Сведения Теории Вероятностей и Матетатической Статистики

Сюй Минчуань

20 ноября 2020 г.

1. **Относительная частота появления события**
   Относительной частотой появления события A в N испытаниях называется число
   $$h_N(A) = \frac{N(A)}{N} \tag{1}$$

2. **Статистическая устойчивость частот**
   Для некоторого эксперимента выполнено свойство статистической устойчивости частот, если выполнены:

   (a) Относительная частота $h_N(A)$ события A в длинной серии испытаний "тяготеет" к некоторому постоянному неслучайному числу.

   (b) В разных сериях испытаний, но проводимых в одинаковых условиях, относительные частоты приблизительно равны.

   (c) Если мы из данной серии испытаний выберем некоторую подсерию, не используя информацию о результатах эксперимента, то новая относительная частота тяготеет к тому же числу.

3. **Определение дискретного вероятностного пространства**
   Дискретным вероятностным пространством называется пара (Ω, P), где Ω - конечное или счетное множество, P - вещественная функция, заданная на Ω такая, что

   (a) $P(\omega) > 0, \quad \forall\omega \in \Omega.$

   (b) $\sum_{\omega\in\Omega} P(\omega) = 1.$

   Ω - Пространство элементарных исходов
   ω - Элементарные исходы
   $P(\omega)$ - Вероятность появления ω

4. **Задача на классическое определение вероятности**
   Говорят, что мы имеем задачу на классическое определение вероятности, если $\Omega = \{\omega_1, \ldots, \omega_n\}$ - конечное множество и для всех $\omega_i, P(\omega_i) = 1/n$, т.е. все исходы равновозможны.

5. **Определение события в дискретном вероятностном пространстве**
   Случайным событием назовем произвольное подмножество A пространства элементарных исходов Ω.

6. **Благоприятный элементарный исход**
   Те ω, которые приводят к появления события A.

7. **Что значит, что событие произошло (формально)?**
   Событие произошло, если появился благоприятный ему элементарный исход.

8. **Операции над событиями (формально и не формально)**

   (a) **Достоверное и невозможное события**
   Достоверное - происходит всегда, невозможный - никогда не происходит. Формально: $\Omega, \emptyset$.

   (b) **Объединение**
   Происходит хотя бы одно из этих двух событий. Формально: $C = A \cup B$.

   (c) **Пересечение**
   Происходит оба эти события одновременно. Формально: $C = A \cap B$ или $C = AB$.

   (d) **Несовместные события**
   Не могут происходить одновременно. Формально: $AB = \emptyset$.

   (e) **Сумма**
   Объеденение в случае они совместны. Формально: $A + B$.

   (f) **Противоположное**
   Не происходит событие A. Формально: $\bar{A}$.

   (g) **Разность**
   Происходит A и не происходит B. Формально: $C = A \setminus B$.

   (h) **Одно событие влечет другое**
   При появлении событии A обязательно происходит и событие B. Формально: $A \subset B$.

9. **Алгебра событий**
   Некоторый класс $\mathcal{A}$ событий называется алгеброй событий, если

   (a) $\Omega \in \mathcal{A}$,

   (b) если $A \in \mathcal{A}$, то $\bar{A} \in \mathcal{A}$,

   (c) если $A, B \in \mathcal{A}$, то $A \cup B \in \mathcal{A}$.

10. **Определение вероятности события**

$$P(A) = \sum_{\omega \in A} P(\omega). \tag{2}$$

11. **Основные свойства вероятностей (1-3)**
Пусть выделены некоторая алгебра $\mathcal{A}$ событий, для которых определены вероятности по формуле (2). Тогда справедливы:
   (a) $P(A) \geq 0$.
   (b) $P(\Omega) = 1$.
   (c) Если $A_1, A_2, \ldots, A_n$- попарно несовместны, то
$$P(\sum_{k=1}^{n} A_k) = \sum_{k=1}^{n} P(A_k). \tag{3}$$

12. **Геометрическое определение вероятности**
Говорят, что мы имеем задачу на геометрическое определение вероятности, если Ω есть ограниченное борелевское подмножество в $\mathbb{R}^d$, $\mathcal{A}$ - алгебра всех его борелевских подмножеств, а вероятность событий задается по правилу.
$$P(A) = \frac{L(A)}{L(\Omega)}, \tag{4}$$
где $L(A)$ - мера Лебега множества A.

13. **Условная вероятность**
Условная вероятность события A при условии, что произошло событие $B(P(B) \neq 0)$, называется число
$$P(A|B) = \frac{P(AB)}{P(B)} \tag{5}$$

14. **Теорема умножения**
Пусть A и B - два события и $P(B) > 0$, тогда
$$P(AB) = P(B)P(A|B) \tag{6}$$

15. **Независимость событий**
События A и B называются независимыми, если
$$P(AB) = P(A)P(B) \tag{7}$$

16. **Полная группа событий**
События $H_1, \ldots, H_n$ образуют полную группу событий, если
   (a) $H_i H_j = \emptyset, \quad i \neq j$,
   (b) $H_1 + \ldots + H_n = \Omega$.

