Solution for Note 1 of 275A Probability Theory

shuyu qin

August 2023

1 Selected Exercises in Note 1

Exercise 3

Proof. We mimic the proof of Lemma 1.3.4 (Well-definedness of simple integral) of "M" and use a Venn diagram argument. The $n+m$ sets $E_1, \dots, E_n, F_1, \dots, F_m$ divide the space Ω into 2^{n+m} disjoint sets, each of which is the intersection of some of the $E_1, \dots, E_n, F_1, \dots, F_m$ and their complements. We throw away any sets that are empty, leaving us with a partition of Ω into k non-empty disjoint sets $A_1, \dots, A_k$ for some $0 \le k \le 2^{n+m}$. As the $E_1, \dots, E_n, F_1, \dots, F_m$ are measurable, the $A_1, \dots, A_k$ are too. By construction, each of the $E_1, \dots, E_n, F_1, \dots, F_m$ arise as unions of some of the $A_1, \dots, A_k$, thus we can write:

$E_i = \bigcup_{j \in J_i} A_j$ and $F_{i'} = \bigcup_{j \in J'_{i'}} A_j$ for all $i = 1, \dots, n$ and $i' = 1, \dots, m$, and some subsets $J_i, J'_{i'} \subset \{1, \dots, k\}$.

By the finite additivity of μ, we thus have

$$\mu(E_i) = \sum_{j \in J_i} \mu(A_j) \text{ and } \mu(F_{i'}) = \sum_{j \in J'_{i'}} \mu(A_j).$$

Thus, our objective is now to show that

$\sum_{i=1}^n a_i \sum_{j \in J_i} \mu(A_j) = \sum_{i'=1}^m b_{i'} \sum_{j \in J'_{i'}} \mu(A_j)$.

To obtain this, we fix $1 \le j \le k$ and evaluate $f = \displaystyle\sum_{i=1}^n a_i 1_{E_i} = \sum_{i'=1}^m b_{i'} 1_{F_{i'}}$ at a point x in the non-empty set A_j. At such a point, $1_{E_i}(x) = 1_{J_i}(j)$, and similarly $1_{F_{i'}}(x) = 1_{J'_{i'}}(j)$. From condition we conclude that

$\sum_{i=1}^n a_i 1_{J_i}(j) = \sum_{i'=1}^m b_{i'} 1_{J'_{i'}}(j)$.

Multiplying the LHS by $\mu(A_j)$ and summing over all $j = 1, \dots, k$, we get

$\sum_{j=1}^k \mu(A_j) \sum_{i=1}^n a_i 1_{J_i}(j) = \sum_{i=1}^n \sum_{j=1}^k \mu(A_j) a_i 1_{J_i}(j) = \sum_{i=1}^n a_i \sum_{j \in J_i} \mu(A_j)$

$= \sum_{i=1}^{n} a_i \mu(E_i)$. Similarly, multiplying the RHS by $\mu(A_j)$ and summing over all $j = 1, \ldots, k$, we get $\sum_{i'=1}^{m} b_{i'} \mu(F_{i'})$. So $\sum_{i=1}^{n} a_i \mu(E_i) = \sum_{i'=1}^{m} b_{i'} \mu(F_{i'})$. $\square$

Exercise 4

Proof. (1). See Exercise 1.3.1 of "M". Suppose that $f = \sum_{i=1}^{n} a_i 1_{E_i}$ and $g = \sum_{j=1}^{m} b_j 1_{F_j}$ are two decomposition of f and g, where the E_i's and F_j's are all measurable, and the a_i's and b_j's all lie in $[0, +\infty]$. By definition, we have

Simp $\int_\Omega f + g \, d\mu =$ Simp $\int_\Omega \sum_{t=1}^{n+m} c_t 1_{G_t} \, d\mu := \sum_{t=1}^{n+m} c_t \mu(G_t)$

$= \sum_{i=1}^{n} a_i \mu(E_i) + \sum_{j=1}^{m} b_j \mu(F_j) =$ Simp $\int_\Omega f \, d\mu +$ Simp $\int_\Omega g \, d\mu$.

Where $c_t = a_t$ and $G_t = E_t$ for $t = 1, \ldots, n$, $c_{n+t} = b_t$ and $G_{n+t} = F_t$ for $t = 1, \ldots, m$.

Likewise, we have

Simp $\int_\Omega cf \, d\mu := \sum_{i=1}^{n} c a_i \mu(E_i) = c \sum_{i=1}^{n} a_i \mu(E_i) = c$ Simp $\int_\Omega f \, d\mu$. $\square$

(2). As above, let f admits a decomposition $f = \sum_{i=1}^{n} a_i 1_{E_i}$. By definition of the simple integral, the values of the integrand on a null set contribute nothing to the value of the integral, and if we let E be the null set on which f differs from g, then we can join E to one of the E_i's and assume for the purpose of integration that g also admits the same decomposition $\sum_{i=1}^{n} a_i 1_{E_i}$, which clearly gives Simp $\int_\Omega f \, d\mu =$ Simp $\int_\Omega g \, d\mu$. $\square$

(3). By definition, the simple integral of f equals $\sum_{i=1}^{n} a_i \mu(E_i)$ for a_i's lie in $[0, +\infty]$ and $\mu(E_i) \geq 0$, thus Simp $\int_\Omega f \, d\mu \geq 0$. If Simp $\int_\Omega f \, d\mu = 0$, then in particular we have $a_i = 0$ for any i such that E_i is not a null set (otherwise the product $a_i \mu(E_i) > 0$), giving $f = 0$ almost everywhere. Conversely, if $f = 0$ almost everywhere, then again we get $a_i = 0$ for any i such that $\mu(E_i) > 0$ (otherwise $f > 0$ on a set of positive measure), giving Simp $\int_\Omega f \, d\mu = 0$. $\square$

(4). The main obstacle here is that f and g may not admit a decomposition on the same collection of sets, but it can be solved by using a common refinement.

Let f be specified by the sets $E_1, \ldots, E_n$ and g by $F_1, \ldots, F_m$, with the E_i's and F_j's being measurable sets. Then as in the proof of Exercise 3, by taking the intersection of the $n + m$ sets $E_1, \ldots, E_n$, $F_1, \ldots, F_m$ and their complements, throwing away empty sets, we get a common refinement $A_1, \ldots, A_k$ for some

$1 \leq k \leq 2^{n+m}$. In particular, as each E_i and each F_i is the union of some of the A_j's, we can write $f = \sum_{i=1}^{k} a_i 1_{A_i}$ and $g = \sum_{i=1}^{k} b_i 1_{A_i}$, with $a_i \leq b_i$ for all i. Then Simp $\int_\Omega f \; d\mu = \sum_{i=1}^{k} a_i \mu(A_i) \leq \sum_{i=1}^{k} b_i \mu(A_i) =$ Simp $\int_\Omega g \; d\mu$, as desired. $\square$

(5). The proof is largely the same as Lemma 1.3.15 of "M". $\forall 0 < t < \infty$, we have the trivial point-wise inequality $t 1_{\{\omega \in \Omega : f(\omega) \geq t\}} \leq f(\omega)$. Since the set $\{\omega \in \Omega : f(\omega) \geq t\} = f^{-1}([t, \infty))$ is measurable, the function $t 1_{\{\omega \in \Omega : f(\omega) \geq t\}}$ is an unsigned simple function bounded point-wise above by f. Hence by monotonicity we get

Simp $\int_\Omega t 1_{\{\omega \in \Omega : f(\omega) \geq t\}} \; d\mu = t\mu(\{\omega \in \Omega : f(\omega) \geq t\}) \leq$ Simp $\int_\Omega f \; d\mu$.

