Bayesian Prediction

Recall $m(x)=\int f(x \mid \theta) \pi(\theta) d \theta$ is the marginal distribution, which is sometimes called the **prior predictive distribution**.

$$f\left(x_{n+1} \mid x_1, \ldots, x_n\right)=\int f\left(x_{n+1} \mid \theta\right) \pi\left(\theta \mid x_1, \ldots, x_n\right) d \theta$$

The above is called the **posterior predictive distribution**.

$$\hat{X}_{n+1}=\int x_{n+1} \times f\left(x_{n+1} \mid x_1, \ldots, x_n\right) d x_{n+1}=\mathbb{E}\left(X_{n+1} \mid X_1, \ldots, X_n\right)$$

The above is called the **predictive mean** (prediction for X_{n+1}).

$$\int\left(x_{n+1}-\hat{X}_{n+1}\right)^2 f\left(x_{n+1} \mid x_1, \ldots, x_n\right) d x_{n+1}$$

The above is called the **predictive variance**.

Example 1

Observations from Exponential distribution with Gamma prior on λ:
$x_1, \ldots, x_n \sim \operatorname{Exp}(\lambda), f\left(x_i\right)=\lambda e^{-\lambda x_i}, \pi(\lambda)=\frac{\beta^\alpha \lambda^{\alpha-1}}{\Gamma(\alpha)} e^{-\beta \lambda}, \lambda \geq 0$
Likelihood:

$$\begin{aligned}
L\left(\lambda \mid x_1, \ldots, x_n\right) & =\prod_{i=1}^n \lambda e^{-\lambda x_i} \\
& =\lambda^n e^{-\lambda \sum_{i=1}^n x_i}
\end{aligned}$$

Posterior:

$$\begin{aligned}
L\left(\lambda \mid x_1, \ldots, x_n\right) \pi(\lambda) & =\lambda^n e^{-\lambda \sum_{i=1}^n x_i} \frac{\beta^\alpha \lambda^{\alpha-1}}{\Gamma(\alpha)} e^{-\beta \lambda} \\
& =\frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} \lambda^n e^{-\lambda \sum_{i=1}^n x_i} e^{-\beta \lambda} \\
& =C \lambda^{\alpha+n-1} e^{-\lambda\left(\sum_{i=1}^n x_i+\beta\right)} \\
& \propto \operatorname{Gamma}\left(\alpha+n, \sum_{i=1}^n x_i+\beta\right)
\end{aligned}$$

The pdf of the posterior is then

$$\pi\left(\lambda \mid x_1, \ldots, x_n\right) = \frac{\left(\sum_{i=1}^n x_i + \beta\right)^{\alpha+n}}{\Gamma(\alpha+n)} \lambda^{\alpha+n-1} e^{-\lambda\left(\sum_{i=1}^n x_i + \beta\right)} \text{ for } \lambda > 0$$

We also need the help of the Gamma function (not distribution) to help solve the next problem:

$$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx \text{ for } z > 0. \text{ We also have the result that } \Gamma(z+1) = z\Gamma(z)$$

