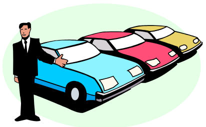
Inférence : deux approches

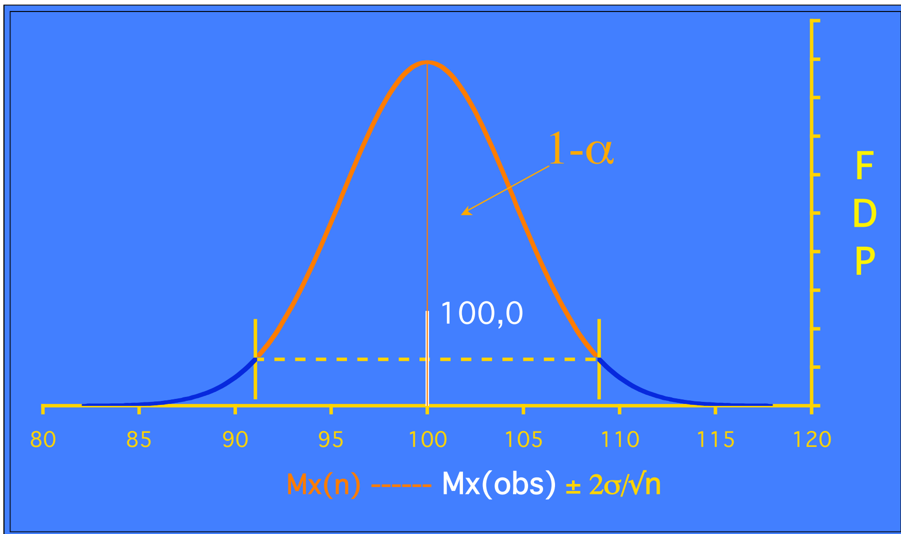
Intervalle de confiance

Test d'hypothèse



L'inférence statistique est un ensemble de techniques qui permet de contrôler le risque d'erreur de la généralisation des observations réalisées dans un échantillon à la population dont il est issu. Deux grandes familles de techniques se complètent : l'intervalle de confiance et le test d'hypothèse. Si l'on prend l'exemple du client qui se rend dans un garage avec sa voiture en panne, l'intervalle de confiance répond à la question : combien la réparation va-t-elle coûter? (*Très probablement* entre 2.000 et 2.500 euros), ou : combien de kms vais-je pouvoir rouler après cette réparation (*Très probablement* entre 10.000 et 15.000 Kms). Il s'agit donc d'une fourchette qui fournit une base de réflexion. Le test d'hypothèse est une prise de décision : soit je répare, soit je ne répare pas et j'achète une nouvelle voiture.

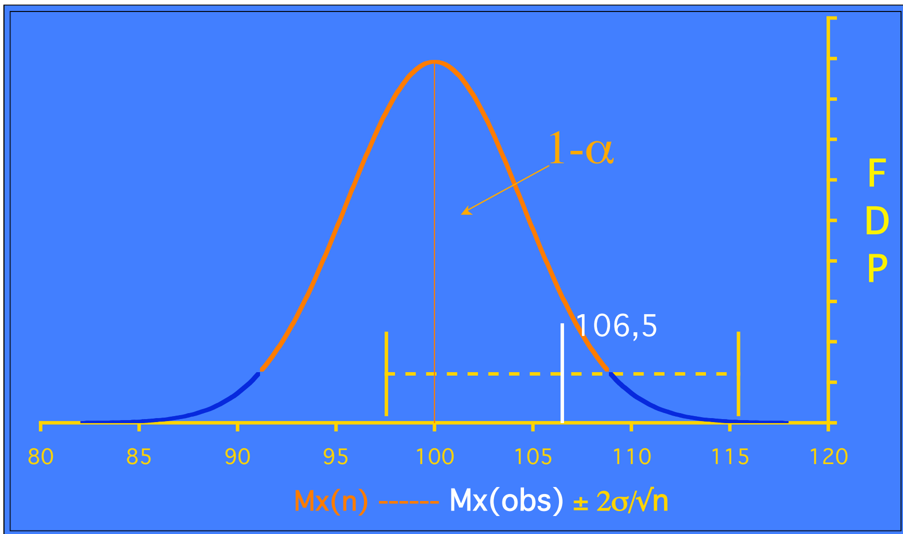
Les barres d'erreurs sont essentiellement une représentation graphique de l'intervalle de confiance. L'utiliser comme « test d'hypothèse graphique » est nettement plus périlleux.



