Résumé de cours Test d'hypothèses

• Soit X une caractéristique dont la distribution dépend de paramètre 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 (à partir d'un échantillon tiré de cette population)

• Dans la plupart des situation 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 : $H_0 : \theta = \theta_0$

Définitions :

• **test d'hypothèse :** un test d'hypothèse ou test statistique 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 et hypothèse alternative :**

L'hypothèse selon laquelle on fixe à 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" (ou contre-hypothèse) et est notée H_1

Scanné avec CamScanner

c'est l'hypothèse nulle (H_0) qui est soumise au test et toute la démarche du test s'effectue en considérant cette hypothèse comme vraie

On a :

si H_0 est vraie alors H_1 est fausse.

et si H_0 est fausse alors H_1 est vraie.

· Catégories des tests d'hypothèses :

1 - test simple : un test est dit simple si on veut choisir entre deux valeurs d'un paramètre θ (θ_1 et θ_2)

On a : $\begin{cases} H_0 : \theta = \theta_1 \\ H_1 : \theta = \theta_2 \end{cases}$

2 - test multiple

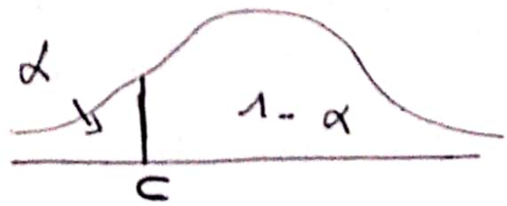
a) test unilatéral :

a-1) unilatéral à droite (test de supériorité)

$\begin{cases} H_0 : \theta = \theta_1 \\ H_1 \quad \theta > \theta_1 \end{cases}$

· a-2) unilatéral à gauche (test d'infériorité)

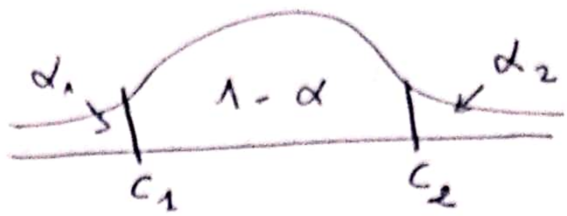
$\begin{cases} H_0 : \theta = \theta_1 \\ H_1 : \theta < \theta_1 \end{cases}$



Scanné avec CamScanner

b] test bilatéral

$$\begin{cases} H_0 : \theta = \theta_1 \\ H_1 : \theta \neq \theta_1 \end{cases}$$



Seuil de signification et erreurs de test

- α = le seuil de signification
  = risque de 1ère espèce = erreur de 1ère espèce
  = P (H_0 est fausse / H_0 est vraie) = P(rejeter H_0 / H_0 vraie)
  = P(H_1 est vraie / H_0 est vraie)
- $\beta(\theta)$ = risque ou erreur de 2ème espèce
  = P(H_1 est fausse / H_1 est vraie)
  = P(H_0 est vraie / H_0 fausse)
- Si $\alpha \nearrow$ alors $\beta \searrow$

Remarque : les seuils de signification les plus utilisés sont $\alpha = 5\%$ et $\alpha = 1\%$.

- la puissance d'un test d'hypothèse

$\varphi(\theta) = 1 - \beta(\theta)$.
  = P(H_0 est fausse / H_1 est vraie) = P(rejeter H_0 / H_1 est vraie)

plus $\beta(\theta)$ est petit plus le test est plus puissant.

Scanné avec CamScanner

· à seuil de signification, on fait correspondre sur la distribution d'echantillonnage de la statistique une région de rejet de l'hypothèse H_0 (appellée également région critique).

· α est choisi à priori
· $\beta(\theta)$ dépend de l'hypothèse H_1 et on ne peut le calculer que si on spécifie des valeurs particulières du paramètre θ dans H_1

Les test usuels

(cas Normale = $X \sim N$ormale = la population est de loi N ou bien la population est de loi quelconque mais $n \geq 30$)

I) tester une moyenne:

$X \sim N(m, \sigma^2)$

Cas ou σ^2 connu

. L'estimateur de m est $\bar{X} \sim N(m, \frac{\sigma^2}{n})$
. La statistique de test

$$\frac{\bar{X} - m}{\sigma}\sqrt{n} \sim N(0,1)$$

Cas ou σ^2 inconnu

L'estimateur de m est $\bar{X} \sim N(m, \sigma^2)$
S^2 et S'^2 estimateurs de σ^2

La statistique de test

$$\frac{\bar{X} - m}{S}\sqrt{n-1} = \frac{\bar{X} - m}{S'}\sqrt{n} \sim St_{(n-1)}$$

de loi de student a $(n-1)$ degré de liberté.

