Министерство образования и науки Российской Федерации

РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ НЕФТИ И ГАЗА
им. И. М. ГУБКИНА

Кафедра прикладной математики и компьютерного моделирования

В. К. Малиновский

Задачи и вопросы по курсу «Математическая статистика»

Рекомендовано в качестве учебного пособия
для студентов специальности
Прикладная математика

Москва 2004

Оглавление

Глава 1

Вероятностные основы статистики

1. Аксиоматика теории вероятностей

Задача 1.1. Дайте определение σ-алгебры на точечном множестве Ω. Приведите примеры.

Задача 1.2. Дайте определение измеримого пространства. Приведите примеры.

Задача 1.3. Дайте определение меры P на измеримом пространстве $(\Omega, \mathcal{F})$. Дайте определение конечной, σ-конечной, вероятностной меры. Приведите примеры.

Задача 1.4. Дайте определение вероятностного пространства $(\Omega, \mathcal{F}, \mathsf{P})$. Приведите примеры.

Задача 1.5. Дайте определение сингулярных, абсолютно непрерывных и эквивалентных мер на пространстве $(\Omega, \mathcal{F}, \mathsf{P})$. Приведите примеры. Дайте определение меры Лебега и считающей меры на вещественной прямой.

2. Случайные величины и измеримость

Задача 1.6. Пусть $(\mathrm{X}, \mathcal{B})$ — произвольное измеримое пространство. Дайте определение измеримого отображения $X : \Omega \to \mathrm{X}$.

Задача 1.7. Дайте определение случайной величины X и ее функции распределения.

Задача 1.8. Пусть $(\mathsf{R}, \mathcal{B})$ — вещественная прямая и заданная на ней борелевская σ-алгебра. Приведите пример отображения $X : \Omega \to \mathsf{R}$, которое измеримо при любом выборе $\mathcal{F}$.

Задача 1.9. Пусть $(\mathsf{R}, \mathcal{B})$ — вещественная прямая и заданная на ней борелевская σ-алгебра. Приведите пример отображения $X : \Omega \to \mathsf{R}$ и двух σ-алгебр $\mathcal{F}_1$ и $\mathcal{F}_2$ на Ω таких, что X измеримо относительно $\mathcal{F}_1$ и неизмеримо относительно $\mathcal{F}_2$.

3. Независимость

Заданные на вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P})$ случайные величины $X_1(\omega), \ldots, X_n(\omega)$ называются независимыми, если для любых измеримых множеств $B_1, \ldots, B_n$ на R

$$\mathsf{P}\{\omega : X_1(\omega) \in B_1, \ldots, X_n(\omega) \in B_n\} = \mathsf{P}\{X_1(\omega) \in B_1\} \ldots \mathsf{P}\{X_n(\omega) \in B_n\}.$$

Задача 1.10. Пусть X и Y — независимые случайные величины с функциями распределения F_X и F_Y. Докажите, что функция распределения суммы $X+Y$ представляется формулой свертки функций распределения F_X и F_Y:

$$[F_X * F_Y](x) = \int_{-\infty}^{\infty} F_X(x-z)dF_Y(z) = \int_{-\infty}^{\infty} F_Y(x-z)dF_X(z).$$

Задача 1.11. Выпишите формулу плотности распределения $[F_X * F_X](x)$, если распределение $F_X(x)$ имеет плотность $f_X(x)$ относительно меры Лебега. Как изменится эта формула, если с.в. X сосредоточена на положительной полуоси?

4. Теоретические характеристики распределений

Задача 1.12. Приведите определения степенных (центральных, нецентральных, абсолютных) моментов.

Задача 1.13. Приведите определения характеристической функции распределения вероятностей.

Задача 1.14. Приведите определение α-квантили распределения вероятностей. Чему равна 0,5-квантиль стандартного нормального распределения?

Задача 1.15. Приведите определения медианы, моды и эксцесса.

Задача 1.16. Приведите пример распределения, имеющего неединственную медиану.

Задача 1.17. Приведите пример распределения, имеющего единственную нулевую медиану, но не имеющего математического ожидания.

5. Виды сходимости

Пусть ξ_n — последовательность случайных величин, определенных на вероятностном пространстве $(\Omega, \mathcal{F}, \mathbf{P})$. Последовательность называется

(1) сходящейся по вероятности к случайной величине ξ, если для любого $\varepsilon > 0$

$$\mathbf{P}\{\omega : |\xi_n(\omega) - \xi(\omega)| > \varepsilon\} \to 0 \text{ при } n \to \infty,$$

(2) сходящейся почти наверное ($\mathbf{P}$-п.н.) к случайной величине ξ, если существует множество $N \in \mathcal{F}$ такое, что $\mathbf{P}(N) = 0$ и

$$\xi_n(\omega) \to \xi(\omega) \text{ при } n \to \infty \text{ для любого } \omega \notin N,$$

(3) сходящейся в среднем порядка r ($r > 0$) к случайной величине ξ, если $E|\xi_n|^r < \infty$ и

$$\mathbf{E}|\xi_n - \xi|^r \to 0 \text{ при } n \to \infty.$$

Если $r = 2$, то говорят, что последовательность ξ_n сходится к ξ в среднеквадратичном,

(4) сходящейся по распределению к случайной величине ξ с функцией распределения $F(x)$, если для $F_n(x) = \mathbf{P}\{\xi_n \leqslant x\}$

$$\lim_{n\to\infty} F_n(x) = F(x)$$

во всех точках непрерывности $F(x)$. Говорят также, что последовательность ξ_n слабо сходится к ξ.

Задача 1.18. Докажите, что если ξ_n сходится к ξ по вероятности, то ξ_n сходится к ξ по распределению.

Задача 1.19. Докажите, что если с.в. ξ постоянная, то ξ_n сходится к ξ по вероятности тогда и только тогда, когда ξ_n сходится к ξ по распределению.

Задача 1.20. Докажите, что если ξ_n сходится к ξ почти наверное, то ξ_n сходится к ξ по вероятности.

Задача 1.21. Докажите, что из сходимости ξ_n к ξ по вероятности не следует сходимость ξ_n к ξ почти наверное.

Задача 1.22. Докажите, что из сходимости ξ_n к ξ в средне-квадратичном (в среднем порядка $r > 0$) следует сходимость ξ_n к ξ по вероятности.

Литературные ссылки: [**6**], [**2**], [**7**], [**8**], [**10**], [**12**].

Глава 2

Специальные распределения

1. Одномерные распределения

Нормальное распределение. Случайная величина X имеет нормальное (гауссовское) распределение с параметрами μ и $\sigma^2 > 0$, если

$$F_X(x) = \int_{-\infty}^{x} f_X(z)dz,\ x \in \mathsf{R},$$

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\Big\{ -\frac{(x-\mu)^2}{2\sigma^2} \Big\},\ x \in \mathsf{R}.$$

Эти выражения обозначаются обычно через $\Phi_{(\mu,\sigma^2)}(x)$ и $\varphi_{(\mu,\sigma^2)}(x)$.

Задача 2.1. Пусть X имеет нормальное распределение с параметрами μ и $\sigma^2 > 0$. Докажите, что $\mathbf{E}(X) = \mu$, $\mathbf{E}(X^2) = \sigma^2 + \mu^2$, $\mathbf{D}(X) = \sigma^2$.

