Tests d'hypothèse

I Introduction à la notion de tests

I.1 Principe d'un test d'hypothèse

Un test d'hypothèses est un procédé qui permet de trancher entre deux hypothèses au vu des résultats d'un échantillon, en quantifiant le risque associé à la décision prise.

- On étudie une population dont les éléments possèdent un caractère (mesurable ou qualitatif) et dont la valeur du paramètre relative au caractère étudié est inconnue,
- une hypothèse est formulée sur la valeur du paramètre,
- on veut porter un jugement sur la base des résultats d'un échantillon prélevé de cette population.

La variable aléatoire d'échantillonnage servant d'estimateur au paramètre de la population ne prendra pas une valeur rigoureusement égale à la valeur théorique proposée dans l'hypothèse. Cette variable aléatoire comporte des fluctuations d'échantillonnage qui sont régies par des distributions connues.
La construction d'un test d'hypothèse consiste à déterminer entre quelles valeurs peut varier la variable aléatoire, en supposant l'hypothèse vraie, sur la seule considération du hasard de l'échantillonnage.

I.2 Définition des concepts utiles à l'élaboration des tests d'hypothèse

Hypothèse statistique.
Une *hypothèse statistique* est une affirmation concernant les caractéristiques (valeurs des paramètres, forme de la distribution des observations) d'une population.
Test d'hypothèse.
Un *test d'hypothèse* est une démarche qui a pour but de fournir une règle de décision permettant, sur la base de résultats d'échantillon, de faire un choix entre deux hypothèses statistiques.
Hypothèse nulle (H_0) **et hypothèse alternative** (H_1).
L'hypothèse selon laquelle on fixe a priori un paramètre de la population à une valeur particulière s'appelle l'**hypothèse nulle** et est notée H_0. N'importe quelle autre hypothèse qui diffère de l'hypothèse H_0 s'appelle l'**hypothèse alternative** et est notée H_1.
C'est l'hypothèse nulle qui est soumise au test et toute la démarche du test s'effectue en considérant cette hypothèse comme vraie. Cette hypothèse de base est celle dont on considère que le rejet à tort aura les conséquences les plus fâcheuses.

Exemple : un constructeur informatique annonce que ses portables ont une autonomie de 10 heures. Un échantillon de 100 ordinateurs donne une moyenne de 9h30, avec un écart-type de 45 minutes. Peut-on croire le fabricant ?

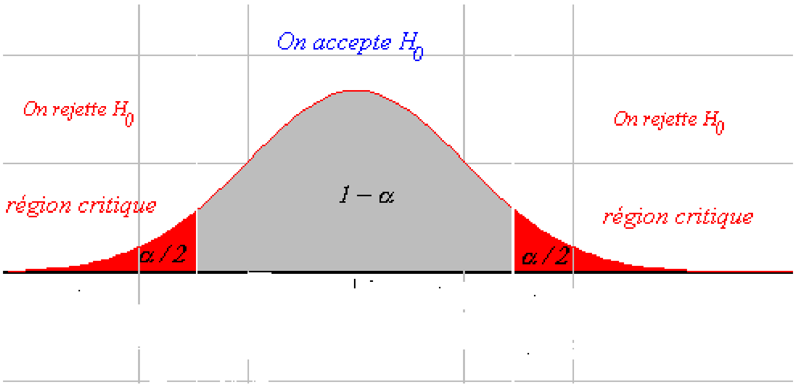
- ▷ on peut tester $H_0 : \mu = 10$ contre $H_1 : \mu \neq 10$: test bilatéral
- ▷ on peut aussi tester $H_0 : \mu = 10$ contre $H_1 : \mu < 10$: test unilatéral.

