§ 1. Разные задачи

1. Пусть X_1 и X_2 — независимые с. в., $\begin{array}{c|c|c|}X_1 & 0 & 1 \\ \hline & 1/2 & 1/2\end{array}$, $\begin{array}{c|c|c|}X_2 & 1 & 4 \\ \hline & 1/3 & 2/3\end{array}$, $Z = X_1 + X_2$, $Y = X_1$. Найти $\mathsf{E}\{Y/Z\}$.

2. Пусть точка с координатами ξ и η имеет равномерное распределение в треугольнике с вершинами (1,0), (0,1), (1,1). Найти условную плотность $f(x/y)$ с. в. η относительно с. в. ξ.

3. Пусть X_1, X_2, $\ldots$ — независимые, одинаково распределенные с. в., $\mathsf{E}X_1^2 < \infty$, $\mathsf{E}X_1 = 1$, $\mathsf{D}X_1 = \sigma^2 > 0$ и $S_n = X_1 + \ldots + X_n$ при $n = 1, 2, \ldots$. Показать, что случайные величины

$\psi_n \equiv \dfrac{S_n^3}{n^{5/2}} - \sqrt{n}$ сходятся слабо к некоторой случайной величине, и найти ее распределение.

4. Пусть $X = (X_1, X_2)$ — выборка объема 2 из нормального распределения $N(1, 2)$. Обозначим $S_1 = X_1 + X_2$, $S_2 = X_1^2 + X_2^2$. Указать число K, при котором с. в. S_1 и $KS_2 - S_1^2$ независимы.

5. Пусть $X = (X_1, X_2)$ — выборка объема 2 из нормального распределения $N(2, 3)$. Указать число a, при котором с. в. $X_1 - aX_2$ и $X_1 + X_2$ независимы.

6. Пусть $\vec{X}$ — выборка из показательного распределения E_1. При каждом $n \geqslant 1$ через ν_n обозначим количество X_i, $1 \leqslant i \leqslant n$, не превышающих 2. Указать по крайней мере две различных последовательности чисел c_n таких, что последовательность $\dfrac{\nu_n - c_n}{\sqrt{n}}$ сходится слабо к некоторому нормальному распределению. Найти параметры этого распределения.

7. Вытекает ли из слабой сходимости $\xi_n \Rightarrow \xi$ сходимость $I(\xi_n \geqslant 0) \Rightarrow I(\xi \geqslant 0)$?

8. Пусть $X_1, \ldots, X_n$ — выборка из распределения F с конечной и отличной от нуля дисперсией. Описать класс всех распределений F, для которых имеет место сходимость $\mathsf{P}(\overline{X} > 3) \to \mathsf{P}(\mathsf{E}X_1 > 3)$.

9. Найти условное математическое ожидание $\mathsf{E}(2X_1 | X_{(n)})$ для выборки объема n из равномерного распределения $U(0, 3)$.

Указание. Можно использовать теорему об улучшении оценок с помощью полных и достаточных статистик.

10. Пусть дана выборка $X_1, \ldots, X_n$ из распределения $\Pi(\lambda)$, где $\lambda \in (0, \infty)$. Является ли статистика $S = n\overline{X} - 5$ достаточной? Будут ли достаточными следующие статистики: а) $2S$? б) S^2? в) S/n^2? г) $\sin S$? д) $\exp\{S\}$? е) $-S$?

§ 2. Характеристики эмпирического распределения

11. Пусть $\vec{X}$ — выборка из распределения Бернулли с параметром 1/4, $F_n^*(y)$ — эмпирическая функция распределения и $F(y)$ — функция распределения для $B_{1/4}$. Доказать теорему Гливенко — Кантелли: при $n \to \infty$

$$\sup_{-\infty < y < \infty} |F_n^*(y) - F(y)| \to 0 \quad \text{п.н.}$$

12. Пусть $X_1, \ldots, X_n$ — выборка из распределения с функцией распределения F. Найти $\mathsf{E}F_n^*(y)$ и $\mathsf{D}F_n^*(y)$.

