

Autoregressive Moving Average Models Under Exponential Power Distributions

Roger W. Barnard^a, A. Alexandre Trindade ^{†a}, R. Indika P. Wickramasinghe^b

^a*Department of Mathematics & Statistics, Texas Tech University, Lubbock, TX 79409-1042, U.S.A.*

^b*Department of Mathematical Sciences, Eastern New Mexico University, Portales, NM 88130, U.S.A.*

Abstract. Distributional and inferential results are derived for ARMA models under the multivariate exponential power (EP) distribution, which includes the Gaussian and Laplace as special cases. Marginal distributions are obtained in the important Laplace case, both under a multivariate Laplace distributed process, as well as a process driven by univariate Laplace noise. Asymptotic results are established for the maximum likelihood estimators under a full EP likelihood, and under a conditional likelihood resulting from a driving univariate EP noise distribution. Whenever tractability permits, results are fully worked out with respect to the MA(1) model. The methodology is illustrated on a fund of real estate returns.

1. Introduction

The modeling of time series data, $X_t, t = 1, \dots, n$, under different distributional assumptions has generated perennial attention over the years. In the classical parametric autoregressive moving average (ARMA) framework, maximization of a Gaussian likelihood is implemented by default in most packages. This approach can be viewed as assuming a multivariate normal joint distribution for the data vector. Of course, a Gaussian likelihood is not a bad default choice as it corresponds closely to least squares; both estimators enjoying the same asymptotic regime by requiring that the driving white noise sequence, Z_t , be only independent and identically distributed (iid) (Brockwell and Davis, 1991).

An alternative to the full likelihood specification that yields maximum likelihood estimates (MLEs) is to maximize a conditional likelihood, resulting in *quasi* or *conditional* MLEs (QMLEs). This approach treats the (unobserved iid) white noise sequence as the data, and as such the likelihood is formed by multiplying together the density function of Z_t evaluated at the empirical values, z_t , which can be related back to the (observed) x_t through the ARMA recursions, after appropriate truncation to account for unobserved initial values. In this vein, the milestone paper by Davis and Dunsmuir (1997) establishes consistency and asymptotic normality for ARMA model parameters under minimization of a least absolute deviations (LAD) criterion, which corresponds to QMLE with Laplace noise.

There are certainly instances when such alternative likelihoods or optimality criteria may be warranted. Damsleth and El-Shaarawi (1988) employed an ARMA model driven by Laplace noise to fit weekly data on sulphate concentration in a Canadian watershed. Anderson and Arnold (1993) observed that IBM daily stock price returns are adequately modeled by Linnik processes. Recent work by Andrews *et al.* (2009) suggests that stable distributions provide useful models for certain types of financial data like asset returns

Keywords. Asymptotic distribution; elliptically contoured; QMLE; marginal distribution; maximum likelihood estimation; multivariate Laplace.

Received: 28 February 2014; Accepted: 13 August 2014

[†]Corresponding author

Email address: alex.trindade@ttu.edu (A. Alexandre Trindade [†])

and traded stock volume, and they propose a method of maximum likelihood estimation for stable AR processes. Trindade *et al.* (2010) investigated ARMA and GARCH models driven by asymmetric Laplace innovations, and found them particularly suited for processes that are skewed, peaked, and leptokurtic, but which appear to have some higher order moments.

To this end, this paper seeks to derive some distributional results for ARMA models under a flexible multivariate family, the *exponential power* (EP) distribution, first proposed in a multivariate setting by Gomez *et al.* (1998). A member of the elliptically contoured family, the shape parameter β of the EP provides flexible control over the thickness of the tail, such that the classical Gaussian ($\beta = 1$) and Laplace ($\beta = 1/2$) are obtained as special cases. One of the key results we will establish, is that maximizing a full EP likelihood is exactly equivalent to a full Gaussian likelihood, thus further strengthening the case for software packages to use it by default. However, in the case of QMLE the estimates will differ depending on β .

Although we strive for general ARMA results, in certain cases these will be illustrated with respect to the first-order moving average model, MA(1),

$$X_t = \theta_0 Z_{t-1} + Z_t, \quad Z_t \sim \text{iid} (0, \sigma^2), \quad (1)$$

where $\{Z_t\}$ is sequence of zero-mean independent and identically distributed (iid) random variables with variance σ^2 . Seemingly a simple model, for n observations the (Gaussian) MLE for the (true) coefficient θ_0 is the root of a polynomial of degree (approximately) $2n$ in θ . In fact, the distribution of the MLE, and simpler estimators like method of moments and least squares, is a mixture of a continuous density on $(-1, 1)$ and point masses at ± 1 (Paige *et al.*, 2014). Basic issues concerning asymptotic behavior in the unit-root (and near unit-root) case, $\theta_0 = -1$, were not settled until the *Koopmans Econometrics Theory Prize* winning paper by Davis and Dunsmuir (1996).

The rest of this paper is organized as follows. Section 2 discusses the EP family of distributions, presents some new results concerning the family, and sets the stage for subsequent sections. To this end, section 3 derives results concerning: (i) the one and two-dimensional marginal distributions of ARMA models under a joint Laplace distribution, and (ii) the one-dimensional marginal distribution of an MA(1) model driven by iid EP noise for select values of β . Section 4 establishes general asymptotic results for the estimates of the parameters of an ARMA model under these two regimes: (i) MLEs corresponding to a full EP likelihood, and (ii) QMLEs corresponding to a conditional likelihood resulting from a driving EP noise distribution. We conclude in sections 5 and 6 with an illustrative application and some summary remarks.

2. The Exponential Power Family

The EP family of distributions was first proposed in a multivariate setting by Gomez *et al.* (1998). It is a member of the elliptically contoured family, which is extensively documented in Fang, Kotz, and Ng (1990), and Fang and Zhang (1990). Fang and Anderson (1990) contains further studies on distributional properties and applications in statistical inference. The n -dimensional random vector $\mathbf{y} \sim EC_n(\boldsymbol{\mu}, \Sigma, \phi)$ is *elliptically contoured* with location and dispersion parameters $\boldsymbol{\mu}$ and Σ , if its characteristic function is of the form $\Xi(\mathbf{s}) = e^{i\mathbf{s}'\boldsymbol{\mu}} \phi(\mathbf{s}'\Sigma\mathbf{s})$, for some function $\phi(u)$, $u \in \mathbb{R}$. An equivalent characterization is through the “generator” function of the pdf, $h(\cdot)$, in which case we write $\mathbf{y} \sim EC_n(\boldsymbol{\mu}, \Sigma, h)$, with pdf

$$f(\mathbf{y}) = |\Sigma|^{-1/2} c_n h(z) \equiv |\Sigma|^{-1/2} g(z), \quad \text{with } z = (\mathbf{y} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{y} - \boldsymbol{\mu}) \in \mathbb{R}^+, \quad (2)$$

where the normalizing constant c_n is given by

$$c_n = \frac{\Gamma(n/2)}{\pi^{n/2} \iota_n}, \quad \iota_n = \int_0^\infty z^{n/2-1} h(z) dz = 2 \int_0^\infty u^{n-1} h(u^2) du. \quad (3)$$

If the first and second moments exist, then $\mathcal{E}(\mathbf{y}) = \boldsymbol{\mu}$ and $\text{Var}(\mathbf{y}) = -2\phi'(0)\Sigma$.