Seuil de signification du test

Le risque d'erreur, consenti à l'avance et que nous notons α, de rejeter à tort l'hypothèse nulle H_0 alors qu'elle est vraie, s'appelle le *seuil de signification* du test et s'énonce en probabilité ainsi,

$$\alpha = P(\text{rejeter } H_0 \mid H_0 \text{ vraie}).$$

À ce seuil de signification, on fait correspondre sur la distribution d'échantillonnage de la statistique une zone de rejet de l'hypothèse nulle, appelée également **région critique**. L'aire de cette région correspond à la probabilité α. Si par exemple on choisit $\alpha = 0,05$, cela signifie que l'on admet d'avance que la variable d'échantillonnage peut prendre, dans 5% des cas, une valeur se situant dans la zone de rejet de H_0, bien que H_0 soit vraie et ceci uniquement d'après le hasard de l'échantillonnage. Sur la distribution d'échantillonnage correspondra aussi une région complémentaire, dite **zone d'acceptation** de H_0 (ou région de non-rejet) de probabilité $1 - \alpha$.



Remarques :

(i) Les seuils de signification les plus utilisés sont $\alpha = 0,05$ et $\alpha = 0,01$, dépendant des conséquences de rejeter à tort l'hypothèse H_0.

(ii) La statistique qui convient pour le test est donc une variable aléatoire X dont la valeur observée sera utilisée pour décider du «rejet» ou du «non-rejet» de H_0. La loi de probabilité de cette variable aléatoire X sera déterminée en supposant que l'hypothèse H_0 est vraie.

Définition 1

- On appelle risque de première espèce, noté α, la probabilité de rejeter à tort H_0 (faux négatif).
- On appelle risque de deuxième espèce, et on note β la probabilité d'accepter à tort H_0 (faux positif).
- On appelle puissance du test le nombre $1 - \beta$. La puissance du test 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.

	H_0 vraie	H_1 vraie
H_0 décidée	$1 - \alpha$ vrai acceptation	β fausse acceptation
H_1 décidée	α faux rejet	$1 - \beta$ vrai rejet

Ces deux risques sont antagonistes, plus α est petit, plus β est grand (et vice-versa) : le fait d'imposer un α faible conduit à une règle de décision plus stricte, qui aboutit à rejeter H_0 que dans des cas rarissimes, et donc à la conserver parfois à tort (hypothèse conservatrice).

I.3 Démarche à suivre pour la mise en place d'un test

Pour chaque test d'hypothèses, on peut procéder en quatre étapes :

(i) Formuler l'hypothèse nulle H_0, et l'hypothèse alternative H_1 qui détermine la forme de la région critique.

(ii) Choisir la variable de test Z et déterminer sa loi de probabilité.

(iii) Déterminer les valeurs critiques de la variable de test, qui délimitent les zones d'acceptation et de rejet.

(iv) Calculer la valeur prise dans l'échantillon, énoncer la règle de décision et conclure le test.

II Test de conformité d'une moyenne à une norme

II.1 Démarche

Nous voulons déterminer si l'échantillon de taille n dont nous disposons appartient à une population de moyenne μ_0 au seuil de signification α.

1ère étape : formulation des deux hypothèses.
L'échantillon provient d'une population de moyenne μ. Nous voulons savoir si $\mu = \mu_0$. On va donc tester l'hypothèse H_0 contre l'hypothèse H_1 : $\begin{cases} H_0 & : \quad \mu = \mu_0 \\ H_1 & : \quad \mu \neq \mu_0. \end{cases}$.

2ème étape : choix de la variable de test.

- On détermine la variable aléatoire qui convient pour ce test. Ici, l'estimateur usuel de la moyenne μ, c'est-à-dire $\overline{X}$, semble tout indiqué.
- On détermine la loi de probabilité de $\overline{X}$ en se plaçant sous l'hypothèse H_0. Deux cas peuvent se produire.

(i) **Échantillon de grande taille** et de variance σ^2 connue.

$\overline{X}$ suit approximativement la loi normale de moyenne μ_0 (puisqu'on se place sous H_0) et d'écart-type $\dfrac{\sigma}{\sqrt{n}}$:

$$\overline{X} \hookrightarrow$$

On pose $Z = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$. Alors $Z \hookrightarrow \mathcal{N}(0, 1)$.

(ii) **Échantillon de taille quelconque** et de variance σ^2 inconnue dans le cas où $X \hookrightarrow \mathcal{N}(\mu\,,\,\sigma)$.

