Chapitre 9

Tests d'hypothèse

Tests d'hypothèse

1. Notion de test

- Tests d'ajustement
- Tests de conformité
- Tests d'égalité
- Tests d'indépendance

Principe des tests d'hypothèse

2. Notion d'hypothèse

Deux types d'hypothèses:

- **H_0: hypothèse nulle. C'est l'hypothèse qu'on veut tester**

  Elle postule la non différence et permet de fixer les paramètres de la distribution de la variable aléatoire étudiée

- **H_1: hypothèse alternative**

  C'est l'hypothèse qu'on retient si H_0 est rejetée.

Principe des tests d'hypothèse

3. Notion de statistique

Une statistique est une variable aléatoire S dont la valeur numérique obtenue pour le test considéré, s_{obs}, permet de décider si H_0 est vraie ou fausse

Principe des tests d'hypothèse

4. Notion de règle de décision

Pour décider d'accepter ou de rejeter H_0, il faut une règle de décision

On compare s_{obs} à deux bornes de rejet

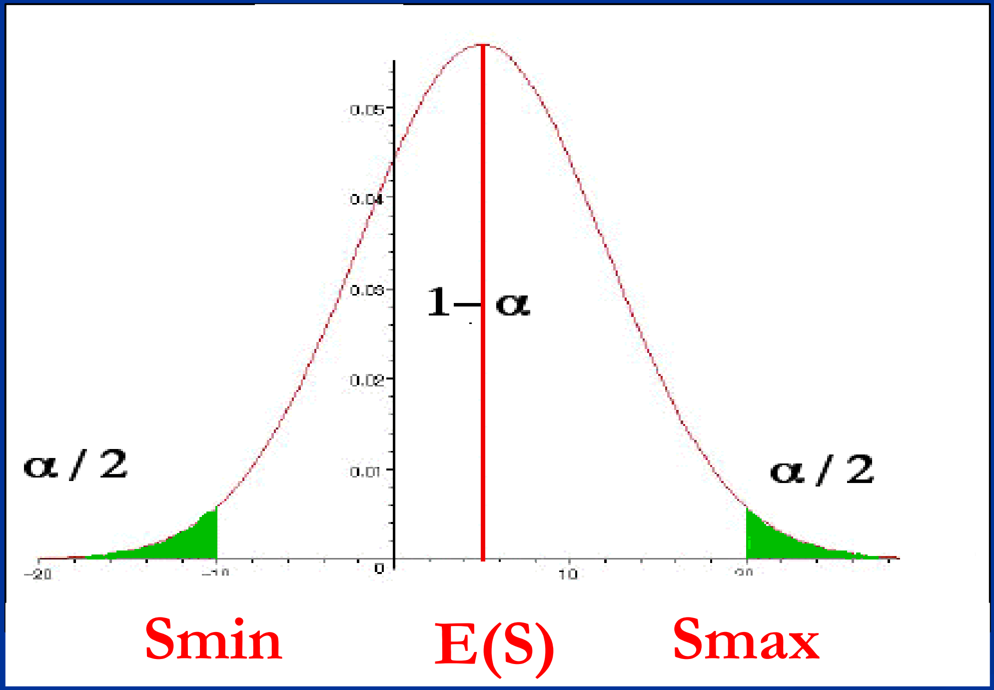
On acceptera H_0 si s_{obs} est supérieur à S_{min} et inférieur à S_{max}

On rejetera H_0 si s_{obs} est inférieur à S_{min} ou supérieur à S_{max}

Principe des tests d'hypothèse

5. Notion de risque

- Risque de première espèce α



Sous H0
Distribution de S connue

⇩

$P(S_{min}<S<S_{max})=1-\alpha$ et donc

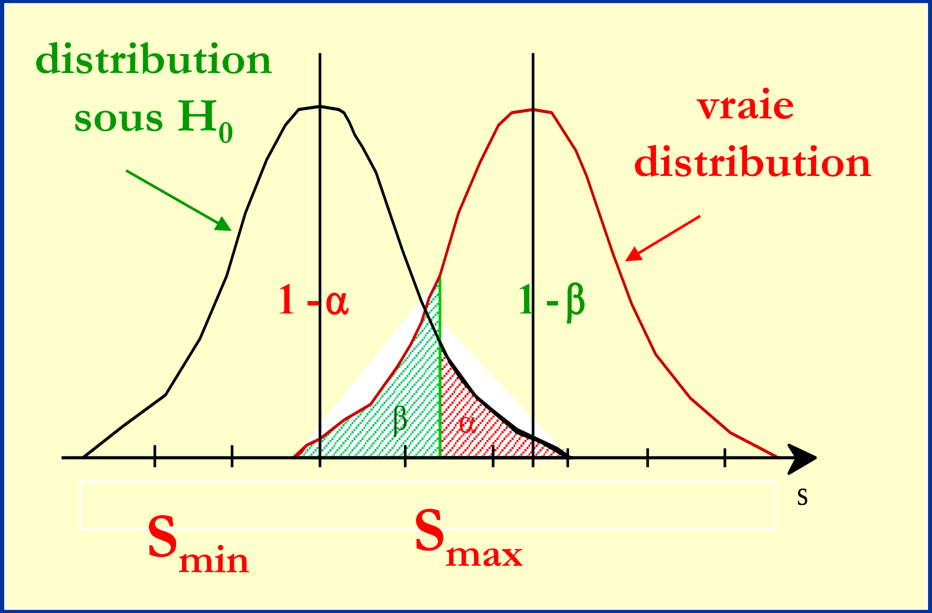
$\alpha = P(S>S_{max}) + P(S<S_{min})$

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

Principe des tests d'hypothèse

5. Notion de risque

- Risque de deuxième espèce β



$\beta = P(\text{accepter } H_0/H_1 \text{ vraie})$

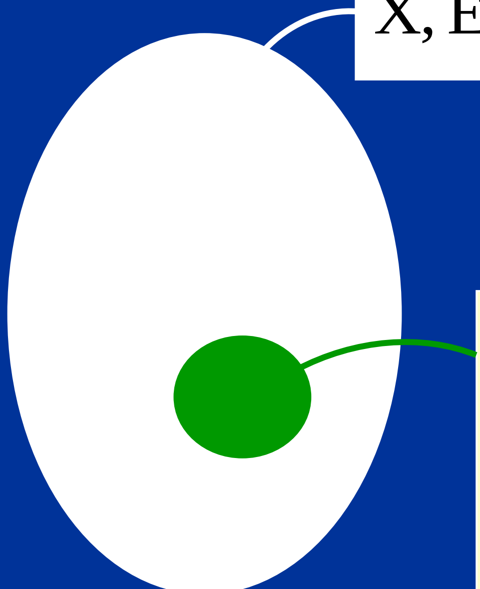
$1-\beta$ est la puissance du test

$\beta \neq 1-\alpha$

Tests de conformité

1. Comparaison d'une moyenne observée et d'une moyenne théorique

$$X, E(X) = \mu; V(X) = \sigma^2$$



n

$\overline{x}$

s^2

Question : $\mu = \mu_{th}$?

Hypothèse H_0 : $\mu = \mu_{th}$

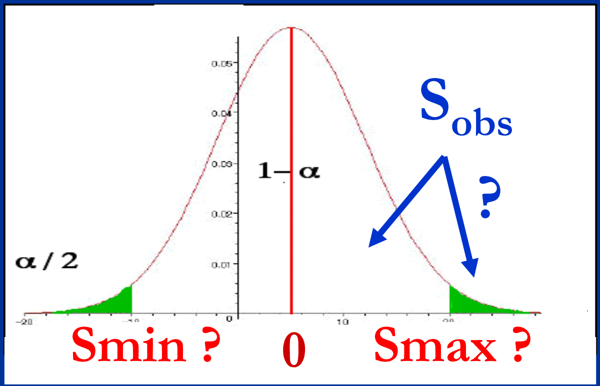
$$\overline{X}, E(\overline{X}) = \mu_{th}; V(\overline{X}) = \frac{\sigma^2}{n}$$

Tests de conformité

1. Comparaison d'une moyenne observée et d'une moyenne théorique

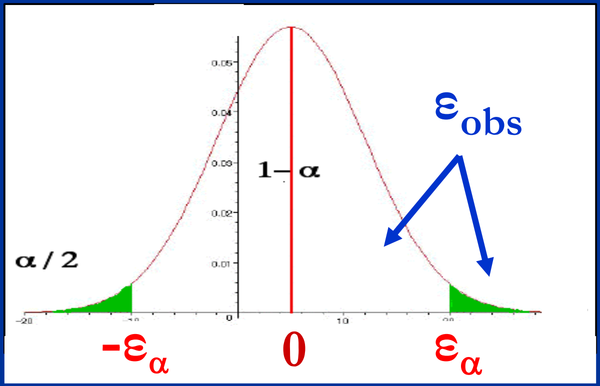
- $n > 30$, σ^2 connue

$$\overline{X} \sim N\left(\mu_{th}, \frac{\sigma}{\sqrt{n}}\right)$$



$$S = \overline{X} - \mu_{th} \quad S \sim N\left(0, \frac{\sigma}{\sqrt{n}}\right) \quad S_{obs} = \bar{x} - \mu_{th}$$

