

Variables aléatoires continues

Rappel

La fonction de répartition

Soit $f(x)$ la fonction de densité, alors **la fonction de répartition** est donnée par :

$$F_X(x) = P(X \leq x) = \int_{-\infty}^x f(t)dt$$

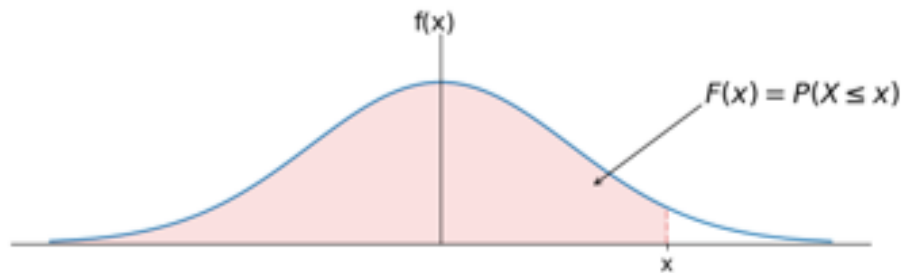


Figure 1: Fig. Probabilité $P(X \leq x)$

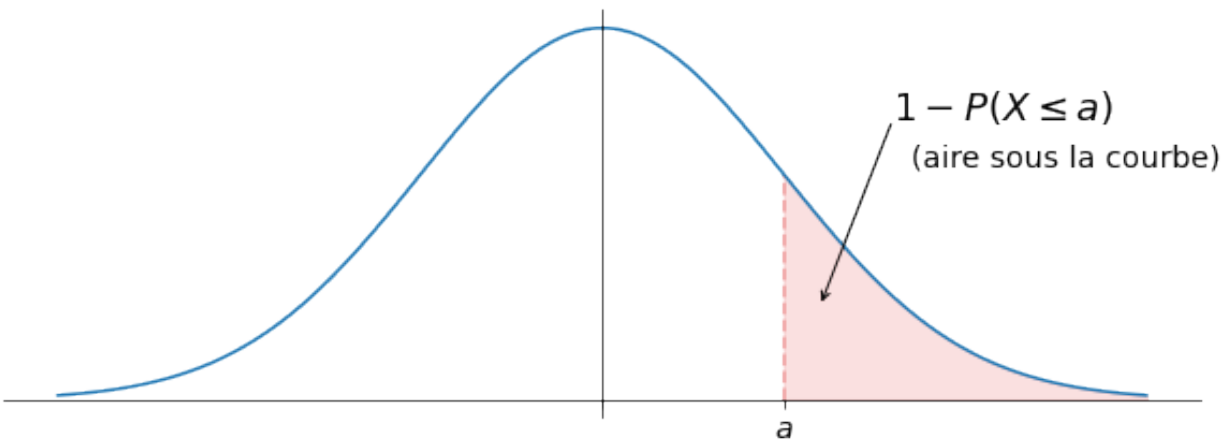


Figure 2: Fig. Probabilité $P(X > a)$

Espérance

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} xf(x)dx$$

- $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$
- $\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$

Variance

$$\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}X)^2] = \mathbb{E}[X^2] - (\mathbb{E}X)^2 = \int_a^b (x - \mathbb{E}X)^2 f(x) dx$$

- $\text{Var}(X + Y) = \text{Var}[X] + \text{Var}[Y]$, (si X et Y indép.)
- $\text{Var}(aX + b) = a^2 \text{Var}(X)$

Normalisation Soit $X \sim \mathcal{N}(m, \sigma^2)$, introduisons une v.a.

$$Y = \frac{X - m}{\sigma} \sim \mathcal{N}(0, 1)$$

Les valeurs de la fonction de répartition de $Y \sim \mathcal{N}(0, 1)$ peuvent être retrouvées dans les tables.

Notons que :

$$\mathbb{P}(X \leq x) = \mathbb{P}\left(\frac{X - m}{\sigma} \leq \frac{x - m}{\sigma}\right) = \mathbb{P}(Y \leq \frac{x - m}{\sigma})$$

Les valeurs de la fonction de répartition de $Y \sim \mathcal{N}(0, 1)$ sont connues et données dans la table.

