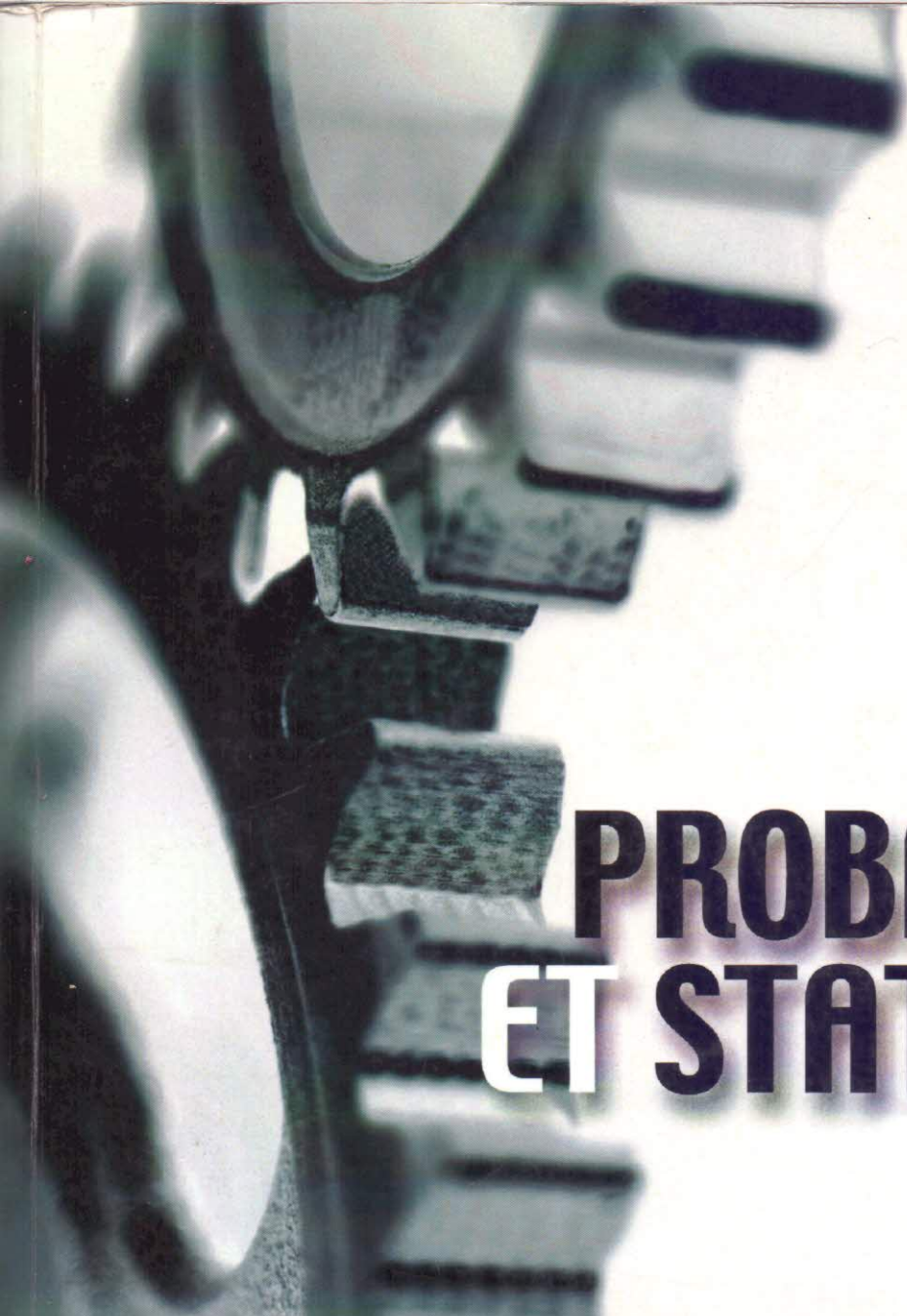




GÉRALD BAILLARGEON

PROBABILITÉS ET STATISTIQUE

AVEC APPLICATIONS EN TECHNOLOGIE ET EN INGÉNIERIE

Traitement de données avec Excel

Les Éditions SMG



**CD - ROM
INCLUS**

Fichiers : **Excel**
Minitab
SPSS et
Statistica

GÉRALD BAILLARGEON

PROBABILITÉS ET STATISTIQUE

avec applications en technologie et en ingénierie

L'ouvrage

Nous présentons dans cet ouvrage les principaux outils statistiques et modèles probabilistes appliqués à différents secteurs de la technologie et de l'ingénierie. L'acquisition des connaissances est facilitée par 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'ouvrage est particulièrement riche en applications industrielles et se veut axé sur l'application pratique des méthodes statistiques.

Nous avons cru bon également d'intégrer à ce livre l'utilisation d'un logiciel (le tableur *Excel* de Microsoft Office) qui permet le traitement de données de façon simple, tout en couvrant pratiquement tous les outils qui sont présentés ici.

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 majorité des exemples ou exercices sont constitués de données réelles provenant de cas d'entreprises, d'articles de revues spécialisées, d'organismes, de journaux ou de rapports de recherche.
- d) 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 la formation en classe.
- e) Le CD-Rom d'accompagnement comporte, en plus des fichiers *Excel*, les fichiers de données permettant le traitement avec les logiciels statistiques *Minitab*, *SPSS* et *Statistica*.

Sommaire de l'ouvrage

1. Introduction : l'information numérique et les outils statistiques
2. Analyse descriptive des données
3. Caractéristiques de tendance centrale et de dispersion
4. Calcul des probabilités
5. Variables aléatoires, loi binomiale, loi de Poisson et loi hypergéométrique
6. Lois continues : loi uniforme, loi normale et loi exponentielle
7. Échantillonnage et estimation de paramètres
8. Tests d'hypothèses sur une moyenne, une proportion et une variance
9. Comparaisons de moyennes et de variances
