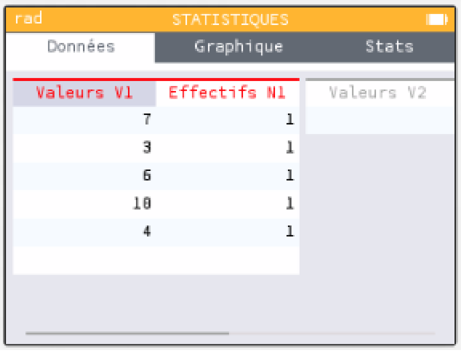
Statistiques Inférentielles (correction)

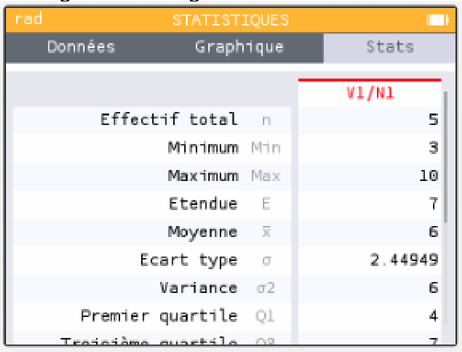
I - Un exemple pour découvrir

Une population est constituée de 5 étudiants de BTS. Leur professeur de mathématiques adoré et vénéré veut estimer le temps moyen hebdomadaire consacré au travail personnel.

Étudiant	Temps de travail (secondes)
Albert	7
Bernard	3
Charles-Henri	6
Dolf	10
Évariste	4

Calculez moyenne et écart-type : $\mu = \frac{7+3+6+10+4}{5} = \frac{30}{5} = 6 \;, \sigma \approx 2{,}45$.





Le problème, c'est qu'on n'a pas souvent accès à toute la population et on se contente d'échantillons.
Voyons ce qui se passe en prenant des échantillons de taille 3 de la population précédente.
Tout d'abord, combien y a-t-il d'échantillons de taille 3 ?
Remplissez le tableau suivant :

Echantillon	Données	Moyenne de l'échantillon	Ecart type échantillon
1	7 3 6	5,33	1,70
2	7 3 10	6,67	2,87
3	7 6 10	7,67	1,70
4	3 6 10	6,33	2,87
5	7 3 4	4,67	1,70
6	7 6 4	5,67	1,25
7	3 6 4	4,33	1,24
8	7 10 4	7	2,45
9	3 10 4	5,67	3,09
10	6 10 4	6,67	2,49

Calculez la moyenne des moyennes échantillonnales : $\mu_{\bar{x}} = 6{,}001$ Pour les écart-types : $\sigma_{\bar{x}} = 1{,}001$

II - Estimation ponctuelle

On ne considère qu'une seule mesure. On ne cherche pas à calculer le risque d'erreur.
Dans ce cas

$\mu_{\bar{x}}$ est une estimation ponctuelle de la moyenne

et on admettra que

$\sigma_{\bar{x}}\sqrt{\frac{n}{n-1}}$ est une estimation ponctuelle de l'écart-type σ

Exercice 1 *Estimation*

Afin de mieux gérer les demandes de crédits de ses clients, un directeur d'agence bancaire réalise une étude relative à la durée de traitement des dossiers, supposée suivre une distribution normale. Un échantillon non exhaustif de 30 dossiers a donné :

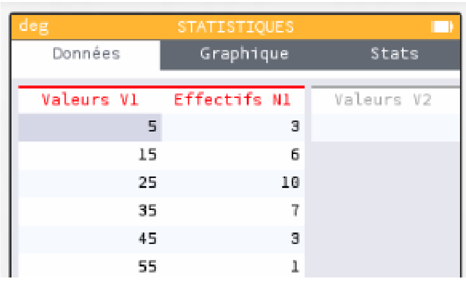
Durée mn	0 – 10	10 –20	20 – 30	30 – 40	40 – 50	50 – 60
Effectif	3	6	10	7	3	1

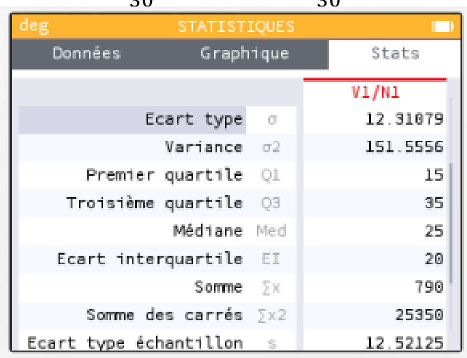
1. Calculer la moyenne $\bar{x}$ et l'écart type σ des durées de traitement des dossiers de cet échantillon.
2. En déduire les estimations ponctuelles de la moyenne m et de l'écart type s de la population des dossiers.

