PROBLEMS 1.

1.

Let $\Omega = [0,1] \cap \mathbb{Q} = \{r \text{ rational} \mid 0 \le r \le 1\}$ and $\mathcal{A}$ the collection of sets which can be represented as finite unions of intervals of type $]a,b] \cap \mathbb{Q}$, $[a,b] \cap \mathbb{Q}$, $]a,b[\cap \mathbb{Q}$ or $[a,b[\cap \mathbb{Q}$ with $0 \le a \le b \le 1$.

Define for all $0 \le a \le b \le 1$:

$$P(]a,b] \cap \mathbb{Q}) = P([a,b] \cap \mathbb{Q}) = P(]a,b[\cap \mathbb{Q}) = P([a,b[\cap \mathbb{Q}) = b-a$$

Claim: $\mathcal{A}$ is an algebra, also, $\Omega \in \mathcal{A}$ and when $A \in \mathcal{A}$ also $A^c := \Omega \setminus A \in \mathcal{A}$ and if $A, B \in \mathcal{A}$ also $A \cup B \in \mathcal{A}$.

Proof:

Direct proof:

Obviously $\Omega = [0,1] \cap \mathbb{Q} \in \mathcal{A}$.

Let $A = \bigcup_{i=1}^{n} I_i \cap \mathbb{Q} \in \mathcal{A}$ where I_i is an interval of type $]a_i; b_i]$ $[a_i; b_i]$, $]a_i; b_i[$ or $[a_i; b_i[$ for some $0 \le a_i \le b_i \le 1$. It can be assumed that $0 \le a_1 \le b_1 \le a_2 \le b_2 \le \dots \le b_n \le 1$. Then

$$A^c = \bigcup_{j=0}^{n} J_j \cap \mathbb{Q}$$

where J_j is an interval of type $]b_j; a_{j+1}[$, $[b_j; a_{j+1}]$, $]b_j; a_{j+1}]$ or $[b_j; a_{j+1}[$ with $b_0 = 0$ and $a_{n+1} = 1$ such that $a_j \in I_j$ $\Leftrightarrow a_j \notin J_{j-1}$ and $b_j \notin I_j \Rightarrow b_j \in J_j$ (we agree $I_0 = \emptyset$).

Thus $A_c \in \mathcal{A}$.

Let then $A, B \in \mathcal{A}$. As they are representable as finite unions of intervals of given type, so is their union $A \cup B$, so $A \cup B \in \mathcal{A}$.

So $\mathcal{A}$ is an algebra. $\square$

Let us extend the function P to a finitely additive propability on the algebra $\mathcal{A}$.

For $A=\bigcup_{i=1}^{n} I_i \cap \mathbb{Q} \in \mathcal{A}$ as in previous proof (with I_i separate intervals with endpoints $a_i \leq b_i$) we define

$$P(A)=\sum_{i=1}^{n}(b_i-a_i) \geq 0$$

Obviously this is an extension of P as a non-negative function to whole of $\mathcal{A}$. As $P([0,1]\cap\mathbb{Q})=P(\Omega)=1-0=1$, and since for $A,B\in\mathcal{A}$,

$$A=\bigcup_{i=1}^{n} I_i \cap \mathbb{Q}$$

$$B=\bigcup_{j=1}^{k} J_j \cap \mathbb{Q}$$

with $A\cap B=\emptyset$, I_i and J_j separate intervals with endpoints $a_i \leq b_i$; $c_j \leq d_j$, we have

$$P(A\cup B)=P\left(\bigcup_{i=1}^{n} I_i \cap \mathbb{Q} \cup \bigcup_{j=1}^{k} J_j \cap \mathbb{Q}\right)=\sum_{i=1}^{n}(b_i-a_i)+\sum_{j=1}^{k}(d_i-c_i)$$

$$=P(A)+P(B)$$

so P is (by simple induction) finitely additive.

Claim: such additive P is not σ-additive

Proof:

Proof by contradiction.

Counterassumption: Suppose P is σ-additive.

Let $\Omega=[0,1]\cap\mathbb{Q}=\{x_1;x_2;\dots\}$ be an enumeration of Ω (Ω is countable). Note $\{x_j\}=[x_j;x_j]\cap\mathbb{Q}\in\mathcal{A}$, $0\leq x_j\leq 1$. Now

$$1=P(\Omega)=P\left(\bigcup_{j=1}^{\infty}\{x_j\}\right)=\sum_{j=1}^{\infty}P(\{x_j\})=\sum_{j=1}^{\infty}(x_j-x_j)=\sum_{j=1}^{\infty}0=0 \ \lightning$$

Thus counterassumption is false. So P cannot be σ-additive.

Thus the claim is true. $\square$

2.

Let us consider an abstract set Ω, and let us define the collection (we assume that A and A^c both can be finite for $A\in\mathcal{A}$)

$$\mathcal{A}=\{A\subseteq\Omega \mid \text{either } A \text{ or its complement } A^c=\Omega\setminus A \text{ is finite}\}$$

Claim: $\mathcal{A}$ is an algebra but it is not a σ-algebra.

Proof:

Direct proof:

$\Omega\in\mathcal{A}$ as $\Omega^c=\emptyset$ is finite.

If $A\in\mathcal{A}$, then as either A^c or $(A^c)^c=A$ is finite, $A^c\in\mathcal{A}$.

If $A,B\in\mathcal{A}$ then: if A^c or B^c is finite, since $(A\cup B)^c\subseteq A^c$ and $(A\cup B)^c\subseteq B^c$, also $(A\cup B)^c$ is finite and $A\cup B\in\mathcal{A}$. If A and B are both finite, $A\cup B$ is finite and $A\cup B\in\mathcal{A}$.

