ADVANCED PROBABILITY: SOLUTIONS TO SHEET 1

Last compiled: November 6, 2013

1. Conditional expectation

Exercise 1.1. To start with, note that

$$\begin{aligned}\mathbb{P}(X \neq Y) &= \mathbb{P}(\exists c \in \mathbb{R} : X > c, Y \leq c \text{ or } X \leq c, Y > c) \\ &= \mathbb{P}(\exists c \in \mathbb{Q} : X > c, Y \leq c \text{ or } X \leq c, Y > c) \\ &\leq \sum_{c \in \mathbb{Q}} \big[\mathbb{P}(X > c, Y \leq c) + \mathbb{P}(X \leq c, Y > c)\big]\end{aligned}$$

where the second equality follows by the density of $\mathbb{Q}$ in $\mathbb{R}$. So it is enough to show that for all $c \in \mathbb{Q}$ it holds $\mathbb{P}(X > c, Y \leq c) = \mathbb{P}(X \leq c, Y > c) = 0$. To this end, fix any $c \in \mathbb{Q}$. Since $\{Y \leq c\} \in \sigma(Y)$, by definition of conditional expectation $\mathbb{E}[X\mathbf{1}_{\{Y \leq c\}}] = \mathbb{E}[Y\mathbf{1}_{\{Y \leq c\}}]$. It follows that

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

and, by reversing the roles of X and Y, that

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

By adding these equations we then see that

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

Both of the above summands are nonnegative and so they must both equal 0. It follows that $\mathbb{P}(X > c, Y \leq c) = 0$ and $\mathbb{P}(X \leq c, Y > c) = 0$.
If we only know that $\mathbb{E}[X \mid Y] = Y$ a.s. then we cannot conclude as we have above. For example, suppose that $Y = 0$ and that X takes values in $\{-1, 1\}$ with equal probability. Then, trivially, X and Y are integrable. Moreover, since $\sigma(Y) = \{\emptyset, \Omega\}$, from

$$\begin{aligned}\mathbb{E}(X\mathbf{1}(\Omega)) &= \mathbb{E}(X) = 0 = Y \\ \mathbb{E}(X\mathbf{1}(\emptyset)) &= 0 = Y\end{aligned}$$

we conclude that $\mathbb{E}[X \mid Y] = \mathbb{E}[X] = 0 = Y$ a.s. So the assumptions hold, but $X \neq Y$ with probability 1.

Comments and corrections should be sent to Vittoria Silvestri, vs358@cam.ac.uk.

Exercise 1.2. Let X and Y be independent Bernoulli random variables of parameter $p \in (0,1)$ and let us define $Z := \mathbf{1}_{\{X+Y=0\}}$. Since $Z \in \{0,1\}$ almost surely by definition, it follows that

$$\begin{aligned}\mathbb{E}[X \mid Z] &= \mathbb{E}[X \mid Z=0]\mathbf{1}_{\{Z=0\}} + \mathbb{E}[X \mid Z=1]\mathbf{1}_{\{Z=1\}} \\ &= \frac{\mathbb{E}[X\mathbf{1}_{\{Z=0\}}]}{\mathbb{P}(Z=0)}\mathbf{1}_{\{Z=0\}} + \frac{\mathbb{E}[X\mathbf{1}_{\{Z=1\}}]}{\mathbb{P}(Z=1)}\mathbf{1}_{\{Z=1\}} \quad \text{a.s.}\end{aligned}$$

If $Z=1$, then $X=0$, and so the second summand equals 0. For the first summand, observe that

$$\mathbb{P}(Z=0) = 1 - \mathbb{P}(Z=1) = 1 - \mathbb{P}(X=0, Y=0) = 1-(1-p)^2 = p(2-p)$$

and, additionally, that $\mathbb{E}[X\mathbf{1}_{\{Z=0\}}] = \mathbb{P}(Z=0, X=1) = \mathbb{P}(X=1) = p$. It follows that $\mathbb{E}[X \mid Z] = \mathbf{1}_{\{Z=0\}}/(2-p)$ a.s. Finally, by symmetry, $\mathbb{E}[Y \mid Z] = \mathbb{E}[X \mid Z]$ a.s.

Exercise 1.3. Let X and Y be independent exponential random variables of parameter θ and define $Z := X+Y$. In order to see that Z has a $\Gamma(2,\theta)$ distribution, it will suffice for us to show that the density of Z, f_Z, is such that $f_Z(z) = \theta^2 z \mathrm{e}^{-\theta z}\mathbf{1}_{\{z\geq 0\}}$. If $z \leq 0$ then $\mathbb{P}(Z \leq z) = 0$ as X and Y are a.s. positive, so let us consider $z > 0$. As (X,Y) has a density given by $f_{X,Y}(x,y) = f_X(x)f_Y(y) = \theta^2\mathrm{e}^{-\theta(x+y)}\mathbf{1}_{\{x\geq 0, y\geq 0\}}$, we have that

$$\begin{aligned}\mathbb{P}(Z \leq z) = \mathbb{P}(X+Y \leq z) &= \int_{\{x+y\leq z\}} \theta^2 \mathrm{e}^{-\theta(x+y)}\mathbf{1}_{\{x\geq 0,y\geq 0\}}\,\mathrm{d}x\,\mathrm{d}y \\ &= \theta^2 \int_0^z \left(\mathrm{e}^{-\theta y}\int_0^{z-y}\mathrm{e}^{-\theta x}\mathrm{d}x\right)\mathrm{d}y \\ &= \theta\int_0^z (1-\mathrm{e}^{-\theta(z-y)})\mathrm{e}^{-\theta y}\,\mathrm{d}y \\ &= 1 - \mathrm{e}^{-\theta z} - \theta z \mathrm{e}^{-\theta z}.\end{aligned}$$

So the distribution function of Z is $F_Z(z) = (1-\mathrm{e}^{-\theta z} - \theta z\mathrm{e}^{-\theta z})\mathbf{1}_{\{z\geq 0\}}$. Differentiating gives the p.d.f. of Z:

$$F_Z'(z) = f_Z(z) = \theta^2 z\mathrm{e}^{-\theta z}\mathbf{1}_{\{z\geq 0\}}$$

and hence $Z \sim \Gamma(2,\theta)$.

