Билет №1.

Пространство элементарных событий.

Случайное испытание - действие, результат которого нельзя предсказать заранее, даже если известен комплекс условий проведения испытания.
Событие - результат случайного испытания.
Пространство элементарных событий $\Omega = \{w\}$ - множество всех возможных результатов случайного испытания, где каждый возможный результат эксперимента представлен один и только один раз (т.е. это полная система событий).
Например:
бросание монеты - эксперимент, в результате которой могут произойти только два события: "орел" и "решка". Т.е. $\Omega = \{\omega_1; \omega_2\}$, где ω_1 - выпал орел, ω_2 - выпала решка.
Бросание игральной кости дает 6 возможных исходов, т.е. $\Omega = \{\omega_1; \omega_2; \omega_3; \omega_4; \omega_5; \omega_6\}$.
Но если наш эксперимент - взвешивание рыбы, то все возможные результаты взвешивания будут элементарными событиями данного множества элементарных событий. И количество их в этом пространстве будет определяться точностью весов.
В случае весов с бесконечной точностью $\Omega = [0; +\infty)$. т.к. множество действительных чисел континуально.
И пусть мы не сможем поймать рыбу весом в 1000кг, но включение невозможных событий не противоречит определению пространства элементарных событий.
Если при вылове рыбы мы измеряем вес и длину, то пространство элементарных событий будет состоять из пар действительных чисел. Геометрически это будет плоскость.

Билет №2.

Операции над событиями. Несовместные события. Привести примеры.

Случайное (составное) событие, или просто событие - любое подмножество пространства элементарных событий.
Достоверное событие - включает в себя все элементарные события пространства элементарных событий. Оно про-

исходит в каждом испытании.

Невозможное (пустое) событие – не содержащее ни одного элементарного события. Не происходит никогда.

Событие происходит, когда происходит одно из составляющих его элементарных событий.

Например, при бросании кости достоверным событием Ω будет событие $\Omega = \{\omega_1; \omega_2; \omega_3; \omega_4; \omega_5; \omega_6\}$ (т. е. выпало число от 1 до 6. Невозможным будет выпадение семерки. А если у нас есть событие "выпадение нечетного номера" $\{\omega_1; \omega_3; \omega_5\}$, то при наступлении любого из этих событий (выпадение 1, 3 или 5) означает наступление события "выпал нечетный номер".

Операции над событиями.

1) Сумма (объединение) двух событий A и B – событие $A \cup B$, которое состоит из всех элементарных событий, входящих хотя бы в одно из событий A или B.

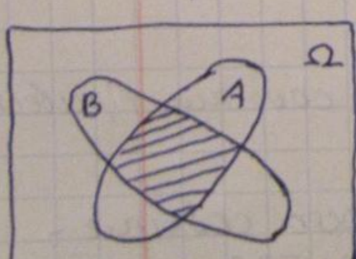
Событие $A \cup B$ происходит при наступлении хотя бы одного из событий A или B.

Геометрическая интерпретация:

Событие $A \cup B$ объединяет все элементарные события, входящие в A и в B, эти элементарные события заштрихованы.

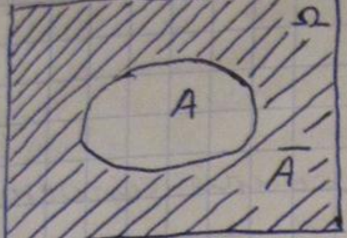
2) Произведение (пересечение) событий A и B называется событие AB, состоящее из всех элементарных событий, входящих одновременно и в A, и в B. Событие AB происходит, когда происходят одновременно и A, и B.

Геометрическая интерпретация:

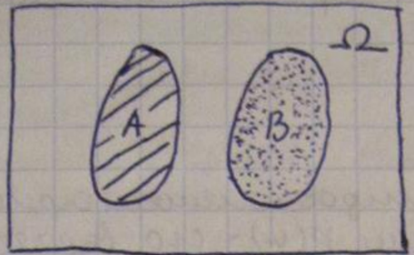


Заштриховано AB.

3) Дополнение к событию A (отрицание события A) – это событие $\overline{A}$, включающее в себя все события, не входящие в A. Событие $\overline{A}$ происходит, когда не происходит A.



Геометрическая интерпретация $\bar{A}$.



Геометрическая интерпретация несовместных событий

Полная система событий

4) События A и B называются несовместные (непересекающиеся), если у них нет общих элементарных событий и их пересечение есть невозможное событие ($AB = \varnothing$).

5) Система событий $A_1; \dots; A_n$ называется полной системой событий, если $A_1 + \dots + A_n = \Omega$, и $A_i A_j = \varnothing$ для всех $i \neq j$. Например, любая пара несовместных событий будет полной системой событий.

Примеры.

- Объединение трех событий A и B и C: н-р, событие, состоящее в том, что выпадет четное число при бросании кости, состоит из событий A - выпала 2, B - выпала 4, C - выпала 6.
- Пересечение событий A и B: это событие "выпала двойка" при A - "выпала четная цифра" и B - "выпала одна из первых двух цифр".
- Дополнением к событию A - "выпала четная цифра" будет событие $\bar{A}$ - "выпала нечетная цифра".
- Несовместными событиями будут A - "выпала двойка" и B - "выпала тройка".
- Полной системой событий будет система следующих событий: A_1 - выпала 1; A_2 - выпала 2; A_3 - выпала 3; ...; A_6 - выпала 6.

Билет №3.

Задание вероятностного пространства.

Пусть есть пространство элементарных событий Ω. Примем за аксиому возможность сопоставить каждому элементарному событию число (вероятность этого элементарного события), такое, что:

1) $\sum\limits_{w \in \Omega} P(w) = 1$;

2) $P(w) \geq 0$;

3) $w \to P(w)$ т.е. каждое элементарное событие сопоставляется некому числу $P(w)$ - его вероятность

4) $P(A) \stackrel{!}{=} \sum\limits_{w \in A} P(w)$ вероятность любого события A равна сумме вероятностей всех элементарных событий w, входящих в A.

Однако такое введение вероятностного пространства годится только для счетных множеств, поскольку для континуальных множеств невозможно придумать такие положительные числа, что, сопоставив их континуальному множеству, мы в итоге получим 1.

Аксиоматическое введение вероятности.

Пусть есть пространство элементарных событий Ω. За аксиому примем возможность сопоставить каждому событию из Ω число (вероятность этого события), такое, что:

1) $P(A) \geq 0$

2) $P(\Omega) = 1$

3) Если $A_i A_j = \varnothing$ (попарно несовместны), $i \neq j$, то $P(A_1 + A_2 + A\ldots) = P(A_1) + P(A2) + P(A\ldots) + \ldots$

Таким образом вводится определение вероятности.

Теорема Бернулли

Вероятность отклонения частоты появления события в n независимых испытаниях от вероятности этого события $P(A)$ на величину, большую, чем заданное число ε, стремится к 0 при n, стремящемся к бесконечности. Причем частота стремится к вероятности по вероятности:

$$\lim_{n \to \infty} P\left\{\left|\frac{m}{n} - P(A)\right| > \varepsilon\right\} = 0,$$ где m - число появления события A в n испытаниях.

Примеры задания вероятностного пространства

Модель равновероятных элементарных событий.

- Число элементарных событий конечно, они считаются равновероятными.
- Вероятность события - отношение числа элементарных событий, составляющих это событие, к числу всех элементар-

ных событий.
Задание вероятностного пространства: необходимо знать число элементарных событий, составляющих данное событие, и общее число событий.

1) Выбор без возвращения, порядок выбора важен

$A_n^m = \frac{n!}{(n-m)!}$ выбор m элементов из n элементов ("размещения").

$A_n^n = n!$ число способов выбора n элементов из n элементов ("перестановки").

2) Выбор без возвращения, порядок выбора не важен.

$C_n^m = \frac{n!}{m!(n-m)!}$ число сочетаний из n элементов по m элементов.

3) Выбор с возвращением, порядок важен

$A_n^m = n^m$ размещения из n элементов по m элементов ($n<m$; $n>m$).

4) Выбор с возвращением, порядок не важен

C_{n+m-1}^m · число сочетаний из n элементов по m элементов.

Можно задать вероятностное пространство геометрически. Рассмотрим область Ω на плоскости площадью S_Ω. Тогда вероятность участка A площадью S_A в этой области: доля его площади в площади всей области:

$$P(A) = \frac{S_A}{S_\Omega}.$$

Такое определение вероятности согласуется с аксиоматическим введением вероятности, поскольку из св-в площадей $P(A) \geq 0$; $P(\Omega) = 1$; $P(A+B) = P(A) + P(B)$, если $AB = \varnothing$.

Билет №4. Свойства вероятности.

Введем аксиоматическое определение вероятности: Пусть есть пространство элементарных событий Ω. За аксиому примем возможность сопоставить каждому событию из Ω некое число - вероятность этого события, такое, что:

1) $P(A) \geq 0$;
2) $P(\Omega) = 1$;
3) Если $A_i A_j = \varnothing$; $i \neq j$, то $P(A_1 + A_2 + \dots) = P(A_1) + P(A_2) + \dots$.

Из этой системы аксиом вытекают следующие свойства вероятности:

1) $0 \leq P(A) \leq 1$
2) $P(\emptyset) = 0$ вероятность невозможного события
3) $P(A \cup B) = P(A) + P(B) - P(AB)$ (A и B - не несовместны)

А так же:

4) $P(\Omega) = 1$ вероятность достоверного события
5) $P(\overline{A}) = 1 - P(A)$ вероятность дополнения к событию A
6) $P(A \cup B) = P(A) + P(B)$, если A и B - несовместны.

Теорема Бернулли.

Пусть m - число появлений события A в n испытаниях. Тогда

$$\forall \varepsilon > 0 \quad \lim_{n \to \infty} P\left\{ \left| \frac{m}{n} - P(A) \right| > \varepsilon \right\} = 0,$$

т. е. частота появления события A (m/n) мало отличается от вероятности события A при большом числе испытаний.

Стремление происходит по вероятности.

Билет № 5. Условная вероятность. Теорема умножения.

Условная вероятность события A при условии, что произошло событие B, называется число $P(A/B)$, такое, что

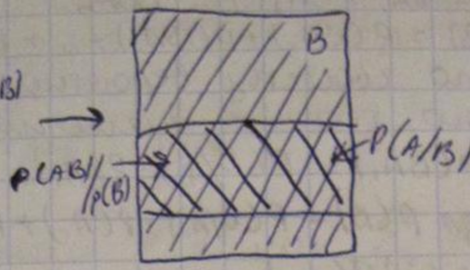
$$P(A/B) = \frac{P(AB)}{P(B)},$$ где $P(AB)$ - вероятность пересечения A и B, $P(B)$ - вероятность B, не равная нулю.

Т.е. мы переходим к новому пространству элементарных событий, которое является частью первоначального и включает только элементарные события, соответствующие событию B. Доля таких элементов в исходном пространстве равна $P(B)$.

Событие A/B - извлечение элементарного события, соответствующего событию A, из совокупности элементарных событий, соответствующих событию B.

Доля таких событий (AB) в исходном пространстве элементарных событий равна $P(AB)$

Доля их в B - $P(AB)/P(B)$.



Пример: пусть при бросании игральной кости выпала нечетная цифра. Какова вероятность, что это 3?
A - событие, состоящее в том, что выпала 3
B - событие, состоящее в том, что выпала нечетная цифра
Пространство элементарных событий при бросании кости состоит из 6 элементарных событий: $\Omega = \{\omega_1; \omega_2; \omega_3; \omega_4; \omega_5; \omega_6\}$
Событие A $= \omega_3$.
Событие B $\omega_1; \omega_3; \omega_5$.
$p(\omega_i) = 1/6$

$$P(B) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2}. \quad P(AB) = \frac{1}{6}.$$

$$P(AB) / P(B) = \frac{1}{6} : \frac{1}{2} = \frac{1}{3}.$$

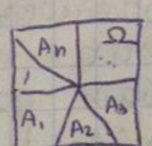
Теорема умножения

$$P(AB) = P(A/B) \cdot P(B).$$

Непосредственно следует из определения условной вероятности.

