Klein, Ingo; Grottke, Michael

Working Paper

On J.M. Keynes' The principal averages and the laws of error which lead to them: refinement and generalisation

IWQW discussion paper series, No. 07/2008

IWQW

Institut für Wirtschaftspolitik und Quantitative
Wirtschaftsforschung

Diskussionspapier
Discussion Papers

No. 07/2008

On J.M. Keynes' "The Principal Averages and the Laws of Error which Lead to Them" - Refinement and Generalisation

Ingo Klein
University of Erlangen-Nuremberg

Michael Grottke
University of Erlangen-Nuremberg

ISSN 1867-6707

On J.M. Keynes' "The Principal Averages and the Laws of Error which Lead to Them" – Refinement and Generalisation

Ingo Klein and Michael Grottke

Department of Statistics and Econometrics
University of Erlangen-Nürnberg
ingo.klein@wiso.uni-erlangen.de
michael.grottke@wiso.uni-erlangen.de

Abstract

Keynes (1911) derived general forms of probability density functions for which the "most probable value" is given by the arithmetic mean, the geometric mean, the harmonic mean, or the median. His approach was based on indirect (i.e., posterior) distributions and used a constant prior distribution for the parameter of interest. It was therefore equivalent to maximum likelihood (ML) estimation, the technique later introduced by Fisher (1912).
Keynes' results suffer from the fact that he did not discuss the supports of the distributions, the sets of possible parameter values, and the normalising constants required to make sure that the derived functions are indeed densities. Taking these aspects into account, we show that several of the distributions proposed by Keynes reduce to well-known ones, like the exponential, the Pareto, and a special case of the generalised inverse Gaussian distribution.
Keynes' approach based on the arithmetic, the geometric, and the harmonic mean can be generalised to the class of quasi-arithmetic means. This generalisation allows us to derive further results. For example, assuming that the ML estimator of the parameter of interest is the exponential mean of the observations leads to the most general form of an exponential family with location parameter introduced by Dynkin (1961) and Ferguson (1962, 1963).

Keywords: ML estimator, criterion function, median, quasi-arithmetic mean, exponential family

JEL classification: C13, C16

1 Introduction

In his dissertation *The Principles of Probability*, submitted in two parts in 1907 and 1908, Keynes derived the most general forms of distributions for which the "most probable value" is given by a "principal average." Keynes used the latter term to refer to the arithmetic mean, the geometric mean, the harmonic mean, and the median. He published his results in the Journal of the Royal Statistical Society in 1911, and he also included them as a chapter in his monograph *A Treatise on Probability* (1921).

For deriving the distributions, Keynes used a technique already employed by Gauss (1809) to show that the arithmetic mean is the "most probable value" of a normal distribution. Following Bayes' theorem, Gauss solved his task by maximising the "indirect probability," which has been known as the "posterior density function" since the mid-twentieth century. Moreover, Gauss assumed an improper prior distribution for the location parameter, more specifically a constant prior density function on the entire domain $\mathbb{R}$. This approach (referred to as the "indifference principle," or the "principle of insufficient reason") was also adopted by Keynes (1907, 1908, 1911). However, when applying the indifference principle, determining the "most probable value" by maximising the posterior distribution is formally identical with deriving the maximum likelihood (ML) estimator by maximising the likelihood function. Kendall & Stuart (1967, p. 677) therefore mention Keynes' work in the context of the "characterization of distributions by forms of maximum likelihood estimators." Following Kendall & Stuart, throughout this paper we will refer to Keynes' problem as that of trying to find distributions connected with given forms of ML estimators.

The concept of a likelihood function was first formally introduced by Fisher (1912). Ten years later, Fisher (1922) presented a complete theory of ML estimation. Fisher's work is thus considered the origin of this theory – although Stigler (1978, 1986) was able to show that ML estimation had occasionally been used before. Conniffe (1992) argued that Fisher must have known Keynes' 1911 paper; he conjectured that Fisher was inspired by the fact that the distributions derived by Keynes are all members of the exponential family, for which properties like sufficiency, completeness, and monotony of the likelihood ratio are easily established. However, Aldrich (2008) casts doubts on this assumption.

Keynes' results on the laws of errors following from specific averages as "most probably values" were hardly taken into account in the literature. Among the few exceptions are biographical accounts on J.M. Keynes' lifework (for example, Skidelsky (1983, 1992, 2000) and O'Donnell (1989)), the above-mentioned paragraph in Kendall & Stuart (1967), and a remark in Patel et al. (1976). Keynes himself considered them as minor work paling in comparison with his contributions to the logic of probability. On page 186 of his monograph *A Treatise on Probability* (1921) he remarked that the content of chapter 17 "is without philosophical interest and should probably be omitted by most readers." This is an appropriate recommendation for those readers who are exclusively interested in the philosophical foundations of probability. However, readers who are into the theory of probability distributions can gain valuable insight from this work and generalise its findings – after refining some of Keynes' results.

The necessary refinements are mainly related to the discussion of the supports, the parameter domains, and the normalising constants of the derived density functions. Keynes did not explicitly specify these aspects in his mathematical expressions. As we will show, only in a few of the cases suggested by him the related ML estimator does indeed turn out to be a principal average. Calculating the normalising constants reveals that in most cases this constant depends on the parameter of interest; the ML estimator of this parameter is then a function of a principle average. Moreover, by including the respective normalising constants, we can show that some of the proposed distributions reduce to ones that are well-known at least nowadays, like the exponential, the Pareto, and a special case of the generalised inverse Gaussian distribution – the latter one may have been unknown when Keynes published his dissertation.

As far as the arithmetic, the geometric, and the harmonic mean are concerned, Keynes' approach can easily be generalised by considering the ML estimator to be a quasi-arithmetic mean. In addition to the three types of means discussed by Keynes, this class of mean values also includes the exponential mean and the power mean. It can be shown that the location exponential family and a special case of the scale exponential family can be derived when assuming that the ML estimator is an exponential, and a power mean, respectively.

This paper is structured as follows: In Section 2, we sketch Keynes' approach to deriving density functions based on a general criterion function that links the observations with the parameter of interest. Moreover, we list Keynes' results and comment on those distributions for which the ML estimator of the parameter is indeed given by a mean, or the median. While Keynes introduced a general criterion function, he did not discuss it any further. Section 3 is devoted to a general criterion function connected with the quasi-arithmetic mean as the ML estimator of the parameter of interest. After deriving general results in Section 3.1, we discuss distributions connected with specific quasi-arithmetic means in Sections 3.2 – 3.6. Among the concrete distributions examined are those derived by Keynes that we did not discuss in Section 2. Calculating the respective normalising constants, we show that in each of these cases the ML estimator of the parameter of interest is not a mean, as Keynes assumed, but the function of a mean. Moreover, the generalisation allows us to derive further results. For example, we demonstrate that the location and scale exponential family can be obtained via Keynes' approach by choosing the criterion functions linked to the exponential mean, and the power mean.

