STAT 578: Advanced Bayesian Modeling

Week 2 – Lesson 2

Two-Parameter Normal Sample

Fall 2019

Conjugate Prior Analysis

Normal Sampling Distribution

$n > 1$ observations of a $\mathrm{N}(\mu, \sigma^2)$ variable:

$$y = (y_1, \ldots, y_n)$$

$$y_1, \ldots, y_n \mid \mu, \sigma^2 \sim \text{iid } \mathrm{N}(\mu, \sigma^2)$$

Realistically, both μ and σ^2 are unknown:

$$\theta = (\mu, \sigma^2) \qquad\qquad \sigma^2 > 0$$

Recall likelihood:

$$p(y \mid \mu, \sigma^2) \quad \propto \quad \frac{1}{(\sigma^2)^{n/2}} \, \exp\left(-\frac{(n-1)s^2}{2\sigma^2}\right) \, \exp\left(-\frac{n}{2\sigma^2}(\mu - \bar{y})^2\right)$$

where

$$\bar{y} \quad = \quad \frac{1}{n}\sum_{i=1}^{n} y_i \qquad\qquad s^2 \quad = \quad \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$$

Which factors contain σ^2? $\quad \mu$? $\quad$ Both?

Conjugate Prior

Strategy: Find (joint) prior density of form similar to likelihood.

Anticipating later results, for $\sigma^2 > 0$,

$$p(\mu, \sigma^2) \quad \propto \quad \frac{1}{(\sigma^2)^{(\nu_0+3)/2}} \exp\left(-\frac{\nu_0 \sigma_0^2}{2\sigma^2}\right) \exp\left(-\frac{\kappa_0}{2\sigma^2}(\mu - \mu_0)^2\right)$$

for some constants $\nu_0 > 0$, $\sigma_0^2 > 0$, $\kappa_0 > 0$, μ_0.

Note:

- μ and σ^2 are dependent under this prior.
- Conditional prior on μ given σ^2 is $\mathrm{N}(\mu_0,\ \sigma^2/\kappa_0)$.

Convenient to integrate out μ:

$$\begin{aligned} p(\sigma^2) &= \int p(\mu, \sigma^2)\, d\mu \\ &\propto \frac{1}{(\sigma^2)^{(\nu_0+3)/2}} \exp\left(-\frac{\nu_0 \sigma_0^2}{2\sigma^2}\right) \int \exp\left(-\frac{\kappa_0}{2\sigma^2}(\mu - \mu_0)^2\right) d\mu \\ &\propto \frac{1}{(\sigma^2)^{(\nu_0+3)/2}} \exp\left(-\frac{\nu_0 \sigma_0^2}{2\sigma^2}\right) \sqrt{2\pi\sigma^2/\kappa_0} \\ &\propto \frac{1}{(\sigma^2)^{\nu_0/2+1}} \exp\left(-\frac{\nu_0 \sigma_0^2}{2\sigma^2}\right) \qquad\qquad \sigma^2 > 0 \end{aligned}$$

Density of what distribution?

Inverse Chi-Square

Prior distribution for σ^2 is **scaled inverse chi-square** (BDA3, Table A.1):

$$\sigma^2 \;\sim\; \text{Inv-}\chi^2(\nu_0,\, \sigma_0^2)$$

Name refers to

$$\nu_0\sigma_0^2/\sigma^2 \;\sim\; \chi^2_{\nu_0}$$

Also called **inverse gamma**: $\quad \text{Inv-gamma}(\nu_0/2,\, \nu_0\sigma_0^2/2)$

The full prior is then

$$\begin{aligned} \mu \mid \sigma^2 &\sim \mathrm{N}(\mu_0,\, \sigma^2/\kappa_0) \\ \sigma^2 &\sim \text{Inv-}\chi^2(\nu_0,\, \sigma_0^2) \end{aligned}$$

Now use Bayes' rule

$$\begin{aligned} p(\mu, \sigma^2 \mid y) &\propto p(\mu, \sigma^2)\, p(y \mid \mu, \sigma^2) \\ &\propto p(\sigma^2)\, p(\mu \mid \sigma^2)\, p(y \mid \mu, \sigma^2) \end{aligned}$$

and some algebra (BDA3, Sec. 3.3) to find that ...

... the posterior is

$$\begin{aligned} \mu \mid \sigma^2, y &\sim \mathrm{N}(\mu_n,\, \sigma^2/\kappa_n) \\ \sigma^2 \mid y &\sim \text{Inv-}\chi^2(\nu_n,\, \sigma_n^2) \end{aligned}$$

where

$$\mu_n = \frac{\kappa_0}{\kappa_0+n}\mu_0 + \frac{n}{\kappa_0+n}\bar{y} \qquad \kappa_n = \kappa_0 + n \qquad \nu_n = \nu_0 + n$$

$$\nu_n\sigma_n^2 = \nu_0\sigma_0^2 + (n-1)s^2 + \frac{\kappa_0 n}{\kappa_0+n}(\bar{y}-\mu_0)^2$$

So the prior specification is conjugate (with joint distribution sometimes called **normal-inverse-gamma**).

