Bayesian semiparametric mixed-effects joint models for analysis of longitudinal-competing risks data with skew distribution

Tao Lu

Appendix A. Multivariate Skew Distributions

Different versions of multivariate skew distributions have been introduced in the literature (Sahu *et al.* , 2003; Azzalini & Capitanio, 2003; Azzalini & Genton, 2008; Jara *et al.* , 2008). A new class of distributions by introducing skewness in multivariate elliptically distributions were developed in publication (Sahu *et al.* , 2003). The class, which is obtained by using transformation and conditioning, contains many standard families including the multivariate skew-normal (SN) and skew-t (ST) distributions as special cases. A k-dimensional random vector $\boldsymbol{Y}$ follows a k-variate skew-elliptical (SE) distribution if its probability density function (pdf) is given by

$$f(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta}; m_\nu^{(k)}) = 2^k f(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{A}; m_\nu^{(k)}) P(\boldsymbol{V} > \boldsymbol{0}), \tag{A.1}$$

where $\boldsymbol{A} = \boldsymbol{\Sigma} + \boldsymbol{\Delta}^2$, $\boldsymbol{\mu}$ is a location parameter vector, $\boldsymbol{\Sigma}$ is a $k \times k$ positive (diagonal) covariance matrix, $\boldsymbol{\Delta} = \text{diag}(\delta_1, \delta_2, \ldots, \delta_k)$ is a $k \times k$ skewness matrix with the skewness parameter vector $\boldsymbol{\delta} = (\delta_1, \delta_2, \ldots, \delta_k)^T$; $\boldsymbol{V}$ follows the elliptical distribution $El(\boldsymbol{\Delta} A^{-1}(\boldsymbol{y} - \boldsymbol{\mu}), \boldsymbol{I}_k - \boldsymbol{\Delta} A^{-1} \boldsymbol{\Delta}; m_\nu^{(k)})$ and the density generator function $m_\nu^{(k)}(\zeta) = \frac{\Gamma(k/2)}{\pi^{k/2}} \frac{m_\nu(\zeta)}{\int_0^\infty r^{k/2-1} m_\nu(\zeta) dr}$, with $m_\nu(\zeta)$ being a function such that $\int_0^\infty r^{k/2-1} m_\nu(\zeta) dr$ exists. The function $m_\nu(\zeta)$ provides the kernel of the original elliptical density and may depend on the parameter ν. This SE distribution is denoted by $SE(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta}; m^{(k)})$. Two examples of $m_\nu(\zeta)$, leading to important special cases used throughout the paper, are $m_\nu(\zeta) = \exp(-\zeta/2)$ and $m_\nu(\zeta) = (1 + \zeta/\nu)^{-(\nu+k)/2}$, where $\nu > 0$. These two expressions lead to the multivariate SN and ST distributions, respectively. In the latter case, ν corresponds to the degrees of freedom parameter.

As we know, a normal distribution is a special case of an SN distribution when the skewness parameter is zero, and the ST distribution reduces to the SN distribution when degrees of freedom are large. For completeness, this Appendix briefly summarizes the multivariate ST distribution introduced by (Sahu *et al.* , 2003) to be suitable for a Bayesian inference since it is built using the conditional method. For detailed discussions on properties of ST distribution, see Reference (Sahu *et al.* , 2003). Assume a k-dimensional random vector $\boldsymbol{Y}$ follows a k variate ST distribution with location vector $\boldsymbol{\mu}$, $k \times k$ positive (diagonal) covariance matrix $\boldsymbol{\Sigma}$ and $k \times k$ skewness matrix $\boldsymbol{\Delta} = \text{diag}(\delta_1, \delta_2, \ldots, \delta_k)$ or the degrees of freedom ν.