2 Keynes' approach and results

Keynes considers the general criterion function $\psi(x;\theta)$ linking the observation x with the unknown parameter θ. If the observations $x_1, x_2, \ldots, x_n$ are realisations of random variables $X_1, X_2, \ldots, X_n$ that are independent and identically distributed with probability density function $f(x;\theta)$, then the ML estimator of the unknown parameter θ is derived by solving the equation

$$\sum_{i=1}^{n} \psi(x_i;\theta) = 0 \tag{1}$$

for θ. Possible choices for $\psi(x;\theta)$ include the following:

$$\psi(x;\theta) = x - \theta, \tag{2}$$

$$\psi(x;\theta) = \ln\left(\frac{x}{\theta}\right) = \ln x - \ln\theta, \tag{3}$$

$$\psi(x;\theta) = \frac{1}{x} - \frac{1}{\theta}, \tag{4}$$

$$\psi(x;\theta) = \operatorname{sign}(x - \theta). \tag{5}$$

It is well known that the criterion functions (2), (3), and (4) result in the ML estimator of θ being the arithmetic, the geometric, and the harmonic mean of the observations, respectively; criterion function (5) leads to their median.[1]

[1] This is the general approach, also used by Huber (1964) for M estimation. However, in robust statistics, the choice of ψ depends on the desired robustness and efficiency properties of the estimator; see Huber (1981).

Multiplying (1) with some function $\phi''(\theta) \neq 0$ for all $\theta \in \Theta$ with a suitable parameter space Θ does not affect the root of the left-hand side; i.e., we can alternatively obtain the ML estimator of θ as the solution of

$$\sum_{i=1}^{n} \psi(x_i; \theta)\phi''(\theta) = 0. \tag{6}$$

$\phi''(\theta)$ represents the second derivative of a function $\phi(\theta)$ which is twice differentiable, but which we otherwise do not specify any further for the time being. Defining the function with which (1) is multiplied as a second derivative will simplify the following expressions.

If density f is differentiable with respect to parameter θ, then under certain regularity conditions the ML estimator can be determined by solving the following necessary condition for maximising the log-likelihood function with respect to θ:

$$\sum_{i=1}^{n} \frac{\partial \ln f(x_i; \theta)}{\partial \theta} = 0. \tag{7}$$

Conditions (6) and (7) are equivalent if density f satisfies the differential equation

$$\frac{\partial \ln f(x; \theta)}{\partial \theta} = \psi(x; \theta)\phi''(\theta). \tag{8}$$

For the related density f the ML estimator of θ is determined by (1). The solution of differential equation (8) is immediately obtained as

$$f(x; \theta) = K \exp\left(\int \psi(x; \theta)\phi''(\theta) d\theta + b(x)\right), \tag{9}$$

where K is a normalising constant, which Keynes (1911) assumed to be independent of both x and θ. Using integration by parts, (9) can be transformed into

$$f(x; \theta) = K \exp\left(\psi(x; \theta)\phi'(\theta) - \int \frac{\partial \psi(x; \theta)}{\partial \theta}\phi'(\theta) d\theta + b(x)\right). \tag{10}$$

Appropriately choosing $\psi(x; \theta)$ as well as $\phi''(\theta)$ and $b(x)$, making sure that K does not depend on θ, it is thus possible to derive a density function for which the ML estimator of the unknown parameter θ is the one connected with the criterion function $\psi(x; \theta)$.

Keynes made a number of implicit assumptions about the mathematical properties of the function ψ and the parameter space Θ. For example, he restricted his considerations to absolutely continuous functions; moreover, he implied that the ML estimator is necessarily determined by the root of the derivative of the log-likelihood function. Teicher (1961) later characterised distributions for which the ML estimator is given by the arithmetic mean, without setting these strict requirements for differentiability.

Besides (2), (3) and (4), Keynes considered the criterion function[2]

$$\psi(x;\theta) = \frac{x-\theta}{|x-\theta|}, \tag{11}$$

which is identical to (5) for $x \neq \theta$. For each of these criterion functions, Table 1 lists the densities derived by Keynes based on different choices for $\phi'(\theta)$ and $b(x)$. The numbers in the first column refer to the numbering in the concluding §9 on page 331 of Keynes (1911). (Items related to numbers not included here are not specific densities, but general results.) Keynes did not discuss the supports of the distributions. Moreover, he did not specify the normalising constant for any of these cases.

#	$\psi(x;\theta)$	$\phi'(\theta)$	$b(x)$	Density
3	(2)	$2k^2\theta$	$-k^2x^2$	$K\exp\left(-k^2(x-\theta)^2\right)$
6	(2)	$-k^2e^\theta$	0	$K\exp\left(-k^2e^\theta(x-\theta)-k^2e^\theta\right)$
8	(3)	$-k^2\theta$	0	$K\left(\frac{\theta}{x}\right)^{k^2\theta}\exp\left(-k^2\theta\right)$
9	(3)	$2k^2\ln\theta$	$-k(\ln x)^2$	$K\exp\left(-k^2\left(\ln\frac{x}{\theta}\right)^2\right)$
11	(4)	$-k^2\theta^2$	$-k^2x$	$K\exp\left(-\frac{k^2}{x}(x-\theta)^2\right)$
4	(11)	$k^2\theta$	$-k^2x\frac{x-\theta}{\lvert x-\theta\rvert}$	$K\exp\left(-k^2\lvert x-\theta\rvert\right)$

Table 1: Exemplary cases discussed by Keynes (1911)

The distributions # 3 and # 4 are well-known as normal distribution and Laplace (or double exponential) distribution.

For the normal distribution, the usual parameterisation is $k^2 = 1/(2\sigma^2)$ with $\sigma > 0$, which implies the normalising constant $K = 1/\sqrt{2\pi\sigma^2}$. Support and parameter domain are both $\mathbb{R}$. Of course, the ML estimators for θ is given by the arithmetic mean.

[2]In fact, the first and the last criterion function considered by Keynes are (2) and (11) multiplied by (-1), respectively. The forms used here are more common nowadays. Employing them will also simplify the following discussion.

With respect to Keynes' derivation of density # 4, it should be noted that plugging his choices of $\psi(x;\theta)$, $\phi'(\theta)$, and $b(x)$ into (10) does indeed lead to the Laplace density; this can easily be seen since the derivative of criterion function (11) is equal to zero, wherever it exists. However, the function $b(x)$ used by Keynes depends on θ, which was explicitly excluded when solving differential equation (8).