Remark. Alternatively, we could use the general fact that, if $(X_k : k = 1,\dots,n)$ is a sequence of *independent* random variables with respective densities $(f_k : k=1,\dots,n)$, then the sum $X_1 + \cdots + X_n$ has a density given by the convolution $f_1 \star \cdots \star f_n$.

For the second part of the exercise, take $h\colon \mathbb{R}\to\mathbb{R}$ to be a non-negative Borel function. We aim to show that almost surely,

$$\mathbb{E}[h(X) \mid Z] = \frac{1}{Z}\int_0^Z h(x)\,\mathrm{d}x.$$

Let us define $\Phi\colon \mathbb{R}^2\to\mathbb{R}^2\colon (x,y)\mapsto(x,x+y)$. This is a $\mathscr{C}^1$-diffeomorphism and, further, $\Phi^{-1}\colon \mathbb{R}^2\to\mathbb{R}^2\colon (x,z)\mapsto(x,z-x)$ is such that $|\det(D\Phi^{-1})|\equiv 1$. The change of variables formula with Φ applies and tells us that the density of (X,Z) is given by

$$f_{X,X+Y}(x,z) = f_{X,Y}(x,z-x) = f_X(x)f_Y(z-x) = \theta^2\mathrm{e}^{-\theta z}\mathbf{1}_{\{z\geq x\geq 0\}}.$$

Therefore

$$f_{X|Z}(x|z) = \frac{f_{X,Z}(x,z)}{f_Z(z)}\mathbf{1}_{\{z : f_Z(z)>0\}}$$

and

$$\begin{aligned}\mathbb{E}[h(X) \mid Z] &= \left(\int_{\mathbb{R}} h(x)\frac{f_{X,Z}(x,Z)}{f_Z(Z)}\,\mathrm{d}x\right)\mathbf{1}_{\{f_Z(Z)>0\}}\\ &= \int_{\mathbb{R}} h(x)\frac{\theta^2\mathrm{e}^{-\theta Z}}{\theta^2 Z\mathrm{e}^{-\theta Z}}\mathbf{1}_{\{Z\geq x>0\}}\,\mathrm{d}x\\ &= \frac{1}{Z}\int_0^Z h(x)\,\mathrm{d}x \quad \text{a.s.}\end{aligned}$$

To answer the third part of the question, suppose that $Z\sim\Gamma(2,\theta)$ and that, for every non-negative Borel function $h\colon\mathbb{R}\to\mathbb{R}$, it holds

$$\mathbb{E}[h(X) \mid Z] = \frac{1}{Z}\int_0^Z h(u)\,\mathrm{d}u \quad \text{a.s.}$$

Then we aim to show that X and $Z-X$ are independent exponential random variables of parameter θ. To this end, it is enough to show that the joint distribution of (X,Y) with $Y:=Z-X$ factorizes, and the marginal densities are Gamma densities with parameter θ. To this end we can apply the change of variables formula with Φ^{-1}, which yields

$$\begin{aligned}f_{X,Z-X}(x,y) &= f_{X,Z}(x,x+y) = \theta^2\mathrm{e}^{-\theta(x+y)}\mathbf{1}_{\{x+y\geq x\geq 0\}}\\ &= \theta\mathrm{e}^{-\theta x}\mathbf{1}_{\{x\geq 0\}}\theta\mathrm{e}^{-\theta y}\mathbf{1}_{\{y\geq 0\}} = f_X(x)f_Y(y)\end{aligned}$$

with f_X, f_Y being p.d.f. of Gamma(θ).

Exercise 1.4. Let X be a non-negative random variable on the probability space $(\Omega,\mathscr{F},\mathbb{P})$ and $\mathscr{G}$ be a sub-σ-algebra of $\mathscr{F}$. We first show that if $X>0$, then $\mathbb{E}[X \mid \mathscr{G}]>0$ a.s. As $A:=\{\mathbb{E}[X \mid \mathscr{G}]\leq 0\}\in\mathscr{G}$, we have that $0\leq\mathbb{E}[X\mathbf{1}_A]=\mathbb{E}[\mathbb{E}[X \mid \mathscr{G}]\mathbf{1}_A]\leq 0$ and hence that $\mathbb{E}[X\mathbf{1}_A]=0$. Since $X>0$ on A, we conclude that $\mathbb{P}(A)=0$.

To see that the event $\{\mathbb{E}[X \mid \mathscr{G}]>0\}$ is the smallest element of $\mathscr{G}$ that contains the event $\{X>0\}$ (up to null events), assume the contrary. Then there exists a $\mathscr{G}$-measurable event B such that $\{X > 0\} \subseteq B \subseteq \{\mathbb{E}[X \mid \mathscr{G}] > 0\}$ and $C := \{\mathbb{E}[X|\mathscr{G}] > 0\} \setminus B$ has positive measure. Then $0 \geq \mathbb{E}[X\mathbf{1}_C] = \mathbb{E}[\mathbb{E}[X \mid \mathscr{G}]\mathbf{1}_C] \geq 0$ and so $\mathbb{E}[\mathbb{E}[X \mid \mathscr{G}]\mathbf{1}_C] = 0$. As $\mathbb{E}[X \mid \mathscr{G}] > 0$ on C, it must be the case that $\mathbb{P}(C) = 0$, which contradicts the assumptions. It follows that $\{\mathbb{E}[X \mid \mathscr{G}] > 0\} \subseteq B$ up to a null event.

