J. Appl. Prob. **54,** 657–669 (2017)
*doi:*10.1017/jpr.2017.25

TAYLOR'S LAW, VIA RATIOS, FOR SOME DISTRIBUTIONS WITH INFINITE MEAN

MARK BROWN,* *Columbia University*
JOEL E. COHEN,** *The Rockefeller University and Columbia University*
VICTOR H. DE LA PEÑA,* *Columbia University*

Abstract

Taylor's law (TL) originated as an empirical pattern in ecology. In many sets of samples of population density, the variance of each sample was approximately proportional to a power of the mean of that sample. In a family of nonnegative random variables, TL asserts that the population variance is proportional to a power of the population mean. TL, sometimes called fluctuation scaling, holds widely in physics, ecology, finance, demography, epidemiology, and other sciences, and characterizes many classical probability distributions and stochastic processes such as branching processes and birth-and-death processes. We demonstrate analytically for the first time that a version of TL holds for a class of distributions with infinite mean. These distributions, a subset of stable laws, and the associated TL differ qualitatively from those of light-tailed distributions. Our results employ and contribute to the methodology of Albrecher and Teugels (2006) and Albrecher *et al.* (2010). This work opens a new domain of investigation for generalizations of TL.

Keywords: Taylor's law; stable distribution; Laplace transform; Tweedie distribution; heavy tail; cumulant generating function; approximate exponentiality; log linear regression

2010 Mathematics Subject Classification: Primary 60E07; 60F05
Secondary 62E20; 62P12; 62P35

1. Introduction

A block of observations is a set of independent observations from a given probability distribution. We assume different blocks are independent. Taylor (1961) and many others before and since 1961 showed empirically that, in a set of blocks of observations, each block coming from a different but related probability distribution, the sample variance of each block was approximately a power-law function of the corresponding sample mean of that block. Equivalently, there was an approximately linear relationship, across all blocks, between the log sample variance and the log sample mean. According to a survey by Eisler *et al.* (2008), more than a thousand papers have been published on the empirical support and theoretical foundations for what has become known as Taylor's law (TL). Many papers on TL have been published since 2008.

Our work contrasts with prior studies that may appear to be related. Since the pioneering explorations of Tweedie (1946), (1947), some statisticians have explored derivations of TL for

Received 14 January 2016; revision received 25 January 2017.
* Postal address: Department of Statistics, Columbia University, New York, NY 10027, USA.
** Postal address: Laboratory of Populations, The Rockefeller University, New York, NY 10065, USA.
Email address: cohen@mail.rockefeller.edu

probability distributions, such as stable laws of index β with $0 < \beta < 1$, which have infinite mean (for example, Jørgensen (1987), (1997), Jørgensen *et al.* (2009), Kendal (2004), (2013), and Kendal and Jørgensen (2011a), (2011b)). These studies sometimes assumed or imposed sequential dependence on observations, by contrast with our assumption that observations are independent. Kendal and Jørgensen used an exponential tilting to transform stable laws into light-tailed distributions with moments of all order. They obtained TL for families of such distributions with the tilting value serving as parameter.

By further contrast with the Tweedie stable-law approach, we work with the actual infinite-mean stable distributions. We show that a TL based on the sample mean and the sample variance of blocks of observations of a single infinite-mean nonnegative random variable holds for stable laws with $0 < \beta < 1$. Here the parameter is a discrete label, such as $1, 2, \dots,$ affixed to each block of observations.

We now sketch the main ideas. Suppose we have a family of nonnegative random variables $X(\theta) \geq 0$ indexed by a parameter θ and each $X(\theta)$ has finite second moment, population mean $\mu(\theta)$, and population variance $\sigma^2(\theta)$. If, for all θ, we have

$$\sigma^2(\theta) = a\mu^b(\theta) \tag{1}$$

with $a > 0$ and b independent of θ, then we have, by definition, a *population* TL.

If (1) holds with $\sigma^2(\theta) > 0$ for all θ, and if $\bar{X}_n(\theta)$ and $\hat{\sigma}_n^2(\theta)$ denote respectively the sample mean and sample variance based on a block of size n from $X(\theta)$, then

$$\log \hat{\sigma}_n^2(\theta) - b \log \bar{X}_n(\theta) \to \log a \tag{2}$$

almost surely (a.s.) for all θ as $n \to \infty$. Thus, the *population* TL (1) implies the *sample* TL (2).

When multiple blocks labeled $1, 2, \dots,$ are drawn from a *single* nonnegative distribution with finite population mean $\mu > 0$, population variance $\sigma^2 > 0$, and third central moment μ_3, obviously $\sigma^2 = (\sigma^2/\mu^b)\mu^b$ holds trivially for all real b, so the *population* TL (1) has little interest. However, because the sample mean and the sample variance are correlated, their logarithms are also correlated among the blocks. Consequently, the *sample* TL (2) holds for the blocks from a single distribution when the parameter that varies from block to block is the block label, not the underlying distribution, and in (2) the slope $b = \mu \times \mu_3/\sigma^4$ has the same sign as the skewness or third central moment μ_3 (Cohen and Xu (2015)). The parameters a and b of this sample TL are independent of the block labels, so we are justified in calling this relationship a kind of TL.

Here we show, surprisingly, that a sample TL (5), (6) holds if blocks of increasing size n are drawn from a *single* nonnegative distribution in a class of stable distributions with no mean. In this case, the limit on the right-hand side of (5) is a random variable with finite mean and variance, unlike the constant which is the limit on the right-hand side of (2).

Specifically, suppose that $X_1, X_2, \dots, X_n$ are independent and identically distributed (i.i.d.) random variables with the distribution of (henceforth abbreviated '$\sim$') X with Laplace transform

$$\mathcal{L}(s) = \mathbb{E}(\mathrm{e}^{-sX}) = \mathrm{e}^{-(cs)^\beta}, \qquad s \geq 0,\ 0 < \beta < 1,\ c > 0. \tag{3}$$

We label the distribution with Laplace transform (3) as $F(c, \beta)$. This stable distribution with index β has infinite mean. Consequently, $\bar{X}_n$ and $\hat{\sigma}_n^2$ are no longer sample estimators of population quantities. So we re-express the sample TL (2) in terms of a ratio of sample moments. For $\alpha > 0$, define

$$W_n(\alpha) = \frac{\hat{\sigma}_n^2}{\bar{X}_n^\alpha}.$$

https://doi.org/10.1017/jpr.2017.25 Published online by Cambridge University Press

For $X > 0$ with $\mathbb{E}X = \infty$, we define the single random variable X to follow a TL with exponent α if $W_n(\alpha)$ converges in distribution to a random variable W with $\operatorname{var}(\log W) < \infty$. Under this definition, $X \sim F(c, \beta)$ follows a TL with exponent $\alpha = (2-\beta)/(1-\beta)$. To avoid confusion, we emphasize that β is fixed. The parameter that varies from one block of observations to another is the block label.

