Probability in the Engineering and Informational Sciences, page 1 of 12, 2017.
doi:10.1017/S0269964817000304

SQUARED COEFFICIENT OF VARIATION OF TAYLOR'S LAW FOR RANDOM ABSOLUTE DIFFERENCES

MARK BROWN
Department of Statistics, Columbia University,
New York, NY 10027 USA
E-mail: mb2484@columbia.edu

JOEL E. COHEN
Laboratory of Populations, Rockefeller University,
New York, NY 10065 USA;
Earth Institute and Department of Statistics, Columbia University,
New York, NY 10027 USA;
Department of Statistics, University of Chicago,
Chicago, IL 60637 USA
E-mail: cohen@rockefeller.edu

In a family, parameterized by θ, of non-negative random variables with finite, positive second moment, Taylor's law (TL) asserts that the population variance is proportional to a power of the population mean as θ varies: $\sigma^2(\theta) = a[\mu(\theta)]^b, a > 0$. TL, sometimes called fluctuation scaling, holds widely in science, probability theory, and stochastic processes. Here we report diverse examples of TL with $b = 2$ (equivalent to a constant coefficient of variation) arising from a difference of random variables in normed vector spaces of dimension 1 and larger. In these examples, we compute a exactly using, in some cases, a simple, new technique. These examples may prove useful in future models that involve differences of random variables, including models of the spatial distribution and migration of human populations.

Keywords: coefficient of variation, geometric probability, migration, signal-to-noise ratio, Taylor's law

1. INTRODUCTION

Ecological studies of the variation of species' population density in space and time led several ecologists to recognize, and Taylor [10] to bring to widespread attention, an empirical pattern now often called Taylor's law (TL) or fluctuation scaling (in the physical sciences). More than a thousand papers were published on the empirical support and theoretical foundations for TL according to a review by Eisler, Bartos, and Kertész [6], and many papers on TL have been published since then.

TL may be described mathematically as a property of a family of non-negative random variables $X(\theta) \geq 0$ indexed by a parameter θ. If, for all θ, each $X(\theta)$ has finite second moment, population mean $\mu(\theta)$ and population variance $\sigma^2(\theta) > 0$, and if, for real a and b independent of θ,

$$\sigma^2(\theta) = a[\mu(\theta)]^b, a > 0, \quad \text{or} \quad \frac{\sigma^2(\theta)}{[\mu(\theta)]^b} = a > 0,$$

then, by definition, TL holds. (We do not deal in this paper with the *sample* version of TL, in which the population moments are replaced by their corresponding sample moments.)

Cohen and Courgeau [2] observed that TL with $b = 2$ exactly describes the distances between pairs of randomly chosen points in various geometric shapes parameterized by their dimension, and under more general conditions. For example, they noted that if $X(\mu, c^2)$ is defined as the difference between two independent random variables, each of which is identically normally distributed with mean μ and variance c^2, then $X(\mu, c^2)$ satisfies TL with $a = \pi/2 - 1, b = 2$ for all μ and for all c. As empirical motivation, they showed that in Réunion Island and France, at some spatial scales, the empirical frequency distributions of inter-individual distances are predicted accurately by the theoretical frequency distributions of inter-point distances in models of geometric probability under a uniform distribution of points. This empirical analysis follows earlier analysis by Courgeau [3,4] and Courgeau and Baccaïni [5] of the spatial distribution of a human population within a region by means of the frequency distribution of the distance between a pair of individuals chosen at random. The larger goal of this work is to model a human population's spatial distribution and migration between points within a region.

Cases of TL with $b = 2$ have particular theoretical and practical interest because $b = 2$ if and only if the coefficient of variation (standard deviation divided by mean) and the signal-to-noise ratio (mean divided by standard deviation) are constant (and equal to $a^{\pm 1/2}$, respectively), regardless of the parameter θ. The family of exponential random variables is a familiar example of TL with $b = 2$ and $a = 1$.

Here we report diverse examples of TL with $b = 2$ arising from a difference of random variables. These examples are distinguished by our ability to compute a exactly. These examples may prove useful for future models that involve differences of random variables.

In Section 2, we develop methods for computing the ratio of the variance to the square of the mean of the difference between univariate random variables. Because $b = 2$ in all these examples, this ratio is the squared coefficient of variation. We give examples. In Section 3, we give examples of exact computations for vector-valued and Banach-space-valued random variables.

