Chapitre III

Test du rapport de vraisemblance pour des lois classiques

Plan

Test unilatéral à gauche

On considère $X_1, \ldots, X_n$ i.i.d. de loi $\mathcal{B}(\theta)$.
On souhaite tester

$H_0 : \theta = \theta_0$ contre $H_1 : \theta = \theta_1$ avec $\theta_1 < \theta_0$ fixé,
ou bien
$H_0 : \theta \geq \theta_0$ contre $H_1 : \theta < \theta_0$

Statistique de test

$S_n = \sum_{i=1}^n X_i$ (nombre de succès) ou bien $\overline{X}_n = \frac{S_n}{n}$ (fréquence des succès).

Région de rejet

$\mathcal{R} = \{S_n \leq k_\alpha\} = \{\overline{X}_n \leq c_\alpha\}$ avec $\mathbb{P}_{\theta_0}(\mathcal{R}) \leq \alpha$

Exemple

Une entreprise de cours particuliers prétend que les étudiants qu'elle suit ont un taux de réussite au BAC de plus de 80%. Une enquête auprès d'un échantillon aléatoire de 100 élèves montre que 75% d'entre eux ont eu le BAC. Soit

$$X_i = \begin{cases} 1 \text{ si le } i\text{-ème étudiant réussit son bac} \\ 0 \text{ sinon.} \end{cases}$$

Si on doute d'un tel taux de succès on teste :

$$\begin{cases} H_0 : p \geq 0.8 \\ H_1 : p < 0.8. \end{cases}$$

On a $X_i \rightsquigarrow \mathcal{B}(p)$ et la statistique de test

$$S_{100} = \sum_{i=1}^{100} X_i \rightsquigarrow \mathcal{B}(100, p)$$

Test unilatéral à droite

On considère $X_1, \ldots, X_n$ i.i.d. de loi $\mathcal{B}(\theta)$.
On souhaite tester

$H_0 : \theta = \theta_0$ contre $H_1 : \theta = \theta_1$ avec $\theta_1 > \theta_0$ fixé,
ou bien
$H_0 : \theta \leq \theta_0$ contre $H_1 : \theta > \theta_0$

Statistique de test

$S_n = \sum_{i=1}^n X_i$ (nombre de succès) ou bien $\overline{X}_n = \frac{S_n}{n}$ (fréquence empirique des succès).

Région de rejet

$\mathcal{R} = \{S_n \geq k_\alpha\} = \{\overline{X}_n \geq c_\alpha\}$ avec $\mathbb{P}_{\theta_0}(\mathcal{R}) \leq \alpha$

Test bilatéral : $H_0 : \theta = \theta_0$ contre $H_1 : \theta \neq \theta_0$ avec θ_0 fixé

On considère $X_1, \ldots, X_n$ i.i.d. de loi $\mathcal{B}(\theta)$.
On souhaite tester

$H_0 : \theta = \theta_0$ contre $H_1 : \theta \neq \theta_0$ avec θ_0 fixé

Statistique de test

$S_n = \sum_{i=1}^n X_i$ (nombre de succès) ou bien $\overline{X}_n = \frac{S_n}{n}$ (fréquence empirique des succès).

Région de rejet

$\mathcal{R} = \{\overline{X}_n \leq c_\alpha$ ou $\overline{X}_n \geq c'_\alpha\}$ avec $\mathbb{P}_{\theta_0}(\overline{X}_n \leq c_\alpha) \leq \frac{\alpha}{2}$ et $\mathbb{P}_{\theta_0}(\overline{X}_n \geq c'_\alpha) \geq \frac{\alpha}{2}$

Résumé : Tests pour une proportion.

Soient $X_1, \dots, X_n$ i.i.d. $\mathcal{B}(\theta)$ avec θ_0 connu. On a

$$S_n = \sum_{i=1}^{n} X_i \rightsquigarrow \mathcal{B}(n, \theta_0) \text{ sous } \mathbb{P}_{\theta_0}$$

$\mathbf{H_0}$	$\mathbf{H_1}$	**Région de rejet**
$\theta \leq \theta_0$	$\theta > \theta_0$	$\{S_n \geq k_\alpha\}$
$\theta \geq \theta_0$	$\theta < \theta_0$	$\{S_n \leq k_\alpha\}$
$\theta = \theta_0$	$\theta \neq \theta_0$	$\{S_n \leq k_\alpha \text{ ou } S_n \geq k'_\alpha\}$

Résumé :Tests asymptotiques pour une proportion.