Intervalle de confiance lorsque σ est connu

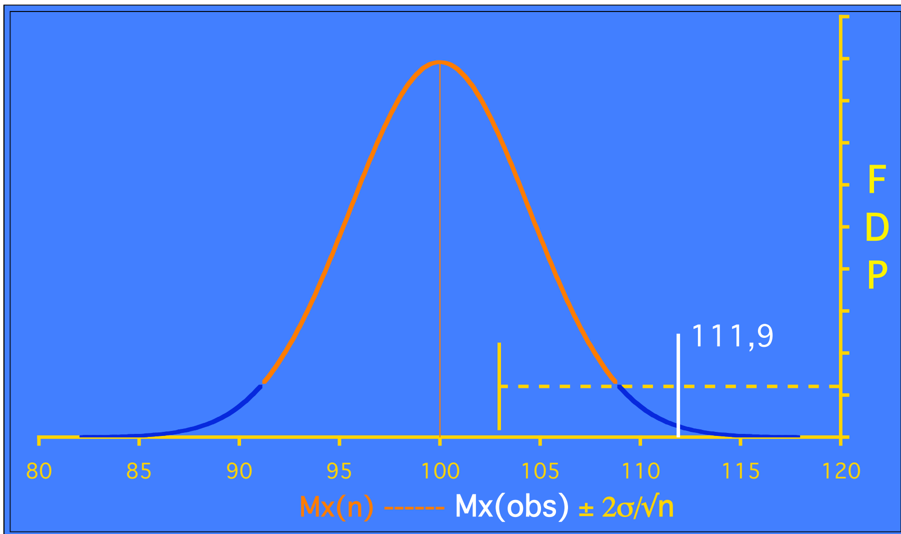
La moyenne d’un échantillon de taille n se trouve dans 95% des cas (1-α) dans l’intervalle $\mu \pm 2\sigma$ (si la variance est connue et que l’on arrondit 1,96 à 2)

(1,96 est la valeur $Z_{0,975}$)



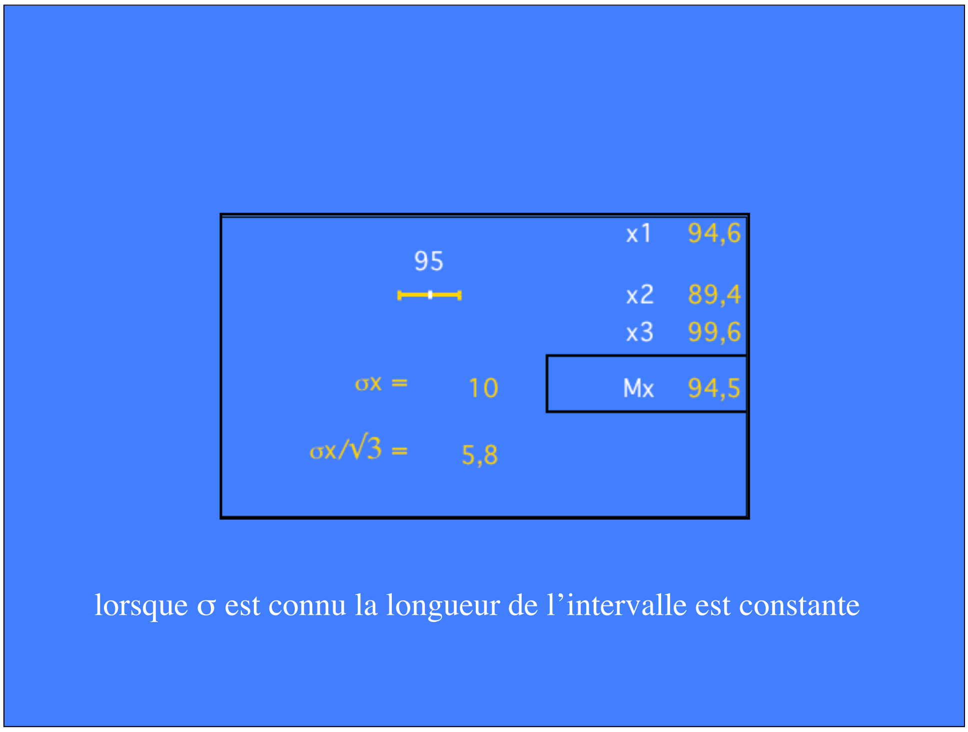
Intervalle de confiance lorsque σ est connu

