Sehldon ross a first course in probability 10th ed 2018 page 355/848

5.6.5 The Pareto Distribution

If X is an exponential random variable with rate λ and $a > 0$, then

$$Y = ae^X$$

is said to be a *Pareto* random variable with parameters a and λ. The parameter $\lambda > 0$ is called the index parameter, and a is called the minimum parameter (because $P\{Y > a\} = 1$). The distribution function of Y is derived as follows: For $y \geq a$,

$$\begin{aligned} P(Y > y) &= P(ae^X > y) \\ &= P(e^X > y/a) \\ &= P(X > \log(y/a)) \\ &= e^{-\log(y/a)} \\ &= e^{-\log((y/a)^\lambda)} \\ &= (a/y)^\lambda \end{aligned}$$

Hence, the distribution function of Y is

$$F_Y(y) = 1 - P(Y > y) = 1 - a^\lambda y^{-\lambda}, \quad y \geq a$$

Differentiating the distribution function yields the density function of Y:

$$f_Y(y) = \lambda a^\lambda y^{-(\lambda+1)}, \quad y \geq a$$

When $\lambda \leq 1$ it is easily checked that $E[Y] = \infty$. When $\lambda > 1$,

$$\begin{aligned} E[Y] &= \int_a^\infty \lambda a^\lambda y^{-\lambda}\, dy \\ &= \lambda a^\lambda \frac{y^{1-\lambda}}{1-\lambda}\bigg|_a^\infty \\ &= \lambda a^\lambda \frac{a^{1-\lambda}}{\lambda-1} \\ &= \frac{\lambda a}{\lambda-1} \end{aligned}$$

$E[Y^2]$ will be finite only when $\lambda > 2$. In this case,

$$\begin{aligned} E[Y^2] &= \int_a^\infty \lambda a^\lambda y^{1-\lambda}\, dy \\ &= \lambda a^\lambda \frac{y^{2-\lambda}}{2-\lambda}\bigg|_a^\infty \\ &= \frac{\lambda a^2}{\lambda - 2} \end{aligned}$$

Hence, when $\lambda > 2$

$$\mathrm{Var}(Y) = \frac{\lambda a^2}{\lambda - 2} - \frac{2^\lambda a^2}{(\lambda-1)^2} = \frac{\lambda a^2}{(\lambda-2)(\lambda-1)^2}$$

Remarks (a) We could also have derived the moments of Y by using the representation $Y = ae^X$, where X is exponential with rate λ . This yields, for $\lambda > n$,

$\lambda e^{-\lambda x}$ $\lambda e^{-(-n+\lambda)x}$

$\dfrac{\lambda a^n}{\lambda - n}$

$$E[Y^n] = a^n E[e^{nX}] = a^n \int_0^\infty e^{nx} e^{-x}\, dx = a^n \int_0^\infty e^{-(-n)x}\, dx = \frac{a^n}{-n}$$

(b) Where the density function $f(y)$ of the Pareto is positive (that is, when $y > a$) it is a constant times a power of y, and for this reason it is called a *power law density*.

(c) The Pareto distribution has been found to be useful in applications relating to such things as

i. the income or wealth of members of a population;
ii. the file size of internet traffic (under the TCP protocol);
iii. the time to compete a job assigned to a supercomputer;
iv. the size of a meteorite;
v. the yearly maximum one day rainfalls in different regions.

Further properties of the Pareto distribution will be developed in later chapters.