

TD ANALYSE CONJOINTE

CARTOGRAPHIE DE PREFERENCE

Relations attributs sémantiques – attributs de design

Les données nécessaires au TD se trouvent dans le fichier Excel TD_AC-avec_ANOVA.xls.

Le TD s'effectue sur Matlab. Les fonctions utiles sont indiquées pour information.

NB : Utiliser l'AIDE Matlab pour la mise en œuvre des fonctions.

Exercice 1 : exemple simple d'application de l'analyse conjointe en utilisant l'analyse de la variance pour déterminer les paramètres du modèle.

On considère le produit « montre » avec deux facteurs : le bracelet (à trois niveaux : cuir – plastique – métal) et le type de cadran (à deux niveaux : aiguille – digital).

On considère une notation de préférence effectuée sur le plan complet (à $2 \times 3 = 6$ designs) répété une fois (2 évaluations), c'est à dire 12 observations (onglet Exercice 1).

1) Les résultats de la modélisation analyse conjointe de la préférence par un modèle linéaire sont les suivants (onglet ANOVA-exo1) : les éléments importants sont surlignés en jaune.

Coefficients d'ajustement :

Observations	12,000
Somme des poids	12,000
DDL	8,000
R²	0,932
R ² ajusté	0,906
MCE	0,896
RMCE	0,946
MAPE	26,498
DW	2,647
Cp	4,000
AIC	1,814
SBC	3,754
PC	0,137
Press RMCE	1,420

ANALYSE DE LA VARIANCE

Source	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
Modèle	3	97,833	32,611	36,403	< 0,0001
Erreur	8	7,167	0,896		
Total corrigé	11	105,000			

Analyse Type III Sum of Squares :

Source	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
Bracelet	2	1,500	0,750	0,837	0,468
cadran	1	96,333	96,333	107,535	< 0,0001

Paramètres du modèle :

Source	Valeur
Constante	4,500
Bracelet-m	-0,250

CEOAC

Bracelet-c	-0,250
Bracelet-p	0,500
cadran-a	-2,833
cadran-d	2,833

Equation du modèle :

$$\text{préférence } y = 4,5 - 0,25 \cdot \text{Bracelet-m} - 0,25 \cdot \text{Bracelet-c} + 0,5 \cdot \text{Bracelet-p} - 2,83 \cdot \text{cadran-a} + 2,83 \cdot \text{cadran-d}$$

Interprétation des résultats

- la constante (4,5) représente la préférence moyenne de toutes les évaluations (cette propriété est due au fait qu'on a utilisé le paramétrage centré : on vérifiera que, pour chaque facteur, $\sum a_i = 0$ pour les coefficients du modèle)
- La part de variance expliquée par le modèle est de 93% ($R^2=0,932$). Le modèle est très bien adapté pour décrire les données. On peut donc interpréter les résultats sans risque d'erreur.
- La régression est significative dans son ensemble (p-value = <0,0001). Au moins un facteur apporte une contribution significative pour expliquer la préférence.
- Le facteur « bracelet » n'est pas significatif (P-value = 0,468). Il n'apporte pas une contribution significative pour expliquer la préférence. Mais on peut quand même interpréter son influence, et même le laisser dans le modèle.
- Le facteur « cadran » est significatif (P-value < 0,00001). Ce facteur explique la variation de préférence.
- Les utilités partielles¹ des niveaux des facteurs sont directement les paramètres du modèle

On peut alors construire le tableau de synthèse qui calcule les amplitudes et les importances de chaque facteur :

Coefficient R2	Facteur	Niveau	utilité partielle	Amplitude	Importance (%)
R ² = 0,932	Bracelet	m	-0,25	0,750	11,69%
		c	-0,25		
		p	0,5		
	Cadran	a	-2,833	5,667	88,31%
		d	2,833		

Rappel : l'importance d'un facteur j est donnée par :

$$I_j(\%) = \frac{MAX_k(a_{jk}) - MIN_k(a_{jk})}{\sum_{i=1}^p (MAX_k(a_{ik}) - MIN_k(a_{ik}))}$$