Solution

Quand on a une classe , par exemple dans la première colonne $0 - 10$, on considère que les valeurs associées sont toutes au milieu de la classe : $\frac{0+10}{2} = 5$ ainsi :

$$\bar{x} = \frac{5\times3+15\times6+25\times10+35\times7+45\times3+55\times1}{3+6+10+7+3+1} = \frac{15+90+250+245+135+55}{30} = \frac{790}{30} \approx 26{,}33$$





Pour l'écart type on utilisera directement la calculatrice $\sigma \approx 12{,}31$

2. d'après le cours $m \approx \bar{x}$ et $s \approx \sqrt{\frac{n}{n-1}}\sigma = \sqrt{\frac{30}{29}}\sigma \approx 12{,}52$

Exercice 2 : On roule pour vous

Une entreprise utilise des camions pour transporter sa production. Elle dispose de 100 camions.

Elle repère sur un échantillon de 30 jours choisis au hasard, le nombre de camions en panne. Voici les résultats :

5,6, 4, 6, 6, 8, 3, 5, 5, 5, 4, 3, 6, 5, 6, 4, 7, 6, 6, 5, 4, 3, 6, 5, 4, 5, 4, 5, 5, 4

Véhicules en panne	3	4	5	6	7	8	Total
Nbr de jours	3	7	10	8	1	1	30

1. Calculer la moyenne $\bar{x}$ et l'écart-type σ du nombre de camions en panne chaque jour pour l'échantillon étudié.

Avec la calculatrice on obtient $\bar{x} = 5$ et $\sigma \approx 1{,}15$.

2. À partir des résultats obtenus pour cet échantillon, proposer une estimation ponctuelle de la moyenne μ et de l'écart-type s du nombre de camions en panne chaque jour pour la population correspondant aux jours ouvrables de l'année.

$\mu \approx \bar{x} = 5$ et de l'écart-type $s \approx \sqrt{\frac{n}{n-1}}\sigma \approx \sqrt{\frac{30}{29}}1{,}15 \approx 1{,}17$

Exercice 3

En vue de réaliser un programme de rééducation, des chercheurs ont soumis un questionnaire de neuropsychologie cognitive à 150 enfants dyslexiques tirés au sort. Le questionnaire comporte 20 questions et les chercheurs ont recueilli pour chaque enfant dyslexique le nombre x_i de bonnes réponses. Les résultats ainsi récoltés sont tels que :

$\sum x_i = 1502$, $\sum {x_i}^2 = 19\,486$

1) Identifier la population, la variable, son type et son/ses paramètre(s).
2) Donner une estimation ponctuelle du nombre moyen de bonnes réponses dans la population étudiée.
3) Donner une estimation ponctuelle de l'écart-type de la variable.

Aide : la moyenne de l'échantillon est $\bar{x} = \frac{\sum x_i}{n}$ et son écart type sera $\sigma = \sqrt{\frac{\sum x_i^2}{n} - \bar{x}^2}$

1.1) Identifier la population, la variable, son type et son/ses paramètre(s).
Réponse :
- Population : Enfants dyslexiques.
- Variable quantitative X = "Nombre de bonnes réponses au questionnaire".
- Deux paramètres inconnus : moyenne μ et écart-type s.

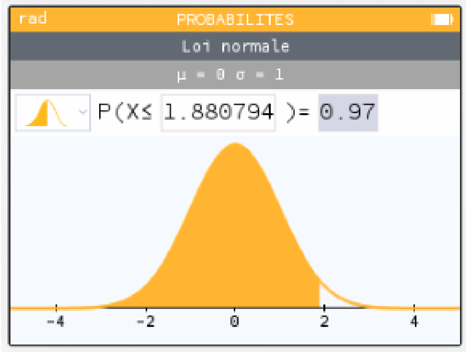
1.2) Donner une estimation ponctuelle du nombre moyen de bonnes réponses dans la population étudiée.
On estime la moyenne μ par la moyenne observée $\bar{x} = \frac{1502}{150} \approx 10{,}01$

