

Session 5 Créer des variables parasitologiques

5.1 Introduction

Maintenant que nous en savons davantage sur les données et savons repérer les valeurs manquantes, nous pouvons commencer à créer des variables parasitologiques. Ces variables dépendent de l'espèce étudiée et du test utilisé :

Pour *S. mansoni* et les helminthes transmis par le sol (STH) évalués à l'aide de Kato-Katz :

1. Créer un décompte des œufs en œufs par gramme
2. Créer des variables d'intensité moyenne de l'infection
3. Créer des variables binaires (0/1) pour indiquer un résultat positif ou négatif
4. Créer des variables binaires pour indiquer si une personne présentait ou non une infection massive

Pour *S. haematobium* :

1. Créer un décompte d'œufs en œufs par 10 ml d'urine
2. Créer une variable binaire pour indiquer un résultat positif ou négatif
3. Créer une variable binaire pour indiquer si une personne présentait ou non une infection massive

Pour les autres variables parasitologiques (CCA, analyse d'urine par bandelettes, hématurie visible) :

1. Créer des variables binaires pour indiquer un résultat positif ou négatif

5.1.1 Classeur de fonctions

Les fonctions sont un concept important d'Excel, et mieux vaut les découvrir à partir d'un exemple.

1. Ouvrir le classeur nommé 'Session 5 – Classeur de fonctions'.
2. Résoudre les problèmes de la feuille 'à compléter' à l'aide des fonctions indiquées.
3. Si vous avez besoin d'aide, les réponses se trouvent dans la feuille 'Réponses'.

5.2 Créer des variables parasitologiques pour *S. mansoni* et les STH évalués par Kato-Katz

5.2.1 Créer un décompte des œufs en œufs par gramme

S. mansoni et les STH sont souvent évalués à l'aide de Kato-Katz où un ou plusieurs petits échantillons de selles sont examinés sous microscope pour détecter des œufs. Chaque échantillon de selles pèse environ $1/24^{\text{ème}}$ de gramme et on multiplie en général le nombre d'œufs trouvés par 24 pour exprimer le résultat en œufs par gramme (opg). Pour cela :

1. Ouvrir 'données 4' dans le dossier '2 Nettoyage des données'.
2. Faire défiler jusqu'à la première colonne vide à droite – colonne U
3. Masquer les lignes O à T qui ne contiennent pas des données Kato-Katz. Conserver *S. mansoni*, *Ascaris*, *Ankylostomiasis* et *Trichuris*.
4. Créer une nouvelle colonne appelée 'opg S. mans 1'. Vérifier que vos résultats ressemblent à ceci :

d'une valeur non manquante comme égale à la valeur non manquante. Le tableau ci-dessous indique comment Excel gère les valeurs manquantes pour calculer des moyennes.

Var 1	Var 2	Moyenne(Var1, Var 2)
6	4	5
6		6
	4	4
		#DIV/0!

3. Pour contourner ce problème, on doit utiliser une fonction SI indiquant
 - a. Si les deux valeurs manquent, afficher manquante
 - b. Sinon, afficher la moyenne des deux valeurs
4. Dans la cellule AC2, saisir la formule '=SI(ET(V2="",Z2=""), "", MOYENNE(V2,Z2))' et appuyer sur Entrée

V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG
opg S. mans 1	opg Ascaris 1	opg Ankylo 1	opg Trichuris 1	opg S. mans 2	opg Ascaris 2	opg Ankylo 2	opg Trichuris 2	S. mans opg moyenne			
0	0		0	24	0		0	=SI(ET(V2="",Z2=""), "", MOYENNE(V2,Z2))			
0	0		0	0	0		0	SI(test_logique, [valeur_si_vrai], [valeur_si_faux])			

5. Copier la formule jusqu'en bas de la page et vérifier que les cellules sont correctement remplies.
6. Trouver quelques élèves dont les données *S. mansoni* manquent et vérifier que la formule fait ce qu'il faut.
7. Nommer les colonnes AE à AG pour les autres STH dans l'ordre qui convient puis utiliser la poignée de recopie pour compléter les colonnes. Vérifier qu'Excel fait ce qu'il faut.
Créer des variables binaires pour indiquer un résultat positif ou négatif

