

Likelihood-ratio test for samples from gamma-distributed populations

Brenton Horne

11th October 2021

Contents

Definitions	2
Hypotheses	2
Derivation of the maximum likelihood under the null	2
Using Newton’s method to obtain $\hat{\alpha}$	4
Derivation of the unrestricted maximum likelihood	5
Using Newton’s method to obtain $\hat{\alpha}_k$	6
Likelihood ratio	7

Definitions

Let Y_{ij} denote the j th observation from the i th treatment group, where $i = 1, 2, 3, \dots, m$ and $j = 1, 2, 3, \dots, n_i$.

Let:

$$n = \sum_{i=1}^m n_i$$

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

$$\bar{Y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij}.$$

Hypotheses

H_0 : $Y_{ij} \sim \text{Gamma}(\alpha, \beta)$

H_A : $Y_{ij} \sim \text{Gamma}(\alpha_i, \beta_i)$, where $\alpha_i \neq \alpha_k$ or $\beta_i \neq \beta_k$ for at least one combination of i and k values.

Let us denote the parameter space under the null hypothesis as $\Omega_0 = \{(\alpha, \beta) : 0 < \alpha, \beta < \infty\}$ and the parameter space under the alternative hypothesis as

$\Omega_a = \{(\alpha_i, \beta_i) : 0 < \alpha_i, \beta_i < \infty, \alpha_i \neq \alpha_k \text{ or } \beta_i \neq \beta_k \text{ for at least one pair of } i \text{ and } k \text{ values}\}$. The unrestricted parameter space is $\Omega = \Omega_0 \cup \Omega_a = \{(\alpha_i, \beta_i) : 0 < \alpha_i, \beta_i < \infty\}$.

Derivation of the maximum likelihood under the null

$$\begin{aligned} L(\Omega_0) &= \prod_{i=1}^m \prod_{j=1}^{n_i} \frac{1}{\Gamma(\alpha)\beta^\alpha} Y_{ij}^{\alpha-1} \exp\left(-\frac{Y_{ij}}{\beta}\right) \\ &= (\Gamma(\alpha)\beta^\alpha)^{-n} \left(\prod_{i=1}^m \prod_{j=1}^{n_i} Y_{ij}\right)^{\alpha-1} \exp\left(-\frac{n\bar{Y}}{\beta}\right) \\ &= \left(\Gamma(\alpha)\beta^\alpha \exp\left(\frac{\bar{Y}}{\beta}\right)\right)^{-n} \left(\prod_{i=1}^m \prod_{j=1}^{n_i} Y_{ij}\right)^{\alpha-1}. \end{aligned} \tag{1}$$

Taking the natural logarithm yields:

$$\ln L(\Omega_0) = -n \ln \left(\Gamma(\alpha) \beta^\alpha \exp \left(\frac{\bar{Y}}{\beta} \right) \right) + (\alpha - 1) \sum_{i=1}^m \sum_{j=1}^{n_i} \ln Y_{ij}.$$

Next we will take the partial derivative with respect to α and set it to zero:

$$\begin{aligned} \frac{\partial \ln L(\Omega_0)}{\partial \alpha} \Big|_{\alpha=\hat{\alpha}, \beta=\hat{\beta}} &= -n \left(\frac{\Gamma'(\hat{\alpha}) \hat{\beta}^{\hat{\alpha}} \exp \left(\frac{\bar{Y}}{\hat{\beta}} \right) + \Gamma(\hat{\alpha}) (\ln \hat{\beta}) \hat{\beta}^{\hat{\alpha}} \exp \left(\frac{\bar{Y}}{\hat{\beta}} \right)}{\Gamma(\hat{\alpha}) \hat{\beta}^{\hat{\alpha}} \exp \left(\frac{\bar{Y}}{\hat{\beta}} \right)} \right) + \sum_{i=1}^m \sum_{j=1}^{n_i} \ln Y_{ij} = 0 \\ \therefore 0 &= -n(\psi^{(0)}(\hat{\alpha}) + \ln \hat{\beta}) + \sum_{i=1}^m \sum_{j=1}^{n_i} \ln Y_{ij}. \end{aligned} \quad (2)$$

