Chapitre 5 - Tests d'hypothèses

Marouane Il Idrissi
il_idrissi.marouane@uqam.ca

STT 1000 - Automne 2025

Département de Mathématiques, Université du Québec à Montréal

Prise de décision sous incertitudes

Jusqu'à présent, on s'est **cantonnés à l'estimation de paramètres**:

- Chapitre 3: **Estimation ponctuelle chiffrée** du paramètre inconnu
- Chapitre 4: **Prise en compte de l'incertitude** dans l'estimation

Les **tests d'hypothèses** permettent de **répondre à des questions différente**:
☞ **La prise de décision sous incertitude**

Par exemple, on pourra répondre **avec un risque de se tromper de** $\alpha \in [0, 1]$ à:

- Est-ce qu'une politicienne va gagner l'élection?
- Est-ce que les pommes d'un verger font en moyenne 200gr?
- Est-ce qu'un nouveau traitement est plus efficace que l'ancien?
- Est-ce que deux variables aléatoires sont indépendantes?
- Est-ce que j'ai bien fait de modéliser ce phénomène par une loi normale?

Plan du chapitre

Sommaire

Tests d'hypothèses - Premier formalisme

On considère que l'on a accès à un échantillon i.i.d. de loi populationnelle ayant $\theta \in \Theta$ comme paramètre

On **partitionne** Θ en **deux parties disjointes et complémentaires** Θ_0 et Θ_1, c'est à dire:

$$\Theta = \Theta_0 \cup \Theta_1 \quad \text{et} \quad \Theta_0 \cap \Theta_1 = \emptyset$$

Au lieu **d'estimer** θ, on va **essayer de déterminer si** $\theta \in \Theta_0$ ou alors si $\theta \in \Theta_1$

Définition *(Hypothèse nulle et alternative)*. ★

L'hypothèse $\mathcal{H}_0$ que $\theta \in \Theta_0$ est appelée l'**hypothèse nulle**: $\mathcal{H}_0 : \theta \in \Theta_0$.
L'hypothèse $\mathcal{H}_1$ que $\theta \in \Theta_1$ est appelée l'**hypothèse alternative**: $\mathcal{H}_1 : \theta \in \Theta_1$.

Comme Θ_0 et Θ_1 sont **disjoints**, **seulement une des deux hypothèses peut être vraie**.

Tests d'hypothèses - Exemple

Exemple *(Elections municipales)*.

Une politicienne est candidate à une élection municipale. Ainsi, elle gagnera l'élection si elle obtient au moins 50% des votes.

☞ On modélise le phénomène "**A voté pour la candidate**" par une $\mathcal{B}(p)$, où $p \in [0,1]$ est inconnu

☞ Si $p \geq 0.5$, alors la candidate va gagner. Si $p < 0.5$ alors la candidate ne gagne pas

Ainsi, on peut formuler l'**hypothèse nulle**:

$$\mathcal{H}_0 : p \leq 0.5 \quad \text{(La candidate ne gagne pas)}$$
$$\mathcal{H}_1 : p > 0.5 \quad \text{(La candidate gagne)}$$

Dans ce cas là: $\Theta = [0,1]$, $\Theta_0 = [0\ ;\ 0.5]$ et $\Theta_1 = (0.5\ ;\ 1]$.

Types de tests

De manière générale, on peut **construire plusieurs types de tests**, pour une valeur $\theta_0 \in \Theta$

- **Unilatéral gauche**: $\mathcal{H}_0 : \theta \geq \theta_0$ contre $\mathcal{H}_1 : \theta < \theta_0$
- **Unilatéral droit**: $\mathcal{H}_0 : \theta \leq \theta_0$ contre $\mathcal{H}_1 : \theta > \theta_0$
- **Bilatéral**: $\mathcal{H}_0 : \theta = \theta_0$ contre $\mathcal{H}_1 : \theta \neq \theta_0$

Théorème *(Equivalence des hypothèses nulles pour les tests unilatéraux (admis)).* ★

*Le **test unilatéral gauche** $\mathcal{H}_0 : \theta \geq \theta_0$ contre $\mathcal{H}_1 : \theta < \theta_0$ est "équivalent" au test:*

$$\mathcal{H}_0 : \theta = \theta_0 \text{ vs. } \mathcal{H}_1 : \theta < \theta_0$$

*Le **test unilatéral droit** $\mathcal{H}_0 : \theta \leq \theta_0$ contre $\mathcal{H}_1 : \theta > \theta_0$ est "équivalent" au test:*

$$\mathcal{H}_0 : \theta = \theta_0 \text{ vs. } \mathcal{H}_1 : \theta > \theta_0$$

Choix de $\mathcal{H}_1$: ☞ C'est l'hypothèse **que l'on cherche à ne pas rejeter!**

Ex. précédent: Si on s'attend à ce que la candidate gagne, $\mathcal{H}_1 : p > 0.5$, sinon $\mathcal{H}_1 : p < 0.5$.

Tests d'hypothèses - Règle de décision

Comment décider entre $\mathcal{H}_0$ et $\mathcal{H}_1$?

A l'aide d'un **échantillon** $\mathbf{X} = (X_1, \ldots, X_n)$ i.i.d. de loi populationnelle paramétrée par θ

Ex. précédent:

$\mathbf{X}$ est un échantillon i.i.d. de loi $\mathcal{B}(p)$, et $\overline{X}_n$ est un "bon" estimateur de p.

Pour $n = 1000$:

- Si on observe $\overline{x}_n = 0.98$, pensez-vous que la candidate soit susceptible de gagner?
- Si on observe $\overline{x}_n = 0.53$, pensez-vous que la candidate soit susceptible de gagner?
- Si on observe $\overline{x}_n = 0.11$, pensez-vous que la candidate soit susceptible de gagner?

☞✻ **A partir de quelle valeur p^* est-ce que l'on va commencer à rejeter $\mathcal{H}_0$?**

On appelle cette valeur **la valeur critique d'un test**

Elle permet de construire **la région de rejet du test** : $\mathcal{R} = [p^* \ ; \ 1]$

☞ **Si $\overline{X}_n$ est dans la région de rejet $\mathcal{R}$, alors on rejette $\mathcal{H}_0$**

Tests d'hypothèses - Valeur critique

Comment peut-on déterminer la valeur critique?

En contrôlant **le risque de rejeter $\mathcal{H}_0$ alors que $\mathcal{H}_0$ est vraie**:

$$\mathbb{P}(\text{rejeter } \mathcal{H}_0 \mid \mathcal{H}_0 \text{ est vraie}) = \alpha$$

On veut trouver θ^*, θ_I^* et θ_S^* tels que: ★

- **Unilatéral gauche**: $\mathcal{H}_0 : \theta = \theta_0 \quad \text{contre} \quad \mathcal{H}_1 : \theta < \theta_0$
$$\mathbb{P}\left(\widehat{\theta}(\mathbf{X}) \leq \theta^* \mid \theta = \theta_0\right) = \alpha \qquad \Longrightarrow \mathcal{R} = (-\infty\ ;\ \theta^*]$$
- **Unilatéral droit**: $\mathcal{H}_0 : \theta = \theta_0 \quad \text{contre} \quad \mathcal{H}_1 : \theta > \theta_0$
$$\mathbb{P}\left(\widehat{\theta}(\mathbf{X}) \geq \theta^* \mid \theta = \theta_0\right) = \alpha \qquad \Longrightarrow \mathcal{R} = [\theta^*\ ;\ \infty)$$
- **Bilatéral**: $\mathcal{H}_0 : \theta = \theta_0 \quad \text{contre} \quad \mathcal{H}_1 : \theta \neq \theta_0$
$$1 - \mathbb{P}\left(\theta_I^* \leq \widehat{\theta}(\mathbf{X}) \leq \theta_S^* \mid \theta = \theta_0\right) = \alpha \qquad \Longrightarrow \mathcal{R} = (-\infty\ ;\ \theta_I^*] \cup [\theta_S^*\ ;\ \infty)$$

☞ **Pour trouver θ^*, θ_I^* et θ_S^*, on a besoin de connaître (ou d'approcher) la loi des estimateurs!**

Tests d'hypothèses - Statistique de test

Une **statistique de test** $T(\mathbf{X})$ est **une fonction de l'échantillon**: **c'est une v.a.!**
☞ Aider à **trouver la valeur critique** et **la région de rejet** pour un test

Comment construire une statistique de test?

- Transformation (centrage-reduction, rapport...) **d'un estimateur de** θ
- Dont la loi **lorsque $\mathcal{H}_0$ est supposée vraie** soit simple et explicite.

Par exemple:
Cadre de travail: X i.i.d. **Gaussien** $\mathcal{N}(\mu, 9)$, μ inconnu
Test unilatéral droit: $\mathcal{H}_0 : \mu = 500$ contre $\mathcal{H}_1 : \mu > 500$

Si on suppose que $\mathcal{H}_0$ est vrai ($\mu = 500$), alors on sait que:

$$T(\mathbf{X}) = \sqrt{n}\left(\frac{\overline{X}_n - 500}{3}\right) \sim \mathcal{N}(0, 1)$$

☞ **C'est notre statistique de test!**

Tests d'hypothèses - Statistique de test

Et donc:

$$\mathbb{P}\left(\overline{X}_n \geq \mu^* \mid \mu = 500\right) = \alpha \iff \mathbb{P}\left(\sqrt{n}\left(\frac{\overline{X}_n - 500}{3}\right) \leq \sqrt{n}\left(\frac{\mu^* - 500}{3}\right) \mid \mu = 500\right) = 1 - \alpha$$

Mais comme sous $\mathcal{H}_0$, on sait que $\sqrt{n}\left(\frac{\overline{X}_n - 500}{3}\right) \sim \mathcal{N}(0, 1)$, alors:

$$z_{1-\alpha} = \sqrt{n}\left(\frac{\mu^* - 500}{3}\right) \quad \Longrightarrow \quad \mu^* = 500 + z_{1-\alpha}\frac{3}{\sqrt{n}}$$

Donc, par exemple, pour:

- Un échantillon de taille 10 ;
- Un niveau de risque $\alpha = 0.05$ de rejeter $\mathcal{H}_0$ à tort ($z_{1-\alpha} = 1.645$).

On **rejettera** $\mathcal{H}_0 : \mu \leq 500$ si:

$$\overline{X}_n \geq 500 + 1.645 \times \frac{3}{\sqrt{10}} \approx 501.56, \quad \text{et donc} \quad \mathcal{R} = [501.56\ ;\ \infty)$$

☞ **Avec un échantillon observé, on comparera la valeur critique à l'estimation $\overline{x}_n$**

Tests d'hypothèses - Statistique de test

En pratique, on va préférer **comparer la statistique de test plutôt que l'estimateur**.