V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG
opg S. mans 1	opg Ascaris 1	opg Ankylo 1	opg Trichuris 1	opg S. mans 2	opg Ascaris 2	opg Ankylo 2	opg Trichuris 2	S. mans opg moyenne	Ascaris opg moyenne	Ankylo opg moyenne	Trichuris opg moyenne
0	0		0	24	0		0	12	0		0
0	0		0	0	0		0	0	0		0
312	0		0	192	0		0	252	0		0

Une fois l'intensité moyenne d'infection mesurée, on peut créer une variable indiquant le résultat positif ou négatif de chaque personne. On utilise 1 pour indiquer si le résultat était positif et 0 s'il était négatif. Ainsi, nous pourrions simplement prendre la moyenne des 1 et des 0 pour obtenir la prévalence. Nous verrons cela en détail ultérieurement. Pour créer les indicateurs de prévalence :

1. Créer une nouvelle variable dans la colonne AG2 appelée 'S. mans inf'.
2. Nous allons utiliser la fonction Excel 'SIGNE'. SIGNE affiche 1 lorsqu'un résultat est positif, 0 lorsqu'un résultat est 0 et -1 lorsqu'un résultat est négatif. Comme précédemment, il va nous falloir gérer les valeurs manquantes. Saisir la formule '=SI(AD2="", "", SIGNE(AD2))'

AH	AI	AJ	AK
S. mans inf			
=SI(AD2="", "", SIGNE(AD2))			
SI(test_logique, [valeur_si_vrai], [valeur_si_faux])			

3. Utiliser la poignée de recopie pour compléter toutes les lignes
4. Vérifier que les données sont telles que prévues. Vérifier en particulier les valeurs manquantes
5. Procéder de la même manière pour les STH dans les colonnes suivantes

5.2.3 Créer des variables binaires pour indiquer un résultat d'infection massive

Nous utiliserons 1 pour indiquer qu'une personne présente une infection massive et 0 pour indiquer qu'il n'y a pas d'infection massive (pas d'infection ou infection légère ou modérée). Les infections massives sont déterminées selon les directives de l'OMS :

Tableau 2.3 *Seuils pour le classement des infections par Ascaris lumbricoides, Trichuris trichiura, Ancylostomes et Schistosoma : infections légères, modérées ou d'intensité élevée*

Helminthe	Densité de l'infection		
	Légère	Modérée	Massive
<i>A. lumbricoides</i>	1–4999 oeufs/g	5000–49999 oeufs/g	≥50 000 oeufs/g
<i>T. trichiura</i>	1–999 oeufs/g	1000–9999 oeufs/g	≥10 000 oeufs/g
<i>Ancylostomes</i>	1–1999 oeufs/g	2000–3999 oeufs/g	≥4 000 oeufs/g
<i>S. mansoni</i>	1–99 oeufs/g	100–399 oeufs/g	≥400 oeufs/g
<i>S. haematobium</i>	1–50 oeufs/10ml d'urine		≥50 oeufs/10ml d'urine ou hématurie visible
<i>S. japonicum</i>	Toute infection est considérée comme massive		

Pour créer une variable d'intensité massive pour *S. mansoni* :

1. Créer une nouvelle variable dans la colonne AK appelée 'S. mans massive'.
2. Saisir la formule ' = SI(AC2="","", "", SI(AC2>400, 1, 0))'

AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN
S. mans opg moyenne	Ascaris opg moyenne	Ankylo opg moyenne	Trichuris opg moyenne	S. mans inf	Ascaris inf	Ankylo inf	Trichuris inf	S. mans massive		
12	0		0	1	0		0 =SI(AD2="","", "", SI(AD2>=400,1,0))			

3. Utiliser la poignée de recopie pour compléter toutes les lignes
4. Vérifier que les données sont telles que prévues. Vérifier en particulier les valeurs manquantes
5. Procéder de la même manière pour les STH dans les colonnes suivantes en modifiant la définition d'infection massive pour :
 - a. Ascaris 50 000
 - b. Ankylostomiase 4 000
 - c. Trichuris 10 000
6. Vérifier que les données sont telles que prévues, en particulier les valeurs manquantes

