

La section statistique qu'un relecteur ne peut pas démonter

Sept leçons ont produit onze nombres. Celle-ci les met dans un rapport — ce qui va dans le corps, ce qui va en annexe, ce qu'un relecteur demandera, et les quatre phrases qui répondent à chaque question avant qu'elle soit posée.

Pierre Robentz Cassion

Leçon 8 sur 8

Statistiques appliquées aux programmes — Ce que le rapport affirme

Ce que couvre cette leçon

- Ce que fait réellement un relecteur
- Les quatre phrases
- Corps, annexe et script
- Le tableau d'annexe
- Le paragraphe de limites, écrit depuis le cours
- Que faire quand le nombre refuse de trancher
- Rapportez-le en entier
- La suite

Ce que fait réellement un relecteur

Le relecteur demande	Répondu par
« À quel point êtes-vous sûr de ce nombre ? »	L'intervalle — leçon 2
« Cette différence est-elle réelle ? »	Le test et ses hypothèses — leçon 3
« Est-ce que cela compte ? »	La taille d'effet dans l'unité de la décision — leçon 4
« Combien de choses avez-vous regardées ? »	Le nombre de tests et la correction — leçon 5

Ce que fait réellement un relecteur

- L'ordre n'est pas celui dans lequel vous les avez calculées — L'intervalle vient en premier parce que c'est le seul...

- Un : le nombre et son intervalle — « L'aboutissement des orientations parmi les cas signalant un handicap vaut 26,7 %...
- Deux : la comparaison et son test — « C'est 19,4 points en dessous des cas sans handicap signalé (46,2 %, $n = 1\,436$;...
- Trois : l'ampleur, dans l'unité de la décision — « Rapportée à la charge de cas 2024, cette différence vaut 39 cas qui...
- Quatre : ce que l'analyse ne peut pas dire — « La comparaison est observationnelle

Les quatre phrases — En Python

```
import pandas as pd
import numpy as np

protection = pd.read_csv("protection-referrals-2024.v1.csv")
consenting = protection[protection["consent_to_refer"]]
disability = consenting["disability_reported"].isin([True, "true", "Yes"])
reached = consenting["referral_accepted"] & consenting["days_to_first_service"].notna()

k, n = reached[disability].sum(), disability.sum()
comparator = reached[~disability].mean()
print(f"cases short of the comparator rate: {comparator * n - k:.0f}")
```

Les quatre phrases — En R

Sentence three is arithmetic, and it is the sentence a manager reads.

Les quatre phrases

- La troisième phrase est celle que les analystes sautent et dont les responsables ont besoin — Des points de pourcentage. . .

- Le corps porte l'affirmation — L'estimation, son intervalle, la comparaison, l'effet en unités de programme, et la...
- L'annexe porte la machinerie — Noms des tests, statistiques, degrés de liberté, vérifications des hypothèses, nombre de...
- Le script porte la reproduction — Chaque nombre des deux, produit par du code qui tourne à partir du fichier versionné...

Corps, annexe et script — En Python

```
def claim(label, k, n, digits=1):
    p = k / n
    z = 1.96
    denom = 1 + z**2 / n
    centre = (p + z**2 / (2 * n)) / denom
    half = z * np.sqrt(p * (1 - p) / n + z**2 / (4 * n**2)) / denom
    return (f"{label}: {p:.{digits}%} "
            f"(95% CI {centre - half:.{digits}%} to {centre + half:.{digits}%}, n={n})")

print(claim("Referral completion, disability reported", 54, 202))
```

Generate the sentence, do not type it. A typed interval drifts from its number.

- Un nombre saisi à la main dans un document est un nombre qui sera faux après la prochaine correction de données —...

