1. Мардиа К., Земроч П., Таблицы F -распределений. М., 1984. 2. Пирсон К. Таблицы неполной бета-функции. М., 1974. 3. O'Reilly F.J., Stephens M.A., Transforming Censored Samples for Testing Fit // Technometrics. 1988. Vol. 30, p. 79-86.

Стаття надійшла до редколегії 17.04.90

УДК 519.21

О.П.Гнатишин, Є.В.Москвяк

МЕТОД МАКСИМАЛЬНОЇ ПРАВДОПОДІБНОСТІ ДЛЯ ЗРІЗАНОЇ ВИБІРКИ

Нехай абсолютно неперервна випадкова змінна ξ з функцією розподілу ймовірностей $F(t;\theta)$, залежною від s невідомих параметрів $\theta=(\theta_1,\dots,\theta_s)$, $(s=1,2)$, описує напрацювання технічної одиниці до відмови. Тоді надійність $R(t;\theta)$, густина ймовірностей $f(t;\theta)$, інтенсивність відмов $h(t;\theta)$ і риск $H(t;\theta)$ відповідно дорівнюють

$$R=1-F=e^{-H}, \quad f=F', \quad h=\frac{f}{R}=-\frac{d}{dt}\ln R, \quad H(t;\theta)=\int_0^t h(x;\theta)\,dx. \qquad /1/$$

1. <u>Незгруповані відмови та зупинки</u>. Нехай k з n однотипних технічних одиниць працювали до відмови F, а решта $n-k$ до зупинки S. Позначимо моменти відмов через t_{iF} $(i=1,\dots,k)$, а моменти зупинок через t_{jS}, $(j=1,\dots,n-k)$. У припущенні, що напрацювання до відмови описується розподілом $F(t;\theta)$, потрібно оцінити невідомі параметри $\theta=(\theta_1,\dots,\theta_s)$.

Згідно з методом максимальної правдоподібності /ММП/ записуємо функцію правдоподібності

$$L=\prod_{i=1}^{k} f(t_{iF};\theta)\prod_{j=1}^{n-k} R(t_{jS};\theta).$$

Звідси

$$\ln L=\sum_{i=1}^{k}\ln f(t_{iF};\theta)+\sum_{j=1}^{n-k}\ln R(t_{jS};\theta), \qquad /2/$$

логарифм правдоподібності виражено через густину та надійність.

Але, згідно з /1/,

$$\ln f(t_{i_F};\theta) = \ln h(t_{i_F};\theta) + \ln R(t_{i_F};\theta).$$

Отже,

$$\ln L = \sum_{i=1}^{k} \ln h(t_{i_F};\theta) + \sum_{l=1}^{n} \ln R(t_l;\theta),$$

де t_l – момент відмови або зупинки. Знову, згідно з /1/,

$$\ln R(t;\theta) = -H(t;\theta).$$

Таким чином,

$$\ln L = \sum_{i=1}^{k} \ln h(t_{i_F};\theta) - \sum_{l=1}^{n} H(t_l;\theta), \qquad /3/$$

логарифм правдоподібності виражено через інтенсивність відмов і риск.

Невідомі параметри θ знаходимо як розв"язок системи рівнянь правдоподібності

$$\frac{\partial \ln L}{\partial \theta_m} = 0, \quad (m = 1, \ldots, s), \qquad /4/$$

де $\ln L$ задається виразом /2/ або /3/.

Приклад. Нехай випадкова змінна ξ має функцію розподілу ймовірностей

$$F(t;\sigma,\nu) = 1 - e^{-\left(\frac{t}{\sigma}\right)^{\nu}}, \ t > 0, \quad (\sigma > 0, \ \nu > 0). \qquad /5/$$

Тоді

$$h(t;\sigma,\nu) = \frac{\nu t^{\nu-1}}{\sigma^{\nu}}, \qquad H(t;\sigma,\nu) = \left(\frac{t}{\sigma}\right)^{\nu}.$$

