Université de Caen	**Examen du 19 décembre 2017. Durée 2h**	**Probabilités et Statistiques**
UFR des Sciences	Par Faïcel Chamroukhi	**L3 Maths-MIASHS**

Consignes :
- Sont interdits : Documents, calculettes, téléphones, écouteurs, ordinateurs, tablettes.
- Il est interdit de composer avec un crayon.
- Votre feuille double d'examen doit porter, à l'emplacement réservé, vos nom, prénom, et signature.
- Cette zone réservée doit être cachée par collage.
- Vos feuilles intercalaires doivent être toutes numérotées.
- Le barème (sur 22) est donné à titre indicatif.

Exercice 1 *(8 pts)* Une urne contient une boule blanche et une boule noire indiscernables au toucher. On y prélève une boule, on note sa couleur, et on la remet dans l'urne en rajoutant α ($\alpha \neq 0$) autres boules de la couleur de la boule tirée (toutes ces boules sont indiscernables au toucher). On répète cette épreuve n fois ($n > 2$). Soit les variables aléatoires $(X_i)_{1 \leqslant i \leqslant n}$ définies par $X_i = 1$ si on obtient une boule blanche au i-ème tirage, et $X_i = 0$ sinon.

1. Donner la loi de X_1 et $\mathbb{E}[X_1]$
2. Déterminer la loi du couple (X_1, X_2)
3. En déduire la loi de X_2 puis $\mathbb{E}[X_2]$. Que peut-on remarquer ?

On définit pour $2 \leqslant m \leqslant n-1$, la variable aléatoire $Y_m = \sum_{i=1}^{m} X_i$.

4. Que représente la variable Y_m ? Donner $Y_m(\Omega)$
5. Déterminer la loi de Y_2
6. Sachant qu'au cours des m premiers tirages on a tiré k boules blanches, quel est le nombre de boules dans l'urne avant le $(m+1)$-ième tirage ?
7. Déterminer pour tout $k \in Y_m(\Omega)$, $\mathbb{P}(X_{m+1} = 1 | Y_m = k)$
8. Montrer que $\mathbb{P}(X_{m+1} = 1) = \frac{1+\alpha\mathbb{E}[Y_m]}{2+\alpha m}$. *Indication : On pourra utiliser la formule des probabilités totales*

Exercice 2 *(8 pts)* On rappelle que si (X, Y) est un couple de v.a de fonction de densité de probabilité jointe définie pour $|\rho| < 1$ (ρ étant le coefficient de corrélation) par

$$f_{(X,Y)}(x,y) = \frac{1}{2\,\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right)\right\} \tag{1}$$

alors X est de loi $\mathcal{N}(\mu_X, \sigma_X^2)$ et Y de loi $\mathcal{N}(\mu_Y, \sigma_Y^2)$ avec $(\mu_X, \mu_Y) \in \mathbb{R}^2$ et $(\sigma_X, \sigma_Y) \in {\mathbb{R}_+^\star}^2$. Soit (X, Y) un couple de v.a. de fonction de densité de probabilité jointe définie pour $|\rho| < 1$ par

$$f_{(X,Y)}(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right). \tag{2}$$

1. Déterminer (aucun calcul n'est nécessaire) la loi de X et la loi de Y
2. Déterminer la matrice de variance-covariance $\mathbf{\Sigma}$ du vecteur $\boldsymbol{V} = (X, Y)^t$
3. Quelle est la condition nécessaire est suffisante pour que le couple (X, Y) soit indépendant ?
4. Trouver la fonction de densité de probabilité conditionnelle de Y sachant X
5. Reconnaître la loi correspondante
6. Trouver la loi du vecteur $\boldsymbol{W} = (X - 2Y + 1, 2X - Y)^t$
7. En déduire la loi de $H = X - 2Y + 1$ et celle de $T = 2X - Y$
8. Quelle est la condition nécessaire est suffisante pour que le couple (H, T) soit indépendant ?

Exercice 3 *(6 pts)* Soit θ le paramètre d'une loi de probabilité et soit $\widehat{\Theta}_n$ un estimateur de ce paramètre que l'on cherche à construire à partir d'un n-échantillon de v.a. i.i.d selon cette loi.

1. On note par $b(\widehat{\Theta}_n, \theta)$ le biais de $\widehat{\Theta}_n$ comme estimateur de θ. Donner l'expression de $b(\widehat{\Theta}_n, \theta)$.
2. On appelle risque quadratique de $\widehat{\Theta}_n$ comme estimateur de θ la quantité définie par : $\ell(\widehat{\Theta}_n, \theta) = \mathbb{E}\big[(\widehat{\Theta}_n - \theta)^2\big]$. Montrer que ce risque peut être décomposé selon la "décomposition biais-variance" suivante :

$$\ell(\widehat{\Theta}_n, \theta) \quad = \quad \mathrm{Var}(\widehat{\Theta}_n) + (b(\widehat{\Theta}_n, \theta))^2.$$

   On cherche donc à trouver l'estimateur qui minimise ce risque.
3. Quel est, parmi tous les estimateurs sans biais de θ, celui que l'on doit choisir ?
4. Qu'appelle-t-on un tel estimateur ?

On dispose d'un n-échantillon $(X_1, \cdots, X_n)$ i.i.d selon la loi de Bernoulli $\mathcal{B}(\theta)$ où $\theta = \mathbb{P}(X_i = 1)$ est le paramètre à estimer. On considère les deux estimateurs suivants pour θ : $\widehat{\Theta}_1 = X_1$ et $\widehat{\Theta}_n = \frac{\sum_{i=1}^n X_i}{n}$.

5. Quel est, au sens du risque quadratique, le meilleur estimateur parmi les deux ?
6. En déduire un estimateur convergent de θ