The Gamma/Poisson Bayesian Model

- If our data $X_1, \ldots, X_n$ are iid Poisson($\lambda$), then a gamma($\alpha, \beta$) prior on $\lambda$ is a **conjugate** prior.

Likelihood:

$$L(\lambda | x) = \prod_{i=1}^{n} \frac{e^{-\lambda} \lambda^{x_i}}{x_i!} = \frac{e^{-n\lambda} \lambda^{\sum x_i}}{\prod_{i=1}^{n}(x_i!)}$$

Prior:

$$p(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta \lambda}, \quad \lambda > 0.$$  

$\implies$ Posterior:

$$\pi(\lambda | x) \propto \lambda^{\sum x_i + \alpha - 1} e^{-(n+\beta)\lambda}, \quad \lambda > 0.$$  

$\implies$ $\pi(\lambda | x)$ is gamma($\sum x_i + \alpha, n + \beta$).  

(Conjugate!)
The posterior mean is:

\[ \hat{\lambda}_B = \frac{\sum x_i + \alpha}{n + \beta} \]

\[ = \frac{\sum x_i}{n + \beta} + \frac{\alpha}{n + \beta} \]

\[ = \left[ \frac{n}{n + \beta} \right] \left( \frac{\sum x_i}{n} \right) + \left[ \frac{\beta}{n + \beta} \right] \left( \frac{\alpha}{\beta} \right) \]

Again, the data get weighted more heavily as \( n \to \infty \).
We can use the Bayesian approach to update our information about the parameter(s) of interest sequentially as new data become available.

Suppose we formulate a prior for our parameter $\theta$ and observe a random sample $x_1$.

Then the posterior is

$$
\pi(\theta|x_1) \propto p(\theta)L(\theta|x_1)
$$

Then we observe a new (independent) sample $x_2$. 
We can use our previous posterior as the **new prior** and derive a **new** posterior:

\[
\pi(\theta|x_1, x_2) \propto \pi(\theta|x_1)L(\theta|x_2) \\
\quad \propto p(\theta)L(\theta|x_1)L(\theta|x_2) \\
\quad = p(\theta)L(\theta|x_1, x_2) \\
(\text{since } x_1, x_2 \text{ independent})
\]

Note this is the same posterior we would have obtained had \(x_1\) and \(x_2\) arrived at the same time!

This “sequential updating” process can continue indefinitely in the Bayesian setup.
Why Normal Models?

- Why is it so common to model data using a normal distribution?
- Approximately normally distributed quantities appear often in nature.
- CLT tells us any variable that is basically a sum of independent components should be approximately normal.
- Note $\bar{X}$ and $S^2$ are independent when sampling from a normal population — so if beliefs about the mean are independent of beliefs about the variance, a normal model may be appropriate.
Why Normal Models?

- The normal model is analytically convenient (exponential family, sufficient statistics $\bar{X}$ and $S^2$)
- Inference about the population mean based on a normal model will be correct as $n \to \infty$ even if the data are truly non-normal.
- When we assume a normal likelihood, we can get a wide class of posterior distributions by using different priors.
A Conjugate analysis with Normal Data (variance known)

- Simple situation: Assume data $X_1, \ldots, X_n$ are iid $N(\mu, \sigma^2)$, with $\mu$ unknown and $\sigma^2$ known.
- We will make inference about $\mu$.
- The likelihood is

$$L(\mu|x) = \prod_{i=1}^{n} \left(2\pi\sigma^2\right)^{-1/2} e^{-\frac{1}{2\sigma^2}(x_i - \mu)^2}$$

- A conjugate prior for $\mu$ is $\mu \sim N(\delta, \tau^2)$:

$$p(\mu) = (2\pi\tau^2)^{-1/2} e^{-\frac{1}{2\tau^2}(\mu - \delta)^2}$$
A Conjugate analysis with Normal Data (variance known)

