Задачи

Задача №1. Пусть $X_1,\dots,X_n$ независимы и имеют биномиальное распределение $Bi(1,\theta)$, $0<\theta<1$. Доказать, что $T(X)=\sum_{i=1}^n X_i$ – полная и достаточная статистика.

Решение. Покажем, что $P_\theta(X{=}x\mid T(X){=}t)$ не зависит от θ:

$$P_\theta\big(X{=}x\mid T(X){=}t\big)=\frac{P\big(X_1{=}x_1,\dots,X_n{=}x_n,\,T(X)=t\big)}{P\big(T(X){=}t\big)}=\Big\{\sum_{i=1}^n X_i\sim Bi(n,\theta)\Big\}=$$
$$=\frac{\theta^{\sum_{i=1}^n x_i}\big(1{-}\theta\big)^{n-\sum_{i=1}^n x_i}}{C_n^t\cdot\theta^t\big(1{-}\theta\big)^{n-t}}\times\mathbb{I}_{\left\{\sum_{i=1}^n x_i=t\right\}}=\frac{1}{C_n^t}\times\mathbb{I}_{\left\{\sum_{i=1}^n x_i=t\right\}}$$

Достаточность статистики $T(X)$ доказана.
Пусть функция $\varphi(\cdot)$ такова, что $\mathbb{E}_\theta\varphi\big(T(X)\big)=0\quad\forall\theta\in(0,1)$.

$$\mathbb{E}_\theta\varphi\big(T(X)\big)=\sum_{k=0}^n\varphi(k)C_n^k\theta^k(1{-}\theta)^{n-k}=\Big\{\frac{\theta}{1-\theta}{=}\tau,\ \tau\in(0,+\infty)\Big\}=$$
$$=(1{-}\theta)^n\sum_{k=0}^n\varphi(k)C_n^k\tau^k\equiv0\iff\sum_{k=0}^n\varphi(k)C_n^k\tau^k\equiv0\quad\forall\tau\in(0,+\infty).$$

$\sum\limits_{k=0}^n\varphi(k)C_n^k\tau^k$ – многочлен степени не выше n – имеет континуум корней, следовательно, все его коэффициенты равны нулю: $\varphi(k){=}0,\ k=\overline{0,n}$.
Итак, из $\mathbb{E}_\theta\varphi\big(T(X)\big){=}0\quad\forall\theta\in(0,1)$ следует $\varphi{\equiv}0$ по распределнию $T(X)$, что и означает полноту статистики $T(X)$. □

Задача №2. Пусть $X_1,\dots,X_n$ независимы и имеют равномерное распределение на отрезке $[a;b]$. Найти оценку максимального правдоподобия для a и b.

Решение. Запишем функцию правдоподобия:

$$L(X;a,b)=\prod_{k=1}^n f(X_k)=\frac{1}{(b{-}a)^n}\mathbb{I}_{\left\{a\le X_{(1)}\le X_{(n)}\le b\right\}}$$

Зафиксируем b. $L(X;a,b)\to\max$, при $a\le X_{(1)}$ и $\frac{1}{(b-a)^n}\to\max$. Значит, оценкой максимального правдоподобия для a будет $\widehat{a}=X_{(1)}$.
Аналогично, оценка максимального правдоподобия для b: $\widehat{b}=X_{(n)}$. □

Задача №3. Пусть $X_1,\dots,X_n$ независимы и имеют равномерное распределение на отрезке $[0;\theta]$. Исследовать несмещенность и состоятельность оценки $T(X){=}X_{(n)}$ параметра θ.

Решение. Поскольку функция распределения $X_{(n)}$ представляется в виде

$$F_{X_{(n)}}(x)=\begin{cases}0, & y<0,\\ \dfrac{y^n}{\theta^n}, & y\in[0,\theta],\\ 1, & y>0,\end{cases}$$

то плотность $X_{(n)}$ имеет следующий вид:

$$f_{X_{(n)}}(x)=\begin{cases}\dfrac{ny^{n-1}}{\theta^n}, & y\in[0,\theta]\\ 0, & \text{иначе.}\end{cases}$$

значит,

$$\mathbb{E}X_{(n)}=\int_0^\theta y\cdot\frac{ny^{n-1}}{\theta^n}dy=\frac{n\theta}{n{+}1}\ne\theta$$

Таким образом, $T(X)$ несмещенной оценкой для θ не является[1].
Проверим теперь состоятельность оценки. Для любого достаточно малого $\varepsilon>0$

$$P\big(X_{(n)}<\theta-\varepsilon\big)=\Big(\frac{\theta{-}\varepsilon}{\theta}\Big)^n\xrightarrow[n\to\infty]{}0,$$

следовательно,

$$P\big(X_{(n)}\in[\theta{-}\varepsilon,\theta]\big)\xrightarrow[n\to\infty]{}1.$$

То есть, $T(X)=X_{(n)}\xrightarrow{P}\theta$, что означает состоятельность оценки $T(X)$. □

Задача № 4. Пусть $X_1,\dots,X_n$ независимы и имеют пуассоновское распределение $\Pi(\theta)$, $\theta{>}0$. Доказать, что $T(X)=\sum_{i=1}^n X_i$ – достаточная и полная статистика.

Решение. Покажем, что $P\big(X{=}x\mid T(X){=}t\big)$ не зависит от θ:

$$P\big(X{=}x\mid T(X){=}t\big)=\frac{P\big(X{=}x,T(X){=}t\big)}{P\big(T(X){=}t\big)}=\Big\{T(X){=}\sum_{i=1}^n X_i\sim\Pi(n\theta)\Big\}=$$
$$=\frac{\prod\limits_{k=1}^n\exp(-\theta)\dfrac{\theta^{x_k}}{x_k!}\times\mathbb{I}_{\left\{\sum_{k=1}^n x_k=t\right\}}}{e^{-n\theta}\dfrac{(n\theta)^t}{t!}}=\frac{t!}{n^t\cdot x_1!\dots x_n!}\times\mathbb{I}_{\left\{\sum_{k=1}^n x_k=t\right\}}$$

Достаточность статистики $T(X)$ доказана. Покажем, что она является полной.

$$\mathbb{E}_\theta\varphi\big(T(X)\big)=\sum_{k=0}^\infty\varphi(k)\exp(-n\theta)\frac{(n\theta)^k}{k!}\equiv0\qquad\forall\theta\in\Theta.\tag{$*$}$$

При $\theta{=}0\quad\mathbb{E}_\theta\varphi\big(T(X)\big)=\varphi(0)$, значит, $\varphi(0)=0$. Разделив ($*$) на θ и устремив θ к нулю, получим $\varphi(1)=0$.
Повторяя эту процедуру (разделить ($*$) на θ и перейти к пределу в нуле), придем к $\varphi(k){=}0,\quad k=0,1,2,\dots,$ Таким образом, $\varphi\equiv0$ по распределению $T(X)$, что означает полноту статистики $T(X)=\sum_{i=1}^n X_i$. □

Задача №5. Пусть $X_1,\dots,X_n$ независимы и имеют равномерное распределение на отрезке $[a,b]$. Найти оценку методом моментов для a и b по первым 2 моментам.

[1] Однако $T(X)$ – асимптотически несмещенная оценка θ.

Решение. Обозначим $M_1{=}\overline{X}$, $M_2{=}\overline{X^2}$ – эмпирические моменты первого и второго порядков соответственно.
Для равномерно распределенной на $[a,b]$ случайной величины X теоретические моменты первого и второго порядков следующие:

$$\mu_1=\mathbb{E}X=\frac{a{+}b}{2},$$
$$\mu_2=\mathbb{E}X^2=\frac13\left(a^2+ab+b^2\right).$$

Из условия равенства теоретических моментов эмпирическим получим оценку методом моментов для a и b:

$$a=M_1-\sqrt{3\left(M_2{-}M_1^2\right)},$$
$$b=M_1+\sqrt{3\left(M_2{-}M_1^2\right)}.$$ □

Задача № 6. Пусть слуыайные величины $X_1,\dots,X_n$ независимы и имеют нормальное распределение $\mathcal{N}(\theta,1)$. Исследовать несмещенность и состоятельность оценки $T(X)=\overline{X}$ параметра θ.

