[ТВиМС] Теормин.
(2 курс, IV семестр)

Часть 1. **Теория вероятностей**

1. Парадокс Бертрана
2. Классическое опр. вероятности. Задача о «Спортлото»
3. Задача об оптимальном дележе ставки
4. Геом. вероятности. Задача о встрече. Задача Бюффона
5. Вероятностное пространство.
6. Свойства вероятности. Эквивалентности счетной аддитивности и непрерывности вероятности
7. Условные вероятности
8. Формула полной вероятности
9. Формула Байеса. Задача о детекторе лжи

10. Определение случайной величины
11. Замкнутость СЛВ относительно операций минимума, максимума и предельного перехода
12. Замкнутость СЛВ относительно арифметических операций
13. Функции распределения. Простейшие свойства
14. Понятие меры Лебега
15. Понятие интеграла Лебега
16. Классификация СЛВ. Дискретные случайные величины
17. Классификация СЛВ. Абсолютно непрерывные СЛВ
18. Классификация СЛВ. Сингулярные СЛВ

19. Мат. ожидание. Простейшие свойства. Задача о пенсионере
20. Медиана. Простейшие св-ва. Задача о предпринимателе
21. Задача о среднем годовом доходе
22. Дисперсия. Простейшие свойства
23. Ковариация. Коэффициент корреляции
24. Неравенства Маркова и Чебышева
25. Неравенство Гаусса. «Правило трех сигм»
26. Закон больших чисел в форме Чебышева
27. Неравенства Иенсена и Ляпунова
28. Коэффициенты эксцесса и асимметрии

29. Схема испытаний Бернулли. Биномиальное распределение
30. Теорема Пуассона. Распр. Пуассона. Задача о булочке
31. Задача об инсектициде
32. Теорема Муавра—Лапласа. Задача о докторе Споке
33. Нормальное распределение. Его коэффициент эксцесса
34. Нормальное распределение. Неравенства для хвостов
35. Задача о рейтинге. Решение по неравенству Чебышева
36. Задача о рейтинге. Решение по теореме Муавра—Лапласа

37. Схема испытаний Бернулли. Геом. распр. Его моменты
38. Предельная теорема для геом. распределение. Показательное распределение. Его моменты.
39. Показательное распр. Св-во отсутствия последействия.
40. Совместные распределения. Совместные плотности
41. Условные плотности и условные математические ожидания
42. Виды сходимости СЛВ. Соотношения между ними

43. Характеристические функции. Их простейшие свойства
44. Теорема непрерывности соответствия между распределениями и хар. функциями (без док-ва)
45. ЗБЧ в форме Хинчина
46. ЦПТ
47. ЦПТ как оценка скорости сходимости в ЗБЧ
48. Условие Линдеберга. Теорема Линдеберга—Феллера
49. Теорема Ляпунова. Взаимосвязь условия Ляпунова и условия Линдеберга
50. Оценки скорости сходимости в ЦПТ

Часть 2. **Математическая статистика**

51. Статистическая структура. Выборка. Вариационный ряд. Порядковые статистики
52. Эмпирическая функция распределения. Независимость равномерного расстояния между теоретической и эмпирической функциями распределения от теоретической функции распределения
53. Эмпирическая функция распределения. Теорема Гливенко. Теорема Колмогорова. Репрезентативность выборки
54. Выборочные моменты, их свойства
55. Точечное оценивание. Неравенство Рао—Крамера. Эффективные оценки.
56. Достаточные статистики. Критерий факторизации
57. Оптимальные оценки. Теорема Блэкуэлла—Колмогорова
58. Полные статистики. Теорема Лемана—Шеффе
59. ОМП. Их асимптотические свойства
60. Оценки метода моментов
61. Интервальное оценивание. Метод центральной СЛВ
62. Интервальное оценивание с помощью точечной оценки
63. Основные понятия проверки статистических гипотез.
64. Критерий согласия Колмогорова
65. Критерий согласия хи-квадрат
66. Байесовский критерий различения двух простых гипотез
67. Лемма Неймана—Пирсона
68. Статистический анализ нормальных выборок. Распределения хи-квадрат и Стьюдента. Теорема Фишера

this page left blank intentionally

1. Парадокс Бертрана
2. Классическое опр. вероятности. Задача о «Спортлото»
3. Задача об оптимальном дележе ставки
4. Геом. вероятности. Задача о встрече. Задача Бюффона
5. Вероятностное пространство
6. Свойства вероятности. Эквивалентности счетной аддитивности и непрерывности вероятности
7. Условные вероятности
8. Формула полной вероятности
9. Формула Байеса. Задача о детекторе лжи

10. Определение случайной величины
11. Замкнутость СЛВ относительно операций min,max,lim
12. Замкнутость СЛВ относительно арифм. операций
13. Функции распределения. Простейшие свойства
14. Понятие меры Лебега
15. Понятие интеграла Лебега
16. Классификация СЛВ. Дискретные случайные величины
17. Классификация СЛВ. Абсолютно непрерывные СЛВ
18. Классификация СЛВ. Сингулярные СЛВ

19. Мат. ожидание, свойства. Задача о пенсионере
20. Медиана, свойства. Задача о предпринимателе
21. Задача о среднем годовом доходе
22. Дисперсия. Простейшие свойства
23. Ковариация. Коэффициент корреляции
24. Неравенства Маркова и Чебышева
25. Неравенство Гаусса. «Правило трех сигм»
26. Закон больших чисел в форме Чебышева
27. Неравенства Иенсена и Ляпунова

28. Коэффициенты эксцесса и асимметрии
29. Схема Бернулли. Биномиальное распределение
30. Теорема Пуассона. Распр. Пуассона. Задача о булочке
31. Задача об инсектициде
32. Теорема Муавра—Лапласа. Задача о докторе Споке
33. Нормальное распределение. Его коэффициент эксцесса
34. Нормальное распределение. Неравенства для хвостов
35. Задача о рейтинге. Решение по неравенству Чебышева
36. Задача о рейтинге. Реш. по теореме Муавра—Лапласа

37. Геом. распределение. Его моменты
38. ПТ для геом. распределения. Показательное распределение. Его моменты
39. Показательное распр. Св-во отсутствия последействия
40. Совместные распределения. Совместные плотности
41. Условные плотности и условные МО
42. Виды сходимости СЛВ. Соотношения между ними

43.Характеристические функции. Их простейшие свойства
44.Теорема непр. соответствия между распределениями и хар. функциями (без док-ва)
45.ЗБЧ в форме Хинчина
46.ЦПТ
47.ЦПТ как оценка скорости сходимости в ЗБЧ
48.Условие Линдеберга. Теорема Линдеберга—Феллера
49.Теор. Ляпунова. Связь (у) Ляпунова и (у) Линдеберга
50.Оценки скорости сходимости в ЦПТ

51.Статистическая структура
52.Эмпирическая функция распределения
53.Т. Гливенко. Т. Колмогорова. Репрезент-ть выборки
54.Выборочные моменты, их свойства
55.Точечное оценивание. Неравенство Рао—Крамера. Эффективные оценки.
56.Достаточные статистики. Критерий факторизации
57.Оптимальные оценки. Т. Блэкуэлла—Колмогорова
58.Полные статистики. Т. Лемана—Шеффе
59.ОМП. Их асимптотические свойства
60.Оценки метода моментов

61.Интервальное оценивание. Метод центральной СЛВ
62.Интервальное оценивание точечной оценкой
63.Основные понятия проверки статистических гипотез
64.Критерий согласия Колмогорова
65.Критерий согласия хи-квадрат Пирсона
66.Байесовский кр. различения двух простых гипотез
67.Лемма Неймана—Пирсона
68.Статистический анализ нормальных выборок.
69.Распределения хи-квадрат и Стьюдента. Т. Фишера

Теория вероятностей

1 Парадокс Бертрана

- задача на геом. вероятность: выберем наудачу хорду AB в круге, треб. найти P (дл. хорды > дл. стороны вписанного РС треугольника).

Парадокс в том, что **Р опр-ся в зависимости от метода.**

2 Классическое опр. вероятности. Задача о «Спортлото»

Элем. исходом (событием) ω наз. элемент **пр-ва элементарных исходов** (непустое мн-во Ω).

Говорят **событие A произошло,** если произошло $\forall$ЭС $\omega \in A$ – благоприятствующее A.

Вероятность – всегда ф-ция, опр. на мн-ве случайных (не элем.) событий ($\sum_{\omega\in\Omega} P(\{w\}) = 1$). Пусть мн-во $\Omega = \{\omega_1, \dots, \omega_n\}$, причем все события равновероятны, т.е. $P(\{\omega_i\}) = 1/n \Rightarrow$ вер-ть $\forall A$ опр. ф-лой $P(A) = \frac{|A|}{|\Omega|} = \frac{\text{число элем.A}}{n}$. **Основные св-ва:**

1) $P(\theta) = 0, P(\Omega) = 1;$
2) $P(A \cup B) = P(A) + P(B) - P(AB);$
3) $P(\bar{A}) = 1 - P(A).$

Спортлото «6 из 49»: в урне 49 занумерованных шаров, извлекаются наудачу первые 6 и наз. их номера. Если играющий угадал 4,5,6 из номеров, то билет выигрышн. $P = \{\text{игрок угадает 5 номеров}\}$?

3 Задача об оптимальном дележе ставки

- два игрока начали игру из неск. партий. Каждая непременно выигрывается одним. Тот, кто выигрывает первым 6 партий забирает обе равные ставки, внесенные вначале игры. Игру прекратили на счете 5:3. Как нужно разделить ставки безобидной игры?

4 Геом. вероятности. Задача о встрече. Задача Бюффона

Пусть имеется нек. обл. G, а в ней g с квадрируемой границей. Наудачу бросаем точку в $G \to$ вероятность попадания в g равна $p = \frac{mes\ g}{mes\ G}$ – **геом. вероятность** (задает равномерное в G распределение).

Задача о встрече: Два лица X, Y условились встретиться между 12 и 13 часами. Пришедший ждет 20 мин. и уходит. Найти Р встречи, если приход каждого в течение часа может произойти наудачу и моменты независимы.

