Probability, homework 8, solutions.

Exercise 1. Let (X_1, X_2) be a Gaussian vector with mean (m_1, m_2) and non-degenerate covariance matrix $(C_{ij})_{1\leq i,j\leq 2}$. Prove that

$$\mathbb{E}[X_1 \mid X_2] = m_1 + \frac{C_{12}}{C_{22}}(X_2 - m_2).$$

Solution. The Gaussian random variables X_2 and $X_1 - \frac{C_{12}}{C_{22}}X_2$ are independent because they are uncorrelated:

$$\mathbb{E}\left[X_2 \cdot (X_1 - \frac{C_{12}}{C_{22}}X_2)\right] = \mathbb{E}[X_2X_1] - \frac{C_{12}}{C_{22}}\,\mathbb{E}[X_2^2] = C_{12} - \frac{C_{12}}{C_{22}}C_{22} = 0.$$

This implies

$$\mathbb{E}[X_1 \mid X_2] = \mathbb{E}\left[X_1 - \frac{C_{12}}{C_{22}}X_2 + \frac{C_{12}}{C_{22}}X_2 \mid X_2\right] = \mathbb{E}\left[X_1 - \frac{C_{12}}{C_{22}}X_2\right] + \frac{C_{12}}{C_{22}}X_2,$$

which gives the result.

Exercise 2. Let X be a random variable such that $\mathbb{P}(X > t) = \exp(-t)$ for any $t \geq 0$. Let $Y = \min(X, s)$, where $s > 0$ is fixed. Prove that, almost surely,

$$\mathbb{E}[X \mid Y] = Y\mathbb{1}_{Y<s} + (1+s)\mathbb{1}_{Y=s}.$$

Solution. We need to check that for any measurable bounded function f one has

$$\mathbb{E}[X \cdot f(Y)] = \mathbb{E}[(Y\mathbb{1}_{Y<s} + (1+s)\mathbb{1}_{Y=s}) \cdot f(Y)].$$

Without loss of generality we can assum that $f(u) = \mathbb{1}_{u<t}$ for some fixed $t \in \mathbb{R}$. If $s < t$, the previous equation becomes

$$\mathbb{E}[X] = \mathbb{E}[X\mathbb{1}_{X<s} + (1+s)\mathbb{1}_{s\leq X}],$$

which is true by a direct calculation: $\int_0^\infty ue^{-u}\mathrm{d}u = \int_0^s ue^{-u}\mathrm{d}u + (1+s)\int_s^\infty e^{-u}\mathrm{d}u$.

If $s \geq t$, the original equation becomes

$$\mathbb{E}[X\mathbb{1}_{X\leq t}] = \mathbb{E}[X\mathbb{1}_{X\leq t} + 0],$$

which is obvious.

Exercise 3. Let μ and ν be two probability measures such that $\mu \ll \nu$ and $\nu \ll \mu$ (usually abbreviated $\mu \sim \nu$). Let $X = \frac{\mathrm{d}\mu}{\mathrm{d}\nu}$.

(i) Prove that $\nu(X = 0) = 0$.

(ii) Prove that $\frac{1}{X} = \frac{\mathrm{d}\nu}{\mathrm{d}\mu}$ almost surely (for μ or ν).

Solution. (i) By Radon-Nikodym, $\int_{X=0}\mathrm{d}\mu = \int_{X=0}\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\mathrm{d}\nu = \int_{X=0}X\mathrm{d}\nu = 0$. By absolute continuity we then also have $\nu(X = 0) = 0$.

(ii) We have $\mu(A) = \int X\mathrm{d}\nu$ for all measurable A, which implies that

$$\int g\mathrm{d}\mu = \int gX\mathrm{d}\nu \tag{0.1}$$

for all $g \in \mathscr{L}^1(\mu)$, by considering monotone limits of simple functions.

