Problems Sheet 8

Statistics 2

Dom Hutchinson

```
xtab<-c(109,65,22,3,1,0)
xobs<-rep(0:5,xtab)
n<-length(xobs)
xobs.bar<-mean(xobs)
```

Question 1

Let $\mathbf{X} := (X_1, \dots, X_{200}) \overset{\text{iid}}{\sim} \text{Poisson}(\lambda)$ be a random vector representing the number of deaths each year & $\mathbf{x} := (x_1, \dots, x_{200})$ be a realisation of $\mathbf{X}$.
Here we shall test $H_0 : \lambda \leq 0.5$ against $H_1 : \lambda > 0.5$.
Define test statistic $T(\mathbf{X}) = \frac{1}{n}\sum_{i=1}^n X_i$, from the data we have

$$T(\mathbf{x}) = -\frac{1}{200}\sum_{i=1}^{200} x_i = \frac{122}{200} = -0.61$$

Since $\sum_{i=1}^n X_i$ is an equivalent test statistic to the Neyamn-Pearson Test statistic then $T(\mathbf{X})$ is an equivalent test statistic too.
Thus $T(\mathbf{X})$ is the uniformly most powerful test statistic for H_0 & H_1 and has associated p-value

$$\begin{array}{rcl}
p(\mathbf{X}) & = & \mathbb{P}(T(\mathbf{X}) \geq T(\mathbf{x}); 0.5) \\
& = & \mathbb{P}\left(\frac{1}{200}\sum_{i=1}^{200} X_i \geq 0.61; 0.5\right) \\
& = & \mathbb{P}\left(\sum_{i=1}^{200} X_i \geq 122; 0.5\right) \\
& = & \mathbb{P}\left(\sum_{i=1}^{200} X_i \geq 122; 0.5\right) \\
& = & \mathbb{P}(Y \geq 122) \text{ by independence where } Y \sim \text{Poisson}(200 \times 0.5) = \text{Poisson}(100) \\
& = & 1 - \mathbb{P}(Y < 122) \\
& = & 0.0180734
\end{array}$$

Question 2

a)

We can approximate $Y \sim \text{Poisson}(100)$ as $\tilde{Y} \approx \text{Normal}(100, 100)$. Then

$$p(\mathbf{X}) = 1 - \mathbb{P}(Y < 122) \approx 1 - \mathbb{P}(\tilde{Y} < 122) = 0.4129356$$

b)

```
lambda<-.5
trials<-1000

samples.raw<-sapply(1:trials, function(i) sum(rpois(200,lambda)))
count<-sum(samples.raw<=122)
p<-1-count/trials
p
```

```
## [1] 0.017
```

Question 3

Let $\mathbf{X} := (X_1, \ldots, X_{200}) \overset{\text{iid}}{\sim} \text{Poisson}(\lambda)$ be a random vector representing the number of deaths each year & $\mathbf{x} := (x_1, \ldots, x_{200})$ be a realisation of $\mathbf{X}$.
I shall test $H_0 : \lambda = 0.55$ against $H_1 : \lambda \neq 0.55$ at a significance level of 5% where $\Theta := \mathbb{R}^{>0}$.
Thus $\Theta_0 := \{0.55\}$ and is nested in Θ. Note that $\dim(\Theta) = 1$ and $\dim(\Theta_0) = 0$.
Since Θ_0 is nested within Θ we have

$$T_n(\mathbf{X}) := -2\ln\Lambda_n(\mathbf{X}) \to_{\mathcal{D}(\cdot;\theta)} \chi^2_{\dim(\Theta)-\dim(\Theta_0)} = \chi^2_{1-0} = \chi^2_1 \text{ where } \Lambda_n(\mathbf{x}) := \frac{\sup_{\lambda\in\Theta_0} f_n(\mathbf{x};\lambda)}{\sup_{\lambda\in\Theta} f_n(\mathbf{x};\lambda)}$$

We shall us $T_n(\mathbf{X})$ as our test statistic.
We have

$$\begin{aligned} T(\mathbf{X}) &:= -2\ln\Lambda_n(\mathbf{X}) \\ &= -2\ln\tfrac{\sup_{\lambda=0.55} f_n(\mathbf{x};\lambda)}{\sup_{\lambda\in\mathbb{R}^{>0}} f_n(\mathbf{x};\lambda)} \\ &= -2\left[\ell_n(\mathbf{x};\lambda) - \ell_n(\mathbf{x};\hat\lambda_{MLE})\right] \end{aligned}$$

Since I am constructing $\alpha = 0.95$ confidence interval we wish to retain λ if

$$T_n(\mathbf{x}) := -2\left[\ell_n(\lambda;\mathbf{x}) - \ell_n(\hat\lambda_{MLE};\mathbf{x})\right] < \chi^2_{1,\alpha}$$

This is the confidence set

$$\begin{aligned} C(\mathbf{x}) &= \left\{\theta : T_n(\mathbf{x}) < \chi^2_{1,\alpha}\right\} \\ &= \left\{-2\left[\ell_n(\lambda;\mathbf{x}) - \ell_n(\hat\lambda_{MLE};\mathbf{x})\right] < \chi^2_{1,\alpha}\right\} \\ &= \left\{\ell_n(\lambda;\mathbf{x}) > \ell_n(\hat\lambda_{MLE};\mathbf{x}) - \tfrac{1}{2}\chi^2_{1,\alpha}\right\} \end{aligned}$$

We can calculate the $\chi^2_{1,\alpha}$ and $\ell_n(\hat\lambda_n;\mathbf{x})$ using r

```r
ell <- function(lambda) {
  stopifnot(all(lambda > 0))
  n <- length(xobs)
  n * (-lambda + mean(xobs) * log(lambda))
}

alpha=0.95

ell.mle<-optimise(ell,interval=c(0,1),maximum=TRUE)$objective
quantile<-qchisq(alpha,1)
cat("ell.mle=",ell.mle,"\nchisq_alpha=",quantile,sep="")
```

```
## ell.mle=-182.3042
## chisq_alpha=3.841459
```

Giving

$$\begin{aligned} C(\mathbf{x}) &= \left\{\ell_n(\lambda;\mathbf{x}) > \ell_n(\hat\lambda_{MLE};\mathbf{x}) - \tfrac{1}{2}\chi^2_{1,\alpha}\right\} \\ &= \left\{\ell_n(\lambda;\mathbf{x}) > -182.3041513 - \tfrac{1}{2}3.8414588\right\} \\ &= \left\{\ell_n(\lambda;\mathbf{x}) > -184.2248807\right\} \end{aligned}$$

```r
x<-seq(0.4,.8,length.out=100)
y<-sapply(x,function(p) ell(p))
z<-x[which(y>ell.mle-.5*quantile)] # The values where the inequality holds
cat("lower:",min(z),"\nupper:",max(z))
```

```
## lower: 0.5090909
## upper: 0.7232323
```

Thus

$$C(\mathbf{x}) = \{\theta : \theta \in [0.5090909, 0.7232323]\}$$

We accept H_0 in this case.

b)

```
alpha=0.9

quantile<-qchisq(alpha,1)

z<-x[which(y>ell.mle-.5*quantile)] # The values where the inequality holds
cat("chisq_alpha=",quantile,"\nlower:",min(z),"\nupper:",max(z))
```

```
## chisq_alpha= 2.705543
## lower: 0.5252525
## upper: 0.7030303
```

Thus

$$C(\mathbf{x}) = \{\theta : \theta \in [0.5252525, 0.7030303]\}$$

We accept H_0 in this case.