On ne connaît pas S_{min} et S_{max}



$$\mathcal{E} = \frac{\overline{X} - \mu_{th}}{\frac{\sigma}{\sqrt{n}}} \quad \mathcal{E} \sim N(0,1) \quad \varepsilon_{obs} = \frac{\bar{x} - \mu_{th}}{\frac{\sigma}{\sqrt{n}}}$$

$P(|\varepsilon| \geq \varepsilon_\alpha) = \alpha$ table de l'écart-réduit

Tests de conformité

1. Comparaison d'une moyenne observée et d'une moyenne théorique

- $n > 30$, σ^2 connue

Règle de décision:

On compare $|\varepsilon_{obs}|$ à ε_α lue dans la table de l'écart-réduit

- Si $|\varepsilon_{obs}| < \varepsilon_\alpha$, on accepte H_0 et on considère $\mu = \mu_{th}$

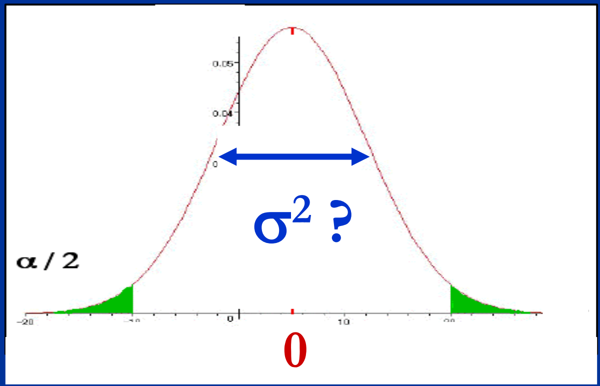
- Si $|\varepsilon_{obs}| > \varepsilon_\alpha$, on rejette H_0 et on considère $\mu \neq \mu_{th}$

Tests de conformité

1. Comparaison d'une moyenne observée et d'une moyenne théorique

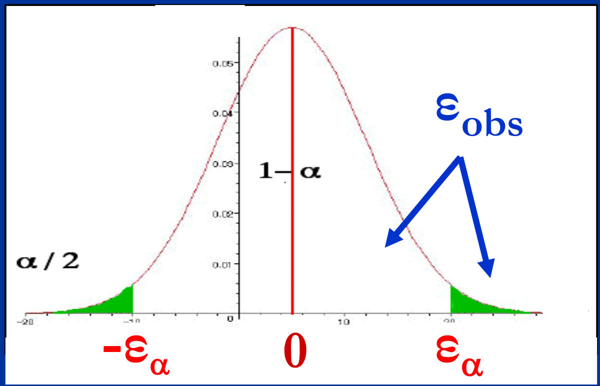
- $n > 30, \sigma^2$ inconnue

$$\overline{X} \sim N\left(\mu_{th}, \frac{\sigma}{\sqrt{n}}\right)$$



$$S = \overline{X} - \mu_{th} \quad S \sim N\left(0, \frac{\sigma}{\sqrt{n}}\right) \quad S_{obs} = \bar{x} - \mu_{th}$$

On ne connaît pas σ^2



$$\mathcal{E} = \frac{\overline{X} - \mu_{th}}{\frac{\hat{\sigma}}{\sqrt{n}}} \quad \mathcal{E} \sim N(0,1) \quad \varepsilon_{obs} = \frac{\bar{x} - \mu_{th}}{\frac{\hat{\sigma}}{\sqrt{n}}} = \frac{\bar{x} - \mu_{th}}{\frac{s}{\sqrt{n-1}}}$$

$P(|\varepsilon| \geq \varepsilon_\alpha) = \alpha$ table de l'écart-réduit

Tests de conformité

1. Comparaison d'une moyenne observée et d'une moyenne théorique

- $n > 30$, σ^2 inconnue

Règle de décision:

On compare $|\varepsilon_{obs}|$ à ε_α lue dans la table de l'écart-réduit

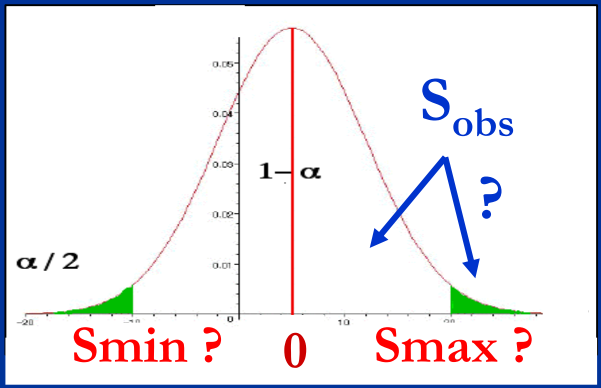
- Si $|\varepsilon_{obs}| < \varepsilon_\alpha$, on accepte H_0 et on considère $\mu = \mu_{th}$
- Si $|\varepsilon_{obs}| > \varepsilon_\alpha$, on rejette H_0 et on considère $\mu \neq \mu_{th}$

Tests de conformité

1. Comparaison d'une moyenne observée et d'une moyenne théorique

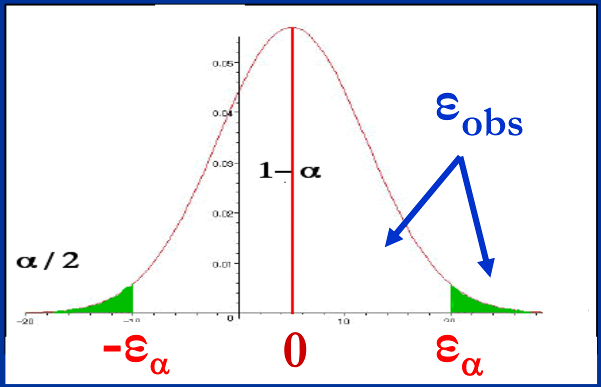
- $n < 30$, $X \sim N$, σ^2 connue

$$\overline{X} \sim N\left(\mu_{th}, \frac{\sigma}{\sqrt{n}}\right)$$



$$S = \overline{X} - \mu_{th} \quad S \sim N\left(0, \frac{\sigma}{\sqrt{n}}\right) \quad S_{obs} = \bar{x} - \mu_{th}$$

On ne connaît pas S_{min} et S_{max}



$$\mathcal{E} = \frac{\overline{X} - \mu_{th}}{\frac{\sigma}{\sqrt{n}}} \quad \mathcal{E} \sim N(0,1) \quad \varepsilon_{obs} = \frac{\bar{x} - \mu_{th}}{\frac{\sigma}{\sqrt{n}}}$$

$P(|\varepsilon| \geq \varepsilon_\alpha) = \alpha$ table de l'écart-réduit

Tests de conformité

1. Comparaison d'une moyenne observée et d'une moyenne théorique

- $n < 30$, $X \sim N$ σ^2 connue

Règle de décision:

On compare $|\varepsilon_{obs}|$ à ε_α lue dans la table de l'écart-réduit

- Si $|\varepsilon_{obs}| < \varepsilon_\alpha$, on accepte H_0 et on considère $\mu = \mu_{th}$

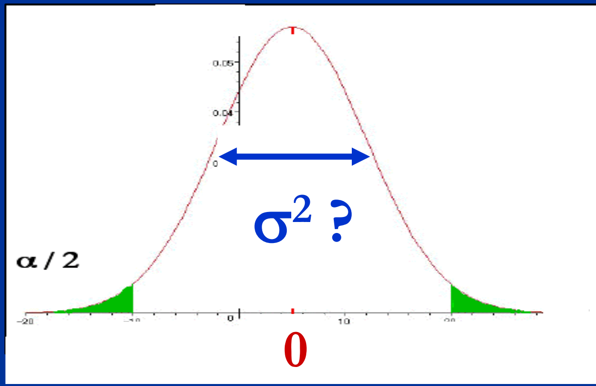
- Si $|\varepsilon_{obs}| > \varepsilon_\alpha$, on rejette H_0 et on considère $\mu \neq \mu_{th}$

Tests de conformité

1. Comparaison d'une moyenne observée et d'une moyenne théorique

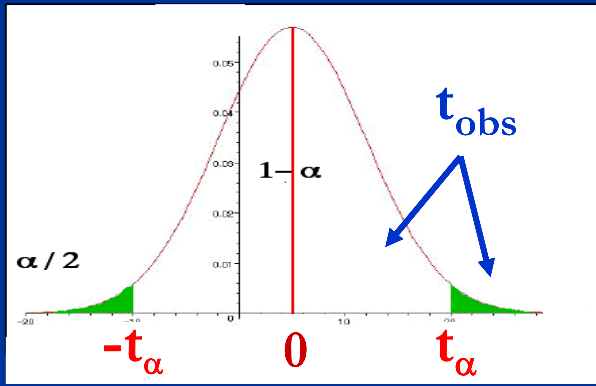
- $n < 30$, $X \sim N$, σ^2 inconnue

$$\overline{X} \sim N\left(\mu_{th}, \frac{\sigma}{\sqrt{n}}\right)$$



$$S = \overline{X} - \mu_{th} \quad S \sim N\left(0, \frac{\sigma}{\sqrt{n}}\right) \quad S_{obs} = \bar{x} - \mu_{th}$$

