

ÉDITION
AOÛT
2022

Florence Grandchamp
Drita Neziri
Abdelkader Amara
Raymond Thériault

COLLECTE DE DONNÉES

MAT

3052 2

FORMATION DE BASE DIVERSIFIÉE



Graphismes, notations et symboles

Graphismes, notations
et symboles utilisés
dans ce module



\approx	environ
%	pour cent
60°	60 degrés
Mo	mode
Md	médiane
\bar{x}	moyenne
$[16, 18[$	l'intervalle fermé-ouvert de 16 à 18
x_i	i° donnée d'une distribution
f_i	fréquence de la i° donnée
Σ	sommation
Q_1	premier quartile
Q_2	deuxième quartile ou médiane
Q_3	troisième quartile
EI	étendue interquartile
\overline{AB}	le segment AB
$P(A)$	probabilité de l'événement A
Ω	univers des résultats possibles
$\{1, 2, 3\}$	l'ensemble des éléments 1, 2 et 3
$A \cap B$	A intersection B
$A \cup B$	A union B
\emptyset	l'ensemble vide
A'	le complémentaire de l'événement A
(x, y)	couple
(x, y, z)	triplet

Tableaux

Rappel de quelques notions



Tableau de données condensées

NOMBRE D'ENFANTS DE 50 COUPLES

Nombre d'enfants	Effectif
0	12
1	19
2	11
3	5
4	2
5	1
Total	50

Tableau de données

REVENU ANNUEL DE 50 COUPLES

Classe (milliers de \$)	Effectif
[0, 20[2
[20, 40[6
[40, 60[11
[60, 80[13
[80, 100[9
[100, 120[5
120 ou plus	4

Mesures de tendance centrale

Mode: la donnée la plus fréquente

Médiane: la donnée au milieu de la distribution dont les données sont classées par ordre croissant

Moyenne: la somme de toutes les données divisée par le nombre de données

Diagrammes

Histogramme

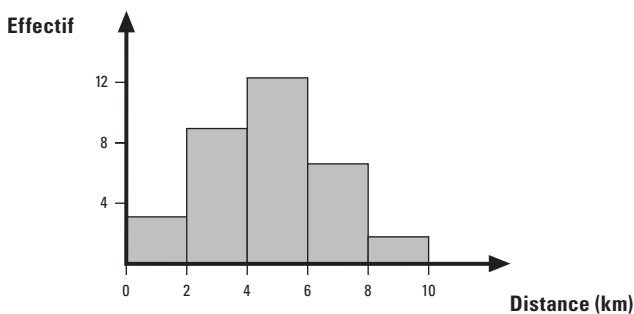


Diagramme de quartiles

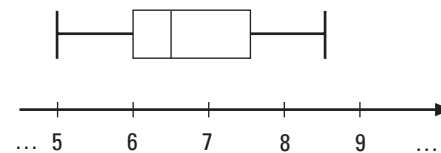


Diagramme en arbre

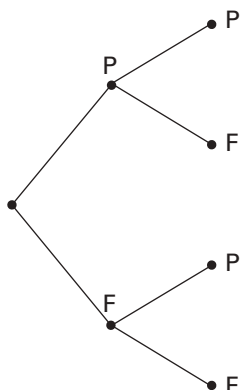
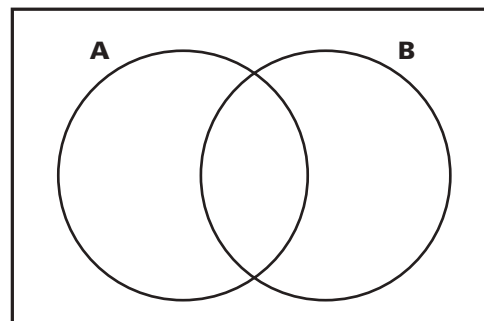
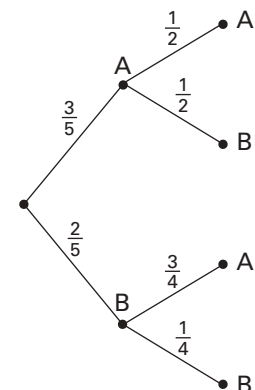


Diagramme de Venn



Arbre de probabilité



COLLECTE DE DONNÉES

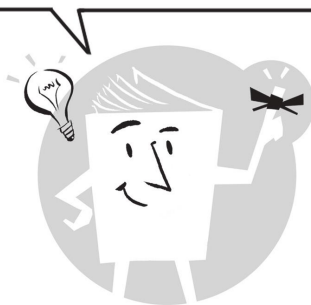
Conforme au Programme



MAT A 3052 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



Tous les 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

MAT 1101 3 MAT 1102 3

Secondaire 2

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

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

FORMATION À DISTANCE :

Secondaire 1, 2 et 3

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

Secondaire 4 et 5 — *En préparation*

Ouvrages déjà parus au catalogue :

MAT 1005 2	MAT 1006 2	MAT 1007 2	MAT 2006 2	MAT 2007 2	MAT 2008 2
MAT 3015 2	MAT 3016 2	MAT 3017 2			
MAT 4101 2	MAT 4102 1	MAT 4103 1	MAT 4104 2	MAT 4105 1	MAT 4106 1
MAT 4107 1	MAT 4108 1	MAT 4109 1	MAT 4110 1	MAT 4111 2	
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
MAN 1000	MAN 2000	MAN 3000		MAT 1005 FAD à MAT 5112 FAD	



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
	Jonathan Crête
Correction	
Direction de la collection	Célestin de La Grange
• contenu éditorial	Annie Lopez
• contenu mathématique	Florence Grandchamp
• infographie et production	Francine Plante
Ideatrice	Marianne Delaroché
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	Louis Bouchard
Révision mathématique	Sylvain Gervais
Révision pédagogique	Mohamed-Seghir Ghellache

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.

Tout usage en location ou prêt est interdit sans autorisation écrite octroyée par Kinésis éducation inc.

Impression Imprimerie Héon & Nadeau

Éditrice déléguée Francine Plante / Les Éditions Jules Châtelain

Page des crédits



Pour en savoir plus sur l'illustrateur et sur les illustrations de votre module, voir p. 401



© 2017-2019, Kinésis éducation inc. Tous droits réservés.

Dépôt légal — Bibliothèque et Archives nationales du Québec, Bibliothèque et Archives Canada, 2019.

ISBN 978-2-7615-0899-5 (2^e édition, 2019)

ISBN 978-2-7615-0645-8 (1^{re} édition, 2017)

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

Vous avez en main la deuxième édition revue et corrigée du module MAT 3052, deuxième 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 relations avec les enseignants. Nous accusons toujours réception de la correspondance reçue des enseignants. 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	
Tableaux	page 3 de couverture
Mesures de tendance centrale	page 3 de couverture
Diagrammes	page 3 de couverture
À l'étudiant et à l'enseignant	V
Présentation	VIII
Comment est construit votre MAT 3052	X
Attentes de fin de cours	XII

01. DISTRIBUTIONS STATISTIQUES À UN CARACTÈRE

Mise en situation:	
UNE ÉTUDE DE MARCHÉ POUR UN MICRO-ORDINATEUR DESTINÉ AUX GENS ACTIFS	2
1.1. Organisation et interprétation de données statistiques	4
1.2. Construction et interprétation de tableaux à données condensées	17
1.3. Construction et interprétation de tableaux à données groupées en classes	28
1.4. Interprétation et construction d'un histogramme	38
Pour en savoir un peu plus... : Loi normale et QI	46
1.5. Le mode d'une distribution	47
1.6. La médiane d'une distribution	56
1.7. Calcul de la moyenne d'une distribution	68
En remontant le cours des siècles: La statistique à travers les âges	80
1.8. Choix d'une mesure de tendance centrale	81
1.9. Les quartiles d'une distribution	91
1.10. Interprétation et représentation d'un diagramme de quartiles	97
1.11. Vue d'ensemble : synthèse des savoirs	109
Consolidation des savoirs	114
1.12. Situations de vie	131
Situations d'évaluation de fin de chapitre SÉ	163
Évaluation des connaissances	164
Évaluation des compétences	166

02. PROBABILITÉS

Mise en situation :

RENTABILISER UNE INVENTION GRÂCE AU DÉNOMBREMENT **172**

2.1.	Variable aléatoire	174
	Pour en savoir un peu plus... : Le jeu de cartes	179
2.2.	Représentation d'événements	180
2.3.	Le dénombrement	194
	Amusons-nous: Le café	204
2.4.	Expériences aléatoires simples	205
2.5.	Expériences aléatoires composées	215
	Amusons-nous: Les coïncidences qui n'en sont pas	231
2.6.	Probabilité géométrique	232
	En remontant le cours des siècles: Blaise Pascal (1623–1662)	247
2.7.	Vue d'ensemble: synthèse des savoirs	248
	Consolidation des savoirs	250
2.8.	Situations de vie	262
	Situations d'évaluation de fin de chapitre SÉ	283
	Évaluation des connaissances	284
	Évaluation des compétences	287

Prêt pour l'évaluation de fin de module ? **291**Révision des connaissances **291**Révision des compétences **308**Glossaire des termes mathématiques **326**Corrigé **337**Index **395**À propos de l'illustrateur et des illustrations... **401****Nos petits plus...**Amusons-nous **204, 231**En remontant le cours des siècles **80, 247**Pour en savoir un peu plus... **46, 179**

COLLECTE DE DONNÉES

Le module MAT 3052, intitulé **Collecte de données**, touchera plusieurs à la 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. Le module **Collecte de données** vous donnera l'occasion de poser des actions en vue d'effectuer ou de comparer des collectes de données.

En traitant les situations-problèmes de ce module, vous serez amené, entre autres, à organiser les données d'un échantillon afin de décrire une population et d'en tirer des conclusions, à interpréter des données provenant d'une étude statistique ou d'une expérience aléatoire ou encore, à prendre position à partir d'un portrait statistique ou probabiliste.

