Loi normale

I/ Propriétés

1) Théorème central-limite

La somme de plusieurs variables aléatoires indépendantes suit une loi qui tend vers une loi normale. Ses paramètres sont l'espérance et l'écart-type.

Il en résulte que la moyenne de nombreuses variables aléatoires de même loi mais indépendantes entre elles suit aussi une loi normale. Son espérance est celle des termes de la somme, mais son écart-type est le quotient de leur écart-type par la racine carrée du nombre de termes.

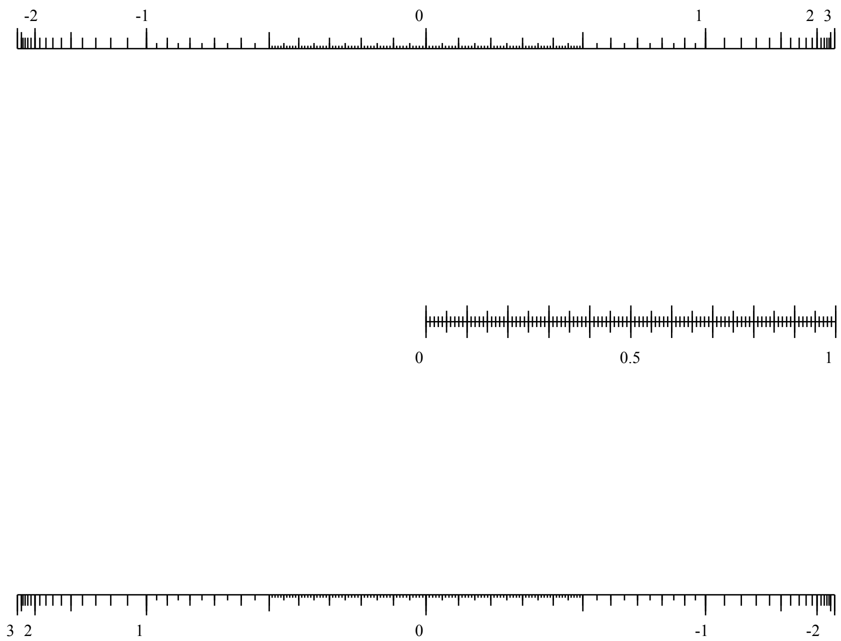
2) Somme de variables aléatoires normales

La somme (ou la moyenne) de variables aléatoires normales indépendantes entre elles suit une loi normale, même s'il y a peu de termes.

II/ Calcul de probabilités

1) Avec le nomogramme

Ce nomogramme permet de calculer la probabilité qu'une variable aléatoire normale de paramètres 0 et 1 soit entre deux valeurs :