Задача 2.2. Пусть X имеет нормальное распределение с параметрами μ и $\sigma^2 > 0$. Докажите, что случайная величина $(X-\mu)/\sigma$ имеет стандартное нормальное распределение; выпишите равенства, связывающие функции распределения и плотности стандартного нормального закона и нормального закона с параметрами μ и σ^2.

Задача 2.3. Каково распределение суммы двух стандартных нормальных случайных величин?

Показательное (экспоненциальное) распределение. Случайная величина X имеет показательное (экспоненциальное) распределение с параметром $\lambda > 0$, если

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geqslant 0, \\ 0, & x < 0. \end{cases}$$

Плотность показательного (экспоненциального) распределения с параметром λ имеет вид

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geqslant 0, \\ 0, & x < 0. \end{cases}$$

Задача 2.4. Пусть с.в. X имеет показательное распределение с параметром $\lambda > 0$. Докажите, что $\mathbf{E}(X) = \frac{1}{\lambda}$, $\mathbf{E}(X^2) = \frac{2}{\lambda^2}$, $\mathbf{D}(X) = \frac{1}{\lambda^2}$.

Задача 2.5 (отсутствие последействия). Пусть с.в. X имеет показательное распределение, $U_X(x) = \mathbf{P}\{X > x\} = 1 - F_X(x)$. Докажите, что

$$U_X(x+y) = U_X(x)U_X(y), \quad x, y > 0.$$

Задача 2.6 (гамма-плотность). Пусть $X_1, \dots, X_n$ — независимые, показательно распределенные с параметром $\lambda > 0$ случайные величины. Докажите, что плотность $X_1 + \dots + X_n$ имеет вид

$$g_n(x) = \begin{cases} \lambda \frac{(\lambda x)^{n-1}}{(n-1)!} e^{-\lambda x}, & x \geqslant 0 \\ 0, & x < 0. \end{cases}$$

Задача 2.7 (гамма-распределение). Пусть $X_1, \dots, X_n$ — независимые, показательно распределенные с параметром $\lambda > 0$ случайные величины. Докажите, что функция распределения $X_1 + \dots + X_n$ имеет вид

$$G_n(x) = \begin{cases} 1 - e^{-\lambda x}\big(1 + \frac{\lambda x}{1!} + \dots + \frac{(\lambda x)^{n-1}}{(n-1)!}\big), & x \geqslant 0 \\ 0, & x < 0. \end{cases}$$

Равномерное распределение. Случайная величина X имеет равномерное распределение в интервале $[a, b]$ $(a < b)$, если ее плотность имеет вид

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b], \\ 0, & x \notin [a, b]. \end{cases}$$

Задача 2.8. Пусть случайная величина X имеет равномерное распределение в интервале $[a, b]$. Докажите, что $\mathbf{E}(X) = \frac{b+a}{2}$, $\mathbf{E}(X^2) = \frac{b^3-a^3}{3(b-a)}$, $\mathbf{D}(X) = \frac{(b-a)^2}{12}$.

Задача 2.9. Докажите, что если X — случайная величина, равномерно распределенная в интервале $[a, b]$, то $(X-a)/(b-a)$ равномерно распределена в интервале $[0, 1]$.

Задача 2.10. Докажите, что если случайная величина X с функцией распределения $F_X(x)$ имеет непрерывное распределение, то случайная величина $U = F_X(X)$ имеет равномерное распределение в интервале $[0, 1]$.

Треугольное распределение (Симпсона). Случайная величина X имеет треугольное распределение в интервале $[a, b]$ $(a < b)$, если ее плотность имеет вид

$$f_X(x) = \begin{cases} \frac{2}{b-a} - \frac{2}{(b-a)^2}|a + b - 2x|, & x \in [a, b], \\ 0, & x \notin [a, b]. \end{cases}$$

Задача 2.11. Докажите, что если X_1 и X_2 — независимые случайные величины, равномерно распределенные в интервале $[a/2, b/2]$, то случайная величина $X = X_1 + X_2$ имеет треугольное распределение в интервале $[a, b]$.

Задача 2.12. Найдите распределение суммы трех независимых случайных величин, равномерно распределенных в интервале $[0, 1]$.

Гамма-распределение. Случайная величина X имеет Гамма-распределение с параметрами $\alpha > 0$, $\lambda > 0$, если ее плотность имеет вид

$$f_X(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, & x > 0, \\ 0, & x \leqslant 0. \end{cases}$$

Задача 2.13. Пусть с.в. X имеет Гамма-распределение с параметрами α и λ. Докажите, что $\mathbf{E}(X) = \frac{\alpha}{\lambda}$, $\mathbf{E}(X^2) = \frac{\alpha(\alpha+1)}{\lambda^2}$, $\mathbf{D}(X) = \frac{\alpha}{\lambda^2}$.

Задача 2.14. Покажите, что при $\alpha = 1$ Гамма-распределение совпадает с показательным, а при $\alpha = n/2$ и $\lambda = 1/2$ — с χ^2-распределением с n степенями свободы.

Распределение Коши. Случайная величина X имеет распределение Коши с параметрами α, $\lambda > 0$, если ее плотность имеет вид

$$f_X(x) = \frac{1}{\pi} \frac{\lambda}{\lambda^2 + (x-\alpha)^2}.$$

Задача 2.15. Докажите, что математическое ожидание распределения Коши бесконечно.

Задача 2.16. Докажите, что параметр α распределения Коши является его модой и медианой.

Задача 2.17. Докажите, что нормированная сумма $\frac{1}{n}\sum_{i=1}^n X_i$ независимых, распределенных по закону Коши с параметрами α, λ случайных величин имеет то распределение Коши с параметрами α, λ.

Распределение Парето. Случайная величина X имеет распределение Парето с параметрами $x_0 > 0$, $\alpha > 0$, если

$$f_X(x) = \begin{cases} \frac{\alpha}{x_0}\left(\frac{x_0}{x}\right)^{\alpha+1}, & x > x_0, \\ 0, & x \leqslant x_0. \end{cases}$$

Задача 2.18. Докажите, что $\mathbf{E}(X^k) = \frac{\alpha}{\alpha-k}\, x_0^k$, если $k < \alpha$ и $\mathbf{E}(X^k) = +\infty$, если $k \geqslant \alpha$.

Задача 2.19. Докажите, что $\mathbf{D}(X) = \frac{\alpha}{(\alpha-1)(\alpha-2)}\, x_0^2$, если $\alpha > 2$ и $\mathbf{D}(X) = +\infty$, если $\alpha \leqslant 2$.

Распределение χ^2 с α степенями свободы. Случайная величина X имеет χ^2-распределение с α степенями свободы ($\alpha > 0$), если

$$f_X(x) = \begin{cases} (2^{\alpha/2}\Gamma(\alpha/2))^{-1}\, x^{\alpha/2-1} e^{-x/2}, & x > 0, \\ 0, & x \leqslant 0. \end{cases}$$

Пусть $X_1, \ldots, X_n$ — независимые случайные величины, имеющие стандартное нормальное распределение. Случайная величина $\sum_{i=1}^n X_i^2$ имеет χ^2-распределение с n степенями свободы.

