



- 1 Exemple d'utilisation de l'outil Statistiques à une variable
- 2 Les données sources
- 3 Présentation de la fenêtre Analyse des données
- 4 La zone de graphique
- 5 Histogramme
- 6 Diagramme en barres
- 7 Boîte à moustaches
- 8 Nuage de points
- 9 Arbre - Tige - Feuilles
- 10 Quantile loi normale
- 11 Calculs et tests statistiques

http://url.univ-irem.fr/f335



La vue **Tableur** intègre un ensemble d'outils permettant d'effectuer une analyse statistique de base sur les valeurs contenues dans une ou plusieurs plages de cellules. GeoGebra permet de gérer les séries statistiques brutes mais aussi les séries statistiques classées en modalités-effectifs ou encore réparties en classe. Le logiciel offre également le choix entre différents modes de représentation graphique de ces données.

Dans cette fiche, nous nous attarderons sur les outils relatifs aux statistiques à une variable.

1 Exemple d'utilisation de l'outil Statistiques à une variable

On considère la série des notes (sur 20) obtenues par 40 étudiants à un examen.

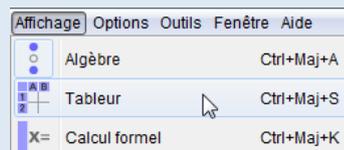
Note	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Effectif	1	0	0	1	2	0	3	2	5	2	2	7	3	2	1	2	1	4	0	2	0

Il s'agit d'une série statistique dont le caractère est quantitatif discret.

La première étape consiste à saisir les données dans le tableur intégré à GeoGebra.

Méthode

- Utiliser le menu Affichage ► Tableur pour afficher la vue **Tableur**.



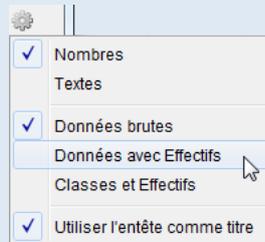
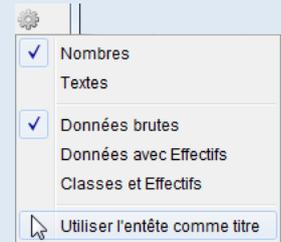
- Dans la colonne A saisir les notes.
- Dans la colonne B saisir les effectifs correspondants.

Tableur			
	A	B	
1	Note	Effectif	
2	0	1	
3	1	0	
4	2	0	
5	3	1	
6	4	0	
7	5	2	
8	6	0	
9	7	3	
10	8	2	
11	9	5	
12	10	2	
13	11	7	
14	12	3	
15	13	2	
16	14	1	
17	15	2	
18	16	1	
19	17	4	
20	18	0	
21	19	2	
22	20	0	

En second lieu, il s'agit d'indiquer à GeoGebra la zone contenant les données à analyser ainsi que leur nature.

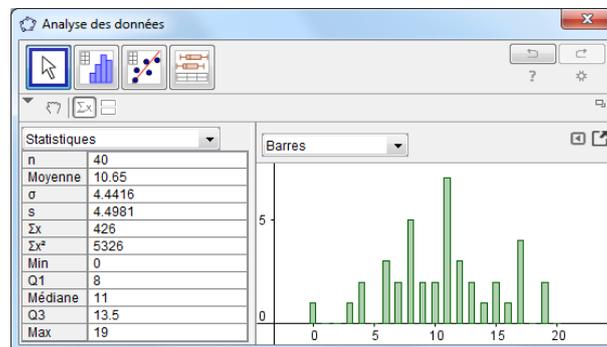
Méthode

- Sélectionner la plage de cellules A1 : A22.
- Cliquer sur l'icône .
- Dans la boîte de dialogue **Source des données**, cliquer sur l'outil **Options**  puis choisir Utiliser l'entête comme titre .
- Cliquer à nouveau sur  puis choisir Données avec Effectifs : une seconde colonne apparaît dans la boîte de dialogue **Source des données**.



- Pour remplir la seconde colonne, sélectionner la plage de cellules B1 : B22 puis cliquer sur l'icône  située dans la colonne **Effectifs**.
- Cliquer sur le bouton .

La fenêtre **Analyse de données** qui s'affiche alors permet à l'utilisateur de connaître un certain nombre de caractéristiques de la série étudiée et de la représenter graphiquement.

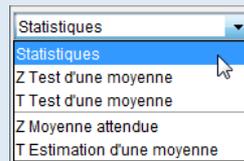


Méthode

- Cliquer sur l'icône  pour afficher la barre d'outils.

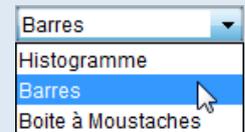


-  permet de faire apparaître la fenêtre **Source des données** et de modifier la sélection des données;
- Σx permet d'afficher de nombreuses caractéristiques de la série et d'effectuer des tests statistiques;



Statistiques	
n	40
Moyenne	10.65
σ	4.4416
s	4.4981
Σx	426
Σx^2	5326
Min	0
Q1	8
Médiane	11
Q3	13.5
Max	19

-  permet d'afficher un second graphique représentatif de la série en cours.
- Plusieurs manipulations sont possibles dans la zone du graphique :
 - utiliser la liste déroulante pour sélectionner le type de graphique ;



- cliquer sur le bouton  pour modifier les options du graphique ;
- cliquer sur la bouton  pour exporter le graphique vers le presse-papier ou vers la vue **Graphique** ou bien en tant que fichier.

