

STAT 578: Advanced Bayesian Modeling

## Week 2 – Lesson 2

# Two-Parameter Normal Sample

Fall 2019

# Noninformative Prior Analysis

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

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

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

with both  $\mu$  and  $\sigma^2$  unknown:

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

## A Noninformative Mean Prior

Recall: For the mean-only normal sample, flat prior

$$p(\mu) \propto 1 \quad -\infty < \mu < \infty$$

did not define an actual distribution (because improper), but still formally produced a proper posterior.

Can regard as result of using a normal prior distribution and letting prior variance tend to infinity.

Note: Invariant to location shift and (up to proportionality) change in scale

## A Noninformative Variance Prior

Recall (partially) conjugate  $\text{Inv-}\chi^2(\nu_0, \sigma_0^2)$  prior for  $\sigma^2$  (BDA3, Table A.1):

$$p(\sigma^2) \propto (\sigma^2)^{-(\nu_0/2+1)} e^{-\nu_0\sigma_0^2/(2\sigma^2)} \quad \sigma^2 > 0$$

Recall  $\nu_0$  represents prior level of certainty about  $\sigma^2$ .

Formally letting  $\nu_0$  go to zero suggests

$$p(\sigma^2) \propto (\sigma^2)^{-1} \quad \sigma^2 > 0$$

The “density”

$$p(\sigma^2) \propto (\sigma^2)^{-1} \quad \sigma^2 > 0$$

is improper – infinite area near zero and infinity.

May still be useful if posterior is proper.

Remark: Formal transformation of variables shows equivalence to flat prior on log scale,

$$p(\log \sigma^2) \propto 1$$

hence scale invariance.

## A Noninformative Joint Prior

Treating  $\mu$  and  $\sigma^2$  as if independent under the prior gives

$$p(\mu, \sigma^2) \propto p(\mu) p(\sigma^2) \propto (\sigma^2)^{-1} \quad \sigma^2 > 0$$

Can regard as a flat prior on  $(\mu, \log \sigma^2)$ .

Formally related to conjugate prior: Choose

$$\nu_0 = -1 \quad \sigma_0^2 = 0 \quad \kappa_0 = 0$$

Formally applying Bayes' rule,

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

which implies (similarly to conjugate case)

$$\begin{aligned} \mu \mid \sigma^2, y &\sim N(\bar{y}, \sigma^2/n) \\ \sigma^2 \mid y &\sim \text{Inv-}\chi^2(n-1, s^2) \end{aligned}$$

So posterior is proper (when  $n > 1$  and  $s^2 > 0$ ).



# Simulation

First, simulate  $X$  from  $\chi_{n-1}^2$  and let

$$\sigma_{\text{sim}}^2 = (n-1)s^2/X$$

Then simulate

$$\mu_{\text{sim}} \text{ from } N(\bar{y}, \sigma_{\text{sim}}^2/n)$$

# Prediction

Consider new single observation  $\tilde{y}$  from the sampling distribution, but (conditionally) independent of the data:

$$\tilde{y} \mid \mu, \sigma^2, y \sim N(\mu, \sigma^2)$$

(Note: Doesn't depend on  $y$ )

Its marginal posterior distribution is a posterior predictive distribution – useful for assessing aspects of the sampling distribution (e.g., quantiles).

Easily simulated:

$$\tilde{y}_{\text{sim}} \text{ from } N(\mu_{\text{sim}}, \sigma_{\text{sim}}^2)$$

## Example: Flint Data

Explicit formulas are possible, but let's use simulation ...

```
> post.sigma.2.sim <- (n-1) * s.2 / rchisq(1000, n-1)

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

> quantile(post.mu.sim, c(0.025,0.975)) # 95% post. int. for mu
      2.5%      97.5%
1.246809 1.554445

> quantile(post.sigma.2.sim, c(0.025,0.975)) # 95% post. int. for sigma.2
      2.5%      97.5%
1.433607 2.025813
```

(Remember these are for data on the log scale.)

Consider a hypothetical new observation sampled from the same population (households in Flint).

Let's simulate from the posterior predictive distribution of log lead level from the first draw ...

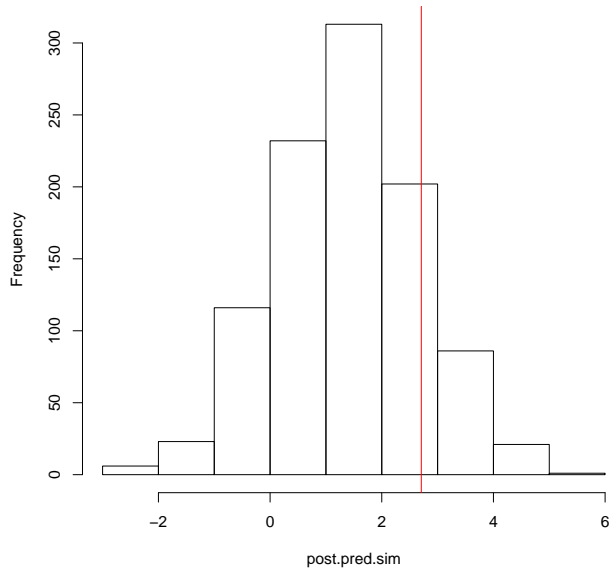
```
> post.pred.sim <- rnorm(1000, post.mu.sim, sqrt(post.sigma.2.sim))
```

... and display as a histogram (with line at 15 ppb level):

```
> hist(post.pred.sim)
> abline(v=log(15), col="red")
```

(May be a bit wider than data because of uncertainty in parameter values)

**Histogram of post.pred.sim**



Recall: By law, no more than 10% of households sampled can exceed action level of 15 ppb.

Approximate posterior probability that level is exceeded for a random household:

```
> mean(post.pred.sim > log(15))  
[1] 0.153
```

Compare with fraction in original sample exceeding level:

```
> mean(Flintdata$FirstDraw > 15)  
[1] 0.1660517
```

Warning:

What if sampling distribution is incorrectly specified?

Inference about some parameters may remain acceptable if sample size is large enough (central limit theorem).

But predictive distributions can be misleading!