

1. But du cours

Le but du cours Collecte de données est de rendre l'adulte apte à traiter des situations qui requièrent la collecte ou le traitement de données exprimées sous forme de distribution à un caractère.

Au terme de ce cours, l'adulte sera en mesure d'effectuer une collecte de données. Il pourra aussi comparer les résultats d'une expérience statistique à l'aide de divers instruments pour valider ses observations relativement à un problème qu'il a lui-même cerné. La présentation des résultats de son analyse sera faite dans le respect des règles et des conventions mathématiques. Il utilisera des stratégies de résolution de situations-problèmes afin de déterminer la solution la plus juste. De plus, il interprétera, à l'aide du raisonnement mathématique, des données probabilistes issues d'une expérience aléatoire et il prendra des décisions.

2. Savoirs prescrits

Procédés intégrateurs

En vue de traiter efficacement les situations d'apprentissage proposées dans ce cours, l'adulte développe 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.

Ces procédés, mis en valeur dans les situations d'apprentissage du présent cours, favorisent l'intégration des savoirs mathématiques et des compétences disciplinaires. Les situations d'apprentissage traitées doivent toucher à l'un ou l'autre de ces procédés intégrateurs. Toutefois, l'ensemble des situations choisies doit être assez vaste pour couvrir les trois procédés.

Savoirs mathématiques

Savoirs mathématiques	Limites et précision
<p>Distributions statistiques à un caractère</p> <ul style="list-style-type: none"> • Organisation et interprétation de données statistiques 	<p>Les méthodes d'échantillonnage à l'étude dans ce cours sont :</p> <ul style="list-style-type: none"> • l'échantillon stratifié • l'échantillon par grappes
<p>Distributions statistiques à un caractère (suite)</p> <ul style="list-style-type: none"> • Construction et interprétation de tableaux de distribution 	<p>L'interprétation de données et la construction des tableaux à l'étude dans ce cours se réalisent à l'aide :</p> <ul style="list-style-type: none"> • d'un tableau à données condensées • d'un tableau à données groupées en classes
<p>Distributions statistiques à un caractère (suite)</p> <ul style="list-style-type: none"> • Représentation et interprétation de graphiques 	<p>Les représentations graphiques à l'étude dans ce cours sont :</p> <ul style="list-style-type: none"> • l'histogramme • le diagramme de quartiles
<p>Distributions statistiques à un caractère (suite)</p> <ul style="list-style-type: none"> • Calcul de mesures de tendance centrale et de dispersion 	<p>Les mesures de tendance centrale à l'étude dans ce cours sont :</p> <ul style="list-style-type: none"> • le mode • la médiane • la moyenne pondérée <p>La mesure de dispersion à l'étude dans ce cours se limite à l'étendue des quarts (y compris l'étendue interquartile).</p>
<p>Probabilité</p> <ul style="list-style-type: none"> • Dénombrement et calcul de probabilités 	<p>Les variables aléatoires à l'étude dans ce cours sont de deux types :</p> <ul style="list-style-type: none"> • discrète • continue <p>Le dénombrement et le calcul de probabilités peuvent être faits dans des situations variées, incluant des contextes de mesure (dont les probabilités géométriques).</p> <p><i>Les calculs (arrangement, permutation et combinaison) se faisant par raisonnement, il n'est pas nécessaire de recourir à des formules de dénombrement.</i></p>
<p>Probabilité (suite)</p> <ul style="list-style-type: none"> • Représentation d'événements 	<p>Les représentations d'événements se font à l'aide :</p> <ul style="list-style-type: none"> • de tableaux • d'arbres • de diagrammes • de figures géométriques

<p>3052-2</p> <p>4104 (+++) 2008 (+)</p> <p>N : • Échantillonnage par grappes • Représentation d'événements à l'aide de figures géométriques</p>

1. L'échantillonnage stratifié et par grappes

Un échantillon est représentatif d'une population s'il possède les mêmes caractéristiques que cette population. Pour former un échantillon, chaque élément d'une population doit avoir la même chance d'être choisi.

Voici quatre méthodes pour former un **échantillon représentatif**.

Méthode d'échantillonnage		Exemple :
Aléatoire	Méthode qui consiste à choisir au hasard les éléments qui forment l'échantillon	On mène un sondage pour connaître le coût d'achat des livres des 3 000 élèves d'un cégep qui offre 105 programmes d'études. On met dans une boîte tous les noms des élèves écrits sur des bouts de papier et on tire 300 noms au hasard pour former l'échantillon.
Systematique	Méthode qui consiste, à l'aide d'une liste de tous les éléments d'une population, à choisir chaque n ^e élément suivant un premier élément choisi au hasard pour former l'échantillon	Au hasard, on choisit la 8 ^e personne sur la liste complète des élèves, puis chaque 10 ^e personne subséquente (18 ^e , 28 ^e , etc.)
Stratifié	Méthode utilisée lorsqu'une population généralement hétérogène est fragmentée en catégories, appelées strates. Chaque strate est représentée dans l'échantillon selon le rapport : $\frac{\text{taille de la strate}}{\text{taille de la population}}$ Les éléments de chaque strate sont choisis au hasard	On sélectionne 10% des élèves (300 élèves) de chacun des programmes d'études pour former l'échantillon (en respectant le rapport).
Par grappes	Méthode utilisée lorsqu'une population généralement homogène est composée de groupes, appelées grappes, qui sont des sous-ensembles de la population. L'échantillon est formé de tous les éléments qui composent certaines grappes choisies au hasard.	Parmi tous les programmes d'études, on en choisit 11 aléatoirement. L'ensemble des élèves de ces programmes forment l'échantillon.

