7 | Variables aléatoires à densité

I – Rappels sur la fonction de répartition

Définition 7.1 – Soit X une variable aléatoire. On appelle **fonction de répartition** de la variable aléatoire X la fonction F_X définie par

$$\forall x \in \mathbb{R}, \quad F_X(x) = P(X \leqslant x).$$

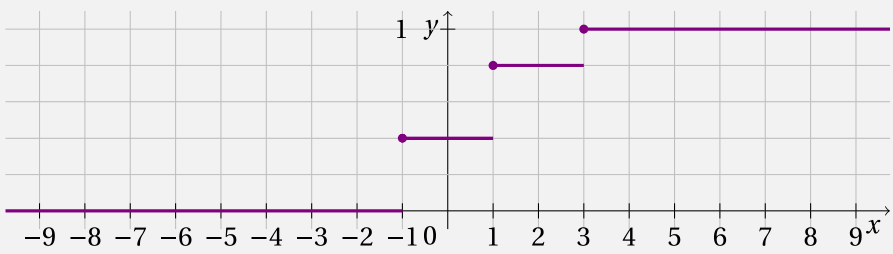
Proposition 7.2

Soit X une variable aléatoire discrète. On note le support $X(\Omega) = \{x_1, x_2, \dots\}$ avec $x_1 < x_2 < \dots$.
Alors

$$F_X(x) = \begin{cases} 0 & \text{si} \quad x < x_1, \\ P(X = x_1) + \dots + P(X = x_k) & \text{si } x_k \leqslant x < x_{k+1}, \\ 1 & \text{si } x \geqslant \max\limits_{i \in \mathbb{N}} x_i. \end{cases}$$

En particulier F_X est constante sur chaque intervalle $[x_k, x_{k+1}[$.

Exemple 7.3 – Un sac contient cinq jetons numérotés de 1 à 5. Pour jouer une partie, on doit miser 1€. On tire au hasard un jeton. Si on tire le numéro 1 on gagne 4€, si on tire un numéro pair on reçoit 2€ et rien sinon. On note X le gain algébrique.



II – Généralités

1 – Notion de variable aléatoire à densité

Définition 7.4 – Une fonction f définie sur $\mathbb{R}$ est une **densité de probabilité** (ou simplement **densité**) si et seulement si elle vérifie les trois conditions suivantes :

1. f est positive : pour tout $x \in \mathbb{R}$, $\quad f(x) \geqslant 0$.
2. f n'admet qu'un nombre **fini** de points de discontinuité.
3. L'intégrale impropre $\displaystyle\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t$ converge et vaut $\displaystyle\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = 1$.

Exemple 7.5 – On considère la fonction f définie sur $\mathbb{R}$ par $f(x) = \dfrac{e^x}{\left(1+e^x\right)^2}$.
Montrer que la fonction f est une densité de probabilité.

Définition 7.6 – Soient X une variable aléatoire et F_X sa fonction de répartition. On dit que X est une **variable aléatoire à densité** si et seulement s'il existe une densité de probabilité f telle que

$$\forall x \in \mathbb{R}, \quad F_X(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t.$$

Il est souvent nécessaire de distinguer plusieurs cas en fonction des différentes expressions de $f(x)$.

Exemple 7.7 – Soit f la fonction définie sur $\mathbb{R}$ par $f(t) = \begin{cases} \dfrac{2}{t^3} & \text{si } t \geqslant 1, \\ 0 & \text{sinon.} \end{cases}$.
On admet que f est une densité de probabilité et on note X une variable aléatoire admettant f pour densité. Déterminer alors la fonction de répartition F_X de X.

Théorème 7.8

Si X est une variable aléatoire à densité, de fonction de répartition F_X et de densité f, alors en chaque réel x où la fonction f est continue, $f(x) = F_X'(x)$.

Exemple 7.9 – Soit X une variable aléatoire de fonction de répartition $F_X(x) = \begin{cases} 0 & \text{si } x < 1, \\ 1 - \dfrac{1}{x} & \text{si } x \geqslant 1. \end{cases}$
On admet que X est une variable aléatoire à densité. Déterminer une densité de X.

Remarque 7.10 – Il n'y a pas unicité de la densité pour une variable aléatoire donnée : si f est une densité de X, alors toute fonction g positive égale à f sauf en un nombre fini de points est aussi une densité de X.

