Exercícios sobre Verossimilhança - VI

Henrique Ap. Laureano

henriquelaureano@outlook.com $\wedge$ http://mynameislaure.github.io/

/UFPR/DEST/LEG/

Junho de 2019

---

1. Seja X uma v.a. com distribuição uniforme $X \sim U(0, \theta)$.
   Uma amostra aleatória forneceu os seguintes valores: 5,5; 3,2; 4;8, 5,3; 3,8 e 5,0. Obtenha a função de verossimilhança e um (ou mais) intervalo(s) adequado(s) especificando a forma de obtenção.

```r
# <r code> -----
data <- c(5.5, 3.2, 4.8, 5.3, 3.8, 5)
# </r code> -----
```

Solution:

The likelihood of a Uniform$(0, \theta)$ is

$$
\begin{aligned}
L(\theta) &= \theta^{-n}, \quad \text{for } x_i < \theta \text{ for all } i \\
&= \theta^{-n}, \quad \text{for } \theta > x_{(n)},
\end{aligned}
$$

and equal to zero otherwise.
Given the data above, we get $x_{(n)} = 5.5$ and the likelihood is shown in Figure 1, together with some intervals.

```r
# <r code> -----
lkl.unif <- function(theta, data) {
  n <- length(data) ; xn <- max(data)
  # this next step, in the way presented, isn't very efficient, since I'll
  # compute the likelihood for all theta's and then, if necessary, convert to
  # zero. however, since we have here a very simple likelihood (and this
  # notation is extremely clear, btw), I keep in this way
  lkl = theta**(-n)*{theta >= xn}
  return(lkl)
}
theta.seq <- seq(4, 12, length.out = 100)
lklseq.unif <- lkl.unif(theta.seq, data)

par(mfrow = c(1, 2), mar = c(4, 4, 2, 2) + .1) # cosmetics
# plotting the normalized likelihood function to have unit maximum
plot(theta.seq, lklseq.unif/max(lklseq.unif), type = "l",
     xlab = expression(theta), ylab = "Likelihood")

# probability-based interval ------------------------------------------------
# cutoff's corresponding to a 95\% and 99\% confidence interval for the mean
cuts <- c(.15, .04) * lkl.unif(max(data), data) / max(lklseq.unif)
abline(h = cuts, lty = 2:3)

# uniroot.all finds the zeros, so we need to subtract the cutoff point from
# the likelihood to be able to find the points where the likelihood is cut
ic.lkl.unif <- function(theta, cut, ...) {
  lkl.unif(theta, ...)/max(lklseq.unif) - cut
}
ic.95.prob <- rootSolve::uniroot.all(ic.lkl.unif, range(theta.seq),
                                     data = data, cut = cuts[1])
ic.99.prob <- rootSolve::uniroot.all(ic.lkl.unif, range(theta.seq),
                                     data = data, cut = cuts[2])
abline(v = ic.95.prob, lty = 2) ; abline(v = ic.99.prob, lty = 3)

plot(theta.seq, lklseq.unif/max(lklseq.unif), type = "l",
     xlab = expression(theta), ylab = "Likelihood")
# pure likelihood interval --------------------------------------------------
# cutoff's corresponding to a 95\% and 99\% confidence interval for the mean
cuts <- c(.05, .01) * lkl.unif(max(data), data) / max(lklseq.unif)
abline(h = cuts, lty = 2:3)

ic.95.pure <- rootSolve::uniroot.all(ic.lkl.unif, range(theta.seq),
                                     data = data, cut = cuts[1])
ic.99.pure <- rootSolve::uniroot.all(ic.lkl.unif, range(theta.seq),
                                     data = data, cut = cuts[2])
abline(v = ic.95.pure, lty = 2) ; abline(v = ic.99.pure, lty = 3)
# </r code> -----
```

