An ecologist is interested in testing if the density of a certain plant species at a site in northern Wisconsin has changed in the past 50 years. Data comes from sampled one-meter-square quadrats where in each quadrat the number of individuals of the plant species in each quadrat is recorded. In a sample from fifty years ago, the mean number of plants per quadrat was 1.3 in a sample of 20 quadrats. In a recent sample of 50 quadrats, the mean was 0.98. The biologist is willing to assume a Poisson model for the number of plants in each quadrat and assumes that these counts are independent.

Conduct a likelihood ratio test of the null hypothesis that the means are the same for the two times versus the alternative that the more recent time has a lower mean.

Solution: We begin by writing an expression for the likelihood. Let $s_1 = (x_1, \ldots, x_{20})$ be the data from the first sample and let $s_2 = (y_1, \ldots, y_{50})$ be the data from the second sample. Let θ_1 and θ_2 be the two means of the Poisson distributions for the respective samples. Then, the likelihood for the complete data set is

$$\begin{aligned} L(\theta_1, \theta_2 \mid s_1, s_2) &= \left(\prod_{i-1}^{20} \frac{\mathrm{e}^{-\theta_1} \theta_1^{x_i}}{x_i!} \right) \left(\prod_{i-1}^{50} \frac{\mathrm{e}^{-\theta_2} \theta_2^{y_i}}{y_i!} \right) \\ &= \frac{\mathrm{e}^{-20\theta_1 - 50\theta_2} \theta_1^{20\bar{x}} \theta_2^{50\bar{y}}}{\left(\prod_{i-1}^{20} x_i! \right) \left(\prod_{i-1}^{50} y_i! \right)} \end{aligned}$$

for the two-dimensional parameter space $\Omega = \{(\theta_1, \theta_2) : \theta_1, \theta_2 > 0\}$. The corresponding log-likelihood is

$$\ell(\theta_1, \theta_2 \mid s_1, s_2) = -20\theta_1 - 50\theta_2 + 20\bar{x} \log \theta_1 + 50\bar{y} \log \theta_2 - \sum_{i-1}^{20} \log(x_i!) - \sum_{i-1}^{50} \log(y_i!) \ .$$

Under the null hypothesis restriction $\theta_1 = \theta_2$, this simplifies to

$$\ell_0(\theta \mid s_1, s_2) = -70\theta + (20\bar{x} + 50\bar{y}) \log \theta - \sum_{i-1}^{20} \log(x_i!) - \sum_{i-1}^{50} \log(y_i!) \ .$$

Note that

$$\frac{\partial \ell(\theta_1, \theta_2 \mid s_1, s_2)}{\partial \theta_1} = -20 + \frac{20\bar{x}}{\theta_1} = 0$$

which has solution $\hat{\theta}_1 = \bar{x} = 1.3$ and a similar calculation shows that $\hat{\theta}_2 = \bar{y} = 0.98$. The corresponding log-likelihood evaluated at the MLE is

$$\max_{\theta_1, \theta_2} \ell(\theta_1, \theta_2 \mid s_1, s_2) = -75 + 26 \log(1.3) + 49 \log(0.98) - \sum_{i-1}^{20} \log(x_i!) - \sum_{i-1}^{50} \log(y_i!) = -69.16846 + C$$

Under the null hypothesis, the maximum likelihood estimate is $\hat{\theta} = (20\bar{x} + 50\bar{y})/70 = (26 + 49)/70 = 1.07$ and the likelihood is maximized at

$$\max_{\theta} \ell_0(\theta \mid s_1, s_2) = -75 + 75 \log(75/70) - \sum_{i-1}^{20} \log(x_i!) - \sum_{i-1}^{50} \log(y_i!) = -69.82553 + C.$$

The likelhood ratio test statistic is then $-2\Lambda = 2(69.82553 - 69.16846) = 1.314$.

Here is a formal statement of the hypothesis test. A test of the null hypothesis $\theta_1 = \theta_2$ versus the alternative hypothesis $\theta_1 > \theta_2$ by a likelihood ratio test has p-value 0.126 which is half the area to the right of 1.314 under a chi-square distribution with 1 degree of freedom. There is very little evidence in this data that the density of the plant species has decreased in the past fifty years — the observed difference in sample means is consistent with chance variation.