On ne connaît pas σ^2



$$T = \frac{\overline{X} - \mu_{th}}{\frac{\hat{\sigma}}{\sqrt{n}}} \quad T \sim \underset{(n\text{-}1)\text{ddl}}{\text{Student}} \quad t_{obs} = \frac{\bar{x} - \mu_{th}}{\frac{\hat{\sigma}}{\sqrt{n}}} = \frac{\bar{x} - \mu_{th}}{\frac{s}{\sqrt{n-1}}}$$

$P(|T| \geq t_\alpha) = \alpha$ table de t de Student

Tests de conformité

1. Comparaison d'une moyenne observée et d'une moyenne théorique

- $n < 30$, $X \sim N$ σ^2 inconnue

Règle de décision:

On compare $|t_{obs}|$ à t_α à (n-1) ddl lue dans la table de Student

- Si $|t_{obs}| < t_\alpha$, on accepte H_0 et on considère $\mu = \mu_{th}$

- Si $|t_{obs}| > t_\alpha$, on rejette H_0 et on considère $\mu \neq \mu_{th}$

Tests de conformité

1. Comparaison d'une moyenne observée et d'une moyenne théorique

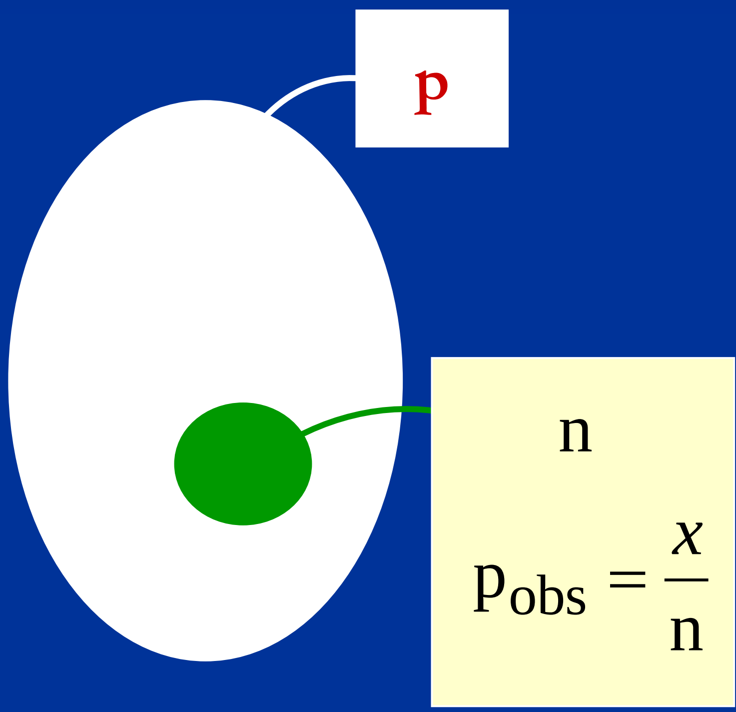
- $n < 30$, $X \sim$ loi quelconque

$$\overline{X} \sim ?$$

Tests non paramétriques !

Tests de conformité

2. Comparaison d'une proportion observée et d'une proportion théorique



Question : $p = p_{th}$?

Hypothèse H_0 : $p = p_{th}$

$\rightarrow X \sim B(n, p_{th})$

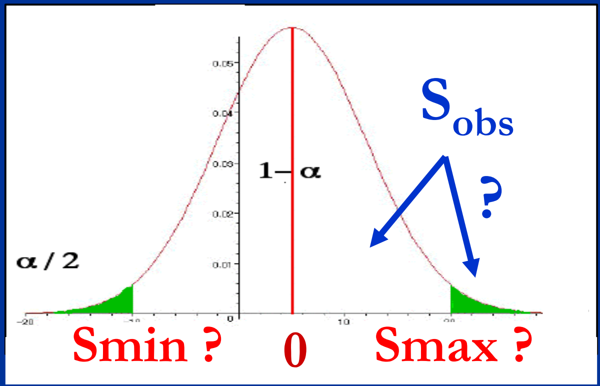
$\rightarrow \quad \Psi = \frac{X}{n}$

si $n > 30$ $\rightarrow$ $\Psi \sim N\left(p_{th}, \sqrt{\frac{p_{th} q_{th}}{n}}\right)$

Tests de conformité

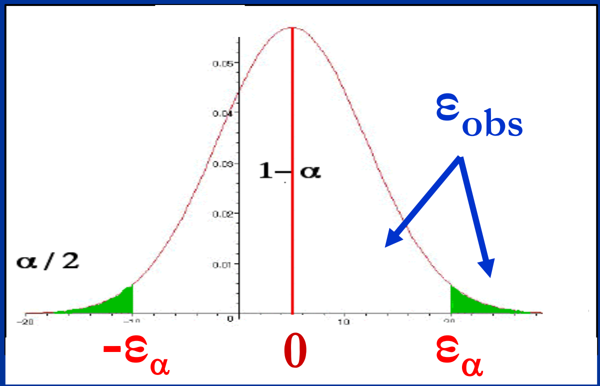
2. Comparaison d'une proportion observée et d'une proportion théorique

- n > 30 et np > 5



$$S = \Psi - p_{th} \quad S \sim N\left(0, \sqrt{\frac{p_{th}q_{th}}{n}}\right) \quad S_{obs} = p_{obs} - p_{th}$$

On ne connaît pas S_{min} et S_{max}



$$\mathcal{E} = \frac{\Psi - p_{th}}{\sqrt{\frac{p_{th}q_{th}}{n}}} \quad \mathcal{E} \sim N(0,1) \quad \varepsilon_{obs} = \frac{p_{obs} - p_{th}}{\sqrt{\frac{p_{th}q_{th}}{n}}}$$

$P(|\varepsilon| \geq \varepsilon_\alpha) = \alpha$ table de l'écart-réduit

Tests de conformité

2. Comparaison d'une proportion observée et d'une proportion théorique

- $n > 30$ et $np > 5$

Règle de décision:

On compare $|\varepsilon_{obs}|$ à ε_α lue dans la table de l'écart-réduit

- Si $|\varepsilon_{obs}| < \varepsilon_\alpha$, on accepte H_0 et on considère $p = p_{th}$
- Si $|\varepsilon_{obs}| > \varepsilon_\alpha$, on rejette H_0 et on considère $p \neq p_{th}$

Tests de conformité

2. Comparaison d'une proportion observée et d'une proportion théorique

- n < 30

$\rightarrow X \sim B(n, p_{th})$

$\rightarrow \quad \Psi = \frac{X}{n}$

si n < 30 $\rightarrow \quad \Psi \sim ?$

On utilise un test non paramétrique

Tests de conformité

2. Comparaison d'une proportion observée et d'une proportion théorique

- test du χ^2

Effectifs \ Modalités	Succès	Echecs	Total
Observés (O_i)	x	n-x	n
Théoriques (T_i)	np_{th}	$n(1-p_{th})$	n

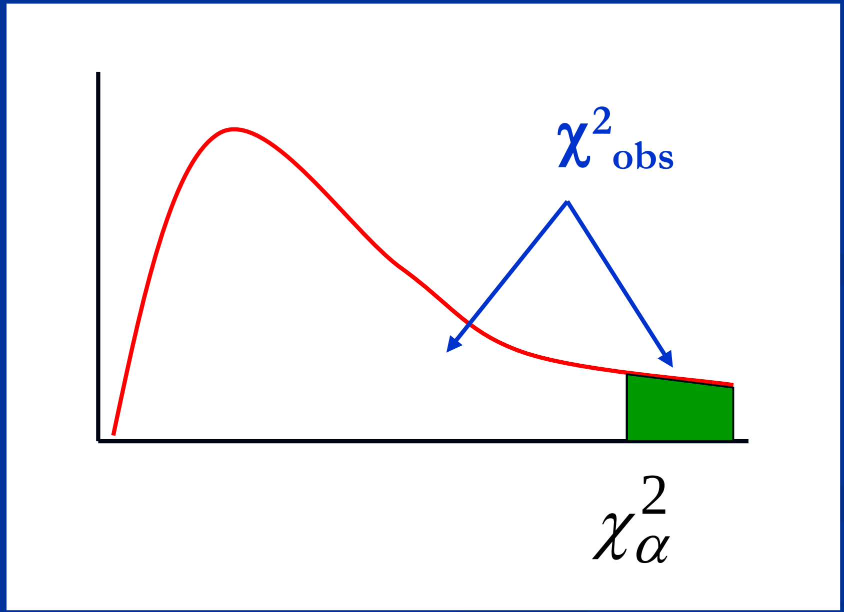
Tests de conformité

2. Comparaison d'une proportion observée et d'une proportion théorique

- test du χ^2

$$S = \sum_{i=1}^{2} \frac{(O_i - T_i)^2}{T_i} \qquad S \sim \chi^2 \text{ à 1 ddl}$$

$$S_{obs} = \chi^2_{obs} = \sum_{i=1}^{2} \frac{(o_i - t_i)^2}{t_i}$$



Les t_i doivent être > 5

$P(\chi^2 \geq \chi^2_\alpha \text{ à 1 ddl}) = \alpha$

Tests de conformité

2. Comparaison d'une proportion observée et d'une proportion théorique

- test du χ^2

Règle de décision:

