TEST D'HYPOTHÈSE

I. Test d'hypothèse :

1°) Le principe :

On se place dans le cas où on ne connaît pas la valeur d'un paramètre sur une population (par exemple la probabilité qu'une punaise lancée retombe pointe en bas). On émet une hypothèse sur ce caractère (par exemple, on suppose que la probabilité que la punaise retombe pointe en bas est de 40 %). À partir de cette hypothèse, on donne un intervalle de fluctuation pour une taille n d'échantillon. On prélève un échantillon de taille n, on regarde alors si le paramètre étudié est dans l'intervalle de fluctuation pour rejeter ou non l'hypothèse.

2°) Vocabulaire :

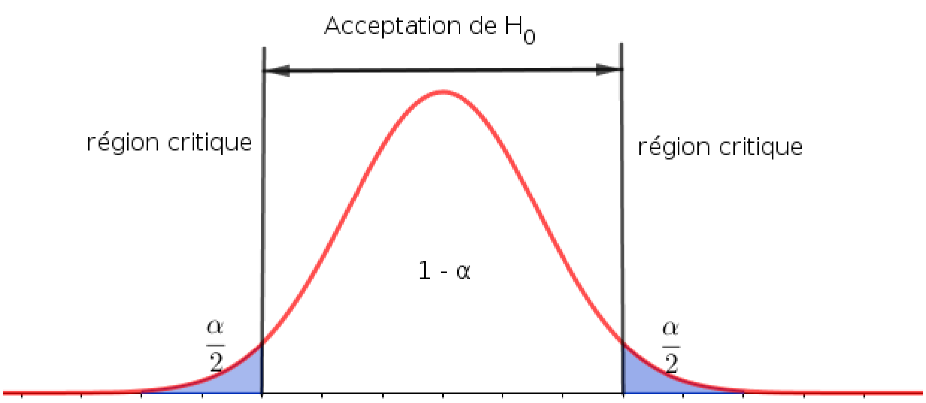
L'hypothèse formulée est appelée hypothèse nulle et est notée H_0. On formule également une autre hypothèse, appelée hypothèse alternative et noté H_1.

Exemple :

Reprenons le cas de la punaise lancée, on appelle p la probabilité que la punaise retombe pointe en bas. On émet l'hypothèse nulle $H_0 : p = 0{,}4$ et l'hypothèse alternative : $H_1 : p \neq 0{,}4$. Comme il s'agit d'une proportion, l'intervalle de fluctuation à 95 % pour un échantillon de taille 100 est donc

$$I = \left[0{,}4 - 1{,}96\sqrt{\frac{0{,}6\times 0{,}4}{100}} ; 0{,}4 + 1{,}96\sqrt{\frac{0{,}6\times 0{,}4}{100}}\right] \approx [0{,}3 ; 0{,}5].$$

Si on s'intéresse au changement de ce paramètre dans les deux directions ($p > 0{,}4$ et $p < 0{,}4$), on dit qu'on effectue un test bilatéral.

