

Exercice de calage

Modèle de répartition modale fer-avion "Prix-temps"

L'exercice de calage du modèle prix-temps sert de support à la présentation du cours « apports méthodologiques ». Ce cours, pour des étudiants n'ayant pas encore une grande pratique de la modélisation, peut paraître assez théorique. C'est pourquoi, il est important de s'appuyer sur un cas concret qui permet de faire comprendre les différentes notions qui sont introduites. Le cours peut être approfondi à l'aide du chapitre 2 de Bonnel P (2004), *Prévoir la demande de transport*, éditions des Ponts et Chaussées, Paris, qui est disponible en version électronique sur le réseau (pédagogie prof/MOTUS).

Ce document présente l'origine des données (section 1), le principe du modèle et la méthode de calage de ce modèle (section 2). Enfin, nous récapitulons les données nécessaires (section 3). Ces données sont déjà saisies dans le fichier exocalag.xls.

L'exercice consiste à caler le modèle prix-temps avec le jeu de données fourni.

1. DONNEES DE BASE

Nous travaillons sur le modèle prix-temps appliqué au partage modal entre le fer et l'air. Nous avons choisi ce modèle pour plusieurs raisons :

- c'est un modèle qui a été utilisé pour des prévisions de trafic en non-urbain, notamment à la SNCF pour les prévisions TGV ;
- c'est un modèle utilisé également en urbain, même s'il est appliqué non pas pour le choix de mode, mais plutôt pour l'étape d'affectation (choix d'itinéraire) ;
- c'est un modèle très simple dans son principe, donc facilement accessible par des étudiants, nécessitant peu de données pour être appliqué. Il fait référence à des notions mathématiques et statistiques peu sophistiquées pour ceux qui ont des bases en mathématiques et statistiques ;
- les résultats se prêtent à une analyse facile à faire partager ;
- le modèle est une très bonne base pour illustrer tous les éléments théoriques et méthodologiques introduits dans le cours.

Les données de base proviennent pour l'essentiel de la SNCF et de Air France (Air Inter à l'époque où les données ont été collectées). Toutefois, pour des raisons commerciales, certaines données ont été modifiées. Les modifications ont été introduites de telle sorte que les données conservent leur signification et leur cohérence permettant de mener l'analyse de manière pertinente. En revanche, il est exclu de vouloir utiliser ces données pour faire de la prévision (même s'il est possible d'utiliser les données à des fins de simulations pédagogiques).

2. EXERCICE

L'exercice proposé consiste à caler le modèle prix-temps. Cela suppose de rentrer les formules permettant de calculer d'une part le logarithme népérien de la valeur d'indifférence du temps et d'autre part la transformée de la part de marché du fer

Il faut ensuite faire la régression sur ces deux vecteurs. L'activation du module de régression linéaire s'effectue de la manière suivante :

- bouton Office/options excel/compléments, puis sélectionner Analysis toolpak/atteindre ; dans la nouvelle fenêtre, sélectionner de nouveau Analysis toolpak, puis valider (confirmer de nouveau l'installation si le module n'a pas été déjà installé). Dans la barre de menu sélectionner Données/utilitaire d'analyse et choisir régression linéaire.

Enfin, il faut faire l'analyse des résultats de l'application du modèle théorique en analysant les résidus (comparaison des parts de marché théoriques fournies par le modèle prix-temps avec les parts de marché observées calculées sur les données disponibles).

2.1. Principe du modèle

Ce modèle repose sur l'hypothèse que le choix d'un voyageur entre le train et l'avion s'effectue en fonction de la valeur que celui-ci attribue à son temps, et des caractéristiques de coût et de temps de transport de chacun des modes (Abraham et al., 1969, Cf. bibliographie du manuel). Ainsi, l'utilisateur k choisit le mode dont le coût généralisé, compte tenu de sa valeur du temps h_k , est le plus faible.

