Введение в теорию вероятностей

Лектор: проф. Булинский Александр Вадимович

15 декабря 2025 г.

Содержание

1 Лекция 1

1.1 Алгебра и сигма-алгебра

Определение. Множество Ω называется множеством элементарных исходов. Множество $A \in 2^{\Omega}$ называется событием.

Определение. Множество $\mathcal{A} \in 2^{\Omega}$ такое, что $\mathcal{A} \neq \varnothing$ называется алгеброй, если

1. $\Omega \in \mathcal{A}$
2. $A \in \mathcal{A} \Rightarrow \bar{A} = \Omega \setminus A \in \mathcal{A}$
3. $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$

Утверждение. (Следствия из определения алгебры)

1. $\varnothing \in \mathcal{A}$, так как $\Omega \in \mathcal{A} \Rightarrow \bar{\Omega} = \varnothing \in \mathcal{A}$
2. $A_1, \ldots, A_n \in \mathcal{A} \Rightarrow \bigcup\limits_{i=1}^{n} A_i \in \mathcal{A}$
3. $A \cap B \in \mathcal{A}$, если $A, B \in \mathcal{A}$, так как $A \cap B = \overline{\overline{A} \cup \overline{B}}$
4. $A_1, \ldots, A_n \in \mathcal{A} \Rightarrow \bigcap\limits_{i=1}^{n} A_i \in \mathcal{A}$
5. $A \setminus B \in \mathcal{A}$, так как $A \setminus B = A \cap \bar{B}$

Определение. Множество $\mathcal{F} \in 2^{\Omega}$ такое, что $\mathcal{F} \neq \varnothing$ называется σ-алгеброй, если

1. $\Omega \in \mathcal{F}$
2. $A \in \mathcal{F} \Rightarrow \bar{A} \in \mathcal{F}$
3. $\forall i \in \mathbb{N} : A_i \in \mathcal{F} \Rightarrow \bigcup\limits_{n=1}^{\infty} A_i \in \mathcal{F}$

Замечание. $\mathcal{F}$ - σ-алгебра $\Rightarrow \mathcal{F}$ - алгебра.

Пример. (Алгебры не являющейся σ-алгеброй)
Возьмем $\Omega = \mathbb{R}$. Скажем, что множество $A \in \mathcal{A}$, если для некоторого n оно является конечным объединеним полуинтрвалов $(a_i, b_i]$

$$A = \bigcup_{i=1}^{n} (a_i, b_i]$$

где $-\infty \leq a_1 \leq b_1 \leq a_2 \leq b_2 \leq \ldots \leq a_n \leq b_n \leq +\infty$.

Покажем, что $\mathcal{A}$ является алгеброй:

1. $(-\infty, +\infty] = \mathbb{R} \in \mathcal{A}$

2. Дополнение к полуинтервалу — объединение двух полуинтервалов.

3. По определению

Покажем, что $\mathcal{A}$ не является σ-алгеброй. Рассмотрим последовательность множеств из $\mathcal{A}$: $A_1 = (0, \frac{1}{2}],\ A_2 = (\frac{1}{2}, \frac{3}{4}],\ A_3 = (\frac{3}{4}, \frac{7}{8}], \ldots, A_n = (1 - \frac{1}{2^{n-1}}, 1 - \frac{1}{2^n}], \ldots$

$$\bigcup_{i=1}^{\infty} A_i = (0, 1) \notin \mathcal{A}$$

Замечание. Наименьшая по включению σ-алгебра, содержащая M, обозначается $\sigma\{M\} = \bigcap\limits_{\alpha} g_\alpha$, где g_α - σ-алгебра, содержащая все элементы M.

1.2 Вероятностная мера

Определение. Мерой на системе множеств U называется функция
$\mu : U \to [0, +\infty]$ такая, что $\forall i \neq j : A_i \cap A_j = \varnothing$:

$$\bigcup_{n=1}^{\infty} A_n \in U$$

выполнено свойство счетной аддитивности:

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$$

Замечание. Если U — σ-алгебра, то условие

$$\bigcup_{n=1}^{\infty} A_n \in U$$

можно упустить.

Пример. (Мера Дирака)
Пусть $B \subset S$

$$\delta_x(B) = \begin{cases} 1,\ x \in B, \\ 0,\ x \notin B \end{cases}$$

Упражнение: доказать, что $\delta_x(.)$ является мерой на 2^S

Определение. Мера P на пространстве $(\Omega, \mathcal{F})$ такая, что $P(\Omega) = 1$ называется вероятностью.

1.3 Дискретные вероятностные пространства

Определение. Пусть $\Omega = \{\omega_n\}_{n \in J}$ не более чем счетно, $\mathcal{F} = 2^\Omega$, причем

$$p_n = P(\{\omega_n\}) \geq 0, \; \sum_{n \in J} p_n = 1$$

Пусть $A \subset \Omega$, определим вероятность так:

$$P(A) = \sum_{k: \omega_k \in A}^{n} p_k$$

такое вероятностное пространство называется дискретным.

Упражнение. Доказать, что определенное выше P является вероятностью.

Определение. (Классическое определение вероятности)
Пусть $|\Omega| = N < \infty$ и положим $p_k = P(\{\omega_k\}) = \frac{1}{N}$. Тогда

$$P(A) = \sum_{k: \omega_k \in A}^{n} p_k = \frac{|A|}{N} = \frac{|A|}{|\Omega|}$$

1.4 Пи-системы и лямбда-системы

Определение. Система M подмножеств множества S называется π-системой, если $A, B \in M \Rightarrow A \cap B \in M$

Определение. Сисмтема M подмножеств множества S называется λ-системой, если

1. $S \in M$
2. $A, B \in M$ и $A \subset B \Rightarrow B \setminus A \in M$
3. $A_1, A_2, \dots \in M$ и $A_n \nearrow A$, то $A \in M$.
   ($A_n \nearrow A \Leftrightarrow A_n \subset A_{n+1}, \; \forall n \in \mathbb{N}$ и $A = \bigcup\limits_{n=1}^{\infty} A_n$)

Теорема. Система $\mathcal{F}$ подмножеств S является σ-алгеброй $\Leftrightarrow$ $\mathcal{F}$ одновременно π-система и λ-система.

Доказательство.

($\Rightarrow$): По следствию из определения алгебры: $A, B \in \mathcal{F} \Rightarrow A \cap B \in \mathcal{F}$, значит, $\mathcal{F}$ является π-системой. Теперь проверим условия λ-системы:

1. $S \in \mathcal{F}$ выполнено по проеделению алгебры.

2. $A, B \in \mathcal{F},\ A \subset B$, причем $B \setminus A = B \cap \bar{A} \Rightarrow B \setminus A \in \mathcal{F}$.

3. $A_1, A_2, \dots \in \mathcal{F},\ A = \bigcup\limits_{n=1}^{\infty} A_n \Rightarrow$ по свойству σ-алгебры $A \in \mathcal{F}$.

($\Leftarrow$): Проверим определению σ-алгебры:

1. $S \in \mathcal{F}$ выполнено по первому свойству λ-системы.
2. $S \in \mathcal{F},\ A \subset S \Rightarrow S \setminus A \in \mathcal{F}$ выполнено по второму свойству λ-системы.
3. Пусть $B_1, B_2, \dots \in \mathcal{F},\ A_m := \bigcup\limits_{n=1}^{m} B_n$ при этом

$$A_m = \bigcup_{n=1}^{m} B_n = \overline{\left(\bigcap_{n=1}^{m} \overline{B}_n\right)} \in \mathcal{F} \Rightarrow A_m \nearrow \bigcup_{n=1}^{\infty} B_n \Rightarrow \bigcup_{n=1}^{\infty} B_n \in \mathcal{F}$$

$\square$

Теорема. Пусть M — π-система, D — λ-система и $M \subset D$. Тогда

$$\sigma\{M\} = \lambda\{M\} \subset D$$

где $\lambda\{M\}$ — наименьшая λ-система, содержащая M.

2 Лекция 2

2.1 Простейшие свойства конечно-аддитивной вероятностной меры на алгебре

Теорема. Пусть P - конечно-аддитивная вероятностная мера на алгебре $\mathcal{A}$, $P(\Omega) = 1,\ P(A \cup B) = P(A) + P(B),\ A \cap B = \varnothing$. Пусть $A, B \in \mathcal{A}$, тогда

1. $A \subset B \Rightarrow P(B \setminus A) = P(B) - P(A)$
2. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
3. Субаддитивность:
$$P\left(\bigcup_{k=1}^{n} A_k\right) \leq \sum_{k=1}^{n} P(A_k)$$
4. Если P - вероятностная мера на σ-алгебре $\mathcal{F}$, то
$$P\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k=1}^{\infty} P(A_k)$$

Доказательство.

1. $B = A \cup (B \setminus A)$ и $A \cap (B \setminus A) = \varnothing \Rightarrow P(B) = P(A) + P(B \setminus A)$

2. $A \cup B = A \cup (B \setminus A)$ и $A \cap (B \setminus A) = \varnothing \Rightarrow P(A \cup B) = P(A) + P(B \setminus A)$,
   $P(B \setminus A) = P(B) - P(A \cap B) \Rightarrow P(A \cup B) = P(A) + P(B) - P(A \cap B)$

3. По индукции. База $n = 2$:
   $P(A \cup B) = P(A) + P(B) - P(A \cup B) \leq P(A) + P(B)$, так как $P(A \cup B) \geq 0$
   Пусть верно для $n - 1$, тогда

$$P\left(\bigcup_{k=1}^{n} A_k\right) = P\left(\bigcup_{k=1}^{n-1} A_k \cup A_n\right) \leq \sum_{k=1}^{n-1} P(A_k) + P(A_n) = \sum_{k=1}^{n} P(A_k)$$

4. *будет позже*

$\square$

2.2 Непрерывность вероятностной меры

Определение. Конечная неотрицательная функция μ, заданная на алгебре $\mathcal{A}$, называется непрерывной в $\varnothing$, если $\forall A_n$:

$$A_n \downarrow \varnothing \ (A_{n+1} \subset A_n, \ \forall n \in \mathbb{N} : \bigcap_{n=1}^{\infty} A_n = \varnothing) \Rightarrow \mu(A_n) \to 0$$

Определение. Конечная неотрицательная функция μ на алгебре $\mathcal{A}$ называется

1. непрерывной сверху на $A \in \mathcal{A}$, если $\forall A_n : A_n \downarrow A \Rightarrow \mu(A_n) \to \mu(A)$.

$$A_n \downarrow A \Leftrightarrow A_{n+1} \subset A_n, \ \forall n \in \mathbb{N} : \bigcap_{n=1}^{\infty} A_n = A$$

2. непрерывной снизу на $A \in \mathcal{A}$, если $\forall A_n : A_n \uparrow A \Rightarrow \mu(A_n) \to \mu(A)$.

$$A_n \uparrow A \Leftrightarrow A_n \subset A_{n+1}, \ \forall n \in \mathbb{N} : \bigcup_{n=1}^{\infty} A_n = A$$

Лемма. Пусть μ - конечная неотрицательная конечно-аддитивная функция на алгебре $\mathcal{A}$, причем μ непрерывна в $\varnothing$. Тогда μ непрерывно сверху и снизу на любом $A \in \mathcal{A}$.

