# Основные понятия теории вероятностей, случайные величины и их распределения

Лекция №1

Статистические методы в ядерном эксперименте, ИЯФ, 2024 г.

1. Определение вероятности
2. Независимые события, сложение и умножение вероятностей
3. Случайные процессы
   - дискретные случайные величины
   - непрерывные случайные величины
   - многомерные распределения

# Определение вероятности I

## Классическое (комбинаторное) определение

Если испытание может приводить к $N$ различным равновозможным исходам, причём признаком А обладают $k$ исходов, то вероятность появления признака А есть:

$$P(A) = \frac{k}{N}$$

Условие равновозможности элементарных исходов, и правильная идентификация элементарного исхода в классическом определении вероятности являются критически важными.

Пример:

Чему равны вероятности выпадения сумм очков 9 или 10 двух(трёх) игральных кубиков ?

Замечание:

Недостатком такого подхода является то, что его применение ограничено задачами с набором равновозможных исходов.

# Определение вероятности II

## Частотное определение

Пусть испытание повторяется большое число раз $M$ и в $n$ случаях его исход обладает признаком А. Если исход каждого последующего испытания не зависит от предыдущих результатов, то вероятность $P(A)$ появления признака А:

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

Замечание:

Недостатком такого подхода является то, что он основан на экспериментировании и подразумевает неосуществимые эксперименты $(M \to \infty)$. Кроме того, такое определение ограничивает применимость понятия вероятности только наблюдаемыми величинами и воспроизводимыми опытами.

# Определение вероятности III

## Аксиоматическое определение

Рассмотрим множество $\Omega$, элементы $\omega_i$ которого назовём элементарными исходами, причём $\omega_i$ взаимно исключают друг друга, то есть появление одного из них означает, что ни один другой не произошел. Определим функцию $P(\omega_i)$ (вероятность исхода $\omega_i$) такую, что:

1. $P(\omega_i) \geq 0,\ \forall \omega_i \in \Omega$;
2. $P(\Omega) = 1$ - условие нормировки;
3. Если множества исходов $A$ и $B$ не имеют общих элементов $(A \cap B = \varnothing)$, то $P(A \cup B) = P(A) + P(B)$.

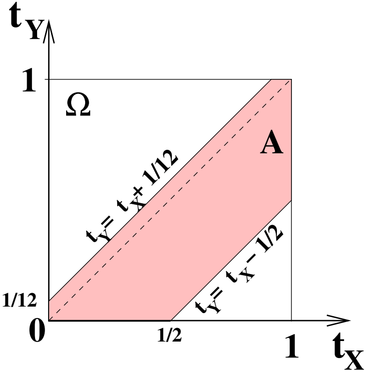
Говорят, что на множестве $\Omega$ задана вероятностная мера, а свойства $1 \div 3$ называются аксиомами вероятности.

Следствия:

- $0 \leq P(A) \leq 1, \quad P(\varnothing) = 0$
- $\bar{A} = \Omega \setminus A, \quad P(A \cup \bar{A}) = 1, \quad P(\bar{A}) = 1 - P(A)$
- $A \subset B, \quad P(A) \leq P(B)$

# Определение вероятности IV

Геометрическая вероятность, как одно из применений аксиоматического подхода:
Задача о встрече: X и Y договорились встретиться между 14-ю и 15-ю часами. Если Y пришел первым, то он ждет X в течении 30 минут (или до 15:00, меньшее из двух) и потом уходит. Если X пришла первой, то она ждет Y только 5 минут и уходит. Какова вероятность встречи, если каждый из них может придти в любой момент в течение условленного часа ?



# Сложение вероятностей

Чтобы найти вероятность того, что исход принадлежит событию $A$ или $B$, нужно сложить вероятности $A$ и $B$, однако исходы, принадлежащие одновременно как $A$, так и $B$, оказываются сосчитаны дважды, поэтому из суммы вероятности нужно вычесть $P(AB)$.

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

или

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

Задача: Вытащим из колоды, содержащей 52 карты, одну. Пусть $A$ обозначает "король", $B$ обозначает "бубновая масть". Чему равна вероятность того, что данная карта окажется либо королём, либо бубновой мастью ?