Nevertheless, the ML estimator of the parameter θ in a Laplace distribution is in fact the median. Consider criterion function (5). Using integration by parts, the density

$$f(x;\theta) = K \exp\left(\int \operatorname{sign}(x-\theta)\phi''(\theta)d\theta + b(x)\right)$$

following from (9) can also be written as

$$f(x;\theta) = K \exp\left(|x-\theta|\phi''(\theta) + \int |x-\theta|\phi'''(\theta)d\theta + b(x)\right).$$

Choosing $\phi'(\theta) = -k^2\theta$, as Keynes did, and $b(x) = 0$ results in

$$f(x;\theta) = K \exp\left(-k^2|x-\theta|\right),$$

the Laplace density with $K = k^2/2$. Support and parameter domain are both $\mathbb{R}$. It is commonly known that the median is the ML estimator of the unknown parameter of a Laplace distribution.

Keynes assumed that for all the densities he derived the ML estimator of θ is the one connected with the respective criterion function chosen. As we have seen, he was correct with respect to densities # 3 and # 4. However, in the next section we will demonstrate that for the remaining cases the ML estimator is not a mean, as Keynes supposed, but the function of a mean. Moreover, we will show that three of the distributions identified by Keynes are in fact the exponential, the Pareto, and a specific type of generalised inverse Gaussian distribution.

Since the arithmetic, the geometric, and the harmonic mean are special cases of the quasi-arithmetic mean, we will first derive results for this more general class of means. This will allow us to discuss further distributions, in addition to Keynes' examples.

3 Quasi-arithmetic means

3.1 General results

As a general criterion function including the cases (2), (3) and (4) considered by Keynes, we choose

$$\psi(x;\theta) = u(x) - u(\theta), \quad x \in \mathcal{X}, \ \theta \in \Theta, \tag{12}$$

where u is a function that is strictly monotonic a.e. in $D \subset \mathbb{R}$, and $\mathcal{X}$, $\Theta \subset D$.

Note that based on (10) a particularly simple expression for the density f related to a criterion function is obtained if $\partial\psi(x;\theta)/\partial\theta$ is independent of x. In this case, we can define

$$C'(\theta) = \left(-\frac{\partial\psi(x;\theta)}{\partial\theta}\right)\phi'(\theta)$$

and write the density as

$$f(x;\theta) = K \exp\left(\psi(x;\theta)\frac{C'(\theta)}{-\partial\psi(x;\theta)/\partial\theta} + C(\theta) + b(x)\right). \tag{13}$$

For criterion functions of the form (12), $\partial\psi(x;\theta)/\partial\theta = -u'(\theta)$ is independent of x. The related density is thus given by

$$f(x;\theta) = K \exp\left((u(x) - u(\theta))\frac{C'(\theta)}{u'(\theta)} + C(\theta) + b(x)\right), \quad x \in \mathcal{X}, \ \theta \in \Theta, \tag{14}$$

with K such that $\int_{\mathcal{X}} f(x;\theta) = 1$. In the following, we will specify all densities by choosing $u(x)$ and $C(\theta)$ instead of $\psi(x;\theta)$ and $\phi'(\theta)$, in addition to $b(x)$.

If K does not depend on θ, as assumed by Keynes, then the ML estimator for θ can be derived by applying the invariance property of ML estimators (Zehna (1966)):

$$\hat{\theta}_{ML} = u^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} u(X_i)\right), \tag{15}$$

where u^{-1} is the inverse function of u defined on its range. This estimator is referred to as quasi-arithmetic mean; see for example Aczél (1966, p. 276) and Bullen et al. (1988, p. 215).

The exact definition of a quasi-arithmetic mean is as follows (see for example Jarczyk (2007, p. 3)):

Definition 1 *Let $D \subset \mathbb{R}$ be an interval, and $u : D \to \mathbb{R}$ be a continuous and strictly monotonic function. Then*

$$u^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} u(x_i)\right)$$

is called quasi-arithmetic mean of $x_1, \ldots, x_n \in D$. u is referred to as the generator of the quasi-arithmetic mean.

Note that there is no one-to-one correspondence between function u and density f. An affine linear transformation of $u(x)$, like $v(x) = a + cu(x)$, results in

$$f(x;\theta) = K \exp\left((v(x) - v(\theta))\frac{C'(\theta)}{v'(\theta)} + C(\theta) + b(x)\right), \quad x \in \mathcal{X},\ \theta \in \Theta,$$

which is equivalent with density (14). Therefore, every affine linear transformation of a generator u leads to the same quasi-arithmetic mean. This relationship is especially important for the Box-Cox transformation (see Box & Cox (1964), (1982)), as illustrated in the following example:

Example 1 *Consider the power function $u(x) = x^\gamma$ for $x > 0$ and $\gamma \neq 0$. Using the generator*

$$v(x) = \frac{x^\gamma - 1}{\gamma} = \frac{u(x)}{\gamma} - \frac{1}{\gamma}, \quad x > 0, \gamma \neq 0,$$

the Box-Cox transformation of x, we obtain the same quasi-arithmetic mean as based on the power function $u(x)$.

Keynes implicitly assumed that the normalising constant K never depends on the parameter of interest. However, for an arbitrary selection of functions $u(\theta)$, $C(\theta)$, and $b(x)$ it may indeed be a function of θ:

$$K(\theta) \equiv \int \exp\left(-(u(x) - u(\theta))\frac{C'(\theta)}{u'(\theta)} - C(\theta) - b(x)\right) dx.$$

Therefore, the general form of the density obtained is

$$f(x;\theta) = K(\theta) \exp\left((u(x) - u(\theta))\frac{C'(\theta)}{u'(\theta)} + C(\theta) + b(x)\right), \quad x \in \mathcal{X},\ \theta \in \Theta. \tag{16}$$

Only if $K(\theta)$ does not depend on the parameter θ (the case discussed above), then the ML estimator of θ is a quasi-arithmetic mean. Otherwise, it is a function of

$1/n\sum_{i=1}^n u(X_i)$, and thus a *function* of a quasi-arithmetic mean, as the following reasoning shows: With K depending on θ, we need to find the solution of

$$\sum_{i=1}^n \frac{\ln \partial f(x_i;\theta)}{\partial\theta} = \sum_{i=1}^n \left(\frac{\partial \ln K(\theta)}{\partial\theta} + (u(x_i) - u(\theta)) \frac{\partial (C'(\theta)/u'(\theta))}{\partial\theta} \right) = 0.$$

