20. Fitting of Normal distribution and Exponential distribution

1. Normal distribution

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad -\infty < x < \infty$$

```
#H0: Fit of normal distribution is good.
#H1: Fit of normal distribution is not good.
midx=seq(62.5,97.5,length = 8)
midx
```

Ex1: Fit a Normal Distribution to the following data.

```
[1] 62.5 67.5 72.5 77.5 82.5 87.5 92.5 97.5
```

```
f=c(3,21,150,335,326,135,26,4)
meanx=sum(f*midx)/sum(f)
meanx
```

```
[1] 79.945
```

```
sd=(sum(f*(midx-meanx)^2)/sum(f))^0.5
sd
```

```
[1] 5.444904
```

```
l=seq(60,100,length=9)
l
```

```
[1]  60  65  70  75  80  85  90  95 100
```

```
z=(l-meanx)/sd
z
```

```
[1] -3.66305844 -2.74476853 -1.82647862 -0.90818872  0.01010119  0.92839110
[7]  1.84668100  2.76497091  3.68326082
```

```
cdf=round(pnorm(l,meanx,sd),4)
cdf
```

```
[1] 0.0001 0.0030 0.0339 0.1819 0.5040 0.8234 0.9676 0.9972 0.9999
```

```r
cdf=c(0,cdf)
pcf=diff(cdf)
f=c(0,f)
ex=round(pcf*sum(f),0)
fr.dist=data.frame(f,ex)
fr.dist
```

```
    f  ex
1   0   0
2   3   3
3  21  31
4 150 148
5 335 322
6 326 319
7 135 144
8  26  30
9   4   3
```

```r
o=c(sum(f[1:3]),c(f[4:7]),sum(f[8:9]))
o
```

```
[1]  24 150 335 326 135  30
```

```r
e=c(sum(ex[1:3]),c(ex[4:7]),sum(ex[8:9]))
e
```

```
[1]  34 148 322 319 144  33
```

```r
chcal=sum((o-e)^2/e)
chcal
```

```
[1] 4.481881
```

```r
df=length(o)-2-1
df
```

```
[1] 3
```

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

```
[1] 7.814728
```

```r
if (chcal <= chtab) {
  cat("Normal distribution fits the data")
} else {
  cat("Normal distribution does not fit the data")
}
```

```
Normal distribution fits the data
```

```r
#H0: Fit of normal distribution is good.
#H1: Fit of normal distribution is not good.
midx=seq(155,235,length = 9)
midx
```

Ex2: Fit a Normal Distribution to the following data.

```
[1] 155 165 175 185 195 205 215 225 235
```

```r
f=c(9,24,51,66,72,48,21,6,3)
meanx=sum(f*midx)/sum(f)
meanx
```

```
[1] 189.8
```

```r
sd=(sum(f*(midx-meanx)^2)/sum(f))^0.5
sd
```

```
[1] 16.15426
```

```r
l=seq(150,240,length=10)
l
```

```
 [1] 150 160 170 180 190 200 210 220 230 240
```

```r
z=(l-meanx)/sd
z
```

```
 [1] -2.46374695 -1.84471506 -1.22568316 -0.60665126  0.01238064  0.63141254
 [7]  1.25044443  1.86947633  2.48850823  3.10754013
```

```r
cdf=round(pnorm(l,meanx,sd),4)
cdf
```

```
 [1] 0.0069 0.0325 0.1102 0.2720 0.5049 0.7361 0.8944 0.9692 0.9936 0.9991
```

```r
cdf=c(0,cdf)
pcf=diff(cdf)
f=c(0,f)
ex=round(pcf*sum(f),0)
fr.dist=data.frame(f,ex)
fr.dist
```

```
    f ex
1   0  2
2   9  8
3  24 23
4  51 49
```

```
5  66 70
6  72 69
7  48 47
8  21 22
9   6  7
10  3  2
```

```
o=c(sum(f[1:2]),c(f[3:8]),sum(f[9:10]))
o
```

```
[1]  9 24 51 66 72 48 21  9
```

```
e=c(sum(ex[1:2]),c(ex[3:8]),sum(ex[9:10]))
e
```

```
[1] 10 23 49 70 69 47 22  9
```

```
chcal=sum((o-e)^2/e)
chcal
```

```
[1] 0.6508483
```

```
df=length(o)-2-1
df
```

```
[1] 5
```

```
chtab=qchisq(0.95,df)
chtab
```

```
[1] 11.0705
```

```
if (chcal <= chtab) {
  cat("Normal distribution fits the data")
} else {
  cat("Normal distribution does not fit the data")
}
```

```
Normal distribution fits the data
```

2. Exponential distribution

$f(a < x < b) = e^{-\theta * a} - e^{-\theta * b}$, **(where a = lower limit and b = upper limit)**

```r
#H0: Fit of exponential distribution is good.
#H1: Fit of exponential distribution is not good.
midx=seq(1.5,16.5,3)
midx
```

Ex: Fitting of Exponential distribution

```
[1]  1.5  4.5  7.5 10.5 13.5 16.5
```

```r
f=c(190,70,25,10,4,1)
m=sum(f*midx)/sum(f)
m
```

```
[1] 3.21
```

```r
theta=1/m
theta
```

```
[1] 0.3115265
```

```r
l=seq(0,12,3)
u=seq(3,15,3)
pr=exp(-theta*l)-exp(-theta*u)
pr=c(pr,1-sum(pr))
ex=round(pr*sum(f),0)
ex
```

```
[1] 182  72  28  11   4   3
```

```r
fr.dist=data.frame(f,ex)
fr.dist
```

```
    f  ex
1 190 182
2  70  72
3  25  28
4  10  11
5   4   4
6   1   3
```

```r
o=c(c(f[1:4]),sum(c(f[5:6])))
o
```

```
[1] 190  70  25  10   5
```

```r
e=c(c(ex[1:4]),sum(c(ex[5:6])))
e
```

```
[1] 182  72  28  11   7
```

```r
chcal=sum((o-e)^2/e)
chcal
```

```
[1] 1.39097
```

```r
df=length(o)-1-1
df
```

```
[1] 3
```

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

```
[1] 7.814728
```

```r
if (chcal <= chtab) {
  cat("Exponential distribution fits the data")
} else {
  cat("Exponential distribution does not fit the data")
}
```

```
Exponential distribution fits the data
```