HAL
open science

UN DISPOSITIF DE FORMATION DES PROFESSIONNELS DE LA SANTE A LA STATISTIQUE INCLUANT UNE SIMULATION INFORMATIQUE D'UNE SITUATION PROFESSIONNELLE

Annie Bessot

To cite this version:

Annie Bessot. UN DISPOSITIF DE FORMATION DES PROFESSIONNELS DE LA SANTE A LA STATISTIQUE INCLUANT UNE SIMULATION INFORMATIQUE D'UNE SITUATION PROFESSIONNELLE. [Rapport de recherche] Laboratoire LIG. 2012. hal-01528261

HAL Id: hal-01528261

https://hal.science/hal-01528261v1

Submitted on 1 Jun 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UN DISPOSITIF DE FORMATION DES PROFESSIONNELS DE LA SANTE A LA STATISTIQUE INCLUANT UNE SIMULATION INFORMATIQUE D'UNE SITUATION PROFESSIONNELLE

D'après le rapport de la mission AUF, ESM et UP HCM

Annie Bessot, équipe MeTAH, Laboratoire LIG, Université Joseph Fourier

Ho Chi Minh ville du 28 février au 1er mai 2012

Introduction

Le cadre du projet de recherche coopérative est celui de deux collaborations :

- une collaboration nouvelle entre l'Université Joseph Fourier (UJF, Grenoble 1) et l'Université des Sciences Médicales de Ho Chi Minh Ville (USM) : signature de l'accord de coopération le 23 mars 2011 (validité 5 ans)
- une collaboration ancienne avec l'Université de pédagogie (UP) : renouvellement de l'accord de coopération le 2 septembre 2011 (validité 5 ans)

Le projet de recherche coopérative, *soutenu par la région Rhône-Alpes (Coopera),* bénéficie d'une expertise acquise dans le projet Loé (UJF) que nous présentons maintenant.

I. Le projet Loé[1]

Ce projet est une recherche-action à laquelle ont participé Muriel Ney, Marie-Caroline Crozet, Jean-Luc Bosson (Croset, Ney et al. 2011). Il vise à modifier le module d'enseignement de biostatistique à la faculté de médecine de Grenoble. Ces équipes ont mis en place un environnement immersif : le Laboratorium of Epidemiology (Loé, http://www.tel-laboratorium.fr).

Figure 1. Vue du site web du projet Loé

Dans cet environnement immersif, les étudiants de médecine doivent réaliser une mission où ils jouent le rôle de médecin en santé publique. Ils doivent concevoir et effectuer une étude épidémiologique puis écrire un article scientifique qu'ils devront présenter à un congrès simulé (http://www.tel-laboratorium.fr).

[1] D'après Muriel Ney

La simulation d'une enquête épidémiologique a été conçue comme un jeu de rôle, dit *jeu sérieux* (à l'origine jeu vidéo qui ont une finalité autre que s'amuser, ici enseignement – apprentissage).
L'image est celle que les étudiants voient à l'entrée dans le Jeu Loé.

Figure 2. Entrée dans le jeu Loé

L'enseignement du module de *biostatistique* se déroule sur un semestre pédagogique : il comporte 6 cours de 2h et 8 séances de travaux Pratiques (TP), chaque séance de TP durant 4h. Il se situe en seconde année de médecine à Grenoble et concerne 170 étudiants répartis en 6 groupes de travaux pratiques. Le jeu Loé a lieu durant les 8 séance de TP. Il est suivi par 8 à 10 tuteurs. Une même famille de maladies a été choisie, *les maladies thromboemboliques veineuses* (MTEV).

Ce dispositif pédagogique est une réponse à un *besoin professionnel affiché nationalement dans la formation des médecins en France*, celui de la lecture et l'utilisation critique d'articles statistiques.

Le jeu Loé vise donc en premier lieu l'enseignement / apprentissage de la lecture critique d'un article médical en faisant entrer les étudiants *dans le processus de l'écriture d'un article médical*. Ce processus nécessite d'accomplir une suite de types de tâches :
- poser le problème que l'on veut traiter,
- écrire les hypothèses à tester,
- définir la situation à observer,
- recueillir les données,
- effectuer le traitement statistique des données,
- écrire l'article
- le présenter à un congrès simulé (démarche hypothético-déductive : voir plus loin l'analyse du manuel français).

A travers ce processus, un deuxième objectif, spécifique à Loé, est de donner du sens à certains savoirs statistiques, au travers de leur fonctionnement dans le processus.

- Environnement informatique et enquête épidémiologique : le recueil des données

Revenons sur le jeu Loé. Une fois le problème posé et les hypothèses formulées, les étudiants passent au recueil des données (même si le processus n'est pas aussi linéaire !). Ils doivent alors rentrer en contact avec différents organismes *via* des sites web imités des vrais sites web (figure 3) comme :
- la « Commission du médicament (local) » qui peut être commanditaire de l'enquête (le projet de l'étudiant doit contribuer à cette enquête qui est plus vaste) ;
- le « Comité de protection des personnes » qui sont les experts qui valident les protocoles d'enquête : objectifs, échantillonnage, méthode de recueil, méthode d'analyse, biais, questions d'éthique ;
- Le Site du congrès ;

- le Site de la bibliothèque : ressources statistiques (qui peuvent être des vrais sites), ressources sur la maladie (idem) ;