Soient $X_1, \dots, X_n$ i.i.d. $\mathcal{B}(\theta)$ avec θ_0 connu. Si $\theta = \theta_0$ et n assez grand, on a

$$Z = \sqrt{n}\frac{\overline{X}_n - \theta_0}{\sqrt{\theta_0(1-\theta_0)}} \approx \mathcal{N}(0,1)$$

H_0	**H_1**	**Région de rejet : $\mathcal{R}$**
$\theta \leq \theta_0$	$\theta > \theta_0$	$\{Z \geq z_{1-\alpha}\}$
$\theta \geq \theta_0$	$\theta < \theta_0$	$\{Z \leq -z_{1-\alpha}\}$
$\theta = \theta_0$	$\theta \neq \theta_0$	$\left\{\lvert Z\rvert \geq z_{1-\alpha/2}\right\}$

Approximation valable dès que $n \geq 30$, $n\theta_0 \geq 5$ et $n(1-\theta_0) \geq 5$

Plan

Test sur la moyenne à variance connue

Contexte : On considère $x_1, \ldots, x_n$ des réalisations de v.a. i.i.d $X_1, \ldots, X_n$ avec $X_1 \rightsquigarrow \mathcal{N}(m, \sigma^2)$ avec m inconnu et σ^2 connue. (correspond à $\theta = m$ et $\Theta = \mathbb{R}$)
Soit m_0 une valeur donnée, on souhaite tester

$$H_0 : m = m_0 \quad \text{contre} \quad H_1 : m = m_1 \quad \text{avec} \quad m_1 > m_0$$

$$V(X_1, \ldots, X_n) = e^{\frac{n}{2\sigma^2}\left(m_0^2 - m_1^2 + 2(m_1 - m_0)\overline{X}_n\right)}$$

Région de rejet

$$\mathcal{R} = \left\{ \overline{X}_n \geq m_0 + \frac{\sigma}{\sqrt{n}} z_{1-\alpha} \right\} = \left\{ \sqrt{n}\frac{\overline{X}_n - m_0}{\sigma} \geq z_{1-\alpha} \right\}$$

où $z_{1-\alpha}$ désigne le quantile d'ordre $1 - \alpha$ de $\mathcal{N}(0, 1)$

Test de $H_0 : m = m_0$ contre $H_1 : m = m_1$ avec $m_1 < m_0$ fixés.

Région de rejet

$$\mathcal{R} = \left\{ \overline{X}_n \leq m_0 - \frac{\sigma}{\sqrt{n}} z_{1-\alpha} \right\} = \left\{ \sqrt{n} \frac{\overline{X}_n - m_0}{\sigma} \leq -z_{1-\alpha} \right\}$$

où $z_{1-\alpha}$ désigne le quantile d'ordre $1-\alpha$ de $\mathcal{N}(0,1)$

Règle de décision :

- ☛ Si $\overline{X}_n(\omega) \leq m_0 - \frac{\sigma}{\sqrt{n}} z_{1-\alpha}$, on rejette H_0.
- ☛ Si $\overline{X}_n(\omega) > m_0 - \frac{\sigma}{\sqrt{n}} z_{1-\alpha}$, on ne rejette pas H_0.

Test de $H_0 : m \geq m_0$ contre $H_1 : m < m_0$ avec m_0 fixé.

Sous H_1, $\overline{X}_n$ a une plus forte probabilité d'être petit, la zone de rejet est

$$\mathcal{R} = \{\overline{X}_n \leq t_\alpha\} \quad \text{avec} \quad \sup_{m \leq m_0} \mathbb{P}_m\left(\overline{X}_n \leq t_\alpha\right) \leq \alpha$$

Comme $m \mapsto \mathbb{P}_m(\overline{X}_n \leq t_\alpha)$ est décroissante, on obtient :

$$\sup_{m \leq m_0} \mathbb{P}_m\left(\overline{X}_n \leq t_\alpha\right) = \mathbb{P}_{m_0}\left(\overline{X}_n \leq t_\alpha\right)$$