Задача 2.20. Пусть с.в. X имеет χ^2 с α степенями свободы. Докажите, что $\mathbf{E}(X) = \alpha$, $\mathbf{E}(X^2) = \alpha(\alpha+2)$, $\mathbf{D}(X) = 2\alpha$.

Распределение Стьюдента (t-распределение). Случайная величина X имеет распределение Стьюдента с $\alpha > 0$ степенями свободы, если

$$f_X(x) = \frac{\Gamma((\alpha+1)/2)}{\sqrt{\alpha\pi}\,\Gamma(\alpha/2)}(1 + \frac{x^2}{\alpha})^{-(\alpha+1)/2}.$$

Если Z и Y — независимые случайные величины, причем Z имеет стандартное нормальное распределение, а Y — χ^2-распределение с n степенями свободы, то $X = Z\sqrt{n/Y}$ имеет распределение Стьюдента с n степенями свободы.

Задача 2.21. Пусть с.в. X имеет распределение Стьюдента с α степенями свободы. Докажите, что $\mathbf{E}(X^{2k-1}) = 0$, если $2k < \alpha$, и $\mathbf{E}(X^{2k-1}) = \infty$ в противном случае.

Распределение Бернулли. Случайная величина X имеет распределение Бернулли с параметром $p \in (0,1)$, если

$$\mathbf{P}\{X=1\}=p,\ \mathbf{P}\{X=0\}=1-p.$$

Функция распределения X имеет вид

$$F_X(x)=\begin{cases}0, & x<0,\\ 1-p, & 0\leqslant x<1,\\ 1, & x\geqslant 1.\end{cases}$$

Задача 2.22. Пусть с.в. X имеет распределение Бернулли с параметром $p\in(0,1)$. Докажите, что $\mathbf{E}(X^k)=p$, $\mathbf{D}(X)=p(1-p)$.

Геометрическое распределение. Случайная величина X имеет геометрическое распределение с параметром $p\in(0,1)$, если

$$\mathbf{P}\{X=k\}=p(1-p)^k,\ k=0,1,2,\ldots.$$

Задача 2.23. Пусть с.в. X имеет геометрическое распределение с параметром $p\in(0,1)$. Докажите, что $\mathbf{E}(X)=(1-p)/p$, $\mathbf{D}(X)=(1-p)/p^2$.

Биномиальное распределение. Случайная величина X имеет биномиальное распределение с параметрами n, p (n натуральное, $p\in(0,1)$), если

$$\mathbf{P}\{X=k\}=\binom{n}{k}p^k(1-p)^k,\ k=0,1,\ldots,n.$$

Функция распределения X имеет вид

$$F_X(x)=\begin{cases}0, & x<0,\\ \sum_{k=1}^{l}\binom{n}{k}p^k(1-p)^{n-k}, & l\leqslant x<l+1,\\ 1, & x\geqslant n.\end{cases}$$

Задача 2.24. Пусть X — биномиальная с.в. с параметрами n и p. Докажите, что $\mathbf{E}(X)=np$, $\mathbf{E}(X)^2=np+n(n-1)p^2$, $\mathbf{D}(X)=np(1-p)$.

Распределение Пуассона. Случайная величина X имеет распределение Пуассона с параметром $\lambda>0$, если

$$\mathbf{P}\{X=k\}=\frac{\lambda^k}{k!}e^{-\lambda},\ k=0,1,\ldots.$$

Задача 2.25. Пусть X — пуассоновская с.в. с параметром λ. Докажите, что $\mathbf{E}(X)=\lambda$, $\mathbf{E}(X)^2=\lambda^2+\lambda$, $\mathbf{D}(X)=\lambda$.

2. Распределения независимых выборок

Задача 2.26. Пусть $\mathbf{X}=(X_1,\ldots,X_n)$ — независимая выборка объема n, где X_i подчиняется нормальному распределению с параметрами μ и σ^2. Покажите, что

$$f_{\mathbf{X}}(x_1,\ldots,x_n)=(2\pi)^{-n/2}\sigma^{-n}\exp\Big\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\Big\}.$$

Задача 2.27. Пусть $\mathbf{X}=(X_1,\ldots,X_n)$ — независимая выборка объема n, где X_i подчиняется закону Пуассона с параметром $\lambda>0$. Покажите, что

$$\mathbf{P}\{X_1=x_1,\ldots,X_n=x_n\}=\frac{1}{\prod_{i=1}^{n}x_i!}\lambda^{\sum_{i=1}^{n}x_i}e^{-n\lambda}.$$

3. Некоторые специальные семейства распределений

Задача 2.28 ([**8**], стр. 62). Дайте определение экспоненциального семейства распределений .

Задача 2.29 ([**8**], стр. 62). Покажите, что биномиальное семейство $\binom{n}{x}p^x(1-p)^{n-x}$ является экспоненциальным семейством.

Задача 2.30. Дайте определение сдвигового семейства распределений. Приведите примеры.

Задача 2.31. Дайте определение масштабного семейства распределений. Приведите примеры.

Задача 2.32 ([**8**], стр. 82). Дайте определение семейства плотностей с монотонным отношением правдоподобия относительно некоторой статистики T.

Задача 2.33 ([**8**], стр. 372). Пусть $f_\theta(x) = g(x-\theta)$. Пользуясь критерием, согласно которому семейство f_θ имеет монотонное отношение правдоподобия относительно x тогда и только тогда, когда функция $u(x) = -\ln g(x)$ является выпуклой, привести пример семейства, не являющегося семейством с монотонным отношением правдоподобия относительно x.

4. Достаточные статистики

Задача 2.34 ([**8**], стр. 59). Дайте определение достаточной статистики.

Задача 2.35. Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — независимая выборка объема n из распределения Бернулли с параметром $p \in (0,1)$. Показать непосредственно (из определения), что выборочное среднее $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ является достаточной статистикой для параметра p.

Задача 2.36 ([**8**], стр. 61). Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — независимая выборка объема n. Сформулировать теорему Неймана (необходимое и достаточное условие достаточности статистики $T(\mathbf{X})$; теорема факторизации).

Задача 2.37. Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — независимая выборка объема n из распределения Бернулли с параметром $p \in (0,1)$. Показать с использованием теоремы Неймана, что выборочное среднее $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ является достаточной статистикой для параметра p.

Задача 2.38. Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — независимая выборка объема n из нормального закона $\Phi_{(\mu,\sigma^2)}(x)$. Показать с использованием теоремы Неймана, что пара $(\bar{X}, s_n^2)$, где $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ — выборочное среднее и $s_n^2 = \frac{1}{n}\sum_{i=1}^n (X_i^2 - \bar{X})^2$ — выборочная дисперсия, является достаточной статистикой для параметра (μ, σ^2).

Задача 2.39. Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — независимая выборка объема n из равномерного на интервале $[0,\theta]$ распределения, $\theta > 0$. Показать с использованием теоремы Неймана, что статистика $X_{(n)} = \max\{x_1, \ldots, X_n\}$ является достаточной статистикой для параметра θ.

Литературные ссылки: [**1**], [**7**], [**8**], [**11**].

Глава 3

Математическая статистика с точки зрения теории принятия решений

1. Игра двух лиц

Задача 3.1. Дайте определение игры двух лиц (т.е. пространства состояний природы Θ, множества решений статистика $\mathfrak{A}$, функции потерь статистика $L(\theta, a)$, $\theta \in \Theta$, $a \in \mathfrak{A}$). Приведите примеры.

