

## 5. Lois continues usuelles

### 5.1 Utilisation directe de l'unité TI-Nspire

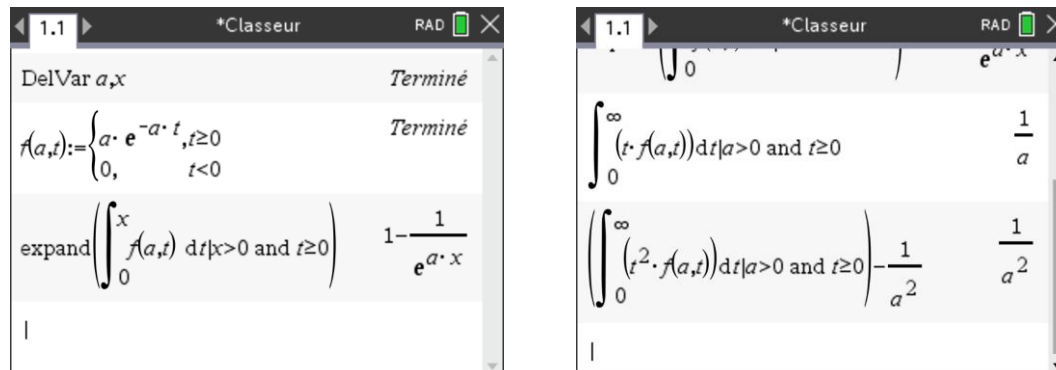
Rappelons pour commencer la définition de deux lois continues classiques, non existantes dans les menus, mais facilement programmables.

1. La loi uniforme sur  $[a, b]$  :  $d(t) = \begin{cases} \frac{1}{b-a} & t \in [a, b] \\ 0 & t \notin [a, b] \end{cases}$ .
2. La loi exponentielle de paramètre  $a$  :  $d(t) = \begin{cases} a e^{-at} & t \geq 0 \\ 0 & t < 0 \end{cases}$ .

La détermination des fonctions de répartition des deux premières ne pose pas de problème :

1. La loi uniforme sur  $[a, b]$  :  $F(t) = \begin{cases} 0 & t \in ]-\infty, a] \\ \frac{t-a}{b-a} & t \in [a, b] \\ 1 & t \in [b, +\infty[ \end{cases}$ .
2. La loi exponentielle de paramètre  $a > 0$  :  $F(t) = \begin{cases} 0 & t \in ]-\infty, 0] \\ 1 - e^{-at} & t \in [0, +\infty[ \end{cases}$ .

On peut vérifier ce dernier résultat avec la calculatrice. On peut de la même façon calculer l'espérance et la variance de cette loi, à condition de bien préciser le signe de  $a$  (écran de droite).



☞ Dans la dernière version du logiciel il faut ajouter la condition sur  $t$  ( $t \geq 0$ ) pour obtenir le résultat, chose qui n'était pas nécessaire dans les versions précédentes...

Des calculs de ce type pourront être faits de la même façon avec d'autres lois que vous rencontrerez dans des exercices.

### 5.2 Lois continues présentes sur la TI-Nspire

- Loi normale de moyenne  $\mu$  et d'écart type  $\sigma$

**normPdf** (densité de probabilité de la loi normale)

calcule la valeur de la densité de probabilité de la loi normale de moyenne  $\mu$  et d'écart type  $\sigma$ , en un réel  $x$  spécifié. Les valeurs par défaut sont  $\mu = 0$  et  $\sigma = 1$ . La densité de probabilité est définie par :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, (\sigma > 0)$$

syntaxe : **normPdf**( $x$  [,  $\mu$ ,  $\sigma$ ])

fonction de répartition : **normCdf**( $a$ ,  $b$ , [,  $\mu$ ,  $\sigma$ ]) donne la probabilité  $P(a \leq X \leq b)$ .

**InvNorm** permet de calculer la valeur  $x$  telle que  $P(X \leq x) = \alpha$ ,  $\alpha \in [0,1]$ , lorsque  $X$  suit la loi normale de moyenne  $\mu$  et d'écart type  $\sigma$ .

Syntaxe : **invNorm**( $\alpha$  [,  $\mu$ ,  $\sigma$ ]).

- Loi de Student à  $df$  degrés de liberté ( $df \in \mathbb{N}^*$ ).