Exercise 1.5. Recall that

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

for $n \geq 0$ integer, $t \geq 0$ real. In order to compute $\mathbb{E}[h(Y)\mathbf{1}_{\{X=n\}}]$ we first compute the p.d.f. of Y conditional on the event $X = n$.

$$\begin{aligned} f_{Y|X=n}(y|X=n) &= \frac{\mathrm{d}}{\mathrm{dt}}\mathbb{P}(Y \leq t|X=n)\Big|_{t=y} = \frac{\mathrm{d}}{\mathrm{dt}}\left(\frac{\mathbb{P}(X=n, Y \leq t)}{\mathbb{P}(X=n)}\right)\Bigg|_{t=y} \\ &= \frac{b}{\mathbb{P}(X=n)}\frac{(ay)^n}{n!}\mathrm{e}^{-(a+b)y}. \end{aligned}$$

Moreover

$$\begin{aligned} \mathbb{P}(X=n) = \mathbb{P}(X=n, Y<\infty) &= b\int_0^\infty \frac{(ay)^n}{n!}\mathrm{e}^{-(a+b)y}\,\mathrm{d}y \\ &= \frac{ba^n}{(a+b)^{n+1}}\underbrace{\int_0^\infty \frac{(a+b)^{n+1}}{n!}y^n\mathrm{e}^{-(a+b)y}\,\mathrm{d}y}_{1} = \frac{ba^n}{(a+b)^{n+1}} \end{aligned}$$

where in the last equality we have used that the p.d.f. of a Gamma$(n+1, a+b)$ integrates to 1, recalling that $\Gamma(n+1) = n!$ for n positive integer. Hence

$$f_{Y|X=n}(y|X=n) = b\frac{(a+b)^{n+1}}{ba^n}\frac{(ay)^n}{n!}\mathrm{e}^{-(a+b)y} = \frac{(a+b)^{n+1}}{\Gamma(n+1)}y^n\mathrm{e}^{-(a+b)y}\mathbf{1}_{(0,+\infty)}(y).$$

In other words, the law of Y conditioned to the event $\{X = n\}$ is $\Gamma(n+1, a+b)$. Therefore if $h\colon (0,\infty) \to [0,\infty)$ is a Borel function, then $\mathbb{E}[h(Y) \mid X = n] = \mathbb{E}[h(G)]$ where G is another random variable, defined on the same probability space, with law Gamma$(n+1, a+b)$. Hence

$$\mathbb{E}[h(Y) \mid X=n] = \frac{(a+b)^{n+1}}{n!}\int_0^\infty h(y)y^n\mathrm{e}^{-(a+b)y}\,\mathrm{d}y.$$

Next we compute $\mathbb{E}(Y/(X+1))$:

$$\begin{aligned} \mathbb{E}\left(\frac{Y}{X+1}\right) &= \sum_{n=0}^\infty \mathbb{E}\left(\frac{Y}{X+1}\middle| X=n\right)\mathbb{P}(X=n) = \sum_{n=0}^\infty \mathbb{E}\left(\frac{Y}{n+1}\right)\mathbb{P}(X=n) \\ &= \sum_{n=0}^\infty \left(\frac{n+1}{a+b}\right)\frac{1}{n+1}\mathbb{P}(X=n) = \frac{1}{a+b}\underbrace{\sum_{n=0}^\infty \mathbb{P}(X=n)}_{1} = \frac{1}{a+b}. \end{aligned}$$

For $\mathbb{E}(\mathbf{1}_{\{X=n\}}|Y)$ we have:

$$\mathbb{E}(\mathbf{1}_{\{X=n\}}|Y) = \mathbb{P}(X=n|Y) = \frac{f_{X,Y}(n,Y)}{f_Y(Y)}\,.$$

Since

$$f_{X,Y}(n,y) = \frac{\mathrm{d}}{\mathrm{dt}}\mathbb{P}(X=n, Y\le t)\big|_{t=y} = b\frac{(ay)^n}{n!}e^{-(a+b)y}\mathbf{1}_{(0,\infty)}(y)$$

and

$$f_Y(y) = \sum_{n=0}^{\infty} f_{X,Y}(n,y) = b\sum_{n=0}^{\infty}\frac{(ay)^n}{n!}e^{-(a+b)y}\mathbf{1}_{(0,\infty)}(y) = be^{-by}\mathbf{1}_{(0,\infty)}(y)\,,$$

from which $Y \sim \text{exponential}(b)$, we conclude

$$\mathbb{P}(X=n|Y) = b\frac{(aY)^n}{n!}e^{-(a+b)Y}\cdot\frac{1}{be^{-by}} = \frac{(aY)^n e^{-aY}}{n!}\,.$$

That is, the law of X conditional on Y is Poisson(aY). This also implies that $\mathbb{E}(X|Y) = aY$ (recall that the expected value of a Poisson random variable of parameter λ is λ).