Доказательство. Докажем непрерывность сверху. Рассмотрим последовательность $A_n \downarrow A \Rightarrow A_n \setminus A \downarrow \varnothing \Rightarrow \mu(A_n \setminus A) = \mu(A_n) - \mu(A) \to 0$. Аналогично, рассмотрим $A_n \uparrow A \Rightarrow A \setminus A_n \downarrow \varnothing \Rightarrow \mu(A \setminus A_n) = \mu(A) - \mu(A_n) \to 0$. $\square$

Теорема. (Критерий счетной-аддитивности)
Пусть μ - конечная неотрицательная функция на алгебре $\mathcal{A}$. Тогда μ является счетно-аддитивной на $\mathcal{A}$ тогда и только тогда, когда

1. μ является конечно-аддитивной.

2. μ непрерывна в $\varnothing$.

Доказательство.

($\Rightarrow$): Пусть μ - счетно-аддитивная на $\mathcal{A}$. Рассмотрим $A_n \downarrow \varnothing,\ n \to \infty$ и введем $B_n = A_n \setminus A_{n+1},\ n \in \mathbb{N}$. Эти слои не пересекаются и $\bigcup\limits_{n=N}^{\infty} B_n = A_N$. Применим счетную аддитивность:

$$\mu(A_N) = \sum_{n=N}^{\infty} \mu(B_n) < \infty$$

Этот ряд сходится, значит последовательность (остаточных рядов)

$$S_N = \sum_{n=N}^{\infty} \mu(B_n) \to 0,\ N \to \infty$$

($\Leftarrow$): Рассмотрим $A \in \mathcal{A}$, причем

$$A = \bigcup_{k=1}^{\infty} A_k$$

где $A_k \in \mathcal{A}$ и $A_i \neq A_j$ при $i \neq j$. Введем C_n:

$$C_n = \bigcup_{k=n}^{\infty} A_n, \quad C_n \downarrow \varnothing$$

Заметим, что

$$A = \bigcup_{k=1}^{\infty} A_k = A_1 \cup \dots \cup A_{n-1} \cup C_n$$

причем $A_1 \cup \dots \cup A_{n-1} \in \mathcal{A}$, поскольку $\mathcal{A}$ — алгебра и $A \in \mathcal{A}$ по условию. Тогда $C_n \in \mathcal{A}$.

$$\mu(A) = \mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \mu(A_1) + \dots + \mu(A_{n-1}) + \mu(C_n) = \sum_{k=1}^{n-1} A_k + \mu(C_n)$$

Поскольку $C_n \downarrow \varnothing$, то $\mu(C_n) \to 0$ при $n \to \infty$. Тогда, после предельного перехода, получим

$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} A_k$$

$\square$

Теорема. Пусть P, Q - меры на $(\Omega, \mathcal{F})$ и $P = Q$ на алгебре $\mathcal{A}$. Тогда $P = Q$ на алгебре $\sigma\{\mathcal{A}\}$.

Доказательство. Сведем к теореме из прошлой лекции. Алгебра $\mathcal{A}$ в частности является π-системой, рассмотрим

$$D = \{B \in \mathcal{F} : P(B) = Q(B)\}$$

$\mathcal{A} \subset D$. Проверим, что D является λ-системой:

1. $P(\Omega) = Q(\Omega) = 1 \Rightarrow \Omega \in D$
2. $P(B \setminus A) = P(B) - P(A),\ Q(B \setminus A) = Q(B) - Q(A)$, причем $P(A) = Q(A),\ P(B) = Q(B) \Rightarrow P(B \setminus A) = Q(B \setminus A) \Rightarrow B \setminus A \in D$.
3. $A_n \in D,\ A_n \uparrow A = \bigcup\limits_{n=1}^{\infty} A_n$. По свойству непрерывности $P(A_n) \to P(A)$, $Q(A_n) \to Q(A) \Rightarrow A \in D$.

Значит $\sigma\{\mathcal{A}\} \subset D$, что и является утверждением теоремы. $\square$

2.3 Условные вероятности

Определение. Пусть $(\Omega, \mathcal{F}, P)$ - вероятностное пространство и $A, B \in \mathcal{F}$, $P(B) \neq 0$. Тогда вероятностью события A при условии B называется

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

Определение. (Условная вероятность в классическом определении)
Пусть $(\Omega, \mathcal{F}, P)$ вероятностное пространство, $|\Omega| = N < \infty,\ P(\{\omega_k\}) = \frac{1}{N}$ и $P(B) = \frac{|B|}{N}$. Тогда

$$P(A|B) = \frac{|A \cap B|}{|B|} = \frac{\frac{1}{N} \cdot |A \cap B|}{\frac{1}{N} \cdot |B|} = \frac{P(A \cap B)}{P(B)}$$

Пример. Три раза бросается правильная монетка. Рассмотрим события:
A - при первом броске выпал герб, B - при трех бросаниях выпало два герба.

$$\Omega = \{(0,0,0),\ (1,0,0),\ (0,1,0),\ (0,0,1),\ (0,1,1),\ (1,0,1),\ (1,1,0),\ (1,1,1)\}$$

$$A = \{(1,0,0),\ (1,1,0),\ (1,0,1),\ (1,1,1)\}$$

$$B = \{(1,1,0),\ (1,0,1),\ (0,1,1)\}$$

$$A \cap B = \{(1,1,0),\ (1,0,1)\}$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{2}{8}}{\frac{3}{8}} = \frac{2}{3}$$

Теорема. (Формула полной вероятности)
Пусть для всех $k : B_k \in \mathcal{F},\ P(B_k) > 0$ и пространство разбивается в их объединение:

$$\Omega = B_1 \sqcup B_2 \sqcup B_3 \sqcup \ldots$$

Тогда определены $P(A|B_k)$, причем для $A \in \mathcal{F}$ верна формула

$$P(A) = \sum_k P(A|B_k)P(B_k)$$

Доказательство.

$$P(A) = P\left(\bigcup_k (A \cap B_k)\right) = \sum_k P(A \cap B_k),\ (A \cap B_k) \cap (A \cap B_m) = \varnothing,\ k \neq m$$

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

отсюда

$$P(A) = \sum_k P(A|B_k)P(B_k)$$

$\square$

Следствие. (Формула Байеса)
Пусть $P(A) \neq 0$, тогда верна формула

$$P(B_n|A) = \frac{P(A|B_n)P(B_n)}{\sum\limits_k P(A|B_k)P(B_k)}$$

Доказательство. По определению условной вероятности и формуле полной вероятности, получим

$$P(B_n|A) = \frac{P(B_n \cap A)}{P(A)} = \frac{P(A|B_n)P(B_n)}{\sum\limits_k P(A|B_k)P(B_k)}$$

$\square$

3 Лекция 3

3.1 Независимые события

Определение. Если $P(A|B) = P(A)$, то при $P(B) \neq 0$ выполнено

$$P(A \cap B) = P(A)P(B)$$

В этом случае события A и B называются независимыми.

Пример. В колоде 36 карт. Выбираем одну карту из колоды. Рассмотрим события: A - вытянули карту масти треф, B - вытянули туз.

$$P(A) = \frac{1}{4},\ P(B) = \frac{4}{36} = \frac{1}{9},\ P(A \cap B) = \frac{1}{36}$$

$$P(A)P(B) = \frac{1}{4} \cdot \frac{1}{9} = \frac{1}{36} = P(A \cap B)$$

значит события независимы.

Определение. Пусть $A_1, \dots, A_n$ - события. Они называются независимыми в совокупности, если $\forall i_1 < i_2 < \dots < i_n \leq n$:

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2}) \dots P(A_{i_k})$$

Определение. Пусть $A_1, \dots, A_n$ - события. Они называются попарно независимыми, если $\forall i, j \in \{1, \dots, n\},\ i \neq j$:

$$P(A_i \cap A_j) = P(A_i)P(A_j)$$

Пример. (Попарной независимости событий недостаточно для их независимости в совокупности)
Рассмотрим $(\Omega, \mathcal{F}, P)$ в рамках классического определения:

$$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\},\ A_1 = \{\omega_1, \omega_2\},\ A_2 = \{\omega_1, \omega_3\},\ A_3 = \{\omega_1, \omega_4\},\ A_i \cap A_j = \{\omega_1\}$$

Тогда

$$P(A_i \cap A_j) = P(\{\omega_1\}) = \frac{1}{4}$$

но с другой стороны

$$P(A_1 \cap A_2 \cap A_3) = \frac{1}{4} \neq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}$$

Определение. Система $\{A_t,\ t \in T\}$ состоит из независимых событий, если для любого конечного $F \subset T,\ F = \{t_1, \dots, t_n\}$, события $A_{t_1}, \dots, A_{t_n}$ независимы в совокупности.