Билет №6. Формула полной вероятности

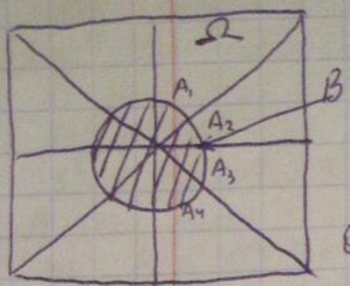
Множество событий $A_1; A_2; A_3; \ldots; A_n$ называется полной системой событий, если:



1) Все они попарно непересекающиеся
$A_i A_j = \varnothing \quad i \neq j$
2) $A_1 + A_2 + \ldots + A_n = \Omega$.

Формула полной вероятности:
Если $A_1; A_2; \ldots; A_n$ - полная система событий, то вероятность любого события B $P(B) = \sum_{i=1}^{n} P(B/A_i) P(A_i)$.

Доказательство.



$B = BA_1 + BA_2 + \ldots + BA_n$

$P(B) = P(BA_1) + P(BA_2) + \ldots + P(BA_n)$

по свойству вероятности

но по теореме умножения

$P(BA_1) = P(B/A_1) \cdot P(A_1)$.

Отсюда $P(B) = P(B/A_1) \cdot P(A_1) + P(B/A_2) P(A_2) + \ldots + P(B/A_n) P(A_n)$.

т.е. $P(B) = \sum\limits_{i=1}^{n} P(B/A_i) P(A_i)$.

Т.е. если даны условные вероятности, а надо узнать безусловные - используется формула полной вероятности.

Пример.

Есть три ружья, вероятность попадания в цель первого ружья 0,5, второго 0,7, третьего 0,8. Наугад берется одно ружье, и делается один выстрел. Найти вероятность попадания.

A_1 - событие, состоящее в том, что выбрано 1-е ружье

A_2 - ... 2-е ружье

A_3 - ... 3-е ружье

B - попадание в цель.

$P(B/A_1) = 0,5$;

$P(B/A_2) = 0,7$;

$P(B/A_3) = 0,8$

Даны условные вероятности

$P(B) = \sum P(B/A_n) A_n$

$P(A_1) = 1/3$

$P(A_2) = 1/3$

$P(A_3) = 1/3$

События $A_1; A_2; A_3$ составляют полную систему событий

$P(B) = 0,5 \cdot 1/3 + 0,7 \cdot 1/3 + 0,8 \cdot 1/3 = \frac{2}{3}$.

Билет №7. Формула Байеса.

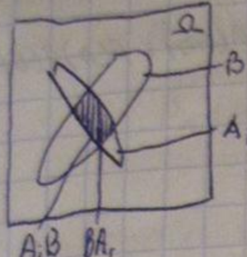
Пусть $A_1; A_2; \ldots; A_n$ - полная система событий. Чтобы найти вероятность любого из этих событий $P(A_i)$ при условии, что произошло некое событие B, необходимо использовать формулу Байеса:

$$P(A_i/B) = \frac{P(B/A_i) P(A_i)}{\sum\limits_{j=1}^{n} P(B/A_j) P(A_j)}.$$

Доказательство.

По теореме умножения найдем $P(A_i/B)$.

$$P(A_i/B) = \frac{P(A_i B)}{P(B)} = \frac{P(BA_i)}{P(B)}.$$



Мы имеем право записать, что $P(A_iB)=P(BA_i)$.
Но $P(BA_i)=P(B/A_i)P(A_i)$ по теореме умножения, тогда

$$P(A_i/B)=\frac{P(B/A_i)P(A_i)}{P(B)}.$$

Запишем $P(B)$ как $\sum_{i=1}^{n} P(B/A_i)P(A_i)$ по ф-ле полной вероятности.

Тогда $P(A_i/B)=\dfrac{P(B/A_i)P(A_i)}{\sum_{i=1}^{n} P(B/A_i)P(A_i)}$.

Применяется формула Байеса, если даны безусловные вероятности, а надо найти условные.

Пример.
Есть 3 ружья, вероятность попадания в цель первого ружья 0,5, второго - 0,7, третьего 0,8. Наугад было взято одно ружье, и сделан выстрел, поразивший цель. Какова вероятность, что выстрел производился из 1-го ружья?

A_1 - событие, состоящее в том, что выбрано 1-е ружье
A_2 - ... 2-е ружье
A_3 - ... 3-е ружье
B - попадание в цель.

$P(B/A_1)=0,5$ $\quad P(A_1)=1/3$ $\quad P(A_1/B)$ - ?
$P(B/A_2)=0,7$ $\quad P(A_2)=1/3$ События $A_1; A_2; A_3$ составляют
$P(B/A_3)=0,8$. $\quad P(A_3)=1/3$ полную систему событий.

$$P(A_1/B)=\frac{0,5\cdot 1/3}{0,5\cdot 1/3+0,7\cdot 1/3+0,8\cdot 1/3}=\frac{1}{6}:\frac{2}{3}=\frac{1}{6}\cdot\frac{3}{2}=\frac{1}{4}.$$

Билет №8. Независимые события

Два события называются независимыми, если вероятность их пересечения равна произведению их вероятностей:

$$P(AB)=P(A)\cdot P(B).$$

Другой вариант определения:
Два события A и B называются независимыми, если происхождение события B не зависит от происхождения события A: $P(A/B)=P(A)$.
Эти утверждения равносильны.

$P(A/B) = P(A)$; по т. умножения $\frac{P(AB)}{P(B)} = P(A)$; $P(AB) = P(A)P(B)$

И наоборот:

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

$P(A/B)P(B) = P(A)P(B)$ по т. умножения,

$P(A/B) = P(A)$.

Примеры.

1) Зависимы ли события "наличие доминантного признака" (AA; Aa; aA) и гомозиготность (AA, aa)?

Вероятность 1-го события $3/4$.

Вероятность 2-го $1/2$.

Если они независимы, то $P(AB) = P(A)P(B)$.

Но $P(AB) = 1/4$; а $P(A)P(B) = 3/4 \cdot 1/2 = 3/8$.

События зависимы.

2) Зависимы ли события "Рождение девочки-первенца" и "Рождение второго ребенка мальчика?"

В семье 2 ребенка. Возможные распределения в порядке рождения: ДД, ДМ, МД, ММ.

А - родилась девочка первой ДД, ДМ $P(A) = 1/2$

В - родился мальчик вторым ДМ, ММ $P(B) = 1/2$.

$P(AB)$ - ДМ - $1/4$.

$P(AB) = P(A) \cdot P(B)$. События независимы.

3) Могут ли быть 2 несовместных события с положительной вероятностью быть независимыми?

Несовместные события: $AB = \emptyset$;

$P(AB) = 0$.

$P(AB) = P(A) \cdot P(B)$: предположим, что они независимы.

$P(A) \cdot P(B) = 0$;

тогда или $P(A) = 0$, или $P(B) = 0$, или $P(A) = P(B) = 0$.

Но по условию $P(A) > 0$ и $P(B) > 0$.

Поэтому 2 несовместных события с положительной вероятностью не могут быть независимыми.

Билет №9. Формула Бернулли

Вероятность того, что в n независимых испытаниях некоторое событие, имеющее вероятность появления в каждом испытании P, наступит ровно k раз, равна:

$P_n(k) = C_n^k p^k q^{n-k}$, где $q = 1-p$

$P_n(k) = C_n^k p^k (1-p)^{n-k}$. $\quad C_n^k = \dfrac{n!}{k!\,(n-k)!}$

Определить вероятность выпадения ровно одного герба при пяти бросаниях идеальной монеты.

$n=5$
$p = 1/2$
$k = 1$

$$P_5(1) = C_5^1 \cdot \frac{1}{2} \cdot \left(1-\frac{1}{2}\right)^{5-1} = 5 \cdot \frac{1}{2} \cdot \frac{1}{2^4} = \frac{5}{32}$$

Вероятность наступления события не менее k_1 раз и не более k_2 раз ($k_1 \le k_2$) равна

$$P_n(k_1;k_2) = \sum_{i=k_1}^{k_2} P_n(i).$$

Билет №10. Случайная величина. Привести 3 примера.

Одномерная случайная величина ξ - любая числовая функция, определенная на пространстве элементарных событий Ω.

Любая числовая функция $\varphi(\xi)$ от случайной величины тоже является случайной величиной.

Случайная величина бывает дискретная и непрерывная.

Дискретная задается таблицей.

Пример.

Пусть есть монета, которую мы бросаем 2 раза. При этом возникает пространство элементарных событий $\Omega = \{\omega_1; \omega_2; \omega_3; \omega_4\}$: ГГ; ГР; РГ; РР. $P(\omega_i) = 0,25$.

Найдем случайную величину ξ, равную числу гербов, возникающих при 2-х бросаниях:

ξ:

0	1	2	← значение случайной величины (количество выпавших гербов)
$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	вероятность каждого из значений этой случайной величины.

Еще пример.

Пусть имеется очень большое число некоторых животных N, среди них доля альбиносов 0,05. Из общего количества животных случайно выбираются любые 40. Пространство элементарных событий этого эксперимента $\Omega = \{ C_N^{40} \}$. Найдем случайную величину ξ - число альбиносов из выбранных 40 животных.

ξ:	1	2	...	39	40
	$P_{40}(1)$	$P_{40}(2)$	...	$P_{40}(39)$	$P_{40}(40)$

Значения ξ - вероятность появления данного числа альбиносов при выборе 40 животных.

$n = 40$

$p = 0{,}05$ $\quad P_{40}(i) = C_{40}^{i} (0{,}05)^{i} (1 - 0{,}05)^{40-i}$.

$k = i$

Еще пример.
Мы взвешиваем рыбу на весах бесконечной точности. Пространство элементарных событий этого эксперимента $\Omega = [0; +\infty)$. Значения ξ - значения веса рыбы. При этом ξ является непрерывной случайной величиной.

Билет №11. Функция распределения и ее св-ва.

Функция распределения случайной величины ξ - функция $F(x)$, равная для любого x (x - значение вероятности события $\xi \le x$):

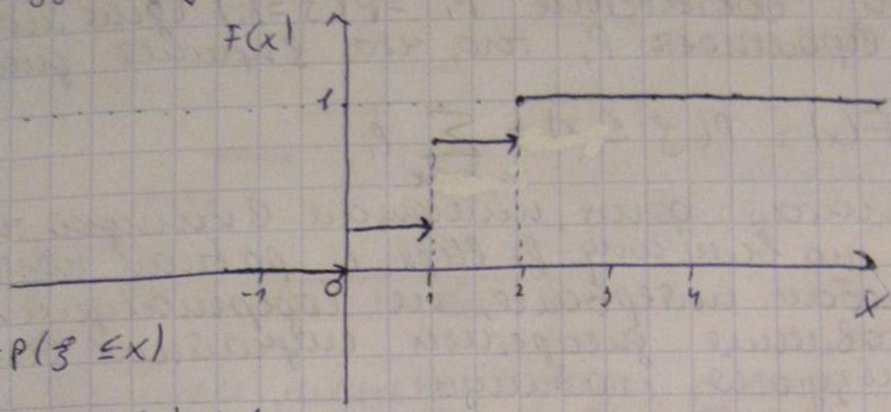
$$F(x) = P(\xi \le x)$$

Свойства функции распределения.

1) $0 \le F(x) \le 1$
2) $F(x)$ - неубывающая функция
   Если $x_1 \le x_2$, то $F(x_1) \le F(x_2)$.
3) $\lim\limits_{x \to -\infty} F(x) = 0$; 4) $\lim\limits_{x \to +\infty} F(x) = 1$;
5) $P(x_1 < \xi \le x_2) = F(x_2) - F(x_1)$.
   поскольку $P(\xi \le x_2) = F(x_2)$;
   $P(\xi \le x_1) = F(x_1)$.

Эти все свойства выводятся из определения функции распределения.