A k-dimensional random vector $\boldsymbol{Y}$ follows an m-variate ST distribution if its probability density function (pdf) is given by

$$f(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta}, \nu) = 2^k t_{k,\nu}(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{A}) P(\boldsymbol{V} > \boldsymbol{0}), \tag{A.2}$$

we denote the k-variate t distribution with parameters $\boldsymbol{\mu}$, $\boldsymbol{A}$ and degrees of freedom ν by $t_{k,\nu}(\boldsymbol{\mu}, \boldsymbol{A})$ and the corresponding pdf by $t_{k,\nu}(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{A})$ henceforth, $\boldsymbol{V}$ follows the t distribution $t_{k,\nu+k}$. We denote this distribution by $ST_{k,\nu}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta})$. In particular, when $\boldsymbol{\Sigma} = \sigma^2\boldsymbol{I}_k$ and $\boldsymbol{\Delta} = \delta\boldsymbol{I}_k$, the equation (A.2) simplifies to

$$f(\boldsymbol{y}|\boldsymbol{\mu}, \sigma^2, \delta, \nu) = 2^k(\sigma^2 + \delta^2)^{-k/2}\frac{\Gamma((\nu+k)/2)}{\Gamma(\nu/2)(\nu\pi)^{k/2}}\left\{1 + \frac{(\boldsymbol{y}-\boldsymbol{\mu})^T(\boldsymbol{y}-\boldsymbol{\mu})}{\nu(\sigma^2+\delta^2)}\right\}^{-(\nu+k)/2}$$
$$\times T_{k,\nu+k}\left[\left\{\frac{\nu+(\sigma^2+\delta^2)^{-1}(\boldsymbol{y}-\boldsymbol{\mu})^T(\boldsymbol{y}-\boldsymbol{\mu})}{\nu+k}\right\}^{-1/2}\frac{\delta(\boldsymbol{y}-\boldsymbol{\mu})}{\sigma\sqrt{\sigma^2+\delta^2}}\right],$$

where $T_{k,\nu+k}(\cdot)$ denotes the cumulative distribution function (cdf) of $t_{k,\nu+k}(\boldsymbol{0}, \boldsymbol{I}_k)$. However, unlike in the SN distribution, the ST density can not be written as the product of univariate ST densities. Here $\boldsymbol{Y}$ is dependent but uncorrelated. It is noted that when $\boldsymbol{\delta} = \boldsymbol{0}$, the ST distribution reduces to usual the t-distribution. It can be shown that the mean and covariance matrix of the ST distribution $ST_{k,\nu}(\boldsymbol{\mu}, \sigma^2\boldsymbol{I}_k, \boldsymbol{\Delta})$ are given by

$$\begin{aligned} E(\boldsymbol{Y}) &= \boldsymbol{\mu} + (\nu/\pi)^{1/2}\frac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)}\boldsymbol{\delta}, \\ \operatorname{cov}(\boldsymbol{Y}) &= \left[\sigma^2\boldsymbol{I}_k + \boldsymbol{\Delta}^2(\boldsymbol{\delta})\right]\frac{\nu}{\nu-2} - \frac{\nu}{\pi}\left[\frac{\Gamma\{(\nu-1)/2\}}{\Gamma(\nu/2)}\right]^2\boldsymbol{\Delta}^2(\boldsymbol{\delta}). \end{aligned} \tag{A.3}$$

In order to have a zero mean, we should assume the location parameter $\boldsymbol{\mu} = -(\nu/\pi)^{1/2}\frac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)}\boldsymbol{\delta}$. According to Lemma 1 of (Azzalini & Capitanio, 2003), if $\boldsymbol{Y}$ follows $ST_{k,\nu}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta})$, it can be represented by

$$\boldsymbol{Y} = \boldsymbol{\mu} + \zeta^{-1/2}\boldsymbol{X} \tag{A.4}$$

where ζ follows a Gamma distribution $\Gamma(\nu/2, \nu/2)$, which is independent of $\boldsymbol{X}$, and $\boldsymbol{X}$ follows a k-dimensional skew-normal (SN) distribution, denoted by $SN_k(\boldsymbol{0}, \boldsymbol{\Sigma}, \boldsymbol{\Delta})$. It follows from (A.4) that $\boldsymbol{Y}|\zeta \sim SN_k(\boldsymbol{\mu}, \zeta^{-1}\boldsymbol{\Sigma}, \zeta^{-1/2}\boldsymbol{\Delta})$. Following studies by (Azzalini & Genton, 2008), the SN distribution of $\boldsymbol{Y}$, conditional on ζ, has a convenient stochastic representation as

