**Quantum Probability: solutions for exercise set 1**

Typically the abstract probability space is denoted by $(\Omega, \mathcal{F}, \mathsf{P})$. Unless otherwise stated, we assume that a r.v. $X \to S$, where $S$ is a metric space, is measurable w.r.t. Borel $\sigma$ algebra $\mathcal{B}(S)$ corresponding the metric topology of $S$, i.e., $X^{-1}(B) \in \mathcal{F}$ for any $B \in \mathcal{B}(S)$. Note that there are many ways to write the same things:

$$\mathsf{P}_X(A) \;\equiv\; \mathsf{P} \circ X^{-1}(A) \;\equiv\; \mathsf{P}(X \in A) \equiv\; \mathsf{P}(\{\omega \in \Omega : X(\omega) \in A\})\,.$$

Recall also that $\sigma(X) \;:=\; \{X^{-1}(B) : B \in \mathcal{B}(S)\}$. Since $\mathsf{E}$ is a linear operator we often drop the parenthesis, e.g., $\mathsf{E}u(X) \equiv \mathsf{E}[u(X)] \equiv \int_\Omega u(X(\omega))\mathsf{P}(\mathrm{d}\omega)$. We denote by $\chi_A$ the indicator function the event $A$, e.g., $\chi_A : \Omega \to \{0, 1\}$ with $\chi_A(\omega) = 1$ if and only if $\omega \in A$.

**Question 1.1:** *Let $f : S \to T$ be an arbitrary function between two sets $S$ and $T$. Recall, that $f^{-1}(B) := \{s \in S : f(s) \in B\}$, for $B \subset T$. Proof that $f^{-1}$ preserves the set operations in the sense that for any subsets $B$ and $B_k$ of $T$ the following hold: (a) $f^{-1}(B^{\mathrm{c}}) = (f^{-1}(B))^{\mathrm{c}}$; (b) $f^{-1}(\cup_k B_k) = \cup_k f^{-1}(B_k)$; (c) $f^{-1}(\cap_k B_k) = \cap_k f^{-1}(B_k)$.*

**Answer 1.1:** The function $f$ induces the map $f^{-1} : 2^T \to 2^S$, where $2^S$ means the family of all the subsets of $S$. (a) Directly from the definitions:

$$f^{-1}(B^{\mathrm{c}}) = \{s : f(s) \in B^{\mathrm{c}}\} = \{s : f(s) \notin B\} = \{s : f(s) \in B\}^{\mathrm{c}} = [f(B)]^{\mathrm{c}}\,.$$

(b) Let $I$ denote the index set that may be uncountable. Then, again directly from the definitions:

$$\begin{aligned} f^{-1}(\cup_k B_k) &= \{s : f(s) \in \cup_k B_k\} = \{s : \exists k \in I \text{ s.t. } f(s) \in B_k\} \\ &= \bigcup_k \{s : f(s) \in B_k\} = \bigcup_k f^{-1}(B_k)\,. \end{aligned}$$

(c) can be done just like (b). One can also do it by writing $\cap_k B_k = (\cup_k B_k^{\mathrm{c}})^{\mathrm{c}}$, and then using parts (a) and (b).

**Question 1.2:** *Let $\{\mathcal{G}_\alpha : \alpha \in A\}$ be a family of $\sigma$-algebras. Show that $\mathcal{G} := \cap_{\alpha\in A}\mathcal{G}_\alpha$ is $\sigma$-algebra. This result implies that $\sigma(X)$, the smallest $\sigma$-algebra, in which r.v. $X$ is measurable always exists.*

**Answer 1.2:** One has to show that $\mathcal{G}$ satisfies the three properties (i-iii) of a $\sigma$-algebra, namely that (i) $\emptyset \in \mathcal{G}$, (ii) $G \in \mathcal{G}$ implies $G^{\mathrm{c}} \in \mathcal{G}$, (iii) $(G_j)_{j\in\mathbb{N}} \subset \mathcal{G}$ implies $\cup_{j\in\mathbb{N}} G_j \in \mathcal{G}$. To see that these hold, write

$$\mathcal{G} \;=\; \{G \,:\, G \in \mathcal{G}_\alpha \text{ for every } \alpha \in A\}\,. \tag{0.1}$$