кудр. 21

Задача (игла Бюффона): пл-ть разграфлена паралл. прямыми на $\rho = 2a$. На пл-ть бросают иглу дл.$2l < 2a$. Найти Р={игра пересечет какую-нибудь прямую}.

5 Вероятностное пространство

σ-алгеброй $\boldsymbol{\mathcal{F}}$ наз. класс (мн-во) подмножеств Ω, обл: след. св-вами (аксиомами):

1) $\Omega \in \mathcal{F};$
2) *если* $A_1, A_2, \dots \in \mathcal{F} \Rightarrow \cup A_i \in \mathcal{F}, \cap A_i \in \mathcal{F};$
3) *если* $A \in \mathcal{F} \Rightarrow \bar{A} \in \mathcal{F}.$

Событием (случайным) наз. эл-т σ-алгебры $\mathcal{F}$.

Борелевской σ-алгеброй $\boldsymbol{\mathcal{B}}$ наз. σ-алгебра порожд. мн-вом всех открытых интервалов (эл-ты – **бор. мн-ва**).

Пара $(\boldsymbol{\Omega}, \boldsymbol{\mathcal{F}})$ наз. **измеримым пространством.**

Тройка $(\boldsymbol{\Omega}, \boldsymbol{\mathcal{F}}, \boldsymbol{P})$ наз. **вероятностным пространством** (вероятность – распределение вер-тей на изм. пр-ве) – не пр-во в топологическом смысле, а модель, в кот. опр. каждого след. элем. тройки базируется на пред.

6 Свойства вероятности. Эквивалентности счетной аддитивности и непрерывности вероятности

Вероятностью наз. действительная ф-ция случ. события $P: \mathcal{F} \to \mathbb{R}$, удовл. след. аксиомам:

1) неотриц: $P(A) \geq 0, \forall A \in \mathcal{F};$
2) нормированность: $P(\Omega) = 1;$

3) σ-аддитивность (счетная):
если $A_1, A_2 \in \mathcal{F}, A_iA_j = \emptyset \Rightarrow P(\cup A_i) = \sum P(A_i)$.

Основные св-ва:

1) $P(\emptyset) = 0; P(\bar{A}) = 1 - P(A)$;
2) $A \subset B \Rightarrow P(A) \le P(B)$;
3) $P(A) \le 1 \ \forall A \in \mathcal{F}$;
4) $P(A \cup B) = P(A) + P(B) - P(AB)$;
5) $P(\cup A_n) \le \sum P(A_n)$;

Аксиома непрерывности: пусть посл-ть событий $\{B_n\}: B_{n+1} \subset B_n, B = \cap B_n \Rightarrow P(B_n) \to P(B), n \to \infty$.

Т. Треб. аксиомы нерп-ности ~ σ-аддитивности.

7 Условные вероятности

Пусть A, B – события, причем $P(B) > 0$. **Условной вероятностью** события A при условии, что произошло B, называется число $P(A|B) = P(AB)/P(B)$.

События A, B - **независимые,** если $P(AB) = P(A)P(B)$ (не путать с *несовместными* $AB = \emptyset$). Их **св-ва:**

1) $P(B) > 0$, незав. $A, B \sim P(A|B) = P(A)$;
2) A, B – незав. $\Rightarrow \bar{A}, B$ – незав.
3) A, B_1 – незав, A, B_2 – незав $\Rightarrow A, B_1 \cup B_2$ – незав.

8 Формула полной вероятности

Пусть A - событие, $\{B_n\}$ – попарно несовместные: $P(B_i) > 0, A \subset \cup B_i \Rightarrow \boldsymbol{P(A) = \sum P(B_i) P(A|B_i)}$.

//площадь фигуры = сумме площ. фигур-составляющих

Если $\cup B_i = \Omega \rightarrow \{B_n\}$ образуют **полную группу соб.**

9 Формула Байеса. Задача о детекторе лжи

Пусть A - событие, $\{B_n\}$ – попарно несовместные: $P(B_i) > 0, A \subset \cup B_i$ и $P(A) > 0 \Rightarrow$ справедлива **ф-ла Байеса:** $P(B_k|A) = \frac{P(B_k)P(A|B_k)}{\sum P(B_i)P(A|B_i)}$.

Задача о детекторе лжи: есть хар-ки детектора джи
$P(z = false|y = false) = 0.99 = \alpha$,
$P(z = false|y = true) = 0{,}01 = \beta$.
Пусть $P(y = false) = 0.01 = \gamma$.
$\Rightarrow P(y = false|z = false) = \frac{\alpha\gamma}{\alpha\gamma + \beta(1-\gamma)} = \frac{1}{2}$.

10 Определение случайной величины

Сл.в. наз. ф-ция эл. события $\xi: \Omega \to \mathbb{R}$, обл. **св-вом измеримости:** $\xi^{-1}(B) = \{\omega | \xi(\omega) \in B\} \in \mathcal{A} \ \forall B \in \mathcal{B}$.

Т. (критерий) Пусть $\mathcal{E}: \delta(\mathcal{E}) = \mathcal{B}$. Тогда $\xi(\omega)$ – сл.в. $\Leftrightarrow \forall E \in \mathcal{E} \ \{\omega: \xi(\omega) \in \mathcal{E}\} \in \mathcal{A}$.

11 Замкнутость СЛВ относительно операций min,max,lim

Числ. ф-ция $g: \mathbb{R} \to \mathbb{R}$ наз-ся **борелевской ф-цией,** если $\forall B \in \mathcal{B} \ g^{-1}(B) \in \mathcal{B}$.

Т. Пусть $g(x)$ – БФ, ξ – СЛВ $\Rightarrow \eta(\omega) = g(\xi(\omega))$ – СЛВ.

→[1] Если ξ – СЛВ $\Rightarrow \xi^n, \max(0,\xi), -\min(0,\xi)$ – СЛВ.

Т. Пусть $\{\xi_n\}$ – СЛВ $\Rightarrow sup, inf, \overline{lim}, \underline{lim}$ – СЛВ.

12 Замкнутость СЛВ относительно арифм. операций

Пусть $\exists x_i \in \mathbb{R}: A_iA_j = \emptyset, A_i \in \mathcal{A}, \cup A_i = \Omega \rightarrow$
$\xi(\omega) = \sum x_i \mathbb{I}_{Ai}(\omega)$ – **дискретная СЛВ.**

Если $\cup A_i = \Omega \to \xi(\omega)$ – **простая СЛВ** (может принимать конечное число значений).

Т. Пусть $\xi(\omega) \ge 0 \Rightarrow \exists \{\xi_n(\omega)\}: \forall n \ \xi_n$ – простая и $\xi_n \uparrow \xi$.

[$\xi_n = \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \mathbb{I}_{\left\{\omega: \frac{k-1}{2^n} \le \xi(\omega) \le \frac{k}{2^n}\right\}}(\omega) + n \cdot \mathbb{I}_{\{\omega: \xi(\omega) \ge n\}}(\omega)$]

→ Сущ простые $\{\xi_n\}: |\xi_n| \le |\xi|, \xi_n \to \xi$.

Т. Пусть $\xi(\omega)$ –СЛВ $\Rightarrow \xi \pm \eta, \xi \times \eta, \xi/\eta$ – СЛВ.

13 Функции распределения. Простейшие свойства

Распределением сл.в ξ наз. ф-цию $P_\xi(B) = P(\xi \in B)$.

Ф-цией распределения сл.в. ξ наз. $F_\xi: \mathbb{R} \to \mathbb{R}$, опр-мая $F_\xi(x) = P(\{\omega: \xi(\omega) < x\}) = P_\xi((-\infty, x)) \ \forall x \in \mathbb{R}$.

Простейшие св-ва:

1) $0 \le F_\xi(x) \le 1$;
2) неубывание: $x < y \Rightarrow F_\xi(x) \le F_\xi(y)$;
3) непр. слева: $\lim_{x \to x_0} F_\xi(x) = F(x_0)$;
4) огр. вариация: $\lim_{x \to 0-\infty} F = 0, \lim_{x \to 0+\infty} F = 1$.

14 Понятие меры Лебега

Мерой Лебега наз. мера, явл. **продолжением длины** с наим. алгебры, сод. мн-во всех откр. интервалов, **на борелевскую σ-алг.**

Т. Сущ. ед. мера λ на $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, значение кот. на $\forall$ инт-ле равно его длине: $\lambda(a,b) = b - a$. Эта мера наз. мерой Лебега.

Наименьшей σ-алг, сод. B (явл. пересечением всех σ-алг, сод. B) наз. σ-алг, порожденной B: $\sigma(B)$.

Т. (**Каратеодори,** о продолжении вер. меры) Пусть $(\Omega, \mathcal{A}, P)$ – вер. пр-во в широком смысле $\Rightarrow$ $!\exists$ вер. мера Q, опр. на $\sigma(\mathcal{A})$: $Q(A) = P(A), \forall A \in \mathcal{A}$.

15 Понятие интеграла Лебега

Есть измеримое пр-во $(\boldsymbol{S}, \boldsymbol{\mathcal{H}})$ с полной неотриц. счетно-аддитивной мерой μ. Рассм. **простую ф-цию** $\boldsymbol{g}: \boldsymbol{S} \to \mathbb{R}$ (которая принимает не более счетного числа значений $g(s) = y_n \neq y_k (n \neq k)$, при $s \in S_n \in \mathcal{H}$, причем $\cup S_n = S$).

Ф-ция $f: S \to \mathbb{R}$ наз. **суммируемой по Лебегу** на S, если $\exists$ посл-ть простых суммируемых ф-ций $g_n \rightrightarrows f$ и $\lim_{n\to\infty} \int_S g_n d\mu = I$ - **интеграл Лебега.** (пример ф-ции, инт-мой по Лебегу, но не по Риману – ф-ция Дирихле)

16 Классификация СЛВ. Дискретные случайные величины

СЛВ наз. **дискретной,** если сущ. не более чем счетное мн-во B, т.ч. $P_\xi(B) = 1$.

Вырожденное $\xi \overset{\text{п.н.}}{=} a$, если $P(\xi = a) = 1$.

Классическое: $P(\xi = x_i) = \frac{1}{n}, \forall x_i \in \mathbb{R}$.

