac{1}{\sigma x\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\log x - \mu}{\sigma}\right)^2}, \quad x > 0,$$

където $\sigma > 0, \quad -\infty < \mu < \infty$. Функцията на разпределение на X се определя от

$$F(x) = \Phi\left(\frac{\log x - \mu}{\sigma}\right), \quad x > 0,$$

където $\Phi(.)$ е стандартната нормална функция на разпределение.

3.8.8 Обратно Гаусово разпределение

Ако X е зададена с плътност на разпределение

$$f(x) = \frac{\mu}{\sqrt{2\pi\beta x^3}} e^{-\frac{(x-\mu)^2}{2\beta x}}, \ x \geq 0, \tag{3.12}$$

където $\mu > 0$, то X е обратно Гаусово разпределена. В специалния случай $\mu = 1$ (3.12) се нарича разпределение на Валд (Wald).

Функцията на разпределение на обратното Гаусово разпределение е

$$F(x) = \Phi\left(\frac{x-\mu}{\sqrt{\beta x}}\right) + e^{\frac{2\mu}{\beta}}\Phi\left(-\frac{x+\mu}{\sqrt{\beta x}}\right), \quad x > 0.$$

Глава 4

Интегриране относно вероятностна мярка

В тази глава се въвежда едно основно понятие във вероятностите, наречено Математическо очакване на случайна величина.

4.1 Елементарни случайни величини

Разглеждаме съвкупността от елементарните сл.в., дефинирани в дадено вероятностно пространство $(\Omega, \mathcal{F}, P)$. Това са сл.в., които приемат краен брой стойности $x_1, x_2, \ldots, x_n$ и имат вида

$$X = \sum_{k=1}^{n} x_k I_{A_k}.$$

В този случай

$$p_k = P(X = x_k) = P(A_k), \quad \sum_{k=1}^{n} p_k = 1.$$

Определение 4.1 *Математическо очакване (очаквана стойност, средна стойност) на случайната величина X се нарича функционала*

$$EX = \sum_{k=1}^{n} x_k p_k. \tag{4.1}$$

Пример 4.1 *За произволно събитие A, съгласно (4.1) се получава $EI_A = P(A)$, т.е. вероятността на едно събитие може да се разглежда като математическото очакване на неговия индикатор.*

Пример 4.2 *Нека $p_k = \frac{1}{n}$. Съгласно (4.1) $EX = \frac{1}{n}(x_1 + \ldots + x_n)$. В този случай математическото очакване е средното аритметично от стойностите на случайната величина.*

Теорема 4.1 *Математическото очакване EX е единственият нормиран, линеен, позитивен и непрекъснат функционал в множеството от елементарни сл.в., за който $EI_A = P(A)$.*

4.2 Интегрируеми случайни величини

Да разгледаме съвкупността от всички неотрицателни сл.в., дефинирани в пространството $(\Omega, \mathcal{F}, P)$. Показахме, че всяка такава сл.в. X може да се представи като граница на редица от неотрицателни елементарни сл.в., т.е. $X = \lim \uparrow X_n$, където $X_n \geq 0$, за всяко n.

Определение 4.2 *За всяка неотрицателна сл.в. X, съотношението*

$$EX = \lim \uparrow EX_n, \tag{4.2}$$

където X_n са елементарни сл.в. и $X_n \uparrow X$, се нарича **математическо очакване.**

Теорема 4.2 *Формула (4.2) дефинира еднозначно продължение на функционала EX от множеството на елементарните сл.в. в множеството на неотрицателните сл.в.*

Доказателство. Ще докажем, че стойността на EX в (4.2) не зависи от избора на редицата. Нека

$$X = \lim \uparrow X_n \leq Y = \lim \uparrow Y_n,$$

където X_n, Y_n са редици от елементарни случайни величини.

Известно е, че ако X_m и Y_n са елементарни случайни величини, то $\min(X_m, Y_n)$ е също елементарна случайна величина и $\min(X_m, Y_n) \leq Y_n$, от където

$$X_m = \lim_{n\to\infty} \uparrow \min(X_m, Y_n) \leq \lim_{n\to\infty} \uparrow Y_n = Y.$$

Тогава от (4.2) се получава

$$EX_m = \lim_{n\to\infty} \uparrow E\min(X_m, Y_n) \leq \lim_{n\to\infty} \uparrow EY_n,$$

от където

$$\lim_{m\to\infty} \uparrow EX_m = EX \leq EY = \lim_{n\to\infty} \uparrow EY_n.$$

Ако $X = \lim \uparrow X_n = \lim \uparrow Y_n$, то

$$EX = \lim \uparrow EX_n = \lim \uparrow EY_n,$$

т.е.стойността на (4.2) не зависи от избора на редицата. □

Свойства на математическото очакване

1. За неотрицателните сл.в. $0 \leq EX \leq \infty$;

2. За положителна константа a, $EaX = aEX$;

3. $E(X + Y) = EX + EY$;

4. Ако $X \leq Y$, то $EX \leq EY$ (монотонност);

5. Ако X_n е растяща редица от неотрицателни сл.в., то съществува $\lim \uparrow X_n = X$, X е неотрицателна и $EX = \lim \uparrow EX_n \leq \infty$ (монотонна непрекъснатост).

Нека за всяка сл.в. X дефинираме

$$X^+ = \max(X, 0) \text{ и } X^- = \max(-X, 0).$$

Ясно е, че X^+ и X^- са неотрицателни и $X = X^+ - X^-$ и $|X| = X^+ + X^-$.

Определение 4.3 *Ако $EX^+ < \infty$ и $EX^- < \infty$, то случайната величина X се нарича* **интегрируема.** *Тогава математическото очакване се определя от*

$$EX = EX^+ - EX^-.$$

Означаваме с $L^1(\Omega, \mathcal{F}, P)$ или просто L^1 множеството от всички интегрируеми случайни величини. Ясно е, че

$$0 \le E|X| = EX^+ + EX^- < \infty.$$

Определение 4.4 *Ако поне едно от числата EX^+ или EX^- е крайно, случайната величина X се нарича* **квазиинтегрируема** *и*

$$EX = EX^+ - EX^-. \tag{4.3}$$

Теорема 4.3 *Дефинираното в (4.3) математическо очакване запазва свойствата линейност и монотонна непрекъснатост върху множеството от квазиинтегрируемите случайни величини (следователно и върху интегрируемите).*

За непрекъснатата случайна величина X, с функция на разпределение $F(x)$ и плътност $f(x)$, математическото очакване се определя от

$$EX = \int_{-\infty}^{\infty} x dF(x) = \int_{-\infty}^{\infty} x f(x) dx,$$

при условие че интегралите са сходящи.

Следващата теорема показва при какви условия може да се извършва граничен преход под знака на интеграла.

Теорема 4.4 *(на Фату - Лебег) Нека $\{X_n\}$ е редица от случайни величини. Тогава при $n \to \infty$*

а) Ако $X_n \leq X \in L^1$, то $\limsup_{n\to\infty} EX_n \leq E(\limsup X_n)$;

б) Ако $X_n \geq Y \in L^1$, то $E(\liminf_{n\to\infty} X_n) \leq \lim_{n\to\infty} E \inf X_n$;

в) Ако $X_n \to X$ и $|X| \leq Z \in L^1$, то $X \in L^1$ и $EX = \lim_{n\to\infty} EX_n$.

4.3 Моменти на случайните величини

Нека X е случайна величина с плътност $f(x)$ и $h(x)$ е или положителна функция, или $E(|h(X)|) < \infty$. Тогава $E(h(X)) = \int h(x)f(x)dx$.

В чатност, за $h(x) = x^k$, очакването $E(X^k)$, $k = 1, 2, \dots$ се нарича $k-$ ти начален момент на случайната величина. Да отбележим, че EX е константа. За функцията $h(x) = (x - EX)^k$, $E[(X - EX)^k]$, $k = 1, 2, \dots$ се нарича $k-$ ти централен момент.

Пространството от случайни величини с краен $k-$ ти момент се означава L^k.

Ако случайната величина $X \in L^2$, то дефинираме

$$Var(X) = E[(X - EX)^2]. \tag{4.4}$$

$Var(X)$ се нарича дисперсия на случайната величина и се означава още DX. Да отбележим, че ако $X \in L^2$, то $X \in L^1$, следователно съществува и математическото очакване. Нека означим $\mu = EX$. Тогава

$$\begin{aligned} E[(X - \mu)^2] &= E(X^2) - 2\mu EX + \mu^2 \\ &= E(X^2) - 2\mu^2 + \mu^2 \\ &= E(X^2) - \mu^2. \end{aligned}$$

$\sigma = \sqrt{Var(X)}$ се нарича стандартно отклонение на случайната величина. Двете характеристики, дисперсия и стандартно отклонение са мерки за разсейване на стойностите на случайната величина.

От определението следват основните свойства на дисперсията:

1. $Var(C) = 0$,
2. $Var(CX) = C^2 Var(X)$,
3. $Var(X \pm Y) = Var(X) + Var(Y)$, когато X и Y са независими.

Теорема 4.5 *(Неравенство на Cauchy - Schwarz) Ако $X, Y \in L^2$, то $XY \in L^1$ и е в сила неравенството*

$$|E(XY)| \leq \sqrt{E(X^2)E(Y^2)}. \tag{4.5}$$

Доказателство. От очевидното неравенство $|XY| \leq \frac{X^2}{2} + \frac{Y^2}{2}$ следва, че ако $X, Y \in L^2$, то $XY \in L^1$. За всяко $x \in \mathbb{R}$ е в сила

$$0 \leq E[(xX+Y)^2] = x^2 E(X^2) + 2xE(XY) + E(Y^2).$$

Дясната част е неотрицателна с дискриминанта равна на $4[(E(XY))^2 - E(X^2)E(Y^2)]$, от където следва

$$E(XY)^2 - E(X^2)E(Y^2) \leq 0,$$

и неравенството (4.4).