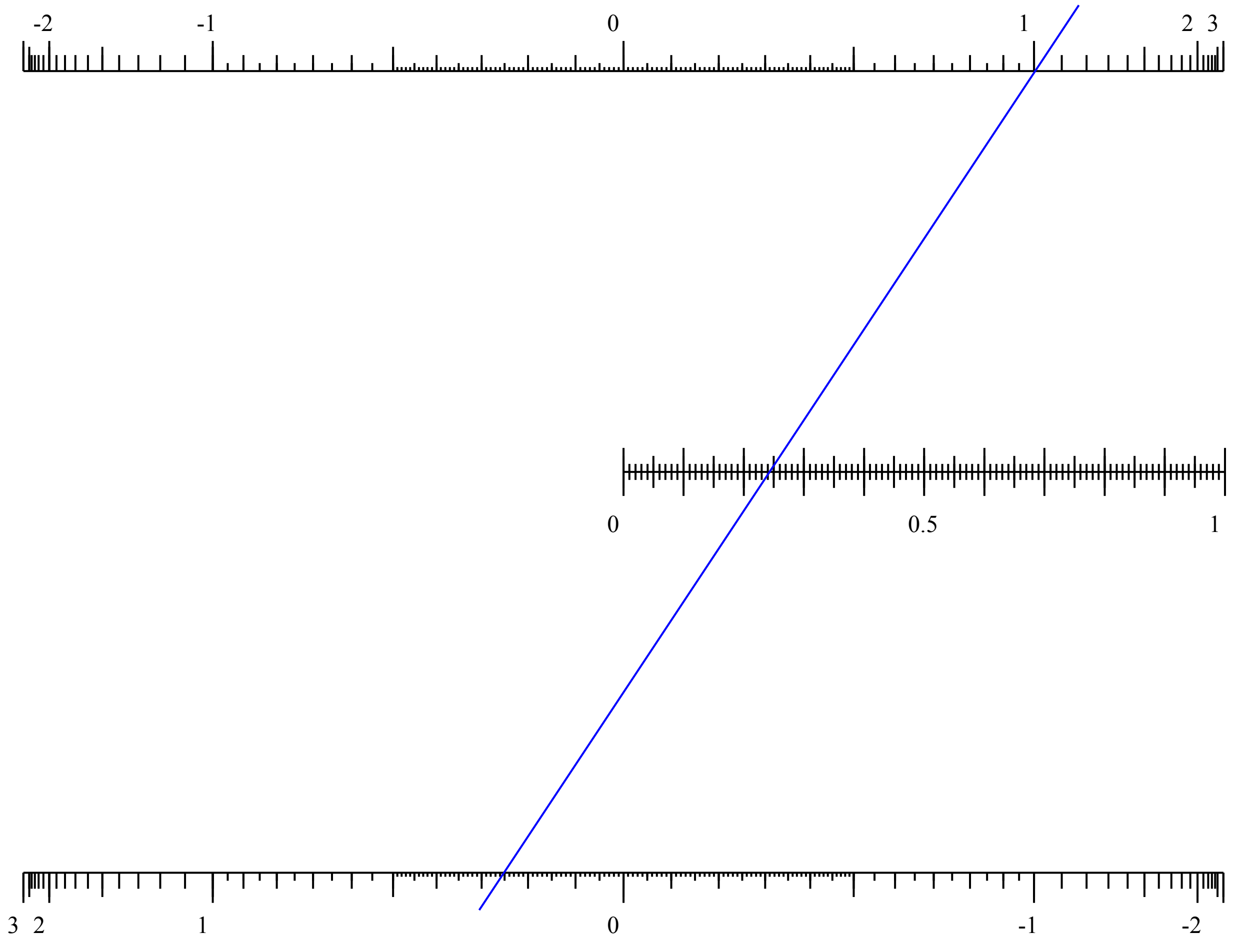


a) Exemple d'utilisation

On suppose que X suit une loi normale d'espérance 120 et d'écart-type 16. On souhaite connaître la probabilité que X soit compris entre 124 et 136. Pour cela on transforme l'inégalité $124 \leq X \leq 136$

- en centrant (soustraire 120 aux trois membres) : $4 \leq X-120 \leq 16$
- en réduisant (diviser les trois membres par 16) : $4/16 \leq (X-120)/16 \leq 16/16$
- soit $0{,}25 \leq Z \leq 1$ avec $Z=(X-120)/16$

On cherche donc nomographiquement la probabilité que Z soit compris entre 0,5 et 1, en posant une règle sur la graduation 0,25 en bas et la graduation 1 en haut :

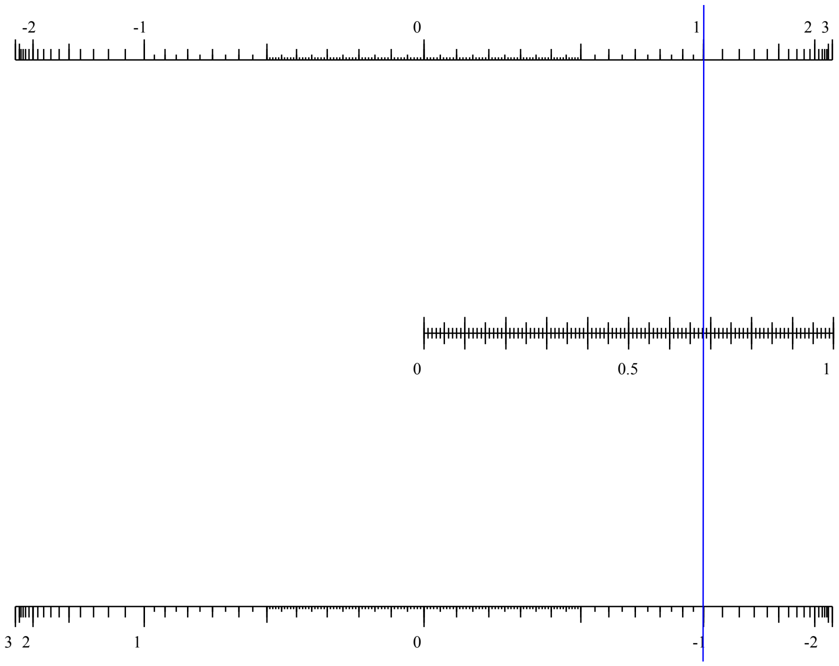


On lit la probabilité sur l'axe du milieu : environ 0,24.

