Bayesian Statistics
Comparing Two Proportions or Means

Michael Anderson, PhD
Hélène Carabin, DVM, PhD

Department of Biostatistics and Epidemiology
The University of Oklahoma Health Sciences Center

May 20, 2016

Outline

One vs. Two Binomial Proportions

- Previously we have looked at estimating a single proportion.
- In many cases, however, we will be interested in comparing proportions from two or more groups.
- Estimation of the proportions will still be important but...
- Inference will typically focus on differences or ratios of these proportions.

Two Binomial Proportions

- Let θ_1 and θ_2 represent the proportions of two populations.
- We must first specify our prior beliefs about these proportions $p(\theta_1, \theta_2)$.
-

$$p(\theta_1, \theta_2 | x_1, \ldots, x_n) = p(x_1, \ldots, x_n | \theta_1, \theta_2) p(\theta_1, \theta_2) / p(x_1, \ldots, x_n)$$

$$= p(x_1, \ldots, x_n | \theta_1, \theta_2) p(\theta_1, \theta_2) / \int \int p(x_1, \ldots, x_n | \theta_1, \theta_2) p(\theta_1, \theta_2) d\theta_1 d\theta_2$$

- If θ_1 and θ_2 are independent then $p(\theta_1, \theta_2) = p(\theta_1) p(\theta_2)$.
- So we can specify priors for θ_1 and θ_2 separately.
- Because the observations are independent, the likelihood for both groups can be specified separately.

Example

In a clinical trial 100 subjects were randomized to drug A (standard+placebo) and 100 to drug B (standard+treatment). For drug A, 60 subjects experienced an effect of interest whereas 75 subjects experienced this effect for drug B. How does the standard+treatment compare to the standard+placebo in regards to the effect of interest? Using non informative priors obtain 95% BCI for the ratio of these proportions.

- Obtain a Bayesian p-value for the test: $H_0 : \pi_B/\pi_A \leq 1$ vs. $H_1 : \pi_B/\pi_A > 1$.
- Assume a ROPE of .9 to 1.1. Obtain a Bayesian p-value for the test: $H_0 : \pi_B/\pi_A = 1$ vs. $H_1 : \pi_B/\pi_A \neq 1$.
- See file "Drug Trial pvalue.odc"

In Class Practice Problem

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
propA	0.5984	0.04869	5.508E-4	0.501	0.5991	0.6916	1001	10000
propB	0.7456	0.04342	3.695E-4	0.6569	0.7474	0.8251	1001	10000
pval	0.0135	0.1154	0.001218	0.0	0.0	0.0	1001	10000
rope.area	0.0989	0.2985	0.002933	0.0	0.0	1.0	1001	10000
rrBA	1.254	0.1272	0.00131	1.03	1.245	1.529	1001	10000

In Class Practice Problem

Node statistics

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
propA	0.5984	0.04869	5.508E-4	0.501	0.5991	0.6916	1001	10000
propB	0.7456	0.04342	3.695E-4	0.6569	0.7474	0.8251	1001	10000
pval	0.0135	0.1154	0.001218	0.0	0.0	0.0	1001	10000
ropp_area	0.0989	0.2985	0.002933	0.0	0.0	1.0	1001	10000
rrBA	1.254	0.1272	0.00131	1.03	1.245	1.529	1001	10000

In Class Practice Problem

| node	mean	sd	MC error	2.5%	median	97.5%	start	sample
propA	0.5984	0.04869	5.508E-4	0.501	0.5991	0.6916	1001	10000
propB	0.7456	0.04342	3.695E-4	0.6569	0.7474	0.8251	1001	10000
pval	0.0135	0.1154	0.001218	0.0	0.0	0.0	1001	10000
rope.area	0.0989	0.2985	0.002933	0.0	0.0	1.0	1001	10000
rrBA	1.254	0.1272	0.0013	1.03	1.245	1.529	1001	10000
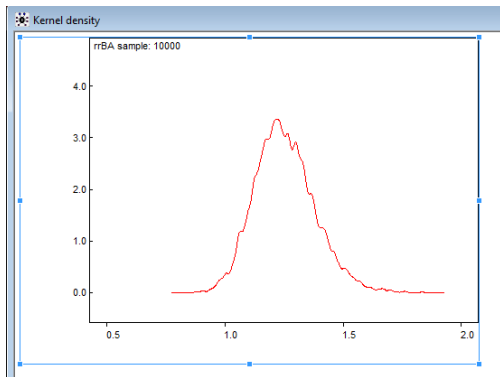
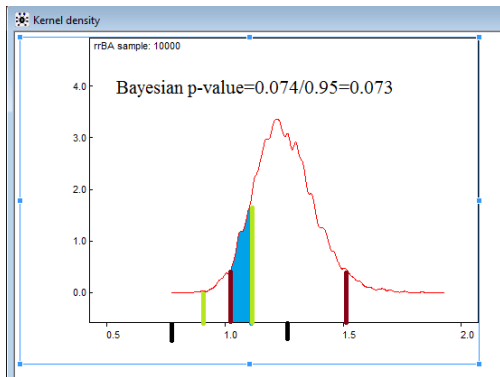# In Class Practice Problem

