LES TESTS STATISTIQUES
AVEC UNE CALCULATRICE OU GEOGEBRA
- Partie 2 -

Cet article est la suite de l'article *Tests statistiques avec une calculatrice ou Geogebra*, dans lequel nous avions abordé les tests de conformité d'une moyenne et d'une proportion.

Ici, nous présentons l'utilisation et l'interprétation des résultats affichés par les outils de calcul pour 3 types de tests de comparaison (de proportions, de variances et de moyennes) sans détailler le lien avec le travail mené habituellement à l'écrit. Pour cela, il peut être intéressant de lire d'abord la partie 1 dans le bulletin 22…

Test de comparaison de deux proportions

Exemple

Deux types de publicité A et B sont envisagés pour lancer un nouveau produit. Après avoir visionné les deux publicités mises au point par des spécialistes en communication, la direction de l'entreprise émet l'hypothèse que la publicité de type A sera plus efficace que celle de type B.

Question : La direction de l'entreprise a-t-elle raison ?

Pour répondre à cette interrogation, on élabore un test de comparaison de proportions.

Deux régions, considérées comme marché-test (possédant les mêmes caractéristiques de consommation), sont choisies pour évaluer l'efficacité des deux types de publicités. La publicité de type A est utilisée dans une région et celle de type B dans l'autre.

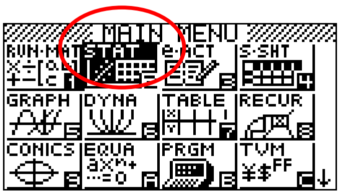
Deux sondages indiquent ensuite que sur 125 personnes ayant vu la publicité A, 47 se sont procuré le produit alors que sur 100 personnes ayant vu la publicité B, 34 ont acheté le produit.

Soit p_1 et p_2 les proportions de personnes intéressées par l'achat du produit dans chacune des deux régions. On cherche à savoir si p_1 est supérieure à p_2.

Avec une calculatrice

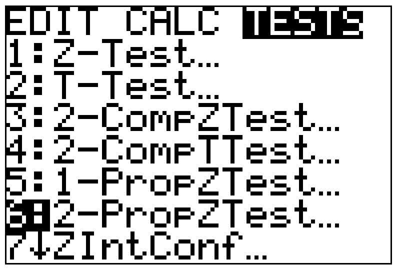
<table>
<tr><th>sur CASIO 85</th><th>sur TI 83 PLUS .fr</th></tr>
<tr><td>

touche **MENU**, option **STAT**



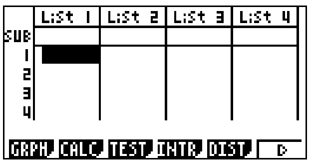
</td><td>

touche **stats**, onglet **TESTS**



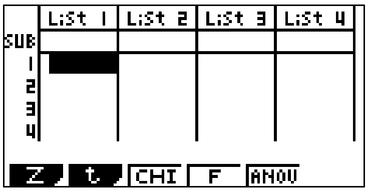
6:2-PropZTest...

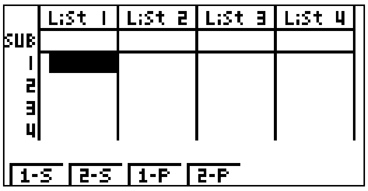
</td></tr>
<tr><td>

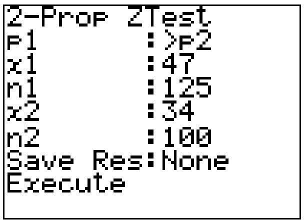
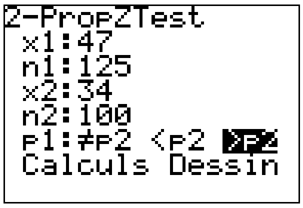


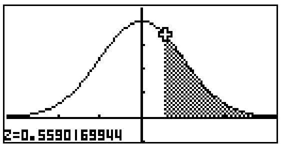
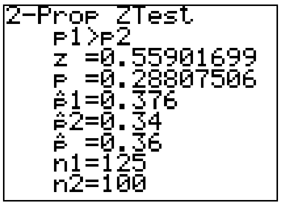
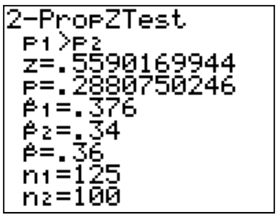
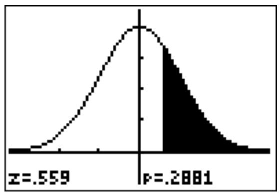
touche **F3** pour l'onglet **TEST**

</td><td></td></tr>
</table>

<table>
<tr><th colspan="2">sur CASIO 85</th><th>sur TI 83 PLUS .fr</th></tr>
<tr><td colspan="2">

Test Z : touche F1 pour l'onglet Z</td><td rowspan="2"></td></tr>
<tr><td colspan="2">

Comparaison de deux proportions :
touche F4 pour l'onglet 2-P</td></tr>
<tr><td>
</td><td>Ici, on complète la fenêtre en fonction des valeurs obtenues avec les échantillons prélevés.</td><td>
</td></tr>
</table>

<table>
<tr><th>DRAW sur CASIO</th><th>CALC sur CASIO</th><th>Calculs sur TI</th><th>Dessin sur TI</th></tr>
<tr><td>
</td><td>
</td><td>
</td><td>
</td></tr>
</table>

La variable de décision est $Z = \dfrac{F_1 - F_2}{\sqrt{F\,(1-F)\left(\dfrac{1}{125}+\dfrac{1}{100}\right)}}$ avec $F = \dfrac{125\,F_1 + 100\,F_2}{125+100}$ où F_1 et F_2 sont les variables aléatoires qui, aux échantillons de tailles respectivement 125 et 100 personnes, issus des régions 1 et 2, associent le pourcentage de personnes ayant acheté le produit. Sous l'hypothèse d'égalité de p_1 et p_2, la loi de Z est approchée par la loi normale centrée réduite.

$\hat{\mathbf{p}}_{\mathbf{1}}$ et $\hat{\mathbf{p}}_{\mathbf{2}}$ sont les valeurs observées de F_1 et F_2 respectivement, c'est-à-dire $\hat{\mathbf{p}}_{\mathbf{1}} = \dfrac{47}{125} = 0{,}376$ et $\hat{\mathbf{p}}_{\mathbf{2}} = \dfrac{34}{100} = 0{,}34$.

$\hat{\mathbf{p}}$ est la valeur observée de F, c'est-à-dire $\dfrac{125\,\frac{47}{125} + 100\,\frac{34}{100}}{125+100} = \dfrac{81}{225} = 0{,}36$.

Z est la valeur observée de Z c'est-à-dire $\dfrac{\dfrac{47}{125}-\dfrac{34}{100}}{\sqrt{0{,}36\,(1-0{,}36)\left(\dfrac{1}{125}+\dfrac{1}{100}\right)}} \approx 0{,}559$

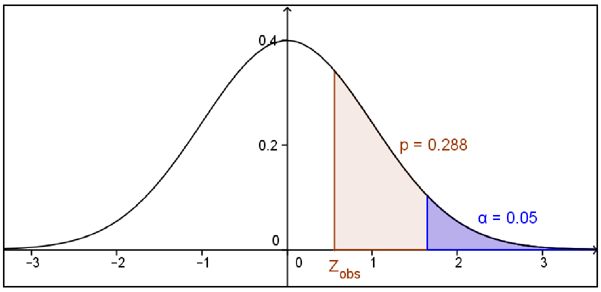
p est la probabilité que, sous l'hypothèse d'égalité de p_1 et p_2, la variable aléatoire Z prenne une valeur supérieure à **Z**. Ici, **p** = $P(Z > 0{,}559\,016\,99) \simeq 0{,}288\,075\,06$.

Exploitons les résultats donnés par la calculatrice pour rédiger une règle de décision en comparant α avec la *p-value* **p**.