Integration

$$\int x^n dx = \frac{x^{n+1}}{n+1} + C$$

Exercice 1

La note obtenue par des étudiants à un examen est une v.a.r. normale $X \sim \mathcal{N}(7, 3^2)$.

1. Calculer le pourcentage d'individus ayant plus de 10, et la note en dessous de laquelle se trouvent 10% des étudiants.
2. Compte tenu de ces résultats, on décide de revaloriser l'ensemble des notes par une transformation linéaire $Z = aX + b$. Quelles valeurs doit-on donner à a et b pour que les valeurs précédentes passent respectivement à 50% et 7 ? (*Indication : calculer $\mathbb{E}(Z)$ et $\text{Var}(Z)$ en fonction de $\mathbb{E}(X)$ et $\text{Var}(X)$*).

Solution

Lorsque $X \sim \mathcal{N}(7, 3^2)$, alors $m = 7$ et $\sigma^2 = 3^2$ ($\sigma = 3$), où m et σ^2 sont des paramètres de la loi normale $\mathcal{N}(m, \sigma^2)$.

Introduisons une v.a. $Y \sim \mathcal{N}(0, 1)$ en normalisant X :

$$Y = \frac{X - m}{\sigma} = \frac{X - 7}{3} \sim \mathcal{N}(0, 1)$$

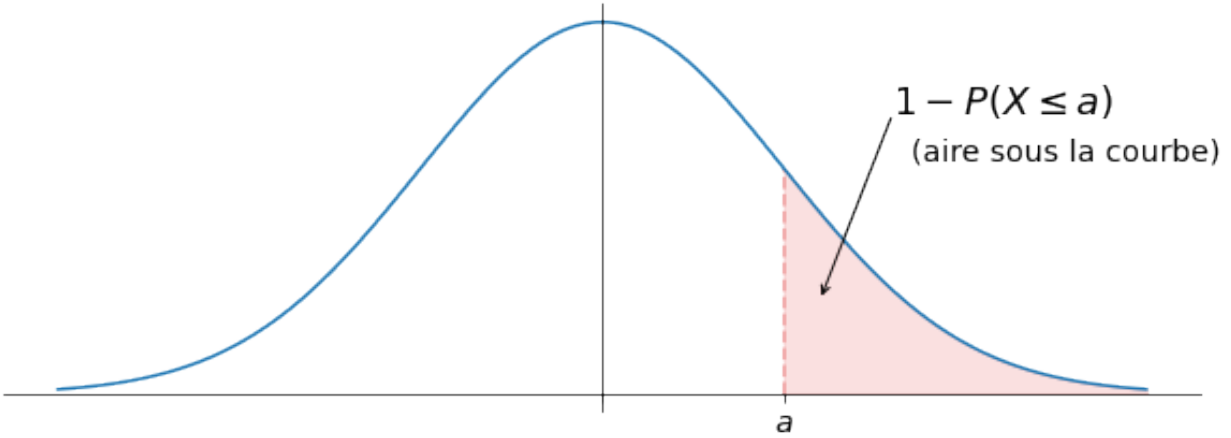
1. Calculer le pourcentage d'individus ayant plus de 10, et la note en dessous de laquelle se trouvent 10% des étudiants.

(a) le pourcentage d'individus ayant plus de 10 On cherche $P(X > 10)$, ce qui est équivalent de chercher :

$$P(X > 10) = P\left(\frac{X - 7}{3} > \frac{10 - 7}{3}\right) = \left[\text{par contruction } Y = \frac{X - 7}{3}\right] = P\left(Y > \frac{10 - 7}{3}\right) = P(Y > 1)$$

.

Pour rappel :



Notons que $P(Y > 1) = 1 - P(Y \leq 1)$

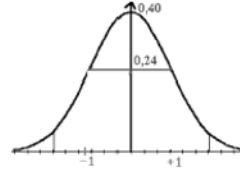
Cherchons la valeur de $P(Y \leq 1)$ dans la table :

Table de la fonction de répartition de la loi normale

Si X est une variable aléatoire suivant une loi normale centrée réduite $\mathcal{N}(0, 1)$, la table donne la valeur de la fonction de répartition de X en x , $F(x) = \mathbb{P}(X \leq x)$.

La valeur de x s'obtient par addition des nombres inscrits en marge.

