Gaussian-Gaussian Conjugate Priors

Donald E. Brown

School of Data Science
University of Virginia
Charlottesville, VA 22904

Gaussian with Unknown mean and known Variance

- Likelihood with N trials, $x = (x_1, \ldots, x_N)$ with unknown mean M and known variance σ^2

$$f(x|m, \sigma^2) \propto \frac{1}{\sigma} exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(x_i - m)^2\right)$$

- Prior for M is $N(\mu_0, \sigma_0^2)$
- Posterior is $N(\mu_N$ and $\sigma_{\text{post}}^2)$ where

$$\mu_N = \frac{\mu_0\sigma^2 + N\bar{x}\sigma_0^2}{\sigma^2 + N\sigma_0^2}$$

$$\sigma_{\text{post}}^2 = \frac{\sigma_0^2\sigma^2}{\sigma^2 + N\sigma_0^2}$$

Gaussian Posterior with Unknown mean and Known Variance

- Mean

$$\begin{aligned} E[M|x,\sigma^2] &= \mu_N \\ &= \frac{\sigma_0^2}{(\sigma_0^2+\sigma^2/N)}\bar{x} + \frac{\sigma^2/N}{(\sigma_0^2+\sigma^2/N)}\mu_0 \\ &\to \bar{x}, N \to \infty \end{aligned}$$

- Variance

$$\begin{aligned} Var[M|x,\sigma^2] &= \sigma^2_{\text{post}} \\ &= \frac{\sigma_0^2\sigma^2}{\sigma^2+N\sigma_0^2} \\ &\to 0, N \to \infty \end{aligned}$$

Using Precision

- Precision defined as inverse variance, so

$$\tau = \frac{1}{\sigma^2} \text{ and } \tau_0 = \frac{1}{\sigma_0^2}$$

- Mean

$$E[M|x,\tau] = \frac{N\tau}{(\tau_0 + N\tau)}\bar{x} + \frac{\tau_0}{(\tau_0 + N\tau)}\mu_0$$

- Precision

$$\tau_{post} = N\tau + \tau_0$$

Prediction of a Gaussian with Known Precision

- Let $\boldsymbol{x} = (x_1, \ldots, x_N)$

$$f(x_{new}|\boldsymbol{x}, \tau) = \int f(x_{new}|\boldsymbol{x}, m, \tau) f(m|\boldsymbol{x}, \tau) dm$$

- Mean

$$E[X_{new}] = \mu_N$$

- Variance

$$Var[X_{new}] = \sigma^2_{post} + \sigma^2$$

Priors for Gaussian Likelihood with Unknown Mean and Unknown Precisions

- Likelihood with N trials, $x = (x_1, \ldots, x_N)$ with unknown mean, M, and unknown precision, W

$$f(x|m,w) \propto w^{\frac{N}{2}} exp\left(-\frac{w}{2}\sum_{i=1}^{N}(x_i - m)^2\right)$$

- Prior for M given $W = w$ is $N(\mu_0, \frac{1}{\tau_0 w}), \tau_0 > 0$
- τ_0 is a scaling parameter, roughly the number of “observations” in the prior
- Prior for W is gamma $(\frac{\alpha_0}{2}, \frac{\beta_0}{2})$

Marginal Prior Distribution of M: t Distribution

- Prior for M found by integrating the joint prior

$$
\begin{aligned}
g(m) &= \int_0^\infty f(m,w)dw \\
&\propto \int_0^\infty w^{\frac{1}{2}} e^{-\frac{\tau_0 w}{2}(m-\mu_0)^2} w^{\alpha_0 - 1} e^{-\beta_0 w} dw \\
&\propto \left[\alpha_0 + \frac{(m-\mu_0)^2}{\frac{\beta_0}{\alpha_0 \tau_0}}\right]^{-(\alpha_0+1)/2}
\end{aligned}
$$

- Marginal prior of M is a t distribution with $\underline{\alpha_0}$ d.o.f., location parameter $= \mu_0$, and scale $= \sqrt{\frac{\beta_0}{\alpha_0 \tau_0}}$
- Since d.o.f. $= \alpha_0$, set $\alpha_0 = \tau_0 - 1$, so $\tau_0 \geq 2$

Posterior Distributions for Gaussian Likelihood with Unknown Mean and Unknown Precision

- Conditional posterior of M given $W = w$ and N observations, $\boldsymbol{X} = \boldsymbol{x}$, is $N(\mu_N, \frac{1}{\tau_N w})$ where $\tau_N = \tau_0 + N$ and

$$\mu_N = \frac{\tau_0 \mu_0 + N\bar{x}}{\tau_0 + N}$$

- Posterior of the precision, W given $\boldsymbol{x}$ is gamma with parameters, α_N and β_N

$$\alpha_N = \alpha_0 + N$$

$$\beta_N = \beta_0 + \sum_{i=1}^{N}(x_i - \bar{x})^2 + \frac{\tau_0 N(\bar{x} - \mu_0)^2}{(\tau_0 + N)}$$

- Marginal posterior for M is a t distribution with α_N d.o.f, location parameter $= \mu_N$, and scale $= \sqrt{\frac{\beta_N}{\alpha_N \tau_N}}$

Examples of Gaussian Posteriors: Unknown σ^2

Priors: $N(20, 1/4W), G(5, 200)$;
$N(30, 1/9W), G(10, 450)$;
$N(32, 1/9W), G(10, 1800)$