So $\mathcal{A}$ is an algebra.

If Ω is finite, $\mathcal{A}=\mathcal{P}(\Omega)$ is a σ-algebra.

If Ω is infinite, let $\{x_1;x_2;\ldots\}\subseteq\Omega$ be an infinite set so that $\{x_1;x_2;\ldots\}^c$ is also infinite. Then $\{x_j\}\in\mathcal{A}$ $\forall j\in\mathbb{N}$ but $\{x_1;x_2;\ldots\}=\bigcup_{j=1}^{\infty}\{x_j\}\notin\mathcal{A}$, so $\mathcal{A}$ is not a σ-algebra.

Hence the claim is true. $\square$

For $A\in\mathcal{A}$, define $Q(A)=0$ when A is finite and $Q(A)=1$ when A is infinite.

Claim: Q is finitely additive on $\mathcal{A}$ but not σ-additive.

Proof:

Direct proof:

If Ω is finite, $Q(A)=0$ $\forall A\in\mathcal{A}$ and Q is σ-additive, and thus also additive, on $\mathcal{A}$.

If Ω is infinite, let $A, B \in \mathcal{A}$, $A \cap B = \emptyset$. We study 3 cases:

1° A, B both finite:

$$Q(A \cup B) = 0 = 0 + 0 = Q(A) + Q(B)$$

2° One of A, B finite and the other infinite:

$$Q(A \cup B) = 1 = 0 + 1 = Q(A) + Q(B)$$

3° Both A, B infinite. Now as $A \cap B = \emptyset$, $B \subseteq A^c$. But A^c is finite, as $A \in \mathcal{A}$ is infinite. So B is finite ↯.

So case 3° is impossible.

So in all cases, for $A, B \in \mathcal{A}$, $A \cap B = \emptyset$: $Q(A \cup B) = Q(A) + Q(B)$.
So (by easy induction) Q is additive on $\mathcal{A}$.

Now, if Ω is infinite, then there exists a set $\{x_1; x_2; \ldots\} \subseteq \Omega$; then if Q would be σ-additive ($\{x_j\} \in \mathcal{A}$ $\forall j \in \mathbb{N}$)

$$1 = Q(\{x_1; x_2; \ldots\}) = Q\Big(\bigcup_{j=1}^{\infty} \{x_j\}\Big) = \sum_{j=1}^{\infty} Q(\{x_j\}) = \sum_{j=1}^{\infty} 0 = 0 \text{ ↯}$$

So Q cannot be σ-additive when Ω is infinite. Above we assume that it is not necessary that $\{x_1; x_2; \ldots\} \in \mathcal{A}$ as Q is well-defined on whole $\mathcal{P}(\Omega)$. If we further demand $\{x_1; x_2; \ldots\} \in \mathcal{A}$ then, if Ω is countably infinite, we choose $\{x_1; x_2; \ldots\} = \Omega \in \mathcal{A}$ and the argument above shows Q is not σ-additive on $\mathcal{A}$. However, if Ω is uncountably finite, $\mathcal{A}$ contains no countable set, and actually Q is σ-additive on $\mathcal{A}$: If $A_1; A_2; \ldots \in \mathcal{A}$ are disjoint and $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$, then $Q\big(\bigcup_{n=1}^{\infty} A_n\big) = \sum_{n=1}^{\infty} Q(A_n)$ [$1 = 1 + \sum_{n=2}^{\infty} 0$] as maximum one of the sets A_n is infinite by disjointness, and if A_n are finite but $\bigcup_{n=1}^{\infty} A_n$ is infinite then $\bigcup_{n=1}^{\infty} A_n \notin \mathcal{A}$. Case A_n finite, $\bigcup_{n=1}^{\infty} A_n$ finite: $0 = \sum_{n=1}^{\infty} 0$.

Thus the claim is almost true. □

3. Let Ω be an abstract set and 2^{Ω} its power set, which is the collection of subsets $A \subseteq \Omega$.

Let us define the symmetric difference of $A, B \subseteq \Omega$ as

$$A \triangle B = (A \cup B) \setminus (A \cap B) = \{\omega \mid \omega \in A \text{ or } \omega \in B \text{ but not in both}\}.$$

Claim: 2^{Ω} is a ring with respect to the operations $\triangle$ (sum) and $\cap$ (product), which means

- There exists an identity element with respect to the operation $\triangle$

- There exists an identity element with respect to the operation $\cap$
- Every element $A \subseteq \Omega$ has an additive inverse
- Δ is associative and the distributive property holds between Δ and $\cap$

Proof:

Direct proof:

Let us assume all other properties of ring are in this case obvious, except the ones stated above.

Let us prove these claims by indicator manipulations. We note that

$$\mathbb{1}_{A\cap B} = \mathbb{1}_A \mathbb{1}_B$$

$$\mathbb{1}_{A\cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_A \mathbb{1}_B$$

$$\mathbb{1}_{A\Delta B} = (\mathbb{1}_A + \mathbb{1}_B) \bmod 2 = \mathbb{1}_A + \mathbb{1}_B - 2\mathbb{1}_A \mathbb{1}_B = \mathbb{1}_A \mathbb{1}_{B^c} + \mathbb{1}_B \mathbb{1}_{A^c}$$

$\forall A, B \subseteq \Omega$. Now we see that, for $A \subseteq \Omega$

$$\mathbb{1}_{A\Delta\phi} = \mathbb{1}_A \mathbb{1}_{\phi^c} + \mathbb{1}_\phi \mathbb{1}_{A^c} = \mathbb{1}_A \mathbb{1}_\Omega + 0 \cdot \mathbb{1}_{A^c} = \mathbb{1}_A$$

$$\mathbb{1}_{\phi\Delta A} = \mathbb{1}_\phi \mathbb{1}_{A^c} + \mathbb{1}_A \mathbb{1}_{\phi^c} = 0 \cdot \mathbb{1}_{A^c} + \mathbb{1}_A \mathbb{1}_\Omega = \mathbb{1}_A$$

since $\mathbb{1}_\phi \equiv 0$, $\mathbb{1}_\Omega \equiv 1$. Thus ϕ is the identity element with respect to operation Δ.

