Normal Likelihood Ratio

Christopher Gillies

2/8/2018

Model

We want to compare the likelihood of two models. One were we assume the means are the same, and the other where we assume the means are different. We assume the variance within each group is different but known.

Normal distribution density function

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \tag{1}$$

$H_0 : \mu_0 = \mu_1$

$H_1 : \mu_0 \neq \mu_1$

Assume we have n_0 observations in group 0 and n_1 observations in group 1 and given the group status the observations are IID with the same variance . Under H_0 we have the following likelihood:

$$L(Y|H_0) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^{n_0} \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^{n_1} \exp\left[-\sum_{i=0}^{n_0} \frac{(Y_{0i} - \hat{\mu}_{0+1})^2}{2\sigma^2}\right] \exp\left[-\sum_{i=0}^{n_1} \frac{(Y_{1i} - \hat{\mu}_{0+1})^2}{2\sigma^2}\right] \tag{2}$$

$$L(Y|H_0) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^{n} \exp\left[-\sum_{i=0}^{n} \frac{(Y_i - \hat{\mu}_{0+1})^2}{2\sigma^2}\right] \tag{3}$$

where Y_{0i} is the value of the ith example from group 0, the same for Y_{1i}, σ^2 is the variance is the overall variance and μ_{0+1} is the pooled mean estimate.

$$L(Y|H_1) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^{n_0} \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^{n_1} \exp\left[-\sum_{i=0}^{n_0} \frac{(Y_{0i} - \hat{\mu}_0)^2}{2\sigma^2}\right] \exp\left[-\sum_{i=0}^{n_1} \frac{(Y_{1i} - \hat{\mu}_1)^2}{2\sigma^2}\right] \tag{4}$$

$$L(Y|H_1) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^{n} \exp\left[-\sum_{i=0}^{n_0} \frac{(Y_{0i} - \hat{\mu}_0)^2}{2\sigma^2}\right] \exp\left[-\sum_{i=0}^{n_1} \frac{(Y_{1i} - \hat{\mu}_1)^2}{2\sigma^2}\right] \tag{5}$$

where $\hat{\mu}_0$ is the MLE of the mean of group 0, and $\hat{\mu}_1$ is the MLE of the mean of group 1.

$$\frac{L(Y|H_0)}{L(Y|H_1)} = \frac{\exp\left[\frac{-1}{2\sigma^2}\sum_{i=0}^{n}(Y_i - \hat{\mu}_{0+1})^2\right]}{\exp\left[\frac{-1}{2\sigma^2}\sum_{i=0}^{n_0}(Y_{0i} - \hat{\mu}_0)^2\right]\exp\left[\frac{-1}{2\sigma^2}\sum_{i=0}^{n_1}(Y_{1i} - \hat{\mu}_1)^2\right]} = \frac{\exp(\frac{-n\hat{\sigma}^2_{0+1}}{2\sigma^2})}{\exp(\frac{-n_0\hat{\sigma}^2_0 - n_1\hat{\sigma}_1}{2\sigma^2})} \tag{6}$$

Taking the log we have

$$-2\log\left(\frac{L(Y|H_0)}{L(Y|H_1)}\right) = -2\left(\frac{n_0\hat{\sigma}^2_0 + n_1\hat{\sigma}_1}{2\sigma^2} - \frac{n\hat{\sigma}^2_{0+1}}{2\sigma^2}\right) = \frac{-1}{\sigma^2}\left(n_0\hat{\sigma}^2_0 + n_1\hat{\sigma}_1 - n\hat{\sigma}^2_{0+1}\right) \tag{7}$$

We can estimate σ^2 using a pooled variance estimate S_p^2.

$$S_p^2 = \frac{(n_0-1)S_0^2 + (n_1-1)S_1^2}{n_0+n_1-2} \tag{8}$$

where S_0^2 and S_1^2 are sample variance estiamtes for group 0 and group 1.

