Statistical Science
2014, Vol. 29, No. 1, 106–112
DOI: 10.1214/13-STS417

The Two-Piece Normal, Binormal, or Double Gaussian Distribution: Its Origin and Rediscoveries

Kenneth F. Wallis

Abstract. This paper traces the history of the two-piece normal distribution from its origin in the posthumous *Kollektivmasslehre* (1897) of Gustav Theodor Fechner to its rediscoveries and generalisations. The denial of Fechner's originality by Karl Pearson, reiterated a century later by Oscar Sheynin, is shown to be without foundation.

Key words and phrases: Gustav Theodor Fechner, Gottlob Friedrich Lipps, Francis Ysidro Edgeworth, Karl Pearson, Francis Galton, Oscar Sheynin.

1. INTRODUCTION

The two-piece normal distribution came to public attention in the late 1990s, when the Bank of England and the Sveriges Riksbank began to publish probability forecasts of future inflation, using this distribution to represent the possibility that the balance of risks around the central forecast might not be symmetric. The forecast probabilities that future inflation would fall in given intervals could be conveniently calculated by scaling standard normal probabilities, and the resulting density forecasts were visualised in the famous forecast fan charts. In both cases the authors of the supporting technical documentation (Britton, Fisher and Whitley, 1998; Blix and Sellin, 1998) refer readers to Johnson, Kotz and Balakrishnan (1994) for discussion of the distribution. These last authors state (page 173) that "the distribution was originally introduced by Gibbons and Mylroie (1973)," a reference that post-dates the first edition of *Distributions in Statistics* (Johnson and Kotz, 1970), in which the two-piece normal distribution made no appearance, under this or any other name. On the contrary, the distribution was originally introduced in Fechner's *Kollektivmasslehre* (1897) as the *Zweispaltiges* or *Zweiseitige Gauss'sche Gesetz*. In his monumental history of statistics, Hald (1998) prefers the latter name, which translates as the "two-sided Gaussian law," and refers to it as "the Fechner distribution" (page 378). However Fechner's claim to originality had been disputed by Pearson (1905), whose denial of Fechner's originality has recently been repeated by Sheynin (2004). In this paper we reappraise the source and nature of the various claims, and record several rediscoveries of the distribution and extensions of Fechner's basic ideas. As a prelude to the discussion, there follows a brief technical introduction to the distribution.

A random variable X has a two-piece normal distribution with parameters μ, σ_1 and σ_2 if it has probability density function (PDF)

$$(1)\quad f(x) = \begin{cases} A\exp[-(x-\mu)^2/2\sigma_1^2], & x \le \mu, \\ A\exp[-(x-\mu)^2/2\sigma_2^2], & x \ge \mu, \end{cases}$$

where $A = (\sqrt{2\pi}(\sigma_1+\sigma_2)/2)^{-1}$. The distribution is formed by taking the left half of a normal distribution with parameters (μ, σ_1) and the right half of a normal distribution with parameters (μ, σ_2), and scaling them to give the common value $f(\mu) = A$ at the mode, μ, as in (1). The scaling factor applied to the left half of the $N(\mu, \sigma_1)$ PDF is $2\sigma_1/(\sigma_1+\sigma_2)$ while that applied to the right half of $N(\mu, \sigma_2)$ is $2\sigma_2/(\sigma_1+\sigma_2)$, so the probability mass under the left or right piece is $\sigma_1/(\sigma_1+\sigma_2)$ or $\sigma_2/(\sigma_1+\sigma_2)$, respectively. An example with $\sigma_1 < \sigma_2$, in which the two-piece normal distribution is positively skewed, is shown in Figure 1. The skewness becomes extreme as $\sigma_1 \to 0$ and the distribution collapses to the half-normal distribution, while the skewness is reduced as $\sigma_1 \to \sigma_2$, reaching zero when

Kenneth F. Wallis is Emeritus Professor of Econometrics, Department of Economics, University of Warwick, Coventry CV4 7AL, United Kingdom (e-mail: K.F.Wallis@warwick.ac.uk).

