**4.** Случайная величина распределена равномерно на отрезке $[\theta, 2\theta]$. ($\theta > 0$ очевидно)

a) По выборке объема $n$ найти оценки параметра $\theta$ методом моментов и методом максимального правдоподобия.

b) Проверить оценки на несмещенность и состоятельность. Исправить эти оценки, если необходимо.

c) Сравнить асимптотическую эффективность оценок?

d) Построить точный доверительный интервал для параметра $\theta$.

e) Построить асимптотический доверительный интервал для параметра $\theta$.

f) Сгенерируйте выборку объема $n = 100$ для некоторого значения параметра $\theta$. Вычислите указанные выше доверительные интервалы для доверительной вероятности 0.95.

g) Численно постройте бутстраповский доверительный интервал.

h) Сравнить все интервалы.

a) $\xi \sim p(x, \theta) = \frac{1}{\theta}\{[\theta;2\theta]\}\ \theta \in [\theta; 2\theta]$ $\quad \vec{X_n}$ - выборка $\quad F(x) = \frac{x}{\theta} - 1$

Метод моментов:

$\alpha_k = M[\xi^k] = \int_{-\infty}^{+\infty} x^k p(x,\theta)\,dx \qquad \tilde{\alpha}_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k$

$M[\xi] = \int_{\theta}^{2\theta} x\frac{1}{\theta}\,dx = \frac{3}{2}\theta \qquad \tilde{\alpha}_1 = \overline{X} \qquad \frac{3}{2}\theta = \overline{X} \Rightarrow \hat{\theta}_1 = \frac{2}{3}\overline{X}$

Метод максимального правдоподобия:

$L(x, \theta) = \prod_{i=1}^{n} p(x_i, \theta) = \left(\frac{1}{\theta}\right)^n \qquad \ln L = -n\ln(\theta)$

$\frac{\delta \ln L}{\delta \theta} = \frac{-n}{\theta}$ $\quad$ sup на $[\theta; 2\theta]$ при $\theta \to \max$ (3, 6)

$\begin{cases} 2\theta \geq X_{max} \\ \theta \leq X_{min} \end{cases} \to \frac{1}{2}X_{max} \leq \theta \leq X_{min} \to$ sup при $\tilde{\theta}_2 = \frac{1}{2}X_{max}$

b) $M[\tilde{\theta}_1] = M\left[\frac{2}{3n}\sum_{i=1}^{n} x_i\right] = \frac{2}{3}M\xi = \frac{2}{3}\cdot\frac{3}{2}\theta = \theta$ - несмещенная

$D[\tilde{\theta}_1] = \frac{4}{9n^2}\sum D\xi \qquad M[\xi^2] = \int_{\theta}^{2\theta} x^2\frac{1}{\theta}\,dx = \frac{7\theta^2}{3} \qquad D\xi = \left(\frac{7}{3}\theta^2 - \frac{9}{4}\theta^2\right)n = \frac{\theta^2}{12}\cdot n$

$D[\tilde{\theta}_1] = \frac{4}{9n^2}\cdot n\cdot\frac{\theta^2}{12} = \frac{\theta^2}{27n} \xrightarrow[n\to\infty]{} 0$ - несмещенная (по теоремке :<)

$$M[\hat{\theta}_2] = M\left[\tfrac{1}{2}x_{max}\right] = \frac{1}{2}\int_{\theta}^{2\theta} x\cdot n\left(\frac{x}{\theta}-1\right)^{n-1}\frac{1}{\theta}\,dx = \left|t=\frac{x}{\theta}-1\right| = \theta\int_{t_1}^{t_2} n(t+1)\cdot t^{n-1}\,dt = \frac{2n+1}{2(n+1)}\theta$$

$$g(y) = n(F(y))^{n-1}\cdot F'(y) = n\cdot\left(\frac{x}{\theta}-1\right)^{n-1}\frac{1}{\theta}$$

