Research Article

Open Access CrossMark

Linear inference under alpha–stable errors

Abstract

Linear inference remains pivotal in statistical practice, despite errors often having excessive tails and thus deficient of moments required in conventional usage. Such errors are modeled here via spherical α –stable measures on $\mathbb{R}^n$ with stability index $\alpha\in(0,2]$, arising in turn through multivariate central limit theory devoid of the second moments required for Gaussian limits. This study revisits linear inference under α –stable errors, focusing on aspects to be salvaged from the classical theory even without moments. Critical entities include Ordinary Least Squares (OLS) solutions, residuals, and conventional F ratios in inference. Closure properties are seen in that OLS solutions and residual vectors under α –stable errors also have α –stable distributions, whereas F ratios remain exact in level and power as for Gaussian errors. Although correlations are undefined for want of second moments, corresponding scale parameters are seen to gauge degrees of association under α –stable symmetry.

AMS subject classification: 62E15, 62H15, 62J20

Keywords: excessive errors, central limit theory, stable laws, linear inference

Volume 7 Issue 3 - 2018

Donald R Jensen

Virginia Polytechnic Institute and State University, USA

Correspondence: Donald R Jensen, Professor Emeritus, Virginia Polytechnic Institute and State University, Blacksburg VA, 24061, USA, Tel 540 639 0865, Email djensen@vt.edu

Received: May 02, 2018 | **Published:** May 23 2018

Introduction

Models here are $\{\mathbf{Y}=\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}\}$ with error vector $\boldsymbol{\varepsilon}\in\mathbb{R}^n$. Classical linear inference rests heavily on means, variances, correlations, skewness and kurtosis parameters, these requiring moments to fourth order. To the contrary, distributions having excessive tails, and devoid of moments even of first or second order, arise in a variety of circumstances. These encompass acoustics, image processing, radar tracking, biometrics, portfolio analysis and risk management in finance, and other venues in contemporary practice. Supporting references include[1–6] monographs of note are,[7–9] together with the recent work of Nolan.[10] In these settings the classical foundations necessarily must be reworked.

To place this study in perspective, alternatives to Gaussian laws long have been sought in theory and practice, culminating in the class $\{S_n(\mathbf{0},\Sigma)\}$ consisting of elliptically contoured distributions in $\mathbb{R}^n$ centered at $\mathbf{0}$ with scale parameters Σ. These typically are taken to be rich in moments, and to provide alternatives to the use of large–sample approximate Gaussian distributions under conditions for central limit theory. Comprehensive treatises on the theory and applications of these models are.[11–13]

In contrast, errors having excessive tails are modeled on occasion via spherically symmetric α –stable $(S\alpha S)$ distributions in $\mathbb{R}^n$ with index $\alpha\in(0,2]$. These comprise the limit distributions of standardized vector sums, specifically, Gaussian limits at $\alpha=2$, Cauchy limits at $\alpha=1$, and corresponding stable limits otherwise. These distributions are contained in the class $\{S_n(\mathbf{0},\mathbf{I}_n)\}$. thus sharing its essential geometric features, but instead are deficient in moments usually ascribed to $\{S_n(\mathbf{0},\mathbf{I}_n)\}$. Despite the venues cited, α –stable errors have seen limited usage for want of closed expressions for stable density functions, known only in selected cases but topics of continuing research. Nonetheless, findings reported here rest on well defined *characteristic functions* (chfs), on critical representations for these, and on the inversion of the latter in order to represent the α –stable densities themselves. Even here a divide emerges between independent, identically distributed (iid) α –stable sequences, and dependent $S\alpha S$ variables, as reported in Jensen[14] and as summarized here for completeness in an Appendix. In addition, many findings of the present study are genuinely nonparametric, in applying for all or portions of distributions in the range $\alpha\in(0,2]$, and thus remaining distribution–free within that class. An outline follows.

Notation and technical foundations are provided in the next major section, Preliminaries, to include Notation and accounts of Special Distributions, Central Limit Theory and Essentials of $S\alpha S$ Distributions as subsections. The principal sections following these address Linear Models under $S\alpha S$ Errors, with a separate subsection on Models Having Cauchy Errors, and Conclusions. Collateral topics are contained for completeness in Appendix A.

Preliminaries

Notation

Spaces of note include $\mathbb{R}^n$ as Euclidean n –space, with $\mathbb{S}_n$ as the real symmetric $(n\times n)$ matrices and $\mathbb{S}_n^+$ as their positive definite varieties. Vectors and matrices are set in bold type; the transpose, inverse, trace, and determinant of $\mathbf{A}$ are $\mathbf{A}'$, $\mathbf{A}^{-1}$, $tr(\mathbf{A})$, and $|\mathbf{A}|$; the unit vector in $\mathbb{R}^n$ is $\mathbf{1}_n=[1,\ldots,1]'$; and $\mathbf{I}_n$ is the $(n\times n)$ identity.

Moreover, $Diag(\mathbf{A}_1,\ldots,\mathbf{A}_k)$ is a block–diagonal array, and $\Sigma^{\frac{1}{2}}$ is the spectral square root of $\Sigma\in\mathbb{S}_n^+$.

