

TRAITEMENT DE DONNÉES AVEC EXCEL • MINITAB • SPSS

2^e édition

GÉRALD BAILLARGEON

MÉTHODES STATISTIQUES

avec applications en gestion,
production, marketing,
relations industrielles et
sciences comptables



CD-ROM
INCLUS

Les Éditions SMG

STATISTE

Les Odyssées
du livre 2002

Prix du ministre
de l'Éducation
2002

.....

MÉTHODES STATISTIQUES

**avec applications
en gestion, production,
marketing, relations industrielles
et sciences comptables**

.....

Table des matières

Avant-propos

I Analyse descriptive des données

Sommaire 1

Objectifs pédagogiques 2

1.1 Introduction: l'information numérique 3

1.2 Étude statistique, statistique descriptive et statistique inférentielle 7

1.3 Les probabilités 8

1.4 Quelques notions fondamentales 8

1.4.1 Ensemble statistique - Population statistique - Unité statistique 8

1.4.2 Caractères 9

1.4.3 Caractère quantitatif - Caractère qualitatif 9

1.4.4 Modalités de caractères 9

1.4.5 Variable statistique - Variable discrète - Variable continue 10

1.4.6 Les échelles de mesure 11

1.4.7 Échantillon - Échantillon aléatoire 14

1.4.8 Unité de mesure 14

1.4.9 Fréquence absolue - Fréquence relative 14

1.5 Présentation des données recueillies 14

1.6 Dépouillement des données et distribution de fréquences absolues 15

1.7 Considérations pratiques dans l'élaboration d'une distribution de fréquences absolues 16

1.8 Dépouillement de données avec Excel : Utilitaire d'analyse 19

1.9 Autre forme élémentaire de dépouillement : le diagramme en feuilles 23

1.10 Distribution de fréquences absolues : cas où le caractère étudié est discret 25

1.11 Distribution de fréquences absolues avec classes ouvertes 25

1.12 Principales représentations graphiques 26

1.12.1 Diagramme en bâtons 26

1.12.2 Histogramme et polygone de fréquences 27

1.13 Tracé de l'histogramme avec Excel 29

1.14 Courbe de fréquences cumulées 33

1.15 Dépouillement de données nominales et ordinales 35

1.16 Diagrammes à secteurs et diagrammes à barres 38

1.17 Autres types de diagrammes 42

1.18 Série de données évoluant dans le temps: série chronologique 45

1.19 Présentation des résultats dans un tableau croisé 46

1.20 Élaboration d'un tableau croisé: Étude sur l'impact de la bureautique 47

1.21 Exercices 50

1.22 Élaboration de diagrammes avec Excel: l'assistant graphique 67

1.23 Élaboration d'un tableau croisé avec Excel 79

Annexe 1 - Traitement avec Minitab et SPSS 83

2 Caractéristiques de tendance centrale et de dispersion

Sommaire 93

Objectifs pédagogiques 94

2.1 Introduction 94

2.2 Définitions des mesures descriptives 95

2.3 Propriétés et remarques sur la moyenne arithmétique, la variance, l'écart-type et l'étendue 99

2.4 Notation indicée et l'opérateur somme (Σ) 101

2.5 Calcul de la moyenne et de la variance: valeurs individuelles et groupées 104

2.6 Calcul simplifié de la moyenne et de la variance: changements d'origine et d'échelle 106

2.7 La règle de Tchebycheff 109

2.8 Autre mesure de tendance centrale: la médiane 110

2.9 Le mode 113

2.10 Autres mesures de position: les quantiles 115

2.11 Calcul des quartiles 116

2.12 Quartiles et diagramme en boîte 118

2.13 Caractéristiques de forme: asymétrie et aplatissement 121

2.14 Statistiques descriptives et échelles de mesure 122

2.15 Calcul de diverses statistiques selon le type de données 123

2.16 Statistiques descriptives avec Excel et l'utilitaire d'analyse 128

2.17 Fonctions statistiques de Excel 131

2.18 Exercices 134

2.19 Exercices de révision et de synthèse 141

Annexe 2 - Traitement avec Minitab et SPSS 151

3 Calcul des probabilités

Sommaire 155

Objectifs pédagogiques 156

3.1 Introduction: phénomènes probabilistes 156

3.2 Notions d'épreuve, de résultat, d'espace échantillonnal et d'événement 157

3.3 La notion de probabilité 160

3.4 Propriétés régissant le calcul des probabilités 162

3.5 Probabilités totales: événements ne s'excluant pas 165

3.6 Événements liés, probabilités conditionnelles, probabilités composées et événements indépendants 166

