Содержание

Лекция 1

В теории вероятности обычно изучают случайные события
Обычно наука занимается закономерностями, но так как в случайных экспериментах нет закономерностей, теория вероятности занимается поисков закономерности в сериях случайных экспериментах
Итак, в XVI веке начали с экспериментов бросков монеты:

число бросков	число гербов	частота
4040	2048	0.5069
12000	6019	0.5016
24000	12012	0.5005

Как можно видеть, частота стремится к 0.5 - появляется статистическая закономерность

Статистическое определение вероятности

Пусть проводится n реальных экспериментов, при которых событие A появилось n_A раз
Отношение $\dfrac{n_A}{n}$ называется частотой события A
Эксперименты показывают, что при увеличении числа n частота стабилизируется у некоторого числа, при котором мы понимаем статистическую вероятность: $P(A) \approx \dfrac{n_A}{n}$ при $n \to \infty$

Пространство элементарных исходов. Случайные события

Def. Пространством элементарных исходов Ω называется множество, содержащее все возможные исходы экспериментов, из которых при испытании происходит ровно один. Элементы этого множества называются элементарными исходами и обозначаются ω

Def. Случайными событиями называется подмножество $A \subset \Omega$. События A наступают, если произошел один из элементарных исходов из множества A

Ex. 1. Бросок монеты: $\Omega = \{\text{Г}, \text{Р}\}$, $A = \{\text{Г}\}$ - выпал герб

Ex. 2. Игральная кость: $\Omega = \{1, 2, 3, 4, 5, 6\}$, $A = \{\text{выпало четное число}\} = \{2, 4, 6\}$

Ex. 3. Монета бросается дважды.
а) Учитываем порядок: $\Omega = \{\text{ГГ}, \text{РР}, \text{РГ}, \text{ГР}\}$
а) Не учитываем порядок: $\Omega = \{\text{ГГ}, \text{РР}, \text{ГР}\}$

Ex. 4. Кубик дважды: $\Omega = \{\langle i, j\rangle \mid 1 \le i, j \le 6\}$
$A = \{\text{разность} \vdots 3\} = \{\langle 1, 4\rangle; \langle 4, 1\rangle; \langle 2, 5\rangle; \langle 5, 2\rangle; \dots\}$

Ex. 5. Монета бросается до первого герба: $\Omega = \{\text{Г}, \text{РГ}, \text{РРГ}, \dots\}$ - счетно-бесконечное множество

Ex. 6. Монета бросается на плоскость: $\Omega = \{\langle x, y\rangle \mid x, y \in \mathbb{R}, \langle x, y\rangle$ - центр монеты$\}$ - несчетное число исходов

Операции над событиями

Ω - достоверные события (наступают всегда)
$\varnothing$ - невозможное события (никогда не наступает, так как не содержит ни одного элем. исхода)
Введем операции:

Def. 1. Суммой $A + B$ называется событие, состоящее в том, что произошло события A или событие B (хотя бы одно из них)

Def. 2. Произведением $A \cdot B$ называется событие, состоящее в том, что произошло событие A и событие B (оба из них)
Nota. $A_1 + A_2 + \dots + A_n + \dots$ - произошло хотя бы одно из этих событий
$A_1 \cdot A_2 \cdot \dots \cdot A_n \cdot \dots$ - произошли все эти события

Def. 3. Противоположным A событием называется событие $\overline{A}$, состоящее в том, что событие A не произошло
Nota. $\overline{\overline{A}} = A$

Def. 4. Дополнение (разность) $A \setminus B$ называется событие $A \cdot \overline{B}$

Def. 5. События A и B называются несовместными, если их произведение - пустое множество (не могут произойти одновременно при одном эксперименте)

Def. 6. События A влекут события B, если $A \subset B$ (если наступает A, то наступит B)

Вероятность

Мы хотим присвоить какую-то числовую характеристику к каждому событию, отражающее его частоту наступления: $0 \le P(A) \le 1$ - вероятность наступления события A

Классическое определение вероятности

Пусть пространство случайных событий Ω содержит конечное число равновозможных исходов, тогда применимо классическое определение вероятности

Def. $\boxed{P(A) = \dfrac{|A|}{|\Omega|} = \dfrac{m}{n}}$, где n - число всех возможных исходов, m - число благоприятных исходов

В частности, если $\Omega = n$ и A_i - элем. исх., то $P(A_i) = \dfrac{1}{n}$

Свойства:

1) $0 \le P(A) \le 1$

2) $P(\Omega) = 1 \quad (m = n)$

3) $P(\varnothing) = 0 \quad (m = 0)$

4) Если события A и B несовместны, то $P(A + B) = P(A) + P(B)$

Геометрическое определение вероятности (граф де Бюффон)

Пусть $\Omega \subset \mathbb{R}^n$ - замкнутая ограниченная область

$\mu(\Omega)$ - мера Ω в $\mathbb{R}^n$ (например, длина отрезка, площадь области на плоскости, объем тела в пространстве)

В эту область наугад бросаем точку. «Наугад» означает, что вероятность попадания в A зависит только от меры A и не зависит от ее расположения

В этом случае применимо геометрическое определение вероятности

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

Ex. 1. Монета диаметром в 6 см бросается на пол, вымощенной квадратной плиткой со стороной 20 см, какова вероятность, что монета окажется целиком внутри одной плитки

$\mu(\Omega) = 20^2 = 400$

$\mu(A) = (20 - 3 - 3)^2 = 196$

$P(A) = \dfrac{\mu(A)}{\mu(\Omega)} = \dfrac{196}{400} = 0.49$

Ex. 2. Задача Бюффона об игле: пусть пол вымощен ламинатом, $2l$ - ширина доски, на пол бросается игла длины, равной ширине доски, найти вероятность того, что игла пересечет стык доски

Определим положение иглы координатами центра и углом, между иглой и стыком доски, причем можно считать, что эти величины независимы

$\exists x \in [0; l]$ - расстояние от центра до ближайшего края,

$\varphi \in [0; \pi]$ - угол

$\Omega = [0; l] \times [0; \pi]$

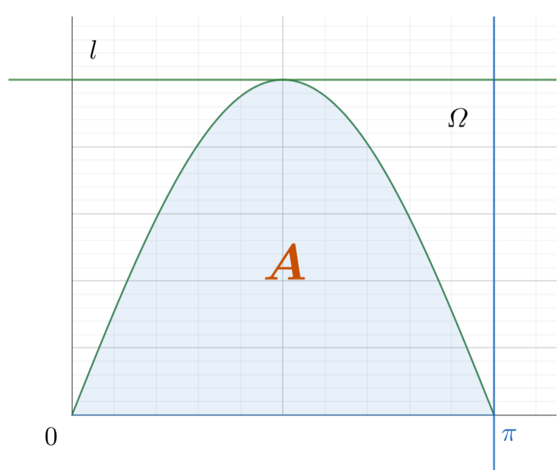
Событие A (пересечет стык) наступает, если $x \le l \sin\varphi$

$P(A) = \dfrac{S(A)}{S(\Omega)}$

$S(\Omega) = \pi l$

$S(A) = \displaystyle\int_0^\pi l \sin\varphi d\varphi = -l\cos\varphi\Big|_0^\pi = -l(-1-1) = 2l$

$P(A) = \dfrac{2l}{\pi l} = \dfrac{2}{\pi}$



Лекция 2

Построение модели случайных явлений

1. Задаем пространство элементарных исходов Ω
2. **Def.** Система $\mathcal{F}$ подмножеств Ω называется σ-алгеброй событий, если:

   1) $\Omega \in \mathcal{F}$;

   2) $A \in \mathcal{F} \Longrightarrow \overline{A} \in \mathcal{F}$;

   3) $A_1, A_2, \dots, A_n, \dots \in \mathcal{F} \Longrightarrow \bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{F}$

   Свойства:

   (a) $\varnothing \in \mathcal{F}$, так как $\Omega \in \mathcal{F} \Longrightarrow \overline{\Omega} = \varnothing \in \mathcal{F}$

   (b) $A_1, A_2, \dots \in \mathcal{F} \Longrightarrow \bigcap\limits_{i=1}^{\infty} A_i \in \mathcal{F}$

   $\square \quad A_1, A_2, \dots \in \mathcal{F} \Longrightarrow \overline{A}_1, \overline{A}_2, \dots \in \mathcal{F} \Longrightarrow \bigcup\limits_{i=1}^{\infty} \overline{A}_i \in \mathcal{F} \Longrightarrow \overline{\bigcup\limits_{i=1}^{\infty} \overline{A}_i} = \bigcap\limits_{i=1}^{\infty} A_i \in \mathcal{F} \quad \square$

   (c) $A, B \in \mathcal{F} \Longrightarrow A \setminus B \in \mathcal{F}$

   $\square \quad A, B \in \mathcal{F} \Longrightarrow A, \overline{B} \in \mathcal{F} \Longrightarrow A \setminus B = A \cdot \overline{B} \in \mathcal{F} \quad \square$

   Ex. 1. $\mathcal{F} = \{\varnothing, \Omega\}$

   Ex. 2. $\mathcal{F} = \{\varnothing, \Omega, A, \overline{A}\}$

   Ex. 3. **Def.** Борелевская σ-алгебра $\mathcal{B}(\mathbb{R})$ - минимальная σ-алгебра, содержащая все возможные интервалы на прямой

3. **Def.** $\sqsupset$ Ω - пространство элементарных исходов, $\mathcal{F}$ - его σ-алгебра событий. *Вероятностью* на $(\Omega, \mathcal{F})$ называется функция $P : \mathcal{F} \to \mathbb{R}$ со свойствами:

   (a) $P(A) \geq 0 \quad \forall A \in \mathcal{F}$ (неотрицательность)

   (b) Если $A_1, A_2, \dots, A_n, \dots \in \mathcal{F}$ - несовместное, то $P(\sum\limits_{i=1}^{\infty} A_i) = \sum\limits_{i=1}^{\infty} P(A_i)$ (свойство счетной аддитивности)

   (c) $P(\Omega) = 1$ (условие нормированности)

Def. Из этого тройка $(\Omega, \mathcal{F}, P)$ называется вероятностным пространством

Свойства вероятности

1. Так как $\varnothing$ и Ω - несовместные, то $1 = P(\Omega) = P(\Omega + \varnothing) = 1 + P(\varnothing) \Longrightarrow P(\varnothing) = 0$

2. Формула обратной вероятности: $P(A) = 1 - P(\overline{A})$

> $\square$ A и $\overline{A}$ - несовместные и $A + \overline{A} = \Omega \Longrightarrow P(A + \overline{A}) = P(\Omega) = 1$ $\square$

3. $P(A) = 1 - P(\overline{A}) \leq 1$

Аксиома непрерывности

> **Th.** Пусть имеется убывающая цепочка событий $A_1 \supset A_2 \supset A_3 \supset \dots \supset A_n \supset \dots$ и $\bigcap\limits_{i=1}^{\infty} A_n = \varnothing$
> Тогда $P(A_n) \underset{n\to\infty}{\longrightarrow} 0$
> При непрерывном изменении области $A \subset \Omega \subset \mathbb{R}^n$ соответствующая вероятность $P(A)$ также должна изменятся непрерывно

Аксиома непрерывности следует из аксиомы счетной аддитивности

> $\square$
> Ясно, что $A_n = \sum\limits_{i=n}^{\infty} A_i \overline{A}_{i+1} + \prod\limits_{i=n}^{\infty} A_i$
> $\prod\limits_{i=n}^{\infty} A_i = A_n \cdot \prod\limits_{i=n+1}^{\infty} A_i = \prod\limits_{i=1}^{n} \cdot \prod\limits_{i=n+1}^{\infty} A_i = \prod\limits_{i=1}^{\infty} = \varnothing \Longrightarrow A_n = \sum\limits_{i=n}^{\infty} A_n \overline{A_{n+1}}$ и так как эти события несовместны, то по свойству счетной аддитивности $P(A_n) = \sum\limits_{i=n}^{\infty} P(A_i \overline{A_{i+1}})$ - это остаток (хвост) сходящегося ряда
> $P(A_1) = \sum\limits_{i=1}^{\infty} P(A_i \overline{A_{i+1}}) = \sum\limits_{i=1}^{n-1} P(A_i \overline{A_{i+1}}) + P(A_n)$ и $P(A_n) \underset{n\to\infty}{\longrightarrow} 0$ по необходимому признаку сходимости
> $\square$

Nota. Аксиому счетной аддитивности можно вывести из конечной аддитивности и аксиомы счетной непрерывности

Свойства операций сложения и умножения

1. Свойство дистрибутивности: $A \cdot (B + C) = AB + AC$
2. Формула сложения: если A и B несовместны, то $P(A + B) = P(A) + P(B)$
3. Формула сложения вероятностей: $P(A + B) = P(A) + P(B) - P(AB)$

□

$A+B=A\overline{B}+AB+\overline{A}B$ - несовместные события $\Longrightarrow P(A+B)=P(A\overline{B})+P(AB)+P(\overline{A}B)=(P(A\overline{B})+P(AB))+(P(AB)+P(\overline{A}B))-P(AB)=P(A)+P(B)-P(AB)$

□

Ex. Из колоды в 36 карт достали одну карту. Какова вероятность того, что будет дама или пика

Пусть Д - дама, П - пика, $P(\text{Д}+\text{П})=P(\text{Д})+P(\text{П})-P(\text{ДП})=\frac{4}{36}+\frac{9}{36}-\frac{1}{36}=\frac{1}{3}$

Формула сложения при $N=3$: $P(A_1+A_2+A_3)=P(A_1)+P(A_2)+P(A_3)-P(A_2A_3)-P(A_1A_3)-P(A_1A_2)+P(A_1A_2A_3)$

Общий случай: $P(A_1+A_2+\dots+A_n)=\sum\limits_{i=1}^{n}P(A_i)-\sum\limits_{i<j}P(A_iA_j)+\sum\limits_{i<j<k}P(A_iA_jA_k)+(-1)^{n-1}\cdot P(A_1A_2\dots A_n)$ - формула включения и исключения

Ex. n писем случайно раскладывается по n конвертам. Найти вероятность того, что хотя бы одно письмо окажется в своем конверте

$\sqsupset A_i$ - i-ое письмо в своем конверте

$P(A_i)=\frac{1}{n}; P(A_iA_j)=\frac{1}{A_n^2}; P(A_iA_jA_k)=\frac{1}{A_n^3}; P(A_1A_2\dots A_n)=\frac{1}{n!}$

Слагаемых вида A_i - n штук; A_iA_j - C_n^2; $A_iA_jA_k$ - C_n^3; $A_1A_2\dots A_n$ - 1 штука

$P(A)=P(A_1+A_2+\dots+A_n)=n\cdot\frac{1}{n}-C_n^2\frac{1}{A_n^2}+C_n^3\frac{1}{A_n^3}-\dots+(-1)^{n-1}\frac{1}{n!}=1-\frac{1}{2}+\frac{1}{3!}-\dots+(-1)^{n-1}\frac{1}{n!}$

Так как $e^{-1}=1-1+\frac{1}{2}-\frac{1}{3!}+\dots$, то при $n\to\infty \quad P(A)\underset{n\to\infty}{\to}1-e^{-1}\approx 0.63$

Независимые события

Под независимыми событиями логично подразумевать события, не связанные причинно-следственной связью (то есть когда факт наступления одного не влияет на оценку вероятности другого)

$\sqsupset |\Omega|=n; |A|=m_1; |B|=m_2$

Проведем пару независимых испытаний. Тогда получаем пространство элементарных исходов $\Omega\times\Omega$ и $|\Omega\times\Omega|=n^2$

По основному принципу комбинаторики $|A\cdot B|=m_1\cdot m_2$

$P(AB)=\frac{|A\cdot B|}{|\Omega\times\Omega|}=\frac{m_1m_2}{n^2}=P(A)\cdot P(B)$

Def. События A и B называются независимыми, если $P(A\cdot B)=P(A)\cdot P(B)$

Lab. $\sqsupset P(A),P(B)\neq 0$, доказать, что если A и B несовместны, то они зависимы

Свойство: Если A и B независимы, то независимы $\overline{A}$ и $\overline{B}$, A и $\overline{B}$, $\overline{A}$ и B

Доказательство: $A = A \cdot (B+\overline{B}) = AB + A\overline{B}$ - несовместные события $\Longrightarrow P(A) = P(AB) + P(A\overline{B}) \Longrightarrow$ $P(A\overline{B}) = P(A) - P(AB) = P(A) - P(A) \cdot P(B) = P(A)(1-P(B)) = P(A)P(\overline{B}) \Longrightarrow$ независимы

Def. События $A_1, A_2, \dots A_n$ - независимы в совокупности, если для любого набора $i_1, i_2, \dots, i_k$ $(2 \leq k \leq n)$ $P(A_{i_1} \cdot A_{i_2} \cdot \dots \cdot A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \dots \cdot P(A_{i_k})$

Nota. Из независимости в совокупности при $k = 2$ получаем попарную независимость. Обратное утверждение неверно

Ex. (С. Бернштейн)
Пусть имеется правильный тетраэдр, одна грань окрашена в красный, вторая в синий, третья в зеленый, а четвертая во все эти три цвета.
Подбросили тетраэдр, $\sqsupset A$ - грань, которая содержит красный цвет, B - синий, C - зеленый.
$P(A) = P(B) = P(C) = \frac{2}{4} = \frac{1}{2}$
Так как $P(AB) = P(AC) = P(BC) = \frac{1}{4}$
$P(AB) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(A)P(B)$ - попарная независимость
$P(ABC) = \frac{1}{4} \neq P(A)P(B)P(C)$ - но вот независимость в совокупности не соблюдается

Ex. (Шевалье де Мере, Паскаль, Ферма, $\approx$ 1650 г.)
Какова вероятность того, что при 4 бросании кости выпадет одна шестерка
A_1 - при первом броске шестерка, A_2 - при втором, A_3 - при третьем, A_4 - при четвертом
B - выпала хотя бы одна шестерка при 4 бросках
$B = A_1 + A_2 + A_3 + A_4$ - совместные события, но независимые
Найдем обратную вероятность: $\overline{B}$ - ни разу не выпала шестерка
$\overline{B} = \overline{A_1} \cdot \overline{A_2} \cdot \overline{A_3} \cdot \overline{A_4}$
$P(\overline{A_1}) = P(\overline{A_2}) = P(\overline{A_3}) = P(\overline{A_4}) = \frac{5}{6}$
$\overline{B} = P(\overline{A_1})P(\overline{A_2})P(\overline{A_3})P(\overline{A_4}) = \left(\frac{5}{6}\right)^4 \approx 0.482$
$P(B) = 1 - P(\overline{B}) \approx 0.52$

Лекция 3

Условная вероятность

Условная вероятность $P(A|B)$ (или $P_B(A)$) - вероятность события A, вычисленная в предположении, что событие B уже произошло

Ex. Бросается кость один раз, известно, что выпало больше 3 очков. Найти вероятность того, что выпало четное число очков

A - выпало четное число очков

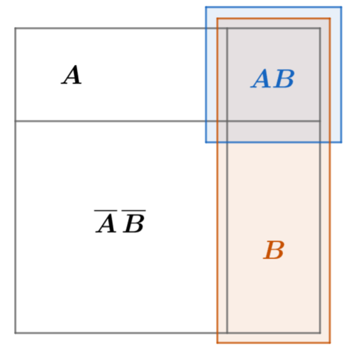
B - выпало больше трех очков

$\Omega = \{1,2,3,4,5,6\}; |\Omega| = 6; A = \{2,4,6\}; B = \{4,5,6\}$

$P(A|B) = \frac{2}{3} = \frac{\frac{2}{6}}{\frac{3}{6}} = \frac{P(AB)}{P(B)}$

Интерпретация с помощью геометрической вероятности:

$P(A|B) = \frac{S_{AB}}{S_B} = \frac{\frac{S_{AB}}{S_\Omega}}{\frac{S_B}{S_\Omega}}$



Def. Условной вероятностью события A при условии, что имело место событие B, называется величина $P(A|B) = \frac{P(AB)}{P(B)}$

Ex. Известно, что среди населения 1% воров. В комнате, где находилось 10 гостей, у хозяина пропал кошелек. Какова вероятность того, что произвольный гость является вором.

A - гость является вором $P(A) = 0.01$

B - пропал кошелек (хотя бы один вор среди гостей есть)

$P(A|B) = \frac{P(AB)}{P(B)} = \frac{P(AB)}{1 - P(\overline{B})} = \frac{P(A)}{1 - 0.99^{10}} = \frac{0.01}{1 - 0.99^{10}} = 0.105$

Формула умножения:

В качестве следствия условной вероятности получаем:

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

Общий случай:

$P(A_1A_2A_3 \dots A_n) = P(A_1)P(A_2|A_1)P(P_3|A_1A_2) \dots P(A_n|A_1A_2 \dots A_{n-1})$

□

База индукции $P(AB) = P(B)P(A|B)$

Шаг индукции: пусть верно при $n-1$:

$P(A_1A_2A_3 \dots A_{n-1}) = P(A_1)P(A_2|A_1)P(A_3|A_1A_2) \dots P(A_n|A_1A_2 \dots A_{n-2})$

$P(A_1A_2A_3 \dots A_n) = P(A_1A_2A_3 \dots A_{n-1}) \cdot P(A_n|A_1A_2 \dots A_{n-1}) =$

$P(A_1)P(A_2|A_1)P(A_3|A_1A_2) \dots P(A_n|A_1A_2 \dots A_{n-1})$

□

Ex. Студент выучил 1 билет из n, в группе n студентов. Каким по очереди ему нужно зайти, чтобы вероятность сдать экзамен была наибольшей

Пусть A_i - билет, вытянутый на i-ом шаге $(1 \le i \le n)$

A - студент сдал экзамен

$$P(A) = P(\overline{A_1} \cdot \overline{A_2} \cdot \dots \cdot \overline{A_{i-1}} \cdot A_i) = \frac{n-1}{n} \cdot \frac{n-2}{n-1} \cdot \dots \cdot \frac{n-(i-1)}{n-(i-2)} \cdot \frac{1}{n-(i-1)} = \frac{1}{n}$$

Полная группа событий

Def. События $H_1, H_2, \dots, H_n, \dots$ образуют полную группу событий, если они попарно несовместны и содержат все возможные элементарные исходы

$H_i \cap H_j = \varnothing \ \forall i, j \qquad \bigcup\limits_{i=1}^{\infty} H_i = \Omega$

Следствие: $\sum\limits_{i=1}^{\infty} P(H_i) = 1$

Th. Формула полной вероятности. $\sqsupset H_1, H_2, \dots, H_n, \dots$ - полная группа событий. Тогда $P(A) = \sum\limits_{i=1}^{\infty} P(H_i) P(A|H_i)$

□

$P(A) = P(\Omega A) = P((H_1 + H_2 + H_3 + \dots)A) = P(H_1 A + H_2 A + H_3 A + \dots) = [H_i \cdot A \cdot H_j \cdot A = \varnothing \cdot A] = P(H_1 A) + P(H_2 A) + \dots = P(H_1)P(A|H_1) + P(H_2)P(A|H_2) + \dots$

□

Th. Формула Байеса. $\sqsupset H_1, H_2, \dots, H_n$ - полная группа событий, и известно, что событие A уже произошло

Тогда $P(H_k|A) = \dfrac{P(H_k)P(A|H_k)}{\sum_{i=1}^{\infty} P(H_i)P(A|H_i)}$

□

$P(H_k|A) = \dfrac{P(H_k A)}{P(A)} = \dfrac{P(H_k)P(A|H_k)}{\sum_{i=1}^{\infty} P(H_i)P(A|H_i)}$

□

Ex. 1. В первой коробке 4 белых и 2 черных шара, во второй 1 белый и 2 черных. Из первой коробки во вторую переложили 2 шара, затем из второй коробки достали шар. Какова вероятность того, что он оказался белым

$\sqsupset H_1$ - переложили 2 белых H_2 - 2 черных

H_3 - разного цвета

A - из второй коробки достали белый шар

$P(H_1) = \dfrac{4}{6} \cdot \dfrac{3}{5} = \dfrac{6}{15}$

$P(H_2) = \dfrac{2}{6} \cdot \dfrac{1}{5} = \dfrac{1}{15}$

$P(H_3) = \frac{4}{6} \cdot \frac{2}{5} + \frac{2}{6} \cdot \frac{4}{5} = \frac{4}{15} + \frac{4}{15} = \frac{8}{15}$

$P(A) = P(H_1) \cdot P(A|H_1) + P(H_2) \cdot P(A|H_2) + P(H_3) \cdot P(A|H_3) = \frac{6}{15} \cdot \frac{3}{5} + \frac{1}{15} \cdot \frac{1}{5} + \frac{8}{15} \cdot \frac{2}{5} = \frac{18}{75} + \frac{1}{75} + \frac{16}{75} = \frac{35}{75} = \frac{7}{15}$

Ex. 2. Вероятность попадания первого стрелка в цель 0.9, а второго 0.3. Наугад вызванный стрелок попал в цель. Какова вероятность того, что это бы первый стрелок?

H_1 - вызван первый стрелок

H_2 - вызван второй стрелок

A - стрелок попал

$P(H_1) = P(H_2) = \frac{1}{2}$

$P(A|H_1) = 0.9 \qquad P(A|H_2) = 0.3$

$P(H_1|A) = \frac{P(H_1)P(A|H_1)}{P(H_1)P(A|H_1) + P(H_2)P(A|H_2)} = \frac{\frac{1}{2}0.9}{\frac{1}{2}0.9 + \frac{1}{2}0.3} = \frac{9}{9+3} = 0.75$

Ex. 3. По статистике раком болеет 1% населения. Тест дает правильный результат в 99% случаев. Тест оказался положительный. Найти вероятность того, что человек болен.

H_1 - человек болен

H_2 - человек здоров

A - анализ положительный

$P(H_1) = 0.01$

$P(H_2) = 0.99$

$P(A|H_1) = 0.99$

$P(A|H_2) = 0.01$

$P(H_1|A) = \frac{P(H_1)P(A|H_1)}{P(H_1)P(A|H_1) + P(H_2)P(A|H_2)} = \frac{0.01 \cdot 0.99}{0.01 \cdot 0.99 + 0.99 \cdot 0.01} = \frac{1}{2} = 0.5$

Допустим, что второй независимый с первым анализ также оказался положительным. Найти вероятность того, что человек болен.

$P(H_1) = 0.01 \qquad P(H_2) = 0.99$

$P(AA|H_1) = 0.99^2 \qquad P(AA|H_2) = 0.01^2$

$P(H_1|AA) = \frac{0.01 \cdot 0.99^2}{0.01 \cdot 0.99^2 + 0.99 \cdot 0.01^2} = \frac{0.99}{0.99 + 0.01} = 0.99$

Интуитивно вероятность $\frac{1}{2}$ может поддаваться непониманию, однако можно рассуждать так: пусть в городе живут 10000 человек, из них 100 болеют, а у 99 из них положительный анализ; у других 9900 положительный анализ всего лишь у 99, отсюда выходит $\frac{1}{2}$

Ex. 4. В телевизионной студии 3 двери 🚪🚪🚪, за одной из них приз 🚗. Игрок выбрал наугад одну из 3 дверей, после чего ведущий открывает одну из двух оставшихся дверей и показывает, что там приза нет 🛴. После чего предлагает игроку поменять свой выбор. Стоит ли игроку соглашаться?

H_1 - игрок угадал

H_2 - игрок не угадал

A - ведущий открыл дверь без приза

$P(H_1) = \frac{1}{3} \qquad P(H_2) = \frac{2}{3}$

$P(A|H_1) = 1 \qquad P(A|H_2) = \frac{1}{2}$

$P(H_1|A) = \frac{\frac{1}{3} \cdot 1}{\frac{1}{3} \cdot 1 + \frac{1}{3} \cdot \frac{1}{2}} = \frac{1}{2}$

Но это неправильно, так как действия ведущего неслучайны - он всегда откроет дверь без приза

В этом случае, если мы гипотетически выберем 300 дверей, в 100 случаях мы отгадаем, ведущий откроет любую дверь без приза; но в 200 случаях мы не отгадаем, ведущий откроет вторую дверь без приза, и в этом случае мы сможем поменяться на дверь с призом, отсюда шанс $\frac{2}{3}$, если мы поменяем свой выбор

Ex. 5. Вероятность того, что в семье с детьми ровно k детей, равна $\frac{1}{2^k}$, $k = 1, 2, \ldots$ Какова вероятность того, что в семье один мальчик, если известно, что нет девочки? Рождения мальчиков и девочек равновероятны.

