STAC51 Assignment 1

Joseph Wang and Willy Chan

2023-02-03

Question 1

a)

The PMF of a Multinomial$(n, \pi_1, \dots, \pi_k)$ is

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{n!}{y_1! \dots y_k!} \pi_1^{y_1} \dots \pi_k^{y_k}$$

Thus is the multivariate moment generating function is

$$\begin{aligned}
M_{\mathbf{Y}}(\mathbf{t}) &= E(e^{\mathbf{t}'\mathbf{y}}) \\
&= \Sigma_{y_1}^{n} \Sigma_{y_2}^{n-y_1} \dots \Sigma_{y_k}^{n-y_1 \dots - y_{k-1}} e^{t_1 y_1 + \dots + t_k y_k} \frac{n!}{y_1! \dots y_k!} \pi_1^{y_1} \dots \pi_k^{y_k} \\
&= \Sigma_{y_1}^{n} \Sigma_{y_2}^{n-y_1} \dots \Sigma_{y_k}^{n-y_1 \dots - y_{k-1}} \frac{n!}{y_1! \dots y_k!} (e^{t_1}\pi_1)^{y_1} \dots (e^{t_k}\pi_k)^{y_k} \\
&= (\Sigma_i^k e^{t_i} \pi_i)^n
\end{aligned}$$

By the Multinomial Theorem.

b)

The moment generating function for the ith var is

$$\begin{aligned}
M_{Y_i}(t) &= M_{\bar{Y}}(0, \dots, t, \dots, 0) \\
&= (\pi_1 + \dots + e^t \pi_i + \dots + \pi_k)^n
\end{aligned}$$

Where t is in the ith position.

The mean is

$$\begin{aligned}
E(Y_i) &= \frac{\partial}{\partial t} M_{Y_i}(0) \\
&= n(\pi_1 + \dots + e^t \pi_i + \dots + \pi_k)^{n-1} e^t \pi_i |_{t=0} \\
&= n\pi_i
\end{aligned}$$

Since $(\pi_1 + \dots + \pi_k) = 1$.

c)

The second moment is

$$\begin{aligned} E(Y_i^2) &= \frac{\partial^2}{\partial t^2} M_{Y_i}(0) \\ &= \frac{\partial}{\partial t} n(\pi_1 + \cdots + e^t\pi_i + \cdots + \pi_k)^{n-1} e^t \pi_i|_{t=0} \\ &= n(\pi_1 + \cdots + e^t\pi_i + \cdots + \pi_k)^{n-1} e^t\pi_i + e^t\pi_i * n(n-1)(\pi_1 + \cdots + e^t\pi_i + \cdots + \pi_k)^{n-2} e^t\pi_i|_{t=0} \\ &= n\pi_i + n^2\pi_i^2 - n\pi_i^2 \end{aligned}$$

Since $(\pi_1 + \cdots + \pi_k) = 1$.

The variance is

$$\begin{aligned} Var(Y_i) &= E(Y_i^2) - E(Y_i)^2 \\ &= n\pi_i + n^2\pi_i^2 - n\pi_i^2 - n^2\pi_i^2 \\ &= n\pi_i(1 - \pi_i) \end{aligned}$$

Since $E(Y_i)^2 = (n\pi_i)^2$

d)

The joint moment generating function for the ith and jth var is

$$\begin{aligned} M_{Y_iY_j}(t_i, t_j) &= M_{\bar{Y}}(0, \ldots, t_i, \ldots, t_j, \ldots, 0) \\ &= (\pi_1 + \cdots + e^{t_i}\pi_i + \cdots + e^{t_j}\pi_j + \cdots + \pi_k)^n \end{aligned}$$

The expected value of Y_iY_j is

$$\begin{aligned} E(Y_iY_j) &= \frac{\partial^2}{\partial t_i t_j} M_{Y_iY_j}(0,0) \\ &= \frac{\partial}{\partial t_j} n(\pi_1 + \cdots + e^{t_i}\pi_i + \cdots + e^{t_j}\pi_j + \cdots + \pi_k)^{n-1} e^{t_i}\pi_i|_{t_i,t_j=0} \\ &= n(n-1)(\pi_1 + \cdots + e^{t_i}\pi_i + \cdots + e^{t_j}\pi_j + \cdots + \pi_k)^{n-2} e^{t_i}\pi_i e^{t_j}\pi_j|_{t_i,t_j=0} \\ &= n(n-1)\pi_i\pi_j \end{aligned}$$

Since $(\pi_1 + \cdots + \pi_k) = 1$.

