Министерство образования и науки Российской Федерации

Федеральное государственное бюджетное образовательное учреждение высшего образования
«НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

Кафедра прикладной математики

РГЗ по курсу «ТВиМС»

Часть I

Факультет:	ПМИ
Группа:	ПМ-63
Студент:	Шепрут И.И.
Вариант:	22
Преподаватели:	Постовалов С.Н.
	Веретельникова И.В.

Новосибирск
2018

1. Постановка задачи

Пусть $X_1, X_2, \ldots, X_n$ — выборка из распределения Вейбулла, с неизвестным параметром λ и известным α, обладающее плотностью:

$$f(x) = \frac{\alpha x^{\alpha-1}}{\lambda^\alpha} \exp\left\{-\left(\frac{x}{\lambda}\right)^\alpha\right\},\ x \geqslant 0,\ \lambda > 0,\ \alpha > 0 \tag{1}$$

1. Найти точечную оценку неизвестного параметра θ (или некоторой функции $\tau(\theta)$) по методу моментов или по методу максимального правдоподобия. Проверить полученную оценку на несмещенность, состоятельность и эффективность.
2. Найти достаточную статистику.
3. Найти функцию $\tau(\theta)$, допускающую эффективную оценку.
4. Построить точный доверительный интервал.
5. Построить асимптотический доверительный интервал.

2. Решение

2.1. Нахождение точечной оценки

Вычисляем момент произвольного порядка для (1):

$$\begin{aligned}
\mathsf{M}X_i^k &= \int\limits_0^\infty x^k f(x)\,dx = \int\limits_0^\infty x^k \frac{\alpha x^{\alpha-1}}{\lambda^\alpha} \exp\left\{-\left(\frac{x}{\lambda}\right)^\alpha\right\}\,dx = \frac{\alpha}{\lambda^a}\int\limits_0^\infty x^{\alpha+k-1} \exp\left\{-\left(\frac{x}{\lambda}\right)^\alpha\right\}\,dx = \\
&= \left| \begin{matrix} t = \left(\frac{x}{\lambda}\right)^\alpha, x = \lambda t^{\frac{1}{\alpha}} \\ a_t = 0, b_t = \infty \\ dx = \frac{\lambda}{\alpha} t^{\frac{1}{\alpha}-1}\,dt \end{matrix} \right| = \frac{\alpha}{\lambda^a}\int\limits_0^\infty \left(\lambda t^{\frac{1}{\alpha}}\right)^{\alpha+k-1} \cdot e^{-t} \cdot \frac{\lambda}{\alpha} t^{\frac{1}{\alpha}-1}\,dt = \lambda^k \int\limits_0^\infty t^{\left(1+\frac{k}{\alpha}\right)-1} e^{-t}\,dt = \\
&= \boxed{\lambda^k \Gamma\left(1+\frac{k}{\alpha}\right) = \mathsf{M}X_i^k}
\end{aligned} \tag{2}$$

Вычисляем математическое ожидание для (1):

$$\boxed{\mathsf{M}X_i = \lambda\Gamma\left(1+\frac{1}{\alpha}\right)} \tag{3}$$

Вычисляем дисперсию для (1):

$$\begin{aligned}
\mathsf{D}\,X_i &= \mathsf{M}X_i^2 - (\mathsf{M}X_i)^2 = \lambda^2\Gamma\left(1+\frac{2}{\alpha}\right) - \lambda^2\Gamma^2\left(1+\frac{1}{\alpha}\right) = \\
&= \boxed{\lambda^2\left[\Gamma\left(1+\frac{2}{\alpha}\right) - \Gamma^2\left(1+\frac{1}{\alpha}\right)\right] = \mathsf{D}\,X_i}
\end{aligned} \tag{4}$$

2.1.1. Нахождение оценки методом моментов

Приравниваем теоретический (3) и выборочный момент:

$$\mathsf{M}X_i = \bar{X}$$

$$\hat{\theta}\Gamma\left(1 + \frac{1}{\alpha}\right) = \bar{X}$$

$$\boxed{\hat{\theta} = \frac{\bar{X}}{\Gamma\left(1 + \frac{1}{\alpha}\right)}} \tag{5}$$

$$\boxed{T(\mathbb{X}_n) = \frac{\bar{X}}{\Gamma\left(1 + \frac{1}{\alpha}\right)}} \tag{6}$$

