Exemple, Distribution de permutation

- Considérons deux échantillons:
- $m = 4$, $n = 2$;
- Les mesures sont les suivantes pour le premier groupe: $X_1 = 2$, $X_2 = -1$, $X_3 = 4$, $X_4 = 0$
- Pour le second groupe: $Y_1 = 3$, $Y_2 = 1$;
- Le test est basé sur la statistique-t:

$$t = \frac{\bar{Y} - \bar{X}}{s_p \sqrt{\frac{1}{n} + \frac{1}{m}}}$$

- Compte tenu de la forme de la distribution, il y a donc:

$$\binom{6}{4}\binom{2}{2} = 15$$

cas à considérer.

- Dans la table qui suit, chaque ligne correspond, compte tenu des permutations à l'intérieur des groupes, à

$$2!\ 4! = 48 \text{ permutations.}$$

- On a alors la distribution suivante:

-1	0	1	2	3	4	-2.954
-1	1	0	2	3	4	-1.585
-1	2	0	1	3	4	-0.910
-1	3	0	1	2	4	-0.423
-1	4	0	1	2	3	0.000
0	1	-1	2	3	4	-0.910
0	2	-1	1	3	4	-0.423
0	3	-1	1	2	4	0.000
0	4	-1	1	2	3	0.423
1	2	-1	0	3	4	0.000
1	3	-1	0	2	4	0.423
1	4	-1	0	2	3	0.910
2	3	-1	0	1	4	0.910
2	4	-1	0	1	3	1.585
3	4	-1	0	1	2	2.954

- Ceci donne la distribution de permutation suivante:

-2.954	$\frac{1}{15}$
-1.585	$\frac{1}{15}$
-0.910	$\frac{2}{15}$
-0.423	$\frac{2}{15}$
0	$\frac{3}{15}$
0.423	$\frac{2}{15}$
0.91	$\frac{2}{15}$
1.585	$\frac{1}{15}$
2.594	$\frac{1}{15}$

- Prenons $\alpha = 10\%$ et effectuons un test unilatéral à droite.
- Puisque:

$$\mathcal{P}(T \geq 2.954) = \frac{1}{15} = 0.067 < 0.1;$$
$$\mathcal{P}(T \geq 1.585) = \frac{2}{15} = 0.133 > 0.1,$$

  le quantile est alors $q^{>}_{0.10} = 2.954$.
- Puisque $s = 0.423$ est une valeur inférieure à 2.954, on ne rejette pas l'hypothèse nulle au niveau $\alpha = 0.10$.

Comparaisons avec le t-test classique

- On note que si l'on avait fait le test classique, le quantile serait provenu de la loi t_{n+m-2}. Or avec $n = 2$, $m = 4$, on trouve que $t_{n+m-2;0.90} = 1.533$ et dans ce cas précis on aurait donc pris la même décision (dans le langage R: qt(0.90, 4)).
- Sous la distribution de permutation, la valeur-p vaut:

$$\mathcal{P}(T \geq 0.423) = \frac{6}{15} = 0.4;$$

Discussion

- Dans cet exemple, on a calculé la distribution de permutation de la statistique-t classique;
- En général, on a entière liberté de choisir la statistique désirée;
- L'idée est de choisir une statistique que l'on estime sensible aux hypothèses;
- Autrement dit, une statistique qui devrait prendre des valeurs centrales sous H_0 et éloignées sous l'alternative H_1.

Avantages d'un test de permutation

- L'avantage primordial de recourir à un test de permutation est de pouvoir contrôler le risque de première espèce α (autrement dit le niveau du test).
- Plus précisément, sous des hypothèses minimes sur la loi (autrement dit ce que l'on a noté F), le niveau est contrôlé, quelque soit la loi régissant les observations.