

# Normal Likelihood Ratio

*Christopher Gillies*

*2/8/2018*

## Model

We want to compare the likelihood of two models. One where 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

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

$$H_0 : \mu_0 = \mu_1$$

$$H_1 : \mu_0 \neq \mu_1$$

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. Under  $H_0$  we have the following likelihood:

$$L(Y|H_0) = \left( \frac{1}{\sqrt{2\pi\sigma_0^2}} \right)^{n_0} \left( \frac{1}{\sqrt{2\pi\sigma_1^2}} \right)^{n_1} \exp \left[ -\sum_{i=0}^{n_0} \frac{(Y_{0i} - \hat{\mu}_{0+1})^2}{2\sigma_0^2} \right] \exp \left[ -\sum_{i=0}^{n_1} \frac{(Y_{1i} - \hat{\mu}_{0+1})^2}{2\sigma_1^2} \right] \quad (2)$$

where  $Y_{0i}$  is the value of the  $i$ th example from group 0, the same for  $Y_{1i}$ ,  $\sigma_0^2$  is the variance in group 0,  $\sigma_1^2$  is the estimated variance from group 1, and  $\hat{\mu}_{0+1}$  is the pooled mean estimate.

$$L(Y|H_1) = \left( \frac{1}{\sqrt{2\pi\sigma_0^2}} \right)^{n_0} \left( \frac{1}{\sqrt{2\pi\sigma_1^2}} \right)^{n_1} \exp \left[ -\sum_{i=0}^{n_0} \frac{(Y_{0i} - \hat{\mu}_0)^2}{2\sigma_0^2} \right] \exp \left[ -\sum_{i=0}^{n_1} \frac{(Y_{1i} - \hat{\mu}_1)^2}{2\sigma_1^2} \right] \quad (3)$$

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.

$$\frac{L(Y|H_0)}{L(Y|H_1)} = \frac{\exp \left[ -\sum_{i=0}^{n_0} (Y_{0i} - \hat{\mu}_{0+1})^2 \right] \exp \left[ -\sum_{i=0}^{n_1} (Y_{1i} - \hat{\mu}_{0+1})^2 \right]}{\exp \left[ -\sum_{i=0}^{n_0} (Y_{0i} - \hat{\mu}_0)^2 \right] \exp \left[ -\sum_{i=0}^{n_1} (Y_{1i} - \hat{\mu}_1)^2 \right]} = \frac{\exp(-n\hat{\sigma}_{0+1})}{\exp(-n_0\hat{\sigma}_0 - n_1\hat{\sigma}_1)} \quad (4)$$

Taking the log we have

$$-2 \log \left( \frac{L(Y|H_0)}{L(Y|H_1)} \right) = 2(n\hat{\sigma}_{0+1} - n_0\hat{\sigma}_0 - n_1\hat{\sigma}_1) \quad (5)$$

$$-2 \log \left( \frac{L(Y|H_0)}{L(Y|H_1)} \right) \sim \chi_1^2 \quad (6)$$

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

lrt = function(s0,s1) {
  pooled = c(s0,s1)
  n = length(pooled)
```

```

n0 = length(s0)
n1 = length(s1)
chi = 2 * ((n * pop.var(pooled)) - n0 * pop.var(s0) - n1 * pop.var(s1))
pchisq(chi,df=1,lower.tail=F)
}

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=0.1)
  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)
  alt = lm(y ~ x)
  p_lms = c(p_lms, anova(null,alt,test="Chisq")[2,5])
}

sum(p_ts < 0.05) / nsim

## [1] 0.112

sum(p_lms < 0.05 ) / nsim

## [1] 0.117

sum(p_lrts < 0.05) / nsim

## [1] 0.284

```