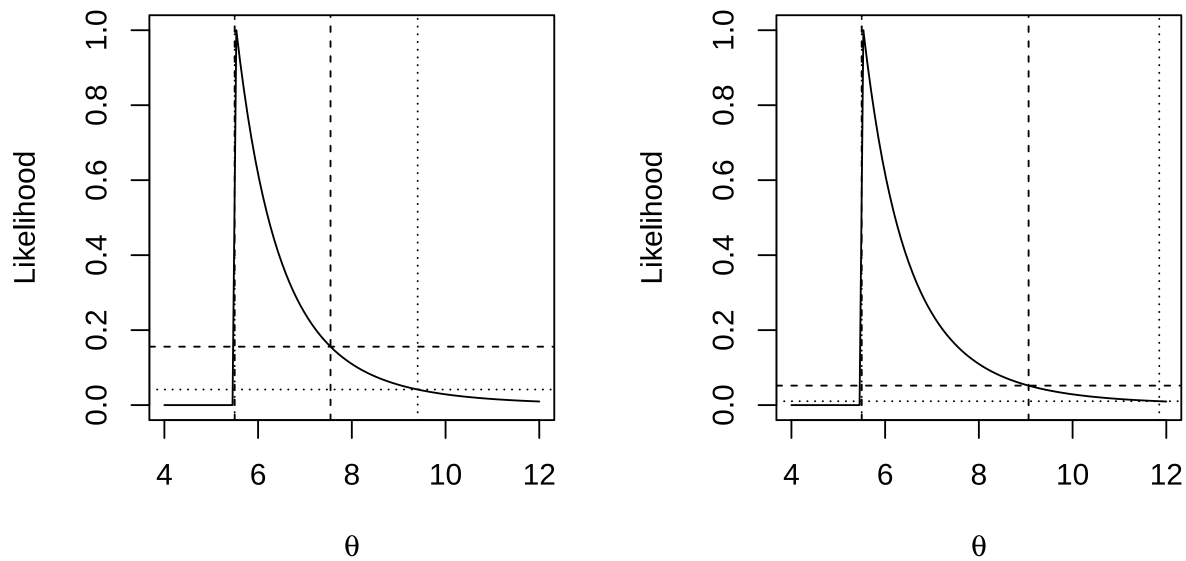

Figure 1: Likelihood function of θ in Uniform(0, θ) based on $x_{(6)} = 5.5$. In the left, likelihood intervals at 15% (dashed line) and 4% (dotted line) cutoff. In the right, likelihood intervals at 5% (dashed line) and 1% (dotted line) cutoff. These cutoffs correspond, in both sides, to a 95% and 99% confidence interval, respectively.

In both graphs of Figure 1 we have two confidence intervals for θ, one corresponding to a 95% confidence interval (dashed line) and one corresponding to a 99% confidence interval for the mean (dotted line). In the left we have a probability-based likelihood interval, in the right we have a pure likelihood interval.

The left ones are bad intervals, since they're based in a large-sample theory that results in an exact interval for the Gaussian case, in a good approximation for a reasonably regular case, and as we can see by the figure... here we doesn't have a regular likelihood (the likelihood isn't well approximated by a quadratic function), thus, here this interval should not be so good.

Explaining better how we arrived in these intervals:

For a normalized likelihood, $L(\theta)/L(\hat{\theta})$, we have the following Wilk's likelihood ratio statistic, defined as W

$$W \equiv 2 \log \frac{L(\hat{\theta})}{L(\theta)} \sim \chi_1^2.$$

Its χ^2 distribution is exact only in the normal mean model, and approximately true when the likelihood is reasonably regular.

Based in this Wilk's statistic we are able to find the probability that the likelihood interval covers θ,

$$\Pr\left\{\frac{L(\theta)}{L(\hat{\theta})} > c\right\} = \Pr\left\{2\log\frac{L(\theta)}{L(\hat{\theta})} < -2\log c\right\} = \Pr\{\chi_1^2 < -2\log c\}.$$

So, if for some $0 < \alpha < 1$ we choose a cutoff

$$c = e^{-\frac{1}{2}\chi^2_{1,(1-\alpha)}},$$

where $\chi^2_{1,(1-\alpha)}$ is the $100(1-\alpha)$ percentile of χ_1^2.
For $\alpha = 0.05$ and 0.01 we have a cutoff $c = 0.15$ and 0.04, that corresponds to a 95% and 99% confidence interval, respectively. $\square$

Having the cutoff values we need now to find the interval values, i.e., in which points the cutoff horizontal line cuts the likelihood. For this purpose we use the function `rootSolve::uniroot.all`.
A likelihood interval at 15% and 4% cutoff for θ are (5.5, 7.545) and (5.5, 9.405).

Already on the right side of Figure 1 we have a more coherent, let's say, confidence interval for the Uniform likelihood. While the likelihood isn't regular, it is still possible to provide an exact theoretical justification for a confidence interval interpretation. Now

$$\Pr\left\{\frac{L(\theta)}{L(\hat{\theta})} > c\right\} = \Pr\left\{\frac{X_{(n)}}{\theta} > c^{1/n}\right\} = 1 - \Pr\left\{\frac{X_{(n)}}{\theta} < c^{1/n}\right\} = 1 - (c^{1/n})^n = 1 - c.$$

So the likelihood interval with cutoff c is a $100(1-c)\%$ confidence interval.
A likelihood interval at 15% and 4% cutoff for θ are (5.5, 9.062) and (5.5, 11.849).

Comparing the obtained intervals we see a broader range with the pure likelihood intervals. In other words, with the pure likelihood intervals we see a higher uncertainty than with the (not so recommended here) probability-based likelihood intervals.

For doing this exercise I read and used it Pawitan's book:

```
@book{pawitan,
  author    = {Yudi Pawitan},
  title     = {In All Likelihood:
               Statistical Modelling and Inference Using Likelihood},
  year      = {1991},
  publisher = {Oxford University Press},
  address   = {Great Clarendon Street, Oxford OX2 6DP},
}
```

2. Seja X uma v.a. de uma distribuição de Poisson ($X \sim P(\lambda)$)) para a qual foi obtida a seguinte amostra aleatória: (3, 1, 0, 2, 1, 1, 0, 0).

```
# <r code> -----
data <- c(3, 1, 0, 2, 1, 1, 0, 0)
# </r code> -----
```

(a) Obtenha a função de verossimilhança, sua aproximação quadrática e intervalos para λ.

Solution:

The likelihood and the log-likelihood of a Poisson(λ) is

$$L(\lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!} \quad \text{and} \quad \log L(\lambda) = l(\lambda) = -n\lambda + \log\lambda \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \log x_i!.$$

We work here with the log-likelihood to make our life easier, since the computations are much simpler in the log scale.

By computing the Score (derivative of $l(\lambda)$ wrt to λ) and making it equal to zero, we find the MLE $\hat{\lambda} = \bar{x}$. Computing the second derivative we find the observed information

$$I_O(\lambda) = \frac{n\bar{x}}{\lambda^2} \quad \rightarrow \quad I_O(\hat{\lambda}) = \frac{n}{\bar{x}}.$$

A graph of the Poisson log-likelihood is provided in black in Figure 2.