Бернулли $\xi \sim Bi(n,p)$: $P(\xi = k) = C_n^k p^k (1-p)^{n-k}$.

Пуассоновское $\xi \sim Pois(\lambda), \lambda > 0$: $P(\xi = k) = \frac{e^{-\lambda}\lambda^k}{k!}$.

Геом. $\xi \sim \mathcal{G}(p), p \in (0,1)$: $P(\xi = k) = (1-p)p^k$.

17 Классификация СЛВ. Абсолютно непрерывные СЛВ

Распр. СЛВ P_ξ наз. **абс. непр,** если $\exists$ такая $f(x) \geq 0$: $\forall B \in \mathcal{B}: P_\xi(B) = \int_B f(x)dx$, где $f(x)$ – плотность.

Равномерное $\xi \sim U[a,b]$: $f(x) = \begin{cases} 1/(b-a), x \in^{[a,b]}; \\ 0, x \notin^{[a,b]}. \end{cases}$

По Коши $\xi \sim K(a,\sigma)$: $f(x) = 1/\pi(1+x^2), x \in (-\infty, \infty)$.

Экспоненциальное $\xi \sim E(\lambda)$: $f(x) = \lambda\, e^{-\lambda x}, x \geq 0$.

Нормальное $\xi \sim N(a, \sigma^2)$: $f(x) = \frac{1}{sqrt(2\pi\sigma^2)} e^{-\frac{(x-\sigma)^2}{2\sigma^2}}$, причем $N(a,\sigma^2)(x) = N(0,1)\left(\frac{x-a}{\sigma}\right)$ – **стандартное.**

Гамма $\xi \in \Gamma(\alpha, \lambda)$: $f(x) = \begin{cases} 0, x \leq 0, \\ c \cdot x^{\lambda-1} e^{-\alpha x}, x > 0, \end{cases}$ где c выч. из св-в плотности $c = \alpha^\lambda \Gamma(\lambda), \Gamma = \int_0^\infty x^{\lambda-1} e^{-x} dx$.

//для счета ф-ция Лапласа - $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$

18 Классификация СЛВ. Сингулярные СЛВ

СЛВ ξ им. сингулярное распр., если $\exists B \in \mathcal{B}, \lambda(B) = 0$: $P(\xi \in B) = 1$, но при этом $P(\xi = x) = 0\ \forall x \in B$. Такое распр. сосредоточено на несчетном мн-ве с МЛ нуль (например, на лестнице Кантора). **Св-ва:**

1) $F_\xi(x) \in C$; $F'(x) = 0$ по мере Лебега;
2) $F(+\infty) - F(-\infty) = 1$.

Кривая Кантора на $K = [0,1] \backslash \cup_{n=1}^\infty \cup_{k=1}^n \left[\frac{3k-2}{3^n}, \frac{3k-1}{3^n}\right]$.

Т. Лебега о разложении: $F_\xi(x) = \alpha_1(\text{дскр}) + \alpha_2(\text{непр}) + \alpha_3(\text{синг}), \alpha_1 + \alpha_2 + \alpha_3 = 1$. Синг F, мн-во т. роста - меры 0.

19 Мат. ожидание, свойства. Задача о пенсионере

Мат. ожиданием СЛВ ξ наз. $\mathbb{E}\xi = \int_\Omega \xi(\omega) P(d\omega) = \int_\infty^\infty x f_\xi(x) dx$ [дискретный сл. - $\mathbb{E}\xi = \sum_{i=1}^\infty x_i P(\xi = x_i)$], если этот инт-л (сумма) сх-ся абсолютно.

Смысл: коорд. центра тяжести прямой. $//m(a_i) = p_i$

Свойства $\mathbb{E}\boldsymbol{\xi}$ (для простых СЛВ):
1) $\xi \geq 0 \Rightarrow \ \mathbb{E}\xi \geq 0$;
2) $\mathbb{E}(a\xi + b\eta) = a\mathbb{E}\xi + b\mathbb{E}\eta$;
3) $\xi \geq \eta \ (\forall\omega) \Rightarrow \mathbb{E}\xi \geq \mathbb{E}\eta$;
4) $|\mathbb{E}\xi| \leq \mathbb{E}|\xi|$;
5) ξ, η – незав $\Rightarrow \mathbb{E}(\xi\eta) = \mathbb{E}\xi \times \mathbb{E}\eta$.

Также число $E\xi = \int_{-\infty}^{+\infty} x \, dF_\xi(x)$, $F_\xi(x)$ – ФР ξ.

Задача о пенсионере: человек тратит на перемещение из $a \to b$ усилия $(b-a)^2$. Пусть X – коорд. точки, в кот. нужно попасть. Где он должен жить, чтобы мин. усилия на перемещение?
[Средние усилия на перем. из неслуч. $x \to$ случ. X им. вид: $\mathbb{E}(X-x)^2 = x^2 - 2x\mathbb{E}X + \mathbb{E}X^2$ – min в $x_0 = \mathbb{E}X$.] //Королев, с.41

20 Медиана, свойства. Задача о предпринимателе

Медианой распределения СЛВ ξ наз. $\boldsymbol{\mu}$: $\boldsymbol{P(\xi \leq \mu) \geq \frac{1}{2}}$, $P(\xi \geq \mu) \geq 1/2$. *Смысл:* точка, левее и правее которой сосредоточено ровно по 1/2 всей вер. массы (см. f(x)). Для распр. с непр. монот. F: μ – реш. $F(\mu) = 1/2$.

Задача о предпренимателе: аналог зад. о пенсионере, но усилия $= |b-a|$. Как затратить мин. усилий?
[Докажем, что минимальны в медиане μ, т.е. решение $\to$ к док-ву $\mathbb{E}|x-\mu| \leq \mathbb{E}|x-d| \ \forall d \in \mathbb{R}$.
Пусть $d > \mu \Rightarrow$

$$|x-d| - |x-\mu| = \begin{cases} \mu - d, x \geq d, \\ d + \mu - 2x, \mu < x < d, \Rightarrow \\ d - \mu, x \leq \mu. \end{cases}$$

$\mathbb{E}|x-d| - \mathbb{E}|x-\mu| = (\mu - d)\left(\mathbb{E}\mathbb{I}_{[d,\infty)}(x) + \mathbb{E}\mathbb{I}_{(\mu,d)}(x)\right) + (d-\mu)\mathbb{E}\mathbb{I}_{(-\infty,\mu]}(x) = (d-\mu)[P(x \leq \mu) - P(x \geq \mu)] \geq 0$ по опр. медианы.] //Королев, с.43

21 Задача о среднем годовом доходе

- имеется $0 < \alpha < 1$, а вероятность того, что доход выб. наугад человека $> x$ есть α^x. Каков средний доход?
[Пусть X – СЛВ, равная сред. доходу наугад выб. человека $\to F(x) = P(X < x) = 1 - \alpha^x$, $f(x) = c\alpha^x \to$ находим c из усл: $\int_0^\infty f(x)dx = 1, c = \ln 1/\alpha$, $\mathbb{E}X = c\int_0^\infty x\alpha^x dx = 1/c$. Медиана – корень ур-я $1 - \alpha^x = \frac{1}{2}$ $\to$ $\boldsymbol{\mu = \log_\alpha \frac{1}{2}}$, мода – точка с max плотностью $\to$ $f(x) = c\alpha^x - max \Rightarrow mod\ x = 0$.] // Королев, с.45

22 Дисперсия. Простейшие свойства

Моментом порядка $\boldsymbol{\alpha}$ > 0 СЛВ ξ наз-ся $\mathbb{E}\xi^\alpha$.
Центральным моментом порядка α наз. $\mathbb{E}(\xi - \mathbb{E}\xi)^\alpha$.
Дисперсией СЛВ наз. центральный момент **порядка 2:** $\mathbb{D}\xi = \mathbb{E}(\xi - \mathbb{E}\xi) = \mathbb{E}\xi^2 - (\mathbb{E}\xi)^2$. *Смысл:* хар-т степень разброса значений СЛВ вокруг её МО (центра тяжести).
Простейшие св-ва:
1) $\mathbb{D}(c - \xi) = c^2\mathbb{D}\xi$; $\mathbb{D}(\xi + c) = \mathbb{D}\xi$;
2) $\mathbb{D}\xi \geq 0$; $\mathbb{D}\xi = 0 \Leftrightarrow \exists c$: $\mathbb{D}(\xi = c) = 1$;
3) Если ξ, η – незав. $\Rightarrow \mathbb{D}(\xi + \eta) = \mathbb{D}\xi + \mathbb{D}\eta$;
4) Для $\forall \xi, \eta$ с конеч. 2ми моментами: $\mathbb{D}(\xi + \eta) = \mathbb{D}\xi + \mathbb{D}\eta + 2(\mathbb{E}(\xi\eta) - \mathbb{E}\xi\mathbb{E}\eta)$;
5) минимум СКО СЛВ от точек числовой прямой есть СКО от её МО: $\min_a \mathbb{E}(\xi - a)^2 = D\xi$.

Среднеквадратическое отклонение $\boldsymbol{\sigma}$ $= \sqrt{\mathbb{D}\xi}$.

23 Ковариация. Коэффициент корреляции

Ковариацией $\boldsymbol{cov}$ СЛВ ξ, η наз. число $cov(\xi, \eta) = \mathbb{E}((\xi - \mathbb{E}\xi)(\eta - \mathbb{E}\eta)) = \mathbb{E}(\xi\eta) - \mathbb{E}\xi\mathbb{E}\eta$. *Смысл:* исп. как *индикатор* зависимости СЛВ (т.к если СЛВ незав. то по 3му св-ву дисперсии $cov = 0$) (зав-ть при $cov \neq 0$).
Св-во: $D(\xi_1 + \dots + \xi_n) = \sum_{i,j} cov(\xi_i, \xi_j)$.

Коэффициентом корреляции ρ СЛВ ξ,η: $\exists\mathbb{D} > 0$ наз-ся число $\rho(\xi,\eta) = \frac{cov(\xi,\eta)}{\sqrt{\mathbb{D}\xi}\sqrt{\mathbb{D}\eta}}$ – та же ковариация, нормир. так, что при × на число не меняется её абс. значение.
Свойства: 1) Если ξ,η незав. $\Rightarrow \rho(\xi,\eta) = 0$;
2) всегда $|\rho(\xi,\eta)| \le 1$; $= 1 \Leftrightarrow P(\eta = a\xi + b) = 1$.

