Prior sensitivity analysis

When communicating results from any analysis, a responsible statistician will report and justify modeling decisions, especially assumptions. In a Bayesian analysis, there is another assumption that is open to scrutiny: the choices of prior distributions. In the models considered so far in this course, there are an infinite number of prior distributions we could have chosen from.

How do you justify the priors you choose? If they truly represent your beliefs about the parameters before analysis and the model is appropriate, then the posterior distribution truly represents your updated beliefs. If you don't have any strong beliefs beforehand, there are often default, reference, or noninformative prior options, and you will have to select one. However, a collaborator or a boss (indeed, somebody somewhere) may not agree with your choice of prior. One way to increase the credibility of your results is to repeat the analysis under a variety of priors, and report how the results differ as a result. This process is called *prior sensitivity analyis*.

At a minimum you should always report your choice of model and prior. If you include a sensitivity analysis, select one or more alternative priors and describe how the results of the analysis change. If they are sensitive to the choice of prior, you will likely have to explain both sets of results, or at least explain why you favor one prior over another. If the results are not sensitive to the choice of prior, this is evidence that the data are strongly driving the results. It suggests that different investigators coming from different backgrounds should come to the same conclusions.

If the purpose of your analysis is to establish a hypothesis, it is often prudent to include a ``skeptical" prior which does not favor the hypothesis. Then, if the posterior distribution still favors the hypothesis despite the unfavorable prior, you will be able to say that the data substantially favor the hypothesis. This is the approach we will take in the following example, continued from the previous lesson.

Poisson regression example

Let's return to the example of number of doctor visits. We concluded from our previous analysis of these data that both bad health and increased age are associated with more visits. Suppose the burden of proof that bad health is actually associated with more visits rests with us, and we need to convince a skeptic.

First, let's re-run the original analysis and remind ourselves of the posterior distribution for the `badh` (bad health) indicator.

```
library("COUNT")
```

```
## Loading required package: msme
```

```
## Loading required package: MASS
```

```
## Loading required package: lattice
```

```
## Loading required package: sandwich
```

```
library("rjags")
```

```
## Loading required package: coda
```

```
## Linked to JAGS 4.2.0
```

```
## Loaded modules: basemod,bugs
```

```
data("badhealth")

mod_string = " model {
    for (i in 1:length(numvisit)) {
        numvisit[i] ~ dpois(lam[i])
        log(lam[i]) = int + b_badh*badh[i] + b_age*age[i] + b_intx*age[i]*badh[i]
    }

    int ~ dnorm(0.0, 1.0/1e6)
    b_badh ~ dnorm(0.0, 1.0/1e4)
    b_age ~ dnorm(0.0, 1.0/1e4)
    b_intx ~ dnorm(0.0, 1.0/1e4)
} "

set.seed(102)

data_jags = as.list(badhealth)

params = c("int", "b_badh", "b_age", "b_intx")

mod = jags.model(textConnection(mod_string), data=data_jags, n.chains=3)
update(mod, 1e3)

mod_sim = coda.samples(model=mod,
                       variable.names=params,
                       n.iter=5e3)
mod_csim = as.mcmc(do.call(rbind, mod_sim))
```
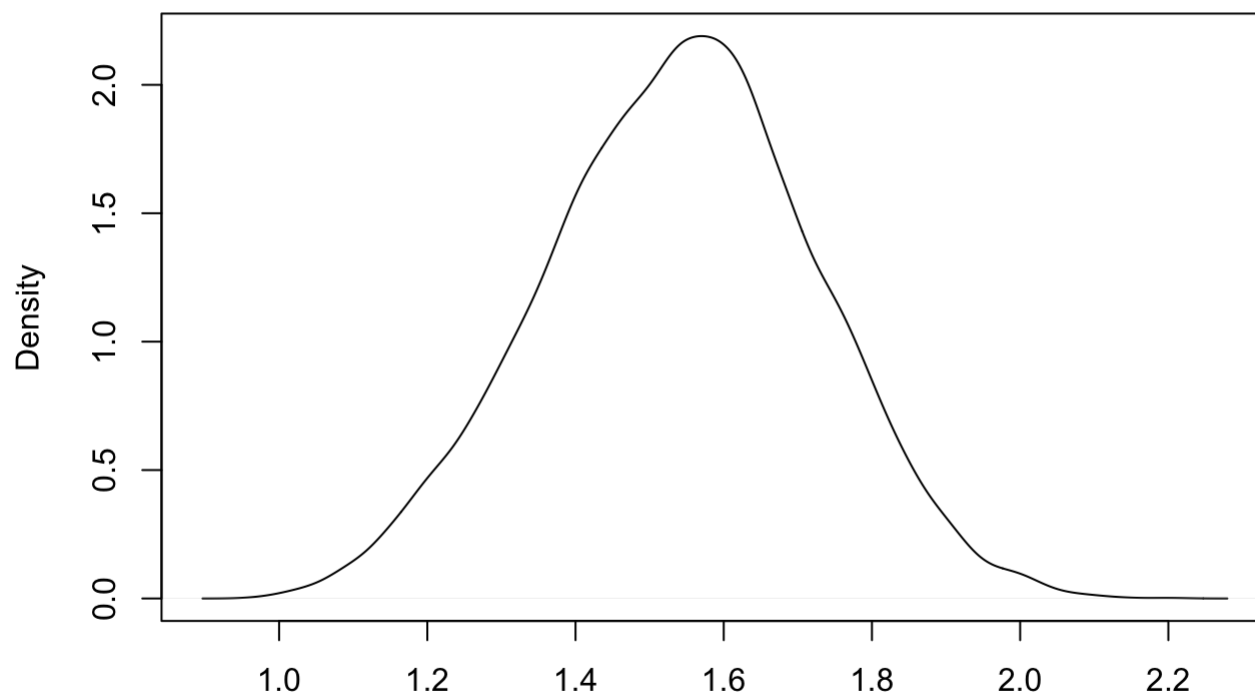
```
plot(density(mod_csim[,"b_badh"]))
```