Special distributions

Given $\mathbf{Y}=[Y_1,\ldots,Y_n]'\in\mathbb{R}^n$, its distribution, expected value, and dispersion matrix are designated as $\mathcal{L}(\mathbf{Y})$, $E(\mathbf{Y})=\mu$, and $V(\mathbf{Y})=\Sigma$, with variance $Var(Y)=\sigma^2$ on $\mathbb{R}^1$. Specifically, $\mathcal{L}(\mathbf{Y})=N_n(\mu,\Sigma)$ is Gaussian on $\mathbb{R}^n$ with parameters (μ,Σ). Distributions on $\mathbb{R}^1$ of note include the $\chi^2(u;\nu,\lambda)$ and related $\chi(u;\nu,\lambda)$ distributions, together with the Snedecor–Fisher $F(u;\nu_1,\nu_2,\lambda)$, these having (ν,ν_1,ν_2) as degrees of freedom and λ a noncentrality parameter. The characteristic function (chf) for $\mathbf{Y}\in\mathbb{R}^n$ is the expectation $\phi_{\mathbf{Y}}(\mathbf{t})=E[e^{i\mathbf{t}'\mathbf{Y}}]$ with argument $\mathbf{t}'=[t_1,\ldots,t_n]$ and $\iota=\sqrt{-1}$; a standard source is Lukacs & Laha.[15] Attention is drawn subsequently to probability density (pdf) and cumulative distribution (cdf) functions. Moreover, the class $\{\mathcal{L}(\mathbf{Z})\in S_n(\mathbf{0},\Sigma)\}$ consists of elliptically contoured distributions in $\mathbb{R}^n$ centered at $\mathbf{0}$ and having chf's of type $\phi_{\mathbf{Z}}(\mathbf{t})=\psi(\mathbf{t}'\Sigma\mathbf{t})$. We adopt the following.

Definition 1 A distribution P on $\mathbb{R}^n$ is said to be monotone unimodal about $\mathbf{0}\in\mathbb{R}^n$ if for every $\mathbf{y}\in\mathbb{R}^n$ and every convex set C symmetric about $\mathbf{0}\in\mathbb{R}^n$, $P[C+k\mathbf{y}]$ is non increasing in $k\in[0,\infty)$. See reference.[16]

Central limit theory

For iid vectors $\{\mathbf{Z}_1,\mathbf{Z}_2,\mathbf{Z}_3,\ldots\}$ in $\mathbb{R}^n$, let $\overline{\mathbf{Z}}_N=N^{-1}[\mathbf{Z}_1+\ldots+\mathbf{Z}_N]$, and consider limit distributions of type $\{\mathcal{L}_\infty(c\overline{\mathbf{Z}}_N)=\liminf \mathcal{L}(c\overline{\mathbf{Z}}_N)\}$ for suitably chosen c. On specializing from the elliptical class $S_n(\mathrm{d},\Sigma)$ having location–scale parameters $(\mathbf{d},\Sigma)$, we consider α–stable limit distributions as follow on identifying $\mathcal{L}_\infty(c\overline{\mathbf{Z}}_N)$ with $\mathcal{L}(\mathbf{Z})$.

Definition 2 Let $\mathcal{L}(\mathbf{Z})\in S_n^\alpha(\mathrm{d},\Sigma)$ designate an elliptical α–Stable law on $\mathbb{R}^n$ centered at $\mathrm{d}\in\mathbb{R}^n$ with scale parameters Σ and stable index $\alpha\in(0,2]$, having the chf $\phi_{\mathbf{Z}}(\mathbf{t})=\exp\{\iota\mathbf{t}'\mathrm{d}-\frac{1}{2}(\mathbf{t}'\Sigma\mathbf{t})^{\frac{\alpha}{2}}\}$. Each marginal distribution of $S_n^\alpha(\delta\mathbf{1}_n,\mathbf{I}_n)$ on $\mathbb{R}^1$, namely $S_1^\alpha(\delta,1)$, has the chf $\phi_{Z_i}(t)=\exp\{\iota t\delta-\frac{1}{2}|t|^\alpha\}$. Let $S\alpha S=\{S_n^\alpha(\mathrm{d},\Sigma);(\mathrm{d},\Sigma)\in(\mathbb{R}^n\otimes S_n)\}$ designate the class of all such distributions.

Remark 1 $\mathcal{L}(\mathbf{Z})$ is of full rank and has a density in $\mathbb{R}^n$ if and only if Σ is of full rank in $\mathbb{S}_n^+$; otherwise $\mathcal{L}(\mathbf{Z})$ is concentrated in a subspace of $\mathbb{R}^n$ of dimension equal to the rank of Σ.

To continue, designate by $\mathfrak{D}_\alpha$ the domain of attraction of each element $\mathbf{Z}_i$ in $\{\mathbf{Z}_1,\mathbf{Z}_2,\mathbf{Z}_3,\ldots\}$ in $\mathbb{R}^n$ having $\liminf \mathcal{L}(c\overline{\mathbf{Z}}_N)$ in $S\alpha S$. That is, their chfs satisfy $\{\liminf \phi_{c\overline{\mathbf{Z}}_N}(\mathbf{t})=\exp[\iota\mathbf{t}'\mathrm{d}-\frac{1}{2}(\mathbf{t}'\Sigma\mathbf{t})^{\frac{\alpha}{2}}]\}$ when scaled suitably. Specifically, the distributions $\mathfrak{D}_2$ attracted to Gaussian limits comprise all distributions $\mathcal{L}(\mathbf{Z}_i)$ in $\mathbb{R}^n$ having second moments. More generally, domains of attraction to distributions in $S\alpha S$ have been studied in references,[17–20] to include Lindeberg conditions in Barbosa & Dorea,[21] together with rates of convergence to stable limits in Paulauskas.[22]

Remark 2 That $\phi_{\mathbf{Z}}(\mathbf{t})=\exp[\iota\mathbf{t}'\mathrm{d}-\frac{1}{2}(\mathbf{t}'\Sigma\mathbf{t})^{\frac{\alpha}{2}}]$ has elliptical contours derives from the spherical chf $\phi_{\mathrm{U}}(\mathbf{t})=\exp[\iota\mathbf{t}'\mathrm{q}-\frac{1}{2}(\mathbf{t}'\mathbf{t})^{\frac{\alpha}{2}}]$ through the transformation a $Z=\Sigma^{1/2}U$.