COMPÉTENCES DISCIPLINAIRES

Dans ce cours, la résolution de situations-problèmes implique le 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 intellectuel : *Exercer son jugement critique ;*

Compétence de l'ordre de la communication : *Communiquer de façon appropriée.*

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 complèterez votre formation en construisant et en vous appropriant les savoirs suivants.

Savoirs prescrits

En vue de traiter efficacement les situations d'apprentissage proposées dans ce cours, vous développerez trois **procédés intégrateurs** :

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

SAVOIRS MATHÉMATIQUES**Distributions statistiques à un caractère**

SM-1 Organisation et interprétation de données statistiques

SM-2 Construction et interprétation de tableaux de distribution

SM-3 Représentation et interprétation de graphiques

SM-4 Calcul de mesures de tendance centrale et de dispersion

SM-5

SM-5 Recensement et calcul de probabilités

SM-6 Représentation d'événements

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



COLLECTE DE DONNÉES

PRÉSENTATION

Présentation des *compétences disciplinaires*, des *compétences transversales*, et du contenu disciplinaire visés par le MAT 3052. ➔ page VIII

COMMENT EST CON

Les deux pages

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



Votre MAT 3052 est divisé en chapitres :

01

DISTRIBUTIONS STATISTIQUES À UN CARACTÈRE

En début de chapitre une *mise en situation*, ici : **UNE ÉTUDE DE MARCHÉ POUR UN MICRO-ORDINATEUR DESTINÉ AUX GENS ACTIFS**. Elle est tirée de la vie courante réelle ou virtuelle, et illustre l'utilité de la matière qui sera abordée. **DANS CE CHAPITRE**, 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



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

Les chapitres de votre MAT 3052 sont divisés en sections :

1.1. Organisation et interprétation de données statistiques

1.11. 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 109 et suivantes

En conclusion du chapitre, des

1.12. 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 131

MAT 3052

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 291

MAT 3052 GLOSSAIRE DES TERMES MATHÉMATIQUES

Un mini-dictionnaire : tous les termes apparaissant en **italique rouge gras** dans le module. ➔ page 326

Et des petits plus....

Amusons-nous

Les mathématiques, un divertissement ? Eh oui... on peut aussi s'amuser en faisant des mathématiques.
➔ page 204

En remontant le cours des siècles XVIII^e à nos jours

Un peu d'histoire pour mieux comprendre les mathématiques.
➔ page 80

ATTENTES DE FIN DE COURS

MAT 3052

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 11 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 3052. Ces situations se repèrent grâce à la bande gris foncé sur la tranche du module.

➔ page 142 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 à la fin de chaque chapitre. 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 337 et suivantes

MAT 3052

INDEX

Une table alphabétique des mots-clés et leurs références. ➔ page 395 et suivantes

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

Pour en savoir un peu plus...

Pour les curieux... un prolongement des connaissances, et de l'enrichissement.

➔ page 46

ATTENTES DE FIN DE COURS

Objectifs visés
par ce cours



Au terme de ce cours, vous serez en mesure d'effectuer une collecte de données. Vous pourrez aussi comparer les résultats d'une expérience statistique à l'aide de divers instruments pour valider vos observations relativement à 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. Vous utiliserez des stratégies de résolution de situations-problèmes afin de déterminer la solution la plus juste. De plus, vous interpréterez, à l'aide du raisonnement mathématique, des données probabilistes issues d'une expérience aléatoire et vous prendrez des décisions.

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, orale ou écrite, de la compréhension de la situation-problème
- 1.2 Mobilisation des stratégies et des savoirs mathématiques appropriés
- 1.3 Élaboration d'une solution appropriée

2. DÉPLOYER UN RAISONNEMENT MATHÉMATIQUE

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

Votre MAT 3052
est divisé en 2 chapitres
dont voici les titres:



**01. DISTRIBUTIONS STATISTIQUES
À UN CARACTÈRE**

02. PROBABILITÉS

Dans ce chapitre, vous aurez l'occasion d'approfondir vos connaissances en statistique. Vous explorerez des outils qui faciliteront l'analyse des données, leur interprétation et leur présentation.

Mise en situation:

UNE ÉTUDE DE MARCHÉ POUR UN MICRO-ORDINATEUR DESTINÉ AUX GENS ACTIFS

Un soir, devant votre ordinateur portable, vous consultez un site *Internet participatif*, aussi connu sous l'appellation anglaise de *crowdfunding*. Sur ce site, des personnes ou des entreprises font appel au public afin de financer de nouveaux projets. Pour chacun de ces projets, une somme à amasser est fixée, ainsi que le temps alloué pour le réaliser. La personne qui désire investir dans un projet le fait par l'intermédiaire de ce site Internet participatif, et la taille de sa contribution est laissée à son entière discrétion. Elle peut même décider de verser une somme aussi petite que 5 \$, et aucun plafond n'est fixé. Lorsqu'on ne réussit pas à recueillir les fonds prévus avant la date butoir, le projet est annulé et les investisseurs n'ont pas à verser l'argent qu'ils avaient promis.

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.



Un projet de micro-ordinateur a capté votre attention, mais la somme désirée n'a pu être amassée à temps. Conçu pour les gens friands d'exercices, l'ordinateur peut afficher de nombreux paramètres liés à l'exercice que vous êtes en train de faire (que ce soit de la natation, de la course à pied, du cyclisme ou de la culture physique) et enregistre des données, telles que les répétitions ou le nombre de kilomètres parcourus. Puis, à la fin de la séance d'exercice, il transmet ces données à vos réseaux sociaux, par l'intermédiaire d'une connexion sans fil.

Le financement a échoué, notamment parce que l'on n'a pas procédé à une étude de marché, c'est-à-dire une analyse statistique qui permettrait de cerner le public cible d'un tel produit, la clientèle qui aurait intérêt à acheter ce micro-ordinateur.

Vous prenez donc contact avec les inventeurs et leur proposez de faire cette analyse avant qu'ils ne soumettent de nouveau le projet sur le site de financement participatif. Par chance, ils ont déjà recueilli quelques données, mais ne savent pas comment les interpréter. Sauriez-vous leur venir en aide ?

Grâce à la statistique, vous serez en mesure, à la fin de ce chapitre, d'effectuer une cueillette de données, puis de présenter et d'interpréter ces données.

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 CHAPITRE

Quoi de nouveau ?

- La présentation et l'analyse des résultats d'une collecte de données

Qu'est-ce que c'est ?

- Pour présenter et analyser les résultats d'une collecte de données, on fait appel à différents outils, tels que : les méthodes d'échantillonnage, les tableaux, les diagrammes, les mesures de tendance centrale et de dispersion.

À quoi ça sert en mathématiques ?

- La présentation et l'analyse des résultats d'une collecte de données permettent non seulement de tirer des conclusions sur une distribution statistique, mais aussi de comparer plusieurs distributions.

À quoi ça servira dans la vie ?

- Les conclusions tirées des résultats des études statistiques aident à prendre des décisions appropriées dans divers domaines : marketing, démographie, finances, etc.

1.1. Organisation et interprétation de

Chaque chapitre est divisé en sections.



- CHAQUE JOUR, LES MÉDIAS NOUS PRÉSENTENT DES ÉCHANGES SUR LESQUELS DES GROUPES ET DES INDIVIDUS PRENDENT DES DÉCISIONS DANS DIFFÉRENTES SPHÈRES D'ACTIVITÉ.

FORMER UN ÉCHANTILLON REPRÉSENTATIF DE LA POPULATION EST L'UNE DES PREMIÈRES ÉTAPES DE TOUTE ÉTUDE STATISTIQUE.

CETTE SECTION VOUS PERMETTRA D'ÉTUDIER LES MÉTHODES D'ÉCHANTILLONNAGE UTILISÉES POUR LA COLLECTE DE DONNÉES STATISTIQUES.



SM-1

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



Outils mathématiques

Rappel des définitions de base en statistique – Méthodes d'échantillonnage

1. Rappel des définitions de base en statistique

Une **population** est l'ensemble des individus ou des objets sur lesquels porte une étude statistique et qui sont regroupés par des critères qui permettent de les regrouper. Lorsque l'étude statistique se fait sur une population, ces individus ou ces objets forment ce qu'on appelle un **échantillon**. Pour former un **échantillon représentatif**, il doit refléter le plus possible les caractéristiques de la population.

Tous les termes apparaissant en italique rouge gras se retrouvent au glossaire des termes mathématiques.

Il existe plusieurs méthodes d'échantillonnage :



Le **recensement**, qui est un dénombrement qu'on effectue sur **toute une population d'individus**, ou l'**inventaire**, qu'on effectue sur **toute une population d'objets** ;

Le **sondage**, qui est une étude ponctuelle qu'on effectue sur un **échantillon** de la population et dont les résultats sont généralisés à cette population ;

L'**enquête statistique**, qui est une étude portant sur un sujet donné, au cours de laquelle la compétence d'experts est nécessaire pour tirer des conclusions fiables.

Dans tout sondage, il est important de sélectionner un **échantillon représentatif** de la population. Une mauvaise méthode d'échantillonnage peut devenir une **source de biais**, c'est-à-dire une cause qui affecte le résultat d'une étude.

2. Méthodes d'échantillonnage

Vous verrez ici quatre méthodes d'échantillonnage : l'**échantillonnage aléatoire** et l'**échantillonnage systématique**, que vous connaissez déjà ; puis deux nouvelles méthodes, l'**échantillonnage stratifié** et l'**échantillonnage par grappes**.

L'échantillonnage aléatoire

Constituer un **échantillon aléatoire** consiste à choisir au **hasard** un certain nombre d'individus dans la population visée. On doit s'assurer que, pour chaque individu de la population, la probabilité de faire partie de l'échantillon soit la même.