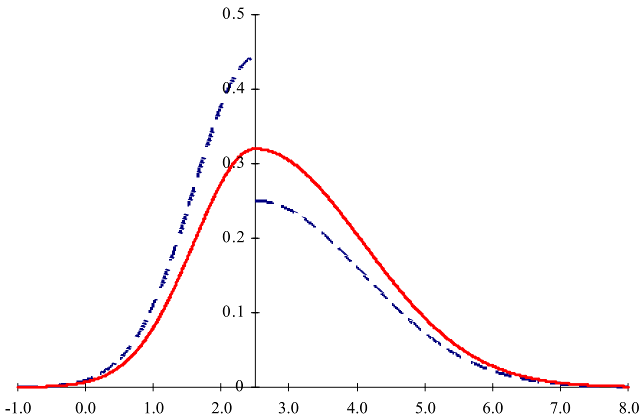

FIG. 1. *The probability density function of the two-piece normal distribution. Dashed line: left half of $N(\mu, \sigma_1)$ and right half of $N(\mu, \sigma_2)$ distributions with $\mu = 2.5$ and $\sigma_1 < \sigma_2$. Solid line: the two-piece normal distribution.*

$\sigma_1 = \sigma_2$ and the distribution is again the normal distribution.

The mean and variance of the distribution are

$$(2)\qquad E(X) = \mu + \sqrt{\frac{2}{\pi}}(\sigma_2 - \sigma_1),$$

$$(3)\qquad \operatorname{var}(X) = \left(1 - \frac{2}{\pi}\right)(\sigma_2 - \sigma_1)^2 + \sigma_1\sigma_2.$$

Expressions for the third and fourth moments about the mean are increasingly complicated and uninformative. Skewness is more readily interpreted in terms of the ratio of the areas under the two pieces of the PDF, which is σ_1/σ_2, or a monotone transformation thereof such as $(\sigma_2 - \sigma_1)/(\sigma_2 + \sigma_1)$, which is the value taken by the skewness measure of Arnold and Groeneveld (1995). With only three parameters there is a one-to-one relation between (the absolute value of) skewness and kurtosis. The conventional moment-based measure of kurtosis, β_2, ranges from 3 (symmetry) to 3.8692 (the half-normal extreme asymmetry), hence the distribution is leptokurtic.

Quantiles of the distribution can be conveniently obtained by scaling the appropriate standard normal quantiles. For the respective cumulative distribution functions (CDFs) $F(x)$ and $\Phi(z)$ we define quantiles $x_p = F^{-1}(p)$ and $z_p = \Phi^{-1}(p)$. Then in the left piece of the distribution we have $x_\alpha = \sigma_1 z_\beta + \mu$, where $\beta = \alpha(\sigma_1 + \sigma_2)/2\sigma_1$. And in the right piece of the distribution, defining quantiles with reference to their upper tail probabilities, we have $x_{1-\alpha} = \sigma_2 z_{1-\delta} + \mu$, where $\delta = \alpha(\sigma_1 + \sigma_2)/2\sigma_2$. In particular, with $\sigma_1 < \sigma_2$, as in Figure 1, the median of the distribution is $x_{0.5} = \sigma_2\Phi^{-1}(1 - (\sigma_1 + \sigma_2)/4\sigma_2) + \mu$. In this case the three central values are ordered mean>median>mode; with negative skewness this order is reversed.

Although the two-piece normal PDF is continuous at μ, its first derivative is not and the second derivative has a break at μ, as first noted by Ranke and Greiner (1904). This has the disadvantage of making standard asymptotic likelihood theory inapplicable, nevertheless standard asymptotic results are available by direct proof for the specific example.

The remainder of this paper is organised as follows. In Section 2 we revisit the distribution's origin in Gustav Theodor Fechner's *Kollektivmasslehre*, edited by Gottlob Friedrich Lipps and published in 1897, ten years after Fechner's death. In Section 3 we note an early rediscovery, two years later, by Francis Ysidro Edgeworth. In Section 4 we turn to the first discussion in the English language of Fechner's contribution, in a characteristically long and argumentative article by Karl Pearson (1905). Pearson derives some properties of "Fechner's double Gaussian curve," but asserts that it is "historically incorrect to attribute [it] to Fechner." We re-examine Pearson's evidence in support of this position, in particular having in mind its reappearance in Oscar Sheynin's (2004) appraisal of Fechner's statistical work. Pearson also argues that "the curve is not general enough," especially in comparison with his family of curves. The overall result was that the Fechner distribution was overlooked for some time, to the extent that there have been several independent rediscoveries of the distribution in more recent years; these are noted in Section 5, together with some extensions.