Решение. $\mathbb{E}T(X){=}\mathbb{E}X_1{=}\theta$. $T(X)=\overline{X}$ – несмещенная оценка параметра θ. Воспользуемся неравенством Чебышева:

$$\forall\varepsilon>0\quad P\left(\big|\overline{X}-\mathbb{E}\overline{X}\big|<\varepsilon\right)>1-\frac{\mathbb{D}\overline{X}}{\varepsilon^2}\xrightarrow[n\to\infty]{}1,\tag{$*$}$$

поскольку дисперсия оценки

$$\mathbb{D}\overline{X}=\frac1n\mathbb{D}X_1=\frac1n\xrightarrow[n\to\infty]{}0.$$

($*$) означает, что

$$T(X)=\overline{X}\xrightarrow{P}\theta.$$

что, в свою очередь, означает состоятельность оценки. □

Задача №7. Пусть $X_1,\dots,X_n$ независимы и имеют равномерное распределение на отрезке $[0,\theta]$. Доказать, что $T(X){=}\max\limits_{1\le i\le n}X_i$ - достаточная и полная статистика.

Решение. Проверим достаточности статистики $T(X)$. Функция правдоподобия выборки $X{=}(X_1,\dots,X_n)$ имеет следующий вид:

$$L(X;\theta)=\prod_{k=1}^n\frac1\theta\mathbb{I}_{\left\{0\le X_k\le\theta\right\}}=\frac{1}{\theta^n}\mathbb{I}_{\left\{0\le X_{(1)}\right\}}\mathbb{I}_{\left\{X_{(n)}\le\theta\right\}}=g\big(T(X),\theta\big)\times h(X),\quad\text{где}$$
$$g\big(T(X),\theta\big)=\frac{1}{\theta^n}\mathbb{I}_{\left\{X_{(n)}\le\theta\right\}},\quad h(X)=\mathbb{I}_{\left\{0\le X_{(1)}\right\}}.$$

Выполнен критерий факторизации, значит, статистика $T(X){=}X_{(n)}$ – достаточная.

Пусть непрерывная функция $\varphi(\cdot)$ такова, что $\mathbb{E}_\theta\varphi\big(T(X)\big){\equiv}0\quad\forall\theta\in(0,+\infty)$.

$$f_{X_{(n)}}(y)=\frac{d}{dy}F_{X_{(n)}}(y)=\begin{cases}\dfrac{ny^{n-1}}{\theta^n}, & y\in[0;\theta]\\ 0, & \text{иначе.}\end{cases}$$
$$\mathbb{E}_\theta\varphi(X_{(n)})=\int_0^\theta\varphi(y)\frac{ny^{n-1}}{\theta^n}dy\equiv0\quad\Longrightarrow\quad\int_0^\theta\varphi(y)y^{n-1}dy\equiv0.$$

Дифференцируем по θ:

$$\varphi(\theta)\cdot\theta^{n-1}\equiv0,$$

откуда следует

$$\varphi(\theta)\equiv0\ \text{на}\ (0,+\infty).$$

Итак, из $\mathbb{E}_\theta\varphi\big(T(X)\big){\equiv}0$ следует $\varphi{\equiv}0$ по распределению $T(X)$, что означает полноту статистики $T(X){=}\max\limits_{1\le i\le n}X_i$. □

Задача №8. Пусть X имеет биномиальное распределение $Bi(n,\frac12)$. Найти оценку максимального правдоподобия для n.

Решение. Функция правдоподобия случайной величины X

$$L(x,\theta)=C_n^x\frac{1}{2^n}.$$

Найдем точки, в которых функция правдоподобия достигает своего максимума. Обозначим

$$a_n=C_n^x\frac{1}{2^n}.$$

Исследуем последовательность $\{a_n\}$ на монотонность.

$$\frac{a_{n+1}}{a_n}=\frac12\cdot\frac{C_{n+1}^x}{C_n^x}=\frac12\cdot\frac{(n{+}1)!\,x!\,(n{-}x)!}{k!\,(n{+}1{-}x)!\,n!}=\frac{n{+}1}{2(n{-}x{+}1)}\vee1\iff 2x{-}1\vee n.$$

Получаем, $a_1{<}\dots{<}a_{2x-2}{<}a_{2x-1}=a_{2x}{>}a_{2x+1}{>}\dots{>}a_n$. Функция правдоподобия достигает максимума в точках $2x{-}1$ и $2x$, которые и будут оценками максимального правдоподобия параметра n. □

Задача №9. Пусть $X_1,\dots,X_n$ независимы и $X_i=\begin{cases}1, & \theta,\\ 2, & \theta,\\ 3, & 1{-}2\theta.\end{cases}$ Найти одномерную достаточную статистику.

Решение. Функцию правдоподобия случайной величины X_1 можно записать, например, так[2]:

$$L_1(x;\theta)=P(X_1{=}x)=\theta^{\frac{(x-2)(x-3)}{2}}\theta^{\frac{(x-1)(x-3)}{-1}}(1{-}2\theta)^{\frac{(x-1)(x-2)}{2}}=\theta^{-\frac{x^2}{2}+\frac{3x}{2}}(1{-}2\theta)^{\frac{x^2}{2}-\frac{3x}{2}+1}.$$

[2] Разумеется, можно ее представить и в другом виде, важно лишь, чтобы в точках $1,2$ и 3 эта функция принимала значения θ,θ и $1{-}2\theta$ соответственно.

Функция правдоподобия выборки $X{=}(X_1,\dots,X_n)$:

$$L_n(X,\theta)=\theta^{-\sum_{k=1}^n\frac{X_k^2-3X_k}{2}}\times(1{-}2\theta)^{\sum_{k=1}^n\left(\frac{X_k^2-3X_k}{2}+1\right)}=\Big(\frac{1{-}2\theta}{\theta}\Big)^{\frac12\sum_{k=1}^n(X_k^2-3X_k)}\times(1{-}2\theta)^n=$$
$$=g\big(T(X),\theta\big)\times h(X),\qquad\text{где}\ T(X){=}\sum_{k=1}^n(X_k^2{-}3X_k),\quad h(X){=}1.$$

Для функции $T(X)$ выполнен критерий факторизации, значит, она и будет достаточной статистикой. □

Задача №10. Пусть $X_1,\dots,X_n$ независимы и имеют равномерное распределение на отрезке $[\theta,\theta{+}1]$. Найти несмещенную оценку максимального правдоподобия для θ.

Решение. Запишем функцию правдоподобия:

$$L(X,\theta)=\mathbb{I}_{\left\{\theta\le X_{(1)}\le X_{(n)}\le\theta+1\right\}}=\mathbb{I}_{\left\{X_{(n)}-1\le\theta\le X_{(1)}\right\}}.$$

Оценка максимального правдоподобия для θ заключена на сегменте $[X_{(n)}{-}1,X_{(1)}]$. Для $x\in[\theta,\theta{+}1]$

$$F_{X_{(1)}}(x)=1{-}P\big(X_{(1)}\ge x\big)=1-(\theta{+}1{-}x)^n,\qquad F_{X_{(n)}}(x)=P\big(X_{(n)}\le x\big)=(x{-}\theta)^n,$$
$$f_{X_{(1)}}(x)=\frac{d}{dx}F_{X_{(1)}}(x)=n(\theta{+}1{-}x)^{n-1},\qquad f_{X_{(n)}}(x)=\frac{d}{dx}F_{X_{(n)}}(x)=n(x{-}\theta)^{n-1},$$

значит, математические ожидания случайных величин $X_{(1)}$ и $X_{(n)}$

$$\mathbb{E}X_{(1)}=\int_\theta^{\theta+1}x\cdot f_{X_{(1)}}(x)\,dx=\theta{+}\frac{1}{n{+}1},\qquad\mathbb{E}X_{(n)}=\int_\theta^{\theta+1}x\cdot f_{X_{(n)}}(x)\,dx=\theta{+}\frac{n}{n{+}1},$$

Следовательно, $\mathbb{E}\left(\dfrac{X_{(1)}{+}X_{(n)}{-}1}{2}\right){=}\theta$.
Наконец, $\dfrac{X_{(1)}{+}X_{(n)}-1}{2}\in[X_{(n)}{-}1,X_{(1)}]$, поэтому функция $T(X){=}\dfrac{X_{(1)}{+}X_{(n)}{-}1}{2}$ является несмещенной оценкой максимального правдоподобия для θ. □

Задача №11. Пусть $X_1,\dots,X_n$ независимы и имеют гамма-распределение $\Gamma(\theta,2)$. Исследовать на несмещенность и состоятельность оценку $T(X)=\overline{X}$ для функции $\tau(\theta)={}^2\!/_\theta$.

