

Modélisation statistique

#1.a Tests d'hypothèses

Dr. Léo Belzile
HEC Montréal

Tests d'hypothèse

Analogie du procès



Capture d'écran du drame judiciaire *Douze hommes en colère* (1957)

Recette

Un test d'hypothèse est une règle de décision binaire.

Voici les étapes de la démarche.

1. Définir les variables d'intérêt
2. Formuler les hypothèses alternatives et nulles, \mathcal{H}_a et \mathcal{H}_0
3. Choisir la statistique de test
4. Comparer la valeur numérique obtenue à la loi nulle
5. Obtenir une valeur- p
6. Conclure dans le contexte du problème

Le Tech3Lab est le plus important laboratoire en expérience utilisateur (UX) en Amérique du Nord

Le Tech3Lab est un laboratoire de recherche appliquée en sciences de la gestion qui se spécialise dans l'analyse des interactions entre les interfaces technologiques des organisations et leurs employés ou consommateurs



Les dangers de texter en marchant

Le risque de texter en marchant : La lenteur du désengagement

OBJECTIF : COMPRENDRE LES PROCESSUS COGNITIFS
SOUS-JACENTS AU MULTITÂCHE MOBILE

HYPOTHÈSE

L'engagement cognitif ralentit la
capacité attentionnelle du piéton

MESURES

Électroencéphalographie



Avec la permission de
ICI Radio-Canada

Détails de l'étude

- + 35 sujets ont participé à l'étude.
- + Chaque personne devait marcher sur un tapis roulant et un écran projetait des obstacles.
- + Au cours d'une séance, le sujet marchait en parlant au cellulaire tandis que lors d'une autre séance, il devait marcher en textant.
- + L'ordre entre les séances a été déterminé *au hasard*.
- + Différents obstacles étaient présentés durant une séance selon un ordre aléatoire.
- + Nous allons nous intéresser à l'un d'eux, l'apparition d'un cycliste dans le champ visuel

Caractéristiques

- + Population: adultes (18 ans et plus)
- + Échantillon: 35 individus
- + Variables:
 - + temps pour percevoir un obstacle: quantitatif
 - + type de distraction (cellulaire ou textos): nominale

On s'intéresse au temps (en secondes) que prend une personne pour apercevoir cet obstacle (mesuré à l'aide d'un encéphalogramme).

#1. Définir les variables d'intérêt

- + μ_c : moyenne du temps de réaction (en secondes) lors d'un appel (c)
- + μ_t : moyenne du temps de réaction (en secondes) lorsqu'on texte (t)

#2. Formuler les hypothèses alternatives et nulles

- + Hypothèse d'intérêt: est-ce que texter est plus distrayant?
 - + $\mathcal{H}_a : \mu_t > \mu_c$ (unilatéral)
- + Hypothèse nulle (avocat du Diable)
 - + $\mathcal{H}_0 : \mu_t \leq \mu_c$

On exprime l'hypothèse en fonction de la différence:

$$\mathcal{H}_a : \mu_t - \mu_c > 0.$$

#3. Choisir la statistique de test

On s'intéresse à la différence de temps de réaction (moyenne)

- + Test- t pour un échantillon pour $t - c$ (un test- t pour données appariées)

$$T_D = \frac{\bar{D} - \mu_0}{\text{se}(\bar{D})}$$

- + \bar{D} est la différence des temps moyens de réaction.
- + On postule $\mu_0 = \mu_t - \mu_c = 0$.
- + L'erreur-type de \bar{D} est $\text{se}(\bar{D}) = S_D / \sqrt{n}$, où S_D est l'écart-type des variables D_i et n est la taille de l'échantillon.