Also for $A \subseteq \Omega$ we see:

$$\mathbb{1}_{A\cap\Omega} = \mathbb{1}_A \mathbb{1}_\Omega = \mathbb{1}_A = \mathbb{1}_\Omega \mathbb{1}_A = \mathbb{1}_{\Omega\cap A}$$

so that Ω is the identity element with respect to operation $\cap$

Then let $A \subseteq \Omega$. We note

$$\mathbb{1}_{A\Delta A} = \mathbb{1}_A \mathbb{1}_{A^c} + \mathbb{1}_A \mathbb{1}_{A^c} \equiv 0 + 0 \equiv 0 \equiv \mathbb{1}_\phi$$

so the additive inverse of $A \subseteq \Omega$ is A.

Then let $A, B, C \subseteq \Omega$. We have

$$\mathbb{1}_{A \Delta (B \Delta C)} = \mathbb{1}_A + \mathbb{1}_{B \Delta C} - 2 \mathbb{1}_A \mathbb{1}_{B \Delta C} = \mathbb{1}_A + (\mathbb{1}_B + \mathbb{1}_C - 2 \mathbb{1}_B \mathbb{1}_C)$$

$$- 2 \mathbb{1}_A (\mathbb{1}_B + \mathbb{1}_C - 2 \mathbb{1}_B \mathbb{1}_C) = (\mathbb{1}_A + \mathbb{1}_B - 2 \mathbb{1}_A \mathbb{1}_B) + \mathbb{1}_C - 2 (\mathbb{1}_A + \mathbb{1}_B - 2 \mathbb{1}_A \mathbb{1}_B) \mathbb{1}_C$$

$$= \mathbb{1}_{A \Delta B} + \mathbb{1}_C - 2 \mathbb{1}_{A \Delta B} \mathbb{1}_C = \mathbb{1}_{(A \Delta B) \Delta C}$$

so Δ is associative.

Let then $A, B, C \subseteq \Omega$. Let us note $\mathbb{1}_{A \cap B} = \mathbb{1}_A \mathbb{1}_B = \mathbb{1}_B \mathbb{1}_A = \mathbb{1}_{B \cap A}$; also, $\cap$ is commutative. Hence it is enough to show left-distributivity. We have

$$\mathbb{1}_{A \cap (B \Delta C)} = \mathbb{1}_A \mathbb{1}_{B \Delta C} = \mathbb{1}_A (\mathbb{1}_B + \mathbb{1}_C - 2 \mathbb{1}_B \mathbb{1}_C) = \mathbb{1}_A \mathbb{1}_B + \mathbb{1}_A \mathbb{1}_C - 2 \mathbb{1}_A \mathbb{1}_B \mathbb{1}_C$$

$$= \mathbb{1}_{A \cap B} + \mathbb{1}_{A \cap C} - 2 (\mathbb{1}_A \mathbb{1}_B)(\mathbb{1}_A \mathbb{1}_C) = \mathbb{1}_{A \cap B} + \mathbb{1}_{A \cap C} - 2 \mathbb{1}_{A \cap B} \mathbb{1}_{A \cap C} = \mathbb{1}_{(A \cap B) \Delta (A \cap C)}$$

Thus the distributive property holds. (The distributivity $A \Delta (B \cap C) = (A \Delta B) \cap (A \Delta C)$ does not even hold.)
Hence the claim is true. $\square$

4.

Let us consider an arbitrary collection of σ-algebrae $\{\mathcal{G}_\alpha \mid \alpha \in I\}$ on the same set Ω.

Claim: the intersection of σ-algebrae

$$\mathcal{G} := \bigcap_{\alpha \in I} \mathcal{G}_\alpha$$

is a σ-algebra.

Proof:

Direct proof:

Since $\Omega \in \mathcal{G}_\alpha \; \forall \alpha \in I$ ($\mathcal{G}_\alpha$'s are σ-algebrae), $\Omega \in \mathcal{G}$.

Let $A \in \mathcal{G}$; then $A \in \mathcal{G}_\alpha \; \forall \alpha \in I$, and as $\mathcal{G}_\alpha$'s are σ-algebrae,

$A^c = \Omega \setminus A \in \mathcal{G}_\alpha \ \forall \alpha \in I$, so $A^c \in \mathcal{G}$.

Let $A_1; A_2; \ldots \in \mathcal{G}$ (infinite sequence); then $A_1; A_2; \ldots \in \mathcal{G}_\alpha \forall \alpha \in I$ and thus, as $\mathcal{G}_\alpha$'s are σ-algebras, $\bigcup_{n=1}^{\infty} A_n \in \mathcal{G}_\alpha \ \forall \alpha \in I$ so that $\bigcup_{n=1}^{\infty} A_n \in \mathcal{G}$.

Hence $\mathcal{G}$ is a σ-algebra and the claim is true. $\square$

5.

About countable and uncountable sets:

a) Claim: in the blackboard represented as $[0,1]^2$ there is place for an uncountable amount of mutually non-intersecting zero symbols "O" (circles), where the circles can be also inside each other but they should not touch each other.

