11 TESTS DE CONFORMITE
(GRANDS ECHANTILLONS)

«Réfléchir c'est nier
ce que l'on croit.»
ALAIN

MARCHE D'APPROCHE

1. GENERALITES

L'information du statisticien, sauf exception, ne porte que sur un nombre limité de valeurs qui composent un échantillon. Or ce qui est intéressant, ce n'est pas l'échantillon lui-même, mais la population dont il est extrait.
Le problème de la statistique inductive est donc de fournir des réponses générales à partir d'informations partielles.
Une telle démarche ne permet jamais d'acquérir une certitude. Elle doit cependant permettre de formuler des jugements, assortis d'un degré mesurable de crédibilité.

1. 2. Tricheur ou non ?

« Sur 20 lancers successifs d'un dé, un joueur obtient 8 fois le nombre 6 ».
Doit-on considérer que ce résultat est dû au seul hasard ou bien, au contraire, qu'il est la conséquence d'une tricherie ?

La fréquence théorique de sortie du nombre 6 est évidemment, pour un dé normal, $\frac{1}{6}$ alors qu'ici on a une fréquence de $\frac{2}{5}$.

Il est logique de « présumer l'accusé innocent » donc de formuler l'hypothèse que le joueur ne triche pas. Cette hypothèse, dite **hypothèse nulle**, est notée(H_0).
Sous cette hypothèse, nous pouvons calculer rigoureusement la probabilité de l'événement A : « apparition de 8 fois le nombre 6 sur 20 lancers ».
La variable aléatoire X qui prend pour valeur le nombre d'apparitions du nombre 6 sur 20 lancers, suit la loi binomiale $\mathscr{B}\left(20\,;\frac{1}{6}\right)$. Alors $p(A) = C_{20}^{12}\left(\frac{1}{6}\right)^{8}\left(\frac{5}{6}\right)^{12} \approx 0{,}0084$.

Nous avons alors tous les éléments pour prendre notre décision :

- Ou bien conclure que (H_0) est vraie.
- Ou bien conclure que (H_0) est fausse donc que le joueur a triché... par exemple, en utilisant un dé pipé.

Notons que, quelle que soit notre décision, elle peut être erronée :

- Soit en refusant (H_0) alors qu'elle est vraie. Ce risque d'erreur est appelé **risque de première espèce** et est généralement noté α.
- Soit en acceptant (H_0) alors qu'elle est fausse. Ce risque d'erreur est appelé **risque de seconde espèce** et noté β.

Ici le risque α est inférieur à 1% donc il paraît clair que nous refusons (H_0). Le tableau suivant résume la situation dans laquelle nous sommes :

(H_0) / Décision	Vraie	Fausse
Accepter	Décision correcte (risque $1 - \alpha$)	Erreur **de seconde espèce** (risque β)
Refuser	Erreur **de première espèce** (risque α)	Décision correcte (risque $1 - \beta$)

1. 2. Deux exemples de la vie courante

■ Acheteur et vendeur

etc...

■ Le magistrat

Lors d'un procès, le magistrat doit se prononcer sur l'hypothèse (H_0) : « Le prévenu est innocent » en ne disposant que d'informations partielles, fournies par le juge d'instruction et les témoignages. Là encore, il peut être amené à commettre deux types d'erreurs :

etc...

CAMP DE BASE

1. TEST DE CONFORMITE D'UNE MOYENNE

Population P

X suit une loi de moyenne μ et d'écart-type σ.

Normes de fabrication

Elles définissent **la conformité d'un objet**.

Loi d'échantillonnage des échantillons de taille n

Pour $n \geq 30$, $\overline{X}$ suit la loi normale de moyenne μ et d'écart-type $\dfrac{\sigma}{\sqrt{n}}$. Alors la variable aléatoire $T = \dfrac{\overline{X}-\mu}{\dfrac{\sigma}{\sqrt{n}}}$ suit la loi normale centrée réduite.

Echantillon, de taille n, à tester

On calcule la moyenne m_e de cet échantillon.
Si σ est inconnu on l'estime par $\hat{\sigma} = \sigma_e\sqrt{\dfrac{n}{n-1}}$ où σ_e est l'écart-type de l'échantillon.

Construction d'un test bilatéral

■ Hypothèse à tester **(H_0) :** $\boldsymbol{\mu = \mu_0}$

■ Hypothèse alternative (ou contradictoire **) (H_1) :** $\boldsymbol{\mu \neq \mu_0}$

■ Condition de rejet de (H_0), au risque α, fixé à priori :

- ♦ On note t_α le réel positif tel que $(|T| \leq t_\alpha) = 1 - \alpha$, d'où : $t_\alpha = \Pi^{-1}\left(1 - \dfrac{\alpha}{2}\right)$.
- ♦ On rejette (H_0) si $|T| > t_\alpha$

Mise en œuvre du test

♦ Calcul de $t = \dfrac{m_e - \mu_0}{\dfrac{\sigma}{\sqrt{n}}}$

Comparaison de $|t|$ avec t_α.

Décision

Cette décision concerne **l'échantillon testé** :

- ♦ Si $|T| \leq t_\alpha$ on accepte (H_0) : l'échantillon est conforme (ou représentatif).
- ♦ Sinon, il ne l'est pas.