смещенная

$$\tilde{\theta}_2^{*} = \frac{1}{2}x_{max}\cdot\frac{2n+2}{2n+1} = \frac{n+1}{2n+1}x_{max}$$ – несмещенная

$$M[x_{max}] = \int_{\theta}^{2\theta} x\, n\cdot\left(\frac{x}{\theta}-1\right)^{n-1}\frac{1}{\theta}\,dx = \frac{2n+1}{n+1}\theta$$

$$M[x_{max}^2] = \int_{\theta}^{2\theta} x^2 n\left(\frac{x}{\theta}-1\right)^{n-1}\frac{1}{\theta}\,dx = \frac{n(4n^2+8n+2)\theta^2}{n(n+1)(n+2)}$$

$$D[\tilde{\theta}_2^{*}] = \frac{(n+1)^2}{(2n+1)^2}\cdot D[x_{max}] = \frac{(n+1)^2}{(2n+1)^2}\left[\frac{4n^2+8n+2}{(n+1)(n+2)} - \left(\frac{2n+1}{n+1}\right)^2\right]\theta^2 =$$

$$\frac{\theta^2 n}{4n^3+12n^2+9n+2} = \frac{\theta^2 n}{(2n+1)^2(n+2)} \xrightarrow[n\to\infty]{} 0$$

$\tilde{\theta}_2^{*}$ – состоятельная (по теореме ?<)

c) Посмотрим, что там с моделью:

$$I(\theta) = M\left[\left(\frac{-1}{\theta}\right)^2\right] = \frac{1}{\theta^2} > 0 \;\forall\, \theta\in[\theta,2\theta],\quad g'(\theta) = \frac{-1}{\theta^2},$$ $\rho$ непр. дифф-мо по $\theta$ на $\Theta$

$$0 = \frac{\partial}{\partial\theta}\int_{\theta}^{2\theta}\rho(x,\theta)\,dx = \int_{\theta}^{2\theta}\frac{\partial}{\partial\theta}\rho(x,\theta)\,dx = \int_{\theta}^{2\theta}\frac{-1}{\theta^2}\,dx = \frac{-1}{\theta} \neq 0$$ – модель нерегулярна

$$nD[\hat{\theta}_1] = \frac{\theta^2}{27} \xrightarrow[n\to\infty]{} \frac{\theta^2}{27}$$

$$nD[\tilde{\theta}_2^{*}] = \frac{\theta^2 n^2}{(2n+1)^2(n+2)} \xrightarrow[n\to\infty]{} 0$$

$\tilde{\theta}_2^{*}$ более ас. эффективна чем $\hat{\theta}_1$

d) $P(t_1 < F(x_n,\theta) < t_2) \geq \beta$

$t_2 - t_1 \to \min$

Магически образом (нет) выбираем $f(X_n, \theta)$

Как сказали умные дяди (или тёти) с этой ссылки

http://new.math.msu.su/department/probab/io/teorver-online/teorver57.html

**Пример 8.1**

Пусть $x = (x_1, \dots, x_n)$ -- независимая выборка из равномерного распределения в отрезке $[0, \theta]$ с неизвестным параметром $\theta > 0$:

$$F_\theta(t) = \begin{cases} 0, & t < 0, \\ t/\theta, & 0 \le t \le \theta, \\ 1, & t > \theta. \end{cases}$$

Пусть задана доверительная вероятность $\gamma$. Построим доверительный интервал для $\theta$.

1. Рассмотрим функцию $G(x_1, \dots, x_n, \theta) = \theta^{-1} \max(x_1, \dots, x_n)$. Вычислим ее

берем $f(X_n, \theta) = \dfrac{X_{max}}{\theta}$ с $F_f(x) = (F(\theta x))^n = (x-1)^n$

$t_1 = q_{\frac{1-\beta}{2}}$ $\quad t_2 = q_{\frac{1+\beta}{2}}$ $\quad t_1 < \dfrac{X_{max}}{\theta} < t_2$

$\dfrac{X_{max}}{t_1} < \theta < \dfrac{X_{max}}{t_2}$ $\quad t_2 = \sqrt[n]{\dfrac{1+\beta}{2}} + 1$ $\quad t_1 = \sqrt[n]{\dfrac{1-\beta}{2}} + 1$

