STAT 520 Homework 2

Yifan Zhu

September 15, 2017

1. Y_{ij} are independent, and $Y_{11}, Y_{12}, \dots, Y_{1n_1} \overset{i.i.d}{\sim} Gamma(\alpha_1, \beta_1)$, $Y_{21}, Y_{22}, \dots, Y_{2n_2} \overset{i.i.d}{\sim} Gamma(\alpha_2, \beta_2)$. We use α and β to note the parameter of Gamma distribution for one group and n be the size of that group. Then the log likelihood is

$$\ell(\alpha, \beta) = n(\alpha \log \beta) - \log \Gamma(\alpha) + (\alpha - 1)\sum_{i=1}^{n} \log y_i - \beta \sum_{i=1}^{n} y_i$$

2. By the property of mle, we can obtain the standard error for $\hat{\alpha}$ and $\hat{\beta}$ in this way.

By taking second derivative, we have

$$\frac{\partial^2 \ell}{\partial \alpha^2} = -n\frac{\Gamma''(\alpha)\Gamma'(\alpha) - \Gamma'(\alpha)^2}{(\Gamma'(\alpha))^2}$$
$$\frac{\partial^2 \ell}{\partial \beta^2} = -\frac{n\alpha}{\beta^2}$$
$$\frac{\partial \ell}{\partial \alpha \partial \beta} = \frac{n}{\beta}$$

Then we can get Fisher Information

$$I_n = -\mathbb{E}_{\alpha,\beta}(Hessian(\ell))$$

and

$$I_n^{-1} = \begin{bmatrix} \sigma_\alpha^2 & \sigma_{\alpha,\beta} \\ \sigma_{\alpha,\beta} & \sigma_\beta^2 \end{bmatrix}$$

Then we can get the 95% confidence interval for α and β from $\hat{\alpha} \pm 1.96\hat{\sigma}_\alpha$ and $\hat{\beta} \pm 1.96\hat{\alpha}_\beta$. The results are listed in the table.

	$\hat{\alpha}$	$\hat{\beta}$	CI for $\hat{\alpha}$	CI for $\hat{\beta}$
Group 1	4.56	2.41	(2.12, 7.01)	(1.04, 3.78)
Group 2	1.04	0.54	(0.53, 1.55)	(0.20, 0.88)
Combined	1.65	0.87	(1.06, 2.25)	(0.50, 1.23)

3. Let ℓ_1 and ℓ_2 be log likelihood for group 1 and group 2. Then the log likelihood for the full model is

$$\ell_{full} = \ell_1(\hat{\alpha}_1, \hat{\beta}_1) + \ell_2(\hat{\alpha}_2, \hat{\beta}_2)$$

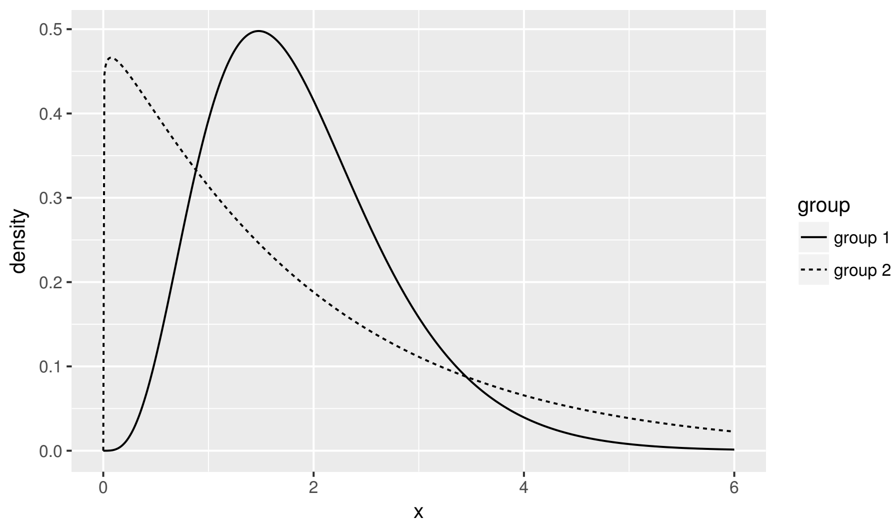
The log likelihood for redued model $\ell_{reduced}$ is that for the combined data. Thus we can get the deviance G^2

$$G^2 = -2(\ell_{reduced} - \ell_{full}) = 14.33$$

The difference in dimension of parameter is 2, thus $G^2 \dot\sim \chi_2^2$.

Meanwhile, $\chi^2_{1,0.9} = 2.7$, which is less than 14.33. Hence we reject the reduced model.

4. Graph as shown below.



5. Let μ_i be the expected value for group i, then $\mu_i = \frac{\alpha_i}{\beta_i} = g(\alpha_i, \beta_i)$. Here g is a continuous function. Then $(\widehat{\mu_1 - \mu_2}) = \hat{\mu}_1 - \hat{\mu}_2 = \frac{\hat{\alpha}_1}{\hat{\beta}_1} - \frac{\hat{\alpha}_2}{\hat{\beta}_2}$.

Then we need $V\hat{a}r(\hat{\mu}_1 - \hat{\mu}_2)$. Because of independece, $Var(\hat{\mu}_1 - \hat{\mu}_2) = Var(\hat{\mu}_1) + Var(\hat{\mu}_2)$. For any group, let μ be its mean, $\mu = \alpha/\beta = g(\alpha, \beta)$. Let

$$D = \begin{bmatrix} \frac{\partial g}{\partial \alpha} \\ \frac{\partial g}{\partial \beta} \end{bmatrix} = \begin{bmatrix} \frac{1}{\beta} \\ -\frac{\alpha}{\beta^2} \end{bmatrix}$$

Then by Delta Method, $Var(\hat{\mu}) = D^T I_n^{-1} D$. By pluging in α_i, β_i, we have $V\hat{a}r(\hat{\mu}_1) = 0.0314, V\hat{a}r(\hat{\mu}_2) = 0.1426$. Thus from $(\hat{\mu}_1 - \hat{\mu}_2) \pm 1.96(V\hat{a}r(\hat{\mu})_2 + V\hat{a}r(\hat{\mu}_2))$, we know

$$CI = (-0.854, 0.781)$$

6. Using the pooled sample variance by assuming equal variance of the two groups

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Then

$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{s_p^2/(n_1 + n_2)}} = -0.1782$$

The t statistic here has a degree of freedom of 48, and p-value is 0.5704. We fail to reject the null hypothesis that $\mu_1 = \mu_2$. Thus it disagree with the results in 3, but since the confidence interval in 5 contains 0, thus it agrees with the result in 5.

7. In this problem, we want to test if these two groups follow the same Gamma distribution. But in 5 and 6, the test is only for the difference in mean of these two groups. It restricts the cause of difference to be the difference in means, which I don't think is reasonable. In 3 we do not have such restriction. It tests if these two Gamma distribution is actually one Gamma distribution. This may also be the reason why 3 do not gave the same conclusion like 5 and 6. Hence I think the result from 3 is better for this problem and conclude that there is evidence taht these two Gamma distributions are different.