



Théorie Statistique

Cours proposé par
Edmundo NAVA SAUCEDO
Aux étudiants de LicPro I4A

Utilisons-nous à bon escient les statistiques ?

- Mesures de centralisation / Mesures de dispersion
- Population / Echantillon
- Distribution Normal (Gauss) / Distribution de Student (Gosset)
- Intervalles de Confiance / Certitude
- Test de signification
- Régression linéaire
 - Ligne droite
 - Courbe logarithmique
 - Hyperbole équilatérale
- Plans d'expériences factoriels

Théorie Statistique

Mesures de centralisation / Mesures de dispersion

Distributions discrètes / Distributions continues

Loi Normal (Gauss) ----> Méthodes paramétriques

Différencier : Population / Echantillon

Grandes Populations / Petites Populations

Grands Echantillons / Petits Echantillons

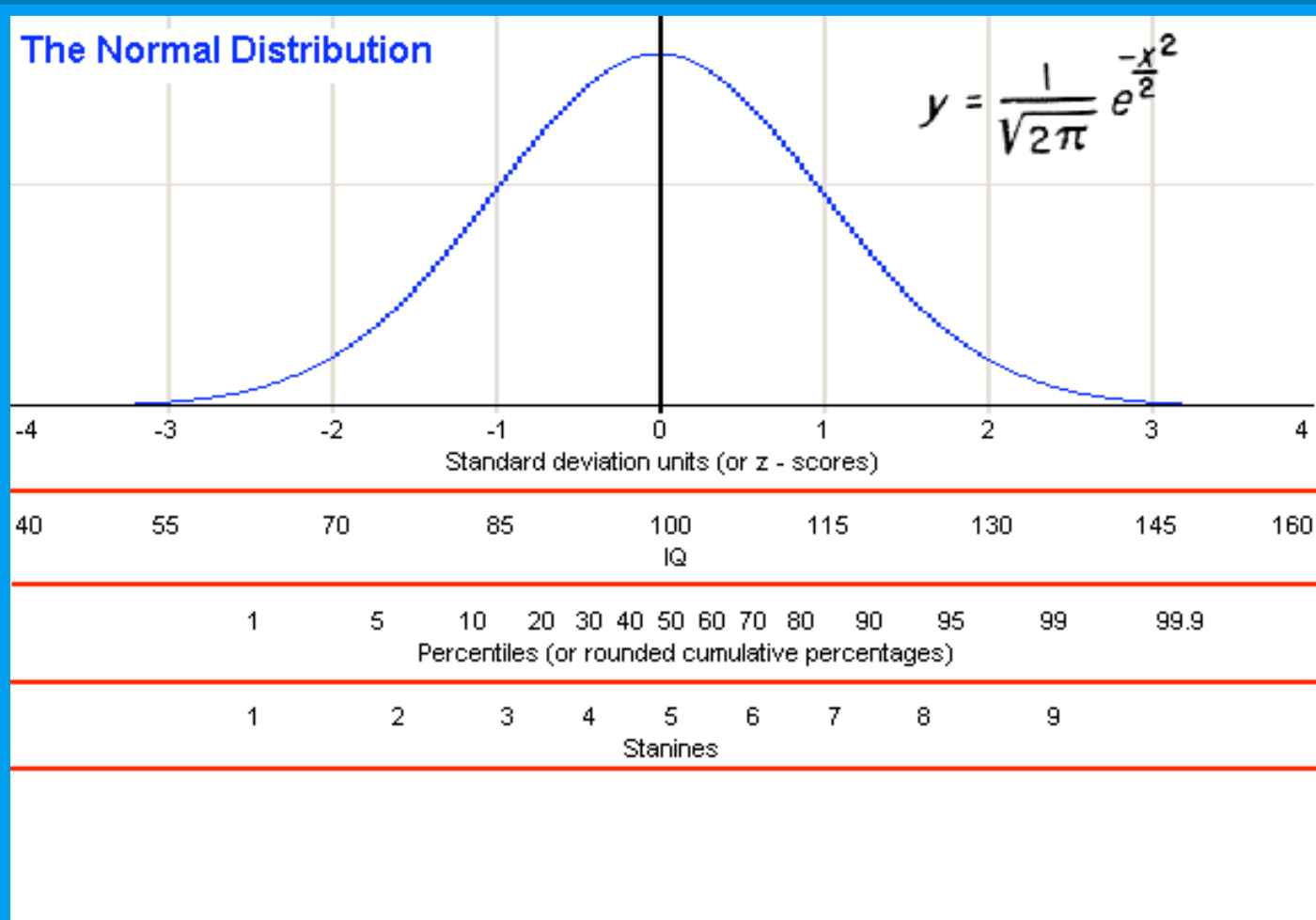
$N \geq 100$ ----> Grande Population

$N < 100$ ----> Petite Population

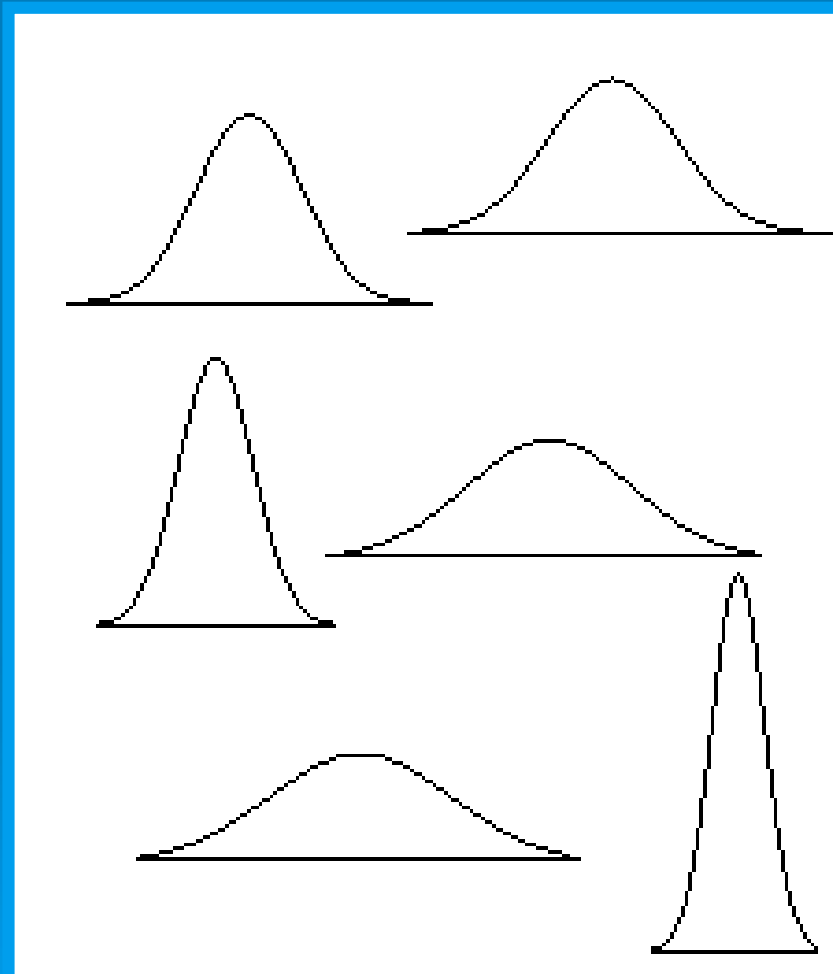
$n \geq 30$ ----> Grand Echantillon

$n < 30$ ----> Petit Echantillon

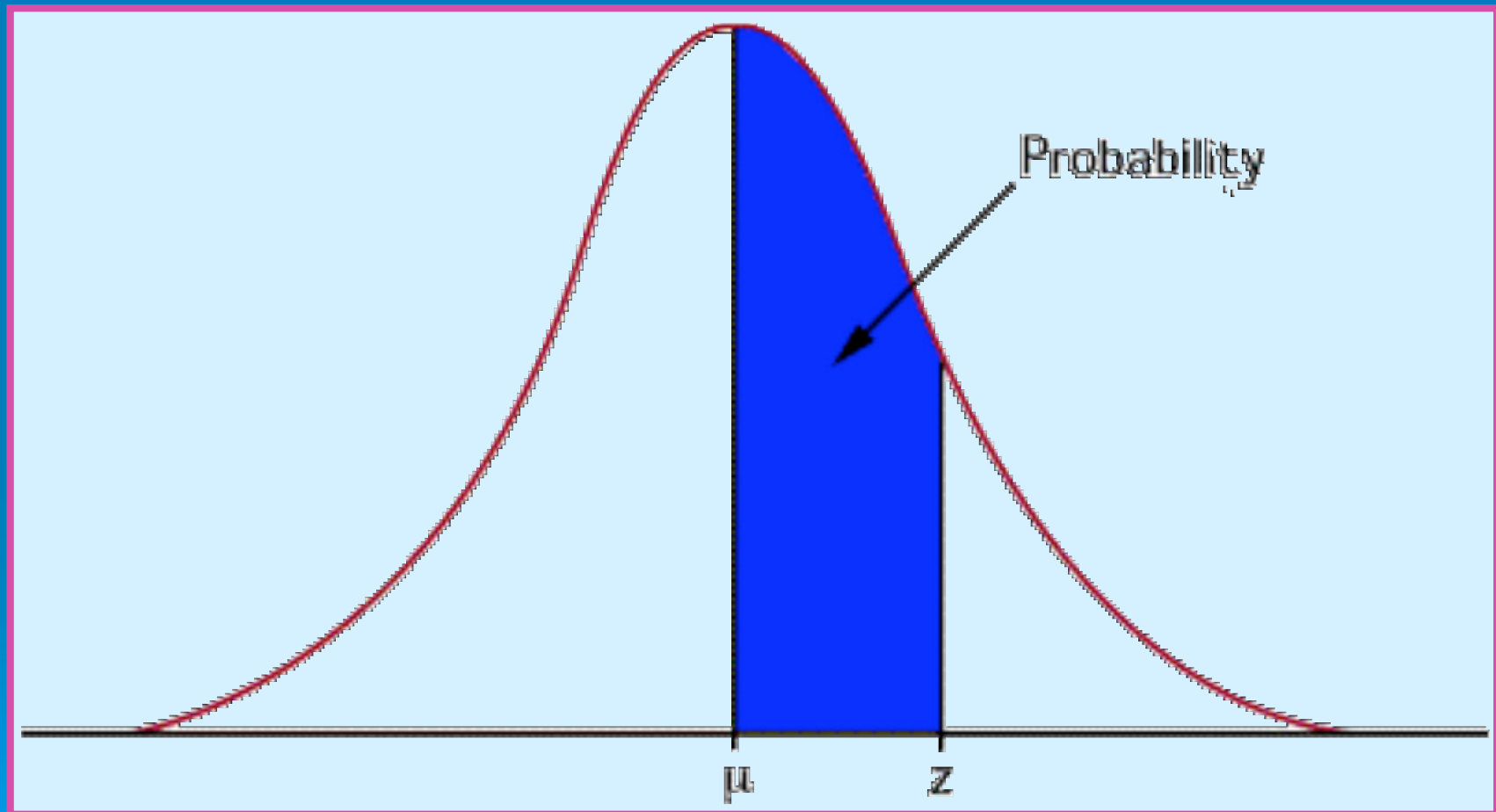
Distribution Normale (Gauss)



