

Florence Grandchamp
Drita Neziri
Abdelkader Amara
Raymond Thériault

COLLECTE DE DONNÉES EN CONTEXTE FONDAMENTAL

MAT_{SN} 4272 2

FORMATION DE BASE DIVERSIFIÉE



Graphismes, notations
et symboles utilisés
dans ce module



Graphismes, notations et symboles

(x, y)	couple de coordonnées x et y
\bar{x}	moyenne des valeurs de x
r	coefficient de corrélation
$ r $	valeur absolue de r
\approx	est approximativement égal à

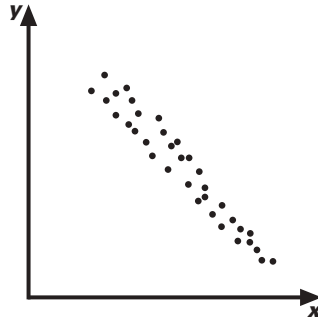
Rappel de quelques notions



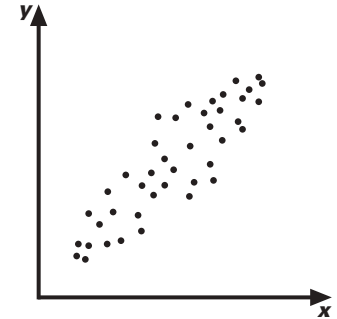
Qualifier une corrélation linéaire



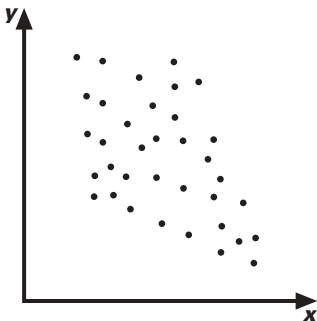
La relation est **parfaite** entre les deux caractères.



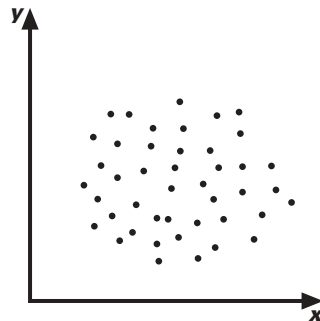
La relation est **forte** entre les deux caractères.



La relation est **moyenne** entre les deux caractères.



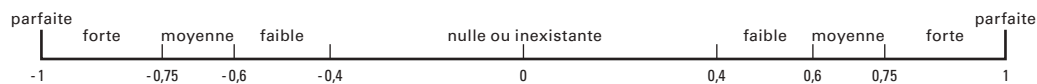
La relation est **faible** entre les deux caractères.



La relation est **inexistante** entre les deux caractères.

Quantifier une corrélation linéaire

Valeur absolue de r	Intensité de la corrélation
$ r = 1$	Parfaite
$0,75 \leq r \leq 1$	Forte
$0,6 \leq r \leq 0,75$	Moyenne
$0,4 \leq r \leq 0,6$	Faible
$ r \leq 0,4$	Inexistante



COLLECTE DE DONNÉES EN CONTEXTE FONDAMENTAL

Conforme au Programme



MAT_{SN} 4272 2

FORMATION DE BASE DIVERSIFIÉE

NE ME JETEZ PAS !
GARDEZ-MOI
COMME AIDE-MÉMOIRE



Car « *la mémoire est une faculté qui oublie* »
... en maths comme en toutes choses.

CE LIVRE APPARTIENT À : _____

La collection



Des titres
de la collection MAT
au catalogue



FORMATION DE BASE COMMUNE :

Présecondaire

MAT P101 4 MAT P102 3 MAT P103 2 MAT P104 4

Secondaire 1 et 2

MAT 1101 3 MAT 1102 3

MAT 2101 3 MAT 2102 3

Mise À Niveau

MAN P100 MAN 1100 MAN 2100

FORMATION DE BASE DIVERSIFIÉE :

Secondaire 3

MAT 3051 2 MAT 3052 2 MAT 3053 2

Secondaire 4

CST MAT 4151 1 MAT 4152 1 MAT 4153 2

TS MAT 4261 2 MAT 4262 2 MAT 4263 2

SN MAT 4271 2 **MAT 4272 2** MAT 4273 2

Secondaire 5 — *En préparation*

CST *MAT 5150 2* *MAT 5151 1* *MAT 5152 1*

TS *MAT 5160 2* *MAT 5161 2* *MAT 5163 2*

SN *MAT 5170 2* *MAT 5171 2* *MAT 5173 2*

MATHÉMATIQUES :

Secondaire 5

MAT 5101 1 MAT 5102 1 MAT 5103 1 MAT 5104 1 MAT 5105 1 MAT 5106 1

MAT 5107 2 MAT 5108 2 MAT 5109 1 MAT 5110 1 MAT 5111 2 MAT 5112 1

FORMATION À DISTANCE

Secondaire 1, 2, 3 et 5

Tous les guides d'apprentissage du secondaire 1, 2, 3 et 5 ont été adaptés pour les besoins de la formation à distance. Pour en savoir plus: voyez notre site www.ebbp.ca

**Florence Grandchamp
Drita Neziri
Abdelkader Amara
Raymond Thériault**

COLLECTE DE DONNÉES EN CONTEXTE FONDAMENTAL

MAT
A_{SN}
4272 2

FORMATION DE BASE DIVERSIFIÉE





L'ensemble des titres admissibles de notre production bénéficie du soutien financier du gouvernement du Canada.

Communication et pédagogie	Christiane Beullac
Composition et index	Audrey d'Amboise Francisca Martinez Galvez Valérie Tardif
Conseiller en mathématiques	Raymond Thériault
Correction	Jonathan Crête
Direction de la collection	Célestin de La Grange
• contenu éditorial	Annie Lopez
• contenu mathématique	Florence Grandchamp
• infographie et production	Francine Plante
Idéatrice	Marianne Delaroche
Illustrations	Paul Bordeleau
Informatique éditoriale	Francisca Martinez Galvez
Maquette de la couverture	Jean-Sébastien Lajeunesse Michel Lajeunesse
Maquette de l'ouvrage	Célestin de La Grange Francine Plante
Réécriture	Jonathan Crête
Révision mathématique	Sylvain Gervais

À propos de photocopie

Photocopier sans permission un imprimé — une œuvre complète ou un passage d'une œuvre —, c'est aussi plagier. C'est aussi s'approprier indûment le fruit du travail d'un auteur.

Et, la plupart du temps, la photocopie gâte l'œuvre, et fait perdre le bénéfice de cinq cents ans de pratique de l'imprimerie: c'est un péché contre l'esprit, en plus d'être un acte malhonnête.

Photocopier sans permission: c'est voler.

Méprisons la photocopie sauvage. Méprisons le vol.

Droits d'auteur et droits de reproduction
Toutes les demandes de reproduction doivent être acheminées à:
Copibec (reproduction papier) 514 288-1664 1 800 717-2022
licences@copibec.qc.ca

© Œuvre protégée par le droit d'auteur.
Toute reproduction interdite sans autorisation de l'éditeur.

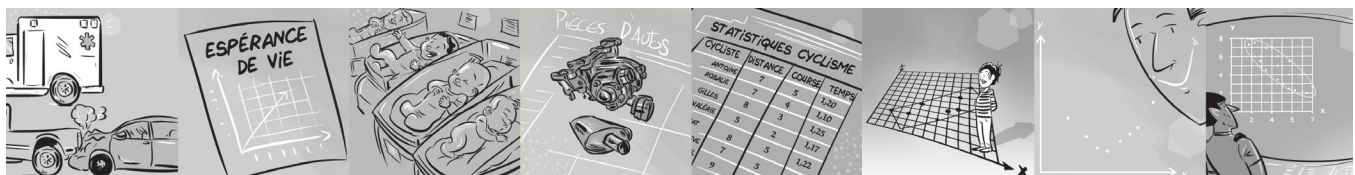
Page des crédits



Impression Imprimerie Héon & Nadeau

Éditrice déléguée Francine Plante / Les Éditions Jules Châtelain

Pour en savoir plus sur l'illustrateur et sur les illustrations de votre module, voir p. 263



À L'ÉTUDIANT ET À L'ENSEIGNANT POUR CETTE PREMIÈRE ÉDITION 2019

Vous avez en main la première édition du module MAT 4272, onzième module de notre collection MAT FORMATION DE BASE DIVERSIFIÉE.

Les auteurs, les correcteurs, les réviseurs et toute l'équipe éditoriale et technique ont fait de leur mieux pour que cet ouvrage respecte l'esprit et la lettre du programme, et réponde à vos attentes et à vos besoins. Mais nul, ni rien, n'est parfait sur terre: moins que quiconque, nous prétendons avoir atteint la perfection, même après révision et correction.

Les auteurs et l'éditeur demandent aux utilisateurs – étudiants et enseignants – de leur faire part de leurs commentaires et de leurs suggestions le plus tôt possible pour que nous puissions dès la prochaine impression apporter les retouches, les modifications ou les ajouts qui se révéleraient nécessaires.

D'autre part, n'hésitez pas à nous signaler coquilles ou erreurs si vous en trouvez: **nous ne procédons jamais à une réimpression sans avoir d'abord effectué les corrections ou les retouches nécessaires.** Un ouvrage didactique n'est pas une œuvre immuable, au contraire, c'est un outil perfectible et en perpétuel devenir.

Avec la collaboration de toutes et de tous, nous pourrons ensemble améliorer et raffiner, au fil des ans, un document dont nous voudrions qu'il soit pour vous l'outil rêvé. Nous ferons tout pour qu'il le devienne.

Écrivez-nous, téléphonez-nous, ou adressez-nous un courriel à l'adresse **cbeullac@ebbp.ca**, la responsable des communications et notre responsable des médias sociaux. Nous accusons toujours réception de la correspondance reçue des utilisateurs. Vous pouvez aussi nous visiter sur le site www.ebbp.ca.

N'hésitez surtout pas!



Depuis plus de soixante-cinq ans, nous n'avons jamais cessé de travailler en étroite collaboration avec le monde de l'enseignement, et nous voulons continuer de le faire: que vous soyez étudiant ou enseignant, merci de garder le contact avec nous par le moyen qui vous est le plus commode: téléphone, télécopieur, courriel.