H_i - в семье i детей $(1 \le i < \infty)$

$P(H_i) = \frac{1}{2^i}$

A - в семье нет девочки

$P(A|H_1) = \frac{1}{2}$

$P(A|H_2) = \frac{1}{4}$

$P(A|H_i) = \frac{1}{2^i}$

$P(H_1|A) = \frac{\frac{1}{2}\frac{1}{2}}{\sum_{i=1}^{\infty} \frac{1}{2^i} \cdot \frac{1}{2^i}} = \frac{\frac{1}{4}}{\frac{\frac{1}{4}}{1-\frac{1}{4}}} = \frac{3}{4} = 0.75$

Лекция 4

Серия испытаний Бернулли

Схемой Бернулли - называется серия одинаковых независимых экспериментов, каждый из которых имеет 2 исхода: произошло интересующее нас событие или нет

$p = p(A)$ - вероятность успеха при одном испытании

$q = 1 - p$ - вероятность неудачи

v_n - число успехов в серии из n испытаний

$p(v_n = k) = p_n(k)$

Из этого получаем формулу Бернулли:

Th. Вероятность того, что при n испытаниях произойдет ровно k успехов, равна $p_n(k) = C_n^k p^k q^{n-k}$

□

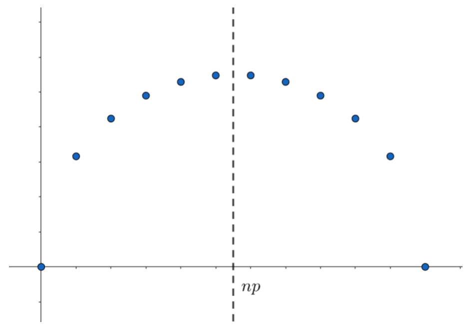
Рассмотрим один из элементарных исходов, благоприятных данному событию:
$A_n = \underbrace{\text{УУУ}\dots\text{У}}_{k}\underbrace{\text{Н}\dots\text{ННН}}_{n-k}$ - k успехов, $n-k$ неудачи
$p(\text{У}) = p, p(\text{Н}) = q$
Так как испытания независимы, то $p(A_n) = p^k q^{n-k}$
Остальные элементарные исходы имеют ту же вероятность, перебираем все расстановки исходов, получаем C_n^k, в итоге, получаем формулу Бернулли

□

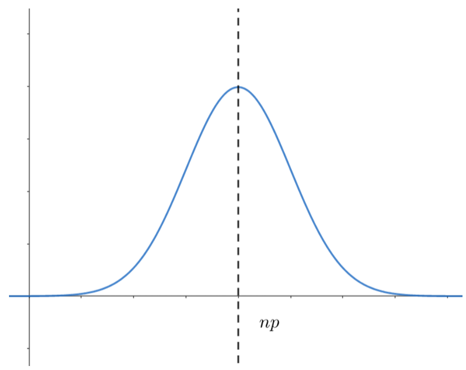
Ex. Вероятность попадания стрелка при одном выстреле - 0.8. Какова вероятность того, что из пяти выстрелов точными будут три
$n = 5 \quad p = 0.8 \quad q = 1 - p = 0.2 \quad k = 3$
$p_5(3) = C_5^3 p^3 q^2 = 0.2048$

Наиболее вероятное число успехов

Выясним, при каком значении k вероятность предшествующего числа успехов $k-1$ будет не более, чем вероятность k успехов
$p_n(k-1) \le p_n(k)$
$C_n^{k-1} p^{k-1} q^{n-k+1} \le C_n^k p^k q^{n-k}$
$\frac{n!}{(k-1)!(n-k+1)!} q \le \frac{n!}{(k)!(n-k)!} p$
$\frac{q}{(k-1)!(n-k+1)!} \le \frac{p}{(k)!(n-k)!}$
$\frac{q}{n-k+1} \le \frac{p}{k}$
$k(1-p) \le p(n-k+1)$
$k \le np + p$
Отсюда $np + p - 1 \le k \le np + p$
Рассмотрим 3 ситуации:
1) np - целое, тогда $np + p$ - нецелое, и $k = np$ - наиболее вероятное
2) $np + p$ - нецелое, тогда $k = \lfloor np + p \rfloor$
3) $np + p$ - целое, тогда $np + p - 1$ - целое, и 2 наиболее вероятных числа успеха
Геометрическая интерпретация:



При увеличении числа n точки превращаются в кривую Гаусса



При увеличении числа испытаний n формула Бернулли вырождается в следующие асимптотические формы (применяем, если требуется найти вероятность точного числа успеха)

1) Локальная формула Муавра-Лапласа

$p_n(k) \underset{n\to\infty}{\longrightarrow} \frac{1}{\sqrt{npq}}\varphi(x)$, где $\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ - функция Гаусса

$x = \frac{k - np}{\sqrt{npq}}$

Свойства $\varphi(x)$:

1. $\varphi(x) = \varphi(-x)$ - функция четная
2. при $x > 5 \quad \varphi(x) \approx 0$

2) Интегральная формула Муавра-Лапласа (если требуется найти вероятность того, что число успехов в данном диапазоне)

$p_n(k_1 \le k \le k_2) \underset{n\to\infty}{\longrightarrow} \Phi(x_2) - \Phi(x_1)$, где $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{z^2}{2}}dz$ - функция Лапласа

$x_1 = \frac{k_1 - np}{\sqrt{npq}}$ - отклонение от левой границы, $x_2 = \frac{k_2 - np}{\sqrt{npq}}$ - отклонение от правой

Свойства $\Phi(x)$

1. $\Phi(-x) = -\Phi(x)$ - функция нечетная
2. при $x > 5 \quad \Phi(x) \approx 0.5$

Nota. Эти формулы обычно можно применять при $n \ge 100$ и $0.1 \le p \le 0.9$

Nota. В некоторых источниках под функцией Лапласа подразумевают другую функцию: $F_0(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt$ - стандартное отклонение. Эта функция отличается от $F_0 = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{0} e^{-\frac{t^2}{2}}dt + \Phi(x) = \frac{1}{2} + \Phi(x)$

Так как $\int_{-\infty}^{\infty} e^{-x^2}dx = \sqrt{\pi}$ - интеграл Пуасона

Ex. Вероятность попадания стрелка в цель 0.8, стрелок сделал 400 выстрелов. Найти вероятность того, что:

а) произошло ровно 330 попаданий

б) произошло от 312 до 336 попаданий

а) $x = \frac{k-np}{\sqrt{npq}} = \frac{330 - 400\cdot 0.8}{\sqrt{400\cdot 0.8\cdot 0.2}} = \frac{330-320}{8} = 1.25$

$p_{400}(330) \approx \frac{1}{\sqrt{npq}}\varphi(1.25) = \frac{1}{8}\varphi(1.25) \approx \frac{1}{8}\cdot 0.1826 \approx 0.0228$

б) $x_1 = \frac{312-320}{8} = -1$, $x_2 = \frac{336-320}{8} = 2$

$p_{400}(312 \le k \le 336) \approx \Phi(2) - \Phi(-1) = \Phi(2) + \Phi(1) \approx 0.4772 + 0.3413 = 0.8185$

Статистическое понятие вероятности

Пусть проводим n реальных экспериментов, n_A - число появления события A, $\frac{n_A}{n}$ - относительная частота события A.

Эксперименты с монетой показали, что при больших n, $\frac{n_A}{n} \approx p(A)$ - явление стабилизации

Вероятность отклонения относительной частоты от вероятности события

n - число испытаний, $p = p(A)$, $\frac{n_A}{n}$ - экспериментальная частота

$p\left(|\frac{n_A}{n} - p| \le \varepsilon\right) = p\left(-\varepsilon \le \frac{n_A}{n} - p \le \varepsilon\right) = p(-n\varepsilon \le n_A - np \le n\varepsilon) = p(np - n\varepsilon \le n_A \le n\varepsilon + np) \underset{n\to\infty}{\longrightarrow}$ [по интегральной формуле Лапласа] $\underset{n\to\infty}{\longrightarrow} \Phi\left(\frac{n\varepsilon}{\sqrt{npq}}\right) - \Phi\left(-\frac{n\varepsilon}{\sqrt{npq}}\right)$

$= \Phi\left(\frac{\sqrt{n}\varepsilon}{\sqrt{pq}}\right) - \Phi\left(-\frac{\sqrt{n}\varepsilon}{\sqrt{pq}}\right)$

$= 2\Phi\left(\frac{\sqrt{n}\varepsilon}{\sqrt{pq}}\right)$

Итак, получили, что нужная нам вероятность $p\left(|\frac{n_A}{n} - p| \le \varepsilon\right) \approx 2\Phi\left(\frac{\sqrt{n}\varepsilon}{\sqrt{pq}}\right)$

Закон больших чисел Бернулли

Итак, $p\left(|\frac{n_A}{n} - p| \le \varepsilon\right) \underset{n\to\infty}{\longrightarrow} 2\Phi\left(\frac{\varepsilon}{\sqrt{pq}}\sqrt{n}\right)$

при $n \to \infty$, $\sqrt{n} \to \infty$, $\frac{\varepsilon}{\sqrt{pq}}\sqrt{n} \to \infty$, $\Phi\left(\frac{\varepsilon}{\sqrt{pq}}\sqrt{n}\right) \to 0.5$, $p\left(|\frac{n_A}{n} - p| \le \varepsilon\right) \to 2\cdot 0.5 = 1$ - закон больших чисел показывает, что вероятность попадания относительной частоты в ε-трубу приближается к 1

$\lim\limits_{n\to\infty} p\left(|\frac{n_A}{n} - p| \leq \varepsilon\right) = 1$ или $\frac{n_A}{n} \xrightarrow[n\to\infty]{} p$ - сходимость по вероятности

Ex. Для оценки доли p курящих людей берется выборка объема n, и делается оценка доли курящих людей по формуле $p^* = \frac{n_A}{n}$. Каким должен быть объем n, чтобы с вероятностью $\gamma = 0.95$ данная оценка отличалась от истинного значения не более, чем на $\varepsilon = 0.01$

По формуле вероятности отклонения частоты от вероятности $p(|p^* - p| \leq \varepsilon) = p\left(|\frac{n_A}{n} - p| \leq \varepsilon\right) \approx$
$2\Phi\left(\frac{\varepsilon}{\sqrt{pq}}\sqrt{n}\right) = 0.95$

$\Phi\left(\frac{\varepsilon}{\sqrt{pq}}\sqrt{n}\right) = 0.475$

$\frac{\varepsilon}{\sqrt{pq}}\sqrt{n} = 1.96$

$\frac{1}{\sqrt{pq}}\sqrt{n} = 196$

$\frac{n}{pq} = 38416$

$n \geq 38416pq$

В самое худшей ситуации $pq \leq 0.5^2 = \frac{1}{4}$

$n \geq \frac{38416}{4} = 9604$

Лекция 5

Схема испытаний и соответствующее распределение

Введем обозначения:

n - число испытаний

p - вероятность успеха при одном испытании

$q = 1 - p$ - вероятность неудачи

I. Схема Бернулли

$\exists \nu_n$ - число успехов в серии из n испытаний

$P_n(\nu_n = k) = C_n^k p^k q^{n-k}, \qquad k = 0, 1, \ldots, n$

Def. Соответствие $k \to C_n^k p^k q^{n-k}, \quad k = 0, \ldots, n$ называется биномиальным распределением (обозначается $B_{n,p}$ или $B(n, p)$)

II. Схема до первого успешного испытания

Пусть проводится бесконечная серия испытаний, которая заканчивается после первого успешного испытания под номером τ

Th. $P(\tau = k) = q^{k-1}p, \qquad k = 1, 2, \dots$

□

$$P(\tau = k) = P(\underbrace{\text{Н}\dots\text{Н}}_{k-1}\text{У}) = q^{k-1}p$$

□

Def. Соответствие $k \to q^{k-1}p, k \in \mathbb{N}$ называется геометрическим распределение вероятности (обозначается G_p или $G(p)$)

Nota. Геометрическое распределение обладает свойством нестарения или свойством отсутствия последействия

Th. $\exists P(\tau = k) = q^{k-1}p, k \in \mathbb{N}$. Тогда $\forall n, k \geq 0 \quad P(\tau > n + k \mid \tau > n) = P(\tau > k)$

□

Заметим, что $P(\tau > m) = q^m$, первые m - неудачи

$$P(\tau > n + k | \tau > n) = \frac{P(\tau > n + k, \tau > n)}{P(\tau > n)} = \frac{P(\tau > n + k)}{P(\tau > n)} = \frac{q^{n+k}}{q^n} = q^k$$

□

Nota. $P(\tau = n + k \mid \tau > n) = p(\tau = k)$ - Lab. доказать

III. Схема испытаний с несколькими исходами

Пусть при n независимых испытаний могут произойти m исходов (несовместных)
p_i - вероятность i-ого исхода при одном испытании

Th. Вероятность того, что при n испытаниях первый исход появится n_1 раз, второй - n_2 раз, m-ый - n_m $(\sum\limits_{i=1}^{m} n_i = n)$ равно

$$P_n(n_1, n_2, \dots, n_m) = \frac{n!}{n_1!n_2!\dots n_m!} p_1^{n_1} p_2^{n_2} \dots p_m^{n_m}$$

При $m = 2$ получаем формулу Бернулли

□

Рассмотрим следующий благоприятный исход, обозначим A_1

$A_1 = \underbrace{11\ldots1}_{n_1}\underbrace{22\ldots2}_{n_2}\ldots\underbrace{mm\ldots m}_{n_m}$

$p(A_1) = p_1^{n_1}p_2^{n_2}\ldots p_m^{n_m}$

Все остальные благоприятные исходы имеют ту же вероятность и отличаются лишь расположением i-ых исходов на n позициях, получаем мультиномиальную теорему: $\dfrac{n!}{n_1!n_2!\ldots n_m!}$

В итоге получаем требуемую формулу

□

Ex. Два одинаковых сильных шахматиста играют шесть партий

Вероятность ничьи в партии - 0.5. Какова вероятность того, что второй игрок выиграет две партии, а еще три сведет к ничьей

1-ый исход - выиграл 1 игрок

2-ой исход - выиграл 2 игрок

3-ий исход - ничья

$n = 6; \quad p_3 = 0.5; \quad p_1 = p_2 = \dfrac{1-p_3}{2} = 0.25$

$P_6(1;2;3) = \dfrac{6!}{1!2!3!}\left(\dfrac{1}{4}\right)^1\left(\dfrac{1}{4}\right)^2\left(\dfrac{1}{2}\right)^3 = \dfrac{4\cdot5\cdot6}{2}\dfrac{1}{2^9} \approx 0.12$

IV. Урновая схема

В урне N шаров, из которых K шаров белые, $N-K$ - черные

Из урны вынимаем (без учета порядка) n шаров. Найти вероятность, что из них k белых

а) Схема с возвратом (после каждого раза кладем шар обратно). В этом случае вероятность вынуть белый шар одинакова и равна $\dfrac{K}{N}$. Получаем схему Бернулли: $P_n(k) = C_n^k\left(\dfrac{K}{N}\right)^k\left(1-\dfrac{K}{N}\right)^{n-k}$

б) Схема без возврата - вынутый шар мы выбрасываем, тогда $P_{N,K}(n,k) = \dfrac{C_K^kC_{N-K}^{n-k}}{C_N^n}$

Def. Соответствие $k \to \dfrac{C_K^kC_{N-K}^{n-k}}{C_N^n}, k = 0,\ldots,n$ называется гипергеометрическим распределением

Nota. Если $K, N \to \infty$ так, что $\dfrac{K}{N} \approx p$ (не меняется), а n и k зафиксировать, то после выбора n шаров пропорции состава шаров не сильно изменятся, поэтому логично предположить, что гипергеометрическое распределение будет сходиться к биномиальному

Th. Если $K, N \to \infty$ таким образом, что $\frac{K}{N} \to p \in (0;1)$, а n и $0 \le k \le n$ фиксированы, то вероятность при гипергеометрическом распределении будет стремиться к биномиальному:

$P_{N,K}(n,k) = \frac{C_K^k C_{N-K}^{n-k}}{C_N^n} \to C_n^k \left(\frac{K}{N}\right)^k \left(1 - \frac{K}{N}\right)^{n-k}$

Воспользуемся леммой: $C_n^k \sim \frac{n^k}{k!}$ при $n \to \infty$ и фиксированном k

Доказательство леммы: $C_n^k = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\dots(n-k+1)}{n^k}\frac{n^k}{k!} = 1\left(1 - \frac{1}{n}\right)\dots\left(1 - \frac{k-1}{n}\right)\frac{n^k}{k!} \sim \frac{n^k}{k!}$

□

$P_{N,K}(n,k) = \frac{C_K^k C_{N-K}^{n-k}}{C_N^n} \sim \frac{K^k}{k!}\frac{(N-K)^{n-k}}{(n-k)!}\frac{n!}{N^n} = \frac{n!}{k!}\frac{(N-K)^{n-k}}{(n-k)!}\frac{K^k}{N^n} = C_n^k \left(\frac{K}{N}\right)^k \left(1 - \frac{K}{N}\right)^{n-k} \to C_n^k \left(\frac{K}{N}\right)^k \left(1 - \frac{K}{N}\right)^{n-k}$

□

V. Схема Пуассона. Теорема Пуассона для схемы Бернулли

Nota. Если вероятность успеха p в схеме Бернулли мала или близка к 1, то предельная формула Лапласа при недостаточно большом числе испытаний дает достаточно большую погрешность. В этой ситуации следует использовать формулу Пуасоона (формула редких событий)

Схема: вероятность числа успеха при одном испытании p_n зависит от числа испытаний n, причем таким образом, что $np_n \approx \lambda = const$

λ - интенсивность появления редких событий в единицу времени в потоке испытаний

Th. 1. (формула Пуассона) Пусть $n \to \infty, p_n \to 0$ таким образом, что $np_n \to \lambda = const > 0$

Тогда вероятность k успехов при n испытаниях: $P_n(k) = C_n^k p_n^k (1-p_n)^{n-k} \underset{n\to\infty}{\to} = \frac{\lambda^k}{k!}e^{-\lambda}$

□

Обозначим $\lambda_n = np_n$. Тогда $p_n = \frac{\lambda_n}{n}$ и

$P_n(k) = C_n^k \left(\frac{\lambda_n}{n}\right)^k \left(1 - \frac{\lambda_n}{n}\right)^{n-k} \underset{n\to\infty}{\to} \frac{n^k}{k!}\frac{\lambda_n^k}{n^k}\left(1 - \frac{\lambda_n}{n}\right)^n \cancelto{1}{\left(1 - \frac{\lambda_n}{n}\right)^{-k}} \underset{n\to\infty}{\to} \frac{\lambda_n^k}{k!}\left(1 - \frac{\lambda_n}{n}\right) =$

$$\frac{\lambda_n^k}{k!}\left(\left(1-\frac{\lambda_n}{n}\right)^{-\frac{n}{\lambda_n}}\right)^{-\frac{\lambda_n}{n}n} \underset{n\to\infty}{\longrightarrow} \frac{\lambda_n^k}{k!}e^{-\lambda_n} \underset{n\to\infty}{\longrightarrow} \frac{\lambda^k}{k!}e^{-\lambda}$$

□

Th. 2. (оценка погрешности в формуле Пуассона) Пусть v_n - число успехов при n испытаниях в схеме Бернулли

p - вероятность успеха при одном испытании, $\lambda = np$, $A \subset \{0, 1, \dots, n\}$ - произвольное подмножество чисел

Тогда $|P_n(v_n \in A) - \sum\limits_{k\in A}\frac{\lambda^k}{k!}e^{-\lambda}| \le \min(p, np^2) = \min(p, p\lambda)$

(без доказательства)

Def. Соответствие $k \longrightarrow \dfrac{\lambda^k}{k!}e^{-\lambda}, k = 0, 1, \dots$ называется распределением Пуассона с параметром $\lambda > 0$ (обозначается Π_λ)

Ex. Прибор состоит из 1000 элементов, вероятность отказа каждого элемента равна 0.001. Какова вероятность отказа больше двух элементов

$P_n(k) \approx \dfrac{\lambda^k}{k!}e^{-\lambda}$

$n = 1000, p = 0.001, \lambda = 1$

$P_n(k > 2) = 1 - P_n(k \le 2) = 1 - P(0) - P(1) - P(2) \approx 1 - \left(\dfrac{1^0}{0!}e^{-1} + \dfrac{1^1}{1!}e^{-1} + \dfrac{1^2}{2!}e^{-1}\right) = 1 - \left(1 + 1 + \dfrac{1}{2}\right)e^{-1} \approx 0.0803$

Лекция 6

Случайные величины

Примеры случайных величин:

Ex. 1. Бросаем кость, может выпасть 6 граней, здесь случайная величина ξ - число выпавших очков

Ex. 2. ξ - время работы микросхемы, в этом случае время может быть:

а) дискретным - $\xi \in \{0, 1, 2, \dots\}$

б) непрерывным - $\xi \in [0; \infty)$

Ex. 3. Температура за окном: $\xi \in (-50, +50)$

Def. На вероятностном пространстве $(\Omega, \mathcal{F}, p)$ функция $\xi : \Omega \to \mathbb{R}$ называется $\mathcal{F}$-измеримой, если $\forall x \in \mathbb{R}\ \{\omega \in \Omega \mid \xi(\omega) < x\} \in \mathcal{F}$ (то есть $\xi^{-1}(y) \in \mathcal{F}$, где $y \in (-\infty; x)$)

Def. Случайной величиной, заданной на вероятностном пространстве $(\Omega, \mathcal{F}, p)$, называется $\mathcal{F}$-измеримая функция $\xi\ : \Omega \to \mathbb{R}$, которая сопоставляет каждому элементарному исходу некоторое вещественное число

Nota. Не все функции являются $\mathcal{F}$-измеримыми
Ex. Кость: $\Omega = \{1, 2, 3, 4, 5, 6\}; \mathcal{F} = \{\varnothing, \Omega, \{2, 4, 6\}, \{1, 3, 5\}\}$
Пусть $\xi(\omega) = i$ - число выпавших очков. Тогда при $x = 4 : \{\omega \in \Omega \mid \xi(\omega) < 4\} = \{1, 2, 3\} \notin \mathcal{F} \Longrightarrow$ случайная величина не является $\mathcal{F}$-измеримой
В данном случае следует сделать ξ таким, что $\xi(2) = \xi(4) = \xi(6) = 1$, $\xi(1) = \xi(3) = \xi(5) = 0$

Nota. Смысл измеримости: если задана случайная величина ξ, то мы можем задать вероятность попадания случайной величины в интервал $(-\infty; x)$: $p(\xi \in (-\infty; x)) = p(\{\omega \in \Omega \mid \xi(\omega) < x\})$
А из интервалов $(-\infty; x)$ с помощью операций пересечения, объединения и дополнения можно получить все другие интервалы (включая точки) и также приписать им вероятности
Из матанализа известно, что мера из интервалов однозначно продолжается до меры на всей Борелевской σ-алгебры на $\mathbb{R}$ и, таким образом, с помощью случайной величины каждому Борелевскому множеству B также приписывается вероятность $p(\xi \in B)$
Итак, пусть ξ задана на вероятностном пространстве $(\Omega, \mathcal{F}, p)$, с помощью нее получаем новой вероятностное пространство $(\mathbb{R}, \mathcal{B}(\mathbb{R}), p_\xi)$
Получая новое вероятностное пространство, мы упрощаем и формализуем работу, так как можем не учитывать природу и структуру исходного пространства

Def. Функция $p(B), B \in \mathcal{B}(\mathbb{R})$, ставящая в соответствие каждому Борелевскому множеству вероятность, называется распределением случайной величины ξ

Основные типы распределения

a) Дискретное
b) Абсолютно непрерывное
c) Сингулярное
d) Смешанное

Дискретная случайная величина

Def. Случайная величина ξ имеет дискретное рапределение, если она принимает не более, чем счетное число значений. То есть существует конечный или счетный набор чисел

$\{x_1, x_2, \dots, x_n, \dots\}$ такой, что $p(\xi = x_i) = p_i > 0$ и $\sum\limits_{i=0}^{\infty} p_i = 1$

Таким образом, дискретная случайная величина (ДСВ) задается законом распределения:

ξ	x_1	x_1	$\dots$	x_n	$\dots$
p	p_1	p_1	$\dots$	p_n	$\dots$

- значения случайной величины
- вероятности этих значений

($\sum\limits_{i=0}^{\infty} p_i = 1$ - условие нормировки)

Ex. 1. кость, $\xi(\omega) = i$ - число выпавших очков

ξ	1	2	3	4	5	6
p	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Ex. 2. все распределения из предыдущих лекций (биномиальное, геометрическое, гипергеометрическое, Пуассона)

Ex. 3. индикатор события A: $I_A(\omega) = \begin{cases} 0, & \omega \notin A \text{ - событие } A \text{ не происходит} \\ 1, & \omega \in A \text{ - событие } A \text{ происходит} \end{cases}$

Числовые характеристики дискретных случайных величин

I. Математическое ожидание (среднее значение, полезность)

Def. Математическим ожиданием $E\xi$ случайной величины ξ называется число

$$E\xi = \sum_{i=1}^{\infty} x_i p_i$$

при условии, что данный ряд сходится абсолютно

Nota. Если $E\xi = \sum\limits_{i=1}^{\infty} x_i p_i = \infty$, то говорят, что матожидание не существует

При условной сходимости ряда при перестановке членов сумма изменяется, поэтому необходима абсолютная

Физический смысл: Среднее значение - число, вокруг которого группируются значения случайной величины, центр тяжести точек x_i с весами p_i

Статистический смысл: среднее арифметическое наблюдаемых значений случайной величины при большом числе реальных экспериментов

II. Дисперсия

Def. Дисперсией $D\xi$ случайной величины ξ называют среднее квадратов ее отклонения от математического ожидания:

$D\xi = E(\xi - E\xi)^2$ или $D\xi = \sum\limits_{i=0}^{\infty}(x_i - E\xi)^2 p_i$ при условии, что данный ряд сходится
В противном случае говорится, что дисперсии не существует

Nota. Дисперсию обычно удобно считать по формуле $D\xi = E\xi^2 - (E\xi)^2 = \sum\limits_{i=1}^{n} x_i^2 p_i - E\xi^2$
Смысл - квадрат среднего разброса (рассеивания) значения случайной величины относительно ее математического ожидания

III. Среднее квадратическое отклонение

Def. Средним квадратическим отклонением (СКО) σ_ξ называется величина $\sigma_\xi = \sqrt{D\xi}$
Смысл - средний разброс

Ex. 1. Кость

ξ	1	2	3	4	5	6
p	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

$E\xi = \sum\limits_{i=1}^{6} x_i p_i = 3.5$ (в данном случае ср. арифм.)
$D\xi = \sum\limits_{i=1}^{6}(x_i - E\xi)^2 p_i = 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + 3^2 \cdot \frac{1}{6} + 4^2 \cdot \frac{1}{6} + 5^2 \cdot \frac{1}{6} + 6^2 \cdot \frac{1}{6} - 3.5^2 = \frac{35}{12}$
$\sigma_\xi = \sqrt{D\xi} \approx 1.79$

Ex. 2. Индикатор события A: $I_A(\omega) = \begin{cases} 0, \omega \notin A \text{ - событие } A \text{ не происходит} \\ 1, \omega \in A \text{ - событие } A \text{ происходит} \end{cases}$

ξ	0	1
p	$1 - P(A)$	$P(A)$

$E\xi = 0 \cdot (1 - P(A)) + 1 \cdot P(A) = P(A)$
$D\xi = 0^2 \cdot (1 - P(A)) + 1^2 P(A) - P(A)^2 = P(A)(1 - P(A)) = pq$
$\sigma_\xi = \sqrt{pq}$

Свойства матожидания и дисперсии

Th. 1. Случайная величина ξ имеет вырожденное распределение, если $\xi(\omega) = \text{const} \quad \forall \omega \in \Omega$

ξ	C
p	1

$E\xi = C \qquad D\xi = 0$

Th. 2. Свойство сдвига: $E(\xi+C) = E\xi+C; D(\xi+C) = D\xi$

Th. 3. Свойство растяжения:
$E(C\xi) = CE\xi$
$D(C\xi) = C^2 D\xi$

Lab. 2-3 доказать

Th. 4. $E(\xi+\eta) = E\xi+E\eta$ (из третьего свойства матожидание - линейная функция)

□

$\sqsupset x_i, y_i$ - значения случайных величин ξ, η, а p_i и q_i - их соответствующие вероятности
$E(\xi+\eta) = \sum\limits_{i,j}(x_i+y_j)p(\xi = x_i \text{ и } \eta = y_j) = \sum\limits_{i} x_i \sum\limits_{j} p(\xi = x_i \text{ и } \eta = y_j) + \sum\limits_{j} y_j \sum\limits_{i} p(\xi = x_i \text{ и } \eta = y_j) = \sum\limits_{i} x_i p(\xi = x_i) + \sum\limits_{j} y_j p(\eta = y_j) = E\xi + E\eta$

□

Def. Дискретные случайные величины ξ и η независимы, если $p(\xi = x_i, \eta = y_i) = p(\xi = x_i)\cdot p(\eta = y_i)$ $\forall i, j$
То есть случайные величины принимают свои величины независимо друг от друга

Th. 5. Если случайные величины ξ и η независимы, то $E(\xi\eta) = E\xi\cdot E\eta$; обратное неверно

□

$E(\xi\eta) = \sum\limits_{i,j} x_i y_i p(\xi = x_i, \eta = y_j) = \sum\limits_{i} x_i \sum\limits_{j} y_j p(\xi = x_i, \eta = y_j) = \sum\limits_{i} x_i \sum\limits_{j} y_j p(\xi = x_i)p(\eta = y_j) = \sum\limits_{i} x_i p(\xi = x_i) \sum\limits_{j} y_j p(\eta = y_j) = E\xi\cdot E\eta$

□

Th. 6. $D\xi = E\xi^2 - (E\xi)^2$

□

$D\xi = E(\xi - E\xi)^2 = E(\xi^2 - 2\xi E\xi + (E\xi)^2) = E\xi^2 - 2E\xi E\xi + E((E\xi)^2) = E\xi^2 - 2(E\xi)^2 + (E\xi)^2 = E\xi^2 - (E\xi)^2$

□

Def. $D(\xi+\eta) = D\xi + D\eta + 2\operatorname{cov}(\xi,\eta)$, где $\operatorname{cov}(\xi,\eta) = E(\xi\eta) - E\xi E\eta$ - ковариация случайных величин (равна 0 при независимых величинах) - индикатор наличия связи между случайными величинами

□

$D(\xi+\eta) = E(\xi+\eta)^2 - (E(\xi+\eta))^2 = E\xi^2 + 2E(\xi\eta) + E\eta^2 - (E\xi + E\eta)^2 = E\xi^2 + E\eta^2 + 2E(\xi\eta) - (E\xi)^2 - (E\eta)^2 - 2E\xi E\eta = D\xi + D\eta + 2\mathrm{cov}(\xi,\eta)$

□

Th. 7. Если случайные величины ξ и η независимы, то $D(\xi+\eta) = D\xi + D\eta$

□

Если ξ и η независимы, то $\mathrm{cov}(\xi,\eta) = 0$ и $D(\xi+\eta) = D\xi + D\eta$

□

Th. 8. Общая формула дисперсии суммы: $D(\xi_1 + \xi_2 + \cdots + \xi_n) = \sum\limits_{i=1}^{n} D\xi_i + 2\sum\limits_{i,j(i\neq j)} \mathrm{cov}(\xi_i, \xi_j)$

Другие числовые характеристики

Моменты старших порядков

а) $m_k = E\xi^k$ - момент k-ого порядка случайной величины ξ (также называют начальным моментом)

б) $\mu_k = E(\xi - E\xi)^k$ - центральный момент k-ого порядка

$E\xi = m_1$ - момент первого порядка

$E\xi^2 = m_2$ - момент второго порядка

$D\xi = E(\xi - E\xi)^2$ - центральный момент второго порядка

Nota. Центральные моменты можно выразить через обычный момент:

$\mu_2 = D\xi = E\xi^2 - (E\xi)^2 = m_2^2 - m_1^2$

$\mu_3 = m_3 - 3m_2m_1 + 2m^3$

$\mu_4 = m_4 - 4m_3m_2 + 6m_2m_1^2 - 3m_1^4$

Ex. Разберем задачу Бюффона с точки зрения матожидания (для простоты l - ширина доски): пусть $p(A)$ - пересечет стык, $\xi = I_A$ - число пересечений. Тогда матожидание $E\xi = EI_A = P(A)$ Заметим, что при изменении длины иглы с l до $2l$ матожидание пересекаемых стыков увеличивается в два раза. Помимо этого можно составить из k игл ломаную, матожидание стыков которой будет равно $kE\xi$

Заметим, что такое работает и в обратную сторону: при уменьшении иглы в k раз матожидание равно $\frac{E\xi}{k}$
Теперь сделаем замкнутый многоугольник из игл, получим, что матожидание в таком случае $P\frac{E\xi}{l}$, где P - периметр
В пределе строим круг диаметра l - он всегда пересечет линии стыка 2 раза, значит матожидание $E_o = P_o\frac{E\xi}{l} = 2$
Длина окружность $P_o = \pi l$, получаем $E\xi = \frac{2l}{P_o} = \frac{2l}{\pi l} = \frac{2}{\pi}$

Лекция 7

Стандартное дискретное распределение

I. Распределение Бернулли

Распределение Бернулли B_p (с параметром $0 < p < 1$)
ξ - число успехов при одном испытании, p - вероятность успеха при одном испытании

ξ	0	1
p	$1 - P(A)$	$P(A)$

Матожидание: $E\xi = p$
Дисперсия: $D\xi = p(1-p) = pq$

Ex. Индикатор события $I_A \in B_p$ как раз имеет распределение Бернулли, где $p = P(A)$

II. Биномиальное распределение

Биномиальное распределение $B_{n,p}$ (с параметрами n, p)
ξ - число успехов в серии из n испытаний, p - вероятность успеха при одном испытании
$p(\xi = k) = C_n^k p^k q^{n-k},\ k = 0, 1, \ldots, n \Longleftrightarrow \xi \in B_{n,p}$

ξ	0	1	$\ldots$	k	$\ldots$	n
p	q^n	$nq^{n-1}p$	$\ldots$	$C_n^k p^k q^{n-k}$	$\ldots$	p^n

Заметим, что $\xi = \xi_1 + \xi_2 + \cdots + \xi_n$, где $\xi_i \in B_p$ - число успехов при i-ой испытании
$E\xi_i = p; \quad D\xi_i = pq$
$E\xi = E\xi_1 + \cdots + E\xi_n = p + \cdots + p = \boxed{np}$
$D\xi = D\xi_1 + \cdots + D\xi_n = pq + \cdots + pq = \boxed{npq}$

III. Геометрическое распределение

Геометрическое распределение G_p (с параметром p)

ξ - номер 1-ого успешного испытания в бесконечной серии

$p(\xi = k) = q^{k-1}p,\ k = 1, 2, 3, \dots \iff \xi \in G_p$

ξ	1	2	...	k	...
p	p	qp	...	$q^{k-1}p$	...