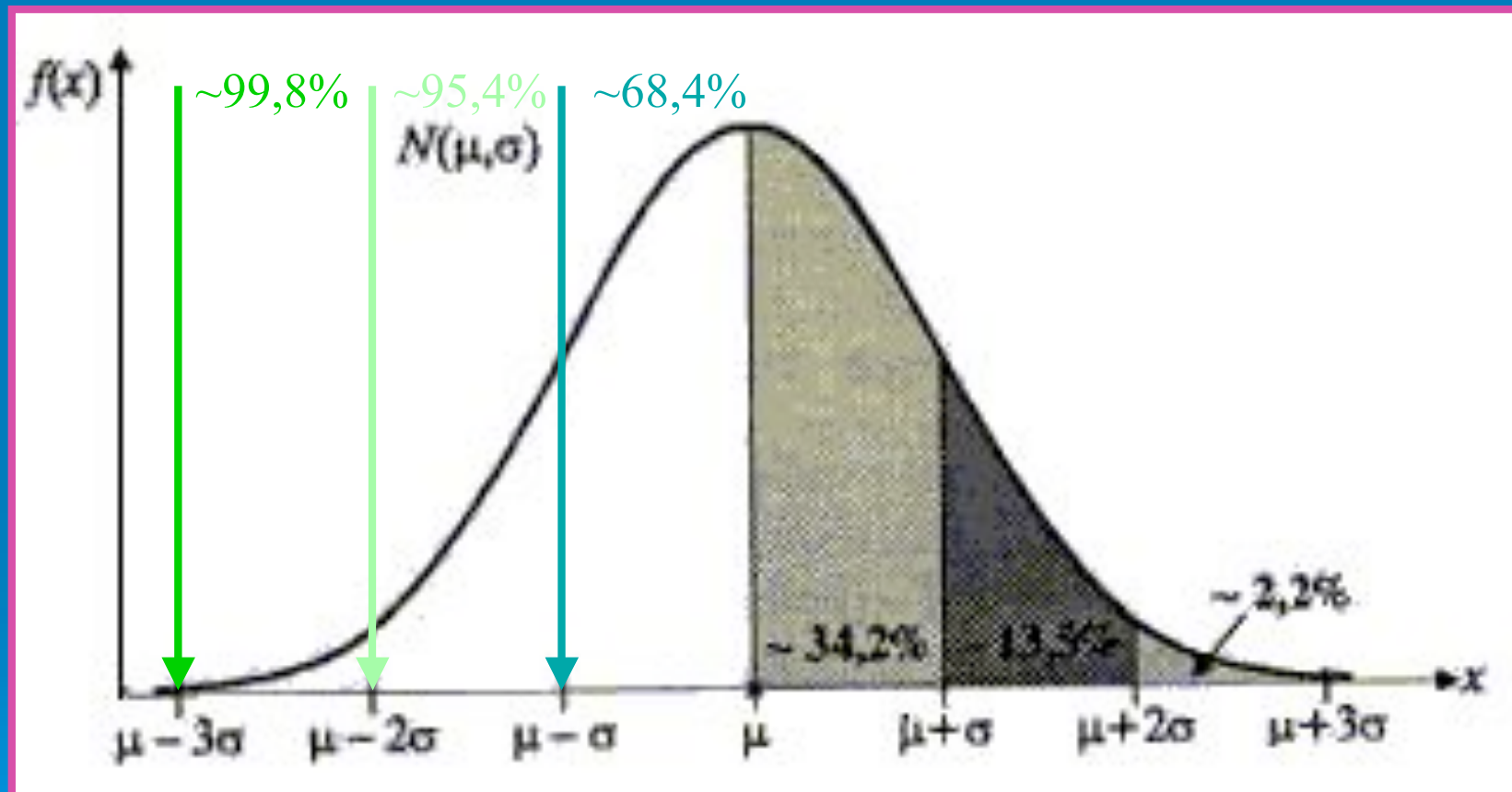
La forme de la courbe dépend de μ (moyenne) et σ (écart type)