- Si $\alpha > 0{,}288\,075\,06$, on rejette l'hypothèse d'égalité de p_1 et p_2
- Si $\alpha \leq 0{,}288\,075\,06$, on n'est pas en mesure de rejeter l'hypothèse d'égalité de p_1 et p_2

Revenons à notre exemple :

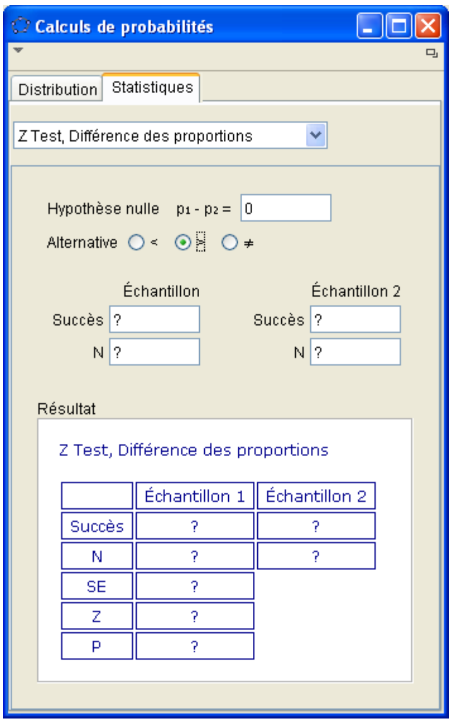
Si $\alpha = 0{,}05$: $\alpha <$ **p**, on n'est pas en mesure de rejeter l'hypothèse d'égalité de p_1 et p_2.

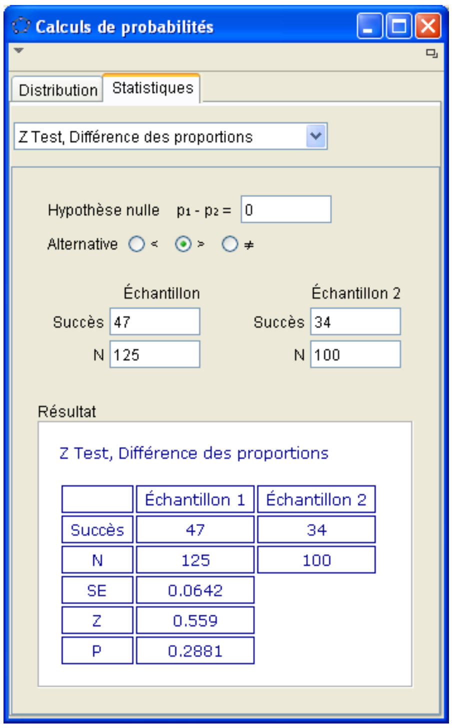


Avec Geogebra

Dans le cas où on connaît les échantillons par les proportions de la modalité étudiée, on utilise l'onglet **Statistiques** proposé dans le module **Calculs de probabilités**.

On choisit la commande **ZTest, Différence de proportions**.



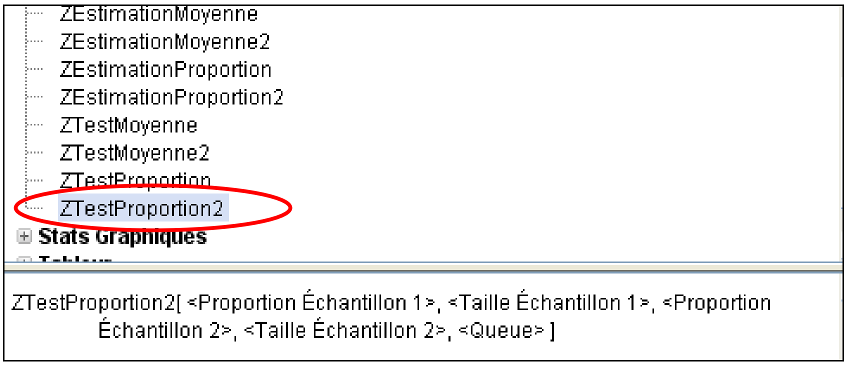


SE = $\sqrt{0{,}36\,(1-0{,}36)\left(\dfrac{1}{125}+\dfrac{1}{100}\right)} \approx 0{,}0642$.

Pour conclure, il reste à comparer α avec la p-value p.

Avec Geogebra (bis)

On peut aussi utiliser la commande **ZTestProportion2**. On affiche la liste des commandes de la rubrique **Statistiques**.



La syntaxe est :

ZTestProportion2[Proportion1,Taille1,proportion2, Taille2,Hypothèse]

Proportion1 et ***Taille1*** désignent la proportion d'individu de l'échantillon 1 ayant la modalité étudiée et la taille de l'échantillon 1.

Proportion2 et ***Taille2*** désignent la proportion d'individu de l'échantillon 2 ayant la modalité étudiée et la taille de l'échantillon 2.

Hypothèse prend une des trois formes : "<", ">" ou "≠" selon que l'hypothèse alternative est "$p_1 < p_2$", "$p_1 > p_2$" ou "$p_1 \neq p_2$".

La commande donne le résultat sous la forme d'une liste : $\{p\text{-}value, Z_{obs}\}$ où *p-value* est l'une des trois probabilités $P(T \leq Z_{obs})$, $P(T \geq Z_{obs})$ ou $P(|Z| \geq |Z_{obs}|)$ selon que l'hypothèse alternative est "$p_1 < p_2$", "$p_1 > p_2$" ou "$p_1 \neq p_2$" où la loi de Z est approchée par la loi normale centrée réduite.

Reprenons notre exemple :

Ici, la formule est ***ZTestProportion2[0.376,125,0.34,100,">"]*** à écrire dans la ligne de saisie. On obtient $\{0.288, 0{,}559\}$.

Pour conclure, il reste à comparer α avec la p-value p.

Test de comparaison de deux variances

Exemple

Une entreprise agroalimentaire qui conditionne des frites surgelées souhaite comparer le taux de matière sèche (M.S.) de deux variétés de pommes de terre A et B.
On admet que les taux de matière sèche des pommes de terre de l'une et l'autre des variétés sont distribués selon des lois normales.

On va comparer les taux de matière sèche à l'aide d'un test de comparaison de moyennes. La nature du test à utiliser dépend de l'égalité ou non des variances σ_1^2 et σ_2^2 des taux de matière sèche des pommes de terre des variétés A et B sont différentes.
Dans un premier temps, on élabore un test permettant de savoir si σ_1^2 et σ_2^2 sont égales.

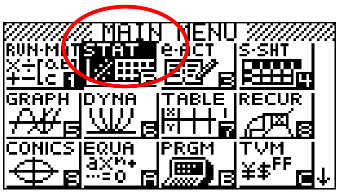
Les résultats des échantillons prélevés et analysés sont donnés dans le tableau suivant.

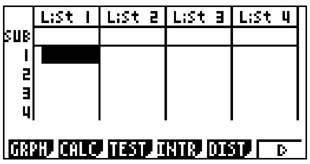
Variété	Taille de l'échantillon	Taux moyen de M.S.	Écart type corrigé du taux de M.S.
A	$n_1 = 16$	$\overline{x_1} = 0{,}23$	$s_1 = 0{,}075$
B	$n_2 = 31$	$\overline{x_2} = 0{,}19$	$s_2 = 0{,}057$

