**Exercice 1.** Soit $X$ une v.a. gaussienne de densité : $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, \forall x \in \mathbb{R}$. On note par $\Phi(x)$ sa fonction de répartition.

1. Calculer $\mathbb{E}(X+1)$ et $\mathbb{V}(-2X)$
2. On pose $Y = e^{|X|}$.
   (a) Déterminer la fonction de répartition de $Y$, notée $F_Y(x)$, en fonction de $\Phi(.)$.
   (b) Déterminer la densité de $Y$, notée $f_Y(x)$.
3. Soit $Z$ la v.a réelle de densité

$$f_Z(x) = \begin{cases} \frac{1}{x}\varphi(\ln(x)) & \text{si } x > 0, \\ 0 & \text{sinon.} \end{cases}$$

   (a) Déterminer la fonction de répartition de $Z$, notée $F_Z(x)$, en fonction de $\Phi(.)$
   (b) Calculer $\mathbb{E}(Z)$

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

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

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

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

1. Déterminer (aucun calcul n'est nécessaire) la loi de $X$ et la loi de $Y$
2. Déterminer la matrice de variance-covariance $\boldsymbol{\Sigma}$ 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.** Soient $X$ et $Y$ deux v.a i.i.d suivant la loi normale $\mathcal{N}(0, 1)$ et soient les v.a $U = X + Y$ et $V = X - Y$.

1. Montrer que $(U, V)$ est un couple Gaussien.
2. Montrer que les v.a $U$ et $V$ sont indépendantes.

**Exercice 4.** Soit $V = (X, Y)^t$ un vecteur gaussien centré tel que $\mathbb{E}\left(X^2\right) = 4$ et $\mathbb{E}\left(Y^2\right) = 1$, et les v.a $2X + Y$ et $X - 3Y$ sont indépendantes.

1. Déterminer la matrice de covariance de $V$.
2. Montrer que le vecteur $W = (X + Y, 2X - Y)^t$ est gaussien.
3. Déterminer sa matrice de covariance.

**Exercice 5.** Soient $X_1$ et $X_2$ deux v.a i.i.d suivant chacune la loi normale $\mathcal{N}(0, 1)$. Soit $\mathbf{Y} = (Y_1, Y_2)^t$ le vecteur aléatoire tel que

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{B},$$

où $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix}$, $\mathbf{X} = (X_1, X_2)^t$ et $\mathbf{B} = (2, 3)^t$.

1. Déterminer la loi de $\mathbf{Y}$.
2. Déterminer la loi de $\mathbf{Z} = (Y_1 + Y_2 + 1, 3Y_1 - Y_2)^t$.
3. Déterminer la loi de $Y_1 + Y_2 + 1$, et la loi de $3Y_1 - Y_2$.

$\star$

**Exercice 6.** Une densité de probabilité (ou loi de probabilité dans le cas discret) d'une v.a réelle $X$, notée $f(x; \theta)$ appartient à la *famille exponentielle* si $f(x; \theta)$ peut s'écrire sous la forme

$$f(x; \theta) = \exp[\eta(\theta)T(x) + a(\theta) + b(x)]\cdot \tag{3}$$

où les fonctions $\eta$, $T$, $a$ et $b$ sont à valeurs dans $\mathbb{R}$. La statistique $T(x)$ est appelée statistique exhaustive naturelle et le paramètre $\eta = \eta(\theta)$ est appelé paramètre naturel. Soit $X$ une variable aléatoire de densité $f(x; \theta)$ appartenant à la famille exponentielle.
Déterminer une statistique exhaustive pour le paramètre $\theta$ pour un $n$-échantillon i.i.d $(X_1, \ldots, X_n)$ de v.a de densité $f(x; \theta)$.

**Exercice 7.** Soit $X$ une v.a suivant une loi Poisson de paramètre $\theta > 0$, i.e, $\forall x \in \mathbb{N}$ :

$$p_X(x; \theta) = \mathbb{P}(X = x) = e^{-\theta}\frac{\theta^x}{x!}\cdot$$

On considère un $n$-échantillon $(X_1, \ldots, X_n)$ généré suivant une loi de Poisson de paramètre $\theta$, $p_X(x; \theta)$

1. Déterminer une statistique exhaustive pour $\theta$.
2. Calculer l'estimateur du maximum de vraisemblance (EMV) de $\theta$.
3. En déduire son expression en fonction de la statistique exhaustive calculée dans la question précédente.

**Exercice 8.** *Risque quadratique et décomposition biais-variance* Soit $\theta$ 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 $\theta$. Donner l'expression de $b(\widehat{\Theta}_n, \theta)$.
2. On appelle risque quadratique de $\widehat{\Theta}_n$ comme estimateur de $\theta$ 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 $\theta$, celui que l'on doit choisir ?
4. Qu'appelle-t-on un tel estimateur ?

