Reference Priors under Partial Invariance

Xiaoyan(Iris) Lin, University of South Carolina

O-Bayes June 7, 2009

Joint work with James Berger and Dongchu Sun

Outline

- Motivated Examples
- Invariant Models and Priors
- Semi-invariant Structures and Priors
- Simulations
- Comments

Motivated Example I: Weibull Distribution

Weibull Distribution,

$$f(x \mid \psi, \eta, \beta) = \frac{\beta(x-\psi)^{\beta-1}}{\eta^{\beta}} \exp\left\{ -\left(\frac{x-\psi}{\eta}\right)^{\beta} \right\}, \quad x > \psi, \quad (1)$$

is used in many areas:

- Reliability and Survival analysis;
- Weather forecasting, e.g., to describe wind speed distributions;
- General insurance, e.g., to model the size of reinsurance claims;
- Psychological research, e.g., to model the distribution of response time;
- Extreme value theory. (Extreme value program at SAMSI in Spring of 2008.)

Two Applications of the Three-parameter Weibull Distribution

1. It models the distribution of the response times in psychological research (Rouder et al. 2003). The time taken to complete a task in a psychology experiment is called the response time.

$$
\begin{aligned}
y_{ij} &\sim \textit{Weibull}(\psi_i, \beta_i, \eta_i), \;\; 1 \le i \le I, \; 1 \le j \le J_i, \\
(\beta_i \,|\, \alpha_1, \alpha_2) &\overset{iid}{\sim} \textit{Gamma}(\alpha_1, \alpha_2) \;\; \text{restricted to } \beta_i > .01, \\
(\eta_i^{-\beta_i} \,|\, \xi_1, \xi_2) &\overset{iid}{\sim} \textit{Gamma}(\xi_1, \xi_2), \\
\alpha_k &\sim \textit{Gamma}(a_k, b_k), \;\; k = 1, 2, \\
\xi_k &\sim \textit{Gamma}(c_k, d_k), \;\; k = 1, 2.
\end{aligned}
$$

Two Applications of the Three-parameter Weibull Distribution

1. It models the distribution of the response times in psychological research (Rouder et al. 2003). The time taken to complete a task in a psychology experiment is called the response time.

$$
\begin{aligned}
y_{ij} &\sim \mathit{Weibull}(\psi_i, \beta_i, \eta_i), \quad 1 \le i \le I, \ 1 \le j \le J_i, \\
(\beta_i \,|\, \alpha_1, \alpha_2) &\overset{iid}{\sim} \mathit{Gamma}(\alpha_1, \alpha_2) \ \text{ restricted to } \beta_i > .01, \\
(\eta_i^{-\beta_i} \,|\, \xi_1, \xi_2) &\overset{iid}{\sim} \mathit{Gamma}(\xi_1, \xi_2), \\
\alpha_k &\sim \mathit{Gamma}(a_k, b_k), \ k = 1, 2, \\
\xi_k &\sim \mathit{Gamma}(c_k, d_k), \ k = 1, 2.
\end{aligned}
$$

2. Pang et al. (2001) used the three-parameter Weibull distribution to model the wind speed distribution. They applied the constant prior to the three parameters. However, it can be shown easily that the corresponding posterior is improper.

What are good priors for the three-parameter Weibull?

What are good priors for the three-parameter Weibull?

- Jeffreys prior does not exist. Fisher Information of $\boldsymbol{\theta} = (\psi, \eta, \beta)$ does not exist since no common support!
- Reference prior? Berger & Bernardo (1992) algorithm does not work.
- An ad hoc reference prior for (ψ, η, β) might be used:

$$\pi(\psi, \eta, \beta) \propto \frac{1}{\eta\beta}.$$

Two-parameter Weibull Distribution

Priors for Weibull Distribution were investigated in Sun (1997).

Two-parameter Weibull Distribution

Priors for Weibull Distribution were investigated in Sun (1997).

