Теория вероятности

Храбров Александр Игоревич

2 марта 2023 г.

Содержание

1. Элементарная теория вероятностей

1.1. Основные понятия

***Определение* 1.1.** $\Omega = \{\omega_1, \dots, \omega_n\}$ – пространство элементарных событий (исходов).

1. равновозможные
2. несовместные
3. одно всегда реализуется

***Определение* 1.2.** Событие $A \subset \Omega$

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

Свойства. вероятности

1. $P(\varnothing) = 0,\ P(\Omega) = 1,\ P(A) \in [0, 1]$
2. Если $A \cap B = \varnothing$, то $P(A \cup B) = P(A) + P(B)$
3. $\underbrace{P(A \cup B)}_{=P(A)+P(B\setminus(A\cap B))} = P(A) + P(B) - P(A \cap B)$
4. $P(\overline{A}) = 1 - P(A)$, где $\overline{A} = \Omega \setminus A$
5. $P(A_1 \cup A_2 \dots \cup A_m) = \sum_{i=1}^{m} P(A_i) - \sum_{i\neq j} P(A_i \cap A_j) + \sum_{i\neq j,\ i\neq k,\ j\neq k} P(A_i \cap A_j \cap A_k) - \dots + (-1)^{m-1} \cdot P(A_1 \cap \dots \cap A_m)$ – формула включений-исключений.

   Доказательство. Индукция по m.

   База $m = 2$.

   Переход $m \to m + 1$:

   $B_j = A_j \cup A_{m+1}$

   $P(\underbrace{A_1 \cup \dots \cup A_m}_{=:B} \cup A_{m+1}) = P(B \cup A_{m+1}) = \underbrace{P(B)}_{\text{это умеем расписывать по инд. предп.}} + P(A_{m+1}) - P(B \cap A_{m+1}) =$

   $= \sum_{j=1}^{m+1} P(A_j) - sum_{i\neq j}^{m} P(A_i \cap A_j) + \sum_{i\neq j\neq k}^{m} P(A_i \cap A_j \cap A_k) - \underbrace{P(A_{m+1} \cap B)}_{=P(B_1 \cup B_2 \cdots \cup B_m)}$, где $B_i := A_i \cap A_{m+1}$. □

6. $P(A \cup B) \leq P(A) + P(B)$

   $P(A_1 \cup \dots \cup A_m) \leq \sum_{j=1}^{m} P(A_j)$

***Определение* 1.3. Условная вероятность**.

$B \neq \varnothing,\ P(B) > 0$.

Знаем, что выполнилось событие B, хотим узнать вероятность наступления A.

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

Свойства.

1. $P(A|A) = 1$, если $B \subset A$, то $P(A|B) = 1$
2. Если $A_1 \cap A_2 = \varnothing$, то $P(A_1 \cup A_2 | B) = P(A_1|B) + P(A_2|B)$

   В частности: $P(A|B) + P(\overline{A}|B) = 1$

`Замечание.` $P(A|B) + P(A|\overline{B})$ не обязана быть 1.

Пример: игральный кубик, B – выпало четное число, A – выпало кратное трем.

$P(A|B) = \frac{1}{3},\ P(A|\overline{B}) = \frac{1}{3}$

Теорема 1.1. Формула полной вероятности.

Пусть $\Omega = \bigsqcup_{j=1}^{m} B_j,\ P(B_j) > 0.$

Тогда $P(A) = \sum_{j=1}^{m} P(A|B_j) \cdot P(B_j)$

Доказательство. $\sum_{j=1}^{m} \underbrace{P(A|B_j)}_{=\frac{P(A\cap B_j)}{P(B_j)}} \cdot P(B_j) = \sum_{j=1}^{m} P(A \cap B_j) = P(A \cap \bigsqcup_{j=1}^{m} B_j) = P(A)$ □

Пример. I. 3 белых шара, 5 черных шаров

II. 5 белых, 5 черных

2 шара из I положили в II, затем вынули 1 шар из II, P(вынули белый) =?

A – вынули из II белый шар.

$B_0,\ B_1,\ B_2$, где B_j – переложили j белых шаров из I в II.

Тогда $P(A|B_0) = \frac{5}{12},\ P(A|B_1) = \frac{1}{2},\ P(A|B_2) = \frac{7}{12}$.

$P(B_0) = \frac{C_5^2}{C_8^2} = \frac{5}{14}$

$P(B_1) = \frac{15}{C_8^2} = \frac{15}{28}$

$P(B_2) = \frac{C_3^2}{C_8^2} = \frac{3}{28}$

Подставляем в формулу:

$P(A) = \frac{331}{336}$

Теорема 1.2. Формула Байеса.

Пусть $P(A) > 0,\ P(B) > 0$, тогда $P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)}$

Доказательство. Расписываем $P(A|B)$, получаем в правой части: $\frac{P(A\cap B)}{P(B)} \cdot P(B) \cdot \frac{1}{P(A)}$. □

Теорема 1.3. Байеса.

Пусть $P(A) > 0,\ P(B_j) > 0,\ \Omega = \bigsqcup_{j=1}^{m} B_j$, тогда

$P(B_j|A) = \frac{P(A|B_j) \cdot P(B_j)}{P(A|B_1)P(B_1) + \dots + P(A|B_m)P(B_m)}$

Пример. Есть 2 монеты (одна симметричная, вторая $P(\text{орла}) = \frac{1}{3},\ P(\text{решка}) = \frac{2}{3}$). Взялу наугад монету, побросили и выпал орел. Какова вероятность, что мы взяли симметричную монету?

A – выпал орел, B – монета симметричная ($\overline{B}$ – монета кривая).

$P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\overline{B})P(\overline{B})} = \frac{\frac{1}{2} \cdot \frac{1}{2}}{\frac{1}{2} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{2}} = \frac{3}{5}$

***Определение* 1.4. Независимые события.**

Рассуждения: A не зависит от B, если $P(A) = P(A|B) = \frac{P(A\cap B)}{P(B)}$.

Опр. $A,\ B$ независимые события, если $P(A \cap B) = P(A) \cdot P(B)$

***Определение* 1.5.** События $A_1, A_2, \dots, A_m$ – независимы в совокупности, если

$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \dots \cdot P(A_{i_k})$ – для любых индексов i_j.

`Замечание.` Независимость в совокупности $\implies$ попарная независимость.

Наоборот неверно.

Пример. Есть два игральных кубика.

A – на первом кубике выпало четное число.

B – на втором выпало четное число.

C – сумма на кубиках четная.

Пространство элементарных исходов это все пары (i,j), где $i,j \in \{1,2,3,4,5,6\}$, $\#\Omega = 36$.

$P(A) = \frac{1}{2},\ P(B) = \frac{1}{2},\ P(C) = \frac{1}{2}$.

$A \cap B = A \cap C = B \cap C = A \cap B \cap C$.

$P(A \cap B) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(A) \cdot P(B)$, остальные равенства тоже выполняются $\implies$ попарная независимость.

$P(A \cap B \cap C) = \frac{1}{4} \neq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = P(A) \cdot P(B) \cdot P(C) \implies$ нет независимости в совокупности.

Упражнение. Д-ть, что $A_1, \dots A_m$ независимы в совокупности $\Leftrightarrow P(B_1 \cap B_2 \cap \dots \cap B_m) = P(B_1) \dots P(B_m)$, где $B_j = A_j$ или $\overline{A_j}$ (все 2^m равенств).

`Замечание.` Небольшое обобщение.

$\Omega = \{\omega_1, \dots, \omega_n\}$ – пр-во элементарных исходов.

Также у нас есть $p_1, \dots p_n : \ \sum_{i=1}^n p_i = 1,\ \forall i :\ p_i \geq 0$.

$P(A) = \sum_{j:\ \omega_j \in A} p_j$.

Теорема 1.4. Схема Бернулли.

орел = успех = 1.

решка = неудача = 0.

$P(\text{орел}) = p,\ 0 \leq p \leq 1$

$P(\text{решка}) = 1 - p$

Бросаем монету n раз, получаем последовательность исходов:

$\Omega = \{x_1, x_2, \dots, x_n\} :\ x_j = 0$ или 1.