-
- Code SAS + Sortie SAS + Code R + Sortie R
-

```
proc ttest data=modstat.distraction side=u;  
paired t*c;  
run;
```

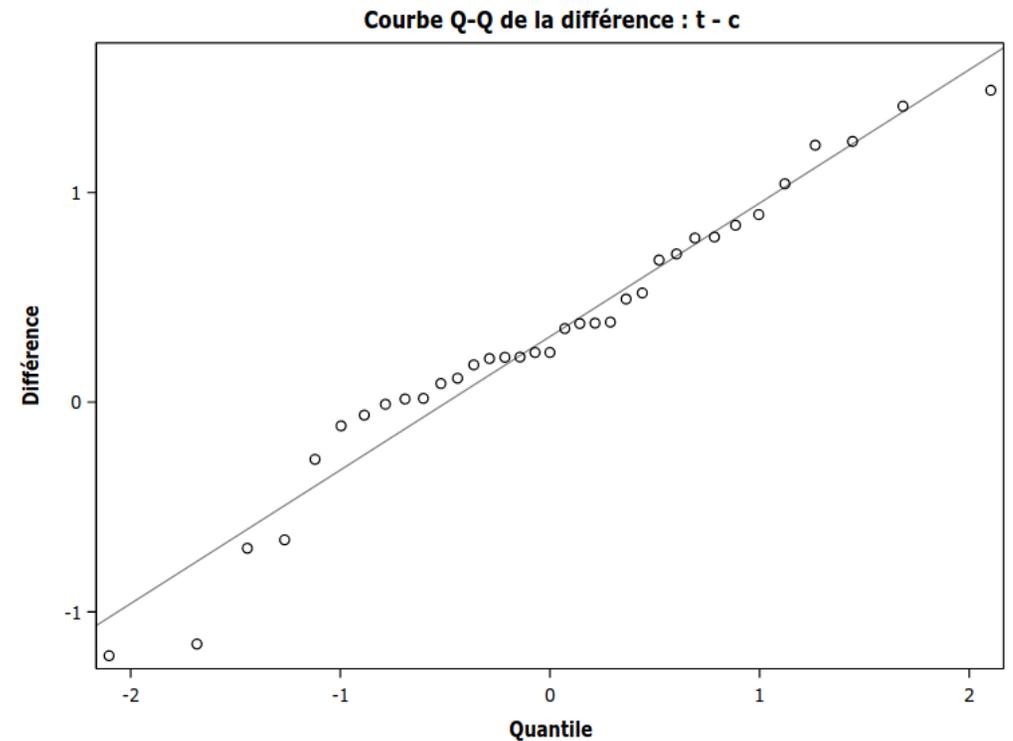
- Code SAS + Sortie SAS + Code R + Sortie R

Différence : t - c

N	Moyenne	Ec-type	Err. type	Minimum	Maximum
35	0.3130	0.6369	0.1077	-1.2100	1.4880

Moyenne	IC à95% - Moyenne	Ec-type	Ec.-type de l'IC à95%		
0.3130	0.1309	Infini	0.6369	0.5152	0.8345

DDL	Valeur du test t	Pr > t
34	2.91	0.0032



- Code SAS + Sortie SAS + Code R + Sortie R

```
data(distractio, package = "hecmoostat")  
with(distractio, t.test(t-c, alternative = "greater", mu = 0))
```

- Code SAS + Sortie SAS + Code R + Sortie R

```
##  
##      One Sample t-test  
##  
## data:  t - c  
## t = 2.91, df = 34, p-value = 0.0032  
## alternative hypothesis: true mean is greater than 0  
## 95 percent confidence interval:  
##  0.13092      Inf  
## sample estimates:  
## mean of x  
##  0.31297
```

#4. Comparer la valeur numérique obtenue à la loi nulle

La loi nulle est Student- t avec 34 degrés de liberté, St_{34} .

On s'intéresse uniquement à la probabilité d'obtenir $T_D > 2.91$ sous \mathcal{H}_0 .

#5. Obtenir une valeur- p

La valeur- p est 0.0032, inférieure à $\alpha = 5\%$.

#6. Conclure dans le contexte du problème

On rejette \mathcal{H}_0 : le temps de réaction est significativement plus élevé lorsqu'on texte que lorsqu'on parle au cellulaire en marchant.

La différence moyenne estimée est 0.313 secondes.