Or equivalently, $\mu(\{\omega \in \Omega : f(\omega) \geq t\}) \leq \frac{1}{t}$Simp $\int_\Omega f \; d\mu$. $\square$

Exercise 5

Proof. (1). See Exercise 1.3.10 of "M". For any unsigned simple functions $h_1 \leq f$ and $h_2 \leq g$, the function $h_1 + h_2$ is an unsigned simple function such that $h_1 + h_2 \leq f + g$. Hence we have

Simp $\int_\Omega h_1 \; d\mu +$ Simp $\int_\Omega h_2 \; d\mu =$ Simp $\int_\Omega h_1 + h_2 \; d\mu \leq \int_\Omega f + g \; d\mu$

Taking the supremum on the LHS of the above inequality over all $0 \leq h_1 \leq f$, followed by taking the supremum on the LHS of the above inequality over all $0 \leq h_2 \leq g$ gives the desired result.

Now let $c \in (0, +\infty)$. For any unsigned simple $h \leq f$, we have

c Simp $\int_\Omega h \; d\mu =$ Simp $\int_\Omega ch \; d\mu \leq \int_\Omega cf \; d\mu$.

Taking the supremum on the LHS of the above inequality over all $h \leq f$ then gives $c \int_\Omega f \; d\mu \leq \int_\Omega cf \; d\mu$.

Similarly, for any unsigned simple $h \leq cf$, we have

Simp $\int_\Omega h \; d\mu =$ Simp $\int_\Omega c \cdot \frac{h}{c} \; d\mu = c$ Simp $\int_\Omega \frac{h}{c} \; d\mu \leq c \int_\Omega f \; d\mu$.

Taking the supremum on the LHS of the above inequality over all $h \leq cf$ gives $\int_\Omega cf \; d\mu \leq c \int_\Omega f \; d\mu$. Hence $\int_\Omega cf \; d\mu = c \int_\Omega f \; d\mu$. The case when $c = 0$ or $c = +\infty$ is obvious from definition. $\square$

(2). For any unsigned simple $h \leq f$, $h \leq g$ almost everywhere. By adjusting h on a null set we see that Simp $\int_\Omega h\ d\mu \leq \int_\Omega g\ d\mu$. Taking the supremum on the LHS of this inequality over all $h \leq f$ gives $\int_\Omega f\ d\mu \leq \int_\Omega g\ d\mu$. Since the roles of f and g in this proof are symmetric, we also get $\int_\Omega g\ d\mu \leq \int_\Omega f\ d\mu$ and thus $\int_\Omega g\ d\mu = \int_\Omega f\ d\mu$. $\square$

(3). That $\int_\Omega f\ d\mu \geq 0$ is obvious from the definition of unsigned integral. Furthermore, if $\int_\Omega f\ d\mu = 0$ and $f > 0$ on a set E of positive measure, we can find some $c > 0$ such that $f \geq c > 0$ on E. Then $\int_\Omega f\ d\mu \geq c\mu(E) > 0$, a contradiction. So $f = 0$ almost everywhere. Conversely, if $f = 0$ almost everywhere, then any simple function minorizing f must be zero almost everywhere, giving $\int_\Omega f\ d\mu = 0$. $\square$

(4). For any unsigned simple $h \leq f$, $h \leq g$ almost everywhere. Adjusting h on a null set to make it smaller than both f and g everywhere, we get by definition that Simp $\int_\Omega h\ d\mu \leq \int_\Omega g\ d\mu$. Taking the supremum on the LHS of the above inequality over all $h \leq f$, we get $\int_\Omega f\ d\mu \leq \int_\Omega g\ d\mu$. $\square$

(5). This is basically the same as the proof for part (5) of Exercise 4, using monotonicity for unsigned functions. By taking $t \to +\infty$, the second conclusion also follows. $\square$

(6). If f is simple, then $\int_\Omega f\ d\mu := \sup_{0 \leq h \leq f, h \text{ simple}} \text{Simp}\int_\Omega h\ d\mu = \text{Simp}\int_\Omega f\ d\mu$.$\square$

(7). Again, this follows directly from the definition of unsigned simple integral.$\square$

Exercise 6

Proof. See Exercise 1.4.40 of "M". Suppose first that $\sum_{\omega\in\Omega} f(\omega)p_\omega < \infty$. By the Archemedian principle, the set $A := \{\omega \in \Omega : f(\omega)p_\omega > 0\}$ is at most finite.

Pick some $c > 0$. Define the unsigned simple function $g : \Omega \to [0, +\infty]$ by setting $g := \sum_{\omega\in A, f(\omega)<+\infty} f(\omega)1_{\{\omega\}} + \sum_{\omega\in A, f(\omega)=+\infty} c1_{\{\omega\}}$.

Then we see that $g \leq f$, and for any other unsigned simple function $g' \leq f$, we have Simp $\int_\Omega g'\ d\mu \leq$ Simp $\int_\Omega g\mu = \sum_{\omega\in\Omega} f(\omega)p_\omega$. So $\int_\Omega f\ d\mu = \sum_{\omega\in\Omega} f(\omega)p_\omega$.

Now if $\sum_{\omega\in\Omega} f(\omega)p_\omega = \infty$. Then either f attains $+\infty$ on some points in Ω or that $f > 0$ on countably many points in Ω (Or both). In the former case, we can choose an unsigned simple $g \leq f$ by letting $g := +\infty$ on some finite set in Ω on which $f = +\infty$ and $g := 0$ elsewhere, Then clearly

Simp $\int_\Omega g\ d\mu = \infty = \int_\Omega f\ d\mu = \sum_{\omega\in\Omega} f(\omega)p_\omega$.