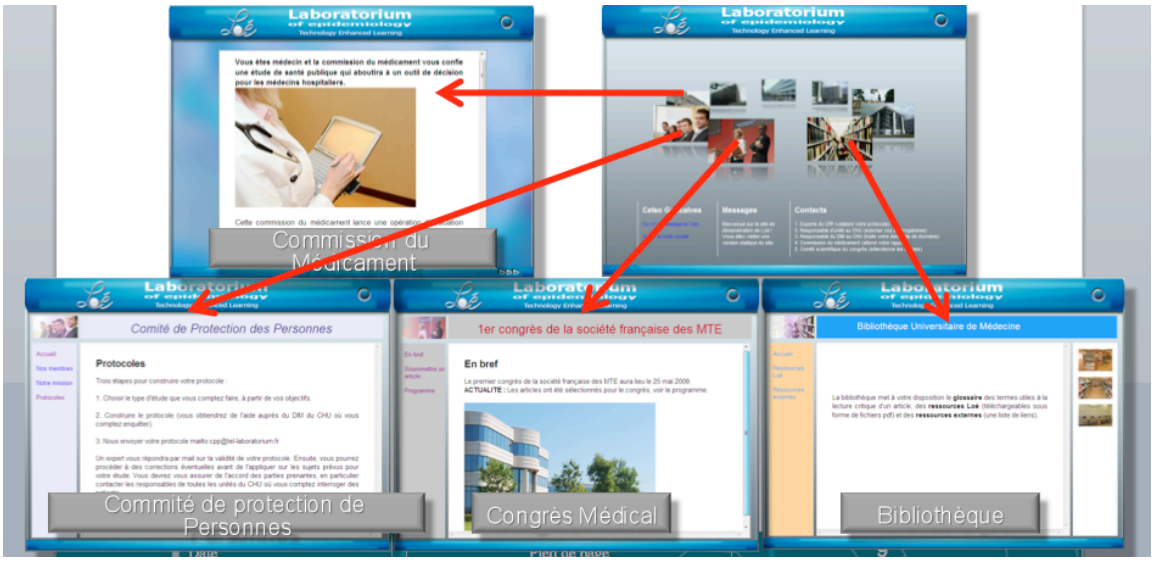

Figure 3. Différents organismes et leurs sites web imités des vrais sites

Les autres images représentent différents hôpitaux virtuels en France. Actuellement il y a trois hôpitaux virtuels sur le site et 45 patients (vidéos).

-> Exemple de recueil de données auprès d'un hôpital (suivre sur la figure 4):

1. Entrée dans l'hôpital H1 dans lequel il y a 3 services.
2. Entrée dans l'un des services dans lequel il y a 5 chambres avec les dossiers médicaux des patients
3. On rentre dans une chambre avec une vidéo du patient et une fiche de questions prévus à l'avance et qui sont typiques : on ne peut pas faire répéter le patient (ce patient est un acteur qui simule la maladie et qui les réponses)
4. Les étudiants doivent s'entretenir avec 10 patients

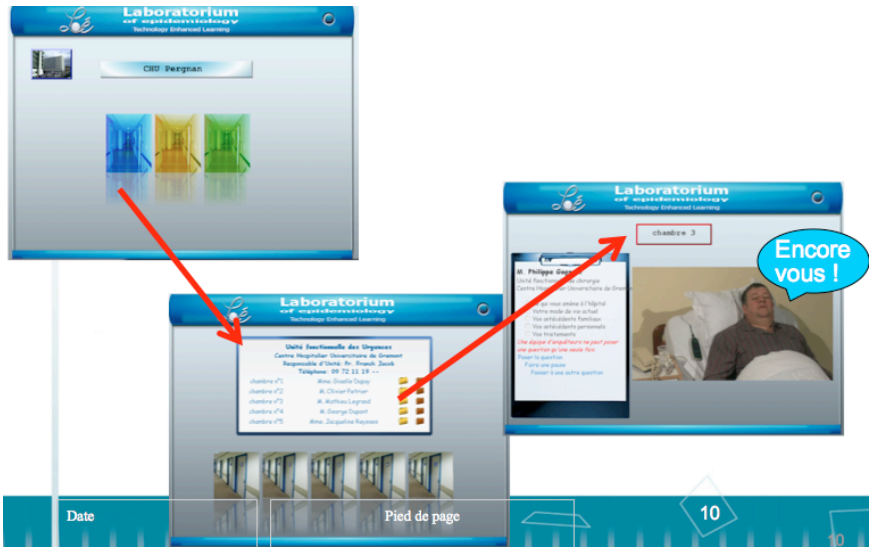

Figure 4. Exemple de recueil de données auprès d'un hôpital H1

Puis l'enquête s'effectue sur une base beaucoup plus large, celle d'une base de données médicale sur les maladies thromboemboliques veineuses. Les étudiants ont accès à cette base de données en déposant une requête *via* le site Loé.

Les étudiants construisent donc leurs données. L'enseignant de statistique a aussi accès à ces données, mais sa démarche est à l'inverse de la démarche pédagogique habituelle où l'enseignant fournit à ses étudiants les données. Ici, l'enseignant ne connait pas à l'avance les données des étudiants !

De plus le site Loé met à la disposition des étudiants pour le traitement de ces données des outils sélectionnés parmi ceux offert par le logiciel libre R : la sélection est le résultat du travail des enseignants.

II. Présentation du projet de recherche coopérative

Au Viêt Nam, comme en France, l'enseignement de biostatistique est obligatoire en 1ère année de médecine au Viêt Nam et en 2ème année en France. Un premier objectif de la recherche est de déterminer quelles sont les généricités et les spécificités de ces enseignements dans le but de concevoir un environnement informatique pour l'apprentissage des statistiques commun aux deux pays tout en respectant les caractéristiques propres à chacun. Le projet s'articule autour de deux grands axes : une analyse comparative des deux institutions d'enseignement et un cahier des charges pour l'élaboration d'un environnement informatique.

II.1 Une analyse comparative des deux systèmes d'enseignement des statistiques

Comme l'exposent Bessot et Comiti (Bessot & Comiti, 2008), une analyse comparative d'institutions d'enseignement permet, entre autre, de prendre du recul et de dénaturaliser le regard du chercheur sur le fonctionnement d'une institution qui, à force d'habitude, rend inapparente l'utilisation de certaines pratiques. De plus, une telle comparaison permet de faire émerger ce qui est commun et spécifique aux deux systèmes.

II.1.a Analyse praxéologique de deux manuels ou documents d'enseignement

L'analyse des manuels et des programmes permet à un premier niveau de comprendre le fonctionnement d'un système éducatif. Il est aujourd'hui devenu classique d'utiliser pour cela le cadre théorique de la Théorie Anthropologique du Didactique (TAD) développée par Chevallard (Chevallard, 1992 ; Bosch & Chevallard, 1999). Ce cadre permet de caractériser les attentes institutionnelles sous forme de praxéologies (Chaachoua & Comiti, 2007). L'objet de savoir ciblé dans la recherche sera les statistiques en jeu lors de l'étude épidémiologique ou d'un essai thérapeutique qui consiste le plus souvent *à comparer deux échantillons et à faire un test d'hypothèse.* Pour cela deux ouvrages de biostatisques, l'un français et l'autre vietnamien, seront sélectionnés en collaboration avec les équipes enseignantes des pays respectifs.

II.1.b Entretiens avec des enseignants de mathématiques de l'Université de Médecine et Pharmacie d'Ho Chi Minh Ville

Certains enseignants, dont Dao Hong Nam, partenaire de ce projet, s'interrogent sur une éventuelle perte de sens des savoirs statistiques abordés pour les futurs médecins auxquels ils enseignent. Une enquête auprès d'enseignants de mathématiques de l'UMP permettra de préciser cette inquiétude.