Doing a quadratic approximation (Taylor expansion of second order) in $l(\lambda)$ around $\hat{\lambda}$ we have

$$\begin{aligned} l(\lambda) &\approx l(\hat{\lambda}) + (\lambda - \hat{\lambda})l'(\hat{\lambda}) + \frac{1}{2}(\lambda - \hat{\lambda})^2 l''(\hat{\lambda}) \\ &\quad \text{(the Score is zero at the MLE)} \\ &= l(\hat{\lambda}) + \frac{1}{2}(\lambda - \hat{\lambda})^2 l''(\hat{\lambda}) = l(\hat{\lambda}) + \frac{1}{2}(\lambda - \hat{\lambda})^2 I_O(\hat{\lambda}). \end{aligned}$$

A graph of the quadratic approximation of the Poisson log-likelihood is provided in red in Figure 2. There we can see how good is the approximation around the maximum likelihood estimator (MLE). Here the sample size is very small, the idea is that as the sample size increase the quality, in this case the range, of the approximation also increase. This should happen because as the sample size increase the Poisson likelihood should be more symmetric.

In the topright graph of Figure 2 we have two intervals for λ.

```r
# <r code> -----
# likelihood -----------------------------------------------------------
lkl.poi <- function(lambda, data) {
  n <- length(data)
  # log-likelihood ignoring irrelevant constant terms
  lkl = -n * lambda + sum(data) * log(lambda)
  return(lkl)
}
lambda.seq <- seq(0, 4.5, length.out = 100)
lklseq.poi <- lkl.poi(lambda.seq, data)
par(mar = c(4, 4, 2, 2) + .1) # cosmetics
layout(matrix(c(1, 3, 2, 3), 2, 2), heights = c(1.5, 1)) # more cosmetics
plot(lambda.seq, lklseq.poi, type = "l",
     xlab = expression(lambda), ylab = "log-likelihood")
lambda.est <- mean(data) # MLE
abline(v = lambda.est, lty = 3, col = 4)
# quadratic approximation ----------------------------------------------
quadprox.poi <- function(lambda, lambda.est, data) {
  n <- length(data)
  obs.info <- n / lambda.est # observed information
  lkl.poi(lambda.est, data) - .5 * obs.info * (lambda - lambda.est)**2
}
curve(quadprox.poi(x, lambda.est, data), col = 2, add = TRUE)
legend(2.4, -6.25, c("log-like", "Quadratic\napprox.", "MLE"),
       col = c(1, 2, 4), lty = c(1, 1, 3), bty = "n")
# intervals for lambda -------------------------------------------------
plot(lambda.seq, lklseq.poi, type = "l",
     xlab = expression(lambda), ylab = "log-likelihood")
curve(quadprox.poi(x, lambda.est, data), col = 2, add = TRUE)
abline(v = lambda.est, col = 4, lty = 3)
## probability-based interval ------------------------------------------
# cutoff's corresponding to a 95\% confidence interval for the mean
cut <- log(.15) + lkl.poi(lambda.est, data) ; abline(h = cut, lty = 2)
ic.lkl.poi <- function(lambda, cut, ...) {
  lkl.poi(lambda, ...) - cut
}
ic.95.prob <- rootSolve::uniroot.all(ic.lkl.poi, range(lambda.seq),
                                     data = data, cut = cut)
arrows(x0 = ic.95.prob, y0 = rep(cut, 2),
       x1 = ic.95.prob, y1 = rep(-25, 2), lty = 2, length = .1)
## wald confidence interval  -------------------------------------------
cut <- log(.15) + quadprox.poi(lambda.est, lambda.est, data)
abline(h = cut, lty = 2, col = 2)
se.lambda.est <- sqrt(lambda.est / length(data))
wald.95 <- lambda.est + qnorm(c(.025, .975)) * se.lambda.est
arrows(x0 = wald.95, y0 = rep(cut, 2),
       x1 = wald.95, y1 = rep(-25, 2), lty = 2, length = .1, col = 2)
```

```r
# a better look in the approximation around the MLE ----------------------
lambda.seq <- seq(.8, 1.2, length.out = 25)
plot(lambda.seq, lkl.poi(lambda.seq, data), type = "l",
     xlab = expression(lambda), ylab = "log-likelihood")
curve(quadprox.poi(x, lambda.est, data), col = 2, add = TRUE)
abline(v = lambda.est, lty = 3, col = 4)
# </r code> -----
```