Explications tirées du guide « Intervalle » CEC, MAT-3052

1.1.1 > Échantillon représentatif et méthodes d'échantillonnage

- Un échantillon est **représentatif** d'une population s'il a les mêmes caractéristiques que cette population.
- Pour former un échantillon représentatif, on utilise une **méthode d'échantillonnage**. En voici deux :
 - l'échantillonnage par grappes ;
 - l'échantillonnage stratifié.

Échantillonnage par grappes

Description :

- Méthode qui consiste à diviser la population en groupes que l'on appelle **grappes**.
- On choisit **au hasard** un certain nombre de grappes et on forme un échantillon qui comporte **tous les éléments** contenus dans ces grappes.

Avantage :

Permet de réduire les coûts de sondage si la population est trop dispersée.

Contextes d'utilisation :

Usines, établissements d'enseignement, régions géographiques, etc.

Exemple : On veut savoir combien d'employés de 18 à 25 ans qui travaillent dans les entreprises du Québec ont terminé leurs études secondaires. On choisit donc au hasard 70 entreprises réparties sur le territoire de la province et on interroge tous les employés âgés de 18 à 25 ans qui travaillent dans ces 70 entreprises.

Échantillonnage stratifié

Description :

- Méthode qui consiste à diviser la population en groupes homogènes ayant des caractéristiques propres que l'on appelle **strates**.
- Chaque strate est représentée dans l'échantillon selon le **rapport** suivant.



$$\frac{\text{Taille de la strate}}{\text{Taille de la population}}$$

Ensuite, on choisit **au hasard** des éléments de chaque strate.

Avantages :

- Rend la stratégie d'échantillonnage plus efficace et l'**estimation** plus exacte.
- Assure une meilleure représentativité de la population.

Contextes d'utilisation :

Dans les situations où les caractères utilisés pour former les strates sont :

- faciles à observer ;
- étroitement liés au thème du sondage (ou du recensement).

Exemple : On cherche à connaître les activités de fin d'année préférées des élèves d'une école en réalisant un sondage auprès d'un échantillon représentatif de 60 de ces élèves. Pour ce faire, on divise les élèves en strates selon leur sexe et leur niveau scolaire. Par exemple, la 1^{re} strate regroupe les élèves de sexe féminin de 1^{re} secondaire, la 2^e strate regroupe les élèves de sexe masculin de 1^{re} secondaire, et ainsi de suite. Ensuite, on choisit au hasard des élèves de chaque strate selon le rapport $\frac{\text{taille de la strate}}{\text{taille de la population}} \times \text{taille de l'échantillon}$.

En pratique >>>

Voici un problème à résoudre.

Dans une école, on souhaite former un comité consultatif composé de 20 élèves représentatifs de la population étudiante décrite ci-contre en procédant à un échantillonnage stratifié.

Population étudiante

Niveau	Effectif
3 ^e secondaire	135
4 ^e secondaire	75
5 ^e secondaire	90

Voici un exemple de démarche possible.

- Calculez le nombre total d'élèves formant la population étudiante.

$$135 + 75 + 90 = 300 \text{ élèves}$$

- Déterminez le nombre d'élèves qu'il faut sélectionner dans chaque groupe pour former l'échantillon.

$$\begin{array}{l} 3^{\circ} \text{ secondaire :} \\ \frac{\text{taille de la strate}}{\text{taille de la population}} \times 20 = \frac{135}{300} \times 20 = 9 \text{ élèves} \end{array}$$

$$\begin{array}{l} 4^{\circ} \text{ secondaire :} \\ \frac{\text{taille de la strate}}{\text{taille de la population}} \times 20 = \frac{75}{300} \times 20 = 5 \text{ élèves} \end{array}$$

$$\begin{array}{l} 5^{\circ} \text{ secondaire :} \\ \frac{\text{taille de la strate}}{\text{taille de la population}} \times 20 = \frac{90}{300} \times 20 = 6 \text{ élèves} \end{array}$$

$$\text{Total} = 9 + 5 + 6 = 20 \text{ élèves}$$

Le comité sera composé de 9 élèves de 3^e secondaire, de 5 élèves de 4^e secondaire et de 6 élèves de 5^e secondaire.

2. Moyenne pondérée

MOYENNE PONDÉRÉE

- La **moyenne pondérée** est utilisée dans des situations où les valeurs n'ont pas toutes la même importance. On attribue alors une pondération à chacune des valeurs.
- La moyenne pondérée peut être calculée de la façon suivante.



$$\text{Moyenne pondérée} = \frac{\text{somme des produits de chaque valeur par sa pondération}}{\text{somme des pondérations}}$$

Exemple: Le tableau suivant présente les résultats d'une danseuse à un concours de danse où on note les participants selon trois critères.

Résultats obtenus à un concours de danse

Critère	Note obtenue (%)	Pondération des critères (%)
Originalité	86	20
Performance	82	50
Respect des exigences	90	30
Total:		100

$$\text{Moyenne pondérée de cette danseuse} = \frac{86 \times 20 + 82 \times 50 + 90 \times 30}{100} = 85,2$$

La moyenne pondérée de cette danseuse est de 85,2 %.

3. Probabilités : pages du guide « Intervalle » CEC

Dans le guide, on recourt à certaines formules pour les cas de dénombrement suivants :

- Dénombrement : importance ou non de l'ordre des choix
 - pages 108-109
- Permutation : on tient compte de l'ordre, la répétition n'est pas permise
 - pages 112 à 114
- Arrangement : avec et sans répétition
 - pages 117 et 119-120
- Combinaison
 - pages 123-124

Dans le programme :

« Les calculs (arrangement, permutation et combinaison) se faisant par raisonnement, il n'est pas nécessaire de recourir à des formules de dénombrement. »