2.1.2. Проверка на несмещенность

Статистика $T(\mathbb{X}_n)$ называется *несмещенной* оценкой параметра θ, если выполняется условие $\mathsf{M}[T(\mathbb{X}_n)] = \theta,\ \forall\theta \in \Theta$.
Проверяем это для статистики (6):

$$\mathsf{M}[T(\mathbb{X}_n)] \overset{?}{=} \theta \Rightarrow \quad \mathsf{M}\left[\frac{\bar{X}}{\Gamma\left(1 + \frac{1}{\alpha}\right)}\right] \overset{?}{=} \theta \Rightarrow$$

$$\Rightarrow \frac{\mathsf{M}\bar{X}}{\Gamma\left(1 + \frac{1}{\alpha}\right)} \overset{?}{=} \theta \Rightarrow \quad \frac{\frac{1}{n}\sum_{i=1}^{n}\mathsf{M}X_i}{\Gamma\left(1 + \frac{1}{\alpha}\right)} \overset{?}{=} \theta \Rightarrow$$

$$\Rightarrow \frac{\frac{1}{n}n\theta\Gamma\left(1 + \frac{1}{\alpha}\right)}{\Gamma\left(1 + \frac{1}{\alpha}\right)} \overset{?}{=} \theta \Rightarrow \quad \boxed{\theta = \theta}$$

Таким образом, оценка является несмещенной.

2.1.3. Проверка на состоятельность

Критерий состоятельности: если оценка $\hat{\theta}$ является несмещенной и $\mathsf{D}\,T(\mathbb{X}_n) \xrightarrow{n\to\infty} 0$, то она также является и состоятельной.
Проверяем (5) на состоятельность:

$$\mathsf{D}\,T(\mathbb{X}_n) = \mathsf{D}\left[\frac{\bar{X}}{\Gamma\left(1 + \frac{1}{\alpha}\right)}\right] = \frac{\frac{1}{n}\mathsf{D}\,X_i}{\Gamma^2\left(1 + \frac{1}{\alpha}\right)} = \boxed{\frac{1}{n}\cdot\frac{\theta^2\left[\Gamma\left(1 + \frac{2}{\alpha}\right) - \Gamma^2\left(1 + \frac{1}{\alpha}\right)\right]}{\Gamma^2\left(1 + \frac{1}{\alpha}\right)} = \mathsf{D}\,T(\mathbb{X}_n)} \tag{7}$$

$$\mathsf{D}\,T(\mathbb{X}_n) \xrightarrow{n\to\infty} 0$$

Таким образом, оценка является состоятельной.

2.1.4. Проверка на эффективность

Если распределение является регулярным, и достигается нижняя граница в неравенстве Рао-Крамера (8), то оценка называется *эффективной.*

$$\mathsf{D}\,T(\mathbb{X}_n) = \frac{\left[\tau'(\theta)\right]^2}{n\cdot i(\theta)} \tag{8}$$

Распределение, очевидно, является регулярным, т. к. плотность (1) дифференцируема по θ, и множество $\{x : f(x,\theta) = 0\}$ не зависит от θ. Поэтому проверим неравенство Рао-Крамера. Но для начала вычислим $i(\theta)$:

$$\begin{aligned} i(\theta) &= -\mathsf{M}\left[\frac{\partial^2 \ln f(X,\theta)}{\partial\theta^2}\right] = -\mathsf{M}\left[\frac{\partial^2\left((\alpha-1)\ln x - \alpha\ln\theta - \left(\frac{X}{\theta}\right)^\alpha\right)}{\partial\theta^2}\right] = \mathsf{M}\left[\frac{\partial\left(\frac{\alpha}{\theta} - \frac{\alpha}{\theta}\left(\frac{X}{\theta}\right)^\alpha\right)}{\partial\theta}\right] = \\ &= \mathsf{M}\left[-\frac{\alpha}{\theta^2} + \frac{\alpha(\alpha+1)}{\theta^2}\left(\frac{X}{\theta}\right)^\alpha\right] = -\frac{\alpha}{\theta^2} + \frac{\alpha(\alpha+1)}{\theta^{\alpha+2}}\mathsf{M}X^\alpha = -\frac{\alpha}{\theta^2} + \frac{\alpha(\alpha+1)}{\theta^{\alpha+2}}\theta^\alpha\Gamma\left(1+\frac{\alpha}{\alpha}\right) = \\ &= \boxed{\frac{\alpha^2}{\theta^2} = i(\theta)} \end{aligned} \tag{9}$$