La moyenne de l'échantillon (106,5) n'est pas égale à celle de la population (100) mais la moyenne de la population est comprise dans l'intervalle approximatif $\pm 2\sigma$ de part et d'autre de la moyenne de l'échantillon.

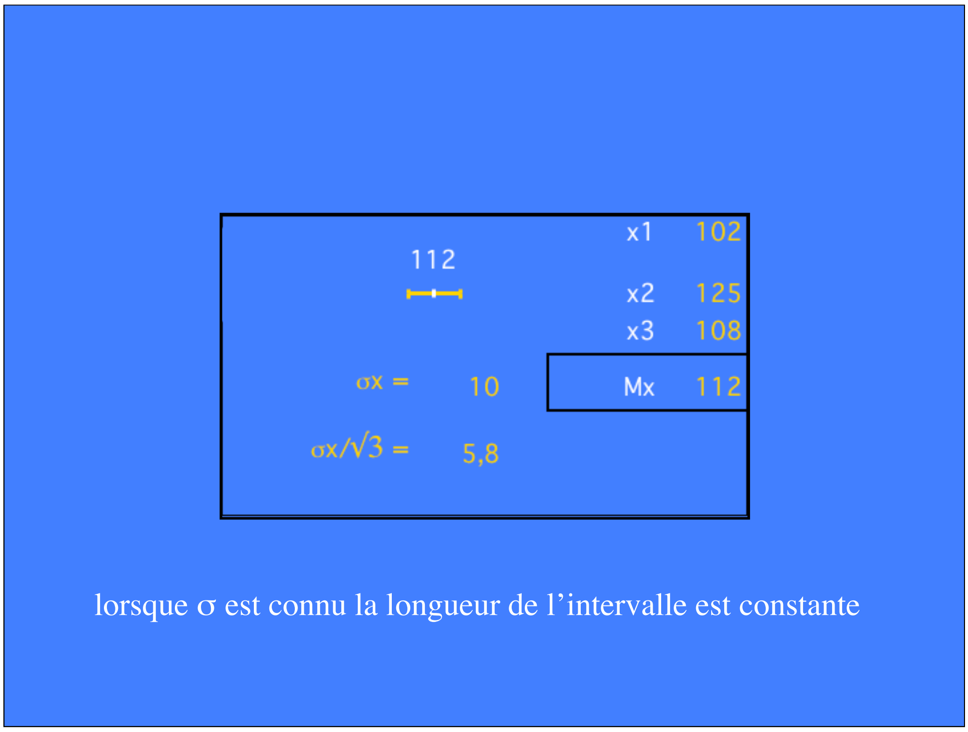


Intervalle de confiance lorsque σ est connu

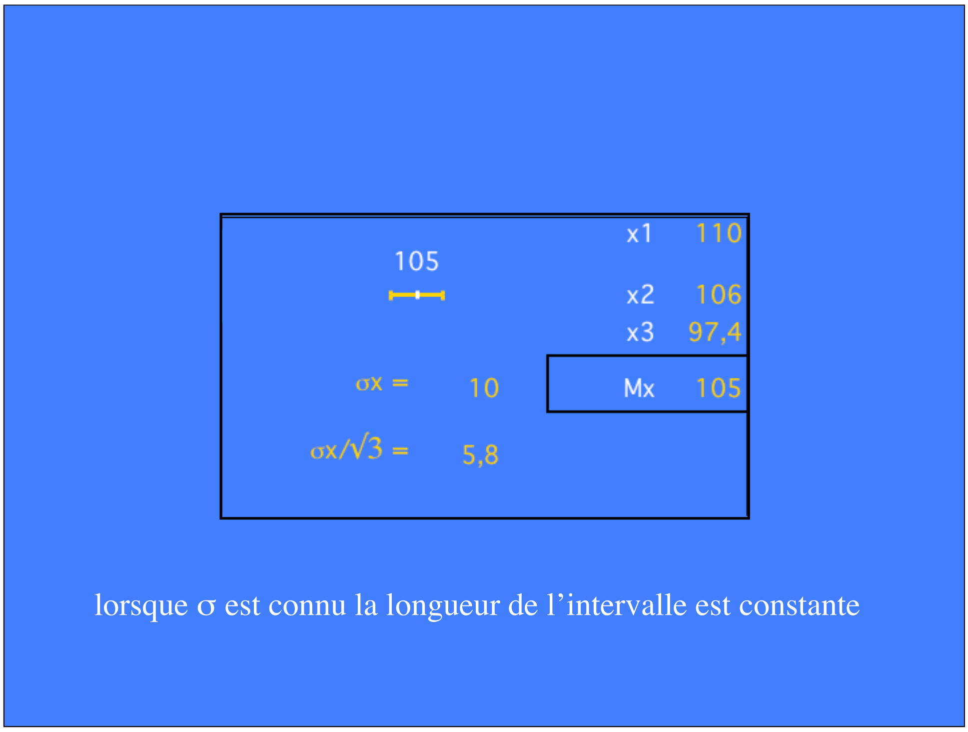