$\omega = (x_1, x_2, \dots, x_n),\ P(\{\omega\}) = p^{\#i:\ x_i=1} \cdot q^{\#i:\ x_i=0} = p^{\sum x_i} \cdot q^{n - \sum x_i}$

Хотим узнать:

$P(\text{выпало ровно } k \text{ орлов}) = C_n^k p^k q^{n-k}$

$P(i\text{-ое подбрасывание } = \text{ орел})$ – независимые в совокупности по $i = 1, 2, \dots, n$.

Теорема 1.5. Полиномиальная схема.

$p_1, p_2, \dots, p_m : \sum p_i = 1$.

$P(x_i = k) = p_k$, где $x_i \in \{1, 2, \dots, m\}$

$\Omega = \{(x_1, x_2, \dots, x_n)\},\ \omega = (x_1, x_2, \dots, x_n)$

$P(\{\omega\}) = p_1^{\#\{i:x_i=1\}} \cdot \dots \cdot p_m^{\#\{i:x_i=m\}}$

$k_1 + k_2 + \dots + k_m = n$

$$P(k_1 \text{ раз выпало 1, } k_2 \text{ раз выпало 2, ...}) = \underbrace{\binom{n}{k_1, k_2, \dots, k_m}}_{= \frac{n!}{k_1! \dots k_m!}} \cdot p_1^{k_1} \cdot \dots \cdot p_m^{k_m}$$

Теорема 1.6. Эрдёша-Мозера

Рассмотрим турнир на n команд. При каком наибольшем k можно всегда выбрать команды $A_1, A_2 \dots A_k$, так, что A_i выиграла у A_j, если $i < j$? При $k \leqslant 1 + [2\log_2 n]$

Доказательство. Рассмотрим случайный турнир(Всего $\binom{n}{2}$, тогда Всего $2^{\binom{n}{2}}$ разных турниров. Случайный - берём из этой кучи наугад).

P(А выиграла у В) $= \frac{1}{2}$. Рассмотрим $A_1, A_2, \dots A_k$ команды.

$P(A_1, A_2 \dots A_k$ подходят$) = (\frac{1}{2})^{\binom{k}{2}}$.

$P(A_1, A_2 \dots A_k$ можно переименовать, так, что они подошли$) \leqslant \frac{k!}{2^{\binom{k}{2}}}$

P(какие-то k команд подошли) $\leqslant \binom{n}{k} \cdot \frac{k!}{2^{\binom{k}{2}}}$

Нужно понять, что если $k \geqslant [2+2\log_2 n]$, то $\binom{n}{k}\frac{k!}{2^{\binom{k}{2}}} < 1$. $\binom{n}{k}\frac{k!}{2^{\binom{k}{2}}} = \frac{n(n-1)(n-2)\dots(n-k+1)}{2^{\frac{k(k-1)}{2}}} < \frac{n^k}{(2^{\frac{k-1}{2}})^k} = (\frac{n}{2^{\frac{k-1}{2}}})^k$

Надо понять, что $2^{\frac{k-1}{2}} \geqslant n \Leftrightarrow \frac{k-1}{2} \geqslant \log_2 n$ □

1.2. Предельные теоремы для схем Бернулли

***Определение* 1.6.** Схема Бернулли с вероятностью успеха $p \in (0,1)$. S_n - число успехов при n испытаниях. $P(S_n = k) = \binom{n}{k}p^k q^{n-k}$

Что будет больше $P(S_{1000} = 220)$ при $p = \frac{1}{5}$ или $P(S_{2000}) = 360$ при $p = \frac{1}{6}$. Точные вычисления дают 0.008984 и 0.006625 соответственно.

Теорема 1.7. Пуассона

Схема Бернулли с n испытаниями и вероятностью успеха p_n - зависит от n. Если $np_n \to \lambda > 0$. Тогда $P(S_n = k) \to \frac{\lambda^k}{k!}e^{-\lambda}$

`Замечание`. Если $np_n = \lambda$, то теорема верна при $k = o(\sqrt{n})$

Доказательство. $P(S_n = k) = \binom{n}{k}p^k(1-p)^{n-k} = \frac{n(n-1)\dots(n-k+1)}{k!}p_n^k(1-p_n)^{n-k}. \sim \frac{n^k}{k!}p_n^k(1-p)^{n-k} \sim \frac{\lambda^k}{k!}(1-p_n)^n$. Прологарифмируем: $\ln(1-p_n)^n = n\ln(1-p_n) \sim -np_n \sim -\lambda$

Доказательство замечания:

$\ln(1-p_n)^k = k\ln(1-p_n) \sim -kp_n \to 0$

Осталось понять, что $\frac{n(n-1)\dots(n-k+1)}{n^k} \to 1$ при $k = o(\sqrt{n})$.

$\frac{n(n-1)\dots(n-k+1)}{n^k} = 1 \cdot (1-\frac{1}{n}) \cdot \dots \cdot (1-\frac{k-1}{n}) \geqslant 1 - \frac{1}{n} - \dots - \frac{k-1}{n} = 1 - \frac{k(k-1)}{2n} \to 1$

Неравенство $(1-x_1)\dots(1-x_k) \geqslant 1 - x_1 - x_2 - \dots - x_k$ при $0 \leqslant x_i \leqslant 1$ - индукция. □

Теорема 1.8. Прохорова

Если $\lambda = np$, то $\sum_{i=0}^{+\infty} |P(S_n = k) - \frac{\lambda^k}{k!}e^{-\lambda}| \leqslant \frac{2\lambda}{n} \cdot \min(2, \lambda)$

Пример. Игра в рулетку: 36 чисел и ноль.

$p = \frac{1}{37}, n = 111, np = 3 = \lambda$.

$P(S_{111} = 3) = \binom{111}{3}(\frac{1}{37})^3(1-\frac{1}{37})^{111-3} = 0.227127$

Из Пуассона $\frac{\lambda^3}{3!}e^{-\lambda} = 0.224$

Видим, что приближение хорошее.

$P(\text{выигрыш}) = 1 - P(S_{111} = 0) - P(S_{111} = 1) - P(S_{111} = 2) - P(S_{111} = 3) = 1 - \frac{\lambda^0}{0!}e^{-\lambda} - \frac{\lambda^1}{1!}e^{-\lambda} - \frac{\lambda^2}{2!}e^{-\lambda} - \frac{\lambda^3}{3!}e^{-\lambda} = 0.352754$

А по формулам 0.352768

Теорема 1.9. Локальная предельная теорема Муавра-Лапласа

Схема Бернулии с вероятностью успеха $p \in (0,1),\ q = 1-p,\ x = \frac{k-np}{\sqrt{npq}}$.

$P(S_n = k) \sim_{n\to+\infty} \frac{1}{\sqrt{2\pi npq}} e^{\frac{-x^2}{2}}$

Если $|x| \leqslant T$, то есть равномерность.

Доказательство. $k = np + x\sqrt{npq} \geqslant np - T\sqrt{npq} \to +\infty$

$n-k = nq - x\sqrt{npq} \geqslant nq - T\sqrt{npq} \to +\infty$

$P(S_n = k) = \binom{n}{k} p^k q^{n-k} = \frac{n!}{k!(n-k)!} p^k q^{n-k}$

Напишем формулу Стирлинга:

$\frac{n^n e^{-n}\sqrt{2\pi n} p^k q^{n-k}}{k^k e^{-k}\sqrt{2\pi n}(n-k)^{n-k} e^{-(n-k)}\sqrt{2\pi(n-k)}} = \frac{p^k q^{n-k}}{(\frac{k}{n})^k \cdot (\frac{n-k}{n})^{n-k} \cdot \sqrt{2\pi \frac{k}{n}(1-\frac{k}{n})n}}$

$\frac{k}{n} \to p, \frac{n-k}{n} \to q$

Поэтому остаётся доказать, что $\frac{p^k q^{n-k}}{(\frac{k}{n})^k \cdot (\frac{n-k}{n})^{n-k}} \to e^{\frac{-x^2}{2}}$