Решение. $\mathbb{E}T(X)=\mathbb{E}X_1={}^2\!/_\theta=\tau(\theta)$. $T(X)$ – несмещенная оценка для $\tau(\theta)$.
Вычислим дисперсию оценки: $\mathbb{D}\overline{X}=\dfrac1n\mathbb{D}X_1=\dfrac{2}{n\theta^2}$.
Воспользовавшись неравенством Чебышева, получим

$$\forall\varepsilon>0\quad P\big(\big|\overline{X}-{}^2\!/_\theta\big|{<}\varepsilon\big)\ge1-\frac{\mathbb{D}\overline{X}}{\varepsilon^2}=1-\frac{1}{\varepsilon^2}\cdot\frac{2}{n\theta^2}\xrightarrow[n\to\infty]{}1,$$

то есть $T(X)\xrightarrow{P}{}^2\!/_\theta$, что по определению доказывает состоятельность оценки $T(X)$. □

Задача №12. Пусть $X_1,\dots,X_n$ независимы и $X_i=\begin{cases}1, & \theta_1,\\ 2, & \theta_2,\\ 3, & 1{-}\theta_1{-}\theta_2.\end{cases}$ Найти двумерную достаточную статистику.

Решение. Функцию правдоподобия случайной величины X_1:

$$L_1(x;\theta_1,\theta_2)=P(X_1{=}x)=\theta_1^{\frac{(x-2)(x-3)}{2}}\theta_2^{\frac{(x-1)(x-3)}{-1}}(1{-}\theta_1{-}\theta_2)^{\frac{(x-1)(x-2)}{2}}=$$
$$=\theta_1^{\frac{x^2}{2}-\frac{5x}{2}+3}\theta_2^{-x^2+4x+3}(1{-}\theta_1{-}\theta_2)^{\frac{x^2}{2}-\frac{3x}{2}+1}.$$

Функция правдоподобия выборки $X{=}(X_1,\dots,X_n)$:

$$L_n(X;\theta_1,\theta_2)=\theta_1^{\sum_{k=1}^n\frac{X_k^2}{2}-\frac{5X_k}{2}+3}\theta_2^{\sum_{k=1}^n-X_k^2+4X_k+3}(1{-}\theta_1{-}\theta_2)^{\sum_{k=1}^n\frac{X_k^2}{2}-\frac{3X_k}{2}+1}=$$
$$=g\big(T(X),\theta_1,\theta_2\big)\times h(X),$$

где $T(X){=}\big(\sum_{k=1}^n X_k^2,\ \sum_{k=1}^n X_k\big)$, $\quad h(X){=}1$.
Для функции $T(X)$ выполнен критерий факторизации, значит, она и будет достаточной статистикой. □

Задача №13. Пусть $X_1,\dots,X_n$ независимы и имеют гамма-распределение $\Gamma(\theta,\lambda)$, λ - известно. Найти оценку максимального правдоподобия для θ.

Решение.

$$L(X,\theta)=\prod_{k=1}^n\frac{\theta^\lambda}{\Gamma(\lambda)}\exp(-\theta X_k)\cdot X_k^{\lambda-1}=\frac{\theta^{n\lambda}}{(\Gamma(\lambda))^n}\exp\Big(-\theta\sum_{k=1}^n X_k\Big)\cdot\Big(\prod_{k=1}^n X_k\Big)^{\lambda-1},$$
$$\frac{\partial}{\partial\theta}\ln L(X,\theta)=\frac{\partial}{\partial\theta}\Big(n\lambda\ln\theta-n\ln\Gamma(\lambda)-\theta\sum_{k=1}^n X_k+(\lambda{-}1)\sum_{k=1}^n\ln X_k\Big)=\frac{n\lambda}{\theta}-\sum_{k=1}^n X_k.$$

В точке экстремума функции правдоподобия

$$\frac{\partial}{\partial\theta}\ln L(X,\widehat{\theta})=0\iff\widehat{\theta}=\frac{\lambda}{\overline{X}}.$$

В точке $\widehat{\theta}$ достигается максимум, поскольку

$$\frac{\partial^2}{\partial\theta^2}\ln L(X,\theta)=-\frac{n\lambda}{\theta^2}<0.$$

Значит, $\widehat{\theta}={}^\lambda\!/_{\overline{X}}$ – оценка максимального правдоподобия для θ. □

Задача №14. Пусть $X_1,\dots,X_n$ независимы и имеют нормальное распределение $\mathcal{N}(0,\theta^2)$. Доказать, что $T(X)=\dfrac1n\sum\limits_{i=1}^n X_i^2$ – эффективная оценка функции $\tau(\theta){=}\theta^2$.

Решение. В неравенстве Рао – Крамера (22) равенство достигается, если и только если найдется фукция $a_n(\theta)$ такая, что

$$T(X)-\tau(\theta)=a_n(\theta)\times\frac{\partial}{\partial\theta}\ln L(X,\theta).\tag{$*$}$$

Проверим это условие.

$$L(X,\theta)=\prod_{k=1}^n\frac{1}{\sqrt{2\pi}\cdot\theta}\exp\Big(-\frac{X_k^2}{2\theta^2}\Big)=\frac{1}{(\sqrt{2\pi}\cdot\theta)^n}\exp\Big(-\frac{1}{2\theta^2}\sum_{k=1}^n X_k^2\Big)$$
$$\frac{\partial}{\partial\theta}\ln L(X,\theta)=-\frac n\theta+\frac{1}{\theta^3}\sum_{k=1}^n X_k^2=\frac{n}{\theta^3}\Big(\sum_{k=1}^n X_k^2-\theta^2\Big).$$

Полагая $a_n(\theta)={}^{\theta^3}\!/_n$, получим равенство ($*$). Эффективность оценки доказана. □

Задача №15. Пусть $X_1,\dots,X_n$ независимы и имеют гамма-распределение $\Gamma(\theta,\lambda)$. Найти оценку методом моментов для θ и λ по первым двум моментам.

Решение. Аналогично задаче №5, из системы уравнений

$$\begin{cases}\overline{X}=M_1=\mu_1=\mathbb{E}X=\dfrac\lambda\theta,\\ \overline{X^2}=M_2=\mu_2=\mathbb{E}X^2=\dfrac{\lambda(\lambda+1)}{\theta^2};\end{cases}$$

получим оценку методом моментов для θ и λ:

$$\theta=\frac{M_1}{M_2-M_1^2},\qquad\lambda=\frac{M_1^2}{M_2-M_1^2}.$$ □

Задача №16. Пусть $X_1,\dots,X_n$ независимы и имеют биномиальное распределение $Bi(1,\theta)$. Требуется исследовать несмещенность и состоятельность оценки $T(X){=}\overline{X}$ для параметра θ.

Решение. $\mathbb{E}T(X){=}\mathbb{E}X_1=\theta$ – оценка является несмещенной.
Вычислим дисперсию оценки: $\mathbb{D}\overline{X}{=}\dfrac1n\mathbb{D}X_1{=}\dfrac{\theta(1{-}\theta)}{n}$.
Воспользовавшись неравенством Чебышева, получим

$$\forall\varepsilon>0\qquad P\big(\big|\overline{X}-\theta\big|{<}\varepsilon\big)>1-\frac{\mathbb{D}\overline{X}}{\varepsilon^2}=1-\frac{\theta(1{-}\theta)}{n\varepsilon}\xrightarrow[n\to\infty]{}1,$$

то есть $T(X){=}\overline{X}\xrightarrow{P}\theta$, что по определению доказывает состоятельность оценки $T(X)$. □

Задача №17. Пусть $X_1,\dots,X_n$ независимы и имеют нормальное распределение $\mathcal{N}(\theta,1)$. Доказать, что $\overline{X}{-}\dfrac1n$ является оптимальной оценкой функции $\tau(\theta){=}\theta^2$.

Решение. Функция правдоподобия

$$L(X,\theta)=\prod_{i=1}^n\frac{1}{\sqrt{2\pi}}\exp\Big(-\frac{(X_i{-}\theta)^2}{2}\Big)=K(\theta)\exp\big(n\theta\overline{X}\big)h(X),$$
$$\text{где}\quad K(\theta){=}\exp\Big(-\frac{n\theta^2}{2}\Big),\quad h(X){=}\frac{1}{\big(\sqrt{2\pi}\big)^n}\exp\Big(-\frac12\sum_{i=1}^n X_i^2\Big).$$

По теореме о полноте экспоненциальных семейств $T(X){=}\overline{X}$ – полная достаточная статистика. Любая измеримая функция от полной достаточной статистики $\overline{X}$ является оптимальной оценкой своего математического ожидания.
В частности, $\overline{X}^2{-}{}^1\!/_n$ – оптимальная оценка для

$$\mathbb{E}\Big(\overline{X}^2{-}\frac1n\Big)=\mathbb{E}\overline{X}^2{-}\frac1n=\mathbb{D}\overline{X}{+}\big(\mathbb{E}\overline{X}\big)^2{-}\frac1n=\frac1n\mathbb{D}X_1+\theta^2-\frac1n=\theta^2,$$

что и требовалось доказать. □

Задача №18. Пусть $X_1,\dots,X_n$ независимы и распределены с плотностью

$$f(x,\theta)=\begin{cases}\exp\{-(x{-}\theta)\}, & x>\theta,\\ 0, & x\le\theta.\end{cases}$$

Найти оценку максимального правдоподобия для θ.