In the latter case, $\forall M > 0$, we can choose a simple $g \leq f$ such that

Simp $\int_\Omega g \ d\mu > M$. By including as finitely many points in the domain of f as needed, suggesting again that $\int_\Omega f \ d\mu = \sum_{\omega\in\Omega} f(\omega)p_\omega$. $\square$

Exercise 7

Proof. (1). See also Exercise 1.3.11 and Theorem 1.4.38 of "M". Let $E \subset \Omega$ be the finite measure support of f. Let $\overline{\int_\Omega} f \ d\mu := \inf_{h\geq f, h \text{ simple}}$ Simp $\int_\Omega h \ d\mu$ and $\underline{\int_\Omega} f \ d\mu := \sup_{0\leq g\leq f, g \text{ simple}}$ Simp $\int_\Omega g \ d\mu$ denote the upper and lower unsigned integral, respectively. We want to show that $\underline{\int_\Omega} f \ d\mu = \overline{\int_\Omega} f \ d\mu$.

Pick an $\varepsilon > 0$. Let f^ε be f rounded up to the nearest integer multiple of ε, and f_ε be f rounded down to the nearest integer multiple of ε. Clearly, we have the point-wise bounds

$f_\varepsilon(\omega) \leq f(\omega) \leq f^\varepsilon(\omega)$, and $f^\varepsilon(\omega) - f_\varepsilon(\omega) \leq \varepsilon$.

Since f is bounded, f_ε and f^ε are simple. Note that if f_ε attains $k\varepsilon$ for some integer k at $\omega \in \Omega$, f^ε attains either $k\varepsilon$ or $(k+1)\varepsilon$. Conversely, if f^ε attains $k\varepsilon$ at ω, f_ε attains either $k\varepsilon$ or $(k-1)\varepsilon$. Thus, f^ε and f_ε admit a decomposition on the same collection of set $E_1, \ldots, E_n$ (the pre-images of all the distinct multiples $k\varepsilon$).

Then by definition we have

$|\overline{\int_\Omega} f \ d\mu - \underline{\int_\Omega} f \ d\mu| \leq$ Simp $\int_\Omega f^\varepsilon \ d\mu -$ Simp $\int_\Omega f_\varepsilon \ d\mu \leq \varepsilon\mu(E)$.

Sending $\varepsilon \to 0$ gives the desired result. $\square$

(2). Define f^ε, f_ε, g^ε, g_ε as in part (1). In view of superadditivity, it suffices to establish the subadditivity property

$\int_\Omega (f+g) \ d\mu \leq \int_\Omega f \ d\mu + \int_\Omega g \ d\mu$.

By construction, we have the point-wise bound

$f + g \leq f^\varepsilon + g^\varepsilon \leq f^\varepsilon + g_\varepsilon + 2\varepsilon$,

hence by monotonicity along with the linearity of the simple integral,

$\int_\Omega f + g \ d\mu \leq \int_\Omega f^\varepsilon + g_\varepsilon + 2\varepsilon \ d\mu =$ Simp $\int_\Omega f^\varepsilon \ d\mu +$ Simp $\int_\Omega g_\varepsilon \ d\mu + 2\varepsilon\mu(E)$, suggesting that $\int_\Omega f+g\ d\mu \leq \int_\Omega f\ d\mu + \int_\Omega g\ d\mu + 2\varepsilon\mu(E)$. Sending $\varepsilon \to 0$ and using the assumption that $\mu(E) < \infty$, we obtain the claim. $\square$

(3). (Horizontal truncation) See Exercise 1.4.36 of "M". First assume that $\int_\Omega f\ d\mu < \infty$, then f is finite almost everywhere. In particular, the functions $f_n = \min(f, n)$ agree with f almost everywhere for sufficiently large n. Thus we get $\lim_{n\to\infty} \int_\Omega \min(f,n)\ d\mu = \int_\Omega f\ d\mu$.

Next assume that $\int_\Omega f\ d\mu = \infty$. Then either $f = +\infty$ on a set E of positive measure, or the set $\{f > 0\}$ has infinite measure.

In the former case, pick $M > 0$. If we let g be the simple function define by $g := M1_E$, then $\int_\Omega \min(f, M)\ d\mu \geq \text{Simp} \int_\Omega g\ d\mu = M\mu(E)$. Sending $M \to \infty$, and using monotonicity of the unsigned integral, we get $\lim_{M\to\infty} \int_\Omega \min(f, M)\ d\mu = \infty$.

In the latter case, we can further assume that $f < \infty$ almost everywhere. Then again for sufficiently large M, we have by equivalency that $\int_\Omega \min(f, M)\ d\mu = \int_\Omega f\ d\mu = \infty$. That is, $\lim_{M\to\infty} \int_\Omega \min(f, M)\ d\mu = \infty$, as desired. $\square$

(4). By the bounded case, for any natural number n, we have

$\int_\Omega \min(f,n) + \min(g,n)\ d\mu \leq \int_\Omega \min(f,n)\ d\mu + \int_\Omega \min(g,n)\ d\mu$.

As $\min(f+g, n) \leq \min(f,n) + \min(g,n)$, we get

$\int_\Omega \min(f+g,n)\ d\mu \leq \int_\Omega \min(f,n)\ d\mu + \int_\Omega \min(g,n)\ d\mu$. Taking $n \to \infty$ and using horizontal truncation, we obtain the claim. $\square$

(5). (Vertical truncation) We prove the more general claim that if $E_1 \subset E_2 \subset \dots$ is an increasing sequence of measurable sets, then

$\lim_{n\to\infty} \int_\Omega f1_{E_n}\ d\mu = \int_\Omega f1_{\bigcup_{n=1}^\infty E_n}\ d\mu$.

First assume that f is simple. Then $f = \sum_{i=1}^n a_i 1_{A_i}$, a finite combination of indicators of measurable sets. Let $E = \bigcup_{n=1}^\infty E_n$. For any natural number m, we have

$|\text{Simp} \int_\Omega f1_E\, d\mu - \text{Simp} \int_\Omega f1_{E_m}\, d\mu| = |\sum_{i=1}^n a_i\mu(E \cap A_i) - \sum_{i=1}^n a_i\mu(E_m \cap A_i)|$
$= \sum_{i=1}^n a_i(\mu(E \cap A_i) - \mu(E_m \cap A_i))$.

Note that for each i, $(E_1 \cap A_i) \subset (E_2 \cap A_i) \subset \dots$ is an increasing sequence of measurable sets. Taking $m \to \infty$ and using continuity from below give the claim for the case when f is simple.