2. THE ORIGINATORS: FECHNER AND LIPPS

Gustav Theodor Fechner (1801–1887) is known as the founder of psychophysics, the study of the relation between psychological sensation and physical stimulus, through his 1860 book *Elemente der Psychophysik*. Stigler's (1986, pages 242–254) assessment of this "landmark" contribution concludes that "at a stroke, Fechner had created a methodology for a new quantitative psychology." However, his final work, *Kollektivmasslehre*, is devoted more generally to the study of mass phenomena and the search for empirical regularities therein, with examples of frequency distributions taken from many fields, including aesthetics, anthropology, astronomy, botany, meteorology and zoology. In his Foreword, Fechner mentions the long gestation period of the book, and states its main objective as the establishment of a generalisation of the Gaussian law of random errors, to overcome its limitations of symmetric probabilities and relatively small positive and negative deviations from the arithmetic mean. He also appeals to astronomical and statistical institutes to use their mechanical calculation powers to produce accurate tables of the Gaussian distribution, which he had desperately missed during his work on the book. But the book had not been completed when Fechner died in November 1887.

The eventual publication of *Kollektivmasslehre* in 1897 followed extensive work on the incomplete manuscript by Gottlob Friedrich Lipps (1865–1931). In his Editor's Preface, Lipps says that he received the manuscript in early 1895 and that material he has worked on is placed in square brackets in the published work. It is not clear how much unfinished material was left behind by Fechner or to what extent Lipps had to guess at Fechner's intentions. It would appear that the overall structure of the book had already been set out by Fechner, since most of the later chapters have early paragraphs by Fechner, before square-bracketed paragraphs begin to appear. Also, some earlier chapters by Fechner have forward references to later material that appears in square brackets. In general, Lipps' material is more mathematical: he was more of a mathematician than Fechner, who perhaps had set some sections aside for attention later, only to run out of time. Lipps also has a lighter style: for example, Sheynin (2004, page 54) complains about some earlier work that "Fechner's style is troublesome. Very often his sentences occupy eight lines, and sometimes much more—sentences of up to 16 lines are easy to find." The same is true of the present work.

The origin of the two-piece normal distribution is in Chapter 5 of *Kollektivmasslehre*, titled "The Gaussian law of random deviations and its generalisations." Here Fechner uses very little mathematics, postponing more analytical treatment to later chapters. He first presents a numerical example of the use of the Gaussian distribution to calculate the probability of an observation falling in a given interval. The measure of location is the arithmetic mean, A, and the measure of dispersion is the mean absolute deviation, ε (related to the standard deviation, in the Gaussian distribution, by $\varepsilon = \sigma\sqrt{2/\pi}$). Tables of the standard normal distribution are not yet available, and his calculations proceed via the error function (see Stigler, 1986, pages 246–248, e.g.), and prove to be remarkably accurate.

In previous work Fechner had introduced other "main values" of a frequency distribution, the *Zentralwert* or "central value" C, and the *Dichteste Wert* or "densest value" D, subsequently known in English as the median and the mode. Arguing that the equality of A, C and D is the exception rather than the rule, he next introduces the *Zweispaltiges Gauss'sche Gesetz* to represent this asymmetry. Calculating mean absolute deviations from the mode separately for positive and negative deviations from D, the "law of proportions" is invoked, that these should be in the same ratio as the numbers of observations on which they are based. On converting from relative frequencies of observations to probabilities, and from subset mean absolute deviations to subset standard deviations, it is seen that this is exactly the requirement discussed above, that the probabilities below and above the mode are in the ratio σ_1/σ_2, to give a curve that is continuous at the mode. Fechner says that he first discovered this law empirically, and warns that determination of the mode from raw data is not straightforward. He goes on to show that, in this distribution, the median lies between the mean and the mode.

