

## CHAPITRE 4 - INTERVALLE DE CONFIANCE

Marouane IL Idrissi  
il.idrissi.marouane@uqam.ca

STT 1000 - Automne 2025  
Département de Mathématiques, Université du Québec à Montréal



### Intervalle de confiance

Dans le **Chapitre 3**, on cherchait à faire une **estimation ponctuelle** d'un paramètre  $\theta$  en se basant sur un échantillon  $\mathbf{X} = (X_1, \dots, X_n)$

Après avoir trouvé un **bon estimateur**  $\hat{\theta}(\mathbf{X})$ , et calculé une **estimation**  $\hat{\theta}(x)$  à l'aide d'un échantillon observé  $x$ , **quelle confiance peut-on avoir en  $\hat{\theta}(x)$** ?

On sait que  $\hat{\theta}(x) \neq \theta$  à moins d'avoir **observé toute la population** (qui est supposée infinie)...

#### Intuition:

Construire un **intervalle**, qui contiendrait **la vraie valeur de  $\theta$**  avec **une grande probabilité**

Par exemple, si on accepte un **risque de se de se tromper de 5%**, est-ce qu'on peut **trouver un intervalle  $[a(\mathbf{X}), b(\mathbf{X})]$  à partir de l'échantillon** tel que:

$$\mathbb{P}(a(\mathbf{X}) \leq \theta \leq b(\mathbf{X})) = 0.95$$

☞ **C'est exactement la définition d'un intervalle de confiance!**

### Notes

## Plan du chapitre

1. Préliminaires probabilistes
2. Intervalles de confiance exacts
3. Intervalles de confiance asymptotiques
4. Intervalles de confiance pour deux échantillons Gaussiens
5. Précision et taille d'échantillon

Notes

## Sommaire

1. Préliminaires probabilistes
  - 1.1 Loi du Chi-Carré
  - 1.2 Loi de Student
2. Intervalles de confiance exacts
3. Intervalles de confiance asymptotiques
4. Intervalles de confiance pour deux échantillons Gaussiens
5. Précision et taille d'échantillon

[illegible]

## Loi du Chi-Carré

Une v.a. de **loi du Chi-Carré** (ou Khi-Carré, ou  $\chi^2$ ) est une v.a. **continue positive**

Elle est paramétrée par **un degré de liberté**  $k \in \mathbb{N} \setminus \{0\}$

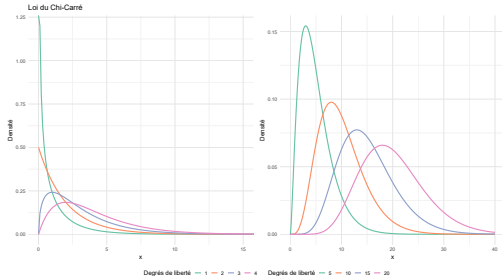
On note  $X \sim \chi^2(k)$ , et son support est  $(0, \infty)$

**Caractéristiques** (*Loi du Chi-Carré*). ★

- $f_X(t) = \frac{t^{k/2-1} \exp(-t/2)}{2^{k/2} \Gamma(k/2)} \mathbb{1}_{(0,\infty)}(t)$
- $\mathbb{E}[X] = k$
- $\mathbb{V}(X) = 2k$

4/58

## Loi du Chi-Carré - Illustration



5/58

## Notes

This image shows a single sheet of white paper with horizontal blue or grey ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.

## Loi du Chi-Carré - Propriétés

**Propriétés** (Chi-Carré = Somme de carré de Gaussiennes standard). ★

Soient  $Z_1, \dots, Z_n$  des **variables aléatoires Gaussiennes standard et indépendantes**. Alors,

$$V = \sum_{i=1}^n Z_i^2 \sim \chi^2(n)$$

Solent  $Z_1, \dots, Z_n$  des **variables aléatoires Gaussiennes standard et indépendantes**. Alors,

$$V = \sum_{i=1}^n Z_i^2 \sim \chi^2(n)$$

$$V = \sum_{i=1}^B Z_i^2 \sim \chi^2(n)$$

- Si  $Z \sim \mathcal{N}(0, 1)$ , quelle est la loi de  $Z^2$ ? 
- En déduire la valeur de  $\mathbb{E}[Z^4]$ . 

• En déduire la valeur de  $\mathbb{E}[Z^4]$ . 

**Propriétés** (Somme de Chi-Carrés = Chi-Carré). ★

Soient  $Y_1 \sim \chi^2(n_1)$  et  $Y_2 \sim \chi^2(n_2)$  deux v.a.s **indépendantes**, alors

$$V = Y_1 + Y_2 \sim \chi^2(n_1 + n_2).$$

Soient  $Y_1 \sim \chi^2(n_1)$  et  $Y_2 \sim \chi^2(n_2)$  deux v.a.s **indépendantes**, alors

$$V = Y_1 + Y_2 \sim \chi^2(n_1 + n_2).$$

$$V = Y_1 + Y_2 \sim \chi^2(n_1 + n_2).$$

6/58

## Loi de Student

Une v.a. de **loi de Student** est une v.a. **continue**

Elle est paramétrée par **un degré de liberté**  $k \in \mathbb{N} \setminus \{0\}$

On note  $X \sim \mathcal{T}(k)$  (des fois  $t(k)$ ), et son support est  $\mathbb{R}$

**Caractéristiques (Loi de Student) :** ★

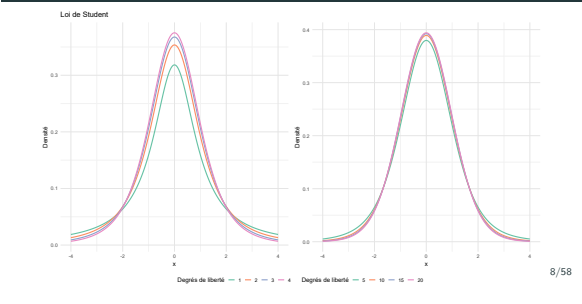
- $f_X(t) = \frac{\Gamma(\frac{k+1}{2})}{\sqrt{k\pi} \Gamma(\frac{k}{2})} \left(1 + \frac{t^2}{k}\right)^{-\frac{k+1}{2}}$
- $\mathbb{E}[X] = 0$  pour  $k \geq 2$ , n'existe pas sinon
- $\mathbb{V}(X) = \frac{k}{k-2}$  pour  $k > 2$ , n'existe pas sinon

- $f_X(t) = \frac{r^{\frac{k+1}{2}}}{\sqrt{k\pi} r^{\frac{k}{2}}} \left(1 + \frac{t^2}{k}\right)^{-\frac{k+1}{2}}$
- $\mathbb{E}[X] = 0$  pour  $k \geq 2$ , n'existe pas sinon
- $\mathbb{V}(X) = \frac{k}{k-2}$  pour  $k > 2$ , n'existe pas sinon

Notes

7/58

## Loi de Student - Illustration



8/58

## Loi de Student - Propriétés

**Propriétés** (*Student = Gaussienne divisée par Chi-Carré*). ★

Soit  $Z \sim \mathcal{N}(0, 1)$  et  $V \sim \chi^2(k)$  deux v.a.s **indépendantes**. Alors:

$$T = Z \sqrt{\frac{k}{V}} \sim \mathcal{T}(k)$$

**Propriétés** (*Comportement asymptotique de la loi de Student*). ★

Soit  $f_T(s; k)$  la densité de la loi de Student, et  $\phi(s)$  la densité **normale standard**. Alors,

$$f_T(s; k) \xrightarrow[k \rightarrow \infty]{} \phi(s).$$

Plus le degré de liberté augmente, plus la loi de Student se rapproche d'une  $\mathcal{N}(0, 1)$ .

## Notes

This image shows a single sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.