1.3) Donner une estimation ponctuelle de l'écart-type de la variable.
Commençons par calculer l'écart type observé : $\sigma = \sqrt{\frac{\sum x_i^2}{n} - \bar{x}^2} = \sqrt{\frac{19\,486}{150} - \left(\frac{1502}{150}\right)^2} \approx \sqrt{129{,}91 - 100{,}27} \approx 5{,}44$

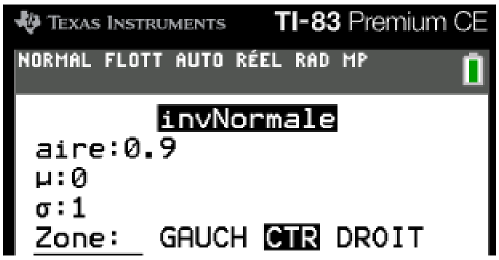
L'écart type estimé sera $s = \sqrt{\frac{250}{249}}\sigma \approx 5{,}45$

III - Estimation de la moyenne par intervalle de confiance

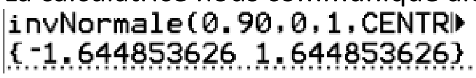
α	1%	5%	10%
t	2,58	1,96	1,645



Recherche de t quand on connait α : Avec la calculatrice Numworks, on sélectionne probabilité dans le menu, puis on sélectionne la loi normale
Chercher le t associé à 10% ça correspond à trouver les valeurs -t et t tel que $P(-t \leq T \leq t) = 0{,}90$
La partie interne a pour aire 0,90 et les zones externes correspondent au 10% restant.
Il y a 5% d'un côté et 5% de l'autre.
Ce qui veut dire que l'on aura $P(-\infty \leq T \leq t) = 0{,}05 + 0{,}90 = 0{,}95$ je vais donc fixer la probabilité à 0,95 et la calculatrice me donnera le $t \approx 1{,}645$.



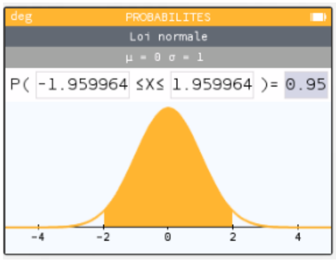
Avec la Ti 83 on utilisera le menu distrib (2nd + var), on sélectionne la fonction invNormale, et on va indiquer qu'on veut avoir 90% au centre, donc aire = 0,9 ; $\mu = 0$; $\sigma = 1$ et Zone = CTR
La calculatrice nous communique alors les bornes -t et t.

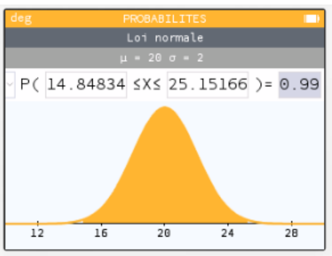


Exercice 4 : Introduction aux intervalles de confiance

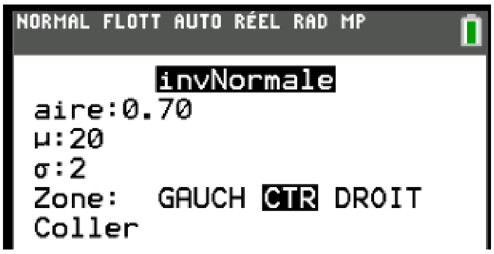
1. Soit X une variable aléatoire suivant la loi $N(0,1)$.

a) Déterminer le réel *a* tel que $P(-a \leq X \leq a) = 0{,}99$ $\quad a \approx 2{,}576$
b) Déterminer le réel *a* tel que $P(-a \leq X \leq a) = 0{,}95$ $\quad a \approx 1{,}960$
c) Déterminer le réel *a* tel que $P(-a \leq X \leq a) = 0{,}90$ $\quad a \approx 1{,}645$
d) Déterminer le réel *a* tel que $P(-a \leq X \leq a) = 0{,}70$ $\quad a \approx 1{,}036$