density.default(x = mod_csim[, "b_badh"])

N = 15000   Bandwidth = 0.02404

Essentially all of the posterior probability mass is above 0, suggesting that this coefficient is positive (and consequently that bad health is associated with more visits). We obtained this result using a relatively noninformative prior. What if we use a prior that strongly favors values near 0? Let's repeat the analysis with a normal prior on the `badh` coefficient that has mean 0 and standard deviation 0.2, so that the prior probability that the coefficient is less than 0.6 is > 0.998. We'll also use a small variance on the prior for the interaction term involving `badh` (standard deviation 0.01 because this coefficient is on a much smaller scale).

```
mod2_string = " model {
    for (i in 1:length(numvisit)) {
        numvisit[i] ~ dpois(lam[i])
        log(lam[i]) = int + b_badh*badh[i] + b_age*age[i] + b_intx*age[i]*badh[i]
    }

    int ~ dnorm(0.0, 1.0/1e6)
    b_badh ~ dnorm(0.0, 1.0/0.2^2)
    b_age ~ dnorm(0.0, 1.0/1e4)
    b_intx ~ dnorm(0.0, 1.0/0.01^2)
} "

mod2 = jags.model(textConnection(mod2_string), data=data_jags, n.chains=3)
update(mod2, 1e3)

mod2_sim = coda.samples(model=mod2,
                        variable.names=params,
                        n.iter=5e3)
mod2_csim = as.mcmc(do.call(rbind, mod2_sim))
```

How did the posterior distribution for the coefficient of `badh` change?

```
curve(dnorm(x, mean=0.0, sd=sqrt(1e4)), from=-3.0, to=3.0, ylim=c(0.0, 3.0), lty=2,
      main="b_badh", ylab="density", xlab="b_badh")
curve(dnorm(x, mean=0.0, sd=0.2), from=-3.0, to=3.0, col="red", lty=2, add=TRUE)
lines(density(mod_csim[,"b_badh"]))
lines(density(mod2_csim[,"b_badh"]), col="red")
legend("topleft", legend=c("noninformative prior", "posterior", "skeptical prior", "posterior"),
       lty=c(2,1,2,1), col=rep(c("black", "red"), each=2), bty="n")
```