Where $\psi^{(0)}(\hat{\alpha})$ is the digamma function. Next we will take the partial derivative with respect to β and set it to zero:

$$\begin{aligned} \frac{\partial \ln L(\Omega_0)}{\partial \beta} \Big|_{\alpha=\hat{\alpha}, \beta=\hat{\beta}} &= -n \left(\frac{\Gamma(\hat{\alpha}) \hat{\beta}^{\hat{\alpha}-1} \hat{\alpha} \exp \left(\frac{\bar{Y}}{\hat{\beta}} \right) - \Gamma(\hat{\alpha}) \hat{\beta}^{\hat{\alpha}} \exp \left(\frac{\bar{Y}}{\hat{\beta}} \right) \frac{\bar{Y}}{\hat{\beta}^2}}{\Gamma(\hat{\alpha}) \hat{\beta}^{\hat{\alpha}} \exp \left(\frac{\bar{Y}}{\hat{\beta}} \right)} \right) = 0 \\ \therefore 0 &= -n \left(\frac{\hat{\alpha}}{\hat{\beta}} - \frac{\bar{Y}}{\hat{\beta}^2} \right). \end{aligned} \quad (3)$$

Multiplying both sides of Equation 3 by $-\frac{\hat{\beta}^2}{n}$ yields:

$$\begin{aligned} 0 &= \hat{\alpha} \hat{\beta} - \bar{Y} \\ \therefore \hat{\beta} &= \frac{\bar{Y}}{\hat{\alpha}}. \end{aligned} \quad (4)$$

Re-writing Equation 2 using Equation 4 yields:

$$0 = -n \left(\psi^{(0)}(\hat{\alpha}) + \ln \frac{\bar{Y}}{\hat{\alpha}} \right) + \sum_{i=1}^m \sum_{j=1}^{n_i} \ln Y_{ij}. \quad (5)$$

This equation cannot be analytically solved, so $\hat{\alpha}$ must be numerically approximated using Newton's method and $\hat{\beta}$ must be estimated from our approximation of $\hat{\alpha}$ using Equation 4.

Before we apply Newton's method to get an algorithm for estimating $\hat{\alpha}$, let us use our MLEs to find our expression for maximum likelihood under the null. Substituting Equation 4 into Equation 1 yields:

$$\begin{aligned} L(\widehat{\Omega}_0) &= \left(\Gamma(\hat{\alpha}) \left(\frac{\bar{Y}}{\hat{\alpha}} \right)^{\hat{\alpha}} \exp(\hat{\alpha}) \right)^{-n} \left(\prod_{i=1}^m \prod_{j=1}^{n_i} Y_{ij} \right)^{\hat{\alpha}-1} \\ &= \left(\Gamma(\hat{\alpha}) \left(\frac{\bar{Y}e}{\hat{\alpha}} \right)^{\hat{\alpha}} \right)^{-n} \prod_{i=1}^m \prod_{j=1}^{n_i} Y_{ij}^{\hat{\alpha}-1}. \end{aligned} \tag{6}$$

Using Newton's method to obtain $\hat{\alpha}$

To use Newton's method, we must find the Jacobian for our problem, which should in this case be a scalar as we have only one equation to solve (Equation 5). Labelling the right-hand side of Equation 5 as $f(\hat{\alpha})$, and differentiating it with respect to $\hat{\alpha}$ yields:

$$\begin{aligned} \frac{\partial f}{\partial \hat{\alpha}} &= -n \left(\psi^{(1)}(\hat{\alpha}) - \frac{1}{\hat{\alpha}} \right) \\ &= n \left(\frac{1}{\hat{\alpha}} - \psi^{(1)}(\hat{\alpha}) \right). \end{aligned}$$

Therefore we can refine our estimate of $\hat{\alpha}$ from an initial guess $\hat{\alpha}^{(0)}$ using:

$$\begin{aligned} \hat{\alpha}^{(k+1)} &= \hat{\alpha}^{(k)} - \frac{f(\hat{\alpha}^{(k)})}{\left. \frac{\partial f}{\partial \hat{\alpha}} \right|_{\hat{\alpha}=\hat{\alpha}^{(k)}}} \\ &= \hat{\alpha}^{(k)} - \frac{-n \left(\psi^{(1)}(\hat{\alpha}^{(k)}) + \ln \frac{\bar{Y}}{\hat{\alpha}^{(k)}} \right) + \sum_{i=1}^m \sum_{j=1}^{n_i} \ln Y_{ij}}{n \left(\frac{1}{\hat{\alpha}^{(k)}} - \psi^{(1)}(\hat{\alpha}^{(k)}) \right)}. \end{aligned}$$

Where $\psi^{(1)}(\hat{\alpha}^{(k)})$ is the derivative of the digamma function, also known as the trigamma function.