The covariance of Y_iY_j is

$$\begin{aligned} Cov(Y_i, Y_j) &= E(Y_iY_j) - E(Y_i)E(Y_j) \\ &= n(n-1)\pi_i\pi_j - n^2\pi_i\pi_j \\ &= -n\pi_i\pi_j \end{aligned}$$

Since $E(Y_i)E(Y_j) = (n\pi_i)(n\pi_j)$

e)

Given $c = 2$, $1 - \pi_i - \pi_j = 0$

$$\begin{aligned}
Cor(Y_i, Y_j) &= \frac{Cov(Y_i, Y_j)}{\sqrt{Var(Y_i)Var(Y_j)}} \\
&= \frac{-n\pi_i\pi_j}{\sqrt{n\pi_i(1-\pi_i)n\pi_j(1-\pi_j)}} \\
&= \frac{-\pi_i\pi_j}{\sqrt{\pi_i\pi_j(1-\pi_i-\pi_j+\pi_i\pi_j)}} \\
&= -1
\end{aligned}$$

This makes sense given n and $c = 2$ as if the number of successes one category increases, the number of successes in the other category has to decrease. The change is 1:1 and thus has a linear relationship of factor 1.

Question 2

```
y = 5
n = 30
pihat = y/n
alpha = 0.1
```

a)

Test P-values

```
pi0 = 0.1
yfit = pi0*n

sewald= sqrt(pihat*(1-pihat)/n)
sescore = sqrt(pi0*(1-pi0)/n)

wald = (pihat - pi0)/sewald
score = (pihat - pi0)/sescore
lrt = 2*(y*log(y/yfit) + (n-y)*log((n-y)/(n-yfit)))

wald
```

```
## [1] 0.9797959
```

```
score
```

```
## [1] 1.217161
```

```r
lrt
```

```
## [1] 1.260204
```

```r
2*pnorm(wald, lower.tail = FALSE)
```

```
## [1] 0.3271869
```

```r
2*pnorm(score, lower.tail = FALSE)
```

```
## [1] 0.2235429
```

```r
pchisq(lrt,1,lower.tail = FALSE)
```

```
## [1] 0.2616124
```

The test statistics for Wald, Score and LRT are 0.9797959, 1.217161 and 1.260204. The P-Values for Wald, Score and LRT are 0.3271869, 0.2235429 and 0.2616124.

b)

Wald CI

```r
z = qnorm(1-alpha/2)

c(pihat - z*sewald, pihat + z*sewald)
```

```
## [1] 0.05474855 0.27858478
```

The Wald CI is (0.05474855, 0.27858478).

c)

Score CI

```r
c(((n*pihat+z^2/2)-z*sqrt(n*pihat*(1-pihat)+z^2/4))/(n+z^2),((n*pihat+z^2/2)+z*sqrt(n*pihat*(1-pihat)+z
```

```
## [1] 0.08356237 0.30492050
```

The Score CI is (0.08356237, 0.30492050).

d)

Agresti-Coull

```r
pis = (y+z^2/2)/(n+z^2)
ns = n+z^2

c(pis-z*sqrt(pis*(1-pis)/ns), pis+z*sqrt(pis*(1-pis)/ns))
```

```
## [1] 0.08045514 0.30802774
```

The Agresti-Coull CI is (0.08045514 0.30802774).

e)

```r
prop.test(y,n,p=0.1, correct=F,conf.level=1-alpha)
```

```
## Warning in prop.test(y, n, p = 0.1, correct = F, conf.level = 1 - alpha):
## Chi-squared approximation may be incorrect

##
##  1-sample proportions test without continuity correction
##
## data:  y out of n, null probability 0.1
## X-squared = 1.4815, df = 1, p-value = 0.2235
## alternative hypothesis: true p is not equal to 0.1
## 90 percent confidence interval:
##  0.08356237 0.30492050
## sample estimates:
##         p
## 0.1666667
```

