Теория вероятностей и математическая статистика—1

Винер Даниил @danya_vin

Версия от 11 ноября 2024 г.

Содержание

1 Дискретное вероятностное пространство. Базовые теоремы вероятности

Определение. $\Omega = \{\omega_1, \ldots, \omega_k, \ldots\}$ называется *пространством элементарных исходов*, где w_i — элементарный исход

Определение. A — любое подмножество Ω

Определение. Событие называется *достоверным*, если $A = \Omega$

Примечание. К A применимы те же опреации, что используются с множествами

Определение. Полная группа несовместных событий — такой набор событий, для которого выполняются такие условия:

$$A_i \cap A_j = \varnothing \ \forall i \neq j$$

$$\bigcup_i A_i = \Omega$$

Аксиома. $\forall \omega_i \ \exists p_i \geqslant 0$, при этом $\sum\limits_i p_i = 1$

Следствие. $0 \leqslant p_i \leqslant 1$

Определение. $P(A) = \sum\limits_{w_i \in A} P(w_i)$, где $P(w_i) = p_i$

(Ω, P) — вероятностное пространство в дискретном случае

Подходы к определению вероятностей

1. Априорный (предварительное знание)
2. Частотный (предел ряда частот)
3. Модельный (математическая модель)

1.1 Классическое определение вероятности

Имеет место, когда исходы равновероятны

Определение.

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

Определение. $P(A) = \sum\limits_{\omega_i \in A} p(\omega_i)$

1.2 Теорема сложения

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$P(A_1 \cup \ldots \cup A_n) = \sum\limits_{i=1}^{n} P(A_i) - \sum\limits_{i<j} (A_i \cap A_j) + \ldots + (-1)^{n-1} P(A_1 \cap \ldots \cap A_n)$

1.3 Условная вероятность

$P(A|B) = \dfrac{P(A \cap B)}{P(B)} \ \forall B : P(B) > 0$

1.4 Теорема умножения

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

2 Независимые события. Апостериорная и условная вероятности

Напомним классическое определение вероятности

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

Примечание. $P(\overline{A}) = 1 - P(A),\ P(\Omega) = 1$

2.1 Задача об авариях

Известно, что в 40% аварий виноваты пьяные водители. Утверждается, что из этого следует, что в 60% случаев виноваты трезвые водители

Пусть есть такие события: A — водитель пьян, B — водитель трезв, C — авария случилась

Формально, задача выглядит так: $P(A|C) = 0.4 \Longrightarrow P(B|C) = 0.6$

Пусть $P(B) = 0.05$

$\dfrac{P(C|A)}{P(C|B)} = \dfrac{P(C \cap A)/P(A)}{P(C \cap B)} = \dfrac{P(A|C)P(C)}{P(B|C)P(C)} \cdot \dfrac{P(B)}{P(A)} = \dfrac{0.4}{0.6} \cdot \dfrac{0.95}{0.05} \approx 12.7$

2.2 Независимость событий

Определение. (интуитивное) События A и B **независимы**, если $P(A|B) = P(A)$

Определение. События A и B называются попарно независимыми, если:

$$P(A \cap B) = P(A) \cdot P(B)$$
$$P(A|B)P(B) = P(A) \cdot P(B) \text{ — вытекает интуитивное определение}$$

Определение. События $A_1, \ldots, A_n$ независимы в совокупности, если:

$$\forall i_1 < \ldots < i_k < \ldots < i_n \ \forall k = 1, \ldots, n:$$
$$P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \ldots \cdot P(A_{i_k})$$

Примечание. Для A_1, A_2, A_3:

$$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2)$$
$$P(A_2 \cap A_3) = P(A_2) \cdot P(A_3)$$
$$P(A_1 \cap A_3) = P(A_1) \cdot P(A_3)$$
$$P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2) \cdot P(A_3)$$

2.3 Пример зависимых событий в совокупности

Положим, что у нас есть тетраэдр, каждая сторона которого покрашена в некоторые комбинации цветов: A — красный, B — желтый, C — зеленый, а четвертая покрашена во все три цвета, тогда вероятности:

$$\frac{1}{4} = P(A \cap B \cap C) \neq P(A) \cdot P(B) \cdot P(C) = \frac{1}{8}$$

Таким образом, события зависимы в совокупности

Примечание. Если $A_1, \ldots, A_n$ независимы в совокупности, то над любым из событий можно поставить знак отрицания, тогда система останется независимой

Пример. Есть события A_1, A_2, A_3 — независимы в совокупности, тогда $\overline{A_1}, A_2, A_3$ — тоже независимы в совокупности

Проверим данный пример.

$$
\begin{aligned}
P(\overline{A_1} \cap A_2 \cap A_3) &= P((A_2 \cap A_1) \setminus A_1) \\
&= P(A_2)P(A_3) - P(A_1 \cap A_2 \cap A_3) \\
&= P(A_2)P(A_3) - P(A_1)P(A_2)P(A_3) \\
&= P(\overline{A_1})P(A_2)P(A_3)
\end{aligned}
$$

Теперь разберемся с двумя событиями

$$
\begin{aligned}
P(\overline{A_1} \cap A_2) &= P(\overline{A_1})P(A_2) \\
P(\overline{A_2} \cap A_1) &= P(A_2 \setminus A_1) \\
&= P(A_2) - P(A_2 \cap A_1) = P(A_2) - P(A_2)P(A_1) \\
&= P(A_2)(1 - P(A_1))
\end{aligned}
$$

2.4 Простейший варинат ЗБЧ. Неизбежность технологических катастроф

Имеется n узлов, а $A_1, \ldots, A_n$ — события, где A_i означает, что i-ый узел вышел из строя

