

INTRODUCTION AUX PLANS D'EXPÉRIENCES

Tout le catalogue sur
www.dunod.com



Jacques Goupy

INTRODUCTION AUX PLANS D'EXPÉRIENCES

Avec applications

5^e édition

**L'USINE
NOUVELLE**

DUNOD

Pratiquer les plans d'expériences, Dunod, 2005

Pratiquer les plans d'expériences : les mélanges, Dunod, 2001

Le pictogramme qui figure ci-contre mérite une explication. Son objet est d'alerter le lecteur sur la menace que représente pour l'avenir de l'écrit, particulièrement dans le domaine de l'édition technique et universitaire, le développement massif du photocopillage.

Le Code de la propriété intellectuelle du 1^{er} juillet 1992 interdit en effet expressément la photocopie à usage collectif sans autorisation des ayants droit. Or, cette pratique s'est généralisée dans les établissements

d'enseignement supérieur, provoquant une baisse brutale des achats de livres et de revues, au point que la possibilité même pour

les auteurs de créer des œuvres nouvelles et de les faire éditer correctement est aujourd'hui menacée.

Nous rappelons donc que toute reproduction, partielle ou totale, de la présente publication est interdite sans autorisation de l'auteur, de son éditeur ou du Centre français d'exploitation du

droit de copie (CFC, 20, rue des Grands-Augustins, 75006 Paris).



© Dunod, Paris, 2001, 2006, 2009, 2013

© Bordas, Paris, 1988 pour la 1^{re} édition

ISBN 978-2-10-059296-8

Le Code de la propriété intellectuelle n'autorisant, aux termes de l'article L. 122-5, 2° et 3° a), d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective » et, d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (art. L. 122-4).

Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles L. 335-2 et suivants du Code de la propriété intellectuelle.

TABLE DES MATIÈRES

Table des matières	V
Avant-propos	IX
1 • Faites connaissance avec la méthode des plans d'expériences	1
1.1 Processus d'acquisition des connaissances	1
1.2 Étude d'un phénomène	3
1.3 Terminologie	4
1.4 Variables centrées réduites	8
1.5 Points expérimentaux	9
1.6 Plans d'expériences	10
2 • Votre premier plan d'expériences	15
2.1 Exemple 01 : je maîtrise la consommation de ma voiture	15
3 • Plan d'expériences pour 3 facteurs	31
3.1 Exemple 02 : des bijoux en or	31
3.2 Conclusion de l'étude	47
4 • Plans factoriels complets à 4 facteurs	49
4.1 Exemple 03 : la galette des rois	50
4.2 Fonction désirabilité	60
4.3 Application à l'exemple 03	64
4.4 Conclusion de l'étude	65
5 • Notions de statistique appliquées aux plans d'expériences	67
5.1 Exemple 04 : durée de vie des outils de coupe jetables	68
5.2 Notions de statistique appliquées aux plans d'expériences	71
5.3 Facteurs et interactions influents	81
5.4 Analyse de la variance	82
5.5 Application à l'exemple 04	85
5.6 Conclusion de l'étude	91

6 • Plans factoriels fractionnaires	93
6.1 Exemple 05 : le tellure	94
6.2 Théorie des alias	97
6.3 Calcul de Box	99
6.4 Relation d'équivalence	102
6.5 Générateurs d'alias	104
6.6 Construction pratique d'un plan fractionnaire	105
6.7 Nombre maximal de facteurs étudiés sur un plan de base	112
6.8 Application de la théorie des alias à l'exemple 05	112
6.9 Conclusion de l'étude	113
7 • Exemples de plans factoriels fractionnaires	115
7.1 Exemple 06 : la sulfonation	116
7.2 Exemple 07 : le spectrofluorimètre	124
7.3 Exemple 08 : les pommes de terre chips	140
8 • Ordre des essais	147
8.1 Nature des erreurs	148
8.2 Exemple 09 : le <i>Penicillium chrysogenum</i> (exemple de blocking)	151
8.3 Exemple 10 : les haricots de Yates	161
8.4 Exemple 11 : le broyeur (exemple de plan anti-dérive)	168
8.5 Avantages et dangers de la randomisation	175
9 • Plans pour surfaces de réponse	179
9.1 Présentation des plans composites	179
9.2 Présentation des plans de Box-Behnken	180
9.3 Présentation des plans de Doehlert	181
9.4 Exemple 12 : la rectification du contremaître (exemple de plan composite)	184
9.5 Exemple 13 : un yoghourt doux (exemple de plan de Box-Behnken)	194
9.6 Exemple 14 : l'insecticide (exemple de plan de Doehlert)	199
10 • Plans de mélanges	207
10.1 Contrainte fondamentale des mélanges	207
10.2 Représentation géométrique des mélanges	208
10.3 Plans de mélanges classiques	212
10.4 Modèles mathématiques des plans de mélanges	215
10.5 Exemple 15 : les trois polymères	217

11 • Notion de plan optimal	221
11.1 Exemple d'Hotelling	221
11.2 Pesées et plans d'expériences	223
11.3 Optimalité	225
11.4 Plans optimaux avec un modèle linéaire	229
11.5 Quand utiliser des plans optimaux ?	232
11.6 Adaptabilité des plans optimaux	232
11.7 Exemple 16 : développement d'un détecteur de fissures	235
11.8 Exemple 17 : les comprimés du pharmacien	240
11.9 Exemple 18 : la crème à bronzer	248
12 • Carrés latins et plans associés	255
12.1 Exemple 19 : les salaires – un facteur discret à plusieurs niveaux	255
12.2 Suite de l'exemple 19 – deux facteurs discrets à plusieurs niveaux	257
12.3 Carrés latins	264
12.4 Carrés gréco-latins	265
12.5 Carrés de Youden	266
12.6 Exemple 20 : étude de la pénétrométrie	267
12.7 Interaction entre facteurs discrets	270
12.8 Analyse de la variance d'un modèle avec interactions	274
13 • Les plans à facteurs mixtes	279
13.1 Plans à un facteur	280
13.2 Plans à deux facteurs	283
13.3 Plans à trois facteurs	284
13.4 Plans à quatre facteurs	285
13.5 Plans à n facteurs	286
13.6 Exemple 21 : les deux catalyseurs	287
13.7 Qu'aurait donné un plan fractionnaire ?	296
13.8 Conclusion sur les plans à facteurs mixtes	302
14 • Les plans en parcelles divisées ou plans split-plot	303
14.1 Introduction	303
14.2 Exemple 22 : le maïs	304
14.3 Exemple 23 : les revêtements anti-corrosions	313
14.4 Exemple 24 : céréale	319
14.5 Conclusion sur les plans split-plot	327

15 • Les plans uniformes	329
15.1 Introduction	329
15.2 Les différents plans uniformes	330
15.3 La modélisation mathématique de la réponse	340
15.4 Application	343
15.5 Plans uniformes pour plusieurs facteurs	350
15.6 Utilisation des plans uniformes en expérimentation classique	351
16 • Synthèse et conseils	353
16.1 Choix d'une méthode d'expérimentation	353
16.2 Expérimentation	359
16.3 Analyse des résultats	359
16.4 Acquisition progressive des connaissances	361
16.5 Recommandations	362
16.6 Limites de l'expérimentique	362
A • Origine des exemples	365
B • Comparaison de deux moyennes indépendantes	369
C • Séries d'essais laissant les facteurs principaux insensibles à une dérive linéaire	371
D • Introduction au calcul matriciel	377
E • Télécharger un logiciel de plans d'expériences	383
Bibliographie	385
Index	397
Liste des exemples	401