## Sommaire

1. Préliminaires probabilistes
2. Intervalles de confiance exacts
  - 2.1 Motivation et définition
  - 2.2 IC unilatéraux et bilatéraux
  - 2.3 Recette
  - 2.4 Intervalles de confiance avec un échantillon Gaussien
3. Intervalles de confiance asymptotiques
4. Intervalles de confiance pour deux échantillons Gaussiens
5. Précision et taille d'échantillon

Intervalles de confiance

## Intervalles de confiance

L'idée des **intervalles de confiance** (IC) est de **construire un intervalle aléatoire**, qui **contiendrait le paramètre que l'on souhaite estimer avec une grande probabilité**

Un **intervalle aléatoire**, c'est simplement un intervalle  $[X, Y]$ , où  $X$  et  $Y$  sont deux v.a.s.

Cet intervalle aléatoire est construit à l'aide de **deux statistiques**  $\hat{\theta}_1(X)$  et  $\hat{\theta}_2(X)$

Une statistique, c'est **une fonction d'un échantillon**

**Définition** (*Intervalle de confiance*). ★

Soit  $\mathbf{X} = (X_1, \dots, X_n)$  un échantillon i.i.d. de loi populationnelle  $f(t, \theta)$ . Soit  $\alpha \in (0, 1)$  un **niveau de risque**. Un **intervalle de confiance au niveau  $1 - \alpha$  pour le paramètre  $\theta$**  est un intervalle aléatoire de la forme  $IC_\theta(\mathbf{X}) = [\hat{\theta}_l(\mathbf{X}), \hat{\theta}_s(\mathbf{X})]$ , tel que:

$$\mathbb{P}(\theta \in IC_\theta(\mathbf{X})) = \mathbb{P}(\hat{\theta}_l(\mathbf{X}) \leq \theta \leq \hat{\theta}_s(\mathbf{X})) = 1 - \alpha$$

Soit  $\mathbf{X} = (X_1, \dots, X_n)$  un échantillon i.i.d. de loi populationnelle  $f(t, \theta)$ . Soit  $\alpha \in (0, 1)$  un **niveau de risque**. Un **intervalle de confiance au niveau  $1 - \alpha$  pour le paramètre  $\theta$**  est un intervalle aléatoire de la forme  $IC_\theta(\mathbf{X}) = [\hat{\theta}_l(\mathbf{X}), \hat{\theta}_s(\mathbf{X})]$ , tel que:

$$\mathbb{P}(\theta \in IC_\theta(\mathbf{X})) = \mathbb{P}(\hat{\theta}_l(\mathbf{X}) \leq \theta \leq \hat{\theta}_s(\mathbf{X})) = 1 - \alpha$$
[illegible]

## Intervalles de confiance

Le niveau de risque  $\alpha$  contrôle le risque que  $\theta$  ne soit pas dans cet intervalle:

- Pour  $\alpha \times 100\%$  des **échantillons observés**,  $\theta$  ne sera pas dans cet intervalle
- pour  $(1 - \alpha) \times 100\%$  des **échantillons observés**,  $\theta$  sera effectivement dans cet intervalle

On dit qu'un **intervalle de confiance** pour  $\theta$  est **informatif** si: ★

1.  $\alpha$  est petit ;
2. La largeur de l'intervalle est petite.

**Par exemple:**

L'intervalle  $[0, 1]$  pour le  $p$  d'une Bernoulli **n'est pas informatif**:

1.  $\alpha = 0$ , donc on est sûr que  $\theta$  est toujours dedans ;
2. La largeur de l'intervalle est trop grande pour être utile.

☞ Généralement, on fixe le risque  $\alpha$  à 0.1, 0.05 ou 0.01

Quels sont les **niveaux de confiance** associés? 🚩

12/58

## Intervall de confiança - Unilateral, bilateral

Il y a trois types d'IC: ★

- **Unilatéral à droite**, de la forme  $[-\infty, \hat{\theta}_S(\mathbf{X})]$ , tel que:  $\mathbb{P}(\theta \leq \hat{\theta}_S(\mathbf{X})) = 1 - \alpha$
- **Unilatéral à gauche**, de la forme  $[\hat{\theta}_I(\mathbf{X}), \infty]$ , tel que:  $\mathbb{P}(\hat{\theta}_I(\mathbf{X}) \leq \theta) = 1 - \alpha$
- **Bilatéral**, de la forme  $[\hat{\theta}_I(\mathbf{X}), \hat{\theta}_S(\mathbf{X})]$ , tel que:  $\mathbb{P}(\hat{\theta}_I(\mathbf{X}) \leq \theta \leq \hat{\theta}_S(\mathbf{X})) = 1 - \alpha$

☞ Un intervalle de confiance bilatéral n'est pas nécessairement symétrique autour de  $\theta$ !

Pourquoi utiliser chaque type d'IC: ★

- **Unilatéral à droite** : Borne supérieure pour  $\theta$
- **Unilatéral à gauche** : Borne inférieure pour  $\theta$
- **Bilatéral** : Encadrer la valeur précise de  $\theta$

13/58

## Notes

This image shows a full page of white paper with horizontal blue or grey ruling lines. The lines are evenly spaced and run across the width of the page, providing a template for handwriting practice. There are no margins, text, or other markings on the paper.

## Intervalle de confiance - Recette pour construire un IC exact

Pour construire un IC pour  $\theta$  inconnu à partir d'un échantillon  $\mathbf{X} = (X_1, \dots, X_n)$ :

1. Proposer un estimateur  $\hat{\theta}(\mathbf{X})$  de  $\theta$
2. Déterminer la loi exacte de  $\hat{\theta}(\mathbf{X})$
3. Construire un pivot  $\Delta$  entre  $\hat{\theta}(\mathbf{X})$  et  $\theta$ , par exemple:

$$\Delta = \hat{\theta}(\mathbf{X}) - \theta \quad \text{ou} \quad \Delta = \hat{\theta}(\mathbf{X})/\theta.$$

Un pivot  $\Delta$  est une v.a. dont la loi ne dépend pas de  $\theta$ !!

⚠ Ce dernier point est **crucial** pour la construction d'un IC

4. Choisir un **niveau de risque**  $\alpha$ , et donc un **niveau de confiance**  $1 - \alpha$ . Se baser sur la **loi exacte** de  $\Delta$  (table de probabilités) pour trouver **deux nombres**  $d_l$  et  $d_s$  tels que

$$\mathbb{P}(d_l \leq \Delta \leq d_s) = 1 - \alpha$$

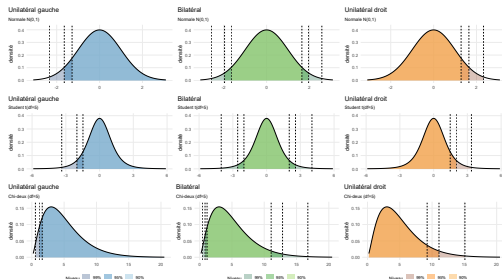
5. Utiliser la définition de  $\Delta$ , et les quantités  $d_l$  et  $d_s$  pour **isoler**  $\theta$  dans le calcul de probabilité au-dessus, par exemple pour  $\Delta = \hat{\theta}(\mathbf{X}) - \theta$ :

$$\mathbb{P}(d_l \leq \hat{\theta}(\mathbf{X}) - \theta \leq d_s) = 1 - \alpha, \quad \Leftrightarrow \quad \mathbb{P}(\hat{\theta}(\mathbf{X}) - d_s \leq \theta \leq \hat{\theta}(\mathbf{X}) - d_l) = 1 - \alpha,$$

14/58

Notes

## Intervalle de confiance - Illustration, loi du pivot



15/58



### Échantillon Gaussien - IC bilatéral pour $\mu$ , variance connue

On a  $X_1, \dots, X_n$  i.i.d. de loi populationnelle  $\mathcal{N}(\mu, \sigma_*^2)$ , avec  $\mu$  **inconnu**, mais  $\sigma_*^2$  est connu  
(Par exemple, on sait que  $\sigma_*^2 = 2$ ).

On veut construire un IC bilatéral pour  $\theta = \mu$ .