Outils mathématiques suite

Exemple

Au gala des *mérites* d'un centre d'éducation des adultes, on doit remettre un cadeau à deux élèves choisis au hasard parmi les 25 qui se sont distingués durant l'année. On inscrit le nom de chacun des candidats sur un bout de papier, on dépose le tout dans un chapeau et on invite une personne à tirer au sort les noms des élèves qui recevront un cadeau. Ces gagnants forment un **échantillon aléatoire**, puisqu'ils sont déterminés par le **hasard**.

L'échantillonnage systématique

La méthode d'**échantillonnage systématique**, ou méthode d'**échantillonnage par intervalles**, consiste à constituer un échantillon en sélectionnant à **intervalle fixe** des individus ou des objets à partir d'une liste ordonnée contenant tous les individus ou tous les objets de la population concernée par l'étude.

Exemple

On veut choisir 25 élèves d'une liste de 150 élèves qui fréquentent un centre d'éducation des adultes. Divisons d'abord 150 par 25, ce qui nous donne 6. Nous devons donc sélectionner un élève sur 6. Par exemple, si on commence par le 3^e, on sélectionne ensuite le 9^e, le 15^e, le 21^e, le 27^e, le 33^e, le 39^e, le 45^e, le 51^e, le 57^e, le 63^e, le 69^e, le 75^e, le 81^e, le 87^e, le 93^e, le 99^e, le 105^e, le 111^e, le 117^e, le 123^e, le 129^e, le 135^e, le 141^e et le 147^e. Dès lors, on aura constitué un **échantillon systématique** de 25 élèves parmi les 150 élèves.

En procédant par échantillonnage systématique à intervalle fixe de 6, il y a six échantillons possibles :

1^{er}, 7^e, 13^e, 19^e, 25^e, 31^e, , 145^e
2^e, 8^e, 14^e, 20^e, 26^e, 32^e, , 146^e
3^e, 9^e, 15^e, 21^e, 27^e, 33^e, , 147^e
4^e, 10^e, 16^e, 22^e, 28^e, 34^e, , 148^e
5^e, 11^e, 17^e, 23^e, 29^e, 35^e, , 149^e
6^e, 12^e, 18^e, 24^e, 30^e, 36^e, , 150^e

Vous constaterez qu'au bout du compte, chaque élève du centre n'apparaît que dans un seul des 6 échantillons et qu'il a autant de chances d'être sélectionné que les autres élèves.

L'échantillonnage stratifié

La méthode d'**échantillonnage stratifié** est utilisée quand la population **n'est pas homogène**, c'est-à-dire lorsqu'une ou plusieurs caractéristiques des individus peuvent influencer les résultats d'un sondage. L'**échantillonnage stratifié** consiste à diviser la population en **strates**, c'est-à-dire en des **sous-ensembles homogènes et mutuellement exclusifs**, et ce, à partir de quelques critères : sexe, âge, langue parlée, pays d'appartenance, etc. On prélève ensuite un nombre d'individus de chaque strate **proportionnel** au nombre d'individus dans la strate correspondante de la population. Ces individus forment un échantillon représentant toutes les strates de la population visée par l'étude.

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





Outils mathématiques suite

Exemple

Un centre d'éducation des adultes compte 400 élèves; 250 sont de sexe masculin, 150 de sexe féminin. On veut choisir un échantillon représentatif de 40 élèves. De combien d'hommes et de femmes sera composé l'échantillon ?

1. La population concernée est l'ensemble des élèves. Désignons par N l'effectif de la population :
 $N = 400$.

2. On considère deux strates : celle qui est formée par les élèves de sexe masculin et celle qui est formée par les élèves de sexe féminin. Désignons par N_1 l'effectif de la première strate et par N_2 l'effectif de la deuxième strate :

Élèves de sexe masculin : **$N_1 = 250$;**

Élèves de sexe féminin : **$N_2 = 150$.**

3. On calcule le pourcentage que représente chaque strate. Désignons par P_1 le pourcentage de représentativité de la première strate et par P_2 le pourcentage de représentativité de la deuxième strate :

$$P_1 = N_1 \div N \times 100 \%$$

$$P_1 = 250 \div 400 \times 100 \%$$

$$P_1 = \mathbf{62,5 \%$$

$$P_2 = N_2 \div N \times 100 \%$$

$$P_2 = 150 \div 400 \times 100 \%$$

$$P_2 = \mathbf{37,5 \%$$

4. On détermine l'effectif de chaque strate de l'échantillon en multipliant chacun des pourcentages précédemment calculés par l'effectif de la strate correspondante de l'échantillon :

$$n_1 = 62,5 \% \text{ de } 40$$

$$n_1 = (62,5 \div 100) \times 40$$

$$n_1 = \mathbf{25}$$

$$n_2 = 37,5 \% \text{ de } 40$$

$$n_2 = (37,5 \div 100) \times 40$$

$$n_2 = \mathbf{15}$$

5. Parmi les 250 élèves de sexe masculin, on en choisit **25** au hasard (ou on utilise tout autre mode de sélection) et, parmi les 150 élèves de sexe féminin, on en choisit **15**.

De façon générale, voici la marche à suivre pour former un échantillon stratifié.

Marche à suivre pour former un échantillon stratifié

1. On établit la liste de la population concernée par l'étude statistique. On désigne par **N l'effectif total** de la population.

2. On subdivise en strates la population concernée, c'est-à-dire en sous-groupes relativement homogènes qui ont des caractéristiques communes et qui sont mutuellement exclusifs. On note **N_1 l'effectif de la première strate**; **N_2 l'effectif de la deuxième strate**; **N_3 l'effectif de la troisième strate**, etc.

3. On calcule le pourcentage de représentativité de chacune des strates.

Pour la **première strate**, le pourcentage de représentativité est **$P_1 = N_1 \div N \cdot 100 \%$** .

Pour la **deuxième strate**, le pourcentage de représentativité est **$P_2 = N_2 \div N \cdot 100 \%$** .

Pour la **troisième strate**, le pourcentage de représentativité est **$P_3 = N_3 \div N \cdot 100 \%$** .

Etc.



Outils mathématiques suite

4. On choisit un échantillon qui, composé d'autant de strates que la population, est représentatif de la population. La taille de l'échantillon est l'effectif de celui-ci, c'est-à-dire le nombre d'individus ou le nombre d'objets qu'on sélectionne dans la population. Supposons **n** la **taille de l'échantillon** désiré: on détermine **n_1** l'effectif de la **première strate** de l'échantillon, **n_2** l'effectif de la **deuxième strate** de l'échantillon, **n_3** l'effectif de la **troisième strate** de l'échantillon, etc., en multipliant le taux de représentativité calculé précédemment par n .

$$n_1 = P_1 \cdot n$$

$$n_2 = P_2 \cdot n$$

$$n_3 = P_3 \cdot n$$

etc.

5. Dans chacune des strates de la population, on choisit au hasard le nombre d'individus précédemment déterminé. Ces individus formeront chacune des strates de l'échantillon.

L'échantillonnage par grappes

L'**échantillonnage par grappes** consiste à séparer en groupes appelés **grappes** la population concernée par l'étude statistique.

On choisit au hasard un **échantillon** composé d'un **certain nombre de grappes** pour représenter la population totale.

Tous les éléments qui constituent ces grappes, et uniquement ceux-là, font partie de l'échantillon.

Exemple

On veut réaliser une étude statistique portant sur les activités sportives pratiquées par les élèves des écoles secondaires au Québec. Le Québec compte environ 550 écoles secondaires publiques et privées. Quelle est la méthode d'échantillonnage la plus appropriée pour réaliser cette étude ?

Il serait trop coûteux de faire un sondage à l'échelle de la province, compte tenu du nombre considérable d'élèves inscrits au secondaire et de la dispersion géographique des écoles.

L'**échantillonnage par grappes** est la méthode la plus appropriée. Chaque **école** constitue une **grappe**. On sélectionnera au hasard quelques écoles, et on réalisera l'étude sur chacun des élèves de ces écoles.

La méthode d'échantillonnage par grappes présente quelques avantages :

Il n'est pas nécessaire de connaître tous les membres de la population ;

Elle permet de réduire les coûts de déplacement, par exemple quand la population est géographiquement très dispersée.

Il est important de ne pas confondre les strates et les grappes. Dans cet exemple, avec un échantillonnage stratifié, on aurait dû sélectionner quelques élèves dans chacune des écoles du Québec, alors qu'avec un échantillonnage par grappes, on sélectionne tous les élèves de quelques écoles.

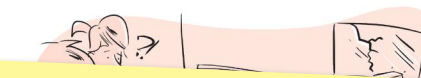
Si on appliquait cette théorie?

- VOICI QUELQUES EXEMPLES DE SITUATIONS QUI VOUS PERMETTRONT DE VOUS FAMILIARISER AVEC LES DIFFÉRENTES MÉTHODES D'ÉCHANTILLONNAGE QUE VOUS AVEZ ÉTUDIÉES.

Exemple 1

2 000 élèves fréquentent une polyvalente.

Le comité étudiant veut connaître l'opinion des élèves au sujet de la qualité, de la quantité et de la diversité des équipements du gymnase dans leur école.



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



Voici la répartition des élèves par niveau:

Niveau	1 ^{re} secondaire	2 ^e secondaire	3 ^e secondaire	4 ^e secondaire	5 ^e secondaire
Nombre d'élèves	550	450	350	400	250

N'ayant pas les ressources suffisantes pour interroger l'ensemble des élèves qui fréquentent l'école, le comité décide de choisir un échantillon représentatif de 320 élèves.

Déterminer le nombre d'élèves de chaque niveau que le comité étudiant doit choisir.

Solution

Les élèves étant répartis en cinq niveaux, on privilégie donc la méthode d'**échantillonnage stratifié**.

