Estimation

Instructions:
Chaque dia contient une question. Réfléchissez à la réponse qui vous semble juste, ensuite cliquez pour que la réponse apparaisse.

Rappel: l'examen se fera sous forme de questions à choix multiple (QCM)

Une étude examine les prédicteurs de la capacité vitale (volume d'air maximal qu'on peut expirer après une inspiration maximale: VEMS) chez des fumeurs. La variable indépendante est la consommation journalière de cigarettes (nombre de cigarettes). L'association entre la quantité de tabac et le VEMS (exprimé en litre) est analysée par une régression linéaire. On trouve un intercept à 3 et une pente à -0.05 (significativement différente de 0). D'après ce modèle:

1) Est-ce que le VEMS moyen augmente ou diminue lorsque la consommation de tabac augmente ? Pourquoi ?
2) Quel est la valeur moyenne de VMS prédite pour une personne consommant 10 cigarettes par jour ?
3) De combien change le VEMS moyen lorsque la consommation journalière de cigarettes est incrémentée de 10 cigarettes ?

1. La pente de la droite de régression (-0.05) est négative donc le VEMS moyen diminue lorsque la consommation de tabac augmente
2. L'équation de la droite de régression est :

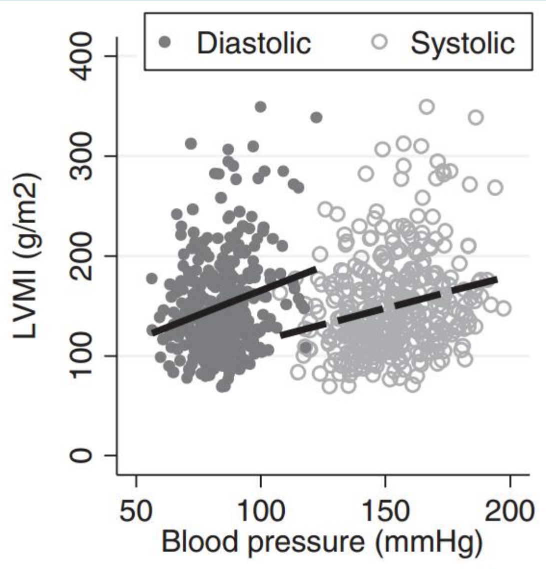
   VEMS = 3 - 0.05 * Nombre de cigarettes / jour

   En remplaçant dans cette équation le nombre de cigarettes par 10. on obtient la valeur moyenne de VEMS prédite par le modèle linéaire:

   3 – 0.05*10=3 – 0.5 = **2.5**

3. La pente de la droite de régression indique que le VEMS diminue en moyenne de 0.05 litres lorsque la consommation de tabac est incrémentée de 1 cigarette. Donc pour un incrément de 10 cigarettes, la baisse de VEMS est en moyenne de 0.05x10 = 0.5 litres

Dans un article, des chercheurs s'intéressent à la relation entre la tension artérielle et l'index de masse ventriculaire gauche (LVMI). Ils utilisent des modèles de régression linéaire simples dont les résultats sont les suivants:



Pente des droites de régression pour 10 mmHg de tension artérielle:
- diastolique: 10.1, 95 % IC: 5.1 to 15.0
- systolique: 7.3, 95 % IC: 4.3 to 10.2

1. Au vu des scatter plots, vous pouvez affirmer que les coefficients de détermination r^2 de ces 2 modèles sont plutôt proches de 0,10 ou de 0,90 ?

1. Pour que le coefficient r^2 soit proche de 1, il faut que les points du scatter plot soient proches de la droite de régression (cf diapo 41 du cours 6). Ici, les points sont très dispersés autour des droites de régression, les coefficients r^2 sont donc faibles. Ils sont plus proches de 0,10 que de 0,90.