Suggests interpretation of constants in the prior:

- $\mu_0 \;=\;$ prior guess for μ
- $\sigma_0^2 \;=\;$ prior guess for σ^2
- $\kappa_0 \;=\;$ prior certainty about μ (equivalent number of "prior observations")
- $\nu_0 \;=\;$ prior certainty about σ^2

Example: Flint Data

Recall:

$$y_i \quad = \quad \textit{logarithm} \text{ of first-draw lead level (ppb), for observation } i$$

```
> (n <- nrow(Flintdata))
[1] 271

> (ybar <- mean(log(Flintdata$FirstDraw)))
[1] 1.402925

> (s.2 <- var(log(Flintdata$FirstDraw)))
[1] 1.684078
```

So

$$n \;=\; 271 \qquad\qquad \bar{y} \;\approx\; 1.40 \qquad\qquad s^2 \;\approx\; 1.684$$

As before, choose prior mean

$$\mu_0 \;=\; \log(3) \;\approx\; 1.10$$

Choose

$$\sigma_0^2 \;=\; 1.17$$

(With μ_0, this puts the 90th percentile at about 12 ppb, its official estimate.)

Not very certain about these prior estimates, therefore let

$$\kappa_0 \;=\; \nu_0 \;=\; 1$$

Compute posterior:

```
> mu0 <- log(3)
> sigma.2.0 <- 1.17
> kappa0 <- 1
> nu0 <- 1

> (mun <- (kappa0*mu0 + n*ybar) / (kappa0 + n))
[1] 1.401807

> (kappan <- kappa0 + n)
[1] 272

> (nun <- nu0 + n)
[1] 272

> (sigma.2.n <- (nu0*sigma.2.0 + (n-1)*s.2 +
+                kappa0*n*(ybar-mu0)^2/(kappa0+n))/nun)
[1] 1.676336
```

Simulation

There are formulas for posterior means and variances (BDA3, Table A.1), but simulation from posterior is easier ...

First,

$$\nu_n \sigma_n^2 / \sigma^2 \,\Big|\, y \;\sim\; \chi^2_{\nu_n}$$

so simulate X from $\chi^2_{\nu_n}$ and let

$$\sigma^2_{\textsf{sim}} \;=\; \nu_n \sigma_n^2 / X$$

Then simulate

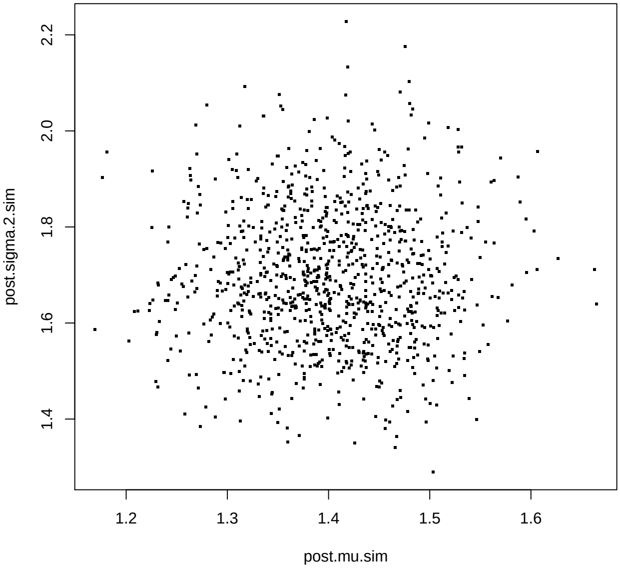
$$\mu_{\textsf{sim}} \;\text{ from }\; \mathrm{N}(\mu_n,\, \sigma^2_{\textsf{sim}}/\kappa_n)$$

In R, simulating 1000 times:

```
> post.sigma.2.sim <- nun * sigma.2.n / rchisq(1000, nun)

> post.mu.sim <- rnorm(1000, mun, sqrt(post.sigma.2.sim / kappan))

> summary(post.sigma.2.sim)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  1.290   1.595   1.686   1.696   1.791   2.227

> summary(post.mu.sim)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  1.169   1.350   1.401   1.402   1.456   1.665
```

Plot the two-dimensional posterior samples $(\mu_{\mathsf{sim}}, \sigma^2_{\mathsf{sim}})$:

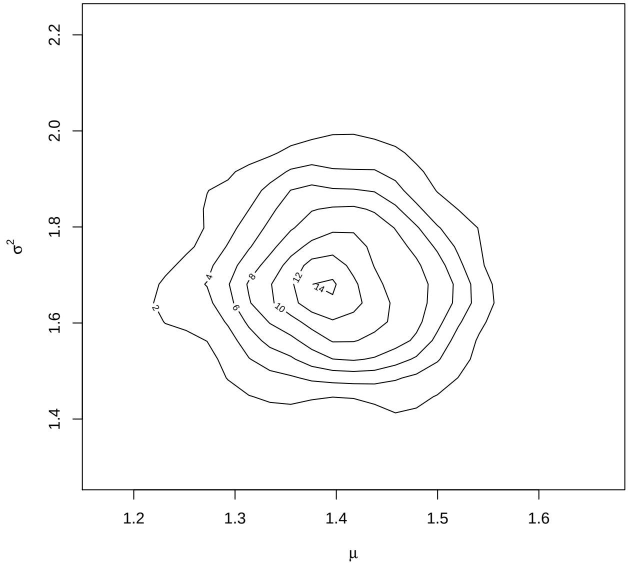
```
> plot(post.mu.sim, post.sigma.2.sim, pch=".", cex=3)
```


post.sigma.2.sim
post.mu.sim
1.4
1.6
1.8
2.0
2.2
1.2
1.3
1.4
1.5
1.6

Approximating the posterior joint density with a contour plot:

```
> library(MASS)  # provides kde2d

> contour(kde2d(post.mu.sim, post.sigma.2.sim),
          xlab=expression(mu), ylab=expression(sigma^2))
```

(Would look better with more samples)



2.2
2.0
1.8
1.6
1.4
σ^2
1.2
1.3
1.4
1.5
1.6
μ
2
4
6
8
10
12
14