**tPdf** (densité d'une loi de Student)

calcule la densité de probabilité de la loi de Student à  $df$  degrés de liberté en un réel  $x$  spécifié. La densité de probabilité est définie par :

$$f(x) = \frac{\Gamma\left(\frac{df+1}{2}\right)}{\Gamma(df/2)} \frac{(1+x^2/df)^{-(df+1)/2}}{\sqrt{\pi df}}$$

où  $\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt$

$$E(X) = 0 \text{ et } V(X) = \frac{df}{df-2} \text{ si } df > 2$$

syntaxe : **tPdf**( $x$ ,  $df$ )

fonction de répartition : **tCdf**( $a$ ,  $b$ ,  $df$ ) donne la probabilité  $P(a \leq X \leq b)$ .

- Loi de Fisher à  $df_1$  et  $df_2$  degrés de liberté ( $(df_1, df_2) \in \mathbb{N}^{*2}$ )

**FPdf** (densité d'une loi de Fisher)

calcule la densité de probabilité de la distribution de Fisher en un réel  $x \in \mathbb{R}_+^*$  spécifié. La densité de probabilité s'exprime sous la forme :

$$f(x) = \frac{\Gamma\left(\frac{n+d}{2}\right)}{\Gamma(n/2)\Gamma(d/2)} \left(\frac{n}{d}\right)^{n/2} x^{n/2-1} (1+nx/d)^{-(n+d)/2}$$

avec :  $n = df_1$  et  $d = df_2$ .

$$E(X) = \frac{df_2}{df_2-2} \text{ si } df_2 > 2 \text{ et}$$

$$V(X) = \left(\frac{df_2}{df_2-2}\right)^2 \frac{2(df_1+df_2-2)}{df_1(df_2-4)} \text{ si } df_2 > 4.$$

syntaxe : **FPdf**( $x$ ,  $df_1$ ,  $df_2$ )

fonction de répartition : **FCdf**( $a$ ,  $b$ ,  $df_1$ ,  $df_2$ ) donne la probabilité  $P(a \leq X \leq b)$ .

- Loi du Khi 2 ( $\chi^2$ ) (ou de Pearson) à  $df$  degrés de liberté ( $df \in \mathbb{N}^*$ )

**chi2Pdf** (densité de probabilité d'une loi du Khi 2)

calcule la densité de probabilité de la loi du Khi 2, en une valeur spécifiée  $x \in \mathbb{R}_+^*$ . La densité de probabilité est définie par :

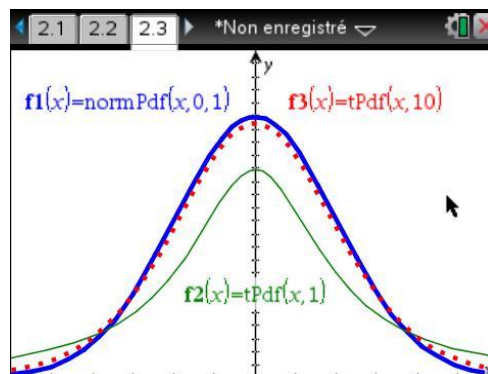
$$f(x) = \frac{1}{\Gamma(df/2)} (1/2)^{df/2} x^{df/2-1} e^{-x/2}, x > 0$$

$$E(X) = df \text{ et } V(X) = 2df$$

syntaxe : **chi2Pdf**( $x, df$ )

fonction de répartition : **chi2Cdf**( $a, b, df$ ) donne la probabilité  $P(a \leq X \leq b)$ .

**Exemple** : représentation de la loi normale centrée réduite (trait épais bleu) et des lois de Student de, respectivement, 1 (trait fin vert) et 10 (en pointillés gras rouge) degrés de liberté.



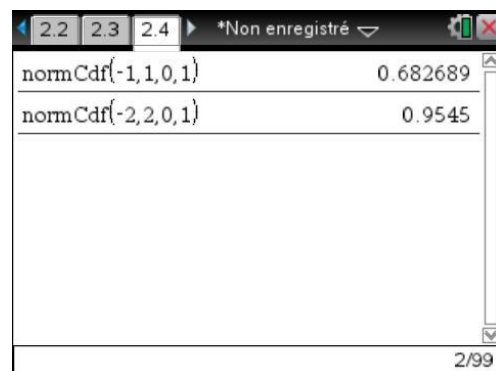
### 5.3 Quelques résultats classiques sur les lois normales

On considère une loi normale de paramètres  $m$  et  $\sigma$ .

1. On demande de calculer  $P(m - \sigma \leq X \leq m + \sigma)$  et  $P(m - 2\sigma \leq X \leq m + 2\sigma)$ .
2. On demande également de déterminer, en fonction de  $\sigma$ , la valeur de  $a$  telle que  $P(m - a \leq X \leq m + a) = 85/100$ .

