MA8109
Stochastic Processes in Systems Theory
Autumn 2013

Exercise set 1 – solutions

1 **a)** $x \in (A \cap B)^c \Leftrightarrow x \notin A \cap B \Leftrightarrow x \notin A$ and $x \notin B \Leftrightarrow x \in A^c$ and $x \in B^c \Leftrightarrow x \in A^c \cap B^c$

b) $x \in (\bigcup_i A_i)^c \Leftrightarrow x \notin \bigcup_i A_i \Leftrightarrow x \notin A_i$ for all $i \Leftrightarrow x \in A_i^c$ for all $i \Leftrightarrow x \in \bigcup_i A_i^c$

c) $x \in f^{-1}(A^c) \Leftrightarrow f(x) \in A^c \Leftrightarrow f(x) \notin A \Leftrightarrow x \notin f^{-1}(A) \Leftrightarrow x \in f^{-1}(A)^c$

d) $x \in f^{-1}(\bigcup_i A_i) \Leftrightarrow f(x) \in \bigcup_i A_i \Leftrightarrow f(x) \in A_i$ for all $i \Leftrightarrow x \in f^{-1}(A_i)$ for all $i \Leftrightarrow x \in \bigcup_i f^{-1}(A_i)$

2 **a)** Note that $X = A \cup B \cup (A \cup B)^c$ and this union is disjoint since $A \cap B = \emptyset$. $\mathcal{F}_{\{A,B\}} = \{\emptyset, A, B, (A \cup B)^c, A \cup B, A \cup (A \cup B)^c, B \cup (A \cup B)^c, X\}$

b) Since any subset of $\mathbb{N}$ is a countable union of elements of $\mathcal{A}$, it follows that $\mathcal{F}_{\mathcal{A}} = \mathcal{P}(\mathbb{N})$, the set of all subsets of $\mathbb{N}$.

3 It is enough to check all axioms: $\varnothing$ and Ω are in all $\mathcal{H}_i$-s , and hence in $\mathcal{H}$. If $A \in \mathcal{H}$, then A (and A^C!) is in all $\mathcal{H}_i$-s, therefore $A^C \in \mathcal{H}$. For the last axiom, we observe that if $\{A_n\}$ is in $\mathcal{H}$, then it is in all $\mathcal{H}_i$ and so will the limit A be since all $\mathcal{H}_i$-s are σ-algebras . But then $A \in \mathcal{H}$.

4 **a)** Since $A_1 \subset A_2$, $A_2 = A_1 \cup (A_2 \cap A_1^c)$, a union of disjoint measurable sets. By using the definition of a measure (σ-additivity) we may conclude the following:

$$m(A_2) = m(A_1 \cup (A_2 \cap A_1^c)) = m(A_1) + \underbrace{m(A_2 \cap A_1^c)}_{\geq 0} \geq m(A_1).$$

b) We would now like to show that $m\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} m(A_i)$. To do this we define the following disjoint sets B_i:

$$\begin{aligned} B_1 &= A_1, \\ B_2 &= A_2 \setminus A_1, \\ B_3 &= A_3 \setminus (A_1 \cup A_2), \\ &\vdots \\ B_k &= A_k \setminus \left(\bigcup_{i=1}^{k-1} A_i \right). \end{aligned}$$

We first show by induction that $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i$. Trivially, $A_1 = B_1$. Now assume $\bigcup_{i=1}^{k} A_i = \bigcup_{i=1}^{k} B_i$ for some $k \in \mathbb{N}$ and check if the same holds for $k+1$,

$$\left(\bigcup_{i=1}^{k} B_i\right) \cup B_{k+1} = \left(\bigcup_{i=1}^{k} A_i\right) \cup B_{k+1},$$

$$\bigcup_{i=1}^{k+1} B_i = \left(\bigcup_{i=1}^{k} A_i\right) \cup \left(A_{k+1} \setminus \left(\bigcup_{i=1}^{k} A_i\right)\right) = \bigcup_{i=1}^{k+1} A_i.$$