17. **Формула полной вероятности**
Пусть события $H_1, \ldots, H_n$ образуют полную группу событий, $P(H_k) > 0$ для всех k и A - произвольное событие. Тогда
$$P(A) = \sum_{k=1}^{n} P(H_k)P(A|H_k) \tag{8}$$

18. **Формула Байеса**
Если $H_1, \dots, H_n$ образуют полную группу событий, A - произвольное событие и $P(H_k) > 0, k = 1, \dots, n, P(A) > 0$, то

$$P(H_k|A) = \frac{P(H_k)P(A|H_k)}{P(A)} = \frac{P(H_k)P(A|H_k)}{\sum_{j=1}^{n} P(H_j)P(A|H_j)} \tag{9}$$

19. **Схема Бернулли (формально и неформально)**
Формально: Схемой Бернулли с параметрами n и p называется дискретное вероятностное пространство (Ω, P), где Ω состоит из элементарных исходов вида $\omega = (\omega_1, \dots, \omega_n), \omega_k = 0, 1, k = 1, \dots, n$, а вероятности элементарных исходов ω задаются по правилу.

$$P(\omega) = p^m(1-p)^{n-m}, \tag{10}$$

где m - число единиц в исходе ω.
Неформально: Схемой Бернулли или последовательностью n независимых одинаковых испытаний с двумя исходами называется случайный эксперимент, в котором:

   (a) проводится n независимых испытаний,
   (b) каждое испытание кончается одним из двух исходов (один исход называется успех и обозначается 1, а второй - неуспех и обозначается 0),
   (c) вероятность появления успеха одна и та же в каждом испытании и равна p.

20. **Биномиальное распределение (где возникает и формула)**
Нас интересует число успехов.
Биномиальный модель с параметрами n и p называется вероятностное пространство (Ω, P), где $\Omega = \{\omega_0, \dots, \omega_n\}$, и

$$P(\omega_m) = b(n, p, m) = C_n^m p^m (1-p)^{n-m}, m = 0, \dots, n, 0 < p < 1. \tag{11}$$

21. **Интегральная теорема Муавра-Лапласа**
Пусть мы имеем схему Бернулли с параметрами n и p. Если $n \to \infty$ p - фиксировано, то равномерно по всем $m_1 < m_2$

$$P(m_1 \leq S_n < m_2) \equiv \Phi(x_{n,m_2}) - \Phi(x_{n,m_1}), \tag{12}$$

где S_n - число успехов в n испытаниях, а

$$\Phi(x) = \int_{-\infty}^{x} \varphi(y)dy \tag{13}$$

Более того, для любых $m_1 < m_2$ имеем место оценка

$$|P(m_1 \leq S_n < m_2) - \Phi(x_{n,m_2}) + \Phi(x_{n,m_1]})| \leq \frac{p^2 + (1-p)^2}{\sqrt{np(1-p)}} \tag{14}$$

22. **Теорема Пуассона**
Пусть мы имеем схему Бернулли с параметрами n и p. Пусть $n \to \infty, p \to 0$, так, что $np \to \lambda, 0 < \lambda < \infty$. Тогда для любого фиксированного m

$$b(n, p, m) \to \pi_m(\lambda) = \frac{\lambda^m}{m!} e^{-\lambda}. \tag{15}$$

Более того, имеем место оценка

$$\sum_{m=0}^{\infty} |b(n, p, m) - \pi_m(\lambda)| \leq np^2 \tag{16}$$

23. **Определение вероятностного пространства в общем случае**
Вероятностным пространством называется тройка $(\Omega, \mathcal{A}, P)$, где Ω - произвольное множество, $\mathcal{A}$ - некоторая σ-алгебра его подмножеств, P - вещественная функция на $\mathcal{A}$:

    (a) $P(A) \geq 0, \forall A \in \mathcal{A}$,
    (b) $P(\Omega)$,
    (c) если $A_1, A_2, \dots$ - попарно несовместны, то

$$P(\sum_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k) \tag{17}$$

24. **Порожденная σ-алгебра**
Пусть $\mathcal{M}$ - некоторая система подмножеств пространства Ω. Класс $\mathcal{A} = \sigma(\mathcal{M})$ подмножеств пространствв Ω называется σ-алгебра, порожденной системой $\mathcal{M}$, если выполнены следующие свойства:

    (a) $\mathcal{M} \in \mathcal{A}$,
    (b) σ-алгебра
    (c) если $\mathcal{A}_1$ - некоторая σ-алгебра, содержащая $\mathcal{M}$, то $\mathcal{A} \subset \mathcal{A}_1$

25. **Борелевская σ-алгебра**
Пусть $\Omega = \mathbb{R}^1, \mathcal{M}$- класс всех интервалов. $\mathcal{A} = \sigma(\mathcal{M})$ называется борелевской σ-алгебра, а элементы A из $\mathcal{A}$ называются борелевскими множествами.

26. **Определение случайной величины**
Пусть $(\Omega, \mathcal{A}, P)$ - вероятностное пространство, а $(\mathbb{R}^1, \mathcal{B})$ - вещественная прямая с выделенной на ней борелевской σ-алгеброй подмножеств. Случайной величиной называется функция $\xi : \Omega \to \mathbb{R}^1$, которая обладает следующим свойством: $\forall B \in \mathcal{B}$

$$\xi^{-1}(B) = \{\omega \in \Omega : \xi(\omega) \in B\} \in \mathcal{A} \tag{18}$$

27. **Распределение случайной величины**
Распределение случайной величины ξ называется функция P_ξ, заданная на борелевской σ-алгебре $\mathcal{B}$ по правилу: $\forall B \in \mathcal{B}$
$$P_\xi(B) = P(\xi \in B) \tag{19}$$

28. **Определение функции распределения и ее основые свойства (1-5)**
Функции распределения $F_\xi(x)$ случайной величины ξ определяется по правилу: $\forall x \in \mathbb{R}^1$
$$F_\xi(x) = P(\xi < x) \tag{20}$$
Если $F_\xi(x)$ - функция распределения случайной величины ξ, то
   (a) $\forall x \in \mathbb{R}^1, 0 \le F_\xi(x) \le 1$
   (b) если $x_1 \le x_2$, то $F_\xi(x_1) \le F_\xi(x_2)$
   (c) $F_\xi(x)$ - непрерывна слева
   (d) $\lim_{x\to-\infty} F_\xi(x) = 0, \lim_{x\to\infty} F_\xi(x) = 1$
   (e) $P(a \le x < b) = F_\xi(b) - F_\xi(a)$