$$-2\log\left(\frac{L(Y|H_0)}{L(Y|H_1)}\right) \sim \chi_1^2 \tag{9}$$

```
pop.var <- function(x) var(x) * (length(x)-1) / length(x)

pooled.var <- function(s0,s1) {
  s0.var = var(s0)
  s1.var = var(s1)

  n0 = length(s0)
  n1 = length(s1)

  ( (n0 - 1) * s0.var + (n1 - 1) * s1.var ) / (n0 + n1 - 2)
}

lrt = function(s0,s1) {
  pooled = c(s0,s1)
  n = length(pooled)
  n0 = length(s0)
  n1 = length(s1)
  pooled.var.est = pooled.var(s0,s1)

  chi =  -1/pooled.var.est * ( n0 * pop.var(s0) + n1 * pop.var(s1) - n * pop.var(pooled)  )
  pchisq(chi,df=1,lower.tail=F)
}

lrt.simple = function(s0,s1) {
  pooled = c(s0,s1)
  pooled.mean = mean(pooled)
  pooled.var.est = pooled.var(s0,s1)
  s0.mean = mean(s0)
  s1.mean = mean(s1)

  chi = -2 * (sum(dnorm(pooled,mean=pooled.mean,sd=sqrt(pooled.var.est),log = T)) -
    sum(dnorm(s0,mean=s0.mean,sd=sqrt(pooled.var.est),log = T)) - sum(dnorm(s1,mean=s1.mean,sd=sqrt(poo
  pchisq(chi,df=1,lower.tail=F)
}
```

Compare LRT simple verus formula

```
n0 = 100
n1 = 100
s0 = rnorm(n0,sd=2)
s1 = rnorm(n1,sd=2)
lrt.simple(s0,s1)
```

```
## [1] 0.02806679
```

```r
lrt(s0,s1)
```

```
## [1] 0.02806679
```

Check Type I Error

```r
p_lrts = c()
p_ts = c()
nsim = 1000
n0 = 100
n1 = 100
p_lms = c()
for(i in seq(1,nsim)) {
  s0 = rnorm(n0,sd=2)
  s1 = rnorm(n1,sd=2)
  y = c(s0,s1)
  x = c( rep(1,length(s0)), rep(0,length(s1)) )
  p_ts = c(p_ts,  t.test(s0,s1,var.equal = TRUE)$p.val)
  p_lrts = c(p_lrts,  lrt.simple(s0,s1))
  null = lm(y ~ 1)
  alt = lm(y ~ x)
  p_lms = c(p_lms, anova(null,alt,test="Chisq")[2,5])
}

sum(p_ts < 0.05) / nsim
```

```
## [1] 0.049
```

```r
sum(p_lms < 0.05 ) / nsim
```

```
## [1] 0.05
```

```r
sum(p_lrts < 0.05) / nsim
```

```
## [1] 0.05
```

Check Power

```r
p_lrts = c()
p_ts = c()
nsim = 1000
n0 = 100
n1 = 100
p_lms = c()
for(i in seq(1,nsim)) {
  s0 = rnorm(n0)
  s1 = rnorm(n1,mean=0.2)
  y = c(s0,s1)
  x = c( rep(1,length(s0)), rep(0,length(s1)) )
  p_ts = c(p_ts,  t.test(s0,s1)$p.val)
  p_lrts = c(p_lrts,  lrt(s0,s1))
  null = lm(y ~ 1)
```

```r
  alt = lm(y ~ x)
  p_lms = c(p_lms, anova(null,alt,test="Chisq")[2,5])
}

sum(p_ts < 0.05) / nsim
```

```
## [1] 0.299
```

```r
sum(p_lms < 0.05 ) / nsim
```

```
## [1] 0.302
```

```r
sum(p_lrts < 0.05) / nsim
```

```
## [1] 0.302
```

So we did all that stuff and got the same thing as a t-test and linear regression! What was the point? To learn how to do a likelihood ratio test so that we can then do a Bayes factor.

Bayes Factor Based Approach

To compute the Bayes Factor we need $p(y|H_0)$ and $p(y|H_1)$, where these are the probabilities of the data given the corresponding models.

$$p(y) = \int p(y|\theta)p(\theta)d\theta \tag{10}$$

This integral can be challenging to compute. But if we can compute the normalizing constant for the posterior distribution ($p(\theta|y)$), the we can compute $p(y)$ without the need to compute the integral.

$$p(\theta|y) \propto p(y|\theta)p(\theta) = p(y, \theta) \tag{11}$$

$$p(\theta|y)p(y) = p(y, \theta) = p(y|\theta)p(\theta) \tag{12}$$

$$p(y) = \frac{p(y|\theta)p(\theta)}{p(\theta|y)} \tag{13}$$

So if we know the posterior distribution of θ given y, then we can compute $p(y)$ using any value of θ.