On suppose que le coefficient de détermination r^2 vaut 0,09 pour le modèle de régression simple avec la tension artérielle diastolique de l'exercice précédent. Vous pouvez affirmer que (vrai/faux):

1) La tension artérielle diastolique explique 30% de la variance de LVMI
2) Le coefficient de corrélation de Pearson entre LVMI et tension artérielle diastolique vaut 0,30
3) Le coefficient de corrélation de Pearson entre tension artérielle diastolique et LVMI vaut 0,30
4) Le LVMI augmente en moyenne de 0,09 unité lorsque la tension artérielle est incrémentée d'une unité

1. FAUX. Le coefficient r^2 est la proportion de variance de LVMI expliquée par la tension artérielle diastolique. C'est 0,09 (soit 9%).
2. VRAI. Le coefficient r^2 est le carré du coefficient de corrélation de Pearson. La racine carrée de 0,09 est 0,30: le coefficient de corrélation de Pearson entre LVMI et la tension artérielle diastolique vaut donc 0,30.
3. VRAI. Le coefficient de corrélation de Pearson entre LVMI et tension artérielle diastolique est égal au coefficient de corrélation de Pearson entre tension artérielle diastolique et LMVI. Il est donc égal à 0,30
4. FAUX. Le coefficient r^2 n'est pas la pente de la droite de régression et ne doit pas être interprété de cette manière.

Dans l'exemple vu en cours de la modélisation du poids en fonction de la taille (slide 33), l'équation de la droite de régression était:

poids prédit = **-86.3** + **0.88***taille

La taille était exprimée en centimètre et le poids en kg

1) Quelle serait la valeur de la pente si la taille était exprimée en mètre ?
2) Quelle serait la valeur de la constante (intercept) si la taille était exprimée en mètre ?

1. Lorsque la taille est exprimée en centimètre, la pente est 0.88 kg/cm: pour un incrément de taille de 1 cm, le poids augmente en moyenne de 0.88 kg. Donc pour un incrément de 1 mètre (=100 cm) de taille, le poids augmente en moyenne de 0.88*100=88 kg. Si la taille était exprimée en mètre, la pente de la droite de régression serait 88 kg/m.
2. Lorsque la taille est exprimée en centimètre, la constante (intercept du modèle), qui vaut -86.3 kg, est le poids prédit par le modèle pour une personne mesurant 0 cm. Si la taille était exprimée en mètre, la constante du modèle serait le poids prédit pour une personne mesurant 0 m. Ce serait le même poids prédit puisque 0 cm = 0 m: la constante du modèle serait toujours égale à -86.3 kg.

Dans l'exemple vu en cours de la modélisation du poids en fonction de la taille (slide 33), l'équation de la droite de régression était:

poids prédit = **-86.3** + **0.88***taille

La taille était exprimée en centimètre et le poids en kg

1) Quelle serait la valeur de la pente si le poids était exprimé en centigramme (1 cg = 0,01 g = 0,00001 kg) ?
2) Quelle serait la valeur de la constante (intercept) si le poids était exprimé en centigramme?

1. Lorsque le poids est exprimé en kg, la pente est 0.88 kg/cm: pour un incrément de taille de 1 cm, le poids augmente en moyenne de 0.88 kg. Comme 1 kg= 100 000 cg, une augmentation de 0,88 kg est égale à une augmentation de 88 000 centigrammes. Si le poids était exprimé en gramme, la pente serait 88 000 cg/cm.
2. Lorsque le poids est exprimé en kg, la constante (intercept du modèle) est le poids prédit par le modèle pour une personne mesurant 0 cm: -86.3 kg, soit -8 630 000 centigrammes. Si le poids était exprimé en gramme, l'intercept du modèle serait -8 630 000 cg.

Dans l’exemple vu en cours de la modélisation du poids en fonction de la taille et du genre (modèle multiple, slide 56), l’équation de prédiction du poids était:

poids prédit = **-72.9** + **0.8***taille + **3.2*** si Genre = homme

La taille était exprimée en centimètre et le poids en kg

Quelle le poids prédit par ce modèle pour:

1) une femme mesurant 160 cm ?
2) un homme mesurant 160 cm ?
3) un homme mesurant 170 cm ?

1. Dans l’équation, il faut remplacer la taille par la valeur en centimètre (160): -72.9+0.8x160 et ajouter 3.2 au poids prédit si la personne est un homme (et 0 si la personne est une femme). Finalement, le poids prédit est -72.9 + 0.8 x 160 + 0 = 55.1 kg
2. La taille est la même que dans la question 1), mais il faut cette fois rajouter 3.2 au poids prédit car la personne est un homme. Le poids prédit est égal à -72.9 + 0.8 x 160 + 3.2 = 58.3 kg
3. Le poids prédit est égal à -72.9+0.8x170+3.2 = **66.3** kg