Билет №12. График функции распределения для какой-нибудь случайной величины



$F(x) = P(\xi \le x)$

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

Область значений от 0 до 1

2) $F(x)$ - неубывающая ф-ия

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

Две асимптоты: $y = 0$ и $y = 1$.

5) $P(x_1 < \xi \le x_2) = F(x_2) - F(x_1)$.

Пусть мы бросаем монету 2 раза. Пространство элементарных событий состоит из 4 событий; $P(\omega_i) = 0{,}25$!

ГГ, ГР; РГ; РР.

Если ξ - число гербов, то имеем:

ξ	0	1	2
P	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

при $x = -1$ $P(\xi \le -1) = 0$

при $x = 0$ $P(\xi \le 0) = 1/4$.

при $x = 1$ $P(\xi \le 1) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$

при $x = 2$ $P(\xi \le 2) = \frac{1}{4} + \frac{3}{4} = \frac{4}{4} = 1$

при $x = 3$ $P(\xi \le 3) = 1$.

Билет №13. Дискретная случайная величина.

Случайная величина, принимающая конечное или счетное число значений.

Обозначим множество всех возможных значений дискретной случайной величины ξ через $x_0; x_1; x_2; \ldots;$ а вероятности, с которыми ξ принимает эти значения через $p_0; p_1; p_2; \ldots$

Тогда $\sum_i p_i = 1$.

Если ввести соответствие $p_i = P(\xi = x_i)$ (для любого i указать вероятность p_i того, что ξ примет значение x_i), то

$$F(x) = P(\xi \le x) = \sum_{x_i \le x} p_i.$$

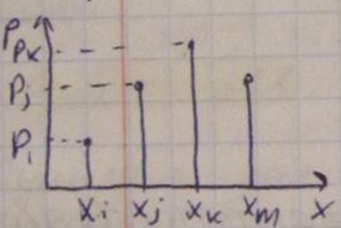
$F(x)$ – ступенчатая ф-ия, имеющая в каждой точке x_i скачок на величину p_i вверх, и равная постоянной на любом интервале, не содержащем x_i.

Для представления дискретных случайных величин используются таблицы:

ξ:

x_0	x_1	…
p_0	p_1	…

и графики:



(графики функций распределения вероятностей дискретной случайной величины – зависимости $p_i = P(\xi = x_i)$ от значения случайной величины x_i).

Чтобы задать дискретную случайную величину ξ, необходимо описать множество всех возможных ~~их~~ значений случайной величины $x_0; x_1; x_2; \ldots$, и указать числа p_i, такие, что

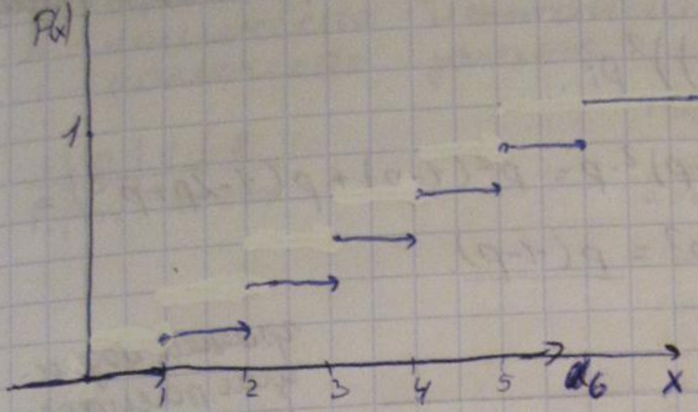
1) $p_i \ge 0$; $i = 0, 1, 2 \ldots$
2) $\sum_i p_i = 1$.

Примеры.

1) Мы бросаем игральную кость 1 раз.
Множество всех возможных значений случайной величины ξ имеет значения 1, 2, 3, 4, 5, 6.
$P(1) = P(2) = \ldots = P(6) = 1/6$.
$\sum_i p_i = 1$.

x_ξ	1	2	3	4	5	6
P	1/6	1/6	1/6	1/6	1/6	1/6

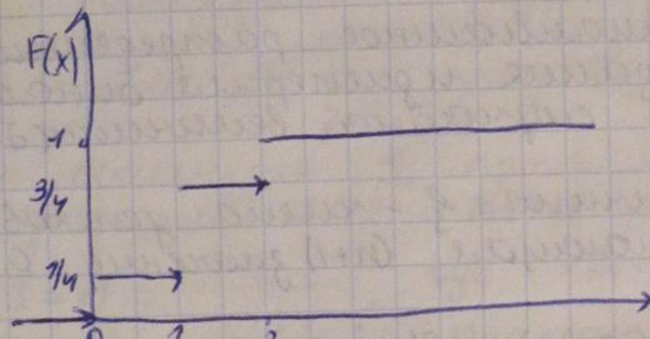
График функции распре-деления этой случай-ной величины.



2) Мы бросаем монету 2 раза; пространство элемен-тарных событий состоит из 4 событий: ГГ, ГР, РГ, РР. Пусть ξ - число выпавших гербов. Возможные значения ξ: 0; 1; 2. $P(0) = 1/4$; $P(1) = 1/2$; $P(2) = 1/4$.

ξ	0	1	2
p	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

График функции распре-деления



Билет №14. Распределение Бернулли. Математическое ожидание и дисперсия слу-чайной величины, распределенной по закону Бернулли.

Случайная величина, распределенная по Бернулли:

$$\xi: \begin{array}{c|c} 0 & 1 \\ \hline 1-p & p \end{array} \quad (0<p<1).$$

p - параметр случайной ве-личины.

Распределение Бернулли описывает процессы, име-ющие 2 исхода: бросание монеты, наличие или от-сутствие признака, пол детеныша и пр.

Математическое ожидание дискретной случай-ной величины вычисляется по формуле:

$$M(\xi) = \sum_i x_i p_i \qquad M(\xi) = 0(1-p) + 1\cdot p = p.$$

Дисперсия случайной величины (дискретной) вычисляется по формуле:

$$D(\xi) = \sum_i (x_i - M(\xi))^2 p_i .$$

$$D(\xi) = (0-p)^2\cdot(1-p) + (1-p)^2\cdot p = p^2(1-p) + p(1-2p+p^2) =$$

$$= p^2 - p^3 + p - 2p^2 + p^3 = p - p^2 = \underline{p(1-p)}.$$

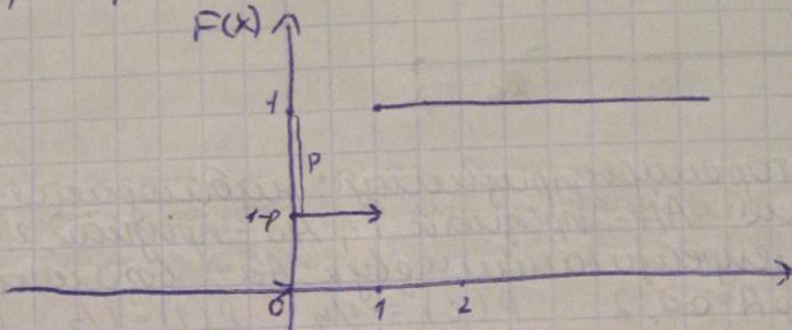


график функции распределения случайной величины, распределенной по Бернулли.

Билет № 15. Биномиальное распределение. Математическое ожидание и дисперсия биномиально распределенной случайной величины (без док-ва).

Случайная величина ξ - число успехов в n испытаниях, принимающая (n+1) значение 0, 1, 2, 3..., n с вероятностями

$$p_i = P(\xi = i) = C_n^i p^i q^{n-i}$$

или $p_i = C_n^i p^i (1-p)^{n-i}$; при $0<p<1$,

где $i = 0, 1, 2, \dots, n$.

n и p - параметры распределения

p_i - вероятность того, что в n испытаниях событие наступит i раз

при n = 1 биномиально распределенная случайная величина превращается в случайную величину, распределенную по Бернулли.

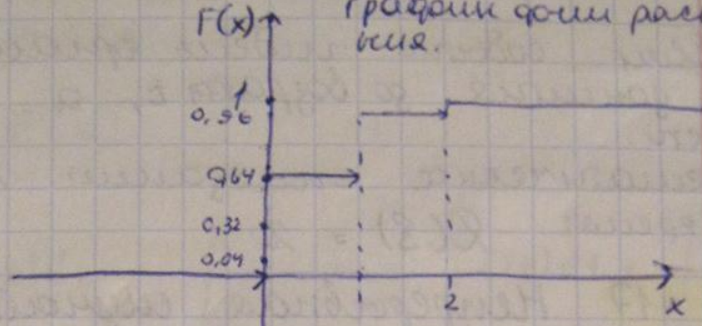
Пример случайной величины, распределенной по биномиальному закону.

Пусть у нас $n=2$, $p=0,2$.

$$p = C_n^i p^i (1-p)^{n-i}.$$

p - вероятность успеха в единичном испытании
n - количество испытаний
i - количество успехов.

i	0	1	2
p_i	0,64	0,32	0,04



Математическое ожидание

$$M(\xi) = np$$

Дисперсия: $D(\xi) = np(1-p)$.

Билет №16. Распределение Пуассона. Математическое ожидание и дисперсия случайной величины, распределенной по Пуассону (без док-ва)

Случайная величина ξ, принимающая счетное множество значений $0, 1, 2, \ldots$ с вероятностями

$$P_i = P(\xi = i) = \frac{\lambda^i e^{-\lambda}}{i!}, \text{ где } i = 0, 1, 2, \ldots, \lambda > 0.$$

λ - параметр распределения Пуассона

ξ:	0	1	2	...	$+\infty$
	p_0	p_1	p_2	...	...

Является предельным случаем биномиальной случайной величины при $n \to \infty$, $p \to 0$, $np = const = \lambda$.
Является моделью числа появления некоторого события за единицу времени, численность чего-либо в единице площади и др. Хорошо описывает ситуацию случайно и независимо друг от друга появляющихся событий в течение заданного периода времени.
Если λ относится к единице времени, то периоду времени t будет соответствовать распределение с параметром λt.

Вероятность того, что в течение периода t не произойдет ни одного события

$$P_0 = P(\xi = 0) = e^{-\lambda t}$$

Если событие - гибель организма, то P_0 - вероятность дожития до возраста t, а λ - интенсивность смертности.

Математическое ожидание: $M(\xi) = \lambda$.

Дисперсия: $D(\xi) = \lambda$.

Билет №17. Непрерывная случайная величина. Свойства ф-ии плотности.

Случайная величина ξ, для которой функция распределения $F(x)$ непрерывна, и существует функция $f(x)$ (функция плотности), такая, что

$$F(x) = \int_{-\infty}^{x} f(t)dt.$$

$f(x)$ непрерывна почти всюду, за исключением множеств, не имеющих предельных конечных точек.

Свойства ф-ии плотности.

1. $f(x) \geq 0$

Поскольку $0 \leq F(x) \leq 1$, т.е. $F(x) \geq 0$, то

$$\int_{-\infty}^{x} f(t)dt \geq 0, \text{ а значит, } f(t) \geq 0.$$

2. $$\int_{-\infty}^{+\infty} f(x)dx = 1.$$

По св-ву ф-ии распределения $0 \leq \int_{-\infty}^{x} f(t)dt \leq 1$.

Но $F(x) = \int_{-\infty}^{x} f(t)dt$ неубывающая ф-ия, поэтому

$$\lim_{x \to +\infty} \int_{-\infty}^{x} f(t)dt = 1; \quad \int_{-\infty}^{+\infty} f(t)dt = 1.$$

3. $$P(a < \xi \leq b) = \int_{a}^{b} f(x)dx.$$

по св-ву ф-ии распределения

$P(x_1 < \xi \le x_2) = F(x_2) - F(x_1)$; т.е.

$P(a < \xi \le b) = F(b) - F(a)$.