2. Статистическая игра двух лиц

Задача 3.2. Дайте определение статистической игры двух лиц (т.е. тройки $(\Theta, \mathfrak{A}, L(\theta, a))$, множества исходов статистического эксперимента $(\mathrm{X}, \mathcal{B})$ и $\mathcal{P} = \{\mathbf{P}_\theta, \theta \in \Theta\}$, семейства распределений $\mathbf{X}(\omega) \in \mathrm{X}$ на $(\Omega, \mathcal{F})$). Дайте определение решающего правила $d : \mathrm{X} \to \mathfrak{A}$, зависящего от исхода статистического эксперимента. Приведите примеры.

Задача 3.3. Дайте определение функции потерь статистика $L(\theta, d(\mathbf{X}(\omega)))$, $\mathbf{X}(\omega) \in \mathrm{X}$, в статистической игре двух лиц. Приведите примеры.

Задача 3.4. Дайте определение функции риска $R(\theta, d) = \mathbf{E}_\theta L(\theta, d(\mathbf{X}))$ в статистической игре двух лиц. Приведите примеры.

Задача 3.5. Дайте определение рандомизированного (из множества D^*) и нерандомизированного (из множества D) решающего правила. Приведите примеры.

Задача 3.6. Дайте определение задачи проверки гипотез (принятия одного из двух решений) как статистической игры двух лиц. Приведите примеры.

Задача 3.7. Дайте определение задачи последовательной проверки гипотез (принятия одного из трех решений) как статистической игры двух лиц. Приведите примеры.

Задача 3.8. Дайте определение задачи точечного оценивания (принятия одного из бесконечного числа решений) как статистической игры двух лиц. Приведите примеры.

3. Принципы сравнения решений. Байесовский принцип

Задача 3.9. Дайте определение априорного распределения на Θ, $\tau \in \Theta^*$, и байесовского риска $r(\tau, \delta)$, $\tau \in \Theta^*$, $\delta \in D^*$.

Задача 3.10. Дайте определение байесовского относительно априорного распределения $\tau \in \Theta^*$ решающего правила.

Задача 3.11. Дайте определение ε-байесовского относительно априорного распределения $\tau \in \Theta^*$ решающего правила.

Задача 3.12. Пусть $\Theta = \{\theta_1, \dots, \theta_k\}$. Пусть

$$\mathfrak{S} = \{(y_1, \dots, y_k) : \text{для некоторого } \delta \in D^*,\ y_i = R(\theta_i, \delta),\ i = 1, 2, \dots, k\} \subset \mathsf{R}^k.$$

Докажите, что множество $\mathfrak{S}$ выпукло, т.е. для любых $y \in \mathfrak{S}$ и $y' \in \mathfrak{S}$ справедливо включение $\alpha y + (1-\alpha)y' \in \mathfrak{S}$.

Задача 3.13 ([**14**], стр. 34–42). Пусть $\Theta = \{\theta_1, \theta_2\}$ и

$$\mathfrak{S} = \{(y_1, y_2) : \text{для некоторого } \delta \in D^*,\ y_i = R(\theta_i, \delta), i = 1, 2\}.$$

На плоскости с координатными осями $y_1 = R(\theta_1, \delta)$ и $y_2 = R(\theta_2, \delta)$ построить

(1) множество точек одинакового байесовского риска относительно некоторого априорного распределения $\tau = (p_1, p_2)$, $p_1 + p_2 = 1$, $0 \leqslant p_i \leqslant 1$, $i = 1, 2$,

(2) множество точек, соответствующих ε-байесовским решающим правилам $\delta \in \mathfrak{S}$,

(3) множество точек, соответствующих байесовским решающим правилам $\delta \in \mathfrak{S}$.

4. Принципы сравнения решений. Минимаксный принцип

Задача 3.14. Дайте определение минимаксного решающего правила.

Задача 3.15. Дайте определение ε-минимаксного решающего правила.

Задача 3.16 ([**14**], стр. 34–42). Пусть $\Theta = \{\theta_1, \theta_2\}$ и

$$\mathfrak{S} = \{(y_1, y_2) : \text{для некоторого } \delta \in D^*,\ y_i = R(\theta_i, \delta), i = 1, 2\}.$$

На плоскости с координатными осями $y_1 = R(\theta_1, \delta)$ и $y_2 = R(\theta_2, \delta)$ построить множество точек, соответствующих минимаксным решающим правилам $\delta \in \mathfrak{S}$.

Задача 3.17. Дайте определение наименее благоприятного априорного распределения.

Литературные ссылки: [**4**], [**8**], гл.1, [**14**].

Глава 4

Задача проверки параметрических гипотез

1. Постановка задачи проверки гипотез

Задача 4.1. ([**8**], гл. 3, §1) Сформулируйте задачу проверки гипотезы H против альтернативы K на основе некоторой величины $\mathbf{X}$, распределение которой принадлежит некоторому параметрическому семейству. Дайте определение критической области, рандомизированного и нерандомизированного критерия $\varphi(\mathbf{X})$, ошибок первого и второго родов.

Задача 4.2. Сформулируйте задачу проверки гипотезы H против альтернативы K в случае независимой выборки $\mathbf{X} = (X_1, \ldots, X_n)$ объема n, где X_i подчиняется распределению из параметрического семейства $\mathcal{P} = \{F_\theta, \theta \in \Theta\}$.

Задача 4.3. ([**8**], стр. 74) Дайте определение функции мощности, размера и уровня значимости критерия $\varphi(\mathbf{X})$.

Задача 4.4. ([**8**], стр. 23) Дайте определение функции правдоподобия. Выпишите функцию правдоподобия для независимой выборки $\mathbf{X} = (X_1, \ldots, X_n)$ объема n, где X_i подчиняется

(1) нормальному распределению с параметрами μ и σ^2,
(2) закону Пуассона с параметром $\lambda > 0$,
(3) экспоненциальному распределению с параметром λ.

Задача 4.5. ([**8**], стр. 76) Дайте определение равномерно наиболее мощного критерия.

Задача 4.6. Дайте определение огибающей функций мощности критериев заданного размера.

2. Простая гипотеза и простая альтернатива. Фундаментальная лемма Неймана–Пирсона

Задача 4.7. ([**8**], стр. 78) Дайте определение критерия Неймана–Пирсона уровня значимости α для проверки простой гипотезы $H : \theta = \theta_0$ против простой альтернативы $K : \theta = \theta_1$.

Задача 4.8. ([**8**], стр. 78) Сформулируйте лемму Неймана–Пирсона.

Задача 4.9. Выпишите явный вид критерия Неймана–Пирсона уровня значимости α для проверки

(1) простой гипотезы $H : \mu = \mu_0$ против простой альтернативы $K : \mu = \mu_1$ для независимой выборки $\mathbf{X} = (X_1, \ldots, X_n)$ объема n, где X_i подчиняется нормальному распределению с единичной дисперсией,

(2) простой гипотезы $H : \sigma^2 = \sigma_0^2$ против простой альтернативы $K : \sigma^2 = \sigma_1^2$ для независимой выборки $\mathbf{X} = (X_1, \ldots, X_n)$ объема n, где X_i подчиняется нормальному распределению с нулевым средним,

(3) простой гипотезы $H : \lambda = \lambda_0$ против простой альтернативы $K : \lambda = \lambda_1$ для независимой выборки $\mathbf{X} = (X_1, \ldots, X_n)$ объема n, где X_i подчиняется закону Пуассона с параметром λ,

(4) простой гипотезы $H : \lambda = \lambda_0$ против простой альтернативы $K : \lambda = \lambda_1$ для независимой выборки $\mathbf{X} = (X_1, \ldots, X_n)$ объема n, где X_i подчиняется экспоненциальному распределению с параметром λ.