Without loss of generality, take $c = 1$ in (3). We derive an identity (Proposition 2 of Section 4) from which it follows that $\mathbb{E}W_n(\alpha) \equiv \infty$ for $\alpha \le 2-\beta$, and, on the other hand, for $\alpha > 2-\beta$,

$$\mathbb{E}W_n(\alpha) = (1-\beta)\frac{\Gamma((\alpha+\beta-2)/\beta)}{\Gamma(\alpha)} n^{((1-\beta)/\beta)[(2-\beta)/(1-\beta)-\alpha]}. \tag{4}$$

Define $\alpha^* = (2-\beta)/(1-\beta)$. From (4),

$$\lim_{n\to\infty} \mathbb{E}W_n(\alpha) = \begin{cases} \infty & \text{for } 2-\beta < \alpha < \alpha^*, \\ 1-\beta & \text{for } \alpha = \alpha^*, \\ 0 & \text{for } \alpha > \alpha^*. \end{cases}$$

Moreover, in Proposition 2(ii), we show that, as $n \to \infty$,

$$\operatorname{var}[W_n(\alpha^*)] = (1-\beta)^2\left(1 + \frac{2\beta}{n-1}\right) \to (1-\beta)^2.$$

Applying a heavy-tailed bivariate limit theorem of Albrecher *et al.* (2010, Theorem 2.1), we demonstrate that $W_n(\alpha^*)$ converges in distribution as $n \to \infty$ to a random variable W such that $\operatorname{var} W = (\mathbb{E}W)^2 = (1-\beta)^2$ and $\operatorname{var}(\log W) < \infty$. Thus, an analog of (2) holds with constant β, while the parameter that varies from one block to another is the block label: as $n \to \infty$,

$$\log \hat{\sigma}_n^2 - \alpha^* \log \bar{X}_n \to \log W \tag{5}$$

in distribution, and, for large n,

$$\log \hat{\sigma}_n^2 \approx \alpha^* \log \bar{X}_n + \mathbb{E}(\log W) + e, \tag{6}$$

where $e = \log W - \mathbb{E}(\log W)$ has mean 0 and finite variance $\operatorname{var}(\log W)$. Consequently, as we show in Section 6, for any $\epsilon > 0$,

$$\mathbb{P}\left(\left|\frac{\log \hat{\sigma}_n^2}{\log \bar{X}_n} - \alpha^*\right| > \epsilon\right) \to 0,$$

so that, for large n, with high probability, $\log \hat{\sigma}_n^2 / \log \bar{X}_n$ will be close to $\alpha^* = (2-\beta)/(1-\beta)$. This linear relationship of the log sample variance $\log \hat{\sigma}_n^2$ to the log sample mean $\log \bar{X}_n$ is Taylor's law or power-law fluctuation scaling.

Since $\bar{X}_n \to \infty$ a.s., α^* is the unique exponent for which $W_n(\alpha)$ converges in distribution to a random variable W with $\mathbb{P}(0 < W < \infty) = 1$.

A referee kindly pointed out that if $X \sim F$ with survival function $\bar{F} = 1 - F$ satisfying

$$\lim_{t\to\infty} \frac{\bar{F}(t)}{l(t)t^{-\beta}} = 1, \qquad 0 < \beta < 1,$$

https://doi.org/10.1017/jpr.2017.25 Published online by Cambridge University Press

with l slowly varying and $\lim_{t\to\infty} l(t) = [\Gamma(1-\beta)]^{-1}$, then (5) and (6) hold with the same α^* and W as with $X \sim F(1, \beta)$. This issue is discussed in Section 4. The result does not extend, however, to general regularly varying tail distributions of index β. Thus, the TL for $F(1, \beta)$ also holds for a restricted family of regularly varying tail distributions of index β with a class of slowly varying functions $l \in \boldsymbol{L}$ serving as parameter. However, even for a single distribution with infinite mean, it is meaningful to describe the behavior of the sample mean and sample variance as a kind of TL.

Equation (6) gives an analog of (2) but with an asymptotic log-log linear regression with noise, represented by the term e, rather than an asymptotically perfect log-log linear relationship. Thus, the light-tailed and heavy-tailed distributions have differing versions of TL. In Section 6 we observe that, relative to $\log \bar{X}_n$, this noise vanishes in probability as $n \to \infty$.

The fact that $\mathbb{E}W = \mathrm{SD}(W)$ (where 'SD' denotes standard deviation) suggests the possibility that W is approximately exponentially distributed. This possibility is supported by our simulations. We conjecture that as $\beta \to 0$, $W(\beta)$ (the distribution of W corresponding to index β) converges in distribution to an exponential distribution. We discuss this issue in Section 5.

We hope that our efforts will encourage others to investigate TL for infinite-mean random variables empirically and mathematically. Given the fascinating history and extensive applications of TL, we believe that this is a potentially fruitful area of research.

2. A useful identity

The useful identity in this section is Proposition 1 below. It will be applied to obtain our main results. First we state an auxiliary result and some definitions. For $x > 0$, $\alpha > 0$,

$$\frac{1}{x^\alpha} = \frac{1}{\Gamma(\alpha)} \int_{\lambda=0}^{\infty} \lambda^{\alpha-1} \mathrm{e}^{-\lambda x}\, \mathrm{d}\lambda. \tag{7}$$

This follows since the probability density function of a gamma distribution with shape parameter α and scale parameter x integrates to 1.

Define C_r to be the class of distributions on $R_+ := (0, \infty)$ with finite rth moment. For a function T on R_+^n to R_+, we say that $T \in C_{r,n}$ if, for every $G \in C_r$ and i.i.d. $X_1, X_2, \dots, X_n$ distributed as G, we have $\mathbb{E}_G T(\underline{X}) := \mathbb{E}_G(T(X_1, X_2, \dots, X_n)) := \theta_T(G) < \infty$.