2. DISTANCES BETWEEN INDEPENDENT AND IDENTICALLY DISTRIBUTED (IID) RANDOM VARIABLES

PROPOSITION 2.1: *Let V_1, V_2 be iid $\sim V$ with $0 < Var(V) < \infty$, $R = |V_1 - V_2|$, and $X_i = cV_i + d, i = 1, 2, c \neq 0$. Then*

$$\frac{Var(|X_1 - X_2|)}{(E|X_1 - X_2|)^2} = \frac{Var(|V_1 - V_2|)}{(E|V_1 - V_2|)^2} = \frac{Var R}{(ER)^2}.$$

In this section, we compute $Var R/(ER)^2$ explicitly for several examples.

Example 2.2: Gamma Distribution
Let $V \sim \Gamma(k, \lambda)$, a gamma distribution with shape parameter k (a positive integer) and scale parameter λ (a positive real number). As λ is a scale parameter, without loss of generality

we set $\lambda = 1$. By definition, the probability density function (pdf) of $V \sim \Gamma(k, 1)$ is

$$f_V(v) = \frac{v^{k-1} e^{-v}}{(k-1)!}. \quad (1)$$

In this case, $EV = VarV = k$.

PROPOSITION 2.3: *Define*

$$p(k) = P\left(Bin\left(2k, \frac{1}{2}\right) = k\right) = \binom{2k}{k}\left(\frac{1}{2}\right)^{2k}, \quad k = 1, 2, \ldots.$$

Let $V_i \sim V, i = 1, 2$. *Then, continuing the notation of Proposition 2.1,*

$$\frac{VarR}{(ER)^2} = \frac{1}{2kp^2(k)} - 1. \quad (2)$$

As $k \to \infty, (VarR)/(ER)^2 \to (\pi/2) - 1$, *which is the value of* $(VarR)/(ER)^2$ *when* V *is normally distributed.*

PROOF: Since V is distributed as the time of the kth event in a Poisson process of rate 1, V exceeds v if and only if $N(v)$, the number of events in $[0, v]$, is at most $k - 1$. Hence, the survival function (complement of the cumulative distribution function) of V is

$$\overline{F}_V(v) = \sum_{r=0}^{k-1} \frac{v^r e^{-r}}{r!}, v \geq 0. \quad (3)$$

$N(v)$ is Poisson with parameter v. Define $m = \min(V_1, V_2)$. From (1) and (3), the pdf of m by time t is

$$f_m(t) = 2f_V(t)\overline{F}_V(t) = \frac{2}{(k-1)!}\sum_{r=0}^{k-1} \frac{t^{r+k-1} e^{-2t}}{r!}. \quad (4)$$

Thus, using the gamma identity $\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx$ and a change of variables, we have

$$\begin{aligned} Em &= \int_{t=0}^{\infty} t f_m(t) dt = \frac{2}{(k-1)!}\sum_{r=0}^{k-1}\frac{1}{r!}\int_{t=0}^{\infty} t^{r+k} e^{-2t} dt \\ &= \frac{2}{(k-1)!}\sum_{r=0}^{k-1}\frac{1}{r!}\int_{t=0}^{\infty} t^{(r+k+1)-1} e^{-2t} dt = \frac{2}{(k-1)!}\sum_{r=0}^{k-1}\frac{1}{r!}\frac{\Gamma(r+k+1)}{2^{r+k+1}} \quad (5) \\ &= 2k\sum_{r=0}^{k-1}\binom{r+k}{k}\left(\frac{1}{2}\right)^{r+k+1} = 2kP(Y \leq k-1), \end{aligned}$$

where Y is distributed as the number of failures prior to the $(k + 1)$st success in Bernoulli trials with success probability $1/2$, so that $P(Y = r) = \binom{r+k}{r}(1/2)^{r+k+1}$.

Since

$$P(Y \leq k) = P(k+1 \text{ successes before } k+1 \text{ failures}) = 1/2 \tag{6}$$

and

$$P(Y = k) = \binom{2k}{k}\left(\frac{1}{2}\right)^{2k+1} = \frac{1}{2}p(k), \tag{7}$$

where (as above) $p(k) = P(Bin(2k,(1/2)) = k)$, it follows from (4)–(7) that

$$Em = 2k[P(Y \leq k) - P(Y \leq k-1)] = 2k\left[\frac{1}{2} - \frac{1}{2}p(k)\right] = k(1 - p(k)). \tag{8}$$

Define $M = \max(V_1, V_2)$. Then

$$ER = E|V_1 - V_2| = EM - Em \tag{9}$$

and

$$2EV = E[V_1 + V_2] = EM + Em. \tag{10}$$

From (9) and (10),

$$ER = 2(EV - Em). \tag{11}$$

We will use (11) in some of our other examples. A consequence of (11) is that in examples where Em is computable, then ER is also computable.