Essentials for $S\alpha S$ distributions

As noted, closed expressions for $S\alpha S$ densities are known in selected cases only, to be complemented by results to follow. Here $g_n(\mathbf{u};\delta,\Sigma)$ is the Gaussian density on $\mathbb{R}^n$ having parameters (δ,Σ), and $f_n^\alpha(z;\delta,\Sigma)$ is the provisional $S\alpha S$ density corresponding to $\phi_{\mathbf{Z}}(\mathbf{t})=\exp[\iota\mathbf{t}'\boldsymbol{\delta}-\frac{1}{2}(\mathbf{t}'\Sigma\mathbf{t})^{\frac{\alpha}{2}}]$. The following properties are essential.

Theorem 1 Let $\mathcal{L}(\mathbf{Z})\in S_n^\alpha(\boldsymbol{\delta},\boldsymbol{\Sigma})$ have the chf $\phi_{\mathbf{Z}}(\mathbf{t})=\exp[\iota\mathbf{t}'\boldsymbol{\delta}-\frac{1}{2}(\mathbf{t}'\mathbb{S}\mathbf{t})^{\frac{\alpha}{2}}]$ and density function $f_n^\alpha(\mathbf{z};\delta,\Sigma)$ if defined. Then the following properties hold.

i. For $\boldsymbol{\Sigma}$ nonsingular, $\mathcal{L}(\mathbf{Z})\in S_n^\alpha(\delta,\Sigma)$ is absolutely continuous in $\mathbb{R}^n$, having a density function $f_n^\alpha(\mathbf{z};\delta,\Sigma)$;

ii. The Gaussian mixture $\phi_{\mathbf{Z}}(\mathbf{t})=\int_0^\infty e^{\iota\mathbf{t}'\boldsymbol{\delta}-\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}/2s}\,d\Psi(s;\alpha)$ holds with $\Psi(s;\alpha)$ as a mixing cdf on $\mathbb{R}^1$;

iii. The Gaussian mixture $f_n^\alpha(\mathbf{z};\delta,\Sigma)=\int_0^\infty g_n(\mathbf{z};\delta,s^{-1}\Sigma)\,d\Psi(s;\alpha)$ holds with $\Psi(s;\alpha)$ as a mixing cdf as before;

iv. $\mathcal{L}(\mathbf{Z})\in S_n^\alpha(\delta,\Sigma)$ is monotone unimodal with mode at $\boldsymbol{\delta}$ for each $\alpha\in(0,2)$;

v. Let $T(\mathbf{Z})=\mathbf{U}\in\mathbb{R}^k$ be scale–invariant; then for $\mathcal{L}(\mathbf{Z})\in S_n^\alpha(\delta,\Sigma)$, the distribution $\mathcal{L}(\mathbf{U})$ is identical to its normal–theory form under $\mathcal{L}(\mathbf{Z})=N_n(\mathbf{z};\delta,\Sigma)$.

Proof: Conclusion (i) is Theorem 6.5.4 of Press.[23] Conclusion (ii) invokes a result of Hartman et al.[24] namely, the process $\{Z_t;t=1,2,\ldots\}$ is spherically invariant if and only if, for each n and $\mathbf{Z}=[Z_1,\ldots,Z_n]$, the chf $\phi_{\mathbf{Z}}(\mathbf{t})$ is a scale mixture of spherical Gaussian chf s on $\mathbb{R}^n$, to give conclusion (ii) on transforming from spherical to elliptical symmetry. To continue, $f_{\mathbf{Z}}(\mathbf{z})=(2\pi)^{-n}\int_{\mathbb{R}^n} e^{-\iota\mathbf{t}'\mathbf{z}}\phi_{\mathbf{Z}}(\mathbf{t})\Lambda(d\mathbf{t})$ is the standard inversion formula from chfs to densities in $\mathbb{R}^n$ with $\Lambda(\cdot)$ as Lebesgue measure, so that from conclusion (ii) we recover

$$f_n^\alpha(z;\delta,\mathbf{I}_n)=\frac{1}{(2\pi)^n}\int_{\mathbb{R}^k} e^{-\iota\mathbf{t}'\mathbf{x}}\int_0^\infty e^{\iota\mathbf{t}'\delta-\mathbf{t}'\mathbf{t}/2s}\,d\Psi(s;\alpha)\Lambda(d\mathbf{t}). \quad (1)$$

Reversing the order of integration inverts the Gaussian chf to give conclusion (iii). Conclusion (iv) follows as in Wolfe[25] in conjunction with conclusion (iii). Finally observe from conclusion (iii), with $\int_0^\infty g_n(Z;\delta,\Sigma/s)d\Psi(s;\alpha)$, that the change of variables $\mathbf{Z}\to\mathbf{U}=T(\mathbf{Z})$ behind the integral is independent of $\Psi(s;\alpha)$ since $T(\mathbf{Z})$ is scale–invariant independently of s, to give conclusion (v).

It remains to reconsider degrees of association in $S\alpha S$ distributions, as distinct from the classical second–moment correlation parameters $\{\rho_{ij}=\sigma_{ij}/(\sigma_{ii}\sigma_{jj})^{\frac{1}{2}}\}$. For $\mathcal{L}(\mathbf{Z})\in S_k^\alpha(\delta,\Sigma)$ with $\alpha<2$, the elements of $\boldsymbol{\Sigma}$ serve instead as scale parameters, since $\mathbf{U}=\Sigma^{-\frac{1}{2}}\mathbf{Z}$ and $\mathbf{U}'\mathbf{U}=\mathbf{Z}'\Sigma^{-1}\mathbf{Z}$ are dimensionless. As to whether $\{\rho_{ij}\}$ again might quantify associations for $\alpha<2$, a definitive answer is supplied in the following.