II.1.c Entretiens avec des médecins, professionnels

Le projet de recherche cherche à contribuer à la modification du rapport des enseignants des mathématiques à l'objet « statistique » pour les rapprocher des besoins des professionnels de la santé, en particulier des médecins, et donner du sens aux statistiques par leur fonctionnement dans une situation professionnelle. Par exemple, un des besoins reconnu dans la formation médicale française est celui de pouvoir porter un regard critique sur la littérature médicale. Quels sont les besoins professionnels en statistique perçus par les médecins au Viêt Nam ?

II.1.d Questionnaire auprès des étudiants

Un questionnaire identique pour les deux pays sera conçu autour des principaux types de tâches repérés lors de l'analyse praxéologique des manuels. Des étudiants des deux pays y répondront. Une comparaison des résultats permettra de caractériser le rapport institutionnel au savoir (au sens de la TAD) propre à chaque pays ; des explications tant épistémologiques que didactiques des différences pouvant ressortir de l'analyse praxéologique précédente. Le dépouillement et le classement des différences se fera à l'aide d'un modèle qui permet de décrire les praxéologies correctes et erronées (dites « en-acte » ou « personnelles ») des étudiants.

II.2 Un cahier des charges pour l'élaboration d'un environnement informatique d'enseignement des statistiques autour d'une situation d'essai thérapeutique

La création d'un environnement informatique pour l'enseignement des statistiques au Viêt Nam fera suite à l'analyse comparative précédente. Son cahier des charges se fera en trois temps.

II.2.a Rédaction d'une mission

Centrée sur une situation d'essai thérapeutique ou d'étude épidémiologique, une mission sera rédigée en collaboration avec les équipes enseignantes de l'Université de médecine et de pharmacie d'Ho Chi Minh Ville et J-L Bosson, responsable de l'enseignement des statistiques en 2ème année à Grenoble et partenaire du projet de coopération. La mission sera conçue pour conduire les étudiants à se poser un certain nombre de problèmes qu'ils devront résoudre à l'aide des savoirs statistiques en jeu.

II.2.b Réflexion autour des espaces d'activités

Pour répondre à cette mission, les étudiants utiliseront un environnement informatique leur donnant accès à des espaces d'activités (accès à des simulations d'hôpitaux ou de patients, des outils statistiques, etc.). Ces espaces auront été définis grâce à l'analyse comparative et plus particulièrement à l'enquête auprès des enseignants.

II.2.c Ecriture d'un scénario possible d'usage

Pour sélectionner puis organiser ces espaces d'activités, un ou plusieurs scénarios d'usage seront proposés à destination des enseignants. Ces scénarios seront conçus en collaboration avec les chercheurs des trois équipes partenaires du projet de coopération, *des enseignants de mathématiques et des médecins de l'USM.*

III. Quelques résultats de l'analyse comparative de deux manuels vietnamien et français

Nous avons retenu deux ouvrages *:*
– J. Bénichou, R. Beusccart, P. Roy, C. Quantin, avec le Collège national des enseignants des facultés de médecine (2009) Biostatistique (397 pages) - ouvrage de référence en France, *noté par la suite Mfr*
– Chu Dinh Thanh, Pham Minh Buu, Tran Dinh Thanh, Nguyen Van Lieng (2011) Pobabilité Statistique (260 pages) – ouvrage que doivent suivre les enseignants de probabilité et statistique de l'USM, *noté par la suite Mvn*

La traduction des chapitres VIII et IX, y compris les exercices, ont nécessité plus de 60 h de travail d'Annie Bessot, Lê Thi Hoai Chau et Vu Nhu Thu Huong. C'est sur cette traduction que s'appuie l'analyse comparative.

Le cadre théorique est celui de la Théorie Anthropologique du Didactique (TAD) : quelles organisations mathématiques autour des tests d'hypothèse ?

III.1. Comparaison globale des deux manuels

Le tableau 1 résume quantitativement la place et l'importance de différents domaines présents dans le deux manuels.

Manuels		**Probabilité**	**Statistique**	**Méthodologie des situations médicales**	**Total**
Mvn	Chap. (nbre)	5	**5**	1	11
	pages (%)	91 (39,2%)	**119 (51,3%)**	22 (9,5%)	232
Mfr	Chap. (nbre)	1	**8**	4	13
	pages (%)	34 (9,6%)	**220 (61,8%)**	102 (28,6%)	356

Tableau 1. Comparaison globale des deux manuels, vietnamien (Mvn) et français (Mfr)

Ce qui est semblable
- L'existence de trois domaines : Probabilité, Statistique, Méthodologie spécifique aux situations médicales ;
- L'ordre des domaines ;

De plus, l'analyse des deux ouvrages montre l'existence de 2 genres de tâche statistique :

Genre 1 : Estimer …

Genre 2 : Comparer …

ainsi que l'étude de la corrélation entre variables quantitatives.

Ce qui est différent
- Les « probabilités » occupent une place plus importante dans Mvn que dans Mfr:

5 chapitres dans Mvn (90 pages sur 260)
1 seul chapitre dans Mfr (34 pages sur 397)
- Par contre, la « méthodologie spécifique aux situations médicales » occupe une place plus importante dans Mfr que dans Mvn :

Mvn : un chapitre (XI) intitulé : « Probabilité dans le diagnostic »
Mfr : 4 chapitres : chap. 10 « Les principes méthodologiques des essais cliniques », Chap. 11 : « Les statistiques pour l'épidémiologie », chap. 12 « L'analyse de la survie », Chap. 13 « L'analyse de la décision médicale », ce dernier chapitre correspond au chapitre 11 de Mvn.

Maintenant nous allons nous intéresser à la partie des manuels concernant les tests d'hypothèse : Chap. VIII et IX dans Mvn (57 pages sur 232 soit 24,6%) ; Chap. 5, 6, 7 et 8 dans Mfr (101 pages sur 356 soit 28,4%)

III.2. Structuration du chap. VIII du manuel vietnamien et des chapitres correspondants du manuel français (Tests d'hypothèses)

Les tests d'hypothèse interviennent pour le genre de tâche « Comparer … ». Ils font l'objet de deux démarches différentes dans Mvn et dans Mfr qui sont présentées dans le tableau 2.

