The Controversy over Tailweight of Distributions and a Robust Measure of Tailweight

C. C. Heyde[1] and S. G. Kou[2]

[1] Department of Statistics, Columbia University, New York, NY 10027, USA
[2] Department of Industrial Engineering and Operations Research, Columbia University, New York, NY 10027, USA

Keywords: heavy tails, tail probability plots, Hill ratio plots, mean excess function, finance.

1 Introduction

The behavior of most probabilistic models is critically influenced by the tails of the distribution(s) which drive the models. This is especially important in such areas as asset pricing, Value-at-Risk, reliability, insurance risk, queueing theory and telecommunication network analysis. The literature, however, reveals considerable uncertainty, controversy, and error in connection with the distinctions between tail weights. In this paper we will explain why the distinctions are surprisingly difficult, why particularly large samples, perhaps in the tens of thousands or even hundreds of thousands, are necessary for clear discrimination and why theoretical errors have confused the discrimination. Then we shall propose a robust way to measure tailweight.

The idea that there is a true statistical model for any real data set is a (convenient) theoretical idealization. D.R. Cox has usefully pointed to the role and purpose of statistical models being to provide a "concise description of the aspects of the data judged relevant for interpretation". We should not expect more than a parsimonius description. Indeed, any assumed distributional form is likely to be rejected in a test of adequacy when one has a (sufficiently) large data set, this usually being a reflection of departures, which may be quite small, in the middle part of the distribution. Perhaps even 5,000 observations may be sufficient for rejection, and in some contexts, such as telecommunications, 50,000 observations may be routinely available. Introducing additional model parameters, up to three or four, may appear to help a little, but does not resolve the intrinsic difficulty. In fact, if a true distribution is assumed, it is known to be impossible to consistently estimate its density in the extreme tail.

Therefore, it seems to us to be more appropriate to ask what type of distribution is suitable for the data than to seek some best-fitting distribution; and we will make this issue the focus of the paper. This apparently straightforward distinction is surprisingly difficult to make, and requires a substantially larger sample than what is intuitively expected to achieve good statistical precision. It turns out that the distinction rests on an extremely small percentage of the sample, perhaps less than 0.01%!

First we need to be clear about the terminology, because different specialist areas have established somewhat different norms. For example, what may be described as a light tail in the queueing literature could be called a semi-heavy tail in the finance literature. In this paper we will settle on the simple dichotomy of calling tails light or heavy according to whether or not a finite moment generating function exists.

Heavyness of tails is a subject of controversy in some areas, but especially in finance. For example it is accepted that real returns data has tails which are heavier than those of the normal distribution, but one school of thought believes the tails to be light (typically exponential) and the other believes that they are heavy (typically powers). Some authors conclude that both tail types are required. Differences of opinion also appear in studies on World Wide Web traffic.

prob.	Laplace	t7	t6	t5	t4	t3
1%	**2.77**	2.53	2.57	2.61	2.65	2.62
0.1%	**4.39**	4.04	4.25	**4.57**	5.07	5.90
0.01%	**6.02**	5.97	**6.55**	7.50	9.22	12.82
0.001%	**7.65**	**8.54**	9.82	12.04	16.50	27.67

TABLE 1. The (right) quantiles for the Laplace and normalized t densities.

prob.	Laplace	t7	t6	t5	t4	t3
1%	**2.82**	2.11	2.19	2.31	2.53	**2.97**
0.1%	**4.48**	3.37	3.63	4.02	**4.81**	6.66
0.01%	**6.14**	4.93	5.56	**6.63**	8.78	14.51
0.001%	**7.80**	7.08	**8.36**	10.68	15.74	31.33

TABLE 2. The (right) quantiles for the Laplace and t densities normalized by their interquartile ranges.

2 Motivation from Standard Methods of Discrimination and Estimation

We shall begin by mentioning six of the most widely used methods, together with their strengths and weaknesses. The methods are (1) the tail-probability plot; (2) the mean excess function method; (3) the moment generating function method; (4) the *max-sum ratio plot*; (5) the *generalized Hill ratio plot*; (6) the likelihood method. These methods, and variants of them, essentially span the spectrum of standard theory. The first five are discussed in considerable detail in the literature, but insufficient attention has been given to their intrinsic capacity to discriminate. The sixth method, although widely used in practice, has a serious theoretical flaw, and should not be employed except under very special circumstances.

After showing that the arguably most popular methods may fail to distinguish the difference between power type tails and exponential type tails for samples of size 5000, one may reasonably ask how large the sample size should be in order to make the distinction with appropriate confidence.

A reasonable impression may be obtained by simply looking at the quantile tables for both standardized Laplace and standardized t distributions. The quantiles for the Laplace and normalized t densities are given in Table 1, which shows that the Laplace distribution may have higher tail probabilities than those of t distributions with low degrees of freedom. For example, the 99.9% quantile of the Laplace distribution is actually larger than that of t distribution with d.f. 6 and 7! Thus, regardless of the sample size, the Laplace distribution may appear to be heavier tailed than a t-distribution with d.f. 6 or 7, up to the 99.9% quantile. In order to distinguish the distributions it is necessary to use quantiles with very low p values and correspondingly large samples, typically in the tens of thousands or even hundreds of thousands, as revealed in Table 1.

It may be thought that standardization using the variance is not an inocuous choice since the variance itself is sensitive to the extreme tail of the distribution. To this end, an illustration is given in Table 2 which adopts standardization using unit interquartile range. This standardization has the virtue of being able to be used if some of the comparator distributions have infinite variance, such as is the case for non-normal stable laws. However, the results turn out to be very similar in Tables 1 and 2.

Overlapping of quantiles even for very low p values makes the job of distinguishing power and exponential type tails extremely difficult in practice. And the problem is further exacerbated by the fact that most large data sets are likely to be contaminated by departures from stationarity or autocorrelation structures.

3 Detailed Theoretical Analysis

(1) In the paper we provides a criterion to calculate minimum the required to distinguish two tails.
(2) In the paper we provides a robust measure of tailweight.