Si P_F et P_A sont les prix respectifs du fer et de l'avion, et si T_F et T_A sont les durées de trajet (y compris terminaux), les coûts généralisés pour l'utilisateur k sont définis par :

$$Cg^k_A = P_A + h_k T_A \quad (1)$$

$$Cg^k_F = P_F + h_k T_F \quad (2)$$

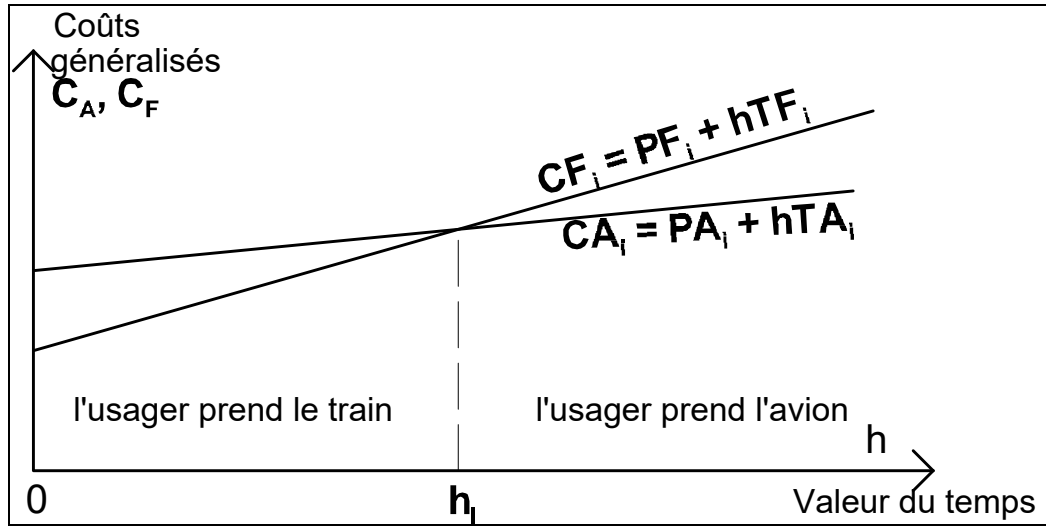
Ces coûts généralisés sont illustrés graphiquement (graphique 1).

Sur une relation donnée i , il existe une valeur du temps h_i telle que :

$$Cg^i_A = Cg^i_F$$

$$\text{soit : } h_i = \frac{P_A - P_F}{T_F - T_A} \quad (3)$$

qui est appelée valeur d'indifférence du temps sur la liaison i . Si $h_k < h_i$, le voyageur k choisit le fer, sinon l'avion.



Graphique 1 : Courbe de coût généralisé fer-avion

On suppose que la population des voyages est caractérisée par une distribution de la valeur du temps des voyageurs $f(h)$. La fonction de répartition :

$$F(h) = \int_0^h f(x)dx \quad (4)$$

est égale à la proportion de voyages dont la valeur du temps est inférieure à h .

Dans ces conditions, si h_i désigne la valeur d'indifférence sur la liaison i , la proportion X_i d'utilisateurs du train dans le trafic total sera donnée par :

$$X_i = F(h_i) = \int_0^{h_i} f(x)dx \quad (5)$$

et la proportion Y_i d'utilisateurs de l'avion dans le trafic total sera donnée par :