3. Простая гипотеза и односторонняя альтернатива

Задача 4.10. ([**8**], стр. 82–83) Пусть $\mathbf{X}$ — с.в., распределение которой принадлежит параметрическому семейству (θ — действительный параметр) с монотонным относительно статистики $T(\mathbf{X})$ отношением правдоподобия. Выпишите равномерно наиболее мощный критерий для проверки гипотезы $H : \theta \leqslant \theta_0$ против альтернативы $K : \theta > \theta_0$.

Задача 4.11. ([**8**], стр. 84) Пусть $\mathbf{X}$ — с.в., распределение которой принадлежит однопараметрическому экспоненциальному семейству. Выпишите равномерно наиболее мощный критерий для проверки гипотезы $H : \theta \leqslant \theta_0$ против альтернативы $K : \theta > \theta_0$.

Задача 4.12. Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — независимая выборка объема n, где X_i подчиняется нормальному распределению $\Phi_{(\mu,1)}(x)$ с единичной дисперсией и неизвестным средним. Выпишите явный вид равномерно наиболее мощного критерия для проверки гипотезы $H : \mu \leqslant 0$ против альтернативы $K : \mu > 0$.

4. Сложная гипотеза и двусторонняя альтернатива

Задача 4.13. ([**8**], стр. 105) Пусть $\mathbf{X}$ — с.в., распределение которой принадлежит однопараметрическому экспоненциальному семейству. Выпишите равномерно наиболее мощный критерий для проверки гипотезы $H : \theta \leqslant \theta_1$ или $\theta \geqslant \theta_2$ ($\theta_1 < \theta_2$) против альтернативы $K : \theta_1 < \theta < \theta_2$.

Задача 4.14. Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — независимая выборка объема n, где X_i подчиняется нормальному распределению $\Phi_{(\mu,1)}(x)$ с единичной дисперсией и неизвестным средним. Выпишите явный вид равномерно наиболее мощного критерия для проверки гипотезы $H : \mu \leqslant \mu_1$ или $\mu \geqslant \mu_2$ ($\mu_1 < \mu_2$) против альтернативы $K : \mu_1 < \mu < \mu_2$.

5. Асимптотическая постановка. Сближающиеся альтернативы

Задача 4.15 ([**2**], стр. 284–293). Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — независимая выборка объема n, где X_i подчиняется распределению из параметрического семейства $\mathcal{P} = \{F_\theta, \theta \in [0, 1]\}$.

(1) Выпишите функцию логарифма отношения правдоподобия $\Lambda_{t,n}$ для проверки гипотезы $H : \theta = 0$ против альтернативы $K : \theta_{t,n} = \frac{t}{\sqrt{n}}$, $t > 0$.

(2) Пусть $\frac{\partial^k}{\partial \theta^k} \ln f_\theta(x) \mid_{\theta=0}$ обозначается через $l_0^{(k)}(x)$. Докажите, что

$$\mathbf{E}_0 l_0^{(1)}(X_i) = 0, \quad \mathbf{E}_0 l_0^{(2)}(X_i) = -\mathbf{E}_0 [l_0^{(1)}(X_i)]^2.$$

(3) Пусть $\mathsf{E}_0[l_0^{(1)}(X_i)]^2 > 0$ обозначается через D^2. При каких условиях распределение

$$\mathsf{P}_0(\Lambda_{t,n} \leqslant x)$$

логарифма отношения правдоподобия $\Lambda_{t,n}$ асимптотически нормально с параметрами $(-\frac{t^2}{2}D^2, t^2D^2)$?

(4) Воспользовавшись неравенством Иенсена, докажите, что

$$\mathsf{E}_0 \ln \frac{f_{\theta_{t,n}}(X_i)}{f_0(X_i)} < 0, \quad \mathsf{E}_{\theta_{t,n}} \ln \frac{f_{\theta_{t,n}}(X_i)}{f_0(X_i)} > 0.$$

Задача 4.16 ([**2**], стр. 292). Дайте определение асимптотически наиболее мощного критерия в задаче различения сближающихся альтернатив.

Литературные ссылки: [**1**], [**2**], [**5**], [**7**], [**8**], [**13**].

Глава 5

Задача проверки непараметрических гипотез

1. Эмпирическая функция распределения

Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — независимая выборка объема n, где X_i подчиняется распределению $F_X(x)$.

Задача 5.1. Дайте определение вариационного ряда $X_{(k)}$, $k = 1, 2, \ldots, n$.

Задача 5.2. Докажите, что $\mathsf{P}\{X_{(n)} < x\} = F_X(x)^n$ и $\mathsf{P}\{X_{(1)} < x\} = 1 - (1 - F_X(x))^n$.

Задача 5.3. Докажите, что

$$\mathsf{P}\{X_{(n)} - X_{(1)} < x\} = n \int_{-\infty}^{\infty} [F_X(x+z) - F_X(z)]^{n-1} dF_X(z).$$

Задача 5.4. Докажите эквивалентность представлений эмпирической функции распределения

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{(-\infty, x]}(X_i)$$

и

$$F_n(x) = \begin{cases} 0, & x \leqslant X_{(1)}, \\ \frac{k}{n}, & X_k < x \leqslant X_{(k+1)},\ 1 \leqslant k \leqslant n-1, \\ 1, & x > X_{(n)}. \end{cases}$$

Задача 5.5. Докажите, что $\mathsf{E}(F_n(x)) = F_X(x)$.

Задача 5.6. Докажите, что $\mathsf{D}(F_n(x)) = \frac{1}{n} F_X(x)[1 - F_X(x)]$.

Задача 5.7. Докажите, что $\mathsf{P}\big\{F_n(x) = \frac{k}{n}\big\} = \binom{n}{k} [F_X(x)]^k [1 - F_X(x)]^{n-k}$.

2. Критерии согласия Колмогорова–Смирнова

Задача 5.8. Докажите, что для любого $x \in \mathsf{R}$ распределение статистики $\sqrt{n} D_n(x) = \sqrt{n}(F_n(x) - F_X(x))$ при $n \to \infty$ асимптотически нормально с нулевым математическим ожиданием и с дисперсией $F_X(x)[1 - F_X(x)]$. Как использовать этот факт для проверки гипотезы $F_X(x)$?

Задача 5.9. Сформулируйте теорему Гливенко.

Задача 5.10. Сформулируйте теорему Колмогорова.

Задача 5.11. Пусть

$$D_n = \sup_{-\infty < x < \infty} |F_n(x) - F_X(x)|$$

и d_α — квантиль распределения Колмогорова $K(z) = \sum_{k=-\infty}^{\infty}(-1)^k \exp\{-2k^2z^2\}$, т.е.

$$1 - K(d_\alpha) = \alpha.$$

Опишите последовательность действий для построения критерия, проверяющего гипотезу $F_X(x)$ на основе статистики Колмогрова и независимой выборки $\mathbf{X} = (X_1, \ldots, X_n)$ большого объема n.