On **rejette** $\mathcal{H}_0$ si

$$\overline{X}_n \geq 500 + z_{1-\alpha}\frac{3}{\sqrt{n}} \quad \Longleftrightarrow \quad T(\mathbf{X}) = \sqrt{n}\left(\frac{\overline{X}_n - 500}{3}\right) \geq z_{1-\alpha}$$

On peut ainsi construire **la région de rejet pour** $T(\mathbf{X})$: $\mathcal{R}_T = [z_{1-\alpha}; \infty)$

☞ **En pratique, avec un échantillon observé** $\mathbf{x}$**, on vérifie si** $T(\mathbf{x})$ **appartient à** $[t_\alpha; \infty)$

Tests d'hypothèses - Interprétation et double négation

Le raisonnement d'un test d'hypothèse repose sur **une démarche de réfutation**:

☞ On **suppose $\mathcal{H}_0$ vraie** et on **vérifie si les données sont compatibles avec cette hypothèse**

- Si les données sont très improbables sous $\mathcal{H}_0$, on la rejette (en faveur de $\mathcal{H}_1$)
- Sinon, on ne la rejette pas, faute de preuve suffisante.

☞ **Un test ne prouve jamais que $\mathcal{H}_0$ est vraie!**

Il cherche seulement à voir si on a assez d'éléments dans les données pour la rejeter

Il y a **deux situations possibles**:

- On rejette $\mathcal{H}_0$ en faveur de $\mathcal{H}_1$

  Ceci ne veut pas dire que $\mathcal{H}_1$ est vraie, elle est juste plus vraisemblable que $\mathcal{H}_0$
- On ne peux pas rejeter $\mathcal{H}_0$ en faveur de $\mathcal{H}_1$

  Ceci ne veut pas dire que $\mathcal{H}_0$ est vraie, elle est juste plus vraisemblable que $\mathcal{H}_1$

☞ **On "acceptera" jamais $\mathcal{H}_0$ ou $\mathcal{H}_1$**

☞ Le **niveau de risque d'un test d'hypothèse** s'interprète comme **le niveau de confiance d'un IC**

Tests d'hypothèses - Première illustration

Exemple *(Paquets de biscuits)*. Vous êtes contrôleur qualité dans une compagnie qui manufacture des biscuits. Vous devez montrer qu'en moyenne les paquets font **au moins 500g** avec une risque toléré de 5%. La machine de remplissage des paquets indique une variabilité de **3**g^2. On suppose que le poids des paquets suivent une loi normale $\mathcal{N}(\mu, 3)$.

A l'aide d'un échantillon **x** de taille 10, on observe $\overline{x}_n = 501.2$.

Test d'hypothèse: $\mathcal{H}_0 : \mu \leq 500,$ contre $\mathcal{H}_1 : \mu > 500$

Statistique de test: Sous $\mathcal{H}_0$, $T(\mathbf{X}) = \sqrt{10}\left(\frac{\overline{X}_n - 500}{\sqrt{3}}\right) \overset{\mathcal{H}_0}{\sim} \mathcal{N}(0,1)$, et donc

$$T(\mathbf{x}) \approx 2.19$$

Région critique pour la statistique de test: $z_{0.95} = \Phi^{-1}(0.95) = 1.645$, et donc

$$\mathcal{R}_T = [1.645 \ ; \ \infty)$$

☞ **Conclusion?** ✍

Notations et nomenclature

En général, on **nomme les tests en fonction de la loi de la statistique de test**

Loi de la statistique de test	Nom du test
Normale centrée réduite $\mathcal{N}(0,1)$	Test z (z-test)
Student $\mathcal{T}$	Test t (t-test)
Chi-Carré χ^2	Test du χ^2 (χ^2-test)
Fisher $\mathcal{F}$	Test F (F-test)

Des tests différents peuvent avoir une statistique de test avec la même loi

On verra des exemples plus tard

☞ **Il est PRIMORDIAL de toujours spécifier clairement $\mathcal{H}_1$!**

Tests d'hypothèses - Probabilité critique

On a vu **deux manières équivalentes** de choisir entre $\mathcal{H}_0$ et $\mathcal{H}_1$:

1. On compare un estimateur $\widehat{\theta}(\mathbf{X})$ à sa valeur critique θ^* en fonction de α
2. On compare une statistique de test $T(\mathbf{X})$ à un des quantiles de sa loi

Il existe **une troisième manière équivalente**: la **probabilité critique** (ou **p-valeur**)
C'est la manière la plus utilisée en pratique

☞ On sait que la valeur du risque α **influence la région de rejet**

Intuition: Quelle est le **niveau de risque le plus petit** qui permettrait quand même de rejeter $\mathcal{H}_0$?

C'est exactement **ce que quantifie la probabilité critique**.

Tests d'hypothèses - Probabilité critique

Définition *(Probabilité critique).* ★

La probabilité critique (ou p-valeur) est le plus petit niveau de risque α^* tel qu'on rejette $\mathcal{H}_0$. Pour un échantillon observé $\mathbf{x}$, elle se définit par:

- **Unilatéral gauche**: $\alpha^* = \mathbb{P}\left(\widehat{\theta}(\mathbf{X}) \leq \widehat{\theta}(\mathbf{x}) \mid \mathcal{H}_0 \text{ est vraie}\right)$
- **Unilatéral droit**: $\alpha^* = \mathbb{P}\left(\widehat{\theta}(\mathbf{X}) \geq \widehat{\theta}(\mathbf{x}) \mid \mathcal{H}_0 \text{ est vraie}\right)$
- **Bilatéral**: $\alpha^* = 2\min\left\{\mathbb{P}\left(\widehat{\theta}(\mathbf{X}) \leq \widehat{\theta}(\mathbf{x}) \mid \mathcal{H}_0 \text{ est vraie}\right), \mathbb{P}\left(\widehat{\theta}(\mathbf{X}) \geq \widehat{\theta}(\mathbf{x}) \mid \mathcal{H}_0 \text{ est vraie}\right)\right\}$

Théorème *(Probabilité critique et statistique de test (admis)).* ★

Pour un test d'hypothèse avec statistique de test $T(\mathbf{X})$, et échantillon observé $\mathbf{x}$:

- ***Unilatéral gauche***: $\alpha^* = \mathbb{P}\left(T(\mathbf{X}) \leq T(\mathbf{x}) \mid \theta = \theta_0\right)$
- ***Unilatéral droit***: $\alpha^* = \mathbb{P}\left(T(\mathbf{X}) \geq T(\mathbf{x}) \mid \theta = \theta_0\right)$
- ***Bilatéral***: $\alpha^* = 2\min\left\{\mathbb{P}\left(T(\mathbf{X}) \leq T(\mathbf{x}) \mid \theta = \theta_0\right), \mathbb{P}\left(T(\mathbf{X}) \geq T(\mathbf{x}) \mid \theta = \theta_0\right)\right\}$ *ou* $\mathbb{P}\left(-|T(\mathbf{x})| \leq T(\mathbf{X}) \leq |T(\mathbf{x})| \mid \theta = \theta_0\right)$ *si la loi de $T(\mathbf{X})$ est symétrique autour de* 0.

☞ **Pour un risque α, si $\alpha^* < \alpha$ alors on rejette $\mathcal{H}_0$**

Exemple *(Paquets de biscuits)*. Vous êtes contrôleur qualité dans une compagnie qui manufacture des biscuits. Vous devez montrer qu'en moyenne les paquets font **au moins 500g** avec une risque toléré de 5%. La machine de remplissage des paquets indique une variabilité de **3**g^2. On suppose que le poids des paquets suivent une loi normale $\mathcal{N}(\mu, 3)$.

A l'aide d'un échantillon **x** de taille 10, on observe $\overline{x}_n = 501.2$.

Test d'hypothèse: $\mathcal{H}_0 : \mu \leq 500, \quad$ contre $\quad \mathcal{H}_1 : \mu > 500$

Statistique de test: Sous $\mathcal{H}_0$, $T(\mathbf{X}) = \sqrt{10}\left(\frac{\overline{X}_n - 500}{\sqrt{3}}\right) \underset{\mathcal{H}_0}{\sim} \mathcal{N}(0,1)$, et $T(\mathbf{x}) \approx 2.19$

Calcul de la probabilité critique:

$$\mathbb{P}\left(T(\mathbf{X}) \geq 2.19 \mid \mu = 500\right) \underset{\mathcal{H}_0}{=} 1 - \Phi(2.19) \approx 0.014$$

☞ **Conclusion?** ✍

Tests d’hypothèses - Recette

Pour **utiliser un test d’hypothèse sur un paramètre de la loi de la population**: ★

1. Formuler l’hypothèse nulle $\mathcal{H}_0$ et son alternative $\mathcal{H}_1$
2. Choisir un niveau de risque
3. Identifier le type du test (unilatéral, bilatéral...)
4. Choisir un **estimateur du paramètre**
5. En déduire **une statistique de test et sa loi sous** $\mathcal{H}_0$
6. **Rejeter ou non l’hypothèse nulle** selon l’une des **trois manières équivalentes**:
   - L’estimation appartient à sa **la région de rejet**, construite avec **la valeur critique**
   - La statistique de test observée appartient à sa **la région de rejet**, construite avec **un quantile de sa loi sous** $\mathcal{H}_0$
   - La p-valeur du test est plus petite que le niveau de risque initialement choisi

Tests d'hypothèses - Deux types de risques

Deux types de risque de se tromper pour un test statistique: ★

- **Le risque de première espèce α (Type I)**: $\alpha = \mathbb{P}\,(\text{Rejeter } \mathcal{H}_0 \mid \mathcal{H}_0 \text{ est vraie})$
  ☞ C'est cette erreur que l'on contrôle pour construire la **région de rejet**
- **Le risque de seconde espèce β (Type II)**: $\beta = \mathbb{P}\,(\text{Ne pas rejeter } \mathcal{H}_0 \mid \mathcal{H}_1 \text{ est vraie})$

À partir de ces deux risques, on peut construire deux notions supplémentaires: ★

- Le **niveau de confiance d'un test statistique**: $1 - \alpha$
- La **puissance d'un test statistique**: $1 - \beta$

	H_0 **vraie**	H_1 **vraie**
Pas rejet de H_0	**Niveau de confiance** $(1-\alpha)$	**Risque de type II** (β)
Rejet de H_0	**Risque de type I** (α)	**Puissance** $(1-\beta)$

☞ Le risque α permet de **construire la règle de décision** (sous $\mathcal{H}_0$)
☞ La puissance $1-\beta$ mesure **la qualité du test** (sous $\mathcal{H}_1$)

On peut contrôler la puissance pour **calculer une taille d'échantillon requise**

Cas d'étude: Échantillon **X** i.i.d. Gaussien de loi $\mathcal{N}(\mu, \sigma^2)$, où σ^2 est connu

$$\mathcal{H}_0 : \mu = \mu_0 \quad \text{contre} \quad \mathcal{H}_1 : \mu > \mu_0$$

On choisit une valeur d'intérêt $\mu_1 > \mu_0$.

