Psy1004 – Section 5:

Le théorème central limite

Plan du cours:

- Varia
- Le TP1
- 5.0: Tests de moyennes
- 5.1: Le test *z*
  - a) Survol
  - b) Exemple 1
  - c) Exemple 2
- 5.2: Les limites du test *z*
- 5.3: Le test *t*
  - a) Le théorème central limite
  - b) La distribution de *t*
  - c) comparaison avec le test *z*
- 5.4: Tests paramétriques vs. tests non paramétriques
- 5.5: Biais et efficacité
  - a) Définition
  - b) Maximum de vraisemblance
  - c) Application pour la normale

Disponible sur: http://mapageweb.umontreal.ca/cousined/home/course/PSY1004

Varia

- TP1 corrigés disponibles.
- Semaine de lecture la semaine prochaine:
  - pas de cours…
  - les assistants seront à leur bureau jeudi
- Pour le TP2:
  - Quand vous utilisez SPSS, mettez en annexe la syntaxe ET le listing si vous voulez du feedback
  - N'oubliez pas votre code permanent ET une page titre.

Le TP1
Les résultats

- Les résultats:
  - moyenne:
  - écart type:
  - asymétrie:

Le TP1
Quelques solutions (1/3)

- De façon générale:
  - Respectez le nombre de lignes allouées
  - Un output SPSS n'est pas une réponse, le mettre en annexe.
  - Les graphiques doivent avoir un titre.

Le TP1
Quelques solutions (2/3)

- Question 5a)
  - La probabilité de trois succès est la probabilité d'obtenir trois fois un  , soit 1/6 x 1/6 x 1/6 = 1/216 = 0.00463
  - Idem pour les 7 autres lancés, soit $(5/6)^7$ = 78125/279936 = 0.2790
- Question 5b)
  - Le nombre de façon d'avoir trois succès et 7 échecs:
    - {+ + + − − − − − − −}   {+ − + + − − − − − −}
    - {+ + − + − − − − − −}   {+ − + − + − − − − −}
    - etc. donc 8 fois   etc. donc 7 fois   etc.
    - Grand total: 120
- Question 5c)
  - Les deux précédents ensembles: 120 x 0.00463 x 0.2790 = .155

-ou-

  - La probabilité d'avoir trois succès étant donné la distribution binomiale $\mathcal{B}(10, 1/6)$:

$$\binom{N}{r} p^r (1-p)^{N-r} = \binom{10}{3}\left(\frac{1}{6}\right)^3\left(1-\frac{1}{6}\right)^{10-3}$$
$$= 120 \times 0.0043 \times 0.2790$$
$$= 0.155$$

Le TP1
Quelques solutions (3/3)

- Question 9: Pour transformer la moyenne et l'écart type de données
  - a) normalisez les données en $\mathcal{N}(0,1)$ avec l'équation:

$$score' = (score - \mu)/\sigma$$

  - Dé normalisez les données en utilisant un autre μ et σ:

$$score'' = score' \;\; \sigma' + \mu'$$

  - Plusieurs façons de procéder existent: Excel, SPSS, etc. Dans SPSS:
    - Compute x1= (x – 100) / 5.
    - Compute x2= 10 * x1 + 20.

5.0: Tests de la tendance centrale

- Nous avons un test de la médiane:
  - Un test non-paramétrique, bidirectionnel, basé sur la distribution binomiale.
- Cependant, la plupart du temps, nous considérons la moyenne
  - La moyenne est plus connue
  - Est un meilleur indicateur de la tendance centrale (plus intuitif)
  - (la moyenne est non biaisée et efficace; voir 5.5)
- Peut-on avoir un test qui opère directement sur une moyenne?
  - Oui, il existe le test *z* et le test *t*.

5.1: Test *z* sur une moyenne a) survol

Postulats pour que le test soit valable:

- population normale (ou un grand nombre de facteurs…)
- écart type de la population σ connu ← Sic!