AVANT-PROPOS

Pour la cinquième édition de ce livre nous avons complété le chapitre 12 sur les facteurs discrets et nous avons ajouté un chapitre sur les plans en parcelles divisés. Le chapitre 12 comprend maintenant un paragraphe sur les interactions entre facteurs discrets et un paragraphe sur l'analyse de la variance de ces facteurs. Le nouveau chapitre, le chapitre 14, concerne les plans split-plot (ou plans en parcelles divisés) qui ont toujours été très utilisés en agriculture. Depuis une vingtaine d'années les expérimentateurs industriels ont pris conscience que beaucoup d'analyses de leurs plans ne tenaient pas compte des restrictions à la randomisation souvent présente dans leur expérimentation. Or dès que le niveau d'un facteur est difficile à changer, les essais sont regroupés pour éviter ces changements. La randomisation est alors partielle et le plan devient un plan split-plot dont l'analyse doit être traitée différemment des plans classiques.

La première et la deuxième édition de ce livre ne traitaient que des plans factoriels, complets ou fractionnaires. La troisième édition comportait, en plus, des chapitres sur les plans pour surfaces de réponse, sur les plans de mélanges, sur les plans D-optimaux et sur les plans pour variables discrètes. La quatrième édition introduisait un chapitre sur les plans contenant des facteurs continus et des facteurs discrets qui soulèvent des difficultés particulières. Cette quatrième édition introduisait également un chapitre sur les plans uniformes qui sont importants dans le domaine des simulations numériques et sont maintenant utilisés en expérimentations classiques.

Ce livre est complété par un blog (<http://www.plansdexperiences.com>) sur lequel vous trouverez des explications supplémentaires sur les exemples, des manuels d'introduction aux logiciels et les fichiers téléchargeables des exemples.

Nous tenons à remercier Bruno Bournique, Michel Genetay, Emmanuel Grenier, Rémi Longera, Alain Morineau, Lalatiana Rakotozafy qui ont relu les manuscrits des deux nouveaux chapitres de cette quatrième édition et qui nous ont fait bénéficier de leurs suggestions et de leurs conseils.

Nous conseillons aux lecteurs qui veulent approfondir leurs connaissances, d'accompagner leur lecture par l'exécution des calculs avec un logiciel. Par exemple MINITAB®, Statgraphics®, JMP® ou Modde® qui peuvent être téléchargé en version complète généralement valable un mois gratuitement.

Ils pourront ainsi approfondir les analyses et procéder à leur propre interprétation. Nous avons essayé de fournir aux lecteurs tous les outils et toutes les explications pour qu'ils puissent utiliser facilement la méthode des plans d'expériences pour leurs propres travaux. Il nous reste à leur souhaiter : « Bons plans d'expériences ! ».

Jacques Goupy, Paris (France), Mars 2013

1 • FAITES CONNAISSANCE AVEC LA MÉTHODE DES PLANS D'EXPÉRIENCES

Si vous lisez ce livre c'est que, probablement, vous faites des expériences et que vous cherchez à mieux les organiser. Vous cherchez surtout à ne faire que les bonnes expériences, c'est-à-dire celles qui vont vous apporter la meilleure information expérimentale que vous puissiez espérer.

Ce livre est écrit pour vous aider. Il va vous décrire la méthode et les outils pour que vous ne fassiez plus que les bonnes expériences et que vous en tiriez le maximum de renseignements. Vous prendrez vos décisions en connaissance de cause et vous trouverez des solutions à vos problèmes.

Nous allons d'abord regarder comment s'intègrent les plans d'expériences dans le processus d'acquisition des connaissances. Puis, nous verrons les concepts de base qui permettent de bien définir une étude et comment l'on interprète les résultats des essais.

1.1 Processus d'acquisition des connaissances

Augmenter ses connaissances, c'est trouver la réponse à une question posée. On commence donc par se poser une ou plusieurs questions (Figure 1.1). Par exemple, si l'on désire connaître l'influence d'un engrais sur le rendement en blé d'une parcelle de terrain, on pourra se poser les questions suivantes qui, bien sûr, ne sont pas limitatives :

- « Peut-on obtenir 10 quintaux de blé de plus à l'hectare en augmentant la quantité d'engrais ? »
- « Quelle est l'incidence de la pluie sur l'efficacité de l'engrais ? »
- « La qualité du blé restera-t-elle bonne si j'utilise cet engrais ? »
- « Combien dois-je utiliser d'engrais pour obtenir la quantité maximale de blé par hectare ? »

Ces questions délimitent le problème à résoudre et fixent les travaux à exécuter. Il est donc important de se poser les questions qui correspondent réellement au problème. Bien sûr, avant d'entreprendre des expériences, il est prudent de vérifier que l'information cherchée n'existe pas déjà ailleurs. Dans ce but, on effectue un inventaire des connaissances, soit une bibliographie, soit une consultation d'experts, soit un calcul théorique, soit toute autre démarche ayant pour but de répondre,

sans expérience, aux questions posées. Après cette phase d'enquête, les questions initiales peuvent être entièrement résolues et il n'y a plus de problème. Dans le cas contraire, certaines questions restent entières ou elles sont modifiées : il est alors nécessaire d'entreprendre des expériences pour obtenir les réponses complètes que l'on souhaite. Ce travail préalable fait partie du métier d'expérimentateur et nous n'insisterons pas. Ce n'est pas sur cette phase du processus que nous voulons faire porter notre effort de réflexion et d'amélioration.

N'ayant plus que les questions sans réponse immédiate, il faut faire des expériences. Comment peut-on choisir ces expériences pour :

- arriver rapidement aux meilleurs résultats possibles,
- éviter de réaliser des expériences inutiles,
- obtenir la meilleure précision possible sur les résultats,
- permettre d'avancer à coup sûr,
- établir la modélisation du phénomène étudié,
- découvrir la solution optimale.

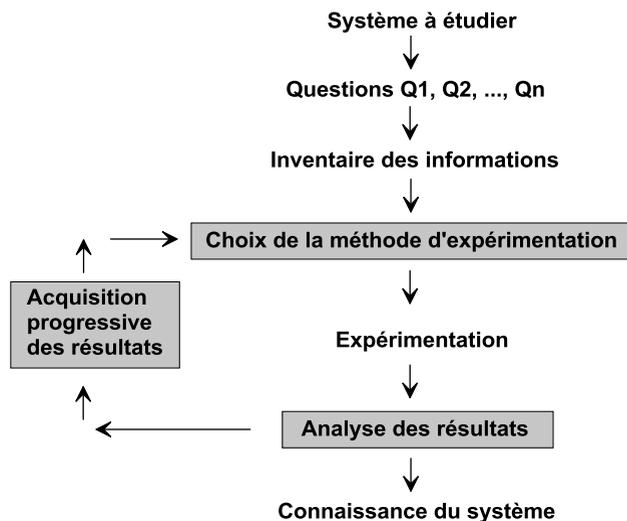


Figure 1.1 – Les plans d'expériences optimisent les trois parties encadrées du processus d'acquisition des connaissances.