2 – Calculs de probabilités avec des variables aléatoires à densité

Proposition 7.11

Soit X une variable aléatoire à densité, de fonction de répartition F_X et de densité f.
Soient a et b deux réels tels que $a < b$. On rappelle que $P(X \leqslant a) = F_X(a) = \int_{-\infty}^{a} f(t)\,\mathrm{d}t$.

- $P(X \leqslant a) = F_X(a) = \int_{-\infty}^{a} f(t)\,\mathrm{d}t$,
- $P(X \geqslant a) = 1 - F_X(a) = \int_{a}^{+\infty} f(t)\,\mathrm{d}t$,
- $P(X = a) = 0$,
- $P(a \leqslant X \leqslant b) = F_X(b) - F_X(a) = \int_{a}^{b} f(t)\,\mathrm{d}t$.

Si on connaît la fonction de répartition de X, alors on utilise les formules mettant en jeu F_X (les intégrales sont déjà calculées). Sinon, on utilise les formules avec les intégrales mettant en jeu la densité f.

Exemple 7.12 –

1. Soit X une variable aléatoire à densité, de fonction de répartition $F_X(x) = \begin{cases} 0 & \text{si } x < 2, \\ 1 - \dfrac{8}{x^3} & \text{si } x \geqslant 2. \end{cases}$
   Calculer $P(X \geqslant 0)$, $P(-1 \leqslant X < 3)$ et $P(X < 4)$.

2. Soit X une variable aléatoire à densité, de densité $f(t) = \begin{cases} 0 & \text{si } t \leqslant 0, \\ e^{-t} & \text{si } t > 0. \end{cases}$
   Calculer $P(X \leqslant 2)$, $P(2 < X \leqslant 3)$ et $P(X \geqslant 1)$.

Remarque 7.13 – Pour une variable aléatoire à densité, les probabilités ponctuelles $P(X = a)$ sont nulles. Ainsi les symboles $\leqslant$ (ou $\geqslant$) peuvent être remplacés par $<$ (ou $>$) sans modifier les formules ci-dessus.

3 – Espérance d'une variable à densité

Définition 7.14 – Soient X une variable aléatoire et f sa densité.
Alors X admet une espérance si et seulement si l'intégrale impropre $\displaystyle\int_{-\infty}^{+\infty} t f(t)\,\mathrm{d}t$ converge.
Dans ce cas, l'espérance de X est le réel noté $E(X)$ défini par

$$E(X) = \int_{-\infty}^{+\infty} t f(t)\,\mathrm{d}t.$$

Exemple 7.15 – Soit X une variable aléatoire de densité $f(t) = \begin{cases} \dfrac{3}{t^4} & \text{si } t \geqslant 1, \\ 0 & \text{si } t < 1. \end{cases}$
X admet-elle une espérance ? Si oui, la calculer.

4 – Variance d'une variable aléatoire à densité

Définition 7.16 – Soient X une variable aléatoire et f sa densité. Sous réserve d'existence, la variance de X est le réel noté $V(X)$ défini par $\displaystyle V(X) = \int_{-\infty}^{+\infty} \big(t - E(X)\big)^2 f(t)\,\mathrm{d}t.$

En pratique, X admet une variance si et seulement si l'intégrale impropre $\displaystyle\int_{-\infty}^{+\infty} t^2 f(t)\,\mathrm{d}t$ converge.

Dans ce cas, l'intégrale précédente est l'espérance de X^2, *i.e.* $\displaystyle E\big(X^2\big) = \int_{-\infty}^{+\infty} t^2 f(t)\,\mathrm{d}t,$
et la variance s'obtient grâce à la formule de König-Huygens :

$$V(X) = E\big(X^2\big) - E(X)^2.$$

Exemple 7.17 – On reprend la variable aléatoire X de l'exemple précédent, de densité

$$f(t) = \begin{cases} \dfrac{3}{t^4} & \text{si } t \geqslant 1, \\ 0 & \text{si } t < 1. \end{cases}$$

X admet-elle une variance ? Si oui, la calculer.

Proposition 7.18

Soient X une variable aléatoire à densité et a et b deux réels. Les résultats rencontrés dans les chapitres précédents se prolongent aussi dans ce cas, sous réserve que toutes les quantités impliquées existent :

- Par linéarité de l'espérance, $E(aX+b) = aE(X)+b$.
- La variance n'est pas linéaire, mais $V(aX+b) = a^2V(X)$
- L'écart-type d'une variable aléatoire X est défini comme la racine carrée de la variance de X.