That is, we have to solve the equation

$$\frac{1}{n}\sum_{i=1}^n u(x_i) = u(\theta) - \frac{\partial \ln K(\theta)}{\partial\theta} \frac{1}{\frac{\partial (C'(\theta)/u'(\theta))}{\partial\theta}} \equiv \tau(\theta). \tag{17}$$

Given that τ is an injective function, the ML estimator of θ is obtained as

$$\hat{\theta}_{ML} = \tau^{-1}\left(\frac{1}{n}\sum_{i=1}^n u(X_i)\right) = (\tau^{-1}\circ u)\left(u^{-1}\left(\frac{1}{n}\sum_{i=1}^n u(X_i)\right)\right). \tag{18}$$

3.2 Arithmetic mean

In this section, we discuss special cases of density (16) with $u(x) = x$ for $x \in \mathcal{X}$. Making this choice results in the general form

$$f(x;\theta) = K(\theta)\exp\left((x-\theta)C'(\theta) + C(\theta) + b(x)\right) \tag{19}$$

for $x \in \mathcal{X}$. If $K(\theta)$ is independent of θ, then the ML estimator of θ is the arithmetic mean of the observations.

3.2.1 Density # 6: Exponential distribution

Plugging the specific choice $C(\theta) = -k^2 e^\theta$ and $b(x) = 0$ into (19) immediately leads to Keynes' density # 6:

$$f(x;\theta) = K(\theta)\exp\left(-k^2 e^\theta (x-\theta) - k^2 e^\theta\right).$$

Since the exponent is a linear function of x, f can only be a density if the support $\mathcal{X}$ has a lower bound. We choose $\mathcal{X} = \mathbb{R}^+$. As a consequence, $\Theta = \mathcal{X} = \mathbb{R}^+$ as well.

Lemma 1 *The normalising constant of density # 6 is given by*

$$K(\theta) = \frac{k^2 e^\theta}{\exp\left(k^2\left(e^\theta\theta - e^\theta\right)\right)},$$

and the ML estimator of θ is

$$\hat{\theta}_{ML} = -\ln\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) - \ln k^2,$$

provided that k is known.

Proof: Since

$$\begin{aligned} K(\theta)^{-1} &= \int_0^\infty \exp\left(-k^2\left(e^\theta\left(x-\theta\right)+e^\theta\right)\right)dx \\ &= \exp\left(k^2 e^\theta(\theta-1)\right)\int_0^\infty \exp\left(-k^2 e^\theta x\right)dx = \exp\left(k^2 e^\theta(\theta-1)\right)\frac{1}{k^2 e^\theta}, \end{aligned}$$

the normalising constant $K(\theta)$ is indeed a function of θ, which needs to be taken into account for ML estimation. Since

$$\frac{\partial \ln K(\theta)}{\partial\theta} = 1 - k^2\theta e^\theta \quad \text{and} \quad \frac{\partial(C'(\theta)/u'(\theta))}{\partial\theta} = -k^2 e^\theta,$$

(17) implies that

$$\tau(\theta) = \theta - \frac{1-k^2\theta e^\theta}{-k^2 e^\theta} = \frac{1}{k^2 e^\theta},$$

with inverse function $\tau^{-1}(y) = -\ln y - \ln k^2$ for $y > 0$. □

In fact, density # 6 specified by Keynes is the density of an exponential distribution. This can be shown by plugging in the normalising constant:

$$f(x;\theta) = k^2 e^\theta \exp\left(-k^2 e^\theta x\right), \quad x > 0,\ \theta > 0,$$

which is the density of an exponential distribution with parameter $k^2 e^\theta$. If k is unknown, then k^2 and θ are not identified. ML estimation of both θ and k^2 is thus impossible.

3.2.2 Inverse Gaussian distribution

Setting $C(\theta) = -k^2/\theta$ and $b(x) = -k^2/x - 3/2\ln x$ for $x \in \mathcal{X}$ and $\theta \in \Theta$ with $\mathcal{X} = \Theta = \mathbb{R}^+$ in (19) results in

$$\begin{aligned} f(x;\theta) &= K(\theta)\frac{1}{\sqrt{x^3}}\exp\left(-k^2(x-\theta)\frac{1}{\theta^2} - k^2\frac{1}{\theta} - k^2\frac{1}{x}\right) \\ &= K(\theta)\frac{1}{\sqrt{x^3}}\exp\left(-k^2\frac{x^2-2x\theta+\theta^2}{x\theta^2}\right) \\ &= K(\theta)\frac{1}{\sqrt{x^3}}\exp\left(-k^2\frac{(x-\theta)^2}{x\theta^2}\right). \end{aligned}$$

This is the density of the inverse Gaussian distribution (cf. Patel et al. (1976, page 155)). The normalising constant $K(\theta)$ does in fact not dependent on θ; for the usual parameterisation $k^2 = 1/(2\sigma^2)$, it is given by

$$K(\theta) = \frac{1}{\sqrt{2\pi\sigma^2}}.$$

As a consequence, similar to the normal distribution the arithmetic mean of the observations is the ML estimator of the unknown parameter θ. However, note that θ is not a *location* parameter of the inverse Gaussian distribution.

3.3 Geometric mean

Let $\mathcal{X} = \Theta \subseteq \mathbb{R}^+$. For $u(x) = \ln x$ with $x \in \mathbb{R}^+$ we obtain from (16) the general density

$$f(x;\theta) = K(\theta) \exp\left((\ln x - \ln\theta)\theta C'(\theta) + C(\theta) + b(x)\right). \tag{20}$$

The ML estimator of θ is the geometric mean of the observations if $K(\theta)$ does not depend on θ.

3.3.1 Density # 8: Pareto distribution

Plugging the choice $C(\theta) = -k^2\theta$ and $b(x) = 0$ into (20) leads to Keynes' density # 8,

$$f(x;\theta) = K(\theta) \exp\left(-(\ln x - \ln\theta)\theta k^2 - k^2\theta\right) = K(\theta) \left(\frac{\theta}{x}\right)^{k^2\theta} \exp\left(-k^2\theta\right).$$

Assuming the support to be $\mathcal{X} = \{x \in \mathbb{R} | x > 1\}$ guarantees that f can indeed feature the properties of a density.