If $X \sim Weibull(\eta, \beta)$, $Y = -\log(X)$ has the Type I Extreme Value Distribution, with the CDF,

$$F(y \mid \eta, \beta) = \exp\Big\{ -\exp\Big[-\beta(y + \log \eta)\Big]\Big\}, \quad y \in \mathbb{R}.$$

- It is a location-scale family with location parameter $\mu = -\log \eta$ and scale parameter $\sigma = 1/\beta$.
- The Jeffreys and the reference priors for (μ, σ) are

$$\pi^J(\mu, \sigma) \propto \frac{1}{\sigma^2}, \quad \pi^R(\mu, \sigma) \propto \frac{1}{\sigma}.$$

- $\pi^J(\mu, \sigma)$ is equivalent to $\pi^J(\eta, \beta) = 1/\eta$.
  $\pi^R(\mu, \sigma)$ is equivalent to $\pi^R(\eta, \beta) = 1/(\eta\beta)$.
- The corresponding Bayesian credible sets of (μ, σ) or (β, η) based on π^R are exact!

Motivated Example II: Three-parameter log-normal

Let $x_1, \ldots, x_n \sim$ log-normal(γ, μ, σ), with pdf,

$$f(x \mid \gamma, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}(x-\gamma)} \exp\left\{ -\frac{1}{2\sigma^2}\Big[\log(x-\gamma)-\mu\Big]^2\right\}, \quad x > \gamma. \tag{2}$$

- in socialogical and biological processes such as income of a population (Altchison & Brown, 1957).
- in the dosage requisite to produce a given response in biological assay (Finney, 1952).

Motivated Example III: Three-parameter Gamma Distribution

Consider one kind of the three parameter Gamma density,

$$f(x \mid \mu, \lambda, \alpha) = \frac{1}{\lambda\Gamma(\alpha)}\left(\frac{x-\mu}{\lambda}\right)^{\alpha-1}\exp\left(-\frac{x-\mu}{\lambda}\right), \quad x > \mu. \quad (3)$$

Weibull:

$$f(x \mid \psi, \eta, \beta) = \frac{\beta(x-\psi)^{\beta-1}}{\eta^{\beta}} \exp\left\{ -\left(\frac{x-\psi}{\eta}\right)^{\beta} \right\}, \quad x > \psi.$$

Log-normal:

$$f(x \mid \gamma, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}(x-\gamma)} \exp\left\{ -\frac{1}{2\sigma^2}\Big[\log(x-\gamma) - \mu\Big]^2 \right\}, \quad x > \gamma.$$

Gamma:

$$f(x \mid \mu, \lambda, \alpha) = \frac{1}{\lambda\Gamma(\alpha)}\left(\frac{x-\mu}{\lambda}\right)^{\alpha-1} \exp\left(-\frac{x-\mu}{\lambda}\right), \quad x > \mu.$$

Invariant Models

Definition. The family of densities $\mathcal{F}$, $f(x \mid \theta), \theta \in \mathbf{\Theta}$, is said to be invariant under the group of transformations $\mathcal{G}$ if for every $g \in \mathcal{G}$ and $\theta \in \mathbf{\Theta}$, there exists a unique $\theta^* \in \mathbf{\Theta}$ such that $Y = g(X)$ has density $f(y \mid \theta^*)$.

In such a situation, θ^* is denoted $\bar{g}(\theta)$. $\bar{\mathcal{G}} = \{\bar{g} : g \in \mathcal{G}\}$ is the induced group of transformations on $\mathbf{\Theta}$.

Invariant Models

Definition. The family of densities $\mathcal{F}$, $f(x \mid \theta), \theta \in \boldsymbol{\Theta}$, is said to be invariant under the group of transformations $\mathcal{G}$ if for every $g \in \mathcal{G}$ and $\theta \in \boldsymbol{\Theta}$, there exists a unique $\theta^* \in \boldsymbol{\Theta}$ such that $Y = g(X)$ has density $f(y \mid \theta^*)$.

In such a situation, θ^* is denoted $\bar{g}(\theta)$. $\bar{\mathcal{G}} = \{\bar{g} : g \in \mathcal{G}\}$ is the induced group of transformations on $\boldsymbol{\Theta}$.