Manuel Vietnamien Mvn *Application des probabilités*	Manuel Français Mfr « démarche hypothético-déductive » (cf. figure 1) *Modélisation par les probabilités*
1. La situation, les données expérimentales et la question sont données soit explicitement soit implicitement (cas). 2. L'hypothèse H est - soit déduite de la question	« 1. Examiner le problème médical à étudier de manière à formuler sous la forme d'une question simple l'hypothèse qui sera mis à l'épreuve. 2. Énoncer cette hypothèse H_0 et une hypothèse

- soit donnée dans la partie cours avec chaque cas. 3. Aux différents cas du cours, sont attachés différents tests statistiques. Ces différents cas font office des conditions d'application. 4. On donne le résultat : on accepte ou on rejette l'hypothèse sous un seuil d'erreur α *a posteriori* et qui semble arbitraire 5. Conclusion : on traduit le résultat dans les termes des observations.	H_1. 3. Fixer à l'avance le risque d'erreur α pour décider l'acceptation ou le rejet de H_0. 4. Choisir le test approprié et en vérifier les conditions d'application. 5. Calculer la statistique appropriée, conclure sur le plan statistique – comparaison de la valeur de la statistique calculée au seuil de signification correspondant au risque α choisi – et donner le degré de signification du test (que l'hypothèse H_0 soit rejetée ou non). 6. Donner une interprétation médicale des résultats du test, en cohérence avec les connaissances scientifiques. (Mfr, p. 165)

Tableau 2. Deux démarches sur les tests d'hypothèses

La démarche propre au manuel français est illustrée dans le manuel par un schéma qui va organiser le chap. 5 « Les principes des tests d'hypothèses » (voir figure 5).

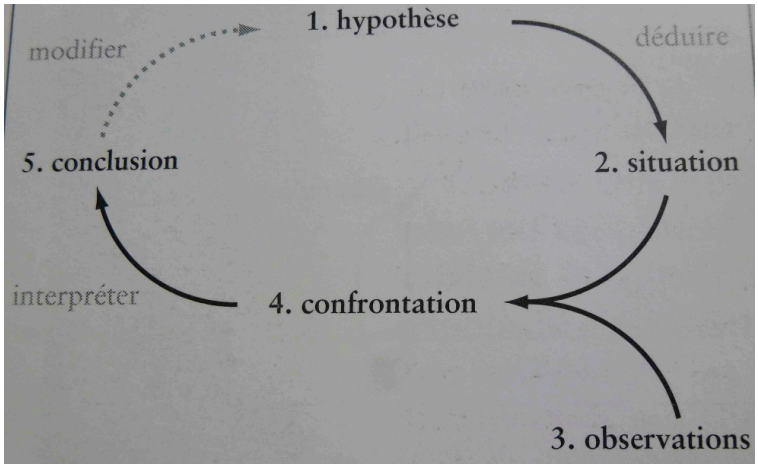

Figure 5. Schéma de la démarche hypothético-déductive des tests d'hypothèse (Mfr, pp. 148-156)

A propos de l'étape 5 (choix d'un test d'hypothèse) de la démarche « hypothético-déductive », voici ce qu'écrivent les auteurs de Mfr.

> Le choix d'un test d'hypothèse nécessite au préalable de porter une réflexion sur les données du problème à analyser. Il existe différents points qui sont communs à l'ensemble des tests et qu'il convient de relever afin de choisir le bon test :
> 1• Le type de variables mises en relation
> - une variable qualitative/une variable qualitative
> - une variable quantitative/une variable qualitative
> - une variable quantitative/une variable qualitative
> 2• La taille de l'échantillon
> - Taille de l'échantillon inférieure à 30 ;
> - Taille de l'échantillon supérieure ou égale à 30 (cette limite permet alors d'utiliser une statistique fondée sur la distribution normale).
> 3• Les conditions d'applications
> Les tests d'hypothèse sont utilisables dans certaines conditions bien définies qui leur sont souvent spécifiques. Ces conditions seront spécifiées lors de la présentation de chaque test. Cela permet également d'orienter vers le choix d'un test paramétrique ou d'un test non paramétrique.
> 4• Échantillons indépendants ou appariés
> - Échantillons indépendants : les mesures sont effectuées sur des individus différents.

- Échantillons appariés : les différents échantillons proviennent de mesures effectuées sur des mêmes individus. (p. 164)

Le tableau 3 reproduit le tableau récapitulatif des tests d'hypothèse présenté dans Mfr, page 65.

Type de variables		échantillons indépendants		échantillons appariés	
		paramétrique	non paramétrique	paramétrique	non paramétrique
qual/qual			test du χ^2		test du χ^2
			test exact de Fisher		test exact de Fisher
quant/qual	2 échantillons	test *t* non apparié	test *U* de Mann-Whitney	test *t* apparié	Test de McNemar
		test *Z*	test de Komogorov-Smimov		test de Wilcoxon
	3 échantillons et plus	ANOVA	test de Kruskall-Wall		
quant/quant[2]				*r* de pearson	*r* de Spearman
Qual : qualitative – quant : quantitative					

Tableau 3. Tableau récapitulatif des différents tests (Mfr p. 65)

III.2.1 Manuel vietnamien

Deux chapitres sont consacrés aux tests d'hypothèse : chap. VIII : « Test d'hypothèse statistique » ; chap. IX « Analyse de la variance ». Nous présentons sous la forme d'un arbre la structure du chap. VIII du manuel vietnamien (figure 6).

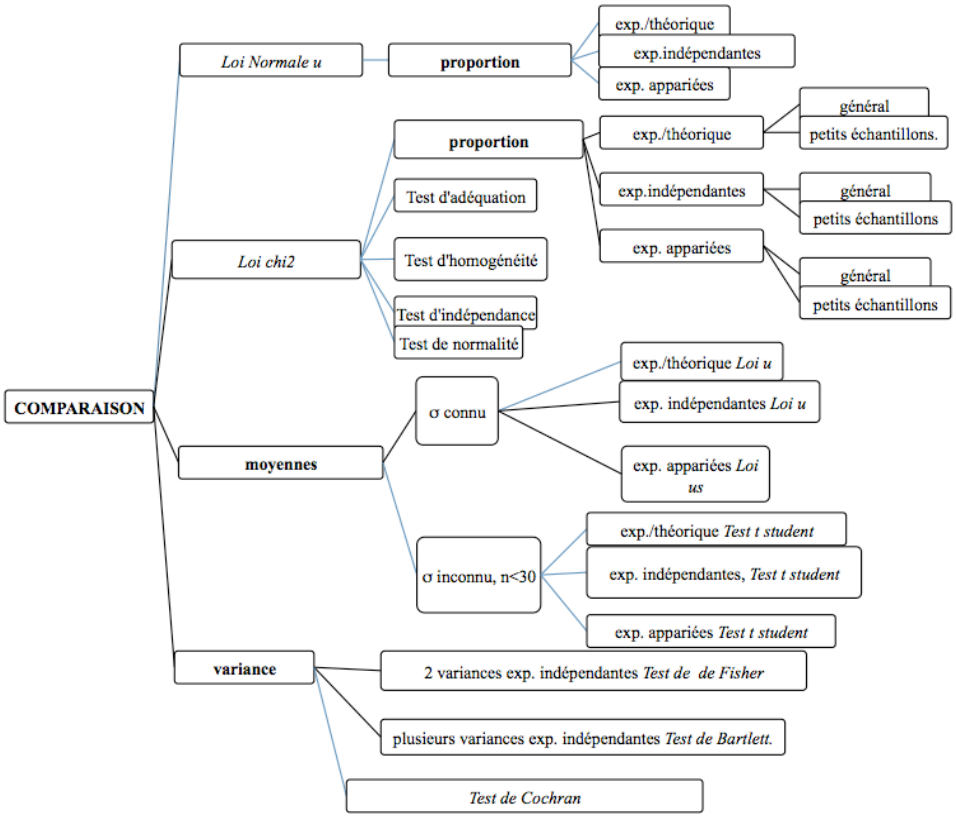

Figure 6. Arbre présentant la structure du chap. VIII de Mvn

[2] Étude de la corrélation entre variables quantitatives.