3.6.1 Événements liés 166

3.6.2 Probabilité conditionnelle 166

3.6.3 Probabilités composées 167

3.6.4 Événements indépendants 168

3.7 Probabilités des causes: formules de Bayes 173

3.8 L'analyse combinatoire, l'échantillonnage et les probabilités 177

3.9 Génération de nombres aléatoires avec Excel 181

3.10 Cheminement de réflexion pour résoudre les exercices sur les probabilités 183

3.11 Exercices 183

Annexe 3 - Traitement avec Minitab et SPSS 193

4 Variables aléatoires, loi binomiale et loi de Poisson

Sommaire 197

Objectifs pédagogiques 198

- 4.1 Généralités sur les notions probabilistes et les notions statistiques 198
- 4.2 Les notions de variable aléatoire et de loi de probabilité 199
- 4.3 Représentation graphique de la loi de probabilité 201
- 4.4 Espérance mathématique et variance d'une variable aléatoire discrète 202
- 4.5 Lois de probabilité importantes 206
- 4.6 Conditions d'application de la loi binomiale 207
- 4.7 La loi binomiale 209
- 4.8 Fonction de Microsoft Excel pour calculer les probabilités binomiales 212
- 4.9 Représentation graphique de la distribution binomiale 214
- 4.10 La loi de Poisson 216
- 4.11 Représentation graphique de la distribution de Poisson 218
- 4.12 Approximation de la loi binomiale par la loi de Poisson 220
- 4.13 Fonction de Microsoft Excel pour calculer les probabilités de la loi de Poisson 223
- 4.14 Notions sur le processus de Poisson 225
- 4.15 Cheminement de réflexion pour résoudre les exercices 228
- 4.16 Exercices 228
- 4.17 Exercice de révision et de synthèse 239

Annexe 4 - Traitement avec Minitab et SPSS 240

5 Lois continues: loi uniforme, loi normale et loi exponentielle

Sommaire 243

Objectifs pédagogiques 244

- 5.1 Variable aléatoire continue et loi de probabilité continue 244
- 5.2 Caractéristique de qualité et la distribution normale 246
- 5.3 Variable centrée réduite et la loi normale 248
- 5.4 L'aire sous la courbe normale et seuil de probabilité 254
- 5.5 Approximation de la loi binomiale par la loi normale 259
- 5.6 Fonction de Microsoft Excel pour calculer les probabilités de la loi normale 262
- 5.7 Autre loi continue: la loi exponentielle 266
- 5.8 Fonction de Microsoft Excel pour calculer les probabilités de la loi exponentielle 268
- 5.9 Tableau-synthèse des lois continues 271
- 5.10 Exercices 271
- 5.11 Exercices de révision et de synthèse 279

ANNEXE : Appréciation de la normalité d'une série de données avec Excel 280

Annexe 5 - Traitement avec Minitab et SPSS 290

6 Échantillonnage et estimation de paramètres

Sommaire 297

Objectifs pédagogiques 298

- 6.1 Introduction 298
- 6.2 Les sondages: échantillonnage d'une population 299
- 6.3 Exemples d'un sondage 299
- 6.4 L'échantillon aléatoire 300
- 6.5 Principe de la construction d'un échantillon 300
- 6.6 Construction de l'échantillon à l'aide d'une table de nombres aléatoires 301
- 6.7 Construction de l'échantillon par tirage systématique 303
- 6.8 Fluctuations d'échantillonnage d'une moyenne 305