Example. Location-scale group operation on a normal distribution.

- Suppose $X \sim N(\mu, \sigma^2)$, where $\boldsymbol{\Theta} = \{(\mu, \sigma) : \mu \in \mathbb{R},\ \sigma > 0\}$.
- Then $Y = aX + b \sim N(\mu^*, (\sigma^*)^2)$, where $\mu^* = a\mu + b$ and $\sigma^* = a\sigma$.
- It is clear that the group of transformations $\mathcal{G} = \{g : a > 0, b \in \mathbb{R}\}$, where $g(x) = ax + b$, induces the group of transformations $\bar{\mathcal{G}} = \{\bar{g} : a > 0, b \in \mathbb{R}\}$, where $\bar{g}_{a,b}((\mu, \sigma)) = (a\mu + b, a\sigma)$.

Invariant Priors

Definition. A right invariant Haar density on $\bar{\mathcal{G}}$, $h^r(\bar{g})$, is a density (with respect to Lebesgue measure) which, for $\mathbf{A} \subset \bar{\mathcal{G}}$ and all $\bar{g}_0 \in \bar{\mathcal{G}}$, satisfies

$$\int_{\mathbf{A}\bar{g}_0} h^r(y)dy = \int_{\mathbf{A}} h^r(x)dx,$$

where $\mathbf{A}\bar{g}_0 = \{\bar{g}\bar{g}_0 : \bar{g} \in \mathbf{A}\}$.

Definition. A left invariant Haar density on $\bar{\mathcal{G}}$, $h^l(\bar{g})$, is a density (with respect to Lebesgue measure) which, for $\mathbf{A} \subset \bar{\mathcal{G}}$ and all $\bar{g}_0 \in \bar{\mathcal{G}}$, satisfies

$$\int_{\bar{g}_0\mathbf{A}} h^l(y)dy = \int_{\mathbf{A}} h^l(x)dx,$$

where $\mathbf{A}\bar{g}_0 = \{\bar{g}_0\bar{g} : \bar{g} \in \mathbf{A}\}$.

Invariant Priors

Definition. A right invariant Haar density on $\bar{\mathcal{G}}$, $h^r(\bar{g})$, is a density (with respect to Lebesgue measure) which, for $\mathbf{A} \subset \bar{\mathcal{G}}$ and all $\bar{g}_0 \in \bar{\mathcal{G}}$, satisfies

$$\int_{\mathbf{A}\bar{g}_0} h^r(y)dy = \int_{\mathbf{A}} h^r(x)dx,$$

where $\mathbf{A}\bar{g}_0 = \{\bar{g}\bar{g}_0 : \bar{g} \in \mathbf{A}\}$.

Definition. A left invariant Haar density on $\bar{\mathcal{G}}$, $h^l(\bar{g})$, is a density (with respect to Lebesgue measure) which, for $\mathbf{A} \subset \bar{\mathcal{G}}$ and all $\bar{g}_0 \in \bar{\mathcal{G}}$, satisfies

$$\int_{\bar{g}_0\mathbf{A}} h^l(y)dy = \int_{\mathbf{A}} h^l(x)dx,$$

where $\mathbf{A}\bar{g}_0 = \{\bar{g}_0\bar{g} : \bar{g} \in \mathbf{A}\}$.

Note: For invariant decision problems, the best invariant decision rule is often the formal Bayes rule with respect to the right Haar measure on the group $\bar{\mathcal{G}}$.

Invariant Priors (Cont.)