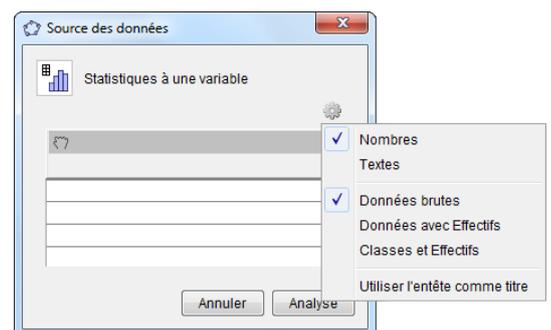


[Ouvrir le fichier exemple](#) 

2 Les données sources

Une fois les données à analyser préalablement entrées dans le tableur de GeoGebra, il convient, en règle générale, de les sélectionner et d'ouvrir la boîte de dialogue **Source des données** à l'aide de l'outil .

Les options alors accessibles depuis cette boîte de dialogue permettent de spécifier le type de variable.



Pour analyser les données brutes d'une variable (quantitative ou qualitative) discrète :

Méthode

- Sélectionner les cellules contenant les valeurs de la série.
- Cliquer sur l'icône .
- Cliquer sur le bouton  :

- si le caractère de la série est quantitatif, cliquer sur l'item Nombres ;
- si le caractère de la série est qualitatif, cliquer sur l'item Textes ;
- si un entête est présent, cliquer sur l'item Utiliser l'entête comme titre ;
- vérifier que l'item Données brutes est coché.

<input checked="" type="checkbox"/>	Nombres
	Textes
<input checked="" type="checkbox"/>	Données brutes
	Données avec Effectifs
	Classes et Effectifs
	Utiliser l'entête comme titre

- Au besoin, sélectionner de nouvelles valeurs dans la feuille de calcul et cliquer sur le bouton  pour ajouter ces valeurs aux données présélectionnées.
- Valider la saisie en cliquant sur le bouton .

Pour analyser les données issues d'une variable discrète dont les différentes valeurs sont regroupées par effectifs :

Méthode

- Sélectionner les cellules contenant les différentes modalités de la série.

- Cliquer sur l'icône .

- Cliquer sur le bouton  :

- si le caractère de la série est quantitatif, cliquer sur l'item Nombres ;
- si le caractère de la série est qualitatif, cliquer sur l'item Textes ;
- si un entête est présent, cliquer sur l'item Utiliser l'entête comme titre ;
- cliquer sur l'item Données avec Effectifs .

	Nombres
<input checked="" type="checkbox"/>	Textes
	Données brutes
<input checked="" type="checkbox"/>	Données avec Effectifs
	Classes et Effectifs
<input checked="" type="checkbox"/>	Utiliser l'entête comme titre

- Dans la vue **Tableur**, sélectionner les cellules contenant les effectifs puis cliquer sur le bouton  de la colonne **Effectifs**.

- Valider la saisie en cliquant sur le bouton .

Pour que le graphique s'affiche sans aberration, il faut que la saisie des données soit cohérente :

Remarque :

- chaque modalité doit se voir attribuer un effectif;
- il ne doit pas y avoir de modalité en double;
- dans le cas d'un caractère quantitatif, il est préférable que les modalités soient ordonnées (mais GeoGebra les réordonne automatiquement).

Pour analyser les données issues d'une variable continue dont les différentes valeurs sont regroupées par classes (les classes ont nécessairement la même amplitude) :

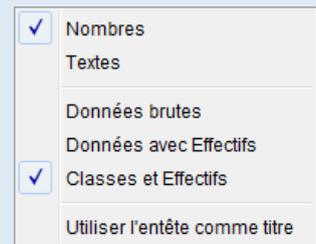
Méthode

- Sélectionner au moins une cellule non vide.

- Cliquer sur l'icône .

- Cliquer sur le bouton  :

- vérifier que l'item Nombres est coché;
- vérifier que l'item Utiliser l'entête comme titre est décoché;
- cliquer sur l'item Classes et Effectifs .

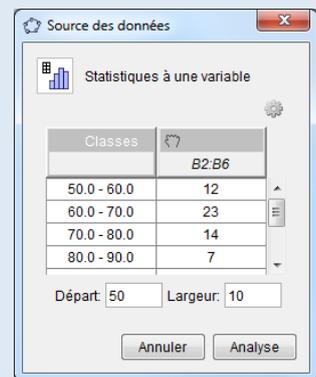


- Dans la vue **Tableur**, sélectionner les cellules contenant les effectifs puis cliquer sur le bouton de la colonne **Effectifs**.

- La colonne **Classes** est automatiquement remplie. Par défaut, la borne inférieure de la première classe est 0 et l'amplitude des classes vaut 1.

Pour modifier les classes :

- modifier la valeur de la borne inférieure de la première classe dans le champ **Départ** ;
- modifier l'amplitude des classes dans le champ **Largeur**.



- Valider la saisie en cliquant sur le bouton .

3 Présentation de la fenêtre Analyse des données

La fenêtre **Analyse des données** est divisée en plusieurs zones.