Pour la question 1, il suffit d'utiliser les fonctions **normCdf**, avec les bornes  $-1, 1$  puis  $-2, 2$  et les paramètres  $m = 0$  et  $\sigma = 1$ , car :

$$P(m - a.\sigma \leq X \leq m + a.\sigma) = P\left(-a \leq \frac{X - m}{\sigma} \leq a\right)$$



Deuxième question.

$$P(m - a \leq X \leq m + a) = 85/100 \text{ équivaut à } P(|X - m| < a) = P\left(\left|\frac{X - m}{\sigma}\right| < \frac{a}{\sigma}\right) \simeq 2\Phi\left(\frac{a}{\sigma}\right) - 1 = 0.85$$

où  $\Phi$  est la fonction de répartition de la loi normale centrée réduite. La fonction **invNorm** permet d'obtenir le résultat :  $a \simeq 1.44\sigma$ .

normCdf(-1,1,0,1)	0.682689
normCdf(-2,2,0,1)	0.9545
solve(2·x-1=0.85,x)	x=0.925
invNorm(0.925,0,1)	1.43953

## 5.4 Approximations usuelles

On peut visualiser à l'aide de la calculatrice les approximations usuelles des diverses lois.

Si  $\begin{cases} n \geq 30 \\ np \geq 15 \\ npq > 5 \end{cases}$  la loi binomiale  $\mathcal{B}(n, p)$  peut être approximée par la loi normale  $\mathcal{N}(np, \sqrt{npq})$ .

Si  $\begin{cases} p \leq 0,1 \\ n \geq 30 \\ np < 15 \end{cases}$  la loi binomiale  $\mathcal{B}(n, p)$  peut être approximée par la loi de Poisson  $\mathcal{P}(np)$ .

Si  $\lambda \geq 15$  la loi de Poisson  $\mathcal{P}(\lambda)$  peut être approximée par la loi normale  $\mathcal{N}(\lambda, \sqrt{\lambda})$ .

Si  $df \geq 30$  la loi de Student à  $df$  degrés de liberté peut être approximée par la loi normale centrée réduite (voir représentation graphique page 15).

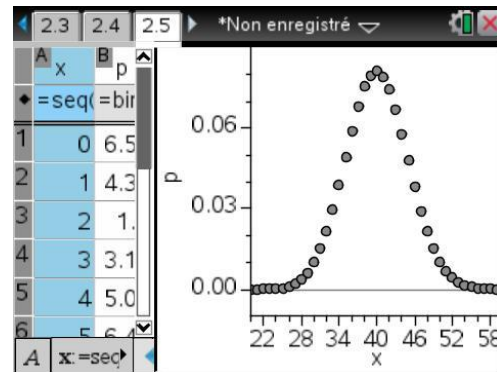
Les graphiques suivants illustrent ces problèmes d'approximation. On travaille avec l'application **Tableur & listes**, dans la colonne A on place les entiers de 0 à 100 (= **seq(i,i,0,100)**), on la nomme x. Dans la colonne B, on place les probabilités de la loi binomiale de paramètres  $n=100$  et  $p=0,4$ ,  $\mathcal{B}(100;0,4)$  (= **binomPdf(100,0.4)**), on la nomme p, et dans la colonne C les probabilités de la loi de Poisson de paramètre 40,  $\mathcal{P}(40)$  (= **poissPdf(40,a[i])**), on la nomme q.

	A x	B p	C q
	=seq(i,i,0,100)	=binompdf(100,0.4)	=poisspdf(40,a[i])
1	0	6.53319E-23	4.24835E-18
2	1	4.35546E-21	1.69934E-16
3	2	1.4373E-19	3.39868E-15
4	3	3.13012E-18	4.53158E-14
5	4	5.06036E-17	4.53158E-13
6	5	6.47727E-16	2.62526E-12