Проверка равенства Рао-Крамера (8), используя (9) и (7):

$$\begin{aligned} \mathsf{D}\,T(\mathbb{X}_n) &\stackrel{?}{=} \frac{\left[\tau'(\theta)\right]^2}{n\cdot i(\theta)} \\ \frac{1}{n}\cdot\frac{\theta^2\left[\Gamma\left(1+\frac{2}{\alpha}\right) - \Gamma^2\left(1+\frac{1}{\alpha}\right)\right]}{\Gamma^2\left(1+\frac{1}{\alpha}\right)} &\stackrel{?}{=} \frac{\theta^2}{n\alpha^2} \\ \frac{\Gamma\left(1+\frac{2}{\alpha}\right)}{\Gamma^2\left(1+\frac{1}{\alpha}\right)} &\stackrel{?}{=} \frac{1}{\alpha^2}+1 \end{aligned}$$

Это равенство выполняется при $\alpha = 1$ и $\alpha \approx -0.6694444982\ldots$. Первое подходит по условию, а второе нет.

Таким образом, оценка является эффективной только при $\alpha = 1$.

2.2. Нахождение достаточной статистики

Воспользуемся *критерием факторизации:* для того, чтобы статистика $T(\mathbb{X}_n)$ была достаточной для модели $F = \{F(x,\theta),\ \theta\in\Theta\}$, необходимо и достаточно, чтобы функция правдоподобия имела вид:

$$L(\mathbb{X}_n,\theta) = h(\mathbb{X}_n)\cdot g(T(\mathbb{X}_n),\theta) \tag{10}$$

Приводим к необходимому виду:

$$L(\mathbb{X}_n,\theta) = \prod_{i=1}^n \frac{\alpha X_i^{\alpha-1}}{\theta^\alpha}\exp\left\{-\frac{X_i^\alpha}{\theta^\alpha}\right\} = \underbrace{\alpha^n\theta^{-\alpha n}\exp\left\{-\frac{1}{\theta^\alpha}\overbrace{\sum_{i=1}^n X_i^\alpha}^{T(\mathbb{X}_n)}\right\}}_{g(T(\mathbb{X}_n),\theta)}\cdot\underbrace{\prod_{i=1}^n X_i^{\alpha-1}}_{h(\mathbb{X}_n)}$$

Таким образом, статистика $T(\mathbb{X}_n) = \sum_{i=1}^n X_i^\alpha$ является достаточной для модели $F = \{F(x,\theta),\ \theta\in\Theta\}$.

2.3. Нахождение функции, допускающей эффективную оценку

Нам уже известно, что распределение является регулярным, поэтому воспользуемся критерием *эффективности:* $T(\mathbb{X}_n)$ — эффективная оценка $\tau(\theta)$, если

$$U(\mathbb{X}_n,\theta)\cdot a(\theta) = T(\mathbb{X}_n) - \tau(\theta),$$

где $a(\theta)$ — некоторая функция от θ, $U(\mathbb{X}_n, \theta) = \frac{\partial \ln L(\mathbb{X}_n,\theta)}{\partial \theta}$.

$$U(\mathbb{X}_n, \theta) = \frac{\partial \ln \left[\prod_{i=1}^n \frac{\alpha X_i^{\alpha-1}}{\theta^\alpha} \exp\left\{-\frac{X_i^\alpha}{\theta^\alpha}\right\}\right]}{\partial \theta}$$
$$U(\mathbb{X}_n, \theta) = \frac{\partial \left[n\ln\alpha - n\alpha\ln\theta + (\alpha-1)\sum_{i=1}^n \ln X_i - \frac{1}{\theta^\alpha}\sum_{i=1}^n X_i^\alpha\right]}{\partial \theta}$$
$$U(\mathbb{X}_n, \theta) = -\frac{n\alpha}{\theta} + \frac{\alpha}{\theta^{\alpha+1}}\sum_{i=1}^n X_i^\alpha$$
$$\boxed{U(\mathbb{X}_n, \theta) = \frac{n\alpha}{\theta^{\alpha+1}}\left[\frac{1}{n}\sum_{i=1}^n X_i^\alpha - \theta^\alpha\right]} \tag{11}$$
$$\underbrace{U(\mathbb{X}_n, \theta) \cdot \frac{\theta^{\alpha+1}}{n\alpha}}_{a(\theta)} = \underbrace{\frac{1}{n}\sum_{i=1}^n X_i^\alpha}_{T(\mathbb{X}_n)} - \underbrace{\theta^\alpha}_{\tau(\theta)}$$

