ADVANCED BAYESIAN MODELING

Noninformative Prior Analysis

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

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

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

with both μ and $\sigma^2 > 0$ unknown:

$$\theta = (\mu, \sigma^2)$$

A Noninformative Mean Prior

For the mean-only normal sample, recall flat prior

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

Not an actual distribution (improper), but still formally produced a proper posterior in mean-only case.

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 $\mathrm{Inv}\text{-}\chi^2(\nu_0,\, \sigma_0^2)$ prior for σ^2 (BDA3, Table A.1):

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

Recall ν_0 represents prior level of certainty about σ^2.

Formally letting ν_0 go to zero suggests

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

The “density”

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

is improper – infinite area near zero and infinity.

May still be useful if posterior is proper.

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

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

hence scale invariance.

A Noninformative Joint Prior

Treating μ and σ^2 as if independent under the prior gives

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

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

Formally related to conjugate prior: Choose

$$\nu_0 \;=\; -1 \qquad \sigma_0^2 \;=\; 0 \qquad \kappa_0 \;=\; 0$$

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)

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

So posterior is proper (when $n > 1$).

Simulation

First, simulate X from χ^2_{n-1} and let

$$\sigma^2_{\mathsf{sim}} = (n-1)s^2 \big/ X$$

Then simulate

$$\mu_{\mathsf{sim}} \text{ from } \mathrm{N}(\bar{y},\, \sigma^2_{\mathsf{sim}}/n)$$

Prediction

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

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

(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:

$$\tilde{y}_{\mathsf{sim}} \;\; \text{from} \;\; \mathrm{N}(\mu_{\mathsf{sim}},\, \sigma^2_{\mathsf{sim}})$$

Example: Flint Data

Explicit formulas are available, 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 on the log scale.)

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

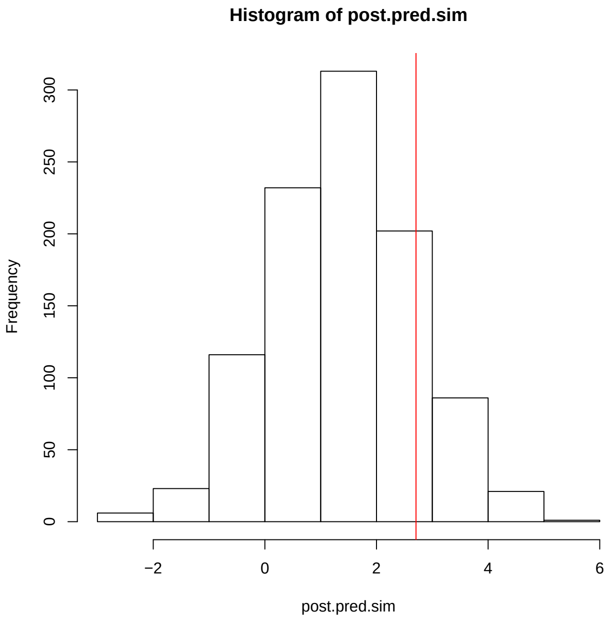
Let's simulate from the posterior predictive distribution of log lead level from the first draw ...

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

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

```
> 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
Frequency
0
50
100
150
200
250
300
−2
0
2
4
6
post.pred.sim

Approximate posterior probability that limit is exceeded:

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

Recall: By law, no more than 10% of households sampled can exceed limit.

Compare with fraction in original sample that exceed limit:

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

Warning:

What if sampling distribution is incorrectly specified?

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

But predictive distributions can be misleading!