For a cumulative distribution function (CDF) F on R_+ with Laplace transform $\mathcal{L}$, and, for $\lambda > 0$, denote by F_λ the *tilted* version of F,

$$\mathrm{d}F_\lambda(x) = \frac{\mathrm{e}^{-\lambda x}\, \mathrm{d}F(x)}{\mathcal{L}(\lambda)}. \tag{8}$$

Tilted distributions are widely applied to areas including large deviations (Billingsley (1986, pp. 142–154)), exponential families (Barndorff-Nielsen (2014, pp. 103–137)), and simulation theory (Ross (2002, pp. 275–279)). One important property is that, for every $\lambda > 0$, $r > 0$, we have $F_\lambda \in C_r$. This is true because F_λ has a finite moment generating function on $(-\infty, \lambda)$. It is not necessary that F have any finite moments.

Proposition 1. *Suppose that $X_1, X_2, \dots, X_n$ are i.i.d. nonnegative random variables with distribution F and Laplace transform $\mathcal{L}$. Define $\bar{X}$ to be the sample mean. For $T \in C_{r,n}$, if either*

(i) $\mathbb{P}_F(X = 0) = 0$, *or*

(ii) $T(\underline{0}) = T(0, 0, \dots, 0) = 0$ *(with $T(\underline{0})/0$ interpreted as 0),*

https://doi.org/10.1017/jpr.2017.25 Published online by Cambridge University Press

then, for $\alpha > 0$,

$$\mathbb{E}_F\left(\frac{T(\underline{X})}{\bar{X}^\alpha}\right) = \frac{n^\alpha}{\Gamma(\alpha)} \int_0^\infty \lambda^{\alpha-1} \mathcal{L}^n(\lambda)\theta_T(F_\lambda)\,\mathrm{d}\lambda, \tag{9}$$

where $\theta_T(F_\lambda) = \mathbb{E}_{F_\lambda}(T(\underline{X}))$. *In (9), either both sides are infinite or both are finite and equal.*

Proof. First we derive the result analogous to (9), in (11) below, for the special case $S = \sum_{i=1}^n X_i$. The stated result then follows from (11) by multiplying both sides by n^α. Assume that (i) holds. By (7),

$$\mathbb{E}_F\left(\frac{T}{S^\alpha}\right) = \mathbb{E}_F\left(T\left(\frac{1}{\Gamma(\alpha)} \int_0^\infty \lambda^{\alpha-1} \mathrm{e}^{-\lambda S}\,\mathrm{d}\lambda\right)\right) = \frac{1}{\Gamma(\alpha)} \int_0^\infty \lambda^{\alpha-1} \mathbb{E}_F(T(\underline{X})\mathrm{e}^{-\lambda S})\,\mathrm{d}\lambda. \tag{10}$$

Since

$$\theta_T(F_\lambda) = \int T(\underline{x}) \frac{\mathrm{e}^{-\lambda S}}{\mathcal{L}^n(\lambda)}\,\mathrm{d}F(\underline{x}),$$

it follows that the right-hand side of (10) reduces to

$$\frac{1}{\Gamma(\alpha)} \int_0^\infty \lambda^{\alpha-1} \mathcal{L}^n(\lambda)\theta_T(F_\lambda)\,\mathrm{d}\lambda.$$

Thus,

$$\mathbb{E}_F\left(\frac{T}{S^\alpha}\right) = \frac{1}{\Gamma(\alpha)} \int_0^\infty \lambda^{\alpha-1} \mathcal{L}^n(\lambda)\theta_T(F_\lambda)\,\mathrm{d}\lambda, \tag{11}$$

from which (9) follows.

If (i) does not hold, but (ii) holds, then

$$\begin{aligned}
\mathbb{E}_F\left(\frac{T(\underline{X})}{S^\alpha}\right) &= \int_{\underline{x}\neq\underline{0}} \frac{T(\underline{x})}{S^\alpha}\left(\prod \mathrm{d}F(x_i)\right) \\
&= \frac{1}{\Gamma(\alpha)} \int_0^\infty \lambda^{\alpha-1} \mathbb{E}_F(T\mathrm{e}^{-\lambda S}\,\mathbf{1}_{\{\underline{x}\neq\underline{0}\}})\,\mathrm{d}\lambda \\
&= \frac{1}{\Gamma(\alpha)} \int_0^\infty \lambda^{\alpha-1} \mathbb{E}_F(T\mathrm{e}^{-\lambda S})\,\mathrm{d}\lambda \\
&= \frac{1}{\Gamma(\alpha)} \int_0^\infty \lambda^{\alpha-1} \mathcal{L}^n(\lambda)\theta_T(F_\lambda)\,\mathrm{d}\lambda,
\end{aligned}$$

since $T(\underline{0})\mathrm{e}^{-\lambda S(\underline{0})} = 0$. In the above, the interchange of order of integration holds by Tonelli's theorem, as the integrands are nonnegative. □

The above proof proceeded by reduction to the *Laplace–Tauber* case which involves a similar integration strategy. It is a key tool in the proof of Karamata's theorem (Karamata (1931); see Bingham *et al.* (1987, p. 4)). This approach is also used in de la Peña and Yang (1999). More closely related to our work, in the context of self-normalized processes, are Giné *et al.* (1997), Fuchs and Joffe (1997), Albrecher and Teugels (2006), and Albrecher *et al.* (2010). Proposition 1 is perhaps novel in pointing out the potential use of tilted distributions.

If neither (i) nor (ii) of Proposition 1 holds then defining $T(\underline{0}) = C > 0$, and $\mathbb{P}_F(T = 0) = p > 0$, it follows that $\mathbb{E}_F(T\mathrm{e}^{-\lambda S}) \geq Cp^n$, thus,

$$\int_0^\infty \lambda^{\alpha-1} \mathbb{E}_F(T\mathrm{e}^{-\lambda S})\,\mathrm{d}\lambda \geq Cp^n \int_0^\infty \lambda^{\alpha-1}\,\mathrm{d}\lambda = \infty.$$

https://doi.org/10.1017/jpr.2017.25 Published online by Cambridge University Press

In this case, both sides of (9) are infinite. Thus, (9) holds but is of little interest. We will examine situations for which the two sides of (9) are finite and equal in Section 7.

3. Application to ratios

As before, suppose that $X_1, X_2, \dots, X_n$ are i.i.d. nonnegative random variables with CDF F and Laplace transform $\mathcal{L}(s) = \mathbb{E}_F(\mathrm{e}^{-sX})$. We do not assume that F has any finite moments.

