Solutions for lab session 1

- **Exercises**

A

1. Model : Y is a random variable (r.v.), the number of events, Y ~ Bin (n, π), events are independent with the same success probability for each individual.

   Probability function:

$$\Pr(Y=k)=\binom{n}{k}\pi^{k}(1-\pi)^{n-k}, k=0,1,....,n.$$

   Likelihood function

$$L(\pi|k)=\binom{n}{k}\pi^{k}(1-\pi)^{n-k}, 0<\pi<1.$$

   Log-Likelihood function

$$l(\pi|k)=\ln\left\{\binom{n}{k}\right\}+k\ln\left(\frac{\pi}{1-\pi}\right)+n\ln(1-\pi), 0<\pi<1$$

2. Model: X is a r.v., systolic blood pressure (mmHg), X ~ N(μ,5).

   Probability density function:

$$f(\chi|\mu)=\frac{1}{\sqrt{2\pi\times 5}}\exp\left\{-\frac{(\chi-\mu)^2}{2\times 5}\right\}, -\infty<\chi<\infty.$$

   Likelihood function

$$L(\mu|\chi)=\frac{1}{\sqrt{2\pi\times 5}}\exp\left\{-\frac{(\chi-\mu)^2}{2\times 5}\right\}, -\infty<\mu<\infty.$$

   Log-Likelihood function

$$l(\mu|\chi)=-\frac{1}{2}\ln(10\pi)-\frac{(\chi-\mu)^2}{10}$$

3. Model: Y_i, (I=1,…..n) are i.i.d. r.v.'s, cholesterol (mmol/l)
$Y_i \sim N(\mu, 4)$.

Probability density function:

$$f\left(y_1, y_2, ...y_n \mid \mu\right) = f\left(y \mid \mu\right) = \prod_{i=1}^{n} f\left(y_i \mid \mu\right) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\times4}} \exp\left\{-\frac{(y_i-\mu)^2}{2\times4}\right\}$$

$$= \left(\frac{1}{\sqrt{2\pi\times4}}\right)^n \exp\left\{-\frac{\sum_{i=1}^{n}(y_i-\mu)^2}{2\times4}\right\}, -\infty < y_i < \infty.$$

Likelihood function:

$$L(\mu \mid y) = \left(\frac{1}{\sqrt{2\pi\times4}}\right)^n \exp\left\{-\frac{\sum_{i=1}^{n}(y_i-\mu)^2}{2\times4}\right\}, -\infty < \mu < \infty.$$

Log-likelihood function:

$$l(\mu \mid y) = -\frac{n}{2}\ln\left(2\pi\times4\right) - \frac{1}{2\times4}\sum_{i=1}^{n}(y_i-\mu)^2, -\infty < \mu < \infty.$$

B.

1. k=2, n=6

Log-likelihood function:

$l(\pi \mid k=2)$ = constant + 2 ln (π) + 4 ln (1-π)

$$l'\left(\pi\right) = \frac{2}{\pi} - \frac{4}{1-\pi},$$

$$l'\left(\hat{\pi}\right) = 0 \Rightarrow \hat{\pi} = \frac{1}{3}$$

2. n = 10

Log-likelihood function:

$$l(\mu \mid y) = \text{constant} - \frac{1}{8}\sum_{i=1}^{10}(y_i-\mu)^2$$

$$l'(\mu) = \frac{1}{4}\sum_{i=1}^{10}(y_i - \mu),$$

$$l'(\hat{\mu}) = 0 \Leftrightarrow \frac{1}{4}\sum_{i=1}^{10}(y_i - \hat{\mu}) = 0 \Leftrightarrow \sum_{i=1}^{10} y_i - 10\hat{\mu} = 0 \Leftrightarrow \hat{\mu} = \sum_{i=1}^{10} y_i / 10 = 6.15$$

A Stata program to sketch the likelihood:

```
drop _all
set obs 451

*Enter data
gen y=6 if _n==1
replace y=6.2 if _n==2
replace y=6.8 if _n==3
replace y=5.3 if _n==4
replace y=5.9 if _n==5
replace y=6.1 if _n==6
replace y=6 if _n==7
replace y=7 if _n==8
replace y=5.9 if _n==9
replace y=6.3 if _n==10

*Variance = 4
local s2=4

*Create sequence of possible m values (-6 to 12 ;step 0.05)
gen m=-6+(_n-1)/20

gen y_mi_sq=.
gen loglike=.

*For all possible m values calculate the log-likelihood

local i=1
while(`i'<=451){
replace  y_mi_sq=(y-m[`i'])^2
egen Sy_mi_sq=sum( y_mi_sq)
replace  loglike= -(1/(2*`s2'))*Sy_mi_sq[1] if _n==`i'
drop Sy_mi_sq
local i=`i'+1
}

*Store maximum likelihood in `y' and MLE in `x'
gsort -loglike
local y=loglike[1]
local x=m[1]

sort m
*Graph the log-likelihood vs. m values
sc  loglike m,xlab(-6(1)18) xline(`x') yline(`y') ms(i) c(l)
```