$k\ln\frac{k}{n} + (n-k)\ln\frac{n-k}{n} - k\ln p - (n-k)\ln q \to \frac{x^2}{2}$

$\alpha = \frac{k}{n} \to p, \beta = \frac{n-k}{n} \to q$

$n\alpha\ln\alpha + n\beta\ln\beta - n\alpha\ln p - n\beta\ln q \to \frac{x^2}{2} = n\alpha\ln\frac{\alpha}{p} + n\beta\ln\frac{\beta}{q}$

Напишем Тейлора:

$\frac{\alpha}{p} = 1 + x\sqrt{\frac{q}{p}}\frac{1}{\sqrt{n}}$

$\frac{\beta}{q} = 1 - x\sqrt{\frac{p}{q}}\frac{1}{\sqrt{n}}$

$\ln\frac{\alpha}{p} = \ln(1 + x\sqrt{\frac{q}{p}\frac{1}{\sqrt{n}}}) = x\sqrt{\frac{q}{p}} - \frac{1}{2}x^2\frac{q}{p}\frac{1}{n} + o(\frac{1}{n})$

$\ln\frac{\beta}{q} = \ln(1 - x\sqrt{\frac{p}{q}\frac{1}{\sqrt{n}}}) = -x\sqrt{\frac{p}{q}\frac{1}{\sqrt{n}}} - \frac{1}{2}x^2\frac{p}{q}\frac{1}{n} + o(\frac{1}{n})$

Сумма равна: $x\sqrt{pq}\sqrt{n} + x^2q - \frac{1}{2}x^2q + o(\frac{1}{n}) - x\sqrt{pq}\sqrt{n} + x^2p - \frac{1}{2}x^2p + o(\frac{1}{n}) = x^2(\frac{q}{2} + \frac{p}{2}) + o(1) = \frac{x^2}{2} + o(1)$

Извините, это было очень больно... □

`Замечание`. Если $\varphi(n) = o(n^{\frac{2}{3}})$ и $|k - np| \leqslant \varphi(n)$, то теорема тоже верна

Пример. Всё та же рулетка. $n = 222, k = 111$. Пытаемся ставить на четное/нечётное(кроме 0). $p = \frac{18}{37}$

$P(S_{222} = 111) \approx \frac{1}{\sqrt{2\pi npq}} e^{\frac{-x^2}{2}} \approx 0.049395\ldots$

Если считать точно, то получим $0.0493228\ldots$

Теорема 1.10. Интегральная теорема Муавра-Лапласа

$0 < p < 1$. $P(a < \frac{S_n - np}{\sqrt{npq}} \leqslant b) \to_{n\to\infty} \frac{1}{\sqrt{2\pi}} \int_a^b e^{\frac{-t^2}{2}}\,dt$

Стремление равномерно по $a,\ b \in \mathbb{R}$.

Теорема 1.11. Берри-Эссеена

Обозначение: $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}}\,dt$

$\left|P(\frac{S_n-np}{\sqrt{npq}} \leqslant x) - \Phi(x)\right| \leqslant \frac{p^2+q^2}{\sqrt{npq}} \cdot \frac{1}{2}$

`Замечание.` Константа лучше, чем $\frac{c}{\sqrt{n}}$ не бывает.

Пример. $p = q = \frac{1}{2}$. Вопрос: $P(S_{2n} = n) = \binom{2n}{n}\frac{1}{2^{2n}} \sim \frac{4^n}{\sqrt{\pi n}}\frac{1}{4^n} = \frac{1}{\sqrt{\pi n}}$.

Но $P(S_{2n} < n) = P(S_{2n} > n)$.

Тогда $P(S_{2n} \leqslant n) = \frac{1+P(S_{2n}=n)}{2} = \frac{1}{2} + \frac{1}{2\sqrt{\pi n}} + o(\frac{1}{\sqrt{n}})$

Муавра-Лаплас нам говорит, что $P(S_{2n} \leqslant n) \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{0} e^{-\frac{t^2}{2}}\,dt = \frac{1}{2}$

Но $P(S_{2n} \leqslant n) = \frac{1}{2} + \frac{1}{2\sqrt{\pi n}} + o(\frac{1}{\sqrt{n}})$

`Замечание.` $P(a < S_n \leqslant b) = P(\frac{a-np}{\sqrt{npq}} < \frac{S_n-np}{\sqrt{npq}} \leqslant \frac{b-np}{\sqrt{npq}}) \to \Phi(\frac{b-np}{\sqrt{npq}}) - \Phi(\frac{a-np}{\sqrt{npq}})$

Отсюда получили, что лучше всего писать полуцелые a и b.

Пример. Задача о театре

Есть театр и 2 входа. У каждого входа расположен гардероб. В театре $n = 1600$ мест. Хотим сделать размер гардероба как можно меньше, но чтобы переполнения случались как можно реже.

Пусть c мест в итоге в гардеробе.

$p = q = \frac{1}{2}$. Нужно, чтобы $n - c \leqslant S_n \leqslant c$.

$P(n-c \leqslant S_n \leqslant c) = P(\frac{n-c-\frac{n}{2}}{\sqrt{n\cdot\frac{1}{4}}}) \leqslant \frac{S_n-\frac{n}{2}}{\sqrt{n\cdot\frac{1}{4}}} \leqslant \frac{c-\frac{n}{2}}{\sqrt{n\cdot\frac{1}{4}}} = P(\frac{800-c}{20}) \leqslant \frac{S_n-800}{20} \leqslant \frac{c-800}{20} \to \Phi(\frac{800-c}{20}) - \Phi(\frac{c-800}{20}) = \frac{1}{\sqrt{2\pi}}\int_{\frac{800-c}{20}}^{\frac{c-800}{20}} e^{-\frac{t^2}{2}}\,dt > \frac{29}{30}$

$\Phi_0(\frac{c-800}{20}) > \frac{29}{60}$. Тогда $c = 843$.

Пример. Случайное блуждание на прямой

Есть прямая, будем считать, что у нас блуждания исключительно по целым точкам.

В каждой точке подбрасываем монетку. С вероятностью p идём вперёд, q - идём назад.

$a_{n+1} = a_n + 1$ с вероятностью p

$a_{n+1} = a_n - 1$ с вероятностью q

$a_n \equiv n \mod 2$

Это почти похоже на схему Бернулли: $2S_n - n = a_n$

$P(a_n = k) = P(S_n = \frac{n+k}{2}) = \begin{cases} 0, \text{ если } n \not\equiv k \mod 2 \\ \binom{n}{\frac{n+k}{2}} p^{\frac{n+k}{2}} q^{\frac{n-k}{2}}, \text{ иначе} \end{cases}$

Теорема 1.12. ван дер Вардена

Рассмотрим числа $1, 2 \ldots k$ и покрасим их в 2 цвета.

Тогда существует k_n, такое, что, если $k > k_n$, то при любой раскраске найдётся одноцветная n-членная арифметическая прогрессия.

Теорема 1.13. Эрдеша-Радо

$k_{n+1} \geqslant \sqrt{n \cdot 2^{n+1}}$

Доказательство. $A_1, A_2 \ldots A_m$ - все арифметические прогрессии длины $n+1$ из чисел $1, 2 \ldots k$.

С разностью 1 $k - n$ прогрессий.

С разностью 2 $k-2n$ прогрессий.

$\ldots$

С разностью $k-[\frac{k}{n}]\cdot n$ прогрессий с разностью $[\frac{k}{n}]$

Тогда $m=(k-n)+(k-2n)+\ldots=k\cdot[\frac{k}{n}]-n\cdot\frac{[\frac{k}{n}]\cdot([\frac{k}{n}]+1)}{2}=[\frac{k}{n}](k-\frac{1}{2}n([\frac{k}{n}]+1))<\frac{k}{n}(k-\frac{1}{2}\cdot n\cdot\frac{k}{n})=\frac{k^2}{2n}$ - это оценка сверху.

$P(A_i$ - одноцветная$)=2\cdot\frac{1}{2^{n+1}}=\frac{1}{2^n}$ (2 - выбор цвета).