Матожидание $E\xi = \sum\limits_{k=1}^{\infty} kp(\xi = k) = \sum\limits_{k=1}^{\infty} kq^{k-1}p = p\sum\limits_{k=1}^{\infty} kq^{k-1} = p\sum\limits_{k=1}^{\infty}(q^k)' = p\left(\sum\limits_{k=1}^{\infty}(q^k)\right)' = p\left(\frac{1}{1-q}\right)' = \frac{p}{p^2} = \frac{1}{p}$

$E\xi^2 = \sum\limits_{k=1}^{\infty} k^2q^{k-1}p = p\sum\limits_{k=1}^{\infty} k(k-1)q^{k-1} + \sum\limits_{k=1}^{\infty} kq^{k-1}p = pq\sum\limits_{k=1}^{\infty} k(k-1)q^{k-2} + E\xi = pq(\sum\limits_{k=1}^{\infty} q^k)'' + \frac{1}{p} = pq\left(\frac{1}{1-q}\right)'' + \frac{1}{p} = 2pq\frac{1}{(1-q)^3} + \frac{1}{p} = 2pq\frac{1}{p^3} + \frac{1}{p} = \frac{2q}{p^2} + \frac{1}{p}$

$D\xi = E\xi^2 - (E\xi)^2 = \frac{2q}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{q}{p^2}$

IV. Распределение Пуассона

Распределение Пуассона Π_λ (с параметром $\lambda > 0$)

Def. Случайная величина ξ имеет распределение Пуассона с параметром $\lambda > 0$, если $p(\xi = k) = \frac{\lambda^k}{k!}e^{-\lambda},\ k = 0, 1, 2, \dots$

ξ	0	1	...	k	...
p	$e^{-\lambda}$	$\lambda e^{-\lambda}$	...	$\frac{\lambda^k}{k!}e^{-\lambda}$	...

Покажем корректность определения - докажем, что сумма нижней строки равна 1:

$\sum\limits_{k=0}^{\infty} p_k = \sum\limits_{k=0}^{\infty}\frac{\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}\underbrace{\sum\limits_{k=0}^{\infty}\frac{\lambda^k}{k!}}_{\text{ряд Тейлора для } e^x} = e^{-\lambda}e^{\lambda} = 1$

$E\xi = \sum\limits_{k=0}^{\infty} k\cdot\frac{\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}\sum\limits_{k=0}^{\infty}\frac{\lambda^k}{(k-1)!} = \lambda e^{-\lambda}\sum\limits_{k=0}^{\infty}\frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda}e^{\lambda} = \lambda = np$

$E\xi^2 = \sum\limits_{k=0}^{\infty} k^2\cdot\frac{\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}\sum\limits_{k=2}^{\infty} k(k-1)\frac{\lambda^k}{k!} + e^{-\lambda}\sum\limits_{k=1}^{\infty} k\frac{\lambda^k}{k!} = \lambda^2e^{-\lambda}\sum\limits_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!} + \lambda e^{-\lambda}\sum\limits_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!} = \lambda^2e^{-\lambda}e^{\lambda} + \lambda e^{-\lambda}e^{\lambda} = \lambda^2 + \lambda$

$D\xi = E\xi^2 - (E\xi)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$

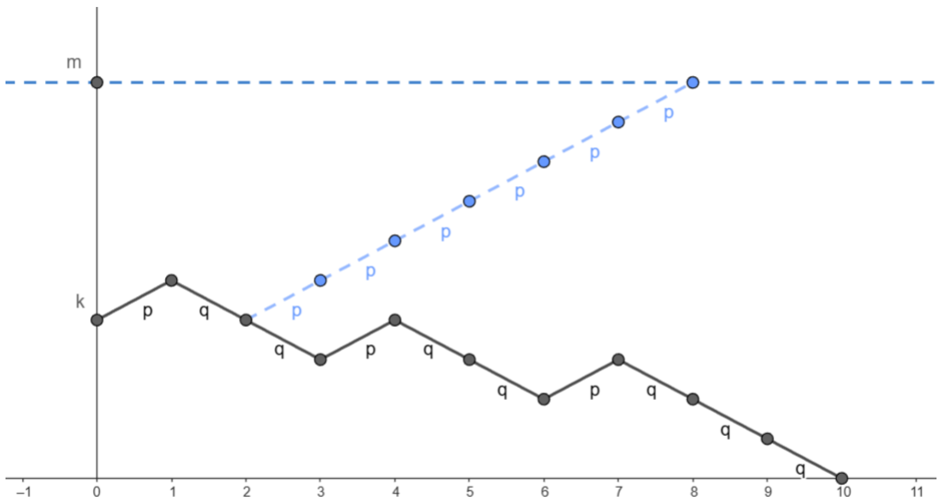
Задача о разорении игрока

Постановка задачи: играют 2 игрока, вероятность выигрыша первого игрока в одной игре равна p, $q = 1 - p$ - вероятность его проигрыша (выигрыш второго)

В каждой игре разыгрывается 1 биткоин. Капитал первого игрока - k биткоинов, $m - k$ биткоинов - капитал второго

Найти вероятность разорения первого игрока

Траектория капитала первого игрока будет выглядить как-то так:



Пусть r_k - интересующая нас вероятность разорение игрока при капитале k (то есть достижения оси абсцисс на графике)

$r_k = p \cdot r_{k+1} + qr_{k-1}$

$pr_{k+1} - r_k + (1-p)r_{k-1} = 0, \quad r_0 = 1, r_m = 0$

$p\lambda^2 - \lambda + (1-p) = 0$

$D = 1 - 4p(1-p) = 4p^2 - 4p + 1 = (2p-1)^2$

$\lambda_{1,2} = \dfrac{1 \pm (2p-1)}{2p}; \quad \lambda_1 = 1; \lambda_2 = \dfrac{2-2p}{2p} = \dfrac{q}{p}$

Обозначим $\lambda = \dfrac{q}{p}$

Рассмотрим два случая:

- $p \neq \dfrac{1}{2}$

  Тогда общее решение: $r_k = C_1\lambda_1^k + C_2\lambda_2^k = C_1 + C_2\lambda^k$

  Найдем частное решение:

  $$\begin{cases} 1 = C_1 + C_2 \\ 0 = C_1 + C_2\lambda^m \end{cases} \iff \begin{cases} C_1 = 1 - C_2 \\ 1 - C_2 + C_2\lambda_m = 0 \end{cases} \iff \begin{cases} C_1 = 1 - C_2 \\ C_2(1-\lambda_m) = 1 \end{cases} \iff \begin{cases} C_1 = 1 - \frac{1}{1-\lambda^m} = \frac{-\lambda^m}{1-\lambda^m} \\ C_2 = \frac{1}{1-\lambda^m} \end{cases}$$

  $$r_k = \frac{-\lambda^m}{1-\lambda^m} + \frac{1}{1-\lambda^m}\lambda^k = \frac{\lambda^k - \lambda^m}{1-\lambda^m}$$

Посмотрим, что будет происходит при бесконечной игре (то есть когда $m \to \infty$ - капитал неограничен)

1) $p < q$, то есть $\lambda > 1$. Тогда $\lambda^m \to \infty$, $r_k = \dfrac{\lambda^k - \lambda^m}{1 - \lambda^m} = \dfrac{\frac{\lambda^k}{\lambda_m} - 1}{\frac{1}{\lambda^m} - 1} \underset{n\to\infty}{\longrightarrow} 1$ - то есть первый игрок гарантированно разорится

2) $p > q$, то есть $\lambda < 1$. Тогда $\lambda^m \to 0$, $r_k = \dfrac{\lambda^k - \lambda^m}{1 - \lambda^m} \underset{n\to\infty}{\longrightarrow} \lambda^k$ - то есть $r_k = \left(\dfrac{q}{p}\right)^k$

- $p = \dfrac{1}{2} \Longrightarrow D = 0$
  Тогда $\lambda_1 = \lambda_2 = 1$
  Общее решение: $r_k = C_1\lambda^k + C_2 k\lambda_k = C_1 + C_2 k$
  Частное решение:
  $\begin{cases} 1 = C_1 \\ 0 = C_1 + C_2 m \end{cases} \Longleftrightarrow \begin{cases} 1 = C_1 \\ -1 = C_2 m \end{cases} \Longleftrightarrow \begin{cases} 1 = C_1 \\ C_2 = -\frac{1}{m} \end{cases}$
  $r_k = 1 - \dfrac{k}{m}$
  При бесконечной игре:
  $r_k = 1 - \dfrac{k}{m} \underset{m\to\infty}{\longrightarrow} 1$ - то есть при равной игре игрок неминуемо разорится

Случайное блуждание на прямой

Пусть в начальный момент времени находимся в начале координат. С вероятностью p идем на единицу вправо, с вероятностью q - влево

При $p = \dfrac{1}{2}$ мы рано или поздно попадем в любую точку числовой прямой

Можно привести аналогию с орлянкой: рано или поздно каждый игрок будет при сколь угодно большом выигрыше

Посмотрим на орлянку как на распределение Бернулли:

ξ	-1	1
p	$\frac{1}{2}$	$\frac{1}{2}$

$E\xi = 0; \quad D\xi = 1$

Пусть ξ - выигрыш первого после n игр.

$E\xi = \sum\limits_{i=1}^{n} E\xi_i = 0$

$D\xi = \sum\limits_{i=1}^{n} D\xi_i = n$

$\sigma_\xi = \sqrt{n}$ - среднее квадратическое отклонение

Это означает, что при большом n СКО поглотит всю числовую прямую

$\dfrac{S_n}{n} \to E\xi$

Закон больших чисел в этой ситуации говорит, что точка останется у 0, однако в то же время она может оказаться на любой точке на числовой прямой

Ex. По n конвертам случайным образом раскладывается m писем. Случайная величина ξ - число писем в своих конвертах

$\sqsupset A_i$ - число i письма в своем конверте, $\xi_i = I_A = \begin{cases} 0, & i\text{-ое письмо в не своем конверте} \\ 1, & i\text{-ое письмо в своем конверте} \end{cases}$

$\xi = \sum\limits_{i=1}^{n} \xi_i$

$E\xi_i = P(A_i) = \dfrac{1}{n}$

$D\xi_i = pq = \dfrac{1}{n}(1 - \dfrac{1}{n}) = \dfrac{n-1}{n^2}$

$E\xi = \sum\limits_{i=1}^{n} E\xi_i = n\dfrac{1}{n} = 1$ - в среднем будет одно письмо в своем конверте

$D\xi = D(\xi_1 + \dots + \xi_n) = \sum\limits_{i=1}^{n} D\xi_i + 2\sum\limits_{i<j} \operatorname{cov}(\xi_i, \xi_j)$

Найдем ковариацию:

$\operatorname{cov}(\xi_i, \xi_j) = E\xi_i\xi_j - E\xi_i E\xi_j = \dfrac{1}{n(n-1)} - \dfrac{1}{n}\dfrac{1}{n} = \dfrac{n-(n-1)}{n^2(n-1)} = \dfrac{1}{n^2(n-1)}$

Заметим, что для любых $i, j, i < j$: $\xi_i\xi_j = \begin{cases} 0, & \text{если хотя бы одно не в своем} \\ 1, & \text{если оба в своем} \end{cases}$

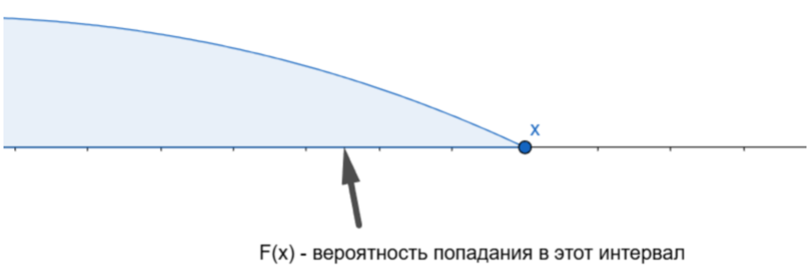
То есть $\xi_i\xi_j \in B_p$ и $E\xi_i\xi_j = P(\text{оба письма в своих}) = \dfrac{1}{n(n-1)}$

Получаем: $D\xi = n\dfrac{n-1}{n^2} + 2\dfrac{n(n-1)}{2}\dfrac{1}{n^2(n-1)} = \dfrac{n-1}{n} + \dfrac{1}{n} = 1$

Лекция 8

Функция распределения

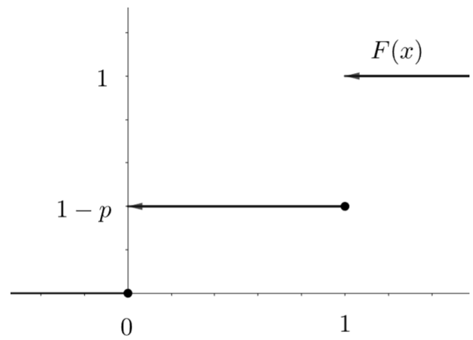
Def. Функция распределения $F_\xi(x)$ случайной величины ξ называется функция $F_\xi(x) = P(\xi < x)$



Ex. $\xi \in B_p$

ξ	0	1
p	$1-p$	p

$$F_\xi(x) = \begin{cases} 0 & x \le 0, \\ 1-p & 0 < x \le 1, \\ 1 & x > 1 \end{cases}$$



Свойства функции распределения

1) $F(x)$ ограничена $0 \le F(x) \le 1$

2) $F(x)$ неубывающая функция: $x_1 < x_2 \Longrightarrow F(x_1) \le F(x_2)$

$x_1 < x_2 \Longrightarrow \{\xi < x_1\} \subset \{\xi < x_2\} \Longrightarrow p(\xi < x_1) \le p(\xi < x_2)$, то есть $F(x_1) \le F(x_2)$

3) $p(\alpha \le \xi < \beta) = F(\beta) - F(\alpha)$

$p(\xi < \beta) = p(\xi < \alpha) + p(\alpha \le \xi < \beta) \Longrightarrow F(\beta) = F(\alpha) + p(\alpha \le \xi < \beta)$

Nota. Функция распределения $F(x)$ - вероятность попадания в интервал $(-\infty; x)$. Так как Борелевская σ-алгебра порождается такими интервалами, то распределение полностью задается этой функцией

4) $\lim\limits_{x\to-\infty} F(x) = 0; \quad \lim\limits_{x\to+\infty} F(x) = 1$

Так как $F(x)$ монотонна и ограничена, то эти пределы существуют. Поэтому достаточно доказать эти пределы для некоторой последовательности $x_n \to \pm\infty$

$\exists A_n = \{n-1 \le \xi < n, n \in Z\}$ - несовместные события, так как $\mathbb{R} = \bigcup\limits_{n=-\infty}^{\infty} A_n$, то по аксиоме счетной аддитивности, вероятность $p(\xi \in \mathbb{R}) = 1 = \sum\limits_{n=-\infty}^{\infty} P(A_n) = \lim\limits_{N\to\infty} \sum\limits_{n=-N}^{N} p(n-1 \le \xi < n) =$

$\lim\limits_{N\to\infty} \sum\limits_{n=-N}^{N} (F(n) - F(n-1)) = \lim\limits_{N\to\infty} (F(N) - F(-N-1)) = \lim\limits_{N\to\infty} F(N) - \lim\limits_{N\to-\infty} F(N) = 1$

$\Longrightarrow \lim\limits_{N\to\infty} F(N) = 1 + \lim\limits_{N\to-\infty} F(N)$

Так как $\lim\limits_{N\to\infty} F(N) \le 1$ и $\lim\limits_{N\to-\infty} F(N) \ge 0$, то $\lim\limits_{N\to\infty} F(N) = 1$ и $\lim\limits_{N\to-\infty} = 0$

5) $F(x)$ непрерывна слева: $F(x_0 - 0) = F(x_0)$

Этот предел существует в силу монотонности и ограниченности функции, поэтому рассмотрим последовательность событий $B_n = \{x_0 - \dfrac{1}{n} \le \xi < x_0, n \in Z\}$

Так как $B_1 \supset B_2 \supset \dots \supset B_n \supset \dots$ и $\bigcap\limits_{n=1}^{\infty} B_n = \varnothing$
То по аксиоме непрерывности $p(B_n) \to 0$
$P(B_n) = F(x_0) - F(x_0 - \frac{1}{n}) \to 0$
$F(x_0 - \frac{1}{n}) \to F(x_0)$
$\lim\limits_{x \to x_0 - 0} F(x) = F(x_0)$

6) Скачок в точке x_0 равен вероятности попадания в данную точку: $F(x_0+0) - F(x_0) = p(\xi = x_0)$ или $F(x_0+0) = p(\xi = x_0) + p(\xi < x_0) = p(\xi \le x_0)$

Этот предел существует в силу монотонности и ограниченности функции, поэтому рассмотрим последовательность событий $C_n = \{x_0 \le \xi < x_0 + \frac{1}{n}, n \in \mathrm{Z}\}$
Так как $C_1 \supset C_2 \supset \dots \supset C_n \supset \dots$ и $\bigcap\limits_{n=1}^{\infty} C_n = \varnothing$
То по аксиоме непрерывности $p(C_n) \to 0$
$P(C_n) = F(x_0 + \frac{1}{n}) - F(x_0) \to 0$
$p(x_0 \le \xi < x_0 + \frac{1}{n}) + p(\xi = x_0) \to p(\xi = x_0)$
$F(x_0 + \frac{1}{n}) - F(x_0) \to p(\xi = x_0)$
$F(x_0 + 0) - F(x_0) \to p(\xi = x_0)$

7) Если функция распределения непрерывна в точке $x = x_0$, то очевидно, что вероятность попадания в эту точка $p(\xi = x_0) = 0$ (следствие из 6 пункта)
8) Если $F(x)$ непрерывна $\forall x \in \mathbb{R}$, то $p(\alpha \le \xi < \beta) = p(\alpha < \xi < \beta) = p(\alpha \le \xi \le \beta) = p(\alpha < \xi \le \beta) = F(\beta) - F(\alpha)$

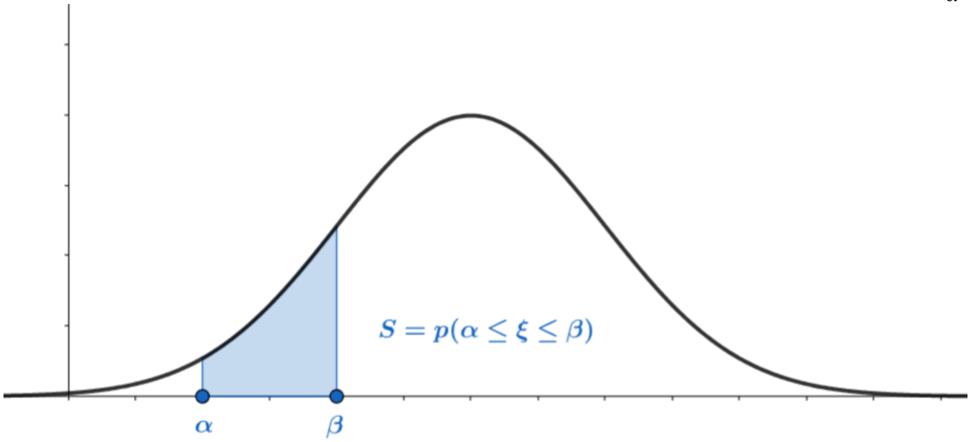
Th. Случайная величина ξ имеет дискретное распределение тогда и только тогда, когда ее функция распределения имеет ступенчатый вид

Абсолютно непрерывное распределение

Def. Случайная величина ξ имеет абсолютно непрерывное распределение, если существует $f_\xi(x)$ такая, что $\forall B \in \mathcal{B}(\mathbb{R})$ $p(\xi \in B) = \int_B f_\xi(x)dx$
Функция f_ξ называется плотностью распределения случайной величины
(в определении использует интеграл Лебега, так как B может быть не просто интервалом на $\mathbb{R}$)

Свойства плотности абсолютно непрерывного распределения

1) Вероятносто-геометрический смысл плотности: $p(\alpha \le \xi < \beta) = \int_{\alpha}^{\beta} f_{\xi}(x)dx$



2) Условие нормировки: $\int_{-\infty}^{+\infty} f_{\xi}(x)dx = 1$

Из определения, если $B = \mathbb{R}$

3) $F_{\xi}(x) = \int_{-\infty}^{x} f_{\xi}(x)dx$

Если $B = (-\infty; x)$, то $F_{\xi}(x) = p(\xi \in (-\infty; x)) = \int_{-\infty}^{x} f_{\xi}(x)dx$

4) $F_{\xi}(x)$ непрерывна

Из свойства непрерывности интеграла с верхним переменным пределом

5) $F_{\xi}(x)$ дифференцируема почти везде и $f_{\xi}(x) = F'_{\xi}(x)$ для почти всех x

По теореме Барроу

6) $f_{\xi}(x) \ge 0$ по определению и как производная неубывающей $F_{\xi}(x)$

7) $p(\xi = x) = 0\ \forall x \in \mathbb{R}$ - так как $F_{\xi}(x)$ непрерывна

8) $p(\alpha \le \xi < \beta) = p(\alpha < \xi < \beta) = p(\alpha \le \xi \le \beta) = p(\alpha < \xi \le \beta) = F(\beta) - F(\alpha)$

9) **Th.** Если $f(x) \ge 0$ и $\int_{-\infty}^{\infty} f(x)dx = 1$ (выполнены свойства 2 и 6), то $f(x)$ - плотность некоторого распределения

Числовые характеристики

Def. Математическим ожиданием $E\xi$ случайной абсолютно непрерывной величины ξ называется величина $E\xi = \int_{-\infty}^{\infty} x f_\xi(x)dx$ при условии, что данный интеграл сходится абсолютно, то есть $\int_{-\infty}^{\infty} |x| f_\xi(x)dx < \infty$

Def. Дисперсией $D\xi$ случайной величины ξ называется величина $D\xi = E(\xi - E\xi)^2 = \int_{-\infty}^{\infty}(x - E\xi)^2 f_\xi(x)dx$ при условии, что данный интеграл сходится

Nota. Вычислять удобно по формуле $D\xi = E\xi^2 - (E\xi)^2 = \int_{-\infty}^{\infty} x^2 f_\xi(x)dx - (E\xi)^2$

Def. Среднее квадратическое отклонение $\sigma_\xi = \sqrt{D\xi}$ определяется, как корень дисперсии

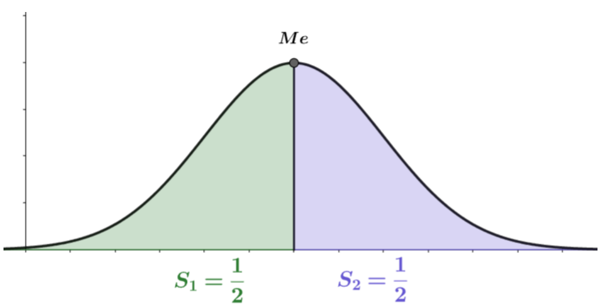
Смысл этих величин такой же, как и при дискретном распределении. Также свойства аналогичны тем, что и при дискретном распределении

Другие числовые характеристики

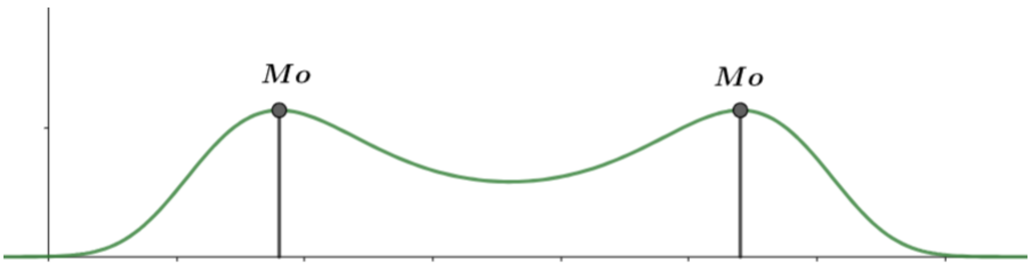
$m_k = E\xi^k = \int_{-\infty}^{\infty} x^k f_\xi(x)dx$ - момент k-ого порядка

$\mu_k = E(\xi - E\xi)^k = \int_{-\infty}^{\infty} (x - E\xi)^k f_\xi(x)dx$ - центральный момент k-ого порядка

Def. Медианой Me абсолютно непрерывной случайной величины ξ называется значение случайной величины ξ, такое что $p(\xi < Me) = p(\xi > Me) = \dfrac{1}{2}$



Def. Модой Mo случайной величины ξ называется точка локального максимума плотности



Сингулярное распределение

Def. Случайная величина ξ имеет сингулярное распределение, если $\exists B$ - Борелевское множество с нулевой мерой Лебега $\lambda(B) = 0$, такое что $p(\xi \in B) \in 1$, но $P(\xi = x) = 0 \ \ \forall x \in B$

Nota. Такое Борелевское множество состоит из несчетного множества точек, так как в противном случае по аксиоме счетной аддитивности $p(\xi \in B) = 0$. То есть при сингулярном распределении случайная величина ξ распределена на несчетном множестве меры 0

Nota. Так как $p(\xi = x) = 0 \ \forall x$, F_ξ непрерывна.

Ex. Сингулярное распределение получим, если возьмем случайную величину, функция распределения которой - лестница Кантора

$$F_\xi(x) = \begin{cases} 0 \quad x \le 0, \\ \frac{1}{2}F(3x) \quad 0 < x \le \frac{1}{3}, \\ \frac{1}{2} \quad \frac{1}{3} < x \le \frac{2}{3}, \\ \frac{1}{2} + \frac{1}{2}F(3x-2) \quad \frac{2}{3} < x \le 1, \\ 1 \quad x > 1 \end{cases}$$

Th. Лебега.
$\sqsupset F_\xi(x)$ - функция распределения ξ. Тогда $F_\xi(x) = p_1F_1(x) + p_2F_2(x) + p_3F_3(x)$, где $p_1 + p_2 + p_3 = 1$
F_1 - функция дискретного распределения
F_2 - функция абсолютно непрерывного распределения
F_3 - функция сингулярного распределения
То есть существуют только дискретное, абсолютно непрерывное, сингулярное распределения и их смеси

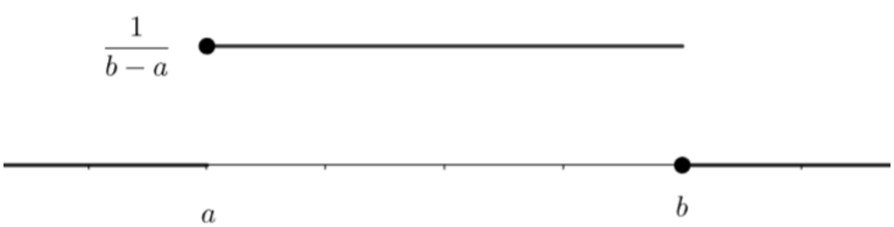
Лекция 9

Стандартное абсолютно непрерывное распределение

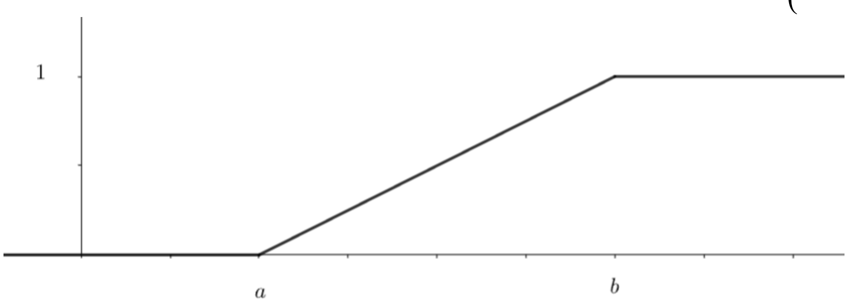
I. Равномерное распределение

Def. Случайная величина ξ имеет равномерное распределение $\xi \in U(a,b)$, если ее плотность на этом отрезке постоянна

Получаем функцию плотности $f_\xi(x) = \begin{cases} 0, & x < a \\ \frac{1}{b-a}, & a \le x < b \\ 0 & x \ge b \end{cases}$ $\frac{1}{b-a}$ из усл. нормировки



Из этого функция распределения $F(x) = \int_{-\infty}^{\infty} f(x)dx = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \le x < b \\ 1 & x \ge b \end{cases}$



Числовые характеристики:

$E\xi = \int_{-\infty}^{\infty} xf(x)dx = \int_a^b x\frac{1}{b-a}dx = \frac{1}{b-a}\frac{x^2}{2}\Big|_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}$

$E\xi^2 = \int_{-\infty}^{\infty} x^2f(x)dx = \int_a^b x^2\frac{1}{b-a}dx = \frac{1}{b-a}\frac{x^3}{3}\Big|_a^b = \frac{b^3-a^3}{3(b-a)} = \frac{b^2+ab+a^2}{3}$

$D\xi = E\xi^2 - (E\xi)^2 = \frac{b^2+ab+a^2}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{b^2-2ab+a^2}{12} = \frac{(b-a)^2}{12}$

$\sigma = \sqrt{D\xi} = \frac{b-a}{2\sqrt{3}}$

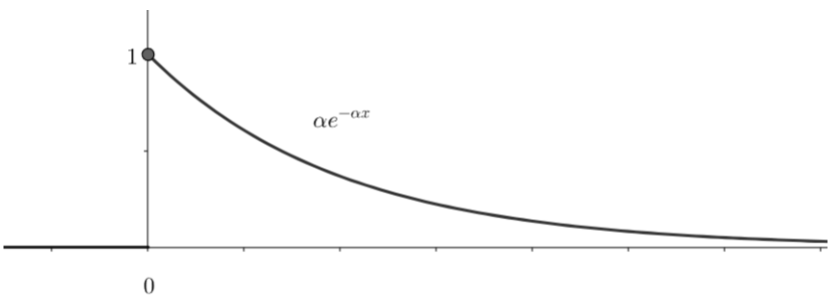
$p(\alpha < \xi < \beta) = \frac{\beta-\alpha}{b-a}$ при условии, что $\alpha, \beta \in [a,b]$

Nota. Примеры равномерного распределения: задача со временем, датчики случайных чисел имеют стандартное равномерное распределение $U(0,1)$

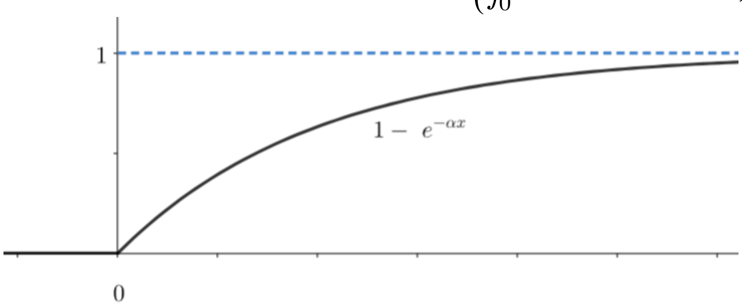
II. Показательное распределение

Def. Случайная величина ξ имеет показательное (или экспоненциальное) распределение с параметром $\alpha > 0$ (обозн. $\xi \in E_\alpha$), если ее плотность имеет вид:

$$f_\xi(x) = \begin{cases} 0, & x < 0 \\ \alpha e^{-\alpha x}, & x \geq 0 \end{cases}$$



Функция распределения $F_\xi(x) = \begin{cases} 0, & x < 0 \\ \int_0^x \alpha e^{-\alpha x} = 1 - e^{-\alpha x}, & x \geq 0 \end{cases}$



Числовые характеристики:

$E\xi = \int_{-\infty}^{\infty} x f(x) dx = \int_0^\infty x\alpha e^{-\alpha x} dx = \begin{bmatrix} u = x & du = dx \\ dv = \alpha e^{-\alpha x} \alpha x & v = -e^{-\alpha x} \end{bmatrix} = -xe^{-\alpha x}\Big|_0^\infty + \int_0^\infty e^{-\alpha x} dx = -\lim\limits_{x\to\infty} \frac{x}{e^{\alpha x}} - \frac{1}{\alpha} e^{-\alpha x}\Big|_0^\infty = -\lim\limits_{x\to\infty} \frac{1}{\alpha e^{\alpha x}} - \frac{1}{\alpha}(\lim\limits_{x\to\infty} e^{-\alpha x} - 1) = \frac{1}{\alpha}$

$E\xi^2 = \int_{-\infty}^{\infty} x^2 f(x) dx = \int_0^\infty x^2 \alpha e^{-\alpha x} dx = \begin{bmatrix} u = x^2 & du = 2xdx \\ dv = \alpha e^{-\alpha x} \alpha x & v = -e^{-\alpha x} \end{bmatrix} = -x^2 e^{-\alpha x}\Big|_0^\infty + 2\int_0^\infty xe^{-\alpha x} dx = \frac{2}{\alpha} \int_0^\infty \alpha x e^{-\alpha x} = \frac{2}{\alpha} E\xi = \frac{2}{\alpha^2}$

$D\xi = E\xi^2 - (E\xi)^2 = \frac{2}{\alpha^2} - \left(\frac{1}{\alpha}\right)^2 = \frac{1}{\alpha^2}$

$\sigma = \sqrt{D\xi} = \frac{1}{\alpha}$

$p(\alpha < \xi < \beta) = F(b) - F(a) = e^{-a\alpha} - e^{-b\alpha} \qquad a, b \geq 0$

Nota. Из непрерывных случайных величин только показательная обладает свойством нестарения

Th. $\exists \xi \in E_\alpha$. Тогда $p(\xi > x+y \mid \xi > x) = p(\xi > y) \qquad \forall x, y > 0$

□

$p(\xi > x+y \mid \xi > x) = \dfrac{p(\xi > x+y, \xi > x)}{p(\xi > x)} = \dfrac{1 - p(\xi < x+y)}{1 - p(\xi < x)} = \dfrac{1 - F(x+y)}{1 - F(x)} = \dfrac{e^{-\alpha(x+y)}}{e^{-\alpha x}} = e^{-\alpha y} = 1 - (1 - e^{-\alpha y}) = 1 - p(\xi < y) = p(\xi > y)$

□

Ex. 1. Время работы надежной микросхемы до поломки

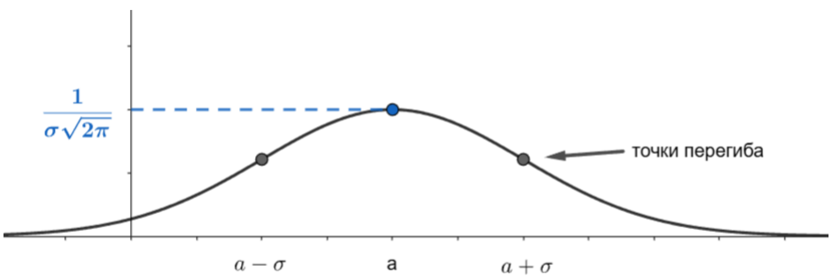
Ex. 2. Время между появлениями двух редких событий (через схему Пуассона)

Nota. Применется в системах массового обслуживания, теория надежности

III. Нормальное распределение (Гауссовское)

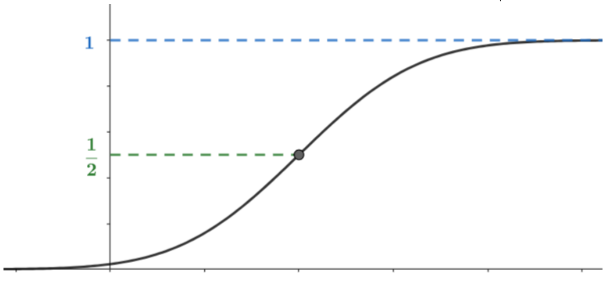
Def. Случайная величина ξ имеет нормальное распределение с параметрами a и σ^2 (обозн. $\xi \in N(a, \sigma^2)$), если ее плотность имеет вид:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad -\infty < x < \infty$$



Смысл параметров распределения: $a = E\xi$ - матожидание и медиана, σ - СКО, а $D\xi = \sigma^2$

Функция распределения: $F(x) = \dfrac{1}{\sigma\sqrt{2\pi}} \displaystyle\int_{-\infty}^{x} e^{-\frac{(t-a)^2}{2\sigma^2}} dt$



Проверим корректность определения - условие нормировки. Покажем, что $\displaystyle\int_{-\infty}^{\infty} f(x)dx = 1$

$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \begin{bmatrix} t = \frac{x-a}{\sigma\sqrt{2}} & dt = \frac{dx}{\sigma\sqrt{2}} \\ t(\pm\infty) = \pm\infty & dx = \sigma\sqrt{2}dt \end{bmatrix} = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-t^2}\sigma\sqrt{2}dt = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-t^2}dt =$
$\frac{1}{\sqrt{\pi}}\sqrt{\pi} = 1$ - верно

Ясно, что $m_k = \int_{-\infty}^{\infty} x^k f(x)dx = \int_{-\infty}^{\infty} x^k \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx$ - интеграл сходится абсолютно для любого k (степень e задавит полином)

$E\xi = m_1 = \int_{-\infty}^{\infty} x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = a$ в силу симметрии

Найдем дисперсию при помощи дифференцирования интеграла по параметру:

Из условия нормировки $\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = 1$

$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \sigma$

$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-a)^2}{2\sigma^2}} \left(-\frac{(x-a)^2}{2}(-2\sigma^{-3})\right) dx = 1$

$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-a)^2 e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \sigma^2 = D\xi$, получаем, что σ - СКО

Стандартное нормальное распределение

Def. Стандартным нормальным распределением называется нормальное распределение с параметрами $a = 0, \sigma^2 = 1$: $\xi \in N(0, 1)$

Плотность: $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ - функция Гаусса

$E\xi = 0;\ D\xi = 1$

Распределение: $F_0(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{z^2}{2}} dz$ - функция стандартного нормального распределения

Заметим, что $F_0(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{-\frac{z^2}{2}} dz + \frac{1}{\sqrt{2\pi}} \int_{0}^{x} e^{-\frac{z^2}{2}} dz = \frac{1}{2} + \Phi(x)$, где $\Phi(x)$ - функция Лапласа

Функция Лапласа нечетная и из соображения симметрии легко вычисляется для отрицательных x, однако большинство ПО используют $F_0(x)$

Связь между нормальным и стандартным нормальным распределениями

1) $\urcorner \xi \in N(a, \sigma^2)$. Тогда $F_\xi(x) = F_0\left(\frac{x-a}{\sigma}\right)$

□

$F_\xi(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(t-a)^2}{2\sigma^2}} dt = \begin{bmatrix} z = \frac{t-a}{\sigma} & t = \sigma z + a & dt = \sigma dz \\ z(-\infty) = -\infty & z(x) = \frac{x-a}{\sigma} & \end{bmatrix} = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\frac{x-a}{\sigma}} e^{-\frac{z^2}{2}} \sigma dz =$
$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-a}{\sigma}} e^{-\frac{z^2}{2}} dz = F_0\left(\frac{x-a}{\sigma}\right)$

□

2) Если $\xi \in N(a, \sigma^2)$, то $\eta = \dfrac{\xi - a}{\sigma} \in N(0,1)$ (процесс $\xi \to \eta$ называется стандартизацией)

□
$F_\eta(x) = p(\eta < x) = p\left(\dfrac{\xi - a}{\sigma} < x\right) = p(\xi < \sigma x + a) = F_\xi(\sigma x + a) = F_0\left(\dfrac{\sigma x + a - a}{\sigma}\right) = F_0(x)$, так как $F_\eta(x) = F_0(x)$, то $\eta \in N(0,1)$
□

3) $\exists \xi \in N(a, \sigma^2)$. Тогда $p(\alpha < \xi < \beta) = \Phi\left(\dfrac{\beta - a}{\sigma}\right) - \Phi\left(\dfrac{\alpha - a}{\sigma}\right)$

$$p(\alpha < \xi < \beta) = F_\xi(\beta) - F_\xi(\alpha) = F_0\left(\frac{\beta - a}{\sigma}\right) - F_0\left(\frac{\alpha - a}{\sigma}\right) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right)$$

4) Вероятность попадания в симметричный интервал (вероятность отклонения случайной величины от матожидания) $p(|\xi - a| < t) = 2\Phi\left(\dfrac{t}{\sigma}\right)$

$$p(|\xi - a| < t) = p(-t < \xi - a < t) = p(a - t < \xi < a + t) = \Phi\left(\frac{a + t - a}{\sigma}\right) - \Phi\left(\frac{a - t - a}{\sigma}\right) = \Phi\left(\frac{t}{\sigma}\right) - \Phi\left(-\frac{t}{\sigma}\right) = 2\Phi\left(\frac{t}{\sigma}\right)$$

Nota. Если через $F_0(x)$, то $p(|\xi - a| < t) = 2F_0\left(\dfrac{t}{\sigma}\right) - 1$
5) Правило 3 «сигм»: $p(|\xi - a| < 3\sigma) \approx 0.9973$ - попадание случайной величины нормального распределения в интервал $(a - 3\sigma, a + 3\sigma)$ близко к 1

$$p(|\xi - a| < 3\sigma) = 2\Phi\left(\frac{3\sigma}{\sigma}\right) = 2\Phi(3) = 2 \cdot 0.49685 = 0.9973$$

6) Свойство линейности: если случайная величина $\xi \in N(a, \sigma^2)$, то $\eta = \gamma\xi + b \in N(a\gamma + b, \gamma^2\sigma^2)$ (можем доказать при помощи свойств ранее, но мы докажем позже, используя другие методы)
7) Устойчивость относительно суммирования: если случайные величины $\xi_1 \in N(a_1, \sigma_1^2), \xi_2 \in N(a_2, \sigma_2^2)$, и они независимы, то $\xi_1 + \xi_2 \in N(a_1 + a_2, \sigma_1^2 + \sigma_2^2)$

Коэффициенты асимметрии и эксцесса

Def. 1. Асимметрией распределения называется число $A_s = E\left(\dfrac{\xi - a}{\sigma}\right)^3 = \dfrac{\mu^3}{\sigma^3}$

Def. 2. Эксцессом распределения называется число $E_s = E\left(\dfrac{\xi - a}{\sigma}\right)^4 - 3 = \dfrac{\mu^4}{\sigma^4} - 3$

Nota. Если случайная величина $\xi \in N(a, \sigma^2)$, то $A_s = E_s = 0$, таким образом, отличие этих характеристик от нуля характеризирует степень отклонения распределения. Благодаря этим и другим параметрам, можно проверять на практике, является ли распределение нормальным

Лекция 10

Преобразование случайных величин

Стандартизация случайной величины

Def. Пусть имеется случайная величина ξ. Соответствующей ей стандартной величиной называется случайная величина $\eta = \dfrac{\xi - E\xi}{\sigma}$

Свойства:

$E\eta = 0; D\eta = 1$

$$E\eta = E\frac{\xi - E\xi}{\sigma} = \frac{1}{\sigma}(E\xi - E\xi) = 0$$
$$D\eta = D\frac{\xi - E\xi}{\sigma} = \frac{1}{\sigma^2}D\xi = 1$$

Стандартизованная случайная величина не имеет единиц измерения, таким образом, ее свойства от них не зависят

Задача: пусть имеется функция $g(x)$ и случайная величина ξ, $\eta = g(\xi)$. Определить ее характеристики

Nota. Если ξ - дискретная случайная величина, то ее законы распределения находятся просто: значения x_i в верхней строке заменяем $g(x_i)$, вероятности остаются прежние. Поэтому будем рассматривать непрерывной случайной величины ξ

Nota. Возможна ситуация, когда ξ - абсолютно непрерывная случайная величина, $g(x)$ - непрерывна, но $g(\xi)$ имеет дискретное распределение

Линейное преобразование

Th. Пусть ξ имеет плотность $f_\xi(x)$, тогда $\eta = a\xi + b$, где $a \neq 0$, имеет плотность $f_\eta(x) = \dfrac{1}{|a|} f_\xi\left(\dfrac{x - b}{a}\right)$

Пусть $a>0$, тогда $F_\eta(x)=p(\eta<x)=p(a\xi+b<x)=p(\xi<\dfrac{x-b}{a})=\displaystyle\int_{-\infty}^{\frac{x-b}{a}} f_\xi(t)dt=$ $\left[\begin{matrix} t=\frac{y-b}{a} & dt=\frac{1}{a}dy & y=at+b \\ y(-\infty)=-\infty & y(\frac{x-b}{a})=x & \end{matrix}\right]=\displaystyle\int_{-\infty}^{x}\frac{1}{a}f_\xi(\frac{y-b}{a})dy\Longrightarrow \eta=\frac{1}{|a|}f_\xi(\frac{x-b}{a})$

Пусть $a<0$, тогда $F_\eta(x)=p(\eta<x)=p(a\xi+b<x)=p(\xi>\dfrac{x-b}{a})=\displaystyle\int_{\frac{x-b}{a}}^{\infty} f_\xi(t)dt=$ $\left[\begin{matrix} t=\frac{y-b}{a} & dt=\frac{1}{a}dy & y=at+b \\ y(\infty)=-\infty & y(\frac{x-b}{a})=x & \end{matrix}\right]=-\displaystyle\int_{-\infty}^{x}\frac{1}{a}f_\xi(\frac{y-b}{a})dy\Longrightarrow \eta=\frac{1}{|a|}f_\xi(\frac{x-b}{a})$

Следствие

1) Если $\xi\in N(a,\sigma^2)$, то $\eta=\gamma\xi+b\in N(a\gamma+b;\gamma^2\sigma^2)$

Так как $\xi\in N(a,\sigma^2)$, то $f_\xi(x)=\dfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)}{2\sigma^2}}, x\in\mathbb{R}$

Тогда $f_\eta(x)=\dfrac{1}{|\gamma|}\cdot\dfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(\frac{x-b}{\gamma}-a)^2}{2\sigma^2}}=\dfrac{1}{|\gamma|\sigma\sqrt{2\pi}}e^{-\frac{(x-b-a\gamma)^2}{2\sigma^2\gamma^2}}\Longrightarrow\eta\in N(b+a\gamma;\sigma^2\gamma^2)$

2) Если $\eta\in N(0,1)$ - стандартное нормальное распределение, то $\xi=\sigma\eta+a\in N(a,\sigma^2)$

3) Если $\eta\in U(0,1)$ - стандартное равномерное распределение и $a>0$, то $\xi=a\eta+b\in U(b,a+b)$

4) Если $\xi\in E_\alpha$, то $\alpha\xi\in E_1$

Монотонное преобразование

Th. Пусть $f_\xi(x)$ - плотность случайной величины ξ, $g(x)$ - строго монотонная функция. Тогда случайная величина $\eta=g(\xi)$ имеет плотность

$$f_\eta(x)=|h'(x)|f_\xi(h(x)),\qquad \text{где } h(g(x))=x$$

Если $g(x)$ не является монотонной функцией, то поступаем следующим образом: разбиваем $g(x)$ на интервалы монотонности, для каждого i-ого интервала находим $h_i(x)$ и плотность случайной величины ищем по *формуле Смирнова*: $f_\eta(x)=\displaystyle\sum_{i=0}^{n}|h_i'(x)|f_\xi(h_i(x))$

Квантильное преобразование

Th. 1. Пусть функция распределения случайной величины ξ $F_\xi(x)$ - непрерывная функция. Тогда $\eta=F(\xi)\in U(0,1)$ - стандартное равномерное распределение

Ясно, что $0 \le \eta \le 1$

а) $F(x)$ - строго возрастающая функция. Тогда $\exists F^{-1}(x)$ - обратная, $F_\eta(x) = p(\eta < x) = p(F(\xi) < x) = \begin{cases} 0, & x < 0 \\ p(\xi < F^{-1}(x)) = F(F^{-1}(x)) = x, & 0 \le x \le 1 \text{ - функция распределения } U(0,1) \\ 1, & x > 1 \end{cases}$

б) $F(x)$ - не является строго возрастающей функцией - то есть существуют участки постоянства, в этом случае определим F^{-1} как $F^{-1}(x) = \min\limits_t (t \mid F(t) = x)$ - то есть берем самую левую точку такого интервала

Тогда снова будет при $0 \le x \le 1$ $F_\eta(x) = p(\eta < x) = p(F(\xi) < x) = F(F^{-1}(x)) = x$

Сформулируем обратную теорему: пусть $F(x)$ - функция распределения (необязательно непрерывная) случайной величины ξ, обозначим $F^{-1}(x) = \inf\limits_t (t \mid F(t) \ge x)$.

В случае непрерывной $F(x)$ это определение совпадет с предыдущем

Th. 2. Пусть $\eta \in U(0,1)$ - стандартное равномерное распределение, $F(x)$ - произвольная функция распределения. Тогда $\xi = F^{-1}(\eta)$ имеет функцию распределения $F(x)$

Данное преобразование $\xi = F^{-1}(\eta)$ называют квантильным

Доказательство аналогично предыдущей теореме

Смысл: датчики случайных чисел имеют стандартное равномерное распределение, из теоремы следует, что при помощи датчика случайных чисел и квантильного преобразования мы сможем смоделировать любое нужно распределение

Ex. 1. Смоделируем показательное распределение E_α : $F_\alpha(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-\alpha x}, & x \ge 0 \end{cases}$

$\eta = 1 - e^{-\alpha x}$, $e^{-\alpha x} = 1 - \eta$, $x = -\dfrac{1}{\alpha}\ln(1-\eta)$ - функция, обратная к $F_\alpha(x)$

Если $\eta \in U(0,1)$, то $\xi = -\dfrac{1}{\alpha}\ln(1-\eta) \in E_\alpha$

Ex. 2. $\xi \in N(0,1)$, $F_0(x) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{x} e^{-\frac{z^2}{2}} dz$

Пусть $F_0^{-1}(x)$ - функция обратная к $F_0(x)$

Если $\eta \in U(0,1)$, то $F_0^{-1}(\eta) \in N(0,1)$

Характеристики преобразованной случайной величины

Th. Если ξ - дискретная случайная величина, то $Eg(\xi) = \sum\limits_{i=1}^{\infty} g(x_i) \cdot p_i = \sum\limits_{i=1}^{\infty} g(x_i) \cdot p(\xi = x_i)$

Для непрерывной случайной величины $Eg(\xi) = \int_{-\infty}^{\infty} g(x) f_\xi(x) dx$

Свойства моментов

1) Если $\xi \geq 0$, то $E\xi \geq 0$
2) Если $\xi \leq \eta$, то $E\xi \leq E\eta$

$\xi \leq \eta \Longrightarrow \eta - \xi \geq 0 \Longrightarrow E(\eta - \xi) \geq 0 \Longrightarrow E\eta - E\xi \geq 0 \Longrightarrow E\eta \geq E\xi$

3) Если $|\xi| \leq |\eta|$, то $E|\xi|^k \leq E|\eta|^k$
4) Если существует момент m_t случайной величины ξ, то существует m_s при $s < t$ (при условии, что интеграл/сумма сходятся)

Пусть $s < t$. Тогда $|x|^s \leq \min(1, |x|^t) \leq 1 + |x|^t$, так как при $|x| < 1$, $|x|^s \leq 1$ и при $|x| \geq 1$, $|x|^s \leq |x|^t$
$E|\xi|^s \leq E|\xi|^t + 1$ и если $E|\xi|^t$ существует (конечно), то $\exists E|\xi|^s$

Th. Неравенство Йенсена. Пусть функция $g(x)$ выпукла вниз, тогда для любой случайной величины ξ

$$Eg(\xi) \geq g(E\xi)$$

Nota. Если $g(x)$ выпукла вверх, знак неравенства меняется

Если $g(x)$ выпукла вниз, то в любой ее точке, можно провести прямую, лежащую не выше графика функции. То есть для любой x_0 существует $k(x_0)$ такой, что $g(x) \geq g(x_0) + k(x_0)(x - x_0)$
Пусть $x_0 = E\xi$, $g(E\xi) \geq g(E\xi) + k(E\xi)(x - E\xi)$
$Eg(\xi) \geq Eg(E\xi) + k(E\xi)\underbrace{(E\xi - E\xi)}_{=0}$
$Eg(\xi) \geq g(E\xi)$

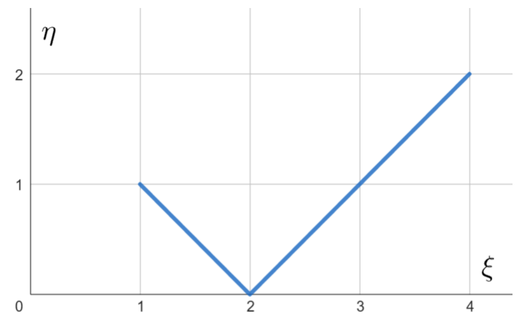
Следствие:

$Ee^{\xi} \geq e^{E\xi}, \quad E\xi^2 \geq (E\xi)^2, \quad E|\xi| \geq |E\xi|, \quad E\ln(\xi) \leq \ln(E\xi), \quad E\frac{1}{\xi} \geq \frac{1}{E\xi}$ при $\xi > 0$

Ex. на формулу Смирнова: дана плотность распределения

$$f_\xi(x) = \begin{cases} 0, & x < 1 \\ \frac{4}{3x^2}, & 1 \leq x \leq 4, 0 \\ x > 4 \end{cases}$$

Найти f_η для $\eta = |\xi - 2|$



Решение

$\xi \in [1, 4], \quad \eta \in [0, 2]$

$$\begin{cases} 0 \leq \eta \leq 1 \Longrightarrow h_1(\eta) = \eta + 2 \text{ и } h_2(\eta) = 2 - \eta & \text{- 2 ветви} \\ 1 < \eta \leq 2 \Longrightarrow h_1(\eta) = \eta + 2 & \text{- 1 ветвь} \end{cases}$$

$h_1'(\eta) = 1, h_2'(\eta) = -1 \qquad |h_1'(\eta)| = |h_2'(\eta)| = 1$

$$f_\eta(x) = \sum_i |h_i'(x)| f_\xi(h_i(x))$$

$$f_\eta(x) = \begin{cases} 0, & x < 0 \\ \frac{4}{3}\left(\frac{1}{(x+2)^2} + \frac{1}{(2-x)^2}\right), & 0 \leq x \leq 1 \\ \frac{4}{3}\frac{1}{(x+2)^2}, & 1 < x \leq 2 \\ 0 & x > 2 \end{cases}$$

Лекция 11

Сходимость случайных величин

Рассмотрим 3 вида сходимости:

- Сходимость «почти наверное»

  Def. Случайная величина ξ имеет свойство Cond «почти наверное», если вероятность $p(\xi$ имеет свойство Cond$) = 1$

  Nota. То есть $p(\xi$ не имеет свойство Cond$) = 0$
  $p(\omega \in \Omega \mid \xi(\omega)$ не имеет св-во Cond$)$

  Def. Последовательность случайных величин $\{\xi_n\}$ сходится «почти наверное» к случайной величине ξ при $n \to \infty$ ($\xi_n \xrightarrow{\text{п. н.}} \xi$), если $p(\omega \in \Omega \mid \xi_n(\omega) \underset{n\to\infty}{\longrightarrow} \xi(\omega)) = 1$
- Сходимость по вероятности

  Def. Последовательность случайных величин $\{\xi_n\}$ сходится по вероятности к случайной величине ξ при $n \to \infty$ ($\xi_n \xrightarrow{p} \xi$), если $\forall \varepsilon > 0 \quad p(|\xi_n - \xi| < \varepsilon) \underset{n\to\infty}{\longrightarrow} 1$

Nota. Не надо думать, что из сходимости по вероятности следует сходимости математического ожидания $\xi_n \xrightarrow{p} \xi \not\Longrightarrow E\xi_n \longrightarrow E\xi$

Th. Пусть $|\xi_n| \leq C = \text{const} \quad \forall n$
Тогда $\xi_n \xrightarrow{p} \xi \Longrightarrow E\xi_n \longrightarrow E\xi$

- Слабая сходимость
  Def. Последовательность случайных величин ξ_n слабо сходится к случайной величине ξ при $n \to \infty$ ($\xi_n \rightrightarrows \xi$), если $F_{\xi_n}(x) \longrightarrow F_\xi(x) \forall x$, где $F_\xi(x)$ - непрерывна

Связь между видами сходимости

Th. $\xi_n \xrightarrow{\text{п. н.}} \xi \Longrightarrow \xi_n \xrightarrow{p} \xi \Longrightarrow \xi_n \rightrightarrows \xi$

Th. Если $\xi_n \rightrightarrows C = \text{const}$, то $\xi_n \xrightarrow{p} C$

Если $\xi_n \rightrightarrows C$, то по определению $F_{\xi_n}(x) \longrightarrow F_C(x) = \begin{cases} 0, & x \leq C \\ 1, & x > C \end{cases} \quad \forall x \neq C$

$\forall \varepsilon > 0 \quad p(|\xi_n - C| < \varepsilon) = p(-\varepsilon < \xi_n - C < \varepsilon) = p(C - \varepsilon < \xi_n < C + \varepsilon) \geq p\left(C - \frac{\varepsilon}{2} < \xi_n < C + \varepsilon\right) = F_{\xi_n}(C + \varepsilon) - F_{\xi_n}\left(C - \frac{\varepsilon}{2}\right) = 1 - 0 = 1$
Так как $p(|\xi_n - C| < \varepsilon) \leq 1$, то по теореме о 2 милиционерах $p(|\xi_n - C| < \varepsilon) \underset{n\to\infty}{\longrightarrow} 1$ то есть по определению $\xi_n \xrightarrow{p} C$

Nota. В общем случае не только из слабой сходимости не следует сходимость по вероятности, но и бессмысленно говорить об этом, так как слабая сходимость - это сходимость не случайных величин, а их распределений

Ex. $\exists \xi_n \rightrightarrows \xi \in N(0, 1)$, тогда $\eta = -\xi \in N(0, 1)$, но ясно, что $\xi_n \xrightarrow{p} \eta = -\xi$ - неверно

Ключевые неравенства

В дальнейшем будем считать, что у случайных величин первый момент существует

I. Неравенство Маркова

Th. $p(|\xi| \geq \varepsilon) \leq \dfrac{E|\xi|}{\varepsilon} \quad \forall \varepsilon > 0$

$I_A(\omega) = \begin{cases} 0, & \omega \notin A \quad - A \text{ нет} \\ 1, & \omega \in A \quad - A \text{ есть} \end{cases}$

$EI_A = p(A)$

$|\xi| \geq |\xi| \cdot I(|\xi| \geq \varepsilon) \geq \varepsilon I(|\xi| \geq \varepsilon)$

$E|\xi| \geq E(\varepsilon \cdot I(|\xi| \geq \varepsilon))$

$E|\xi| \geq \varepsilon \cdot E(\varepsilon I(|\xi| \geq \varepsilon)) = \varepsilon \cdot p(|\xi| \geq \varepsilon) \Longrightarrow p(|\xi| \geq \varepsilon) \leq \dfrac{E|\xi|}{\varepsilon}$

II. Неравенство Чебышева

Th. $P(|\xi - E\xi| \geq \varepsilon) \leq \dfrac{D\xi}{\varepsilon^2}$

$p(|\xi - E\xi| \geq \varepsilon) = p((\xi - E\xi)^2 \geq \varepsilon^2) \leq \dfrac{E(\xi - E\xi)^2}{\varepsilon^2} = \dfrac{D\xi}{\varepsilon}$

III. Правило «трех сигм»

Th. $P(|\xi - E\xi| \geq 3\sigma) \leq \dfrac{1}{9}$

По неравенству Чебышева $P(|\xi - E\xi| \geq 3\sigma) \leq \dfrac{D\xi}{(3\sigma)^2} = \dfrac{D\xi}{9\sigma^2} = \dfrac{1}{9}$

Среднее арифметическое независимых одинаково распределенных случайных величин

Пусть $\xi_1, \xi_2, \dots, \xi_n$ - независимые одинаково распределенные случайные величины с конечным вторым моментом

Обозначим $a = E\xi_i, d = D\xi_i, \sigma = \sigma_{\xi_i}, \quad 1 \leq i \leq n$

$S_n = \xi_1 + \dots + \xi_n$ - их сумма

$\dfrac{S_n}{n} = \dfrac{\xi_1 + \dots + \xi_n}{n}$ - среднее арифметическое

$E\left(\frac{S_n}{n}\right) = \frac{1}{n}(E\xi_1 + \dots + E\xi_n) = \frac{1}{n}na = a = E\xi_1$ - математическое ожидание не меняется
$D\left(\frac{S_n}{n}\right) = \frac{1}{n^2}(D\xi_1 + \dots + D\xi_n) = \frac{1}{n^2}nd = \frac{d}{n} = \frac{D\xi_1}{n}$ - дисперсия уменьшилась в n раз
$\sigma\left(\frac{S_n}{n}\right) = \frac{\sigma}{\sqrt{n}}$ - СКО уменьшилось в $\sqrt{n}$ раз

Законы больших чисел

I. Закон больших чисел Чебышева

Th. Пусть $\xi_1, \dots, \xi_n, \dots$ - последовательность независимых одинаково распределенных с конечным вторым моментом, тогда $\frac{\xi_1 + \dots + \xi_n}{n} \xrightarrow[n\to\infty]{p} E\xi_1$

Обозначим $a = E\xi_i, d = D\xi_i, \sigma = \sigma_{\xi_i}, \quad 1 \le i \le n$
$S_n = \sum\limits_{i=1}^{n} \xi_i$

Тогда по неравенству Чебышева $p\left(\left|\frac{S_n}{n} - a\right| \ge \varepsilon\right) = p\left(\left|\frac{S_n}{n} - E\left(\frac{S_n}{n}\right)\right| \ge \varepsilon\right) \le \frac{D\left(\frac{S_n}{n}\right)}{\varepsilon^2} = \frac{d}{n\varepsilon^2} \xrightarrow[n\to\infty]{} 0 \Longrightarrow p\left(|\frac{S_n}{n} - a| < \varepsilon\right) \xrightarrow[n\to\infty]{} 1$, то есть $\frac{S_n}{n} \xrightarrow{p} a$

Среднее арифметическое большое числа независимых одинаковых случайных величин «стабилизируется» около математического ожидания, «при $n \to \infty$ случайность переходит в закономерность»
Статистический смысл: при большом объеме n статистических данных среднее арифметическое данных дает достаточно точную оценку теоретического математического ожидания

Nota. При доказательстве получили полезную, хотя и грубую оценку: $p\left(\left|\frac{S_n}{n} - a\right| \ge \varepsilon\right) \le \frac{D\xi_i}{n\varepsilon^2}$

II. Закон больших чисел Бернулли

Th. Пусть v_n - число успехов из n независимых испытаний, $p = P(A)$ - вероятность успеха при одном испытании. Тогда $\frac{v_n}{n} \xrightarrow{p} P(A)$

При этом $P\left(\left|\frac{v_n}{n} - p\right| \le \varepsilon\right) \le \frac{p(1-p)}{n\varepsilon^2}$

$v_n = \xi_1 + \dots + \xi_n$, где $\xi_i \in B_p$ - число успехов при i-ом испытании
$E\xi_i = p; D\xi_i = pq$
$\frac{v_n}{n} \xrightarrow{p} E\xi_1 = p$
$p\left(\left|\frac{v_n}{n} - p\right| \geq \varepsilon\right) \leq \frac{D\xi_1}{n\varepsilon^2} = \frac{pq}{n\varepsilon^2}$

III. Закон больших чисел Хинчина

Th. $v_n = \xi_1 + \dots + \xi_n$ последовательность независимых одинаково распределенных случайных величин с конечным первым моментом, тогда $\frac{\xi_1 + \dots + \xi_n}{n} \xrightarrow{p} E\xi_i$