Define $K(s) = \log \mathcal{L}(s)$ and $K_r(s)$ to be the rth derivative of K at s. Then K is known as the cumulant generating function of F. For the tilted distribution F_λ,

$$\mathcal{L}_{F_\lambda}(s) = \mathbb{E}_{F_\lambda}(\mathrm{e}^{-sX}) = \frac{\mathcal{L}(s+\lambda)}{\mathcal{L}(\lambda)}.$$

The cumulant generating function of F_λ is, thus,

$$K_{F_\lambda} = \log \mathcal{L}_{F_\lambda}(s) = \log \mathcal{L}(s+\lambda) - \log \mathcal{L}(\lambda) = K(s+\lambda) - K(\lambda).$$

Consequently,

$$K_{2,F_\lambda}(s) = \frac{\mathrm{d}^2}{\mathrm{d}s^2} K_{F_\lambda}(s) = K_2(s+\lambda), \qquad \operatorname{var}_{F_\lambda}(X) = K_{2,F_\lambda}(0) = K_2(\lambda).$$

Since $\operatorname{var}_G(X) < \infty$ for $G \in C_2$ and F_λ has moments of all order, it follows that $\operatorname{var}_{F_\lambda}(X) < \infty$. It is well known that the variance of a random variable equals the second derivative of its cumulant generating function at 0 (Billingsley (1986, p. 144)).

In Proposition 1, choose

$$T_1(\underline{X}) = \hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2,$$

the sample variance. Since, for $G \in C_2$, the sample variance is unbiased for $\sigma^2(G)$,

$$\theta_{T_1}(F_\lambda) = \mathbb{E}_{F_\lambda}(\hat{\sigma}^2) = \operatorname{var}_{F_\lambda}(X) = K_{2,F_\lambda}(0) = K_2(\lambda).$$

Applying Proposition 1 to $T_1 = \hat{\sigma}^2$,

$$\mathbb{E}_F(W_n(\alpha)) = \frac{n^\alpha}{\Gamma(\alpha)} \int_0^\infty \lambda^{\alpha-1} K_2(\lambda) \mathcal{L}^n(\lambda)\, \mathrm{d}\lambda. \tag{12}$$

Another well-known result (Neter *et al.* (1990)) is that if $G \in C_4$ then

$$\mathbb{E}_G((\hat{\sigma}^2)^2) = \mathbb{E}_G(\hat{\sigma}^4) = \frac{K_{4,G}(0)}{n} + \frac{n+1}{n-1} K_{2,G}^2(0).$$

Letting $G = F_\lambda$, it follows that

$$\mathbb{E}_{F_\lambda}(\hat{\sigma}^4) = \frac{K_4(\lambda)}{n} + \frac{n+1}{n-1} K_2^2(\lambda).$$

Thus, if we define

$$T_2(\underline{X}) = \hat{\sigma}^4 = \left(\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2 \right)^2,$$

https://doi.org/10.1017/jpr.2017.25 Published online by Cambridge University Press

then

$$\theta_{T_2}(F_\lambda) = \frac{K_4(\lambda)}{n} + \frac{n+1}{n-1} K_2^2(\lambda),$$

and from Proposition 1,

$$\mathbb{E}_F([W_n(\alpha)]^2) = \frac{n^{2\alpha}}{\Gamma(2\alpha)} \int_0^\infty \lambda^{2\alpha-1} \bigg(\frac{K_4(\lambda)}{n} + \frac{n+1}{n-1} K_2^2(\lambda) \bigg) \mathcal{L}^n(\lambda) \, \mathrm{d}\lambda. \tag{13}$$

From (12) and (13), we can compute the variance of $W_n(\alpha)$.

4. Application to stable distributions

We will work with the family of distributions with Laplace transform

$$\mathcal{L}(\lambda) = \mathrm{e}^{-(c\lambda)^\beta}, \qquad c > 0,\ 0 < \beta < 1. \tag{14}$$

This widely studied class of stable laws encompasses all stable laws with support $[0, \infty)$. Feller (1971, pp. 448–449) is still a good reference for this topic. Distributions in this class have infinite mean and, with the exception of $\beta = \frac{1}{2}$, unwieldy probability density functions. For $\beta = \frac{1}{2}$, the distribution with Laplace transform (14) is the first-passage time in standard Brownian motion from 0 to the point $\sqrt{c/2}$. It is distributed as $c/(2Z^2)$, where Z is standard normal.

Denote the distribution with Laplace transform (14) for a given c and β as $F(c, \beta)$.

Proposition 2. *Let $X \sim F(c, \beta)$ and define $\alpha^* = \alpha^*(\beta) = (2-\beta)/(1-\beta)$ and $W_n = \hat{\sigma}_n^2 / \bar{X}_n^{\alpha^*}$ corresponding to $X_1, X_2, \ldots, X_n$. Then*

(i) $\mathbb{E} W_n = c^{-(\beta/(1-\beta))}(1 - \beta)$,

(ii) $\operatorname{var}(W_n) = (\mathbb{E} W_n)^2 (1 + 2\beta/(n-1))$,

(iii) $\sup_n \mathbb{E} W_n^k < \infty$, *for all* $k \geq 1$,

(iv) *W_n converges in distribution to W, with $\mathbb{E} W^k < \infty$ for all $k \geq 1$, and $\lim_{n\to\infty} \mathbb{E} W_n^k = \mathbb{E} W^k$. Moreover,* $\mathrm{SD}(W) = \mathbb{E} W = c^{-(\beta/(1-\beta))}(1 - \beta)$.

(v) $\mathbb{E}[(\log W)^2] < \infty$.

Proof. If $X \sim F(1, \beta)$ then

$$cX \sim F(c, \beta) \quad \text{and} \quad W_n(cX_1, cX_2, \ldots, cX_n) = c^{-\beta/(1-\beta)} W_n(X_1, X_2, \ldots, X_n).$$

Thus (i) and (ii) will follow from the corresponding result for $F(1, \beta)$. From (14),

$$K_2(\lambda) = \beta(1 - \beta)\lambda^{\beta-2}.$$

Thus, from (11) (recalling that $\alpha^* = (2 - \beta)/(1 - \beta)$),

$$\mathbb{E} W_n = \frac{n^{\alpha^*}}{\Gamma(\alpha^*)} \int \lambda^{\alpha^*-1} (\beta(1-\beta)\lambda^{\beta-2}) \mathrm{e}^{-n\lambda^\beta} \, \mathrm{d}\lambda.$$

Make the change of variable $\lambda = z^{1/\beta}$. Then using $(\alpha^* - 2)/\beta = \alpha^* - 1$, we obtain

$$\mathbb{E} W_n = (1 - \beta) \frac{n^{\alpha^*}}{\Gamma(\alpha^*)} \int z^{\alpha^*-1} \mathrm{e}^{-nz} \, \mathrm{d}z = 1 - \beta,$$

which proves (i).

