Random Variables

Exploring Relationships between Mean and Variance

Arka Mukhopadhyay   Ayuj Aryan

March 19, 2024

1 When Mean and Variance both do not exist:

1. **Standard Cauchy Distribution**
   Cauchy distribution has a Probability Density Function:

   $$f(x) = \frac{1}{\pi(1 + x^2)}$$

   where $-\infty < x < \infty$.
   It has an undefined mean and an undefined variance. The tails of the Cauchy distribution decrease very slowly, leading to infinite moments. Mathematically, the tails of the Cauchy distribution decay as slowly as $\frac{1}{x^2}$ which causes the integrals defining the mean and variance to diverge. Therefore, both the mean and variance of the Cauchy distribution do not exist.

   MEAN: The mean of Cauchy distribution, calculated by integrating $x \cdot P(x)$ over $-\infty$ to $+\infty$, turns out to be $\ln\left(\frac{1}{x^2}\right)$ without putting the limits. Clearly, when we put the limits, the integral will tend to infinity, which implies that it diverges, hence the mean does not exist

   VARIANCE: The variance of Cauchy distribution involves a term containing $\ln\left(\frac{1}{x^2}\right)$ and a term containing $\arctan x$, both of which diverge to ∞. Therefore, the variance of the standard Cauchy distribution also does not exist.

2. **Pareto Distribution with Shape Parameter (a) $\in$ (0,1]:**
   A Pareto distribution has a Probability Density Function:

   $$f(x) = 1 - \frac{1}{x^a}$$

where $'a'$ is a positive real constant and is called the shape parameter.
Typically, the value of a decides how slowly the tail of the distribution decays, which eventually decides where mean and variance exist or not.

MEAN: The value of mean of Pareto distribution comes out to be $\frac{a}{a-1}$. Clearly a has to be greater than 1 for the mean to exist. If $a = 1$, denominator is equal to 0 which is not possible whereas for $a < 1$, the tail of distribution decays very slowly and function diverges.

VARIANCE: The value of variance of Pareto distribution comes out to be $\frac{a}{(a-1)^2(a-2)}$. If $a < 2$, then variance will be negative, which is not possible since variance is a positive value. Therefore, variance exists only for $a > 2$. Therefore, in the interval (0,1], both variance and mean do not exist.

3. **Levy Distribution** The Probability Distribution Function of levy distribution is given by:

$$f(x) = \sqrt{\frac{c}{2\pi}} \frac{\exp\left(\frac{-c}{2(x-\mu)}\right)}{(a-\mu)^{\frac{3}{2}}}$$

where μ is location parameter (mean) and $c > 0$ is scale parameter. It has a heavy tail, i.e. the rate of decay is such that at infinity the series diverges, hence both the mean and the variance do not exist.

MEAN: The mean of the distribution is the value of integral $x \cdot \sqrt{\frac{c}{2\pi}} \frac{\exp\left(\frac{-c}{2(x-\mu)}\right)}{(a-\mu)^{\frac{3}{2}}}$ from $-\infty$ to ∞, which tends to ∞, therefore the mean is not defined.

VARIANCE: The variance of the distribution is the value of the integral $x^2 \cdot \sqrt{\frac{c}{2\pi}} \frac{\exp\left(\frac{-c}{2(x-\mu)}\right)}{(a-\mu)^{\frac{3}{2}}}$ from $-\infty$ to ∞, which again tends to ∞,, therefore the variance is not defined as well.

2 WHEN MEAN EXISTS BUT VARIANCE DOES NOT:

1. **Pareto Distribution with Shape Parameter (a) $\in$ (1,2]:**

A Pareto distribution has a Probability Density Function:

$$f(x) = 1 - \frac{1}{x^a}$$

where $'a'$ is a positive real constant and is called the shape parameter.
Typically, the value of a decides how slowly the tail of the distribution decays, which eventually decides where mean and variance exist or not.

MEAN: The value of mean of Pareto distribution comes out to be $\frac{a}{a-1}$. Clearly a has to be greater than 1 for the mean to exist. If $a = 1$, denominator is equal to 0 which is not possible whereas for $a < 1$, the tail of distribution decays very slowly.

VARIANCE: The value of variance of Pareto distribution comes out to be $\frac{a}{(a-1)^2(a-2)}$. If $a < 2$, then variance will be negative, which is not possible since variance is a positive value. Therefore, variance exists only for $a > 2$.
Therefore, in the interval (1,2], variance does not exist whereas mean does exist.

2. **Hypothetical Case**

   Let's consider a probability mass function:

   $$f(x) = \frac{k}{x^3}$$

   where $k = \zeta(x)$ which is referred to as Apéry's constant.

   MEAN: The mean of the distribution is $\sum_{-\infty}^{\infty} x \cdot \frac{k}{x^3}$ which converges to a real number.

   VARIANCE: The variance of the distribution is $\sum_{-\infty}^{\infty} x^2 \cdot \frac{k}{x^3}$ which diverges to ∞. Hence, mean exists while variance does not.

3 When Mean does not exist but Variance does:

There is no such distribution (that we found!) in which this criteria holds. We can try to justify this from the fact that:

$$Var(X) = E[X^2] - (E[X])^2$$

If Var(X) exists and is finite, it follows that E[X] and hence the mean must exist and be finite.

4 References

1. Cauchy Distribution
2. Pareto Distribution
3. Levy Distribution