10. Comparaisons de proportions, test d'ajustement et tableau croisé
11. Corrélation linéaire et régression
12. Introduction à la régression multiple

Corrigé des exercices d'apprentissage et réponses aux exercices

Tables statistiques

Fonctions statistiques et modèles probabilistes dans Excel

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 2-89094-137-X

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

Table des matières

Avant-propos

1 Introduction: l'information numérique et les outils statistiques

Sommaire	1
Objectifs pédagogiques	2
1.1 Introduction: l'information numérique	2
1.2 Outils statistiques et probabilités en milieu industriel	6
1.3 Détermination des causes d'un problème à l'aide d'un diagramme de Pareto	7
1.3.1 Diagramme de Pareto	7
1.3.2 Analyse de Pareto: Étapes à suivre et tracé du diagramme	7
1.4 Diagramme d'Ishikawa	12
1.5 Expérimentation industrielle	13
1.6 Étude statistique, statistique descriptive et statistique inférentielle	13
1.7 Les probabilités	14
1.8 La méthodologie «SIX SIGMA»	14
1.9 Exercices	15

2 Analyse descriptive des données

Sommaire	17
Objectifs pédagogiques	18
2.1 Introduction	18
2.2 Quelques notions fondamentales	18
2.2.1 Ensemble statistique - Population statistique - Unité statistique	18
2.2.2 Caractères	19
2.2.3 Caractère quantitatif - Caractère qualitatif	19
2.2.4 Modalités de caractères	19
2.2.5 Variable statistique - Variable discrète - Variable continue	20
2.2.6 Variable indépendante et variable dépendante	21
2.2.7 Les échelles de mesure	23
2.2.8 Échantillon - Échantillon aléatoire	25
2.2.9 Unité de mesure	26
2.2.10 Fréquence absolue - Fréquence relative	26
2.3 Présentation des données recueillies	26
2.4 Dépouillement des données et distribution de fréquences absolues	27
2.5 Considérations pratiques dans l'élaboration d'une distribution de fréquences absolues	28
2.6 Autre forme élémentaire de dépouillement : le diagramme en feuilles	31
2.7 Distribution de fréquences absolues : cas où le caractère étudié est discret	34
2.8 Distribution de fréquences absolues avec classes ouvertes	34