# Условная вероятность, умножение вероятностей

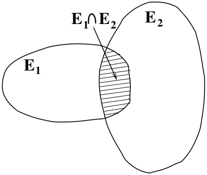
Пусть событию $E_1$ соответствуют $k_1$ элементарных исходов, событию $E_2$ — $k_2$ исходов, событию $E_1E_2$ — $k_{12}$ исходов. Пусть полное число элементарных исходов равно $n$. Тогда вероятность $P(E_1E_2)$ события $E_1E_2$ равна:

$$P(E_1) = \frac{k_1}{n}, \quad P(E_2) = \frac{k_2}{n}, \quad P(E_1E_2) = \frac{k_{12}}{n}$$

$$P(E_1E_2) = \frac{k_{12}}{n} = \frac{k_1}{n} \cdot \frac{k_{12}}{k_1} = P(E_1) \cdot \frac{k_{12}}{k_1} = P(E_1) \cdot P(E_2|E_1)$$

$$P(E_1E_2) = P(E_2) \cdot P(E_1|E_2)$$

$P(E_2|E_1) = \frac{k_{12}}{k_1}$ - вероятность события $E_2$ при условии, что событие $E_1$ уже произошло; $P(E_2|E_1)$ называется условной вероятностью. В общем случае $P(E_2|E_1) \neq P(E_1|E_2)$.




События $E_1$ и $E_2$ называют <u>независимыми</u> если $P(E_1E_2) = P(E_1)P(E_2)$, или: $P(E_2|E_1) = P(E_2)$. Т.е. если событие $E_1$ произошло, и это не влияет на вероятность события $E_2$, то события $E_1$ и $E_2$ независимы.
События называются <u>несовместными</u> если они не могут появиться одновременно в результате однократного испытания ($E_1E_2 = \varnothing : k_{12} = 0$).

<u>Несовместность и независимость событий - разные свойства !</u>

Пример:

Пусть из колоды 52-ух карт вынимают одну. Событием $E_1$ назовем - вынута карта "туз", а событием $E_2$ - вынута карта пиковой масти. Тогда $P(E_1) = 4/52 = 1/13$, $P(E_2) = 13/52 = 1/4$, $P(E_1E_2) = P(\text{"вынуть туз пик"}) = 1/52$. Видим, что выполняется соотношение $P(E_1E_2) = P(E_1)P(E_2)$, и значит события $E_1$ и $E_2$ независимы.

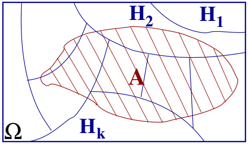
Рассмотрим эту же ситуацию при условии, что в колоду добавлен джокер (не имеет масти). Тогда $P(E_1) = 4/53$, $P(E_2) = 13/53$, $P(E_1E_2) = P(\text{"вынуть туз пик"}) = 1/53$. Однако теперь $P(E_1)P(E_2) = 1/53 - 1/53^2 \neq P(E_1E_2)$, т.е. события $E_1$ и $E_2$ перестали быть независимыми. Действительно, если мы знаем, что вынутая карта есть "туз", то это нам уже дает некоторую информацию о том, какой она может быть масти (точно не безмастный джокер).

В случае набора более чем двух событий, независимости каждой пары событий из этого набора для "полной" независимости всех событий (т.е. для того, чтобы знание того, что произошла какая-то комбинация событий из данного набора, не влияло на вероятности случиться остальным событиям) недостаточно.

# Формула полной вероятности

Набор попарно несовместных событий $H_1, H_2, \ldots$ $(H_i H_j = \varnothing, \forall i \neq j)$ таких, что $P(H_k) > 0, \ \forall k$ и $\cup H_k = \Omega$, называется полной группой событий или разбиением пространства $\Omega$.

Тогда вероятность любого события $A$ может быть записана как:
$P(A) = \sum\limits_k P(H_k)P(A|H_k)$ - формула полной вероятности




Действительно:

$$P(A) = P(A \cdot \Omega) = P(A \cdot (\cup H_k)) = P(\cup(AH_k)) = \sum_k P(AH_k) = \sum_k P(H_k)P(A|H_k)$$

Из $P(AH_k) = P(H_k)P(A|H_k) = P(A)P(H_k|A)$, получаем:

**Формула Байеса**

$$P(H_k|A) = \frac{P(H_k)P(A|H_k)}{P(A)} = \frac{P(H_k)P(A|H_k)}{\sum\limits_i P(H_i)P(A|H_i)}$$

# Дискретные и непрерывные случайные величины

Случайной величиной $\xi$ будем называть вещественную функцию, определенную на пространстве элементарных исходов $\Omega$, т.е.: $\xi : \Omega \to R$.

