Задачи по математической статистике

1. Построить оценки неизвестных параметров по методу моментов для следующих распределений: а) B_p, $0 < p < 1$; б) Π_λ, $\lambda > 0$; в) G_p, $0 < p < 1$; г) $U[0,\theta]$, $\theta > 0$; д) $U[\theta-1,\theta+1]$, $-\infty < \theta < \infty$; е) $U[-\theta,\theta]$, $\theta > 0$; ж) E_α, $\alpha > 0$; з) $\Phi_{\alpha,1}$, $-\infty < \alpha < \infty$; и) Φ_{0,σ^2}, $\sigma^2 > 0$.
Исследовать полученные оценки на несмещенность и состоятельность.

2. 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) E_α, $\alpha > 0$; ж) $\Phi_{\alpha,1}$, $-\infty < \alpha < \infty$; з) Φ_{0,σ^2}, $\sigma^2 > 0$.
Исследовать полученные оценки на несмещенность и состоятельность.

3. Построить оценки максимального правдоподобия параметра θ, если распределение выборки имеет плотность

а) $\theta y^{\theta-1}$ при $y \in [0,\theta]$, $\theta > 0$;

б) $\dfrac{2y}{\theta^2}$ при $y \in [0,\theta]$, $\theta > 0$;

в) $\dfrac{1}{2}e^{-|y-\theta|}$, $y \in \mathbb{R}$, $-\infty < \theta < \infty$;

г) $\dfrac{\theta e^{-\theta^2/2y}}{\sqrt{2\pi y^3}}$ при $y \geq 0$, $\theta > 0$.

4. Пусть X_1, …, X_n — выборка из смещённого показательного распределения с плотностью

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

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

5. Пусть X_1, …, X_n — выборка из равномерного распределения на отрезке $[0,\theta]$. С помощью неравенства Чебышева доказать состоятельность следующих оценок параметра $\theta > 0$: а) $2\overline{X}$; б) $X_{(n)}$.

6. Пусть X_1, …, X_n — выборка из некоторого распределения с конечной дисперсией. Доказать, что

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

является состоятельной оценкой параметра $\sigma^2 = \mathsf{D}X_1$. Является ли S^2 несмещённой оценкой дисперсии σ^2? Построить оценку, являющуюся одновременно состоятельной и несмещённой оценкой параметра σ^2.

7. Пусть X_1, …, X_n — выборка из распределения с конечным вторым моментом. Пусть значение $a = \mathsf{E}X_1$ известно. Проверить на несмещённость и состоятельность следующие оценки неизвестной дисперсии:
а) $\frac{1}{n-1}\sum\limits_{i=1}^{n}\left(X_i - \overline{X}\right)^2$; б) $\overline{X^2} - a^2$; в) $\frac{1}{n}\sum\limits_{i=1}^{n}(X_i - a)^2$; г) $\frac{1}{n-1}\sum\limits_{i=1}^{n}(X_i - a)^2$.

8. Пусть $X_1, \ldots, X_n$ — выборка из показательного распределения с параметром α. Будет ли оценка $\alpha_n^* = 1/\overline{X}$ несмещённой? Если «нет», найти смещение. Является ли оценка состоятельной?

9. Пусть θ^* — оценка параметра θ со смещением $b(\theta) = 2\theta$. Построить несмещённую оценку параметра θ.

10. Пусть θ_n^* — асимптотически несмещённая оценка для θ и $\mathsf{D}\theta_n^* \to 0$ при $n \to \infty$ для любого $\theta \in \Theta$. Доказать, что оценка θ_n^* состоятельна.

11. Пусть $X_1, \ldots, X_n$ — выборка из равномерного распределения на отрезке $[0, \theta]$. Сравнить оценки $2\overline{X}$, $X_{(n)}$ и $\frac{n+1}{n}X_{(n)}$ параметра θ в среднеквадратичном смысле.

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

13. Пусть $X_1, \ldots, X_n$ — 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 $\overline{X} - 1$, $X_{(1)}$ и $X_{(1)} - 1/n$ параметра сдвига β.

14. Пусть $X_1, \ldots, X_n$ — выборка из равномерного распределения на отрезке $[0, \theta]$, где $\theta \in (0, 1]$. Используя неравенство Чебышева, построить доверительный интервал для θ с помощью а) оценки $2\overline{X}$; б) оценки $X_{(n)}$.

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

16. С помощью оценки $X_{(1)}$ по выборке объёма n из смещённого показательного распределения с параметром сдвига β построить точный доверительный интервал для параметра β.

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

18. С помощью статистики $\overline{X}$ построить асимптотический доверительный интервал уровня $1 - \varepsilon$ для параметра λ распределения Пуассона.

19. Пусть $X_1, \ldots, X_n$ — выборка из смещённого показательного распределения с параметром сдвига β. С помощью статистики $\overline{X}$ построить асимптотический доверительный интервал для параметра β уровня $1 - \varepsilon$. Сравнить его с точным доверительным интервалом из задачи **16**. Какой из интервалов следует предпочесть?

20. Пусть $X_1, \ldots, X_n$ — 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.

21. Пусть $X_1, \ldots, X_n$ — 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 $\overline{X} < -n^{\gamma}$; в противном случае принимается альтернативная гипотеза. Здесь γ — заранее выбранное вещественное число. Определить все числа γ, при которых критерий является состоятельным.

22. 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.

23. 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?

24. 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?

25. Используя конструкции доверительного интервала, построить критерий с (точной или асимптотической) ошибкой первого рода ε для проверки гипотезы $\theta = 1$ по выборке из

а) нормального распределения со средним θ и дисперсией 1;

б) нормального распределения со средним 1 и дисперсией θ;

в) показательного распределения с параметром θ;

г) распределения Бернулли с параметром $\theta/2$;

д) распределения Пуассона с параметром θ.