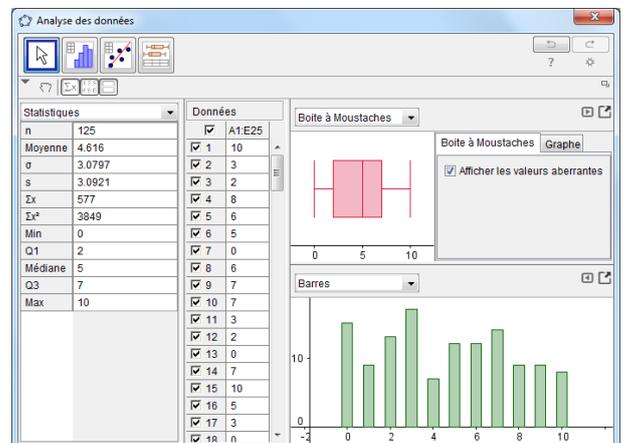


La barre principale propose les outils suivants :

- : l'outil Statistiques à une variable pour saisir des données issues d'une série statistique à une variable ;
- : l'outil Statistiques à deux variables pour saisir des données issues de séries statistiques à deux variables ;
- : l'outil Statistiques à plusieurs variables pour saisir des données issues de plusieurs séries statistiques et les comparer.



La barre d'outils (qui se déploie en cliquant sur le bouton) propose les outils suivants :



-  : l'outil Afficher la source de données pour afficher la fenêtre **Source des données** et modifier les données à analyser;
- Σx : l'outil Afficher statistiques pour afficher dans la vue les calculs statistiques concernant les données;
- $\frac{123}{345}$: l'outil Afficher données pour afficher dans la vue toutes les valeurs de la série (cet outil ne s'affiche que lorsque les données brutes sont disponibles) et il est alors possible, en décochant certaines valeurs, de les ignorer lors de l'analyse statistique;
-  : l'outil Afficher 2d graphique pour afficher un second graphique dans la vue.

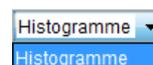
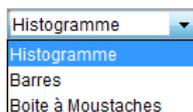
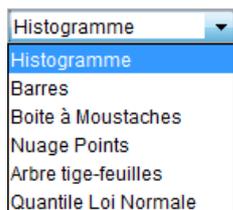
Dans la barre d'outils, le bouton  permet d'ancrer la vue **Analyse des données** dans la fenêtre principale de GeoGebra.

La zone de graphique ainsi que l'outil Afficher statistiques sont étudiés dans les paragraphes suivants.

4 La zone de graphique

En fonction du type de données sources choisi, GeoGebra propose différents modes de représentation de la série statistique à sélectionner dans la liste déroulante :

- Pour des données brutes :
- Pour des données avec effectif :
- Pour des classes et effectifs :



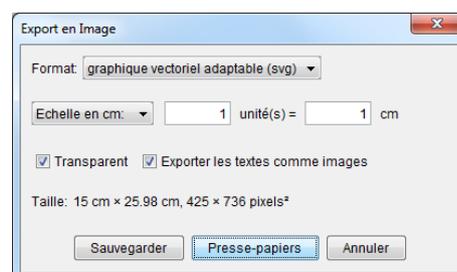
À l'exception du type de graphique Arbre-Tige-Feuilles, au moins deux outils sont toujours disponibles :

-  : cet outil permet de paramétrer le graphique;
-  : cet outil ouvre un menu déroulant qui permet d'exporter le graphique :
 - vers la vue **Graphique** active;
 - vers le presse-papier afin de coller le graphique dans une autre application en tant qu'image;
 - vers un fichier image.



Remarque :

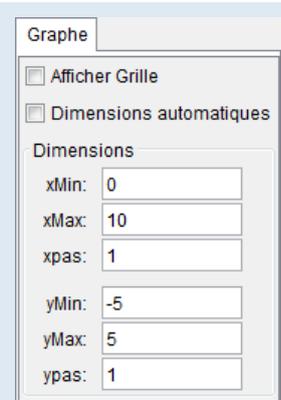
- L'export du graphique vers la vue **Graphique** crée un nouvel objet défini à l'aide d'une commande **Barres**, **BoiteMoustaches**, **DiagrammeBâtons**, ...
- Une fois le graphique copié dans la vue **Graphique**, l'export vers un fichier image (Fichier ► Exporter ►  Graphique en tant qu'image (png, eps) ...) permet également d'exporter vers le presse-papier tout en réglant plus finement certains paramètres comme l'échelle de la figure en particulier.



Le repère de la fenêtre graphique est calculé automatiquement et s'adapte au type de représentation graphique choisi ainsi qu'aux dimensions de la fenêtre **Analyse des données**. Il reste cependant possible de définir le repère manuellement.

Méthode

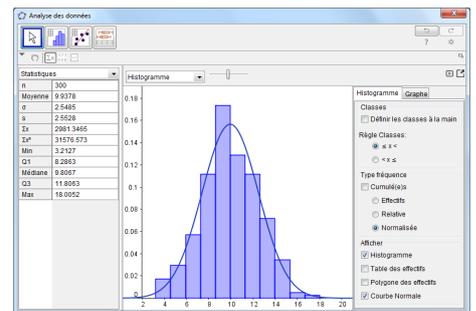
- Cliquer sur le bouton .
- Ouvrir l'onglet **Graphe**.
- Cocher ou décocher la case **Afficher Grille** selon les besoins.
- Décocher la case **Dimensions automatiques** pour modifier les paramètres du repère : les bornes des abscisses et des ordonnées sont modifiables par les champs **xMin**, **xMax**, **yMin** et **yMax**, tandis que les graduations des axes sont modifiables par les champs **xpas** et **ypas**.



