Assignment 8

Ryan Klughart

06/04/2021

Question 1

$H_0 : p_1 = p_2 = ...p_8 = 1/8$
H_A : at least one probabilities is not 1/8

```
N=120
observed = c(12,16,17,15,13,20,17,10)
expected = N*1/8
expected # for all directions
```

```
## [1] 15
```

```
TestStat = sum((observed-expected)^2/(expected))
TestStat
```

```
## [1] 4.8
```

```
chisq = qchisq(.9, df=7)
chisq
```

```
## [1] 12
```

Since the test statistic 4.8 is not in the rejection zone >12.017 we fail to reject the null hypothesis. Therefore, when homing pigeons are disoriented they will exhibit no preference in direction of flight after takeoff.

Question 2

a)

The 5 intervals are solved at bottom of pdf and are: (0,.223),(.223,.511),(.511,.916),(.916,1.61),(1.61,inf)

b)

$H_0 : p_1 = p_2 = ...p_5 = 1/5$
H_A : at least one probability does not equal 1/5

```
xi = c(.1,.99,1.14,1.26,3.24,.12,.26,.8,.79,1.16,1.76,.41,.59,.27,2.22,.66,.71,2.21,.68,.43,.11,.46,.69

a1 = .233
a2 = .511
a3 = .916
a4 = 1.61
a5 = Inf

x1 = xi[xi<a1]
x2 = xi[xi<a2 & xi>a1]
x3 = xi[xi<a3 & xi>a2]
x4 = xi[xi<a4 & xi>a3]
```

```r
x5 = xi[xi<a5 & xi>a4]

Observed = c(length(x1),length(x2),length(x3),length(x4),length(x5))


N =40
ps = 1/5
Expected = N*ps
X2 = sum((Observed-Expected)^2/(Expected))
X2
```

```
## [1] 1.25
```

```r
chi = qchisq(.95, df=4)
chi
```

```
## [1] 9.49
```

Since the test statistic 1.25 is not in the rejection region >9.488 we fail to reject the null hypothesis. Therefore, an exponential distribution with λ =1 does fit the data well.

Question 3

a)

$H_0 : p_0 = .125, p_1 = .375, p_2 = .375, p_3 = .125$ where p_i is the probability of i children
H_A : at least one probability does not equal the value suggested in the null hypothesis

```r
px0 = .125 # .5^4
px1 = .375 # choose(4,2)*.5^4
px2 = .375 # choose(4,3)*.5^4
px3 = .125 # .5^4
ps = c(px0,px1,px2,px3) #probabilities
N=160
Expected = N*ps
Expected
```

```
## [1] 20 60 60 20
```

```r
Observed = c(14,66,64,16)

X2 = sum((Observed-Expected)^2/(Expected))
X2
```

```
## [1] 3.47
```

```r
chi = qchisq(.95, df=3)
chi
```

```
## [1] 7.81
```

Since the test statistic 3.467 is not in the rejection region >7.815 we fail to reject the null hypothesis. Therefore the binomial distribution fits the data of number of male children in families of 3 children well.

Question 4

mle estimate of θ is calculated at the end of the pdf.

```r
Observed = c(26,51,47,16,10)
N = 150
```

```r
x = 0:4


u = sum(Observed*x)
v = sum(rev(x)*Observed)
theta=u/(v+u)
theta
```

```
## [1] 0.388
```

```r
ps =c()
for( i in x){
  ps = c(ps,choose(4,i)*theta^i*(1-theta)^(4-i))
}
ps# probabilities
```

```
## [1] 0.1400 0.3555 0.3385 0.1433 0.0227
```

```r
Expected = N*ps
Expected # expected values
```

```
## [1] 21.00 53.32 50.78 21.49  3.41
```

```r
X2 = sum((Observed-Expected)^2/(Expected))
X2
```

```
## [1] 15.7
```

```r
chi = qchisq(.95, df=4)
chi
```

```
## [1] 9.49
```

Since the test statistic 15.704 is in the rejection region >9.488 we reject the null hypothesis. Therefore the binomial distribution does not fit the data of number of defective batteries

Question 5

a)

Shown at end of pdf

b)

```r
N=130
Observed = c(48,31,20,9,6,5,4,2,1,1,2,1)
x = 1:12
n=12
phat = (sum(Observed)-n)/sum(Observed)
phat
```

```
## [1] 0.908
```

```r
qhat = 1-phat
Expected =c()
for( i in x){
  Expected = c(Expected,N*phat^(i-1)*qhat)
}
Expected
```

```
##  [1] 12.00 10.89  9.89  8.97  8.15  7.39  6.71  6.09  5.53  5.02  4.56  4.14
```

```r
Expected7 = c(Expected[c(1:6)],sum(Expected[c(7:12)]))
Expected7
```

```
## [1] 12.00 10.89  9.89  8.97  8.15  7.39 32.04
```

```r
Observed7 = c(Observed[c(1:6)],sum(Observed[c(7:12)]))
Observed7
```

```
## [1] 48 31 20  9  6  5 11
```

```r
X2 = sum((Observed7-Expected7)^2/(Expected7))
X2
```

```
## [1] 171
```

```r
chi = qchisq(.95, df=11)
chi
```

```
## [1] 19.7
```

Since the test statistic 170.624 is in the rejection region >19.675 we reject the null hypothesis. Therefore the given distribution does not fit the data of number of hops before flight of birds