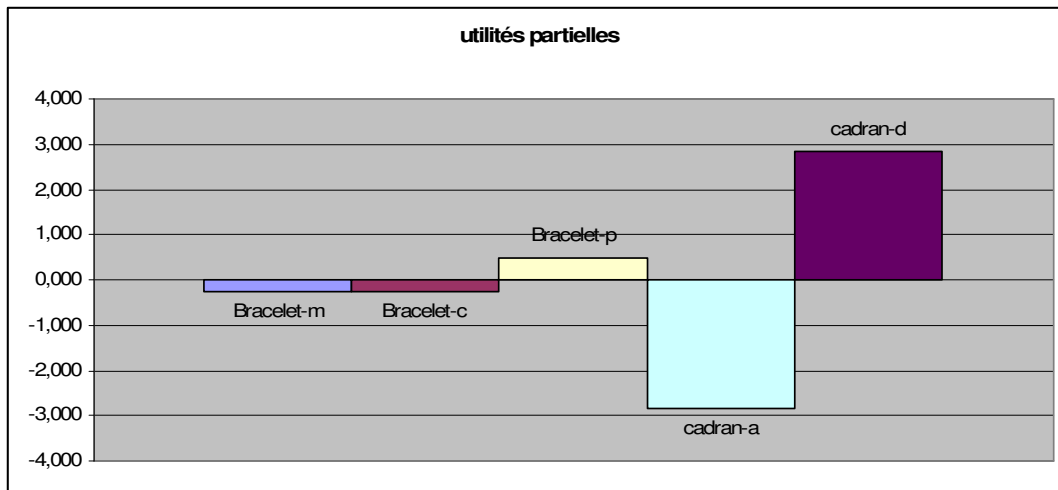
avec p : nombre de facteurs

a_{jk} : utilité partielle du niveau k du facteur j

L'importance du cadran pour expliquer la préférence est primordiale (88,31%) par rapport au bracelet (11,69%). Le cadran digital (d) tire la préférence (utilité partielle de +2,833), le cadran à aiguille (a) la repousse.

Le sujet est principalement sensible au type de cadran, avec une préférence pour le cadran digital. Dans une moindre mesure, le bracelet plastique tire la préférence, les autres modalités sont moins appréciées.

¹ Les utilités partielles sont directement les valeurs données dans le tableau des paramètres du modèle. On choisira les paramètres **non normalisés**. Les coefficients normalisés sont ceux que l'on obtiendrait avec des données d'entrée centrées réduites (de moyenne nulle et d'écart type = 1). Le tableau des coefficients normalisés (aussi appelés coefficients bêta) permet de comparer le poids relatif des facteurs. Plus la valeur absolue d'un coefficient est élevée, plus le poids de la variable correspondante est important. Lorsque l'intervalle de confiance autour des coefficients normalisés comprend la valeur 0 (cela est facilement visible sur le graphique des coefficients normalisés), le poids d'une variable dans le modèle n'est pas significatif.



2) Retrouver tous ces résultats en utilisant Matlab.

Voir la fonction matlab fournie TD_AC.m

Rappel :

- On choisira un paramétrage centré (contrainte $\sum a_i = 0$ pour les coefficients du modèle)
- on choisira un modèle sans interaction
- on activera le test de Fisher pour tous les facteurs

Fonctions Matlab utile

Anova avec n facteurs (n-way ANOVA) :

```
[p,tbl,stats,terms]= anovan(y,group,Name,Value)
```

Exercice 1bis : les données de préférence du sujet sont les suivantes (cf. onglet exercice1) :

Réponse	Facteurs	
preference y	Bracelet	cadran
2	métal	aiguille
8	cuir	digitale
3	plastique	aiguille
2	plastique	digitale
2	métal	digitale
9	cuir	aiguille
2	métal	aiguille
7	cuir	digitale
2	plastique	aiguille
3	plastique	digitale
2	métal	digitale
8	cuir	aiguille

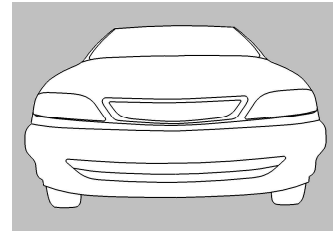
Refaire l'ANOVA et interpréter les résultats. Etait-il possible de prédire les conclusions uniquement en examinant les cotations de préférence ?

VOUS POUVEZ PASSER À L'EXERCICE 2 LORSQUE VOUS AVEZ RETROUVÉ ET COMPRIS TOUS LES RÉSULTATS DE L'EXERCICE 1 ET INTERPRETE L'EXERCICE 1BIS

Exercice 2 : Face avant de voiture

On s'intéresse à la perception d'avant de voitures par un ensemble de sujets. Une étude de préférence a été réalisée sur 12 voitures (voir dernière page pour les images des voitures).

On propose d'analyser les préférences avec l'analyse conjointe en les reliant aux attributs du design.



Une étude morphologique a permis de sélectionner les attributs du design influents (table 1). 3 facteurs ont été sélectionnés.