Now for a general unsigned measurable f. By definition, we can find a simple function $g \leq f$ such that $|\int_\Omega f\ d\mu\ -\ \mathrm{Simp}\int_\Omega g\ d\mu| \leq \varepsilon$, for any $\varepsilon > 0$. By monotonicity and the triangle inequality, we get

$$|\int_\Omega f\ d\mu - \int_\Omega f 1_{f \geq 1/n}| \leq |\int_\Omega f\ d\mu - \int_\Omega f 1_{g \geq 1/n}\ d\mu| \leq |\int_\Omega f\ d\mu - \int_\Omega g 1_{g \geq 1/n}\ d\mu|$$

$$\leq |\int_\Omega f\ d\mu - \int_\Omega g\ d\mu| + |\int_\Omega g\ d\mu - \int_\Omega g 1_{g \geq 1/n}\ d\mu| \leq \varepsilon + |\int_\Omega g\ d\mu - \int_\Omega g 1_{g \geq 1/n}\ d\mu|.$$

Taking $n \to \infty$ and then $\varepsilon \to 0$ gives the desired result. $\square$

(6). See Exercise 1.4.38 of "M". If either $\int_\Omega f\ d\mu$ or $\int_\Omega g\ d\mu$ is infinite, then by monotonicity $\int_\Omega f\ d\mu$ is infinite as well, and the claim follows; so we may assume that $\int_\Omega f\ d\mu$ and $\int_\Omega g\ d\mu$ are both finite. By Markov's inequality, we conclude that for each natural number n, the set $E_n := \{\omega \in \Omega : f(\omega) > 1/n\} \cup \{\omega \in \Omega : g(\omega) > 1/n\}$ has finite measure. These sets are increasing in n, and f, g, $f+g$ are supported on $\bigcup_{n=1}^\infty E_n$, and so by vertical truncation

$$\int_\Omega f + g\ d\mu = \lim_{n\to\infty} \int_\Omega (f+g) 1_{E_n}\ d\mu.$$

By part (4), we have

$$\int_\Omega (f+g) 1_{E_n}\ d\mu = \int_\Omega f 1_{E_n}\ d\mu + \int_\Omega g 1_{E_n}\ d\mu.$$

Letting $n \to \infty$ and using vertical truncation we obtain the claim. $\square$

Exercise 8

Proof. (1). By Exercise 18 in note 0, both $f+g$ and cf are measurable. Write $f = f_+ - f_-$, where $f_+ := \max(f, 0)$ and $f_- := \max(-f, 0)$, and similarly write $g = g_+ - g_-$, as difference of the magnitudes of the positive and negative parts.

First note that if f and g are unsigned measurable functions such that $f - g$ is also unsigned measurable, then by applying linearity to the sum $f = (f-g)+g$, we get $\int_\Omega f - g\ d\mu = \int_\Omega f\ d\mu - \int_\Omega g\ d\mu$.

By the fact that $f + g = (f+g)_+ - (f+g)_- = f_+ - f_- + g_+ - g_-$, we get $f_+ + g_+ - (f+g)_+ = f_- + g_- - (f+g)_-$. Hence we have

$$\int_\Omega f_+ + g_+ - (f+g)_+\ d\mu = \int_\Omega f_- + g_- - (f+g)_-\ d\mu.$$

As $f_+ + g_+ \geq (f+g)_+$ and $f_- + g_- \geq (f+g)_-$, we can apply the above result to conclude that

$\int_\Omega f_+ \ d\mu + \int_\Omega g_+ \ d\mu - \int_\Omega (f+g)_+ \ d\mu = \int_\Omega f_- \ d\mu + \int_\Omega g_- \ d\mu - \int_\Omega (f+g)_- \ d\mu.$

After rearrangement, this becomes

$(\int_\Omega f_+ \ d\mu - \int_\Omega f_- \ d\mu) + (\int_\Omega g_+ \ d\mu - \int_\Omega g_- \ d\mu) = \int_\Omega (f+g)_+ \ d\mu - \int_\Omega (f+g)_- \ d\mu.$

i.e. $\int_\Omega f \ d\mu + \int_\Omega g \ d\mu = \int_\Omega f + g \ d\mu.$

When $c \geq 0$. We have $\int_\Omega cf \ d\mu := \int_\Omega (cf)_+ \ d\mu - \int_\Omega (cf)_- \ d\mu$
$= \int_\Omega cf_+ \ d\mu - \int_\Omega cf_- \ d\mu = c(\int_\Omega f_+ \ d\mu - \int_\Omega f_- \ d\mu) = c \int_\Omega f \ d\mu.$

When $c = -1$. By definition $\int_\Omega -f \ d\mu = \int_\Omega (-f)_+ \ d\mu - \int_\Omega (-f)_- \ d\mu$
$= \int_\Omega f_- \ d\mu - \int_\Omega f_+ \ d\mu = -\int_\Omega f \ d\mu.$

Finally, when $c < 0$. We have $\int_\Omega cf \ d\mu = \int_\Omega (-1)(-c)f \ d\mu = -\int_\Omega (-c)f \ d\mu = c\int_\Omega f \ d\mu.$ □

(2). By definition, $f = g$ almost everywhere if and only if $f_+ = g_+$ and $f_- = g_-$ almost everywhere. So $\int_\Omega f_+ \ d\mu - \int_\Omega f_- \ d\mu = \int_\Omega g_+ \ d\mu - \int_\Omega g_- \ d\mu$. i.e. $\int_\Omega f \ d\mu = \int_\Omega g \ d\mu$. □

(3). Again, we have $\int_\Omega |f| \ d\mu = \int_\Omega |f|_+ \ d\mu - \int_\Omega |f|_- \ d\mu = \int_\Omega |f|_+ \ d\mu \geq 0$. As $|f| = |f|_+$ and the unsigned integral vanishes if and only if the integrand vanishes almost everywhere, we get the claim. □

(4). By condition, we have $f_+ - f_- \leq g_+ - g_-$ almost everywhere. i.e. $f_+ + g_- \leq g_+ + f_-$ almost everywhere. By monotonicity of the unsigned integral, we then have

$\int_\Omega f_+ + g_- \ d\mu \leq \int_\Omega g_+ + f_- \ d\mu.$

That is, $\int_\Omega f_+ \ d\mu + \int_\Omega g_- \ d\mu \leq \int_\Omega g_+ \ d\mu + \int_\Omega f_- \ d\mu$. And thus $\int_\Omega f_+ \ d\mu - \int_\Omega f_- \ d\mu = \int_\Omega g_+ \ d\mu - \int_\Omega g_- \ d\mu$. i.e. $\int_\Omega f \ d\mu = \int_\Omega g \ d\mu$. □

(5). This follows directly from the the compatibility between the Lebesgue integral and the unsigned integral, and Markov's inequality for the unsigned one. □

Exercise 9

Proof. (1). By monotonicity and the triangle inequality,
$\int_\Omega |f+g| \ d\mu \leq \int_\Omega |f| + |g| \ d\mu = \int_\Omega |f| \ d\mu + \int_\Omega |g| \ d\mu < \infty$. And $\int_\Omega |cf| \ d\mu =$

$\int_\Omega |c||f|\ d\mu = |c| \int_\Omega |f|\ d\mu < \infty$. So $f+g$ and cf are both absolutely integrable.