In this example, from (8) and (11),

$$ER = 2[k - k(1 - p(k))] = 2kp(k). \tag{12}$$

Next,

$$ER^2 = E\{[(V_1 - EV) - (V_2 - EV)]^2\} = 2VarV = 2k, \tag{13}$$

so from (12) and (13),

$$\frac{VarR}{(ER)^2} = \frac{ER^2}{(ER)^2} - 1 = \frac{1}{2kp^2(k)} - 1,$$

which is (2). From Stirling's approximation Feller [7, p. 47], $p(k)/(1/\sqrt{\pi k}) \to 1$ as $k \to \infty$. Thus from (2), $(VarR)/(ER)^2 \to (\pi/2) - 1 \approx 0.5708$ as $k \to \infty$. ■

For $k = 5$, $(VarR)/(ER)^2 \approx 0.6512$; for $k = 50$, $(VarR)/(ER)^2 \approx 0.5787$.

Example 2.4: Symmetrized Distributions

If we can compute the coefficient of variation for $R = |V_1 - V_2|$ corresponding to $V \geq 0$ with pdf f_V, then we can also do so for R^* corresponding to V^*, the symmetrized version of V, which has (by definition) pdf

$$f_{V^*}(v) = \frac{1}{2}f_V(|v|), \quad -\infty < v < \infty.$$

PROPOSITION 2.5: *Define* $R^* = |V_1^* - V_2^*|$, *where* V_1^*, V_2^* *are iid* $\sim V^*$. *Then*

$$\frac{VarR^*}{(ER^*)^2} = \frac{2EV^2}{(EV + (ER/2))^2} - 1. \tag{14}$$

If $V \sim \Gamma(k,1)$, *then* (14) *gives*

$$\frac{VarR^*}{(ER^*)^2} = \frac{2(k+1)}{k(1+p(k))^2} - 1. \tag{15}$$

As $k \to \infty$, (15) converges slowly to 1. For $k = 5, 50, 10^2, 10^3, 10^4$, respectively, $(VarR^*)/(ER^*)^2 \approx 0.5465, 0.7503, 0.8102, 0.9324, 0.9778$.

If F is the standard normal cdf and $f = F'$ is the standard normal pdf, so that $f' = -xf$, then $f(0) = (2\pi)^{-1/2}$ and $f'(0) = 0$. Asymptotically as $k \to \infty$,

$$p(k)/[2F((2k)^{-1/2}) - 1] \to 1.$$

Thus,

$$p(k) = 2(1/2 + 1/(2(k\pi)^{1/2}) + o(1/k)) - 1 = (k\pi)^{-1/2} + o(1/k).$$

Hence,

$$(p(k)+1)^{-2} = 1 - 2(k\pi)^{-1/2} + 3/(k\pi) + o(1/k)$$

and

$$\frac{2(k+1)}{k(1+p^2(k))} - 1 = \frac{VarR^*}{(ER^*)^2} = 1 - \frac{4}{\sqrt{k\pi}} + \frac{2}{k}\left[1 + \frac{3}{\pi}\right] + o(k^{-1}).$$

For $k = 10^2, 10^3, 10^4$, the first three terms on the right give, respectively, 0.8134, 0.9325, 0.9778.

PROOF: To prove (14), we construct V_1^* by letting $V_1^* = V_1$ with probability 1/2 and $V_1^* = -V_1$ with probability 1/2. We construct V_2^* similarly by letting $V_2^* = V_2$ with probability 1/2 and $V_2^* = -V_2$ with probability 1/2. To have V_1^*, V_2^* independent, we independently choose their signs, for example with independent flips of a fair coin. If V_1^*, V_2^* have the same sign, which happens with probability 1/2, then $R^* \sim R$. If V_1^*, V_2^* have opposite signs, then $R^* \sim V_1 + V_2$. Thus,

$$ER^* = \frac{1}{2}[ER + 2EV] = EV + \frac{ER}{2} = 2EV - Em, \tag{16}$$

where the last equality follows from (11). Arguing similarly to (16),

$$ER^{*2} = \frac{1}{2}[ER^2 + E(V_1+V_2)^2] = \frac{1}{2}[2VarV + 2EV^2 + 2(EV)^2] = 2EV^2. \tag{17}$$

Result (14) then follows from (16) and (17).