- Le type de phare : facteur **A** à 3 niveaux A1, A2, A3
- Le type de calandre : facteur **B** à 2 niveaux B1, B2
- Le type d'entrée d'air : facteur **C** à 2 niveaux C1, C2

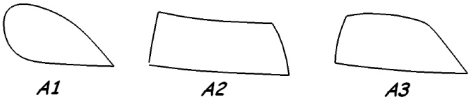
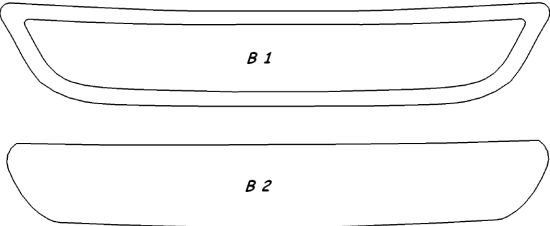
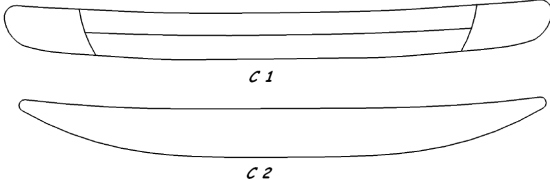
 Facteur A : phares	 Facteur B : calandre
 Facteur C : entrée d'air	

Table 1 : les 3 facteurs et leurs niveaux

Les 12 designs du plan complet (12 produits) ont été modélisés avec un logiciel de DAO (voir dernière page). Une épreuve de notation de la préférence des produits a conduit aux résultats donnés dans l'onglet exercice 2.

- 1) A l'aide des photos présentées en dernière page, évaluez vos préférences pour les 12 designs, sur une échelle de 1 (détesté) à 10 (préféré) (ceci constituera les données du sujet *moi*). Reportez ces notes dans Excel.
- 2) On s'intéresse au modèle explicatif linéaire sans interaction, expliquant la notation de préférence des voitures par les 3 facteurs A, B et C (analyse conjointe).
 - a. Pour le sujet S1, ajuster un modèle linéaire sans interaction expliquant ses notes de préférence par les 3 facteurs caractérisant les designs. Utiliser la fonction `anovan`
 - b. Déterminer la qualité d'ajustement du modèle sur les données (coefficient R^2). Déterminer les utilités partielles relatives à chaque niveau des facteurs ainsi que les importances de chaque facteur. Remplir le tableau de synthèse suivant. Conclusions concernant l'explication de la préférence ?

Coefficient R^2	Facteur	Niveau	utilité partielle	Amplitude	Importance (%)
	A : phares	A1			
		A2			
		A3			
	B : calandre	B1			
		B2			
	C : entrée d'air	C1			
		C2			

- a. Même question pour le sujet S2. L'ajustement des données est-il correct ? proposer si besoin un modèle avec interaction. Interpréter l'interaction et conclure sur l'influence des facteurs sur les préférences de ce sujet.
- b. Mêmes questions pour le sujet *moi*, correspondant à vos préférences. Si vous le souhaitez, vous pouvez « jouer avec vos préférences », en changeant les valeurs en fonction du design, et en observant comment le modèle explique vos choix.

Exercice 3 : Etude de la sportivité d'un tableau de bord numérique²

Les tableaux de bord numériques utilisent un écran TFT (Thin Film Transistor) et sont facilement paramétrables (figure 1).

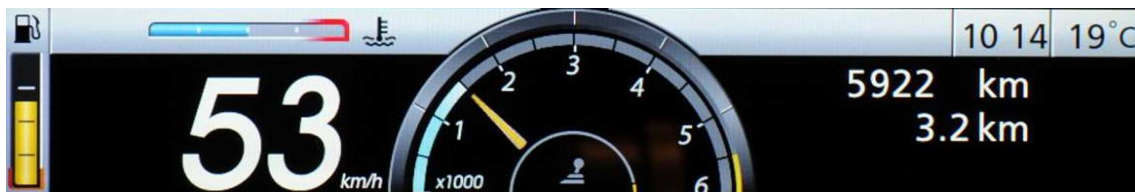


Figure 1 : Tableau de bord en écran TFT : exemple de style de l'affichage

On s'intéresse à l'influence de certains attributs de design sur la sportivité du tableau de bord. 8 facteurs (V1 à V8) ont été déterminés pour constituer le plan expérimental. Leur définition est donnée dans le tableau suivant.