1. Proposez un estimateur pour  $\mu$  🏆
2. Quelle est la loi de cet estimateur? 🏆
3. Trouvez un pivot tel quel  $\Delta \sim \mathcal{N}(0, 1)$  (qui ne dépend pas de  $\mu$ ) 🏆
4. Quels deux nombres  $d_l$  et  $d_s$  permettent d'avoir

$$\mathbb{P}(d_l \leq \Delta \leq d_S) = 1 - \alpha ?$$

- Proposez un intervalle de confiance  $[\hat{\mu}_I(\mathbf{X}), \hat{\mu}_S(\mathbf{X})]$  pour  $\mu$  au niveau  $1 - \alpha$ . 🚩
- Supposons que  $n = 8$ ,  $\alpha = 0.05$ ,  $\sigma_x^2 = 2$  et que nous avons comme échantillon observé

$x = (45.44, 45.68, 45.65, 44.92, 45.88, 48.61, 43.94, 43.75)$ , et donc  $\bar{x}_n = 45.48$ .

Calculez l'intervalle observé  $[\hat{\mu}_I(\mathbf{x}), \hat{\mu}_S(\mathbf{x})]$  et interprétez-le. 📌

16/58

### Échantillon Gaussien - IC unilatéral à droite pour $\mu$ , variance connue

On a  $X_1, \dots, X_n$  i.i.d. de loi populationnelle  $\mathcal{N}(\mu, \sigma_*^2)$ , avec  $\mu$  inconnu, mais  $\sigma_*^2$  est connu  
(Par exemple, on sait que  $\sigma_*^2 = 2$ ).

On veut construire un IC unilatéral à droite pour  $\theta = \mu$ .

1. Proposez un estimateur pour  $\mu$  🚩
2. Quelle est la loi de cet estimateur? 🚩
3. Trouvez un pivot tel quel  $\Delta \sim \mathcal{N}(0, 1)$  (qui ne dépend pas de  $\mu$ ) 🚩
4. Quels nombre  $d_c$  permet d'avoir

$$\mathbb{P}(\Delta \leq d_S) = 1 - \alpha ?$$

- Proposez un intervalle de confiance  $(-\infty, \hat{\mu}_S(\mathbf{X})]$  pour  $\mu$  au niveau  $1 - \alpha$ .
- Calculez l'intervalle observé  $(-\infty, \hat{\mu}_S(\mathbf{x})]$  basé sur les données précédentes et interprétez-le.

17/58

## Notes

This image shows a single sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.

### Échantillon Gaussien - IC unilatéral à gauche pour $\mu$ , variance connue

On a  $X_1, \dots, X_n$  i.i.d. de loi populationnelle  $\mathcal{N}(\mu, \sigma_*^2)$ , avec  $\mu$  **inconnu**, mais  $\sigma_*^2$  est connu  
(Par exemple, on sait que  $\sigma_*^2 = 2$ ).

On veut construire un IC unilatéral à gauche pour  $\theta = \mu$ .

1. Proposez un estimateur pour  $\mu$
2. Quelle est la loi de cet estimateur?
3. Trouvez un pivot tel quel  $\Delta \sim \mathcal{N}(0, 1)$  (qui ne dépend pas de  $\mu$ )
4. Quels nombre  $d_\epsilon$  permet d'avoir

$$\mathbb{P}(\Delta \geq d_I) = 1 - \alpha ?$$

5. Proposez un intervalle de confiance  $[\hat{\mu}_I(\mathbf{X}), \infty)$  pour  $\mu$  au niveau  $1 - \alpha$ . 🍷
6. Calculez l'intervalle observé  $[\hat{\mu}_I(\mathbf{x}), \infty)$  basé sur les données précédentes et interprétez-le. 🍷

18/58

### Échantillon Gaussien - IC bilatéral pour $\mu$ , variance inconnue

On a  $X_1, \dots, X_n$  i.i.d. de loi populationnelle  $\mathcal{N}(\mu, \sigma^2)$ , avec  $\mu$  et  $\sigma^2$  **inconnus**.  
On veut construire un IC bilatéral pour  $\theta = \mu$ .

**Idée:** Remplacer  $\sigma^2$  par un estimateur!

1. Proposez un estimateur pour  $\mu$  et un estimateur pour  $\sigma^2$ .
2. Quelles sont les lois de ces estimateurs?
3. En utilisant le résultat suivant (non démontré):

**Lemme** (*Lemme de Basu*). ★

Soit  $X$  un échantillon Gaussien i.i.d. Alors,  $\bar{X}_n \perp\!\!\!\perp S^2$ .

Trouvez un pivot tel quel  $\Delta \sim \mathcal{T}(n-1)$ .

4. Quels deux nombres  $d_i$  et  $d_s$  permettent d'avoir

$$\mathbb{P}(d_l < \Delta < d_s) = 1 - \alpha ?$$

5. Proposez un intervalle de confiance  $[\hat{\mu}_I(\mathbf{X}), \hat{\mu}_S(\mathbf{X})]$  pour  $\mu$  au niveau  $1 - \alpha$ . 🍷

19/58

## Notes

### Échantillon Gaussien - IC unilatéral à droite pour $\mu$ , variance inconnue

On a  $X_1, \dots, X_n$  i.i.d. de loi populationnelle  $\mathcal{N}(\mu, \sigma_*^2)$ , avec  $\mu$  et  $\sigma^2$  **inconnus**.

On veut construire un IC unilatéral à droite pour  $\theta = \mu$ .

1. Proposez un estimateur pour  $\mu$  et un estimateur pour  $\sigma^2$ .
2. Quelles sont les lois de ces estimateurs?
3. Trouvez un pivot tel que  $\Delta \sim \mathcal{T}(n-1)$ .
4. Quel nombre  $d_c$  permettent d'avoir

$\mathbb{P}(\Delta \leq d_S) = 1 - \alpha$  ? 🍷

5. Proposez un intervalle de confiance  $(-\infty, \hat{\mu}_S(\mathbf{X})]$  pour  $\mu$  au niveau  $1 - \alpha$ . 🍷

20/58

### Échantillon Gaussien - IC unilatéral à gauche pour $\mu$ , variance inconnue

On a  $X_1, \dots, X_n$  i.i.d. de loi populationnelle  $\mathcal{N}(\mu, \sigma^2)$ , avec  $\mu$  et  $\sigma^2$  **inconnus**.

On veut construire un IC unilatéral à gauche pour  $\theta = \mu$ .

1. Proposez un estimateur pour  $\mu$  et un estimateur pour  $\sigma^2$ .
2. Quelles sont les lois de ces estimateurs?
3. Trouvez un pivot tel quel  $\Delta \sim \mathcal{T}(n-1)$ .
4. Quel nombre  $d_i$  permettent d'avoir

$$\mathbb{P}(\Delta \geq d_I) = 1 - \alpha ?$$

5. Proposez un intervalle de confiance  $[\hat{\mu}_I(\mathbf{X}), \infty)$  pour  $\mu$  au niveau  $1 - \alpha$ . 📌

21/58

## Notes

[illegible]

### Échantillon Gaussien - IC bilatéral pour $\sigma^2$ , moyenne connue

On a  $X_1, \dots, X_n$  i.i.d. de loi populationnelle  $\mathcal{N}(\mu_*, \sigma^2)$ , avec  $\mu_*$  connu.

On cherche à construire un IC bilatéral pour le paramètre  $\theta = \sigma^2$ .

1. Proposez un estimateur pour  $\sigma^2$ .
2. Quelle est la loi de cet estimateur?
3. Trouvez un pivot tel quel  $\Delta \sim \chi^2(n)$ .
4. Quels deux nombres  $d_l$  et  $d_s$  permettent d'avoir

$$\mathbb{P}(d_l \leq \Delta \leq d_s) = 1 - \alpha ?$$

5. Proposez un intervalle de confiance  $[\widehat{\sigma}_I^2(\mathbf{X}), \widehat{\sigma}_S^2(\mathbf{X})]$  pour  $\sigma^2$  au niveau  $1 - \alpha$ . 🍷

22/58

### Échantillon Gaussien - IC unilatéral à droite pour $\sigma^2$ , moyenne connue

On a  $X_1, \dots, X_n$  i.i.d. de loi populationnelle  $\mathcal{N}(\mu_*, \sigma^2)$ , avec  $\mu_*$  connu.