Avec une calculatrice

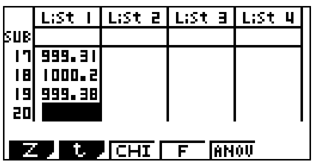
sur CASIO 85

touche **MENU**, option **STAT**





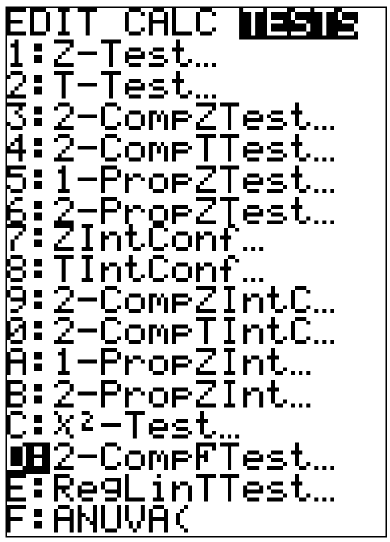
touche **F3** pour l'onglet **TEST**



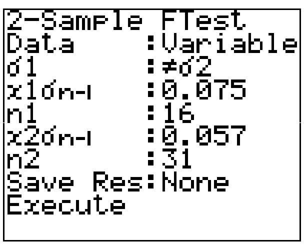
Test F : touche **F4** pour l'onglet **F**

sur TI 83 PLUS .*fr*

touche **stats**, onglet **TESTS**

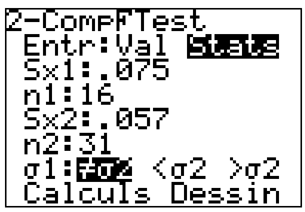


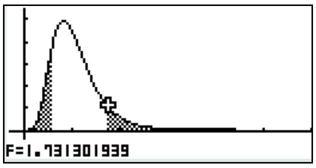
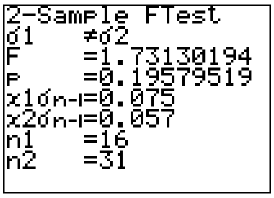
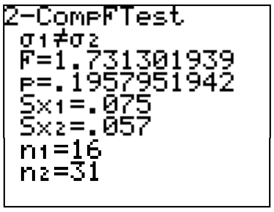
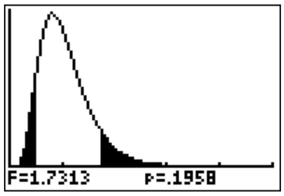
D:2-CompFTest...



Ici, on complète la fenêtre en fonction des valeurs obtenues avec les échantillons prélevés.

Dans le cas où on dispose des valeurs brutes des échantillons prélevés, saisies dans des listes :
Data : List (Casio) ou **Entr : Val** (T.I)



<table>
<tr><th>DRAW sur CASIO</th><th>CALC sur CASIO</th><th>Calculs sur TI</th><th>Dessin sur TI</th></tr>
<tr><td>
On peut aussi faire apparaître $\frac{1}{F}$ en utilisation les flèches de déplacement horizontal de la calculatrice.</td><td></td><td></td><td></td></tr>
</table>

La variable de décision est $F = \dfrac{\widehat{S_1}^2}{\widehat{S_2}^2}$ où $\widehat{S_1}$ et $\widehat{S_2}$ sont les variables aléatoires qui, aux échantillons de tailles respectivement 16 et 31, issus des productions A et B respectivement, associent l'écart type corrigé du caractère dans l'échantillon. Sous l'hypothèse d'égalité de σ_1 et σ_2, la loi de probabilité de F est la loi de Fisher-Snedecor à 15 et 30 degrés de liberté.

$\mathbf{x_1\sigma_{n-1}}$ ou $\mathbf{Sx_1}$ est la valeurs observées de $\widehat{S_1}$.

$\mathbf{x_2\sigma_{n-1}}$ ou $\mathbf{Sx_2}$ est la valeurs observées de $\widehat{S_2}$.

F est la valeur observée de F, c'est-à-dire $\dfrac{0{,}075^2}{0{,}057^2} \simeq 1{,}731\,301\,94$.

p est le double de la probabilité que, sous l'hypothèse d'égalité de σ_1 et σ_2, la variable aléatoire F prenne une valeur supérieure à F_{obs} (ou inférieure à F_{obs} selon la position de F_{obs} par rapport à 1). Le double se justifie par le fait que le test est bilatéral.

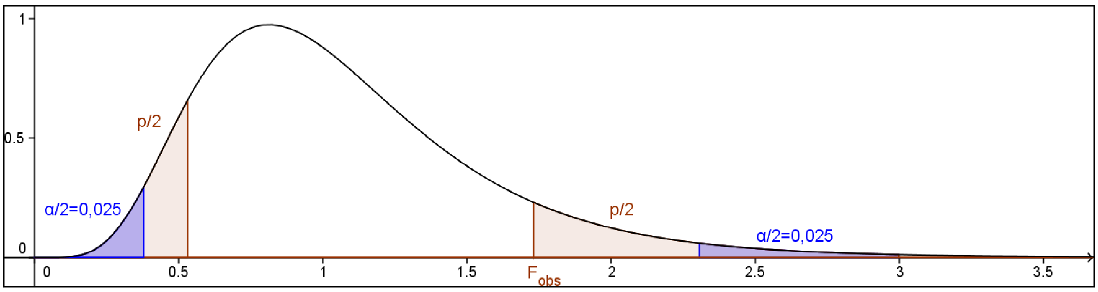
Ici, $\mathbf{p} = 2\,P(F > 1{,}731\,301\,94) \simeq 0{,}195\,795\,19$.

Exploitons les résultats donnés par la calculatrice pour rédiger une règle de décision en comparant α avec la *p-value* **p**.

- Si $\alpha > \mathbf{p}$, on rejette l'hypothèse d'égalité de σ_1 et σ_2
- Si $\alpha \leq \mathbf{p}$, on n'est pas en mesure de rejeter l'hypothèse d'égalité de σ_1 et σ_2

Revenons à notre exemple :

Si $\alpha = 0{,}05$: $\alpha < 0{,}195\,795\,19$, on n'est pas en mesure de rejeter l'hypothèse d'égalité de σ_1 et σ_2.



Avec Geogebra

Il n'est pas programmé de test de comparaison de variances !

Test de comparaison de deux moyennes

Avant d'utiliser la calculatrice, on s'interroge… ! Les variances du caractère dans les deux populations sont-elles connues ou pas ?

Variances inconnues mais égales

Exemple

Reprenons l'exemple précédent où l'on a élaboré un test de comparaison de variances et dont l'application aux échantillons prélevés nous a permis de ne pas rejeter l'hypothèse d'égalité des variances.

On élabore un test permettant de conclure si le taux de M.S. moyen μ_1 des pommes de terre de la variété A est supérieur à μ_2, celui des pommes de terre de la variété B (ou non).

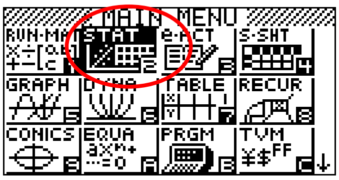
Les résultats des échantillons prélevés et analysés sont donnés dans le tableau suivant.

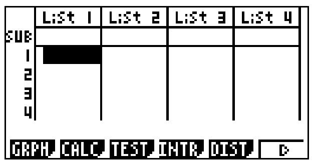
Variété	Taille de l'échantillon	Taux moyen de M.S.	Écart type corrigé du taux de M.S.
A	$n_1 = 16$	$\overline{x_1} = 0{,}23$	$s_1 = 0{,}075$
B	$n_2 = 31$	$\overline{x_2} = 0{,}19$	$s_2 = 0{,}057$

Avec une calculatrice

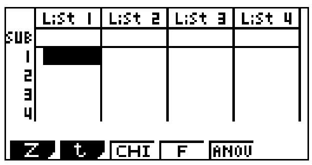
sur CASIO 85

touche **MENU**, option **STAT**



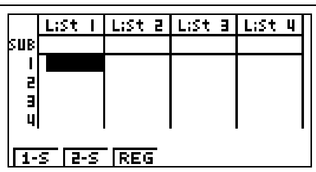


touche **F3** pour l'onglet **TEST**



Les variances du caractères sont inconnues :

Test t : touche **F2** pour l'onglet **t**

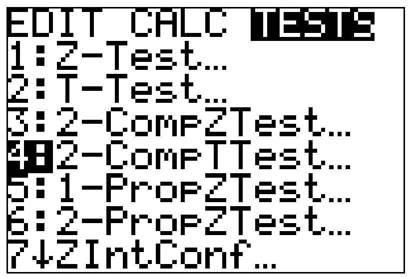


Comparaison de deux moyennes

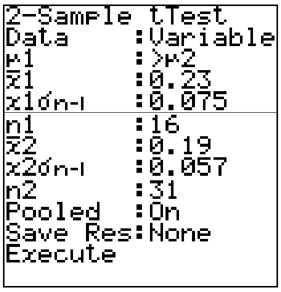
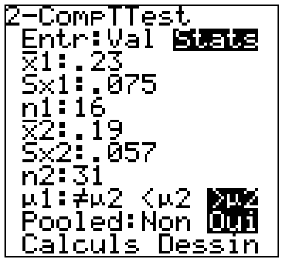
touche **F2** pour l'onglet **2-S**