Sous $\mathcal{H}_0$:

$$\overline{X}_n \sim \mathcal{N}(\mu_0, \sigma^2/n)$$

$$T_0(\mathbf{X}) = \sqrt{n}\left(\frac{\overline{X}_n - \mu_0}{\sigma}\right) \sim \mathcal{N}(0, 1)$$

Sous $\mu = \mu_1$:

$$\overline{X}_n \sim \mathcal{N}\left(\mu_1, \sigma^2/n\right)$$

$$T_1(\mathbf{X}) = \sqrt{n}\left(\frac{\overline{X}_n - \mu_1}{\sigma}\right) \sim \mathcal{N}(0, 1)$$

Tests d'hypothèses - Puissance et taille échantillonnale

- **Étape 1:** Région de rejet
  On rejette $\mathcal{H}_0$ si:
  $$\overline{X}_n - \mu_0 > z_{1-\alpha} \frac{\sigma}{\sqrt{n}}$$
- **Étape 2:** Région de rejet en fonction de $T_1(\mathbf{X})$
  De façon équivalente, on peut écrirel a région de rejet:
  $$T_1(\mathbf{X}) > \frac{\mu_0 - \mu_1}{\sigma/\sqrt{n}} + z_{1-\alpha} \text{ ✍}$$
- **Étape 3:** Calcul de la puissance
  On pose $\Lambda = \sqrt{n}\left(\frac{\mu_1 - \mu_0}{\sigma}\right)$ et on a que:
  $$1 - \beta = \mathbb{P}(T_1(\mathbf{X}) > z_{1-\alpha} - \Lambda \mid \mu = \mu_1) = 1 - \Phi(z_{1-\alpha} - \Lambda)$$

$$\iff \beta = \Phi(z_{1-\alpha} - \Lambda) \iff z_\beta = z_{1-\alpha} - \Lambda \iff \boxed{\Lambda = z_{1-\alpha} + z_{1-\beta}}$$

☞ Ce qui implique que $\boxed{n^* = \sigma^2 \left(\frac{z_{1-\alpha} + z_{1-\beta}}{\mu_1 - \mu_0}\right)^2}$

Tests d'hypothèses - Puissance et taille échantillonnale

Exemple *(Temps de réponse d'un serveur web)*. Une compagnie technologique exploite une plateforme de commerce en ligne. Pour garantir une bonne expérience utilisateur, les ingénieurs surveillent en continu le **temps de réponse moyen** du serveur principal lors du chargement de la page d'accueil.

Un audit interne impose que le temps de réponse moyen ne dépasse pas **200 ms**. Après une mise à jour logicielle censée améliorer les performances, l'équipe souhaite vérifier si la moyenne du temps de réponse a **diminué**.

Les données historiques montrent que la variabilité journalière du temps de réponse est $\sigma = 25$ ms. De plus, les temps de réponse journaliers sont supposées suivre une loi normale.

On souhaite réaliser un test unilatéral :

$$\mathcal{H}_0 : \mu = \mu_0 = 200 \qquad \text{contre} \qquad \mathcal{H}_1 : \mu < \mu_0.$$

L'équipe fixe l'**effet minimal détectable** à une amélioration d'au moins $\mu_0 - \mu_1 = 10$ ms (c-à-d que $\mu_1 = 190$ ms).

Le niveau de test est fixé à $\alpha = 0{,}05$, et la **puissance souhaitée** du test est $1 - \beta = 0{,}90$

☞ Quelle est la taille échantillonnale minimale afin de faire un test d'hypothèse satisfaisant? ✍

Sommaire

Test sur la moyenne d'un échantillon Gaussien, variance connue

Cadre de travail: $\mathbf{X} = (X_1, \ldots, X_n)$ i.i.d. de loi $\mathcal{N}(\mu, \sigma^2)$, σ^2 **connue**, $\mathbf{x}$ observé, Φ f.d.r. normale standard

Hypothèse du test statistique:

$$\mathcal{H}_0 : \mu = \mu_0 \quad \text{contre} \quad \mathcal{H}_1 = \begin{cases} \mu < \mu_0 & \text{(uni. gauche)} \\ \mu \neq \mu_0 & \text{(bilatéral)} \\ \mu > \mu_0 & \text{(uni. droit)} \end{cases}$$

Statistique de test et sa loi sous $\mathcal{H}_0$:

$$T(\mathbf{X}) = \sqrt{n} \left(\frac{\overline{X}_n - \mu_0}{\sigma} \right) \overset{\mathcal{H}_0}{\sim} \mathcal{N}(0, 1)$$

Rejet de $\mathcal{H}_0$ au risque α si:

$$\begin{cases} \overline{x}_n < \mu_0 - z_{1-\alpha} \frac{\sigma}{\sqrt{n}} \\ \overline{x}_n \notin \left[\mu_0 \pm z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \right] \\ \overline{x}_n > \mu_0 + z_{1-\alpha} \frac{\sigma}{\sqrt{n}} \end{cases} \quad \text{ou si} \quad \begin{cases} T(\mathbf{x}) < -z_{1-\alpha} \\ |T(\mathbf{x})| > z_{1-\alpha/2} \\ T(\mathbf{x}) > z_{1-\alpha} \end{cases} \quad \text{ou si} \quad \begin{cases} \alpha^* = \Phi(T(\mathbf{x})) < \alpha \\ \alpha^* = 2\left(1 - \Phi\left(|T(\mathbf{x})|\right)\right) < \alpha \\ \alpha^* = 1 - \Phi\left(T(\mathbf{x})\right) < \alpha \end{cases}$$

Tests d'hypothèses - Illustration

Exemple *(Contrôle de la teneur en caféine)*.

Un fabricant de boissons énergisantes affirme que sa canette contient en moyenne $\mu_0 = 80$mg de caféine.

Pour des raisons réglementaires, l'autorité de santé veut vérifier si la moyenne réelle est bien conforme à cette valeur, ou si elle est différente (trop faible ou trop élevée). L'autorité de santé est prête à accepter un risque d'erreur de type I de 1%

On prélève un échantillon de $n = 17$ canettes au hasard dans la production pour lesquelles on mesure précisément la quantité de caféine qu'elles contiennent.

L'historique du procédé de production et du labo de mesure indique que, la teneur en caféine de chaque canette se modélise par une loi normale de variance $\sigma^2 = 9$.

On observe une teneur en caféine moyenne $\overline{x}_n = 78.9$mg.

☞ Est-ce que la teneur en caféine est réglementaire? ✍

Test sur la moyenne d'un échantillon Gaussien, variance inconnue

Cadre de travail: $\mathbf{X} = (X_1, \dots, X_n)$ i.i.d. de loi $\mathcal{N}(\mu, \sigma^2)$, σ^2 **inconnue**, $\mathbf{x}$ observé

Hypothèse du test statistique:

$$\mathcal{H}_0 : \mu = \mu_0 \quad \text{contre} \quad \mathcal{H}_1 = \begin{cases} \mu < \mu_0 & \text{(uni. gauche)} \\ \mu \neq \mu_0 & \text{(bilatéral)} \\ \mu > \mu_0 & \text{(uni. droit)} \end{cases}$$

Statistique de test et sa loi sous $\mathcal{H}_0$:

$$T(\mathbf{X}) = \sqrt{n}\left(\frac{\overline{X}_n - \mu_0}{S}\right) \overset{\mathcal{H}_0}{\sim} \mathcal{T}(n-1), \quad t_{1-\alpha} := t_{1-\alpha,n-1}, \quad F_T(x) \text{ f.d.r. d'une } \mathcal{T}(n-1)$$

Rejet de $\mathcal{H}_0$ au risque α si:

$$\begin{cases} \overline{x}_n < \mu_0 - t_{1-\alpha}\frac{S}{\sqrt{n}} \\ \overline{x}_n \notin \left[\mu_0 \pm t_{1-\alpha/2}\frac{S}{\sqrt{n}}\right] \\ \overline{x}_n > \mu_0 + t_{1-\alpha}\frac{S}{\sqrt{n}} \end{cases} \quad \text{ou si} \quad \begin{cases} T(\mathbf{x}) < -t_{1-\alpha} \\ |T(\mathbf{x})| > t_{1-\alpha/2} \\ T(\mathbf{x}) > t_{1-\alpha} \end{cases} \quad \text{ou si} \quad \begin{cases} \alpha^* = F_T(T(\mathbf{x})) < \alpha \\ \alpha^* = 2\left(1 - F_T\left(|T(\mathbf{x})|\right)\right) < \alpha \\ \alpha^* = 1 - F_T\left(T(\mathbf{x})\right) < \alpha \end{cases}$$

Tests d'hypothèses - Illustration

Exemple *(Contrôle de la concentration en oxygène d'une salle d'opération)*.

Une équipe hospitalière doit s'assurer que la concentration minimale en oxygène dans une salle d'opération est d'au moins $\mu_0 = 21\%$ pour respecter les normes de sécurité.

Après une panne électrique, les ingénieurs biomédicaux veulent vérifier si la moyenne réelle n'est pas descendue en dessous de cette valeur minimale. Ils acceptent un risque d'erreur de type I de 5%.

On prélève un échantillon de $n = 17$ mesures instantanées de concentration d'oxygène durant une opération. Les données sont historiquement bien modélisées par une loi normale.

Les mesures donnent une moyenne empirique de $\overline{x}_n = 20.6\%$ et un écart-type empirique $s_n = 0.7\%$.

☞ Le niveau d'oxygène respecte-t-il la norme minimale? ✍

Échantillon Gaussien - Test et IC bilatéral sur la moyenne

Théorème *(Lien entre test et IC pour la moyenne).* ★

Soit **X** *i.i.d. de loi* $\mathcal{N}(\mu, \sigma^2)$, μ *inconnue* ***et de variance connue ou non***.

Soit $IC_{1-\alpha}(\mathbf{X})$ *un IC pour* μ *au niveau* $1-\alpha$.