Exercise 1.6. Let us suppose that X and Y are random variables defined on the probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and that $\mathscr{G}$ is a sub-σ-algebra of $\mathscr{F}$. We say that X and Y are *conditionally independent given* $\mathscr{G}$ if, for all Borel functions $f, g\colon \mathbb{R}\to[0,\infty)$,

$$\mathbb{E}[f(X)g(Y) \mid \mathscr{G}] = \mathbb{E}[f(X) \mid \mathscr{G}]\mathbb{E}[g(Y) \mid \mathscr{G}] \tag{1}$$

almost surely. If $\mathscr{G} = \{\emptyset, \Omega\}$ in the above then this implies that

$$\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]$$

for all non-negative Borel functions f and g. In particular, if we take $f = \mathbf{1}_A$ and $g = \mathbf{1}_B$ for $A, B \in \mathscr{B}$, this implies that

$$\mathbb{P}(X\in A, Y\in B) = \mathbb{E}[\mathbf{1}_A(X)\mathbf{1}_B(Y)] = \mathbb{E}[\mathbf{1}_A(X)]\mathbb{E}[\mathbf{1}_B(Y)] = \mathbb{P}(X\in A)\mathbb{P}(Y\in B).$$

That is to say, if X and Y are independent conditionally on $\{\emptyset, \Omega\}$, then they are independent. (The converse is also true by linearity and the monotone convergence theorem.)

We next show that the random variables X and Y are conditionally independent given $\mathscr{G}$ if and only if, for every non-negative $\mathscr{G}$-measurable random variable Z and all non-negative Borel functions f and g,

$$\mathbb{E}[f(X)g(Y)Z] = \mathbb{E}[f(X)Z\mathbb{E}[g(Y) \mid \mathscr{G}]]. \tag{2}$$

Suppose first that X and Y are independent conditionally on $\mathscr{G}$. Then

$$\begin{aligned}\mathbb{E}[f(X)g(Y)Z] = \mathbb{E}[\mathbb{E}[f(X)g(Y)Z \mid \mathscr{G}]] &= \mathbb{E}[Z\mathbb{E}[f(X)g(Y) \mid \mathscr{G}]] \\ &= \mathbb{E}[Z\mathbb{E}[f(X) \mid \mathscr{G}]\mathbb{E}[g(Y) \mid \mathscr{G}]]\end{aligned}$$

with the second equality holding as Z is $\mathscr{G}$-measurable and as Z, f and g are non-negative. Further, as $Z\mathbb{E}[g(Y) \mid \mathscr{G}]$ is $\mathscr{G}$-measurable and as everything is a.s. non-negative,

$$\mathbb{E}[Z\mathbb{E}[f(X) \mid \mathscr{G}]\mathbb{E}[g(Y) \mid \mathscr{G}]] = \mathbb{E}[f(X)Z\mathbb{E}[g(Y) \mid \mathscr{G}]].$$

Now assume that (2) holds for every $\mathscr{G}$-measurable Z and f, g non-negative Borel functions. We aim to show that this implies (1). Let us take $A \in \mathscr{G}$ and $Z = \mathbf{1}_A$. We have that

$$\begin{aligned}\mathbb{E}[\mathbb{E}[f(X)g(Y) \mid \mathscr{G}]\mathbf{1}_A] = \mathbb{E}[f(X)g(Y)\mathbf{1}_A] &= \mathbb{E}[f(X)\mathbf{1}_A\mathbb{E}[g(Y) \mid \mathscr{G}]] \\ &= \mathbb{E}[\mathbb{E}[f(X) \mid \mathscr{G}]\mathbb{E}[g(Y) \mid \mathscr{G}]\mathbf{1}_A].\end{aligned}$$

The second equality follows from our hypothesis; the final equality holds as $\mathbf{1}_A\mathbb{E}[g(Y) \mid \mathscr{G}]$ is $\mathscr{G}$-measurable and as everything is a.s. non-negative. Therefore, as $\mathbb{E}[f(X) \mid \mathscr{G}]\mathbb{E}[g(Y) \mid \mathscr{G}]$ is $\mathscr{G}$-measurable and a.s. non-negative it follows that, with probability 1,

$$\mathbb{E}[f(X)g(Y) \mid \mathscr{G}] = \mathbb{E}\big(\mathbb{E}[f(X) \mid \mathscr{G}]\mathbb{E}[g(Y) \mid \mathscr{G}]\big|\mathscr{G}\big) = \mathbb{E}[f(X) \mid \mathscr{G}]\mathbb{E}[g(Y) \mid \mathscr{G}]\,.$$

For the last part of the exercise, we have to show that

$$\mathbb{E}[f(X)g(Y)Z] = \mathbb{E}[f(X)Z\mathbb{E}[g(Y) \mid \mathscr{G}]] \tag{3}$$

for every $\mathscr{G}$-measurable random variable Z and all Borel functions $f, g \colon \mathbb{R} \to [0, \infty)$ if and only if, for each Borel function $g \colon \mathbb{R} \to [0, \infty)$,

$$\mathbb{E}[g(Y) \mid \sigma(\mathscr{G}, \sigma(X))] = \mathbb{E}[g(Y) \mid \mathscr{G}]. \tag{4}$$

Assume (3). It is immediate that $\mathbb{E}[g(Y) \mid \mathscr{G}]$ is $\sigma(\mathscr{G}, \sigma(X))$-measurable. We are to show that, for all $A \in \sigma(\mathscr{G}, \sigma(X))$, $\mathbb{E}[g(Y)\mathbf{1}_A] = \mathbb{E}[\mathbb{E}[g(Y) \mid \mathscr{G}]\mathbf{1}_A]$. It suffices, by the theorem on the uniqueness of extensions, to prove this for all $A \cap B$, where $A \in \mathscr{G}$ and $B \in \sigma(X)$, as the set $\{A \cap B : A \in \mathscr{G}, B \in \sigma(X)\}$ is a generating π-system for $\sigma(\mathscr{G}, \sigma(X))$ that contains Ω. So let $A \in \mathscr{G}$ and $B \in \sigma(X)$. Then

$$\mathbb{E}[g(Y)\mathbf{1}_{A\cap B}] = \mathbb{E}[\underbrace{\mathbf{1}_B}_{f(X)} g(Y) \underbrace{\mathbf{1}_A}_{Z}] = \mathbb{E}[\mathbf{1}_B\mathbf{1}_A\mathbb{E}[g(Y) \mid \mathscr{G}]] = \mathbb{E}[\mathbb{E}[g(Y) \mid \mathscr{G}]\mathbf{1}_{A\cap B}]$$

where we have used (3) in the second equality.