Очевидно, что P(хотя бы один узел выйдет из строя) $= P(A_1 \cup \ldots \cup A_n)$, тогда $\forall i : 0 < \varepsilon \leqslant P(A_i) < 1$ выполняется:

$$
1 \geqslant P(A_1 \cup \ldots \cup A_n) = 1 - P(\overline{A_1} \cap \ldots \cap \overline{A_n}) = 1 - \prod_{i=1}^{n} \underbrace{P(\overline{A_i})}_{\leqslant 1-\varepsilon} \geqslant \lim_{n \mapsto \infty} 1 - (1-\varepsilon)^n = 1
$$

Это означает, что вероятность технологических катастроф при количестве узлов, стремящемся в бесконечность, равна 100%

2.5 Формула полной вероятности

Пусть $\{H_i\}$ — полная группа несовместных событий (разбиение Ω)

Некоторые свойства:

- $H_i \cap H_j = \varnothing \ \forall i \neq j$ — несовместность
- $\displaystyle\bigcup_{i=1}^{n} H_i = \Omega$ — полнота

Теорема. Тогда, $P(A) = \displaystyle\sum_{i=1}^{n} P(A|H_i)P(H_i)$

Доказательство.

$$
\begin{aligned}
P(A) &= P\left(\bigcup_{i=1}^{n} (A \cap H_i)\right) \\
&= \sum_{i=1}^{n} P(A \cap H_i) \\
&= \sum_{i=1}^{n} P(A|H_i) \cdot P(H_i)
\end{aligned}
$$

$\square$

2.6 Формула Байеса. Апосториорные вероятности

$$
\begin{aligned}
P(H_k|A) &= \frac{P(A|H_k) \cdot P(H_k)}{P(A)} \\
&= \frac{P(A|H_k) \cdot P(H_k)}{\sum\limits_{i=1}^{n} P(A|H_i)P(H_i)}
\end{aligned}
$$

2.7 Задача про неизличимые заболевания

Требуется вычислить, какова вероятность, что человек, который получил положительный тест на СПИД, на самом деле не болен

$P(\text{СПИД}) = 0.03$ — вероятность быть носителем СПИДа

$P(+|\ \text{СПИД}) = 0.98$ — чувствительность теста, т.е. тест будет положительным, если у человека есть СПИД, с вероятностью 0.98

$P(+|\overline{\text{СПИД}}) = 0.01 = 1 -$ специфичность, т.е. тест будет положительным, если у человека нет СПИДа, с вероятностью 0.01

$$
\begin{aligned}
P(\text{СПИД}|+) &= \frac{P(+|\text{СПИД})P(\text{СПИД})}{P(+|\text{СПИД})P(\text{СПИД}) + P(+|\overline{\text{СПИД}})P(\overline{\text{СПИД}})} \\
&= \frac{0.98 \cdot 0.03}{0.98 \cdot 0.03 + 0.01 \cdot 0.97} \\
&\approx 0.75
\end{aligned}
$$

То есть, **каждый четвертый** человек с положительным СПИД-тестом **здоров**

3 Схема Бернулли. Теорема Пуассона. Случайные величины

Рассмотрим задачу проведения социологических опросов. Допустим, проводится опрос «Употребляет ли человек наркотики?». Предполагается, что не все респонденты, даже на условиях анонимности, дадут честный ответ. Для решения этой проблемы иногда задаются косвенные вопросы, мы же рассмотрим следующий подход.

Пусть человек перед ответом берет из урны случайный шарик. В урне лежит N шариков трех цветов: белый, черный и красный. Если человек достал белый, то его ответ на вопрос просто число «0», если черный — «1», если красный — респондент должен сказать правду: употребляет ли он наркотики. Тогда, все ответы разделятся на три категории, которые можно описать, как

$$N = N_{\text{к}} + N_{\text{ч}} + N_{\text{б}}$$

Пусть p — вероятность употребления наркотиков респондентом, тогда по формуле полной вероятности получаем

$$P(1) = P(1|\text{б})P(\text{б}) + P(1|\text{ч})P(\text{ч}) + P(1|\text{к})P(\text{к})$$

Отсюда следует, что $p = \dfrac{P(1)N - N_{\text{ч}}}{N_{\text{к}}}$

Пример. Если оказалось, что $P(1) = 0.6$, это означает, что $p = \frac{1}{2}$

3.1 Схема Бернулли. Биномиальное распредление вероятности

Допустим, что мы проводим n независимых опытов, каждый из которых может закончится успехом (1) или неудачей (0). Пусть вероятность успеха $P(1) = p$, тогда $P(0) = 1 - p = q$

Упорядочим результаты экспериментов в последовательность длины n, тогда всего исходов $|\Omega| = 2^n$, то есть последоватльности длины n, состоящие из 0 и 1. Заметим, что они **НЕ равновероятны**

Пользуясь тем, что события независимы, получаем, что

$$P(\underbrace{1, \ldots, 1}_{k}, \underbrace{0, \ldots, 0}_{n-k}) = p^k (1-p)^{n-k}$$

Заметим, что исходов, в которых k успехов, существует C_n^k, при этом

$$P(k \text{ успехов в } n \text{ испытаниях}) = C_n^k p^k (1-p)^{n-k}$$

Определение. Набор чисел $\{C_n^k p^k (1-p)^{n-k} | k = 0, 1, \ldots, n\}$ называется **биномиальным распределением вероятности**

3.2 Исследование Пуассона

Французский математик Симеон Дэни Пуассон проводил исследование, в котором пытался понять насколько справедливы обвинительные приговоры, которые выносят судьи присяжных во Франции.