The first mathematical expression of the two normal curves with different precision soon appears in what is the first square-bracketed paragraph in the book and the only such paragraph in Chapter 5. More extensive workings by Lipps appear in Chapter 19, "The asymmetry laws," where every paragraph is enclosed in square brackets. Here Lipps traces the development and properties of the distribution more formally, including an expression for the density function [equation (6), page 297] which corresponds to equation (1) on converting between measures of dispersion. Nevertheless, the key steps in that development, in Chapter 5, were Fechner's alone.

We note that the second "generalisation" presented later in Chapter 5 of *Kollektivmasslehre* is a form of log-normal distribution, but this receives less emphasis and is not our present focus of attention.

3. AN EARLY REDISCOVERY: EDGEWORTH

In 1898–1900 Edgeworth contributed a five-part article "On the representation of statistics by mathematical formulae" to the *Journal of the Royal Statistical Society*, each part appearing in a different issue of the journal. His objective was "to recommend formulae which have some affinity to the normal law of error, as being specially suited to represent statistics of frequency." The first two parts deal with the "method of translation," or transformations to normality, and the "method of separation," or mixtures of normals, using modern terminology.

In the third part Edgeworth considers the "method of composition," in which he constructs "a *composite probability-curve*, consisting of two half-probability curves of different types, tacked together at the *mode*, or greatest ordinate, of each, so as to form a continuous whole, as in the accompanying figure" (1899, page 373, emphasis in original; the figure is very similar to the solid line in Figure 1 above). He gives expressions for the two appropriately scaled half-normal curves, as above, using the *modulus*, equal to $\sqrt{2}$ standard deviation, as his preferred measure of spread. He says that this idea of two probability curves with different moduli is suggested by Ludwig (1898); however, its development in the context of the normal distribution is Edgeworth's alone, since Ludwig's comment comes in a discussion of frequency curves based on the binomial distribution.

To "determine the constants," that is, estimate the parameters, given a sample mean and second and third sample moments, Edgeworth rearranges their definitions to obtain a cubic equation in the distance between the mean and the mode; the required parameter estimates follow from the real solution to this equation. He gives a practical example and compares the method of composition to the methods discussed earlier. In his opinion, the "essential attribute" of the new method is its "deficiency of a priori justification," in contrast to the normal distribution itself.

4. THE CRITICS: PEARSON AND SHEYNIN

The first English-language discussion of Fechner's contribution appears in a 44-page article by Karl Pearson, published in 1905 in *Biometrika*, the journal he had co-founded four years earlier. The article is a response to a review of Pearson's and Fechner's works on skew variation by Ranke and Greiner (1904) in the leading German anthropology journal. Pearson's title quotes most of the title of the German article, omitting its reference to anthropology, and adds the words "A rejoinder," although the running head throughout his article is "Skew variation, a rejoinder." He explains that the German journal had provisionally accepted a rejoinder, but when it arrived the editors did not "see fit to publish" his reply, so he placed it in *Biometrika*, of which he was, in effect, managing editor. From a statistical point of view this seems to have been a more appropriate outcome, since his article contains much general statistical discussion and is most often cited for its introduction of the terms platykurtic, leptokurtic and mesokurtic.

However, Pearson's article also contains extensive attacks on Ranke and Greiner, who had argued that, for the anthropologist, only the Gaussian law is of importance. In this respect the article is a good example of his well-documented behaviour. For example, Stigler (1999, Chapter 1) opens by observing that "Karl Pearson's long life was punctuated by controversies, controversies he often instigated, usually pursued with a zealous energy bordering on obsession;" he "was a fighter who vigorously reacted against opinions that seemed to detract from his own theories. Instead of giving room for other methods and seeking cooperation, his aggressive style led to controversy" (Hald, 1998, page 651); he was ever "relentless in controversy" (Cox, 2001, page 5) and "beyond question a fierce antagonist" (Porter, 2004, page 266). Some of this antagonism is directed towards Fechner: although Pearson and Fechner are on the same side of the debate with Ranke and Greiner about asymmetry, Pearson sees "Fechner's double Gaussian curve" as a rival to his family of curves, and criticises it on both statistical and historical grounds.