On compare χ^2_{obs} à χ^2_α à 1 ddl lu dans la table du χ^2

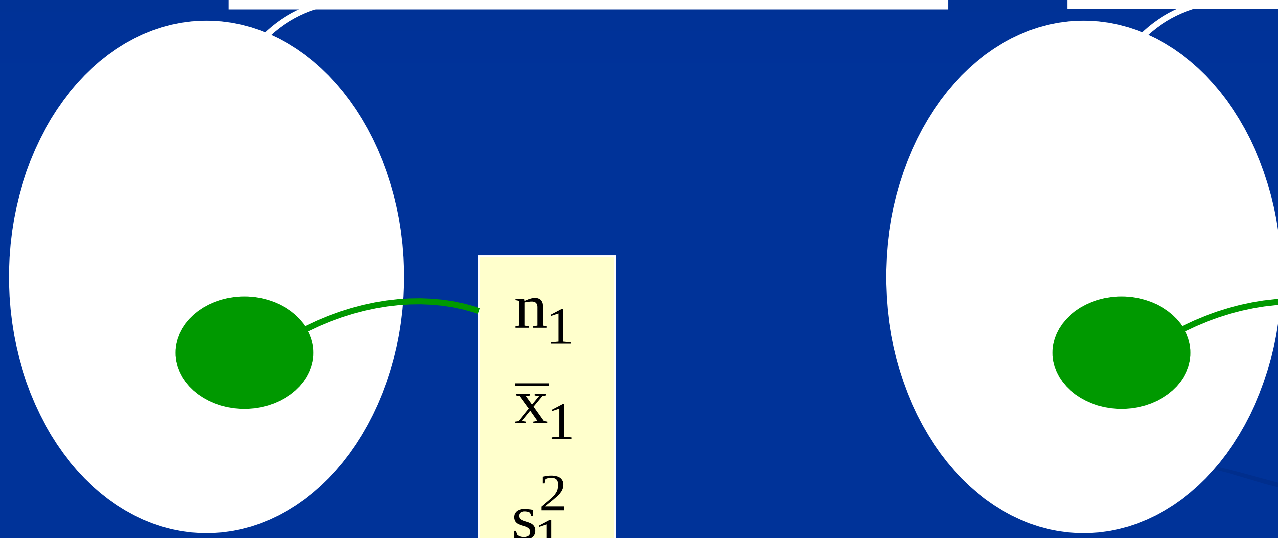
- Si $\chi^2_{obs} < \chi^2_\alpha$, on accepte H_0 et on considère $p = p_{th}$
- Si $\chi^2_{obs} > \chi^2_\alpha$, on rejette H_0 et on considère $p \neq p_{th}$

Tests d'égalité

2. Comparaison de deux moyennes observées

$$X_1 \; E(X_1) = \mu_1, V(X_1) = \sigma_1^2$$

$$X_2 \; E(X_2) = \mu_2, V(X_2) = \sigma_2^2$$



n_1, $\bar{x}_1$, s_1^2

n_2, $\bar{x}_2$, s_2^2

Question : $\mu_1 = \mu_2$?

Hypothèse H_0 : $\mu_1 = \mu_2 = \mu$

Tests d'égalité

2. Comparaison de deux moyennes observées

Sous l'hypothèse H_0

Population 1

$\rightarrow$ $E(X_1) = \mu$

$\rightarrow$ $E(\overline{X}_1) = \mu \,;\, V(\overline{X}_1) = \dfrac{\sigma_1^{\,2}}{n_1}$

Population 2

$\rightarrow$ $E(X_2) = \mu$

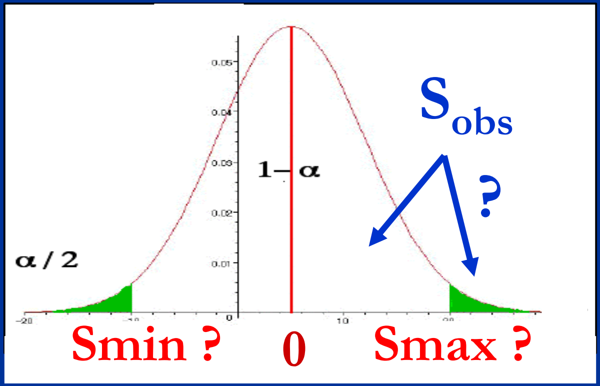
$\rightarrow$ $E(\overline{X}_2) = \mu \,;\, V(\overline{X}_2) = \dfrac{\sigma_2^{\,2}}{n_2}$

Tests d'égalité

2. Comparaison de deux moyennes observées

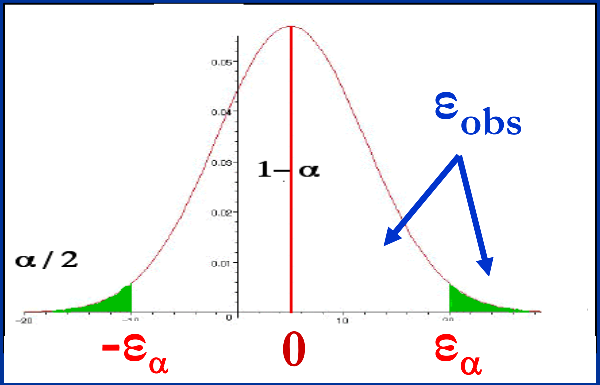
- n_1 et n_2 > 30, σ_1^2 et σ_2^2 connues

$$\overline{X}_1 \; et \; \overline{X}_2 \sim N$$



$$S = \overline{X}_1 - \overline{X}_2 \quad S \sim N\left(0, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right) \quad S_{obs} = \bar{x}_1 - \bar{x}_2$$

On ne connaît pas S_{min} et S_{max}



$$\mathcal{E} = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad \mathcal{E} \sim N(0,1) \quad \varepsilon_{obs} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$P(|\varepsilon| \geq \varepsilon_\alpha) = \alpha$ table de l'écart-réduit

Tests d'égalité

2. Comparaison de deux moyennes observées

- n_1 et $n_2 > 30$, σ_1^2 et σ_2^2 connues

Règle de décision:

On compare $|\varepsilon_{obs}|$ à ε_α lue dans la table de l'écart-réduit

- Si $|\varepsilon_{obs}| < \varepsilon_\alpha$, on accepte H_0 et on considère $\mu_1 = \mu_2$

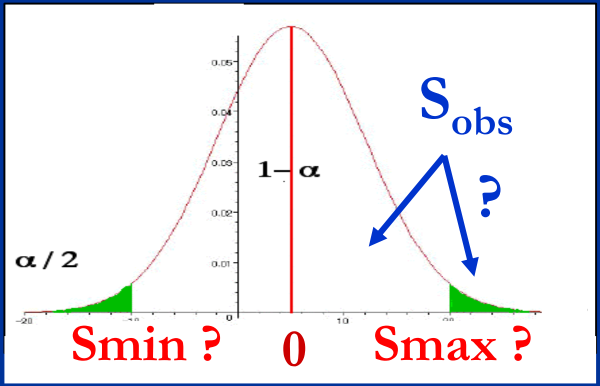
- Si $|\varepsilon_{obs}| > \varepsilon_\alpha$, on rejette H_0 et on considère $\mu_1 \neq \mu_2$

Tests d'égalité

2. Comparaison de deux moyennes observées

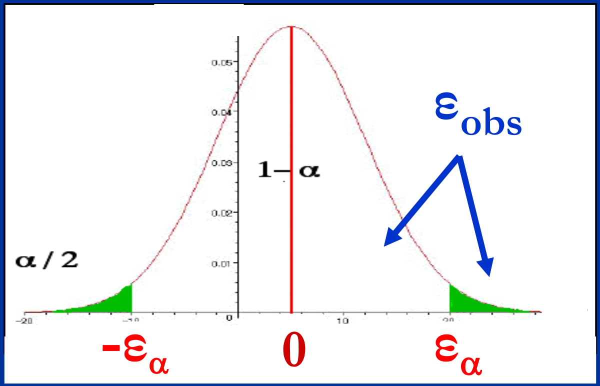
- n_1 et $n_2 > 30$, σ_1^2 et σ_2^2 inconnues

$$\overline{X}_1 \; et \; \overline{X}_2 \sim N$$



$$S = \overline{X}_1 - \overline{X}_2 \quad S \sim N\left(0, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right) \quad S_{obs} = \bar{x}_1 - \bar{x}_2$$

On ne connaît pas σ_1^2 et σ_2^2



$$\mathcal{E} = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}} \quad \mathcal{E} \sim N(0,1) \quad \varepsilon_{obs} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1 - 1} + \frac{s_2^2}{n_2 - 1}}}$$

$P(|\varepsilon| \geq \varepsilon_\alpha) = \alpha$ table de l'écart-réduit

Tests d'égalité

2. Comparaison de deux moyennes observées

- n_1 et $n_2 > 30$, σ_1^2 et σ_2^2 inconnues

Règle de décision:

On compare $|\varepsilon_{obs}|$ à ε_α lue dans la table de l'écart-réduit

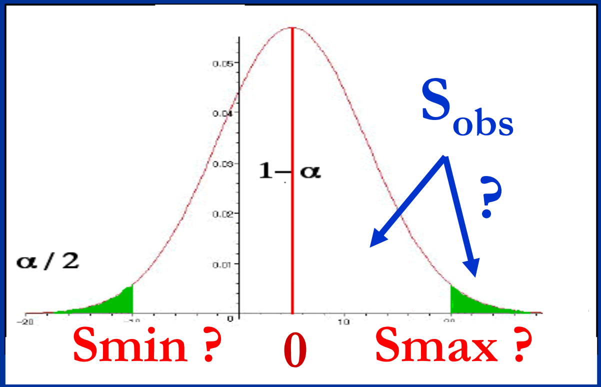
- Si $|\varepsilon_{obs}| < \varepsilon_\alpha$, on accepte H_0 et on considère $\mu_1 = \mu_2$
- Si $|\varepsilon_{obs}| > \varepsilon_\alpha$, on rejette H_0 et on considère $\mu_1 \neq \mu_2$

Tests d'égalité

2. Comparaison de deux moyennes observées

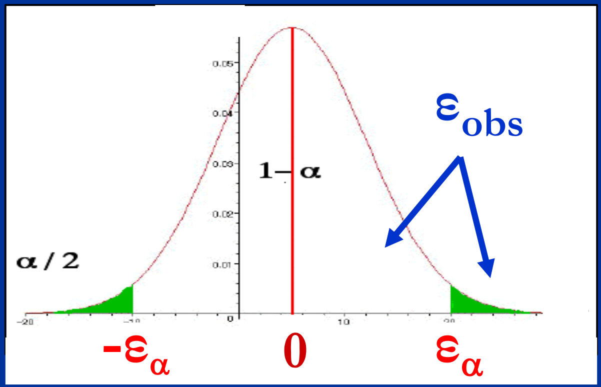
- n_1 et $n_2 < 30$, $X \sim N$, σ_1^2 et σ_2^2 connues

$$\overline{X}_1 \; et \; \overline{X}_2 \sim N$$



$$S = \overline{X}_1 - \overline{X}_2 \quad S \sim N\left(0, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right) \quad S_{obs} = \bar{x}_1 - \bar{x}_2$$

On ne connaît pas S_{min} et S_{max}



$$\mathcal{E} = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad \mathcal{E} \sim N(0,1) \quad \varepsilon_{obs} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$P(|\varepsilon| \geq \varepsilon_\alpha) = \alpha$ table de l'écart-réduit

Tests d'égalité

2. Comparaison de deux moyennes observées

- n_1 et $n_2 < 30$, $X \sim N$, σ_1^2 et σ_2^2 connues

Règle de décision:

On compare $|\varepsilon_{obs}|$ à ε_α lue dans la table de l'écart-réduit

- Si $|\varepsilon_{obs}| < \varepsilon_\alpha$, on accepte H_0 et on considère $\mu_1 = \mu_2$

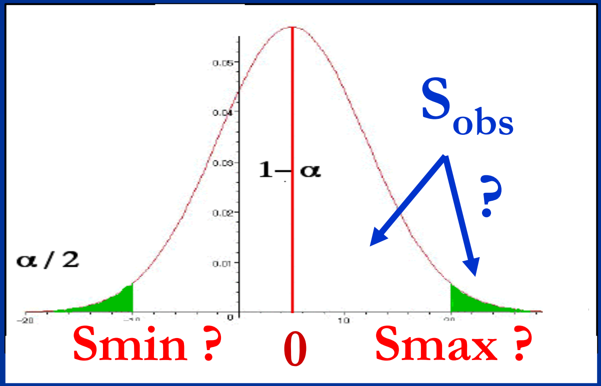
- Si $|\varepsilon_{obs}| > \varepsilon_\alpha$, on rejette H_0 et on considère $\mu_1 \neq \mu_2$

Tests d'égalité

2. Comparaison de deux moyennes observées

- n_1 et $n_2 < 30$, $X \sim N$, σ_1^2 et σ_2^2 inconnues

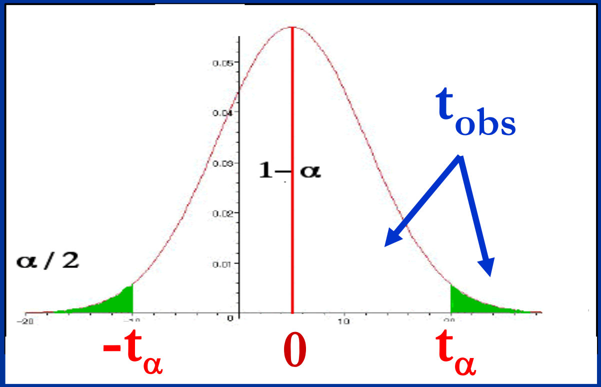
$$\overline{X}_1 \; et \; \overline{X}_2 \sim N$$



$$S = \overline{X}_1 - \overline{X}_2 \quad S \sim N\left(0, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right) \quad S_{obs} = \bar{x}_1 - \bar{x}_2$$

On ne connaît pas σ_1^2 et σ_2^2

Si σ_1^2 et σ_2^2 égales



$$T = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\hat{\sigma}^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad T \sim \underset{(n_1+n_2-2)\text{ddl}}{\text{Student}} \quad t_{obs} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\hat{\sigma}^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \text{ avec } \hat{\sigma}^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}$$

$P(|T| \geq t_\alpha) = \alpha$ table de Student

Tests d'égalité

2. Comparaison de deux moyennes observées

- n_1 et $n_2 < 30$, $X \sim N$, σ_1^2 et σ_2^2 inconnues

Règle de décision:

On compare $|t_{obs}|$ à t_α à (n_1+n_2-2) ddl lue dans la table de Student

- Si $|t_{obs}| < t_\alpha$, on accepte H_0 et on considère $\mu_1 = \mu_2$

- Si $|t_{obs}| > t_\alpha$, on rejette H_0 et on considère $\mu_1 \neq \mu_2$

Tests d'égalité

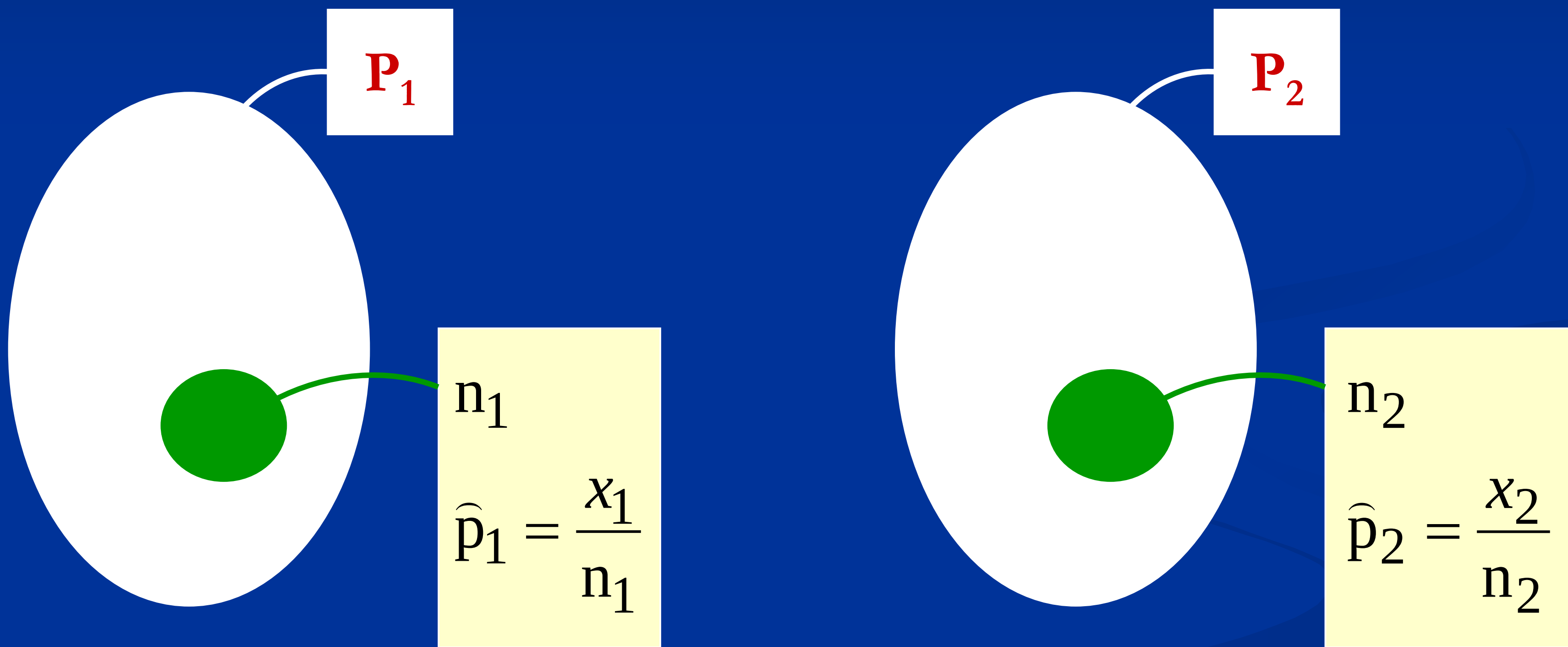
2. Comparaison de deux moyennes observées

- σ_1^2 et σ_2^2 inconnues, n_1 et/ou $n_2 < 30$, X~loi quelconque ou σ_1^2 et σ_2^2 différentes

Tests non paramétriques !

