

# UNIVERSITE PARIS 1 - PANTHEON SORBONNE

## UFR 02

### LICENCE D'ECONOMIE

### Statistiques appliquées

#### Fiche de TD N°2

L'objectif de cette fiche de TD est de poursuivre vos révisions des concepts et outils de probabilité étudiés au cours des deux premières années de Licence. Les exercices de cette fiche vont porter essentiellement sur les distributions multivariées et sur les différentes notions de convergence.

**Exercice 1 :** Reprenons les données de l'exemple considéré en cours sur le lien entre salaire et formation :

Salaire	1000 à 1199 euros (1100)	1200 à 1600 euros (1400)	1600 à 2400 euros (2000)	Plus de 2400 euros (3000)
Niveau de formation				
< Bac (10)	0.25	0.05	0.00	0.00
Bac (12)	0.10	0.05	0.03	0.02
<= Bac+3 (15)	0.07	0.10	0.08	0.05
> bac+3 (18)	0.03	0.05	0.04	0.08

NB Les nombres entre parenthèses correspondent, pour les salaires, aux centres de classe et pour les niveaux de formation, au nombre d'années d'études moyen pour atteindre le niveau considéré. Ce sont ces nombres qui doivent être utilisés dans vos calculs.

1) Calculez les probabilités suivantes :

- $\text{prob}(\text{formation} \leq \text{bac} \ \& \ 1200 \text{ euros} \leq \text{salaire} \leq 2400 \text{ euros})$
- $\text{prob}(\text{formation} > \text{bac}+3 \ \& \ \text{salaire} \leq 1600 \text{ euros})$ .

2) Déterminez la distribution marginale des salaires et celle de la formation.

3) déterminez la distribution des salaires conditionnelle à un niveau de formation strictement supérieur au baccalauréat

4) Calculez le salaire moyen global et le salaire moyen pour chaque niveau de formation. Quelle relation existe-t-il entre ces différentes grandeurs ? Comment s'appelle cette relation ?

5) De même, calculez la variance du salaire et les variances du salaire pour chaque niveau de formation. Quelle relation existe-t-il entre la variance globale et les variances et espérances relatives à chaque niveau de formation ?

6) Peut-on considérer que le salaire est indépendant du niveau de formation ?

**Exercice 2 :** On considère les ventes de Clio réalisées au sein de 300 concessions Renault et on veut analyser dans quelle mesure les différences de chiffres de ventes observés dans ces concessions s'expliquent par les remises qui y sont accordées par les vendeurs, mesurées par le pourcentage moyen des remises accordées dans chaque concession.

On suppose que la distribution jointe des ventes  $V$  et des remises  $R$  peut être supposée normale bivariée:  $(v,r) \sim N(m_v, m_r, \sigma_v^2, \sigma_r^2, \rho)$ .

- Quelle est la distribution marginale des ventes ?
- Quelle est la distribution conditionnelle des ventes aux remises accordées ?
- Quelle est l'espérance conditionnelle des ventes aux remises accordées ?

**Exercice 3 :**

Si  $X$  est un vecteur de dimension  $n$  suivant une loi normale multivariée  $N(m, \Sigma)$ , montrez que la variable aléatoire réelle  $(X-m)' \Sigma^{-1} (X-m)$  suit une loi du Khi-deux à  $n$  degrés de liberté.

**Exercice 4 :** On considère les variables aléatoires indépendantes  $X$  et  $Y_n$ .  $X$  suit une loi normale centrée réduite et les variables  $Y_n$  sont telles que  $E(Y_n)=0$  et  $V(Y_n)=1/n$ .

On s'intéresse aux variables aléatoires définies par  $Z_n = X+Y_n$ . En utilisant l'inégalité de Chebychev, montrer que  $Z_n$  converge en probabilité vers  $X$ .

**Exercice 5 :** En utilisant l'inégalité de Chebychev, montrer que si  $X$  est une variable aléatoire réelle d'espérance  $m$  et de variance  $\sigma^2$ , alors  $\text{prob}(|X-m| < \lambda\sigma) \geq 1-1/\lambda^2$  pour tout  $\lambda > 0$ .

**Exercice 6 :** Rappelez les définitions des différentes notions de convergence (presque sûre, en loi, en probabilité, en moyenne quadratique) et mettez en exergue leurs différences et les relations qui peuvent les lier.