

1 Possible priors

We consider several different possible priors for the parameters of the behavior stream. Depending on the context, it may be more appropriate to express these in terms of (μ, λ) or in terms of (ϕ, ζ) ; we therefore consider the implications of different parameterizations. The goal is to identify distributional families that can be used to express “vague” priors, which regularize the posterior but are dominated by the likelihood for reasonable sample sizes, and that have a sensible substantive interpretation, thus providing a convenient means for expressing stronger prior information when available. We consider several possibilities and parameterizations.

1.1 Gamma priors on (μ, λ)

One obvious candidate for priors on (μ, λ) is the gamma distributions, which has the useful property that the implied priors on ϕ and ζ have a familiar distributional form. Consider the priors

$$\mu \sim \text{Gamma}(\kappa_\mu, \theta_\mu), \quad \lambda \sim \text{Gamma}(\kappa_\lambda, \theta_\lambda),$$

with μ and λ taken to be independent. The penalty function, written in terms of ϕ and ζ , is

$$f_G(\phi, \zeta) = (\kappa_\mu - 1) \log \phi + (\kappa_\lambda - 1) \log(1 - \phi) - (\kappa_\mu + \kappa_\lambda - a) \log(\zeta) - \frac{\frac{\phi}{\theta_\mu} + \frac{1-\phi}{\theta_\lambda}}{\zeta},$$

with $a = 2$.

If we take $\theta_\mu = \theta_\lambda = \theta$, then

$$\phi \sim \text{Beta}(\kappa_\mu, \kappa_\theta), \quad \zeta \sim \Gamma^{-1}(\kappa_\mu + \kappa_\lambda, 1/\theta),$$

where Γ^{-1} denotes the inverse-gamma distribution. Parameterizing the prior in terms of (ϕ, ζ) leads to the same penalty function as above, but with $a = -1$; parameterizing in terms of (ϕ, ζ^{-1}) has $a = 1$. Note that we must take $\kappa_\mu, \kappa_\lambda > 1$ for the penalty function to regularize the likelihood.

1.2 Log-normal priors on (μ, λ)

Another possibility would be to log-normal priors:

$$\log \mu \sim N(\gamma_\mu, \sigma_\mu^2), \quad \log \lambda \sim N(\gamma_\lambda, \sigma_\lambda^2),$$

with μ and λ again taken to be independent. These priors imply that $\text{logit } \phi \sim N(\gamma_\mu - \gamma_\lambda, \sigma_\mu^2 + \sigma_\lambda^2)$, although they do not lead to a tractable form for the prior on ζ . Parameterizing the prior in terms of μ, λ , the penalty function is

$$f_{LN1}(\phi, \zeta) = -\frac{1}{2} \left[\left(\frac{\log(\phi) - \log(\zeta) - \gamma_\mu}{\sigma_\mu} \right)^2 + \left(\frac{\log(1 - \phi) - \log(\zeta) - \gamma_\lambda}{\sigma_\lambda} \right)^2 \right] - \log \phi - \log(1 - \phi) + 2 \log \zeta.$$

Alternately, parameterizing the prior in terms of $\log \mu, \log \lambda$ leads to the final three terms dropping out:

$$f_{LN2}(\phi, \zeta) = -\frac{1}{2} \left[\left(\frac{\log(\phi) - \log(\zeta) - \gamma_\mu}{\sigma_\mu} \right)^2 + \left(\frac{\log(1 - \phi) - \log(\zeta) - \gamma_\lambda}{\sigma_\lambda} \right)^2 \right].$$

A disadvantage of this penalty function is that it is not obvious which choices for the hyperparameters will regularize the likelihood, because the limits (approaching the boundaries of the parameter space) of the derivatives with respect to ϕ and ζ do not always have a constant sign.

1.3 Normal priors on (logit ϕ , log ζ)

A further possibility would be to put normal priors on transformations of ϕ and ζ that have unbounded supports. This would be reasonable, for example, if one were constructing a generalized linear model with a logistic link for ϕ and a log link for ζ . Thus, we might assume

$$\text{logit } \phi \sim N(\gamma_\phi, \sigma_\phi^2), \quad \log \zeta \sim N(\gamma_\zeta, \sigma_\zeta^2),$$

with ϕ and ζ independent. The implied priors on μ, λ are not familiar distributions. The corresponding penalty function is

$$f_N(\phi, \zeta) = -\frac{1}{2} \left[\left(\frac{\text{logit}(\phi) - \gamma_\phi}{\sigma_\phi} \right)^2 + \left(\frac{\log(\zeta) - \gamma_\zeta}{\sigma_\zeta} \right)^2 \right].$$

1.4 Dirichlet prior on $p_0(c+d), p_1(c+d)$

A final alternative would be to put priors on $p_0(c+d), p_1(c+d)$, rather than directly on the parameters of the behavior stream. This approach is computationally convenient for the MTS estimators, in that it leads to closed-form expressions for the posterior mode of $l_{MTS}(\phi, \zeta)$. Specifically, one could assume that

$$p_0(c+d), 1-p_1(c+d) \sim \text{Dirichlet}(\alpha_0, \alpha_1, \alpha_2).$$

From the fact that $\phi = p_0/(p_0 + 1 - p_1)$, the implied prior for prevalence $\phi \sim \text{Beta}(\alpha_0, \alpha_1)$. However, the implied prior for incidence does not have a convenient form, though the marginal distribution of ζ appears to be quite similar to a gamma distribution. Neither do the implied priors on (μ, λ) have a tractable form, which makes it rather difficult to choose sensible hyperparameters.

The penalty function corresponding to the Dirichlet prior distribution is

$$f_{Dir}(\phi, \zeta) = (\alpha_0 - 1) \log \phi + (\alpha_1 - 1) \log(1 - \phi) + (\alpha_0 + \alpha_1 - 2) \log \left[1 - \exp \left(\frac{-c\zeta}{\phi(1 - \phi)} \right) \right] - \frac{(\alpha_2 - 1)c\zeta}{\phi(1 - \phi)}.$$

We must take $\alpha_0, \alpha_1, \alpha_2 > 1$ in order for the penalty function to regularize the likelihood.