13. Пусть $X_1, \ldots, X_n$ — выборка из равномерного распределения на отрезке $[0, \theta]$ и $X_{(n)} = \max(X_1, \ldots, X_n)$. Найти предельное распределение последовательности $\dfrac{n}{\theta}\left(\theta - X_{(n)}\right)$ при $n \to \infty$.

14. Дана выборка $X_1, \ldots, X_n$, $\mathsf{D}X_1 = \sigma^2 < \infty$. Доказать, что $S^2 = \dfrac{1}{n}\sum\limits_{i=1}^{n}(X_i - \overline{X})^2$ является состоятельной (сильно состоятельной) оценкой для σ^2. Является ли S^2 несмещенной оценкой для σ^2? Построить сильно состоятельную и несмещенную оценку.

15. Пусть $X_1, \ldots, X_n$ — выборка из распределения $B(1, p)$. Проверить, является ли статистика $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$ состоятельной оценкой для параметра $\theta = p(1-p)$. Имеет ли статистика $\dfrac{nS^2}{p(1-p)}$ распределение хи-квадрат?

16. Пусть $X_1, \ldots, X_n$ — выборка из распределения $U[0, b]$. Проверить, является ли статистика $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$ состоятельной оценкой для параметра $\theta = b^2/12$. Имеет ли статистика $\dfrac{nS^2}{\theta}$ распределение хи-квадрат?

17. Пусть $X_1, \ldots, X_n$ — выборка из равномерного распределения $U[-\sqrt{3a}, \sqrt{3a}]$. Проверить, является ли статистика $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$ состоятельной оценкой для параметра a. Имеет ли статистика $\overline{X}$ нормальное распределение?

18. Дана выборка $X_1, \ldots, X_n$ из распределения с конечным моментом порядка $2k$. Найти математическое ожидание и дисперсию k-го выборочного момента

$$\overline{X^k} = \frac{X_1^k + \ldots + X_n^k}{n}.$$

Является ли $\overline{X^k}$ состоятельной (сильно состоятельной) оценкой для $m_k = \mathsf{E}\, X_1^k$?

§ 3. Построение оценок

19. Используя функции $g(y) = y^k, k = 1, 2, \ldots$, построить оценки неизвестного параметра для следующих распределений:

а) $U_{0,\theta}$, $\theta > 0$; б) E_α, $\alpha > 0$; в) $E_{1/\alpha}$, $\alpha > 0$; г) $N_{\alpha,1}$, $\alpha \geqslant 0$.

20. Привести пример, когда нельзя построить оценку по методу моментов с помощью функции $g(y) = y$.

21. Построить оценку неизвестного параметра (λ_1, λ_2) по методу моментов при подходящих функциях $g_1(y)$, $g_2(y)$ для смеси двух распределений Пуассона:

$$\mathsf{P}(X_1 = k) = \frac{1}{2}\frac{\lambda_1^k}{k!}e^{-\lambda_1} + \frac{1}{2}\frac{\lambda_2^k}{k!}e^{-\lambda_2},\ k = 1, 2, \ldots$$

22. Привести пример, когда оценка максимального правдоподобия не единственна.

23. Привести пример, когда оценка максимального правдоподобия не совпадает с оценкой по методу моментов, полученной с помощью функции $g(y) = y$.

24. Пусть элементы выборки $X_1, \ldots, X_n$ имеют распределение с плотностью

$$f_\theta(y) = \begin{cases} \dfrac{3y^2}{8\theta^3}, & \text{если } y \in [0, 2\theta], \\ 0, & \text{если } y \notin [0, 2\theta], \end{cases} \quad \theta \in \Theta = (0, \infty).$$

а) Найти оценку максимального правдоподобия $\hat{\theta}^*$ параметра θ. Является ли оценка $\hat{\theta}^*$ несмещенной и состоятельной?

б) Найти оценку по методу моментов θ^*, используя первый момент. Является ли оценка θ^* несмещенной и состоятельной?

в) Проверить оценки $\hat{\theta}^*$ и θ^* на асимптотическую нормальность.

г) Сравнить оценки θ^* и $\hat{\theta}^*$ в среднеквадратическом смысле.