Since $\mathcal{G}_\alpha$ is a $\sigma$-algebra, $\emptyset \in \mathcal{G}_\alpha$ for every $\alpha \in A$. Now, let $G, G_j \in \mathcal{G}$ be arbitrary. By (0.1) $G^{\mathrm{c}}, \cup_j G_j \in \mathcal{G}_\alpha$ for every $\alpha$. Thus (0.1) implies that $\emptyset, G^{\mathrm{c}}, \cup_j G_j \in \mathcal{G}$ and thus we have shown that $\mathcal{G}$ is a $\sigma$-algebra.

**Question 1.3:** *Let $\mathcal{R}$ be a finite partition of $\Omega$. Let $\mathcal{G}$ consists of the empty set $\emptyset$, and all possible unions of $R \in \mathcal{R}$. Show that (a) $\mathcal{G}$ is a $\sigma$-algebra; (b) Any finite $\sigma$-algebra is of this type; (c) Let $X : \Omega \to \mathbb{R}$ be $\mathcal{G}$-measurable r.v. Prove that $X$ can take only finitely many values $\{x_j : 1 \le j \le m\} \subset \mathbb{R}$, $m \in \mathbb{N}$, and express $X$ is terms of $\chi_G$, $G \in \mathcal{G}$.*

**Answer 1.3:**

(a) We must show that properties (i-iii) of $\sigma$-algebra listed in exercise 1.2 hold. Let $\mathcal{R} = \{R_j : j = 1, \dots, n\}$, so that $|\mathcal{R}| = n$. By definition a general element $G \in \mathcal{G}$ is of the form

$G = \cup_{i \in I} R_i$, with $I \subset \{1, \ldots, n\}$ or $I = \emptyset$. Now, (i) is obviously true. To show (ii) we write $G^c = \Omega \backslash G = (\cup_{i=1}^n R) \backslash (\cup_{i \in I} R_i) = \cup_{i \in I^c} R_i$ with $I^c := \{1, \ldots, n\} \backslash I$. But this means $G^c \in \mathcal{G}$. To show (iii) we take an arbitrary sequence of sets $G_j$ and write $\cup_{j \in \mathbb{N}} G_j = \cup_{j \in \mathbb{N}} \cup_{i_j \in I_j} R_{i_j} = \cup_{i \in (\cup_j I_j)} R_i$, which is clearly an element of $\mathcal{G}$.

(b) Let $\mathcal{F}$ be a finite $\sigma$-algebra. Set

$$\mathcal{R} := \{ R \in \mathcal{F} \backslash \{\emptyset\} : \nexists F \in \mathcal{F} \backslash \{\emptyset, R\} \text{ s.t. } F \subset R \} .$$

This is the partition we are looking for. First, it is clearly finite as $\mathcal{R} \subset \mathcal{F}$. Moreover, it is obviously a partition. We have $\tilde{\Omega} := \cup \{R : R \in \mathcal{R}\} = \Omega$ since otherwise $F := \Omega \backslash \tilde{\Omega} \in \mathcal{F} \backslash \{\emptyset\}$ and thus there would have to exist $\tilde{R} \subset F$ such that $\tilde{R}$ is nonempty and not contained in any $F \in \mathcal{F}$ other than itself. But this would imply $\tilde{R} \in \mathcal{R}$ which is a contradiction. This shows that one must have $\tilde{\Omega} = \Omega$. Finally, if $R_1, R_2 \in \mathcal{R}$ are distict elements then $R_1 \cap R_2 \neq \emptyset$. For if $R := R_1 \cap R_2 \in \mathcal{F} \backslash \{\emptyset\}$ then $R \subset R_1$ and $R \subset R_2$ and consequently either $R_1$ or $R_2$ could not belong into $\mathcal{R}$ resulting in the contradiction (note: it could be that $R = R_1$ or $R = R_2$ but not both at the same time).