Случайная величина имеет дискретное распределение, если существует конечный или счётный набор вещественных чисел $a_k$, таких что:

$$P_k = P(\xi = a_k) > 0, \quad \sum_k P_k = 1.$$

Множество $(a_k,\ P_k)$ (таблица) называется дискретным распределением вероятности.

Случайная величина имеет непрерывное распределение, если существует функция $f_\xi(x) \geq 0$ такая, что для "любого" множества $B$ выполняется:

$$P(\xi \in B) = \int_B f_\xi(x)dx, \quad P(\xi \in R) = 1 = \int_R f_\xi(x)dx.$$

Функция $f_\xi(x)$ называется функцией плотности вероятности (в дальнейшем индекс $\xi$ в обозначении $f_\xi(x)$ будем опускать).

# Дискретные распределения I

**Схема Бернулли**

Схемой Бернулли называется последовательность независимых испытаний, в каждом из которых возможны лишь два исхода: "успех" с вероятностью $p$ и "неудача" с вероятностью $q = 1 - p$.

Найдём вероятность $r$ благоприятных исходов при полном числе испытаний $N$. В этом случае $r$ - случайная величина, пробегающая значения от 0 до $N$. Распределение случайной величины $r$ – биномиальное, с двумя параметрами, $p$ и $N$:

$$P(r|p,N) = C_N^r p^r (1-p)^{N-r} = \frac{N!}{r!(N-r)!} p^r (1-p)^{N-r}$$

Пример ($N = 2,\ r = 0, 1, 2$):
Вероятность срабатывания счётчика при попадании в него одной частицы равна $\varepsilon = 0.9$. Чему равна вероятность срабатывания счётчика при попадании в него двух частиц ?
$P(r=1) + P(r=2) = 1 - P(r=0) = 1 - (1-\varepsilon)^2 = 0.99$

# Дискретные распределения II

Замечания: Если имеются две серии испытаний Бернулли длинами $N_1$ и $N_2$ с одинаковым значением $p$, то их сумма также является серией испытаний Бернулли с $N = N_1 + N_2$, и полное число успехов в этой объединенной серии также имеет биномиальное распределение.
С другой стороны, испытание Бернулли можно рассматривать как биномиальное распределение с $N = 1$, тогда биномиальное распределение с $N > 1$ есть сумма $N$ элементарных биномиально-распределённых величин.

- Если $\mu = Np$ (среднее) фиксировано, то при $N \to \infty$ биномиальное распределение стремится к распределению Пуассона: $P(r|\mu) = \frac{\mu^r}{r!} e^{-\mu}$
- Если $p$ фиксировано, то при $N \to \infty$ биномиальное распределение стремится к нормальному (Гауссовому) распределению ($\mu = Np,\ \sigma = \sqrt{Np(1-p)}$): $P(r|\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(r-\mu)^2}{2\sigma^2}}$

# Дискретные распределения III

## Пуассоновский процесс

Допустим нас интересует вероятность того, что за данный промежуток времени $(0, t)$ произойдёт $r$ $(r = 0 \div \infty)$ событий.
При этом выполняются следующие предположения:

- Событие в момент $t$ не зависит от истории, т.е. событий до момента $t$
- Вероятность события за малый интервал времени $\delta t$ пропорциональна длительности интервала: $\delta P_1(t \div t + \delta t) = \nu \delta t + \mathrm{o}(\delta t)$
- Вероятность двух или большего числа событий за тот же промежуток мала: $\delta P_{\geq 2}(t \div t + \delta t) = 0 + \mathrm{o}(\delta t)$

$\underline{r = 0}$: за промежуток времени $(0, t + \delta t)$ не произойдёт ни одного события, если не будет событий в интервалах $(0, t)$ и $(t, t + \delta t)$:

$$P_0(t + \delta t) = P_0(t)(1 - \nu \delta t + \mathrm{o}(\delta t)) \to \frac{P_0(t + \delta t) - P_0(t)}{\delta t} \underset{\delta t \to 0}{=} -\nu P_0(t)$$

$$\frac{dP_0(t)}{dt} = -\nu P_0(t),\ P_0(t) = Ae^{-\nu t},\ A = P_0(0) = 1,\ P_0(t) = e^{-\nu t}$$