Same as the calculated CI.

f)

```r
library(binom)

binom.confint(x=y,n=n,conf.level=1-alpha, methods="all")
```

```
##           method x  n      mean      lower     upper
## 1  agresti-coull 5 30 0.1666667 0.08045514 0.3080277
## 2     asymptotic 5 30 0.1666667 0.05474855 0.2785848
## 3          bayes 5 30 0.1774194 0.06782250 0.2829193
## 4        cloglog 5 30 0.1666667 0.07379893 0.2917865
## 5          exact 5 30 0.1666667 0.06805557 0.3189712
## 6          logit 5 30 0.1666667 0.08201730 0.3092501
## 7         probit 5 30 0.1666667 0.07848132 0.3017123
## 8        profile 5 30 0.1666667 0.07582948 0.2964459
## 9            lrt 5 30 0.1666667 0.07579682 0.2964405
## 10     prop.test 5 30 0.1666667 0.06303555 0.3545108
## 11        wilson 5 30 0.1666667 0.08356237 0.3049205
```

Same as the calculated CIs.

Question 3

The likelihood function is $l(\theta|y) = \binom{n}{y}(\theta)^y(1-\theta)^{n-y}$. With n = 30 and y = 5.

a)

Thus the likelihood under $\pi_0 = 0.1$ is

$$l(0.1|5) = \binom{30}{5}(0.1)^5(0.9)^{25}$$

```
l0 = choose(n,y)*(pi0)^5*(1-pi0)^(n-y)
l0
```

```
## [1] 0.1023048
```

The maximum likelihood under H_0 is 0.1023048.

b)

The MLE is the maximized likelihood over all possible values. $\hat{\pi} = \frac{y}{n} = \frac{5}{30} = \frac{1}{6}$.

$$l(\frac{1}{6}|5) = \binom{30}{5}(\frac{1}{6})^5(\frac{5}{6})^{25}$$

```
l1 = choose(n,y)*(pihat)^5*(1-pihat)^(n-y)
l1
```

```
## [1] 0.1921081
```

The maximum likelihood under the MLE is 0.1921081.

c)

```
lrt2 = -2*log(l0/l1)
lrt2
```

```
## [1] 1.260204
```

The LRT statistic is 1.260204.

d)

```
qchisq(0.9, df=1)
```

```
## [1] 2.705543
```

The test statistic must be at least greater than the critical value at 2.705543.

e)

```
library(rootSolve)

f1 = function(pif){
  -2*(y*log(pif) + (n-y)*log(1-pif)-y*log(pihat) - (n-y)*log(1-pihat)) - qchisq(0.9, df=1)
}

uniroot.all(f=f1,interval=c(0,1))
```

```
## [1] 0.0756936 0.2963735
```

The LRT CI is (0.0756936,0.2963735).

Question 4

a)

```
N = 100000
n = 25
pi = 0.06
alpha = 0.05
za = qnorm(1-alpha/2)

count = 0
for(i in 1:N){
  z = rbinom(n,1,pi)
  y = sum(z)
  pi_hat = y/n

  l = pi_hat - za*sqrt(pi_hat*(1-pi_hat)/n)
  t = pi_hat + za*sqrt(pi_hat*(1-pi_hat)/n)

  if(l<= pi & pi <= t){
    count = count + 1
  }
}

count/N
```

```
## [1] 0.78414
```

As explained in lecture, unless n is very large, the Wald CI's true coverage is lower than $(1-\alpha)\%$ when π approaches 0 or 1.

b)

```r
p = 0
n = 25
for(i in 0:n){
  pi_hat = i/n

  l = pi_hat - za*sqrt(pi_hat*(1-pi_hat)/n)
  t = pi_hat + za*sqrt(pi_hat*(1-pi_hat)/n)
  if(l<= pi & pi <= t){
    p = p + choose(n,i)*(pi)^i*(1-pi)^(n-i)
  }
}

p
```

```
## [1] 0.784026
```