24 Неравенства Маркова и Чебышева

Т. (нер-во Маркова) Если $\mathbb{E}|\xi| < \infty \Rightarrow \forall x > 0$

$$P(|\xi| \ge x) \le \frac{\mathbb{E}|\xi|}{x}.$$

Т. (нер-во Чебышева) Если $\exists\mathbb{D}\xi \Rightarrow \forall x > 0$

$$\mathrm{P}(|\xi - \mathbb{E}\xi| \ge x) \le \frac{\mathbb{D}\xi}{x^2}.$$

Смысл: для оценки $P(|\xi_n - \xi| \ge \varepsilon)$ сверху при больших n ($\rightarrow$для док-ва сх-ти по вероятности устремляем к 0).

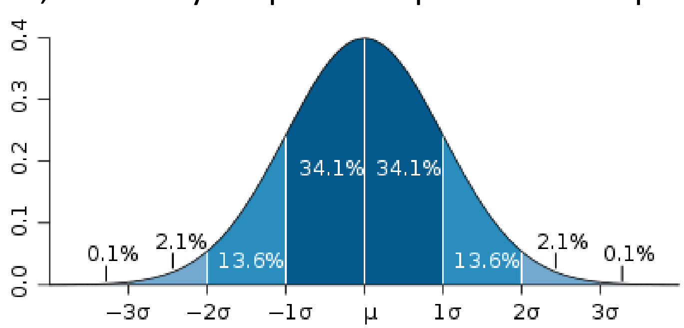
25 Неравенство Гаусса. «Правило трех сигм»

Для унимодального распр. (обл. ед. модой – лок. max плотности распр.) с модой μ, верно **нер-во Гаусса:**

$$\mathrm{P}(|\xi - \mu| \ge x) \le \frac{4}{3}\frac{\mathbb{D}\xi}{\varepsilon}, \varepsilon \ge \frac{2\mathbb{D}\xi}{\sqrt{3}}.$$

(равенство достиг. на равномерном распред.)

Правило 3σ: практически все значения нормально распределённой СЛВ лежат в инт-ле $[\bar{x} - 3\sigma, \bar{x} + 3\sigma]$. Более строго — не менее чем с 99,7 % достоверностью значение нормально распределенной случайной величины лежит в указанном интервале (при усл., что $\bar{x}$ истинная, а не получ. в рез-те обработки выборки).



26 Закон больших чисел в форме Чебышева

Посл-ть СЛВ с конеч. моментами удовл. **ЗБЧ,** если

$$\frac{\xi_1+\dots+\xi_n}{n} - \frac{\mathbb{E}\xi_1+\dots+\mathbb{E}\xi_n}{n} \to^p 0.$$

Если сх-ть «почти наверное» → **УЗБЧ.**

Т. (ЗБЧ Чебышева) Для $\forall$посл-ти $\xi_1,\xi_2,\dots$ попарно незав. и одинак. распр. СЛВ с конеч. 2м моментом $\mathbb{E}\xi_1^2 = \sigma^2 < \infty$ им. место сх-ть

$$\frac{\xi_1+\dots+\xi_n}{n} \to^p \mathbb{E}\xi_1, \text{ или } P\left(\left|\frac{\sum\xi_i}{n} - \mathbb{E}\xi_1\right| \ge \varepsilon\right) \to 0.$$

Смысл: среднее арифметическое большого числа слаг. СЛВ «стабилизируется» с ростом их числа – отклонения СЛВ «взаимно гасятся».

27 Неравенства Иенсена и Ляпунова

Т (нер-во Йенсена) Пусть ф-ция $g(\xi) \in \mathbb{R}$ выпукла $\Rightarrow$ для $\forall$СЛВ ξ с конеч. 1м моментом $\mathbb{E}g(\xi) \ge g(\mathbb{E}\xi)$. Для вогнутых ф-ций знак меняется.

$$\rightarrow \quad \mathbb{E}e^{\xi} \ge e^{\mathbb{E}\xi}, \mathbb{E}\xi^2 \ge (\mathbb{E}\xi)^2, \mathbb{E}|\xi| \ge |\mathbb{E}\xi|,$$
$$\mathbb{E}\ln\xi \le \ln(\mathbb{E}\xi), \mathbb{E}\frac{1}{\xi} \ge \frac{1}{\mathbb{E}\xi}, \mathbb{E}\sqrt{\xi} \le \sqrt{\mathbb{E}\xi}.$$

Т. (нер-во Ляпунова)

$$\left(\mathbb{E}|\xi|^{\beta}\right)^{1/\beta} \ge (\mathbb{E}|\xi|^{\alpha})^{1/\alpha}.$$

28 Коэффициенты эксцесса и асимметрии

Модой абс. непр. распр-я наз. $\forall$ точку лок. max $f(x)$. Распределение с ед. модой – **унимодальное.** Для описания «островершинности», «наклона» плотности УМ распред. используют след. характеристики:
Коэффициент ассиметрии с конечным 3м моментом:

$$\boldsymbol{\beta_1} = \mathbb{E}\left(\frac{\xi - a}{\sigma}\right)^3, a = \mathbb{E}\xi, \sigma = \sqrt{\mathbb{D}\xi}.$$

У симм. распр. $\beta_1 = 0$; если $\beta_1 > 0 \rightarrow$ плотность имеет более крутой наклон слева, пологий справа.
Коэффициент эксцесса с конеч. 4м моментом:

$$\boldsymbol{\beta_2} = \mathbb{E}\left(\frac{\xi - a}{\sigma}\right)^4 - 3.$$

Для всех норм. распр. $\beta_2 = 0 \colon \mathbb{E}\eta^4 = 3$.
При $\beta_2 > 0 \to$ плотность имеет более острую вершину, чем у N_{a,σ^2}; $\beta_2 < 0 \to$ более плоскую.

29 Схема Бернулли. Биномиальное распределение

Схема Бернулли – посл-ть незав. испытаний, в каждом из которых возм. «успех» или «неудача», $P_{\text{успеха}} = p$, $P_{\text{неудачи}} = 1 - p = q$. Тут $\Omega = \{(a_1, \dots) | a_i \in \{1,0\}\}$.

Т. (ф-ла Бернулли) Пусть v_n – число успехов в n испытаниях $\to P(v_n = k) = C_n^k p^k q^{n-k}$.
[Из независимости исп. $\Rightarrow P(A) = p^k q^{n-k}$.
C_n^k – кол-во расположения k успехов на n местах.]
СЛВ $\boldsymbol{\xi \sim Bi(n,p)}$, если $P(\xi = k) = C_n^k p^k q^{n-k}$.

30 Теорема Пуассона. Распр. Пуассона. Задача о булочке

Т. (Пуассона) Пусть $n \to \infty, p_n \to 0 \colon np_n \to \lambda > 0 \Rightarrow \forall k \geq 0$ вер-ть получить k успехов в n исп. Бернулли с $P_{\text{успеха}} = p_n$: $P(v_n = k) = C_n^k p^k q^{n-k} \to \frac{\lambda^k}{k!} e^{-\lambda}$.
СЛВ $\boldsymbol{\xi \sim Pois(\lambda)}$, если $P(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda}$
– распределение числа редких событий.

Задача о булочке: по ГОСТу изюмин в булочке 5шт. Какова вер-ть купить булочку без изюминок?
$n \to \infty$ - число булочек в чане с тестом;
$p_n \to 0$ – вер-ть изюмины в булочке

$$\Rightarrow P(\xi_k = 0) = e^{-5} \frac{5^{10}}{10!} \to e^{-5}.$$

31 Задача об инсектициде

Имеется поле с инсектицидом (n_1 жуков) и жилое поле (n_2 жуков). Нужно чтобы $n_1 \leq n_2$.
N_1, N_2 – реализация СЛВ n_1, n_2 соотв. и $N = Pois(\lambda)$.
Бьем поле на большое число маленьких так, что в отдельном квадрате мб только 1 жук.
Сравним $\lambda_{1,2}$: $\lambda_1 < \lambda_2 \to$ инсектицид действует, $\lambda_1 = \lambda_2$ – не действует. Обозначим $n = n_1 + n_2$. Найдем $P(N_1 = k | N_1 + N_2 = n)$. Считаем $N_{1,2}$ – незав. $\to P = \frac{P(N_1 = k) P(N_2 = n-k)}{P(N_1 + N_2 = n)} = C_n^k \left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^k \left(1 - \frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^{n-k} \Rightarrow$
$N_1 | N_1 + N_2 = n \sim Bi\left(n, \frac{\lambda_1}{\lambda_1 + \lambda_2}\right)$.
Гипотеза H_0: $\lambda_1 = \lambda_2 \to Bi\left(n, \frac{1}{2}\right) \to$
$P_0(N_1 \leq n_1 | N_1 + N_2 = n) = \frac{1}{2^n} \sum_{k=0}^{n1} C_n^k < \varepsilon$.
Если вер-ть очень мала, но мы ее наблюдаем $\to H_0$ неверна. Если не очень мала, вывод можем сделать: будем считать вер-ть невероятного события $\sim 10^{-5}$.

32 Теорема Муавра—Лапласа. Задача о докторе Споке

Как следствие ЦПТ имеем:
Т. (предельная т. Муавра-Лапласа) Пусть A может произойти в $\forall n$ незав. испытаний с одинак. вероятностью p и $v_n(A)$ – число успехов $\Rightarrow$

$$\frac{v_n(A) - np}{\sqrt{np(p-1)}} \Rightarrow N_{0,1} \text{ при } n \to \infty,$$

т.е. $\forall x < y \in \mathbb{R}$:

$$P\left(x \leq \frac{v_n(A) - np}{\sqrt{np(p-1)}} \leq y\right) \to \Phi(y) - \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_x^y e^{-\frac{t^2}{2}} dt.$$

[Тут $\mathbb{E}\xi_1 = p, \mathbb{D}\xi_1 = p(1-p), v_n = \xi_1 + \dots + \xi_n \to$ ЦПТ.]