2.9 Principales représentations graphiques	35
2.9.1 Diagramme en bâtons	35
2.9.2 Histogramme et polygone de fréquences	36
2.10 Courbe de fréquences cumulées	40
2.11 Dépouillement de données nominales et ordinales	42
2.12 Diagrammes à secteurs et diagrammes à barres	45
2.13 Exercices	51
2.14 Exercices de révision et de synthèse	61
Annexe 1 - Traitement avec Excel / Histogramme et diagrammes	65

3 Caractéristiques de tendance centrale et de dispersion

Sommaire	79
Objectifs pédagogiques	80
3.1 Introduction	80
3.2 Définitions des mesures descriptives	81
3.3 La règle de Tchebycheff	87
3.4 Autre mesure de tendance centrale: la médiane	88
3.5 Le mode	92
3.6 Autres mesures de position: les quantiles	93
3.7 Calcul des quartiles	94
3.8 Quartiles et diagramme en boîte	96
3.9 Caractéristiques de forme: asymétrie et aplatissement	99
3.10 Exercices	100
3.11 Exercices de révision et de synthèse	108
Annexe 2 - Traitement avec Excel / Statistiques descriptives avec Excel	113

4 Calcul des probabilités

Sommaire	119
Objectifs pédagogiques	120
4.1 Introduction: phénomènes probabilistes	120
4.2 Notions d'épreuve, de résultat, d'espace échantillonnal et d'événement	121
4.3 La notion de probabilité	123
4.4 Axiomes régissant le calcul des probabilités	125
4.5 Vocabulaire et opérations sur les événements	128
4.6 Probabilités totales: événements ne s'excluant pas	129
4.7 Événements liés, probabilités conditionnelles, probabilités composées et événements indépendants	131
4.7.1 Événements liés	131
4.7.2 Probabilité conditionnelle	131
4.7.3 Probabilités composées	132
4.7.4 Événements indépendants	134
4.8 Probabilités des causes: formules de Bayes	136
4.9 L'analyse combinatoire, l'échantillonnage et les probabilités	142
4.10 Cheminement de réflexion pour résoudre les exercices sur les probabilités	147
4.11 Exercices	148
4.12 Exercices de révision et de synthèse	159

5 Variables aléatoires discrètes, loi binomiale, loi de Poisson et loi hypergéométrique

Sommaire	161
Objectifs pédagogiques	162
5.1 Généralités sur les notions probabilistes et les notions statistiques	162
5.2 Les notions de variable aléatoire et de loi de probabilité	163
5.3 Représentatif graphique de la loi de probabilité	165
5.4 Espérance mathématique et variance d'une variable aléatoire discrète	166
5.5 Lois de probabilité importantes	170
5.6 Conditions d'application de la loi binomiale	171
5.7 La loi binomiale	173
5.8 Représentation graphique de la distribution binomiale	176
5.9 La loi de Poisson	179
5.10 Représentation graphique de la distribution de Poisson	181
5.11 Approximation de la loi binomiale par la loi de Poisson	182
5.12 Notions sur le processus de Poisson	186
5.13 La loi hypergéométrique	189
5.14 Tableau-synthèse des lois	193
5.15 Cheminement de réflexion pour résoudre les exercices	194
5.16 Exercices	194
5.17 Exercice de révision et de synthèse	207
Annexe 3 - Traitement avec Excel / Lois de probabilités discrètes	209

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

Sommaire	217
Objectifs pédagogiques	218
6.1 Variable aléatoire continue et loi de probabilité continue	218
6.2 Loi uniforme	220
6.3 Variable aléatoire continue et la distribution normale	221
6.4 Variable centrée réduite et la loi normale	223
6.5 L'aire sous la courbe normale et seuil de probabilité	230
6.6 Approximation de la loi binomiale par la loi normale	232
6.7 Combinaison de variables aléatoires normales	235
6.8 Théorème central limite	237
6.9 La loi exponentielle	238
6.10 Tableau-synthèse des lois continues	241
6.11 Exercices	241
6.12 Exercices de révision et de synthèse	257
Annexe 4 - Traitement avec Excel / Lois de probabilité continues	259
Annexe 5 : Appréciation de la normalité d'une série de données	273