2. Soit X une variable aléatoire suivant la loi normale centrée réduite $N(20,2)$.
a) Déterminer le réel *a* tel que $P(20 - a \leq X \leq 20 + a) = 0{,}99$
sur la calculatrice la borne de droite est $25{,}1516 = 20 + 5{,}1516$ donc $a \approx 5{,}152$
on remarquera que l'on a bien $a = t_{1\%}\sigma$ $\quad$ ($t_{1\%}$ a été calculée à la question 1a))
b) Déterminer le réel *a* tel que $P(20 - a \leq X \leq 20 + a) = 0{,}95$ $\quad a \approx 3{,}92$
c) Déterminer le réel *a* tel que $P(20 - a \leq X \leq 20 + a) = 0{,}90$ $\quad a \approx 3{,}29$
d) Déterminer le réel *a* tel que $P(20 - a \leq X \leq 20 + a) = 0{,}70$ $\quad a \approx 2{,}072$



avec la Ti83 on utilisera 2nde+Var = distrib , fonction invNormale, on rentre les valeurs associées à la question, ci-contre on a les valeurs pour la 2d)

On obtient : 

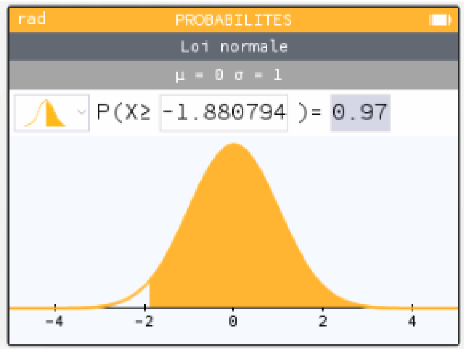
De là il faudra retrancher 20 aux deux valeurs pour pouvoir trouver la valeur de $a \approx 2{,}07286676$

Exercice 5

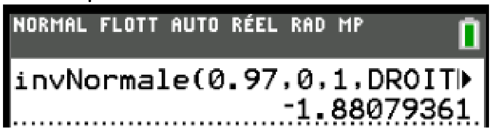
Une entreprise fabrique des sacs en plastique pour les enseignes de distribution. Elle s'intéresse au poids maximal que ces sacs peuvent supporter sans se déchirer.

On suppose ici que le poids maximal que ces sacs peuvent supporter suit une loi normale d'espérance mathématique 58 Kg et d'écart- type 3 Kg.

1. Sur 200 sacs reçus, une grande enseigne de distribution constate un poids moyen de 57,7 Kg.
   1.1. Donner un intervalle de confiance bilatéral de la moyenne des poids sur un échantillon de taille 200, au seuil de risque 1 %.
   $$P\left(\mu - t\frac{\sigma}{\sqrt{n}} \leq \bar{x} \leq \mu + t\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$
   Au seuil de 1% on aura $t \approx 2{,}58$ on sait aussi que $\mu = 58$, $\sigma = 3$ et $n = 200$
   $$P\left(58 - 2{,}58\frac{3}{\sqrt{200}} \leq \bar{x} \leq 58 + 2{,}58\frac{3}{\sqrt{200}}\right) = 0{,}99$$
   On est donc sûr) 99% d'avoir une moyenne entre 57,45 et 58,54
   1.2. Quelle est votre conclusion sur le poids moyen constaté ?
   On a la moyenne empirique qui est bien comprise entre les deux bornes et donc les poids sont dans la zone limite.



2. Donner le poids moyen dépassé dans 97 % des cas, sur un échantillon de taille 200.
   Il n'y aura que 3% des valeurs qui seront inférieures au poids recherché
   Avec la calculatrice TI 83, je cherche la borne t associée telle qu'on ait une probabilité de 97% d'être à droite de cette valeur



Ainsi $P\left(58 - 1{,}881\frac{3}{\sqrt{200}} \leq \bar{x}\right) = 0{,}97$ la borne droite dans notre situation est $58 - 1{,}881\frac{3}{\sqrt{200}} \approx 57{,}6$

IV - Estimation de la fréquence par intervalle de confiance

Si $n \geq 30$ on peut montrer comme pour les moyennes qu'un intervalle de confiance de la fréquence *p* au seuil de α est :

$$\left[f - t\sqrt{\frac{f(1-f)}{n-1}} ; f + t\sqrt{\frac{f(1-f)}{n-1}}\right]$$

Un exemple

Un sondage dans un lycée relève que sur les 500 personnes interrogées, 42% sont mécontentes du menu de la cantine.