- 1) Le nombre d'élèves concernés par l'étude statistique **$N = 2\ 000$** .
- 2) Les élèves sont répartis en **5 strates** homogènes: chacun des niveaux du secondaire.
- 3) On calcule le pourcentage de représentativité de chacun des niveaux.

1^{re} secondaire: $P_1 = (550 \div 2\ 000) \times 100 \%$
 $P_1 = 27,5 \%$

2^e secondaire: $P_2 = (450 \div 2\ 000) \times 100 \%$
 $P_2 = 22,5 \%$

3^e secondaire: $P_3 = (350 \div 2\ 000) \times 100 \%$
 $P_3 = 17,5 \%$

4^e secondaire: $P_4 = (400 \div 2\ 000) \times 100 \%$
 $P_4 = 20 \%$

5^e secondaire: $P_5 = (250 \div 2\ 000) \times 100 \%$
 $P_5 = 12,5 \%$

4) La taille de l'échantillon est $n = 320$

L'échantillon doit être constitué de 5 strates (niveaux) dont les effectifs sont :

1^{re} secondaire : $n_1 = (27,5 \div 100) \times 320$

$n_1 = 88$

2^e secondaire : $n_2 = (22,5 \div 100) \times 320$

$n_2 = 72$

3^e secondaire : $n_3 = (17,5 \div 100) \times 320$

$n_3 = 56$

4^e secondaire : $n_4 = (20 \div 100) \times 320$

$n_4 = 64$

5^e secondaire : $n_5 = (12,5 \div 100) \times 320$

$n_5 = 40$

5) L'échantillon choisi doit être constitué de la façon suivante :

88 élèves de **1^{re} secondaire** ;

72 élèves de **2^e secondaire** ;

56 élèves de **3^e secondaire** ;

64 élèves de **4^e secondaire** ;

40 élèves de **5^e secondaire**.

Exemple 2

À tous les 20^{es} clients qui s'y abonnent pour une période de trois mois ou plus, le gymnase de votre quartier remet un coupon-rabais de 40 % applicable à tout achat de boisson énergétique pour sportifs, en vente au gymnase.

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



Déterminer la méthode d'échantillonnage selon laquelle les coupons-rabais sont distribués à la clientèle du gymnase de votre quartier. Justifier votre réponse.

Solution

De quelle méthode d'échantillonnage s'agit-il? _____

Justification :

Il s'agit, bien sûr, d'un **échantillonnage systématique**. En effet, un individu est sélectionné à tous les intervalles de 20 personnes.

Exemple 3

La province de Québec compte environ 193 centres d'éducation pour adultes. On veut réaliser une étude statistique pour connaître l'opinion des élèves sur le français dispensé dans ces centres.

Troisième exemple:
Encore + de pratique!

Déterminer la méthode d'échantillonnage la plus appropriée pour cette étude. Justifier votre réponse.

Solution

La méthode d'échantillonnage la plus appropriée est:

Justification :

La **méthode d'échantillonnage par grappes** est la plus appropriée.

Les centres d'éducation des adultes sont dispersés sur un **grand territoire** et le nombre d'élèves qui les fréquentent sont très nombreux. On s'assure de **réduire les coûts** et les délais de réalisation de l'étude en privilégiant la méthode d'échantillonnage par grappes, c'est-à-dire en choisissant un **échantillon constitué de quelques centres**.

Dans les **Activités d'apprentissage** qui suivent, vous découvrirez beaucoup d'autres situations dans lesquelles vous aurez la possibilité de déterminer la méthode d'échantillonnage qui est la plus appropriée selon le contexte.

1. Pour chacune des situations qui suivent, déterminer la méthode qui permettra de sélectionner l'échantillon.

Des activités d'apprentissage afin de vous pratiquer à acquérir par étapes la ou les compétences disciplinaires.



a) Dans une épicerie de quartier, on invite les clients à déposer les coupons de caisse dans un grand récipient. À la fin de chaque jour on pige trois coupons de caisse, et les clients gagnent une somme équivalente au montant de leur facture.



b) Le service de contrôle de la qualité d'une entreprise de fabrication de cuisinières électriques sélectionne, à intervalle fixe de 50 unités, un appareil prêt à l'emploi afin d'en vérifier le fonctionnement.

De l'espace fourni afin de vous faciliter la tâche en écrivant à même le module! Aucune feuille volante!



c) Au lotto 6/49, le boulier de *Loto-Québec* éjecte, un à un, les 6 numéros qui vont constituer la combinaison gagnante.

d) Un parti politique provincial demande à une firme de sondage de réaliser une étude statistique qui calculerait ses chances d'obtenir une majorité aux prochaines élections. Les données seront recueillies par une douzaine de téléphonistes et doivent être représentatives de l'ensemble de la population.

Une mention tout au bas vous indique à quelle page vous trouverez le corrigé afin de vous vérifier.



1.11. Vue d'ensemble : synthèse des savoirs

Nous arrivons à la fin du chapitre traitant de statistique. Avant de vous attaquer 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

Les méthodes d'échantillonnage

Les méthodes d'échantillonnage sont : l'échantillonnage **aléatoire**, l'échantillonnage **stratifié** et l'échantillonnage **par grappes**.

- Constituer un **échantillon aléatoire** consiste à choisir au **hasard** des individus dans la population visée. On doit s'assurer que chaque individu de la population a une chance de faire partie de l'échantillon.
- La méthode d'**échantillonnage systématique** ou méthode d'**échantillonnage à intervalle fixe** consiste à sélectionner à **intervalle fixe**, sur une liste ordonnée contenant tous les individus ou tous les objets de la population concernée par l'étude, les individus ou les objets de l'échantillon.
- La méthode d'**échantillonnage stratifié** est utilisée quand la population **n'est pas homogène**, c'est-à-dire qu'une ou plusieurs caractéristiques des individus peuvent influencer les résultats du sondage. L'**échantillonnage stratifié** consiste à diviser la population en **strates**, c'est-à-dire en des **sous-ensembles homogènes et mutuellement exclusifs**. On prélève ensuite un nombre d'individus de chaque strate **proportionnel** au nombre d'individus dans la strate correspondante de la population. Ces individus forment un échantillon représentant toutes les strates de la population visée par l'étude.
- L'**échantillonnage par grappes** permet de séparer la population concernée par l'étude statistique en groupes appelés **grappes**. On choisit au hasard un **échantillon** composé d'un **certain nombre de grappes** pour représenter la population totale. Seuls les éléments qui composent les grappes choisies font partie de l'échantillon.

Le tableau à données condensées

Un **tableau à données condensées** est un tableau qui présente l'effectif ou la fréquence de chacune des valeurs d'une distribution de données.

Un tableau à données condensées contient **une seule valeur par ligne**, sauf, éventuellement, la dernière ligne, qui peut être le total de chaque colonne.

Un tableau à données condensées comporte toujours :

Un **titre** qui indique le thème abordé dans l'étude ;

Des **sous-titres** qui identifient le contenu de chaque colonne ;

Deux colonnes :

dans la **première colonne**, les **valeurs** observées ;

dans la **deuxième colonne**, l'**effectif** (ou la fréquence) correspondant à chacune des modalités ou des valeurs.

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





Résumé des savoirs mathématiques *suite*

Le tableau à données groupées en classes

Un **tableau à données groupées en classes** est utilisé lorsque le caractère considéré est quantitatif et qu'il y a un très grand nombre de données différentes.

Une **classe** est un intervalle généralement fermé-ouvert. Pour décrire une classe, on utilise les **crochets**.

La largeur d'un intervalle se nomme l'**amplitude** de la classe.

Toutes les classes ont la **même amplitude**, sauf la première et la dernière, qui peuvent être des **classes ouvertes**, c'est-à-dire de longueur indéfinie.

Un **tableau à données groupées en classes** contient **plusieurs données par ligne**. Il contient au moins deux colonnes: une colonne contient les **classes** et une autre, l'**effectif** de chacune de ces classes. Au besoin, on ajoute une colonne pour la fréquence relative, une colonne pour la fréquence cumulée et une colonne pour la fréquence relative cumulée de chacune des classes.

Un **tableau à données groupées en classes** comporte toujours:

Un **titre** qui indique le thème abordé dans l'étude;

Deux colonnes:

dans la **première colonne**, les **classes** représentées par des intervalles de valeurs;

dans la **deuxième colonne**, l'**effectif** (ou la fréquence) de chacune des classes.

Pour construire un tableau à données groupées en classes:

On **donne** un **titre** au tableau et aux colonnes;

On **localise** la donnée la **plus grande** et la donnée la **plus petite**;

On **choisit** le **nombre de classes** N . En général, le nombre de classes varie entre 5 et 10;

On **détermine l'étendue**, notée E , qui représente la différence entre la plus grande valeur et la plus petite valeur:

$$E = \text{plus grande valeur} - \text{plus petite valeur}$$

On **détermine l'amplitude des classes** par la formule: Amplitude d'une classe = $E \div N$;

On **détermine** les **intervalles** et l'**effectif** correspondant à chacune des classes;

Les classes doivent couvrir toutes les valeurs possibles. La **première classe** doit obligatoirement contenir la **plus petite valeur** et la **dernière classe** la **plus grande valeur**;

On **indique** les **classes** dans la première colonne du tableau et l'**effectif** de chacune des classes dans la deuxième colonne;

On **indique** la **source** et la **date de publication**, s'il y a lieu;

On **indique**, s'il y a lieu, les **unités utilisées** (\$, €, %) et leur **ordre de grandeur** (milliards, milliers, etc.)



Résumé des savoirs mathématiques *suite*

L'histogramme

L'**histogramme** est un graphique qui sert à représenter une distribution dont les données sont groupées en classes.

Pour représenter par un histogramme une distribution dont les données sont groupées en classes, on suit les étapes suivantes :

On **indique le titre du diagramme** ;

On **trace** les axes et on les **identifie** ;

On **gradue l'axe horizontal** en y indiquant les limites de chacune des classes.