On cherche à construire un IC unilatéral à droite pour le paramètre  $\theta = \sigma^2$ .

1. Proposez un estimateur pour  $\sigma^2$ .
2. Quelle est la loi de cet estimateur?
3. Trouvez un pivot tel quel  $\Delta \sim \chi^2(n)$ .
4. Quel nombre  $d_i$  permettent d'avoir

$$\mathbb{P}(\Delta \geq d_I) = 1 - \alpha ?$$

5. Proposez un intervalle de confiance  $[\hat{\mu}_I(\mathbf{X}), \infty)$  pour  $\mu$  au niveau  $1 - \alpha$ . 🍷

23/58

## Notes

This image shows a full page of white paper with horizontal blue or grey ruling lines. The lines are evenly spaced and run across the width of the page, typical of notebook paper. There are no margins, text, or other markings on the page.

### Échantillon Gaussien - IC unilatéral à gauche pour $\sigma^2$ , moyenne connue

On a  $X_1, \dots, X_n$  i.i.d. de loi populationnelle  $\mathcal{N}(\mu_*, \sigma^2)$ , avec  $\mu_*$  connu.

On cherche à construire un IC unilatéral à gauche pour le paramètre  $\theta = \sigma^2$ .

1. Proposez un estimateur pour  $\sigma^2$ .
2. Quelle est la loi de cet estimateur?
3. Trouvez un pivot tel quel  $\Delta \sim \chi^2(n)$ .
4. Quel nombre  $d_i$  permettent d'avoir

$\mathbb{P}(\Delta \leq d_S) = 1 - \alpha$  ? 🍷

5. Proposez un intervalle de confiance  $(-\infty, \hat{\mu}_S(\mathbf{X})]$  pour  $\mu$  au niveau  $1 - \alpha$ . 🍷

24/58

### Échantillon Gaussien - IC bilatéral pour $\sigma^2$ , moyenne inconnue

On a  $X_1, \dots, X_n$  i.i.d. de loi populationnelle  $\mathcal{N}(\mu, \sigma^2)$ , avec  $\mu$  et  $\sigma^2$  **inconnus**.

On cherche à construire un IC bilatéral pour le paramètre  $\theta = \sigma^2$ .

**Idée: Remplacer  $\mu$  par un estimateur!**

1. Proposez un estimateur pour  $\mu$  et  $\sigma^2$ .
2. Quelle est la loi de cet estimateur?
3. Trouvez un pivot tel quel  $\Delta \sim \chi^2(n-1)$ .
4. Quels deux nombres  $d_t$  et  $d_s$  permettent d'avoir

$$\mathbb{P}(d_l \leq \Delta \leq d_s) = 1 - \alpha ?$$

5. Proposez un intervalle de confiance  $[\hat{\sigma}_{\beta_1}^2(\mathbf{X}), \hat{\sigma}_{\beta_2}^2(\mathbf{X})]$  pour  $\sigma^2$  au niveau  $1 - \alpha$ . 🚩

25/58

## Notes

### Échantillon Gaussien - IC unilatéral à droite pour $\sigma^2$ , moyenne inconnue

On a  $X_1, \dots, X_n$  i.i.d. de loi populationnelle  $\mathcal{N}(\mu, \sigma^2)$ , avec  $\mu$  et  $\sigma^2$  **inconnus**.

On cherche à construire un IC bilatéral pour le paramètre  $\theta = \sigma^2$ .

**Idée:** Remplacer  $\mu$  par un estimateur!

1. Proposez un estimateur pour  $\mu$  et un estimateur pour  $\sigma^2$ .
2. Quelle est la loi de cet estimateur?
3. Trouvez un pivot tel quel  $\Delta \sim \chi^2(n-1)$ .
4. Quel nombre  $d_i$  permettent d'avoir

$\mathbb{P}(\Delta < d_S) = 1 - \alpha$  ? 

5. Proposez un intervalle de confiance  $(\hat{\sigma}^2_I(\mathbf{X}), \infty)$  pour  $\sigma^2$  au niveau  $1 - \alpha$ . 🍷

26/58

### Échantillon Gaussien - IC unilatéral à gauche pour $\sigma^2$ , moyenne inconnue

On a  $X_1, \dots, X_n$  i.i.d. de loi populationnelle  $\mathcal{N}(\mu, \sigma^2)$ , avec  $\mu$  et  $\sigma^2$  **inconnus**.

On cherche à construire un IC bilatéral pour le paramètre  $\theta = \sigma^2$ .

**Idée:** Remplacer  $\mu$  par un estimateur!

1. Proposez un estimateur pour  $\mu$  et un estimateur pour  $\sigma^2$ .
2. Quelle est la loi de cet estimateur?
3. Trouvez un pivot tel quel  $\Delta \sim \chi^2(n-1)$ .
4. Quel nombre  $d_i$  permettent d'avoir

$$\mathbb{P}(d_i \leq \Delta) = 1 - \alpha ?$$

5. Proposez un intervalle de confiance  $(-\infty, \widehat{\sigma}_S^2(\mathbf{X}))$  pour  $\sigma^2$  au niveau  $1 - \alpha$ . 🍷

27/58

## Notes

[illegible]

Paramètre et type	Situation	$\hat{\theta}(\mathbf{X})$	Pivot $\Delta$	IC exact
Bilatéral pour $\mu$	$\sigma^2$ connu	$\bar{X}_n$	$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$	$\left[ \bar{X}_n - z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} ; \bar{X}_n + z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \right]$
Uni. droite pour $\mu$	$\sigma^2$ connu	$\bar{X}_n$	$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$	$\left( -\infty ; \bar{X}_n + z_{1-\alpha} \frac{\sigma}{\sqrt{n}} \right]$
Uni. gauche pour $\mu$	$\sigma^2$ connu	$\bar{X}_n$	$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$	$\left[ \bar{X}_n - z_{1-\alpha} \frac{\sigma}{\sqrt{n}} ; \infty \right)$
Bilatéral pour $\mu$	$\sigma^2$ inconnu	$\bar{X}_n, S^2$	$\frac{\bar{X}_n - \mu}{S/\sqrt{n}} \sim \mathcal{T}(n-1)$	$\left[ \bar{X}_n - t_{1-\alpha/2, n-1} \frac{S}{\sqrt{n}} ; \bar{X}_n + t_{1-\alpha/2, n-1} \frac{S}{\sqrt{n}} \right]$
Uni. droite pour $\mu$	$\sigma^2$ inconnu	$\bar{X}_n, S^2$	$\frac{\bar{X}_n - \mu}{S/\sqrt{n}} \sim \mathcal{T}(n-1)$	$\left( -\infty ; \bar{X}_n + t_{1-\alpha, n-1} \frac{S}{\sqrt{n}} \right]$
Uni. gauche pour $\mu$	$\sigma^2$ inconnu	$\bar{X}_n, S^2$	$\frac{\bar{X}_n - \mu}{S/\sqrt{n}} \sim \mathcal{T}(n-1)$	$\left[ \bar{X}_n - t_{1-\alpha, n-1} \frac{S}{\sqrt{n}} ; \infty \right)$
Bilatéral pour $\sigma^2$	$\mu$ connu	$S_0^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$	$\frac{nS_0^2}{\sigma^2} \sim \chi^2(n)$	$\left[ \frac{nS_0^2}{c_{1-\alpha/2, n}} ; \frac{nS_0^2}{c_{\alpha/2, n}} \right]$
Uni. droite pour $\sigma^2$	$\mu$ connu	$S_0^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$	$\frac{nS_0^2}{\sigma^2} \sim \chi^2(n)$	$\left( 0 ; \frac{nS_0^2}{c_{\alpha, n}} \right]$
Uni. gauche pour $\sigma^2$	$\mu$ connu	$S_0^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$	$\frac{nS_0^2}{\sigma^2} \sim \chi^2(n)$	$\left[ \frac{nS_0^2}{c_{1-\alpha, n}} ; \infty \right)$
Bilatéral pour $\sigma^2$	$\mu$ inconnu	$\bar{X}_n, S^2$	$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$	$\left[ \frac{(n-1)S^2}{c_{1-\alpha/2, n-1}} ; \frac{(n-1)S^2}{c_{\alpha/2, n-1}} \right]$
Uni. droite pour $\sigma^2$	$\mu$ inconnu	$\bar{X}_n, S^2$	$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$	$\left( 0 ; \frac{(n-1)S^2}{c_{\alpha, n-1}} \right]$
Uni. gauche pour $\sigma^2$	$\mu$ inconnu	$\bar{X}_n, S^2$	$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$	$\left[ \frac{(n-1)S^2}{c_{1-\alpha, n-1}} ; \infty \right)$