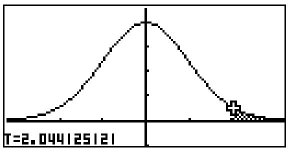
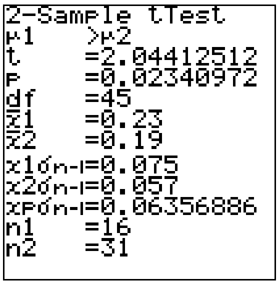
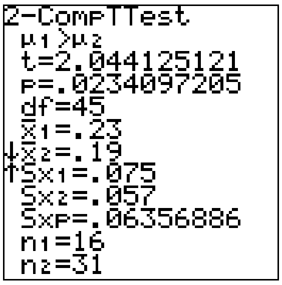
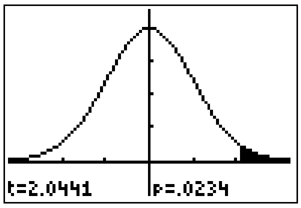
sur TI 83 PLUS *.fr*

touche **stats**, onglet **TESTS**



4:2-CompTTest...

<table>
<tr><th>sur CASIO 85</th><th colspan="2">sur TI 83 PLUS .fr</th></tr>
<tr><td></td><td>On complète la fenêtre avec les paramètres calculés sur les échantillons prélevés.
Si on dispose des valeurs brutes des échantillons prélevés, saisies dans des listes : Data : List (Casio) ou Entr : Val (T.I)
Ici, l'option Pooled est On ou Yes car les deux variances sont égales (dans le cas où elles sont différentes, on choisit Off ou No).</td><td></td></tr>
</table>

DRAW sur CASIO	CALC sur CASIO	Calculs sur TI	Dessin sur TI
			

On est dans le cas où les variances des distributions des deux populations sont égales, la variable de décision est $T = \dfrac{\overline{X}_1 - \overline{X}_2}{\hat{S}\sqrt{\dfrac{1}{16}+\dfrac{1}{31}}}$ où $\hat{S} = \sqrt{\dfrac{15\,\hat{S}_1^{\,2} + 30\,\hat{S}_2^{\,2}}{16+31-2}}$ avec $\overline{X}_1$, $\overline{X}_2$, $\hat{S}_1$ et $\hat{S}_2$ les variables aléatoires qui, aux échantillons de tailles respectives n_1 et n_2, issus des productions des deux machines, associent la moyenne et l'écart type corrigé du taux de M.S. dans l'échantillon. Comme le caractère étudié est distribué selon une loi normale dans les deux populations, la loi de T est la loi de Student à $n_1 + n_2 - 2$ degrés de liberté.

***x*pσn-1** ou **Sxp** est la valeur observée de $\hat{S}$ sur les échantillons prélevés, c'est-à-dire :

$$\sqrt{\frac{15 \times 0{,}075^2 + 30 \times 0{,}057^2}{45}} \simeq 0{,}063\,568\,86$$

t est la valeur de T observée sur l'échantillon : $\mathbf{t} \simeq \dfrac{0{,}23 - 0{,}19}{0{,}063\,568\,86\sqrt{\dfrac{1}{16}+\dfrac{1}{31}}} \simeq 2{,}044\,125\,12$.

p est la probabilité que, sous l'hypothèse alternative "$\mu_1 > \mu_2$", la variable aléatoire T prenne une valeur supérieure à **t**.

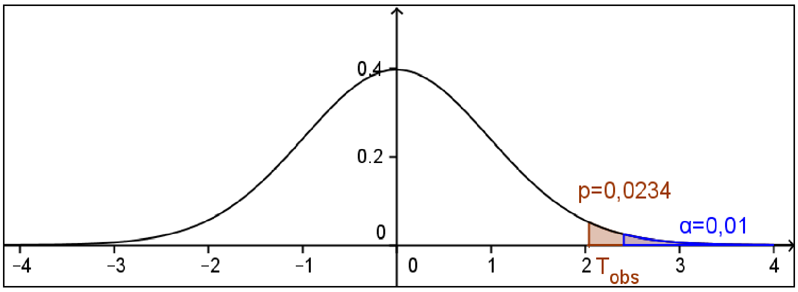
Ici, $\mathbf{p} = \mathrm{T}(T > 2{,}044\,125\,12) \simeq 0{,}023\,409\,72$, la loi de probabilité de T étant la loi de Student à **df** $= 16 + 31 - 2 = 45$ degrés de liberté.

Exploitons les résultats donnés par la calculatrice pour rédiger une règle de décision en comparant α avec la *p-value* **p**.

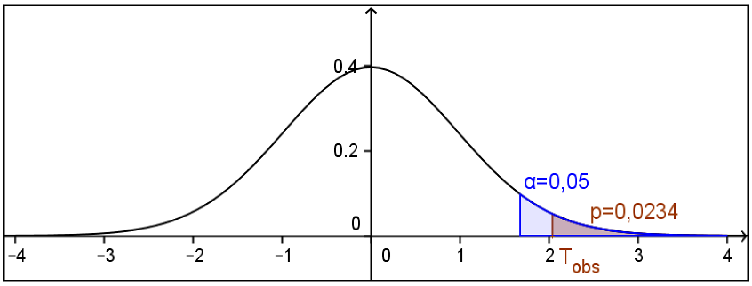
- Si $\alpha > \mathbf{p}$, on rejette l'hypothèse d'égalité de μ_1 et μ_2
- Si $\alpha \leq \mathbf{p}$, on n'est pas en mesure de rejeter l'hypothèse d'égalité de μ_1 et μ_2

Revenons à notre exemple :

Si $\alpha = 0,01$: $\alpha < 0,023\,409\,72$, on n'est pas en mesure de rejeter l'hypothèse d'égalité de μ_1 et μ_2.



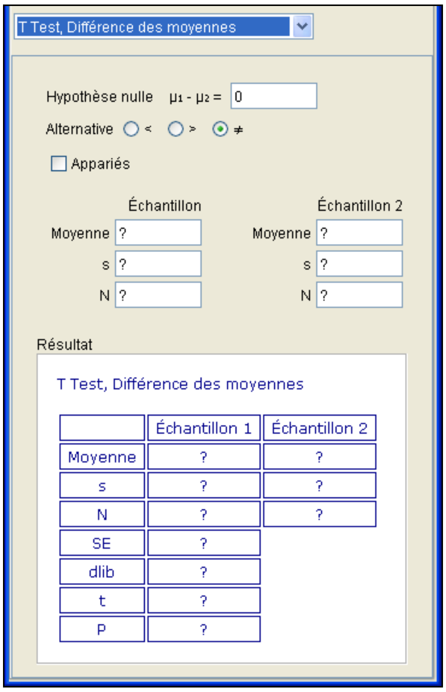
Si $\alpha = 0,05$: $\alpha > 0,023\,409\,72$, on rejette l'hypothèse d'égalité de μ_1 et μ_2.

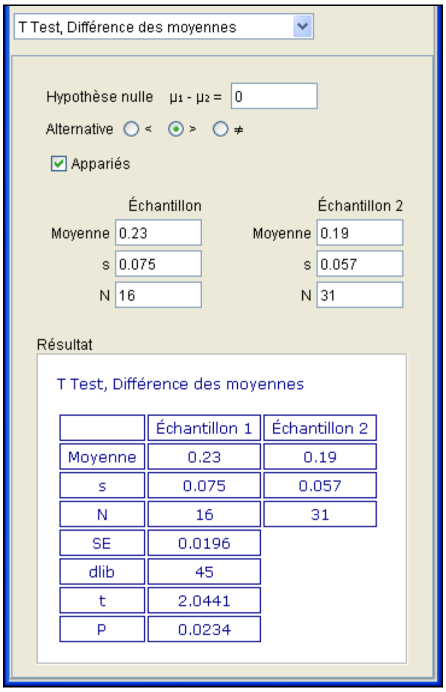


Avec Geogebra

Dans le cas où on connaît les moyennes et les variances corrigées du caractère dans les deux échantillons, on utilise l'onglet **Statistiques** proposé dans le module **Calculs de probabilités**.