Provost and Cheong (2002, Lemma 2) presented an expression for the pdf of $\mathbf{y} \sim EC_n(\boldsymbol{\mu}, \Sigma, \phi)$, as

$$f(\mathbf{y}) = \int_0^\infty w(t) \phi_n(\mathbf{y}; \boldsymbol{\mu}, \Sigma/t) dt, \quad (4)$$

where $\phi_n(\mathbf{y}; \boldsymbol{\mu}, \Sigma/t)$ denotes the pdf of an n -dimensional normal with mean $\boldsymbol{\mu}$ and covariance matrix Σ/t , and $w(t)$ is a “weighting” function that is defined through the inverse Laplace transform of $f(t)$, $t = (\mathbf{y} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{y} - \boldsymbol{\mu})/2 \equiv z/2$, denoted by $\mathcal{L}_f^{-1}(t)$. More explicitly, we have

$$w(t) = (2\pi)^{n/2} |\Sigma|^{1/2} t^{-n/2} \mathcal{L}_f^{-1}(t), \quad \text{with} \quad \int_0^\infty w(t) dt = 1, \tag{5}$$

which results in the following relationship between the pdf at z and $w(t)$,

$$f(z) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \int_0^\infty t^{n/2} e^{-tz/2} w(t) dt.$$

Two noteworthy special cases of the elliptically contoured family with finite moments of all orders are presented in Table 1. The weighting function for the normal is the Dirac delta centered at 1. The Laplace listed here is a scaled version of the “bilateral exponential” in Table 1 of Provost and Cheong (2002). Both the Gaussian and Laplace are special cases of the EP, with shape parameters $\beta = 1$ and $\beta = 1/2$ respectively.

Table 1: Two classical elements of the elliptically contoured family, the Gaussian and Laplace, which are elements of the exponential power sub-family (Gomez et al., 1998), with shape parameters $\beta = 1$ and $\beta = 1/2$ respectively.

Distribution	$\phi(u)$	$g(z)$	$w(t)$
Gaussian	$e^{-u/2}$	$(2\pi)^{-n/2} e^{-z/2}$	$\delta(t - 1)$
Laplace	Not Explicit	$\frac{(2\pi)^{-n/2} \Gamma(n/2)}{\Gamma(n) 2^{1+n/2}} e^{-\sqrt{z}/2}$	$\frac{\Gamma(n/2) e^{-1/(8t)} t^{-(n+3)/2}}{\Gamma(n) 2^{(5+n)/2} \sqrt{\pi}}$

If $\mathbf{y} \sim EP_n(\boldsymbol{\mu}, \Sigma, \beta)$ denotes an n -dimensional EP random vector, the values $\boldsymbol{\mu}$, Σ , and β play the role of location, scale, and shape parameters, with pdf given by

$$f(z) = \frac{n\Gamma(n/2)}{\pi^{n/2} \Gamma(1 + \frac{n}{2\beta}) 2^{1+\frac{n}{2\beta}}} |\Sigma|^{-1/2} \exp\{-z^\beta/2\}, \quad z = (\mathbf{y} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{y} - \boldsymbol{\mu}).$$

The mean and variance are related to the location and scale parameters by the relations

$$\mathcal{E}(\mathbf{y}) = \boldsymbol{\mu}, \quad \text{and} \quad \text{Var}(\mathbf{y}) = c_{n,\beta} \Sigma, \quad \text{where} \quad c_{n,\beta} = \frac{2^{1/\beta} \Gamma\left(\frac{n+2}{2\beta}\right)}{n\Gamma\left(\frac{n}{2\beta}\right)}. \tag{6}$$

The EP is a multivariate generalization of the univariate exponential power distribution (EPD), where $\beta > 0$ controls the thickness of the tails: for $\beta < 1$ ($\beta > 1$) the tails are heavier (lighter) than the normal. Gomez et al. (1998) originally termed this *power exponential*, but later realized the inadvertent switching of the names (Gomez et al., 2002). The EPD is also variously called *Subbotin*, *Generalized Error Distribution* (Mineo and Ruggieri, 2005), and *Generalized Normal Distribution* (Nadarajah, 2005), with slight differences in the parametrizations. However, Gomez et al. (1998) were the first to provide a multivariate extension of the EPD.

It is obvious that $\mathbf{y} \sim EP_n(\boldsymbol{\mu}, \Sigma, \beta)$ implies $\mathbf{y} \sim EC_n(\boldsymbol{\mu}, \Sigma, h)$, with generator function $h(z) = \exp\{-z^\beta/2\}$. Gomez et al. (1998) gave the following expression for the function $\phi(u)$ that defines the characteristic function $\Xi(\mathbf{s}) = e^{i\mathbf{s}'\boldsymbol{\mu}} \phi(\mathbf{s}'\Sigma\mathbf{s})$:

$$\phi(u) = \frac{n}{\Gamma(1 + \frac{n}{2\beta}) 2^{1+\frac{n}{2\beta}}} \int_0^\infty \Phi_n(ru) r^{n-1} \exp\{-r^{2\beta}/2\} dr, \tag{7}$$

where for $x > 0$ and $n \geq 2$

$$\Phi_n(x) = \frac{\Gamma(n/2)}{\sqrt{\pi} \Gamma(\frac{n-1}{2})} \iota_n(x), \quad \text{and} \quad \iota_n(x) = \int_0^\pi e^{ix \cos \theta} \sin^{n-2} \theta d\theta. \tag{8}$$

In this section we state and prove two new results concerning the EP distribution that will be needed in subsequent sections. The following Proposition simplifies the expression for $\phi(u)$.

Proposition 2.1. For $n \geq 2$, the expression for $\phi(u)$ in the characteristic function of $\mathbf{y} \sim EP_n(\boldsymbol{\mu}, \Sigma, \beta)$, becomes

$$\phi(u) = \frac{\beta\Gamma(n/2)}{\Gamma(\frac{n}{2\beta})} 2^{\frac{n(\beta-1)}{2\beta}} u^{1-n/2} \int_0^\infty r^{n/2} J_{n/2-1}(ru) \exp\{-r^{2\beta}/2\} dr,$$

where $J_m(\cdot)$ denotes a Bessel function of the 1st kind of order m .

Proof. From Watson (1944, p.48), it follows that for any $x > 0$ and $n \geq 2$, $\iota_n(x)$ in (8) can be expressed as

$$\iota_n(x) = \sqrt{\pi} \Gamma\left(\frac{n-1}{2}\right) \left(\frac{2}{x}\right)^{n/2-1} J_{n/2-1}(x).$$

Substituting this result into the expression for $\Phi_n(\cdot)$ with $x = ru$ and thence into $\phi(u)$ gives the desired result. \square

Of interest also is an expression for the moments of an EP distribution in the one-dimensional case, $Y \sim EP_1(\mu, \sigma^2, \beta)$.