Les trois aspects essentiels du processus d'acquisition des connaissances sont les suivantes :

- le choix de la méthode d'expérimentation,
- l'analyse des résultats,
- l'acquisition progressive de la connaissance.

Examinons plus en détail ces trois aspects sachant que les expériences sont organisées pour faciliter l'exploitation des résultats et pour permettre l'acquisition progressive des résultats d'intérêt.

1.1.1 Choix de la méthode d'expérimentation

La méthode d'expérimentation choisie doit faciliter l'interprétation des résultats. Elle doit également minimiser le nombre des essais sans toutefois sacrifier la qualité. La théorie des plans d'expériences assure les conditions pour lesquelles on obtient la meilleure précision possible avec le minimum d'essais. On a donc le maximum d'efficacité avec le minimum d'expériences et par conséquent le coût minimum.

1.1.2 Analyse des résultats

L'analyse des résultats d'expériences est facilitée par le choix initial des expériences. Les résultats seront faciles à interpréter et riches d'enseignement si l'on a bien préparé les expériences.

Grâce aux ordinateurs et aux logiciels la construction des plans d'expériences et les calculs nécessaires à leur interprétation sont devenus très simples. Ces outils favorisent également les représentations graphiques qui illustrent de manière spectaculaire les résultats et améliorent la compréhension des phénomènes.

1.1.3 Acquisition progressive de la connaissance

L'expérimentateur qui entreprend une étude ne connaît pas les résultats, il est donc sage d'avancer progressivement pour pouvoir réorienter les essais en fonction des premiers résultats. Une première ébauche permettra de mieux orienter les essais vers les seuls aspects intéressants de l'étude et d'abandonner les voies sans issues.

Une première série d'expériences conduit à des conclusions provisoires ; en fonction de ces conclusions provisoires, on lance une nouvelle série d'essais. L'ensemble des deux séries d'expériences est utilisé pour obtenir un contour précis des résultats d'étude. L'expérimentateur accumule ainsi les seuls résultats dont il a besoin et s'arrête dès qu'il a obtenu ce qu'il cherche.

1.2 Étude d'un phénomène

L'étude d'un phénomène revient souvent à s'intéresser à une grandeur particulière comme la consommation d'essence d'une voiture ou comme le prix de revient d'un produit chimique ou encore comme le rendement en blé d'une parcelle de terrain. Cette grandeur, consommation, prix ou rendement, dépend d'un grand nombre de variables. La consommation de la voiture est fonction de la vitesse du véhicule, de la puissance du moteur, de la manière de conduire, de la direction et de la force du vent, du gonflage des pneumatiques, de la présence ou non d'un porte-bagages, du nombre de personnes transportées, de la marque de la voiture, etc. Le prix du produit chimique dépend de la qualité des matières premières, des rendements des unités de production, des spécifications imposées, des conditions de fabrication, etc. Il en est de même pour le rendement en blé qui est fonction de la nature du terrain, de la quantité d'engrais incorporé, de l'exposition au soleil, du climat, de la variété de blé ensemencé, etc.

Sous une forme mathématique, on peut écrire que la grandeur d'intérêt, y , que nous appellerons également *réponse* par la suite, est une fonction de plusieurs variables x_i (variables que nous appellerons aussi *facteurs* par la suite). On a :

$$y = f(x_1, x_2, x_3, \dots, x_k)$$

L'étude du phénomène se ramène à déterminer la fonction $f()$ qui lie la réponse y aux différents facteurs x_1, x_2, \dots, x_k .

Pour approfondir cette approche il faut introduire quelques notions particulières et une terminologie spécifique aux plans d'expériences.

1.3 Terminologie

La grandeur d'intérêt, qui est généralement notée y , porte le nom de *réponse*. Les variables qui peuvent modifier la réponse sont appelées *facteurs*. On parle donc des facteurs qui influent sur une réponse. Les termes *facteur* et *réponse* sont universellement employés dans le domaine des plans d'expériences.

1.3.1 Les différents types de facteurs

La construction des plans et l'interprétation des résultats dépendent en grande partie des types de facteurs rencontrés dans l'étude. On distingue plusieurs types de facteurs. Nous retiendrons les types de facteurs suivants : les facteurs continus, les facteurs discrets, les facteurs ordonnables, les facteurs booléens.

■ Facteurs continus

La pression est un exemple de facteur continu. Dans un intervalle de pression donné, on peut choisir toutes les valeurs possibles. Il en est de même d'une longueur, d'une concentration ou d'une température. Les valeurs prises par les facteurs continus sont donc représentées par des nombres continus.

■ Facteurs discrets

Au contraire, les facteurs discrets ne peuvent prendre que des valeurs particulières. Ces valeurs ne sont pas forcément numériques : on peut représenter un facteur discret par un nom, une lettre, une propriété ou même par un nombre qui n'a alors en soi aucune valeur numérique mais qu'une signification de repère. Par exemple, on peut s'intéresser aux couleurs d'un produit : bleu, rouge et jaune sont des facteurs discrets.

■ Facteurs ordonnables

Il s'agit de facteurs discrets que l'on peut mettre dans un ordre logique. Par exemple, grand, moyen, petit, ou encore premier, deuxième, troisième et quatrième.

■ Facteurs booléens

Les facteurs booléens sont des facteurs discrets qui ne peuvent prendre que deux valeurs : haut ou bas, ouvert ou fermé, blanc ou noir, etc.

La frontière est parfois floue entre ces différents types de facteurs. Une couleur, facteur apparemment discret, peut être transformée en grandeur ordonnable et même continue si l'on introduit la notion de longueur d'onde. Un facteur continu, comme la vitesse, peut être transformé en facteur ordonnable ou discret : rapide et lent, ou vitesse A et vitesse B. Cette possibilité n'est pas un inconvénient, c'est une souplesse supplémentaire dont l'expérimentateur pourra tirer avantage au moment de l'interprétation. En effet, suivant le choix effectué, on pourra mettre plus facilement en relief un aspect particulier de l'étude. Le changement de types de variables est aussi un moyen d'adapter la réponse à l'objectif de l'étude. Par exemple, on possède l'âge des membres d'une population. Si l'on cherche l'âge moyen, la variable « âge » sera choisie continue. Si l'on cherche le pourcentage des personnes ayant un âge donné, la variable « âge » sera choisie ordonnable : il y aura plusieurs catégories définies par les limites d'âge, les jeunes, les adolescents, les adultes et les seniors. Si l'on cherche la proportion de jeunes de moins de 18 ans, la variable « âge » sera choisie booléenne : les moins de 18 ans et les plus de 18 ans. La compréhension des plans d'expériences s'appuie sur la notion essentielle d'espace expérimental des grandeurs étudiées. Nous allons examiner en détail ce concept fondamental. Nous insisterons sur les facteurs continus qui sont les plus couramment utilisés.

1.3.2 Espace expérimental

Pour présenter l'espace expérimental nous utiliserons un espace à deux dimensions, ce qui facilitera les représentations graphiques. Il est ensuite facile d'étendre les notions introduites à des espaces multidimensionnels.

