Probabilites des lois usuelles en R (1)

October 17, 2018

Luc Adjengue

Contents

1 Introduction

Dans R, chaque loi a une abréviation (*norm* pour la loi normale par exemple). Pour faire des calculs sur une variable aléatoire suivant cette loi, on précède l'abréviation par une lettre indiquant le type de calcul que l'on souhaite faire:

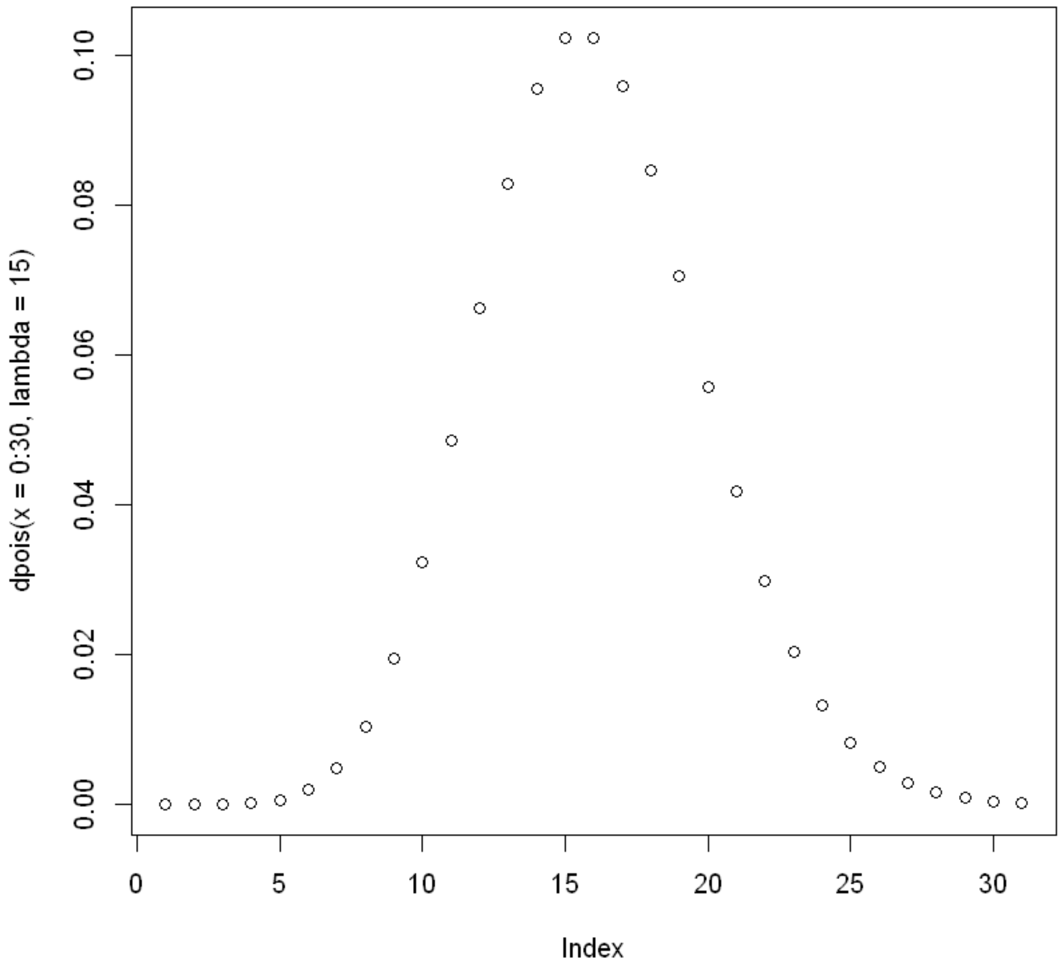
- d pour la densité
- p pour la fonction de répartition (ie le calcul de probabilités)
- q pour les quantiles
- r pour générer aléatoirement des valeurs

Pour les fonctions p et q faisant référence à la fonction de répartition pour la première et à au calcul des quantiles pour la seconde, en plus des paramètres propres à chaque distribution, il y a le paramètre *lower.tail* à ajouter.

Par exemple, pnorm(m = moyenne, sd = ecart_type, lower.tail = TRUE). Sa valeur est *TRUE* par défaut, indiquant que les probabilités calculées sont $P(X \leq x)$.

Il peut être utile parfois d'avoir une idée du graphe de la fonction de densité de la loi de probabilité. La ligne de code à écrire est : *plot(d_fonctionADefinir(x=debut:fin, paramètre de la fonction))*. Ci-dessous un exemple avec la loi de Poisson décrite plus bas.

```
In [15]: plot(dpois( x=0:30, lambda=15 ))
```



2 Loi Binomiale

La loi binomiale admet deux paramètres:

- la taille n de l'échantillon ou le nombre de tirages, appelée *size* dans R
- la probabilité p de succès.

On peut soit préciser chaque nom de paramètre et sa valeur (cas 1 ci-dessous), soit mettre les paramètres dans l'ordre sans préciser leur nom (cas 2).