Суд на доктором Споком: В суде заседают 12 человек, которых выб. из n=300 человек (из них 90 женщин). Подсудимый доктор Спок, подает протест, считая что женщин слишком много. Прав ли он?
[Строим гипотезу: $P(\xi \leq 90) = \frac{1}{2}, \xi \sim Bi(300, p) \to H_0$: вер-ть попасть женщине в число 300 – 1/2. Пусть X – СЛВ, равная числу женщин. Имеем $P(X \leq 90) = P\left(\frac{X - 150}{\sqrt{75}} \leq \frac{90 - 150}{\sqrt{75}}\right) \approx \Phi\left(\frac{90-150}{\sqrt{75}}\right) < 2 \cdot 3 \cdot 10^{-12} \Rightarrow$ Спок прав.]

33 Нормальное распределение. Его коэффициент эксцесса

СЛВ ξ им. **нормальное распр.** $\boldsymbol{\xi \sim N_{a,\sigma^2}}$ $(a \in \mathbb{R}, \sigma > 0)$, если ξ им. плотность $f_\xi(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\{-\frac{(x-a)^2}{2\sigma^2}\}, x \in \mathbb{R}$.

Для него **коэфф. эксцесса** $\boldsymbol{\beta_2} = \mathbb{E}\left(\frac{\xi - \mathbb{E}\xi}{\sigma}\right)^4 - 3 = \mathbf{0}$.

(т.к. 4й момент N_{a,σ^2}: $\mathbb{E}\left(\frac{\xi-\mathbb{E}\xi}{\sigma}\right)^4 = 3$) //чернова (с.99 61)

Ф-ция распр. $\xi \sim N_{a,\sigma^2}$: $\Phi(x) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{(t-a)^2}{2\sigma^2}}dt$.

Св-ва: 1) $\Phi(x) = \Phi_{0,1}\left(\frac{x-a}{\sigma}\right)$;

2) $P(x_1 < \xi < \xi_2) = \Phi(x_2) - \Phi(x_1)$;

34 Нормальное распределение. Неравенства для хвостов

?1//2

35 Задача о рейтинге. Решение по неравенству Чебышева

Пусть $x_1, \dots, x_n$ – мнения n человек.

$x_i = \begin{cases} 1, p \\ 0, 1-p \end{cases}$, $\mathbb{E}x_i = p$, $\mathbb{D}x_i = p(1-p)$.

Рейтинг Пупкина – вер-ть того, что след. человек проголосует за него. $\left|\frac{1}{n}\sum x_i - p\right| \to 0$, сколько человек проголосует за Пупкина, если $P\left(\left|\frac{1}{n}\sum x_i - p\right| \le \varepsilon\right) \ge \gamma$? //сколько нужно чел, чтобы с вер-тью $\ge \gamma$, относительная частота голоса $\frac{1}{n}\sum x_i$ отлич. от абс. вер-ти отдать голос p, не более чем на ε?

Реш. по нер-ву Чебышева ($P(|\xi - \mathbb{E}\xi| \ge \varepsilon) \le \frac{\mathbb{D}\xi}{\varepsilon^2}$):

$P\left(\left|\frac{1}{n}\sum x_i - p\right| \le \varepsilon\right) \ge 1 - \frac{1}{\varepsilon^2}\mathbb{D}\frac{1}{n}\sum x_i = 1 - \frac{p(1-p)}{n\varepsilon^2} \ge \gamma \Rightarrow n \ge \frac{1}{4\varepsilon^2(1-\gamma)}$.

Пусть $\gamma = 0{,}95$, $\varepsilon = 10^{-3}$ $\to$ $n \ge 5 \cdot 10^6 = 1{,}96 \approx 2$ и $n \ge 10^6$.

36 Задача о рейтинге. Реш. по теореме Муавра—Лапласа

Решение через т. М-Л:

$P\left(\left|\frac{1}{n}\sum x_i - p\right| \le \varepsilon\right) = P\left(\left|\frac{\sum x_i - np}{\sqrt{np(1-p)}}\right| \le \frac{\varepsilon\sqrt{n}}{\sqrt{p(1-p)}}\right) \approx$

$2\Phi\left(\frac{\varepsilon\sqrt{n}}{\sqrt{p(1-p)}}\right) - 1 \ge \gamma \to \Phi\left(\frac{\varepsilon\sqrt{n}}{\sqrt{p(1-p)}}\right) \ge \frac{\gamma+1}{2}$.

Пусть $\alpha \in (0,1)$: $\Phi(u_\alpha) = \alpha$, u_α – **квантиль уровня** $\boldsymbol{\alpha}$.

$\frac{\varepsilon\sqrt{n}}{\sqrt{p(1-p)}} \ge u_{\frac{\gamma+1}{2}} \to n \ge \frac{p(1-p)}{\varepsilon^2}u^2_{\frac{\gamma+1}{2}} \Rightarrow n \ge u^2_{\frac{\gamma+1}{2}}/4\varepsilon^2$.

(учли, что $\max p(1-p) = \frac{1}{4}$ при $p = \frac{1}{2}$).

Пусть $\gamma = 0{,}95, \varepsilon = 10^{-3} \Rightarrow \frac{\gamma+1}{2} = u_{0{,}975}$.

37 Геом. распределение. Его моменты

СЛВ ξ им. **геом. распределение** $\boldsymbol{\xi \sim G_p}$, если ξ принимает знач. в вер-тями $P(\xi = k) = p(1-p)^{k-1}$. Смысл СЛВ – номер первого успешного исп. Бернулли.

$\mathbb{E}\xi = \sum_{k=1}^{\infty} kpq^{k-1} = p\frac{d}{dq}\left(\sum_{k=1}^{\infty} q^k\right) = p\frac{d}{dq}\left(\frac{q}{1-q}\right) = \frac{1}{p}$;

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

38 Предельная теорема для геом. распределения. Показательное распределение. Его моменты

СЛВ ξ им. **показательное распр.** $\boldsymbol{\xi \sim E_\alpha}$ $(\alpha > 0)$, если ξ им. плотность $f_\xi(x) = \begin{cases} 0, x < 0, \\ \alpha e^{-\alpha x}, x \ge 0. \end{cases}$

$$\mathbb{E}\xi^k = \int_{-\infty}^{\infty} x^k f_\xi(x)dx = \int_0^{\infty} x^k \alpha e^{-\alpha x}dx = \frac{k!}{\alpha^k}.$$

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

39 Показательное распр. Св-во отсутствия последействия

Ф-ция распр. для $\boldsymbol{\xi \sim E_\alpha}$: $F_\xi(x) = \begin{cases} 0, x < 0, \\ 1 - e^{-\alpha x}, x \ge 0 \end{cases}$.

- для него вып. **св-во отсутствия последействия:**

$$\xi \in \mathbb{E}_\alpha \Rightarrow \forall x, y > 0 \; P(\xi > x + y | \xi > x) = P(\xi > y).$$

Смысл: вероятность появления событий в любом промежутке времени не зависит от того, появлялись или не появлялись они до начала рассм. промежутка.

40 Совместные распределения. Совместные плотности

Ф-ция $P(\xi \in B) = P\{\omega \big| (\xi_1(\omega), \dots, \xi_n(\omega)) \in B\}$ опр. для $\forall B \in \mathcal{B}_n$ наз. распред. сл. вектора ξ или **совместным распр. СЛВ** $\xi_1, \dots \xi_n$. ($\mathcal{B}_n$ – бор. σ-алг. подмн-в $\mathbb{R}^n$)

Ф-ция $F(x_1, \dots, x_n) = P(\xi_1 < x_1, \dots, \xi_n < x_n)$ опр. для $\forall x_i \in \mathbb{R}$ наз. **совместной ф-цией распред.** СЛВ $\{\xi_n\}$.

Если $\exists f(x_1, \dots, x_n) \ge 0 : \forall B \in \mathcal{B}_n \;\; P(\xi \in B) = \int_B f(x)dx$, то распр. наз. абс. непр., а $f(x)$ – **совм. плотностью.**

Маргинальные (частные) распр. СЛВ $\boldsymbol{\xi_i}$ – компонент ξ – ищут след. образом: $F_i(x_i) = F(+\infty, \dots x_i, \dots + \infty)$.

Смысл: совместное распр. нужно для отыскания ф-ций распр. от неск. СЛВ, ибо его невозможно опр. только по распр. слагаемых.

41 Условные плотности и условные МО

Пусть ξ, η – две СЛВ (мб многомерные) и $L(\eta)$ – мн-во, в кот. собраны все СЛВ $g(\eta)$ – борел. ф-ция. Введем ск.П. как $(\varphi, \eta) = \mathbb{E}(\varphi\eta)$, если МО существует.

Условное МО $\mathbb{E}(\boldsymbol{\xi}, \boldsymbol{\eta})$ СЛВ $\boldsymbol{\xi}$ отн. $\boldsymbol{\eta}$ можно предст. как рез-т проектир-я СЛВ ξ на L - СЛВ $\hat{\xi}$: $\mathbb{E}\left(\left(\xi - \hat{\xi}\right)g\right) = 0$ – **тождество ортопроекции** (→УМО опр. неоднозначно).

Св-ва:

1) $\mathbb{E}\xi^2 < \infty \Rightarrow \mathbb{E}\left(\xi - \hat{\xi}\right)^2 = \min \mathbb{E}\left(\xi - g(\eta)\right)^2$;
2) $\mathbb{E}|f(\eta)\xi| < \infty \Rightarrow \mathbb{E}(f(\eta)\xi|\eta) =^{\text{п.н.}} f(\eta)\mathbb{E}(\xi|\eta)$;
3) $f(\eta) \in L, \mathbb{E}|f(\eta)| < \infty \Rightarrow \mathbb{E}(f(\eta)|\eta) =^{\text{п.н.}} f(\eta)$;
4) $\mathbb{E}\xi = \mathbb{E}[\mathbb{E}(\xi|\eta)]$, *т.е.* $\mathbb{E}\xi = \mathbb{E}\hat{\xi}$;
5) ξ, η – незав. СЛВ $\Rightarrow \mathbb{E}(\xi|\eta) = \mathbb{E}\xi$.