Dans ce cas la variable aléatoire discriminante du test sera

$$Z = \frac{\overline{X} - \mu_0}{S/\sqrt{n}} \qquad \text{où} \quad S^2 =$$

Ici $Z \hookrightarrow T_{n-1}$ (loi de Student à $n-1$ degrés de liberté).

3ème étape : détermination des valeurs critiques de Z délimitant les zones d'acceptation et de rejet.
On cherche z_α tel que $P(-z_\alpha \leqslant Z \leqslant z_\alpha) = 1-\alpha$ c'est-à-dire $P(Z \leqslant z_\alpha) =$

d'où $P\left(\mu_0 - z_\alpha \dfrac{S}{\sqrt{n}} \leqslant \overline{X} \leqslant \mu_0 + z_\alpha \dfrac{S}{\sqrt{n}}\right) =$
Ce qui donne le domaine d'acceptation au seuil de signification de $\alpha\%$:

$$D_0 = \left[\mu_0 - z_\alpha \frac{\widehat{\sigma}}{\sqrt{n}} \ , \ \mu_0 + z_\alpha \frac{\widehat{\sigma}}{\sqrt{n}}\right]$$

4ème étape : calcul de la valeur de $\overline{X}$ prise dans l'échantillon et conclusion du test.
On calcule la valeur $\overline{x_e}$ prise par $\overline{X}$ dans l'échantillon.

- Si la valeur $\overline{x_e}$ se trouve dans la zone de rejet, on dira que l'écart observé est statistiquement significatif au seuil α. Cet écart est anormalement élevé et ne permet pas d'accepter H_0.
On rejette H_0 et on décide que H_1 est vraie avec un risque de première espèce α.

- Si la valeur $\overline{x_e}$ se trouve dans la zone d'acceptation, on dira que l'écart observé n'est pas significatif au seuil α. Cet écart est imputable aux fluctuations d'échantillonnage.
On ne peut pas rejeter H_0 avec un risque de deuxième espèce β.

II.2 Risque de deuxième espèce

La règle de décision que nous avons déterminée acceptait un risque α qui était le risque de rejeter à tort l'hypothèse nulle H_0, c'est-à-dire le risque de rejeter l'hypothèse H_0, alors que H_0 est vraie. La règle de décision du test comporte également un deuxième risque, à savoir de celui de ne pas rejeter l'hypothèse nulle H_0 alors que c'est l'hypothèse H_1 qui est vraie. C'est le **risque de deuxième espèce**.

Le risque de première espèce α est choisi a priori. Toutefois le risque de deuxième espèce β dépend de l'hypothèse alternative H_1 et on ne peut le calculer que si on spécifie des valeurs particulières du paramètre dans l'hypothèse H_1 que l'on suppose vraie.

(i) Pour un même risque α et une même taille d'échantillon, si l'écart entre la valeur du paramètre posée en H_0 et celle supposée dans l'hypothèse vraie H_1 augmente, le risque β diminue.

(ii) Une réduction du risque de première espèce (de $\alpha = 0,05$ à $\alpha = 0,01$ par exemple) élargit la zone d'acceptation de H_0. Toutefois, le test est accompagné d'une augmentation du risque de deuxième espèce β. On ne peut donc diminuer l'un des risques qu'en consentant à augmenter l'autre.

(iii) Pour une valeur donnée de α, l'augmentation de la taille d'échantillon peut réduire la zone d'acceptation de H_0, conduisant à une diminution du risque β. Le test est alors plus puissant.

III Test de conformité d'une proportion à une norme

On se propose de tester si la proportion p d'éléments dans la population présentant un certain caractère qualitatif peut être ou non considérée comme égale à une valeur hypothétique p_0.
On dispose de la proportion $\widehat{p} = f_e$ d'éléments possédant ce caractère dans un échantillon de taille n.

1ère étape : formulation des deux hypothèses.
L'échantillon dont nous disposons provient d'une population dont la proportion d'éléments présentant le caractère qualitatif est p. Nous voulons savoir si $p = p_0$.

