Pràctica 3

Variables aleatòries discretes

3.1 Distribucions de probabilitat discretes conegudes

A la taula 3.1, per a les distribucions de probabilitat llistades a la columna de l'esquerra, tenim el seu nom en **R** i els paràmetres corresponents.

Taula 3.1: Noms en R de distribucions de probabilitat discretes

Distribució	Sufix del nom	Paràmetres addicionals
binomial	`binom`	`size, prob`
geomètrica	`geom`	`prob`
hypergeomètrica	`hyper`	`m, n, k`
binomial negativa	`nbinom`	`size, prob`
Poisson	`pois`	`lambda`
Wilcoxon	`wilcox`	`m, n`

Per cadascuna de les distribucions podem cridar quatre funcions diferents afegint al nom el prefix d, p, q o r:

`dnom(valors, paràmetres)` = la funció de massa (o densitat) de probabilitat,
`pnom(valors, paràmetres)` = la funció de distribució de probabilitat (Fdd),
`qnom(valors, paràmetres)` = la funció quantila, és a dir, la (pseudo)inversa de la Fdd,
`rnom(valors, paràmetres)` = nombres aleatoris segons la distribució de probabilitat donada.

Aquestes funcions ens són útils per generar variables amb aquestes distribucions, per dibuixar la funció de massa de probabilitat, per dibuixar la funció de distribució de probabilitat i per calcular probabilitats.

3.2 Generació de variables discretes amb distribució coneguda

Es tracta de generar llistes de nombres aleatoris amb distribucions discretes conegudes que heu vist a teoria com: la Bernoulli, la Binomial la Poisson i la Geomètrica.

1. Si volem generar 10 nombres aleatoris amb distribució Binomial, $\mathrm{Bn}(n, p)$, de paràmetres $n = 20$ i $p = 0.2$, hem de fer:

```
x<-rbinom(10,20,0.2)
```

Podem calcular la mitjana i la variància corregida la mostra simulada amb

```
mean(x)
var(x)
```

Són aproximadament iguals als valors teòrics? Fem el mateix amb 100 repeticions i veieu si l'aproximació és més bona:

```
N<-100
x<-rbinom(N,20,0.2)
mean(x)
var(x)
```

i si augmentem la mostra a 500?

2. Genereu 400 nombres aleatoris amb distribució Bernoulli (Bn$(1, p)$) amb $p = 0.3$ i poseu-los en forma de matriu amb 40 files i 10 columnes. La funció `apply` ens permet fer càlculs paral·lels:

```
N<-400
x<-matrix(rbinom(N,1,0.3),nrow=40)
apply(x,1,mean) # fa mitjanes per files
apply(x,2,mean) # fa mitjanes per columnes
```

Proveu de calcular les desviació estàndard corregida i les medianes per files i columnes.

Exercicis:

1. Feu ara simulacions de la Poisson per exemple amb $N = 50$, i paràmetre $\lambda = 2$. Calculeu la mitjana i la variància corregida de les simulacions i compareu-les amb els valors teòrics. Augmenteu N fins obtenir uns bona aproximació.

2. Per fer simulacions de la Geomètrica cal anar amb més cura. La Geomètrica compta els intents **necessaris** fins a obtenir el primer èxit. En **R** la Geomètrica compta el nombre d'intents **fallits** fins a obtenir el primer èxit, per tant en compta un de menys respecte el que heu vist a teoria. Per exemple, si volem fer 100 simulacions del nombre de llançaments necessaris d'un dau fins que surti un 6:

```
N<-100
x<-rgeom(N,1/6)
x<-x+1
```

Calculeu la mitjana i la variància corregida de les simulacions i compareu-les amb els valors teòrics.

3.3 Funció de massa, F.d.d., càlculs de probabilitats, esperança.

Prenem com exemple la distribució Binomial Bn$(8, 0.4)$ i el llançament d'un dau perfecte per veure com es fan els dibuixos de la funció de massa de probabilitat, de la funció de distribució i com es calculen diferents probabilitats.

- **dibuix de la funció de massa de probabilitat**
  - Hem de tenir un vector amb els valors de la variable i un vector amb les probabilitats. En el cas de la Bn$(8, 0.4)$:

```
x<-c(0:8)
prob<-dbinom(x,8,0.4)
```

Pel llançament d'un dau:

```
x<-c(1:6)
prob<-c(rep(1/6,6))
```

Per fer el dibuix de la massa de probabilitat d'una distribució discreta hem de seguir la instrucció següent: En el cas de la Bn$(8, 0.4)$:

```
plot(x,prob,type="h", xlim=c(0,8), ylim=c(0,1))
```

Pel llançament d'un dau:

```
plot(x,prob,type="h", xlim=c(0,6), ylim=c(0,1))
```

evidentment els límits de `xlim` cal ajustar-los als valors de la variable.

