Теория вероятностей. Вопросы к экзамену

Этот файл — вырезка из лекций, которая составляется для удобства подготовки к экзамену. Можно сказать, что составляется она один раз: всякие ошибки, будь они типографическими или смысловыми, будут, скорее всего, исправлены только в конспекте лекций.

Замечание. *Если строгость/нестрогость знака сравнения не имеет значения, то для обозначения этого используется знак* $\preccurlyeq \in \{<, \leqslant\}$. *При ведении лекций вместо* $\preccurlyeq$ *использовался знак, похожий на* $\leqslant$, *где нижняя черта рисовалась пунктиром. Похожего знака нет в стандартном наборе знаков LaTeX, поэтому была выбрана замена.*

1 Случайные события

1.1 Определение пространства элементарных исходов, примеры. Понятие события (нестрогое), следствие события, невозможное и достоверное событие, примеры. Операции над событиями. Сформулировать классическое определение вероятности и доказать его следствия.

Определение. *Случайным называют эксперимент, результат которого невозможно точно предсказать заранее.*

Определение. *Множество* Ω *всех исходов данного случайного эксперимента называют пространством элементарных исходов.*

Замечание. *Всюду в дальнейшем будем предполагать, что*

1. *Каждый элементарный исход является «неделимым», т. е. он не может быть разбит на более «мелкие» исходы;*
2. *В результате каждого эксперимента обязательно имеет место ровно один из входящих в* Ω *элементарных исходов.*

Пример. *Бросают монетку. Возможные исходы: выпадение герба или решки.*
$\Omega = \{$ *Герб* , *Решка* $\}$ *— множество элементарных исходов.* $|\Omega| = 2$

Пример. *Бросают игральную кость:* $\Omega = \{$ *"1", "2", "3", "4", "5", "6"*$\}$, $|\Omega| = 6$

Нестрогое определение. *Событием будем называть произвольное подмножество множества элементарных исходов множества* Ω.

Определение. *Событие* A *называется следствием события* B, *если из того, что произошло* B, *следует то, что произошло* A, *т. е.* $B \subseteq A$.

Замечание. *Любое множество Ω содержит в себе два подмножества: Ω и $\emptyset$. События, соответствующие данным множествам, называются невозможным и достоверным соответственно. Эти события называются несобственными событиями. Все остальные события называются собственными.*

Пример. *Из урны, содержащей два красных и три синих шара, извлекают один шар. Возможные события: $A = \{$ извлечённый шар является красным или синим $\}$ — является достоверным, $B = \{$ извлечён белый шар $\}$ — невозможным.*

Операции над событиями.

События — множества элементарных исходов. Следовательно, над ними можно выполнять все операции над множествами. При этом вводится следующая терминология:

- Объединение множеств принято называть суммой событий: $A \cup B = A + B$;
- Пересечение множеств называют произведением событий: $A \cap B = A \cdot B$;
- $A \setminus B$ называют разностью событий A и B;
- Дополнение A называют событием, противоположным A: $\overline{A} = \Omega \setminus A$.

Пусть

1. Ω — пространство исходов некоторого случайного эксперимента ($|\Omega| = N < \infty$)[1];
2. По условиям эксперимента нет оснований предпочесть тот или иной элементарный исход остальным (в таком случае говорят, что все элементарные исходы равновозможны);
3. Существует событие $A \subseteq \Omega$, мощность $|A| \overset{\text{(обозначим)}}{=} N_A$

Тогда

Определение. *Вероятностью осуществления события A называют число*

$$P\{A\} = \frac{N_A}{N}$$

Свойства вероятности (в соответствии с классическим определением).

1. Вероятность $P(A) \geqslant 0$ (неотрицательна).
2. $P(\Omega) = 1$.

[1] Запись $x < \infty$ означает, что x конечно. Напротив, запись $x \leqslant \infty$ означает, что x либо конечно, либо бесконечно. — Прим. ред.

3. Если A, B — несовместные события, то $P(A+B) = P(A) + P(B)$.

Доказательства этих свойств:

1. Т. к. $N_A \geqslant 0$, $N > 0$, то следует $P(A) = \frac{N_A}{N} \geqslant 0$.

2. Принимая во внимание, что $N_\Omega = |\Omega| = N$, получается
$$P(\Omega) = \frac{N_\Omega}{N} = \frac{N}{N} = 1$$

3. Т. к. Ω — конечно, $A, B \subseteq \Omega$, то получается, что A, B конечны. Существует формула[2]
$$|A+B| = |A| + |B| - |AB|$$
Т. к. A и B — несовместные, то $AB = \emptyset$, из чего следует, что $N_{a+b} = N_a + B_b$. Таким образом,
$$P(A+B) = \frac{N_{a+b}}{N} = \frac{N_a + N_b}{N} = \frac{N_a}{N} + \frac{N_b}{N} = P(A) + P(B)$$

1.2 Определение пространства элементарных исходов, примеры. Понятие события (нестрогое). Сформулировать геометрическое и статистическое определения вероятности. Достоинства и недостатки этих определений.

Определение. *Случайным называют эксперимент, результат которого невозможно точно предсказать заранее.*

Определение. *Множество Ω всех исходов данного случайного эксперимента называют пространством элементарных исходов.*

Замечание. *Всюду в дальнейшем будем предполагать, что*

1. *Каждый элементарный исход является «неделимым», т. е. он не может быть разбит на более «мелкие» исходы;*

2. *В результате каждого эксперимента обязательно имеет место ровно один из входящих в Ω элементарных исходов.*

Пример. *Бросают монетку. Возможные исходы: выпадение герба или решки.*
$\Omega = \{$ *Герб* , *Решка* $\}$ — *множество элементарных исходов.* $|\Omega| = 2$

Пример. *Бросают игральную кость:* $\Omega = \{$ *“1”, “2”, “3”, “4”, “5”, “6”*$\}$, $|\Omega| = 6$

[2] Её называют формулой включений и исключений. — Прим. лект.

Нестрогое определение. *Событием будем называть произвольное подмножество множества элементарных исходов множества Ω.*

Геометрическое определение вероятности.

Геометрическое определение вероятности является обобщением классического определения на случай, когда $|\Omega| = \infty$.

Пусть

1. $\Omega \subseteq \mathbb{R}^n$;
2. $\mu(\Omega) < \infty$, где μ — некая мера.

   Если $n = 1$, то μ — это длина; если $n = 2$, то μ — площадь; если $n = 3$ — объём. Можно определить меры и при больших n;
3. Возможность принадлежности некоторого элементарного исхода случайного эксперимента событию $A \subseteq Q$ пропорциональна мере этого события и не зависит от формы события A и его расположения внутри Ω.

Тогда

Определение. *Вероятностью случайного события $A \subseteq \Omega$ называют число*

$$P\{A\} = \frac{\mu(A)}{\mu(\Omega)}$$

Геометрическое определение, в отличие от классического, применимо в случае бесконечного пространства исходов.

Недостаток геометрического определения заключается в том, что оно не учитывает возможность того, что некоторые области внутри Ω окажутся более предпочтительными, чем другие.

Статистическое определение вероятности.

Пусть

1. Некоторый случайный эксперимент произведён n раз;
2. При этом некоторое наблюдаемое в этом эксперименте событие A произошло n_A раз.

Определение. *Вероятностью осуществления события A называют эмпирический (т. е. найденный экспериментальным путём) предел:*

$$P(A) = \lim_{n\to\infty} \frac{n_a}{n}$$

Замечание. *У статистического определения полным-полно недостатков:*

1. *Никакой эксперимент не может быть произведён бесконечное много раз;*

2. *С точки зрения современной математики статистическое определение является архаизмом, т. к. не даёт достаточно базы для дальнейшего построения теории.*

1.3 Определение пространства элементарных исходов, примеры. Сформулировать определение сигма-алгебры событий. Доказать простейшие свойства сигма-алгебры. Сформулировать аксиоматическое определение вероятности.

Определение. *Случайным называют эксперимент, результат которого невозможно точно предсказать заранее.*

Определение. *Множество Ω всех исходов данного случайного эксперимента называют пространством элементарных исходов.*

Замечание. *Всюду в дальнейшем будем предполагать, что*

1. *Каждый элементарный исход является «неделимым», т. е. он не может быть разбит на более «мелкие» исходы;*

2. *В результате каждого эксперимента обязательно имеет место ровно один из входящих в Ω элементарных исходов.*

Пример. *Бросают монетку. Возможные исходы: выпадение герба или решки.*
$\Omega = \{$ *Герб* , *Решка* $\}$ — *множество элементарных исходов.* $|\Omega| = 2$

Пример. *Бросают игральную кость:* $\Omega = \{$ *“1”, “2”, “3”, “4”, “5”, “6”*$\}$, $|\Omega| = 6$

Пусть

1. Ω — пространство элементарных исходов, связанных с некоторым случайным экспериментом;

2. $\beta \neq \emptyset$ — система (набор) подмножеств в множестве Ω.

Определение. *β называется сигма-алгеброй*[3] *событий, если выполнены условия:*

1. *Если $A \in \beta$, то $\overline{A} \in \beta$;* [4]

2. *Если $A_1, \ldots, A_n, \ldots \in \beta$, то $A_1 + \ldots + A_n + \ldots \in \beta$.*

[3]При ведении лекций слово «сигма» иногда заменялось на букву δ (дельта — `\delta` в LaTeX). Буква «сигма» выглядит как σ. Лектор говорит, что корректнее всего словосочетание «сигма-алгебра» вообще не сокращать и писать полностью, не используя греческие буквы. — Прим. ред.

[4]Обратите внимание, что $A \subseteq \Omega$, но $A \in \beta$, т. к. элементы множества β — подмножества из Ω. — Прим. лект.

Простейшие следствия из аксиом сигма-алгебры

1. $\Omega \in \beta$;
2. $\emptyset \in \beta$;
3. Если $A_1, \ldots, A_n, \ldots \in \beta$, то $A_1 \cdot \ldots \cdot A_n \cdot \ldots \in \beta$;
4. Если $A, B \in \beta$; то $A \setminus B \in \beta$.

Доказательства этих следствий:

1. По определению $\beta \neq \emptyset \implies \exists A \subseteq Q \colon A \in \beta$; из определения сигма-алгебры (аксиома 1) $\exists A \in \beta \implies \overline{A} \in \beta$; тогда из второй аксиомы следует, что $\exists (A + \overline{A}) \in \beta$; т. к. $A + \overline{A} = \Omega$, то $\Omega \in \beta$.
2. Т. к. $\Omega \in \beta$ (по следствию 1), то, по аксиоме 1, $\overline{\Omega} \in \beta$, а $\overline{\Omega} = \emptyset$. Следовательно, $\emptyset \in \beta$.
3. Из существования событий $A_1, \ldots, A_n, \ldots \in \beta$ по аксиоме 1 следует, что существуют дополнения этих событий $\overline{A_1}, \ldots, \overline{A_n}, \ldots \in \beta$. По аксиоме 2 следует существование объединения $\overline{A_1} + \ldots + \overline{A_n} + \ldots \in \beta$, и из аксиомы 1 — существование дополнения этого объединения: $\overline{\overline{A_1} + \ldots + \overline{A_n} + \ldots} \in \beta$. Из этого, по законам де Моргана, получается $\overline{\overline{A_1}} \cdot \ldots \cdot \overline{\overline{A_n}} \cdot \ldots \in \beta$, что тривиально преобразуется в $A_1 \cdot \ldots \cdot A_n \cdot \ldots \in \beta$.
4. Из свойств операций над множествами можно заключить, что $A \setminus B = A \cdot \overline{B}$. По аксиоме 1, из $B \in \beta \implies \overline{B} \in \beta$. По следствию 3, $A, \overline{B} \in \beta \implies A \cdot \overline{B} \in \beta$, что, собственно, является утверждением $A \setminus B \in \beta$.

Аксиоматическое определение вероятности.
Пусть

1. Ω — пространство элементарных исходов некоторого случайного эксперимента;
2. β — сигма-алгебра, заданная на Ω.

Определение. *Вероятностью (вероятностной мерой) называется функция*

$$P \colon \beta \to \mathbb{R}$$

обладающая следующими свойствами:

1. $\forall A \in \beta \implies P(A) \geqslant 0$ *(аксиома неотрицательности);*
2. $P(\Omega) = 1$ *(аксиома нормированности);*

3. *Если $A_1, \ldots, A_n, \ldots$ — попарно несовместные события, то вероятность осуществления их суммы равна сумме вероятностей осуществления каждого из них по отдельности: $P(A_1 + \ldots + A_n + \ldots) = P(A_1) + \ldots + P(A_n) + \ldots$ (расширенная аксиома сложения).*

Замечание 1. *Аксиомы 1-3 называются аксиомами вероятности.*

1.4 Определение пространства элементарных исходов, примеры. Сформулировать определение сигма-алгебры событий. Сформулировать аксиоматическое определение вероятности и доказать простейшие свойства вероятности.

Определение. *Случайным называют эксперимент, результат которого невозможно точно предсказать заранее.*

Определение. *Множество Ω всех исходов данного случайного эксперимента называют пространством элементарных исходов.*

Замечание. *Всюду в дальнейшем будем предполагать, что*

1. *Каждый элементарный исход является «неделимым», т. е. он не может быть разбит на более «мелкие» исходы;*

2. *В результате каждого эксперимента обязательно имеет место ровно один из входящих в Ω элементарных исходов.*

Пример. *Бросают монетку. Возможные исходы: выпадение герба или решки.*
$\Omega = \{$ *Герб* , *Решка* $\}$ — *множество элементарных исходов.* $|\Omega| = 2$

Пример. *Бросают игральную кость:* $\Omega = \{$ *"1", "2", "3", "4", "5", "6"*$\}$, $|\Omega| = 6$

Пусть

1. Ω — пространство элементарных исходов, связанных с некоторым случайным экспериментом;

2. $\beta \neq \emptyset$ — система (набор) подмножеств в множестве Ω.

Определение. *β называется сигма-алгеброй*[5] *событий, если выполнены условия:*

1. *Если $A \in \beta$, то $\overline{A} \in \beta$;* [6]

[5] При ведении лекций слово «сигма» иногда заменялось на букву δ (дельта — `\delta` в LaTeX). Буква «сигма» выглядит как σ. Лектор говорит, что корректнее всего словосочетание «сигма-алгебра» вообще не сокращать и писать полностью, не используя греческие буквы. — Прим. ред.

[6] Обратите внимание, что $A \subseteq \Omega$, но $A \in \beta$, т. к. элементы множества β — подмножества из Ω. — Прим. лект.

2. *Если* $A_1, \ldots, A_n, \ldots \in \beta$, *то* $A_1 + \ldots + A_n + \ldots \in \beta$.