Lorsque la moyenne de l’échantillon (111,9) est particulièrement grande (ou petite) la moyenne de la population <u>n’est plus</u> comprise dans l’intervalle $\pm 2\sigma$ de part et d’autre de la moyenne de l’échantillon. La probabilité que cela arrive est α.



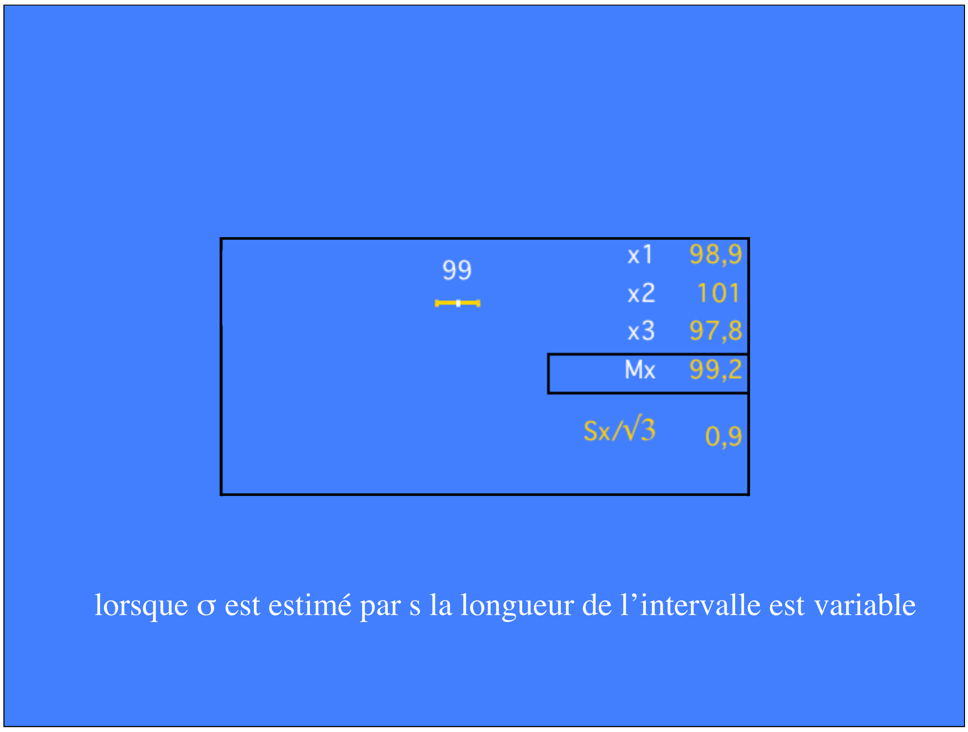
La longueur de l'intervalle de confiance à 95% est constante:
1,96*erreur standard connue (σ est connu et n est connu)