Alors, on a les équivalences suivantes au risque α*:*

$$\text{Rejeter } \mathcal{H}_0 : \mu = \mu_0 \quad \Longleftrightarrow \quad \text{Ne pas rejeter } \mathcal{H}_1 : \begin{cases} \mu < \mu_0 & \text{(uni. gauche)} \\ \mu \neq \mu_0 & \text{(bilatéral)} \\ \mu > \mu_0 & \text{(uni. droit)} \end{cases}$$

$$\Longleftrightarrow \quad \mu_0 \notin \begin{cases} IC_{1-\alpha}(\mathbf{x}) \text{ uni. gauche} \\ IC_{1-\alpha}(\mathbf{x}) \text{ bilatéral} \\ IC_{1-\alpha}(\mathbf{x}) \text{ uni. droit} \end{cases}$$

Test sur la variance d'un échantillon Gaussien, moyenne connue

Cadre de travail: $\mathbf{X} = (X_1, \ldots, X_n)$ i.i.d. de loi $\mathcal{N}(\mu, \sigma^2)$, μ **connue**, $\mathbf{x}$ observé, $S_0^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \mu)^2$

Hypothèse du test statistique:

$$\mathcal{H}_0 : \sigma^2 = \sigma_0^2 \quad \text{contre} \quad \mathcal{H}_1 = \begin{cases} \sigma^2 < \sigma_0^2 & \text{(uni. gauche)} \\ \sigma^2 \neq \sigma_0^2 & \text{(bilatéral)} \\ \sigma^2 > \sigma_0^2 & \text{(uni. droit)} \end{cases}$$

Statistique de test et sa loi sous $\mathcal{H}_0$:

$$T(\mathbf{X}) = \frac{n\,S_0^2}{\sigma_0^2} \overset{\mathcal{H}_0}{\sim} \chi^2(n), \quad c_{1-\alpha} := c_{1-\alpha,n}, \quad F_T(x) \text{ f.d.r. de } \chi^2(n)$$

Rejet de $\mathcal{H}_0$ au risque α si: $\mathcal{R} = \left[\frac{\sigma_0^2}{n}\,c_{\alpha/2}\,;\,\frac{\sigma_0^2}{n}\,c_{1-\alpha/2}\right]$, $\mathcal{R}_T = \left[c_{\alpha/2}\,;\,c_{1-\alpha/2}\right]$

$$\begin{cases} s_0^2 < \frac{\sigma_0^2}{n}\,c_\alpha \\ s_0^2 \notin \mathcal{R} \\ s_0^2 > \frac{\sigma_0^2}{n}\,c_{1-\alpha} \end{cases} \quad \text{ou si} \quad \begin{cases} T(\mathbf{x}) < c_\alpha \\ T(\mathbf{x}) \notin \mathcal{R}_T \\ T(\mathbf{x}) > c_{1-\alpha} \end{cases} \quad \text{ou si} \quad \begin{cases} \alpha^* = F_T\big(T(\mathbf{x})\big) < \alpha \\ \alpha^* = 2\,\min\Big\{F_T\big(T(\mathbf{x})\big),\ 1 - F_T\big(T(\mathbf{x})\big)\Big\} < \alpha \\ \alpha^* = 1 - F_T\big(T(\mathbf{x})\big) < \alpha \end{cases}$$

Tests d'hypothèses - Illustration

Exemple *(Stabilité de la température d'une chaîne du froid)*.

Une entreprise pharmaceutique transporte des vaccins sensibles. Le protocole impose que la variabilité de la température dans les camions réfrigérés ne dépasse pas $\sigma_0^2 = 1.5$ (en °C^2), afin de limiter les risques de dégradation du produit.

Après une mise à jour des systèmes réfrigérants, l'équipe qualité veut vérifier si la variance réelle n'est pas devenue plus grande que cette valeur. Elle accepte un risque d'erreur de type I de 10%.

On enregistre $n = 22$ températures au cours d'un transport. Les capteurs sont précis et la température intérieure est bien modélisée par une loi normale. La température moyenne cible est considérée comme connue : $\mu = 5°$C.

À partir de ces mesures, l'équipe fait une estimation $s_0^2 = 2.1$.

☞ La variabilité de la température est-elle devenue supérieure au seuil acceptable? ✍

Test sur la variance d'un échantillon Gaussien, moyenne inconnue

Cadre de travail: $\mathbf{X} = (X_1, \ldots, X_n)$ i.i.d. de loi $\mathcal{N}(\mu, \sigma^2)$, μ **inconnue**, $\mathbf{x}$ observé, $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \overline{X}_n)^2$

Hypothèse du test statistique:

$$\mathcal{H}_0 : \sigma^2 = \sigma_0^2 \quad \text{contre} \quad \mathcal{H}_1 = \begin{cases} \sigma^2 < \sigma_0^2 & \text{(uni. gauche)} \\ \sigma^2 \neq \sigma_0^2 & \text{(bilatéral)} \\ \sigma^2 > \sigma_0^2 & \text{(uni. droit)} \end{cases}$$

Statistique de test et sa loi sous $\mathcal{H}_0$:

$$T(\mathbf{X}) = \frac{(n-1)\, S^2}{\sigma_0^2} \overset{\mathcal{H}_0}{\sim} \chi^2(n-1), \quad c_{1-\alpha} := c_{1-\alpha,\, n-1}, \quad F_T(x) \text{ f.d.r. de } \chi^2(n-1)$$

Rejet de $\mathcal{H}_0$ au risque α si: $\mathcal{R} = \left[\frac{\sigma_0^2}{n-1}\, c_{\alpha/2} \,;\, \frac{\sigma_0^2}{n-1}\, c_{1-\alpha/2} \right], \quad \mathcal{R}_T = \left[c_{\alpha/2} \,;\, c_{1-\alpha/2} \right]$

$$\begin{cases} s^2 < \frac{\sigma_0^2}{n-1}\, c_\alpha \\ s^2 \notin \mathcal{R} \\ s^2 > \frac{\sigma_0^2}{n-1}\, c_{1-\alpha} \end{cases} \quad \text{ou si} \quad \begin{cases} T(\mathbf{x}) < c_\alpha \\ T(\mathbf{x}) \notin \mathcal{R}_T \\ T(\mathbf{x}) > c_{1-\alpha} \end{cases} \quad \text{ou si} \quad \begin{cases} \alpha^* = F_T\big(T(\mathbf{x})\big) < \alpha \\ \alpha^* = 2 \min\left\{F_T\left(T(\mathbf{x})\right),\ 1 - F_T\left(T(\mathbf{x})\right)\right\} < \alpha \\ \alpha^* = 1 - F_T\big(T(\mathbf{x})\big) < \alpha \end{cases}$$

Exemple *(Volatilité des montants de sinistres en assurance automobile)*.

Une compagnie d'assurance surveille la variance des montants de sinistres liés aux accidents légers. Le service actuariel estime historiquement que la variance réelle est de $\sigma_0^2 = 1\,200$ (en dollars^2).

Après un changement de politique de remboursement, l'équipe de gestion des risques veut vérifier si la variance a changé, à la hausse ou à la baisse. Elle est prête à accepter un risque d'erreur de type I de 10%.

On prélève un échantillon aléatoire de $n = 25$ sinistres récents. Les montants de réparations pour ce type d'accident sont bien modélisés par une loi normale.

À partir de ces observations, l'assureur estimate la variance par : $s^2 = 1\,430$.

☞ Peut-on conclure que la variabilité des sinistres a changé depuis la nouvelle politique ? ✍

Échantillon Gaussien - Test et IC bilatéral sur la variance

Théorème *(Lien entre test et IC pour la variance).* ★

Soit $\mathbf{X}$ *i.i.d. de loi* $\mathcal{N}(\mu, \sigma^2)$, σ^2 *inconnue* ***et de moyenne connue ou non***.

Soit $IC_{1-\alpha}(\mathbf{X})$ *un IC pour* σ^2 *au niveau* $1-\alpha$.

Alors, on a les équivalences suivantes au risque α*:*

$$\text{Rejeter } \mathcal{H}_0 : \sigma^2 = \sigma_0^2 \quad \Longleftrightarrow \quad \text{Ne pas rejeter } \mathcal{H}_1 : \begin{cases} \sigma^2 < \sigma_0^2 & \text{(uni. gauche)} \\ \sigma^2 \neq \sigma_0^2 & \text{(bilatéral)} \\ \sigma^2 > \sigma_0^2 & \text{(uni. droit)} \end{cases}$$

$$\Longleftrightarrow \quad \sigma_0^2 \notin \begin{cases} IC_{1-\alpha}(\mathbf{x}) \text{ uni. gauche} \\ IC_{1-\alpha}(\mathbf{x}) \text{ bilatéral} \\ IC_{1-\alpha}(\mathbf{x}) \text{ uni. droit} \end{cases}$$

Sommaire

Deux échantillons Gaussiens

Jusqu'à présent, nous n'avions accès **qu'à un seul échantillon i.i.d.**
Ça nous a permis de construire des tests statistiques sur un paramètre

Comment faire si l'on veut comparer **deux populations** différentes?

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 μ_X, et le **temps de fabrication par la méthode B** par une Gaussienne d'espérance μ_Y.

Quelle méthode de fabrication est la plus efficace?

Exemple *(Efficacité d'un médicament)*. On cherche à étudier les effets d'un médicament. On mesure la température de patients fiévreux **avant la prise du médicament** et **3h après la prise du médicament.**

Est-ce que ce médicament est efficace contre la fièvre?

Rappels - Deux configurations

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

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

On dit que:

- $\mathbf{X}$ et $\mathbf{Y}$ sont **indépendants** si, pour $i = 1, \ldots, n_X$ et $j = 1, \ldots, n_Y$: $X_i \perp\!\!\!\perp Y_j$ ★
- $\mathbf{X}$ et $\mathbf{Y}$ 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 $i \neq j$.

Tests d'hypothèses - Deux échantillons Gaussiens

Test d'égalité des moyennes pour éch. indep., variances connues

Cadre de travail: $\mathbf{X}, \mathbf{Y}$ indépendants de **variances σ_X^2, σ_Y^2 connues**, $\mathbf{x}, \mathbf{y}$ observés

Hypothèse du test statistique:

$$\mathcal{H}_0 : \mu_X = \mu_Y \quad \text{contre} \quad \mathcal{H}_1 = \begin{cases} \mu_X < \mu_Y & \text{(uni. gauche)} \\ \mu_X \neq \mu_Y & \text{(bilatéral)} \\ \mu_X > \mu_Y & \text{(uni. droit)} \end{cases}$$

Statistique de test et sa loi sous $\mathcal{H}_0$:

$$T(\mathbf{X}, \mathbf{Y}) = \frac{(\overline{X}_n - \overline{Y}_n)}{\sqrt{\sigma_X^2/n_X + \sigma_Y^2/n_Y}} \overset{\mathcal{H}_0}{\sim} \mathcal{N}(0,1), \quad z_{1-\alpha} := \Phi^{-1}(1-\alpha)$$

Rejet de $\mathcal{H}_0$ au risque α si:

$$\begin{cases} \overline{x}_n - \overline{y}_n < -z_{1-\alpha}\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}} \\ \overline{x}_n - \overline{y}_n \notin \left[\pm z_{1-\alpha/2}\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}\right] \\ \overline{x}_n - \overline{y}_n > z_{1-\alpha}\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}} \end{cases} \quad \text{ou si} \quad \begin{cases} T(\mathbf{x}, \mathbf{y}) < -z_{1-\alpha} \\ |T(\mathbf{x}, \mathbf{y})| > z_{1-\alpha/2} \\ T(\mathbf{x}, \mathbf{y}) > z_{1-\alpha} \end{cases} \quad \text{ou si} \quad \begin{cases} \alpha^* = \Phi(T(\mathbf{x}, \mathbf{y})) < \alpha \\ \alpha^* = 2(1 - \Phi(|T(\mathbf{x}, \mathbf{y})|)) < \alpha \\ \alpha^* = 1 - \Phi(T(\mathbf{x}, \mathbf{y})) < \alpha \end{cases}$$

Tests d'hypothèses - Illustration

Exemple *(Comparaison de deux méthodes d'entraînement en course à pied)*.