Аксиоматическое определение вероятности.
Пусть

1. Ω — пространство элементарных исходов некоторого случайного эксперимента;
2. β — сигма-алгебра, заданная на Ω.

Определение. *Вероятностью (вероятностной мерой) называется функция*

$$P\colon \beta \to \mathbb{R}$$

обладающая следующими свойствами:

1. $\forall A \in \beta \implies P(A) \geqslant 0$ *(аксиома неотрицательности);*
2. $P(\Omega) = 1$ *(аксиома нормированности);*
3. *Если* $A_1, \ldots, A_n, \ldots$ *— попарно несовместные события, то вероятность осуществления их суммы равна сумме вероятностей осуществления каждого из них по отдельности:* $P(A_1 + \ldots + A_n + \ldots) = P(A_1) + \ldots + P(A_n) + \ldots$ *(расширенная аксиома сложения).*

Свойства вероятностей (из аксиоматического определения):

1. $P(\overline{A}) = 1 - P(A)$;
2. $P(\emptyset) = 0$;
3. Если $A \subseteq B$, то $P(A) \leqslant P(B)$;
4. $\forall A \in \beta\colon 0 \leqslant P(A) \leqslant 1$;
5. $P(A + B) = P(A) + P(B) - P(AB)$, где $A, B \in \beta$;
6. Для любого *конечного* набора событий $A_1, \ldots, A_n$ верно

$$\begin{aligned}
&P(A_1 + \ldots + A_n) = \\
&+ \sum_{1 \leqslant i_1 \leqslant n} P(A_{i_1}) \\
&- \sum_{1 \leqslant i_1 < i_2 \leqslant n} P(A_{i_1}, A_{i_2}) \\
&+ \sum_{1 \leqslant i_1 < i_2 < i_3 \leqslant n} P(A_{i_1}, A_{i_2}, A_{i_3}) - ... + ...
\end{aligned}$$

Доказательства этих свойств:

1. По аксиомам 1, 2 сигма-алгебры $\exists A+\overline{A} = \Omega$; по аксиоме вероятности №2 $P(\Omega) = 1 = P(A+\overline{A})$; по аксиоме вероятности №3 (A и $\overline{A}$ несовместны), $P(A+\overline{A}) = P(A)+P(\overline{A}) = 1 \implies P(\overline{A}) = 1 - P(A)$.

2. $P(\emptyset) = P(\overline{\Omega})$; по свойству №1 $P(\emptyset) = 1 - \overbrace{P(\Omega)}^{=1 \text{ (по аксиоме 2)}} = 0$

3. $A \subseteq B \overset{\text{(по рисунку)}}{\Longrightarrow} B = A + (B \setminus A)$

   (см. рисунок 15)

   Тогда

$$P(B) = P(A + (B \setminus A)) =$$
$$\overset{A, B\setminus A \text{ несовместны, используем аксиому 3}}{=}$$
$$= P(A) + \overbrace{P(B \setminus A)}^{\geqslant 0 \text{ по аксиоме 1}} \geqslant P(A)$$
$$\implies P(B) \geqslant P(A)$$

4.
   (a) Неравенство $P(A) \geqslant 0$ следует из аксиомы 1.

   (b) Осталось доказать, что $P(A) \leqslant 1$.

$$\forall A \subseteq \Omega \overset{\text{по свойству 3}}{\Longrightarrow} P(A) \leqslant \overset{=1}{P(\Omega)} \implies P(A) \leqslant 1$$

5. Для любых A, B:

   (a) $A + B = A + (B \setminus A)$,

   (см. рисунок 16)

   при этом $A \cdot (B \setminus A) = \emptyset$.

   В соответствии с аксиомой 3,

$$P(A+B) = P(A) + P(B \setminus A) \tag{1}$$

(b) $B = AB + (B \setminus A)$,

(см. рисунок 17)

причём $(AB)(B \setminus A) = \emptyset$.
По аксиоме 3, имеем $P(B) = P(AB) + P(B \setminus A) \implies P(B \setminus A) = P(B) - P(AB)$. Подставим результат в 1 и получим

$$P(A + B) = P(A) + P(B) - P(AB)$$

6. Это свойство доказывать не станем. Оно является обобщением свойства 5 и может быть доказано из 5 с использованием метода математической индукции.

1.5 Сформулировать определение условной вероятности. Доказать, что при фиксированном событии B условная вероятность $P(A|B)$ обладает всеми свойствами безусловной вероятности.

Пусть

1. A и B — два события, связанные с одним случайным экспериментом;

2. Дополнительно известно, что в результате эксперимента произошло событие B.

Определение. *Условной вероятностью осуществления события A при условии, что произошло B, называется число*[7]

$$P(A \mid B) = \frac{P(AB)}{P(B)}, \; P(B) \neq 0$$

Теорема. *Пусть*

1. *Зафиксировано событие B, $P(B) \neq 0$;*

2. *$P(A \mid B)$ рассматривается как функция события A.*

Тогда $P(A \mid B)$ обладает всеми свойствами безусловной вероятности.

Доказательство. *Докажем отдельно соответствие $P(A \mid B)$ трём аксиомам вероятности и следствиям из неё.*

[7] В разговорной речи $P(A \mid B)$ читается как P от A при B. — Прим. лект.

1. *Докажем, что условная вероятность $P(A\,|\,B)$ удовлетворяет трём аксиомам вероятности:*

   (a) $P(A\,|\,B) = \dfrac{\overbrace{P(AB)}^{\geqslant 0}}{\underbrace{P(B)}_{>0}} \implies P(A|B) \geqslant 0.$

   (b) $P(\Omega\,|\,B) = \dfrac{P(\Omega B)}{P(B)} = \dfrac{P(B)}{P(B)} = 1.$

   (c)
$$P(A_1 + \ldots + A_n + \ldots\,|\,B) = \frac{P((A_1 + \ldots + A_n + \ldots)B)}{P(B)} =$$
$$= \frac{1}{P(B)} \cdot P(A_1B + A_2B + \ldots + A_nB + \ldots) \underset{A_i, A_j \text{ несовместны, } i\neq j;\ A_iB\subseteq A_i,\, A_jB\subseteq A_j \implies (A_iB)\cap(A_jB)=\emptyset, \text{ и тогда по аксиоме вероятности №3}}{=}$$
$$= \frac{1}{P(B)} \cdot [P(A_1B) + \ldots + P(A_nB) + \ldots] =$$
$$\underset{\text{(ряд)}}{=} \frac{P(A_1B)}{P(B)} + \ldots + \frac{P(A_nB)}{P(B)} + \ldots =$$
$$= P(A_1\,|\,B) + \ldots + P(A_n\,|\,B) + \ldots$$

2. *Т. к. свойства 1-6 безусловной вероятности являются прямыми следствиями из аксиом 1-3, а условная вероятность этим аксиомам удовлетворяет, то она удовлетворяет свойствам 1-6.*

1.6 Сформулировать определение условной вероятности. Доказать теорему (формулу) умножения вероятностей. Привести пример использования этой формулы.

Пусть

1. A и B — два события, связанные с одним случайным экспериментом;

2. Дополнительно известно, что в результате эксперимента произошло событие B.

Определение. *Условной вероятностью осуществления события A при условии, что произошло B, называется число*[8]

$$P(A\,|\,B) = \frac{P(AB)}{P(B)},\ P(B) \neq 0$$

[8] В разговорной речи $P(A\,|\,B)$ читается как P от A при B. — Прим. лект.

Теорема. ***Формула умножения вероятностей для двух событий.***
Пусть

1. *A, B — события;*
2. *$P(A) > 0$.*

Тогда

$$P(AB) = P(A)\,P(B\,|\,A)$$

Доказательство. *Т. к. $P(A) > 0$, то определена условная вероятность*

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

из чего напрямую следует

$$P(AB) = P(A)\,P(B\,|\,A)$$

Теорема. ***Формула умножения вероятностей для n событий***
Пусть

1. *$A_1, \ldots, A_n$ — события;*
2. *$P(A_1 \cdot \ldots \cdot A_{n-1}) > 0$.*

Тогда

$$P(A_1 \cdot A_2 \cdot \ldots \cdot A_n) = P(A_1)\,P(A_2\,|\,A_1)\,P(A_3\,|\,A_1A_2) \cdot \ldots \cdot P(A_n\,|\,A_1 \cdot \ldots \cdot A_{n-1})$$

Доказательство.

1. *Обозначив $k = \overline{1, n-1}$, имеем $A_1 \cdot \ldots \cdot A_k \supseteq A_1 \cdot \ldots \cdot A_{n-1}$.*

   По свойству 3 вероятности $P(A_1 \cdot \ldots \cdot A_k) \geqslant P(A_1 \cdot \ldots \cdot A_{n-1}) > 0$.

   Следовательно, все условные вероятности, входящие в правую часть доказываемой формулы, определены, и можно задавать условные вероятности по типу $P(A_n\,|\,A_1A_2 \ldots A_{n-1})$, и, следовательно, можно пользоваться формулой умножения вероятностей для двух событий.

2. *Последовательно применим формулу умножения вероятностей для двух событий*[9] $(P(A_{mf}B_{mf}) = P(A_{mf})P(B_{mf}\,|\,A_{mf}))$*:*

$$P(\underbrace{A_1\cdot\ldots\cdot A_{n-1}}_{A_{mf1}}\cdot\underbrace{A_n}_{B_{mf1}}) =$$

$$= P(\overbrace{\underbrace{A_1\cdot\ldots\cdot A_{n-2}}_{A_{mf2}}\cdot\underbrace{A_{n-1}}_{B_{mf2}}}^{A_{mf1}})\cdot P(\overbrace{A_n}^{B_{mf1}}\,|\,\overbrace{A_1\cdot\ldots\cdot A_{n-1}}^{A_{mf1}}) =$$

$$= P(\overbrace{\underbrace{A_1\cdot\ldots\cdot A_{n-3}}_{A_{mf3}}\cdot\underbrace{A_{n-2}}_{B_{mf3}}}^{A_{mf2}})\cdot P(\overbrace{A_{n-1}}^{B_{mf2}}\,|\,\overbrace{A_1\cdot\ldots\cdot A_{n-2}}^{A_{mf2}})\cdot P(A_n\,|\,A_1\cdot\ldots\cdot A_{n-1}) =$$

$$= \ldots =$$

$$= P(A_1)\,P(A_2\,|\,A_1)\,P(A_3\,|\,A_1A_2)\cdot\ldots\cdot P(A_n\,|\,A_1\ldots A_{n-1})$$

Пример. *На семи карточках написаны буквы слова «ШОКОЛАД». Карточки тщательно перемешивают, и по очереди извлекают случайным образом три из них без возвращения первых карточек. Найти вероятность того, что эти три карточки в порядке появления образуют слово «ШОК»:*

$$A = \{\ \textit{три карточки в порядке появления образуют слово «ШОК»}\ \}$$

Давайте введём следующие обозначения:

$$A_1 = \{\ \textit{на первой извлечённой карточке написано «Ш»}\ \}$$
$$A_2 = \{\ \textit{на второй извлечённой карточке написано «О»}\ \}$$
$$A_3 = \{\ \textit{на третьей извлечённой карточке написано «К»}\ \}$$

Тогда $A = A_1\cdot A_2\cdot A_3$.

$$P(A) = P(A_1A_2A_3) \overset{\textit{по ф-ле умножения вероятностей}}{=} P(A_1)\cdot P(A_2\,|\,A_1)\cdot P(A_3\,|\,A_1A_2)$$

Вероятность события A_1 — $\frac{1}{7}$.

Предположим, что в результате эксперимента стало доподлинно известно, что произошло событие A_1*. Тогда вероятность вытащить «О»* — $\frac{2}{6}$.

Потом стало доподлинно известно, что произошло событие A_2*. Тогда вероятность вытащить «К»* — $\frac{1}{5}$.

Тогда $A = A_1\cdot A_2\cdot A_3 = \frac{1}{7}\cdot\frac{2}{6}\cdot\frac{1}{5} = \dfrac{1}{105}$

[9] Т. к. обозначения A, B накладываются на уже используемые, то при иллюстрации применения этой формулы будем использовать индекс $_{mf}$ (multiplication formula). — Прим. ред.

1.7 Сформулировать определение пары независимых событий. Доказать критерий независимости двух событий. Сформулировать определение попарно независимых событий и событий, независимых в совокупности. Обосновать связь этих свойств.

Пусть A и B — два события, связанные с некоторым случайным экспериментом.

Определение. *События A и B называются независимыми, если $P(AB) = P(A)\,P(B)$.*

Теорема.

1. *Пусть $P(B) > 0$.*

   Утверждение «A и B — независимы» равносильно $P(A\,|\,B) = P(A)$;

2. *Пусть $P(A) > 0$.*

   Утверждение «A и B — независимы» равносильно $P(B\,|\,A) = P(B)$.

Доказательство.

1. *Сначала докажем, что если A и B — независимые, то $P(A\,|\,B) = P(A)$. По определению независимых событий, $P(AB) = P(A)P(B)$. По определению условной вероятности,*

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

   Теперь докажем обратное.

   Пусть $P(A\,|\,B) = P(A)$. Докажем, что $P(AB) = P(A)P(B)$.

   $$P(AB) \overset{\text{по ф-ле умножения вероятностей}}{=} P(B) \cdot \overset{=P(A)}{P(A\,|\,B)} = P(B)P(A)$$

2. *Доказательство второго пункта теоремы аналогично.*

Определение. *События $A_1, \ldots, A_n$ называется попарно независимыми, если* [10]

$$\forall\forall\, i \neq j;\ i,\, j \in \{1, \ldots, n\}\ P\{A_i A_j\} = P\{A_i\}P\{A_j\}$$

Определение. *События $A_1, \ldots, A_n$ называются независимыми в совокупности, если*

$$\forall\, k \in \{2, \ldots, n\}\, \forall\forall\, i_1 < i_2 < \ldots < i_k\ P\{A_{i_1}, \ldots, A_{i_k}\} = P\{A_{i_1}\} \cdot \ldots \cdot P\{A_{i_k}\}$$

[10] Обозначение $\forall\forall$ является математическим сленгом и технически некорректно. Тем не менее, это удобный способ обозначения того, что в выражении должно стоять несколько $\forall$ подряд. — Прим. лект.

Замечание. *Очевидно, что если события $A_1, \ldots, A_n$ независимы в совокупности, то они и попарно независимы. Обратное неверно.*

Пример. ***(Бернштейна)***

Рассмотрим правильный тетраэдр[11]*, на одной грани которого «написано» 1, второй — 2, третьей — 3, четвёртой — 1, 2, 3.*

Этот тетраэдр один раз подбрасывают.

Событие A_1 заключается в том, что на нижней грани «написано» 1; также введём A_2 для 2, A_3 для 3. Давайте покажем, что события A_1, A_2, A_3 попарно независимы, но не являются независимыми в совокупности.