Méthode 7.19 – Fonction de répartition d'une variable aléatoire de la forme $Y = aX+b$

Pour déterminer la fonction de répartition F_Y d'une variable aléatoire de la forme $Y = aX+b$, en connaissant la fonction de répartition de X :

1. On exprime la fonction de répartition F_Y de Y en fonction de celle F_X de X, en passant par les probabilités grâce à la formule $F_Y(y) = P(Y \leqslant y)$.
2. On utilise alors l'expression de $F_X(x)$ pour obtenir l'expression de $F_Y(y)$.

Exemple 7.20 – On reprend un des exemples précédents.

On considère la fonction f définie sur $\mathbb{R}$ par $f(t) = \begin{cases} \dfrac{2}{t^3} & \text{si } t \geqslant 1, \\ 0 & \text{sinon.} \end{cases}$

On admet que f est une densité de probabilité et on note X une variable aléatoire admettant f pour densité. On sait déjà que la fonction de répartition de X est donnée par

$$\forall x \in \mathbb{R}, \quad F_X(x) = \begin{cases} 1 - \dfrac{1}{x^2} & \text{si } x \geqslant 1, \\ 0 & \text{sinon.} \end{cases}$$

Soit $Y = 2X + 3$. Déterminer la fonction de répartition F_Y de Y.

III – Lois usuelles à densité

1 – Loi uniforme sur un intervalle

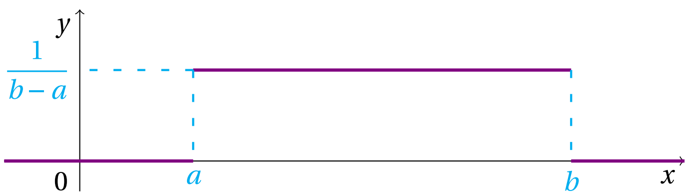
Dans ce paragraphe, a et b désignent deux réels tels que $a < b$.

Définition 7.21 – Une variable aléatoire X suit une **loi uniforme** sur $[a, b]$ lorsque X est la variable aléatoire de densité f définie par

$$f(t) = \begin{cases} \dfrac{1}{b-a} & \text{si } t \in [a, b], \\ 0 & \text{si } t \notin [a, b]. \end{cases}$$

On note $X \hookrightarrow \mathcal{U}\big([a, b]\big)$.

Voici la représentation graphique de la densité f d'une loi uniforme sur $[a, b]$:



Remarque 7.22 –

- La loi uniforme sur $[a, b]$ est la loi du tirage au hasard d'un nombre dans cet intervalle : la variable X a autant de chance de tomber n'importe où dans l'intervalle $[a, b]$.

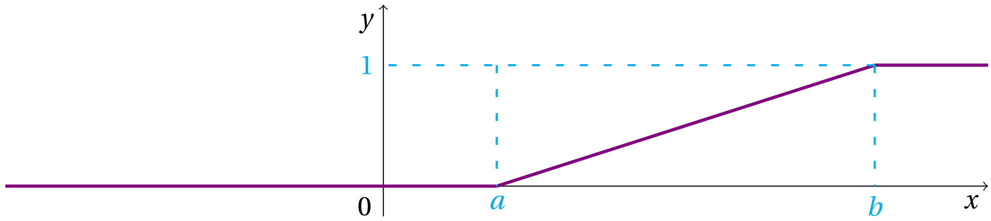
- La fonction f définie sur $[a,b]$ par $f(t) = \dfrac{1}{b-a}$ est bien une densité de probabilité sur l'intervalle $[a,b]$ puisque
  - ▷ f est positive et continue sur $\mathbb{R} \setminus \{a,b\}$,
  - ▷ $\displaystyle\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = \int_a^b \frac{1}{b-a}\,\mathrm{d}t = \left[\frac{t}{b-a}\right]_a^b = \frac{b}{b-a} - \frac{a}{b-a} = 1.$

Proposition 7.23

Si X suit une loi uniforme sur $[a,b]$, alors la fonction de répartition F_X de X est donnée par

$$F_X(x) = \begin{cases} 0 & \text{si } x < a, \\ \dfrac{x-a}{b-a} & \text{si } x \in [a,b], \\ 1 & \text{si } x > b. \end{cases}$$

Voici la représentation graphique de la fonction de répartition F_X d'une loi uniforme sur $[a,b]$:



Proposition 7.24

Si X suit une loi uniforme sur $[a,b]$, alors X admet une espérance et une variance et

$$E(X) = \frac{a+b}{2} \quad \text{et} \quad V(X) = \frac{(b-a)^2}{12}.$$

Exemple 7.25 – Le temps d'attente T, exprimé en minutes, auprès du standard téléphonique du service après-vente d'une entreprise suit une loi uniforme sur l'intervalle $[0.5, 9.5]$.