Let $g = 1/X$ if $X \neq 0$, 0 otherwise. Then $g \in \mathscr{L}^1(\mu)$ because for any $g \geq 0$, $\int g\mathrm{d}\mu \leq \int gX\mathrm{d}\nu$ (again by first picking simple g and then monotony). We can therefore apply (0.1) and obtain that for any measurable A, we have

$$\nu(A) = \int_A g\mathrm{d}\mu$$

for all $A \in \mathscr{F}$, where we have used that $gX = 1$ a.s. from question (i). The previous displayed equation characterizes $\mathrm{d}\nu/\mathrm{d}\mu$ uniquely up to zero measure sets (for μ), so that $g = \mathrm{d}\nu/\mathrm{d}\mu$ a.s.

Exercise 4. Let $(X_n)_{n\geq 0}$ be defined on $(\Omega, \mathcal{A}, \mathbb{P})$. Assume this sequence converges in probability (under $\mathbb{P}$) to X. Let $\mathbb{Q}$ be another probability measure on $(\Omega, \mathcal{A})$ assumed to be absolutely continuous w.r.t. $\mathbb{P}$. Prove that $X_n \to X$ in probability under $\mathbb{Q}$.

Solution. By the Radon-Nikodym theorem denoting $f = \mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P}$, for any $\varepsilon > 0$ we have

$$\mathbb{Q}(|X_n - X| > \varepsilon) = \int f \mathbb{1}_{|X_n - X| > \varepsilon} \mathrm{d}\mathbb{P}.$$

Moreover, $f \in \mathscr{L}^1(\mathbb{P})$, so that we can apply dominated convergence to conclude that $\mathbb{Q}(|X_n - X| > \varepsilon) \to 0$.

Exercise 5. Let Y be an integrable random variable on $(\Omega, \mathcal{A}, \mathbb{P})$ and $\mathcal{G}$ a sub σ-field of $\mathcal{A}$. Suppose that $\mathcal{H} \subset \mathcal{G}$ is a sub σ-field of $\mathcal{G}$. Show that $\mathbb{E}\left(\mathbb{E}(Y \mid \mathcal{G}) \mid \mathcal{H}\right) = \mathbb{E}\left(Y \mid \mathcal{H}\right)$ (almost surely).

Solution. By almost sure characterization of the conditional expectation, we only need to prove that for any $H \in \mathcal{H}$ we have

$$\mathbb{E}[\mathbb{E}[Y \mid \mathcal{H}]\mathbb{1}_H] = \mathbb{E}[\mathbb{E}[Y \mid \mathcal{G}]\mathbb{1}_H]$$

The L.H.S. is also $\mathbb{E}[\mathbb{E}[Y\mathbb{1}_H \mid \mathcal{H}]] = \mathbb{E}[Y\mathbb{1}_H]$, and the R.H.S. is $\mathbb{E}[\mathbb{E}[Y\mathbb{1}_H \mid \mathcal{G}]] = \mathbb{E}[Y\mathbb{1}_H]$ because $\mathbb{1}_H$ is both $\mathcal{H}$ and $\mathcal{G}$-measurable.

Exercise 6. Let $X_1, \ldots, X_n$ be i.i.d. integrable random variables, and $S = \sum_{i=1}^n X_i$. Calculate $\mathbb{E}[S \mid X_1]$ and $\mathbb{E}[X_1 \mid S]$.

Solution. We have

$$\mathbb{E}[S \mid X_1] = \mathbb{E}[X_1 + \cdots + X_n \mid X_1] = X_1 + \mathbb{E}[X_2] + \cdots + \mathbb{E}[X_n] = X_1 + (n-1)\,\mathbb{E}[X_1]\, a.s.,$$

where we have used linearity of the conditional expectation, the fact that $\mathbb{E}[X \mid \mathcal{H}] = X$ a.s. when X is $\mathcal{H}$-measurable, and $\mathbb{E}[X \mid \mathcal{H}] = \mathbb{E}[X]$ a.s. when X is independent of $\mathcal{H}$.