25. Решить предыдущую задачу, если элементы выборки $X_1, \ldots, X_n$ имеют распределение с плотностью

$$f_\theta(y) = \begin{cases} \dfrac{y - \theta}{2}, & \text{если } y \in [\theta, \theta + 2], \\ 0, & \text{если } y \notin [\theta, \theta + 2], \end{cases} \quad \theta \in \Theta = (-\infty, \infty).$$

26. Пусть элементы выборки $X_1, \ldots, X_n$ имеют равномерное распределение на отрезке $[-2\theta, 2\theta]$, где $\theta > 0$ — неизвестный параметр.

а) Найти оценку максимального правдоподобия $\hat{\theta}^*$ параметра θ, проверить несмещенность и состоятельность.

б) Найти оценку θ^* по методу моментов. Является ли оценка θ^* состоятельной? Несмещенной?

в) Проверить оценки $\hat{\theta}^*$ и θ^* на асимптотическую нормальность. Для асимптотически нормальных оценок найти коэффициенты асимптотической нормальности.

27. Пусть $\vec{X} = (X_1, \ldots, X_n)$ — выборка из параметрического семейства $\mathcal{F}_\theta$, $\theta \in \Theta$. Доказать, что если оценка $\theta^* = \theta^*(\vec{X})$ является R-эффективной оценкой в классе $K_{\theta/n}$, то она состоятельна. Построить эффективную оценку в классе K_0.

28. Дана выборка $X_1, \ldots, X_n$ из распределения P_q: $\mathsf{P}_q(X_i = k) = \begin{cases} q^5, & k = 1 \\ 1 - q - q^5, & k = 2, \\ q, & k = 3 \end{cases}$

где $q \in (0, 1/2)$. Является ли оценка $q^* = \sqrt[5]{\dfrac{I(X_1 = 1) + \ldots + I(X_n = 1)}{n}}$ а) состоятельной; б) несмещенной; в) асимптотически нормальной оценкой параметра q?

29. Пусть дана выборка $X_1, \ldots, X_n$ из распределения $\Pi(\lambda)$, где $\lambda > 0$. Найти оценку метода моментов (по первому моменту) для параметра $\theta = \mathsf{P}_\lambda(X_1 = 0)$. Проверить ее несмещенность, состоятельность, асимптотическую нормальность.

30. Пусть элементы выборки $X_1, \ldots, X_n$ имеют равномерное распределение на отрезке $[0, \theta]$, где $\theta > 0$ — неизвестный параметр. Является ли оценка $\hat{\theta}^* = \dfrac{n+1}{n} X_{(n)}$ асимптотически нормальной?

31. Пусть θ^* — асимптотически нормальная оценка параметра θ с коэффициентом $\sigma^2(\theta)$. Доказать, что $\dfrac{1}{(\theta^*)^2}$ — асимптотически нормальная оценка для $\dfrac{1}{\theta^2}$.

32. Пусть дана выборка $X_1, \ldots, X_n$ из распределения $B(1, \sqrt{p})$, где $0 < p < 1$. Является ли оценка метода моментов для параметра p, полученная по первому моменту, а) несмещенной? б) состоятельной? в) асимптотически нормальной?

33. Всегда ли несмещенная оценка является: а) состоятельной? б) асимптотически нормальной? в) R-эффективной? Если «да», то почему, если «нет» — привести пример(ы).

34. Всегда ли состоятельная оценка является: а) несмещенной? б) асимптотически нормальной? в) R-эффективной? Если «да», то почему, если «нет» — привести пример(ы).

35. Пусть $X_1, \ldots, X_{3n}$ — выборка объема $3n$ из распределения $N(a, 1)$. Является ли оценка $a^* = \dfrac{1}{n} \sum\limits_{i=n+1}^{2n} X_i$ параметра a: а) несмещенной? б) состоятельной (при $n \to \infty$)? в) эффективной? г) асимптотически нормальной (при $n \to \infty$)?

36. Пусть $X_1, \ldots, X_n$ — выборка из распределения $\Pi(\lambda)$. Построить оценку параметра λ, которая одновременно является а) состоятельной и смещенной; б) несостоятельной и несмещенной; в) состоятельной, но не асимптотически нормальной.