Pour $x < 0$, on a $F(x) = 1 - F(|x|)$.



x	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7703	.7734	.7764	.7793	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990

Figure 3: Fig. Recherche de la valeur de $P(X \leq x)$ dans la table

Ainsi, $P(Y > 1) = 1 - P(Y \leq 1) = 1 - 0.8413 = 0.1587$

Autrement dit, le pourcentage d'individus ayant plus de 10 est 15.87%.

(b) la note en dessous de laquelle se trouvent 10% des étudiants Nous cherchons x tel que $P(X \leq x) = 0.1$.

On va d'abord chercher y tel que $P(Y \leq y) = 0.1$. On va se servir de la table des valeurs de la fonction de répartition. Notons que 0.1 est une valeur suffisamment petite (les valeurs dans la table commencent avec 0.5). C'est-à-dire que la valeur de y avec cette probabilité est négative $y < 0$.

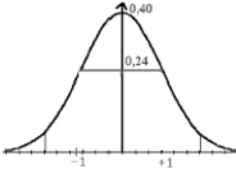
Etant donné la symétrie de la distribution, pour tout $y < 0$, $F_Y(y) = 1 - F_Y(|y|)$. Donc on va chercher dans la table la valeur à la probabilité $1 - 0.1 = 0.9$:

Table de la fonction de répartition de la loi normale

Si X est une variable aléatoire suivant une loi normale centrée réduite $N(0, 1)$, la table donne la valeur de la fonction de répartition de X en x , $F(x) = P(X \leq x)$.

La valeur de x s'obtient par addition des nombres inscrits en marge.

Pour $x < 0$, on a $F(x) = 1 - F(|x|)$.



x	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7703	.7734	.7764	.7793	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990

C'est la valeur $|y| = 1.28$, donc $y = -1.28$.

Maintenant, repassons à x :

$$x = y\sigma + m = -1.28 \times 3 + 7 = -3.84 + 7 = 3.16$$

2. Quelles valeurs doit-on donner à a et b dans $Z = aX + b$ pour que les valeurs précédentes passent respectivement à 50% et 7 ?

$$Z = aX + b \sim \mathcal{N}(m_z, \sigma_z^2).$$

Selon l'énoncé :

1. $P(Z > 10) = 0.5$ (le pourcentage d'individus ayant plus de 10 est 50%)
2. $P(Z \leq 7) = 0.1$ (la note en dessous de laquelle se trouvent 10% des étudiants est égale à 7)

En suivant la même procédure, nous pouvons passer à une v.a.

$$U = \frac{Z - m_Z}{\sigma_Z} \sim \mathcal{N}(0, 1)$$

Notons que les paramètres m_Z et σ_Z sont inconnus.

Reprenons la première condition $P(Z > 10) = 0.5$:

$$P(Z > 10) = P\left(\frac{Z - m_Z}{\sigma_Z} > \frac{10 - m_Z}{\sigma_Z}\right) = P\left(U > \frac{10 - m_Z}{\sigma_Z}\right) = [\text{soit } u = \frac{10 - m_Z}{\sigma_Z}] = P(U > u) = 1 - P(U \leq u) = 0.5$$

Trouvons u tel que $P(U \leq u) = 0.5$:

Donc, $u = 0$, autrement dit $u = \frac{10 - m_Z}{\sigma_Z} = 0$. Si $\sigma_Z \neq 0$, alors $10 - m_Z = 0 \Rightarrow m_Z = 10 = \mathbb{E}[Z]$.

De l'autre côté il est possible d'exprimer $\mathbb{E}[Z]$ via $\mathbb{E}[X]$ en se servant des propriétés de l'espérance :

$$m_z = \mathbb{E}[Z] = \mathbb{E}[aX + b] = [\text{par propr. de } \mathbb{E}X] = a\mathbb{E}[X] + b = 7a + b = \mathbf{10}$$

Utilisons maintenant la deuxième condition (la note en dessous de laquelle se trouvent 10% des étudiants est 7).

$$P(Z \leq 7) = P\left(\frac{Z - m_Z}{\sigma_Z} \leq \frac{7 - m_Z}{\sigma_Z}\right) = P\left(U \leq \frac{7 - m_Z}{\sigma_Z}\right) = [\text{notons } u = \frac{7 - m_Z}{\sigma_Z}] = P(U \leq u) = 0.1$$

