L1-S4-SV&SVN 2006/2007

# Examen de Mathématiques (Probabilités et Statistiques) du 15 mai 2007

Durée: **2** heures; les documents ne sont pas autorisés; les calculatrices type collège sont autorisées

**Exercice 1 :** Sur une population de microorganismes, on veut étudier la corrélation entre un certain caractère quantitatif $Z$ et le pH du milieu dans lequel ils réussissent à survivre. A l'issue d'une étude faite sur une population totale de 500 microorganismes, on obtient le tableau du nombre d'individus en fonction du pH ($X$ notera une transformation affine du caractère "pH") et du caractère $Z$ (dont les valeurs ont été groupées par classes). Le tableau est le suivant :

| $X$ \ $Z$ | [15,30[ | [30,35[ | [35,40[ | [40,45[ | [45,50[ | [50, 55] | Sommes |
|---|---|---|---|---|---|---|---|
| $-2$ | 22 | 8 | | | | | 30 |
| $-1$ | 36 | 41 | 20 | 13 | | | 110 |
| 0 | 12 | 36 | 68 | 50 | 22 | 22 | 210 |
| 1 | | | 12 | 26 | 32 | 30 | 100 |
| 2 | | | | 6 | 14 | 20 | 40 |
| 3,25 | | | | | 2 | 8 | 10 |
| Sommes | 70 | 85 | 100 | 95 | 70 | 80 | 500 |

1. Tracer l'histogramme de la distribution de $Z$ quand $X = 0$.
2. Refaire le tableau à deux entrées de sorte que sur la première ligne on ait une variable *discrete* $Y = (Z - 42,5)/5$.
3. Trouver la moyenne, la variance et l'écart-type pour les caractères $X$ et $Y$.
4. Même question pour le caractère $Z$.
5. Représenter graphiquement le nuage statistique correspondant à la paire de caractères $X$ et $Y$.
6. Calculer la covariance $\mathrm{cov}(X, Y)$.
7. En déduire le coefficient de corrélation des séries de caractères $X$ et $Y$.
8. Calculer les coefficients directeurs (les pantes) des droites d'ajustement de $y$ en $x$ et de $x$ en $y$.
9. Trouver les equations des deux droites d'ajustement, et les tracer sur la figure du nuage statistique.

*Indication* : Pour (1) attention à la longueur des intervalles ! Pour (2) on pensera au milieu de l'intervalle. Pour (3), il est souhaitable (mais pas obligatoire) de faire une "tableau étendu" qui comporte, à part la ligne/colonne des sommes partielles $n_{\cdot j}$ (resp. $n_{i\cdot}$) déjà donnée dans l'énoncé, aussi 2 autres lignes/colonnes, l'une pour $n_{\cdot j}y_j$ (resp. $n_{i\cdot}x_i$) l'autre pour $n_{\cdot j}y_j^2$ (resp. $n_{i\cdot}x_i^2$). Rappeler toujours les formules des objets à calculer (moyenne, variance, etc…).

**Exercice 2 :** On prélève des coquillages sur une zone déterminée d'une plage afin de mesurer leur longueur (en mm). Sur un échantillon de 83 coquillages on trouve une moyenne de la longueur de 52,17 mm et un écart-type de 1,8 mm.

1. L'échantillon peut-il être considéré comme étant grand ?
2. Connaît-on l'écart-type "théorique" $\sigma$ de la longueur (c-à-d pour la population totale de coquillages) ?
3. Quelle est la formule qui relie l'écart-type "théorique" $\sigma$ à l'écart-type $\sigma_e$ observé sur l'échantillon ?
4. Peut-on dire que la variable aléatoire $\overline{X}$ (moyenne des longueurs des coquillages) suit une loi normale $\mathcal{N}(m, s)$ ? Préciser (par rapport aux questions précédentes) qui sont $m$ et $s$.
5. En déduire la variable aléatoire qui suit une loi normale centrée réduite $\mathcal{N}(0, 1)$.
6. Donner un encadrement de $\overline{X}$ qui correspond à un *risque* de 5 %.
7. Même question pour un risque de 0,1 %.
8. Quel devrait être l'effectif de l'échantillon à choisir de sorte que la moyenne $\overline{X}$ se situe dans un intervalle de longueur égale à 0,4 mm avec un coefficient de sécurité de 95 % ?

*Indication* : Utiliser la table de la loi normale (voir verso).

**— tourner la page —**