$$\boldsymbol{Y} = \boldsymbol{\mu} + \zeta^{-1/2}\boldsymbol{\Delta}|\boldsymbol{X}_0| + \zeta^{-1/2}\boldsymbol{\Sigma}^{1/2}\boldsymbol{X}, \tag{A.5}$$

where $\boldsymbol{X}_0$ and $\boldsymbol{X}$ are two independent $N_k(\boldsymbol{0}, \boldsymbol{I}_k)$ random vectors. Note that the expression (A.5) provides a convenience device for random number generation and for implementation purpose. Let $\boldsymbol{w} = \zeta^{-1/2}|\boldsymbol{X}_0|$; then $\boldsymbol{w}$, conditional on ζ, follows a k-dimensional normal distribution $N_k(\boldsymbol{0}, \zeta^{-1}\boldsymbol{I}_k)$ truncated in the space $\boldsymbol{w} > \boldsymbol{0}$ (i.e., the half-normal distribution). Thus, following (Jara *et al.* , 2008), a hierarchical representation of (A.5) is given by

$$\boldsymbol{Y}|\boldsymbol{w}, \zeta \sim N_k(\boldsymbol{\mu} + \boldsymbol{\Delta}\boldsymbol{w}, \zeta^{-1}\boldsymbol{\Sigma}),\ \boldsymbol{w}|\zeta \sim N_k(\boldsymbol{0}, \zeta^{-1}\boldsymbol{I}_k)\boldsymbol{I}(\boldsymbol{w} > \boldsymbol{0}),\ \zeta \sim \Gamma(\nu/2, \nu/2), \tag{A.6}$$

Note that the ST distribution presented in (A.6) can be reduced to the following three special cases: (i) as $\nu \to \infty$ and $\zeta \to 1$ with probability 1 (i.e., the last distributional specification is omitted), then the hierarchical expression (A.6) becomes an SN distribution $SN_k(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta})$; (ii) as $\boldsymbol{\Delta} = \boldsymbol{0}$, then the hierarchical expression (A.6) is a standard multivariate t-distribution; (iii) as $\nu \to \infty$, $\zeta \to 1$ with probability 1, and $\boldsymbol{\Delta} = \boldsymbol{0}$, then the hierarchical expression (A.6) reverts to a standard multivariate normal distribution.