Присяжных $n = 12$ человек. Для того, чтобы человек был признан виновынм, нужно 7 голосов «ЗА». Пуассон знал лишь процент обвинительных приговоров и количество обвинительных приговоров, вынесенных ровно 7 присяжными заседателями. Пусть p — вероятность верного решения присяжного (неизвестный нам параметр). Верное решение — когда улик достаточно для обвинения - присяжный голосует за *виновность*, не достаточно — за *невиновность*

Пусть есть событие A — улик достаточно для обвинения. Заметим, что $P(A)$ нам неизвестно.

Очевидно, что частота обвинительных приговоров= P(обвинение), тогда

$$P(\text{обвинение}) = P(\text{обвинение}|A)P(A) + P(\text{обвинение}|\overline{A})P(\overline{A})$$

Теперь введем вероятность, что человек был обвинен, когда за это проголосовало ровно 7 присяжных:

$$\begin{aligned} P(\text{обв. } 7) &= P(\text{обв. } 7|A)P(A) + P(\text{обв. } 7|\overline{A})P(\overline{A}) \\ &= C_{12}^7 p^7 (1-p)^5 P(A) + C_{12}^5 (1-p)^7 p^5 (1 - P(A)) \end{aligned}$$

Тогда, $P(\text{обвинение}) = \left(\sum_{k=7}^{12} C_{12}^k p^k (1-p)^{12-k}\right) P(A) + \left(\sum_{k=7}^{12} C_7^{12-k} (1-p)^k p^{12-k}\right)(1 - P(A))$

Пуассон решил, что $p \approx \dfrac{2}{3}$. Однако, это было ошибочным решением. Математик считал, что присяжные выносят решения независимо друг от друга, что в корни неверно. А так как Пуассон опирался на схему Бернулли, которая работает только для *независимых событий*, то и результат его исследования получился ложным

При этом, Пуассона не смутило, что в Англии для вынесения обвинительного приговора нужно, чтобы все 12 присяжных проголосавали за обвинение. Если бы решение Пуассона было верным, тогда вероятность обвинительных приговоров должна быть $\approx \left(\frac{2}{3}\right)^{12}$, что не соответствовало действительности, ведь в Англии было больше обвинительных приговоров чем во Франции

3.3 Теорема Пуассона

Теорема. Пусть имеются серии испытаний. В каждой серии n испытаний и p_n — вероятность успеха. Также, пусть $n \to \infty$ и $p_n \to 0$, причем $np_n \to \lambda$

Тогда,

$$\forall k \; P(k \text{ успехов в } n \text{ испытаниях}) = C_n^k p^k (1-p)^{n-k} \to \frac{\lambda^k}{k!} e^{-\lambda}$$

Доказательство. Пусть μ_n — число успехов в n испытаниях Бернулли с вероятностью успеха p_n, а также $\lambda_n = n \cdot p_n$. По условию, $\lambda_n^k \to \lambda > 0$. Подставим $p_n = \dfrac{\lambda_n}{n}$ в формулу Бернулли. Тогда,

$$\begin{aligned}
P(\mu_n = m) &= \frac{n!}{m!(n-m)!} p_n^m (1-p_n)^{n-m} \\
P(\mu_n = m) &= \frac{n \cdot \ldots \cdot (n-m+1)}{n^k} \left(\frac{\lambda_n}{m}\right)^m \frac{(1-\frac{\lambda_n}{n})^n}{(1-\frac{\lambda_n}{n})^m} \\
&= \frac{n \cdot \ldots \cdot (n-m+1)}{n^k} \left(\frac{\lambda_n}{m}\right)^m \left(1 - \frac{\lambda_n}{n}\right)^n \cdot \frac{1}{(1-\frac{\lambda_n}{n})^m} \\
&= \underbrace{\frac{n \cdot \ldots \cdot (n-m+1)}{n^k}}_{\lim\limits_{n\to\infty}=1} \frac{\lambda_n^m}{m!} \underbrace{\left(1 - \frac{\lambda_n}{n}\right)^n}_{e^{-\lambda}} \cdot \underbrace{\left(1 - \frac{\lambda_n}{n}\right)^{-m}}_{\lim\limits_{n\to\infty}=1} \\
&= \frac{\lambda^k}{k!} e^{-\lambda}
\end{aligned}$$

Определение. Набор чисел $\left\{\frac{\lambda^k}{k!} e^{-\lambda} \colon k = 0, 1, \ldots, n\right\}$ называется **распределением Пуассона**

Примечание. Величина $\frac{\lambda^k}{k!} e^{-\lambda}$ устойчива при зависимых событиях, а также при ошибках предпосылок (таких, какие допустил Пуассон)

Примечание. $\forall B$ — множество значений, которые может принимать число успехов

$$\left| P(\mu \in B) - \sum_{m \in B} e^{-\lambda} \frac{\lambda^m}{m!} \right| \leqslant \min(p, np^2)$$

3.4 Случайные величины. Математическое ожидание

Определение. Случайная величина $\xi : \Omega \to \mathbb{R}$

ξ имеет биномиальное распределение, если

$$\xi \in \{0, 1, \ldots, n\} \; P(\xi = k) = C_n^k p^k (1-p)^{n-k},$$

где $\{0, 1, \ldots, n\}$ — число успехов в n испытаниях

$P(\xi \in B) = \sum\limits_{\omega} P(\omega : \xi(\omega) \in B)$

Определение. Математическим ожиданием величины ξ называется величина

$$\mathbb{E}(\xi) = \sum_{\omega} P(\xi) \cdot \xi(\omega),$$

если этот ряд сходится абсолютно, то есть $\sum |\xi(\omega)| \, P(\omega)$

Теорема. $\mathbb{E}(\xi) = \sum\limits_{k} a_k P(\xi = a_k)$

4 Основные дискретные распределения. Математическое ожидание

Напомним определение случайной величины — $\xi : \Omega \to \mathbb{R}$

Определение. Дискретное вероятностное пространство — $(\Omega, P) \to (\mathbb{R}, P(\xi))$, где $P\xi : P(\xi = a)\ \forall a \in \mathbb{R}$