https://doi.org/10.1017/jpr.2017.25 Published online by Cambridge University Press

From (14) with $c = 1$, we calculate

$$K_4(\lambda) = \beta(1-\beta)(2-\beta)(3-\beta)\lambda^{\beta-4}.$$

Then, from (13), $\mathbb{E}W_n^2 = A + B$, where

$$\begin{aligned} A &= \frac{n^{2\alpha^*}}{\Gamma(2\alpha^*)} \int \lambda^{2\alpha^*-1}\left(\frac{K_4(\lambda)}{n}\right) \mathrm{e}^{-n\lambda^\beta}\,\mathrm{d}\lambda \\ &= \beta(1-\beta)(2-\beta)(3-\beta)\frac{n^{2\alpha^*-1}}{\Gamma(2\alpha^*)} \int \lambda^{2\alpha^*-1}\lambda^{\beta-4}\mathrm{e}^{-n\lambda^\beta}\,\mathrm{d}\lambda. \end{aligned}$$

Again with the same change of variable $z = \lambda^\beta$,

$$\begin{aligned} A &= (1-\beta)(2-\beta)(3-\beta)\frac{n^{2\alpha^*-1}}{\Gamma(2\alpha^*)} \int z^{2(\alpha^*-1)}\mathrm{e}^{-nz}\,\mathrm{d}z \\ &= (1-\beta)(2-\beta)(3-\beta)\frac{n^{2\alpha^*-1}}{\Gamma(2\alpha^*)}\frac{\Gamma(2\alpha^*-1)}{n^{2\alpha^*-1}} \\ &= \frac{(1-\beta)(2-\beta)(3-\beta)}{2\alpha^*-1} \\ &= (1-\beta)^2(2-\beta) \end{aligned}$$

and

$$\begin{aligned} B &= \beta^2(1-\beta)^2\left(\frac{n+1}{n-1}\right)\frac{n^{2\alpha^*}}{\Gamma(2\alpha^*)} \int \lambda^{2\alpha^*-1}\lambda^{2\beta-4}\mathrm{e}^{-n\lambda^\beta}\,\mathrm{d}\lambda \\ &= \beta(1-\beta)^2\left(\frac{n+1}{n-1}\right)\frac{n^{2\alpha^*}}{\Gamma(2\alpha^*)} \int z^{2\alpha^*-1}\mathrm{e}^{-nz}\,\mathrm{d}z \\ &= \beta(1-\beta)^2\left(\frac{n+1}{n-1}\right). \end{aligned}$$

Combining, we have

$$\mathbb{E}W_n^2 = A + B = (1-\beta)^2\left(2 + \frac{2}{n-1}\beta\right). \tag{15}$$

From (i) and (15),

$$\operatorname{var}(W_n) = \mathbb{E}W_n^2 - (\mathbb{E}W_n)^2 = (\mathbb{E}W_n)^2\left(1 + \frac{2\beta}{n-1}\right),$$

which proves (ii).

For (iii), define

$$V_n = \frac{\hat{\sigma}^2}{n\bar{X}^2} = \frac{n\hat{\sigma}^2}{S^2} = \frac{n}{n-1}\left(\sum_{i=1}^n\left(\frac{X_i}{S}\right)^2 - \frac{1}{n}\right),$$

where $S = \sum_{i=1}^n X_i$. Define $p_i = X_i/S$. Then $p_i > 0$, $\sum_{i=1}^n p_i = 1$, and

$$V_n = \frac{n}{n-1}\sum_{i=1}^n\left(p_i - \frac{1}{n}\right)^2.$$

https://doi.org/10.1017/jpr.2017.25 Published online by Cambridge University Press

Then V_n is bounded above by 1, equaling 1 if a single p_i equals 1, with the others being 0. (1 is the least upper bound for V_n, but is not achievable.) Recalling that $\alpha^* = (2-\beta)/(1-\beta)$, observe that

$$W_n = \frac{n^{\alpha^*}\hat{\sigma}^2}{S^{\alpha^*}} = \left(\frac{n\hat{\sigma}^2}{S^2}\right)\left(\frac{n^{1/(1-\beta)}}{S^{\beta/(1-\beta)}}\right) = V_n\left(\frac{n^{1/(1-\beta)}}{S^{\beta/(1-\beta)}}\right). \tag{16}$$

Now $S = \sum_{i=1}^n X_i \sim n^{1/\beta}X$ (check by Laplace transform), thus, $S^{\beta/(1-\beta)} \sim (n^{1/\beta}X)^{\beta/(1-\beta)}$. From (16),

$$W_n \sim V_n X^{-\beta/(1-\beta)}. \tag{17}$$

Of course, V_n and $X^{-\beta/(1-\beta)}$ are dependent. Despite this limitation of (17), we find it useful for proving Proposition 2(iii). By (7), for all $k \geq 1$,

$$\begin{aligned}\mathbb{E}(X^{-k\beta/(1-\beta)}) &= \frac{1}{\Gamma(k\beta/(1-\beta))}\int \lambda^{k\beta/(1-\beta)-1}\mathrm{e}^{-\lambda^\beta}\,\mathrm{d}\lambda \\ &= \frac{1}{\beta\Gamma(k\beta/(1-\beta))}\int z^{k/(1-\beta)-1}\mathrm{e}^{-z}\,\mathrm{d}z \\ &= \frac{\Gamma(k/(1-\beta))}{\beta\Gamma(k\beta/(1-\beta))} \\ &< \infty.\end{aligned}$$

Thus,

$$\sup_n \mathbb{E}W_n^k \leq \mathbb{E}(X^{-k\beta/(1-\beta)}) < \infty. \tag{18}$$

(iv) It follows from (18) that if W_n converges in distribution to W, then by uniform integrability

$$\mathbb{E}W^k = \lim_{n\to\infty}\mathbb{E}W_n^k.$$

For convergence in distribution, we review Albrecher *et al.* (2010, Theorem 2.1). Consider $X > 0$ with survival function $\bar{F} = 1 - F$ satisfying

$$\lim_{t\to\infty}\frac{\bar{F}(t)}{l(t)t^{-\beta}} = 1, \qquad 0 < \beta < 1, \tag{19}$$

where l is slowly varying. Define $a_n = \bar{F}^{-1}(1/n)$, the upper $1/n$ quantile. Then

$$(U_n, V_n) := \left(\frac{1}{a_n^2}\sum_{i=1}^n X_i^2, \frac{1}{a_n}\sum_{i=1}^n X_i\right) \to (U, V)$$

in distribution as $n \to \infty$. The joint Laplace transform of (U, V) is derived (Albrecher *et al.* (2010, p. 6)); U and V are dependent, each distributed according to a stable law:

$$U \sim F\left(\left(\Gamma\left(1-\frac{\beta}{2}\right)\right)^{2/\beta}, \frac{\beta}{2}\right), \tag{20}$$

$$V \sim F((\Gamma(1-\beta))^{1/\beta}, \beta). \tag{21}$$

For $X \sim F(1, \beta)$ (Feller (1971, p. 448)),

$$\lim_{x\to\infty} x^\beta \bar{F}(x) = (\Gamma(1-\beta))^{-1}.$$

https://doi.org/10.1017/jpr.2017.25 Published online by Cambridge University Press