Le tableau d'annexe — En Python

```
annex = pd.DataFrame([
    {"comparison": "Completion by disability status", "test": "two-sample z",
     "statistic": "z = -5.21", "p": "<0.001", "n": "202 vs 1,436",
     "effect": "-19.4 pts (CI -26.1 to -12.8)"},
    {"comparison": "Over-age by sex", "test": "two-sample z",
     "statistic": "z = 1.84", "p": "0.066", "n": "510 vs 542",
     "effect": "+5.5 pts (CI -0.4 to +11.3)"},
    {"comparison": "Attendance by sex", "test": "two-sample z",
     "statistic": "z = 2.28", "p": "0.023", "n": "68,267 marks",
     "effect": "+0.56 pts, risk ratio 1.006"},
    {"comparison": "Attendance by school feeding", "test": "Welch t, school level",
     "statistic": "t = 3.11, df 12.7", "p": "0.008", "n": "15 vs 9 schools",
     "effect": "+5.2 pts (CI 1.6 to 8.8)"},
])
print(annex.to_string(index=False))
```

Le tableau d'annexe — En R

One row per test run, including the ones that found nothing.

Le tableau d'annexe

Comparaison	Test	Statistique	p	Effet
Aboutissement selon le handicap	z à deux échantillons	$z = -5,21$	$<0,001$	-19,4 pts
Retard d'âge selon le sexe	z à deux échantillons	$z = 1,84$	0,066	+5,5 pts
Présence selon le sexe	z à deux échantillons	$z = 2,28$	0,023	+0,56 pt, RR 1,006
Présence selon la cantine	t de Welch, niveau école	$t = 3,11$, ddl 12,7	0,008	+5,2 pts
Retard d'âge par sexe, 24 écoles	$z \times 24$	2 sur 24 à $p < 0,05$	—	0 survit à Bonferroni

Le tableau d'annexe

- Trois de ces cinq lignes sont en annexe et non dans le corps — parce qu'un écart non significatif, une significativité. . .
- Lister les tests qui n'ont rien trouvé est ce qui rend crédibles ceux qui ont trouvé quelque chose — Une annexe ne. . .

Le paragraphe de limites, écrit depuis le cours

- Comparaison observationnelle — « Les groupes n'ont pas été randomisés ; les différences peuvent refléter qui entre dans. . .
- Regroupement en grappes — « Le programme est attribué par école, donc l'analyse emploie 24 unités indépendantes
- Comparaisons multiples — « Vingt-quatre tests au niveau école ont été lancés ; 1,2 résultat significatif est attendu. . .
- Restriction de l'étendue — « La présence de cette cohorte va de 86,7 % à 97,8 % pour la moitié centrale, si bien que. . .
- Chacune tient en une phrase et chacune devance une question précise — Un paragraphe de limites qui dit « les données. . .

Que faire quand le nombre refuse de trancher

Résultat	Pourquoi il ne tranche pas	Que faire
Retard d'âge par sexe, $p = 0,066$	Sous-puissant pour un écart de 5 points	Le dire ; regrouper avec le passage suivant plutôt que fermer la question
Présence et littératie, $r = 0,04$	Étendue restreinte	Rapporter le résultat nul et nommer la restriction
Cantine et présence, 24 grappes	Peu de grappes, observationnel	Rapporter l'intervalle ; ne pas revendiquer d'effet

Que faire quand le nombre refuse de trancher

- « Les données ne peuvent pas trancher » est un constat sur lequel un programme peut agir — parce que l'action consiste. . .

Statistical methods

Proportions are reported with Wilson 95% confidence intervals. Group comparisons use two-sample z-tests for proportions and Welch's t-test for means. Where a programme is assigned above the level of the observation, the analysis is conducted at the level of assignment.

Analyses were specified before the data were examined. Where a family of comparisons was run, the number of tests is stated with the result and a Bonferroni correction applied.

All figures are generated from the committed dataset files by the scripts in this annex; no number in this report is typed by hand.

Findings reported in the body: 2. Comparisons run in total: 29.

- La dernière ligne est celle qui change la façon dont la section est lue — Deux constats sur vingt-neuf comparaisons est. . .

- Mais la discipline est complète ici — Chaque nombre d'un rapport de programme peut porter un intervalle, une taille. . .

Lire la leçon complète, avec le code exécutable

`https://data-analysis.cassion.dev/fr/cours/
applied-statistics-for-programmes/stat-08-the-statistics-in-the-report`