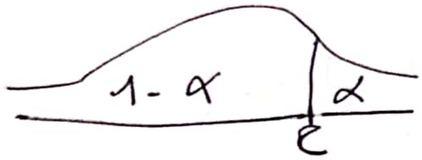
c est la valeur critique
on détermine la région critique (-)
on calcul la valeur de c

Scanné avec CamScanner

test multiple

1) test unilatéral à droite : $H_0 : m = m_0$ vs $H_1 : m > m_0$

$C = ??$ la région critique de la forme : $]C, +\infty[$

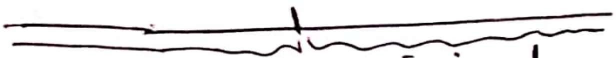


$C = ??$ la région critique de la forme $]C, +\infty[$

si $\bar{X} \in]C, +\infty[\Rightarrow H_0$ est fausse

si $\bar{X} \notin]C, +\infty[$ càd $\bar{X} \in]-\infty, C]$ alors H_0 est acceptée

La valeur critique C



Zone = région d'acceptation de H_0 — région de rejeter H_0

$$\alpha = P(\bar{X} > C / m = m_0)$$
$$= P(H_0 \text{ fausse} / H_0 \text{ est vraie})$$
$$= P(H_1 \text{ est vraie} \mid H_0 \text{ est vraie})$$
$$= P\left(\frac{\bar{X} - m_0}{\sigma}\sqrt{n} > \frac{C - m_0}{\sigma}\sqrt{n}\right)$$

Comme $\frac{\bar{X} - m_0}{\sigma}\sqrt{n} \sim N(0,1)$

$$\Rightarrow \alpha = 1 - P\left(\frac{\bar{X} - m_0}{\sigma}\sqrt{n} \leq \frac{C - m_0}{\sigma}\sqrt{n}\right)$$
$$\alpha = 1 - F_{N(0,1)}\left(\frac{C - m_0}{\sigma}\sqrt{n}\right)$$
$$\Rightarrow F_{N(0,1)}\left(\frac{C - m_0}{\sigma}\sqrt{n}\right) = 1 - \alpha$$
$$\Rightarrow \frac{C - m_0}{\sigma}\sqrt{n} = q_{1-\alpha}$$ = le quantile d'ordre $(1-\alpha)$ d'une $N(0,1)$

$$\alpha = P(\bar{X} > C / m = m_0)$$
$$= P\left(\frac{\bar{X} - m_0}{S}\sqrt{n-1} > \frac{C - m_0}{S}\sqrt{n-1}\right)$$
ou bien
$$= P\left(\frac{\bar{X} - m_0}{S'}\sqrt{n} > \frac{C - m_0}{S'}\sqrt{n}\right)$$

Comme $\frac{\bar{X} - m_0}{S'}\sqrt{n} = \frac{\bar{X} - m_0}{S}\sqrt{n-1}$ suit $St_{(n-1)}$

$$\Rightarrow \alpha = 1 - P\left(\frac{\bar{X} - m_0}{S}\sqrt{n-1} \leq \frac{C - m_0}{S}\sqrt{n-1}\right)$$
$$\alpha = 1 - F_{St(n-1)}\left(\frac{C - m_0}{S}\sqrt{n-1}\right)$$
$$\Rightarrow F_{St(n-1)}\left(\frac{C - m_0}{S}\sqrt{n-1}\right) = 1 - \alpha$$
$$\Rightarrow \frac{C - m_0}{S}\sqrt{n-1} = q_{1-\alpha} = St_{(n-1),\alpha}$$
= le quantile d'ordre $(1-\alpha)$ d'une loi de student à $(n-1)$ d.d.l.