Если $\xi \in a_1 \dots a_k$, а вероятности равны $p_1 \dots p_k$, где $p_i = P(\xi = a)$, тогда

$$P(\xi = a) = P\left(\underbrace{\{\omega : \xi(\omega) = a_i\}}_{A\text{-прообраз } \{\xi = a_i\}}\right) \text{ — распределение вероятностей случайной величины } \xi$$

Пример. $\xi : 0, 1, \dots, n,\ p : P(\xi = k) = C_n^k p^k (1-p)^{n-k}$ — биномиальное распредеденме

Обозначение: $Bi(n, p)$

Определение. $\mathbb{E}(\xi) = \sum\limits_{\omega \in \Omega} \xi(\omega) P(\omega)$ — если ряд сходится абсолютно

Если ряд не сходится, то математического ожидания **не существует**

4.1 Основные дискретные распределения

1. Вырожденное распределение. $\exists c : P(\xi = c) = 1,\ c \in \mathbb{R}$
2. Распределение Бернулли: $Bi(1, p)$. $\xi \in [0, 1]$ с вероятностями p и $1 - p$ соответственно

   Индикатор события A — $I\{A\} = \begin{cases} 1, \text{ если } A \\ 0, \text{ если } \overline{A} \end{cases}$

   Пример. 10 кубиков, A — хотя бы одна 6

   $I\{A\} = \begin{cases} 1, P = 1 - \left(\frac{5}{6}\right)^{10} \\ 0, P = \left(\frac{5}{6}\right)^{10} \end{cases}$
3. Биномиальное распределение $Bi(n, p)$, где μ_n — число успехов в n испытаниях

   $P(\mu_n = k) = C_n^k p^k (1-p)^{n-k}, k = \overline{0, n}$

   $\mathbb{E}(\mu_n) = \sum\limits_{k=0}^{n} k C_n^k p^k (1-p)^{n-k}$
4. Распределение Пуассона: $\Pi(\lambda)$

   $P(\xi = k) = \dfrac{\lambda^k}{k!} e^{-\lambda}, k \in \mathbb{Z}$

   Иногда это распределение называют *распределением редких событий*

   $\mathbb{E}(\xi) = \lambda\ (\lambda = np)$, где $\mathbb{E}(\xi) = \sum\limits_{k=0}^{\infty} k \dfrac{\lambda^k}{k!} e^{-\lambda}$
5. Геометрическое распределение

   Пример. Число неудач до первого успеха $= \xi - 1$ **или** номер первого успеха $= \xi$

   При этом, p — вероятность успеха

   $\xi : 1, \dots .\ P(\xi = k) = (1-p)^{k-1} p$

   $\mathbb{E}(\xi) = \sum\limits_{k=1}^{\infty} k(1-p)^{k-1} p$

   Пусть $\mu = \mathbb{E}(\xi)$ — мат.ожидание номера первого успеха

   Тогда, $\mu = 1 \cdot p + (1-p)(1+\mu) \Longrightarrow \mu = \frac{1}{p}$
6. Отрицательное биномиальное распределение: $\overline{Bi}(r, p)$. Номер r-го успеха, k-е место

   $P(\xi = k) = C_{k-1}^{k-r}\ p^{r-1} (1-p)^{k-r} \cdot p$

7. Гипергеометрическое распределение

   $$P(\xi = k) = \frac{C_M^k C_{N-M}^{n-k}}{C_n^k}$$

4.2 Свойства линейности мат.ожидания

Пусть $\mu_n = \sum\limits_{i=1}^{n} I\{A_i\}$, где A_i — в i-ом опыте был успех

$\mathbb{E}(I\{A_i\}) = P(A_i) = p$, тогда, $\mathbb{E}(\mu_n) = np$

$np - q \leqslant$ (наиболее вероятное число успехов) $\leqslant np + p$

5 Аксиоматика Колмогорова

Утверждение. Всякий непротиворечивый веротяностный эксперимпент может быть заисан на языке теории меры

Определение. Пусть задана непустное множество Ω. Система подмножеств $\mathcal{F}$ называется σ-алгеброй, если

- $\Omega \in \mathcal{F}$
- если $A \in \mathcal{F}$, то $A^c \in \mathcal{F}$
- если $A_1, \ldots, A_n \in \mathcal{F}$, то $\bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{F}$

При этом, Ω называется единицей σ-алгебры $\mathcal{F}$

Пример. $\Omega = \{a, b, c, d\}$, $\mathcal{F} = \{\varnothing, \Omega\}$ — σ-алгебра? Первые два условия выполняются тривиально, а третье

$$\bigcup_{i=1}^{\infty} A_i = \begin{cases} \varnothing \in \mathcal{F}, \text{ если все } A_i = \varnothing \\ \Omega \in \mathcal{F}, \text{ если хотя бы один из } A_i = \Omega \end{cases}$$

Пример. $\Omega = \{a, b, c, d\}$, $\mathcal{F} = \{\varnothing, \Omega, \{a, b\}\}$ — σ-алгебра?

