

Statistique Inférentielle - Examen du 12/01/2015

Durée : 2h - Tous documents autorisés

Consignes : la qualité de la rédaction sera prise en compte dans la note. Tous les intervalles de confiance seront réalisés avec un niveau de confiance de 95% et les tests avec un risque de première espèce de 5%.

Exercice 1

En L3 IDS, il y a 11 femmes sur 33 étudiants. En supposant que cette promotion est représentative de l'ensemble des formations en statistique et informatique de Lyon, pouvez vous affirmer que ce type de formation intéresse autant les hommes que les femmes ?

Exercice 2

Une société de production d'électricité éolienne, cherche à comparer l'efficacité de deux types d'éoliennes : une éolienne à deux pales (E2p) et une éolienne à trois pales (E3p). Pour ce faire, elle a installé sur un même parc éolien une éolienne de chaque type, et a relevé les puissances de chaque éolienne (en kW) toutes les 10 minutes. Afin de comparer les productions des éoliennes, l'ingénieur statisticien a prélevé aléatoirement dans la base de données, et ce de façon indépendante pour chaque éolienne, les 20 productions suivantes :

E2p	5	18	19	11	6	19	20	22	17	12	18	19	4	12	0	2	23	3	8	3
E3p	2	22	28	12	6	18	29	21	24	8	25	22	7	12	0	0	24	7	8	2

Nous supposons que ces échantillons sont distribués suivant une loi normale.

1. Donner une estimation, pour chaque éolienne, de la production moyenne ainsi que de son écart-type.
2. Donner un intervalle de confiance sur la production moyenne de chaque éolienne.
3. A l'aide de ces intervalles de confiances, pouvez-vous préconiser à l'entreprise un des deux types d'éoliennes ?
4. Un parc éolien est jugé intéressant s'il produit en moyenne plus de 16kW. Le parc éolien serait-il rentable avec chacun de ces deux types d'éoliennes ?
5. Et si nous n'avions pas supposé la normalité des échantillons, la réponse à la question précédente aurait-elle été différente ?

Exercice 3

Afin de comparer la rentabilité de quatre parcs éoliens situés dans le Nord, l'Hérault, le Finistère et la Franche-Comté, l'ingénieur statisticien a extrait aléatoirement (et indépendamment) de ses bases de données six productions mensuelles d'électricité (en MW) pour chacun de ces quatre parcs.

Nord	1.50	1.12	1.42	1.43	1.73	1.83
Hérault	1.22	1.01	1.52	1.36	1.24	1.17
Finistère	1.36	1.99	1.38	1.63	1.16	1.25
Franche-Comté	1.16	1.79	1.58	1.36	1.61	1.60

Un parc éolien est jugé rentable s'il produit mensuellement plus de 1.5MW.

1. Chacun des quatre parcs est-il rentable ? Vous utiliserez une méthodologie statistique appropriée pour répondre à la question.
2. Pour chaque parc, quelle est la probabilité de le juger non rentable alors qu'il l'est en réalité.
3. Soit X le nombre de parcs qui, parmi les quatre, sont jugés non rentable à tort. Donner la loi de probabilité de X . Déduisez-en la probabilité que l'on ait jugé non rentable à tort au moins un parc.
4. Quel risque de premier espèce proposeriez-vous pour que cette probabilité ne dépasse pas 5% ?