Proof:

Proof by example:

Let us assume the circles can have arbitrary size (if we demand some size for all circles, the claim is obviously false, as each circle needs finite amount (non-zero amount) of 2D Lebesgue measure to live, and the interiors of the circles would be disjoint). Then the collection of circles C_r:

$$C_r := \{(x,y) \in [0,1]^2 \mid (x-\tfrac{1}{2})^2 + (y-\tfrac{1}{2})^2 = r^2\}$$

where $0 < r < \frac{1}{2}$ proves the claim is true. $\square$

b) Claim: on the blackboard $[0,1]^2$ or on an infinite blackboard $\mathbb{R}^2$ there is place for at most a countable number of mutually non-intersecting "8"-symbols, or ∞-symbols if you like, where the symbols can contain each other but the boundaries of different curves cannot touch each other.

Proof:

Direct proof:

Let us have some configuration of 8-symbols in $\mathbb{R}^2$. Let us define, for $n \in \mathbb{N}$

$A_n = \{$Those 8-symbols that are contained in $[-n,n]^2$ and

whose diameter is more than $\frac{1}{n}\}$

(where the diameter of an 8 is 8 diameter).

Now if we define an 8 to be 2 similar circles intersecting at one point, we see that, if we have an 8 whose diameter is greater than $\frac{1}{n}$, we cannot put another 8 whose diameter is greater than $\frac{1}{n}$ inside it; actually the whole interior of the afore mentioned 8 must be disjoint from the other 8's (whose diameter is more than $\frac{1}{n}$) and their interiors. Thus, for every 8 that is in A_n, we find $2 \cdot \pi(\frac{1}{4}\frac{1}{n})^2$ of 2D-Lebesgue measure that "belongs" to this 8 and this 8 only. Thus, as $[-n,n]^2$ has finite 2D Lebesgue measure, A_n must be finite (A_n contains finite number of 8's).

Now, $\bigcup_{n=1}^{\infty} A_n$ contains all the 8's drawn in the plane $\mathbb{R}^2$. As $\bigcup_{n=1}^{\infty} A_n$ is a countable union of finite sets, it is countable.

This proves that there are at most a countable number of 8's in $\mathbb{R}^2$. For $[0,1]^2 \subseteq \mathbb{R}^2$ the claim follows trivially.

Hence the claim is true. $\square$

PROBLEMS 2.

1.

A finitely additive probability $\mathbb{P}$ on a probability space Ω equipped with a σ-algebra $\mathcal{F}$ is also σ-additive if and only if for any event sequence $(A_n : n \in \mathbb{N}) \subseteq \mathcal{F}$ with $A_n \downarrow \emptyset$, meaning that $\forall n \in \mathbb{N}: A_{n+1} \subseteq A_n$ and $\bigcap_{n \in \mathbb{N}} A_n = \emptyset$, it follows that $\mathbb{P}(A_n) \downarrow 0$.

This does not hold for infinite measures with $\lambda(\Omega) = \infty$, for example for the Lebesgue measure λ on $\mathbb{R}$ equipped with the Borel σ-algebra $\mathcal{B}(\mathbb{R})$, such that $\lambda(]a,b]) = (b-a)^+$. Here $x^+ = \max\{0, x\}$ is the notation for the positive part of $x \in \mathbb{R}$.

Let us find a counterexample as a sequence $(A_n : n \in \mathbb{N}) \subseteq \mathcal{B}(\mathbb{R})$ with $A_n \downarrow \emptyset$ but $\lambda(A_n) \not\downarrow 0$.

We choose $A_n =]n, \infty[\in \mathcal{B}(\mathbb{R})$. Now $\forall n \in \mathbb{N}: A_{n+1} \subseteq A_n$ and $\bigcap_{n \in \mathbb{N}} A_n = \emptyset$ so we have $A_n \downarrow \emptyset$. But also $\lambda(A_n) = \infty$ $\forall n \in \mathbb{N}$, so that $\lambda(A_n) \not\downarrow 0$.

2.

Let $\Omega = \mathbb{R}^d$, the Euclidean space. In general the Borel σ-algebra is the smallest σ-algebra containing the open sets.

For $t \in \mathbb{R}^d$, we introduce the infinite rectangle

$$]-\infty, t] = \{s \in \mathbb{R}^d \mid s_i \le t_i,\ i = 1, \dots, d\}.$$

Claim: the class $\mathcal{I} = \{]-\infty, q] \mid q \in \mathbb{Q}^d\}$ is a π-class with $\sigma(\mathcal{I}) = \mathcal{B}(\mathbb{R}^d)$.

Proof:

Direct proof:

Obviously $\mathcal{I} \neq \emptyset$. Also, for $]-\infty, q],]-\infty, q'] \in \mathcal{I}$ we have

$$]-\infty, q] \cap]-\infty, q'] =]-\infty, \min\{q; q'\}] \in \mathcal{I}$$

where $\min\{q; q'\} = (\min\{q_1; q_1'\}; \dots; \min\{q_d; q_d'\}) \in \mathbb{Q}^d$.

So $\mathcal{I}$ is a π-system.

Also, as the sets $]-\infty,q]\in I$ are closed, also their complements are open sets, so by defining properties of σ-algebra, $I\subseteq B(\mathbb{R}^d)$, so $\sigma(I)\subseteq B(\mathbb{R}^d)$.

Then, let U be an open set in $\mathbb{R}^d$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, $\forall x\in U\ \exists r,q\in\mathbb{Q}^d$ such that $r<q$ (meaning that $r_i<q_i$ for each coordinate $i=1,\dots,d$) and

$$x\in]r,q[:=]r_1;q_1[\times]r_2;q_2[\times\cdots\times]r_d;q_d[\subseteq U$$

id est, there is a small open rectangle containing x which is contained in U.

Now we see that

$$U=\bigcup_{x\in U}]r(x);q(x)[$$

for choices of $r(x);q(x)$ for $x\in U$ as above. As $\mathbb{Q}$ is countable, so is $\mathbb{Q}^{2d}$, so that in fact U can be expressed as a countable union of rectangles $]r,q[$ with $r,q\in\mathbb{Q}^d$.