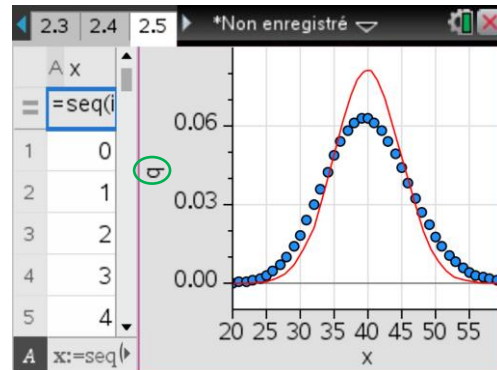
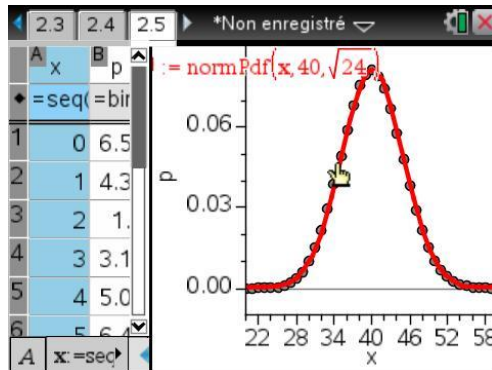
On sélectionne les deux premières colonnes. **menu** **3** **9** (**Graphe rapide**) permet de représenter la première loi. On peut modifier le partage d'écran à l'aide du TouchPad, cliquer sur la barre centrale, la flèche se transforme en  $\blacktriangleleft$ , **ctrl** **menu** **4** **1** permet de régler la fenêtre graphique, **ctrl** **tab** pour passer de l'écran graphique au tableur et réciproquement.

A	x	B	p	C	q
=seq(	=binompdf(100,0.4)	=poisspdf(40,a[])			
1	0	6.53319E-23			4.24835E-18
2	1	4.35546E-21			1.69934E-16
3	2	1.4373E-19			3.39868E-15
4	3	3.13012E-18			4.53158E-14
5	4	5.06036E-17			4.53158E-13
6	5	6.47727E-16			2.62526E-12

1: Actions	2: Insertion	3: Données	4: Statistiques	5: Table des va
1: Générer une suite	2: Capture de données	3: Remplissage	4: Supprimer	5: Aléatoire
6: Liste Maths	7: Opérations sur les listes	8: Résumé graphique	9: Graphe rapide	



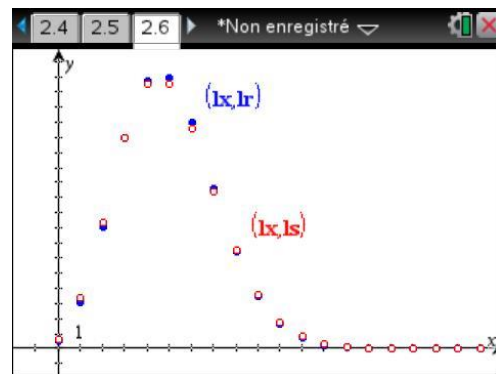
Le menu **4** **4** permet de tracer la courbe représentative de la densité de la loi normale de paramètres  $np = 40$  et  $\sqrt{npq} = \sqrt{24}$ . Pour obtenir l'écran de droite, on clique sur la lettre p représentant l'axe des ordonnées, et on choisit la variable q.



On voit bien que, pour ces valeurs des paramètres, l'approximation par la loi normale  $\mathcal{N}(40, \sqrt{24})$  est bonne, alors qu'elle ne l'est pas par la loi de Poisson.

On peut également étudier un cas où la loi de Poisson donne une bonne approximation de la loi binomiale. On colle dans la première colonne la liste lx définie par **seq(x,x,0,100)** (**seq** s'obtient dans le catalogue, ou si l'on connaît la syntaxe peut être tapé directement), puis dans lr : **binomPdf(100, 0.05)** et dans ls : **poissPdf(5, lx)**. Plutôt que d'utiliser comme ci-dessus **Graphe rapide**, on insère une nouvelle page avec l'application **Graphiques & géométrie**, ce qui permet d'avoir des représentations plus lisibles et on définit les deux nuages de points (lx, lr), (lx, ls).

	lx	lr	ls
	=seq(binompdf(100,0.05)=poisspdf(5,lx)		
1	0	0.005921	0.006738
2	1	0.031161	0.03369
3	2	0.081182	0.084224
4	3	0.139576	0.140374
5	4	0.178143	0.175467
6	5	0.180019	0.175467
A lx=seq(i,i,0,100)			



Ils se superposent presque parfaitement ( $p = 0,05 \leq 0,1$ ,  $n = 100 \geq 30$  et  $np = 5 < 15$ ).

### Exemple 1

On lance une paire de dés non pipés,  $X_i$  représente la somme des deux nombres obtenus au  $i$ -ième lancer. Déterminer le nombre de lancers pour que la moyenne des résultats  $X_i$  diffère de 7 de moins de 0,1 avec une probabilité supérieure à 0,95.

Les  $X_i$  constituent une suite de variables indépendantes, de même loi de moyenne 7 et d'écart type  $\sqrt{\frac{35}{6}}$ .