Derivation of the unrestricted maximum likelihood

$$\begin{aligned}
 L(\Omega) &= \prod_{i=1}^m \prod_{j=1}^{n_i} \frac{1}{\Gamma(\alpha_i) \beta_i^{\alpha_i}} Y_{ij}^{\alpha_i-1} \exp\left(-\frac{Y_{ij}}{\beta_i}\right) \\
 &= \left(\prod_{i=1}^m (\Gamma(\alpha_i) \beta_i^{\alpha_i})^{-n_i} \right) \left(\prod_{i=1}^m \prod_{j=1}^{n_i} Y_{ij}^{\alpha_i-1} \right) \exp\left(-\sum_{i=1}^m \frac{n_i \bar{Y}_i}{\beta_i}\right).
 \end{aligned} \tag{7}$$

Taking the natural logarithm yields:

$$\ln L(\Omega) = -\sum_{i=1}^m n_i \ln(\Gamma(\alpha_i) \beta_i^{\alpha_i}) + \sum_{i=1}^m (\alpha_i - 1) \sum_{j=1}^{n_i} \ln Y_{ij} - \sum_{i=1}^m \frac{n_i \bar{Y}_i}{\beta_i}.$$

Differentiating our log-likelihood with respect to α_k and setting the derivative to zero:

$$\begin{aligned}
 \frac{\partial \ln L(\Omega)}{\partial \alpha_k} \Big|_{\alpha_i=\hat{\alpha}_i, \beta_i=\hat{\beta}_i} &= -\sum_{i=1}^m n_i \left(\frac{\Gamma'(\hat{\alpha}_i) \hat{\beta}_i^{\hat{\alpha}_i} + \Gamma(\hat{\alpha}_i) (\ln \hat{\beta}_i) \hat{\beta}_i^{\hat{\alpha}_i}}{\Gamma(\hat{\alpha}_i) \hat{\beta}_i^{\hat{\alpha}_i}} \right) \delta_{ik} + \sum_{i=1}^m \delta_{ik} \sum_{j=1}^{n_i} \ln Y_{ij} = 0 \\
 \therefore 0 &= -n_k (\psi^{(0)}(\hat{\alpha}_k) + \ln \hat{\beta}_k) + \sum_{j=1}^{n_k} \ln Y_{kj}.
 \end{aligned} \tag{8}$$

Where δ_{ik} is the Kronecker delta. Differentiating our log-likelihood with respect to β_k and setting the derivative to zero:

$$\begin{aligned}
 \frac{\partial \ln L(\Omega)}{\partial \beta_k} \Big|_{\alpha_i=\hat{\alpha}_i, \beta_i=\hat{\beta}_i} &= -\sum_{i=1}^m n_i \left(\frac{\Gamma(\hat{\alpha}_i) \hat{\beta}_i^{\hat{\alpha}_i-1} \hat{\alpha}_i}{\Gamma(\hat{\alpha}_i) \hat{\beta}_i^{\hat{\alpha}_i}} \right) \delta_{ik} + \sum_{i=1}^m \frac{n_i \bar{Y}_i}{\hat{\beta}_i^2} \delta_{ik} = 0 \\
 \therefore 0 &= -\frac{n_k \hat{\alpha}_k}{\hat{\beta}_k} + \frac{n_k \bar{Y}_k}{\hat{\beta}_k^2}.
 \end{aligned} \tag{9}$$

Multiplying Equation 9 by $\frac{\hat{\beta}_k^2}{n_k}$ yields:

$$\begin{aligned}
 -\hat{\alpha}_k \hat{\beta}_k + \bar{Y}_k &= 0 \\
 \hat{\beta}_k &= \frac{\bar{Y}_k}{\hat{\alpha}_k}.
 \end{aligned} \tag{10}$$

Substituting Equation 10 into Equation 8 yields:

$$0 = -n_k \left(\psi^{(0)}(\hat{\alpha}_k) + \ln \left(\frac{\bar{Y}_k}{\hat{\alpha}_k} \right) \right) + \sum_{j=1}^{n_k} \ln Y_{kj}. \quad (11)$$