On choisit la commande **T Test, Différence des moyennes**.



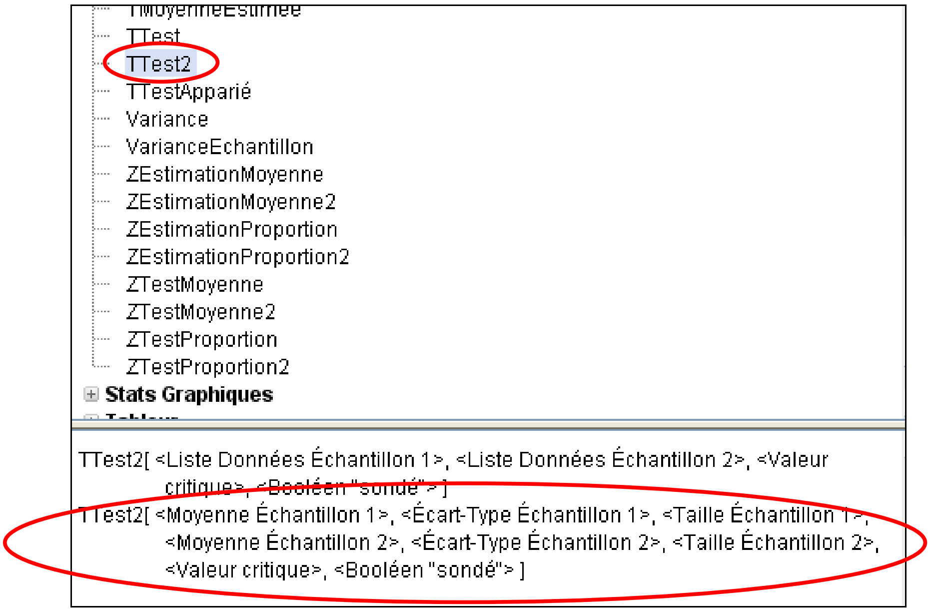


On sélectionne *appariés* car les deux variances du caractère étudié dans les deux populations sont supposées égales et **SE** $= 0,063\,568\,86\sqrt{\frac{1}{16}+\frac{1}{31}} \simeq 0,019\,6$.

Pour conclure, il reste à comparer α avec la p-value *p*.

Avec Geogebra (bis)

On peut aussi utiliser la commande **TTest2**. On affiche la liste des commandes de la rubrique **Statistiques**. Comme on ne connaît des échantillons que leurs moyennes et leurs écarts types corrigés(on ne dispose pas des données brutes), on utilise la deuxième option proposée.



La syntaxe est :

TTest2[Moyenne1,Écart-type1,Taille1,Moyenne2,Écart-type2,Taille2,Hypothèse,Égalité]

Moyenne1, ***Écart-type1*** et ***Taille1*** désignent la moyenne, l'écart-type corrigé de l'échantillon 1 et sa taille.

Moyenne2, ***Écart-type2*** et ***Taille2*** désignent la moyenne, l'écart-type corrigé de l'échantillon 2 et sa taille.

Hypothèse prend une des trois formes : <, > ou ≠ selon que H_1 est "$\mu_1 < \mu_2$", "$\mu_1 > \mu_2$" ou "$\mu_1 \neq \mu_2$".

Égalité est une variable booléenne (prenant les valeurs *True* ou *False*) qui indique si les variances de la variable dans les deux populations sont égales ou non.

La commande donne le résultat sous la forme d'une liste : $\{p\text{-}value, t_{obs}\}$ où *p-value* est l'une des trois probabilités $P(T \leq t_{obs})$, $P(T \geq t_{obs})$ ou $P(|T| \geq |t_{obs}|)$ selon que l'hypothèse alternative est "$\mu_1 < \mu_2$", "$\mu_1 > \mu_2$" ou "$\mu_1 \neq \mu_2$" où T suit la loi de Student avec $n_1 + n_2 - 2$ degrés de liberté.

Reprenons notre exemple :

Ici, la formule est **TTest2[0.23,0.075,16,0.19,0.057,31,">",true]** à écrire dans la ligne de saisie. On obtient {0.023409717,2.0441251206}.

Pour conclure, il reste à comparer α avec la p-value *p*.

Remarque :

Si on dispose des données brutes des deux échantillons, on peut en créer deux listes et utiliser la première option de la commande **TTest2** dont la syntaxe est :

TTest2[Liste1,Liste2,Hypothèse,Égalité].

Variances inconnues mais différentes

Exemple

Dans une coopérative agricole, on désire tester l'effet d'un engrais sur la production de blé. Pour cela, on choisit 24 lots de terrain de même superficie. La moitié de ces parcelles est traitée avec l'engrais et l'autre moitié ne l'est pas (c'est le groupe de référence). On suppose normale la distribution des productions dans les lots.

La production moyenne de blé obtenue sur les lots non traités est de 4,8 t avec un écart type corrigé de 0,48 t, tandis que la moyenne obtenue sur les lots traités est de 5,2 t avec un écart type corrigé de 0,25 t.

Question : Peut-on admettre que l'engrais améliore la production ?

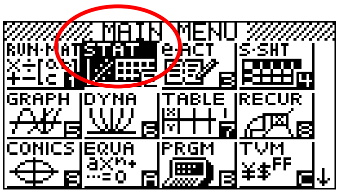
Pour répondre à cette interrogation, on élabore un test de comparaison de moyennes.

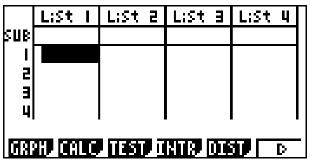
Les variances des productions des parcelles sont inconnues et dans un premier temps, on élabore un test d'égalité de ces variances :

Avec une calculatrice

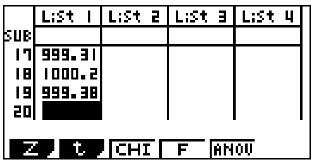
sur CASIO 85

touche **MENU**, option **STAT**





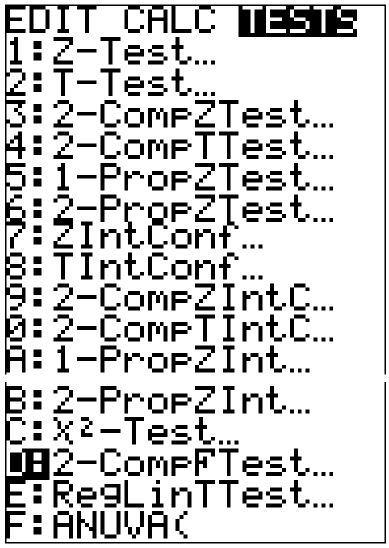
touche **F3** pour l'onglet **TEST**



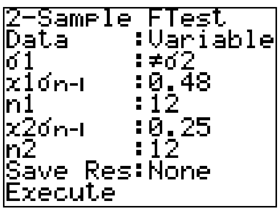
Test F : touche **F4** pour l'onglet **F**

sur TI 83 PLUS .fr

touche **stats**, onglet **TESTS**

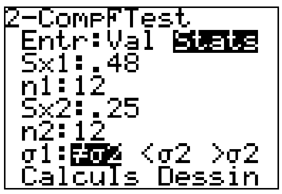


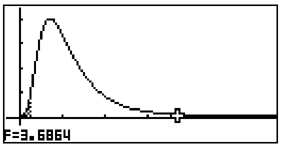
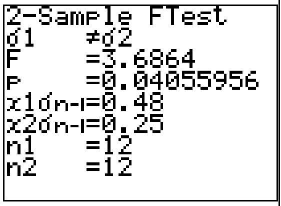
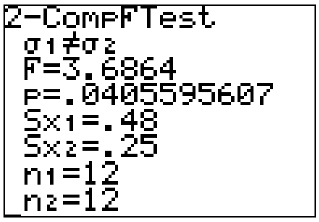
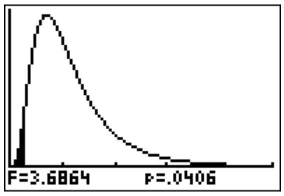
D:2-CompFTest...