Решение. Функция правдоподобия:

$$L(X,\theta)=\prod_{i=1}^n\Big(\exp\big(-(X_i-\theta)\big)\,\mathbb{I}_{\left\{X_i\ge\theta\right\}}\Big)=\exp\Big(n\theta-\sum_{i=1}^n X_i\Big)\mathbb{I}_{\left\{X_{(1)}\ge\theta\right\}}$$

Данная функция достигает максимума в точке $\widehat{\theta}{=}X_{(1)}$, которая и будет оценкой максимального правдоподобия для θ. □

Задача №19. Пусть $X_1,\dots,X_n$ независимы и имеют гамма-распределение $\Gamma\big({}^1\!/_\theta,1\big)$. Доказать, что $T(X){=}\dfrac1n\sum\limits_{i=1}^n X_i$ является эффективной оценкой θ.

Решение. Для доказательства эффективности оценки снова воспользуемся критерием эффективности, а именно, покажем, что

$$T(X)-\tau(\theta)=a_n(\theta)\frac{\partial}{\partial\theta}\ln L(X,\theta).\tag{$*$}$$

Функция правдоподобия

$$L(X,\theta)=\prod_{i=1}^n\frac1\theta\exp\Big(-\frac{X_i}{\theta}\Big),$$
$$\frac{\partial}{\partial\theta}\ln L(X,\theta)=-\frac n\theta+\frac{\sum_{i=1}^n X_i}{\theta}=\frac{n}{\theta^2}\big(\overline{X}-\theta\big).$$

Полагая $a_n(\theta){=}{}^{\theta^2}\!/_n$, получим равенство ($*$), значит, $T(X)$ – эффективная оценка θ. □

Задача №20. Пусть $X_1,\dots,X_n$ независимы и имеют нормальное распределение $\mathcal{N}(\theta,2\theta)$. Найти оценку максимального правдоподобия для θ.

Решение. Функция правдоподобия

$$L(X;\theta)=\prod_{i=1}^n\frac{1}{\sqrt{4\pi\theta}}\exp\Big\{-\frac{(X_i-\theta)^2}{4\theta}\Big\};$$
$$\frac{\partial}{\partial\theta}\ln L(X,\theta)=-\frac{n}{2\theta}+\frac{1}{4\theta^2}\sum_{i=1}^n X_i^2-\frac n4.$$

В точке экстремума функции правдоподобия

$$\frac{\partial}{\partial\theta}\ln L(X,\theta)=0\iff\theta^2+2\theta-\overline{X^2}=0.$$

В силу неотрицательности дисперсии $\theta>0$, значит, возможная точка максимума

$$\widehat{\theta}=-1+\sqrt{1{+}\overline{X^2}}.$$

В этой точке действительно достигается максимум функции правдоподобия, поскольку

$$\frac{\partial^2}{\partial\theta^2}\ln L(X,\theta)=\frac{n}{2\theta^3}\big(\theta-\overline{X^2}\big)\Big|_{\theta=\widehat{\theta}}=\frac{n}{2\theta^3}\Big(\sqrt{1{+}\overline{X^2}}-\big(1{+}\overline{X^2}\big)\Big)<0.$$ □

Задача №21. Пусть случайные величины $X_1,\dots,X_n$ независимы и имеют гамма - распределение $\Gamma\big({}^1\!/_\theta,1\big)$. Доказать, что $T(X){=}\dfrac{n}{n{+}1}\overline{X}^2$ является оптимальной оценкой θ^2.

Решение. Функция правдоподобия

$$L(X;\theta)=\frac{1}{\theta^n}\exp\Big(-\frac n\theta\overline{X}\Big).$$

По теореме о полноте экспоненциальных семейств $\overline{X}$ – полная и достаточная статистика, поэтому любая измеримая функция от $\overline{X}$ является оптимальной оценкой своего математического ожидания. В частности, $T(X)=\dfrac{n}{n+1}\overline{X}^2$ – оптимальная оценка для

$$\mathbb{E}T(X)=\frac{n}{n{+}1}\mathbb{E}\overline{X}^2=\frac{n}{n{+}1}\Big(\mathbb{D}\overline{X}{+}\big(\mathbb{E}\overline{X}\big)^2\Big)=\frac{n}{n{+}1}\Big(\frac1n\mathbb{D}X_1{+}\big(\mathbb{E}X_1\big)^2\Big)=$$
$$=\frac{n}{n{+}1}\Big(\frac{\theta^2}{n}+\theta^2\Big)=\theta^2=\tau(\theta),$$

что и требовалось доказать. □

Задача №22. Пусть случайные величины $X_1,\dots,X_n$ независимы и распределены с плотностью $\exp\big\{-(x{-}\theta){-}\exp[-(x{-}\theta)]\big\}$. Найти оценку максимального правдоподобия для θ.

Решение. Функция правдоподобия:

$$L(X,\theta)=\exp\Big\{n\theta-\sum_{i=1}^n X_i-\sum_{i=1}^n\exp(\theta-X_i)\Big\}.$$

В точках экстремума

$$\frac{\partial}{\partial\theta}\ln L(X,\theta)=n-\sum_{i=1}^n\exp(\theta-X_i)=0,$$

откуда

$$\widehat{\theta}=\ln n-\ln\sum_{i=1}^n\exp(-X_i).$$

Полученная оценка действительно является оценкой максимального правдоподобия, так как

$$\frac{\partial^2}{\partial\theta^2}\ln L(X,\theta) = -\sum_{i=1}^{n}\exp(\theta - X_i) < 0. \qquad \square$$

Задача №23. Пусть случайные величины $X_1,\ldots,X_n$ независимы и имеют пуассоновское распределение $\Pi(\theta)$, $\theta > 0$. Исследовать несмещенность и состоятельность оценки $T(X){=}\overline{X}$ параметра θ.

Решение. $\mathbb{E}T(X) = \mathbb{E}X_1 = \theta$. Несмещенность оценки доказана.

Дисперсия оценки $\mathbb{D}T(X) = \dfrac{1}{n}\mathbb{D}X_1 = \dfrac{\theta}{n}\xrightarrow[n\to\infty]{} 0.$

Воспользовавшись неравенством Чебышева, получим

$$\forall\varepsilon > 0 \quad P\big(\big|\overline{X} - \mathbb{E}\overline{X}\big|{<}\varepsilon\big) > 1 - \frac{\mathbb{D}\overline{X}}{\varepsilon^2}\xrightarrow[n\to\infty]{} 1,$$

то есть $T(X)\xrightarrow{P}\theta$, что по определению доказывает состоятельность оценки $T(X)$. $\square$

Задача №24. Пусть случайные величины $X_1,\ldots,X_n$ независимы и имеют биномиальное распределение $Bi(1,\theta)$, $0{<}\theta{<}1$. Доказать, что $T(X) = \dfrac{\overline{X}(1{-}\overline{X})\cdot n}{n-1}$ является оптимальной оценкой $\tau(\theta) = \theta(1{-}\theta)$.

Решение. Функция правдоподобия выглядит следующим образом:

$$L(X,\theta) = \theta^{\sum\limits_{i=1}^{n}X_i}(1{-}\theta)^{n-\sum\limits_{i=1}^{n}X_i} = (1{-}\theta)^n\exp\left\{n\ln\frac{\theta}{1{-}\theta}\overline{X}\right\}$$

По теореме о полноте экспоненциальных семейств $\overline{X}$ является полной и достаточной статистикой. Значит $\varphi(\overline{X})$ является оптимальной оценкой для $\mathbb{E}\varphi(\overline{X})$, где $\varphi(\cdot)$ – любая измеримая функция от $\overline{X}$.

$$\mathbb{E}T(X) = \frac{n}{n{-}1}\left(\mathbb{E}\overline{X} - \mathbb{E}\overline{X}^2\right) = \frac{n}{n{-}1}\left(\theta - \left(\mathbb{D}\overline{X} + \left(\mathbb{E}\overline{X}\right)^2\right)\right) =$$
$$= \frac{n}{n{-}1}\left(\theta - \frac{1}{n}\theta(1{-}\theta) - \theta^2\right) = \theta(1{-}\theta),$$

поэтому $T(X)$ является оптимальной оценкой для $\tau(\theta) = \theta(1{-}\theta)$. $\square$

Задача №25. Пусть случайные величины $X_1,\ldots,X_n$ независимы и имеют гамма-распределение $\Gamma(\theta,2)$. Исследовать на несмещенность и состоятельность оценку $T(X) = \overline{X}$ для функции $\tau(\theta) = {}^2\!/_\theta$.