Assume now that (4) holds. As $f(X)Z$ is non-negative and $\sigma(\mathscr{G}, \sigma(X))$-measurable and as $g(Y)$ is non-negative,

$$\begin{aligned}\mathbb{E}[g(Y)f(X)Z] &= \mathbb{E}[\mathbb{E}[g(Y)f(X)Z \mid \sigma(\mathscr{G}, \sigma(X))]] \\ &= \mathbb{E}[\mathbb{E}[g(Y) \mid \sigma(\mathscr{G}, \sigma(X))]f(X)Z] \\ &= \mathbb{E}[\mathbb{E}[g(Y) \mid \mathscr{G}]f(X)Z]\end{aligned}$$

where, we have used (4) in the second equality.

Exercise 1.7. Recall that, given a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, two sub-$\sigma$-algebras $\mathscr{G}, \mathscr{H}$ of $\mathscr{F}$ are said to be independent if for every $G \in \mathscr{G}$, $H \in \mathscr{H}$ it holds $\mathbb{P}(G \cap H) = \mathbb{P}(G)\mathbb{P}(H)$. Moreover, we know (see Proposition 1.24 in the lecture notes) that if $\sigma(X, \mathscr{G})$ is independent of $\mathscr{H}$, then $\mathbb{E}(X|\sigma(\mathscr{G}, \mathscr{H})) = \mathbb{E}(X|\mathscr{G})$ a.s.. Hence we seek for an example in which $\sigma(X, \mathscr{G})$ and $\mathscr{H}$ are dependent.

Let X_1, X_2 be independent Bernoulli(1/2), and set $\mathscr{H} = \sigma(X_1)$, $\mathscr{G} = \sigma(X_2)$ and

$$X = \mathbf{1}(X_1 = X_2)\,.$$

Then $\mathscr{H}$ and $\mathscr{G}$ are independent by construction. Moreover, X is itself a Bernoulli(1/2) random variable and it is independent of $\mathscr{H}$, since:

$$\begin{aligned}\mathbb{P}(X = 0, X_1 = 0) &= \mathbb{P}(X_1 = 0, X_2 = 1) = \mathbb{P}(X_1 = 0)\mathbb{P}(X_2 = 1) \\ &= 1/4 = \mathbb{P}(X = 0)\mathbb{P}(X_1 = 0)\end{aligned}$$

and similarly one sees that $\mathbb{P}(X = x, X_1 = y) = \mathbb{P}(X = x)\mathbb{P}(X_1 = y)$ for all $x, y \in \{0, 1\}$. So the assumptions are satisfied. Now notice that $\sigma(\mathscr{G}, \mathscr{H}) = \mathscr{F}$ and therefore X is $\sigma(\mathscr{G}, \mathscr{H})$-measurable, from which

$$\mathbb{E}(X|\sigma(\mathscr{G}, \mathscr{H})) = X \quad a.s.$$

On the other hand,

$$\begin{aligned}\mathbb{E}(X|\mathscr{G}) &= \mathbb{E}(\mathbf{1}(X_1 = X_2)|X_2 = 0)\mathbf{1}(X_2 = 0) + \mathbb{E}(\mathbf{1}(X_1 = X_2)|X_2 = 1)\mathbf{1}(X_2 = 1) \\ &= \mathbb{P}(X_1 = 0)\mathbf{1}(X_2 = 0) + \mathbb{P}(X_1 = 1)\mathbf{1}(X_2 = 1) \\ &= \frac{1}{2}\mathbf{1}(X_2 = 0) + \frac{1}{2}\mathbf{1}(X_2 = 1) = \frac{1}{2}\end{aligned}$$

almost surely, since with probability 1 exactly one of the indicator functions in the last line is non-zero. But then

$$\mathbb{E}(X|\sigma(\mathscr{G}, \mathscr{H})) \neq \mathbb{E}(X|\mathscr{G})$$

on an event of probability 1, since X takes values in $\{0, 1\}$.

2. Discrete-time martingales

Exercise 2.1. Assume that X is a martingale with respect to its natural filtration. Then, since the event $\{X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n\}$ is $\mathscr{F}_n$-measurable for all $n \geq 0$ and $i_0, \ldots, i_n \in E$, we have

$$\begin{aligned}\mathbb{E}(X_{n+1}|X_0 = i_0, \ldots, X_n = i_n) &= \frac{\mathbb{E}(X_{n+1}\mathbf{1}(X_0 = i_0, \ldots, X_n = i_n))}{\mathbb{P}(X_0 = i_0, \ldots, X_n = i_n)} \\ &= \frac{\mathbb{E}(\mathbb{E}(X_{n+1}|\mathscr{F}_n)\mathbf{1}(X_0 = i_0, \ldots, X_n = i_n))}{\mathbb{P}(X_0 = i_0, \ldots, X_n = i_n)} \\ &= \frac{\mathbb{E}(X_n\mathbf{1}(X_0 = i_0, \ldots, X_n = i_n))}{\mathbb{P}(X_0 = i_0, \ldots, X_n = i_n)} \\ &= \frac{\mathbb{E}(i_n\mathbf{1}(X_0 = i_0, \ldots, X_n = i_n))}{\mathbb{P}(X_0 = i_0, \ldots, X_n = i_n)} = i_n\end{aligned}$$

almost surely.