$F(b) = \int_{-\infty}^{b} f(t)dt$; $\quad F(a) = \int_{-\infty}^{a} f(t)dt$.

$$F(b) - F(a) = \int_{-\infty}^{b} f(t)dt - \int_{-\infty}^{a} f(t)dt = \int_{-\infty}^{b} f(t)dt + \int_{a}^{-\infty} f(t)dt =$$

(по теореме о разбиении отрезка интегрирования)

$$= \int_{a}^{b} f(t)dt.$$

Функция плотности описывает непрерывную случайную величину.

график функции плотности непр. случайной величины

Билет №18. Непрерывное равномерное распределение. Математическое ожидание и дисперсия случайной величины, распределенной по равномерному закону на интервале (a, b).

Случайная величина ξ называется равномерно распределенной на отрезке $[a, b]$ (a и b - параметры распределения), если ее ф-ия плотности имеет вид:

$$f(x) = \begin{cases} \frac{1}{b-a}, \text{ если } a \le x \le b \\ 0 \text{ - иначе.} \end{cases}$$

Вероятность того, что равномерно распределенная случайная величина ξ примет значение из отрезка $[x_1; x_2]$, лежащего целиком внутри (a, b), не зависит от расположения этого отрезка и равна

$$\frac{(x_2 - x_1)}{b - a}.$$

Математическое ожидание

$$M(\xi) = \int_{-\infty}^{+\infty} x f(x)\, dx$$

$$M(\xi) = \int_{-\infty}^{+\infty} x \cdot f(x)\, dx = \int_{-\infty}^{a} x f(x)\, dx + \int_{a}^{b} x f(x)\, dx + \int_{b}^{+\infty} x f(x)\, dx =$$

$$= 0 + \int_{a}^{b} \frac{x}{b-a}\, dx + 0 = \frac{1}{b-a} \frac{x^2}{2}\Big|_a^b = \frac{1}{b-a}\cdot\left(\frac{b^2}{2} - \frac{a^2}{2}\right) = \frac{b^2 - a^2}{(b-a)2} =$$

$$= \frac{(b-a)(b+a)}{2(b-a)} = \frac{b+a}{2}.$$

$$D(\xi) = \int_{-\infty}^{+\infty} (x - M(\xi))^2 f(x)\, dx = \int_{-\infty}^{a} (x - M(\xi))^2 f(x)\, dx + \int_{a}^{b} (x - M(\xi))^2 f(x)\, dx +$$

$$+ \int_{b}^{+\infty} (x - M(\xi))^2 f(x)\, dx = \int_{a}^{b} \left(x - \frac{b+a}{2}\right)^2 \cdot \frac{1}{b-a}\, dx = \int_{a}^{b} \left(x^2 - \frac{2x(b+a)}{2} + \frac{(a+b)^2}{4}\right)\cdot$$

$$\cdot \frac{1}{b-a}\, dx = \int_{a}^{b} \frac{x^2}{b-a}\, dx - \int_{a}^{b} \frac{x(a+b)}{b-a}\, dx + \int_{a}^{b} \frac{(a+b)^2}{4(b-a)}\, dx =$$

$$= \frac{1}{b-a} \cdot \frac{x^3}{3}\Big|_a^b - \frac{a+b}{b-a} \cdot \frac{x^2}{2}\Big|_a^b + \frac{(a+b)^2}{4(b-a)}\, x\Big|_a^b =$$

$$= \frac{1}{b-a}\left(\frac{b^3}{3} - \frac{a^3}{3}\right) - \frac{a+b}{b-a}\left(\frac{b^2}{2} - \frac{a^2}{2}\right) + \frac{(a+b)^2}{4(b-a)}(b-a) =$$

$$= \frac{b^3 - a^3}{3(b-a)} - \frac{(a+b)(b^2 - a^2)}{2(b-a)} + \frac{(a+b)^2}{4} =$$

$$= \frac{(b-a)(b^2+ab+a^2)}{3(b-a)} - \frac{(a+b)^2}{2} + \frac{(a+b)^2}{4} =$$

$$= \frac{b^2+ab+a^2}{3} - \frac{(a+b)^2}{2} + \frac{(a+b)^2}{4} = \frac{4b^2+4ab+4a^2-6(a+b)^2+3(a+b)^2}{12} =$$

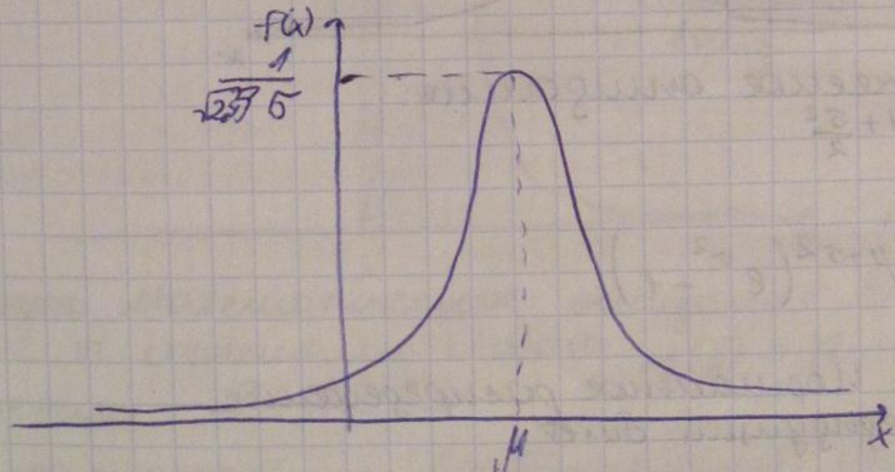
$$= \frac{4b^2+4ab+4a^2-3(a^2+2ab+b^2)}{12} = \frac{4b^2+4ab+4a^2-3a^2-6ab-3b^2}{12} =$$

$$= \frac{b^2-2ab+a^2}{12} = \frac{(b-a)^2}{12}.$$

Билет №19. Лог-нормальное распределение

Для начала рассмотрим нормальное распределение. Случайная величина ξ называется нормально распределенной с параметрами μ и σ^2, если ф-ция плотности $f(x)$ имеет вид:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < +\infty, \quad -\infty < \mu < +\infty, \quad \sigma > 0.$$



Свойства графика функции плотности

1) Максимум в точке $\left(\mu; \frac{1}{\sqrt{2\pi}\sigma}\right)$.
2) Симметрия относительно $x = \mu$
3) При $x \to \infty$ $f(x) \to 0$, при $x \to -\infty$ $f(x) \to 0$.
4) при $\mu \uparrow$ график смещается вправо
при $\sigma \uparrow$ график сплющивается (толще и ниже).

при $\sigma\downarrow$ график тоньше и выше.

5) Площадь графика под интервалом $(\mu-k\sigma; \mu+k\sigma)$ не зависит от μ и σ.

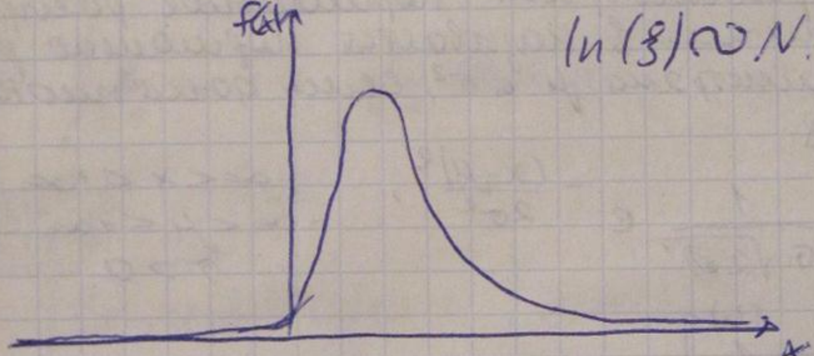
Математическое ожидание:

$$M(\xi) = \mu$$

Дисперсия: $\mathcal{D}(\xi) = \sigma^2$.

Теперь рассмотрим лог-нормальное распределение. Непрерывная случайная величина ξ называется распределенной по логнормальному закону с параметрами μ и σ^2, если случайная величина $\eta = \ln\xi$ распределена нормально.

График функции плотности:



$\ln(\xi) \sim N$.

Математическое ожидание:

$$M(\xi) = e^{\mu+\frac{\sigma^2}{2}}$$

Дисперсия:

$$\mathcal{D}(\xi) = e^{2\mu+\sigma^2}\left(e^{\sigma^2}-1\right).$$

Билет №20. Нормальное распределение.

см. предыдущий билет

Если $\xi \sim N(\mu; \sigma^2)$, то $\frac{\xi-\mu}{\sigma} \sim N(0;1)$.

т.е. $\mu=0$, $\sigma^2=1$

Это стандартное нормальное распределение случайной величины.

Многие случайные величины подчиняются законам нормального распределения.

Центральная предельная теорема
Если есть n независимых случайных величин, то их сумма при $n \to \infty$ стремится к нормальному распределению:

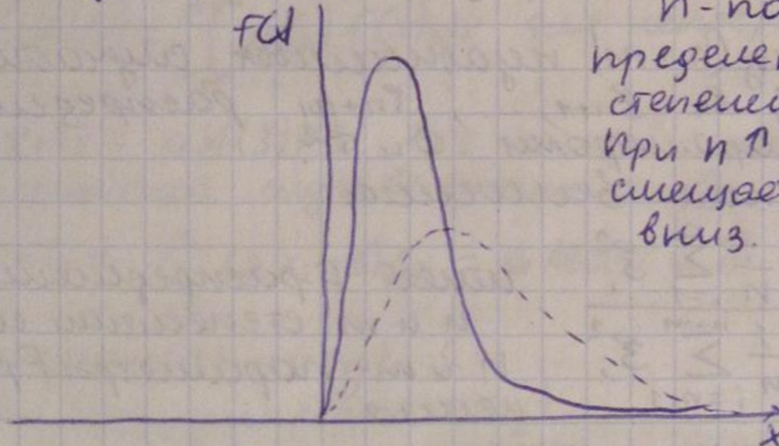
$$\frac{\xi_1 + \xi_2 + \ldots + \xi_n}{C_n} \xrightarrow{n \to \infty} N.$$

Билет №21. χ^2-распределение, t-распределение, F-распределение.

① χ^2-распределение.
Пусть каждая из n независимых случайных величин $\xi_1; \xi_2; \ldots; \xi_n$ распределена нормально с параметрами 0 и 1. Тогда случайная величина

$$y = \xi_1^2 + \xi_2^2 + \ldots + \xi_n^2$$ распределена по закону χ^2.

График функции плотности:



n - параметр χ^2-распределения (число степеней свободы).
При $n \uparrow$ пик графика смещается вправо и вниз.

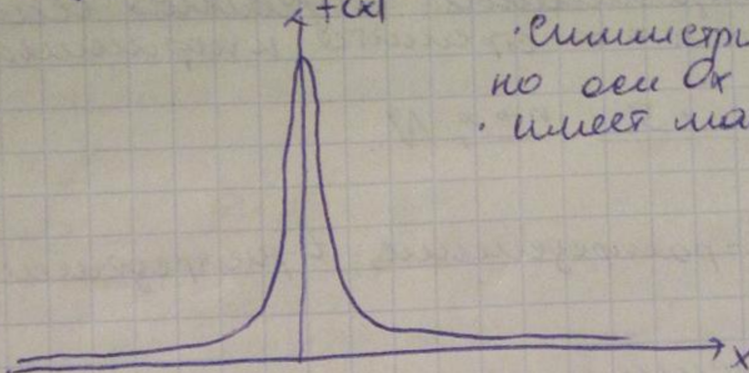
~~Дисперсия~~ Математическое ожидание χ^2-распределения с n степенями свободы $M(\xi) = n$
Дисперсия - $D(\xi) = 2n$

② t-распределение (распределение Стьюдента)

Пусть каждая из $(n+1)$ независимых случайных величин $\xi_0; \xi_1; \ldots; \xi_n$ распределена нормально с параметрами 0 и σ^2. Тогда случайная величина

$$y = \frac{\xi_0}{\sqrt{\frac{1}{n}\sum_{i=1}^{n} \xi_i^2}}$$ имеет распределение Стьюдента с n степенями свободы.