7 Échantillonnage et estimation de paramètres

Sommaire	281
Objectifs pédagogiques	282
7.1 Introduction	282
7.2 Les sondages: échantillonnage d'une population	283
7.3 L'échantillon aléatoire	283

7.4	Principe de la construction d'un échantillon	284
7.5	Construction de l'échantillon à l'aide d'une table de nombres aléatoires	284
7.6	Construction de l'échantillon par tirage systématique	286
7.7	Fluctuations d'échantillonnage d'une moyenne	288
7.8	Théorème central limite: forme de la distribution de $\bar{X}$	292
7.9	Transformation de la variable aléatoire $\bar{X}$ en une variable aléatoire centrée réduite	294
7.10	Propriétés de la distribution d'échantillonnage de $\bar{X}$	295
7.11	L'estimation de paramètres: l'objectif fondamental de l'échantillonnage d'une population	298
7.12	Estimation d'une moyenne par intervalle de confiance	301
7.13	Marge d'erreur associée à l'estimation et taille d'échantillon requise pour ne pas excéder la marge d'erreur	304
7.14	La distribution de Student	306
7.14.1	Propriétés de la loi de Student	306
7.14.2	Valeurs tabulées du T de Student et leur signification	307
7.15	Estimation de μ par intervalle: population normale, variance inconnue et petit échantillon ($n < 30$)	308
7.16	Fluctuations d'échantillonnage d'une proportion	311
7.17	Estimation d'une proportion par intervalle de confiance	313
7.18	Marge d'erreur associée à l'estimation de p et taille d'échantillon requise	314
7.19	Exercices	317
7.20	Exercices de révision et de synthèse	325
	Annexe 6 - Traitement avec Excel / Intervalle de confiance et échantillonnage	329

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

	Sommaire	337
	Objectifs pédagogiques	338
8.1	Principe d'un test d'hypothèse statistique	339
8.2	Formulation d'hypothèses statistiques	339
8.3	Seuil de signification du test	341
8.4	Risques d'erreur et mesure de crédibilité de H_0	342
8.5	Comment exécuter un test d'hypothèse: démarche à suivre	343
8.6	Test sur une moyenne	343
8.7	Valeurs critiques de la moyenne	349
8.8	Utilisation du seuil descriptif du test (valeur p) pour tester une hypothèse	352
8.9	Risque de première espèce et risque de deuxième espèce	354
8.10	Schématisation des deux risques d'erreur sur la distribution d'échantillonnage de la moyenne	355
8.11	Test sur une moyenne avec contrôle sur α et β	361
8.12	Test sur une proportion	364
8.13	La loi du khi-deux	367
8.14	Test sur une variance	369
8.15	Test statistique sur la normalité d'une série de données: test de Shapiro-Wilk	371
8.16	Exercices	375
8.17	Exercices de révision et de synthèse	387
	Annexe 7 - Traitement avec Excel / Tests d'hypothèses sur une moyenne et une variance	391

9 Comparaisons de moyennes et de variances

Sommaire	401
Objectifs pédagogiques	402
9.1 Introduction	402
9.2 Comparaison de deux moyennes	403
9.3 Test sur l'égalité de deux moyennes	404
9.4 Test de comparaison de deux échantillons appariés	411
9.5 Comparaison de variances et la distribution de Fisher-Snedecor	415
9.6 Distribution d'échantillonnage du quotient de deux variances	417
9.7 Test sur l'égalité de deux variances	418
9.8 Test sur l'égalité de deux moyennes avec variances inconnues et inégales: test de Satterhwaite	421
9.9 Comparaison de plusieurs moyennes: introduction à l'analyse de variance	425
9.10 Test de l'influence du facteur: test de Fisher	430
9.11 Exercices	433
9.12 Exercice de révision et de synthèse	446
Annexe 8-Traitement avec Excel / Comparaisons de moyennes	451