Un facteur continu peut être représenté par un axe gradué et orienté. S'il y a un second facteur continu, il est représenté, lui aussi, par un axe gradué et orienté. Ce second axe est disposé orthogonalement au premier. On obtient ainsi un repère cartésien qui définit un espace euclidien à deux dimensions. Cet espace est appelé *l'espace expérimental* (figure 1.2). L'espace expérimental comprend tous les points du plan « facteur 1 \times facteur 2 » et chacun d'eux représente une expérience.

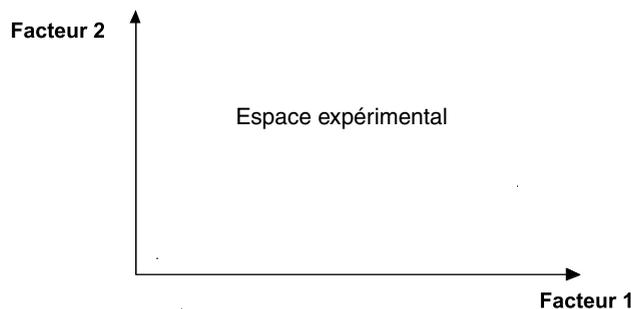


Figure 1.2 – Chaque facteur est représenté par un axe gradué et orienté. Les axes des facteurs sont orthogonaux entre eux. L'espace ainsi défini est l'espace expérimental.

S'il y a un troisième facteur, on le représente aussi par un axe orienté et gradué, et on le positionne perpendiculairement aux deux premiers. À partir de quatre facteurs, on opère de même, mais il n'y a plus de représentation géométrique possible et l'on doit adopter une représentation purement mathématique de l'espace expérimental qui est un hypervolume à quatre dimensions.

1.3.3 Domaine d'un facteur

La valeur donnée à un facteur pour réaliser une expérience est appelée *niveau*.

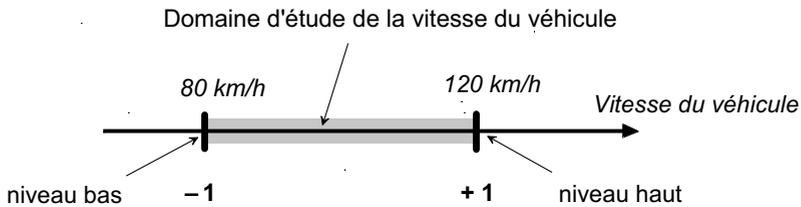


Figure 1.3 – Le domaine de variation de la vitesse est constitué de toutes les vitesses comprises entre 80 et 120 km/h. Le niveau bas du facteur est noté -1 et le niveau haut $+1$.

Lorsqu'on étudie l'influence d'un facteur, en général, on limite ses variations entre deux bornes. La borne inférieure est le *niveau bas*. La borne supérieure est le *niveau haut*. Si l'on étudie l'influence de la vitesse du véhicule sur la consommation, celle-ci peut varier, par exemple, entre 80 et 120 km/h. La vitesse de 80 km/h est le niveau bas et la vitesse de 120 km/h est le niveau haut. C'est l'expérimentateur qui définit ces deux niveaux en fonction des spécificités de l'étude. L'ensemble de toutes les valeurs que peut prendre le facteur entre le niveau bas et le niveau haut, s'appelle le *domaine de variation du facteur* ou plus simplement le *domaine du facteur*.

S'il y a plusieurs facteurs, chacun d'eux à son domaine de variation. Afin d'avoir une représentation commune pour tous les facteurs, on a l'habitude d'indiquer les niveaux bas par -1 et les niveaux hauts par $+1$. La vitesse de 80 km/h est le niveau -1 et celle de 120 km/h est le niveau $+1$.

À l'intérieur du domaine d'un facteur continu toutes les valeurs sont théoriquement possibles. On peut donc y choisir deux, trois ou plusieurs niveaux selon les besoins de l'étude. Par exemple, si l'on veut établir un modèle du second degré, on choisira trois ou quatre niveaux, c'est-à-dire trois ou quatre vitesses différentes.

1.3.4 Domaine d'étude

Dans la pratique, l'expérimentateur sélectionne une partie de l'espace expérimental pour réaliser son étude. Cette zone particulière de l'espace expérimental est le *domaine d'étude* (figure 1.4). Ce domaine est défini par les niveaux hauts et les niveaux bas de tous les facteurs et éventuellement par des contraintes entre les facteurs. Supposons que le second facteur soit la surcharge du véhicule définie comme toute masse supplémentaire à celle du véhicule et du chauffeur. Le niveau

bas de la surcharge est 0 kg et le niveau haut 300 kg, par exemple. S'il n'y a pas de contraintes, le domaine d'étude est représenté par tous les points dont les surcharges sont comprises entre 0 et 300 kg et dont les vitesses sont comprises entre 80 et 120 km/h.

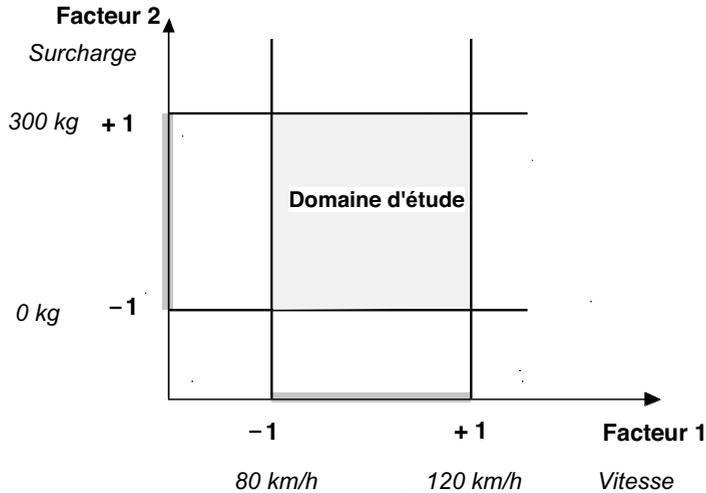


Figure 1.4 – Le domaine d'étude est défini par la réunion des domaines des différents facteurs (ici, il n'y a pas de contraintes).

Il peut y avoir des contraintes sur le domaine d'étude. Par exemple, il peut être impossible d'atteindre la vitesse de 120 km/h avec une surcharge trop élevée. La figure 1.5 illustre une réduction possible du domaine d'étude initial. Une zone du domaine d'étude initial échappe aux expériences.

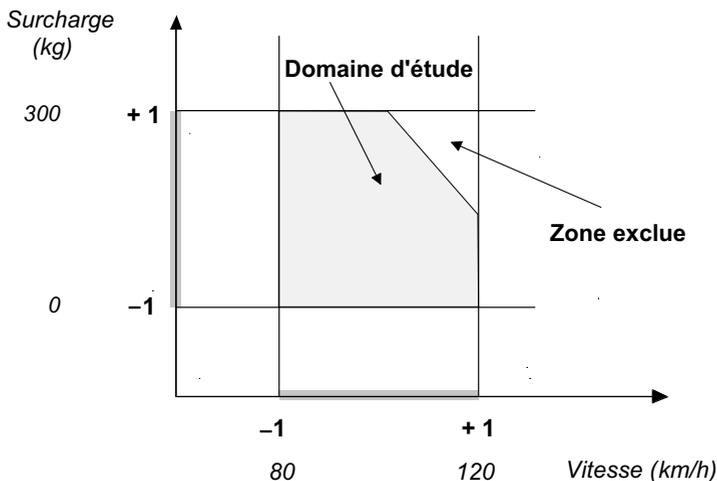


Figure 1.5 – Le domaine d'étude sous contraintes est représenté par la partie grisée.