The true confidence is 0.784026 and is close to the simulated value

Question 5

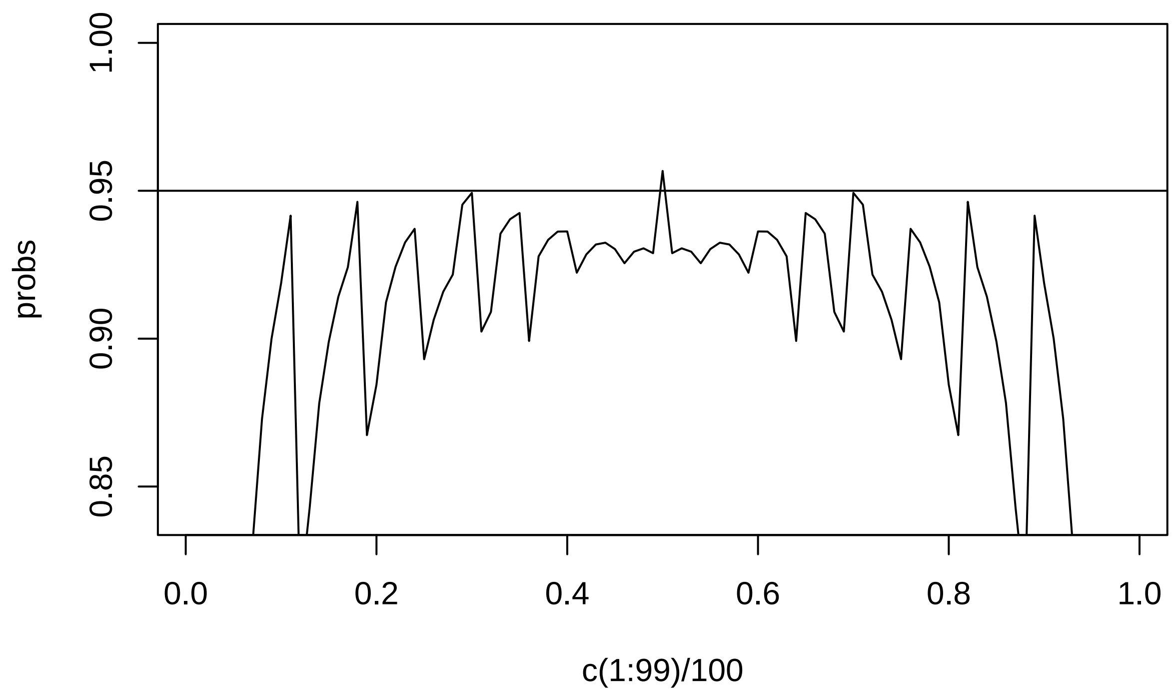
a)

```r
probs = c()
n=25
alpha = 0.05
z = qnorm(1-alpha/2)


for(i in 1:99){
  pi = i/100
  p = 0
  for(j in 0:n){
    pi_hat = j/n

    l = pi_hat - z*sqrt(pi_hat*(1-pi_hat)/n)
    t = pi_hat + z*sqrt(pi_hat*(1-pi_hat)/n)

    if(l<= pi & pi <= t){
      p = p + choose(n,j)*(pi)^j*(1-pi)^(n-j)
    }
  }
  probs = append(probs, p)
}
plot(x=c(1:99)/100, y=probs, type="l", ylim=c(0.84,1))
abline(h=0.95)
```