1. Quelle est la probabilité que le temps d'attente soit inférieur à deux minutes ?
2. Quelle est la probabilité que le temps d'attente soit supérieur à trois minutes ?
3. Quel est le temps d'attente moyen auprès du standard téléphonique ?

2 – Loi exponentielle

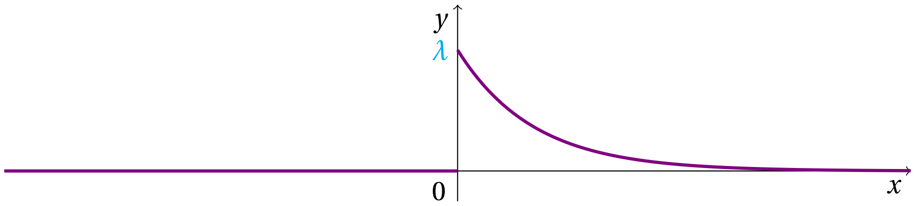
Dans ce paragraphe, λ désigne un réel strictement positif.

Définition 7.26 – Une variable aléatoire X suit une **loi exponentielle** de paramètre λ lorsque X est la variable aléatoire de densité f définie par

$$f(t) = \begin{cases} \lambda e^{-\lambda t} & \text{si } t \geqslant 0, \\ 0 & \text{si } t < 0. \end{cases}$$

On note $X \hookrightarrow \mathcal{E}(\lambda)$.

Voici la représentation graphique de la densité f d'une loi exponentielle de paramètre λ :



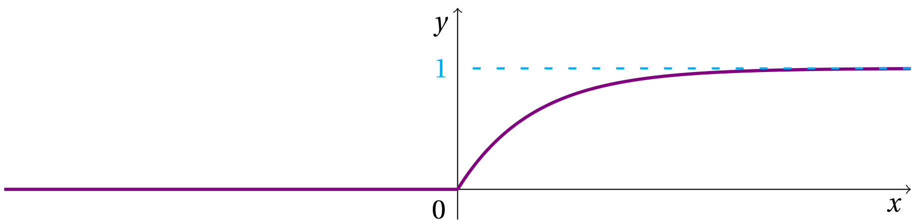
Remarque 7.27 – Les lois exponentielles sont utilisées pour modéliser des durées de vie.

Proposition 7.28

Si X suit une loi exponentielle de paramètre λ, alors la fonction de répartition F_X de X est donnée par

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x} & \text{si } x \geqslant 0, \\ 0 & \text{si } x < 0. \end{cases}$$

Voici la représentation graphique de la fonction de répartition d'une loi exponentielle de paramètre λ :



Exemple 7.29 – Soit X une variable aléatoire dont la fonction de répartition F_X est donnée par

$F_X(x) = \dfrac{e^x - 1}{e^x}$ si $x \geqslant 0$ et $F_X(x) = 0$ sinon. Retrouver la loi de X.

Proposition 7.30

Si X suit une loi exponentielle de paramètre λ, alors X admet une espérance et une variance et

$$E(X) = \frac{1}{\lambda} \quad \text{et} \quad V(X) = \frac{1}{\lambda^2}.$$

3 – Loi normale

Dans ce paragraphe, m désigne un réel et σ un réel strictement positif.

Définition 7.31 – Une variable aléatoire X suit une **loi normale** de paramètres m et σ^2 lorsque X est la variable aléatoire de densité f définie par

$$\forall t \in \mathbb{R}, \quad f(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-m)^2}{2\sigma^2}}.$$

On note $X \hookrightarrow \mathcal{N}\left(m, \sigma^2\right)$.