It follows that, if $X \sim F(1, \beta)$ then

$$\frac{a_n}{(\Gamma(1-\beta))^{-1/\beta} n^{1/\beta}} \to 1 \quad \text{as } n \to \infty,$$

$$W_n(\alpha^*) = \frac{n}{n-1}\bigg[\frac{U_n - V_n^2/n}{V_n^{\alpha^*}}\bigg](\Gamma(1-\beta))^{1/(1-\beta)} \to (\Gamma(1-\beta))^{1/(1-\beta)}\bigg(\frac{U}{V^{\alpha^*}}\bigg)$$

in distribution as $n \to \infty$.

(v) If $X \sim F(c, \beta)$ then, since $\mathbb{E}X^\delta < \infty$ for $-\infty < \delta < \beta$ (Feller (1971, p. 578)), it follows that $\mathbb{E}([\log X]^2) < \infty$. This is true because, for $0 < \delta_1 < \beta$, $(\log x)^2 < x^{\delta_1}$ for all sufficiently large x, and, for $\delta_2 < 0$, $(\log x)^2 < x^{\delta_2}$ for all sufficiently small positive x. Then, by (20) and (21),

$$\mathbb{E}\log^2\bigg(\frac{U}{V^{\alpha^*}}\bigg) \le \mathbb{E}\log^2 U + \alpha^{*2}\mathbb{E}\log^2 V + 2\alpha^*\sqrt{(\mathbb{E}\log^2 U)(\mathbb{E}\log^2 V)} < \infty. \qquad \square$$

It follows from the above discussion of the limit theorem of Albrecher *et al.* (2010) that, if (19) holds with l slowly varying and $\lim_{t\to\infty} l(t) = d$, then $W_n(\alpha^*)$ has the same limiting distribution for $X \sim F$ as for $X \sim F([d \times \Gamma(1-\beta)]^{1/\beta}, \beta)$. This limiting distribution is $d^{-(\beta/(1-\beta))}(U/V^{\alpha^*})$. This justifies the remark in Section 1 that, if $d = (\Gamma(1-\beta))^{-1}$, then $W_n(\alpha^*)$ has the same limiting distribution as it would with $X \sim F(1, \beta)$, and, hence, that the TL for $F(1, \beta)$ also holds with l serving as parameter if l is slowly varying with $\lim_{t\to\infty} l(t) = [\Gamma(1-\beta)]^{-1}$ and the same β. It also follows that

$$\mathrm{SD}\bigg(\frac{U}{V^{\alpha^*}}\bigg) = \mathbb{E}\bigg(\frac{U}{V^{\alpha^*}}\bigg) = (1-\beta)(\Gamma(1-\beta))^{1/(1-\beta)}.$$

5. Approximate exponentiality

Since, by Proposition 2,

$$\operatorname{var}(W_n) = (\mathbb{E}W_n)^2\bigg(1 + \frac{2\beta}{n-1}\bigg), \qquad \operatorname{var}(W) = (\mathbb{E}W)^2,$$

we thought it possible that W (and, thus, W_n for large n) would be approximately exponentially distributed. A tedious calculation of $\mathbb{E}(\hat{\sigma}^6/\bar{X}^{3\alpha^*})$ for $\alpha^* = (2-\beta)/(1-\beta)$ shows that

$$\mathbb{E}W^3 = \bigg(6 - \frac{\beta}{5-2\beta}\bigg)(\mathbb{E}W)^3,$$

while an exponentially distributed random variable Y has $\mathbb{E}Y^3 = 6(\mathbb{E}Y)^3$. We define the quantity $\beta/(5-2\beta)$ to be the *shortfall*. For example, if $\beta = \frac{1}{2}$ then $\mathbb{E}W^3 = (5\frac{7}{8})(\mathbb{E}W)^3$ and the shortfall equals $\frac{1}{8}$.

In simulations with β from 0.125 to 0.875 in increments of 0.125 using various values of n, we found the distribution of W_n to be close to that of an exponential distribution with the same mean. The simulated CDF of W_n starts off larger than that of Y (an exponential with the simulated mean $\mathbb{E}W_n$), then the CDFs cross and the CDF of Y is larger than that of the simulated W_n for an interval, then the CDFs cross again and are close from that point on. The Kolmogorov distance ($\sup_t |\mathbb{P}(W_n \le t) - \mathbb{P}(Y \le t)|$) is achieved in the initial interval (where $\mathbb{P}(W_n \le t) > \mathbb{P}(Y \le t)$). In these simulations, the Kolmogorov distance to exponentiality is

https://doi.org/10.1017/jpr.2017.25 Published online by Cambridge University Press

roughly 0.4 times the shortfall. For example, with $\beta = \frac{1}{2}$, the shortfall is equal to $\frac{1}{8}$ and the Kolmogorov distance is approximately $\frac{1}{20}$. We plan to do more simulations to clarify these issues.

For each $\beta \in (0, 1)$, we define $W(\beta) \equiv \lim_{n\to\infty} W_n(\beta)$ corresponding to $X \sim F(1, \beta)$. We conjecture that the distribution of $W(\beta)$ is exponential with mean 1. From (21), we see that this conjecture is equivalent to the statement that $U(\beta)/(V(\beta))^{\alpha^*}$ converges to an exponential distribution with mean 1 as $\beta \to 0$, where $(U(\beta), V(\beta))$ is the distribution of (U, V) corresponding to β. If this conjecture is true, we know of no interpretation of the exponential distribution as being the distribution of $\hat{\sigma}^2/\bar{X}^{\alpha^*}$ corresponding to any distribution, because $F(1, \beta)$ does not converge in distribution to a proper random variable as $\beta \to 0$, though as $\beta \to 0$, $\alpha^* = (2-\beta)/(1-\beta) \to 2$.

