

Задачи по математической статистике

1. По наблюдавшимся значениям выборки $\vec{X}$ построить эмпирическую функцию распределения, если а) $\vec{X} = (2, 4, 0, 1, 2, 3, 0, 0)$; б) $\vec{X} = (0, 3, 4, 1, 0, 2, 0, 2)$; в) $\vec{X} = (2, 4, 2, 0, 1, 0, 4, 3, 0, 1, 0, 2, 3, 0, 2, 0)$.
2. Можно ли по эмпирической функции распределения восстановить выборку и вариационный ряд, если а) объем выборки известен; б) объем выборки неизвестен.
3. В терминах общей функции распределения элементов выборки найти функцию распределения а) максимального члена вариационного ряда $X_{(n)}$; б) минимального члена вариационного ряда $X_{(1)}$; в) k -ой порядковой статистики $X_{(k)}$.
4. По выборке объема n из распределения Бернулли с параметром p построить график эмпирической функции распределения.
5. Построить оценки неизвестных параметров по методу моментов для следующих распределений: а) B_p , $0 < p < 1$; б) $B_{m,p}$, $0 < p < 1$, m — известно; в) Π_λ , $\lambda > 0$; г) G_p , $0 < p < 1$; д) $U[0, \theta]$, $\theta > 0$; е) $U[\theta - 1, \theta + 1]$, $-\infty < \theta < \infty$; ж) $U[-\theta, \theta]$, $\theta > 0$; з) E_α , $\alpha > 0$; и) $N_{\alpha,1}$, $-\infty < \alpha < \infty$; к) N_{0,σ^2} , $\sigma^2 > 0$; л) N_{α,σ^2} , $-\infty < \alpha < \infty$, $\sigma^2 > 0$; м) $U[a, b]$, $-\infty < a < b < \infty$. Исследовать полученные оценки на несмещенность (если возможно) и состоятельность.
6. Построить оценки максимального правдоподобия неизвестных параметров для следующих распределений: а) B_p , $0 < p < 1$; б) Π_λ , $\lambda > 0$; в) $B_{m,p}$, $0 < p < 1$; г) $U[0, \theta]$, $\theta > 0$; д) $U[-\theta, \theta]$, $\theta > 0$; е) $U[-\theta, 0]$, $\theta > 0$; ж) $U[\theta, \theta + 1]$, $-\infty < \theta < \infty$; з) E_α , $\alpha > 0$; и) $N_{\alpha,1}$, $-\infty < \alpha < \infty$; к) N_{0,σ^2} , $\sigma^2 > 0$. Исследовать полученные оценки на несмещенность и состоятельность.
7. Построить оценку метода моментов и оценку максимального правдоподобия параметра θ , если распределение выборки имеет плотность $f_\theta(y) = \frac{2y}{\theta^2}$ при $y \in [0, \theta]$, $\theta > 0$.
8. Пусть $X_1, \dots, X_n$ — выборка из смещённого показательного распределения с плотностью

$$f_\beta(y) = \begin{cases} e^{\beta-y} & \text{при } y \geq \beta, \\ 0 & \text{при } y < \beta. \end{cases}$$

Для параметра $\beta \in \mathbb{R}$ построить а) оценку по методу моментов; б) оценку максимального правдоподобия. Исследовать полученные оценки на несмещенность и состоятельность.

9. Пусть $X_1, \dots, X_n$ — выборка из некоторого распределения с конечной дисперсией. Доказать, что

$$S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

является состоятельной оценкой параметра $\sigma^2 = DX_1$. Является ли S^2 несмещённой оценкой дисперсии σ^2 ? Построить оценку, являющуюся одновременно состоятельной и несмещённой оценкой параметра σ^2 .

10. Пусть $X_1, \dots, X_n$ — выборка из распределения с конечным вторым моментом. Пусть значение $a = EX_1$ известно. Проверить на несмещенность и состоятельность следующие оценки неизвестной дисперсии: а) $\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$; б) $\bar{X}^2 - a^2$; в) $\frac{1}{n} \sum_{i=1}^n (X_i - a)^2$; г) $\frac{1}{n-1} \sum_{i=1}^n (X_i - a)^2$.
11. Пусть θ^* — оценка параметра θ со смещением $b(\theta) = 2\theta$. Построить несмещённую оценку параметра θ .
12. Пусть $X_1, \dots, X_n$ — выборка из равномерного распределения на отрезке $[0, \theta]$. Сравнить оценки $2\bar{X}$, $X_{(n)}$ и $\frac{n+1}{n} X_{(n)}$ параметра θ в среднеквадратичном смысле.