Лемма. Пусть $A_1, \dots, A_n$ - независимые события. Рассмотрим $B_1, \dots, B_n$ такие, что $B_i = A_i$ или $B_i = \overline{A_i}$. Тогда $B_1, \dots, B_n$ - независимые события.

Доказательство. Достаточно рассмотреть случай $B_j = \overline{B_j},\ B_i = A_i$, $\forall i \neq j$. Возьмем $I = \{i_1, \dots, i_k\},\ 1 \leq i_1 < \dots < i_k \leq n$ и проверим, что

$$P(B_{i_1} \cap \dots \cap B_{i_k}) = P(B_{i_1}) \dots P(B_{i_k}) \qquad (*)$$

Рассмотрим случаи:

1. $j \notin I$. Тогда $B_i = A_i$ и $(*)$ выполнено.

2. $j \in I \Rightarrow \exists\ m : j = i_m$. Тогда

$$P(B_{i_1} \cap \dots \cap B_{i_k}) = P(A_{i_1} \cap \dots \cap \overline{A_{i_m}} \cap \dots \cap A_{i_k}) \overset{(1)}{=}$$
$$\overset{(1)}{=} P((A_{i_1} \cap \dots \cap A_{i_{m-1}} \cap A_{i_{m+i}} \cap A_{i_k}) \setminus (A_{i_1} \cap \dots \cap A_{i_{m-1}} \cap A_{i_m} \cap A_{i_{m+1}} \cap \dots \cap A_{i_k})) \overset{(2)}{=}$$
$$\overset{(2)}{=} P(A_{i_1} \cap \dots \cap A_{i_{m-1}} \cap A_{i_{m+i}} \cap A_{i_k}) - P(A_{i_1} \cap \dots \cap A_{i_k}) =$$
$$= P(A_{i_1}) \dots P(A_{i_{m-1}}) P(A_{i_{m+1}}) \dots P(A_{i_k}) - P(A_{i_1}) \dots P(A_{i_k}) \overset{(3)}{=}$$
$$\overset{(3)}{=} P(A_{i_1}) \dots P(A_{i_{m-1}}) P(A_{i_{m+1}}) \dots P(A_{i_k})(1 - P(A_{i_m})) =$$
$$= P(B_{i_1}) \dots P(B_{i_m}) \dots P(B_{i_k})$$

(1): $A \cap \overline{C} = A \setminus (A \cap C)$
(2): $A \subset B \Rightarrow P(B \setminus A) = P(B) - P(A)$
(3): Выносим общий множитель за скобку.

□

Теорема. Пусть $\varphi(n)$ - функция Эйлера, p_i - i-е простое число. Тогда

$$\varphi(n) = n \cdot \prod_{i=1}^{m} \left(1 - \frac{1}{p_i}\right)$$

Доказательство. Введем $\Omega = \{1, \dots n\},\ \mathcal{F} = 2^{\Omega},\ P(\{i\}) = \frac{1}{n},\ i = 1, \dots, n$. Проведем следующий эксперимент: из чисел $1, \dots, n$ наугад выбирается число. Рассмотрим событие A - выбрано число, взаимно простое с n. Тогда

$$P(A) = \frac{\varphi(n)}{n}$$

Рассмотрим события A_i - выбраное число делится на p_i. Отсюда

$$A_i = \{p_i,\ 2p_i,\ \dots,\ \frac{n}{p_i} \cdot p_i\} \Rightarrow P(A_i) = \frac{\frac{n}{p_i}}{n} = \frac{1}{p_i}$$

Для любых $1 \le i_1 < \dots < i_k \le n$:

$$A_{i_1} \cap \dots \cap A_{i_k} = \{p_{i_1} \dots p_{i_k},\ 2p_{i_1} \dots p_{i_k},\ \dots,\ \frac{n}{p_{i_1} \dots p_{i_k}} \cdot p_{i_1} \dots p_{i_k}\}$$

$$P(A_{i_1} \cap \dots \cap A_{i_k}) = \frac{\frac{n}{p_{i_1} \dots p_{i_k}}}{n} = \frac{1}{p_{i_1} \dots p_{i_k}} = P(A_{i_1}) \dots P(A_{i_k})$$

Значит $A_{i_1}, \dots, A_{i_k}$ независимы $\Rightarrow$ по лемме $\overline{A_{i_1}}, \dots, \overline{A_{i_k}}$ независимы. Тогда

$$P(A) = P(\overline{A_{P_1}} \cap \overline{A_{p_2}} \cap \dots \cap \overline{A_{p_m}}) = P(\overline{A_{p_1}})P(\overline{A_{p_2}}) \dots P(\overline{A_{p_n}}) =$$
$$= (1 - P(A_{p_1}))(1 - P(A_{p_2})) \dots (1 - P(A_{p_m})) = \prod_{i=1}^{m} \left(1 - \frac{1}{p_i}\right)$$

значит

$$\prod_{i=1}^{m} \left(1 - \frac{1}{p_i}\right) = \frac{\varphi(n)}{n} \;\Rightarrow\; \varphi(n) = n \cdot \prod_{i=1}^{m} \left(1 - \frac{1}{p_i}\right)$$

$\square$

3.2 Независимые системы событий

Определение. Системы событий $\mathcal{G}_1, \dots, \mathcal{G}_n$ называются независимыми, если $\forall 1 \le i_1 < \dots < i_k \le n$ и $\forall A_i \in \mathcal{G}_i : A_{i_1}, \dots, A_{i_k}$ независимы.

Определение. Пусть все $\mathcal{G}_k$ содержат Ω (например $\mathcal{G}_k$ - алгебры), то они независимы, если $\forall A_i \in \mathcal{G}_i$:

$$P(A_1 \cap \dots \cap A_n) = P(A_1) \dots P(A_n)$$

Утверждение. Если все $\mathcal{G}_k$ содержат Ω, то эти определения эквивалентны.

Доказательство.

1. Первое определение влечет второе, поскольку в качестве $i_1, \dots, i_k$ можно взять $1, \dots, n$.

2. Докажем, что второе определение влечет первое: Рассмотрим произвольный набор событий $A_{i_1}, \dots, A_{i_k}$ и определим события $B_1, \dots, B_n$:

$$B_m = \begin{cases} A_m, \;\; m \in \{i_1, \dots, i_k\}, \\ \Omega, \;\; m \notin \{i_1, \dots, i_k\}. \end{cases}$$

   Тогда

$$B_1 \cap \dots \cap B_n = A_{i_1} \cap \dots \cap A_{i_k} \cap (\Omega \cap \dots \cap \Omega) = A_{i_1} \cap \dots \cap A_{i_k}$$

   значит

$$P(A_{i_1} \cap \dots \cap A_{i_k}) = P(B_1 \cap \dots \cap B_n) =$$
$$= P(A_{i_1}) \dots P(A_{i_k}) \cdot (1 \dots 1) = P(A_{i_1}) \dots P(A_{i_k})$$

$\square$

Теорема. Пусть π-системы $M_1, \dots, M_n$ (подмножества Ω из $\mathcal{F}$) независимы. Тогда независимы $\sigma\{M_1\}, \dots, \sigma\{M_n\}$.

Доказательство. Рассмотрим все события B_1 такие, что выполнено:

$$P(B_1 \cap B_2 \cap \dots \cap B_n) = P(B_1)P(B_2)\dots P(B_n) \qquad (*)$$

для проиизвольных $B_2 \in M_2, \ \dots, \ B_n \in M_n$. Назовем такие B_1 системой D_1. Покажем, что D_1 является λ-системой:

1. $(\Omega \in D_1)$

   $$P(\Omega \cap B_2 \cap \dots \cap B_n) = P(B_2 \cap B_n) = \\ = P(B_2)\dots P(B_n) = P(\Omega)P(B_2)\dots P(B_n)$$

   Значит $\Omega \in D_1$.

2. $(A_1, A_2 \in D_1, \ A_1 \subset A_2 \Rightarrow A_1 \setminus A_2 \in D_1)$
   Пусть B_i, B_j - события, для которых выполнено $(*)$ и $B_j \subset B_i$

   $$P((B_i \setminus B_j) \cap (B_2 \cap \dots \cap B_n)) \overset{(1)}{=} \\ \overset{(1)}{=} P((B_i \cap B_2 \cap \dots \cap B_n) \setminus (B_j \cap B_2 \cap \dots \cap B_n)) = \\ = P(B_i \cap B_2 \cap \dots \cap B_n) - P(B_j \cap B_2 \cap \dots \cap B_n) = \\ = P(B_i)\dots P(B_n) - P(B_j)\dots P(B_n) = \\ = P(B_2)\dots P(B_n)(P(B_i) - P(B_j)) = \\ = P(B_i \setminus B_j)P(B_2)\dots P(B_n)$$

   (1): $(B \setminus A) \cap C = (B \cap C) \setminus (A \cap C)$

3. $(A_1, A_2 \dots \in D, \ A_i \subset A_{i+1}$ и $A_i \uparrow A \Rightarrow A \in D_1)$
   Пусть $B_i \uparrow B$. Тогда

   $$(B_i \cap B_2 \cap \dots \cap B_n) \uparrow (B \cap B_2 \cap \dots \cap B_n)$$

   Поскольку вероятностная мера непрерывна, то

   $$P(B_i \cap B_2 \cap \dots \cap B_n) \to P(B \cap B_2 \cap \dots \cap B_n)$$

   причем

   $$P(B_i)P(B_2)\dots P(B_n) \to P(B)P(B_2)\dots P(B_n)$$

значит

$$P(B \cap B_2 \cap \dots \cap B_n) = P(B)P(B_2)\dots P(B_n)$$

а это означает, что $B \in D_1$.

