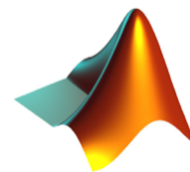




Laboratoire 1 : Simulation de variables aléatoires

Dr Stephan Robert
Octobre 2018

Credit: Dr. Rappos



1

Fonction de densité et Fonction de répartition

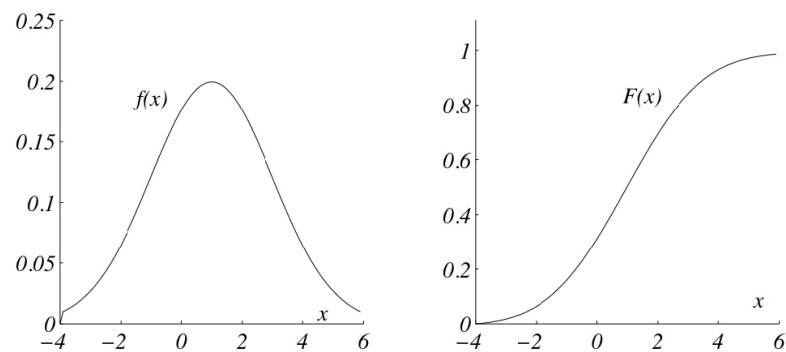


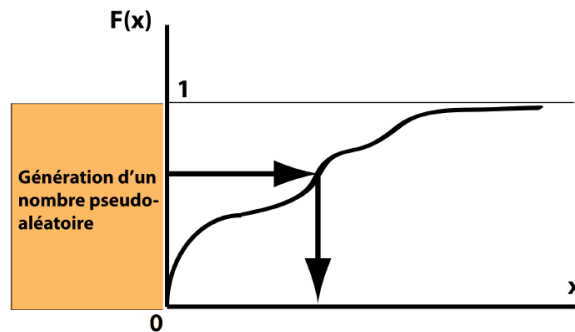
Figure: Densité et fonction de répartition d'une variable aléatoire Gaussienne (ou Normale) avec $\mu = 1$ et $\sigma = 2$

2

Génération de réalisations de variables aléatoires

Méthode de **transformation inverse**

Variable aléatoire discrète : $F(x_n) = \sum_{i=1}^n P(X = x_i)$



3

Exemple 1 : loi uniforme continue

La densité de la loi de probabilité uniforme continue sur $\Omega = [a, b]$ est définie par

$$f_X(x) = \frac{1}{b-a}$$

quand $a \leq x \leq b$ avec $a < b$, $f_X(x) = 0$ ailleurs. La fonction de répartition correspondante est donnée par

$$F_X(x) = \begin{cases} 0 & \text{si } x \leq a \\ (x-a)/(b-a) & \text{si } a < x \leq b \\ 1 & \text{si } x > b. \end{cases}$$

$$E[X] = (a+b)/2, \text{ Var}(X) = (b-a)^2/12$$

4

Exemple 1 : Simulation avec Octave

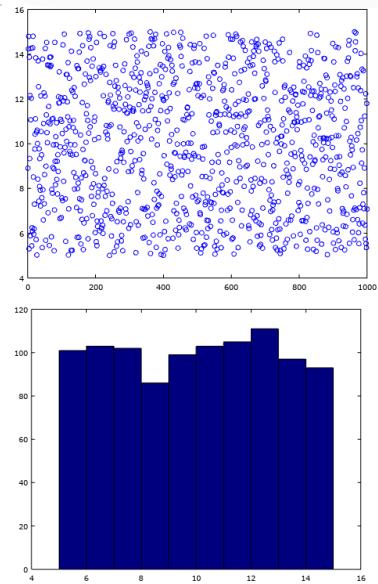
**N = 1000 valeurs
générées**
a = 5 b = 15

```
EDU>> x=rand(1000,1);
```

```
EDU>> x=5+10*x;
```

```
EDU>> plot(x,'o')
```

```
EDU>> hist(x)
```



5

Exemple 1 : Simulation avec Octave (2)

Moyenne

```
EDU>> S=0
EDU>> for t= 1:1000
S = S + x(t)
end
EDU>> m = S / 1000
m = 9.9528
```

Variance + écart-type

```
EDU>> S=0
EDU>> for t= 1:1000
S = S + ( x(t) - m ) ^ 2
end
EDU>> v = S / 1000
v = 8.7699
EDU>> s = sqrt(v)
s = 2.9614
```