So, to prove that $U\in\sigma(I)$, and hence $B(\mathbb{R}^d)\subseteq\sigma(I)$ and thus $\sigma(I)=B(\mathbb{R}^d)$, it is enough to prove that rectangles $]r,q[\in\sigma(I)$, where $r,q\in\mathbb{Q}^d$.

But this is clear as

$$]r,q[=\Big(\bigcup_{n\in\mathbb{N}}]-\infty,(1-\tfrac{1}{n})q]\Big)\setminus\Big(]-\infty,(r_1;q_2;\cdots;q_d)]\cup]-\infty,(q_1;r_2;q_3;\cdots;q_d)]\cup\dots\cup]-\infty;(q_1;\cdots;q_{d-1};r_d)]\Big)$$

Hence $\sigma(I)=B(\mathbb{R}^d)$.

Thus the claim is true. $\square$

3.

Consider a probability triple $(\Omega;\mathcal{F};\mathbb{P})$ and a sequence of events $A_n\in\mathcal{F}$ such that $\mathbb{P}(A_n)=1\ \forall n\in\mathbb{N}$.

Claim: $\mathbb{P}(\bigcap_{n\in\mathbb{N}} A_n)=1$

Proof:

Direct proof:

$$1-\mathbb{P}(\bigcap_{n\in\mathbb{N}} A_n)=\mathbb{P}((\bigcap_{n\in\mathbb{N}} A_n)^c)=\mathbb{P}(\bigcup_{n\in\mathbb{N}} A_n^c)\leq\sum_{n\in\mathbb{N}}\mathbb{P}(A_n^c)=\sum_{n\in\mathbb{N}}(1-\mathbb{P}(A_n))$$

$$=\sum_{n\in\mathbb{N}}(1-1)=0$$

$$\Rightarrow \mathbb{P}(\bigcap_{n\in\mathbb{N}} A_n)=1.$$

Hence the claim is true. $\square$

Consider also a sequence $B_n\in\mathcal{F}$ such that $\mathbb{P}(B_n)=0\ \forall n\in\mathbb{N}$.

Claim: $\mathbb{P}(\bigcup_{n\in\mathbb{N}} B_n)=0$

Proof:

Direct proof:

$$\mathbb{P}(\bigcup_{n\in\mathbb{N}} B_n)\leq\sum_{n\in\mathbb{N}}\mathbb{P}(B_n)=\sum_{n\in\mathbb{N}}0=0$$

$$\Rightarrow \mathbb{P}(\bigcup_{n\in\mathbb{N}} B_n)=0$$

Hence the claim is true. $\square$

5.

Consider the probability space $(\Omega,\mathcal{F})$ and an event sequence $(A_n:n\in\mathbb{N})\subseteq\mathcal{F}$. We denote:

$$\limsup_{n\to\infty} A_n:=\bigcap_{k\in\mathbb{N}}\bigcup_{n\geq k} A_n;\qquad \liminf_{n\to\infty} A_n:=\bigcup_{k\in\mathbb{N}}\bigcap_{n\geq k} A_n$$

a) Claim: $\limsup_{n\to\infty} A_n\in\mathcal{F}$.

Proof:

Direct proof:

For every $k\in\mathbb{N}$, $\bigcup_{n\geq k} A_n\in\mathcal{F}$. Thus $(\bigcup_{n\geq k} A_n)^c\in\mathcal{F}$, and thus

$$\limsup_{n\to\infty} A_n=\bigcap_{k\in\mathbb{N}}\bigcup_{n\geq k} A_n=(\bigcup_{k\in\mathbb{N}}(\bigcup_{n\geq k} A_n)^c)^c\in\mathcal{F}.$$

Hence the claim is true. $\square$

b) Claim: $\liminf_{n\to\infty} A_n = \left(\limsup_{n\to\infty} A_n^c\right)^c$, where $B^c = \Omega \setminus B$ is the complement event.

Proof:

Direct proof:

$\left(\limsup_{n\to\infty} A_n^c\right)^c = \left(\bigcap_{k\in\mathbb{N}} \bigcup_{n\geq k} A_n^c\right)^c = \bigcup_{k\in\mathbb{N}} \left(\bigcup_{n\geq k} A_n^c\right)^c = \bigcup_{k\in\mathbb{N}} \bigcap_{n\geq k} (A_n^c)^c$

$= \bigcup_{k\in\mathbb{N}} \bigcap_{n\geq k} A_n = \liminf_{n\to\infty} A_n$

Hence the claim is true. $\square$

c) Claim: $\liminf_{n\to\infty} A_n \in \mathcal{F}$.

Proof:

Direct proof:

By b)-part, $\liminf_{n\to\infty} A_n = \left(\limsup_{n\to\infty} A_n^c\right)^c$. By a)-part, $\liminf_{n\to\infty} A_n \in \mathcal{F}$.

Thus the claim is true. $\square$

d) For $A \in \mathcal{F}$, let $\mathbb{1}_A(\omega)$ be the indicator function of A, defined for $\omega \in \Omega$.