b_badh
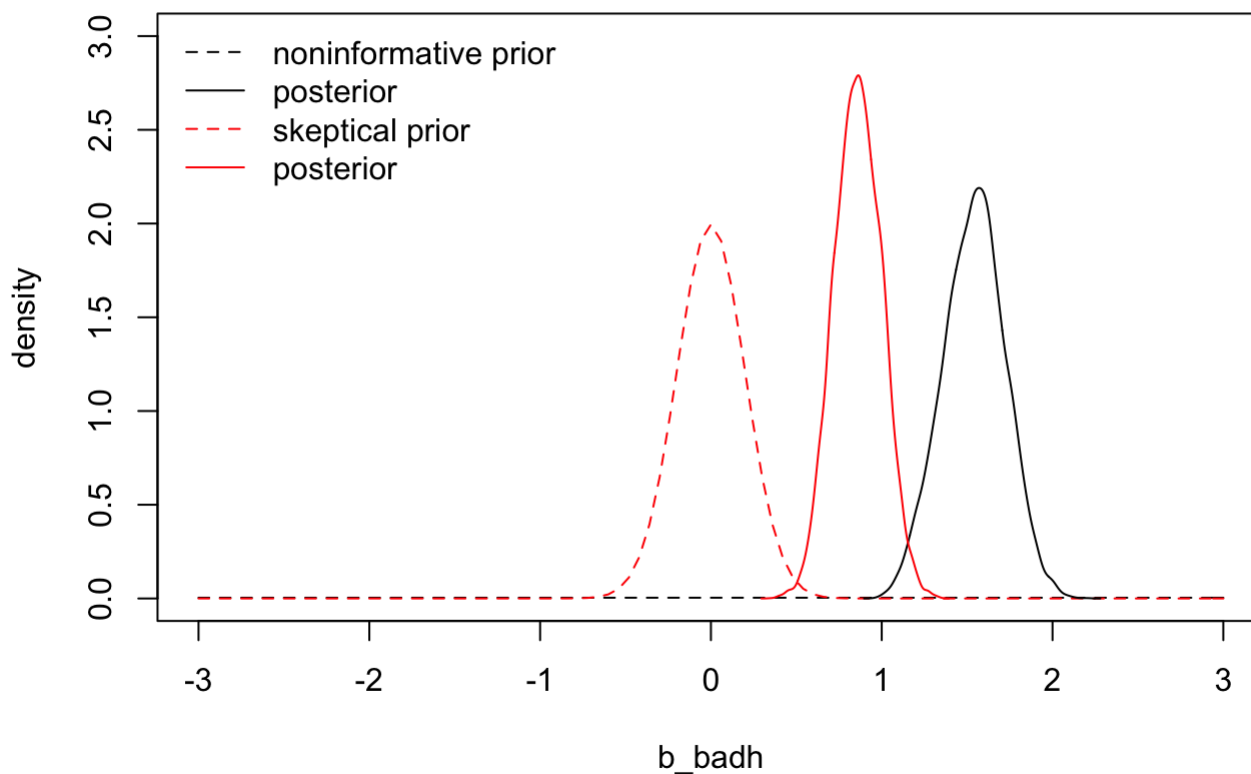
Under the skeptical prior, our posterior distribution for `b_badh` has significantly dropped to between about 0.6 and 1.1. Although the strong prior influenced our inference on the magnitude of the bad health effect on visits, it did not change the fact that the coefficient is significantly above 0. In other words: even under the skeptical prior, bad health is associated with more visits, with posterior probability near 1.

