Likelihood ratio test equivalence to one-way ANOVA F-test

Brenton Horne

27th September 2021

Contents

Hypotheses and notation

Let Y_{ij} denote the jth observation of the ith treatment group. Where $i = 1, 2, 3, ..., m$ and $j = 1, 2, 3, ..., n_i$. Under the null hypothesis: $Y_{ij} \sim \mathrm{N}(\mu, \sigma^2)$. Under the alternative hypothesis: $Y_{ij} \sim \mathrm{N}(\mu_i, \sigma^2)$, where $\mu_i \neq \mu_k$ for at least one pair of i and k values.

Definitions

$$\begin{aligned}
n &= \sum_{i=1}^{m} n_i \\
\overline{Y} &= \frac{1}{n}\sum_{i=1}^{m}\sum_{j=1}^{n_i} Y_{ij} \\
\overline{Y}_i &= \frac{1}{n_i}\sum_{j=1}^{n_i} Y_{ij} \\
\text{Total SS} &= \sum_{i=1}^{m}\sum_{j=1}^{n_i} (Y_{ij} - \overline{Y})^2 \\
\text{SST} &= \sum_{i=1}^{m}\sum_{j=1}^{n_i} (\overline{Y}_i - \overline{Y})^2 \\
\text{SSE} &= \sum_{i=1}^{m}\sum_{j=1}^{n_i} (Y_{ij} - \overline{Y}_i)^2.
\end{aligned}$$

We will later show that:

$$\begin{aligned}
\text{SSE} &= \sum_{i=1}^{m}\sum_{j=1}^{n_i} (Y_{ij} - \overline{Y})^2 - (\overline{Y}_i - \overline{Y})^2 \\
&= \text{Total SS} - \text{SST}.
\end{aligned}$$

We will later use δ_{ik}, which is the Kronecker delta symbol. It equals 0 if $i \neq k$ and 1 otherwise.

F statistic for the one-way ANOVA test is:

$$\begin{aligned}
F &= \frac{\text{MST}}{\text{MSE}} \\
&= \frac{\text{SST}/(m-1)}{\text{SSE}/(n-m)} \\
&= \frac{(n-m)\text{SST}}{(m-1)\text{SSE}}.
\end{aligned}$$

Which has $m-1$ numerator and $n-m$ denominator degrees of freedom.

Derivation of maximum likelihood under the null

Let us denote the parameter space under the null hypothesis as $\Omega_0 = \{(\mu, \sigma^2) : -\infty < \mu < \infty,\ \sigma^2 > 0\}$. The parameter space under the alternative hypothesis is

$\Omega_a = \{(\mu_i, \sigma^2) : \mu_i \neq \mu_j \text{ for at least one pair of } i \text{ and } j \text{ values}, -\infty < \mu_i < \infty,\ \sigma^2 > 0\}$. Our unrestricted parameter space is given by $\Omega = \Omega_0 \cup \Omega_a$.

$$\begin{aligned}
L(\Omega_0) &= \prod_{i=1}^{m}\prod_{j=1}^{n_i} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(Y_{ij}-\mu)^2\right) \\
&= (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij}-\mu)^2\right). \qquad (1) \\
\ln L(\Omega_0) &= -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij}-\mu)^2.
\end{aligned}$$

Setting the partial derivative with respect to σ^2 to zero to maximize the likelihood:

$$\begin{aligned}
\frac{\partial \ln L(\Omega_0)}{\partial \sigma^2}\bigg|_{\sigma^2=\widehat{\sigma^2},\ \mu=\widehat{\mu}} &= -\frac{n}{2\widehat{\sigma^2}} + \frac{1}{2\widehat{\sigma^4}}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij}-\widehat{\mu})^2 \\
&= 0. \qquad (2)
\end{aligned}$$

Multiplying Equation 2 by $2\widehat{\sigma^4}$ yields:

$$\begin{aligned}
-n\widehat{\sigma^2} + \sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij}-\widehat{\mu})^2 = 0 \\
\implies \widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij}-\hat{\mu})^2. \qquad (3)
\end{aligned}$$

