Université Montpellier 2 Année 2013-2014
M2 ICAP

Examen : Traitement statistique des données sensorielles

Durée 1h30. 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.

Exercice 1. Généralités

1. Quelle est la loi de la moyenne empirique d'un échantillon de 10 variables aléatoires $\mathcal{N}(1.1; 4)$ indépendantes et identiquement distribuées.
2. Calculer la loi de $X + Y$ où $X \sim Bin(31; 0.7)$ et $Y \sim Bin(4; 0.7)$ sont indépendantes. *Indication : on pourra exprimer X et Y sous forme de somme d'indicatrices...*

Exercice 2. Classement

Un laboratoire a mis au point 5 arômes artificiels imitant celui de la vanille naturelle. Il cherche à trouver les arômes synthétiques les plus proches du naturel. Pour ce faire, il demande à 20 juges de goûter d'abord la vanille naturelle (témoin), puis les 5 imitations présentées en aveugle et de classer ces dernières par ordre de ressemblance croissante au témoin. On note $R_{k,j}$ le rang du mélange k dans le classement du juge j. On note :

$$S_k = \sum_{j=1}^{20} R_{k,j} \qquad \text{pour tout } k = 1, \cdots, 5$$

On obtient :

$$S_1 = 55 \ ; \ S_2 = ?? \ ; \ S_3 = 65 \ ; \ S_4 = 80 \ ; \ S_5 = 70$$

L'hypothèse nulle des tests est H_0 : "les imitations ont toutes le même degré de ressemblance au témoin". Voici deux statistiques susceptibles d'être utilisées ici :

$$F = \frac{12}{nK(K+1)} \sum_{k=1}^{K} (S_k)^2 - 3n(K+1) \text{ suit approximativement sous } H_0 \text{ une loi du } \chi^2(K-1)$$

$$T = \frac{(12L - 3nK(K+1)^2)^2}{nK^2(K^2-1)(K+1)}, \text{ avec } L = \sum_{k=1}^{K} kS_k \text{ suit approximativement sous } H_0 \text{ une loi du } \chi^2(1)$$

où K est le nombre de rangs dans le classement et n le nombre de classements effectué.

1. Le nouveau stagiaire fraîchement arrivé dans le laboratoire a perdu la donnée S_2. Ce n'est pas grave : démontrer que l'on a nécessairement $S_2 = 30$.
2. Tester H_0 au niveau $\alpha = 0.05$ contre H_1 : "les imitations n'ont pas toutes le même degré de ressemblance au témoin". Peut-on affirmer H_1 au risque 5% de se tromper ?
3. Calculer la plus petite différence significative ($PPDS$). Peut-on dire que l'imitation 2 est significativement plus loin de la vanille naturelle que les autres ? *Indication :* On rappelle que $PPDS_\gamma = u_{1-\frac{\gamma}{2}} \sqrt{\frac{nK(K+1)}{6}}$, où γ est un niveau approprié.
4. Les coûts de fabrication des arômes artificiels ne sont pas identiques. En les classant du moins cher à la plus cher, on obtient :

   $$n°2 < n°1 < n°5 < n°3 < n°4$$

   Peut-on affirmer au risque 5% de se tromper que l'ordre de ressemblance au témoin correspond à celui de la cherté ?

Exercice 3. Discernabilité

On cherche à tester la discernabilité de deux bières stout pression A et B. Pour cela on choisit 24 testeurs dans une population et on pratique 2 types d'épreuves classiques en analyse sensorielle. On vous demande d'analyser les résultats à l'aide de tests statistiques.

Dual-Standard On utilise dans un premier temps l'épreuve dual standard : on présente aux 24 testeurs les témoins A et B, puis les mêmes masqués (X et Y) dans un ordre inconnu des testeurs. On leur demande d'identifier X et Y. On obtient les réponses suivantes :

Echantillons																								
X	A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B	B	B
Y	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	A	A	A	A
Réponse																								
X	A	A	B	A	B	B	A	A	A	B	A	B	B	A	B	B	A	B	B	B	A	B	B	A
Y	B	B	A	B	A	A	B	B	B	A	B	A	A	B	A	A	B	A	A	A	B	A	A	B

1. Justifier l'effectif de testeurs utilisé ainsi que le motif (ordre) des présentations de produits. Quel était l'effectif supérieur immédiatement envisageable ?
2. Donner la relation reliant la proportion théorique π_{ds} de testeurs parvenant à discerner les deux bières et la proportion théorique p_{ds} de bonnes réponses dans la population de testeurs.
3. Calculer un intervalle de confiance (IC) approché de niveau 0.95 pour π_{ds}. Discuter la qualité de l'IC et/ou la valeur des bornes. Comment pourrait on améliorer l'IC ?
4. Réaliser le test exact afin de tester si les deux bières sont discernables (prendre le niveau 0.05 et indiquer la p-valeur). *Indication : Si* $X \sim Bin(24; 0.5)$ *alors*

i	*15*	*16*	*17*	*18*
$\mathbb{P}(X \geq i)$	*0.154*	*0.076*	*0.032*	*0.011*

5. Calculer la puissance de ce test lorsque $\pi = 0.5$. *Indication : Si* $X \sim Bin(24; 0.75)$ *alors*

i	*15*	*16*	*17*	*18*
$\mathbb{P}(X < i)$	*0.055*	*0.121*	*0.234*	*0.393*

Épreuve triangulaire On a ensuite testé la discernabilité de ces deux mêmes bières (auprès des mêmes testeurs), mais cette fois ci à l'aide d'un test triangulaire.

1. Donner la relation reliant la proportion théorique π_{tr} de testeurs parvenant à discerner les deux bières et la proportion théorique p_{tr} de bonnes réponses dans la population de testeurs.
2. On a observé 15 bonnes réponses dans le test triangulaire. Réaliser le test exact afin de tester si les deux bières sont discernables (prendre le niveau 0.05 et indiquer la p-valeur). *Indication : Si* $X \sim Bin(24; 1/3)$ *alors*

i	*12*	*13*	*14*	*15*
$\mathbb{P}(X \geq i)$	*0.068*	*0.028*	*0.01*	*0.003*

3. Lequel de ces 2 tests est-il le plus puissant ? Pourquoi ? (On ne demande pas le calcul, mais justifier votre réponse).

Exercice 4. ACP (Bonus)

On a mesuré 3 variables sensorielles A, B et C pour 6 variétés de tabac différentes. Voici les notes que l'on a relevées :

	A	B	C
tabac 1	2.10	4.20	1.05
tabac 2	2.50	5.00	1.25
tabac 3	2.37	4.74	1.185
tabac 4	2.02	4.04	1.01
tabac 5	2.15	4.30	1.075
tabac 6	2.78	5.56	1.39

Lors d'une ACP normée, justifier quel est le plus petit nombre de facteurs principaux à conserver pour décrire 90% de l'inertie de ce tableau ? 100% de l'inertie ?