D'après la recherche de la valeur dans la Q1, nous savons que $u = -1.28$. Donc

$$u = \frac{z - m_z}{\sigma_z} = \frac{7 - 10}{\sigma_z} = \frac{-3}{\sigma_z} = -1.28$$

$$\text{Donc, } \sigma_z = \frac{-3}{-1.28} = 2.34375$$

$$\text{La variance est alors : } \text{Var}[Z] = \sigma_z^2 = 2.34375^2 = 5.493164$$

De l'autre côté :

$$\text{Var}[Z] = \text{Var}[aX + b] = [\text{par propr. de } \text{Var}X] = a^2 \text{Var}[X] = 9a^2 = (3a)^2 = \mathbf{2.34375^2}$$

Donc,

$$\begin{cases} a \times 7 + b = 10 \\ (3a)^2 = 2.34375^2 \end{cases}$$

A partir de la 2ème équation nous obtenons :

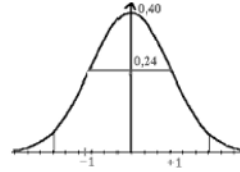
$$a = \frac{2.34375}{3} = 0.78125$$

Table de la fonction de répartition de la loi normale

Si X est une variable aléatoire suivant une loi normale centrée réduite $\mathcal{N}(0, 1)$, la table donne la valeur de la fonction de répartition de X en x , $F(x) = \mathbb{P}(X \leq x)$.

La valeur de x s'obtient par addition des nombres inscrits en marge.

Pour $x < 0$, on a $F(x) = 1 - F(|x|)$.



x	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7703	.7734	.7764	.7793	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990

Figure 4: Fig. Recherche de u

Alors,

$$7 \times 0.78125 + b = 10$$

$$b = 10 - 7 \times 0.78125 = 10 - 5.46875 = 4.53125$$

Réponse : $a = 0.78125$, $b = 4.53125$

Exercice 2 (DS 2018)

On souhaite modéliser les revenus salariés dans les pays européens. Soit X le salaire des individus et r le revenu minimum fixé par le pays. On suppose que X suit la loi de Pareto $\mathcal{P}(\alpha, r)$, de densité

$$f(x) = \begin{cases} \alpha r^\alpha x^{-(\alpha+1)} = \frac{\alpha r^\alpha}{x^{(\alpha+1)}} & \text{si } x > r \\ 0 & \text{sinon} \end{cases}$$

1. Quelle est l'espérance de X ?
2. Donner la fonction de répartition de X , notée F_X . Montrer que $U = (\frac{r}{X})^\alpha$ suit une loi uniforme.
3. D_9 correspond au revenu minimal des 10% les plus riches et D_1 correspond au revenu maximal des 10% les plus pauvres. Ces quantités sont données par $F_X(D_9) = 0.9$ et $F_X(D_1) = 0.1$.

En 2014, le rapport D_9/D_1 observé vaut 2.81 en France et 3.56 au Royaume-Uni (données OCDE sur les gains bruts, https://www.oecd-ilibrary.org/employment/data/gains/rapport-inter-decile-des-gains-bruts_data-00302-fr).

Calculer le rapport D_9/D_1 pour $r = 1$, en fonction de α . Pour quelle valeur de α a-t-on $D_9/D_1 = 2.8$? Pour quelle valeur de α a-t-on $D_9/D_1 = 3.56$? Comment interpréter le paramètre α ?

Solution

1. Quelle est l'espérance de X ?

Considérons le cas où $x > r$:

$$\begin{aligned} \mathbb{E}X &= \int_{-\infty}^{+\infty} x f(x) dx = \int_r^{+\infty} x \alpha r^\alpha x^{-(\alpha+1)} dx = \int_r^{+\infty} \alpha r^\alpha x^{-\alpha} dx = \alpha r^\alpha \int_r^{+\infty} x^{-\alpha} dx = \\ &= \alpha r^\alpha \left[\frac{1}{-\alpha+1} x^{-\alpha+1} \right]_r^{+\infty} = \frac{\alpha r^\alpha}{1-\alpha} [x^{1-\alpha}]_r^{+\infty} = \end{aligned}$$

Notons que si $1-\alpha < 0$, alors $x^{1-\alpha} \xrightarrow{x \rightarrow \infty} 0$ car pour $a > 0$, $x^{-a} = \frac{1}{x^a}$.