(c) Let $n = |\mathcal{G}| < \infty$. Write $X(\Omega) = \{x_j : 1 \leq j \leq m\}$ with $x_i \neq x_j$ for $i \neq j$. Since $X$ is measurable w.r.t. $\mathcal{G}$ the sets $R_j := X^{-1}(\{x_j\})$, $1 \leq j \leq m$ form a measurable partition $\mathcal{R} := \{R_j : 1 \leq j \leq m\} \subset \mathcal{F}$ of $\Omega$. Thus $|X(\Omega)| \equiv m \leq n$. By definition we have $X(\omega) = \sum_{j=1}^m x_j \cdot \chi_{R_j}(\omega)$.

**Question 1.4:** *Let $X, Y$ be random variables. (a) Suppose they both can take only finitely many values. Express $\sigma(X, Y)$ in terms of the elements of $\sigma(X)$ and $\sigma(Y)$. (b) Now, suppose $X$ is general, but $Y$ can take finitely many values. Write $\mathsf{E}(X|Y) := \mathsf{E}(X|\sigma(Y))$ as a deterministic function $f(Y)$ of $Y$.*

**Answer 1.4:** (a) Let $\mathcal{R} = \{R_i : i \leq m\}$ and $\mathcal{S} = \{S_j : j \leq n\}$ be the generating partitions of $\sigma$-algebras $\sigma(A)$ and $\sigma(Y)$ respectively. Then $\sigma(X, Y)$ consists of $\emptyset$ and all the possible unions of the partition $\{R_i \cap S_j : 1 \leq i \leq m, 1 \leq j \leq n\}$.

(b) By considering the partition that generates $\sigma(Y)$ it is easy to show that

$$Z = \sum_{y \in Y(\Omega)} \frac{\mathsf{E}(\chi_{\{Y=y\}} X)}{\mathsf{P}(Y = y)} 1_{\{Y=y\}} = \sum_{j=1}^n \alpha_j \chi_{B_j}(y)$$

where $Y(\Omega) = \{y_1, \ldots, y_n\}$ $B_j := \{y_j\}$ and $\alpha_j := \mathsf{E}(\chi_{\{Y=y_j\}} X) / \mathsf{P}(Y = y_j)$, is the conditional expectation $\mathsf{E}(X|Y)$. In other words, $\mathsf{E}[X|Y](\omega) = f(Y(\omega))$ with $f = \sum_{j=1}^n \alpha_j \chi_{B_j}$. Notice that $\chi_B(Y(\omega)) = 1$ if and only if $Y(\omega) \in B$ or in other words $\omega \in Y^{-1}(B)$. This gives $\chi_B(Y) = 1_{Y^{-1}(B)}$.

**Question 1.5:** *(a) Suppose $\phi : \mathbb{R} \to \mathbb{R}_+$ is increasing function. Proof that for $\mathbb{R}$-valued r.v. $X$ one has:*

$$\mathsf{P}(X \geq x) \leq \frac{\mathsf{E}\phi(X)}{\phi(x)} \qquad \textit{for every} \quad x \in \mathbb{R} .$$

*(b) Assume that $\mathsf{E}\mathrm{e}^{\alpha|X|} < \infty$, for some $\alpha > 0$. Show that $\mathsf{P}(|X| \geq r) \leq M\mathrm{e}^{-\lambda r}$, for some $0 < \lambda, M < \infty$.*

**Answer 1.5:** (a) The starting point for the proof is the basic equality $\mathsf{P}(A) = \mathsf{E}(1_A)$ valid for any $A \in \mathcal{F}$. Now, let's choose $A = \{X \geq x\}$. Since $x < y$ implies $\phi(x) < \phi(y)$, we have $A = \{\phi(X) \geq \phi(x)\}$ and therefore we may approximate:

$$\begin{aligned} 1_A(\omega) &\leq \frac{\phi(X(\omega))}{\phi(x)} 1_A(\omega) + 0 \cdot 1_{A^c}(\omega) \\ &\leq \frac{\phi(X(\omega))}{\phi(x)} 1_A(\omega) + \frac{\phi(X(\omega))}{\phi(x)} 1_{A^c}(\omega) = \frac{\phi(X(\omega))}{\phi(x)} , \end{aligned}$$

where we have used the partition of identity $1 = 1_A(\omega) + 1_{A^c}(\omega)$. Notice also that positivity of $\phi$ has been used to write $\phi(X)/\phi(x) \geq 0$. Taking expectations can comparing the left and right most expressions gives now the bound.

