

Comment utiliser les DGLM?

Martin Storme

24/06/13

Problématique

- Tester l'effet d'un entraînement sur la fidélité (stabilité temporelle) des évaluations de la créativité
- Etudier la stabilité temporelle:
 - Consistance (ex. coefficient de corrélation de Bravais Pearson)
 - Concordance (ex. coefficient de concordance de Lin)
- Utile de pouvoir étudier séparément concordance et consistance (même si les deux sont liées)

Consistance et concordance

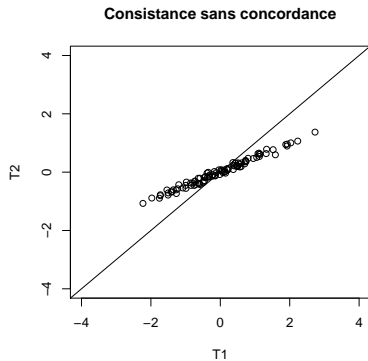
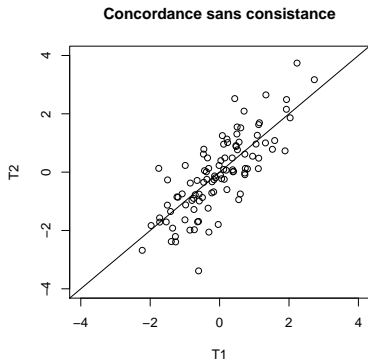


Figure : Consistance et concordance

Exemple avec des données fictives

On simule les évaluations de deux individus fictifs (1 dans chaque condition, 500 évaluations par individu).

Pour simplifier, T2 est construite à partir de T1:

$$T2 = B0 + B1.T1 + \epsilon \quad (1)$$

Où:

- T1 suit une loi normale dont la moyenne est 0 et l'écart-type 1,
- B0=0 et B1=1,
- et ϵ suit une loi normale dont la moyenne est 0 et l'écart-type est fonction de l'entraînement: Avec=0.02, Sans=0.2

Représentation graphique

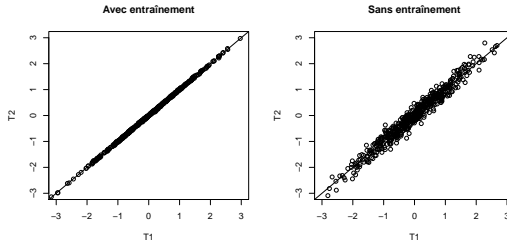


Figure : Représentation graphique 1

Ajuster un modèle par condition

On étudie le modèle ($T2 \sim T1+1$) pour chacun des deux individus
On trouve approximativement les mêmes valeurs pour l'ordonnée à l'origine et la pente.

La différence s'observe dans la variance résiduelle:
 $\text{var}(\text{res}(\text{lm}(T2 \sim T1)))$

- Avec entraînement: 0.0004
- Sans entraînement: 0.0391

Principe du DGLM

- Modéliser simultanément la moyenne et la dispersion
- Utile pour :
 - limiter le biais de l'estimation lorsque les résidus ne sont pas homogènes
 - faire des hypothèses sur la dispersion
- Package R: dhglm (Noh & Lee, 2011)

Première idée

Entraînement	Observation	T1	T2	phi
0	1	0.34	0.57	1
0	2	-2.01	-1.04	1
...
1	501	2.45	2.53	1
1	502	0.28	0.09	1
...

Table : Idée 1: Organisation des données

Sortie R et interprétation

Modèle	Paramètre	Estimation	Err. Std.	Valeur T
Moyenne	Intercept	0.00	.00	0.75
	T1	1.00	.00	1053.96
Dispersion	Intercept	-3.18	.06	
	Entraînement	-4.60	.08	

Table : Estimation du modèle 1 (VD=T2)

Sortie R et interprétation (suite)

- Moyenne: Intercept = 0, Pente = 1
- Dispersion: $\exp(-3.18)=0.0416$ et $\exp(-3.18-4.60)=0.0004$

Deuxième idée

Entraînement	Observation	Temps	Créativité	phi
0	1	1	0.34	1
0	1	2	0.57	1
...
1	501	1	2.45	1
1	501	2	2.53	1
...

Table : Idée 2: Organisation des données

Sortie R et interprétation

Modèle	Paramètre	Estimation	Err. Std.	Valeur T
Moyenne	Intercept	-0.01	0.04	-0.39
	Entraînement	0.03	0.04	0.59
	Temps	-0.00	0.04	-0.00
Dispersion	Intercept	0.03	.05	
	Entraînement	-0.13	.06	
	Temps	0.03	.06	

Table : Estimation du modèle 2 (VD=Note de créativité)

Représentation graphique

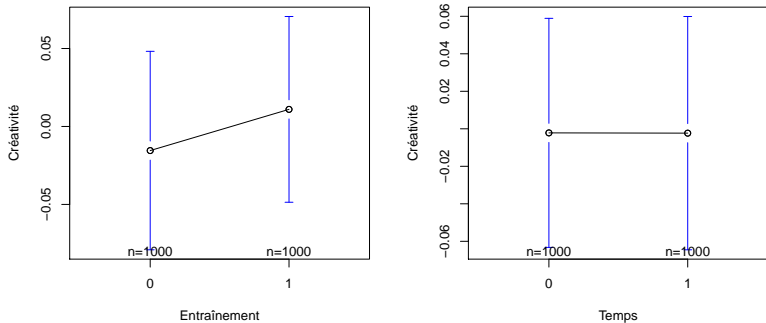


Figure : Représentation graphique 2

Question

Comment utiliser le DGLM pour étudier la question posée dans la première approche (mais sans faire d'erreur)?