$$= \begin{cases} 0 - \frac{\alpha r^\alpha}{1-\alpha} r^{1-\alpha} & \text{si } \alpha > 1 \\ \infty & \text{si } \alpha \leq 1 \end{cases} = \begin{cases} \frac{\alpha r}{\alpha-1} & \text{si } \alpha > 1 \\ \infty & \text{si } \alpha \leq 1 \end{cases}$$

2. Donner la fonction de répartition de X , notée F_X . Montrer que $U = (\frac{r}{X})^\alpha$ suit une loi uniforme.

(a) La fonction de répartition Pour $x > r$:

$$\begin{aligned} F_X(x) &= P(X \leq x) = \int_{-\infty}^x f(t) dt = \int_r^x \alpha r^\alpha t^{-(\alpha+1)} dt = \alpha r^\alpha \int_r^x t^{-(\alpha+1)} dt = \\ &= \alpha r^\alpha \left[\frac{1}{-(\alpha+1)+1} t^{-(\alpha+1)+1} \right]_r^x = \alpha r^\alpha \left[\frac{1}{-\alpha} t^{-\alpha} \right]_r^x = -\frac{\alpha r^\alpha}{\alpha} [t^{-\alpha}]_r^x = -r^\alpha [t^{-\alpha}]_r^x = -\left[\frac{r^\alpha}{t^\alpha} \right]_r^x = \end{aligned}$$

$$= - \left[\left(\frac{r}{t} \right)^\alpha \right]_r^x = - \left(\frac{r}{x} \right)^\alpha - (-1) = 1 - \left(\frac{r}{x} \right)^\alpha$$

Pour $x < r$: $F_X(x) = 0$

En résumant :

$$F_X(x) = P(X \leq x) = \begin{cases} 1 - \left(\frac{r}{x} \right)^\alpha & \text{si } x > r \\ 0 & \text{si } x < r \end{cases}$$

(b) Montrer que $U = \left(\frac{r}{X} \right)^\alpha$ suit une loi uniforme $U = \left(\frac{r}{X} \right)^\alpha$

Remarquons que lorsque $P(X \leq x) = 1 - \left(\frac{r}{x} \right)^\alpha$,

$$P(X > x) = \left(\frac{r}{x} \right)^\alpha$$

Pour $x > r$:

$$\begin{aligned} F_U(u) &= P(U \leq u) = P\left(\left(\frac{r}{X}\right)^\alpha \leq u\right) = P\left(\frac{r}{X} \leq u^{1/\alpha}\right) = P\left(\frac{r}{u^{1/\alpha}} \leq X\right) = P\left(X \geq \frac{r}{u^{1/\alpha}}\right) = \\ &= \left[x = \frac{r}{u^{1/\alpha}} \text{ dans } P(X > x) \right] = \left(\frac{r}{\frac{r}{u^{1/\alpha}}} \right)^\alpha = \left(\frac{ru^{1/\alpha}}{r} \right)^\alpha = \left(u^{1/\alpha} \right)^\alpha = u = \mathcal{U}([0, 1]) \text{ pour } u \in [0, 1] \end{aligned}$$

Pour rappel : la fonction de répartition de la loi uniforme $\mathcal{U}([a, b])$ est donnée par :

$$F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & x \in [a, b] \\ 1 & x > b \end{cases}$$

Donc, pour $U([0, 1])$ si $x \in [0, 1]$: $F(x) = \frac{x-0}{1-0} = x$.

3. Calculer le rapport D_9/D_1 pour $r = 1$, en fonction de α . Pour quelle valeur de α a-t-on $D_9/D_1 = 2.8$? Pour quelle valeur de α a-t-on $D_9/D_1 = 3.56$? Comment interpréter le paramètre α ?