Un club d'athlétisme teste deux programmes d'entraînement pour améliorer la vitesse maximale de ses sprinteurs. L'objectif est de déterminer si les deux méthodes donnent en moyenne les mêmes performances.

Dix athlètes suivent la méthode A et douze suivent la méthode B.

Les vitesses maximales (en km/h) mesurées en fin de programme peuvent être modélisées par deux lois normales, indépendantes, et dont les variances sont historiquement **connues**: $\sigma_A^2 = 1.2$ et $\sigma_B^2 = 1.5$.

Les moyennes estimées pour chaque méthode sont: $\overline{x}_A = 32.4$ et $\overline{x}_B = 31.1$.

Le club accepte un risque d'erreur de type I de 10% pour décider entre les deux méthodes

☞ Les deux programmes d'entraînement ont-ils la même efficacité moyenne? ✍

Test d'égalité des moyennes pour éch. indep., variances inconnues égales

Cadre de travail: $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}$, **variances inconnues égales à** σ^2, $\mathbf{x}, \mathbf{y}$ observés $S_p^2 = \frac{(n_X-1)S_X^2+(n_Y-1)S_Y^2}{n_X+n_Y-2}$,

Hypothèse du test statistique:

$$\mathcal{H}_0 : \mu_X = \mu_Y \quad \text{contre} \quad \mathcal{H}_1 = \begin{cases} \mu_X < \mu_Y & \text{(uni. gauche)} \\ \mu_X \neq \mu_Y & \text{(bilatéral)} \\ \mu_X > \mu_Y & \text{(uni. droit)} \end{cases}$$

Statistique de test et sa loi sous $\mathcal{H}_0$:

$$T(\mathbf{X}, \mathbf{Y}) = \frac{\overline{X}_n - \overline{Y}_n}{S_p\sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}} \overset{\mathcal{H}_0}{\sim} \mathcal{T}(n_X + n_Y - 2), \quad t_{1-\alpha} := t_{1-\alpha,\, n_X+n_Y-2}, \quad F_T(x) \text{ f.d.r. de } \mathcal{T}(n_X + n_Y - 2)$$

Rejet de $\mathcal{H}_0$ au risque α si:

$$\begin{cases} \overline{x}_n - \overline{y}_n < -t_{1-\alpha}\, S_p\sqrt{\frac{1}{n_X} + \frac{1}{n_Y}} \\ \overline{x}_n - \overline{y}_n \notin \left[\pm t_{1-\alpha/2}\, S_p\sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}\right] \\ \overline{x}_n - \overline{y}_n > t_{1-\alpha}\, S_p\sqrt{\frac{1}{n_X} + \frac{1}{n_Y}} \end{cases} \text{ ou si } \begin{cases} T(\mathbf{x}, \mathbf{y}) < -t_{1-\alpha} \\ |T(\mathbf{x}, \mathbf{y})| > t_{1-\alpha/2} \\ T(\mathbf{x}, \mathbf{y}) > t_{1-\alpha} \end{cases} \text{ ou si } \begin{cases} \alpha^* = F_T(T(\mathbf{x}, \mathbf{y})) < \alpha \\ \alpha^* = 2\,(1 - F_T(|T(\mathbf{x}, \mathbf{y})|)) < \alpha \\ \alpha^* = 1 - F_T(T(\mathbf{x}, \mathbf{y})) < \alpha \end{cases}$$

Tests d'hypothèses - Illustration

Exemple *(Influence du médiator sur la vitesse de jeu à la guitare)*.

Une école de musique cherche à savoir si le type de plectre (médiator) utilisé par les guitaristes influence leur vitesse de jeu (notes par seconde).

Deux groupes indépendants sont formés :

- 15 guitaristes professionnels utilisent un **plectre souple** (type S),
- 18 guitaristes professionnels utilisent un **plectre rigide** (type R).

Les vitesses mesurées sont bien modélisées par deux lois normales indépendantes, mais les variances réelles sont **inconnues**.

On observe les moyennes et écarts-types suivants :

$$\overline{x}_S = 6.8, \quad s_S = 0.9 \qquad\qquad \overline{x}_R = 7.4, \quad s_R = 1.1$$

L'école est prête à accepter un risque d'erreur de type I de 5% avant d'investir dans l'un ou l'autre des types de plectre.

☞ Les deux types de médiators conduisent-ils à la même vitesse moyenne de jeu? ✍

Test d'égalité des moyennes pour éch. appariés, variance inconnue

Cadre de travail: $\mathbf{X}, \mathbf{Y}$ appariés, $D_i = (X_i, Y_i)$, et $\mathbf{D}$ de loi $\mathcal{N}(\mu_D = \mu_X - \mu_Y, \sigma_D^2)$, σ_D^2 **inconnu**, $\mathbf{d}$ observé

Hypothèse du test statistique:

$$\mathcal{H}_0 : \mu_D = 0 \iff \mathcal{H}_0 : \mu_X = \mu_Y \quad \text{contre} \quad \mathcal{H}_1 = \begin{cases} \mu_D < 0 \iff \mu_X < \mu_Y & \text{(uni. gauche)} \\ \mu_D \neq 0 \iff \mu_X \neq \mu_Y & \text{(bilatéral)} \\ \mu_D > 0 \iff \mu_X > \mu_Y & \text{(uni. droit)} \end{cases}$$

Statistique de test et sa loi sous $\mathcal{H}_0$:

$$T(\mathbf{D}) = \frac{\overline{D}_n}{S_D/\sqrt{n}} \overset{\mathcal{H}_0}{\sim} \mathcal{T}(n-1), \quad t_{1-\alpha} := t_{1-\alpha,\, n-1}, \quad F_T(x) \text{ f.d.r. de } \mathcal{T}(n-1)$$

Rejet de $\mathcal{H}_0$ au risque α si:

$$\begin{cases} \overline{d}_n < -t_{1-\alpha} \dfrac{S_D}{\sqrt{n}} \\ \overline{d}_n \notin \left[\pm t_{1-\alpha/2} \dfrac{S_D}{\sqrt{n}}\right] \\ \overline{d}_n > t_{1-\alpha} \dfrac{S_D}{\sqrt{n}} \end{cases} \quad \text{ou si} \quad \begin{cases} T(\mathbf{d}) < -t_{1-\alpha} \\ |T(\mathbf{d})| > t_{1-\alpha/2} \\ T(\mathbf{d}) > t_{1-\alpha} \end{cases} \quad \text{ou si} \quad \begin{cases} \alpha^* = F_T(T(\mathbf{d})) < \alpha \\ \alpha^* = 2\,(1 - F_T(|T(\mathbf{d})|)) < \alpha \\ \alpha^* = 1 - F_T(T(\mathbf{d})) < \alpha \end{cases}$$

Exemple *(Comparaison d'une recette classique et allégée en sucre)*.

Une pâtisserie souhaite proposer une version allégée en sucre de son gâteau le plus vendu, sans diminuer le niveau moyen de satisfaction de la clientèle.

Un panel de $n = 18$ clients est invité. Chaque personne goûte la recette classique, puis et la recette allégée, dans un ordre aléatoire, puis note sa satisfaction.

Les différences de notes (recette allégée − recette classique) sont bien modélisées par une loi normale, mais la variance réelle de ces différences est **inconnue**.

À partir des 16 paires de notes, on calcule la moyenne et l'écart-type empirique des différences :

$$\overline{d}_n = -0.2 \quad \text{et} \quad s_d = 0.6.$$

La pâtisserie accepte un risque d'erreur de type I de 1% pour tester si la recette allégée offre le même niveau de satisfaction.

☞ La recette allégée offre-t-elle le même niveau moyen de satisfaction que la recette classique? ✍

Échantillon Gaussien - Test et IC bilatéral sur la différence de moyennes

Théorème *(Lien entre test et IC pour la différence de moyenne).* ★

Soit **X** *et* **Y** *i.i.d. de loi* $\mathcal{N}(\mu, \sigma^2)$, μ *inconnue* ***appariés ou indépendants, et de variance connue ou non***.

Soit $IC_{1-\alpha}(\mathbf{X})$ *un IC pour* $\mu_X - \mu_Y$ *au niveau* $1-\alpha$.

Alors, on a les équivalences suivantes au risque α*:*

$$\text{Rejeter } \mathcal{H}_0 : \mu_X = \mu_Y \quad \Longleftrightarrow \quad \text{Ne pas rejeter } \mathcal{H}_1 : \begin{cases} \mu < \mu_0 & \text{(uni. gauche)} \\ \mu \neq \mu_0 & \text{(bilatéral)} \\ \mu > \mu_0 & \text{(uni. droit)} \end{cases}$$

$$\Longleftrightarrow \quad 0 \notin \begin{cases} IC_{1-\alpha}(\mathbf{x}) \text{ uni. gauche} \\ IC_{1-\alpha}(\mathbf{x}) \text{ bilatéral} \\ IC_{1-\alpha}(\mathbf{x}) \text{ uni. droit} \end{cases}$$

Test de ratio de variances pour éch. indep., paramètres inconnus

Cadre de travail: $\mathbf{X} \perp\!\!\!\perp \mathbf{Y}$, $\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2$ **inconnus**, $\mathbf{x}, \mathbf{y}$ observés.