Also, by definition we have $\int_\Omega f+g\ d\mu := \int_\Omega Re(f+g)\ d\mu + i\int_\Omega Im(f+g)\ d\mu$

$= \int_\Omega Re(f)+Re(g)\ d\mu + i\int_\Omega Im(f)+Im(g)\ d\mu$

$= \int_\Omega Re(f)\ d\mu + \int_\Omega Re(g)\ d\mu + i\int_\Omega Im(f)\ d\mu + i\int_\Omega Im(g)\ d\mu$

$= \int_\Omega Re(f)\ d\mu + i\int_\Omega Im(f)\ d\mu + \int_\Omega Re(g)\ d\mu + i\int_\Omega Im(g)\ d\mu = \int_\Omega f\ d\mu + \int_\Omega g\ d\mu$.

On the other hand, when $c \in \mathbf{R}$, $\int_\Omega cf\ d\mu = \int_\Omega Re(cf)\ d\mu + i\int_\Omega Im(cf)\ d\mu$

$= \int_\Omega Re(c)Re(f) - Im(c)Im(f)\ d\mu + i\int_\Omega Im(c)Re(f) + Re(c)Im(f)\ d\mu$.

On the other hand, $c\int_\Omega f\ d\mu = (Re(c)+iIm(c))(\int_\Omega Re(f)\ d\mu + i\int_\Omega Im(f)\ d\mu)$

$= Re(c)\int_\Omega Re(f)\ d\mu - Im(c)\int_\Omega Im(f)\ d\mu + i(Im(c)\int_\Omega Re(f)\ d\mu + Re(c)\int_\Omega Im(f)\ d\mu)$.

Using linearity of the real-valued absolutely integrable function, we get $\int_\Omega f\ d\mu = c\int_\Omega f\ d\mu$. $\square$

(2). This follows from the equivalency property of real-valued absolutely integrable functions. $\square$

(3). Again, this follows from the positivity and vanishing property of the real-valued counterpart. $\square$

(4). This follows from the Markov's inequality for the real-valued counterpart. $\square$

Exercise 10

Proof. See Lemma 8.3.6 and Proposition 8.4.1 of "Analysis II". Define the upper and lower Lebesgue integral $\overline{\int}_{[a,b]} f$ and $\underline{\int}_{[a,b]} f$ of $f : [a,b] \to \mathbf{R}$ to be such that

$\overline{\int}_{[a,b]} f\ dm := \inf\{\int_{[a,b]} g\ dm,\ g \text{ absolutely integrable},\ g \geq f\}$

$\underline{\int}_{[a,b]} f\ dm := \sup\{\int_{[a,b]} g\ dm,\ g \text{ absolutely integrable},\ g \leq f\}$

We first show that the agreement between upper and lower Lebesgue integrals implies absolute integrability, with their common value being the Lebesgue integral.

Let $A \in \mathbf{R}$, with $\underline{\int}_{[a,b]} f\ dm = \overline{\int}_{[a,b]} f\ dm = A$.

By definition of upper Lebesgue integral, for every $n \geq 1$ we may find an absolutely integrable function $f_n^+ : [a,b] \to \mathbf{R}$ which majorizes f such that $\int_{[a,b]} f_n^+\ dm \leq A + \frac{1}{n}$. Similarly we may find an absolutely integrable function $f_n^- : [a,b] \to \mathbf{R}$ which minorizes f such that $\int_{[a,b]} f_n^-\ dm \geq A - \frac{1}{n}$. Let $F^+ := \inf_n f_n^+$ and $F^- := \sup_n f_n^-$. Then F^+ and F^- are measurable and absolutely integrable (because they are squeezed between the absolutely integrable functions f_1^+ and f_1^-, for instance). Also, F^+ majorizes f and F^- minorizes f. Finally, we have

$\int_{[a,b]} F^+\ dm \leq \int_{[a,b]} f_n^+\ dm \leq A + \frac{1}{n}$ for every n, and hence $\int_{[a,b]} F^+\ dm \leq A$.

Similarly we have $\int_{[a,b]} F^-\ dm \geq A$. But F^+ majorizes F^-, and hence $\int_{[a,b]} F^+\ dm \geq \int_{[a,b]} F^-\ dm$. Hence we must have $\int_{[a,b]} F^+\ dm = \int_{[a,b]} F^-\ dm = A$. In particular, $\int_{[a,b]} F^+ - F^-\ dm = 0$, suggesting that $F^+(x) = F^-(x)$ almost everywhere. But since f is squeezed between F^- and F^+, we thus have $f(x) = F^+(x) = F^-(x)$ for almost every x. In particular, f differs from the absolutely integrable function F^+ only by a null set and is thus measurable (Exercise 1.3.8 of "M") and absolutely integrable, with $\int_{[a,b]} f\ dm = \int_{[a,b]} F^+\ dm = \int_{[a,b]} F^-\ dm = A$.

Now we show that if $f : [a,b] \to \mathbf{R}$ is Riemann integrable, then its' upper and lower Lebesgue integrals agree, with the Riemann integral $A := \int_a^b f(x)\ dx$ being their common value.

Since f is Riemann integrable, the upper and lower Riemann integrals are equal to A. Thus, for every $\varepsilon > 0$, there is a partition $\mathbf{P}$ of $[a,b]$ into smaller intervals J such that

$A - \varepsilon \leq \sum_{J \in \mathbf{P}} |J| \inf_{x \in J} f(x) \leq A \leq \sum_{J \in \mathbf{P}} |J| \sup_{x \in J} f(x) \leq A + \varepsilon$

Let $f_\varepsilon^- : [a,b] \to \mathbf{R}$ and $f_\varepsilon^+ : [a,b] \to \mathbf{R}$ be the functions

$f_\varepsilon^-(x) := \sum_{J \in \mathbf{P}} \inf_{x \in J} f(x) 1_J(x)$ and $f_\varepsilon^+(x) := \sum_{J \in \mathbf{P}} \sup_{x \in J} f(x) 1_J(x)$.

These are simple functions and thus measurable and absolutely integrable, so $\int_{[a,b]} f_\varepsilon^-\ dm = \sum_{J \in \mathbf{P}} |J| \inf_{x \in J} f(x)$, and $\int_{[a,b]} f_\varepsilon^+\ dm = \sum_{J \in \mathbf{P}} |J| \sup_{x \in J} f(x)$, and hence

$A - \varepsilon \leq \int_{[a,b]} f_\varepsilon^-\ dm \leq A \leq \int_{[a,b]} f_\varepsilon^+\ dm \leq A + \varepsilon$

Since f_ε^+ majorizes f, and f_ε^- minorizes f, we thus have

$A - \varepsilon \leq \underline{\int}_{[a,b]} f\ dm \leq \overline{\int}_{[a,b]} f\ dm \leq A + \varepsilon$ for every ε, and thus

$$\underline{\int}_{[a,b]} f\ dm = \overline{\int}_{[a,b]} f\ dm = A$$

By our first result, f is absolutely integrable with $\int_{[a,b]} f\ dm = A$, as desired. The same argument runs through with the interval $[a,b]$ replaced by the box $[a_1, b_1] \times \ldots \times [a_n, b_n]$. □

Exercise 12

Proof. Suppose that X and Y are modeled by some sample space Ω. Since Y is simple, we can find distinct reals $a_1 < a_2 < \ldots < a_n \in [0, +\infty]$, and disjoint measurable sets $E_1, \ldots, E_n$ such that $Y_\Omega = \sum_{i=1}^n a_i 1_{E_i}$. For any $x \in [0, +\infty]$, define the sets L_x and R_x by setting $L_x := \{a_i : a_i \leq x,\ 1 \leq i \leq n\}$ and $R_x := \{a_i : a_i > x,\ 1 \leq i \leq n\}$.