Значит D_1 - это λ-система. Тогда, поскольку $M_1 \subset D_1$, то $\sigma\{M_1\} \subset D_1$. Далее определим D_2 — это будут все события B_2 такие, что

$$P(B_1 \cap B_2 \cap \dots \cap B_n) = P(B_1)P(B_2)\dots P(B_n)$$

для произвольных $B_1 \in \sigma\{M_1\},\ B_3 \in M_3,\ \dots,\ B_n \in M_n$. Аналогично, D_2 — λ-система, причем D_2 независима с $\sigma\{M_1\}$ и $\sigma\{M_2\} \subset D_2$ следовательно $\sigma\{M_1\}$ и $\sigma\{M_2\}$ независимы. Аналогично рассуждаем про $D_3, \dots, D_n$ и получаем, что $\sigma\{M_1\}, \dots, \sigma\{M_n\}$ независимы. □

3.3 Лемма Бореля-Кантелли

Лемма. (Лемма Бореля-Кантелли)

1. Если события $A_1, A_2, \dots$ таковы, что

$$\sum_{n=1}^{\infty} P(A_n) < \infty$$

то вероятность события "произошло бесконечное число A_n" равна 0.

2. Если $A_1, A_2, \dots$ независимые события, такие, что

$$\sum_{n=1}^{\infty} P(A_n) = \infty$$

то вероятность события "произошло бесконечное число A_n" равна 1.

Доказательство. Заметим, что событие "произошло бесконечное число A_n" можно представить в виде

$$A = \bigcap_{n=1}^{\infty} \left(\bigcup_{k=n}^{\infty} A_k \right)$$

Тогда

1. При $n \to \infty$

$$P\left(\bigcap_{n=1}^{\infty} \left(\bigcup_{k=n}^{\infty} A_k \right) \right) \le P\left(\bigcup_{k=n}^{\infty} A_k \right) \le \sum_{k=n}^{\infty} P(A_k) \to 0$$

2. Это равносильно тому, что

$$P\left(\overline{\bigcap_{n=1}^{\infty}\left(\bigcup_{k=n}^{\infty}A_k\right)}\right)=0$$

$$P\left(\overline{\bigcap_{n=1}^{\infty}\left(\bigcup_{k=n}^{\infty}A_k\right)}\right)=P\left(\bigcup_{n=1}^{\infty}\left(\overline{\bigcup_{k=n}^{\infty}A_k}\right)\right)=$$

$$=P\left(\bigcup_{n=1}^{\infty}\left(\bigcap_{k=n}^{\infty}\overline{A_k}\right)\right)\leq\sum_{n=1}^{\infty}P\left(\bigcap_{k=n}^{\infty}\overline{A_k}\right)$$

Покажем, что $\forall n\in\mathbb{N}$:

$$P\left(\bigcap_{k=n}^{\infty}\overline{A_k}\right)=0$$

Поскольку вероятностная мера непрерывна, то

$$P\left(\bigcap_{k=n}^{\infty}\overline{A_k}\right)=\lim_{N\to\infty}P\left(\bigcap_{k=n}^{N}\overline{A_k}\right)$$

$A_1,\dots,A_n$ независимы $\Rightarrow\overline{A_1},\dots,\overline{A_n}$ независимы по лемме. Поэтому

$$P\left(\bigcap_{k=n}^{N}\overline{A_k}\right)=\prod_{k=n}^{N}P(\overline{A_k})=\prod_{k=n}^{N}(1-P(A_k))\overset{(1)}{\leq}$$

$$\overset{(1)}{\leq}\prod_{k=n}^{N}e^{-P(A_k)}=\exp\left(-\sum_{k=n}^{N}P(A_k)\right)$$

(1): Используем неравенство: $\forall x\in\mathbb{R},\ 1-x\leq e^{-x}$ (Можно доказать, разложив экспоненту в степенной ряд).

При $N\to\infty$ получим

$$P\left(\bigcap_{k=n}^{\infty}\overline{A_k}\right)\leq\exp\left(-\sum_{k=n}^{\infty}P(A_k)\right)\to 0$$

$\square$

Лемма. (Лемма о групировке)

Пусть имеется семейство $\mathcal{A}_t,\ t\in T$ независимых σ-алгебр.

Возьмем $I_1,I_2,\dots\subset T: I_k\cap I_m=\varnothing,\ \forall k\neq m$ и для $I\subset T$ введем $\sigma\{I\}=\sigma\{\mathcal{A}_t,\ t\in I\}$. Тогда $\sigma\{I_1\},\sigma\{I_2\},\dots$ - независимые σ-алгербы.

4 Лекция 4

4.1 Функция распределения и плотность

Определение. Пусть (S, τ) - топологическое пространство. Наименьшая σ-алгебра, содержащая все открытые множества называется Борелевской σ-алгебро и обозначается $\mathcal{B}(S)$

Определение. Пусть P - вероятностная мера на $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Функция

$$F(x) = P((-\infty, x])$$

называется функцией распределения вероятностной меры.

Теорема. Функция распределения обладает следующими свойствами:

1. F не убывает
2. F непрерывна справа $\forall x \in \mathbb{R}$
3. $\lim\limits_{x \to -\infty} F(x) = 0$
4. $\lim\limits_{x \to +\infty} F(x) = 1$

Доказательство.

1. $x \le y \Rightarrow (-\infty, x] \subset (-\infty, y] \Rightarrow P((-\infty, x]) \le P((-\infty, y]) \Rightarrow F(x) \le F(y)$
2. Если x_n монотонно стремится в x, то $(-\infty, x_n] \downarrow (-\infty, x]$, значит, по свойству непрерывности вероятностной меры, $P((-\infty, x_n]) \to P((-\infty, x])$ $\Rightarrow F(x_n) \to F(x)$.
3. и 4. напрямую следуют из непрерывности вероятностной меры.

$\square$

Теорема. Если функция F обладается свойствами $(1) - (4)$, то на $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ существует единственная вероятностная мера, такая, что $\forall x \in \mathbb{R}:$

$$F(x) = P((-\infty, x])$$

Это задает взаимно-однозначное соответствие между вероятностными мерами и функциями распределения.

Определение. Пусть $p(x)$ вещественная кусочно-непрерывная функция. Если $\forall x \in \mathbb{R} : p(x) \geq 0$ и интеграл Римана

$$\int\limits_{\mathbb{R}} p(u)\ du = 1$$

то такая $p(x)$ называется плотностью.

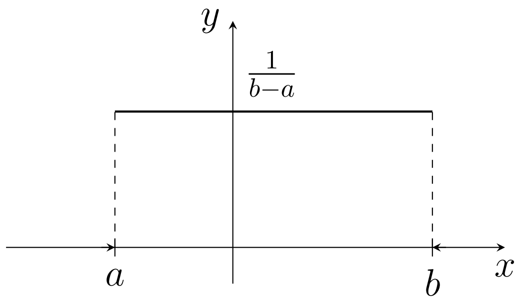
Утверждение.

$$F(x) = \int\limits_{-\infty}^{x} p(u)\ du$$

является функцией распределения.
(для нее выполнены свойства $(1) - (4)$)

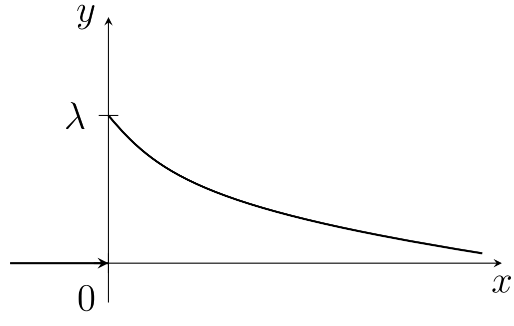
Пример. Равномерное распределение на $[a, b] \subset \mathbb{R}$

$$p(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b], \\ 0, & x \notin [a, b]. \end{cases}$$



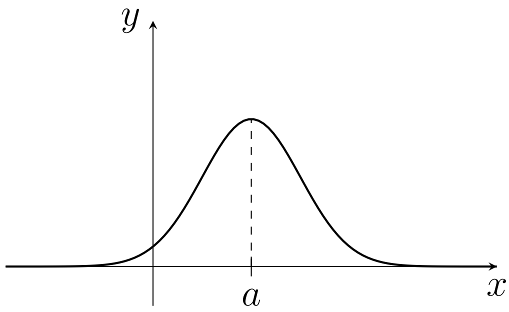
Пример. Экспоненциальное распределение с параметром $\lambda > 0$

$$p(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$



Пример. Гауссовское (нормальное) распределение

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-a)^2}{2\sigma^2}}$$



Упражнение. Доказать, что в каждом из этих случаев действительно выполнено

$$\int\limits_{\mathbb{R}} p(x)\ dx = 1$$

4.2 Случайные величины

Определение. Рассмотрим вероятностные пространства $(\Omega, \mathcal{F})$, $(S, \mathcal{B})$. Отображение $X : \Omega \to S$ называется $\mathcal{F}|\mathcal{B}$ измеримым (или случайным элементом), если $\forall B \in \mathcal{B} : X^{-1}(B) \in \mathcal{F}$. В этом случае, пишут $X \in \mathcal{F}|\mathcal{B}$.

Определение. Рассмотрим вероятностные пространства $(\Omega, \mathcal{F})$ и $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Если $X \in \mathcal{F}|\mathcal{B}(\mathbb{R})$, то X называется случайной величиной.

Теорема. Рассмотрим $f : V \to S$ и систему M подмножеств S. Определим множество $f^{-1}(M) = \{f^{-1}(B) : B \in M\}$ и рассмотрим $\sigma\{f^{-1}(M)\}$. Тогда

$$\sigma\{f^{-1}(M)\} = f^{-1}(\sigma\{M\})$$

Доказательство. Упражнение. $\square$

Следствие. Пусть $f : V \to S$ и $\mathcal{B} = \sigma\{M\}$. Пусть $f^{-1}(B) \in \mathcal{F}$ $(B \in \mathcal{B})$. Тогда $f \in \mathcal{F}|\mathcal{B}$.

Следствие. Пусть (V, τ) и (S, ν) - топологические пространства. Если $f : V \to S$ непрерывна $(f^{-1}(\nu) \subset \tau)$, то $f \in \mathcal{B}(V)|\mathcal{B}(S)$

4.3 Распределение случайного элемента

Определение. Рассмотрим вероятностные пространства $(\Omega, \mathcal{F})$ и $(S, \mathcal{B})$. Пусть $X : \Omega \to S,\ X \in \mathcal{F}|\mathcal{B}$. Распределением случайного элемента X называется вероятностная мера P_X на $(S, \mathcal{B})$ такая, что

$$P_X(B) = P(X^{-1}(B))$$

для любого $B \in \mathcal{B}$.