□

Пример 4.3 *Нека $X \sim N(\mu, \sigma^2)$. Тогава за математическото очакване се получава*

$$EX = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx.$$

След смяна на променливите $x = y + \mu$, се получава

$$\begin{aligned} EX &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (y+\mu) e^{-\frac{y^2}{2\sigma^2}} dy \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} y e^{-\frac{y^2}{2\sigma^2}} dy + \mu \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2\sigma^2}} dy. \end{aligned}$$

Първият интеграл е равен на нула (нечетна функция в симетрични граници), а вторият е $\mu \int_{-\infty}^{\infty} f(x)dx = \mu$. *Така* $EX = \mu$.

Пример 4.4 *Нека* X *има разпределение на Коши с плътност*

$$f(x) = \frac{1}{\pi}\frac{1}{1+x^2}, \quad -\infty < x < \infty.$$

Тогава

$$\begin{aligned} E(X^+) &= \frac{1}{\pi}\int_0^{\infty} \frac{x}{1+x^2}dx + \frac{1}{\pi}\int_{-\infty}^{0} \frac{0}{1+x^2}dx \\ &\geq \frac{1}{\pi}\int_1^{\infty} \frac{1}{2x}dx = \infty, \end{aligned}$$

което следва от това, че $\frac{x}{1+x^2} \geq 0$ *за* $x \geq 0$ *и* $\frac{x}{1+x^2} \geq \frac{1}{2x}$ *за* $x > 1$. *Аналогично се показва, че* $E(X^-) = \infty$, *от където следва* $E(|X|) = \infty$ *и че* $E(X)$ *не съществува.*

Глава 5

Многомерни случайни величини

Нека случайните величини X и Y приемат стойности в $\mathbb{R}$. Може да се предполага, че векторът (X, Y) приема стойности в $\mathbb{R}^2$.

5.1 Дискретни случайни величини

Нека (X, Y) е двумерен дискретен случаен вектор със стойности (x_i, y_j), $i = 1, 2, \dots n$, $j = 1, 2, \dots m$.

Съвместното разпределение на случайните величини X и Y се определя от

$$f_{X,Y}(x_i, y_j) = P(X = x_i, Y = y_j),$$

където

$$\sum_i \sum_j f_{X,Y}(x_i, y_j) = 1.$$

Съвместната функция на разпределение е

$$F_{X,Y}(x, y) = P(X \leq x, Y \leq y) = \sum_{x_i \leq x} \sum_{y_j \leq y} f_{X,Y}(x_i, y_j).$$

Вероятностите

$$p_1(x_i) = \sum_j f_{X,Y}(x_i, y_j), \quad i = 1, 2, \dots n$$

и

$$p_2(y_j) = \sum_i f_{X,Y}(x_i, y_j), \quad j = 1, 2, \ldots m$$

определят разпределенията на случайните величини X и Y и се наричат маргинални разпределения.

5.2 Непрекъснати случайни величини

Непрекъснатите случайни величини X и Y се задават със съвместна функция на разпределение

$$F_{X,Y}(x, y) = P(X \leq x, Y \leq y),$$

която е непрекъсната.

Ако съществува функция $f_{X,Y}(x, y)$, за която

$$F_{X,Y}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(s, t) dt ds,$$

то $f_{X,Y}(x, y)$ се нарича съвместна плътност на разпределение на X и Y. Плътността е неотрицателна функция на две променливи, за която

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(s, t) dt ds = 1$$

и

$$f_{X,Y}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x, y).$$

Плътността $f_{X,Y}(x, y) dx dy$ изразява вероятността

$$P(x \leq X \leq x + dx, y \leq Y \leq y + dy).$$

5.3 Свойства на $F(x, y)$

Съвместната функция на разпределение притежава свойства, аналогични на едномерните функции на разшпределение.

1. $0 \leq F(x, y) \leq 1$.

2. Ако $x_1 \le x_2$ и $y_1 \le y_2$, то

$$F(x_1, y_1) \le F(x_2, y_1) \le F(x_2, y_2))$$

$$F(x_1, y_1) \le F(x_1, y_2) \le F(x_2, y_2)).$$

3. $\lim_{x\to\infty, y\to\infty} F(x, y) = F(\infty, \infty) = 1.$

4.

$$\lim_{x\to-\infty} F(x, y) = F(-\infty, y) = 0$$

$$\lim_{y\to-\infty} F(x, y) = F(x, -\infty) = 0.$$

5.

$$\lim_{x\to a+} F(x, y) = F(a+, y) = F(a, y)$$

$$\lim_{y\to b+} F(x, y) = F(x, b+) = F(x, b).$$

6.

$$P(x_1 < X \le x_2, Y \le y) = F(x_2, y) - F(x_1, y)$$

$$P(X \le x, y_1 < Y \le y_2) = F(x, y_2) - F(x, y_1)$$

7. Ако $x_1 \le x_2$ и $y_1 \le y_2$, то

$$F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1) \ge 0.$$

Пример 5.1 *(Двумерно разпределение на Парето)*

$$F_{X,Y}(x, y) = 1 - \frac{1}{x} - \frac{1}{y} + \frac{1}{x + y - 1}, \ x, y \ge 1.$$

$$F_{X,Y}(1, 1) = 0, \quad \lim_{x,y\to\infty} F_{X,Y}(x, y) = 1.$$

$$f_{X,Y}(x, y) = \frac{2}{(x + y - 1)^3}.$$

5.4 Маргинални функции на разпределение

Тъй като събитието $\{Y \leq \infty\}$ се сбъдва винаги, то

$$\lim_{y \to \infty} P(X \leq x, Y \leq y) = P(X \leq x, Y \leq \infty) = P(X \leq x),$$

т.е.

$$\lim_{y \to \infty} F(x, y) = F(x, \infty) = F_X(x) \tag{5.1}$$

и

$$\lim_{x \to \infty} F(x, y) = F(\infty, y) = F_Y(x) \tag{5.2}$$

(5.1) и (5.2) се наричат маргинални функции на разпределение.

5.5 Условни разпределения и условно математическо очакване

За условната вероятност установихме

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

при условие, че $P(B) > 0$.

Естествено възниква въпросът как да се дефинира $E(X|B)$, т.е. математическо очакване на случайната величина X, при условие събитието B. Да припомним, че условието B може да се разглежда като ново вероятностно пространство с вероятност $\overline{P}(A) = \frac{P(A)}{P(B)}$ за произволно събитие A. Тогава, ако $P(B) > 0$, математическото очакване на X върху B се определя от

$$E(X|B) = \frac{1}{P(B)} \int_B X dP.$$

Каква ще бъде дефиницията на $E(X|Y)$ - математическото очакване на X, при условие случайната величина Y. Това означава, че ако се е сбъднало елементарното събитие $\omega \in \Omega$, всичко което знаем за ω е стойността $Y(\omega)$.

Ще покажем два подхода за определяне на $E(X|Y)$.

Пример 5.2 *Предполагаме, че е дадено вероятностното пространство $(\Omega, \mathcal{F}, P)$, на което е дефинирана случайната величина Y, т.е.*

$$Y = \sum_{i=1}^{m} a_i I_{A_i},$$

където I_{A_i} е индикаторната функция на A_i. Това означава, че

$$Y = a_i \quad \text{върху} \quad A_i, \quad i = 1, 2, \dots m$$

за различни реални числа $a_1, a_2, \dots a_m$ и непресичащи се събития $A_1, A_2, \dots A_m$, всички с положителни вероятности и $\sum_{i=1}^{m} A_i = \Omega$. Нека X е друга реална случайна величина на Ω.

Така, ако знаем стойностите на $Y(\omega)$, можем да кажем кое от събитията $A_1, A_2, \dots A_m$ съдържа ω. При наличието само на тази информация, най - добрата оценка на X, ще бъде средната стойност на X върху всяко събитие, т.е.

$$E(X|Y) = \frac{1}{P(A_i)} \int_{A_i} X dP, \quad \text{върху} \quad A_i, \quad i = 1, 2, \dots m$$

От този пример се вижда, че

1. $E(X|Y)$ е случайна величина.
2. $E(X|Y)$ е $\mathcal{F}(Y)-$ измерима.
3. $\int_A X dP = \int_A E(X|Y) dP$ за всяко $A \in \mathcal{F}(Y)$.

Приемаме тези свойства за определящи:

Определение 5.1 *Нека Y е случайна величина. Тогава $E(X|Y)$ е някаква $\mathcal{F}(Y)-$ измерима случайна величина, такава, че*

$$\int_A X dP = \int_A E(X|Y) dP \quad \text{за всяко} \quad A \in \mathcal{F}(Y).$$

От дефиницията се вижда, че определяща е не случайната величина Y, а $\sigma-$ алгебрата, породена от Y. Така идваме до следната дефиниция:

Определение 5.2 *Нека $(\Omega, \mathcal{F}, P)$ е вероятностно пространство и нека $\mathcal{V}$ е $\sigma-$ алгебра $\mathcal{V} \subseteq \mathcal{F}$. Ако $X : \Omega \to \mathbb{R}^n$ е една интегрируема случайна величина, определяме $E(X|\mathcal{V})$ като случайна величина на Ω, такава, че*

1. $E(X|\mathcal{V})$ е $\mathcal{V}-$ измерима и

2. $\int_A X dP = \int_A E(X|\mathcal{V}) dP$ за всяко $A \in \mathcal{V}$.

