

Consignes pour l'Oral de TP STA401

Tous les documents et matériels de votre choix (Rstudio, calculatrice et tables statistiques) sont bien sûr autorisés et entre le moment où est donné le sujet qui sera constitué d'une ou deux questions à 1, 2 ou 3 points selon le choix du candidat (voir précisions plus loin) le temps de planche et restitution de la réponse est de **10 minutes**. On procèdera de la façon suivante : Un salon est prévu par groupe de 12 d'une durée de 60 minutes. Le groupe et lien de connexion de chaque groupe de 12 sera envoyé par mail la veille du passage de l'oral (pas de panique si vous avez oublié votre numéro de binôme la liste des noms des 12 candidats de chaque groupe sera précisée dans le mail contenant le lien de connexion au salon).

Binômes du lundi 11 mai

- MIN3 (+1 binôme INF1) : lundi 11 mai 9h30-10h30 binômes num. 1 à 6 (jury FL&CM) + INF2 binôme 11 : 12 étudiants
- MIN3 (+2 binôme INF2) : lundi 11 mai 10h30-11h30 binômes num. 7 à 12 (jury FL&CM) INF1 binômes 14(2) et 15(2?) : 12 à 14 étudiants
- INF1 : lundi 11 mai 11h30-12h30 binômes num. 1 à 6 (jury FL&CM) : 10 à 12 étudiants
- INF1 : lundi 11 mai 12h30-13h30 binômes num. 7 à 13 sauf 11 (jury FL&CM) : 12 étudiants
- **derniers binômes des groupes INF1+INF2** : lundi 11 mai 13h30-14h30 binômes num. 14(2) et 16(2) de INF1 et binômes 2(2?), 6(2), 15(2) et 16(1) de INF2 (jury FL&CM) et Khial A. de INF4 : au plus 12 étudiants
- INF2 : lundi 11 mai 14h30-15h30 binômes num. 1 à 7 et pas 2 (jury FL&BL) : 12 étudiants
- INF2 : lundi 11 mai 15h30-16h30 binômes num. 8 à 13 (jury FL&BL) : 12 étudiants
- INF3 : lundi 11 mai 16h30-17h30 binômes num. 1 à 6 (jury FL&BL) : 12 étudiants
- INF3 : lundi 11 mai 17h30-18h30 binômes num. 7 à 12 (jury FL&BL) : 12 étudiants

Binômes du mardi 12 mai

- INF4 : mardi 12 mai 9h30-10h30 binômes num. 1 à 6 (jury FL&CM) : 12 étudiants
- INF4 : mardi 12 mai 10h30-11h30 binômes num. 7 à 12 (jury FL&CM) : 12 étudiants
- MIN2 : mardi 12 mai 11h30-12h30 binômes num. 1 à 6 (jury FL&BL) : 12 étudiants
- MIN2 : mardi 12 mai 12h30-13h30 binômes num. 7 à 12 (jury FL&BL) : 12 étudiants
- **derniers binômes des groupes INF3+INF4+MIN2** : mardi 12 mai 13h30-14h30 binômes num. 13 à 15 de INF3 et INF4 (jury FL&BL) et 14 de MIN2 : 13 étudiants

Binômes du mercredi 13 mai

- MAT1 : mercredi 13 mai 17h00-18h00 binômes num. 1 à 6 (jury FL&CM) : 12 étudiants
- MAT1 + MIN2 : mercredi 13 mai 18h00-19h00 binômes num. 7 à 12 (jury FL&CM) et 13 de MIN2 : 13 étudiants

Vous pourrez éventuellement (avec une bonne raison et pas seulement par confort) changer de créneau à condition de faire un échange de créneaux avec un étudiant d'une autre session et à condition de l'avoir signalé avant dimanche 10 à 18h.