$$Y_i = 1 - F(h_i) = \int_{h_i}^{\infty} f(x)dx \quad (6)$$

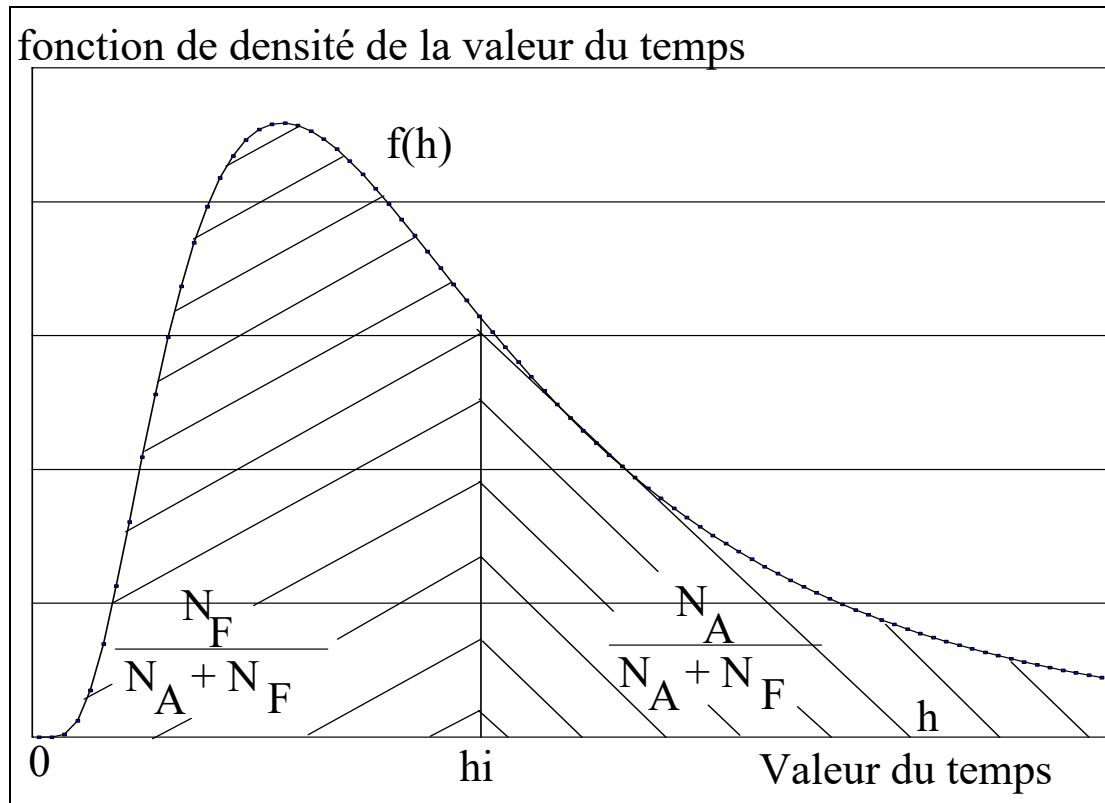
Ces équations sont illustrées par les deux graphiques 2 et 3 (pour une distribution log-normale).

Soit N_A = trafic avion en valeur absolue ;

et N_F = trafic fer 1ère + 2ème classes en valeur absolue,

$$X_i \text{ est donc égal à : } \frac{N_F}{N_A + N_F}, \quad (7)$$

$$\text{et } Y_i \text{ est donc égal à : } \frac{N_A}{N_A + N_F} \quad (8)$$



Graphique 2 : Forme de la fonction de densité log-normale

La forme de la fonction de distribution des valeurs du temps n'étant pas connue, on pose par hypothèse une analogie avec la distribution des revenus dans la population. On a donc retenu une fonction de densité $f(h)$ log-normale (Calot G (1975), *Cours de statistique descriptive*, 2ème édition, Dunod, Paris ; Dagnelie P (1973), *Théorie et méthodes statistiques, vol 1*, 2ème édition, Les presses agronomiques de Gembloux, Bruxelles), c'est-à-dire :

