

TESTS D'HYPOTHESES

Les tests vont permettre de valider ou non une hypothèse considérée comme fondée. Il s'agit de donner dans ce chapitre, outre le principe de construction des tests, une idée de l'erreur commise, c'est à dire du risque pris.

1. Tests classiques

La mise en place d'un test classique revient à utiliser une démarche _____ .

1.1. Test bilatéral

Exemple :

Un ancien procédé de fabrication d'écrans plats produit des écrans dont la durée de vie suit une loi normale de moyenne 10000 h et d'écart-type 1200 h.

Un nouveau procédé est mis en place sur une autre chaîne. On admet que la durée de vie suit aussi une loi normale. Sur un échantillon de 100 écrans issus de cette nouvelle chaîne, la durée de vie moyenne est de 10300 h.

La question posée est : ce nouveau procédé est-il différent de l'ancien, en terme de durée de vie moyenne des écrans ?

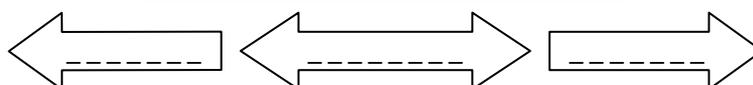
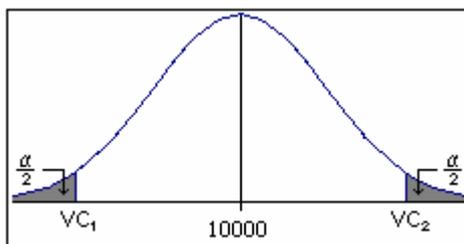
Mise en place du problème :

On précise la taille de l'échantillon sur lequel le test va être construit (ici 100) puis

- On pose tout d'abord une hypothèse H_0^1 , appelée hypothèse nulle, qui est que le nouveau procédé est identique à l'ancien donc _____.
- On pose ensuite une hypothèse H_1 , appelée hypothèse alternative, qui est que le nouveau procédé est différent de l'ancien donc _____.
- On choisit un seuil d'erreur du test (le risque) que l'on note α , ici $\alpha = 0,05$.

On suppose alors provisoirement que l'hypothèse nulle est vraie. Il s'agit alors de donner un domaine de validité [VC1 ; VC2] pour notre hypothèse ou plus exactement le complémentaire de ce domaine appelé *région critique*.

Sous _____



¹ En pratique l'hypothèse nulle est définie par une égalité contrairement à l'hypothèse alternative.

Comme H_0 est supposée vraie, la variable aléatoire \bar{X} égale à la durée de vie des écrans de l'échantillon suit la loi normale _____.

Pour notre seuil d'erreur le test signifie :

$$P(10000 - a \leq \bar{X} \leq 10000 + a) = 0,95.$$

On pose $T = \frac{\bar{X} - 10000}{120}$ qui suit alors la loi normale $\mathcal{N}(0 ; 1)$

$$P\left(-\frac{a}{120} \leq T \leq \frac{a}{120}\right) = 0,95 \text{ donc } a = 1,96 \times 120 = 235,2.$$

En conclusion, si pour un échantillon donné, la moyenne de l'échantillon m_e , est dans l'intervalle $[10000 - 235,2 ; 10000 + 235,2] = [9764,8 ; 10235,2]$ alors on accepte H_0 sinon on rejette H_0 , c'est à dire qu'on accepte H_1 .

Pour notre échantillon de tubes issus de la nouvelle chaîne $m_e = 10300$, on conclut donc qu'il y a _____.

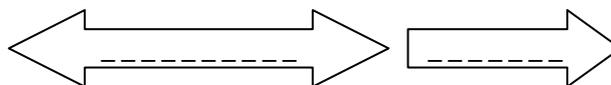
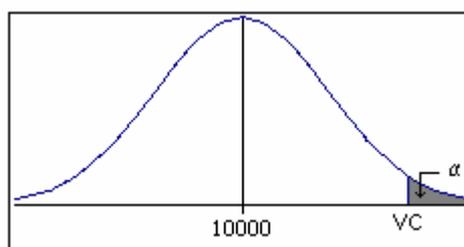
1.2. Test unilatéral

Souvent, on veut s'assurer qu'un procédé est meilleur qu'un autre (ou pire etc.). On utilise alors un test unilatéral.

En reprenant l'exemple ci-dessus, la question que l'on peut se poser est de savoir si le nouveau procédé est meilleur que l'ancien. Pour cela :

- H_0 : le nouveau procédé est identique à l'ancien donc _____.
- H_1 : le nouveau procédé est meilleur que l'ancien donc _____.
- On choisit $\alpha = 0,05$.

Sous _____



Comme H_0 est supposée vraie, la variable aléatoire \bar{X} égale à la durée de vie des écrans de l'échantillon suit la loi normale _____.