$P($какое-то A_i - одноцветно$)=\sum_{i=1}^m P(A_i$ - одноцветно$)=\frac{m}{2^n}<\frac{k^2}{2n}\cdot\frac{1}{2^n}=(\frac{k}{\sqrt{2^{n+1}\cdot n}})^2\leqslant 1$ (если так, то найдётся, на которой не выполнится) $\square$

2. Общая теория вероятностей

2.1. Колмогоровская модель теории вероятности

***Определение* 2.1.** $(\Omega, \mathcal{F}, P)$ - вероятностное пространство.

Ω - множество или пространство элементарных исходов.

$\mathcal{F}$ - σ-алгебра подмножеств Ω. Элементы $\mathcal{F}$ - случайный события.

P - мера на $\mathcal{F}$ с условием $P(\Omega) = 1$.

`Замечание.` Если Ω не более чем счётно, то можно взять $\mathcal{F} = 2^\Omega$

***Определение* 2.2.** Условная вероятность. A - событие, такое, что $P(A) > 0$. Тогда $P(B|A) = \frac{P(B\cap A)}{P(A)}$, где $A, B \in \mathcal{F}$.

***Определение* 2.3.** Независимые события A и B. Если $P(A \cap B) = P(A) \cdot P(B)$

***Определение* 2.4.** Независимость в совокупности $A_1, A_2 \dots A_n$. $P(A_{i_1} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \cdot \dots \cdot P(A_{i_k})$ для всевозможных наборов индексов.

***Определение* 2.5.** Последовательность событий $A_1, A_2 \dots$ независимы - любой конечный набор событий независим в совокупности.

Лемма. Бореля-Кантелли

$A_1, A_2, \dots$ случайные события.

1. Если $\sum_{n=1}^{\infty} P(A_n) < +\infty$, то вероятность, что случилось бесконечное число из них равна 0.
2. Если $A_1, A_2, \dots$ независимы и $\sum_{n=1}^{\infty} P(A_n) = +\infty$, тогда

   $P(\text{случилось бесконечное число из } A_n) = 1$.

Доказательство. $B = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$.

$\omega \in B \iff \omega \in \bigcup_{k=n}^{\infty} A_k \ \forall n \iff w \in A_k$ для бесконечного количества индексов k.

Док-во этого факта:

1. $\Leftarrow$: Лежит в каждом объединении, значит лежит в B.
2. $\Rightarrow$: ω лежит в пересечении. Пусть лежит в конечном - возьмём самый большой номер и получим противоречие.

Док-во теоремы:

1. $P(B) = 0$ - хотим доказать.

   $B \subset \bigcup_{k=n}^{\infty} A_k \Rightarrow P(B) \leqslant P(\bigcup_{k=n}^{\infty} A_k) \leqslant \sum_{k=n}^{\infty} P(A_k)$, а это хвост сходящегося ряда, а он стремится к нулю.

2. Давайте смотреть на $\bar{A}_1, \bar{A}_2, \dots$ - независимые события (следует из упражнения с прошлой лекции).

   $P(\bigcap_{k=1}^{n} \bar{A}_k) = \prod_{k=1}^{n} P(\bar{A}_k) \to_{n\to\infty} \prod_{k=1}^{\infty} P(\bar{A}_k)$

   Но всё вложено по убыванию, по монотонности меры получаем $P(\bigcap_{k=1}^{\infty} \bar{A}_k) = \prod_{k=1}^{\infty} P(\bar{A}_k) = \prod_{k=1}^{\infty}(1 - P(A_k))$

   Прологарифмируем это равенство.

$\ln(P(\bigcap_{k=n}^{\infty}\bar{A}_k)) = \sum_{k=n}^{\infty}\ln(1-P(A_k)) \leqslant \sum_{k=n}^{\infty}(-P(A_k)) = -\infty$ – сумма хвоста расходящегося ряда.

А значит мы логарифмировали $0 \Rightarrow P(\bigcap_{k=n}^{\infty}\bar{A}_k) = 0 \Rightarrow P(\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\bar{A}_k) = 0 \Rightarrow P(\bar{B}) = 0 \Rightarrow$
$\Rightarrow \overline{\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\bar{A}_k} = \bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}A_k = B \Rightarrow P(B) = 1.$

Добавим, что $B = \bigcap_{n=1}^{\infty}B_n$, где $B_1 \supset B_2 \supset \ldots$ и $P(B) = \lim P(B_n) = 1$.

$\square$

Теорема 2.1. Закон нуля и единицы

Если $A_1, A_2 \ldots$ независимы, то $P(B) = 0$ или $P(B) = 1$.

Пример. Испытания Бернулли, успех с вероятностью p,

P(ОРО встречается бесконечное число раз) = ?.

A_n = случилось ОРО на позициях $n, n+1, n+2$.

Тогда $A_1, A_4, A_7, \ldots$ независимы. $P(A_j) = pqp = p^2q > 0$.

Лемма Бореля-Кантелли говорит: бесконечное кол-во A_{3k+1} случится, если $\sum_{k=1}^{\infty}P(A_{3k+1}) = +\infty \implies P$(ОРО встречается бесконечное число раз) = 1.

2.2. Случайные величины

***Определение* 2.6.** $(\Omega, \mathcal{F}, P)$ - вероятностное пространство.

$\xi : \Omega \to \mathbb{R}$ - случайная величина, если это измеримая функция.

***Определение* 2.7.** Распределение случайное величины

P_ξ - вероятностная мера на борелевских подмножествах $\mathbb{R}$

A – борелевское мн-во, $P_\xi(A) = P(\omega \in \Omega \ : \ \xi(\omega) \in A)$

***Определение* 2.8.** Случаный величины ξ и η одинаково распределены, если $P_\xi = P_\eta$

`Замечание.` P_ξ однозначно определяются своими значениями на ячейках.

$P_\xi(a, b] = P_\xi(-\infty, b] - P_\xi(-\infty, a] = P(\xi \leqslant b) - P(\xi \leqslant a)$

***Определение* 2.9.** Функция распределения случайной величины

$F_\xi(x) = P(\xi \leqslant x)$

Свойства.

1. Функция распределения однозначно определяет распределение случайной величины.

   Доказательство: Если у двух случайных величин совпали, то у них одинаковые распределения

2. $0 \leqslant F_\xi(x) \leqslant 1 \, \forall x \in \mathbb{R}$

3. $\lim_{x\to-\infty}F_\xi(x) = 0$

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

   Доказательство: берём $x_n \to -\infty . A_n = \{\xi \leqslant x_n\}$ Тогда $A_{n+1} \subset A_n$. Тогда $\lim_{n\to\infty}P(A_n) = P(\bigcap_{n=1}^{\infty}A_n) = P(\varnothing) = 0$

4. F_ξ монотонно возрастает

5. Непрерывность справа: $\lim_{y\to x+} F_\xi(y) = F_\xi(x)$

   Доказательство: берём y_n убывающие и $y_n \to x$. Тогда $A_n = \{\xi \leqslant y_n\}$. $A_{n+1} \subset A_n$. А тогда $\lim P(A_n) = P(\bigcup_{n=1}^\infty A_n) = P(\xi \leqslant x) = F_\xi(x)$. Но с другой стороны $\lim P(A_n) = \lim P(\xi \leqslant y_n) = \lim F_\xi(y_n)$

6. $\lim_{y\to x-} F_\xi(y) P(\xi < x)$

   Доказательство: берём y_n возрастающие и $y_n \to x$. $B_n = \{\xi \leqslant y_n\}$ и $B_n \subset B_{n+1}$. $\lim P(B_n) = P(\bigcup B_n) = P(\xi < x)$. Но с другой стороны $\lim P(B_n) = \lim F_\xi(y_n)$

7. $F_{\xi+a}(x) = F_\xi(x-a)$

   Доказательство: $\{\xi + a \leqslant x\} = \{\xi \leqslant x - a\}$

8. $F_{c\xi} = F_\xi(\frac{x}{c})$

   Доказательство: $\{c\xi \leqslant x\} = \{\xi \leqslant \frac{x}{c}\}$

`Замечание.` Фукнция, обладающая свойствами 3, 4, 5 - это фукнция распределения некоторой случайной величины.