Tests d'égalité

3. Comparaison de deux proportions observées



Question : $P_1 = P_2$?

Hypothèse H_0 : $P_1 = P_2 = P$

Tests d'égalité

3. Comparaison de deux proportions observées

Sous l'hypothèse H_0

Population 1

$\rightarrow X_1 \sim B(n_1,p)$

$\rightarrow \Psi_1 = \dfrac{X_1}{n_1}$

si $n_1 > 30$ $\rightarrow$ $\Psi_1 \sim N\left(p, \sqrt{\dfrac{pq}{n_1}}\right)$

Population 2

$\rightarrow X_2 \sim B(n_2,p)$

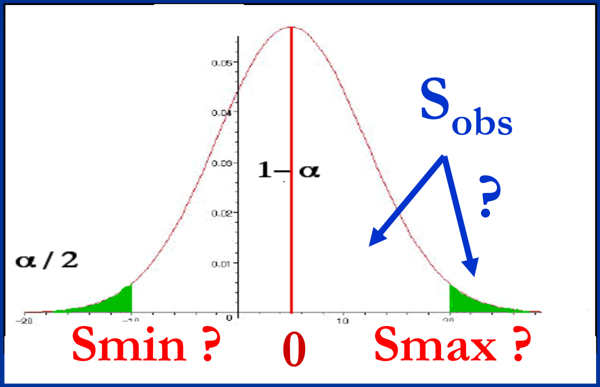
$\rightarrow \Psi_2 = \dfrac{X_2}{n_2}$

si $n_2 > 30$ $\rightarrow$ $\Psi_2 \sim N\left(p, \sqrt{\dfrac{pq}{n_2}}\right)$

Tests d'égalité

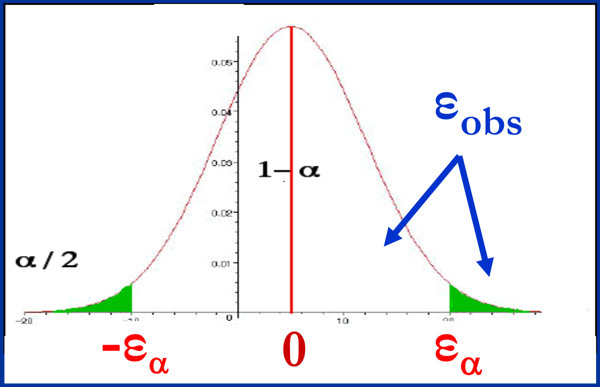
3. Comparaison de deux proportions observées

- n_1 et $n_2 > 30$



$$S = \Psi_1 - \Psi_2 \quad S \sim N\left(0, \sqrt{pq\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}\right) \quad S_{obs} = \hat{p}_1 - \hat{p}_2$$

On ne connaît pas S_{min} et S_{max}



$$\mathcal{E} = \frac{\Psi_1 - \Psi_2}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \quad \mathcal{E} \sim N(0,1) \quad \varepsilon_{obs} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

$$\text{avec} \quad \hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1+n_2} = \frac{x_1+x_2}{n_1+n_2}$$

$P(|\varepsilon| \geq \varepsilon_\alpha) = \alpha$ table de l'écart-réduit

Tests d'égalité

3. Comparaison de deux proportions observées

- n_1 et $n_2 > 30$

Règle de décision:

On compare $|\varepsilon_{obs}|$ à ε_α lue dans la table de l'écart-réduit

- Si $|\varepsilon_{obs}| < \varepsilon_\alpha$, on accepte H_0 et on considère $P_1 = P_2$

- Si $|\varepsilon_{obs}| > \varepsilon_\alpha$, on rejette H_0 et on considère $P_1 \neq P_2$

Tests d'égalité

3. Comparaison de deux proportions observées

- test du χ^2

Tableau des effectifs observés

Echan. \ Modalités	Succès	Echecs	Total
Echantillon 1	x_1	n_1-x_1	n_1
Echantillon 2	x_2	n_2-x_2	n_2
Total	x	$n-x$	n

Tests d'égalité

3. Comparaison de deux proportions observées

- test du χ^2

Tableau des effectifs théoriques

Modalités / Echan.	Succès	Echecs	Total
Echantillon 1	n_1x/n	$n_1(n-x)/n$	n_1
Echantillon 2	n_2x/n	$n_2(n-x)/n$	n_2
Total	x	$n-x$	n

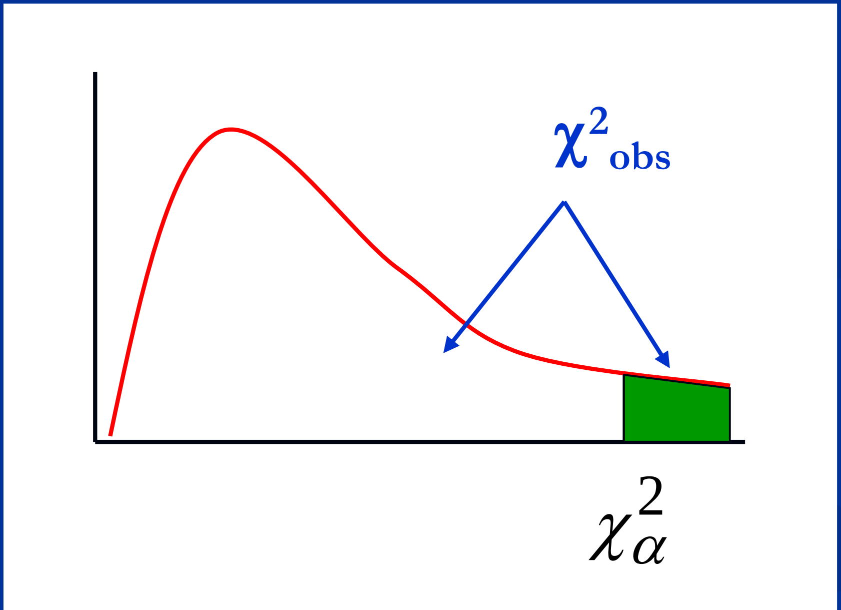
Tests d'égalité

3. Comparaison de deux proportions observées

- **test du χ^2**

$$S = \sum_{i=1}^{4} \frac{(O_i - T_i)^2}{T_i} \qquad S \sim \chi^2 \text{ à 1 ddl}$$

$$S_{obs} = \chi^2_{obs} = \sum_{i=1}^{4} \frac{(o_i - t_i)^2}{t_i}$$



Les T_i doivent être > 5

$P(\chi^2 \geq \chi^2_\alpha$ à 1 ddl$) = \alpha$

Tests d'égalité

3. Comparaison de deux proportions observées

- test du χ^2

Règle de décision:

On compare χ^2_{obs} à χ^2_{α} à 1 ddl lu dans la table du χ^2

- Si $\chi^2_{obs} < \chi^2_{\alpha}$, on accepte H_0 et on considère $P_1 = P_2$

- Si $\chi^2_{obs} > \chi^2_{\alpha}$, on rejette H_0 et on considère $P_1 \neq P_2$

Tests d'égalité

4. Extension à la comparaison de q distributions observées

On considère q échantillons indépendants avec p modalités

Ces échantillons sont extraits de q populations

On désire savoir si les q distributions sont identiques, c'est-à-dire

si les p proportions sont égales dans les q populations

H_0: les q distributions sont égales

Tests d'égalité

4. Extension à la comparaison de q distributions observées

Effectifs observés

	mod1	….	modi	….	modp	Total
E1	n_{11}	….	n_{i1}	….	n_{p1}	$n_{.1}$
:	:		:		:	:
Ej	n_{1j}	….	n_{ij}	….	n_{pj}	$n_{.j}$
:	:		:		:	:
Eq	n_{1q}	….	n_{iq}	….	n_{pq}	$n_{.q}$
Total	$n_{1.}$	…	$n_{i.}$		$n_{p.}$	n

Tests d'égalité

4. Extension à la comparaison de q distributions observées

Effectifs théoriques

$$t_{ij} = \frac{n_{i.}\, n_{.j}}{n}$$

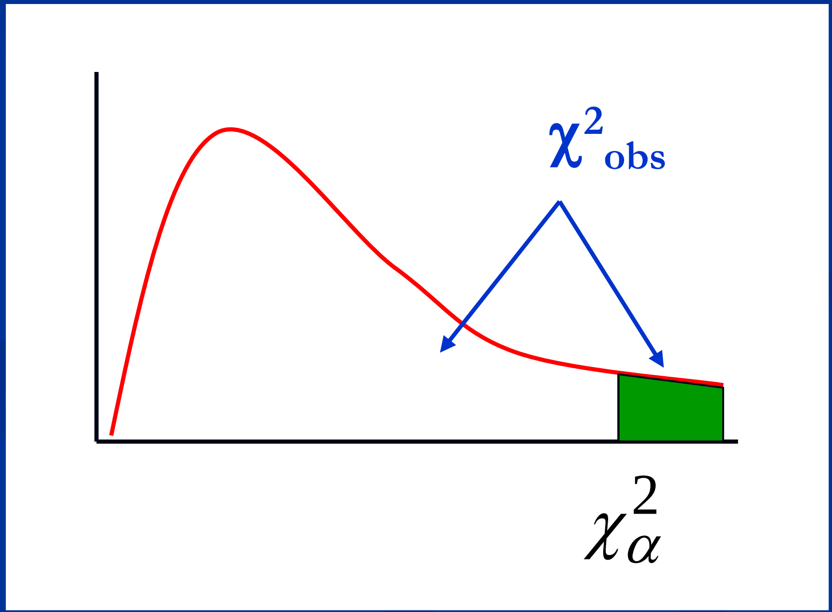
	A_1	....	Ai	....	Ap	Total
B_1	t_{11}	....	t_{i1}	....	t_{p1}	$n_{.1}$
:	:		:		:	:
Bj	t_{1j}	....	t_{ij}	....	t_{pj}	$n_{.j}$
:	:		:		:	:
Bq	t_{1q}	....	t_{iq}	....	t_{pq}	$n_{.q}$
Total	$n_{1.}$	...	$n_{i.}$		$n_{p.}$	n