Les urnes de questions possibles (attention les questions 5 et 6 de l'EX3 sont passées dans C) :

- Urne A (questions à 1 point) :
EX1 questions 1 à 7, 10, 13 ou 14 (donner 2 appariés et 2 indépendants dans 13 et l'effet étudié pour 1 apparié et pour 1 indépendant dans 14)
EX2 questions 1 ou 2 pour un seul échantillon
- Urne B (questions à 2 points) :
EX2 questions 3 à 8
EX3 questions 1 à 4
EX4 questions 7 à 8
- Urne C (question à 3 points) :
EX3 questions 5 à 8
EX4 questions 1 à 4

Déroulement de l'épreuve :

Chaque étudiant à son tour de passage choisira **au plus** ou ou deux urnes de tirage de questions et dont le cumul de points possibles est **au maximum 4**. Par exemple s'il indique AA il aura deux questions à 1 point, s'il indique AB il aura une question à 1 point et une question à 2 points, s'il indique AC une question à 1 point et une question à 3 points ou dernière combinaison possible BB deux questions à deux points. Impossible de demander 3 questions à 1 point. Et si vous choisissez 2 questions d'une même urne elles ne se ressembleront pas (par exemple un même candidat n'aura pas une question sur le lycée A et la même question sur le lycée B).

Ensuite **les questions tirées dans les urnes seront toutes légèrement modifiées** afin que soit mise en oeuvre une procédure de calcul par chaque candidat et pas seulement restitués des résultats préparés à l'avance. Par exemple pour toutes les questions de statistiques un 27^{ième} individu tiré au hasard sera rajouté dans l'échantillon et pour les questions de probabilité un paramètre sera modifié. Voici quelques exemples de questions issues des trois urnes :

- Urne A :
EX1 question 1 avec une Normale d'écart-type 2 au lieu de 3 (les autres paramètres seront inchangés).
EX1 questions 4 avec un ordre $p = 10\%$ au lieu de $p = 2,5\%$.
EX2 questions 1 avec l'individu supplémentaire num. 27 venant du lycée A et ayant eu 15 en maths et 18 en français.
- Urne B :
EX3 question 7 avec l'individu supplémentaire num. 27 venant du lycée A et ayant eu 15 en maths et 18 en français.
- Urne C :
EX3 question 8 avec l'individu supplémentaire num. 27 venant du lycée A et ayant eu 15 en maths et 18 en français.

Pour les étudiants convoqués sur une même session :

Les étudiants d'un créneau se connecteront dès le début de leur session et seront interrogés dans l'ordre de leur numéro de binôme et dans l'ordre alphabétique de leur nom dans chaque binôme. Votre numéro de binôme est indiqué dans le fichier de votre groupe de TP ci joint. Les cinq premières minutes de la session seront utilisées pour rappeler à tous les candidats de la session le déroulement de l'oral et mode de fonctionnement de la discussion publique. Les quinze minutes suivantes seront interrogés les quatre candidats (en simultané) des deux premiers binômes, puis le quart d'heure suivant quatre autres candidats en même temps (les deux binômes suivants),... Chacun à son tour devra à la demande d'un des membres du jury indiquer dans son premier post un des choix suivants : AA; AB; AC ou BB. Il recevra alors dans la minute suivante un post du jury avec deux questions choisies dans les urnes qu'il aura indiquées et **disposera à partir de l'heure du post de ses deux questions, de dix minutes pour retourner ses réponses**. Il faudra *bien préciser* dans les **réponses données en un seul post** Q1 : ... pour la réponse à la question tirée dans la première urne et Q2 : pour la réponse à la question tirée dans la seconde urne. **Relisez vous avant d'envoyer votre post**. Vous ne pourrez modifier le post une fois envoyé. Une fois les dix minutes écoulées l'étudiant sera sorti du salon. **Important : lorsque vous vous connectez renseignez votre nom de famille comme identifiant et choisissez le mode écoute seul.**

La réponse donnée pour chaque question devra être précise mais concise (un exemple de question réponse est proposé à la fin de ce document). Vous devrez y indiquer : les conditions requises pour appliquer la procédure utilisée (dans le cas des IC ou des tests) les valeurs numériques obtenues (dans l'ordre de la question s'il s'agit d'estimations sans biais de paramètres ou de bornes d'intervalles de confiance et pour les tests seulement la pval); l'hypothèse alternative testée dans le cas des tests soit H1 (inutile de préciser H0 qui se déduit automatiquement de H1 puisque c'est sa complémentaire) et une conclusion littérale en français indiquant le risque d'erreur.