On utilise, au besoin, une rupture d'axe ;

On **gradue l'axe vertical** de façon à couvrir tous les effectifs ;

Entre les limites de chaque classe, on **élève un rectangle** dont la hauteur est égale à l'effectif ou à la fréquence de la classe.

Le mode d'une distribution

Le **mode** d'une distribution, noté M_o , est la modalité ou la valeur la plus fréquente de cette distribution.

Une distribution peut comporter plusieurs modes.

Dans le cas d'une distribution dont les données sont groupées en classes, la classe avec l'effectif le plus élevé porte le nom de **classe modale**. Le **milieu** de la classe modale donne une estimation de la valeur du **mode**.

La médiane d'une distribution

La **médiane**, symbolisée par les lettres M_d , est la valeur qui partage la liste ordonnée des données en deux parties de même effectif.

Lorsque la distribution comporte **un nombre impair (n) de données**, la médiane est la valeur unique qui occupe la position au centre de la distribution.

$$\text{Rang de la médiane} = \frac{(n+1)}{2}$$

Lorsque la distribution comporte **un nombre pair (n) de données**, la médiane est la moyenne des deux valeurs qui occupent les positions du centre de la distribution.

$$\text{Rang de la médiane} = \frac{n}{2} \text{ et } \frac{n}{2} + 1$$

Pour une distribution à **données condensées**, on doit considérer la somme des effectifs pour déterminer le rang de la médiane.

Pour une distribution à **données groupées en classes**, la classe comportant la médiane est appelée **classe médiane**. On estime la valeur de la médiane en calculant la moyenne des limites de la classe médiane.

Résumé des savoirs mathématiques suite

La moyenne d'une distribution

La **moyenne**, représentée par le symbole \bar{x} , est la mesure de tendance centrale la plus couramment utilisée.

Elle correspond à la **somme des valeurs** de toutes les données, **divisée** par le **nombre total de données** :

$$\bar{x} = \frac{\sum x_i}{n}$$

où $\sum x_i$ désigne la somme de toutes les données
et n désigne le nombre total de données

Pour calculer la moyenne d'une distribution à partir d'un **tableau à données condensées**, on doit tout d'abord multiplier chacune d'elles par sa fréquence. On obtient ainsi le **produit de fréquences**. On effectue ensuite la **somme de ces produits** et on **divise** le résultat par le **nombre total** de données pour obtenir la moyenne :

$$\bar{x} = \frac{\sum f_i \cdot x_i}{n}$$

où $\sum f_i \cdot x_i$ désigne la somme de tous les produits des données par leur fréquence
et n désigne le nombre total de données

Pour estimer la moyenne d'une distribution dont les **données sont groupées par classes**, on calcule le **milieu de chaque classe**, puis on procède comme dans le cas d'un tableau à données condensées, en calculant la somme des produits de fréquences, qu'on divise enfin par le nombre de données :

$$\bar{x} = \frac{\sum f_i \cdot m_i}{n}$$

où $\sum f_i \cdot m_i$ désigne la somme de tous les produits des milieux des classes par leur fréquence et n désigne le nombre total de données

La **moyenne pondérée** est utilisée dans le cas où les valeurs n'ont pas la même importance, on donne alors une pondération à chacune des valeurs. Toutes les pondérations doivent avoir une somme de 100 %. Dans ce cas, on calcule la moyenne pondérée de la façon suivante :

Moyenne pondérée = Somme des produits des valeurs par leur pondération

Choix de la mesure de tendance centrale la plus représentative

Les mesures de tendance centrale d'une distribution statistique sont : la **moyenne**, la **médiane** et le **mode**. Dans certains cas, l'une des trois mesures de tendance centrale est plus appropriée pour décrire l'ensemble des données d'une distribution.

- La **moyenne** est la mesure de tendance centrale la plus employée, car elle se prête bien à la généralisation des résultats d'une étude. Elle comporte cependant certaines limites :
 - On ne peut calculer la moyenne de données à caractère qualitatif ;
 - On ne peut calculer la moyenne d'une distribution comportant une ou des classes ouvertes ;
 - Lorsque la distribution comporte des données extrêmes qui s'éloignent beaucoup de l'ensemble des données, la moyenne est moins représentative que la médiane.
- La **médiane** est un bon complément de la moyenne. Utilisée moins couramment, elle offre néanmoins une solution de rechange valable lorsque la distribution comporte des données dites « aberrantes », c'est-à-dire des données qui s'éloignent fortement de la concentration des autres données.

On peut calculer la médiane d'une distribution comportant une ou deux **classes ouvertes**, car la classe médiane, située plus au centre, n'est habituellement pas une classe ouverte.

Précisons que cette mesure de tendance centrale convient plutôt à une étude à **caractère quantitatif**. Pour une étude à caractère qualitatif, les données doivent correspondre à une échelle graduée, sinon on ne peut pas calculer sa médiane.



Résumé des savoirs mathématiques suite

- Le **mode** est surtout utilisé dans le cas d'une étude à **caractère qualitatif**. On l'emploie moins lorsque les données présentent un caractère quantitatif, car il ne tient pas compte de toutes les données, mais seulement des plus fréquentes.

Dans une distribution où la fréquence du mode est nettement plus élevée que celle des autres données, le mode pourrait être la mesure à privilégier, selon le contexte.

Les quartiles et les quarts d'une distribution

Les **quartiles** sont des valeurs qui séparent un ensemble de données placées en ordre croissant en quatre parties comprenant le même nombre de données.

Une distribution comporte trois quartiles :

Le **premier quartile**, noté Q_1 , sépare le premier quart des données du reste des données ;

Le **deuxième quartile**, noté Q_2 , coïncide avec la médiane. Il sépare la distribution en deux parties comprenant le même nombre de données ;

Le **troisième quartile**, noté Q_3 , sépare les trois premiers quarts des données du reste des données.

Chacune des quatre parties de l'ensemble de données séparées par les quartiles s'appelle un **quart**.

Les données inférieures à Q_1 forment le **premier quart**.

Les données comprises entre Q_1 et Q_2 forment le **deuxième quart**.

Les données comprises entre Q_2 et Q_3 forment le **troisième quart**.

Les données supérieures à Q_3 forment le **quatrième quart**.

Le diagramme de quartiles

Le **diagramme de quartiles**, aussi appelé **boîte à moustache**, est un diagramme qui résume quelques caractéristiques de position d'une distribution. Le diagramme de quartiles met en évidence le **minimum**, le **maximum** et les **quartiles** d'une distribution. Il permet de tirer des conclusions sur la dispersion et la concentration des données.

Les mesures de dispersion

Les **mesures de dispersion** sont des mesures qui permettent de décrire la manière dont les données d'une distribution se dispersent. Les mesures de dispersion sont l'**étendue des quarts** et l'**étendue interquartile**.

L'**étendue des quarts** est la différence entre les valeurs qui délimitent chacun des quarts :

L'étendue du premier quart = $Q_1 - \text{Min}$;

L'étendue du deuxième quart = $Q_2 - Q_1$;

L'étendue du troisième quart = $Q_3 - Q_2$;

L'étendue du quatrième quart = $\text{Max} - Q_3$.

L'**étendue interquartile** d'une distribution est la distance entre le premier et le troisième quartiles :

$$EI = Q_3 - Q_1$$

Données aberrantes

Dans un diagramme de quartiles, la longueur de chaque partie de la moustache ne doit jamais dépasser **1,5 fois l'étendue interquartile**. Une donnée qui s'éloigne du rectangle de plus de 1,5 fois l'**étendue interquartile** est une **donnée aberrante** ou une **donnée extrême**. Une valeur x est considérée comme une donnée aberrante si :

$$Q_1 - x > 1,5 (Q_3 - Q_1)$$

ou

$$x - Q_3 > 1,5 (Q_3 - Q_1)$$

Dans un diagramme de quartiles, on représente une donnée aberrante par un *.

Consolidation des savoirs

1. Pour chacune des situations qui suivent, déterminer la méthode qui permet de sélectionner l'échantillon.

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



- a) Le ministère de l'Éducation et de l'Enseignement supérieur veut connaître les méthodes les plus pratiquées par les élèves dans les différentes écoles secondaires.

- b) À toutes les demi-heures, une mère prend la température de son enfant malade.

- c) Un infirmier maladroit a laissé tomber les contenants de pilules de ses patients. Pour camoufler sa maladresse, il remet les pilules dans n'importe quel contenant.



- d) Un bateau avec une cargaison de 1 200 000 litres d'huile d'olive extra vierge emballée dans des caisses de 12 bouteilles de 1 litre a accosté au port de Montréal. On veut contrôler la qualité de l'huile de la cargaison.

- e) Une firme fait une étude dans le but de connaître les habitudes de consommation chez les gens selon leurs revenus.

- f) Un facteur distribue un tract publicitaire à toutes les résidences dont le numéro civique est pair.

- g) Un DJ débutant veut connaître les goûts musicaux des individus d'un groupe selon leur âge.



1.12. Situations de vie

Des inventeurs ont développé un micro-ordinateur pour les amateurs d'exercice physique du Québec. Malgré l'absence de financement pour le projet, vous avez proposé votre aide aux inventeurs. Vous leur offrez de réaliser une étude de marché, de dresser le portrait type du client potentiel pour leur produit.

