TD7 : Gaussian vectors, conditionning

Exercise 1 — *Gaussian vectors.*
Let X be a random vector in $\mathbb{R}^n$. We say that it is a Gaussian vector if for every $t \in \mathbb{R}^n$, the random variable $\langle t, X\rangle \in \mathbb{R}$ has a Gaussian distribution (with possibly null variance).

(1) Recall the parameters, the characteristic function, and (when it exists) the p.d.f. of a Gaussian distribution on $\mathbb{R}$.
(2) Show that $t \mapsto \mathbb{E}[\langle t, X\rangle]$ is a linear form, and $(s,t) \mapsto \mathrm{Cov}[\langle s, X\rangle, \langle t, X\rangle]$ is a *positive semi-definite* bilinear form. Let them be represented by $\langle \cdot, m\rangle$ and $\langle \cdot, \Sigma \cdot\rangle$. Give an interpretation of m_i and Σ_{ij} for every $i, j \in \{1, \dots, n\}$.
(3) Let X be a Gaussian vector, for every $t \in \mathbb{R}^n$, compute $\mathbb{E}[e^{i\langle t, X\rangle}]$. Briefly explain why the distribution of X is characterized by the parameters m and Σ.
(4) Let X be a Gaussian vector with parameters (m, Σ) and A be a $p \times n$ matrix, show that $AX \in \mathbb{R}^p$ is a Gaussian vector, and compute its parameters.
(5) We say that two processes A and B are uncorrelated when for every index t, s, $Cov(A_t, B_s) = 0$. Let V_1 and V_2 be two subspaces of $\mathbb{R}^n$ and X a Gaussian vector. Show that the σ-algebras $\sigma(\langle t, X\rangle, t \in V_1)$ and $\sigma(\langle t, X\rangle, t \in V_2)$ are independent if and only if $(\langle t, X\rangle)_{t \in V_1}$ and $(\langle s, X\rangle)_{s \in V_2}$ are uncorrelated.
(6) Build two standard Gaussian variables X and Y that are uncorrelated yet not independent (they obviously do not form a Gaussian vector !)
(7) Show that the vector $(X_1, \dots X_n)$ with $X_1, \dots, X_n$ independent standard Gaussian variables, is Gaussian. Use it to build a Gaussian vector with arbitrary parameters. Deduce its p.d.f. when it has one.

Exercise 2 — *Conditioning and independence.*
Let $\mathcal{G}$ be a σ-algebra, $X \in \mathcal{G}$ and $Y \perp\!\!\!\perp \mathcal{G}$ be two random variables, and $f : \mathbb{R}^2 \to \mathbb{R}$ such that $f(X, Y) \in L^1$. Compute $\mathbb{E}[f(X, Y) \mid \mathcal{G}]$. Deduce the conditional distribution of $f(X, Y)$ given $\mathcal{G}$.

Exercise 3 — *Gaussian conditional distribution and Bayesian statistics 101.*
Let (X, Y) be a non-degenerate centered Gaussian vector in $\mathbb{R}^2$ with covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_x^2 & \rho \\ \rho & \sigma_y^2 \end{pmatrix}.$$

(1) For every $y \in \mathbb{R}$, compute the conditional distribution of X given $Y = y$.
(2) Let $\theta \sim \mathcal{N}(0, \tau^2)$ and $Y_1, \dots, Y_n$ i.i.d. $\sim \mathcal{N}(0, \sigma^2)$ random variables, define $X_i = \theta + Y_i$. What is the conditional distribution of θ given $\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i = \overline{x}$?
(3) Give an interpretation of the situation discribed in the previous question.

(4) Compute the limit of the distribution of θ given $\overline{X} = \overline{x}$ and give an interpretation in each of the following cases.
  (a) $\sigma \to +\infty$
  (b) $\sigma \to 0$
  (c) $\tau \to +\infty$
  (d) $\tau \to 0$
(5) ($\star$) What about the conditional distribution of θ given $(X_1, \dots X_n)$?

Exercise 4 — *Limit in distribution of Gaussian vectors.*
Let $(X_n)_{n\geq 0}$ be a sequence of Gaussian variables $(X_n)_{n\geq 0}$. Give a necessary and sufficient condition for convergence in distribution, show that the limit is always Gaussian, and determine its parameters.
Hint: You can use tightness to show that when $(X_n)_{n\geq 0}$ converges in distribution, the sequence $(\mathbb{E}\, X_n)_{n\geq 0}$ is bounded.

Exercise 5 — *Borel-Kolmogorov paradox.*
Let P denote a uniform point in the sphere $\mathbb{S}^2$, i.e. for every bounded measurable f,

$$\int f(p)\, \mathbb{P}_P(dp) = \frac{1}{\mathrm{Leb}_3(B_{\mathbb{R}^3}(0,1))} \int_{B_{\mathbb{R}^3}(0,1)} f\left(\frac{p}{|p|}\right) \mathrm{Leb}_3(dp).$$

Denote $\phi_P \in (-\pi/2, \pi/2]$ its latitude and $\theta_P \in (-\pi, \pi]$ its (almost surely defined) longitude.
(1) Compute the joint distribution of (θ_P, ϕ_P).
(2) Let $\overline{\theta}_P \in [0, \pi)$ denote a representant of θ_P modulo π. Compute the conditional distribution of P given $\overline{\theta}_P$.
(3) Compute the conditional distribution of P given ϕ_P.
(4) Justify that there is only one "right way" of specializing those answers when computing the conditional distribution of P given $\overline{\theta}_P = 0$ and the conditional distribution of P given $\phi_P = 0$).
(5) What is the paradox ?