Using the parameterisation in terms of σ_1 and σ_2 as in equation (1), Pearson presents expressions for the first four moments of the distribution. Rather than "the rough process by which Fechner determines the mode and obtains the constants of the distribution," he shows that "fitting by my method of moments is perfectly straightforward." To do this, he obtains the cubic equation discussed above, and says in a footnote (page 197) "This cubic was, I believe, first given by Edgeworth," but there is no reference. He observes that the skewness and kurtosis are not independent of one another, so that "we cannot have any form of symmetry but the mesokurtic." He obtains the bounds on β_2 given above, but notes that many empirical distributions with values outside this range have been observed. Hence, Pearson's overall conclusion is that "the double Gaussian curve fails us hopelessly." Curiously, having defined platykurtic as "more flat-topped" and leptokurtic as "less flat-topped" than the normal curve, as has become standard usage, he contrarily describes Fechner's double Gaussian curve as platykurtic, despite having shown its positive excess kurtosis. Similarly, another curve, the symmetrical binomial, is said to be "essentially leptokurtic, that is, $\beta_2 < 3$" (page 175).

Turning to questions of precedence, Pearson's counter claims appear in a footnote (page 196) at the start of the statistical discussion summarised above, which reads as follows:

> Here again it is historically incorrect to attribute these curves to Fechner. They had

> been proposed by De Vries in 1894, and termed "half-Galton curves," and Galton was certainly using them in 1897. See the discussion in Yule's memoir, *R. Statist. Soc. Jour.* Vol. LX, page 45 *et seq.*

Pearson was familiar with De Vries (1894), having used two of his J-shaped botanical frequency distributions as Examples XI and XII in his 1895 article on skew variation. De Vries said that these deserved the name half-Galton (i.e., half-normal) simply on the basis of the appearance of the empirical distributions, and no fitting was attempted, nor did he make any proposal to place two such curves together to give a more general asymmetric distribution. Fechner's curve had *not* been proposed by De Vries. [Edgeworth knew that his composite curve had not either, noting at the outset (1899, page 373) that "It will be observed that the following construction is not much indebted to the "half-Galtonian" curve employed by Professor De Vries."]

Galton comes a little closer, but Pearson is again incorrect. His citation is inaccurate, since he clearly has in mind Yule's paper read at the Royal Statistical Society in January 1896, published with discussion later that year (Yule, 1896a). Galton had opened the discussion at the meeting and mentioned his method of percentiles as an alternative to the method of frequency curves developed by Pearson and applied by Yule. In response to a request at the meeting, he provided a memorandum giving fuller information on his method, which was published in the same issue of the Society's journal (Galton, 1896), together with a reply by Yule (1896b). Galton explains how his method of percentiles, in this example method of deciles, smooths the original frequency table or "frequency polygon" of Yule by interpolating deciles and plotting them. He then mentions another approach, namely

> . . . the extremely rude and scarcely defensible method, but still a sometimes serviceable one, of looking upon skew-curves as made up of the halves of two different normal curves pieced together at the mode. . . . On trying it, again for curiosity's sake, with the present series for all the five years, there was of course no error for the 2nd, 5th, and 8th deciles, . . .

because he had inferred the spread or standard deviation of the lower half-normal distribution from the lower 20% point of the standard normal distribution, and similarly for the upper part; he goes on to discuss the errors of fit at the other deciles. But no "law of proportions" or scaling is applied, and the resulting curve is discontinuous, like the initial two halves of normal curves in Figure 1. Yule (1896b) recognises this in his response, noting that, in contrast, his skew-curve "presents a continuous distribution round the mode." Galton was certainly *not* using Fechner's curves.