1.4 Variables centrées réduites

Lorsqu'on attribue la valeur -1 au niveau bas d'un facteur et la valeur $+1$ au niveau haut, on effectue deux modifications importantes :

- On déplace l'origine des mesures. Dans l'exemple choisi, le milieu de l'intervalle $[-1 ; +1]$ correspond à une valeur de 100 km/h. La valeur numérique de la nouvelle origine, zéro, diffère donc de l'origine exprimée en unité courante.
- On change l'unité de mesure. Par exemple, si le niveau bas du facteur « vitesse du véhicule » est 80 km/h et le niveau haut 120 km/h, il y a 40 km/h entre ces deux valeurs, soit 40 fois l'unité de vitesse. Entre -1 et $+1$ il y a deux unités nouvelles : la nouvelle unité vaut 20 km/h, on lui donne le nom de *Pas*.

Ces deux modifications entraînent l'introduction de nouvelles variables que l'on appelle *variables centrées réduites* (v.c.r) : centrées pour indiquer le changement d'origine et réduites pour signaler la nouvelle unité. On utilise également le terme de variables codées ou d'unités codées.

Le passage des variables d'origine A aux variables codées x , et inversement, est donné par la formule suivante (A_0 est la valeur centrale en unités courantes) :

$$x = \frac{A - A_0}{Pas} \quad \{1.1\}$$

L'intérêt des unités codées est de pouvoir présenter les plans d'expériences de la même manière quels que soient les domaines d'étude retenus et quels que soient les facteurs. La théorie des plans d'expériences présente ainsi une grande généralité.

L'utilisation des v.c.r est très répandue dans les logiciels de plans d'expériences et certaines opérations comme la recherche des meilleurs points d'expériences par le critère de D-optimalité ne sont réalisables qu'avec ces variables.

Les variables codées résultent du rapport de deux grandeurs de même unité physique, elles sont donc sans dimension. La disparition des unités naturelles associée au fait que tous les facteurs ont le même domaine de variation (deux unités codées) permet la comparaison directe des effets des facteurs entre eux.

Application 1

Un expérimentateur choisit pour le facteur « vitesse de la voiture » 80 km/h comme niveau bas et 120 km/h comme niveau haut. Quelle est, en variables codées, la valeur de la vitesse de 90 km/h ?

Calculons le *Pas* du facteur vitesse. Il est égal à la moitié de la différence entre le niveau haut et le niveau bas :

$$Pas = \frac{A_{+1} - A_{-1}}{2} = \frac{120 - 80}{2} = 20$$

A_0 est la valeur centrale entre le niveau haut et le niveau bas, c'est-à-dire la moitié de la somme du niveau haut et du niveau bas :

$$A_0 = \frac{A_{+1} + A_{-1}}{2} = \frac{120 + 80}{2} = 100$$

Appliquons la relation {1.1} :

$$x = \frac{A - A_0}{Pas} = \frac{90 - 100}{20} = -0,5$$

Pour cet exemple, la vitesse de 90 km/h est donc égale à $-0,5$ en variables codées.

Application 2

On peut aussi chercher la valeur normale connaissant la valeur codée. Quelle est, en unités normales, la valeur de la vitesse de +0,5 en unités codées ? Écrivons la relation $\{1.1\}$:

$$+0,5 = \frac{A - 100}{20}$$

D'où : $A = +100 + 0,5 \times 20 = 110$

La vitesse codée de 0,5 correspond à une vitesse de 110 km/h.

1.5 Points expérimentaux

Dans un espace à deux dimensions, le niveau i du facteur 1, noté $x_{1,i}$, et le niveau j du facteur 2, noté $x_{2,j}$, peuvent être considérés comme les coordonnées d'un point de l'espace expérimental ou du domaine d'étude (figure 1.6). Par exemple, si le niveau de la vitesse est 90 km/h et celui de la surcharge est 250 kg, les coordonnées du point expérimental sont :

$$\begin{aligned}x_{1,i} &= 90 \text{ km/h} \\x_{2,j} &= 250 \text{ kg}\end{aligned}$$

Une expérience donnée est alors représentée par un point dans ce système d'axes. C'est la raison pour laquelle une expérience est souvent désignée par l'expression *point expérimental*, *point d'expérience* ou même simplement *point*. Un plan d'expériences est donc représenté par un ensemble de points expérimentaux, eux-mêmes situés dans l'espace expérimental. Dans l'exemple que nous avons pris, l'expérience est conduite sur un véhicule qui roule à 90 km/h avec une surcharge de 250 kg.

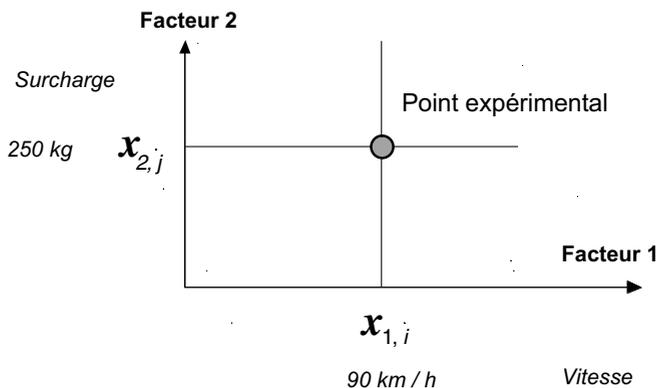


Figure 1.6 – Dans l'espace expérimental, les niveaux des facteurs définissent des points expérimentaux.

Jusqu'à trois facteurs, il est possible de dessiner le domaine d'étude. Au-delà de trois facteurs, on utilise une représentation en tableau, dite *matricielle*, plus générale puisqu'elle permet de représenter les points d'expériences dans un hypervolume à un nombre quelconque de dimensions.

1.6 Plans d'expériences

1.6.1 Méthodologie des plans sans contrainte

Le choix du nombre et de l'emplacement des points d'expériences est le problème fondamental des plans d'expériences. On cherche à réaliser le minimum d'expériences tout en réduisant l'influence de l'erreur expérimentale sur les modélisations mathématiques qui serviront à prendre des décisions. On atteint ce but en considérant les propriétés mathématiques et statistiques qui relient la réponse aux facteurs. Lorsqu'il n'y a pas de contraintes sur le domaine d'étude, il existe des plans classiques qui possèdent d'excellentes qualités statistiques et qui permettent de modéliser les réponses dans les meilleures conditions (figure 1.7). Lorsqu'il existe des contraintes, il faut construire des plans sur mesure en recherchant la position des points expérimentaux qui conduisent, là aussi, à de bonnes qualités statistiques et à une bonne modélisation des réponses.