График функции плотности:



· Симметричен относительно оси Ox
· имеет максимум в точке 0

Математическое ожидание t-распределения с n степенями свободы $M(\xi) = 0$;
дисперсия $D(\xi) = \frac{n}{n-2}$.

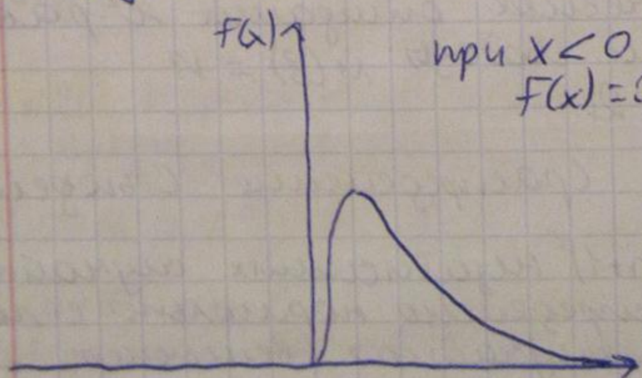
③ F-распределение (распределение Фишера).

Пусть каждая из $(n+m)$ независимых случайных величин $\xi_1; \xi_2; \dots; \xi_n; \xi_{n+1}; \dots; \xi_{n+m}$ распределена нормально с параметрами 0 и σ^2.
Тогда случайная величина

$$\eta = \frac{\frac{1}{n}\sum_{i=1}^{n}\xi_i^2}{\frac{1}{m}\sum_{i=n+1}^{n+m}\xi_i^2}$$

имеет F-распределение с n и m степенями свободы
n и m – параметры F-распределения.

График функции плотности:



при $x<0$
$f(x)=0$.

Математическое ожидание F-распределения с n и m степенями свободы

$$M(\xi) = \frac{m}{m-2}.$$

Дисперсия

$$D(\xi) = \frac{2m^2(n+m-2)}{n(m-2)^2(m-4)}$$

Билет №22. Математическое ожидание случайной величины и его свойства.

Математическое ожидание случайной величины ξ – число $M(\xi)$:

$$M(\xi) = \sum x_i p_i \text{ для дискретной } \xi.$$

$$M(\xi) = \int_{-\infty}^{+\infty} x f(x)\,dx \text{ для непрерывной } \xi.$$

Является характеристикой положения центра распределения.

Свойства математического ожидания

1) $M(C) = C$

2) Если $\varphi(\xi)$ – некоторая функция от случайной величины, то

$$M(\varphi(\xi)) = \sum_i \varphi(x_i) p_i \text{ для дискретной } \xi;$$

$$M(\varphi(\xi)) = \int_{-\infty}^{+\infty} \varphi(x) f(x)\,dx \text{ для непрерывной.}$$

3) $M(a\xi + b) = aM(\xi) + b$, где a и b – константы.

Для дискретной случ. величины док-во:

$$M(a\xi + b) = \sum_{i=1}^{k} (a x_i + b) p_i = a \sum_{i=1}^{k} x_i p_i + b \sum_{i=1}^{k} p_i = aM(\xi) + b.$$

4) $M(\xi + \eta) = M(\xi) + M(\eta)$.

Док-во. (для дискретной ξ и η).

Пусть $p_{ij} = P(\xi = x_i;\ \eta = y_j)$; отсюда $p_i = P(\xi = x_i)$; $p_j = P(\eta = y_j)$, тогда

$$\sum_{j=1} p_{ij} = p_i\ ; \quad \sum_{i=1} p_{ij} = p_j$$

Тогда $M(\xi + \eta) = \sum_{i=1} \sum_{j=1} (x_i + y_j) p_{ij} = \sum_{i=1} x_i \sum_{j=1} p_{ij} + \sum_{j=1} y_j \sum_{i=1} p_{ij} =$

$$= \sum_{i=1} x_i p_i + \sum_{j=1} y_j p_j = M(\xi) + M(\eta).$$

5) Если ξ и η - независимые случайные величины, то $M(\xi\eta) = M(\xi)\cdot M(\eta)$.

Док-во:

$$M(\xi\eta) = \sum_{i=1}\sum_{j=1} x_i y_j p_{ij} = \sum_{i=1}\sum_{j=1} x_i y_j p_i p_j =$$

$$= \sum_{i=1} x_i p_i \sum_{j=1} y_j p_j = M(\xi) M(\eta).$$

Отсюда еще определение:
Две случайные величины ξ_1 и ξ_2 называются независимыми, если $\forall (a, b, c, d)$

$$P[a \le \xi_1 \le b,\ c \le \xi_2 \le d] = P[a \le \xi_1 \le b]\cdot P[c \le \xi_2 \le d].$$

Билет № 23 Дисперсия случайной величины и ее свойства.

Дисперсия случайной величины $D(\xi)$ - число, равное:

$$D(\xi) = \sum_i (x_i - M(\xi))^2 p_i \quad \text{дискр. } \xi$$

$$D(\xi) = \int_{-\infty}^{+\infty} (x - M(\xi))^2 f(x)dx \quad \text{непрерывн. } \xi$$

Характеризует средний квадрат отклонения случайной величины от своего математического ожидания.

Свойства дисперсии

1. $D(\xi) = M[\xi - M(\xi)]^2 = M(\xi^2) - [M(\xi)]^2$

Док-во:

$$D(\xi) = \sum_i (x_i - M(\xi))^2 \cdot p_i = \underset{\text{св-во мат. ожид.}}{M(\xi - M(\xi))^2} = M(\xi^2) - 2[M(\xi)]^2 +$$

$$+ [M(\xi)]^2 = M(\xi^2) - [M(\xi)]^2.$$

поскольку $M(2\xi M(\xi)) = 2M(\xi)\cdot M(\xi)$.

2. $D(C) = 0$

т.к. $M(C) = C$, и $x_i = C$.

3) $D(\xi) \ge 0$, т.к. дисперсия - сумма неотрицательных слагаемых или интеграл от неотрицательной ф-ии по определению.

4) $D(a\xi+b) = a^2 D(\xi)$, где a и b - const.

$D(a\xi+b) = M[a\xi+b-M(a\xi+b)]^2 = M[a\xi+b-aM(\xi)-b]^2 =$

$= M[a\xi-aM(\xi)]^2 = a^2 M[\xi-M(\xi)]^2 = a^2 D(\xi)$.

5) Если ξ и η - независимые случайные величины, то

$D(\xi\pm\eta) = D(\xi)+D(\eta)$.

$D(\xi\pm\eta) = M(\xi\pm\eta)^2 - [M(\xi\pm\eta)]^2 = M(\xi)^2 \pm 2M(\xi\eta) + M(\eta)^2 -$

$- [M(\xi)]^2 \mp 2M(\xi\eta) - [M(\eta)]^2 = M(\xi)^2 - [M(\xi)]^2 + M(\eta)^2 - [M(\eta)]^2 =$

$= D(\xi)+D(\eta)$.

Неравенство Чебышева

Позволяет оценивать вероятность отклонения ξ от своего $M(\xi)$ при известной $D(\xi)$:

$$P(|\xi-M(\xi)| \geq \varepsilon) \leq \frac{D(\xi)}{\varepsilon^2}$$

Вывод формулы для дискретной случайной величины.

Пусть ξ имеет конечную $D(\xi) = \sum_i [x_i - M(\xi)]^2 p_i$ и $\varepsilon > 0$.

Если из суммы выбросить все $|x_i - M(\xi)| < \varepsilon$, а для тех, для кого $|x_i - M(\xi)| \geq \varepsilon$, заменить $[x_i - M(\xi)]^2$ их минимальным значением ε^2, то

$$D(\xi) \geq \varepsilon^2 \sum_j p_j \quad ; \quad j = |x_i - M(\xi)| \geq \varepsilon.$$

Но $P(|\xi-M(\xi)| \geq \varepsilon) = \sum_j p_j$, $j = |x_i - M(\xi)| \geq \varepsilon$.

т.е. $\frac{D(\xi)}{\varepsilon^2} = P(|\xi-M(\xi)| \geq \varepsilon)$.

Билет №24. Среднеквадратичное отклонение и коэффициент вариации случайной величины.

Среднее квадратичное отклонение - корень квадратный из дисперсии: $\sigma(\xi) = \sqrt{D(\xi)}$.

Является, как и дисперсия, мерой рассеяния распределения, но измеряется не в квадратных единицах, а в тех единицах, которые использовались для измерения случайной величины.

Коэффициент вариации случайной величины ξ - число $V(\xi)$, такое, что

$$V(\xi) = \frac{\sigma(\xi)}{M(\xi)} \quad \text{при } M(\xi) > 0.$$

Служит для измерения среднего квадратичного отклонения в долях математического ожидания, является мерой рассеяния распределения, дает возможность получения безразмерной величины.

Билет №25. Моменты, центральные моменты, абсолютные моменты случайной величины.

Момент порядка γ случайной величины ξ - это число μ_γ, равное $\mu_\gamma = M(\xi^\gamma)$.

Центральный момент порядка γ случайной величины ξ - число $\mu_{\gamma,0}$:

$$\mu_{\gamma,0} = M[\xi - M(\xi)]^\gamma.$$

Т.е. математическое ожидание - это момент 1-го порядка: $\mu_1 = M(\xi)$.

Дисперсия - центральный момент второго порядка

$$\mu_{2,0} = M[\xi - M(\xi)]^2$$

Абсолютный момент порядка γ случайной величины ξ - это число μ_γ, равное

$\mu_\gamma = M(|\xi|^\gamma)$, т.е. степень γ модуля случайной величины.

Центральный абсолютный момент порядка δ случайной величины ξ - это число $\mu_{\delta,0}$, равное

$$\mu_{\delta,0} = M[|\xi - M(\xi)|^{\delta}].$$

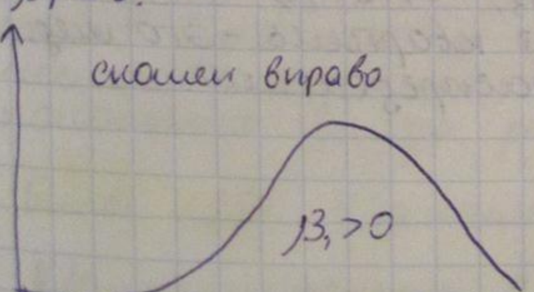
Билет № 26. Коэффициенты асимметрии и эксцесса случайной величины.

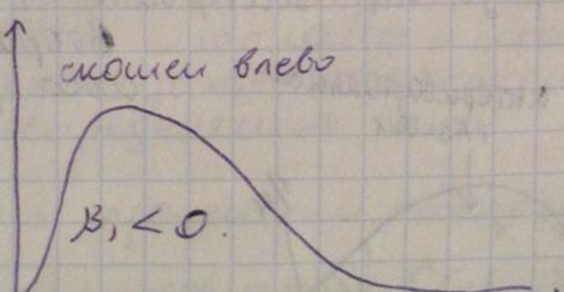
Коэффициент асимметрии случайной величины ξ - величина $\beta_1 = \frac{\mu_{3,0}}{[D(\xi)]^{3/2}}$;

Коэффициент эксцесса случайной величины ξ - величина

$$\beta_2 = \frac{\mu_{4,0}}{[D(\xi)]^2} - 3.$$

Если случайная величина симметрична, то $\beta_1 = 0$. При $\beta_1 > 0$ график функции плотности скошен вправо, при $\beta_1 < 0$ скошен влево. Для нормального распределения $\beta_1 = 0$.



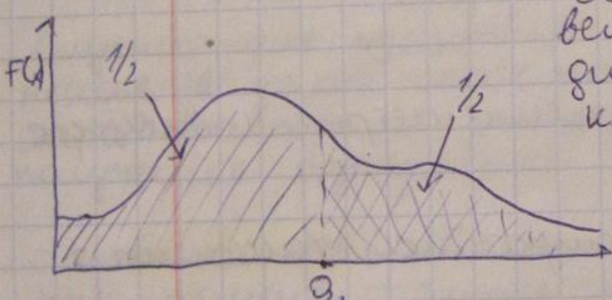


Для нормального распределения $\beta_2 = 0$.
Если $\beta_2 > 0$, то распределение "толще", чем нормальное, а если $\beta_2 < 0$, то распределение "тоньше," теснее

вокруг среднего, чем нормальное.
Коэффициент эксцесса указывает на нормальное распределение

Билет №27. Медиана, нижняя и верхняя квартили случайной величины.