The erroneous assertions in Pearson's footnote may be due to his combativeness. Several authors also discuss the tremendous volume of work he undertook. For example, Cox (2001, page 6) observes that he "wrote more than 90 papers in *Biometrika* in the period up to 1915, few of them brief, and appears to have been the moving spirit behind many more." He founded not only the journal but also the Biometrics Laboratory at University College London at this time. His son Egon remarks that the volume of work "led inevitably to a certain hurry in execution" (E. S. Pearson, 1936, page 222). This remark is made during discussion of one of Pearson's two well-known errors, recently reappraised by Stigler (2008), but it perhaps also applies to the mistakes discussed above, which are of a smaller order of magnitude. Nevertheless, Pearson's assertions in the quoted footnote are mistaken, and his challenge to Fechner's claim to priority is unjustified, and thereby unjust.

Sheynin (2004), in his review of Fechner's statistical work, has a very brief discussion of the double-sided Gaussian law, quoting from sections of *Kollektivmasslehre* that had been worked on by Lipps, and hence underestimating the role of Fechner's law of proportions. In his discussion (page 68) he states that the double-sided Gaussian law was not original to Fechner, this having been pointed out, forcefully, by Pearson (1905). As if quoting from Pearson, and giving no citation for De Vries (1894), Sheynin states "De Vries, in 1894, had applied the double-sided law." In this statement "applied" is somewhat stronger than Pearson's "proposed," hence is further from the truth, and Sheynin's denial of Fechner's originality is similarly inaccurate and unjust.

5. LATER REDISCOVERIES AND EXTENSIONS

The result of Pearson's critique appears to have been that, with two exceptions discussed below, the Fechner distribution, with this attribution, disappeared from the statistical literature until its reappearance in Hald's (1998) history. Meanwhile, three independent rediscoveries occurred.

First, in the physics literature, is Gibbons and Mylroie's (1973) "joined half-Gaussian" distribution, cited by Johnson, Kotz and Balakrishnan (1994), as noted above; the distribution is fitted by what statisticians recognise as the method of moments. Second, in the statistics literature, is the "three-parameter two-piece normal" distribution of John (1982), also cited by Johnson, Kotz and Balakrishnan (1994); John compares estimation by the method of moments and maximum likelihood. In the same journal Kimber (1985) notes that John (1982) is a rediscovery, with reference to Gibbons and Mylroie (1973); he proves the asymptotic normality of ML estimators and provides a likelihood ratio test of symmetry. Finally, in the meteorology literature, Toth and Szentimrey (1990) introduce the "binormal" distribution, again fitted by ML, with a test of symmetry. The same name is used by Garvin and McClean (1997), who nevertheless again attribute the distribution to Gibbons and Mylroie. In all these articles the distribution is parameterised in terms of the mode, using various symbols, and the standard deviations σ_1 and σ_2, as in (1) above. An alternative parameterisation, with a single explicit skewness parameter, is given by Mudholkar and Hutson (2000), who do acknowledge Fechner's priority.

A modern, but pre-Hald (1998) attribution to *Kollektivmasslehre* occurs at the start of an exploration by Runnenburg (1978) of the mean, median, mode ordering. He notes that Fechner had shown this *Lagegesetz der Mittelwerte* for the two-piece normal distribution, and investigates more general conditions in which it holds. The second exceptional appearance of the Fechner distribution in the statistical literature pre-1998 is more substantial. Barnard (1989), seeking a family of distributions "which may be expected to represent most of the types of skewness liable to arise in practice," introduces the distribution

$$f(x) = \begin{cases} K \exp\left[-\frac{1}{2}\left(\frac{-M(x-\mu)}{\sigma}\right)^a\right], & x \le \mu, \\ K \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^a\right], & x \ge \mu, \end{cases}$$

which reparameterises and generalises the two-piece normal distribution in equation (1). He calls it the Fechner family, because by allowing the skewness parameter M ($M > 0$) to differ from 1 it embodies Fechner's idea in *Kollektivmasslehre* of having different scales for positive and negative deviations from the mode, μ. It also allows for nonnormal kurtosis by allowing a ($1 \le a < \infty$) to differ from 2. The scale parameter σ is equal to the standard deviation if $(M, a) = (1, 2)$ but not otherwise, in general. This "Fechner family of unimodal densities" also appears in a later article (Barnard, 1995), which is cited by Hald (1998, page 380). We note that the case $a = 1$, the asymmetric Laplace distribution, has a considerable life of its own, beginning before Barnard's work: see, for example, Kotz, Kozubowski and Podgorski (2001, Chapter 3) and the references therein.