Déterminez, au seuil de risque de 1%, un intervalle de confiance du pourcentage p de personnes mécontentes dans le lycée.

$f = 0{,}42$, $n = 500$, $\alpha = 0{,}01$ donc $t =$

Un intervalle de confiance de *p* à 1% est :

$$\left[f - t_{1\%}\sqrt{\frac{f(1-f)}{n-1}} ; f + t_{1\%}\sqrt{\frac{f(1-f)}{n-1}}\right] = \left[0{,}42 - 2{,}576\sqrt{\frac{0{,}42(1-0{,}42)}{500-1}} ; 0{,}42 + 2{,}576\sqrt{\frac{0{,}42(1-0{,}42)}{500-1}}\right] \approx [0{,}363; 0{,}477]$$

Exercice 6
On considère une grande quantité de pièces devant être livrées à une chaîne d'hypermarchés. On considère un échantillon de 100 pièces prélevées au hasard dans cette livraison. La livraison est assez importante pour que l'on puisse assimiler ce tirage à un tirage avec remise.
On constate que 96 pièces sont sans défaut.
1. Donner une estimation ponctuelle de la fréquence inconnue *p* des pièces de cette livraison qui sont sans aucun défaut.
2. Soit F la variable aléatoire qui, à tout échantillon de 100 pièces prélevées au hasard et avec remise dans cette livraison, associe la fréquence des pièces de cet échantillon qui sont sans défaut.
On suppose que F suit la loi normale de moyenne *p* et d'écart type $\sqrt{\frac{p(1-p)}{100}}$, où *p* est la fréquence inconnue des pièces de la livraison qui sont sans aucun défaut.
Déterminer un intervalle de confiance de la fréquence *p* avec le coefficient de confiance de 95%.

Solution
1. une estimation de la fréquence de défaut dans la population complète est la fréquence de l'échantillon $f \approx \frac{96}{100}$
2. $\left[f - t_{5\%}\sqrt{\frac{f(1-f)}{n-1}}; f + t_{5\%}\sqrt{\frac{f(1-f)}{n-1}}\right] = \left[0{,}96 - 1{,}96\sqrt{\frac{0{,}96(1-0{,}96)}{100-1}}; 0{,}96 + 1{,}96\sqrt{\frac{0{,}96(1-0{,}96)}{100-1}}\right] \approx [0{,}921; 0{,}999]$

Exercice 7
Dans le cadre d'une étude sur la santé au travail, des chercheurs ont interrogé au hasard 500 salariés de différents secteurs et de différentes régions de France. 145 d'entre eux déclarent avoir déjà subi un harcèlement moral au travail.
1) Identifier la population, la variable, son type et son/ses paramètre(s).
Réponse : - Population : Salariés en France.
- Variable qualitative X = "a déjà subi un harcèlement moral".
- 1 paramètre : proportion de la modalité "oui".
2) Donner une estimation ponctuelle de la proportion de salariés ayant déjà subi un harcèlement moral au travail.
On estime p par la fréquence observée de "oui" $f = \frac{145}{500} = 0{,}29$.
2.3) Donner une estimation de cette proportion par un intervalle de confiance à 90 %.
On a tout d'abord $n = 500 \geq 30$. L'estimation par intervalle à 90 % est donnée par :

$$IC_{0,90}(p) = \left[f - t_{10\%}\sqrt{\frac{f(1-f)}{n-1}}; f + t_{10\%}\sqrt{\frac{f(1-f)}{n-1}}\right]$$

$$= \left[0{,}29 - 1{,}645\sqrt{\frac{0{,}29(1-0{,}29)}{500-1}}; 0{,}29 - 1{,}645\sqrt{\frac{0{,}29(1-0{,}29)}{500-1}}\right] \approx [0{,}257; 0{,}323]$$

2.4) Si avec les mêmes données on calculait un intervalle de confiance à 95 %, serait-il plus grand ou plus petit que celui trouvé à la question précédente ? (justifier sans calcul.)
Dans l'intervalle obtenu à la question précédente, il suffirait de remplacer $t_{10\%} = 1.645$ par $t_{5\%} = 1{,}96$ qui est plus grand. On obtiendrait donc un intervalle plus grand.