(b) Choose $\phi(x) := \mathrm{e}^{\lambda x}$, $\lambda \in (0, \alpha]$ and $Y := |X|$. Then (a) yields $\mathsf{P}(|X| \geq r) = \mathsf{P}(Y \geq y) \leq M\mathrm{e}^{-\alpha r}$ with $M = \mathsf{E}\mathrm{e}^{\lambda Y} \leq \mathsf{E}\mathrm{e}^{\alpha|X|}$.

**Question 1.6:** *Construct random variables $X, Y, Z$ such that the pairs $\{X, Y\}, \{X, Z\}$ and $\{Y, Z\}$ are independent, but $\{X, Y, Z\}$ are not independent.*

**Answer 1.6:** Independence means that $\mathsf{P}(X \in A, Y \in B, Z \in C) = \mathsf{P}(X \in A)\mathsf{P}(Y \in B)\mathsf{P}(Z \in C)$ for any proper measurable sets $A, B, C$. Perhaps simplest answer to this question is the following: Let $X, Y, Z : \Omega \to \{-1, 1\}$. Let $X, Y$ be identical but independent random variables such that $\mathsf{P}(X = \pm 1) = 1/2$. Now set $Z := XY$. Clearly, by symmetry $\mathsf{P}(Z = \pm 1) = 1/2$. Moreover, $\mathsf{P}(X = a, Z = b) = \mathsf{P}(X = a, XY = b) = \mathsf{P}(X = a, Y = b/a) = \mathsf{P}(X = a)\mathsf{P}(Y = b/a) = 1/4 = \mathsf{P}(X = a)\mathsf{P}(Z = b)$ for any $a, b \in \{-1, 1\}$. It then follows easily that $\mathsf{P}(X_1 \in A, X_2 \in B) = \mathsf{P}(X_1 \in A)\mathsf{P}(X_2 \in B)$ for any $\{X_1, X_2\} \subset \{X, Y, Z\}$, $X_1 \neq X_2$. Similarly, one obtains $\mathsf{P}(Y = a, Z = b) = \mathsf{P}(Y = a)\mathsf{P}(Z = b)$. However, $X, Y, Z$ are not independent as $\mathsf{P}(X = a, Y = b, Z = -ab) = 0 \neq 1/8 = \mathsf{P}(X = a)\mathsf{P}(Y = b)\mathsf{P}(Z = -ab)$ shows.

**Question 1.7:** *Suppose $\mathcal{G}$ is a sub-$\sigma$-algebra of $\mathcal{F}$, and two random variables satisfy $\mathsf{E}|XY| < \infty$. Show that (a) if $X$ is $\mathcal{G}$-measurable then $\mathsf{E}(XY|\mathcal{G}) = X\mathsf{E}(Y|\mathcal{G})$; (b) if $X$ is independent of $\mathcal{G}$ then $\mathsf{E}(X|\mathcal{G}) = \mathsf{E}X$; (c) if $\tilde{\mathcal{G}} \subset \mathcal{G}$ is another coarser sub-$\sigma$-algebra then $\mathsf{E}(X|\tilde{\mathcal{G}}) = \mathsf{E}(\mathsf{E}(X|\mathcal{G})|\tilde{\mathcal{G}})$.*
*Hint: Use simple functions.*

**Answer 1.7:** Actually the hint here is a bit misleading. Simple functions are only used in the proof of (a). Let us do (b-c) first.

(b) Let $X$ be independent of $\mathcal{G}$ so that $\mathsf{E}(1_G X) = \mathsf{E}(X)\mathsf{E}(1_G)$. But this implies $\mathsf{E}(1_G \mathsf{E}[X|\mathcal{G}]) = \mathsf{E}(1_G X) = \mathsf{E}(1_G) \cdot \mathsf{E}(X) = \mathsf{E}\{1_G \mathsf{E}(X)\}$.