IV. Усиленный закон больших чисел Колмогорова

В условиях теоремы Хинчина $\frac{\xi_1 + \dots + \xi_n}{n} \xrightarrow{\text{п.н.}} E\xi_1$

V. Закон больших чисел Маркова

Th. Пусть имеется последовательность случайных величин $\xi_1, \dots, \xi_n, \dots$ с конечными вторыми моментами, таких что $D(S_n) = o(n^2)$. Тогда $\frac{S_n}{n} \xrightarrow{p} E\left(\frac{S_n}{n}\right)$ или $\frac{\xi_1 + \dots + \xi_n}{n} \xrightarrow{p} \frac{1}{n}(E\xi_1 + \dots + E\xi_n)$

По неравенству Чебышева $p\left(\left|\frac{S_n}{n} - E\left(\frac{S_n}{n}\right)\right| \geq \varepsilon\right) \leq \frac{D\left(\frac{S_n}{n}\right)}{\varepsilon^2} = \frac{D(S_n)}{n^2\varepsilon^2} = \frac{1}{\varepsilon^2}\frac{o(n^2)}{n^2} \longrightarrow 0 \Longrightarrow p\left(\left|\frac{S_n}{n} - E\left(\frac{S_n}{n}\right)\right| \leq \varepsilon\right) \longrightarrow 1$

Центральная предельная теорема

Th. Центральная предельная теорема (ЦПТ Ляпунова, ≈1901 год)
Пусть $\xi_1, \dots, \xi_n, \dots$ - последовательность независимых одинаково распределенных случайных величин с конечной дисперсией ($D\xi_1 < \infty$) и $S_n = \sum_{i=1}^{n} \xi_i$. Тогда имеет место слабая сходимость:

$$\frac{S_n - nE\xi_1}{\sqrt{nD\xi_1}} \rightrightarrows N(0,1)$$

Теорема показывает, что стандартизованная сумма слабо сходится к стандартному нормальному распределению

Nota. Можно представить в ином виде: $\sqsupset a = E\xi_i, \sigma = \sigma_{\xi_i}$, тогда $E\left(\frac{S_n}{n}\right) = a, \sigma\left(\frac{S_n}{n}\right) = \frac{\sigma}{\sqrt{n}}$, а $\frac{\frac{S_n}{n} - a}{\sigma\sqrt{n}} \rightrightarrows N(0,1)$

Nota. Другая, грубая, формулировка: $\frac{S_n}{n} \rightrightarrows N\left(a, \frac{\sigma^2}{n}\right)$

Лекция 12

Совместное распределение случайных величин

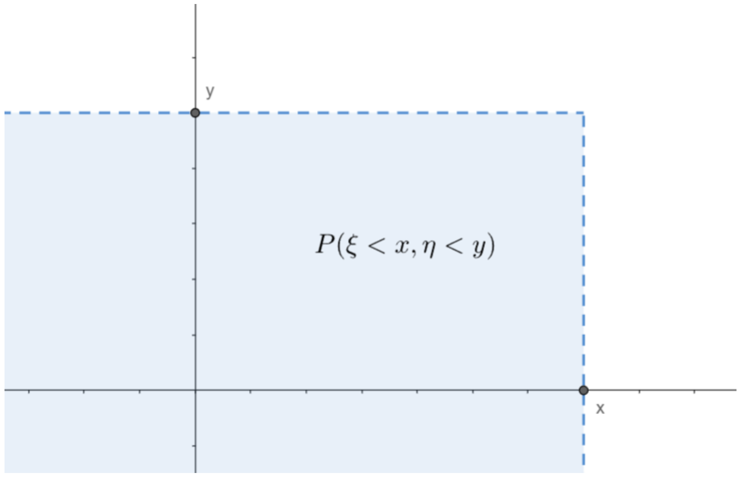
Пусть $\xi_1, \xi_2, \dots, \xi_n$ заданы на одном и том же вероятностном пространстве $(\Omega, \mathcal{F}, p)$

Def. Случайным вектором $\vec{\xi} = (\xi_1, \xi_2, \dots, \xi_n)$ называется упорядоченный набор случайных величин, заданных на одном вероятностном пространстве
Случайный вектор задает отображение $(\xi_1, \dots, \xi_n)(\omega) : \Omega \longrightarrow \mathbb{R}^n$
Поэтому случайный вектор еще называют многомерной случайной величиной, а соответствующее ей распределение многомерным распределением:
$\forall B \in \mathcal{B}(\mathbb{R}^n) \qquad P(B) = P(\omega \in \Omega \mid (\xi_1, \dots, \xi_n) \in B)$
Таким образом, получили новое вероятностное пространство. В качестве элементарных исходов берем точки многомерного пространства, а σ-алгебра - многомерное Борелевская σ-алгебра $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), P(B))$

Функция распределения

Def. Функцией совместного распределения случайных величин $\xi_1, \xi_2, \dots, \xi_n$ называется функция $F_{\xi_1, \xi_2, \dots, \xi_n}(x_1, x_2, \dots, x_n) = P(\xi_1 < x_1, \xi_2 < x_2, \dots, \xi_n < x_n)$
Nota. Распределение полностью задается функцией распределения

Nota. В дальнейшем, в основном, будем рассматривать системы из 2 случайных величин. Функция распределения в данном случае $F_{\xi,\eta}(x, y) = P(\xi < x, \eta < y)$ - вероятность попадания в эту область.



Свойства функции распределения

1. $0 \le F_{\xi,\eta}(x,y) \le 1$
2. $F_{\xi,\eta}(x,y)$ - неубывающая по каждому аргументу
3. $\lim\limits_{x\to-\infty} F_{\xi,\eta}(x,y) = \lim\limits_{y\to-\infty} F_{\xi,\eta}(x,y) = 0$, $\lim\limits_{\substack{x\to\infty\\ y\to\infty}} F_{\xi,\eta}(x,y) = 1$
4. Восстановление маргинального (частного) распределения: $\lim\limits_{x\to\infty} F_{\xi,\eta}(x,y) = F_\eta(y)$, и наоборот - $\lim\limits_{y\to\infty} F_{\xi,\eta}(x,y) = F_\xi(x)$
5. $F_{\xi,\eta}(x,y)$ - непрерывна слева по каждому аргументу
6. $P(x_1 \le \xi < x_2, y_1 \le \eta < y_2) = F_{\xi,\eta}(x_2,y_2) - F_{\xi,\eta}(x_2,y_1) - F_{\xi,\eta}(x_1,y_2) + F_{\xi,\eta}(x_1,y_1)$

Независимость случайных величин

Def. Случайные величины $\xi_1,\dots,\xi_n$ независимы в совокупности, если для любого набора Борелевских множеств из $\mathcal{B}(\mathbb{R}^n)$, $B_1, B_2, \dots, B_n$
$p(\xi_1 \in B_1, \xi_2 \in B_2, \dots, \xi_n \in B_n) = p(\xi_1 \in B_1) \cdot p(\xi_2 \in B_2) \cdot \dots \cdot p(\xi_n \in B_n)$

Def. Случайные величины $\xi_1, \xi_2, \dots, \xi_n$ попарно независимы, если независимы любые две из них
Nota. Из независимости в совокупности следует попарная независимость:
$\xi_1,\dots,\xi_n$ независимы в совокупности, тогда покажем $\forall i,j$ ξ_i и ξ_j - независимы
Возьмем набор $B_i, B_j \in \mathcal{B}(\mathbb{R}^n)$, при $k \ne i,j$ $B_k = \mathbb{R}$ $\quad P(\xi_k \in B_k) = 1$
Тогда $p(\xi_1 \in B_1, \dots, \xi_n \in B_n) = P(\xi_i \in B_i, \xi_j \in B_j) = P(\xi_i \in B_i) \cdot P(\xi_j \in B_j)$

Nota. Из попарной независимости не следует независимость в совокупности, как видно из примера Бернштейна
Под независимыми величинами будем понимать независимые в совокупности

Дискретная система двух случайных величин

Def. Случайные величины ξ, η имеют совместное дискретное распределение, если случайный вектор (ξ, η) принимает не более, чем счетное число значений, то есть существует конечный или счетный набор пар чисел (x_i, y_i), таких что $P(\xi = x_i, \eta = y_i) > 0, \sum\limits_{i,j} P(\xi = x_i, \eta = y_i) = 1$

Таким образом двумерная дискретная случайная величина задается законом распределения - таблице вероятностей

$\xi \backslash \eta$	y_1	y_2	$\ldots$	y_m
x_1	p_{11}	p_{12}	$\ldots$	p_{1m}
x_2	p_{21}	p_{22}	$\ldots$	p_{2m}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
x_n	p_{n1}	p_{n2}	$\ldots$	p_{nm}

Условие нормировки: $\sum\limits_{i,j} p_{i,j} = 1$

Зная общий закон распределения, можно восстановить частное (маргинальное) распределение по формулам:

$$p_i = \sum_{j=1}^{m} p_{i,j} \qquad q_j = \sum_{i=1}^{n} p_{i,j}$$

Def. Дискретные случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ независимы, если для любых $x_1, x_2, \ldots, x_n$
$p(\xi_1 = x_1, \xi_2 = x_2, \ldots, \xi_n = x_n) = p(\xi_1 = x_1) \cdot p(\xi_2 = x_2) \cdot \dots \cdot p(\xi_n = x_n)$
При $n = 2$: $p_{i,j} = p_i \cdot q_j \ \forall i, j$

Ex.

$\xi \backslash \eta$	-1	0	1	p_i
-1	0.1	0.2	0.1	0.4
2	0.2	0.3	0.1	0.6
q_j	0.3	0.5	0.2	$\Sigma = 1$

Найти маргинальное распределение и проверить независимость случайных величин

ξ	-1	2
p_i	0.4	0.6

η	-1	0	1
q_j	0.3	0.5	0.2

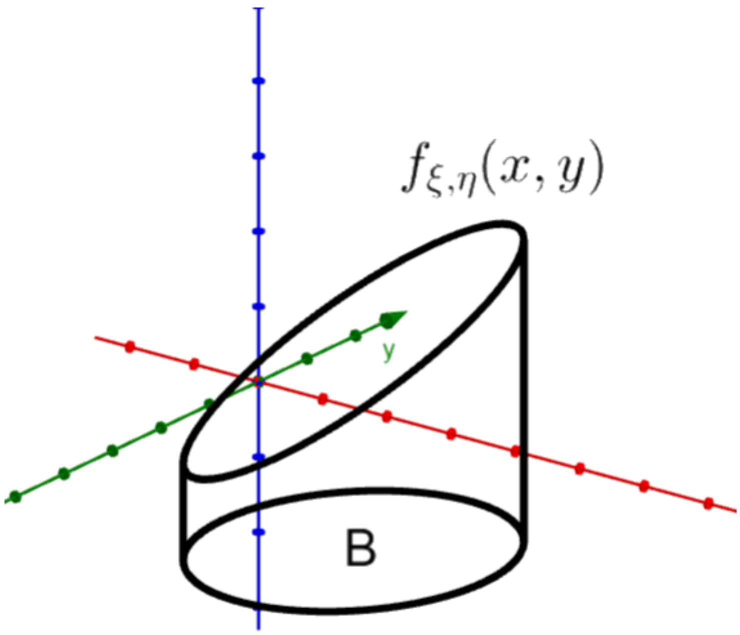
$p_{11} = 0.1 \neq 0.12 = p_1 \cdot q_1 \qquad \Longrightarrow \xi, \eta$ - зависимы

Абсолютно непрерывная система двух случайных величин

Def. Случайные величины ξ и η имеют абсолютно непрерывное совместное распределение, если $\exists f_{\xi,\eta}(x,y)$, такая что $\forall B \in \mathcal{B}(\mathbb{R}^2)$ $P((\xi,\eta)\in B) = \iint_B f_{\xi,\eta}(x,y)dxdy$

Функцию $f_{\xi,\eta}(x,y)$ будем называть функцией плотности совместного распределения случайных величин ξ и η

Геометрический смысл плотности:



Свойства плотности:

1. $f_{\xi,\eta}(x,y) \geq 0$
2. Условие нормировки: $\iint_{\mathbb{R}^2} f_{\xi,\eta}(x,y)dxdy = 1$
3. $F_{\xi,\eta} = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{\xi,\eta}(x,y)dydx$
4. $f_{\xi,\eta}(x,y) = \dfrac{\partial^2 F_{\xi,\eta}(x,y)}{\partial x \partial y}$
5. Если случайные величины ξ,η имеют абсолютно непрерывное совместное распределение с плотностью $f(x,y)$, то маргинальное распределение величин ξ,η также имеют абсолютно непрерывное распределение с плотностями $f_\xi(x) = \int_{-\infty}^{\infty} f_{\xi,\eta}(x,y)dy, f_\eta(y) = \int_{-\infty}^{\infty} f_{\xi,\eta}(x,y)dx$

$$F_\xi(x) = \lim_{y\to\infty} F_{\xi,\eta}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{\infty} f(x,y)dydx$$

Из этого $\int_{-\infty}^{\infty} f(x,y) = f_\xi(x)$

6. Так как вероятность попадания в Борелевские множества полностью задается функцией распределения, то условие независимости случайных величин эквивалентно следующему: $\xi_1, \xi_2, \dots, \xi_n$ независимы, если функция общего распределения распадается в произведение отдельных функцию распределения
$F_{\xi_1,\xi_2,\dots,\xi_n}(x_1, x_2, \dots, x_n) = F_{\xi_1}(x_1) \cdot F_{\xi_2}(x_2) \cdot \dots \cdot F_{\xi_n}(x_n)$
7. *Равносильное определение*: абсолютно непрерывные случайные величины $\xi_1, \dots, \xi_n$ независимы в совокупности тогда и только тогда, когда плотность совместного распределения $f_{\xi_1,\xi_2,\dots,\xi_n}(x_1, x_2, \dots, x_n) = f_{\xi_1}(x_1) \cdot f_{\xi_2}(x_2) \cdot \dots \cdot f_{\xi_n}(x_n)$

При $n = 2$ случайные величины ξ и η независимы $\Longleftrightarrow F_{\xi,\eta}(x,y) = F_\xi(x) \cdot F_\eta(y) = \int_{-\infty}^{x} f_\xi(x)dx \cdot \int_{-\infty}^{y} f_\eta(y)dy = \int_{-\infty}^{x}\int_{-\infty}^{y} f_\xi(x) \cdot f_\eta(y)dxdy \Longrightarrow f_{\xi,\eta}(x,y) = f_\xi(x)f_\eta(y)$
Аналогично для высших размерностей

Nota. Совместное распределение абсолютно непрерывных случайных величин не обязано быть абсолютно непрерывным, оно может быть сингулярным
Ex. Бросаем точку на отрезок прямой $y = x$ $(0 \le x, y \le 1)$, ξ - абсцисса точки, η - ордината точки
Случайный вектор (ξ, η) имеют сингулярное распределение (непрерывное с нулевой областью) - так как число элементарных исходов несчетно, но мера Лебега в $\mathbb{R}^2$ отрезка равна 0

Nota. Совместное распределение $\xi_1, \xi_2, \dots, \xi_n$ будет сингулярным, если одна из координат является функцией других (наблюдается функциональная зависимость)

Многомерное равномерное распределение

Def. $\exists D \subset \mathbb{R}^n$ - Борелевское множество в $\mathbb{R}^n$ с конечной мерой Лебега $(0 < \lambda(D) < \infty)$, случайный вектор $(\xi_1, \dots, \xi_n)$ имеет равномерное распределение, если плотность совместного распределения постоянна в данной области и равна нулю вне данной области

$$f_{\xi_1,\dots,\xi_n}(x_1, \dots, _n) = \begin{cases} \frac{1}{\lambda(D)}, & \text{если } (x_1, \dots, x_n) \in D \\ 0, & \text{если } (x_1, \dots, x_n) \notin D \end{cases}$$

Лекция 13

Математическое ожидание и дисперсия случайного вектора

$\exists \vec{\xi} = (\xi_1, \ldots, \xi_n)$ - случайный вектор, $\forall 1 \leq i \leq n$ ξ_i - случайная величина

Def. Математическим ожиданием случайного вектора называется вектор с координатами из математических ожиданий его компонент: $E\vec{\xi} = (E\xi_1, \ldots, E\xi_n)$

Def. Дисперсией (или матрицей ковариаций) случайного вектора называется матрица $D\vec{\xi} = E(\vec{\xi} - E\vec{\xi})^T \cdot (\vec{\xi} - E\vec{\xi})$, состоящая из элементов $d_{i,j} = (\xi_i, \xi_j)$. В частности $d_{i,i} = (\xi_i, \xi_i) = D\xi_i$

Функции от двух случайных величин

Th. Пусть ξ_1, ξ_2 - случайные величины с общем плотностью $f_{\xi_1,\xi_2}(x, y)$, и есть функция $g(x, y) : \mathbb{R}^2 \to \mathbb{R}$. Тогда случайная величина $\eta = g(\xi_1, \xi_2)$ имеет функцию распределения $F_\eta(z) = \iint_{D_z} f(x, y)dxdy$, где $D_z = \{(x, y) \in \mathbb{R}^2 \mid g(x, y) < z\}$

$$F_\eta = p(\eta < z) = p(g(\xi_1, \xi_2) < z) = p((\xi_1, \xi_2) \in D_z) = \iint_{D_z} f(x, y)dxdy$$

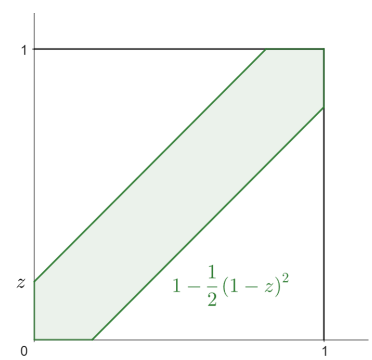
Ex. Задача о встрече. двое договорились встретится между 12:00 и 13:00. Случайная величина η - время ожидания. Найти функцию распределения

ξ_1 - время прихода первого, ξ_2 - второго; $\xi_1, \xi_2 \in U(0, 1)$, они независимы, $\forall x, y \in [0, 1]$ $f_{\xi_1}(x) = 1, f_{\xi_2}(y) = 1$

Поэтому $f_{\xi_1,\xi_2}(x, y) = f_{\xi_1}(x)f_{\xi_2}(y) = 1, (x, y) \in [0, 1] \times [0, 1]$

$\eta = |\xi_1 - \xi_2| \Longrightarrow D_z = \{(x, y) \in \mathbb{R}^2 \mid |x - y| < z\}$

$F_\eta = \iint_{D_z} f_{\xi_1,\xi_2}(x, y)dxdy = \iint_{D_z} dxdy = 1 - 2 \cdot \frac{1}{2}(1 - z)^2 = 2z - z^2,\ z \in [0, 1]$



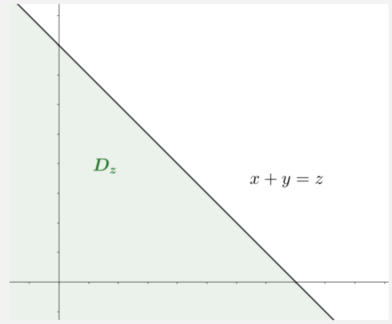
Th. $\exists \xi_1, \xi_2$ - независимые абсолютно непрерывные случайные величины с плотностями $f_{\xi_1}(x)$ и $f_{\xi_2}(y)$

Тогда плотность суммы $\xi_1 + \xi_2$ равна $f_{\xi_1+\xi_2}(t) = \int_{-\infty}^{\infty} \underbrace{f_{\xi_1}(x)f_{\xi_2}(t - x)dx}_{\text{т. н. свертка}}$

Так как случайные величины ξ_1 и ξ_2 независимы, то $f_{\xi_1,\xi_2}(x,y) = f_{\xi_1}(x)f_{\xi_2}(y)$

И согласно предыдущей теореме $F_{\xi_1+\xi_2}(z) = \iint_{D_z} f_{\xi_1,\xi_2}(x,y)dxdy = \iint_{D_z} f_{\xi_1}(x)f_{\xi_2}(y)dxdy$,

где $D_z = \{(x,y) \in \mathbb{R}^2 \mid x+y<z\}$



$$F_{\xi_1+\xi_2}(z) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{z-x} f_{\xi_1}(x)f_{\xi_2}(y)dy = \begin{bmatrix} y = t-x; & dy = dt; & t = y+x \\ t(-\infty) = -\infty; & t(z-x) = z & \end{bmatrix} =$$

$$\int_{-\infty}^{\infty} f_{\xi_1}(x)dx \int_{-\infty}^{z} f_{\xi_2}(t-x)dt =$$

$$\int_{-\infty}^{z}\left(\int_{-\infty}^{\infty} f_{\xi_1}(x)f_{\xi_2}(t-x)dx\right)dt \Longrightarrow f_{\xi_1+\xi_2}(t) = \int_{-\infty}^{\infty} f_{\xi_1}(x)f_{\xi_2}(t-x)dx$$

Следствие: сумма двух независимых абсолютно непрерывных случайных величин также имеет абсолютно непрерывное распределение

Nota. Условие независимости существенно, контр-пример: $\xi_1; \xi_2 = -\xi_1$, тогда $\xi_1 + \xi_2 \equiv 0$

Сумма стандартных распределений. Устойчивость относительно суммирования

Def. Если сумма двух независимых случайных величин одного типа распределения также будет этого же типа, то говорят, что распределение устойчиво относительно суммирования

Ex. 1. $\xi \in B_{n,p}; \eta \in B_{m,p}$. Тогда ясно, что $\xi+\eta \in B_{n+m,p}$ (по определению биномиального распределения $B_{n,p}$ - число успехов из n испытаний, где p - вероятность успеха)

Ex. 2. $\xi \in \Pi_\lambda, \eta \in \Pi_\mu$, они независимы. Тогда $\xi+\eta \in \Pi_{\lambda+\mu}$

$\xi+\eta = 0,1,2,3,\ldots \quad \exists k \geq 0$. Тогда $p(\xi+\eta=k) = \sum_{i=0}^{k} P(\xi=i, \eta=k-i) = \sum_{i=0}^{k} P(\xi=i)P(\eta=k-i) = \sum_{i=0}^{k} \frac{\lambda^i}{i!}e^{-\lambda}\frac{\mu^{k-i}}{(k-i)!}e^{-\mu} = e^{-\lambda-\mu}\sum_{i=0}^{k}\frac{\lambda^i\mu^{k-i}}{i!(k-i)!} = e^{-\lambda-\mu}\frac{1}{k!}\sum_{i=0}^{k}\frac{\lambda^i\mu^{k-i}k!}{i!(k-i)!} = e^{-\lambda-\mu}\frac{1}{k!}\sum_{i=0}^{k}\lambda^i\mu^{k-i}C_k^i = e^{-\lambda-\mu}\frac{(\lambda+\mu)^k}{k!} \Longrightarrow \xi+\eta \in \Pi_{\lambda+\mu}$

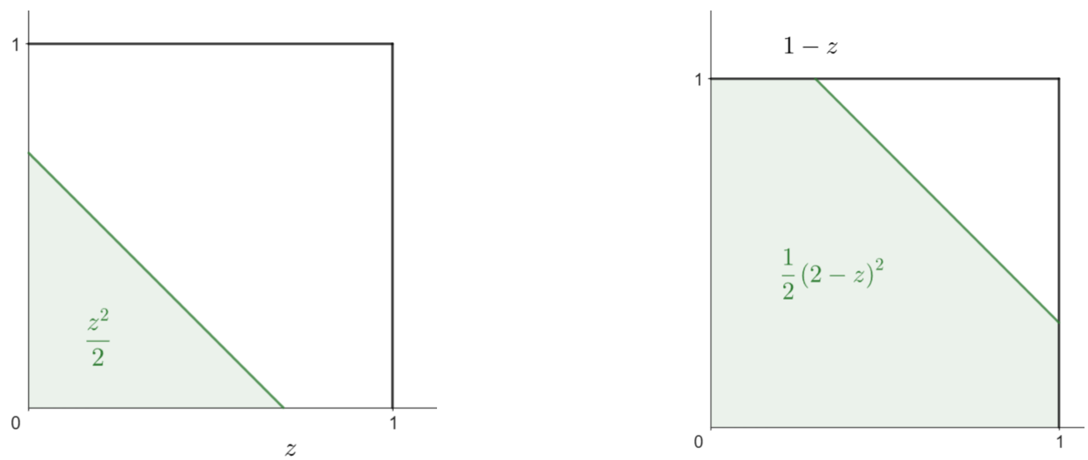
Ex. 3. $\xi, \eta \in N(0,1)$ и независимы. Тогда $\xi+\eta \in N(0,2)$

$f_\xi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}; f_\eta(y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}$

По формуле свертки $f_{\xi+\eta}(t) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}}e^{-\frac{(t-x)^2}{2}} = \frac{1}{2\pi}\int e^{-(x^2-tx+\frac{t^2}{2})} = \frac{1}{2\pi}e^{-\frac{t^2}{4}}\int_{-\infty}^{\infty} e^{-(x^2-tx+\frac{t^2}{4})}dx = \frac{1}{2\pi}e^{-\frac{t^2}{4}}\int_{-\infty}^{\infty} e^{-(x-\frac{t}{2})^2}d(x-\frac{t}{2}) = \frac{1}{2\pi}e^{-\frac{t^2}{4}}\sqrt{\pi} = \frac{1}{\sqrt{2}\sqrt{2\pi}}e^{-\frac{t^2}{2(\sqrt{2})^2}} \Longrightarrow$
$\xi+\eta \in N(0,2)$

Ex. 4. В общности для независимых $\xi \in N(a_1, \sigma_1^2), \eta \in N(a_2, \sigma_2^2)$ $\xi + \eta \in N(a_1 + a_2, \sigma_1^2 + \sigma_2^2)$

Ex. 5. Равномерное распределение неустойчиво относительно суммирования, контрпример:
$\xi, \eta \in U(0,1)$ - независимы
$\forall x, y \in [0,1]$ $f_\xi(x) = 1, f_\eta(y) = 1$ и $f_{\xi,\eta}(x,y) = 1$
По первой теореме $F_{\xi,\eta}(x,y) = \iint_{D_z} f_{\xi,\eta}(x,y)dxdy = \iint_{D_z} dxdy = S_{D_z}$, где $D_z = \{(x,y) \mid x+y<z\}$



а) $0 < z \le 1$ б) $1 < z \le 2$

$$S_{D_z} = \begin{cases} 0, & z<0 \\ \frac{z^2}{2}, & 0 \le z \le 2 \\ 1 - \frac{1}{2}(2-z)^2 & 1 \le z \le 2 \\ 0, & z > 2 \end{cases}$$

$$f_{\xi+\eta}(z) = \begin{cases} 0, & z<0 \\ z, & 0 \le z \le 2 \\ 2-z & 1 \le z \le 2 \\ 0, & z > 2 \end{cases} \not\equiv C \Longrightarrow \xi+\eta \text{ не имеют равномерное распределение}$$

Nota. FUN FACT: сумма нескольких величин с равномерным распределением приближается к нормальному распределению

Условное распределение

Def. Условным распределением случайной величины из системы случайных величин (ξ, η) называется ее распределение, найденное при условии, что другая случайная величина приняла определенное значение. Обозначается $\xi|\eta = y$

Def. А.: Условным математическим ожиданием (обозначается $E(\xi|\eta = y)$) называется математическим ожиданием случайной величины ξ при соответствующем условном распределении

I. Условное распределение в дискретной системе двух случайных величин

Пусть (ξ, η) задана законом распределения:

$\xi \backslash \eta$	y_1	y_2	$\dots$	y_m
x_1	p_{11}	p_{12}	$\dots$	p_{1m}
x_2	p_{21}	p_{22}	$\dots$	p_{2m}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
x_n	p_{n1}	p_{n2}	$\dots$	p_{nm}

Формула условной вероятности: $P(A|B) = \dfrac{P(AB)}{P(B)}$

Вероятности условных распределений считаем по формулам:

$\xi|\eta = y_j$: $p_i = p(\xi = x_i \mid \eta = y_j) = \dfrac{p(\xi = x_i, \eta = y_j)}{p(\eta = y_j)} = \dfrac{p_{ij}}{q_j} = \dfrac{p_{ij}}{\sum_i p_{ij}}$

$\eta|\xi = x_i$: $q_j = p(\eta = y_j \mid \xi = x_i) = \dfrac{p(\xi = x_i, \eta = y_j)}{p(\xi = x_i)} = \dfrac{q_{ij}}{p_i} = \dfrac{q_{ij}}{\sum_j p_{ij}}$

То есть вероятность в соответствующем столбце/строке делим на суммарную вероятность по строке или столбцу, в зависимости от того, какое условие мы рассматриваем

II. Условное распределение в непрерывной системе двух случайных величин

Пусть (ξ, η) задана плотностью $f_{\xi,\eta}(x, y)$ совместного распределения, тогда плотность условного распределения $\xi|\eta = y$:

$$f(x|y) = \frac{f_{\xi,\eta}(x, y)}{\int_{\mathbb{R}} f_{\xi,\eta}(x, y)dx} = \frac{f_{\xi,\eta}(x, y)}{f_\eta(y)}$$

Def. Функция $f(x|y) = \dfrac{f_{\xi,\eta}(x, y)}{f_\eta(y)}$ называется условной плотностью

Def. Условное математические ожидание вычисляется по формуле $E(\xi|\eta = y) = \displaystyle\int_{-\infty}^{\infty} x f(x|y)dx$

Аналогично $E(\eta|\xi = x) = \int_{-\infty}^{\infty} yf(y|x)dy$

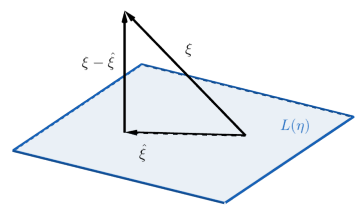
Nota. При фиксированном значении x $f(y|x)$ зависит только от y, а $E(\eta|\xi = x) \in \mathbb{R}$. Если рассматривать x как переменную, то условное математическое ожидание $E(\eta|\xi = x)$ является функцией от x и называется функцией регрессии η на ξ. График такой функции называют линией регрессии

Nota. Так как значение x - значение случайной величины ξ, то условное матожидание $E(\eta|\xi = x)$ можно рассматривать как случайную величину

Лекция 14

Пространство случайных величин

Nota. Если две случайных величин $\xi \overset{\text{п.н.}}{=} \eta$, то считаем, что $\xi = \eta$
Пусть имеется вероятностное пространство $(\Omega, \mathcal{F}, P)$
Введем пространство $L_2(\Omega, \mathcal{F}, P) = \{\xi \mid D\xi < \infty\}$ - множество случайных величин на данном пространстве с конечной дисперсией
Ясно, что L_2 - линейное пространство. Введем на нем скалярное произведение



Def. Скалярным произведением случайных величин ξ и η из $L_2(\Omega, \mathcal{F}, P)$ называется число $(\xi, \eta) = E(\xi\eta)$

Nota. Если (ξ, η) - дискретная система случайных величин $\left(p(\xi = x_i, \eta = y_i) = p_{ij}\right)$, то $E(\xi\eta) = \sum\limits_{i,j} x_i y_j p_{ij}$
Если же (ξ, η) - непрерывная система с плотностью $f_{\xi,\eta}(x, y)$, то $E(\xi\eta) = \iint_{\mathbb{R}^2} xyf_{\xi,\eta}(x, y)dxdy$

Свойства:

1. $(\xi, \eta) = (\eta, \xi)$
2. $(C\xi, \eta) = C(\xi, \eta)$
3. $(\xi_1 + \xi_2, \eta) = (\xi_1, \eta) + (\xi_2, \eta)$
4. $(\xi, \xi) \geq 0$
5. $(\xi, \xi) = 0 \Longrightarrow \xi = 0$ п.н.

