Derivation of a test for equality of means with unequal variances

Brenton Horne

19th September 2021

Contents

Hypotheses

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: $Y_{ij} \sim \mathrm{N}(\mu, \sigma_i^2)$. Under the alternative hypothesis: $Y_{ij} \sim \mathrm{N}(\mu_i, \sigma_i^2)$, where $\mu_i \neq \mu_k$ for at least one pair of i and k values.

Definitions

$$n = \sum_{i=1}^{m} n_i$$
$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{m}\sum_{j=1}^{n_i} Y_{ij}$$
$$\overline{Y}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} Y_{ij}$$

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

Derivation of the maximum likelihood under the null

In this section we will use $L(H_0)$ to denote the likelihood under the null.

$$\begin{aligned} L(H_0) &= \prod_{i=1}^{m}\prod_{j=1}^{n_i} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{1}{2\sigma_i^2}(Y_{ij}-\mu)^2\right) \\ &= (2\pi)^{-n/2}\left(\prod_{i=1}^{m}\sigma_i^{-n_i}\right)\exp\left(-\frac{1}{2}\sum_{i=1}^{m}\frac{1}{\sigma_i^2}\sum_{j=1}^{n_i}(Y_{ij}-\mu)^2\right). \end{aligned} \tag{1}$$

Taking the natural logarithm yields:

$$\ln L(H_0) = -\frac{n}{2}\ln 2\pi - \frac{1}{2}\sum_{i=1}^{m} n_i \ln\sigma_i^2 - \frac{1}{2}\sum_{i=1}^{m}\frac{1}{\sigma_i^2}\sum_{j=1}^{n_i}(Y_{ij}-\mu)^2.$$

Differentiating the log-likelihood with respect to μ and setting to zero to maximize the likelihood:

$$\frac{\partial \ln L(H_0)}{\partial \mu}\Big|_{\mu=\widehat{\mu},\ \sigma_i^2=\widehat{\sigma_i^2}} = -\frac{1}{2}\sum_{i=1}^{m}\frac{1}{\widehat{\sigma_i^2}}\sum_{j=1}^{n_i}2(-1)(Y_{ij}-\widehat{\mu})$$

$$= 0$$

$$\sum_{i=1}^{m}\frac{1}{\widehat{\sigma_i^2}}\sum_{j=1}^{n_i}(Y_{ij}-\widehat{\mu}) = 0$$

$$\sum_{i=1}^{m}\frac{1}{\widehat{\sigma_i^2}}(n_i\overline{Y}_i - n_i\widehat{\mu}) = 0$$

$$\left(\sum_{i=1}^{m}\frac{n_i\overline{Y}_i}{\widehat{\sigma_i^2}}\right) - \left(\sum_{i=1}^{m}\frac{n_i}{\widehat{\sigma_i^2}}\right)\widehat{\mu} = 0$$

$$\widehat{\mu} = \frac{\sum_{i=1}^{m}\frac{n_i\overline{Y}_i}{\widehat{\sigma_i^2}}}{\sum_{i=1}^{m}\frac{n_i}{\widehat{\sigma_i^2}}}. \tag{2}$$

Differentiating the log-likelihood with respect to σ_k^2 and setting to zero to maximize the likelihood:

$$\frac{\partial \ln L(H_0)}{\partial \sigma_k^2}\Big|_{\mu=\widehat{\mu},\ \sigma_i^2=\widehat{\sigma_i^2}} = -\frac{1}{2}\sum_{i=1}^{m}\frac{n_i}{\widehat{\sigma_i^2}}\delta_{ik} - \frac{1}{2}\sum_{i=1}^{m}-\frac{1}{\widehat{\sigma_i^4}}\delta_{ik}\sum_{j=1}^{n_i}(Y_{ij}-\widehat{\mu})^2$$

$$= -\frac{1}{2}\frac{n_k}{\widehat{\sigma_k^2}} + \frac{1}{2\widehat{\sigma_k^4}}\sum_{j=1}^{n_k}(Y_{kj}-\widehat{\mu})^2$$

$$= 0.$$

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

$$-n_k\widehat{\sigma_k^2} + \sum_{j=1}^{n_k}(Y_{kj}-\widehat{\mu})^2 = 0$$

$$\widehat{\sigma_k^2} = \frac{1}{n_k}\sum_{j=1}^{n_k}(Y_{kj}-\widehat{\mu})^2. \tag{3}$$

So here we must numerically approximate $\widehat{\mu}$ and $\widehat{\sigma_i^2}$ using a technique like Newton's method and substitute this into our expression for $L(H_0)$ to get the maximum likelihood under the null. If we use Newton's method, we must have functions we are finding the zeros of. Let:

$$f(\widehat{\mu},\widehat{\sigma_k^2}) = \widehat{\mu} - \frac{\sum_{k=1}^m \frac{n_k\overline{Y}_k}{\widehat{\sigma_k^2}}}{\sum_{k=1}^m \frac{n_k}{\widehat{\sigma_k^2}}}$$

$$g_i(\widehat{\mu},\widehat{\sigma_k^2}) = \widehat{\sigma_i^2} - \frac{1}{n_i}\sum_{j=1}^{n_i}(Y_{ij}-\widehat{\mu})^2.$$

Therefore:

$$\frac{\partial f}{\partial \widehat{\mu}} = 1$$

$$\frac{\partial f}{\partial \widehat{\sigma_i^2}} = -\frac{\sum_{k=1}^m -\frac{n_k\overline{Y}_k}{\widehat{\sigma_k^4}}\delta_{ik}}{\sum_{k=1}^m \frac{n_k}{\widehat{\sigma_k^2}}} - \frac{\sum_{k=1}^m \frac{n_k\overline{Y}_k}{\widehat{\sigma_k^2}}}{-\left(\sum_{k=1}^m \frac{n_k}{\widehat{\sigma_k^2}}\right)^2}\sum_{k=1}^m -\frac{n_k}{\widehat{\sigma_k^4}}\delta_{ik}$$

$$= \frac{\frac{n_i\overline{Y}_i}{\widehat{\sigma_i^4}}}{\sum_{k=1}^m \frac{n_k}{\widehat{\sigma_k^2}}} - \frac{\sum_{k=1}^m \frac{n_k\overline{Y}_k}{\widehat{\sigma_k^2}}}{\left(\sum_{k=1}^m \frac{n_k}{\widehat{\sigma_k^2}}\right)^2}\frac{n_i}{\widehat{\sigma_i^4}}$$

$$= \frac{n_i}{\widehat{\sigma_i^4}\sum_{k=1}^m \frac{n_k}{\widehat{\sigma_k^2}}}\left(\overline{Y}_i - \frac{\sum_{i=1}^m \frac{n_i\overline{Y}_i}{\widehat{\sigma_i^2}}}{\sum_{i=1}^m \frac{n_i}{\widehat{\sigma_i^2}}}\right).$$

As for the derivatives of $g_i(\widehat{\mu},\widehat{\sigma_k^2})$:

$$\frac{\partial g_i(\widehat{\mu},\widehat{\sigma_k^2})}{\partial \widehat{\mu}} = -\frac{1}{n_i}\sum_{j=1}^{n_i}2(-1)(Y_{ij}-\widehat{\mu})$$

$$= \frac{2}{n_i}(n_i\overline{Y}_i - n_i\widehat{\mu})$$

$$= 2(\overline{Y}_i - \widehat{\mu})$$

$$\frac{\partial g_i(\widehat{\mu},\widehat{\sigma_k^2})}{\partial \widehat{\sigma_k^2}} = \delta_{ik}.$$