Appendix B. Winbugs Code:

```
model
{
  ##### competing risk hazard at failure time #####
  ## ncW2R: number of columns of design matrix of the baseline
     hazard for AIDS deaths
  ## ncW2D: number of columns of design matrix of the baseline
     hazard for other deaths
  etaBaselineR[i] <- gamma * surv.t2[ i, 5 ]
  etaBaselineD[i] <- gamma2* surv.t2[ i, 5 ]
  log.h0.TR[i] <- inprod(Bs.gammasR[1:(ncW2R)], baseline.hazard.
     index1[i, 1:ncW2R])
  log.h0.TD[i] <- inprod(Bs.gammasD[1:(ncW2D)], baseline.hazard.
     index2[i, 1:ncW2D])
  log.hazardR[i] <- log.h0.TR[i] + etaBaselineR[i] + inprod(alphasR
     [1:nb], b[i,1:nb])
  log.hazardD[i] <- log.h0.TD[i] + etaBaselineD[i] + inprod(alphasD
     [1:nb], b[i,1:nb])

  for (k in 1:K) {
    log.h0.sR[i, k] <- inprod(Bs.gammasR[1:(ncW2R)], bl.haz.ind.
       quad1[K * (i - 1) + k, 1:ncW2R])
    log.h0.sD[i, k] <- inprod(Bs.gammasD[1:(ncW2D)], bl.haz.ind.
       quad2[K * (i - 1) + k, 1:ncW2D])
    SurvLongR[i, k] <- wk[k] * exp(log.h0.sR[i, k] + inprod(alphasR
       [1:nb], b[i,1:nb]) + gamma * hazard.cova[K * (i - 1) + k, 24
       ] )
    SurvLongD[i, k] <- wk[k] * exp(log.h0.sD[i, k] + inprod(alphasD
       [1:nb], b[i,1:nb]) + gamma2 * hazard.cova[K * (i - 1) + k, 24
        ] )
  }
  P[i] <- surv.t2[ i , 3 ] / 2
  zeros[i] <- 0
  log.survivalR[i] <- - P[i] * sum(SurvLongR[i,])
  log.survivalD[i] <- - P[i] * sum(SurvLongD[i,])
  phi[i] <- C - ((fail[i] * log.hazardR[i]) + (fail2[i] * log.
     hazardD[i])) - (log.survivalR[i] + log.survivalD[i])
  zeros[i] ~ dpois(phi[i])
}
  C <- 10000

  ### longitudinal model
  for (j in 1:N)
  {
    fix[ j ] <- y2[j, 27]  * beta[1] + y2[j, 28]  * beta[2] + y2[j,
       29]  * beta[3] + y2[j,24] * beta[4]
```

```
    ran[ j ] <- y2[j, 27]  * b[y2[j,23],1] + y2[j, 28]  * b[y2[j
       ,23],2] + y2[j, 29] * b[y2[j,23],3] +  y2[j,24] *  b[y2[j
       ,23],4]
    dm5[j] <- fix[ j ]+ ran[ j ]
    w[j]~dt(0,1,nu)I(0,)
    c1[j]<-(nu+w[j]*w[j])/n1
    aau[j]<-tau / c1[j]
    mu[j]<- dm5[j] +delta*(w[j]-mue)   #ST distribution
    y2[j,25]~dt(mu[j],aau[j],n1)
  }
  ### random effects
  for (i in 1:m)
  {
    b2[i,1]<-0
    b2[i,2]<-0
    b2[i,3]<-0
    b2[i,4]<-0
    b[i,1:4]~dmnorm(b2[i,1:4], Omega2[,])
  }

  ## Priors
  #Degree of freedom
  nu0<-0.5
  nu~ dexp(nu0)I(3,)
  n1<-nu+1
  mue<-exp(loggam(0.5*(nu-1))-loggam(0.5*nu))*sqrt(nu/3.14159)

  for (l in 1:4){beta[l]~dnorm(0,1.0E-2)}
  gamma~dnorm(0,1.0E-2)
  gamma2~dnorm(0,1.0E-2)
  for (l in 1:nb){alphasR[l]~dnorm(0,1.0E-2)}
  for (l in 1:nb){alphasD[l]~dnorm(0,1.0E-2)}
  for (l in 1:ncW2R){Bs.gammasR[l]~dgamma(0.1,0.1)}
  for (l in 1:ncW2D){Bs.gammasD[l]~dgamma(0.1,0.1)}

  delta ~ dnorm(0.0, 0.01)
  Omega2[1:4,1:4] ~ dwish(R2[,],4)
  V2[1:4,1:4]<-inverse(Omega2[,])
  tau ~ dgamma(.01,.01)
  sigma.tau <- 1/tau
}
```

References

Azzalini, A., & Capitanio, A. 2003. Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of Royal Statistical Society, Series B*, **65**, 367–389.

Azzalini, A., & Genton, M. 2008. Robust likelihood methods based on the skew-t and related distributions. *International Statistical Review*, **76**, 106–129.

Jara, A., Quintana, F., & Martin, E. 2008. Linear mixed models with skew-elliptical distributions: a Bayesian approach. *Computational Statistics and Data Analysis*, **52**, 5033–5045.

Sahu, S., Dey, D, & Branco, M. 2003. A new class of multivariate skew distributions with applications to Bayesian regression models. *The Canadian Journal of Statistics*, **31**, 129–150.