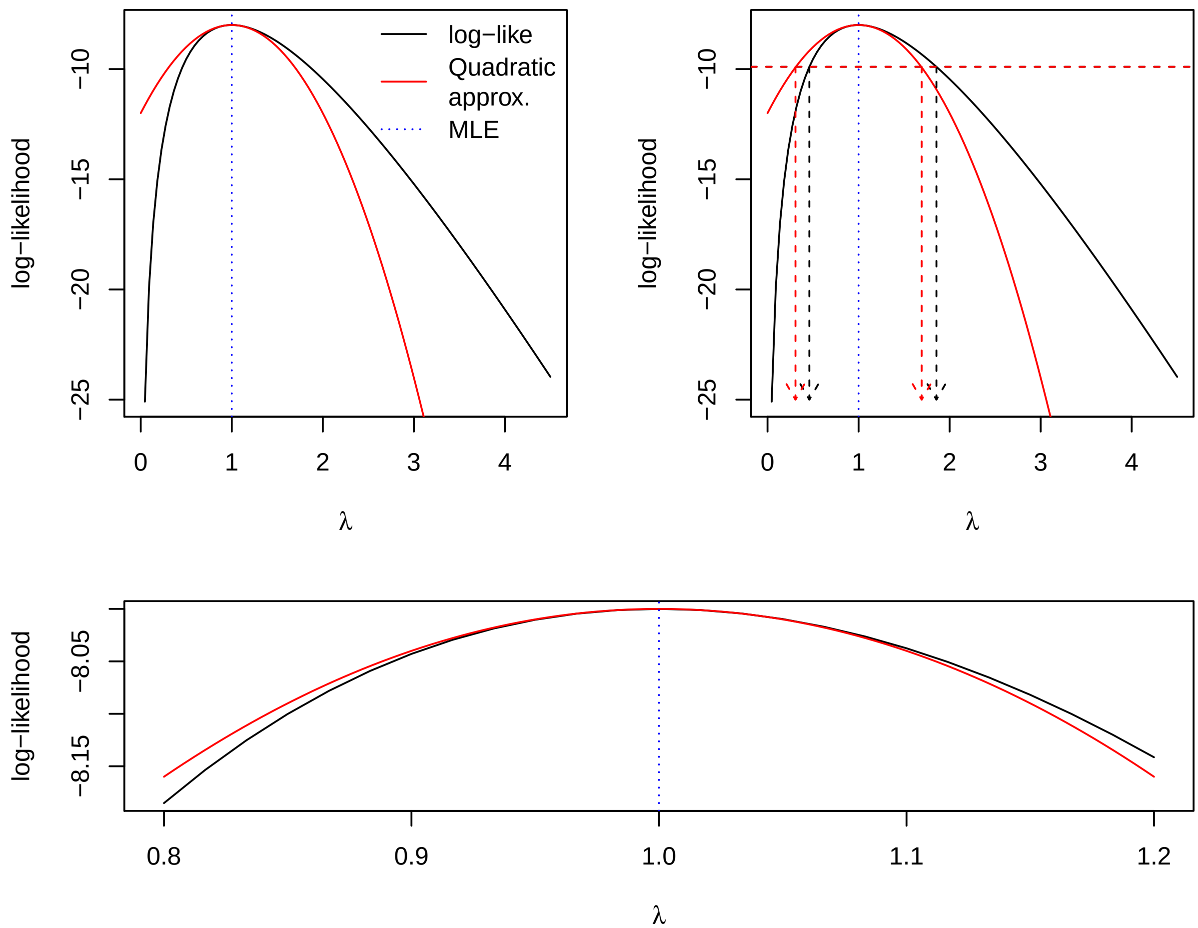

Figure 2: log-likelihood function and MLE of λ in Poisson(λ) based on `data`. In the topleft, a quadratic approximation in red. In the topright, two intervals for λ - one based in the likelihood (in black) and one based in the quadratic approximation (in red). In the bottom we provide a better look in the quadratic approximation.

In the topright, in black, we have a interval based in the likelihood, $\lambda \in (0.459, 1.855)$, but with a cutoff criterion based in a χ^2 distribution. Thus, to have a nominal 95% confidence interval we use a cutoff $c = 15\%$. As we saw before, this interval is based in a large-sample theory. Since we're dealing here with a reasonable regular case, this interval shows as a good approximation.

Still in the topright of Figure 2, in red we have a interval based in the quadratic approximation of the log-likelihood, $\lambda \in (0.307, 1.693)$. From the quadratic approximation we get

$$\log \frac{L(\lambda)}{L(\hat{\lambda})} \approx -\frac{1}{2} I_O(\hat{\lambda})(\lambda - \hat{\lambda})^2,$$

that also follows a χ^2 distribution, since we have a r.v. $\hat{\lambda}$ normalized (with its expected value subtracted and divided by its variance) and squared. From this we get the following exact (in the normal case) 95% confidence interval

$$\hat{\lambda} \pm 1.96 \; I_O(\hat{\lambda})^{-1/2} \qquad (\hat{\lambda} \pm 1.96 \; \text{se}(\hat{\lambda})).$$

In the nonnormal cases this is an approximate 95% CI.

The actual variance is $I_E(\hat{\lambda})^{-1/2}$, but for the Poisson case the Fisher (expected) information is equal to the observed one, $I_O(\hat{\lambda})^{-1/2}$.

A very nice thing that we can see from this intervals is that a Wald interval (a $\pm$ interval) corresponds to a cut in the quadratic approximation exactly in the same point that the probability-based interval cuts the (log-)likelihood. Thus, as more regular the likelihood, better will be the fit of the approximation and more reliable will be the Wald interval.

(b) Repita a questão anterior para a reparametrização $\theta = \log(\lambda)$.

Solution:

By the invariance property of the MLE we have

$$\hat{\lambda} = \bar{x} \quad \Rightarrow \quad g(\hat{\lambda}) = \log \hat{\lambda} = \hat{\theta} = \log \bar{x} = g(\bar{x}).$$

i.e., the MLE of $\hat{\theta}$ is $\log \bar{x}$.