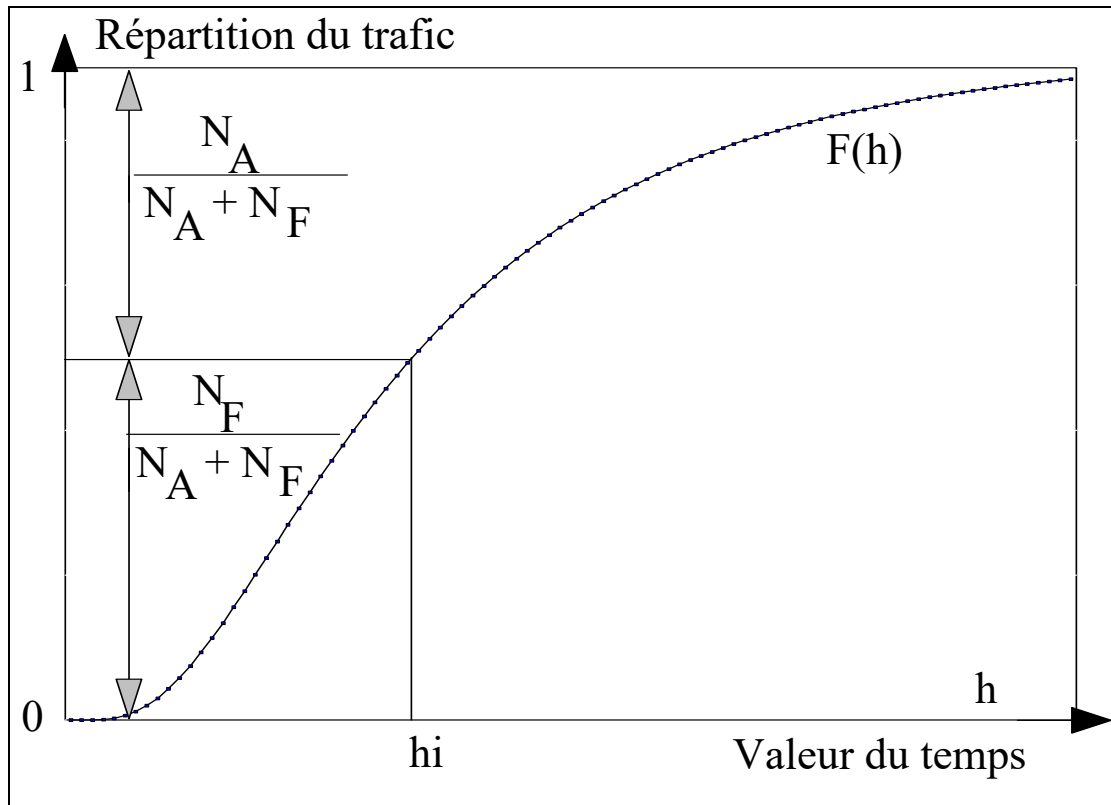
$$f(h) = \frac{1}{h\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln(h)-m)^2}{2\sigma^2}\right) \quad (9)$$

où σ est l'écart-type du logarithme népérien des valeurs individuelles du temps, et m la moyenne du logarithme népérien des valeurs individuelles du temps, sont les deux inconnues de la distribution log-normale.

La fonction de répartition est de la forme suivante :

$$F(h) = \Phi\left(\frac{\ln(h)-m}{\sigma}\right) \quad (10)$$

où Φ est la distribution cumulative de la loi normale standard ;



Graphique 3 : Forme de la fonction de répartition log-normale

2.2. Ajustement du modèle

L'ajustement du modèle consiste à caler les paramètres σ (écart-type du logarithme népérien des valeurs individuelles du temps) et m (moyenne du logarithme népérien des valeurs individuelles du temps) de la distribution log-normale.

Si l'on prend l'inverse de la fonction de répartition l'équation 10 devient :

$$\Phi^{-1}(F(h)) = \frac{(\ln(h) - m)}{\sigma} \quad (11)$$

Cette fonction est appelée la transformée de Laplace-Gauss.