13. Пусть $X_1, \dots, X_n$ — выборка из равномерного распределения на отрезке $[0, \theta]$. Сравнить оценки $\theta_{k,n}^* = \frac{n+k}{n} X_{(n)}$, $k = 0, 1, 2, \dots$, параметра θ в среднеквадратичном смысле.

14. Пусть $X_1, \dots, X_n$ — выборка из смещённого показательного распределения с плотностью

$$f_{\beta}(y) = \begin{cases} e^{\beta-y} & \text{при } y \geq \beta, \\ 0 & \text{при } y < \beta. \end{cases}$$

Сравнить в среднеквадратичном смысле оценки $\bar{X} - 1$, $X_{(1)}$ и $X_{(1)} - 1/n$ параметра сдвига β .

15. Пусть $X_1, \dots, X_n$ — выборка из равномерного распределения на отрезке $[0, \theta]$, где $\theta \in (0, 1]$. Используя неравенство Чебышева, построить доверительный интервал для θ с помощью оценки $2\bar{X}$.

16. Пусть $X_1, \dots, X_n$ — выборка из равномерного распределения на отрезке $[0, \theta]$. С помощью статистики $X_{(n)}$ построить точный доверительный интервал уровня $1 - \varepsilon$ для параметра θ .

17. С помощью статистики $\bar{X}$ построить асимптотический доверительный интервал уровня $1 - \varepsilon$ для неизвестного параметра следующих распределений: а) E_{α} , $\alpha > 0$; б) Π_{λ} , $\lambda > 0$; в) B_p , $0 < p < 1$; г) $B_{m,p}$, $0 < p < 1$, m — известно; д) G_p , $0 < p < 1$; е) $U[0, \theta]$, $\theta > 0$.

18. В результате проверки 400 электрических лампочек 40 штук оказалось бракованными. Найти доверительный интервал уровня 0,99 для вероятности брака.

19. Пусть $X_1, \dots, X_n$ — выборка из нормального распределения со средним a и единичной дисперсией. Для проверки основной гипотезы $a = 0$ против альтернативы $a = 1$ используется следующий критерий: основная гипотеза принимается, если $X_{(n)} < 3$, и отвергается в противном случае. Найти вероятности ошибок первого и второго рода.

20. Пусть $X_1, \dots, X_n$ — выборка из нормального распределения со средним a и единичной дисперсией. Рассматриваются две простые гипотезы: основная $a = -1$ и альтернативная $a = 0$. Предлагается следующий статистический критерий для проверки этих гипотез: основная гипотеза принимается, если $\bar{X} < -n^{\gamma}$; в противном случае принимается альтернативная гипотеза. Здесь γ — заранее выбранное вещественное число. Определить все числа γ , при которых критерий является состоятельным.

21. Есть две гипотезы: основная состоит в том, что элементы выборки имеют нормальное распределение, а альтернатива — в том, что элементы выборки имеют распределение Пуассона. Построить критерий, обладающий нулевыми вероятностями ошибок первого и второго рода.

22. Основная гипотеза состоит в том, что данный человек лишён телепатических способностей и угадывает мысли на расстоянии в каждом единичном эксперименте с вероятностью $1/2$. Гипотеза же о наличии телепатических способностей у данного человека принимается, если в 100 независимых однотипных экспериментах по угадыванию мыслей на расстоянии не менее 70 заканчиваются успехом. Чему равна вероятность признать телепатом человека без телепатических способностей?

23. При $n = 4040$ бросаниях монеты Бюффон получил 2048 выпадений герба и 1992 выпадений решётки. Совместимо ли это с гипотезой о том, что существует постоянная вероятность $p = 1/2$ выпадения герба?

24. Используя конструкции доверительного интервала, построить критерий с (точной или асимптотической) ошибкой первого рода ε для проверки гипотезы $\theta = 1$ по выборке из
а) $N_{\theta,1}$; б) $N_{1,\theta}$; в) E_{θ} ; г) $B_{\theta/2}$; д) Π_{θ} .

25. По официальным данным в Швеции в 1935 г. родилось 88 273 ребенка, причем в январе родилось 7280 детей, в феврале — 6957, марте — 7883, апреле — 7884, мае — 7892, июне — 7609, июле — 7585, августе — 7393, сентябре — 7203, октябре — 6903, ноябре — 6552, декабре — 7132 ребенка. Совместимы ли эти данные с гипотезой, что день рождения наудачу выбранного человека с равной вероятностью приходится на любой из 365 дней года?