Упражнение. Проверить, что P_X является вероятностной мерой.

Определение. Если случайный элемент X распределен по распределению Q, то пишут $X \sim Q$

Определение. Случайные элементы $X_k : \Omega \to S,\ k = 1, \ldots, n$ называются независимыми, если независимы σ-алгебры $\sigma\{X_1\}, \ldots, \sigma\{X_n\}$, где
$\sigma\{X_i\} = \{X_i^{-1}(B),\ B \in \mathcal{B}\}$, что означает

$$P(X_1 \in B_1, \ldots, X_n \in B_n) = \prod_{k=1}^{n} P(X_k \in B_k)$$

Определение. Случайная величина X называется дискретной, если она принимает не более чем счетное число значений.

Определение. (Биномиальное распределение)
Дискретная случайная величина X имеет биномиальное распределение:
$X \sim B(n, p)$, где $n \in \mathbb{N},\ 0 < p < 1$, если

$$P(X = k) = C_n^k p^k (1-p)^{n-k},\ k = 0, 1, \ldots, m$$

Определение. (Распределение Пуассона)
Дискретная случайная величина распределяется по Пуассону:
$X \sim Pois(\lambda)$, если

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!},\ k \in \mathbb{Z}_+$$

Теорема. (Классическая теорема Пуассона)
Если $n \cdot p(n) \to \lambda > 0,\ n \to \infty$, то

$$C_n^k p^k(n)(1 - p(n))^{n-k} \to \frac{\lambda^k e^{-\lambda}}{k!}$$

Доказательство.
$$p^k(n) = \frac{(n\cdot p(n))^k}{n^k} \to \frac{\lambda^k}{n^k}$$
Заметим, что
$$n\cdot p(n) \to \lambda \Rightarrow p(n) \to 0$$
Значит
$$(1-p(n))^{n-k} = \frac{(1-p(n))^n}{(1-p(n))^k} \to (1-p(n))^n$$
Таким обзаром
$$\frac{n!}{k!(n-k)!}\cdot p^k(n)(1-p(n))^{n-k} \to \frac{n!}{k!(n-k)!}\cdot\frac{\lambda^k}{n^k}(1-p(n))^n$$
Нам нужно, чтобы
$$\frac{n!}{n^k(n-k)!}\cdot(1-p(n))^n \to e^{-\lambda}$$
$$(1-p(n))^n = (1-p(n))^{\frac{n\cdot p(n)}{p(n)}} = ((1-p(n))^{-\frac{1}{p(n)}})^{-n\cdot p(n)} \to e^{-\lambda}$$
Значит, остается доказать, что
$$\frac{n!}{n^k(n-k)!} \to 1$$
Действительно
$$\frac{n!}{n^k(n-k)!} = \frac{n(n-1)(n-2)\dots(n-k+1)}{n^k} \to 1$$
Объединив результаты, получим
$$\frac{n!}{k!(n-k)!}\cdot p^k(n)(1-p(n))^{n-k} \to \frac{\lambda^k e^{-\lambda}}{k!}$$
□

тут пропущена теорема которую я не понял

5 Лекция 5

5.1 Современная теорема Пуассона

Теорема. Пусть $X_1,\dots,X_n$ - независимые Бернуллиевские случайные элементы, $P(X_i=1)=p_i,\ P(X_i=0)=1-p_i,\ S_n = X_1+\dots+X_n,\ Y\sim Pois(\lambda)$, где $\lambda = p_1+\dots+p_n$. Тогда
$$\sup_{B\in\mathcal{B}(\mathbb{R})} |P(S_n\in B) - P(Y\in B)| \le \sum_{k=1}^n p_k^2$$
Для доказательства этой теоремы потребуются две леммы

Лемма. Если $Y_1, \dots, Y_n$ независимые случайные величины, причем $Y_k \sim Pois(\lambda_k)$. Тогда

$$\sum_{k=1}^{n} Y_k \sim Pois\left(\sum_{k=1}^{n} \lambda_k\right)$$

Доказательство. Сначала докажем, что $Y_1 + Y_2 \sim Pois(\lambda_1 + \lambda_2)$ при условии, что $Y_1 \sim Pois(\lambda_1),\ Y_2 \sim Pois(\lambda_2)$.

$$P(Y_1 + Y_2 = m) \overset{(1)}{=} \sum_{i=0}^{m} P(Y_1 + Y_2 = m,\ Y_1 = i) =$$
$$= \sum_{i=0}^{m} P(Y_1 = i,\ Y_2 = m - i) = \sum_{i=0}^{m} P(Y_1 = i) P(Y_2 = m - i)$$
$$= \sum_{i=0}^{m} \left(\frac{\lambda_1^i e^{-\lambda_1}}{i!} \cdot \frac{\lambda_2^{m-i} e^{-\lambda_2}}{(m-i)!}\right) \overset{(2)}{=} \frac{e^{-(\lambda_1+\lambda_2)}}{m!} \cdot \sum_{i=0}^{m} \left(\frac{m!}{i!(m-i)!} \lambda_1^i \lambda_2^{m-i} =\right) =$$
$$= \frac{e^{-(\lambda_1+\lambda_2)}}{m!} (\lambda_1 + \lambda_2)^m$$

(1): По формуле полной вероятности

(2): Домножили и поделили на $m!$

Далее, индукция по n: база очевидна, $Y_1 + \dots + Y_n = (Y_1 + \dots + Y_{n-1}) + Y_n$. По предположению индукции $Y_1 + \dots + Y_{n-1} \sim Pois(\lambda_1 + \dots + \lambda_{n-1})$ Поэтому, если $Y_1 + \dots + Y_{n-1}$ и Y_n независимы, то

$$\sum_{k=1}^{n} Y_k \sim Pois\left(\sum_{k=1}^{n} \lambda_k\right)$$

Остается показать, что $Y_1 + \dots + Y_{n-1}$ и Y_n независимы □

Лемма. Пусть X, Y случайные величины. Тогда $\forall B \in \mathcal{B}(\mathbb{R})$ выполнено:

$$|P(X \in B) - P(Y \in B)| \leq P(X \neq Y)$$

Доказательство. Рассмотрим разбиение Ω на четыре части:
$X(\omega) \in B$ и $Y(\omega) \in B$, $X(\omega) \in B$ и $Y(\omega) \notin B$, $X(\omega) \notin B$ и $Y(\omega) \in B$, $X(\omega) \notin B$ и $Y(\omega) \notin B$. Тогда по формуле полной вероятности:

$$P(X \in B) - P(Y \in B) = P(X \in B,\ Y \in B) +$$
$$+ P(X \in B,\ Y \notin B) - P(Y \in B,\ X \in B) - P(Y \in B, X \notin B) \leq$$
$$\leq P(X \in B,\ Y \notin B) \leq P(X \neq Y)$$

Аналогично можно показать, что $P(Y \in B) - P(X \in B) \leq P(X \neq Y)$. Тогда

$$|P(X \in B) - P(Y \in B)| \leq P(X \neq Y)$$

$\square$

Теперь можно перейти к доказательству теоремы. Напомним ее условие

Теорема. Пусть $X_1, \ldots, X_n$ - независимые Бернуллиевские случайные элементы, $P(X_i = 1) = p_i,\ P(X_i = 0) = 1 - p_i,\ S_n = X_1 + \cdots + X_n,\ Y \sim Pois(\lambda)$, где $\lambda = p_1 + \cdots + p_n$. Тогда

$$\sup_{B \in \mathcal{B}(\mathbb{R})} |P(S_n \in B) - P(Y \in B)| \leq \sum_{k=1}^{n} p_k^2$$

Доказательство.

$$|P(S_n \in B) - P(Y \in B)| = |P(X_1 + \cdots + X_n \in B) - P(Y \in B)| \overset{(1)}{\leq}$$
$$\overset{(1)}{\leq} P(X_1 + \cdots + X_n \neq Y) \overset{(2)}{=} P(X_1 + \cdots + X_n \neq Y_1 + \cdots + Y_n) \overset{(3)}{\leq}$$
$$\overset{(3)}{\leq} P\left(\bigcup_{k=1}^{n} \{X_k \neq Y_k\}\right) \overset{(4)}{\leq} \sum_{k=1}^{n} P(X_k \neq Y_k)$$

Пояснения:
(1): По лемме 2
(2): Так как

$$Y \sim Pois\left(\sum_{k=1}^{n} p_k\right)$$

то по лемме 1 $P(Y \in B) = P(Y_1 + \cdots + Y_n \in B)$, где $Y_1, \ldots, Y_n$ независимы и $Y_k \sim Pois(p_k)$
(3): $X_1 + \cdots + X_n \neq Y_1 + \cdots + Y_n \subset \bigcup_{k=1}^{n} \{X_k \neq Y_k\}$ это можно показать, используя тот факт, что $A \subset B \Leftrightarrow \overline{B} \subset \overline{A}$.
(4): Субаддитивность.