Pour $X \sim Binom(n = 20, p = 0.4)$, calcul de P(X = 5):

```
In [1]: # calcul de la probabilité d'avoir x succès parmi les n tirages, CAS 1
        dbinom(x = 5,size = 20,p = 0.4)
```

0.074647019528871

```
In [2]: # calcul de la probabilité d'avoir x succès parmi les n tirages, CAS 2
        dbinom(5, 20, 0.4)
```

0.074647019528871

Pour $X \sim Binom(n = 20, p = 0.4)$, calcul de $P(3 < X < 16)$:

```
In [4]: sum(dbinom(4:15, 20, 0.4))
```

0.983721806088823

Pour $X \sim Binom(n = 20, p = 0.4)$, calcul de $P(3 \leq X \leq 16)$:

On peut faire le calcul de deux façons, soit en utilisant la somme des probabilités de chaque valeur comprise dans l'inégalité, soit en utilisant la fonction de répartition ie la notation **pbinom** de R en utilisant l'égalité $P(3 \leq X \leq 16) = P(X \leq 16) - P(X \leq 2)$

```
In [7]: # CAS 1 avec la somme de la fonction de masse
        sum(dbinom(3:16, 20, 0.4))
```

0.996341182970179

```
In [8]: # CAS 2 avec la fonction de répartition
        pbinom(16,20,0.4) - pbinom(2,20,0.4)
```

0.996341182970179

3 Loi géométrique

La loi géométrique admet un seul paramètre : la probabilité *prob* de succès des épreuves de Bernouilli.

Attention, dans R, la définition de la loi géométrique n'est pas la même que celle vue en cours. Dans R, la loi géométrique compte le nombre d'épreuves qui précèdent le premier succès et les valeurs possibles commencent à 0.

Dans le cours, on a défini la loi géométrique comme le nombre d'épreuves nécessaires avant d'obtenir un premier succès et l'ensemble des valeurs possibles commence à partir de 1.

Pour $X \sim Geom(p = 0.4)$, calcul de P(X = 2):

```
In [9]: dgeom(2,prob = 0.4)
```

0.144

Pour $X \sim Geom(p = 0.4)$, calcul de $P(X \leq 3)$:

```
In [10]: pgeom(3,prob = 0.4)
```

0.8704

4 Loi de Poisson

La loi de Poisson admet un paramètre: c, qui désigne le nombre moyen de réalisations de l'évènement dans l'intervalle de temps considéré ou dans l'espace considéré. Dans R, ce paramètre est appelé *lambda*.

Pour $X \sim Poisson(c = 4)$, calcul de P(X = 2):

```
In [5]: dpois(2,lambda = 4)
```

0.146525111109873

Pour $X \sim Poisson(c = 8)$, calcul de $P(X \geq 10)$:

```
In [6]: 1 - ppois(q = 10, lambda = 8)
```