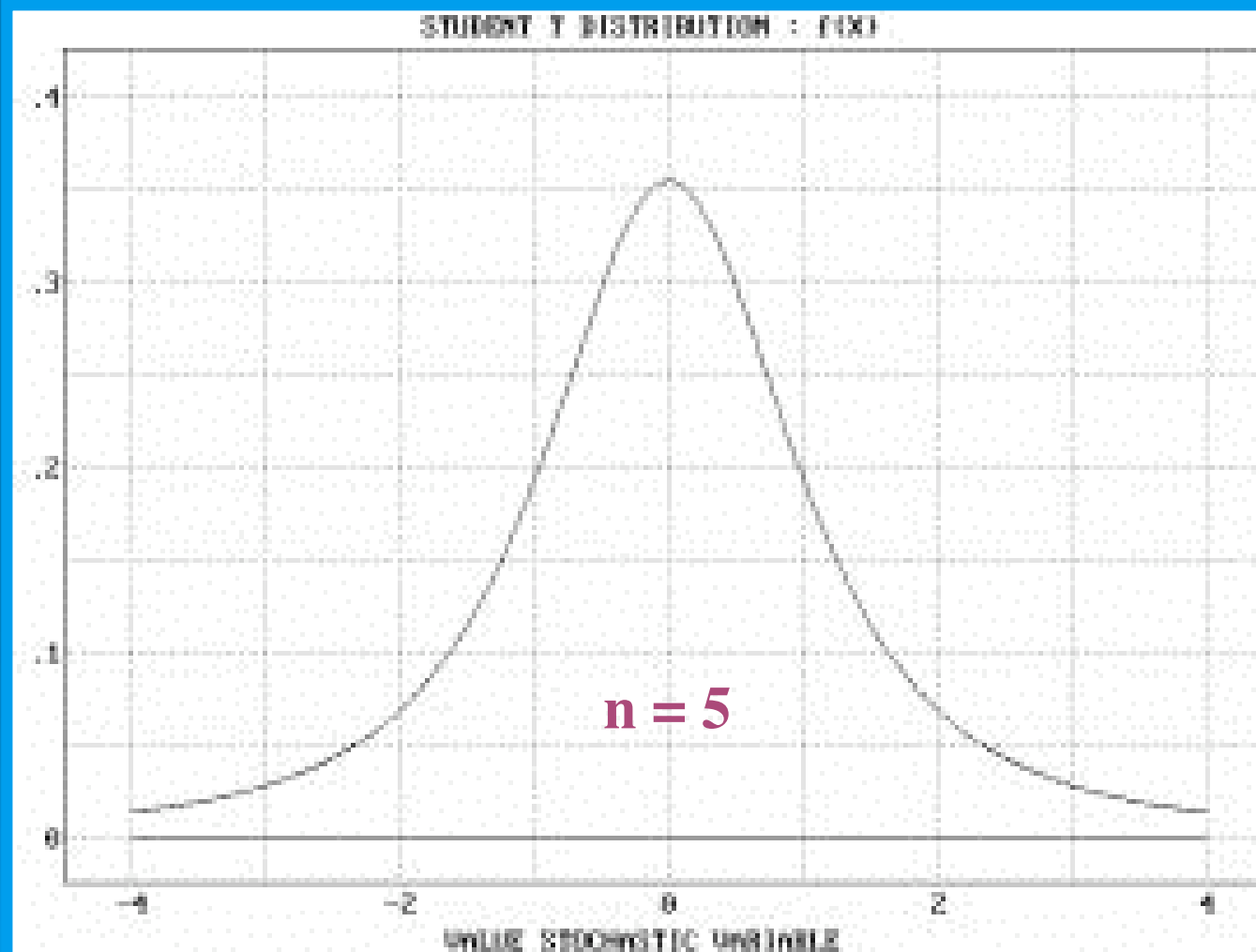
Probabilité



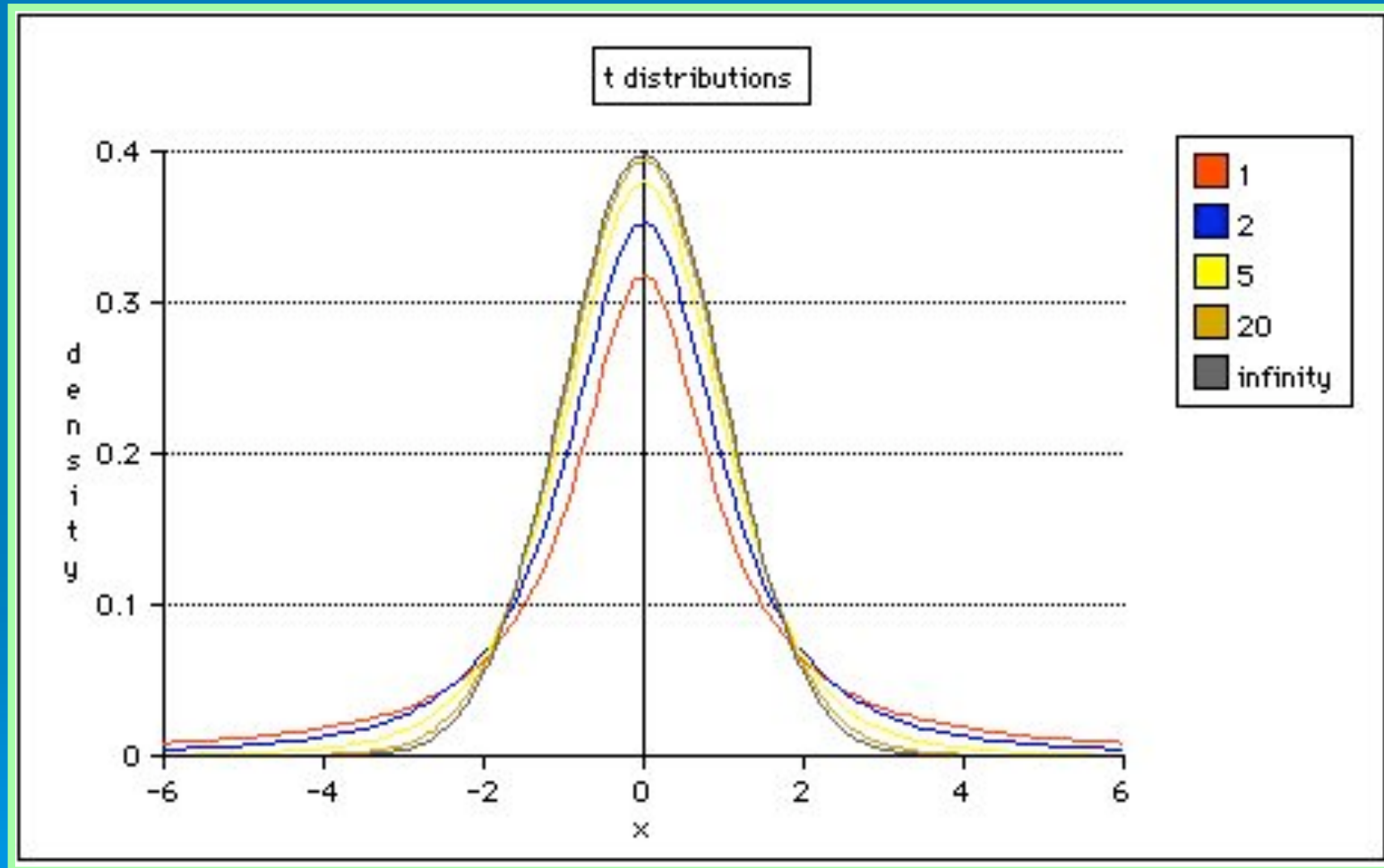
