Lecture (11)

1//

• 4.2.4 Data – Dependent Distributions p. (55)

Def$\underline{n}$ (1) p. 55

A data – dependent distribution is at least as complex as the data or knowledge that produced it and the number of "parameters" increases as the number of data points or amount of knowledge increases.

Def$\underline{n}$ (2) p. 55

The empirical model is a discrete distribution based on a sample of size n that assigns probability $1/n$ to each data point.

Example 4.7 p. 55

Consider a sample of size 8 in which the observed data points were 3, 5, 6, 6, 6, 7, 7, and 10. The empirical model then has probability function

$$p(x) = \begin{cases} 0.125, & x=3 \\ 0.125, & x=5 \\ 0.375, & x=6 \\ 0.25, & x=7 \\ 0.125, & x=10 \end{cases}$$

Note that:

(1) The empirical model is a data – dependent dist$\underline{n}$. Each data point contributes probability $1/n$ to the probability function, so the n parameters are the n observations in the data set that produced the empirical dist$\underline{n}$.

(2) Clearly, the empirical dist$\underline{n}$ is a discrete model with finite support, but we use the term "empirical model" only when there is an actual data sample behind it.

Pb 4.8 p. 57

Let X have a Burr distribution with parameters $\alpha=1$, $\gamma=2$ and $\theta=\sqrt{1,000}$ and let Y have a pareto distn with parameters $\alpha=1$ and $\theta=1000$. Let Z be a mixture of X and Y with equal weight on each Component. Determine the median of Z. Let $W=1.1Z$. Demonstrate that W is also a mixture of a Burr and a pareto distribution and determine the parameters of W.

Ans: $\because X \sim$ Burr $- \alpha=1, \gamma=2, \theta=\sqrt{1000}$

$$\therefore F(x) = 1-u^{\alpha}, \quad u=\frac{1}{1+(x/\theta)^{\gamma}}$$

$$= 1-\frac{1}{1+(x/\sqrt{1000})^2}$$

& $\because Y \sim$ Pareto $- \alpha=1$ and $\theta=1,000$

$$\therefore F(y) = 1-\left(\frac{\theta}{y+\theta}\right)^{\alpha} = 1-\left(\frac{1}{1+y/\theta}\right)^{\alpha}$$

$$= 1-\frac{1}{1+y/1,000}$$

$\Rightarrow$ For mixture distn,

$$F_Z(z) = 0.5\left[1-\frac{1}{1+(z/\sqrt{1,000})^2}\right]+0.5\left[1-\frac{1}{1+z/1,000}\right] \quad (1)$$

$$F(z) = 1-0.5\,\frac{1,000}{1000+z^2}-0.5\,\frac{1000}{1000+z}$$

$$F(z) = 1-\frac{0.5\left[1,000^2+1,000z+1,000^2+1,000z^2\right]}{(1,000+z^2)(1,000+z)}$$

$$\therefore F(z) = 1-\frac{0.5\left[2(1,000)^2+1,000z+1,000z^2\right]}{(1,000+z^2)(1,000+z)} \quad (2)$$

3/ the median m is the solution of the equation

$$F(m) = 0.5 \quad ③$$

Apply ③ in ②, we get

$$\frac{\cancel{0.5}\,[2(1{,}000)^2 + 1{,}000m + 1{,}000m^2]}{(1{,}000+m^2)(1{,}000+m)} = \cancel{0.5}\;1$$

$$\Rightarrow (1{,}000+m^2)(1{,}000+m) = 2(1{,}000)^2 + 1{,}000m + 1{,}000m^2$$

$$\therefore 1{,}000^2 + \cancel{1{,}000m^2} + \cancel{1{,}000m} + m^3 = 2(1{,}000)^2 + \cancel{1{,}000m} + \cancel{1{,}000m^2}$$

$$\therefore m^3 = 1{,}000^2$$

$$m^3 = 10^6$$

$\therefore m = 10^2 = 100$ which is the median.

* For $W = 1.1Z$

$$F_W(w) = pr(W \le w) = pr(1.1Z \le w) = Pr\left(Z \le \frac{w}{1.1}\right)$$

$$= F_Z(w/1.1)$$

$$① \Rightarrow F_W(w) = 0.5\left[1 - \frac{1}{1+(w/1.1\sqrt{1000})^2}\right] + 0.5\left[1 - \frac{1}{1+w/1{,}100}\right]$$

which is a 50/50 mixture of a Burr with parameter $\alpha = 1$, $\gamma = 2$, and $\theta = 1.1\sqrt{1{,}000}$ and a pareto distribution with parameters $\alpha = 1$ and $\theta = 1{,}100$.

#