1. *Докажем, что они независимы попарно. Т. к. $P(A_1) = \frac{1}{2}$, $P(A_2) = \frac{1}{2}$, то*

   $$P(A_1A_2) = P(A_1)\,P(A_2) = \frac{1}{4}$$

   Событие A_1A_2 означает, что на нижней грани присутствуют и 1, и 2.

   Всё аналогично для $P(A_1A_3) = P(A_1)P(A_3)$ и $P(A_2A_3) = P(A_2)P(A_3)$.

2. *Проверим равенство $P(A_1A_2A_3) = P(A_1)\,P(A_2)\,P(A_3)$, которое, казалось бы, должно равняться $\frac{1}{8}$. Но произведение событий A_1, A_2, A_3 означает, что на нижней грани присутствуют и 1, и 2, и 3, вероятность чего равна $\frac{1}{4}$.*

   И выходит, что $\frac{1}{4} \neq \frac{1}{8}$.

   Следовательно, события A_1, A_2, A_3 не являются независимыми в совокупности.

1.8 Сформулировать определение полной группы событий. Доказать теоремы о формуле полной вероятности и о формуле Байеса. Понятия априорной и апостериорной вероятностей.

Пусть Ω — пространство элементарных исходов, связанных с некоторым случайным экспериментом, а (Ω, β, P) — вероятностное пространство этого случайного эксперимента.

Определение. *Говорят. что события $H_1, \ldots, H_n \in \beta$ образуют полную группу событий, если*

1. $P(H_i) > 0,\ i = \overline{1, n}$;

2. $H_iH_j = \emptyset$ *при* $i \neq j$;

[11] Трёхмерная фигура, состоящая из четырёх треугольников. — Прим. ред.

3. $H_1 + \ldots + H_n = \Omega$.

Теорема. ***Формула полной вероятности.***

Пусть

1. $H_1, \ldots, H_n$ — *полная группа событий;*
2. $A \in \beta$ — *событие.*

Тогда (это выражение называется формулой полной вероятности):

$$P(A) = P(A \mid H_1)P(H_1) + \ldots + P(A \mid H_n)P(H_n)$$

Доказательство.

(см. рисунок 19)

1. $A = A\Omega \overset{\Omega = H_1 + \ldots + H_n}{=} A \cdot (H_1 + \ldots + H_n) = AH_1 + \ldots + AH_n$.
   Принимая $i \neq j$: $H_i \neq \emptyset$, $H_j \neq \emptyset$, *но* $(AH_i) \subseteq H_i$, $(AH_j) \subseteq H_j \implies (AH_i)(AH_j) = \emptyset$, *т. е.* AH_i *попарно не пересекаются.*
2. *Тогда*

$$P(A) = P(AH_1 + ... + AH_n) =$$
$$\overset{AH_i \text{ попарно не пересекаются}}{=}$$
$$= P(AH_1) + \ldots + P(AH_n) =$$
$$\overset{\text{т. к. } P(H_i) > 0, \text{ то } P(AH_i) = P(H_i)P(A \mid H_i)}{=}$$
$$= P(A \mid H_1)P(H_1) + \ldots + P(A \mid H_n)P(H_n)$$

Теорема. *Пусть*

1. $H_1, \ldots, H_n$ — *полная группа событий;*
2. $P(A) > 0$.

Тогда

$$P(H_i \mid A) = \frac{P(A \mid H_i)P(H_i)}{P(A \mid H_1)P(H_1) + \ldots + P(A \mid H_n)P(H_n)},\ i = \overline{1, n}$$

Доказательство.

$$P(H_i \mid A) \overset{\text{по опр. условной вероятности}}{=}$$
$$= \frac{P(AH_i)}{P(A)} \overset{\text{по ф-ле умножения в числителе, полной вероятности в знаменателе}}{=}$$
$$= P(H_i \mid A) = \frac{P(A \mid H_i)P(H_i)}{P(A \mid H_1)P(H_1) + \ldots + P(A \mid H_n)P(H_n)},\ i = \overline{1, n}$$

Вероятности $P(H_i)$, $i = \overline{1, n}$ называются априорными, т. к. они известны до опыта; Вероятности $P(H_i \mid A)$, $i = \overline{1, n}$ называются апостериорными — они вычисляются после опыта.

1.9 Сформулировать определение схемы испытаний Бернулли. Доказать формулу для вычисления вероятности реализации ровно k успехов в серии из n испытаний по схеме Бернулли. Доказать следствия этой формулы.

Давайте рассмотрим случайный эксперимент, в результате которого возможна реализация одного из двух элементарных исходов, т. е. пространство элементарных исходов у нас будет состоять из двух элементов ($|\Omega| = 2$).

Один из элементарных исходов условно будем называть успехом, второй — неудачей. Пусть p — вероятность осуществления успеха в случайном эксперименте, а q ($q = 1 - p$) — вероятность неудачи.

Определение. *Схемой испытаний Бернулли называется серия из однотипных экспериментов указанного вида, в которой отдельные испытания независимы, т. е. вероятность реализации успеха в i-ом испытании не зависит от исходов первого, второго, …, $i-1$-ого испытаний.*

Теорема. *Пусть проводится серия из n испытаний по схеме Бернулли с вероятностью успеха p. Тогда $P_n(k)$ есть вероятность того, что в серии из n испытаний произойдёт ровно k успехов:*

$$P_n(k) = C_n^k p^k q^{n-k}$$

Доказательство.

1. *Результат проведения серии из n экспериментов запишем с использованием кортежа $(x_1, \ldots, x_n)$, где*

$$x_i = \begin{cases} 1, \text{ если в испытании имел место успех;} \\ 0, \text{ если в испытании имела место неудача.} \end{cases}$$

2. *Пусть*

$$A = \{ \text{ в серии из } n \text{ испытаний произошло ровно } k \text{ успехов } \}$$

Тогда A состоит из кортежей, в которых будет ровно k единиц и $n - k$ нулей.

В событии A будет столько элементарных исходов, сколькими способами можно расставить k единиц по n позициям. Каждая такая расстановка однозначно определяется номерами позиций, в которых будут записаны единички. В остальные позиции будут записаны нули.

Выбрать k позиций из имеющихся n можно C_n^k способами. Вероятность каждого отдельного исхода равна произведению вероятностей каждого отдельного x_i, и тогда общая вероятность исхода будет равна $p^k q^{n-k}$.

Все испытания независимы; следовательно, все кортежи из A равновероятны, и их C_n^k штук, что означает

$$P_n(k) = C_n^k p^k q^{n-k}$$

Следствие. *Вероятность того, что кол-во успехов в серии из n испытаний по схеме Бернулли с вероятностью успеха p будет заключено между k_1 и k_2:*

$$P_n(k_1 \leqslant k \leqslant k_2) = \sum_{i=k_1}^{k_2} C_n^i p^i q^{n-i}$$

Доказательство.

1. *Пусть*

$$A_i = \{\text{ в серии произошло ровно } i \text{ успехов }\},\ i = \overline{k_1, k_2}$$
$$P(A_i) = P_n(i) = C_n^i p^i q^{n-i}$$

2.

$$A = A_{k_1} + A_{k_1+1} + \ldots + A_{k_2} = P_n(k_1 \leqslant k \leqslant k_2)$$
$$\Longrightarrow$$
$$P(A) = (A_{k_1} + \ldots + A_{k_2}) \overset{A_i \text{ и } A_j \text{ несовместны при } i \neq j}{=} P(A_{k_1}) + \ldots + P(A_{k_2}) =$$
$$= \sum_{i=k_1}^{k_2} P\{A_i\} = \sum_{i=k_1}^{k_2} C_n^i p^i q^{n-i}$$

Следствие. *Вероятность того, что в серии испытаний Бернулли с вероятностью успеха p (и неудачи $q = 1 - p$) произойдёт хотя бы один успех: $P_n(k \geqslant 1) = 1 - q^n$.*

Доказательство. *Пусть $A = \{$ в серии произошёл хотя бы один успех $\}$. В таком случае $\overline{A} = \{$ в серии не будет ни одного успеха $\}$, и тогда*

$$P(A) = 1 - P(\overline{A}) = 1 - P_n(0) = 1 - C_0^i p^0 q^{n-0} = 1 - q^n$$

2 Случайные величины

2.1 Сформулировать определение случайной величины и функции распределения вероятностей случайной величины. Доказать свойства функции распределения.

Пусть $(\Omega,\, \beta,\, P)$ — вероятностное пространство некоторого случайного эксперимента.

Определение. *Случайной величиной называется функция*

$$X\colon \Omega \to \mathbb{R}$$

такая, что для каждого $x \in \mathbb{R}$ множество $\{\omega : X(\omega) < x\} \in \beta$ (т. е. для любого x множество $\{\omega\colon X(\omega) < x\}$ является событием).

Пусть X — случайная величина, связанная с некоторым случайным экспериментом.

Определение. *Функцией распределения вероятностей случайной величины X называется отображение*

$$F_X\colon \mathbb{R} \to \mathbb{R}$$

определённое следующим правилом[12]

$$F_X(x) = P\{X < x\}$$

Свойства функции распределения:

1. $0 \leqslant F(x) \leqslant 1$;
2. F является неубывающей функцией, т. е. если $x_1 \leqslant x_2$, то $F(x_1) \leqslant F(x_2)$;
3. $$\lim_{x\to-\infty} F(x) = 0$$ $$\lim_{x\to+\infty} F(x) = 1$$
4. В каждой точке функция распределения непрерывна слева[13]: $\lim\limits_{x\to x_0-} F(x) = F(x_0)$;
5. $P\{a \leqslant X < b\} = F(b) - F(a)$.

Доказательства.

1. *$F(x)$ определена как вероятность, т. е. $F(x) = P\{...\} \in [0;\, 1]$.*

---

[12] С тем же успехом можно определить F_X как $F_X(x) = P\{X \leqslant x\}$; выбор обусловлен конвенцией (при этом, видимо, в международной литературе используют именно знак $\leqslant$). От того, как определена эта функция, в основном зависят выборы знаков ($<$ или $\leqslant$) в определённых местах в дальнейших формулах, но список отличий этим не ограничивается. — Прим. ред.

[13] Если бы F_X была определена как $F_X = P\{X \leqslant x\}$, то функция распределения была бы непрерывна справа. В учебнике по Теории Вероятностей из серии «Математика в техническом университете» (с римской цифрой на обложке, далее книги из серии будут адресоваться по номерам; для Теории Вероятностей — учебник XVI) написано (издание третье, исправленное), что на этом отличия в свойствах заканчиваются; это, судя по всему, ошибка, т. к. при использовании $\leqslant$ свойство 5 должно записываться как $P\{a < X \leqslant b\} = F(b) - F(a)$. — Прим. ред.

2. *Имея $x_1 \leqslant x_2$, выразим $F(x_2)$:*

$$F(x_2) = P\{X < x_2\} = P\{\underbrace{\{X < x_1\}}_{\text{Событие } A} + \underbrace{\{x_1 \leqslant X < x_2\}}_{\text{Событие } B}\}$$

События A и B несовместны, т. е.

$$F(x_2) = \underbrace{P\{X < x_1\}}_{F(x_1)} + \underbrace{P\{x_1 \leqslant X < x_2\}}_{\geqslant 0} \geqslant F(x_1)$$

3. *Сначала докажем, что $\lim\limits_{x \to +\infty} F(x) = 1$:*

   Рассмотрим последовательность x_1, x_2, x_3, … такую, что

   (a) $x_1 < x_2 < x_3 < \ldots$;

   (b) $\lim\limits_{n \to \infty} x_n = +\infty$.

   Обозначим $A_n = \{X < x_n\}$, $n \in N$; очевидно, что последовательность событий A_n имеет свойство $A_1 \subseteq A_2 \subseteq \ldots \subseteq A_n \subseteq A_{n+1} \subseteq \ldots$, т. е. эта последовательность является неубывающей последовательностью событий.

   Тогда, применяя аксиому непрерывности

$$\lim_{x \to \infty} F(x_n) = \lim_{n \to \infty} P(A_n) \overset{\text{аксиома непрерывности}}{=} \underbrace{P\{X < +\infty\}}_{\text{(достоверное событие)}} = 1$$

   Т. к. x_1, x_2, … — произвольная последовательность (неубывающая и стремящаяся к бесконечности), то в соответствии с определением предела функции по Гейне[14]

$$\lim_{x \to +\infty} F(x) = 1$$

   Другая часть этого свойства, $\lim\limits_{x \to -\infty} F(x) = 0$, доказывается аналогично.

4. *Пусть x_1, x_2, … — возрастающая последовательность такая, что $\lim\limits_{n \to \infty} x_n = x_0$.*

   Пусть $A_i = \{X < x_i\}$, $i \in \mathbb{N}$. Тогда событие

$$\{X < x_n\} = \bigcup_{i=1}^{\infty} A_i$$

   причём последовательность событий A_1, A_2, … является возрастающей;

[14] Значение A называется пределом функции $f(x)$ в точке x_0, если для любой последовательности точек $\{x_n\}_{n=1}^{\infty}$, сходящейся к x_0 (см. понятие предела *последовательности*, не функции), но не содержащей x_0 в качестве одного из своих элементов, последовательность значений функции $\{f(x_n)\}_{n=1}^{\infty}$ сходится к A. — Прим. ред.

$$\lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} P(A_n) \overset{\textit{аксиома непрерывности}}{=} P\{X < x_0\} = F(x_0)$$

Т. к. x_1, x_2, ... — произвольная последовательность, сходящаяся к x_0 слева, то в соответствии с определением предела функции по Гейне

$$\lim_{x\to x_0-} F(x) = F(x_0)$$

5. *$\{X < b\} = \{X < a\} + \{a \leqslant X < b\}$; события в объединении несовместные, поэтому*

$$\underbrace{P\{X < b\}}_{F(B)} = \underbrace{P\{X < a\}}_{F(A)} + P\{a \leqslant X < b\}$$

из чего тривиально следует

$$P\{a \leqslant X < b\} = F(b) - F(a)$$

2.2 Сформулировать определения случайной величины и функции распределения случайной величины. Сформулировать определения дискретной и непрерывной случайной величины. Доказать свойства плотности распределения вероятностей непрерывной случайной величины.

Пусть (Ω, β, P) — вероятностное пространство некоторого случайного эксперимента.

Определение. *Случайной величиной называется функция*

$$X\colon \Omega \to \mathbb{R}$$

такая, что для каждого $x \in \mathbb{R}$ множество $\{\omega : X(\omega) < x\} \in \beta$ (т. е. для любого x множество $\{\omega\colon X(\omega) < x\}$ является событием).

Пусть X — случайная величина, связанная с некоторым случайным экспериментом.