Suppose model $f(\mathbf{x} \mid \boldsymbol{\theta})$, $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, has a group structure $\mathcal{G}$ with the induced group $\bar{\mathcal{G}}$. Assume $\bar{\mathcal{G}}$ is locally compact with the unit e. For any $\bar{g}_0 \in \bar{\mathcal{G}}$, the left and right haar measures:

$$\pi^l(\bar{g}_0) = \frac{1}{\left| det\Big(J(\bar{g} \to \bar{g}_0 \circ \bar{g})\Big) \right|_{\bar{g}=e}},$$

$$\pi^r(\bar{g}_0) = \frac{1}{\left| det\Big(J(\bar{g} \to \bar{g} \circ \bar{g}_0)\Big) \right|_{\bar{g}=e}}.$$

- $\pi^l(\bar{g}_0)$ is often the Jeffreys rule prior on $\boldsymbol{\Theta}$;
- $\pi^r(\bar{g}_0)$ is often a reference prior (Chang and Eavesl, 1990).

Invariant Priors (Cont.)

- For a location family $f(x-\mu)$,

$$\pi^l(\mu) = \pi^r(\mu) = 1.$$

- For a scale family $\sigma^{-1} f(x/\sigma)$,

$$\pi^l(\sigma) = \pi^r(\sigma) = 1/\sigma.$$

- For a location-scale family $\sigma^{-1} f((x-\mu)/\sigma)$,

$$\pi^l(\mu,\sigma) = 1/\sigma^2 \text{ and } \pi^r(\mu,\sigma) = 1/\sigma.$$

A Statistical Models with Semi-invariant Structure

Consider a statistical model $p(\mathbf{x} \mid \boldsymbol{\theta}, \boldsymbol{\xi})$, where $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ and $\boldsymbol{\xi} \in \Xi$. Suppose for each fixed $\boldsymbol{\theta}$, there is the (same) group invariance structure on Ξ, namely $\mathcal{G}$ with unit element $\mathbf{e}$.

- **Case A.** $\boldsymbol{\xi}$ is of interest and $\boldsymbol{\theta}$ is nuisance parameter(s).
- **Case B.** $\boldsymbol{\theta}$ is of interest and $\boldsymbol{\xi}$ is nuisance parameter(s);

A Statistical Models with Semi-invariant Structure

Consider a statistical model $p(\mathbf{x} \mid \boldsymbol{\theta}, \boldsymbol{\xi})$, where $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ and $\boldsymbol{\xi} \in \Xi$. Suppose for each fixed $\boldsymbol{\theta}$, there is the (same) group invariance structure on Ξ, namely $\mathcal{G}$ with unit element $\mathbf{e}$.

- **Case A.** $\boldsymbol{\xi}$ is of interest and $\boldsymbol{\theta}$ is nuisance parameter(s).
- **Case B.** $\boldsymbol{\theta}$ is of interest and $\boldsymbol{\xi}$ is nuisance parameter(s);

Theorem 1

Case A. If $\boldsymbol{\xi}$ is of interest and $\boldsymbol{\theta}$ is nuisance parameter, the reference prior of $(\boldsymbol{\xi}, \boldsymbol{\theta})$ is

$$\pi^R(\boldsymbol{\xi}, \boldsymbol{\theta}) = \pi^r(\boldsymbol{\xi})\pi^R(\boldsymbol{\theta}),$$

where $\pi^r(\boldsymbol{\xi})$ is the right Haar prior of $\boldsymbol{\xi}$ for the group $\mathcal{G}$ on Ξ, and $\pi^R(\boldsymbol{\theta})$ is the reference or Jeffreys prior for $p(\mathbf{x} \mid \boldsymbol{\theta}, \boldsymbol{\xi} = \mathbf{e})$.

Semi-location Family

Fact 1 Consider

$$f(x \mid \mu, \boldsymbol{\theta}) = g(x - \mu \mid \boldsymbol{\theta}), \quad (4)$$

where $g(\cdot \mid \boldsymbol{\theta})$ is a known density depending on $\boldsymbol{\theta}$ only, and $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^k$. Clearly, the right haar prior for μ is a constant prior,

$$\pi^r(\mu) \propto 1.$$

The reference prior when μ is of interest is

$$\pi^R(\mu, \boldsymbol{\theta}) = \pi^R(\boldsymbol{\theta}),$$

where $\pi^R(\boldsymbol{\theta})$ is the reference prior for $\boldsymbol{\theta}$ under the model $g(\cdot \mid \boldsymbol{\theta})$.