Введем $V_1, \ldots, V_n$ независимые с $Y_1, \ldots, Y_n$ такие, что $P(V_k = 0) = (1 - p_k)e^{p_k}$ и $P(V_k = 1) = 1 - P(V_k = 0)$, причем $(1 - p_k)e^{p_k} \leq 1$, так как $e^{-p_k} \geq (1 - p_k)$. Подберем $U_1, \ldots, U_n$ такие, что

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

и $\{U_k = 0\} = \{V_k = 0,\ Y_k = 0\}$, $\{U_k = 1\} = \Omega \setminus \{U_k = 0\}$.
Отсюда

$$P(U_k = 0) = P(V_k = 0, Y_k = 0) = P(V_k = 0)P(Y_k = 0) =$$
$$= (1 - p_k)e^{p_k} \cdot \frac{\lambda^0 e^{-p_k}}{0!} = 1 - p_k$$

а следовательно $P(U_k = 1) = 1 - (1 - p_k) = p_k$, значит, случайная величина X_k совпадает с U_k и нам достаточно оценить $P(U_k \neq Y_k)$.

$$P(U_k \neq Y_k) = P(U_k \neq Y_k,\ Y_k = 0) + P(U_k \neq Y_k,\ Y_k = 1) + P(U_k \neq Y_k,\ Y_k \geq 2)$$

Разберем каждое слагаемое:

1. $$P(U_k \neq Y_k,\ Y_k = 0) = P(U_k = 1,\ Y_k = 0)$$

2. Вспомним, что $\{U_k = 0\} = \{V_k = 0,\ Y_k = 0\}$. Тогда
$$P(U_k \neq Y_k,\ Y_k = 1) = P(U_k = 0,\ Y_k = 1) = P(V_k = 0,\ Y_k = 0,\ Y_k = 1) = 0$$

3. $$P(U_k \neq Y_k,\ Y_k \geq 2) = P(Y_k \geq 2)$$

Итого:
$$P(U_k \neq Y_k) = P(U_k = 1,\ Y_k = 0) + P(Y_k \geq 2)$$

Снова анализируем слагаемые:

1. $$\begin{aligned} P(U_k = 1,\ Y_k = 0) = P(Y_k = 0) - P(U_k = 0,\ Y_k = 0) = \\ = P(Y_k = 0) - P(V_k = 0,\ Y_k = 0,\ Y_k = 0) = \\ = P(Y_k = 0) - P(V_k = 0,\ Y_k = 0) = \\ = e^{-p_k} - (1 - p_k)e^{p_k}e^{-p_k} = p_k + e^{-p_k} - 1 \end{aligned}$$

2. $$P(Y_k \geq 2) = 1 - P(Y_k = 0) - P(Y_k = 1) = 1 - e^{-p_k} - p_k e^{-p_k}$$

Получили:

$$P(U_k \neq Y_k) = p_k + e^{-p_k} - 1 + 1 - e^{-p_k} - p_k e^{-p_k} =$$
$$= p_k - p_k e^{-p_k} = p_k(1 - e^{-p_k}) \leq p_k^2$$

Таким образом:

$$|P(S_n \in B) - P(Y \in B)| \le \sum_{k=1}^{n} P(U_k \ne Y_k) \le \sum_{k=1}^{n} p_k^2$$

$\square$

В доказательстве мы воспользовались теоремой

Теорема. (Теорема Ломницкого-Улама)
Пусть $(S_t, \mathcal{B}_t, Q_t)$ - семейство вероятностных пространств. Тогда на некотором вероятностном пространстве $(\Omega, \mathcal{F}, P)$ можно построить семейство независимых случайных элементов $X_t : \Omega \to S_t,\ X_t \in \mathcal{F}|\mathcal{B}_t$, причем $P_{X_t} = Q_t,\ t \in T$. Это означает, что всегда можно построить семейство независимых случайных элементов, имеющих заданное распределение.

5.2 Расширенные случайные величины

Определение. Рассмотрим $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$. Отображение $X : \Omega \to \overline{\mathbb{R}}$ назовем расширенной случайной величиной, если $X \in \mathcal{F}|\mathcal{B}(\overline{\mathbb{R}})$, где
$\mathcal{B}(\overline{\mathbb{R}}) = \{B,\ B \cup \{+\infty\},\ B \cup \{-\infty\},\ B \cup \{+\infty,\ -\infty\} : B \in \mathcal{B}(\mathbb{R})\}$

Утверждение. X - случайная величина (расширенная) $\Leftrightarrow \forall x \in \mathbb{R}$:
$\{\omega : X(\omega) \le x\} \in \mathcal{F}$

Теорема. Пусть $X_1, X_2, \ldots$ - случайные величины (расширенные). Тогда справедливы следующие утверждения

1. $\sup\limits_n X_n,\ \inf\limits_n X_n$ - случайные величины (расширенные)
2. $\lim\limits_{n\to\infty} \sup\limits_n X_n,\ \lim\limits_{n\to\infty} \inf\limits_n X_n$ - случайные величины (расширенные).
3. Если существует предел $\lim\limits_{n\to\infty} X_n$, то он является случайной величиной (расширенной).

Доказательство. Будем доказывать пользуясь утверждением

1. $\forall x \in \mathbb{R}$:

$$\{\omega : \sup_n X_n \le x\} = \bigcap_{n=1}^{\infty} \{\omega : X_n(\omega) \le x\} \in \mathcal{F}$$

$$\{\omega : \inf_n X_n \le x\} = \bigcup_{n=1}^{\infty} \{\omega : X_n(\omega) \le x\} \in \mathcal{F}$$

2. Сведем к первому пункту

$$\lim_{n\to\infty}\sup_n X_n(\omega) = \inf_n \sup_{k\leq n} X_k(\omega)$$

$$\lim_{n\to\infty}\inf_n X_n(\omega) = \sup_n \inf_{k\leq n} X_k(\omega)$$

3. Если существует предел, то он соппадает с верхним и нижним пределами, значит

$$\lim_{n\to\infty} X_n = \lim_{n\to\infty}\sup_n X_n$$

а это, в свою очередь, является случайной величиной по пункту 2.

□

6 Лекция 6

6.1 Математическое ожидание

Определение. Пусть $(\Omega, \mathcal{F}, P)$ - вероятностное пространство, $X : \Omega \to \mathbb{R}$, $X \in \mathcal{F}|\mathcal{B}(\mathbb{R})$. Математическим ожиданием случайной величины X назовем интеграл Лебега от случайной величины по вероятностной мере

$$EX = \int\limits_{\Omega} X\ d(P)$$

Таким образом, для определения понятия математического ожидания случайной величины, необходимо ввести понятие интеграла Лебега от случайной величины по вероятностной мере.

6.2 Построение интеграла Лебега

Построение интеграла Лебега происходит в три этапа:

1. Для простых случайных величин.

2. Для неотрицательных случайных величин.

3. Для произвольных случайных величин.

6.2.1 Первый этап

Определение. Случайная величина X называется простой, если

$$X(\omega) = \sum_{k=1}^{n} a_k I_{A_k}$$

где $A_1, \dots, A_n \in \mathcal{F}$ образуют разбиение Ω.

Определение. Если X - простая случайная величина, то

$$EX = \int\limits_{\Omega} X\ dP = \sum_{k=1}^{n} a_k P(A_k) = \sum_{k=1}^{n} a_k P(X = a_k)$$

Утверждение. Приведенное выше определение корректно.

Доказательство. Пусть $B_1, \dots, B_m$ - другое разбиение Ω и

$$X = \sum_{j=1}^{m} b_j I_{B_j}$$

необходимо показать, что

$$\sum_{k=1}^{n} a_k P(A_k) = \sum_{j=1}^{m} b_j P(B_j)$$

Рассмотрим $C_{kj} = A_k \cap B_j$ - новое разбиение Ω. Тогда, по формуле полной вероятности:

$$\sum_{k=1}^{n} a_k P(A_k) = \sum_{k=1}^{n} a_k \sum_{j=1}^{m} P(A_k \cap B_j) = \sum_{k=1}^{n} \sum_{j=1}^{m} a_k P(c_{kj}) =$$

$$= \sum_{(k,j) \in J} a_k P(C_{kj}) = \sum_{(k,j) \in J} b_j P(C_{kj}) = \sum_{j=1}^{m} b_j P(B_j)$$

где $J = \{(k, j) : C_{kj} \neq \varnothing\}$. Если $C_{kj} \neq \varnothing$, то $X(\omega) = a_k = b_j,\ \forall \omega \in C_{kj}$ $\square$

Теорема. Для простых случайных величин X и Y справедливы следующие утверждения:

1. $E(X + Y) = EX + EY$.
2. $\forall c \in \mathbb{R} : E(cX) = c \cdot EX$.
3. Если $X \geq 0$, то $EX \geq 0$.

4. Если $X \leq Y$, то $EX \leq EY$.

Доказательство.

1. Рассмотрим два случая

   (a) Для X и Y одно и то же разбиение $A_1, \dots, A_n$ на Ω. В этом случае, $\forall \omega \in A_k : X + Y(\omega) = X(\omega) + Y(\omega) = a_k + b_k$. Отсюда

   $$E(X+Y) = \sum_{k=1}^{n}(a_k+b_k)P(A_k) = \sum_{k=1}^{n} a_k P(A_k) + \sum_{k=1}^{n} b_k P(A_k) = EX+EY$$

   (b) Если разбиения $A_1, \dots, A_n$ для X и $B_1, \dots, B_m$ для Y оказались разными, то рассмотрим новое разбиение $C_{kj} = A_k \cap B_j$. Таким образом, свели к первому случаю.

2. $\forall c \in \mathbb{R}$:

   $$E(cX) = \sum_{k=1}^{n} c \cdot a_k P(A_k) = c \cdot \sum_{k=1}^{n} a_k P(A_k) = c \cdot EX$$

3. Если $X \geq 0$, то $a_k \geq 0$. Значит

   $$EX = \sum_{k=1}^{n} a_k P(A_k) \geq 0$$

4. Пусть $X \geq Y \Rightarrow X - Y \geq 0 \Rightarrow E(X - Y) \geq 0$. Тогда из доказанных свойств:

   $$E(X - Y) = E(X + (-1) \cdot Y) = EX + E((-1) \cdot Y) = EX - EY \geq 0$$

□

6.2.2 Второй этап

Определение. Пусть $X \geq 0$ случайная величина. Тогда

$$EX = \sup\{EY : 0 \leq Y \leq X,\ Y - \text{простая}\}$$

Лемма. Пусть $X \geq 0$ случайная величина. Тогда существует последовательность простых функций $X_n \geq 0$ таких, что $X_n \to X$ и $X_n(\omega) \leq X_{n+1}(\omega)$, $\forall \omega \in \Omega$

Доказательство. Возьмем последовательность

$$X_n = \min\{(2^{-n}\ [2^n X]\ ),\ n\}$$

Заметим, что

$$X_n(\omega) = \begin{cases} k2^{-n}, & k2^{-n} \le X(\omega) < (k+1)2^{-n},\ k = 0, 1, \dots, n2^n - 1, \\ n \quad , & X(\omega) \ge n. \end{cases}$$

$$|X_n(\omega) - X(\omega)| \le 2^{-n} \to 0$$

Тогда $X_n \to X$. $\square$

Лемма. Пусть $X \ge 0$ случайная величина. Тогда для любой последовательности простых функций $0 \le X_n \le X,\ X_n \to X$ выполнено, что $EX_n \to EX$.