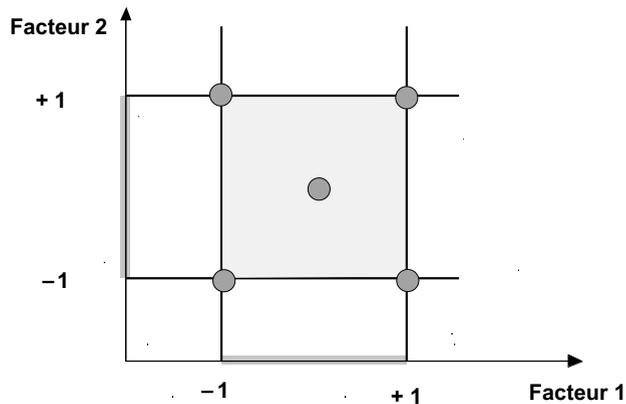


Figure 1.7 – Exemple de disposition des points expérimentaux dans un domaine sans contraintes.

1.6.2 Méthodologie des plans avec contrainte

La procédure de construction des plans dont le domaine est contraint est la suivante :

1. On définit le domaine de chacun des facteurs (niveau bas et niveau haut).
2. On définit les contraintes qui pèsent sur les facteurs. Ces contraintes sont exprimées par des relations d'inégalité entre les facteurs et elles définissent les zones autorisées, c'est-à-dire celles où les expériences sont possibles, et les zones interdites, c'est-à-dire celles où les expériences ne doivent pas être exécutées.
3. On définit les niveaux des facteurs les plus intéressants pour l'étude, autres que les niveaux bas et hauts. Le plus souvent entre 2 et 5 niveaux supplémentaires par facteur.
4. On construit une grille en prenant en compte toutes les combinaisons des niveaux des facteurs. Cette grille ne doit contenir que les points expérimentaux réalistes, c'est-à-dire les points du domaine autorisé. Ces points constituent les points candidats (figure 1.8).

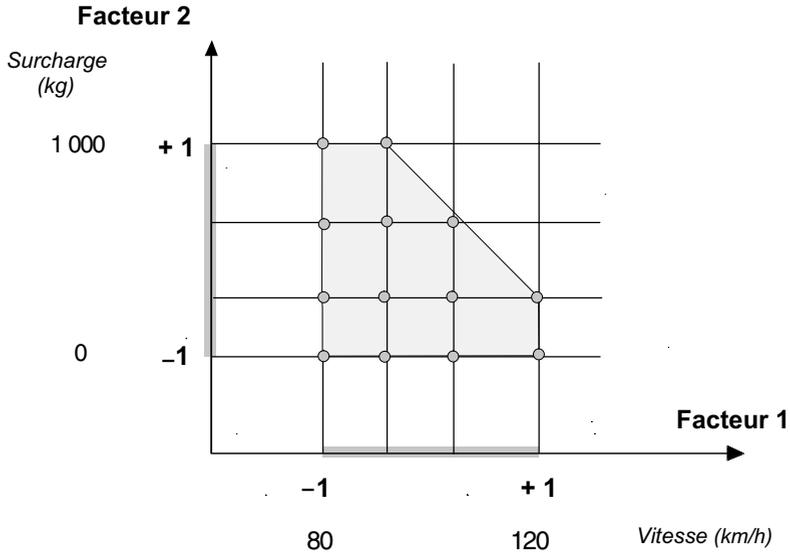


Figure 1.8 – La grille des points candidats est constituée d'expériences possibles dans le domaine d'étude.

5. On choisit une fonction reliant *a priori* la réponse aux facteurs.
6. On sélectionne, en fonction d'un critère d'optimalité choisi, le nombre et l'emplacement des points expérimentaux les plus utiles à la modélisation du phénomène étudié (figure 1.9). Cette sélection exige de longs calculs et n'est possible qu'avec l'aide d'un logiciel de plans d'expériences.

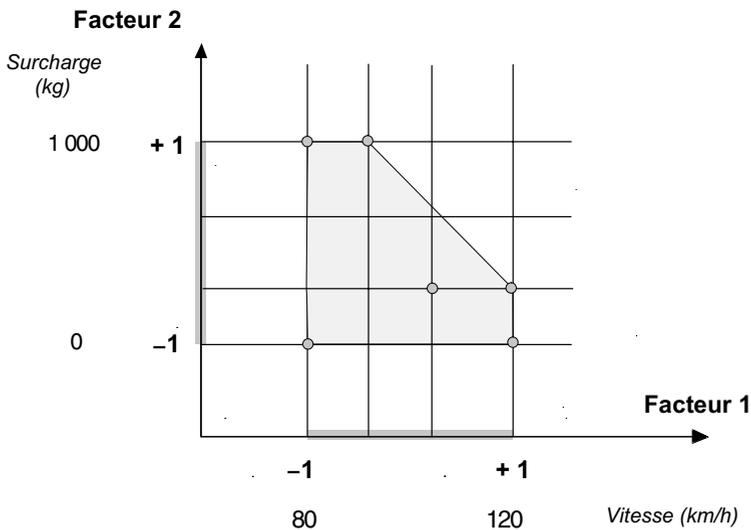


Figure 1.9 – Les meilleurs points sont sélectionnés par le logiciel.

1.6.3 Surfaces de réponse

À chaque point du domaine d'étude correspond une réponse. À l'ensemble de tous les points du domaine d'étude correspond un ensemble de réponses qui se localise sur une surface appelée *surface de réponse* (figure 1.10).

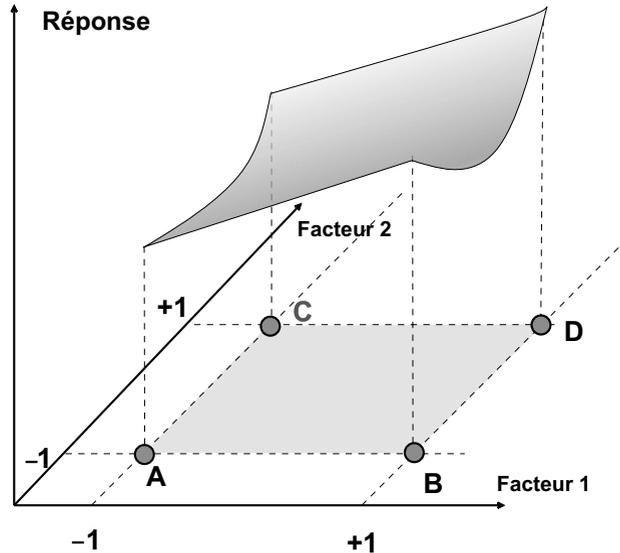


Figure 1.10 – L'ensemble des réponses qui correspond à tous les points du domaine d'étude forme la surface de réponse.

En général, on ne connaît que quelques réponses, celles qui correspondent aux points expérimentaux retenus par l'expérimentateur. On interpole à l'aide d'un modèle mathématique, les réponses inconnues pour obtenir la surface de réponse. Les points d'expériences retenus par la théorie des plans d'expériences assurent la meilleure précision possible sur la forme et la position de la surface de réponse.