(c) Take $\tilde{G} \in \tilde{\mathcal{G}} \subset \mathcal{G}$. By using the definition of conditional expectations we get:

$$\int_{\tilde{G}} \mathsf{E}[\mathsf{E}(X|\mathcal{G})|\tilde{\mathcal{G}}]\mathrm{d}\mathsf{P} = \int_{\tilde{G}} \mathsf{E}(X|\mathcal{G})\mathrm{d}\mathsf{P} = \int_{\tilde{G}} X\mathrm{d}\mathsf{P}\,.$$

(a) Let us first consider $X = 1_G$ with $G \in \mathcal{G}$. Let $H \in \mathcal{G}$. Then

$$\int_H \mathsf{E}(1_G Y|\mathcal{G})\mathrm{d}\mathsf{P} = \int_H 1_G Y\mathrm{d}\mathsf{P} = \int_{G\cap H} Y\mathrm{d}\mathsf{P} = \int_{G\cap H} \mathsf{E}(Y|\mathcal{G})\mathrm{d}\mathsf{P} = \int_H 1_G \mathsf{E}(Y|\mathcal{G})\mathrm{d}\mathsf{P}$$

Since $H$ was arbitrary and the conditional expectation is a linear operator this proves (i) for simple functions, i.e., random variables $X : \Omega \to \mathbb{R}$ of the form $X = \sum_{j=1}^k x_j 1_{A_j}$, where $A_j \in \mathcal{F}$, $x_j \in \mathbb{R}$, and $k \in \mathbb{N}$.

Extension to general $X$ and $Y$ is now done in two steps. First, we write $X = X^+ - X^-$ and $Y = Y^+ - Y^-$ where $Z^\pm = \pm\max(\pm Z, 0) \geq 0$. It is enough to prove (a) for positive random variables as can be seen by assuming such a result and writing

$$\mathsf{E}(XY|\mathcal{G}) = \sum_{a,b\in\pm} ab\mathsf{E}(X^a Y^b|\mathcal{G}) = \sum_{a,b\in\pm} abX^a\mathsf{E}(Y^b|\mathcal{G}) = X\mathsf{E}(Y|\mathcal{G})\,,$$

So we may assume that $X, Y \geq 0$. Now *Monotone Convergence theorem* (M.C.) says that if a sequence $(Z_n)_{n\in\mathbb{N}}$ of random variables satisfies $0 \leq Z_n \leq Z_{n+1}$ and there is a limit $Z = \lim_{n\to\infty} Z_n$ satisfying $\mathsf{E}Z < \infty$ then $\lim_{n\to\infty} \mathsf{E}Z_n = \mathsf{E}Z$. To use this theorem we approximate positive random variables $X$ by simple functions. This we do with the help of functions $\phi_n : [0, \infty) \to \{2^{-n}j : j = 0, 1, 2, \ldots, n2^n\}$, $n \in \mathbb{N}$, defined by $\phi_n(r) := 2^{-n}\lfloor 2^n \max(r, n)\rfloor$, where $\lfloor x\rfloor$ is the integer part of a real $x$. By setting $X_n := \phi_n(X)$ we get a sequence of simple functions

$(X_n)_{n\in\mathbb{N}}$ which satisfy $\lim_n X_n = X$ and $X_n \leq X$. Now, we can prove (a) by using monotone convergence theorem and the fact that the result holds for the simple functions $(X_n)$. Indeed, let $G \in \mathcal{G}$

$$\int_G XY \mathrm{dP} = \lim_n \int_G X_n Y \mathrm{dP} = \lim_n \int_G \mathsf{E}(X_n Y|\mathcal{G}) \mathrm{dP} = \lim_n \int_G X_n \mathsf{E}(Y|\mathcal{G}) \mathrm{dP} = \int_G X \mathsf{E}(Y|\mathcal{G}) \mathrm{dP}\,.$$

Notice that M.C. has been used twice - in the first equality and the last equality.