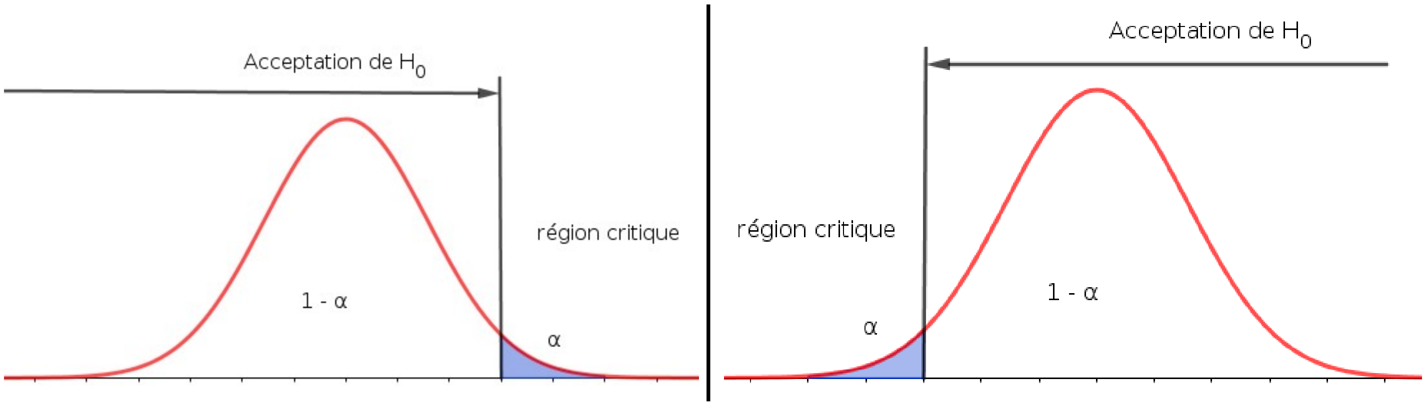


Exemple :

On effectue 100 lancers de la punaise et on calcule la fréquence f de la position « pointe vers le bas » dans l'échantillon ainsi obtenu.

Si $f \notin I$, on rejette l'hypothèse H_0. Par exemple, si dans l'échantillon on trouve $f = 0{,}55$, on considère que l'hypothèse H_0 ne pouvait donc pas être la bonne et on la rejette.

Si on s'intéresse au changement de ce paramètre dans une seule direction, on dit qu'on effectue un test unilatéral.



Exemple :

Ça peut être le cas par exemple si on s'intéresse à une proportion p de pièces défectueuses dans une production. Si $H_0 : p = 0,2$. Si on veut juste tester le fait que la proportion de pièces défectueuses est inférieure à 0,2, on effectuera un test unilatéral. Ici, pour un échantillon de taille 100, avec un seuil de 95 %, on regardera si la proportion f de pièces défectueuses est inférieure à $0,2+1,64\sqrt{\dfrac{0,2\times 0,8}{100}} \approx 0,28$.

Remarque :

1,64 est obtenu en faisant FracNormal(0.95,0,1) ou InvNormal(0.95,0,1) suivant les calculatrices.

II. Test de comparaison :

Comme ci-dessus, les tests de comparaison peuvent être unilatéraux ou bilatéraux.

1°) Comparaison de moyennes :

Dans cette partie, on veut tester s'il existe une différence significative entre deux échantillons, (par exemple si deux usines fabriquent une même pièce, s'il y a un écart entre les deux productions).

On prend donc deux échantillons :

Un échantillon A, de taille n_A, de moyenne m_A et d'écart type σ_A, extrait d'une population P de moyenne μ et d'écart type σ.

Un échantillon B, de taille n_B, de moyenne m_B et d'écart type σ_B, extrait d'une population P' de moyenne μ' et d'écart type σ'.

On ne connaît pas μ, μ', σ et σ', on veut comparer μ et μ'.

On note $\overline{X_A}$ (resp. $\overline{X_B}$) la variable aléatoire qui a tout échantillon de taille n_A (resp. n_B) associe sa moyenne. Elle suit approximativement la loi normale de moyenne m_A et d'écart type $\dfrac{\sigma}{\sqrt{n_A}}$, soit $\mathcal{N}(m_A ; \dfrac{\sigma^2}{n_A})$ (resp. $\mathcal{N}(m_B ; \dfrac{\sigma'^2}{n_B})$).

On suppose que les variables aléatoires $\overline{X_A}$ et $\overline{X_B}$ sont indépendantes. On note la variable aléatoire $D = \overline{X_A} - \overline{X_B}$. Cette dernière suit $\mathcal{N}(m_A - m_B; \dfrac{\sigma^2}{n_A} + \dfrac{\sigma'^2}{n_B})$.