De plus il est impératif d'utiliser les noms de variables et notations pour leurs paramètres définies dans le script R ci-joint et rappelées à la fin de ce document. Les valeurs numériques seront données à 10^{-2} pour les moyennes et écart-type à 10^{-3} pour les proportions, et pour les p-valeurs à 10^{-3} (par ex : pval = 0.104) ou en écriture scientifique si elle est inférieure à 10^{-3} (par ex : pval = $3,2 \cdot 10^{-6}$).

Avant l'épreuve pour qu'elle se déroule au mieux il est aussi conseillé à chacun d'avoir une version imprimée de l'énoncé du TP où auront été surlignées dans des couleurs différentes les questions selon leur urne. Ses tables statistiques, un script R solution du TP ou des notes précises sur les fonctions à employer si vous travaillez avec une calculatrice. D'avoir testé le chat et les passages à la ligne dans un seul post. Les salons de lundi, mardi et mercredi seront ouverts une heure dimanche à partir de 14h pour ceux de lundi, 16h pour ceux de mardi et 17h pour ceux de mercredi. Entraînez-vous entre vous à l'avance (faites comme nous l'avons fait avant vous... un jeu de rôles !).

Pendant l'épreuve il est aussi conseillé de rédiger vos réponses sur un document word et sa rédaction terminée de retourner vos réponses avec un copié collé de votre document word dans le post de la discussion.

Si vous n'avez ni la 4G ni la wifi nous ferons l'oral par téléphone et dans ce cas vous devrez me retourner un numéro où vous joindre la veille de l'épreuve.

Vous utilisez l'outils de calcul de votre choix : vous devriez tous savoir utiliser R et/ou la calculatrice mais si vous préférez travailler avec un autre logiciel ou un tableur c'est à vous de voir.

Cet oral n'est pas obligatoire dans le sens où si vous ne le passez pas vous n'aurez pas de points en plus que la somme des deux colles multipliée par deux pour votre CC de TP. Dans ce cas si vous êtes surs de ne pas souhaiter le passer merci de me prévenir au plus vite que j'ajuste les effectifs des groupes.

Notations :

- X_A variable aléatoire modélisant la note de maths d'un candidat venant du lycée A, μ_{X_A} sa moyenne théorique et σ_{X_A} son écart-type théorique.
- X_B variable aléatoire modélisant la note de maths d'un candidat venant du lycée B, μ_{X_B} sa moyenne théorique et σ_{X_B} son écart-type théorique.
- Y_A variable aléatoire modélisant la note de français d'un candidat venant du lycée A, μ_{Y_A} sa moyenne théorique et σ_{Y_A} son écart-type théorique.
- Y_B variable aléatoire modélisant la note de français d'un candidat venant du lycée B, μ_{Y_B} sa moyenne théorique et σ_{Y_B} son écart-type théorique.
- p_A la proportion de candidats venant du lycée A dans la population des candidats des deux lycées et $p_B = 1 - p_A$.
- p_{sA} la proportion de candidats venant du lycée A obtenant une note de maths supérieure à 12 (> 12) et p_{sB} lorsque le candidat vient du lycée B.

Script R et autres documents:

Un script R permettant d'extraire correctement les échantillons après avoir chargé les données est également fourni avec les solutions des questions traitées lors de la dernière séance de TP. Vous trouverez aussi dans ce script toutes les informations permettant d'enrichir facilement les échantillons d'un individu supplémentaire. L'enregistrement et une sauvegarde de la discussion publique associée d'une séance en visio est aussi fourni.