Two further extensions of note, independent of Fechner, can be found in Bayesian statistics. For the application of Monte Carlo integration with importance sampling to Bayesian inference, Geweke (1989) uses "split" (i.e., two-piece) multivariate normal and Student-t distributions as importance sampling densities. The generalisation by Fernandez and Steel (1998) is also cast in a Bayesian setting: as in Barnard's Fechner family, there is a single skewness parameter, which is convenient whenever it is desired to assign priors to skewness; nevertheless, it has general applicability. For any univariate PDF $f(x)$ which is unimodal and symmetric around 0, Fernandez and Steel's class of two-piece or split distributions $p(x|\gamma)$, indexed by a skewness parameter γ ($\gamma > 0$), is

$$(4) \qquad p(x|\gamma) = \begin{cases} Kf(\gamma x), & x \le 0, \\ Kf\left(\frac{x}{\gamma}\right), & x \ge 0, \end{cases}$$

where $K = 2(\gamma + \gamma^{-1})^{-1}$. If $\gamma > 1$ there is positive skewness, and inverting γ produces the mirror image of the density function around 0. Unlike Barnard's Fechner family there is no explicit kurtosis parameter; kurtosis is introduced, if desired, by the choice of $f(x)$, most commonly as Student-t. An extension with two tail parameters to allow different tail behaviour in an asymmetric two-piece t-distribution is developed by Zhu and Galbraith (2010).

ACKNOWLEDGMENTS

I am grateful to Sascha Becker, David Cox, Chris Jones, Malte Knüppel, Kevin McConway, Mary Morgan, Oscar Sheynin, Mark Steel, Stephen Stigler and two referees for comments and suggestions at various stages of this work. I am especially indebted to Karl-Heinz Tödter for his expert assistance with the German-language works cited. Thanks also to the Resource Delivery team at the University of Warwick Library.

REFERENCES

ARNOLD, B. C. and GROENEVELD, R. A. (1995). Measuring skewness with respect to the mode. *Amer. Statist.* **49** 34–38. MR1341197

BARNARD, G. A. (1989). Sophisticated theory and practice in quality improvement. *Philos. Trans. R. Soc. Lond. A* **327** 581–589.

BARNARD, G. A. (1995). Pivotal models and the fiducial argument. *Internat. Statist. Rev.* **63** 309–323.

BLIX, M. and SELLIN, P. (1998). Uncertainty bands for inflation forecasts. Working Paper 65, Sveriges Riksbank, Stockholm.

BRITTON, E., FISHER, P. G. and WHITLEY, J. D. (1998). The *Inflation Report* projections: Understanding the fan chart. *Bank of England Quarterly Bulletin* **38** 30–37.

COX, D. R. (2001). Biometrika: The first 100 years. *Biometrika* **88** 3–11. MR1841263

DE VRIES, H. (1894). Ueber halbe Galton-Curven als Zeichen discontinuirlicher Variation. *Berichte der Deutschen Botanischen Gesellschaft* **12** 197–207.

EDGEWORTH, F. Y. (1899). On the representation of statistics by mathematical formulae (Part III). *J. Roy. Statist. Soc.* **62** 373–385.

FECHNER, G. T. (1860). *Elemente der Psychophysik*. Breitkopf and Hartel, Leipzig.

FECHNER, G. T. (ed. LIPPS, G. F.) (1897). *Kollektivmasslehre* Engelmann, Leipzig.

FERNÁNDEZ, C. and STEEL, M. F. J. (1998). On Bayesian modeling of fat tails and skewness. *J. Amer. Statist. Assoc.* **93** 359–371. MR1614601

GALTON, F. (1896). Application of the method of percentiles to Mr. Yule's data on the distribution of pauperism. *J. Roy. Statist. Soc.* **59** 392–396.

GARVIN, J. S. and MCCLEAN, S. I. (1997). Convolution and sampling theory of the binormal distribution as a prerequisite to its application in statistical process control. *The Statistician* **46** 33–47.