6.9	<i>Théorème central limite: forme de la distribution de $\bar{X}$</i>	309
6.10	<i>Transformation de la variable aléatoire $\bar{X}$ en une variable aléatoire centrée réduite</i>	311
6.11	<i>Propriétés de la distribution d'échantillonnage de $\bar{X}$</i>	272
6.12	<i>L'estimation de paramètres: l'objectif fondamental de l'échantillonnage d'une population</i>	315
6.13	<i>Estimation d'une moyenne par intervalle de confiance</i>	317
6.14	<i>Marge d'erreur associée à l'estimation et taille d'échantillon requise pour ne pas excéder la marge d'erreur</i>	321
6.15	<i>La distribution de Student</i>	323
6.15.1	<i>Propriétés de la loi de Student</i>	324
6.15.2	<i>Valeurs tabulées du T de Student et leur signification</i>	324
6.16	<i>Estimation de μ par intervalle: population normale, variance inconnue et petit échantillon ($n < 30$)</i>	325
6.17	<i>Fluctuations d'échantillonnage d'une proportion</i>	328
6.18	<i>Estimation d'une proportion par intervalle de confiance</i>	330
6.19	<i>Marge d'erreur associée à l'estimation de p et taille d'échantillon requise</i>	332
6.20	<i>Détermination de la taille d'échantillon avec Excel pour une marge d'erreur fixée</i>	334
6.21	<i>Exercices</i>	335
6.22	<i>Exercices de révision et de synthèse</i>	342
	Annexe 6 - Traitement avec Minitab et SPSS	348

7 Tests d'hypothèses sur une moyenne, une proportion et une variance

<i>Sommaire</i>	355
<i>Objectifs pédagogiques</i>	356
7.1	<i>Principe d'un test d'hypothèse</i> 356
7.2	<i>Formulation d'hypothèses statistiques</i> 357
7.3	<i>Seuil de signification du test</i> 359
7.4	<i>Risques d'erreur et mesure de crédibilité de H_0</i> 359
7.5	<i>Comment exécuter un test d'hypothèse: démarche à suivre</i> 360
7.6	<i>Test sur une moyenne</i> 361
7.7	<i>Valeurs critiques de la moyenne</i> 367
7.8	<i>Utilisation du seuil descriptif du test («p-value») pour tester une hypothèse</i> 370
7.9	<i>Test sur la moyenne avec Excel et calcul du seuil descriptif du test</i> 372
7.10	<i>Risque de première espèce et risque de deuxième espèce</i> 374
7.11	<i>Schématisation des deux risques d'erreur sur la distribution d'échantillonnage de $\bar{X}$</i> 376
7.12	<i>Calcul du risque de 2e espèce avec Excel et tracé de la courbe</i> 383
7.13	<i>Test sur une proportion</i> 386
7.14	<i>La loi de khi-deux</i> 389
7.15	<i>Test sur une variance</i> 390
7.16	<i>Exercices</i> 391
7.17	<i>Exercice de révision et de synthèse</i> 398
Annexe 7 - Traitement avec Minitab et SPSS 401	

8 Tests sur deux moyennes, deux proportions et deux variances

<i>Sommaire</i>	407
<i>Objectifs pédagogiques</i>	408
8.1 <i>Introduction</i>	408
8.2 <i>Comparaison de deux moyennes</i>	408
8.3 <i>Tests sur deux moyennes</i>	410
8.4 <i>Test de comparaison de deux moyennes avec Excel</i>	415
8.5 <i>Test de comparaison de deux échantillons appariés</i>	419

- 8.6 Test de deux échantillons appariés avec Excel 422
- 8.7 Comparaison de variances et la distribution de Fisher-Snedecor 425
- 8.8 Distribution d'échantillonnage du quotient de deux variances 427
- 8.9 Test sur l'égalité de deux variances 428
- 8.10 Test sur l'égalité de deux moyennes avec variances inconnues et inégalités: test de Satterthwaite 431
- 8.11 Comparaison de deux proportions (pourcentages) 435
- 8.12 Exercices 440
- 8.13 Exercice de révision et de synthèse 454
- Annexe 8 - Traitement avec Minitab et SPSS 457**