Test du genre:

rejet de H_0 si $\frac{|\bar{\mathbf{X}}-\mu|}{SE_\mu} > s(\alpha/2)$ (bilatéral) $\frac{\bar{\mathbf{X}}-\mu}{SE_\mu} > s(\alpha)$ (unilatéral)

- où SE_μ est l'erreur type de la moyenne $\frac{\sigma}{\sqrt{n}}$
- La valeur $\frac{\bar{\mathbf{X}}-\mu}{SE_\mu}$ est une moyenne normalisée (notée $\mathcal{N}(0,1)$) et se distribue normalement (par ex. est inférieure à 2 dans 97.3%)
- On choisi donc la valeur critère $s(\alpha)$ à l'aide d'une table $\mathcal{N}(0,1)$

5.1: Test *z* sur une moyenne b) Exemple 1

Soit un test de Q. I. chez les gauchers (on suppose que $\sigma = 10$)

a) H_0: $\mu = 100$, H_1: $\mu \neq 100$ {donc, bicaudal}

b) Seuil à 5% {a priori, avant de voir les données}

Collecte des données: $n = 16$, $\bar{\mathbf{X}} = 107$

c) Le Q. I. est normal et σ connu → $s(\alpha/2)$ lu sur $\mathcal{N}(0,1) = 1.96$, on utilise le test z: Rejet de H_0 si $\frac{|\bar{\mathbf{X}}-\mu|}{SE_\mu} > s(\alpha/2)$

d) Calcul et conclusion: "Le Q.I. moyen des gauchers est de 107 ± 2.5. Cette valeur diffère significativement de 100 ($\underline{z}$ = 2.80, $\underline{p}$ < .05)"

Rappel des résultats

Seuil α: < = significatif; > = non significatif

Un test a été fait

Il s'agit d'un test utilisant une table normale

Résultat du calcul

5.1: Test *z* sur une moyenne
c) Exemple 2

1 N = 1 newton, l'unité de force de frappe en physique.

Soit un entraînement sur le "uppercut" (on suppose que σ = 10 N)

a) H_0: μ = 70 N, H_1: μ > 70 N {donc, unicaudal}

b) Seuil à 1% {a priori, avant de voir les données}

Collecte des données: n = 25, $\overline{X} = 120$

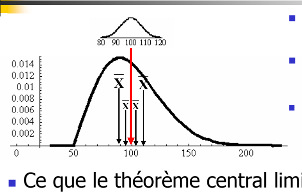
c) La force est normale et σ connu ➔ s(α) lu sur $\mathcal{N}$(0,1) = 2.32, on utilise le test *z*: Rejet de H_0 si $\frac{|\overline{X}-\mu|}{SE_\mu} > s(\alpha)$

d) Calcul et conclusion: "La force de frappe de l'uppercut est significativement supérieure à 70 après entraînement (z = 25.0, p < .01)"

5.2: Les limites du test *z*

- En recherche, 90% des expériences utilisent $\overline{X}$ comme statistique de la tendance centrale. Peut-on tester une moyenne avec le test *z*:
  - si σ inconnue? non
  - si pas distribution normale? non
  - Quand les **X** ne sont pas exactement normal, le test *z* n'est qu'approximativement valable.
  - De plus, σ ne fait pas parti de l'hypothèse, donc inconnu; cependant, on sait que $_{n-1}\overleftrightarrow{X}$ est un bon estimateur de σ (voir section 5.4). Peut-on utiliser l'écart type de l'échantillon?
- Le test *t* résout ces deux problèmes grâce à deux extensions au test *z* (toutes deux résolues au début du XX^{ième} siècle):
  - a) le Théorème central limite
  - b) la distribution *t*

5.3: Le test *t*
a) Théorème central limite



- Soit une distribution hypothétique $\mathcal{D}$, avec paramètres μ et σ.
- $\mathbf{X} \sim \mathcal{D}(\mu, \sigma)$
- $\mathcal{D}$ n'est pas normale, mais aussi pas trop asymétrique ($\overset{\nwarrow}{X}$ ~ 0) et pas trop de données extrêmes ($\overset{\curvearrowright}{X}$ ~ 3).

- Ce que le théorème central limite prouve, c'est que
  - même si les données brutes d'une population ne sont pas normales,
  - les moyennes d'échantillons tirées de cette population se répartissent de façon symétrique autour de la vraie moyenne
  - À la limite (quand la taille de l'échantillon est grande, n > 30), une moyenne se distribue exactement de façon normale.

5.3: Le test *t*
b) La distribution *t* (1/2)

Est normal en vertu du théorème central limite.

