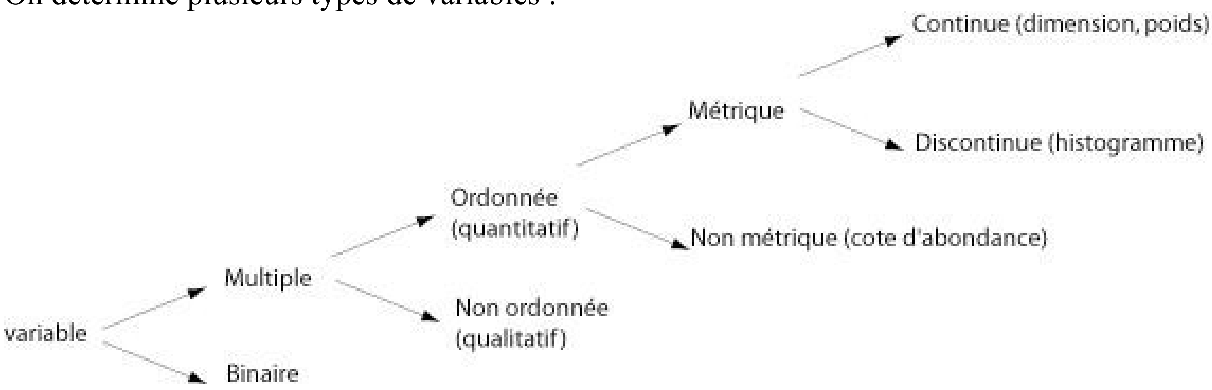
Mathématiques	
Loi Normale et Estimation	**Niveau 3**

I] Introduction

1°) Différents types de variables :

On détermine plusieurs types de variables :



2°) Notion de populations et d'échantillons :

On ne peut pas connaître une population dans son ensemble, mais on peut l'estimer grâce à l'étude d'échantillons : c'est l'inférence statistique.

On suppose un échantillonnage aléatoire simple : tous les individus de la population ont la même probabilité de faire partie de l'échantillon.

La moyenne de l'échantillon est un estimateur de la moyenne théorique de la population.

II] Statistique descriptive :

Il existe plusieurs manières de représenter les données :

_ par histogramme
_ Fonction de densité

1°) Paramètres de position

a) Moyenne :

$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ ou $\bar{x} = \sum_{j=1}^{p} f_j.c_j$ avec p le nombre de classe, f la fréquence à la classe j et c le centre de la classe j.

b) Moyenne géométrique :

$$\bar{x}_g = \sqrt[n]{x_1.x_2...x_n}$$

c) Médiane :

Impaire : $\tilde{x} = x_{\frac{n+1}{2}}$

Paire : $\tilde{x} = \dfrac{(x_{\frac{n}{2}} + x_{\frac{n+1}{2}})}{2}$. Cette valeur est la valeur centrale.

La médiane est un estimateur robuste.

d) Mode :

$M_0 = L + i(\dfrac{\Delta i}{\Delta 1 + \Delta S})$ elle représente la valeur de la variable qui a la plus forte fréquence.

L est la borne inférieure de la classe modale (classe ayant la fréquence la plus élevée)
i est l'intervalle de classe.

2°) Paramètres de dispersion

a) Variance

La variance représente la somme des carrés des écarts à la moyenne

$\sigma^2 = \dfrac{1}{N}\sum (x_i - \bar{x})^2$ au niveau de la population, valeur théorique inconnue.

$\hat{\sigma}^2 = S_x^2 = \dfrac{1}{n-1}\sum (x_i - \bar{x})^2$ Estimateur sans biais de la population : Variance de l'échantillon.

$$S_x^2 = \frac{1}{n-1}\left[\sum x_i^2 - \frac{1}{n}(\sum x_i)^2\right]$$

b) Ecart-type

$\sigma = \sqrt{\sigma^2}$ et $S_x = \sqrt{S_x^2}$

L'écart-type traduit la dispersion de l'échantillon.

c) Moments :

$$M = \frac{1}{n}\sum x_i - c)^k$$

Si c=0 et k=1, on obtient la moyenne (moment par rapport à l'origine)

Si $c = \bar{x}$ et k=2, on obtient la variance
Si $c = \bar{x}$ et k=3, M traduit l'asymétrie de la distribution
Si $c = \bar{x}$ et k=4, M traduit l'aplatissement de la distribution.

III] Loi Normale :

La loi normale représente des distributions de variables aléatoires.

N(0,1) ; pour N (μ, σ) on réduit la fonction a $\frac{x-\mu}{\sigma}$

Si $x_1 \in \aleph(\mu_1, \sigma_1)$ et $x_2 \in \aleph(\mu_2, \sigma_2)$ alors $x_1 \pm x_2 \in \aleph(\mu_1 \pm \mu_2 ; \sqrt{\sigma_1^2 + \sigma_2^2})$

Et E(aZ)=a E(Z) et Var (aZ)=a² Var (Z)

Intervalle de confiance :

Pour un échantillon appartenant à $\aleph(M, Var(M))$, les moyennes vont suivre une loi normale N (μ, $\frac{\sigma}{\sqrt{n}}$)

a) Intervalle de confiance de la moyenne (n>30)

$$P(-Z_{\frac{\alpha}{2}} < Z < Z_{\frac{\alpha}{2}}) = 1 - \alpha$$

$$P(-Z_{\frac{\alpha}{2}} < \frac{M-\mu}{\sqrt{VarM}} < Z_{\frac{\alpha}{2}}) = 1 - \alpha$$

Soit $P(\bar{x} - Z_{\frac{\alpha}{2}} S_{\bar{x}}) < \mu < \bar{x} + Z_{\frac{\alpha}{2}} S_{\bar{x}}) = 1 - \alpha$

Pour $\alpha = 0.05 \quad Z_{\frac{\alpha}{2}} = 1.96$
$\alpha = 0.01 \text{'} \; Z_{\frac{\alpha}{2}} = 2.575$

b) Intervalle de confiance de la moyenne (n<30)

$P(\bar{x} - t_{\frac{\alpha}{2}} S_{\bar{x}}) < \mu < \bar{x} + t_{\frac{\alpha}{2}} S_{\bar{x}}) = 1 - \alpha$ avec $S_{\bar{x}} = \frac{S_x}{\sqrt{n-1}}$