9 Tests d'ajustement et tableau de contingence

- Sommaire 463
- Objectifs pédagogiques 464
- 9.1 Introduction 464
- 9.2 Application du test de Pearson: test de conformité entre deux distributions (test d'ajustement) 466
- 9.3 Test de conformité: hypothèses statistiques et règle de décision 467
- 9.4 Vérification de la qualité d'ajustement pour les distributions binomiale, de Poisson, normale et exponentielle 470
- 9.5 Utilisation de Microsoft Excel pour l'ajustement d'une loi binomiale et le test du χ^2 473
- 9.6 Utilisation de Microsoft Excel pour l'ajustement d'une loi de Poisson 476
- 9.7 Utilisation de Microsoft Excel pour l'ajustement d'une loi normale 480
- 9.8 Le test de khi-deux et la représentativité de l'échantillon 481
- 9.9 Tableau croisé et test d'indépendance 484
- 9.10 Mesures d'association entre deux caractères 492
 - 9.10.1 Coefficient de contingence 492
 - 9.10.2 Le V de Cramer 492
 - 9.10.3 Coefficient phi (ϕ) 492
- 9.11 Utilisation de Excel pour un test d'indépendance 493
- 9.12 Test sur l'égalité de k proportions 499
- 9.13 Sommaire concernant le test d'indépendance et test d'égalité de k proportions 507
- 9.14 Exercices 508
- 9.15 Exercices de révision et de synthèse 523
- Annexe 9 - Traitement avec Minitab et SPSS 530**

10 Comparaison de plusieurs moyennes: l'analyse de variance

- Sommaire 535
- Objectifs pédagogiques 536
- 10.1 Comparaison de plusieurs moyennes: analyse de variance à un seul facteur 536
- 10.2 Détermination des sommes de carrés et tableau d'analyse de variance 538
- 10.3 Test de l'influence du facteur A: test de Fisher 542
- 10.4 Analyse de variance avec Excel 545
- 10.5 Comparaisons multiples: méthode de Tukey-Kramer 547
 - 10.5.1 Utilisation de Excel pour la méthode de comparaison de Tukey 550
- 10.6 Analyse de variance suivant deux facteurs 551
- 10.7 Plan expérimental avec restriction: plan en blocs randomisés 551
- 10.8 Analyse de variance pour le plan avec blocs à l'aide de Excel 558
- 10.9 Plan d'expérience à deux facteurs contrôlés: plan factoriel 560

- 10.10 Plan factoriel sans interaction 560
- 10.11 Plan factoriel avec interaction: expérience à deux facteurs contrôlés avec n répétitions par traitement 564
- 10.12 Analyse de variance pour un plan factoriel à l'aide de Excel 569
- 10.13 Exercices 574
- 10.14 Exercices de révision et de synthèse 585
- Annexe 10 - Traitement avec Minitab et SPSS 290**

11 Corrélation linéaire et régression linéaire simple

- Sommaire 593
- Objectifs pédagogiques 594
- 11.1 Étude de corrélation et analyse de régression 594
- 11.2 Calcul du coefficient de corrélation linéaire 598
- 11.3 La corrélation est-elle significative? 602
- 11.4 Test d'hypothèse et estimation par intervalle de confiance 606
 - 11.4.1 Transformation de Fisher 606
 - 11.4.2 Test de l'hypothèse $H_0: \rho = \rho_0$ et valeurs critiques 607
 - 11.4.3 Intervalle de confiance pour ρ 607
- 11.5 La régression linéaire simple 609
- 11.6 Choix de la variable dépendante 609
- 11.7 Détermination de la droite de régression: méthode des moindres carrés 609
- 11.8 Décomposition de la variation dans la variable dépendante 613
- 11.9 Test de signification sur β_1 avec le t de Student 616
- 11.10 Autre façon de tester si la régression est significative: l'analyse de variance 618
- 11.11 La régression avec Excel 621
- 11.12 Tracé du diagramme de dispersion avec Excel 624
- 11.13 Estimation et prévision par intervalle de confiance à l'aide de la régression 625
- 11.14 Prévision de y pour une nouvelle observation de x 626
- 11.15 Quelques considérations pratiques dans l'application des méthodes d'analyse de régression 630
- 11.16 Vérification des hypothèses fondamentales: analyse des résidus 631
- 11.17 Exercices 640
- 11.18 Exercices de révision et de synthèse 659
- Annexe 11 - Traitement avec Minitab et SPSS 664**

