1

lecture (22)
Mathematical statistics
Ch 11 p. 236

Sec 11.4 Truncated or Censored Data

Defn (11.2) : right censoring
An observation is censored from above (also called right censored) at u if when it is at or above u it is recorded as being equal to u, but when it is below u it is recorded at its observed value.

Defn (11.3) : truncation
An observation is truncated from below (also called left truncated) at d if when it is at or below d it is not recorded, but when it is above d it is recorded at its observed value.

* For truncated data, there are two ways to proceed.

shifting approach　　　unshifted approach

(Ex) 11.4 p 237
Assume that the values in Data Set B had been truncated from below at 200. Using both methods, estimate the value of α for a Pareto distribution with $\theta = 800$ known. Then use the model to estimate the cost per payment with deductibles of 0, 200, and 400.

Ans: Data Set B {27 82 115 126 155 161 243 294 340 384
457 680 855 877 974 1,193 1,340 1,884 2,558 15,743}

(I) shifting approach
Using the shifting approach, the values become

After truncation from below at 200

⇒ 43 94 140 184 257 480 655 677
774 993 1140 1684 2358 15543

Because the data have been shifted, it is not possible to estimate the cost with no deductible ($d = 0$).

لا يكتب في
هذا الهامش

2

$$L(\alpha) = \prod_{j=1}^{n} f_{X_j}(x_j|\alpha)$$

For pareto α,θ p.d.f is $f(x) = \frac{\alpha\theta^{\alpha}}{(x+\theta)^{\alpha+1}}$

See p. 494

$$\therefore L(\alpha) = \prod_{j=1}^{14} \frac{\alpha(800^{\alpha})}{(800+x_j)^{\alpha+1}}$$

$$\therefore l(\alpha) = \sum_{j=1}^{14} \left[\ln\alpha + \alpha\ln 800 - (\alpha+1)\ln(x_j+800)\right]$$

$$\therefore l(\alpha) = 14\ln\alpha + 93.5846\alpha - 103.969(\alpha+1)$$

where $14\ln 800 = 93.5846$, $\sum_{j=1}^{14}\ln(x_j+800) = 103.969$

$$\Rightarrow l(\alpha) = 14\ln\alpha - 10.384\alpha - 103.969$$

For pareto α,θ cdf is $F(x) = 1-\left(\frac{\theta}{x+\theta}\right)^{\alpha}$

To get $\hat{\alpha}$, let $l'(\alpha) = 0$

$$\Rightarrow l'(\alpha) = \frac{14}{\alpha} - 10.384$$

$$\therefore \hat{\alpha} = \frac{14}{10.384} = 1.3482$$

For pareto α,θ $E(X^k) = \frac{\theta^k k!}{(\alpha-1)\dots(\alpha-k)}$, k is a +ve integer

See p. 494

* The cost per payment with deductible $d=200$ is the expected value of the estimated pareto distn. $= E(X) = \frac{\theta}{\alpha-1}$

$$= \frac{800}{1.3482-1} = \frac{800}{0.3482} \simeq 2298$$

* The cost per payment with deductible $d=400$, can be calculated by the following formula

$$\frac{E(X) - E(X\wedge 200)}{1-F(200)}$$

Revise theorem 8.3 p.130

For pareto α,θ, the limited expected value $E(X\wedge x) = \frac{\theta}{\alpha-1}\left[1-\left(\frac{\theta}{x+\theta}\right)^{\alpha-1}\right]$, $\alpha \neq 1$

See p. 494

$$= \frac{\cancel{\frac{800}{0.3482}} - \frac{800}{0.3482}\left[\cancel{1} - \left(\frac{800}{200+800}\right)^{0.3482}\right]}{\left(\frac{800}{200+800}\right)^{1.3482}}$$

$$= 1000/0.3482 \simeq 2,872$$

لا يكتب في هذا الهامش

(II) Unshifted approach

The probability of observing each data value knowing that values under 200 are omitted from the data set

This conditional probability leads to

$$\Rightarrow L(\alpha) = \prod_{j=1}^{14} \frac{f(x_j|\alpha)}{1-F(200|\alpha)}$$

For pareto dist^n

$$L(\alpha) = \prod_{j=1}^{14}\left[\frac{\alpha(\cancel{800^{\alpha}})}{(800+x_j)^{\alpha+1}} \Big/ \left(\frac{\cancel{800}}{800+200}\right)^{\alpha}\right]$$

$$L(\alpha) = \prod_{j=1}^{14} \frac{\alpha\,(1000)^{\alpha}}{(800+x_j)^{\alpha+1}}$$

$$\therefore \ell(\alpha) = \sum_{j=1}^{14} \log\left[\frac{\alpha\,(1000)^{\alpha}}{(800+x_j)^{\alpha+1}}\right]$$

$$\therefore \ell(\alpha) = 14\ln\alpha + 14\alpha\ln(1000) - (\alpha+1)\sum_{j=1}^{14}\ln(800+x_j)$$

$\because 14\ln 1000 = 96.709$

$\because \sum_{j=1}^{14} \ln(800+x_j) = 105.810$

where x_j: 243, 294, 340, ..., 15743

$$\ell(\alpha) = 14\ln\alpha + 96.709\alpha - (\alpha+1)\,105.810$$

$$\therefore \ell'(\alpha) = \frac{14}{\alpha} - 9.101 = 0 \Rightarrow \hat{\alpha} = \frac{14}{9.101} = 1.5383$$

This model is used for losses with no deductible and the expected payment without a deductible is

$$E(X) = \frac{\theta}{\alpha-1} = \frac{800}{0.5383} = 1486$$

4/

The expected payment with a deductible $d = 200$ (As before)

$$= \frac{E(X) - E(X \wedge 200)}{1 - F(200)}$$

$$= \frac{\cancel{\frac{800}{0.5383}} - \frac{800}{0.5383}\left[\cancel{1} - \left(\frac{800}{200+800}\right)^{0.5383}\right]}{\left(\frac{800}{200+800}\right)^{1.5383}}$$

$$= \frac{1,000}{0.5383} = 1,858$$

Similarly

The expected payment with a deductible $d = 400$

$$= \frac{E(X) - E(X \wedge 400)}{1 - F(400)}$$

$$= \frac{1200}{0.5383} = 2,229$$

#