Test

```
EDU>> m2 = mean(x)
m2 = 9.9528

EDU>> v2 = var(x)
v2 = 8.7787 (!!!!!!!!!!)

EDU>> help var

EDU>> v2 = var(x,1)
v2 = 8.7699
```

6

Exemple 1 : Valeurs théoriques

- Espérance mathématique

$$M = (b - a) / 2 = 10$$

- Variance

$$V = (b - a)^2 / 12 = 8.3333$$

- Ecart-type

$$s = \sqrt{V} = 2.8868$$

7

Différence entre variance de la population et d'un échantillon

• Population

- Moyenne, espérance mathématique

$$\mu = \frac{X_1 + X_2 + \dots + X_N}{N}$$

- Variance

$$Var = \frac{(X_1 - \mu)^2 + (X_2 - \mu)^2 + \dots + (X_N - \mu)^2}{N}$$

Echantillon

- Estimation de la moyenne

$$\hat{\mu} = \frac{x_1 + x_2 + \dots + x_N}{N}$$

- Estimation de la variance

$$\widehat{Var} = \frac{(x_1 - \hat{\mu})^2 + (x_2 - \hat{\mu})^2 + \dots + (x_N - \hat{\mu})^2}{N - 1}$$

8

Exemple 2 : échantillon

```

N = 10000
a = 5      b = 15

EDU>> x=rand(10000,1);

EDU>> x=5+10*x;

EDU>> m_pop = mean(x)
m_pop = 9.9863

EDU>> v_pop = var(x,1)
v_pop = 8.3842

Échantillon N=500 :

EDU>> S=0
EDU>> for t= 1:500
    S = S + x(t);
end
EDU>> m = S / 500
m = 9.8832

EDU>> S=0
EDU>> for t= 1:500
    S = S + ( x(t) - m) ^ 2
end
EDU>> v = S / 500
v = 8.2859
(la variance des 500
valeurs)

EDU>> v = S / 499
v = 8.3025 (une meilleure
estimation de la variance
de la population)

```

9

Exemple 3 : La loi exponentielle

- ▶ Un paramètre : λ ($\lambda > 0$)
- ▶ Fonction de densité

$$f_X(x) = \lambda e^{-\lambda x} \text{ pour } x \geq 0, \text{ nulle ailleurs}$$

- ▶ Fonction de répartition

$$F_X(x) = 1 - e^{-\lambda x} \text{ pour } x \geq 0, \text{ nulle ailleurs}$$

- ▶ $E[X] = 1/\lambda$, $Var[X] = 1/\lambda^2$.

10

Exemple 3 : La loi exponentielle (2)

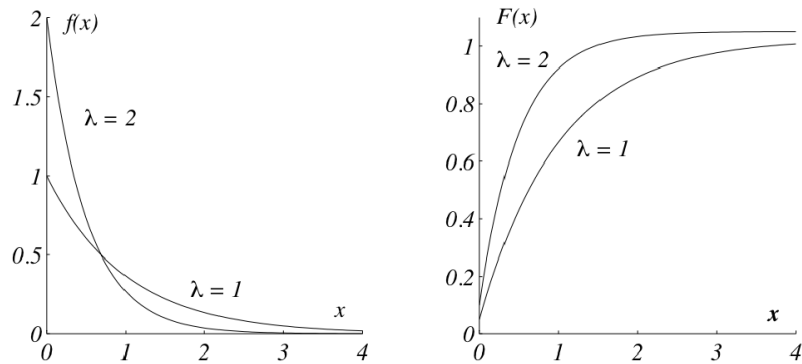


Figure: Densité et fonction de répartition d'une variable aléatoire exponentielle