5.3 Créer des variables parasitologiques pour *S. haematobium* par filtration de l'urine

5.3.1 Créer des variables d'intensité moyenne de l'infection par *S. haematobium*

Lorsque l'on teste *S. haematobium* par filtration de l'urine, on filtre en général 10 ml d'urine. Toutefois, si moins de 10 ml d'urine est disponible, on estime le nombre d'œufs présents dans 10 ml

avec le calcul : $\text{oeufs par 10 ml d'urine} = \frac{\text{nombre d'oeufs détectés}}{\text{volume d'urine filtré en ml}} * 10$

Pour estimer le nombre d'œufs par 10 ml d'urine :

1. Faire défiler jusqu'à la prochaine colonne disponible (AP). Masquer toutes les variables entre AP et les informations sur la filtration d'urine (entre T et AN). Vous aurez peut-être à afficher

les données masquées dans la section précédente. Intituler la colonne AP 'S. haem op10ml' et vérifier que vos données ressemblent à ceci.

S	T	U	AP
maturie visible	Volume d'urine	S. haem Lame	S. haem op10ml
0	10	0	

2. Nous allons ajouter une formule qui dit :
 - a. Si le nombre d'œufs manque, afficher manquant
 - b. Si le nombre d'œufs manque ou est 10, afficher le nombre d'œufs (supposer que le volume d'urine recueilli était de 10 ml)
 - c. Si moins de 10 ml d'urine a été recueilli, utiliser la formule ci-dessus
 - d. Noter que cette formule n'inclut pas les fois où le volume d'urine a été stocké mais où le nombre d'œufs manquait, il faudrait donc étudier ça plus en détail
3. Essayer de trouver par vous-même quelle formule utiliser
4. On peut utiliser la formule suivante
 '=SI(ET(ESTVIDE(R2),ESTVIDE(S2)), "", SI(OU(R2=10,ESTVIDE(R2)),S2,10*S2/R2))'

	T	U	AP	AQ	AR	AS	AT	AU	AV	A
rie	Volume d'urine	S. haem Lame	S. haem op10ml							
	10	0	=SI(ET(ESTVIDE(T2),ESTVIDE(U2)), "", SI(OU(T2=10,ESTVIDE(T2)),U2,10*U2/T2))							

5. Utiliser la poignée de recopie vers le bas et vérifier que les données sont telles que prévues.

5.3.2 Créer des variables binaires pour indiquer un résultat positif ou négatif pour *S. haematobium*

Créer une variable 'S. haem inf' dans la colonne AP. Utiliser la fonction SIGNE de la section 2.3.3 pour créer une variable binaire pour *S. haematobium* et vérifier qu'elle ressemble à ceci

AP	AQ
S. haem op10ml	S. haem inf
0	0
14	1
861	1

5.3.3 Créer un indicateur d'infection massive pour *S. haematobium*

Suivre les étapes de la section 2.2.4 pour créer un indicateur d'infection massive pour *S. haematobium*. Les décomptes d'œufs de 50 et plus par 10 ml d'urine sont considérés comme des infections massives. Vérifier que vos résultats ressemblent à ceci :

U	AP	AQ	AR
S. haem Lame	S.haem op10ml	S.haem inf	S. haem massive
0	0	0	0
14	14	1	0
861	861	1	1
0	0	0	0

5.4 Créer des indicateurs d'infection pour CCA, analyse d'urine par bandelettes et hématurie visible

Utiliser la fonction SIGNE pour créer des indicateurs pour CCA, analyse d'urine par bandelettes et hématurie visible. Vérifier que vos résultats ressemblent à ceci :

O	P	Q	AR	AS	AT
CCA	Dipstick	Visible Haematuri a	CCA pos	Dipstick pos	Visible haematuri a pos
1	0	0	1	0	0
2	3	0	1	1	0
2	5	0	1	1	0
	0	0		0	0
3	2	0	1	1	0
0	4	0	0	1	0

5.4.1 Enregistrer votre travail.

Enregistrer l'ensemble de données actuel sous le nom 'données 5' dans '2 nettoyage des données'. Ainsi, vous pourrez voir ce que vous avez modifié lors de ces étapes.