Moreover, one can write by exchangeability that

$$\mathbb{E}[X_1 \mid S] = \frac{1}{n}\,\mathbb{E}[X_1 + \cdots + X_n \mid S] = \frac{S}{n}.$$

Exercise 7. For fixed $a, b > 0$, let (X, Y) be a $\mathbb{N} \times \mathbb{R}_+$-valued random variable such that

$$\mathbb{P}(X = n, Y \leq t) = b \int_0^t \frac{(ay)^n}{n!} e^{-(a+b)y} \mathrm{d}y.$$

For $h : \mathbb{R}_+ \to \mathbb{R}$ continuous and bounded, calculate $\mathbb{E}[h(Y) \mid X]$. Calculate $\mathbb{E}[\frac{Y}{X+1}]$. Calculate $\mathbb{P}(X = n \mid Y)$. Calculate $\mathbb{E}[X \mid Y]$.

Solution. Note that $\mathbb{P}[X = n] = \mathbb{P}[X = n, Y \leq \infty] = \int_{\mathbb{R}_+} b\frac{(at)^n}{n!}e^{-(a+b)t}\mathrm{d}t$. We therefore have, for any u continuous bounded,

$$\mathbb{E}[h(Y)u(X)] = \int_{\mathbb{N}\times\mathbb{R}_+} b\frac{(at)^n}{n!}e^{-(a+b)t}h(t)u(n)\mathrm{d}n\mathrm{d}t = \int_{\mathbb{N}} g(n)u(n)\mathbb{P}[X = n]\mathrm{d}n = \mathbb{E}[g(X)u(X)]$$

where $g(n) = \int_{\mathbb{R}_+} b\frac{(at)^n}{n!}e^{-(a+b)t}h(t)\mathrm{d}t / \int_{\mathbb{R}_+} b\frac{(at)^n}{n!}e^{-(a+b)t}\mathrm{d}t$ (the denominator can be simplified). We can conclude that $\mathbb{E}[h(Y) \mid X] = g(X)$.

We have $\mathbb{E}[\frac{Y}{X+1}] = \mathbb{E}[\mathbb{E}[Y \mid X]\frac{1}{X+1}]$. From the previous question, we have

$$\mathbb{E}[Y \mid X] = \int_{\mathbb{R}_+} b\frac{(at)^X}{X!}e^{-(a+b)t}t\mathrm{d}t / \int_{\mathbb{R}_+} b\frac{(at)^X}{X!}e^{-(a+b)t}\mathrm{d}t = \frac{X+1}{a+b}.$$

This gives $\mathbb{E}[\frac{Y}{X+1}] = \frac{1}{a+b}$.

Note that $\mathbb{P}[Y \leq t] = \sum_{n\geq 0} b\int_0^t \frac{(ay)^n}{n!}e^{-(a+b)y}\mathrm{d}y = b\int_0^t e^{-by}\mathrm{d}y$, so that Y has density be^{-by}. We therefore have, for any u continuous bounded,

$$\mathbb{E}[\mathbb{1}_{X=n}u(Y)] = \int_{\mathbb{N}\times\mathbb{R}_+} b\frac{(at)^m}{m!}e^{-(a+b)t}u(t)\mathbb{1}_{m=n}\mathrm{d}m\mathrm{d}t = \int_{\mathbb{R}_+} g(t)u(t)be^{-bt}\mathrm{d}t = \mathbb{E}[g(Y)u(Y)]$$

where $g(t) = e^{-at}(at)^n/n!$ (the denominator can be simplified). We can conclude that $\mathbb{P}(X = n \mid Y) = e^{-aY}(aY)^n/n!$ Finally, we have

$$\mathbb{E}[X \mid Y] = \sum_{n\geq 0} n\mathbb{P}(X = n \mid Y) = \sum_{n\geq 0} ne^{-aY}(aY)^n/n! = aY.$$

Exercise 8. Let X and Y be random variables on $(\Omega, \mathcal{F}, \mathbb{P})$, and $\mathcal{G}, \mathcal{H}$ sub σ-fields of $\mathcal{F}$ such that $\sigma(\mathcal{G}, \mathcal{H}) = \mathcal{F}$. Find counterexamples to the following assertions:

(i) If $\mathbb{E}[X \mid Y] = \mathbb{E}[X]$ then X and Y are independent.
(ii) If $\mathbb{E}[X \mid \mathcal{G}] = \mathbb{E}[X \mid \mathcal{H}] = 0$ then $X = 0$.
(iii) If X and Y are independent then so are $\mathbb{E}[X \mid \mathcal{G}]$ and $\mathbb{E}[Y \mid \mathcal{G}]$.

