TD statistiques. 15/1/98

Exercice 1

Soit $(X_1, X_2, ..., X_n)$ un échantillon de n mesures.
On peut écrire:

$$X_i = \mu + \epsilon_i,$$

avec ϵ_i suivant une loi normale $\mathcal{N}(0, \sigma^2)$, et μ une constante (la mesure exacte de la longueur). On peut aussi dire que les X_i suivent une loi normale $\mathcal{N}(\mu, \sigma^2)$.

1. Soit $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. (longueur moyenne estimée). On cherche n tel que:

$$P(\left|\overline{X} - \mu\right| \leq 1.4) = .95.$$

Si les X_i suivent la loi normale $\mathcal{N}(\mu, \sigma^2)$, alors:

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}},$$

suit une loi normale centrée réduite. (une partie du th de Fisher...à demontrer à l'aide des fonctions génératrices par exemple).
Connaissant la loi de Z, on va chercher à utiliser cette VA dans l'encadrement.

$$\begin{aligned}
& P(\left|\overline{X} - \mu\right| \leq 1.4) = .95. \\
\Leftrightarrow \quad & P(\frac{-1.4}{\sigma/\sqrt{n}} \leq Z \leq \frac{1.4}{\sigma/\sqrt{n}}) = .95 \\
\Leftrightarrow \quad & \int_{-a}^{a} f(x)dx = .95,
\end{aligned}$$

où $a = \frac{1.4}{\sigma/\sqrt{n}}$, et f est la loi normale centrée réduite.
Comme la loi normale est symétrique, on cherche donc n tq:

$$\int_0^a f(x)dx = .475$$

On trouve $a = 1.96$, c'est-à-dire $n = 18$.
Il faut donc au moins 18 mesures pour que μ soit compris dans l'intervalle $[\overline{X} - 1.4, \overline{X} + 1.4]$ avec un taux de confiance de 95%.

2. On cherche a tq:

$$P(\left|\overline{X} - \mu\right| \leq a) = .95$$

Ici, on ne connait pas σ^2, mais on l'a estimé par $\widehat{\sigma^2} = S^2 = 22.9mm^2$.
On sait que $W = \frac{S^2(n-1)}{\sigma^2}$ suit une loi du χ^2 à n-1 degrés de liberté (re- th

de Fisher !!). On sait aussi que $T = \frac{Z}{\sqrt{\frac{W}{n-1}}}$ suit une loi de Student à n-1 deg de lib.

$$T = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}\sqrt{S^2/\sigma^2}} = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

On va donc se servir de T pour trouver un majorant de $\left|\overline{X} - \mu\right|$.

$$P(\frac{-a}{s/\sqrt{n}} \leq \frac{\overline{X} - \mu}{S/\sqrt{n}} \leq \frac{a}{s/\sqrt{n}}) = .95$$

$$\Leftrightarrow \quad \int_{-b}^{b} f(x)dx = .95,$$

où $b = \frac{a}{S/\sqrt{n}}$, et f est la loi de Student à 9 deg de lib.
Comme la loi de Student est symétrique, on a:

$$\int_0^b f(x)dx = .475$$

Les tables ne nous donnent que les résultats des intégrales du type $\int_{-\infty}^{x}$.
On s'arrange donc pour passer par la borne $-\infty$:

$$\int_0^b f(x)dx = \int_0^{-\infty} + \int_{-\infty}^{b} = .475$$

$$\Leftrightarrow \quad \int_{-\infty}^{b} f(x)dx = .975$$

D'où: $b = 2.262$, *ie* $a = \frac{S}{\sqrt{10}} \times 2.262 = 3.42$ L'intervalle de confiance à 95% est donc beaucoup plus large que dans la question précédente...les résultats sont moins précis.

3. On a vu que $W = \frac{S^2(n-1)}{\sigma^2}$ suit une loi du χ^2 à n-1 degrés de liberté. On cherche a et b tq:

$$P(a \leq \frac{S^2 \times 9}{\sigma^2} \leq b) = A,$$

A valant tour à tour .95 et .99. Il y a une infinité de couples (a, b) vérifiant cette équation. Par convention, on les choisit tq:

$$P(W \leq a) = P(W > b) = \frac{1 - A}{2}$$

On trouve:
Si $A = .95$, $a = 2.7 et b = 19$
Si $A = .99$, $a = 1.73 et b = 23.6$

$$a \leq \frac{S^2 \times 9}{\sigma^2} \leq b \Leftrightarrow \sqrt{\frac{9 \times 22.9}{b}} \leq \sigma \leq \sqrt{\frac{9 \times 22.9}{a}}$$

Si $A = .95$, $\sigma \in [3.29, 8.74]$
Si $A = .99$, $\sigma \in [2.95, 10.91]$
En fait, on sait que σ vaut 3, donc il ne rentre que dans l'intervalle de confiance à 99%...cela veut dire que l'estimation S est relativement mauvaise, et donc que l'échantillon des 10 mesures de l'étudiant est assez peu représentative de sa population parente.

Exercice 2

cf cours...