Решение. $\mathbb{E}T(X) = \mathbb{E}X_1 = {}^2\!/_\theta$. Несмещенность доказана.

Вычислим дисперсию оценки: $\mathbb{D}\overline{X} = \dfrac{1}{n}\mathbb{D}X_1 = {}^2\!/_{n\theta}\xrightarrow[n\to\infty]{} 0.$

Воспользовавшись неравенством Чебышева, получим

$$\forall\varepsilon > 0 \quad P\big(\big|\overline{X} - \mathbb{E}\overline{X}\big|{<}\varepsilon\big) > 1 - \frac{\mathbb{D}\overline{X}}{\varepsilon^2}\xrightarrow[n\to\infty]{} 1,$$

то есть $T(X)\xrightarrow{P}\theta$, что по определению доказывает состоятельность оценки $T(X)$. $\square$

Задача №26. Пусть $X_1,\ldots,X_n$ независимы и имеют биномиальное распределение $Bi(1,\theta)$. Доказать, что не существует оптимальной оценки для $\tau(\theta){=}\theta^{n+1}$.

Решение. Достаточно доказать, что не существует несмещенной оценки для $\tau(\theta)$. Предположим обратное: пусть найдется функция $T(X)$ такая, что

$$\mathbb{E}T(X) = \theta^{n+1}. \qquad (*)$$

Однако

$$\mathbb{E}T(X_1,\ldots,X_n) =$$
$$= T(0,\ldots,0)\cdot(1{-}\theta)^n + T(1,0,\ldots,0)\cdot\theta(1{-}\theta)^{n-1} + \ldots + T(1,\ldots,1)\cdot\theta^n$$

– многочлен степени не выше n, поэтому равенство $(*)$ невозможно.
Таким образом, не существует оптимальной оценки для $\tau(\theta){=}\theta^{n+1}$. $\square$

Задача №27. Пусть $X_1,\ldots,X_n$ независимы и имеют биномиальное распределение $Bi(1,\theta)$. Доказать, что не существует оптимальной оценки для $\tau(\theta) = \theta^{-1}$.

Решение. Решение аналогично решению задачи №26. $\square$

Задача №28. Пусть $X_1,\ldots,X_n$ независимы и имеют пуассоновское распределение $\Pi(\theta)$. Доказать, что не существует оптимальной оценки для θ^{-2}.

Решение. Как и в задаче №26, доказав отсутствие несмещенной оценки, мы докажем отстутствие оптимальной.
Пусть существует функция $T(X) = T(X_1,\ldots,X_n)$ такая, что

$$\mathbb{E}T(X) = \theta^{-2} \qquad (*)$$

Её математическое ожидание

$$\mathbb{E}T(X_1,\ldots,X_n) = \sum_{\substack{(i_1,\ldots,i_n)\\ i_k\geq 0,\ k=0..n}} T(i_1,\ldots,i_n)\exp(-n\theta)\frac{\theta^{i_1+\ldots+i_n}}{i_1!\cdot\ldots\cdot i_n!} = \sum_{i=0}^{\infty}\exp(-n\theta)a_i\theta^i.$$

В силу несмещенности оценки

$$\mathbb{E}T(X) = \sum_{i=0}^{\infty}\exp(-n\theta)a_i\theta^i = \theta^{-2}.$$

Устремим θ к нулю:

$$\lim_{\theta\to+0}\mathbb{E}T(X) = a_0, \qquad \lim_{\theta\to+0}\theta^{-2} = +\infty,$$

Значит, равенство $(*)$ не выполнено ни для какой функции $T(X)$. $\square$

Задача №29. Пусть $X_1,\ldots,X_n$ независимы и имеют равномерное распределение на отрезке $[a;b]$. Доказать, что $T(X){=}\dfrac{X_{(1)}{+}X_{(n)}}{2}$ является несмещенной и состоятельной оценкой функции $\tau(a,b) = \dfrac{a+b}{2}$.

Решение. Для доказательства несмещенности необходимо, чтобы математическое ожидание оценки было равно оцениваемой величине. Для вычисления матожидания необходимо знать плотность случайной величины, которую можно посчитать из функции распределения.

$$F_{X_{(1)}}(y) = 1 - \left(1 - \frac{y-a}{b-a}\right)^n, \qquad F_{X_{(n)}}(y) = \left(\frac{y-a}{b-a}\right)^n,$$
$$f_{X_{(1)}}(y) = \frac{n(y-a)^{n-1}}{(b-a)^n}, \qquad f_{X_{(n)}}(y) = \frac{n(y-a)^{n-1}}{(b-a)^n},$$
$$\mathbb{E}X_{(1)} = \int_a^b y\cdot f_{X_{(1)}}y\,dy = \frac{na}{n{+}1} + \frac{b}{n{+}1}, \qquad \mathbb{E}X_{(n)} = \int_a^b y\cdot f_{X_{(n)}}(y)\,dy = \frac{nb}{n{+}1} + \frac{a}{n{+}1}.$$

Отсюда

$$\mathbb{E}T(X) = \frac{1}{2}(\mathbb{E}X_{(1)}{+}\mathbb{E}X_{(n)}) = \frac{a{+}b}{2}.$$

Несмещенность оценки доказана.
Для произвольно малого $\varepsilon > 0$

$$P\big(X_{(n)} > b{-}\varepsilon\big) = 1 - P\big(X_{(n)} \leq b{-}\varepsilon\big) = 1 - \left(\frac{b{-}\varepsilon{-}a}{b{-}a}\right)^n\xrightarrow[n\to\infty]{} 1, \quad \text{то есть } X_{(n)}\xrightarrow{P} b.$$

Аналогично,

$$X_{(1)}\xrightarrow{P} a.$$

Значит,

$$T(X){=}\frac{X_{(1)}{+}X_{(n)}}{2}\xrightarrow{P}\frac{a{+}b}{2},$$

что и означает состоятельность оценки. $\square$

Задача №30. Пусть $X_1,\ldots,X_n$ независимы и имеют равномерное распределение на отрезке $[a;b]$. Доказать, что $T(X){=}\dfrac{n{+}1}{n{-}1}\cdot\big(X_{(n)}{-}X_{(1)}\big)$ является несмещенной и состоятельной оценкой функции $\tau(a,b){=}b{-}a$.

Решение. Из предыдущей задачи:

$$\mathbb{E}X_{(1)} = \frac{na}{n{+}1} + \frac{b}{n{+}1}, \qquad \mathbb{E}X_{(n)} = \frac{nb}{n{+}1} + \frac{a}{n{+}1},$$

следовательно,

$$\mathbb{E}T(X) = b - a.$$

Несмещенность доказана.

$$X_{(1)}\xrightarrow{P} a, \qquad X_{(n)}\xrightarrow{P} b,$$

откуда

$$T(X){=}\frac{n{+}1}{n{-}1}\cdot\big(X_{(n)}{-}X_{(1)}\big)\xrightarrow{P} b{-}a.$$

Состоятельность доказана. $\square$

Задачи

Задача №31. Пусть $X_1,\ldots,X_n$ независимы и имеют равномерное распределение на отрезке $[0;\theta]$. Построить кратчайший доверительный интервал для θ с коэффициентом доверия α, основанный на центральной статистике $G(X,\theta) = \dfrac{\max\limits_{1\leq i\leq n}X_i}{\theta}$.

Решение. Разрешив неравенство $g_1 < G(X,\theta) < g_2$ относительно θ, получим доверительный интервал (θ_1,θ_2):

$$g_1 < \frac{X_{(n)}}{\theta} < g_2 \iff \frac{X_{(n)}}{g_2} < \theta < \frac{X_{(n)}}{g_1}: \quad \theta_1 = \frac{X_{(n)}}{g_2},\ \theta_2 = \frac{X_{(n)}}{g_1}.$$

Длину его $X_{(n)}{\times}\Big(\dfrac{1}{g_1} - \dfrac{1}{g_2}\Big)$ нужно минимизировать при заданном уровне доверия

$$\alpha = \mathsf{P}(\theta_1 < \theta < \theta_2) = \mathsf{P}\Big(g_1 < \frac{X_{(n)}}{\theta} < g_2\Big) = \mathsf{P}\Big(\frac{X_{(n)}}{\theta} < g_2\Big) - \mathsf{P}\Big(\frac{X_{(n)}}{\theta} < g_1\Big) = g_2^n - g_1^n,$$
$$0 \leq g_1 < g_2 \leq 1.$$

Поскольку длина интервала должна быть минимальной, иные значения g_1 и g_2 рассматривать не имеет смысла.