Posterior predictive distribution

$$\begin{aligned}
f\left(x_{n+1} \mid x_1, \ldots, x_n\right) &= \int_0^\infty \lambda e^{-\lambda x_{n+1}} \pi\left(\lambda \mid x_1, \ldots, x_n\right) d\lambda \\
&= \int_0^\infty \lambda e^{-\lambda x_{n+1}} \frac{\left(\sum_{i=1}^n x_i + \beta\right)^{\alpha+n}}{\Gamma(\alpha+n)} \lambda^{\alpha+n-1} e^{-\lambda\left(\sum_{i=1}^n x_i + \beta\right)} d\lambda \\
&= \frac{\left(\sum_{i=1}^n x_i + \beta\right)^{\alpha+n}}{\Gamma(\alpha+n)} \int_0^\infty \lambda^{\alpha+n} e^{-\lambda\left(x_{n+1} + \sum_{i=1}^n x_i + \beta\right)} d\lambda \\
&\text{Substituting } u = \lambda\left(x_{n+1} + \sum_{i=1}^n x_i + \beta\right), du = \left(x_{n+1} + \sum_{i=1}^n x_i + \beta\right) d\lambda \\
&= \frac{\left(\sum_{i=1}^n x_i + \beta\right)^{\alpha+n}}{\Gamma(\alpha+n)} \int_0^\infty \left(\frac{u}{x_{n+1} + \sum_{i=1}^n x_i + \beta}\right)^{\alpha+n} e^{-u} \frac{du}{\left(x_{n+1} + \sum_{i=1}^n x_i + \beta\right)} \\
&= \frac{\left(\sum_{i=1}^n x_i + \beta\right)^{\alpha+n}}{\Gamma(\alpha+n)\left(x_{n+1} + \sum_{i=1}^n x_i + \beta\right)^{\alpha+n+1}} \int_0^\infty u^{\alpha+n} e^{-u} du \\
&= \frac{\left(\sum_{i=1}^n x_i + \beta\right)^{\alpha+n} \Gamma(\alpha+n+1)}{\Gamma(\alpha+n)\left(x_{n+1} + \sum_{i=1}^n x_i + \beta\right)^{\alpha+n+1}} \\
&= \frac{\left(\sum_{i=1}^n x_i + \beta\right)^{\alpha+n} (\alpha+n)\Gamma(\alpha+n)}{\Gamma(\alpha+n)\left(x_{n+1} + \sum_{i=1}^n x_i + \beta\right)^{\alpha+n+1}} \\
&= \frac{(\alpha+n)\left(\sum_{i=1}^n x_i + \beta\right)^{\alpha+n}}{\left(x_{n+1} + \sum_{i=1}^n x_i + \beta\right)^{\alpha+n+1}}
\end{aligned}$$

Thus $x_{n+1} + \sum_{i=1}^n x_i + \beta$ has a Pareto distribution with parameters $\sum_{i=1}^n x_i + \beta$ and $\alpha + n$.

Example 2

If $X \sim Pa(c, \alpha)$, then $f(x) = \frac{\alpha}{c}\left(\frac{c}{x}\right)^{\alpha+1}, x \geq c$. We have

$$E[X] = \frac{\alpha c}{\alpha - 1}, \alpha > 1$$

and

$$\operatorname{Var}(X) = \frac{\alpha c^2}{(\alpha-1)^2(\alpha-2)}, \quad \alpha > 2$$

We have $x_{n+1} + \sum_{i=1}^n x_i + \beta \sim \mathrm{Pa}\left(\sum_{i=1}^n x_i + \beta, \alpha + n\right)$
Then

$$\begin{aligned} E\hat{X}_{n+1} &= EX_{n+1} \\ &= \frac{\left(\sum_{i=1}^n x_i + \beta\right)(\alpha + n)}{\alpha + n - 1} - \sum_{i=1}^n x_i - \beta \\ &= \frac{\sum_{i=1}^n x_i + \beta}{\alpha + n - 1} \end{aligned}$$

Exercise for reader: show

$$\hat{\sigma}^2_{x_{n+1}} = \frac{\left(\sum_{i=1}^n x_i + \beta\right)^2 (n + \alpha)}{(\alpha + n - 1)^2(\alpha + n - 2)}$$

For example if $x_1 = 2.1, x_2 = 5.5, x_3 = 6.4, x_4 = 8.7, x_5 = 4.9, x_6 = 5.1, x_7 = 2.3$ and $\lambda \sim Ga(2, 1)$, then

$$\hat{X}_8 = \frac{9}{2}, \hat{\sigma}^2_{x_8} = 26.0357$$

This is easier if only $\hat{X}_{n+1}$ is wanted:

$$\hat{X}_{n+1} = \int_\theta \mu(\theta)\pi\left(\theta \mid x_1, \ldots, x_n\right) d\theta$$

where $\mu(\theta) = \mathbb{E}[X] = \int x f(x \mid \theta) dx$ is the mean of X.