If W were exponential with mean $c^{-1/(1-\beta)}(1-\beta)$, then $\mathbb{E}\log W$ would be equal to $\log(c^{-1/(1-\beta)}(1-\beta)) - \gamma$, where γ is the Euler–Mascheroni constant (≈ 0.5772), and $\operatorname{var}(\log W)$ would be equal to $\pi^2/6$, while $-\log W$ would be Gumbel distributed.

Although W is approximately exponential, this may not be well reflected in $\mathbb{E}\log W$ or $\operatorname{var}(\log W)$, as the log is sensitive to values close to 0.

6. Connection to Taylor's law

Consider a block of n observations from the distribution $F(c, \beta)$ with Laplace transform (3). From Proposition 2(iv) we see that, for $W_n = \hat{\sigma}_n^2/\bar{X}_n^{\alpha^*}$ and $\alpha^* = (2-\beta)/(1-\beta)$,

$$\log W_n = \log \hat{\sigma}_n^2 - \alpha^* \log \bar{X}_n \to \log W$$

in distribution as $n \to \infty$. By the strong law of large numbers for random variables with infinite mean (Çinlar (2011, Proposition 6.3, p. 119)), $\bar{X}_n \to \infty$ a.s. and, hence, $\log \bar{X}_n \to \infty$ a.s. Now divide the displayed equation and limit above by $\log \bar{X}_n$. By a variant of Slutsky's theorem (Arnold (1990, Corollary 6.8(c), p. 242)), it follows that

$$\frac{\log W_n}{\log \bar{X}_n} = \frac{\log \hat{\sigma}_n^2}{\log \bar{X}_n} - \frac{2-\beta}{1-\beta} \to 0$$

in probability. Thus, for any $\epsilon > 0$,

$$\mathbb{P}\left(\left|\frac{\log \hat{\sigma}_n^2}{\log \bar{X}_n} - \alpha^*\right| > \epsilon\right) \to 0,$$

so that, for large n, with high probability, $\log \hat{\sigma}_n^2/\log \bar{X}_n$ will be close to $\alpha^* = (2-\beta)/(1-\beta)$. This is a form of Taylor's law.

7. Sufficient conditions for finite expectations

Referring to the definition of T in Section 2, we first examine $T = 1$. By Proposition 1,

$$\mathbb{E}((\bar{X})^{-\alpha}) = \frac{n^\alpha}{\Gamma(\alpha)} \int_0^\infty \lambda^{\alpha-1} \mathcal{L}^n(\lambda)\, \mathrm{d}\lambda.$$

For $\alpha > 0$, the integral will not blow up at 0 as $\lambda^{\alpha-1}$ is integrable in a neighborhood of 0, and $\mathcal{L}(\lambda) \to \mathcal{L}(0) = 1$ as $\lambda \to 0$. The behavior of $\mathcal{L}^n(\lambda)$ for large λ will determine whether the integral converges. In Proposition 3 we have simple checkable sufficient conditions on α and n for $\mathbb{E}((\bar{X})^{-\alpha})$ to be finite.

Proposition 3. *If, for some $\delta > 0$,*

$$\lim_{\lambda \to \infty} (\lambda^{\delta} \mathcal{L}(\lambda)) = 0, \tag{22}$$

then $\mathbb{E}((\bar{X})^{-\alpha}) < \infty$ for $\alpha > 0$ and $n \geq 1 + [\alpha/\delta]$, where $[x]$ is the greatest integer $\leq x$. Thus, for any $\alpha > 0$, $\mathbb{E}((\bar{X})^{-\alpha})$ is finite for all sufficiently large n.

Proof. Given $\alpha > 0$, for a unique nonnegative integer K, $K\delta \leq \alpha < (K+1)\delta$. This K has nothing to do with the K of Section 3. Define $n_0 = K + 1$ and $\varepsilon = (K+1)\delta - \alpha > 0$. For $n \geq n_0 = K + 1$,

$$\frac{\alpha + \varepsilon}{n} \leq \frac{\alpha + \varepsilon}{K + 1} = \delta.$$

Thus, for $n \geq n_0$ and $\lambda \geq 1$, by (22),

$$\lambda^{\alpha+\varepsilon} \mathcal{L}^n(\lambda) = (\lambda^{(\alpha+\varepsilon)/n} \mathcal{L}(\lambda))^n \leq (\lambda^{\delta} \mathcal{L}(\lambda))^n \to 0 \quad \text{as } \lambda \to \infty.$$

Next, since

$$\lambda^{\alpha+\varepsilon} \mathcal{L}^n(\lambda) = \frac{\lambda^{\alpha-1} \mathcal{L}^n(\lambda)}{\lambda^{-(1+\varepsilon)}} \to 0 \quad \text{as } \lambda \to \infty,$$

there exists λ_0 such that $\lambda \geq \lambda_0$ implies that

$$\lambda^{\alpha-1} \mathcal{L}^n(\lambda) \leq \lambda^{-(1+\varepsilon)}.$$

Consequently, for $\lambda \geq \lambda_0$,

$$\int_{\lambda}^{\infty} s^{\alpha-1} \mathcal{L}^n(s) \, \mathrm{d}s \leq \int_{\lambda}^{\infty} s^{-(1+\varepsilon)} \, \mathrm{d}s = \frac{1}{\varepsilon \lambda^{\varepsilon}} \to 0 \quad \text{as } \lambda \to \infty.$$

Thus, $\lambda^{\alpha-1} \mathcal{L}^n(\lambda)$ is integrable and

$$\mathbb{E}((\bar{X})^{-\alpha}) = \frac{n^{\alpha}}{\Gamma(\alpha)} \int_0^{\infty} \lambda^{\alpha-1} \mathcal{L}^n(\lambda) \, \mathrm{d}\lambda < \infty. \qquad \square$$

An interpretation for F_{λ} which follows from (8) is that F_{λ} is the conditional distribution of $X \sim F$ given $X \leq \varepsilon/\lambda$, where ε is independent of X and is exponential with mean 1. As $\lambda \to \infty$, F_{λ} converges to a point distribution at 0, while as $\lambda \to 0$, $F_{\lambda} \to F$. For $F \in C_2$, $K_2(\lambda) \to K_2(0) = \operatorname{var}_F(X) < \infty$. Thus, in this case the presence of $K_2(\lambda)$ in (12) does not affect the integrability. Under (22), it follows that, for $n \geq 1 + [\alpha/\delta]$, we have $\mathbb{E}(W_n(\alpha)) < \infty$. Similarly, if $F \in C_4$ then, for $n \geq 1 + [2\alpha/\delta]$, we have $\operatorname{var}(W_n(\alpha)) < \infty$.