This leads to

$$m\left(\bigcup_{i=1}^{\infty} A_i\right) = m\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} m(B_i) \leq \sum_{i=1}^{\infty} m(A_i),$$

since $B_i \subset A_i$ and thus $m(B_i) \leq m(A_i)$ as shown in a).

c) Assuming $A_1 \subset A_2 \subset A_3 \subset \ldots$ and $A = \bigcup_{i=1}^{\infty} A_i$, we want to show that $m(A) = \lim_{i\to\infty} m(A_i)$. If we construct disjoint sets B_i like in b), the following holds:

$$m\left(\bigcup_{i=1}^{n} B_i\right) = m(A_n)$$

for all $n \in \mathbb{N}$. By using the definition of A and the fact that $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i$ we get:

$$\begin{aligned} m(A) &= m\left(\bigcup_{i=1}^{\infty} A_i\right) = m\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} m(B_i) \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} m(B_i) = \lim_{n\to\infty} m\left(\bigcup_{i=1}^{n} B_i\right) = \lim_{n\to\infty} m(A_n). \end{aligned}$$

5 Let m_L be the Lebesgue measure on $\mathbb{R}$ and let $\epsilon > 0$. Let $\{q_n\}_{n\in\mathbb{N}}$ be a sequence of the elements in $\mathbb{Q}$. Define the open intervals $A_n = (q_n - 2^{-n}\epsilon, q_n + 2^{-n}\epsilon)$. Notice that $\mathbb{Q} \subset \cup_{n\in\mathbb{N}} A_n$, thus

$$m_L(\mathbb{Q}) \leq m_L(\cup_{n\in\mathbb{N}} A_n) \leq \sum_{n\in\mathbb{N}} m_L(A_n) = \sum_{n\in\mathbb{N}} \frac{\epsilon}{2^n} = \epsilon.$$

Since ϵ is chosen arbitrarily small, we have $m_L(\mathbb{Q}) = 0$.

6 Assuming f_1, f_2 are $\mathcal{F}$-measurable, we know

$$f_1^{-1}((a,\infty)) \in \mathcal{F} \text{ and } f_2^{-1}((a,\infty)) \in \mathcal{F},\ \forall\, a \in \mathbb{R}.$$

From this, we find that

$$\begin{aligned} f^{-1}((a,\infty)) =& \{x \in X : f(x) \in (a,\infty)\} \\ =& \{x \in X : \max(f_1(x), f_2(x)) \in (a,\infty)) \\ =& \{x \in X : f_1(x) \in (a,\infty) \vee f_2(x) \in (a,\infty)\} \\ =& f_1^{-1}((a,\infty)) \cup f_2^{-1}((a,\infty)) \in \mathcal{F}\, \forall\, a \in \mathbb{R}, \end{aligned}$$

since unions of measurable sets are measurable ($\mathcal{F}$ is a σ-algebra). Since the collection of sets (a,∞), $a \in \mathbb{R}$, generate the Borel σ-algebra $\mathcal{B}$, it follows that f is $(\mathcal{F},\mathcal{B})$-measurable, i.e. $\mathcal{F}$-measurable.

7 1. We first prove the result for simple functions. So let $s_1 = \sum_{k=1}^n a_k \chi_{A_k}$ and $s_2 = \sum_{l=1}^m b_l \chi_{B_l}$, where $\{A_i\}_i$ and $\{B_i\}_i$ are partitions of the underlying space.[1] We want to prove that $\int (s_1 + s_2) = \int s_1 + \int s_2$.

Since $s_1 + s_2 = \sum_{k=1}^n \sum_{l=1}^m (a_k + b_l)\chi_{A_k \cap B_l}$ we see that $s_1 + s_2$ is also a simple function, and thus

$$\begin{aligned}
\int (s_1 + s_2) &= \sum_{k=1}^n \sum_{l=1}^m (a_k + b_l) m(A_k \cap B_l) \\
&= \sum_{k=1}^n a_k \sum_{l=1}^m m(A_k \cap B_l) + \sum_{l=1}^m b_l \sum_{k=1}^n m(A_k \cap B_l) \\
&= \sum_{k=1}^n a_k m(A_k) + \sum_{l=1}^m b_l m(B_l) \\
&= \int s_1 + \int s_2.
\end{aligned}$$