La longueur de l'intervalle de confiance à 95% est constante:
1,96*erreur standard connue (σ est connu et n est connu)

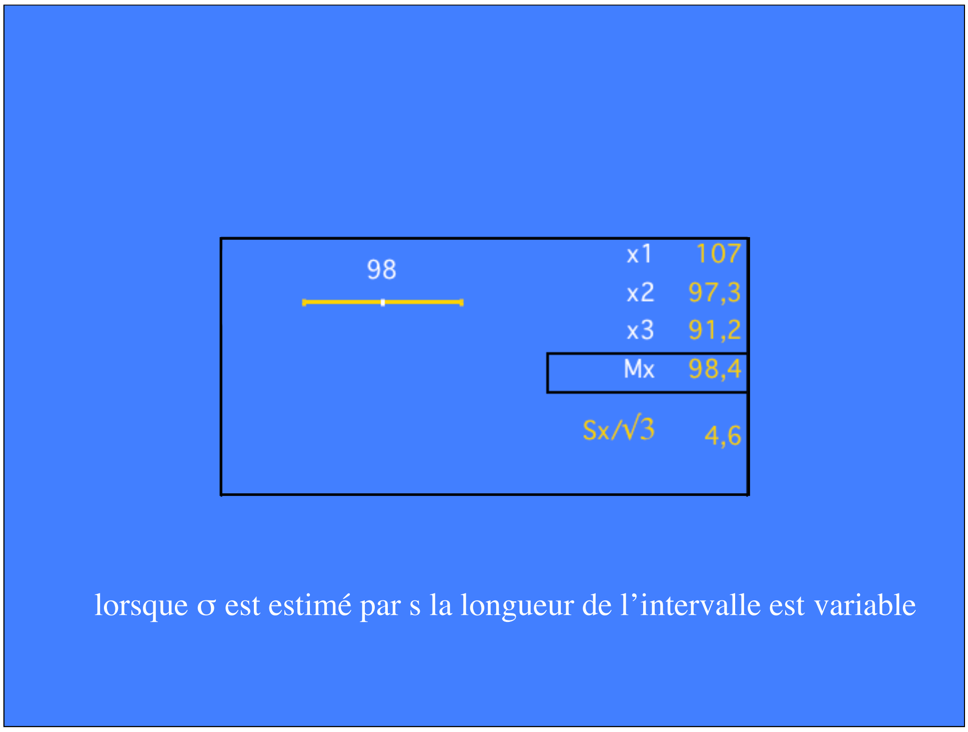


La longueur de l'intervalle de confiance à 95% est constante:
1,96*erreur standard connue (σ est connu et n est connu)