Pourcentages sous la courbe de Gauss



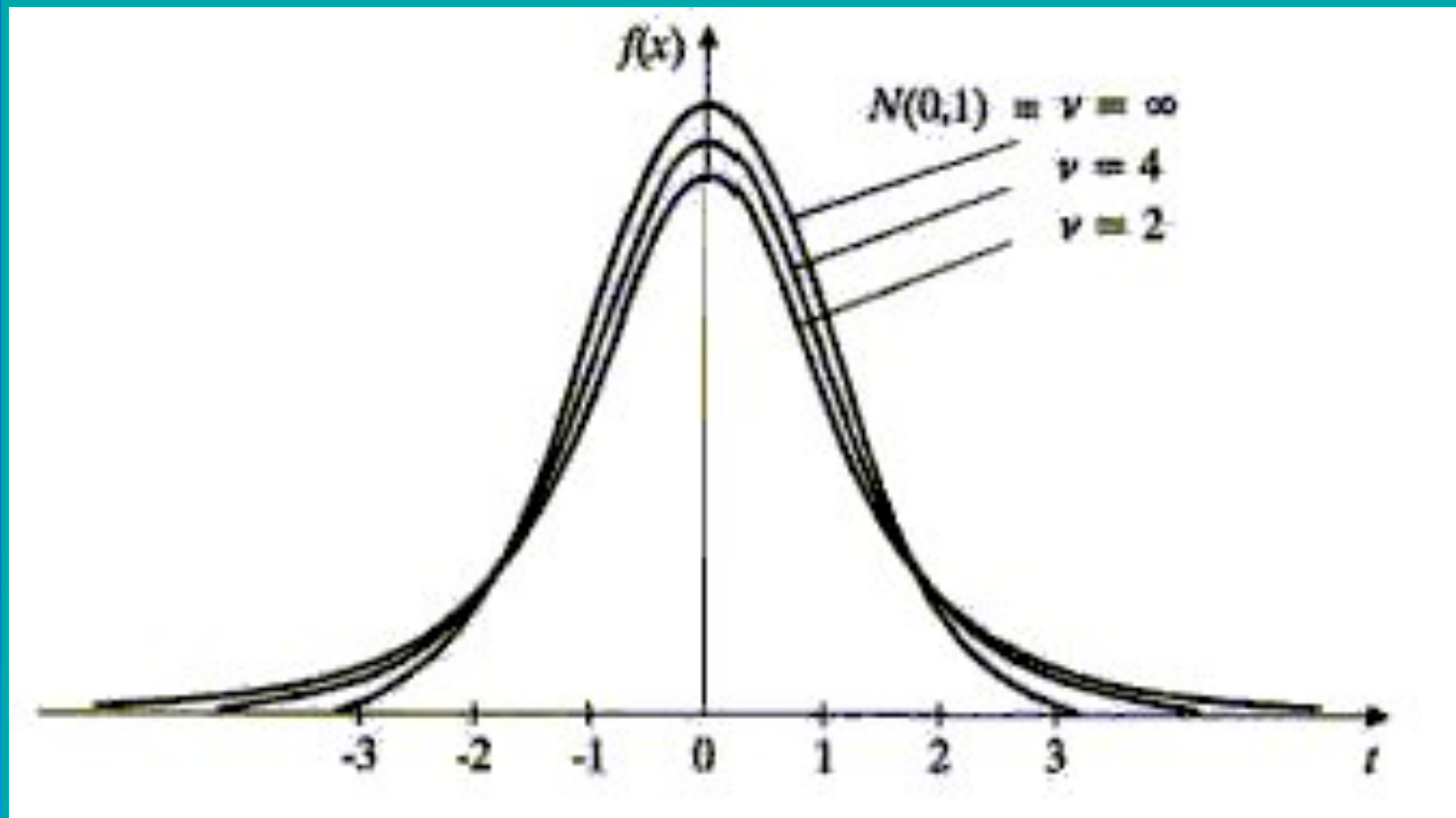
Distribution de Student (Gosset)