Lemma 1 Let $\mathcal{L}(\mathbf{Z})\in S_n^\alpha(\delta,\Sigma)$. For $\alpha<2$, the parameters $\{\rho_{ij}=\sigma_{ij}/(\sigma_{ii}\sigma_{jj})^{\frac{1}{2}}\}$ serve to quantify degrees of association between (Z_i,Z_j), the extent of their association increasing with ρ_{ij}.

Proof: It suffices to consider (Z_1,Z_2) centered at (0,0) with $\Sigma=\begin{bmatrix}1 & \rho\\ \rho & 1\end{bmatrix}$. On taking $U=(Z_1-Z_2)$, $\mathcal{L}(U)$ clearly is symmetric about 0 with scale parameter $\sigma_U=2(1-\rho)$. A result of Fefferman et al.[26] shows for each $c>0$ that $P(U\in(-c,c))$ is decreasing in σ_U thus increasing in ρ. Equivalently, $P(|Z_1-Z_2|\le c)\uparrow 1$ as $\rho\uparrow 1$, identifying the sense in which (Z_1,Z_2) become increasingly indistinguishable, thus associated, with increasing values of ρ.

Definition 3 For $\mathcal{L}(\mathbf{Z})\in S_n^\alpha(\delta,\Sigma)$ with $\alpha<2$, the entities $\{\rho_{ij}=\sigma_{ij}/(\sigma_{ii}\sigma_{jj})^{\frac{1}{2}}\}$ are called pseudo–correlation, specifically, α–association parameters

Linear models under $S\alpha S$ errors

The principal findings

Take $\mathcal{L}(\mathbf{Y})\in S_n^\alpha(\mathbf{X}\beta,\sigma^2\mathbf{I}_n)$ with $(\mathbf{X}\beta,\sigma^2\mathbf{I}_n)$ as centering and scale parameters, where $\{\mathbf{Y}\in\mathbb{R}^n,\mathbf{X}\in\mathbb{F}_{n\times k},\beta\in\mathbb{R}^k\}$. *OLS* solutions $\hat\beta=(\mathbf{X'X})^{-1}\mathbf{X'Y}$, as minimally dispersed unbiased linear estimates, are available here only for $\alpha=2$, whereas alternative moment criteria necessarily are subject to moment constraints. Specifically, for scalars $(\hat\theta,\theta)\in\mathbb{R}^1$ under loss $L(\hat\theta,\theta)=|\hat\theta-\theta|$, the risk $R(\hat\theta)=E[L(\hat\theta,\theta)]$ is undefined for $\alpha<1$ as for Cauchy errors at $\alpha=1$. Moreover, risk functions $\{R(\hat\theta)=E(|\hat\theta-\theta|^\kappa)\}$ are defined but concave for $\{\kappa<\alpha<1\}$, and for $\{1<\kappa<\alpha\le 2\}$ are convex, at issue in attaining global optima. Versions of these apply also for vector parameters; however, minimal risk estimation would require not only knowledge regarding α, but also optimizing algorithms. Instead we seek what might be salvaged from classical linear models under the constraints of *OLS* errors. In addition, portions of our findings extend beyond Gauss–Markov theory and *OLS* to include the much larger class of equivariant estimators.

Definition 4 An estimator $\delta(\mathbf{Y})$ for $\beta\in\mathbb{R}^k$ is translation–equivariant if for $\{\mathbf{Y}\to\mathbf{Y}+\mathbf{Xb}\}$, then $\{\delta(\mathbf{Y}+\mathbf{Xb})=\delta(\mathbf{Y})+\mathbf{b}\}$ for every $\mathbf{b}\in\mathbb{R}^k$.

On taking $\mathbf{P}=[\mathbf{I}_n-\mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'}]$, the elements of $\mathbf{e}=\mathbf{PY}$ comprise the observed residuals and $S^2=\mathbf{e'e}/(n-k)$ the residual mean square. Normal–theory tests for $H_0{:}\beta=\beta_0$ against $H_1{:}\beta\ne\beta_0$ utilize $F=(\hat\beta-\beta_0)'\mathbf{X'X}(\hat\beta-\beta_0)/S^2$ having the distribution $F(u;k,n-k,\lambda)$ with $\lambda=(\beta-\beta_0)'\mathbf{X'X}(\beta-\beta_0)/\sigma^2$. We proceed to examine essential properties of $S_n^\alpha(\mathbf{X}\beta,\sigma^2\mathbf{I}_n)$ as α ranges over (0,2), where some expressions simplify on taking $\sigma^2=1$, then reinstating σ^2 as needed. The following properties are fundamental.

Theorem 2 Given $\mathcal{L}(\mathbf{Y})=S_n^\alpha(\mathbf{X}\beta,\sigma^2\mathbf{I}_n)$, consider $[\hat\beta,\mathbf{e}]$ with $\mathbf{e}=\mathbf{PY}$ as the residual vector, and $U=(n-k)S^2/\sigma^2$. Then

(i) $\mathcal{L}(\hat\beta,\mathbf{e})=S_{n+k}^\alpha([\beta,\mathbf{0}],\Sigma)$, with $\Sigma=\sigma^2 Diag((\mathbf{X'X})^{-1},\mathbf{P})$, a distribution on $\mathbb{R}^{n+k}$ of rank s

(ii) The marginal's are $\mathcal{L}(\hat\beta)=S_k^\alpha(\beta,\sigma^2(\mathbf{X'X})^{-1})$ centered at β with scale parameters $\sigma^2(\mathbf{X'X})^{-1}$, and

(iii) $\mathcal{L}(\mathbf{e})=S_n^\alpha(\mathbf{0},\sigma^2\mathbf{P})$ on $\mathbb{R}^n$ of rank $n-k$ centered at $\mathbf{0}$ with scale parameters $\sigma^2\mathbf{P}$;

(iv) $U=(n-k)S^2/\sigma^2$ has density $f(u;\nu,\alpha)=\int_0^\infty h(u;\nu,s)d\Psi(s;\alpha)$ with $h(u;\nu,s)$ as the central chi–squared density on $\nu=(n-k)$ degrees of freedom, scaled by s, and with $\Psi(s;\alpha)$ as a mixing distribution.