2. We prove the result for measurable functions $f_1, f_2 : X \to [0, \infty]$. Pick two increasing sequences of nonnegative simple functions $\{s_n^1\}_n$ and $\{s_n^2\}_n$ converging pointwise to f_1 and f_2, respectively. Then $s_n^1 + s_n^2$ is a nonnegative sequence of increasing simple functions converging to $f_1 + f_2$. Thus we get, by the Monotone Convergence Theorem and the corresponding result for simple functions, that

$$\int (f_1 + f_2) = \lim_{n\to\infty} \int (s_n^1 + s_n^2) = \lim_{n\to\infty} \left(\int s_n^1 + \int s_n^2 \right) = \int f_1 + \int f_2.$$

8 We first note that $\mu(A) = \int_A \phi \, dm$ is well-defined and non-negative for all $A \in \mathcal{F}$ since ϕ is non-negative and $\mathcal{F}$-measurable.

Next we check that μ is σ-additive. If $A = \cup_{i=1}^\infty A_i$ with $A_i \cap A_j = \emptyset$ $(i \neq j)$, then

$$\mu(\cup_{i=1}^\infty A_i) = \int_{\cup_{i=1}^\infty A_i} \phi \, dm = \int_X \sum_{i=1}^\infty \phi \chi_{A_i} \, dm. \tag{1}$$

Define $f_n = \sum_{i=1}^n \phi \chi_{A_i}$. Obviously, $\{f_n\}_{n\in\mathbb{N}}$ is a non-decreasing sequence of non-negative functions. Hence by (1) and the monotone convergence theorem,

$$\mu(\cup_{i=1}^\infty A_i) = \int_X \lim_{n\to\infty} f_n \, dm = \lim_{n\to\infty} \int_X f_n \, dm.$$

[1] In the note *Measure and Probability*, a simple function is defined to be a linear combination of characteristic functions of disjoint sets $\{A_i\}_i$. We can demand that $\{A_i\}_i$ is a partition of X – meaning that in addition to disjointness we have that $\cup A_i = X$ – since we can always add $0 \cdot \chi_{s^{-1}(\{0\})}$ as a term without changing the function.

Now σ-additivity follows from linearity of the integral since

$$\begin{aligned}\lim_{n\to\infty}\int_X f_n\,dm &= \lim_{n\to\infty}\int_X \sum_{i=1}^{n}\phi\chi_{A_i}\,dm\\ &= \lim_{n\to\infty}\sum_{i=1}^{n}\int_X \phi\chi_{A_i}\,dm\\ &= \sum_{i=1}^{\infty}\int_{A_i}\phi\,dm\\ &= \sum_{i=1}^{\infty}\mu(A_i).\end{aligned}$$

Since $\mu(\emptyset)=0$, we then conclude that $\mu(A)$ is a measure on $(X,\mathcal{F})$.

9 **a)** We have $\mu_X(\varnothing)=P(X^{-1}(\varnothing))=P(\varnothing)=0$ and also $\mu_X(A)=P(X^{-1}(A))\geq 0$ for every $A\in\mathcal{B}_{\mathbb{R}}$ since P is a measure. Lastly, if $\{A_i\}_{i=1}^{\infty}\subseteq\mathcal{F}$ are pairwise disjoint, then

$$f^{-1}(A_i)\cap f^{-1}(A_j)=f^{-1}(A_i\cap A_j)=f^{-1}(\emptyset)=\emptyset \quad\text{for}\quad i\neq j,$$

and $\{f^{-1}(A_i)\}_{i=1}^{\infty}$ is also pairwise disjoint. It follows that

$$\begin{aligned}\mu_X\left(\bigsqcup_{i=1}^{\infty}A_i\right) &= P\left(X^{-1}\left(\bigsqcup_{i=1}^{\infty}A_i\right)\right)\\ &= P\left(\bigsqcup_{i=1}^{\infty}X^{-1}(A_i)\right)\\ &= \sum_{i=1}^{\infty}P(X^{-1}(A_i))\\ &= \sum_{i=1}^{\infty}\mu_X(A_i),\end{aligned}$$

again since P is a measure. So μ_X is a measure on $(\mathbb{R},\mathcal{B}_{\mathbb{R}})$. It is is a probability measure since $\mu_X(\mathbb{R})=P(X^{-1}(\mathbb{R}))=P(\Omega)=1$.