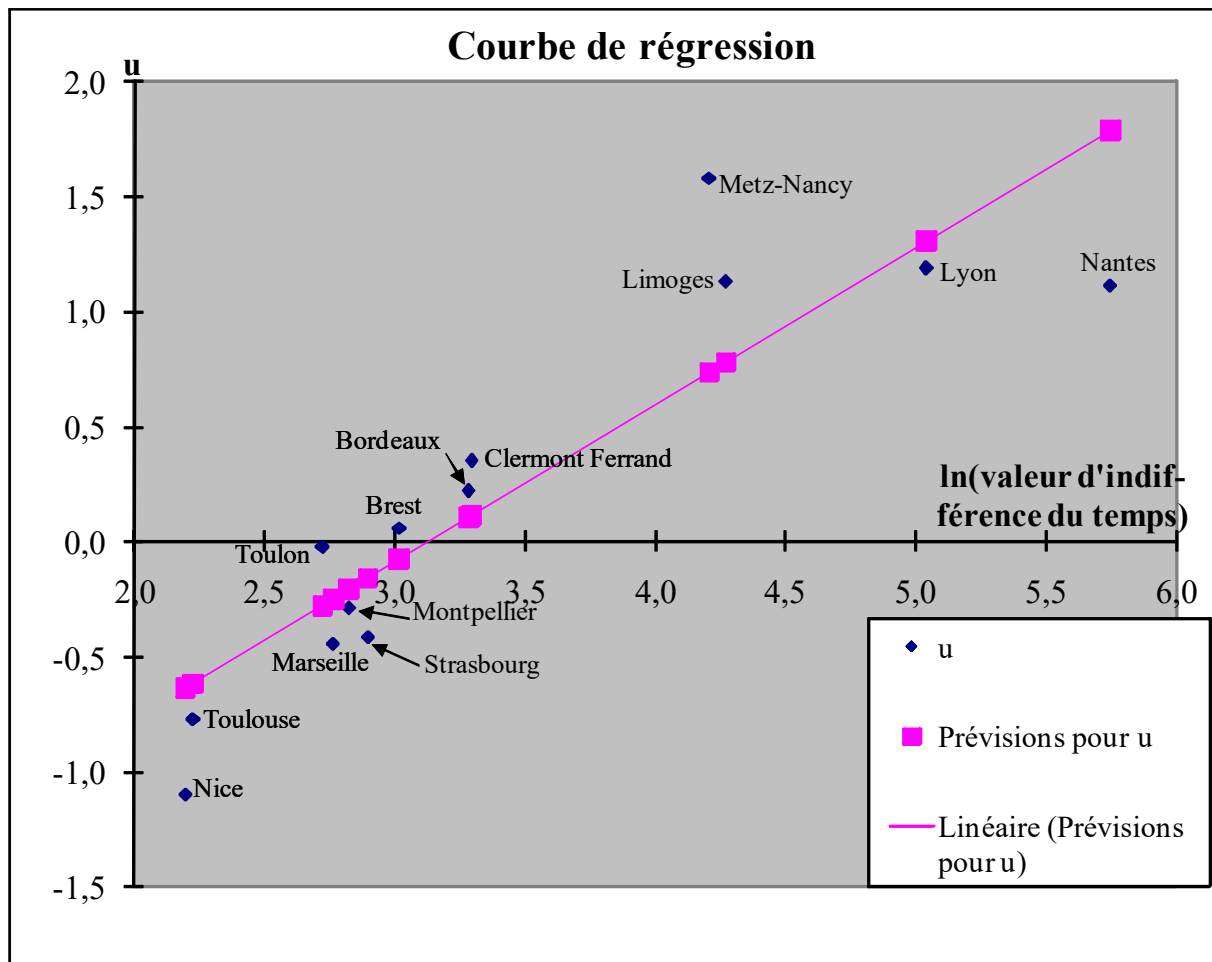
La part de marché du fer est fournie par $F(h)$ pour la valeur $h = h_i$. La transformée de la part de marché du fer est donc :

$$\Phi^{-1}(F(h_i)) = \frac{(\ln(h_i) - m)}{\sigma} \quad (12)$$

Pour calculer m et σ , il faut déterminer les paramètres de la droite (équation 12). Pour cela, nous disposons, sur un certain nombre de liaisons où l'avion et le train sont en concurrence, des données de trafic. Elles nous permettent de calculer la part de marché du fer $F(h_i)$, puis la valeur de $\Phi^{-1}(F(h_i))$:

- utiliser sous tableur la fonction statistique « loi.normale.standard.inverse » qui fournit $\Phi^{-1}(F(h_i))$ en fonction de $F(h_i)$ pour le train.