On considère 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 $\theta$ : $\widehat{\Theta}_1 = X_1$ et $\widehat{\Theta}_n = \frac{\sum_{i=1}^n X_i}{n}$.

1. Quel est, au sens du risque quadratique, le meilleur estimateur parmi les deux ?
2. En déduire un estimateur convergent de $\theta$

**Exercice 9.** *Estimateurs et Borne Inférieure de Cramér-Rao (CRLB)* Soit $X$ une v.a. gaussienne univariée de densité :

$$f(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\,\pi}}\, \mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}.$$

On considère un $n$-échantillon i.i.d $(X_1, \ldots, X_n)$ de v.a de densité $f(.; \mu, \sigma^2)$

1. Calculer la borne inférieure de Cramér-Rao pour un estimateur sans biais de l'espérance $\mu$ (variance $\sigma^2$ connue)
2. Calculer l'estimateur du maximum de vraisemblance de l'espérance $\mu$. Que peut-on remarquer ?
3. Montrer qu'il est sans biais
4. En déduire qu'il est efficace
5. En déduire qu'il est convergent
6. Calculer la la borne inférieure de Cramér-Rao pour un estimateur sans biais de la variance $\sigma^2$ (espérance $\mu$ connu)
7. Calculer l'estimateur du maximum de vraisemblance de la variance $\sigma^2$ qu'on notera $S^2$.
8. On suppose que $\mu$ est connu. Montrer que que $S^2$ est efficace et consistent pour $\sigma^2$.
9. Maintenant mn suppose que $\mu$ est inconnu. En remplaçant dans l'expression de $S^2$ l'espérance $\mu$ par celle de son estimateur calcuée en .2, que peut-on dire sur le biais et l'efficacité de $S^2$ ?
10. Montrer que la variance de l'estimateur $S^2$ est $\frac{2\sigma^4}{n-1}$.
11. En déduire l'efficacité de cet estimateur.
12. Que se passe-t-il si on considère $\frac{n}{n-1}S^2$ comme estimateur de la variance au lieu de $S^2$ ?

⋆

**Exercice 10.** *Régression et maximum de vraisemblance vs moindres carrés* Considérons le modèle de régression linéaire simple

$$Y_i = \beta_0 + \beta_1 X_i + E_i, \tag{4}$$

les erreurs $E_i$ sont supposés centrées (d'espérance nulle) et de variance $\sigma^2$ connue. On cherche à estimer les cœfficients de régression $(\beta_0, \beta_1)$ à partir d'un échantillon *i.i.d.* $\{(x_1, y_1), \ldots, (x_n, y_n)\}$. L'objectif est de montrer que, dans le cas Gaussien, l'estimateur du maximum de vraisemblance est celui des moindres carrés. Pour ce faire

1. Donner l'expression de la fonction *somme des carrés des résidus (RSS)* minimisée par la méthode des moindres carrés. On notera cette fonction $RSS(\beta_0, \beta_1)$
2. On montre que sous le modèle (4) la densité conditionnelle des observations $y_i$ est Gaussienne

   $$f(y_i|x_i; \beta_0; \beta_1)$$

   d'espérance conditionnelle $\mu = \mathbb{E}[Y_i|X_i]$ et de variance $\sigma^2$.

   Calculer cette espérance à partir de (4)
3. Décrire la fonction de *log-vraisemblance conditionnelle*[1] $L(\beta_0, \beta_1; y_1, \ldots, y_n|x_1, \ldots, x_n)$. On notera cette fonction $\ln L(\beta_0, \beta_1)$, où $L$ représente la vraisemblance conditionnelle
4. En déduire que l'estimateur du maximum de vraisemblance est identique à celui des moindres carrés.

**Exercice 11.** *Régression linéaire et moindres carrés* On rappelle le modèle de régression linéaire simple sous sa forme matricielle :

$$\boldsymbol{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{E} \tag{5}$$

où $\boldsymbol{Y}$ est le vecteur aléatoires des observations, $\mathbf{X}$ la matrice de design (matrice des prédicteurs), $\boldsymbol{\beta}$ le vecteur des cœfficients de régression et $\boldsymbol{E}$ le vecteur aléatoire représentant le bruit supposé centré (d'espérance nulle) et de matrice de covariance $\sigma^2\mathbf{I}$, $\mathbf{I}$ étant la matrice identité.

1. Donner l'expression de l'estimateur des moindres carrés (EMC) $\hat{\boldsymbol{\beta}}$ de $\boldsymbol{\beta}$
2. Montrer qu'il est sans biais
3. Calculer sa matrice de covariance

1. Le terme "conditionnelle" ici vient du fait que, comme on est dans un contexte de régression où la variable expliquée $Y$ qui est conditionné par la variable explicative $X$, on a donc une densité conditionnelle $f(y|x; \beta_0, \beta_1)$ plutôt qu'une densité (non conditionnelle) $f(y; \boldsymbol{\beta}_0, \boldsymbol{\beta}_1)$ comme en estimation de densité classique.