D'où la zone de rejet suivante

Région de rejet

$$\mathcal{R} = \left\{\overline{X}_n \leq m_0 - \frac{\sigma}{\sqrt{n}} z_{1-\alpha}\right\} = \left\{\sqrt{n}\frac{\overline{X}_n - m_0}{\sigma} \leq -z_{1-\alpha}\right\}$$

où $z_{1-\alpha}$ désigne le quantile d'ordre $1-\alpha$ de $\mathcal{N}(0,1)$

Test de $H_0 : m = m_0$ contre $H_1 : m \neq m_0$

La zone de rejet du test de niveau α est donné par

$$\mathcal{R} = \left\{\overline{X}_n \leq m_0 - \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}}\right\} \cup \left\{\overline{X}_n \geq m_0 + \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}}\right\} = \left\{\sqrt{n}\left|\frac{\overline{X}_n - m_0}{\sigma}\right| \geq z_{1-\alpha/2}\right\}$$

où $z_{1-\frac{\alpha}{2}}$ désigne le quantile d'ordre $1 - \frac{\alpha}{2}$ de $\mathcal{N}(0,1)$

Remarque 6

Faire ce test revient à vérifier que m_0 est dans l'intervalle de confiance de niveau $1-\alpha$,

$$IC_{1-\alpha}(m) = \left[\overline{X}_n - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}, \overline{X}_n + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$

Si $m_0 \in IC_{1-\alpha}(m)$, on rejette pas H_0 et inversement.

Résumé

Soient $X_1, \ldots, X_n$ i.i.d. $\mathcal{N}(m, \sigma^2)$ avec σ^2 connu. Si $m = m_0$, on a

$$Z = \sqrt{n}\frac{\overline{X}_n - m_0}{\sigma} \rightsquigarrow \mathcal{N}(0,1)$$

$\mathbf{H_0}$	$\mathbf{H_1}$	**Région de rejet :** $\mathcal{R}$
$m \leq m_0$	$m > m_0$	$\{Z \geq z_{1-\alpha}\}$
$m \geq m_0$	$m < m_0$	$\{Z \leq -z_{1-\alpha}\}$
$m = m_0$	$m \neq m_0$	$\left\{\lvert Z \rvert \geq z_{1-\alpha/2}\right\}$

Exemple 15

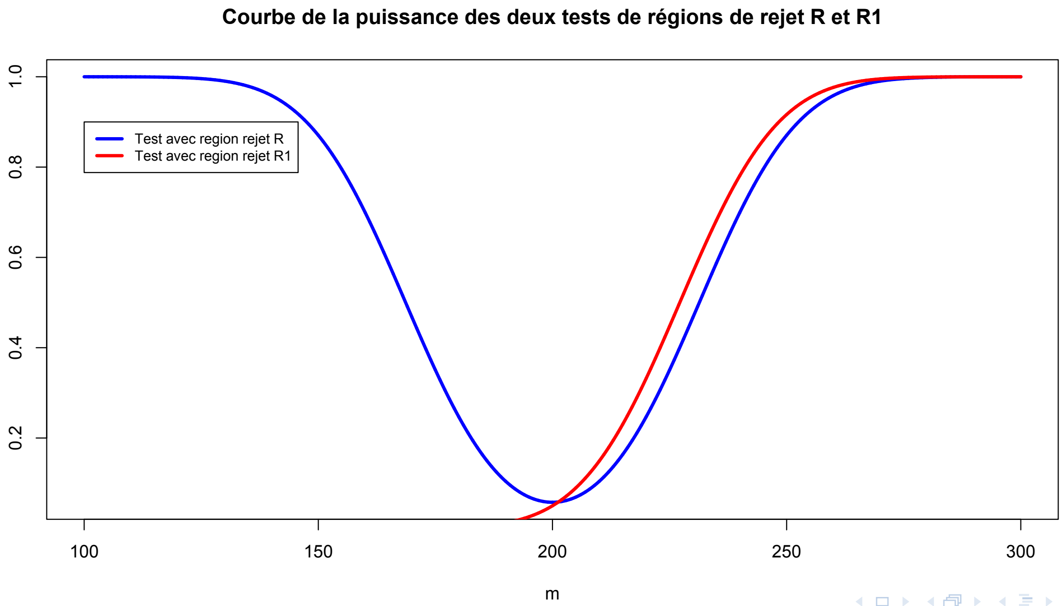
Dans une banque, les dépôts mensuels d'une certaine catégorie socio-professionnelle se répartissent autour de leur moyenne suivant des lois normales de même écart-type $\sigma = 64$ euros. On s'intéresse seulement aux dépôts dont la moyenne est $m_0 = 200$ euros. Afin de choisir entre les hypothèses $H_0 : m = 200$ contre $H_1 : m \neq 200$ on prélève un échantillon aléatoire de $n = 16$ clients pour lesquels on observe un dépôt moyen de 190 euros. Effectuer un test de niveau $\alpha = 5\%$ pour prendre une décision.

