Assignment 6

Varenya Upadhyaya EP20BTECH11026

Download all python codes from:

https://github.com/varenya27/AI1103/blob/main/Assignment6/codes

and all latex-tikz codes from:

https://github.com/varenya27/AI1103/blob/main/Assignment6/main.tex

1 Problem

Let Y_1 denote the first order statistic in a random sample of size n from a distribution that has the pdf,

$$f(x) = \begin{cases} e^{-(x-\theta)} & \text{when } \theta < x < \infty \\ 0 & \text{otherwise} \end{cases}$$

Obtain the distribution of $Z_n = n(Y_1 - \theta)$

2 Solution

The first order statistic for any sample is the the minimum of the given sample In order to calculate the distribution for Y_1, we need the the cumulative distribution function $F(x)$:

$$F(x) = \int_{-\infty}^{x} f(t)\,dt \tag{2.0.1}$$

$$= \int_{-\infty}^{\theta} 0\,dt + \int_{\theta}^{x} e^{\theta-t}\,dt \tag{2.0.2}$$

$$= \left(e^{\theta-t}\right)_x^{\theta} \tag{2.0.3}$$

$$= 1 - e^{\theta-x} \tag{2.0.4}$$

$$= 1 - f(x) \tag{2.0.5}$$

This gives,

$$F(x) = \begin{cases} 1 - e^{-(x-\theta)} & \text{when } \theta < x < \infty \\ 0 & \text{otherwise} \end{cases} \tag{2.0.6}$$

Now,

$$F(x) = \Pr(X \leq x) \tag{2.0.7}$$

$$\implies 1 - F(x) = \Pr(X > x) \tag{2.0.8}$$

In the random sample of size n, Y_1 can take n values. For $Y_1 = x$, all the other terms must be greater than x. Let $p_{Y_1}(y)$ and F_{Y_1} be the pdf and cdf for Y_1 respectively:

$$Y_1 = min\{X_1, X_2 ..., X_n\} \tag{2.0.9}$$

Calculating the cdf of Y_1:

$$F_{Y_1}(y) = \Pr(Y_1 \leq y) \tag{2.0.10}$$

$$= 1 - \Pr(Y_1 > y) \tag{2.0.11}$$

Since Y_1 is the minimum order statistic,

$$F_{Y_1}(y) = 1 - \Pr(X_1 > y, X_2 > y, ..., X_n > y)$$

$$= 1 - \prod_{i=1}^{n} \Pr(X_i > y) \tag{2.0.12}$$

$$= 1 - (1 - F(y))^n \tag{2.0.13}$$

Using the cdf, the pdf can be calculated as:

$$p_{Y_1}(y) = \frac{d}{dy} F_{Y_1}(y) \tag{2.0.14}$$

$$= \frac{d}{dy}(1 - (1 - F(y))^n) \tag{2.0.15}$$

$$= n \times (1 - F(y))^{n-1} \times \frac{d}{dy} F(y) \tag{2.0.16}$$

$$= n \times (f(y))^{n-1} \times f(y) \tag{2.0.17}$$

$$= n(f(y))^n \tag{2.0.18}$$

Thus,

$$p_{Y_1}(y) = \begin{cases} n(e^{-n(y-\theta)}) & \text{when } \theta < y < \infty \\ 0 & \text{otherwise} \end{cases} \tag{2.0.19}$$

Fig. 0 and Fig. 0 are the plots for the function in (2.0.19) when $\theta = 20$. It can be seen that as n approaches infinity, the function $p_{Y_1}(x)$ attains a spike at $x = \theta$