Equation 11 has no closed-form solution and hence $\hat{\alpha}_k$ must be numerically approximated using a technique like Newton's method. Before we obtain an iterative formula based on Newton's method to approximate $\hat{\alpha}_k$, we will use Equation 10 to obtain the unrestricted maximum likelihood by substituting it into Equation 7:

$$\begin{aligned} L(\hat{\Omega}) &= \left(\prod_{i=1}^m \left(\Gamma(\hat{\alpha}_i) \left(\frac{\bar{Y}_i}{\hat{\alpha}_i} \right)^{\hat{\alpha}_i} \right)^{-n_i} \right) \left(\prod_{i=1}^m \prod_{j=1}^{n_i} Y_{ij}^{\hat{\alpha}_i-1} \right) \exp \left(- \sum_{i=1}^m \frac{n_i \bar{Y}_i}{\hat{\alpha}_i} \right) \\ &= \left(\prod_{i=1}^m \left(\Gamma(\hat{\alpha}_i) \left(\frac{\bar{Y}_i}{\hat{\alpha}_i} \right)^{\hat{\alpha}_i} \right)^{-n_i} \right) \left(\prod_{i=1}^m \prod_{j=1}^{n_i} Y_{ij}^{\hat{\alpha}_i-1} \right) \exp \left(- \sum_{i=1}^m n_i \hat{\alpha}_i \right) \\ &= \left(\prod_{i=1}^m \left(\Gamma(\hat{\alpha}_i) \left(\frac{\bar{Y}_i e}{\hat{\alpha}_i} \right)^{\hat{\alpha}_i} \right)^{-n_i} \right) \prod_{i=1}^m \prod_{j=1}^{n_i} Y_{ij}^{\hat{\alpha}_i-1}. \end{aligned} \quad (12)$$

Using Newton's method to obtain $\hat{\alpha}_k$

Calling the right-hand side of Equation 11, $g_k(\{\hat{\alpha}_j\})$, and taking its partial derivative with respect to $\hat{\alpha}_i$ so as to create the Jacobian matrix:

$$\begin{aligned} \frac{\partial g_k(\{\hat{\alpha}_j\})}{\partial \hat{\alpha}_i} &= -n_k \left(\psi^{(1)}(\hat{\alpha}_k) - \frac{1}{\hat{\alpha}_k} \right) \delta_{ik} \\ &= n_k \delta_{ik} \left(\frac{1}{\hat{\alpha}_k} - \psi^{(1)}(\hat{\alpha}_k) \right). \end{aligned}$$

Which means our Jacobian $\mathbf{J} = \left(\frac{\partial g_k(\{\hat{\alpha}_j\})}{\partial \hat{\alpha}_i} \right)$ will be a diagonal matrix. Hence $\mathbf{J}^{-1} = \left(\left(\frac{\partial g_k(\{\hat{\alpha}_j\})}{\partial \hat{\alpha}_i} \right)^{-1} \right)$. If we let $\hat{\boldsymbol{\alpha}} = (\hat{\alpha}_i)$ and $\mathbf{F} = (g_k(\{\hat{\alpha}_j\}))$, then:

$$\hat{\boldsymbol{\alpha}}^{(l+1)} = \hat{\boldsymbol{\alpha}}^{(l)} - \mathbf{J}^{-1} \mathbf{F}.$$

Likelihood ratio

Therefore the likelihood ratio is:

$$\begin{aligned} \lambda &= \frac{L(\widehat{\Omega}_0)}{L(\widehat{\Omega})} \\ &= \frac{\left(\Gamma(\widehat{\alpha}) \left(\frac{\overline{Y}e}{\widehat{\alpha}}\right)^{\widehat{\alpha}}\right)^{-n} \prod_{i=1}^m \prod_{j=1}^{n_i} Y_{ij}^{\widehat{\alpha}-1}}{\left(\prod_{i=1}^m \left(\Gamma(\widehat{\alpha}_i) \left(\frac{\overline{Y}_ie}{\widehat{\alpha}_i}\right)^{\widehat{\alpha}_i}\right)^{-n_i}\right) \prod_{i=1}^m \prod_{j=1}^{n_i} Y_{ij}^{\widehat{\alpha}_i-1}} \\ &= \left(\Gamma(\widehat{\alpha}) \left(\frac{\overline{Y}}{\widehat{\alpha}}\right)^{\widehat{\alpha}}\right)^{-n} \left(\prod_{i=1}^m \left(\Gamma(\widehat{\alpha}_i) \left(\frac{\overline{Y}_i}{\widehat{\alpha}_i}\right)^{\widehat{\alpha}_i}\right)^{n_i}\right) \prod_{i=1}^m \prod_{j=1}^{n_i} Y_{ij}^{\widehat{\alpha}-\widehat{\alpha}_i} \end{aligned}$$

And we know that asymptotically under the null hypothesis $-2 \ln \lambda \sim \chi^2_{2m-2}$, so we will test $-2 \ln \lambda$ against the χ^2_{2m-2} distribution to find our p-value.