1.6.4 Modélisation mathématique *a priori* de la réponse

■ Modélisation mathématique

En l'absence de toute information sur la fonction qui lie la réponse aux facteurs, on se donne *a priori* une loi d'évolution dont la formulation la plus générale est la suivante :

$$y = f(x_1, x_2, x_3, \dots, x_n) \quad \{1.2\}$$

Cette fonction est trop générale et il est d'usage d'en prendre un développement limité de Taylor-Mac Laurin, c'est-à-dire une approximation. Si les dérivées peuvent être considérées comme des constantes, le développement précédent prend la forme d'un polynôme de degré plus ou moins élevé :

$$y = a_0 + \sum a_i x_i + \sum a_{ij} x_i x_j + \sum a_{ii} x_i^2 + \dots \quad \{1.3\}$$

où :

- y est la grandeur à laquelle s'intéresse l'expérimentateur ; c'est la réponse ou la grandeur d'intérêt,
- x_i représente un niveau du facteur i ,
- x_j représente un niveau du facteur j ,
- a_0, a_i, a_{ij}, a_{ii} sont les *coefficients* du polynôme.

Ce modèle est appelé le modèle *a priori* ou le modèle *postulé*.

Les modèles établis sont des modèles de prévision valables dans le domaine d'étude, domaine que l'on doit toujours préciser. Ce ne sont pas des modèles théoriques basés sur des lois physico-chimiques ou mécaniques. Dans quelques rares cas, il est possible d'utiliser des lois physiques théoriques connues.

■ Modélisation expérimentale

Deux compléments doivent être apportés au modèle purement mathématique précédemment décrit.

Le premier complément est le manque d'ajustement. Cette expression traduit le fait que le modèle choisi par l'expérimentateur avant les expériences est probablement un peu différent du modèle réel qui régit le phénomène étudié. Il y a un écart entre ces deux modèles. Cet écart est le *manque d'ajustement* (*lack of fit* en anglais), on le note par la lettre Δ .

Le second complément est la prise en compte de la nature aléatoire de la réponse. En effet, dans le cas général, si l'on mesure plusieurs fois une réponse en un même point expérimental, on n'obtiendra pas exactement le même résultat. Il y a une dispersion des résultats. Les dispersions ainsi constatées sont appelées *erreurs aléatoires* ou *erreurs expérimentales* (*pure error* en anglais) et on les note par la lettre ε .

La relation générale {1.2} doit être modifiée ainsi :

$$y = f(x_1, x_2, x_3, \dots, x_n) + \Delta + \varepsilon \quad \{1.4\}$$

Cette relation sera exploitée au cours du chapitre 5 où l'on verra comment on estime le manque d'ajustement Δ et l'erreur aléatoire ε .

■ Système d'équations

Chaque point expérimental apporte une valeur de la réponse. Or cette réponse est modélisée par un polynôme dont les coefficients sont les inconnues qu'il faut déterminer. À la fin du plan d'expériences, on a un système de n équations (s'il y a n essais) à p inconnues (s'il y a p coefficients dans le modèle choisi *a priori*). Ce système s'écrit d'une manière simple en notation matricielle :

$$y = X a + e \quad \{1.5\}$$

où :

- y est le *vecteur des réponses*,
- X est la *matrice de calcul des coefficients* ou *matrice du modèle* qui dépend des points expérimentaux choisis pour exécuter le plan et du modèle postulé,
- a est le *vecteur des coefficients*,
- e est le *vecteur des écarts*.

Ce système ne peut pas, en général, être résolu simplement car le nombre d'équations est inférieur au nombre d'inconnues. En effet, il y a n équations et $p + n$ inconnues. Cette résolution ne peut être menée à bien que si l'on utilise une méthode de régression. La plupart du temps cette méthode est basée sur le critère d'optimisation des moindres carrés. On obtient ainsi les estimations des coefficients que l'on note

$$\hat{\mathbf{a}}$$

Le résultat de ce calcul est :

$$\hat{\mathbf{a}} = (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{y} \quad \{1.6\}$$

formule dans laquelle la matrice \mathbf{X}' est la matrice transposée de \mathbf{X} (voir l'annexe D sur le calcul matriciel). Il existe de nombreux logiciels qui exécutent ce calcul et qui donnent directement les valeurs des coefficients.

Deux matrices interviennent constamment dans la théorie des plans d'expériences :

- la matrice d'information $\mathbf{X}' \mathbf{X}$,
- la matrice de dispersion $(\mathbf{X}' \mathbf{X})^{-1}$.

2 • VOTRE PREMIER PLAN D'EXPÉRIENCES

Si vous avez lu le premier chapitre, vous en savez assez pour attaquer votre premier plan d'expériences. Nous allons suivre la démarche suivante dont les principales étapes sont :

- la préparation de l'étude,
- le choix du plan d'expériences,
- l'expérimentation,
- l'interprétation des résultats,
- l'arrêt ou la poursuite de l'étude.

Nous allons examiner ces différentes phases sur un premier exemple.

2.1 Exemple 01 : je maîtrise la consommation de ma voiture

2.1.1 Préparation du plan d'expériences

Cette étape se décompose en plusieurs parties dont les principales sont décrites dans les paragraphes suivants.

■ Définition de l'objectif de l'étude

Nous allons suivre la réalisation d'une étude que tout le monde peut mettre sur pieds. L'objectif de l'étude est de connaître la consommation d'essence d'une voiture lorsqu'on roule plus ou moins chargé et plus ou moins vite.

L'étude sera réussie si on peut répondre à des questions telles que : « Combien ma voiture consomme-t-elle : si elle est à vide ou non ? si je roule vite ou non ? si elle est chargée et si je roule vite ? »

■ Description des éléments sur lesquels va porter l'expérimentation

Je vais décrire les essais que j'ai effectués sur ma propre voiture. Je vais être le seul conducteur et j'effectuerai les essais sur une autoroute. Si j'en avais la possibilité, je ferais les essais sur un circuit automobile où il est facile de respecter la plupart des conditions imposées à l'expérimentation. Il est en effet plus facile de réaliser des essais sur un circuit, dans un laboratoire, un atelier pilote ou tout lieu spécialement conçu pour ce genre d'activité.

■ Choix de la réponse permettant d'atteindre l'objectif

La réponse sera la consommation d'essence en litres pour 100 kilomètres parcourus.

■ Recherche des facteurs qui pourraient être influents sur la réponse

Quels sont les facteurs qui peuvent modifier la consommation d'une voiture ?

Il y a bien sûr les deux facteurs que je vais étudier, la charge et la vitesse de la voiture. Mais il y en a d'autres. Par exemple, la marque et la pression de gonflage des pneumatiques, la présence ou non d'une galerie, la direction du vent, la pluie, le réglage du moteur, le nombre d'arrêts et de démarrages, la route plus ou moins accidentée. Il est prudent de noter par écrit tous les facteurs possibles. Pour remplir la liste des facteurs et essayer de ne pas en oublier, on peut aller voir son garagiste pour lui demander s'il ne connaît pas d'autres facteurs pouvant augmenter ou réduire la consommation. On a toujours intérêt à bien faire le tour du problème. On ne sera pas obligé d'étudier tous les facteurs mais on connaîtra ceux qui pourraient être influents. Les facteurs qui ne seront pas étudiés dans le plan d'expériences seront, en général, fixés à un niveau constant pendant toute l'expérimentation.