b) First suppose that $f=\chi_A$ for some $A\in\mathcal{B}_{\mathbb{R}}$. We deduce that $\chi_A\circ X=\chi_{X^{-1}(A)}$, which yields

$$\begin{aligned}E(\chi_A(X)) &= \int_\Omega \chi_A(X(\omega))dP(\omega)\\ &= \int_\Omega \chi_{X^{-1}(A)}(\omega)dP(\omega)\\ &= P(X^{-1}(A))\\ &= \mu_X(A)\\ &= \int_{\mathbb{R}}\chi_A(x)d\mu_X(x).\end{aligned}$$

By the linearity of the integral, this will also hold for all simple functions. Now take any $\mathcal{B}_{\mathbb{R}}$-measurable $f:\Omega\to[0,\infty]$. Then we can find a sequence (s_n) of simple functions such that $0 \le s_n \uparrow f$ (pointwise everywhere!). Clearly this also means that $0 \le s_n \circ X \uparrow f \circ X$. By the above we have

$$E(s_n(X)) = \int_\Omega s_n(X(\omega))dP(\omega) = \int_\mathbb{R} s_n(x)d\mu_X(x)$$

for all $n \in \mathbb{N}$. Letting $n \to \infty$ and using the monotone convergence theorem, we obtain

$$E(f(X)) = \int_\Omega f(X(\omega))dP(\omega) = \int_\mathbb{R} f(x)d\mu_X(x),$$

which is what we wanted to show.

10 **(a)** is simple: Since f_n is 0 on the interval $[2\omega_n, 1]$ and $\omega_n \to 0$, the variable $\omega > 0$ will not be in $[2\omega_n, 1]$ when n is large enough. Thus the limit is 0 for all ω-s. Clearly, since $g(\omega) = \lim_{n\to\infty} f_n(\omega) = 0$, also $\int g dP = 0$.

For **(b)** we observe that $\int f_n dP$ is the same as the familiar integral $\int_0^1 f_n(\omega)\,d\omega$. For $a_n = \omega_n^\alpha$ we obtain

$$\int_0^1 f_n(\omega)\,d\omega = \omega_n^\alpha \cdot \frac{(2\omega_n)\cdot 1}{2} = \omega_n^{\alpha+1}, \tag{2}$$

showing that the limit when $n \to \infty$ may be 0, 1, or ∞ depending on the value of α.

In **(c)** the function h is a dominating function for all f_n-s and in fact the smallest such function. If $\int h dP < \infty$, h is integrable, and in that case (recalling the Lebesgue Dominated Convergence Theorem) we *must* have

$$\int \left(\lim_{n\to\infty} f_n\right) dP = \lim_{n\to\infty} \int f_n dP. \tag{3}$$

Thus, we conclude that $\int h dP$ has to be ∞ when $a_n = \omega_n^{-1}$, and ω_n^{-2}.

Comment: For the first case, $a_n = \omega_n^{-1/2}$, it is tempting to say that

$$h(\omega) = \max_n f_n(\omega) \le \omega^{-1/2}. \tag{4}$$

But this function is integrable, and $\int_0^1 \omega^{-1/2} d\omega = 2$. Therefore, Eqn. 4 must be wrong (why?). *Challenge:* Prove that an integrable majorant nevertheless exists by finding a larger function, g, such that $h(\omega) < g(\omega)$ and $\int g dP < \infty$.

It could also be remarked that it is possible to find more complicated examples, with similar functions, such that $\int (\lim_{n\to\infty} f_n)\, dP = \lim_{n\to\infty} \int f_n dP$ even if there is no integrable majorant.