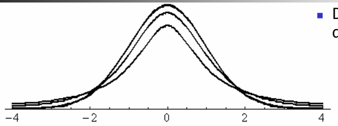
$$\frac{|\overline{\mathbf{X}}-\mu|}{SE_{\overline{\mathbf{X}}}} > s(\alpha)$$

N'est pas une valeur fixe, mais aléatoire. Elle varie d'un échantillon à l'autre, à cause d'erreurs d'échantillonnage.

$$où\ SE_{\overline{\mathbf{X}}} = \frac{{}_{n-1}\overleftrightarrow{\mathbf{X}}}{\sqrt{n}}$$

- Parce que l'erreur type de la population est estimé par l'erreur type de l'échantillon, le résultat de la division n'est pas exactement $\mathcal{N}(0,1)$.
  - Le résultat est proche, et d'autant plus proche que *n* >>.
  - Comme l'estimé de l'erreur type est parfois plus petit, parfois plus grand que dans la population, la division donne parfois des résultats plus petits, parfois plus grand, tendant à accroître l'épaisseur des queues de la distribution (distribution platycurtique, i.e. kurtose > 3).

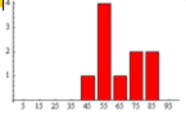
5.3: Le test *t*
b) La distribution *t* (2/2)



- Distribution *t* pour trois degrés de libertés:
  - $t(1)$
  - $t(3)$
  - $t(30)$ (kurtose ≈ 3) ≈ $\mathcal{N}(0,1)$
  - $t(\infty) = \mathcal{N}(0,1)$

- La distribution *t* est une distribution normale $\mathcal{N}(0,1)$ corrigée pour la kurtose selon le "degré de liberté" dl (soit le nombre de donnée) ➔ $t(\upsilon) \sim \mathcal{N}(0,1)(\upsilon)$
- Pris ensemble, le théorème central limite ET la correction de Student pour la division permettent de tester des hypothèses sur une moyenne dans un grand nombre de situations (postulats beaucoup plus généraux):
  - quand σ est inconnu (comme c'est presque toujours le cas);
  - quand *n* n'est pas très grand (comme c'est souvent le cas).

5.3: Le test *t*
Exemple 1

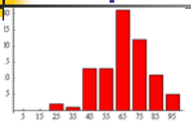


Moyenne	= 65.3
écart type	= 13.7
n	= 10

- a) Hypothèses H_0: $\mu = 70$, H_1: $\mu \neq 70$
- b) Seuil 5%
- c) Choisir le test:
  - Pas d'indication que les données sont normales, mais pas de déviation majeure…
  - σ est inconnue

  } Le test *z* ne s'applique pas
  - Le test *t* est parfait ici; on utilise $n - 1$ dl ➔ $s(\alpha/2) = 2.262$.
- On calcule: $\frac{|\overline{\mathbf{X}}-\mu|}{SE_{\overline{\mathbf{X}}}} \left(où\ SE_{\overline{\mathbf{X}}} = \frac{{}_{n-1}\overleftrightarrow{\mathbf{X}}}{\sqrt{n}} \right)$ = 1.08, on conclue

5.3: Le test *t*
Exemple 2



Moyenne = 65.7
écart type = 15.2
n = 100

- a) Hypothèses H_0: μ = 70, H_1: $\mu \neq 70$
- b) Seuil 5%
- c) Choisir le test:
  - Pas exactement normales, mais pas de déviation majeure…
  - Le test *t* est parfait ici; on utilise *n* – 1 dl ➔ s($\alpha/2$) = 1.984.
- On calcule: $\frac{|\bar{\mathbf{X}} - \mu|}{SE_{\bar{\mathbf{X}}}}$ $\left(où\ SE_{\bar{\mathbf{X}}} = \frac{{}_{n-1}\overleftrightarrow{\mathbf{X}}}{\sqrt{n}} \right)$ = 2.829 et on conclue.

5.3: Le test *t*
c) Comparaison avec le test *z*

- La différence entre les deux tests se trouve dans le choix de la valeur critique *s*(α). Regardons par exemple pour un test bicaudal avec un seuil de 5%:

n	test *t*	test *z*
2	12.70	1.96
5	2.78	1.96
10	2.26	1.96
30	2.04	1.96
100	1.98	1.96

- Même avec 100 données brutes dans l'échantillon, les valeurs critères ne sont pas encore les mêmes…

➔ Toujours préférable d'utiliser un test *t* lorsque les conditions du test *z* ne sont pas exactement satisfaits (autrement dit, ne jamais utiliser un test *z* sur une moyenne).