Assume, on the other hand, that for all $n \geq 0$ and all $i_0, i_1, \ldots, i_n \in E$ it holds

$$\mathbb{E}(X_{n+1}|X_0 = i_0, \ldots, X_n = i_n) = i_n \quad a.s.$$

Then, being E countable, we can write

$$\begin{aligned}\mathbb{E}(X_{n+1}|\mathscr{F}_n) &= \sum_{i_0,\ldots,i_n \in E} \underbrace{\mathbb{E}(X_{n+1}|X_0 = i_0, \ldots, X_n = i_n)}_{i_n} \mathbf{1}(X_0 = i_0, \ldots, X_n = i_n) \\ &= \sum_{i_n \in E} i_n \mathbf{1}(X_n = i_n) = X_n \quad a.s.\end{aligned}$$

To conclude that this implies that X is a martingale, take any $m > n$. Then using the tower property of the expectation we get

$$\begin{aligned}\mathbb{E}(X_m|\mathscr{F}_n) &= \mathbb{E}(\mathbb{E}(X_m|\mathscr{F}_{m-1})|\mathscr{F}_n) = \mathbb{E}(X_{m-1}|\mathscr{F}_n) \\ &= \ldots = \mathbb{E}(X_{n+1}|\mathscr{F}_n) = X_n \quad a.s.\end{aligned}$$

Exercise 2.2. We say that a process $C = (C_n : n \geq 0)$ is *previsible* with respect to the filtration $(\mathscr{F}_n : n \geq 0)$ if C_{n+1} is $\mathscr{F}_n$-measurable for all $n \geq 0$. Suppose that $X = (X_n)_{n \geq 0}$ is a martingale on the filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_n : n \geq 0), \mathbb{P})$, and the previsible process C is bounded. We aim to show that $Y := C \bullet X$ is a martingale.

It is clear that Y is adapted and integrable. Moreover, for each $n \geq 0$ we have:

$$\begin{aligned}\mathbb{E}[Y_{n+1} - Y_n \mid \mathscr{F}_n] &= \mathbb{E}[(C \bullet X)_{n+1} - (C \bullet X)_n \mid \mathscr{F}_n] \\ &= \mathbb{E}[C_{n+1}(X_{n+1} - X_n) \mid \mathscr{F}_n] = C_{n+1}\mathbb{E}[X_{n+1} - X_n \mid \mathscr{F}_n] = 0\end{aligned}$$

where the second equality holds as C_{n+1} is bounded and $\mathscr{F}_n$-measurable and $X_{n+1} - X_n$ is integrable, and the final equality holds by the martingale property of X. This shows that Y is itself a martingale. If, on the other hand, X is a supermartingale and C is bounded and non-negative, then

$$\mathbb{E}[Y_{n+1}-Y_n \mid \mathscr{F}_n] = \mathbb{E}[C_{n+1}(X_{n+1}-X_n) \mid \mathscr{F}_n] = C_{n+1}\mathbb{E}[X_{n+1}-X_n \mid \mathscr{F}_n] \leq 0$$

where the second equality holds as C_{n+1} is bounded and $\mathscr{F}_n$-measurable and $X_{n+1} - X_n$ is integrable, while the inequality holds as C is non-negative and as X is a supermartingale. It follows that Y is a supermartingale.

Exercise 2.3. Let $(X_n : n \geq 1)$ be a sequence of independent random variables such that $\mathbb{P}(X_n = -n^2) = n^{-2}$ and $\mathbb{P}(X_n = n^2/(n^2-1)) = 1 - n^{-2}$. We aim to show that if $S_n = X_1 + \cdots + X_n$, then $S_n/n \to 1$ a.s. To this end, define the collection of events $A_n := \{X_n = -n^2\}$. Then, by the first Borel–Cantelli lemma, $\mathbb{P}(A_n \text{ i.o.}) = 0$. It follows that $\mathbb{P}(A_n^c \text{ a.a.}) = 1$ and hence that $\mathbb{P}(X_n = n^2/(n^2-1) \text{ a.a.}) = 1$. It is thus enough to prove that $S_n/n \to 1$ on this lattermost set. By definition, for each ω in this set there is some $N_\omega \in \mathbb{Z}_{>0}$ such that $X_n(\omega) = n^2/(n^2-1)$ for all $n \geq N_\omega$. We thus see that

$$\begin{aligned}\frac{S_n(\omega)}{n} &= \underbrace{\frac{1}{n}\sum_{k=1}^{N_\omega-1} X_k(\omega)}_{C_\omega} + \frac{1}{n}\sum_{k=N_\omega}^{n} \frac{k^2}{k^2-1} = \frac{C_\omega}{n} + \frac{1}{n}\sum_{k=N_\omega}^{n}\left(1 + \frac{1}{k^2-1}\right) \\ &= \frac{C_\omega}{n} + \frac{n-N_\omega}{n} + \frac{1}{n}\sum_{k=N_\omega}^{n}\frac{1}{k^2-1} = 1 + o(1)\end{aligned}$$

and therefore $S_n(\omega)/n \to 1$ as $n \to \infty$. Since this holds for all ω in a set o measure 1, we conclude $S_n/n \to 1$ almost surely.

We now want to show that the process $(S_n : n \geq 1)$ is a martingale with respect to its natural filtration, $(\mathscr{F}_n : n \geq 1)$, and that it converges to ∞ a.s. It is clear that $(S_n : n \geq 1)$ is adapted to its natural filtration. Moreover, as S_n is a finite sum of integrable random variables, it is integrable. For the martingale property we note that

$$\mathbb{E}[S_{n+1} - S_n \mid \mathscr{F}_n] = \mathbb{E}[X_{n+1} \mid \mathscr{F}_n] = \mathbb{E}[X_{n+1}] = 0,$$

where the second equality follows from the independence of X_{n+1} from $\mathscr{F}_n$. It follows that $(S_n : n \geq 1)$ is a martingale. Finally, for all ω such that $S_n(\omega)/n \to 1$, it is immediate that $S_n(\omega) \to \infty$. As the set of all such ω has full measure, $S_n \to \infty$ a.s.