Нет, так как $\{a, b\} \in \mathcal{F}$, но $\{a, b\} = \{c, d\} \notin \mathcal{F}$

Утверждение. Пусть $\mathcal{F}$ — σ-алгебра, тогда если $B, C \in \mathcal{F}$, то

- $B \cup C \in \mathcal{F}$

  Доказательство. $B \cup C \cup \varnothing \cup \varnothing \ldots \cup \varnothing \ldots \in \mathcal{F}$
- $B \cap C \in \mathcal{F}$

  Доказательство. $B \cap C = (B^c \cup C^c)^c \in \mathcal{F}$ по предыдущему пункту и свойству σ-алгебры №2
- $B \setminus C \in \mathcal{F}$

  Доказательство. $B \cap C^c \in \mathcal{F}$

Утверждение. Пусть $\mathcal{F}$ — σ-алгебра, тогда если $A_1, \ldots, A_n \ldots \in \mathcal{F}$, то $\bigcap\limits_{i=1}^{\infty} A_i \in \mathcal{F}$

Доказательство. $\bigcap\limits_{n=1}^{\infty} A_n = \underbrace{\bigcup\limits_{n=1}^{\infty} \underbrace{A_n^c}_{\in \mathcal{F}}}_{\in \mathcal{F}} \in \mathcal{F}$

Пусть $\omega \in \bigcap A_i \iff (\forall i \in \mathbb{N}\ \omega \in A_i) \iff (\forall i \in \mathbb{N}\ \omega \notin A_i^c) \iff (\omega \notin \bigcup\limits_{i=1}^{\infty} A_i^c) \iff \omega \in (\bigcup A_i^c)^c$

Определение. Пусть $\Omega \neq \varnothing$ и $\mathcal{S}$ — это непустая система подмножеств множества Ω

Минимальной σ-алгеброй, содержащей систему $\mathcal{S}$ называется такая σ-алгебра $\sigma(\mathcal{S})$, что

- $\mathcal{S} \subseteq \sigma(\mathcal{S})$
- $\forall$ σ-алгебры $\mathcal{G}$, которая содержит систему $\mathcal{S}$ ($\mathcal{S} \subseteq \mathcal{G}$) справедливо, что $\sigma(\mathcal{S}) \subseteq \mathcal{G}$

Определение. Минимальная σ-алгебра, содержащая все полуинтервалы вида $(a; b] \in \mathbb{R}$, где a и b, такие что $a, b \in (-\infty; +\infty)$ называется **борелевской σ-алгеброй** на числовой прямой и обозначается $\mathcal{B}(\mathbb{R})$

Элементы борелевской σ-алгебры называются **борелевскими множествами**

Пример. Докажите, что следующие помдмножества числовой прямой являются борелевскими

a) $\{b\} = \bigcap_{n=1}^{\infty} \underbrace{\left(b - \frac{1}{b}, b\right]}_{\in \mathcal{B}(\mathbb{R}) \text{ по опр.}} \in \mathcal{B}(\mathbb{R})$

b) $(a, b) = \underbrace{(a, b]}_{\substack{\in \mathcal{B}(\mathbb{R}) \\ \text{по опр.}}} \setminus \underbrace{\{b\}}_{\substack{\in \mathcal{B}(\mathbb{R}) \\ \text{по 1.}}} \in \mathcal{B}(\mathbb{R})$

c) $[a, b) = \underbrace{\{a\}}_{\substack{\in \mathcal{B}(\mathbb{R}) \\ \text{по 1.}}} \cup \underbrace{(a, b)}_{\substack{\in \mathcal{B}(\mathbb{R}) \\ \text{по 2.}}} \in \mathcal{B}(\mathbb{R})$

d) $[a, b] = \underbrace{\{a\}}_{\substack{\in \mathcal{B}(\mathbb{R}) \\ \text{по 1.}}} \cup \underbrace{(a, b]}_{\substack{\in \mathcal{B}(\mathbb{R}) \\ \text{по опр.}}} \in \mathcal{B}(\mathbb{R})$

e) $\mathbb{Q} = \bigcup_{k=1}^{\infty} \underbrace{\{r_k\}}_{\in \mathcal{B}(\mathbb{R})} \in \mathcal{B}(\mathbb{R})$

f) $\underbrace{\mathbb{R}}_{=\Omega} \setminus \underbrace{\mathbb{Q}}_{\in \mathcal{B}(\mathbb{R})} \in \mathbb{Q}^c \in \mathcal{B}(\mathbb{R})$

Определение. Пусть задано непустое множество Ω и некоторая σ-алгебра $\mathcal{F}$ подмножества в множестве Ω. Тогда упорядоченная пара $(\Omega, \mathcal{F})$ называется **измеримым пространством**

Определение. Пусть задано измеримое пространство $(\Omega, \mathcal{F})$. Функция $\xi : \Omega \to \mathbb{R}, \xi = \xi(\omega), \omega \in \Omega$ называется измеримой относительно σ-алгебры $\mathcal{F}$ функцией, если

$$\forall c \in \mathbb{R} \text{ множество } \{\omega \in \Omega : \xi(\omega) > c\} \in \mathcal{F}$$

Функцию, измеримую относительно σ-алгебры $\mathcal{F}$, также называют $\mathcal{F}$-измеримой функцией

Определение. В теории вероятностей $\mathcal{F}$-измеримые функции также называются слчайными величинами, или $\mathcal{F}$-измеримыми случайными величинами, если хотят подчернкуть тот факт, что они измеримы относительно σ-алгебры $\mathcal{F}$

Пример. $\Omega = \{a, b, c, d\}$, $\mathcal{F} = \{\varnothing, \Omega, \{a, b\}, \{c, d\}\}$, $\xi : \Omega \to \mathbb{R}, \eta : \Omega \to \mathbb{R}$