Хотим найти $h(y) = \mathbb{E}(\xi|\eta = y)$:

1) если ξ, η – дискр. $h(y) = \sum_i a_i P(\xi = a_i | \eta = y)$ – условное распр. ξ при усл. $\eta = y$.
2) если ξ, η – абс. непр. и $f_{\xi,\eta}(x, y)$ – совм. плотность, то $h(y) = \int_{\mathbb{R}} x \frac{f_{\xi,\eta}(x,y)}{f_\eta} dx$ – при фикс. y, $h(y)$ – МО усл. распр. с **усл. плотностью $\boldsymbol{f(x|y)}$** $= \frac{f_{\xi,\eta}(x,y)}{f_\eta(y)}$.

42 Виды сходимости СЛВ. Соотношения между ними

Пусть задано ВП $(\Omega, \mathcal{F}, P)$ и посл-ть СЛВ $\xi, \xi_1, \xi_2, \dots$

Виды сходимости к СЛВ ξ при $n \to \infty$:

почти наверное - $\boldsymbol{\xi \to^{\text{п.н.}} \xi_n}$ - $P\{\omega | \xi_n(\omega) \to \xi(\omega)\} = 1$;
по вероятности - $\boldsymbol{\xi \to^P \xi}$ - $\forall \varepsilon > 0 \; P(|\xi_n - \xi| > \varepsilon) \to 0$;
в среднем порядка $\boldsymbol{r}$ - $\boldsymbol{\xi_n \to^{(r)} \xi}$ – $\mathbb{E}|\xi_n - \xi|^r \to 0$;
по распределению - $\boldsymbol{\xi_n \to^d \xi}$ - $\{F_{\xi n}(x)\} \to F_\xi(x)$;
слабо - $\boldsymbol{\xi_n \Rightarrow \xi}$ - $\mathbb{E}f(\xi_n) \to \mathbb{E}f(\xi)$;

Соотношения:

$$\begin{array}{ccccccc}
\boldsymbol{\xi \to^{\text{п.н.}} \xi_n} & \Lsh & & & \boldsymbol{\xi_n \Rightarrow \xi} & \leftrightarrow & F_n \Rightarrow F \\
& & \boldsymbol{\xi \to^P \xi} & - & \updownarrow & & \updownarrow \\
\boldsymbol{\xi_n \to^{(r)} \xi} & \Rsh & & & \boldsymbol{\xi_n \to^d \xi} & \leftrightarrow & \varphi_n \to \varphi
\end{array}$$

43 Характеристические функции. Их простейшие свойства

Характеристической ф-цией СЛВ ξ наз. $\varphi_\xi : \mathbb{R} \to \mathbb{C}$, опр. $\forall t \in \mathbb{R}$: $\varphi_\xi(t) = \mathbb{E}e^{it\xi} = \int e^{itx} dF_\xi(x)$. Также наз. **преобразованием Фурье-Стилтьеса** и $= \int e^{it\xi} f_\xi(x)dx$.

Существует всегда. **Основные свойства:**

1) $\forall \xi \; \varphi_\xi(0) = 1, \left|\varphi_\xi(t)\right| \le 1 \; \forall t$;
2) $\forall \xi, a, b \; \varphi_{\xi a + b}(t) = e^{itb}\varphi_\xi(at)$;
3) если ξ, η – незав. $\Rightarrow \varphi_{\xi+\eta}(t) = \varphi_\xi(t)\varphi_\eta(t)$;
4) $\varphi_\xi(t)$ – равномерно непрерывная;
5) $\mathbb{E}|\xi|^k < \infty, k \ge 1 \Rightarrow \exists \varphi_\xi^{(k)}(t) : \varphi_\xi^{(k)}(t)|_{t=0} = i^k \mathbb{E}\xi^k$;
6) компл.-сопряженная $\bar{\varphi}_\xi(t) = \varphi_\xi(-t) = \varphi_{-\xi}(t)$.

Т. Хар. ф-ция взаимноодн. опред. её распределение.

44 Теорема непрерывности соответствия между распределениями и хар. функциями (без док-ва)

Т. (непрерывности) Слабая сх-ть посл-ти СЛВ $\{\xi_n\} \to \xi$ эквивалентна сх-ти посл-ти соотв. характеристических ф-ций $\{\varphi_{\xi n}(t)\} \to \varphi_\xi(t)$ при $n \to \infty$ в кажд. $t \in \mathbb{R}$.
//обычно исп. для док-ва слабой сх-ти

45 ЗБЧ в форме Хинчина

Т. (ЗБЧ Хинчина) Для $\forall$посл-ти $\xi_1, \xi_2, \ldots$ незав. в совокупности и одинак. распр. СЛВ с конеч. 1м моментом $\mathbb{E}|\xi_1| < \infty$ им. место сх-ть

$$\frac{\xi_1 + \cdots + \xi_n}{n} \to^p \mathbb{E}\xi_1.$$

***Т. (УЗБЧ Колмогорова)** Выполнен УЗБЧ (сх-ть почти наверное) $\Leftrightarrow$ существование конечного $\mathbb{E}\xi_1$.

46 ЦПТ

Возьмем ЗБЧ Чебышёва: $\frac{S_n}{n} \to^p \mathbb{E}\xi_1$ или $\frac{S_n - n\mathbb{E}\xi_1}{n} \to^p 0$. Хотим домножить на что-то растущее, чтобы «погасить» стремление к нулю, получить что-то конечное. Оказывается: если $\times \sqrt{n}$

Т. (ЦПТ Ляпунова) Для $\forall$посл-ти $\xi_1, \xi_2, \ldots$ незав. и одинак. распр. СЛВ с конеч. дисперсией $0 < \mathbb{D}\xi_1 < \infty$ имеет место *слабая сходимость:* $\frac{S_n - n\mathbb{E}\xi_1}{\sqrt{n\mathbb{D}\xi_1}} \Rightarrow N_{0,1}$ посл-ти центрир. и нормир. сумм СЛВ к $N_{0,1} = \int_{-\infty}^{x} e^{-u^2/2} du$.

47 ЦПТ как оценка скорости сходимости в ЗБЧ

Если вып. ЦПТ, то с**корость сх-ти в ЗБЧ – порядка $\frac{1}{\sqrt{n}}$**.

В случае, если $\exists\mathbb{E}, \nexists\mathbb{D}$ (вып ЗБЧ, ЦПТ не вып.) – $n^{-\frac{\delta}{2}}$.

[Введем $S_n = \xi_1 + \cdots + \xi_n$, $a = \mathbb{E}\xi_1$.

Из ЗБЧ Чебышева $\frac{S_n}{n} - a \to^p 0$. Рассм. $\eta_n = \left|\frac{S_n - na}{n}\right|$.

Ищем такие $c_n \in \mathbb{R}$: $\frac{\eta_n}{c_n} \to^d \xi$, ξ – невыр. в 0 и конечна.

Пусть $c_n = \frac{1}{\sqrt{n}} \to P\left(\frac{\eta_n}{c_n} < x\right) = P\left(\left|\frac{S_n - na}{\sigma\sqrt{n}}\right| < \frac{x}{\sigma}\right) \to_{n\to\infty}$ $\Phi\left(\frac{x}{\sigma}\right) - \Phi\left(-\frac{x}{\sigma}\right) = 2\Phi\left(\frac{x}{\sigma}\right) - 1$ – что-то конечное $\Rightarrow$ такая СЛВ нам подходит.]

48 Условие Линдеберга. Теорема Линдеберга—Феллера

Рассм. ЦПТ в случае, когда НСЛВ , $\xi_1, \xi_2, \ldots$ имеют разные распределения. Введем $a_i = \mathbb{E}\xi_i, b_i = \mathbb{D}\xi_i$, $A_n = a_1 + \cdots + a_n = \mathbb{E}S_n, B_n^2 = b_1^2 + \cdots + b_n^2 = \mathbb{D}S_n$ $\to$ ЦПТ имеет вид **(*)** $P\left(\frac{S_n - A_n}{B_n} < x\right) \to \Phi(x)$, но если $S_n = \xi_1 + \theta_2 + \cdots + \theta_n$ (почти 0), то $S_n \sim \xi_1$ $\Rightarrow$ необходимо доп. условие:

Т. (Линдеберга-Феллера)

Если вып. **усл. Линдеберга:** $\forall\tau > 0$

$\frac{1}{B_n^2}\sum_{i=1}^{n} \int_{|x - a_i| > \tau B_n} (x - a_i)^2 dF_i(x) \to 0$ при $n \to \infty$, где $F_i(x) = P(\xi_i < x) \Leftrightarrow$

1) вып. ЦПТ: $P\left(\frac{S_n - A_n}{B_n} < x\right) \to \Phi(x) = \int_{-\infty}^{x} e^{-u^2/2} du$;

2) $\max_{1\le i\le n} P\left(\frac{|\xi_i - a_i|}{B_n}\right) \to 0\ \forall\varepsilon > 0$. (из н-ва Чебышева)

(если СЛВ одинаково распределены, то усл. сущ. инт-ла ~ усл. сущ. дисперсии).

Практический вывод: если СЛВ одинак. распр., то $\exists\mathbb{D}$ $\Leftrightarrow$ вып. ЦПТ (иначе в пределе не получ. норм. распр.).

49 Теор. Ляпунова. Связь (у) Ляпунова и (у) Линдеберга

Введем $m_i^3 = \mathbb{E}|\xi_i - a_i|^3, M_n^3 = m_1^3 + \dots + m_n^3,$
$a_i = \mathbb{E}\xi_i, b_i = \mathbb{D}\xi_i, A_n = a_1 + \dots + a_n = \mathbb{E}S_n,$
$B_n^2 = b_1^2 + \dots + b_n^2 = \mathbb{D}S_n \rightarrow B_n^3 = \left(\sqrt{B_n^2}\right)^3.$
Ляпуновской дробью называется

Т. Если **ляпуновская дробь** $\frac{M_n^3}{B_n^3} \to 0 \Rightarrow P\left(\frac{S_n - A_n}{B_n} < x\right) \to \Phi(x) \Rightarrow$ **условие Ляпунова сильнее** усл. Линдеберга.

Усл. Ляпунова вып. для одинак. распр. СЛВ ($\frac{M_n^3}{B_n^3} = \frac{m^3}{b^3\sqrt{n}}$).