When $V \sim \Gamma(k,1)$, it follows from (12) and (16) that

$$ER^* = k(1+p(k)), \tag{18}$$

and, from (17), that

$$ER^{*2} = 2EV^2 = 2(VarV + (EV)^2) = 2k(k+1). \tag{19}$$

Result (15) then follows from (18) and (19). ∎

Example 2.6: Half-normal Distributions

Let $Z, Z_i \sim N(0,1), V_i = |Z_i|, i = 1,2$. Then $ER = E||Z_1| - |Z_2||$, which appears difficult to compute directly. Instead, we compute the simpler $ER^* = E|Z_1 - Z_2|$ and then invert (16) to get ER.

Since $R^* = |Z_1 - Z_2| \sim \sqrt{2}|Z|$,

$$ER^* = \sqrt{2}E|Z| = \sqrt{2}\sqrt{\frac{2}{\pi}}. \tag{20}$$

Then from (16) and (20),

$$ER = E\big||Z_1| - |Z_2|\big| = 2[ER^* - EV] = 2[ER^* - E|Z|] = 2\sqrt{\frac{2}{\pi}}[\sqrt{2} - 1]. \tag{21}$$

From (12),

$$ER^2 = 2VarV = 2Var(|Z|) = 2\left[1 - \frac{2}{\pi}\right]. \tag{22}$$

From (21) and (22), we obtain

$$\frac{VarR}{(ER)^2} = \frac{Var[\big||Z_1| - |Z_2|\big|]}{(E\big||Z_1| - |Z_2|\big|)^2} = \frac{\pi - 2}{2(2 - \sqrt{2})^2} - 1 \approx 0.6634. \tag{23}$$

Example 2.7: Beta Distributions

Suppose $V \sim$ Beta(a,1) with pdf $f_V(t) = at^{a-1}, 0 < t < 1, a > 0$. Then $\overline{F}_V(t) = 1 - t^a, 0 \leq t \leq 1$. We calculate

$$Em = \int_0^1 (1 - 2t^a + t^{2a})dt = \frac{2a^2}{(a+1)(2a+1)}, \tag{24}$$

$$ER = 2(EV - Em) = \frac{2a}{(a+1)(2a+1)}, \tag{25}$$

$$ER^2 = \frac{2a}{(a+1)^2(a+2)}. \tag{26}$$

It follows that

$$\frac{VarR}{(ER)^2} = \frac{2a^2+1}{2a(a+2)} \stackrel{def}{=} g(a). \tag{27}$$

The global minimum of $g(a)$ is $1/2$ at $a = 1$, when Beta(1,1) = Uniform(0,1) distribution. $g(a)$ decreases from ∞ to $1/2$ as a goes from 0 to 1, and increases from $1/2$ to 1 as a goes from 1 to ∞; $g(a) = 1$ at the single point $a = 1/4$.

The method developed in Example 2.4 yields in this Beta(a,1) case that

$$\frac{VarR^*}{(ER^*)^2} = \frac{VarR}{(ER)^2} = g(a).$$

This equality between the squared coefficient of variation of R and R^* does not hold in any of our other examples. It is a curiosity.

Example 2.8: Completely Monotone Distributions

Completely (or totally) monotone distributions are mixtures of exponential distributions (Bernstein's theorem). We can represent a completely monotone distribution as

$$V = U\epsilon$$

where U and ϵ are independent, ϵ is exponentially distributed with mean 1, and U is a non-negative random variable (not necessarily Uniform(0,1) here). If $EU < \infty$, then $EV = EU$, and if $EU^2 < \infty$, then

$$EV^2 = 2(EU^2), \quad VarV = 2VarU + (EU)^2. \tag{28}$$

Completely monotone distributions form a subclass of distributions with decreasing failure rate (DFR) (e.g., Barlow and Proschan [1], Marshall and Olkin [9]). Completely monotone distributions are studied in Feller [7], and applied to first-passage times for time-reversible Markov chains in Keilson [8].