L'éditeur

KINÉSIS ÉDUCATION
Bureau 275, 4823, rue Sherbrooke Ouest, Westmount, Québec H3Z 1G7
Téléphone: 514 932-9466 Télécopieur: 514 932-5929
Courriel: cbeullac@ebbp.ca Site: www.ebbp.ca



Graphismes, notations et symboles	page 3 de couverture
Qualifier une corrélation linéaire	page 3 de couverture
Quantifier une corrélation linéaire	V
À l'étudiant et à l'enseignant	VIII
Présentation	X
Comment est construit votre MAT 4272	XII
Attentes de fin de cours	

01. DISTRIBUTIONS STATISTIQUES À DEUX CARACTÈRES

Mise en situation:	
LE MAGASIN DE PIÈCES DE VOITURES	2
1.1. Construction et interprétation de tableaux de distributions à deux caractères	4
1.2. Représentation graphique à l'aide d'un nuage de points	21
1.3. Approximation et interprétation du coefficient de corrélation	37
Pause technologie: Utilisation de la technologie pour calculer le coefficient de corrélation linéaire	56
1.4. Représentation de la droite de régression à l'aide d'un graphique	65
1.5. Interpolation ou extrapolation à l'aide de la droite de régression	82
1.6. Représentation de la droite de régression à l'aide de la méthode de la droite médiane-médiane	97
1.7. Représentation de la droite de régression à l'aide de la méthode de Mayer	108
1.8. Interpolation et extrapolation à l'aide du modèle fonctionnel le mieux ajusté à la situation-problème	121
1.9. Vue d'ensemble: synthèse des savoirs	130
Consolidation des savoirs	135
1.10. Situations de vie	148
Situations d'évaluation de fin de chapitre SÉ	163
Évaluation des connaissances	164
Évaluation des compétences	167
Prêt pour l'évaluation de fin de module ?	173
Révision des connaissances	173
Révision des compétences	180
Glossaire des termes mathématiques	202
Corrigé	206
Index	261
À propos de l'illustrateur et des illustrations...	263

Nos petits plus...

Pause technologie	56
-------------------	-----------

COLLECTE DE DONNÉES EN CONTEXTE FONDAMENTAL

Le module MAT 4272, intitulé **Collecte de données en contexte fondamental**, aborde plusieurs aspects d'une grande famille de situations d'apprentissage : *Traitement de données*. Cette famille regroupe les situations qui comportent un problème pouvant être traité en partie par la collecte ou le traitement de données, dans une perspective fondamentale. Le module **Collecte de données en contexte fondamental** vous fournira l'occasion de poser des actions qui visent à vous rendre apte à effectuer ou à comparer des collectes de données. En traitant les situations-problèmes de ce cours, vous serez amené, entre autres, à décoder le sens des symboles, des termes et des notations utilisés, à interpréter correctement l'intensité et le signe du coefficient de corrélation.

COMPÉTENCES DISCIPLINAIRES

Pour résoudre les situations-problèmes de ce cours, vous aurez recours aux trois compétences disciplinaires, soit :

- Utiliser des stratégies de résolution de situations-problèmes ;
- Déployer un raisonnement mathématique ;
- Communiquer à l'aide du langage mathématique.

COMPÉTENCES TRANSVERSALES

Plusieurs compétences transversales peuvent contribuer au traitement de situations de la famille *Traitement de données*. Le programme d'études en propose deux qui apparaissent les plus appropriées pour ce cours :

Compétence d'ordre méthodologique : *Se donner des méthodes de travail efficaces ;*

Compétence d'ordre intellectuel : *Exercer son jugement critique.*

CONTENU DISCIPLINAIRE

Dans ce cours, vous réactiveriez et approfondirez l'ensemble des savoirs propres à la statistique, acquis précédemment. Afin de traiter efficacement les situations-problèmes, vous complétez votre formation en vous appropriant les savoirs propres à ce cours.

Savoirs prescrits

En vue de traiter efficacement les situations proposées dans ce cours, vous développerez trois **procédés intégrateurs** énoncés comme suit :

- La réalisation d'une collecte de données ;
- La comparaison de collectes de données ;
- L'interprétation de données issues d'une expérience.

SAVOIRS MATHÉMATIQUES**Distribution à deux caractères**

SM-1 Construction et interprétation de tableaux de distribution à deux caractères

SM-2 Représentation graphique à l'aide d'un nuage de points

Tous les savoirs
mathématiques : SM.
On le reconnaît
à ce picto associé
aux Outils mathématiques.

représentation et détermination de l'équation de la droite de régression

interpolation ou extrapolation à l'aide de la droite de régression

interprétation qualitative et quantitative d'une corrélation

interpolation et extrapolation à l'aide du modèle fonctionnel

mieux ajusté à la situation-problème



Présentation des *compétences disciplinaires*, des *compétences transversales*, et du contenu disciplinaire visés par le MAT 4272. ➔ page VIII

Les deux pages

Comment est construit votre module.
Vous retrouverez des pages +détaillées un peu +loin à cet extrait.



01

DISTRIBUTIONS STATISTIQUES À DEUX CARACTÈRES

En début de module une *mise en situation*, ici: **LE MAGASIN DE PIÈCES DE VOITURES**.

Elle est tirée de la vie courante réelle ou virtuelle, et illustre l'utilité de la matière qui sera abordée.

DANS CE MODULE, vous dit ce que vous verrez comme nouvelles notions, à quoi cela sert en mathématique et dans la vie de tous les jours. ➔ page 2



1.1. Construction et interprétation de tableaux de distributions à deux caractères

Au début de chaque section: les

Outils mathématiques nécessaires à l'acquisition des *savoirs mathématiques*. Présentation succincte, niveau de langue simple, exemples concrets, illustrations au besoin.

➔ page 4 et suivantes



1.9. Vue d'ensemble: synthèse des savoirs

Un résumé des *savoirs mathématiques* est présenté sous forme de tableau. Il est suivi de *consolidations des savoirs* pour vous aider à maîtriser les nouveaux *savoirs mathématiques*.

➔ page 130 et suivantes

En conclusion du chapitre, des

1.10. Situations de vie

font un *retour sur la mise en situation du début*, laquelle peut maintenant être résolue grâce aux savoirs et compétences acquis dans ce chapitre.

➔ page 148



MAT 4272

PRÊT POUR L'ÉVALUATION DE FIN DE MODULE ?

PREMIÈRE PARTIE Révision des connaissances

Banque de questions portant chacune sur l'un des *savoirs mathématiques* du module.

DEUXIÈME PARTIE Révision des compétences

Banque de *situations-problèmes* permettant de vérifier l'acquisition de toutes les compétences liées à ce module.

➔ page 173

MAT 4272 GLOSSAIRE DES TERMES MATHÉMATIQUES

Un mini-dictionnaire: tous les termes apparaissant en **italique rouge gras** dans le module. ➔ page 202



Et des petits plus....



Pause technologie

Les amateurs de technologie apprendront à utiliser certains logiciels ou les fonctions particulières de certains modèles de calculatrice.

➔ page 56

Pour savoir où vous allez: la liste des *critères d'évaluation* de ce cours.

➔ page XII

Si on appliquait cette théorie?

Ensuite, des cas concrets en relation avec les *savoirs mathématiques* que vous avez découverts dans les **Outils mathématiques**.

➔ page 8 et suivantes

Activités d'apprentissage

Puis, de la pratique, pour vous aider à acquérir par étapes la ou les *compétences disciplinaires* à atteindre. Vous pouvez facilement repérer ces *activités d'apprentissage* grâce à la bande gris pâle sur la tranche du module.

➔ page 13 et suivantes

UN PEU DE PRATIQUE

Situations-problèmes

Viennent ensuite des situations plus globales et plus complexes, les *situations-problèmes* qui vous amèneront à maîtriser les *compétences transversales* visées par le MAT 4272.

Ces situations se repèrent grâce à la bande gris foncé sur la tranche du module.

➔ page 153 et suivantes

UN PEU PLUS DE PRATIQUE

Situations d'évaluation de fin de chapitre

PREMIÈRE PARTIE Évaluation des connaissances

DEUXIÈME PARTIE Évaluation des compétences

Ces *SÉ* se trouvent vers la fin du module. Elles sont signalées par une bande rouge à rayures blanches sur la tranche. Elles sont en deux parties: la première vous permet de vérifier l'acquisition des connaissances, ou *savoirs mathématiques*; la seconde, l'acquisition des *compétences dites transversales*. ➔ page 163 et suivantes

Corrigé

Il vous donne les solutions de toutes les *activités d'apprentissage*, des *situations-problèmes* et des *consolidations des savoirs*.

Ce corrigé se repère grâce à la bande rouge sur la tranche du module.

➔ page 206 et suivantes

MAT 4272

INDEX

Une table alphabétique des mots-clés et leurs références. ➔ page 261

En tiré à part pour l'enseignant

- Corrigé des **SÉ de fin de chapitre**
- Corrigé du **Prêt pour l'évaluation de fin de module?**
- Grilles d'évaluation

Au terme de ce cours, vous serez en mesure d'effectuer et de comparer à un ou deux caractères en vue de répondre à un questionnement lié à un problème que vous aurez vous-même cerné. Vous pourrez aussi comparer d'autres collectes de données de même type pour résoudre un problème que vous aurez vous-même cerné. La présentation des résultats de votre analyse sera faite dans le respect des règles et des conventions mathématiques. Des stratégies de résolution de situations-problèmes seront mises à profit pour déterminer la solution la plus efficiente. De plus, vous serez à même d'éprouver votre façon de traiter une situation à partir d'une analyse statistique en menant une expérience à l'aide d'un support technologique.

CRITÈRES D'ÉVALUATION

- Utiliser des stratégies de résolution de situations-problèmes
- Déployer un raisonnement mathématique
- Communiquer à l'aide du langage mathématique*

1. UTILISER DES STRATÉGIES DE RÉOLUTION DE SITUATIONS-PROBLÈMES

- 1.1 Manifestation, oralement ou par écrit, d'une compréhension adéquate de la situation-problème
- 1.2 Mobilisation de stratégies et de savoirs mathématiques appropriés à la situation-problème

2. DÉPLOYER UN RAISONNEMENT MATHÉMATIQUE

- 2.1 Utilisation correcte des concepts et des processus mathématiques appropriés
- 2.2 Mise en œuvre convenable d'un raisonnement mathématique adapté à la situation
- 2.3 Structuration adéquate des étapes d'une démarche pertinente

* La compétence 3 « Communiquer à l'aide du langage mathématique » ne fait pas l'objet d'une évaluation spécifique au regard de la sanction et de la reconnaissance. Toutefois, puisqu'elle se manifeste nécessairement dans toute activité mathématique, elle a été prise en compte dans les outils d'évaluation élaborés pour aider les enseignants à porter leur jugement.