5 Histogramme

Dans le cadre de ce paragraphe consacré aux histogrammes, nous étudierons les modifications permises lorsque l'utilisateur a sélectionné Données brutes dans la boîte de dialogue **Source de données**, ce qui nous permettra d'aborder l'ensemble des réglages possibles (les réglages accessibles depuis Données avec Effectifs ou Classes et Effectifs sont moins nombreux).

Par défaut, GeoGebra affiche un histogramme dans la zone de graphique. Les classes et le mode de calcul des effectifs sont modifiables.



Pour modifier le nombre de classes :

Méthode

- Déplacer le curseur  situé à droite de la liste déroulante.

L'amplitude des classes est alors automatiquement recalculée et le graphique actualisé (GeoGebra gère lui-même le nombre minimal et maximal de classes).

Pour modifier les classes manuellement :

Méthode

- Cliquer sur le bouton .
- Dans l'onglet **Histogramme**, rubrique **Classes**, cocher **Définir les classes à la main**.

Le curseur  cède alors la place à deux champs de saisie.

Départ Largeur

- Dans le champ **Départ**, saisir la borne inférieure de la première classe (GeoGebra ne génère aucun histogramme si la valeur saisie ici est supérieure à la valeur minimale des données brutes).
- Dans le champ **Largeur**, saisir l'amplitude désirée pour les classes.

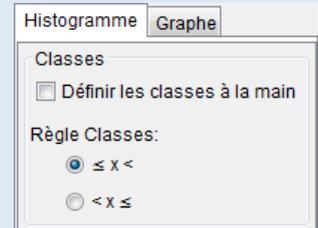


Par défaut, dans un histogramme, les bornes inférieures des classes sont fermées tandis que les bornes supérieures sont ouvertes. Il est possible d'inverser ce choix.

Pour modifier les règles de calculs sur les classes :

Méthode

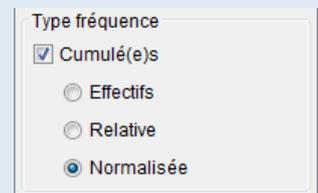
- Cliquer sur le bouton .
- Dans l'onglet **Histogramme**, rubrique **Classes**, cocher dans **Règle Classes** :
 - $\leq x <$ pour une borne inférieure fermée et une borne supérieure ouverte ;
 - $< x \leq$ pour une borne inférieure ouverte et une borne supérieure fermée ;



Pour modifier le type d'histogramme :

Méthode

- Cliquer sur le bouton .
- La rubrique **Type de fréquence** de l'onglet **Histogramme** permet de régler les valeurs en ordonnées de l'histogramme .
- Les boutons radio offrent trois choix possibles :
 - **Effectifs** permet la création d'un histogramme des effectifs : les hauteurs des rectangles représentent le nombre de valeurs de la série statistique se trouvant dans chaque classe ;
 - **Relative** permet la création d'un histogramme des fréquences relatives : les hauteurs des rectangles sont données par le quotient de l'effectif de la classe divisé par l'effectif total (ainsi la somme des fréquences vaut toujours 1) ;
 - **Normalisée** permet la création d'un histogramme des fréquences normalisées : les hauteurs des rectangles sont modifiées de telle sorte que l'aire de l'histogramme soit toujours égale à 1. Ceci permet de donner une estimation de la densité à condition que la série statistique soit supposée provenir d'une variable à caractère continu.
- Si la case **Cumulé(e)s** est cochée, on obtient un histogramme des fréquences cumulées. Les significations des trois boutons radio sont alors les suivantes :
 - **Effectifs** : les hauteurs des rectangles représentent les effectifs cumulés croissants de la série statistique se trouvant dans chaque classe ;
 - **Relative** : en ordonnées, les hauteurs des rectangles s'obtiennent par le calcul du quotient de l'effectif cumulé croissant de la classe divisé par l'effectif total ;
 - **Normalisée** : l'histogramme obtenu donnera une estimation de la fonction de répartition. Si les valeurs de la série statistique étudiée sont considérées comme des réalisations indépendantes d'une seule variable aléatoire X et si on appelle F la fonction de répartition de la variable X , l'utilisation de cette option permet d'avoir les propriétés suivantes vérifiées :

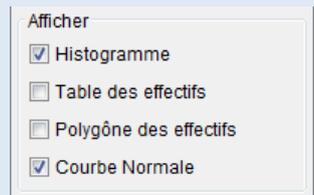


$$\lim_{x \rightarrow -\infty} F(x) = 0 \quad \text{et} \quad \lim_{x \rightarrow +\infty} F(x) = 1$$

Pour modifier l'affichage de l'histogramme :

Méthode

- Cliquer sur le bouton .
- Dans l'onglet **Histogramme**, rubrique **Afficher**, cocher ou décocher les cases selon les besoins :
 - **Histogramme** permet d'afficher ou non l'histogramme avec les différents réglages déjà effectués ;
 - **Table des effectifs** affiche ou masque un tableau avec les différentes classes ainsi que les valeurs associées à chaque classe (en fonction du choix effectué dans la rubrique **Type de fréquence**) ;
 - **Polygone des effectifs** affiche ou masque le polygone des effectifs ou des fréquences, cumulé(e)s ou non ;
 - **Courbe Normale** permet de représenter la fonction de densité de la loi normale de paramètres la moyenne et l'écart-type de la série statistique (cette option n'est accessible que si la case **Normalisée** de la rubrique **Type de fréquence** est cochée). GeoGebra représente donc la fonction f ainsi définie :