By the Delta Method we compute the variance of $\hat{\theta}$,

$$V[\theta] = V[g(\lambda)] = \left[\frac{\partial}{\partial \lambda} g(\lambda)\right]^2 V[\lambda] = \left[\frac{1}{\lambda}\right]^2 \frac{\lambda}{n} = \frac{1}{\lambda \; n} \quad \rightarrow \quad V[\hat{\theta}] = \frac{1}{\bar{x} \; n}.$$

From this we can take the observed information for the reparametrization

$$V[\hat{\theta}] = I_O^{-1}(\hat{\theta}) = (\bar{x} \; n)^{-1}.$$

Now we do, in the same manner, everything that we did in the previous letter.

```
# <r code> -----
# likelihood -------------------------------------------------------------
## we can still use the lkl.poi function, but now we do
## lambda = exp(theta),
## with theta being a real rate (positive or negative)
theta.seq <- seq(-2, 2, length.out = 100)
lklseq.poi <- lkl.poi(exp(theta.seq), data)
```

```r
par(mar = c(4, 4, 2, 2) + .1) # cosmetics
layout(matrix(c(1, 3, 2, 3), 2, 2), heights = c(1.5, 1)) # more cosmetics
plot(theta.seq, lklseq.poi, type = "l",
     xlab = expression(theta), ylab = "log-likelihood")
theta.est <- log(mean(data)) # MLE
abline(v = theta.est, lty = 3, col = 4)
# quadratic approximation for the reparametrization ----------------------
quadprox.poi.repa <- function(theta, theta.est, data) {
  obs.info <- sum(data) # observed information
  lkl.poi(exp(theta.est), data) - .5 * obs.info * (theta - theta.est)**2
}
curve(quadprox.poi.repa(x, theta.est, data), col = 2, add = TRUE)

legend(-2.15, -27.5, c("log-like", "Quadratic\napprox.", "MLE"),
       col = c(1, 2, 4), lty = c(1, 1, 3), bty = "n")
# intervals for lambda ----------------------------------------------------
plot(theta.seq, lklseq.poi, type = "l",
     xlab = expression(theta), ylab = "log-likelihood")
curve(quadprox.poi.repa(x, theta.est, data), col = 2, add = TRUE)
abline(v = theta.est, col = 4, lty = 3)
## probability-based interval ---------------------------------------------
# cutoff's corresponding to a 95\% confidence interval for the mean
cut <- log(.15) + lkl.poi(exp(theta.est), data) ; abline(h = cut, lty = 2)
ic.lkl.poi <- function(theta, cut, ...) {
  lambda <- exp(theta)
  lkl.poi(lambda, ...) - cut
}
ic.95.prob <- rootSolve::uniroot.all(ic.lkl.poi, range(theta.seq),
                                     data = data, cut = cut)
arrows(x0 = ic.95.prob, y0 = rep(cut, 2),
       x1 = ic.95.prob, y1 = rep(-43, 2), lty = 2, length = .1)
## wald confidence interval  ----------------------------------------------
cut <- log(.15) + quadprox.poi.repa(theta.est, theta.est, data)
abline(h = cut, lty = 2, col = 2)

se.theta.est <- sqrt(1 / sum(data))
wald.95 <- theta.est + qnorm(c(.025, .975)) * se.theta.est

arrows(x0 = wald.95, y0 = rep(cut, 2),
       x1 = wald.95, y1 = rep(-43, 2), lty = 2, length = .1, col = 2)
# a better look in the approximation around the MLE ----------------------
theta.seq <- seq(-.75, .75, length.out = 25)
plot(theta.seq, lkl.poi(exp(theta.seq), data), type = "l",
     xlab = expression(theta), ylab = "log-likelihood")
curve(quadprox.poi.repa(x, theta.est, data), col = 2, add = TRUE)
abline(v = theta.est, lty = 3, col = 4)
# </r code> -----
```