On va donc tester l'hypothèse H_0 contre l'hypothèse H_1 : $\begin{cases} H_0 & : \quad p = p_0 \\ H_1 & : \quad p \neq p_0. \end{cases}$.

2ème étape : choix de la variable de test.

- On détermine la variable aléatoire qui convient pour ce test. Ici, l'estimateur usuel de la proportion p, c'est-à-dire F, semble tout indiqué.

- On détermine la loi de probabilité de F en se plaçant sous l'hypothèse H_0.

On suppose que l'on dispose d'un grand échantillon et que p n'est pas trop petit (de manière à avoir $n > 30$, $np \geqslant 15$ et $n(1-p) > 5$). F suit approximativement la loi normale de moyenne p_0 (puisqu'on se place sous H_0) et d'écart-type $\sqrt{\dfrac{p_0(1-p_0)}{n}}$:

$$F \hookrightarrow$$

On pose $Z = \dfrac{F - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$. Alors $Z \hookrightarrow \mathcal{N}(0,1)$.

3ème étape : détermination des valeurs critiques de F délimitant les zones d'acceptation et de rejet.
On cherche z_α tel que $P(-z_\alpha \leqslant Z \leqslant z_\alpha) = 1 - \alpha$ c'est-à-dire $P(Z \leqslant z_\alpha) =$

Ce qui donne le domaine d'acceptation

$$D_0 = \left[p_0 - z_\alpha \sqrt{\frac{p_0(1-p_0)}{n}} \, , \, p_0 + z_\alpha \sqrt{\frac{p_0(1-p_0)}{n}} \right]$$

4ème étape : calcul de la valeur de F prise dans l'échantillon et conclusion du test.
On calcule la valeur $\widehat{p} = f_e$ prise par F dans l'échantillon.

- Si la valeur $\widehat{p}$ se trouve dans la zone de rejet, on dira que l'écart observé est statistiquement significatif au seuil α. On rejette H_0 avec un risque α de se tromper.

- Si la valeur $\widehat{p}$ se trouve dans la zone d'acceptation, on dira que l'écart observé n'est pas significatif au seuil α. Cet écart est imputable aux fluctuations d'échantillonnage.
  On ne peut pas rejeter H_0.

IV Test de conformité d'une variance à une norme

Soit X une variable aléatoire qui suit une loi normale d'espérance μ et de variance σ^2 inconnues. On dispose d'un échantillon de n réalisations indépendantes de la variable aléatoire X.

1ère étape : formulation des deux hypothèses.

On va tester l'hypothèse H_0 contre l'hypothèse H_1 : $\begin{cases} H_0 & : \quad \sigma^2 = \sigma_0^2 \\ H_1 & : \quad \sigma^2 > \sigma_0^2. \end{cases}$.

2ème étape : choix de la variable de test.

L'estimateur corrigé de la variance est $S^2 = \dfrac{1}{n-1}\sum_{k=1}^{n}(X_k - \overline{X})^2$. On a vu que la variable de test est

$$Z = \frac{(n-1)S^2}{\sigma_0^2} = \sum_{k=1}^{n}\left(\frac{X_k - \overline{X}}{\sigma_0}\right)^2 \hookrightarrow$$

3ème étape : détermination de la valeur critique de S^2 délimitant les zones d'acceptation et de rejet.

On cherche z_α tel que $P(Z \leqslant z_\alpha) = 1 - \alpha$ ce qui donne $P(S^2 \leqslant \dfrac{z_\alpha \sigma_0^2}{n-1}) = 1 - \alpha$,

d'où le domaine d'acceptation

$$D_0 = \left[0\ ,\ \frac{z_\alpha\ \sigma_0^2}{n-1}\right]$$

4ème étape : calcul de la valeur de S^2 prise dans l'échantillon et conclusion du test.
On calcule la valeur $s^2 = \widehat{\sigma}^2$ prise par S^2 dans l'échantillon.

- Si la valeur s^2 se trouve dans la zone de rejet, alors le test est significatif et on décide de rejeter l'hypothèse nulle H_0 avec un risque de première espèce α.