Задача 5.12. Сформулируйте теорему Смирнова. Опишите альтернативы и последовательность действий при применении критериев Смирнова.

3. Критерий χ^2

Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — независимая выборка объема n. Для проверки простой гипотезы о том, что с.в. X_i подчиняется дискретному распределению $(p_1, \ldots, p_m)$, $\sum_{r=1}^{m} p_r = 1$, $p_r \in [0,1]$, $r = 1, 2, \ldots, m$, или $p_r = \mathsf{P}\{X_i = r\}$, $r = 1, 2, \ldots, m$, используется статистика

$$\chi^2 = \sum_{k=1}^{m} \frac{(\nu_k - np_k)^2}{np_k},$$

где $\nu_1, \ldots, \nu_m$ — частоты результатов наблюдений в выборке $\mathbf{X}$,

$$\nu_r = \sum_{k=1}^{n} \mathbf{1}_{\{X_k = r\}},\ r = 1, 2, \ldots, m.$$

Задача 5.13. Показать, что $\sum_{r=1}^{m} \nu_r = n$.

При $n \to \infty$ распределение статистики χ^2 стремится к χ^2-распределению с $m-1$ степенью свободы. Плотность χ^2-распределения с $m-1$ степенью свободы имеет вид (см. стр. 6)

$$f_X(x) = \begin{cases} (2^{(m-1)/2}\Gamma((m-1)/2))^{-1}\, x^{(m-1)/2-1} e^{-x/2}, & x > 0, \\ 0, & x \leqslant 0 \end{cases}$$

и не зависит от распределения $(p_1, \ldots, p_m)$.

Критерий χ^2 определяется критической областью $\{\chi^2 > \chi^2_\alpha\}$, где χ^2_α выбирается из условия на размер критерия

$$\mathsf{P}\{\chi^2 > \chi^2_\alpha\} = \alpha.$$

Задача 5.14. Воспользовавшись критерием χ^2, на основе последовательности из n независимых подбрасываний монеты проверить гипотезу о том, что эта монета «правильная».

Задача 5.15. Воспользовавшись критерием χ^2, на основе n независимых выбрасываний игральной кости проверить гипотезу о том, что эта кость «правильная».

Литературные ссылки: [**1**], [**2**], [**7**].

Глава 6

Задача точечного оценивания

1. Несмещенные и асимптотически несмещенные оценки

Пусть $\mathbf{X} = (X_1, \dots, X_n)$ — независимая выборка объема n, где X_i подчиняется распределению из параметрического семейства $\mathcal{P} = \{F_\theta, \theta \in \Theta\}$. Статистическая оценка $\delta(\mathbf{X})$ параметра θ называется несмещенной, если

$$\mathsf{E}_\theta \delta(\mathbf{X}) = \theta \text{ при всех } \theta \in \Theta.$$

Последовательность оценок $\delta_n(\mathbf{X})$ параметра θ называется асимптотически несмещенной[1], если

$$\lim_{n\to\infty} \mathsf{E}_\theta \delta_n(\mathbf{X}) = \theta \text{ при всех } \theta \in \Theta.$$

Задача 6.1. Привести пример несмещенной оценки.

Задача 6.2. Привести пример асимптотически несмещенной последовательности оценок.

Задача 6.3. Пусть $\mathbf{X} = (X_1, \dots, X_n)$ — независимая выборка объема n из нормального закона $\Phi_{(\mu,\sigma^2)}(x)$. Показать, что выборочная дисперсия $s_n^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2$ является смещенной оценкой σ^2 со смещением $-\sigma^2/n$. Показать, что последовательность s_n^2 является асимптоически несмещенной для σ^2. Показать, что $\lim_{n\to\infty} \mathsf{D}(s_n^2) = 0$.

2. Состоятельные оценки

Пусть $\mathbf{X} = (X_1, \dots, X_n)$ — независимая выборка объема n, где X_i подчиняется распределению из параметрического семейства $\mathcal{P} = \{F_\theta, \theta \in \Theta\}$. Последовательность статистических оценок $\delta_n(\mathbf{X})$ параметра θ называется состоятельной[2], если

$$\lim_{n\to\infty} \mathsf{P}_\theta\{|\delta_n(\mathbf{X}) - \theta| > \varepsilon\} = 0 \quad \text{для любого } \theta \in \Theta.$$

Задача 6.4. Пусть T_n — последовательность несмещенных оценок параметра θ и пусть $\mathsf{D}_\theta(T_n) \to 0$ при $n \to \infty$. Доказать, что последовательность T_n является состоятельной для θ.

Задача 6.5. Пусть T_n — асимптотически несмещенная последовательность оценок параметра θ и пусть $\mathsf{D}_\theta(T_n) \to 0$ при $n \to \infty$. Доказать, что посдедовательность T_n является состоятельной для θ.

Задача 6.6. Пусть $\mathbf{X} = (X_1, \dots, X_n)$ — независимая выборка объема n из распределения Бернулли с параметром $p \in (0, 1)$. Показать, что выборочное среднее $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ является состоятельной оценкой параметра p.

[1] Для краткости говорят, что сама оценка δ_n является асимптотически несмещенной.
[2] Для краткости говорят, что сама оценка δ_n является состоятельной.

Задача 6.7. Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — независимая выборка объема n из распределения Пуассона с параметром $\lambda > 0$. Показать, что выборочное среднее $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ является состоятельной оценкой параметра λ.

Задача 6.8. Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — независимая выборка объема n из нормального закона $\Phi_{(\mu,1)}(x)$. Показать, что выборочное среднее $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ является состоятельной оценкой для μ.

Задача 6.9. Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — независимая выборка объема n из распределения Коши с плотностью

$$f_X(x) = \frac{1}{\pi}\frac{1}{1+(x-\theta)^2}.$$

Показать, что выборочное среднее $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ не является состоятельной оценкой параметра θ, т.к.

$$\mathbf{P}\{|\bar{X}_n - \theta| \geqslant \varepsilon\} = \mathbf{P}\{|X_1 - \theta| \geqslant \varepsilon\} = 1 - \frac{2}{\pi}\operatorname{arctg}\varepsilon.$$

Задача 6.10. Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — независимая выборка объема n из нормального закона $\Phi_{(\mu,\sigma^2)}(x)$. Показать, что последовательность выборочных дисперсий $s_n^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2$ является состоятельной последовательностью смещенных оценок параметра σ^2.

Задача 6.11. Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — независимая выборка объема n из нормального закона $\Phi_{(\mu,\sigma^2)}(x)$. Показать, что $S_n^2 = \frac{n}{n-1}s_n^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$ является состоятельной последовательностью несмещенных оценок параметра σ^2.

Задача 6.12. Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — независимая выборка объема n из нормального закона $\Phi_{(\mu,\sigma^2)}(x)$. Показать, что $\mathbf{E}(S_n^2 - \sigma)^2 = \frac{2}{n-1}\sigma^4$ и $\mathbf{E}(s_n^2 - \sigma)^2 = \frac{2n-1}{n^2}\sigma^4$. Проверить, что $\frac{2}{n-1} > \frac{2n-1}{n^2}$ при $n \geqslant 2$. Интерпретировать в терминах качества оценок.