Commentaires
• 3 catégories d'objets comparés
*1. Proportion*s (tests : u et Chi2)
2. *moyennes* (tests : u et t Student)
3. *Variances* (tests Fisher, Bartlett, Cochran)
• 3 sous catégories :
a. Expérimental &Théorique
b. Expérimentales indépendantes
. Expérimentales appariées
• 2 sous catégories / à la taille de l'échantillon
Général/ petit échantillon (qui dépendent de nombres précis : d'où viennent ces nombres ?)
⇨ 18 cas : analogue aux cas de l'arithmétique élémentaire
Trois types de tâche dans le genre « Comparer … » : « comparer des proportions » ; « comparer des moyennes » ; « comparer des variances »

III.2.2. Manuel français Biostatistique (2009)

Les tests d'hypothèse sont présents dans 4 chapitres : Chap. 5 « Principes des tests d'hypothèse » ; Tests paramétriques : Chap. 6 « Comparaison des fréquences » ; Chap. 7 Comparaison des moyennes » ; Chap. 8 « Tests non paramétriques ».
Nous présentons sous la forme d'un arbre la structure des chap. 6 et 7 du manuel français (figure 7) qui correspondent aux chapitres VIII et IX de Mvn.

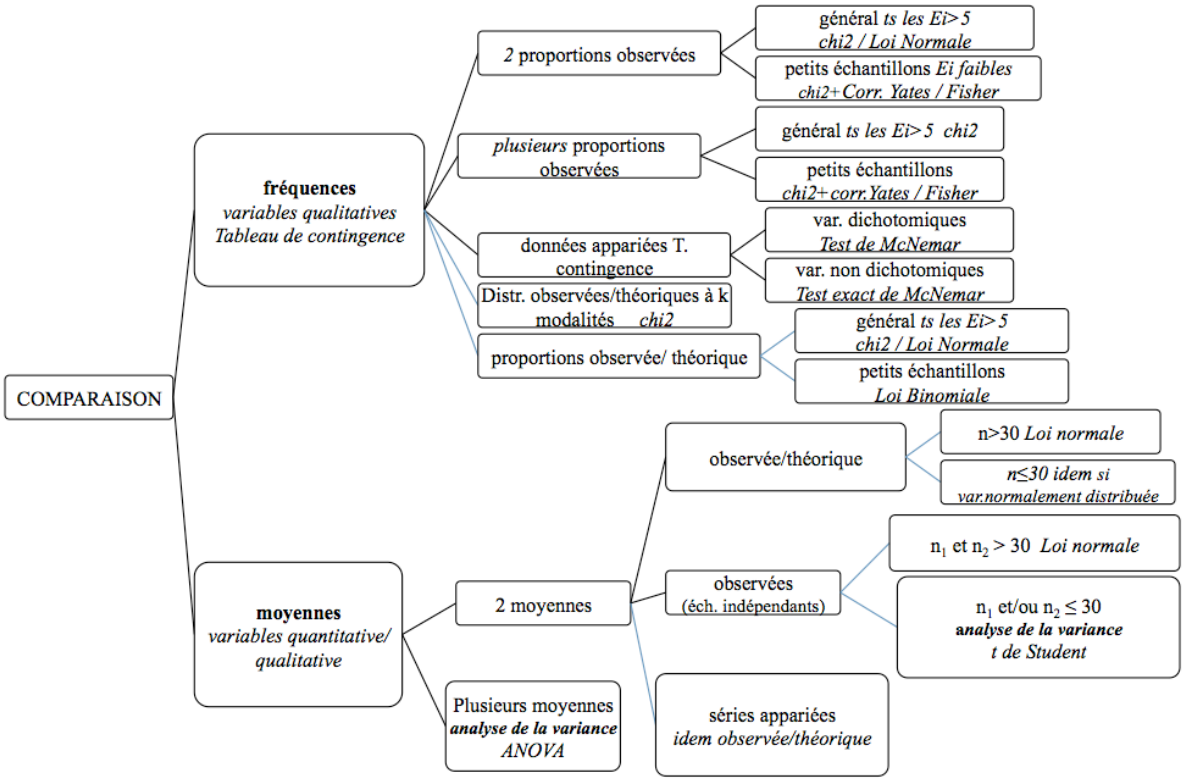

Figure 7. Arbre la structure des chap. 6 et 7 de Mfr

Commentaires
• 2 catégories d'objets comparés (pas de comparaison de variance en intitulé)

1. Fréquences ou proportions tests : Chi2 ou loi normale[3], Fisher, correction de Yates, McNemar, Loi Binomiale
2. *moyennes* tests : Loi Normale, t Student, ANOVA
• 3 sous catégories :
a. Expérimental &Théorique
b. Expérimentales indépendantes
c. Expérimentales appariées
• 2 sous catégories / à la taille de l'échantillon
Général/ petit échantillon (qui dépendent de nombres précis : d'où grands types de tâches « Comparer des fréquences » et « comparer des moyennes ». Par exemple la séparation entre fréquence et moyenne pour les tests paramétriques est justifiée par les traitements de variables qualitatives ou quantitatives.
L'analyse de la variance est présente pour comparer les moyennes (cas de variables quantitatives) quand petits échantillons, ou plusieurs moyennes (on pose alors l'hypothèse nulle : les k moyennes sont égales)

III. 3. Comparaison de Mvn et Mfr - Les principales différences

III. 3.1. Hypothèse(s)

•Mvn : 1 seule hypothèse H
•Mfr : 2 hypothèses H0 *dite hypothèse nulle* que l'on cherche à rejeter et H1 *dite hypothèse alternative* (négation logique de H0)
Justification pour le choix de H0, hypothèse à tester pour son rejet : au niveau logique plus facile de rejeter (un seul CE suffit) que d'accepter

III. 3.2. Deux distinctions absentes dans Mvn, présentes dans Mfr

– Variables aléatoires qualitative / quantitative (présentées dans Mfr au chap. 3 « L'analyse descriptive » pp. 79-110)
Or cette distinction justifie la séparation : comparer des fréquences / comparer des moyennes (voir plus loin).
– Tests paramétrique / non paramétrique :

Test paramétrique : test qui repose sur des hypothèses concernant les distributions sous jacentes des variables étudiées (p. 381)

Test non paramétrique (basé sur les rangs) : un test est dit non paramétrique s'il ne fait aucune hypothèse sur la distribution du ou des paramètre(s) étudié(s). Ces tests sont, dans les mêmes conditions, moins puissants que les tests paramétriques (plus faible probabilité de détecter une différence ou un effet qui existe dans la population).

