A New Class of Multivariate Prior Distributions with an Application to Reliability Engineering

UCA Universidad de Cádiz

GOBIERNO DE ESPAÑA — MINISTERIO DE ECONOMÍA, INDUSTRIA Y COMPETITIVIDAD

F. Ruggeri[1], M. Sánchez-Sánchez[2], M. A. Sordo[3] and A. Suárez-Llorens[4]

ISIPTA2019, 3-6 July (2019), Ghent (Belgium)

11th International Symposium on Imprecise Probabilities: Theories and Applications

1. Bayesian framework

- Let X be the underlying observation with probability density function (PDF) $f_{\theta}(x)$.
- $\boldsymbol{\theta}$ represents the unknown parameter defined in the set of states $\Theta \subseteq \mathbb{R}^n$, $n \in \mathbb{N}, n \geq 1$.
- Let $\boldsymbol{\pi}$ be the specific prior state of knowledge over Θ with PDF $\pi(\boldsymbol{\theta})$.
- Let $\boldsymbol{\pi}_{\mathbf{x}}$ be the posterior state of knowledge after observing data, $\mathbf{x}$, with PDF given by

$$\boldsymbol{\pi}_{\mathbf{x}}(\boldsymbol{\theta}) = \frac{l(\boldsymbol{\theta} \mid \mathbf{x})\boldsymbol{\pi}(\boldsymbol{\theta})}{m_{\boldsymbol{\pi}}(\mathbf{x})},$$

  where $l(\boldsymbol{\theta} \mid \mathbf{x})$ and $m_{\boldsymbol{\pi}}(\mathbf{x})$ denote the likelihood function and the marginal density, respectively.
- Objective: To make inference in some quantity of interest, $H_X(\boldsymbol{\theta})$, by using $\boldsymbol{\pi}_{\mathbf{x}}$.

2. The classical criticism

Why a unique prior? A Bayesian analysis is robust if it does not depend sensitively on the initial assumptions -Bayesian sensitivity-.

- A solution. To replace the specific prior distribution by a class of priors Γ.

Example: Given a specific prior π and ϵ in $(0, 1)$, the classical ϵ-contamination class is defined as

$$\Gamma_{\epsilon} = \{\boldsymbol{\pi}' : \boldsymbol{\pi}' = (1-\epsilon)\boldsymbol{\pi} + \epsilon \mathbf{Q},\ \mathbf{Q} \in \mathcal{Q}\},$$

where $\boldsymbol{\pi}'$ is given by a mixture and $\mathcal{Q}$ is a family of priors called the class of contaminations.

- Problem 1. The class Γ implies a class of posterior distributions Γ_x.
- Problem 2. The class Γ_x implies a range of the posterior quantity of interest.
- Problem 3. It is difficult in practice to compute that range. We consider a new class of priors.

3. The main results

- Key definition 1. Let $\mathbf{X}$ and $\mathbf{Y}$ be two absolutely continuous [discrete] n-dimensional random vectors with distribution functions $F_{\mathbf{X}}$ and $F_{\mathbf{Y}}$ and densities [discrete densities] $f_{\mathbf{X}}$ and $f_{\mathbf{Y}}$, respectively, such that

  $f(\mathbf{x})g(\mathbf{y}) \leq f(\mathbf{x} \wedge \mathbf{y})g(\mathbf{x} \vee \mathbf{y})$ for every $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$.

  Then $\mathbf{X}$ is said to be smaller than $\mathbf{Y}$ in the multivariate likelihood ratio order, denoted by $\mathbf{X} \leq_{lr} \mathbf{Y}$.
- Key definition 2. A function $l : \mathbb{R}^n \mapsto \mathbb{R}^+$, ($n \in \mathbb{N}, n \geq 2$) is said to be multivariate totally positive of order 2 (MTP_2), ($TP2$ when $n = 2$) if satisfies

$$l(\mathbf{x})l(\mathbf{y}) \leq l(\mathbf{x} \wedge \mathbf{y})l(\mathbf{x} \vee \mathbf{y}),\ \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

  Additionally, a n-dimensional random vector $\mathbf{X}$ with PDF f is said to be MTP_2 if its density f is MTP_2 or, equivalently, if $\mathbf{X} \leq_{lr} \mathbf{X}$.
- Key definition 3. Let $\boldsymbol{\theta}$ a random vector with $\pi(\boldsymbol{\theta})$ its density (probability) function. Let $w(\boldsymbol{\theta})$ be a non-negative weight function and assume that $E[w(\boldsymbol{\theta})]$ exists. A new density (probability) function will be denoted by

$$\boldsymbol{\pi}_w(\boldsymbol{\theta}) = \frac{w(\boldsymbol{\theta})\boldsymbol{\pi}(\boldsymbol{\theta})}{E[w(\boldsymbol{\theta})]}$$