COLLECTE DE DONNÉES EN CONTEXTE FONDAMENTAL

Votre MAT 4272
est présenté en 1 chapitre
dont voici le titre:

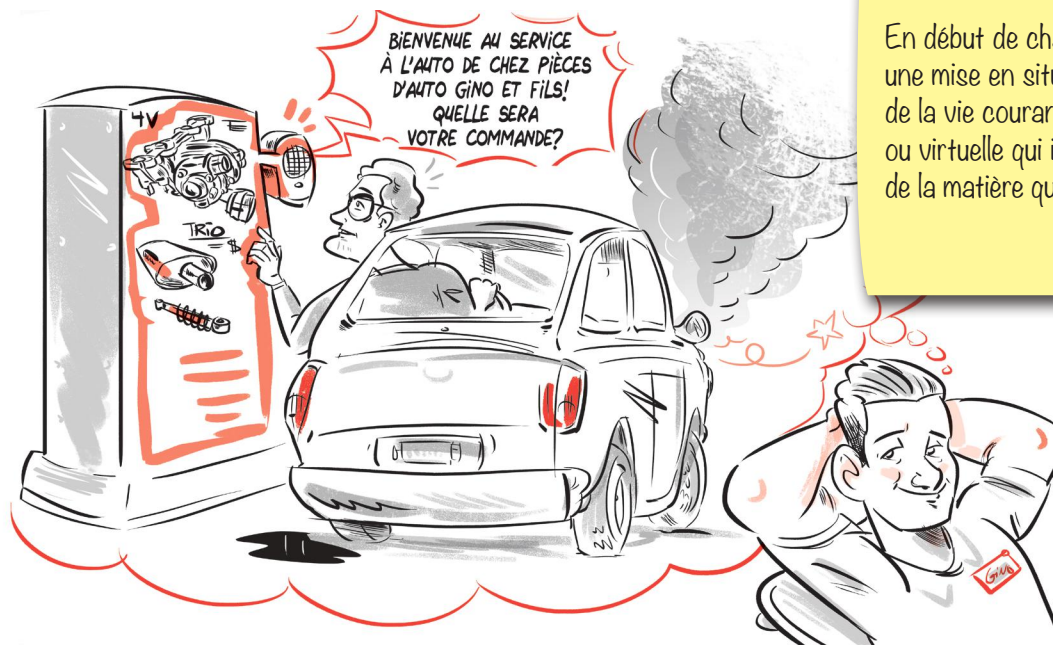


01. DISTRIBUTIONS STATISTIQUES À DEUX CARACTÈRES

Dans ce chapitre, vous étudierez les distributions à deux caractères, très utilisées dans les analyses sociologiques, économiques, politiques, etc. Vous apprendrez à construire et à interpréter un tableau de distribution à deux caractères, à représenter la droite de régression à l'aide d'un graphique, à l'aide de la méthode de la droite médiane-médiane, ou à l'aide de la méthode de Mayer. Vous aborderez aussi le coefficient de corrélation ainsi que l'interprétation quantitative et qualitative d'une corrélation.

Mise en situation:

LE MAGASIN DE PIÈCES DE VOITURES



En début de chapitre, une mise en situation tirée de la vie courante réelle ou virtuelle qui illustre l'utilité de la matière qui sera abordée.

KINÉSIS
EDUCATION

Votre famille possède un magasin de pièces de voitures depuis 60 ans. Votre grand-père a légué son commerce à votre père, et il est prévu que vous repreniez le flambeau au moment où vos parents prendront leur retraite. Pour le moment, vous faites de menus travaux dans l'entreprise: inventaire, comptabilité, caisse, ménage, etc.

Votre grand-père a toujours fait la publicité de son magasin par la poste, mais vos parents pensent que cette méthode a une portée limitée. Selon eux, les gens ne sont pas prêts à se déplacer plus de 50 km pour acheter des pièces d'automobiles dans leur magasin.

Histoire de vérifier si l'hypothèse de vos parents est valide, vous profiterez de vos moments à la caisse pour demander aux clients leur adresse.

De plus, vous formulez votre propre hypothèse : plus le client vient de loin, plus il achète d'articles afin de rentabiliser son déplacement. Voici un tableau de quelques données que vous avez relevées au cours de cette période.

DONNÉES CONCERNANT LES CLIENTS SÉLECTIONNÉS

Numéro du client	Distance du magasin en kilomètres	Nombre de pièces achetées	Type de client
1	65	8	nouveau
2	46	4	mécanicien
3	31	7	habituel
4	75	11	nouveau
5	3	2	mécanicien
6	12	1	mécanicien
7	22	6	nouveau
8	102	13	nouveau
9	16	2	mécanicien
10	98	9	habituel
11	47	5	habituel
12	77	7	nouveau
13	22	3	mécanicien
14	33	6	nouveau

Les données que vous avez consignées vous permettront-elles de confirmer votre hypothèse ?

Vous découvrirez, dans ce module, tout ce qu'il faut savoir sur la corrélation des variables statistiques et pourrez ainsi confirmer ou infirmer votre hypothèse.

Le bloc *Dans ce chapitre* vous indique les nouvelles notions que vous apprendrez et quelles seront leurs utilités en mathématiques et dans la vie de tous les jours.



DANS CE MODULE

Quoi de nouveau ?

- La corrélation entre deux variables statistiques

Qu'est-ce que c'est ?

- La corrélation statistique est une méthode qui permet de déterminer si deux variables sont liées.

À quoi ça sert en mathématiques ?

- Étudier la corrélation entre deux variables permet de déterminer la relation qui les unit.

À quoi ça servira dans la vie ?

- La corrélation entre deux variables permet d'extrapoler la valeur d'une variable à partir d'une valeur donnée de l'autre variable.

1.1. Construction et interprétation de tableaux de distributions à deux caractères

Le chapitre est divisé en sections.



- DANS CETTE SECTION, VOUS APPRENDREZ À CONSTRUIRE UN TABLEAU DE DISTRIBUTION À DEUX CARACTÈRES ET À INTERPRÉTER LES DONNÉES D'UN TEL TABLEAU.



SM-1

Les outils mathématiques nécessaires à l'acquisition des savoirs mathématiques: **SM**.



Outils mathématiques

Construction d'un tableau de distribution à deux caractères – Interprétation d'un tableau de distribution à deux caractères

1. Construction d'un tableau de distribution à deux caractères

Un **tableau de distribution à deux caractères** appelé aussi **tableau à double entrée**, est un mode de

de **tableau de distribution à deux caractères** pour organiser les données d'une distribution à deux caractères. Les caractères étudiés apparaissent dans les colonnes et les valeurs de l'autre caractère dans les lignes. Dans un tableau de distribution à deux caractères, la lecture se fait par lignes et par colonnes.

Le tableau à deux caractères:

		Colonnes		
Caractère 1				
Caractère 2		C1	C2	C3
Lignes	L1			
	L2			
	L3			

Pour construire un tableau de distribution à deux caractères, on suit les étapes suivantes.

Étape 1:

On détermine le **nombre de lignes et de colonnes**. Pour ce faire, on identifie les différents types d'informations à classer ainsi que les éléments que l'on veut comparer et les critères de comparaison.

Étape 2:

On inscrit les **titres des lignes et des colonnes** ainsi que la **source en bas du tableau**.

Les éléments à comparer sont placés en **titre au début de chaque colonne** tandis que les critères de comparaison sont placés en titre **au début de chaque ligne**.

Étape 3:

On complète les cases du tableau. Pour ce faire, on choisit dans les données celles qui correspondent au titre de chaque colonne ou de chaque ligne.

Étape 4:

On donne un titre au tableau. « Tableau représentant »





Outils mathématiques suite

Exemple

La distribution suivante représente les ventes effectuées par chacun d'un concessionnaire *Nissan* durant une semaine selon le modèle de véhicule. Pierre a vendu 1 *Murano* et 2 *Rogue*; Jean a vendu 2 *Murano* et 2 *Altima*; Sami, quant à lui, vendu 1 *Murano* et 3 *Altima*.

Construire un tableau représentant cette distribution.

Notre tableau va comporter cinq colonnes et cinq lignes. La première colonne représente le nom des vendeurs avec comme caractère **nom du vendeur**. La première ligne représente les modèles de véhicules avec comme caractère **modèle de véhicule**.

La dernière ligne et colonne représentent les totaux.

Nom du vendeur \ Modèle de véhicule	<i>Murano</i>	<i>Rogue</i>	<i>Altima</i>	Total
Pierre	1			
Jean				
Sami				
Total				

Source: interne

Une fois le tableau dessiné, on complète les cases par le nombre de véhicules vendus par chacun des vendeurs selon le modèle.

Pierre a vendu 1 *Murano*. À l'intersection de la **ligne Pierre** et de la **colonne *Murano***, on inscrit le nombre de voitures vendues, c'est à dire **1**. On complète les autres cases en suivant le même raisonnement. On n'oublie pas de remplir les cases représentant les totaux.

Une fois les cases complétées, on assigne le titre suivant au tableau: **Tableau représentant les ventes de véhicules de trois vendeurs selon le modèle.**


On obtient finalement le tableau complété suivant.

TABLEAU REPRÉSENTANT LES VENTES DE VÉHICULES DE TROIS VENDEURS SELON LE MODÈLE

Nom du vendeur \ Modèle de véhicule	<i>Murano</i>	<i>Rogue</i>	<i>Altima</i>	Total
Pierre	1	2	0	3
Jean	2	0	2	4
Sami	1	0	3	4
Total	4	2	5	11

Source: interne

Cet outil comprend des exemples, des démarches détaillées et leurs résolutions.





Outils mathématiques suite

2. Interprétation d'un tableau de distribution à deux caractères

Un tableau représentant une distribution statistique peut être analysé et interprété pour tirer certaines conclusions.

Voici les points auxquels on doit porter attention pour répondre aux questions posées sur l'interprétation d'un tableau à deux caractères.

Titre du tableau : Le titre du tableau nous renseigne sur l'objet de l'étude.

La source : En général, la source est inscrite en bas du tableau. Elle nous renseigne sur la provenance des informations du tableau.

Le champ : Le champ précise le domaine d'étude : population concernée, pays, périodes, etc.

Les unités utilisées pour exprimer les données (effectifs, pourcentage, monnaie, indice, etc.)

Les titres des lignes et des colonnes : Les définitions des deux caractères, la relation qui peut exister entre les deux caractères, etc.

Analyse des données : On regarde s'il y a un caractère dominant dans le tableau ainsi que l'évolution des caractères. On regarde aussi les totaux des lignes et des colonnes pour tirer des conclusions.

On détermine s'il y a des différences en observant la structure globale des données chiffrées pour dégager des tendances, par exemple la concentration des données selon une diagonale.

L'exemple suivant vous permettra de comprendre comment interpréter un tableau de distribution à deux caractères.

Exemple

POPULATION FÉMININE, EN POURCENTAGE, ÂGÉE DE 15 ANS ET PLUS, VIVANT DANS LES MÉNAGES PRIVÉS SELON LE PLUS HAUT DIPLÔME OBTENU, ET LE GROUPE D'ÂGE, DE L'ENSEMBLE DU QUÉBEC, 2011

		Taux de niveau d'études %					
Âge \ Diplôme	Diplôme	Aucun diplôme	DES	DEP	DEC	Certificat	Diplôme universitaire
	15-19 ans		48,8	39,1	2,8	8,8	0,4
20-24 ans		11,0	24,4	11,6	35,5	4,3	15,2
25-34 ans		8,1	13,4	15,2	22,8	5,2	35,3
35-44 ans		9,8	14,2	16,3	22,4	6,1	31,2
45-54 ans		15,1	23,7	16,5	19,9	5,9	18,9
55-64 ans		20,7	28,9	12,7	15,9	6,3	15,4
65 et plus		45,0	24,5	7,0	9,3	6,1	8,1
15 ans et plus		21,9	22,8	12,5	18,3	5,4	19,1
Immigrantes		22,0	18,9	9,1	14,1	7,2	28,8

Source : Statistique Canada. Recensement de 2011
Compilation : Ministère de l'Éducation, du Loisir et du Sport



Outils mathématiques *suite*

Interprétation du tableau :

Titre : Le titre nous renseigne sur le **sujet traité**, c'est-à-dire les **diplômes obtenus par les femmes** au Québec selon le groupe d'âge.

La source : La source nous indique que les données proviennent du recensement de 2011 par Statistique Canada, compilées par le Ministère de l'Éducation, du Loisir et du Sport.

Le champ : Le tableau nous précise que l'étude est basée sur les données de **l'année 2011**, et concerne les femmes au Québec et leur niveau d'étude (sans diplôme, DES, DEP, DEC, certificat et diplôme universitaire).