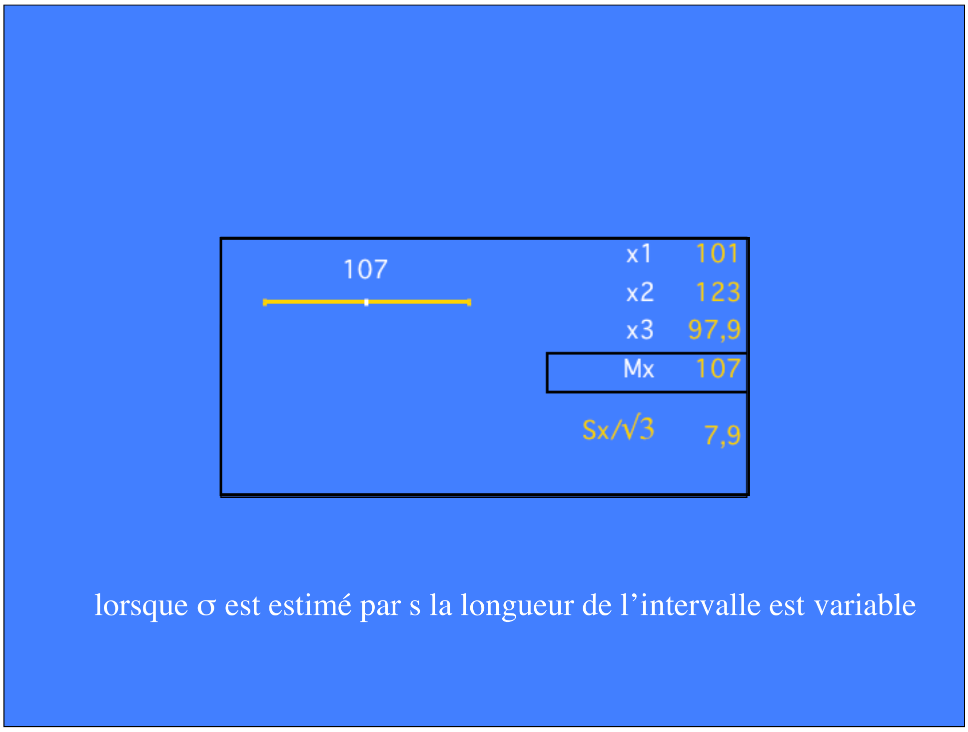
La longueur de l'intervalle de confiance à 95% est variable car Sx est variable:

$t_{1-\alpha/2}$ * erreur standard estimée



La longueur de l'intervalle de confiance à 95% est variable car Sx est variable:

$t_{1-\alpha/2}$ * erreur standard estimée



La longueur de l'intervalle de confiance à 95% est variable car Sx est variable:

$t_{1-\alpha/2}$ * erreur standard estimée

ATTENTION !!

Utiliser l’intervalle de confiance comme test d’hypothèse:

“les moyennes sont significativement différentes lorsque les intervalles de confiance ne se recouvrent pas”

Est extrêmement dangereux car ce raisonnement n’est valable que dans des conditions très strictes, rarement réunies sur un graphique

ATTENTION !!

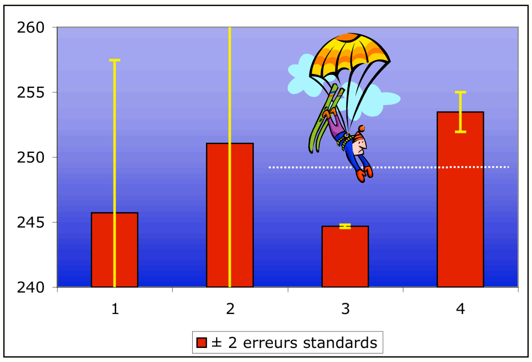
-On compare seulement 2 moyennes
-Le test d'hypothèse est bi directionnel
-Les barres d'erreur du graphique représentent bien l'intervalle de confiance
    - ces barres représentent le plus souvent ±S, ±2S, ±2ES, et non ± t * ES

-Cet intervalle de confiance est calculé:

    -Sur la variance globale des échantillons
    -En référence à une variable t
    -Les variances des populations sont égales

Bref,

-Pour faire un test d'hypothèse en toute sécurité, il faut utiliser les techniques de test d'hypothèse