Медиана случайной величины ξ - число Q_2, такое, что $P(\xi \le Q_2) \ge 1/2$; и $P(\xi \ge Q_2) \ge 1/2$.
т.е. вероятность того, что случайная величина примет значение меньше Q_2, равна $1/2$, и вероятность того, что случайная величина примет значение больше Q_2, равна $1/2$.

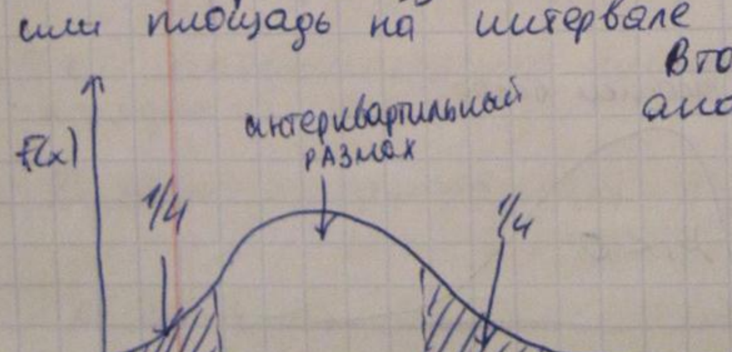


Если распределение случайной величины симметрично, то медиана совпадет с математическим ожиданием.
Для несимметричных распределений совпадения нет.

Первая (нижняя) квартиль распределения случайной величины ξ - число Q_1, такое, что
$$P(\xi \le Q_1) = 1/4.$$
или площадь на интервале $[-\infty; Q_1]$ под графиком функции плотности равна $1/4$.

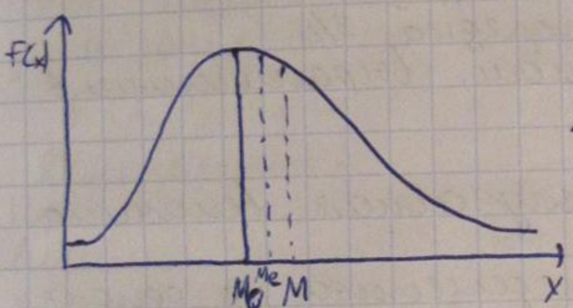
Третья (верхняя) квартиль распределения случайной величины ξ - число Q_3, такое, что
$$P(\xi > Q_3) = 1/4.$$
или площадь на интервале $[Q_3; +\infty) = 1/4$.

Вторая квартиль - это медиана распределения.



Билет №28. Интерквартильный размах, мода случайной величины.

Мода непрерывной случайной величины - такое значение x, в котором $f(x)$ достигает своего локального максимума. В интервал вокруг моды вероятнее всего попадет случайная величина.
Мода - "центр сгущения" распределения.
Унимодальное распределение - 1 мода
Мультимодальное - несколько.
Для симметричного унимодального распределения мода совпадает с математическим ожиданием и медианой.



Mo - мода
M - математическое ожидание
Me - медиана
Для несимметричного распределения они не совпадают.
Если ввести Q_1 и Q_3 - первую и третью квартиль, где
Q_1: $P(\xi \le Q_1) = 1/4$;
Q_3: $P(\xi > Q_3) = 1/4$,

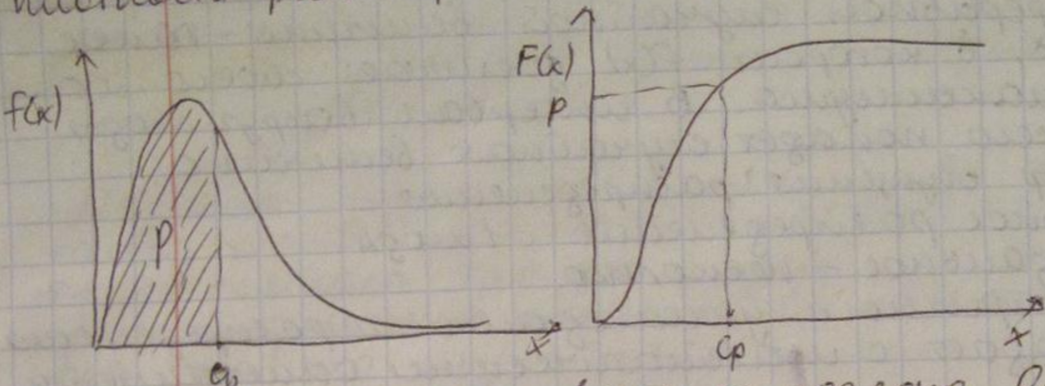
то вероятность того, что случайная величина примет значение в промежутке $[Q_1; Q_3]$ равна
$$P(Q_1 < \xi \le Q_3) = 1/2.$$

Разность $Q_3 - Q_1$ - интерквартильный размах.
Это мера рассеяния значений случайной величины, в отличие от моды, медианы и математического ожидания, являющихся характеристиками положения центра распределения.

Билет №29. Квантили и процентили распределения.

Квантиль порядка p распределения $F(x)$ - это число C_p, такое, что $F(C_p) = p$.
или $F(C_p) = P\{\xi \le C_p\} = p_к$
вероятность
порядок квантили

Квантиль порядка p - такое число c_p, что вероятность попасть левее этого числа на графике функции плотности равна p.



Т.е. медиана - это квантиль порядка 0,5;
первая квартиль - квантиль порядка 1/4;
третья квартиль - квантиль порядка 3/4.
Процентиль - квантиль с порядком, выраженным в процентах.

Билет №30. Многомерная случайная величина

Совокупность m функций, определенных на одном и том же множестве элементарных событий - m-мерная случайная величина (совокупность m одномерных случайных величин).
Определяется функцией распределения:
$$F(x_1; \dots; x_m) = p(\xi_1 \le x_1; \dots; \xi_m \le x_m).$$
Свойства функции распределения:
1) $0 \le F(x_1; \dots; x_m) \le 1$
2) $F(x_1; \dots; x_m)$ не убывает по каждому аргументу
3) $\lim\limits_{x_i \to -\infty} F(x_1; \dots; x_i; \dots; x_m) = 0$ для любого i
4) $\lim\limits_{x_i \to +\infty,\ i=\overline{1,m}} F(x_1; \dots; x_m) = 1$
5) $\lim\limits_{x_j \to \infty,\ j \ne i} F(x_1; \dots; x_i; \dots; x_m) = F_i(x_i)$, где
$F_i(x_i)$ - функция распределения одномерной случайной величины ξ_i

Дискретные многомерные случайные величины - составляющие их случайные величины являются дискретными. Принимают конечное или счетное множество значений.

Непрерывные многомерные случайные величины.
Многомерная случайная величина называется непрерывной, если непрерывна ее функция распределения $F(x_1; \dots; x_m)$, и существует функция плотности $f(x_1; \dots; x_m)$,

такая, что
$$F(x_1; x_2; \dots; x_m) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} \dots \int_{-\infty}^{x_m} f(x_1; x_2; \dots; x_m)\,dx_1; dx_2 \dots dx_m.$$

Две случайные величины ξ_1 и ξ_2 независимы тогда и только тогда, когда
$$F(x_1; x_2) = F(x_1) \cdot F(x_2).$$

Доказательство.
По определению, две случайных величины независимы, когда
$$p(a \le \xi_1 \le b;\ c \le \xi_2 \le d) = p(a \le \xi_1 \le b) \cdot p(c \le \xi_2 \le d).$$

А значит, по свойству функции плотности
$$\int_a^b \int_c^d f(x_1; x_2)\,dx_1 dx_2 = \int_a^b f(x_1)\,dx_1 \cdot \int_c^d f(x_2)\,dx_2.$$

Устремим $a \to -\infty$ и $c \to -\infty$, а тогда
$$\int_{-\infty}^{b}\int_{-\infty}^{d} f(x_1; x_2)\,dx_1 dx_2 = \int_{-\infty}^{b} f(x_1)\,dx_1 \cdot \int_{-\infty}^{d} f(x_2)\,dx_2.$$

Теперь если b и d будут переменными, то причем $b = x_1$; а $d = x_2$, то по определению функции распределения многомерной случайной величины
$$F(x_1; x_2) = F(x_1) \cdot F(x_2).$$

Билет № 31. Коэффициент корреляции случайной величины и его свойства.

Коэффициент корреляции двух случайных величин ξ_1 и ξ_2 - это число $\rho(\xi_1, \xi_2)$, равное:

$$\rho(\xi_1; \xi_2) = \frac{M[(\xi_1 - M(\xi_1)) \cdot (\xi_2 - M(\xi_2))]}{\sqrt{D(\xi_1) \cdot D(\xi_2)}}.$$

является мерой взаимосвязи ξ_1 и ξ_2.

Здесь $cov(\xi_1; \xi_2) = M[(\xi_1 - M(\xi_1)) \cdot (\xi_2 - M(\xi_2))]$.

(коэффициент ковариации).

Свойства коэф. корреляции.

1. Для любых ξ_1 и ξ_2 $-1 \leq \rho(\xi_1; \xi_2) \leq 1$.

2. Если ξ_1 и ξ_2 - независимые случайные величины, то $\rho(\xi_1; \xi_2) = 0$.

Обратное неверно.

Только в случае нормального распределения случайных величин ξ_1 и ξ_2 обратное верно.

3. Если $\xi_2 = a + b\xi$, где a и b - const, то

$$\rho(\xi_1; \xi_2) = \begin{cases} +1 & b > 0 \\ -1 & b < 0 \end{cases}$$

Обратное верно.

Т.е. коэффициент корреляции - мера линейности зависимости.

Две случайные величины называется некоррелируемыми, если коэффициент корреляции между ними равен нулю.

Из независимости следует некоррелируемость.

Из некоррелируемости следует независимость только в случае нормального распределения.

II. Математическая статистика

Билет № 1. Случайная выборка. Выборочные значения. Объем выборки.

Случайная выборка $X_1; X_2; \dots; X_n$ - совокупность наблюдений случайной величины ξ, полученных в результате n независимых повторений случайного эксперимента. Здесь n - объем выборки (количество повторений случайного эксперимента), отдельное X_i - выборочное значение.

Случайный эксперимент - эксперимент, результатом которого являются X_i - случайные величины, каждая из которых распределена по закону $F(x)$. Это распределение охватывает все возможные X_i.

$X_1; X_2; X_n$ - независимые случайные величины.

Генеральная совокупность - совокупность всех возможных выборочных значений (результатов случайного эксперимента), содержащихся в пропорциях, соответствующих распределению случайной величины.

Конкретная случайная выборка - извлечение из генеральной совокупности некого набора значений случайной величины ξ; одна из бесконечного числа потенциально возможных выборок объема n.

Билет №2. Гистограмма. График гистограммы для какого-либо примера.

Пусть имеется случайная выборка $X_1; X_2; \dots; X_n$ значений случайной величины ξ, распределенной по некоторому закону $F(x)$.

Тогда гистограмма выборки - функция

$$\hat{f}(x) = \frac{\beta}{n \cdot h},$$ где n - объем выборки; h - длина интервала разбиения; β - число значений выборки, попавших в данный интервал разбиения.

Функция $\hat{f}(x)$ определена для всех x и является аналогом функции плотности для случайной величины ξ:

$$\hat{f}(x) \xrightarrow[n \to \infty]{P} f(x)$$ при стремлении объема выборки к бесконечности функция гистограммы стремится по вероятности к функции плотности.

Такая гистограмма использует частоты значений в выборке. Можно еще построить гистограмму для численностей, при этом диапазон изменения значений выборки разбивается на K интервалов равной длины h, и по оси y откладывается не частота, а количество значений, попавших в этот интервал. Но такие гистограммы от гистограмм для частот различаются только масштабом шкалы на Oy.