Доказательство: пусть g - такая функция. Тогда $\nu_g(a,b] = g(b) - g(a)$. Случайная величина $\xi(x) = x$. Тогда $F_\xi = g$

***Определение* 2.10.** Случайная величина имеет дискретное распределение, если её множество значений не более чем счётное.

`Замечание.`

1. $\xi \to \{y_1, y_2, \ldots\}$

   Если $x \neq y_k$, то $P(\xi = x) = 0$, т.е. $P_\xi(\{x\}) = 0$

2. $P_\xi(A) = \sum_{k: y_k \in A} P(\xi = y_k)$. Тут счётное число слагаемых, поэтому сумма корректно определена.

   Распределение однозначно определяется набором вероятностей $P(\xi = y_k)$

3. $F_\xi(x) = \sum_{k: y_k \leqslant x} P(\xi = y_k)$

***Определение* 2.11.** Случайная величина имеет непрерывное распределение, если $P(\xi = x) = 0$

`Замечание.`

1. Это значит, что фукнция распределения непрерывна.

2. Непрерывные распределения бывают не очень хорошими, например Канторова лестница.

***Определение* 2.12.** Случайная величина имеет абсолютно непрерывное распределение, если существует $p_\xi(t) \geqslant 0$, измеримая, т.ч. $F_\xi(x) = \int_{-\infty}^x p_\xi(t)\,dt$ ($p_\xi(t)$ – плотность распределения).

Свойства.

1. $A \subset \mathbb{R}$ – борелевское, то $P_\xi(A) = \int_A p_\xi(t)\,dt$

   Доказательство: слева мера и справа мера. Нужно понять, почему они совпадают на ячейках.

   $P_\xi(a,b] = F_\xi(b) - F_\xi(a) = \int_a^b p_\xi(t)\,dt$

2. $\int_{-\infty}^{+\infty} p_\xi(t)\,dt = 1$

3. p_ξ определена однозначно с точностью до почти везде (из теории меры)

4. F_ξ почти везде диффиренцируема и $F'_\xi(x) = p_\xi(x)$

   Доказательство: а его не будет

Пример. Вероятностные распределения

1. Биномиальное распределение: $\xi \sim Binom(p, n), 0 < p < 1$
   $\xi : \Omega \to \{0, 1, \dots n\}$. $P(\xi = k)\binom{n}{k}p^k(1-p)^{n-k}$
2. Распределение Пуассона: $\xi \sim Poisson(\lambda), \lambda > 0$.
   $\xi : \Omega \to \{0, 1, \dots\}$. $P(\xi = k) = \frac{\lambda^k}{k!}e^{-\lambda}$
3. Геометрическое распределение: $\xi \sim Geom(p), 0 < p < 1$.
   $\xi : \Omega \to \{1, 2, \dots\}$. $P(\xi = k) = p(1-p)^{k-1}$.
4. Дискретные равномерные распределения: $\xi \sim U(\dots)$
   $\xi : \Omega \to \{1, 2, \dots n\}$. $P(\xi = k) = \frac{1}{n}$
5. Непрерывно равномерное распределение: $\xi \sim U([a, b])$
   $\xi : \Omega \to [a, b]$. $p_\xi(t) = \frac{1}{b-a} \cdot \mathbb{1}_{[a,b]}(t)$
6. Нормальное распределение: $\xi \sim \mathcal{N}(a, \sigma^2), a \in \mathbb{R}, \sigma > 0$
   $\xi : \Omega \to \mathbb{R}$. $p_\xi(t) = \frac{1}{\sqrt{2\pi}\cdot\sigma}e^{-\frac{(t-a)^2}{2\sigma^2}}$
   Стандартное нормальное распределение: $\mathcal{N}(0, 1)$
7. Экспонециальное распределение: $\xi \sim Exp(\lambda), \lambda > 0$.
   $\xi : \Omega \to [0, +\infty]$. $p_\xi(t) = \begin{cases} \lambda e^{-\lambda t}, \text{ при } t \geqslant 0 \\ 0, \text{ в других точках} \end{cases}$

`Замечание.`

1. $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}\,dt$.
   На самом деле это функция распределения стандартной нормальной случайной величины.
2. Если $\nu \sim \mathcal{N}(0, 1)$, то $\xi = \sigma\nu + a$. $\xi \sim \mathcal{N}(a, \sigma^2)$
   $F_\xi(x) = P(\sigma\nu + a \leqslant x) = P(\nu \leqslant \frac{x-a}{\sigma}) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{x-a}{\sigma}} e^{-\frac{t^2}{2}}\,dt$
   Замена $t = \frac{s-a}{\sigma}$. Тогда $dt = \frac{ds}{\sigma}$
   Тогда: $\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{x-a}{\sigma}} e^{-\frac{t^2}{2}}\,dt = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x} e^{-\frac{(s-a)^2}{2\sigma^2}}\,ds$

2.3. Совместное распределение

***Определение* 2.13.** Совместное(многомерное) распределение.
$\bar{\xi} = (\xi_1, \xi_2, \dots, \xi_n) : \Omega \to \mathbb{R}^n$
$P_{\bar{\xi}}(A) = P(\bar{\xi} \in A)$, где A - борелевское подмножество $\mathbb{R}^n$

`Замечание.` Совместное распределение однозначно определяет распределение случайной величины, но не наоборот

Пример. $\xi, \eta : \Omega \to \{0, 1\}$ с равными вероятностями.

Если это были независимые подбрасывания: $(\xi, \eta) : \Omega \to \{(0, 0), (0, 1), (1, 0), (1, 1)\}$ с равными вероятностями.

Если $\xi = \eta$, то $(\xi, \eta) : \Omega \to \{(0, 0), (1, 1)\}$.

***Определение* 2.14.** Случайные величины $\xi_1, \xi_2 \dots \xi_n$ независимы, если для любых борелевских подмножеств $A_1, A_2 \dots A_n \subset \mathbb{R}$, события $\{\xi_1 \in A_1\}, \dots, \{\xi_n \in A_n\}$ независимы

`Замечание.` $P(\xi_1 \in A_1, \dots, \xi_n \in A_n) = P(\xi_1 \in A_1) \cdot \ldots \cdot P(\xi_n \in A_n)$

Теорема 2.2. $\xi_1, \xi_2 \dots \xi_n$ независимы $\iff P_{\bar{\xi}} = P_{\xi_1} \times \ldots \times P_{\xi_n}$

Доказательство. 1. $\Leftarrow$ очевидно. $P_{\bar{\xi}}(A_1 \times \ldots \times A_n) = P_{\xi_1}(A_1) \dots P_{\xi_n}(A_n)$

2. $\Rightarrow$. На множествах $A_1 \times \ldots \times A_n$ есть равенство + единственность продолжения.

$\square$

***Определение* 2.15.** Совместная (многомерная) функция распределения.
$\bar{\xi} = (\xi_1 \dots \xi_n)$. $F_{\bar{\xi}} : \mathbb{R}^n \to \mathbb{R}$. и $F_{\bar{\xi}}(\bar{x}) = P(\xi_1 \leqslant x_1, \dots, \xi_n \leqslant x_n)$

Свойства. 1. $0 \leqslant F_{\bar{\xi}} \leqslant 1$

2. Монотонно возрастает по каждой координате

3. $\lim_{x_i \to -\infty} F_{\bar{\xi}}(\bar{x}) = 0$
   $\lim_{x_1, \dots, x_n \to +\infty} F_{\bar{\xi}}(\bar{x}) = 1$

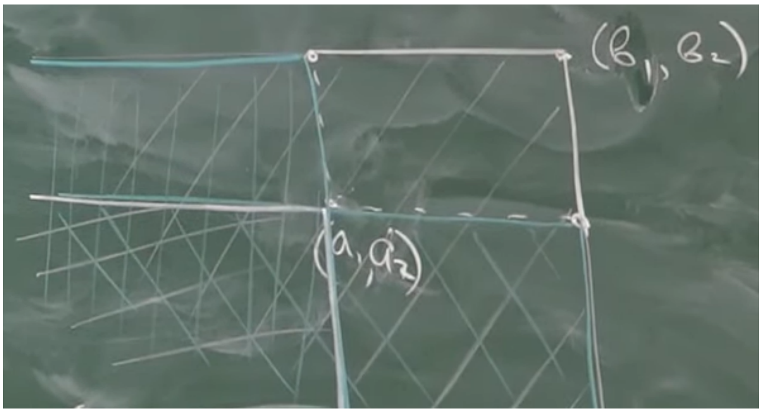
4. $\lim_{x_i \to +\infty} F_{\bar{\xi}}(\bar{x}) = F_{\xi_1, \dots, \xi_{i-1}, \xi_{i+1}, \dots}$