- Key result 1. Let π be a specific prior belief which is MTP_2 in the variables. Let w be an increasing (decreasing) weight function. Then $\boldsymbol{\pi} \leq_{lr} (\geq_{lr})\boldsymbol{\pi}_w$.
- Class of priors. Let π be a specific MTP_2 prior belief. We will define the weighted band $\Gamma_{w_1,w_2,\boldsymbol{\pi}}$ associated with $\boldsymbol{\pi}$ based on w_1 and w_2, a decreasing weight function and an increasing weight function, respectively, (weighted band, for short), as

$$\Gamma_{w_1,w_2,\boldsymbol{\pi}} = \{\boldsymbol{\pi}' : \boldsymbol{\pi}_{w_1} \leq_{lr} \boldsymbol{\pi}' \leq_{lr} \boldsymbol{\pi}_{w_2}\}.$$

- Key result 2 Let π be a specific MTP_2 prior and let $\Gamma_{w_1,w_2,\boldsymbol{\pi}}$ be a weighted band associated with $\boldsymbol{\pi}$ based on w_1 and w_2. Given the observed data $\mathbf{x}$, if $l(\boldsymbol{\theta} \mid \mathbf{x})$ is MTP_2 in $\boldsymbol{\theta}$, then for all $\boldsymbol{\pi}' \in \Gamma_{w_1,w_2,\boldsymbol{\pi}}$ we obtain that $\boldsymbol{\pi}_{w_1,\mathbf{x}} \leq_{lr} \boldsymbol{\pi}'_{\mathbf{x}} \leq_{lr} \boldsymbol{\pi}_{w_2,\mathbf{x}}$.
- Key result 3. Let X be the underlying random variable and let H be a functional of interest such that $H_X(\boldsymbol{\theta})$ is non-decreasing in $\boldsymbol{\theta}$. Given $\boldsymbol{\pi}$, a specific MTP_2 prior, and the corresponding weighted band $\Gamma_{w_1,w_2,\boldsymbol{\pi}}$ based on w_1 and w_2, if $l(\boldsymbol{\theta} \mid \mathbf{x})$ is MTP_2 in $\boldsymbol{\theta}$, then the univariate random variables obtained by mapping the posterior distributions by the functional $H_X(\boldsymbol{\theta})$ satisfy

$$H_X(\boldsymbol{\pi}_{w_1,\mathbf{x}}) \leq_{st} H_X(\boldsymbol{\pi}'_{\mathbf{x}}) \leq_{st} H_X(\boldsymbol{\pi}_{w_2,\mathbf{x}}),$$

  $\forall \boldsymbol{\pi}' \in \Gamma_{w_1,w_2,\boldsymbol{\pi}}$.
  - Consequence 1. The predictive expectations are ordered:

$$E^{\boldsymbol{\pi}_{w_1,\mathbf{t}}}[H_X(\boldsymbol{\theta})] \leq E^{\boldsymbol{\pi}'_{\mathbf{t}}}[H_X(\boldsymbol{\theta})] \leq E^{\boldsymbol{\pi}_{w_2,\mathbf{t}}}[H_X(\boldsymbol{\theta})].$$

  - Consequence 2. Bayesian credible quantile-based intervals are ordered:

$$F^{-1}_{H_X(\boldsymbol{\pi}_{w_1,\mathbf{t}})}(p) \leq F^{-1}_{H_X(\boldsymbol{\pi}'_{\mathbf{t}})}(p) \leq F^{-1}_{H_X(\boldsymbol{\pi}_{w_2,\mathbf{t}})}(p).$$

4. A real example

Let's consider failure data from 40 underground trains associated with electrical opening commands. Failure monitoring started on 19th September 1991 ended on 31st December 1998. When a failure took place, both odometer reading and failure date were recorded. Finally, 530 failures were recorded.

