Lecture - 6

Instructor: Dr. Arabin Kumar Dey

1 Heavy tail Distributions :

Distribution with high tail probabilities compared to a reference distribution (say normal) with same mean and standard deviation are called heavy-tailed.

Since kurtosis is more sensitive to tail-weight, high kurtosis is nearly with having heavy tails. A heavy tailed distribution is more prove to extreme values, which are sometime called outliers.

In financial application one is especially concerned when the return distribution has heavy tails because of the possibility of an extremely large negative return which could entirely deplete the capital reserves of a firm.

Examples :

Double exponential distribution Double exponential has slightly heavier tails than normal distributions. This fact can be appreciates by comparing their densities, the density of double exponential with scale parameter θ is proportional to $\exp(-|\frac{x}{\theta}|)$ and the density of the $N(0, \sigma^2)$ distribution is proportional to $\exp(-0.5(\frac{x}{\sigma})^2)$. The term $-x^2$ converges to $-\infty$ much faster than $-|x|$ as $|x| \longrightarrow \infty$. Therefore the normal density converges to 0 much faster than the double exponential density as $x \longrightarrow \infty$. In fact, the density that is proportional to

$$\exp\big(-\Big|\frac{x}{\theta}\Big|^{\alpha}\big) \tag{1}$$

where $0 < \alpha < 2$ is a shape parameter and θ scale parameter, will have heavier tails than a normal distribution, however no density of the form (1) will have truly heavy tails, and in particular $E|X|^n < \infty$ for all n. no matter how large, and in addition there is finite mean and variance

To achieve very heavy tails, the density must behave like $|x|^{-(a+1)}$ for some $a > 0$, which we call polynomial tails, rather than like (1), which we call exponential tails. The parameter of polynomial tail is called tail index.

Pareto Distribution : Swiss economics professor *Vilfredo Pareto (1848 - 1923)* formulates an eponymous law which states that the function of a population with income exceeding an amount x is equal to

$$cx^{-a} \qquad for \quad all \quad x \tag{2}$$

Where c and a are positive constant independent of x but depending on the population. If $F(x)$ is the CDF of the income distribution, then (2) implies that $F(x) = 1 - \left(\frac{c}{x}\right)^a; \quad x > c$. where, c is the minimum income.

Therefore the p.d.f. corresponding to above type is

$$f(x) = \frac{ac^a}{x^{a+1}}, \quad x > c$$

So a Pareto distribution has polynomial tails. As mentioned before constant a is called the tail-index. It is also called the Pareto constant

Distribution with Pareto tail : A cumulative distribution function $F(x)$ is said to have **Pareto right tail** if its survival function satisfies

$$(1 - F(x)) = L(x)x^{-a}$$

For some $a > 0$ where $L(x)$ is slowly varying function.

A function is said to be slowly varying at ∞ if

$$\frac{L(tx)}{L(x)} \longrightarrow 1 \quad as \;\; x \longrightarrow \infty \;\; \text{for} \quad t > 0.$$

For example $\log x$ is slowly varying as

$$\frac{\log(tx)}{\log(x)} = \frac{\log t + \log x}{\log x} \longrightarrow 1 \quad x \longrightarrow \infty \;\; \text{for} \quad t > 0.$$

t-distribution have heavy tails :

t-distribution's density is proportional to

$$\frac{1}{\left[1+\left(\frac{x^2}{\nu}\right)\right]^{\frac{\nu+1}{2}}}$$

which for large values of $|x|$ is approximately

$$\frac{1}{\left[1+\left(\frac{x^2}{\nu}\right)\right]^{\frac{\nu+1}{2}}} \quad \propto \quad \frac{1}{|x|^{-(\nu+1)}}$$

Therefore, t-distribution has polynomial tails with $a = \nu$. Any distribution with polynomial tails has heavy tails, and the smallest value of $'\nu'$ the heavier the tails. From modeling perspective, a problem with the t-distribution is that the tail index is **integer valued**, rather than **a continuous parameter**, which limits the flexibility of the t-distribution as a model for financial market data.

Mixture Models :

Another class of heavy tailed distribution is mixture models.

Consider a distribution which is 90% $N(0, 1)$ and 10% $N(0, 25)$. this is an example of a normal mixture distribution since it is a mixture of two different normal distributions called the components. The variance of this distribution is $(0.9)(1) + (0.1)(25) = 3.4$ so its standard deviation is $\sqrt{3.4} = 1.84$ This distribution is much different than an $N(0, 3.4)$ distribution, even though both distributions have same mean and variance.

In general p.d.f. can be written as

$$f(x) = pf_1(x;\theta_1) + (1-p)f_2(x;\theta_2)$$

Where $f_1(x;\theta_1)$ and $f_2(x;\theta_2)$ are two different p.d.f. let $F_1(x;\theta_1)$ and $F_2(x;\theta_2)$ are two corresponding c.d.f's.

Then $F(x) = pF_1(x;\theta_1) + (1-p)F_2(x;\theta_2)$.

Now we explore how much more probability the normal mixture distribution has in the outlier range $|x| > 6$ compared to normal distribution.

For a $N(0,\sigma^2)$ random variable X,

$$P\Big[|X| > x\Big] = 2\Big[1 - \Phi\Big(\frac{x}{\sigma}\Big)\Big]$$

therefore, for the normal distribution with variance 3.4,

$$P_N\Big[|X| > 6\Big] = 2\Big[1 - \Phi\Big(\frac{6}{\sqrt{3.4}}\Big)\Big] = 0.0011$$

For the normal mixture population which has variance 1 with probability 0.9 and variance 25 with probability 0.1 we have that

$$P_{MN}\Big[|X| > 6\Big] = 2\Big[0.9(1-\Phi(6)) + 0.1(1-\Phi(6/5))\Big] = 2[0.9(0) + (0.1)(0.115)] = 0.023$$

Since, $\frac{0.023}{0.0011} \approx 21$, the normal mixture distribution is 21 times more likely to be in this outlier range than the $N(0,3.4)$ population, even though both have a variance of 3.4.

It is not difficult to compute the Kurtosis of this mixture.

If $Z \sim N(\mu,\sigma^2)$, then $E(Z-\mu)^4 = 3\sigma^4$. Therefore if X has normal mixture distribution. then

$$E\Big(X^4\Big) = 3(0.9 + (0.1)25^2) = 190.2$$

and Kurtosis of X is $\frac{190.2}{3.4^2} = 16.453$, which is clearly very high indicating the heavy tail behavior.

Despite the heavy tails of a normal mixture, the tails are exponential, not polynomial.