***Определение* 2.16.** Совместная плотность $p_{\bar{\xi}}(\bar{t})$ - неотрицательная измеримая функция, такая, что $F_{\bar{\xi}}(\bar{\xi}) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} p_{\bar{\xi}}(\bar{t})\, dt_n \dots dt_1$

Теорема 2.3. $\xi_1 \dots \xi_n$ независимы $\iff F_{\bar{\xi}}(\bar{x}) = F_{\xi_1}(x_1) \cdot \ldots \cdot F_{\xi_n}(x_n)$

Доказательство. 1. Докажем $\Rightarrow$. Независимость $\Rightarrow (*) P_{\bar{\xi}} = P_{\xi_1} \times \ldots \times P_{\xi_n} \Rightarrow P_{\bar{\xi}}((-\infty, x_1] \times \ldots \times (-\infty, x_n]) = P_{\xi_1}(-\infty, x_1] \cdot \ldots \cdot P_{\xi_n}(-\infty, x_n]$

2. Хотим проверить совпадение на ячейках, чтобы доказать $(*)$ ещё и в другую сторону.



$$P_{\bar{\xi}}((a_1, b_1] \times (a_2, b_2]) = F_{\bar{\xi}}(b_1, b_2) + F_{\bar{\xi}}(a_1, a_2) - F_{\bar{\xi}}(a_1, b_2) - F_{\bar{\xi}}(a_2, b_1) =$$
$$= (F_{\xi_1}(b_1) - F_{\xi_1}(a_1)) \cdot (F_{\xi_2}(b_2) - F_{\xi_2}(a_2)) = P_{\xi_1}(a_1, b_1] \cdot P_{\xi_2}(a_2, b_2]$$

$\square$

Следствие. $\xi_1 \dots \xi_n$ - абсолютно непрерывные случайные величины. Тогда $\xi_1 \dots \xi_n$ независимы $\iff p_{\bar{\xi}}(\bar{t}) = p_{\xi_1}(t_1) \cdot \ldots \cdot p_{\xi_n}(t_n)$

В частности, в случае независимости $\bar{\xi}$ абсолютно непрерывна.

Доказательство. 1. Докажем $\Rightarrow$.

Независимость $\Rightarrow F_{\bar{\xi}}(\bar{x}) = F_{\xi_1}(x_1) \cdot \ldots \cdot F_{\xi_n}(x_n) = \int_{-\infty}^{x_1} p_{\xi_1}(t_1)\,dt_1 \cdot \ldots \cdot \int_{-\infty}^{x_n} p_{\xi_n}(t_n)\,dt_n = \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_n} p_{\xi_1}(t_1) \ldots p_{\xi_n}(t_n)\,dt_n \ldots dt_1$.

Запихали всё под один интеграл, то что под интегралом и есть совместная плотность.

2. Докажем $\Leftarrow$.

Просто проинтегрируем равенство.

$$\int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_n} p_{\bar{\xi}}(\bar{t})\,dt_n \ldots dt_1 = \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_n} p_{\xi_1}(t_1) \ldots p_{\xi_n}(t_n)\,dt_n \ldots dt_1 \underbrace{=}_{\text{по т. Тонелли можно выносить интегралы}} F_{\xi_1}(x_1) \cdot \ldots \cdot F_{\xi_n}(x_n)$$

□

`Замечание.` Напоминание.

Свертка последовательностей: $\{a_n\}, \{b_n\}$ это $\{c_n\}$, такая что $c_n = a_0 b_n + a_1 b_{n-1} + \ldots + a_n b_0$.

Мотивировка: $(\sum_{n=0}^{\infty} a_n z^n) \cdot (\sum_{n=0}^{\infty} b_n z^n) = \sum_{n=0}^{\infty} c_n z^n$ (при наличии хоть каких-нибудь кругов сходимости у обоих рядов).

`Замечание.` **Свертки мер**

μ и ν - конечные меры на борелевских подмножествах $\mathbb{R}$.

$\mu * \nu(A) = \int_{\mathbb{R}} \mu(A - x)\,d\nu(x)$ - это свертка мер, где $(A - x) := \{a - x \mid a \in A\}$.

Свойства. **Свойства свёртки**

1. $\mu * \nu(A) = \int_{\mathbb{R}^2} \mathbb{1}_A(x+y) d\mu(x)\,d\nu(y)$

   Доказательство: $\mu * \nu(A) = \int_{\mathbb{R}} \mu(A - x)\,d\nu(x) = \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbb{1}_{A-x}(y) d\mu(y)\,d\nu(x)$

2. $\mu * \nu = \nu * \mu$

3. $\mu_1 * \ldots * \mu_n(A) = \int_{\mathbb{R}^n} \mathbb{1}_A(x_1 + \ldots + x_n)\,d\mu_1(x_1) \ldots\, d\mu_n(x_n)$

4. $(\mu_1 * \mu_2) * \mu_3 = \mu_1 * (\mu_2 * \mu_3)$

5. $(\mu_1 + \mu_2) * \nu = \mu_1 * \nu + \mu_2 * \nu$

6. δ_x - мера с единичной нагрузкой в точке x. Тогда $\mu * \delta_0 = \mu$.

   Получили линейное пространство относительно $+$ и $*$

   Доказательство: $\mu * \delta_0(A) = \int_R \mu(A - x)\,d\delta_0(x) = \mu A$ - значение подынтегральной функции в точке $x = 0$.

Теорема 2.4. Пусть μ и ν имеют плотности p_μ и p_ν

Тогда $\mu * \nu$ имеет плотность $p(t) = \int_{\mathbb{R}} p_\mu(t - s) p_\nu(s)\,ds$

Доказательство. Возьмём функцию, определяемую этой формулой и проверим, что подходит.

$\int_A p(t)\,dt = \int_A \int_{\mathbb{R}} p_\mu(t-s) p_\nu(s)\,ds\,dt = \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbb{1}_A(t) p_\mu(t-s) p_\nu(s)\,ds\,dt = (*)$.

Положим $u = t - s$. Тогда $(*) = \int_{\mathbb{R}^2} \mathbb{1}_A(u+s) p_\mu(u) p_\nu(s)\,ds\,du = \int_{\mathbb{R}^2} \mathbb{1}_A(u+s)\,d\nu(s)\,d\mu(u) = \mu * \nu(A)$ □

Теорема 2.5. Если ξ и η независимые случайный величины, то $P_{\xi+\eta} = P_\xi * P_\eta$

Доказательство. Нужно взять какое-то борелевское множество и понять как устроено там распределение суммы.

Пусть $B = \{(x, y) : x + y \in A\}$

$P_{\xi+\eta}(A) = P(\xi + \eta \in A) = P((\xi, \eta) \in B) = P_{\xi,\eta}(B) = \int_{\mathbb{R}^2} \mathbb{1}_B(x, y) dP_\xi(x)\, dP_\eta(y) = \int_{\mathbb{R}^2} \mathbb{1}_A(x + y) dP_\xi(x)\, dP_\eta(y) = P_\xi * P_\eta(A)$ ☐

Пример. 1. **Свертка с дисректным распределением**

$\nu = \sum_{k=1}^{\infty} p_k \delta_{x_k}$. Тогда $\mu * \nu(A) = \int_{\mathbb{R}} \mu(A - x)\, d\nu(x) = \sum_{k=1}^{\infty} \mu(A - x_k) p_k$

2. $\xi_i \sim Poisson(\lambda_i)$. ξ_1 и ξ_2 независимы.