Entre les deux axes du haut et du bas il y a une symétrie centrale, ce qui permet de lire assez facilement la probabilité que Z soit entre deux valeurs opposées.

b) À σ près

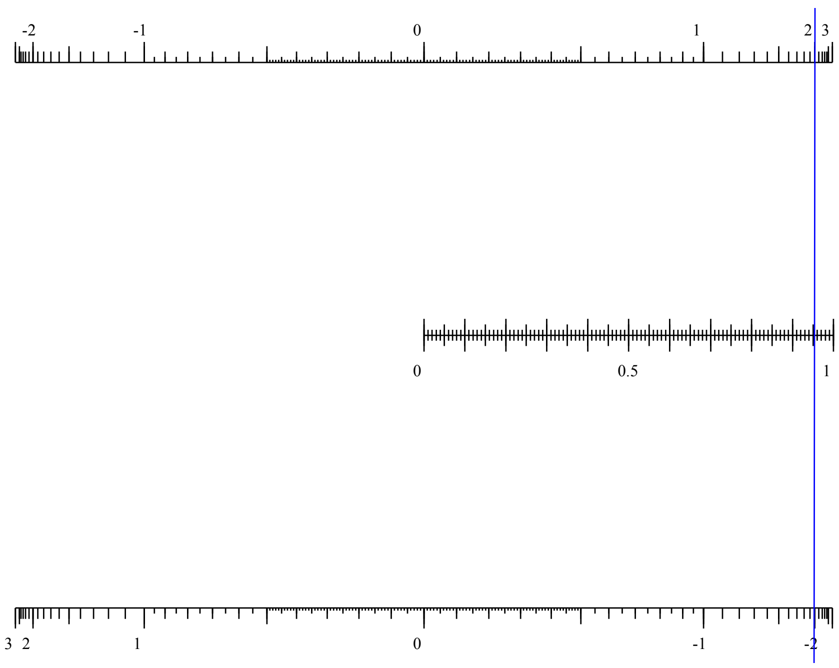
La probabilité que X (variable aléatoire normale d'espérance μ et d'écart-type σ) soit comprise entre $\mu-\sigma$ et $\mu+\sigma$ est égale à la probabilité que Z soit compris entre -1 et 1, que l'on lit sur le nomogramme :



On lit environ 0,68.

c) À 2×σ près

De même, la probabilité que X soit comprise entre $\mu-2\times\sigma$ et $\mu+2\times\sigma$ est égale à la probabilité que Z soit comprise entre -2 et 2. Celle-ci vaut environ 0,95 comme on le voit sur le nomogramme :



En fait, la probabilité que Z soit entre -2 et 2 est légèrement supérieure à 0,95. La probabilité que Z soit comprise entre -1,96 et 1,96 (ou, ce qui revient au même, que X soit entre $\mu-1{,}96\times\sigma$ et $\mu+1{,}96\times\sigma$) par contre est très proche de 0,95 (soit 19 chances sur 20).

d) À 3×σ près

La probabilité que Z soit entre -3 et 3 est très proche de 1 (de 0,9973 pour être plus précis) et la probabilité que Z soit entre -4 et 4 est encore plus proche de 1 : concrètement pour une variable aléatoire normale, $\mu+4\times\sigma=\infty$.

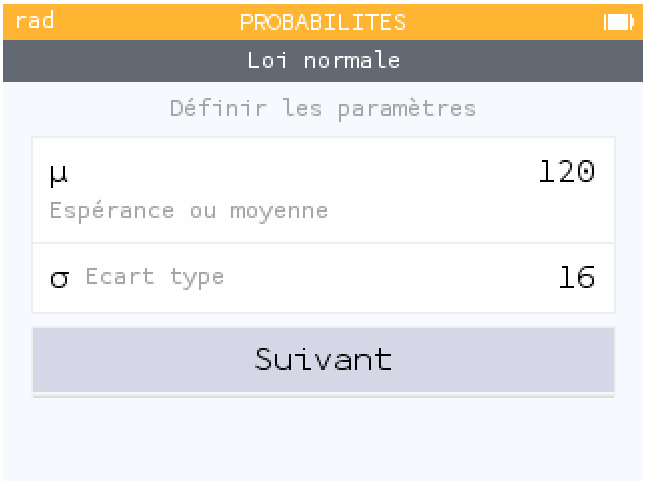
2) Avec la calculatrice

On rappelle que X suit une loi normale d'espérance 120 et d'écart-type 16. On souhaite connaître la probabilité que X soit compris entre 124 et 136.

