Solutions for $\mathbb{E}$xercise sheet 1 : Review of Gaussian vectors and conditional expectation, and a first approach of Brownian Motion.

Solution 1 — *Gaussian vectors.* (1) The parameters are the mean $\mu \in \mathbb{R}$ and the variance $\sigma^2 \geq 0$. When $\sigma^2 = 0$, the distribution is just the Dirac in μ, and when $\sigma^2 > 0$, it has pdf $f(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-t^2/(2\sigma^2)}$. In both cases the characteristic function is $\phi(t) = e^{i\mu t - \sigma^2/2t^2}$.

(2) This is immediate to check. By decomposing on the standard Euclidean basis it turns out that $m_i = \mathbb{E}[X_i]$ and $\Sigma_{i,j} = \mathrm{Cov}(X_i, X_j)$. We call those the mean vector and the covariance matrix of X.

(3) We have that $\langle t, X\rangle$ is a Gaussian of mean $\langle t, m\rangle$ and variance $\langle t, \Sigma t\rangle$. So by taking the characteristic function of $\langle t, X\rangle$ at point 1 we get $\mathbb{E}[e^{i\langle t,X\rangle}] = \exp(i\langle t, m\rangle - \frac{1}{2}\langle t, \Sigma t\rangle)$. So the distribution of X is completely characterized by the parameters m and Σ.

(4) Compute $\mathbb{E}[e^{i\langle t,Ax\rangle}] = \mathbb{E}[e^{i\langle {}^\intercal At,x\rangle}] = \exp(i\langle {}^\intercal At, m\rangle - \frac{1}{2}\langle {}^\intercal At, \Sigma {}^\intercal At\rangle) = \exp(i\langle t, Am\rangle - \frac{1}{2}\langle t, A\Sigma {}^\intercal At\rangle)$. Gaussianity and identification of the parameters follows.

(5) If we have the independence condition, then for $t \in V_1$ and $s \in V_2$, we have $\mathrm{Cov}[\langle t, X\rangle, \langle s, X\rangle] = 0$ by Fubini's theorem (justified since everybody is in L^2). But the converse is also true: Suppose that for every $t \in V_1$ and $s \in V_2$, we have $\mathrm{Cov}[\langle t, X\rangle, \langle s, X\rangle] = 0$. Let $f_1, \dots f_m$ be a finite family in V_1 followed by a finite family in V_2. Set $Y = (\langle f_1, X\rangle, \dots, \langle f_m, X\rangle) = (Y_1, Y_2)$. Then, by computing covariances, we see that the covariance matrix of Y is block-diagonal. This means that we have a product decomposition $\mathbb{E}[e^{i(\langle t_1, Y1\rangle + \langle t_2, Y_2\rangle}] = \mathbb{E}[e^{i\langle t_1,Y_1\rangle}]\,\mathbb{E}[e^{i\langle t_2,Y_2\rangle}]$. By injectivity of the characteristic distribution, we have identified the distribution of (Y_1, Y_2) as one of an independent couple of two Gaussian vectors. Now because by definition the σ-algebra spanned by a family of variables is generated by the finite subfamilies, we get the independence of the two σ-algebras.

(6) The classic example : set (X, A) to be an independent couple of a standard Gaussian and a Rademacher variable (uniform on $\{\pm 1\}$). Set $Y = AX$. Then Y is not independent of X ($\mathbb{P}(X > 0, Y > 0) = 0 \neq 1/4$). Yet $\mathrm{Cov}(X, Y) = \mathbb{E}[AX^2] = \mathbb{E}[A]\,\mathbb{E}[X^2] = 0 \times 1 = 0$.

(7) If $X = (X_1, \dots X_n)$ then we compute $\mathbb{E}[e^{i\langle t,X\rangle}] = e^{-\frac{1}{2}\langle t,t\rangle}$. So it's Gaussian. For m a vector and Σ a semi-definite positive matrix, use the spectral theorem to write $\Sigma = {}^\intercal ODO$, and consider $Y = m + {}^\intercal O\sqrt{D}X$. It should have the prescribed parameters.

Solution 2 — *Central Limit Theorem and random walks.* (1) We have

$$\widetilde{S}_n(t_i) - \widetilde{S}_n(t_{i-1}) = \frac{1}{\sigma\sqrt{n}} \sum_{k=\lfloor nt_{i-1}\rfloor+1}^{\lfloor nt_i\rfloor} X_k.$$

These form independent random variables thanks to independence of the $(X_k)_k$ and the grouping lemma.