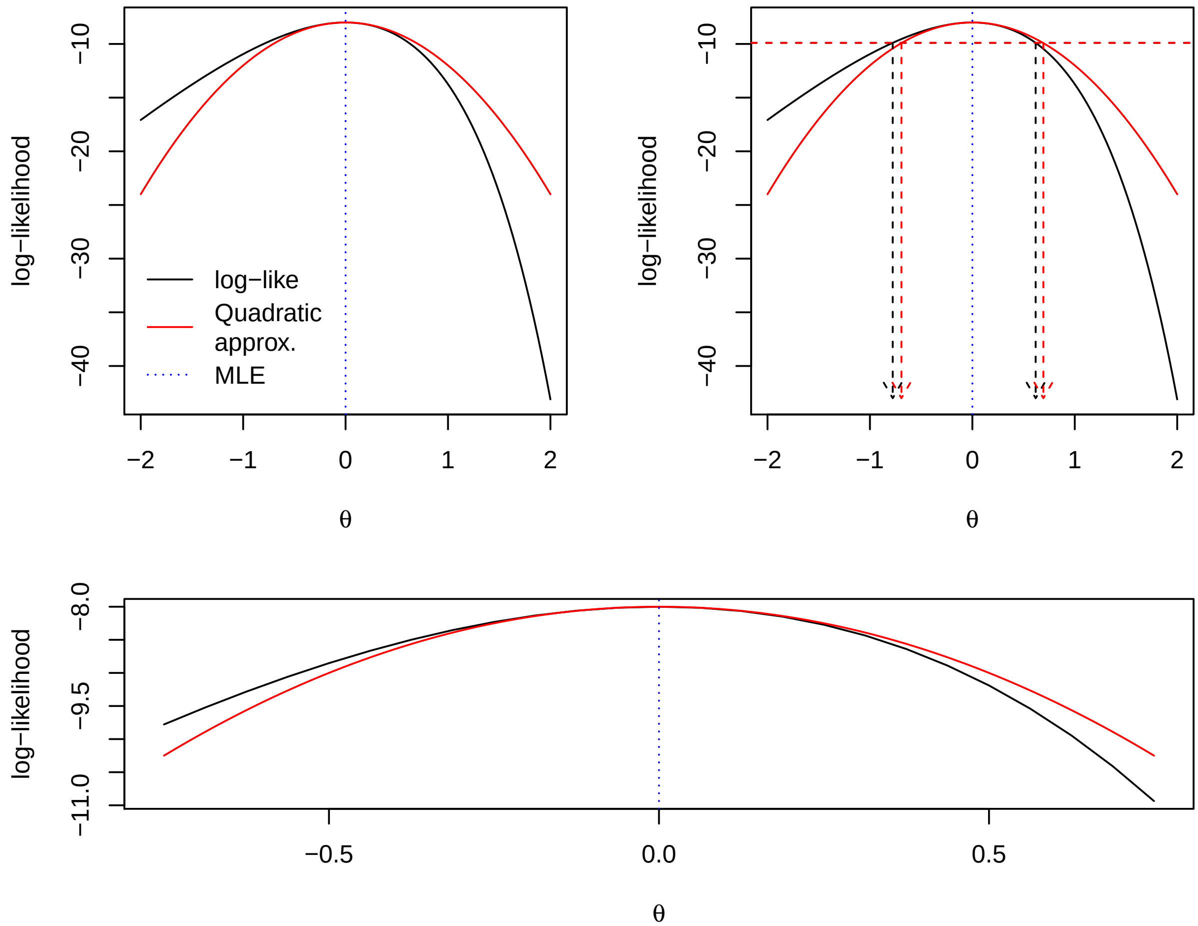

Figure 3: log-likelihood function and MLE of θ in Poisson($\lambda = e^{\theta}$) based on `data`. In the topleft, a quadratic approximation in red. In the topright, two intervals for θ - one based in the likelihood (in black) and one based in the quadratic approximation (in red). In the bottom we provide a better look in the quadratic approximation.

We obtained here two intervals for θ. One based in a cut in the likelihood, $\theta \in (-0.778, 0.618)$ - in black on the topright of Figure 3, and one based in a cut in the quadratic approximation of the likelihood, $\theta \in (-0.693, 0.693)$ - in red on the topright of Figure 3.

With the parametrization $\theta = \log \lambda$ the two intervals are closer than the ones obtained for λ. i.e., for θ (with the use of the log) we get a more regular likelihood.

(c) Obtenha ainda (por pelo menos dois métodos diferentes) intervalos de confiança para o parâmetro λ a partir da função de verossimilhança (aproximada ou não) de θ.

Solution:

Since the MLE has the invariance property, a very simple idea is: take the obtained interval for θ and apply a transformation, $\lambda = e^{\theta}$. This simple idea is true for the interval obtained via likelihood function, as we can see in Figure 4. The obtained interval is the same that the one in letter `a)`

However, this idea doesn't work for the interval based in the quadratic approximation.

```
# <r code> -----
# likelihood ------------------------------------------------------------
theta.seq <- seq(-2, 2, length.out = 100)
lklseq.poi <- lkl.poi(exp(theta.seq), data)
par(mar = c(4, 4, 2, 2) + .1) # cosmetics
plot(theta.seq, lklseq.poi, type = "l",
     xlab = expression(theta), ylab = "log-likelihood")
abline(v = lambda.est, lty = 3, col = 4)
# quadratic approximation for the reparametrization ----------------------
curve(quadprox.poi.repa(x, theta.est, data), col = 2, add = TRUE)
legend(-2.15, -25, c("log-like", "Quadratic\napprox.", "MLE"),
       col = c(1, 2, 4), lty = c(1, 1, 3), bty = "n")
# intervals for lambda ---------------------------------------------------
## probability-based interval --------------------------------------------
arrows(x0 = exp(ic.95.prob), y0 = c(-9, -36.5),
       x1 = exp(ic.95.prob), y1 = rep(-43, 2), lty = 2, length = .1)
## wald confidence interval  ---------------------------------------------
arrows(x0 = exp(wald.95), y0 = c(-9, -24),
       x1 = exp(wald.95), y1 = rep(-43, 2), lty = 2, length = .1, col = 2)
# </r code> -----
```

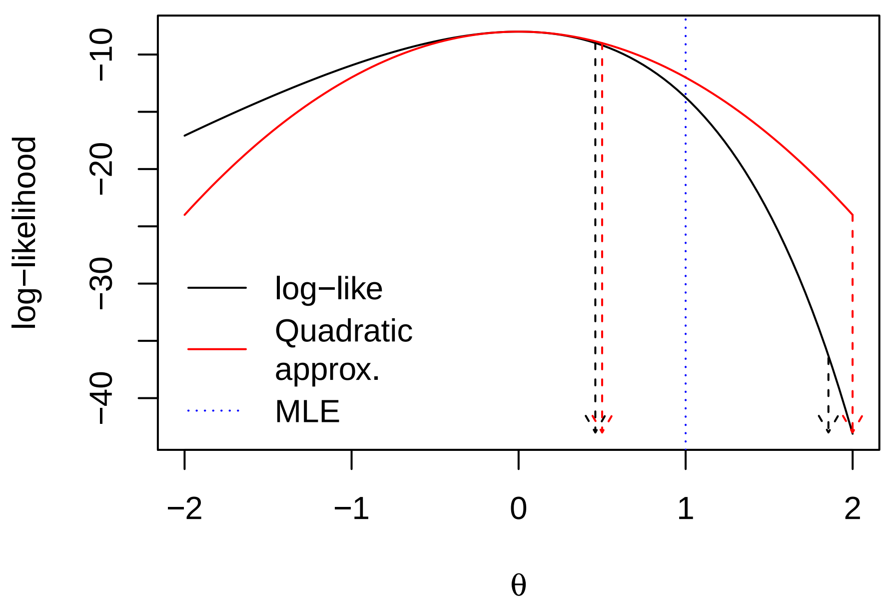

Figure 4: log-likelihood function and quadratic approximation of θ, and MLE of λ in Poisson($\lambda = e^{\theta}$) based on `data`. In dashed, 95% confidence intervals for λ.