V - Test de validité d'hypothèse

Exercice 8
Une entreprise fabrique en grande quantité des tiges en plastique de longueur théorique 100 mm. Un client reçoit un lot important de tiges de ce type. Il veut vérifier que la moyenne µ de l'ensemble des longueurs, en mm, des tiges constituant ce lot est égale à la longueur théorique.
On note L la variable aléatoire qui, à chaque tige prélevée au hasard dans le lot, associe sa longueur en mm. La variable aléatoire L suit la loi normale de moyenne inconnue µ et d'écart-type 0,16.
On désigne par L la variable aléatoire qui, à chaque échantillon aléatoire de 90 tiges prélevé dans le lot, associe-la moyenne des longueurs de ces tiges (le lot est assez important pour que l'on puisse assimiler ces prélèvements à des tirages avec remise). L suit la loi normale de moyenne µ et d'écart-type $\sigma = \frac{0{,}16}{\sqrt{90}} \approx 0{,}017$.

Le client construit un test d'hypothèse :

- L'hypothèse nulle est $H_0 : \mu = 100$.
- L'hypothèse alternative est $H_1 : \mu \neq 100$.
- Le seuil de signification est fixé à 5 %.

1. Sous l'hypothèse nulle H0 déterminer le réel positif *h* tel que :

$$P(100 - h < \overline{L} < 100 + h) = 0{,}95.$$

A priori on doit avoir $h = t_{5\%} \times \sigma = 1{,}96 \times 0{,}017 = 0{,}333$

2. Énoncer la règle de décision permettant d'utiliser ce test.

Règle de décision

C'est un test bilatéral, on ne veut pas que la longueur dépasse dans un sens ou dans l'autre (trop ou trop peu)

– Si la moyenne de l'échantillon appartient à $[100 - 0{,}333; 100 + 0{,}333] = [99{,}667; 100{,}333]$, on accepte H_0 ;

– sinon, on rejette H_0

Le seuil de 5% est la probabilité de rejeter H_0 alors que H_0 est vraie.

3. Le client prélève un échantillon aléatoire de 90 tiges dans la livraison et il constate que la moyenne des longueurs de l'échantillon est de 100,04mm. Le client estime que le fournisseur n'a pas respecté ses engagements et renvoie tout le lot.

Le client a-t-il raison ? Justifier votre réponse.

Ici, on a mesuré une moyenne de 100,004 sur l'échantillon. On a $100{,}004 \in [99{,}667; 100{,}333]$. On rejette donc l'hypothèse H_0. Au seuil de 5%, on considère que le stock entier tige a la moyenne annoncée par le fabricant de 100 et on accepte la livraison.

Exercice 9

Une entreprise spécialisée produit des boules de forme sphérique en grande série.

Le responsable de la qualité cherche à analyser la production. Il mesure pour cela le diamètre des boules d'un échantillon (E) de 50 pièces. La moyenne obtenue sur l'échantillon (E) 72,80 amène à se poser la question : « Le diamètre moyen *m* des boules fabriquées est-il strictement inférieur à 73mm? ».

Pour cela, on construit un test d'hypothèse au risque de 5%.

L'hypothèse nulle H_0 est : $m = 73$;

L'hypothèse alternative H_1 est : $m < 73$.

On admet que la variable aléatoire D, qui mesure le diamètre moyen sur un échantillon de 50 boules prélevées au hasard et avec remise, suit une loi normale de moyenne 73 et d'écart type $\frac{0{,}2}{\sqrt{50}}$.

1. Calculer le nombre réel *a* tel que $p(\overline{D} > 73 - a) = 0{,}95$.

On veut éliminer les 5 premiers pourcents, donc on n'utilisera pas $t_{5\%}$ qui élimine $\frac{5}{2} = 2{,}5\%$ sur les extrémités mais plutôt $t_{10\%}$ qui élimine 5% d'un côté (et 5% de l'autre… mais on n'utilisera pas l'autre)

Ici $a = t_{10\%}\sigma \approx 1{,}645 \times \frac{0{,}2}{\sqrt{50}} \approx 0{,}0465$

2. Énoncer la règle de décision du test.

Règle de décision

C'est un test unilatéral, on ne veut pas être en dessous d'une valeur butoir mais si on a une valeur trop grande ça n'est pas un souci.

– Si la moyenne de l'échantillon appartient à $[73 - 0{,}0465; +\infty[= [72{,}9535; +\infty[$, on accepte H_0 ;

– sinon, on rejette H_0

Le seuil de 5% est la probabilité de rejeter H_0 alors que H_0 est vraie.

