Jared Yu

Data Mining

Module 2 Assignment

1. Let have an exponential distribution
   1. Sketch versus for a fixed value of the parameter .

Ans:

The plot of versus is shown below in Figure 1 for . It was created in RStudio, where the range of -values are from -1 to 5. It shows the exponential distribution and how it is nonzero only for values of . There is a solid circle at coordinates and an unfilled circle at coordinates to indicate that for . The different values for for 0.5, 1, and 1.5 are denoted in blue, red, and black respectively. The plot shows how as increases, the density increases near zero, but decreases for larger values of . Also, as increases, the density tends to decrease.

A screenshot of a video game

Description automatically generated

Figure 1 A plot of vs. , where and . At point is an unfilled circle and at points , , and there are filled circles. This is to indicate that when .

* 1. Sketch versus , for a fixed value of .

Ans:

The plot of versus is shown below in Figure 2 for . It was created in RStudio, where the range of -values are from . An unfilled circle at is used to indicate that at that coordinate, the value of . The different values for for 1, 2, and 3 are denoted in black, red, and blue respectively. The plot shows how as increases, the density tends to peak and the curve down. The point at which it peaks is different depending on the value of . Furthermore, for larger values of , the density will be lower for smaller values of , before seeming to converge at larger values.

A close up of a map

Description automatically generated

Figure 2 A plot of vs. , where and . An unfilled circle is included at the coordinate (0,0) to indicate that at that point, the value of .

* 1. Suppose that samples are drawn independently according to . Show that the maximum likelihood estimate for is given by

Ans:

Let

Equation (1.1) shows the likelihood formula for the exponential function.

Equation (1.2) shows the log-likelihood version of equation (1).

In equation (1.3), the derivative of is taken and the goal is to find either its minimum or maximum by setting it to 0.

In equation (1.4), the optimum value of is shown to be . Then, given that at least one is nonzero:

In equation (1.5), the second derivative of is taken. This it to confirm whether it is a minimum or maximum. By plugging in the estimate from equation (1.4), it can be seen that this value is always less than zero. This implies that is a maximum point and so the MLE is .

1. Let have a uniform distribution
   1. Sketch versus for an arbitrary value of .

Ans:

The plot of versus is shown below in Figure 3 for , , and . It was created in RStudio, where the range of -values are from . There are solid circles at coordinates , , and . There are unfilled circles at coordinates , , and . These are to indicate that for respectively. The plot shows that for a fixed , the density is maximum when and starts to decrease as increases.

A close up of a map

Description automatically generated

Figure 3 A plot of vs. , where and . There are filled circles at coordinates , , and . There are unfilled circles at coordinates , , and . This is to indicate that when respectively.

* 1. Suppose that samples are drawn independently according to . Show that the maximum likelihood estimate for is .

Ans: (used reference [1])

Let .

The support of , , contains the parameter , therefore the method of finding the log likelihood will not work. Instead, it is possible to look at the following joint probability density function (p.d.f.),

where is an indicator function defined as follows,

When , where is the order statistic from the sample, then because that would violate the rules of the support. Therefore, the only case considered is . Then the joint p.d.f. will become,

It can be seen that when decreases, increases, therefore to maximize requires choosing the minimum value for , where . The MLE is , where is the largest value in the sample, or .

* 1. Find the method of moments estimator for .

Ans: (used reference [2])

Let .

In equation (2.1), the first theoretical moment is derived. Only one is needed, since only has a single parameter.

In equation (2.2), the first sample moment is calculated.

Equation (2.3) comes from equation the first theoretical moment about the origin with the first sample moment.

In Equation (2.4), it shows that the method of moments estimator for is .

1. Let be a binary vector with multivariate Bernoulli distribution

where is an unknown parameter vector, being the probability that . Show that the maximum likelihood estimate for is

Ans: (used reference [3])

Let .

Equation (3.1) shows the likelihood function for .

Equations (3.2) to (3.4) show the log-likelihood of equation (3.1).

where . Equation (3.5) shows the partial derivative of equation (3.4) w.r.t. a single , for . Then, setting , it is possible to find the optimum value for each .

Equations (3.6) to (3.10) show that the optimum value, for . To prove that it is a maximum, the Hessian matrix must be shown to be negative definite.

Equation (3.11) and (3.12) show the second derivatives for the diagonal and off-diagonal elements.

In equations (3.13), the second derivatives are seen in the Hessian matrix denoted . In equation (3.14), the values are inserted, where only the diagonal elements have distinctly nonzero values.

Let the vector , and .

Equation (3.15) then shows that , which means that the Hessian is negative-definite. So, the MLE of is . In other words, the MLE of is as follows,

This completes the proof of this problem.

1. Let have a Gamma distribution
   1. Suppose that samples are drawn independently according to . Find the method of moments estimator for and .

Ans: (used reference [4], [5], [6], [7])

The first step will be to find the moment generating function (M.G.F.) for the Gamma distribution.

Equation (4.1) shows the formula for the M.G.F. along with plugging in the p.d.f. of the Gamma distribution. It is important to note that the integral is finite only when , otherwise the integrand will increase towards infinity.

In equation (4.2), the is used to replace the part in the exponential function, so it can be seen that the integrand is similar to another p.d.f.

In equation (4.3), the previous is integrated to , however since it lacks the constant term , will be left as a constant.

In equation (4.4), the term is rewritten so that it is more convenient to substitute back into equation (4.3).

