

Groupe de recherche en économie
et politique agricoles
Département d'économie agroalimentaire
et des sciences de la consommation

98-09

**RELÈVE EN AGRICULTURE
ET FORMATION AGRICOLE :
UNE ÉQUATION À PLUSIEURS VARIABLES**

Jacques Tondreau
Michel Morisset

Université Laval
© Droits réservés GREPA

Septembre 1998

**Ce document a été déposé par chapitre.
Pour consulter la section désirée, cliquez sur le lien approprié :**

[Table des matières, liste des tableaux, des schémas et des graphiques, résumé, avant-propos, introduction](#)
[CHAPITRE 1 – La formation agricole : l'histoire d'un enjeu](#)
[CHAPITRE 2 – Agriculture spécialisée et discours sur la formation agricole](#)
[CHAPITRE 3 – Rapport à l'école et formation professionnelle](#)
[CHAPITRE 4 – Cadre d'analyse et méthodologie](#)
[CHAPITRE 5 – Relève en agriculture et formation agricole : analyse quantitative](#)
[CHAPITRE 6 – Relève en agriculture et formation agricole : analyse qualitative](#)
[Conclusion et Bibliographie](#)
[ANNEXE 1 – Grilles d'entrevue de groupe et questionnaires 1 et 2](#)
[ANNEXE 2 – Création de variantes pour la variable scol. 1](#)
[ANNEXE 3 – Analyses statistiques utilisées](#)

ANNEXE 3

Analyses statistiques utilisées

Avant d'entreprendre les analyses statistiques, la première étape a été de valider les données afin de s'assurer le plus possible l'exactitude des informations. De nouvelles variables ont été créées à partir des informations contenues dans la base de données. Pour répondre aux objectifs de l'étude et pour les fins des analyses statistiques, une variable dépendante a été créée. Cette variable classe en trois catégories la scolarité de la relève identifiée. Une brève explication des différentes méthodes statistiques appliquées permet de mieux comprendre les résultats obtenus et leurs limites.

Tests d'association

Afin d'épurer la base de données, des analyses de tableaux de contingence ont été utilisées. Les tests d'association de variables s'effectuent sur des données catégoriques. L'objectif de ces tests est de déterminer quelles sont les facteurs ou les caractéristiques de type discret (variables qui se présentent sous forme de choix de réponse) qui exercent une certaine influence sur la classification de la relève identifiée selon la scolarité. Ces tests ont aussi été utilisés pour mettre en évidence les diverses relations pouvant exister entre deux variables nominales contenues dans la base de données. Les hypothèses testées se présentent alors sous la forme suivante :

H_0 : Il n'y a pas d'association entre les deux variables (indépendance)

H_1 : Il y a une association entre les deux variables (dépendance)

Pour la plupart de ces tests, l'une des deux variables testées était le groupe dans lequel se retrouve chaque relève identifiée classée selon sa scolarité. L'autre variable était l'un des facteurs testés. La statistique du Chi-carré, qui mesure la force de l'évidence qu'une association existe entre deux variables, a été utilisée avec un seuil de signification de 5 %. Le rejet de H_0 ($(p > X_2) < 0.05$) indique la présence d'une association entre les deux variables. C'est donc dire que la distribution du facteur testé est différente dépendamment de la classe de scolarité. Ces analyses univariées ont permis de faire une première sélection de variables.

Régression logistique

La régression logistique est un modèle de régression dans lequel la variable réponse (Y) est dichotomique (de type 0,1) ou polytomique (de type 0,1,2,...) et les variables explicatives sont continues et/ou qualitatives, c'est-à-dire qu'elles ne peuvent prendre qu'un nombre fini - généralement faible - de valeurs. La régression linéaire est une technique statistique utilisée pour étudier les liaisons entre les variables. Cette méthode, effectuée sur un échantillon d'observations aléatoires, permet d'étudier la liaison stochastique entre une variable aléatoire Y (la variable dépendante) et p variables certaines (les variables indépendantes). Le but d'une telle analyse est d'expliquer les variations de la variable dépendante. L'usage de la régression permet d'estimer les valeurs des coefficients de régression et de la constante. La relation existant entre les variables s'exprime alors sous la forme d'une équation mathématique.

Afin de déterminer quels facteurs expliquent le mieux la possession d'un diplôme agricole, d'un diplôme général ou de l'absence de diplôme, la méthode de sélection *stepwise* a été appliquée. Dans la détermination du modèle final, un choix statistique judicieux s'avère être le modèle le plus parcimonieux (avec le moins de variables). Le modèle résultant sera ainsi plus stable.