Pràctica 6

Teorema central del límit

6.1 Teorema central del límit

Recordem que el Teorema central del límit ens diu que la distribució de la suma de moltes variables aleatòries independents amb la mateixa distribució, independentment de quina sigui aquesta, és aproximadament una llei normal.

La versió general del teorema és:

Sigui $X_1,\ X_2, \ldots,\ X_n$ una mostra aleatòria d'una distribució amb esperança μ i variància σ^2. Aleshores, quan n és prou gran ($n > 30$),

$$\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} = \frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}},$$

es comporta com una distribució normal.

Comprovarem el resultat d'aquest teorema a través de les simulacions dels següents problemes.

6.2 Problemes

1. Fixem un N, per exemple $N = 200$ i un B, per exemple $B = 400$. Realitzeu les següents operacions:

   (a) Genereu N mostres de mida B, d'una distribució de Bernoulli amb paràmetre $p = 0.5$, en forma d'una matriu A, de dimensió $N \times B$.

   Calculeu la llista a de les N sumes de les N files de A, tal i com es va explicar al guió de la Pràctica 3.

   ```
   p<-0.5
   N<-200
   B<-400
   A<-matrix(rbinom(N*B,1,p),N)
   a<-apply(A,1,sum)
   ```

   (b) Calculeu l'esperança teòrica μ i la variància teòrica σ^2 (en aquest cas de la llei Bernouilli(p)).

   (c) De cada fila de la matriu calculeu la mitjana, obtenint així un vector de mitjanes.

   (d) Centreu i normalitzeu el vector de mitjanes. (Recordeu que cal restar la mitjana teòrica μ i dividir per la desviació teòrica $\sigma/\sqrt{n}$)

   (e) Dibuixeu l'histograma del vector de mitjanes normalitzat amb la corba de la densitat de la normal, què observem? Cal augmentar el valor de N per obtenir una aproximació millor?

2. Genereu una matriu $N \times B$ corresponent a N mostres de mida B d'una variable aleatòria amb funció de massa de probabilitat :

$$P(X = -2) = 1/8,\ P(X = -1) = 1/8,\ P(X = 0) = 1/2$$

$$P(X = 1) = 1/8,\ P(X = 2) = 1/8.$$

   Podeu prendre N=100 i B=100. Calculeu i responeu els mateixos apartats que en l'exercici 1 per aquesta variable aleatòria i per aquests valors de N i B.

3. Genereu una matriu $N \times B$ corresponent a N mostres de mida B d'una variable aleatòria amb funció de massa de probabilitat

$$P(X = 5) = 1/3,\ P(X = 7) = 1/2,\ P(X = 9) = 1/6.$$

   Podeu prendre N=200 i B= 400. Calculeu i responeu els mateixos apartats que en l'exercici 1 per aquesta variable aleatòria.

4. Genereu una matriu $N \times B$ corresponent a N mostres de mida B d'una variable aleatòria contínua amb funció de densitat:

$$f(x) = 3x^2 1_{(0,1)}(x).$$

   Podeu prendre N=200 i B= 400. Calculeu i responeu els mateixos apartats que en l'exercici 1 per aquesta variable aleatòria.

5. Fixem un N, per exemple $N = 200$ i un B, per exemple $B = 400$, i un t una mica més petit que N, per exemple $t = 186$.
   Realitzeu les següents operacions:

   (a) Genereu N mostres de mida B, d'una distribució de Bernoulli amb paràmetre $p = 0.5$, en forma d'una matriu A, de dimensió $N \times B$. Calculeu la llista a de les N sumes de les N files de A

   (b) Quins són els paràmetres μ i σ de la llei normal que aproxima bé la distribució de a?

   (c) Calculeu utilitzant l'aproximació de la Normal de l'apartat anterior el nombre d'elements de la llista a que haurien de ser menors que t. Comproveu que heu obtingut una bona aproximació.

   (d) Quina és la distribució exacta de a? (suposant que els elements de les files de A es poden considerar realitzacions de variables de Bernoulli independents, igualment distribuïdes). Calculeu utilitzant la distribució exacta el nombre d'elements de la llista a que haurien de ser menors que t, compareu-lo amb els resultats obtinguts a l'apartat anterior.

   (e) Dibuixeu la funció de massa de probabilitat (exacta) de a. Superposeu-hi la funció de densitat de probabilitat de la llei normal de l'apartat 5b.

   (f) Repetiu l'experiment amb diferents valors de N i variant t.

6. Un poble té dues sales de cinema d'igual cabuda i 1000 habitants cinèfils. En un dia donat tots ells van al cinema, cadascun a una de les dues sales, indiferentment i sense influir-se els uns als altres.

   Calculeu la cabuda mínima que han de tenir les sales per tal que la probabilitat d'*overbooking* sigui del u per cent com a màxim.

7. En una elecció un partit obté el 45% dels vots en una població homogènia. Si extraïem a l'atzar dues mostres de $n = 100$ votants cadascuna, quina és la probabilitat que la diferència $X - Y$ dels nombres de votants en les dues mostres sigui superior a 5?