37. Пусть $X_1, \ldots, X_n$ — выборка из распределения $B(10, p)$. Построить оценку параметра p, которая одновременно является а) асимптотически нормальной и смещенной; б) несостоятельной и несмещенной; в) несмещенной, но не асимптотически нормальной.

38. Для всякого ли параметрического семейства распределений найдется $c > 0$ такое, что для любой несмещенной оценки θ^* неизвестного параметра θ выполняется неравенство $\mathsf{E}(\theta^* - \theta)^2 \geqslant \dfrac{c}{n}$?

39. Пусть элементы выборки $X_1, \ldots, X_n$, $n \geqslant 2$, имеют геометрическое распределение $G(p)$ с параметром $p \in (0, 1)$, то есть $\mathsf{P}_p(X_1 = k) = p(1-p)^{k-1}$ при $k = 1, 2, \ldots$.

а) Доказать, что статистика $S = n\overline{X}$ имеет распределение $\mathsf{P}_p(S = k) = C_{k-1}^{n-1} p^n (1-p)^{k-n}$ при $k = n, n+1, \ldots$.

б) Доказать, что статистика $S = n\overline{X}$ является достаточной и полной статистикой.

в) Какому классу K_b принадлежит оценка $p^* = I(X_1 = 1)$? Используя (а) и (б), построить эффективную оценку в этом классе.

40. Пусть $X_1, \ldots, X_n$ — выборка из нормального распределения $N(\theta, 1)$, где θ может принимать лишь значения 1 и 2. Построить оценку максимального правдоподобия для параметра θ.

§ 4. Доверительные интервалы

41. Пусть $X_1, \ldots, X_n$ — выборка из распределения $B(1, p)$. Имеется доверительный интервал $I_n = [\overline{X} - 1/\sqrt{n}, \overline{X} + 1/n]$.

а) Найти предел при $n \to \infty$ уровня доверия $\delta_n = \mathsf{P}_p(I_n \ni p)$

б) Используя доверительный интервал I_n, построить критерий для проверки гипотезы $H_1 : p = p_1$ против альтернативы $H_2 : p \neq p_1$. Найти предел вероятности ошибки первого рода этого критерия при $n \to \infty$.

42. Имеется выборка объема 3 из нормального распределения с неизвестными математическим ожиданием и дисперсией. Несмещенная выборочная дисперсия равна 1. Не пользуясь таблицами, построить точный доверительный интервал для неизвестной дисперсии уровня доверия 0.9.

Указание. Считать, что $\ln 0.05 = -3$; $\ln 0.95 = -0.05$.

§ 5. Гипотезы и критерии

43. Дана выборка $X_1, \ldots, X_n$. Гипотеза H_1: X_i имеют распределение с плотностью f_1. Альтернатива H_2: X_i имеют распределение с плотностью f_2. Здесь:

$$f_1(y) = 2\mathrm{e}^{-2(y-3)} \cdot I(y \in [3, \infty)); \quad f_2(y) = \mathrm{e}^{-(y-6)} \cdot I(y \in [6, \infty)).$$

а) Критерий δ предписывает принимать гипотезу H_1, если $\overline{X} \geqslant 7$; альтернативу H_2, если $\overline{X} < 7$. Найти пределы вероятностей ошибок первого и второго рода критерия δ при $n \to \infty$.

б) Построить наиболее мощный критерий асимптотического уровня $\varepsilon = 0{,}06$. Проверить состоятельность этого критерия.

44. Дана выборка $X_1, \ldots, X_n$. Гипотеза H_1: X_i имеют равномерное распределение на отрезке $[0, 2]$. Альтернатива H_2: X_i имеют равномерное распределение на отрезке $[1, 4]$.

а) Критерий $\delta_1 = \delta_1(X_1, \ldots, X_n)$ (при $n \geqslant 21$) предписывает принимать гипотезу H_1, если $X_2 \leqslant 1$; альтернативу H_2, если $X_2 > 1$. Найти размер и мощность этого критерия.

б) Построить (при $n = 1$) НМК размера $\alpha = 1/2$. Найти мощность этого критерия.

