Cours 6 : TESTS

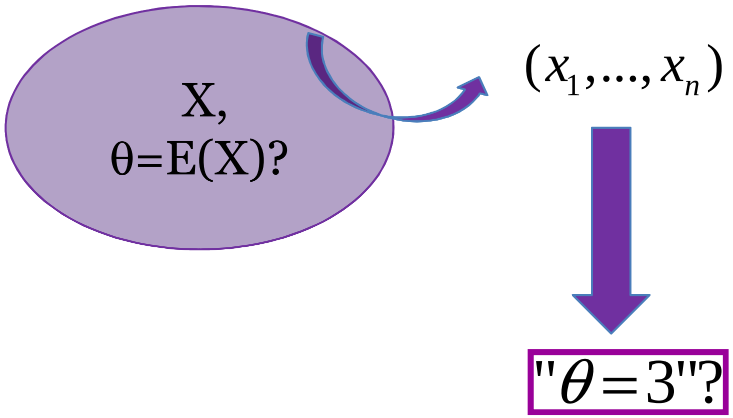
A- Généralités

B- Tests optimaux

A-1 Introduction

- ✓ Soit X une caractéristique dont la distribution dépend de θ inconnu.
- ✓ Faire un test de la valeur du paramètre θ consiste à prendre une décision concernant la valeur de ce paramètre à partir d'un sondage de la population.

Ex : P



Au vu de l'échantillon, peut-on confirmer cette hypothèse?

A-1 Introduction

✓ **Principe**

- Soit X une caractéristique de la population dont la loi dépend d'un paramètre inconnu θ.

- Dans la plupart des situations réelles, il arrive que l'on ait une idée de la valeur de θ . On peut dès lors faire l'hypothèse de cette valeur. On cherchera à valider une hypothèse de type :

$$\boxed{H_0 : \theta = \theta_0}$$

- Les techniques de test vont permettre de confirmer ou d'infirmer cette hypothèse.

A-2 Principe d'un test

- ✓ **Exemple :** Lors du recensement de l'INSEE , Le ministre du travail a affirmé après calcul :

H0 : » le salaire moyen m des français est de 1000 euros par mois ».

Un statisticien est chargé de vérifier ces dires au vu d'un échantillon de la population . On note le salaire X. Des relevés des salaires depuis de nombreuses années ont permis d'établir que la dispersion des salaires français vaut σ. Il prélève un échantillon aléatoire de taille 100 et calcule la valeur de la moyenne d'échantillonnage $\overline{x}$

1. $\overline{x} = 100 \Rightarrow$ rejet de l'hypothèse du ministre que la véritable moyenne m=E(X) est de 1000, étant donné l'écart important existant entre $\overline{x}$ et la valeur hypothétique de m.

2. $\overline{x} = 1001 \Rightarrow$ il semble raisonnable d'accepter l'hypothèse du ministre.

A-2 Principe d'un test

3. $\overline{x} = 980 \text{ ou } 1010 \Rightarrow$ la moyenne d'échantillonnage n'est ni très grande ni très petite par rapport à la valeur hypothétique , de telle sorte que la décision ne s'impose pas d'elle-même.

- Il arrive fréquemment que la valeur $\overline{x}$ ne permette pas de trancher la décision comme dans 3.
- De plus, même lorsqu'elle parait s'imposer (1. et 2.) *on n'est jamais sûr de ne pas être tombé sur un échantillon ayant très peu de chances de se réaliser*.

Comment être sur de prendre la « bonne » décision ? jamais. Tout au plus, on peut prendre la **décision la plus probable.**

A-2 Principe d'un test

Base de décision: combien grande ou « significative » doit être la différence entre $\overline{x}$ et l'affirmation du ministre pour rejeter légitimement cette affirmation?

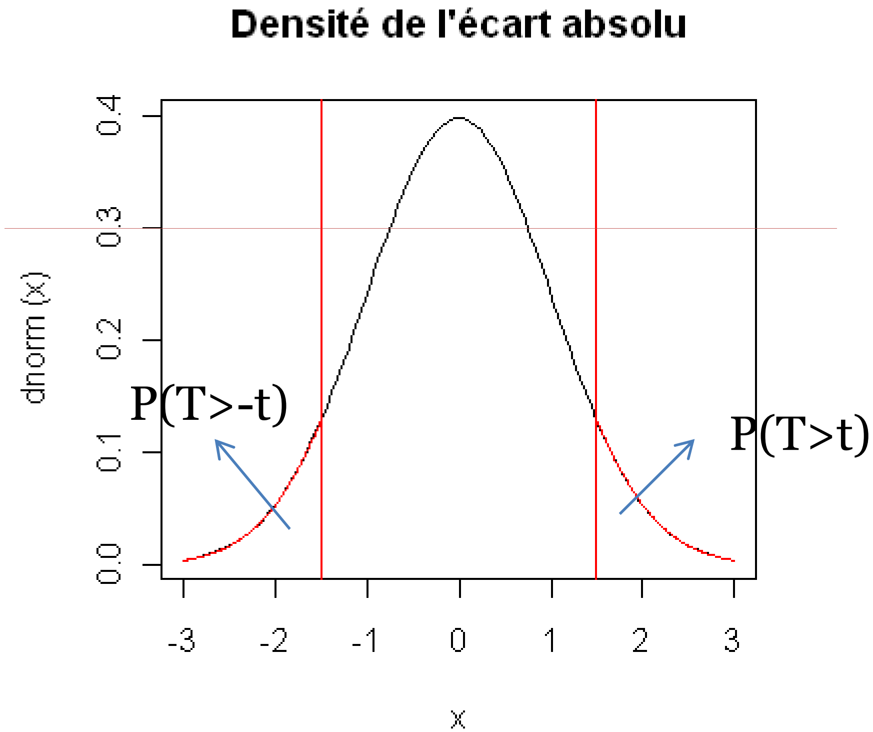
Réponse : on dispose de la loi de l'écart

$$\left(\bar{X} - 1000\right)$$

sous Ho ou de façon équivalente de celle de l'écart réduit :

$$\text{Lorsque n>30, T} = \frac{\bar{X} - 1000}{\sigma / \sqrt{n}} \approx N(0{,}1)$$

On peut alors déterminer la probabilité d'observer un écart absolu >= |t| (valeur de l'écart observé) lorsqu'on est effectivement sous Ho (si cet écart des fréquent ou non).