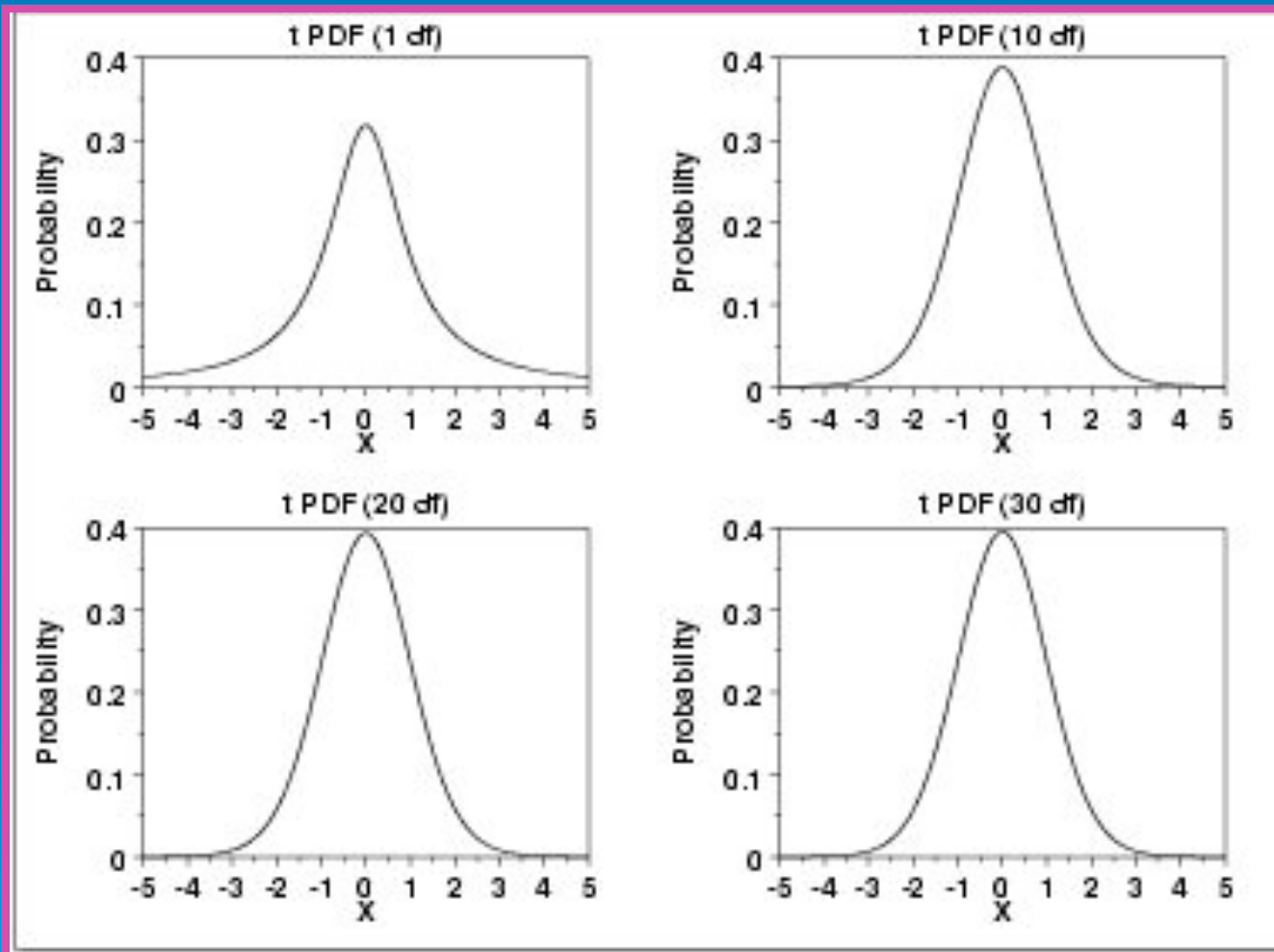
Distributions t



Degrés de liberté



Distribution t (fonction de **df**)



Intervalles de Confiance

Souvent la distribution est normale et continue

Souvent on a une Population **N** éléments

Souvent **N** tend vers l'infini

~~~~

On est obligé de prendre un Echantillon de **n** éléments

~~~~

On calcule la moyenne de l'Echantillon $\bar{X}$ ---> meilleure estimation de la moyenne de la Population

~~~~

On calcule l'écart type de l'Echantillon  $S_x$  ---> meilleure estimation de l'écart type de la Population

~~~~

Calcul des Intervalles de Confiance

~~~~

Si  $n \geq 30$  --->  $\mu = \bar{X} \pm [s_x / \sqrt{n}] \cdot Z$

Z valeurs de confiance de la distribution normale

~~~~

Si $n < 30$ ---> $\mu = \bar{X} \pm [s_x / \sqrt{n}] \cdot t$

T valeurs de confiance de la distribution de Student (pour n-1 degrés de liberté)

Intervalles de Confiance

*Beaucoup de gens confondent **l'écart-type** et **l'erreur-type** de la moyenne et ne savent pas lequel utiliser pour présenter les données sous forme graphique ou tabulaire.*

***L'écart-type** indique la variabilité des données originales et devrait être mentionné pour toutes les études.*

***L'erreur-type** montre la variabilité des valeurs moyennes, comme si l'étude avait été reprise de nombreuses fois.*

Intervalles de Confiance

En soi, l'erreur-type n'a pas d'utilité particulière; toutefois, on s'en sert pour créer les intervalles de confiance à 95% et à 99% utilisés pour établir la fourchette de valeurs dans laquelle se situe la valeur «réelle».