28/58

## Intervalle de confiance - Illustration

**Exemple** (Contrôleur de paquets cafés). Vous êtes contrôleur qualité dans une compagnie qui vend du café. Pour assurer la conformité des paquets de cafés, ils doivent en moyenne peser **500g**, avec une variabilité acceptée de **3g<sup>2</sup>**. On suppose que les masses (en g) suivent une loi Normale. Vous prélevez un échantillon de taille 10 :

498.7, 501.2, 499.8, 500.5, 497.9, 502.1, 499.3, 500.8, 501.0, 498.9

On observe  $\bar{x}_n = 500.2$ ,  $s^2 = 1.73$ ,  $s = 1.32$ ,  $s_0^2 = 1.56$

- ☞ Quelle est la loi de la population?
- ☞ Calculez les ICs bilatéraux aux niveaux 90%, 95% et 99% pour  $\mu$  et  $\sigma^2$ , dans toutes les configurations possibles.
- ☞ Peut-on conclure que la production de café est conforme à la qualité attendue?

## Notes

## Intervalle de confiance - Interprétation

Le niveau de confiance **ne garantie pas** que  $\theta$  soit à l'intérieur de l'intervalle!

Interprétation d'un **intervalle de confiance**  $IC_{\theta}(X)$  au niveau  $(1 - \alpha)$ :

- On suppose que l'on peut simuler **beaucoup d'échantillons observés de même taille  $n$**
- A partir de chacun échantillon, on **calcule un IC** au niveau  $1 - \alpha\%$
- Alors,  $\theta$  **sera dans approximativement  $1 - \alpha\%$  de ces échantillons**
- **Si on augmente le nombre d'échantillons simulés, alors la proportion d'IC qui contiennent  $\theta$  va tendre vers  $1 - \alpha\%$**

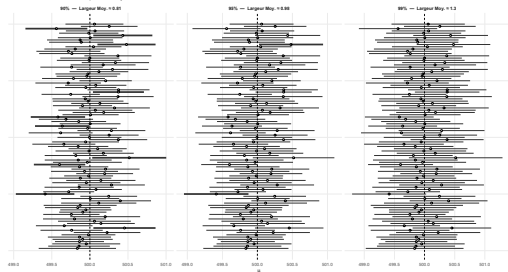
☞ **Pour un seul échantillon observé et un seul IC, on ne sait pas si  $\theta$  appartient effectivement à cet échantillon**

30/58

Notes

## Intervalle de confiance - Interprétation

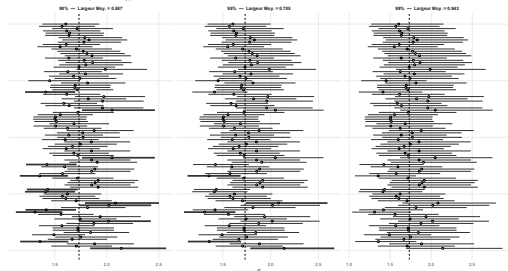
100 intervalles de confiance pour la moyenne



31/58



100 intervalles de confiance pour l'écart-type



32/58

## Sommaire

1. Préliminaires probabilistes
2. Intervalles de confiance exacts
3. Intervalles de confiance asymptotiques
  - 3.1 Rappel TLC
  - 3.2 Recette
  - 3.3 IC pour un échantillon quelconque
  - 3.4 Illustration
4. Intervalles de confiance pour deux échantillons Gaussiens
5. Précision et taille d'échantillon

33/58

Jusqu'à présent, on a **supposé que les échantillons étaient Gaussiens**

☞ C'était bien pratique pour **déterminer la loi du pivot**

Mais c'est un cadre de travail très limitant

On ne peut pas modéliser **TOUS** les phénomènes par les v.a. Gaussiennes

Par exemple, si on s'intéresse **un échantillon de durées de vie d'ampoules**

☞ C'est un phénomène qui se modélise généralement avec une loi exponentielle.

**Idée:** Utiliser le **Théorème de la Limite Centrale** pour approcher la loi du pivot

C'est la base des **ICs asymptotiques**

34/58

## Théorème de la Limite Centrale

**Théorème** (Théorème de la limite centrale (TLC)). ★

Soient  $X_1, \dots, X_n$  des v.a.s **i) indépendantes, ii) ayant toutes la même loi iii) avec**  $\mathbb{E}[X_i] = \mu \in \mathbb{R}$ ,  
**et**  $\forall (X_i) = \sigma^2 < \infty$ .

Alors la fonction (de  $n$ ) suivante admet comme limite

$$\sqrt{n} \left( \frac{\bar{X}_n - \mu}{\sigma} \right) \xrightarrow{n \rightarrow \infty} Z \sim \mathcal{N}(0, 1)$$

**Théorème** (Loi asymptotique du pivot pour  $\mu$ ). ★

Dans le cadre du TLC, on a que:

$$\sqrt{n} \left( \frac{\bar{X}_n - \mu}{S} \right) \xrightarrow{n \rightarrow \infty} Z \sim \mathcal{N}(0, 1)$$

☞ Quelle est la différence entre ces deux résultats? 📌

35/58

## Intervalle de confiance asymptotique pour $\mu$ - Recette

Pour construire un IC **asymptotique** pour  $\mathbb{E}[X_1] = \mu$  à partir d'un échantillon  $\mathbf{X} = (X_1, \dots, X_n)$ :

1. Utiliser l'estimateur  $\bar{X}_n$  pour  $\mu$
2. Utiliser le pivot:

$$\Delta = \sqrt{n} \left( \frac{\bar{X}_n - \mu}{S} \right)$$

3. Vérifier que l'on est dans un régime asymptotique:  $n > 30$

☞ Si (et seulement si) c'est le cas, supposer l'approximation suivante:

$$\Delta = \sqrt{n} \left( \frac{\bar{X}_n - \mu}{S} \right) \sim \mathcal{N}(0, 1)$$

4. Choisir un **niveau de risque**  $\alpha$ , et donc un **niveau de confiance**  $1 - \alpha$ . Se baser sur la loi **asymptotique** de  $\Delta$  (table de probabilités) qui donne, pour un IC bilatéral:

$$\mathbb{P}\left(-z_{1-\alpha/2} \leq \sqrt{n} \left( \frac{\bar{X}_n - \mu}{S} \right) \leq z_{1-\alpha/2} \right) = 1 - \alpha$$




5. Utiliser la définition de  $\Delta$  pour **isoler**  $\mu$  dans le calcul de probabilité au-dessus.

36/58

## IC pour l'espérance - Échantillon quelconque

On a  $X_1, \dots, X_n$  i.i.d. de loi populationnelle inconnue, avec:

- $\mu = \mathbb{E}[X_1] < \infty$
- $\mathbb{E}[X_1^2] < \infty$
- $n > 30$

1. Quelle est la loi asymptotique du pivot  $\Delta = \sqrt{n} \left( \frac{\bar{X}_n - \mu}{s} \right)$ ? 
2. Sommes-nous dans un régime asymptotique? 
3. Quelle approximation pouvons-nous accepter? 
4. Sous cette approximation, quels nombres  $d_t$  et  $d_s$  permettent d'avoir:

$$\mathbb{P}(d_j \leq \Delta \leq d_\xi) = 1 - \alpha?$$



5. Proposez un IC asymptotique bilatéral  $[\hat{\mu}_S(\mathbf{X}), \hat{\mu}_S(\mathbf{X})]$  pour  $\mu$  au niveau  $1 - \alpha$ . 
6. Proposez des ICs asymptotiques unilatéraux pour  $\mu$  au niveau  $1 - \alpha$ . 