Scanné avec CamScanner

le quantile $q_{1-\alpha}$ lue sur la table de la $N(0,1)$

$$\Rightarrow \boxed{C = q_{1-\alpha}\frac{\sigma}{\sqrt{n}} + m_0}$$

$$\Rightarrow \boxed{C = m_0 + \frac{\sigma}{\sqrt{n}}\, q_{1-\alpha}}$$

le quantile $q_{1-\alpha} = t_{\alpha,(n-1)}$ lue sur la table de la student pour α et $(n-1)$ d.d.l

$$\Rightarrow \boxed{C = t_{\alpha,(n-1)} \frac{S}{\sqrt{n-1}} + m_0}$$

on peut utiliser S' au lieu S on trouve

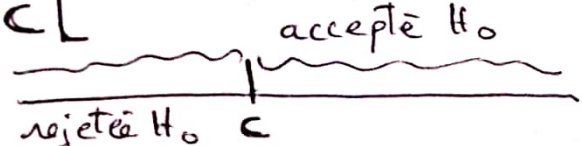
$$\boxed{C = t_{\alpha,(n-1)} \frac{S'}{\sqrt{n}} + m_0}$$

2] test unilatérale à gauche

$H_0: m = m_0$ vs $H_1: m < m_0$

c = valeur critique

la région critique = la région de rejet de H_0 est de la forme : $]-\infty, c[$



$$\alpha = P(\bar{X} < c \,/\, m = m_0)$$

$$= P\left(\frac{\bar{X} - m_0}{\sigma}\sqrt{n} < \frac{c - m_0}{\sigma}\sqrt{n}\right)$$

$$\alpha = F_{N(0,1)}\left(\frac{c - m_0}{\sigma}\sqrt{n}\right)$$

$$\Rightarrow \frac{c - m_0}{\sigma}\sqrt{n} = q_\alpha = -q_{1-\alpha}$$

q_α = le quantile d'ordre α d'une $N(0,1)$ lue sur la table

$$\boxed{C = -q_{1-\alpha}\frac{\sigma}{\sqrt{n}} + m_0}$$

$$\boxed{C = m_0 - \frac{\sigma}{\sqrt{n}}\, q_{1-\alpha}}$$

$$\alpha = P(\bar{X} < c \,/\, m = m_0)$$

$$= P\left(\frac{\bar{X} - m_0}{S}\sqrt{n-1} < \frac{c - m_0}{S}\sqrt{n-1}\right)$$

$$= P\left(\frac{\bar{X} - m_0}{S'}\sqrt{n} < \frac{c - m_0}{S'}\sqrt{n}\right)$$

$$\Rightarrow \alpha = F_{St_{(n-1)}}\left(\frac{c - m_0}{S}\sqrt{n-1}\right)$$

donc

$$\frac{c - m_0}{S}\sqrt{n-1} = \frac{c - m_0}{S'}\sqrt{n}$$

$= t_{1-\alpha,(n-1)} = -t_{\alpha,(n-1)}$ = le quantile d'ordre α d'une student à $(n-1)$ d.d.l lue sur la table

$$\Rightarrow \boxed{C = m_0 - \frac{S}{\sqrt{n-1}}\, t_{\alpha,(n-1)}}$$

Scanné avec CamScanner

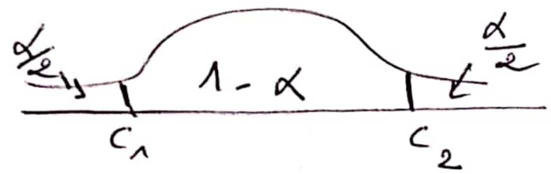
ou bien

$$\boxed{c = m_0 - \frac{s'}{\sqrt{n}} t_{\alpha,(n-1)}}$$

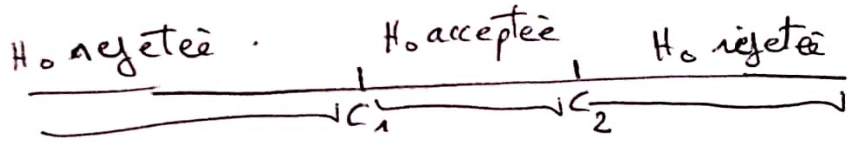
si la valeur de c est donnée on peut calculer α = le seuil de signification