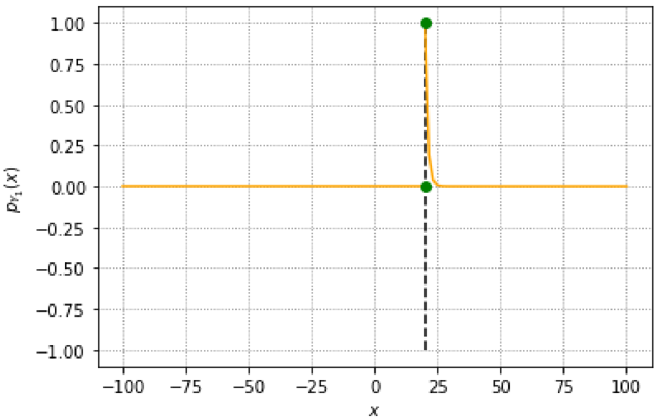

Fig. 0: Distribution when $n = 1$

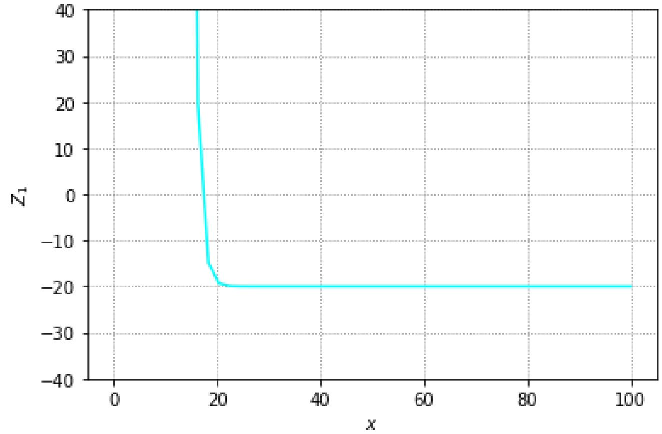

Fig. 0: Z_n when $n = 1$

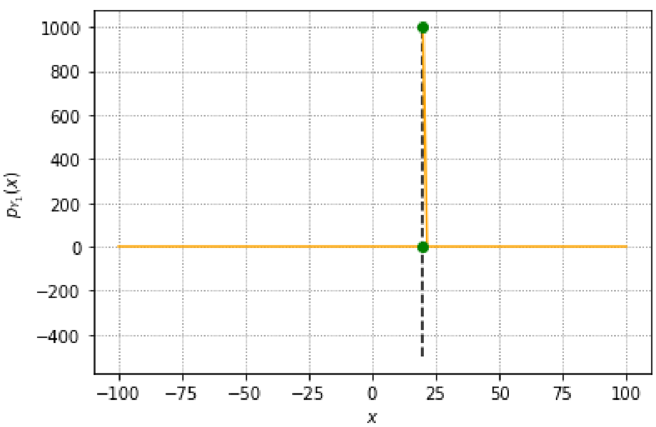

Fig. 0: Distribution as n increases to 1000

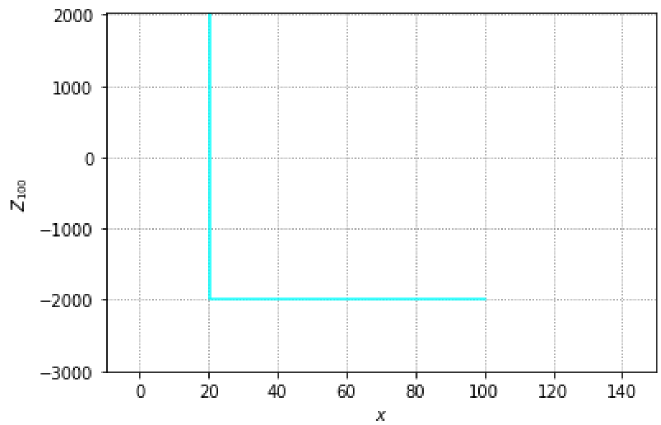

Fig. 0: Z_n when $n = 100$

The required distribution Z_n is:

$$Z_n = n(Y_1 - \theta) \tag{2.0.20}$$

$$= \begin{cases} n(n(e^{-n(y-\theta)}) - \theta) & \text{when } \theta < y < \infty \\ 0 & \text{otherwise} \end{cases} \tag{2.0.21}$$

Fig. 0 and Fig. 0 are the plots for the distribution in (2.0.21) for $\theta = 20$