## Notes

[illegible]

37/58

Paramètre et type	Condition	$\theta(\mathbf{X})$	Pivot $\Delta$	Loi asympto.	IC asympto
Bilatéral pour $\mu$	$n > 30 + \text{TLC}$	$\bar{X}_n$	$\sqrt{n} \left( \frac{\bar{X}_n - \mu}{s} \right)$	$\mathcal{N}(0, 1)$	$\left[ \bar{X}_n - z_{1-\alpha/2} \frac{s}{\sqrt{n}} ; \bar{X}_n + z_{1-\alpha/2} \frac{s}{\sqrt{n}} \right]$
Unil. droite pour $\mu$	$n > 30 + \text{TLC}$	$\bar{X}_n$	$\sqrt{n} \left( \frac{\bar{X}_n - \mu}{s} \right)$	$\mathcal{N}(0, 1)$	$\left( -\infty ; \bar{X}_n + z_{1-\alpha} \frac{s}{\sqrt{n}} \right]$
Unil. gauche pour $\mu$	$n > 30 + \text{TLC}$	$\bar{X}_n$	$\sqrt{n} \left( \frac{\bar{X}_n - \mu}{s} \right)$	$\mathcal{N}(0, 1)$	$\left[ \bar{X}_n - z_{1-\alpha} \frac{s}{\sqrt{n}} ; \infty \right)$

38/58

## Intervalle de confiance asymptotique pour $\mu$ - Illustration

**Exemple** (*Contrôleur d'ampoules*). Vous êtes contrôleur qualité dans une compagnie qui produit des ampoules. Pour assurer la qualité des produits, elles doivent en moyenne durer **20000h**. On ne connaît pas la loi de la durée de vie de ces ampoules, mais on sait que son espérance et sa variance existent. Vous prélevez un échantillon de taille 50 avec:

$$\bar{x}_n = 20026.94, \quad s = 164.49$$

« Sommes-nous dans un régime asymptotique? Si oui, quel est la loi asymptotique du pivot? »

« Calculez les ICs asymptotiques bilatéraux et unilatéraux aux niveaux 90%, 95% et 99% pour  $\mu$ . »

« Peut-on conclure que la durée de vie des ampoules est conforme? »

## Notes

39/58

1. Préliminaires probabilistes
2. Intervalles de confiance exacts
3. Intervalles de confiance asymptotiques
4. Intervalles de confiance pour deux échantillons Gaussiens
  - 4.1 Deux configurations
  - 4.2 Différence de moyennes
  - 4.3 Ratios de variances
5. Précision et taille d'échantillon

40/58

## Deux échantillons Gaussiens

Jusqu'à présent, nous n'avions accès **qu'à un seul échantillon i.i.d.**  
Ça nous a permis d'inférer de l'information sur les paramètres de la **loi de la population**

Comment faire si l'on a **deux populations différentes** que l'on souhaite comparer?

**Exemple** (*Méthodes de manufacture*). Vous êtes en charge d'une usine de construction de vélo. Vous voulez comparer **deux méthodes de fabrication**. On modélise le **temps de fabrication par la méthode A** par une Gaussienne d'espérance  $\mu_X$ , et le **temps de fabrication par la méthode B** par une Gaussienne d'espérance  $\mu_Y$ .

On cherche **une estimation par IC** du gain de productivité moyen  $\mu_X - \mu_Y$ .

Objectif de cette section:

**Construire des outils pour répondre à ce genre de problématique dans le cas Gaussien**

41/58

## Deux configurations

On a accès à **deux échantillons** issus de **deux populations différentes**:

- $\mathbf{X} = (X_1, \dots, X_{n_X})$  i.i.d. de taille  $n_X$  et de loi populationnelle  $\mathcal{N}(\mu_X, \sigma_X^2)$
- $\mathbf{Y} = (Y_1, \dots, Y_{n_Y})$  i.i.d. de taille  $n_Y$  et de loi populationnelle  $\mathcal{N}(\mu_Y, \sigma_Y^2)$

On dit que:

- X et Y sont **indépendants** si, pour  $i = 1, \dots, n_x$  et  $j = 1, \dots, n_y$ :  $X_i \perp\!\!\!\perp Y_j$  ★
- X et Y sont **appariés** si les individus forment des paires: ★
  - Les deux échantillons ont la même taille:  $n_x = n_y = n$
  - $X_i$  et  $Y_i$  proviennent du **même individu** et forment une **paire**  $(X_i, Y_i)$
  - Les paires sont **indépendantes entre elles**:  $(X_i, Y_i) \perp\!\!\!\perp (X_j, Y_j)$  pour  $i \neq j$ .

**Exemple (Échantillons appariés).** On veut mesurer l'effet d'un médicament sur la fièvre:

- $X_i$ : température corporelle de l'individu  $i$  **avant la prise du médicament**
- $Y_i$ : température corporelle du même individu  $i$  **4h après la prise du médicament**

42/58

## Différence de moyennes - Échantillons indépendants, variances connues

On a  $\mathbf{X} = (X_1, \dots, X_{n_X})$  i.i.d. de loi populationnelle  $\mathcal{N}(\mu_X, \sigma_X^2)$  et  $\mathbf{Y} = (Y_1, \dots, Y_{n_Y})$  i.i.d. de loi populationnelle,  $\mathcal{N}(\mu_Y, \sigma_Y^2)$  où  $\sigma_X^2$  et  $\sigma_Y^2$  **sont supposés connus**, et  $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}$ .

On cherche à construire des ICs pour  $\mu_X - \mu_Y$ .

1. Proposez un estimateur pour  $\mu_X$  et pour  $\mu_Y$ .
2. Quelle est la loi de la différence de ces estimateurs?
3. Trouvez un pivot tel quel  $\Delta \sim \mathcal{N}(0, 1)$ .
4. En déduire les intervalles de confiance unilatéraux et bilatéraux au niveau  $1 - \alpha$ .

## Notes

[illegible]

43/58

## Différence de moyennes - Échantillons indépendants, variances égales, inconnues

On a  $\mathbf{X} = (X_1, \dots, X_{n_X})$  i.i.d. de loi populationnelle  $\mathcal{N}(\mu_X, \sigma^2)$  et  $\mathbf{Y} = (Y_1, \dots, Y_{n_Y})$  i.i.d. de loi populationnelle,  $\mathcal{N}(\mu_Y, \sigma^2)$  où  $\sigma^2$  est **inconnu** et  $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}$ .

**Très important:** On suppose que les lois populationnelles des deux échantillons ont la même variance!

On cherche à construire des ICs pour  $\mu_X - \mu_Y$ .

1. Proposez un estimateur pour  $\mu_X$  et pour  $\mu_Y$ .
2. Quelle est la loi de la différence de ces estimateurs?
3. En utilisant le pivot

$$\Delta = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_p \sqrt{1/n_X + 1/n_Y}} \sim \mathcal{T}(n_X + n_Y - 2),$$

où

$$S_p^2 = \frac{(n_X - 1)S_X^2 + (n_Y - 1)S_Y^2}{n_X + n_Y - 2},$$

déduire les intervalles de confiance unilatéraux et bilatéraux au niveau  $1 - \alpha$ . 📌

44/58

## Différence de moyennes - Échantillons appariés

On a  $\mathbf{X} = (X_1, \dots, X_n)$  i.i.d. de loi populationnelle  $\mathcal{N}(\mu_X, \sigma_X^2)$  et  $\mathbf{Y} = (Y_1, \dots, Y_n)$  i.i.d. de loi populationnelle,  $\mathcal{N}(\mu_Y, \sigma_Y^2)$  où  $\sigma^2$  est **inconnu** et pour tout  $i \neq j$ ,  $(X_i, Y_i) \perp\!\!\!\perp (X_j, Y_j)$ .