1. ξ — $\mathcal{F}$-измерима
   $c = 10^6 : \ \{\omega \in \Omega : \xi(\omega) = c\} = \varnothing \in \mathcal{F}$
   $c > 1 : \ \{\omega \in \Omega : \xi(\omega) > c\} = \varnothing \in \mathcal{F}$
   $c = 1 : \ \{\omega \in \Omega : \xi(\omega) > 1\} = \varnothing \in \mathcal{F}$
   $-1 < c < 1 : \ \{\omega \in \Omega : \xi(\omega) > c\} = \{c, d\} \in \mathcal{F}$
2. η — не $\mathcal{F}$-измерима
   $c = 3.5 : \ \{\omega \in \Omega : \eta(\omega) > 3.5\} = \{d\} \notin \mathcal{F} \Longrightarrow$ не $\mathcal{F}$-измерима

Теорема. Пусть $\xi : \Omega \to \mathbb{R}$ — $\mathcal{F}$-измеримая функция. Тогда, $\forall a, b, c \in \mathbb{R}$

1. $\{\omega \in \Omega : \xi(\omega) \leqslant c\} \in \mathcal{F}$
2. **Доказательство.** $\{\omega \in \Omega : \xi(\omega) \leqslant c\} = \underbrace{\Omega}_{\in \mathcal{F}} \setminus \underbrace{\{\omega \in \Omega : \xi(\omega) > c\}}_{\in \mathcal{F}} \in \mathcal{F}$

Определение. Пусть задано измеримое пространство $(\Omega, \mathcal{F})$. Функция $\mathbb{P}\colon \mathcal{F} \to [0, 1]$ называется **вероятностной мерой**, если выполнены следующие условия:

1. **(условие нормировки)** $P(\Omega) = 1$
2. **(условие счётной аддитивности, σ-аддитивность)** Для любой последовательности попарно непересекающихся множеств $A_1, \ldots, A_n, \cdots \subset \mathcal{F}$ верно

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

Определение. Пусть $\xi : \Omega \to \mathbb{R}$ — $\mathcal{F}$-измеримая функция. Функция

$$F_\xi(x) = \mathbb{P}(\{\omega \in \Omega : \xi(\omega) \leqslant x\}), x \in \mathbb{R}$$

называется **функцией распределения** случайной величины ξ

6 Свойства функции распределения. Функция плотности. Квантиль

6.1 Свойства функции распределения

1. $\exists$ пределы $\lim\limits_{x\to\infty} F_\xi(x) = 1,\ \lim\limits_{x\to-\infty} F_\xi(x) = 0$
2. F_ξ не убывает: $F_\xi(x_1) \leqslant F_\xi(x_2)\ \forall x_1 \leqslant x_2$
3. F_ξ непрерывна справа: $\lim\limits_{x\to x_0+} F_\xi(x) = F_\xi(x_0)$

Пример. Случайно бросаем точку в квадрат. Вероятность попадания в какую-то конкретную область:

$$\frac{S(A)}{1}$$

Пусть $\xi = u$ — первая координата

$$F_\xi(x) = \mathbb{P}\left(\{\xi \le x\}\right),\ x \in [0,1]$$
$$\begin{cases} F_\xi(x) = 0, & x < 0 \\ F_\xi(x) = 1, & x > 1 \end{cases}$$

Определение. Случайная величина ξ называется **непрерывной**, если её функция распределения непрерывна

6.2 Функция плотности

Определение. Случайная величина ξ называется **абсолютно непрерывной**, если $\exists f_\xi(x) \geqslant 0$, что функция распределения представима в виде интеграла:

$$F_\xi(x) = \int\limits_{-\infty}^{x} f_\xi(x)\mathrm{d}x$$

Определение. Функция $f_\xi(x)$ называется **функцией плотности вероятности**

Примечание. Во всех точках непрерывности $f\xi(x)$:

$$F'_\xi(x) = f_\xi(x)$$

6.3 Свойства функции плотности

1. $f_\xi(x) \geqslant 0, \forall x \in \mathbb{R}$
2. $\int\limits_{-\infty}^{+\infty} f_\xi(x)\mathrm{d}x = 1$ — условие нормировки

Определение. Сингулярное распределение — непрерывное распределение, для которого $\nexists$ функции плотности

6.4 Теорема Лебега

Теорема. Любая функция распределения $F_\xi(x)$ представима единственным образом в виде суммы

$$F_\xi(x) = c_1F_1(x) + c_2F_2(x) + c_3F_3(x),\ \text{где} \begin{cases} F_1\text{— дискретная} \\ F_2\text{— сингулярная} \\ F_3\text{— абсолютно непрерывная} \end{cases}$$

$$\begin{cases} c_1 + c_2 + c_3 = 1 \\ c_i \geq 0 \end{cases}$$

Примечание. Рассмотрим также такие свойства функции плотности и распределения

1. $\mathbb{P}(\{a < \xi \leqslant b\}) = F_\xi(b) - F_\xi(a)$ для любых распределений

2. $\mathbb{P}(\{a < \xi \leqslant b\}) = \int\limits_a^b f_\xi(x)\mathrm{d}x$ для абсолютно непрерывных распределений

3. Если случайная величина ξ имеет абсолютно непрерывное распределение, то $F_\xi(x)$ непрерывна