Claim: $\mathbb{1}_{\limsup_{n\to\infty} A_n}(\omega) = \limsup_{n\to\infty} \mathbb{1}_{A_n}(\omega)$; $\mathbb{1}_{\liminf_{n\to\infty} A_n}(\omega) = \liminf_{n\to\infty} \mathbb{1}_{A_n}(\omega)$

Proof:

Direct proof:

Obviously $\mathbb{1}_{\limsup_{n\to\infty} A_n}(\omega)$; $\limsup_{n\to\infty} \mathbb{1}_{A_n}(\omega)$; $\mathbb{1}_{\liminf_{n\to\infty} A_n}(\omega)$; $\liminf_{n\to\infty} \mathbb{1}_{A_n}(\omega)$ $\in \{0,1\}$. We have

$\mathbb{1}_{\limsup_{n\to\infty} A_n}(\omega) = 1 \iff \omega \in \limsup_{n\to\infty} A_n = \bigcap_{k\in\mathbb{N}} \bigcup_{n\geq k} A_n \iff \forall k \in \mathbb{N}\ \exists n \geq k : \omega \in A_n$

$\iff \forall k \in \mathbb{N} : \sup_{n\geq k} \mathbb{1}_{A_n}(\omega) = 1 \Rightarrow \limsup_{n\to\infty} \mathbb{1}_{A_n}(\omega) = 1.$

Also, if $\limsup_{n\to\infty} \mathbb{1}_{A_n}(\omega)=1$, since $\forall k\in\mathbb{N}: \sup_{n\geq k} \mathbb{1}_{A_n}(\omega)\in\{0,1\}$,

we have, for some $k'\in\mathbb{N}$, $\forall k\geq k'$: $\sup_{n\geq k} \mathbb{1}_{A_n}(\omega)=1$. But thus

$\sup_{n\geq k} \mathbb{1}_{A_n}(\omega)=1$ $\forall k\in\mathbb{N}$, so $\mathbb{1}_{\limsup_{n\to\infty} A_n}(\omega)=1$. So

$\mathbb{1}_{\limsup_{n\to\infty} A_n}(\omega)=1 \Leftrightarrow \limsup_{n\to\infty} \mathbb{1}_{A_n}(\omega)=1$

So $\mathbb{1}_{\limsup_{n\to\infty} A_n}(\omega) = \limsup_{n\to\infty} \mathbb{1}_{A_n}(\omega)$.

On the other hand,

$$\mathbb{1}_{\liminf_{n\to\infty} A_n}(\omega) = \mathbb{1}_{(\limsup_{n\to\infty} A_n^c)^c}(\omega) = 1-\mathbb{1}_{\limsup_{n\to\infty} A_n^c}(\omega) = 1-\limsup_{n\to\infty} \mathbb{1}_{A_n^c}(\omega)$$

$$=1-\limsup_{n\to\infty}(1-\mathbb{1}_{A_n}(\omega)) = 1-(1-\liminf_{n\to\infty} \mathbb{1}_{A_n}(\omega)) = \liminf_{n\to\infty} \mathbb{1}_{A_n}(\omega)$$

Hence the claim is true. $\square$

e) Claim: $\limsup_{n\to\infty} A_n = \{\omega\in\Omega \mid \omega\in A_n$ infinitely often, which means for infinitely many $n\}$; $\liminf_{n\to\infty} A_n = \{\omega\in\Omega \mid \omega\in A_n$ eventually, which means for all n large enough$\}$.

Proof:

Direct proof:

We see that

$\omega\in \limsup_{n\to\infty} A_n = \bigcap_{k\in\mathbb{N}} \bigcup_{n\geq k} A_n \Leftrightarrow \forall k\in\mathbb{N}\,\exists n\geq k: \omega\in A_n \Leftrightarrow \omega\in\{\omega\in\Omega \mid$ $\omega\in A_n$ infinitely often, which means for infinitely many $n\}$

$\omega\in \liminf_{n\to\infty} A_n = \bigcup_{k\in\mathbb{N}} \bigcap_{n\geq k} A_n \Leftrightarrow \exists k\in\mathbb{N}: \forall n\geq k: \omega\in A_n \Leftrightarrow \omega\in\{\omega\in\Omega \mid \omega\in A_n$ eventually, which means for all n large enough$\}$

Thus the claim is true. $\square$

f) Let $A_n \subseteq B_n \ \forall n$.

Claim: $\limsup_{n\to\infty} A_n \subseteq \limsup_{n\to\infty} B_n$; $\liminf_{n\to\infty} A_n \subseteq \liminf_{n\to\infty} B_n$

Proof:

Direct proof:

We see that $A_n \subseteq B_n \ \forall n \iff \mathbb{1}_{A_n} \leq \mathbb{1}_{B_n} \ \forall n$. Thus, by d)-part

$$\mathbb{1}_{\limsup_{n\to\infty} A_n} = \limsup_{n\to\infty} \mathbb{1}_{A_n} \leq \limsup_{n\to\infty} \mathbb{1}_{B_n} = \mathbb{1}_{\limsup_{n\to\infty} B_n}$$

$$\mathbb{1}_{\liminf_{n\to\infty} A_n} = \liminf_{n\to\infty} \mathbb{1}_{A_n} \leq \liminf_{n\to\infty} \mathbb{1}_{B_n} = \mathbb{1}_{\liminf_{n\to\infty} B_n}$$

Thus we see $\limsup_{n\to\infty} A_n \subseteq \limsup_{n\to\infty} B_n$ and $\liminf_{n\to\infty} A_n \subseteq \liminf_{n\to\infty} B_n$; hence the claim is true. $\square$

g) Suppose that for a probability $\mathbb{P}$ we have $\mathbb{P}(\limsup_{n\to\infty} A_n) = 1$ and $\mathbb{P}(\liminf_{n\to\infty} B_n) = 1$.

Claim: $\mathbb{P}(\limsup_{n\to\infty} A_n \cap B_n) = 1$.