A-2 Principe d'un test

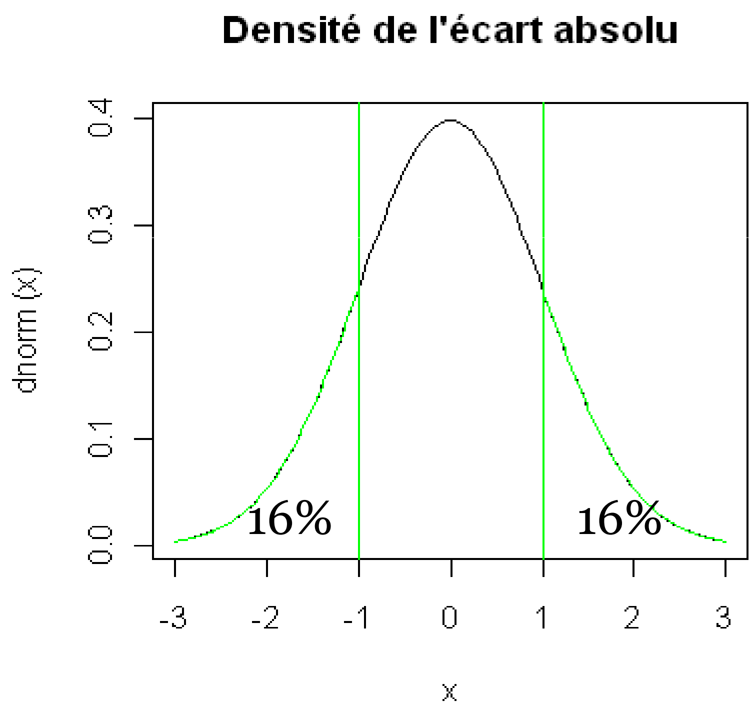
Applications :

- Supposons que l'affirmation du ministre soit vraie (θ=1000) et que la valeur obtenue soit de $\overline{x} = 1010$
  Avec un écart-type de 100, les chances d'obtenir une moyenne qui diffère d'au moins 10 de 1000 , i.e. d'obtenir un écart supérieur ou égal à 10 en valeur absolue, c'est-à-dire un écart réduit de

$$|t| = \frac{|\overline{x} - 1000|}{\sigma / \sqrt{n}} \geq 1$$

est de $P(|T| \geq 1) \approx 32\%$ lorsque l'hypothèse du ministre est vraie..

Cette probabilité est relativement forte et la différence de 10 n'est pas suffisamment significative pour réfuter l'affirmation du ministre.

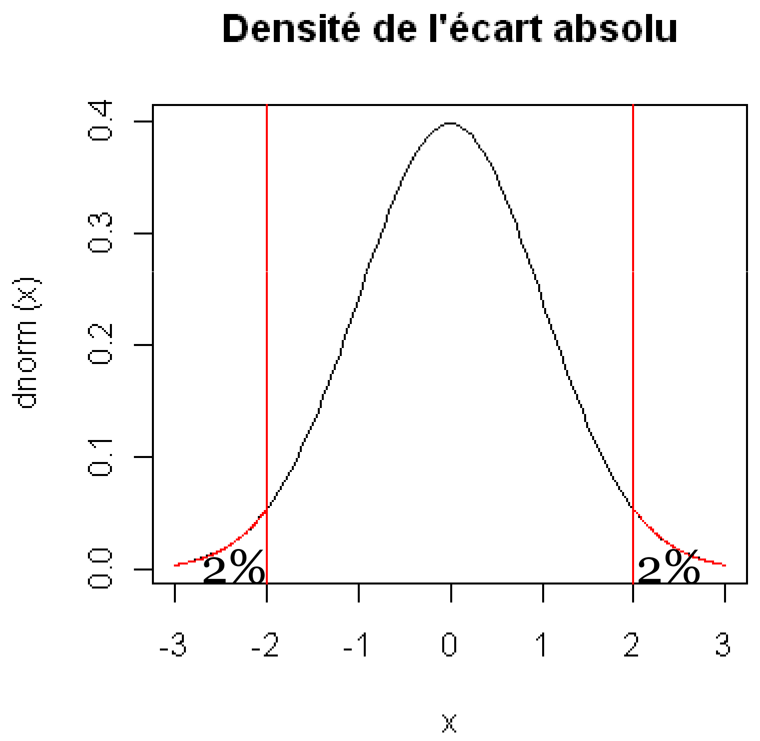


A-2 Principe d'un test

- Supposons l'hypothèse du ministre soit vraie et que la moyenne d'échantillonnage soit de $\overline{x} = 980$ La probabilité d'obtenir une différence d'au moins 20 correspond à un écart-réduit de |z|=2.

$$P(|T| \geq 2) \approx 4\%$$

Les chances de réalisation d'une telle différence sont faibles ; on peut légitiment réfuter l'hypothèse du ministre. Il existe une évidence statistique suffisante pour le faire.



A-2 Principe d'un test

Conclusion : il peut arriver qu'une moyenne d'échantillonnage soit possible mais qu'en même temps, son occurrence soit si faible que l'on soit justifié de ne pas accepter l'hypothèse.
La différence entre la moyenne d'échantillonnage obtenue et la valeur présumée est considérée comme suffisamment grande pour justifier le rejet lorsque la probabilité d'observer une telle différence est *suffisamment faible.*

Que veut dire suffisamment faible ***?*** Ce critère est subjectif. Il dépend des normes fixées par le décideur lui-même. Le seuil de probabilité α toléré dépend du risque qu'il accepte de prendre en rejetant l'hypothèse alors qu'elle est vraie (il est en effet susceptible de se trompe r dans α 100 % des cas s'il fait un grand nombre de tests).

Comme ce seuil est subjectif, il est fixé avant d'entreprendre le test par le décideur lui-même. Cette probabilité s'appelle *seuil de signification du test ou risque de première espèce*.

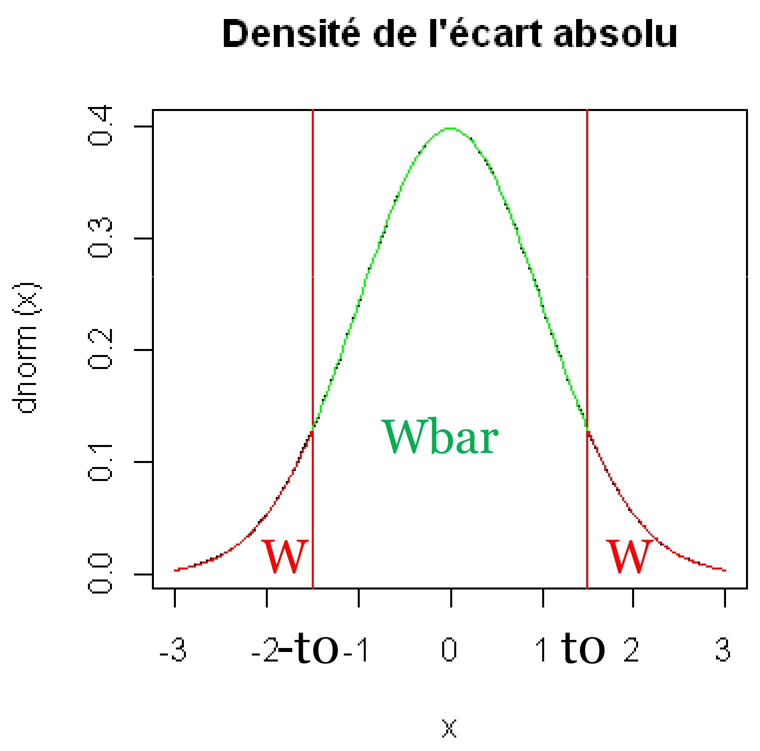
A-2 Principe d'un test

Règle de Décision : Supposons que l'on se fixe un seuil de α=5%. Cela veut dire que l'on considèrera comme rare toute occurrence dont la probabilité d'apparition est inférieure ou égale à 5%.