Proposition 2.2. If $Y \sim EP_1(\mu, \sigma^2, \beta)$, then the n -th order moment is

$$\mathcal{E}(Y^n) = \sum_{k=0}^n \binom{n}{k} \mu^{n-k} \sigma^k \frac{2^{k/(2\beta)} \Gamma\left(\frac{k+1}{2\beta}\right)}{2\Gamma\left(\frac{1}{2\beta}\right)} [1 + (-1)^k].$$

Proof. Nadarajah (2005) computes the n -th moment of a *generalized normal distribution*, which is equivalent to an EP_1 with a slightly different shape parametrization. A consequence of this computation is that we obtain the following result for any $s > 0$:

$$\int_{-\infty}^\infty y^k \exp\{-|y|^s\} dy = \frac{1 + (-1)^k}{s} \Gamma\left(\frac{k+1}{s}\right). \tag{9}$$

Letting $X \sim EP_1(0, 1, \beta)$, a standardized EP_1 , we simplify notation by setting $\alpha = 2\beta$, and compute the k -th order moment as

$$\mathcal{E}(X^k) = \frac{\alpha}{2^{1+1/\alpha} \Gamma\left(\frac{1}{\alpha}\right)} \int_{-\infty}^\infty x^k \exp\left\{-\frac{1}{2}|x|^\alpha\right\} dx.$$

Substituting $x = 2^{1/\alpha}y$, implies $|x|^\alpha = 2|y|^\alpha$, and we obtain

$$\mathcal{E}(X^k) = \frac{2^{k/\alpha+1/\alpha} \alpha}{2^{1+1/\alpha} \Gamma\left(\frac{1}{\alpha}\right)} \int_{-\infty}^\infty y^k \exp\{-|y|^\alpha\} dy = \frac{2^{k/\alpha} \Gamma\left(\frac{k+1}{\alpha}\right)}{2\Gamma\left(\frac{1}{\alpha}\right)} [1 + (-1)^k],$$

which follows by substituting the expression for the integral from (9). Finally, the desired result is obtained by invoking the binomial formula:

$$\mathcal{E}(Y^n) = \mathcal{E}[(\mu + \sigma X)^n] = \sum_{k=0}^n \binom{n}{k} \mu^{n-k} \sigma^k \mathcal{E}(X^k).$$

\square

In particular, we obtain $\mathcal{E}(Y) = \mu$ and $\text{Var}(Y)$ as in (6) with $n = 1$ and $\Sigma = \sigma^2$.

We end this section by stating an ancillary result involving an integral that will be needed in a subsequent section in order to derive the asymptotic covariance matrix of MLEs. This parallels the corresponding result obtained by Nadarajah (2005) in the course of deriving the Fisher information matrix.

Lemma 2.3. For $\alpha > 0$ and integer k :

$$\begin{aligned}
 A(\alpha, k) &\equiv \int_0^\infty x^\alpha (\log x)^k e^{-x^\alpha/2} dx \\
 &= \frac{2^{1+1/\alpha}}{\alpha^{k+1}} \Gamma(1 + 1/\alpha) \begin{cases} 1, & k = 0, \\ [\log 2 + \Psi(1 + 1/\alpha)], & k = 1, \\ \{[\log 2 + \Psi(1 + 1/\alpha)]^2 + \Psi'(1 + 1/\alpha)\}, & k = 2. \end{cases}
 \end{aligned}$$

where $\Psi(z) = \partial \log \Gamma(z) / \partial z$ and $\Psi'(z) = \partial \Psi(z) / \partial z$ are respectively the digamma and trigamma functions.

Proof. See the Appendix. \square

3. Marginals of ARMA Models Under EP Distributions

This section considers what marginals arise from a joint EP distributional specification, and from a process driven by EP noise.

3.1. ARMA marginals for a Laplace process

We first derive the marginal distributions of $\mathbf{X}_n = (X_1, \dots, X_n)' \sim EP_n(\mathbf{0}, \Sigma, \beta = 1/2)$, a zero mean stationary ARMA(p, q) with covariance function $\Gamma_n = 4(n + 1)\Sigma$, whose joint distribution is Laplace.

Proposition 3.1. Let $\mathbf{X}_n = (X_1, \dots, X_n)'$ be a zero mean stationary ARMA(p, q) with covariance function $\Gamma_n = [\text{Cov}(X_i, X_j)]_{i,j=1}^n \equiv [\gamma_{|i-j|}]_{i,j=1}^n$ whose joint distribution follows a Laplace random vector (EP with $\beta = 1/2$). Then, the marginal densities of X_1 and $\mathbf{X}_2 = (X_1, X_2)'$ are given respectively by:

$$\begin{aligned}
 f_{X_1}(x_1) &= \frac{\sqrt{\frac{n+1}{\pi\gamma_0}}}{\Gamma(\frac{n+1}{2})2^n} z^{n/4} K_{n/2}(\sqrt{z}/2), & z &= \frac{4(n+1)x_1^2}{\gamma_0}, \\
 f_{\mathbf{X}_2}(\mathbf{x}_2) &= \frac{(n+1)|\Gamma_2|^{-1/2}}{2^n \pi \Gamma(\frac{n+1}{2})} z^{(n-1)/4} K_{(n-1)/2}(\sqrt{z}/2), & z &= 4(n+1)\mathbf{x}_2' \Gamma_2^{-1} \mathbf{x}_2,
 \end{aligned}$$

where $K_m(\cdot)$ denotes a modified Bessel function of the 2nd kind of order m .

Proof. From (iii) of Proposition 3.2 of Gomez et al. (1998),

$$\mathbf{X}_n \sim EP_n(\boldsymbol{\mu} = \mathbf{0}, \Sigma = [4(n + 1)]^{-1}\Gamma_n, \beta = 1/2),$$

whence a straightforward application of Proposition 5.1 of Gomez et al. (1998) gives the following elliptically contoured marginals

$$\mathbf{X}_1 \sim EC_n(\boldsymbol{\mu} = 0, \Sigma = \gamma_0/(4(n + 1)), h_1), \quad \mathbf{X}_2 \sim EC_n(\boldsymbol{\mu} = 0, \Sigma = \Gamma_2/(4(n + 1)), h_2),$$

with respective generator functions

$$\begin{aligned}
 h_1(z) &= z^{(n-1)/2} \int_0^1 x^{-(n+1)/2} (1-x)^{(n-3)/2} \exp\left\{-\sqrt{\frac{z}{4x}}\right\} dx \\
 &= \frac{\Gamma(\frac{n-1}{2})2^{n-1}}{\sqrt{\pi}} z^{n/4} K_{n/2}(\sqrt{z}/2)
 \end{aligned} \tag{10}$$

$$\begin{aligned}
 h_2(z) &= z^{n/2-1} \int_0^1 x^{-n/2} (1-x)^{n/2-2} \exp\left\{-\sqrt{\frac{z}{4x}}\right\} dx \\
 &= \frac{\Gamma(\frac{n}{2}-1)2^{n-2}}{\sqrt{\pi}} z^{(n-1)/4} K_{\frac{n-1}{2}}(\sqrt{z}/2)
 \end{aligned} \tag{11}$$

The proofs of (10) and (11) involve elementary Bessel and Gamma function manipulations, and are provided in the Appendix. It remains to compute the correct norming constants, c_1 and c_2 , such that $c_1 h_1(z)$ and $c_2 h_2(z)$ are EC_1 and EC_2 pdf's upon multiplication by the appropriate values of $|\Sigma|^{-1/2}$. These norming constants are obtained from (3) with $n = 1$ and $n = 2$, respectively. For the first case, we obtain

$$\frac{1}{c_1} = \iota_1 = \frac{\Gamma(\frac{n-1}{2})2^{n-1}}{\sqrt{\pi}} \int_0^\infty z^{n/4-1/2} K_{n/2}(\sqrt{z}/2) dz = 2^{2n} \Gamma\left(\frac{n+1}{2}\right) \Gamma\left(\frac{n-1}{2}\right),$$

which follows from standard results on integrals of Bessel functions, e.g. equation (A.0.13) in Kotz *et al.* (2001, Appendix). Since $\Sigma = \gamma_0/(4(n+1))$, $|\Sigma|^{-1/2} = 2\sqrt{(n+1)}/\gamma_0$, we thus obtain the given expression for the pdf of X_1 : $f_{X_1}(x_1) = |\Sigma|^{-1/2} c_1 h_1(z)$ with $z = 4(n+1)x_1^2/\gamma_0$.