Construction de test :

On pose :

Hypothèse nulle $H_0 : m_A = m_B$;

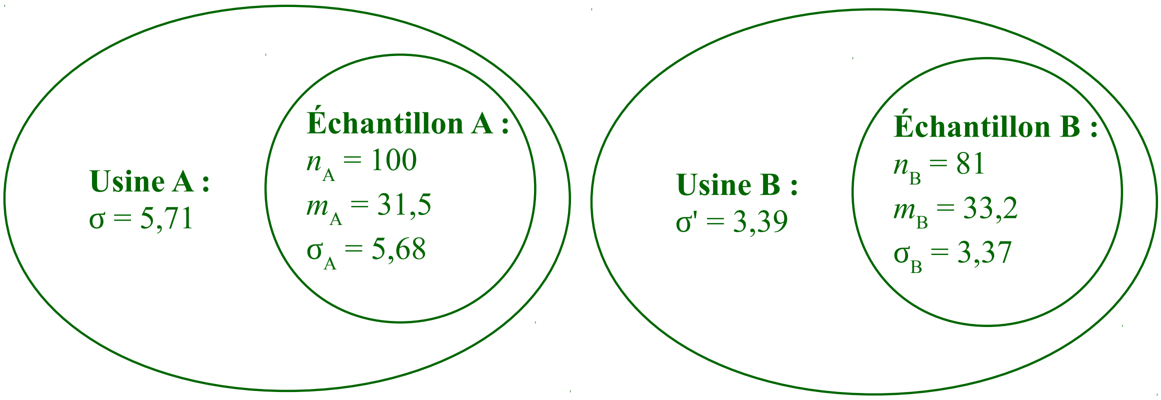
Hypothèse nulle $H_1 : m_A \neq m_B$.

Sous l'hypothèse H_0, on construit donc un intervalle de fluctuation de la variable D (donc centré en 0) au seuil $1 - \alpha$ et nous procédons comme plus haut.

2°) Exemple :

Exemple :

Les usines A et B fabriquent des tuyaux. On a les données suivantes :



Les écarts types σ et σ' sont des estimations ponctuelles à partir de σ_A et σ_B.

L'écart type de D est donc $\sigma_D = \sqrt{\frac{\sigma^2}{n_A}+\frac{\sigma'^2}{n_B}} \approx 0{,}68$.

Au taux de confiance de 95 %, sous l'hypothèse H_0, l'intervalle de fluctuation de D est donc :

$I = [0 - 1{,}96\times0{,}68 \; ; \; 0 + 1{,}96\times0{,}68] \approx [-1{,}33 \; ; \; 1{,}33]$.

Utilisation du test :

Ici, la différence d entre les deux moyennes des échantillons est $d = 33{,}2 - 31{,}5 = 1{,}7 \notin I$, on rejette donc l'hypothèse H_0 et on considère donc qu'il y a une différence entre les deux productions.

3°) Comparaison de proportions :

On s'y prend exactement de la même façon que pour la comparaison de moyennes, à la différence près que, si on nomme p et p' les proportions respectives des échantillons A et B, les variances sont respectivement $\frac{p(1-p)}{n_A}$ et $\frac{p'(1-p')}{n_B}$, et donc la variance de D est $\frac{p(1-p)}{n_A}+\frac{p'(1-p')}{n_B}$.

III. Risques de première et de deuxième espèce :

Définitions :

Les règles de décision que nous énonçons accepte un risque d'erreur (α) de rejeter à tort l'hypothèse H_0, c'est-à-dire de rejeter H_0 alors qu'elle est vraie. C'est ce qu'on appelle le **risque de première espèce**.

Ces règles de décision comporte également un deuxième risque, celui de ne pas rejeter H_0 alors que c'est H_1 qui est vraie. C'est le **risque de deuxième espèce** que l'on note habituellement β.

On peut résumer tout ceci dans un tableau :

	H_0 est vraie	**H_1 est vraie**
Ne pas rejeter H_0	Bonne décision Niveau de confiance $1 - \alpha$.	Manque de puissance Risque β de deuxième espèce
Rejeter H_0	Erreur : rejet à tort Risque α de première espèce	Puissance du test $1 - \beta$

Remarque :

La probabilité complémentaire du risque de deuxième espèce $(1- \beta)$ définit la puissance du test à l'égard de la valeur du paramètre dans l'hypothèse alternative H_1. Cette puissance représente la probabilité de rejeter l'hypothèse nulle H_0 lorsque l'hypothèse vraie est H_1. Plus β est petit, plus le test est puissant.