Le théorème central limite permet de dire que la variable représentant la moyenne  $F_n = \frac{X_1 + X_2 + \dots + X_n}{n}$  centrée, réduite,  $Y_n = \frac{F_n - 7}{\sqrt{\frac{35}{6n}}}$  converge en loi vers la variable normale

centrée réduite, quand  $n$  tend vers  $+\infty$ .

En utilisant cette approximation, nous avons donc :

$$P(|F_n - 7| < 0,1) = P(|Y_n| < 0,1 \cdot \sqrt{\frac{6n}{35}}) \simeq 2\Phi\left(0,1 \cdot \sqrt{\frac{6n}{35}}\right) - 1$$

où  $\Phi$  est la fonction de répartition de la loi normale centrée réduite. Nous voulons donc trouver  $n$  tel que :

$$2\Phi\left(0,1 \cdot \sqrt{\frac{6n}{35}}\right) - 1 \geq 0,95, \text{ soit } \Phi\left(0,1 \cdot \sqrt{\frac{6n}{35}}\right) \geq 0,975$$

invNorm(0.975)	1.95996
solve(0.1 * sqrt(6*n/35) = 1.96, n)	n=2240.93

On doit avoir  $0,1 \cdot \sqrt{\frac{6n}{35}} \geq 1,96$ , il suffit donc de prendre  $n \geq 2241$ .

### Exemple 2 : Approximation de la loi hyper-géométrique par une loi binomiale, puis par une loi normale

Un jour d'élections, dans une ville comptant 25 000 votants, 12 800 votent pour la municipalité sortante. On dépouille 50 bulletins tirés au hasard parmi les 25 000. Quelle est la probabilité que, sur cet échantillon, la majorité sortante soit minoritaire ? (Au maximum 24 bulletins favorables.)

On se trouve ici en présence d'une variable aléatoire suivant une loi hypergéométrique de paramètres  $N = 25\,000$ ,  $p = 12\,800 / 25\,000$  et  $n = 50$ .

Attention, si vous demandez un calcul exact, vous risquez d'obtenir un résultat totalement inutilisable et un temps de calcul important. Il est préférable soit de valider par **ctrl enter**, soit de mettre un point à la fin d'un des nombres (comme le montre l'écran précédent) pour que le calcul se fasse en mode approché.

On peut comparer avec le résultat que l'on obtiendrait en prenant une loi binomiale de paramètres  $n = 50$  et  $p = 12\,800 / 25\,000$  :

$p := \frac{12800.}{25000}$	0.512
$hyp(25000, 50, p, 24)$	0.377492
$binomCdf(50, p, 24)$	0.377609
3/99	

Le résultat obtenu est très proche du précédent, ce qui est conforme à la théorie sur l'approximation par une loi binomiale dans le cas où  $N > 10n$ .

On peut enfin effectuer une nouvelle approximation, en utilisant à présent une loi normale de paramètres  $m = np$  et  $\sigma = \sqrt{np(1-p)}$ , et en utilisant une correction de continuité.

On calcule  $P(X \leq 24,5)$  :

$hyp(25000, 50, p, 24)$	0.377492
$binomCdf(50, p, 24)$	0.377609
$n := 50$	50
$m := n \cdot p$	25.6
$\sigma := \sqrt{n \cdot p \cdot (1-p)}$	3.53452
$normCdf(-\infty, 24.5, m, \sigma)$	0.377818
7/99	

On peut voir que cette approximation est également satisfaisante, ce qui était prévisible puisque les trois conditions que l'on impose en général pour ce type d'approximation :  $n \geq 30$ ,  $np \geq 15$ ,  $np(1-p) > 5$  étaient bien toutes vérifiées.

Vous remarquerez que dans les conditions de l'exercice, un sondage sur un aussi petit nombre de personnes risque de donner un résultat faux plus d'une fois sur trois.

Reprenons rapidement le calcul avec un échantillon de 1 000 personnes :

$\sigma = \sqrt{n \cdot p \cdot (1-p)}$	3.53452
$\text{normCdf}(-\infty, 24.5, m, \sigma)$	0.377818
$n := 1000$	1000
$m := n \cdot p$	512.
$\sigma = \sqrt{n \cdot p \cdot (1-p)}$	15.8068
$\text{normCdf}(-\infty, 499.5, m, \sigma)$	0.214531
	11/99

Cette fois, il n'y a plus qu'environ une "chance" sur cinq de se tromper.  
(Ce qui n'est pas totalement négligeable...)