11 **(a)** This, so-called *monotone property* of conditional expectation, follows easily by utilizing the linearity and that $X \ge 0$ a.s. implies that $E(X|\mathcal{H}) \ge 0$ a.s. (Property 5):

$$Y - X \ge 0 \Longrightarrow E(Y - X|\mathcal{H}) = E(Y|\mathcal{H}) - E(X|\mathcal{H}) \ge 0. \tag{5}$$

(b) From point (a) it follows that also $\{E(X_n|\mathcal{H})\}$ will be an increasing sequence,

$$E(X_n|\mathcal{H})(\omega) \le E(X_{n+1}|\mathcal{H})(\omega) \text{ a.s.} \tag{6}$$

Apart from a possible set of probability 0, we then know that

$$E\left(X_n|\mathcal{H}\right)(\omega) \underset{n\to\infty}{\to} Y(\omega) \leq \infty \tag{7}$$

for some non-negative function Y. First of all, the Y function will be $\mathcal{H}$-measurable since it is a limit of $\mathcal{H}$-measurable functions (this was mentioned briefly in the lectures and is a general result from measure theory). Moreover, by applying the Monotone Convergence Theorem to *both* sides of

$$\int_H E\left(X_n|\mathcal{H}\right) dP = \int_H X_n dP,\ H \in \mathcal{H}, \tag{8}$$

we have

$$\int_H Y dP = \int_H X dP \tag{9}$$

for all $H \in \mathcal{H}$. Thus, $Y = E\left(X|\mathcal{H}\right)$ a.s.

12 **a)** The RHS is the familiar (for some) probability density for a multivariate gaussian variable $\mathbf{X}$ as long as the covariance matrix is non-singular.

Let $\mathbf{y} = \mathbf{\Sigma}^{1/2}\mathbf{u}$. The square root exists since $\mathbf{\Sigma}$ is positive definite. Then $d^n y = \left|\mathbf{\Sigma}^{1/2}\right| d^n u$, where $\left|\mathbf{\Sigma}^{1/2}\right|$, the determinant of $\mathbf{\Sigma}^{1/2}$, is the *Jacobian* of the transformation. Note also that $\mathbf{u}'\mu = \mathbf{y}'\mathbf{\Sigma}^{-1/2}\mu$. By introducing this, we obtain

$$\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{-i\mathbf{u}'\mathbf{x}} \exp\left(i\mathbf{u}'\mu - \frac{1}{2}\mathbf{u}'\mathbf{\Sigma}\mathbf{u}\right) d^n u \tag{10}$$

$$= \frac{1}{(2\pi)^n}\frac{1}{\left|\mathbf{\Sigma}^{1/2}\right|}\int_{\mathbb{R}^n} \exp\left(i\mathbf{y}'\mathbf{\Sigma}^{-1/2}\left(\mu - \mathbf{x}\right) - \frac{1}{2}\mathbf{y}'\mathbf{y}\right) d^n y \tag{11}$$

$$= \frac{1}{(2\pi)^n}\frac{1}{\left|\mathbf{\Sigma}^{1/2}\right|}\int_{\mathbb{R}^n} \exp\left(i\mathbf{y}'\mathbf{a} - \frac{1}{2}\mathbf{y}'\mathbf{y}\right) d^n y, \quad \mathbf{a} = \mathbf{\Sigma}^{-1/2}\left(\mu - \mathbf{x}\right). \tag{12}$$

The integral now splits into a product of n one-dimensional integrals of the form $\int_{\mathbb{R}} \exp\left(iya - \frac{1}{2}y^2\right) dy$, which can be found by observing that

$$\int_{\mathbb{R}} \exp\left(iya - \frac{1}{2}y^2\right) dy \tag{13}$$

$$= e^{-\frac{a^2}{2}}\int_{-\infty}^{\infty} \exp\left(-\frac{(ia-y)^2}{2}\right) dy \tag{14}$$

$$\overset{s=(y-ia)}{=} e^{-\frac{a^2}{2}}\int_{-\infty}^{\infty} e^{-s^2/2} ds = \sqrt{2\pi}e^{-\frac{a^2}{2}} \tag{15}$$

(We are integrating an analytic function in the complex plane along a path parallel to the y-axis. The contributions from the connecting paths at both ends vanish. Therefore,

$$\int_{-\infty}^{\infty} \exp\left(-\left(ia-y\right)^2/2\right) dy = \int_{-\infty}^{\infty} \exp\left(-y^2/2\right) dy = \sqrt{2\pi} \tag{16}$$

for all real a-s).