50 Оценки скорости сходимости в ЦПТ

Можем оценить погрешность приближения в ЦПТ:
Т. (нер-во Берри-Эссеена) В усл. ЦПТ $\forall x \in \mathbb{R}$ и $\forall$ распределения ξ_1 с конечным 3м моментом:

$$\sup_x \left|P\left(\frac{S_n - n\mathbb{E}\xi_1}{\sqrt{n\mathbb{D}\xi_1}} < x\right) - \Phi(x)\right| \le C_0 \cdot \frac{\mathbb{E}|\xi_1 - \mathbb{E}\xi_2|^3}{\sqrt{n}\left(\sqrt{\mathbb{D}\xi_1}\right)^3}.$$

Есть такая $C_0 \ge 0{,}4784$, что $\forall S_n \downarrow 0$ $\exists$ НОРСВ, что sup убывает медленнее S_n. **Оценка скорости сх-ти ЦПТ:**

$$\left|P\left(\frac{S_n - A_n}{B_n} < x\right) - \Phi(x)\right| \le C \cdot \frac{m^3}{\left(\sqrt{\mathbb{D}\xi_1}\right)^3 \sqrt{n}(1+|x|)^3}, C \le 32.$$

Математическая статистика

51 **Статистическая структура**

- так называется совокупность $(\Omega, \mathcal{A}, P)$, где Ω – мн-во элем. исходов, $\mathcal{A}$ – σ-алгебра событий, P – семейство вер. мер на $\mathcal{A}$. Задача МС – на основе проведения неск. раз эксперимента, зная как он заканчивался, выбрать какую-либо вер-ть из P или сузить P.

Проведя эксперимент n раз, получим знач. $X_1, \dots X_n$ СЛВ $\xi: \Omega \to \mathbb{R}$, где $\mathcal{F}_\xi$ частично или полн. неизвестно.
Выборкой объема n из распределения $\mathcal{F}$ наз-ся набор $\bar{X} = (X_1, \dots, X_n)$ из НОСЛВ, им. распр. $\mathcal{F}$.

Вариационный ряд – ряд из СЛВ выборки, упоряд. по возр. на каждом элем. исходе: $X_{(1)} \le \dots \le X_{(n)}$, где $X_{(1)} = \min\{X_i\}, X_{(n)} = \max\{X_i\}$, а $X_{(k)}$ наз-ся k-тым членом ВР или **k-й порядковой статистикой.**

52 Эмпирическая функция распределения

Рассмотрим **реализацию** выборки на одном элем. исходе – числа $X_i = X_i(\omega_0)$, $i = 1..n$. Разыграем новую СЛВ ξ^*, кот. принимает $X_1, \dots, X_n$ с *одинак. вер-тями:*

ξ^*	X_1		X_n
P^*	1/n	…	1/n

Распределение ξ^* наз. **эмпирическим**

Пусть ω_0 меняется $\rightarrow$ все величины будут случайными.
Тогда оценка истинной ф/распр. $F(y) = P(X_1 < y)$ - **эмпирическая ф-ция распред.,** постр. по выборке $\bar{X}$ объема n – случайная ф-ция $F_n^*: \mathbb{R} \times \Omega \to [0,1]$, т.ч. $F_n^*(y) = \frac{\text{колич.} X_i \in (-\infty, y)}{n} = \frac{1}{n}\sum_{i=1}^n \mathbb{I}(X_i < y)\ \ \forall y \in \mathbb{R}.$

Независимость равномерного расстояния между теоретической и эмпирической ф-циями распределения от теоретической ф-ции распределения

Т1 Пусть $X_1, \dots, X_n$ – выборка из $\mathcal{F}$ с ФР F и F_n^* - эмпирическая ФР, построенная по этой выбоке $\Rightarrow$ $F_n^*(y) \to^p F(y)$ при $n \to \infty$ $\forall y \in \mathbb{R}$.
[$\mathbb{E}I(X_1 < y) = F(y) < \infty \Rightarrow$ применяем ЗБЧ Хинча: $F_n^*(y) = \frac{1}{n}\sum_{i=1}^n \mathbb{I}(X_i < y) \to^p \mathbb{E}I(X_1 < y) = F(y).$]

53 Т. Гливенко. Т. Колмогорова. Репрезент-ть выборки

Т. (Гливенко-Кантелли) В усл. Т1 $\sup|F_n^*(y) - F(y)| \to^p 0$ при $n \to \infty$ (характер «равномерный»).
Вообще им. место сх-ть п.н. а скорость порядка $1/\sqrt{n}$:
Т. (Колмогорова) Пусть $X_1, \dots, X_n$ – выборка из $\mathcal{F}$ с ФР F и F_n^* - эмпирическая ФР $\Rightarrow \sqrt{n} \cdot \sup|F_n^*(y) - F(y)| \Rightarrow \eta$, где СЛВ η им. распр. Колмогорова с непр. ФР $K(x) = \left\{\sum_{j=-\infty}^{\infty}(-1)^j e^{-2j^2x^2}, x \ge 0;\ 0, x < 0\right\}.$

Состоятельные хар-ки – т.ч. разница между ними и истинными. хар-ками →0 с ростом объема выборки. Введенные нами хар-ки состоятельны.

Некоторые св-ва эмпирической ФР:

1) $\mathbb{E}F_n^*(y) = F(y)$, т.е. ЭФР– **несмещенная** оц. для ИФР
2) $\mathbb{D}F_n^*(y) = F(y)\big(1 - F(y)\big)/n$;
3) $\sqrt{n}\big(F_n^*(y) = F(y)\big) \Rightarrow N_{0,F(y)(1-F(y))}$ при $F(y) \neq 0{,}1$, т.е. ЭФР – **асимпт. норм.** оценка для ИФР;
4) $nF_n^* \sim Bi_{n,F(y)}$;

54 Выборочные моменты, их свойства

Все оценки явл. СЛВ, если $X_1, \dots, X_n$ – набор СЛВ.

$\mathbb{E}\xi = \mathbb{E}X_1 = a \to \bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ – выб. среднее;

$\mathbb{E}\xi^k = \mathbb{E}X_1^k = m_k \to \overline{X^k} = \frac{1}{n}\sum_{i=1}^n X_i^k$ – выб. k-й момент;

- несмещенные, сост. и асимпт. норм. оценкой для ист.

$\mathbb{D}\xi = \mathbb{D}X_1 = \sigma^2 \to S^2 = \frac{1}{n}\sum_{i=1}^n (X_i - X)^2$ – дисперсия;

- смещенная ($S_0^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - X)^2$- несмещенная), асимпт. нормальные оценки истинной дисперсии.

$\mathbb{E}g(\xi) \to \overline{g(X)} = \frac{1}{n}\sum_{i=1}^n g(X_i)$ – общ. случай.

55 Точечное оценивание. Неравенство Рао—Крамера. Эффективные оценки.

Имеется выб. $X_1, \dots, X_n$ объема n, извлеч. из **распр. $\mathcal{F}_\theta$,** кот. известным образом зависит от неизв. параметра θ.

Параметр $\theta \in \Theta$ – **мн-во всевозм. значений парам.**

Статистикой наз. ∀борелевская ф-ция $T = T(X_1, \dots, X_n)$.

//измеримая, нужна для оценки θ и от него НЕ зависит!

Некоторые св-ва оценок статистики:

- **несмещенная** – если $\mathbb{E}T = \theta\ \forall\theta \in \Theta$;

(отсутствие ошибок в среднем)

- **асимпт. несм.** - $\mathbb{E}T \to 0$ при $n \to \infty$;
- **состоятельная** - $T \to^p \theta$ при $n \to \theta$.

Оценка $T \in K_b$ (класс всех оценок со смещением $K_{b(\theta)} = \{T | \mathbb{E}T = \theta + b(\theta)\}$) наз. **эффективной,** если она лучше всех других оценок класса K_b в смысле **сред-кв. подхода:** $\forall T_1 \in K_b\ \mathbb{E}(T-\theta)^2 \le \mathbb{E}(T_1 - \theta)^2$.

Условия регулярности:

(R) Сущ. **носитель $\boldsymbol{C}$**$\subset\mathbb{R}$ (мн-во: $\forall\theta \in \Theta\ P(X_1 \in C) = 1$) семейства распр. $\mathcal{F}_\theta$: при кажд. $y \in C$ $\sqrt{f_\theta(y)} \in C^1(\theta)$.

(RR) **Инф. Фишера** $I(\theta) = \mathbb{E}\left(\frac{\partial}{\partial\theta}\ln f_\theta(X_1)\right)^2 \in C(\theta), > 0$.

Т. (Рао-Крамера) Пусть семейство распр. $\mathcal{F}_\theta$ удовл. усл. регулярности ⇒ ∀оценки $T \in K_{b(\theta)}$: $\mathbb{D}T$ – огр. на ∀компакте в Θ справедливо нер-во:

$\mathbb{E}(T-\theta)^2 \ge \frac{\left(1+b'(\theta)\right)^2}{nI(\theta)} + b^2(\theta)$, т.е. $\mathbb{D}T \ge \frac{\left(1+b'(\theta)\right)^2}{nI(\theta)}$.

Смысл: поиск эфф. оценки попарным сравнением – долго, нер-во Р-К позв. доказать эфф. (если есть) – в кажд. классе K_b ∃НГ для сред-кв. откл. ∀оценки.

56 Достаточные статистики. Критерий факторизации

Для одного неизв. парам. θ приходится хранить всю выборку (громадная). Хотим сократить хран. инф-ю, не потеряв никакие сведения о параметре.

Статистика $S = S(X_1, \dots, X_n)$ наз. **достаточной** для θ, если $\forall s, B \in \mathcal{B}(\mathbb{R}^n)$ усл. распр. $P(X_1, \dots, X_n \in B | S = s)$ не зависит. от парам. θ.

Смысл: нашли статистику – можем выкинуть выборку.

Т. (факторизационная Неймана-Фишера) Статистика S явл. достаточной ⇔ ф-ция правдоподобия

$$f(X_1, \dots, X_n; \theta) \overset{\text{п.н.}}{=} h(\bar{X}) \cdot \Psi(S, \theta),$$

где кажд. из ф-ций зав. *только* от указ. аргументов.