29. **Дикретное распределение и его свойства (1-3)**
Случайная величина ξ имеет дискретное распределение, если существует такое конечное или счетное множество $X = \{x_1, x_2, \ldots\}$, что $P(\xi \in X) = 1$. Числа $x_1, x_2, \ldots$ называются значениями случайной величины ξ, а - p_k вероятностями этих значений.
Пусть случайная величина ξ имеет дискретное распределение с множеством значений $X = \{x_1, x_2, \ldots\}$ и вероятностями этих значений $\{p_k\}$. Тогда
   (a) $p_k \ge 0$
   (b) $\sum_k p_k = 1$
   (c) $\forall B \in \mathcal{B}, P_\xi(B) = \sum_{x_k \in B} p_k$
   (d) $\forall x \in \mathbb{R}^1, F_\xi(x) = \sum_{x_k < x} p_k$
   (e) $\forall x \in X, p_n = P(\xi = x_n) = F_\xi(x_n + 0) - F_\xi(x_n)$
   (f) $P(a < x < b) = \sum_{a < x_k < b} p_k$

30. **Непрерывное распределение и его свойства (1-7)**
Распределение с.в. ρ называется абсолютно непрерывным, если существует такая вещественная функция $\rho_\xi(x)$ что $\forall B \in \mathcal{B}$
$$P_\xi(B) = P(\xi \in B) = \int_B \rho_\xi(x)dx \tag{21}$$
Функция $\rho_\xi(x)$ называется плотностью распределения с.в. ξ.
Справедливы следующие свойства

(a) $\forall x \in \mathbb{R}^1, \rho_\xi(x) \geq 0$
(b) $\int_{-\infty}^{\infty} \rho_\xi(x)dx = 1$
(c) $\forall B \in \mathcal{B}, P(\xi \in B) = \int_B \rho_\xi(x)dx$
(d) $\forall x \in \mathbb{R}^1, F_\xi(x) = \int_{-\infty}^{x} \rho(x)dx$
(e) $P(a \leq \xi < b) = F_\xi(b) - F_\xi(a) = \int_a^b \rho_\xi(x)dx$
(f) $\forall x \in \mathbb{R}^1$ где $\rho_\xi(x)$ непрерывна, $\rho_\xi(x) = \frac{d}{dx}F_\xi(x)$
(g) $\forall x \in \mathbb{R}^1, P(\xi = x) = 0$

31. **Примеры стандартных распределений**

(a) **Бернулли**
$P(\xi = 1) = p, \quad P(\xi = 0) = 1 - p$

(b) **Биномиальное**
$P(\xi = m) = C_n^m p^m (1-p)^{n-m}, \quad m = 0, 1, \ldots$

(c) **Геометрическое**
$P(\xi = m) = p(1-p)^{m-1}, \quad m = 1, 2, \ldots$

(d) **Пуассоновское**
$P(\xi = m) = \frac{\lambda^m}{m!}e^{-\lambda}, \quad m = 0, 1, \ldots$

(e) **Равномерное**

$$\rho_\xi(x) = \begin{cases} \frac{1}{b-a}, & x \in (a, b), \\ 0, & x \notin (a, b). \end{cases} \tag{22}$$

(f) **Показательное**

$$\rho_\xi(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{23}$$

(g) **Нормальное**
$\rho_\xi(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad x \in R^1$

32. **Определение случайного вектора**
n-мерным слу. вектором ξ называется набор $\xi = (\xi_1, \ldots, \xi_n)$ случайных величин, заданных на одном и том же вероятностном пространстве $(\Omega, \mathcal{A}, P)$.

33. **Распределение случайного вектора**
Распределением сл. вектора ξ называется функция P_ξ, заданная на σ-алгебра $\mathcal{B}_n$ по правилу

$$P_\xi(B) = P(\xi \in B) \tag{24}$$

34. **Дискретный случайный вектор, таблица его распределения**
    Случайный вектор $\xi = (\xi_1, \ldots, \xi_n)$ имеет дискретное распределение, если существует конечное или счетное множество $X \subset \mathbb{R}^n$, такое что $P(\xi \in X) = 1$.
    Для $n = 2$ распределение дискретного слу. вектора обычно задают в виде таблицы, называемой таблицей распределения.

35. **Случайный вектор с непрерывным распределением. Плотность распределения случайного вектора и ее свойства** Случайный вектор $\xi = (\xi_1, \ldots, \xi_n)$ имеет абсолютно непрерывное распределение, если существует вещественная функция $\rho_\xi(x), x \in R^n$, такая что $\forall B \in \mathcal{B}_n$
    $$P_\xi(B) = P(\xi \in B) = \int_B \rho_\xi(x)dx \tag{25}$$
    Функция $\rho_\xi(x)$ называется плотностью распределения с.в. ξ.
    Справедливы следующие свойства
    (a) $\forall x \in \mathbb{R}^n, \rho_\xi(x) \geq 0$
    (b) $\int_{\mathbb{R}^n} \rho_\xi(x)dx = 1$
    (c) $\forall B \in \mathcal{B}_n, P(\xi \in B) = \int_B \rho_\xi(x)dx$
    (d) $\forall x = (x_1, \ldots, x_n) \in \mathbb{R}^n$
    $$F_\xi(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_n} \rho_\xi(y_1, \ldots, y_n)dy_n \ldots dy_1 \tag{26}$$
    (e) Если $(x_1, \ldots, x_n)$ - точка нерерывности плотности $\rho_\xi(x)$, то
    $$\rho_\xi(x_1, \ldots, x_n) = \frac{\partial^n F_\xi(x_1, \ldots, x_n)}{(\partial x_1, \ldots, \partial x_n)} \tag{27}$$
    (f) Плотность с.веткtора $\tilde{\xi} = (\xi_1, \ldots, \xi_{k-1}, \xi_{k+1}, \ldots, \xi_n)$ можно вычислить по формуле
    $$\rho_\xi(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) = \int_{-\infty}^{+\infty} \rho_\xi(x_1, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_n)dx_k \tag{28}$$

36. **Маргинальные (одномерные) распределения и их вычисление**
    Распределение отдельно взятой координаты ξ_i вектора ξ называется одномерным или маргинальным.