Solution. (i) Let X be a standard Gaussian. Then $\mathbb{E}[X \mid X^2] = 0 = \mathbb{E}[X]$, but X and X^2 are not independent.

(ii) Just take $\mathcal{G} = \mathcal{F}$, and $\mathcal{H} = \{\emptyset, \Omega\}$.

(iii) As in Exercise 4, X_1 andX_2 are independent but $\mathbb{E}[X_1 \mid X_1 + X_2]$ and $\mathbb{E}[X_2 \mid X_1 + X_2]$ are not (both equal $(X_1 + X_2)/2$).

Exercise 9. Let $(X_n)_{n\geq 1}$ be a sequence of nonnegative random variables on $(\Omega, \mathcal{A}, \mathbb{P})$, and $(\mathcal{F}_n)_{n\geq 0}$ a sequence of sub σ-fields of $\mathcal{F}$. Assume that $\mathbb{E}(X_n \mid \mathcal{F}_n)$ converges to 0 in probability.

(i) Show that X_n converges to 0 in probability.
(ii) Show that the reciprocal is wrong.

Solution. (i) We have

$$\mathbb{E}\left[\frac{X_n}{1+X_n}\right] = \mathbb{E}\left[\mathbb{E}\left[\frac{X_n}{1+X_n} \mid \mathcal{F}_n\right]\right] \leq \mathbb{E}\left[\frac{\mathbb{E}[X_n \mid \mathcal{F}_n]}{1+\mathbb{E}[X_n \mid \mathcal{F}_n]}\right],$$

where the inequality follows from concavity of $x \mapsto x/(1+x)$ on $\mathbb{R}_+$. We then conclude by dominated convergence.

(ii) For $\mathcal{F}_n = \{\emptyset, \Omega\}$, we have $\mathbb{E}[X_n \mid \mathcal{F}_n] = \mathbb{E}[X_n]$, so that with the choice $X_n = X$ with X with mean 0 and non-identically 0 we have a counterexample.

Exercise 10. On the same probability space, let X, Y be positive random variables such that $\mathbb{E}[X \mid Y] = Y$ and $\mathbb{E}[Y \mid X] = X$ (almost surely). Prove that $X = Y$ almost surely.

Solution. For any $c \geq 0$, we have

$$0 = \mathbb{E}[\mathbb{E}[(X-Y)\mathbb{1}_{X\leq c}|X]] = \mathbb{E}[(X-Y)\mathbb{1}_{X\leq c}] = \mathbb{E}[(X-Y)\mathbb{1}_{X\leq c, Y\leq c}]+\mathbb{E}[(X-Y)\mathbb{1}_{X\leq c, Y>c}],$$

so that

$$\mathbb{E}[(X-Y)\mathbb{1}_{X\leq c, Y\leq c}] = -\mathbb{E}[(X-Y)\mathbb{1}_{X\leq c, Y>c}] \geq 0.$$

Similarly, by permuting the roles of X and Y,

$$\mathbb{E}[(Y-X)\mathbb{1}_{X\leq c, Y\leq c}] \geq 0.$$

These two equations give $\mathbb{E}[(X-Y)\mathbb{1}_{X\leq c, Y\leq c}] = 0$, hence $\mathbb{E}[(X-Y)\mathbb{1}_{X\leq c, Y>c}] = 0$, i.e. $\mathbb{E}[(Y-X)\mathbb{1}_{X\leq c<Y}] = 0$. Thus

$$0 \leq \mathbb{E}[(Y-X)\mathbb{1}_{X<Y}] \leq \sum_{q\in\mathbb{Q}_+} \mathbb{E}[(Y-X)\mathbb{1}_{X\leq q<Y}] = 0.$$

This gives $\mathbb{P}(Y > X) = 0$. By symmetry, this concludes $X = Y$ almost surely.