For the second case, we obtain similarly

$$\frac{1}{c_2} = \pi \iota_2 = \pi \frac{\Gamma(\frac{n}{2}-1)2^{n-2}}{\sqrt{\pi}} \int_0^\infty z^{(n-1)/4} K_{(n-1)/2}(\sqrt{z}/2) dz = \pi \frac{\Gamma(\frac{n}{2}-1)2^{n-2}}{\sqrt{\pi}} 2^{2+n} \Gamma\left(\frac{n+1}{2}\right),$$

whence the desired result follows by noting that since $\Sigma = \Gamma_2/(4(n+1))$, $|\Sigma|^{-1/2} = 4(n+1)|\Gamma_2|^{-1/2}$, and the pdf of \mathbf{X}_2 is: $f_{\mathbf{X}_2}(\mathbf{x}_2) = |\Sigma|^{-1/2} c_2 h_2(z)$ with $z = 4(n+1)\mathbf{x}_2' \Gamma_2^{-1} \mathbf{x}_2$. \square

3.2. Marginals of an MA(1) process driven by EP₁ noise

We now focus on the marginal of an MA(1) process driven by iid EP₁ noise. Specifically, we consider the distribution of $X = \theta Z_1 + Z_2$, where Z_1, Z_2 are iid EP₁(0, σ^2, β) with common pdf

$$f_Z(z|\sigma, \beta) = \frac{1}{c_\beta \sigma} \exp\left\{-\frac{1}{2} \left|\frac{z}{\sigma}\right|^{2\beta}\right\}, \quad c_\beta = 2^{1+\frac{1}{2\beta}} \Gamma\left(1 + \frac{1}{2\beta}\right). \tag{12}$$

The pdf of X then corresponds to the density of an MA(1), and would be given by the convolution formula:

$$f_X(x) = \int_{-\infty}^\infty \frac{1}{|\theta|\sigma} f_Z(y) f_Z\left(\frac{x/\sigma - y}{\theta}\right) dy = \frac{1}{|\theta|c_\beta^2 \sigma} \int_{-\infty}^\infty \exp\left\{-\frac{1}{2} \left[\left|\frac{x/\sigma - y}{\theta}\right|^{2\beta} + |y|^{2\beta}\right]\right\} dy. \tag{13}$$

A general solution of this integral in closed-form is a difficult problem. (A representation of the integral as an infinite double summation involving pochhammer and Gamma functions is available from the authors upon request.) However, for two special cases of β , the solution is readily obtained.

Normal. In the case $\beta = 1$ we obtain (as we must) that $X \sim N(0, (1 + \theta^2)\sigma^2)$.

Laplace. The case $\beta = 1/2$ is solved by first noting that the standardized EP₁ with pdf $f_Z(z) = \frac{1}{4} \exp\{-|z|/2\}$ is related to the standard Laplace with pdf $f_W(w) = \frac{1}{\sqrt{2}} \exp\{-\sqrt{2}|w|\}$ via the simple scale transform: $Z = 2\sqrt{2}W$. Scaling Z by σ then means that the latter filters through as a scaling parameter for X , and hence the pdf of X is given by Corollary 1 of Trindade *et al.* (2010) with skewness parameter $\kappa = 1$ and scale parameter $\tau = 2\sqrt{2}\sigma$.

4. Asymptotics for ARMA MLEs and QMLEs Under EP Likelihoods

We consider the asymptotic regime for the MLEs and QMLEs of the vector of coefficients $\boldsymbol{\alpha} = (\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)'$ in the causal and invertible mean zero ARMA(p, q), where the AR and MA polynomials $\phi(z) = 1 - z - \dots - z^p$ and $\theta(z) = 1 + z + \dots + z^q$, respectively, have no common zeros

$$\phi(B)X_t = \theta(B)Z_t, \quad Z_t \sim \text{iid}(0, \sigma^2), \tag{14}$$

and σ is a scale parameter for the distribution of Z_t . Let $\boldsymbol{\eta} = (\boldsymbol{\alpha}', \beta, \sigma^2)'$ denote the entire vector of unknown parameters, and $\mathbf{x}_n = (x_1, \dots, x_n)'$ the vector of observations.

4.1. Asymptotics under a full EP likelihood

The MLEs in this case are obtained by maximizing a joint EP_n likelihood for the data vector, $\mathbf{X}_n = (X_1, \dots, X_n)' \sim EP_n(\mathbf{0}, \Sigma_n, \beta)$, a zero mean ARMA(p, q) with covariance function $\Gamma_n \equiv \sigma^2 \Omega_{n,\alpha} = c_{n,\beta} \Sigma_n$, where $c_{n,\beta}$ is as defined in (6), $\Omega_{n,\alpha}$ depends only on n and α , and σ^2 is the variance of Z_t . The corresponding log-likelihood for $\boldsymbol{\eta}$ is given by

$$\ell(\boldsymbol{\eta}) = \log(k_{n,\beta}) - \frac{1}{2} \log |\Sigma_n| - \frac{1}{2} (\mathbf{x}'_n \Sigma_n^{-1} \mathbf{x}_n)^\beta, \quad \text{where} \quad k_{n,\beta} = \frac{n\Gamma(n/2)}{\pi^{n/2} \Gamma(1 + \frac{n}{2\beta}) 2^{1 + \frac{n}{2\beta}}}. \quad (15)$$

Substitution of $\Sigma_n = \sigma^2 \Omega_{n,\alpha} / c_{n,\beta}$ then leads to the alternate version of the log-likelihood with terms that allow us to see the dependence on each of the unknown parameters,

$$\ell(\boldsymbol{\alpha}, \beta, \sigma^2) = \log(k_{n,\beta} c_{n,\beta}^{n/2}) - \frac{n}{2} \log(\sigma^2) - \frac{1}{2} \log |\Omega_{n,\alpha}| - \frac{1}{2} \left(\frac{c_{n,\beta}}{\sigma^2} \right)^\beta (\mathbf{x}'_n \Omega_{n,\alpha}^{-1} \mathbf{x}_n)^\beta. \quad (16)$$

Computing $\partial \ell / \partial \sigma^2$ and setting equal to zero gives the profile MLE for σ^2

$$\hat{\sigma}^2 = c_{n,\beta} \left(\frac{\beta}{n} \right)^{1/\beta} Q_{n,\alpha}, \quad \text{where} \quad Q_{n,\alpha} = (\mathbf{x}'_n \Omega_{n,\alpha}^{-1} \mathbf{x}_n), \quad (17)$$

which upon substitution into (16) gives the profile log-likelihood

$$\ell(\boldsymbol{\alpha}, \beta, \hat{\sigma}^2) = \frac{n}{2\beta} \left[\log \left(\frac{nk_{n,\beta}^{2\beta/n}}{\beta} \right) - 1 \right] - \frac{1}{2} \log |\Omega_{n,\alpha}| - \frac{n}{2} \log(Q_{n,\alpha}) \equiv \ell_1(\beta) + \ell_2(\boldsymbol{\alpha}). \quad (18)$$

The separability of this log-likelihood in the parameters β and $\boldsymbol{\alpha}$, means that the MLEs of the latter are obtained by maximizing

$$\ell_2(\boldsymbol{\alpha}) = -\frac{1}{2} \log |\Omega_{n,\alpha}| - \frac{n}{2} \log(Q_{n,\alpha}),$$

which coincides with the objective function for the usual Gaussian likelihood. Hence, the asymptotics of the resulting MLEs ($\hat{\boldsymbol{\alpha}}, \hat{\sigma}^2$) are identical to that case, e.g. Brockwell and Davis (1991, §8.8). Note that $\ell_1(\beta)$ is monotone increasing in β , and so it does not appear that (useful) estimation of β is possible under this paradigm. On the other hand, this finding sheds new light on the importance of maximizing a Gaussian likelihood, in particular of its robustness with respect to an EP criterion.