Proof. Let $\mathbf{L'}=(\mathbf{X'X})^{-1}\mathbf{X'}$ and $\mathbf{P}=\left[\mathbf{I}_n-\mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'}\right]$ to project onto the error space, so that $\mathbf{G}=[\mathbf{L},\mathbf{P}]$ operates on $\mathbf{Y}$ to give

$$\mathbf{Z}=\mathbf{G'Y}=\begin{bmatrix}\hat\beta\\ \mathbf{e}\end{bmatrix}=\begin{bmatrix}\mathbf{L'}\\ \mathbf{P'}\end{bmatrix}\mathbf{Y}\in\mathbb{R}^{n+k}\ \textit{and}\ \mathbf{G'G}=\begin{bmatrix}(\mathbf{X'X})^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{P}\end{bmatrix},\qquad(2)$$

the latter of order $[(n+k)\times(n+k)]$ and rank n. The chf with argument $\mathbf{s}=[s_1,\ldots,s_{n+k}]$ is $E[\exp(\iota\mathbf{s'Z})]=E[\exp(\iota\mathbf{s'G'Y})]=E[\exp(\iota\mathbf{v'Y})]=\phi_{\mathbf{Y}}(\mathbf{v})$ with argument $\mathbf{v}=\mathbf{Gs}$ replacing $\mathbf{t}$, to give conclusion

(i). Next partition $\mathbf{s'}=[\mathbf{s}_1',\mathbf{s}_2']$ with $\mathbf{s}_1'=[s_1,\ldots,s_k]$, to obtain

$$\phi_{\mathbf{Z}}(\mathbf{s})=\exp[\iota\mathbf{s'G'X}\beta-\frac{1}{2}(\mathbf{s'G'Gs})^{\frac{\alpha}{2}}]=\exp[\iota\mathbf{s}_1'\beta-\frac{1}{2}(\mathbf{s}_1'(\mathbf{X'X})^{-1}\mathbf{s}_1+\mathbf{s}_2'\mathbf{P}\mathbf{s}_2)^{\frac{\alpha}{2}}].$$

The marginal chfs of $\hat\beta$ and $\mathbf{e}$ follow on setting $\mathbf{s}_2=\mathbf{0}$, then $\mathbf{s}_1=\mathbf{0}$ in succession, to give conclusions (ii) and (iii). Conclusion (iv) attributes to Hartman et al.[24] through Theorem 1. Specifically, a change of variables $\mathbf{u}\to\mathbf{e}\to\mathbf{e'e}=(n-k)S^2$ behind the integral on the right of Theorem 1(iii) gives the conditional density for $\mathcal{L}((n-k)S^2|s)$, namely the scaled chi–squared density $h(u;\nu,s)$ depending on s, so that integrating with respect to $d\Psi(s;\alpha)$ gives conclusion (iv).

Remark 3 That $\Sigma=\sigma^2 Diag(\mathbf{X'X},\mathbf{P})$ is block–diagonal in conclusion (i), assures under $S\alpha S$ errors that $(\hat\beta,\mathbf{e})$ are α–unassociated as in Definition 3, well known to be mutually uncorrelated under second moments.

It remains to reexamine topics in inference under $S\alpha S$ errors. The following are germane.

Definition 5 An estimator $\hat\theta$ for $\theta\in\mathbb{R}^k$ is said to be linearly median unbiased if and only if the median $\text{med}(\mathbf{a}'\hat\theta)=\mathbf{a}'\theta$ for each $\mathbf{a}\in\mathbb{R}^k$; and to be modal unbiased provided that the mode $M(\hat\theta)=\theta$.

Definition 6 An estimator $\hat\theta$ for θ is said to be more concentrated about θ than $\tilde\theta$ provided that $P((\hat\theta-\theta)\in C_0)\ge P((\tilde\theta-\theta)\in C_0)$ for every convex set C_0 in $\mathbb{R}^k$ symmetric under reflection about $\mathbf{0}\in\mathbb{R}^k$.

Essential properties under $S\alpha S$ errors include the following.

Theorem 3 For $\mathcal{L}(\mathbf{Y})=S_n^\alpha(\mathbf{X}\beta,\sigma^2\mathbf{I}_n)$, consider properties of the *OLS* solutions $\hat\beta=(\mathbf{X'X})^{-1}\mathbf{X'Y}$, and of the equivariant estimators $\tilde\beta=\delta(\mathbf{Y})$. of Definition 4.

i. $\hat\beta$ is unbiased for β for each $\{1<\alpha\le 2\}$;

ii. $\hat\beta$ is linearly median unbiased for β;

iii. $\hat{\beta}$ is most concentrated about β among all median–unbiased linear estimators;

iv. $\tilde{\beta}$ is modal unbiased for β;

v. $\hat{\beta}_N$ is consistent for β in a sequence of N identical but dependent experiments $\{\mathbf{Y}_i=\mathbf{X}\beta+\mathbf{e}_i; i=1,2,\ldots,N\}$;

vi. The null distribution of $F=(\hat{\beta}-\beta_0)'\mathbf{X'X}(\hat{\beta}-\beta_0)/S^2$ has exactly its normal–theory form; the power increases with increasing $\lambda=(\beta-\beta_0)'\mathbf{X'X}(\beta-\beta_0)/\sigma^2$; and such tests are unbiased;

vii. $\hat{\beta}$ is most concentrated about β among all modal–unbiased linear estimators.

viii. $\hat{\beta}$ is most concentrated about β among all equivariant estimators $\tilde{\beta}=\delta(\mathbf{Y})$.