Tests d'égalité

4. Extension à la comparaison de q distributions observées

$$S = \sum_{i=1}^{p'} \sum_{j=1}^{q'} \frac{\left(N_{ij} - T_{ij}\right)^2}{T_{ij}} \quad S \sim \chi^2 \text{ à } (p'-1)(q'-1) \text{ ddl}$$

$$S_{obs} = \chi^2_{obs} = \sum_{i=1}^{p'} \sum_{j=1}^{q'} \frac{\left(n_{ij} - t_{ij}\right)^2}{t_{ij}}$$



q' et p' = nb de lignes et colonnes après regroupement car les T_{ij} doivent être > 5

$P[\chi^2 \geq \chi^2_{\alpha}$ à (p'-1)(q'-1) ddl$] = \alpha$

Tests d'égalité

4. Extension à la comparaison de q distributions observées

Règle de décision:

On compare χ^2_{obs} à χ^2_{α} à (p'-1)(q'-1) ddl lu dans la table du χ^2

- **Si $\chi^2_{obs} < \chi^2_{\alpha}$, on accepte H_0, les q distributions sont égales**
- **Si $\chi^2_{obs} > \chi^2_{\alpha}$, on rejette H_0, les q distributions sont différentes**

Tests d'ajustement

1. Ajustement à une loi binomiale

X: nombre de filles dans fratries de 5 enfants

Question: $X \sim B\ (5;\ 0{,}5)$? Hypothèse H_0: $X \sim B\ (5;\ 0{,}5)$

x_i	0	1	2	3	4	5	T
n_i	18	56	110	88	40	8	320(N)
P_i							
t_i							

Tests d'ajustement

1. Ajustement à une loi binomiale

X: nombre de filles dans fratries de 5 enfants

$$P_i = P(X = x_i) = C_n^{xi} p^{xi} q^{n-xi} \text{ avec } n = 5 \text{ et } p = q = 0,5$$

$$\text{effectifs théoriques} : t_i = N.P_i \text{ avec } N = 320 \text{ familles}$$

Tests d'ajustement

1. Ajustement à une loi binomiale

X: nombre de filles dans fratries de 5 enfants $X \sim B(5; 0,5)$

x_i	0	1	2	3	4	5(k)	T
n_i	18	56	110	88	40	8	320(N)
P_i	0.03125	0.15625	0.3125	0.3125	0.15625	0.03125	1
t_i	10	50	100	100	50	10	320

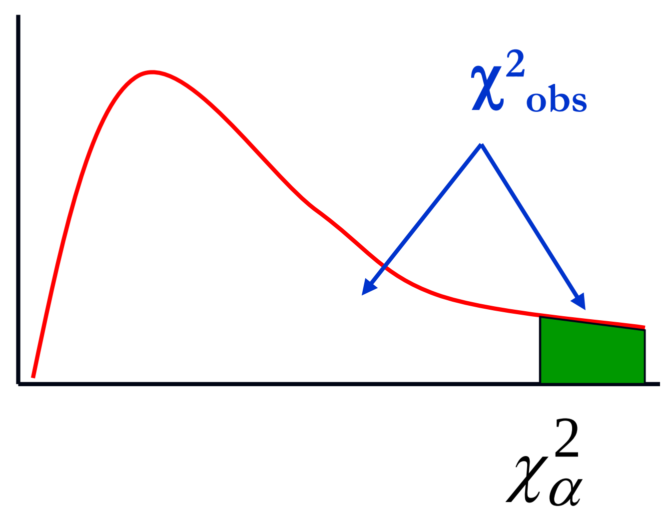
Tests d'ajustement

1. Ajustement à une loi binomiale

$$S = \sum_{i=1}^{k'} \frac{(N_i - T_i)^2}{T_i}$$

$$S \sim \chi^2 \text{ à } (k'\text{-}1\text{-}\text{nb paramètres estimés})\,\text{ddl}$$

$$S_{obs} = \chi^2_{obs} = \sum_{i=1}^{k'} \frac{(n_i - t_i)^2}{t_i} = 11{,}96$$



Les T_i doivent être > 5

$P[\chi^2 \geq \chi^2_\alpha$ à (k'-1-nbpar est) ddl$] = \alpha$

k' = nb modalités après regroupement

Tests d'ajustement

1. Ajustement à une loi binomiale

Règle de décision:

On compare χ^2_{obs} à χ^2_{α} à (k'-1-nb par. est.) ddl lu dans la table du χ^2

- **Si $\chi^2_{obs} < \chi^2_{\alpha}$, on accepte H_0, X suit une loi Binomiale B(n; p)**
- **Si $\chi^2_{obs} > \chi^2_{\alpha}$, on rejette H_0, X ne suit pas une loi binomiale**

Tests d'ajustement

2. Ajustement à une loi de Poisson

X: nombre de merles à plastron capturés par jour

Question: X ~ P ?   Hypothèse H_0: X~P (λ)

x_i	0	1	2	3	4	≥5	T
n_i	56	22	9	1	0	1	89(N)
P_i							
t_i							

Tests d'ajustement

2. Ajustement à une loi de Poisson

X: nombre de merles à plastron capturés par jour

λ est inconnu et est estimé par : $\hat{\lambda} = \bar{x}$

$$P_i = P(X = x_i) = e^{-\lambda} \frac{\lambda^{xi}}{x_i!} \text{ avec } \lambda = \bar{x} = 0{,}539$$

$$\text{effectifs théoriques} : t_i = N.P_i \text{ avec } N = 89 \text{ jours}$$

Tests d'ajustement

2. Ajustement à une loi de Poisson

X: nombre de merles à plastron capturés par jour $X \sim P(0,539)$

x_i	0	1	≥ 2	T
n_i	56	22	11	89(N)
P_i	0,5833	0,3144	0,1023	1
t_i	51,91	27,98	9,11	89

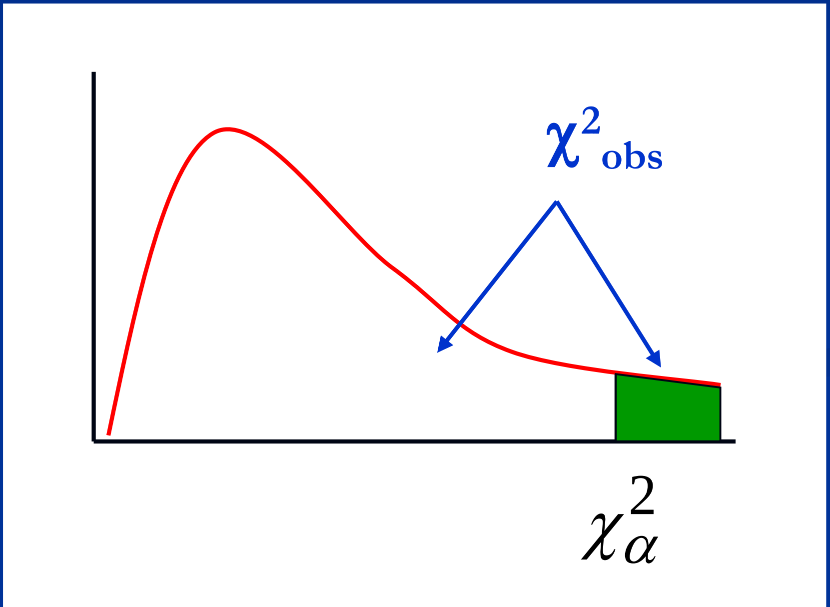
Tests d'ajustement

2. Ajustement à une loi de Poisson

$$S = \sum_{i=1}^{k'} \frac{(N_i - T_i)^2}{T_i}$$

$$S \sim \chi^2 \text{ à } (k'\text{-}1\text{ - nb paramètres estimés}) \text{ ddl}$$

$$S_{obs} = \chi^2_{obs} = \sum_{i=1}^{k'} \frac{(n_i - t_i)^2}{t_i} = 1{,}99$$



Les T_i doivent être > 5

$P[\chi^2 \geq \chi^2_\alpha$ à (k'-1-nbpar est) ddl$] = \alpha$

k' = nb modalités après regroupement. Ici k' = 3

Tests d'ajustement

2. Ajustement à une loi de Poisson

Règle de décision:

On compare χ^2_{obs} à χ^2_{α} à (k'-1-nb par. est.) ddl lu dans la table du χ^2

- Si $\chi^2_{obs} < \chi^2_{\alpha}$, on accepte H_0, X suit une loi de Poisson $P(\lambda)$
- Si $\chi^2_{obs} > \chi^2_{\alpha}$, on rejette H_0, X ne suit pas une loi de Poisson