$\left( \int_1^{t_1} n(x-1)^{n-1} dx = \dfrac{1-\beta}{2} \rightarrow (x-1)^n \Big|_1^{t_1} = \dfrac{1-\beta}{2} \right)$

$$\underbrace{\frac{X_{max}}{\sqrt[n]{\frac{1+\beta}{2}} + 1} < \theta < \frac{X_{max}}{\sqrt[n]{\frac{1-\beta}{2}} + 1}}_{\text{точный доверительный интервал}}$$

е) Ищем по ОМП (нет, не ищем, наша модель не является сильно регулярной → нет асимпт. нормал. → нет тех формул с лекции)

Ищем по ОММ: $\quad f(\hat\alpha) = \hat\theta = \dfrac{2}{3}\bar{X}$ $\quad \nabla f = \dfrac{2}{3}$

$\sqrt{n}\,\dfrac{g(\hat\alpha) - g(\alpha)}{\sigma(\alpha)} \rightsquigarrow N(0,1)$ $\quad f(\alpha) = \theta = \dfrac{2}{3}\alpha_1$ $\quad \alpha_1 = \bar{x}$ $\quad \alpha_2 = \overline{x^2}$

$K = \overline{x^2} - \bar{x}^2$ $\quad \sigma(\alpha) = \dfrac{2}{3}\sqrt{\overline{x^2} - \bar{x}^2}$

$$\frac{\hat{\theta}-\theta}{\frac{2}{3}\sqrt{\overline{x^2}-\overline{x}^2}} \rightsquigarrow N(0,1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \qquad u_{\frac{1+\beta}{2}} = -u_{\frac{1-\beta}{2}}$$

(тут можно было просто использовать $u_{\frac{1-\beta}{2}} = -1{,}96$, но я не я если не полезу в дебри)

$$u_{\frac{1-\beta}{2}} = u_1$$

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u_{\frac{1-\beta}{2}}} e^{-\frac{x^2}{2}}\,dx = \frac{1-\beta}{2} = \left(\text{по вольфраму}\right) = \frac{1}{2}\left(\operatorname{erf}\left(\frac{u_1}{\sqrt{2}}\right)+1\right) = \frac{1-\beta}{2}$$

$$\operatorname{erf}\left(\frac{u_{\frac{1-\beta}{2}}}{\sqrt{2}}\right) = -\beta \rightarrow \qquad \operatorname{erf}^{-1}(-\beta) = \frac{u_{\frac{1-\beta}{2}}}{\sqrt{2}}$$

(из Inverse functions Error function из википедии)

$$\operatorname{erf}(\operatorname{erf}^{-1} x) = x$$

$$-\sqrt{2}\operatorname{erf}^{-1}(-\beta) > \sqrt{n}\,\frac{\tilde{\theta}-\theta}{\frac{2}{3}\sqrt{\overline{x^2}-\overline{x}^2}} > \sqrt{2}\operatorname{erf}^{-1}(-\beta), \qquad \tilde{\theta} = \frac{2}{3}\overline{x}$$

$$\underbrace{-\frac{1}{\sqrt{n}}\sqrt{2}\operatorname{erf}^{-1}(-\beta)\cdot\frac{2}{3}\sqrt{\overline{x^2}-\overline{x}^2} + \frac{2}{3}\overline{x} > \theta > \frac{\sqrt{2}}{\sqrt{n}}\operatorname{erf}^{-1}(-\beta)\cdot\frac{2}{3}\sqrt{\overline{x^2}-\overline{x}^2} + \frac{2}{3}\overline{x}}_{\text{асимптотический доверительный интервал}}$$

**5.** Случайная величина имеет распределение Парето:

для википедии это $x_m = 1$, $k = \theta - 1$

$$p(x) = \begin{cases} \dfrac{\theta - 1}{x^\theta}, & x \geq 1 \\ 0, & x < 1 \end{cases}, \quad \theta > 1.$$

a) По выборке объема $n$ найти оценку параметра $\theta$ методом максимального правдоподобия.

b) Построить доверительный интервал для медианы.

c) Построить асимптотический доверительный интервал для параметра $\theta$.

d) Сгенерируйте выборку объема $n = 100$ для некоторого значения параметра $\theta$. Вычислите указанные выше доверительные интервалы для доверительной вероятности 0.95.