La zone de rejet du test de niveau 5% est

$$\mathcal{R} = \left\{ \left| \overline{X}_{200} - 200 \right| \geq 31.36 \right\}.$$

Ce test n'est pas UPP !

Courbe de la puissance du test de l'exemple 15



Contexte : On considère $x_1, \ldots, x_n$ des réalisations de v.a. i.i.d $X_1, \ldots, X_n$ avec $X_1 \rightsquigarrow \mathcal{N}(m, \sigma^2)$ avec m connu et σ^2 inconnue (correspond à $\theta = \sigma^2$ et $\Theta = \mathbb{R}_+$). Soient σ_0 et σ_1 des valeurs données, on souhaite tester

$$H_0 : \sigma^2 = \sigma_0^2 \quad \text{contre} \quad H_1 : \sigma^2 = \sigma_1^2 \quad \text{avec} \quad \sigma_1 < \sigma_0$$

Le rapport de vraisemblance

$$V(X_1, \ldots, X_n) = \left(\frac{\sigma_0}{\sigma_1}\right)^n \exp\left(-\frac{1}{2}\left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_0^2}\right)\sum_{i=1}^n (X_i - m)^2\right)$$

On rejette H_0 si

$$V(X_1, \ldots, X_n) \geq v_\alpha \iff V_n^2 \leq t_\alpha \text{ avec } V_n^2 = \frac{1}{n}\sum_{i=1}^n (X_i - m)^2.$$

Statistique de test

$$V_n^2 = \frac{1}{n}\sum_{i=1}^n (X_i - m)^2$$

Test de $H_0 : \sigma^2 = \sigma_0^2$ contre $H_1 : \sigma^2 = \sigma_1^2$ avec $\sigma_1 < \sigma_0$

Région de rejet

$$\mathcal{R} = \{V(X_1, \dots, X_n) \geq v_\alpha\} = \left\{ V_n^2 \leq \frac{\sigma_0^2}{n} \chi^2_{n,\alpha} \right\} = \left\{ \underbrace{\frac{nV_n^2}{\sigma_0^2}}_{K} \leq \chi^2_{n,\alpha} \right\}$$

où $\chi^2_{n,\alpha}$ désigne le quantile d'ordre α de la loi du Khi-deux à n degrés de liberté.

Test de $H_0 : \sigma^2 \geq \sigma_0^2$ contre $H_1 : \sigma^2 < \sigma_0^2$ avec σ_0^2 connu.

Statistique de test

$$V_n^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - m)^2$$

On rejette H_0 si

$$V(X_1, \ldots, X_n) \geq v_\alpha \iff V_n^2 \leq t_\alpha \quad \text{avec} \quad \sup_{\sigma^2 \geq \sigma_0^2} \mathbb{P}_{\sigma^2}\left(V_n^2 \leq t_\alpha\right) \leq \alpha$$

Région de rejet

$\mathcal{R} = \left\{ V_n^2 \leq \frac{\sigma_0^2}{n} \chi_{n,\alpha}^2 \right\} = \left\{ \frac{nV_n^2}{\sigma_0^2} \leq \chi_{n,\alpha}^2 \right\}$ où $\chi_{n,\alpha}^2$ désigne le quantile d'ordre α de $\chi^2(n)$.

Test de $H_0 : \sigma^2 = \sigma_0^2$ contre $H_1 : \sigma^2 \neq \sigma_0^2$ avec σ_0^2 connu.

TD ?

La zone de rejet du test de RV de niveau α est donné par

Zone de rejet

non α.

avec $\boxed{m}$

$$\mathcal{R} = \left\{ V_n^2 \leq \frac{\sigma_0^2}{n} \chi^2_{n,\alpha/2} \right\} \cup \left\{ V_n^2 \geq \frac{\sigma_0^2}{n} \chi^2_{n,1-\alpha/2} \right\}$$

où $\chi^2_{n,\alpha}$ désigne le quantile d'ordre α de $\chi^2(n)$.