То есть это действительно скалярное произведение

Def. Норма вектора равна числу $\|\xi\| = \sqrt{(\xi, \xi)}$

Def. Метрикой (расстоянием) между случайными величинами называют число $d(\xi,\eta) = \|\xi-\eta\|$

Th. Неравенство Коши-Буняковского-Шварца
Пусть случайные величины ξ и η имеют конечный второй момент, тогда $|E(\xi,\eta)| \le \sqrt{E\xi^2 \cdot E\eta^2}$ (или $|(\xi,\eta)| \le \|\xi\| \cdot \|\eta\|$)
Причем $|E(\xi,\eta)| = \sqrt{E\xi^2 \cdot E\eta^2} \Longleftrightarrow \eta = C\xi$, где $C = \mathrm{const}$

$P_2(x) = E(x\xi-\eta)^2 = x^2E\xi^2 - 2xE(\xi\eta) + E\eta^2 \ge 0 \Longrightarrow D = 4(E(\xi\eta))^2 - 4E\xi^2 - E\eta^2 \le 0 \Longrightarrow |E(\xi\eta)| \le \sqrt{E\xi^2 \cdot E\eta^2}$
$|E(\xi,\eta)| = \sqrt{E\xi^2 - E\eta^2} \Longrightarrow D = 0 \Longrightarrow \exists$ какая-либо точка касания C, из этого $E(C\xi-\eta)^2 = 0 \Longrightarrow C\xi - \eta = 0 \Longleftrightarrow \eta = C\xi$ п.н.

Условное математическое ожидание

В $L_2(\Omega, \mathcal{F}, P)$ возьмем линейное подпространство $L(\eta) = \{g(\eta) \mid Dg(\eta) < \infty\}$

Def. B. Условным математическим ожиданием (УМО, обозначается $E(\xi|\eta) = \hat{\xi}$) случайной величины ξ относительно случайной величины η называется ортогональная проекция случай ной величины ξ на $L(\eta)$

Свойства:

1. Тождество ортопроекций: $\exists\hat{\xi} \in L(\eta)$, тогда $\hat{\xi} = E(\xi|\eta) \Longleftrightarrow E(\xi \cdot g(\eta)) = E(\hat{\xi} \cdot g(\eta))\ \forall g(\eta) \in L(\eta)$

   $\hat{\xi} = E(\xi|\eta) \Longleftrightarrow (\xi - \hat{\xi}) \perp L(\eta) \Longleftrightarrow (\xi - \hat{\xi}, g(\eta)) = 0\ \forall g(n) \in L(\eta) \Longleftrightarrow E(\xi \cdot g(\eta)) = E(\hat{\xi} \cdot g(\eta))$

2. Формула полного математического ожидания
   $E\xi = E(E(\xi|\eta))$ или $E\xi = E\hat{\xi}$

   Nota. При распределении Бернулли получаем обычную формулу полной вероятности

   Верно из тождества ортопроекций при $g(\eta) = 1$

3. Линейность: $E(C_1\xi_1 + C_2\xi_2 \mid \eta) = C_1E(\xi_1|\eta) + C_2E(\xi_2|\eta)$
4. Если ξ и η независимы, то $E(\xi|\eta) = E\xi$

   ξ, η независимы $\Longrightarrow \xi$ и $g(\eta)$ независимы
   Из этого $E(\xi \cdot g(\eta)) = E\xi \cdot E(g(\eta)) = E(E\xi \cdot g(\eta)) \Longrightarrow E\xi = \hat{\xi}$

5. Если ξ и η независимы, то $(\xi - E\xi) \perp g(\eta)\ \forall g(\eta) \in L(\eta)$, в частности $(\xi - E\xi) \perp \eta$

Докажем, что **Def. A.** согласуется с **Def. B.**

По **Def. A.** $E(\xi|\eta) = h(\eta)$, где $h(y) = E(\xi|\eta = y)$
Рассмотрим случай абсолютно непрерывной системы (ξ, η) с плотностью $f_{\xi,\eta}(x, y)$. Тогда $h(y) = \int_{-\infty}^{\infty} x f(x|y)dx$, где $f(x|y) = \dfrac{f_{\xi,\eta}(x, y)}{f_\eta(y)}$
Следует доказать, что функция $h(y)$ удовлетворяет тождеству ортопроекций $E(\xi g(\eta)) = E(h(\eta)g(\eta))\ \forall g(\eta) \in L(\eta)$
$E(\xi \cdot g(\eta)) = \iint_{\mathbb{R}^2} xg(y)f_{\xi,\eta}(x, y)dxdy$
$E(h(\eta)g(\eta)) = \int_{-\infty}^{\infty} h(y)g(y)f_\eta(y)dy = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} x\dfrac{f_{\xi,\eta}(x, y)}{f_\eta(y)}dx\right) g(y)f_\eta(y) = dy = \iint_{\mathbb{R}^2} xg(y)f_{\xi,\eta}(x, y)dxdy = E(\xi g(\eta))$

Числовые характеристики. Зависимости случайных величин

Мет. Если случайные величины ξ и η, то $E(\xi\eta) = E\xi E\eta \Longrightarrow E(\xi\eta) - E\xi E\eta = 0$
Поэтому в качестве индикатора наличия связи берем величину $E(\xi\eta) - E\xi E\eta = \operatorname{cov}(\xi, \eta)$

Def. Ковариацией (ξ, η) называется величина $\operatorname{cov}(\xi, \eta) = E((\xi - E\xi)(\eta - E\eta))$

Свойства:

1. $\operatorname{cov}(\xi, \eta) = E(\xi\eta) - E\xi E\eta$

   $\operatorname{cov}(\xi, \eta) = E((\xi - E\xi)(\eta - E\eta)) = E(\xi\eta - \eta E\xi - \xi E\eta + E\xi E\eta) = E(\xi\eta) - E\xi E\eta$

2. $\operatorname{cov}(\xi, \xi) = D\xi$

   $\operatorname{cov}(\xi, \xi) = E\xi^2 - (E\xi)^2$

3. $\operatorname{cov}(\xi, \eta) = \operatorname{cov}(\eta, \xi)$
4. $\operatorname{cov}(C_1\xi_1 + C_2\xi_2, \eta) = C_1\operatorname{cov}(\xi_1, \eta) + C_2\operatorname{cov}(\xi_2, \eta)$
5. $D(\xi + \eta) = D\xi + D\eta + 2\operatorname{cov}(\xi, \eta)$
6. $D(\xi_1 + \cdots + \xi_n) = \sum_{i=1}^{n} D\xi_i + 2\sum_{i<j} \operatorname{cov}(\xi_i, \xi_j) = \sum_{i,j=1}^{n} \operatorname{cov}(\xi_i, \xi_j)$
7. (a) Если ξ и η - независимы, то $\operatorname{cov}(\xi, \eta) = 0$
   (b) Если $\operatorname{cov}(\xi, \eta) \neq 0$, то ξ и η - зависимы
   (c) Если $\operatorname{cov}(\xi, \eta) = 0$, то неясно
8. Если $\operatorname{cov}(\xi, \eta) > 0$, то зависимость прямая, если $\operatorname{cov}(\xi, \eta) < 0$, то обратная

Nota. Ковариация зависит от единиц измерения случайных величин, поэтому по ее величине нельзя судить о силе зависимости

Коэффициент линейной корреляции

Def. Коэффициентом корреляции случайных величин ξ и η с конечными вторыми моментами, называется величина $r_{\xi,\eta} = \dfrac{\operatorname{cov}(\xi,\eta)}{\sqrt{D\xi}\sqrt{D\eta}} = \dfrac{E(\xi\eta) - E\eta}{\sigma_\xi\sigma_\eta}$

Можно записать в другой форме: $r_{\xi,\eta} = \dfrac{E((\xi - E\xi)(\eta - E\eta))}{\sqrt{E(\xi - E\xi)^2}\sqrt{E(\eta - E\eta)^2}} = \dfrac{(\xi - E\xi, \eta - E\eta)}{\|\xi - E\xi\|\|\eta - E\eta\|} = \cos(\xi - \widehat{E\xi, \eta -} E\eta)$ - косинус угла между величинами (грубая интерпретация)

Свойства:

1. $r_{\xi,\eta} = r_{\eta,\xi}$
2. (a) Если ξ и η - независимы, то $r_{\xi,\eta} = 0$
   (b) Если $r_{\xi,\eta} \neq 0$, то ξ и η - зависимы
   (c) Если $r_{\xi,\eta} = 0$, то неясно
3. $|r_{\xi,\eta}| \leq 1$

   По неравенству Коши-Буняковского-Шварца $|E((\xi - E\xi)(\eta - E\eta))| \leq \sqrt{E(\xi - E\xi)^2 E(\eta - E\eta)^2}$

4. $|r_{\xi,\eta}| = 1 \iff \eta = a\xi + b$ п.н.

   По неравенству Коши-Буняковского-Шварца $|r_{\xi,\eta}| = 1 \iff |E((\xi - E\xi)(\eta - E\eta))| = \sqrt{E(\xi - E\xi)^2 E(\eta - E\eta)^2} \implies \eta - E\eta = C(\xi - E\xi) \implies \eta = C\xi + (E\eta - CE\xi)$ п.н.

5. (a) Если $r_{\xi,\eta} = 1$, то $\eta = a\xi + b$ и $a > 0$ (прямая линейная зависимость)
   (b) Если $r_{\xi,\eta} = -1$, то $\eta = a\xi + b$ и $a < 0$ (обратная линейная зависимость)

   Так как $|r_{\xi,\eta}| = 1$, то по свойству 4) $\eta = a\xi + b$ и $r_{\xi,\eta} = \dfrac{E(\xi\eta) - E\xi E\eta}{\sigma_\xi\sigma_\eta} = \dfrac{E(\xi(a\xi + b)) - E\xi E(a\xi + b)}{\sqrt{D\xi D(a\xi + b)}} = \dfrac{aE\xi^2 + bE\xi - a(E\xi)^2 - bE\xi}{\sqrt{D\xi a^2 D\xi}} = \dfrac{a(E\xi^2 - (E\xi)^2)}{|a|D\xi} = \dfrac{a}{|a|} = \operatorname{sign} a$

Def. Если $r_{\xi,\eta} \neq 0$, то говорят, что случайные величины коррелированы друг с другом. Если $r_{\xi,\eta} > 0$, то имеет прямая корреляция, если $r_{\xi,\eta} < 0$ - обратная

Nota. Корреляция не транзитивна: $r_{\xi_1,\xi_2} > 0 \wedge r_{\xi_2,\xi_3} > 0 \not\Rightarrow r_{\xi_1,\xi_3} > 0$

Лекция 15

Характеристические функции

Мет. i - комплексная единица
Мет. $e^{it} = \cos t + i \sin t$
Пусть $\xi + i\eta$ - комплексная случайная величина, где ξ - вещественная часть, а η - мнимая часть

Def. $E(\xi + i\eta) = E\xi + iE\eta$

Def. Характеристической функций случайной величины ξ называется функция

$$\varphi_\xi(t) = Ee^{it\xi}, t \in \mathbb{R}$$

Свойства:

1. Любая случайная величина ξ имеет характеристическую функцию, причем $|\varphi_\xi(t)| \le 1$

   Характеристическая функция существует по теореме об абсолютной сходимости интеграла от произведения ограниченной и нормированной функций
   Докажем неравенство:
   $|\varphi_\xi(t)|^2 = |Ee^{it\xi}|^2 = |E\cos t\xi + iE\sin t\xi|^2 = (E\cos\xi t)^2 + (E\sin\xi t)^2 \le$ [по неравенству Йенсена] $\le E\cos^2\xi t + E\sin^2\xi t = E(\cos^2\xi t + \sin^2\xi t) = E1 = 1$

2. Пусть $\varphi_\xi(t)$ - характеристическая функция случайной величины ξ. Тогда характеристическая функция случайной величины $a + b\xi$ равна $\varphi_{a+b\xi}(t) = e^{ita}\varphi_\xi(bt)$

   $\varphi_{a+b\xi}(t) = Ee^{it(a+b\xi)} = E(e^{ita} \cdot e^{itb\xi}) = e^{ita}Ee^{itb\xi} = e^{ita}\varphi_\xi(bt)$

3. Характеристическая функция суммы независимых случайных величин равна произведению их характеристических функций

   Пусть случайные величины ξ и η - независимы. Тогда
   $\varphi_{\xi+\eta}(t) = E(e^{it\xi} \cdot e^{it\eta}) =$ [так как они независимы] $= Ee^{it\xi} \cdot Ee^{it\eta} = \varphi_\xi(t) \cdot \varphi_\eta(t)$
   Аналогично для большего числа величин

4. Пусть $E\xi^k < \infty$. Тогда

$$\varphi_\xi(t) = 1 + itE\xi - \frac{t^2}{2}E\xi^2 + \dots + \frac{(it)^k}{k!}E\xi^k + o(|t|^k)$$

$\varphi_\xi(t) = Ee^{it\xi} = E(1 + it\xi + \frac{(it\xi)^2}{2!} + \dots + \frac{(it\xi)^k}{k!} + o(|t|^k)) = 1 + itE\xi\frac{i^2t^2}{2}E\xi^2 + \dots + \frac{(it)^k}{k!}E\xi^k + o(|t|^k)$

5. Пусть $E\xi^k < \infty$. Тогда $\varphi_\xi^{(k)}(0) = i^k E\xi^k$

$E\xi^k < \infty \Longrightarrow$ существует k членов разложения в ряд Маклорена: $\frac{\varphi_\xi^{(k)}(0)}{k!}t^k = \frac{i^k E\xi^k}{k!}t^k$; $\frac{\varphi_\xi^{(k)}(0)}{k!}t^k = i^k E\xi^k$

6. Существует взаимно-однозначное соответствие между распределениями и характеристическими функциями. Зная характеристическую функцию можно восстановить распределение.

   Ex. Если распределение абсолютно непрерывное, то его можно восстановить по преобразованию Фурье
   $f_\xi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-itx}\varphi_\xi(t)dt$
7. Теорема о непрерывном соответствии

Th. Последовательность случайных величин $\{\xi_n\}$ слабо сходится к ξ тогда и только тогда, когда соответствующая последовательность характеристических функций сходится поточечно к $\varphi_\xi(t)$
$\{\xi_n\} \rightrightarrows \xi \Longleftrightarrow \varphi_{\xi_n}(t) \longrightarrow \varphi_\xi(t) \forall t \in \mathbb{R}$

Характеристические функции стандартных распределений

- Распределение Бернулли

ξ	0	1
p	$1-p$	p

  $\varphi_\xi(t) = Ee^{i\xi t} = e^{i\cdot 0\cdot t}p(\xi = 0) + e^{i\cdot 1\cdot t}p(\xi = 1) = 1 - p + pe^{it}$
- Биномиальное распределение
  $P(\xi = k) = C_n^k p^k q^{n-k}, \quad k = 0, 1, \dots, n$
  Если $t \in B_{n,p}$, то $\xi = \xi_1 + \xi_2 + \xi_3 + \dots + \xi_n$, где $\xi_i \in B_p$ - независимы

$\varphi_\xi(t) = (\varphi_{\xi_n}(t))^n = (1 - p + pe^{it})^n$

- Распределение Пуассона

  $P(\xi = k) = \dfrac{\lambda^k}{k!}e^{-\lambda}, \quad k = 0, 1, \dots, n$

  $\varphi_\xi(t) = Ee^{it\xi} = \sum\limits_{k=0}^{\infty} e^{itk}p(\xi = k) = \sum\limits_{k=0}^{\infty} e^{itk}\dfrac{\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}\sum\limits_{k=0}^{\infty}\dfrac{(\lambda e^{it})^k}{k!} = e^{-\lambda}e^{\lambda e^{it}} = e^{\lambda(e^{it}-1)}$

  Следствие: распределение Пуассона устойчиво относительно суммирования

  > $\exists \xi \in \Pi_\lambda, \eta \in \Pi_\mu$, они независимы. Тогда $\xi + \eta \in \Pi_{\lambda+\mu}$

  > По третьему свойству $\varphi_{\xi+\eta}(t) = \varphi_\xi(t) \cdot \varphi_\eta(t) = e^{\lambda(e^{it}-1)}e^{\mu(e^{it}-1)} = e^{(\lambda+\mu)(e^{it}-1)}$ - характеристическая функция распределения Пуассона $\Pi_{\lambda+\mu}$

- Стандартное нормальное распределение

  $f_\xi(x) = \dfrac{1}{2\pi}e^{-\frac{x^2}{2}}$

  $\varphi_\xi(t) = Ee^{it\xi} = \int_{-\infty}^{\infty} e^{itx}f_\xi(x)dx = \int_{-\infty}^{\infty} e^{itx}\dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx = \dfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(x^2-2itx)}dx =$
  $\dfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(x^2-2itx-t^2)}e^{-\frac{t^2}{2}}dx = \dfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(x-it)^2}{2}}d(x-it) = \dfrac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\sqrt{2\pi} = e^{-\frac{t^2}{2}}$

- Нормальное распределение

  $\xi \in N(a, \sigma^2)$

  Если $\eta \in N(0, 1)$, то $\xi = a + \sigma\eta \in N(a, \sigma^2)$

  По второму свойству $\varphi_\xi(t) = e^{ita}\varphi_\eta(\sigma t) = e^{ita - \frac{\sigma^2 t^2}{2}}$

  Следствие: нормальное распределение устойчиво относительно суммирования

  > Если $\xi \in N(a_1, \sigma_1^2), \eta \in N(a_2, \sigma_2^2)$ и они независимы, то $\xi + \eta \in N(a_1 + a_2, \sigma_1^2 + \sigma_2^2)$

  > $\varphi_{\xi+\eta}(t) = \varphi_\xi(t)\varphi_\eta(t) = e^{ita_1 - \frac{\sigma_1^2 t^2}{2}}e^{ita_2 - \frac{\sigma_2^2 t^2}{2}} = e^{it(a_1+a_2) - \frac{(\sigma_1^2+\sigma_2^2)t^2}{2}}$ - характеристическая функция $N(a_1 + a_2, \sigma_1^2 + \sigma_2^2)$

Доказательства теорем через свойства характеристических функций

Докажем некоторые теоремы с помощью характеристических функций

Закон больших чисел Хинчина

Для доказательства закона больших чисел Хинчина докажем такую лемму:

$\left(1+\frac{x}{n}+o\left(\frac{1}{n}\right)\right)^n \xrightarrow[n\to\infty]{} e^x$

$\left(1+\frac{x}{n}+o\left(\frac{1}{n}\right)\right)^n = e^{n\ln\left(1+\frac{x}{n}+o\left(\frac{1}{n}\right)\right)} = e^{n\left(\frac{x}{n}+o\left(\frac{1}{n}\right)+o\left(\frac{x}{n}+o\left(\frac{1}{n}\right)\right)\right)} = e^{n\left(\frac{x}{n}+o\left(\frac{1}{n}\right)+o\left(\frac{1}{n}\right)\right)} = e^{x+no\left(\frac{1}{n}\right)} \xrightarrow[n\to\infty]{} e^x$

Th. Закон больших чисел Хинчина

Пусть $\xi_1, \xi_2, \dots, \xi_n$ - последовательность независимых одинаково распределенных случайных величин с конечным матожиданием. Тогда $\frac{S_n}{n} = \frac{\xi_1+\dots+\xi_n}{n} \xrightarrow{p} E\xi_1$

Обозначим $a = E\xi_1$

Ранее было доказано, что сходимость по вероятности к константе эквивалентно к слабой сходимости. Поэтому достаточно доказать, что $\frac{S_n}{n} \rightrightarrows a$

По теореме о непрерывном соответствии остается доказать, что $\varphi_{\frac{S_n}{n}}(t) \longrightarrow \varphi_a(t) = e^{ita}$

По четвертому свойству $\varphi_{\xi_1}(t) = 1 + itE\xi_1 + o(|t|) = 1 + ita + o(|t|)$

$\varphi_{\frac{S_n}{n}}(t)$ = [по второму свойству] = $\varphi_{S_n}\left(\frac{t}{n}\right) = \left(\varphi_{\xi_1}\left(\frac{t}{n}\right)\right)^n = \left(1+ia\frac{t}{n}+o\left(\left|\frac{t}{n}\right|\right)\right)^n \xrightarrow[\text{по лемме}]{} e^{ita} = \varphi_a(t)$

Центральная предельная теорема

Th. Центральная предельная теорема Ляпунова, 1901 г.

Пусть $\xi_1, \xi_2, \dots, \xi_n$ - последовательность независимых одинаково распределенных случайных величин с конечным вторым моментом ($D\xi_1 < \infty$)

Обозначим $a = E\xi_1, \sigma^2 = D\xi_1$. Тогда

$$\frac{S_n - na}{\sigma\sqrt{n}} \rightrightarrows N(0,1)$$

Пусть $\eta_i = \frac{\xi_i - a}{\sigma}$ - стандартизованная случайная величина

$E\eta_i = 0, D\eta_i = 1$

Обозначим $Z_n = \eta_1 + \dots + \eta_n = \frac{(\xi_1+\dots+\xi_n) - na}{\sigma} = \frac{S_n - na}{\sigma}$

Надо доказать, что если $\frac{Z_n}{\sqrt{n}} \rightrightarrows N(0,1)$

По четвертому свойству $\varphi_{\eta_1}(t) = 1 + itE\eta_1 - \frac{t^2}{2}E\eta_1^2 + o(t^2) = 1 - \frac{t^2}{2} + o(t^2)$

$\varphi_{\frac{Z_n}{\sqrt{n}}} = \varphi_{Z_n}\left(\frac{t}{\sqrt{n}}\right) = \left(\varphi_{\eta_1}\left(\frac{t}{\sqrt{n}}\right)\right)^n = \left(1 - \frac{\left(\frac{t}{\sqrt{n}}\right)^2}{2} + o\left(\left(\frac{t}{\sqrt{n}}\right)^2\right)\right)^n = \left(1 - \frac{t^2}{2n} + o\left(\left(\frac{t}{\sqrt{n}}\right)^2\right)\right)^n \xrightarrow[n\to\infty]{} e^{-\frac{t^2}{2}}$ - характеристическая функция $N(0,1)$

Предельная теорема Муавра-Лапласа

Th. Пусть $v_n(A)$ - число появления события A при n независимых испытаний, p - вероятность успеха при одном испытании, $q = 1 - p$. Тогда $\frac{v_n(A) - np}{\sqrt{npq}} \rightrightarrows N(0,1)$

$v_n(A) = \xi_1 + \xi_2 + \cdots + \xi_n = S_n$, где $\xi_i \in B_p$ и независимы, $E\xi_1 = p, D\xi_1 = pq$
По ЦПТ $\frac{v_n(A) - np}{\sqrt{npq}} = \frac{S_n - nE\xi_1}{\sqrt{nD\xi_1}} \rightrightarrows N(0,1)$

Следствие. Интегральная формула Лапласа:

$p(k_1 \le v_n \le k_2) = p\left(\frac{k_1 - np}{\sqrt{npq}} \le \frac{v_n - np}{\sqrt{npq}} \le \frac{k_2 - np}{\sqrt{npq}}\right)$. Обозначим $\eta = \frac{v_n - np}{\sqrt{npq}}$

$p\left(\frac{k_1 - np}{\sqrt{npq}} \le \frac{v_n - np}{\sqrt{npq}} \le \frac{k_2 - np}{\sqrt{npq}}\right) = F_\eta\left(\frac{k_2 - np}{\sqrt{npq}}\right) - F_\eta\left(\frac{k_1 - np}{\sqrt{npq}}\right) \xrightarrow[n\to\infty]{} F_0\left(\frac{k_2 - np}{\sqrt{npq}}\right) - F_0\left(\frac{k_1 - np}{\sqrt{npq}}\right)$, где $F_0(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$

Nota. Аналогичным образом ЦПТ применяется для приближенного вычисления вероятностей, связанных с суммами большого числа независимых одинаковых случайных величин, заменяя стандартизованную сумму на стандартное нормальное распределение. Возникает вопрос: какова погрешность данного вычисления?

Th. Неравенство Берри-Эссеена
В условиях ЦПТ для ξ_1 с конечным третьим моментом можно оценить так:
$\left| p\left(\frac{S_n - nE\xi_1}{\sqrt{nD\xi_1}} < x\right) - F_0(x) \right| \le C \frac{E|\xi_1 - E\xi_1|^3}{\sqrt{n(D\xi_1)^3}} \quad \forall x \in \mathbb{R}$

Nota. На практике берут $C = 0.4$, точная оценка сверху $C < 0.77$

Лекция 16

Условная дисперсия

Def. Условной дисперсией случайной величины ξ относительно случайной величины η называется случайная величина $D(\xi|\eta) = E((\xi - E(\xi|\eta))^2|\eta)$

Nota. То есть дисперсия соответствующего условного распределения

Свойства

1. $D(\xi|\eta) = E(\xi^2|\eta) - E^2(\xi|\eta)$
2. Закон полной дисперсии

   Th. $D\xi = E(D(\xi|\eta)) + D(E(\xi|\eta))$

   Из первого свойства $E(\xi^2|\eta) = D(\xi|\eta) + E^2(\xi|\eta)$
   $D\xi = E\xi^2 - (E\xi)^2 = E(E\xi^2|\eta) - E^2(E(\xi|\eta)) = E(D(\xi|\eta) + E^2(\xi|\eta)) - E^2(E(|\eta)) = E(D(\xi|\eta)) + E(E^2(\xi|\eta)) - E^2(E(\xi|\eta)) = E(D(\xi|\eta)) + D(E(\xi|\eta))$

   Следствие и смысл:

   - Если ξ и η независимы (некоррелированы), то $D(E(\xi|\eta)) = D(E\xi) = 0$ и $D\xi = E(D(\xi|\eta))$
   - Если имеется функциональная зависимость (то есть $\xi = g(\eta)$), то $D(E(\xi|\eta)) = D(E(g(\eta)|\eta)) = D(g(\eta)) = D\xi$
   - Таким образом по величине $R^2 = \dfrac{D(E(\xi|\eta))}{D\xi}$ $(0 \le R^2 \le 1)$ можно судить о силе корреляционной зависимости. Такая величина называется корреляционным отношением

Энтропия

Пусть ξ - результат эксперимента с исходами $A_1, A_2, \dots, A_N$, вероятности которых $p_1, p_2, \dots, p_N$

Def. Энтропией эксперимента называется величина $H(\xi) = -\sum\limits_{i=1}^{N} p_i \cdot \log_2 p_i$

Свойства энтропии:

1. Очевидно, что $H(\xi) \ge 0$, так как $p \ge 0$, а $\log_2 p_i \le 0$
2. $H(\xi) = 0 \iff \exists i$, такой что $p_i = 1, p_j = 0 \forall j \ne i$ - то есть эксперимент заканчивается всегда одним исходом, нет неопределенности
3. Максимум $H(\xi) = \log_2 N = H_0$ достигается при $p_1 = p_2 = \dots = \dfrac{1}{N}$ - то есть когда все вероятности одинаковы, ни одному исходу нельзя отдать предпочтение, и результат эксперимента получается максимально неопределенным

Рассмотрим $\varphi(x) = x\log_2 x$. Так как $\varphi''(x) = \dfrac{1}{x\ln 2} > 0$ при $x > 0$, следовательно $\varphi(x)$ выпукла вниз

Рассмотрим случайную величину η

η	p_1	p_2	$\dots$	p_n
p	$\frac{1}{N}$	$\frac{1}{N}$	$\dots$	$\frac{1}{N}$

По неравенству Йенсена $\varphi(E\eta) = \varphi(\sum_{i=1}^{N}\frac{p_i}{N}) = \varphi(\frac{1}{N}\sum_{i=1}^{N}p_i) = \varphi(\frac{1}{N}) = \dfrac{\log_2\frac{1}{N}}{N} \le E(\varphi(\eta)) =$

$\dfrac{1}{N}\sum_{i=1}^{N}p_i\log_2 p_i = -\dfrac{1}{N}H(\eta)$

Получаем $\dfrac{\log_2\frac{1}{N}}{N} \le -\dfrac{1}{N}H(\eta)$, то есть $H \le \log_2 N$

Следствие: Энтропию можно рассматривать как меру неопределенности эксперимента

Ex. $\xi \in B_p$

ξ	0	1
p	$1-p$	p

$H(\xi) = -(1-p)\log_2(1-p) - p\log_2 p$

Ex. 1. Психолог Р. Хайман проводил такой эксперимент: перед человеком загорались с некоторой частотой лампочки, замерялась время реакции на загоревшуюся лампочку. Если лампочки загорались с одинаковой частотой, то энтропия была пропорциональна H_0

Ex. 2. Также с помощью энтропии определен второй закон термодинамики

Ex. 3. Теория кодирования информации

Если алфавит сообщения состоит из N символов, то каждому символу присваиваем последовательность одинаковой длины из 0 и 1, причем ее длина будет $\lceil\log_2 N\rceil$

Для передачи n символов потребуется последовательность длиной $n\lceil\log_2 N\rceil$

Цель: сократить длину последовательности

Для больших по объему сообщений можно заметно уменьшить эту величину, используя, что разные символы встречаются с разными частотами.