5.6 Функции от случайни величини

Теорема 5.1 *Нека X е непрекъсната случайна величина и нека $\varphi(x)$ е строго растяща функция, дефинирана върху стойностите на X. Определяме случайната величина $Y = \varphi(X)$ и означаваме с F_X и F_Y функциите на разпределение на случайните величини X и Y. Тогава тези функции са свързани със съотношението*

$$F_Y(y) = F_X(\varphi^{-1}(y)).$$

Ако $\varphi(x)$ е строго намаляваща върху множеството от стойности на X, то

$$F_Y(y) = 1 - F_X(\varphi^{-1}(y)).$$

Доказателство. Тъй като $\varphi(x)$ е строго растяща върху стойностите на X, то събитията $\{X \leq \varphi^{-1}(y)\}$ и $\{\varphi(X) \leq y\}$ са еквивалентни. Така

$$F_Y(y) = P(Y \leq y) = P(\varphi(X) \leq y) = P(X \leq \varphi^{-1}(y)) = F_X(\varphi^{-1}(y)).$$

Ако $\varphi(x)$ е строго намаляваща, то

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(\varphi(X) \leq y) = \\ &= P(X > \varphi^{-1}(y)) = 1 - P(X \leq \varphi^{-1}(y)) = 1 - F_X(\varphi^{-1}(y)). \end{aligned}$$

□

Следствие 5.1 *Ако f_X и f_Y са вероятностните плътности на X и Y, и $\varphi(x)$ е строго растяща, то*

$$f_Y(y) = f_X(\varphi^{-1}(y))\frac{d}{dy}\varphi^{-1}(y).$$

Ако $\varphi(x)$ е строго намаляваща, то

$$f_Y(y) = -f_X(\varphi^{-1}(y))\frac{d}{dy}\varphi^{-1}(y).$$

Ако функцията $\varphi(x)$ не е нито растяща, нито намаляваща, има случаи, при които този метод също е приложим.

Пример 5.3 *Например, при $Y = X^2$, функцията е $\varphi(x) = x^2$. Тогава*

$$F_Y(y) = P(Y \leq y)$$

$$= P(-\sqrt{y} < X \leq \sqrt{y}) = P(X \leq \sqrt{y}) - P(X \leq -\sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y}).$$

За плътността се получава

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}\left(F_X(\sqrt{y}) - F_X(-\sqrt{y})\right) = \left(f_X(\sqrt{y}) + f_X(-\sqrt{y})\right)\frac{1}{2\sqrt{y}}.$$

5.7 Сума на случайни величини

Нека X и Y са независими случайни величини. Функцията на разпределение на сумата $Z = X + Y$ се означава

$$F_Z(z) = F_X * F_Y(z)$$

и се нарича **конволюция (композиция)** на F_X и F_Y и се определя от

$$\begin{aligned} F_Z(z) &= \int P(Z \leq z|Y = y)dF_Y(y) \\ &= \int P(X \leq z - y|Y = y)dF_Y(y) \\ &= \int F_{X|Y}(z - y|y)dF_Y(y) = \int F_X(z - y)dF_Y(y). \end{aligned} \tag{5.3}$$

Ако X и Y са непрекъснати случайни величини, плътността на Z се получава чрез диференциране на (5.3) относно z:

$$f_Z(z) = f_X * f_Y(z) = \int f_X(z-y)f_Y(y)dy.$$

Ако X и Y са дискретни случайни величини

$$P(Z=k) = \sum_{i=0}^{k} P(X=k-i)P(Y=i), \quad k = 0,1,2,\dots.$$

В многомерния случай, ако $X_1, X_2, \dots X_n$ са независими случайни величини, то разпределението на сумата $S_n = X_1+X_2+\dots+X_n$ се получава чрез рекурентна формула. Означаваме $S_1 = X_1$ и $S_j = S_{j-1} + X_j, \quad j = 2,3,\dots n$. функцията на разпределение на сумата се означава

$$F_{S_n}(x) = F_{X_1} * F_{X_2} * \dots * F_{X_n}(x)$$

и плътността

$$f_{S_n}(x) = f_{X_1} * f_{X_2} * \dots * f_{X_n}(x).$$

Ако $X_1, X_2, \dots X_n$ са еднакво разпределени с функция на разпределение $F_X(x)$ и плътност $f_X(x)$, то

$$F_{S_n}(x) = F_X^{*n}(x)$$

и

$$f_{S_n}(x) = f_X^{*n}(x).$$

Тук $*n$ означава $n-$ та конволюция на X. В този случай, от (5.3) следва, че трансформациите на $S_n = X_1 + X_2 + \dots + X_n$ са $n-$та степен на съответната трансформация на X. Например за пораждащата функция се получава

$$P_{S_n}(t) = [P_X(t)]^n.$$

5.7.1 Отрицателно биномно разпределение

Нека $X_1, X_2, \dots X_r$ са независими $Ge(p)$ разпределени случайни величини. Случайната величина $N = X_1 + X_2 + \dots X_r$ има функция на вероятностите

$$p_k = \binom{k-1}{r-1} p^r (1-p)^{k-r}, \quad k = r, r+1, \dots$$

и се нарича отрицателно биномно разпределена ($N \sim NB(r, p)$). Интерпретацията на отрицателно биномната случайна величина е брой на опитите до първа поява на r успеха в схемата на Бернули. Пораждащата функция има вида

$$P_N(t) = \left(\frac{pt}{1-(1-p)t}\right)^r, \quad t < 1.$$

Математическото очакване и дисперсията са

$$EN = \frac{r}{p} \quad Var(N) = \frac{r(1-p)}{p^2}.$$

Алтернативно представяне на отрицателно биномното разпределение е чрез случайната величина Y равна на брой на неуспехите преди появата на $r-$ тия успех. Така $Y = N - r$ и функцията на вероятностите е

$$p_k = \binom{r+k-1}{k} p^r (1-p)^k, \quad k = 0, 1, 2, \dots$$

Пораждащата функция има вида

$$P_Y(t) = \left(\frac{p}{1-(1-p)t}\right)^r, \quad t < 1.$$

Математическото очакване и дисперсията са

$$EY = \frac{r(1-p)}{p} \quad Var(Y) = \frac{r(1-p)}{p^2}.$$

5.7.2 Разпределение на Ерланг

Нека $X_1, X_2, \dots X_n$ са независими, експоненциално разпределени случайни величини с параметър λ. Да се намери разпределението на случайната величина $S_n = X_1 + \ldots + X_n$. Известно е, че експоненциалното разпределение е $\Gamma(1, \lambda)$.

Предполагаме, че

$$f_{X_1+\ldots+X_{n-1}}(t) = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-2}}{(n-2)!}.$$

За S_n се получава

$$\begin{aligned} f_{X_1+\ldots+X_n}(t) &= \int_0^\infty f_{X_n}(t-s) f_{X_1+\ldots+X_{n-1}}(s) ds \\ &= \int_0^\infty \lambda e^{-\lambda(t-s)} \lambda e^{-\lambda s} \frac{(\lambda s)^{n-2}}{(n-2)!} ds = \lambda \frac{(\lambda t)^{n-1}}{(n-1)!} e^{-\lambda t}. \end{aligned}$$

Определение 5.3 *Случайна величина с плътност на разпределение*

$$f_X(t) = \lambda \frac{(\lambda t)^{n-1}}{(n-1)!} e^{-\lambda t}, \quad t \geq 0, \tag{5.4}$$

се нарича Ерланг разпределена с параметри n и λ, $X \sim Erl(n, \lambda)$, $n =$ 1, 2, ...

Функцията на разпределение има вида

$$F(t) = 1 - \sum_{j=0}^{k-1} \frac{(\lambda t)^j}{j!} e^{-\lambda t}, \quad t \geq 0.$$

Параметърът λ се нарича локационен параметър, n е параметър на формата.

Математическото очакване и дисперсията са

$$EX = \frac{n}{\lambda} \quad \text{и} \quad Var(X) = \frac{n}{\lambda^2}.$$

Известно е, че (5.4), когато n е реален параметър е плътността на Гама разпределена случайна величина ($X \sim \Gamma(n, \lambda)$.

5.8 Двумерно нормално разпределение

Нека плътността на случайния вектор (X, Y) е

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y(1-\rho^2)^{\frac{1}{2}}} e^{-\frac{Q(x,y)}{2}}, \ -\infty < x, y < \infty,$$

където

$$Q(x,y) = \frac{1}{1-\rho^2}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right].$$

Означаваме $u = \frac{x-\mu_X}{\sigma_X}$ и $v = \frac{y-\mu_Y}{\sigma_Y}$. В термините на u и v, $Q(x,y)$ приема вида

$$\begin{aligned} Q(x,y) &= q(x,y) = \frac{1}{1-\rho^2}(u^2 - 2\rho uv + v^2) \\ &= u^2 + \frac{(v-\rho u)^2}{1-\rho^2} = \left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \frac{(y-\alpha(x))^2}{\sigma_X^2(1-\rho^2)}, \end{aligned}$$

където

$$\alpha(x) = \mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X).$$

Така се получава

$$f(x,y) = h(x)g(x,y),$$

където

$$h(x) = f_X(x) = \frac{1}{\sigma_X\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2}$$

е плътността на нормално разпределена случайна величина ($X \sim N(\mu_X, \sigma_X^2)$). Аналогично, от симетрията на съвместната плътност следва, че $Y \sim N(\mu_Y, \sigma_Y^2)$. За $g(x,y)$ се получава

$$g(x,y) = \frac{1}{\sigma_Y\sqrt{2\pi(1-\rho^2)}} e^{-\frac{1}{2}\frac{(y-\alpha(x))^2}{\sigma_Y^2(1-\rho^2)}},$$

т.е., за всяко $x, g(x,y)$ е плътност на нормално разпределена случайна величина ($N(\alpha(x), \sigma_Y^2(1-\rho^2))$).