4.2. Asymptotics under a driving EP noise distribution

In this case the driving iid noise sequence $Z_t \sim EP_1(0, \sigma^2, \beta)$ is assumed to follow a univariate $EP_1 \sim f_Z(z|\sigma, \beta)$ pdf as in (12), and we maximize the conditional log-likelihood

$$\ell(\boldsymbol{\alpha}, \beta, \sigma) = \sum_{t=p+1}^n \log f_Z(z_t|\sigma, \beta).$$

This expression can be profiled in σ similarly to the above case, to yield

$$\ell(\boldsymbol{\alpha}, \beta) = \ell(\boldsymbol{\alpha}, \beta, \hat{\sigma}) = -n \log c_\beta - \frac{n}{2\beta} \left[\log \left(\frac{\beta S_{\boldsymbol{\alpha}, \beta}}{n} \right) + 1 \right], \quad (19)$$

where

$$\hat{\sigma} = \left[\frac{\beta}{n} S_\beta(\boldsymbol{\alpha}) \right]^{1/(2\beta)}, \quad \text{and} \quad S_\beta(\boldsymbol{\alpha}) = \sum_{t=p+1}^n |z_t(\boldsymbol{\alpha})|^{2\beta}, \quad (20)$$

Using the ARMA equations (14) to obtain the values of $z_t(\boldsymbol{\alpha})$ recursively from the data vector \mathbf{x}_n , the $\ell(\boldsymbol{\alpha}, \beta)$ criterion can be numerically optimized to give $\tilde{\boldsymbol{\eta}} = (\tilde{\boldsymbol{\alpha}}', \tilde{\beta}, \tilde{\sigma})'$, the corresponding QMLEs. An application of the Theorem in Li and McLeod (1988), allows us to conclude that $\tilde{\boldsymbol{\eta}}$ is asymptotically normal with mean $\boldsymbol{\eta}$ and covariance matrix $n^{-1}\Omega^{-1}$, that is, with Ω in block form, we have

$$\sqrt{n}(\tilde{\boldsymbol{\eta}} - \boldsymbol{\eta}) \xrightarrow{d} N(\mathbf{0}, \Omega^{-1}), \quad \text{where} \quad \Omega = \begin{bmatrix} J_{p+q} & 0 \\ 0 & I_2 \end{bmatrix}. \tag{21}$$

I_2 is the 2×2 Fisher Information matrix (per observation) corresponding to n iid observations from the $EP_1(0, \sigma^2, \beta)$ pdf, whose elements are given by the following Theorem.

Theorem 4.1. *If $Z \sim EP_1(0, \sigma^2, \beta)$, then the elements of the Fisher Information matrix (per observation) are given by*

$$I_2(2, 2) = \frac{2^{1+\frac{1}{2\beta}}(2\beta+1)}{\sigma^2 c_\beta} \Gamma\left(1 + \frac{1}{2\beta}\right) - \frac{1}{\sigma^2}, \quad I_2(1, 2) = \frac{2^{1+\frac{1}{2\beta}} \Gamma\left(1 + \frac{1}{2\beta}\right)}{\sigma c_\beta \beta} \left[\Psi\left(1 + \frac{1}{2\beta}\right) + \log(2) - 1 \right],$$

and

$$I_2(1, 1) = \frac{2^{\frac{1}{2\beta}} \Gamma\left(1 + \frac{1}{2\beta}\right)}{c_\beta \beta^3} \left\{ \left[\log(2) + \Psi\left(1 + \frac{1}{2\beta}\right) \right]^2 + \Psi'\left(1 + \frac{1}{2\beta}\right) \right\} + \frac{1}{4\beta^4} \left\{ \Psi'\left(1 + \frac{1}{2\beta}\right) + 4\beta \left[\log(2) + \Psi\left(1 + \frac{1}{2\beta}\right) \right] \right\}.$$

Proof. Straightforward differentiation gives (Wickramasinghe, 2012):

$$\frac{\partial^2 \log f(z|\sigma, \beta)}{\partial \sigma^2} = \frac{1}{\sigma^2} - \frac{\beta(2\beta+1)}{\sigma^2} \left| \frac{z}{\sigma} \right|^{2\beta}, \quad \frac{\partial^2 \log f(z|\sigma, \beta)}{\partial \sigma \partial \beta} = -\frac{2\beta}{\sigma} \left| \frac{z}{\sigma} \right|^{2\beta} \log |z/\sigma| + \frac{1}{\sigma} \left| \frac{z}{\sigma} \right|^{2\beta},$$

and

$$\frac{\partial^2 \log f(z|\sigma, \beta)}{\partial \beta^2} = -\frac{\Psi'(1 + \frac{1}{2\beta}) + 4\beta[\log(2) + \Psi(1 + \frac{1}{2\beta})]}{4\beta^4} - 2 \left| \frac{z}{\sigma} \right|^{2\beta} (\log |z/\sigma|)^2.$$

Since $Z \sim EP_1(0, \sigma^2, \beta)$, it follows immediately that $X \equiv Z/\sigma \sim EP_1(0, 1, \beta)$, and thus to compute expectations in the above we need expressions for

$$\mathcal{E} \left[|X|^\alpha (\log |X|)^k \right], \quad \alpha = 2\beta > 0 \text{ and } k = 0, 1, 2. \tag{22}$$

Now, noticing that the integrand in (22) is an even function, we obtain

$$\mathcal{E} \left[|X|^\alpha (\log |X|)^k \right] = \int_{-\infty}^{\infty} \frac{|x|^\alpha}{c_\beta} (\log |x|)^k e^{-|x|^\alpha/2} dx = \frac{2}{c_\beta} \int_0^{\infty} x^\alpha (\log x)^k e^{-x^\alpha/2} dx = \frac{2}{c_\beta} A(\alpha, k),$$

where $A(\alpha, k)$ is as defined in Lemma 2.3, whence an application of the expressions therein yields the desired result for I_2 . \square

J_{p+q} has dimension $p+q$, its elements being functions of the driving noise pdf $f_Z(z|\boldsymbol{\lambda})$, parametrized by $\boldsymbol{\lambda}$, and the autocovariance of the process $\{X_t\}$. Let $\{U_t\}$ and $\{V_t\}$ be the auxiliary autoregressive processes $\phi(B)U_t = W_t$ and $\theta(B)V_t = W_t$, where $\{W_t\}$ is iid with mean zero and variance one. If $\mathbf{U} := (U_1, \dots, U_p)'$ and $\mathbf{V} := (V_1, \dots, V_q)'$, then define

$$\Upsilon_{p+q}(\boldsymbol{\alpha}) = \begin{bmatrix} \mathcal{E}(\mathbf{U}\mathbf{U}') & \mathcal{E}(\mathbf{U}\mathbf{V}') \\ \mathcal{E}(\mathbf{V}\mathbf{U}') & \mathcal{E}(\mathbf{V}\mathbf{V}') \end{bmatrix} \quad \text{and} \quad \varpi(\boldsymbol{\lambda}) = \int_{\mathbb{R}} \left(\frac{\partial f_Z(z|\boldsymbol{\lambda})}{\partial z} \right)^2 \frac{1}{f_Z(z|\boldsymbol{\lambda})} dz.$$

From the Theorem in Li and McLeod (1988), we then have that

$$J_{p+q} = \sigma^2 \Upsilon_{p+q}(\boldsymbol{\alpha}) \varpi(\boldsymbol{\lambda}).$$

Remark 4.2. If $\{Z_t\} \sim N(0, \sigma^2)$, a straightforward computation gives $\varpi(\sigma) = 1/\sigma^2$, so that $J_{p+q} = \Upsilon_{p+q}(\boldsymbol{\alpha})$, which coincides with the classical result for Gaussian MLEs (Brockwell and Davis, 1991, §8.8).