Hypothèse du test statistique:

$$\mathcal{H}_0 : \sigma_X^2 = \sigma_Y^2 \iff \frac{\sigma_X}{\sigma_Y} = 1 \quad \text{contre} \quad \mathcal{H}_1 = \begin{cases} \sigma_X^2 < \sigma_Y^2 & \text{(uni. gauche)} \\ \sigma_X^2 \neq \sigma_Y^2 & \text{(bilatéral)} \\ \sigma_X^2 > \sigma_Y^2 & \text{(uni. droit)} \end{cases}$$

Statistique de test et sa loi sous $\mathcal{H}_0$:

$$T(\mathbf{X}, \mathbf{Y}) = \frac{S_X^2}{S_Y^2} \overset{\mathcal{H}_0}{\sim} \mathcal{F}(n_X - 1,\ n_Y - 1), \quad f_{1-\alpha} := f_{1-\alpha,\, n_X-1,\, n_Y-1}, \quad F_T(x) \text{ f.d.r. de } \mathcal{F}(n_X - 1,\ n_Y - 1)$$

Rejet de $\mathcal{H}_0$ au risque α si: $\mathcal{R} = \left[f_{\alpha/2}\ ;\ f_{1-\alpha/2}\right]$

$$\begin{cases} \frac{s_X^2}{s_Y^2} < f_\alpha \\ \frac{s_X^2}{s_Y^2} \notin \mathcal{R} \\ \frac{s_X^2}{s_Y^2} > f_{1-\alpha} \end{cases} \quad \text{ou si} \quad \begin{cases} T(\mathbf{x}, \mathbf{y}) < f_\alpha \\ T(\mathbf{x}, \mathbf{y}) \notin \mathcal{R} \\ T(\mathbf{x}, \mathbf{y}) > f_{1-\alpha} \end{cases} \quad \text{ou si} \quad \begin{cases} \alpha^* = F_T(T(\mathbf{x}, \mathbf{y})) < \alpha \\ \alpha^* = 2 \min \left\{ F_T(T(\mathbf{x}, \mathbf{y})),\ 1 - F_T(T(\mathbf{x}, \mathbf{y})) \right\} < \alpha \\ \alpha^* = 1 - F_T(T(\mathbf{x}, \mathbf{y})) < \alpha \end{cases}$$

Tests d'hypothèses - Illustration

Exemple *(Comparaison de vibrations entre deux matériaux pour un panneau spatial)*.

Une équipe d'ingénierie aérospatiale doit choisir entre deux matériaux pour construire un panneau extérieur d'un satellite. Un critère crucial est la **stabil ité vibratoire** : le matériau ne doit pas présenter une variabilité trop élevée dans ses fréquences de vibration, afin d'éviter des résonances dangereuses en orbite.

Deux séries de tests indépendants sont menées : $n_A = 14$ mesures pour le **matériau A**, $n_B = 18$ mesures pour le **matériau B**.

Les fréquences de vibration (en Hz) sont bien modélisées par deux lois normales indépendantes, les paramètres sont **inconnus**. L'équipe veut vérifier si le matériau B possède une variance **plus élevée** que celle du matériau A.

Les variances empiriques observées sont: $s_A^2 = 0.42$ et $s_B^2 = 0.78$.

L'équipe accepte un risque d'erreur de type I de 1%.

☞ Le matériau B présente-t-il une variabilité vibratoire significativement plus élevée? ✍

Sommaire

Tests asymptotiques

Jusqu'à présent, on a **supposé que les échantillons étaient Gaussiens**
☞ C'était bien pratique pour **déterminer la loi de la statistique de test**

Mais c'est un cadre de travail très limitant
On ne peut pas modéliser **TOUS** les phénomènes par des 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 de la statistique de test
C'est l'idée derrière les **tests asymptotiques**

Rappel TLC

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*** $\mathbb{V}(X_i) = \sigma^2 < \infty$.

Alors la fonction (de n) suivante admet comme limite

$$\sqrt{n}\left(\frac{\overline{X}_n - \mu}{\sigma}\right) \underset{n \to \infty}{\longrightarrow} Z \sim \mathcal{N}(0,1)$$

Théorème *(Loi asymptotique du pivot pour μ).* ★

Dans le cadre du TLC, on a que:

$$\sqrt{n}\left(\frac{\overline{X}_n - \mu}{S}\right) \underset{n \to \infty}{\longrightarrow} Z \sim \mathcal{N}(0,1)$$

Test asymptotiques - Recette

Pour construire un **test asymptotique sur l'espérance d'un échantillon**: ★

1. **Formuler l'hypothèse nulle $\mathcal{H}_0$ et son alternative $\mathcal{H}_1$**
2. **Choisir un niveau de risque α**
3. **Identifier le type du test (unilatéral, bilatéral...)**
4. **Choisir $\overline{X}_n$ comme estimateur de l'espérance**
5. **Vérifier que nous sommes bien dans le régime asymptotique ($n > 30$)**
6. **Accepter l'approximation normale, sous $\mathcal{H}_0$ pour la statistique de test:**

$$\sqrt{n}\left(\frac{\overline{X}_n - \mu_0}{S}\right) \overset{\mathcal{H}_0}{\approx} \mathcal{N}(0,1)$$

7. **Rejeter ou non l'hypothèse nulle selon l'une des trois manières équivalentes (z-test)**

Test asymptotique sur la moyenne pour un éch. non-gaussien

Cadre de travail: $\mathbf{X} = (X_1, \ldots, X_n)$ i.i.d., $\mu = \mathbb{E}[X_1] < \infty$, $\sigma^2 = \mathbb{V}(X_1) < \infty$, **$\mu$ et σ^2 inconnus,** $n > 30$

Hypothèse du test statistique:

$$\mathcal{H}_0 : \mu = \mu_0 \quad \text{contre} \quad \mathcal{H}_1 = \begin{cases} \mu < \mu_0 & \text{(uni. gauche)} \\ \mu \neq \mu_0 & \text{(bilatéral)} \\ \mu > \mu_0 & \text{(uni. droit)} \end{cases}$$

Statistique de test et loi asymptotique sous $\mathcal{H}_0$: Dans le régime asymptotique, on accepte l'approximation:

$$T(\mathbf{X}) = \sqrt{n}\,\frac{\overline{X}_n - \mu_0}{S} \overset{\mathcal{H}_0}{\approx} \mathcal{N}(0,1), \quad z_{1-\alpha} := \Phi^{-1}(1-\alpha)$$

Rejet de $\mathcal{H}_0$ au risque α si:

$$\begin{cases} \overline{x}_n < \mu_0 - z_{1-\alpha}\,\frac{s}{\sqrt{n}} \\ \overline{x}_n \notin \left[\mu_0 \pm z_{1-\alpha/2}\,\frac{s}{\sqrt{n}}\right] \\ \overline{x}_n > \mu_0 + z_{1-\alpha}\,\frac{s}{\sqrt{n}} \end{cases} \quad \text{ou si} \quad \begin{cases} T(\mathbf{x}) < -z_{1-\alpha} \\ |T(\mathbf{x})| > z_{1-\alpha/2} \\ T(\mathbf{x}) > z_{1-\alpha} \end{cases} \quad \text{ou si} \quad \begin{cases} \alpha^* = \Phi(T(\mathbf{x})) < \alpha \\ \alpha^* = 2\,(1 - \Phi(|T(\mathbf{x})|)) < \alpha \\ \alpha^* = 1 - \Phi(T(\mathbf{x})) < \alpha \end{cases}$$

Tests d'hypothèses - Illustration

Exemple *(Surveillance du niveau de pollution d'une rivière)*.

Une agence environnementale surveille le niveau moyen de concentration en oxyde d'azote (NO_2) dans une rivière proche d'une zone industrielle. Les normes exigent que la concentration moyenne reste au-dessus d'un seuil minimal écologique de $\mu_0 = 3.5$ mg/L pour garantir un niveau acceptable d'oxygénation pour la faune aquatique.

Après plusieurs plaintes, l'agence souhaite tester si la moyenne réelle n'est pas descendue en dessous de ce seuil. Elle accepte un risque d'erreur de type I de 5%.

Un échantillon de $n = 314$ mesures hebdomadaires est collecté. La concentration peut être modélisée par une loi quelconque non-normale. La variance réelle est **inconnue**.

Les mesures donnent une moyenne empirique $\overline{x}_n = 3.41$ mg/L et un écart-type empirique $s_n = 0.62$ mg/L.

☞ Le niveau moyen de NO_2 est-il significativement passé sous le seuil écologique? ✍

Sommaire

Évaluer une modélisation

Dans cette section, nous allons construire des tests pour:

- **Test d'ajustement**: Décider à partir de données observées si un choix de modélisation est bon **(ajustement)**
- **Test d'indépendance**: Décider, à partir de deux échantillons, si les deux phénomènes modélisés sont **indépendants**

Exemple *(Triche à un jeu de société)*.

Pendant les fêtes, vous jouez à une longue partie de Monopoly. À plusieurs reprises, un joueur semble obtenir des 6 anormalement souvent, ce qui lui permet d'acheter pratiquement tout ce qu'il veut.

La suspicion monte: **le dé est-il pipé?**

Pour en avoir le cœur net, tout le monde se tourne vers vous et vous demande de trancher. Vous proposez de lancer n fois le dé pour construire un échantillon.

Échantillon $\mathbf{X} = (X_1, \ldots, X_n)$ i.i.d. de loi **discrète inconnue** dont le support est $\{1, \ldots, 6\}$.

On cherche à construire le test:

$$\mathcal{H}_0 : \text{le dé n'est pas pipé} \quad \text{contre} \quad \mathcal{H}_1 : \text{le dé est pipé}$$

ou encore

$$\mathcal{H}_0 : \mathbf{X} \text{ est de loi } \mathcal{U}(\{1, \ldots, 6\}) \quad \text{contre} \quad \mathcal{H}_1 : \mathbf{X} \text{ n'est pas de loi } \mathcal{U}(\{1, \ldots, 6\})$$

Test d'ajustement - Formalisation

En toute généralité, on formule un test d'ajustement de la façon suivante:

- Soit **X** un échantillon i.i.d. de loi populationnelle à densité/fonction de masse de probabilité (fmp) $f_X(t)$ **inconnue**

- On souhaite savoir si:

$$\mathcal{H}_0 : f_X = f_0 \quad \text{contre} \quad \mathcal{H}_1 : f_X \neq f_0$$

  où f_0 est la densité/fmp **que l'on souhaite ajuster**

  Par ex: f_0 peut être la densité Gaussienne, la fmp uniforme discrète...

Problème: On ne connaît pas la "vraie" densité/fmp f_X!

Avant, on a testé la valeur que pouvait prendre le paramètre θ inconnu
☞ Pour le test d'ajustement, c'est la loi populationnelle toute entière f_X qui est inconnue!

Test d'ajustement - Vers une première réponse

Intuitivement, on aimerait **comparer nos observations issues de f_X à la loi f_0**.