Define $f : [0, +\infty] \to [0, +\infty]$ by

$$f(x) := \begin{cases} \max(L_x) & \text{when } L_x \neq \emptyset \\ 0 & \text{when } L_x = \emptyset \end{cases}$$

We first show that f is measurable. $\forall t \in [0, +\infty]$, either $R_t \neq \emptyset$ or $R_t = \emptyset$. In the former case, by letting $c_t := \min(R_t)$ one can check that $f^{-1}([0,t]) = [0, c_t)$. In the latter case, $f^{-1}([0,t]) = [0, a_n]$. Hence in particular we see that $f^{-1}([0,t])$ is Lebesgue measurable for any $t \in [0, +\infty]$, by Exercise 24 of note 0 (or Lemma 1.3.9 of "M"), we thus see that f is Lebesgue measurable.

Now we show that f has the desired properties. By construction, f can take at most $n+1$ values. As f is measurable, $f(X)$ is a random variable taking values in $[0, +\infty]$. Since $Y \leq X$ is surely true, we see from the definition of f that $Y \leq f(X) \leq X$ is surely true.

This implies that any simple unsigned random variable Y modeled by Ω such that $0 \leq Y \leq X$ can be replaced by the simple unsigned random variable $f(X)$ modeled by the same space. And for any choice of model Ω, the expectation $\mathbf{E}X$ will be equal to $\sup_{f,\ 0 \leq f(X) \leq X} \mathrm{Simp}\mathbf{E}X$, where the supremum ranges over all simple unsigned functions $f : [0, +\infty] \to [0, +\infty]$. □

Some note on the equivalent form of Fatou's Lemma: That the original form implies the second is obvious. To show the other direction, let $F_N := \inf_{n \geq N} f_n$.

The case is trivial when $M := \liminf_{n\to\infty} \int_\Omega f_n \, d\mu$ is infinite, so we let $M := \liminf_{n\to\infty} \int_\Omega f_n \, d\mu$ be finite. By monotonicity, $\int_\Omega F_N \, d\mu \le \int_\Omega f_n \, d\mu$ for all $n \ge N$, hence in particular $\int_\Omega F_N \, d\mu \le \inf_{n\ge N} \int_\Omega f_n \, d\mu \le M$. Applying the second form to F_N then gives the desired result.

Exercise 20

The only quantitative bound on probability available to us at this point, besides the one provided by the condition, is the Markov's inequality, which gives upper rather than lower bound, hence we should involve the compliment event to flip the upper bound to the lower one.

Proof. Since $\inf_n \mathbf{P}(E_n) > 0$, the sequence $\mathbf{P}(E_n)$ is bounded away from 0. i.e. There exists some $\delta > 0$ such that $\mathbf{P}(E_n) \ge \delta > 0$ for all n. Without loss of generality, we make $0 < \delta < 1$. Notice that

$$\mathbf{P}(\textstyle\sum_{n\le N} 1_{E_n} \ge \delta N/2) = 1 - \mathbf{P}(\sum_{n\le N} 1_{E_n} < \delta N/2) = 1 - \mathbf{P}(\frac{1}{N}\sum_{n\le N} 1_{E_n} < \delta/2)$$

Then observe that $\frac{1}{N}\sum_{n\le N} 1_{E_n} < \delta/2$ holds if and only if at most $\lfloor \frac{N\delta}{2} \rfloor$ of the E_n hold. Or equivalently, if and only if at least $N - \lfloor \frac{N\delta}{2} \rfloor$ of the $\overline{E_n}$ hold. By Markov's inequality, we get

$$\mathbf{P}(\textstyle\sum_{n\le N} 1_{\overline{E_n}} \ge N - \lfloor \frac{N\delta}{2} \rfloor) \le \mathbf{P}(\sum_{n\le N} 1_{\overline{E_n}} \ge N - \frac{N\delta}{2}) \le \frac{2}{N(2-\delta)} \mathbf{E} \sum_{n\le N} 1_{\overline{E_n}}$$

$$= \tfrac{2}{N(2-\delta)} \textstyle\sum_{n\le N} \mathbf{P}(1_{\overline{E_n}}) \le \frac{2(1-\delta)}{2-\delta}$$

Thus we see that $\mathbf{P}(\sum_{n\le N} 1_{E_n} \ge \delta N/2) \ge 1 - \frac{2(1-\delta)}{2-\delta} \ge \frac{\delta}{2}$. Sending $N \to \infty$ gives the desired result.

For the alternative approach, we see by the zero test that if $\limsup_{n\to\infty} 1_{E_n} = 1$, then $\sum_{n\to\infty} 1_{E_n} = \infty$. Thus we have

$$\mathbf{P}(\textstyle\sum_{n\to\infty} 1_{E_n} = \infty) \ge \mathbf{P}(\limsup_{n\to\infty} 1_{E_n} = 1).$$

From the identity $\limsup_{n\to\infty} 1_{E_n} = 1 - \liminf_{n\to\infty} 1_{\overline{E_n}}$, we see that

$$\mathbf{P}(\limsup_{n\to\infty} 1_{E_n} = 1) = \mathbf{P}(\liminf_{n\to\infty} 1_{\overline{E_n}} = 0)$$

$$= 1 - \mathbf{P}(\liminf_{n\to\infty} 1_{\overline{E_n}} = 1) = 1 - \mathbf{E} \liminf_{n\to\infty} 1_{\overline{E_n}} \ge 1 - \liminf_{n\to\infty} \mathbf{E} 1_{\overline{E_n}}$$

$= 1 - \liminf_{n\to\infty} \mathbf{P}(\overline{E_n}) \ge 1 - (1-\delta) = \delta > 0$, where we've used Fatou's lemma in exchanging limit and expectation. $\square$

Exercise 21

Consider $(\Omega, \mathcal{F}, \mu) = ([0,1], \mathcal{B}([0,1]), m)$, the Borel σ-algebra of the unit interval with the Lebesgue measure. The idea is that if $\sum_{n=1}^{\infty} p_n = +\infty$, one should be able to get a covering of $[0,1]$ with arbitrary "thickness".