- The model. A nonhomogeneous Poisson process (NHPP) with a Power Law intensity function:

$$\lambda(t|\boldsymbol{\theta}) = M\beta t^{\beta-1},\ \ \boldsymbol{\theta} = (M, \beta) \in \mathbb{R}^+ \times \mathbb{R}^+.$$

- The likelihood function. Let $\mathbf{t}^* = (T_1, \ldots, T_n)$ be the vector of failures times recorded in the interval $(0, T]$. By changing the scale, i.e., $\mathbf{t} = (T_1/T, \ldots, T_n/T)$. Then

$$l(\boldsymbol{\theta}|\mathbf{t}) = \prod_{i=1}^{n} \lambda(T_i) \cdot \exp(-m(T|\boldsymbol{\theta})),$$

$$= M^n \beta^n \exp((\beta - 1)\sum_{i=1}^{n} \ln(\frac{T_i}{T})) \exp(-M).$$

- The prior. The prior belief $\boldsymbol{\pi}$ over $\Theta = \mathbb{R}^+ \times \mathbb{R}^+$ is a bivariate random vector having independent exponential marginal distributions:

$$\pi(\boldsymbol{\theta}) = \lambda\mu \exp(-\lambda M)\exp(-\mu\beta),\ (M, \beta) \in \Theta,$$

  where the hyperparameters λ and μ are assumed to be known: from the initial values $M_0 = 495.5$ and $\beta_0 = 0.79$, we take the values $\lambda = 1/M_0$ and $\mu = 1/\beta_0$.
- The weighted functions.
  - $w_1(\boldsymbol{\theta}) = \theta_1^{a-1}\theta_2^{b-1}\exp[-c\theta_1\theta_2]$.
  - $w_2(\boldsymbol{\theta}) = \theta_1^{a'-1}\theta_2^{b'-1}(\theta_1^{c'} + \theta_2^{c'})$.
- The metrics. Degree of uncertainty.
  - Hellinger metric (H).

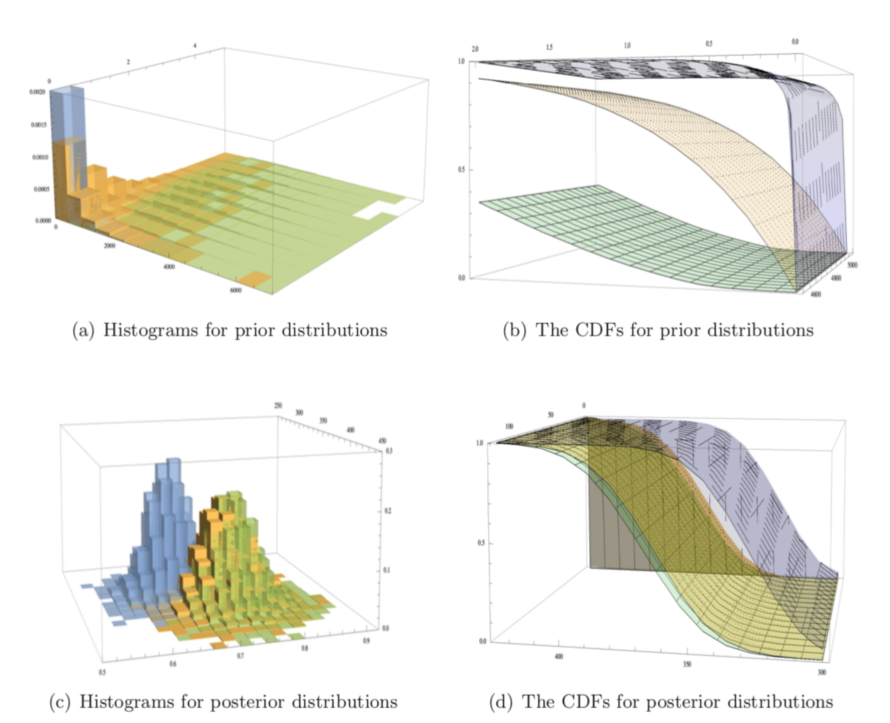
$$H(\mathbf{X}, \mathbf{Y}) = \frac{1}{2}\int (\sqrt{f_{\mathbf{X}}(\mathbf{x})} - \sqrt{f_{\mathbf{Y}}(\mathbf{x})})^2 d\mathbf{x}.$$

  - Kullback-Leibler divergence (KL).

$$KL(\mathbf{X}, \mathbf{Y}) = \int f_{\mathbf{X}}(\mathbf{x}) \ln \frac{f_{\mathbf{X}}(\mathbf{x})}{f_{\mathbf{Y}}(\mathbf{x})} d\mathbf{x}.$$

- Summary of the uncertainty. $a = 0.8$, $b = 0.4$, $c = 0.17$ and $a' = 3.8$, $b' = 3.4$, $c' = 1.17$.