Доказательство. X_n - простые и $X_n \le X_{n+1} \Rightarrow EX_n \le EX_{n+1}$. При этом

$$EX = \sup\{EY : 0 \le Y \le X,\ Y - \text{простая}\}$$

значит существует предел

$$\lim_{n\to\infty} EX_n = a \le EX$$

Покажем, что выполнено и обратное неравенство. Для этого докажем, что для любой простой фунции $0 \le Y \le X$ выполнено, что $EY \le a$. Если $Y \equiv 0$, то утверждение верно. Пусть

$$Y = \begin{cases} 0 = b_0 < b_1 < \dots < b_k \\ \qquad p_0,\ \ p_1\ \ ,\ \dots\ ,\ p_k \end{cases}$$

Тогда

$$EY = \sum_{j=1}^{k} b_j p_j$$

Введем $Y_n = (1-\varepsilon)Y \cdot I\{(1-\varepsilon)Y \le X_n\},\ n \in \mathbb{N},\ \varepsilon \in (0,1)$. Из определения $Y_n \le X_n \Rightarrow EY_n \le EX_n \le a$. Пусть

$$B_j = \{\omega : Y(\omega) = b_j\},\ C_{jn} = \{(1-\varepsilon)Y \le X_n\}$$

Тогда $Y_n(\omega) = (1-\varepsilon)b_j$, если $\omega \in B_j \cap C_{jn}$. Получаем, что

$$EY_n = \sum_j (1-\varepsilon) b_j P(B_j \cap C_{jn})$$

При этом, по свойству непрерывности вероятностной меры

$$P(B_j \cap C_{jn}) \to P(B_j)$$

следовательно

$$EY_n = \sum_j (1-\varepsilon) b_j P(B_j \cap C_{jn}) \to (1-\varepsilon)\sum_j b_j P(B_j) = (1-\varepsilon)EY$$

Получили, что $\forall \varepsilon \in (0,1)$:

$$(1-\varepsilon)EY \le a$$

тогда, устремив $\varepsilon \to 0$, получим для любой $Y : EY \le a \Rightarrow EX \le a$. Из полученных неравентв, следует, что

$$\lim_{n\to\infty} EX_n = EX$$

□

Лемма. Пусть $X, Y \ge 0$ случайные величины. Тогда

1. $E(X+Y) = EX + EY$.
2. $\forall c \ge 0 : E(cX) = c \cdot EX$

Доказательство.

1. Возьмем неубывающие последовательности простых функций $X_n \to X$, $Y_n \to Y$. Тогда $X_n + Y_n \to X + Y$ и по лемме 2:

$$E(X_n + Y_n) = EX_n + EY_n \to E(X+Y)$$

2. Возьмем неубывающую последовательность простых функций $X_n \to X$. Тогда $cX_n \to cX$ и $E(cX_n) \to E(cX)$ и по лемме 2:

$$E(cX_n) = c \cdot E(X_n) \to c \cdot E(X)$$

□

6.2.3 Третий этап

Замечание. Заметим, что любую функцию $X : \Omega \to \mathbb{R}$ можно представить в виде $X = X^+ - X^-$, где

$$X^+ = \max\{X,\ 0\} = X \cdot I\{X \ge 0\}, \quad X^- = -\min\{X, 0\} = -X \cdot I\{X \le 0\}$$

Определение. Пусть X - произвольная случайная величина. Тогда

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

по определению положим, что EX не существует, если $EX^+ = +\infty$ и $EX^- = +\infty$ одновременно. Также по определению для любой константы $+\infty - c = +\infty$ и $c - \infty = -\infty$

Определение. Если $EX \in \mathbb{R}$, то случайная величина X называется интегрируемой по Лебегу и обозначается $X \in \mathcal{L}^1$

Теорема. Для произвольной случайной величины справедливы следующие утверждения:

1. Если $Y \leq X$ и $X \in \mathcal{L}^1$, то $Y \in \mathcal{L}^1$.
2. Если $X, Y \in \mathcal{L}^1$ и $Y \leq X$, то $EY \leq EX$.
3. Если $X \in \mathcal{L}^1$, то $|EX| \leq E|X|$.
4. $\mathcal{L}^1$ - линейное пространство.
5. E - линейный функционал.

6.3 Дополнение (материалы с семинара)

Утверждение. Если $X \sim B(n, p)$, то $EX = np$.

Утверждение. Если $X \sim Pois(\lambda)$, то $EX = \lambda$

Утверждение. Если X и Y независимые случайные величины, то $E(XY) = EX \cdot EY$

7 Лекция 7

7.1 Теорема о монотонной сходимости

Теорема. (Теорема о монотонной сходимости)
Пусть $X_n \geq 0$ случайные величины и $X_n \nearrow X$. Тогда $EX_n \to EX$.

Доказательство. Пусть

$$\begin{gathered} 0 \leq Y_{n,1} \nearrow X_1 \\ 0 \leq Y_{n,2} \nearrow X_2 \\ \vdots \\ 0 \leq Y_{n,k} \nearrow X_k \end{gathered}$$

где $X_1 \leq X_2 \leq \dots \leq X_k \to X,\ Y_{n,i}$ - простые. Рассмотрим

$$\begin{gathered} Y_{1,1} \ \leq Y_{2,1} \leq \dots \leq Y_{k,1} \leq \dots \leq Y_{n,1} \nearrow X_1 \\ \vdots \\ Y_{1,k} \ \leq Y_{2,k} \leq \dots \leq Y_{k,k} \leq \dots \leq Y_{n,k} \nearrow X_k \end{gathered}$$

и определим

$$0 \le Z_k := \max_{1\le i,j\le k} Y_{i,j} \nearrow Z$$

Таким образом

$$0 \le Y_{n,k} \le Z_k \le X_k \le X \tag{1}$$

После предельного перехода при $k \to \infty$, получим

$$0 \le X_n \le Z \le X$$

После предельного перехода при $n \to \infty$, получим

$$0 \le Z_k \nearrow X \tag{2}$$

Так как $X_n \le X$, то $EX_n \le EX$, а значит

$$\lim_{n\to\infty} EX_n \le EX$$

Также из (1) и (2) имеем:

$$EZ_n \le EX_n$$

и как следствие

$$EX \le \lim_{n\to\infty} EX_n$$

$\square$

Утверждение. Рассмотрим вероятностные пространства $(\Omega, \mathcal{F}, P)$, $(S, \mathcal{B}, P_X)$, $X \in \mathcal{F}|\mathcal{B}$, где $\forall B \in \mathcal{B} : P_X(B) = P(X^{-1}(B))$, т.е P_X - мера на $(S, \mathcal{B})$, индуцированная X. Также рассмотрим вероятностное пространство $(\mathbb{R},\ \mathcal{B}(\mathbb{R}))$ и $h \in \mathcal{B}|\mathcal{B}(\mathbb{R})$. Теперь рассмотрим композицию отображений $h(X(\omega)) \in \mathcal{F}|\mathcal{B}(\mathbb{R})$. Тогда оба интеграла существуют или не существуют одновременно, и если существуют, то равны:

$$Eh(X) = \int\limits_\Omega h(X(\omega))P(d\omega) = \int\limits_S h(x)P_X(dx)$$

Доказательство. Сначала пусть h - индикатор:

$$h(x) = I_B(x) = \begin{cases} 1,\ x \in B, \\ 0,\ x \notin B \end{cases}$$

$$\int\limits_S h(x)P_X(dx) = 1 \cdot P_X(B) = P(X^{-1}(B))$$

$$\int\limits_{\Omega} h(X(\omega))P(d\omega) = 1 \cdot P(X^{-1}(B))$$

Тогда формула будет верна для линейной комбинации индикаторов, а значит для простых $h \geq 0$. По теореме о монотонной сходимости, формула будет верна для произвольной

$$h = h^+ - h^-$$

$\square$

Замечание. Пусть дискретная случайная величина

$$X = \begin{cases} x_1,\ x_2, \ldots \\ p_1,\ p_2, \ldots \end{cases}$$

$$Eh(X) = \sum_{h(x_n) \geq 0} h(x_n)P(X = x_n) + \sum_{h(x_n) < 0} h(x_n)P(X = x_n)$$

То есть конечное $Eh(X)$ существует, если

$$\sum_n |h(x_n)| \cdot P(X = x_n) < \infty$$

7.2 Дисперсия случайной величины

Определение. Рассмотрим случайную величину X. Пусть $X \in \mathcal{L}^1$, $(X - EX)^2 \in \mathcal{L}^1$. Тогда дисперсией случайной величины X называется выражение

$$\operatorname{var} X = E(X - EX)^2$$

Замечание. $(X - EX)^2 = X^2 - 2XEX + (EX)^2 \in \mathcal{L}^1 \Rightarrow X \in \mathcal{L}^2$

Замечание. Если X - дискретная случайная величина, то

$$\operatorname{var} X = \sum_n (X_n - EX)^2 P(X = x_n)$$

Определение. Рассмотрим случайные величины X и Y. Пусть $X, Y \in \mathcal{L}^1$, $(X - EX) \in \mathcal{L}^1$. Тогда ковариацией случайных величин X и Y называется выражение

$$\operatorname{cov}(X, Y) = E((X - EX)(Y - EY))$$

Замечание.

$$\operatorname{cov}(X, X) = \operatorname{var} X$$

Теорема. Пусть $X_1, X_2, \dots \in \mathcal{L}^2$ - случайные величины. Тогда

$$\operatorname{var}\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} \operatorname{var} X_k + 2 \sum_{1 \le i < j \le n} \operatorname{cov}(X_i, X_j)$$

Доказательство.