Finally,

$$\frac{|\mathbf{a}|^2}{2} = \frac{1}{2}\sum_{k=1}^{n}\left(\mathbf{\Sigma}^{-1/2}(\mu - \mathbf{x})\right)_k^2 = \frac{1}{2}(\mathbf{x}-\mu)'\,\mathbf{\Sigma}^{-1}(\mathbf{x}-\mu)\,, \tag{17}$$

which is just what we need.

(b) *Proof of (i)*: This is rather surprising since it holds *regardless* the variables are independent or not! The proof is simple if we use the series expansion for the characteristic function,

$$\exp\left(-\frac{1}{2}\mathbf{u}'\mathbf{\Sigma}\mathbf{u}\right) = 1 - \frac{1}{2}\mathbf{u}'\mathbf{\Sigma}\mathbf{u} + \frac{1}{2}\left(\frac{1}{2}\mathbf{u}'\mathbf{\Sigma}\mathbf{u}\right)^2 + \cdots\,, \tag{18}$$

and the formula

$$E\left(X_1X_2X_3\right) = i^3\frac{\partial^3\phi}{\partial u_1\partial u_2\partial u_3}(\mathbf{0})\,. \tag{19}$$

Observe that the third term in Eqn. 18 contains a product of 4 u-components and all later terms contain even more than that. At the end we are going to put all u-components equal to zero. After three derivations, the third term and all later terms will either already be 0, or have remaining u-components and will be 0 for $\mathbf{u} = \mathbf{0}$. The derivatives of the second term are all of the form

$$\sigma_{ij}\frac{\partial^3\left(u_iu_j\right)}{\partial u_1\partial u_2\partial u_3}, \tag{20}$$

and are all equal to 0 as well.

Proof of (ii):

In this case, we need to take four derivatives,

$$E\left(X_1X_2X_3X_4\right) = \frac{\partial^4\phi}{\partial u_1\partial u_2\partial u_3\partial u_4}(\mathbf{0})\,, \tag{21}$$

and all terms in the expansion in Eqn. 18 will vanish, apart from some of the derivatives of the third term. The third term may be written as

$$\frac{1}{8}\sum_{i,j,k,l=1}^{4}\sigma_{ij}\sigma_{kl}u_iu_ju_ku_l, \tag{22}$$

and we see that

$$\frac{\partial^4\left(u_iu_ju_ku_l\right)}{\partial u_1\partial u_2\partial u_3\partial u_4} \tag{23}$$

will be non-zero *only* if all four components in the numerator are different. Since $\sigma_{ij} = \sigma_{ji}$, there will be 8 such terms for $\sigma_{12}\sigma_{34}$, and likewise, 8 terms for $\sigma_{13}\sigma_{24}$ and $\sigma_{14}\sigma_{23}$. This makes up the identity.

For info: The *cumulants* are coefficients in the Taylor expansion of $\log(\phi(\mathbf{u}))$. The main k-th order cumulant is the coefficient in front of the $\frac{i^k}{k!}u_1 \cdot u_2 \cdots u_n$ -term of the multidimensional Taylor expansion. Since the Taylor expansion of $\log(\phi(\mathbf{u}))$ for multivariate Gaussian variables is simply $\log(\phi(\mathbf{u})) = i\mathbf{u}'\mu - \frac{1}{2}\mathbf{u}'\mathbf{\Sigma}\mathbf{u}$, all cumulants larger that the second vanish. The fourth cumulant is in general

$$\kappa_4 = E(X_1X_2X_3X_4) - E(X_1X_2)E(X_3X_4) - E(X_1X_3)E(X_2X_4) - E(X_1X_4)E(X_2X_3), \tag{24}$$

hence the name of the identity. The identity is useful in all situations involving Gaussian signals.