Proof:

Direct proof: by d)-part above

$$\mathbb{1}_{\limsup_{n\to\infty} A_n \cap B_n} = \limsup_{n\to\infty} \mathbb{1}_{A_n \cap B_n} = \limsup_{n\to\infty} \mathbb{1}_{A_n} \mathbb{1}_{B_n} \geq \limsup_{n\to\infty} \mathbb{1}_{A_n} \cdot \liminf_{n\to\infty} \mathbb{1}_{B_n} = \mathbb{1}_{(\limsup_{n\to\infty} A_n) \cap (\liminf_{n\to\infty} B_n)}$$

$$\Rightarrow \mathbb{P}(\limsup_{n\to\infty} A_n \cap B_n) \geq \mathbb{P}((\limsup_{n\to\infty} A_n) \cap (\liminf_{n\to\infty} B_n)) = 1$$

by exercise 3 first part.

Thus $\mathbb{P}(\limsup_{n\to\infty} A_n \cap B_n) = 1$.

Hence the claim is true. $\square$

5.

Let $(\Omega, \mathcal{F})$ be a probability space, with a sequence of probability measures $(P_n : n \in \mathbb{N})$.

Suppose that $\forall A \in \mathcal{F}$ the limits

$$P(A) := \lim_{n\to\infty} P_n(A)$$

exist.

a) Claim: the map $P : \mathcal{F} \to [0,1]$ is a probability measure.

Proof: (Continued $\Rightarrow$)

Direct proof:

By assumptions (P_n are probability measures) it is clear that $P : \mathcal{F} \to [0,1]$ is a well-defined function.

We have

$$P(\emptyset) = \lim_{n\to\infty} P_n(\emptyset) = \lim_{n\to\infty} 0 = 0$$

$$P(\Omega) = \lim_{n\to\infty} P_n(\Omega) = \lim_{n\to\infty} 1 = 1$$

and, for $A_1; \dots; A_K \in \mathcal{F}$ pairwise disjoint

$$P\left(\bigcup_{k=1}^{K} A_k\right) = \lim_{n\to\infty} P_n\left(\bigcup_{k=1}^{K} A_k\right) = \lim_{n\to\infty} \sum_{k=1}^{K} P_n(A_k) = \sum_{k=1}^{K} \lim_{n\to\infty} P_n(A_k) = \sum_{k=1}^{K} P(A_k)$$

so P is finitely additive. Let $A_1; A_2; \dots \in \mathcal{F}$ be pairwise disjoint; set $A := \bigcup_{k=1}^{\infty} A_k$. We have $P(A \cup A^c) = P(A) + P(A^c) = P(\Omega) = 1$. We also have $\alpha := P(A^c) + \sum_{k=1}^{\infty} P(A_k) \leq P(\Omega) = 1$ (all partial sums $\leq P(\Omega)$ by finite additivity). If $\alpha = 1$, we have $P(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k)$, also σ-additivity. Assume towards a contradiction that $\alpha < 1$. Let $A_0 = A^c$; $a_{n,k} = \frac{1}{1-\alpha}(P_n(A_k) - P(A_k))$. Now $\sum_{k=0}^{\infty} a_{n,k} = \frac{1}{1-\alpha}(1-\alpha) = 1$ $\forall n$, $\sum_{k=0}^{\infty} |a_{n,k}| \leq \frac{1}{|1-\alpha|}\left(\sum_{k=0}^{\infty} P_n(A_k) + \sum_{k=0}^{\infty} P(A_k)\right) = \frac{2}{|1-\alpha|} < \infty$ $\forall n$, and $a_{n,k} \to 0$ $\forall k$. Thus by Steinhaus' lemma there exist $(x_k)_{k\in\mathbb{N}_0} \in \{0,1\}^{\mathbb{N}_0}$ so that $t_n = \sum_{k=1}^{\infty} a_{n,k} x_k$ does not converge to any limit. Thus if we let $(A'_\ell)_{\ell=0,1,\dots}$ be the sequence of A_k's with $x_k = 1$, we get that $\frac{1}{1-\alpha}\sum_{\ell=0}^{\infty}(P_n(A'_\ell) - P(A'_\ell))$ does not converge, but $P(\bigcup_\ell A'_\ell) = \lim_{n\to\infty} \sum_{\ell=0}^{\infty} P_n(A'_\ell)$ by assumption that the convergence happens for all $A \in \mathcal{F}$, and $0 \leq \sum_{\ell=0}^{\infty} P(A'_\ell) \leq P(\Omega) = 1$ so $\frac{1}{1-\alpha}\sum_{\ell=0}^{\infty}(P_n(A'_\ell) - P(A'_\ell)) = \frac{1}{1-\alpha}\left(\sum_{\ell=0}^{\infty} P_n(A'_\ell) - \sum_{\ell=0}^{\infty} P(A'_\ell)\right)$ should converge $\lightning$. So we must have $\alpha = 1$ and P is σ-additive. Thus P is a probability measure. Hence the claim is see. $\square$

b) Claim: For each event sequence $(A_k : k \in \mathbb{N}) \subseteq \mathcal{F}$ such that $A_k \downarrow \emptyset$ we have

$$\sup_{n \in \mathbb{N}} P_n(A_k) \downarrow 0 \text{ as } k \to \infty.$$

Proof:

Direct proof:

Because $A_k \supseteq A_{k+1}$, $P_n(A_k) \geq P_n(A_{k+1})$. Thus it is clear that

$\sup_{n \in \mathbb{N}} P_n(A_k) \geq \sup_{n \in \mathbb{N}} P_n(A_{k+1})$, so if the sequence $\sup_{n \in \mathbb{N}} P_n(A_k)$ converges

to 0, the convergence is monotone from above.

Also, as $\sup_{n \in \mathbb{N}} P_n(A_k)$ is a monotonous decreasing bounded

(is in interval $[0,1]$) sequence of real numbers, it

converges to $L \in [0,1]$ say.

