

## Fiche UE4 : Biostatistiques :

Type de donnée	Caractérisation des données	Représentation graphique	Données à 2 dimensions
<p><b>Qualitatives :</b> Une donnée est qualitative si elle peut se présenter sous plusieurs <b>aspects</b> ou <b>modalités</b>.</p> <p>Ces données donnent lieu à des <b>dénombrements</b>.</p> <p><i>Exemples :</i> Race d'un chien, couleur des yeux...</p>	<ul style="list-style-type: none"> <li>• Fréquence <b>absolue</b> (effectif de la classe concernée)</li> <li>• Fréquence <b>relative</b> (= pour chaque classe, le rapport de son effectif à l'effectif total, la somme des fréquences relatives étant = à 1). On l'exprime également en %.</li> </ul>	<p>✓ <b>Diagramme Camembert :</b> Il représente les effectifs absolus ou relatifs par des secteurs de cercles correspondants aux classes proportionnels à l'effectif de chacune des classes.</p>	<p>✓ <b>Tableau de contingence :</b> Notez qu'on appelle la dernière ligne et colonne "les marginales".</p> <p>✓ <b>Regroupement libre</b> (cité)</p>
<p><b>Ordinales :</b> Une donnée est dite ordinale s'il existe entre les différentes classes de la donnée une relation d'ordre.</p> <p><i>Exemples :</i> Stade d'un cancer, Stade histologique d'une tumeur..</p>	<ul style="list-style-type: none"> <li>• Fréquences <b>absolues</b></li> <li>• Fréquences <b>relatives</b></li> <li>• Fréquences <b>cumulées</b> (relatives et absolues) qui présentent le cumul de chaque classe avec les classes précédentes.</li> </ul>	<p>✓ <b>Diagramme en bâtons :</b> La longueur du bâton représente l'effectif de la classe. Attention, dans le cas de fréquences cumulées, le bâton le plus grand n'est pas forcément celui de plus grand effectif.</p>	<p>✓ <b>Tableau de contingence</b></p> <p>✓ <b>Regroupement de classes contigües</b></p>
<p><b>Quantitatives :</b> Une donnée est quantitative quand elle prend pour valeur un nombre résultant d'une mesure, avec une unité, d'un caractère chez des individus.</p> <p>Un caractère quantitatif peut être continu (prend toutes les valeurs de l'intervalle ou il est défini), ou discret (prend des valeurs isolées sur cet intervalle).</p> <p><i>Exemple :</i> Taille d'un individu, Glycémie...</p>	<ul style="list-style-type: none"> <li>• <b>Tableaux de mesure</b></li> <li>• <b>Paramètres de tendance centrale</b> (certains caractérisent des données ordinales ou qualitatives, on y reviendra)</li> <li>• <b>Paramètres de dispersion</b> (certains peuvent être utilisés pour des données ordinales, on y reviendra)</li> </ul>	<p>✓ Pour les variables <b>continues :</b> <b>Histogramme :</b> C'est la surface du rectangle qui est proportionnelle à l'effectif de la classe (faire attention à l'échelle sur l'axe des abscisses formant des intervalles).</p> <p>✓ Pour les variables <b>discontinues :</b> <b>Diagramme en bâtons.</b></p>	<ul style="list-style-type: none"> <li>• <b>Nuage de points :</b> On place le couple (x;y) à un "endroit" sur le graphique. On retrouve "G", centre de gravité, point moyen de coordonnées(<math>\bar{x}</math>; <math>\bar{y}</math>).</li> </ul>