(2) The increment $\widetilde{S}_n(t_i) - \widetilde{S}_n(t_{i-1})$ is distributed like $\frac{1}{\sigma\sqrt{n}} \sum_{p=1}^{\lfloor nt_i\rfloor - \lfloor nt_{i-1}\rfloor} X_p$, which we rewrite as $\frac{\lfloor nt_i\rfloor - \lfloor nt_{i-1}\rfloor}{\sqrt{n}} \times \frac{1}{\lfloor nt_i\rfloor - \lfloor nt_{i-1}\rfloor} \sum_{p=1}^{\lfloor nt_i\rfloor - \lfloor nt_{i-1}\rfloor} X_p$. Now the first factor is a deterministic sequence of numbers that converges to $\sqrt{t_i - t_{i-1}}$, and the second one is a sequence of random variables that converge in distribution to a standard Gaussian Z thanks to the CLT. The convergence in distribution of the increment to $\sqrt{t_i - t_{i-1}}Z \sim \mathcal{N}(0, t_i - t_{i-1})$ follows from the following lemma.

Lemma. If c_n is deterministic, X_n is random, $c_n \to c$ and $X_n \xrightarrow{d} X$, then $c_n X_n \xrightarrow{d} cX$.

Proof. The joint convergence in distribution $(c_n, X_n) \xrightarrow{d} (c, X)$ follows from either Slutsky's lemma (look it up!) or the "basic fact" about convergence in distribution of independent variables stated below. From that we get the convergence $c_n X_n \xrightarrow{d} cX$ by the continuous mapping property of the convergence in distribution (indeed multiplication is continuous). □

To prove convergence in distribution of $(X_1, \dots X_n)$, we use:

Basic fact. If for every n, $(X_n^1, \dots, X_n^k)$ form an independent vector of random variables, and for every $1 \le i \le k$ we have $X_n^i \xrightarrow{d} X^i$, then $(X_n^1, \dots, X_n^k) \xrightarrow{d} (Y^1, \dots, X^k)$, where $Y^1 \stackrel{d}{=} X^1, \dots, Y^k \stackrel{d}{=} X^k$, and the $(Y^i)_i$ are independent.

Proof. Characteristic functions □

From there we get that the vector of increments converges to a vector $(Y_1 \dots Y_k)$ of k independent Gaussians, of respective variances $t_1 - t_0, \dots, t_k - t_{k-1}$.

(3) $(\widetilde{S}_n(t_0), \dots, \widetilde{S}_n(t_k))$ is the linear transform by, say, A (the lower triangular matrix of ones) of the vector of increments. So by the continuous mapping theorem it converges to AY (which we call$(B_{t_0}, \dots, B_{t_k})$). It is a centered Gaussian vector as the linear transform of the centered Gaussian vector Y. Now we compute covariances: $\mathrm{Cov}(B_{t_i}, B_{t_j}) = \mathrm{Cov}(\sum_{p=1}^{i} Y_p, \sum_{p=1}^{j} Y_p) = \sum_{p=1}^{i\wedge j} \mathrm{Var}(Y_p) = t_{i\wedge j} = t_i \wedge t_j$.

(4) Indeed$(B_{1/2}, B_1)$ is distributed like $(U, V) = (\frac{X}{\sqrt{2}}, \frac{X}{\sqrt{2}} + \frac{Y}{\sqrt{2}})$, because this is a centered Gaussian vector with the desired covariances. To rewrite this distribution as the distribution of (something, X), we project the vector $(\frac{1}{\sqrt{2}}, 0)$ onto $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$. This yields

$$(\frac{1}{\sqrt{2}}, 0) = \frac{1}{2}(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) + \frac{1}{2}(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}). \tag{†}$$

We set $W = (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$. Equation (†) translates immediately into $X = \frac{1}{2}V + \frac{1}{2}W$. But we can show that (V, W) is also a standard Gaussian and $(X, Y) = (\frac{1}{2}V + \frac{1}{2}W, V) \stackrel{d}{=} (\frac{1}{2}X + \frac{1}{2}Y, X)$.

Solution 3 — *Limit in distribution of Gaussian vectors.*
We restrict ourselves to gaussian **variables**. It is rather easy to lift this up to vectors afterwards. Let μ_n and σ_n be the parameters of X_n If we have convergence in distribution, then we have convergence of the characteristic functions to the one of the limit. So there exists a characteristic function $f : \mathbb{R} \to \mathbb{R}$ such that for all $t \in \mathbb{R}$, $f_n(t) = e^{i\mu_n t - \frac{\sigma_n^2}{2}t^2} \to f(t)$. Now taking the modulus then the log yields $\sigma_n^2 \to -\frac{2}{t^2}\log(|f(t)|) = \sigma^2 \geq 0$. We deduce that $|f(t)| = e^{-\frac{\sigma^2}{2}t^2}$. Now $e^{i\mu_n t} = e^{\frac{\sigma_n^2}{2}t^2} f_n(t) \to e^{\frac{\sigma^2}{2}t^2} f(t) = u(t)$, which is a continuous function in $\mathbb{C}$ of modulus 1 (with $u(0) =: 1$). So it can be lifted up to a continuous real function, *i.e.* there exists h continuous with $h(0) = 0$ such that $u(t) = e^{ih(t)}$ for all t. We have