5.4: Tests paramétriques vs. tests non paramétriques (1/3)

- Tests paramétriques:
  - Évalue un paramètre de la population
  - Par exemple:
    - Si population affectée par un grand nombre de facteurs… donc, la population est $\mathcal{N}(\mu, \sigma)$ et H_0 : μ = 100
    - Si population est binaire… donc, la population est $\mathcal{B}(n, p)$ et H_0 : p = ½.
- Tests non paramétriques
  - N'évalue pas un paramètre de la population. En fait, suppose que la nature de la population (normal, binomial, Weibull) est inconnue ou visiblement étrange (très asymétrique, multimodale, beaucoup de données extrêmes)
  - Évalue une propriété de la population ou de la statistique
  - Par exemple:
    - Peu importe la population, 50% des données sont supérieures à la médiane (propriété de la médiane)
    - Peu importe la population, si je mesure deux fois le même individu dans les mêmes conditions, il aura plus ou moins le même score (50% plus/50% moins).

5.4: Tests paramétriques vs. tests non paramétriques (2/3)

	Tests paramétriques	Tests non paramétriques
Avantages:	Plus puissants (nécessitent un nombre d'observation n plus petits).	Moins puissants (nécessitent plus d'observations n pour être significatifs).
Désavantages:	Basés sur des postulats restrictifs (ex: σ connue)	Basés sur des postulats très généraux (plus facilement valides).

Privilégier les tests paramétriques à toutes les fois que possible, même si utiliser un test non paramétrique n'est pas une erreur.

Cependant (la contraposée), utiliser un test paramétrique si les postulats ne le permettent pas est une erreur.

5.4: Tests paramétriques vs. tests non paramétriques (3/3)

- Une grosse partie du travail de statistique est de savoir quel test est le meilleur dans une situation donnée.
  - Toujours utiliser le test s'appliquant le plus précisément à votre situation.
    - Par exemple, si population normale et σ connue ➔ Le test z est le meilleur (meilleur que le test t).
  - Un test paramétrique est toujours préférable sur un test non paramétrique, mais moins souvent applicable…
    - Par exemple, pour tester la tendance centrale, préférer un test t à moins que l'asymétrie soit très forte, qu'il y ait multimodalité, ou beaucoup de données extrêmes, auquel cas faire un test de la médiane.

5.5: Biais et efficacité a) définition (1/2)

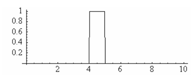
- Supposons une distribution $\mathcal{D}$ avec paramètre Ω: $\mathcal{D}(\Omega)$.
- Biais: Une statistique $\overset{?}{\mathbf{X}}$ est un estimateur sans biais du paramètre Ω si la valeur attendue de la statistique égale le paramètre (on note: $E(\overset{?}{\mathbf{X}}) = \mathbf{\Omega}$).
- Par exemple,
  - La moyenne $\overline{\mathbf{X}}$ est une statistique sans biais de μ pour la normal $\mathcal{N}(\mu, \sigma)$, (on note: $E(\overline{\mathbf{X}}) = \mu$).
  - L'écart type ${}_n\overleftrightarrow{\mathbf{X}}$ est une statistique biaisée de σ car à un n donné, le résultat est toujours trop petit. Cependant, on peut corriger le biais en utilisant ${}_{n-1}\overleftrightarrow{\mathbf{X}}$.

5.5: Biais et efficacité
a) définition (2/2)

- Supposons une distribution $\mathcal{D}$ avec paramètre Ω: $\mathcal{D}(\Omega)$.
- Efficacité: Une statistique $\overset{?}{\mathbf{X}}$ est un estimateur efficace d'un paramètre si la variance de cette statistique est minimale (On note: Var($\overset{?}{\mathbf{X}}$) est minimal).
- Par exemple:
  - La médiane et le mode sont aussi des statistiques sans biais de μ pour la normal $\mathcal{N}(\mu, \sigma)$ puisque la normal est symétrique..
    $E(\breve{\mathbf{X}}) = \mu \quad et \quad E(\dot{\mathbf{X}}) = \mu$
  - Cependant, le mode est extrêmement variable, et la médiane l'est aussi beaucoup. $Var(\dot{\mathbf{X}}) >> Var(\breve{\mathbf{X}}) > Var(\overline{\mathbf{X}})$
  - Pour cette raison, on préfère $\overline{\mathbf{X}}$ et $\sqrt{Var(\overline{\mathbf{X}})} = SE_{\overline{\mathbf{X}}} = \frac{{}_{n-1}\overleftrightarrow{\mathbf{X}}}{\sqrt{n}}$

5.5: Biais et efficacité
b) Maximum de vraisemblance (1/2)

- Pour évaluer si une statistique est efficace, nous avons la fonction de vraisemblance.
- Soit une population avec une distribution des données très simple: La probabilité d'observer une donnée entre Ω et $\Omega+1$ est de 100%.:



$\mathbf{X} \sim \mathcal{D}(\Omega)$ où Ω, le paramètre de la population est 4 ici.