Les unités utilisées pour exprimer les données : Les unités utilisées sont le pourcentage (%) et les groupes d'âges.

Les titres des lignes et des colonnes : Les colonnes représentent le type de diplôme obtenu (sans diplôme, DES, DEP, DEC, certificat et diplôme universitaire) et les lignes les groupes d'âges.

On remarque que le pourcentage de diplômées change selon le type de diplôme et le groupe d'âge.

Analyse des données : Voici quelques conclusions qu'on peut tirer de l'observation des données de ce tableau : Les femmes du groupe d'âge 25-34 ans qui ont un diplôme universitaire, représentent le plus haut pourcentage (35,3 %).

On remarque aussi que le pourcentage de femmes immigrantes ayant un diplôme universitaire (28,8 %) est plus grand que le pourcentage de l'ensemble des femmes (15 ans et plus) du Québec ayant le diplôme universitaire (19,1 %).

On remarque aussi que 48,8 % des femmes du groupe d'âge 15-19 ans sont sans diplôme.

Si on appliquait cette théorie?

- LES EXEMPLES CI-APRÈS VOUS PERMETTRONT DE MIEUX COMPRENDRE COMMENT CONSTRUIRE ET INTERPRÉTER UN TABLEAU DE DISTRIBUTION À DEUX CARACTÈRES.

Exemple 1

Claude observe les résultats des Jeux olympiques d'été de Rio de Janeiro sur Wikipédia. Les États-Unis ont obtenu 46 médailles d'or, 37 d'argent et 38 de bronze. La Chine a récolté 26 médailles d'or, 18 d'argent et 26 de bronze. Les athlètes de la Grande-Bretagne ont gagné 27 médailles d'or, 23 d'argent et 17 de bronze. La Russie, quant à elle, a raflé 19 médailles d'or, 18 d'argent et 20 de bronze. L'Allemagne a décroché 17 médailles d'or, 10 d'argent et 15 de bronze.

Des cas concrets en relation avec les savoirs mathématiques. Celui-ci comprend au moins 2 exemples: Le premier est détaillé avec une démarche élaborée.



Présenter ces données à l'aide d'un tableau à double entrée.

Solution

Le tableau comportera 7 lignes et 5 colonnes. Les lignes représenteront les pays et les colonnes la nature des médailles.

Nature des médailles \ Pays	Or	Argent	Bronze	Total
États-Unis				
Grande-Bretagne				
Chine				
Russie				
Allemagne				
Total				

On complète chacune des cases par le nombre de médailles. Par exemple, à l'intersection de la ligne **États-Unis** et de la colonne **Or**, on inscrit le nombre de médailles d'or gagné par les États-Unis, soit **46**. On complète de la même manière les autres cases ainsi que les totaux.

On n'oublie pas d'inscrire la source « Wikipédia » au bas du tableau.

On donne finalement un titre au tableau.

Exemple de titre: **Gain en médailles de cinq pays lors des Jeux olympiques d'été de Rio en 2016**

On obtient le tableau complété suivant :

GAIN EN MÉDAILLES DE CINQ PAYS LORS DES JEUX OLYMPIQUES D'ÉTÉ DE RIO EN 2016

Nature des médailles \ Pays	Or	Argent	Bronze	Total
États-Unis	46	37	38	121
Grande-Bretagne	27	23	17	67
Chine	26	18	26	70
Russie	19	18	20	57
Allemagne	17	10	15	42
Total	135	106	116	357

Source : Wikipédia

Exemple 2

La direction a procédé à un recensement des élèves qui fréquentent la répartition suivante des élèves selon le niveau et le sexe.

Première année du secondaire : 120 élèves dont 45 sont des garçons et 75 filles.

Deuxième année du secondaire : 55 garçons et 40 filles.

Troisième année du secondaire : 80 élèves dont 30 sont des garçons et 50 filles.

Quatrième année du secondaire : 36 garçons et 42 filles.

Cinquième année du secondaire : 76 élèves dont 35 sont des garçons et 41 filles.

Le deuxième exemple : à vous de démontrer votre savoir en effectuant la démarche proposée!



Construire un tableau de distribution à deux caractères représentant cette distribution.

Solution

Le tableau comportera colonnes et lignes.

Les lignes représenteront le sexe des élèves et les colonnes le des élèves.

Complétez le tableau représentant la distribution.

Titre : _____

	1 ^{re} secondaire	2 ^e secondaire	_____	4 ^e secondaire	_____	_____
Masculin	45	55	30	_____	35	201
_____	_____	40	_____	42	_____	_____
Total	_____	_____	_____	_____	_____	_____

Source : _____



Si vous avez bien compris les notions étudiées, vous allez arriver aux réponses suivantes.

Le tableau comportera **7 colonnes** et **4 lignes**.

Les lignes représenteront le sexe des élèves et les colonnes le **niveau** des élèves.

Voici le tableau, une fois complété.

RECENSEMENT DES ÉLÈVES D'UNE POLYVALENTE SELON LE NIVEAU ET LE SEXE

	1 ^{re} secondaire	2 ^e secondaire	3 ^e secondaire	4 ^e secondaire	5 ^e secondaire	Total
Masculin	45	55	30	36	35	201
Féminin	75	40	50	42	41	248
Total	120	95	80	78	76	449

Source: Direction de la polyvalente

Exemple 3

Interpréter le tableau de données ci-après.

CONDUCTEURS IMPLIQUÉS DANS LES ACCIDENTS MORTELS DE 2008 À 2015, SELON L'ÂGE

Année \ Âge	16 à 24 ans	25 à 44 ans	45 à 64 ans	65 ans ou plus
2008	170	270	255	73
2009	138	255	230	76
2010	152	221	210	92
2011	130	240	210	98
2012	126	181	188	82
2013	100	197	209	81
2014	77	163	149	84
2015	74	184	173	106

Troisième exemple:
Une autre interprétation!



Source: Société de l'assurance automobile du Québec

Le tableau nous renseigne sur _____.

Le **titre** nous indique que le tableau porte sur **les accidents mortels selon l'âge des conducteurs**.

La **première ligne** indique _____ tandis que la **première**

colonne indique _____.

La première ligne indique **l'âge des conducteurs** tandis que la première colonne indique **l'année**.

Les **données** sont **fiables**, car elles sont produites par la Société d'assurance automobile du Québec qui est une société d'état.

Analyse des données

Voici quelques exemples de conclusions possibles après l'observation des résultats :

Sauf pour l'année 2010, les accidents mortels commis par les conducteurs de 16 à 24 ans diminuent avec le temps ;

C'est en 2015 qu'on a enregistré le plus grand nombre d'accidents pour les conducteurs de 65 ans et plus ;

Le nombre d'accidents mortels est à la baisse entre 2008 et 2012 pour les conducteurs âgés de 45 ans à 64 ans ;

Etc.

Exemple 4

Répondre aux questions en vous basant sur les données.

REVENU D'EMPLOI MOYEN ET RATIO ENTRE LES REVENUS DES FEMMES ET DES HOMMES CHEZ LES TRAVAILLEURS À TEMPS PLEIN, QUÉBEC, ONTARIO ET AU CANADA EN 2002

	Québec		Ontario	
	Revenu (\$ de 2002)	Ratio F/H en %	Revenu (\$ de 2002)	Ratio F/H en %
Femmes	34 400	75,3	38 800	70,0
Hommes	45 700		55 400	

Quatrième exemple:
Encore + de pratique!



Source: Statistique Canada, Enquête sur les finances des consommateurs; Enquête sur la dynamique du travail et du revenu; céderom Tendances du revenu au Canada, tableau 202-0102

Le tableau nous renseigne sur quoi ?

Le titre indique que le tableau porte sur **le revenu moyen des femmes et des hommes qui travaillent à temps plein dans deux provinces et au Canada en 2002.**

Quel organisme a émis les données ? _____

La source qui a émis les données est: **Statistique Canada.**

Au Québec, les femmes sont-elles mieux payées que les hommes ?

Non, les femmes ne sont pas mieux payées que les hommes.



De combien le revenu des hommes est supérieur à celui des femmes au Québec ?

Le revenu des hommes dépasse celui des femmes de **11 300 \$**.

Les hommes du Québec sont-ils mieux payés que leurs homologues ontariens ?

Non, les hommes de l'Ontario sont mieux payés que leurs homologues québécois. Les travailleurs masculins du Québec touchent **9 700 \$ de moins** que leurs homologues ontariens.

De combien le salaire des hommes de l'Ontario dépasse-t-il le revenu moyen des hommes à l'échelle du pays ?

Le revenu moyen des hommes de l'Ontario dépasse celui du pays de **4 900 \$**.

Est-il vrai que les femmes au Québec sont mieux payées que leurs homologues ontariennes par rapport aux hommes ?

Oui, car au **Québec**, le revenu d'emploi des femmes équivaut à **75,3 %** de celui des hommes alors qu'en **Ontario**, il équivaut à 70,0 % de celui des hommes.

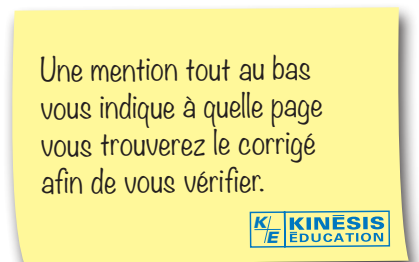
Pour vous permettre de mieux maîtriser la construction et l'interprétation d'un tableau de distribution à deux caractères, nous vous proposons de traiter les **Activités d'apprentissage** qui suivent.

1. Construire le tableau demandé.

- a) Selon le département des ventes, une entreprise de montage de produits électroniques et d'électroménagers a exporté en 2009, 10 500 réfrigérateurs, 31 000 cuisinières au gaz, 2 300 cuisinières électriques, 76 900 téléviseurs et 32 000 robots. En 2010, l'entreprise a exporté 12 200 réfrigérateurs, 26 500 cuisinières au gaz, 3 200 cuisinières électriques, 78 100 téléviseurs et 29 800 robots.



Construire le tableau de distribution à deux caractères représentant cette situation.



1.9. Vue d'ensemble: synthèse des savoirs

Nous arrivons à la fin du chapitre traitant des distributions statistiques à deux caractères. Avant de passer aux **Situations-problèmes** plus globales qui vont conclure ce chapitre, voici un résumé des *savoirs mathématiques* que vous avez appris jusqu'ici.

Résumé des savoirs mathématiques

Tableaux de distributions statistiques à deux caractères

Un **tableau de distribution à deux caractères** appelé aussi **tableau à double entrée** est un tableau à deux lignes et des colonnes. Les deux caractères étudiés apparaissent dans la première case du tableau. Dans un tableau de distribution à deux caractères, les données sont classées par croisement des lignes et des colonnes.

Nuage de points

On utilise un plan cartésien pour représenter graphiquement des données **quantitatives** d'une distribution statistique à deux caractères par un **nuage de points**.

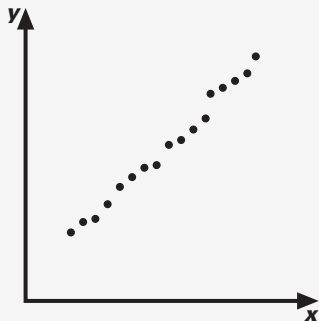
Chacun des couples de la distribution est représenté par un point dans le plan cartésien. L'ensemble des points obtenus dans le plan cartésien constitue le **nuage de points**.