37. **Независимость случайных величин (в общем случае и для непрерывных и дискретных с.в.)**
    Случайные величины $\xi_1, \ldots, \xi_n$ называются независимыми, если для любых борелевских $B_1, \ldots, B_n \in \mathcal{B}_1$
    $$P(\xi_1 \in B_1, \ldots, \xi_n \in B_n) = P(\xi_1 \in B_1) \ldots P(\xi_n \in B_n) \tag{29}$$

38. **Формула свертки в дискретном и непрерывном случаях**
Дискретный случай:

$$P(\xi_1 + \xi_2 = z) = \sum_x P_1(x)P_2(z-x) \tag{30}$$

Непрерывный случай:

$$\rho_{\xi_1+\xi_2}(y) = \int_{-\infty}^{\infty} \rho_{\xi_1}(y-x)\rho_{\xi_2}(x)dx \tag{31}$$

39. **Математичекое ожидание дискретной случайной величины**
$M\xi = \sum_n x_n p_n$

40. **Математичекое ожидание непрерывной случайной величины**
Пусть ξ - неотрицательная с.в., $\{\xi_n\}$ - последовательность дискретных с.в., обладающих свойствами:

    (a) $\xi_n(\omega) \geq 0, \quad \forall n$
    (b) $\xi_n(\omega) \leq \xi_{n+1}(\omega)$
    (c) $\xi_n(\omega) \to \xi_(\omega)$ равномерно по $\omega \in \Omega$ при $n \to \infty$

    Математическим ожиданием с.в. ξ называется число

$$M\xi = \lim_{\xi} M\xi_n \tag{32}$$

Математическим ожиданием произвольной случайной величины ξ называется число

$$M\xi = M\xi^+ - M\xi^- \tag{33}$$

если хотя бы одно из чисел в правой части этого равенства конечно.

41. **Основные свойства математического ожидания (1-5)**

    (a) Если $P(\xi = C) = 1$, то $M\xi = C$
    (b) $M(C\xi) = C \cdot M\xi$
    (c) $M(\xi_1 + \xi_2) = M\xi_1 + M\xi_2$
    (d) Если $\xi \geq 0$, то $M\xi \geq 0$, причем $M\xi = 0$ тогда и только тогда, когда $P(\xi = 0) = 1$
    (e) Если $\xi_1 \geq \xi_2$, то $M\xi_1 \geq M\xi_2$

42. **Дисперсия и ее основные свойства**
Дисперсией с.в. ξ называется число

$$D(\xi) = M[(\xi - M\xi)^2] \tag{34}$$

    (a) $D(\xi) \geq 0, D(\xi) = 0$ тогда и только тогда, когда $\xi = C$.

(b) $D(\xi + C) = D(\xi)$

(c) $D(C\xi) = C^2 D(\xi)$

(d) $D(\xi) = M(\xi^2) - (M\xi)^2$

43. **Ковариация. Некоррелированные случайные величины**
Ковариацией с.в. ξ_1 и ξ_2 называется число

$$\operatorname{cov}(\xi_1, \xi_2) = M[(\xi_1 - M\xi_1)(\xi_2 - M\xi_2)] \tag{35}$$

С.в. ξ_1 и ξ_2 называются некоррелированными, если

$$\operatorname{cov}(\xi_1, \xi_2) = M[(\xi_1 - M\xi_1)(\xi_2 - M\xi_2)] = 0 \tag{36}$$

44. **Определение коэффициента корреляции**
Коэффициентом корреляции с.в. ξ_1 и ξ_2 называется число

$$\rho(\xi_1, \xi_2) = \frac{\operatorname{cov}(\xi_1, \xi_2)}{\sqrt{D(\xi_1)D(\xi_2)}} \tag{37}$$

45. **Определение момента случайной величины**
Моментом порядка k относительно точки a с.в. ξ называется число

$$M(\xi - a)^k \tag{38}$$

Если $a = 0$ - начальный момент
Если $a = M\xi$ - центральный момент
$\beta_k = M(|\xi|^k)$ - абсолютные моменты

46. **Квантиль порядка p для с.в.. Медиана**
Квантиль порядка p, $0 < p < 1$, с.в. ξ (или ее распределения) называется число $x_p \in \mathbb{R}^1$:

$$P(\xi \le x_p) \ge p, \quad P(\xi \ge x_p) \ge 1 - p \tag{39}$$

Число $x_{1/2}$ - медиана. Оно определяет центр распределения.