Определение. *Функцией распределения вероятностей случайной величины X называется отображение*

$$F_X\colon \mathbb{R} \to \mathbb{R}$$

определённое следующим правилом[15]

$$F_X(x) = P\{X < x\}$$

[15] С тем же успехом можно определить F_X как $F_X(x) = P\{X \leqslant x\}$; выбор обусловлен конвенцией (при этом, видимо, в международной литературе используют именно знак $\leqslant$). От того, как определена эта функция, в основном зависят выборы знаков ($<$ или $\leqslant$) в определённых местах в дальнейших формулах, но список отличий этим не ограничивается. — Прим. ред.

Определение. *Случайная величина называется дискретной, если множество её значений конечно или счётно.*

Определение. *Случайная величина X называется непрерывной, если существует функция*

$$f(x)\colon \mathbb{R} \to \mathbb{R}$$

такая, что $\forall x \in \mathbb{R}$ *функция* $F(x) = \int\limits_{-\infty}^{x} f(t)dt$ *(F — функция распределения в X).*

Замечание 1. *При этом f называют функцией плотности распределения вероятности случайной величины X.*

Свойства непрерывных случайных величин[16].

1. $f(x) \geqslant 0$;

2. $$P\{a \leqslant X < b\} = \int\limits_{a}^{b} f(x)dx$$

   где X — непрерывная случайная величина, а f – её функция плотности[17];

3. Условие нормировки:

   $$\int\limits_{-\infty}^{+\infty} f(x)\,dx = 1$$

   где f — функция плотности некоторой случайной величины;

4. $$P\{x_0 \leqslant X < x_0 + \triangle x\} \approx f(x_0)\triangle x$$

   где X — непрерывная случайная величина, f — её функция плотности, x_0 — точка непрерывности функции f, а $\triangle x$ — мало;

5. Если X — непрерывная случайная величина, то для любого наперёд заданного $x_0 \in \mathbb{R}$

   $$P\{X = x_0\} = 0$$

Доказательства.

1. Почти всюду $f(x) = F'(x) \overset{F\text{ — неубывающая}}{\geqslant} 0.$

[16] Свойств именно функции плотности непрерывных случайных величин «в чистом виде» в лекциях нет. — Прим. ред.

[17] Чуть позже по лекциям утверждается, что в этом свойстве не важно, какие знаки стоят в условии при P, т. е. $P\{a \leqslant X < b\} = P\{a < X < b\} = P\{a < X \leqslant b\} = P\{a \leqslant X \leqslant b\} = P\{a \preccurlyeq X \preccurlyeq b\} = \int\limits_{a}^{b} f(x)dx$. — Прим. ред.

2. *По свойству функции распределения*

$$P\{x \leqslant X < b\} = F(b) - F(a)$$

Т. к. F — первообразная для f, то по формуле Ньютона-Лейбница

$$P\{x \leqslant X < b\} = F(b) - F(a) = \int\limits_a^b f(x)dx$$

3.

$$\int\limits_{-\infty}^{+\infty} f(x)\,dx = \lim_{x_1 \to -\infty,\ x_2 \to +\infty} \int\limits_{x_1}^{x_2} f(x)\,dx \overset{\text{свойство 2}}{=} \lim_{x_2 \to +\infty} F(x_2) - \lim_{x_1 \to -\infty} F(x_1) =$$

$$= \cancelto{1}{F(+\infty)} - \cancelto{0}{F(-\infty)} = 1$$

4.

$$P\{x_0 \leqslant X < x_0 + \triangle x\} \overset{\text{свойство 2}}{=} F(x_0 + \triangle x) - F(x_0)$$

Т. к. x_0 — точка непрерывности f, а $\triangle x$ мало, то можно считать, что в окрестности $(x_0,\, x_0 + \triangle x)$ функция $F' = f$ непрерывна. Тогда применим к функции f на $[x_0,\, x_0 + \triangle x]$ теорему Лагранжа[18] [19]*:*

$$F(x_0 + \triangle x) - F(x_0) = \underbrace{F'(\xi)}_{f(\xi)} \triangle x$$

где $\xi \in (x_0,\, x_0 + \triangle x)$. Т. к. $\triangle x$ мало, а f непрерывна в некоторой окрестности x_0, то можно считать, что $f(\xi) \approx f(x_0)$. Таким образом,

$$P\{x_0 \leqslant X < x_0 + \triangle x\} \approx f(x_0)\triangle x$$

5.

$$P\{X = x_0\} = \lim_{\triangle x \to 0} P\{x_0 \leqslant X < x_0 + \triangle x\} =$$

$$\overset{\text{свойство 2}}{=}$$

$$= \lim_{\triangle x \to 0} [F(x_0 + \triangle x) - F(x_0)] \overset{F \text{ непрерывна, см. замечание выше}}{=} 0$$

[18] Пусть функция $y = f(x)$ непрерывна на отрезке $[a,\, b]$ и дифференцируема в интервале $(a,\, b)$; тогда между точками a и b найдётся хотя бы одна такая точка c $(a < c < b)$, для которой справедливо равенство $f(b) - f(a) = f'(c)(b - a)$. — Прим. ред.

[19] ξ — строчная буква «кси» греческого алфавита. — Прим. ред.

2.3 Сформулировать определение нормальной случайной величины, указать геометрический смысл параметров. Понятие стандартного нормального закона. Доказать формулу для вычисления вероятности попадания нормальной случайной величины в интервал.

Определение. *Говорят, что случайная величина X имеет нормальное распределение с параметрами m и σ^2 ($\sigma > 0$), если её функция плотности имеет вид*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\dfrac{(x-m)^2}{2\sigma^2}},\ x \in \mathbb{R}$$

Обозначается $X \sim N(m,\, \sigma^2)$.

Функция плотности нормального распределения имеет характерную колоколообразную форму; m является координатой x «центра» этого колокола (центра симметрии), а σ характеризует разброс значений случайной величины; чем меньше σ, тем выше экстремум функции плотности[20].

Замечание 1.

Определение. *Распределение $N(0,\, 1)$ называет стандартным нормальным распределением; для него функция плотности равна*

$$f_{0,\,1}(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}},\ x \in \mathbb{R}$$

Часто для вычисления вероятностей из стандартного нормального распределения рассматривают функцию

$$\Phi(x) = F_{0,\,1}(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt$$

$\Phi(x)$ называют функцией Лапласа; для нахождения её значений используйте заранее высчитанные таблицы[21].

Замечание 2. *Часто вместо функции $\Phi(x)$ удобнее рассмотреть функцию*

$$\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{0}^{x} e^{-\frac{t^2}{2}} dt$$

[20] В статье «Нормальное распределение» в Википедии присутствуют неплохие иллюстрации. — Прим. ред.

[21] В конце учебника XVI (приложение П.3) находится таблица значений для функции $\Phi_0(x)$ (см. следующее замечание). Лектор говорит, что в домашнем задании в обязательном порядке необходимо посчитать вероятности при помощи этой функции; на контрольную или экзамен можно принести распечатку таблицы значений, но там высчитывание конечного значения вероятности не обязательно. — Прим. ред.

(см. рисунок 28)

Свойства функций Φ *и* Φ_0*:*

1. $\Phi(x) = \frac{1}{2} + \Phi_0(x)$;
2. $\Phi_0(x) = -\Phi_0(x)$ *(функция чётная);*
3. $\Phi_0(+\infty) = \frac{1}{2}$;
4. $\Phi_0(-\infty) = -\frac{1}{2}$.

Замечание 3. *Пусть* $x \sim N(m,\, \sigma^2)$; *чему равно* $P\{a \preccurlyeq X \preccurlyeq b\}$*?*
Рассмотрим

$$P\{a \leqslant X < b\} = \langle \text{ по свойству плотности распределения } \rangle =$$

$$= \int\limits_a^b f_X(x)\,dx = \frac{1}{\sigma\sqrt{2\pi}} \int\limits_a^b e^{-\frac{(x-m)^2}{2\sigma^2}} =$$

$$= \left\langle t = \frac{x-m}{\sigma};\ dx = \sigma\,dt;\ \begin{cases} x = a \implies t = \frac{a-m}{\sigma}, \\ x = b \implies t = \frac{b-m}{\sigma} \end{cases} \right\rangle =$$

$$= \frac{1}{\cancel{\sigma}\sqrt{2\pi}} \cdot \cancel{\sigma} \cdot \int\limits_{\frac{a-m}{\sigma}}^{\frac{b-m}{\sigma}} e^{-\frac{t^2}{2}}\,dt = \langle \text{ т. к. } \Phi(t) \text{ — первообразная } f_{0,1}(t) \ \rangle =$$

$$= \Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right) =$$

$$= \left\langle \Phi(t) \equiv \frac{1}{2} + \Phi_0(t) \right\rangle =$$

$$= \Phi_0\left(\frac{b-m}{\sigma}\right) - \Phi_0\left(\frac{a-m}{\sigma}\right)$$

Т. к. $X \sim N(m,\, \sigma^2)$ *— непрерывная случайная величина, то*

$$P\{a \preccurlyeq X \preccurlyeq b\} = \Phi_0\left(\frac{b-m}{\sigma}\right) - \Phi_0\left(\frac{a-m}{\sigma}\right) = \Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right)$$

2.4 Сформулировать определение случайного вектора и функции распределения вероятностей случайного вектора. Сформулировать свойства функции распределения двумерного случайного вектора. Доказать предельные свойства.

Пусть

1. $(\Omega,\, \beta,\, P)$ — вероятностное пространство;
2. $X_\omega = X_1(\omega),\, \ldots,\, X_n(\omega)$ — случайные величины, заданные на этом вероятностном пространстве.

Определение. *n-мерным случайным вектором называется кортеж*[22]

$$\overrightarrow{X} = (X_1,\, \ldots,\, X_n)$$

Определение. *Функцией распределения вероятностей случайного вектора*

$$X = (X_1,\, \ldots,\, X_n)$$

называется отображение

$$F\colon \mathbb{R}^n \to \mathbb{R}$$

определённое правилом

$$F(x_1,\, \ldots,\, x_n) = P\{X_1 < x_1,\, \ldots,\, X_n < x_n\}$$

Свойства функции распределения случайного вектора (для n = 2):

1. $0 \leqslant F(x_1,\, x_2) \leqslant 1$;
2. (a) При фиксированном x_2 функция $F(x_1,\, x_2)$ как функция переменного x_1 является неубывающей функцией;

   (b) При фиксированном x_1 функция $F(x_1,\, x_2)$ как функция переменного x_2 является неубывающей функцией.
3. $\lim\limits_{x_1 \to -\infty,\, x_2 = const} F(x_1,\, x_2) = 0$

   $\lim\limits_{x_1 = const,\, x_2 \to -\infty} F(x_1,\, x_2) = 0$
4. $\lim\limits_{x_1 \to +\infty,\, x_2 \to +\infty} F(x_1,\, x_2) = 1$
5. $\lim\limits_{x_1 = const,\, x_2 \to +\infty} F(x_1,\, x_2) = F_{X_1}(x_1)$

   $\lim\limits_{x_1 \to +\infty,\, x_2 = const} F(x_1,\, x_2) = F_{X_2}(x_2)$

   где $F_{X_i}(x_i)$ — функция распределения случайной величины X_i;
6. Вероятность того, что реализация попадёт в похожую на прямоугольник область $D = \{(x,\, y)\colon x \in [a_1,\, b_1),\, y \in [a_2,\, b_2)\}$:

   $P\{a_1 \leqslant X < b_1,\, a_2 \leqslant X_2 < b_2\} = F(b_1,\, b_2) - F(a_1,\, b_2) - F(b_1,\, a_2) + F(a_1,\, a_2)$

[22] Обратите внимание, что векторы обозначаются стрелочкой ($\overrightarrow{X}$), а не прямой ($\overline{X}$). Это важно, т. к. далее в курсе появится величина, которая будет обозначаться прямой. За использование прямой для обозначения вектора будут снижаться баллы. — Прим. лект.

7. (a) При фиксированном x_2 функция $F(x_1, x_2)$ как функция переменной x_1 является непрерывной слева в каждой точке;

   (b) При фиксированном x_1 функция $F(x_1, x_2)$ как функция переменной x_2 является непрерывной слева в каждой точке.

Доказательства.

1. ...

2. ...

3. *Покажем, что* $\lim\limits_{x_1 \to -\infty,\, x_2 = const} F(x_1, x_2) = 0.$

По определению, $F(x_1, x_2) = P\{\{X_1 < x_1\} \cdot \{X_2 < x_2\}\}$*; при* $x_1 \to -\infty$ *событие* $\{X_1 < -\infty\}$ *является невозможным. Произведение невозможного события на событие* $\{X_2 < x_2\}$ *является невозможным событием, поэтому* $F(x_1, x_2)$ *стремится к нулю при* $x_1 \to -\infty$, $x_2 = const$.

$\lim\limits_{x_1 = const,\, x_2 \to -\infty} F(x_1, x_2) = 0$ *доказывается аналогично.*

4. *По определению,* $F(x_1, x_2) = P\{\{X_1 < x_1\} \cdot \{X_2 < x_2\}\}$.

Событие $\{X_1 < +\infty\}$ *является достоверным,* $\{X_2 < +\infty\}$ *также является достоверным, а произведение достоверных событий — достоверное событие; таким образом,*

$$\lim_{x_1 \to +\infty,\, x_2 \to +\infty} F(x_1, x_2) = 1$$

5. *Покажем, что* $\lim\limits_{x_1 = const,\, x_2 \to +\infty} F(x_1, x_2) = F_{X_1}(x_1)$

По определению,

$$F(x_1, x_2) = P\{\{X_1 < x_1\} \cdot \{X_2 < x_2\}\}$$

Событие $\{X_2 < +\infty\}$ *является достоверным; произведение события* $\{X_1 < x_1\}$ *на достоверное равно* $\{X_1 < x_1\}$ *(т. е. равно ему же), поэтому*

$$\lim_{x_1 = const,\, x_2 \to +\infty} F(x_1, x_2) = P\{X_1 < x_1\} = F_{X_1}(x_1)$$

$\lim\limits_{x_1 \to +\infty,\, x_2 = const} F(x_1, x_2) = F_{X_2}(x_2)$ *доказывается аналогично.*

6. ...

7. *Доказывается аналогично одномерному случаю.*

2.5 Сформулировать определение случайного вектора и функции распределения вероятностей случайного вектора. Сформулировать свойства функции распределения двумерного случайного вектора. Доказать формулу для вычисления $P\{a_1 \leqslant X_1 < b_1,\, a_2 \leqslant X_2 < b_2\}$.

Пусть

1. $(\Omega,\, \beta,\, P)$ — вероятностное пространство;
2. $X_\omega = X_1(\omega),\, \ldots,\, X_n(\omega)$ — случайные величины, заданные на этом вероятностном пространстве.