3] test bilatérale

$H_0 : m = m_0$ vs $H_1 : m \neq m_0$



$]-\infty, c_1[\cup]c_2, +\infty[$ = la région critique (de rejet de H_0)



$$\alpha = P(\bar{X} < c_1 \mid m = m_0) + P(\bar{X} > c_2 \mid m = m_0)$$

$$= P\left(\frac{\bar{X} - m_0}{\sigma}\sqrt{n} < \frac{c_1 - m_0}{\sigma}\sqrt{n}\right)$$

$$+ P\left(\frac{\bar{X} - m_0}{\sigma}\sqrt{n} > \frac{c_2 - m_0}{\sigma}\sqrt{n}\right)$$

$$\Rightarrow \alpha = F_{N(0,1)}\left(\frac{c_1 - m_0}{\sigma}\sqrt{n}\right) + 1 - F_{N(0,1)}\left(\frac{c_2 - m_0}{\sigma}\sqrt{n}\right)$$

$$\frac{c_1 - m_0}{\sigma}\sqrt{n} = -q_{1-\frac{\alpha}{2}}$$

$$\frac{c_2 - m_0}{\sigma}\sqrt{n} = q_{1-\frac{\alpha}{2}}$$

lues sur la table de la $N(0,1)$

$$\Rightarrow \boxed{\begin{array}{l} c_1 = -\frac{\sigma}{\sqrt{n}} q_{1-\frac{\alpha}{2}} + m_0 \\ c_2 = \frac{\sigma}{\sqrt{n}} q_{1-\frac{\alpha}{2}} + m_0 \end{array}}$$

$$c_1 = m_0 - \frac{\sigma}{\sqrt{n}} q_{1-\frac{\alpha}{2}}$$

c_2 = " + "

$$\alpha = P(\bar{X} < c_1 \mid m = m_0) + P(\bar{X} > c_2 \mid m = m_0)$$

on utilise soit s ou s'

$$= P\left(\frac{\bar{X} - m_0}{s}\sqrt{n-1} < \frac{c_1 - m_0}{s}\sqrt{n-1}\right)$$

$$+ P\left(\frac{\bar{X} - m_0}{s}\sqrt{n-1} > \frac{c_2 - m_0}{s}\sqrt{n-1}\right)$$

$$\Rightarrow \alpha = F_{St_{(n-1)}}\left(\frac{c_1 - m_0}{s}\sqrt{n-1}\right) + 1 - F_{St_{(n-1)}}\left(\frac{c_2 - m_0}{s}\sqrt{n-1}\right)$$

$$\frac{c_1 - m_0}{s}\sqrt{n-1} = -t_{\frac{\alpha}{2},(n-1)}$$

$$\frac{c_2 - m_0}{s}\sqrt{n-1} = t_{\frac{\alpha}{2}, n-1}$$

lue sur la table de la students

$$\boxed{\begin{array}{l} c_1 = -\frac{s}{\sqrt{n-1}} t_{\frac{\alpha}{2},(n-1)} + m_0 \\ c_2 = \frac{s}{\sqrt{n-1}} t_{\frac{\alpha}{2},(n-1)} + m_0 \end{array}}$$

Scanné avec CamScanner

II] tester une variance

1er cas la moyenne m = connue

la statistique de test est

$$\frac{nT}{\sigma^2} \rightsquigarrow \chi^2_{(n)}$$

2ème cas : la moyenne m = inconnue

la statistique de test est

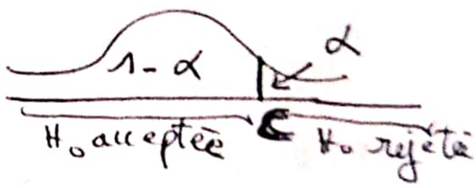
$$\frac{n s^2}{\sigma^2} = \frac{(n-1)s'^2}{\sigma^2} \rightsquigarrow \chi^2_{(n-1)}$$

car s et s' estimateur de σ^2

test multiple

1] $H_0 : \sigma^2 = \sigma_0^2$ VS H_1 $\sigma^2 > \sigma_0^2$

$]c, +\infty[$ la région critique



$$\alpha = P(T > c \mid \sigma^2 = \sigma_0^2)$$

$$= P\left(\frac{nT}{\sigma_0^2} > \frac{nc}{\sigma_0^2}\right)$$

$$\Rightarrow \alpha = 1 - P\left(\frac{nT}{\sigma_0^2} < \frac{nc}{\sigma_0^2}\right)$$