Задача 6.13. Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — независимая выборка объема n из нормального закона $\Phi_{(\mu,\sigma^2)}(x)$. Пусть $\mathcal{K}$ — класс оценок вида kS_n^2, $k > 0$. Показать, что оценка $\frac{1}{n+1}\sum_{i=1}^n (X_i - \bar{X})^2$ имеет минимальный в классе $\mathcal{K}$ квадратичный риск $\frac{2}{n+1}\sigma^4$. Показать, что смещение этой оценки равно $-\frac{2}{n+1}\sigma^2$. Сравнить со смещением и квадратичным риском оценок s_n^2 и S_n^2.

3. Эффективные и асимптотически эффективные оценки

Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — независимая выборка объема n, пусть $L(\mathbf{X}, \theta) = \prod_{i=1}^n f_\theta(X_i)$ — функция правдоподобия, $\Lambda(\mathbf{X}, \theta) = \frac{\partial}{\partial\theta}\ln L(\mathbf{X}, \theta)$. Неравенством информации Рао–Крамера для статистики $T_n(\mathbf{X})$ называется неравенство

$$\mathbf{D}_\theta T_n(\mathbf{X}) \geqslant I_n(\theta)^{-1},$$

где $I_n(\theta) = \mathbf{D}_\theta \Lambda(\mathbf{X}, \theta)$ — количество информации Фишера.

Несмещенная оценка $T_n = T(\mathbf{X})$ параметра θ называется эффективной, если

$$\mathbf{D}_\theta T_n(\mathbf{X}) = I_n(\theta)^{-1}.$$

Задача 6.14. Показать, что если $\lim_{n\to\infty} I_n(\theta) = 0$, то последовательность $T_n = T(\mathbf{X})$ является состоятельной оценкой параметра θ.

Задача 6.15. Сформулировать условия, при которых для $\mathbf{X} = (X_1, \ldots, X_n)$, независимой выборки объема n,

$$I_n(\theta) = n\mathsf{D}_\theta\Big(\frac{\partial}{\partial\theta}\ln f_\theta(X_1)\Big).$$

Задача 6.16. Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — независимая выборка объема n из распределения Бернулли с параметром $p \in (0, 1)$. Проверить,

(1) что выборочное среднее $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ является несмещенной оценкой параметра p,
(2) что $I_n(p) = \frac{n}{p(1-p)}$,
(3) что $\bar{X}$ является эффективной оценкой параметра p.

Литературные ссылки: [**1**], [**2**], [**3**], [**5**], [**7**], [**9**], [**13**].

Глава 7

Методы построения точечных оценок

1. Метод моментов

Пусть $\mathbf{X} = (X_1 \ldots, X_n)$ — независимая выборка объема n, где X_i подчиняется распределению, принадлежащего семейству $\mathcal{P} = \{F_\theta, \theta \in \Theta\}$ и пусть $\Theta \subset \mathsf{R}^d$. Пусть существуют первые d степенных моментов,

$$m_r(\theta) = \int_{\mathsf{R}} x^r dF_\theta(x), \quad r = 1, 2, \ldots, d.$$

Соответствующие выборочные моменты — это

$$\bar{m}_r = \frac{1}{n}\sum_{k=1}^{n} X_k^r, \quad r = 1, 2, \ldots, d.$$

Метод моментов состоит в приравнивании теоретических и выборочных моментов. Это дает систему уравнений

$$m_r(\theta) = \bar{m}_r, \quad r = 1, 2, \ldots, d,$$

откуда ищется d неизвестных координат параметра $\theta = (\theta_1, \ldots, \theta_d) \in \mathsf{R}^d$. Если существует единственное решение $\hat{\theta}_r = \theta_r(\bar{m}_1, \ldots, \bar{m}_d)$, $r = 1, 2, \ldots, d$, то оно называется оценкой по методу моментов.

2. Метод максимального правдоподобия

Пусть $\mathbf{X} = (X_1 \ldots, X_n)$ — независимая выборка объема n, где X_i подчиняется распределению, принадлежащего семейству $\mathcal{P} = \{F_\theta, \theta \in \Theta\}$, и пусть f_θ — соответствующие плотности относительно доминирующей меры ν. Метод максимального правдоподобия состоит в нахождении такой оценки $\hat{T}(\mathbf{X})$, что

$$L(\mathbf{X}, \hat{T}(\mathbf{X})) = \sup_{\theta \in \Theta} L(\mathbf{X}, \theta),$$

где $L(\mathbf{X}, \theta) = \prod_{i=1}^n f_\theta(X_i)$ является функцией правдоподобия. Уравнением правдоподобия называется уравнение $\frac{\partial}{\partial\theta} L(\mathbf{X}, \theta) = 0$ или эквивалентное уравнение $\frac{\partial}{\partial\theta} \ln L(\mathbf{X}, \theta) = 0$.

При определенных условиях регулярности (см. [**7**], стр. 558) решение уравнения правдоподобия $\hat{T}(\mathbf{X})$ представляет собой состоятельную, асимптотически эффективную и асимптотически нормальную оценку параметра θ. Последнее означает, что при $n \to \infty$ распределение

$$\mathsf{P}_\theta(I(\theta)\sqrt{n}(\hat{T}(\mathbf{X}) - \theta) \leqslant x),$$

где $I(\theta) = \mathsf{E}_\theta[\frac{\partial}{\partial\theta} \ln f_\theta(X_1)]^2$, сближается с $\Phi_{(0,1)}(x)$ равномерно по x.

Задача 7.1. Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — независимая выборка объема n, где X_i подчиняется нормальному закону $\Phi_{(\mu,\sigma^2)}(x)$. Найдите оценки максимального правдоподобия для (μ, σ^2).

Задача 7.2. Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — независимая выборка объема n, где X_i подчиняется закону Пуассона с параметром $\lambda > 0$. Покажите, что оценкой максимального правдоподобия является выборочное среднее $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Задача 7.3. Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — независимая выборка объема n, где X_i подчиняется геометрическому распределению с параметром $p \in [0, 1)$, т.е. $\mathbf{P}\{X_i = k\} = p(1-p)^{k-1}$, $k = 1, 2, \ldots$. Покажите, что оценкой максимального правдоподобия является $(\bar{X})^{-1} = n / \sum_{i=1}^{n} X_i$.

Задача 7.4. Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — независимая выборка объема n, где X_i подчиняется распределению с плотностью

$$f_\theta(x) = \begin{cases} 1-\theta, & -\frac{1}{2} \leqslant x < 0, \\ 1+\theta, & 0 \leqslant x \leqslant \frac{1}{2} \end{cases}$$

и $|\theta| < 1$. Покажите, что оценкой максимального правдоподобия является $\frac{1}{n}\Big(2\sum_{i=1}^{n} \mathbf{1}_{\{X_i \geqslant 0\}} - n\Big)$.

