Problem Section 6

Likelihood Inference

Exercises

1. Suppose the number of students who arrive late to each of 10 consecutive lectures can be modeled as independent draws from a Poisson distribution with rate λ_0:

$$f(x) = \frac{e^{-\lambda_0}\,\lambda_0^x}{x!} \quad x = 0, 1, 2, \ldots$$

The data are observed as follows: $x_1 = 0, x_2 = 0, x_3 = 0, x_4 = 1, x_5 = 1, x_6 = 1, x_7 = 2, x_8 = 2, x_9 = 3, x_{10} = 3$.

a. Find the MLE $\widehat{\lambda}_0^{mle}$. (Hint: write the log likelihood function $\ell(\lambda)$ in terms of the x's and only plug in the values after you have solved the first order equation.)

The expression for the MLE of λ_0 is $\widehat{\lambda}_0^{mle} = \frac{\sum\limits_{i=1}^{10} x_i}{10} = \bar{x}$.

```
x <- c(0, 0, 0, 1, 1, 1, 2, 2, 3, 3)
lambda_mle <- mean(x)
```

The MLE of λ_0 is $\widehat{\lambda}_0^{mle} = 1.3$.

b. Find the observed information $-\ell''(\widehat{\lambda}_0^{mle})$.

As computed $\ell''(\lambda) = \frac{-10\bar{x}}{\lambda^2}$. Thus $-\ell''(\widehat{\lambda}_0^{mle}) = \frac{10\bar{x}}{(\widehat{\lambda}_0^{mle})^2} = \frac{10}{\bar{x}} = \frac{10}{1.3} \approx 7.692 = \frac{100}{13}$.

c. Use the large sample normality of the MLE to find an approximate 95% Wald confidence interval for λ_0.

As found above, $\widehat{\lambda}_0^{mle} = 1.3$, $\hat{Var}(\widehat{\lambda}_0^{mle}) = \frac{1}{\frac{100}{13}} = \frac{13}{100}$. Thus the asymptotic standard error is $\hat{SD} = \sqrt{\frac{13}{100}} = \frac{\sqrt{13}}{10}$.

```
asp_se <- sqrt(13) / 10
z_quant <- qnorm(0.975)

data.frame(lower = lambda_mle - z_quant * asp_se,
           upper = lambda_mle + z_quant * asp_se)
```

```
##       lower    upper
## 1 0.5933249 2.006675
```