Определение. *n-мерным случайным вектором называется кортеж*[23]

$$\overrightarrow{X} = (X_1,\, \ldots,\, X_n)$$

Определение. *Функцией распределения вероятностей случайного вектора*

$$X = (X_1,\, \ldots,\, X_n)$$

называется отображение

$$F\colon \mathbb{R}^n \to \mathbb{R}$$

определённое правилом

$$F(x_1,\, \ldots,\, x_n) = P\{X_1 < x_1,\, \ldots,\, X_n < x_n\}$$

Свойства функции распределения случайного вектора (для n = 2):

1. $0 \leqslant F(x_1,\, x_2) \leqslant 1$;
2. (a) При фиксированном x_2 функция $F(x_1,\, x_2)$ как функция переменного x_1 является неубывающей функцией;

   (b) При фиксированном x_1 функция $F(x_1,\, x_2)$ как функция переменного x_2 является неубывающей функцией.
3. $\lim\limits_{x_1 \to -\infty,\, x_2 = const} F(x_1,\, x_2) = 0$

   $\lim\limits_{x_1 = const,\, x_2 \to -\infty} F(x_1,\, x_2) = 0$
4. $\lim\limits_{x_1 \to +\infty,\, x_2 \to +\infty} F(x_1,\, x_2) = 1$

[23] Обратите внимание, что векторы обозначаются стрелочкой ($\overrightarrow{X}$), а не прямой ($\overline{X}$). Это важно, т. к. далее в курсе появится величина, которая будет обозначаться прямой. За использование прямой для обозначения вектора будут снижаться баллы. — Прим. лект.

5. $\lim\limits_{x_1 = const,\, x_2 \to +\infty} F(x_1,\, x_2) = F_{X_1}(x_1)$

   $\lim\limits_{x_1 \to +\infty,\, x_2 = const} F(x_1,\, x_2) = F_{X_2}(x_2)$

   где $F_{X_i}(x_i)$ — функция распределения случайной величины X_i;

6. Вероятность того, что реализация попадёт в похожую на прямоугольник область $D = \{(x,\, y) \colon x \in [a_1,\, b_1),\, y \in [a_2,\, b_2)\}$:

   $P\{a_1 \leqslant X < b_1,\, a_2 \leqslant X_2 < b_2\} = F(b_1,\, b_2) - F(a_1,\, b_2) - F(b_1,\, a_2) + F(a_1,\, a_2)$

7. (a) При фиксированном x_2 функция $F(x_1,\, x_2)$ как функция переменной x_1 является непрерывной слева в каждой точке;

   (b) При фиксированном x_1 функция $F(x_1,\, x_2)$ как функция переменной x_2 является непрерывной слева в каждой точке.

Доказательство. ***(формулы для вычисления*** $P\{a_1 \leqslant X_1 < b_1,\, a_2 \leqslant X_2 < b_2\}$***)***

1. Найдём вероятность попадания случайного вектора $(X_1,\, X_2)$ в полосу

$$\{X_1 < x_1, a_2 \leqslant X < b_2\}$$

(a) $\{X_1 < x_1,\, X_2 < b_2\} = \{X_1 < x_1,\, a_2 \leqslant X_2 < b_2\} + \{X_1 < x_1,\, X_2 < a_2\}$

(b) По теореме сложения (события объединения несовместны):

$$\underbrace{P\{X_1 < x_1,\, X_2 < b_2\}}_{F(x_1, b_2)} = P\{X_1 < x_1,\, a_2 \leqslant X_2 < b_2\} + \underbrace{P\{X_1 < x_1,\, X_2 < a_2\}}_{F(x_1, a_2)}$$

Таким образом,

$$P\{X_1 < x_1,\, a_2 \leqslant X_2 < b_2\} = F(x_1,\, b_2) - F(x_1,\, a_2)$$

2. (a)

$$\begin{aligned}\{X_1 < b_1,\, a_2 \leqslant X_2 < b_2\} =& \{a_1 \leqslant X_1 < b_1,\, a_2 \leqslant X_2 < b_2\} + \\ & \{X_1 < a_1,\, a_2 \leqslant X_2 < b_2\}\end{aligned}$$

(b) По формуле сложения (события объединения несовместны):

$$\begin{aligned}\underbrace{P\{X_1 < b_1,\, a_2 \leqslant X_2 < b_2\}}_{\text{(из пункта а) } F(b_1, b_2) - F(b_1, a_2)} =& P\{a_1 \leqslant X_1 < b_1,\, a_2 \leqslant X_2 < b_2\} + \\ & \underbrace{P\{X_1 < a_1,\, a_2 \leqslant X_2 < b_2\}}_{\text{(из пункта а) } F(a_1, b_2) - F(a_1, a_2)}\end{aligned}$$

Таким образом,

$$P\{a_1 \leqslant X_1 < b_1,\, a_2 \leqslant X_2 < b_2\} = F(b_1,\, b_2) - F(b_1,\, a_2) - F(a_1,\, b_2) + F(a_1,\, a_2)$$

2.6 Сформулировать определение случайного вектора и функции распределения вероятностей случайного вектора. Сформулировать определение непрерывного случайного вектора и доказать свойства плотности распределения вероятностей для двумерного случайного вектора.

Пусть

1. $(\Omega,\, \beta,\, P)$ — вероятностное пространство;
2. $X_\omega = X_1(\omega),\, \ldots,\, X_n(\omega)$ — случайные величины, заданные на этом вероятностном пространстве.

Определение. *n-мерным случайным вектором называется кортеж*[24]

$$\overrightarrow{X} = (X_1,\, \ldots,\, X_n)$$

Определение. *Функцией распределения вероятностей случайного вектора*

$$X = (X_1,\, \ldots,\, X_n)$$

называется отображение

$$F\colon \mathbb{R}^n \to \mathbb{R}$$

определённое правилом

$$F(x_1,\, \ldots,\, x_n) = P\{X_1 < x_1,\, \ldots,\, X_n < x_n\}$$

Определение. *Случайный вектор $(X_1,\, \ldots,\, X_n)$ называется непрерывным, если существует функция*

$$f\colon \mathbb{R}^n \to \mathbb{R}$$

такая, что для каждой точки $(x_1,\, \ldots,\, x_n)$ выполняется

$$F(x_1,\, \ldots,\, x_n) = \int\limits_{-\infty}^{x_1} dt_1 \int\limits_{-\infty}^{x_2} dt_2 \ldots \int\limits_{-\infty}^{x_i} dt_i \ldots \int\limits_{-\infty}^{x_n} f(t_1,\, \ldots,\, t_n)\, dt_n$$

где F — функция распределения плотности случайного вектора $(X_1,\, \ldots,\, X_n)$. При этом f называется функцией плотности распределения вероятностей этого вектора.

Замечание 1. *В определении предполагается. что несобственный интеграл сходится в каждой точке $(x_1,\, \ldots,\, x_n) \in \mathbb{R}^n$.*

[24] Обратите внимание, что векторы обозначаются стрелочкой ($\overrightarrow{X}$), а не прямой ($\overline{X}$). Это важно, т. к. далее в курсе появится величина, которая будет обозначаться прямой. За использование прямой для обозначения вектора будут снижаться баллы. — Прим. лект.

Замечание 2. *При* $n = 2$*:*

$$F(x_1,\, x_2) = \int\limits_{-\infty}^{x_1} dt_1 \int\limits_{-\infty}^{x_2} f(t_1,\, t_2)\, dt_2$$

Свойства непрерывных случайных векторов (для n=2)Свойств именно функции плотности непрерывных случайных векторов «в чистом виде» в лекциях нет. — Прим. ред..

1. Если f — функция плотности двумерного случайного вектора, то $f(x_1,\, x_2) \geqslant 0$.

2. Если $(X_1,\, X_2)$ — непрерывный случайный вектор, а $f(x_1,\, x_2)$ — его функция плотности, то

$$P\{a_1 \preccurlyeq X_1 \preccurlyeq b_1,\, a_2 \preccurlyeq X_2 \preccurlyeq b_2\} = \int\limits_{a_1}^{b_1} dx_1 \int\limits_{a_2}^{b_2} f(x_1,\, x_2)\, dx_2$$

3. Условие нормировки: $\iint\limits_{R^2} f(x_1,\, x_2)\, dx_1\, dx_2 = 1$.

4. Если $f(x_1,\, x_2)$ — функция плотности вектора (X_1, X_2), а (x_1^0, x_2^0) — точка непрерывности функции f, то

$$P\{x_1^0 \preccurlyeq X_1 \preccurlyeq x_1^0 + \triangle x_1, x_2^0 \preccurlyeq X_2 \preccurlyeq x_2^0 + \triangle x_2\} \approx f(x_1^0,\, x_2^0) \triangle x_1 \triangle x_2$$

   если $\triangle x_1$, $\triangle x_2$ достаточно малы.

5. Если $(X_1,\, X_2)$ — непрерывный случайный вектор, то для любых наперёд заданных x_1^0, x_2^0

$$P\{(X_1,\, X_2) = (x_1^0,\, x_2^0)\} = 0$$

6.
$$P\{(X,\, Y) \in D\} = \iint\limits_{D} f(x_1,\, x_2)\, dx_1\, dx_2$$

7.
$$\int\limits_{-\infty}^{+\infty} f(x_1,\, x_2)\, dx_2 = f_{X_1}(x_1)$$
$$\int\limits_{-\infty}^{+\infty} f(x_1,\, x_2)\, dx_1 = f_{X_2}(x_2)$$

где f_{X_1}, f_{X_2} — маргинальные функции плотности случайных величин X_1 и X_2, f — совместная функция плотности случайных величин X_1 и X_2 ($\equiv$ функция плотности случайного вектора $(X_1,\, X_2)$).

Обратите внимание, что из функции плотности можно получить обе маргинальные.

$$f(x_1,\, x_2) \implies \begin{cases} f_{X_1}(x_1) \\ f_{X_2}(x_2) \end{cases}$$

Доказательства. *Доказательства свойств 1-5 аналогичны одномерному случаю. Свойство 6 является обобщением свойства 2 на случай произвольной области D (без доказательства).*

Доказательство свойства 7.

Докажем, что $f_{X_1}(x_1) = \int\limits_{-\infty}^{+\infty} f(x_1,\, x_2)\, dx_2$.

По свойству двумерной функции распределения $F(x_1,\, +\infty) = F_{X_1}(x_1)$; таким образом (подставим определение функции распределения для двумерного вектора),

$$F_{X_1}(x_1) = \int\limits_{-\infty}^{x_1} dt_1 \int\limits_{-\infty}^{+\infty} f(t_1,\, t_2)\, dt_2$$

$$f_{X_1}(x_1) = \frac{dF_{X_1}(x_1)}{dx_1} =$$

$$= \Big\langle x_1 \text{ — } \textit{точка непрерывности функции } f_{X_1}(x_1), \textit{ и тогда по теореме}$$

$$\textit{о производной интеграла с перменным верхним пределом} \Big\rangle =$$

$$= \int\limits_{-\infty}^{+\infty} f(x_1,\, t_2)\, dt_2$$

Вторая формула доказывается аналогично.

2.7 Сформулировать определение пары независимых случайных величин. Доказать свойства независимых случайных величин. Понятия попарно независимых случайных величин и случайных величин, независимых в совокупности.

Определение. *Случайные величины X и Y называются независимыми, если*

$$F(x,\, y) = F_X(x)\, F_Y(y)$$

где F — совместная функция распределения X и Y ($\equiv$ функция распределения случайного вектора $(X,\,Y)$); F_X, F_Y — маргинальные функции распределения случайных величин X и Y.

Свойства независимых случайных величин:

1.
$$\text{Случайные величины } X \text{ и } Y \text{ независимы}$$
$$\Longleftrightarrow$$
$$\forall\forall\, x,\, y \in \mathbb{R} \text{ события } \{X < x\} \text{ и } \{Y < y\} \text{ независимы.}$$

2.
$$\text{Случайные величины } X \text{ и } Y \text{ независимы}$$
$$\Longleftrightarrow$$
$$\forall\forall x_1,\, x_2 \in \mathbb{R},\, \forall\forall y_1,\, y_2 \in \mathbb{R} \text{ события } \{x_1 \leqslant X < x_2\} \text{ и } \{y_1 \leqslant Y < y_2\} \text{ независимы.}$$

3.
$$\text{Случайные величины } X \text{ и } Y \text{ независимы}$$
$$\Longleftrightarrow$$
$$\forall\forall M_1,\, M_2 \text{ события } \{X \in M_1\} \text{ и } \{Y \in M_2\} \text{ независимы,}$$

   где M_1, M_2 — промежутки или объединения промежутков в $\mathbb{R}$.

4. Если X и Y — дискретные случайные величины, то
$$X,\, Y \text{ — независимые } \iff p_{ij} \equiv P_{X_i} P_{Y_j}$$
   где $p_{ij} = P\{(X,\,Y) = (x_i,\,y_j)\}$, $P_{X_i} = P\{X = x_i\}$, $P_{Y_j} = P\{Y = y_j\}$.

5. Если X и Y — непрерывные случайные величины, то
$$X,\, Y \text{ — независимы } \iff f(x,\,y) \equiv f_X(x)\, f_Y(y)$$
   где f — совместная плотность распределения случайных величин X и Y ($\equiv$ функция плотности распределения случайного вектора $(X,\,Y)$); f_X, f_Y — маргинальные плотности распределения случайных величин X и Y соответственно.

Доказательства.