Exercise 2.4. Let $T := m\mathbf{1}_A + m'\mathbf{1}_{A^c}$. We see that $\{T = m\} = A \in \mathscr{F}_n \subseteq \mathscr{F}_m$ and $\{T = m'\} = A^c \in \mathscr{F}_n \subseteq \mathscr{F}_{m'}$. Moreover, if k is neither m nor m', then $\{T = k\} = \emptyset \in \mathscr{F}_k$. It follows that T is a stopping time.

For the second part of the exercise, assume that X is a martingale. Then, if T is a bounded stopping time, $\mathbb{E}[X_T] = \mathbb{E}[X_0]$ by the Optional Stopping Theorem. Assume, on the other hand, that for every T bounded stopping time it holds $\mathbb{E}[X_T] = \mathbb{E}[X_0]$. Then, since n is a bounded stopping time, it holds $\mathbb{E}[X_n] = \mathbb{E}[X_0]$ for all $n \geq 0$. Moreover, we can consider any $A \in \mathscr{F}_n$ and define $T := (n+1)\mathbf{1}_A + n\mathbf{1}_{A^c}$. By the previous proposition this is a bounded stopping time so, again by our hypothesis,

$$\mathbb{E}[X_0] = \mathbb{E}[X_T] = \mathbb{E}[X_{n+1}\mathbf{1}_A + X_n\mathbf{1}_{A^c}] = \mathbb{E}[(X_{n+1} - X_n)\mathbf{1}_A] + \mathbb{E}[X_n].$$

As $\mathbb{E}[X_0] = \mathbb{E}[X_n]$, the above implies that $\mathbb{E}[X_{n+1}\mathbf{1}_A] = \mathbb{E}[X_n\mathbf{1}_A]$. As we have assumed that X is integrable and adapted and as $A \in \mathscr{F}_n$ was arbitrary, $\mathbb{E}[X_{n+1} \mid \mathscr{F}_n] = X_n$ a.s. We conclude that X is a martingale.

Exercise 2.5. Suppose that $X = (X_n : n \geq 0)$ is a martingale (respectively, a supermartingale) and that T is an a.s. finite stopping time. We aim to show that if there is some $M \geq 0$ such that $|X| \leq M$ a.s. then $\mathbb{E}[X_T] = \mathbb{E}[X_0]$ (respectively, $\mathbb{E}[X_T] \leq \mathbb{E}[X_0]$). To see this, recall that X^T is a martingale (respectively, a supermartingale). Since $T < \infty$ a.s., it follows that $T \wedge n \to T$ a.s. and therefore

$$\mathbb{E}[X_T] = \mathbb{E}\left[\lim_{n\to\infty} X_{T\wedge n}\right] = \lim_{n\to\infty} \mathbb{E}[X_{T\wedge n}] = \lim_{n\to\infty} \mathbb{E}[X_{T\wedge 0}] = \mathbb{E}[X_0].$$

The second equality holds by the dominated convergence theorem with M as the dominating (degenerate) random variable, the third equality holds as X^T is a martingale. In the case where X is a supermartingale, the third equality becomes '$\leq$' by the supermartingale property.

For the second part of the exercise, assume that $\mathbb{E}(T) < \infty$ and that X is a martingale (respectively, a supermartingale) with bounded increments. We aim to show that $\mathbb{E}[X_T] = \mathbb{E}[X_0]$ (respectively, $\mathbb{E}[X_T] \leq \mathbb{E}[X_0]$).

Being T integrable, it is a.s. finite, so $T \wedge n \to T$ a.s. Now, for all $n \geq 0$, we a.s. have that

$$|X_{T\wedge n}| = \left| X_0 + \sum_{k=1}^{T\wedge n} (X_k - X_{k-1}) \right| \leq |X_0| + \sum_{k=1}^{T\wedge n} |X_k - X_{k-1}| \leq |X_0| + MT.$$

The far right-hand side of the above is integrable, so we can apply the dominated convergence theorem to get

$$\mathbb{E}[X_T] = \mathbb{E}\left[\lim_{n\to\infty} X_{T\wedge n}\right] = \lim_{n\to\infty} \mathbb{E}[X_{T\wedge n}] = \lim_{n\to\infty} \mathbb{E}[X_{T\wedge 0}] = \mathbb{E}[X_0]\,,$$

where the third equality holds as X^T is a martingale. As in the proof of the previous proposition, in the case where X is a supermartingale the third equality becomes '$\leq$' by the supermartingale property.

Exercise 2.6. Let T be a $\mathscr{F}_n$-stopping time, and suppose that for some integer $N > 0$ and some $\varepsilon > 0$ it holds

$$\mathbb{P}(T \leq N + n|\mathscr{F}_n) \geq \varepsilon \qquad \forall n \geq 0 .$$

We aim to show that $\mathbb{E}(T) < \infty$.