Pour $r = 1$ et $x > r$:

$$F_X(x) = P(X \leq x) = 1 - \left(\frac{r}{x} \right)^\alpha = 1 - \left(\frac{1}{x} \right)^\alpha$$

D_9	D_1
$F_X(D_9) = 1 - \left(\frac{1}{D_9} \right)^\alpha = 0.9$	$F_X(D_1) = 1 - \left(\frac{1}{D_1} \right)^\alpha = 0.1$
$\left(\frac{1}{D_9} \right)^\alpha = 0.1$	$\left(\frac{1}{D_1} \right)^\alpha = 0.9$
$\frac{1}{D_9} = 0.1^{1/\alpha}$	$\frac{1}{D_1} = 0.9^{1/\alpha}$
$\frac{1}{0.1^{1/\alpha}} = D_9$	$\frac{1}{0.9^{1/\alpha}} = D_1$
$\left(\frac{1}{0.1} \right)^{1/\alpha} = D_9$	$\left(\frac{1}{0.9} \right)^{1/\alpha} = D_1$

Cas France

Si $D_9/D_1 = 2.81$, alors

$$\frac{\left(\frac{1}{0.1} \right)^{1/\alpha}}{\left(\frac{1}{0.9} \right)^{1/\alpha}} = 2.81$$

$$\left(\frac{\frac{1}{0.1}}{\frac{1}{0.9}}\right)^{1/\alpha} = 2.81$$

$$\left(\frac{0.9}{0.1}\right)^{1/\alpha} = 9^{1/\alpha} = 2.81$$

$$1/\alpha = \log_9 2.81 = 0.4702225 \approx 1/2$$

```
log(2.81, base=9)
```

```
## [1] 0.4702225
```

```
1/log(2.81, base=9)
```

```
## [1] 2.126653
```

Donc, $\alpha = 2.126653$

Cas Royaume-Uni

$$\left(\frac{0.9}{0.1}\right)^{1/\alpha} = 9^{1/\alpha} = 3.56$$

$$1/\alpha = \log_9 3.56 = 0.5778929$$

```
log(3.56, base=9)
```

```
## [1] 0.5778929
```

```
1/log(3.56, base=9)
```

```
## [1] 1.730424
```

Donc, $\alpha = 1.730424$

```
require("EnvStats")
```

```
## Loading required package: EnvStats
```

```
##
```

```
## Attaching package: 'EnvStats'
```

```
## The following objects are masked from 'package:stats':
```

```
##
```

```
##      predict, predict.lm
```

```
library(EnvStats)
```

```
alphaFr <- 1/log(2.81, base=9)
```

```
alphaUK <- 1/log(3.56, base=9)
```

```
xx <- seq(1.0,5,0.1)
```

```
xFr <- dpareto(xx, location=1, shape=alphaFr)
```

```
xUK <- dpareto(xx, location=1, shape=alphaUK)
```

```
x5 <- dpareto(xx, location=1, shape=5)
```

```
x1 <- dpareto(xx, location=1, shape=1)
```

```
x10 <- dpareto(xx, location=1, shape=10)
```

```
plot(xx, xFr, col="darkgreen",type="l", xlab="x", ylab="f(x)")
```

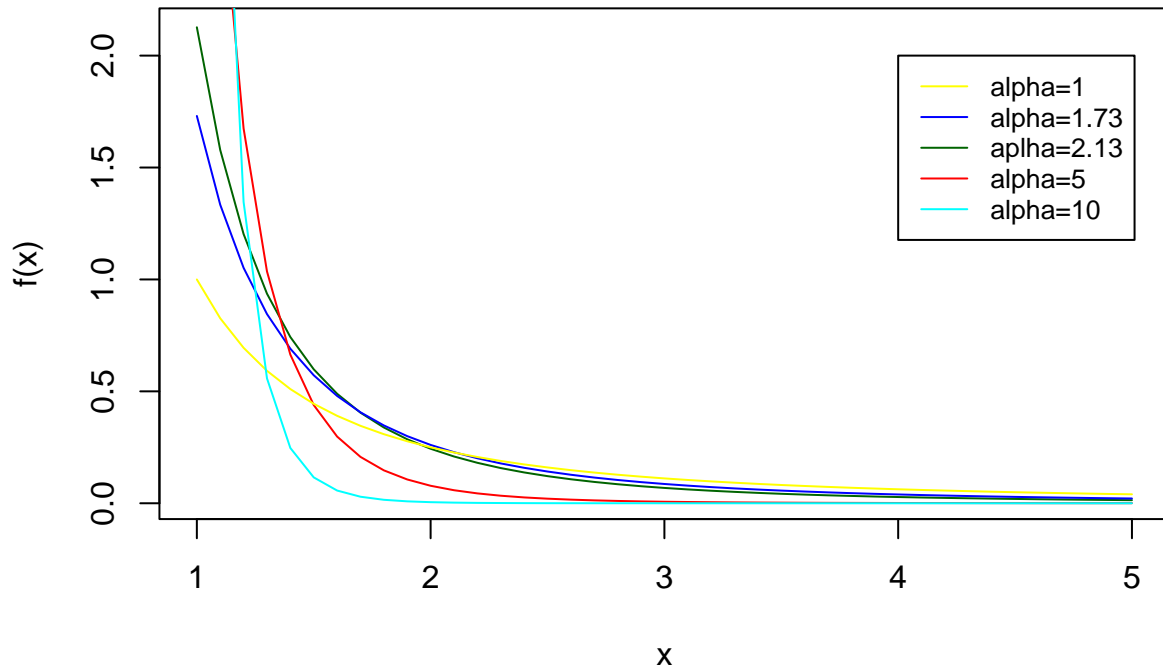
```
lines(xx, xUK, col="blue",type="l")
```

```
lines(xx, x5, col="red",type="l")
```

```

lines(xx, x1, col="yellow",type="l")
lines(xx, x10, col="cyan",type="l")
#abline(v=1)
# Add a legend
legend(4, 2, legend=c("alpha=1","alpha=1.73","alpha=2.13","alpha=5","alpha=10"),
      col=c("yellow", "blue", "darkgreen", "red", "cyan"), lty=1, cex=0.8)