26. 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?

Ответы

1. а) $\overline{X}$, несмещенная, состоятельная; б) $\overline{X}$, несмещенная, состоятельная; в) $1/\overline{X}$, смещенная, состоятельная; г) $2\overline{X}$, несмещенная, состоятельная; д) $\overline{X}$, несмещенная, состоятельная; е) $\sqrt{3\overline{X^2}}$, смещенная, состоятельная; ж) $1/\overline{X}$, смещенная, состоятельная; з) $\overline{X}$, несмещенная, состоятельная; и) $\overline{X^2}$, несмещенная, состоятельная.

2. а) $\overline{X}$, несмещенная, состоятельная; б) $\overline{X}$, несмещенная, состоятельная; в) $\overline{X}/m$, несмещенная, состоятельная; г) $X_{(n)}$, смещенная, состоятельная; д) $\max\{-X_{(1)}, X_{(n)}\} = \max\{|X_i|\}$, смещенная, состоятельная; е) $1/\overline{X}$, смещенная, состоятельная; ж) $\overline{X}$, несмещенная, состоятельная; з) $\overline{X^2}$, несмещенная, состоятельная.

3. а) $-1/\overline{\ln X}$; б) $X_{(n)}$; в) $\theta^* = \begin{cases} X_{(m)}, & \text{если } n = 2m-1 \text{ (нечётно)}, \\ \dfrac{X_{(m)} + X_{(m+1)}}{2}, & \text{если } n = 2m \text{ (чётно)}; \end{cases}$

г) $1/\sqrt{\overline{X^{-1}}}$.

4. а) $\overline{X} - 1$, несмещенная, состоятельная; б) $X_{(1)}$, смещенная, состоятельная.

6. нет; $S_0^2 = \dfrac{n}{n-1} S^2$.

7. а), б), в) несмещённая и состоятельная; г) смещённая и состоятельная.

8. нет, $\alpha/(n-1)$; да.

9. $\theta^*/3$.

11. Среднеквадратические отклонения: $\theta^2/12n$, $2\theta^2/(n+1)(n+2)$, $\theta^2/n(n+2)$.

12. $\theta^*_{1,n}$ — наилучшая; $\theta^*_{0,n}$ лучше, чем $\theta^*_{2,n}$; $\theta^*_{k,n}$ лучше, чем $\theta^*_{k+1,n}$ при $k \geq 2$.

13. $\mathsf{E}(\overline{X} - 1 - \theta)^2 = 1/n$, $\mathsf{E}(X_{(1)} - \theta)^2 = 2/n^2$, $\mathsf{E}(X_{(1)} - 1/n - \theta)^2 = 1/n^2$.

14. а) $(2\overline{X} - \sqrt{1/3n\varepsilon},\, 2\overline{X} + \sqrt{1/3n\varepsilon})$; б) $(X_{(n)},\, X_{(n)} + /(n+1)\varepsilon)$.

15. $(X_{(n)},\, X_{(n)}/\sqrt[n]{\varepsilon})$.

16. $(X_{(1)} + (\ln\varepsilon)/n,\, X_{(1)})$.

17. (0,061; 0,139).

18. $\big(\overline{X} - t_{1-\varepsilon/2}\sqrt{\overline{X}}/\sqrt{n},\ \overline{X} + t_{1-\varepsilon/2}\sqrt{\overline{X}}/\sqrt{n}\big)$, где $t_{1-\varepsilon/2}$ — квантиль уровня $1-\varepsilon/2$ распределения $\Phi_{0,1}$.

19. $\big(\overline{X} - 1 - t_{1-\varepsilon/2}/\sqrt{n},\ \overline{X} - 1 + t_{1-\varepsilon/2}/\sqrt{n}\big)$, где $t_{1-\varepsilon/2}$ — квантиль уровня $1-\varepsilon/2$ распределения $\Phi_{0,1}$. Предпочесть следует точный, так как его длина есть величина порядка $1/n$, а не $1/\sqrt{n}$.

20. $1 - \big(1 - \overline{\Phi}(3)\big)^n = 1 - 0{,}99865^n$; $\big(1 - \overline{\Phi}(2)\big)^n = 0{,}977^n$.

21. $\gamma > -1/2$.

22. Основная гипотеза отвергается, если значение хотя бы одного элемента выборки целое.

23. $\overline{\Phi}(4) = 0{,}00003167$.

24. Вероятность получить такое же или еще большее число гербов (реально достигнутый уровень значимости) при верной основной гипотезе равна 0,189.

25. Гипотеза принимается, если: а) $|\overline{X} - 1| < t_{1-\varepsilon/2}/\sqrt{n}$, где $t_{1-\varepsilon/2}$ — квантиль уровня $1 - \varepsilon/2$ распределения $\Phi_{0,1}$; б) $t_{\varepsilon/2} < n\overline{(X-1)^2} < t_{1-\varepsilon/2}$, где t_δ — квантиль уровня δ χ^2-распределения с n степенями свободы; в) $X_{(1)} < -(\ln\varepsilon)/n$; г) $|2\overline{X} - 1| < t_{1-\varepsilon/2}\sqrt{\overline{X}(2-\overline{X})}/\sqrt{n}$, где $t_{1-\varepsilon/2}$ — квантиль уровня $1 - \varepsilon/2$ распределения $\Phi_{0,1}$; д) $|\overline{X} - 1| < t_{1-\varepsilon/2}\sqrt{\overline{X}}/\sqrt{n}$, где $t_{1-\varepsilon/2}$ — квантиль уровня $1 - \varepsilon/2$ распределения $\Phi_{0,1}$.

26. нет; реально достигнутый уровень значимости равен $2{,}7 \cdot 10^{-49}$.