Пример построения гистограммы.

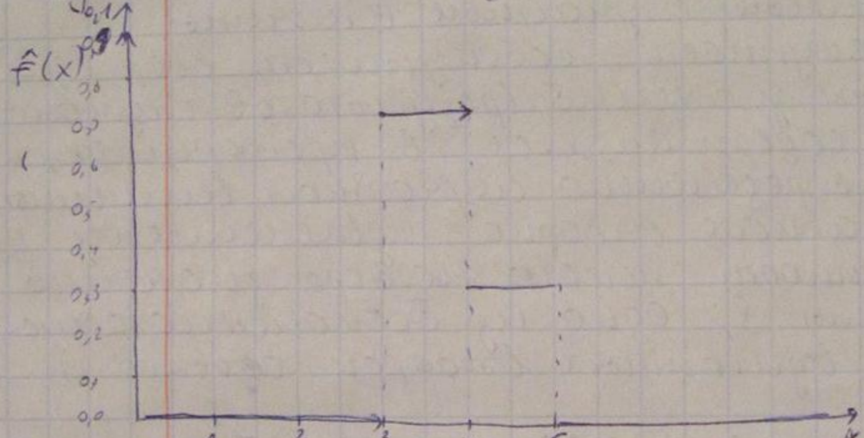
Случайная выборка:

3,0 3,9 4,5 3,5 3,9 4,7 3,3 3,0 4,1 3,4

$n = 10$

пусть $h = 1$. Тогда

Сумма площадей всех столбиков гистограммы равна 1.



$\hat{f}(x) = \frac{7}{10} = 0,7$ $\hat{f}(x) = 0,3$

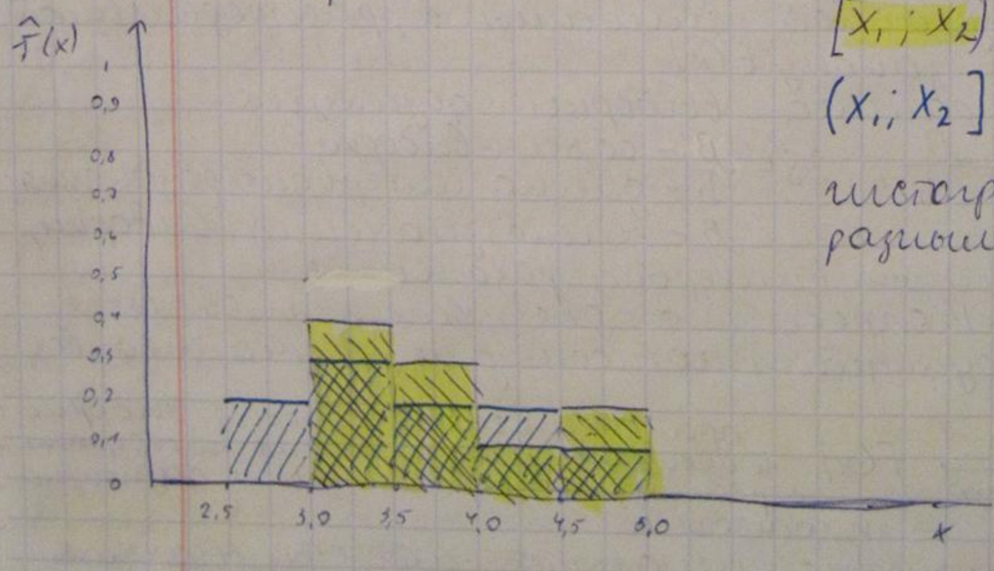
Если выбрать $h = 0,5$

$\cancel{[x_1, x_2]}$

$[x_1; x_2)$

$(x_1; x_2]$

гистограммы с разными интервалами:



Билет №3. Эмпирическая функция распределения. График этой функции для какого-нибудь примера.

Пусть $x_1; x_2; \dots; x_n$ – случайная выборка. Тогда эмпирической функцией распределения называется функция $\hat{F}(x) = \alpha/n$, где α – число значений выборки, меньших либо равных x.

Каждой выборке соответствует своя $\hat{F}(x)$, т.е. $\hat{F}(x)$ – случайная ф-ия.

Это ступенчатая функция, возрастающая от 0 до 1 скачками высотой в $1/n$ в точках $x_1; x_2; \dots; x_n$.

При $n \to \infty$ (n – объем выборки) эмпирическая ф-ия стремится по вероятности к функции распределения случайной величины ξ. Т.е. $\hat{F}(x)$ – аналог функции распределения.

Пример $\hat{F}(x)$.

Случайная выборка: 2,1 3,2 2,8 4,1 5,8 6,0 2,4

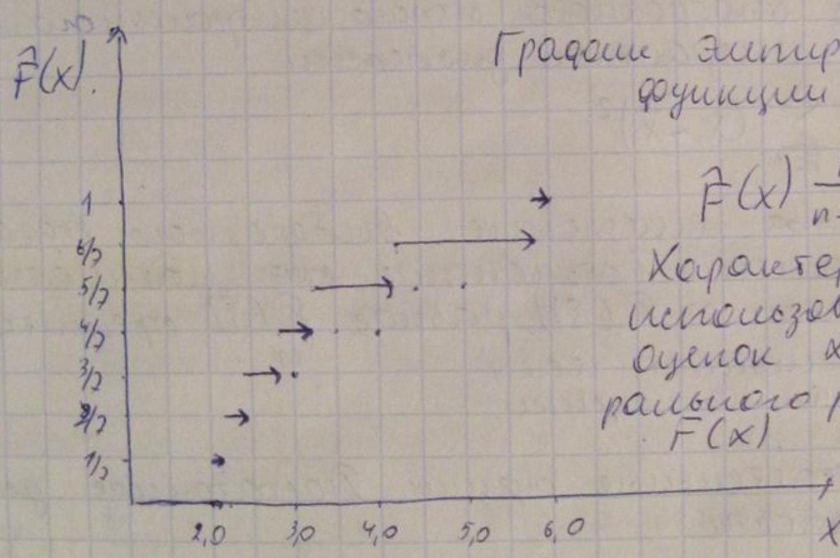


График эмпирической функции распределения

$$\hat{F}(x) \xrightarrow[n \to \infty]{P} F(x)$$

Характеристики $\hat{F}(x)$ можно использовать в качестве оценок характеристик генерального распределения $F(x)$.

Билет №4. Точечное оценивание. Несмещенные оценки.

Пусть есть выборка $x_1; x_2; \dots; x_n$ из случайной величины, распределенной по некоторому закону, и есть ее некая характеристика θ.

Оценка неизвестного параметра θ – любая ф-ия от случайной выборки: $\hat{\theta} = \hat{\theta}(x_1; x_2; \dots; x_n)$.

Точечная оценка θ — за оценку неизвестного параметра принимается конкретное значение (точка), вычисляемое по выборке.

Оценка должна быть близка к оцениваемому параметру. Однако оценка $\hat{\theta}$ — ф-ия от случайной выборки, т.е. случайная величина, и поэтому для разных выборок из одной генеральной совокупности $\hat{\theta}$ может быть разными.

Можно потребовать, чтобы в среднем оценки для разных выборок группировались вокруг ~~среднего~~ пара оцениваемого параметра, т.е

$$M(\hat{\theta}) = \theta.$$

Оценка $\hat{\theta}$ неизвестного параметра θ называется несмещенной, если $M(\hat{\theta}) = \theta$.

Выборочное математическое ожидание (выборочное ср. еднее) является несмещенной оценкой своего истинного математического ожидания. Но математическое ожидание выборочного второго центрального момента не равно генеральной дисперсии:

$$\hat{\mu}_{2,0} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2.$$

Однако, при $n \to \infty$ смещение выборочного второго центрального момента (отклонение ~~его~~ математич ожидания оценки от $D(\xi)$), равное $\frac{D(\xi)}{n}$, стремится к нулю. Это — асимптотически несмещенные оценки.

Билет №5. Состоятельные оценки. Достаточное условие состоятельности

Оценка $\hat{\theta}$ неизвестного параметра θ называется состоятельной оценкой для θ, если

$$\hat{\theta} \xrightarrow[n\to\infty]{P} \theta$$

$$\left(\forall \varepsilon > 0 \; P(|\hat{\theta} - \theta| > \varepsilon) \xrightarrow[n\to\infty]{} 0\right).$$

Этот критерий оценивания позволяет учитывать наиболее полно всю информацию, содержащуюся в выборке, ведь существуют несмещенные оценки, которые плохи тем, что они используют только часть этой информации.

Достаточное условие состоятельности:

Если $\hat{\theta} = \hat{\theta}(x_1; x_2; \ldots; x_n)$ - несмещенная оценка неизвестного параметра θ, и $D(\hat{\theta}) \to 0$ при $n \to \infty$, то $\hat{\theta}$ - состоятельная оценка для θ.

Доказательство:

Т.к. $\hat{\theta}$ - несмещенная оценка, то $M(\hat{\theta}) = \theta$.

Т.е. $\forall \varepsilon > 0 \; P(|\hat{\theta} - M(\hat{\theta})| > \varepsilon) \xrightarrow[n \to \infty]{} 0$, а поскольку $M(\hat{\theta}) = \theta$, то

$$\forall \varepsilon > 0 \; P(|\hat{\theta} - \theta| > \varepsilon) \xrightarrow[n \to \infty]{} 0.$$

Т.е. $P(|\hat{\theta} - M(\hat{\theta})| > \varepsilon) = P(|\hat{\theta} - \theta| > \varepsilon)$.

Но по неравенству Чебышева $P(|\xi - M(\xi)| > \varepsilon) \leq \frac{D(\xi)}{\varepsilon^2}$

$$P(|\hat{\theta} - \theta| > \varepsilon) < \frac{D(\hat{\theta})}{\varepsilon^2},$$

$D(\hat{\theta}) \to 0$ по условию, тогда имеем определение состоятельной оценки:

$$\lim_{n \to \infty} P(|\hat{\theta} - \theta| > \varepsilon) = 0.$$

Билет № 6. Выборочное среднее, выборочная дисперсия, выборочное среднеквадратичное уклонение, выборочный коэффициент вариации.

Выборочное среднее - оценка для математического ожидания.

$$\overline{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$ Несмещенная и состоятельная.

Выборочная дисперсия - оценка для дисперсии.

$$s^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n-1}$$

Если разделить на n, ~~то оценка~~ т.е. принять за выборочную дисперсию второй центральный момент, $\hat{\mu}_{2b} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$, то его математическое ожидание не будет равно дисперсии.

Для получения несмещенности оценки выборочной дисперсии ее надо поделить на $(n-1)$

Выборочное среднеквадратичное отклонение

$\varsigma = \sqrt{S^2}$. Однако такая оценка будет асимптотически несмещенной.

Выборочный коэффициент вариации

$$\bar{V}(\xi) = \frac{S}{\bar{x}}$$

Билет № 7. Доказать несмещенность и состоятельность выборочного среднего, как оценки математического ожидания.

Пусть есть некоторая случайная выборка $x_1; x_2; \dots; x_n$.

Тогда

$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$ выборочное среднее.

$$M\left(\frac{x_1 + x_2 + \dots + x_n}{n}\right) = \frac{1}{n} M(x_1 + x_2 + \dots + x_n) = \frac{1}{n}\left[M(x_1) + M(x_2) + \dots + M(x_n)\right] =$$

выносим константу мат. ожидание суммы равно сумме мат. ожид.

$$= \frac{1}{n}\left[m + m + \dots + m\right] = \frac{1}{n} n \cdot m = m.$$

у всех случ. величин $x_1; x_2; \dots x_n$ одно и то же распределение, а значит, и одно и то же мат. ожид.

$M(\xi) = m$.

т.е. $M(\bar{x}) = M(\xi) = m$.

Доказана несмещенность оценки.