For the next results, we recall that if X and Y are independent exponentials with means c_1 and c_2, respectively, then $E\min(X,Y)$ is exponential with mean $1/(1/c_1 + 1/c_2) = c_1c_2/(c_1+c_2)$.

Let $V_i = U_i\epsilon_i, i = 1,2$, be iid. Then $U_1\epsilon_1, U_2\epsilon_2$, given $U_1 = u_1, U_2 = u_2$, are independent exponentials with parameters u_1, u_2. Then, by the preceding facts,

$$E(m|U_1,U_2) = E(\min(U_1\epsilon_1, U_2\epsilon_2)|U_1,U_2) = E\left[\frac{U_1U_2}{U_1+U_2}\right], \tag{29}$$

$$ER = 2[EV - Em] = 2E\left[U_1 - \frac{U_1U_2}{U_1+U_2}\right] = 2E\left[\frac{U_1^2}{U_1+U_2}\right].$$

From (28) and (29),

$$\frac{VarR}{(ER)^2} = \frac{2\sigma_U^2 + (EU)^2}{2\left(E\left[\frac{U_1^2}{U_1+U_2}\right]\right)^2} - 1. \tag{30}$$

Expression (30) is difficult to compute in general. We compute it in a few special cases and then derive two-sided bounds in (37) by exploiting the DFR property.

Case 2.9: $P(U=1) = p = 1 - P(U=2)$, where $P(U=2) = q$ and $0 < p < 1$, a mixture of two exponential distributions with means 1 and 1/2, respectively. In this case,

$$E\left(\frac{U_1^2}{U_1+U_2}\right) = \frac{1}{6}(3p^2 + 10pq + 6q^2), \tag{31}$$

$$ER = \frac{1}{3}(6 - 2p - p^2). \tag{32}$$

From (28), (30), and (32),

$$\frac{VarR}{(ER)^2} = \frac{36 - 12p - 10p^2 - 4p^3 - p^4}{(6-2p-p^2)^2}, \tag{33}$$

which is maximized at $p = \sqrt{3} - 1$, where the maximum equals 5/4; (33) exceeds 1 for all p, converging to 1 as p goes to 0 or 1.

Case 2.10: $U \sim$ Uniform(0,1), $V = U\epsilon$. In this case,

$$ER = 2\left[\frac{2}{3}\log 2 - \frac{1}{6}\right],$$

$$\frac{VarR}{(ER)^2} = \frac{5/6}{4\left[(2/3)\log 2 - (1/6)\right]^2} - 1 \approx 1.3870.$$

Case 2.11: $U \sim (1/\Gamma(\alpha, \lambda)), \alpha > 2, \lambda > 0, V = U\epsilon$. In this case,

$$ER = \frac{2\alpha\lambda}{(\alpha-1)(2\alpha+1)},$$

$$\frac{VarR}{(ER)^2} = \frac{2\alpha^2+1}{2\alpha(\alpha-2)}.$$

The above distribution has survival function

$$\overline{F}(t) = \left(\frac{\lambda}{\lambda+t}\right)^{\alpha},$$

the survival function of $U\epsilon$ coinciding with the Laplace transform of U^{-1}.

Bound: Define $K(t) = E[V - t|V > t]$, the mean residual life or remaining life expectancy at t. Then

$$ER = E[M - m] = EK(m),$$

where (as above) $m = \min(V_1, V_2)$. This is true since given $m = y, M - m \sim (V - y|V > y)$, and $E[M - m|m = y] = K(y)$. Thus, $ER = E[E(R|m)] = E[K(m)]$.

If V has DFR with hazard rate $h(s) = f(s)/\overline{F}(s)$, the hazard rate of $(V - t|V > t)$ at s equals $h(t + s)$; thus $(V - t|V > t)$ has DFR, and R is thus a mixture of random variables with DFR, which is DFR. As R has DFR,

$$VarR \geq (ER)^2. \tag{34}$$

Further, R is stochastically larger than V. Thus,

$$ER \geq EV = EU(\text{if } V = U\epsilon). \tag{35}$$

Thus

$$\frac{VarR}{(ER)^2} = \frac{2VarV}{(ER)^2} - 1 \leq \frac{2VarV}{(EV)^2} - 1 = \frac{2[2VarU + (EU)^2]}{(EU)^2} - 1 = 1 + \frac{4VarU}{(EU)^2}. \tag{36}$$

From (35) to (36) for V DFR,

$$1 \leq \frac{VarR}{(ER)^2} \leq \frac{2VarV}{(EV)^2} - 1. \tag{37}$$

If $V = U\epsilon$ (which implies that V completely monotone, a subclass of DFR),

$$\frac{2VarV}{(EV)^2} - 1 = 1 + \frac{4VarU}{(EU)^2}. \tag{38}$$

Thus (37) gives a two-sided bound for the squared coefficient of variation of a DFR distribution, and (38) an equivalent bound for completely monotone distributions.