Motivated Example II: Three-parameter log-normal–Continued

$$f(x \mid \gamma, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}(x-\gamma)} \exp\left\{ -\frac{1}{2\sigma^2}\left[\log(x-\gamma)-\mu\right]^2\right\}, \ x > \gamma$$

Let $\xi = \gamma$ and $\boldsymbol{\theta} = (\mu, \sigma)$. When parameter γ is of interest,

$$\pi^R(\gamma, \mu, \sigma) = \frac{1}{\sigma}.$$

Here, we apply the one-at-a-time reference priors $\pi^R(\mu, \sigma) = 1/\sigma$.

Motivated Example II: Three-parameter log-normal–Continued

$$f(x \mid \gamma, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}(x-\gamma)} \exp\left\{ -\frac{1}{2\sigma^2}\Big[\log(x-\gamma)-\mu\Big]^2 \right\}, \ x > \gamma$$

Let $\xi = \gamma$ and $\boldsymbol{\theta} = (\mu, \sigma)$. When parameter γ is of interest,

$$\pi^R(\gamma, \mu, \sigma) = \frac{1}{\sigma}.$$

Here, we apply the one-at-a-time reference priors $\pi^R(\mu, \sigma) = 1/\sigma$.

Hill (1963) considered the "independent" Jeffreys prior:

$$\pi^{IJ}(\gamma, \mu, \sigma) = \frac{1}{\sigma}.$$

From Hill (1963),

- if $\gamma \in \mathbb{R}$, the joint posterior is always improper!
- if $\gamma \in (\gamma_0, \infty)$, the joint posterior is proper if $n \geq 3$.

Semi-location-scale Family

Fact 2 Consider

$$f(x \mid \mu, \sigma, \boldsymbol{\theta}) = \frac{1}{\sigma} g\left(\frac{x-\mu}{\sigma} \mid \boldsymbol{\theta}\right), \tag{5}$$

where $\sigma > 0$ and $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^k$. Here μ is a location parameter, σ is a scale parameter, and $g(\cdot, \boldsymbol{\theta})$ is a known density depending on $\boldsymbol{\theta}$ only. We let $\boldsymbol{\xi} = (\mu, \sigma)$. Clearly,

$$\pi^r(\mu, \sigma) \propto \frac{1}{\sigma}.$$

Then

$$\pi^R(\mu, \sigma, \boldsymbol{\theta}) = \frac{1}{\sigma} \pi^R(\boldsymbol{\theta}),$$

where $\pi^R(\boldsymbol{\theta})$ is the reference prior for $\boldsymbol{\theta}$ under the model $g(\cdot \mid \boldsymbol{\theta})$.

Motivated Example I: Three-parameter Weibull–Continued

$$f(x \mid \psi, \eta, \beta) = \frac{\beta(x-\psi)^{\beta-1}}{\eta^{\beta}} \exp\left\{ -\left(\frac{x-\psi}{\eta}\right)^{\beta} \right\}, \quad x > \psi,$$

It is a semi-location-scale family, with $\theta = \beta$,

$$f(x \mid \psi, \eta, \beta) = \frac{1}{\eta} g\left(\frac{x-\psi}{\eta} \;\Big|\; \beta\right),$$

where $g(y \mid \beta) = \beta y^{\beta-1} e^{-y^{\beta}} 1_{(0,\infty)}(y)$.
If $Y \sim g(y \mid \beta)$, $\log(Y)$ is a scale family with scale parameter $1/\beta$.
The reference prior for β is of the form, $\pi^R(\beta) = 1/\beta, \ \beta > 0$.
Consequently, the reference prior for (ψ, η, β) is

$$\pi^R(\psi, \eta, \beta) = \frac{1}{\eta\beta}, \quad \psi \in \mathbb{R}, \eta > 0, \beta > 0.$$

Motivated Example I: Three-parameter Weibull–Continued

$$f(x \mid \psi, \eta, \beta) = \frac{\beta(x-\psi)^{\beta-1}}{\eta^{\beta}} \exp\left\{ -\left(\frac{x-\psi}{\eta}\right)^{\beta} \right\}, \quad x > \psi,$$