Les intervalles de confiance signalent au lecteur la précision des estimations des valeurs scientifiques (ou démographiques). Lorsqu'on se sert de graphiques, la barre d'erreur représente un intervalle de plus à moins 2 (1,96) erreurs-types (ce qui correspond à l'intervalle de confiance de 95 %).

Elle devrait entourer la valeur moyenne. Les épreuves de signification statistique et les intervalles de confiance présentent une grande utilité, car ils aident le lecteur à établir l'importance des constatations.

Ecart-type

Nombre
d'éléments

POPULATION

ECHANTILLON

N

n

Moyenne

μ

$\bar{x}$

Ecart-type

$$\sigma = \sqrt{[\sum (\bar{x} - x)^2 / N]}$$

$$S_x = \sqrt{[\sum (\bar{x} - x)^2 / (n-1)]}$$

Intervalle de Confiance

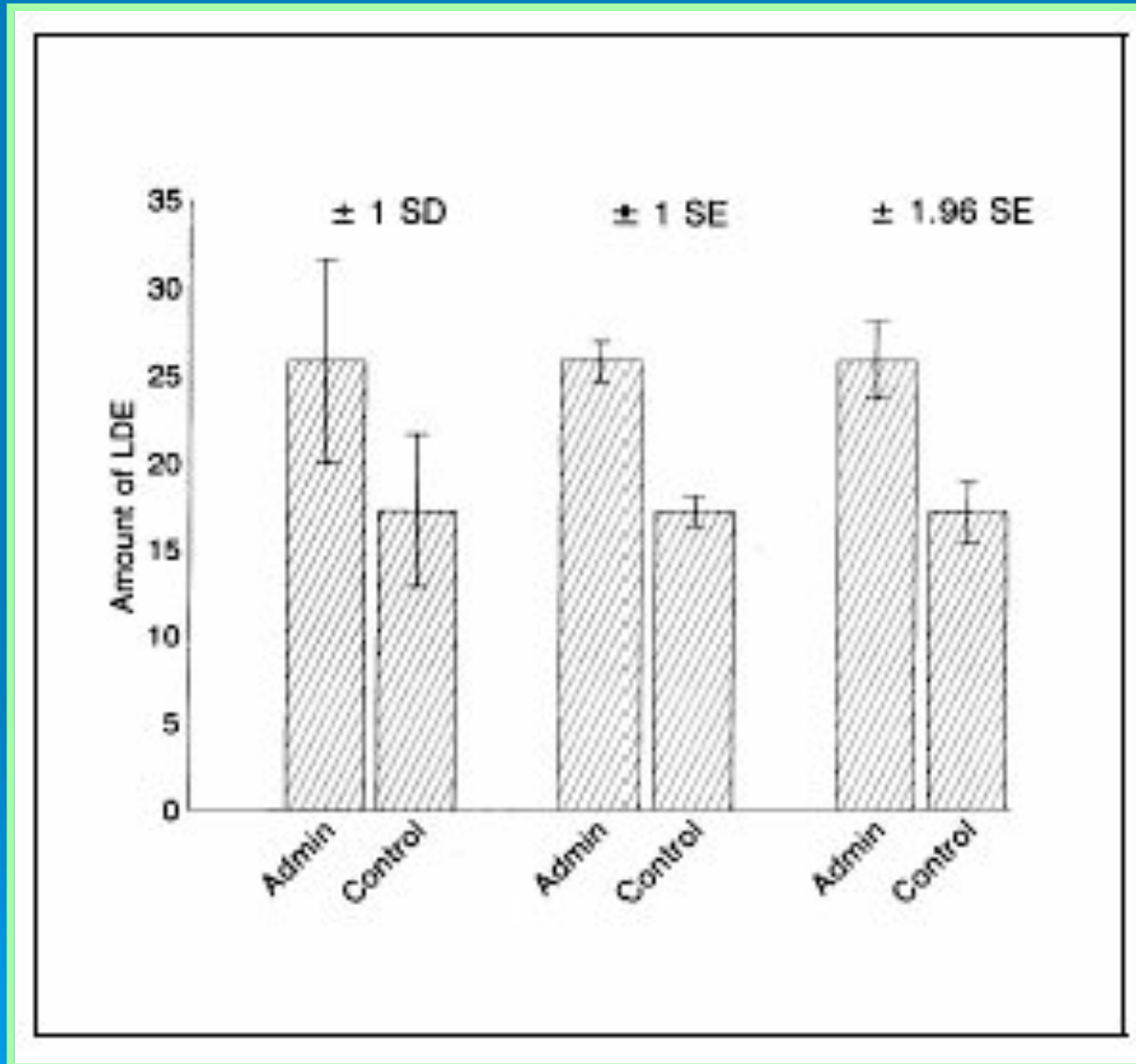


Table de Student

r	90%	95%	97.5%	99.5%
1	3.07766	6.31371	12.7062	63.656
2	1.88562	2.91999	4.30265	9.92482
3	1.63774	2.35336	3.18243	5.84089
4	1.53321	2.13185	2.77644	4.60393
5	1.47588	2.01505	2.57058	4.03212
10	1.37218	1.81246	2.22814	3.16922
30	1.31042	1.69726	2.04227	2.74999
100	1.29007	1.66023	1.98397	2.62589
∞	1.28156	1.64487	1.95999	2.57584

Intervalles de Confiance

Ecart-type

$$s_x$$

Erreur type

$$s_x/\sqrt{n}$$

Intervalles de confiance

$$\pm [s_x/\sqrt{n}] \cdot z$$

ou

$$\pm [s_x/\sqrt{n}] \cdot t$$

Test de Signification

Université A

16 étudiants moyenne QI = 107 écart type = 10

Université B

14 étudiants moyenne QI = 112 écart type = 8

Y a-t-il une différence significative entre les quotients intellectuels des deux groupes aux seuils de signification :

a) 0,01 b) 0,05

Test de Signification

Si μ_1 et μ_2 sont les moyennes dans la population des quotients intellectuels des deux groupes, on doit opter entre les deux hypothèses :

$H_0 : \mu_1 = \mu_2$, il n'y a pas de différence apparente entre les deux groupes

$H_1 : \mu_1 \neq \mu_2$, il y a une différences significative entre les deux groupes

Sous l'hypothèse H_0 ,

$$t = (X_1 - X_2) / \sigma \sqrt{(1/N_1 + 1/N_2)} \quad \sigma = \sqrt{[(N_1 s_1^2 + N_2 s_2^2) / (N_1 + N_2 - 2)]}$$

D'où : $\sigma = 9,44$ et $t = 1,45$

Test de Signification

$\sigma = 9,44$ et $t = 1,45$

a) Avec un test bilatéral au seuil 0,01, on rejette H_0 si t est extérieur à l'intervalle $(-t_{0,995} ; +t_{0,995})$ qui pour $(N_1 + N_2 - 2) = 28$ degrés de liberté est l'intervalle $(-2,76 ; 2,76)$.