6.5 Квантиль

Определение. Квантилью уровня $u \in (0; 1)$ случайной величины ξ называется наименьшее число $q \in R$:

$$\int\limits_{-\infty}^{q} f_\xi(t)dt = u$$

7 Непрерывные распределения. Дисперсия. Преобразования случайных величин

7.1 Смешанное дискретно-непрерывное распределение

Определение. Смешанное дискретно-непрерывное распределение:

$$\exists a_1 \dots, a_n,\ p_1 \dots p_n \dots,\ \exists f_\xi(x) \geqslant 0$$
$$\mathbb{P}\left(\{\xi = a_k\}\right)$$
$$\mathbb{P}\underset{B \subset \mathcal{B}}{\left(\{\xi \in B\}\right)} = \sum_{a_k \in B} \mathbb{P}\left(\{\xi = a_k\}\right) + \int\limits_B f_\xi(x)\mathrm{d}x$$

7.2 Преобразование случайных величин

Примечание. $\forall$ борелевской (измеримой) функции $g(x)$, $x \in \mathbb{R}$ $g(\xi)$ — случайная величина

Теорема. Пусть ξ — абсолютно непрерывна, $g(x)$ монотонна, дифференцируема, $g'(x) \neq 0, \forall x \in \mathbb{R}$. $\eta = g(\xi)$. Тогда

$$f_\eta(y) = \frac{f_\xi(g^{-1}(y))}{|g'(g^{-1}(y))|}$$

Доказательство. Пусть $(g(x) \uparrow)$. Найдём $F_\eta(y) = \mathbb{P}(\eta \leq y) = \mathbb{P}(g(\xi) \leq y) = \mathbb{P}(\xi \leq g^{-1}(y)) = F_\xi(g^{-1}(y))$

$$F'_\eta = f_\eta(y) = \frac{f_\xi(g^{-1}(y))}{g'(g^{-1}(y))}$$

7.3 Свойства математического ожидания

Для $\forall$ борелевской функции $g(x)$ выполняется:

1. $\mathbb{E}\left[g(\xi)\right] = \begin{cases} \sum\limits_k g(xk)\mathbb{P}(\xi = x_k), & \text{если } \xi \text{ — дискретная} \\ \int\limits_{-\infty}^{\infty} g(x) f_x(x)\mathrm{d}x, & \text{если } \xi \text{ абсолютно непрерывна} \end{cases}$

   $\eta = g(\xi), \quad \int\limits_{-\infty}^{\infty} y f_\eta(y)\mathrm{d}y = \int\limits_{-\infty}^{\infty} g(x) f_\xi(x)\mathrm{d}x$
2. $\mathbb{E}\left[c\right] = c$
3. $\mathbb{E}\left[c\xi\right] = c \cdot \mathbb{E}\left[\xi\right]$
4. $\mathbb{E}\left[\xi + \eta\right] = \mathbb{E}\left[\xi\right] + \mathbb{E}\left[\eta\right]$
5. Если $\mathbb{P}\left(\{\xi \geqslant 0\}\right) = 1$, $(\xi \geqslant 0)$, то $\mathbb{E}\left[\xi\right] \geqslant 0$
6. Если $\xi \geqslant \eta$ почти наверное, то $\mathbb{E}\left[\xi\right] \geqslant \mathbb{E}\left[\eta\right]$, $(\xi - \eta \geqslant 0)$
7. $a \leqslant \xi \leqslant b$ почти наверное, $a \leqslant E(\xi) \leqslant b$

Примечание. Почти наверное означает «с вероятностью 1»

Если $\xi \geqslant 0$ почти наверное $(\mathbb{P}\left(\{\xi \geqslant 0\}\right) = 1)$ и $\mathbb{E}\left[\xi\right] = 0$, то $\xi = 0$ почти наверное

7.4 Свойства дисперсии

1. $\mathbb{D}\left[\xi\right] = \mathbb{E}\left[\xi^2\right] - (\mathbb{E}\left[\xi\right])^2$
2. $\mathbb{D}\left[c\right] = 0$
3. $\mathbb{D}\left[c \cdot \xi\right] = c^2 \mathbb{D}\left[\xi\right]$
4. $\mathbb{D}\left[\xi\right] \geqslant 0$, при этом, если $\mathbb{D}\left[\xi\right] = 0$, то $\mathbb{P}\left(\{\xi = c\}\right) = 1$

7.5 Основные непрерывные распределения

1. $U[a,b]$ — равномерное

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

$$\mathbb{E}[\xi] = \frac{a+b}{2},\ \mathbb{D}[\xi] = \frac{(b-a)^2}{12}$$

2. Показательное (экпоненциальное)

$$f_\xi(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & x \leqslant 0 \end{cases}$$

$$\mathbb{E}[\xi] = \frac{1}{2},\ \mathbb{D}[\xi] = \frac{1}{\lambda^2}$$

$$F_\xi(x) = 1 - e^{-\lambda x},\ x > 0$$

$$\begin{aligned} \mathbb{P}(\{a+\tau < \xi < b \mid \xi < \tau\}) &= \mathbb{P}(\{a < \xi < b\}) \\ &= \frac{\mathbb{P}(\{a+\tau < \xi < b+\tau\})}{\mathbb{P}(\{\xi > \tau\})} \\ &= \mathbb{P}(\{a < \xi < b\}) \end{aligned}$$

8 Многомерное распределение

8.1 Случайные величины

Пусть $\xi_1, \ldots, \xi_n$ определены на $(\Omega, \mathcal{F}, \mathbb{P})$

Определение. $\xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$ — случайный вектор ИЛИ n-мерная случайная величина

Распределение случайного вектора $\forall B \in \mathcal{B}(\mathbb{R}^n)$ $\mathbb{P}\left(\{\xi \in B\}\right)$