Les valeurs de h_i (valeur d'indifférence du temps) sont obtenues à partir des données de temps et de tarif (équation 4). Il suffit donc de déterminer la droite (appelée droite de Henry) passant au plus près de ces points par régression linéaire (graphique 4).



Graphique 4 : Calage de la droite de Henry avec $u = \Phi^{-1}(F(h))$

3. DONNEES NECESSAIRES AU CALAGE DU MODELE "PRIX-TEMPS"

3.1. Trafics en 1996

entre Paris et en milliers de passagers	SNCF 1996			AERIEN 1996		
	1ère classe	2ème classe	total	Air France entre Orly et	Air France entre CDG et	Total y compris autres compagnies
Bordeaux	413	1592	2005	860	274	1400
Brest	67	404	471	376	51	428
Clermont-Ferrand	76	353	429	227	17	243
Limoges	116	359	475	63	0	70
Lyon	2024	3345	5369	410	300	710
Marseille	253	1069	1322	1406	375	2700
Metz	156	384	540			
Metz/Nancy				0	0	75
Montpellier	144	584	728	707	165	1150
Nancy	252	454	706			
Nantes	583	1712	2295	168	125	350
Nice	105	351	456	1273	601	2900
Strasbourg	99	542	641	795	198	1250
Toulon	139	490	629	461	0	650
Toulouse	115	634	749	1272	358	2650

Les trafics sont exprimés en milliers de voyageurs aller + retour. CDG = Charles de Gaulle. Pour l'aérien le trafic total comprend les deux aéroports et l'ensemble des compagnies assurant la liaison.

3.2. Temps de parcours en minutes en 1996

Entre Paris et	S N C F		AIR FRANCE (ORLY et ...)	
	trajet	temps terminal	trajet	temps terminal
Bordeaux	177	65	60	105
Brest	243	65	65	105
Clermont-Ferrand	201	65	50	90
Limoges	170	65	70	110
Lyon	120	65	55	115
Marseille	254	65	70	120
Metz/Nancy			50	90
Metz	163	65		
Montpellier	254	65	70	115
Nancy	157	65		
Nantes	119	65	50	125
Nice	386	65	80	130
sans TGV	231	65		
Strasbourg			55	105
avec TGV projet 1C	110	65		
avec TGV projet 3CP	136	65		
Toulon	303	65	75	110
Toulouse	300	65	70	120

Le temps terminal = temps d'accès à l'aéroport (ou à la gare + temps d'accès à la destination finale à partir de l'aéroport ou de la gare).

3.3. Tarifs de base Automne-Hiver 1996 (converti en équivalent euros 96)

entre Paris et	Tarif moyen (moyenne sur les résa)		Air France tarif moyen entre Orly et
	1ère	2ème	
Bordeaux	62,20	44,71	82,61
Brest	72,57	48,07	98,60
Clermont-Ferrand	44,96	30,06	89,44
Limoges	43,15	28,77	98,02
Lyon	64,27	43,41	90,02
Marseille	72,90	50,77	89,10
Metz/Nancy	39,39	26,31	125,83
Montpellier	72,76	50,67	92,69
Nantes	57,53	38,36	90,37
Nice	84,68	60,18	101,96
Strasbourg sans TGV	51,57	34,34	78,09
Toulon	77,66	54,15	105,78
Toulouse	80,21	56,76	87,36

Pour le calcul du tarif moyen SNCF 1ère + 2ème classe, vous prendrez la moyenne des deux classes au prorata des trafics fournis dans le tableau 1.

Remarques. Pour le trafic aérien, compte de la structure tarifaire, le tarif moyen effectivement payé par l'utilisateur est obtenu en appliquant un coefficient de 76 % du plein tarif. Pour la SNCF, ce pourcentage est de 85 % du plein tarif (obtenu en faisant la moyenne du plein tarif pour chaque niveau de réservation pondéré par le poids du trafic de chaque niveau de réservation). C'est ce tarif qui figure dans le tableau ci-dessus.