$$e^{i(\mu_n t - h(t))} \to 0.$$

We shall now show that $(\mu_n)_n$ is bounded. This important step is treated with a probabilistic proof: we use the fact that the distribution of X_n is symmetric about its mean[1]. Suppose there is an increasing subsequence $m_{k_n} \to \infty$. Then $\mathbb{P}(X_{k_n} \geq m_{k_n}) = 1/2$ for all n, and $\mathbb{P}(X_{k_n} \geq m_{k_p}) \geq 1/2$ for all $n \geq p$. So by taking $n \to \infty$ with fixed p we get $\mathbb{P}(X \geq m_{k_p}) \geq 1/2$ for all p, which is absurd as $m_{k_p} \to \infty$.
So $(m_n)_n$ is bounded above and the symmetric argument allows to show that it is bounded below.
Back to our problem, we shall now show that $A = \{t \in \mathbb{R} : \mu_n t \to h(t)\}$ is the whole of $\mathbb{R}$.

- It is nonempty as it contains 0.
- It is closed because of the uniform control of μ_n in n.
- It is open: let $t \in A$. For $s \in \mathbb{R}$ we have $e^{i(\mu_n t - h(t) - \mu_n s + h(s))}) \to 0$. By the bound on μ_n and continuity of h we can find $\epsilon > 0$ such that for all $s \in (t - \epsilon, t + \epsilon)$ and all n, $|\mu_n t - h(t) - \mu_n s + h(s)| < \pi/2$. But for $|\theta| < \pi/2$, $\theta \mapsto e^{i\theta}$ is an homeomorphism. We deduce $\mu_n t - h(t) - \mu_n s + h(s) \to 0$ and hence $s \in A$.

We conclude by connectedness of $\mathbb{R}$. We get that for every $t \neq 0$, $\mu_n \to h(t)/t$, so μ_n converges to some μ and $h(t) = \mu t$. This proves that $f(t) = e^{i\mu t - \frac{\sigma^2}{2}t^2}$, so X is a Gaussian with parameters $\mu = \lim \mu_n$ and $\sigma^2 = \lim \sigma_n^2$. Conversely these convergences directly imply convergence in distribution.

Solution 4 — *Conditional Fubini's theorem.*
Set $u(x) = \mathbb{E}[f(x, Y)] = \int f(x, y) d\,\mathbb{P}_Y(y)$. According to Fubini's theorem, $u(x)$ is defined $\mathbb{P}_X$-a.e. Let us check that the almost-surely defined random variable $u(X)$ satisfies the universal property required from the conditional expectation $\mathbb{E}[f(X, Y) \mid \mathcal{G}]$.

[1]Since we know that σ_n is bounded, we could as well use the fact that X_n concentrates around its mean

Let Z be a $\mathcal{G}$-measurable bounded random variable. Then $Zf(X,Y) \in L^1$, and since Y is independent of (X,Z), which means $\mathbb{P}_{(X,Z,Y)} = \mathbb{P}_{(X,Z)} \otimes \mathbb{P}_Y$. We deduce

$$\begin{aligned}\mathbb{E}[Zf(X,Y)] &= \int zf(x,y)d\,\mathbb{P}_{(X,Z,Y)}(x,z,y) = \int zf(x,y)d(\mathbb{P}_{(X,Z)} \otimes \mathbb{P}_Y)(x,z,y) \\ &= \int z\left(\int f(x,y)d\,\mathbb{P}_Y(y)\right)d\,\mathbb{P}_{(X,Z)}(x,z) \text{ (Fubini)} \\ &= \mathbb{E}[Zu(X)].\end{aligned}$$

This proves the claim. I often write this very basic claim about conditional expectations as follows :

$$\mathbb{E}[f(X,Y) \mid \mathcal{G}] = \mathbb{E}[f(x,Y)]_{x=X}.$$

Solution 5 — *"Conditional probability".* (1) We know from the first exercise that $(B_{1/2}, B_1)$ is distributed as $(\frac{B_1}{2} + \frac{Y}{2})$, where Y is a standard Gaussian independent of B_1. Then $\mathbb{E}[f(B_{1/2}, B_1) \mid B_1] = \mathbb{E}[f(\frac{B_1}{2} + \frac{Y}{2}, B_1) \mid B_1] = \mathbb{E}[f(\frac{u}{2} + \frac{Y}{2}, u)]_{u=B_1}$ by the previous exercise. We sum this up by saying that the conditional distribution of $(B_{1/2}, B_1)$ given $B_1 = u$ is that of a $(u/2 + \mathcal{N}(0, 1/4), u)$.