$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

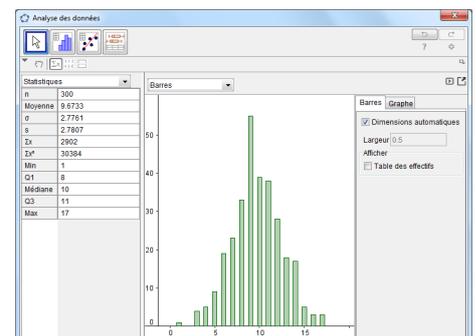
où σ est l'écart-type de la série statistique et μ la moyenne des valeurs.

Remarque :

Dans le cas où l'on affiche le polygone des effectifs de la série et où la case **Normalisée** de la rubrique **Type de fréquence** est cochée, alors le polygone des effectifs donnera une estimation de la densité associée aux données. En effet, l'aire sous la ligne brisée est égale aussi à 1.

6 Diagramme en barres

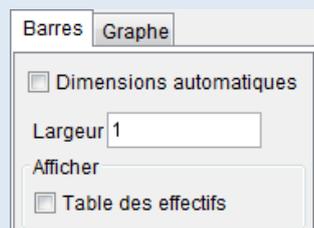
Lorsque le caractère de la série est discret et qu'il n'est pas nécessaire de recourir à un regroupement en classes, la représentation graphique de la série par un diagramme en barres est plus adaptée.



Pour modifier la largeur des barres :

Méthode

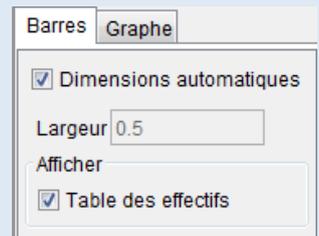
- Cliquer sur le bouton .
- Dans l'onglet **Barres**, décocher la case **Dimensions automatiques**.
- Dans le champ **Largeur** saisir la largeur désirée pour les barres.



Pour afficher la table des effectifs contenant les modalités et les effectifs :

Méthode

- Cliquer sur le bouton .
- Dans l'onglet **Barres**, rubrique **Afficher**, cocher la case **Table des effectifs**.

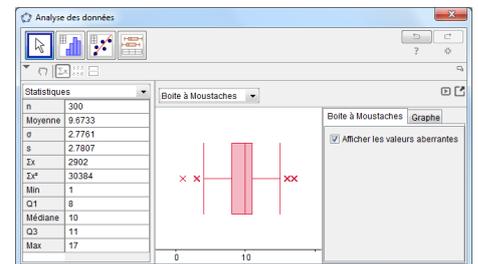


7 Boîte à moustaches

Pour résumer visuellement quelques caractéristiques de position du caractère étudié, le diagramme en boîte (appelé encore boîte à moustaches ou bien boîte de TUKEY) semble particulièrement adapté.

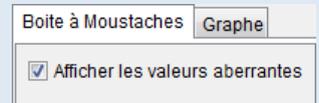
Si on appelle EI l'écart interquartile ($EI = Q_3 - Q_1$), GeoGebra considère qu'une valeur est aberrante lorsqu'elle est plus petite que $Q_1 - 1,5 \times EI$ ou lorsqu'elle est plus grande que $Q_3 + 1,5 \times EI$.

On peut alors demander au logiciel de représenter la boîte à moustaches en tenant compte ou non des valeurs aberrantes.



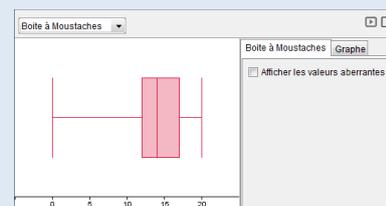
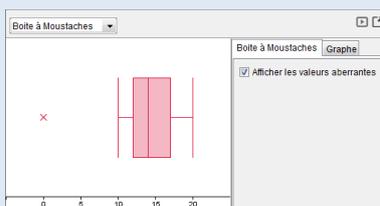
Méthode

- Cliquer sur le bouton .
- Dans l'onglet **Boîtes à moustaches**, cocher ou décocher la case **Afficher les valeurs aberrantes**.



Avec affichage des valeurs aberrantes

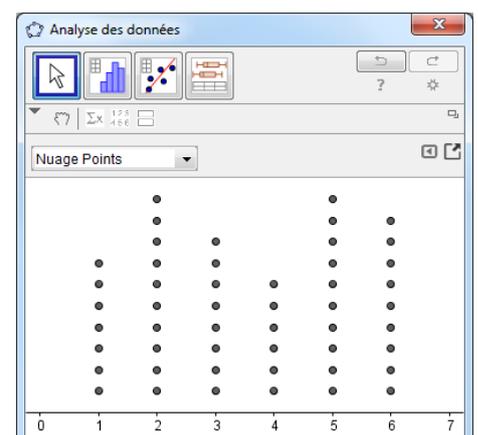
Sans affichage des valeurs aberrantes



8 Nuage de points

Dans une représentation en nuage de points, les points ont pour abscisse les différentes modalités de la série statistique et pour ordonnée des valeurs entières qui représentent le nombre d'apparitions d'une modalité sur l'ensemble des données analysées.