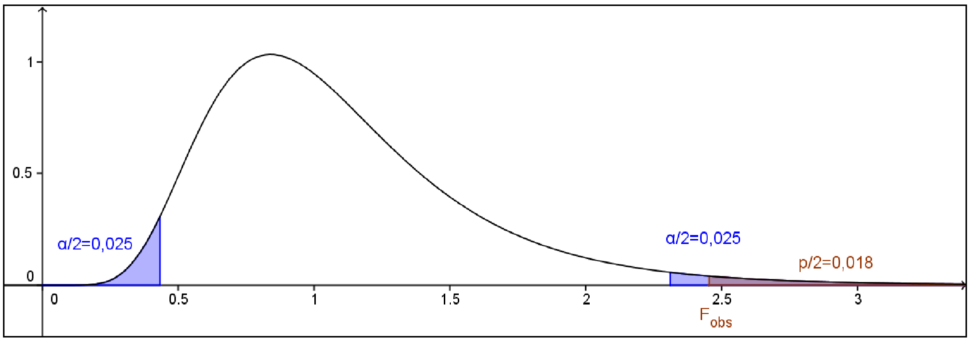
Ici, on complète la fenêtre en fonction des valeurs obtenues avec les échantillons prélevés.

Dans le cas où on dispose des valeurs brutes des échantillons prélevés, saisies dans des listes : **Data : List** (Casio) ou **Inpt : Data** (T.I)



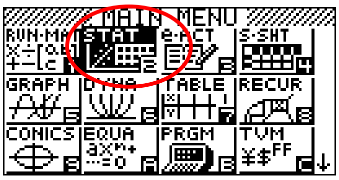
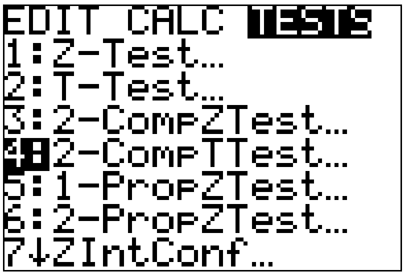
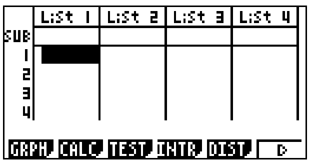
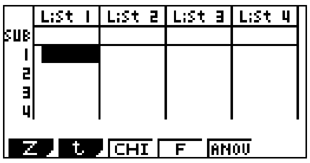
DRAW sur CASIO	CALC sur CASIO	Calculs sur TI	Dessin sur TI
 On peut aussi faire apparaître $\frac{1}{F}$ en utilisation les flèches de déplacement horizontal de la calculatrice.			

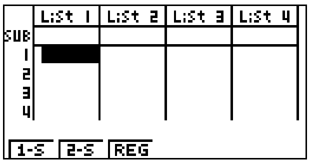
Comme α = 0,05 : α > 0,040 956 7, on rejette l'hypothèse d'égalité de σ_1 et σ_2.

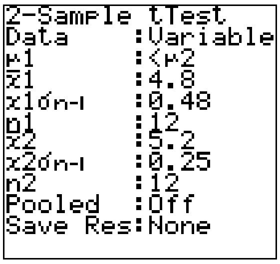


Dans un second temps, on élabore un test d'égalité des moyennes :

Avec une calculatrice

sur CASIO 85	sur TI 83 PLUS *.fr*
touche **MENU**, option **STAT** 	touche **stats**, onglet **TESTS**  **4:2-CompTTest...**
 touche **F3** pour l'onglet **TEST**	
 Les variances du caractères sont inconnues : **Test t :** touche **F2** pour l'onglet **t**	

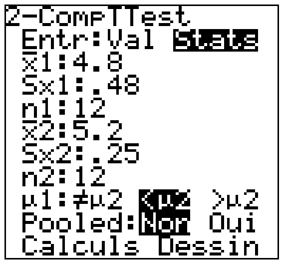
sur CASIO **85**	**sur** TI **83** PLUS ***.fr***
 **Comparaison de deux moyennes** touche $\boxed{\textbf{F2}}$ pour l'onglet $\boxed{\textbf{2-S}}$	

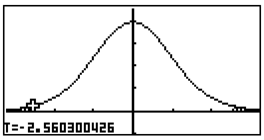
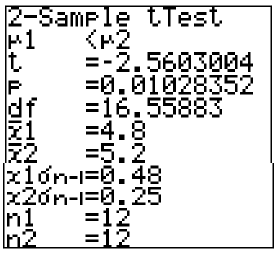
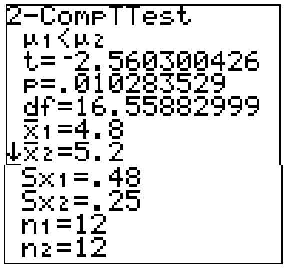
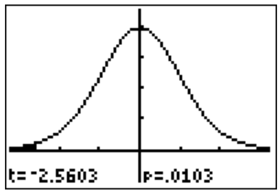


On complète la fenêtre avec les paramètres calculés sur les échantillons prélevés.

Si on dispose des valeurs brutes des échantillons prélevés, saisies dans des listes : **Data : List** (Casio) ou **Inpt : Data** (T.I)

Ici, l'option **Pooled** est **Off** ou **No** car les variances des deux populations sont différentes.



DRAW sur CASIO	**CALC sur** CASIO	**Calculs sur** TI	**Dessin sur** TI
			

On est dans le cas où les variances des distributions du rendement des deux populations de parcelles sont différentes, la variable de décision est $T = \dfrac{\overline{X_1} - \overline{X_1}}{\sqrt{\dfrac{\widehat{S_1}^2}{12} + \dfrac{\widehat{S_2}^2}{12}}}$ où $\overline{X_1}$, $\overline{X_2}$, $\widehat{S_1}$ et $\widehat{S_2}$ sont les variables aléatoires qui, aux échantillons de 12 parcelles respectivement traitées et non traitées en engrais, associent la moyenne et l'écart type corrigé du rendement.

Sous l'hypothèse d'égalité des moyennes, la loi de T est approchée par la loi de Student à ν degrés de liberté avec ν l'entier le plus proche de $\mathbf{df} = \dfrac{\left(\dfrac{0{,}48^2}{12} + \dfrac{0{,}25^2}{12}\right)^2}{\dfrac{\left(\dfrac{0{,}48^2}{12}\right)^2}{11} + \dfrac{\left(\dfrac{0{,}25^2}{12}\right)^2}{11}}$.

t est la valeur de T observée sur l'échantillon : $\mathbf{t} \simeq \dfrac{4{,}8 - 5{,}2}{\sqrt{\dfrac{0{,}48^2}{12} + \dfrac{0{,}25^2}{12}}} \simeq -2{,}560\,300\,4$.

p est la probabilité que, sous l'hypothèse alternative "$\mu_1 > \mu_2$", la variable aléatoire T prenne une valeur inférieure à **t**.

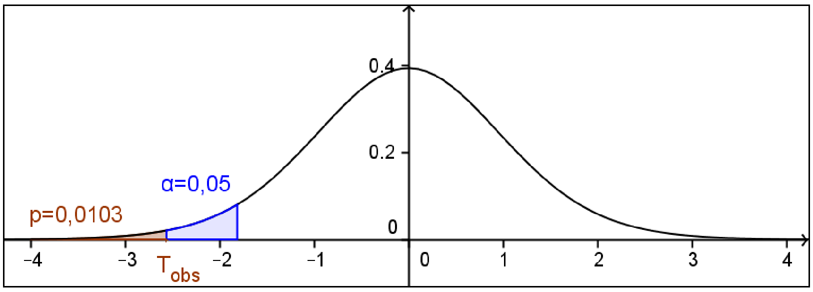
Ici, $\mathbf{p} = P(T < -2{,}560\,300\,4) \simeq 0{,}010\,283\,52$, la loi de probabilité de T étant la loi de Student à 17 degrés de liberté.