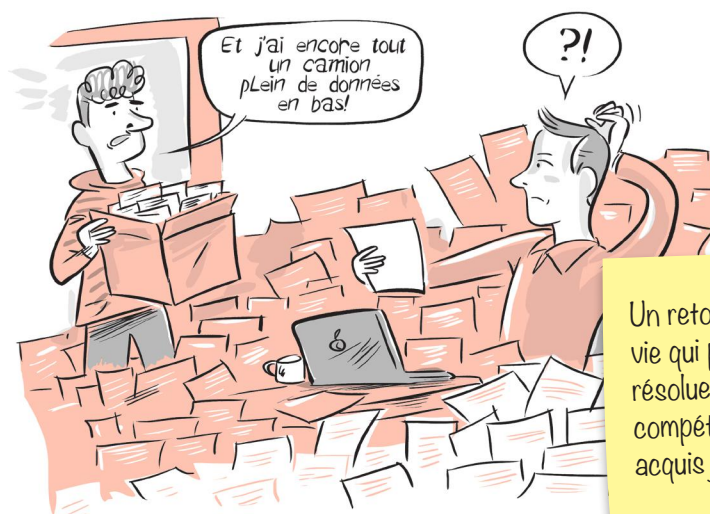
Toutes les personnes qui ont fait connaître leur intérêt pour le micro-ordinateur pour sportifs ont rempli un questionnaire de sondage en ligne portant sur leur sexe, leur âge, leur lieu de résidence, leurs revenus annuels et sur l'activité physique principale qu'ils pratiquent.

On vous a transmis les données de l'étude faite par les développeurs du logiciel et, comme vous l'aviez anticipé, les données sont un fouillis. Vous avez dû faire un tri pour y voir clair.

Afin de dresser un tableau de la situation, vous avez réparti le marché québécois par secteurs géographiques dominés par une ville majeure: la grande région de la ville de Québec, de Trois-Rivières, de Montréal, de Gatineau, etc. Aussi, pour simplifier l'analyse des différentes activités, vous avez réparti les différentes activités physiques en cinq grandes catégories.

Retour à la mise en situation :

ANALYSER ET INTERPRÉTER LES DONNÉES



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

- 1) Activités en groupe
Exemples: danse, aérobic, zumba, danse sociale, danse folklorique, danse spectacle.
- 2) Musculation
Exemples: poids et haltères libres, poids avec appareils.
- 3) Exercices avec appareil
Exemples: rameur, vélo stationnaire, cardiovélo, course à pied intérieur (tapis roulant), elliptique, appareil hybride.
- 4) Exercices extérieurs (seul ou dans une ligue)
Exemples: marche, course à pied, cyclisme, football, soccer, baseball, frisbee.
- 5) Exercices méditatifs
Exemples: yoga, tai-chi, pilates.

Voici, dans un tableau, les données que vous avez classées :

Région	Activité en groupe	Musculation	Exercice avec appareil	Exercice extérieur	Exercice méditatif
Montréal	896	380	256	348	160
Québec	284	192	112	104	68
Gatineau	148	84	52	44	72
Sherbrooke	100	48	60	68	84
Saguenay	48	36	28	36	12
Trois-Rivières	28	12	20	20	0
Drummondville	32	8	20	16	4
Baie-Comeau	20	20	8	8	24
Rivière-du-Loup	12	8	4	4	12

Les personnes qui ont participé à l'étude ont donné une estimation de leurs revenus annuels. Voici les données recueillies, regroupées en classes.

REVENUS ANNUELS DES PERSONNES QUI ONT REMPLI LE QUESTIONNAIRE

Revenus annuels (milliers de \$)	Effectif
[0, 20[140
[20, 40[980
[40, 60[1 760
[60, 80[650
[80, 100[230
[100, 120[150
120 ou plus	90

Les données sur l'âge et sur le sexe figurent dans le tableau suivant.

ÂGE DES PERSONNES QUI ONT REMPLI LE QUESTIONNAIRE

Âge (années)	Nombre de femmes	Nombre d'hommes
[20, 30[880	330
[30, 40[1 410	390
[40, 50[600	200
[50, 60[105	60
[60, 70[5	15
[70, 80[0	5

1. Le financement participatif.

Votre rôle consiste à décrire le portrait de la clientèle cible pour l'invention en déterminant les mesures de tendance centrale pour le lieu de résidence, les revenus annuels, l'âge, le sexe et le type d'activité physique que pratiquent les participants à l'étude.

1^{re} tâche

Déterminer les mesures de tendance centrale des revenus annuels des personnes formant la population concernée par cette étude.

Les mesures de tendance centrale sont: la **moyenne**, la **médiane** et le **mode**. On considère, tour à tour, chacune de ces mesures.

a) Estimation de la moyenne

Est-il possible d'estimer la moyenne des revenus annuels des personnes qui ont participé

à l'étude? _____

Si oui, calculez cette moyenne, sinon, dites pourquoi c'est impossible.

b) Estimation de la médiane

Est-il possible d'estimer la médiane des revenus annuels des personnes qui ont participé

à l'étude? _____

Si oui, calculez cette médiane, sinon, dites pourquoi c'est impossible.

c) Estimation du mode

Est-il possible d'estimer le mode des revenus annuels des personnes qui ont participé

à l'étude? _____

Si oui, déterminez ce mode, sinon, dites pourquoi c'est impossible.

Toujours de l'espace pour écrire vos développements tout au long des tâches!



1. L'usine de pâtes alimentaires.

Vous avez récemment été embauché à la gestion de la qualité dans une usine de pâtes alimentaires sèches.

L'usine produit plusieurs kilogrammes de pâtes alimentaires à l'heure qui sont ensuite insérées dans des sachets de 250 g, ou des boîtes de 1 kg ou 2 kg. Le mécanicien a calibré une machine pour qu'elle remplisse des sachets de 250 g.

Votre rôle consiste à vous assurer que la calibration des machines soit la bonne pour que le contenu de chaque sachet ait une masse voisine de celle qui est indiquée sur l'emballage et à détecter toute variation dans le calibrage de la machine.

Pour vérifier l'exactitude du calibrage de la machine, vous prélevez un sachet de pâtes alimentaires à toutes les 50 unités produites. Vous avez pesé individuellement chaque sachet.

Voici les résultats, en grammes, que vous avez obtenus :

249	252	259	247	257	251	250	252	253
255	252	254	255	254	255	254	255	256
256	256	249	245	252	254	242	250	255
256	254	250	249	249	251	243	247	249
247	251	249	247	252	255	258	246	244
243	246	244	253	256	247	242	247	245
245	246	244	254	245	245	252	247	247
247	245	244						

1^{re} tâche

En regard des mesures de tendance centrale de l'échantillon considéré, diriez-vous que la machine est bien calibrée? Justifier votre réponse.

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



Avant de continuer et pour conclure cette première étape

Pour terminer ce chapitre, traitant des **distributions statistiques à un caractère**, et pour vous assurer de bien maîtriser 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.

Présentez une solution claire et complète et ne demandez l'aide de personne. Cela 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 cette étape du MAT 3052, et si vous êtes prêt à aborder la prochaine étape. Étape par étape, vous arriverez à la fin du cours. Avec succès, n'en doutez pas.

Bon travail !

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



01 PREMIÈRE PARTIE

Évaluation des connaissances

1. Représenter les...

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



01 DEUXIÈME PARTIE

Évaluation des compétences

4. L'usine XYZ.

Un représentant...

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. Pour chacune...

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 les probabilités et les statistiques. Faites-en bon usage!

1. L'âge des joueurs.

Un lundi de...

amplitude d'une classe

L'amplitude d'une classe est la largeur de cette classe. Elle correspond à la différence entre la limite supérieure et la limite inférieure de cette classe.

arbre de probabilité

On appelle arbre de probabilité le diagramme qu'on obtient en ajoutant une probabilité sur chacune des branches du diagramme en arbre.

arrangement

Un arrangement est une sélection ordonnée d'un certain nombre d'éléments d'un ensemble.

biais

Une source de biais est une cause qui affecte le résultat d'une étude statistique.

boîte à moustache

La boîte à moustache, aussi appelée diagramme de quartiles, est un diagramme qui résume quelques caractéristiques de position d'une distribution. Le diagramme de quartiles met en évidence le minimum, le maximum et les quartiles d'une distribution. Il nous permet aussi de tirer des conclusions sur la dispersion et la concentration des données.

caractère

Un caractère représente l'aspect des individus de la population étudiée sur lequel porte une étude statistique.

caractère qualitatif

On dit qu'un caractère est qualitatif lorsque les données recueillies ne sont pas des nombres.

caractère quantitatif continu

Un caractère est quantitatif continu si les données recueillies sont des nombres qui peuvent prendre n'importe quelle valeur d'un intervalle.

caractère quantitatif discret

Un caractère est quantitatif discret lorsque les données recueillies peuvent être uniquement des nombres entiers.

1.1. Organisation et interprétation de données statistiques

1. p. 11

- a) Aléatoire
- b) Systématique
- c) Aléatoire
- d) Stratifié
- e) Grappes
- f) Grappes
- g) Systématique
- h) Stratifié
- i) Grappes

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



2. p. 14

- a) Population = 560
 Nombre de strates = 2 (femmes et hommes)
 Nombre de femmes: 350
 Nombre d'hommes: 560 – 350 = 210
 Pourcentage de représentativité des femmes = $(350 \div 560) \times 100 \% = 62,5 \%$
 Pourcentage de représentativité des hommes = $(210 \div 560) \times 100 \% = 37,5 \%$

Comme les pourcentages de représentativité doivent s'appliquer à l'échantillon, ce dernier sera constitué de:

- Nombre de femmes = $(62,5 \div 100) \times 80 = 50$
- Nombre d'hommes = $(37,5 \div 100) \times 80 = 30$

L'échantillon sera constitué de 50 étudiantes et de 30 étudiants.