La représentation graphique en nuage de points n'est disponible que pour des données brutes et n'offre la modification d'aucun paramètre.

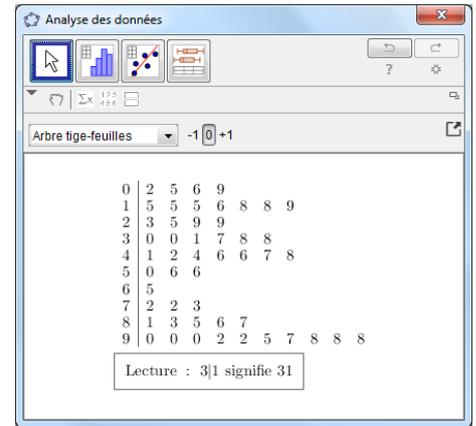


9 Arbre - Tige - Feuilles

Le type de graphique tige et feuilles est un procédé pour organiser et représenter un ensemble de données numériques. Il sert, en particulier, à organiser des données au fur et à mesure de leur collecte.

On range sur une même colonne (appelée la tige) tous les nombres ayant la même partie principale (tous les chiffres du nombre sauf le dernier) et on ventile sur les lignes (les feuilles), dans l'ordre croissant, les valeurs collectées en inscrivant uniquement le dernier chiffre.

Dans GeoGebra, le type de graphique tige et feuilles possède trois options, accessibles en cliquant sur l'un des boutons situés à droite de la liste déroulante, qui permettent de régler la précision de la tige :



- **-1** : pour avoir trois chiffres significatifs (la tige va de 1 à 99) ;
- **0** : pour avoir deux chiffres significatifs (la tige va de 1 à 9) ;
- **+1** : pour avoir un chiffre significatif (la tige vaut 0).

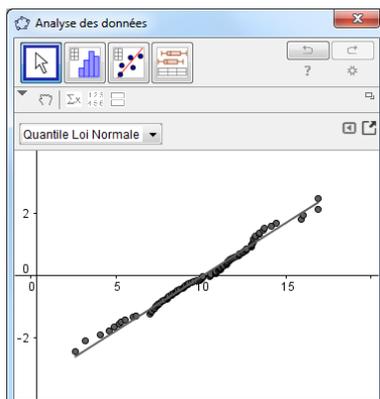
10 Quantile loi normale

Ce diagramme compare les valeurs de la série statistique avec les quantiles de la loi normale centrée-réduite. Si les données sont issues d'une loi normale, alors les points de la courbe sont parfaitement alignés sur la droite attendue (droite d'ajustement de HENRY).

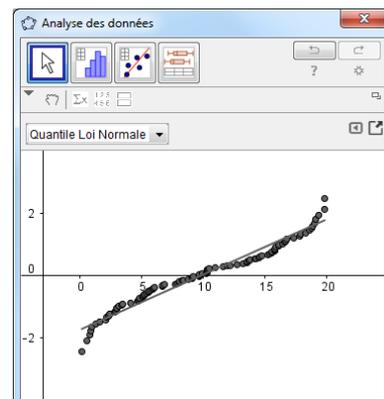
Les points du nuage de points ont pour abscisses les valeurs de la série (x_i) et pour ordonnées leur score normal attendu (score z). On suppose ici que les valeurs de la fonction de répartition empirique ($F_i = \frac{\text{Effectif} < x_i}{n}$) diffèrent peu de celles de la fonction de répartition théorique. Donc cette courbe met en relation les x_i de l'échantillon avec les quantiles théoriques obtenus à partir des valeurs de la fonction de répartition empirique. Si l'échantillon suit une loi normale de paramètres m et σ , alors le graphique obtenu doit être à peu près une droite de pente $\frac{1}{\sigma}$ et coupant l'axe des abscisses en m .

Ce type de représentation graphique ne dispose pas de paramètre modifiable dans GeoGebra.

La distribution suit une loi normale



La distribution ne suit apparemment pas une loi normale



Les tests statistiques peuvent être utilisés dans plusieurs situations :

- pour vérifier si une chaîne de production de pièces industrielles fabrique bien des pièces ayant des caractéristiques suffisamment proches par rapport aux exigences/tolérances ;
- pour vérifier si dans une même usine deux chaînes de fabrication fournissent des pièces ayant les mêmes caractéristiques à une tolérance près ;
- pour vérifier, par exemple, l'efficacité d'un nouveau traitement thérapeutique ;
- pour tester l'efficacité d'un appareil de mesure.

Lorsqu'on effectue une analyse statistique des données, on suppose que :

- les données représentent un échantillon prélevé d'une population possédant plusieurs caractéristiques connues ou non (comme la moyenne et l'écart-type) ;
- l'échantillon est représentatif des caractéristiques de la population, c'est-à-dire que les valeurs le constituant ont été choisies de manière systématique ou aléatoire ;
- cet échantillon suit une loi dont on connaît ou non tous les paramètres.

Pour une série donnée, GeoGebra fournit un ensemble de caractéristiques et de tests de moyenne d'hypothèses simples et/ou composites.