$P_{\xi_1+\xi_2}(\{n\}) = \sum_{k=0}^{+\infty} P_{\xi_1}(\{n - k\}) \cdot \frac{\lambda_2^k e^{-\lambda_2}}{k!} = \sum_{k=0}^{n} \frac{\lambda_1^{n-k} e^{-\lambda_1}}{(n-k)!} \cdot \frac{\lambda_2^k e^{-\lambda_2}}{k!} = e^{-\lambda_1} e^{-\lambda_2} \sum_{k=0}^{n} \frac{\lambda_1^{n-k} \lambda_2^k}{k!(n-k)!} = \frac{(\lambda_1+\lambda_2)^n e^{-\lambda_1-\lambda_2}}{n!}$

$\xi_1 + \xi_2 \sim Poisson(\lambda_1 + \lambda_2)$

2.4. Математическое ожидание и дисперсия

***Определение* 2.17.** $\xi : \Omega \to \mathbb{R}$ - случайная величина ($\xi \geq 0$, либо суммируемая функция). $\mathbb{E}\xi = \int_{\mathbb{R}} \xi(\omega)\, dP(\omega)$ - математическое ожидание (среднее значение случайной величины).

Свойства. 1. $a, b \in \mathbb{R} :\ \mathbb{E}(a\xi + b\eta) = a\mathbb{E}\xi + b\mathbb{E}\eta$

2. Если $\xi \geqslant 0$, с вероятностью 1, то $\mathbb{E}\xi \geqslant 0$ (по сути написано, что если функция почти везде неотрицательна, то интеграл неотрицателен).

3. Если $\xi \geqslant \eta$ с вероятностью 1, то $\mathbb{E}\xi \geqslant \mathbb{E}\eta$

4. $\mathbb{E}\xi = \int_{\mathbb{R}} x\, dP_\xi(x)$

5. Если $f : \mathbb{R}^n \to \mathbb{R}$ - измерима относительно борелевской $\sigma-$алгебры.

   Тогда $\mathbb{E}f(\xi_1, \xi_2 \ldots \xi_n) = \int_{\mathbb{R}^n} f(x_1, \ldots, x_n) dP_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n)$

   Доказательство: $f = \mathbb{1}_A$. Тогда $\mathbb{E}\mathbb{1}_A(\xi_1, \ldots \xi_n) = \int_\Omega \mathbb{1}_A(\xi_1(w), \ldots, \xi_n(w)) dP(\omega) = P(\omega \in \Omega : \bar{\xi} \in A) = P_{\bar{\xi}}(A) = \int_{\mathbb{R}^n} \mathbb{1}_A(x_1, \ldots, x_n) dP_{\bar{\xi}}(x_1, \ldots, x_n)$.

   Тогда по линейности верно для простых.

   Теперь берём f_j неотрицательный простые, такие, что возрастают и $\to f$. И предельный переход по теореме Леви.

6. Если ξ_1 и ξ_2 независимы, то $\mathbb{E}(\xi \cdot \eta) = \mathbb{E}\xi \cdot \mathbb{E}\eta$

   Доказательство: $\mathbb{E}(\xi\eta) = \int_{\mathbb{R}^2} xy dP_{\xi,\eta}(x, y) =$

   $$\underbrace{=}_{\text{независимость сл. вел.}} \int_{\mathbb{R}} \int_{\mathbb{R}} xy dP_\xi(x)\, dP_\eta(y) = \int_{\mathbb{R}} y \int_{\mathbb{R}} x dP_\xi(x)\, dP_\eta(y) = \mathbb{E}\xi \cdot \mathbb{E}\eta$$

7. Если $\xi \geqslant 0$, то $\mathbb{E}\xi = \int_0^{+\infty} P(\xi \geqslant t)\, dt$ - из теории меры.

8. Если $p, q > 1$ и $\frac{1}{p} + \frac{1}{q} = 1$, то $\mathbb{E}|\xi\eta| \leqslant (\mathbb{E}|\xi|^p)^{\frac{1}{p}} (\mathbb{E}|\eta|^q)^{\frac{1}{q}}$ - неравенство Гёльдера

9. Неравенство Ляпунова

   $0 < r < s$, тогда $(\mathbb{E}|\xi|^r)^{\frac{1}{r}} \leqslant (\mathbb{E}|\xi|^s)^{\frac{1}{s}}$.

   Доказательство: $p = \frac{s}{r} > 1,\ \frac{1}{q} = 1 - \frac{1}{p} = \frac{s-r}{s} < 1$.

Тогда запишем Гельдера для ξ и $\eta = 1$:

$\mathbb{E}|\xi|^r|1| \leq (\mathbb{E}(|\xi|^r)^p)^{\frac{1}{p}} \cdot (\mathbb{E}1^q)^{\frac{1}{q}} = (\mathbb{E}|\xi|^s)^{\frac{r}{s}}$.

`Замечание.` $\mathbb{E}(\xi\eta) = \mathbb{E}\xi \cdot \mathbb{E}\eta$ без независимости неверно. Пример.

Теорема 2.6. Неравенство Маркова

Если $\xi \geqslant 0, p, t > 0$, то $P(\xi \geqslant t) \leqslant \frac{\mathbb{E}\xi^p}{t^p}$.

Доказательство. Неравенство Чебышёва из теории меры. □

***Определение* 2.18.** 1. Моменты случайной величины. $\mathbb{E}(\xi^k)$ - k-ый момент.

2. Центральный момент. $\mathbb{E}(\xi - \mathbb{E}\xi)^k$ - k-ый центральный момент.

3. Абсолютный момент. $\mathbb{E}|\xi|^k$ - k-ый абсолютный момент.

***Определение* 2.19.** Медиана случайной величины. m - медиана ξ, если $P(\xi \geqslant m) \geqslant \frac{1}{2}$ и $P(\xi \leqslant m) \geqslant \frac{1}{2}$.

`Замечание.` Медиана не единственна.

Возьмём кубик. $\xi = 1, 2, \ldots, 6$ с вероятностью $\frac{1}{6}$. Тогда любое число $m \in [3, 4]$ подходит.

Чаще всего всё равно берут середину, чтобы была единственность.

Пример. Есть организация из 1000 человек. 1 начальник и 999 подчиненных.

Зарплата начальника 1.000.000$, а подчинённых 1000$.

$\mathbb{E} = \frac{999}{1000} \cdot 1000 + \frac{1}{1000} \cdot 1000000 = 1999$

$m = 1000$ - медиана лучше характеризует ситуацию в этом случае.

***Определение* 2.20.** Дисперсия. $\mathbb{D}\xi = \mathbb{E}(\xi - \mathbb{E}\xi)^2$ - второй центральный момент.

Обозначение в англоязычной литературе: $Var\xi$

Свойства. 1. $\mathbb{D}\xi = \mathbb{E}\xi^2 - (\mathbb{E}\xi)^2$

Доказательство: Пусть $a = \mathbb{E}\xi$.

Тогда $\mathbb{D}\xi = \mathbb{E}(\xi - a)^2 = \mathbb{E}\xi^2 - 2a\mathbb{E}\xi + a^2$

2. $\mathbb{D}\xi \geqslant 0$ и если $\mathbb{D}\xi = 0$, то $P(\xi = c) = 1$

   Доказательство: Если $\mathbb{D}\xi = 0$, то $\int_\Omega (\xi - a)^2\, dP = 0$, значит $(\xi - a)^2 = 0$ почти везде.

3. $\mathbb{D}(\xi + a) = \mathbb{D}\xi$

   Доказательство: $\mathbb{E}(\xi + a) = \mathbb{E}\xi + a$. А тогда $(\xi + a) - \mathbb{E}(\xi + a) = \xi - \mathbb{E}\xi$

4. $\mathbb{D}(c\xi) = c^2\mathbb{D}\xi$

   Доказательство: $\mathbb{D}(c\xi) = \mathbb{E}(c\xi)^2 - (\mathbb{E}(c\xi))^2$

5. Если ξ и η независимы, то $\mathbb{D}(\xi + \eta) = \mathbb{D}\xi + \mathbb{D}\eta$

   Доказательство: $\mathbb{D}(\xi + \eta) = \mathbb{E}(\xi + \eta)^2 - (\mathbb{E}(\xi + \eta))^2 = \mathbb{E}\xi^2 + 2\mathbb{E}(\xi\eta) + \mathbb{E}\eta^2 - (\mathbb{E}\xi)^2 - 2\mathbb{E}\xi\mathbb{E}\eta - (\mathbb{E}\eta)^2 = \mathbb{D}\xi + \mathbb{D}\eta$

6. Аналогично предыдущему, но для n случайных величин.