In the letter `a)`, via the quadratic approximation we got $\lambda \in (0.307, 1.693)$. Here, applying the relation $\lambda = e^{\theta}$ we get $\lambda \in (0.5, 2)$. i.e., this shows that the invariance property applies to the likelihood, not to the quadratic approximation of the likelihood.

3. A fim de se obter uma estimativa do público de um jogo sem utilizar dados de venda de ingressos ou registros das roletas do estádio, foram distribuídas camisas especiais para 300 torcedores sob condição que estes a utilizassem durante um jogo. Durante o jogo foram selecionados ao acaso 250 torcedores verificando-se 12 destes possuiam a camisa.

   a) Obtenha a função de verossimilhança para o número total de torcedores.

   Solution:

   We have here a random variable, let's say, X, represeting the number of observed successes, k. Here an observed success is select a fan using a special shirt. So, we have

$$X \sim \text{Hypergeometric}(N, K = 300, n = 250),$$

   with probability $p = K/N$.

   N is the population size, that we want estimate. K is the unknown number of fans using a special shirt, and n is the number of, randomly, selected fans.

$$\begin{aligned}
\Pr[X = k] = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}} \quad \rightarrow \quad \Pr[X = 12] &= \frac{\binom{300}{12}\binom{N-300}{250-12}}{\binom{N}{250}} \\
&= \frac{300!\ 250!}{288!\ 238!\ 12!} \frac{(N-300)!\ (N-250)!}{(N-538)!\ N!} \\
&= \text{constant} \times \frac{(N-300)!\ (N-250)!}{(N-538)!\ N!}.
\end{aligned}$$

   Since we already have a k, the likelihood is equal $\Pr[X = k]$.

$$\begin{aligned}
L(N) = \Pr[X = 12] &= \text{constant} \frac{(N-300)!\ (N-250)!}{(N-538)!\ N!} \\
l(N) = \log L(N) &= \log\ \text{constant} + \log\ \frac{(N-300)!\ (N-250)!}{(N-538)!\ N!} \\
&\approx \log\ \frac{(N-300)!\ (N-250)!}{(N-538)!\ N!}.
\end{aligned}$$

```r
# <r code> -----
# likelihood ------------------------------------------------------------
lkl.hypergeo <- function(N = 539, M) {
  size <- M - N
```

```r
  n1 <- n2 <- n3 <- n4 <- numeric(size)
  for (i in 1:size) {
    n1[i] = log(N - 300)
    n2[i] = log(N - 250)
    n3[i] = log(N - 538)
    n4[i] = log(N)
    N = N + 1
  }
  lkl <- sum(n1) + sum(n2) - sum(n3) - sum(n4)
  return(lkl)
}
compute.lkl <- function(grid, ...) {
  size.grid = length(grid)
  lkl.grid = numeric(size.grid)
  i = 1
  for (j in grid) {
    lkl.grid[i] = lkl.hypergeo(M = j)
    i = i + 1
  }
  return(lkl.grid)
}
par(mar = c(4, 4, 2, 2) + .1)
layout(matrix(c(0, 1, 1, 0,
                2, 2, 3, 3), nrow = 2, byrow = TRUE))
m.grid <- seq(3000, 17000, by = 20)
plot(m.grid, compute.lkl(m.grid), "l", xlab = "N", ylab = "log-like")
abline(v = 6250, lty = 2)
m.grid <- seq(539, 1e4, by = 20)
plot(m.grid, compute.lkl(m.grid), "l", xlab = "N", ylab = "log-like")
abline(v = 6250, lty = 2)
m.grid <- seq(5000, 8000, by = 10)
plot(m.grid, compute.lkl(m.grid), "l", xlab = "N", ylab = "log-like")
abline(v = 6250, lty = 2)
# </r code> -----
```