Proposition 7.32

Si X suit une loi normale de paramètres m et σ^2, alors X admet une espérance et une variance et

$$E(X) = m \quad \text{et} \quad V(X) = \sigma^2.$$

4 – Loi normale centrée réduite

Définition 7.33 – On appelle **loi normale centrée réduite** la **loi normale** de paramètres $m = 0$ et $\sigma^2 = 1^2 = 1$. On note $X \hookrightarrow \mathcal{N}(0,1)$.
Une variable aléatoire X suit la **loi normale centrée réduite** lorsque X est la variable aléatoire de densité f définie par

$$\forall t \in \mathbb{R}, \quad f(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}.$$

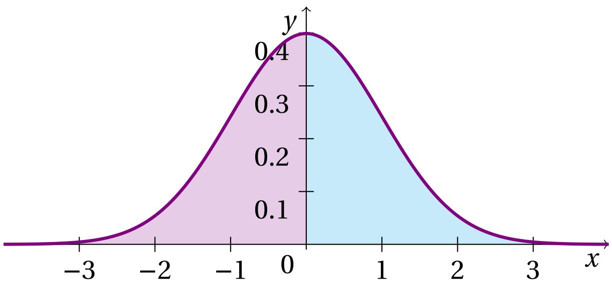
La fonction de répartition de X est la fonction notée Φ définie par

$$\forall x \in \mathbb{R}, \quad \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, \mathrm{d}t.$$

Remarque 7.34 – On ne sait pas expliciter Φ à l'aide des fonctions usuelles, mais on dispose d'un tableau de valeurs de $\Phi(x)$ pour différentes valeurs de $x \geqslant 0$.

La courbe de la fonction de densité de la loi normale centrée réduite est symétrique par rapport à l'axe des ordonnées, donc les probabilités $P(X \leqslant 0)$ et $P(X \geqslant 0)$ sont égales.
Comme $P(X \leqslant 0) + P(X \geqslant 0) = 1$, alors

$$P(X \leqslant 0) = P(X \geqslant 0) = \frac{1}{2}.$$



Théorème 7.35

On sait déjà que $\Phi(0) = \frac{1}{2}$. Plus généralement, pour tout réel x, la fonction Φ vérifie

$$\Phi(-x) = 1 - \Phi(x).$$

Théorème 7.36

Soit X une variable aléatoire. Alors

$$X \text{ suit une loi normale } \mathcal{N}(m,\sigma^2) \iff \frac{X-m}{\sigma} \text{ suit la loi normale centrée réduite } \mathcal{N}(0,1).$$

Méthode 7.37 – Calculer des probabilités pour une loi normale

Pour calculer les probabilités d'une loi normale, on utilise les formules donnant les probabilités à l'aide de la fonction de répartition :

$$P(X \leqslant a) = F_X(a), \qquad P(X \geqslant b) = 1 - F_X(b) \quad \text{et} \quad P(a \leqslant X \leqslant b) = F_X(b) - F_X(a).$$

Deux cas se présentent alors :

- Si X suit la loi normale centrée réduite, alors on utilise ces formules et le tableau des valeurs de $\Phi(x)$ pour $x \geqslant 0$, en se ramenant si besoin à x positif grâce à la formule $\Phi(-x) = 1 - \Phi(x)$.
- Si X suit une loi normale quelconque, avec $m \neq 0$ ou $\sigma \neq 1$, alors on se ramène au cas d'une variable aléatoire suivant la loi normale centrée réduite :

$$X \text{ suit la loi } \mathcal{N}(m,\sigma^2) \iff \frac{X-m}{\sigma} \text{ suit la loi } \mathcal{N}(0,1).$$

  On conclut alors à l'aide du premier point.

Exemple 7.38 –

1. Soit X une variable aléatoire qui suit la loi normale centrée réduite $\mathcal{N}(0,1)$.
   Calculer les probabilités $P(X \leqslant 0.65)$, $P(X > 0.23)$ et $P(-0.5 < X \leqslant 1.23)$.

2. Soit X une variable aléatoire qui suit la loi normale $\mathcal{N}(1,4)$.
   Calculer les probabilités $P(X < 3)$, $P(X \geqslant -1)$ et $P(0 \leqslant X < 5)$.

Fonction de répartition Φ d'une variable aléatoire X
suivant la loi normale centrée réduite $\mathcal{N}(0,1)$.

$$\Phi(x) = P(X \leqslant x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\,\mathrm{d}t \quad \text{et} \quad \Phi(-x) = 1 - \Phi(x).$$

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Pour trouver la valeur de $\Phi(x)$ dans le tableau, on repère les dixièmes de x en ligne et les centièmes en colonne.
Par exemple, la valeur de $\Phi(1.72)$ se lit sur la ligne débutant par 1.7 et la colonne 0.02. On trouve $\Phi(1.72) = 0.9573$.