# Дискретные распределения IV

$\underline{r = 1:}$ за промежуток времени $(0, t + \delta t)$ произойдёт одно событие, либо в интервале $(0, t)$ либо в $(t, t + \delta t)$:

$$P_1(t + \delta t) = P_1(t)(1 - \nu\delta t + \mathrm{o}(\delta t)) + P_0(t)(\nu\delta t + \mathrm{o}(\delta t))$$

$$\frac{dP_1(t)}{dt} = -\nu P_1(t) + \nu e^{-\nu t},\ P_1(t) = \nu t e^{-\nu t}$$

$\underline{r > 1:}$

$$\frac{dP_r(t)}{dt} = -\nu P_r(t) + \nu P_{r-1}(t),\ P_r(\nu t) = \frac{(\nu t)^r}{r!} e^{-\nu t}$$

**Распределение Пуассона**

Распределение Пуассона для случайной величины $r$ содержит единственный параметр $\mu$ и описывается формулой:

$$P_r(\mu) \equiv P(r|\mu) = \frac{\mu^r}{r!} e^{-\mu}$$

# Дискретные распределения V

Замечание 1: Если 2 независимые случайные величины распределёны по Пуассону, $P^{(1)}(r_1|\mu_1)$ и $P^{(2)}(r_2|\mu_2)$, то случайная величина $r = r_1 + r_2$ тоже распределена по Пуассону $P(r|\mu_1 + \mu_2)$:

$$P(r) = \sum_{r_1+r_2=r} P^{(1)}(r_1)P^{(2)}(r_2) = \sum_{r_1=0}^{r} P^{(1)}(r_1)P^{(2)}(r-r_1) = \sum_{r_1=0}^{r} e^{-\mu_1}\frac{\mu_1^{r_1}}{r_1!}\cdot e^{-\mu_2}\frac{\mu_2^{r-r_1}}{(r-r_1)!} =$$

$$= \frac{e^{-(\mu_1+\mu_2)}}{r!}\sum_{r_1=0}^{r}\underbrace{\frac{r!}{r_1!(r-r_1)!}}_{C_r^{r_1}}\mu_1^{r_1}\mu_2^{r-r_1} = \frac{(\mu_1+\mu_2)^r}{r!}e^{-(\mu_1+\mu_2)} = P(r|\mu_1+\mu_2)$$

Замечание 2: Пусть в интервале времени $(0, t)$ источник в среднем испускает $\mu$ частиц. Разобьем этот интервал на (достаточно большое число) $N$ равных интервалов, в каждом из которых есть вероятность $p$ наблюдать одну частицу (вероятность наблюдать более одной частицы мала). Тогда вероятность наблюдать $r$ частиц за $N$ интервалов равна $P(r|p, N) = C_N^r p^r (1-p)^{N-r}$, при этом $p = \mu/N$. При фиксированном $\mu = Np$ и $r$ рассмотрим предел $N \to \infty$:

$$P(r|p,N) = \frac{N(N-1)\dots(N-r+1)}{r!}\left(\frac{\mu}{N}\right)^r\left(1-\frac{\mu}{N}\right)^{N-r} =$$

$$= \frac{(1-\frac{1}{N})\dots(1-\frac{r-1}{N})}{(1-\frac{\mu}{N})^r}\frac{\mu^r}{r!}\left(1-\frac{\mu}{N}\right)^N \underset{N\to\infty}{\to} \frac{\mu^r}{r!}e^{-\mu}$$

# Непрерывные распределения I

Пусть $X$ – случайная величина, принимающая любые значения в интервале $[a, b]$. Вероятность того, что значение $X$ заключено в малом интервале $[x, x + dx]$:

$$P(x \leq X \leq x + dx) = f(x)dx \geq 0, \ \int\limits_a^b f(x)dx = 1$$

где $f(x)$ - функция плотности вероятности (ф.п.в.).
Функцией распределения $F(x)$ (или функцией распределения накопленной вероятности) случайной величины $X$ называется неубывающая функция $F : [a, b] \to [0, 1]$, при каждом $x \in [a, b]$ равная вероятности $X$ принять значение $a < X < x$:

$$F(x) = P(a < X < x), \ F(x) = \int\limits_a^x f(x')dx', \ f(x) = \frac{dF(x)}{dx},$$

причём $F(a) = 0$ и $F(b) = 1$.
<u>Замечание:</u> Любая отнормированная в определённой области $D \subset R$ функция $f(x)$ ($\int\limits_D f(x)dx = 1$) может рассматриваться как ф.п.в. для некоторой случайной величины $X$ определённой в области $D$.