In the stable case examined in Section 4, $K_2(\lambda) = \beta(1-\beta)\lambda^{\beta-2}$. The presence of this factor in (12) can cause the integral in (12) to blow up at $\lambda = 0$ even though $\mathbb{E}((\bar{X}_n)^{-\alpha})$ is finite. However, for sufficiently large α, $\mathbb{E}W_n(\alpha)$ will be finite. A similar remark holds for $\operatorname{var}(W_n(\alpha))$. From Proposition 1, we find that $\mathbb{E}W_n(\alpha) < \infty$ if and only if $\alpha > 2 - \beta$, and $\operatorname{var}(W_n(\alpha)) < \infty$ if and only if $\alpha > 2 - \beta/2$. This result is independent of the value of $n > 1$. For all $\alpha > 0$ and $n \geq 1$, $\mathbb{E}(\bar{X}_n)^{-\alpha} = \beta^{-1}(\Gamma(\alpha/\beta)/\Gamma(\alpha))n^{\alpha(1-1/\beta)} < \infty$, as follows from (7).

For example, suppose that $X_1, X_2, \dots, X_n$ are i.i.d. with stable distribution $F(c, \beta)$ with index $\beta = \frac{1}{2}$. Then $\mathbb{E}((\bar{X}_n)^{-\alpha}) < \infty$ for all $\alpha > 0$, $n \geq 1$. For $0 < \alpha \leq \frac{3}{2}$ and $n \geq 2$, $\mathbb{E}((\bar{X}_n)^{-\alpha}) < \infty$ but $\mathbb{E}(W_n(\alpha)) = \infty$. For $\frac{3}{2} < \alpha \leq \frac{7}{4}$, $\mathbb{E}(W_n(\alpha)) < \infty$ but $\operatorname{var}(W_n(\alpha))$ is infinite. For $\alpha > \frac{7}{4}$, both the mean and variance of $W_n(\alpha)$ are finite.

Acknowledgements

J. E. C. acknowledges with thanks the support of the U.S. National Science Foundation (grant number DMS-1225529) and the assistance of Priscilla K. Rogerson. All authors thank Xiaopei Zhang for performing the simulations summarized here and Xiaopei Zhang and Weichi Yao for help in preparing the manuscript. The paper has benefitted from the insightful comments of two anonymous referees, whom we gratefully acknowledge.

References

Albrecher, H. and Teugels, J. L. (2006). Asymptotic analysis of a measure of variation. *Theory Prob. Math. Statist.* **74,** 1–9.

Albrecher, H., Ladoucette, S. A. and Teugels, J. L. (2010). Asymptotics of the sample coefficient of variation and the sample dispersion. *J. Statist. Planning Inf.* **140,** 358–368.

Arnold, S. F. (1990). *Mathematical Statistics*. Prentice-Hall, Englewood Cliffs, NJ.

Barndorff-Nielsen, O. (2014). *Information and Exponential Families in Statistical Theory*. John Wiley, Chichester.

Billingsley, P. (1986). *Probability and Measure*, 2nd edn. John Wiley, New York.

Bingham, N. H., Goldie, C. M. and Teugels, J. L. (1987). *Regular Variation* (Encyclopedia Math. Appl. **27**). Cambridge University Press.

Çinlar, E. (2011). *Probability and Stochastics*. Springer, New York.

Cohen, J. E. and Xu, M. (2015). Random sampling of skewed distributions implies Taylor's power law of fluctuation scaling. *Proc. Nat. Acad. Sci. USA* **112,** 7749–7754.

de la Peña, V. H. and Yang, M. (1999). Probabilistic criteria for the order of the moment of the first passage time. Preprint, Department of Statistics, Columbia University.

Eisler, Z., Bartos, I. and Kertész, J. (2008). Fluctuation scaling in complex systems: Taylor's law and beyond. *Adv. Physics* **57,** 89–142.

Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*, Vol. II, 2nd edn. John Wiley, New York.

Fuchs, A. and Joffe, A. (1997). Exact formula for the expectation of the ratio of the sum of squares by the square of the sum. *C. R. Acad. Sci. Paris Ser. I Math.* **325,** 907–909.

Giné, E., Götze, F. and Mason, D. M. (1997). When is the Student t-statistic asymptotically standard normal? *Ann. Prob.* **25,** 1514–1531.

Jørgensen, B. (1987). Exponential dispersion models. *J. R. Statist. Soc. B* **49,** 127–162.

Jørgensen, B. (1997). *The Theory of Dispersion Models*. Chapman & Hall, London.

Jørgensen, B., Martínez, J. R. and Vinogradov, V. (2009). Domains of attraction to Tweedie distributions. *Lithuanian Math. J.* **49,** 399–425.

Karamata, J. (1931). Neuer Beweis und Verallgemeinerung der Tauberschen Sätze, welche die Laplacesche und Stieltjessche Transformation betreffen. *J. Reine Angew. Math.* **164,** 27–39.

Kendal, W. S. (2004). Taylor's ecological power law as a consequence of scale invariant exponential dispersion models. *Ecological Complexity* **1,** 193–209.

Kendal, W. S. (2013). Fluctuation scaling and $1/f$ noise: shared origins from the Tweedie family of statistical distributions. *J. Basic Appl. Physics* **2,** 40–49.

Kendal, W. S. and Jørgensen, B. (2011a). Taylor's power law and fluctuation scaling explained by a central-limit-like convergence. *Phys. Rev. E* **83,** 066115.

Kendal, W. S. and Jørgensen, B. (2011b). Tweedie convergence: a mathematical basis for Taylor's power law, $1/f$ noise, and multifractality. *Phys. Rev. E* **84**, 066120.

Neter, J., Wasserman, W. and Kutner, M. H. (1990). *Applied Linear Statistical Models: Regression, Analysis of Variance, and Experimental Designs*, 3rd edn. Irwin, Homewood, IL.

Ross, S. M. (2002). *Probability Models for Computer Science*. Harcourt Academic Press, San Diego, CA.

Taylor, L. R. (1961). Aggregation, variance and the mean. *Nature* **189,** 732–735.

Tweedie, M. C. K. (1946). The regression of the sample variance on the sample mean. *J. London Math. Soc.* **21,** 22–28.

Tweedie, M. C. K. (1947). Functions of a statistical variate with given means, with special reference to Laplacian distributions. *Proc. Camb. Phil. Soc.* **43,** 41–49.