Correction de l'Examen de juin 2010

STA109

Exercice 2

1a) Médiane= 34

Quartile 1 = 32

Quartile 2 = 37

1b) Moyenne = 34,18 cm

Estimation non biaisée de la variance = 8,56 cm^2

1c) Intervalle de confiance de la moyenne

la variance n'étant pas connue on utilise la loi de Student à 11-1=10 degrés de liberté.

Fractile de la loi de Student à 10 ddl pour un risque de 5 % : t=2,228

Donc :

$$34{,}18 - 2{,}228*\sqrt{(8{,}56/11)} \leq \mu \leq 34{,}18 + 2{,}228*\sqrt{(8{,}56/11)}$$

$$32{,}21 \leq \mu \leq 36{,}15$$

2.a) H_0 : la moyenne du périmètre crânien de la population des filles dont est issue l'échantillon est égale à celle de la population des garçons dont est issu l'échantillon étudié.

H_1 : les moyennes des périmètres crâniens sont différentes.

2.b) La mesure du périmètre crânien doit suivre une loi normale, hypothèse préalable car l'effectif est inférieur à 30 dans chaque groupe.

2.c) Les variances étant estimées on doit d'abord tester l'égalité des variances à partir de leur estimation. La statistique du test s'écrit :

$F = S^2_g / S^2_f$

S^2_g : estimation de la variance chez les garçons

S^2_f : estimation de la variance chez les filles

La statistique du test F, suit sous H_0 (hypothèse nulle d'égalité des variances) une loi de Fisher à (16-1, 11-1) degrés de liberté (ddl).

Le fractile théorique de la loi de Fisher à (15, 10) ddl pour un risque de 5 % est : f_{th} = 3,52

$f_{obs} = 2,4 / 2,1 = 1,14 < 3,52$

l'hypothèse H_0 ne peut être rejetée, les variances sont supposées égales.

Pour le test de comparaison de moyenne on utilise la loi de Student à (16+11-2) ddl, car on dispose des estimations des variances et les effectifs sont inférieurs à 30 dans chaque groupe. On calcule une variance commune à partir des variances des périmètres crâniens des filles et des garçons :

$S = \sqrt{((10*2,1 + 15*2,4)/(11+16-2))} = \sqrt{2,28} = 1,51$

Statistique du test : $t_{obs} = (35,1 – 34,7) / \sqrt{(2,28*(1/11 + 1/16))} = 0,676$

Le fractile théorique de la loi de Student à 25 ddl pour un risque de 5 % est : $t_{th}(25) = 2,06$

$t_{obs} > t_{th}(25)$

on ne peut rejeter H_0, les moyennes sont statistiquement les mêmes.

Exercice 3

1) Fréquence de l'exposition à l'amiante :

chez les cas : 36 / 46 = 0,7826

chez les témoins : 196 / 402 = 0,4876

2) Si le cancer du poumon est indépendant de l'exposition à l'amiante, la fréquence de l'exposition à l'amiante chez les cas ne doit pas être statistiquement différente de la fréquence à l'exposition chez les témoins.

3) Calcul des effectifs attendus sous H_0 :

		Cas (personne avec un cancer)	Témoin	
Exposition à l'amiante	Oui	23,82	208,18	232
	Non	22,18	193,82	216
	Total	46	402	448

Calcul du Chi2 :

$X= (36-23,82)^2/23,82 + (196-208,18)^2/208,18 + (10-22,18)^2/22,18 + (206 – 193,82)^2/193,82$

$X^2 = 6,2281 + 0,7126 + 6,6886 + 0,7654$

$X = 14,3947$

Chi 2 théorique à 1 ddl, au risque 5 % : 3,84

Le Chi2 observé est supérieur au Chi2 théorique, donc on rejette H_0, le cancer du poumon est lié à l'exposition à l'amiante.

4) Calcul de l'OR :

OR = 36*206 / (196*10) = 3,78

5) L'OR étant de 3,78, l'exposition à l'amiante est liée « positivement » au risque de cancer. Il y a 3,78 fois plus de personnes avec un cancer par rapport à ceux sans cancer chez les exposés à l'amiante que de personnes avec un cancer par rapport à celles sans cancer chez les non-exposés à l'amiante. Pour simplifier on dit que le risque d'avoir un cancer du poumon est environ 3,78 fois plus grand chez les personnes exposées à l'amiante par rapport à celles qui ne sont pas exposées à l'amiante.

Exam CNAM : correction bJ.

M affecte 1% de la pop.

80% possèdent l'allèle A contre 10% de la pop totale.

1) Proba qu'un sujet qui a A développe M ?

$P(M) = 0{,}01$

$P(A|M) = 0{,}80$

$P(A) = 0{,}1$

$$P(M|A) = \frac{P(A|M)\,P(M)}{P(A)}$$

$$= \frac{0{,}8 \times 0{,}01}{0{,}1}$$

$$= 0{,}08$$

2) Proba qu'un sujet qui n'a pas A dév M :

$$P(M|\bar{A}) = \frac{P(\bar{A}|M)\,P(M)}{P(\bar{A})}$$

$$= \frac{(1 - P(A|M))\,P(M)}{(1 - P(A))}$$

$$= \frac{(1 - 0{,}8)\,0{,}01}{(1 - 0{,}1)}$$

$$= 0{,}0022$$

3) Proportion des sujets qui présentent les 2 caractères : $\overline{\Pi}$ et A.

$$\boxed{\begin{aligned} P(A \cap \overline{\Pi}) &= \cancel{P(A|\overline{\Pi})\, P(\overline{\Pi})} \\ &= P(\overline{\Pi}|A)\, P(A). \\ &= (1 - P(\Pi|A))\, P(A) \\ &= (1 - 0{,}08)\ 0{,}1 \\ &= 0{,}092 \end{aligned}}$$

4) (indép)

échantillon : 90 ne présentent ~~pas~~ pas A

10 le présentant.

On sélectionne 3 pers au hasard. Quelle est la proba qu'au moins un sujet présente A.

nb trios possibles : C_{100}^{3}

nb cas défavorables : C_{90}^{3}

$\Rightarrow P(\text{au moins } 1\,A) = 1 - \dfrac{C_{90}^{3}}{C_{100}^{3}}$

$= 0{,}27$