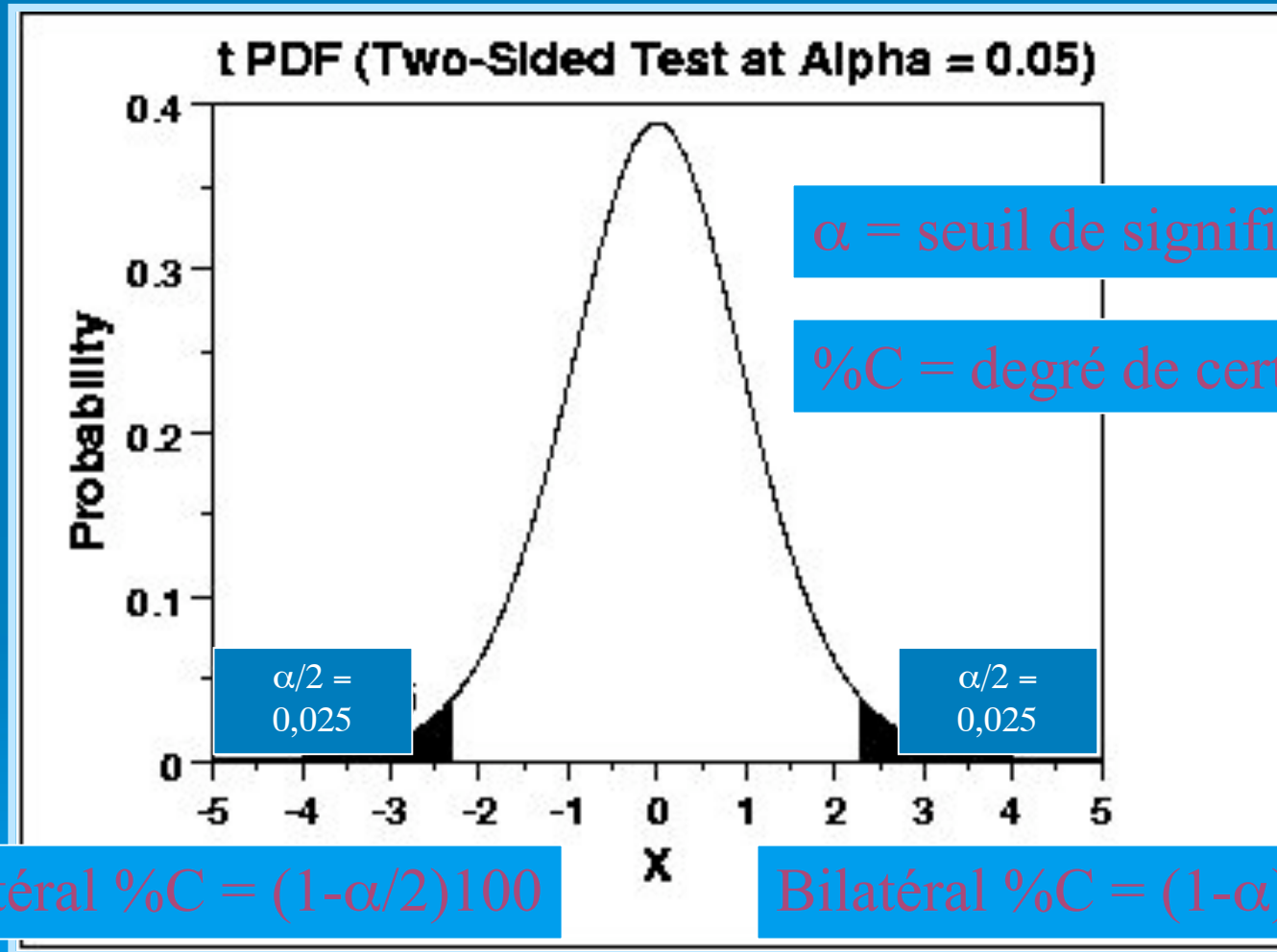
On ne peut pas rejeter au seuil de signification 0,01.

b) Avec un test bilatéral au seuil 0,05, on rejette H_0 si t est extérieur à l'intervalle $(-t_{0,975} ; +t_{0,975})$ qui pour 28 degrés de liberté est l'intervalle $(-2,05 ; 2,05)$.