When U is a constant, or equivalently V is exponentially distributed, the lower and upper bounds coincide, giving the value 1, from Example 2.2 with $k = 1$.

Example 2.12: Truncated Exponential Distributions

Let $T > 0$ be a constant and let X be exponential with parameter λ. Then λX is exponential with parameter 1. Define $V = \min(X, T)$. Then

$$\lambda V = \min(\lambda X, \lambda T) = \min(\text{Exponential}(1), \lambda T).$$

Since $VarR/(ER)^2$ is the same for λV as for V, we can without loss of generality work with

$$V = \min(\epsilon, w), \epsilon \sim \text{Exponential}(1), \text{constant } w > 0.$$

In any given example (λ, T), replace w by λT to get the value of $(VarR)/(ER)^2$. Now

$$EV = \int_0^w e^{-x} dx = 1 - e^{-w},$$
$$Em = E\min(V_1, V_2) = \int_0^w e^{-2x} dx = \frac{1 - e^{-2w}}{2}.$$

Thus by (11),

$$ER = 2(EV - Em) = (1 - e^{-w})^2.$$

Next, $EV^2 = \int_0^w 2xe^{-x}dx = 2(1 - (w+1)e^{-w})$, so $VarV = 1 - 2we^{-w} - e^{-2w}$. From (10) and (30),

$$\frac{VarR}{(ER)^2} = \frac{2(1 - 2we^{-w} - e^{-2w})}{(1 - e^{-w})^4} - 1 = \frac{1 - 4(w-1)e^{-w} - 8e^{-2w} + 4e^{-3w} - e^{-4w}}{(1 - e^{-w})^4}. \quad \textbf{(39)}$$

This ratio converges to 1 (the value for an untruncated exponential distribution) as $w \to \infty$, and converges to ∞ as $w \downarrow 0$.

Example 2.13: Weibull Distributions

The Weibull distribution is used in survival analysis, extreme value theory, industrial and reliability engineering. One of its two parameters is a scale parameter, which we ignore because it has no effect on $VarR/(ER)^2$.

Let ϵ be exponential with mean 1. Define

$$Y_\alpha = \epsilon^{1/\alpha}, \quad \alpha > 0.$$

The survival function of Y_α is

$$\overline{F}_\alpha(t) = P(Y_\alpha > t) = P(\epsilon > t^\alpha) = e^{-t^\alpha}, \quad t \geq 0.$$

This distribution is Weibull with shape parameter α and scale parameter 1.

The hazard function of the Weibull distribution is $h_\alpha(t) \stackrel{def}{=} f_\alpha(t)/\overline{F}_\alpha(t) = \alpha t^{\alpha-1}$. For $\alpha < 1, Y_\alpha$ has DFR, for $\alpha > 1$ has increasing failure rate (IFR), and for $\alpha = 1$ is exponential. Some well-known Weibull properties are (40)–(42) below:

$$EY_\alpha = \Gamma\left(\frac{1}{\alpha}+1\right) \quad \text{(by definition of the gamma function).} \tag{40}$$

$$EY_\alpha^2 = \Gamma\left(\frac{2}{\alpha}+1\right). \tag{41}$$

$$VarY_\alpha = \Gamma\left(\frac{2}{\alpha}+1\right) - \Gamma^2\left(\frac{1}{\alpha}+1\right) \quad \text{(by (40) and (41)).} \tag{42}$$

$$\begin{aligned} Em_\alpha &= \int_0^\infty e^{-2t\alpha} dt = \frac{1}{\alpha}\int_0^\infty z^{(1/\alpha)-1} e^{-2z} dz \quad (\text{where } z = t^\alpha) \\ &= \frac{(1/\alpha)\Gamma(1/\alpha)}{2^{1/\alpha}} = \frac{\Gamma((1/\alpha)+1)}{2^{1/\alpha}}. \end{aligned} \tag{43}$$