**Très important:** Les deux échantillons sont appariés et donc pas indépendants. Leur taille est donc nécessairement la même.

On cherche à construire des ICs pour  $\mu_X - \mu_Y$

1. Pour une paire  $(X_i, Y_i)$ , déterminer la loi de  $D_i = X_i - Y_i$ . 📌
2. En déduire la loi de  $\bar{D}_n$ . 📌
3. Proposez un estimateur pour la variance de  $D_i$ . Quelle est sa loi? 📌
4. Trouvez un pivot tel quel  $\Delta \sim \mathcal{T}(n - 1)$ .
5. En déduire les intervalles de confiance unilatéraux et bilatéraux au niveau  $1 - \alpha$ . 📌

## Notes

## Différence de deux moyennes - Récapitulatif des ICs

Variane "poolée":  $S_p^2 = \frac{(n_X-1)S_X^2 + (n_Y-1)S_Y^2}{n_X + n_Y - 2}$

Ech.	Type	Situation	Estimateurs	Pivot $\Delta$
Indép.	Bilat.	V.C	$\bar{X}, \bar{Y}$	$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\sigma_X^2/n_X + \sigma_Y^2/n_Y}} \sim \mathcal{N}(0, 1)$
Indép.	Droite	V.C	$\bar{X}, \bar{Y}$	$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\sigma_X^2/n_X + \sigma_Y^2/n_Y}} \sim \mathcal{N}(0, 1)$
Indép.	Gauche	V.C	$\bar{X}, \bar{Y}$	$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\sigma_X^2/n_X + \sigma_Y^2/n_Y}} \sim \mathcal{N}(0, 1)$
Indép.	Bilat.	V.E.I	$\bar{X}, \bar{Y}, S_p^2$	$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_p \sqrt{1/n_X + 1/n_Y}} \sim T(n_X + n_Y - 2)$
Indép.	Droite	V.E.I	$\bar{X}, \bar{Y}, S_p^2$	$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_p \sqrt{1/n_X + 1/n_Y}} \sim T(n_X + n_Y - 2)$
Indép.	Gauche	V.E.I	$\bar{X}, \bar{Y}, S_p^2$	$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_p \sqrt{1/n_X + 1/n_Y}} \sim T(n_X + n_Y - 2)$
Appariés	Bilat.	V.I	$\bar{D}_n, S_D^2$	$\frac{\bar{D}_n - \mu_D}{S_D / \sqrt{n}} \sim T(n-1)$
Appariés	Droite	V.I	$\bar{D}_n, S_D^2$	$\frac{\bar{D}_n - \mu_D}{S_D / \sqrt{n}} \sim T(n-1)$
Appariés	Gauche	V.I	$\bar{D}_n, S_D^2$	$\frac{\bar{D}_n - \mu_D}{S_D / \sqrt{n}} \sim T(n-1)$

V.C = Variances connues; V.E.I = Variances égales inconnues; V.I = Variances inconnues

46/58

## Notes

## Différence de deux moyennes - Récapitulatif des ICs

Ech.	Type	Situation	Estimateurs	IC exact
Indép.	Bilat.	V.C	$\bar{X}, \bar{Y}$	$(\bar{X} - \bar{Y}) \pm z_{1-\alpha/2} \sqrt{\sigma_X^2/n_X + \sigma_Y^2/n_Y}$
Indép.	Droite	V.C	$\bar{X}, \bar{Y}$	$(-\infty; (\bar{X} - \bar{Y}) + z_{1-\alpha} \sqrt{\sigma_X^2/n_X + \sigma_Y^2/n_Y})$
Indép.	Gauche	V.C	$\bar{X}, \bar{Y}$	$(\bar{X} - \bar{Y}) - z_{1-\alpha} \sqrt{\sigma_X^2/n_X + \sigma_Y^2/n_Y}; \infty)$
Indép.	Bilat.	V.E.I	$\bar{X}, \bar{Y}, S_p^2$	$(\bar{X} - \bar{Y}) \pm t_{1-\alpha/2, n_X+n_Y-2} S_p \sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}$
Indép.	Droite	V.E.I	$\bar{X}, \bar{Y}, S_p^2$	$(-\infty; (\bar{X} - \bar{Y}) + t_{1-\alpha, n_X+n_Y-2} S_p \sqrt{\frac{1}{n_X} + \frac{1}{n_Y}})$
Indép.	Gauche	V.E.I	$\bar{X}, \bar{Y}, S_p^2$	$(\bar{X} - \bar{Y}) - t_{1-\alpha, n_X+n_Y-2} S_p \sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}; \infty)$
Appariés	Bilat.	V.I	$\bar{D}_n, S_D^2$	$\bar{D}_n \pm t_{1-\alpha/2, n-1} \frac{S_D}{\sqrt{n}}$
Appariés	Droite	V.I	$\bar{D}_n, S_D^2$	$(-\infty; \bar{D}_n + t_{1-\alpha, n-1} \frac{S_D}{\sqrt{n}})$
Appariés	Gauche	V.I	$\bar{D}_n, S_D^2$	$(\bar{D}_n - t_{1-\alpha, n-1} \frac{S_D}{\sqrt{n}}; \infty)$

V.C = Variances connues; V.E.I = Variances égales inconnues; V.I = Variances inconnues

47/58



## Loi Fisher-Snedecor

Une v.a. de **loi de Fisher-Snedecor** (ou simplement Fisher) est une v.a. **continue positive**

Elle est paramétrée par **deux degrés de libertés**  $\nu_1, \nu_2 \in \mathbb{N} \setminus \{0\}$

On note  $X \sim \mathcal{F}(\nu_1, \nu_2)$ , et son support est  $[0, \infty)$  et sa densité est donnée par:

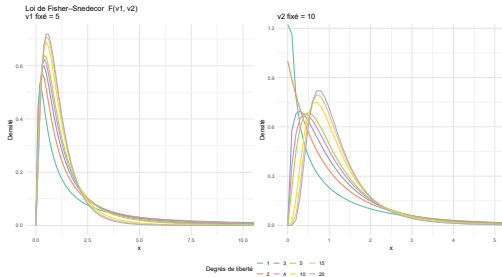
$$f_X(t) = \frac{\left(\frac{\nu_1 t}{\nu_1 t + \nu_2}\right)^{\nu_1/2} \left(1 - \frac{\nu_1 t}{\nu_1 t + \nu_2}\right)^{\nu_2/2}}{t B(\nu_1/2, \nu_2/2)} \mathbb{1}_{[0, \infty)}(t), \quad \text{ou} \quad B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$$

**Caractéristiques** (Loi de Fisher-Snedecor). ★

- $\mathbb{E}[X] = \frac{\nu_2}{\nu_2 - 2}$  pour  $\nu_2 > 2$
- $\mathbb{V}(X) = \frac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 2)^2(\nu_2 - 4)}$  pour  $\nu_2 > 4$

48/58

## Loi Fisher-Snedecor - Illustration



49/58

## Notes

This image shows a full page of blank, lined paper. It features approximately 20 evenly spaced horizontal blue or grey lines across its entire width. The lines are uniform in thickness and color, providing a standard template for handwriting practice or general note-taking. There are no margins, text, or other markings present on the page.

