Université Montpellier
M2 ICAP

Année 2015-2016

Examen : Traitement statistique des données sensorielles

Durée 1h45. Les documents ainsi que les téléphones portables, tablettes, ordinateurs ne sont pas autorisés. La calculatrice est autorisée. La qualité de la rédaction sera prise en compte. Le sujet comporte deux pages.

Sujet I

Exercice 1. Généralités

1. Calculer $\binom{100}{98}$ puis $\mathbb{P}(Bin(100, 0.9) \geq 98)$.
2. Quelle est la loi de la moyenne empirique d'un échantillon de 10 variables aléatoires $\mathcal{N}(1.1; 4)$ indépendantes et identiquement distribuées.
3. Soit $X_1, X_2, ..., X_n$ des variables aléatoires indépendantes gaussiennes centrées et réduites.
   (a) On note $C_n^2 = X_1^2 + X_2^2 + ... + X_n^2$. Déterminer la loi de C_n^2 puis sa moyenne et sa variance. *Indication : on admettra que* $\mathbb{E}(X_1) = 1$ *et* $\mathbb{V}(X_1) = 2$.
   (b) Soit A_1 et A_2 deux variables aléatoires indépendantes suivant respectivement des lois du $\chi^2(K_1)$ et $\chi^2(K_2)$. Determiner (en justifiant) la loi de la somme $A_1 + A_2$.

Exercice 2. Discernabilité

On cherche à tester la discernabilité de deux vins rouges A et B vinifiés avec des raisins récoltés à deux stades de maturité différents. Pour cela, on choisit 20 testeurs dans une population d'experts et on pratique l'épreuve dite “2 parmi 5” : on présente aux testeurs 5 échantillons anonymes issus de A ou de B. Un des deux produits est doublé, l'autre est triplé. On demande aux juges d'identifier les deux échantillons correspondant au produit doublé.

Partie I On considère qu'une réponse au test “2 parmi 5” est bonne si le juge a trouvé les 2 échantillons correspondant au produit doublé. Il y a $N = 6$ juges qui ont fait de bonnes réponses.

1. Justifier l'effectif de testeurs utilisé pour ce test.
2. Donner la relation reliant la proportion théorique π de testeurs parvenant à discerner les deux vins et la proportion théorique p_{br} de bonnes réponses dans la population de testeurs.
3. Calculer un intervalle de confiance (IC) approché de niveau 0.95 pour π. Discuter la qualité de l'IC et/ou la valeur des bornes. Si l'on voulait diviser par 4 la longueur de l'IC, de combien devrait être l'effectif ? *Indication : On pouura se servir du tableau de quantiles suivant : si* $X \sim \mathcal{N}(0; 1)$ *alors*

i	*1.645*	*1.96*	*2.326*	*2.576*
$\mathbb{P}(X \geq i)$	*0.95*	*0.975*	*0.99*	*0.995*

4. Réaliser le test exact en utilisant le nombre N de bonnes réponses afin de tester si les deux vins sont discernables (prendre le niveau 0.05 et indiquer la p-valeur). *Indication : On pouura se servir de l'un de ces deux tableaux de quantiles. Si* $X \sim Bin(24; 2/5)$ *alors*

i	*4*	*5*	*6*	*7*
$\mathbb{P}(X \geq i)$	*0.984*	*0.949*	*0.874*	*0.750*

Si $X \sim Bin(20; 1/10)$ *alors*

i	*4*	*5*	*6*	*7*
$\mathbb{P}(X \geq i)$	*0.133*	*0.043*	*0.011*	*0.002*

5. Calculer la puissance de ce test lorsque $\pi = 0.2$. *Indication : Si* $X \sim Bin(24; 0.28)$ *alors*

i	*4*	*5*	*6*	*7*
$\mathbb{P}(X \geq i)$	*0.853*	*0.698*	*0.505*	*0.317*

Partie II On dit qu'un juge a donné une "bonne réponse partielle" s'il a trouvé au moins un des deux échantillons du produit doublé. On s'intéresse dans cette partie au test basé sur le nombre M de bonnes réponses partielles.

1. Comparer M et N.
2. On cherche la relation reliant la proportion théorique π de testeurs parvenant à discerner les deux vins et la proportion théorique p_{rp} de réponses dans la population de testeurs.
   (a) Calculer la probabilité qu'un juge ne détectant aucune différence entre A et B donne une bonne réponse partielle.
   (b) Donner la relation reliant la proportion théorique π de testeurs parvenant à discerner les deux vins et la proportion théorique p_{rp} de réponses partielles dans la population de testeurs.
3. Durant le test décrit plus haut, on a observé $M = 16$ bonnes réponses partielles. Réaliser un test exact utilisant M afin de tester si les deux vins sont discernables (prendre le niveau 0.05 et indiquer la p-valeur). *Indication : Si* $X \sim Bin(20, .7)$

i	*15*	*16*	*17*	*18*	*19*
$\mathbb{P}(X \geq i)$	*0.416*	*0.238*	*0.107*	*0.035*	*0.00763*

4. On pose $\pi = 0.2$ et on fixe $\alpha = 0.05$:
   (a) Calculer la puissance de ce test pour ces valeurs des paramètres. *Indication : Si* $X \sim Bin(20, .76)$

i	*15*	*16*	*17*	*18*	*19*
$\mathbb{P}(X \geq i)$	*0.657*	*0.456*	*0.257*	*0.109*	*0.030*

Conclusion

1. Comparer les deux procédures de test de la partie I et II. Laquelle des deux utiliseriez vous ? Expliquer pourquoi les conclusions différentes des ces tests ne sont, qu'en apparance, contradictoires.
2. Comparer la procédure de la partie I avec le test triangulaire et quadragulaire. Expliquer lequel vous parait le meilleur ?

Exercice 3. Trois jurys composés d'experts, d'amateurs et de consommateurs goutent un vin blanc du Jura assemblage de chardonnay et de savagnin. On note J la variable représentant le jury. La premiere question du qestionnaire porte sur la présence ou non d'un goût de noix (variable notée GN). Voici les résultats :

GN \ J	Expert	Amateur	Consomateur
oui	34	12	20
non	6	6	20

1. Donner les distributions conditionnelles et la distribution marginale de la variable GN. Les représenter sur un même graphique. Les variables J et GN vous semblent-elles indépendantes ?
2. Calculer la statistique de χ^2.
3. On estime que J et GN sont indépendante (au risque 1% de se tromper) si la statistique du χ^2 est plus petite que $c_{0.99}(2) = 9.210$. D'où vient cette valeur ?