$$\operatorname{var}\left(\sum_{k=1}^{n} X_k\right) = E\left(\sum_{k=1}^{n} X_k - E\left(\sum_{k=1}^{n} X_k\right)\right)^2 \overset{(1)}{=}$$

$$\overset{(1)}{=} E\left(\sum_{k=1}^{n} (X_k - EX_K)\right)^2 \overset{(2)}{=} E\left(\sum_{i,j=1}^{n} (X_i - EX_i)(X_j - EX_j)\right) \overset{(1)}{=}$$

$$\overset{(1)}{=} \sum_{i,j=1}^{n} E((X_i - EX_i)(X_j - EX_j)) = \sum_{i,j=1}^{n} \operatorname{cov}(X_i, X_j)$$

(1): По линейности математического ожидания:

$$E\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} EX_k$$

(2): Поскольку

$$\left(\sum_{k=1}^{n} a_k\right)^2 = (a_1 + \dots + a_n)(a_1 + \dots + a_n)$$

Итак, поскольку $\operatorname{cov}(X, X) = \operatorname{var} X$, то

$$\operatorname{var}\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} \operatorname{var} X_k + 2 \sum_{1 \le i < j \le n} \operatorname{cov}(X_i, X_j)$$

□

Следствие. Пусть $X_1, X_2, \dots \in \mathcal{L}^2$ - независимые случайные величины. Тогда

$$\operatorname{var}\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} \operatorname{var} X_k$$

Доказательство. Если $X_1, \dots, X_n$ - независимы в совокупности, то они попарно независимы, тогда независимы величины $X_i - EX_i$ и $X_j - EX_j$. Тогда

$$E((X_i - EX_i)(X_j - EX_j)) = E(X_i - EX_i) \cdot E(X_j - EX_j) = 0$$

так как

$$E(X_i - EX_i) = EX_i - EX_i = 0$$

□

7.3 Неравенство Маркова

Теорема. (Неравенство Маркова)
Пусть $f : R_+ \to R_+,\ f$ - не убывает. Тогда для любой случайной величины X и $\forall t \geq 0$:

$$Ef(|X|) \geq f(t) \cdot P(|X| \geq t)$$

Доказательство.

$$f(|X|) \geq f(|X|) \cdot I\{|X| \geq t\} \geq f(t) \cdot I\{|X| \geq t\}$$

Значит

$$Ef(|X|) \geq E(f(t) \cdot I\{|X| \geq t\}) = f(t) \cdot E(I\{|X| \geq t\}) = f(t) \cdot P(|X| \geq t)$$

□

Следствие. (Неравенство Бьенеме-Чебышёва)

$$P(|X - EX| \geq \varepsilon) \leq \frac{\operatorname{var} X}{\varepsilon^2}$$

Доказательство. Возьмем $f(t) = t^2$. По неравенству Маркова:

$$P(|X - EX| \geq \varepsilon) \leq \frac{Ef(|X - EX|)}{f(\varepsilon)} = \frac{E(X - EX)^2}{\varepsilon^2} = \frac{\operatorname{var} X}{\varepsilon^2}$$

□

7.4 Бернуллиевские случайные величины

Определение. Случайная величина X называется бернуллиевской, если она принимает всего два значения 0 и 1 с вероятностями

$$P(X = 1) = p, \quad P(X = 0) = 1 - p$$

Теорема. (Теорема Бернулли)
Пусть $X_1, X_2, \ldots$ независимые случайные Бернуллиевские случайные величины, $S_n = X_1 + \cdots + X_n$. Тогда $\forall \varepsilon > 0$

$$P\left(\left| \frac{S_n}{n} - p \right| \geq \varepsilon \right) \to 0,\ n \to \infty$$

Это означает, что $\frac{S_n}{n}$ при больших n имеют как-бы неслучайный порядок поведения.

Доказательство.

$$p\left(\left|\frac{S_n}{n}-\frac{1}{n}\cdot\sum_{k=1}^{n}EX_k\right|\geq\varepsilon\right)=P(|S_n-ES_n|\geq n\varepsilon)\overset{(1)}{\leq}\frac{\operatorname{var}S_n}{(n\varepsilon)^2}=$$

$$=\frac{1}{n^2\varepsilon^2}\cdot\sum_{k=1}^{n}\operatorname{var}X_k\leq\frac{C\cdot n}{n^2\varepsilon^2}=\frac{C}{n\varepsilon^2}\to 0,\ n\to\infty$$

(1): По неравенству Бьенеме-Чебышёва. □

Определение. Случайные величины Z_n сходятся по вероятности в Z, если $\forall\varepsilon>0:$

$$P(|Z_n-Z|\geq\varepsilon)\to 0$$

обозначают

$$Z_n\overset{P}{\to}Z$$

Таким образом, теорема Бернулли показывает, что

$$\frac{S_n}{n}\overset{P}{\to}p,\ n\to\infty$$

Это простейшая (слабая) форма закона больших чисел.

7.5 Теорема Вейерштрасса об аппроксимации

Теорема. (Теорема Вейерштрасса об аппроксимации)
Пусть $f(x)$ непрерывна на $[a,b]$. Тогда $\forall\varepsilon>0$ существует многочлен $P_n(x)$ такой, что

$$\sup_{x\in[a,b]}|f(x)-P_n(x)|<\varepsilon$$

Доказательство. Временно пропущено □

8 Лекция 8

8.1 Стабильность случайных величин

Лемма. $E|X|=0\Leftrightarrow X=0$ почти наверное, то есть $\forall\omega\notin A$, $P(A)=0:X(\omega)=0$

Доказательство. Достаточно доказать для неотрицательных случаных величин, поскольку $X=X^+-X^-$ и

$$X=0\Leftrightarrow\begin{cases}X^+=0\\X^-=0\end{cases}$$

Далее $X\geq 0$.

($\Rightarrow$): Пусть $E|X| = 0$. По неравенству Маркова $\forall n \in \mathbb{N}$:

$$P\left(|X| \geq \frac{1}{n}\right) \leq \frac{E|X|}{\frac{1}{n}} = 0$$

$$\{|X| > 0\} = \bigcup_{n=1}^{\infty} \{|X| \geq \frac{1}{n}\}$$

$$P(|X| > 0) \leq \sum_{n=1}^{\infty} P(|X| \geq \frac{1}{n}) = 0$$

($\Leftarrow$): Пусть $X = 0$ почти наверное, $X \geq 0 \Rightarrow EX \geq 0$. Рассмотрим

$$EX_n \to EX,\ 0 \leq X_n \nearrow X,\ X_n = 2^{-n}[2^n X] \wedge n$$

Заметим, что $\{X_n \in [0, 2^{-n})\} \subset \{X_n = 0\}$

$$EX_n = \sum_{k=0}^{n2^n} k2^{-n} P(k2^{-n} \leq X \leq (k+1)2^{-n}) = 0$$

Значит $\forall n \in \mathbb{N} : EX_n = 0 \Rightarrow EX = 0.$

□

Следствие. Если $X = Y$ почти наверное и $X \in \mathcal{L}^1$, то $Y \in \mathcal{L}^1$ и $EX = EY$.

Доказательство. $Y = X + Y - X,\ X = Y$ почти наверное $\Rightarrow Y - X = 0$ почти наверное. Отсюда существует

$$EY = EX + E(Y - X) = EX$$

□

Утверждение. Отношение $X \sim Y \Leftrightarrow X = Y$ почти наверное - отношение эквивалентности.

Доказательство.

1. $P(X = X) = 1 \Rightarrow X \sim X$.

2. $X \sim Y \Rightarrow P(X \neq Y) = 0$ и так как $\{X \neq Y\} = \{Y \neq X\}$, то $P(Y \neq X) = 0$ $\Rightarrow Y \sim X$.

3. $X \sim Y$ и $Y \sim Z \Rightarrow P(X \neq Y) = 0$ и $P(Y \neq Z) = 0$. Заметим, что если $X \neq Z$, то либо $X \neq Y$, либо $Y \neq Z$, то есть

$$\{X \neq Z\} \subset \{X \neq Y\} \cup \{Y \neq Z\}$$

Отсюда

$$P(X \neq Z) \leq P(X \neq Y) + P(Y \neq Z) = 0 + 0 = 0$$

значит $X \sim Z$.

□

Определение. Определим пространство L^p как совокупность классов эквивалентности величин из $\mathcal{L}^p$.

Утверждение. Если $p \geq 1$, то L^p - нормированное пространство, причем $\forall X \in L^p:$

$$\|X\| = (E|X|^p)^{\frac{1}{p}}$$

Утверждение. L^2 является гильбертовым пространством, причем

$$(X, Y) = E(XY), \quad \|X\|_2 = (EX^2)^{\frac{1}{2}}$$

то есть

$$\rho(X, Y) = \|X - Y\|_2$$

по этой норме L_2 - полно.

Лемма. (Неравенство Коши-Буняковского-Шварца)
Пусть H - гильбертово пространство. Тогда $\forall X, Y \in H:$

$$|(X, Y)| \leq \|X\| \cdot \|Y\|$$

Доказательство. Для $t \in \mathbb{R}$ рассмотрим

$$(tX + Y, tX + Y) = t^2(X, X) + 2(X, Y)t + (Y, Y) \geq 0$$

отсюда

$$D = (X, Y)^2 - (X, X)(Y, Y) \leq 0$$

□

Определение. Пусть $X, Y \in L^2, \ \operatorname{var} X \neq 0, \ \operatorname{var} Y \neq 0$. Коэффициентом корреляции Пирсона называется

$$\rho(X, Y) = \frac{\operatorname{cov}(X, Y)}{\sqrt{\operatorname{var} X \cdot \operatorname{var} Y}}$$

$$\operatorname{cov}(X, Y) = E((E - EX)(Y - EY))$$

отсюда, по неравенству КБШ: $\rho(X, Y) \leq 1$

Упражнение. $|\rho(X, Y)| = 1 \Leftrightarrow X$ и Y - линено зависимы.

Замечание. Далее равенство случайных величин всегда подразумеваем как равенство почти наверное.