UFR Sciences et Techniques **Année 2007-2008**

Master 1 de Mathématiques **Statistiques**

Feuille de TD n° 6

Statistiques d'ordre, distributions d'échantillonnage

Exercice 1 (loi de $X_{(i)}$)

Soit $X_{(i)}$ la ie statistique d'ordre d'un n-échantillon $X = (X_1, \ldots, X_n)$ issu d'une loi de fonction de répartition continue F.

1) Pour tout $x \in \mathbb{R}$, on considère la variable aléatoire $N_x = \displaystyle\sum_{k=1}^{n} \mathbf{1}_{[X_k \leq x]}$.

a) Montrer que $[X_{(i)} \leq x] = [N_x \geq i]$

b) En déduire la valeur de $P(X_{(i)} \leq x)$ en fonction de $F(x), n$ et i.

c) Si la loi F a une densité de probabilité $f(x)$, montrer que $X_{(i)}$ a une densité de la forme

$$f_{(i)}(x) = n\ C_{n-1}^{i-1}\ (F(x))^{i-1}(1-F(x))^{n-1}f(x)$$

(*indication* : on pourra utiliser sans la démontrer, la formule : $\displaystyle\sum_{j=i}^{n} C_n^j\ a^j(1-a)^{n-j} = n\ C_{n-1}^{i-1}\int_0^a\ t^{i-1}(1-t)^{n-i}dt \quad (0 \leq a \leq 1)$).

Exercice 2 (intervalle de confiance du p-quantile)

Soit $X = (X_1, \ldots, X_n)$ un n-échantillon issu d'une loi de fonction de répartition continue F. On rappelle que le p-quantile de F est le nombre $q_p = \inf\{x; F(x) \geq p\}$.

1) Montrer que si i et j sont des entiers tels que $1 \leq i\ < j \leq n$, on a :

$$(1) \qquad P(X_{(i)} \leq q_p \leq X_{(j)}) = \sum_{r=i}^{j-1} \frac{n!}{r!(n-r)!}\ p^r(1-p)^{n-r}.$$

(*indication* : on utilisera l'exercice précédent).

2) a) Décrire précisément une méthode qui, à partir d'un n-échantillon X d'une loi F, permette de déterminer un intervalle de confiance q_p au niveau de confiance $1-\alpha$ donné. Y-a-t'il unicité de cet intervalle de confiance ?

b) A partir d'un échantillon de taille $n = 20$, préciser, à l'aide d'une table de la loi binomiale, un intervalle de confiance de la médiane de F au niveau de confiance $1-\alpha = 0,95$.

Exercice 3 (intervalle de confiance approché)

On suppose qu'on dispose d'un n-échantillon d'une loi continue F ayant une densité de probabilité f strictement positive sur $\mathbb{R}$. On suppose n grand. Donner une expression approchée d'un intervalle de confiance pour le quantile q_p d'ordre p (*indication* : on pourra approcher $f(q_p)$ par $f(X_{[np]+1})$).

Exercice 4 (un test de comparaison)

On veut comparer les rendements de blé avec deux types d'engrais : un engrais A traditionnel et un engrais B d'un nouveau type. On a sélectionné 10 terrains qu'on a divisé en 2 parcelles ; dans la première on a utilisé l'engrais A et dans la deuxième l'engrais B. Voici les résultats obtenus (en tonnes récoltées à l'hectare)

terrain	1	2	3	4	5	6	7	8	9	10
engrais A	6,51	6,62	6,41	6,91	6,54	6,71	6,85	6,82	6,64	6,97
engrais B	6,80	6,70	6,72	6,63	6,66	6,60	6,75	6,91	6,75	6,78

On ne fait aucune hypothèse de normalité sur les distributions de probabilité associées aux deux différentes productions. Proposer un modèle permettant de comparer les deux engrais deux à deux pour dire si l'engrais B améliore le rendement.