12 Régression linéaire multiple

- Sommaire 673
- Objectifs pédagogiques 674
- 12.1 Introduction 674
- 12.2 Modèle linéaire à plusieurs variables explicatives 675
- 12.3 Démarche à suivre dans l'élaboration d'un modèle de régression multiple 678
- 12.4 Détermination de l'équation de régression: estimation des paramètres du modèle 679
- 12.5 Analyse de variance en régression multiple 681
- 12.6 Comment déterminer si la régression est significative dans son ensemble? 685
- 12.7 La régression multiple avec Excel 688
- 12.8 Autres tests de signification en régression multiple 692
- 12.9 Contribution marginale de chaque variable explicative 692
- 12.10 Autre façon de tester la contribution marginale: le «F partiel» 694

- 12.11 Estimation de $E(y_h)$ par intervalle de confiance 697
- 12.12 Intervalle de prévision pour une valeur individuelle y_h 699
- 12.13 Utilisation de l'information qualitative dans un modèle de régression multiple: variables auxiliaires 701
- 12.14 Comment tester la nullité d'un sous-ensemble de paramètres β_i du modèle? 707
- 12.15 Colinéarité 710
- 12.16 Autres modèles statistiques linéaires 712
- 12.17 La notion de corrélation partielle 713
- 12.18 Corrélation partielle et test de contribution marginale 715
- 12.19 Méthodes pratiques pour élaborer une équation de régression multiple 718
- 12.20 Exercices 726
- 12.21 Exercices de révision et de synthèse 745
- Annexe 12 - Traitement avec Minitab et SPSS 751**

13 Introduction aux séries chronologiques

- Sommaire 755
- Objectifs pédagogiques 756
- 13.1 Séries chronologiques: introduction 756
 - 13.1.1 Représentation graphique d'une série chronologique 757
 - 13.1.2 Composantes d'une série chronologique 759
 - 13.1.3 Modèle descriptif d'une série chronologique 759
- 13.2 Recherche de la tendance 760
 - 13.2.1 Recherche de la tendance à l'aide de la régression 760
 - 13.2.2 Estimation de la tendance par la méthode des moyennes mobiles 764
 - 13.2.3 Estimation de la tendance par le lissage exponentiel 764
- 13.3 Utilisation de Excel pour les moyennes mobiles et le lissage exponentiel 771
- 13.4 Variations saisonnières: comment tenir compte de la composante saisonnière dans une série chronologique 774
- 13.5 Exercices 780
- 13.6 Exercice de révision et de synthèse 786
- Annexe 13 - Traitement avec Minitab et SPSS 787**

Corrigé des exercices d'apprentissage et réponses aux exercices 795

Tables statistiques 845

- Table A1 : Table de la loi binomiale 846
- Table A2 : Table de la loi de Poisson 850
- Table A3 : Table de la loi normale centrée réduite 855
- Table A4 : Table de la loi de Student 856
- Table A5 : Table de la loi du khi-deux 858
- Table A6 : Table de la loi de Fisher-Snedecor 859
- Table A7: Table de Tukey 866
- Table A8 : Table de valeurs critiques r_c pour le coefficient de corrélation de Pearson 868
- Table A9 : Table de la transformation de Fisher 869
- Table A10 : Table de Durbin-Watson 870
- Table A11 : Table de nombres aléatoires 872

Annexe 14: Fonctions statistiques et modèles probabilistes dans Excel 873

Annexe 15: Introduction au logiciel Microsoft Excel: Office 97 881

Index analytique 891

.....

MÉTHODES STATISTIQUES

avec applications en gestion, production, marketing,
relations industrielles et sciences comptables

2^e édition

GÉRALD BAILLARGEON

L'objectif premier de cet ouvrage est de présenter les outils statistiques utilisés en sciences de la gestion à l'aide de données réelles provenant d'entreprises, de journaux, de magazines, de revues spécialisées, de rapports de recherche, de sondages... et ceci avec une variété d'applications provenant des domaines de la gestion, des ressources humaines, de la psychologie industrielle, de la production, du marketing, de l'économie, des relations industrielles et des sciences comptables.