We should also check the effect of our skeptical prior on the interaction term involving both age and health.

```
curve(dnorm(x, mean=0.0, sd=sqrt(1e4)), from=-0.05, to=0.05, ylim=c(0.0, 140.0), lty=2,
      main="b_intx", ylab="density", xlab="b_intx")
curve(dnorm(x, mean=0.0, sd=0.01), from=-0.05, to=0.05, col="red", lty=2, add=TRUE)
lines(density(mod_csim[,"b_intx"]))
lines(density(mod2_csim[,"b_intx"]), col="red")
legend("topleft", legend=c("noninformative prior", "posterior", "skeptical prior", "posterior"),
       lty=c(2,1,2,1), col=rep(c("black", "red"), each=2), bty="n")
```
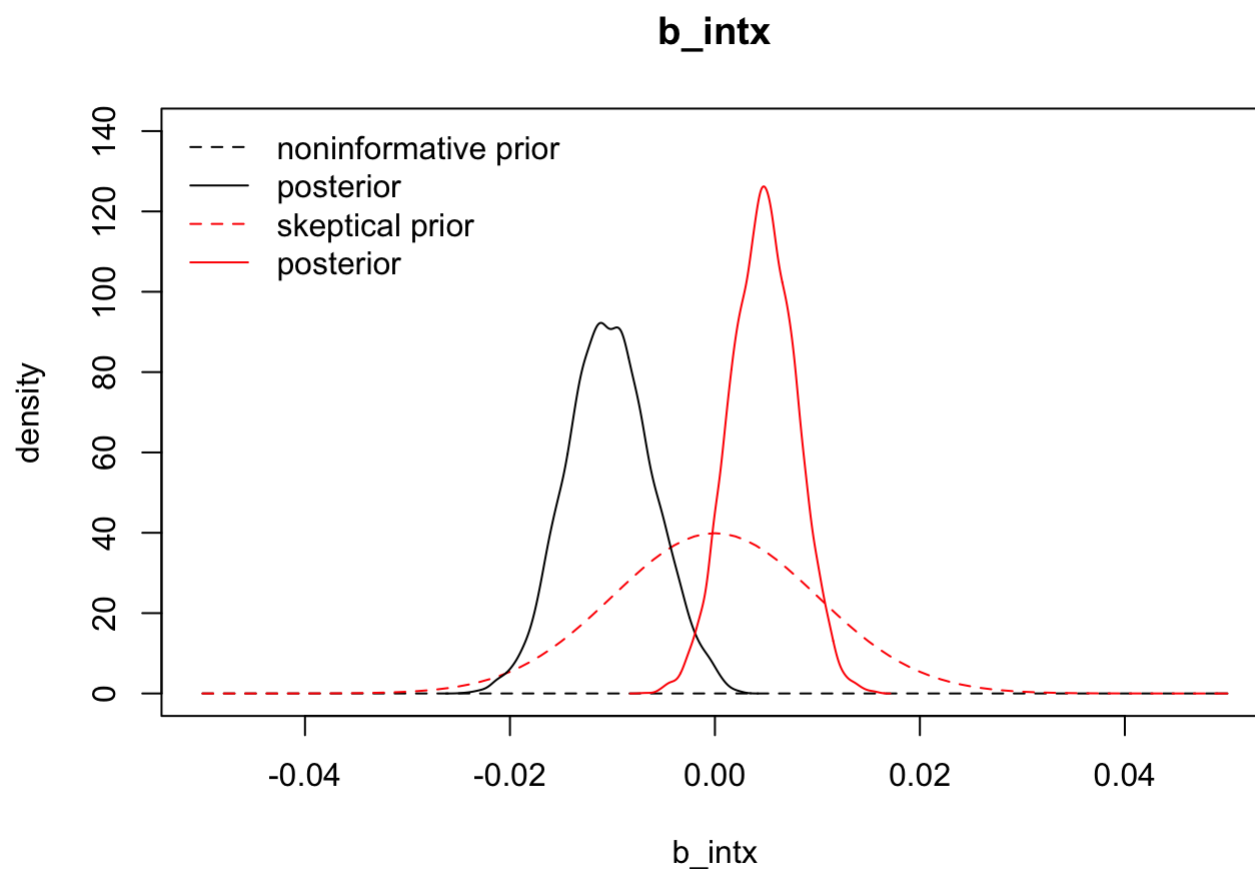
b_intx


```
mean(mod2_csim[,"b_intx"] > 0) # posterior probability that b_intx is positive
```

```
## [1] 0.9290667
```

The result here is interesting. Our estimate for the interaction coefficient has gone from negative under the noninformative prior to positive under the skeptical prior, so the result is sensitive. In this case, because the skeptical prior shrinks away much of the bad health main effect, it is likely that this interaction effect attempts to restore some of the positive effect of bad health on visits. Thus, despite some observed prior sensitivity, our conclusion that bad health positively associates with more visits remains unchanged.