$$1-\alpha = F_{\chi^2_{(n)}}\left(\frac{nc}{\sigma_0^2}\right)$$

$\Rightarrow \frac{nc}{\sigma_0^2}$ = le quantile d'ordre $1-\alpha$ d'une Khi-deux à n d.d.l lue sur la table de Khi-deux pour $\underline{\alpha}$ et $\underline{n}$ d.d.l

$$\alpha = P(s^2 > c \mid \sigma^2 = \sigma_0^2)$$

$$\alpha = P\left(\frac{n s^2}{\sigma_0^2} > \frac{nc}{\sigma_0^2}\right)$$

$$\Rightarrow 1-\alpha = P\left(\frac{n s^2}{\sigma_0^2} < \frac{nc}{\sigma_0^2}\right) = F_{\chi^2_{(n-1)}}\left(\frac{cn}{\sigma_0^2}\right)$$

$\Rightarrow \frac{cn}{\sigma_0^2} = q_{1-\alpha}$ = quantile d'ordre $1-\alpha$ d'une χ^2 à $(n-1)$ d.d.l

lue sur la table pour α et $(n-1)$ d.d.l de la loi de Khi-deux.

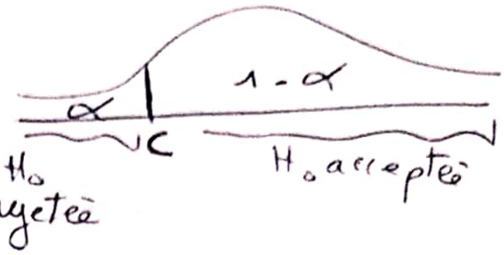
Scanné avec CamScanner

$\Rightarrow \boxed{C = \frac{\sigma_0^2}{n}\chi^2_{\alpha,(n)}}$

$\boxed{C = \frac{\sigma_0^2}{n}\chi^2_{\alpha,(n-1)}}$

2] $H_0: \sigma^2 = \sigma_0^2$ vs $\sigma^2 < \sigma_0^2$

tst unilatéraleà gauche.



la région critique de la forme:

$]-\infty, C[$

$\alpha = P(T < c \,/\, \sigma^2 = \sigma_0^2)$

$= P\left(\frac{nT}{\sigma_0^2} < \frac{cn}{\sigma_0^2}\right)$

$\Rightarrow \alpha = F_{\chi^2_{(n)}}\left(\frac{nc}{\sigma_0^2}\right)$

$\Rightarrow \frac{nc}{\sigma_0^2} = q_\alpha = \chi^2_{(1-\alpha),n}$

$\Rightarrow \boxed{C = \frac{\sigma_0^2}{n}\chi^2_{(1-\alpha),n}}$

$\alpha = P(S^2 < c \,/\, \sigma^2 = \sigma_0^2)$

$= P\left(\frac{nS^2}{\sigma_0^2} < \frac{nc}{\sigma_0^2}\right)$

on peut utiliser S'^2 au lieu S^2

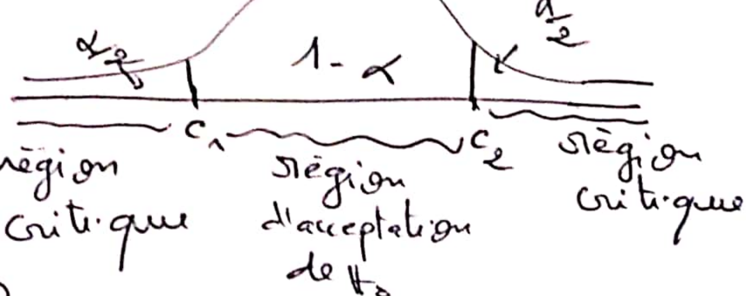
$\Rightarrow \alpha = F_{\chi^2_{(n-1)}}\left(\frac{nc}{\sigma_0^2}\right)$