Теперь

$$D(\bar{x}) = D\left(\frac{x_1 + x_2 + \dots + x_n}{n}\right) = \sum\left(\frac{x_i - M(\xi)}{n}\right)^2 = \frac{1}{n^2}\sum(x_i - M(\xi))^2 =$$

по ф-ле дисперсии выносим константу

$$= \frac{1}{n^2} D(x_1 + x_2 + \dots + x_n) = \frac{1}{n^2}(D(x_1) + D(x_2) + \dots + D(x_n)) =$$

дисперсия суммы равна сумме дисперсий

$$= \frac{1}{n^2}(D(\xi) + D(\xi) + \dots + D(\xi)) = \frac{n D(\xi)}{n^2} = \frac{D(\xi)}{n}$$

все x_n - значения одной и той же случ. величины

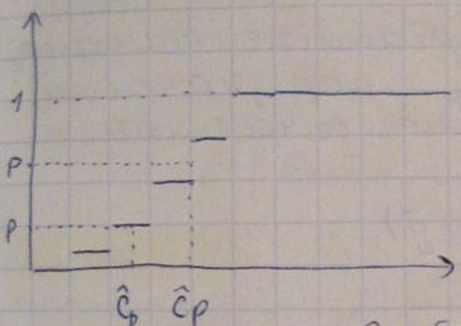
при $n \to \infty$ $\frac{D(\xi)}{n} \to 0$

т.е. доказана состоятельность.

Билет №8. Выборочная квантиль, выборочная медиана, выборочные нижняя и верхняя квартили, выборочная мода.

Для получения значений этих характеристик необходима функция распределения $\hat{F}(x)$ (эмпирическая) и ее график.

Выборочная квантиль $\hat{c}_p$ порядка p — абсцисса точки пересечения прямой $y=p$ с графиком ~~[illegible]~~ эмпирической функции распределения $\hat{F}(x)$.



Если пересечением является отрезок, то в качестве квантили берется абсцисса середины этого отрезка.

Выборочная квартиль, и верхняя, и нижняя, получаются как квантили соответствующих порядков: $\hat{c}_{0,25}$ $\hat{c}_{0,75}$.

Выборочная медиана — квантиль порядка $1/2$ $\hat{c}_{0,5}$

Выборочная мода оценивается двумя методами.

· Дискретная случайная величина — значение, встретившееся наибольшее число раз.

· Непрерывная случайная величина — середина интервала, содержащего локальный максимум на гистограмме (аналоге функции плотности).

Билет №9. Выборочные коэффициенты асимметрии и эксцесса.

Выборочный коэффициент асимметрии

$$\hat{\beta}_1 = \frac{\hat{\mu}_{3,0}}{\hat{\mu}_{2,0}^{3/2}}$$

, где $(\hat{\mu}_{2,0})^{3/2}$ — выборочный центральный момент порядка 2 в степени 3/2

$$\hat{\mu}_{2,0} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

Не выборочная дисперсия!

$$\hat{\mu}_{3,0} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^3$$

выборочный центральный момент порядка 3.

Выборочный коэффициент эксцесса $\hat{\beta}_2 = \frac{\hat{\mu}_{4,0}}{\hat{\mu}_{2,0}^{2}}$, где

$\hat{\mu}_{4,0}$ - выборочный центральный момент порядка 4

$$\hat{\mu}_{4,0} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^4$$

$\hat{\mu}_{2,0}$ - выборочный центральный момент порядка 2.

$$\hat{\mu}_{2,0} = \frac{1}{n}\sum_{i=1}^{n}(x_i - x)^2$$

Не является оценкой выборочной дисперсии, вернее, является, но это будет смещенная оценка выборочной дисперсии.

Билет №10. Выборочный коэффициент корреляции

Если у нас есть двумерная случайная величина, $(\xi_1; \xi_2)$, т.е. выборка состоит из n пар значений $(x_1; y_1), (x_2; y_2), \dots (x_n; y_n)$, то можно вычислить выборочную ковариацию:

$$\hat{cov}(\xi_1; \xi_2) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}),$$

где $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$; $\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$.

Тогда выборочный коэффициент корреляции

$$r = \frac{\hat{cov}(\xi_1; \xi_2)}{\sqrt{S^2_{\xi_1} \cdot S^2_{\xi_2}}}, \text{ где}$$

$$S^2_{\xi_1} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2, \quad S^2_{\xi_2} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{y})^2$$

выборочные дисперсии для ξ_1 и ξ_2. Которые смещенные.

Билет №11. Доверительные интервалы.

Пусть есть случайная выборка $x_1; x_2; \dots x_n$ из неизвестной случайной величины, распределенной по некоторому закону $F(x)$ и пусть есть неизвестная характеристика распределения θ. Тогда доверительный интервал для θ - это интервал $(\theta_{нижн}; \theta_{верхн})$, причем $\theta_{нижн} = \theta_{нижн}(x_1; x_2; \dots x_n)$ и

$\theta_{верх}(x_1;\ x_2;\ \dots;\ x_n)$, и

$P[\theta_{нижн} \le \theta \le \theta_{верх}] = 1-\varepsilon$, где $1 > \varepsilon > 0$

Вероятность того, что доверительный интервал накроет неизвестный параметр, равна $1-\varepsilon$.
Если $\varepsilon = 0{,}05$, то доверительный интервал 95%-ный.
ε - уровень доверия
С вероятностью ε случайная величина не попадает в доверительный интервал.
$\theta_{нижн}$ и $\theta_{верх}$ - доверительные пределы, зависят от выборочных значений и заданной доверительной вероятности $1-\varepsilon$.
Центральный доверительный интервал - обычно используется. Это интервал, для которого

$$P(\theta < \theta_{нижн}) = P(\theta > \theta_{верх}) = \varepsilon/2.$$

Свойство всех доверительных интервалов:

$$P(\theta < \theta_{нижн}) + P(\theta > \theta_{верх}) = \varepsilon.$$

Билет № 12. Доверительный интервал для математического ожидания в случае нормального распределения.

Возможно два случая: известна дисперсия и неизвестна дисперсия.

1. Известна дисперсия.

Пусть ξ - нормально распределенная случайная величина с известной дисперсией.

Тогда $M(\xi) = \mu$, $D(\xi) = \sigma^2$

Имеется выборка значений этой случайной величины $x_1;\ x_2;\ \dots;\ x_n$.

Поскольку $M(\bar{x}) = M(\xi)$, а $D(\bar{x}) = \dfrac{D(\xi)}{n}$

$$\left(D(\bar{x}) = D\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} D(x_i) = \frac{1}{n^2}\, n\, D(\xi) = \frac{D(\xi)}{n} \right), \text{ то}$$

$$M(\bar{x}) = \mu; \quad D(\bar{x}) = \frac{\sigma^2}{n}.$$

Стандартное нормальное распределение $\dfrac{\xi - \mu}{\sigma}$.

Логика построения доверительных интервалов в том, чтобы, по сути, выбрать известные квантили порядков

$\frac{\varepsilon}{2}$ и $1-\frac{\varepsilon}{2}$, которые и будут $\theta_{нижн.}$ и $\theta_{верхн.}$, а для этого надо найти такие распределения, т.е. такую функцию от выборки, ~~расп~~ чтобы она была распределена по закону с известными квантилями, и являлась оценкой величины, для которой строится доверительный интервал.

Для математического ожидания (по сути $\overline{x}$) можно применить статистику (функцию от выборки)

$$u = \frac{\overline{x}-\mu}{\sigma/\sqrt{n}}$$

распределенную по стандартному нормальному закону.

Тогда

$$P\left(-u_{1-\frac{\varepsilon}{2}} < u < u_{1-\varepsilon/2}\right) = 1-\varepsilon;$$

$$P\left(-u_{1-\frac{\varepsilon}{2}} < \frac{\overline{x}-\mu}{\sigma/\sqrt{n}} < u_{1-\varepsilon/2}\right) = 1-\varepsilon,$$

$$P\left(\overline{x} - u_{1-\varepsilon/2}\cdot\frac{\sigma}{\sqrt{n}} < \mu < \overline{x} + u_{1-\varepsilon/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\varepsilon.$$

Так будут рассчитываться доверительные интервалы. При этом $u_{1-\varepsilon/2}$ – квантиль стандартного нормального распределения порядка $1-\varepsilon/2$.

2. Дисперсия не известна.

Тогда в формуле статистики σ заменяется на его выборочную оценку s.

$$t = \frac{\overline{x}-\mu}{s/\sqrt{n}}.$$

Эта функция распределена по t-закону.

Тогда

$$\mu_{нижн.} = \overline{x} - t_{n-1;\,1-\varepsilon/2}\cdot\frac{s}{\sqrt{n}},$$

$$\mu_{верхн.} = \overline{x} + t_{n-1;\,1-\varepsilon/2}\cdot\frac{s}{\sqrt{n}}$$

$\overline{x}$ – выборочное среднее

s – выборочное среднеквадратичное отклонение

n – объем выборки

$t_{n-1;\,1-\frac{\varepsilon}{2}}$ – квантиль распределения Стьюдента с $(n-1)$ степенью свободы порядка $1-\varepsilon/2$

При увеличении n доверительный интервал сужается.

При уменьшении ε доверительный интервал расширяется.

Эти доверительные интервалы нечувствительны к умеренным отклонениям от нормальности. При больших n ($n \geq 15$) можно использовать эти формулы для любого распределения.

$\frac{s}{\sqrt{n}}$ - ошибка среднего.

Билет № 13. Доверительный интервал для дисперсии в случае нормального распределения.

Пусть есть выборка $x_1; x_2; ...; x_n$, из случайной величины, распределенной по нормальному закону с неизвестными дисперсией и математическим ожиданием. Для нахождения доверительного интервала используется статистика

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}.$$

Распределена по χ^2-распределению с n-1 степенями свободы.

Тогда

$$\sigma^2_{нижн} = \frac{(n-1)s^2}{\chi^2_{n-1;\,1-\varepsilon/2}}, \quad \sigma^2_{верх.} = \frac{(n-1)s^2}{\chi^2_{n-1;\,\varepsilon/2}},$$

где $\chi^2_{n-1,\,1-\varepsilon/2}$ и $\chi^2_{n-1;\,\varepsilon/2}$ - квантили χ^2-распределения с (n-1) степенями свободы порядка $1-\varepsilon/2$ и $\varepsilon/2$.

Этот доверительный интервал является центральным, но несимметричным, поскольку χ^2-распределение несимметрично.

И еще этот доверительный интервал чувствителен к отклонениям от нормальности распределения, даже в случае большой выборки. Поэтому он применяется только для нормального распределения.

Билет № 14. Доверительный интервал для коэффициента корреляции.

Пусть есть некоторая случайная выборка $(x_1; y_1)$, $(x_2; y_2)$, ... $(x_n; y_n)$ из двумерного нормального распределения. И ρ - коэффициент корреляции случайных величин ξ_1 и ξ_2, а r - выборочный коэффициент корреляции

Чтобы $u(r)$ была распределена по нормальному закону (так удобно считать квантили), с r производится преобразование Фишера:

$$z = \frac{1}{2}\ln\frac{1+r}{1-r}$$

$$M(z) = \frac{1}{2}\ln\frac{1+\rho}{1-\rho}\ ; \quad D(z) = \frac{1}{n-3}.$$

Тогда

$$u = \sqrt{n-3}\left(z - \frac{1}{2}\ln\frac{1+\rho}{1-\rho}\right) \sim N(0;1)$$

распределена по стандартному нормальному закону.

Она с вероятностью $1-\varepsilon$ будет заключена в пределах $\pm u_{1-\varepsilon/2}$. Если решить полученное неравенство

$$P\left(-u_{1-\varepsilon/2} < \sqrt{n-3}\left(z - \frac{1}{2}\ln\frac{1+\rho}{1-\rho}\right) < u_{1-\varepsilon/2}\right) = 1-\varepsilon,$$

то получатся страшные и огромные формулы для вычисления $\rho_{нижн.}$ и $\rho_{верхн.}$

Этот доверительный интервал чувствителен к отклонениям от нормальности и годится только для нормально распределенных двумерных случайных величин.