Lemma 2 *Provided that $k^2 > 1/\theta$, the normalising constant in density # 8 is derived as*

$$K(\theta) = \left(k^2\theta - 1\right) \exp\left(k^2\theta(1 - \ln\theta)\right).$$

The ML estimator of θ is then given by

$$\hat{\theta}_{ML} = \frac{1}{k^2}\left(1 + \frac{1}{\ln\left(\prod_{i=1}^n X_i\right)^{1/n}}\right).$$

Proof: Using the substitution $y = \ln x$, we obtain

$$\int_1^\infty \exp\left(-k^2\theta\left(\ln x - \ln\theta + 1\right)\right) dx = \exp\left(-k^2\theta(1-\ln\theta)\right) \int_0^\infty e^y e^{-k^2\theta y} dy$$
$$= \exp\left(-k^2\theta(1-\ln\theta)\right) \frac{1}{k^2\theta} m_Y(1),$$

where $m_Y(t) = k^2\theta/(k^2\theta - t)$ with $t < k^2\theta$ represents the moment generating function of an exponential distribution with parameter $k^2\theta$. Therefore,

$$K(\theta) = \left(k^2\theta - 1\right)\exp\left(k^2\theta(1-\ln\theta)\right) = \frac{k^2\theta - 1}{\theta^{k^2\theta}}\exp\left(k^2\theta\right).$$

Again, the normalising constant $K(\theta)$ depends on the parameter θ. Combining

$$\frac{\partial \ln K(\theta)}{\partial\theta} = \frac{k^2}{k^2\theta - 1} - k^2\ln\theta \quad \text{and} \quad \frac{\partial(C'(\theta)/u'(\theta))}{\partial\theta} = -k^2$$

with (17) results in the transformation $\tau(\theta) = 1/(k^2\theta - 1)$ for $\theta > 1/k^2$; thus, $\tau^{-1}(y) = (1 + 1/y)/k^2$ for $y > 0$. □

Plugging the normalising constant into the density results in a Pareto density of the form

$$f(x;\theta) = (k^2\theta - 1)\frac{1}{x^{(k^2\theta-1)+1}}, \quad x > 1,$$

with parameter $k^2\theta - 1$. It is well known that the ML estimator of this parameter of the Pareto distribution is determined by the logarithm of the geometric mean; see for example Johnson et al. (1994, p. 581). The values of k^2 and θ cannot be identified.

3.3.2 Density # 9

Keynes suggested another distribution for which the geometric mean is supposedly the ML estimator of the unknown parameter θ. Setting $C(\theta) = k^2(\ln\theta)^2$ and $b(x) = -k^2(\ln x)^2$ in (20) results in Keynes' density # 9:

$$f(x;\theta) = K(\theta)\exp\left(-k^2\left(-2\ln x\ln\theta + 2(\ln\theta)^2 - (\ln\theta)^2 + (\ln x)^2\right)\right)$$
$$= K(\theta)\exp\left(-k^2\left(\ln x - \ln\theta\right)^2\right) = K(\theta)\exp\left(-k^2\left(\ln\frac{x}{\theta}\right)^2\right).$$

Due to the logarithm, the support $\mathcal{X}$ and the parameter domain Θ are both $\mathbb{R}^+$.

Lemma 3 *The normalising constant in density # 9 is*

$$K(\theta) = \frac{1}{\sqrt{\pi/k^2}} \frac{1}{\theta} \exp\left(-\frac{1}{4k^2}\right),$$

and the ML estimator of θ is given by

$$\hat{\theta}_{ML} = \left(\prod_{i=1}^{n} X_i\right)^{1/n} e^{-1/(2k^2)}.$$

Proof: We have

$$\begin{aligned} K(\theta) &= \left(\int_0^\infty \exp\left(-k^2 (\ln x - \ln\theta)^2\right) dx\right)^{-1} \\ &= \left(\int_{-\infty}^\infty \exp(y) \exp\left(-\frac{1}{2}\frac{1}{1/(2k^2)}(y - \ln\theta)^2\right)\right)^{-1} \\ &= \left(\sqrt{2\pi\frac{1}{2k^2}} m_Y(1)\right)^{-1}. \end{aligned}$$

In this expression, Y represents a random variable following a normal distribution with mean $\ln\theta$ and variance $1/(2k^2)$. $m_Y(1)$ denotes the value of its moment generating function evaluated at 1, which is given by

$$m_Y(1) = \exp\left(\ln\theta + \frac{1}{2}\frac{1}{2k^2}\right).$$

Combining results, we obtain

$$K(\theta) = \frac{1}{\sqrt{\pi/k^2}} \exp\left(-\left(\ln\theta + \frac{1}{2}\frac{1}{2k^2}\right)\right).$$

This normalising constant depends on θ. Plugging

$$\frac{\partial \ln K(\theta)}{\partial\theta} = -\frac{1}{\theta} \quad \text{and} \quad \frac{\partial(C'(\theta)/u'(\theta))}{\partial\theta} = 2k^2\frac{1}{\theta}$$

into (17) results in the transformation $\tau(\theta) = \ln\theta + 1/2k^2$ for $\theta > 0$ and the inverse transformation $\tau^{-1}(y) = e^y e^{-1/(2k^2)}$. $\square$

3.3.3 Lognormal distribution

Keynes (1911) noticed the similarity between his density # 9 and "Sir Donald McAlister's law of error." Nowadays, the distribution discussed by McAlister (1879) is better known as the "lognormal distribution."

In fact, starting out with (20) and choosing $C(\theta) = k^2(\ln\theta)^2$, $b(x) = -k^2(\ln x)^2 - \ln x$ leads to the lognormal density

$$\begin{aligned} f(x;\theta) &= K(\theta)\exp\left(-k^2\left(-2\ln x\ln\theta + 2(\ln\theta)^2 - (\ln\theta)^2 + (\ln x)^2\right) - \ln x\right) \\ &= \frac{K(\theta)}{x}\exp\left(-k^2\left(\ln x - \ln\theta\right)^2\right), \end{aligned}$$

where

$$K(\theta) = \frac{k}{\sqrt{\pi}}.$$

For the usual parameterisation $k = 1/(2\sigma^2)$ with $\sigma > 0$, we obtain $K(\theta) = 1/\sqrt{2\pi\sigma^2}$. Since the normalising constant does not depend on θ, the ML estimator of this parameter of interest is given by the geometric mean, which is a well-known result.

3.4 Harmonic mean

Let $\mathcal{X}$, $\Theta \subseteq \mathbb{R}^+$. Plugging $u(x) = 1/x$ into (16) results in

$$f(x;\theta) = K(\theta)\exp\left(-\left(\frac{1}{x} - \frac{1}{\theta}\right)\theta^2 C'(\theta) + C(\theta) + b(x)\right). \tag{21}$$

If the normalising constant is independent of θ, then this density is the most general form of a density leading to the harmonic mean as the ML estimator of θ.

