Web Appendix A: Bayesian Hypothesis Testing

Consider the hypotheses $H_0 : \theta \in \Theta_0$ versus $H_1 : \theta \in \Theta_1$ where $\Theta_0 \bigcup \Theta_1 = \Theta$ and $\Theta_0 \bigcap \Theta_1 = \emptyset$. Formal Bayesian hypothesis testing requires the specification of prior probabilities on the hypotheses (e.g. $p(H_i)$ for $i = 0, 1$) and prior distributions for θ specified over the parameter space defined with respect to each of the hypotheses (e.g. $\pi(\theta \big| H_i)$ for $i = 0, 1$).

The posterior probability of hypothesis H_i is given by

$$p(H_i|\boldsymbol{D}) = \frac{p(\boldsymbol{D}|H_i) \cdot p(H_i)}{p(\boldsymbol{D}|H_0) \cdot p(H_0) + p(\boldsymbol{D}|H_1) \cdot p(H_1)}, \tag{1}$$

where $p(\boldsymbol{D}|H_i) = \int_{\Theta_i} p(\boldsymbol{D}|\theta)\pi(\theta|H_i)d\theta$ is the marginal likelihood associated with hypothesis H_i. In practice, most Bayesian hypothesis testing methods are based on the posterior probability of the *event defining* H_i. For this approach, one simply needs to specify a prior $\pi(\theta)$ representing belief about θ and compute the posterior distribution. The posterior probability that $\theta \in \Theta_i$ is given by

$$P(\theta \in \Theta_i|\boldsymbol{D}) = \frac{\int_{\Theta_i} p(\boldsymbol{D}|\theta)\pi(\theta|\theta \in \Theta_i)d\theta \cdot P(\theta \in \Theta_i)}{\sum_{j=0,1} \int_{\Theta_j} p(\boldsymbol{D}|\theta)\pi(\theta|\theta \in \Theta_j)d\theta \cdot P(\theta \in \Theta_j)} \tag{2}$$

where $P(\theta \in \Theta_i) = \int_{\Theta_i} \pi(\theta)d\theta$. We can readily see that the $P(\theta \in \Theta_i|\boldsymbol{D})$ is equal to $p(H_i|\boldsymbol{D})$ if one takes $p(H_i) = P(\theta \in \Theta_i)$ and $\pi\left(\theta \big| H_i\right) = \pi\left(\theta \big| \theta \in \Theta_i\right)$ for $i = 0, 1$. If in fact $\pi(\theta)$ does represent belief about θ, these choices are perhaps the most intuitive and thus we should have no reservation referring to $P(\theta \in \Theta_i|\boldsymbol{D})$ as the probability that hypothesis H_i is true.

Web Appendix B: Parameterizing Flattened and Concentrated Monitoring Priors

Recall the value of the normal density at the mode is $\frac{1}{\sqrt{2\pi}\sigma}$ and note that the value of a generalized normal density at the mode is $\frac{\beta}{2\alpha\Gamma(1/\beta)}$. These are equivalent when $\beta = 2$ and $\alpha = \sqrt{2}\sigma$ (i.e. the normal density is a special case of the generalized normal density at these parameter values). Let $F_{\mu,\alpha,\beta}$ denote the cumulative distribution function of the generalized

normal distribution $\mathcal{GN}(\mu, \alpha, \beta)$, which can be expressed as (Griffin, 2018)

$$P(\theta \leqslant q|\mu, \alpha, \beta) = \frac{1}{2} + \frac{\text{sign}(q-\mu)}{2} \int_0^{|q-\mu|^\beta} \frac{w^{1/\beta-1}}{\alpha\Gamma(1/\beta)} \exp\left\{-\left(\frac{1}{\alpha}\right)^\beta w\right\} dw.$$

A flattened or concentrated enthusiastic monitoring prior in the generalized normal family of distributions has density at the mode equal to $k \times \frac{1}{\sqrt{2\pi}\sigma}$. The parameters for the generalized normal distribution $\mathcal{GN}(\mu, \alpha, \beta)$ are derived as follows: μ remains equal to the mode value of θ_1 and α and β are determined to minimize the function

$$\left(F_{\mu,\alpha,\beta}(\theta_0) - \epsilon\right)^2 + \left(\frac{\beta}{2\alpha\Gamma(1/\beta)} - k\frac{1}{\sqrt{2\pi}\sigma}\right)^2$$

with box-constrained optimization (Byrd et al., 1995), where $\sigma = \frac{\theta_1-\theta_0}{\Phi^{-1}(1-\epsilon)}$ is the standard deviation of the default normally distributed enthusiastic monitoring prior. The first term reflects the residual uncertainty that $\theta < \theta_0$, and the second term reflects the density at the mode value. Similarly, the parameters for a flattened or concentrated skeptical monitoring prior are as follows: μ remains equal to the mode value of θ_0 and α and β are determined to minimize the function

$$\left((1 - F_{\mu,\alpha,\beta}(\theta_1)) - \epsilon\right)^2 + \left(\frac{\beta}{2\alpha\Gamma(1/\beta)} - k\frac{1}{\sqrt{2\pi}\sigma}\right)^2,$$

where $\sigma = \frac{\theta_0-\theta_1}{\Phi^{-1}(\epsilon)}$.

This parameterizing procedure is applicable to a generalized normal distribution truncated to an interval domain (e.g. when θ is a response probability with domain $[0, 1]$). In this case, the generalized normal distribution truncated to an interval domain $\Theta = (\theta_{min}, \theta_{max})$ has density equal to $f(\theta) = c \cdot \exp\left\{-\frac{|\theta-\mu|}{\alpha}^\beta\right\} I(\theta \in \Theta)$ where $c = \frac{\beta}{2\alpha\Gamma(1/\beta)}(F_{\mu,\alpha,\beta}(\theta_{max}) - F_{\mu,\alpha,\beta}(\theta_{min}))^{-1}$.

References

Byrd, R. H., Lu, P., Nocedal, J., and Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing* **16,** 1190–1208.

Griffin, M. (2018). Working with the exponential power distribution using gnorm.