Глава 6

Трансформации върху случайни величини

Една възможност за изучаване на случайните величини е чрез трансформиране на функцията на разпределение по подходящ начин. Разглеждаме

$$Eg(X) = \int g(x)dF_X(x), \tag{6.1}$$

където $F_X(x)$ е функция на разпределение на случайната величина X. Ще разгледаме следните възможности за функцията $g(x)$:

а) $g(x) = e^{sx}$;

б) $g(x) = s^x$;

в) $g(x) = e^{isx}$;

г) $g(x) = e^{-sx}$.

Във всеки един от случаите (6.1) е функция на s и се нарича трансформация на $F_X(x)$ (или трансформация на случайната величина X).

Да разгледаме трансформациите поотделно.

а) **Функция, пораждаща моментите (mgf)** се нарича

$$M_X(s) = Ee^{sX} = \int e^{sx}dF_X(x),$$

при условие, че интеграла в дясно е сходящ. Когато интеграла не е сходящ mgf не съществува.

Функцията, пораждаща моментите може да се развие в степенен ред

$$\begin{aligned} M_X(s) &= \int (1 + sx + \tfrac{(sx)^2}{2!} + \ldots) dF_X(x) = \\ &= 1 + \mu'_1 s + \mu'_2 \tfrac{s^2}{2!} + \mu'_3 \tfrac{s^3}{3!} + \ldots, \end{aligned}$$

от където

$$\mu'_k = EX^k = \frac{d^k}{dx^k} M_X(s)|_{s=0} = M_X^{(k)}(0),\ k = 1, 2, \ldots$$

б) **Пораждаща функция** (п.ф.) на случайната величина X се нарича

$$\psi_X(s) = Es^X = \int s^x dF_X(x).$$

Най-често пораждащата функция се прилага върху случайни величини с цели, неотрицателни стойности. В този случай

$$\psi_X(s) = Es^X = \sum_{k=0}^{\infty} s^k p_k,$$

където

$$p_k = P(X = k), \quad k = 0, 1, 2, \ldots.$$

Известно е, че степенният ред с радиус на сходимост r е диференцируем безбройно много пъти в интервала $(-r, r)$. Диференцираме $\psi_X(s)$ за $|s| < 1,\ k$ пъти:

$$\psi^{(k)}(s) = \sum_{n=k}^{\infty} n(n-1)\ldots(n-k+1)P(X = n)s^{n-k}. \tag{6.2}$$

В частност

$$\psi_X^{(k)}(0) = k!P(X = k).$$

Ако предположим, че $\psi^{(k)}(s)$ е сходящ ред за $|s| < 1$, то при $s = 1$ от (6.2) се получава

$$\sum_{n=k}^{\infty} n(n-1)\dots(n-k+1)P(X=n) = E[X(X-1)\dots(X-k+1)]. \quad (6.3)$$

Ако $E[X(X-1)\dots(X-k+1)] < \infty$, то $\psi^{(k)}(s)$ е непрекъсната за $|s| \le 1$. Ако редът в ляво на (**??**) е разходящ, то $\psi^{(k)}(s) \to \infty$ при $s \to 1$ и тогава $E[X(X-1)\dots(X-k+1)] = \infty$.

в) **Характеристична функция** на случайната величина X се нарича

$$\varphi_X(s) = Ee^{isX} = \int e^{isx} dF_X(x).$$

Характеристичната функция винаги съществува. Моментите на случайната величина могат да се получат по начин, подобен на този при mgf.

г) **Лапласова трансформация (LT)** на случайната величина X или на нейното разпределение се нарича

$$LT_X(s) = Ee^{-sX} = \int e^{-sx} dF_X(x), \quad (6.4)$$

Обикновено Лапласовата трансформация се прилага върху случайни величини, приемащи неотрицателни стойности. За такива случайни величини LT винаги съществува за комплексни стойности на s с положителна реална част. За нашите цели е достатъчно да се предположи, че s приема реални стойности.

LT (6.4) на случайната величина се нарича още **Лаплас - Стилтесова трансформация (LST)** на функцията на разпределение $F_X(x)$ и в същност е LT на плътността на разпределение $f_X(x)$, ако тя съществува.

Всяка от тези трансформации определя еднозначно разпределението на случайната величина.

В сила са следните връзки:

$$\psi_X(s) = M_X(\ln s), \quad \varphi_X(s) = M_X(is), \quad LT_X(s) = M_X(-s).$$

Ако X и Y са независими случайни величини, то всяка една трансформация на сумата $X + Y$ е равна на произведение от съответните трансформации на X и Y. Например, в случая на пораждаща функция

$$P_{X+Y}(s) = P_X(s)P_Y(s).$$

6.1 Пораждаща функция

Ще докажем следната теорема.

Теорема 6.1 *Пораждащата функция определя еднозначно функцията на разпределение.*

Доказателство. Нека ψ и φ са две пораждащи функции, такива, че $\psi(s) = \varphi(s)$ за $|s| < h,\ h > 0$. Нека

$$\psi(s) = \sum_{k=0}^{\infty} s^k P(X = k) \quad \text{и} \quad \varphi(s) = \sum_{k=0}^{\infty} s^k P(Y = k).$$

Тогава, за всяко $k = 0, 1, \ldots$

$$P(X = k) = \frac{1}{k!}\psi^{(k)}(0) = \frac{1}{k!}\varphi^{(k)}(0) = P(Y = k).$$

Следователно X и Y имат едно и също разпределение.

$\square$

Като частен случай на (6.3) в сила е следния резултат.

Твърдение 6.1 *Ако* $E(X) < \infty$, *то* $E(X) = \psi'(1)$. *Ако* $E(X^2) < \infty$, *то*

$$Var(X) = \psi"(1) + \psi'(1) - [\psi'(1)]^2. \tag{6.5}$$

В общия случай, ако $E[X(X-1)\ldots(X-k+1)] < \infty$, *то*

$$E[X(X-1)\ldots(X-k+1)] = \psi^{(k)}(1).$$

Доказателство. Достатъчно е да се провери (6.5). В сила е

$$\begin{aligned} Var(X) &= E(X^2) - (E(X))^2 \\ &= E[X(X-1)] + E(X) - (E(X))^2 \\ &= \psi"(1) + \psi'(1) - [\psi'(1)]^2. \end{aligned}$$

□

Твърдение 6.2 *Нека $X_1, X_2, \ldots,$ са независими случайни величини. Тогава, поне за $|s| < 1$,*

$$\psi_{\sum X_i}(s) = \Pi_{i=1}^n \psi_{X_i}(s), \tag{6.6}$$

където $\psi_{\sum X_i}(s)$ е п.ф. на $\sum_{i=1}^n X_i$ и $\psi_{X_i}(s)$ е п.ф. на X_i, $i = 1, 2, \ldots$. В частност, когато X_i са еднакво разпределени

$$\psi_{\sum X_i}(s) = [\psi_{X_1}(s)]^n, \tag{6.7}$$

поне за $|s| \leq 1$.

Доказателство. От независимостта на $X_1, X_2, \ldots$ следва, че

$$\psi_{\sum X_i}(s) = E[s^{\sum X_i}] = E[\Pi_{i=1}^n s^{X_1}] = \Pi_{i=1}^n E(s^{X_i}) = \Pi_{i=1}^n \psi_{X_i}(s).$$

Когато $\psi_{X_1}(s) = \psi_{X_i}(s)$ за всяко $i = 1, 2, \ldots,$ (6.7) следва от (6.6). □

6.1.1 Свойства

Нека $\psi_X(s) = \sum_{k=0}^\infty p_k s^k$, $\psi_Y(s) = \sum_{k=0}^\infty q_k s^k$ и $\psi_Z(s) = \sum_{k=0}^\infty r_k s^k$. Тогава

1. $r_k = \alpha p_k + \beta q_k$, $k = 0, 1, 2, \ldots$ тогава и само тогава, когато

$$\psi_Z(s) = \alpha\psi_X(s) + \beta\psi_Y(s).$$

2. $r_k = \sum_{i=0}^k p_i q_{k-i}$, $k = 0, 1, 2, \ldots$ тогава и само тогава, когато

$$\psi_Z(s) = \psi_X(s)\psi_Y(s).$$

3. Може да се покаже, че ако редът, представящ $\psi_X(s)$ е сходящ за всяко $|s| < R$, то производната $\psi'_X(s)$ е също сходящ ред за всяко $|s| < R$ и

$$s\psi'_X(s) = \sum_{k=0}^{\infty} k p_k s^k.$$

Да разгледаме следния полезен пример. Нека

$$\psi_Y(s) = [\psi_X(s)]^\alpha, \tag{6.8}$$

където α е реално число. Да се изрази редицата $q_k,\ k = 0, 1, 2, \dots$, такава, че $\psi_Y(s) = \sum_{k=0}^{\infty} q_k s^k$, при условие, че $p_0 \neq 0$. В някои случаи, например, когато $\psi_X(s) = e^s$ или при α цяло число, това е по-лесна задача и въпреки това се решава числено. Тук ще дадем една полезна апроксимация. От (6.8) следва, че

$$\psi'_Y(s) = \alpha[\psi_X(s)]^{\alpha-1}\psi'_X(s),$$

което може да се запише по следния начин

$$\psi_X(s)[s\psi'_Y(s)] = \alpha\psi_Y(s)[s\psi'_X(s)].$$

Приравняваме коефициентите пред s^k от двете страни на това равенство. Така, за всяко k е в сила

$$\sum_{i=0}^{k}(k-i)q_{k-i}p_i = \alpha\sum_{i=0}^{k} i q_{k-i}p_i.$$

От това равенство се получава рекурентната формула

$$q_k = \frac{1}{kp_0}\sum_{i=1}^{k}[(\alpha+1)i - k]q_{k-i}p_i, \quad k = 1, 2, \dots,$$

която заедно с

$$q_0 = p_0^\alpha$$

определя вероятностното разпределение на случайната величина Y.

6.1.2 Теорема за непрекъснатост

Преди да дадем теоремата ще разгледаме един пример.

