Maîtrise Sciences Co : Projet noté 2

1 Définition du traitement des données

L'objectif de ce sujet est d'automatiser l'analyse statistique des résultats obtenus à l'issue d'une expérience de psychologie sur l'amorçage intermodal. Chaque participant était confronté à 7 conditions expérimentales, dans lesquelles l'intervalle inter-stimuli (ISI en anglais) était manipulé. Pour chacune des conditions, 100 essais ont été réalisés. 20 sujets ont pris part à l'expérience. Vous disposez de 7 fichiers (ISI[1...7].txt), présentant les résultats moyens obtenus pour chaque sujet dans chacune des conditions. Le nombre de lignes par fichier est variable, on a retiré un certain nombre de données pour tenir compte d' erreurs lors de l'expérience. Chaque fichier se compose de 4 colonnes, les deux premières colonnes indiquent le taux de reconnaissance (pourcentage) et le temps de réponse (en seconde) pour la condition de congruence, les deux dernières colonnes indiquent les résultats correspondant pour la condition de non-congruence.

On s'intéresse uniquement à l'effet de congruence pour chacune des valeur d'ISI. Nous souhaitons procéder à un test t pour comparer les moyennes des échantillons appariés. On rappelle la procédure :

- Constitution d'un échantillon de différences : pour chaque sujet, on calcule l'écart entre la performance en condition "congruente" et en condition "non-congruente"
- Calcul des paramètres définissant l'échantillon des différences (moyenne $\overline{X}$ et écart-type S) pour N sujets.
- Calculer le t correspondant
$$t = \frac{\overline{X} * \sqrt{N}}{S}$$
- Lire dans la table des t (tTable.txt, fourni) la valeur limite de t, pour un degré de liberté (N-1) et un risque α donné.
- Comparer la valeur de t calculé et la valeur lue dans la table du t. Si le t calculé est supérieur (en valeur absolue) au t lu, l'écart entre les moyennes est significatif

2 Etapes de développement

2.1 Initialisation du programme

- Définition de l'architecture de fichiers (main.c, define.h + makefile).

- Initialisation d'un tableau double** pour le stockage et la gestion des données. La taille maximale du tableau (en fonction du nombre de sujets max par fichier de données) est demandée à l'utilisateur avant allocation.

2.2 Lecture des données

- Création d'une méthode permettant d'ouvrir un fichier (parmi ISI[0...7].txt, fournis) et de remplir le tableau de données en lisant la première ligne du fichier. Attention, nécessite de faire une converstion entre int et double pour le pourcentage de reconnaissance.
- Lecture d'un fichier complet (ISI1.txt ayant 20 lignes), en le remplissant ligne par ligne.
- Stockage du nombre de lignes lues.
- Gestionnnaire de fichiers, pour lire plusieurs fichiers de manière incrémentale.
- Remplissage d'un tableau comportant la table en t (tTable.txt, fourni). Attention, la première ligne est un cas particulier qui indique le risque α correspondant à chaque colonne.

2.3 Calcul du test t

- Définition des méthodes nécessaires au calcul du test t dans mathTools.c (moyenne, écart-type).
- Interrogation de l'utilisateur sur le risque α souhaité et lecture dans le tableau du t limite en fonction de degré de liberté (déterminé au moment de la lecture du fichier par le nombre de tests réellement lus)

2.4 Ecriture des résultats

- Décision sur la significativité des résultats en fonction du t limite, et écriture des paramètres motivant la décision dans un fichier texte.
- Traitement séquentiel des fichiers ISI[0...7].txt et affichage des résultats dans le même fichier texte.

2.5 Test du programme

- Approximation des résultats multifacteurs : significativité sur l'ensemble des résultats...
- Modification "manuelle" du nombre de lignes dans les fichiers et influence de la diminution du nombre de lignes sur les résultats