To start with, we show that $\mathbb{P}(T > kN) \leq (1-\varepsilon)^k$ for all $k \geq 0$.

$$\begin{aligned}
\mathbb{P}(T > kN) &= \mathbb{P}(T > N, T > 2N, \dots, T > kN) \\
&= \mathbb{E}(\mathbb{P}(T > N, T > 2N, \dots, T > kN|\mathscr{F}_N)) \\
&= \mathbb{E}\big(\mathbf{1}(T > N)\underbrace{\mathbb{P}(T > 2N, \dots, T > kN|\mathscr{F}_N)}_{\text{iterate}}\big) \\
&= \mathbb{E}\Big(\mathbf{1}(T > N)\mathbb{E}\big(\mathbf{1}(T > 2N, \dots, T > kN)|\mathscr{F}_N\big)\Big) \\
&= \mathbb{E}\Big(\mathbf{1}(T > N)\mathbb{E}\big(\mathbb{E}(\mathbf{1}(T > 2N, \dots, T > kN)|\mathscr{F}_{2N})|\mathscr{F}_N\big)\Big) \\
&= \mathbb{E}\Bigg(\mathbf{1}(T > N)\mathbb{E}\Bigg(\mathbf{1}(T > 2N)\underbrace{\mathbb{P}(T > 3N, \dots, T > kN|\mathscr{F}_{2N})}_{\text{iterate}}\Bigg|\mathscr{F}_N\Bigg)\Bigg) \\
&= \dots \\
&= \mathbb{E}\Bigg(\mathbf{1}(T > N)\mathbb{E}\Bigg(\mathbf{1}(T > 2N)\cdots\underbrace{\mathbb{P}\Bigg(T > kN\Bigg|\mathscr{F}_{(k-1)N}\Bigg)}_{\leq 1-\varepsilon}\cdots\Bigg|\mathscr{F}_N\Bigg) \\
&\leq (1-\varepsilon)^k .
\end{aligned}$$

From this we can conclude, since

$$\mathbb{E}(T) = \sum_{n=0}^{\infty}\mathbb{P}(T \geq n) \leq \sum_{n=0}^{\infty}\mathbb{P}(T \geq k_n N) = N\sum_{k=0}^{\infty}\mathbb{P}(T \geq kN) \leq N\sum_{k=0}^{\infty}(1-\varepsilon)^k = \frac{N}{\varepsilon} < \infty$$

where we have set $k_n = \lfloor n/N \rfloor$.

Exercise 2.7. We are playing a game in which our winnings per unit stake on the game at time $n \geq 1$ is ε_n, where $(\varepsilon_n : n \geq 1)$ is an i.i.d. sequence of random variables such that

$$\mathbb{P}(\varepsilon_n = 1) = \begin{cases} 1\,, & \text{with probability } p \\ -1\,, & \text{with probability } q = 1-p \end{cases}$$

where $p \in (1/2, 1)$. Let Z_n denote our fortune at time n and let C_n be our stake on the game at time n, with $0 \leq C_n < Z_{n-1}$. Our goal in this game is to choose a strategy that maximises our expected interest rate $\mathbb{E}[\log(Z_N/Z_0)]$, where Z_0 is our initial fortune and N is some fixed time corresponding to how long we are to play the game.

Assume that C is a previsible process. Then we aim to show that $(\log Z_n - n\alpha : n \geq 1)$ is a supermartingale, where α denotes the *entropy*

$$\alpha := p \log p + q \log q + \log 2.$$

To see this, note that for each $n \geq 0$, $Z_{n+1} = Z_n + \varepsilon_{n+1} C_{n+1}$ by definition. Therefore:

$$\begin{aligned}
\mathbb{E}[\log Z_{n+1} \mid \mathscr{F}_n] &= \mathbb{E}[\log(Z_n + \varepsilon_{n+1} C_{n+1}) | \mathscr{F}_n] \\
&= \mathbb{E}[\log(Z_n + C_{n+1}) \mathbf{1}(\varepsilon_{n+1} = 1) | \mathscr{F}_n] + \mathbb{E}[\log(Z_n - C_{n+1}) \mathbf{1}(\varepsilon_{n+1} = -1) | \mathscr{F}_n] \\
&= p \log(Z_n + C_{n+1}) + q \log(Z_n - C_{n+1}) \\
&= p \log\left(Z_n \big(1 + \frac{C_{n+1}}{Z_n}\big)\right) + q \log\left(Z_n \big(1 - \frac{C_{n+1}}{Z_n}\big)\right) \\
&= \log Z_n + p \log\left(1 + \frac{C_{n+1}}{Z_n}\right) + q \log\left(1 - \frac{C_{n+1}}{Z_n}\right).
\end{aligned}$$

almost surely, where we have used that the quantity $\log(Z_n \pm C_{n+1})$ is $\mathscr{F}_n$-measurable. Let us now define the function

$$f(x) := p \log(1 + x) + q \log(1 - x)$$

for $x \in [0, 1)$. By elementary calculus, f is maximised at $p - q$ and $f(p - q) = \log 2 + p \log p + q \log q = \alpha$. Therefore

$$\mathbb{E}[\log Z_{n+1} \mid \mathscr{F}_n] = \log Z_n + f(C_{n+1}/Z_n) \leq \log Z_n + \alpha \qquad a.s. \tag{5}$$

from which

$$\mathbb{E}[\log Z_{n+1} - (n + 1)\alpha \mid \mathscr{F}_n] \leq \log Z_n - n\alpha \qquad a.s.$$

The function $\log Z_n - n\alpha$ is clearly $\mathscr{F}_n$-measurable and integrable for each $n \geq 0$, and hence we conclude that the process $(\log Z_n - n\alpha : n \geq 0)$ is a supermartingale. Note that $\mathbb{E}[\log(Z_N / Z_0)] = \mathbb{E}[\log Z_N] - \log Z_0 \leq N\alpha$.

Finally, it is clear that if we take $C_{n+1} := (p - q) Z_n$ in (5) then all inequalities above become equalities, and so $(\log Z_n - n\alpha : n \geq 0)$ is in fact a martingale for this strategy. As $p - q$ is the unique maximiser of f, this strategy is the unique optimal strategy.

References

[Wil91] David Williams, *Probability with martingales*, Cambridge Mathematical Textbooks, Cambridge University Press, Cambridge, 1991.