

Quick Notes on Exponential Family Conjugate Priors

Sébastien Lahaie

September 16, 2013

In Sindhu's notes it's assumed that the agent has a belief \mathbf{q} , and moves the market state from initial state \mathbf{q}_{init} to a convex combination $\tilde{\mathbf{q}} = \lambda\mathbf{q} + (1 - \lambda)\mathbf{q}_{init}$. This note tries to justify this based on Bayesian updating, rather than budget limitations. (But it doesn't quite succeed.)

Let $p(x; \theta)$ denote a probability density drawn from an exponential family with sufficient statistic $\phi : \mathcal{X} \rightarrow \mathbf{R}^d$, where θ is the natural parameter:

$$p(x; \theta) = \exp [\langle \theta, \phi(x) \rangle],$$

and

$$g(\theta) = \log \int_{\mathcal{X}} \exp \langle \phi(x), \theta \rangle dx.$$

Recall that $\nabla g(\theta) = \mathbf{E}[\phi(x)]$ and $\nabla^2 g(\theta) = \text{Var}[\phi(x)]$. The family of conjugate priors is also an exponential family and takes the form

$$p(\theta; n, \nu) = \exp [\langle n\nu, \theta \rangle - ng(\theta) - h(\nu, n)].$$

Here the feature map is $\psi(\theta) = (\theta, -g(\theta))$, the natural parameter is $(n\nu, n)$ where $n \in \mathbf{R}$ and $\nu \in \mathbf{R}^d$. The normalizer $h(\nu, n)$ is convex in $(n\nu, n)$. It is helpful to think of the prior as being based on a 'phantom' sample of size n and mean ν . The justification for this is that

$$\mathbf{E}_{\theta} [\mathbf{E}_x [\phi(x)|\theta]] = \mathbf{E}_{\theta} [\nabla g(\theta)] = \nu.$$

(The proof is straightforward but not obvious. I have a reference for this but it's impossible to read—it's a mathematical statistics paper.)

Suppose we draw a sample $X = (x_1, \dots, x_m)$ of size m , and denote the empirical mean by $\mu[X] = \sum_{i=1}^m \phi(x_i)$. The posterior distribution is then

$$p(\theta|X) \propto p(X|\theta)p(\theta|n, \nu) \propto \exp [\langle \mu[X] + n\nu, \theta \rangle - (m+n)g(\theta)],$$

and so the posterior mean is

$$\frac{m\mu[X] + n\nu}{m + n}. \quad (1)$$

Thus the posterior mean is a convex combination of the prior and posterior means, and their relative weights depend on the phantom and empirical sample sizes.

In the context of prediction markets, one could imagine an agent who uses the current market estimate as its prior, and draws a empirical sample of size m . Its belief then takes the form (1), where n depends on the importance the agent places on the market estimate (perhaps based on how long the market has been running). However, note that the *mean parameter* becomes a convex combination of market state and empirical belief, and this does not translate to the *natural parameter*, which is what we would have liked. The new natural parameter is

$$\nabla g^{-1} \left(\frac{m\mu[X] + n\nu}{m + n} \right) = \nabla g^* \left(\frac{m\mu[X] + n\nu}{m + n} \right).$$

where g^* is the convex conjugate of g .