Разглеждаме случайна величина X, която е отрицателно биномно разпределена с параметри (m, p). Означаваме $X \sim NB(m, p)$. Пораждащата функция на тази сл. в. е

$$\psi(s) = p^m(1 - qs)^{-m},$$

където $q = 1 - p$.

Нека λ е произволно положително число. Да разгледаме редицата $\{X_m\}$ от случайни величини $NB(m, 1 - \frac{\lambda}{m})$ разпределени, за $m > \lambda$. Съответната редица от п.ф. има общ член

$$\psi_m(s) = \left(1 - \frac{\lambda}{m}\right)^m \left(1 - \frac{\lambda}{m}s\right)^{-m}, \quad m > \lambda.$$

Оставаме $m \to \infty$. Получава се, че редицата от п.ф. е сходяща за всяко фиксирано $s \in [0, 1)$ и

$$\lim_{m\to\infty} \psi_m(s) = e^{-\lambda}e^{\lambda s} = e^{-\lambda(1-s)}.$$

От друга страна, нека X е Поасоново разпределена с параметър $\lambda > 0$, т.е.

$$P(X = k) = \frac{\lambda^k}{k!}e^{-\lambda}, \quad k = 0, 1, 2, \dots$$

Пораждащата функция на X е

$$\psi(s) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}e^{-\lambda}s^k = e^{-\lambda(1-s)}.$$

Сравнявайки двата резултата можем да кажем, че редицата от п.ф. на случайните величини X_m клони към п.ф. на сл.в. с Поасоново разпределение. Можем ли от тук да твърдим, че редицата от разпределения на X_m клони към Поасоновото разпределение, т.е. в сила ли е

$$\lim_{m\to\infty} F_m(x) = \lim_{m\to\infty} P(X_m \le x) = P(X \le x) = \sum_{k=0}^{x} \frac{\lambda^k}{k!}e^{-\lambda}$$

във всяка точка на непрекъснатост на функцията в дясно. Отговор на този въпрос дава следната теорема.

Теорема 6.2 *Нека $F_n(x), \quad n = 1, 2, \ldots$ е редица от функции на разпределение на целочислени случайни величини:*

$$F_n(x) = \sum_{k \leq x} p_{k,n},$$

където $p_{k,n} \geq 0$ за $n = 1, 2, \ldots$ и $\sum_{k=0}^{\infty} p_{k,n} = 1$. Нека $\psi_n(s)$ са съответните п.ф.:

$$\psi_n(s) = \sum_{k=0}^{\infty} p_{k,n} s^k, \quad 0 \leq s < 1.$$

Ако $F(x)$ е някакво дискретно разпределение

$$F(x) = \sum_{k \leq x} p_k,$$

то необходимото и достатъчно условие за

$$\lim_{n \to \infty} F_n(x) = F(x), \tag{6.9}$$

в точките на непрекъснатост за $F(x)$ е

$$\lim_{n \to \infty} \psi_n(s) = \psi(s), \tag{6.10}$$

за всяко $s \in [0, 1)$, където $\psi(s) = \sum_{k=0}^{\infty} p_k s^k$.

Доказателство. Необходимост: Нека $s \in (0, 1)$ е фиксирано и $\varepsilon > 0$ е произволно. Тогава, за достатъчно големи m, за които $\frac{s^{m+1}}{1-s} < \varepsilon$

$$|\psi_n(s) - \psi(s)| < \sum_{k=0}^{m} |p_{k,n} - p_k| s^k. \tag{6.11}$$

Тъй като (6.9) е в сила по предположение, то за всяко фиксирано k, $p_{k,n} \to p_k$, при $n \to \infty$. От (6.11) следва, че $\lim_{n \to \infty} |\psi_n(s) - \psi(s)| \leq \varepsilon$. Тъй като ε е произволно, следва, че (6.10) е в сила.

Достатъчност: Допускаме, че (6.10) е в сила. Разглеждаме матрицата

$$\begin{pmatrix} p_{11} & p_{12} & p_{13} & \dots \\ p_{21} & p_{22} & p_{23} & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix},$$

съставена от вероятностите $p_{k,n} = F_n(k) - F_n(k-0)$. Тъй като $p_{k,n} \leq 1$ за всяко k и n, то от редицата $p_{11}, p_{12}, p_{13}, \dots$ може да се избере сходяща подредица $p_{1n_1} \to p_1^*$ при $n_1 \to \infty$.

Аналогично, от редицата $\{p_{2n_2}\}$ може да се избере сходяща подредица $p_{2n_2} \to p_2^*$. Но индексите n_2 са измежду стойностите на n_1. Следователно $p_{1n_2} \to p_1^*$. Така, при дадено $\varepsilon > 0$, за достатъчно голямо n_2 са в сила едновременно $|p_{1n_2} - p_1^*| < \varepsilon$ и $|p_{2n_2} - p_2^*| < \varepsilon$. Процедурата продължава неограничено. Така осигуряваме, че за произволно $\varepsilon > 0$ и r, за всяко $k \leq r$ и достатъчно голямо n_r е в сила

$$|p_{kn_r} - p_k^*| < \varepsilon.$$

Да разгледаме редицата от функции

$$\psi_{n_r}(s) = \sum_{k=0}^{\infty} p_{kn_r} s^k$$

и функцията $\psi^*(s) = \sum_{k=0}^{\infty} p_k^* s^k$. Нека $\varepsilon > 0$ и r са произволни. При $s < 1$ е в сила следната оценка

$$|\psi_{n_r}(s) - \psi^*(s)| < \sum_{k=0}^{r} |p_{kn_r} - p_k^*| s^k + \sum_{k=r+1}^{\infty} s^k < \varepsilon r + \frac{s^{r+1}}{1-s}.$$

Оставаме $\varepsilon \to 0$, след това $r \to \infty$. Получаваме, че $\lim_{n_r \to \infty} \psi_{n_r}(s) = \psi^*(s)$. Ако допуснем, че (6.9) не е вярно, то могат да се изберат поне две редици $\{p_{kn_r}\}$ и $\{p_{kn_r}^*\}$ клонящи съответно към $\{p_k^*\}$ и $\{p_k^{**}\}$ като поне за едно k, $p_k^* \neq p_k^{**}$. Но тогава, съответните подредици на $\psi_n(s)$, т.е. ψ_{n_r} и $\psi_{n_r}^*$ ще имат граници $\psi^*(s)$ и $\psi^{**}(s)$, които са различни. Това противоречи на (6.10). От тук следва, че (6.9) е вярно.

$\square$

6.2 Характеристични функции (от учебника)

Теорема 2.

Примери 1, 2 и 3.

Теорема 3.

Глава 7

Гранични теореми

7.1 Сходимост на редици от случайни величини

В курсовете по математически анализ се изучава сходимост на редица от функции $f_n : \mathbb{R} \longrightarrow \mathbb{R}$. Редицата $\{f_n\}$ е сходяща и има за граница функцията f ($\lim_{n\to\infty} f_n = f$), ако за всяко $x \in \mathbb{R}$, $\lim_{n\to\infty} f_n(x) = f(x)$. Тази сходимост се нарича поточкова сходимост. Аналогично се дефинира поточкова сходимост на редица от случайни величини. Редицата от случайни величин $X_n : \Omega \longrightarrow \mathbb{R}$ клони поточково към случайната величина X, ако за всяко $\omega \in \Omega$, $\lim_{n\to\infty} X_n(\omega) = X(\omega)$. Тази дефиниция се оказва не много полезна във вероятностите. Това се вижда от следния пример.

Пример 7.1 *Нека X_n, $n = 1, 2, \dots$ е редица от независими еднакви Бернулиеви случайни величини с разпределение $P(X_n = 1) = p$ и $P(X_n = 0) = 1 - p$. При достатъчно голямо n може да се очаква, че относителната част на единиците ще бъде p, т.е. за всяко $\omega \in \Omega$*

$$\lim_{n\to\infty} \frac{X_1(\omega) + \dots X_n(\omega)}{n} = p.$$

Да разгледаме редицата от нули $\omega_0 = \{0, 0, \ldots\}$. За тази редица

$$\lim_{n\to\infty} \frac{X_1(\omega_0) + \ldots X_n(\omega_0)}{n} = 0.$$

По-общо, нека $A = \{\omega : \text{краен брой } 1\}$. Тогава, за всяко $\omega \in A$

$$\lim_{n\to\infty} \frac{X_1(\omega) + \ldots X_n(\omega)}{n} = 0.$$

Може да се покаже, че $P(A) = 0$. Означаваме $A_n = \{X_n = 1\}$. Редицата $A_1, A_2, \ldots$ е сходяща и има за граница събитието $\overline{A} = \{\text{безбройно много } 1\}$.

За тази редица $\sum_{n=1}^{\infty} P(A_n) = \infty$. Съгласно Лемата на Борел-Кантели $P(\overline{A}) = 1$, от където $P(A) = 0$.

Това, което ще покажем в този случай е, че

$$P\left(\left\{\omega : \lim_{n\to\infty} \frac{X_1(\omega) + \ldots X_n(\omega)}{n} = p\right\}\right) = 1.$$

Тук сходимостта не е за всяко ω, а за почти всяко ω.

7.1.1 Сходимост с вероятност 1 и сходимост по вероятност

Голяма част от резултатите във вероятностите са свързани с някакъв граничен преход. Съществуват няколко вида сходимост. Ние ще се спрем на четири от тях.

Нека $X_1, X_2, \ldots$ е редица от случайни величини, дефинирани върху $(\Omega, \mathcal{F}, P)$.