8.2 Функция распределения

Определение. Многомерной (совместной) функцией распределения $F_\xi(x_1, \ldots, x_n)$ называется

$$\mathbb{P}\left(\{\xi_1 \leqslant x_1, \ldots, \xi_n \leqslant x_n\}\right)$$

1. $F_\xi(x) \in [0;1]$. Здесь и далее $x = (x_1, \ldots, x_n)$
2. $\lim\limits_{x_1 \to -\infty} F_\xi(x_1, x_2) = 0$
   $\lim\limits_{x_1 \to +\infty} F_\xi(x_1, x_2) = F_{\xi_2}(x_2)$
   $\lim\limits_{\substack{x_1 \to +\infty, \\ x_2 \to +\infty}} F_\xi(x_1, x_2) = 1$
3. $F_\xi(x_1, x_2)$ не убывает по каждому из аргументов
4. $F_\xi(x_1, x_2)$ непрерынвва справа по каждому из аргументов

Определение. Случайные величины ξ_1, ξ_2 имеют совместное дискретное распределение, если $\exists \{a_i, b_j\}$ не более, чем счетный набор чисел

$$\sum_i \sum_j \mathbb{P}\left(\{\xi_1 = a_i, \xi_2 = b_j\}\right) = 1$$

Пример.

	-1	0	1	$\mathbb{P}\left(\{\xi_2\}\right)$
-2	0.1	0.1	0.1	0.3
0	0.1	0.1	0.1	0.3
2	0.1	0.1	0.2	0.4
$\mathbb{P}\left(\{\xi_1\}\right)$	0.3	0.3	0.4	1

Определение. Случайные величины ξ_1, ξ_2 имеют совместное абсолютно непрерывное распределение, если $\exists f_\xi(x_1, x_2) \geqslant 0$ такая, что $\forall B \in \mathcal{B}(\mathbb{R}^2)$:

$$\mathbb{P}\left(\{\xi \in B\}\right) = \iint\limits_B f_\xi(x_1, x_2) \mathrm{d}x_1 \mathrm{d}x_2$$

Пример. $\begin{pmatrix} \xi \\ \xi \end{pmatrix}$ — не имеет совместной плотности

Примечание. Если $f_\xi(x_1, x_2)$ непрерына, то

$$\frac{\partial^2 F_\xi(x_1, x_2)}{\partial x_1 \partial x_2} = f_\xi(x_1, x_2)$$

8.3 Свойства функции плотности

1. $f_\xi \geqslant 0$ $\forall x_1, x_2 \in \mathbb{R}$

2. $\iint\limits_{\mathbb{R}^2} f_\xi(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2 = 1$

3. $f_{\xi_1}(x_1) = \int\limits_{-\infty}^{\infty} f_\xi(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2$

$$\begin{aligned} F_{\xi_1}(x_1) &= \mathbb{P}\left(\{\xi_1 \leqslant x_1 \cap \xi_2 \in \mathbb{R}\}\right) \\ &= \int\limits_{\infty}^{x_1}\int\limits_{-\infty}^{+\infty} f_\xi(t_1, x_2)\mathrm{d}x_2\mathrm{d}t_1 \end{aligned}$$

Пример. Равномерное распределение в области $S \in \mathbb{R}^n$

1. Имеем куб со стороной 1. Тогда $f_\xi(x_1, x_2, x_3) = \begin{cases} \frac{1}{V} = 1, & x_1, x_2, x_3 \in \text{куб} \\ 0, & \text{иначе} \end{cases}$

2. Имеем единичный круг в $\mathbb{R}^2$, тогда

$$f_\xi(x_1, x_2) = \begin{cases} \frac{1}{\Pi}, & x_1^2 + x_2^2 \leqslant 1 \\ 0, & \text{иначе} \end{cases}$$

3. В общем случае, для $S \in \mathbb{R}^n$

$$f_\xi(x) = \begin{cases} \frac{1}{V(S)}, & x \in S \\ 0, & x \notin S \end{cases}$$

8.4 Независимость случайных величин

Определение. $\xi_1, \ldots, \xi_n$ независимы, если $\forall B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R})$:

$$\mathbb{P}\left(\{\xi_1 \in B_1, \ldots, \xi_n \in B_n\}\right) = \mathbb{P}\left(\{\xi_1 \in B_1\}\right) \cdot \ldots \cdot \mathbb{P}\left(\{\xi_n \in B_n\}\right)$$

Определение. $\xi_1, \ldots, \xi_n$ независимы, если

$$F_\xi(x_1, \ldots, x_n) = F_{\xi_1}(x_1) \cdot \ldots \cdot F_{\xi_n}(x_n)$$

Определение. Дискретные случайные величины ξ_1, ξ_n независимы, если

$$\forall a_i, b_j \ \mathbb{P}\left(\{\xi_1 = a_i, \xi_2 = b_j\}\right) = \mathbb{P}\left(\{\xi_1 = a_i\}\right) \cdot \mathbb{P}\left(\{\xi_2 = b_j\}\right)$$

Определение. Для абсолютно непрерывных

$$f_\xi(x_1, \ldots, \xi_n) = f_{\xi_1}(x_1) \cdot \ldots \cdot f_{\xi_n}(x_n)$$

Определение. Копулой называется функция $C(x, y)$, опредедленная на единичном квадрате, $\begin{cases} x \in [0; 1] \\ y \in [0; 1] \end{cases}$

$C(x, 1) = \begin{cases} x, & x \in [0; 1] \\ 0, & x \notin [0; 1] \end{cases}$

Теорема. (Склара) Для произвольной двумерной функции распределения $F(x, y)$ с маржинальными распределениями $F_1(x)$ и $F_2(y)$ $\exists$ копула, такая что

$$F(x, y) = C(F_1(x), F_2(y))$$