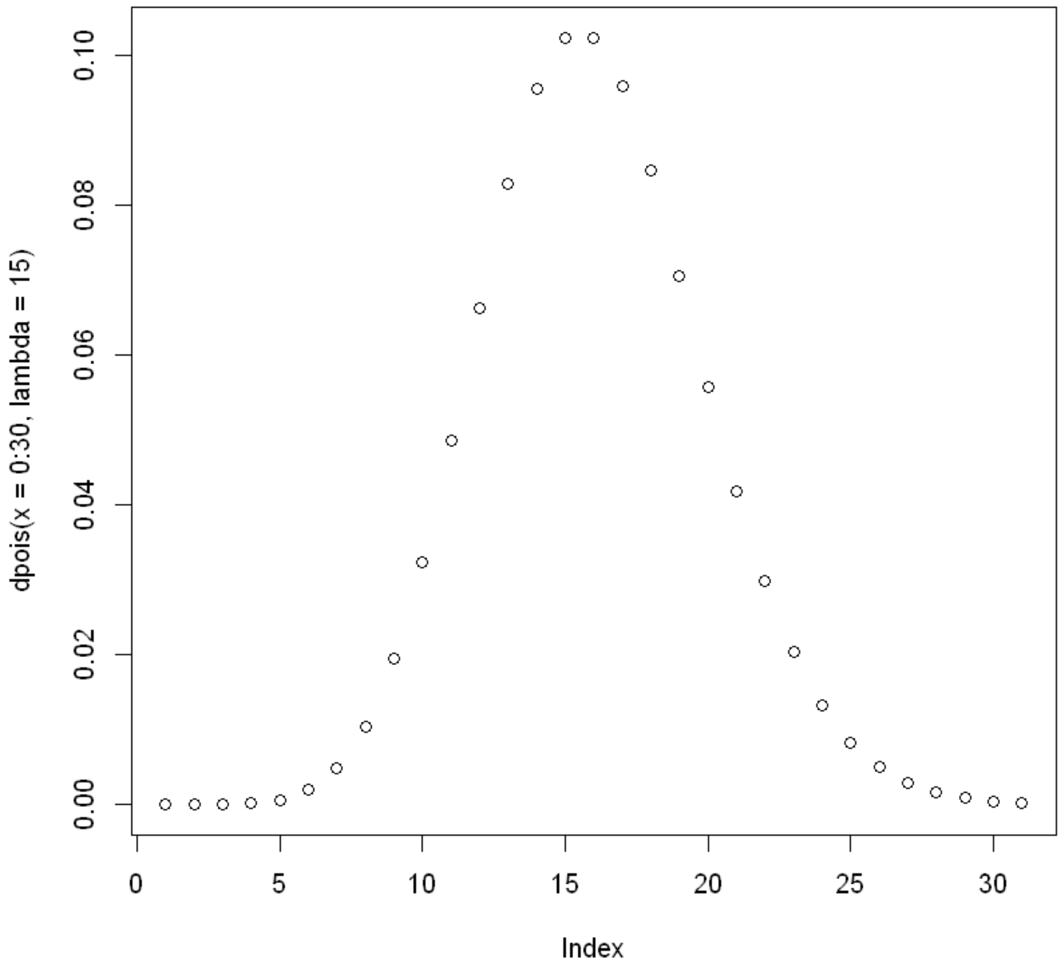
0.184114207441454

Si on cherche la plus petite valeur k telle que $P(X \leq k) \geq 0,95$, avec $X \sim Poisson(c = 15)$:

```
In [8]: qpois(0.95,lambda = 15,lower.tail = TRUE)
```

22

```
In [14]: plot( dpois( x=0:30, lambda=15 ))
```



Si on cherche k tel que $P(X > k) \geq 0,90$, avec $X \sim Poisson(c = 15)$:

```
In [9]: qpois(0.9,lambda = 15,lower.tail = FALSE)
```

10

5 Loi Normale

La loi normale admet deux paramètres: la moyenne notée m dans R et l'écart-type noté sd dans R.

- dnorm(x) : densité de probabilité en x de la loi normale centrée réduite
- dnorm(x, mean = m, sd = s) : densité de probabilité en x de la loi normale d'espérance m et d'écart-type s

- pnorm(x) : fonction de répartition en x de la loi normale centrée réduite
- pnorm(x, mean = m, sd = s) : fonction de répartition en x pour la loi normale d'espérance m et d'écart-type s

Pour $X \sim N(0,1)$, calcul de $P(X \leq 0.6)$:

```
In [10]: pnorm(.6)
```

0.725746882249926
Pour $X \sim N(0,1)$, calcul de $P(X \geq 10)$:

```
In [11]: # première méthode
         1-pnorm(10)
```

0

```
In [12]: # seconde méthode
         pnorm(10,lower.tail = FALSE)
```

7.61985302416053e-24
Pour $X \sim N(10,4)$, calcul de $P(X \leq 8)$:

```
In [13]: pnorm(8,m = 10,sd = 2)
```

0.158655253931457
Pour $X \sim N(10,4)$, on cherche k tel que $P(X \leq k) = .9$

```
In [14]: qnorm(0.9,m = 10,sd = 2)
```

12.5631031310892
Pour $X \sim N(10,4)$, on cherche k tel que $P(X > k) = .8$

```
In [15]: qnorm(0.8,m = 10,sd = 2)
```

11.6832424671458

6 Loi de Student

La loi de Student admet un seul paramètre: le degré de liberté noté *df* dans R.
Pour $X \sim t_2$, calcul de $P(X \leq 3)$:

```
In [16]: pt(3,df = 2)
```

0.952267016866645
Calcul du quantile $t_{0.05;10}$:

```
In [17]: qt(1-0.05,10)
```

1.81246112281168

7 Loi du khi-carré

La loi du khi-deux admet un unique paramètre: le nombre de degrés de liberté, noté *df* dans R.

Pour $X \sim \chi^2_4$, calcul de $P(X \leq 3)$:

```
In [18]: pchisq(3,df = 4)
```

0.442174599628925

Calcul de $\chi^2_{0.05,10}$

```
In [19]: qchisq(0.05,10,lower.tail = FALSE)
```

18.3070380532751

8 Loi F de Fisher

La loi de Fisher admet deux paramètres :

- *df1* le premier nombre de degrés de liberté
- *df2* le second nombre de degrés de liberté

Pour $X \sim F_{3,4}$, calcul de $P(X \leq 5)$:

```
In [20]: pf(5,df1 = 3,df2=4)
```

0.922981284596924

Pour $X \sim F_{3,4}$, calcul de $P(X \geq 4)$:

```
In [12]: pf(4,df1 = 3,df2=4,lower.tail = FALSE)
```

0.106911302347298

Vous pouvez vérifier que $F_{1-\alpha;u,v} = \frac{1}{F_{\alpha;v,u}}$

```
In [22]: qf(0.9,df1 = 4,df2 = 5)
```

3.52019624553412

```
In [23]: 1/ qf(0.1,df1 = 5, df2 = 4)
```

3.52019624553412