Lemma 4.3. If $Z \sim EP_1(0, \sigma^2, \beta)$, then for $\beta > 1/4$, $\varpi(\sigma, \beta)$ is given by

$$\varpi(\sigma, \beta) = \int_{\mathbb{R}} \left(\frac{\partial f_Z(z|\sigma, \beta)}{\partial z} \right)^2 \frac{1}{f_Z(z|\sigma, \beta)} dz = \frac{\beta 2^{2-1/(2\beta)}}{\sigma^2 c_\beta} \Gamma\left(2 - \frac{1}{2\beta}\right).$$

Proof. Straightforward computations give the integrand as

$$\left(\frac{\partial f_Z(z|\sigma, \beta)}{\partial z} \right)^2 \frac{1}{f_Z(z|\sigma, \beta)} = \frac{\beta^2}{\sigma^{4\beta}} |z|^{4\beta-2} f_Z(z|\sigma, \beta),$$

whence we need to develop an expression for the k -th absolute moment, $\mathcal{E}|Z|^k$. Noticing that $|z|^k f_Z(z|\sigma, \beta)$ is an even function, this can be computed as

$$\mathcal{E}|Z|^k = \frac{2}{\sigma c_\beta} \int_0^\infty z^k \exp\left\{-\frac{1}{2}\left(\frac{z}{\sigma}\right)^{2\beta}\right\} dz,$$

which, by means of the substitution $x = \frac{1}{2}\left(\frac{z}{\sigma}\right)^{2\beta}$, can be written as

$$\mathcal{E}|Z|^k = \frac{2(2^{1/(2\beta)}\sigma)^{k+1}}{2\sigma\beta c_\beta} \int_0^\infty x^{\frac{k+1}{2\beta}-1} e^{-x} dx = \frac{2^{(k+1)/(2\beta)}\sigma^k}{\beta c_\beta} \Gamma\left(\frac{k+1}{2\beta}\right),$$

and is convergent provided $k > -1$. Evaluating this at $k = 4\beta - 2$, then yields the desired result. \square

We are now ready to state the main result in this section, which specifies the asymptotic distribution of the QMLEs under EP_1 white noise.

Theorem 4.4. Consider the MA(1) model (1) where the driving white noise sequence $\{Z_t\}$ has the EP_1 density given by (12) with $\beta > 1/4$. Then the vector of QMLEs, $\tilde{\boldsymbol{\eta}}$, has the asymptotic distribution given by (21) with the constituent matrices in Ω given by

$$J_{p+q} \equiv J_1 = \begin{pmatrix} \beta 2^{1-1/\beta} \\ 1 - \theta^2 \end{pmatrix} \frac{\Gamma(2 - \frac{1}{2\beta})}{\Gamma(1 + \frac{1}{2\beta})},$$

and I_2 as in Theorem 4.1.

Proof. In the case of an MA(1), $\Upsilon_{p+q}(\boldsymbol{\alpha})$, computed with $p = 0$ and $q = 1$, is just the variance of the process V_t , i.e. $(1 - \theta^2)^{-1}$. Hence, invoking Lemma 4.3, we obtain

$$J_{p+q} \equiv J_1 = \frac{\beta 2^{1-1/\beta}}{1 - \theta^2} \frac{\Gamma(2 - \frac{1}{2\beta})}{\Gamma(1 + \frac{1}{2\beta})}.$$

\square

5. Application: Modeling the TIAA Real Estate Account

In this section we consider the TIAA Real Estate Account³⁾ over the period 15 June 2004 to 31 December 2006; a total of 650 daily values for which returns were computed by taking differences of successive log values. Trindade *et al.* (2010) found an MA(1) driven by iid asymmetric Laplace noise to be a plausible

³⁾ www.tiaa-cref.org/

model for the returns, and obtained the corresponding QMLEs by maximizing the appropriate conditional likelihood. The estimates for this model are displayed in the first row of Table 2, where $0 < \kappa < \infty$ controls skewness, and standard errors for non-fixed parameters appear in parentheses. Repeating this, but using iid $EP_1(0, \sigma^2, \beta)$ noise as in subsection 4.2 instead, results in the estimates in the second row of Table 2. The standard errors in the latter were calculated by plugging the QMLEs into the formulas of Theorems 4.1 and 4.4. Noting that the $\hat{\beta} = 0.4625$ shape parameter in the latter is very close to 0.5, suggests one might approximate the EP noise model with a (symmetric, $\kappa = 1$) Laplace distribution, where the shape parameter could then be computed from the relationship discussed in subsection 3.2,

$$\tau = 2\sqrt{2}\sigma = 2\sqrt{2}(2.20\text{E-}04) = 6.22\text{E-}04,$$

obtained by plugging in the estimate $\hat{\sigma} = 2.20\text{E-}04^4$. Finally, we entertained the usual Gaussian QMLE model, iid $EP_1(0, \sigma^2, \beta = 1)$ noise, which is routinely obtained with a software package (e.g., `arima` function in R with the option “method=CSS”) since it corresponds to minimizing a least-squares criterion. The estimates of these last two models appear in the third and fourth rows of Table 2.

Table 2: QMLEs (and standard errors) for four MA(1) models fitted to the TIAA-CREF returns with different noise structures. (“n/a”: standard errors “not applicable”).

IID Noise Structure	Parameter QMLEs (Standard Errors)				AIC
	MA (θ)	Skewness (κ)	Scale (σ)	Shape (β)	
Asymmetric Laplace	0.0621 (2.38E-02)	0.664 (3.68E-02)	6.31E-04 (4.28E-05)	0.5 (fixed)	-7703.7
Exponential Power	0.1350 (8.487E-02)	1.0 (fixed)	2.20E-04 (9.000E-06)	0.4625 (7.852E-03)	-7603.5
Laplace (approx.)	0.1350 (n/a)	1.0 (fixed)	6.22E-04 (n/a)	0.5 (fixed)	-7580.7
Gaussian	0.1488 (3.52E-02)	1.0 (fixed)	8.202E-04 (2.277E-05)	1.0 (fixed)	-7377.8

The corresponding Akaike Information Criterion (AIC) values for each model are given in the rightmost column of Table 2. If these are taken as suggestive of model adequacy, then we would conclude that the asymmetric Laplace noise structure is more plausible than the competing EP model (which in turn fares better than the remaining two models). It is interesting to consider the trade-off between these two competing models with the same number of non-fixed parameters; the inclusion of a skewness parameter appears to be more important than that of a shape parameter in this particular case, but that is because the Laplace shape of $\beta = 0.5$ is already close to optimal, as evidenced by the EP fit. There is some confirmatory evidence for this finding, in that an inspection of the marginal histogram of returns suggests marked peakedness and right-skewness, consistent with the asymmetric Laplace distribution.

6. Summary

Distributional and inferential results were derived for ARMA models under the multivariate EP distribution. These were obtained by entertaining two different regimes: (i) a full multivariate EP distribution for the process, and (ii) a process driven by iid univariate EP noise. Under the first regime we provided closed-form expressions for the marginal distributions of ARMA processes in the Laplace case ($\beta = 1/2$),

⁴Of course, one would need to compute the appropriate QMLEs in this case, but for illustrative purposes only we retained the values obtained from the EP model, making this only an *approximate* symmetric Laplace model for which standard errors are not applicable.

and discovered that maximization of a full likelihood (general β) is identical to maximization of the usual Gaussian likelihood ($\beta = 1$) routinely implemented by default in software packages; a result which further strengthens the rationale for this default criterion. Under the second regime we were only able to provide closed-form expressions for the marginal distribution of the MA(1) process in the Laplace case (the Gaussian case being trivial), but presented a complete development of the asymptotics for a general ARMA and any β . The methodology was illustrated on a real dataset, where it was shown that allowing for a shape parameter (β) by using the EP to drive the process, provides a useful extension to the QMLE method.