# Непрерывные распределения II

### Равномерное на отрезке $[a, b]$

$$f(x|a,b) = \begin{cases} 0, & \text{если } x < a \\ \frac{1}{b-a}, & \text{если } a \le x \le b \\ 0, & \text{если } x > b \end{cases} \qquad F(x|a,b) = \begin{cases} 0, & \text{если } x < a \\ \frac{x-a}{b-a}, & \text{если } a \le x \le b \\ 0, & \text{если } x > b \end{cases}$$

Замечание: Равномерное на отрезке $[0, 1]$ распределение ($\gamma(0, 1)$ или просто $\gamma$) выступает в роли основного строительного блока при генерировании случайных величин с произвольной ф.п.в.

# Непрерывные распределения III

## Экспоненциальное, $x \in [0, +\infty)$

$$f(x|\lambda) = \begin{cases} 0, & \text{если } x < 0 \\ \frac{1}{\lambda} e^{-\frac{x}{\lambda}}, & \text{если } x \geq 0 \end{cases} \quad F(x|\lambda) = \begin{cases} 0, & \text{если } x < 0 \\ 1 - e^{-\frac{x}{\lambda}}, & \text{если } x \geq 0 \end{cases}$$

Замечание: Экспоненциальное распределение описывает закон уменьшения частиц в пучке (за счёт поглощения или рассеяния) при его прохождении через слой вещества:

$$N(x) = N_0 e^{-\frac{x}{\lambda}},\ x > 0,$$

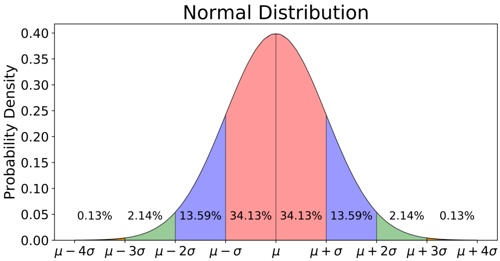
где $\lambda$ - длина ослабления.

## Нормальное (Гауссово), $x \in (-\infty, +\infty)$

$$f(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$F(x|\mu,\sigma) = \Phi\left(\frac{x-\mu}{\sigma}\right),\ \Phi(z) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{z} e^{-t^2/2} dt$$

- $\mu$ - среднее (математическое ожидание)
- $\sigma$ - стандартное отклонение



# Многомерные распределения

Событие может описываться определённым числом ($n$) случайных величин $\vec{x} = (x_1, ..., x_n)$, в этом случае совместная ф.п.в. $f(x_1, ..., x_n) : R^n \to R$ становится функцией многих переменных, $f(\vec{x}) \geq 0,\ \forall \vec{x} \in R^n$ и $\int\limits_{R^n} f(\vec{x'})d\vec{x'} = 1$. Если все случайные величины являются независимыми, то $f(\vec{x})$ распадается на произведение ф.п.в. каждой из $n$ случайных величин:

$$f(\vec{x}) = \prod_{i=1}^{n} f_i(x_i),\ \ f_i(x_i) > 0\ \forall x_i \in R,\ \int\limits_R f_i(x'_i)dx'_i = 1$$

Пример: Распределение атомов идеального газа по скоростям $\vec{v} = (v_x, v_y, v_z)$ (распределение Максвелла):

$$f(v_x, v_y, v_z) = \frac{1}{\sqrt{2\pi}\sqrt{\frac{kT}{m}}} e^{-\frac{v_x^2}{2\left(\sqrt{\frac{kT}{m}}\right)^2}} \times \frac{1}{\sqrt{2\pi}\sqrt{\frac{kT}{m}}} e^{-\frac{v_y^2}{2\left(\sqrt{\frac{kT}{m}}\right)^2}} \times \frac{1}{\sqrt{2\pi}\sqrt{\frac{kT}{m}}} e^{-\frac{v_z^2}{2\left(\sqrt{\frac{kT}{m}}\right)^2}} =$$

$$= \left(\frac{m}{2\pi kT}\right)^{\frac{3}{2}} e^{-\frac{m\vec{v}^2}{2kT}}$$