```



α définit la forme (*shape*) de la courbe, la descente.

Exercice 3

Montrer que si $X \sim \mathcal{E}(\lambda)$, alors pour tous réels positifs a et b , on a

$$P(X > a + b \mid X > b) = P(X > a)$$

Si X représente par exemple la durée de vie d'une imprimante, que signifie cette propriété ?

Solution

Si $X \sim \mathcal{E}(\lambda)$, alors pour $x \geq 0$:

$$f(x) = \lambda e^{-\lambda x}$$

Sinon $f(x) = 0$.

Calculons la fonction de répartition pour $x \geq 0$:

$$\begin{aligned}
F_X(x) &= P(X \leq x) = \int_{-\infty}^x f(t)dt = \int_0^x \lambda e^{-\lambda t} dt = \lambda \int_0^x e^{-\lambda t} dt = \\
&= \lambda \left[e^{-\lambda t} \cdot \frac{-1}{\lambda} \right]_0^x = -\frac{\lambda}{\lambda} \left[e^{-\lambda t} \right]_0^x = -e^{-\lambda t} \Big|_0^x = -e^{-\lambda x} - (-1) = 1 - e^{-\lambda x}
\end{aligned}$$

Dans ce cas là $P(X > x) = 1 - P(X \leq x) = 1 - (1 - e^{-\lambda x}) = e^{-\lambda x}$ ou :

$$\begin{aligned}
P(X > x) &= \int_x^{+\infty} f(t)dt = \int_x^{+\infty} \lambda e^{-\lambda t} dt = \lambda \int_x^{\infty} e^{-\lambda t} dt = \\
&= \lambda e^{-\lambda t} \cdot \frac{-1}{\lambda} \Big|_x^{+\infty} = -e^{-\lambda t} \Big|_x^{+\infty} = 0 - (-e^{-\lambda x}) = e^{-\lambda x}
\end{aligned}$$

Nous nous intéressons à la probabilité conditionnelle $P(X > a + b \mid X > b)$. Pour rappel, la probabilité conditionnelle est donnée par :

$$P(A \cap B) = P(A) \cdot P(B|A)$$

Alors :

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

Dans notre cas $B = X > a + b$ et $A = X > b$:

$$P(X > a + b \mid X > b) = \frac{P(X > a + b \cap X > b)}{P(X > b)}$$

Lorsque $a \geq 0$ et $b \geq 0$, alors $a + b \geq a$ et $a + b \geq b$. Donc : $X > a + b \cap X > b = X > a + b$