References

- Anderson, D.N., Arnold, B.C. (1993) Linnik distributions and processes, *J. Appl. Prob.* 30, 330–340.
- Andrews, B., Calder, M., Davis, R.A. (2009) Maximum likelihood estimation for α -stable autoregressive processes, *Annals of Statistics*, 37, 1946–1982.
- Brockwell, P.J., Davis, R.A. (1991) *Time Series: Theory and Methods*, 2nd ed., Springer, New York.
- Damsleth, E., El-Shaarawi, A.H. (1989) ARMA models with double-exponentially distributed noise, *J. Royal Statist. Soc. B* 51, 61–69.
- Davis, R.A., Dunsmuir, W.T.M. (1996) Maximum Likelihood Estimation for MA(1) Processes with a Root on or near the Unit Circle, *Econometric Theory*, 12, 1–29.
- Davis, R.A., Dunsmuir, W.T.M. (1997) Least absolute deviation estimation for regression with ARMA errors, *Journal of Theoretical Probability*, 10, 481–497.
- Davidson, R., MacKinnon, J.G. (1993) *Estimation and Inference in Econometrics*, Oxford University Press, New York.
- Fang, K.T., Anderson, T.W. (Eds.) (1990) *Statistical Inference in Elliptically Contoured and Related Distributions*, Allerton Press: New York.
- Fang, K., Kotz, S. Ng, K.W. (1990) *Symmetric multivariate and related distributions*, Monographs on Statistics & Applied Probability, Chapman and Hall/CRC: New York.
- Fang, K.T., Zhang, Y.T. (1990) *Generalized Multivariate Analysis*, Springer-Verlag: Berlin.
- Gomez, E., Gomez-Villegas, M.A., Marin, J.M. (1998) A multivariate generalization of the power exponential family of distributions, *Communications in Statistics*, A27, 589–600.
- Gomez, E., Gomez-Villegas, M.A., Marin, J.M. (2002) Continuous elliptical and exponential power linear dynamic models, *Journal of Multivariate Analysis*, 83, 22–36.
- Gradshteyn, I.S., Ryzhik, I.M. (2000) *Table of Integrals, Series, and Products*, A. Jeffrey and D. Zwillinger (eds.), 6th ed., Academic Press: San Diego.
- Kotz, S., Kozubowski, T., Podgórski, K. (2001) *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance*, Birkhäuser: Boston.
- Li, W.K., McLeod, A.I. (1988) ARMA modelling with non-Gaussian innovations, *Journal of Time Series Analysis*, 9, 155–168.
- Mineo, A.M., Ruggieri, M. (2005) A Software Tool for the Exponential Power Distribution: The normalp package, *Journal of Statistical Software*, 12(4).
- Nadarajah, S. (2005) A generalized normal distribution, *Journal of Applied Statistics*, 32, 685–694.
- Paige, R.L., Trindade, A.A., Wickramasinghe, R.I.P. (2014) Extensions of Saddlepoint-Based Bootstrap Inference, *Annals of the Institute of Statistical Mathematics*, to appear.
- Provost, S.B., Cheong, Y-H. (2002) The Distribution of Hermitian Quadratic Forms in Elliptically Contoured Random Vectors, *Journal of Statistical Planning and Inference*, 102, 303–316.
- Trindade, A.A., Zhu, Y., Andrews, B. (2010) Time Series Models With Asymmetric Laplace Innovations, *Journal of Statistical Computation and Simulation*, 80, 1317–1333.
- Watson, G.N. (1944) *A Treatise on the Theory of Bessel Functions*, Cambridge University Press, Cambridge.
- Wickramasinghe, R.I.P. (2012) Extensions of Saddlepoint-Based Bootstrap Inference with Application to the Moving Average Model of Order One, *PhD Dissertation*, Texas Tech University.

7. Appendix

7.1. Proof of Lemma 2.3

In all cases we use the substitution $y = x^\alpha/2$, which implies $x = (2y)^{1/\alpha}$ and $dx = 2^{1/\alpha}y^{(1/\alpha)-1}/\alpha dy$.

Case $k = 0$:

$$\begin{aligned} \int_0^\infty x^\alpha e^{-x^\alpha/2} dx &= \int_0^\infty (2y)e^{-y}(1/\alpha)2^{1/\alpha}y^{1/\alpha-1} dy = \frac{2^{(1+1/\alpha)}}{\alpha} \int_0^\infty e^{-y}y^{(1+1/\alpha)-1} dy \\ &= \frac{2^{1+1/\alpha}}{\alpha} \Gamma(1 + 1/\alpha). \end{aligned}$$

Case $k = 1$:

$$\begin{aligned} \int_0^\infty x^\alpha (\log x) e^{-x^\alpha/2} dx &= \frac{2^{1/\alpha}}{\alpha^2} \int_0^\infty x^\alpha e^{-y} (2y) \frac{\log(2y)}{\alpha} y^{1/\alpha-1} dx \\ &= \frac{2^{1+1/\alpha}}{\alpha^2} \int_0^\infty x^\alpha e^{-y} y^{1/\alpha} \{\log(2) + \log(y)\} dy \\ &= \frac{2^{1+1/\alpha}}{\alpha^2} \left\{ \log(2) \Gamma(1 + 1/\alpha) + \int_0^\infty e^{-y} y^{(1/\alpha+1)-1} \log(y) dy \right\} \\ &= \frac{2^{1+1/\alpha}}{\alpha^2} \{ \log(2) \Gamma(1 + 1/\alpha) + \Gamma'(1 + 1/\alpha) \}, \end{aligned}$$

where the last line follows from equation (4.352) of Gradshteyn and Ryzhik (2000) with $\mu = 1$.

Case $k = 2$:

$$\begin{aligned} \int_0^\infty x^\alpha (\log x)^2 e^{-x^\alpha/2} dx &= \frac{2^{1/\alpha}}{\alpha^2} \int_0^\infty x^\alpha e^{-y} (2y) \left\{ \frac{\log(2y)}{\alpha} \right\}^2 y^{1/\alpha-1} dy \\ &= \frac{2^{1/\alpha}}{\alpha^3} \left\{ (\log 2)^2 \int_0^\infty e^{-y} y^{(1+1/\alpha)-1} dy + 2 \log 2 \int_0^\infty e^{-y} y^{(1+1/\alpha)-1} (\log y) dy \right\} \\ &\quad + \frac{2^{1/\alpha}}{\alpha^3} \int_0^\infty e^{-y} y^{(1+1/\alpha)-1} (\log y)^2 dy \\ &= \frac{2^{1/\alpha}}{\alpha^3} \{ (\log 2)^2 \Gamma(1 + 1/\alpha) + 2 \log 2 \Gamma'(1 + 1/\alpha) \} \\ &\quad + \frac{2^{1/\alpha}}{\alpha^3} \Gamma(1 + 1/\alpha) [\Psi(1 + 1/\alpha)^2 + \zeta(2, 1 + 1/\alpha)], \end{aligned}$$

where $\zeta(\cdot, \cdot)$ is the Generalized Riemann-Zeta function, and the last step follows by §4.358 of Gradshteyn and Ryzhik (2000). Finally, substituting the expression $\zeta(2, 1 + 1/\alpha) = \Psi'(1 + 1/\alpha)$ from §8.363, equation (8), of Gradshteyn and Ryzhik (2000), gives the desired result.