Soit z0, la valeur telle que P(|T|>t0)=5% On établira la *règle de décision* suivante :

- on rejettera l'hypothèse du ministre dès lors que l'écart réduit observé |t| est supérieur à t0, c'est-à-dire si z se situe dans $W = (-\infty, -t0) \cup (t0, +\infty)$
W est la *zone de rejet du test ou région critique*.

- on l'acceptera si |t| est inférieur à z0 en valeur absolue, c'est-à-dire z se situeen dehors de la région critique : dans la zone d'acceptation $\bar{W} = [-t0, t0]$



A-3 Notions générales

Soit X une v.a. dont la loi dépend d'un paramètre $\theta \in \vartheta \subset R^p$

✓ **Définition d'un test**: Un test est un mécanisme permettant de trancher entre deux hypothèses sur ce paramètre, au vu des résultats d'un sondage :

$$H_0 : \theta \in \vartheta_0$$
$$H_1 : \theta \in \vartheta_1$$

avec $\vartheta_0 \in \vartheta \,, \vartheta_1 \in \vartheta \,, \vartheta_0 \cap \vartheta_1 = \varnothing$

A-3 Notions générales

- **Hypothèse nulle, hypothèse alternative**

- L'hypothèse que l'on cherche à vérifier s'appelle *hypothèse nulle* (notée H0) . Elle porte sur la loi de probabilité *(ou de facon équivalente le paramètre θ de cette loi)* ayant donné naissance à l'échantillon disponible.

$$\boxed{H_0 : \theta \in \vartheta_0}$$

Dans l'exemple précédent, l'hypothèse nulle est $H_0 : m = 1000$

- Si le résultat d'échantillonnage conduit au rejet de l'hypothèse nulle, nous devrons alors en arriver à une autre conclusion. La conclusion acceptée dans ce cas s'appelle *hypothèse alternative* (notée H1) . $\boxed{H_1 : \theta \in \vartheta_1}$

RQ : Il peut y en avoir plusieurs pour la même hypothèse nulle suivant le choix de ϑ_1
Dans l'exemple précédent, $H_1 : m \neq 1000$

- La décision aboutira à choisir entre ces deux hypothèses.

RQ : il se peut cependant que les hypothèses, nulle et alternative, soient fausses toutes les deux , ex: $\mathbf{H}_0$: $m = m_0$ contre $\mathbf{H}_1$: $\mu m > m_0$ alors qu'en réalité $m < m_0$).

A-3 Notions générales

- **Les hypothèses H0 et H1 ne sont pas symétriques** :

o de même que dans un procès d'assises il y a présomption d'innocence, en théorie des tests il y a présomption de *H0: il faut montrer que H0 est peu probable pour décider H1.*

o Par contre accepter H0 peut venir du fait que soit effectivement H0 est vraie, soit *on n'a pas les moyens nécessaires de prouver H1*. (ex: pas assez d'observations).

o H0 est généralement une hypothèse à laquelle on tient particulièrement, pour des raisons qui peuvent être subjectives, ou que l'on aimerait avoir des raisons de réfuter.

✓ - H0 est une hypothèse solidement établie et qui n'a pas été contredite par l'expérience jusqu'alors.

- H0 correspond à une hypothèse de prudence.

  Ex: test d'innocuité d'un vaccin : il est prudent de partir d'une hypothèse défavorable au nouveau produit.

- H0 est la seule hypothèse facile à formuler.

A-3 Notions générales

- H1 s'élabore en fonction de ce que l'on espère découvrir du phénomène.

Ex : si les résultats expérimentaux d'un nouveau traitement de l'hypertension artérielle sont soumis à un test, l'hypothèse nulle est :

H_0 : le nouveau traitement n'a aucun effet.

On espère que les données ne vont pas seulement démentir l'hypothèse nulle, mais vont aussi suggérer une réduction de la tension artérielle des patients. On formulera alors l'hypothèse alternative :

H1 : le nouveau traitement réduit la tension artérielle.

plutôt que l'hypothèse plus générale mais moins utile :

H1 : le nouveau traitement a un effet sur la tension artérielle.

A-3 Notions générales

- **Vocabulaire sur les hypothèses:**

- *Hypothèse simple, hypothèse composite* : H0 (resp. H1) est appelée hypothèse simple si l'ensemble ϑ_0 (resp. ϑ_1) est réduit à un seul point.
  Ex : $H_0 : \theta = \theta_0$. Elle est dite composite dans le cas contraire.

- *Test unilatéral, bilatéral* : lorsque H0 est une hypothèse simple $H_0 : \theta = \theta_0$ le test est dit unilatéral si H1 est de type $H_1 : \theta > \theta_0$ ou $H_1 : \theta < \theta_0$
  et bilatéral si $H_1 : \theta \neq \theta_0$

A-3 Notions générales

- **Exemples d'hypothèses**

o Cas pratique : $\begin{cases} H_0 : E(X) = 1000 \\ H_1 : E(X) \neq 1000 \end{cases}$

o Test bilatéral d'une moyenne : $\begin{cases} H_0 : E(X) = m_0 \\ H_1 : E(X) \neq m_0 \end{cases}$

o Test unilatéral d'une moyenne $\begin{cases} H_0 : E(X) = m_0 \\ H_1 : E(X) \neq m_0 \end{cases}$

o Test d'une variance : $\begin{cases} H_0 : V(X) = \sigma_0^2 \\ H_1 : V(X) > \sigma_0^2 \end{cases}$

o Test d'une distribution : $\begin{cases} H_0 : X \sim N(m,\sigma) \\ H_1 : X \sim E(\lambda) \end{cases}$

A-3 Notions générales

✓ **Risques et probabilités d'erreurs**

La décision conduit à choisir entre H0 et H1, toute décision comportant une part de risque de se tromper. Il y a 4 cas possibles :