En ce qui concerne les statistiques descriptives, GeoGebra propose un certain nombre de paramètres :

Méthode

- Dans la barre d'outils de la vue **Analyse de données**, cliquer sur le bouton Σx .
- Sélectionner **Statistiques** dans la liste déroulante.
- GeoGebra propose les indicateurs suivants (de haut en bas) :

- n : l'effectif total de la série statistique (le nombre de données étudiées)
- Moyenne : la moyenne de la série statistique
- σ : l'écart-type standard de la série statistique (racine carrée de la variance)

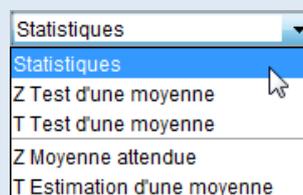
$$\sigma = \frac{\sum X^2}{n} - \left(\frac{\sum X}{n}\right)^2$$

- s : l'écart-type sans biais de la série statistique (racine carrée de la variance sans biais)

$$s = \sigma \times \sqrt{\frac{n}{n-1}}$$

Cet estimateur de l'écart-type est utilisé dans les tests de moyenne où la variance est supposée inconnue ou encore dans les intervalles de confiance.

- $\sum x$: la somme des valeurs de la série statistique
- $\sum x^2$: la somme des carrés des valeurs de la série statistique
- Min : la valeur minimale de la série statistique
- Q_1 : le premier quartile de la série statistique



n	408
Moyenne	24.6836
σ	14.627
s	14.6449
Σx	10070.9027
Σx^2	335876.9944
Min	0.0226
Q1	12.7619
Médiane	23.9881
Q3	38.4954
Max	49.9351

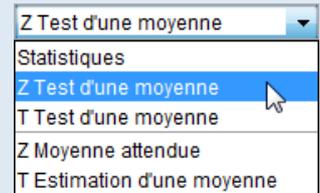
- Médiane : la médiane de la série statistique
- Q_3 : le troisième quartile de la série statistique
- Max : la valeur maximale de la série statistique

En ce qui concerne les tests de moyenne, deux cas de figures se présentent : soit on connaît l'écart-type de la population, soit on ne le connaît pas. On suppose aussi dans tous les cas que l'échantillon suit une loi normale.

Le test z (ou z test) d'une moyenne permet de rejeter ou non une moyenne hypothétique μ_0 connaissant la taille de l'échantillon et l'écart-type de la population :

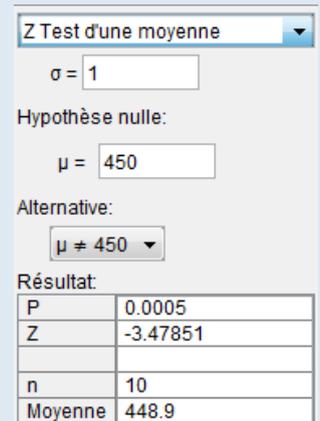
Méthode

- Dans la barre d'outils de la vue **Analyse de données**, cliquer sur le bouton Σx .
- Sélectionner **Z Test d'une moyenne** dans la liste déroulante.
- Dans le champ $\sigma =$ inscrire l'écart-type supposé connu de la série.
- Dans le champ $\mu =$ inscrire la valeur de la moyenne μ_0 à tester.



On suppose ici que la série suit une loi normale de paramètres μ_0 et $\frac{\sigma}{\sqrt{n}}$.

- La liste déroulante **Alternative** permet de sélectionner différentes hypothèses alternatives :
 - $\mu \neq \dots$: pour un test bilatéral;
 - $\mu > \dots$: pour un test unilatéral à droite;
 - $\mu < \dots$: pour un test unilatéral à gauche.
- La rubrique **Résultat** affiche le résultat du test (seule la valeur de P change) :
 - P : probabilité de rejet de l'hypothèse nulle (couramment, on rejette l'hypothèse nulle si la probabilité est inférieure à 0,05);
 - Z : la valeur de Z avec $Z = \frac{m - \mu_0}{\sigma / \sqrt{n}}$ si m désigne la moyenne de la série;
 - n : l'effectif de l'échantillon;
 - Moyenne : la moyenne de l'échantillon.



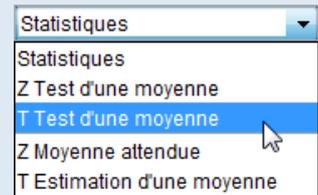
Pour un test bilatéral ($\mu \neq \dots$), il y a une probabilité de 0,95 que la fluctuation d'échantillonnage place Z entre -1,96 et 1,96. Hors de cet intervalle, on peut affirmer avec 5 % d'erreur qu'un phénomène autre que la fluctuation d'échantillonnage entre en jeu.

Pour un test unilatéral, Z doit être supérieur à -1,645 (pour $\mu < \dots$) ou inférieur à 1,645 (pour $\mu > \dots$) pour accepter l'hypothèse initiale. En dehors de ces valeurs, un phénomène autre que la fluctuation d'échantillonnage entre en jeu (avec 5 % d'erreur).

Le test de STUDENT (ou encore test t ou bien t test) d'une moyenne permet de rejeter ou non une moyenne hypothétique μ_0 lorsque l'on ne connaît pas la variance (et donc l'écart-type) de la population. Dans ce cas, GeoGebra va utiliser l'écart-type corrigé s extrapolé à partir de l'échantillon

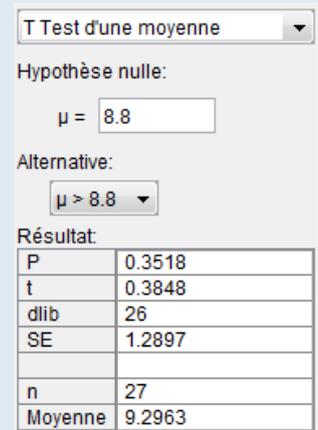
Méthode

- Dans la barre d'outils de la vue **Analyse de données**, cliquer sur le bouton Σx .
- Sélectionner **T Test d'une moyenne** dans la liste déroulante.
- Dans le champ $\mu =$ inscrire la valeur de la moyenne μ_0 à tester.