■ Définition des niveaux des facteurs

Il s'agit de choisir les niveaux haut et bas de chaque facteur.

La vitesse ne devra pas être trop faible et elle ne devra pas dépasser les limitations imposées par la sécurité. Dans notre exemple, le niveau bas sera de 80 km/h et le niveau haut de 120 km/h.

La charge à vide est définie par la présence du conducteur seul, sans aucun bagage. La surcharge est définie par le poids ajouté à la charge à vide. S'il y a 3 personnes supplémentaires et 25 kg de bagage par personne, la surcharge est d'environ $3 \times 70 = 210$ pour les passagers et $4 \times 25 = 100$ kg pour les bagages des quatre personnes à bord, soit environ 300 kg de surcharge. Au lieu de faire appel à ses voisins pour les essais, on pourra utiliser des gueuses de fonte régulièrement réparties sur les sièges et dans le coffre. Le poids total des gueuses est de 300 kg.

Le niveau bas du facteur charge sera : la voiture et le conducteur seul et sans bagage.

Le niveau haut du facteur charge sera : la voiture, le conducteur et une charge supplémentaire de 300 kg régulièrement répartie.

Il est bon de résumer ces niveaux dans un tableau (Tableau 2.1).

Tableau 2.1 – Facteurs et domaine d'étude.

Facteur	Niveau bas (-)	Niveau haut (+)
Vitesse (1)	80 km/h	120 km/h
Surcharge (2)	0	300 kg

Ce tableau est important car les conclusions de l'expérimentation ne seront valables qu'à l'intérieur de ce domaine d'étude. Si l'on charge la voiture à 400 kg et que l'on roule à 130 km/h, on ne pourra pas utiliser les conclusions de cette étude.

■ Examen des contraintes

Lorsque l'on roule, la quantité d'essence diminue. Le poids du véhicule est donc modifié au cours de l'expérimentation. Il faudra commencer chaque essai avec le réservoir plein pour que le facteur « diminution d'essence au cours de l'essai » n'ait pas trop d'influence. Cela oblige à débiter le parcours dans une station service.

Il faudra qu'il fasse le même temps pour tous les essais (température, vent, pluie). On choisira donc de faire les essais le même jour et, si possible, un jour de beau temps sans vent ni pluie. La longueur du parcours devra être compatible avec cette contrainte.

On choisira le même parcours et dans le même sens pour chaque essai pour que l'influence des côtes et des descentes soit toujours la même.

On vérifiera que les pneus sont bien gonflés à la même pression avant chaque essai.

On peut ainsi faire une liste de précautions à prendre pour tenir compte des contraintes.

2.1.2 Choix du plan d'expériences

On sait qu'il y a deux facteurs à étudier. Les niveaux bas et haut de chaque facteur ont été définis. Les facteurs à conserver constants pendant l'expérimentation sont : le parcours et par conséquent le kilométrage, le sens du parcours, le point de départ, le point d'arrivée et la pression des pneus.

Ayant deux facteurs prenant chacun deux niveaux, le plus simple est de choisir un plan d'expériences factoriel complet 2^2 . La dénomination 2^2 a la signification suivante : le 2 en exposant indique le nombre de facteurs, l'autre 2 indique les nombres de niveaux des facteurs. Ce plan est bien adapté à notre problème puisqu'il correspond exactement à deux facteurs prenant chacun deux niveaux. Les points d'expériences ont pour coordonnées les niveaux bas et les niveaux hauts des facteurs.

On peut présenter ce plan de plusieurs manières qui se complètent les unes les autres. On peut d'abord dessiner le domaine d'étude dans l'espace expérimental, puis ajouter les points d'expériences en tenant compte de leurs coordonnées (Figure 2.1).

On peut aussi représenter les expériences à faire sous forme de tableaux, en utilisant soit les grandeurs habituelles ou légales (km/h et kg), soit les grandeurs codées. Avec les grandeurs légales, le tableau prend le nom de *tableau d'expérimentation* ou de *matrice d'expérimentation*. Avec les grandeurs codées, le tableau prend le nom de *plan d'expériences* ou de *matrice d'expériences*.

La première colonne de la matrice d'expérimentation est utilisée pour indiquer les noms des essais (Tableau 2.2). On peut soit les numéroter, soit leur donner un nom. La deuxième colonne est celle du premier facteur, on y indique successivement les niveaux qu'il faut donner à ce facteur. La troisième colonne est celle du deuxième facteur et on y indique également les niveaux de ce facteur. Le premier essai, essai n° 1 ou essai A, sera exécuté avec une vitesse de 80 km/h et sans surcharge. Le deuxième essai, essai n° 2 ou essai B, sera exécuté avec une vitesse de 120 km/h et sans surcharge. Le troisième essai, essai n° 3 ou essai C, sera exécuté avec une vitesse de 80 km/h et avec une surcharge de 300 kg. Enfin le quatrième essai, essai n° 4 ou essai D, sera exécuté avec une vitesse de 120 km/h et avec une surcharge de 300 kg. Ce tableau est très utile pour l'exécution des essais.

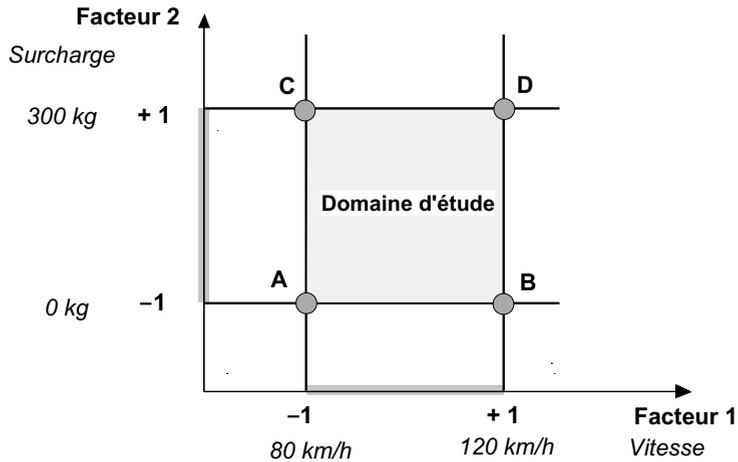


Figure 2.1 – Représentation du plan d'expérimentation.

Tableau 2.2 – Matrice d'expérimentation.

Essai n°	Vitesse Facteur 1	Surcharge Facteur 2
1 (A)	80 km/h	0 kg
2 (B)	120 km/h	0 kg
3 (C)	80 km/h	300 kg
4 (D)	120 km/h	300 kg

Tableau 2.3 – Matrice d'expériences.

Essai n°	Vitesse Facteur 1	Surcharge Facteur 2
1 (A)	-1	-1
2 (B)	+1	-1
3 (C)	-1	+1
4 (D)	+1	+1
Niveau -1	80 km/h	0 kg
Niveau +1	120 km/h	300 kg

La première colonne de la matrice d'expériences (Tableau 2.3) est utilisée de la même manière pour indiquer les noms des essais. La deuxième colonne est celle du premier facteur, on y indique successivement les niveaux qu'il faut donner à ce