Exemple dé pipé:

On peut **calculer la fréquence d'occurrence de chaque chiffre** et le comparer à **celle qu'on devrait observer si le dé n'était pas pipé**:

Notons: $\widehat{p}_j(\mathbf{X}) = \frac{1}{n}\sum_{j=1}^{n} \mathbb{1}_{\{j\}}(X_i)$

☞ $\widehat{p}_j(\mathbf{x})$ = la fréquence d'occurrence de la valeur j après avoir lancé le dé n fois

Valeur du dé	1	2	3	4	5	6
Fréq. du dé (f_X)	$\widehat{p}_1(\mathbf{x})$	$\widehat{p}_2(\mathbf{x})$	$\widehat{p}_3(\mathbf{x})$	$\widehat{p}_4(\mathbf{x})$	$\widehat{p}_5(\mathbf{x})$	$\widehat{p}_6(\mathbf{x})$
Fréq. dé non-pipé (f_0)	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

☞ **Si chaque $\widehat{p}_j(x)$ est "proche" de $1/6$, on voudrait ne pas rejeter $\mathcal{H}_0$ (dé non-pipé)**

Test d'ajustement - Généralisation

Comment formaliser tout ça?

1. **On commence par partitionner le support $\mathcal{X}$ de f_X**:
   On crée k classes disjointes $C_1, \ldots, C_k$ **telles que** $\bigcup_{i=1}^{k} C_i = \mathcal{X}$.
2. **On calcule la fréquence d'individu dans chaque classe**: ★
   $$p_j = \mathbb{P}(X \in C_j) = \int_{C_j} f_X(t)dt \quad \text{et} \quad \widehat{p}_j(\mathbf{X}) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{C_j}(X_i)$$
   ☞ $\widehat{p}_j(\mathbf{x})$ est **le pourcentage des individu de l'échantillon observé $\mathbf{x}$ qui sont dans la classe** C_j
3. **On calcule les probabilités d'appartenir à chaque classe, sous $\mathcal{H}_0$**
   Soit X une v.a. ayant f_0 comme densité/fmp:
   $$p_j^0 = \mathbb{P}(X \in C_j \mid f_X = f_0) = \int_{C_j} f_0(t)dt = \int_{-\infty}^{\infty} \mathbb{1}_{C_j}(t) f_0(t)dt$$

On va **reformuler notre test:**

$$\mathcal{H}_0 : p_j = p_j^0, \text{ pour } j = 1, \ldots k \quad \text{contre} \quad \mathcal{H}_1 : \exists j \in \{1, \ldots, k\} \text{ tel que } p_j \neq p_j^0$$

Test d'ajustement - Créer les classes

Comment créer les classes $C_1, \ldots, C_k$?

- **Si la loi de X est discrète à k modalités:** ★

  ☞ **Chaque classe correspond à une modalité**

  Par ex. $\mathcal{U}(\{1, \ldots, k\})$: $C_j = \{j\}$ et $\mathbb{P}(X \in C_j) = \mathbb{P}(X = j)$

- **Si la loi de X est continue:** ★

  ☞ **On prend k intervalles disjoints et complémentaires $I_1, \ldots, I_k$**

  Par ex. $\mathcal{U}([0, 1])$: $C_1 = [0\ ;\ 0.5)$, $C_2 = [0.5\ ;\ 1]$ et $\mathbb{P}(X \in C_1) = \mathbb{P}(0 \leq X < 0.5)$

Test d'ajustement - Statistique de test

Pour construire un test, **il nous faut une statistique de test**, et **connaître sa loi sous $\mathcal{H}_0$!**

Théorème *(Test d'ajustement (admis)).* ★

Soit $\mathbf{X} = (X_1, \ldots, X_n)$ *un échantillon i.i.d. de loi populationnelle à densité inconnue* $f_X(t)$*. On dispose d'une discrétisation* $C_1, \ldots, C_k$ *du support de la loi de populationnelle.*

Pour une densité f_0*, on s'intéresse au test:*

$$\mathcal{H}_0 : p_j = p_j^0, \text{ pour } j = 1, \ldots k \quad \text{contre} \quad \mathcal{H}_1 : \exists j \in \{1, \ldots, k\} \text{ tel que } p_j \neq p_j^0.$$

Si $np_j^0 > 5$ *pour* $j = 1, \ldots, k$*, alors:*

$$T(\mathbf{X}) = \sum_{j=1}^{k} \frac{\left(\widehat{p}_j(\mathbf{X}) - p_j^0\right)^2}{p_j^0} \overset{\mathcal{H}_0}{\sim} \chi^2(k-1)$$

Test d'ajustement du Chi-carré, densité quelconque

Cadre de travail: $\mathbf{X}$ i.i.d. de densité f_X, partitionnement du support $C_1, \dots, C_k$, $p_j = \mathbb{P}(X \in C_j)$, $\widehat{p}_j(\mathbf{X}) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{C_j}(X_i)$, $p_j^0 = \mathbb{P}(X \in C_j \mid f_X = f_0)$.

Hypothèse du test statistique:

$$\mathcal{H}_0 : p_j = p_j^0 \text{ pour tout } j \quad \text{contre} \quad \mathcal{H}_1 : \exists j,\ p_j \neq p_j^0.$$

Statistique de test et loi sous $\mathcal{H}_0$:

Sous la condition que $np_j^0 > 5$ pour tout j

$$T(\mathbf{X}) = \sum_{j=1}^k \frac{\left(\widehat{p}_j(\mathbf{X}) - p_j^0\right)^2}{p_j^0} \overset{\mathcal{H}_0}{\sim} \chi^2(k-1), \quad c_{1-\alpha} := c_{1-\alpha,\,k-1}, F_T(x) \text{ f.d.r. de } \chi^2(k-1)$$

Rejet de $\mathcal{H}_0$ au risque α si:

$$T(\mathbf{x}) > c_{1-\alpha} \quad \text{ou si} \quad \alpha^* = 1 - F_T\big(T(\mathbf{x})\big) < \alpha$$

(Il n'y a pas de versions unilatérale gauche ou bilatérale pertinentes)

Tests d'hypothèses - Illustration

Exemple *(Répartition des directions de vagues autour d'un navire)*.

Un navire océanographique utilise un radar pour enregistrer la direction d'arrivée des vagues, codée entre 0 et 1 (où 0 correspond à l'avant du navire, 0.5 à l'arrière).

Pour des raisons de sécurité, l'équipe de navigation veut vérifier si les directions sont **uniformément réparties**. Une direction non uniforme peut signaler un risque accru de chavirement.

Un échantillon de $n = 200$ directions est collecté, supposé i.i.d. mais sans loi connue a priori. On souhaite tester si ces données suivent une $\mathcal{U}([0, 1])$. L'équipe accepte un risque d'erreur de type I de 5%.

On partitionne l'intervalle $[0, 1]$ en 5 classes de même longueur et le tableau suivant donne les fréquences observées :

Classe	Partitions	Effectif observé
Nord-Est	$[0,\ 0.2)$	28
Est	$[0.2,\ 0.4)$	45
Sud	$[0.4,\ 0.6)$	52
Ouest	$[0.6,\ 0.8)$	39
Nord-Ouest	$[0.8,\ 1]$	36

☞ Les directions de vagues peuvent-elles être considérées comme uniformément réparties? ✍

Test d'ajustement - Ajuster une loi paramétrée

Dans le test précédent, **on a supposé que la densité f_0 n'avait pas de paramètre**

Si jamais $f_0(t;\theta)$ était paramétrée par θ inconnu, alors **on aurait pas pu calculer les probabilités p_j^0!**

Mais on peut adapter le test précédent au cas où f_0 est paramétrée!

Intuition: Remplacer le paramètre θ inconnu **par une estimation** $\widehat{\theta}(\mathbf{X})$.

Théorème *(Test d'ajustement, densité paramétrée (admis)).* ★

Soit $\mathbf{X} = (X_1, \ldots, X_n)$ *un échantillon i.i.d. de loi populationnelle à densité inconnue* $f_X(t)$*. On dispose d'une discrétisation* $C_1, \ldots, C_k$ *du support de la loi de populationnelle. Soit* $f_0(.;\theta)$ *une densité paramétrée par* d *paramètres* θ *inconnus et soit* $\widehat{\theta}(\mathbf{X})$ *un estimateur de* θ*.*

On note, pour $j = 1, \ldots, k$*:*

$$\widehat{p}_j^0 = \mathbb{P}\left(X \in C_j \mid f_X(.) = f_0\left(. \, ; \, \widehat{\theta}(\mathbf{x})\right)\right) = \int_{C_j} f_0\left(t, \widehat{\theta}(\mathbf{x})\right) dt$$

Pour une densité $f_0(.;\theta)$*, on s'intéresse au test:*

$$\mathcal{H}_0 : p_j = p_j^0, \text{ pour } j = 1, \ldots k \quad \text{contre} \quad \mathcal{H}_1 : \exists j \in \{1, \ldots, k\} \text{ tel que } p_j \neq p_j^0.$$

Si $n\widehat{p}_j^0 > 5$ *pour* $j = 1, \ldots, k$*, alors:*

$$T(\mathbf{X}) = \sum_{j=1}^{k} \frac{\left(\widehat{p}_j(\mathbf{X}) - \widehat{p}_j^0\right)^2}{\widehat{p}_j^0} \overset{\mathcal{H}_0}{\sim} \chi^2(k - d - 1).$$

Test d'ajustement du Chi-carré, densité paramétrée

Cadre de travail: $\mathbf{X}$ i.i.d. de densité f_X, partitionnement du support $C_1, \dots, C_k$, $p_j = \mathbb{P}(X \in C_j)$, $\widehat{p}_j(\mathbf{X}) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{C_j}(X_i)$, $p_j^0 = \mathbb{P}(X \in C_j \mid f_X = f_0)$.

$f_0(t, \theta)$ est supposée paramétrée par $\theta \in \mathbb{R}^d$, et $\widehat{\theta}(\mathbf{X})$ est un estimateur de θ. On approche p_j^0 par $\widehat{p}_j^0$ en supposant que $\theta = \widehat{\theta}(\mathbf{X})$

Hypothèse du test statistique:

$$\mathcal{H}_0 : p_j = p_j^0 \text{ pour tout } j \quad \text{contre} \quad \mathcal{H}_1 : \exists j,\ p_j \neq p_j^0.$$

Statistique de test et loi sous $\mathcal{H}_0$:
Sous la condition que $n\widehat{p}_j^0 > 5$ pour tout j

$$T(\mathbf{X}) = \sum_{j=1}^{k} \frac{\left(\widehat{p}_j(\mathbf{X}) - \widehat{p}_j^0\right)^2}{\widehat{p}_j^0} \overset{\mathcal{H}_0}{\sim} \chi^2(k-d-1), \quad c_{1-\alpha} := c_{1-\alpha,\,k-d-1}, F_T(x) \text{ f.d.r. de } \chi^2(k-d-1)$$

Rejet de $\mathcal{H}_0$ au risque α si:

$$T(\mathbf{x}) > c_{1-\alpha}, \quad \text{ou si} \quad \alpha^* = 1 - F_T\big(T(\mathbf{x})\big) < \alpha$$

Exemple *(Production d'un parc éolien)*.

Un opérateur énergétique surveille la production quotidienne (en MWh) d'un parc éolien. Il souhaite vérifier si la production journalière peut être modélisée par une **loi normale**.