Comment qualifier la relation entre deux caractères d'une distribution à l'aide d'un nuage de points

En observant l'allure du nuage de points, on peut qualifier le degré (force) du lien qui existe entre les deux caractères.

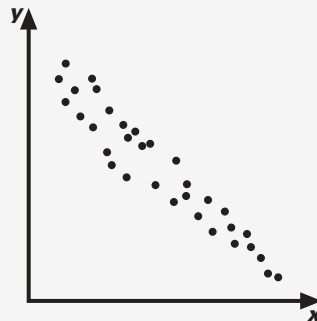
Les modèles suivants vous aideront à qualifier le degré du lien existant entre les deux caractères d'une distribution représentée par un nuage de points.

Modèle 1



La relation est **parfaite** entre les deux caractères.

Modèle 2



La relation est **forte** entre les deux caractères.

Un résumé des savoirs mathématiques de ce chapitre vous est présenté.



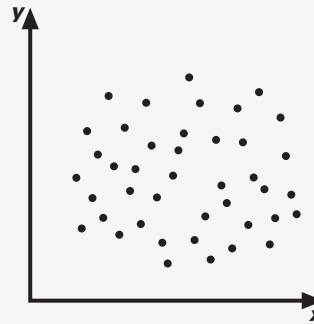
Résumé des savoirs mathématiques *suite*

Modèle 3



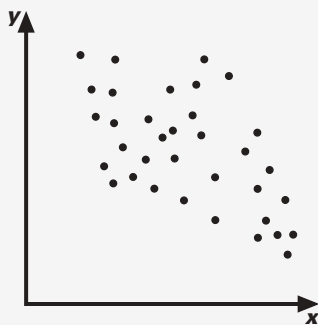
La relation est **moyenne** entre les deux caractères.

Modèle 5



La relation est **inexistante** entre les deux caractères.

Modèle 4

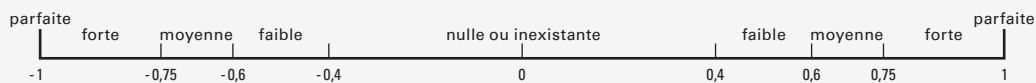


La relation est **faible** entre les deux caractères.

Approximation et interprétation du coefficient de corrélation

Pour **quantifier l'intensité** de la corrélation linéaire associée à un nuage de points, on utilise le **coefficient de corrélation linéaire** noté r qui est un nombre compris entre -1 et 1.

On utilise la ligne suivante pour qualifier une corrélation linéaire selon la valeur du coefficient r .



Estimation du coefficient de corrélation linéaire

On peut estimer la valeur du coefficient de corrélation linéaire par deux méthodes : la méthode du rectangle et la méthode de l'ellipse.

Méthode du rectangle

La **méthode du rectangle** permet d'**estimer** la valeur du coefficient de corrélation directement à partir du nuage de points. Même si cette méthode n'est pas précise, elle donne tout de même une bonne idée de **l'intensité de la corrélation**.

Résumé des savoirs mathématiques suite

La méthode à suivre pour **estimer le coefficient de corrélation linéaire** d'une distribution statistique représentée par un nuage de points est la suivante :

On **trace**, autour de l'ensemble des points, un rectangle le plus proche possible du nuage de points en respectant l'orientation générale des points.

On **mesure**, à l'aide de la règle, la longueur et la largeur du rectangle.

On **calcule** une approximation du coefficient de corrélation linéaire à l'aide de la formule suivante :

$$r \approx \pm \left(1 - \frac{\text{mesure du petit côté}}{\text{mesure du grand côté}} \right)$$

On **détermine** le signe du coefficient de corrélation selon le sens du nuage de points.

Méthode de l'ellipse

La **méthode de l'ellipse** permet d'**estimer** la valeur du coefficient de corrélation d'une distribution statistique à deux variables. La méthode consiste à entourer le nuage de points d'une ellipse se rapprochant le plus possible de l'ensemble des points en respectant l'orientation générale des points.

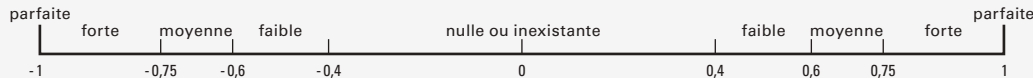
On calcule le coefficient de corrélation linéaire par la formule suivante :

$$r \approx \pm \left(1 - \frac{\text{mesure du petit axe}}{\text{mesure du grand axe}} \right)$$

Interprétation quantitative d'une corrélation linéaire

La valeur du coefficient de corrélation linéaire r comprise entre -1 et 1 permet de **quantifier l'intensité** de la corrélation linéaire qui existe entre deux variables.

Lorsque le coefficient de corrélation linéaire est calculé, reportez-vous à la droite suivante pour interpréter le résultat.



Plus la valeur du coefficient se **rapproche de zéro**, plus la corrélation est **faible**. À l'inverse, plus la valeur du coefficient se **rapproche de -1** ou **de 1**, plus la corrélation est **forte**. Notez que le **signe** du coefficient indique le **sens**, positif ou négatif, du nuage de points selon son orientation.

On peut aussi représenter les différents degrés d'intensité de corrélation linéaire dans un tableau en ne considérant que la valeur absolue du coefficient de corrélation.

Valeur absolue de r	Intensité de la corrélation
$ r = 1$	Parfaite
$0,75 \leq r \leq 1$	Forte
$0,6 \leq r \leq 0,75$	Moyenne
$0,4 \leq r \leq 0,6$	Faible
$ r \leq 0,4$	Inexistante

Résumé des savoirs mathématiques *suite*

Représentation de la droite de régression à l'aide d'un graphique et détermination de son équation

La **droite de régression** est une droite qui représente le mieux possible le nuage de points et qui le partage en deux parties égales, c'est-à-dire, le nombre de points de part et d'autre de la droite est à peu près le même.

Tracer la droite de régression à partir de deux points du nuage de points

On prend un point du nuage de points et on choisit un deuxième point de sorte que la droite qui relie ces deux points respecte l'orientation générale du nuage de points et partage les points du nuage en deux parties égales ou presque.

En joignant les deux points, on a la droite de régression qui représente approximativement le nuage de points.

Tracer la droite de régression à partir de la méthode du point des moyennes des coordonnées et d'un autre point

La méthode consiste à déterminer les coordonnées d'un point à l'aide de la moyenne des abscisses et de la moyenne des ordonnées. Ensuite, on choisit un deuxième point de sorte que la droite passant par ces deux points respecte l'orientation générale du nuage de points et le partage le plus également possible.

En joignant les deux points, on a la droite qui représente approximativement le nuage de points.

Détermination de l'équation de la droite de régression

Une droite peut être représentée par une équation de la forme $y = ax + b$ où a représente le taux de variation (pente) de la droite et b son ordonnée à l'origine. Soient $A(x_1, y_1)$ et $B(x_2, y_2)$ les deux points qui permettent de tracer la droite de régression.

On utilise la formule suivante pour calculer la pente de la droite de régression : $a = \frac{y_2 - y_1}{x_2 - x_1}$.

En substituant la valeur de la pente calculée précédemment et les coordonnées d'un des deux points dans l'équation de la droite de régression $y = ax + b$, on détermine la valeur de l'ordonnée à l'origine b .

En substituant la valeur de la pente a et la valeur de l'ordonnée à l'origine b dans l'équation $y = ax + b$; on obtient l'équation de la droite de régression.

Représentation de la droite de régression à l'aide de la méthode de la droite médiane-médiane

La médiane joue un rôle central pour déterminer l'équation de la droite de régression.

La démarche à suivre pour déterminer l'équation de la droite de régression est la suivante.

On **ordonne** selon un ordre croissant les coordonnées de la distribution selon la variable **indépendante x** .

On **sépare** la distribution en **trois groupes égaux** si le nombre de données est divisible par 3.

Si ce n'est pas possible, on s'assure que le premier et le troisième groupe comportent le même nombre de données. La différence entre le nombre de données dans chacun des groupes doit être la moindre possible.



Résumé des savoirs mathématiques *suite*

On **détermine** le **point médian** de chacun des groupes en calculant la médiane en x et la médiane en y de chacun qu'on désigne par M_1 , M_2 , M_3 .

On **détermine** les coordonnées du **point moyen P** en calculant la **moyenne** en x et en y des coordonnées des points médians M_1 , M_2 , M_3 .

On **calcule** la **pente** de la droite à l'aide des coordonnées des points M_1 et M_3 ;

On **détermine** l'équation de la droite de régression en utilisant la valeur de la pente et les coordonnées du point **P**.

Représentation de la droite de régression à l'aide de la méthode de Mayer

La moyenne joue un rôle central pour déterminer l'équation de la droite de régression.

La démarche à suivre pour déterminer l'équation de la droite de régression est la suivante.

On **ordonne** selon un ordre croissant les coordonnées de la distribution selon la variable **indépendante x** .

On **sépare** la distribution en **deux groupes** égaux G_1 et G_2 . Si le nombre de données est impair, un groupe aura une donnée de plus.

On **calcule la moyenne en x et en y** du groupe G_1 et G_2 . Les coordonnées des points sont: $G_1 (\bar{x}_1, \bar{y}_1)$; $G_2 (\bar{x}_2, \bar{y}_2)$.

La droite qui joint les deux points est la droite de régression dont l'équation est $y = ax + b$.

On **détermine** la pente de la droite par la formule: $a = \frac{y_2 - y_1}{x_2 - x_1}$.

En substituant la valeur calculée de la pente et les coordonnées d'un des deux points, dans l'équation $y = ax + b$, on détermine l'ordonnée à l'origine b .

Interpolation ou extrapolation à l'aide de la droite de régression

L'**interpolation** consiste à déterminer la valeur d'une variable à partir d'une valeur connue de l'autre variable à l'**intérieur** du nuage de points.

L'**extrapolation** consiste à déterminer la valeur d'une variable à partir d'une valeur connue de l'autre variable à l'**extérieur** du nuage de points.

Extrapolation à l'aide d'une fonction quadratique

Lorsque la parabole est la meilleure approximation de la forme d'un nuage de points que la droite, on peut déterminer les paramètres a , b et c de la règle de la fonction quadratique $y = ax^2 + bx + c$, à l'aide d'une calculatrice munie des fonctions statistiques. La règle nous permet d'extrapoler des valeurs.

Extrapolation à l'aide d'une fonction en escalier

Lorsque la courbe d'une fonction en escalier est la meilleure approximation de la forme d'un nuage de points, on utilise la représentation graphique du nuage de points pour faire la meilleure approximation possible des valeurs extrapolées.