10 Comparaison de proportions, test d'ajustement et tableau croisé

Sommaire	457
Objectifs pédagogiques	458
10.1 Introduction	458
10.2 Comparaison de deux proportions	459
10.3 Test sur l'égalité de k proportions	462
10.4 Test d'ajustement	465
10.5 Test de conformité: hypothèses statistiques et règle de décision	467
10.6 Vérification de la qualité d'ajustement pour les distributions binomiale, Poisson, normale et exponentielle	469
10.7 Tableau croisé et test d'indépendance	478
10.8 Mesures d'association entre deux caractères	484
10.8.1 Coefficient de contingence	484
10.8.2 Le V de Cramer	485
10.8.3 Coefficient phi (ϕ)	485
10.9 Exercices	488
10.10 Exercices de révision et de synthèse	502
Annexe 9-Traitement avec Excel / Test d'ajustement et tableau croisé	508

11 Corrélation linéaire et régression

Sommaire	515
Objectifs pédagogiques	516
11.1 Étude de corrélation et analyse de régression	516
11.2 Calcul du coefficient de corrélation linéaire	520
11.3 La corrélation est-elle significative?	523

11.4	Test de l'hypothèse $H_0: \rho = \rho_0$ et estimation par intervalle de confiance	526
11.4.1	Transformation de Fisher	526
11.4.2	Test de l'hypothèse $H_0: \rho = \rho_0$ et valeurs critiques	526
11.4.3	Intervalle de confiance pour ρ	527
11.5	La régression linéaire simple	529
11.6	Choix de la variable dépendante	529
11.7	Détermination de la droite de régression: méthode des moindres carrés	529
11.8	Décomposition de la variation dans la variable dépendante	533
11.9	Test de signification sur β_1 avec le t de Student	536
11.10	Autre façon de tester si la régression est significative: l'analyse de variance	538
11.11	Estimation et prévision par intervalle de confiance à l'aide de la régression	541
11.12	Prévision de y pour une nouvelle observation de x	543
11.13	Quelques considérations pratiques dans l'application des méthodes d'analyse de régression	544
11.14	Vérification des hypothèses fondamentales: analyse des résidus	546
11.15	Régression avec transformations sur les variables	555
11.16	Exercices	558
11.17	Exercices de révision et de synthèse	582
Annexe 10-Traitement avec Excel / Corrélation linéaire et régression		590

12 Introduction à la régression multiple

Sommaire	601	
Objectifs pédagogiques	602	
12.1 Introduction	602	
12.2 Détermination de l'équation de régression: estimation des paramètres du modèle		605
12.3 Analyse de variance en régression multiple	607	
12.4 Comment déterminer si la régression est significative dans son ensemble?		611
12.5 Test de signification sur chaque paramètre de régression	613	
12.6 Exercices	619	
12.7 Exercices de révision et de synthèse	631	
Annexe II-Traitement avec Excel / Régression multiple	637	

Corrigé des exercices d'apprentissage et réponses aux exercices	643
--	------------

Tables statistiques 687

Table A1 : Table de la loi binomiale	688
Table A2 : Table de la loi de Poisson	692
Table A3 : Table de la loi normale centrée réduite	697
Table A4 : Table de la loi de Student	698
Table A5 : Table de la loi du khi-deux	700
Table A6 : Table de la loi de Fisher-Snedecor	701
Table A7 : Table Shapiro-Wilk	708
Table A8 : Table de valeurs critiques r_c pour le coefficient de corrélation de Pearson	709
Table A9 : Table de la transformation de Fisher	710
Table A10 : Table de Durbin-Watson	711
Table A11 : Table de nombres aléatoires	713

Annexe 12: Fonctions statistiques et modèles probabilistes dans Excel	715
--	------------

Index analytique	723
-------------------------	------------