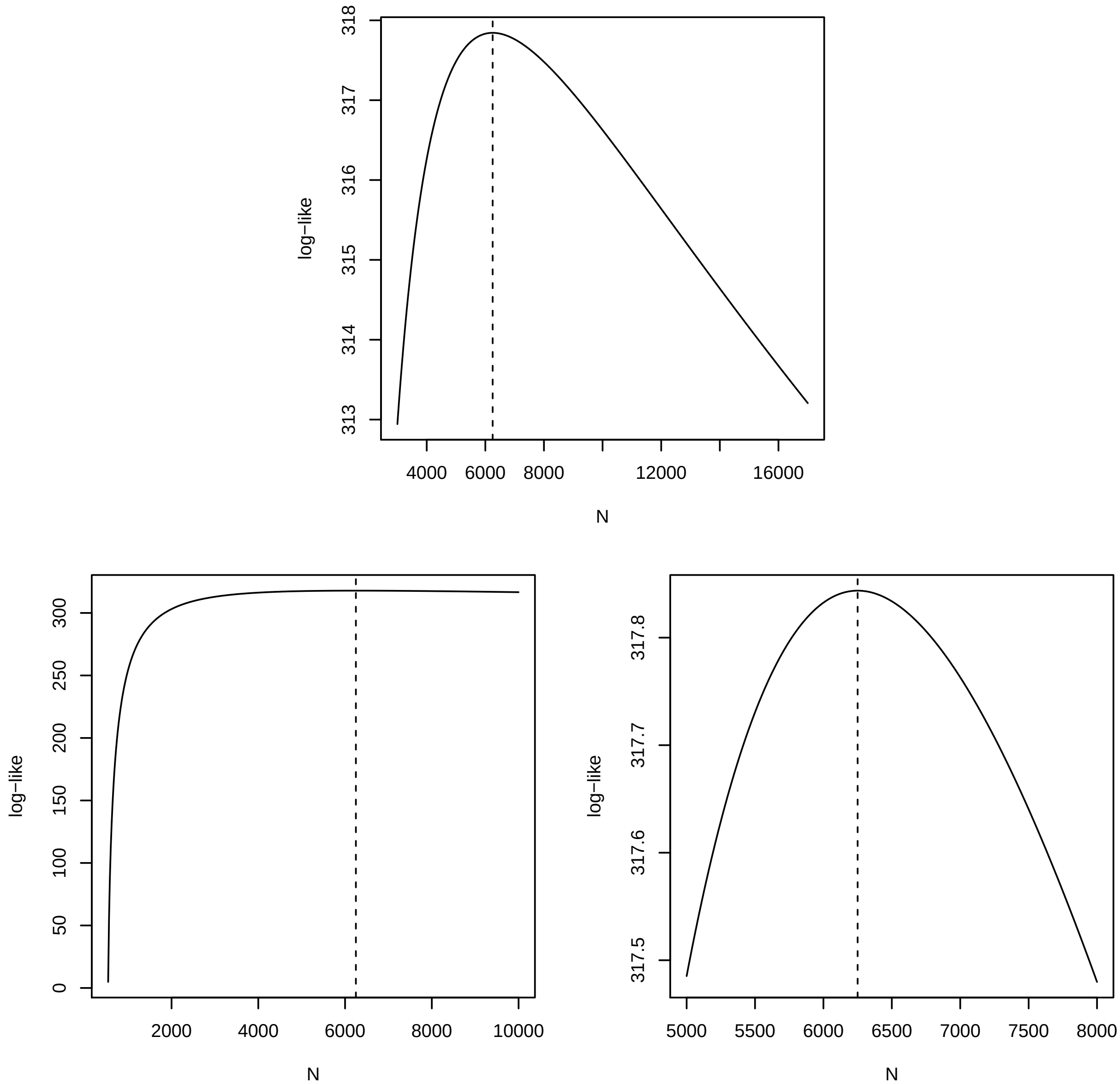

Figure 5: Hypergeometric likelihood.

b) Obtenha a estimativa pontual e a intervalar, esta última por pelo menos dois métodos diferentes.

c) Repita e compare os resultados caso fossem 500 camisas e 20 com camisas dentre os 250.

4. Foram tomadas as seguintes observações independentes de uma v.a. $X \sim N(\mu, \sigma^2)$.

- ```
  [1] 56.4 54.2 65.3 49.2 50.1 56.9 58.9 62.5 70.0 61.0
  ```

- sabe-se que outras 5 observações são menores que 50.
  ```

- sabe-se que outras 3 observações são maiores que 65.

(a) Escrever a função de verossimilhança.

(b) Obter as estimativas de máxima verossimilhança.

(c) Obter as verossimilhanças perfilhadas.

(d) Obter os erro-padrão das estimativas.

5. Considere os dados da tabela a seguir (adaptados/modificados de Montgomery & Runger, 1994) aos quais deseja-se ajustar um modelo de regressão linear simples relacionando a variável resposta Y (pureza em %) a uma variável explicativa X (nível de hidrocarbonetos).

X	0.99	1.02	1.15	1.29	1.46	1.36	0.87	1.23	1.55	1.40
Y	99.01	89.05	91.43	93.74	96.73	94.45	87.59	91.77	99.42	93.65
X	1.19	1.15	0.98	1.01	1.11	1.20	1.26	1.32	1.43	0.95
Y	93.54	92.52	90.56	89.54	89.85	90.39	93.25	93.41	94.98	87.33

(a) Encontre a função de verossimilhança.

(b) Encontre as estimativas de máxima verossimilhança.

(c) Obtenha a verossimilhança conjunta para os parâmetros β_0 e β_1:

i. considerando σ fixo com valor igual à sua estimativa.

ii. obtendo a verossimilhança (conjunta - 2D) perfilhada em relação a σ.

(d) Obtenha a verossimilhança perfilhada para os parâmetros β_0 e β_1 individualmente.

6. Considere uma amostra de uma v.a. Y em que assume-se que $Y_i \sim P(\lambda_i)$ em que o parâmetro λ_i é descrito por uma função de uma variável explicativa $\log(\lambda_i) = \beta_0 + \beta_1 x_i$ com valores conhecidos de x_i.
Dados simulados com $(\beta_0 = 2, \beta_1 = 0.5)$ são mostrados a seguir.

Y	10	15	11	37	70	19	12	12	13	88
X	1.7	1.5	0.5	2.8	4.4	1.8	0.4	0.7	1.3	5.1

(a) Obtenha o gráfico da função de verossimilhança indicando a posição dos valores verdadeiros dos parâmetros. text

(b) Obtenha os perfis de verossimilhança dos parâmetros.

(c) Obtenha estimativas.

(d) Obtenha intervalos (use diferentes métodos).

(e) Compare os resultados anteriores com os fornecidos pela função `glm()` e discuta os achados.

last modification on ...

```
[1] "2019-06-17 22:06:06 -03"
```