Nous présentons les divers outils statistiques et probabilistes selon une approche pédagogique bien éprouvée, associée à une démarche participative qui permet l'apprentissage rapide et efficace et ceci à travers de nombreuses applications provenant de situations réelles.

L'analyse des données est supportée par une utilisation importante du tableur *Excel* de Microsoft et de deux logiciels statistiques, *Minitab* et *SPSS* (en annexe de chaque chapitre).

Structure pédagogique de l'ouvrage

La conception pédagogique de cet ouvrage se caractérise par les points suivants :

- a) Chaque chapitre débute par un sommaire, ce qui permet d'avoir une vue d'ensemble et très détaillée des notions à acquérir; on y présente également les objectifs pédagogiques (objectif général et objectifs spécifiques) associés au contenu du chapitre.
- b) Les outils statistiques sont présentés avec une approche intuitive, souvent à partir d'applications réelles; cette présentation se veut structurée et très visuelle.
- c) La participation active à travers de nombreux exercices d'apprentissage permet l'assimilation pratique et rapide des outils présentés. Ces exercices d'apprentissage peuvent servir de session de travail lors de formation en classe.
- d) Tous les exemples et les exercices d'apprentissage que nous présentons sont titrés pour donner une idée immédiate du concept ou de l'outil que nous voulons mettre en évidence.
- e) L'utilisation de *Excel* de Microsoft Office (version 8.0 et version 7.0) est fréquente et permet de réaliser rapidement et avec facilité le traitement informatique des données pour mieux se concentrer sur l'analyse; il permet également de réaliser des diagrammes et graphiques de présentation professionnelle. Nous présentons également en annexe de chaque chapitre, le traitement de divers exemples avec les logiciels statistiques *Minitab* et *SPSS*.
- f) La présentation des procédures requises pour le traitement de données avec les logiciels *Excel* et *Minitab* est très visuelle et facile à suivre puisqu'on y retrouve imprimées les unes à la suite des autres les fenêtres et les zones de dialogue telles qu'elles apparaissent à l'écran. De plus, les fenêtres et les zones de dialogue sont souvent commentées pour assurer une meilleure compréhension de la démarche à suivre et de l'interprétation des résultats.
- g) Chaque chapitre se termine par une série d'exercices qui permettent de mettre en application les connaissances ainsi que des exercices de révision et de synthèse; on y trouve à la fin de l'ouvrage les réponses aux nombreux exercices de fin de chapitre.

Sommaire de l'ouvrage

1. Analyse descriptive des données
2. Caractéristiques de tendance centrale et de dispersion
3. Calcul des probabilités
4. Variables aléatoires, loi binomiale et loi de Poisson
5. Lois continues : la loi uniforme, la loi normale et la loi exponentielle
6. Échantillonnage et estimation de paramètres
7. Tests sur une moyenne, une proportion et une variance
8. Tests sur deux moyennes, deux proportions et deux variances
9. Test d'ajustement et tableau de contingence
10. Comparaison de plusieurs moyennes : l'analyse de variance
11. Corrélation linéaire et régression linéaire simple
12. Régression linéaire multiple
13. Introduction aux séries chronologiques
- Annexe : Fonctions statistiques et modèles probabilistes dans Excel
- Annexe : Introduction au logiciel Microsoft Excel : Office 97
- Tables statistiques - Réponses aux exercices - Index

Les Éditions SMG

5365, boulevard Jean XXIII, bureau 203
Trois-Rivières-Ouest (Québec) G8Z 4A6

Téléphone : (819) 376-5650
Télécopieur : (819) 373-2904

DISTRIBUTEUR EXCLUSIF POUR TOUTS LES PAYS
(SAUF LE CANADA) :

Technique et Documentation Lavoisier
14, rue de Provigny
F-94236 Cachan cedex France

ISBN 2-89094-136-1

© Excel de Microsoft, MINITAB et SPSS sont des marques déposées.