Define $R_\alpha = |Y_{\alpha,1} - Y_{\alpha,2}|$. Then by (11),

$$ER_\alpha = 2(EY_\alpha - Em_\alpha) = 2\Gamma\left(\frac{1}{\alpha}+1\right)(1-2^{-1/\alpha}). \tag{44}$$

Thus,

$$\frac{VarR_\alpha}{(ER_\alpha)^2} = \frac{2VarY_\alpha}{(ER_\alpha)^2} - 1 = \frac{\Gamma((2/\alpha)+1) - \Gamma^2((1/\alpha)+1)}{2\Gamma^2(\frac{1}{\alpha}+1)(1-2^{-1/\alpha})^2} - 1.$$

For example, if $\alpha = 2$,

$$\begin{aligned} \frac{VarR_2}{(ER_2)^2} &= \frac{\Gamma(2) - \Gamma^2(3/2)}{2\Gamma^2(3/2)(1-2^{-1/2})^2} - 1 = \frac{1-(\pi/4)}{(\pi/2)(1-2^{-1/2})^2} - 1 \\ &= \frac{4-\pi}{\pi(\sqrt{2}-1)^2} - 1 = \frac{4-\pi}{\pi(3-\sqrt{8})} - 1 = \frac{4-\pi(4-\sqrt{8})}{\pi(3-\sqrt{8})} \approx 0.5926. \end{aligned}$$

Example 2.14: Correlated Normal Distributions

The iid univariate normal case was considered in Example 2.4. Here we consider two correlated normals. Assume

$$\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right), \quad \rho \neq 1. \tag{45}$$

Then $V_1 - V_2 \sim N(0, 2(1-\rho)) \sim \sqrt{2(1-\rho)}Z$, with $Z \sim N(0,1)$. Define $R = |V_1 - V_2|$. Then

$$ER = \sqrt{2(1-\rho)}\sqrt{\frac{2}{\pi}} = 2\sqrt{\frac{1-\rho}{\pi}},$$

$$ER^2 = 2(1-\rho).$$

Because $\rho \neq 1$ by assumption in (45), we have $ER > 0$, $ER^2 > 0$, and

$$\frac{VarR}{(ER)^2} = \frac{\pi}{2} - 1. \tag{46}$$

More generally, if either $\rho \neq 1$ or $\sigma_1 \neq \sigma_2$ and

$$\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu \\ \mu \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix} \right), \quad \sigma_1 > 0, \sigma_2 > 0$$

then (46) holds.

3. VECTOR AND BANACH-VALUED RANDOM VARIABLES

Let $V_1^{m\times 1}$ and $V_2^{m\times 1}$ be iid random vectors of m components, $c \neq 0$ be a scalar, $\Gamma^{m\times m}$ be orthogonal (i.e., $\Gamma\Gamma' = \Gamma'\Gamma = I$), and

$$X_i = \mu + c\Gamma V_i, \quad i = 1, 2.$$

Define $R = V_1 - V_2$ and $\|x^{m\times 1}\| = \sqrt{x'x} = \sum x_i^2$. Then

$$D \overset{def}{=} \|X_1 - X_2\| = \|c\Gamma(V_1 - V_2)\| = |c|\|\Gamma R\| = |c|\|R\|.$$

Thus, $ED = |c|E(\|R\|)$, and $ED^2 = c^2 E(\|R\|^2)$, so

$$\frac{VarD}{(ED)^2} = \frac{E(\|R\|^2)}{(E(\|R\|))^2} - 1 = \frac{Var(\|R\|)}{(E\|R\|)^2}, \tag{47}$$

independently of μ, c, and Γ.

Example 3.1: Uncorrelated Normal Vectors

Let $Z_i^{m\times 1} \sim N(0, I), i = 1, 2, Z_1, Z_2$ independent. Then $R = Z_1 - Z_2 \sim \sqrt{2}N(0, I) \sim \sqrt{2}Z$ with $Z \sim N(0, I)$. Thus,

$$\|R\| \sim \sqrt{2}\sqrt{\chi_m^2} \tag{48}$$

where χ_m^2 is the chi-squared distribution with m degrees of freedom. Since

$$E(\sqrt{\Gamma(\alpha, \lambda)}) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha + (1/2))}{\lambda^{\alpha+1/2}} = \frac{1}{\sqrt{\lambda}} \frac{\Gamma(\alpha + (1/2))}{\Gamma(\alpha)},$$

it follows that

$$E\sqrt{\chi_m^2} = E\sqrt{\Gamma\left(\frac{m}{2}, \frac{1}{2}\right)} = \sqrt{2}\frac{\Gamma((m+1)/2)}{\Gamma(m/2)}.$$