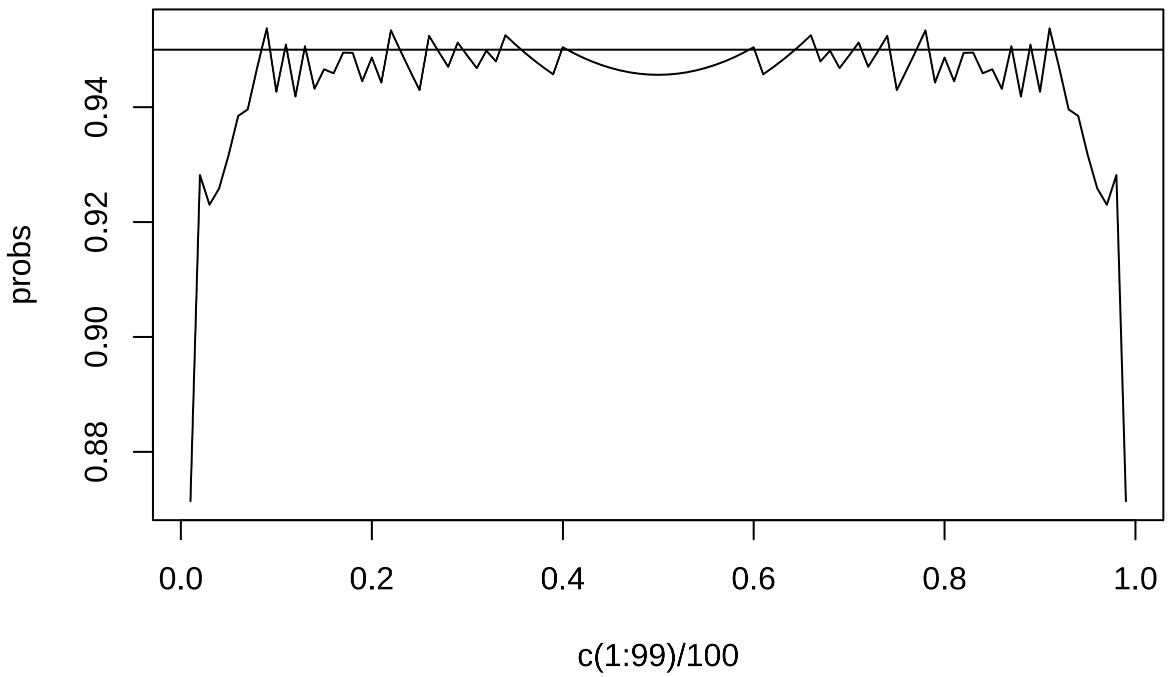


Given a small n, the function does not consistently hit the 95% inclusion mark and that probability decreases as π approaches 0 or 1. This is consistent with what we found in question 4a).

b)

```
probs = c()
n=500
for(i in 1:99){
  pi = i/100
  p = 0
  for(j in 0:n){
    pi_hat = j/n

    l = pi_hat - z*sqrt(pi_hat*(1-pi_hat)/n)
    t = pi_hat + z*sqrt(pi_hat*(1-pi_hat)/n)
    if(l<= pi & pi <= t){
      p = p + choose(n,j)*(pi)^j*(1-pi)^(n-j)
    }
  }
  probs = append(probs, p)
}
plot(x=c(1:99)/100, y=probs, type="l")
abline(h=0.95)
```



Just like a) the probability decreases as π approaches 0 or 1 but since we have a relatively larger n = 500, the CI hits the true coverage probability more often.

c)

```r
probs = c()
n=25
for(i in 1:99){
  pi = i/100
  p = 0
  for(j in 0:n){
    pi_hat = j/n

    l = pi_hat - z*sqrt(pi_hat*(1-pi_hat)/n)
    t = pi_hat + z*sqrt(pi_hat*(1-pi_hat)/n)
    if(l<= pi & pi <= t){
      p = p + choose(n,j)*(pi)^j*(1-pi)^(n-j)
    }
  }
  probs = append(probs, p)
}
plot(x=c(1:99)/100, y=probs, type="l", ylim=c(0.85,1), col="red")
abline(h=0.95)
```