Proof. Conclusions (i)–(vi) carry over from reference Jensen[27] without benefit of moments, regardless of membership in the $S\alpha S$ class. To consider concentration properties of modal–unbiased estimators, begin with $\phi_{\mathbf{Y}}(\mathbf{t})=\exp[\iota\mathbf{t'X}\beta-\frac{1}{2}(\mathbf{t't})^{\frac{\alpha}{2}}]$, and consider $\tilde{\beta}=\mathbf{L'Y}$ with $\mathbf{L'}=[(\mathbf{X'X})^{-1}\mathbf{X'},\mathbf{G'}]$, so that

$$\phi_{\tilde{\beta}}(\mathbf{s})=\exp[\iota\mathbf{s'L'X}\beta-\frac{1}{2}(\mathbf{s'L'Ls})^{\frac{\alpha}{2}}];$$

$$\mathbf{s'L'X}\beta=\mathbf{s'}[(\mathbf{X'X})^{-1}\mathbf{X'},\mathbf{G'}]\mathbf{X}\beta.$$

That $\tilde{\beta}$ should have mode at β, it is necessary that $\mathbf{s'L'X}\beta=\mathbf{s'}\beta$, i.e. $\mathbf{G'X=0}$. accordingly, $\phi_{\tilde{\beta}}(\mathbf{s})=\exp[\iota\mathbf{s'}\beta-\frac{1}{2}[\mathbf{s'}\Omega\mathbf{s}]^{\frac{\alpha}{2}}]$, with $\Omega=\mathbf{L'L}=[(\mathbf{X'X})^{-1}+\mathbf{G'G}]$. Clearly the matrix $[\mathbf{L'L}-(\mathbf{X'X})^{-1}]=\mathbf{G'G}$ is positive semi definite, giving conclusion (vii) from Jensen.[28] Conclusion (viii) follows from Theorem 2.7 of Burk et al.[29] since $S\alpha S$ distributions are unimodal from Theorem 1(iv).

Spherical cauchy errors

Spherical multivariate t errors on ν degrees of freedom trace to Zellner[30] to include Cauchy errors at $\nu=1$, equivalently, at $\alpha=1$ in the class $S\alpha S$. Specializing from Theorem 1(ii), the spherical Cauchy chf is $\phi_{\mathbf{Z}}(\mathbf{t})=\exp[\iota\mathbf{t'}\mathrm{d}-\frac{1}{2}(\mathbf{t't})^{\frac{1}{2}}]$. Recast in terms of linear inference, we have the following specialization of Theorems 1 and 2.

Corollary 1 Under the conditions of Theorems 1 and 2, the following properties hold under spherical Cauchy errors.

i. The spherical Cauchy density on $\mathbb{R}^n$ at $\alpha=1$ is

$$f_n^1(\mathbf{z};\delta,\mathbf{I}_n)=\int_0^\infty g_n(\mathbf{z};\delta,s^{-2}\mathbf{I}_n)d\Psi(s;1)$$

$$=c(n)\left[1+(\mathbf{z}-\delta)'(\mathbf{z}-\delta)\right]^{-\frac{n+1}{2}}$$

$$c(n)=\Gamma(\frac{n+1}{2})/\pi^{\frac{n+1}{2}}$$

where $d\Psi(s;1)=e^{-\frac{s^2}{2}}/(2\pi)^{\frac{1}{2}}$, the mixing $\chi(s;1)$ density.

ii. The elliptical Cauchy density for $\hat{\beta}$ on $\mathbb{R}^k$ is

$$f_k^1(\hat{\beta};\beta,\mathbf{X'X})=c(k)\left[1+(\hat{\beta}-\beta)'\mathbf{X'X}(\hat{\beta}-\beta)\right]^{-\frac{k+1}{2}} \quad (3)$$

Proof. The multivariate t–distribution on $\mathbb{R}^n$ is that of $\{T_i=Y_i/S; 1\le i\le n\}$ from $\mathcal{L}(\mathbf{Y})=N_n(\delta,\sigma^2\mathbf{I}_n)$, with S as a sample standard deviation on ν degrees of freedom, known to be spherical Cauchy at $\nu=1$. This gives conclusion (i) on specializing the conventional multivariate t density. Conclusion (ii) follows directly on specializing Theorem 2(ii) at $\alpha=1$

Case study

The viability (Y_i) for each of $n=13$ biological specimens was recorded after storage under additives X_{i1} and X_{i2} as listed in Table 1, p. 408 of Walpole & Myers.[31] The model is $\{Y_i=\beta_0+\beta_1X_{i1}+\beta_2X_{i2}+\varepsilon_i\}$, where the errors are taken to be spherical Cauchy. The conventional OLS solutions are $\hat{\beta}_0=36.094$, $\hat{\beta}_1=1.031$, $\hat{\beta}_2=-1.870$, as elements of $\hat{\beta}=[\hat{\beta}_0,\hat{\beta}_1,\hat{\beta}_2]'$. The matrix $\mathbf{X'X}$, its inverse $(\mathbf{X'X})^{-1}$, and the transition of the latter into its α–association form of Definition 3 are given respectively by

$$\begin{bmatrix} 13 & 59.43 & 81.82 \\ 59.43 & 394.7255 & 360.6621 \\ 81.82 & 360.6621 & 576.7264 \end{bmatrix}^{-1}=\begin{bmatrix} 1.0114 & -0.0494 & -0.1126 \\ -0.0494 & 0.0083 & 0.0018 \\ -0.1126 & 0.0018 & 0.0166 \end{bmatrix}\rightarrow\begin{bmatrix} 1 & -0.5392 & -0.8690 \\ -0.5392 & 1 & 0.1533 \\ -0.8690 & 0.1533 & 1 \end{bmatrix}.$$