Определение 7.1 *Редицата X_n се нарича сходяща* **с вероятност 1** *или* **почти сигурно (п.с.)** *към случайната величина X при $n \to \infty$, ако*

$$N = \{\omega : \lim_{n\to\infty} X_n(\omega) \neq X(\omega)\} \quad \longrightarrow \quad P(N) = 0. \qquad (7.1)$$

Означаваме $\lim_{n\to\infty} X_n = X$, *п.с.* ($X_n \stackrel{п.с.}{\longrightarrow} X$ *при* $n \to \infty$).

Множеството N се нарича пренебрежимо.

Определение 7.2 X_n *клони* **по вероятност** *към случайната величина* X *при* $n \to \infty$, *ако за всяко* $\varepsilon > 0$,

$$\lim_{n\to\infty} P(\{\omega : |X_n(\omega) - X(\omega)| > \varepsilon\}) = 0. \tag{7.2}$$

Означаваме $X_n \stackrel{P}{\longrightarrow} X$ *при* $n \to \infty$.

Очевидно е, че от сходимост п.с. следва сходимост по вероятност. Обратното не е вярно. Но ако X_n е монотонно намаляваща или монотонно растяща редица, то от сходимост по вероятност следва и сходимост п.с.

Теорема 7.1 *За монотонна редица от случайни величини, сходимостта почти сигурно съвпада със сходимост по вероятност.*

Доказателство. Нека $X_n \uparrow X$. Тогава $\eta = (X - X_n) \downarrow 0$ и за всяко $\varepsilon > 0$ е в сила

$$\{\eta_{n+1} > \varepsilon\} \subset \{\eta_n > \varepsilon\},$$

от където следва

$$\lim_{n\to\infty} \{\eta_n > \varepsilon\} = \bigcap_{n=1}^{\infty} \{\eta_n > \varepsilon\}.$$

и

$$P(\lim_{n\to\infty} \{\eta_n > \varepsilon\}) = \lim_{n\to\infty} P(\{\eta_n > \varepsilon\}) = 0.$$

□

Сходимостта по вероятност се определя от поведението на числовата редица $P(|X_n - X| > \varepsilon)$. Може ли по аналогичен начин да се характеризира и сходимостта с вероятност 1. Означаваме случайната величина $\eta_n = \sup_{k\geq n} |X_n - X|$, която може да бъде и несобствена.

Теорема 7.2 $X_n \xrightarrow{n.c.} X$ *тогава и само тогава, когато за всяко* $\varepsilon > 0$

$$\lim_{n\to\infty} P(\sup_{k\geq n} |X_k - X| > \varepsilon) = 0. \tag{7.3}$$

Доказателство. Ясно е, че $X_n \longrightarrow X$ почти сигурно тогава и само тогава, когато $\eta_n \longrightarrow 0$ почти сигурно. Редицата η_n е монотонно намаляваща и $\eta_n \xrightarrow{P} 0$. Следователно $\eta_n \longrightarrow 0$ почти сигурно.

□

Оказва се, че ако $P(|X_n - X| > \varepsilon)$ клони към нула достатъчно бързо, от сходимост по вероятност следва сходимост почти сигурно.

Теорема 7.3 *Ако за всяко* $\varepsilon > 0$ *редът*

$$\sum_{n=1}^{\infty} P(|X_n - X| > \varepsilon) \tag{7.4}$$

е сходящ, то $X_n \longrightarrow X$ *почти сигурно.*

Доказателство. Използваме свойството полуадитивност на вероятностните мерки и получаваме

$$P(\sup_{k\geq n} |X_k - X| > \varepsilon) = P(\bigcup_{k=n}^{\infty} \{|X_k - X| > \varepsilon\}) \leq \sum_{k=n}^{\infty} P(|X_n - X| > \varepsilon) \longrightarrow 0.$$

при $n \to \infty$. Неравенството следва от (7.4), а твърдението от (7.3).

□

Обратното твърдение не е вярно.

Следствие 7.1 *Нека* $X_n \xrightarrow{P} X$. *Тогава съществува подредица* n_k, *такава че* $\lim_{k\to\infty} X_{n_k} = X \quad n.c.$

Доказателство. Достатъчно е да се вземат такива n_k, за които $P(|X_{n_k} - X| > \varepsilon) \leq \frac{1}{k^2}$ и да се приложи Теорема 7.3.

□

Забележка 7.1 *При сходимост почти сигурно, разглеждаме всяко $\omega \in \Omega$ и проверяваме сходимостта на числовата редица $X_n(\omega)$ към реалното число $X(\omega)$ при $n \to \infty$. Сходимост п.с. е в сила, ако множеството от ω, за които съществува сходимост, има вероятност 1.*

Пример 7.2 *Нека $X_n \in \Gamma(n, \frac{1}{n})$. Да се покаже, че $X_n \xrightarrow{P} 1$ при $n \to \infty$.*

Първо да отбележим,че $EX_n = 1$ и $VarX_n = \frac{1}{n}$. Съгласно неравенството на Чебишев, за всяко $\varepsilon > 0$ е в сила

$$P(|X_n - 1| > \varepsilon) \leq \frac{1}{n\varepsilon^2} \to 0 \text{ при } n \to \infty.$$

Пример 7.3 *Нека $X_1, X_2, \ldots$ е редица от независими, еднакво $U(0,1)$ разпределени случайни величини и*

$$Y_n = \min\{X_1, \ldots X_n\}.$$

Когато се появи стойност X_n, която е по - малка от предходните, Y_n намалява. Следователно редицата от стойности на Y_n, $n = 1, 2, \ldots$ не може да нараства. За произволно $\varepsilon > 0$, от независимостта на $\{X_n\}$ следва

$$P(|Y_n| \geq \varepsilon) = P(X_1 \geq \varepsilon, \ldots X_n \geq \varepsilon) = \prod_{i=1}^{n} P(X_i \geq \varepsilon) = (1-\varepsilon)^n,$$

от където

$$\lim_{n \to \infty} P(|Y_n - 0| \geq \varepsilon) = 0,$$

т.е. редицата $Y_n \xrightarrow{P} 0$.

7.1.2 L^p - сходимост

Определение 7.3 *X_n клони към случайната величина X* **в смисъл на момент от ред** *p, $p > 0$ ако $E|X_n|^p < \infty$, $E|X|^p < \infty$ и*

$$\lim_{n\to\infty} E|X_n - X|^p = 0. \tag{7.5}$$

Означаваме $X_n \xrightarrow{L_p} X$ при $n \to \infty$ и наричаме още L^p - сходимост.

Сходимост в смисъл на момент от ред p се нарича още **сходимост в пространството** L^p.

Особена роля играе L^p - сходимост при $p = 2$. Нарича се средноквадратична сходимост. Нека $p = 1$. От това, че $||x| - |y|| \le |x - y|$ следва

$$|E(X_n - X)| \le E|X_n - X|$$

и

$$|E|X_n| - E|X|| \le E|X_n - X|,$$

от където $X_n \xrightarrow{L_1} X$. От това следва, че $EX_n \to EX$ и $E|X_n| \to E|X|$.

По аналогичен начин, ако $X_n \xrightarrow{L^p} X$, $p \in (1, \infty)$, то $E|X_n|^p \to E|X|^p$.

Теорема 7.4 *Ако X_n е редица от случайни величини, то*

а) $(X_n \overset{п.с.}{\to} X) \longrightarrow (X_n \overset{P}{\to} X)$;

б) $(X_n \overset{L^p}{\to} X) \longrightarrow (X_n \overset{P}{\to} X)$

Доказателство. а) Известно е, че ако $\omega \in \{|X_n - X| > \varepsilon\}$, то $\omega \in \{\sup_{k\ge n} |X_k - X| > \varepsilon\}$. Следователно

$$P(\{|X_n - X| > \varepsilon\}) \le P(\{\sup_{k\ge n} |X_k - X| > \varepsilon\}),$$

от където следва а). б) следва от неравенството

$$P(|X_n - X| > \varepsilon) \le \frac{E|X_n - X|^p}{\varepsilon^p}.$$

$\square$

Пример 7.4 *Нека* $\nu_n \in Bi(n,p)$. *Тогава*

$$E\left(\frac{\nu_n}{n} - p\right)^2 = \frac{E(\nu_n - np)^2}{n^2} = \frac{D\nu_n}{n^2} = \frac{npq}{n^2} = \frac{pq}{n} \to 0$$

при $n \to \infty$, *т.е.* $\frac{\nu_n}{n} \xrightarrow{L_2} p$.

7.1.3 Слаба сходимост

Нека $C(F_X) = \{x : F_X(x)$ е непрекъсната в $x\}$ е множеството от точки, където F_X е непрекъсната.

Определение 7.4 *Редицата* X_n *клони* **по разпределение** *към случайната величина* X *при* $n \to \infty$, *ако за всяко* $x \in C(F_X)$

$$F_{X_n}(x) \to F_X(x), \quad n \to \infty. \tag{7.6}$$

Означаваме $X_n \xrightarrow{d} X$ при $n \to \infty$.

Содимостта по разпределение се нарича още **слаба сходимост.**

Теорема 7.5 *Редицата* $\{X_n\}$ *е сходяща по разпределение към случайната величина* X, *тогава и само тогава, когато за всяка ограничена и непрекъсната функция* $f(x),\ x \in R^+$ *е в сила*

$$\lim_{n\to\infty} Ef(X_n) = Ef(X).$$

Доказателство. Твърдението следва от дефиницията и от изразите

$$Ef(X_n) = \int_{-\infty}^{\infty} f(x)dF_n(x) \quad Ef(X) = \int_{-\infty}^{\infty} f(x)dF_(x)$$

□

Теорема 7.6 *Нека* $X, X_1, X_2, \ldots$ *са случайни величини на* $(\Omega, \mathcal{F}, P)$. *Ако* $X_n \xrightarrow{P} X$ *то* $X_n \xrightarrow{d} X$.