Задача 7.5. Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — независимая выборка объема n, где X_i подчиняется гамма-распределению с параметрами $p > 0$ и $\sigma > 0$,

$$f_X(x) = \begin{cases} (\sigma^p \Gamma(p))^{-1} x^{p-1} \exp\big\{-\frac{x}{\sigma}\big\}, & x > 0, \\ 0, & x \leqslant 0, \end{cases}$$

так что $\mathbf{E} X_i = p\sigma$, $\mathbf{D}\, X_i = p\sigma^2$, $i = 1, 2, \ldots, n$. Докажите, что функция правдоподобия имеет вид

$$L(\mathbf{X}, \sigma, p) = (\sigma^p \Gamma(p))^{-n} \Big(\prod_{i=1}^{n} X_i\Big)^{p-1} \exp\Big\{-\frac{1}{\sigma}\sum_{i=1}^{n} X_i\Big\} h(X_{(1)}).$$

Выпишите уравнение правдоподобиядля σ при известном p. Является ли оценка максимального правдоподобия $\frac{1}{p}\bar{X}$ для параметра σ несмещенной? Является ли последовательность этих оценок (при $n \to \infty$) состоятельной для σ?

3. Метод наименьших квадратов

Пусть $(X_1, \ldots, X_n)$ — наблюдения за числом θ,

$$X_i = \theta + \varepsilon_i, \ \ i = 1, 2, \ldots, n.$$

Оценкой по методу наименьших квадратов ($\mathbf{E}\varepsilon_i = 0$, $\mathbf{E}\varepsilon_i^2 < \infty$) называется такая оценка $\hat{T}(\mathbf{X})$, что

$$\sum_{i=1}^{n} (X_i - \hat{T}(\mathbf{X}))^2 = \min_{T(\mathbf{X})} \sum_{i=1}^{n} (X_i - T(\mathbf{X}))^2.$$

Задача 7.6. Пусть $\mathbf{E}\varepsilon_i = 0$, $\mathbf{E}\varepsilon_i^2 = \sigma^2$, $i = 1, 2, \ldots, n$ (прямые равноточные измерения θ). Найти оценку по методу наименьших квадратов для θ.

Литературные ссылки: [**2**], [**3**], [**5**], [**7**], [**9**], [**13**].

Глава 8

Задача интервального оценивания

Пусть $\mathcal{P} = \{F_\theta, \theta \in \Theta\}$ — однопараметрическое семейство функций распределения, где Θ — интервал вещественной прямой R.

Интервал $[\underline{T}(\mathbf{X}), \bar{T}(\mathbf{X})] \subset \Theta$, границы которого зависят от выборки $\mathbf{X} = (X_1, \ldots, X_n)$, где X_i подчиняется закону из $\mathcal{P}$, называют доверительным интервалом с доверительным уровнем $1 - \alpha$, если

$$\mathsf{P}_\theta\{\theta \in [\underline{T}(\mathbf{X}), \bar{T}(\mathbf{X})]\} \geqslant 1 - \alpha \text{ для всех } \theta \in \Theta.$$

Задача 8.1 ([**8**], гл. 3, §5). Сформулируйте результаты о связи равномерно наиболее точных доверительных границ с равномерно наиболее мощными критериями.

Задача 8.2 ([**8**], стр. 96). Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — независимая выборка объема n, где X_i подчиняется распределению Пуассона с параметром $\lambda > 0$. Пусть $\alpha \in (0, 1)$. Построить равномерно наиболее точную верхнюю доверительную границу для λ с уровнем $1 - \alpha$.

Задача 8.3 ([**8**], стр. 97). Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — независимая выборка объема n, где X_i подчиняется распределению Бернулли с параметром $p \in (0, 1)$. Пусть $\alpha \in (0, 1)$. Построить равномерно наиболее точную верхнюю доверительную границу для p с уровнем $1 - \alpha$.

Литературные ссылки: [**1**], [**2**], [**5**], [**7**], [**8**], [**13**].

Глава 9

Вопросы для подготовки к экзамену

1. Вероятностные основы статистики (вероятностное пространство, случайные величины, функции распределения, независимость, моменты, характеристические функции и проч.). Различные постановки задач математической статистики. Примеры.
2. Специальные распределения (нормальное, пуассоновское, гамма, Коши, показательное, бернуллиевское, биномиальное, равномерное, Фишера, Парето, χ^2 и проч.), их свойства. Основные характеристики распределений (моменты, медиана, мода и проч.) и соотношения между ними. Параметрические (в том числе – сдвиговые и масштабные) семейства.
3. Постановка задач математической статистики с точки зрения теории принятия решения. Примеры.
4. Оптимальные решающие правила. Различные подходы к определению оптимальности. Геометрическая интерпретация байесовских и минимаксных правил. Примеры.
5. Задача проверки гипотез. Определения. Лемма Неймана–Пирсона (с доказательством). Примеры.
6. Задача проверки сложных гипотез. Определения. Равномерно наиболее мощные критерии. Примеры.
7. Достаточные статистики. Теорема Неймана. Экспоненциальные семейства. Критерии достаточности. Монотонное отношение правдоподобия. Примеры.
8. Асимптотические постановки задачи проверки гипотез. Огибающая функций мощности. Примеры.
9. Эмпирическое распределение, вариационный ряд, выборочные моменты и их свойства. Теоремы Гливенко, Колмогорова и Смирнова. Непараметрические критерии проверки гипотез. Подход к доказательству этих теорем путем аппроксимации винеровскими процессами (в т.ч. броуновским мостом). Примеры.
10. Задача точечного оценивания. Состоятельность, несмещенность. Примеры.
11. Задача точечного оценивания. Асимптотическая состоятельность и асимптотическая несмещенность. Примеры.
12. Задача точечного оценивания. Неравенство информации. Эффективность. Асимптотическая эффективность. Примеры.
13. Метод максимального правдоподобия, метод моментов. Состоятельность и асимптотическая нормальность о.м.п. Примеры.
14. Задача интервального оценивания. Доверительные интервалы. Асимптотические доверительные интервалы. Примеры.

Литература

[1] Большев Л.Н., Смирнов Н.В. (1965) *Таблицы математической статистики.* М.: Наука.

[2] Боровков А.А. (1984) *Математическая статистика. Оценка параметров, проверка гипотез.* М.: Наука.

[3] Воинов В.Г., Никулин М.С. (1989) *Несмещенные оценки и их применения.* М.: Наука.

[4] Де Гроот М. (1974) *Оптимальные статистические решения.* М.: Наука.

[5] Закс Ш. (1975) *Теория статистических выводов.* М.: Мир.

[6] Колмогоров А.Н., Фомин С.В. (1976) *Элементы теории функций и функционального анализа.* Изд. 4-е. М.: Наука.

[7] Королюк В.С., Портенко Н.И., Скороход А.В., Турбин А.Ф. (1985) *Справочник по теории вероятностей и математической статистике.* М.: Наука.

[8] Леман Э. (1979) *Проверка статистических гипотез.* М.: Наука.

[9] Леман Э. (1991) *Теория точечного оценивания.* М.: Наука.

[10] Неве Ж. (1969) *Математические основы теории вероятностей.* М.: Мир.

[11] Феллер В. (1967) *Введение в теорию вероятностей и ее приложения.* т. 2, М.: Мир.

[12] Ширяев А.Н. (1980) *Вероятность.* М.: Наука.

[13] Шметтерер Л. (1976) *Введение в математическую статистику.* М.: Наука.

[14] Ferguson T.S. (1967) *Mathematical Statistics. A Decision Theoretic Approach.* Academic Press.