## 5.5 Estimations et intervalles de confiance

### • Estimateur sans biais

Soient  $\mu$  (moyenne) et  $\sigma$  (écart type) les paramètres d'un caractère  $X$  dans une population  $\mathcal{P}$  de taille  $N$ . On prélève tous les échantillons de taille  $n$ , on note  $(n, \bar{x}_i, s_i)$  les paramètres de  $X$ .

Si  $(n, \bar{x}, s)$  sont les paramètres d'un échantillon pris au hasard,  $\bar{x}$  est un estimateur sans biais de  $\mu$ , par contre  $s$  n'est pas un estimateur sans biais de  $\sigma$ , il faut prendre pour cela  $S$  défini par

$$S^2 = \frac{n}{n-1} s^2, \text{ soit } S = \sqrt{\frac{1}{n-1} \sum_{j=1}^n (x_j - \bar{x})^2} \text{ (les } x_j \text{ étant les valeurs de } X \text{ dans l'échantillon).}$$

### • Distribution d'échantillonnage des moyennes

On note  $\bar{X}$  la variable qui suit la loi de probabilité déterminée par la distribution d'échantillonnage des moyennes  $\{\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k\}$  ( $k$  est le nombre d'échantillons,  $N = nk$ ).

$$E(\bar{X}) = E(X) = \mu$$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \text{ dans le cas d'un tirage non exhaustif (avec remise)}$$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} \text{ dans le cas d'un tirage exhaustif (sans remise)}$$

☞ Si  $N$  est grand  $\sqrt{\frac{N-n}{N-1}} \simeq 1$ , on peut supposer le tirage non exhaustif,  
si  $\sigma$  est inconnu on le remplace dans le calcul par son estimateur  $S$ .

### • Distribution d'échantillonnage des fréquences (ou proportions)

On se place dans le cas où le caractère  $X$  ne peut prendre que les valeurs 0 et 1 avec respectivement les probabilités  $p$  et  $q = 1 - p$ . Soit  $F$  la variable prenant pour valeur la fréquence  $f_i$  d'apparitions de "l'événement favorable" dans le  $i$ -ième échantillon.

$$E(F) = p$$

$$\sigma_F = \sqrt{\frac{pq}{n}} \text{ dans le cas d'un tirage non exhaustif (avec remise)}$$



$$\sigma_F = \sqrt{\frac{pq}{n}} \sqrt{\frac{N-n}{N-1}} \text{ dans le cas d'un tirage exhaustif (sans remise)}$$

☞ Si  $N$  est grand  $\sigma_F \simeq \sqrt{\frac{pq}{n}}$ , on peut supposer le tirage non exhaustif. Si  $p$  est inconnue on la remplace par son estimation  $f$ , fréquence observée sur un échantillon, et  $n$  par  $n-1$ , si  $n$  petit.

### • Intervalle de confiance d'une moyenne

Si  $n \geq 30$ , que la population mère soit normalement distribuée ou non, on peut supposer que la variable centrée réduite  $T = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$  suit approximativement une loi normale centrée réduite.

$$\text{Si } P(|T| \leq t_{\alpha/2}) = P\left(\mu - t_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \bar{X} \leq \mu + t_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

ce qui revient à dire que la probabilité qu'une valeur  $\bar{x}$  de  $\bar{X}$  appartienne à l'intervalle  $\left[\mu - t_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \mu + t_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right]$  est égale à  $1 - \alpha$ , ou encore que  $\mu \in \left[\bar{x} - t_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{x} + t_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right]$  avec la probabilité  $1 - \alpha$  ou le risque d'erreur  $\alpha$ .

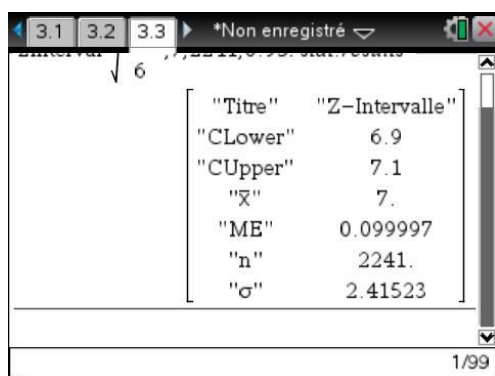
L'intervalle  $\left[\bar{x} - t_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{x} + t_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right]$  est appelé intervalle de confiance de l'estimateur de  $\mu$  au niveau de confiance  $1 - \alpha$  ou au risque d'erreur  $\alpha$ .

☞ Si le tirage est exhaustif on remplace  $\frac{\sigma}{\sqrt{n}}$  par  $\frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$   
si  $\sigma$  est inconnu on le remplace par son estimateur  $S$ .