Define $L_0 := 0$, for each $n \geq 1$, let $L_n := L_{n-1} + p_n$. Define E_n for all n by setting $E_n := [L_{n-1}, L_n]$ if $L_n \leq 1$ and $E_n := [L_{n-1}, 1] \cup [1, L_n]$ if $L_n > 1$. In this way we wrap around $[0,1]$ by the E_n's infinitely many times, so almost surely infinitely many of the E_n occur. By construction, we clearly have $\mathbf{P}(E_n) = p_n$ for all n. Refer also to Exercise 1.4.45 of "M".

Exercise 27

Proof. See also Exercise 1.4.46 (Almost dominated convergence) of "M". From the hint, once we obtain the decomposition $X_n = X_{n,1} + X_{n,2}$ with the desired properties, by the triangle inequality we have

$\mathbf{E}|X - X_n| \leq \mathbf{E}|X - X_{n,1}| + \mathbf{E}|X_{n,2}|.$

Note that $|X - X_{n,1}| \leq |X| + |X_{n,1}| \leq 2|X|$, hence we can apply the dominated convergence theorem and conclude that $\mathbf{E}|X - X_{n,1}| \to 0$. On the other hand, from the property $|X_n| = |X_{n,1}| + |X_{n,2}|$, we have $\mathbf{E}|X_n| = \mathbf{E}|X_{n,1}| + \mathbf{E}|X_{n,2}|$, which shows by the fact that $\mathbf{E}|X_n|$ converges to $\mathbf{E}|X|$ and dominated convergence theorem, that $\mathbf{E}|X_{n,2}| \to 0$, establishing our claim.

Define $X_{n,1} := \begin{cases} X_n & \text{when } |X_n| \leq |X| \\ X & \text{when } X_n \text{ and } X \text{ have same signs and } |X_n| > |X| \\ -X & \text{when } X_n \text{ and } X \text{ have opposite signs and } |X_n| > |X| \end{cases}$

And $X_{n,2} := X_n - X_{n,1}$. We verify that the given decomposition obey the given properties. Clearly $X_{n,1}$ is dominated by X and converges almost surely to X by construction. In all cases we also have $|X_n| = |X_{n,1}| + |X_{n,2}|$, as desired. $\square$

Exercise 36

Proof. Let $E_1, E_2, \ldots$ be a disjoint sequence of Borel sets in $\mathbf{R}$. By the monotone convergence theorem, we have

$m_f(\bigcup_n E_n) := \int_{\cup_n E_n} f(x)\ dx = \int_{\mathbf{R}} f 1_{\cup_n E_n}(x)\ dx = \int_{\mathbf{R}} \sum_{n=1}^{\infty} f 1_{E_n}(x)\ dx = \sum_{n=1}^{\infty} \int_{\mathbf{R}} f 1_{E_n}(x)\ dx = \sum_{n=1}^{\infty} \int_{E_n} f(x)\ dx = \sum_{n=1}^{\infty} m_f(E_n)$.

Clearly $m_f(\emptyset) := \int_{\emptyset} f(x)\ dx = 0$, and $m_f(\mathbf{R}) := \int_{\mathbf{R}} f(x)\ dx = 1$. So m_f is a probability measure on $\mathbf{R}$. By definition, the Stieltjes measure function associated to m_f is given by $F(t) := m_f((-\infty, t]) = \int_{(-\infty,t]} f(x)\ dx$ for any $t \in \mathbf{R}$.

By the change of variables formula, $\mathbf{E}G(X) := \int_{\mathbf{R}} G(x)\ dm_f(x)$. Hence we want to show that $\int_{\mathbf{R}} G(x)\ dm_f(x) = \int_{\mathbf{R}} G(x) f(x)\ dx$, when either $G : \mathbf{R} \to [0, +\infty]$ is an unsigned measurable function, or $G : \mathbf{R} \to \mathbf{C}$ is measurable with $G(X)$ absolutely integrable.

By taking real and imaginary parts, we can assume that G is real valued, by writing $G = G_1 - G_2$, where G_2 and G_2 are both unsigned, we can further assume that G is unsigned.

Finally, for simple unsigned $G = \sum_{i=1}^{n} a_i 1_{E_i}$, we have $\int_{\mathbf{R}} G(x)\ dm_f(x) = \int_{\mathbf{R}} \sum_{i=1}^{n} a_i 1_{E_i}(x)\ dm_f(x)$

$= \sum_{i=1}^{n} a_i \int_{\mathbf{R}} 1_{E_i}(x)\ dm_f(x) = \sum_{i=1}^{n} a_i m_f(E_i) = \int_{\mathbf{R}} G(x) f(x)\ dx$. As desired. ☐

Exercise 37

Proof. Let $G : \mathbf{R} \to \mathbf{R}$ be such that $G(x) := xF(x)$, and $n : \mathbf{R} \to [0, +\infty]$ be the probability density function $n(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. Then G is measurable and of polynomial growth. In particular, it can be shown that both $G(X)$ and $F'(X)$ are absolutely integrable. i.e. $\int_{\mathbf{R}} |G(x)| n(x)\ dx$ and $\int_{\mathbf{R}} |F'(x)| n(x)\ dx$ are both finite (as $e^{-x^2/2}$ decays much faster than any polynomial grows). Hence by Exercise 36, we have:

$\mathbf{E}XF(X) = \mathbf{E}G(X) = \int_{\mathbf{R}} G(x) n(x)\ dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbf{R}} xF(x) e^{-x^2/2}\ dx.$

By integration by parts, and the compatibility between Lebesgue and Riemann integrals, this becomes

$-\frac{1}{\sqrt{2\pi}} F(x) e^{-x^2/2} \Big|_{-\infty}^{+\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} F'(x)\ dx = 0 + \mathbf{E}F'(X) = \mathbf{E}F'(X).$

Finally, define $F : \mathbf{R} \to \mathbf{R}$ by $F(x) := x^{k-1}$ for any $k > 0$, by the Stein identity $\mathbf{E}X^k = \mathbf{E}X \cdot X^{k-1} = (k-1)\mathbf{E}X^{k-2}$. When $k = 1$, we clearly have $\mathbf{E}X = 0\mathbf{E}X^{-1} = 0$. Suppose inductively that $\mathbf{E}X^k = 0$ for some $k = 2n - 1$ with $n \geq 1$, then $\mathbf{E}X^{2n+1} = 2n\mathbf{E}X^{2n-1} = 2n \cdot 0 = 0$, induction closed.