² Ce travail a été réalisé dans le cadre d'un contrat d'étude avec RENAULT

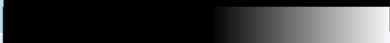
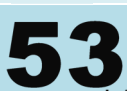
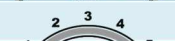
Facteur	Niveau	Description	Image
V1 : Fond	a1	Clair	
	a2	Foncé	
	a3	Dégradé	
V2 : Bandeau	b1	Clair	
	b2	Foncé	
	b3	Dégradé	
V3 : Thème	c1	Neutre	
	c2	Orange	
	c3	Turquoise	
V4 : Police	g1	Épaisse (Arial Black)	
	g2	Fine (Bookman Old Style)	
V5 : Essence	d1	Barre	
	d2	Analogique simple	
	d3	Analogique plein	
V6 : Vitesse	f1	Numérique	
	f2	Analogique simple	
	f3	Analogique détaillé	
V7 : Compte-tours	h1	Barre graphe	
	h2	Analogique simple	
	h3	Analogique détaillé	
	h4	Absent	
V8 : Température	e1	Barre	
	e2	Barre sectionnée	
	e3	Absent	

Tableau 1 : Niveaux des différents facteurs du tableau de bord

En supposant un modèle linéaire additif sans interaction entre la réponse et les niveaux des facteurs, un plan d'expérience D-optimal à 32 designs a été déterminé. La cotation de sportivité de ces 32 designs sur une échelle quantitative de 0 (pas sportif) à 10 (très sportif) a été effectuée par un panel de 30 sujets (onglet exercice 3).

- a) Déterminer, pour chaque sujet, les utilités partielles relatives à chaque niveau des facteurs, ainsi que les importances des facteurs. Tableau de synthèse. Le modèle s'ajuste-t-il correctement sur les données pour tous les sujets ?
- b) Visualiser le consensus dans le groupe de sujets, sur les utilités partielles, puis sur les importances. Pour cela, tracer les box plot des importances, et des utilités partielles.
- c) Ecarter les sujets pour lesquels le modèle linéaire additif sans interaction est inadapté (coefficient R^2 inférieur à 0,7). En considérant les sujets restants :
 - a. Calculer la moyenne des utilités de chaque niveau des facteurs. Sous l'hypothèse d'additivité des effets, comment se caractériserait le produit idéal (le plus sportif) pour l'ensemble du groupe ? le produit « anti-idéal (le moins sportif) ? Quelles sont les conclusions concernant la sportivité du tableau de bord ?
 - b. Calculer la moyenne des importances à partir des importances individuelles³. Tirer toutes les conclusions pertinentes concernant l'influence des facteurs de design sur la sportivité. Quelles sont les limites d'une étude sur les valeurs moyennes des utilités et des importances ?
 - c. On considère deux produits donnés par la valeur de leurs modalités : produit A(a1, b1, c1, d1, e1, f1, g1, h1) et produit B (a2, b2, c2, d2, e2, f2, g2, h2). En utilisant chaque modèle individuel, déterminer le nombre de fois que le produit A surclasse (utilité totale supérieure) le produit B. Donner la réponse en pourcentage. Conclusion ?

Voir le cours .ppt où une analyse similaire a été effectuée sur les jus d'orange (slides 182 à 193).

- d) Effectuer une Classification ascendante hiérarchique des sujets, selon les 8 variables « importance des facteurs » ; « méthode d'agrégation », choisir « méthode de Ward ».
- e) Interpréter les différences entre les classes
Pour cela, calculer la moyenne des importances par classe, la moyenne générale, et analyser les différences pour typer les classes.

Fonctions Matlab utile

Anova avec n facteurs (n-way ANOVA) :

```
p, tbl, stats, terms] = anovan(y, group, Name, Value)
```

tracé de boxplot :

```
boxplot(y, group)
```

Classification ascendante hiérarchique :

```
Z = linkage(X, 'ward');  
dendrogram(Z, 'Labels', marque)
```

³ En AC, il est conseillé de calculer les importances moyennes à partir des importances individuelles. En effet, un calcul des importances moyennes à partir des utilités moyennes risquerait d'annuler l'effet d'un facteur si celui-ci est l'objet d'un important désaccord dans le panel de sujets (si on suppose par exemple que la moitié des sujets dans le panel préfère fortement la marque A par rapport à B alors que l'autre moitié préfère fortement la marque B/A, la moyenne des utilités partielles individuelles des deux marques sera nulle, donc l'importance sera nulle aussi. Alors que la moyenne des importances individuelles sera non nulle, ce qui correspond à la réalité.