7.2. Proof of Equations (10) and (11)

By definition,

$$K_\nu(z) = \frac{\sqrt{\pi} (2z)^\nu e^{-z}}{\Gamma(\nu + \frac{1}{2})} \int_0^\infty e^{-2zt} t^{\nu-\frac{1}{2}} (1+t)^{\nu-\frac{1}{2}} dt, \tag{23}$$

so that upon replacing z with $\frac{\sqrt{z}}{2}$ and substituting $\nu = \frac{n}{2}$ gives

$$\begin{aligned} K_{\frac{n}{2}}\left(\frac{\sqrt{z}}{2}\right) &= \frac{\sqrt{\pi} (\sqrt{z})^{\frac{n}{2}} e^{-\frac{\sqrt{z}}{2}}}{\Gamma\left(\frac{n}{2} + \frac{1}{2}\right)} \int_0^\infty e^{-\sqrt{z}t} t^{\frac{n-1}{2}} (1+t)^{\frac{n-1}{2}} dt \\ &= \frac{\sqrt{\pi} z^{\frac{n}{4}}}{\Gamma\left(\frac{n-1}{2} + 1\right)} \int_0^\infty e^{-\sqrt{z}(t+\frac{1}{2})} t^{\frac{n-1}{2}} (1+t)^{\frac{n-1}{2}} dt \\ &= \frac{\sqrt{\pi} z^{\frac{n}{4}}}{\left(\frac{n-1}{2}\right) \Gamma\left(\frac{n-1}{2}\right)} \int_0^\infty e^{-\sqrt{z}(t+\frac{1}{2})} t^{\frac{n-1}{2}} (1+t)^{\frac{n-1}{2}} dt, \end{aligned}$$

using basic properties of the Gamma function. Thus,

$$\frac{\Gamma\left(\frac{n-1}{2}\right)}{\sqrt{\pi}} K_{\frac{n}{2}}\left(\frac{\sqrt{z}}{2}\right) = \frac{2z^{\frac{n}{4}}}{n-1} \int_0^\infty e^{-\sqrt{z}(t+\frac{1}{2})} t^{\frac{n-1}{2}} (1+t)^{\frac{n-1}{2}} dt.$$

Now, letting $t + \frac{1}{2} = \frac{1}{2\sqrt{x}}$ implies $t = \frac{1-\sqrt{x}}{2\sqrt{x}}$, $dt = -\frac{1}{4}x^{-\frac{3}{2}}dx$, and the limits become $t \rightarrow 0 \Rightarrow x \rightarrow 1$, and $t \rightarrow \infty \Rightarrow x \rightarrow 0$, which upon substitution into the previous step yields,

$$\begin{aligned} \frac{\Gamma\left(\frac{n-1}{2}\right)}{\sqrt{\pi}} K_{\frac{n}{2}}\left(\frac{\sqrt{z}}{2}\right) &= \frac{2^{-1}2^{1-n}z^{\frac{n}{4}}}{n-1} \int_0^1 e^{-\sqrt{\frac{z}{4x}}} \left(\frac{1-x}{4x}\right)^{\frac{n-1}{2}} \left(-\frac{1}{4}\right) x^{-\frac{3}{2}} dx \\ &= \frac{2^{-n}z^{\frac{n}{4}}}{n-1} \int_0^1 e^{-\sqrt{\frac{z}{4x}}} (1-x)^{\frac{n-1}{2}} (x)^{-\frac{n+2}{2}} dx. \end{aligned}$$

Now consider the derivative of $e^{-\sqrt{\frac{z}{4x}}}$.

$$\begin{aligned} \frac{d\left(e^{-\sqrt{\frac{z}{4x}}}\right)}{dx} &= e^{-\sqrt{\frac{z}{4x}}} \left(-\frac{1}{2}\right) \left(\frac{z}{4x}\right)^{-\frac{1}{2}} \left(-\frac{z}{4x^2}\right) = \frac{e^{-\sqrt{\frac{z}{4x}}} x^{-\frac{3}{2}} z^{\frac{1}{2}}}{2^2} \\ \implies e^{-\sqrt{\frac{z}{4x}}} &= 2^2 x^{\frac{3}{2}} z^{-\frac{1}{2}} \frac{d\left(e^{-\sqrt{\frac{z}{4x}}}\right)}{dx} \end{aligned}$$

Substitution of this result into the integrand of the previous equation gives,

$$\begin{aligned} \frac{\Gamma\left(\frac{n-1}{2}\right)}{\sqrt{\pi}} K_{\frac{n}{2}}\left(\frac{\sqrt{z}}{2}\right) &= \frac{2^{-1}2^{1-n}z^{\frac{n}{4}}}{n-1} \int_0^1 \frac{d\left(e^{-\sqrt{\frac{z}{4x}}}\right)}{dx} 2^2 x^{\frac{3}{2}} z^{-\frac{1}{2}} (1-x)^{\frac{n-1}{2}} (x)^{-\frac{n+2}{2}} dx \\ &= \frac{2^{-n+2}z^{\frac{n-2}{4}}}{n-1} \int_0^1 \frac{d\left(e^{-\sqrt{\frac{z}{4x}}}\right)}{dx} \left(\frac{1-x}{x}\right)^{\frac{n-1}{2}} dx \\ &= \frac{2^{-n+2}z^{\frac{n-2}{4}}}{n-1} \int_0^1 e^{-\sqrt{\frac{z}{4x}}} \left(\frac{n-1}{2}\right) \left(\frac{1-x}{x}\right)^{\frac{n-3}{2}} x^{-2} dx. \end{aligned}$$

A final simplifying step then establishes the required result of (10):

$$\begin{aligned} \frac{\Gamma\left(\frac{n-1}{2}\right)2^{n-1}z^{\frac{n}{4}}}{\sqrt{\pi}} K_{\frac{n}{2}}\left(\frac{\sqrt{z}}{2}\right) &= z^{\frac{n}{4}+\frac{n-2}{4}} \int_0^1 e^{-\sqrt{\frac{z}{4x}}} (1-x)^{\frac{n-3}{2}} (x)^{-\frac{n+1}{2}} dx \\ &= z^{\frac{n-1}{2}} \int_0^1 e^{-\sqrt{\frac{z}{4x}}} (1-x)^{\frac{n-3}{2}} (x)^{-\frac{n+1}{2}} dx. \end{aligned}$$

As for the proof of (11), replacing z with $\frac{\sqrt{z}}{2}$ and substituting $\nu = \frac{n-1}{2}$ into (23) gives,

$$K_{\frac{n-1}{2}}\left(\frac{\sqrt{z}}{2}\right) = \frac{\sqrt{\pi}(\sqrt{z})^{\frac{n-1}{2}}}{\Gamma\left[\left(\frac{n}{2}-1\right)+1\right]} \int_0^\infty e^{-\sqrt{z}(t+\frac{1}{2})} (t)^{\frac{n-2}{2}} (1+t)^{\frac{n-1}{2}-\frac{1}{2}} dt.$$

Following steps very similar to the proof of (10), we obtain the desired result,

$$z^{n/2-1} \int_0^1 x^{-n/2}(1-x)^{n/2-2} \exp\left\{-\sqrt{\frac{z}{4x}}\right\} dx = \frac{\Gamma\left(\frac{n}{2}-1\right)2^{n-2}}{\sqrt{\pi}} z^{(n-1)/4} K_{\frac{n-1}{2}}(\sqrt{z}/2).$$