Par exemple, le test du χ^2 est un test non paramétrique.

III. 3.3. Le genre de tâche « Comparer ... »

•Mvn : … des proportions, des moyennes, des variances
– L'analyse de la variance est présente de façon spécifique dans le chap. VIII et dans le chap. IX consacré au test de Duncan
•Mfr : … des fréquences/proportions, des moyennes
– L'analyse de la variance est un *sous-type de tâche de la comparaison des moyennes* dans 2 cas :

• Comparer 2 moyennes observées pour de petits échantillons (n1 ou/et n2 < 30)

---

[3] « Le test du χ^2 a l'avantage de se généraliser à plusieurs proportions observées […]. Toutefois, de nombreux enseignants ou auteurs préfèrent une autre formulation du test de comparaison de deux proportions observées, spécifique à ce problème et basée sur la loi normale. Cette formulation conduit à un test équivalent au test du χ^2 […]. (Mfr p. 177)

• Comparer plus de 2 moyennes

Cependant soulignons que, au-delà de ces différences, les tests présentés sont à peu près identiques. Nous allons maintenant examiner plus en détail l'analyse de la variance dans les 2 manuels.

III. 4. Le cas de l'analyse de la variance

III.4.1. Analyse praxéologique - Mfr

« Comparer des variances » (Intitulé « Analyse de la variance ») est un sous type de tâches de 2 types de tâches « comparer les moyennes » Ta et Tb.

Ta « Comparer 2 moyennes observées pour de petits échantillons (n_1 ou/et $n_2 < 30$) » (p. 209)
$H_O : \mu_1 = \mu_2$
« Il est alors nécessaire de supposer de plus que la distribution de la variable étudiée suit une loi normale et que les variances inconnues sont égales à σ^2. Il faut donc vérifier, dans un premier temps, que les variances ne sont pas statistiquement différentes. » (p. 209) discours explicatif qui sera repris dans Tb.

Tb « Comparer plusieurs moyennes observées » (p. 211) ANOVA

Nous allons étudier la praxéologie autour de ce type de tâche (praxéologie ponctuelle) de façon détaillée *car ce type de tâche fait l'objet d'un chapitre spécifique (chap.. IX) dans Mvn.*

Technique, c'est-à-dire ce qui permet d'accomplir le type de tâche « Comparer plusieurs moyennes observées » (cf. Mfr, p. 216),
- Tracer le tableau *[aide aux calculs]*

E_i	E_1	E_2	…	E_k	Total
n_i	n_1	n_2	…	n_k	N
T_i	T_1	T_2	…	T_k	T
$\sum x_i^2$	$\sum_{j=1}^{n_1} x_{1j}^2$	$\sum_{j=1}^{n_2} x_{2j}^2$	…	$\sum_{j=1}^{n_k} x_{kj}^2$	$\sum_{i=1}^{k}(\sum_{j=1}^{n_i} x_{ij}^2)$

[Légende : **n_i** : « chacun des k échantillons E_i contient n_i observations, que nous noterons x_{ij}, dont la moyenne est m_i et la variance $\sigma_{X_i}^2$ » (p. 212) **T_i** : somme des valeurs de chaque groupe : $T_i = \sum_{j=1}^{n_i} x_{ij}$]

- Calculer les moyennes $m_i = \frac{T_i}{n_i}$;

- Calculer la somme des carrés résiduelles SCR [SCR = $\sum_{i=1}^{k}(\sum_{j=1}^{n_i} x_{ij}^2) - \sum_{i=1}^{k}\frac{T_i^2}{n_i}$] ;

- Calculer la somme des carrés entre colonnes SCC [SCC = $\sum_{i=1}^{k}\frac{T_i^2}{n_i} - \frac{\left(\sum_{i=1}^{k} T_i\right)^2}{N}$] ;

- Calculer $S^2 = \frac{SCR}{N-k}$;
- Calculer $S_C^2 = \frac{SCC}{k-1}$;
- Calculer $F = \frac{S_C^2}{S^2}$;
- Lire sur la table de Fisher pour $\alpha/2 = 2{,}5\%$;

- Comparer F et la valeur lue ;
- Conclure : H_0 (les moyennes sont égales) accepter ou rejeter
- Interpréter

Eléments technologiques, c'est-à-dire éléments qui permettent de justifier, comprendre et produire la technique précédente (cf. Mfr, pp. 212 – 215)
Nous citons ci-après les principales justifications présentes dans le manuel.

> Soit k échantillons E_i (i = 1, 2, …, k). Chacun de ces échantillons issus d'une population P_i dans laquelle la variable étudiée suit une loi normale de moyenne μ_i et de variance σ_i^2 inconnues.
> La première hypothèse que nous devons formuler est que les variances sont égales ($\sigma_i^2 = \sigma^2$, σ^2 désignant la variance commune). Un test statistique, celui de Bartlett par exemple, nous permet de tester l'hypothèse l'égalité des variances (…).
> Le problème à résoudre consiste ici à choisir entre deux hypothèses :
> H_0 : $\mu_1 = \mu_2 = ... = \mu_k$
> H_1 : 'au moins deux de ces moyennes différent'. (p. 212)
>
> La méthode qui va nous permettre de résoudre le problème est simple : nous allons calculer une estimation de la variance commune σ^2 dans deux cas :
> - sans tenir compte de l'hypothèse H_0
> - en supposant l'hypothèse H_0 vraie.
> Puis nous allons comparer ces deux estimations par le test de comparaison de deux variances. Si ces variances différent significativement, on rejettera l'hypothèse nulle et on conclura que deux moyennes au moins différent (H_1 est acceptée). (p. 213)
>
> $S^2 = \frac{SCR}{N-k}$ et $S_C^2 = \frac{SCC}{k-1}$ sont deux estimations de la variance :
> $S^2 = \frac{SCR}{N-k}$ est une estimation de la variance pour H_0 vraie ou non. Cette estimation, la meilleure, est obtenue en additionnant les numérateurs et les dénominateurs des $S_i^2 = \frac{\sum_{j=1}^{n_i} x_{ij}^2 - \frac{T_i^2}{n_i}}{n_i - 1}$. (p. 213)
>
> $S_C^2 = \frac{SCC}{k-1}$ est une estimation indépendante de la précédente, appelée variance entre colonnes. (p. 214)
> On compare ces deux variance en posant $F = \frac{S_C^2}{S^2}$. F suit une loi de Fisher à (k-1, N-K) degrés de liberté. (p. 215)

La théorie de cette technologie (qui permet de fonder les éléments technologiques précédents) est la théorie des probabilités et en particulier la distribution normale d'une variable aléatoire quantitative.