- Si j'observe l'échantillon **X** = {4, 4, 4, 4, 4, 4, 4}, quelle est la vraisemblance que Ω = 4?
- Si j'observe l'échantillon **X** = {6, 6, 6, 6, …, 6}, quelle est la vraisemblance que Ω = 4?
- Si j'observe l'échantillon **X** = {6, 4, 4, 4, 4, 4} ?

5.5: Biais et efficacité
b) Maximum de vraisemblance (2/2)

- La fonction de vraisemblance donne un résultat entre 0 (paramètre totalement non plausible) et 1 (paramètre absolument adéquat)
- Pour y arriver, il faut une formule multiplicative, et utiliser la probabilité prédite par la distribution $\mathcal{D}(\Omega)$:

$$Vraisemblance(\Omega \mid \mathbf{X}) = \prod_{i=1}^{n} f(\mathbf{X}_i \mid \Omega)$$

- Dans l'exemple 1, **X** = {4, 4, 4, 4, 4, 4}, et
  Vraisemblance(4| **X**) = 1 × 1 × 1 × 1 × 1 × 1 = 1
- Dans l'exemple 3, **X** = {6, 4, 4, 4, 4, 4}, et
  Vraisemblance(4| **X**) = 0 × 1 × 1 × 1 × 1 × 1 = 0
- Pour trouver la meilleur valeur possible pour le paramètre Ω, on essaie toutes les valeurs possibles et retient celle qui maximise la Vraisemblance (qui l'amène le plus près de 1).

5.5: Biais et efficacité
c) Application pour la normal (1/2)

- Plutôt qu'une distribution $\mathcal{D}$, prenons une distribution plus utile, la distribution normal $\mathcal{N}(\mu, \sigma)$.
- On a que:

$$f(\mathbf{X}_i \mid \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{\mathbf{X}_i - \mu}{\sigma}\right)^2}$$

- Et donc, pour l'ensemble des données brutes de **X**:

$$Vraisemblance(\mu \mid \mathbf{X}) = \prod_{i=1}^{n} f(\mathbf{X}_i \mid \mu, \sigma)$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{\mathbf{X}_i - \mu}{\sigma}\right)^2}$$

$$= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{\mathbf{X}_i - \mu}{\sigma}\right)^2}$$

5.5: Biais et efficacité
c) Application pour la normal (2/2)

- On peut maximiser l'équation en utilisant une dérivée et en solvant pour une pente égale à zéro. On obtient que le plus efficace μ est donné par:

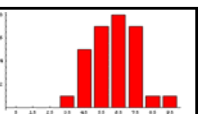
$$\mu = \frac{1}{n}\sum_{i=1}^{n} \mathbf{X}_i$$

- Si on fait la même chose pour le paramètre σ, on obtient que le plus efficace σ est donné par:

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n} (\mathbf{X}_i - \mu)^2$$

- Malheureusement, la fonction de vraisemblance ne se résout pas pour toutes les distributions (ex. la Weibull).

5.6: Un dernier exemple



- Soit ces données **X** = {35, 42, 43, 49, 50, 50, 53, 54, 54, 55, 56, 57, 57, 62, 62, 63, 63, 65, 66, 68, 70, 72, 73, 78, 78, 79, 80, 80, 89, 90}
- Explorez les données; sont-elles loufoques?
- Faites un test pour savoir si la moyenne peut être de 70?
  - a) Écrivez vos hypothèses — H_0: $\mu = 70$; H_1: $\mu \neq 70$
  - b) Choisissez le seuil de décision — $\alpha = .05$
  - Calculez moyenne/écart type pour le test — 63.27, 13.85
  - c) Choisissez le test et la valeur critique — test t, s($\alpha/2$) = 2.045
  - d) Calculez et Concluez. — "La moyenne obtenue est de 63.27 ± 2.53. Elle diffère significativement de 70 ($t(29) = 2.66$, $p < .05$)."

Nouveauté ici: Quand un test possède des degrés de liberté, il faut les écrire entre parenthèse