So the posterior is:

\[ \pi(\mu|\mathbf{x}) \propto L(\mu|\mathbf{x})p(\mu) \]
\[ \propto \prod_{i=1}^{n} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2} e^{-\frac{1}{2\tau^2}(\mu-\delta)^2} \]
\[ = \exp\left\{ -\frac{1}{2} \left[ \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2 + \frac{1}{\tau^2} (\mu - \delta)^2 \right] \right\} \]
\[ = \exp\left\{ -\frac{1}{2} \left[ \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i^2 - 2x_i\mu + \mu^2) + \frac{1}{\tau^2} (\mu^2 - 2\mu\delta + \delta^2) \right] \right\} \]
A Conjugate analysis with Normal Data (variance known)

So the posterior is:

\[ \pi(\mu|x) \propto \exp \left\{ -\frac{1}{2} \frac{1}{\sigma^2 \tau^2} \left( \tau^2 \sum x_i^2 - 2 \tau^2 \mu n\bar{x} + n\mu^2 \tau^2 ight. \\
+ \sigma^2 \mu^2 - 2 \sigma^2 \mu \delta + \sigma^2 \delta^2 \right) \right\} \\
= \exp \left\{ -\frac{1}{2} \frac{1}{\sigma^2 \tau^2} \left[ \mu^2 (\sigma^2 + n \tau^2) - 2 \mu (\delta \sigma^2 + \tau^2 n \bar{x}) \right. \\
+ \left( \delta^2 \sigma^2 + \tau^2 \sum x_i^2 \right) \right\} \\
\propto \exp \left\{ -\frac{1}{2} \left[ \mu^2 \left( \frac{1}{\tau^2} + \frac{n}{\sigma^2} \right) - 2 \mu \left( \frac{\delta}{\tau^2} + \frac{n\bar{x}}{\sigma^2} \right) + k \right] \right\} \\
(\text{where } k \text{ is some constant}) \]
Hence $\pi(\mu|x) \propto \exp\left\{-\frac{1}{2} \left[ \left( \frac{1}{\tau^2} + \frac{n}{\sigma^2} \right) (\mu^2 - 2\mu \left( \frac{\delta}{\tau^2} + \frac{n\bar{x}}{\sigma^2} \right) + k) \right]\right\}$

$\propto \exp\left\{-\frac{1}{2} \left[ \left( \frac{1}{\tau^2} + \frac{n}{\sigma^2} \right) \left( \mu - \frac{\delta}{\tau^2} + \frac{n\bar{x}}{\sigma^2} \right)^2 \right]\right\}$
Hence the posterior for $\mu$ is simply a normal distribution with mean

$$\frac{\delta}{\tau^2} + \frac{n\bar{x}}{\sigma^2}$$

and variance

$$\left(\frac{1}{\tau^2} + \frac{n}{\sigma^2}\right)^{-1} = \frac{\tau^2 \sigma^2}{\sigma^2 + n\tau^2}$$

The \textbf{precision} is the reciprocal of the \textbf{variance}.

Here, $\frac{1}{\tau^2}$ is the \textbf{prior precision} . . .

$\frac{n}{\sigma^2}$ is the \textbf{data precision} . . .

. . . and $\frac{1}{\tau^2} + \frac{n}{\sigma^2}$ is the \textbf{posterior precision}. 

...
A Conjugate analysis with Normal Data (variance known)

- Note the posterior mean $E[\mu|x]$ is simply
  \[
  \frac{1/\tau^2}{1/\tau^2 + n/\sigma^2} \delta + \frac{n/\sigma^2}{1/\tau^2 + n/\sigma^2} \bar{x},
  \]
  a combination of the **prior mean** and the **sample mean**.

- If the prior is highly precise, the weight is large on $\delta$.
- If the data are highly precise (e.g., when $n$ is large), the weight is large on $\bar{x}$.
- Clearly as $n \to \infty$, $E[\mu|x] \approx \bar{x}$, and $\text{var}[\mu|x] \approx \frac{\sigma^2}{n}$ if we choose a large prior variance $\tau^2$.
- This implies that for $\tau^2$ large and $n$ large, Bayesian and frequentist inference about $\mu$ will be nearly identical.