Pour notre seuil d'erreur le test signifie :

$$P(\bar{X} \leq 10000 + a) = 0,95.$$

On pose $T = \frac{\bar{X} - 10000}{120}$ qui suit alors la loi normale $\mathcal{N}(0 ; 1)$

$$P\left(T \leq \frac{a}{120}\right) = 0,95 \text{ donc } a = 1,65 \times 120 = 198.$$

En conclusion, si pour un échantillon donné, la moyenne de l'échantillon m_e , est inférieure à $10000 + 198 = 10198$ alors on accepte H_0 sinon on rejette H_0 , c'est à dire qu'on accepte H_1 .

Pour notre échantillon d'écrans issus de la nouvelle chaîne $m_e = 10300$, on conclut donc _____.

2. Erreurs de 1^{ère} et de 2^{ème} espèce

Lors de la prise de décision, on prend le risque de commettre une erreur que l'on choisit au départ, c'est α le seuil d'erreur du test. Ce risque est celui de rejeter à tort H_0 qui est vraie.

Mais il y a un autre type d'erreur que l'on peut commettre, c'est d'accepter H_0 à tort.

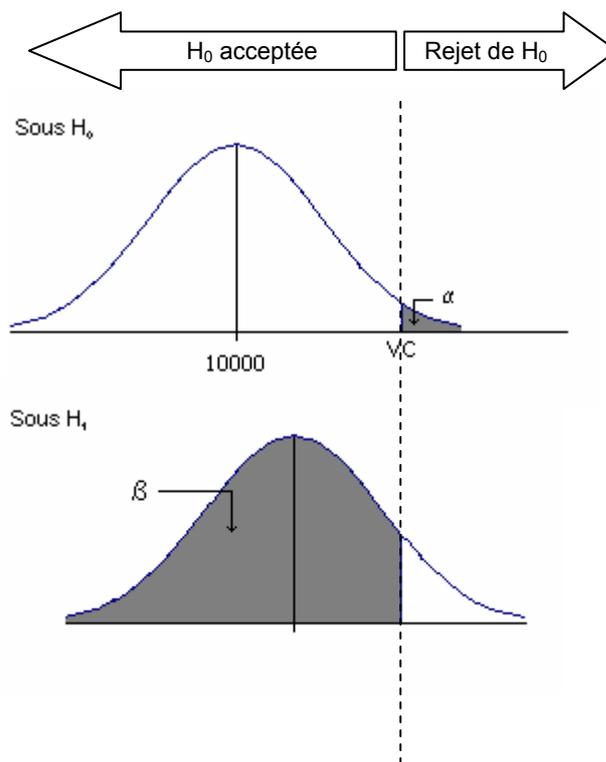
En fait rejeter H_0 alors que H_0 est vraie c'est commettre une erreur de _____ espèce, de probabilité _____ et accepter à tort H_0 (c'est à dire si H_1 est vraie) c'est commettre une erreur de _____ espèce, de probabilité _____.

Résumons :

	Décision	
	H_0 acceptée	H_0 rejetée
H_0 est vraie	Décision correcte Probabilité $1 - \alpha$	Erreur 1 ^{ère} espèce Probabilité α
H_1 est vraie	Erreur de 2 ^{ème} espèce Probabilité β	Décision correcte Probabilité $1 - \beta$

Remarque : lors d'un test, on pose l'erreur de 1^{ère} espèce α mais on ne peut connaître l'erreur de 2^{ème} espèce β que si l'hypothèse alternative H_1 est définie par une égalité.

Graphiquement, il est possible de représenter "l'état du monde" sous H_0 et "l'état du monde" sous H_1 et de voir apparaître les deux types d'erreurs :



Le problème est que pour, à la fois, minimiser α et β la seule solution est d'augmenter la taille de l'échantillon, c'est à dire n . Sinon si α diminue β augmente et vice versa.

3. Probabilité critique

Il peut sembler arbitraire de fixer, à priori, le seuil de risque (erreur de 1^{ère} espèce α) du test. En effet, il est préférable, dans bien des cas², de se forger sa propre opinion sur le rejet hypothétique de l'hypothèse H_0 . C'est le rôle de la probabilité critique.

Pour calculer la probabilité critique, on utilise l'hypothèse nulle H_0 et on regarde dans quelle mesure elle est compatible avec les données de notre échantillon.

Reprenons notre exemple : $n = 100$ et $m_e = 10300$

- H_0 : le nouveau procédé est identique à l'ancien donc _____.
- H_1 : le nouveau procédé est meilleur que l'ancien donc _____.