- Si la valeur s^2 se trouve dans la zone d'acceptation, le test n'est pas significatif et on conserve l'hypothèse nulle H_0 avec un risque de deuxième espèce β.

Exemple :
soit un échantillon d'une distribution normale, composé de 4 observations

$$2, 5, 10, 11.$$

Tester l'hypothèse nulle $H_0 : \sigma^2 = 10$ contre $H_1 : \sigma^2 > 10$. Ici $n = 4$ et $\sigma_0^2 = 10$. On obtient les estimations $\overline{x_e} = 7$ et $s^2 = 18$. Soit la variable de test $Z = \dfrac{3S^2}{\sigma_0^2} \hookrightarrow \chi_3^2$. On cherche z_α tel que $P(Z \leqslant z_\alpha) = 0,95$, ce qui donne $z_\alpha =$. Ainsi $P(\dfrac{3S^2}{\sigma_0^2} \leqslant z_\alpha) = 0,95$, ce qui donne

$$D_0 =$$

V Test du Khi-deux :

ajustement d'une distribution empirique à une loi théorique

V.1 Introduction

Lorsque nous avons accumulé suffisamment de données sur une variable statistique, on peut examiner si la distribution des observations semble s'apparenter à une distribution théorique connue (comme une loi binomiale, de Poisson, normale...). Un outil statistique qui permet de vérifier la concordance entre une distribution expérimentale et une distribution théorique est le **test du khi-deux**.
On cherche donc à déterminer si un modèle théorique est susceptible de représenter le comportement probabiliste de la variable observée, comportement fondé sur les fréquences des résultats obtenus sur l'échantillon.

Répartitions expérimentales

On répartit les observations suivant k classes (si le caractère est continu) ou k valeurs (si le caractère est discret). On dispose alors des effectifs des k classes : n_1, $n_2, \ldots, n_k$. On a la relation :

$$\sum_{i=1}^{k} n_i = n$$

où n est le nombre total d'observations effectuées. En pratique, on se placera dans le cas où $n \geqslant 50$.

Répartitions théoriques

En admettant comme plausible une distribution théorique particulière, on peut construire une répartition idéale des observations de l'échantillon de taille n en ayant recours aux probabilités tabulées (ou calculées) du modèle théorique : p_1, $p_2, \ldots, p_k$.

On obtient alors les effectifs théoriques c_i en écrivant

$$\forall\, i \in [\![1\,, k]\!],$$

Chaque effectif théorique c_i est supérieur ou égal à 5. Si cette condition n'est pas satisfaite, il y a lieu de regrouper deux ou plusieurs classes adjacentes. Il arrive fréquemment que ce regroupement s'effectue sur les classes aux extrémités de la distribution. k représente donc le nombre de classes après regroupement. On dispose automatiquement de la relation $\displaystyle\sum_{i=1}^{k} c_i = \ldots$

Définition de l'écart entre les deux distributions

La règle de décision se fonde sur la distance entre la distribution des fréquences observées dans l'échantillon, et celle qu'elle devrait être si elle était identique à celle de la loi théorique. Pour évaluer l'écart entre les effectifs observés n_i et les effectifs théoriques c_i, on utilise la somme des écarts normalisés entre les deux distributions, à savoir

$$\chi^2_{obs} = \frac{(n_1 - c_1)^2}{c_1} + \frac{(n_2 - c_2)^2}{c_2} + \cdots + \frac{(n_k - c_k)^2}{c_k} = \sum_{i=1}^{k}$$

Plus le nombre χ^2_{obs} ainsi calculé est grand, plus la distribution étudiée diffère de la distribution théorique.