On considère alors la variable statistique $T_{n-1} = \frac{\mu - \mu_0}{s} \sqrt{n-1}$ qui suit une loi de STUDENT à $n - 1$ degrés de liberté.

- Dans la liste déroulante **Alternative**, sélectionner le type de test à effectuer :
 - $\mu \neq \dots$: pour un test bilatéral;
 - $\mu > \dots$: pour un test unilatéral à droite;
 - $\mu < \dots$: pour un test unilatéral à gauche.
- La rubrique **Résultat** affiche le résultat du test :
 - P : la probabilité de rejet de l'hypothèse nulle;
 - t : la valeur de T_{n-1} obtenue en remplaçant par les données fournies;
 - dlib : le nombre de degrés de liberté associé à la valeur de t ;
 - SE : le quotient de s (écart-type corrigé) par la racine carrée de n (nombre d'observations), ce coefficient (erreur standard) permet de contrôler la fiabilité de l'estimation d'un paramètre d'une population en supposant qu'il n'y ait pas d'erreur de mesure;
 - n : l'effectif de l'échantillon;
 - Moyenne : la moyenne de l'échantillon.

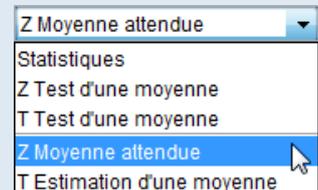


GeoGebra propose la construction d'un intervalle de confiance pour la moyenne de la population sous l'hypothèse de la connaissance ou non de son écart-type.

Le test Z Moyenne attendue est utilisé lorsque l'écart-type de la population est connu :

Méthode

- Dans la barre d'outils de la vue **Analyse de données**, cliquer sur le bouton Σx .
- Sélectionner **Z Moyenne attendue** dans la liste déroulante.
- Dans le champ $\sigma =$ inscrire l'écart-type de la population supposé connu.
- Dans le champ **Niveau de confiance** inscrire la probabilité que la moyenne de la population appartienne à l'intervalle dont les bornes sont données par **Limite inférieure** et **Limite supérieure** dans la rubrique du dessous.



- La rubrique **Résultat** affiche le résultat du test :
 - Intervalle : l'intervalle dans lequel se trouve la moyenne de la population au niveau de confiance choisi sous l'hypothèse de la connaissance de l'écart-type;
 - Limite inférieure : borne inférieure de l'intervalle de confiance;
 - Limite supérieure : borne supérieure de l'intervalle de confiance;
 - Marge d'erreur : moitié de l'étendue de l'intervalle;
 - n : l'effectif de l'échantillon;
 - Moyenne : la moyenne de l'échantillon.

Z Moyenne attendue	
$\sigma =$	3
Niveau de confiance:	
	0.95
Résultat:	
Intervalle	6.1 ± 1.0735
Limite inférieure	5.0265
Limite supérieure	7.1735
Marge d'erreur	1.0735
n	30
Moyenne	6.1

Le test T Estimation d'une moyenne est utilisé lorsque l'écart-type de la population est inconnu.

Méthode

- Dans la barre d'outils de la vue **Analyse de données**, cliquer sur le bouton Σx .
- Sélectionner **T Estimation d'une moyenne** dans la liste déroulante.
- Dans le champ **Niveau de confiance** inscrire la probabilité que la moyenne de la population appartienne à l'intervalle dont les bornes sont données par **Limite inférieure** et **Limite supérieure** dans la rubrique du dessous.

T Estimation d'une moyenne	
Statistiques	
Z Test d'une moyenne	
T Test d'une moyenne	
Z Moyenne attendue	
T Estimation d'une moyenne	

- La rubrique **Résultat** affiche le résultat du test :
 - Intervalle : l'intervalle dans lequel se trouve la moyenne de la population au niveau de confiance choisi sous l'hypothèse de ne pas connaître l'écart-type;
 - Limite inférieure : borne inférieure de l'intervalle de confiance;
 - Limite supérieure : borne supérieure de l'intervalle de confiance;
 - Marge d'erreur : moitié de l'étendue de l'intervalle;
 - dlib : le nombre de degrés de liberté ($n - 1$) qui correspond au paramètre de la loi de STUDENT, loi suivie par l'estimateur;
 - SE : le quotient de l'écart-type σ par la racine carrée de la taille de l'échantillon n (il permet de quantifier la fluctuation d'échantillonnage : plus la taille de l'échantillon est grand, plus petite sera l'erreur standard, et plus proche de la réalité sera l'estimation de la moyenne);
 - n : l'effectif de l'échantillon;
 - Moyenne : la moyenne de l'échantillon.

T Estimation d'une moyenne	
Niveau de confiance:	
	0.95
Résultat:	
Intervalle	6.1 ± 1.1196
Limite inférieure	4.9804
Limite supérieure	7.2196
Marge d'erreur	1.1196
dlib	29
SE	0.5474
n	30
Moyenne	6.1