It is a semi-location-scale family, with $\theta = \beta$,

$$f(x \mid \psi, \eta, \beta) = \frac{1}{\eta} g\left(\frac{x-\psi}{\eta} \;\middle|\; \beta \right),$$

where $g(y \mid \beta) = \beta y^{\beta-1} e^{-y^{\beta}} 1_{(0,\infty)}(y)$.
If $Y \sim g(y \mid \beta)$, $\log(Y)$ is a scale family with scale parameter $1/\beta$.
The reference prior for β is of the form, $\pi^R(\beta) = 1/\beta,\ \beta > 0$.
Consequently, the reference prior for (ψ, η, β) is

$$\pi^R(\psi, \eta, \beta) = \frac{1}{\eta\beta}, \quad \psi \in \mathbb{R}, \eta > 0, \beta > 0.$$

Fact
(a) if $\psi \in \mathbb{R}$, the joint posterior is always improper!
(b) if $\psi \in (\mu_0, \infty)$, the joint posterior is proper if $n \geq 3$.

Motivated Example III: Three-parameter Gamma – Continued

$$f(x \mid \mu, \lambda, \alpha) = \frac{1}{\lambda\Gamma(\alpha)}\Big(\frac{x-\mu}{\lambda}\Big)^{\alpha-1}\exp\Big(-\frac{x-\mu}{\lambda}\Big), \quad x > \mu.$$

Let $y = (x-\mu)/\lambda$. Then y depends on α only with the pdf,

$$g(y \mid \alpha) = \frac{1}{\Gamma(\alpha)} y^{\alpha-1}\exp(-y), \quad y > 0.$$

The fisher information for α is $-E\Big(\frac{\partial^2 l}{\partial\alpha^2}\Big) = \frac{\partial^2 \log\Gamma(\alpha)}{\partial\alpha^2}$. Then $\pi^R(\alpha) = \sqrt{\frac{\partial^2}{\partial\alpha^2}\log\Gamma(\alpha)}$ and

$$\pi^R(\mu, \lambda, \alpha) = \frac{1}{\lambda}\sqrt{\frac{\partial^2}{\partial\alpha^2}\log\Gamma(\alpha)}, \quad \mu \in \mathbb{R}, \lambda > 0, \alpha > 0.$$

Motivated Example III: Three-parameter Gamma – Continued

$$f(x \mid \mu, \lambda, \alpha) = \frac{1}{\lambda \Gamma(\alpha)} \left(\frac{x-\mu}{\lambda} \right)^{\alpha-1} \exp\left(-\frac{x-\mu}{\lambda} \right), \quad x > \mu.$$

Let $y = (x-\mu)/\lambda$. Then y depends on α only with the pdf,

$$g(y \mid \alpha) = \frac{1}{\Gamma(\alpha)} y^{\alpha-1} \exp(-y), \quad y > 0.$$

The fisher information for α is $-E\left(\frac{\partial^2 l}{\partial \alpha^2}\right) = \frac{\partial^2 \log \Gamma(\alpha)}{\partial \alpha^2}$. Then $\pi^R(\alpha) = \sqrt{\frac{\partial^2}{\partial \alpha^2} \log \Gamma(\alpha)}$ and

$$\pi^R(\mu, \lambda, \alpha) = \frac{1}{\lambda} \sqrt{\frac{\partial^2}{\partial \alpha^2} \log \Gamma(\alpha)}, \quad \mu \in \mathbb{R}, \lambda > 0, \alpha > 0.$$

Fact

Under $\pi^R(\mu, \lambda, \alpha)$, the joint posterior exists iff $n \geq 3$.