Tests d'ajustement

3. Ajustement à une loi de normale

X: longueur de l'aile en mm

Question: X ~ N ? Hypothèse H_0: X~N (μ;σ)

x_i	]-∞;150]	]150;155]	]155;160]	]160;165]	]165;+∞[	T
u_i						
P_i						
t_i						
n_i	2	9	17	16	6	50

Tests d'ajustement

3. Ajustement à une loi de normale

X: longueur de l'aile en mm

μ et σ^2 sont inconnus et sont estimés par : $\hat{\mu} = \bar{x}$ et $\hat{\sigma}^2 = \frac{n}{n-1} s^2$

$$P_i = P(x_i < X < x_{i+1}) = P(u_i < U < u_{i+1})$$

$$\text{avec } U = \frac{X - \mu}{\sigma} \quad u_i = \frac{x_i - \mu}{\sigma} \quad u_{i+1} = \frac{x_{i+1} - \mu}{\sigma} \quad U \sim N(0;1)$$

$$\text{effectifs théoriques} : t_i = n.P_i \quad \text{avec } n = 50 \text{ mesures}$$

Tests d'ajustement

3. Ajustement à une loi de normale

x_i	]-∞;155]	]155;160]	]160;165]	]165;+∞[	T
u_i	]-∞;-0,7]	]-0,7;0,13]	]0,13;0,96]	]0,96;+∞]	
P_i	0,242	0,3097	0,2798	0,1685	1
t_i	12,0	15,48	13,99	8,43	50
n_i	11	17	16	6	50

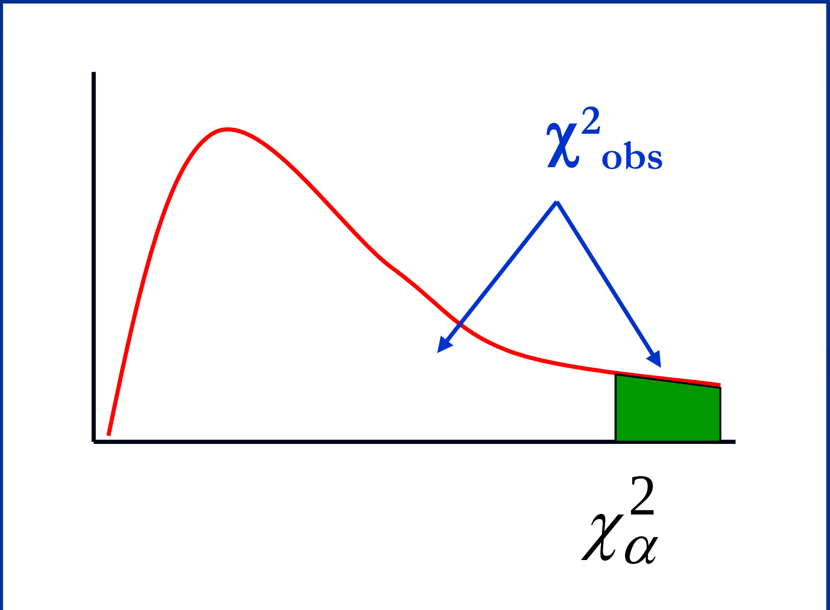
Tests d'ajustement

3. Ajustement à une loi de normale

$$S = \sum_{i=1}^{k'} \frac{(N_i - T_i)^2}{T_i}$$

$$S \sim \chi^2 \text{ à (k'-1 - nb paramètres estimés) ddl}$$

$$S_{obs} = \chi^2_{obs} = \sum_{i=1}^{k'} \frac{(n_i - t_i)^2}{t_i} = 1{,}24$$



Les T_i doivent être > 5

$P[\chi^2 \geq \chi^2_\alpha$ à (k'-1-nbpar est) ddl$] = \alpha$

k' = nb classes après regroupement. Ici k' = 4

Tests d'ajustement

3. Ajustement à une loi de normale

Règle de décision:

On compare χ^2_{obs} à χ^2_{α} à (k'-1-nb par. est.) ddl lu dans la table du χ^2

• Si $\chi^2_{obs} < \chi^2_{\alpha}$, on accepte H_0, X suit une loi Normale $N(\mu;\sigma)$

• Si $\chi^2_{obs} > \chi^2_{\alpha}$, on rejette H_0, X ne suit pas une loi Normale

Tests d'indépendance

1. Indépendance de deux caractères

On extrait d'une population un échantillon aléatoire simple sur lequel

on mesure 2 caractères: A avec p modalités et B avec q modalités

Question: A et B indépendants ? **H_0: A et B indépendants**

Tests d'indépendance

1. Indépendance de deux caractères

Effectifs observés

	A_1	….	Ai	….	Ap	Total
B_1	n_{11}	….	n_{i1}	….	n_{p1}	$n_{.1}$
:	:		:		:	:
Bj	n_{1j}	….	n_{ij}	….	n_{pj}	$n_{.j}$
:	:		:		:	:
Bq	n_{1q}	….	n_{iq}	….	n_{pq}	$n_{.q}$
Total	$n_{1.}$	…	$n_{i.}$		$n_{p.}$	n

Tests d'indépendance

1. Indépendance de deux caractères

$$\text{Sous } H_0 : \quad P(A_i \cap B_j) = P(A_i).P(B_j)$$

$$P(A_i) = \frac{n_{i.}}{n} \text{ et } P(B_j) = \frac{n_{.j}}{n} \Leftrightarrow P(A_i \cap B_j) = \frac{n_{i.}n_{.J}}{n^2}$$

$$t_{ij} \text{ est donc} : n\,P(A_i \cap B_j) = n\,\frac{n_{i.}n_{.J}}{n^2} = \frac{n_{i.}n_{.J}}{n}$$

Tests d'indépendance

1. Indépendance de deux caractères

Effectifs théoriques

$$t_{ij} = \frac{n_{i.}\, n_{.j}}{n}$$

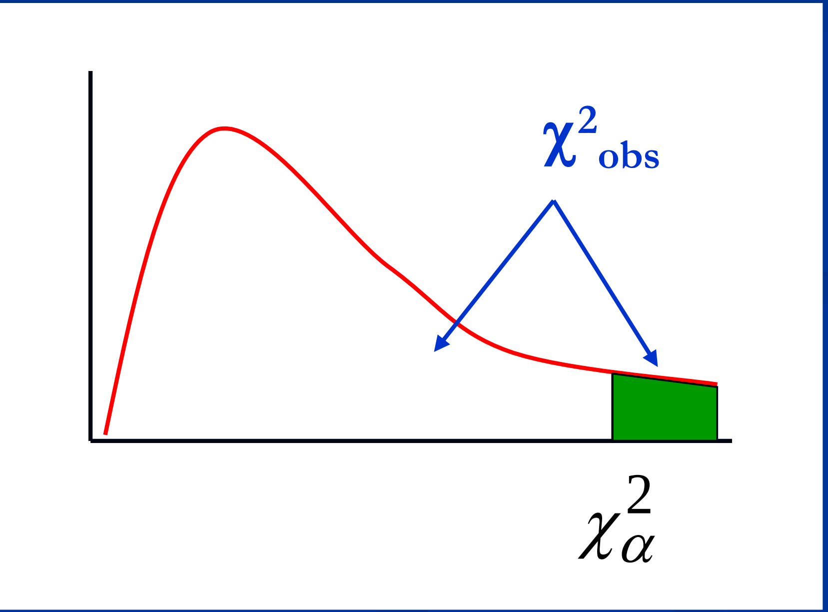
	A_1	....	Ai	....	Ap	Total
B_1	t_{11}	....	t_{i1}	....	t_{p1}	$n_{.1}$
:	:		:		:	:
Bj	t_{1j}	....	t_{ij}	....	t_{pj}	$n_{.j}$
:	:		:		:	:
Bq	t_{1q}	....	t_{iq}	....	t_{pq}	$n_{.q}$
Total	$n_{1.}$	...	$n_{i.}$		$n_{p.}$	n

Tests d'indépendance

1. Indépendance de deux caractères

$$S = \sum_{i=1}^{p'} \sum_{j=1}^{q'} \frac{(N_{ij} - T_{ij})^2}{T_{ij}} \quad S \sim \chi^2 \text{ à } (p'-1)(q'-1) \text{ ddl}$$

$$S_{obs} = \chi^2_{obs} = \sum_{i=1}^{p'} \sum_{j=1}^{q'} \frac{(n_{ij} - t_{ij})^2}{t_{ij}}$$



q' et p' = nb de lignes et colonnes après regroupement car les T_{ij} doivent être > 5

$$P[\chi^2 \geq \chi^2_\alpha \text{ à } (p'-1)(q'-1) \text{ ddl}] = \alpha$$

Tests d'indépendance

1. Indépendance de deux caractères

Règle de décision:

On compare χ^2_{obs} à χ^2_{α} à (p'-1)(q'-1) ddl lu dans la table du χ^2

• Si $\chi^2_{obs} < \chi^2_{\alpha}$, on accepte H_0, les caractères sont indépendants

• Si $\chi^2_{obs} > \chi^2_{\alpha}$, on rejette H_0, caractères pas indépendants