Таким образом, статистика $T(\mathbb{X}_n) = \frac{1}{n}\sum_{i=1}^n X_i^\alpha$ является эффективной оценкой функции $\tau(\theta) = \theta^\alpha$.

2.4. Построение точного доверительного интервала

Для построения доверительного интервала используем центральную статистику. Так как функция распределения:

$$F(x, \lambda) = \int_{-\infty}^{x} f(t, \lambda)\, dt = \begin{cases} 0, & x < 0, \\ 1 - e^{-\left(\frac{x}{\lambda}\right)^\alpha}, & x \geqslant 0, \end{cases}$$

монотонно убывает по λ, то можем взять в качестве центральной статистики функцию

$$G(\mathbb{X}_n, \theta) = -\sum_{i=1}^n \ln F(X_i, \theta) = -\sum_{i=1}^n \ln\left(1 - e^{-\left(\frac{X_i}{\theta}\right)^\alpha}\right) \succ \Gamma(\beta, n),\ \beta > 0$$

Тогда границы доверительного интервала $(\tilde{T}_1,\ \tilde{T}_2)$ определяются при численном решении уравнений $G(\mathbb{X}_n, T_1) = g_1$, $G(\mathbb{X}_n, T_2) = g_2$, где $g_1 < g_2$, $\tilde{T}_1 = \min\{T_1, T_2\}$, $\tilde{T}_2 = \max\{T_1, T_2\}$ и $\mathsf{P}\{g1 \leqslant G(\mathbb{X}_n, \theta) \leqslant g_2\} = \gamma$, т. е. $g_1,\ g_2$ должны удовлетворять уравнению:

$$\int_{g_1}^{g_2} \frac{x^{n-1}e^{-x/\beta}}{\Gamma(n)\beta^n}\, dx = \gamma$$

2.5. Построение асимптотического доверительного интервала

Оценки максимального правдоподобия являются асимптотически эффективными и асимптотически нормальными, следовательно сходятся к стандартному нормальному распределению. Следовательно, случайный интервал $\left(\hat{\theta} - {}^{c_\gamma}\!/\!\sqrt{ni(\hat{\theta})}, \hat{\theta} + {}^{c_\gamma}\!/\!\sqrt{ni(\hat{\theta})}\right)$, где $c_\gamma = \Phi^{-1}\left(\frac{\gamma+1}{2}\right)$, является асимптотическим γ-доверительным интервалом.

Поэтому для начала построим ОМП-оценку параметра λ. Воспользуемся (11):

$$U(\mathbb{X}_n, \theta) = 0 \quad \Rightarrow \quad \frac{n\alpha}{\theta^{\alpha+1}} \left[\frac{1}{n} \sum_{i=1}^{n} X_i^{\alpha} - \theta^{\alpha} \right] = 0 \quad \Rightarrow \quad \boxed{\hat{\theta} = \sqrt[\alpha]{\frac{1}{n} \sum_{i=1}^{n} X_i^{\alpha}}}$$

А также, используя (9), получаем: $i(\hat{\theta}) = \frac{\alpha^2}{\hat{\theta}^2}$. Строим асимптотический доверительный интервал:

$$\left(\hat{\theta} - \frac{c_\gamma}{\sqrt{ni(\hat{\theta})}}, \hat{\theta} + \frac{c_\gamma}{\sqrt{ni(\hat{\theta})}} \right)$$

$$\left(\sqrt[\alpha]{\frac{1}{n} \sum_{i=1}^{n} X_i^{\alpha}} \left(1 - \frac{c_\gamma}{\alpha\sqrt{n}} \right),\ \sqrt[\alpha]{\frac{1}{n} \sum_{i=1}^{n} X_i^{\alpha}} \left(1 + \frac{c_\gamma}{\alpha\sqrt{n}} \right) \right)$$