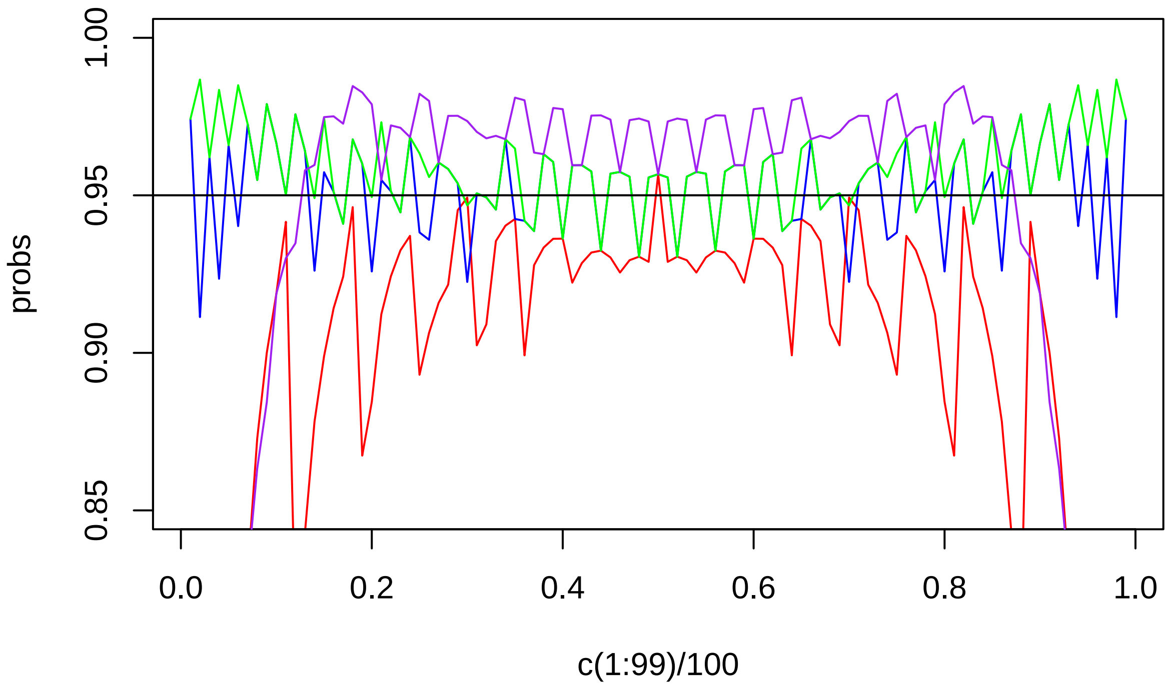
```
probs = c()
for(i in 1:99){
  pi = i/100
  p = 0
  for(j in 0:n){
    pi_hat = j/n

    l = ((n*pi_hat+z^2/2)-z*sqrt(n*pi_hat*(1-pi_hat)+z^2/4))/(n+z^2)
    t = ((n*pi_hat+z^2/2)+z*sqrt(n*pi_hat*(1-pi_hat)+z^2/4))/(n+z^2)
    if(l<= pi & pi <= t){
      p = p + choose(n,j)*(pi)^j*(1-pi)^(n-j)
    }
  }
  probs = append(probs, p)
}
lines(x=c(1:99)/100, y=probs, type="l",col="blue")
abline(h=0.95)


probs = c()
ns = n+z^2
for(i in 1:99){
  pi = i/100
  p = 0
  for(j in 0:n){
    pis = (j+z^2/2)/(n+z^2)

    l = pis-z*sqrt(pis*(1-pis)/ns)
    t = pis+z*sqrt(pis*(1-pis)/ns)
    if(l<= pi & pi <= t){
      p = p + choose(n,j)*(pi)^j*(1-pi)^(n-j)
    }
  }
  probs = append(probs, p)
}
lines(x=c(1:99)/100, y=probs, type="l",col="green")
abline(h=0.95)

probs = c()
for(i in 1:99){
  pi = i/100
  p = 0
  for(j in 1:(n-1)){
    l = 1/(1 + (n-j+1)/(j*qf(alpha/2,2*j,2*(n-j+1))))
    t = 1/(1 + (n-j)/((j+1)*qf(1-alpha/2,2*(j+1),2*(n-j))))
    if(l<= pi & pi <= t){
      p = p + choose(n,j)*(pi)^j*(1-pi)^(n-j)
    }
  }
  probs = append(probs, p)
}
lines(x=c(1:99)/100, y=probs, type="l",col="purple")
abline(h=0.95)
```



The order of inclusion rates seems to be Clopper-Preason, Agresti-Coull, Score and Wald. The higher inclusion rates could be because of wider CIs.