Next we will set the partial derivative with respect to μ to zero to find the MLE of μ:

$$\begin{aligned}
\frac{\partial \ln L(\Omega_0)}{\partial \mu}\bigg|_{\sigma^2=\widehat{\sigma^2},\ \mu=\widehat{\mu}} &= -\frac{1}{2\widehat{\sigma^2}}\sum_{i=1}^{m}\sum_{j=1}^{n_i} 2\cdot(-1)\cdot(Y_{ij}-\widehat{\mu}) \\
&= \frac{1}{\widehat{\sigma^2}}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij}-\widehat{\mu}) \\
&= 0. \qquad (4)
\end{aligned}$$

Multiplying Equation 4 by $\widehat{\sigma^2}$ yields:

$$\begin{aligned}
\therefore \quad \sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij}-\widehat{\mu}) &= 0 \\
n\overline{Y} - n\widehat{\mu} &= 0 \\
\implies \widehat{\mu} &= \overline{Y}. \qquad (5)
\end{aligned}$$

Substituting Equations 3 and 5 into Equation 1 yields:

$$\begin{aligned}
L(\widehat{\Omega_0}) &= (2\pi\widehat{\sigma^2})^{-n/2}\exp\left(-\frac{1}{2\widehat{\sigma^2}}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij}-\widehat{\mu})^2\right) \\
&= (2\pi\widehat{\sigma^2})^{-n/2}\exp\left(-\frac{n}{2}\right) \\
&= (2\pi e\widehat{\sigma^2})^{-n/2} \\
&= \left(\frac{2\pi e}{n}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij}-\overline{Y})^2\right)^{-n/2}. \qquad (6)
\end{aligned}$$

Derivation of the unrestricted maximum likelihood

$$\begin{aligned}
L(\Omega) &= \prod_{i=1}^{m}\prod_{j=1}^{n_i}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}(Y_{ij}-\mu_i)^2\right) \\
&= (2\pi\sigma^2)^{-n/2}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij}-\mu_i)^2\right) \qquad (7) \\
\ln L(\Omega) &= -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij}-\mu_i)^2
\end{aligned}$$

Setting the partial derivative with respect to σ^2 to zero to find the MLE of σ^2:

$$\begin{aligned}
\frac{\partial \ln L(\Omega)}{\partial\sigma^2}\Big|_{\sigma^2=\widehat{\sigma^2},\ \mu_i=\widehat{\mu}_i} &= -\frac{n}{2\widehat{\sigma^2}} + \frac{1}{2\widehat{\sigma^4}}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij}-\widehat{\mu}_i)^2 \\
&= 0. \qquad (8)
\end{aligned}$$

Multiplying Equation 8 by $2\widehat{\sigma^4}$ yields:

$$\begin{aligned}
&-n\widehat{\sigma^2} + \sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij} - \widehat{\mu}_i)^2 = 0 \\
&\qquad\qquad\qquad \implies \widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij} - \widehat{\mu}_i)^2 \qquad (9)
\end{aligned}$$

Setting the partial derivative with respect to μ_i to zero to find the MLE of μ_k:

$$\begin{aligned}
\frac{\partial \ln L(\Omega)}{\partial \mu_k}\Big|_{\sigma^2=\widehat{\sigma^2},\ \mu_i=\widehat{\mu}_i} &= -\frac{1}{2\widehat{\sigma^2}}\sum_{i=1}^{m}\sum_{j=1}^{n_i} 2\cdot(-1)\cdot(Y_{ij} - \widehat{\mu}_i)\delta_{ik} \\
&= 0 \\
\frac{1}{\widehat{\sigma^2}}\sum_{j=1}^{n_k}(Y_{kj} - \widehat{\mu}_k) = 0. & \qquad (10)
\end{aligned}$$

Multiplying both sides of Equation 10 by $\widehat{\sigma^2}$ yields:

$$\begin{aligned}
\sum_{j=1}^{n_k}(Y_{kj} - \widehat{\mu}_k) &= 0 \\
n_k\overline{Y}_k - n_k\widehat{\mu}_k &= 0 \\
\implies \widehat{\mu}_k &= \overline{Y}_k. \qquad (11)
\end{aligned}$$