The following properties are evident.

i. The elliptical Cauchy density for $\hat{\beta}$ is given by equation (3) with $k=3$ and $\mathbf{X'X}$ as listed for these data.

ii. The solution $\hat{\beta}$ is both linear median–unbiased and modal–unbiased, and among all such estimators is most concentrated about β.

iii. The normal–theory confidence set $\{\beta\in(\hat{\beta}-\beta)'\mathbf{X'X}(\hat{\beta}-\beta)\le S^2c_\gamma\}$ holds exactly with confidence coefficient $1-\gamma=0.95$, where $S^2=4.001$ is the residual mean square on $\nu=10$ degrees of freedom, and $c_\gamma=3.71$ is the upper 0.95 percentile for $F(3,10,0)$.

iv. As correlations are undefined, elements of the α–association matrix nonetheless do serve to quantify the degrees of association among $[\hat{\beta}_0,\hat{\beta}_1,\hat{\beta}_2]$ as in Definition 3, on taking $\alpha=1$ in Lemma 1.

v. In particular, $\hat{\beta}_0$ is negatively associated with $(\hat{\beta}_1,\hat{\beta}_2)$, whereas $(\hat{\beta}_1,\hat{\beta}_2)$ are themselves positively associated.

Table 1 The viability (Y_i) of $n=13$ biological specimens after storage under additives X_{i1} and X_{i2}

Y_i	X_{i1}	X_{i2}	Y_i	X_{i1}	X_{i2}
25.5	1.74	3.3	31.2	6.32	5.42
25.9	6.22	8.41	38.4	10.52	4.63
18.4	1.19	11.6	26.7	1.22	5.85
26.4	4.1	6.62	25.9	6.32	8.72
32.0	4.08	4.42	25.2	4.15	7.6
39.7	10.15	4.83	35.7	1.72	3.12
26.5	1.7	5.3			

Summary and discussion

This study offers further insight into the class $S\alpha S$ comprising the spherical α–stable laws as limit distributions under conditions for central limit theory. In addition to their essential properties, expanded here to include representations for density functions, this study focuses on models of type $\{\mathbf{Y}=\mathbf{X}\beta+\mathbf{e}\}$ when devoid of moments undergirding the classical theory. Recall that normal–theory procedures routinely are applied in practice as large–sample approximations in distributions attracted to Gaussian laws. Specifically, Berry–Esséen bounds on rates of convergence to Gaussian limits are given Jensen,[32,33] with special reference to linear models in Jensen.[34,35] Results here validate corresponding large–sample approximations for distributions attracted to $S\alpha S$ laws as cited in references.[17–21] Of similar importance are rates of convergence to stable limits as in Paulauskas.[22] By showing that many standard properties carry over in essence under significantly weakened assumptions, this study gives further credence to the widely and correctly held view that Gauss–Markov estimation and normal theory inferences extend considerably beyond the confines of the classical theory.

A appendix

The preceding study has developed exclusively around spherically dependent $S\alpha S$ errors, as alternative to iid stable errors. This choice is prompted by discrepancies encountered in the simplest case $\{Z_i \to Z_i+\delta; i=1,2,\ldots,N\}$ with common location parameter. Essential details from Jensen[14] may be summarized as follows. To distinguish the disparate properties of iid vs spherical $S\alpha S$ models, sequences $\mathbb{Z}=\{Z_1,Z_2,Z_3,\ldots\}$ are fundamental in order to take limits. Of significance is that averages of $S\alpha S$ sequences with $\alpha<2$ may be inconsistent for iid sequences but consistent under $S\alpha S$ symmetry. Accordingly, let $\mathcal{L}_\infty(Z_N)=\liminf \mathcal{L}(Z_N)$. Essentials follow.

Lemma 2 Given $\mathbb{Z}=\{Z_1,Z_2,Z_3,\ldots\}$, consider the case that $\mathbf{Z}'=[Z_1,\ldots,Z_N]$ either are iid $S_1^\alpha(\delta,1)$, with chf $\phi_{Z_i}(t)=\exp\{\iota t\delta-|t|^\alpha\}$, or are $S\alpha S$ on $\mathbb{R}^N$ with chf $\phi_{\mathbf{Z}}(\mathbf{t})=\exp\{\iota\delta\mathbf{t}'\mathbf{1}_N-(\mathbf{t}'\mathbf{t})^{\frac{\alpha}{2}}\}$. Let $S_N=(Z_1+\ldots+Z_N)$ and $\overline{Z}_N=N^{-1}S_N$, and consider the standardized variables $U_N=N^{\frac{1}{2}}(\overline{Z}_N-\delta)$.

i. Consistent and inconsistent properties of $\overline{Z}_N$ for iid sequences are as follow.

For $0<\alpha<1$: $\phi_{\overline{Z}_N}(t)=e^{\iota t\delta-N^{\varepsilon}|t|^\alpha}$ for $\varepsilon>0$, so that $\overline{Z}_N$ is inconsistent for δ.

For $\alpha=1$, $\phi_{\overline{Z}_N}(t)=e^{\iota t\delta-|t|^\alpha}\equiv\phi_{Z_i}(t)$, so that $\overline{Z}_N$ is inconsistent for δ.