Un exemple de post pour l'oral passé par Frédérique Leblanc avec Cécile Martinet comme jury1 :

Martinet (jury 1) : Frédérique ton choix ?; Frédérique Leblanc : Sujet AC ; Martinet 20h15 : Frédérique : Q1:..... Q2:.....; Frédérique Leblanc 20h20: Q1 : (saut de ligne dans le poste avec touches majuscule+entrée) Q2:.....

Ou mieux peut-être, préparer ses réponses dans un bloc notes *se relire... puis envoyer ses deux réponses dans le même post*. Voici ce que ça donne (ne tenez compte que de votre nom et de Q1 et Q2; sujet AC10 est une indication repère uniquement destinée au jury):

Ma

Martinet 20:15
Frédérique : sujet AC10
Q1 : EX1 question 1 avec l'intervalle $[-1, 2]$ au lieu de $[-1, 5]$
Q2 : EX4 question 1 avec l'individu supplémentaire ayant eu 10 en maths et 18 en français et venant de lycée A.

Fr

Frederique Leblanc 20:15
Sujet AC

Fr

Frederique Leblanc 20:20
Q1 : 0.34
Q2 : comp de moy ech. appariés
hyp modele : $DA = X_A - Y_A$ de loi normale de moyenne théo $\mu_{X_A} - \mu_{Y_A}$
 $H1 : \mu_{X_A} - \mu_{Y_A} > 0$; $pval = 0.103$
A moins d'un risque d'erreur supérieur à 10,3% on ne peut déclarer les candidats du lycée A meilleur en moyenne en français qu'en maths

Evaluation :

Chaque membre du jury donne de 0 à 2 points pour AA; 0 à 3 points pour AB et de 0 à 4 points pour BB ou AC. Note d'oral = note jury 1 + note jury 2.

Dans l'exemple, Q1 est juste et rapporte 1 et Q2 est presque juste : en effet FL dans sa précipitation a répondu "...on ne peut déclarer les candidats du lycée A meilleur en moyenne en *français qu'en maths*" alors que son hypothèse H1 étant : $\mu_{XA} > \mu_{YA}$, elle aurait dû répondre "...on ne peut déclarer les candidats du lycée A meilleurs en moyenne *en maths qu'en français*" (ou "on ne peut, conclure H1 (ou de façon équivalente rejeter H0), c'est à dire que le résultat en maths est en moyenne meilleur qu'en français de façon significative puisque la pval est assez élevée, ici $> 10\%$ "). *Et oui ! Elle a été trop vite et n'a pas pris le temps de se relire.....* Pour cette seconde réponse elle n'obtiendra donc que 3,5 points. Et si les deux membres du jury sont d'accord sa note d'oral sera de 7 points qui se rajouteront aux points cumulés de ses deux colles eux même multipliés par deux. Si elle avait eu (*colle1, colle2*) = (3, 2) sa note finale de TP sera de 17.

Barème indicatif :

Dans les questions de test : 1 point pour la conclusion littérale si elle est juste et complète (-0.5 si le risque dans la conclusion ou sa significativité n'est pas mentionnée).

Pour les questions de l'urne C : 1 point si le problème est correctement posé (hyp. de modélisation et H1) et 1 point pour la p-val.

Pour les questions de l'urne B, si IC ou test 0.5 point pour les hyp de modélisation, 0.5 point pour pval et 1 pour conclusion si c'est un test, 1.5 pour les bornes si c'est un intervalle de confiance.

Pour l'urne A, 1 ou 0 pour les questions de proba, et 0.5 par val numérique des stats demandées si c'est une question de stat (sauf pour la question 10 de EX1 où sont demandées 4 valeurs : 0.25 point pour chaque).