Звідси вираз /3/ набуває вигляду

$$\ln L = k\ln\nu + (\nu-1)\sum_{i=1}^{k} \ln t_{i_F} - k\nu\ln\sigma - \sum_{l=1}^{n}\left(\frac{t_l}{\sigma}\right)^{\nu}.$$

Отже, система рівнянь правдоподібності /4/

$$\begin{cases} \dfrac{\partial \ln L}{\partial \sigma} = -\dfrac{k\nu}{\sigma} + \dfrac{\nu}{\sigma^{\nu+1}}\sum\limits_{l=1}^{n} t_l^{\nu} = 0, \\ \dfrac{\partial \ln L}{\partial \nu} = \dfrac{k}{\nu} + \sum\limits_{i=1}^{k} \ln t_{i_F} - k\ln\sigma - \sum\limits_{l=1}^{n}\left(\dfrac{t_l}{\sigma}\right)^{\nu}\ln\dfrac{t_l}{\sigma} = 0. \end{cases} \qquad /6/$$

З першого рівняння /6/ виходить, що

$$\sigma^{\nu} = \frac{1}{k}\sum_{l=1}^{n} t_l^{\nu}. \qquad /7/$$

Тоді з другого дістаємо співвідношення для параметра

$$\frac{\nu}{k}\sum_{i=1}^{k} \ln t_{i_F} = \frac{\sum_{\ell=1}^{n} t_\ell^{\nu} \ln t_\ell^{\nu}}{\sum_{\ell=1}^{n} t_\ell^{\nu}} - 1. \qquad /8/$$

У випадку розподілу Релея з /7/ одержуємо оцінку

$$\sigma^2 = \frac{1}{k}\sum_{\ell=1}^{n} t_\ell^2 , \qquad /9/$$

а для експонентного

$$\sigma = \frac{1}{k}\sum_{\ell=1}^{n} t_\ell . \qquad /10/$$

Для ілюстрації дано напрацювання у відповідних одиницях $n = 15$ однотипних технічних пристроїв із популяції Релея до відмов F і зупинок S

20 S	23 F	36 F	46 F	55 F	60 S	60 S	60 S
66 F	77 F	88 F	90 S	100 S	110 F	140 F.	

Оцінити параметр масштабу σ.

За формулою /9/

$$\hat{\sigma}^2 = \frac{1}{9}\cdot 85995, \qquad \hat{\sigma} = 97{,}74968.$$

Контролюємо, чи можна прийняти $\nu = 2$. Ліва частина формули /8/ набуває значення

$$\frac{2}{9}\cdot 37{,}20789 = 8{,}268421,$$

а права – досить близьке значення

$$\frac{773705{,}7}{85995} - 1 = 7{,}997100.$$

Отже, залишається перевірити гіпотезу про те, що повнща вибірка взята з генеральної сукупності, керованої розподілом

$$F(t) = 1 - e^{-\left(\frac{t}{100}\right)^2}, \qquad t > 0.$$

2. Згруповані відмови та зупинки. Нехай з n однотипних технічних одиниць n_F працювали до відмови F, а решта $n_S = n - n_F$ до зупинки S, причому в момент t_{i_F} було n_{i_F} відмов, $/ i = 1, \ldots, k$ /, $n_{1_F} + \ldots + n_{k_F} = n_F$, а в момент t_{j_S} було n_{j_S} зупинок, $/ j = 1, \ldots, \ell$ /, $n_{1_S} + \ldots + n_{\ell_S} = n_S$, $n_F + n_S = n$.

У припущенні, що напрацювання до відмови керується розподілом $F(t;\theta)$, оцінити невідомі параметри $\theta=(\theta_1,\dots,\theta_s)$.