Keynes' density # 10 is obtained when additionally choosing $C(\theta) = k^2\theta$ and $b(x) = -k^2x$:

$$\begin{aligned} f(x;\theta) &= K(\theta)\exp\left(-\left(\frac{1}{x} - \frac{1}{\theta}\right)\theta^2 k^2 + k^2\theta - k^2 x\right) \\ &= K(\theta)\exp\left(-\frac{k^2}{x}(x-\theta)^2\right) = K(\theta)e^{2k^2\theta}\exp\left(-k^2\left(\frac{\theta^2}{x} + x\right)\right). \end{aligned} \tag{22}$$

To determine the normalising constant, we assume the support to be $\mathbb{R}^+$ and make use of the similarity between this density and the density of the so-called generalised inverse Gaussian (GIG) distribution (see Johnson et al. (1994, p. 284f.)),

$$f(x;\psi,\chi,r) = \frac{(\psi/\chi)^{r/2}}{2K_r(\sqrt{\psi\chi})}x^{r-1}\exp\left(-\frac{1}{2}\left(\chi\frac{1}{x} + \psi x\right)\right) \tag{23}$$

for $x > 0$, $\psi, \chi > 0$. K_r denotes a modified Bessel function of the third kind.

Lemma 4 *The normalising constant in density # 10 is given by*

$$K(\theta) = \frac{1/\theta^2}{2K_1(2k^2\theta)} e^{-2k^2\theta}.$$

The ML estimator for $\tau(\theta)$ takes the form

$$\widehat{\tau(\theta)}_{ML} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i} = \left(\frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i}}\right)^{-1},$$

with transformation

$$\tau(\theta) = \frac{1}{\theta} - \frac{k}{\theta}\left(1 + \frac{\partial \ln K_1(2k^2\theta)}{\partial(2k^2\theta)}\right) - \frac{1}{2k\theta^2}.$$

Proof: Setting $r = 1$, $\psi = 2k^2$, $\chi = 2k^2\theta^2$ in (23) results in the following density of a GIG distribution:

$$f(x; k, r) = \frac{1/\theta}{2K_1(2k^2\theta)} \exp\left(-k^2\left(x + \frac{\theta^2}{x}\right)\right).$$

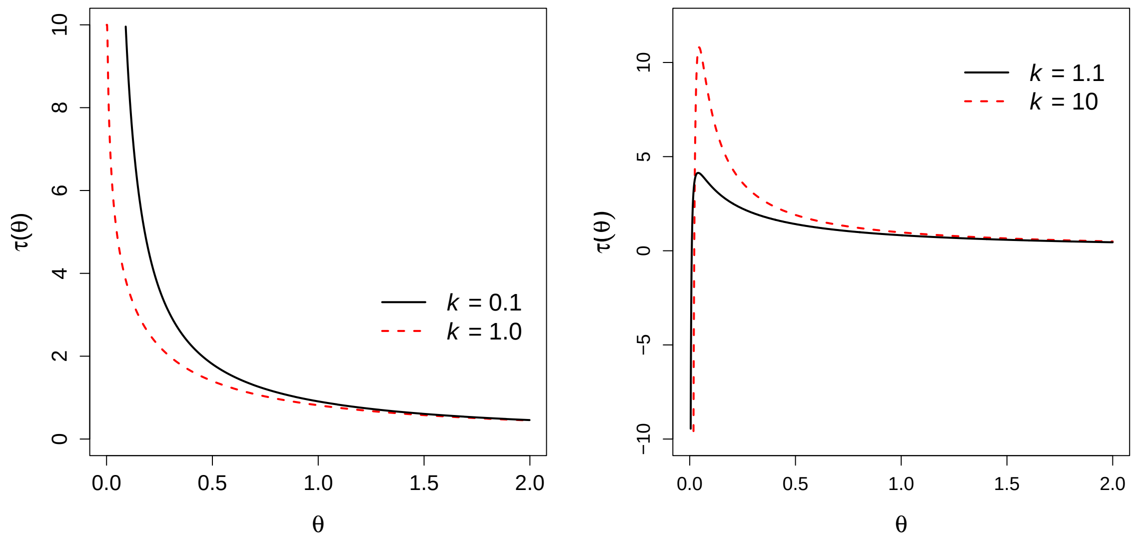
Comparing this density with (22) reveals that

$$K(\theta) = \frac{1/\theta}{2K_1(2k^2\theta)} e^{-2k^2\theta}.$$

This normalising constant does indeed depend on θ. According to (17),

$$\begin{aligned}
\tau(\theta) &= u(\theta) - \frac{\partial \ln K(\theta)}{\partial\theta}\frac{1}{\frac{\partial(C'(\theta)/u'(\theta))}{\partial\theta}} \\
&= \frac{1}{\theta} - \left[-2k^2\left(1 + \frac{\partial \ln K_1(2k^2\theta)}{\partial(2k^2\theta)}\right) - \frac{1}{\theta}\right]\cdot\frac{1}{-2k\theta} \\
&= \frac{1}{\theta} - \frac{k}{\theta}\left(1 + \frac{\partial \ln K_1(2k^2\theta)}{\partial(2k^2\theta)}\right) - \frac{1}{2k\theta^2}. \ \square
\end{aligned}$$

Note that due to the modified Bessel function of the third kind τ can only be calculated numerically. For two specific cases of $k \leq 1$, Figure 1(a) shows the strictly monotonic function $\tau(\theta)$. Figures 1(b) gives two examples for a non-monotonic function $\tau(\theta)$ with $k > 1$.



(a) $k \leq 1$ (b) $k > 1$

Figure 1: $\tau(\theta)$ for density # 10

3.5 Exponential mean

Let $\mathcal{X}$, $\Theta \subseteq \mathbb{R}$. Choosing $u(x) = e^{\gamma x}$ turns (16) into

$$f(x;\theta) = K(\theta) \exp\left(\left(e^{\gamma x} - e^{\gamma\theta}\right) \frac{C'(\theta)}{\gamma e^{\gamma\theta}} + C(\theta) + b(x) \right). \tag{24}$$

For densities of that form with $K(\theta)$ independent of θ, the ML estimator of this parameter of interest is given by the quasi-arithmetic mean

$$u^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} u(X_i)\right) = \frac{1}{\gamma} \ln\left(\frac{1}{n} \exp\left(\sum_{i=1}^{n} \gamma X_i\right)\right),$$

which is known as the "exponential mean."

For $\gamma \to 0$ the exponential mean converges towards the arithmetic mean. Due to the invariance of the quasi-arithmetic mean with respect to affine linear transformations, using the generator $v(x) = (e^{\gamma x} - 1)/\gamma$ instead of $u(x) = e^{\gamma x}$ leads to the same general density (24). Note that the transformation $v(x)$ was used by Hoaglin (1985) to generate a family of skewed distributions (the so-called "g distributions") from a normal distribution.