Simulations for Three-parameter Weibull

Table: Frequentist coverage probabilities of one-side .05 posterior credible sets of ψ, η and β for sample size $n = 10.(\psi^* = 1, \eta^* = 1)$

β^*	$Q(.05; \psi)$	$Q(.05; \eta)$	$Q(.05; \beta)$
.5	0.0637	0.0563	0.0686
1	0.0256	0.0508	0.0484

Table: Frequentist coverage probabilities of one-side .95 posterior credible sets of ψ, η and β for sample size $n = 10.(\psi^* = 1, \eta^* = 1)$

β^*	$Q(.95; \psi)$	$Q(.95; \eta)$	$Q(.95; \beta)$
.5	0.9963	0.9128	0.9782
1	0.9751	0.9794	0.9678

Table: Frequentist coverage probabilities of one-side .05 posterior credible sets of ψ, η and β for sample size $n = 5.(\psi^* = 1, \eta^* = 1)$

β^*	$Q(.05;\psi)$	$Q(.05;\eta)$	$Q(.05;\beta)$
.5	0.0314	0.0345	0.0438
1	0.0164	0.0062	0.0594

Table: Frequentist coverage probabilities of one-side .95 posterior credible sets of ψ, η and β for sample size $n = 5.(\mu^* = 1, \eta^* = 1)$

β^*	$Q(.95;\psi)$	$Q(.95;\eta)$	$Q(.95;\beta)$
.5	0.9935	0.9369	0.9755
1	0.9932	0.9369	0.9793

Case B. $\boldsymbol{\theta}$ is of Interest

Suppose θ is the parameter of interest and $\boldsymbol{\xi}$ is nuisance parameter. There is a group structure $\mathcal{G}$ for Ξ with the right haar prior $\pi^r(\Xi)$. Assume that there is a maximum invariance statistic $\mathbf{T}$ for Ξ. Since

$$f(\mathbf{x} \mid \boldsymbol{\theta}, \boldsymbol{\xi}) = f(\mathbf{T} \mid \theta) f(\mathbf{x} \mid \mathbf{T}, \boldsymbol{\theta}, \boldsymbol{\xi}),$$

the marginal likelihood funciton of $\boldsymbol{\theta}$ is

$$\int_{\Xi} f(\mathbf{x} \mid \boldsymbol{\theta}, \boldsymbol{\xi}) \pi^r(\boldsymbol{\xi}) d\boldsymbol{\xi} = f(\mathbf{T} \mid \theta)\ H,$$

where

$$H = \int_{\Xi} f(\mathbf{x} \mid \mathbf{T}, \boldsymbol{\theta}, \boldsymbol{\xi}) \pi^r(\boldsymbol{\xi}) d\boldsymbol{\xi}. \tag{6}$$

Theorem 2

(a) H in (6) does not depend on $\boldsymbol{\theta}$.

(b) The reference prior for $\boldsymbol{\theta}$ is based on the model $f(\mathbf{T} \mid \theta)$ only.

Example IV: Bivariate Normal

The bivariate normal distribution of (x_1, x_2) has mean parameters (μ_1, μ_2) and variance matrix $\mathbf{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$, where ρ is the correlation between x_1 and x_2.

The priors when ρ is of interest have been facinating since 1930. See Berger and Sun (2008). Let $\boldsymbol{\xi} = (\mu_1, \mu_2, \sigma_1, \sigma_2)$ and consider a group of transformations

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \to \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} b_1 & 0 \\ 0 & b_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

where $a_1, a_2 \in \mathbb{R}$, $b_1, b_2 > 0$. The right haar is $\pi^r(\boldsymbol{\xi}) = 1/(\sigma_1\sigma_2)$. One can show that $\pi^R(\rho) = 1/(1-\rho^2)$.
The reference prior for $\{\rho, \boldsymbol{\xi}\}$ is the Lindley-Bayarri prior

$$\pi^R(\rho, \mu_1, \mu_2, \sigma_1, \sigma_2) = \frac{1}{\sigma_1\sigma_2(1-\rho^2)}.$$

Comments

- Invariant priors, which are reference priors, are quite successful in practice.
- For a model with semi-invariance structure, we should take the advantage of the invariance component.
- It can be applied to many examples.
- Frequentist matching under a semi-invariance structure?
- *Weakness?*