1. *Очевидно следует из определения независимых случайных величин.*

2. (a) ***Необходимость ($\Longrightarrow$).***
   Пусть $F(x,\, y) = F_X(x)\, F_Y(y)$.
   Тогда

$$
\begin{gathered}
P\{x_1 \leqslant X < x_2,\, y_1 \leqslant Y < y_2\} = \\
= \langle \text{ свойство функции распределения случайного вектора } \rangle = \\
= F(x_1,\, y_1) + F(x_2,\, y_2) - F(x_1,\, y_2) - F(x_2,\, y_1) = \\
= F_X(x_1)\, F_Y(y_1) + F_X(x_2)\, F_Y(y_2) - F_X(x_1)\, F_Y(y_2) - F_X(x_2)\, F_Y(y_1) = \\
= [F_X(x_2) - F_X(x_1)]\, [F_Y(y_2) - F_Y(y_1)] = \\
= \langle \text{ свойство одномерной функции распределения } \rangle = \\
= P\{x_1 \leqslant X < x_2\}\, P\{y_1 \leqslant Y < y_2\}
\end{gathered}
$$

   (b) ***Достаточность ($\Longleftarrow$).***
   Пусть $\forall\forall x_1,\, x_2,\, y_1,\, y_2 \in \mathbb{R}$,

$$
P\{x_1 \leqslant X < x_2,\, y_1 \leqslant Y < y_2\} = P\{x_1 \leqslant X < x_2\}\, P\{y_1 \leqslant Y < y_2\}
$$

   Тогда

$$
\begin{gathered}
F(x,\, y) = P\{X < x,\, Y < y\} = P\{-\infty < X < x,\, -\infty < Y < y\} = \\
= \langle x_1 = -\infty,\, x_2 = x,\; y_1 = -\infty,\, y_2 = y \rangle = \\
= P\{-\infty < X < x\}\, P\{-\infty < Y < y\} = F_X(x)\, F_Y(y)
\end{gathered}
$$

3. *Является обобщением свойств 1 и 2 (без доказательства).*

4. (a) ***Достаточность ($\Longleftarrow$).*** *Достаточность была доказана выше, в рассуждениях перед определением независимых случайных величин.*

   (b) ***Необходимость ($\Longrightarrow$).*** *Необходимость студентам предлагается доказать самостоятельно.*

5. (a) ***Необходимость ($\Longrightarrow$).***
   Пусть $F(x,\, y) \equiv F_X(x)\, F_Y(y)$. *По свойству двумерной плоскости*

$$
f(x,\, y) = \frac{\delta^2 F(x,\, y)}{\delta x\, \delta y} = \frac{\delta^2}{\delta x\, \delta y}\, [F_X(x)\, F_Y(y)] = \frac{dF_X(x)}{dx} \cdot \frac{dF_Y(y)}{dy} = f_X(x)\, f_Y(y)
$$

   (b) ***Достаточность ($\Longleftarrow$).***
   Пусть $f(x,\, y) = f_X(x)\, f_Y(y)$.

Тогда

$$F(x,\,y) = \int\limits_{-\infty}^{x} dt \int\limits_{-\infty}^{y} f(t,\,v)\,dv = \int\limits_{-\infty}^{x} dt \int\limits_{-\infty}^{y} f_X(t)\,f_Y(v)\,dv =$$

$$\underbrace{\int\limits_{-\infty}^{x} f_X(t)\,dt}_{F_X(x)} \underbrace{\int\limits_{-\infty}^{y} f_Y(v)dv}_{F_Y(y)} = F_X(x)\,F_Y(y)$$

Определение. *Случайные величины $X_1,\,\ldots,\,X_n$, заданные на одном вероятностном пространстве, называются*

- *Попарно независимыми, если X_i и X_j независимы при $i \neq j$;*
- *Независимыми в совокупности, если*

$$F(x_1,\,\ldots,\,x_n) = F_{X_1}(x_1)\cdot\ldots\cdot F_{X_n}(x_n)$$

где F — совместная функция распределения случайных величин $X_1,\,\ldots\,X_n$ ($\equiv$ функция распределения случайного вектора $(X_1,\,\ldots\,X_n)$), F_{X_i} — маргинальные функции распределения случайных величин X_i, $i = \overline{1;\,n}$.

Замечание 1. *Можно доказать, что*

1. *Если $X_1,\,\ldots,\,X_n$ независимы в совокупности, то они попарно независимы. Обратное неверно.*
2. *Обобщения свойств 4 и 5 будут справедливы для любого числа n случайных величин, независимых в совокупности. К примеру, обобщение свойства 5:*

$$X_1,\,\ldots,\,X_n \text{ — независимы в совокупности}$$

$$\Longleftrightarrow$$

$$f(x_1,\,\ldots,\,x_n) = f_{X_1}(x_1)\cdot\ldots\cdot f_{X_n}(x_n)$$

2.8 Понятие условного распределения случайной величины. Сформулировать определение условного ряда распределения компоненты двумерного дискретного случайного вектора. Привести рассуждения, приводящие к такому определению. Сформулировать определение условной плотности распределения компоненты двумерного непрерывного случайного вектора. Сформулировать критерии независимости случайных величин в терминах условных распределений.

Давайте рассмотрим случайный вектор (X, Y). Предположим, известно, что случайная величина Y приняла значение y_0. Что в этом случае можно сказать о возможных значениях случайной величины X и что можно сказать о законе распределения случайной величины X при условии $Y = y_0$?

Случай дискретного случайного вектора Пусть

1. (X, Y) — дискретный случайный вектор;
2. $X \in \{x_1, \dots, x_m\}$, $Y \in \{y_1, \dots, y_n\}$;
3. $p_{ij} = P\{(X, Y) = (x_i, y_j)\}$, $i = \overline{1;\, m}$, $j = \overline{1;\, n}$,
   $P_{X_i} = P\{X = x_i\}$, $i = \overline{1;\, m}$,
   $P_{Y_j} = P\{Y = y_j\}$, $j = \overline{1;\, n}$;
4. Известно, что $Y = y_j$ для некоторого фиксированного j.

Тогда

$$P\{X = x_i \,|\, Y = y_j\} = \langle \text{ из определения условной вероятности } \rangle = \\ = \frac{P\{\{X = x_i\} \cdot \{Y = y_j\}\}}{P\{Y = y_j\}} = \frac{P\{(X, Y) = (x_i, y_j)\}}{P_{Y_j}} = \frac{p_{ij}}{P_{Y_j}}$$

Определение. *В случае двумерного дискретного случайного вектора (X, Y) условной вероятностью того, что случайная величина X приняла значение x_i при условии $Y = y_j$ называют число*

$$\pi_{ij} = \frac{p_{ij}}{P_{Y_j}}$$

Определение. *Набор вероятностей π_{ij}, $i = \overline{1;\, m}$, для данного фиксированного j называются уловным распределением случайного вектора X при условии $Y = y_j$.*

Замечание 1. *Условная вероятность того, что случайная величина Y приняла значение y_j при условии $X = x_i$ определяется аналогично*[25]*:*

$$\tau_{ij} = P\{Y = y_j \,|\, X = x_i\} = \frac{p_{ij}}{P_{X_i}}$$

Замечание 2. *Набор вероятностей τ_{ij}, $j = \overline{1;\, n}$, для фиксированного i называют условным распределением случайной величины Y при условии $X = x_i$.*

Случай непрерывного случайного вектора

В случае непрерывного случайного вектора $(X,\, Y)$ рассуждения, аналогичные проведённым для дискретного случайного вектора рассуждениям, приводят к следующему определению.

Определение. *Условной плотностью распределения случайного вектора X при условии $Y = y$ называется функция*

$$f_X(x \,|\, Y = y) = \frac{f(x,\, y)}{f_Y(y)}$$

где f — совместная плотность распределения случайных величин X и Y ($\equiv$ плотность распределения вектора $(X,\, Y)$), f_Y — маргинальная плотность распределения случайной величины Y.

Замечание. *Аналогичным образом определяется условная плотность распределения случайной величины Y при условии $X = x$:*

$$f_Y(y \,|\, X = x) = \frac{f(x,\, y)}{f_X(x)}$$

где $f_X(x)$ — маргинальная плотность распределения случайной величины X.

Теорема. ***Критерии независимости случайных величин в терминах условных распределений.***

1. *Пусть $(X,\, Y)$ — двумерный случайный вектор. Тогда следующие условия эквивалентны:*
   - *$X,\, Y$ — независимые.*
   - *$F_X(x \,|\, Y = y) \equiv F_X(x)$ для всех y, в которых определена $F_X(x \,|\, Y = y)$.*
   - *$F_Y(y \,|\, X = x) \equiv F_Y(y)$ для всех x, в которых определена $F_Y(y \,|\, X = x)$.*
2. *Если $(X,\, Y)$ — непрерывный случайный вектор, то следующие условия эквивалентны. (для условной плотности)*

[25] τ — строчная буква «тау» греческого алфавита. — Прим. ред.

- *X, Y — независимые.*
- *$f_X(x \mid Y=y) \equiv f_X(x)$ для всех y, в которых определена $f_X(x|Y=y)$.*
- *$f_Y(y \mid X=x) \equiv f_Y(y)$ для всех x, в которых определена $f_Y(y|X=x)$.*

3. *Если (X, Y) — дискретный случайный вектор, то следующие утверждения эквиваленты.*
   - *X, Y — независимые.*
   - *$P\{X=x_i \mid Y=y_j\} \equiv P\{X=x_i\}$ для всех $j=\overline{1;\, n}$.*
   - *$P\{Y=y_j \mid X=x_i\} \equiv P\{Y=y_j\}$ для всех $i=\overline{1;\, m}$.*

   (здесь $X \in \{x_1, \ldots, x_m\}$, $Y \in \{y_1, \ldots, y_n\}$).

Доказательство. *Без доказательства.*

2.9 Понятие функции скалярной случайной величины. Доказать теорему о формуле для вычисления плотности $f_Y(y)$ случайной величины $Y=\phi(X)$, если X – непрерывная случайная величина, а φ – монотонная непрерывно дифференцируемая функция. Сформулировать аналогичную теорему для кусочно-монотонной функции φ.

Пусть

1. X — некоторая случайная величина;
2. $\varphi\colon \mathbb{R} \to \mathbb{R}$ — некоторая известная функция.

   Тогда $\varphi(X) = Y$ — некоторая случайная величина.

Теорема. *Пусть*

1. *X — непрерывная случайная величина;*
2. *$\varphi\colon \mathbb{R} \to \mathbb{R}$;*
3. *φ монотонна и непрерывно дифференцируема;*
4. *ψ — функция*[26]*, обратная к φ (т. к. φ — монотонная, то $\exists \psi = \varphi^{-1}$);*
5. *$Y = \varphi(X)$.*

   Тогда

[26] ψ — строчная буква «пси» греческого алфавита. — Прим. ред.

1. *Y также является непрерывной случайной величиной;*

2. $f_Y(y) = f_X(\psi(y))|\psi'(y)|$.

Доказательство.

1. $F_Y(y) = P\{Y < y\} = P\{\varphi(X) < y\}$

   (a) *Если φ — монотонно возрастающая функция, то $\varphi(X) < y \iff X < \varphi^{-1}(y) = \psi(y)$;*

   (b) *Если φ — монотонная убывающая функция, то $\varphi(X) < y \iff X > \varphi^{-1}(y) = \psi(y)$*

2. *В случае случая а,* $F_Y(y) = P\{X < \psi(y)\} = F_X(\psi(y))$; *в случае б* $F_Y(y) = P\{X > \varphi(y)\} = 1 - P\{X \preccurlyeq \psi(y)\} = \langle$ *X — непрерывная* $\rangle = 1 - P\{X < \psi(y)\} = 1 - F_X(\psi(y))$.

3.

$$f_Y(y) = F'_Y(y) = \begin{cases} \dfrac{d}{dy}[F_X(\psi(y))] = F'_X(\psi(y)) \cdot \psi'(y), & \text{если а} \\ \dfrac{d}{dy}[1 - F_X(\psi(y))] = -F'_X(\psi(y)) \cdot \psi'(y), & \text{если б} \end{cases}$$
$$= f_X(\psi(y)) \cdot |\psi'(y)|$$

Теорема. *Пусть*

1. *X — непрерывная случайная величина;*

2. *$\varphi\colon \mathbb{R} \to \mathbb{R}$ является кусочно-монотонной функцией, имеющей n интервалов монотонности;*

3. *φ дифференцируема;*

4. *Для данного $y \in R$, $x_1 = x_1(y), \ldots, x_k = x_k(y)$ $(k \leqslant n)$ — это все решения уравнения $y = \varphi(x)$, принадлежащие интервалам $I_1, \ldots, I_k$ монотонности функции φ.*

   Тогда для данного в условии 4 значения y

$$f_Y(y) = \sum_{j=1}^{k} f_X(\psi_j(y)) * |\psi'_j(y)|$$

   где $\psi_j(y)$ — функция, обратная к $\varphi(x)$ на интервале I_j, $j = \overline{1; k}$

2.10 Понятие скалярной функции случайного вектора. Обосновать формулу для вычисления функции распределения случайной величины Y, функционально зависящей от случайных величин X_1 и X_2, если (X_1, X_2) – непрерывный случайный вектор. Доказать теорему о формуле свёртки.

Пусть

1. (X_1, X_2) — двумерный случайный вектор;
2. $\varphi : \mathbb{R}^2 \to \mathbb{R}$
3. $Y = \varphi(X_1, X_2)$ — некоторая одномерная случайная величина.

Как, зная закон распределения случайного вектора (X_1, X_2), найти закон распределения случайной величины Y?

Рассмотрим два случая. **Случай дискретного случайного вектора:** Пусть (X_1, X_2) — дискретный случайный вектор. В таком случае Y — дискретная случайная величина.

Случай непрерывного случайного вектора:

Если (X_1, X_2) — непрерывный случайный вектор, то функцию распределения случайной величины $Y = \varphi(X_1, X_2)$ можно найти по формуле

$$F_Y(y) = \iint\limits_{D(y)} f(x_1, x_2)\, dx_1\, dx_2$$

где f — совместная плотность распределения случайных величин X_1 и X_2, $D(y) = \{(x_1, x_2)\colon \varphi(x_1, x_2) < y\}$.

Доказательство.

$$\begin{gathered} F_Y(y) = P\{Y < y\} = \\ = \langle\ \textit{события } \{Y < y\} \textit{ и } \{(X_1, X_2) \in D(y)\} \textit{ эквивалентны}\ \rangle = \\ = \langle\ \textit{свойство непрерывного случайного вектора}\ \rangle = \\ = \iint\limits_{D(y)} f(x_1, x_2)\, dx_1\, dx_2 \end{gathered}$$

Формула свёртки.

Теорема. *Пусть*

1. *X_1, X_2 — непрерывные случайные величины;*
2. *X_1, X_2 — независимые случайные величины;*

3. $Y = X_1 + X_2$ *(т. е.* $\varphi(x_1, x_2) = x_1 + x_2$*).*

Тогда

$$f_Y(y) = \int\limits_{-\infty}^{+\infty} f_{X_2}(y - x_1)\, f_{X_1}(x_1)\, dx_1$$

Доказательство.

1.

$$F_Y(y) = P\{Y < y\} = P\{X_1 + X_2 < y\} = \iint\limits_{D(y)} f(x_1,\, x_2)\, dx_1\, dx_2 =$$

$$= \left\langle X_1,\, X_2 - \textit{независимые} \implies f(x_1,\, x_2) = f_{X_1}(x_1) f_{X_2}(x_2) \right\rangle =$$

$$= \int\limits_{-\infty}^{+\infty} dx_1 \int\limits_{-\infty}^{y-x_1} f_{X_1}(x_1) f_{X_2}(x_2)\, dx_2 = \int\limits_{-\infty}^{+\infty} f_{X_1}(x_1) \left[F_{X_2}(x_2)\Big|_{-\infty}^{y-x_1} \right] dx_1 =$$

$$= \langle F_{X_2}(-\infty) = 0 \rangle = \int\limits_{-\infty}^{+\infty} f_{X_1}(x_1) F_{X_2}(y - x_1)\, dx_1$$

2.