11

Exemple 3 : Simulation avec Octave

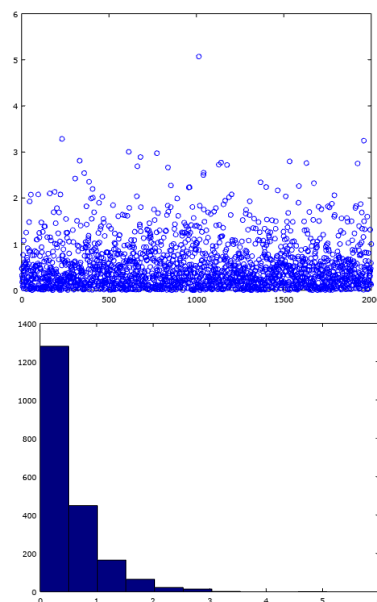
**N = 2000 valeurs
générées
 $\lambda = 2.0$**

```
EDU>> x=rand(2000,1);
```

```
EDU>> x= -log(1-x)/  
2.0;
```

```
EDU>> plot(x,'o')
```

```
EDU>> hist(x)
```



12

Exemple 3 : Simulation avec Octave (2)

Moyenne

```
EDU>> S=0
EDU>> for t= 1:2000
S = S + x(t);
end
EDU>> m = S / 2000
m = 0.51327
```

Variance

```
EDU>> S=0
EDU>> for t= 1:2000
S = S + ( x(t) - m) ^ 2
end
EDU>> v = S / 2000
v = 0.26528
EDU>> s = sqrt(v)
s = 0.51506
```

Test

```
EDU>> m2 = mean(x)
m2 = 0.51327
EDU>> v2 = var(x)
v2 = 0.26542
EDU>> v2 = var(x,1)
v2 = 0.26528
```

Valeurs théoriques

```
m = 1 / λ = 0.5
var = 1 / λ^2 = 0.25
```

13

Exemple 4 : La loi Gaussienne (ou Normale)

- ▶ Deux paramètres : μ et $\sigma^2 > 0$.
- ▶ Fonction de densité

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma)}$$

- ▶ Fonction de répartition

$$F_X(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-(u-\mu)^2/(2\sigma)} du$$

- ▶ $E[X] = \mu$, $Var(X) = \sigma^2$.

14

Exemple 4 : La loi Gaussienne (ou Normale)

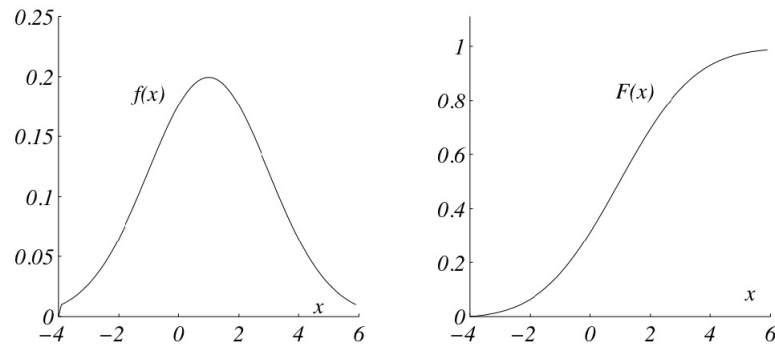


Figure: Densité et fonction de répartition d'une variable aléatoire Gaussienne (ou Normale) avec $\mu = 1$ et $\sigma = 2$

15

Exemple 4 : La loi Gaussienne (ou Normale)

Méthode de **Box-Miller** : Les deux variables aléatoires ci-dessous sont iid et normalement distribuées :

$$X_1 = \sqrt{-2 \ln(U_1)} \sin(2\pi U_2) \sim \mathcal{N}(0, 1)$$

$$X_2 = \sqrt{-2 \ln(U_1)} \cos(2\pi U_2) \sim \mathcal{N}(0, 1)$$

avec U_1, U_2 générés dans $[0, 1]$.

Pour obtenir une distribution de $Y \sim \mathcal{N}(\mu, \sigma^2)$ il suffit de savoir que $Y = \mu + \sigma X$ avec $X \sim \mathcal{N}(0, 1)$.

16

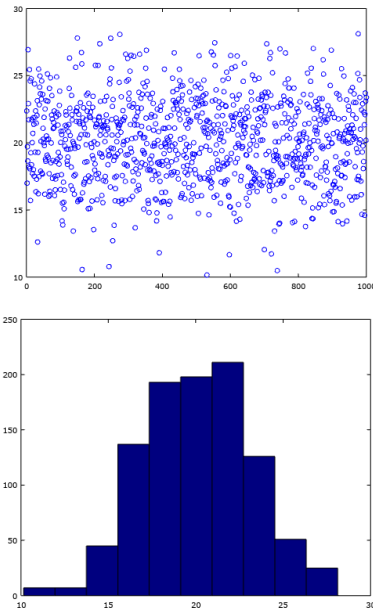
Exemple 4 : Simulation avec Octave

**N = 1000 valeurs
générées**
 $\mu = 20.0$ $\sigma = 3.0$

```
EDU>>
u1=rand(1000,1);
EDU>>
u2=rand(1000,1);
EDU>>
x1=sqrt(-2*log(u1)).*
sin(2*pi*u2);

EDU>> x= 20.0 +
3.0*x1;

EDU>> plot(x,'o')
EDU>> hist(x)
```