## Fiche UE4 : Biostatistiques :

Paramètre central	Type de donnée	Définition - Calcul	Unité
Moyenne	Quantitatif	<p>Elle correspond au <b>rapport</b> de la <b>somme</b> des mesures au <b>nombre</b> de mesures effectuées.</p> $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ <p><b>Propriété</b> : Soit une série de mesure <math>x_i</math> a pour moyenne <math>\bar{x}</math>, et deux constantes <b>a</b> et <b>b</b>. S'il existe une série de mesure <math>y_i</math> telle : <math>y_i = ax_i + b</math>, alors <math>\bar{y} = a\bar{x} + b</math>.</p>	<i>Même unité que les <math>x_i</math> : unité de la mesure.</i>
Médiane	Quantitatif & Ordinal.	<p>C'est une <b>valeur séparant l'ensemble des données préalablement ordonnées en deux sous-ensembles de même effectif</b>, on a donc la nécessité d'ordonner les mesures pour estimer la médiane.</p> <ul style="list-style-type: none"> <li>• Si le nombre de mesures <b>n</b> est impair (<math>n = 2k + 1</math>), alors on détermine la médiane comme la valeur située <b>au milieu</b> de la série de mesures ordonnées (c'est à dire qu'on aura un nombre égal de mesures de part et d'autre de la médiane).</li> <li>• Si <b>n</b> est pair (<math>n = 2k</math>), alors on appelle médiane toute valeur comprise entre <math>x_k</math> et <math>x_{k+1}</math>, c'est à dire entre les deux valeurs les plus "médianes" de la série.</li> <li>• La médiane est <b>bien moins influencée</b> que la <b>moyenne</b> par les valeurs extrêmes.</li> </ul>	<i>Même unité que les <math>x_i</math> : unité de la mesure.</i>
Quantiles		<p>Ce sont des <b>valeurs séparant l'ensemble des données préalablement ordonnées en sous-ensembles d'effectifs égaux</b>. On retrouve notamment les <b>quartiles et les percentiles</b>.</p> <ul style="list-style-type: none"> <li>• <b>Les Quartiles</b> séparent les valeurs de la variable en 4 groupes d'effectif identique, on a donc 3 valeurs appelées quartiles : Q1, Q2, Q3. (Notez que Q2 correspond tout simplement à la médiane).</li> </ul>	<i>Même unité que les <math>x_i</math> : unité de la mesure.</i>
Mode	Qualitatif, Quantitatif, Ordinal.	<p>Elle correspond à la <b>classe de plus grand effectif</b> ou de <b>plus grande fréquence</b>. Il existe également des distributions à plusieurs modes, appelées multimodales. Il est important de distinguer <b>mode</b> et <b>plus « grande » classe de la série de mesure</b>.</p>	<i>Même unité que les <math>x_i</math> : unité de la mesure.</i>

## Fiche UE4 : Biostatistiques :

Paramètre de dispersion	Type de donnée adapté	Définition et Calcul	Unité
Variance	Quantitatif	<p>Elle se définit ainsi : la variance est la <b>moyenne arithmétique des carrés des écarts à la moyenne de l'échantillon.</b></p> $Var(X) = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n}$ <p>Quand on parle de variance, on parle de variance de l'échantillon et non pas de celle dans la population (ça prendra son sens dans les chapitres suivants).</p>	Celle des <b><math>x_i</math> au carré</b> , vu qu'on multiplie une valeur par une autre mais on divise que par un chiffre.
Ecart-type		<p>L'écart type correspond « simplement » à la racine carrée de la variance :</p> $\sigma = \sqrt{Var(X)}$	<i>Même unité que les <math>x_i</math> :</i> unité de la mesure.
Etendue		<p>L'étendue correspond à la <b>différence entre la plus grande et la plus petite valeur de la série de mesure.</b></p> $e = x_{max} - x_{min}$	<i>Même unité que les <math>x_i</math> :</i> unité de la mesure.
Intervalle inter-quartile	Quantitatif & Ordinal	<p>Il correspond à la <b>différence entre les valeurs de Q3 et Q1</b>, soit « Q3 – Q1 ». L'intervalle inter-quartile, comme vous pouvez le deviner (on se trouve entre 25% et 75% des mesures) contient <b>50% des mesures.</b></p> <ul style="list-style-type: none"> <li>• Plus la <b>largeur</b> de l'intervalle interquartile est <b>faible</b>, plus les données sont dites <b>groupées.</b></li> <li>• Plus la <b>largeur</b> de l'intervalle interquartile est <b>importante</b>, plus les données sont <b>dispersées.</b></li> </ul>	<i>Même unité que les <math>x_i</math> :</i> unité de la mesure.

## Fiche UE4 : Biostatistiques :