$\Rightarrow \frac{nc}{\sigma_0^2} = q_\alpha = \chi^2_{(1-\alpha),(n-1)}$

$\Rightarrow \boxed{C = \frac{\sigma_0^2}{n}\chi^2_{(1-\alpha),(n-1)}}$

3] tst bilatérale

$H_0: \sigma^2 = \sigma_0^2$ vs $H_1: \sigma^2 \neq \sigma_0^2$



La région critique de la forme

$]-\infty, c_1[\cup]c_2, +\infty[$

Scanné avec CamScanner

$$\alpha = P(T < c_1 \mid H_0 \text{ vraie}) + P(T > c_2 \mid H_0 \text{ vraie})$$

$$= P(T < c_1 \mid \sigma^2 = \sigma_0^2) + P(T > c_2 \mid \sigma^2 = \sigma_0^2)$$

$$= \underbrace{P\left(\frac{nT}{\sigma_0^2} < \frac{nc_1}{\sigma_0^2}\right)}_{"\frac{\alpha}{2}} + \underbrace{P\left(\frac{nT}{\sigma_0^2} > \frac{nc_2}{\sigma_0^2}\right)}_{"\frac{\alpha}{2}}$$

$$\alpha = \underbrace{F_{\chi^2_{(n)}}\left(\frac{nc_1}{\sigma_0^2}\right)}_{"\frac{\alpha}{2}} + \underbrace{1 - F_{\chi^2_{(n)}}\left(\frac{nc_2}{\sigma_0^2}\right)}_{\frac{\alpha}{2}}$$

$$\Rightarrow \frac{nc_1}{\sigma_0^2} = q_{\frac{\alpha}{2}} = \chi^2_{(1-\frac{\alpha}{2}), n}$$

et $\frac{nc_2}{\sigma_0^2} = q_{1-\frac{\alpha}{2}} = \chi^2_{\frac{\alpha}{2}, n}$

$$\Rightarrow \boxed{\begin{aligned} c_1 &= \frac{\sigma_0^2}{n} \chi^2_{(1-\frac{\alpha}{2}), n} \\ c_2 &= \frac{\sigma_0^2}{n} \chi^2_{\frac{\alpha}{2}, n} \end{aligned}}$$

$$\alpha = P(S^2 < c_1 \mid H_0 \text{ vraie}) + P(S^2 > c_2 \mid H_0 \text{ vraie})$$

$$= P(S^2 < c_1 \mid \sigma^2 = \sigma_0^2) + P(S^2 > c_2 \mid \sigma^2 = \sigma_0^2)$$

$$= P\left(\frac{nS^2}{\sigma_0^2} < \frac{nc_1}{\sigma_0^2}\right) + P\left(\frac{nS^2}{\sigma_0^2} > \frac{nc_2}{\sigma_0^2}\right)$$

$$\alpha = F_{\chi^2_{(n-1)}}\left(\frac{nc_1}{\sigma_0^2}\right) + 1 - F_{\chi^2_{(n-1)}}\left(\frac{nc_2}{\sigma_0^2}\right)$$

$$\Rightarrow \frac{nc_1}{\sigma_0^2} = q_{\frac{\alpha}{2}} \quad \text{et} \quad \frac{nc_2}{\sigma_0^2} = q_{1-\frac{\alpha}{2}}$$

Où $q_{\frac{\alpha}{2}}$ = quantile d'ordre $\frac{\alpha}{2}$ d'une χ^2 lue sur la table de la Khi-deux pour $(n-1)$ d.d.l et pour $(1-\frac{\alpha}{2})$ noté $\chi^2_{1-\frac{\alpha}{2}, (n-1)}$

et $q_{1-\frac{\alpha}{2}}$ = quantile d'ordre $1-\frac{\alpha}{2}$ d'une χ^2 lue sur la table pour $(n-1)$ d.d.l et pour $(\frac{\alpha}{2})$ noté $\chi^2_{\frac{\alpha}{2}, (n-1)}$

$$\Rightarrow \boxed{\begin{aligned} c_1 &= \frac{\sigma_0^2}{n} \chi^2_{1-\frac{\alpha}{2}, (n-1)} \\ c_2 &= \frac{\sigma_0^2}{n} \chi^2_{\frac{\alpha}{2}, (n-1)} \end{aligned}}$$

Scanné avec CamScanner