Priors	$\boldsymbol{\pi}, \boldsymbol{\pi}_{w_1}$	$\boldsymbol{\pi}, \boldsymbol{\pi}_{w_2}$	$\boldsymbol{\pi}_{w_1}, \boldsymbol{\pi}_{w_2}$
H metric	0.6991	0.6959	0.9998
KL divergence	66.724	7.5520	27.101
Posteriors	$\boldsymbol{\pi}_{\mathbf{t}}, \boldsymbol{\pi}_{w_1,\mathbf{t}}$	$\boldsymbol{\pi}_{\mathbf{t}}, \boldsymbol{\pi}_{w_2,\mathbf{t}}$	$\boldsymbol{\pi}_{w_1,\mathbf{t}}, \boldsymbol{\pi}_{w,\mathbf{t}}$
H metric	0.7219	0.0071	0.7709
KL divergence	5.1661	0.0287	4.4334



(a) Histograms for prior distributions

(b) The CDFs for prior distributions

(c) Histograms for posterior distributions

(d) The CDFs for posterior distributions

Figure 2: The weighted prior distributions and the posterior weighted distributions.

- The functional of interest The expected number of failures in future time intervals $T_u = [T, T+u]$.

$$H_X(\boldsymbol{\theta}) = E[N(1+h) - N(1)] = M((1+h)^{\beta} - 1).$$

- Making inference. True Value vs Forecasts.

TIME		$\boldsymbol{\pi}_{w_1,\mathbf{t}}$		$\boldsymbol{\pi}_{\mathbf{t}}$		$\boldsymbol{\pi}_{w_2,\mathbf{t}}$	
T_u	TRUE	$IC_{95\%}$ CRED.	MEAN	$IC_{95\%}$ CRED.	MEAN	$IC_{95\%}$ CRED.	MEAN
92_1	83	[47.63, 78.53]	63	[66.25, 102]	84	[70.95, 107.8]	89
92_2	72	[42.81, 72.30]	57	[63.20, 98.12]	80	[68.45, 104.64]	86
92_3	62	[39.83, 68.37]	54	[61.25, 95.66]	78	[66.84, 102.62]	84
93_1	72	[35.90, 63.29]	49	[48.22, 79.32]	63	[50.67, 82.44]	66
93_2	62	[35.52, 58.77]	45	[45.28, 75.51]	60	[47.82, 78.78]	63
93_3	42	[30.15, 55.58]	42	[43.16, 72.78]	57	[45.76, 76.13]	60
94_1	62	[31.59, 57.54]	44	[42.01, 71.31]	56	[43.53, 73.29]	58
94_2	42	[29.12, 54.22]	41	[39.77, 68.38]	54	[41.35, 70.45]	55
94_3	35	[27.28, 51.68]	39	[38.05, 66.12]	52	[39.66, 68.24]	53
95_1	42	[30.84, 56.54]	43	[40.92, 69.91]	55	[42.10, 71.45]	56
95_2	35	[28.92, 53.94]	41	[39.21, 67.66]	53	[40.42, 69.26]	54
95_3	23	[27.40, 51.86]	39	[37.82, 62.84]	51	[39.08, 67.48]	53
96_1	35	[26.20, 50.21]	38	[34.49, 61.42]	47	[35.35, 62.58]	48
96_2	23	[24.65, 48.05]	36	[32.99, 59.43]	46	[33.88, 60.61]	47
97_1	23	[22.75, 45.40]	34	[29.73, 55.04]	42	[30.46, 56.03]	43

5. Concluding remarks

- Elicitation and interpretation are easy.
- Prior uncertainty reflected by metrics.
- Bounds for the range of quantities of interest.

Basic references

- Arias-Nicolás, Ruggeri and Suárez-Llorens (2016). *New classes of priors based on stochastic orders and distortion functions. Bayesian Analysis, 11 (4): 1107–1136.*.
- Ruggeri, Sánchez-Sánchez, Sordo and Suárez-Llorens (2019). *On a new class of multivariate prior distributions: theory and application in reliability. Submitted to Bayesian Analysis.*

emails: fabrizio@mi.imati.cnr.it[1], marta.sanchez@uca.es[2], mangel.sordo@uca.es[3], alfonso.suarez@uca.es[4]. 11th International Symposium on Imprecise Probabilities: Theories and Applications.