Consolidation des savoirs

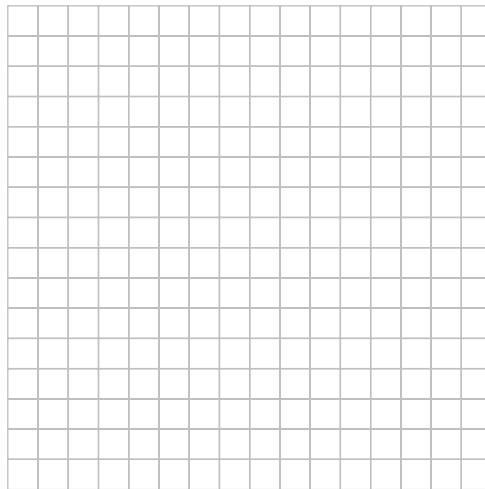
1. Qualifier le lien qui existe entre les deux variables de chaque tableau ci-dessous et dire si la corrélation est positive ou négative.

Des consolidations des savoirs vous sont offertes afin de mieux les maîtriser.



a) Avant de passer l'examen final d'un cours, les étudiants s'entraîment. La distribution suivante montre la note obtenue à l'examen final et le nombre de prétests faits par 6 étudiants.

Nombre de prétests	0	1	2	3	5	6
Note finale en %	50	65	75	80	85	90

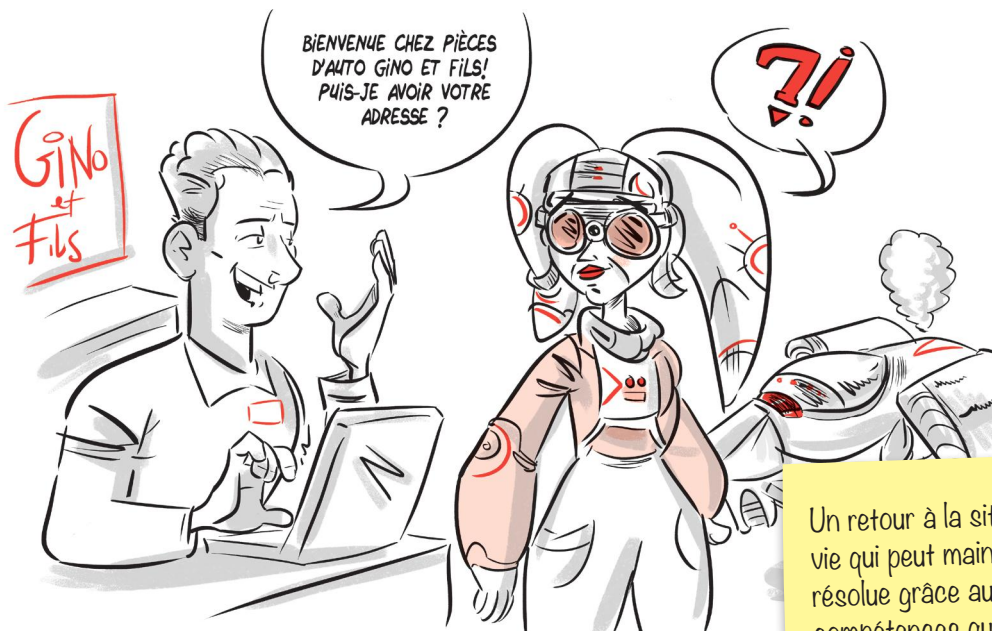


1.10. Situations de vie

Revenons à la situation dont il était question au début de ce chapitre: vous travaillez dans le commerce familial de pièces de voitures.

Retour à la mise en situation:

UNE HYPOTHÈSE À VÉRIFIER...



Un retour à la situation de vie qui peut maintenant être résolue grâce aux savoirs et compétences que vous avez acquis jusqu'à présent.



Selon vos parents, la publicité par la poste a une portée limitée. Les gens vivant à une distance de 50 km ou plus du magasin ne vont pas se déplacer pour y acheter des pièces d'automobiles. D'après vous, plus le client vient de loin, plus il achète d'articles afin de rentabiliser son déplacement.

Voici venu le moment de vérifier si votre hypothèse est défendable.

1. Votre hypothèse est-elle défendable?

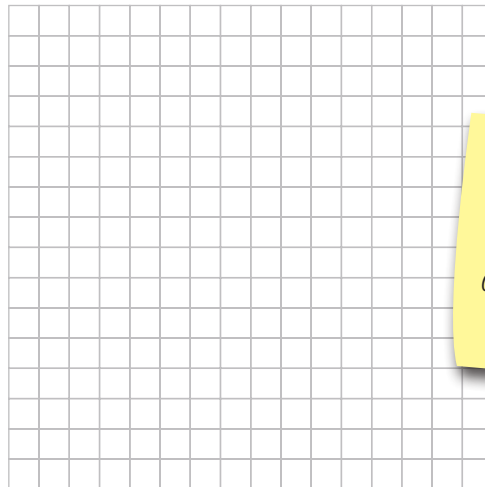
Pendant quelque temps, vous avez demandé aux clients de vous donner leur adresse. Vous avez consigné les données dans le tableau suivant:

DONNÉES CONCERNANT LES CLIENTS SÉLECTIONNÉS

Numéro du client	Distance du magasin en kilomètres	Nombre de pièces achetées	Type de client
1	65	8	nouveau
2	46	4	mécanicien
3	31	7	habituel
4	75	11	nouveau
5	3	2	mécanicien
6	12	1	mécanicien
7	22	6	nouveau
8	102	13	nouveau
9	16	2	mécanicien
10	98	9	habituel
11	47	5	habituel
12	77	7	nouveau
13	22	3	mécanicien
14	33	6	nouveau

1^{re} tâche

Les données que vous avez consignées permettent-elles de confirmer votre hypothèse?



Des éléments graphiques, tel qu'ici une grille vous évitant les feuilles quadrillées volantes.



1. À La Glacière.

Ricardo, l'un de vos amis les plus chers, a ouvert un petit commerce dont la vocation est, vous l'aurez deviné, la vente de friandises g

Ricardo a remarqué que plus la température extérieure augmente au comptoir de *La Glacière* s'étire.

Dans le but de savoir les moments propices à engager du personnel supplémentaire, Ricardo surveille de près le nombre de ventes par heure selon la température extérieure. Voici un échantillon de ses observations.

Température (°C)	14	18	16	12	13	15	15	17	18	19
Nombre de glaces vendues par heure	10	16	11	9	8	10	12	13	17	20

L'été approche à grands pas, et Ricardo veut estimer la température extérieure qui lui assurerait une trentaine de ventes à l'heure.

Déterminer la température à laquelle on peut s'attendre à une trentaine de ventes à l'heure à *La Glacière*.

Toujours de l'espace
fourni afin d'écrire
vos développements!



Ces situations-problèmes sont plus globales et plus complexes afin de maîtriser les compétences transversales visées par ce module.



Pour conclure ce module

Pour terminer cette dernière étape, traitant des **distributions statistiques à deux caractères**, et pour vous assurer que vous maîtrisez bien les notions que vous y avez découvertes, vous traiterez maintenant des **SÉ**. Les solutions de ces situations ne sont pas dans votre module : votre enseignante ou votre enseignant en fera la correction.

Avant d'aborder ces **SÉ**, nous vous recommandons de noter, sur une feuille, les formules, les énoncés et même des exemples que vous jugez importants. Vous pouvez utiliser cette feuille comme aide-mémoire.

Assurez-vous de présenter une solution claire et complète. Vous ne devez demander l'aide de personne. Ce qui vous permettra de vous évaluer, et de connaître les exigences et les attentes de fin d'étape. Ce faisant, vous pourrez, si vous constatez certaines lacunes, les corriger avant de poursuivre.

Cette auto-évaluation vous permettra aussi de savoir si vous répondez aux attentes fixées pour ce module MAT 4272.

Une banque de situations-problèmes supplémentaires vous permettra d'augmenter encore plus vos compétences en seconde partie du **Prêt pour l'évaluation de fin de module?**

Bon travail !

Ces situations d'évaluation se trouvent à la fin du chapitre et sont divisées en 2 parties. Votre enseignant(e) en fera la correction.



01 PREMIÈRE PARTIE

Évaluation des connaissances

1. Estimer...

Ces situations d'évaluation vous permettent de vérifier l'acquisition des connaissances et des compétences dites transversales.



01 DEUXIÈME PARTIE

Évaluation des compétences

5. Marcher pour une bonne cause.

Un de vos amis...

Félicitations, vous êtes près de la fin, le questionnaire qui suit a été préparé pour vous permettre d'évaluer vos forces et vos faiblesses dans ce module. Le corrigé de ce questionnaire ne se trouve pas dans votre module. Votre enseignant en fera la correction.

La première partie de ce questionnaire porte sur les savoirs mathématiques de ce cours. Dans la deuxième partie de cette rubrique, vous trouverez dix situations-problèmes pour démontrer vos compétences liées à ce module: utiliser des stratégies de résolution de situations-problèmes et déployer un raisonnement mathématique. Bonne révision!

PREMIÈRE PARTIE

Révision des connaissances

1. Présenter...

Cette section est constituée de 2 banques d'exercices dont votre enseignant(e) en fera la correction: ceci dans le but d'évaluer vos forces et vos faiblesses.



DEUXIÈME PARTIE

Révision des compétences

Voici enfin le dernier virage avant l'examen: une banque de 10 situations-problèmes portant sur la collecte de données en contexte fondamental. Faites-en bon usage!

1. Des cônes orange.

Vous travaillez...

coefficient de corrélation linéaire

Le coefficient de corrélation linéaire est un nombre entre -1 et 1 qui quantifie l'intensité de la corrélation linéaire associée à un nuage de points.

corrélation

La corrélation entre deux variables statistiques numériques est l'intensité de la liaison qui existe entre ces variables. Le type le plus simple de liaison est la relation affine. Lorsque la liaison entre les variables se traduit par une droite, on dit que la corrélation est linéaire.

droite de régression

La droite de régression est une droite qui représente le mieux possible le nuage de points.

extrapolation

L'extrapolation consiste à déterminer la valeur d'une variable à partir d'une valeur connue de l'autre variable à l'extérieur d'un nuage de points.

interpolation

L'interpolation consiste à déterminer la valeur d'une variable à partir d'une valeur connue de l'autre variable à l'intérieur d'un nuage de points.

médiane

La médiane, symbolisée par les lettres Md , est la valeur située au centre de la distribution lorsque les données sont classées en ordre croissant.

méthode de la droite médiane-médiane

La méthode de la droite médiane-médiane permet d'obtenir une approximation de la droite de régression en utilisant les valeurs médianes de trois sous-groupes des données d'une distribution à deux caractères.

méthode de l'ellipse

La méthode de l'ellipse est une méthode graphique qui permet d'estimer le coefficient de corrélation entre deux variables. Cette méthode consiste à tracer une ellipse autour d'un nuage de points et à appliquer la formule: $r \approx \pm \left(1 - \frac{\text{mesure du petit axe}}{\text{mesure du grand axe}} \right)$.