47. **Сходимость в среднем квадратическом**
Будем говорить, что последовательность с.в. $\{\xi_n\}$ сходится в среднем квадратическомк к с.в., если

$$\|\xi_n - \xi\|^2 = M(|\xi_n - \xi|^2) \to 0, \quad n \to \infty \tag{40}$$

48. **Постановка задачи о наилучшей линейной оценке**
Пусть $\mathcal{L} \subset L_2$ - некоторое линейное подпространство, которое замкнуто относительно сходимости в среднем квадратическом, а η - произвольный элемент из L_2.
С.в. $\hat{\eta} \in L_2$ называется наилучшим приближением с.в. η в пространстве $\mathcal{L}$, если

(a) $\hat{\eta} \in \mathcal{L}$

(b) $\|\eta - \hat{\eta}\|^2 = M|\eta - \hat{\eta}|^2 \le M|\eta - \xi|^2 = \|\eta - \xi\|^2, \quad \forall \xi \in \mathcal{L}.$

49. **Лемма о перпендикуляре**
С.в. $\hat{\eta}$ является наилучшим риюлижением с.в. η в линейном пространстве $\mathcal{L}$ тогда и только тогда, когда

(a) $\hat{\eta} \in \mathcal{L}$

(b) $(\eta - \hat{\eta}, \xi) = M[(\eta - \hat{\eta})\xi] = 0, \quad \forall \xi \in \mathcal{L}.$

50. **Условное распределение и условное математическое ожидание в дискретном случае**
Условное распределение

$$P(x_{m+1}, \dots, x_{m+n}|x_1, \dots, x_m) = \frac{P(\xi_1 = x_1, \dots, \xi_m = x_m, \xi_{m+1} = x_{m+1}, \dots, \xi_{m+n} = x_{m+n}}{P(\xi_1 = x_1, \dots, \xi_m = x_m)} \tag{41}$$

Условное математическое ожидание

$$M(\xi_{m+1}|\xi_1 = x_1, \dots, \xi_m = x_m) = \sum_y P(\xi_{m+1} = y|\xi_1 = x_1, \dots, \xi_m = x_m) \tag{42}$$

51. **Условное распределение и условное математическое ожидание в непрервном случае**
Условное распределение

$$\rho_{\xi''|\xi'}(x_{m+1}, \dots, x_{m+n}|x_1, \dots, x_m) = \frac{\rho_\xi(x_1, \dots, x_m, x_{m+1}, \dots, x_{m+n})}{\rho_{\xi'}(x_1, \dots, x_m)} \tag{43}$$

Условное математическое ожидание

$$M[\xi_{m+1}|\xi' = x'] = \int_{-\infty}^{\infty} y\rho_{\xi_{m+1}|\xi'}(y|x')dy = g(x') \tag{44}$$

52. **Функция регрессии и ее экстремальное свойство**
$g = (\xi_1, \dots, \xi_m) = M(\xi_{m+1}|\xi_1, \dots, \xi_m)$ - Функция регрессии с.в. ξ_{m+1} на с.в. $\xi_1, \dots, \xi_m$.
$y = g(x') = M[\xi_{m+1}|\xi_1, \dots, \xi_m]$- Функция регрессии с.в. ξ_{m+1} на с.вектор ξ'.
Экстремальное свойство:
Пусть ξ и η - две с.в., $y = f(x)$ - некоторая борелевская функция, причем $M(\eta^2) < \infty$ и $M([f(\xi)]^2) < \infty$. Если $y = g(x) = M[\eta|\xi = x]$ есть функция регрессии с.в. η на с.в. ξ, то

$$M|\eta - g(\xi)|^2 \le M|\eta - f(\xi)|^2 \tag{45}$$

53. **Сходимость по вероятности**
Последовательность с.в. $\{\xi_n\}$ сходится по вероятности к с.в. ξ, если $\forall \varepsilon > 0$

$$P(|\xi_n - \xi| > \varepsilon) \to 0 \tag{46}$$

или

$$P(|\xi_n - \xi| \leq \varepsilon) \to 1 \tag{47}$$

при $n \to \infty$. Обозначение: $\xi_n \xrightarrow{P} \xi, n \to \infty$

54. **Что такое закон больших чисел?**
Говорят, что к последовательности с.в. $\{\xi_n\}$ применим закон больших чисел (ЗБЧ), если

$$\frac{\xi_1 + \ldots + \xi_n}{n} - \frac{M\xi_1 + \ldots + M\xi_n}{n} \xrightarrow{P} 0, \quad n \to \infty \tag{48}$$

55. **Неравенство Чебышева**

$$P(|\xi - M\xi| > \varepsilon) \leq \frac{D(\varepsilon)}{\varepsilon^2} \tag{49}$$

56. **З.Б.Ч. для н.о.р.с.в**
Пусть $\{\xi_n\}$ - последовательность с.в.:

    (a) $\{\xi_n\}$ - независимы
    (b) $\{\xi_n\}$ - одинаково распределены
    (c) $\exists M\xi_n = a, D(\xi_n) = \sigma^2 < \infty$

    Тогда применим ЗБЧ, т.е.

$$\frac{\xi_1 + \ldots + \xi_n}{n} \xrightarrow{P} a, \quad n \to \infty \tag{50}$$

57. **Определение характеристической функции и ее вычисление в дискретном и непрерывном случае**
Характеристической функции с.в. ξ называется комплекснозначная функция $\varphi_\xi(t), t \in \mathbb{R}^1$, определяемая по правилу

$$\varphi_\xi(t) = M(e^{it\xi}) = \int_{-\infty}^{\infty} e^{itx} dF(x) \tag{51}$$

Для дискретной с.в. имеем формулу

$$\varphi_\xi(t) = \sum_n e^{itx_n} \cdot p_n \tag{52}$$

Для непрерывной с.в. имеем формулу

$$\varphi_\xi(t) = \int_{-\infty}^{\infty} e^{itx_n} \rho_\xi(x) dx \tag{53}$$

58. **Вычисление характеристической функции для суммы независимых с.в.**
$\varphi_{\xi_1+\xi_2}(t) = \varphi_{\xi_1}(t) \cdot \varphi_{\xi_2}(t)$