t) Численно постройте бутстраповский доверительный интервал двумя способами, используя параметрический бутстрап и непараметрический бутстрап.

f) Сравнить все интервалы.

a)

Метод максимального правдоподобия

$$L(x, \theta) = \prod_{i=1}^{n} p(x_i, \theta) = \frac{(\theta - 1)^n}{\prod_{i=1}^{n} x_i^\theta} \qquad \ln L = n \ln(\theta - 1) - \theta \sum_{i=1}^{n} \ln x_i$$

$$\frac{\delta \ln L}{\delta \theta} = \frac{n}{\theta - 1} - \sum_{i=1}^{n} \ln x_i = 0 \qquad \sum_{i=1}^{n} \ln x_i = \frac{n}{\theta - 1} \qquad \tilde{\theta} = \frac{n}{\sum_{i=1}^{n} \ln x_i} + 1$$

b) med - медиана

$$\int_{-\infty}^{med} p(x)\,dx = \frac{1}{2} \qquad \int_{1}^{med} (\theta - 1) \frac{1}{x^\theta} dx = -x^{-\theta+1} \Big|_1^{med} = \frac{1}{2}$$

$$med = 2^{\frac{1}{\theta - 1}} \qquad \frac{f(\tilde{\theta}) - f(\theta)}{\sigma(\theta)} \sqrt{n} \sim N(0, 1)$$

$$f(\theta) = 2^{\frac{1}{\theta - 1}} \rightarrow \nabla f = 2^{\frac{1}{\theta - 1}} \ln 2 \cdot \left( \frac{1}{(\theta - 1)^2} \right)$$

$$f(\tilde{\theta}) = 2^{\frac{1}{\tilde{\theta} - 1}} \qquad \tilde{\theta} = \frac{n}{\sum_{i=1}^{n} \ln x_i} + 1$$

$$I(\theta) = M\left[ \left( \frac{\delta \ln p}{\delta \theta} \right)^2 \right] = \int_1^\infty \left( \frac{1}{\theta - 1} - \ln x \right)^2 \frac{\theta - 1}{x^\theta} dx = \frac{1}{(\theta - 1)^2}$$

$$\sigma(\tilde{\theta}) = \left( 2^{\frac{1}{\theta - 1}} \ln 2 \cdot \left( \frac{1}{(\tilde{\theta} - 1)^2} \right) \cdot (\tilde{\theta} - 1)^2 \cdot 2^{\frac{1}{\tilde{\theta} - 1}} \ln 2 \cdot \left( \frac{1}{(\tilde{\theta} - 1)^2} \right) \right)^{\frac{1}{2}} = 2^{\frac{1}{\tilde{\theta} - 1}} \ln 2 \cdot \frac{1}{(\tilde{\theta} - 1)}$$

$$-\sqrt{2}\,\mathrm{erf}^{-1}(-\beta) > \frac{2^{\frac{1}{\hat\theta-1}} - med}{2^{\frac{1}{\hat\theta-1}}\cdot \ln 2 \cdot \frac{1}{(\hat\theta-1)^2}}\sqrt{n} > \sqrt{2}\,\mathrm{erf}^{-1}(-\beta)$$

$$\underbrace{\frac{\sqrt{2}}{\sqrt{n}}\mathrm{erf}^{-1}(-\beta)\cdot 2^{\frac{1}{\hat\theta-1}}\cdot \ln 2\cdot\frac{1}{(\hat\theta-1)^2} + 2^{\frac{1}{\hat\theta-1}} < med < -\frac{\sqrt{2}}{\sqrt{n}}\mathrm{erf}^{-1}(-\beta)\cdot\frac{1}{2^{\frac{1}{\hat\theta-1}}}\frac{\ln 2}{(\hat\theta-1)^2} + 2^{\frac{1}{\hat\theta-1}}, \quad \text{где } \hat\theta = \frac{n}{\sum_{i=1}^{n}\ln x_i} + 1}_{\text{доверительный интервал для med}}$$