   Доказательство: индукция

7. $\mathbb{E}|\xi - \mathbb{E}\xi| \leqslant \sqrt{\mathbb{D}\xi}$

   Доказательство: $\mathbb{E}|\xi - \mathbb{E}\xi| \leqslant (\mathbb{E}|\xi - \mathbb{E}\xi|^2)^{\frac{1}{2}} = \sqrt{\mathbb{D}\xi}$ - написали Ляпунова.

8. **Неравенство Чебышёва**

   $P(|\xi - \mathbb{E}\xi| \geqslant t) \leqslant \frac{\mathbb{D}\xi}{t^2}$, где $t > 0$

   Доказательство: $P(|\xi - \mathbb{E}\xi| \geqslant t) \leqslant \frac{\mathbb{E}|\xi - \mathbb{E}\xi|^2}{t^2} = \frac{\mathbb{D}\xi}{t^2}$ - неравенство Маркова для $p = 2$.

***Определение* 2.21.** Стандартное отклонение $\sigma = \sqrt{\mathbb{D}\xi}$

Пример. 1. $\xi \sim U[0,1]$.

   Тогда $\mathbb{E}\xi = \int_0^1 x\,dx = \frac{x^2}{2}\Big|_0^1 = \frac{1}{2}$.

   $\mathbb{E}\xi^2 = \int_0^1 x^2\,dx = \frac{x^3}{3}\Big|_0^1 = \frac{1}{3}$. А тогда $\mathbb{D}\xi = \mathbb{E}\xi^2 - (\mathbb{E}\xi)^2 = \frac{1}{12}$

2. $\xi \sim U[a,b]$.

   Если $\eta \sim U[0,1]$ и $\xi = (b-a)\eta + a \sim U[a,b]$. Тогда $\mathbb{E}\xi = \mathbb{E}((b-a)\eta + a) = \frac{a+b}{2}$

   $\mathbb{D}((b-a)\eta + a) = \mathbb{D}((b-a)\eta) = (b-a)^2\mathbb{D}\eta = \frac{(b-a)^2}{12}$

3. $\xi \sim \mathcal{N}(0,1)$

   $\mathbb{E}\xi = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} xe^{\frac{-x^2}{2}}\,dx = 0$, так как функция нечётная.

   Значит $\mathbb{D}\xi = \mathbb{E}\xi^2 = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} x^2 e^{-\frac{x^2}{2}}\,dx = -\frac{e^{\frac{-x^2}{2}}x}{\sqrt{2\pi}}\Big|_{-\infty}^{+\infty} + \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-\frac{x^2}{2}}\,dx = 1$

4. $\xi \sim \mathcal{N}(a,\sigma^2)$

   Если $\eta \sim \mathcal{N}(0,1)$, то $\xi = \sigma\eta + a \sim \mathcal{N}(a,\sigma^2)$.

   $\mathbb{E}\xi = \mathbb{E}(\sigma\eta + a) = \sigma\mathbb{E}\eta + a = a$

   $\mathbb{D}\xi = \mathbb{D}(\sigma\eta + a) = \sigma^2\mathbb{D}\eta = \sigma^2$

***Определение* 2.22.** Пусть $\mathbb{E}\xi^2 < +\infty$ и $\mathbb{E}\eta^2 < +\infty$.

Ковариация $cov(\xi,\eta) = \mathbb{E}((\xi - \mathbb{E}\xi)(\eta - \mathbb{E}\eta))$

Свойства. 1. $cov(\xi,\xi) = \mathbb{D}\xi$

2. $cov(\xi,\eta) = cov(\eta,\xi)$

3. $cov(c\xi,\eta) = c \cdot cov(\xi,\eta)$

4. $cov(\xi_1 + \xi_2,\eta) = cov(\xi_1,\eta) + cov(\xi_2,\eta)$

5. $cov(\xi,\eta) = \mathbb{E}(\xi\eta) - \mathbb{E}\xi\mathbb{E}\eta$

   Доказательство: $\mathbb{E}\xi = a, \mathbb{E}\eta = b$

   $cov(\xi,\eta) = \mathbb{E}((\xi - a)(\eta - b)) = \mathbb{E}(\xi\eta) - a\mathbb{E}\eta - b\mathbb{E}\xi + ab$

6. Если ξ и η независимы, то $cov(\xi,\eta) = 0$

7. $\mathbb{D}(\xi + \eta) = \mathbb{D}\xi + \mathbb{D}\eta + 2cov(\xi,\eta)$

8. $\mathbb{D}(\xi_1 + \xi_2 + \ldots + \xi_n) = \mathbb{D}\xi_1 + \mathbb{D}\xi_2 + \ldots + \mathbb{E}\xi_n + 2\sum_{i<j} cov(\xi_i,\xi_j)$.

Пример. $P(\text{успех}) = p$. Делаем n подбрасываний. $\eta =$ количество переходов от орла к решке.

Пусть $\xi_i = 1$, если на i позиции орёл, на $i+1$ позиции решка, иначе $\xi_i = 0$.

$\eta = \xi_1 + \ldots + \xi_{n-1}$. Тогда $\mathbb{E}\eta = \sum_{i=1}^{n-1} \mathbb{E}\xi_i = (n-1)pq$.

$\mathbb{D}\eta = \sum_{i=1}^{n-1} \mathbb{D}\xi_i + 2\sum_{i<j} cov(\xi_i, \xi_j)$.

Если $i + 1 < j$, то ξ_i и ξ_j независимы, поэтому в сумме почти везде нули.

Значит $\mathbb{D}\eta = \sum_{i=1}^{n-1} \mathbb{D}\xi_i + 2\sum_{i=1}^{n-1} cov(\xi_i, \xi_{i+1})$.

$\mathbb{D}\xi_i = \mathbb{E}\xi_i^2 - (\mathbb{E}\xi_i)^2 = pq - p^2q^2$.

$cov(\xi, \xi_{i+1}) = \mathbb{E}(\xi_i\xi_{i+1}) - \mathbb{E}\xi_i\mathbb{E}\xi_{i+1} = -p^2q^2$

`Замечание.` 1. $\{\xi : \mathbb{E}\xi^2 < +\infty\}$

$\langle \xi, \eta \rangle = \mathbb{E}(\xi\eta)$ - скалярное произведение.

$\mathbb{E}\xi$ - ортогональная проекция на константы.

2. $\langle \xi, \eta \rangle = cov(\xi, \eta)$ - тоже скалярное произведение.

Норма - это стандартное отклонение.

Теорема 2.7. Выбор двудольного подграфа

Есть граф G с n вершинами и m рёбрами. Хотим стереть некоторое количество рёбер(как можно меньше) так, чтобы остался двудольный подграф.

Тогда G содержит двудольный подграф с $\geqslant \frac{m}{2}$ рёбрами.

Доказательство. A - те вершины, на которых выпал орёл, B - на которых выпала решка.

Будем интересоваться матожидание количества рёбер в такой ситуации.

$$\xi_{xy} = \begin{cases} 1, & \text{если x, y из разных долей} \\ 0, & \text{иначе} \end{cases}$$

$\mathbb{E}\xi = \sum_{xy \in E} \mathbb{E}\xi_{xy} = \frac{m}{2}$, а значит есть реализация с $\frac{m}{2}$. □

***Определение* 2.23.** Коэффициент корреляции. $\rho(\xi, \eta) = \frac{cov(\xi,\eta)}{\sqrt{\mathbb{D}\xi}\sqrt{\mathbb{D}\eta}} \in [-1, 1]$

***Определение* 2.24.** Если $cov(\xi, \eta) = 0$, то это некоррелирующие случайные величины.