Substituting Equations 9 and 11 into Equation 7 yields:

$$\begin{aligned}
L(\widehat{\Omega}) &= (2\pi\widehat{\sigma^2})^{-n/2}\exp\left(-\frac{1}{2\widehat{\sigma^2}}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij} - \widehat{\mu}_i)^2\right) \\
&= (2\pi\widehat{\sigma^2})^{-n/2}\exp\left(-\frac{n}{2}\right) \\
&= \left(\frac{2\pi e}{n}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij} - \overline{Y}_i)^2\right)^{-n/2}. \qquad (12)
\end{aligned}$$

Likelihood ratio

$$\begin{aligned}
\lambda &= \frac{L(\widehat{\Omega_0})}{L(\widehat{\Omega})} \\
&= \left(\frac{\sum_{i=1}^{m} \sum_{j=1}^{n_i} (Y_{ij} - \overline{Y}_i)^2}{\sum_{i=1}^{m} \sum_{j=1}^{n_i} (Y_{ij} - \overline{Y})^2} \right)^{n/2} \\
\lambda^{-2/n} &= \frac{\sum_{i=1}^{m} \sum_{j=1}^{n_i} (Y_{ij} - \overline{Y})^2}{\sum_{i=1}^{m} \sum_{j=1}^{n_i} (Y_{ij} - \overline{Y}_i)^2}. \qquad (13)
\end{aligned}$$

Let us try to simplify our denominator in Equation 13:

$$\begin{aligned}
\sum_{i=1}^{m} \sum_{j=1}^{n_i} (Y_{ij} - \overline{Y}_i)^2 &= \sum_{i=1}^{m} \sum_{j=1}^{n_i} (Y_{ij} - \overline{Y} + \overline{Y} - \overline{Y}_i)^2 \\
&= \sum_{i=1}^{m} \sum_{j=1}^{n_i} (Y_{ij} - \overline{Y})^2 + 2(Y_{ij} - \overline{Y})(\overline{Y} - \overline{Y}_i) + (\overline{Y} - \overline{Y}_i)^2 \qquad (14)
\end{aligned}$$

Next let us try to simplify the second term in Equation 14.

$$\sum_{i=1}^{m} \sum_{j=1}^{n_i} (Y_{ij} - \overline{Y})(\overline{Y} - \overline{Y}_i) = \sum_{i=1}^{m} \sum_{j=1}^{n_i} (Y_{ij} - \overline{Y})\overline{Y} - (Y_{ij} - \overline{Y})\overline{Y}_i \qquad (15)$$

Let us now simplify the first term in Equation 15.

$$\begin{aligned}
\sum_{i=1}^{m} \sum_{j=1}^{n_i} (Y_{ij} - \overline{Y})\overline{Y} &= \overline{Y} \sum_{i=1}^{m} \sum_{j=1}^{n_i} Y_{ij} - \overline{Y} \\
&= \overline{Y}(n\overline{Y} - \sum_{i=1}^{m} \sum_{j=1}^{n_i} \overline{Y}) \\
&= \overline{Y}(n\overline{Y} - n\overline{Y}) \\
&= 0. \qquad (16)
\end{aligned}$$

Substituting Equation 16 into Equation 15 yields:

$$\begin{aligned}
\therefore \sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij}-\overline{Y})(\overline{Y}-\overline{Y}_i) &= \sum_{i=1}^{m}\sum_{j=1}^{n_i}-(Y_{ij}-\overline{Y})\overline{Y}_i \\
&= -\sum_{i=1}^{m}\overline{Y}_i\sum_{j=1}^{n_i}(Y_{ij}-\overline{Y}) \\
&= -\sum_{i=1}^{m}\overline{Y}_i(n_i\overline{Y}_i-n_i\overline{Y}) \\
&= -\sum_{i=1}^{m}n_i\overline{Y}_i(\overline{Y}_i-\overline{Y}).
\end{aligned} \tag{17}$$