59. **Теорема единственности**
Соответствие между функциями распределения F и характеристическими функциями φ является взаимно однозначным. Более того, если x_1 и x_2 - точки непрерывности ф.р. F, то

$$F(x_2) - F(x_1) = \lim_{A\to\infty} \frac{1}{2\pi} \int_{-A}^{A} \frac{e^{-itx_2} - e^{-itx_1}}{-it} \varphi(t) dt \tag{54}$$

В частности, если существует плотность $\rho(x)$, то

$$\rho(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi(t) dt \tag{55}$$

60. **Теорема непрерывности**
Пусть $F_0, F_1, \ldots$ - последовательность ф.р., а $\varphi_0, \varphi_1, \varphi_2, \ldots$ - соответствующая ей последовательность х.ф. Тогда следующие утверждения эквивалентны:

    (a) F_n слабо сходится к F_0
    (b) $\varphi_n(t) \to \varphi_0(t)$ при $n \to \infty$ для любого $t \in R^1$

61. **Характеристическая функция для стандартного нормального распределения**
$\varphi_\xi(t) = e^{-\frac{1}{2}t^2}$

62. **Слабая сходимость функций распределений**
Последовательность функций распределения $\{F_n\}$ слабо сходится к функции распределения F_0, если

$$F_n(x) \to F_0(x) \tag{56}$$

при $n \to \infty$ для всех точек, где предельная ф.р. F_0 непрерывна. Для соответствующих с.в. ξ_n будем говорить, что они сходятся к с.в. ξ_0 по распределению. Обозначение: $\xi_n \xrightarrow{d} \xi_0$

63. **Что такое Центральная предельная теорема?**
Говорят, что к последовательности с.в. $\{\xi_n\}$ применима центральная предельная теорема (ЦПТ), если для любого $n \geq N_0$ существуют центрирующие и нормирующие константы $A_n \in \mathbb{R}^1$ и $B_n > 0$ такие, что для

$$S_n^* = \frac{\xi_1 + \ldots + \xi_n - A_n}{B_n} \tag{57}$$

имеем место сходимость

$$P(S_n^* < x) = F_{S_n^*}(x) \to \Phi(x) \tag{58}$$

при $n \to \infty$ для любого $x \in \mathbb{R}^1$.

64. **Ц.П.Т. для н.о.р.с.в.**
Пусть $\{\xi_n\}$ - последовательность с.в.:

(a) $\{\xi_n\}$ - независимы

(b) $\{\xi_n\}$ - одинаково распределены

(c) $\exists M\xi_n = a, D(\xi_n) = \sigma^2$

Тогда к этой последовательности применима ЦПТ, т.е., если

$$S_n^* = \frac{\xi_1 + \ldots + \xi_n - na}{\sigma\sqrt{n}}, \tag{59}$$

то

$$F_{S_n^*}(x) \to \Phi(x), \quad n \to \infty \tag{60}$$

65. **Основная задача математической статистики**
На основе экспериментальных данных сузить класс $\mathcal{P}$ априорно заданных вероятностных мер до некоторого более узкого подкласса $\mathcal{P}_0 \subset \mathcal{P}$(в идеале выбрать одно распределение)

66. **Определение статистической структуры**
Статистической структурой называется тройка $(\Omega, \mathcal{A}, \mathcal{P})$, где

(a) Ω - произвольное множество = пространство элементарных исходов

(b) $\mathcal{A}$ - σ-алгебра подмножеств Ω = события, доступных наблюдению.

(c) $\mathcal{P}$ - некоторый набор вероятностных мер на $(\Omega, \mathcal{A})$.

67. **Параметрические и непараметрические структуры**
Если существует конечное число числовых параметров $(\theta_1, \ldots, \theta_m) = \theta, \theta \in \Theta \subset \mathbb{R}^m$, с помощью которых удается занумеровать все распределения P из класса $\mathcal{P}$, то статистическая структура называется параметрической. В противном случае мы имеем непараметрическую структуру.

68. **Примеры параметрических структур**

(a) **Биноминальная**
$P(\xi = m) = C_n^m p^m (1-p)^{n-m}$. $n \geq 0$ и $p \in (0,1)$ - параметры.

(b) **Пуассоновская**
$P(\xi = m) = \frac{\lambda^m}{m!} e^{-\lambda}$. $\lambda > 0$ - параметр.

(c) **Показательное**

$$\rho_\xi(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{61}$$

$\lambda > 0$ - параметр.

(d) **Нормальное**

$\rho_\xi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad x \in \mathbb{R}^1$. $a \in \mathbb{R}^1$ и $b > 0$ - параметры.

69. **Выборка, повторная выборка**
Набор чисел $x = (x_1, \ldots, x_n)$, получаемый из экспреимента - выборка. С.в. $X_1, \ldots, X_N$ одинаковы распределены (однородная выборка) и независимы, тогда X - повторная выборка.

70. **Вариационный ряд, порядковые статистики**
Вариационный ряд: $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(N)}$.
$x_{(k)}$ - k-ой порядковая статистика

71. **Эмпирическое распределение, эмпирическая функция распределения, гистограмма**
Эмпирическое распределение:
Имеем выборку $x = (x_1, \ldots, x_N)$. Рассмотрим новую с.в. ξ^*, множеством значений которой являются числа $x_1, \ldots, x_N$, каждому из которых приписывается вероятность $1/N$ (если некоторое значение появляется несколько раз, то его вероятность увеличивается в то же число раз). С.в. ξ^* является дискретной и ее распределение P_N^* называется эмпирическим распределением, построенным по выборке x.

$$P_N^*(B) = \frac{v_N(B)}{N} \tag{62}$$

где $v_N(B)$ - число элементов выборки, которые попали во множество B. Эмпирическая функция распределения:

$$F_N^*(y) = \frac{N(y)}{N} \tag{63}$$

где $N(y)$ - число элементов x_k выборки x, для которых $x_k < y$.
Гистограмма:
Пусть мы сгруппировали все элементы x_i выборки x в r интервалов, длина -го интервала равна Δ_k, а N_k есть число элементов выборки, попавших в k-ый интервал. Гистограмма есть функция $\rho_N^*(y)$, определяемая по правилу

$$\rho_N^*(y) = \frac{N_k}{N \cdot \Delta_k} \tag{64}$$

если принадлежит k-му интервалу, и равная нулю в противном случае.