Если $p_1, p_2, \dots, p_N$ - эти частоты, то в сообщении длиной N i-ый символ появляется $v_i \approx np_i$ раз

Def. Сообщение длины N называется типичным с параметрами n и δ, если $|v_i - np_i| < \delta\ \forall 1 \le i \le N$

Пусть $M_{n,\delta}$ - число таких сообщений

Th. (частный случай теоремы Макмиллана)
$\frac{1}{n}\log_2 M_{n,\delta} \underset{n\to\infty}{\longrightarrow} H = -\sum_{i=1}^{N} p_i \log_2 p_i$

Следствие: существует $\varepsilon > 0 \mid \frac{1}{n}\log_2 M_{n,\delta} < H + \varepsilon$ (или $M_{n,\delta} < 2^{n(H+\varepsilon)}$)
Если можно занумеровать эти типичные сообщения, то для них потребуется число символов $\log_2 2^{n(H+\varepsilon)} = n\cdot(H+\varepsilon)$
И поэтому с вероятностью приблизительно 1 можно сократить длины сообщение с коэффициентом сжатия $\gamma \approx \frac{nH}{nH_0} = \frac{H}{H_0}$, где $H_0 = \log_2 N$
Если все символы встречаются независимо, то дальнейшее сжатие невозможно, но так как буквы встречаются в определенных сочетаниях, то можно сжать информации дальше, используя этот факт
Пусть γ_∞ - коэффициент итогового сжатия
В русском языке $\gamma \approx 0.87$. Если считать слова символами нашего алфавита, то получится $\gamma_\infty \approx 0.24$ для литературного языка и $\gamma_\infty \approx 0.17$ для делового языка

Def. $1-\gamma_\infty$ называют коэффициентом избыточности языка

Энтропия при непрерывном распределении

Def. Пусть ξ абсолютно непрерывная случайная величина с плотностью $f(x)$ и носителем $A = \{x \mid f(x) > 0\}$. Энтропией $H(\xi)$ называется величина $-\int_A f(x)\log_2 f(x)dx$

Th. Следующие распределения имеют наибольшую энтропию:

1. Если $A = [0,1]$, то $U(0,1)$
2. Если $A = [0,\infty)$ и $E\xi = 1$, то показательное E_1
3. Если $A = \mathbb{R}$ и $E\xi = 0$, а $D\xi = 1$, то $N(0,1)$

X. Программа экзамена в 2024/2025

1. *Пространство элементарных исходов. Случайные события. Операции над событиями.*
   Пространство элементарных исходов: Пространством элементарных исходов Ω называется множество, содержащее все возможные исходы экспериментов, из которых при испытании происходит ровно один. Элементы этого множества называются элементарными исходами и обозначаются ω
   Случайное событие: Случайными событиями называется подмножество $A \subset \Omega$. События A наступают, если произошел один из элементарных исходов из множества A
   Операции над событиями: Суммой $A + B$ называется событие, состоящее в том, что произошло события A или событие B (хотя бы одно из них)
   Произведением $A \cdot B$ называется событие, состоящее в том, что произошло событие A и событие B (оба из них)
   Противоположным A событием называется событие $\overline{A}$, состоящее в том, что событие A не произошло
   Дополнение (разность) $A \setminus B$ называется событие $A \cdot \overline{B}$
   События A и B называются несовместными, если их произведение - пустое множество (не могут произойти одновременно при одной эксперименте)
   События A влечет события B, если $A \subset B$ (если наступает A, то наступит B)
2. *Статистическое определение вероятности. Классическое определение вероятности.*
   Статистическое определение вероятности: Пусть проводится n реальных экспериментов, при которых событие A появилось n_A раз. Отношение $\frac{n_A}{n}$ называется частотой события A. Эксперименты показывают, что при увеличении числа n частота стабилизируется у некоторого числа, при котором мы понимаем статистическую вероятность: $P(A) \approx \frac{n_A}{n}$ при $n \to \infty$
   Классическое определение вероятности: Пусть пространство случайных событий Ω содержит конечное число равновозможных исходов, тогда применимо классическое определение вероятности: $\boxed{P(A) = \frac{|A|}{|\Omega|} = \frac{m}{n}}$, где n - число всех возможных исходов, m - число благоприятных исходов
3. *Геометрическое определение вероятности. Задача Бюффона об игле.*
   Геометрическое определение вероятности: Пусть $\Omega \subset \mathbb{R}^n$ - замкнутая ограниченная область, $\mu(\Omega)$ - мера Ω в $\mathbb{R}^n$ (например, длина отрезка, площадь области на плоскости, объем тела в пространстве), в этом случае применимо геометрическое определение вероятности: $P(A) = \frac{\mu(A)}{\mu(\Omega)}$
   Задача Бюффона об игле: пусть пол вымощен ламинатом, $2l$ - ширина доски, на пол бросается игла длины, равной ширине доски, найти вероятность того, что игла пересечет стык доски

Определим положение иглы координатами центра и углом, между иглой и стыком доски, причем можно считать, что эти величины независимы

$\exists x \in [0; l]$ - расстояние от центра до ближайшего края, $\varphi \in [0; \pi]$ - угол

$\Omega = [0; l] \times [0; \pi]$

Событие A (пересечет стык) наступает, если $x \leq l \sin\varphi$

$P(A) = \dfrac{S(A)}{S(\Omega)} = \dfrac{\int_0^\pi l \sin\varphi d\varphi}{\pi l} = \dfrac{2l}{\pi l} = \dfrac{2}{\pi}$

4. *Аксиоматическое определение вероятности. Вероятностное пространство. Свойства вероятности.*

   Аксиоматическое определение вероятности: $\exists\ \Omega$ - пространство элементарных исходов, $\mathcal{F}$ - его σ-алгебра событий. *Вероятностью* на $(\Omega, \mathcal{F})$ называется функция $P : \mathcal{F} \to \mathbb{R}$ со свойствами:

   (a) $P(A) \geq 0 \quad \forall A \in \mathcal{F}$ (неотрицательность)

   (b) Если $A_1, A_2, \dots, A_n, \dots \in \mathcal{F}$ - несовместное, то $P(\sum\limits_{i=1}^{\infty} A_i) = \sum\limits_{i=1}^{\infty} P(A_i)$ (свойство счетной аддитивности)

   (c) $P(\Omega) = 1$ (условие нормированности)

   Вероятностное пространство: Вероятностное пространство - тройка $(\Omega, \mathcal{F}, P)$

   Свойства вероятности:

   (a) Так как $\varnothing$ и Ω - несовместные, то $1 = P(\Omega) = P(\Omega + \varnothing) = 1 + P(\varnothing) \Longrightarrow P(\varnothing) = 0$

   (b) Формула обратной вероятности: $P(A) = 1 - P(\overline{A})$

   (c) $P(A) = 1 - P(\overline{A}) \leq 1$

5. *Аксиома непрерывности. Ее смысл и вывод.*

   Аксиома непрерывности: **Th.** Пусть имеется убывающая цепочка событий $A_1 \supset A_2 \supset A_3 \supset \dots \supset A_n \supset \dots$ и $\bigcap\limits_{i=1}^{\infty} A_n = \varnothing$

   Тогда $P(A_n) \underset{n\to\infty}{\longrightarrow} 0$

   При непрерывном изменении области $A \subset \Omega \subset \mathbb{R}^n$ соответствующая вероятность $P(A)$ также должна изменятся непрерывно

   Ясно, что $A_n = \sum\limits_{i=n}^{\infty} A_i\overline{A}_{i+1} + \prod\limits_{i=n}^{\infty} A_i$

   $\prod\limits_{i=n}^{\infty} A_i = A_n \cdot \prod\limits_{i=n+1}^{\infty} A_i = \prod\limits_{i=1}^{n} \cdot \prod\limits_{i=n+1}^{\infty} A_i = \prod\limits_{i=1}^{\infty} = \varnothing \Longrightarrow A_n = \sum\limits_{i=n}^{\infty} A_n\overline{A_{n+1}}$ и так как эти события несовместны, то по свойству счетной аддитивности $P(A_n) = \sum\limits_{i=n}^{\infty} P(A_i\overline{A_{i+1}})$ - это остаток (хвост) сходящегося ряда

   $P(A_1) = \sum\limits_{i=1}^{\infty} P(A_i\overline{A_{i+1}}) = \sum\limits_{i=1}^{n-1} P(A_i\overline{A_{i+1}}) + P(A_n)$ и $P(A_n) \underset{n\to\infty}{\longrightarrow} 0$ по необходимому признаку сходимости

6. *Свойства операций сложения и умножения. Формула сложения вероятностей.*

   Свойства операций сложения и умножения:

   (a) Свойство дистрибутивности: $A \cdot (B + C) = AB + AC$

   (b) Формула сложения: если A и B несовместны, то $P(A + B) = P(A) + P(B)$

   (c) Формула сложения вероятностей: $P(A + B) = P(A) + P(B) - P(AB)$

7. *Независимость событий. Независимые события в совокупности и попарно. Пример Бернштейна.*

   Независимые события: События A и B называются независимыми, если $P(A \cdot B) = P(A) \cdot P(B)$

   События $A_1, A_2, \dots A_n$ - независимы в совокупности, если для любого набора $i_1, i_2, \dots, i_k$ $(2 \le k \le n)$ $P(A_{i_1} \cdot A_{i_2} \cdot \dots \cdot A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \dots \cdot P(A_{i_k})$

   Пример Бернштейна: Пусть имеется правильный тетраэдр, одна грань окрашена в красный, вторая в синий, третья в зеленый, а четвертая во все эти три цвета.

   Подбросили тетраэдр, $\sqsupset A$ - грань, которая содержит красный цвет, B - синий, C - зеленый. $P(A) = P(B) = P(C) = \dfrac{2}{4} = \dfrac{1}{2}$

   Так как $P(AB) = P(AC) = P(BC) = \dfrac{1}{4}$

   $P(AB) = \dfrac{1}{4} = \dfrac{1}{2} \cdot \dfrac{1}{2} = P(A)P(B)$ - попарная независимость

   $P(ABC) = \dfrac{1}{4} \neq P(A)P(B)P(C)$ - но вот независимость в совокупности не соблюдается

8. *Условная вероятность. Формула умножения событий.*

   Условная вероятность $P(A|B)$ (или $P_B(A)$) - вероятность события A, вычисленная в предположении, что событие B уже произошло. $P(A|B) = \dfrac{P(AB)}{P(B)}$

   Формула умножения событий:

   Для двух событий: $P(AB) = P(B) \cdot P(A|B) = P(A) \cdot P(B|A)$

   В общем случае: $P(A_1 A_2 A_3 \dots A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\dots P(A_n|A_1A_2\dots A_{n-1})$

9. *Полная группа событий. Формула полной вероятности. Формула Байеса.*

   Полная группа событий: События $H_1, H_2, \dots, H_n, \dots$ образуют полную группу событий, если они попарно несовместны и содержат все возможные элементарные исходы

   Формула полной вероятности: $\sqsupset H_1, H_2, \dots, H_n, \dots$ - полная группа событий. Тогда $P(A) = \sum\limits_{i=1}^{\infty} P(H_i)P(A|H_i)$

   Формула Байеса: $\sqsupset H_1, H_2, \dots, H_n$ - полная группа событий, и известно, что событие A уже произошло

   Тогда $P(H_k|A) = \dfrac{P(H_k)P(A|H_k)}{\sum_{i=1}^{\infty} P(H_i)P(A|H_i)}$

10. *Последовательность независимых испытаний. Формула Бернулли. Наиболее вероятное число успехов в схеме Бернулли.*

    Схемой Бернулли называется серия одинаковых независимых экспериментов, каждый из которых имеет 2 исхода: произошло интересующее нас событие или нет

Формула Бернулли: Вероятность того, что при n испытаниях произойдет ровно k успехов, равна $p_n(k) = C_n^k p^k q^{n-k}$

Наиболее вероятное число успехов:

(a) np - целое, тогда $np+p$ - нецелое, и $k = np$ - наиболее вероятное

(b) $np+p$ - нецелое, тогда $k = \lfloor np+p \rfloor$

(c) $np+p$ - целое, тогда $np+p-1$ - целое, тогда $k \in \{np+p-1, np+p\}$

11. *Локальная и интегральная формулы Муавра-Лапласа (без док-ва).*

Локальная формула: $p_n(k) \underset{n\to\infty}{\longrightarrow} \frac{1}{\sqrt{npq}}\varphi(x)$, где $\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ - функция Гаусса, $x = \frac{k-np}{\sqrt{npq}}$

Интегральная формула: $p_n(k_1 \le k \le k_2) \underset{n\to\infty}{\longrightarrow} \Phi(x_2) - \Phi(x_1)$, где $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{z^2}{2}}dz$ - функция Лапласа, $x_1 = \frac{k_1-np}{\sqrt{npq}}$ - отклонение от левой границы, $x_2 = \frac{k_2-np}{\sqrt{npq}}$ - отклонение от правой

12. *Вероятность отклонения относительной частоты от вероятности события. Закон больших чисел Бернулли.*

Вероятность отклонения относительной частоты от вероятности события

n - число испытаний, $p = p(A)$, $\frac{n_A}{n}$ - экспериментальная частота

$$p\left(|\frac{n_A}{n} - p| \le \varepsilon\right) = p\left(-\varepsilon \le \frac{n_A}{n} - p \le \varepsilon\right) \underset{n\to\infty}{\longrightarrow} 2\Phi\left(\frac{\sqrt{n}\varepsilon}{\sqrt{pq}}\right)$$

Закон больших чисел Бернулли: $p\left(|\frac{n_A}{n} - p| \le \varepsilon\right) \underset{n\to\infty}{\longrightarrow} 2\Phi\left(\frac{\varepsilon}{\sqrt{pq}}\sqrt{n}\right) \to 1$ - закон больших чисел показывает, что вероятность попадания относительной частоты в ε-трубу приближается к 1

13. *Схемы испытаний: Бернулли, до первого успеха. Биномиальное и геометрическое распределения. Свойство отсутствия последействия.*

Схема Бернулли: $\exists v_n$ - число успехов в серии из n испытаний; $P_n(v_n = k) = C_n^k p^k q^{n-k}, \qquad k = 0, 1, \dots, n$

Биномиальное распределение: Соответствие $k \to C_n^k p^k q^{n-k}, \quad k = 0, \dots, n$ называется биномиальным распределением (обозначается $B_{n,p}$ или $B(n,p)$)

Схема до первого успеха: Пусть проводится бесконечная серия испытаний, которая заканчивается после первого успешного испытания под номером τ, тогда вероятность $P(\tau = k) = q^{k-1}p, \qquad k = 1, 2, \dots$

Геометрическое распределение: Соответствие $k \to q^{k-1}p, k \in \mathbb{N}$ называется геометрическим распределение вероятности (обозначается G_p или $G(p)$)

Геометрическое распределение обладает свойством нестарения или свойством отсутствия последействия: **Th.** $\exists P(\tau = k) = q^{k-1}p, k \in \mathbb{N}$. Тогда $\forall n, k \ge 0 \quad P(\tau > n+k \mid \tau > n) = P(\tau > k)$

14. *Урновая схема с возвратом и без возврата. Гипергеометрическое распределение. Теорема об его асимптотическом приближении к биномиальному.*

Урновая схема: В урне N шаров, из которых K шаров белые, $N-K$ - черные. Из урны вынимаем (без учета порядка) n шаров. Найти вероятность, что из них k белых

а) Схема с возвратом (после каждого раза кладем шар обратно). В этом случае вероятность вынуть белый шар одинакова и равна $\frac{K}{N}$. Получаем схему Бернулли: $P_n(k) = C_n^k \left(\frac{K}{N}\right)^k \left(1-\frac{K}{N}\right)^{n-k}$

б) Схема без возврата - вынутый шар мы выбрасываем, тогда $P_{N,K}(n,k) = \frac{C_K^k C_{N-K}^{n-k}}{C_N^n}$

Гипергеометрическое распределение: Соответствие $k \to \frac{C_K^k C_{N-K}^{n-k}}{C_N^n}, k = 0, \dots, n$ называется гипергеометрическим распределением

Теорема о приближении к биномиальному: **Th.** Если $K, N \to \infty$ таким образом, что $\frac{K}{N} \to p \in (0;1)$, а n и $0 \le k \le n$ фиксированы, то вероятность при гипергеометрическом распределении будет стремиться к биномиальному: $P_{N,K}(n,k) = \frac{C_K^k C_{N-K}^{n-k}}{C_N^n} \to C_n^k \left(\frac{K}{N}\right)^k \left(1-\frac{K}{N}\right)^{n-k}$

15. *Схема Пуассона. Формула Пуассона. Оценка погрешности в формуле Пуассона.*

Схема Пуассона: вероятность числа успеха при одном испытании p_n зависит от числа испытаний n, причем таким образом, что $np_n \approx \lambda = const$, λ - интенсивность появления редких событий в единицу времени в потоке испытаний. Применимо при p близком к 0 или к 1.

Формула Пуассона: **Th.** Пусть $n \to \infty, p_n \to 0$ таким образом, что $np_n \to \lambda = const > 0$. Тогда вероятность k успехов при n испытаниях: $P_n(k) = C_n^k p_n^k (1-p_n)^{n-k} \underset{n\to\infty}{\longrightarrow} = \frac{\lambda^k}{k!} e^{-\lambda}$

Оценка погрешности: **Th.** Пусть v_n - число успехов при n испытаниях в схеме Бернулли p - вероятность успеха при одном испытании, $\lambda = np$, $A \subset \{0, 1, \dots, n\}$ - произвольное подмножество чисел

Тогда $|P_n(v_n \in A) - \sum\limits_{k\in A} \frac{\lambda^k}{k!} e^{-\lambda}| \le \min(p, np^2) = \min(p, p\lambda)$

16. *Случайные величины, определение. Измеримость функции, ее смысл. Вероятностное пространство* $(\mathbb{R}, B, P)$. *Распределение случайной величины.*

Случайной величиной, заданной на вероятностном пространстве $(\Omega, \mathcal{F}, p)$, называется $\mathcal{F}$-измеримая функция $\xi : \Omega \to \mathbb{R}$, которая сопоставляет каждому элементарному исходу некоторое вещественное число

Измеримость: На вероятностном пространстве $(\Omega, \mathcal{F}, p)$ функция $\xi : \Omega \to \mathbb{R}$ называется $\mathcal{F}$-измеримой, если $\forall x \in \mathbb{R}$ $\{\omega \in \Omega \mid \xi(\omega) < x\} \in \mathcal{F}$ (то есть $\xi^{-1}(y) \in \mathcal{F}$, где $y \in (-\infty; x)$)

Смысл измеримости: если задана случайная величина ξ, то мы можем задать вероятность попадания случайной величины в интервал $(-\infty; x)$: $p(\xi \in (-\infty; x)) = p(\{\omega \in \Omega \mid \xi(\omega) < x\})$

Вероятностное пространство $(\mathbb{R}, B, P)$: Пусть ξ задана на вероятностном пространстве $(\Omega, \mathcal{F}, p)$, с помощью нее получаем новой вероятностное пространство $(\mathbb{R}, \mathcal{B}(\mathbb{R}), p_\xi)$, с которым проще работать

Распределение случайной величины: Функция $p(B), B \in \mathcal{B}(\mathbb{R})$, ставящая в соответствие каждому Борелевскому множеству вероятность, называется распределением случайной величины ξ

17. *Дискретные случайные величины. Определение, закон распределения, числовые характеристики.*

Дискретная случайная величина: Случайная величина ξ имеет дискретное рапределение, если она принимает не более, чем счетное число значений. То есть существует конечный или счетный набор чисел $\{x_1, x_2, \ldots, x_n, \ldots\}$ такой, что $p(\xi = x_i) = p_i > 0$ и $\sum\limits_{i=0}^{\infty} p_i = 1$

Таким образом, дискретная случайная величина (ДСВ) задается законом распределения:

ξ	x_1	x_1	$\ldots$	x_n	$\ldots$	- значения случайной величины
p	p_1	p_1	$\ldots$	p_n	$\ldots$	- вероятности этих значений

Характеристики дискретной случайной величины:

Математическим ожиданием $E\xi$ случайной величины ξ называется число $E\xi = \sum\limits_{i=1}^{\infty} x_i p_i$

Дисперсией $D\xi$ случайной величины ξ называют среднее квадратов ее отклонения от математического ожидания: $D\xi = E(\xi - E\xi)^2$ или $D\xi = \sum\limits_{i=0}^{\infty}(x_i - E\xi)^2 p_i$ при условии, что данный ряд сходится

Дисперсию обычно удобно считать по формуле $D\xi = E\xi^2 - (E\xi)^2 = \sum\limits_{i=1}^{n} x_i^2 p_i - E\xi^2$

Средним квадратическим отклонением (СКО) σ_ξ называется величина $\sigma_\xi = \sqrt{D\xi}$

$m_k = E\xi^k$ - момент k-ого порядка случайной величины ξ (также называют начальным моментом)

$\mu_k = E(\xi - E\xi)^k$ - центральный момент k-ого порядка

18. *Свойства математического ожидания и дисперсии дискретной случайной величины.*

Свойства:

Th. 1. Случайная величина ξ имеет вырожденное распределение, если $\xi(\omega) = \text{const} \;\; \forall \omega \in \Omega$

ξ	C
p	1

$E\xi = C \qquad D\xi = 0$

Th. 2. Свойство сдвига: $E(\xi + C) = E\xi + C; D(\xi + C) = D\xi$

Th. 3. Свойство растяжения: $E(C\xi) = CE\xi$, $D(C\xi) = C^2 D\xi$

Th. 4. $E(\xi + \eta) = E\xi + E\eta$ (из третьего свойства матожидание - линейная функция)

Def. Дискретные случайные величины ξ и η независимы, если $p(\xi = x_i, \eta = y_i) = p(\xi = x_i) \cdot p(\eta = y_i) \; \forall i, j$. То есть случайные величины принимают свои величины независимо друг от друга

Th. 5. Если случайные величины ξ и η независимы, то $E(\xi\eta) = E\xi \cdot E\eta$; обратное неверно

Th. 6. $D\xi = E\xi^2 - (E\xi)^2$

Def. $D(\xi+\eta) = D\xi + D\eta + 2\mathrm{cov}(\xi,\eta)$, где $\mathrm{cov}(\xi,\eta) = E(\xi\eta) - E\xi E\eta$ - ковариация случайных величин (равна 0 при независимых величинах) - индикатор наличия связи между случайными величинами

Th. 7. Если случайные величины ξ и η независимы, то $D(\xi+\eta) = D\xi + D\eta$

Th. 8. Общая формула дисперсии суммы: $D(\xi_1+\xi_2+\dots+\xi_n) = \sum\limits_{i=1}^{n} D\xi_i + 2\sum\limits_{i,j(i\neq j)} \mathrm{cov}(\xi_i,\xi_j)$

19. *Стандартные дискретные распределения и их числовые характеристики (Бернулли, биномиальное, геометрическое, Пуассона).*

Распределение Бернулли: B_p (с параметром $0 < p < 1$), ξ - число успехов при одном испытании, p - вероятность успеха при одном испытании

ξ	0	1
p	$1-P(A)$	$P(A)$

$E\xi = p$ $\qquad D\xi = p(1-p) = pq$

Биномиальное распределение $B_{n,p}$ (с параметрами n,p), ξ - число успехов в серии из n испытаний, p - вероятность успеха при одном испытании

$p(\xi=k) = C_n^k p^k q^{n-k},\ k=0,1,\dots,n \Longleftrightarrow \xi \in B_{n,p}$

$E\xi_i = p; \quad D\xi_i = pq$

$E\xi = E\xi_1 + \dots + E\xi_n = p + \dots + p = np$

$D\xi = D\xi_1 + \dots + D\xi_n = pq + \dots + pq = npq$

Геометрическое распределение G_p (с параметром p), ξ - номер 1-ого успешного испытания в бесконечной серии

$p(\xi=k) = q^{k-1}p,\ k=1,2,3,\dots \Longleftrightarrow \xi \in G_p$

$E\xi = \frac{1}{p}, D\xi = \frac{q}{p^2}$

Распределение Пуассона Π_λ (с параметром $\lambda > 0$)

Случайная величина ξ имеет распределение Пуассона с параметром $\lambda > 0$, если $p(\xi = k) = \frac{\lambda^k}{k!}e^{-\lambda},\ k=0,1,2,\dots$

$E\xi = \lambda = np, D\xi = \lambda$

20. *Функция распределения и ее свойства (в свойствах 4, 5, 6 достаточно привести одно из доказательств).*

Функция распределения $F_\xi(x)$ случайной величины ξ называется функция $F_\xi(x) = P(\xi < x)$

Свойства:

1) $F(x)$ ограничена $0 \le F(x) \le 1$

2) $F(x)$ неубывающая функция: $x_1 < x_2 \Longrightarrow F(x_1) \le F(x_2)$

3) $p(\alpha \le \xi < \beta) = F(\beta) - F(\alpha)$

4) $\lim\limits_{x\to-\infty} F(x) = 0; \quad \lim\limits_{x\to+\infty} F(x) = 1$

5) $F(x)$ непрерывна слева: $F(x_0 - 0) = F(x_0)$

6) Скачок в точке x_0 равен вероятности попадания в данную точку: $F(x_0+0) - F(x_0) = p(\xi = x_0)$ или $F(x_0+0) = p(\xi=x_0) + p(\xi<x_0) = p(\xi \le x_0)$

7) Если функция распределения непрерывна в точке $x = x_0$, то очевидно, что вероятность

попадания в эту точка $p(\xi = x_0) = 0$ (следствие из 6 пункта)

8) Если $F(x)$ непрерывна $\forall x \in \mathbb{R}$, то $p(\alpha \le \xi < \beta) = p(\alpha < \xi < \beta) = p(\alpha \le \xi \le \beta) = p(\alpha < \xi \le \beta) = F(\beta) - F(\alpha)$

21. *Абсолютно непрерывные случайные величины. Плотность и ее свойства.*

Абсолютно непрерывне случайные величины: Случайная величина ξ имеет абсолютно непрерывное распределение, если существует $f_\xi(x)$ такая, что $\forall B \in \mathcal{B}(\mathbb{R})$ $p(\xi \in B) = \int_B f_\xi(x)dx$

Функция плотности: Функция f_ξ называется плотностью распределения случайной величины

Свойства:

1) Вероятносто-геометрический смысл плотности: $p(\alpha \le \xi < \beta) = \int_\alpha^\beta f_\xi(x)dx$

2) Условие нормировки: $\int_{-\infty}^{+\infty} f_\xi(x)dx = 1$

3) $F_\xi(x) = \int_{-\infty}^{x} f_\xi(x)dx$

4) $F_\xi(x)$ непрерывна

5) $F_\xi(x)$ дифференцируема почти везде и $f_\xi(x) = F'_\xi(x)$ для почти всех x

6) $f_\xi(x) \ge 0$ по определению и как производная неубывающей $F_\xi(x)$

7) $p(\xi = x) = 0$ $\forall x \in \mathbb{R}$ - так как $F_\xi(x)$ непрерывна

8) $p(\alpha \le \xi < \beta) = p(\alpha < \xi < \beta) = p(\alpha \le \xi \le \beta) = p(\alpha < \xi \le \beta) = F(\beta) - F(\alpha)$

9) **Th.** Если $f(x) \ge 0$ и $\int_{-\infty}^{\infty} f(x)dx = 1$ (выполнены свойства 2 и 6), то $f(x)$ - плотность некоторого распределения

22. *Числовые характеристики абсолютно непрерывной случайной величины, их свойства.*

Характеристики:

Математическим ожиданием $E\xi$ случайной абсолютно непрерывной величины ξ называется величина $E\xi = \int_{-\infty}^{\infty} x f_\xi(x)dx$ при условии, что данный интеграл сходится абсолютно, то есть $\int_{-\infty}^{\infty} |x| f_\xi(x)dx < \infty$

Дисперсией $D\xi$ случайной величины ξ называется величина $D\xi = E(\xi - E\xi)^2 = \int_{-\infty}^{\infty} (x - E\xi)^2 f_\xi(x)dx$ при условии, что данный интеграл сходится. Вычислять удобно по формуле $D\xi = E\xi^2 - (E\xi)^2 = \int_{-\infty}^{\infty} x^2 f_\xi(x)dx - (E\xi)^2$

Среднее квадратическое отклонение $\sigma_\xi = \sqrt{D\xi}$ определяется, как корень дисперсии

$m_k = E\xi^k = \int_{-\infty}^{\infty} x^k f_\xi(x)dx$ - момент k-ого порядка

$\mu_k = E(\xi - E\xi)^k = \int_{-\infty}^{\infty} (x - E\xi)^k f_\xi(x)dx$ - центральный момент k-ого порядка

Медианой Me абсолютно непрерывной случайной величины ξ называется значение случайной величины ξ, такое что $p(\xi < Me) = p(\xi > Me) = \frac{1}{2}$

Модой Mo случайной величины ξ называется точка локального максимума плотности

23. *Равномерное распределение.*

Равномерное распределение: Случайная величина ξ имеет равномерное распределение $\xi \in U(a,b)$, если ее плотность на этом отрезке постоянна. Получаем функцию плотности

$$f_\xi(x) = \begin{cases} 0, & x < a \\ \frac{1}{b-a}, & a \le x < b \\ 0 & x \ge b \end{cases}$$

$\frac{1}{b-a}$ из усл. нормировки

$$F(x) = \int_{-\infty}^{\infty} f(x)dx = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \le x < b \\ 1 & x \ge b \end{cases}$$

$$E\xi = \frac{a+b}{2}, \quad D\xi = \frac{(b-a)^2}{12}, \quad \sigma = \frac{b-a}{2\sqrt{3}}$$

$p(\alpha < \xi < \beta) = \dfrac{\beta-\alpha}{b-a}$ при условии, что $\alpha, \beta \in [a,b]$

24. *Показательное распределение. Свойство нестарения.*

Показательное распределение: Случайная величина ξ имеет показательное (или экспоненциальное) распределение с параметром $\alpha > 0$ (обозн. $\xi \in E_\alpha$), если ее плотность имеет вид: $f_\xi(x) = \begin{cases} 0, & x < 0 \\ \alpha e^{-\alpha x}, & x \ge 0 \end{cases}$

$$F_\xi(x) = \begin{cases} 0, & x < 0 \\ \int_0^x \alpha e^{-\alpha x} = 1 - e^{-\alpha x}, & x \ge 0 \end{cases}$$

$$E\xi = \frac{1}{\alpha}, \quad D\xi = \frac{1}{\alpha^2}, \quad \sigma = \frac{1}{\alpha}$$

$$p(\alpha < \xi < \beta) = F(b) - F(a) = e^{-a\alpha} - e^{-b\alpha} \qquad a, b \ge 0$$

Из непрерывных случайных величин только показательная обладает свойством нестарения: **Th.** $\sqsupset \xi \in E_\alpha$. Тогда $p(\xi > x+y \mid \xi > x) = p(\xi > y) \qquad \forall x, y > 0$

25. *Нормальное распределение. Стандартное нормальное распределение, его числовые характеристики.*

Нормальное распределение: Случайная величина ξ имеет нормальное распределение с параметрами a и σ^2 (обозн. $\xi \in N(a,\sigma^2)$), если ее плотность имеет вид: $f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(t-a)^2}{2\sigma^2}} dt$$

$$E\xi = a, \quad D\xi = \sigma^2, \quad \sigma = \sigma$$

Стандартным нормальным распределением называется нормальное распределение с параметрами $a = 0, \sigma^2 = 1$: $\xi \in N(0,1)$

Плотность: $\phi(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ - функция Гаусса

Распределение: $F_0(x) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{x} e^{-\frac{z^2}{2}} dz$ - функция стандартного нормального распределения

$E\xi = 0;\ D\xi = 1$

26. *Связь между стандартным нормальным и нормальным распределениями. Следствия.*

    Связь:

    1) $\sqsupset \xi \in N(a, \sigma^2)$. Тогда $F_\xi(x) = F_0\left(\dfrac{x-a}{\sigma}\right)$

    2) Если $\xi \in N(a, \sigma^2)$, то $\eta = \dfrac{\xi - a}{\sigma} \in N(0, 1)$ (процесс $\xi \to \eta$ называется стандартизацией)

    3) $\sqsupset \xi \in N(a, \sigma^2)$. Тогда $p(\alpha < \xi < \beta) = \Phi\left(\dfrac{\beta - a}{\sigma}\right) - \Phi\left(\dfrac{\alpha - a}{\sigma}\right)$

    4) Вероятность попадания в симметричный интервал (вероятность отклонения случайной величины от матожидания) $p(|\xi - a| < t) = 2\Phi\left(\dfrac{t}{\sigma}\right)$

    5) Правило 3 «сигм»: $p(|\xi - a| < 3\sigma) \approx 0.9973$ - попадание случайной величины нормального распределения в интервал $(a - 3\sigma, a + 3\sigma)$ близко к 1

    6) Свойство линейности: если случайная величина $\xi \in N(a, \sigma^2)$, то $\eta = \gamma\xi + b \in N(a\gamma + b, \gamma^2\sigma^2)$

    7) Устойчивость относительно суммирования: если случайные величины $\xi_1 \in N(a_1, \sigma_1^2)$, $\xi_2 \in N(a_2, \sigma_2^2)$, и они независимы, то $\xi_1 + \xi_2 \in N(a_1 + a_2, \sigma_1^2 + \sigma_2^2)$

27. *Сингулярные распределения. Теорема Лебега (без док-ва).*

    Сингулярное распределение: Случайная величина ξ имеет сингулярное распределение, если $\exists B$ - Борелевское множество с нулевой мерой Лебега $\lambda(B) = 0$, такое что $p(\xi \in B) \in 1$, но $P(\xi = x) = 0\ \ \forall x \in B$

    Теорема Лебега: **Th. Лебега.**

    $\sqsupset F_\xi(x)$ - функция распределения ξ. Тогда $F_\xi(x) = p_1F_1(x) + p_2F_2(x) + p_3F_3(x)$, где $p_1 + p_2 + p_3 = 1$

    F_1 - функция дискретного распределения

    F_2 - функция абсолютно непрерывного распределения

    F_3 - функция сингулярного распределения

    То есть существуют только дискретное, абсолютно непрерывное, сингулярное распределения и их смеси

28. *Преобразования случайных величин. Стандартизация случайной величины.*

    Стандартизация: Пусть имеется случайная величина ξ. Соответствующей ей стандартной величиной называется случайная величина $\eta = \dfrac{\xi - E\xi}{\sigma}$

    $E\eta = 0; D\eta = 1$

    Преобразование: Если ξ - дискретная случайная величина, то ее законы распределения находятся просто: значения x_i в верхней строке заменяем $g(x_i)$, вероятности остаются прежние.