**Propriétés** (Fisher = Ratio de deux Chi-Carrés indépendantes). ★

Pour  $U_1 \sim \chi^2(\nu_1)$ ,  $U_2 \sim \chi^2(\nu_2)$ , et  $U_1 \perp U_2$ :

$$\frac{U_1/\nu_1}{U_2/\nu_2} \sim \mathcal{F}(\nu_1, \nu_2)$$

**Propriétés** (*Transformation et lien avec d'autres lois*). ★

- Si  $X \sim \mathcal{F}(\nu_1, \nu_2)$ , alors  $\frac{1}{X} \sim \mathcal{F}(\nu_2, \nu_1)$
- Si  $T \sim \mathcal{T}(\nu)$ , alors  $T^2 \sim \mathcal{F}(1, \nu)$

**Propriétés** (Quantile d'ordre de la loi de Fisher). ★

Soit  $f_{\alpha, \nu_1, \nu_2}$  le quantile d'ordre  $\alpha$  de  $X \sim \mathcal{F}(\nu_1, \nu_2)$ . Alors, pour tout  $\alpha \in (0, 1)$ :

$$f_{1-\alpha, \nu_2, \nu_1} = \frac{1}{f_{\alpha, \nu_1, \nu_2}}$$

50/58

## Ratio de variances - Échantillons indépendants

On a  $\mathbf{X} = (X_1, \dots, X_{n_X})$  i.i.d. de loi populationnelle  $\mathcal{N}(\mu_X, \sigma_X^2)$  et  $\mathbf{Y} = (Y_1, \dots, Y_{n_Y})$  i.i.d. de loi populationnelle,  $\mathcal{N}(\mu_Y, \sigma_Y^2)$  où **aucun des paramètres n'est connu** et  $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}$ .

On cherche à construire des ICs pour  $\frac{\sigma_x^2}{\sigma_y^2}$ .

1. Proposez un estimateur pour  $\sigma_X^2$  et pour  $\sigma_Y^2$ . Quelles sont leurs lois? 🍷
2. Trouvez un pivot tel quel  $\Delta \sim \mathcal{F}(n_X - 1, n_Y - 1)$ . 🍷
3. En déduire les intervalles de confiance unilatéraux et bilatéraux au niveau  $1 - \alpha$ . 🍷

This image shows a full page of white paper with horizontal blue or grey ruling lines. The lines are evenly spaced and run across the width of the page, typical of notebook paper. There are no margins, text, or other markings on the page.

Éch.	Type	Situation	Estimateurs	Pivot $\Delta$	IC exact
Indép.	Bilatéral	Moys. inconnues	$S_X^2, S_Y^2$	$\frac{S_X^2/S_Y^2}{\sigma_X^2/\sigma_Y^2} \sim F_{n-1, m-1}$	$\left[ \frac{S_X^2/S_Y^2}{F_{1-\alpha/2, n_X-1, n_Y-1}}; \frac{S_X^2/S_Y^2}{F_{\alpha/2, n_X-1, n_Y-1}} \right]$
Indép.	Uni. droite	Moys. inconnues	$S_X^2, S_Y^2$	$\frac{S_X^2/S_Y^2}{\sigma_X^2/\sigma_Y^2} \sim F_{n-1, m-1}$	$\left( 0; \frac{S_X^2/S_Y^2}{F_{1-\alpha, n_X-1, n_Y-1}} \right]$
Indép.	Uni. gauche	Moys. inconnues	$S_X^2, S_Y^2$	$\frac{S_X^2/S_Y^2}{\sigma_X^2/\sigma_Y^2} \sim F_{n-1, m-1}$	$\left[ \frac{S_X^2/S_Y^2}{F_{\alpha, n_X-1, n_Y-1}}; \infty \right)$

## Sommaire

1. Préliminaires probabilistes
2. Intervalles de confiance exacts
3. Intervalles de confiance asymptotiques
4. Intervalles de confiance pour deux échantillons Gaussiens
5. Précision et taille d'échantillon

## Notes

This image shows a single sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.

## Précision d'un intervalle de confiance

Dans plusieurs domaines, il est crucial de déterminer **combien d'observations** sont nécessaires pour que l'**intervalle de confiance** soit suffisamment précis

Par ex: Essais cliniques, sondages...

On doit d'abord définir la notion de **précision** d'un IC bilatéral

**Définition** (Précision d'un intervalle de confiance). ★

La **précision** d'un IC bilatéral  $[\hat{\theta}_S(\mathbf{X}), \hat{\theta}_I(\mathbf{X})]$  correspond à sa **demi-largeur**:

$$\delta = \frac{\hat{\theta}_S(\mathbf{X}) - \hat{\theta}_I(\mathbf{X})}{2} \quad (\text{pour un IC bilatéral}).$$

☞ Plus  $\delta$  est petit, plus l'IC est précis

54/58

## Précision d'un IC - Exemple Gaussien

Exemple — moyenne d'une population gaussienne (variance connue) :

$$\left[ \bar{X}_n - z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}; \bar{X}_n + z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \right]$$

La **précision** est alors :

$$\delta = z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}.$$

Remarques importantes : ★

- $\delta$  **diminue quand  $n$  augmente**: plus d'observations = intervalle plus serré
- $\delta$  **augmente avec  $1 - \alpha$** : plus le niveau de confiance est grand, plus l'intervalle est large ;
- $\delta$  **augmente avec  $\sigma$** : plus le phénomène est incertain, moins on est précis.

## Notes

55/58

## Calcul de taille échantillonnale

**Objectif :** fixer une précision  $\delta$  désirée et en déduire la taille minimale d'échantillon  $n$  nécessaire.

Pour l'exemple précédent, à partir de la relation:

$$\delta = z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}$$

on obtient: ★

$$n^* = \left( \frac{z_{1-\alpha/2} \sigma}{\delta} \right)^2$$

✎ L'arrondi à l'entier supérieur de  $n^*$  assure d'avoir une précision d'au moins  $\delta$  pour cet IC

56/58

## Calcul de taille échantillonnale - Exemple

**Example :**

- On veut estimer la moyenne  $\mu$  d'une population avec un IC à 95% de confiance
- On connaît  $\sigma = 3$
- On souhaite une précision  $d = 0.5$ .

Alors :

$$n = \left( \frac{1.96 \times 3}{0.5} \right)^2 = 138.3 \Rightarrow \boxed{n = 139.}$$

Ainsi, pour une précision donnée:

- Doubler la précision (diviser  $d$  par 2)  $\Rightarrow$  **multiplier  $n$  par 4;**
- Augmenter le niveau de confiance  $\Rightarrow$  **Augmente la taille nécessaire.**

## Notes

This image shows a full page of white paper with horizontal blue or grey ruling lines. The lines are evenly spaced and run across the width of the page, typical of notebook paper. There are no margins, text, or other markings on the page.

57/58

**Exemple (Sondage).** Durant des élections municipales une personne candidate souhaite estimer **par un IC bilatéral asymptotique à 95%** la proportion  $p$  de citoyens qui voteront pour elle. Elle souhaite une précision de 3 points de pourcentages ( $\delta = 0.03$ ).

Combien de personnes doit-il interroger au minimum?

Durant des élections municipales une personne candidate souhaite estimer **par un IC bilatéral asymptotique à 95%** la proportion  $p$  de citoyens qui voteront pour elle. Elle souhaite une précision de 3 points de pourcentages ( $\delta = 0.03$ ).

Combien de personnes doit-il interroger au minimum?

Combien de personnes doit-il interroger au minimum?

🔍 Comment modéliser l'échantillon? Quelle est sa loi populationnelle? 🚩

Donnez un estimateur de  $p$ .

✎ Proposez un pivot asymptotique pour cet estimateur. En déduire un IC asymptotique bilatéral. 📌

🔗 Quelle est la précision de cet IC? Quelle valeur de  $p$  la maximise? 🚩

☞ Dans le pire cas, quelle taille d'échantillon permettrait d'avoir un IC avec la précision demandée? 📌

This image shows a full page of white paper with horizontal blue or grey ruling lines. The lines are evenly spaced and run across the width of the page, typical of notebook paper. There are no margins, text, or other markings on the page.