Similarly, we have $\mathbf{E}X^0 = \mathbf{E}1 = 1$. Suppose inductively that $\mathbf{E}X^k = (k-1)(k-3)\dots 1$ for some $k = 2n$ with $n \geq 1$, then $\mathbf{E}X^{2n+2} = (2n+1)\mathbf{E}X^{2n} = (k+1)(k-1)(k-3)\dots 1$, induction closed. $\square$

Exercise 38

There are multiple ways to proceed. For instance, one can use the Fubini–Tonelli theorem, or one can approximate the exponential function by a piecewise constant function and use dominated convergence to justify passing to the limit. One can show that PDF is the derivative of CDF and use Exercise 36. For this problem a PDF is not assumed to exist (except in a distributional or measure theoretic sense), so care would have to be taken if one wished to proceed by that route. -T

Proof. We first show that if X is a non-negative random variable, then $\mathbf{E}X = \int_0^\infty \mathbf{P}(X \geq x)\ dx$. Note that $X = \int_0^X 1\ dx = \int_0^\infty 1_{X \geq x}\ dx$. Hence $\mathbf{E}X = \mathbf{E}(\int_0^\infty 1_{X\geq x}\ dx) = \int_0^\infty \mathbf{P}(X \geq x)\ dx$, where we use Fubini-Tonelli theorem in the last step.

Then we have $\mathbf{E}e^{tX} = \int_0^\infty \mathbf{P}(e^{tX} \geq x)\ dx = \int_0^\infty \mathbf{P}(X \geq \ln x/t)\ dx$,

which equals $\int_{-\infty}^{+\infty} \mathbf{P}(X \geq u) te^{tu}\ du = \int_{\mathbf{R}} (1 - F_X(u)) te^{tu}\ du$ by substituting $u = \ln x/t$.

By a similar argument, if X is non-negative, we can show that $\mathbf{E}X^p = \int_0^\infty \mathbf{P}(X \geq u) pu^{p-1}\ du$ for all $p > 0$. $\square$

Exercise 40

Proof. We view f as a convex function on $\mathbf{R^2}$ and let $x_0 \in \mathbf{R^2}$. As in the real case, the graph of f must be supported by some hyperplane at $(x_0, f(x_0))$, that is to say there exists a vector v in $\mathbf{R^2}$ depending on x_0 such that $f(x) \geq f(x_0) + v \cdot (x - x_0)$ for all $x \in \mathbf{R^2}$, in the usual sense of vector dot product. In particular, $f(X) \geq f(x_0) + v \cdot (X - x_0)$. Let $X = U + iV$, where U, V are real random variables, $a = (a, b)$ and $x_0 = (a_0, b_0)$, we see that

$$f(U + iV) \geq f(x_0) + a(U - a_0) + b(V - b_0).$$

Taking expectations and using linearity of expectation, we conclude

$$\mathbf{E}f(X) \geq f(x_0) + a(\mathbf{E}U - a_0) + b(\mathbf{E}V - b_0) = f(x_0) + v \cdot (\mathbf{E}X - x_0)$$

and the claim follows from setting $x_0 := \mathbf{E}X$. $\square$

Exercise 41

Proof. Let $q > p$. Since the function $x \mapsto x^c$ is convex for any $c > 0$, we can apply Jensen's inequality to obtain

$(\mathbf{E}|X|^p)^{q/p} \leq \mathbf{E}|X|^q$,

taking the q-th root on both sides gives the desired result.

Alternatively, let $0 < r < s$, and $p = s/r$, q be such that $1/p + 1/q = 1$. Apply Hölder inequality to the scalars $|X|^r$ and 1, we get

$\mathbf{E}|X|^r \leq (\mathbf{E}|X|^s)^{r/s}$,

Taking the r-th root on both sides gives the desired result. $\square$

Exercise 42

Proof. Equivalently, we want to show that $\mathbf{P}(X \neq 0)\mathbf{E}(|X|^2) \geq (\mathbf{E}|X|)^2$. By Cauchy-Schwarz inequality, we get LHS $= \mathbf{E}1_{X\neq 0} \cdot \mathbf{E}(|X|^2) = (\|1_{X\neq 0}\|_2 \cdot \||X|\|_2)^2 \geq |\mathbf{E}(|X| \cdot 1_{X\neq 0})|^2 = (\mathbf{E}|X|)^2$, as desired. $\square$

Exercise 43

Proof. Let q be such that $1/p + 1/q = 1$. Then $q = \frac{p}{p-1}$. By Hölder inequality, we have:

$\mathbf{E}|X||X+Y|^{p-1} \leq \|X\|_p \|(X+Y)^{p-1}\|_q = \|X\|_p \|X+Y\|_p^{p-1}$,

using the fact that the p-norm of $|X|$ is equal to the p-norm of X. Similarly we can show that $\mathbf{E}|Y||X+Y|^{p-1} \leq \|Y\|_p \|X+Y\|_p^{p-1}$.

Now, for any any $1 < p < \infty$, we have:

$\|X+Y\|_p^p = \mathbf{E}|X+Y|^p = \mathbf{E}|X+Y||X+Y|^{p-1}$

$\leq \mathbf{E}|X||X+Y|^{p-1} + \mathbf{E}|Y||X+Y|^{p-1} \leq \|X+Y\|_p^{p-1}(\|X\|_p + \|Y\|_p)$.

The case when $\|X+Y\| = 0$ is trivial. For $\|X+Y\|_p > 0$, dividing both sides of the above inequality by $\|X+Y\|_p^{p-1}$ gives the desired result. The edge case for $p = 1$ follows from triangle inequality and linearity, while the edge case for $p = \infty$ follows from definition. $\square$

Exercise 44

Proof. By definition of expectation and the Cauchy-Schwarz inequality, we can write

$$\mathbf{E}X = \mathbf{E}X1_{X \le \theta\mathbf{E}X} + \mathbf{E}X1_{X > \theta\mathbf{E}X} \le \theta\mathbf{E}X + (\mathbf{E}|X|^2)^{1/2}(\mathbf{P}(X > \theta\mathbf{E}X))^{1/2},$$

which after rearrangement gives the desired inequality. $\square$

Exercise 45

Proof. As in the proof of the Paley-Zygmund inequality, we have

$$\mathbf{E}X = \mathbf{E}X1_{X<1/2\mathbf{E}X} + \mathbf{E}X1_{X \ge 1/2\mathbf{E}X} \le \frac{1}{2}\mathbf{E}X + \|X\|_2(\mathbf{P}(X > \mathbf{E}X))^{1/2},$$

from which follows that $\frac{1}{2}\mathbf{E}X \le \|X\|_2(\mathbf{P}(X > \mathbf{E}X))^{1/2}$, thus

$\mathbf{P}(X > \mathbf{E}X) \ge (\mathbf{P}(X > \mathbf{E}X))^{1/2} \ge \frac{1}{2}\frac{\mathbf{E}X}{\|X\|_2} \ge \frac{1}{2}\frac{\mathbf{E}X}{\|X\|_\infty}$ by the monotonicity of p-norm in p (note that the condition that X is a.s bounded but not identically zero ensures that the denominator $0 < \|X\|_\infty < \infty$). $\square$