29. *Теорема о монотонном преобразовании. Линейное преобразование случайной величины. (без док-ва).*

Теорема о монотонном преобразовании: **Th.** Пусть $f_\xi(x)$ - плотность случайной величины ξ, $g(x)$ - строго монотонная функция. Тогда случайная величина $\eta = g(\xi)$ имеет плотность $f_\eta(x) = |h'(x)|f_\xi(h(x))$, где $h(g(x)) = x$

Если $g(x)$ не является монотонной функцией, то поступаем следующим образом: разбиваем $g(x)$ на интервалы монотонности, для каждого i-ого интервала находим $h_i(x)$ и плотность случайной величины ищем по *формуле Смирнова*: $f_\eta(x) = \sum\limits_{i=0}^{n} |h_i'(x)|f_\xi(h_i(x))$

Линейное преобразование: **Th.** Пусть ξ имеет плотность $f_\xi(x)$, тогда $\eta = a\xi + b$, где $a \neq 0$, имеет плотность $f_\eta(x) = \dfrac{1}{|a|} f_\xi\left(\dfrac{x-b}{a}\right)$

30. *Квантильное преобразование. Моделирование случайной величины с помощью датчика случайных чисел.*

Квантильное преобразование: Пусть функция распределения случайной величины ξ $F_\xi(x)$ - непрерывная функция. Тогда $\eta = F(\xi) \in U(0,1)$ - стандартное равномерное распределение

Пусть $\eta \in U(0,1)$ - стандартное равномерное распределение, $F(x)$ - произвольная функция распределения. Тогда $\xi = F^{-1}(\eta)$ имеет функцию распределения $F(x)$

Преобразование $\xi = F^{-1}(\eta)$ называют квантильным

Смысл: датчики случайных чисел имеют стандартное равномерное распределение, из теоремы следует, что при помощи датчика случайных чисел и квантильного преобразования мы сможем смоделировать любое нужно распределение

31. *Виды сходимостей случайных величин, связь между ними. Теорема об эквивалентности сходимостей к константе (все без док-ва).*

Виды сходимостей:

- Сходимость «почти наверное»

  Def. Последовательность случайных величин $\{\xi_n\}$ сходится «почти наверное» к случайной величине ξ при $n \to \infty$ ($\xi_n \xrightarrow{\text{п. н.}} \xi$), если $p(\omega \in \Omega \mid \xi_n(\omega) \underset{n\to\infty}{\longrightarrow} \xi(\omega)) = 1$
- Сходимость по вероятности

  Def. Последовательность случайных величин $\{\xi_n\}$ сходится по вероятности к случайной величине ξ при $n \to \infty$ ($\xi_n \xrightarrow{p} \xi$), если $\forall \varepsilon > 0 \quad p(|\xi_n - \xi| < \varepsilon) \underset{n\to\infty}{\longrightarrow} 1$
- Слабая сходимость

  Def. Последовательность случайных величин ξ_n слабо сходится к случайной величине ξ при $n \to \infty$ ($\xi_n \rightrightarrows \xi$), если $F_{\xi_n}(x) \longrightarrow F_\xi(x) \forall x$, где $F_\xi(x)$ - непрерывна

Связь:

Th. $\xi_n \xrightarrow{\text{п. н.}} \xi \Longrightarrow \xi_n \xrightarrow{p} \xi \Longrightarrow \xi_n \rightrightarrows \xi$

Th. Если $\xi_n \rightrightarrows C = \text{const}$, то $\xi_n \xrightarrow{p} C$

Nota. В общем случае не только из слабой сходимости не следует сходимость по вероятности, но и бессмысленно говорить об этом, так как слабая сходимость - это сходимость не случайных величин, а их распределений

32. *Математическое ожидание преобразованной случайной величины. Свойства моментов.*
Матожидание: **Th.** Если ξ - дискретная случайная величина, то $Eg(\xi) = \sum_{i=1}^{\infty} g(x_i) \cdot p(\xi = x_i)$
Для непрерывной случайной величины $Eg(\xi) = \int_{-\infty}^{\infty} g(x) f_\xi(x) dx$
Свойства моментов: 1) Если $\xi \geq 0$, то $E\xi \geq 0$
2) Если $\xi \leq \eta$, то $E\xi \leq E\eta$
3) Если $|\xi| \leq |\eta|$, то $E|\xi|^k \leq E|\eta|^k$
4) Если существует момент m_t случайной величины ξ, то существует m_s при $s < t$ (при условии, что интеграл/сумма сходятся)
33. *Неравенство Йенсена, следствие.*
Неравенство Йенсена: **Th.** Пусть функция $g(x)$ выпукла вниз, тогда для любой случайной величины ξ $Eg(\xi) \geq g(E\xi)$

Nota. Если $g(x)$ выпукла вверх, знак неравенства меняется
Следствие: $Ee^\xi \geq e^{E\xi}$, $E\xi^2 \geq (E\xi)^2$, $E|\xi| \geq |E\xi|$, $E\ln(\xi) \leq \ln(E\xi)$, $E\frac{1}{\xi} \geq \frac{1}{E\xi}$ при $\xi > 0$
34. *Неравенства Маркова, Чебышева, правило трех сигм.*
Для ξ, у которой существует матожидание, верно:
Неравенство Маркова: **Th.** $p(|\xi| \geq \varepsilon) \leq \frac{E|\xi|}{\varepsilon}$ $\forall \varepsilon > 0$
Неравенство Чебышева: **Th.** $P(|\xi - E\xi| \geq \varepsilon) \leq \frac{D\xi}{\varepsilon^2}$
Правило «трех сигм»: **Th.** $P(|\xi - E\xi| \geq 3\sigma) \leq \frac{1}{9}$
35. *Среднее арифметическое одинаковых независимых случайных величин. Закон больших чисел Чебышева.*
Среднее арифметическое: $\frac{S_n}{n} = \frac{\xi_1 + \dots + \xi_n}{n}$
$E\left(\frac{S_n}{n}\right) = \frac{1}{n}(E\xi_1 + \dots + E\xi_n) = \frac{1}{n}na = a = E\xi_1$ - математическое ожидание не меняется
$D\left(\frac{S_n}{n}\right) = \frac{1}{n^2}(D\xi_1 + \dots + D\xi_n) = \frac{1}{n^2}nd = \frac{d}{n} = \frac{D\xi_1}{n}$ - дисперсия уменьшилась в n раз
$\sigma\left(\frac{S_n}{n}\right) = \frac{\sigma}{\sqrt{n}}$ - СКО уменьшилось в $\sqrt{n}$ раз
Закон больших чисел Чебышев: **Th.** Пусть $\xi_1, \dots, \xi_n, \dots$ - последовательность независимых одинаково распределенных с конечным вторым моментом, тогда $\frac{\xi_1 + \dots + \xi_n}{n} \xrightarrow[n\to\infty]{p} E\xi_1$
36. *Вывод закона больших чисел Бернулли из закона больших чисел Чебышева. Законы больших чисел Хинчина и Колмогорова (только формулировки).*
ЗБЧ Бернулли: **Th.** Пусть v_n - число успехов из n независимых испытаний, $p = P(A)$ - вероятность успеха при одном испытании. Тогда $\frac{v_n}{n} \xrightarrow{p} P(A)$

$v_n = \xi_1 + \cdots + \xi_n$, где $\xi_i \in B_p$ - число успехов при i-ом испытании
$E\xi_i = p; D\xi_i = pq$
$\frac{v_n}{n} \xrightarrow{p} E\xi_1 = p$
$p\left(\left|\frac{v_n}{n} - p\right| \geq \varepsilon\right) \leq \frac{D\xi_1}{n\varepsilon^2} = \frac{pq}{n\varepsilon^2}$

ЗБЧ Хинчина: **Th.** $v_n = \xi_1 + \cdots + \xi_n$ последовательность независимых одинаково распределенных случайных величин с конечным первым моментом, тогда $\frac{\xi_1 + \cdots + \xi_n}{n} \xrightarrow{p} E\xi_i$

ЗБЧ Колмогорова: В условиях теоремы Хинчина $\frac{\xi_1 + \cdots + \xi_n}{n} \xrightarrow{\text{п.н.}} E\xi_1$

37. *Совместные распределения случайных величин. Функция совместного распределения, ее свойства. Независимость случайных величин.*

Совместное распределение: Случайным вектором $\vec{\xi} = (\xi_1, \xi_2, \ldots, \xi_n)$ называется упорядоченный набор случайных величин, заданных на одном вероятностном пространстве

Случайный вектор задает отображение $(\xi_1, \ldots, \xi_n)(\omega) : \Omega \longrightarrow \mathbb{R}^n$

Функция совместного распределения: Функцией совместного распределения случайных величин $\xi_1, \xi_2, \ldots, \xi_n$ называется функция $F_{\xi_1, \xi_2, \ldots, \xi_n}(x_1, x_2, \ldots, x_n) = P(\xi_1 < x_1, \xi_2 < x_2, \ldots, \xi_n < x_n)$

Свойства:

(a) $0 \leq F_{\xi,\eta}(x, y) \leq 1$
(b) $F_{\xi,\eta}(x, y)$ - неубывающая по каждому аргументу
(c) $\lim\limits_{x \to -\infty} F_{\xi,\eta}(x, y) = \lim\limits_{y \to -\infty} F_{\xi,\eta}(x, y) = 0$, $\lim\limits_{\substack{x \to \infty \\ y \to \infty}} F_{\xi,\eta}(x, y) = 1$
(d) Восстановление маргинального (частного) распределения: $\lim\limits_{x \to \infty} F_{\xi,\eta}(x, y) = F_\eta(y)$, и наоборот - $\lim\limits_{y \to \infty} F_{\xi,\eta}(x, y) = F_\xi(x)$
(e) $F_{\xi,\eta}(x, y)$ - непрерывна слева по каждому аргументу
(f) $P(x_1 \leq \xi < x_2, y_1 \leq \eta < y_2) = F_{\xi,\eta}(x_2, y_2) - F_{\xi,\eta}(x_2, y_1) - F_{\xi,\eta}(x_1, y_2) + F_{\xi,\eta}(x_1, y_1)$

Независимость величин: Случайные величины $\xi_1, \ldots, \xi_n$ независимы в совокупности, если для любого набора Борелевских множеств из $\mathcal{B}(\mathbb{R}^n)$, $B_1, B_2, \ldots, B_n$ верно $p(\xi_1 \in B_1, \xi_2 \in B_2, \ldots, \xi_n \in B_n) = p(\xi_1 \in B_1) \cdot p(\xi_2 \in B_2) \cdot \cdots \cdot p(\xi_n \in B_n)$

Случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ попарно независимы, если независимы любые две из них

38. *Дискретная система двух случайных величин. Закон совместного распределения. Маргинальные распределения.*

Дискретная система: Случайные величины ξ, η имеют совместное дискретное распределение, если случайный вектор (ξ, η) принимает не более, чем счетное число значений, то есть существует конечный или счетный набор пар чисел (x_i, y_i), таких что $P(\xi = x_i, \eta = y_i) > 0, \sum\limits_{i,j} P(\xi = x_i, \eta = y_i) = 1$

Таким образом двумерная дискретная случайная величина задается законом распределения - таблице вероятностей

$\xi\backslash\eta$	y_1	y_2	$\dots$	y_m
x_1	p_{11}	p_{12}	$\dots$	p_{1m}
x_2	p_{21}	p_{22}	$\dots$	p_{2m}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
x_n	p_{n1}	p_{n2}	$\dots$	p_{nm}

Зная общий закон распределения, можно восстановить частное (маргинальное) распределение по формулам:

$p_i = \sum_{j=1}^{m} p_{i,j} \qquad q_j = \sum_{i=1}^{n} p_{i,j}$

39. *Абсолютно непрерывная система двух случайных величин. Плотность совместного распределения, ее свойства.*

Непрерывная система: Случайные величины ξ и η имеют абсолютно непрерывное совместное распределение, если $\exists f_{\xi,\eta}(x,y)$, такая что $\forall B \in \mathcal{B}(\mathbb{R}^2)$ $P((\xi,\eta) \in B) = \iint_B f_{\xi,\eta}(x,y)dxdy$
Функцию $f_{\xi,\eta}(x,y)$ будем называть функцией плотности совместного распределения случайных величин ξ и η

Свойства:

(a) $f_{\xi,\eta}(x,y) \geq 0$

(b) Условие нормировки: $\iint_{\mathbb{R}^2} f_{\xi,\eta}(x,y)dxdy = 1$

(c) $F_{\xi,\eta} = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{\xi,\eta}(x,y)dydx$

(d) $f_{\xi,\eta}(x,y) = \dfrac{\partial^2 F_{\xi,\eta}(x,y)}{\partial x \partial y}$

(e) Если случайные величины ξ,η имеют абсолютно непрерывное совместное распределение с плотностью $f(x,y)$, то маргинальное распределение величин ξ,η также имеют абсолютно непрерывное распределение с плотностями $f_\xi(x) = \int_{-\infty}^{\infty} f_{\xi,\eta}(x,y)dy, f_\eta(y) = \int_{-\infty}^{\infty} f_{\xi,\eta}(x,y)dx$

(f) Так как вероятность попадания в Борелевские множества полностью задается функцией распределения, то условие независимости случайных величин эквивалентно следующему:
$\xi_1,\xi_2,\dots,\xi_n$ независимы, если функция общего распределения распадается в произведение отдельных функцию распределения
$F_{\xi_1,\xi_2,\dots,\xi_n}(x_1,x_2,\dots,x_n) = F_{\xi_1}(x_1)\cdot F_{\xi_2}(x_2)\cdot\dots\cdot F_{\xi_n}(x_n)$

(g) *Равносильное определение*: абсолютно непрерывные случайные величины $\xi_1,\dots,\xi_n$ независимы в совокупности тогда и только тогда, когда плотность совместного распределения $f_{\xi_1,\xi_2,\dots,\xi_n}(x_1,x_2,\dots,x_n) = f_{\xi_1}(x_1)\cdot f_{\xi_2}(x_2)\cdot\dots\cdot f_{\xi_n}(x_n)$

40. *Функции от двух случайных величин. Теорема о функции распределения. Формула свертки.*

Функция от двух случайных величин: **Th.** Пусть ξ_1, ξ_2 - случайные величины с общем плотностью $f_{\xi_1,\xi_2}(x,y)$, и есть функция $g(x,y) \ : \ \mathbb{R}^2 \to \mathbb{R}$. Тогда случайная величина $\eta = g(\xi_1, \xi_2)$ имеет функцию распределения $F_\eta(z) = \iint_{D_z} f(x,y)dxdy$, где $D_z = \{(x,y) \in \mathbb{R}^2 \mid g(x,y) < z\}$

Плотность суммы: **Th.** $\sqsupset \xi_1, \xi_2$ - независимые абсолютно непрерывные случайные величины с плотностями $f_{\xi_1}(x)$ и $f_{\xi_2}(y)$

Тогда плотность суммы $\xi_1 + \xi_2$ равна $f_{\xi_1+\xi_2}(t) = \int_{-\infty}^{\infty} \underbrace{f_{\xi_1}(x) f_{\xi_2}(t-x)}_{\text{т. н. свертка}} dx$

41. *Суммы стандартных распределений, устойчивость по суммированию (биномиальное, Пуассона, стандартное нормальное).*

Суммы стандартных распределений: *Ex. 1.* $\xi \in B_{n,p}; \eta \in B_{m,p}$. Тогда ясно, что $\xi + \eta \in B_{n+m,p}$ (по определению биномиального распределения $B_{n,p}$ - число успехов из n испытаний, где p - вероятность успеха)

Ex. 2. $\xi \in \Pi_\lambda, \eta \in \Pi_\mu$, они независимы. Тогда $\xi + \eta \in \Pi_{\lambda+\mu}$

Ex. 3. $\xi, \eta \in N(0,1)$ и независимы. Тогда $\xi + \eta \in N(0,2)$

Ex. 4. В общности для независимых $\xi \in N(a_1, \sigma_1^2), \eta \in N(a_2, \sigma_2^2)$ $\xi + \eta \in N(a_1 + a_2, \sigma_1^2 + \sigma_2^2)$

Ex. 5. Равномерное распределение неустойчиво относительно суммирования, контрпример:

$\xi, \eta \in U(0,1)$ - независимы

$\forall x, y \in [0,1]$ $f_\xi(x) = 1, f_\eta(y) = 1$ и $f_{\xi,\eta}(x,y) = 1$

По первой теореме $F_{\xi,\eta}(x,y) = \iint_{D_z} f_{\xi,\eta}(x,y)dxdy = \iint_{D_z} dxdy = S_{D_z}$, где $D_z = \{(x,y) \mid x+y < z\}$

42. *Условные распределения и условные математические ожидания. Случаи дискретной и абсолютно непрерывной систем двух случайных величин.*

Условным распределением случайной величины из системы случайных величин (ξ, η) называется ее распределение, найденное при условии, что другая случайная величина приняла определенное значение. Обозначается $\xi | \eta = y$

Условным математическим ожиданием (обозначается $E(\xi | \eta = y)$) называется математическим ожиданием случайной величины ξ при соответствующем условном распределении

Дискретная система: Пусть (ξ, η) задана законом распределения:

$\xi \backslash \eta$	y_1	y_2	$\dots$	y_m
x_1	p_{11}	p_{12}	$\dots$	p_{1m}
x_2	p_{21}	p_{22}	$\dots$	p_{2m}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
x_n	p_{n1}	p_{n2}	$\dots$	p_{nm}

Вероятности условных распределений считаем по формулам:

$\xi|\eta = y_j$: $p_i = p(\xi = x_i \mid \eta = y_j) = \dfrac{p(\xi = x_i, \eta = y_j)}{p(\eta = y_j)} = \dfrac{p_{ij}}{q_j} = \dfrac{p_{ij}}{\sum_i p_{ij}}$

$\eta|\xi = x_i$: $q_j = p(\eta = y_j \mid \xi = x_i) = \dfrac{p(\xi = x_i, \eta = y_j)}{p(\xi = x_i)} = \dfrac{q_{ij}}{p_i} = \dfrac{q_{ij}}{\sum_j p_{ij}}$

Матожидание $E(\xi|\eta = y_j) = \sum\limits_i x_i p(\xi = x_i, \eta = y_j)$

Непрерывная система: Пусть (ξ, η) задана плотностью $f_{\xi,\eta}(x, y)$ совместного распределения, тогда плотность условного распределения $\xi|\eta = y$: $f(x|y) = \dfrac{f_{\xi,\eta}(x, y)}{\int_{\mathbb{R}} f_{\xi,\eta}(x, y)dx} = \dfrac{f_{\xi,\eta}(x, y)}{f_\eta(y)}$

Def. Функция $f(x|y) = \dfrac{f_{\xi,\eta}(x, y)}{f_\eta(y)}$ называется условной плотностью

Def. Условное математические ожидание вычисляется по формуле $E(\xi|\eta = y) = \displaystyle\int_{-\infty}^{\infty} x f(x|y)dx$

43. *Пространство случайных величин. Скалярное произведение, неравенство Коши-Буняковского-Шварца.*

    Пространство $L_2(\Omega, \mathcal{F}, P) = \{\xi \mid D\xi < \infty\}$ - множество случайных величин на данном пространстве с конечной дисперсией

    Скалярным произведением случайных величин ξ и η из $L_2(\Omega, \mathcal{F}, P)$ называется число $(\xi, \eta) = E(\xi\eta)$

    Неравенство Коши-Буняковского-Шварца: **Th.** Пусть случайные величины ξ и η имеют конечный второй момент, тогда $|E(\xi, \eta)| \le \sqrt{E\xi^2 \cdot E\eta^2}$ (или $|(\xi, \eta)| \le \|\xi\| \cdot \|\eta\|$)

    Причем $|E(\xi, \eta)| = \sqrt{E\xi^2 \cdot E\eta^2} \iff \eta = C\xi$, где $C = \text{const}$

44. *Условное математическое ожидание как случайная величина, его свойства. Формула полного математического ожидания.*

    Условным математическим ожиданием (УМО, обозначается $E(\xi|\eta) = \hat{\xi}$) случайной величины ξ относительно случайной величины η называется ортогональная проекция случайной величины ξ на $L(\eta)$

    Свойства:

    (a) Тождество ортопроекций: $\exists \hat{\xi} \in L(\eta)$, тогда $\hat{\xi} = E(\xi|\eta) \iff E(\xi \cdot g(\eta)) = E(\hat{\xi} \cdot g(\eta))\ \forall g(\eta) \in L(\eta)$

    (b) Формула полного математического ожидания

    $E\xi = E(E(\xi|\eta))$ или $E\xi = E\hat{\xi}$

    Nota. При распределении Бернулли получаем обычную формулу полной вероятности

    (c) Линейность: $E(C_1\xi_1 + C_2\xi_2 \mid \eta) = C_1 E(\xi_1|\eta) + C_2 E(\xi_2|\eta)$

    (d) Если ξ и η независимы, то $E(\xi|\eta) = E\xi$

    (e) Если ξ и η независимы, то $(\xi - E\xi) \perp g(\eta)\ \forall g(\eta) \in L(\eta)$, в частности $(\xi - E\xi) \perp \eta$

45. *Условная дисперсия. Закон полной дисперсии. Смысл второго слагаемого в разложении дисперсии.*

    Условной дисперсией случайной величины ξ относительно случайной величины η называется случайная величина $D(\xi|\eta) = E((\xi - E(\xi|\eta))^2|\eta)$

Закон полной дисперсии: **Th.** $D\xi = E(D(\xi|\eta)) + D(E(\xi|\eta))$

Следствие и смысл:

- Если ξ и η независимы (некоррелированы), то $D(E(\xi|\eta)) = D(E\xi) = 0$ и $D\xi = E(D(\xi|\eta))$
- Если имеется функциональная зависимость (то есть $\xi = g(\eta)$), то $D(E(\xi|\eta)) = D(E(g(\eta)|\eta)) = D(g(\eta)) = D\xi$

46. *Числовые характеристики зависимости случайных величин. Ковариация, ее свойства. Коэффициент корреляции, его свойства. Корреляция случайных величин.*

Ковариацией (ξ, η) называется величина $\text{cov}(\xi, \eta) = E((\xi - E\xi)(\eta - E\eta))$

Свойства:

(a) $\text{cov}(\xi, \eta) = E(\xi\eta) - E\xi E\eta$

(b) $\text{cov}(\xi, \xi) = D\xi$

(c) $\text{cov}(\xi, \eta) = \text{cov}(\eta, \xi)$

(d) $\text{cov}(C_1\xi_1 + C_2\xi_2, \eta) = C_1\text{cov}(\xi_1, \eta) + C_2\text{cov}(\xi_2, \eta)$

(e) $D(\xi + \eta) = D\xi + D\eta + 2\text{cov}(\xi, \eta)$

(f) $D(\xi_1 + \cdots + \xi_n) = \sum\limits_{i=1}^{n} D\xi_i + 2\sum\limits_{i<j} \text{cov}(\xi_i, \xi_j) = \sum\limits_{i,j=1}^{n} \text{cov}(\xi_i, \xi_j)$

(g) i. Если ξ и η - независимы, то $\text{cov}(\xi, \eta) = 0$

ii. Если $\text{cov}(\xi, \eta) \neq 0$, то ξ и η - зависимы

iii. Если $\text{cov}(\xi, \eta) = 0$, то неясно

(h) Если $\text{cov}(\xi, \eta) > 0$, то зависимость прямая, если $\text{cov}(\xi, \eta) < 0$, то обратная

Коэффициентом корреляции случайных величин ξ и η с конечными вторыми моментами, называется величина $r_{\xi,\eta} = \dfrac{\text{cov}(\xi, \eta)}{\sqrt{D\xi}\sqrt{D\eta}} = \dfrac{E(\xi\eta) - E\eta}{\sigma_\xi\sigma_\eta}$

Свойства:

(a) $r_{\xi,\eta} = r_{\eta,\xi}$

(b) i. Если ξ и η - независимы, то $r_{\xi,\eta} = 0$

ii. Если $r_{\xi,\eta} \neq 0$, то ξ и η - зависимы

iii. Если $r_{\xi,\eta} = 0$, то неясно

(c) $|r_{\xi,\eta}| \leq 1$

(d) $|r_{\xi,\eta}| = 1 \Longleftrightarrow \eta = a\xi + b$ п.н.

(e) i. Если $r_{\xi,\eta} = 1$, то $\eta = a\xi + b$ и $a > 0$ (прямая линейная зависимость)

ii. Если $r_{\xi,\eta} = -1$, то $\eta = a\xi + b$ и $a < 0$ (обратная линейная зависимость)

Если $r_{\xi,\eta} \neq 0$, то говорят, что случайные величины коррелированы друг с другом. Если $r_{\xi,\eta} > 0$, то имеет прямая корреляция, если $r_{\xi,\eta} < 0$ - обратная

47. *Характеристическая функция случайной величины, ее свойства. Теорема о непрерывном соответствии (формулировка).*

Характеристической функций случайной величины ξ называется функция $\varphi_\xi(t) = Ee^{it\xi}, t \in$

$\mathbb{R}$

Свойства:

(a) Любая случайная величина ξ имеет характеристическую функцию, причем $|\varphi_\xi(t)| \le 1$

(b) Пусть $\varphi_\xi(t)$ - характеристическая функция случайной величины ξ. Тогда характеристическая функция случайной величины $a+b\xi$ равна $\varphi_{a+b\xi}(t) = e^{ita}\varphi_\xi(bt)$

(c) Характеристическая функция суммы независимых случайных величин равна произведению их характеристических функций

(d) Пусть $E\xi^k < \infty$. Тогда $\varphi_\xi(t) = 1 + itE\xi - \dfrac{t^2}{2}E\xi^2 + \dots + \dfrac{(it)^k}{k!}E\xi^k + o(|t|^k)$

(e) Пусть $E\xi^k < \infty$. Тогда $\varphi_\xi^{(k)}(0) = i^k E\xi^k$

(f) Существует взаимно-однозначное соответствие между распределениями и характеристическими функциями. Зная характеристическую функцию можно восстановить распределение.

(g) Теорема о непрерывном соответствии

Th. Последовательность случайных величин $\{\xi_n\}$ слабо сходится к ξ тогда и только тогда, когда соответствующая последовательность характеристических функций сходится поточечно к $\varphi_\xi(t)$

$\{\xi_n\} \rightrightarrows \xi \Longleftrightarrow \varphi_{\xi_n}(t) \longrightarrow \varphi_\xi(t) \forall t \in \mathbb{R}$

48. *Характеристические функции стандартных распределений (Бернулли, биномиальное, Пуассона, нормальное). Следствия.*

Характеристические функции:

- Распределение Бернулли

ξ	0	1
p	$1-p$	p

$$\varphi_\xi(t) = Ee^{i\xi t} = e^{i\cdot 0\cdot t}p(\xi=0) + e^{i\cdot 1\cdot t}p(\xi=1) = 1-p+pe^{it}$$

- Биномиальное распределение

$P(\xi = k) = C_n^k p^k q^{n-k}, \quad k = 0, 1, \dots, n$

Если $t \in B_{n,p}$, то $\xi = \xi_1 + \xi_2 + \xi_3 + \dots + \xi_n$, где $\xi_i \in B_p$ - независимы

$\varphi_\xi(t) = (\varphi_{\xi_n}(t))^n = (1-p+pe^{it})^n$

- Распределение Пуассона

$P(\xi = k) = \dfrac{\lambda^k}{k!}e^{-\lambda}, \quad k = 0, 1, \dots, n$

$\varphi_\xi(t) = e^{\lambda(e^{it}-1)}$

Следствие: распределение Пуассона устойчиво относительно суммирования: $\exists \xi \in \Pi_\lambda, \eta \in \Pi_\mu$, они независимы. Тогда $\xi + \eta \in \Pi_{\lambda+\mu}$

- Стандартное нормальное распределение

$f_\xi(x) = \dfrac{1}{2\pi}e^{-\frac{x^2}{2}}$

$\varphi_\xi(t) = e^{-\frac{t^2}{2}}$

- Нормальное распределение

$\xi \in N(a, \sigma^2)$

$\varphi_\xi(t) = e^{ita}\varphi_\eta(\sigma t) = e^{ita - \frac{\sigma^2 t^2}{2}}$

Следствие: нормальное распределение устойчиво относительно суммирования: если $\xi \in N(a_1, \sigma_1^2), \eta \in N(a_2, \sigma_2^2)$ и они независимы, то $\xi + \eta \in N(a_1 + a_2, \sigma_1^2 + \sigma_2^2)$

49. *Доказательство закона больших чисел Хинчина.*

Закон больших чисел Хинчина

Пусть $\xi_1, \xi_2, \dots, \xi_n$ - последовательность независимых одинаково распределенных случайных величин с конечным матожиданием. Тогда $\dfrac{S_n}{n} = \dfrac{\xi_1 + \dots + \xi_n}{n} \xrightarrow{p} E\xi_1$

Обозначим $a = E\xi_1$

Ранее было доказано, что сходимость по вероятности к константе эквивалентно к слабой сходимости. Поэтому достаточно доказать, что $\dfrac{S_n}{n} \rightrightarrows a$

По теореме о непрерывном соответствии остается доказать, что $\varphi_{\frac{S_n}{n}}(t) \longrightarrow \varphi_a(t) = e^{ita}$

По четвертому свойству $\varphi_{\xi_1}(t) = 1 + itE\xi_1 + o(|t|) = 1 + ita + o(|t|)$

$\varphi_{\frac{S_n}{n}}(t) = [\text{по второму свойству}] = \varphi_{S_n}\left(\dfrac{t}{n}\right) = \left(\varphi_{\xi_1}\left(\dfrac{t}{n}\right)\right)^n = \left(1 + ia\dfrac{t}{n} + o\left(\left|\dfrac{t}{n}\right|\right)\right)^n \underset{\text{по лемме}}{\longrightarrow} e^{ita} = \varphi_a(t)$

50. *Центральная предельная теорема. Вывод из нее предельной теоремы Муавра-Лапласа. Неравенство Берри-Ессеена (формулировка).*

ЦПТ Ляпунова **Th.** Пусть $\xi_1, \dots, \xi_n, \dots$ - последовательность независимых одинаково распределенных случайных величин с конечной дисперсией ($D\xi_1 < \infty$) и $S_n = \sum_{i=1}^{n} \xi_i$. Тогда имеет место слабая сходимость:

$$\frac{S_n - nE\xi_1}{\sqrt{nD\xi_1}} \rightrightarrows N(0, 1)$$

Предельная теорема Муавра-Лапласа: Пусть $v_n(A)$ - число появления события A при n независимых испытаний, p - вероятность успеха при одном испытании, $q = 1 - p$. Тогда $\dfrac{v_n(A) - np}{\sqrt{npq}} \rightrightarrows N(0, 1)$

$v_n(A) = \xi_1 + \xi_2 + \dots + \xi_n = S_n$, где $\xi_i \in B_p$ и независимы, $E\xi_1 = p, D\xi_1 = pq$

По ЦПТ $\dfrac{v_n(A) - np}{\sqrt{npq}} = \dfrac{S_n - nE\xi_1}{\sqrt{nD\xi_1}} \rightrightarrows N(0, 1)$

Неравенство Берри-Эссеена: В условиях ЦПТ для ξ_1 с конечным третьим моментом можно оценить так:

$\left| p\left(\dfrac{S_n - nE\xi_1}{\sqrt{nD\xi_1}} < x\right) - F_0(x)\right| \leq C\dfrac{E|\xi_1 - E\xi_1|^3}{\sqrt{n(D\xi_1)^3}} \quad \forall x \in \mathbb{R}$