Минимум выражения $X_{(n)}{\times}\Big(\dfrac{1}{g_1} - \dfrac{1}{g_2}\Big)$ при условии $g_2^n - g_1^n = \alpha,\ 0 \leq g_1 < g_2 \leq 1$ достигаетя на $g_2 = 1,\ g_1 = \sqrt[n]{1-\alpha}$.

Итак, кратчайший доверительный интервал, основанный на центральной статистике $G(X,\theta)$, имеет вид $\left(X_{(n)}, \dfrac{X_{(n)}}{\sqrt[n]{1{-}\alpha}}\right)$. $\square$

Задача №32. Пусть $X_1,\ldots,X_n$ независимы и имеют биномиальное распределение $Bi(1,\theta)$, $0{<}\theta{<}1$. Построить равномерено наиболее мощный критерий размера α для проверки гипотезы $H_0:\theta = \theta_0$ при альтернативе $H_1:\theta < \theta_0$. Найти функцию мощности.

Решение. Построим наиболее мощный критерий для проверки H_0 при простой альтернативе $H_1:\theta = \theta_1$, $\theta_1 < \theta_0$, используя лемму Неймана–Пирсона:

$$\frac{L_1}{L_0} = \frac{\theta_1^{\sum\limits_{i=1}^{n}X_i}(1{-}\theta_1)^{n-\sum\limits_{i=1}^{n}X_i}}{\theta_0^{\sum\limits_{i=1}^{n}X_i(1-\theta_0)^{n-\sum\limits_{i=1}^{n}X_i}}} = \left(\frac{1{-}\theta_1}{1{-}\theta_0}\right)^n\left(\frac{\frac{\theta_1}{1-\theta_1}}{\frac{\theta_0}{1-\theta_0}}\right)^{\sum\limits_{i=1}^{n}X_i} \geq c_\alpha \iff T(X) = \sum_{i=1}^{n}X_i \leq c'_\alpha.$$

Существует c'_α для которого выполняется следующее неравенство:

$$\alpha''{=}\sum_{i=0}^{c'_\alpha-1}C_n^i\theta_0^i(1{-}\theta_0)^{n-i} < \alpha \leq \sum_{i=0}^{c'_\alpha-1}C_n^i\theta_0^i(1{-}\theta_0)^{n-i}{=}\alpha'.$$

При $\alpha = \alpha'$ критическая функцмя имеет вид:

$$\varphi(X) = \begin{cases}1, & T(X) \leq c'_\alpha\\ 0, & T(X) > c'_\alpha\end{cases}.$$

В случае $\alpha < \alpha'$, критерий является рандомизированным и из

$$\alpha = \mathbb{E}_{\theta_0}\varphi(X) = \mathsf{P}_{\theta_0}\big(T(X){<}c'_\alpha\big) + \varepsilon_\alpha\mathsf{P}_{\theta_0}\big(T(X){=}c'_\alpha\big) \implies \varepsilon_\alpha = \frac{\alpha{-}\alpha''}{\alpha'{-}\alpha''}$$

получаем

$$\varphi(X) = \begin{cases}1, & T(X) < c'_\alpha,\\ \varepsilon_\alpha, & T(X) = c'_\alpha,\\ 0, & T(X) > c'_\alpha;\end{cases}$$

где

$$\varepsilon_\alpha = \frac{\alpha - \sum_{i=0}^{c'_\alpha-1}C_n^i\theta_0^i(1{-}\theta_0)^{n-i}}{C_n^{c'_\alpha}\theta_0^{c'_\alpha}(1{-}\theta_0)^{n-c'_\alpha}}.$$

Функция мощности

$$W(\theta) = \mathbb{E}_\theta\varphi(X) = \mathsf{P}_\theta\big(T(X){<}c'_\alpha\big){+}\varepsilon_\alpha\mathsf{P}_\theta\big(T(X){=}c'_\alpha\big) = \mathsf{P}_\theta\big(T(X){<}c'_\alpha\big){+}\varepsilon_\alpha\theta^{c'_\alpha}(1{-}\theta)^{n-c'_\alpha} =$$
$$= \sum_{i=0}^{c'_\alpha-1}C_n^i\theta^i(1{-}\theta)^{n-i} + \left(\alpha - \sum_{i=0}^{c'_\alpha-1}C_n^i\theta_0^i(1{-}\theta_0)^{n-i}\right)\frac{\theta^{c'_\alpha}(1{-}\theta)^{n-c'_\alpha}}{\theta_0^{c'_\alpha}(1{-}\theta_0)^{n-c'_\alpha}}.$$

Построенный критерий – наиболее мощный, если гипотеза H_1 – простая (то есть $H_1:\theta{=}\theta_1$). При построении критерия значение θ_1 используется неявно: важно лишь, что $\theta_1 < \theta_0$. Значит, построенный критерий – равномерно наиболее мощный. $\square$

Задача №33. Пусть $X_1,\ldots,X_n$ независимы и имеют нормальное распределение $\mathcal{N}(\theta,1)$. Построить кратчайший доверительный интервал для θ с коэффициентом доверия α, основанный на центральной статистике $\sqrt{n}(\overline{X}{-}\theta)$.

Решение.

$$G(X,\theta) = \sqrt{n}\left(\overline{X} - \theta\right) = \frac{\sum_{i=1}^{n}X_i{-}n\theta}{\sqrt{n}} \sim \mathcal{N}(0,1).$$
$$\mathsf{P}\big(\tau_1 < G(X,\theta) < \tau_2\big) = \mathsf{P}\Big(\overline{X} - \frac{\tau_1}{\sqrt{n}} < \theta < \overline{X} - \frac{\tau_2}{\sqrt{n}}\Big).$$

Доверительный интервал имеет вид (θ_1,θ_2), где

$$\theta_1 = \overline{X} - \frac{\tau_1}{\sqrt{n}}, \qquad \theta_2 = \overline{X} - \frac{\tau_2}{\sqrt{n}}.$$

Длину его $\theta_2{-}\theta_1 = \dfrac{\tau_2{-}\tau_1}{\sqrt{n}}$ нужно минимизировать при условии

$$\alpha = \mathsf{P}\big(\tau_1 < G(X,\theta) < \tau_2\big) = \Phi(\tau_2) - \Phi(\tau_1),$$

где $\Phi(x)$ – функция распределения стандартного нормального закона.

Применив метод множителей Лагранжа, получим $\tau_2 = -\tau_1 = \tau_{\frac{1+\alpha}{2}}$, где $\tau_{\frac{1+\alpha}{2}}$ – квантиль порядка $\frac{1+\alpha}{2}$ функции $\Phi(x)$.

Итак, кратчайший доверительный интервал, основанный на центральной статистике $G(X,\theta)$, имеет вид $\left(\overline{X} - \dfrac{\tau_{\frac{1+\alpha}{2}}}{\sqrt{n}}, \overline{X} + \dfrac{\tau_{\frac{1+\alpha}{2}}}{\sqrt{n}}\right)$. $\square$

Задача №34. Пусть $X_1,\ldots,X_n$ независимы и имеют равномерное распределение на отрезке $[0;\theta]$. Построить наиболее мощный критерий размера α для проверки гипотезы $H_0:\theta = \theta_0$ при альтернативе $H_1:\theta = \theta_1 > \theta_0$. Найти мощность критерия.