01 DISTRIBUTIONS STATISTIQUES À DEUX CARACTÈRES Activités d'apprentissage

1.1. Construction et interprétation de tableaux de distributions à deux caractères

1. p. 13

a) **PRODUCTION EXPORTÉE SELON LE TYPE DE PRODUIT**

Année	2009	2010	Total
Production exportée			
Réfrigérateurs	10 500	12 200	22 700
Cuisinières au gaz	31 000	26 500	57 500
Cuisinières électriques	2 300	3 200	5 500
Téléviseurs	76 900	78 100	155 000
Robots	32 000	29 800	61 800
Total	152 700	149 800	302 500

Source: Département des ventes de l'entreprise

b) **RÉPARTITION DES ÉTUDIANTS D'UN CENTRE D'ÉDUCATION AUX ADULTES, SELON LE SEXE ET L'OCCUPATION D'UN EMPLOI**

Sexe	Masculin	Féminin	Total
Occupation d'un emploi			
Oui	3	15	18
Non	7	5	12
Total	10	20	30

Source: Sondage de la direction

c) Exemple de tableau:

CHOIX DES ACTIVITÉS DE LOISIR CHEZ LES ÉLÈVES D'UNE ÉCOLE SECONDAIRE, SELON LE SEXE

Sexe	Garçons	Filles	Total
Activité			
Sport	4	5	9
Musique	10	12	22
Jeux vidéo	7	13	20
Lecture	5	3	8
Télévision	7	9	16
Total	33	42	75

Source: Comité étudiants

On a interrogé au total **75** élèves.

Un corrigé aéré, élaboré avec une démarche détaillée, qui vous permet de vous vérifier de façon autonome, pour toutes les Activités d'apprentissage.

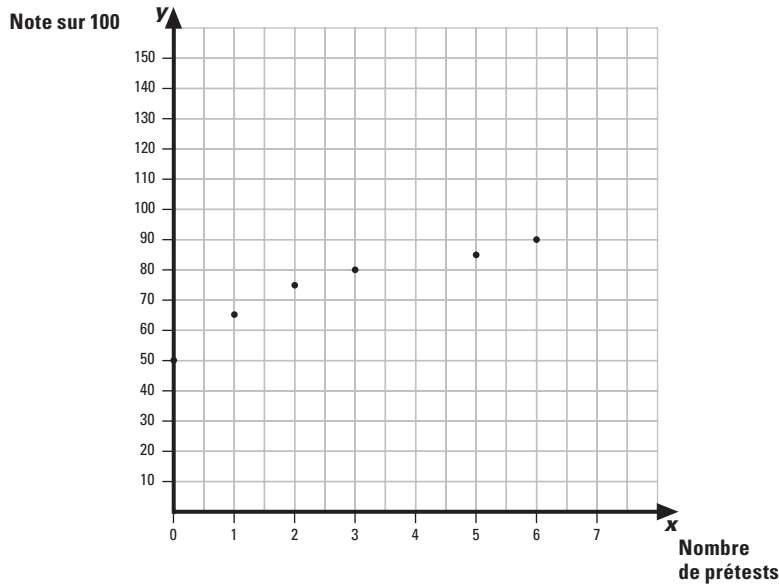


1.9. Vue d'ensemble : synthèse des savoirs

1. p. 135

a)

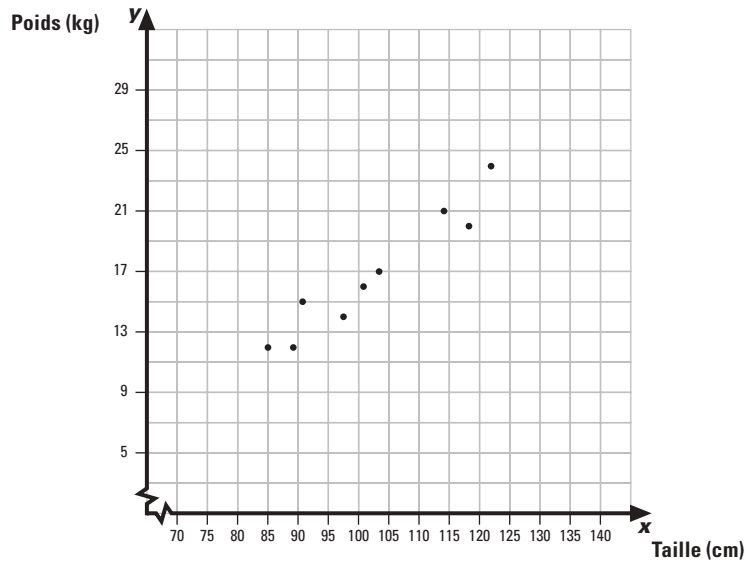
NOTES À L'EXAMEN FINAL SELON LE NOMBRE DE PRÉTESTS



La corrélation est **forte** et **positive** entre le nombre de prétests faits et la note obtenue à l'examen.

b)

POIDS SELON LA TAILLE DE 9 ENFANTS D'UNE GARDERIE FAMILIALE



La corrélation est **forte** et **positive** entre la taille et le poids des enfants.

Un corrigé aéré, élaboré avec une démarche détaillée, qui vous permet de vous vérifier de façon autonome, pour toutes les Consolidations des savoirs.

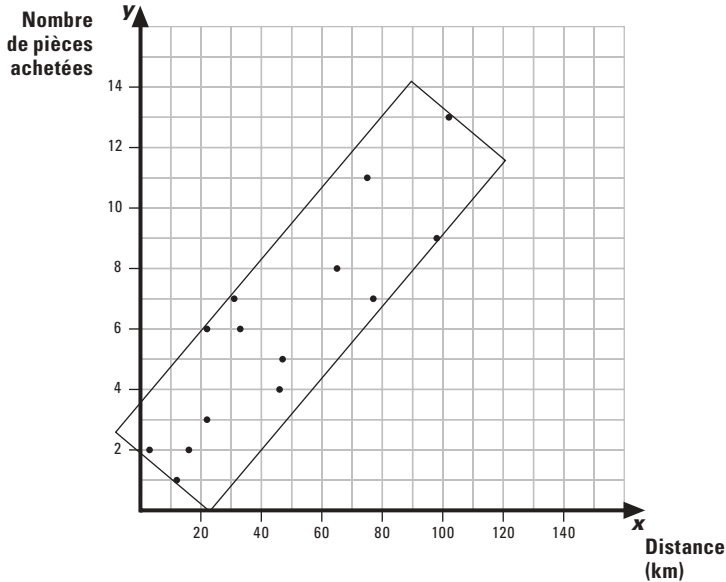


1.10. Situations de vie

1. Votre hypothèse est-elle défendable?

p. 149

1^{re} tâche



$$r \approx \pm \left(1 - \frac{\text{mesure du petit côté}}{\text{mesure du grand côté}} \right)$$

$$r \approx \pm \left(1 - \frac{1,6}{6} \right)$$

$$r \approx \pm 0,73$$

$$r \approx 0,73$$

Le coefficient de corrélation de la distribution est de 0,73.

La corrélation est moyenne et positive.

Votre hypothèse est vérifiée: plus un client demeure loin du magasin, plus il achète de pièces, mais la corrélation demeure moyenne.

2^e tâche

Le coefficient de corrélation linéaire est moyen, on peut utiliser la droite de régression pour obtenir une extrapolation valable.

Exemple de solution, par la méthode de Mayer:

Les données sont: {(3, 2), (12, 1), (16, 2), (22, 3), (22, 6), (31, 7), (33, 6), (46, 4), (47, 5), (65, 8), (75, 11), (77, 7), (98, 9), (102, 13)}

Groupe 1: {(3, 2), (12, 1), (16, 2), (22, 3), (22, 6), (31, 7), (33, 6)}

$$G_1: \left(\frac{3 + 12 + 16 + 22 + 22 + 31 + 33}{7}, \frac{2 + 1 + 2 + 3 + 6 + 7 + 6}{7} \right)$$

$$G_1: \left(\frac{139}{7}, \frac{27}{7} \right)$$

$$G_1: (19,86; 3,86)$$

Groupe 2: {(46, 4), (47, 5), (65, 8), (75, 11), (77, 7), (98, 9), (102, 13)}

$$G_2: \left(\frac{46 + 47 + 65 + 75 + 77 + 98 + 102}{7}, \frac{4 + 5 + 8 + 11 + 7 + 9 + 13}{7} \right)$$

$$G_2: \left(\frac{510}{7}, \frac{57}{7} \right)$$

$$G_2: (72,86; 8,14)$$

Un corrigé aéré, élaboré avec une démarche détaillée, qui vous permet de vous vérifier de façon autonome, pour toutes les Situations de vie.



1. À La Glacière.

p. 153

Exemple de solution, par la méthode de la droite médiane-médiane :

Température (°C)	14	18	16	12	13	15	15	1
Nombre de glaces vendues par heure	10	16	11	9	8	10	12	1

On ordonne les points par ordre croissant des valeurs de x :
 (12, 9), (13, 8), (14, 10), (15, 10), (15, 12), (16, 11), (17, 13), (18, 16), (18, 17), (19, 20)

Groupe 1 = {(12, 9), (13, 8), (14, 10)}

M₁: (13, 9)

Groupe 2 = {(15, 10), (15, 12), (16, 11), (17, 13)}

$$M_2: \left(\frac{15+16}{2}, \frac{11+12}{2} \right)$$

M₂: (15,5; 11,5)

Groupe 3 = {(18, 16), (18, 17), (19, 20)}

M₃: (18, 17)

Calcul des coordonnées du point P:

$$P: \left(\frac{13+15,5+18}{3}, \frac{9+11,5+17}{3} \right)$$

P: (15,5; 12,5)

Pente de la droite de régression:

$$a = \frac{Y_2 - Y_1}{x_2 - x_1}$$

$$a = \frac{17 - 9}{18 - 13}$$

a = 1,6

Équation de la droite de régression:

$$y = 1,6x + b$$

$$12,5 = 1,6 \cdot 15,5 + b$$

$$12,5 = 24,8 + b$$

$$12,5 - 24,8 = b$$

$$\mathbf{-12,3 = b}$$

L'équation de la droite de régression est $y = 1,6x - 12,3$.Si $y = 30$:

$$30 = 1,6x - 12,3$$

$$30 + 12,3 = 1,6x$$

$$42,3 = 1,6x$$

$$x = \frac{42,3}{1,6}$$

$$x \approx 26,4$$

À une température d'environ 26 °C, Ricardo peut s'attendre à une trentaine de ventes par heure.

Un corrigé aéré, élaboré avec une démarche détaillée, qui vous permet de vous vérifier de façon autonome, pour toutes les Situations-problèmes.