Vérité **Décision**	**H0 est vrai** $\boxed{\theta \in \vartheta_0}$	**H1 est vrai** $\boxed{\theta \in \vartheta_0}$
On choisit H0	1-α (OK)	β
On choisit H1	α	1-β (OK)

$$P(\text{choisir } \mathrm{H}_0 / \mathrm{H}_0 \text{ vraie}) = P_{\vartheta_0}(\theta \in \vartheta_0) = 1 - \alpha$$

$$P(\text{choisir } \mathrm{H}_0 / \mathrm{H}_1 \text{ vraie}) = P_{\vartheta_1}(\theta \in \vartheta_0) = \beta$$

$$P(\text{choisir } \mathrm{H}_1 / \mathrm{H}_1 \text{ vraie}) = P_{\vartheta_1}(\theta \in \vartheta_1) = 1 - \beta$$

$$P(\text{choisir } \mathrm{H}_1 / \mathrm{H}_0 \text{ vraie}) = P_{\vartheta_0}(\theta \in \vartheta_1) = \alpha$$

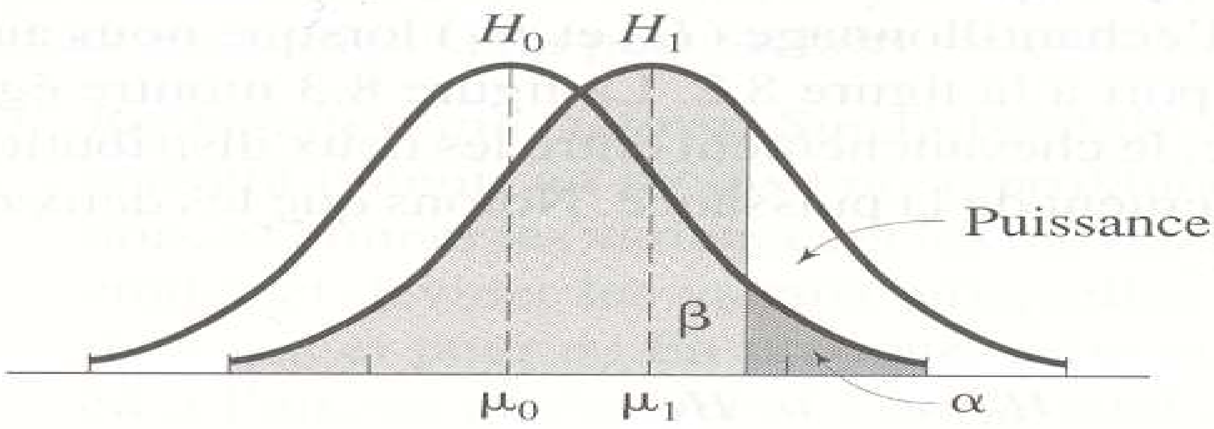
- α est le risque de première espèce ou niveau de signification : c'est le risque que l'on prend en rejetant à tort H0 : le risque de penser que le ministre n'est pas honnête.
- β est le risque de deuxième espèce : c'est le risque que l'on prend en acceptant à tort H0 : le risque de ne pas avoir vu que le ministre est un menteur

A-3 Notions générales

- Dans la pratique des tests statistiques, il est de règle de fixer α (ex : 5%, 1%,10%) car c'est le risque que l'on veut contrôler, que l'on est prêt à prendre en rejetant à tort H0. Ce choix est basé sur la perception que l'on a de la gravité des conséquences d'un rejet injustifié de H0.

- α étant fixé, β sera déterminé comme le résultat d'un calcul . β varie en sens contraire de α : si l'on veut diminuer α on est conduit à ne rejeter H0 que dans des cas rares donc à conserver H0 bien souvent à tort, donc on augmente β (proba sous H1 d'accepter H0)

- 1-β est la probabilité d'opter pour H1 en ayant raison et s'appelle la puissance du test.



A-3 Notions générales

✓ **La statistique du test : variable de décision**

- La conception d'un test passe par l'identification d'une statistique $T = h(X_1,..X_n)$ de l'échantillon, appelée statistique de test**,** jouant le rôle de variable de décision :

- elle doit apporter le maximum d'information sur le problème posé.

  RQ : Pour des tests portant sur un paramètre , la statistique de test est souvent basée sur une fonction d'une statistique exhaustive du paramètre (souvent un écart entre un estimateur exhaustif du paramètre et la valeur de ce paramètre sous H_0.)

- Sa loi doit être différente sous H_0 et sous H_1.

- Il faut que sa loi soit entièrement connue au moins sous H_0.

A-3 Notions générales

- Remarques :

- Concevoir une statistique de test n'est pas une question simple. L*a plupart des tests* classiques reposent sur des statistiques dont l'identification a demandé à leurs auteurs beaucoup d'efforts et d'imagination.

- Une statistique de test n'a aucune raison d'être unique, et le choix entre plusieurs statistiques candidates est une question difficile., qui déterminera entre autre la puissance du test.

- La statistique de test choisie et sa distribution sous H0 dépendent généralement de la distribution de la population et/ou de la taille de l'échantillon. Il s'agit de préciser ces hypothèses et de les vérifier ultérieurement.

  Ex : lorsque X est gaussienne de variance inconnue et n>30, le test de H0 : E(X)=m0 utilise la statistique

$$T = \sqrt{n}\,\frac{\overline{X} - m_0}{S_n} \sim N(0,1) \text{ sous H0}$$

A-3 Notions générales

❑ Exemples de statistiques de test :

- Cas pratique : Si X gaussienne ou n>30 $T = \frac{\bar{X} - 1000}{10}$ T ~ $N(0,1)$ sous H0

- Test simple d'une moyenne : si X gaussienne et n>30

✓ σ connu : $T = \sqrt{n}\frac{\bar{X} - m_0}{\sigma_X}$. T ~ N(0,1) sous H0

✓ σ inconnu : $T = \sqrt{n}\frac{\bar{X} - m_0}{S_n}$. T ~ N(0,1) sous H0

- Test d'une variance : si X est gaussienne

$$T = \frac{(n-1)S_n^2}{\sigma_0^2}.\ \text{T} \sim \chi^2\text{(n-1) sous H0}$$

A-3 Notions générales

✓ **Région critique, zone d'acceptation**

❑ On appelle région critique W du test **toute** région de R^n contenant l'ensemble des réalisations de l'échantillon aléatoire conduisant au rejet de l'hypothèse nulle.

$$W = \{(x_1, \ldots x_n) \in R^n \ / \text{on choisit } \mathrm{H}_1 \ \}$$

Lorsqu'on dispose d'un statistique de test T et de sa loi sous Ho, les réalisations de Wsont celles pour lesquelles la valeur correspondante t de la statistique de test conduit à rejeter Ho. $W = \{t \in R \ \ / \text{on choisit } \mathrm{H}_1 \ \}$

La détermination d'une région critique W dépend de α et se fait en écrivant : $P(W / H_0) = \alpha$