Exploitons les résultats donnés par la calculatrice pour rédiger une règle de décision en comparant α avec la *p-value* **p**.

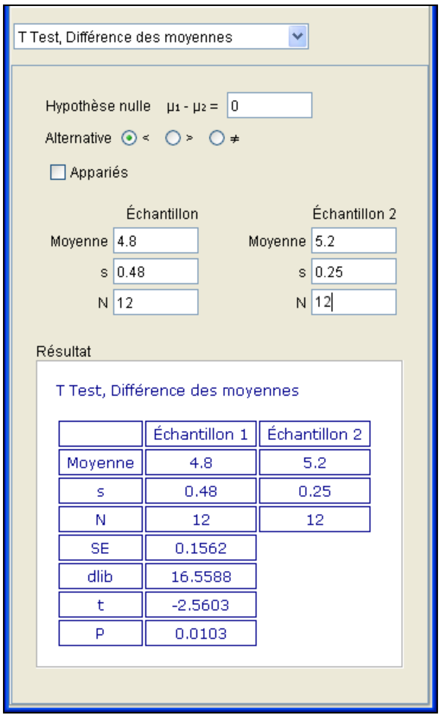
- Si $\alpha > \mathbf{p}$, on rejette l'hypothèse d'égalité de μ_1 et μ_2.
- Si $\alpha \leq \mathbf{p}$, on n'est pas en mesure de rejeter l'hypothèse d'égalité de μ_1 et μ_2

Revenons à notre exemple :

Si $\alpha = 0{,}05$: $\alpha > 0{,}010\,283\,52$, on rejette l'hypothèse d'égalité de μ_1 et μ_2.



Avec Geogebra



Sans hypothèse d'égalité des variances du caractère étudié dans les deux populations et dans le cas où on connaît les moyennes et les variances corrigées du caractère dans les deux échantillons, on utilise l'onglet **Statistiques** proposé dans le module **Calculs de probabilités**.

On choisit la commande :

TTest, Différence des moyennes.

$$\mathbf{SE} = \sqrt{\frac{0{,}48^2}{12} + \frac{0{,}25^2}{12}} \approx 0{,}156\,2.$$

Pour conclure, il reste à comparer α avec la p-value p.

Avec Geogebra (bis)

On peut aussi utiliser la commande **TTest2**.

Ici, la formule est **TTest2[0.23,0.075,16,0.19,0.057,31,"<",false]** à écrire dans la ligne de saisie. On obtient {0.963,1.872}.

Pour conclure, il reste à comparer α avec la p-value p.

Variances connues

On utilise un test de Student.

Exemple (D'après un sujet du B.T.S.A I.A.A)

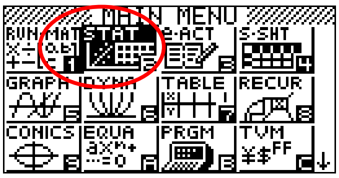
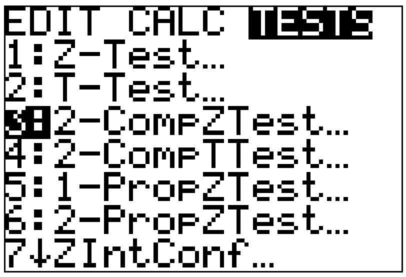
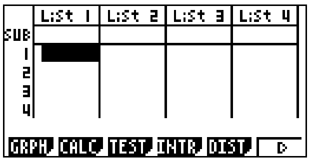
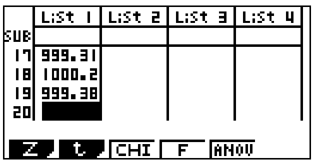
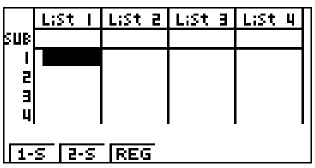
On désire comparer le travail de deux doseuses pour boîtes de haricots verts de quantité nominale égale à 800 g. Les deux distributions sont supposées normales et les écarts types des deux machines sont : $\sigma_1 = 20$ g et $\sigma_2 = 16$ g. On prélève un échantillon de 20 éléments sur chacune des deux machines, ce qui donne les deux valeurs moyennes suivantes :

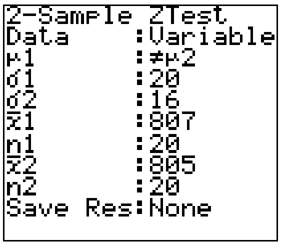
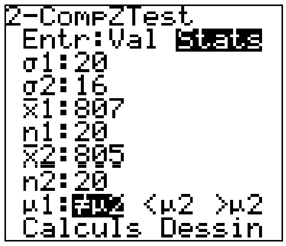
$\bar{x}_1 = 807$ g pour la première machine et

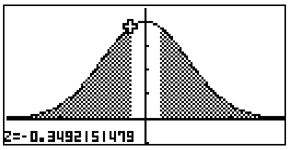
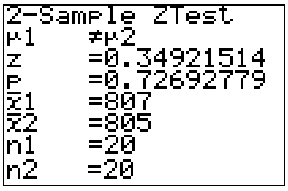
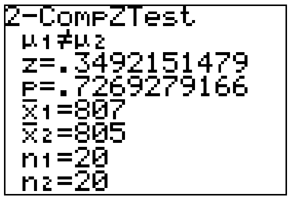
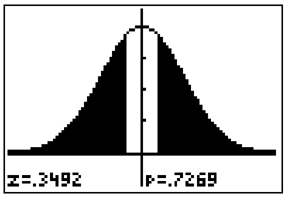
$\bar{x}_2 = 805$ g pour la deuxième machine

Pour cette comparaison, on élabore un test de comparaison des moyennes μ_1 et μ_2 des remplissages des deux doseuses.

Avec une calculatrice

<table>
<tr><th>sur CASIO 85</th><th>sur TI 83 PLUS .fr</th></tr>
<tr><td>touche MENU, option STAT
</td>
<td rowspan="4">touche stats, onglet TESTS

3:2-CompZTest...</td></tr>
<tr><td>
touche F3, pour l'onglet TEST</td></tr>
<tr><td>
Les variances du caractères sont connues :
Test Z : touche F1, pour l'onglet Z</td></tr>
<tr><td>
Comparaison de deux moyennes :
touche F2, pour l'onglet 2-S</td></tr>
</table>

sur CASIO 85		sur TI 83 PLUS *.fr*
	Ici, on complète la fenêtre en fonction des valeurs obtenues avec les échantillons prélevés. Dans le cas où on dispose des valeurs brutes des échantillons prélevés, saisies dans des listes : **Data : List** (Casio) ou **Inpt : Data** (T.I)	

DRAW sur CASIO	CALC sur CASIO	Calculs sur TI	Dessin sur TI
			

On est dans le cas où les variances des distributions des deux populations sont connues, la variable de décision est $Z = \dfrac{\overline{X_1} - \overline{X_2}}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$, soit ici $\dfrac{\overline{X_1} - \overline{X_2}}{\sqrt{\dfrac{20^2}{20} + \dfrac{16^2}{20}}}$ où $\overline{X_1}$ et $\overline{X_2}$ sont les variables aléatoires qui, aux échantillons de tailles 20, issus des productions des deux machines, associent la moyenne du caractère dans l'échantillon.

Sous l'hypothèse d'égalité de μ_1 et μ_2, la loi de probabilité de Z est la loi normale centrée réduite car le caractère est distribué selon une loi normale dans les deux populations.

Z est la valeur de Z observée sur l'échantillon : $Z_{obs} = \dfrac{807 - 805}{\sqrt{\dfrac{20^2}{20} + \dfrac{16^2}{20}}} \simeq 0{,}349\,215\,14$.

p est la probabilité que, sous l'hypothèse d'égalité de μ_1 et μ_2, la variable aléatoire $|Z|$ prenne une valeur supérieure à $|Z_{obs}|$.