//Если выборка из $Bi_p, \Pi_\lambda, E_\alpha$ – ДС $S = n\bar{X}$ или $\bar{X}$.

57 Оптимальные оценки. Т. Блэкуэлла—Колмогорова

58 Полные статистики. Т. Лемана—Шеффе

Пусть $X_i \in \mathcal{F}_\theta, \theta \in \Theta$.

Статистика S наз. **полной,** если из $\mathbb{E}g(S) = 0 \Rightarrow g(S) =^{\text{п.н.}} 0$ ($g(x)$ – борелевская ф-ция).

Смысл: чтобы в классе K_b оценка (ф-ция от S) была единственна (если $\exists$), а в $\forall$ классе эфф.оценка – одна.

Т. (Лемана-Шеффе) Пусть S – полн. и дост. статистика. Если оценка $T_S \in K_b$ явл. ф-цией от $S \Rightarrow$ она эфф. в K_b.

59 ОМП. Их асимптотические свойства

- выбираем такое θ, при кот. вер-ть получить данную выборку наибольшая.

Обозначим ф-цию $f_\theta(y) = \begin{cases} \text{пл} - \text{ть}, \mathcal{F}_\theta - \text{ абс. непр.} \\ P_\theta(X_1 = y), \mathcal{F}_\theta - \text{дискр.} \end{cases}$

Ф-ция $f(\bar{X};\theta) = f_\theta(X_1) \cdot \ldots \cdot f_\theta(X_n) = \prod_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{f_\theta(X_i)}$ наз. **ф-цией правдоподобия.** *Смысл:* вер-ть с кот. выборка в данной серии прин. знач. $x_1, \ldots, x_n$ (или вокруг них).

Оценкой макс. правдоподобия $\boldsymbol{\widehat{\theta}}$ для неизв. парам. θ наз. значение θ, при кот. достигается $\max f(\bar{X};\theta)$.

Асимпт. св-ва?

60 Оценки метода моментов

Сам метод: ∀момент СЛВ X_1 явл. ф-цией от $\theta \Rightarrow$ параметр. явл. ф-цией от теор. k-го момента.

Выберем нек. $g(y)\colon \mathbb{R} \to \mathbb{R}\colon \mathbb{E}g(X_1) = h(\theta)$ и $\exists h^{-1}$.

Решим и пост. выб. момент: $\theta = h^{-1}\big(\mathbb{E}g(X_1)\big)$,

$\theta^* = h^{-1}\big(\overline{g(X)}\big) = h^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} g(X_i)\right)$ – **оценка метода моментов (ОММ).** //обычно берут $g(y) = y^k$.

Т. Оценка, полученная по ОММ – состоятельна.

61 Интервальное оценивание. Метод центральной СЛВ

Есть выборка, мы указ. интервал, накрывающий параметр с зад. наперед вероятностью – **интервальное оценивание.** (чем шире инт-л, тем выше P).

Инт-л со случ. концами (θ^-, θ^+) наз. доверительным для парам. θ уровня доверия $1 - \varepsilon$, если $\forall\theta \in \Theta$ $P(\theta^- < \theta < \theta^+) \geq 1 - \varepsilon$.

62 Интервальное оценивание точечной оценкой

Интервал наз. **точным,** когда вер-ть того, что доверительный инт-л накроет параметр = $1 - \varepsilon$.

63 Основные понятия проверки статистических гипотез

Дана выборка – набор НОРСЛВ $X_1, \ldots X_n$ из $\mathcal{F}$.

Гипотезой (H) наз. любое предполож. о распределении наблюдений: **простая** $H = \{\mathcal{F} = \mathcal{F}_1\}$ или **сложная** $H = \{\mathcal{F} \in \mathbb{F}\}$, $\mathbb{F}$ – подмн-во всех распр-й.

Критерием $\boldsymbol{\delta}$ $= \delta(X_1, \ldots, X_n)$ наз. измеримое отображ. $\delta\colon \mathbb{R}^n \to \{H_1, \ldots, H_k\}$ из мн-во знач. выб. → мн-во гип.

Произошла **ошибка i-го рода,** если кр. отверг H_i, а вероятность её ошибки $\alpha_i(\delta) = P_{H_i}(\delta(\bar{X}) \neq H_i)$.

Подходы к сравнению критериев:

- минимаксный: кр. δ не хуже ρ в смысле мм подхода, если $\max\{\alpha_1(\delta), \alpha_2(\delta)\} \leq \max\{\alpha_1(\rho), \alpha_2(\rho)\}$;

- байесовский.

64 Критерий согласия Колмогорова

Пусть $\rho(\bar{X})$ – борел. ф-ция, обл. св-вами:
(К) если H_1 верна $\rightarrow \rho(\bar{X}) \Rightarrow \mathcal{G}$ – известное непр. распр.
(КК) если H_1 неверна $\rightarrow |\rho(\bar{X})| \rightarrow^p \infty$ при $n \rightarrow \infty$.
Для данной $\eta \in \mathcal{G}$ опр. $C: \varepsilon = P(|\eta| \geq C)$ и построим

$$\delta(\bar{X}) = \begin{cases} H_1, |\rho(\bar{X})| < C, \\ H_2, |\rho(\bar{X})| \geq C. \end{cases}$$ – **критерий Согласия.**

Имеется выборка $\bar{X} = (X_1, \dots, X_n)$ из $\mathcal{F}$. Проверяется простая гипотеза $H_1 = \{\mathcal{F} = \mathcal{F}_1\}$ против сложной альтернативы $H_2 = \{\mathcal{F} \neq \mathcal{F}_1\}$. Если $\mathcal{F}_1$ им. $F_1 \in C \Rightarrow$ можем пользоваться **критерием Колмогорова:**
пусть $\rho(\bar{X}) = \sqrt{n} \sup_y |F_n^*(y) - F_1(y)|$. Тогда если СЛВ им. распр. $K(y) \rightarrow \delta(\bar{X}) = \begin{cases} H_1, \rho(\bar{X}) < C, \\ H_2, \rho(\bar{X}) \geq C. \end{cases}$

65 Критерий согласия хи-квадрат Пирсона

Область знач. предп. распр. $\mathcal{F}_1$ делят на интервалы, после чего строят ф-цию откл. ρ по разностям теор. вер-тей попадания в интервалы и эмп. частот.
Пусть $\{A_n\}$ – попарно непересек. инт-лы группировки, разб. всю обл. значений $\mathcal{F}_1$, v_j – число эл-тов выборки, попавший в $A_j: v_j = \sum_{i=1}^n I(X_i \in A_j)$ и $p_j > 0$ – теор. вер. $P_{H_1}(X_1 \in A_j)$. Пусть $\rho(\bar{X}) = \sum_{j=1}^k \frac{(v_j - np_j)^2}{np_j}$.

Т. (Пирсона) Если верна H_1 или $H_1'\{P(X_1 \in A_j) = p_j\} \Rightarrow$ при фикс. k $\rho(\bar{X}) \Rightarrow H_{k-1}, n \rightarrow \infty$.

66 Байесовский кр. различения двух простых гипотез

А

67 Лемма Неймана—Пирсона

Есть 2 гипотезы: $H_1 = \{X_i \in \mathcal{F}_1\}$, $H_2 = \{X_i \in \mathcal{F}_2\}$. Пл-ти распр-й $f_1(y), f_2(y)$. Строим ф-ции правдоп. для ФР.

Пусть вып. предположения:
(I) $\mathcal{F}_1, \mathcal{F}_2$ либо оба дискретны, либо оба абс.непр.
(II) Ф-ция $R(c) = P_{H_1}(T(\bar{X}) \geq c) \in (0, +\infty)$ по c, где R – хвост ф-ции распр. СЛВ T.

Отношение правдоподобия - $\boldsymbol{T(\bar{X})} = \frac{f_2(x_1, \dots x_n)}{f_1(x_1, \dots x_n)}$, его **критерий (КОП):** $\delta_c(\bar{X}) = \begin{cases} H_1, T(\bar{X}) < c, \\ H_2, T(\bar{X}) \geq c. \end{cases}$

Т. (лемма Н-П) Критерий отн. правдоподобия является
1) минимаксным при $c: \alpha_1(\delta_c) = \alpha_2(\delta_c)$;
2) байесовским при зад. априор. вер. r, s, если $c = {}^r/_s$;
3) НМК разм. $\varepsilon, 0 < \varepsilon \leq P_{H_1}(f_2(\bar{X}) > 0), c: \alpha_1(\delta_c) = \varepsilon$.

68 Статистический анализ нормальных выборок. Распределения хи-квадрат и Стьюдента. Т. Фишера

Основной вопрос – построение точных доверительных инт-лов для σ при неизв. a и наоборот в N_{a,σ^2}.
Берем НСЛВ $\{\xi_k\}$ со станд. $N_{0,1}$.

Распределение χ^2 Пирсона:
СЛВ $\chi^2 = \sum \xi_k^2 \sim \Gamma_{1/2, k/2}$. Обозн. H_k.

$$H_k = \begin{cases} \frac{1}{2^{\frac{k}{2}} \Gamma\left(\frac{k}{2}\right)} y^{\frac{k}{2}-1} e^{-\frac{y}{2}}, y \geq 0, \\ 0, y \leq 0. \end{cases} \quad \mathbb{E}\chi^2 = k, \mathbb{D}\chi^2 = 2k.$$

Распределение Стьюдента:
СЛВ $t_k = \xi_0 / \sqrt{\frac{\xi_1^2 + \dots + \xi_k^2}{k}}$ – k степ. свободы, обозн. T_k.
$T_k = \frac{\xi}{\sqrt{\chi^2/k}}, \xi \in N_{0,1}$ и $\chi^2 \in H_k$ – независимы.

Т. (Фишера) Пусть $\bar{X}$ сост. из НСЛВ с $N_{0,1}$, C – ортогональная м-ца, $\bar{Y} = C\bar{X} \Rightarrow \forall k = 1..n-1$ СЛВ $T(\bar{X}) = \sum_{i=1}^n X_i^2 - Y_1^2 - \dots - Y_k^2$ не завис. от Y_j и им. распределение H_{n-k}.