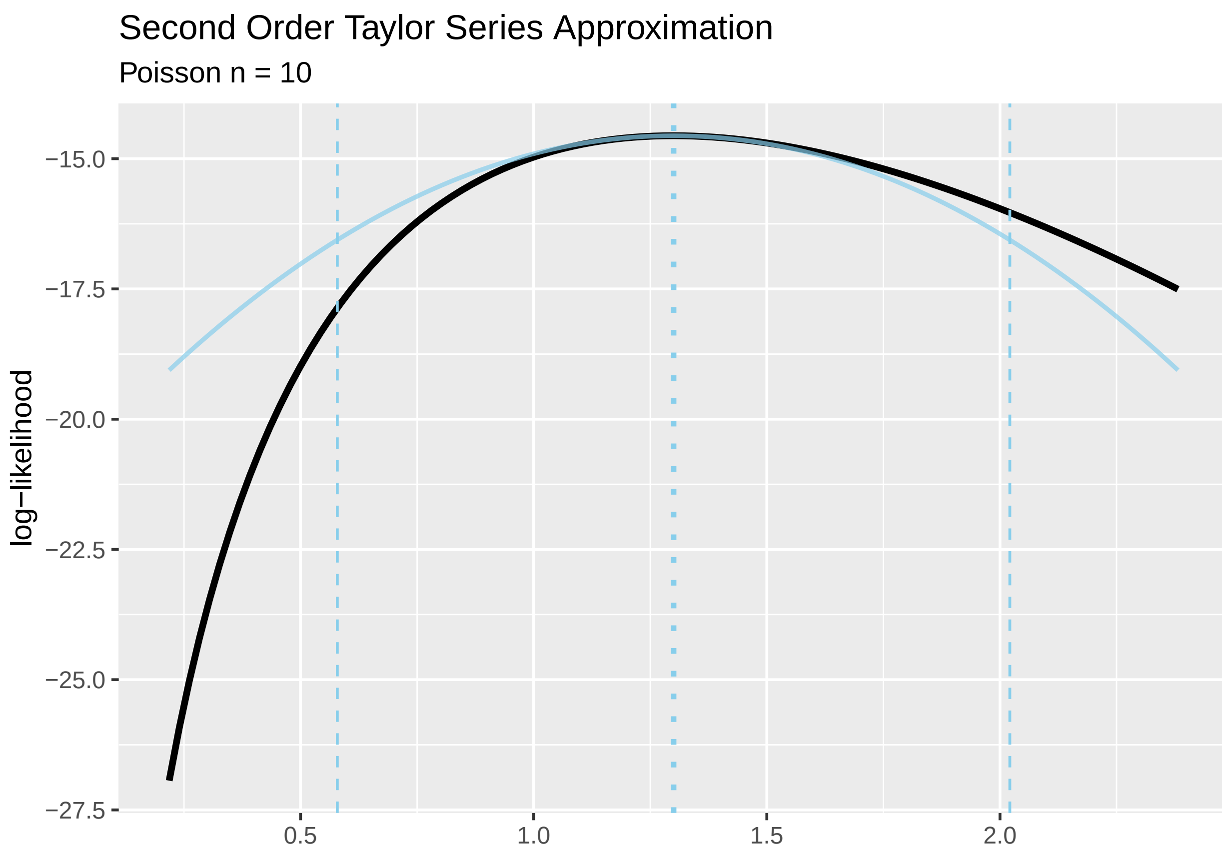
Thus as computed above the 95% wald confidence interval for λ_0 is [0.593, 2.007].

d. Examine the quality of the second-order approximation to the log-likelihood.

```
loglik_pois <- function(lambda) {
  ifelse(lambda < 0,
         NA,
         -10 * lambda + sum(x)*log(lambda) - log(prod(factorial(x)))
         )
}
```

```
maxlik_pois <- maxLik2(loglik = loglik_pois,
                       start = lambda_mle,
                       method = "NR")

plot(maxlik_pois) %>% gf_labs(title = "Second Order Taylor Series Approximation",
                              subtitle = "Poisson n = 10")
```



e. Suppose we wish to test $H_0 : \lambda_0 = \lambda_0^{null}$ versus $H_1 : \lambda_0 \neq \lambda_0^{null}$. Give an expression for

$$W = 2\ln\left[\frac{L(\widehat{\lambda}_0^{mle})}{L(\widehat{\lambda}_0^{null})}\right],$$

the likelihood ratio test statistic.

The expression for W is $2\ln(e^{-13}(\frac{13}{10})^{13}) - 2\ln(e^{-10\lambda_0^{null}}(\lambda_0^{null})^{13})$

```
W <- 2 * (loglik_pois(lambda_mle) - loglik_pois(1))
```

The likelihood ratio test statistic for this hypothesis test is 0.8214709.

f. Calculate the (large sample) P-value testing $\lambda_0^{null} = 1$. That is assume $W \sim \chi_1^2$ under the null hypothesis.

```
pval <- pchisq(q = W, df = 1, lower.tail = F)
```

The p-value is thus 0.3647505.

g. Since $n = 10$ is not really a large sample, an alternative approach is to calculate an empirical P-value. Take a look at chapter 20.4 where I show you how to calculate one for a Poisson model. Write code below to find the empirical P-value for the likelihood ratio test statistic W from part e.

```
set.seed(175)
```