Проверяем, что модель регулярна: $\exists\, p$ непр. дифф-до по $\theta$ в $\Theta$

$$0 = \frac{\partial}{\partial\theta}\int_1^{\infty} p(x,\theta)\,dx = \int_1^{\infty}\frac{\partial}{\partial\theta}p(x,\theta)\,dx = \int_1^{\infty}\frac{x^{\theta} - x^{\theta}\ln x\,(\theta-1)}{x^{2\theta}}\,dx = x^{1-\theta}\ln x\Big|_1^{\infty} = 0$$

модель регулярна, поэтому можно ОМП и асимп. эфф. и все остальное

c)

По ОМП: $f(\theta) = \theta \qquad \nabla f = 1$

$f(\hat\theta) = \hat\theta \qquad I^{-1} = (\theta-1)^2 \qquad \sigma = \theta - 1, \quad \sigma(\hat\theta) = \hat\theta - 1$

$$-\sqrt{2}\,\mathrm{erf}^{-1}(-\beta) > \frac{\hat\theta-\theta}{\hat\theta-1}\sqrt{n} > \sqrt{2}\,\mathrm{erf}^{-1}(-\beta)$$

$$\underbrace{\frac{\sqrt{2}}{\sqrt{n}}(\hat\theta-1)\,\mathrm{erf}^{-1}(-\beta) + \hat\theta < \theta < -\frac{\sqrt{2}}{\sqrt{n}}\mathrm{erf}^{-1}(-\beta)(\hat\theta-1) + \hat\theta}_{\text{асимп. дов. интервал по ОМП}} \quad \text{где } \hat\theta = \frac{n}{\sum_{i=1}^{n}\ln x_i} + 1$$

По ОММ:

оценка $\alpha_1 = M_\xi = \int_1^{\infty} x\,\frac{\theta-1}{x^{\theta}}\,dx = \frac{\theta-1}{\theta-2}$ для $\theta > 2$.

$\tilde{\alpha}_1 = \bar{x}$ $\qquad 1 + \frac{1}{\theta-2} = \bar{x} \to \tilde{\theta} = \frac{1}{\bar{x}-1}+2$ при $\theta > 2$

$f(\hat{\alpha}_1) = \hat{\theta} = \frac{1}{\bar{x}-1}+2$ $\qquad \nabla f(\hat{\alpha}_1) = \frac{-1}{(\bar{x}-1)^2}$ $\qquad \kappa = \hat{\alpha}_2 - \hat{\alpha}_1^2 = \overline{x^2} - \bar{x}^2$

$$-\sqrt{2}\,\mathrm{erf}^{-1}(-\beta) > \frac{\frac{1}{\bar{x}-1}+2-\theta}{\frac{1}{(\bar{x}-1)^2}\sqrt{(\overline{x^2}-\bar{x}^2)}}\sqrt{n} > \sqrt{2}\,\mathrm{erf}^{-1}(-\beta)$$

$$\underbrace{+\frac{\sqrt{2}}{\sqrt{n}}\mathrm{erf}^{-1}(-\beta)\left(\frac{\sqrt{\overline{x^2}-\bar{x}^2}}{(\bar{x}-1)^2}\right)+\frac{1}{\bar{x}-1}+2 < \theta < -\frac{\sqrt{2}}{\sqrt{n}}\mathrm{erf}^{-1}(-\beta)\left(\frac{\sqrt{\overline{x^2}-\bar{x}^2}}{(\bar{x}-1)^2}\right)+\frac{1}{\bar{x}-1}+2}_{\text{асимпт. дов. инт. по ОММ}}$$

можно хорошую оценочку пожалуйста

$$0 = \frac{\partial^2}{\partial\theta^2}\int_1^{\infty} p(x,\theta)\,dx = \int_1^{\infty}\frac{\partial^2}{\partial\theta^2}p(x,\theta)\,dx = \int_1^{\infty}\frac{\ln x\cdot((\theta-1)\ln x-2)}{x^{\theta}}\,dx = 0$$

сказал Вольфрам про

$$\int \frac{\ln x\cdot((\theta-1)\ln x-2)}{x^{\theta}}\,dx = -x^{(1-\theta)}\cdot\ln^2 x$$

$\to$ модель мега (целых 2 раза) регулярна $\to$ сильнорегулярна $\to$ теорема работает и все в шоколаде