- b) Nombre d'employés = 36 + 44 + 360 + 260 + 280 + 20 = 1 000
 Nombre de strates homogènes = 6

Pourcentages de représentativité de chaque strate:

- Administration: $(36 \div 1\ 000) \times 100 \% = 3,6 \%$
- Médecins: $(44 \div 1\ 000) \times 100 \% = 4,4 \%$
- Infirmières: $(360 \div 1\ 000) \times 100 \% = 36 \%$
- Infirmières auxiliaires: $(260 \div 1\ 000) \times 100 \% = 26 \%$
- Préposés aux bénéficiaires: $(280 \div 1\ 000) \times 100 \% = 28 \%$
- Autres: $(20 \div 1\ 000) \times 100 \% = 2 \%$

Composition de l'échantillon:

Taille de l'échantillon = 250

- Administration: $(3,6 \div 100) \times 250 = 9$
- Médecins: $(4,4 \div 100) \times 250 = 11$
- Infirmières: $(36 \div 100) \times 250 = 90$
- Infirmières auxiliaires: $(26 \div 100) \times 250 = 65$
- Préposés aux bénéficiaires: $(28 \div 100) \times 250 = 70$
- Autres: $(2 \div 100) \times 250 = 5$

L'échantillon sera composé de 9 employés de l'administration, de 11 médecins, de 90 infirmières, de 65 infirmières auxiliaires, de 70 préposés aux bénéficiaires et de 5 employés affectés à d'autres postes.

- c) 1)

NOMBRE D'HABITANTS EN ÂGE DE VOTER

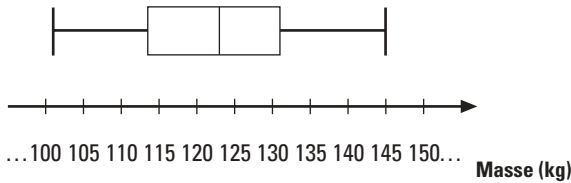
Quartier	Nombre d'habitants en âge de voter	Nombre de femmes	Nombre d'hommes
La Tulipe	1 200	800	400
Le Narcisse	3 400	1 400	2 000
Le Frêne	2 700	1 800	900

16. p. 107 suite

c) 101, 102, 105, **112 | 115**, 120, 121, **122 | 124**, 125, 128, **130 | 132**, 140, 140, 145

$Q_1 = 113,5$ $Q_2 = 123$ $Q_3 = 131$

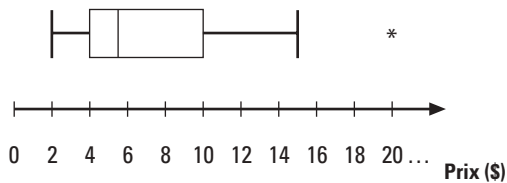
**MASSE DE CERTAINS CANDIDATS
RETENUS POUR UNE ÉTUDE**



d) 2, 2, 2, 2, 3, 3, **4**, 5, 5, 5, 5, 5, **5 | 6**, 6, 6, 6, 8, 8, **10**, 10, 12, 12, 15, 15, 20

$Q_1 = 4$ $Q_2 = 5,5$ $Q_3 = 10$

**PRIX DES ARTICLES VENDUS
LORS DE LA VENTE DE GARAGE
DE LA FAMILLE LACASSE**



1.11. Vue d'ensemble: synthèse des savoirs

1. p. 114

- a) Échantillonnage par grappes
- b) Échantillonnage systématique
- c) Échantillonnage aléatoire
- d) Échantillonnage par grappes
- e) Échantillonnage par grappes
- f) Échantillonnage systématique
- g) Échantillonnage aléatoire

2. p. 115

Le nombre d'employés concerné par l'étude statistique: $N = 5\,000$
Nombre de strates: 3 strates (catégories d'employés)

Pourcentage de représentativité de chacune des strates:

Employés sur des chaînes de production: $P_1 = 4\,000 \div 5\,000 \times 100\%$
 $P_1 = 80\%$

Employés du développement: $P_2 = 600 \div 5\,000 \times 100\%$
 $P_2 = 12\%$

Employés de la gestion: $P_3 = 400 \div 5\,000 \times 100\%$
 $P_3 = 8\%$

Taille de l'échantillon choisi: $n = 400$

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



1.12. Situations de vie

1. Le financement participatif.

p. 133

1^{re} tâche

a) Estimation de la moyenne

Est-il possible de calculer la moyenne des revenus annuels de

On ne peut calculer la moyenne des revenus annuels, car il y a une classe ouverte.

b) Estimation de la médiane

Est-il possible de calculer la médiane des revenus annuels des personnes qui ont participé à l'étude? **Oui.**

$$Md = \frac{40\,000 + 60\,000}{2}$$

$$Md = 50\,000$$

La médiane des revenus annuels est de 50 000 \$.

c) Estimation du mode

Est-il possible de déterminer le mode des revenus annuels des personnes qui ont participé à l'étude? **Oui.****La classe modale est [40 000, 60 000[**

$$Mo = (40\,000 + 60\,000) \div 2$$

$$Mo = 50\,000$$

Le mode des revenus annuels est de 50 000 \$.2^e tâche

a) Estimation de la moyenne

Est-il possible de calculer l'âge moyen des personnes qui ont participé à l'étude? **Oui.**

$$\bar{x} = \frac{\sum f_i \cdot m_i}{n}$$

$$\bar{x} = [(880 + 330) \cdot 25 + (1\,410 + 390) \cdot 35 + (600 + 200) \cdot 45 + (105 + 60) \cdot 55 + (5 + 15) \cdot 65 + (0 + 5) \cdot 75] \div 4\,000$$

$$\bar{x} = 140\,000 \div 4\,000$$

$$\bar{x} = 35$$

L'âge moyen est de 35 ans.

b) Estimation de la médiane

Est-il possible de calculer l'âge médian des personnes qui ont participé à l'étude? **Oui.**

$$Md = \frac{30 + 40}{2}$$

$$Md \approx 35$$

L'âge médian est de 35 ans.

c) Estimation du mode

Est-il possible de déterminer l'âge modal des personnes qui ont participé à l'étude? **Oui.****La classe modale est [30, 40[, et le mode est: (30 + 40) ÷ 2, soit 35 ans.**

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

étude? **Non.**

1. L'usine de pâtes alimentaires.

p. 142

1^{re} tâche**MASSE DES SACHETS DE PÂTES ALIMENTAIRES
D'UN ÉCHANTILLON**

Masse	Effectif
242	2
243	2
244	4
245	6
246	3
247	9
249	6
250	3
251	3
252	6
253	2
254	6
255	6
256	5
257	1
258	1
259	1
Total	66

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

**Calcul de la moyenne :**

$$\bar{x} = \frac{\sum f_i \cdot x_i}{n}$$

$$\bar{x} = (2 \cdot 242 + 2 \cdot 243 + 4 \cdot 244 + 6 \cdot 245 + 3 \cdot 246 + 9 \cdot 247 + 6 \cdot 249 + 3 \cdot 250 + 3 \cdot 251 + 6 \cdot 252 + 2 \cdot 253 + 6 \cdot 254 + 6 \cdot 255 + 5 \cdot 256 + 1 \cdot 257 + 1 \cdot 258 + 1 \cdot 259) \div 66$$

$$\bar{x} = 16\,500 \div 66$$

$$\bar{x} = 250$$

Calcul de la médiane :

Il y a 66 données. La médiane est la moyenne des données qui occupent les 33^e et 34^e rangs de la distribution dont les données sont classées par ordre croissant. La médiane est donc de 250 g.

Calcul du mode :

La donnée la plus fréquente est 247 g.

La machine semble bien calibrée, puisque les mesures de tendance centrale donnent des résultats très proches du résultat attendu :

$$\bar{x} = 250 \text{ g}, Md = 250 \text{ g} \text{ et } Mo = 247 \text{ g}.$$

MOTS	CHAPITRE 1	CHAPITRE 2
Amplitude d'une classe	28, 29, 30, 32, 33, 39, 110	
Arbre de probabilité		216
Arrangement		19 24
Boîte à moustache	97, 100, 102, 113	
Caractère	17, 18, 19, 28, 81, 83, 110, 112, 113	
Caractère qualitatif	17, 81, 83, 112, 113	
Caractère quantitatif continu	17	
Caractère quantitatif discret	17	
Classe	28, 29, 30, 32, 33, 38, 39, 40, 41, 48, 50, 58, 59, 63, 64, 70, 71, 81, 85, 110, 111, 112	
Classe modale	48, 50, 85, 111	
Classe ouverte	28, 81, 85, 110, 112	
Combinaison		195, 196, 197, 198, 199, 217, 248
Deuxième quart	91, 92, 93, 97, 98, 113	
Deuxième quartile	91, 97, 100, 113	
Diagramme de quartiles	97, 98, 99, 100, 101, 102, 103, 113	
Diagramme de Venn		205, 206, 207, 209, 210, 211, 248
Diagramme en arbre		181, 182, 183, 216, 248, 249

Une table alphabétique des mots clés et leurs références.



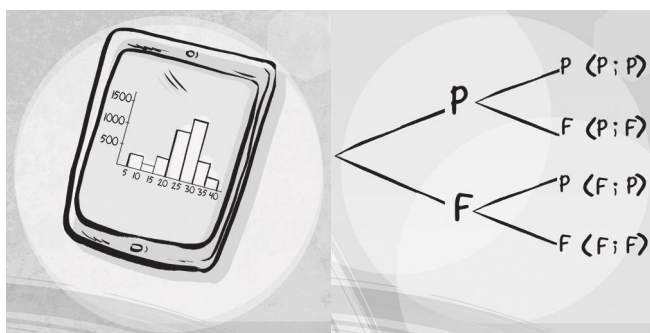
À propos de l'illustrateur et des illustrations...

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.



Une page est consacrée à l'illustrateur afin de vous le présenter.

