MGEC11 A3

Xinrui Wang 1004741078;Bisong Zhou 1004738741

2021/7/22

## Q1

1. Y Poisson distribution
2. P (Y = 0) = 0.5 -> e^(-lambda)=0.5 –> lambda=-log(0.5) in poisson distribution, mean = variance = lambda =

-log(0.5)

## [1] 0.6931472

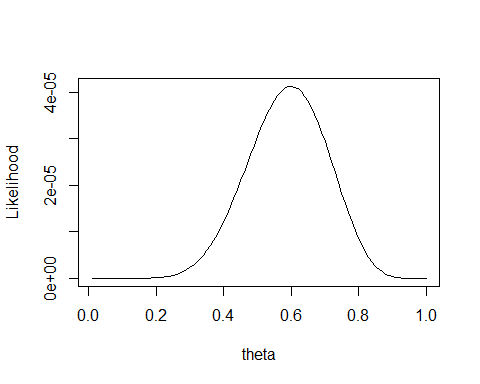
b)P(X=3)=

dpois(3,-log(0.5))

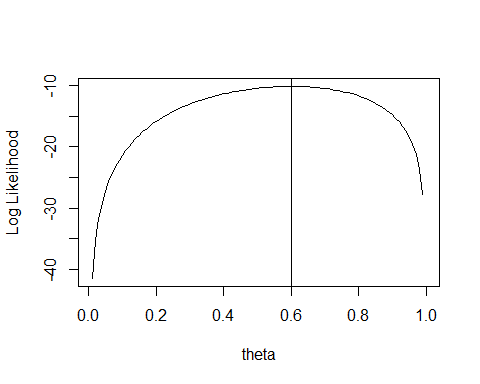
## [1] 0.02775205

1. est P(H) = pi, x=9,n=15,x/n=0.6
2. Bin(n=15,theta=0.5) pdf:f(x|theta)=thetax*(1-theta)^x L(theta|x=9)=theta^9*(1-theta)6

likelihood=function(n,y,theta){return(theta^y\*(1-theta)^(n-y))}  
theta=seq(0.01,1,0.01)  
plot(theta,likelihood(15,9,theta),type = "l",ylab="Likelihood")



loglike=function(n,y,theta){return(y\*log(theta)+(n-y)\*log(1-theta))}  
plot(loglike(15,9,theta)~theta,type="l", ylab="Log Likelihood")  
abline(v=0.6)



MLE

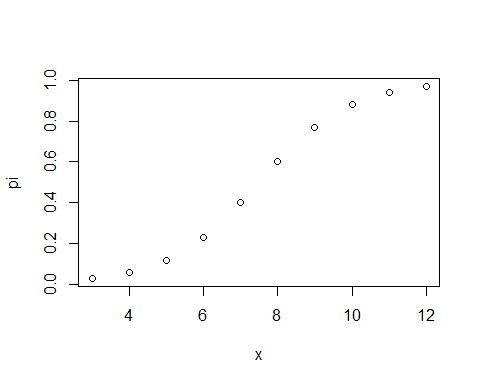
pi=

which.max(loglike(15,9,theta))\*0.01

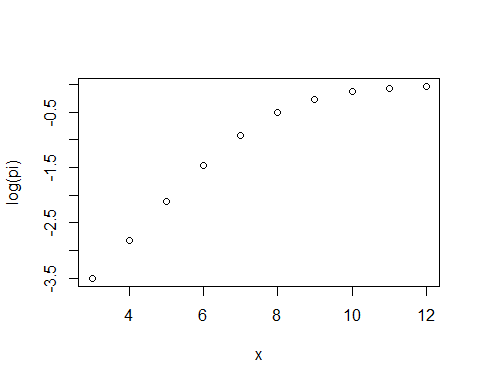
## [1] 0.6

## Q2

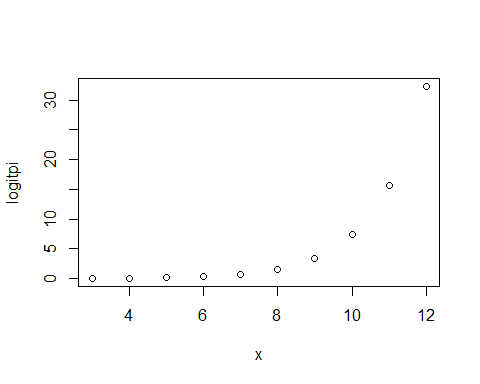
x<-seq(3,12,1)  
pi<-c(0.03,0.06,0.12,0.23,0.4,0.6,0.77,0.88,0.94,0.97)  
plot(pi~x)

 not linear

plot(log(pi)~x)

 not linear

logitpi<-pi/(1-pi)  
plot(logitpi~x)

 not linear

cor(x,logitpi)

## [1] 0.7833222

d) none of the models above is linear though, a) pi~x looks best fit the data

## Q3

d=c(45,15,40,83,90,25,35,65,95,35,75,45,50,75,30,25,20,60,70,30,60,61,65,15,20,45,15,25, 15,30,40,15,135,20,40)  
y=c(0,0,1,1,1,1,1,1,1,1,1,1,0,1,0,1,0,1,1,1,1,0,1,0,0,1,0,1,0,1,1,0,1,0,0)

model<-glm(y~d,family=binomial)  
summary(model)

##   
## Call:  
## glm(formula = y ~ d, family = binomial)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.0964 -0.7392 0.3020 0.8711 1.3753   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -2.21358 0.99874 -2.216 0.02667 \*   
## d 0.07038 0.02667 2.639 0.00831 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 46.180 on 34 degrees of freedom  
## Residual deviance: 33.651 on 33 degrees of freedom  
## AIC: 37.651  
##   
## Number of Fisher Scoring iterations: 5

logit=-2.21+0.07\*d

β=0.07038 means the sore threat on walking will increase by 0.07038 as the minute increase by 1.

b)

exp(-2.21358+0.07038\*1)/(1+exp(-2.21356+0.07038\*1))

## [1] 0.1049681

c)

library(lmtest)

## Loading required package: zoo

##   
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':  
##   
## as.Date, as.Date.numeric

lrtest(model)

## Likelihood ratio test  
##   
## Model 1: y ~ d  
## Model 2: y ~ 1  
## #Df LogLik Df Chisq Pr(>Chisq)   
## 1 2 -16.826   
## 2 1 -23.090 -1 12.528 0.0004008 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

the possibility (p-value) is 0.0004 smaller than 0.05, we will reject null hypothesis:restricted model is good. so unrestricted model is good.