Решение. Воспользуемся леммой Неймана-Пирсона:

$$\frac{L_1}{L_0} = \frac{\prod\limits_{i=1}^{n}\frac{1}{\theta_1}I_{\{0\leqslant X_i\leqslant\theta_1\}}}{\prod\limits_{i=1}^{n}\frac{1}{\theta_0}I_{\{0\leqslant X_i\leqslant\theta_0\}}} = \left(\frac{\theta_0}{\theta_1}\right)^n\frac{I_{\{X_{(n)}\leqslant\theta_1\}}}{I_{\{X_{(n)}\leqslant\theta_0\}}} = \begin{cases}\infty, & X_{(n)} > \theta_0\\ \left(\frac{\theta_0}{\theta_1}\right)^n, & X_{(n)} \leq \theta_0\end{cases}$$

Критическая функция $\varphi(X) = \begin{cases}1, & X_{(n)} > \theta_0,\\ \varepsilon_\alpha, & X_{(n)} \leq \theta_0.\end{cases}$

$\alpha = \mathbb{E}_{\theta_0}\varphi(X) = 1\cdot\mathsf{P}_{\theta_0}(X_{(n)} > \theta_0) + \varepsilon_\alpha\cdot\mathsf{P}_{\theta_0}(X_{(n)} \leq \theta_0) = 1\cdot 0 + \varepsilon_\alpha\cdot 1$, откуда $\alpha = \varepsilon_\alpha$.
Мощность критерия

$$W(\varphi,\theta_1) = \mathbb{E}_{\theta_1}\varphi(X) = 1\cdot\mathsf{P}_{\theta_1}\big(X_{(n)} > \theta_0\big) + \alpha\cdot\mathsf{P}_{\theta_1}\big(X_{(n)} \leq \theta_0\big) =$$
$$= 1 - \mathsf{P}_{\theta_1}\big(X_{(n)} \leq \theta_0\big) + \alpha\cdot\mathsf{P}_{\theta_1}\big(X_{(n)} \leq \theta_0\big) =$$
$$= 1 - (1{-}\alpha)\cdot\prod_{k=1}^{n}\mathsf{P}_{\theta_1}\big(X_k \leq \theta_0\big) = 1 - (1{-}\alpha)\Big(\frac{\theta_0}{\theta_1}\Big)^n. \qquad \square$$

Задача №35. Пусть $X_1,\ldots,X_n$ независимы и имеют пуассоновское распределение $\Pi(\theta)$. Построить центральный доверительный интервал с коэффициентом доверия α, используя точечную оценку $T(X) = \overline{X}$.

Решение.

$$X_i \sim \Pi(\theta), \quad n\overline{X} \sim \Pi(n\theta), \quad \mathsf{P}_\theta\Big(\overline{X} = \frac{k}{n}\Big) = \mathsf{P}_\theta\big(x\overline{X} = k\big) = \exp(-n\theta)\frac{(n\theta)^k}{k!}.$$

Функция распределения $T(X)$

$$F_{T(X)}(t;\theta) = \begin{cases}\sum\limits_{k=0}^{[nt]}\exp(-n\theta)\dfrac{(n\theta)^n}{k!}, & t \geq 0,\\ 0, & t < 0;\end{cases}$$

монотонна по θ при $\theta \geq 0$, поскольку

$$\frac{\partial}{\partial\theta}F_{T(X)}(t;\theta) = -n\exp(-n\theta)\frac{(n\theta)^{[nt]}}{[nt]!} < 0.$$

Границы доверительного интервала (θ_1,θ_2) однозначно задаются уравнениями

$$F_{T(X)}(t;\theta_1) = \frac{1+\alpha}{2}, \qquad F_{T(X)}(t;\theta_2) = \frac{1-\alpha}{2}$$

при $t = T(X)$.

$\square$

Задачи

Задача №36. Пусть $X_1,\ldots,X_n$ независимы и имеют гамма-распределение $\Gamma(\theta,1)$. Построить равномерено наиболее мощный критерий размера α для проверки гипотезы $H_0:\theta{=}\theta_0$ при альтернативе $H_1:\theta{>}\theta_0$. Найти функцию мощности.

Решение. Построим наиболее мощный критерий для проверки H_0 при простой альтернативе $H_1:\theta{=}\theta_1$, $\theta_1{>}\theta_0$, воспользовавшись леммой Неймана-Пирсона:

$$\frac{L_1(X,\theta)}{L_0(X,\theta)} = \frac{\prod\limits_{i=1}^{n}\theta_1\exp(-\theta_1X_i)}{\prod\limits_{i=1}^{n}\theta_0\exp(-\theta_0X_i)} = \left(\frac{\theta_1}{\theta_0}\right)^n\exp\Big\{\big(\theta_0{-}\theta_1\big)\sum_{i=1}^{n}X_i\Big\} \geqslant c_\alpha \iff T(X){=}\sum_{i=1}^{n}X_i \leq c'_\alpha.$$

c'_α найдем из условия

$$\alpha = P_{\theta_0}\left(T(X) \leq c'_\alpha\right) = P_{\theta_0}\Big(2\theta_0\sum_{i=1}^{n}X_i \leq 2\theta_0c'_\alpha\Big) = \left\{\begin{array}{c}2\theta_0X_i \sim \Gamma\left(\frac{1}{2},1\right) = \chi_2^2\\ 2\theta_0\sum_{i=1}^{n}X_i \sim \chi_{2n}^2\end{array}\right\} = F_{2n}(2\theta_0c'_\alpha),$$

где $F_m(a) = \int\limits_{-\infty}^{a}k_m(x)\,dx$, а $k_m(x)$ – плотность распределения случайной величины χ_m^2.

Отсюда $2\theta_0c'_\alpha = \chi_{\alpha,2n}^2 \implies c'_\alpha = \dfrac{\chi_{\alpha,2n}^2}{2\theta_0}$, где $\chi_{\alpha,2n}^2$ – квантиль порядка α функции $F_{2n}(y)$.

Критическая функция $\varphi(X) = \begin{cases}1, & T(X) \leq \chi_{\alpha,2n}^2/2\theta_0,\\ 0, & T(X) > \chi_{\alpha,2n}^2/2\theta_1.\end{cases}$

Функция мощности будет представляться в следующем виде:

$$W(\theta) = \mathbb{E}_\theta\varphi(X) = P\big(T(X) \leq c'_\alpha\big) = P_\theta\big(2\theta T(X) \leq 2\theta c'_\alpha\big) = F_{2n}\Big(\frac{\theta}{\theta_0}\chi_{\alpha,2n}^2\Big).$$

Построенный критерий – наиболее мощный, если гипотеза H_1 – простая, то есть $H_1:\theta{=}\theta_1$. При построении критерия значение θ_1 используется неявно, важно лишь, что $\theta_1{>}\theta_0$. Значит, построенный критерий – равномерно наиболее мощный. $\square$

Задача №37. Пусть $X_1,\ldots,X_n$ независимы и имеют гамма-распределение $\Gamma(\theta,2)$. Построить кратчайший доверительный интервал для θ с коэффициентом доверия α, основанный на центральной статистике $G(X,\theta) = \theta\cdot\sum\limits_{i=1}^{n}X_i$

Решение. Прежде всего заметим, что

$$2\theta X_i \sim \Gamma\Big(\frac{1}{2},2\Big) \implies 2\theta\sum_{i=1}^{n}X_i \sim \Gamma\Big(\frac{1}{2},2n\Big) = \chi_{4n}^2.$$

Написав определение доверительного интервала с уровнем доверия α и воспользовавшись описанными выше соотношениями получим следующую цепочку равенств:

$$\alpha = P\big(g_1 < 2G(X,\theta) < g_2\big) = P\Big(\frac{g_1}{2n\overline{X}} < \theta < \frac{g_2}{2n\overline{X}}\Big) = F_{4n}(g_2) - F_{4n}(g_1),$$

где $F_{4n}(y)$ – функция распределения χ_{4n}^2.

Для построения наименьшего интервала необходимо минимизировать $g_2 - g_1$ при условии $F_{4n}(g_2) - F_{4n}(g_1) = \alpha$.

Запишем функцию Лагранжа:

$$F(\lambda) = g_2 - g_1 - \lambda(F_{4n}(g_2) - F_{4n}(g_1) - \alpha).$$

Продифференцируем ее по переменным g_1, g_2 и λ и приравняем полученные выражения к нулю. В результате получим систему уравнений

$$\begin{cases}F_{4n}(g_2) - F_{4n}(g_1) = \alpha,\\ F'_{4n}(g_2) = F'_{4n}(g_2).\end{cases} \qquad (*)$$

Кратчайший доверительный интервал для θ с коэффициентом доверия α, основанный на центральной статистике $G(X,\theta)$, имеет вид $\left(\dfrac{g_1}{2n\overline{X}}, \dfrac{g_2}{2n\overline{X}}\right)$, где g_1 и g_2 - решения системы $(*)$. $\square$

Задача №38. Пусть $X_1,\ldots,X_n$ независимы и имеют пуассоновское распределение $\Pi(\theta)$. Построить равномерно наиболее мощный критерий размера α для проверки гипотезы $H_0:\theta = \theta_0$ при альтернативе $H_1:\theta < \theta_0$. Найти функцию мощности.