72. **Выборочное среднее, выборочная дисперсия и выборочные моменты**

$$
\begin{aligned}
\bar{x} &= \frac{1}{N}\sum_{k=1}^{N} x_k, \\
S^2 &= \frac{1}{N}\sum_{k=1}^{N} x_k^2 - \bar{x}^2 = \frac{1}{N}\sum_{k=1}^{N}(x_k - \bar{x})^2, \\
\hat{v}_m &= \frac{1}{N}\sum_{k=1}^{N} x_k^m.
\end{aligned}
\tag{65}
$$

73. **Определение точечной оценки**
Оценкой параметра θ называется произвольная функция $\hat{\theta}$ из выборочного пространства X в пространство параметров $\Theta \subset \mathbb{R}^m$.

74. **Свойства оценок**

   (a) **Несмещенность**
   Оценка $\hat{\theta} = \hat{\theta}_N(X_1, \ldots, X_N)$ параметра θ называется несмещенной, если
   $$M_\theta(\hat{\theta}_N(X)) = \theta, \quad \theta \in \Theta \tag{66}$$

   (b) **Состоятельность**
   Последовательность оценок $\{\hat{\theta}_N(X), N \geq N_0\}$ параметра θ называется состоятельной, если
   $$\hat{\theta}_N(X) \xrightarrow{P_\theta} \theta, \quad \theta \in \Theta \tag{67}$$

   (c) **Оптимальность в среднем квадратическом**
   Эффективная оценка является оптимальной в среднем квадратическом оценкой в классе несмещенных оценок. Оптимальная оценка необязательно является эффективной.

   (d) **Асимптотическая нормальность**
   Посоедовательность оценок $\{\hat{\theta}_N, N \geq N_0\}$ называется асимптотически нормальной, если $\forall N \geq N_0$ существуют константы $A_N(\theta) \in \mathbb{R}^1$ и $B_N(\theta) > 0$ такие что, с.в.
   $$\frac{\hat{\theta}_N - A_N(\theta)}{B_N(\theta)} \tag{68}$$
   имеет асимптотически стандартное нормальное распределение, т.е. ее функция распределения сходится к функции распределения стандартного нормального закона.

75. **Неравенство Рао-Крамера. Эффективная оценка**
$\hat{\theta}_N = \hat{\theta}(X)$ является эффективной, если она несмещенная и для любой другой несмещенной оценки $\tilde{\theta}_N = \tilde{\theta}_N(X)$ мы имеем
$$D_\theta(\hat{\theta}_N) \leq D_\theta(\tilde{\theta}_N) \tag{69}$$

Неравенство Рао-Крамера:
Пусть $\hat{g}_N = \hat{g}_N(x)$ - некоторая несмещенная оценка для вещественной функции $g(\theta)$ от параметра θ, построенный по повторной выборке $X = (X_1, \ldots, X_N)$ из генеральной совокупности с функцией распределения $F(y, \theta)$, удовлетворяющей условиям регулярности. Тогда имеет место неравенство

$$D_\theta(\hat{g}_N) \geq \frac{[g'(\theta)]^2}{N \cdot I(\theta)}. \tag{70}$$

В частности, если $g(\theta) \equiv \theta$, то

$$D_\theta \geq \frac{1}{I_N(\theta)}, \quad \forall \theta \tag{71}$$

76. **Определение доверительного интервала. Точность и надежность интервальной оценки**
Пусть чы имеем повторную выборку $X = (X_1, \ldots, X_N)$ из генеральной совокупности с функцией распределения $F(y, \theta)$, где $\theta \in \Theta \subset \mathbb{R}^1$ - скалярный параметр. Доверительный интервал уровня γ для параметра θ называется интервал $(\hat{\theta}^{(1)}(X), \hat{\theta}^{(2)}(X))$ со случайными концами такой, что

$$P_\theta(\hat{\theta}^{(1)}(X) < \theta < \hat{\theta}^{(2)}(X)), \quad \theta \in \Theta \tag{72}$$

Число γ называется доверительным уровнем интервала.
Число γ характеризует надежность этого интервала.
Число $l = M_\theta(\hat{\theta}^{(2)} - \hat{\theta}^{(1)})$ характеризует точность интервала.

77. **Гипотеза о согласии (однородности и независимости)**
Пусть $\mathcal{F}$ - некоторое семейство функций распределения. ξ - случайная величина с неизвестной функцией распределения $F_\xi(y)$. Гипотеза о виде распределения называется предположение:

$$H: \quad F_\xi \in \mathcal{F} \tag{73}$$

Пусть имеем повторные выборки $X = (X_1, \ldots, X_{N_1}), Y = (Y_1, \ldots, Y_{N_2})$.
Пусть F_1 и F_2 есть функции распределения, отвечающие выборкам X и Y соответственно.
Гипотеза однородности имеет вид:

$$H: \quad F_1(y) \equiv F_2(y) \tag{74}$$

где F_1 и F_2 неизвестны.
Пусть $\xi = (\xi_1, \xi_2)$ - двумерный слу. вектор с функцией распределения $F(z_1, z_2), F_1(z_1)$ - функция распределения для $\xi_1, F_2(z_2)$ - функция распределения для ξ_2.
Гипотеза о независимости имеет вид:

$$H: \quad F(z_1, z_2) = F_1(z_1) \cdot F_2(z_2) \tag{75}$$

78. **Основные понятия теории проверки статистических гипотез**

(a) **Статистическая гипотеза**
Статистической гипотезой называется подкласс $\mathcal{P}_0 \in \mathcal{P}$. Условно это записывается в виде

$$H: \quad P \in \mathcal{P}_0 \tag{76}$$

Статистические гипотезы будут обозначаться $H, H_0, H_1, \ldots$.