KINÉSIS
ÉDUCATION

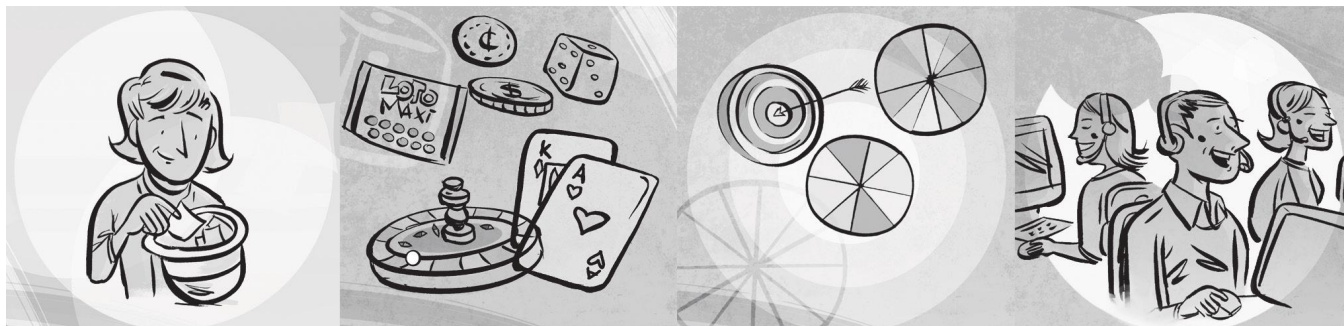


En 2009, il était l'un des bédéistes invités au festival *BoomFest* de Saint-Pétersbourg, en Russie. Il a illustré entre autres le générique de la télésérie *La Galère* à Ici Radio-Canada. En 2016, il a participé au projet *Correspondances* de Lyon.

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.

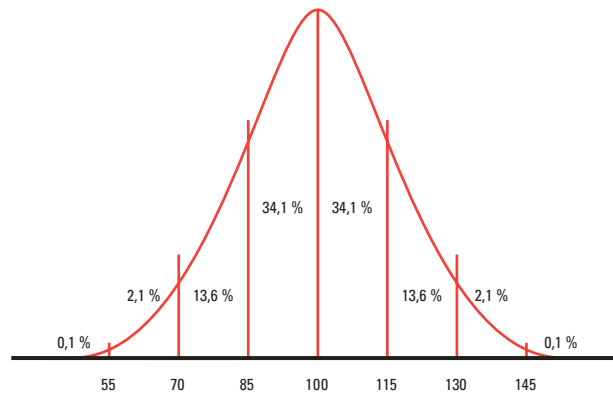
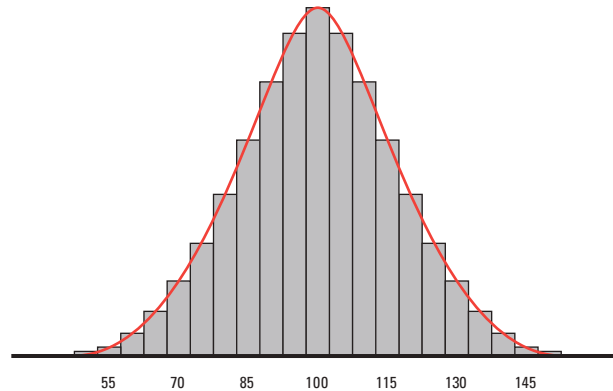


Vous voulez en savoir plus sur Paul Bordeleau ?
Voici ses coordonnées : www.paulbordeleau.com

Pour en savoir un peu plus...

Loi normale et QI

Si on trace l'histogramme des quotients intellectuels de la population, et qu'on considère des classes de plus en plus petites, on obtient une courbe qui a la forme d'une cloche.



Pour les curieux,
un prolongement
des connaissances
et de l'enrichissement.

Le centre de la courbe est 100. Cette valeur est, à la fois le QI moyen et le QI médian. La moitié des gens ont un QI situé au-dessus de 100 et l'autre moitié en dessous de 100. La majorité de la population a un QI voisin de la moyenne : 68,2 % des gens ont un QI entre 85 et 115. À l'extrémité droite de la courbe se trouvent les personnes avec un QI nettement au-dessus de la moyenne (145 ou plus) : seulement 0,1 % de la population se trouvent dans cette catégorie. La courbe étant symétrique, on trouve également 0,1 % de la population dont le QI est inférieur à 55.

De nombreux phénomènes peuvent être représentés graphiquement par une courbe en forme de cloche, si bien qu'on appelle cette courbe la courbe normale ou courbe de Gauss, en l'honneur du mathématicien allemand Carl Friedrich Gauss (1777 – 1855) à qui en appartient la découverte.

Les petits plus...



La statistique à travers les âges

À l'origine, l'utilité de la statistique se limitait au recensement des troupeaux, la plupart du temps pour déterminer le montant à payer en impôt. Puis vint l'ère où on recensait les forces armées pour évaluer la force d'un pays.

Ce n'est qu'au XVIII^e siècle, avec l'arrivée des premières tables de mortalité que la statistique sert à faire des prévisions, et que naissent les compagnies d'assurance. Ne vous êtes-vous jamais demandé comment votre compagnie d'assurance calcule le montant de la prime qu'il vous faudra payer pour prendre une assurance sur votre vie, sur votre maison, sur votre voiture, etc. ? En effet, les compagnies d'assurance utilisent grandement les données statistiques pour faire des affaires. Calculer le facteur de risque de se faire voler sa voiture selon l'endroit où on habite, ou le risque que sa demeure brûle si on est un fumeur ne sont que des exemples de valeurs qui permettent aux compagnies d'assurances d'établir leurs tarifs.

De nos jours, la statistique sert à toutes les sauces, que ce soit dans le monde politique, dans le monde sportif, et même dans le monde virtuel. N'avez-vous jamais fait une recherche d'un produit en ligne, et aperçu, sans aucune raison, une publicité sur un produit connexe à celui que vous aviez cherché ? L'affichage de cette publicité sur votre écran n'est pas le fruit du hasard : on a enregistré les données afin de connaître les produits qui, statistiquement, ont intéressé d'autres clients comme vous.

Un peu d'histoire
pour mieux comprendre
les mathématiques.



Le café

Récemment, un de vos amis, qui est barista au café du coin, vous a raconté sa journée de travail. Un client a commandé un café personnalisé, un latté avec double sirop de caramel, crème fouettée et copeaux de chocolat. Malheureusement, le client est parti subitement et votre ami est demeuré avec la commande. Vous lui demandez pourquoi il n'a pas vendu le café à un autre client, mais il se met à rire et vous dit qu'avec toutes les possibilités de cafés différents, et tous ses ingrédients, il peut s'écouler des mois avant qu'un autre client commande le même café.

Incrédule, vous vous êtes rendu au café où travaille votre ami lors de son jour de congé, et vous avez commandé un café.

- Un café, s'il-vous-plaît.
- Un café americano, un café latté, un café moka, un cappuccino, un macchiato au caramel, un espresso, un espresso macchiato, un café viennois ou un café frappé ?
- Un café moka.
- De quelle origine ? Brésil brun, Costa Rica brun, Colombie Supremo brun, Guatemala brun, Mexique brun, Mexique noir, Pérou brun, Pérou noir, Jamaïque Blue Mountain brun, Éthiopie Harrar Longberry brun, Kenya AA noir, Tanzanie Kilimandjaro AA mi-noir, Java Estate brun, Sumatra Mountain brun ou Sumatra ?
- Costa Rica Brun.
- Avec quel sirop ? Vanille, caramel, caramel brûlé, moka à la menthe poivrée, marron praliné, noisette, moka ou cannelle ?
- Caramel.
- Avec quelle garniture ? Chocolat blanc, chocolat brun, chocolat noir, double chocolat, fraise séchée ou framboise séchée ?
- Chocolat noir.
- Avec quel type de lait ou de crème ? Lait écrémé, lait 2 %, lait 3,25 %, crème 10 %, lait de soya, lait d'amande, lait de soya à la vanille ou crème fouettée ?
- Lait écrémé.
- Régulier, décaféiné ou moitié-moitié ?
- Régulier.
- Petit, moyen ou grand ?
- Moyen.



Pendant que le barista prépare votre café, vous commencez à faire des calculs : si on vend un café à la minute, 24 heures sur 24, vous vous demandez combien de temps il faudrait pour que toutes les combinaisons possibles du menu soient servies. Sauriez-vous trouver cette durée ?

On peut s'amuser
en faisant
des mathématiques !
Et son corrigé !

Au moment où vous pensiez en avoir fini avec la commande de votre café, le barista revient et vous demande :

- Pour ici ou pour emporter ?
- !!!

Amusons-nous / page 204**Le café**

Type de café: 9

Origine du café: 15

Sirop: 8

Garniture: 6

Lait ou crème: 8

Caféine ou non: 3

Taille: 3

Nombre de possibilités: $9 \times 15 \times 8 \times 6 \times 8 \times 3 \times 3 = 466\,560$ $466\,560 \div 60 = 7\,776$ heures $7\,776 \div 24 = 324$ jours $324 \div 30 = 10,8$ mois environ**En vendant un café à la minute, il faut environ 11 mois pour épuiser toutes les possibilités.****Amusons-nous / page 231****Les coïncidences qui n'en sont pas**

La probabilité que tous aient des dates d'anniversaire différentes est:

$$P(A) = \frac{365 \times 364 \times 363 \times \dots \times 341}{365 \times 365 \times 365 \times \dots \times 365} \approx 0,43$$

L'événement B est le complémentaire de l'événement A:

$$P(B) = 1 - P(A)$$

$$P(B) = 1 - 0,43$$

$$P(B) = 0,57$$

Contre toute attente, l'événement B est le plus probable.

Le MAT 3052

Vise l'acquisition de deux grandes compétences transversales: exercer son jugement critique et communiquer de façon appropriée. Au moyen de trois procédés intégrateurs: la réalisation d'une collecte de données; la comparaison de collectes de données; l'interprétation de données issues d'une expérience.

ÉDITION
AOÛT
2022

MAT 3052 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.

ISBN 978-2-7615-0899-5



9 782761 508995

400 1335

Florence Grandchamp
Drita Neziri
Abdelkader Amara
Raymond Thériault

ÉDITION
AOÛT
2022

COLLECTE DE DONNÉES

MAT A 3052 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



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?*