3. Au risque de 5% et au vu de l'échantillon (E), que peut-on conclure ?

On a une moyenne de 72,80 , or $72{,}85 \notin [72{,}9535; +\infty[$ donc on rejette H_0 au profit de H_1, on peut dire que le diamètre des sphères est inférieur à 73 dans la population.

Exercice 10

Une entreprise fabrique, en grande quantité, des tiges métalliques cylindriques pour l'industrie. Leur longueur et leur diamètre sont exprimés en millimètres.

Dans cette question on s'intéresse au diamètre des tiges, exprimé en millimètres.

On prélève au hasard et avec remise un échantillon de 50 tiges dans la production d'une journée.

Soit $\overline{D}$ la variable aléatoire qui, à tout échantillon de 50 tiges prélevées au hasard et avec remise dans la production d'une journée, associe la moyenne des diamètres des tiges de cet échantillon.

On suppose que $\overline{D}$ suit la loi normale de moyenne inconnue µ et d'écart type $\frac{\sigma}{\sqrt{50}}$ avec $\sigma = 0{,}19$.

Pour l'échantillon prélevé, la moyenne obtenue, arrondie à 10^{-2} est $\bar{x} = 9{,}99$.

1. À partir des informations portant sur cet échantillon, donner une estimation ponctuelle de la moyenne µ des diamètres des tiges produites dans cette journée.

$$\mu \approx \bar{x} = 9{,}99$$

2. Déterminer un intervalle de confiance centré sur $\bar{x}$ de la moyenne µ des diamètres des tiges produites pendant la journée considérée, avec le coefficient de confiance 95 %.

$$I_{95\%} = \left[\bar{x} - t_{5\%}\frac{\sigma}{\sqrt{n}} ; \bar{x} + t_{5\%}\frac{\sigma}{\sqrt{n}}\right] = \left[9{,}99 - 1{,}96\frac{0{,}19}{\sqrt{50}} ; 9{,}99 - 1{,}96\frac{0{,}19}{\sqrt{50}}\right] \approx [9{,}937 ; 10{,}043]$$

3. On considère l'affirmation suivante : « la moyenne µ est obligatoirement dans l'intervalle de confiance obtenu à la question 2 ».Est-elle vraie ?

   Non, ça n'a rien d'obligatoire, $I_{95\%}$ s'appelle intervalle de confiance au seuil de 95% , ce qui veut dire qu'on n'est sûr qu'à 95% que μ soit à l'intérieur, et non 100% comme affirmé dans la question.

Exercice 11

La machine se déréglant dans le temps, on veut tester la moyenne *m* des longueurs des barres produites par la machine.

On se demande si on peut accepter, au seuil de risque de 5 %, l'hypothèse selon laquelle la moyenne *m* des longueurs des barres est encore de 92,50 cm.

Pour cela, on construit un test d'hypothèse bilatéral.

On suppose que la variable aléatoire X, qui à tout échantillon de 30 barres de métal prélevées au hasard associe la moyenne des longueurs en centimètres des barres de l'échantillon, suit une loi normale de moyenne *m* et d'écart type 0,03.

On choisit l'hypothèse nulle H_0 : *m* = 92,50.

1. Donner l'hypothèse alternative H_1.

   Le test étant bilatéral on ne veut dépasser dans un sens ou dans l'autre, ainsi H_1 : *m* ≠ 92,50

2. Sous l'hypothèse H_0, calculer le réel *h* tel que $P(92{,}5 - h < X < 92{,}5 + h) = 0{,}95$.

   On sait que l'on aura : $P\left(m - t_{5\%}\frac{\sigma}{\sqrt{n}} \leq \bar{x} \leq m + t_{5\%}\frac{\sigma}{\sqrt{n}}\right) = 0{,}95$

   $\Leftrightarrow P\left(92{,}5 - 1{,}96\frac{0{,}03}{\sqrt{30}} \leq \bar{x} \leq 92{,}5 - 1{,}96\frac{0{,}03}{\sqrt{30}}\right) = 0{,}95$

   Ainsi $h = 1{,}96\frac{0{,}03}{\sqrt{30}} \approx 0{,}01057$

3. Énoncer la règle de décision du test.