Question 6

a)

View each square as an independent Poisson distribution. Thus their joint distribution is,

$$\begin{aligned} P(X_1 = x_1, \ldots, X_{576} = x_{576}|\lambda) &= P(X_1 = x_1|\lambda) \ldots P(X_{576} = x_{576}|\lambda) \\ &= (\frac{e^{-\lambda}\lambda^0}{0!})^{229} * (\frac{e^{-\lambda}\lambda^1}{1!})^{221} \\ &* (\frac{e^{-\lambda}\lambda^2}{2!})^{93} * (\frac{e^{-\lambda}\lambda^3}{3!})^{35} \\ &* (\frac{e^{-\lambda}\lambda^4}{4!})^{7} * (\frac{e^{-\lambda}\lambda^5}{5!})^{1} \\ &= e^{-576\lambda}\lambda^{535}/C \end{aligned}$$

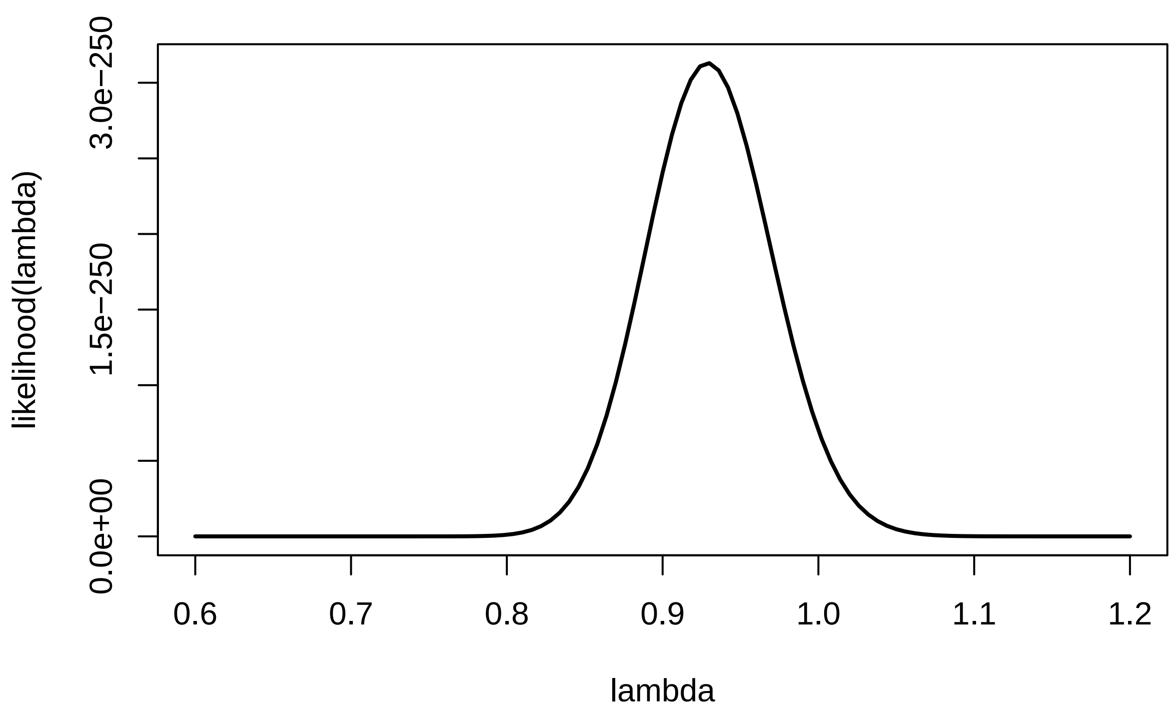
The likelihood function is

$$l(\theta|x) = e^{-576\lambda}\lambda^{535}$$

b)

```
n=576
total = sum(c(1,2,3,4,5)*c(211,93,35,7,1))
likelihood = function(lambda){
  exp(-n*lambda)*lambda^total
}

curve(likelihood, from=0.6, to=1.2, xlab="lambda", ylab="likelihood(lambda)", lwd=2)
```



c)

```
max = optimize(likelihood, interval=c(0.6,1.2), maximum=TRUE)$maximum
max
```

```
## [1] 0.9288118
```

The MLE for lambda is 0.9288118. ## d)

```
test = -2*log(likelihood(1)/likelihood(max))
test
```

```
## [1] 2.990222
```

```r
qchisq(0.95, df=1)
```

```
## [1] 3.841459
```

Cannot reject $H_0 : \lambda = 1$ since test statistic $<$ critical value.