Ex : cas pratique : $W = \{z \in R \ \ / \text{on choisit } \mathrm{H}_1 \ \}$

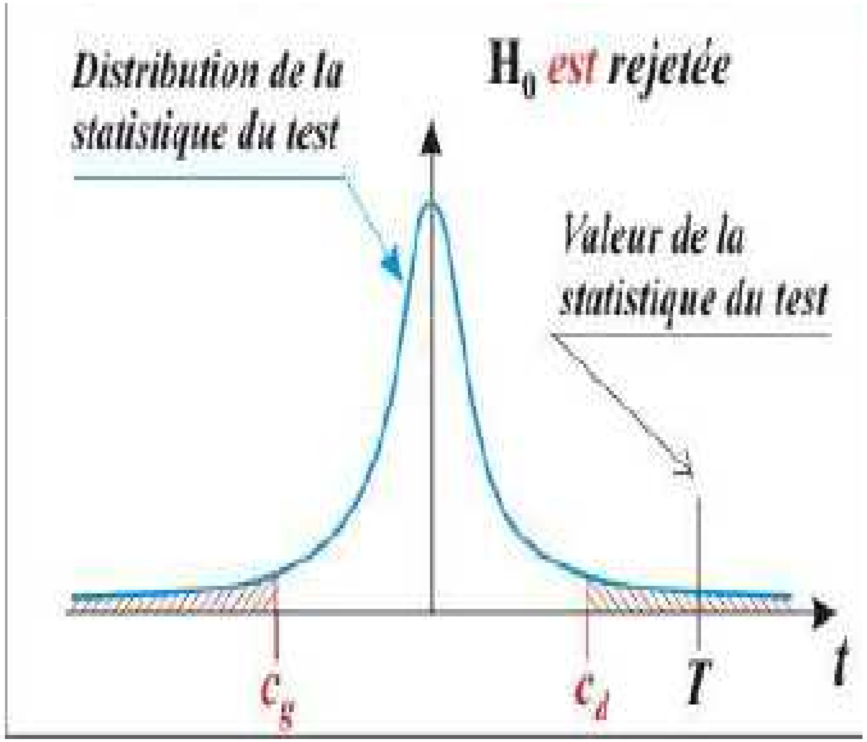
$$P_{H_0}(z \in R / \text{on choisit } \mathrm{H}_1) = \alpha = \left]-\infty, -q_{1-\alpha/2}\right[\cup \left]q_{1-\alpha/2}, \infty\right[$$

❑ A contrario, la zone d'acceptation est le complémentaire dans R^n de la zone de rejet

A-3 Notions générales

- Pour une statistique de test donnée, c'est un intervalle réel (cg, cd) . Cg et cd s'appellent les **valeurs critiques**

- Pour le niveau de signification α choisi, les valeurs critiques sont telles que la somme des aires sous la courbe de la densité de la statistique à gauche de c_g et à droite de c_d est égale à α.



Pour une statistique de test, un seuil
Et une taille d'échantillon donnés, la
Région Critique n'est pas unique . Ex: La
Figure montre que c_g = - c_d , mais
Toute paire de valeurs (c_g , c_d) définissant
Une aire sous la courbe égale à α définit
Une région critique.

A-3 Notions générales

- **Exemple**: On considère un test simple de l'hypothèse

$$H_0 : \theta = \theta_0$$

On suppose que l'on dispose d'une statistique de test T et de sa loi sous H_0.

Une région critique se détermine en cherchant t_0 tel que :

- *Dans le cas d'un test unilatéral* (H_1 de type « $\theta > \theta_0$ » ou « $\theta < \theta_0$ ») $P(T > t_0) = \alpha$ (ou $P(T < -t_0) = \alpha$) s. On a alors

$$W = (t0, +\infty) \text{ ou } W = (-\infty, -t0)$$

- *Dans le cas d'un test bilatéral* (H_1 de type « $\theta \neq \theta_0$ »), $P(|T| > t_0) = \alpha$ On a alors :

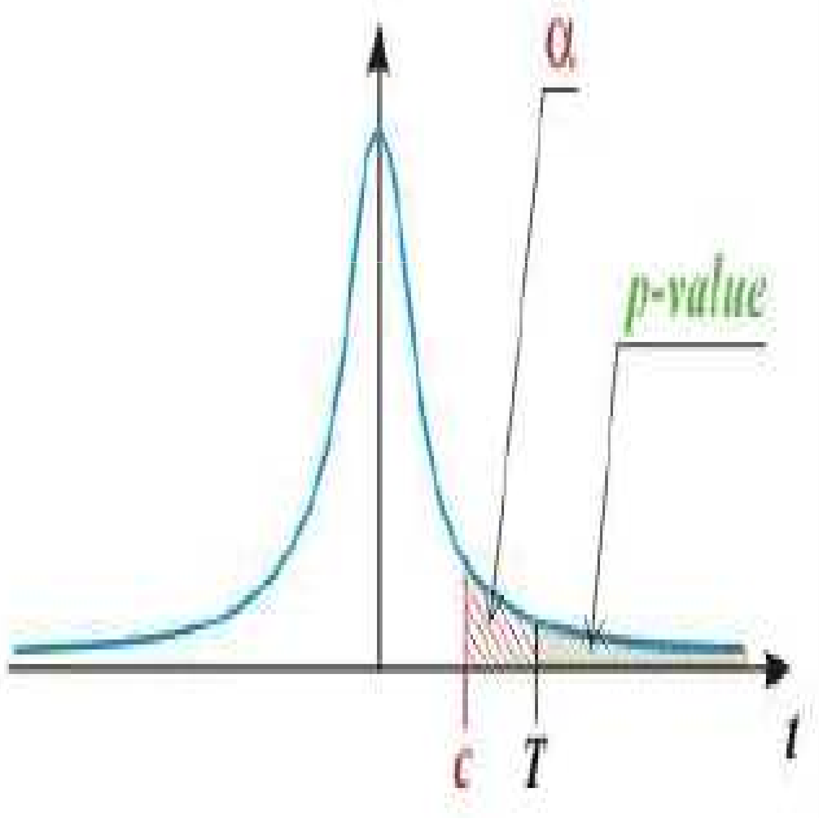
$$W = (-\infty, -t0) \cup (t0, +\infty)$$

A-3 Notions générales

- ***p*-value**

Dans un test unilatéral (disons, à droite), C'est la probabilité sous H0 pour que la statistique prenne une valeur au moins aussi grande que sa valeur calculée T Sur l'échantillon (c'est aire sous la courbe de La densité de la statistique à la droite de T).
H0 est rejetée si la *p*-value est plus petite que le niveau de signification

Dans le cas d'un test bilatéral c'est l'aire sous la courbe de La densité de la statistique à la droite de T et à la gauche de -T,s
T>0.



A-3 Notions générales

- ✓ **Règle de décision :**

Une fois choisie une région critique, l'hypothèse nulle est rejetée comme trop peu Vraisemblable si la valeur de l'échantillon (ou de la statistique de test) est à l'intérieur de cette région critique. Sinon, l'hypothèse nulle n'est pas rejetée. (si elle est dans la zone d'acceptation)

La décision de rejet est donc basée sur le fait que, devant une valeur improbable de l'échantillon nous avons le choix entre deux interprétations :

- $\mathbf{H}_0$ est vraie, et la valeur observée de la statistique est très improbable,
- $\mathbf{H}_0$ est fausse,

et nous favorisons la seconde explication parce que nous ne croyons pas dans les « évènements rares ».