Решение. Построим наиболее мощный критерий для простой альтернативы H_1 : $\theta = \theta_1,\ \theta_1 < \theta_0$. Для этого воспользуемся леммой Неймана–Пирсона:

$$\frac{L_1}{L_0} = \frac{\prod\limits_{i=1}^{n}e^{-\theta_1}\dfrac{\theta_1^{X_i}}{X_i!}}{\prod\limits_{i=1}^{n}e^{-\theta_0}\dfrac{\theta_0^{X_i}}{X_i!}} = \left(\frac{\theta_1}{\theta_0}\right)^{\sum\limits_{i=1}^{n}X_i}\exp\big(-n(\theta_1{-}\theta_0)\big){\geq}\, c_\alpha \iff T(X){=}\sum_{i=1}^{n}X_i \leq c'_\alpha.$$

Учтем, что $T(X) \sim \Pi(\theta n)$. c'_α найдем из условия

$$F_{\theta_0}(c'_\alpha - 1) < \alpha \leq F_{\theta_0}(c'_\alpha), \quad c'_\alpha \in \mathbb{Z},$$

где $F_{\theta_0}(y)$ – функция распределения $T(X)$ при условии, что гипотеза H_0 верна:

$$F_{\theta_0}(y) = \sum_{k=0}^{[y]}\exp(-n\theta_0)\frac{(n\theta_0)^k}{k!}.$$

Критическая функция

$$\varphi(X) = \begin{cases}1, & T(X) < c'_\alpha,\\ \varepsilon_\alpha, & T(X) = c'_\alpha,\\ 0, & T(X) > c'_\alpha;\end{cases} \qquad \text{где } \varepsilon_\alpha = \frac{\alpha - F_{\theta_0}(c'_\alpha - 1)}{F_{\theta_0}(c'_\alpha) - F_{\theta_0}(c'_\alpha - 1)}.$$

Если найдется $c'_\alpha \in \mathbb{Z}$ такое, что $F_{\theta_0}(c'_\alpha) = \alpha$, критерий будет нерандомизированным:

$$\varphi(X) = \begin{cases}1, & T(X) \leq c'_\alpha,\\ 0, & T(X) > c'_\alpha.\end{cases}$$

Построенный критерий – наиболее мощный, если гипотеза H_1 - простая ($H_1:\theta{=}\theta_1$). При построении критерия значение θ_1 используется неявно: важно лишь, что $\theta_1{<}\theta_0$. Значит, построенный критерий – равномерно наиболее мощный.

Функция мощности:

$$W(\varphi,\theta) = \mathbb{E}_\theta\varphi(X) = 1\cdot P_\theta\big(T(X) < c'_\alpha\big) + \varepsilon_\alpha\cdot P_\theta\big(T(X) = c'_\alpha\big) = F_\theta(c'_\alpha{-}1) + \varepsilon_\alpha F_\theta(c'_\alpha),$$
$$\text{где} \quad F_\theta(y) = \sum_{k=0}^{[y]}\exp(-n\theta)\frac{(n\theta)^k}{k!}. \qquad \square$$

Задача №39. Пусть $X_1, \ldots, X_n$ независимы и имеют плотность распределения

$$f(x,\theta) = \begin{cases} \exp\big\{-(x-\theta)\big\}, & x > \theta, \\ 0, & x \le \theta. \end{cases}$$

Построить наиболее мощный критерий размера α для проверки гипотезы $H_0 : \theta{=}\theta_0$ при альтернативе $H_1 : \theta{=}\theta_1 < \theta_0$. Найти мощность критерия.

Решение. Для решения данной задачи воспользуемся леммой Неймана–Пирсона:

$$\frac{L_1}{L_0} = \frac{\prod\limits_{i=1}^{n} \exp(\theta_1 - X_i)\mathbb{I}_{\big\{X_i > \theta_1\big\}}}{\prod\limits_{i=1}^{n} \exp(\theta_0 - X_i)\mathbb{I}_{\big\{X_i > \theta_0\big\}}} = \exp\big(n(\theta_1{-}\theta_0)\big)\frac{\mathbb{I}_{\big\{X_{(1)} > \theta_1\big\}}}{\mathbb{I}_{\big\{X_{(1)} > \theta_0\big\}}} > c_\alpha.$$

Поскольку

$$\frac{L_1}{L_0} = \begin{cases} \infty, & X_{(1)} \le \theta_0, \\ \exp\big(n(\theta_1 - \theta_0)\big), & X_{(1)} > \theta_0; \end{cases}$$

то при $X_{(1)} \le \theta_0$ и $\dfrac{L_1}{L_0} > c_\alpha$ для любого α критическая функция принимает вид

$$\varphi(X) = \begin{cases} 1, & X_{(1)} \le \theta_0, \\ \varepsilon_\alpha, & X_{(1)} > \theta_0. \end{cases}$$

Отыщем ε_α:

$$\alpha = \mathbb{E}_{\theta_0}\varphi(X) = 1 \cdot P_{\theta_0}\big(X_{(1)} \le \theta_0\big) + \varepsilon_\alpha \cdot P_{\theta_0}\big(X_{(1)} > \theta_0\big) = 1 \cdot 0 + \varepsilon_\alpha \cdot 1,$$

откуда $\alpha = \varepsilon_\alpha$.
Мощность критерия

$$W(\varphi,\theta_1) = \mathbb{E}_{\theta_1}\varphi(X) = P_{\theta_1}\big(X_{(1)} \le \theta_0\big) + \alpha P_{\theta_1}\big(X_{(1)} > \theta_0\big) =$$
$$= \left(1 - \left(\int_{\theta_0}^{+\infty} \exp(\theta_1{-}x)\,dx\right)^n\right) + \alpha\left(\int_{\theta_0}^{+\infty} \exp(\theta_1{-}x)\,dx\right)^n = 1 - (1{-}\alpha)\exp\big(n(\theta_1{-}\theta_0)\big).$$

□

Задача №40. Пусть $X_1, \ldots, X_n$ независимы и имеют равномерное распределение на отрезке $[0,\theta]$. Построить наиболее мощный критерий размера α для проверки гипотезы $H_0 : \theta{=}\theta_0$ при альтернативе $H_1 : \theta{=}\theta_1 < \theta_0$. Найти мощность критерия.

Решение. Как и в предыдущих задачах, воспользуемся леммой Неймана–Пирсона (следует отметить, что поведение при $X_{(1)}{<}0$ и $X_{(n)} > \theta_0$ нас не интересует):

$$\frac{L_1}{L_0} = \frac{\prod\limits_{i=1}^{n} \frac{1}{\theta_1}\mathbb{I}_{\big\{0 \le X_i \le \theta_1\big\}}}{\prod\limits_{i=1}^{n} \frac{1}{\theta_0}\mathbb{I}_{\big\{0 \le X_i \le \theta_0\big\}}} = \left(\frac{\theta_0}{\theta_1}\right)^n \frac{\mathbb{I}_{\big\{X_{(n)} \le \theta_1\big\}}}{\mathbb{I}_{\big\{X_{(n)} \le \theta_0\big\}}} = \begin{cases} 0, & X_{(n)} > \theta_1, \\ \left(\dfrac{\theta_0}{\theta_1}\right)^n, & X_{(n)} \le \theta_1. \end{cases}$$

Следовательно, критическая функция имеет вид

$$\varphi(X) = \begin{cases} \varepsilon_\alpha, & X_{(n)} \le \theta_1, \\ 0, & X_{(n)} > \theta_1. \end{cases}$$

Рассмотрим следующее выражения для α:

$$\alpha = \mathbb{E}_{\theta_0}\varphi(X) = \varepsilon_\alpha P_{\theta_0}\left(\frac{X_{(n)}}{\theta_0} \le \frac{\theta_1}{\theta_0}\right) = \varepsilon_\alpha \left(\frac{\theta_1}{\theta_0}\right)^n,$$

откуда $\varepsilon_\alpha = \min\Big\{\alpha{\cdot}\Big(\frac{\theta_0}{\theta_1}\Big)^n; 1\Big\}$.

Возможны два случая:

1. $\varepsilon_\alpha = 1$, мощность критерия $W(\varphi,\theta_1) = \mathbb{E}_{\theta_1}\varphi(X) = 1$;
2. $\varepsilon_\alpha = \alpha{\cdot}\left(\dfrac{\theta_0}{\theta_1}\right)^n$, мощность критерия $W(\varphi,\theta_1) = \mathbb{E}_{\theta_1}\varphi(X) = \varepsilon_\alpha P_{\theta_1}(X_{(n)} \le \theta_1) = \varepsilon_\alpha$. □