в) Построить (при $n = 1$) НМК размера $\alpha = 1/3$. Найти мощность этого критерия.

г) Критерий $\delta_2 = \delta_2(X_1, \ldots, X_n)$ предписывает принимать гипотезу H_1, если $X_{(n)} < 3/2$; альтернативу H_2, если $X_{(n)} \geqslant 3/2$. Найти размер и мощность критерия δ_2.

д) Является ли критерий δ_2 из предыдущего пункта состоятельным?

45. Дана выборка $X_1, \ldots, X_n$ объема n. Построить наиболее мощный критерий размера 1/2 для проверки гипотезы $H_1 : \vec{X} \Subset N_{0,1}$ против альтернативы $H_2 : \vec{X} \Subset B_{1/2}$.

46. По выборке объема n при заданной вероятности ошибки первого рода построить наиболее мощный критерий для различения двух простых гипотез относительно неизвестной дисперсии нормального распределения (математическое ожидание известно и равно нулю).

47. Дана выборка $X_1, \ldots, X_n$ объема n. Гипотеза H_1: X_i имеют распределение с плотностью f_1. Альтернатива H_2: X_i имеют распределение с плотностью f_2. Здесь $f_1(y) = \mathrm{e}^{-(y-6)} I(y \in [6, \infty))$; $f_2(y) = 2\mathrm{e}^{-2(y-3)} I(y \in [3, \infty))$.
Найти пределы вероятностей ошибок первого и второго рода следующих критериев при $n \to \infty$:

$$\delta_1(\vec{X}) = \begin{cases} H_1, & \overline{X} > 3{,}5 + 1/\sqrt{n}, \\ H_2, & \overline{X} \leqslant 3{,}5 + 1/\sqrt{n}; \end{cases} \quad \delta_2(\vec{X}) = \begin{cases} H_1, & \overline{X} > 3{,}5 + 1/n, \\ H_2, & \overline{X} \leqslant 3{,}5 + 1/n; \end{cases} \quad \delta_3(\vec{X}) = \begin{cases} H_1, & \overline{X} > 3{,}5, \\ H_2, & \overline{X} \leqslant 3{,}5. \end{cases}$$

48. Имеется нормальное распределение $N(a, 1)$, где число a неизвестно. Рассматриваются две простые гипотезы: основная $\{a = -1\}$ и альтернативная $\{a = 0\}$. Предлагается следующий статистический критерий для проверки этих гипотез: если $\overline{X} < -n^{\alpha}$, то принимается основная гипотеза; в противном случае принимается альтернативная гипотеза. Здесь n — объем выборки и α — заранее заданное известное вещественное число. Определить все числа α, при которых критерий является состоятельным.

49. Проверяется гипотеза о близости математических ожиданий двух независимых нормальных выборок объемов n и m с единичными дисперсиями: $\vec{X} = (X_1, \ldots, X_n) \Subset N(a_1, 1)$, $\vec{Y} = (Y_1, \ldots, Y_m) \Subset N(a_2, 1)$. Рассматривается основная гипотеза $H_1 : a_1 = a_2$ против альтернативы $H_2 : |a_1 - a_2| > 1$. Предлагается критерий: $\delta(\vec{X}, \vec{Y}) = H_2$, если $|\overline{X} - \overline{Y}| > 1$; $\delta(\vec{X}, \vec{Y}) = H_1$ иначе.

Проверить, является ли этот критерий состоятельным (при $n, m \to \infty$). Посчитать вероятность ошибки первого рода.

50. Дана выборка $X_1, \ldots, X_n$, элементы которой принимают целые неотрицательные значения $k = 0, 1, 2, 3$ и т. д. Рассматриваются две простые гипотезы: $H_1 : \mathsf{P}\{X_i{=}k\} = \dfrac{1}{k!}\, e^{-1}$, $H_2 : \mathsf{P}\{X_i{=}k\} = \dfrac{3^k}{k!}\, e^{-3}$.

Критерий δ предписывает принимать гипотезу H_1, если $\max(X_1, \ldots, X_n) < 2$, и альтернативу H_2 — в противном случае. Найти минимальный размер выборки, при котором мощность этого критерия превышает заданное значение.