Further setting $C(\theta) = -\gamma r\theta$ and $b(x) = \gamma r x$ for $x, \theta \in \mathbb{R}$, $\gamma \neq 0$, we derive the density

$$\begin{aligned} f(x;\theta) &= K(\theta)\exp\left(\left(e^{\gamma x} - e^{\gamma\theta}\right)\frac{-\gamma r}{\gamma e^{\gamma\theta}} - \gamma r\theta + \gamma r x\right) \\ &= K(\theta)\exp\left(-re^{\gamma(x-\theta)} + r + \gamma r(x-\theta)\right) \end{aligned} \tag{25}$$

with

$$K(\theta) = e^{-r}|\gamma|\frac{r^r}{\Gamma(r)}, \tag{26}$$

provided that $r > -1$. The normalising constant is obtained immediately noticing that $e^{\gamma x}$ follows a gamma distribution. As this constant is independent of θ, the ML estimator of θ is given by the exponential mean. Plugging (26) into (25) results in the density derived by Dynkin (1961) and Ferguson (1962, 1963). The problem examined by these authors is similar to the one discussed by Keynes. They searched for the most general representation of an exponential family with location parameter θ and obtained

$$f(x;\theta) = |\gamma|\frac{r^r}{\Gamma(r)}\exp\left(-re^{\gamma(x-\theta)} + r\gamma(x-\theta)\right)$$

for $x \in \mathbb{R}$, $\gamma \neq 0$ and $r > -1$. Denny (1970) and Pfanzagl (1972) later formulated regularity conditions guaranteeing that this special exponential family can be characterised via the existence of a one-dimensional sufficient statistic. Moreover, Takeuchi (1973) and Bondesson (1975) showed that the members of this special exponential family are the only regular distributions for which an UMVUE for the location parameter exists. As we have seen, this family of distribution also results as one special case of the class of distributions for which the exponential mean of the observations is the ML estimator of the unknown parameter.

For $\gamma \to 0$, when the exponential mean approaches the arithmetic mean, (25) converges towards the normal distribution.

3.6 Power mean

Let $\mathcal{X} \subseteq \mathbb{R}$ and $\Theta \subseteq \mathbb{R}^+$. For $u(x) = |x|^\gamma$ density (16) becomes

$$f(x;\theta) = K(\theta)\exp\left(\left(|x|^\gamma - \theta^\gamma\right)\frac{C'(\theta)}{\gamma\theta^{\gamma-1}} + C(\theta) + b(x)\right). \tag{27}$$

Provided that the normalising constant is independent of θ, this is the most general form of a density for which the ML estimator of the unknown parameter is given by

$$u^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} u(X_i)\right) = \left(\frac{1}{n}\sum_{i=1}^{n} |X_i|^{\gamma}\right)^{1/\gamma},$$

a specific type of power mean.

Making the additional choice $C(\theta) = -r\gamma \ln \theta$ and $b(x) = r\gamma \ln |x|$ for $\theta > 0$, $x \in \mathbb{R}$, $\gamma \neq 0$, we obtain the density

$$f(x;\theta) = K(\theta) \exp\left((|x|^{\gamma} - \theta^{\gamma}) \frac{-r\gamma/\theta}{\gamma\theta^{\gamma-1}} - r\gamma \ln \theta + r\gamma \ln |x| \right). \tag{28}$$

Lemma 5 *The normalising constant in density (28) is given by*

$$K(\theta) = \frac{e^{-r}}{2} \frac{|\gamma|}{\theta} \frac{r^r}{\Gamma(r)}. \tag{29}$$

The ML estimator for θ takes the form

$$\hat{\theta}_{ML} = \left(\frac{1}{n}\sum_{i=1}^{n} |X_i|^{\gamma}\right)^{1/\gamma} \cdot \left(\frac{\gamma r}{\gamma r + 1}\right)^{1/\gamma}.$$

Proof: The normalising constant can be derived by noticing that (28) is a special case of the scale exponential family (see Ferguson (1962)). This family of distributions is obtained when trying to find a proper distribution with a scale parameter within the exponential family; its probability density function has the general form

$$f(x;\theta) = p_1 f_X(x;\theta) + p_2 f_{-X}(x;\theta)$$

with $p_1 + p_2 = 1$, $0 \leq p_1, p_2 \leq 1$, $\gamma \neq 0$, $\theta > 0$, $r > -1$, where

$$f_X(x;\theta) = \begin{cases} 0 & \text{for } x < 0, \\ \frac{|\gamma|}{\theta} \frac{r^r}{\Gamma(r)} \left(\frac{x}{\theta}\right)^{r\gamma-1} \exp\left(-r\left(\frac{x}{\theta}\right)^{\gamma}\right) & \text{for } x \geq 0, \end{cases}$$

and

$$f_{-X}(x;\theta) = \begin{cases} \frac{|\gamma|}{\theta} \frac{r^r}{\Gamma(r)} \left(\frac{|x|}{\theta}\right)^{r\gamma-1} \exp\left(-r\left(\frac{|x|}{\theta}\right)^{\gamma}\right) & \text{for } x < 0, \\ 0 & \text{for } x \geq 0. \end{cases}$$

Choosing $p_1 = p_2 = 1/2$ results in the special case

$$f(x;\theta) = \frac{1}{2} \frac{|\gamma|}{\theta} \frac{r^r}{\Gamma(r)} \left(\frac{|x|}{\theta}\right)^{r\gamma-1} \exp\left(-r\left(\frac{|x|}{\theta}\right)^{\gamma}\right)$$

for $x \in \mathbb{R}$, $\gamma \neq 0$, $\theta > 0$ and $r > -1$. Comparing this density with (28) it is immediately seen that the normalising constant in the latter density is (29).

Since this normalising constant depends on θ, we need to derive the transformation $\tau(\theta)$. This is done by plugging

$$\frac{\partial \ln K(\theta)}{\partial \theta} = -\frac{1}{\theta} \quad \text{and} \quad \frac{\partial (C'(\theta)/u'(\theta))}{\partial \theta} = \frac{\gamma r}{\theta^{\gamma+1}}$$

into (17), which results in $\tau(\theta) = \theta^{\gamma}(1 + 1/(\gamma r))$. The inverse transformation is therefore

$$\tau^{-1}(y) = \left(\frac{y\gamma r}{\gamma r + 1}\right)^{1/\gamma}. \qquad \square$$

4 Conclusions

Keynes (1911) derived densities for which the ML estimator of the unknown parameter is given by the arithmetic mean, the geometric mean, the harmonic mean, or the median. In this paper, we have refined his results, mainly by calculating the respective normalising constants of these densities. As a consequence, we have seen that in most cases the ML estimator is not a mean, but a function of a mean, because the normalising constant depends on the parameter to be estimated. Moreover, almost all densities have turned out to be related to simple distributions that are well-known today, like the Pareto distribution and the generalised inverse Gaussian distribution. Applying Keynes' approach to the class of quasi-arithmetic means, we have derived further general results, as well as specific distributions for which the ML estimator of the parameter of interest is (the function of) such a mean. Among these distributions are the location exponential family and a special case of the scale exponential family.