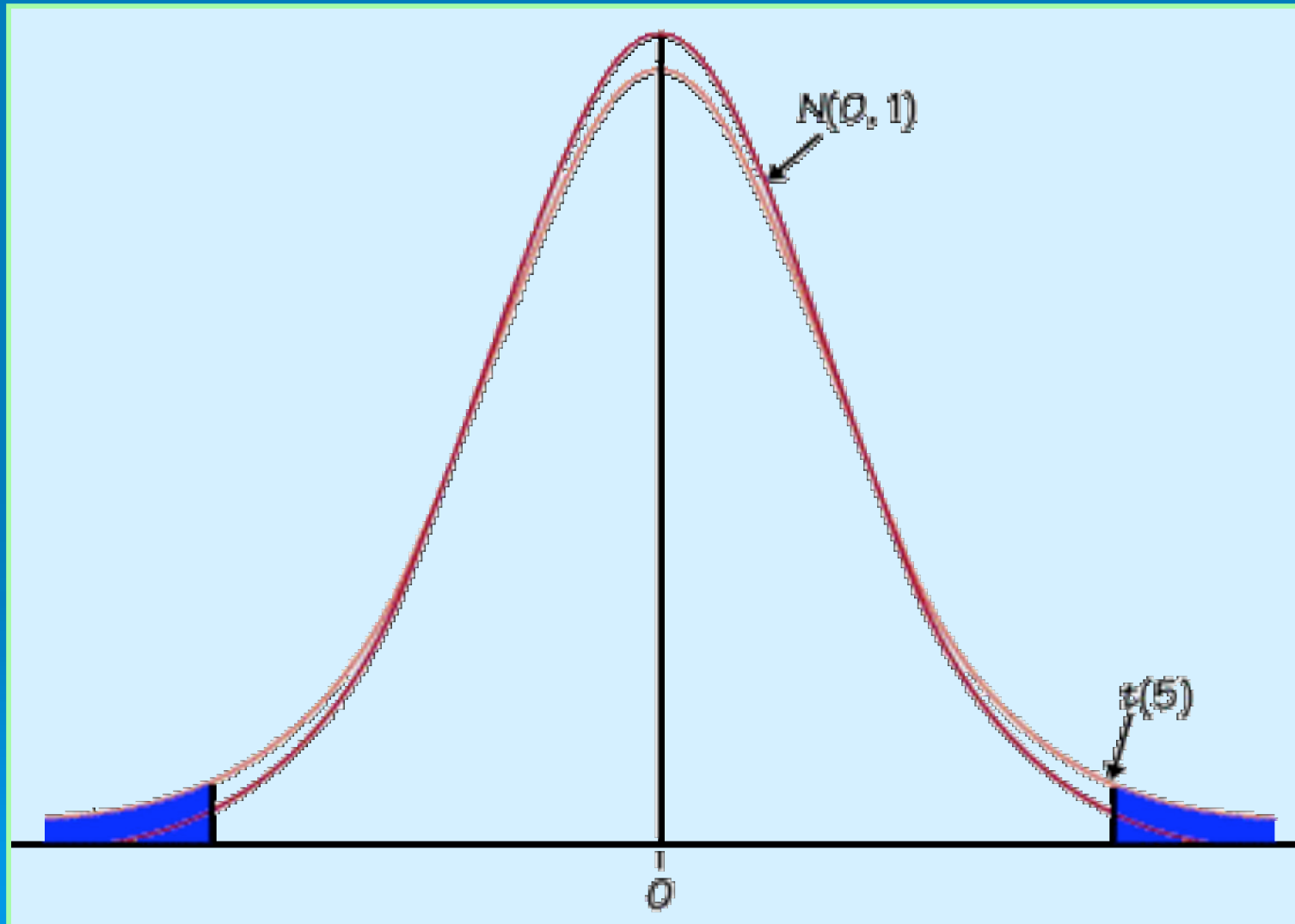
On ne peut pas rejeter au seuil de signification 0,05.

On en conclut qu'il n'y a pas de différence significative entre les quotients intellectuels des deux groupes.

Test de Signification

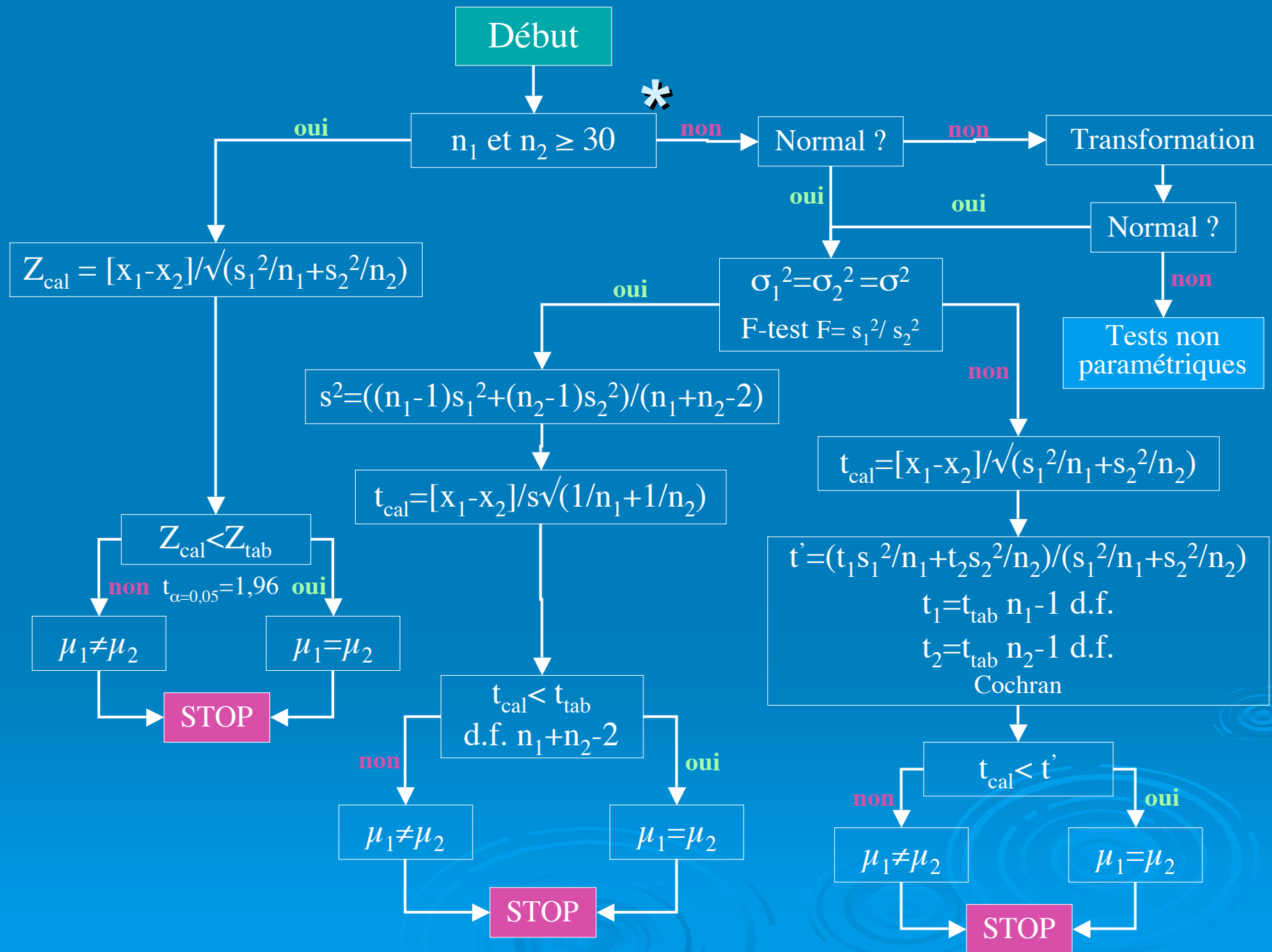


Comparaison Gauss / Student



Arbre décisionnel





fin