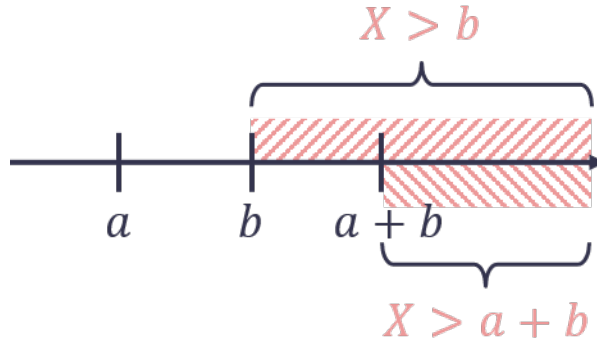


Figure 5: Fig. $X > a + b \cap X > b$

Alors :

$$P(X > a + b \mid X > b) = \frac{P(X > a + b \cap X > b)}{P(X > b)} = \frac{P(X > a + b)}{P(X > b)}$$

Notons que $P(X > x) = e^{-\lambda x}$

Alors :

$$P(X > b) = 1 - P(X \leq b) = e^{-\lambda b}$$

$$P(X > a + b) = 1 - P(X \leq a + b) = e^{-\lambda(a+b)} = e^{-\lambda a} e^{-\lambda b} = P(X > a) \cdot P(X > b)$$

Alors :

$$P(X > a + b | X > b) = \frac{P(X > a + b \cap X > b)}{P(X > b)} = \frac{P(X > a + b)}{P(X > b)} = \frac{P(X > a) \cdot P(X > b)}{P(X > b)} = P(X > a)$$

Cette propriété est appelée **absence de mémoire** ou **perte de mémoire** (*memorylessness*).

Si X représente la durée de vie d'une imprimante, elle peut être interprétée de la façon suivante : sachant que l'imprimante a déjà servi pendant b temps, la probabilité qu'elle dure au moins $a + b$ temps sera la même que la probabilité qu'elle dure a temps à partir de son installation. Autrement dit, le fait qu'elle ne soit pas tombée en panne pendant b temps ne change rien à son espérance de vie à partir du temps b .

Dans le cas discret, une distribution qui a également la propriété de la perte de mémoire est la distribution géométrique pour X défini sur $\{0, 1, 2, \dots\}$ et tout $n, m \in \{0, 1, 2, \dots\}$:

$$P(X > m + n | X \geq m) = P(X > n)$$

Annexe.

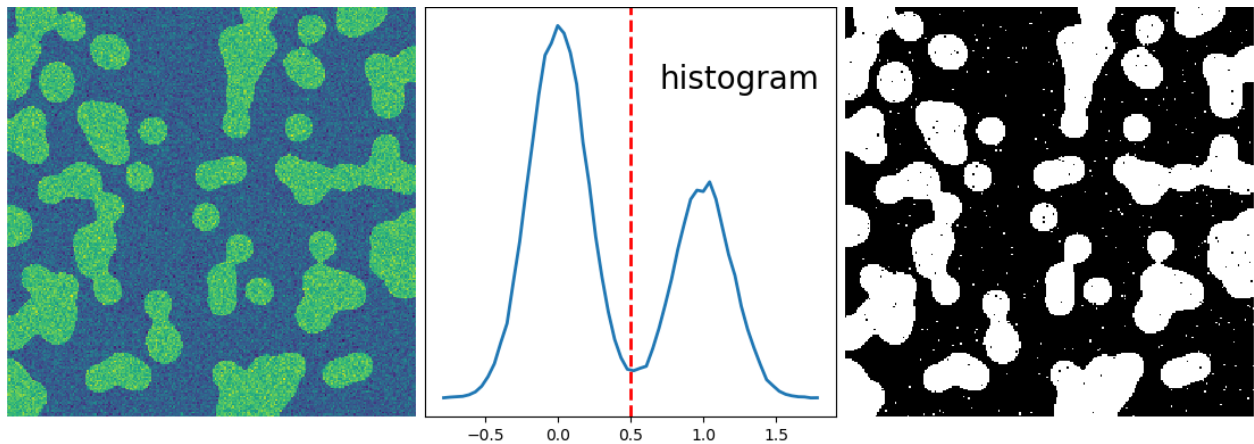


Figure 6: Source : http://scipy-lectures.org/advanced/image_processing/

Des procédures de segmentation reposent sur l'identification d'un mélange de deux lois de probabilités. Ainsi sur l'exemple ci-dessus, la densité sur l'image peut être vue comme une superposition de deux densités gaussiennes : identifier ces deux densités permet une segmentation de l'image.