Il dispose de $n = 250$ mesures de production journaliers. La loi réelle est inconnue, mais on veut tester l'ajustement à une $\mathcal{N}(\mu, \sigma^2)$ où les paramètres sont **inconnus**.

Les estimateurs empiriques donnent : $\overline{X}_n = 145.2$, $S^2 = 18.7$

On partitionne l'espace des valeurs en 6 classes d'effectifs théoriques équilibrés:

Classe	Partitions (MWh)	Effectif observé
Très faible production	$(-\infty,\ 123.0)$	38
Faible production	$[123.0,\ 134.7)$	44
Production moyenne	$[134.7,\ 145.2)$	39
Production standard	$[145.2,\ 155.7)$	43
Forte production	$[155.7,\ 167.4)$	47
Très forte production	$[167.4,\ \infty)$	39

☞ La production journalière peut-elle être considérée comme suivant une loi normale? ✍

Test d'indépendance - Motivation

Exemple *(Âge et moment de la journée pour faire l'épicerie)*.

Une grande chaîne de supermarchés au Québec souhaite savoir si le moment de la journée où les clients font leurs courses est lié à leur tranche d'âge. On effectue un sondage sur un échantillon de 300 clients, en notant pour chacun :

- Sa tranche d'âge: [18 ; 34], [35 ; 54], ou [55+]
- Le moment privilégié pour faire ses courses: matin, après-midi ou soir

Est-ce que le moment de la journée où les gens font l'épicerie est indépendant de la tranche d'âge?

On a deux échantillons **X** (âge) et **Y** (moment des courses) supposés i.i.d.

On cherche à construire le test:

$$\mathcal{H}_0 : \text{le moment des courses ne dépend pas de l'âge}$$

contre

$$\mathcal{H}_1 : \text{le moment des courses dépend de l'âge}$$

Test d'indépendance - Formalisation

En toute généralité, on formule un **test d'indépendance** de la façon suivante:

- Soit **X** et **Y** deux échantillons i.i.d. de loi populationnelle inconnue
- On souhaite savoir si

$$\mathcal{H}_0 : \mathbf{X} \perp\!\!\!\perp \mathbf{Y} \quad \text{contre} \quad \mathcal{H}_1 : \mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}$$

Intuition:

Comparer le tableau de contingence observé au tableau sous hypothèse d'indépendance (sous $\mathcal{H}_0$)

Test d'indépendance - Tableau de contingence

Pour **X** catégorielle avec $M_1, \ldots, M_{m_X}$ modalités, et **Y** catégorielle avec $U_1, \ldots, U_{m_Y}$ modalités

Un **tableau de contingence** est un tableau à **double entrée** qui comptabilise, pour chaque **paire de modalité** (M_i, U_j), le nombre d'individus appartenant simultanément à ces **deux** catégories.

	$Y = U_1$	$Y = U_2$	$Y = U_3$	Total
$X = M_1$	n_{11}	n_{12}	n_{13}	$n_{1\bullet}$
$X = M_2$	n_{21}	n_{22}	n_{23}	$n_{2\bullet}$
$X = M_3$	n_{31}	n_{32}	n_{33}	$n_{3\bullet}$
Total	$n_{\bullet 1}$	$n_{\bullet 2}$	$n_{\bullet 3}$	n

☞ $n_{ij}(\mathbf{x}, \mathbf{y}) = \sum_{l=1}^{n} \mathbb{1}_{M_i}(x_l)\mathbb{1}_{U_j}(y_l)$ c'est le nombre d'individu ayant la modalité M_i **et** U_j

☞ $n_{i\bullet}(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^{m_Y} n_{ij}(\mathbf{x}, \mathbf{y})$ c'est le nombre total d'individu ayant la modalité M_i (**X**)

☞ $n_{\bullet j}(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{m_X} n_{ij}(\mathbf{x}, \mathbf{y})$ c'est le nombre total d'individu ayant la modalité U_j (**Y**)

☞ $\sum_{i=1}^{m_X} n_{i\bullet}(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^{m_Y} n_{\bullet j}(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{m_X} \sum_{j=1}^{m_Y} n_{ij}(\mathbf{x}, \mathbf{y}) = n$ la taille de l'échantillon

Et si une, ou les deux variables sont continues?

☞ **On discrétise la variable comme pour le test d'adéquation**

Test d'indépendance - Comparaison

On veut comparer le **tableau de contingence observé** à celui **sous $\mathcal{H}_0$**

Sous $\mathcal{H}_0$, on a que, le nombre d'individu observé ayant les modalités M_i et U_j:

$$n \times \mathbb{P}(\{X = M_i\} \cap \{Y = U_j\}) \overset{\mathcal{H}_0}{=} n \times \mathbb{P}(\{X = M_i\}) \times \mathbb{P}(\{Y = U_j\}) =: E_{ij}$$

☞ $\frac{n_{i\bullet}(\mathbf{X},\mathbf{Y})}{n}$ est un estimateur de $\mathbb{P}(\{X = M_i\})$ et $\frac{n_{\bullet j}(\mathbf{X},\mathbf{Y})}{n}$ est un estimateur de $\mathbb{P}(\{Y = U_j\})$

Donc, on peut estimer E_{ij} par:

$$\widehat{E}_{ij}(\mathbf{X}, \mathbf{Y}) = n \times \frac{n_{i\bullet}(\mathbf{X}, \mathbf{Y})}{n} \times \frac{n_{\bullet j}(\mathbf{X}, \mathbf{Y})}{n} = \frac{n_{i\bullet}(\mathbf{X}, \mathbf{Y})\, n_{\bullet j}(\mathbf{X}, \mathbf{Y})}{n}$$

En notant $n_{ij} = n_{ij}(\mathbf{x}, \mathbf{y})$ et $\widehat{E}_{ij} = \widehat{E}_{ij}(\mathbf{x}, \mathbf{y})$

Tableau de contingence observé

	$Y = U_1$	$Y = U_2$	$Y = U_3$	Total
$X = M_1$	n_{11}	n_{12}	n_{13}	$n_{1\bullet}$
$X = M_2$	n_{21}	n_{22}	n_{23}	$n_{2\bullet}$
$X = M_3$	n_{31}	n_{32}	n_{33}	$n_{3\bullet}$
Total	$n_{\bullet 1}$	$n_{\bullet 2}$	$n_{\bullet 3}$	n

Tableau de contingence sous $\mathcal{H}_0$

	$Y = U_1$	$Y = U_2$	$Y = U_3$	Total
$X = M_1$	$\widehat{E}_{11}$	$\widehat{E}_{12}$	$\widehat{E}_{13}$	$n_{1\bullet}$
$X = M_2$	$\widehat{E}_{21}$	$\widehat{E}_{22}$	$\widehat{E}_{23}$	$n_{2\bullet}$
$X = M_3$	$\widehat{E}_{31}$	$\widehat{E}_{32}$	$\widehat{E}_{33}$	$n_{3\bullet}$
Total	$n_{\bullet 1}$	$n_{\bullet 2}$	$n_{\bullet 3}$	n

Test d'indépendance - Statistique de test

Théorème *(Test d'indépendance (admis)).* ★

Soient $\mathbf{X}$ *et* $\mathbf{Y}$ *deux échantillons i.i.d. On suppose que* $\mathbf{X}$ *est catégorielle (ou discrétisée) selon* m_X *modalités, et* $\mathbf{Y}$ *est catégorielle (ou discrétisée) selon* m_Y *modalités.*

On s'intéresse au test:

$$\mathcal{H}_0 : \mathbf{X} \perp\!\!\!\perp \mathbf{Y} \quad \text{contre} \quad \mathcal{H}_1 : \mathbf{X} \not\!\perp\!\!\!\perp \mathbf{Y}$$

Si pour tout $i = 1, \ldots, m_X$ *et* $j = 1, \ldots, m_Y$*;* $\widehat{E}_{ij} > 5$*, alors:*

$$T(\mathbf{X}, \mathbf{Y}) = \sum_{i=1}^{m_X} \sum_{j=1}^{m_Y} \frac{\left(n_{ij} - \widehat{E}_{ij}\right)^2}{\widehat{E}_{ij}} \overset{\mathcal{H}_0}{\sim} \chi^2\left((m_X - 1)(m_Y - 1)\right)$$

Test d'indépendance du Chi-carré

Cadre de travail: $\mathbf{X}$ i.i.d. à m_X modalités, $\mathbf{Y}$ i.i.d. à m_Y modalités (ou discrétisés)

Hypothèse du test statistique:

$$\mathcal{H}_0 : \mathbf{X} \perp\!\!\!\perp \mathbf{Y} \quad \text{contre} \quad \mathcal{H}_1 : \mathbf{X} \not\perp\!\!\!\perp \mathbf{Y}$$

Statistique de test et loi sous $\mathcal{H}_0$:

Sous la condition que pour tout $i = 1, \ldots, m_X$ et $j = 1, \ldots, m_Y$; $\widehat{E}_{ij} > 5$

$$T(\mathbf{X}, \mathbf{Y}) = \sum_{i=1}^{m_X} \sum_{j=1}^{m_Y} \frac{\left(n_{ij} - \widehat{E}_{ij}\right)^2}{\widehat{E}_{ij}} \overset{\mathcal{H}_0}{\sim} \chi^2\left((m_X - 1)(m_Y - 1)\right)$$

Rejet de $\mathcal{H}_0$ au risque α si:

$$T(\mathbf{x}, \mathbf{y}) > c_{1-\alpha,(m_X-1)(m_Y-1)} \quad \text{ou si} \quad \alpha^* = 1 - F_T\big(T(\mathbf{x}, \mathbf{y})\big) < \alpha$$

Test d'indépendance (Chi-carré) - Illustration

Exemple *(Âge et moment de la journée pour faire l'épicerie)*.

Une grande chaîne de supermarchés au Québec souhaite savoir si le moment de la journée où les clients font leurs courses est lié à leur tranche d'âge. On effectue un sondage sur un échantillon de 300 clients, en notant pour chacun :

- Sa tranche d'âge: [18 ; 34], [35 ; 54], ou [55+]
- Le moment privilégié pour faire ses courses: matin, après-midi ou soir

La chaîne de supermarchés souhaite savoir si elle a un intérêt à faire de la publicité ciblée en tout au long de la journée. Elle souhaite avoir un niveau de confiance de 90%. On observe la table de contingence suivante:

Tranche d'âge	**Matin**	**Après-midi**	**Soir**	**Total**
18–34 ans	15	45	60	120
35–54 ans	30	50	30	110
55 ans et plus	35	15	20	70
Total	**80**	**110**	**110**	**300**

☞ Quelle serait votre recommandation pour la stratégie publicitaire de la chaîne de supermarché? ✍