   Le seuil de 5% est la probabilité de rejeter H_0 alors que H_0 est vraie. C'est un test bilatéral, on ne veut pas que la longueur dépasse dans un sens ou dans l'autre (trop ou trop peu)

   – Si la moyenne de l'échantillon appartient à $[92{,}5 - 0{,}01057 ; 92{,}5 - 0{,}01057] = [12{,}4894 ; 12{,}51057]$, on ne rejette pas H_0 ;

   – sinon, on rejette H_0 et accepte H_1

4. On prélève un échantillon de 30 barres extraites au hasard dans la production de la machine, on obtient les résultats suivants :

Longueurs (en cm)	92,1	92,2	92,3	92,4	92,5	92,6	92,7	92,8	92,9
Nombre de barres	3	2	6	5	5	3	2	2	2

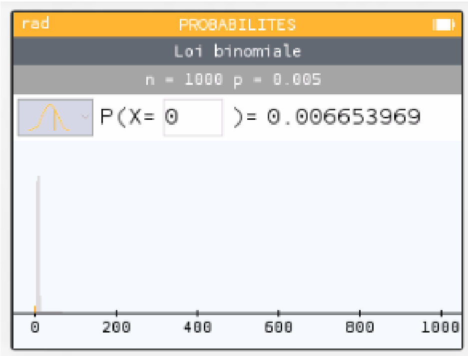
   Au vu des résultats de cet échantillon, peut-on admettre au seuil de risque de 5 %, l'hypothèse selon laquelle la moyenne *m* des longueurs des barres est encore de 92,50 cm?

   La calculatrice nous donne $\bar{x} \approx 92{,}45333 \notin [12{,}4894 ; 12{,}51057]$ donc H_0 est rejetée, on ne peut accepter que la longueur des barres produites par la machine soit de 92,5 et ce, au seuil de risque de 5%.

Bonus :
Sujet d'Annales

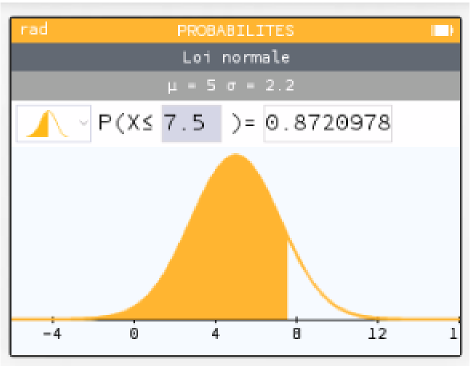
Groupe B : 9 mai 2017
Exercice 2
B loi binomiale => loi normale

1) On répète n=1000 fois de manière identique et indépendante la même épreuve de Bernoulli dont le succès « la bille présente un défaut de fabrication » a pour probabilité $p = 0{,}005$. X la variable aléatoire qui comptera le nombre de succès au cour de cette expérience suivra donc un loi Binomiale de paramètre $n = 1000$ et $p = 0{,}005$.
2) .

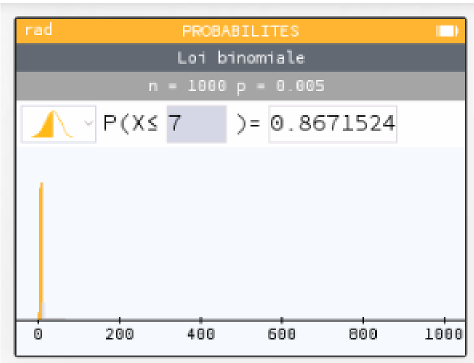


   a. $P(X = 0) \approx 0{,}00665$
   b. $P(X \geq 1) = 1 - P(X = 0) \approx 1 - 0{,}00665 \approx 0{,}99335$

3a)
Si $n \geq 30$ la binomiale peut être approchée à l'aide d'une loi normale ayant la même espérance et le même écart type.
Si X suit une loi $B(n\,;p)$ le cours nous donne $E(X) = np$ et $\sigma = \sqrt{np(1-p)}$
Ici on aura $E(X) = np = 1000 \times 0{,}005 = 5$ et $\sigma = \sqrt{np(1-p)} = \sqrt{1000 \times 0{,}005 \times (1 - 0{,}005)} \approx 2{,}2$
3b)



Si on avait voulu faire le calcul de manière directe on aurait eu :



Ça c'est la vraie valeur , et donc on voit que l'approximation est trop élevée d'à peu près 0,005.