Si  $n < 30$ , si la population mère est normalement distribuée et si l'échantillon est non exhaustif, on peut supposer que la variable centrée réduite  $T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$  ( $S$  estimateur de  $\sigma$  voir page 20) suit une loi

de Student à  $n-1$  degrés de liberté, les calculs sont les mêmes que ci-dessus. Seuls changent les  $t_{\alpha/2}$  qui sont donnés ici par une loi de Student au lieu de la loi normale. Dans la pratique on utilise la fonction **zInterval** (**menu** **6** **6** **1**) dans le premier cas :  $\sigma$  connu ; **tInterval** (**menu** **6** **6** **2**) dans le second cas :  $\sigma$  inconnu.

**Exemple.** On peut contrôler le résultat de l'exemple du paragraphe précédent. On utilise **zInterval** avec l'option Stats et les paramètres  $\sigma = \sqrt{\frac{35}{6}}$ ,  $\bar{x} = 7$ ,  $n = 2241$  et  $C.Level = 0,95$ . On retrouve bien l'intervalle  $[6,9;7,1]$ .



- **Intervalle de confiance d'une fréquence (ou proportion)**

Avec les notations vues page 20, on appelle intervalle de confiance de l'estimateur de  $p$  au niveau de confiance  $1-\alpha$  ou au risque d'erreur  $\alpha$  l'intervalle  $\left[ f - t_{\alpha/2} \sqrt{\frac{pq}{n}}, f + t_{\alpha/2} \sqrt{\frac{pq}{n}} \right]$ . Utiliser **zInterval\_1Prop** (menu **6 6 5**).

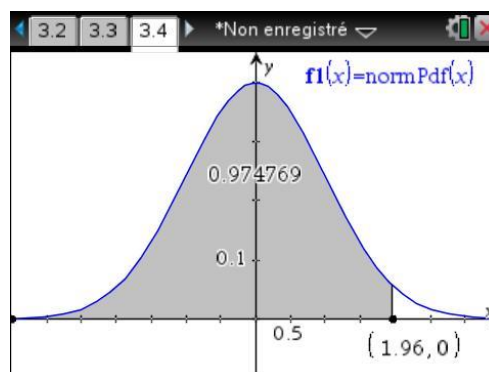
☞ Si le tirage est exhaustif on remplace  $\sqrt{\frac{pq}{n}}$  par  $\sqrt{\frac{pq}{n}} \sqrt{\frac{N-n}{N-1}}$   
 si  $p$  est inconnue on remplace  $\sqrt{\frac{pq}{n}}$  par son estimateur  $\sqrt{\frac{f(1-f)}{n-1}}$ .

- **Calcul des  $t_{\alpha/2}$**

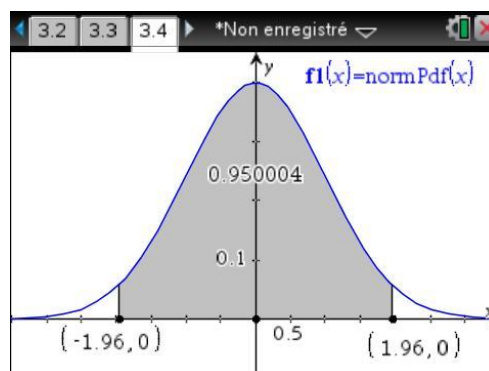
Si  $T$  suit la loi normale centrée réduite  $\mathcal{N}(0,1)$ , notons  $t_{\alpha/2}$  le réel positif tel que  $P(|T| \leq t_{\alpha/2}) = 1 - \alpha$ .

$t_{\alpha/2}$  peut être donné par la fonction **InvNorm** (menu **5 5 3**) en prenant pour valeur de Area  $1 - \alpha/2$ , voir le graphique ci-dessous.

invNorm(0.95,0,1)	1.64485
invNorm(0.975,0,1)	1.95996
invNorm(0.995,0,1)	2.57583
invNorm(0.9975,0,1)	2.80703



invNorm(0.995,0,1)	2.57583
invNorm(0.9975,0,1)	2.80703
normCdf(-1.645,1.645)	0.90003
normCdf(-1.96,1.96)	0.950004
normCdf(-2.58,2.58)	0.99012
normCdf(-2.81,2.81)	0.995046



On peut ainsi dresser le tableau suivant :

Risque $\alpha$	0,5 %	1 %	5 %	10 %
Seuil de confiance $1-\alpha$	99,5 %	99 %	95 %	90 %
$t_{\alpha/2}$	2,81	2,58	1,96	1,645

Pour un effectif faible ( $n < 30$ ) on peut procéder de la même façon pour calculer  $t_{\alpha/2}$  en utilisant la loi de Student. Par exemple au risque d'erreur de 5 % et pour 5 degrés de liberté  $t_{\alpha/2} \simeq 2,57$  (voir calcul ci-dessous).

normCdf(-1.96,1.96)	0.950004
normCdf(-2.58,2.58)	0.99012
normCdf(-2.81,2.81)	0.995046
normCdf(-∞,1.96)	0.975002
invT(0.975,5)	2.57058
tCdf(-2.57,2.57,5)	0.949965

vérification

☞ La valeur est supérieure à celle donnée par la loi normale, ceci vient du fait que la distribution de Student est plus aplatie.