In Class Practice Problem
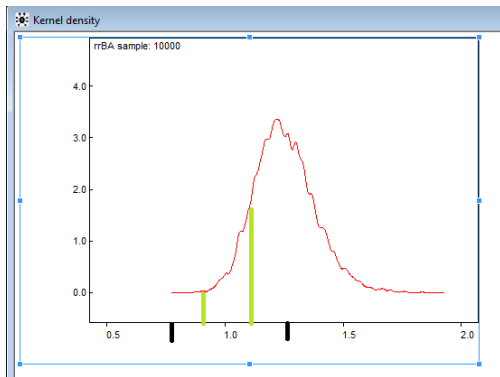
In Class Practice Problem
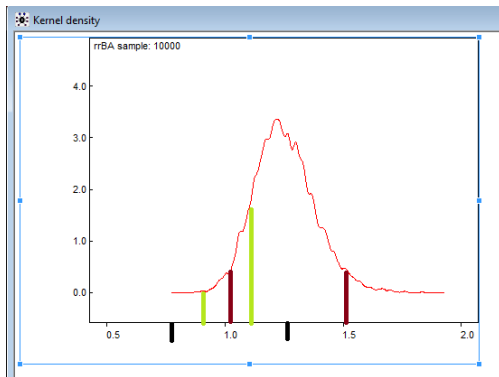
In Class Practice Problem
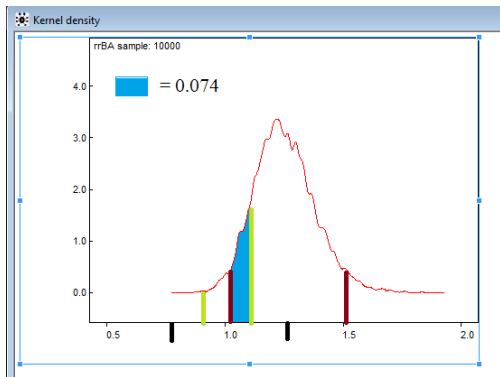
In Class Practice Problem

Comparing Two Means

Recall that the estimates for continuous data usually include a mean as well as a variance.

To compare two means from two populations, we will need to obtain

- likelihoods (two of them) for the two populations.
- priors (four of them) for the two means AND the two variances.

These additional items are easily incorportade in WinBUGS as illustrated in the following example.

Comparing Two Means

Suppose a physician prescribes an exercise regimen in addition to the blood pressure medication and records the systolic blood pressure (SBP) for 14 subjects after 3 months. Let's compare the mean SBP for medication only (TMT A) to the mean SBP for exercise + medication (TMT B).

```
list(N1=19,N2=14,a.tmt=c(121,94,119,122,142,168,116,172,155,
        107,180,119,157,101,145,148,120,147,125),
          b.tmt=c(126,125,130,130,122,118,118,
              111,123,126,127,111,112,121))
```

As before, we we could reasonably assume a Normal likelihood for SBP measures for both treatments.
Here we must estimate the means and variances for both groups separately. Thus we put diffuse priors on the means of both groups and diffuse priors on the precisions of both groups.

Specification of the Likelihood and Prior For Two Means

Putting this altogether we have:

- $p(a_1, \ldots, a_{19}|\mu_1, \sigma_1^2) \sim dnorm(\mu_1, \tau_1)$
- $p(b_1, \ldots, b_{14}|\mu_2, \sigma_2^2) \sim dnorm(\mu_2, \tau_2)$
- $p(\mu_1) \sim dnorm(0, 0.0001)$.
- $p(\mu_2) \sim dnorm(0, 0.0001)$.
- $p(\tau_1) \sim dgamma(.5, .01)$ and $\sigma_1^2 = 1/\tau_1$
- $p(\tau_2) \sim dgamma(.5, .01)$ and $\sigma_2^2 = 1/\tau_2$

Let's do this for the Systolic Blood Pressure Example

Specification of the Likelihood and Prior For Two Means


```
model{

  #likelihood
  for(i in 1:N1){
    a.tmt[i]~dnorm(mu[1],prec[1])  #note this uses precision not var
  }

   for(i in 1:N2){
    b.tmt[i]~dnorm(mu[2],prec[2])  #note this uses precision not var
  }

  #priors
    mu[1]~dnorm(0,0.00001)      #diffuse prior on mean for treatmtent a
    mu[2]~dnorm(0,0.00001)      #diffuse prior on mean for treatment b
    prec[1]~dgamma(0.5,0.01)    #diffuse prior on prec/var for tmt a
    prec[2]~dgamma(0.5,0.01)    #diffuse prior on prec/var for tmt b
    var[1]<-1/prec[1]           #convert to variances for treatment a
    var[2]<-1/prec[2]           #convert to variances for treatment b

  #derived variables
  diff.mu1.mu2<-mu[1]-mu[2]        #difference in the treatment means
  ratio.var1.var2<-var[1]/var[2]   # ratio of the two variances
}
```

In Class Practice Problems

See file "Systolic Blood Pressure Two Means.odc"