² On utilise un seuil fixe lorsque α peut être déterminé rationnellement ou s'il y a beaucoup d'échantillons à tester. Dans un contrôle de qualité on prélève régulièrement des échantillons afin de s'assurer que la production est toujours conforme. On applique alors le test au seuil de risque α qui ne demande pas de calcul supplémentaire.

Si H_0 est vraie, quelle est la probabilité que $\bar{X} \geq 10300$?

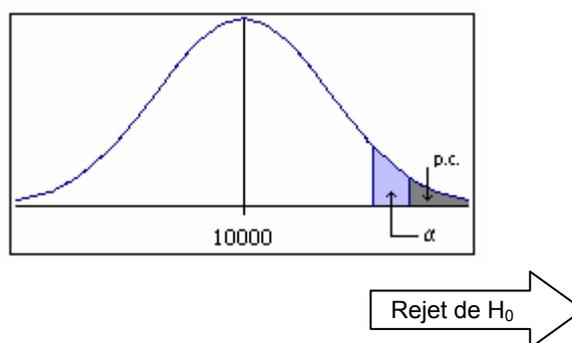
\bar{X} suit la loi _____ donc :

$$P(\bar{X} \geq 10300) = P(T \geq 2,5) = 1 - P(T < 2,5) = 1 - 0,9938 = 0,0062.$$

Cela signifie que sous l'hypothèse H_0 , il n'y a que 0,62% de chance d'observer une moyenne aussi élevée. Ce 0,62% est appelé probabilité critique (unilatérale) pour H_0 .

La probabilité critique résume donc le degré de concordance entre les données de l'échantillon et l'hypothèse H_0 .

Pour conclure, lorsque l'on connaît la probabilité critique associée à une hypothèse H_0 , il est facile de conclure lors d'un test classique. Si la probabilité critique est inférieure au seuil de risque α du test alors l'hypothèse H_0 est à rejeter.



Remarque : on peut calculer de même une probabilité critique bilatérale pour H_0 .

Les tests d'hypothèses classiques peuvent être utilisés pour tester une hypothèse portant sur les proportions (tester l'hypothèse nulle "le taux de défection est de 10% contre l'hypothèse alternative "il est de 15%") mais si l'on veut tester des hypothèses concernant les proportions d'une population multinomiale il faut utiliser un autre type de test, le test du χ^2 .

4. Test du χ^2

On utilise le test du χ^2 (prononcer khi-2) dans deux cas de figure :

- Pour tester si une distribution suit une distribution théorique connue.
- Pour tester l'indépendance de deux variables.

Le test du χ^2 est basé sur la comparaison entre les effectifs observés et les effectifs théoriques, obtenus en supposant l'hypothèse que la distribution est connue. En d'autres termes de combien l'effectif observé s'écarte-t-il de l'effectif théorique ?

Exemple :

Pour étudier l'influence, éventuelle, de la couleur de l'emballage sur un nouveau produit cosmétique, une entreprise a procédé à une étude sur un échantillon de 200 ménages. On a remis à chacun d'eux quatre produits identiques dans des boîtes de couleurs différentes en leur laissant supposer que les produits étaient différents. Après utilisation des produits, on a demandé aux différents ménages leur préférence :

Couleur	Rouge	Blanc	Bleu	Vert
Effectif	51	74	30	45

Le problème est de savoir si la couleur de l'emballage a de l'importance.

Nous allons utiliser le test du χ^2 pour tester l'hypothèse, H_0 , selon laquelle _____

_____.

Si on suppose H_0 vraie, l'effectif de chaque couleur doit être identique (loi uniforme discrète) et égal à _____.

Couleur	Rouge	Blanc	Bleu	Vert
Effectif observé O_i				
Effectif théorique T_i				
$O_i - T_i$	1	24	-20	-5
$(O_i - T_i)^2$	1	576	400	25
$(O_i - T_i)^2 / T_i$	0,02	11,52	8	0,5

Le test du χ^2 consiste à mesurer les écarts³ entre les effectifs observés et les effectifs théoriques et à les comparer aux effectifs théoriques. On note : $\chi^2_{\text{cal}} = \sum_i \frac{(O_i - T_i)^2}{T_i}$.

Ici $\chi^2_{\text{cal}} =$ _____.

Une valeur élevée du χ^2_{cal} traduit un écart important par rapport à la distribution théorique, donc une crédibilité moindre pour H_0 . Les valeurs limites de ces écarts sont données dans la table de la loi du χ^2 , qui nous permet de conclure (pour peu que l'on ait choisi un risque α) :

- Si $\chi^2_{\text{cal}} > \chi^2_{\alpha}$ alors on _____ avec une probabilité _____ de se tromper.
- Si $\chi^2_{\text{cal}} < \chi^2_{\alpha}$ alors on ne _____ donc on accepte _____.