A contrario, "Ne pas rejeter l'hypothèse nulle" ne veut pas dire "l'accepter comme vraie". Cela veut seulement dire que les données ne sont pas en contradiction flagrante avec Cette hypothèse.

A-3 Notions générales

- Exemple : Test bilatéral d'une hypothèse simple sur la moyenne :

  H0 : E(X)=m0 contre H1 : E(X)≠m0

On suppose X gaussien. On utilise la stat de test

$$T = \sqrt{n}\,\frac{\bar{X} - m_0}{S_n}.\ T \sim N(0,1) \text{ sous } H0$$

- On accepte H0 si la valeur t observée de T se situe à l'extérieur de W. ie, si | t| <t0 (t0=quantile d'ordre 1-α/2 de la loi normale centrée réduite)

- On rejette si cette valeur se situe dans W. Dans le cas d'un test bilatéral si |t|> t0

A-3 Notions générales

✓ **Résumé : Démarche d'un test**

1. Choix de H0 et H1
2. Choix de α
3. Détermination de la variable de décision (statistique de test) et de sa loi sous H0 (en fonction de H0 et H1, des hypothèses sur la distribution de X et de la taille de l'échantillon prélevé)
4. Détermination d'une région critique en fonction de α
5. Décision (voir si les observations sont ou pas dans la région critique)
6. Calcul éventuel de la puissance du test

B- Test optimal

B-1 Puissance et région critique

Soit X une v.a. dont la loi dépend d'un paramètre $\theta \in \vartheta \subset R^p$

Soit $\vartheta_0 \in \vartheta, \vartheta_1 \in \vartheta, \vartheta_0 \cap \vartheta_1 = \varnothing$

Au vu des résultats d'un sondage, on veut réaliser un test de niveau α de

$$\begin{cases} H_0 : \theta \in \vartheta_0 \\ \text{contre} \\ H_1 : \theta \in \vartheta_1 \end{cases}$$

Rappels :

$$\boxed{\begin{array}{l} \text{Niveau : } \alpha(\theta) = P(\text{choisir H}_1 / \text{H}_0 \text{ vraie}) = P_{\vartheta_0}(\theta \in \vartheta_1) = P_{\vartheta_0}(W) = P(W / H_1) \\ \text{Puissance : } 1-\beta(\theta) = P(\text{choisir H}_1 / \text{H}_1 \text{ vraie}) = P_{\vartheta_1}(\theta \in \vartheta_1) = P_{\vartheta_1}(W) = P(W / H_1) \end{array}}$$

Lorsque H0 est une hypothèse composite α dépend de la vrai valeur inconnue du paramètre θ étudiée. Le niveau du test est alors défini par : $\boxed{\alpha = \sup_{\theta \in \vartheta_0} P_\theta(W)}$

A moins que H1 soit une hypothèse simple, β dépend de θ. Normalement, la puissance augmente lorsqu'on s'éloigne de l'hypothèse nulle.

B-1 Puissance et région critique

Un test de niveau et taille d’échantillon donnés est d’autant meilleur que sa puissance est grande, i.e. que l’on ne se trompe pas en rejetant H0.

Pour α et n fixés, la qualité d’un test est caractérisée par sa puissance.

La puissance

- ✓ Croit avec niveau de signification α
- ✓ Croit avec la taille n de l’échantillon
- ✓ Dépend de la région critique

Pour une taille d'échantillon et un niveau de signification donnés,

- ✓ la puissance du test ne dépend plus que du choix de la région critique choisie
- ✓ toutes les régions critiques (i.e. tous les tests) ne fournissent pas la même puissance.

B-1 Puissance et région critique

Exemple : Soit X gaussienne de variance inconnue, de moyenne inconnue.
On teste au seuil 5%:

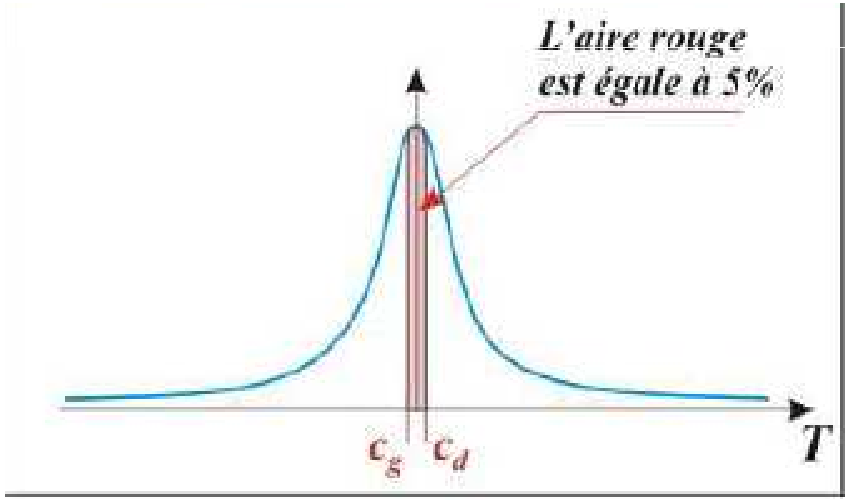
$$\begin{cases} H_0 : E(X) = 0 \\ H_1 : E(X) \neq 0 \end{cases}$$

Statistique de test :

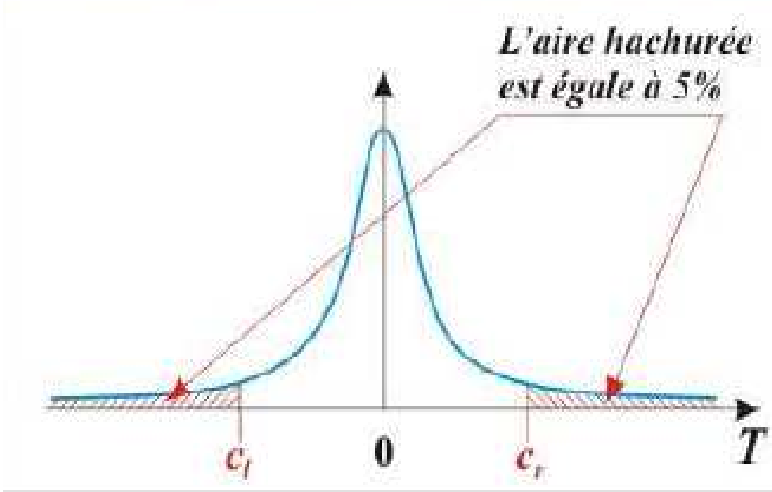
$$T = \sqrt{n}\,\overline{X}_n / S_n \quad \text{Sous } \mathrm{H}_0, T \sim T(n-1)$$

B-1 Puissance et région critique

1° région critique (i.e. premier test): la présence de la valeur de la statistique de test dans la région critique ne suggère pas que H_1 soit vraie : le test a une puissance très faible.