Dans le menu Intervalles de confiance il existe d'autres fonctions, par exemple :

**zInterval\_2Samp** et **tInterval\_2Samp** (menu **6** **6** **3** et **4**) qui calculent un intervalle de confiance pour la différence de deux moyennes lorsque respectivement les écarts types sont connus, ou ne le sont pas ; **zInterval\_2Prop** (menu **6** **6** **6**) calcule un intervalle de confiance pour la différence de deux proportions.

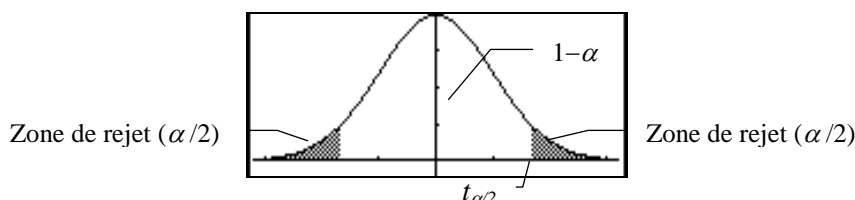
## 5.6 Tests

### • Tests de conformité d'un échantillon à une population

Comparaison d'une moyenne (resp. fréquence) observée à une moyenne (resp. fréquence) théorique.

Sous l'hypothèse nulle  $H_0 : \mu = \mu_0$ , la variable  $\frac{\bar{X} - \mu_0}{\sigma_{\bar{X}}}$  peut être assimilée à une variable normale centrée réduite ou à une variable suivant une loi de Student, suivant la taille de l'échantillon (voir page 21).

**Test bilatéral** : on teste l'hypothèse nulle  $H_0 : \mu = \mu_0$  et l'alternative  $\mu \neq \mu_0$ . Au seuil de signification  $\alpha$  la zone de rejet est la partie grisée.

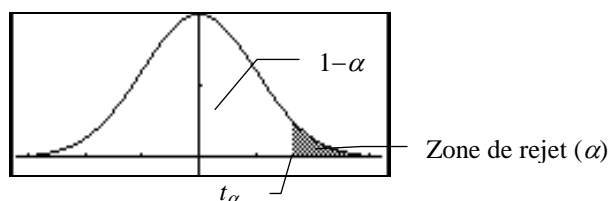


Si la valeur de  $\frac{|\bar{X} - \mu_0|}{\sigma_{\bar{X}}}$  est inférieure à  $t_{\alpha/2}$  on accepte l'hypothèse  $H_0$  ;

sinon on rejette l'hypothèse  $H_0$ .

$t_{\alpha/2}$  est donné par la loi normale centrée réduite quand  $n \geq 30$  et par la loi de Student à  $n-1$  degrés de liberté dans le cas contraire.

**Test unilatéral à droite** : on teste l'hypothèse nulle  $H_0 : \mu = \mu_0$  et l'alternative  $\mu > \mu_0$ . Au seuil de signification  $\alpha$  la zone de rejet est la partie grisée :



Si la valeur de  $\frac{\bar{X} - \mu_0}{\sigma_{\bar{X}}}$  est inférieure à  $t_{\alpha}$  on accepte l'hypothèse  $H_0$  ;<sup>2</sup>

sinon, on rejette l'hypothèse  $H_0$  (avec les mêmes remarques que ci-dessus).

**Test unilatéral à gauche** : on teste l'hypothèse nulle  $H_0 : \mu = \mu_0$  et l'alternative  $\mu < \mu_0$ . La zone de rejet est à gauche, symétrique de celle à droite et la condition d'acceptation de  $H_0$  :

la valeur de  $\frac{\bar{X} - \mu_0}{\sigma_{\bar{X}}}$  est supérieure à  $-t_{\alpha}$ .

<sup>2</sup>  $t_{\alpha}$  est directement donné par **invNorm** (menu **6** **5** **3**) ou **invT** (Student menu **6** **5** **6**).