Note that positivity is only mildly used here, to ensure that all considered expectations are in $\mathbb{R} \cup \{\infty\}$ and no ambiguous $\infty - \infty$ terms ever appear.

Exercise 11. For any $d \geq 1$, we admit that there is only one probability measure μ on $\mathcal{S}_d$, (the $(d-1)$-th dimensional sphere embedded in $\mathbb{R}^d$) that is uniform, in the following sense: for any isometry $A \in \mathrm{O}(d)$ (the orthogonal group in $\mathbb{R}^d$), and any continuous function $f : \mathcal{S}_d \to \mathbb{R}$,

$$\int_{\mathcal{S}_d} f(x)\mathrm{d}\mu(x) = \int_{\mathcal{S}_d} f(Ax)\mathrm{d}\mu(x).$$

Let $X = (X_1, \ldots, X_d)$ be a vector of independent centered and reduced Gaussian random variables.

a) Prove that the random variable $U = X/\|X\|_{L^2}$ is uniformly distributed on the sphere.

b) Prove that, as $d \to \infty$, the main part of the globe is concentrated close to the Equator, i.e. for any $\varepsilon > 0$,

$$\int_{x\in\mathcal{S}_d, |x_1|<\epsilon} \mathrm{d}\mu(x) \to 1.$$

Solution. a) We only need to prove that for any $A \in \mathrm{O}(d)$, AU and U have the same distribution. We will first prove that AX and X have the same distribution, through the characterization by Fourier transform: for any $\lambda \in \mathbb{R}^d$, we have

$$\mathbb{E}(e^{\mathrm{i}\lambda\cdot AX}) = \mathbb{E}(e^{\mathrm{i}\sum_{i,j}\lambda_i A_{ij} X_j}) = e^{-\frac{1}{2}\sum_j(\sum_i \lambda_i A_{ij})^2} = e^{-\frac{1}{2}|A\lambda|^2} = e^{-\frac{1}{2}|\lambda|^2} = \mathbb{E}(e^{\mathrm{i}\lambda\cdot X})$$

where we used the fact that A is an isometry in the fourth equality. Hence for any continuous bounded f we have $\mathbb{E}\, f(AX) = \mathbb{E}\, f(X)$. In particular, picking $f(x) = g(x/|x|)$ with contiunuous bounded g, we obtain $\mathbb{E}\, g(AU) = \mathbb{E}\, g(U)$, which concludes the proof.

b) From the previous question, we have

$$\int_{x\in\mathcal{S}_d,|x_1|<\epsilon} \mathrm{d}\mu(x) = \mathbb{P}(|U_1|<\varepsilon) = \mathbb{P}\left(|X_1|/\sqrt{\sum_{i=1}^{d} X_i^2} < \varepsilon\right) \geq \mathbb{P}\left(|X_1| < \varepsilon\sqrt{\sum_{i=2}^{d} X_i^2}\right).$$

Let $A = \{\sum_{i=2}^{d} X_i^2 > d/100\}$. This RHS is bounded below by

$$\mathbb{P}\left(|X_1| < \varepsilon\sqrt{\sum_{i=2}^{d} X_i^2}, A\right) \geq \mathbb{P}\left(|X_1| < \varepsilon\sqrt{d/100}, A\right) = \mathbb{P}\left(|X_1| < \varepsilon\sqrt{d/100}\right)\cdot\mathbb{P}\left(A\right)$$

By the law of large numbers (or just a second moment computation and Markov's inequality), we know that $\mathbb{P}(A) \to 1$ as $d \to \infty$. This concludes the proof.