2° région critique l(i.e. deuxième test) a présence de la valeur de la statistique dans la région critique suggère que la moyenne de X est plus petite (resp. plus grande) que 0, un argument en faveur de H_1 : La puissance du test est importante.



B-1 Puissance et région critique

Objectif : Etant donnés n et α fixés, on aimerait identifier la région critique qui maximise la puissance du test., c'est-à-dire un domaine de R^n parmi l'ensemble de toutes les réalisations possibles de l'échantillon $(X_1, \ldots X_n)$ de probabilité inférieurs ou égale à α et de puissance maximale.

NB **Définir un test revient à définir sa région critique.**

B-2 Qualité d'un test

Définitions:

- Un test de région critique W* est meilleur qu'un test de région critique W si et seulement si :

$$P(W^* / H_0) \leq P(W / H_0),$$

$$P(W^* / H_1) \geq P(W / H_1)$$

- **Test UPP** : On dira qu'un test W* est uniformément le plus puissant (UPP) dans la classe des tests de niveau α si et seulement si il est meilleur que tous les autres tests de niveau α.

$$\forall W, W^* \in R^n \; / \; \sup_{\theta \in \vartheta_0} P_\theta(W^*) = \sup_{\theta \in \vartheta_0} P_\theta(W \quad) = \alpha,$$

$$P(W^* / H_1) \geq P(W / H_1)$$

- **Meilleure Région Critique (MRC) :** C'est la région critique du test UPP, si celui-ci existe.

B-2 Qualité d'un test

- Test sans biais : Un test est dit sans biais ssi $\alpha < 1 - \beta(\theta)$

- Test convergent : Un test est dit convergent ssi $1 - \beta(\theta) \xrightarrow[n \to \infty]{} 1$

B-2 Qualité d’un test

Existence d’un test UPP:

Le problème du choix de la région critique optimale (ou de facon équivalente du test optimal) a été résolu théoriquement par Neyman et Pearson pour le test entre deux hypothèses simples.

Dans le cas général d’une hypothèse H1 composite, H1 regroupe plusieurs valeurs du paramètre inconnu θ. La puissance d’un test dépend de la vrai valeur de θ, que l’on ne connait pas et que l’on étudie. C’est donc une fonction $(1-\beta(\theta))$ de θ.

Il est alors généralement impossible de trouver un test ayant une puissance optimale pour tous les θ possibles . En général, une région critique fournit une puissance supérieure à une autre pour certaines valeurs de θ et inférieure pour d’autres valeurs de θ.

B3- Test entre deux hypothèses simples: La méthode de Neyman-Pearson

Soit X une variable aléatoire de densité f(x,θ) où $\theta \in R$ est un paramètre inconnu.
On cherche à tester :

$$(0)\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta = \theta_1 \end{cases}$$

Au niveau α, au vu d'un échantillon de données de taille n.

NB : Ici, α et β sont des valeurs et non des fonctions, car la valeur du paramètre est Connue sous chaque hypothèses.

Problème : Existe-t-il une variable de décision et des valeurs de cette variable qui fournissent la région critique optimale ? Quelles sont-elles?

Formalisation : on cherche $W \subset R^n$ satisfaisant $P(W / H_0) = \alpha$ et qui

maximise $1 - \beta = P(W / H_1)$

B3- Test entre deux hypothèses simples: La méthode de Neyman-Pearson

Soit L la vraisemblance de l'échantillon.

$$P(W/H_i) = P_{\theta_i}((X_1,..,X_n) \in W) = \int_W L(x_1,...,x_n,\theta_i)dx_1...dx_n \ , i \in \{0,1\}.$$

Fonction à maximiser

$$1-\beta = P(W/H_1) = \int_W \frac{L(x_1,...,x_n,\theta_1)}{L(x_1,...,x_n,\theta_0)} L(x_1,...,x_n,\theta_0)dx_1...dx_n \ .$$

Contrainte : $\alpha = P(W/H_0) = P(W/H_0) = \int_W L(x_1,...,x_n,\theta_0)dx_1...dx_n$.

B3- Test entre deux hypothèses simples: La méthode de Neyman-Pearson

Solution : Théorème de Neyman-Pearson (admis) : Pour n et α fixés, il existe un test de niveau α UPP (optimal) pour (0) dans la classe des tests de niveau α , défini par la région critique :

$$\boxed{W = \left\{ (x_1,\ldots,x_n) \in R^n \,/\, \frac{L(x_1,\ldots,x_n,\theta_1)}{L(x_1,\ldots,x_n,\theta_0)} > k_\alpha \right\}}$$

où k_α est déterminée de telle sorte que : $P(W/H_0) = \alpha$

- W s'appelle la région critique de Neyman-Pearson
- Le test de (0) ayant cette région critique s'appelle test de Neyman-Pearson ou test du rapport de vraisemblance.

Interprétation : Lorsque H0 est vraie, l'échantillon dont on dispose est plus probable (plus vraisemblable) lorsque il est tiré d'une loi de paramètre θ_0 que de paramètre θ_1 : le rapport des vraisemblances est plus petit que sous H1.

B3- Test entre deux hypothèses simples: La méthode de Neyman-Pearson

Propriétés du test de Neyman-Pearson (admises)

P0 : Le test ainsi constitué est l'unique test UPP

P1 : Le test ainsi constitué est sans biais

P2 : Le test ainsi constitué est convergent, i.e.

P3 : S'il existe une statistique exhaustive T pour θ de densité g(t,θ)

$$W = \left\{ (x_1, \ldots, x_n) \in R^n \,/\, \frac{g(t,\theta_1)}{g(t,\theta_0)} > k_\alpha \right\}$$

B3- Test entre deux hypothèses simples: La méthode de Neyman-Pearson

Exemple Soit X de loi N(m,σ) où σ connu et m inconnu. On veut tester :

$$\begin{cases} H_0 : m = m_0 \\ H_1 : m = m_1 \end{cases} \qquad \text{avec } m_1 > m_0$$

Calcul de la Région critique de Neyman-Pearson (de niveau α):

- **Sol 1 :** $$L(x_1,\ldots,x_n,m) = \frac{e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-m)^2}}{(2\pi\sigma^2)^{n/2}} = \frac{e^{-\frac{1}{2\sigma^2}(n\overline{x^2}-2mn\overline{x}+nm^2)}}{(2\pi\sigma^2)^{n/2}}$$

donc $$W = \left\{(x_1,\ldots,x_n)\in R^n / \overline{x} > \frac{1}{2}\left((m_1+m_0)+\frac{2\sigma^2}{n(m_1-m_0)}\ln k_\alpha\right)\right\}$$

