

# VAGUE DE QUESTIONS POUR LE PROFESSEUR STACCINI :

## INTRODUCTION AUX MODÈLES MULTIVARIÉS

1. Un étudiant n'a « pas compris les formules dans la régression logistique ». Avez-vous une réponse à lui apporter ?

Non pas de réponse, sauf qu'il faut bien comprendre le principe d'une analyse multivariée, la fonction logit (voir question 11)

2. Un étudiant ne comprend pas ce qu'est la statistique inférentielle ?

C'est la suite de la statistique descriptive, c'est la statistique qui permet de tirer des conclusions à partir de comparaisons par exemple

3. Un étudiant ne comprend pas ce que représente X (en tant que variable explicative) ?

$Y = f(X)$  : Y variable expliquée, X variable explicative

Y = réussite au concours OUI.NON

X = nombre d'heures à travailler les cours

4. Un étudiant ne comprend pas ce qu'est la droite d'ajustement ?

C'est une droite qui passe au sein d'un nuage de point. Elle est dite la droite des moindres carrés car elle a pour propriété de faire en sorte que la somme des distances de chaque point observé à la droite (au carré) soit la plus petite possible

5. Un étudiant demande des explications sur la fonction de régression en elle-même (« je pensais qu'on affichait des valeurs puis qu'on traçait une droite qui "regrouperait" ces valeurs pour trouver alpha et bêta, mais on nous dit ensuite que alpha et bêta sont estimés pour avoir le plus petit epsilon possible... ») ?

Oui, c'est exact !

6. Un étudiant se demande ce que représente le pourcentage de variance expliquée ?

C'est la part de variance (c'est-à-dire la dispersion) expliquée par l'ajustement (coefficient de détermination R au carré)

7. Un étudiant n'a pas compris ce qu'était l'odds ratio ou rapport de cotes ?

La cote et non la côte !

La cote est le rapport entre une probabilité et son contraire :  $p / (1 - p)$

Dans une enquête cas témoin on fait le rapport de cote de l'exposition chez les malades avec la cote de l'exposition chez les non malades. On montre que l'OR est une bonne estimation du RR dans le cas de maladies rares

8. Un étudiant se demande ce que représentent les données censurées (« estimation de la survie, qui est un autre cours mais on nous en parle aussi ici... ») ?

Perdu de vue, exclus vivant, ce sont les patients que je ne vois pas à la date de point et qui n'ont pas eu l'événement considéré

9. A propos de l'ACP, un étudiant ne comprend pas pourquoi « si on a p variables, il existe  $p(p+1)/2$  corrélations possibles » ?

C'est classique ! partir d'un exemple : 3 variables A, B, et C, il y a 6 corrélations possibles : AB, AC, BA, BE, CA et CB :  $3 \times (3+1) / 2 = 6$

10. Un étudiant ne comprend pas ce qu'est le cercle des corrélations ? Pouvez-vous lui fournir des explications supplémentaires s'il vous plait ?

Pas simple à expliquer en une phrase ! le doc suivant explique pas mal.

<https://openclassrooms.com/fr/courses/4525281-realisez-une-analyse-exploratoire-de-donnees/5278723-interpretez-le-cercle-des-corrrelations>

11. Les étudiants se demandent sur quelles notions du cours « Introduction aux modèles multivariés » sont-ils susceptibles d'être interrogés à l'examen ?

Le principe de l'analyse multivariée, les conditions d'utilisation, la fonction logit, variable binomiale et multinomiale, variance expliquée coefficient de détermination, adéquation du modèle

## **SANTÉ NUMÉRIQUE**

12. Les étudiants se demandent si l'épreuve d'ECUE8 (Santé Publique et Santé Numérique) sera bien sous forme de QCMs (et non pas de QRUs comme l'ECUE7) ?

# **TOUT EST EN QRU ! ECUE6, ECUE7 et ECUE8**

13. A propos du cours sur les dossiers médicaux, un étudiant demande pourquoi les documents nominatifs sont-ils détruits ? («Je pensais avoir compris que tout devait être stocké pour un temps déterminé...»)

Une fois le temps déterminé passé, que fait-on ? on les détruit !

14. Un étudiant a du mal à différencier le rôle des différents règlements (RGPD, IFL), "en effet, ils me semblent très similaires..." ? Pouvez-vous faire un récap sur ces deux lois, et indiquez si elles sont bien à différencier ou si l'une complète seulement l'autre ?

C'est l'histoire qui fait la différence : on a commencé par l'IFL en 1978, puis d'autres lois sont venues compléter en particulier par rapport à la recherche clinique, puis l'europe est arrivée et a apporté de nouvelles contraintes expliquées dans les dernier slides : il revient au transporteur de la donnée d'assurer la qualité du transport, etc. accountability ! je dois rendre compte.

15. Qu'est-ce que le droit à l'oubli ?

Réponse tutrice : Le droit à l'oubli est un droit qui permet à tout internaute de demander le déréférencement d'une ou de plusieurs pages contenant des informations sur lui. Une personne physique a donc le droit de demander à un moteur de recherche de supprimer de la liste des résultats des liens pointant vers des pages contenant ses données personnelles une fois que l'on saisit son nom dans la barre de recherche. Avec la mise en application du RGPD en 2018 ce droit a été renforcé.

«16. On nous parle d'un délai de réflexion de 48h, s'agit-il d'un délai qu'a le patient pour consulter son dossier ? C'est l'explication que je me suis donné mais je pense qu'elle est fausse... »

Réponse tutrice :

Le délai de réflexion de 48h est un délai qu'on laisse au patient pour être sur de sa décision, de voir son dossier. Par exemple tu as un patient, il a une maladie assez grave. Un jour il devient un peu parano et il demande à voir son dossier car il a l'impression qu'on lui cache un truc. A partir du moment de sa demande on va laisser un délai de 48H pour qu'il soit sur de lui. Si le patient a « fini » sa crise de parano d'ici 48H et qu'il ne veut finalement plus voir son dossier, on aura pas remué ciel et terre pour avoir un dossier pour rien. Mais si après ces 48H il a toujours l'envie de voir son dossier et bien on donnera dans les délais indiqués dans la suite du cours (<8j pour un dossier de moins de 5ans et <2mois pour un dossier de plus de 5ans).

17. « J'ai du mal à différencier le rôle des différents règlements (RGPD, IFL), en effet, ils me semblent très similaires... » ?

Réponse tutrice : Il n'y a pas de différences à proprement parler entre le Règlement Général sur la Protection des Données et la loi Informatique et Libertés, le premier complétant désormais la seconde.

Toutefois, le RGPD amène plusieurs avancées majeures. D'une part, il crée un cadre unifié pour l'ensemble des États membres de l'UE et offre donc une même protection aux Européens quant au traitement de leurs données personnelles. D'autre part, ce nouveau cadre légal simplifie les règles auxquelles sont soumis les acteurs concernés par le RGPD, tout en assurant aux consommateurs une protection optimale de leurs informations personnelles.

Le RGPD attend des acteurs économiques une responsabilisation quant à son application. L'ancien système de contrôle, fondé sur l'appréciation a posteriori du respect des règles en vigueur, n'a plus lieu d'être. Désormais, chaque entreprise concernée par le RGPD doit prouver sa mise en conformité immédiate avec ce cadre juridique.