Remarque 7

Faire ce test revient à vérifier que σ_0^2 est dans l'intervalle de confiance de niveau $1-\alpha$,

$$IC_{1-\alpha}(\sigma^2) = \left[\frac{nV_n^2}{\chi^2_{n,1-\alpha/2}}, \frac{nV_n^2}{\chi^2_{n,\alpha/2}} \right]$$

Same

$|V_n - \sigma^2|$

Si $\sigma_0^2 \in IC_{1-\alpha}(\sigma^2)$, on ne rejette pas H_0 et inversement.

Résumé

Soient $X_1, \ldots, X_n$ i.i.d. $\mathcal{N}(m, \sigma^2)$ où m est connue. L'EMV de σ^2 est donné par $V_n^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - m)^2$.

Si $\sigma^2 = \sigma_0^2$, on a $K = \frac{nV_n^2}{\sigma_0^2} \rightsquigarrow \chi^2(n)$.

$\mathbf{H_0}$	$\mathbf{H_1}$	**Région de rejet :** $\mathcal{R}$
$\sigma^2 \leq \sigma_0^2$	$\sigma^2 > \sigma_0^2$	$\{K \geq \chi^2_{n;1-\alpha}\}$
$\sigma^2 \geq \sigma_0^2$	$\sigma^2 < \sigma_0^2$	$\{K \leq \chi^2_{n;\alpha}\}$
$\sigma^2 = \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$	$\left\{K \leq \chi^2_{n,1-\alpha/2} \text{ ou } K \geq \chi^2_{n,\alpha/2}\right\}$

Exemple 16

On étudie un procédé de fabrication pour lequel la variable de qualité X est la contenance d'un flacon (en cl) normalement distribuée qui doit avoir une espérance m de 100cl. A la suite d'erreurs survenant dans le processus de fabrication, on se demande si la variabilité de la contenance est $\sigma^2 = 1$ ou $\sigma^2 = 1/4$. On prélève un échantillon de 10 flacons sur lesquels on observe $\sum_{i=1}^{10}(X_i - 100)^2 = 3$. Donner la règle de décision du test de niveau $\alpha = 5\%$.

On teste $H_0 : \sigma^2 = 1$ contre $H_1 : \sigma^2 = 1/4$.
La zone de rejet du test de niveau 5% est $\mathcal{R} = \left\{ V_{10}^2 \leq \frac{1}{10} 3.94 \right\}$

Test sur la moyenne à variance inconnue

Contexte : On considère $x_1, \ldots, x_n$ des réalisations de v.a. i.i.d $X_1, \ldots, X_n$ avec $X_1 \rightsquigarrow \mathcal{N}(m, \sigma^2)$ avec m inconnu et σ^2 inconnue (correspond à $\theta = m$ et $\Theta = \mathbb{R}$).
Soit m_0 une valeur donnée, on souhaite tester

Test unilatéral à droite

$$H_0 : m = m_0 \quad \text{contre} \quad H_1 : m > m_0.$$

Statistique de test :

$$T = \sqrt{n}\frac{\overline{X}_n - m_0}{S_n}$$

Sous H_0,

$$T \rightsquigarrow T_{n-1} \text{ où } T_{n-1} \text{ suit la loi de Student à } n-1 \text{ degrés de liberté.}$$

Sous H_1,

$$T = \sqrt{n}\frac{\overline{X}_n - m}{S_n} + \sqrt{n}\frac{m - m_0}{S_n} \text{ prend des valeurs plus grandes que sous } H_0.$$

Test de $H_0 : m = m_0$ contre $H_1 : m > m_0$.

Zone de rejet

La zone de rejet est

$$\mathcal{R} = \{T \geq t_{n-1,1-\alpha}\},$$

où $t_{n-1,1-\alpha}$ est le quantile d'ordre $1-\alpha$ de la loi de Student à $n-1$ degrés de liberté.

Règle de décision

- ☛ Si $T(\omega) \in \mathcal{R}$, on rejette H_0
- ☛ Si $T(\omega) \notin \mathcal{R}$, on ne rejette pas H_0.

Test de $H_0 : m \leq m_0$ contre $H_1 : m > m_0$

Statistique de test :

$$T = \sqrt{n}\frac{\overline{X}_n - m_0}{S_n}$$

Zone de rejet

La zone de rejet est

$$\mathcal{R} = \{T \geq t_{n-1,1-\alpha}\},$$

où $t_{n-1,1-\alpha}$ est le quantile d'ordre $1-\alpha$ de la loi de Student à $n-1$ d.d.l.