Remarque : On connaît le risque de 1^{ère} espèce α , mais pas celui de 2^{ème} espèce β .

³ Pour pouvoir mesurer les écarts, il faut les élever au carré.

Pour pouvoir lire sur la table du χ^2 la valeur de χ_α^2 , il nous faut déterminer le degré de liberté (d.d.l.) c'est à dire le nombre minimal d'informations nécessaires pour conclure.

Ici, sur les 4 modalités (appelées aussi classes) prises par la couleur, la connaissance de l'effectif de 3 classes permet de déterminer le quatrième sachant que l'effectif total est 200. Donc le d.d.l vaut $v = 4 - 1 = 3$.

Par lecture sur la table du χ^2 on a :

$\chi_{\text{cal}}^2 = 20,04 > \chi_{0,001}^2 = 16,27$ donc on _____ l'hypothèse selon laquelle la couleur n'influe pas sur le choix du produit avec moins de 0,1 % de risque d'erreur.

Remarques :

1. La règle pour calculer le nombre de degrés de liberté est la suivante :

$$v = \text{nombre de classes} - 1 - \text{nombre de paramètres estimés sur l'échantillon.}$$

2. En particulier, si on veut tester l'hypothèse selon laquelle une distribution suit une distribution normale $\mathcal{N}(m ; \sigma)$ où m et σ sont calculés sur l'échantillon le nombre de d.d.l. vaut $v = \text{nombre de classes} - 3$.

3. Ce test nécessite qu'aucun effectif théorique soit inférieur à 5. Dans le cas contraire, on procède à des regroupements de classes.

Le test du χ^2 peut aussi être utilisé pour tester l'indépendance de deux variables.

Exemple :

On demande à 200 personnes (100 du type H et 100 du type F) dans quelle tranche se trouve leur salaire.

	5000 – 10000	10000 – 15000	15000 – 20000	20000 – 25000
Type_H	20	30	40	10
Type_F	25	40	30	5

Le salaire est-il indépendant du type des personnes ?

Là aussi, nous allons utiliser le test du χ^2 pour conclure. La démarche est la même que pour l'exemple précédent, avec cette fois-ci un tableau à double entrée.

On suppose que l'hypothèse H_0 , comme quoi le salaire est indépendant du type des personnes, est vraie.

	5000 – 10000	10000 – 15000	15000 – 20000	20000 – 25000	$n_{i\bullet}$
Type_H	20	30	40	10	100
Type_F	25	40	30	5	
$n_{\bullet j}$	45	70	70	15	200

Remarque : Soit le tableau de distribution des variables X et Y

$x_i \setminus y_j$	y_1	y_2	...	y_q	$n_{i\bullet}$
x_1	n_{11}	n_{12}		n_{1q}	$n_{1\bullet}$
x_2	n_{21}			n_{2q}	$n_{2\bullet}$
...					...
x_p	n_{p1}			n_{pq}	$n_{p\bullet}$
$n_{\bullet j}$	$n_{\bullet 1}$	$n_{\bullet 2}$...	$n_{\bullet q}$	n

Les variables X et Y sont indépendantes si et seulement si $n \times n_{ij} = n_{i\bullet} \times n_{\bullet j}$ pour tous i et j.

On forme alors le tableau des effectifs théoriques :

	5000 – 10000	10000 – 15000	15000 – 20000	20000 – 25000
Type_H				
Type_F				

On calcule $(O_i - T_i)^2 / T_i$:

0,28	0,71	0,71	0,83
0,28	0,71	0,71	0,83

Puis on fait la somme : $\chi^2_{\text{cal}} = \text{-----}$.

Reste à trouver le nombre de degrés de liberté. Ici, comme on connaît les effectifs totaux par ligne, il suffit de connaître (nombre de colonnes – 1) valeurs par ligne pour les connaître toutes. Et comme on connaît les effectifs totaux par colonne, il suffit de connaître (nombre de lignes – 1) valeurs par colonne.

	5000 – 10000	10000 – 15000	15000 – 20000	20000 – 25000	$n_{i\bullet}$
Type_H	20	30	40	?	100
Type_F	?	?	?	?	100
$n_{\bullet j}$	45	70	70	15	200

Donc le d.d.l. vaut (nombre de lignes – 1) \times (nombre de colonnes – 1).

Ici $v = \text{-----}$. Par lecture sur la table $\chi^2_{0,05} = \text{-----}$.

Comme $\chi^2_{\text{cal}} = \text{-----} < \chi^2_{0,05} = \text{-----}$ on ne rejette pas l'hypothèse ----- .

On accepte donc l'hypothèse selon laquelle Il n'y a pas de différence de salaire entre les deux types, avec un risque de 2^{ème} espèce inconnu.