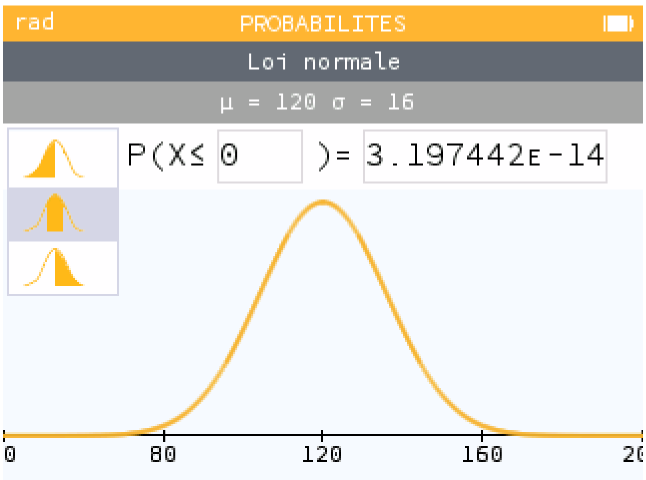
a) Numworks

On se rend dans l'application de probabilités, et on choisit la loi normale.

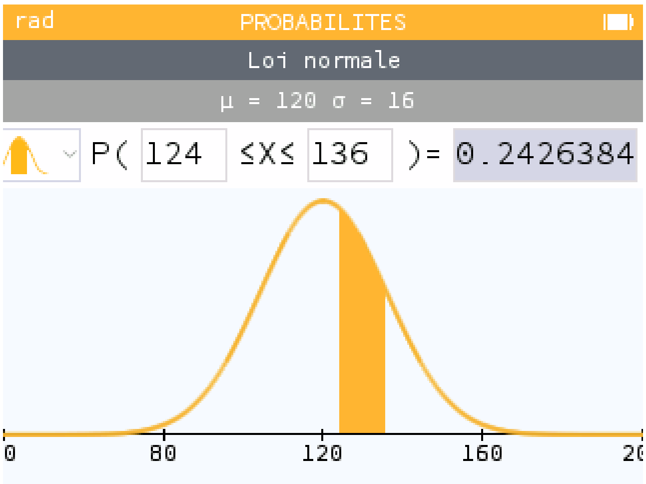
Là, on entre les paramètres :



On choisit la probabilité que la variable aléatoire soit dans un intervalle :



puis on entre les bornes de l'intervalle :



et on lit la probabilité : 0,2426 à 10^{-4} près (le nomogramme donnait environ 0,24).

b) Texas Instruments

En faisant 2nde var, on a la liste des « distributions » (les lois de probabilité). La première entrée `normalFdP` ne sert à rien pour cet exercice. Mais la seconde entrée `normalFRep` donne la probabilité qu'une variable aléatoire normale soit entre deux valeurs. On entre alors les bornes 124 et 136 d'abord, les paramètres ensuite.

III/ Statistiques inférentielles

La moyenne de variables aléatoires indépendantes suit une loi normale, la fréquence empirique[1] aussi. On en déduit les résultats suivants :

1) Intervalles de fluctuation

La moyenne de n variables aléatoires indépendantes entre elles, d'espérance μ et d'écart-type σ, a 19 chances sur 20 de se trouver entre $\mu-1{,}96\times\sigma/\sqrt{n}$ et $\mu+1{,}96\times\sigma/\sqrt{n}$: ce sont les bornes de l'intervalle de fluctuation à 95 % pour la moyenne.

La fréquence empirique d'un caractère à deux valeurs suit aussi une loi approximativement normale, d'espérance p (proportion du caractère dans la population) et d'écart-type $\sigma\times\sqrt{p\times(1-p)/n}$, où n est la taille de l'échantillon. Un intervalle de fluctuation pour la fréquence empirique a donc pour bornes $p-1{,}96\times\sqrt{p\times(1-p)/n}$ et $p+1{,}96\times\sqrt{p\times(1-p)/n}$.

2) Intervalles de confiance

Pour calculer un intervalle de confiance, on remplace l'espérance (inconnue) par la moyenne dans l'échantillon, ou la fréquence (inconnue) par celle dans l'échantillon.

---

1 Quotient d'une variable aléatoire binomiale par son premier paramètre.