- **F:d.d.**
  - Per fer el dibuix de la funció de distribució hem de seguir les instruccions següents:

    ```
    acum<-cumsum(prob)
    s<-stepfun(x, c(0,acum))
    plot(s,verticals=FALSE)
    ```

- **Càlcul de probabilitats**
  - Si volem calcular probabilitats concretes, en el cas que la nostra variable X és una Bn$(8, 0.4)$:
    $P(X \leq 3)$ es calcula mitjançant `pbinom(3,8,0.4)`
    $P(X = 3)$ es calcula mitjançant `dbinom(3,8,0.4)`
    $P(X \geq 3) = 1 - P(X < 3) = 1 - P(X \leq 2)$ es calcula mitjançant `1-pbinom(2,8,0.4)`
    Comproveu que $P(X \leq 3)$ també es pot calcular fent
    `dbinom(0,8,0.4)+dbinom(1,8,0.4)+dbinom(2,8,0.4)+dbinom(3,8,0.4).`
  - Per trobar el valor de la variable en que la funció de distribució assoleix una probabilitat concreta, per exemple, si volem saber el valor k en que $P(X \leq k) = 0.5940864$, fem:

    ```
    k<-qbinom(0.5940864,8,0.4)
    ```

    Com no totes les probabilitats p entre $(0, 1)$ s'assoleixen, el que fa **R** amb aquesta funció és calcular k tal que
    $$P(X \leq k) \geq p \text{ i } P(X \leq k - 1) < p.$$
    Proveu `qbinom(0.59,8,0.4)` i `qbinom(0.595,8,0.4)` i veieu que són diferents.
  - En el cas que la nostra variable X és el resultat del llançament d'un dau no perfecte, i suposant que en `prob` tenim el vector de les probabilitats:
    $P(X \leq 3)$ es calcula mitjançant `prob[1]+prob[2]+prob[3]`
    $P(X = 3)$ es calcula mitjançant `prob[3]`
    $P(X \geq 2) = 1 - P(X = 1)$ es calcula mitjançant `1-prob[1]`

- **Esperança**
  - Recordeu que $\mathbf{E}[X] = \sum_{i=1}^{n} x_i P(X = x_i)$. Per tant el seu càlcul el podem fer facilment:

    ```
    esp<-sum(x*prob)
    ```

    Exercici: Penseu la forma de calcular la variància i la desviació típica.

3.4 Problemes

1. La variable aleatòria X té les probabilitats següents:

valors x_i	2	3	6	7	8	10
probabilitats $P[X = x_i]$	0.2	0.1	0.2	0.1	0.2	0.2

   a) Representeu la funció de massa de probabilitat i la corresponent funció de distribució de probabilitat.

   b) Calculeu l'esperança, la variància i la desviació típica de X.

2. Dibuixeu la funció de massa de probabilitat i la funció de distribució de probabilitat d'una llei $B(p)$ amb $p = 0.5$; $p = 0.005$; $p = 0.90$.

3. Dibuixeu la funció de massa de probabilitat i la funció de distribució de probabilitat d'una llei $Bn(n, p)$ en els casos següents:

   a) $n = 3, \quad p = 0.4$.

   b) $n = 10, \quad p = 0.001$.

   c) $n = 20, \quad p = 0.8$.

4. Un agent d'assegurances ven pòlisses a 5 individus, tots de la mateixa edat. Sabem que la probabilitat que un individu d'aquesta edat visqui 30 anys més és de $3/5$. Determineu la probabilitat que després de 30 anys visquin:

   a) Tots 5 individus. b) Com a màxim 3 individus.
   c) Almenys 3 individus. d) Exactament 2 individus.
   e) Almenys 1 individu.

5. Un examen té 8 preguntes amb 4 respostes possibles de les quals només n'hi ha una de correcta. Per aprovar l'examen cal contestar bé almenys la meitat de les preguntes. Un alumne ha contestat a l'atzar totes 8 preguntes. Calculeu:

   a) La probabilitat d'aprovar l'examen.

   b) La probabilitat d'encertar totes les preguntes.

   c) El nombre mitjà de respostes correctes i la seva variància.

6. Dibuixeu les funcions de massa de probabilitat i les funcions de distribució de probabilitat de les lleis $Pois(\lambda)$ amb $\lambda = 1$ i $\lambda = 20$.

7. En una fàbrica el nombre d'accidents per setmana segueix una llei $Pois(5)$. Calculeu:

   a) Probabilitat que en una setmana hi hagi algun accident.

   b) Probabilitat que en una setmana hi hagi dos accidents.

c) Probabilitat que en una setmana hi hagi almenys tres accidents.

d) El nombre mitjà d'accidents per setmana i la seva variància.

8. El nombre mitjà d'automòbils que arriben a una estació de subministrament de gasolina és de 210 per hora. Si aquesta estació pot atendre a un màxim de 10 automòbils per minut, determineu la probabilitat que en un minut donat arribin a l'estació més automòbils dels que es poden atendre. Suposeu que el nombre d'automòbils que arriben a l'estació durant 1 minut segueix una distribució de Poisson.