De la forme : $$\boxed{W = \left\{(x_1,\ldots,x_n)\in R^n / \overline{x} > k\right\}}$$

B3- Test entre deux hypothèses simples: La méthode de Neyman-Pearson

- La variable de décision (statistique de test) est $\bar{X}$. Sous Ho, $\bar{X} \sim N(m_0, \frac{\sigma}{\sqrt{n}})$
- k se détermine alors en considérant que (avec Φ fdr de la loi N(0,1))

$$P(W / H_0) = \alpha = P_{\theta_0}(\bar{X} > k) = 1 - \Phi\left(\frac{k - m_0}{\sigma / \sqrt{n}}\right).$$

$$\Leftrightarrow k = \frac{\sigma}{\sqrt{n}} q_{1-\alpha} + m_0 \text{ où } q_{1-\alpha} \text{ est le quantile d'ordre } 1-\alpha \text{ d'une N(0,1)}$$

➢ **Sol2** : $\bar{X}$ est une statistique exhaustive pour m, de loi $N(m_0, \frac{\sigma}{\sqrt{n}})$.

D'après P3, la région critique est de la forme

$$W = \left\{ (x_1, \ldots, x_n) \in R^n / \frac{g(t, \theta_1)}{g(t, \theta_0)} > k_\alpha \right\}$$

On retrouve la même que précédemment.

B3- Test entre deux hypothèses simples: La méthode de Neyman-Pearson

Calcul de la puissance :

Sous H1, $\bar{X} \sim N(m_1, \frac{\sigma}{\sqrt{n}})$

On calcule la puissance par :

$$1-\beta = P_{\theta_1}(\bar{X} > k) = 1-\Phi\left(\frac{k-m_1}{\sigma/\sqrt{n}}\right)$$

$\Leftrightarrow \beta$ est la valeur de la fdr d'une N(0,1) au point $\frac{k-m_1}{\sigma/\sqrt{n}}$

Application numérique :

α=5%, m0=0, m1=1, s=1, n=100. $k = k = \frac{\sigma}{\sqrt{n}} q_{1-\alpha} + m_0 = \frac{1}{10} * 1.64$

Région critique : $W = \{\bar{x} > 0.164\}$

Puissance : $1-\beta = P_{\theta_1}(\bar{X} > k) = 1-\Phi(-8.35) = \Phi(8.35) = 1$

B4- Test d'une hypothèse simple contre une alternative composite unilatérale

Soit X une variable aléatoire de densité f(x,θ) où θ est un paramètre inconnu. On cherche à tester :

$$(1)\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta > \theta_0 \end{cases} \quad \text{ou} \quad (1')\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta < \theta_0 \end{cases}$$

Au risque α, au vu d'un échantillon de données de taille n.

NB : Ici, α est une valeur et β est une fonction de la valeur du paramètre θ inconnu.

Problème : Quelle variable de décision et quelles valeurs de cette variable fournit la région critique optimale (en terme de puissance)?

Solution (Propriété): Si la région critique du test de Neyman-Pearson pour (0) ne dépends pas de θ1 , alors ce test est UPP pour (1) (si θ1 > θ0) ou (1')(si θ0 > θ1) parmi tous les tests de niveau α.

Exemple : Dans l'exemple précédent, on a vu que $W = \left\{ \overline{x} > \frac{\sigma}{\sqrt{n}} q_{1-\alpha} + m_0 \right\}$ ne dépend pas de m1. Donc le test de région critique W est UPP pour $\begin{cases} H_0 : m = m_0 \\ H_1 : m > m_0 \end{cases}$

B5- Cas général

- **On considère le cas général du test d'hypothèses**

$$\begin{cases} H_0 : \theta \in \vartheta_0 \\ \text{contre} \\ H_1 : \theta \in \vartheta_1 \end{cases}$$

Au risque α, au vu d'un échantillon de données de taille n.

NB : Lorsque ϑ_0 (resp. ϑ_1) n'est pas réduit à un point , α (resp. β) est une fonction de la valeur du paramètre θ inconnu.

Problème : Quelle variable de décision et quelles valeurs de cette variable fournit la région critique optimale (en terme de puissance)?

Solution : En dehors des cas particuliers (0), (1), (1') et de quelques autres, il n'existe pas en général , de test optimal pour (3)

B5- Cas général

- **Test du rapport des vraisemblances** : on appelle test du rapport de vraisemblance, le test défini par la région critique :

$$\boxed{W = \left\{ (x_1,\ldots,x_n) \in R^n \,/\, \lambda = \frac{\sup_{\theta \in \vartheta_1} L(x_1,\ldots,x_n,\theta)}{\sup_{\theta \in \vartheta_0} L(x_1,\ldots,x_n,\theta)} > k_\alpha \right\}}$$

où k_α est déterminée de telle sorte que : $P(W / H_0) \leq \alpha$

NB : On obtient le même test en remplaçant λ par $\lambda' = \frac{\sup_{\theta \in \vartheta_0 \cup \vartheta_1} L(x_1,\ldots,x_n,\theta)}{\sup_{\theta \in \vartheta_0} L(x_1,\ldots,x_n,\theta)}$

dès lors que le sup dans la définition de λ ‘est atteint.

Propriétés

- La loi asymptotique de 2lnλ est une loi du chi2 à p degrés de liberté (p=dim de θ)
- Le test du rapport de vraisemblances est convergent

B5- Cas général

- **Cas particuliers du test du rapport des vraisemblances :**
- Cas (0) = Le test du rapport des vraisemblances est le test de Neyman-Pearson, optimal.
- Cas (1), (1'): Le test du rapport des vraisemblances est le test de Neyman-Pearson, optimal.

- **Cas du test de**

$$(2)\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta \neq \theta_0 \end{cases}$$

$$\boxed{W = \left\{ (x_1,\dots,x_n) \in R^n \,/\, \lambda = \frac{\sup_{\theta\in\vartheta} L(x_1,\dots,x_n,\theta)}{L(x_1,\dots,x_n,\theta_0)} = \frac{L(x_1,\dots,x_n,\hat{\theta}_n)}{L(x_1,\dots,x_n,\theta_0)} > k_\alpha \right\}}$$

Où $\hat{\theta}_n$ est l'estimateur du maximum de vraisemblance de θ.

- **Exemple :** Soit X une population de loi N(m,σ), σ connu. On veut tester

$$\begin{cases} H_0 : m = m_0 \\ H_1 : m \neq m_0 \end{cases}$$

Au vu d'un échantillon iid de la population. Le test du rapport de vraisemblance donne ¨la région critique:

$$\boxed{W = \left\{ (x_1, \ldots, x_n) \in R^n \,/\, \left| \overline{x} - m_0 \right| > \frac{\sigma}{\sqrt{n}} q_{1-\alpha/2} \right\}}$$