(b) **Простая и сложная гипотезы**
Если $\mathcal{P}_0$ содержит только один элемент, то гипотеза H называется простой. В противном случае называется сложной.

(c) **Основная (нулевая) гипотеза и альтернатива**
Обычно формулируют несколько гипотез. Одну из них выделяют в качестве основной и называют ее нулевой, обозначаемой H_0. Остальные гипотезы называют альтернативами и обозначают $H_1, H_2, \ldots$.

(d) **Критерий, критическая зона, статистика критерия, критическая константа**
Статистическим критерием (тестом, решающим правилом) для проверки гипотезы H_0 против альтернативы H_1 называется произвольное отображение $\varphi : \chi \to \{0, 1\}$. Если $\phi(x) = 0$, то принимаем гипотезу H_0. Если $\varphi(x) = 1$, то принимаем гипотезу H_1.
Обозначим
$$K = \{x \in \mathcal{X} : \quad \varphi(x) = 1\} \tag{77}$$
Множество K - критическая зона теста ϕ.
Очень часто критическая зона теста задается по правилу

$$K = \{x \in \mathcal{X} : \quad T(x) > c\} \tag{78}$$

или
$$K = \{x \in \mathcal{X} : \quad T(x) < c\} \tag{79}$$

$T(x)$ - статистика критерия φ, c - критическая константа.

(e) **Ошибки первого и второго рода**
Ошибки первого рода: Принимаем гипотезу H_1, когда верна H_0.
Ошибки второго рода: Принимаем гипотезу H_0, когда верна H_1.
Вероятность ошибки первого рода: $\alpha(\theta) = P_\theta(X \in K | H_0)$.
Вероятность ошибки второго рода: $1 - \beta(\theta) = P_\theta(X \notin K | H_1)$.

(f) **Уровень значимости и мощность критерия**
Мы выбираем некоторое малое α и рассматриваем только такие критерии φ, для которых

$$\sup_{\theta \in \Theta_0} \alpha(\theta) \leq \alpha, \text{или} \quad P_0(X \in K) = \alpha \tag{80}$$

Значение α - уровень значимости.
Функция мощности:

$$\beta(\theta) = \beta(\theta = P_\theta(X \in K)), \quad \theta \in \Theta_1 \tag{81}$$

(g) **Равномерно наиболее мощный критерий**
Критерий φ_0 уровня α называется равномерно наиболее мощным критерием уровня α, если для любого другого критерия φ уровня α имеет место

$$\beta_{\varphi_0}(\theta) \geq \beta_\varphi(\theta), \quad \forall\theta \in \Theta_1 \tag{82}$$

(h) **Лемма Неймана-Пирсона**
Пусть с.в. ξ имеет дискретное распределение и проверяется простая гипотеза H_0 против простой альтернативы H_1. Тогда среди всех тестов φ уровня α существует наиболее мощный тест φ_0, и он имеет следующий вид:

$$\varphi_0(x) = \begin{cases} 1, & P_1(x) > c \cdot P_0(x); \\ 0, & P_1(x) < c \cdot P_0(x); \\ \gamma, & P_1(x) = c \cdot P_0(x); \end{cases} \tag{83}$$

где константа $c > 0$ и $0 \leq \gamma \leq 1$ определяется из условия:

$$P_0(X \in K) = \alpha \tag{84}$$

(i) **Критерий отношения правдоподобия**
Пусть с.в. ξ имеет функцию распределения $F(y, \theta)$, где $\theta \in \Theta$ - неизвестный параметр. Пусть Θ_0 и Θ_1 - подмножества Θ, для которых $\Theta_0 \cap \Theta_1 = \emptyset, \Theta_0 \cup \Theta_1 = \Theta$. Проверяется гипотеза

$$H_0 : \quad \theta \in \Theta_0 \tag{85}$$

против альтернативы

$$H_1 : \quad \theta \in \Theta_1 \tag{86}$$

Для проверки таких гипотез предлагается использовать статистику

$$\lambda(x) = \frac{\sup_{\theta\in\Theta_1} \mathcal{L}(x, \theta)}{\sup_{\theta\in\Theta_0} \mathcal{L}(x, \theta)} \tag{87}$$

или эквивалентно

$$\lambda_1(x) = \lambda(x) = \frac{\sup_{\theta\in\Theta} \mathcal{L}(x, \theta)}{\sup_{\theta\in\Theta_0} \mathcal{L}(x, \theta)} = \max(\lambda(x), 1) \tag{88}$$

где $\mathcal{L}$ есть функция правдоподобия, т.е. распределение вероятности (плотность) повторной выборки $X = (X_1, \ldots, X_N)$. Если удается найти распределение этих статистик, то соответствующий критерий, называемый критерием отношения правдоподобия, задается с помощью критической области следующего вида:

$$K = \{x : \quad \lambda_1(x) > c\} \tag{89}$$

где критическая константа c находится из условия

$$P_\theta(\lambda_1(x) > c) \leq \alpha, \quad \theta \in \Theta \tag{90}$$