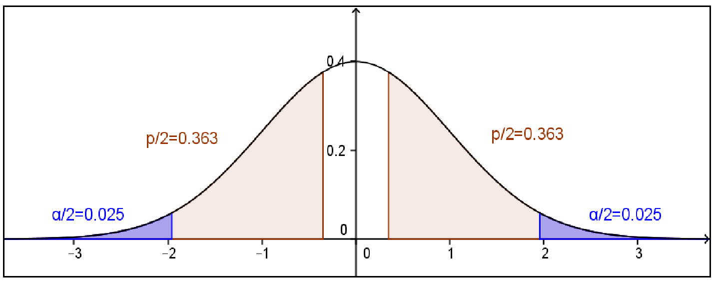
Ici, **p** $= P(|Z| > 0{,}349\,215\,14) \simeq 0{,}726\,927\,79$.

Exploitons les résultats donnés par la calculatrice pour rédiger une règle de décision en comparant α avec la *p-value* **p**.

- Si $\alpha >$ **p**, on rejette l'hypothèse d'égalité de μ_1 et μ_2
- Si $\alpha \leq$ **p**, on n'est pas en mesure de rejeter l'hypothèse d'égalité de μ_1 et μ_2

Revenons à notre exemple :

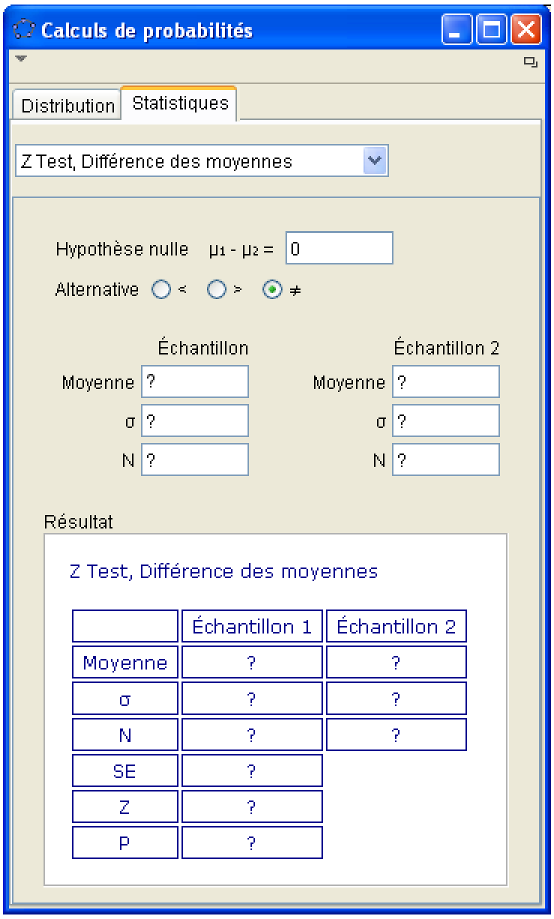
Si $\alpha = 0{,}05$: $\alpha < 0{,}726\,927\,79$, on rejette l'hypothèse d'égalité de μ_1 et μ_2.

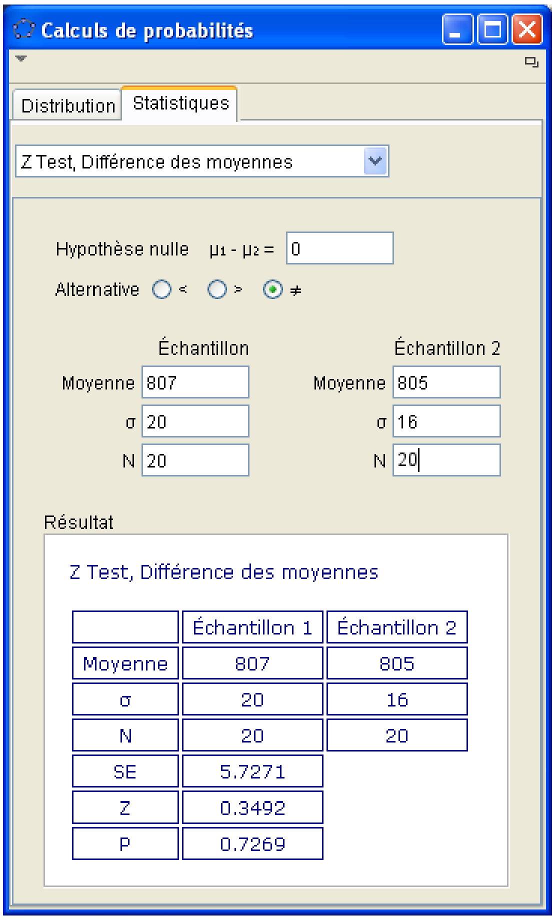


Avec Geogebra

Dans le cas où on connaît les variances du caractère étudié dans les deux populations et les moyennes du caractères dans les deux échantillons, on utilise l'onglet **Statistiques** proposé dans le module **Calculs de probabilités**.

On choisit la commande **ZTest, Différence des moyennes**.



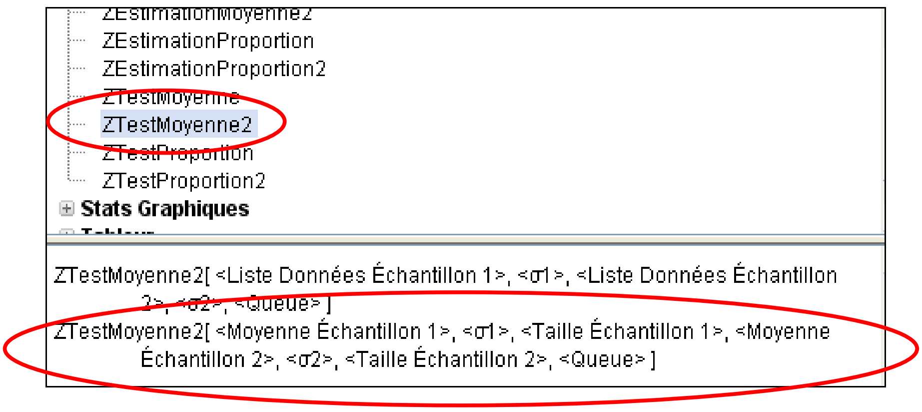


$$\mathbf{SE} = \sqrt{\frac{20^2}{20} + \frac{16^2}{20}} \approx 5{,}727\ 1.$$

Pour conclure, il reste à comparer α avec la p-value *p*.

Avec Geogebra (bis)

On peut aussi utiliser la commande **ZTestMoyenne2**. On affiche la liste des commandes de la rubrique **Statistiques**. Comme on ne connaît des échantillons que leurs moyennes (on ne dispose pas des données brutes), on utilise la deuxième option proposée.



La syntaxe est :

ZTestMoyenne2[Moyenne1,σ1,Taille1,Moyenne2,σ2,Taille2,Hypothèse]

Moyenne1, ***Écart-type1*** et ***Taille1*** désignent la moyenne du caractère dans l'échantillon 1, l'écart type du caractère dans la population 1 et la taille de l'échantillon 1.

Moyenne2, ***Écart-type2*** et ***Taille2***désignent la moyenne du caractère dans l'échantillon 2, l'écart type du caractère dans la population 2 et la taille de l'échantillon 2.

Hypothèse prend une des trois formes : <, > ou ≠ selon que l'hypothèse alternative est "$\mu_1 < \mu_2$", "$\mu_1 > \mu_2$" ou "$\mu_1 \neq \mu_2$".

La commande donne le résultat sous la forme d'une liste : $\{p\text{-}value, t_{obs}\}$ où *p-value* est l'une des trois probabilités $P(Z \leq Z_{obs})$, $P(Z \geq Z_{obs})$ ou $P(|Z| \geq |Z_{obs}|)$ selon que l'hypothèse alternative est "$\mu_1 < \mu_2$", "$\mu_1 > \mu_2$" ou "$\mu_1 \neq \mu_2$" où Z suit la loi normale centrée réduite.

Revenons à notre exemple :

Ici, la formule est **ZTestMoyenne2[807,20,20,805,16,20,"≠"]** à écrire dans la ligne de saisie. On obtient {0.727,0.349}.

Pour conclure, il reste à comparer α avec la p-value *p*.