Доказателство. Нека F_{X_n} и F_X са функциите на разпределение на X_n и X и нека $a < b$ са точки от $C(F_X)$. Тогава

$$\{X < a\} = \{X < a, X_n < b\} \cup \{X < a, X_n \geq b\} \subset \{X_n < b\} \cup \{X < a, X_n \geq b\},$$

от където

$$F_X(a) \leq F_{X_n}(b) + P(X < a, X_n \geq b). \tag{7.7}$$

От това, че $X_n \xrightarrow{P} X$,

$$\{X < a, X_n \geq b\} \subset \{|X_n - X| \geq b - a\}$$

и

$$\lim_{n\to\infty} P(X < a, X_n \geq b) = 0.$$

От (7.7) следва, че

$$F_X(a) \leq \liminf_{n\to\infty} F_{X_n}(b). \tag{7.8}$$

Аналогично, ако поставим $c < d$ на мястото на a и b

$$F_{X_n}(c) \leq F_X(d) + P(X_n < c, X \geq d).$$

При $n \to \infty$, отново втория член клони към нула и

$$\limsup_{n\to\infty} F_{X_n}(c) \leq F_X(d). \tag{7.9}$$

От (7.8) и (7.9) следва, че при $a < b \leq c < d$

$$F_X(a) \leq \liminf_{n\to\infty} F_{X_n}(b) \leq \limsup_{n\to\infty} F_{X_n}(b) \leq \limsup_{n\to\infty} F_{X_n}(c) \leq F_X(d). \tag{7.10}$$

Оставаме $a \uparrow b$ и $d \downarrow c$, където $b = c$ е точка от $C(F_X)$. От (7.10) следва

$$\lim_{n\to\infty} F_{X_n}(b) = F_X(b).$$

Тъй като тези точки на непрекъснатост на F_X са навсякъде гъсто в $\mathbb{R}$, то $X_n \xrightarrow{d} X$.

Ако $X = C$ п.с. тогава за всяко $\varepsilon > 0$, за което $C \pm \varepsilon$ са точки от $C(F_X)$

$$\begin{aligned} P(|X_n - C| \geq \varepsilon) &= P(X_n \geq C + \varepsilon) + P(X_n \leq C - \varepsilon) \\ &= 1 - F_{X_n}(C + \varepsilon) + F_{X_n}(C - \varepsilon) \to 0, \quad n \to \infty \end{aligned}$$

и

$$F_{X_n}(x) \to F_X(x) = \begin{cases} 0, & x < C \\ 1, & x \geq C. \end{cases}$$

От тук следва, че $X_n \xrightarrow{d} C$.

□

Пример 7.5 *Нека $X_1, X_2, \ldots$ са независими, еднакво разпределени случайни величини с плътност на разпределение*

$$f(x) = \alpha x^{-\alpha-1}, \ x > 1, \ \alpha > 0. \tag{7.11}$$

Означаваме $Y_n = n^{-1/\alpha} \max_{1 \leq k \leq n} X_k, \ n \geq 1$. Да се покаже, че Y_n е сходяща по разпределение при $n \to \infty$ и да се намери граничното разпределение.

7.2 Закон за големите числа

Нека $X_1, X_2, \ldots$ е редица от независими, еднакво разпределение случайни величини със средна стойност μ и дисперсия σ^2 и $S_n = X_1 + \ldots + X_n$. Граничните теореми се отнасят основно за S_n. От независимостта следва

$$Var(S_n) = Var(X_1) + \ldots + Var(X_n) = n\sigma^2.$$

Да разгледаме $\overline{X}_n = \dfrac{S_n}{n}$.

$$E\overline{X}_n = \mu \quad \text{и} \quad Var(\overline{X}_n) = \frac{\sigma^2}{n}.$$

От тук следва, че случайните величини $Z_n = \dfrac{S_n - n\mu}{\sigma\sqrt{n}}$ имат средна стойност и дисперсия

$$EZ_n = 0 \quad \text{и} \quad Var(Z_n) = 1.$$

7.2.1 Неравенства

Теорема 7.7 *Нека $h : \mathbb{R} \to [0, \infty)$ е неотрицателна функция и X е случайна величина с реални стойности. Тогава, за всяко $\varepsilon > 0$ е в сила*

$$P(\omega : h(X(\omega)) \geq \varepsilon) \leq \frac{E(h(X))}{\varepsilon}.$$

Доказателство. Тъй като X е случайна величина, то $Y = h(X)$ също е случайна величина. Означаваме

$$A = \{\omega : h(X(\omega)) \geq \varepsilon\} = \{h(X) \geq \epsilon\}.$$

От това, че $h(X) \geq \varepsilon I_A$ следва

$$Eh(X) \geq E(\varepsilon I_A) = \varepsilon E(I_A) = \varepsilon P(A).$$

□

Следствие 7.2 (Неравенство на Марков)

$$P(|X| \geq \varepsilon) \leq \frac{E(|X|}{\varepsilon}.$$

Доказателство. Следва от Теорема 7.7 за функцията $h(x) = |x|$.

□

Нека X е реална случайна величина, за която $X^2 \in \mathcal{L}^1$.

Следствие 7.3 (Неравенства на Чебишев) *Ако* $X^2 \in \mathcal{L}^1$, *то*

а) $P(|X| \geq \varepsilon) \leq \frac{E(X^2)}{\varepsilon^2}, \quad \varepsilon > 0,$

б) $P(|X - EX| \geq \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}, \quad \varepsilon > 0,$

Доказателство. а) следва от Теорема 7.7 за функцията $h(x) = x^2$

$$P(|X| \geq \varepsilon) = P(h(X) \geq \varepsilon^2) \leq \frac{E(X^2)}{\varepsilon^2}.$$

б) Означаваме $Y = |X - EX|$. Тогава

$$P(|X - EX| \geq \varepsilon) = P(Y \geq \varepsilon) = P(Y^2 \geq \varepsilon^2) \leq \frac{EY^2}{\varepsilon^2} = \frac{Var(X)}{\varepsilon^2}.$$

□

7.3 Закон за големите числа на Чебишев

Теорема 7.8 *Нека* $X_1, \ldots, X_n, \ldots$ *са независими случайни величини с крайна дисперсия, ограничена от една и съща константа, т.е.* $Var(X_i) \leq C, \quad i = 1, 2, \ldots$. *Тогава, за всяко* $\varepsilon > 0$

$$\lim_{n \to \infty} P\left(\left|\overline{X}_n - \frac{1}{n}\sum_{k=1}^{n} EX_k\right| < \varepsilon\right) = 1. \tag{7.12}$$

Доказателство. При условието за независимост

$$Var(\overline{X}_n) = \frac{1}{n^2}\sum_{k=1}^{n} Var(X_k) \leq \frac{C}{n}.$$

Съгласно неравенството на Чебишев, за всяко $\varepsilon > 0$

$$P\left(\left|\overline{X}_n - \frac{1}{n}\sum_{k=1}^{n} EX_k\right| \geq \varepsilon\right) \leq \frac{Var(\overline{X}_n)}{\varepsilon^2} = \frac{Var(S_n)}{n^2\varepsilon^2} \leq \frac{C}{n\varepsilon^2},$$

което клони към 0 при $n \to \infty$.

□

Забележка 7.2 *Условието (7.12) означава, че редицата* $\overline{X}_n, \ \ n = 1.2.\ldots$ *клони по вероятност към* $\frac{1}{n}\sum_{k=1}^{n} EX_k$. *В този случай говорим за* **слаб закон за големите числа**. *Когато е налице сходимост почти сигурно имаме* **силен закон за големите числа**.

Тук ще отбележим важни частни случаи на теоремата на Чебишев.

Теорема 7.9 *(на Бернули) Нека p е вероятността за успех в безкрайна редица от Бернулиеви опити и X е случайната величина равна на броя на успехите при n опита. Тогава за всяко* $\varepsilon > 0$,

$$\lim_{n\to\infty} P\left(\left|\frac{X}{n} - p\right| < \varepsilon\right) = 1. \tag{7.13}$$

Доказателство. Въвеждаме случайните величини $X_k, \quad k = 1, 2, \ldots,$ равни на броя на успехите при $k-$ тия опит, т.е.

$$X_k = \begin{cases} 1, & p \\ 0, & 1-p. \end{cases}$$

Тогава $X = X_1 + \ldots + X_n$. Тъй като

$$EX_k = p \ \text{ и } \ Var(X_k) = p(1-p) \le \frac{1}{4},$$

то теоремата на Бернули е частен случай на теоремата на Чебишев.

□

Теорема 7.10 *(на Поасон) Нека* p_k *е вероятността за успех при* $k-$ *тия опит в безкрайна редица от независими опити и X е случайната величина равна на броя на успехите при n опита. Тогава за всяко* $\varepsilon > 0$,

$$\lim_{n\to\infty} P\left(\left|\frac{X}{n} - \frac{p_1 + p_2 + \ldots + p_n}{n}\right| < \varepsilon\right) = 1. \tag{7.14}$$

Доказателство. Въвеждаме случайните величини $X_k, \quad k = 1, 2, \ldots,$ равни на броя на успехите при $k-$ тия опит, т.е.

$$X_k = \begin{cases} 1, & p_k \\ \\ 0, & 1 - p_k. \end{cases}$$

Тогава $X = X_1 + \ldots + X_n$. Тъй като

$$EX_k = p_k \quad \text{и} \quad Var(X_k) = p_k(1 - p_k) \leq \frac{1}{4},$$

то и теоремата на Поасон е частен случай на теоремата на Чебишев.

□

Третият частен случай ще формулираме така. Ако редицата от независими случайни величини $X_k, \quad k = 1, 2, \ldots,$ е такава че $EX_k = a, \quad Var(X_k) \leq C, \quad k = 1, 2, \ldots,$ където a и C са константи, то за всяко $\varepsilon > 0$,

$$\lim_{n \to \infty} P\left(\left|\overline{X}_n - a\right| < \varepsilon\right) = 1.$$

Този частен случай дава основание средното аритметично $\overline{X}_n$ да се приема за оценка на математическото очакване и в ролята на локационен параметър на разпределението играе съществена роля. Ние ще се ограничим с формулиране на теоремата на Марков.