Dans Mfr, « l'analyse de la variance » ou « comparer des variances » a la signification de « Comparer 2 estimations d'une même variance inconnue » (en ayant auparavant pris la précaution de *tester l'hypothèse de l'égalité des variances des k échantillons indépendants*).

III.4.2. Analyse praxéologique - Mvn

« Comparer les variances » est un type de tâche.

Ce type apparaît dans le cours comme séparé du type de tâche « comparer des moyennes, » aussi bien dans le chap. VIII que dans le chap. IX, *alors que dans Mfr il est clairement un sous-type de tache de la comparaison des moyennes.*

Technique pour accomplir l'analyse de la variance (donnée dans le cours, Mvn p. 180)

Calculer :

1) $\overline{y_1} = \frac{1}{n_1}\sum_{j=1}^{n_1} y_{1j}$; $\overline{y_2} = \frac{1}{n_2}\sum_{j=1}^{n_2} y_{2j}$; … ; $\overline{y_k} = \frac{1}{n_k}\sum_{j=1}^{n_k} y_{kj}$;

On déduit : $\sum_{i=1}^{k} n_i \overline{y_i}^2$

2) $\sum_i \sum_j y_{ij}$. On déduit : $\overline{y} = \frac{1}{N}.\sum_{ij} y_{ij}$ et $N(\overline{y})^2$.

3) $\sum_i \sum_j y_{ij}^2$

4) $Q_f = \sum_i n_i(\overline{y_i} - \overline{y})^2 = \sum_{i=1}^{k} n_i \overline{y_i}^2 - N(\overline{y})^2$; $Q_r = \sum_i \sum_j (y_{ij} - \overline{y_i})^2 = \sum_i \sum_j y_{ij}^2 - \sum_{i=1}^{k} n_i \overline{y_i}^2$;

$s_f^2 = \frac{Q_f}{k-1}$; $s_r^2 = \frac{Q_r}{N-k}$; $F = \frac{s_f^2}{s_r^2}$

Tableau de résultats :

Source	TBPDL[4]	Degré de liberté	Variance	F
Entre les modalités du facteur A	Q_f	k-1	s_f^2	$F = \frac{s_f^2}{s_r^2}$
A l'intérieur de chaque modalités (hasard)	Q_r	N-k	s_r^2	

Eléments technologiques (p. 183-184) qui sont surtout des explicitations de notations

Soit : $\overline{y_1} = \frac{1}{n_1}\sum_{j=1}^{n_1} y_{1j}$; $\overline{y_2} = \frac{1}{n_2}\sum_{j=1}^{n_2} y_{2j}$; … ; $\overline{y_k} = \frac{1}{n_k}\sum_{j=1}^{n_k} y_{kj}$;

$N = \sum_{i=1}^{k} n_i$; $\overline{y} = \frac{1}{N}.\sum_{ij} y_{ij}$

$Q_T = \sum_i \sum_j (y_{ij} - \overline{y})^2 = \sum_i \sum_j y_{ij}^2 - N(\overline{y})^2$ Somme des carrés des écarts types. Elle exprime la différence entre les individus et la moyenne de la population.

$Q_f = \sum_i n_i(\overline{y_i} - \overline{y})^2 = \sum_{i=1}^{k} n_i \overline{y_i}^2 - N(\overline{y})^2$ Somme des carrés des écarts entre les modalités. Elle exprime la différence causée par les individus et le facteur A.

$Q_r = \sum_i \sum_j (y_{ij} - \overline{y_i})^2 = \sum_i \sum_j y_{ij}^2 - \sum_{i=1}^{k} n_i \overline{y_i}^2$ Somme des carrés de la différence pour chaque modalité. Elle exprime la différence causée par des individus (facteur hasard).

On déduit : $Q_T = Q_f + Q_r$ (*)

[4] Ces sont les initiales de la formulation vietnamienne « *TONG BINH PHUONG DO LECH* », soit en français : « somme des carrés des écarts types »

On calcule : $s_f^2 = \frac{Q_f}{k-1}$; $s_r^2 = \frac{Q_r}{N-k}$; $F = \frac{s_f^2}{s_r^2}$

- Si F est grand, c'est-à-dire s_f^2 est grand par rapport s_r^2 et donc Q_f est grand par rapport à Q_r. A partir de (*) on déduit que la différence entre modalités du facteur A a un effet sur le résultat.
- Si F est petit, c'est-à-dire s_f^2 est petit par rapport s_r^2 et donc Q_f est petit par rapport à Q_r. A partir de (*) on déduit que la différence entre modalités du facteur A n'a pas d'effet sur le résultat.

On pose l'hypothèse H : La différence des moyennes $\overline{y_1}, \overline{y_2}, \ldots, \overline{y_k}$ n'est pas significative (c'est-à-dire le facteur A n'a pas d'effet sur le résultat). D'après l'hypothèse H, F a la distribution de Fisher (k-1 ; N-K)). On a le test suivant :

- Si $F \leq F_{1-\alpha}$ (k-1, ; N-k) on accepte H, c'est-à-dire le facteur A n'a pas d'effet sur le résultat, au seuil d'erreur α.
- Si $F > F_{1-\alpha}$ (k-1, ; N-k) on rejette H, c'est-à-dire le facteur A a un effet sur le résultat.

Nous pouvons faire les deux commentaires suivants.

1. les variances $s_f^2 = \frac{Q_f}{k-1}$; $s_r^2 = \frac{Q_r}{N-k}$ ne sont jamais présentées comme *des estimations de la variance commune inconnue de la population dont sont issues les observations,* bien que l'on puisse lire (p. 186) :

> • ***Attention*** : Les variances s_1^2, s_2^2, …, s_k^2 ne sont pas significativement différentes. Avant d'effectuer ANOVA on doit vérifier les variances par le test de Bartlett ou Cochran, car quand on pose l'hypothèse H on considère que les y_{ij} suivent la loi normale N(μ_0 ; σ^2).
>
> (*Les méthodes Statistiques en Pharmacie et Chimie, J.Philippe, p. 125*) (Mvn, p. 186)

2. Implicitement l'analyse de la variance est bien « au service » de la comparaison des moyennes, mais cela reste non affirmée. En effet, dans les exemples traités, l'hypothèse H porte sur l'égalité des moyennes et dans la rubrique exercices (p. 195) est écrit : « Comparer plusieurs moyennes – ANOVA ». Les conséquences sont que les conditions de l'utilisation de l'analyse de la variance dans les tests d'hypothèse risquent d'être obscures pour les étudiants.