In equation (4.5), the term is substituted back into equation (4.4). It is then seen that the M.G.F. for the .

The next step is to find the first and second theoretical moments of by using the M.G.F. derived in equation (4.5).

Equation (4.6) shows the first derivative of the M.G.F. Equation (4.7) shows that substituting evaluates to the first theoretical moment being .

Equation (4.8) shows the second derivative of the M.G.F. Equation (4.9) shows that substituting evaluates to the second theoretical moment being .

The next steps are to find values for and based on the sample moments.

where

In equations (4.10) to (4.13), the goal is to solve for and based on and . Equation (4.15) is to show that . Where and are the first and second sample moments. These values were also plugged in at the end in equation (4.14). This is analogous to showing that .

* 1. Show that the exponential distribution is .

Ans:

The Gamma distribution has the following p.d.f.,

Plugging in and leads to the following,

which is equal to the exponential distribution. The constraints of are slightly different, but the probability is still 0 at so it does not make a significant difference.

References

[1] <http://www2.stat.duke.edu/~banks/111-lectures.dir/lect10.pdf>

[2] <https://online.stat.psu.edu/stat415/lesson/1/1.4>

[3] <https://en.wikipedia.org/wiki/Definite_symmetric_matrix>

[4] <https://www.youtube.com/watch?v=TePh29vzVEk>

[5] <https://www.youtube.com/watch?v=-elod4SsOts>

[6] <https://www.stat.berkeley.edu/~vigre/activities/bootstrap/2006/wickham_stati.pdf>

[7] <http://www2.econ.iastate.edu/classes/econ500/hallam/documents/Sample_Moments.pdf>

Code Appendix

library(latex2exp)  
### 1  
### a  
exponential\_distribution <- function(x, theta) {  
 ifelse(x >= 0,  
 theta \* exp(-theta \* x),  
 0)  
}  
  
xs <- seq(-1, 5, length.out = 1e4)  
ys <- exponential\_distribution(x = xs, theta = 1)  
  
thetas <- seq(0.5, 1.5, length.out = 3)  
y\_vec <- sapply(X = thetas, FUN = function(x)  
 exponential\_distribution(x = xs, theta = x))  
  
plot(xs, y\_vec[,3], type = 'l',  
 main = TeX('$p(x|\\theta)\\;vs.\\;x,\\;\\theta =0.5,1,1.5\\;x=\\[-1,\\;5\\]$'),  
 xlab = TeX('$x$'), ylab = TeX('$p(x|\\theta)$'))  
lines(xs, y\_vec[,2], col = 'red'); lines(xs, y\_vec[,1], col = 'blue')  
points(0, 0, pch=1)  
points(0, 0.5, pch=19)  
points(0, 1, pch=19)  
points(0, 1.5, pch=19)  
legend("topright",  
 legend = c(TeX('$\\theta =1.5$'), TeX('$\\theta =1$'), TeX('$\\theta =0.5$')),  
 col = c("black", "red", "blue"), lty = rep(1,3))  
### b  
thetas <- seq(0, 5, length.out = 1e4)  
exp\_dist\_vec <- Vectorize(exponential\_distribution, vectorize.args = "theta")  
  
y\_vec <- sapply(X = seq(1,3,length.out = 3), FUN = function(x)  
 exp\_dist\_vec(x = x, theta = thetas))  
  
plot(thetas, y\_vec[,1], type = 'l',  
 main = TeX('$p(x|\\theta)\\;vs.\\;\\theta,\\;\\theta = (0,\\;5\\],\\;x=1, 2, 3$'),  
 xlab = TeX('$\\theta$'), ylab = TeX('$p(x|\\theta)$'))  
lines(thetas, y\_vec[,2], col = 'red')  
lines(thetas, y\_vec[,3], col = 'blue')  
points(0, 0, pch=1)  
legend("topright", legend = c(TeX('$x=1$'), TeX('$x=2$'), TeX('$x=3$')),  
 col = c('black', 'red', 'blue'), lty = rep(1,3))  
  
### 2  
### a  
uniform\_distribution <- function(x, theta) {  
 ifelse((0<=x) & (x<=theta),  
 1 / theta,  
 0)  
}  
  
thetas <- seq(2, 10, length.out = 1e4)  
ys1 <- uniform\_distribution(x = 3, theta = thetas)  
ys2 <- uniform\_distribution(x = 5, theta = thetas)  
ys3 <- uniform\_distribution(x = 7, theta = thetas)  
  
y\_vec <- sapply(X = seq(1,3,length.out = 3), FUN = function(x)  
 uniform\_distribution(x = x, theta = thetas))  
  
plot(thetas, ys1, type = 'l',  
 main = TeX('$p(x|\\theta)\\;vs.\\;\\theta,\\;\\theta = \\[2,\\;10\\],\\;x=3,5,7$'),  
 xlab = TeX('$\\theta$'), ylab = TeX('$p(x|\\theta)$'))  
lines(thetas, ys2, col = 'red')  
lines(thetas, ys3, col = 'blue')  
points(3, 0, pch=1); points(3, max(ys1), pch=19)  
points(5, 0, pch=1); points(5, max(ys2), pch=19)  
points(7, 0, pch=1); points(7, max(ys3), pch=19)  
legend("topright", legend = c(TeX('$x=3$'), TeX('$x=5$'), TeX('$x=7$')),  
 col = c('black', 'red', 'blue'), lty = rep(1,3))