Теорема 7.11 *(на Марков) Нека редицата от случайни величини $X_k, \quad k = 1, 2, \ldots,$ е такава че при $n \to \infty$,*

$$\lim_{n \to \infty} \frac{1}{n^2} Var(\sum_{k=1}^{n} X_k) = 0, \tag{7.15}$$

то за всяко $\varepsilon > 0$,

$$\lim_{n \to \infty} P\left(\left|\overline{X}_n - \frac{1}{n}\sum_{k=1}^{n} EX_k\right| < \varepsilon\right) = 1. \tag{7.16}$$

Доказателство. Следва от неравенството на Чебишев.

□

Ясно е, че теоремата на Чебишев е частен случай от теоремата на Марков.

Вече отбелязахме, че закона за големите числа е един от основните в теория на вероятностите. Теоремата на Бернули е публикувана през 1713 година след неговата смърт. В началото на 19 век Поасон доказва аналогична теорема при по-широки условия. В 1866 година П.Л. Чебишев открива метода изложен тук. По-късно А.А. Марков забелязва, че по този начин може да се получи по-общ резултат. Теоремата на Марков дава достатъчно условие, без да се налага независимост на случайните величини. През 1923 година А.Я Хинчин показва, че ако случайните величини са не само независими, но и еднакво разпределени, то съществуването на математическото очакване EX_k е достатъчно условие за закона за големите числа. Едва през 1926 година А.Н.Колмогоров получава необходими и достатъчни условия за това независимите случайни величини $X_1, X_2, \ldots$ да се подчиняват н закона за големите числа.

Теорема 7.12 *За редицата от случайни величини* $X_k, \quad k = 1, 2, \ldots,$ *необходимото и достатъчно условие за*

$$\lim_{n\to\infty} P\left(\left|\overline{X}_n - \frac{1}{n}\sum_{k=1}^{n} EX_k\right| < \varepsilon\right) = 1, \tag{7.17}$$

при произволно $\varepsilon > 0$, *е*

$$\lim_{n\to\infty} E\frac{(\sum_{k=1}^{n}(X_k - EX_k))^2}{n^2 + (\sum_{k=1}^{n}(X_k - EX_k))^2} = 0. \tag{7.18}$$

Доказателство. Следва от неравенството на Чебишев.

□

7.4 Централна гранична теорема

Теорема 7.13 *Нека $X_1, \ldots X_n,\quad n \geq 1$ са независими еднакво разпределени случайни величини с функция на разпределение F и крайни средна стойност μ и дисперсия σ^2. Означаваме $S_n = X_1 + \ldots + X_n$ и $\overline{X} = \frac{S_n}{n}$. Тогава при $n \to \infty$*

а) $P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right) \longrightarrow \Phi(x), \quad x \in \mathbb{R}$

б) $P\left(\frac{\sqrt{n}(\overline{X} - \mu)}{\sigma} \leq x\right) \longrightarrow \Phi(x), \quad x \in \mathbb{R}$

Доказателството ще направим с помощта на следната теорема за непрекъснатост на функцията пораждаща моментите (MGF).

Лема 7.1 *(за непрекъснатост) Нека $X_1, X_2, \ldots$ е редица от случайни величини и $M_n(t)$ е функцията пораждаща моментите на X_n в някакъв отворен интервал $(-a, a)$, съдържащ нулата. Ако $M_n(t) \to e^{\frac{t^2}{2}}$ за всяко $t \in (-a, a)$ при $n \to \infty$, то $P(X_n \leq x) \to \Phi(x)$ за всяко $x \in \mathbb{R}$.*

Забележка 7.3 *В доказатеството се използват:*

1) За редица от числа a_n, такава че $a_n \to 0$ при $n \to \infty$, пишем $a_n = o(1)$.

2) Ако a_n и b_n са две редици от числа и $\dfrac{a_n}{b_n} \to 0$ при $n \to \infty$ пишем $a_n = o(b_n)$.

Например $a_n = o(\frac{1}{n})$ означава, че $a_n \to 0,\ n \to \infty$ и че $na_n \to 0$ при $n \to \infty$.

Доказателство на теоремата. а) Може да се предположи, че $\mu = 0$ и $\sigma = 1$. Знаем вече, че случайните величини $Y_i = \frac{X_i - \mu}{\sigma}$ имат средна стойност $EY_i = 0$ и дисперсия $Var(Y_i) = 1$ и $Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{\sum_{i=1}^n Y_i}{\sqrt{n}},\ \ n \geq 1$.

Така условието а) е в сила тогава и само тогава когато

$$P\left(\frac{\sum_{i=1}^n Y_i}{\sqrt{n}} \le x\right) \longrightarrow \Phi(x).$$

Следователно за $\mu = 0$ и $\sigma = 1$ трябва да се покаже, че

$$P\left(\frac{S_n}{\sqrt{n}} \le x\right) \longrightarrow \Phi(x), \quad n \to \infty.$$

От независимостта на X_i за MGF на $Z_n = \frac{S_n}{\sqrt{n}}$ се получава

$$\begin{aligned} M_{Z_n}(t) &= Ee^{tZ_n} = E\left[e^{t\frac{S_n}{\sqrt{n}}}\right] = E\left[e^{t\sum_{i=1}^n \frac{X_i}{\sqrt{n}}}\right] \\ &= \prod_{i=1}^n E\left[e^{t\frac{X_i}{\sqrt{n}}}\right] = \left(Ee^{t\frac{X}{\sqrt{n}}}\right)^n = \left[M_X(\tfrac{t}{\sqrt{n}})\right]^n. \end{aligned}$$

Логаритмуваме и развиваме $M_X(\frac{t}{\sqrt{n}})$ в ред на Тейлор около $t = 0$.

$$\begin{aligned} \log M_{Z_n}(t) &= n \log M_X(\tfrac{t}{\sqrt{n}}) \\ &= n \log[1 + \tfrac{t}{\sqrt{n}} M_X'(0) + \tfrac{t^2}{2n} M_X''(0) + o(n^{-1})] \\ &= n[\tfrac{t}{\sqrt{n}} M_X'(0) + \tfrac{t^2}{2n} M_X''(0) - \tfrac{t^2}{2n}(M_X'(0))^2 + o(n^{-1})]. \end{aligned}$$

Последното равенство следва от това че $\log(1+x) \approx x - \frac{x^2}{2}$ при $x = 0$. Приели сме че $\mu = EX = M_X'(0) = 0$, от където

$$\log M_{Z_n}(t) = \frac{t^2}{2} M_X''(0) + o(1) = \frac{t^2}{2} + o(1).$$

Следователно при $n \to \infty$

$$M_{Z_n}(t) \longrightarrow e^{\frac{t^2}{2}}.$$

Това означава, че

$$P(Z_n \le x) \longrightarrow \Phi(x), \quad x \in \mathbb{R}.$$

б) следва от а)

□

Забележка 7.4 *Твърденията на Теорема 7.13 означават, че при голямо n*

$$S_n \sim N(n\mu, n\sigma^2) \quad \text{и} \quad \overline{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$

7.5 Нормална апроксимация на биномното разпределение

Теорема 7.14 *Нека $X \sim Bi(n,p)$ случайна величина. Тогава за фиксирано p и $x \in \mathbb{R}$*

$$P\left(\frac{X - np}{\sqrt{np(1-p)}} \leq x\right) \longrightarrow \Phi(x), \quad n \to \infty.$$

Доказателство. Биномно разпределената случайна величина X съвпада със $S_n = X_1 + \ldots + X_n$, където $X_i \sim Bi(1,p)$ и $\mu = p, \quad \sigma^2 = p(1-p)$. □

Следствие 7.4 *(Интегрална теорема на Моавър - Лаплас). Нека $X \sim Bi(n,p)$ случайна величина. Тогава*

$$P(X \leq k) = P\left(\frac{X - np}{\sqrt{np(1-p)}} \leq \frac{k - np}{\sqrt{np(1-p)}}\right) \approx \Phi\left(\frac{k - np}{\sqrt{np(1-p)}}\right).$$

Често се налага да знаем вероятността за фиксирана стойност на случайната величина. Това е възможно чрез $P(X = k) = P(X \leq k) - P(X \leq k-1)$. Тази теорема се нарича локална теорема на Моавър - Лаплас.

Теорема 7.15 *(Локална теорема на Моавър - Лаплас). Нека $X \sim Bi(n,p)$ случайна величина. Тогава за фиксирано p и $k = 0, 1, \dots n$*

$$\begin{aligned} P(X = k) &= P\left(\frac{X - np}{\sqrt{np(1-p)}} = \frac{k - np}{\sqrt{np(1-p)}}\right) \\ &\sim \frac{1}{\sqrt{np(1-p)}}\varphi\left(\frac{k - np}{\sqrt{np(1-p)}}\right) = \frac{1}{\sqrt{np(1-p)}} e^{\frac{(k-np)^2}{2np(1-p)}}. \end{aligned}$$

Библиография

[1] Димитров Б. и Янев Н. (2007). *Вероятности и статистика,* СОФТЕХ, София.

[2] Galambos J. and Kotz S. (1978). *Characterizations of Probability Distributions*, Springer - Verlag. 3.8.1

[3] Johnson N.L., Kotz S. and Kemp A.W. (1992). *Univariate Discrete Distributions*, 2nd ed., Wiley Series in Probability and Mathematical Statistics.

[4] Ross S.M. (1998). A First Course in Probability, fifth edition, Prentice Hall, New Jersey.

[5] Ross S.M. (2007). Introduction to Probability Models, 9th edition, Elsevier Inc.