Equation 17 can be shown to be equivalent to $-\sum_{i=1}^{m}\sum_{j=1}^{n_i}(\overline{Y}_i-\overline{Y})^2$. Namely:

$$\begin{aligned}
\sum_{i=1}^{m}\sum_{j=1}^{n_i}(\overline{Y}_i-\overline{Y})^2 &= \sum_{i=1}^{m}n_i(\overline{Y}_i-\overline{Y})^2 \\
&= \sum_{i=1}^{m}n_i(\overline{Y}_i-\overline{Y})(\overline{Y}_i-\overline{Y}) \\
&= \sum_{i=1}^{m}n_i\overline{Y}_i(\overline{Y}_i-\overline{Y})-n_i\overline{Y}(\overline{Y}_i-\overline{Y}) \\
&= \left(\sum_{i=1}^{m}n_i\overline{Y}_i(\overline{Y}_i-\overline{Y})\right)-\overline{Y}\sum_{i=1}^{m}n_i(\overline{Y}_i-\overline{Y}) \\
&= \left(\sum_{i=1}^{m}n_i\overline{Y}_i(\overline{Y}_i-\overline{Y})\right)-\overline{Y}\sum_{i=1}^{m}n_i\overline{Y}_i-n_i\overline{Y}.
\end{aligned} \tag{18}$$

Simplifying the second term in Equation 18 yields:

$$\begin{aligned}
\sum_{i=1}^{m}n_i\overline{Y}_i &= \sum_{i=1}^{m}n_i\frac{1}{n_i}\sum_{j=1}^{n_i}Y_{ij} \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n_i}Y_{ij} \\
&= n\overline{Y}.
\end{aligned} \tag{19}$$

Substituting Equation 19 into Equation 18 yields:

$$\begin{aligned}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(\overline{Y}_i-\overline{Y})^2 &= \left(\sum_{i=1}^{m} n_i\overline{Y}_i(\overline{Y}_i-\overline{Y})\right)-\overline{Y}(n\overline{Y}-\sum_{i=1}^{m} n_i\overline{Y})\\ &= \left(\sum_{i=1}^{m} n_i\overline{Y}_i(\overline{Y}_i-\overline{Y})\right)-\overline{Y}(n\overline{Y}-\overline{Y}\sum_{i=1}^{m} n_i)\\ &= \left(\sum_{i=1}^{m} n_i\overline{Y}_i(\overline{Y}_i-\overline{Y})\right)-\overline{Y}(n\overline{Y}-\overline{Y}n)\\ &= \sum_{i=1}^{m} n_i\overline{Y}_i(\overline{Y}_i-\overline{Y}).\end{aligned}$$

Therefore:

$$\begin{aligned}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij}-\overline{Y})(\overline{Y}-\overline{Y}_i) &= -\sum_{i=1}^{m} n_i(\overline{Y}_i-\overline{Y})^2\\ &= -\sum_{i=1}^{m}\sum_{j=1}^{n_i}(\overline{Y}_i-\overline{Y})^2.\end{aligned}$$

Thus Equation 14 becomes:

$$\begin{aligned}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij}-\overline{Y}_i)^2 &= \sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij}-\overline{Y})^2-2(\overline{Y}_i-\overline{Y})^2+(\overline{Y}_i-\overline{Y})^2\\ &= \sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij}-\overline{Y})^2-(\overline{Y}_i-\overline{Y})^2.\end{aligned}$$

Therefore Equation 13 becomes:

$$\begin{aligned}\lambda^{-2/n} &= \frac{\sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij}-\overline{Y})^2}{\sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij}-\overline{Y})^2-(\overline{Y}_i-\overline{Y})^2}\\ &= \frac{\sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij}-\overline{Y})^2-(\overline{Y}_i-\overline{Y})^2+(\overline{Y}_i-\overline{Y})^2}{\sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij}-\overline{Y})^2-(\overline{Y}_i-\overline{Y})^2}\\ &= 1+\frac{\sum_{i=1}^{m}\sum_{j=1}^{n_i}(\overline{Y}_i-\overline{Y})^2}{\sum_{i=1}^{m}\sum_{j=1}^{n_i}(Y_{ij}-\overline{Y})^2-(\overline{Y}_i-\overline{Y})^2}\\ &= 1+\frac{\text{SST}}{\text{SSE}}\\ &= 1+\frac{F(m-1)}{(n-m)}\\ \implies \lambda &= \left(\frac{1}{1+\dfrac{F(m-1)}{n-m}}\right)^{n/2}.\end{aligned}$$

Therefore as λ is a monotone decreasing function of F, testing whether $\lambda < k$ is equivalent to testing whether $F > k^*$, which forms the basis of the one-way ANOVA F-test.