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \int\limits_{-\infty}^{+\infty} f_{X_1}(x_1) f_{X_2}(y - x_1)\, dx_1$$

2.11 Сформулировать определение математического ожидания для дискретной и непрерывной случайных величин. Механический смысл математического ожидания. Доказать свойства математического ожидания. Записать формулы для вычисления математического ожидания функции случайной величины и случайного вектора.

Определение. *Математическим ожиданием дискретной величины* X *называют число*

$$M[X] = \sum_i p_i x_i$$

где i *пробегает такое множество значений, что* x_i *исчерпывает все возможные значения случайной величины* X*;* $p_i = P\{X = x_i\}$*.*

Определение. *Математическим ожиданием непрерывной случайной величины X называется число*[27]

$$M[X] = \int\limits_{-\infty}^{+\infty} x\, f(x)\, dx$$

где $f(x)$ — функция плотности случайной величины X.

Замечание. ***Механический смысл математического ожидания.***

Будем интерпретировать величину p_i как «вероятностную массу» значения x_i случайной величины X. Т. к. $\sum\limits_i p_i = 1$, то $M[X]$ характеризует положение центра тяжести вероятностной массы.

Свойства математического ожидания:

1. Если $P\{X = x_0\} = 1$ (т. е. если X фактически не является случайной), то $MX = x_0$;
2. $M[aX + b] = a \cdot MX + b$;
3. $M[X_1 + X_2] = MX_1 + MX_2$;
4. Если X_1, X_2 — независимы, то $M[X_1 X_2] = (MX_1)(MX_2)$.

Доказательства.

1. *Ряд распределения:*

$$\begin{array}{cc} X & x_0 \\ P & 1 \end{array}$$

По определению: $MX = \sum\limits_i p_i x_i = 1 \cdot x_0 = x_0$.

2. *Докажем для случая непрерывной случайной величины.*

$$M[aX + b] = \langle \varphi(x) = ax + b;\ aX + b \text{ — непрерывная случайная величина } \rangle =$$
$$= \int\limits_{-\infty}^{+\infty} (ax + b)\, f(x)\, dx = a \int\limits_{-\infty}^{+\infty} x\, f(x)\, dx + b \int\limits_{-\infty}^{+\infty} f(x)\, dx = a \cdot MX + b.$$

[27]Если аргумент достаточно простой, то квадратные скобки в обозначении $M[X]$ часто опускают. Также часто MX обозначают просто буквой m. — Прим. ред.

3. *Доказательство для дискретного случая. Элементы X_1 обозначаются индексами i, пробегающими множество I; для X_2 используются j и J. Запись о:*

$$
\begin{gathered}
M[X_1 + X_2] = \langle \varphi(x_1, x_2) = x_1 + x_2 \rangle = \\
= \sum_{i \in I} \sum_{j \in J} (x_{1,i} + x_{2,j}) p_{ij} = \sum_{i \in I} \sum_{j \in J} x_{1,i}\, p_{ij} + \sum_{i \in I} \sum_{j \in J} x_{2,j}\, p_{ij} = \\
= \sum_{i \in I} x_{1,i} \underbrace{\sum_{j \in J} p_{ij}}_{P\{X_1 = X_{1,i}\}} + \sum_{j \in J} x_{2,j} \underbrace{\sum_{i \in I} p_{ij}}_{P\{X_2 = x_{2,j}\}} = MX_1 + MX_2
\end{gathered}
$$

4. *Докажем для непрерывных случайных величин*

$$
\begin{gathered}
M[X_1 X_2] = \langle \varphi(x_1, x_2) = x_1 x_2 \rangle = \iint\limits_{R^2} x_1\, x_2\, f(x_1, x_2)\, dx_1\, dx_2 = \\
= \langle X_1, X_2 \text{ — независимы } \implies f(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2) \rangle = \\
= \int\limits_{-\infty}^{+\infty} dx_1 \int\limits_{-\infty}^{+\infty} x_1\, x_2\, f_{X_1}(x_1)\, f_{X_2}(x_2)\, dx_2 = \\
\left(\int\limits_{-\infty}^{+\infty} x_1\, f_{X_1}(x_1)\, dx_1 \right) \left(\int\limits_{-\infty}^{+\infty} x_2\, f_{X_2}(x_2)\, dx_2 \right) = (MX_1)(MX_2)
\end{gathered}
$$

Рубрика «Сделай сам». *Доказательства для дискретного случая свойств 2 и 4, непрерывного случая свойства 3 предлагается написать самостоятельно.*

Замечание.

1. *Пусть X — случайная величина, а $\varphi \colon \mathbb{R} \to \mathbb{R}$ — некоторая функция, тогда*

   (a) *Если X — дискретная случайная величина, то*

   $$M[\varphi(X)] = \sum_i \varphi(x_i)\, p_i$$

   (b) *Если X — непрерывная случайная величина, то*

   $$M[\varphi(X)] = \int\limits_{-\infty}^{+\infty} \varphi(x)\, f(x)\, dx$$

2. *Если (X, Y) — случайный вектор, $\varphi \colon \mathbb{R}^2 \to \mathbb{R}$, — функция, то*

(a) *Если* (X, Y) *— дискретный случайный вектор, то*

$$M[\varphi(X, Y)] = \sum_{i,j} \varphi(x_i, y_j)\, p_{ij}$$

где $p_{ij} = P\{(X, Y) = (x_i, y_j)\}$.

(b) *Если* (X, Y) *— непрерывный случайный вектор, то*

$$M[\varphi(X, Y)] = \iint\limits_{R^2} \varphi(x, y)\, f(x, y)\, dx\, dy$$

где f *— совместная плотность распределения* X *и* Y.

2.12 Сформулировать определение дисперсии случайной величины. Механический смысл дисперсии. Доказать свойства дисперсии. Понятие среднеквадратичного отклонения случайной величины.

Определение. *Дисперсией случайной величины* X *называют число*

$$D[X] = M[(X - m)^2]$$

где $m = MX$.

Замечание 1. *Если* X *— дискретная случайная величина, то*

$$DX = \langle DX = M[(X - m)^2],\ \varphi(x) = (x - m)^2 \rangle = \sum_i (x_i - m)^2 p_i$$

где $p_i = P\{X = x_i\}$.

Замечание 2. *Если* X *— непрерывная случайная величина, то*

$$D[X] = \int\limits_{-\infty}^{+\infty} (x - m)^2 f(x)\, dx$$

где f *— функция плотности случайной величины* X.

Замечание. ***Механический смысл дисперсии.***

Дисперсия случайной величины характеризует разброс значений этой случайной величины относительно математического ожидания. Чем больше дисперсия, тем больше разброс значений.

С точки зрения механики дисперсия — момент инерции вероятностной массы относительно математического ожидания.

Свойства дисперсии:

1. $DX \geqslant 0$;
2. Если $P\{X = x_0\} = 1$, то $DX = 0$.
3. $D[aX + b] = a^2 DX$;
4. $D[X] = M[X^2] - (MX)^2$;
5. Если X_1, X_2 — независимы, то $D[X_1 + X_2] = DX_1 + DX_2$.

Доказательства.

1. $DX = MY$, где $Y = (X - m)^2$. Т. к. $Y \geqslant 0$, то следует, что $DX = MY \geqslant 0$.

2.

$$\begin{array}{cc} X & x_0 \\ P & 1 \end{array}$$

Математическое ожидание $MX = m = x_0$.

Дисперсия $DX = \sum_i (x_i - m)^2 p_i = (x_0 - x_0)^2 \cdot 1 = 0$.

3.

$$D[aX + b] = M[[(aX + b) - M(aX + b)]^2] = M[[aX + b - a \cdot MX - b]^2] = \\ = M[(a(X - MX))^2] = a^2 M[(X - MX)^2] = a^2 DX$$

4. Обозначим $m = MX$; тогда

$$DX = M[(X - m)^2] = M[X^2 - 2mX + m^2] = M[X^2] - 2\underbrace{m \cdot M[X]}_{m^2} + m^2 = \\ = M[X^2] - m^2$$

5.

$$D(X_1 + X_2) = \langle \text{ по свойству 4 } \rangle = M[(X_1 + X_2)^2] - (M(X_1 + X_2))^2 = \\ = M[X_1^2] + M[X_2^2] + 2\,M[X_1 X_2] - (MX_1)^2 - (MX_2)^2 - 2\,MX_1 \cdot MX_2 = \\ = \langle X_1,\, X_2 \text{ — независимые, тогда } M[X_1\, X_2] = MX_1 \cdot MX_2 \rangle = \\ = (M[X_1^2] - (MX_1)^2) + (M[X_2^2] - (MX_2)^2) = DX_1 + DX_2$$

Замечание. *DX имеет размерность, равную квадрату размерности случайной величины X. Это не всегда удобно, особенно при решении практических задач. Поэтому рассматривают такую числовую характеристику, как среднеквадратичное отклонение (СКО).*

Определение. *Среднеквадратичным отклонением (СКО) случайной величины X называют число*

$$\sigma_X = \sqrt{D[X]}$$

2.13 Сформулировать определение математического ожидания и дисперсии. Записать законы распределения биномиальной, пуассоновской, равномерной, экспоненциальной и нормальной случайной величин. Найти математические ожидания и дисперсии этих случайных величин.

Определение. *Математическим ожиданием дискретной величины X называют число*

$$M[X] = \sum_i p_i x_i$$

где i пробегает такое множество значений, что x_i исчерпывает все возможные значения случайной величины X; $p_i = P\{X = x_i\}$.

Определение. *Математическим ожиданием непрерывной случайной величины X называется число*[28]

$$M[X] = \int\limits_{-\infty}^{+\infty} x\, f(x)\, dx$$

где $f(x)$ — функция плотности случайной величины X.

Определение. *Дисперсией случайной величины X называют число*

$$D[X] = M[(X-m)^2]$$

где $m = MX$.

Замечание 1. *Если X — дискретная случайная величина, то*

$$DX = \langle DX = M[(X-m)^2],\ \varphi(x) = (x-m)^2 \rangle = \sum_i (x_i - m)^2 p_i$$

где $p_i = P\{X = x_i\}$.

Замечание 2. *Если X — непрерывная случайная величина, то*

$$D[X] = \int\limits_{-\infty}^{+\infty} (x-m)^2 f(x)\, dx$$

где f — функция плотности случайной величины X.

[28]Если аргумент достаточно простой, то квадратные скобки в обозначении $M[X]$ часто опускают. Также часто MX обозначают просто буквой m. — Прим. ред.

Биномиальная случайная величина

Обозначение $X \sim B(n, p)$, X — кол-во успехов в серии из n испытаний по схеме Бернулли с вероятностью успеха p. Рассмотрим случайную величину X_i, $i = \overline{1;\, n}$:

$$X_i = \begin{cases} 1, & \text{если в } i\text{-ом испытании произошёл успех,} \\ 0, & \text{если в } i\text{-ом имела место неудача} \end{cases}$$

Случайные величины $X_1, \ldots, X_n$ независимы, т. к. отдельные испытания в схеме Бернулли независимы.

Каждое $X_i \sim B(1, p)$, $i = \overline{1;\, n}$. Ранее было показано, что $MX_i = p$, $DX_i = pq$. Тогда

$$X = \sum_{i=1}^{n} X_i$$

$$MX = M\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} MX_i = np$$

$$DX = D\left[\sum_{i=1}^{n} X_i\right] = \langle X_i \text{ независимы} \rangle = \sum_{i=1}^{n} DX_i = npq$$

Пуассоновская случайная величина $X \sim \Pi(\lambda)$

$P\{X = k\} = \dfrac{\lambda^k}{k!} e^{-\lambda}$, $k = 0, 1, 2, \ldots$

Тогда математическое ожидание выражается

$$MX = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=1}^{\infty} k \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}\lambda}{(k-1)!} =$$
$$= \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \langle i = k-1 \rangle = \lambda e^{-\lambda} \underbrace{\sum_{i=0}^{\infty} \frac{\lambda^i}{i!}}_{\text{формула Маклорена для } e^{\lambda}} = \lambda$$

Дисперсия выражается как $DX = M[X^2] - (MX)^2$.

$MX^2 = \sum\limits_{k=0}^{\infty} k^2 \dfrac{\lambda^k}{k!} e^{-\lambda} = \ldots = \lambda^2 + \lambda$. Тогда $DX = \lambda$.

Случайная величина X, имеющая геометрическое распределение с параметром p

$P\{X = k\} = pq^k,\ k = 0,\, 1,\, 2,\, \ldots$

$$MX = \sum_{k=0}^{\infty} k \cdot pq^k = \sum_{k=1}^{\infty} k \cdot pq^k = pq \underbrace{\sum_{k=1}^{\infty} kq^{k-1}}_{=1+2q+3q^2+4q^3+\ldots} =$$

$=\Big\langle$ продифференцируем сумму бесконечно убывающей геометрической прогрессии:

$$(1 + q + q^2 + \ldots)' = \left(\frac{1}{1-q}\right)' \implies 1 + 2q + 3q^2 + 4q^3 + \ldots = \frac{1}{(1-q)^2} = \frac{1}{p^2}\Big\rangle =$$

$$= pq\frac{1}{p^2} = \frac{q}{p}$$

Аналогичным образом можно показать, что $DX = \dfrac{q}{p^2}$

Равномерно распределённая случайная величина $X \sim R(0,\, 1)$

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & x \in (a,\, b) \\ 0 & \text{иначе} \end{cases}$$

$$MX = \int\limits_{-\infty}^{\infty} x\, f(x)\, dx = \frac{1}{b-a} \int\limits_{a}^{b} x\, dx = \frac{1}{2(b-a)}(b^2 - a^2) = \frac{a+b}{2}$$

$$DX = \int\limits_{-\infty}^{+\infty} \left(x - \frac{a+b}{2}\right)^2 f(x)\, dx = \frac{1}{b-a} \int\limits_{a}^{b} \left(x - \frac{a+b}{2}\right)^2 dx =$$

$$= \frac{1}{3(b-a)} \left(x - \frac{a+b}{2}\right)^3 \Bigg|_a^b = \frac{1}{3(b-a)} \left(\frac{(b-a)^3}{8} - \frac{(a-b)^3}{8}\right) =$$

$$= \frac{(b-a)^3}{12(b-a)} = \frac{(b-a)^2}{12}$$

Экспоненциальная случайная величина $X \sim Exp(\lambda)$

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & x < 0 \end{cases}$$

$$MX = \int\limits_{-\infty}^{+\infty} x\, f(x) x = \lambda \int\limits_{0}^{+\infty} xe^{-\lambda x} dx = -\int\limits_{0}^{+\infty} x de^{-\lambda x} = \cancel{-xe^{-\lambda x}\Big|_0^{+\infty}} \to 0 + \int\limits_{0}^{+\infty} e^{-\lambda x} dx =$$

$$= -\frac{1}{\lambda} e^{-\lambda x}\Bigg|_0^{+\infty} = \frac{1}{\lambda}$$

$DX = M[X^2] - (MX)^2,$

$$M[X^2] = \int\limits_{-\infty}^{+\infty} x^2 f(x)\,dx = \lambda \int\limits_{0}^{+\infty} x^2 e^{-\lambda x} dx = -\int\limits_{0}^{+\infty} x^2 d e^{-\lambda x} =$$

$$= \cancelto{0}{-x^2 e^{-\lambda x}\Big|_0^{+\infty}} + 2\underbrace{\int_0^{+\infty} x e^{-\lambda x} dx}_{=\frac{MX}{\lambda}} = \frac{2}{\lambda^2}$$