GEWEKE, J. (1989). Bayesian inference in econometric models using Monte Carlo integration. *Econometrica* **57** 1317–1339. MR1035115

GIBBONS, J. F. and MYLROIE, S. (1973). Estimation of impurity profiles in ion-implanted amorphous targets using joined half-Gaussian distributions. *Appl. Phys. Lett.* **22** 568–569.

HALD, A. (1998). *A History of Mathematical Statistics from* 1750 *to* 1930. Wiley, New York. MR1619032

JOHN, S. (1982). The three-parameter two-piece normal family of distributions and its fitting. *Comm. Statist. A—Theory Methods* **11** 879–885. MR0651620

JOHNSON, N. L. and KOTZ, S. (1970). *Distributions in Statistics*. Wiley, New York.

JOHNSON, N. L., KOTZ, S. and BALAKRISHNAN, N. (1994). *Continuous Univariate Distributions*, *Vol. 1*, 2nd ed. Wiley, New York. MR1299979

KIMBER, A. C. (1985). Methods for the two-piece normal distribution. *Comm. Statist. A—Theory Methods* **14** 235–245. MR0788796

KOTZ, S., KOZUBOWSKI, T. J. and PODGÓRSKI, K. (2001). *The Laplace Distribution and Generalizations: A Revisit With Applications to Communications, Economics, Engineering, and Finance*. Birkhäuser, Boston, MA. MR1935481

LUDWIG, F. (1898). Die pflanzlichen Variationscurven und die Gauss'sche Wahrscheinlichkeitscurve (continuation). *Botanisches Centralblatt* **73** 343–349.

MUDHOLKAR, G. S. and HUTSON, A. D. (2000). The epsilon-skew-normal distribution for analyzing near-normal data. *J. Statist. Plann. Inference* **83** 291–309. MR1748020

PEARSON, K. (1895). Contributions to the mathematical theory of evolution.—II. Skew variation in homogeneous material. *Phil. Trans. R. Soc. Lond. A* **186** 343–414.

PEARSON, K. (1905). Das fehlergesetz und seine verallgemeinerungen durch Fechner und Pearson. A rejoinder. *Biometrika* **4** 169–212.

PEARSON, E. S. (1936). Karl Pearson: An appreciation of some aspects of his life and work. Part I: 1857–1906. *Biometrika* **28** 193–257.

PORTER, T. M. (2004). *Karl Pearson: The Scientific Life in a Statistical Age*. Princeton Univ. Press, Princeton, NJ. MR2054951

RANKE, K. E. and GREINER, A. (1904). Das Fehlergesetz und seine Verallgemeinerungen durch Fechner und Pearson in ihrer Tragweite für die Anthropologie. *Archiv fur Anthropologie* **2** 295–332.

RUNNENBURG, J. T. (1978). Mean, median, mode. *Stat. Neerl.* **32** 73–79. MR0501446

SHEYNIN, O. (2004). Fechner as a statistician. *British J. Math. Statist. Psych.* **57** 53–72. MR2086215

STIGLER, S. M. (1986). *The History of Statistics: The Measurement of Uncertainty Before* 1900. The Belknap Press of Harvard Univ. Press, Cambridge, MA. MR0852410

STIGLER, S. M. (1999). *Statistics on the Table: The History of Statistical Concepts and Methods*. Harvard Univ. Press, Cambridge, MA. MR1712969

STIGLER, S. M. (2008). Karl Pearson's theoretical errors and the advances they inspired. *Statist. Sci.* **23** 261–271. MR2446501

TOTH, Z. and SZENTIMREY, T. (1990). The binormal distribution: A distribution for representing asymmetrical but normal-like weather elements. *J. Climate* **3** 128–136.

YULE, G. U. (1896a). Notes on the history of pauperism in England and Wales from 1850, treated by the method of frequency-curves; with an introduction on the method (with discussion). *J. Roy. Statist. Soc.* **59** 318–357.

YULE, G. U. (1896b). Remarks on Mr. Galton's note. *J. Roy. Statist. Soc.* **59** 396–398.

ZHU, D. and GALBRAITH, J. W. (2010). A generalized asymmetric Student-t distribution with application to financial econometrics. *J. Econometrics* **157** 297–305. MR2661602