References

Aczél, J. (1966). *Lectures on Functional Equations and Their Applications.* Academic Press, New York.

Aldrich, J. (2008). Keynes among statisticians. *History of Political Economy* **40**, 265–316.

Bondesson, L. (1975). Uniformly minimum variance estimation in location parameter families. *Annals of Statistics* **3**, 637–660.

Box, G.E.P. & Cox, D.R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society, Series B* **26**, 211–252.

Box, G.E.P. & Cox, D.R. (1982). An analysis of transformations revisited, rebutted. *Journal of the American Statistical Association* **77**, 209–210.

Bullen, P.S., Mitrinović, D.S., Vasić, P.M. (1988). *Means and Their Inequalities.* D. Reidel, Dordrecht.

Conniffe, D. (1992). Keynes on probability and statistical inference and the links to Fisher. *Cambridge Journal of Economics* **16**, 475–489.

Denny, J.L. (1970). Cauchy's equation and sufficient statistics on arcwise connected spaces. *Annals of Mathematical Statistics* **41**, 401–411.

Dynkin, E.B. (1961). Necessary, and sufficient statistics for a family of probability distributions. *Selected Translations in Mathematical Statististics and Probability* **1**, 23–41.

Ferguson, T.S. (1962). Location and scale parameters in exponential families of distribution. *Annals of Mathematical Statistics* **33**, 986–1001.

Ferguson, T.S. (1963). Location and scale parameters in exponential families of distribution. Correction. *Annals of Mathematical Statistics* **34**, 1603.

Fisher, R.A. (1912). On an absolute criterion for fitting curves. *Messenger of Mathematics* **41**, 155–160.

Fisher, R.A. (1922). On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society London, Series A* **222**, 309–368.

Gauss, C.F. (1809). *Theoria motus corporum coelestiumin in sectionibus conicis solem ambientium.* F. Perthes & I.H. Besser, Hamburg.

Hoaglin, D.C. (1985). Summarizing shape numerically: The g-and-h distributions. In Hoaglin, D.C., Mosteller, F. & Tukey, J.W. (eds.). *Exploring Data Tables, Trends, and Shapes.* John Wiley & Sons, New York, 461–513.

Huber, P. (1964). Robust estimation of a location parameter. *Annals of Mathematical Statistics* **35**, 73–101.

Huber, P. (1981). *Robust Statistics.* John Wiley & Sons, New York.

Jarczyk, J. (2007). When Lagrangian and quasi-arithmetic mean coincide. *Journal of Inequalities in Pure and Applied Mathematics* **8**, 1–11.

Johnson, N.L., Kotz, S., Balakrishnan, N. (1994). *Continuous Univariate Distributions, Vol 1.* Wiley, New York.

Kendall, M.G., Stuart, A. (1967). *The Advanced Theory of Statistics.* Hafner, New York.

Keynes, J.M. (1907). *The Principles of Probability.* Submitted as a fellowship dissertation to King's College Cambridge. The Papers of John Maynard Keynes, King's College Archive Centre, Cambridge, JMK/TP/A/1-2.

Keynes, J.M. (1908). *The Principles of Probability.* Submitted as a fellowship dissertation to King's College Cambridge. The Papers of John Maynard Keynes, King's College Archive Centre, Cambridge, JMK/TP/A/1-2.

Keynes, J.M. (1911). The principal averages and the laws of error which lead to them. *Journal of the Royal Statistical Society.* **74**, 322–331.

Keynes, J.M. (1921). *A Treatise on Probability.* Macmillan, London.

McAlister, D. (1879). The law of the geometric mean. *Proceedings of the Royal Society* **29**, 367–376.

O'Donnell, R.M. (1989). *Keynes: Philosophy, Economics, and Politics.* Macmillan, London.

Patel, J.K., Kapadia, C.H., Owen, D.B. (1976). *Handbook of Statistical Distributions.* Marcel Dekker, New York & Basel.

Pfanzagl, J. (1972). Transformation groups and sufficient statistics. *Annals of Mathematical Statistics* **43**, 553–586.

Skidelsky, R. (1983). *John Maynard Keynes, vol. 1: Hopes Betrayed, 1883-1920.* Macmillan, London.

Skidelsky, R. (1992). *John Maynard Keynes, vol. 2: The Economist as Saviour, 1920-1937.* Macmillan, London.

Skidelsky, R. (2000). *John Maynard Keynes, vol. 3: Fighting for Britain, 1937-1946.* Macmillan, London.

Stigler, S.M. (1978). Francis Ysidro Edgeworth, statistician (with discussion). *Journal of the Royal Statistical Society* **141**, 287–322.

Stigler, S.M. (1986). *The History of Statistics. The Measurement of Uncertainty before 1900.* Harvard University Press, Cambridge.

Takeuchi, K. (1973). On location parameter family of distributions with uniformly minimum variance unbiased estimator of location. *Proc. of the Second Japan-USSR Symposium on Probability and Theory.* Lecture Notes in Mathematics, vol. 330, Springer, Berlin, 465–477.

Teicher, H. (1961). Maximum likelihood characterization of distributions. *Annals of Mathematical Statistics* **32**, 1214–1222.

Zehna, P.W. (1966). Invariance of maximum likelihood estimation. *Annals of Mathematical Statistics* **37**, 755.

Diskussionspapiere 2008
Discussion Papers 2008

01/2008 **Grimm, Veronika and Gregor Zoettl**: Strategic Capacity Choice under Uncertainty: The Impact of Market Structure on Investment and Welfare

02/2008 **Grimm, Veronika and Gregor Zoettl**: Production under Uncertainty: A Characterization of Welfare Enhancing and Optimal Price Caps

03/2008 **Engelmann, Dirk and Veronika Grimm**: Mechanisms for Efficient Voting with Private Information about Preferences

04/2008 **Schnabel, Claus and Joachim Wagner**: The Aging of the Unions in West Germany, 1980-2006

05/2008 **Wenzel, Tobias**: On the Incentives to Form Strategic Coalitions in ATM Markets

06/2008 **Herrmann, Klaus**: Models for Time-varying Moments Using Maximum Entropy Applied to a Generalized Measure of Volatility