IV. Entretiens avec des enseignants de mathématiques de l'USM

Les objectifs de ces entretiens sont les suivants :
– Repérer les objectifs pédagogiques du module de statistiques
– Comprendre les difficultés propres à cet enseignement et les difficultés perçues chez les étudiants
– Evaluer avec les enseignants les possibilités de changement dans leur manière d'enseigner

Nous nous sommes entretenus durant plus d'une heure avec 6 des 7 enseignants de mathématiques à l'USM (un enseignant a refusé l'entretien). Ces entretiens ont été enregistrés et dépouillés sous forme de 6 fichiers. Ces 6 fichiers ont été rassemblés question par question en un seul fichier pour permettre l'analyse. L'analyse sera faite au retour en France.

Les entretiens et le dépouillement ont nécessité plus de 30 h de travail d'Annie Bessot, Lê Thi Hoai Chau, Vu Nhu Thu Huong et Dao Hong Nam.

V. Entretiens avec des médecins (de l'USM ou extérieurs à l'USM)

Les objectifs de ces entretiens sont les suivants :
- lister les *besoins professionnels* en statistique, en particulier autour de *l'étude épidémiologique* ou de *l'essai thérapeutique*
– contribuer à la modification du rapport des enseignants de mathématiques à l'objet statistiques pour les rapprocher des besoins des médecins

–donner du sens aux statistiques par leur fonctionnement dans une situation professionnelle *authentique*
Par exemple, un des besoins reconnus dans la formation médicale française est celui de pouvoir porter un regard critique sur la littérature médicale.

Nous nous sommes entretenus entre 30 et 45 minutes avec 9 médecins. Ces entretiens ont été enregistrés et dépouillés sous forme de 9 fichiers. Ils ont été rassemblés question par question en un seul fichier pour permettre l'analyse. L'analyse sera faite au retour en France.

Les entretiens et le dépouillement ont nécessité plus de 35 h de travail d'Annie Bessot, Lê Thi Hoai Chau, Vu Nhu Thu Huong et Dao Hong Nam.

Perspectives

1. Perspectives : au retour en France

1.1. Analyse des entretiens des enseignants de mathématiques et des médecins ;

1.2. Conception d'un questionnaire auprès des étudiants en médecine des deux pays autour des principaux types de tâche qui ont été repérés lors de l'analyse praxéologique des manuels.

2. Perspectives plus lointaines

2.1. Colloque International franco-vietnamien de didactique des mathématiques prévu en janvier 2013 au Viêt Nam à Ho Chi Minh ville. Deux thèmes (sur trois) offrent la possibilité d'exposer le travail de la coopération :

- le thème 1 « Les probabilités et les statistiques dans l'enseignement secondaire, dans l'enseignement supérieur et dans la formation des enseignants.»
- le thème 2 « Interactions entre mathématiques et autres disciplines scientifiques dans l'enseignement général ou professionnel : le cas particulier de la formation des médecins... ».

2.2. Cahier des charges pour l'élaboration d'un environnement informatique : s'appuyant sur l'analyse comparative précédente, trois temps de recherche et de rédaction :

- Mission médicale conduisant les étudiants à se poser des problèmes à résoudre à l'aide de savoirs statistiques ;
- Pour répondre à cette mission, environnement informatique donnant accès à des espaces d'activités définis grâce à l'analyse comparative et plus particulièrement à l'enquête auprès des enseignants ;
- Scénarios possibles d'usage.

2.3. Élaboration d'un site Loé pour l'USM **si des enseignants de mathématiques et des médecins de l'USM acceptent d'y collaborer.**

2.4. Diagnostic des erreurs et des connaissances des apprenants en statistique pour mettre en place des rétroactions informatiques à destination des étudiants.

En guise de conclusion

Je conclurais par une citation extraite de la préface du manuel français : elle me semble caractériser une raison d'être d'un enseignement de statistique en médecine comme réponse à l'incontournable variabilité du fonctionnement de l'homme et par là même des patients.

> Les étudiants de première année anticipent rarement que c'est avec l'enseignement de la biostatistique qu'ils vont avoir leur première prise de contact réelle avec la médecine : celle-ci a en effet pour nature d'être confrontée, à toutes ces étapes, avec le fait qu'un patient n'est jamais

identique à un autre. Si les patients étaient identiques entre eux il n'y aurait pas besoin de biostatistique – mais il n'y aurait pas de médecine non plus : un médecin disposant de la 'notice de fonctionnement de l'homme' y suffirait. (Mfr, préface pp. 5-6)

Bibliographie

Benichou J., Beusccart R., Roy P., Quantin C., avec le Collège national des enseignants des facultés de médecine (2009). *Biostatistique*. Ed. Omniscience.

Bessot A., Comiti C. (2008). Apport des études comparatives aux recherches en didactique des mathématiques : le cas Viêt-Nam / France, *Actes du séminaire national de didactique des mathématiques*, Année 2008, p.171-194, Editions ARDM et IREM de Paris 7.

Bosch, M., Chevallard, Y. (1999). La sensibilité de l'activité mathématique aux ostensifs. Objet d'étude et problématique. *Recherche en Didactique des Mathématiques*, 19/1, 77- 124.

Chaachoua, H., C. Comiti (2007). L'analyse du rôle des manuels dans l'approche anthropologique. *2e Congrès sur la Théorie Anthropologique du Didactique*, Uzès, France.

Chevallard, Y. (1992). Concepts fondamentaux de la didactique : perspectives apportées par un approche anthropologiques. *Recherches en Didactiques des Mathématiques*, 12/1, 83- 121.

Chu Dinh Thanh, Pham Minh Buu, Tran Dinh Thanh, Nguyen Van Lieng (2011). *Probabilité Statistique*. Presses Universitaires de l'Université des Sciences Médicales d'Ho Chi Minh ville.

Croset M-C, Ney M., Gonçalves C., Balacheff N., Schwartz C., Bosson J-L. (2011). Evaluation of immersive experiences in a medical game: student exchanges with characters of the game. *Simulation and gaming : an interdisciplinary journal*.

Vu Nhu T.H. (2009). *Etude didactique sur l'introduction dans l'enseignement mathématique vietnamien de notions statistiques dans leurs liens avec les probabilités.* Thèse en co-tutelle France-Viêt Nam