Тепер функція правдоподібності

$$L=\prod_{i=1}^{k}\left[f(t_{if};\theta)\right]^{n_{if}}\prod_{j=1}^{\ell}\left[R(t_{js};\theta)\right]^{n_{js}}.$$

Звідси

$$\ln L=\sum_{i=1}^{k}n_{if}\ln f(t_{if};\theta)+\sum_{j=1}^{\ell}n_{js}\ln R(t_{js};\theta), \qquad /11/$$

або, враховуючи /1/,

$$\ln L=\sum_{i=1}^{k}n_{if}\ln h(t_{if};\theta)-\sum_{j=1}^{k+\ell}n_j H(t_j;\theta), \qquad /12/$$

де t_j позначає t_{if} або t_{js}, а n_j позначає n_{if} або n_{js}. Невідомі параметри θ знаходимо як розв"язок системи рівнянь /4/, де $\ln L$ задається виразом /11/ або /12/.

Приклад. Якщо випадкова змінна ξ має функцію розподілу /5/, то вираз /12/ набуває вигляду

$$\ln L=n_f\ln\nu+(\nu-1)\sum_{i=1}^{k}n_{if}\ln t_{if}-n_f\nu\ln\sigma-\sum_{j=1}^{k+\ell}n_j\left(\frac{t_j}{\sigma}\right)^{\nu}.$$

Отже, система рівнянь правдоподібності /4/

$$\frac{\partial\ln L}{\partial\sigma}=-n_f\frac{\nu}{\sigma}+\frac{\nu}{\sigma^{\nu+1}}\sum_{j=1}^{k+\ell}n_j t_j^{\nu}=0,$$

$$\frac{\partial\ln L}{\partial\nu}=\frac{n_f}{\nu}+\sum_{i=1}^{k}n_{if}\ln t_{if}-n_f\ln\sigma\sum_{j=1}^{k+\ell}n_j\left(\frac{t_j}{\sigma}\right)^{\nu}\ln\frac{t_j}{\sigma}. \qquad /13/$$

З першого рівняння /13/ випливає, що

$$\sigma^{\nu}=\frac{1}{n_f}\sum_{j=1}^{k+\ell}n_j t_j^{\nu}, \qquad /14/$$

а тоді з другого

$$\frac{\nu}{n_f}\sum_{i=1}^{k}n_{if}\ln t_{if}=\frac{\sum_{j=1}^{k+\ell}n_j t_j^{\nu}\ln t_j^{\nu}}{\sum_{j=1}^{k+\ell}n_j t_j^{\nu}}-1. \qquad /15/$$

Для розподілу Релея з /14/ дістаємо оцінку

$$\sigma^2=\frac{1}{n_f}\sum_{j=1}^{k+\ell}n_j t_j^{2}, \qquad /16/$$

а для експонентного

$$\sigma = \frac{1}{n_F} \sum_{j=1}^{k+l} n_j t_j . \qquad /I7/$$

Проілюструємо наш виклад даними про згруповані напрацювання до відмов і зупинок у відповідних одиницях n = I00 однотипних технічних пристроїв з експонентного розподілу

i	t_{i_F}	n_{i_F}	j	t_{j_S}	n_{j_S}
			I	I7,5	I5
I	35	25			
			2	52,5	9
2	70	I5			
			3	87,5	4
3	I05	9			
			4	I22,5	3
$k = 4$	I40	7			
			$l = 5$	I57,5	I3
	n_F = 56			n_S = 44	

За формулою /I7/

$$\hat{\sigma} = \frac{I}{56} /I5 \cdot I7,5 + 25 \cdot 35 + \ldots + I3 \cdot I57,5/ = \frac{7350}{56} = I3I,25.$$

Ліва частина формули /I5/ при ν = I набуває значення

$$\frac{I}{56} \cdot 229,0882 = 4,09086I,$$

а права

$$\frac{33I26,68}{7350} - I = 3,50703I.$$

Отже, залишається перевірити гіпотезу про те, що дана вибірка згрупованих даних взята з генеральної сукупності керованої розподілом

$$F(t) = 1 - e^{-\frac{t}{130}}, \qquad t > 0.$$

Зазначимо, що при $n_{1F} = \ldots = n_{kF} = n_{1S} = \ldots = n_{lS} = 1$, $k + l = n$, формули /II/–/I7/ переходять відповідно в /2/, /3/, /6/ –/I0/.

Стаття надійшла до редколегії 05.I0.89