Simulation:
n = 2
RH0 : 4,9 %
Test graphique : 46%

$$\pm\ 2\sqrt{\frac{\text{variances par groupes}}{n_i}}$$

Le recouvrement des barres égales à

± 2 erreurs standard

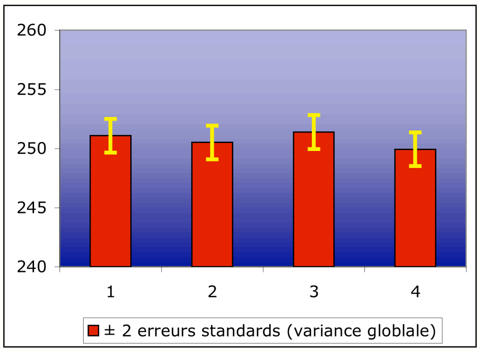
Induit une erreur nettement plus élevée que prévu

Un « test graphique » est réalisé en considérant que les moyennes sont significativement différentes lorsque les barres d’erreur ne se recouvrent pas.

Ce test est très dangereux car il n’est correct que dans des conditions très strictes qui ne sont pas remplies sur cet exemple

Bien que cette pratique ne soit pas recommandée, elle est tellement courante, et publiée, qu’il est indispensable d’en fixer les limites de validité.

Tout d’abord, seul l’intervalle de deux erreurs standards est comparable à un test d’hypothèses bi-directionnel au seuil 5%. Si n est petit et que les variances sont estimées, le test graphique est totalement invalide. Dans la simulation reproduite ci-dessus, le taux d’erreur réel est de 46% au lieu des 5% annoncés.



$$\pm\ 2\sqrt{\frac{variance\ \ globale}{n}}$$

Simulation:
n = 2
RH0 : 7 %
Test graphique : 16%

Egalité des variances

Calcul variance globale

Le recouvrement des barres égales à

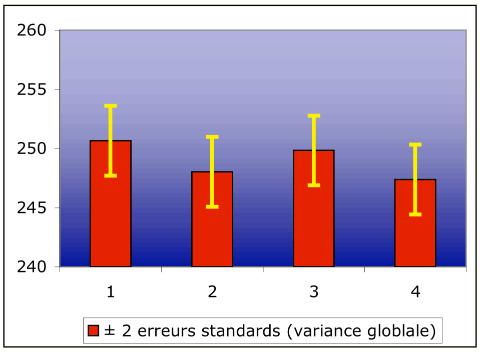
± 2 erreurs standard (variance globale)

erreur plus élevée que prévu

Si l'on calcule une variance globale pour estimer l'erreur standard le test est amélioré.

Un « test d'hypothèse graphique » ne peut pas être réalisé lorsque l'on constate que les barres d'erreurs n'ont pas la même longueur!

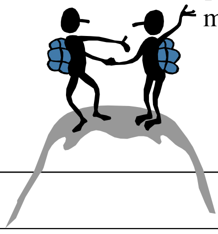
Toutefois lorsque n est petit, la longueur de l'intervalle ne peut pas être estimée par 2 erreurs standards car 2 approxime un $Z_{0,975}$ mais pas un $t_{0,975}$ (le recours au t de Student est dû à l'estimation de la variance).



$$\pm\ t_{1-\alpha/2}\sqrt{\frac{\mathrm{var}\,iance\ \ globale}{n}}$$

n = 2
RH0 ≈ 5 %
Test graphique ≈ 5%
Accord ≈ 100%

Ssi:
Vérifier égalité des variances
Calcul variance globale
Recours à la table de Student
Ne pas comparer plus de 2 moyennes



Enfin, si l'homogénéité des variances est testée, si une variance globale (CMR, d.l. = Nombre d'observations moins nombre de groupes) est estimée et si la longueur de l'intervalle est

$t_{\text{d.l. } 1-a/2}$ alors l'accord entre le test d'hypothèse graphique et le test d'égalité des moyennes est quasi parfait.

Notons cependant qu'en cas de comparaison multiples 1 vs 2, 1 vs 3, 2 vs 3 etc… le seuil alpha n'est plus garanti et que la seule voie possible est celle des contrastes (par exemple contrastes de Scheffé).

En conclusion, la plus grande prudence est recommandée pour interpréter les barres d'erreur comme un test d'hypothèse!