Test de $H_0 : m \geq m_0$ contre $H_1 : m < m_0$

Statistique de test :

$$T = \sqrt{n}\frac{\overline{X}_n - m_0}{S_n}$$

Zone de rejet

La zone de rejet est

$$\mathcal{R} = \{T \leq -t_{n-1,1-\alpha}\},$$

où $t_{n-1,1-\alpha}$ est le quantile d'ordre $1-\alpha$ de la loi T_{n-1} à $n-1$ d.d.l.

Test de $H_0 : m = m_0$ contre $H_1 : m \neq m_0$

Statistique de test :

$$T = \sqrt{n}\frac{\overline{X}_n - m_0}{S_n}$$

Zone de rejet

La zone de rejet est

$$\mathcal{R} = \{|T| \geq t_{n-1,1-\alpha/2}\},$$

où $t_{n-1,1-\alpha/2}$ est le quantile d'ordre $1-\alpha/2$ de la loi T_{n-1} à $n-1$ d.d.l.

Résumé

Soient $X_1, \ldots, X_n$ i.i.d. $\mathcal{N}(m, \sigma^2)$ avec σ^2 inconnue. Si $m = m_0$ avec m_0 connu, on a

$$T = \sqrt{n}\frac{\overline{X}_n - m_0}{S_n} \rightsquigarrow T_{n-1}$$

$\mathbf{H_0}$	$\mathbf{H_1}$	**Région de rejet :** $\mathcal{R}$
$m \leq m_0$	$m > m_0$	$\{T \geq t_{n-1;1-\alpha}\}$
$m \geq m_0$	$m < m_0$	$\{T \leq -t_{n-1;1-\alpha}\}$
$m = m_0$	$m \neq m_0$	$\left\{\lvert T \rvert \geq t_{n-1;1-\alpha/2}\right\}$

Exemple 17

Le temps de sommeil par nuit en période de projet pour une promotion en 4ème année d'une école d'ingénieurs est supposé suivre une loi $\mathcal{N}(m, \sigma^2)$ avec m et σ^2 inconnu. On a interrogé 30 étudiants de la promotion pour lesquels, on a obtenu un temps moyen de sommeil de 6.36h et une variance estimée de $s_n^2 = 1.85h^2$. Les enseignants souhaitent savoir si l'espérance du temps de sommeil est significativement inférieure au temps moyen de sommeil des autres individus de la population, qui est de 7 heures.

Contexte : On considère $x_1, \dots, x_n$ des réalisations de v.a. i.i.d $X_1, \dots, X_n$ avec $X_1 \rightsquigarrow \mathcal{N}(m, \sigma^2)$ avec m inconnue. Soit σ_0^2 une valeur donnée, on souhaite tester

Test unilatéral à gauche

$$H_0 : \sigma^2 = \sigma_0^2 \quad \text{contre} \quad H_1 : \sigma^2 < \sigma_0^2.$$

Statistique de test

$$V_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2 = \frac{n-1}{n} S_n^2$$

Région de rejet

$$\mathcal{R} = \left\{ V_n^2 \leq \frac{\sigma_0^2}{n} \chi^2_{n-1,\alpha} \right\} = \left\{ \frac{n V_n^2}{\sigma_0^2} \leq \chi^2_{n-1,\alpha} \right\}$$

où $\chi^2_{n,\alpha}$ désigne le quantile d'ordre α de $\chi^2(n-1)$.

Test de $H_0 : \sigma^2 = \sigma_0^2$ contre $H_1 : \sigma^2 > \sigma_0^2$ avec σ_0^2 connu.

Statistique de test

$$V_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2 = \frac{n-1}{n}S_n^2$$

Région de rejet

$$\mathcal{R} = \left\{ V_n^2 \geq \frac{\sigma_0^2}{n}\chi^2_{n-1,1-\alpha} \right\} = \left\{ \frac{nV_n^2}{\sigma_0^2} \geq \chi^2_{n-1,1-\alpha} \right\}$$

où $\chi^2_{n-1,1-\alpha}$ désigne le quantile d'ordre $1-\alpha$ de $\chi^2(n-1)$.

Exemple 18

Rappelons que 30 étudiants de 4 ème année avaient donné leur temps moyen de sommeil sur une nuit en période de projet et que nous avions obtenu $\overline{x}_n = 6.36h$ *et* $s_n^2 = 1.86h^2$*. Dans la population, l'écart-type vaut* $\sigma_0 = 1.2$ *heure. Les enseignants voudraient savoir si la variabilité est plus forte au sein de la promotion de que dans le reste de la population.*