For $1<\alpha\leq 2$, $\phi_{\overline{Z}_N}(t)=e^{\iota t\delta-N^{-\varepsilon}|t|^\alpha}$ for $\varepsilon>0$, so that $\overline{Z}_N$ is consistent for δ.

ii. For SaS sequences $\overline{Z}_N$ is consistent for δ. for every $0<\alpha\leq 2$.

iii. For iid sequences with $0<\alpha<2$, $\mathcal{L}_\infty(U_N)$ diverges to an improper distribution.

iv. For $S\alpha S$ sequences $\liminf \mathcal{L}(U_N)=\mathcal{L}(Z_i)$, the limit being identical to each component.

Acknowledgements

None.

Conflict of interest

Authors declare that there is no conflict of interest.

References

1. Cheng BN, Rachev ST. Multivariate stable futures prices. *Math Finance*. 1995;5:133–153.
2. Kuruoglu E, Zerubia J. Modeling SAR Images with a generalization of the Rayleigh distribution. *IEEE Trans Image Processing*. 2009;13(4):527–533.
3. Qiou Z, Ravishanker N, Dey DK. Multivariate survival analysis with positive stable frailties. *Biometrics*. 1999;55(2):637–644.
4. Bonato M. Modeling fat tails in stock returns: A multivariate stable–GARCH approach. *Computational Statist*. 2012;27(3):499–521.
5. Kim YS, Glacometti R, Rachev ST, et al. Measuring financial risk and portfolio optimization with a non–Gaussian multivariate model. *Ann Operations Research*. 2012;201(1):325–343.
6. Tsionas EG. Estimating multivariate heavy tails and principal directions easily, with an application to international exchange rates. *Statist Probability Letters*. 2012;82(11):1986–1989.
7. Samorodnitsky G, Taqqu M. *Stable Non–Gaussian Random Processes*. USA: Chapman–Hall; 1994.
8. Arce GR. *Nonlinear Signal Processing*. New York: Wiley; 2000.
9. Chernobai A, Rachev ST, Fabozzi FJ. *Operational Risk: A Guide to Basel II Capital Requirements, Models and Analysis*. USA: Wiley; 2007.
10. Nolan JP. Multivariate elliptically contoured stable distributions: Theory and estimation. *Computational Statist*. 2013;28(5):2067–2089.
11. Fang KT, Anderson TW. *Statistical Inference in Elliptically Contoured and Related Distributions*. New York: Allerton Press; 1990.
12. Fang KT, Zhang YT. *Generalized Multivariate Analysis*. Beijing: Science Press; 1990.
13. Gupta AK, Varga T. *Elliptically Contoured Models in Statistics*. Dordrecht: Kluwer Academic Publishers; 1993.
14. Jensen DR. Limits and inferences for alpha–stable variables. B*iostat Biometrics Open Acc J*. 2017;4(1):1–4.
15. Lukacs E, Laha RG. *Applications of Characteristic Functions*. New York: Hafner; 1964.
16. Dharmadhikari S, Joag Dev K. *Unimodality, Convexity, and Applications*. New York: Academic; 1988.
17. Rvaceva EL. On domains of attraction of multidimensional dlstributions. *Transl Math Statist Prob*. 1962;2:183–205.
18. Kuelbs J, Mandrekar V. Domains of attraction of stable measures on a Hilbert space. *Stucia Math*. 1974;50:149–162.
19. De Haan L, Resnick SI. Derivatives of regularly varying functions in R^d and domains of attraction of stable distributions. *Stochastic Processes Applic*. 1979;8(3):349–356.
20. Ruschendorf L, Rueschendorf L, Rachev ST, et al. A new ideal metric with applications to multivariate stable limit theorems. *Probability Theory and Related Fields*. 1992;94(2):163–187.
21. Barbosa EG, Dorea CCY. A note on the Lindeberg condition for convergence to stable laws in Mallows distance. *Bernoulli*. 2009;15(3):922–924.

22. Paulauskas V. On the rate of convergence to bivariate stable laws. *Lithuanian Math J*. 2009;49:426–445.

23. Press SJ. *Applied Multivariate Analysis*. USA: Holt, Rinehart, Winston; 1972.

24. Hartman P, Wintner A. On the spherical approach to the normal distribution law. *Amer J Math*.1940;62(1):759–779.

25. Wolfe SJ. On the unimodality of spherically symmetric stable distribution functions. *J Multvariate Anal*. 1975;5(1):236–242.

26. Fefferman C, Jodeit M, Perlman MD. A spherical surface measure inequality for convex sets. *Proc Amer Math Soc*. 1972;33(1):114–119.

27. Jensen DR. Linear models without moments. *Biometrika*. 1979;66(3):611–617.

28. Jensen DR. Ordering ellipsoidal measures: Scale and peakedness orderings. *SIAM J Appl Math*. 1984;44(6):1226–1231.

29. Burk R, Hwang JT. Optimality of the least squares estimator. *J Multivariate Analysis*. 1989;30(2):245–254.

30. Zellner A. Bayesian and non–Bayesian analysis of the regression model with multivariate Student–t error terms. *J Amer Statist Assoc*. 1976;71:400–405.

31. Walpole RE, Myers RH. Probability and Statistics for Engineers and Scientists, 4^{th} ed. USA: Macmillan; 1989.

32. Jensen DR. Monotone bounds on the chi–squared approximation to the distribution of Pearson's X^2 statistics. *Australian J Statist*. 1973;15(2):65–70.

33. Jensen DR. On approximating the distribution of Friedman's X^2_r and related statistics. *Metrika*. 1977;24(1):75–85.

34. Jensen DR, Mayer LS. Normal–theory approximations to tests for linear hypotheses. *Ann Statist*. 1975;3(2):429–444.

35. Jensen DR, Mayer LS, Myers RH. Optimal designs and large–sample tests for linear hypotheses. *Biometrika*. 1975;62(1):71–78.