17

Exemple 4 : Simulation avec Octave (2)

N = 1000 valeurs générées
 $\mu = 20.0$ $\sigma = 3.0$

Moyenne

```
EDU>> S=0
EDU>> for t= 1:1000
S = S + x(t)
end
EDU>> m = S / 1000
m = 20.113
```

Variance

```
EDU>> S=0
EDU>> for t= 1:2000
S = S + ( x(t) - m) ^ 2
end
EDU>> v = S / 1000
v = 8.7450
EDU>> s = sqrt(v)
s = 2.9572
```

Test

```
EDU>> m2 = mean(x)
m2 = 20.113
EDU>> v2 = var(x)
v2 = 8.7538
EDU>> v2 = var(x,1)
v2 = 8.7450
```

18

Le Bootstrap

- Estimation de la variance et/ou intervalle de confiance à travers d'une simulation
- Nous pouvons estimer la variance, mais sa conversion en intervalles de confiance n'est pas facile que dans le cas d'une estimation de la moyenne et/ou une distribution normale
- [https://fr.wikipedia.org/wiki/Bootstrap_\(statistiques\)](https://fr.wikipedia.org/wiki/Bootstrap_(statistiques))
- Certains éléments de l'échantillon sont modifiés, en supprimant quelques observations et en dupliquant certaines autres
- La taille du nouveau échantillon reste la même
- On recalcule le paramètre qui nous intéresse
- Après un nombre d'itérations, on peut estimer la variance et IC

19

Exemple 5 : Bootstrap – loi Normale

N = 100
 $\mu = 20.0$ $\sigma = 3.0$
 itérations: 200

```
EDU>> x=normrnd
(20,3,100,1);
```

```
EDU>> plot(x,'o');
```

```
EDU>> mean(x)
ans = 20.356
```

```
EDU>> var(x)
ans = 10.717
```

90% IC théoriques pour
 normale/moyenne
 20.356-/
 $+1.695 \cdot \sqrt{10.717/100} =$
 [19.801 , 20.911]

Estimation bootstrap:

```
EDU>> for t=1:200
for s=1:100
y(s,t) = x(randperm(100,1));
end;
end;
```

```
EDU>> z = mean(y);
```

```
EDU>> z = z' ;
```

```
EDU>> z = sort(z)
```

```
EDU>> z(10)
```

```
ans= 19.799
```

```
EDU>> z(190)
```

```
ans = 20.853
```

90% intervalle de confiance
 pour la moyenne: [19.799 ,
 20.853]

20

Exemple 6 : Bootstrap – loi Poisson

```

N = 60
λ = 5.0      itérations:
500

EDU>> x=poissrnd (5,60,1);

EDU>> hist(x)

EDU>> mean(x)
ans = 5.100

EDU>> var(x)
ans = 6.2949

90% IC théoriques pour
normale/moyenne
5.1-/+1.695*sqrt(6.2949/60)
= [4.551 , 5.649]

```

```

EDU>> for t=1:500
for s=1:60
y(s,t) = x(randperm(60,1));
end;
end;

EDU>> z = mean(y);
EDU>> z = z' ;
EDU>> z = sort(z)
EDU>> z(25)
ans= 4.600
EDU>> z(475)
ans = 5.6167

90% intervalle de confiance
pour la moyenne:  [ 4.600 ,
5.617]

```

21

Exemple 6 : Bootstrap – loi Poisson (2)

Estimation de l' intervalle
de confiance pour la
médiane:

```

EDU>> z = median(y);
EDU>> z = z' ;
EDU>> z = sort(z)
EDU>> z(25)
ans= 4
EDU>> z(475)
ans = 5.5

```

90% intervalle de confiance
pour la médiane: [4 ,
5.5]

Pas de formule exacte
possible!

22