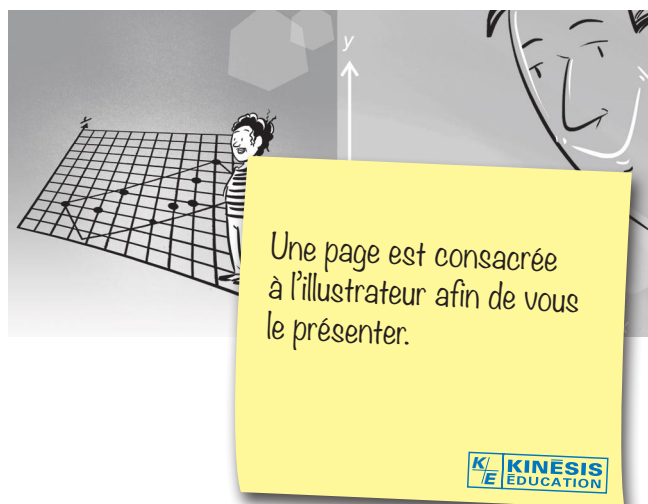
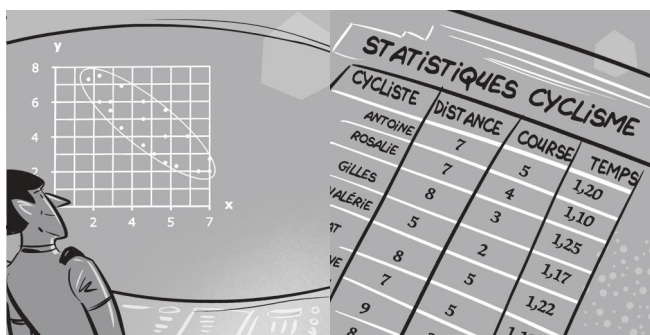
MOTS	CHAPITRE 1
Coefficient de corrélation linéaire	38, 39, 40, 41, 43, 44, 47, 49, 131, 132
Corrélation	24, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49
Droite de régression	65, 66, 67, 68, 70, 71, 72, 73, 74, 82, 83, 84, 85, 86, 87, 88, 101, 102, 103, 108, 109, 110, 111, 112, 113, 114, 115
Extrapolation	84, 87, 90, 122, 123, 134
Extrapolation à l'aide d'une fonction en escalier	123, 134
Extrapolation à l'aide d'une fonction quadratique	122, 134
Interpolation	82, 84, 85, 86, 89, 90, 134
Méthode de la droite médiane-médiane	97, 100, 103, 133
Méthode de l'ellipse	41, 46, 47, 131, 132
Méthode de Mayer	108, 114, 134
Méthode des moyennes	66, 68, 72, 133
Méthode du rectangle	39, 42, 44, 131
Nuage de points	21, 22, 24, 25, 26, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 65, 66, 67, 68, 69, 71, 72, 73, 74, 82, 84, 86, 87, 89, 90, 97, 108, 121, 122, 123, 124, 125, 126, 127, 130, 131, 132, 133, 134
Tableau à double entrée	4, 8, 130
Tableau de distribution à deux caractères	4, 6, 9, 12, 22, 130

Une table alphabétique des mots clés et leurs références.



À propos de l'illustrateur et des illustrations...

Les illustrations des couvertures et les illustrations que vous trouverez au fil des pages de ce module sont des illustrations originales, commandées pour notre collection à Paul Bordeleau, illustrateur québécois, auteur de bandes dessinées et illustrateur-éditorialiste pour l'hebdomadaire *Voir* de 1992 à 2004, et pour le journal *La Presse* en 2001 et 2002. En 2003, il a pris la relève de Garnotte et de Gité comme illustrateur de nos collections.

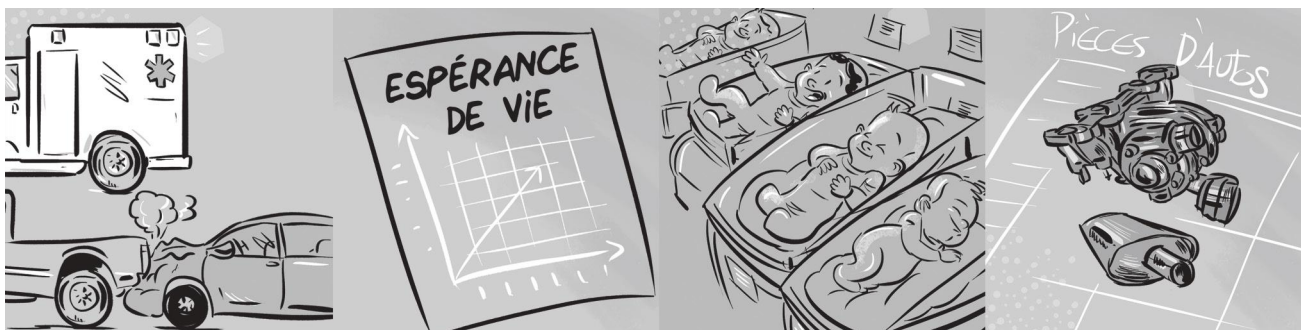


En 2009, il était l'un des bédéistes invités au festival *BoomFest* de Saint-Pétersbourg, en Russie. Il a illustré entre autres le générique de la télésérie *La Galère* à Ici Radio-Canada. En 2016, il a participé au projet *Correspondances* de Lyon.

Dans la collection MAT, ses illustrations sont parfois conçues comme de petites pauses détente au fil des chapitres.

D'autres fois, elles sont des illustrations essentielles à la compréhension et à la résolution des situations qui vous sont présentées.

Dans les pages d'ouverture des chapitres, elles illustrent la situation concrète qui vous amène à vous plonger dans la réalité mathématique des activités d'apprentissage et des situations-problèmes. Ces activités et ces situations vous permettent d'acquérir la maîtrise des savoirs mathématiques visée par le module.



Vous voulez en savoir plus sur Paul Bordeleau ?
Voici ses coordonnées : www.paulbordeleau.com



Utilisation de la technologie pour calculer le coefficient de corrélation

L'utilisation de moyens technologiques facilite grandement le calcul du coefficient de corrélation linéaire d'une distribution statistique à deux variables. Le logiciel *Microsoft Excel*, la calculatrice statistique et la calculatrice graphique permettent de déterminer rapidement et facilement le coefficient de corrélation.

Utilisation de *Microsoft Excel*

Pour déterminer le coefficient de corrélation linéaire à partir des données d'une distribution statistique à deux caractères, on procède de la façon suivante :

Étape 1 :

On **ouvre** une feuille de calcul *Excel*;

Dans la case **A1**, on **inscrit** x ou le libellé de la variable ;

On **reporte** les valeurs de x dans chacune des cases de la colonne **A** ;

Exemple

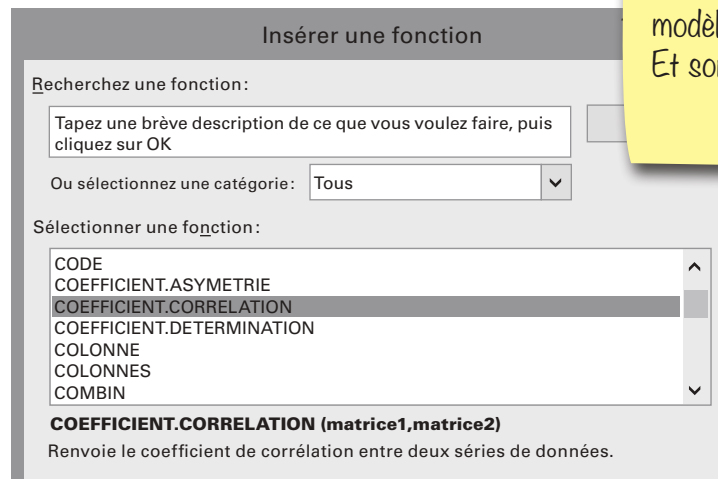
x_1 dans la case A2 ; x_2 dans la case A3, etc...

On suit la **même procédure** pour inscrire les valeurs de la variable y dans la colonne B ;

On **positionne** ensuite le **curseur** dans une case vide pour afficher la valeur du coefficient de corrélation r ;

On **clique** sur le symbole $f(x)$ qui s'affiche dans la barre au-dessus de la cellule ;

La boîte de dialogue suivante s'ouvre.



Pour apprendre à utiliser certains logiciels ou les fonctions de certains modèles de calculatrice. Et son corrigé.



6. p. 160 suite

La fonction qui approxime le mieux le nuage de points est quadratique.

À l'aide d'un outil technologique, on détermine les paramètres de la fonction quadratique:

$$y = 82,047 - 2,296x + 0,088x^2$$

Pour $x = 0$:

$$y = 82,047 - 2,296 \cdot 0 + 0,088 \cdot 0^2$$

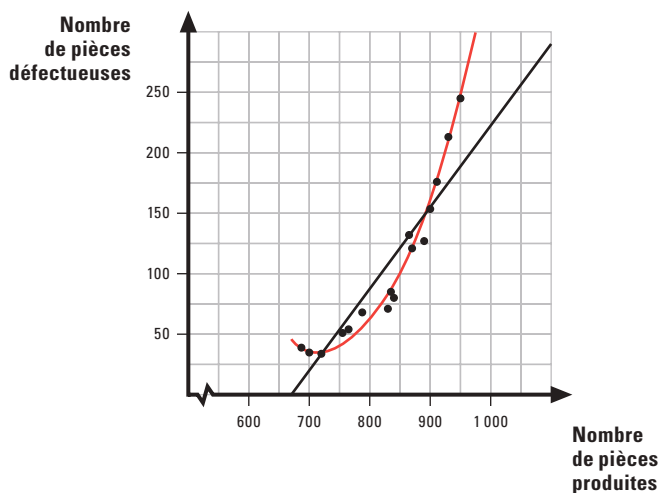
$$y \approx 82$$

À 0 °C, la durée de la randonnée est d'environ 82 minutes.

7. Les pièces de verre artisanal.

p. 161

1^{re} tâche



Le coefficient de corrélation linéaire est 0,92.

Oui, vos soupçons sont fondés: le nuage de points se rapproche davantage d'une parabole que d'une droite car $0,988 > 0,92$.

2^e tâche

À l'aide d'un outil technologique, on peut établir que la règle de la fonction quadratique qui approxime le mieux le nuage de points est: $y = 0,0039x^2 - 5,6492x + 2081,1$.

Pour $x = 1200$:

$$y = 0,0039 \cdot 1200^2 - 5,6492 \cdot 1200 + 2081,1$$

$$y = 918$$

Non, il n'est pas rentable d'avoir autant de stagiaires, car, pour 1200 pièces produites, on peut s'attendre à 918 pièces défectueuses, soit plus des trois quarts.

Pause technologie / page 56

Utilisation de la technologie pour calculer le coefficient de corrélation linéaire

- a) $r = -0,74$
- b) $r = 0,82$
- c) $r = -0,31$

Le MAT 4272

Vise l'acquisition de deux grandes compétences transversales : se donner des méthodes de travail efficace et exercer son jugement critique. Au moyen de trois procédés intégrateurs : la réalisation d'une collecte de données, la comparaison de collectes de données et l'interprétation de données issues d'une expérience.

MAT^{SN} 4272 2

FORMATION DE BASE DIVERSIFIÉE



Notre maison n'a qu'une seule et unique raison d'être depuis sa création il y a plus d'un demi-siècle : publier des ouvrages de qualité irréprochable, de bonne tenue, aux contenus solides, privilégiant des démarches en accord avec les principes des différentes approches pédagogiques, et libres de tout compromis de caractère purement commercial.



401 1524

Florence Grandchamp
Drita Neziri
Abdelkader Amara
Raymond Thériault

ÉDITION
2019

COLLECTE DE DONNÉES EN CONTEXTE FONDAMENTAL

MAT
AT SN
A 4272 2

FORMATION DE BASE DIVERSIFIÉE

Ce document est disponible
gratuitement pour
l'enseignant(e). Il suffit
d'en faire la demande
à editions@ebbp.ca

 KINESIS
EDUCATION

TIRÉ À PART

Corrigé des *Situations d'évaluation de fin de chapitre*

Grilles d'évaluation

Corrigé du *Prêt pour l'évaluation de fin de module?*

 KINESIS
EDUCATION

L'éditeur permet la reproduction
de ce document.