Thus,

$$E(\|R\|) = 2\frac{\Gamma((m+1)/2)}{\Gamma(m/2)}, \tag{49}$$

$$E(\|R\|^2) = 2E\chi_m^2 = 2m. \tag{50}$$

Consequently, from (42) and (43),

$$\frac{Var(\|R\|)}{(E\|R\|)^2} = \frac{m\Gamma^2(m/2)}{2\Gamma^2((m+1)/2)} - 1. \tag{51}$$

For $m = 1$, the squared coefficient of variation of $\|R\|$ equals $(\pi/2) - 1$, agreeing with (46). For $m = 5$, $Var(\|R\|)/(E(\|R\|))^2 = (45\pi/128) - 1 \approx 0.1045$.

Example 3.2: Banach-valued Random Variables

To find a computable case, we deal with a squared norm (rather than a norm, as in previous examples). Suppose that V_1 and V_2 are iid standard Brownian motion processes on $[0, T]$, $T > 0$. V_1 and V_2 take values in the Banach space $L^2([0, T], \beta, l)$, with β the Borel sets of $[0, T]$ and l the Lebesgue measure, which contains real-valued square-integrable functions x on $[0, T]$ with $\|x\|^2 = \int_0^T x^2(s)ds$. Define $R^2 = \|V_1 - V_2\|^2 \sim 2\|V\|^2$, where V is the standard Brownian motion on $[0, T]$. Then

$$ER^2 = 2E\left[\int_0^T V^2(s)ds\right] = \int_0^T 2sds = T^2.$$

Let $\tilde{V}$ be the standard Brownian motion on $[0, T]$, independent of V. Using stationary and independent increments,

$$\begin{aligned} ER^4 &= 8\int_{u=0}^{T}\int_{v=0}^{u} E\left[V^2(v)[V^2(v) + 2V(v)\tilde{V}(u-v) + \tilde{V}^2(u-v)]\right] dvdu \\ &= 8\int_{u=0}^{T}\int_{v=0}^{u}(2v^2 + vu)dvdu = 8\int_0^T \frac{7}{6}u^3 du = \frac{7}{3}T^4. \end{aligned}$$

Thus,

$$\frac{VarR^2}{(ER^2)^2} = \frac{7}{3} - 1 = \frac{4}{3}.$$

If

$$X_i(s) = m(s) + cV_i(s), \quad i = 1, 2, \quad 0 \leq s \leq T,$$

then $X_1(s) - X_2(s) = c(V_1(s) - V_2(s))$, and

$$\frac{Var(\|X_1 - X_2\|^2)}{(E\|X_1 - X_2\|^2)^2} = \frac{4}{3},$$

independently of m, c, and T. Thus, TL holds for $\|X_1 - X_2\|^2$ with exact coefficient $a = 4/3$ and exact exponent $b = 2$.

References

1. Barlow, R.E. & Proschan, F. (1975). *Statistical theory of reliability and life testing*. New York: Holt.
2. Cohen, J.E. & Courgeau, D. (In press). Modeling distances between humans using Taylor's law and geometric probability. *Mathematical Population Studies*.
3. Courgeau, D. (1970). *Les champs migratoires en France*. Paris: Presses Universitaires de France.
4. Courgeau, D. (1973). Migrations et découpages du territoire. *Population* 28(3): 511–537.
5. Courgeau, D. & Baccaïni, B. (1989). Migrations et distances. *Population* 42(1): 57-82.
6. Eisler, Z., Bartos, I., & Kertész, J. (2008). Fluctuation scaling in complex systems: Taylors law and beyond. *Advances in Physics* 57(1): 89–142. doi: 10.1080/00018730801893043.
7. Feller, W. (1971). *An introduction to probability theory and its applications*, vol. II, 2nd ed. New York: John Wiley and Sons.
8. Keilson, J. (1979). *Markov chain models, rarity and exponentiality*. New York: Springer-Verlag.
9. Marshall, A.W. & Olkin, I. (2007). *Life distributions: structure of nonparametric, semiparametric and parametric families. Springer series in statistics*. New York: Springer-Verlag.
10. Taylor, L.R. (1961). Aggregation, variance and the mean. *Nature* 189: 732–735.