Таким образом,

$$DX = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

Нормальная случайная величина $X \sim N(\underbrace{m}_{MX}, \underbrace{\sigma^2}_{DX})$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}},\ x \in \mathbb{R}$$

$$MX = \int\limits_{-\infty}^{+\infty} x f(x)\,dx = \frac{1}{\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} x e^{-\frac{(x-m)^2}{2\sigma^2}}\,dx = \langle x - m = t \rangle =$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} (t+m) e^{-\frac{t^2}{2\sigma^2}}\,dt =$$

$$= \cancelto{0 \text{ (нечётная функция)}}{\frac{1}{\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} t e^{-\frac{t^2}{2\sigma^2}}\,dt} + m \cdot \cancelto{1 \text{ (условие нормировки)}}{\frac{1}{\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} e^{-\frac{t^2}{2\sigma^2}}\,dt} = m$$

$$DX = M[(X - MX)^2] = \int\limits_{-\infty}^{+\infty} (x-m)^2 f(x)\,dx = \frac{1}{\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} (x-m)^2 e^{-\frac{(x-m)^2}{2\sigma^2}}\,dx =$$

$$= \left\langle \frac{x-m}{\sigma} = t,\ dx = \sigma dt \right\rangle =$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} t^2 e^{-\frac{t^2}{2}}\,dt = \frac{\sigma^2}{2\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} t e^{-\frac{t^2}{2}}\,dt^2 = -\frac{\sigma^2}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} t\, de^{-\frac{t^2}{2}}\,dt = \langle \text{ по частям } \rangle =$$

$$= \cancelto{0}{-\frac{\sigma^2}{\sqrt{2\pi}} t e^{-e\frac{t^2}{2}}\Big|_{-\infty}^{+\infty}} + \sigma^2 \cdot \cancelto{1 \text{ (усл. норм. } N(0,\,1))}{\frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} e^{-\frac{t^2}{2}}\,dt} = \sigma^2$$

2.14 Сформулировать определение ковариации и записать формулы для ее вычисления в случае дискретного и непрерывного случайных векторов. Доказать свойства ковариации.

Ковариация является характеристикой случайного вектора.
Пусть (X, Y) — двумерный случайный вектор.

Определение. *Ковариацией случайных величин X и Y называется число*

$$\operatorname{cov}(X, Y) = M[(X - m_X)(Y - m_Y)]$$

где $m_X = MX$, $m_Y = MY$.

Замечание. *Из определения следует:*

1. *В случае дискретного случайного вектора*

$$\operatorname{cov}(X, Y) = \sum_i \sum_j (x_i - m_X)(y_j - m_Y)p_{ij}$$

   где $p_{ij} = P\{(X, Y) = (x_i, y_j)\}$.

2. *В случае непрерывного случайного вектора*

$$\operatorname{cov}(X, Y) = \iint\limits_{R^2} (x - m_X)(y - m_Y)f(x, y)\,dx\,dy$$

   где f — совместная плотность величин X и Y.

Свойства ковариации:

1. $D(X + Y) = DX + DY + 2\operatorname{cov}(X,Y)$;
2. $\operatorname{cov}(X, X) = DX$;
3. Если X, Y — независимые, то $\operatorname{cov}(X, Y) = 0$;
4. $\operatorname{cov}(a_1X + b_1, a_2Y + b_2) = a_1a_2\operatorname{cov}(X, Y)$
5. $|\operatorname{cov}(X, Y)| \leqslant \sqrt{DX \cdot DY}$, причём $|\operatorname{cov}(X, Y)| = \sqrt{DX \cdot DY} \iff \exists\exists a, b \in \mathbb{R}$, $Y = aX + b$ (т. е. X и Y связаны линейной зависимостью);
6. $\operatorname{cov}(X, Y) = M[XY] - (MX)(MY)$.

Доказательства.

1.

$$\begin{gathered}D(X+Y) = M[((X+Y) - M[X+Y])^2] = \langle MX = m_1,\, MY = m_2\rangle = \\ = M[((X-m_1) - M[Y-m_2])^2] = \\ = \underbrace{M[(X-m_1)^2]}_{=DX} + \underbrace{M[(Y-m_2)^2]}_{=DY} + 2\underbrace{M[(X-m_1)(Y-m_2)]}_{\operatorname{cov}(X,Y)} = \\ = DX + DY + 2\operatorname{cov}(X,\, Y)\end{gathered}$$

2. $\operatorname{cov}(X,\, X) = M[(X-m)(X-m)] = M[(X-m)^2] = DX$.

3.

$$\begin{gathered}\operatorname{cov}(X,\, Y) = M[(X-m_1)(Y-m_2)] = \\ = \langle X,\, Y \text{ — независимы} \implies (X-m_1) \text{ и } (Y-m_2) \text{ тоже независимы}\rangle = \\ = [\cancelto{0}{M(X-m_1)}][\cancelto{0}{M(Y-m_2)}] = 0\end{gathered}$$

4.

$$\begin{gathered}\operatorname{cov}(a_1X + b_1,\, a_2X + b_2) = \\ = M[[a_1X + b_1 - M(a_1X + b_1)] \cdot [a_2X + b_2 - M(a_2X + b_2)]] = \\ = M[[a_1X + \cancel{b_1} - a_1m_1 - \cancel{b_1}][a_2X + \cancel{b_2} - a_2m_2 - \cancel{b_2}]] = \\ = M[a_1a_2(X_1 - MX_1)(X_2 - MX_2)]\end{gathered}$$

5. (a) *Выберем произвольное число $t \in R$. Рассмотрим случайную величину $Z(t) = tX - Y$.*

   Тогда $D[Z(t)] = D[tX - Y] = \langle$ свойство 1 + свойство дисперсии $\rangle = t^2DX + DY - 2t\operatorname{cov}(X,\, Y) = DX \cdot t^2 - 2t \cdot \operatorname{cov}(X,\, Y) + DY$ — квадратный трёхчлен относительно t.

   Т. к. $D[Z(t)] \geqslant 0$, следовательно, трёхчлен должен быть параболой вверх, следовательно — дискриминант $D \leqslant 0$.

$$\frac{D}{4} = (\operatorname{cov}(X,\, Y))^2 - DX \cdot DY \leqslant 0 \implies |\operatorname{cov}(X,\, Y)| \leqslant \sqrt{DX \cdot DY}$$

   (b) ***Необходимость ($\implies$).***

   Если

$$\begin{gathered}|\operatorname{cov}(X,\, Y)| = \sqrt{DX \cdot DY} \implies \text{дискриминант} = 0 \implies \\ \implies D[Z(t)] \text{ имеет единственный корень. Обозначим его } t = a \implies \\ \implies D[Z(a)] = 0 \implies \\ \implies Z(a) = aX - Y \text{ принимает единственное значение с вероятностью 1,} \\ \text{обозначим это значение как } -b \implies Z(a) = aX - Y = -b \implies \\ \implies Y = aX + b\end{gathered}$$

(c) ***Достаточность ($\Longleftarrow$).***
Если $Y = aX+b \implies Z(a) = -b \implies D[Z(a)] = 0 \implies$ *дискриминант* $= 0 \implies |\operatorname{cov}(X, Y)| = \sqrt{DX \cdot DY}$.

6.

$$\begin{aligned}\operatorname{cov}(X, Y) = M[(X - m_1)(Y - m_2)] = M[XY - m_1Y - m_2X + m_1m_2] = \\ = M[XY] - m_1 \underbrace{MY}_{m_2} - m_2 \underbrace{MX}_{m_1} + m_1m_2 = M[XY] - m_1m_2\end{aligned}$$

2.15 Сформулировать определение ковариации и коэффициента корреляции случайных величин. Сформулировать свойства коэффициента корреляции. Сформулировать определения независимых и некоррелированных случайных величин, указать связь между этими свойствами. Понятия ковариационной и корреляционной матриц. Записать свойства ковариационной матрицы.

Пусть (X, Y) — двумерный случайный вектор.

Определение. *Ковариацией случайных величин X и Y называется число*

$$\operatorname{cov}(X, Y) = M[(X - m_X)(Y - m_Y)]$$

где $m_X = MX$, $m_Y = MY$.

Замечание. *Недостатком ковариации является то, что она имеет размерность, равную произведению разностей случайных величин X и Y. Часто рассматривают аналогичную безразмерную характеристику, которая называется коэффициентом корреляции случайных величин случайных величин X и Y:*

$$\rho(X, Y) = \rho_{XY} = \frac{\operatorname{cov}(X, Y)}{\sqrt{DX \cdot DY}}$$

где $DX \cdot DY > 0$.

Свойства коэффициента корреляции:

1. $\rho_{XX} = 1$;

2. Если X, Y независимые, то $\rho_{XY} = 0$;

3. $\rho(a_1X + b_1, a_2Y + b_2) = \pm\rho(X, Y)$, причём $\pm$ заменяется на

   - $+$, если $a_1a_2 > 0$;

- $-$, если $a_1 a_2 < 0$.

4. $|\rho_{XY}| \leqslant 1$, причём $\rho_{XY} = \begin{cases} 1, & \text{когда } Y = aX + b, \text{ где } a > 0, \\ -1, & \text{когда } Y = aX + b, \text{ где } a < 0 \end{cases}$

Определение. *Случайные величины X и Y называются независимыми, если*

$$F(x,\, y) = F_X(x)\, F_Y(y)$$

где F — совместная функция распределения X и Y ($\equiv$ функция распределения случайного вектора $(X,\, Y)$); F_X, F_Y — маргинальные функции распределения случайных величин X и Y.

Определение. *Случайные величины X и Y называют некоррелированными, если* $\operatorname{cov}(X,\, Y) = 0$.

Из свойства 3 ковариации следует, что если X, Y — независимые, то X, Y — некоррелированные. Обратное неверно — приведите пример самостоятельно.

Пусть $\overrightarrow{X} = (X_1,\, \ldots,\, X_n)$ — случайный вектор.

Определение. *Ковариантной матрицей случайного вектора $\overrightarrow{X}$ называется матрица*

$$\Sigma_{\overrightarrow{X}} = (\sigma_{ij})_{i,\, j = \overline{1;\, n}}$$

где $\sigma_{ij} = \operatorname{cov}(X_i,\, X_j)$.

Определение. *Корреляционной матрицей вектора $\overrightarrow{X} = (X_1,\, \ldots,\, X_n)$ называют матрицу*

$$P = (\rho_{ij})_{i,\, j = \overline{1;\, n}}$$

где $\rho_{ij} = \rho(X_i,\, X_j)$.

Замечание. *Логично показать некоторые свойства ковариантной матрицы:*

1. $\sigma_{ii} = DX_i$;
2. $\Sigma_{\overrightarrow{X}} = \Sigma_{\overrightarrow{X}}^T$;
3. *Если*

   $$\overrightarrow{Y} = \overrightarrow{X} B + \overrightarrow{c}$$

   где $\overrightarrow{Y} = (Y_1,\, \ldots,\, Y_m)$, $\overrightarrow{X} = (X_1,\, \ldots,\, X_n)$, $B \in M_{n,m}(\mathbb{R})$ (т. е. $\overrightarrow{Y}$ является линейной функцией от вектора $\overrightarrow{X}$), то

   $$\Sigma_{\overrightarrow{Y}} = B^T \Sigma_{\overrightarrow{X}} B$$

4. *Матрица* $\Sigma_{\overrightarrow{X}}$ *является неотрицательно определённой, т. е.* $\forall \overrightarrow{b} \in \mathbb{R}^{\omega}$

$$\overrightarrow{b}^T \Sigma_{\overrightarrow{X}} \overrightarrow{b} \geqslant 0$$

5. *Если все компоненты вектора* $\overrightarrow{X}$ *попарно независимы, то* $\Sigma_{\overrightarrow{X}}$ *— диагональная матрица.*

A Комбинаторика

Пусть X — некое множество. Для примеров определим $X = \{1, 2, 3, 4, 5, 6\}$.

A.1 Сочетания без повторений

Определение. *Сочетанием без повторений из n ($n = |X|$) элементов по m называется любое неупорядоченное подмножество множества X, содержащее m различных элементов.*

Кол-во таких подмножеств обозначается как C_n^m и равно

$$C_n^m = \frac{n!}{m!(n-m)!}$$

A.2 Размещения без повторений

Определение. *Размещением без повторений из n элементов (исходного множества, $n = |X|$) по m (длина кортежа) называется кортеж, состоящий из m различных элементов множества X.*

Примеры. $(1, 2, 4)$, *но не* $(5, 5, 4)$.

Кол-во возможных размещений без повторений:

$$A_n^m = \frac{n!}{(n-m)!}$$

A.3 Перестановки без повторений

Перестановки без повторений — крайний случай размещений без повторений.

Определение. *Перестановкой без повторений называют называют кортеж, состоящий из $n = |X|$ различных элементов множества X.*

Кол-во возможных перестановок без повторений:

$$P_n = A_n^n = n!$$

Замечание. *Три предыдущих понятия связаны как*

$$A_n^m = P_m C_n^m$$

A.4 Размещения с повторениями

Определение. *Размещением с повторениями из n ($n = |X|$) по m элементов называется любой элемент из $X^m = X \times X \times \ldots \times X$.*

Примеры. $(1,\, 2,\, 3,\, 4,\, 5)$, $(1,\, 4,\, 4,\, 4,\, 2)$.

Кол-во возможных размещений с повторениями:

$$\widetilde{A}_n^m = n^m$$

A.5 Перестановки с повторениями

Определение. *Перестановкой с повторениями называют кортеж длины n из элементов множества X, в котором каждый элемент $x_i \in X$ повторяется n_i, $i = \overline{1, k}$ раз ($\sum_i^k n_i = n$).*

Кол-во возможных перестановок с повторениями:

$$P_n(n_1,\, n_2,\, \ldots,\, n_k) = \frac{n!}{n_1! \cdot n_2! \cdot \ldots \cdot n_k!},\; n = n_1 + n_2 + \ldots + n_k$$