Question 6

```r
N = 300
n = 10
x = 0:(n-1)
Observed = c(6,24,42,59,62,44,41,14,6,2)
lambdahat = (1/N)*sum(x*Observed)
lambdahat
```

```
## [1] 3.88
```

```r
ps = c()
for(i in x){
  ps = c(ps,(exp(-lambdahat)*lambdahat^i)/factorial(i))
}
Expected = N*ps
Expected
```

```
##  [1]  6.22 24.10 46.71 60.36 58.50 45.35 29.30 16.23  7.86  3.39
```

```r
Expected89 = c(Expected[c(1:8)],sum(Expected[c(9,10)]))
Expected89
```

```
## [1]  6.22 24.10 46.71 60.36 58.50 45.35 29.30 16.23 11.25
```

```r
Observed89 = c(Observed[c(1:8)],sum(Observed[c(9,10)]))
Observed89
```

```
## [1]  6 24 42 59 62 44 41 14  8
```

```r
X2 = sum((Observed89-Expected89)^2/(Expected89))
X2
```

```
## [1] 6.68
```

```r
chi = qchisq(.95, df=9)
chi
```

```
## [1] 16.9
```

Since the test statistic 6.677 is not in the rejection region >16.919 we fail to reject the null hypothesis. Therefore the Poisson distribution with λ=3.88 does fit the data of sister-chromatid exchanges

$$P_1 \quad P_2 \quad P_3 \quad P_4 \quad P_5$$

2. $(0, a_1)(a_1, a_2)(a_2, a_3)(a_3, a_4)(a_4, \infty)$

want $P_i = 1/5$

$$\frac{1}{5} = \int_0^{a_1} e^{-x}\,dx = -e^{-x}\Big|_0^{a_1}$$

$$= -e^{-a_1} - (-1)$$

$$= 1 - e^{-a_1}$$

$$e^{-a_1} = 4/5$$

$$-a_1 = \ln(4) - \ln(5)$$

$$a_1 = \ln(5) - \ln(4) = 0.223$$

$$\int_{a_1}^{a_2} e^{-x}\,dx = -e^{-a_2} - (-e^{-a_1})$$

$$= -e^{-a_2} + \frac{4}{5} = \frac{1}{5}$$

$$= e^{-a_2} = \frac{3}{5}$$

$$-a_2 = \ln(3) - \ln(5)$$

$$a_2 = 0.511$$

Scanned with CamScanner

$$\int_{a_2}^{a_3} e^{-x}\,dx = -e^{-a_3} - (-e^{-a_2}) = 1/5$$

$$= -e^{-a_3} + \frac{3}{5} = \frac{1}{5}$$

$$= e^{-a_3} = \frac{2}{5}$$

$$a_3 = \ln(5) - \ln(2) = 0.916$$

$$\int_{a_3}^{a_4} f(x)\,dx = -e^{-a_4} - (-e^{-a_3})$$

$$= -e^{-a_4} + \frac{2}{5} = \frac{1}{5}$$

$$a_4 = \ln(5) - \ln(1) = 1.61$$

4. $f(\theta) = \theta^{u}(1-\theta)^{v}$

$\ln f(\theta) = u \ln(\theta) + v \ln(1-\theta)$

$$\frac{d \ln f(\theta)}{d\theta} = \frac{u}{\theta} - \frac{v}{1-\theta} = 0$$

$$(1-\theta)u - \theta v = 0$$

$$u - \theta u - \theta v = 0$$

$$\theta u + \theta v = u$$

$$\hat{\theta} = \frac{u}{u+v}$$

Solving for u & v

$$f(n_0, n_1, n_2, n_3, n_4, \theta) = \frac{n!}{n_0! n_1! \cdots n_4!} P_0^{n_0} \cdots P_4^{n_4}$$

$$= L(\theta)$$

$$= \text{Constant} \times \left(\binom{4}{0}\theta^0 (1-\theta)^{4-0}\right)^{n_0} \left(\binom{4}{1}\theta^1 (1-\theta)^{4-1}\right)^{n_1} \cdots \left(\quad\right)^{n_4}$$

$$= \underbrace{\text{Constant} \times \binom{4}{0}\binom{4}{1} \cdots \binom{4}{4}}_{\text{Constant}} \left(\theta^{0 \cdot n_0 + \cdots 4 \cdot n_4} (1-\theta)^{4 \cdot n_0 + \cdots 0 \cdot n_4}\right)$$

$$L(\theta) = \theta^{u} (1-\theta)^{v}$$

$$u = 0 \cdot n_0 + 1 \cdot n_1 + \cdots 4 \cdot n_4 = 233$$
$$v = 4 \cdot n_0 + 3 \cdot n_1 + \cdots 0 \cdot n_4 = 367$$

$$\hat{\theta} = \frac{233}{233 + 367} = 0.388$$

Scanned with CamScanner

5a) $p^{\sum x_i - n} \cdot (1-p)^n$

$\ln(p^{\sum x_i - n} \cdot (1-p)^n)$

$= (\sum x_i - n)\ln(p) + n\ln(1-p)$

$\frac{d}{dp} = \frac{\sum x_i - n}{p} - \frac{n}{1-p} = 0$

$= (1-p)(\sum x_i - n) - pn = 0$

$= \sum x_i - n - p\sum x_i + \cancel{pn} - \cancel{pn} = 0$

$-p\sum x_i = -\sum x_i + n$

$\hat{p} = \frac{\sum x_i - n}{\sum x_i}$