Quelques considérations théoriques à propos de cet écart
Le nombre d'observations n_i parmi l'échantillon de taille n susceptible d'appartenir à la classe n°i est la réalisation d'une variable aléatoire N_i qui suit la loi binomiale $\mathscr{B}(n\,,\,p_i)$. Comme n est suffisamment grand ($n \geqslant 50$) et que $n\,p_i$ n'est pas trop petit (on a effectué des regroupements de classes pour qu'il en soit ainsi), on peut approcher la loi binomiale par une loi normale, c'est-à-dire

Pour simplifier, on approxime $n\,p_i(1-p_i)$ par $n\,p_i$. Donc $\dfrac{N_i - n\,p_i}{\sqrt{n\,p_i}}$ suit approximativement la loi normale centrée réduite $\mathcal{N}(0,1)$. Lorsqu'on élève au carré toutes ces quantités et qu'on en fait la somme, on obtient une somme de k var «presque indépendantes» suivant toutes la loi normale centrée réduite. On sait que cette somme suit une loi du Khi-deux.

Mais quel est le nombre de degrés de liberté de cette variable du khi-deux ?
Il y a k carrés, donc a priori k degrés de liberté. Mais on perd toujours un degré de liberté car on a fixé l'effectif total de l'échantillon, $N_1 + N_2 + \cdots + N_k =$

On peut perdre d'autres degrés de liberté si certains paramètres de la loi théorique doivent être estimés à partir de l'échantillon.

1. Si la distribution théorique est entièrement spécifiée, c'est-à-dire si on cherche à déterminer si la distribution observée suit une loi dont les paramètres sont connus avant même de choisir l'échantillon, on a $k-1$ degrés de liberté.

2. S'il faut d'abord estimer r paramètres de la loi à partir des observations de l'échantillon (par exemple on cherche si la distribution est normale mais on ne connaît d'avance ni sa moyenne, ni son écart-type), il n'y a plus que $k-1-r$ degrés de liberté.

Dans le cas général, on dira que la loi du khi-deux suivie par l'écart entre les deux distributions a $\boxed{k-1-r}$ degrés de liberté lorsqu'on a estimé r paramètre(s) de la loi théorique à partir des observations de l'échantillon (avec la possibilité pour r de valoir 0).

V.2 Démarche

On doit maintenant décider, à l'aide de cet indicateur qu'est le χ^2, si les écarts entre les effectifs théoriques et ceux qui résultent des observations sont significatifs d'une différence de distribution, ou s'ils sont dus aux fluctuations d'échantillonnage.

1ère étape : formulation des deux hypothèses.
On va tester l'hypothèse H_0 contre l'hypothèse H_1 :

$$\begin{cases} H_0 : & \text{dans la population initiale, la variable parente } X \text{ suit la loi } \mathscr{L}, \\ & \text{les observations suivent la distribution théorique spécifiée,} \\ H_1 : & X \text{ ne suit pas la loi } \mathscr{L}, \\ & \text{les observations ne suivent pas la distribution théorique spécifiée.} \end{cases}$$

2ème étape : choix de la variable de test.
On utilise la variable aléatoire

$$\chi^2 =$$

3ème étape : détermination des valeurs critiques de χ^2 délimitant les zones d'acceptation et de rejet.

On impose à la zone d'acceptation de H_0 concernant la valeur du χ^2 d'être un intervalle dont 0 est la borne inférieure (car un χ^2 est toujours positif).
Il faut donc déterminer dans la table la valeur maximale z_α de l'écart entre les deux distributions imputable aux variations d'échantillonnage au seuil de signification α, c'est-à-dire vérifiant

$$P(\chi^2 > z_\alpha) =$$

z_α représente donc la valeur critique pour un test sur la concordance entre deux distributions et le test sera toujours unilatéral à droite.

4ème étape : calcul de la valeur de χ^2_{obs} prise dans l'échantillon et conclusion du test.
On calcule la valeur χ^2_{obs} prise par χ^2 dans l'échantillon.

- Si la valeur χ^2_{obs} se trouve dans la zone de rejet, on dira que l'écart observé entre les deux distributions est **statistiquement significatif** au seuil α.
  On rejette H_0 avec un risque α de se tromper.

- Si la valeur χ^2_{obs} se trouve dans la zone d'acceptation, on dira que l'écart observé n'est pas significatif au seuil α. Cet écart est imputable aux fluctuations d'échantillonnage.
  On ne peut pas rejeter H_0.