

II. Computing

2.4. Simple simulations in statistics

1. Consistency and asymptotic normality of M.L.E.:

Example:

Let $X_1, X_2, \dots, X_n \sim Poisson(\lambda)$. Then, the M.L.E. for λ is $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$.

Example (Splus):

```
rp30=rpois(30,10)          # random Poisson sample  
mean(rp30)                 # sample mean  
rp1000=rpois(1000,10)  
mean(rp1000)  
rp100000=rpois(100000,10)  
mean(rp100000)
```

Example:

According to central limit theorem, $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$. Please generate the data with sample sizes 10, 100, 1000 from Binomial(3, 0.2). Then, run a simulation study to justify the central limit theorem.

Example (Splus):

```
par(mfrow=c(1,3))  
  
n=3  
p=0.2  
mu=n*p  
s=sqrt(3*p*(1-p))  
stat1=stat2=stat3=rep(0,1000)  
for(i in 1:1000){  
  stat1[i]=(mean(rbinom(10,n,p))-mu)/(s/sqrt(10))  
  stat2[i]=(mean(rbinom(100,n,p))-mu)/(s/sqrt(100))  
  stat3[i]=(mean(rbinom(1000,n,p))-mu)/(s/sqrt(1000))  
}  
qqnorm(stat1)
```

```

qqline(stat1)
qqnorm(stat2)
qqline(stat2)
qqnorm(stat3)
qqline(stat3)

```

2. Comparison of different estimates:

Suppose we want to estimate the mean μ and the variance σ^2 of a normal population. A sample of data x_1, x_2, \dots, x_n can be used to estimate μ and σ^2 .

Let $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ and m be the sample mean and sample median, respectively. Both

\bar{x} and m can be used to estimate μ . Also, let $s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$ and $(s^*)^2 = \frac{\sum_{i=1}^n (x_i - m)^2}{n}$ be the estimates of σ^2 .

Example (Splus):

```

mean1=median1=s2=sstar2=rep(0,1000)
for(i in 1:1000){
  data=rnorm(30,mean=5,sd=2)          # random normal sample
  mean1[i]=mean(data)                 # sample mean
  median1[i]=median(data)            # sample median
  s2[i]=var(data)                   # s^2
  sstar2[i]=var(data)*(29/30)       # (s*)^2
}
mean(mean1)
mean(median1)
mean(s2)
mean(sstar2)
var(mean1)
var(median1)

```

Note:

In the above example, the sample mean and the sample variance perform better than the other estimates. In the following example, some data from a different population are added to the regular data. It turns out the sample median then performs better than the sample mean.

Example (Splus):

```
for(i in 1:1000){  
    data=rnorm(27,mean=5,sd=2)      # the regular data  
    adddata=rnorm(3,mean=15,sd=5)    # the data from a different  
                                    # population  
    data=c(data,adddata)  
    mean1[i]=mean(data)            # sample mean  
    median1[i]=median(data)        # sample median  
    s2[i]=var(data)                #  $s^2$   
    sstar2[i]=var(data)*(29/30)    #  $(s^*)^2$   
}  
  
mean(mean1)  
mean(median1)  
mean(s2)  
mean(sstar2)
```

3. Time consideration:

In Splus, `dos.time()` and `proc.time()` can be used to returns the time in seconds taken to evaluate the expression.

Example (Splus):

```
dos.time(rnorm(1000))  
start=proc.time()  
for(i in 1:1000){  
    print(i)  
}  
end=proc.time()  
end-start
```

Useful links:

Firefox : <http://moztw.org/>

Google Map : <https://www.google.com.tw/>