Let us assume towards a contradiction that $L > 0$. Now, as P is a probability measure (a)-part), by exercise 1 we have $P(A_k) \downarrow 0$. Thus choose $k' \in \mathbb{N}$: $P(A_k) \leq \frac{L}{2}$ $\forall k \geq k'$. Now as $P(A_{k'}) = \lim_{n \to \infty} P_n(A_{k'})$ choose $n' \in \mathbb{N}$: $|P_n(A_{k'}) - P(A_{k'})| < \frac{L}{4}$ $\forall n \geq n'$.

Then we have $\sup_{n \in \mathbb{N}} P_n(A_{k'}) \leq \max\{P_1(A_{k'}); \ldots; P_{n'}(A_{k'}); \frac{3}{4}L\}$

Now as we have the monotonicity $P_n(A_k) \geq P_n(A_{k+1})$ $\forall n, k \in \mathbb{N}$, we see that $P_n(A_k) \leq P_n(A_{k'}) \leq \frac{3}{4}L$ for all $n \geq n'$, $k \geq k'$.

Thus, we can write $\sup_{n \in \mathbb{N}} P_n(A_k) \leq \max\{P_1(A_k); \ldots; P_{n'}(A_k); \frac{3}{4}L\}$

for all $k \geq k'$. Now, as the P_n's are also probability measures, by exercise 1 we have $P_n(A_k) \downarrow 0$ as $k \to \infty$ for all $n \in \mathbb{N}$; thus for k big enough

$$\sup_{n \in \mathbb{N}} P_n(A_k) \leq \max\{P_1(A_k); \ldots; P_{n'}(A_k); \tfrac{3}{4}L\} \leq \tfrac{3}{4}L$$

But $\sup_{n \in \mathbb{N}} P_n(A_k) \downarrow L$ as $k \to \infty$; as $L > 0$, we have the contradiction

$$L=\lim_{k\to\infty}\sup_{n\in\mathbb{N}} P_n(A_k)\le\frac{3}{4}L$$

So the assumption $L>0$ leads to a contradiction; thus we must have

$$\lim_{k\to\infty}\sup_{n\in\mathbb{N}} P_n(A_k)=L=0.$$

Hence the claim is true. □

⇐ Continuation of part a).

Let us prove the Steinhaus' lemma.

Let $(a_{n,k})_{n,k\in\mathbb{N}_0}\subseteq\mathbb{R}$. Let

(1) $\sum_{k=0}^{\infty} a_{n,k}=1\quad\forall n$

(2) $\sum_{k=0}^{\infty} |a_{n,k}|\le C<\infty\quad\forall n$

(3) $a_{n,k}\xrightarrow[n\to\infty]{}0\quad\forall k$

Then $\exists(x_n)_{n\in\mathbb{N}_0}\in\{0,1\}^{\mathbb{N}_0}$ such that $t_n=\sum_{k=0}^{\infty}a_{n,k}x_k$ fails to converge to a finite or infinite limit.

Let $n_0, k_0\in\mathbb{N}_0$ be arbitrary. Choose inductively $n_1; k_1$ and $n_2; k_2$ and so on as follows: having chosen $n_1;\dots;n_r; k_1;\dots;k_r$ choose $n_{r+1}>n_r$ so that $\sum_{k\le k_r}|a_{n_{r+1},k}|<\frac{1}{8}$; by (3) this choice is possible.

Then choose $k_{r+1}>k_r$ such that $\sum_{k>k_{r+1}}|a_{n_{r+1},k}|<\frac{1}{8}$; this is possible by (2). Continue to find n_j's, k_j's.

Now define $x_k=0$ for $k_{2s-1}<k\le k_{2s}$ and $x_k=1$, $k_{2s}<k\le k_{2s+1}$; $s=0,1,2,\dots$ where $k_{-1}:=0$.
Then note that $t_{n_{r+1}}=h_1+h_2+h_3$ where $h_1=\sum_{k\le k_r}a_{n_{r+1},k}x_k$; $h_2=\sum_{k_r<k\le k_{r+1}}a_{n_{r+1},k}x_k$
$h_3=\sum_{k>k_{r+1}}a_{n_{r+1},k}x_k$

If r is odd here, then $x_k=0$ for $k_r<k\le k_{r+1}$ and thus $h_2=0$ and $|t_{n_{r+1}}|<|h_1|+|h_3|<\frac{1}{8}+\frac{1}{8}=\frac{1}{4}$.

If r is even, $x_k=1$ for $k_r<k\le k_{r+1}$ and $h_2=\sum_{k_r<k\le k_{r+1}}a_{n_{r+1},k}$. Hence by (1) we get

$$h_2 = 1 - \sum_{k \in k_r} a_{n_{r+1}; k} - \sum_{k > k_{r+1}} a_{n_{r+1}; k} > 1 - \frac{1}{8} - \frac{1}{8} = \frac{3}{4}.$$

So $t_{n_{r+1}} > h_2 - |h_1| - |h_3| > \frac{3}{4} - \frac{1}{8} - \frac{1}{8} = \frac{1}{2}$.

Thus $(t_n)_{n \in \mathbb{N}_0}$ does not converge. $\square$

PROBLEMS 3.

On the cylinder algebra on an infinite product space

Let S be an abstract probability space equipped with a σ-algebra $\hat{S}$, for example $S=\mathbb{R}^d$ and $\hat{S}=B(\mathbb{R}^d)$, the Borel σ-algebra, and T an (infinite) arbitrary set. Consider the space $\Omega=S^T$, whose elements are the maps $\omega: T\to S$, with $t\mapsto \omega_t\in S$.

We can also understand Ω as the infinite product space $\Omega=\prod_{t\in T} S_t$, where each S_t is a copy of S.

A cylinder is an Ω-subset with representation

(1) $C=\{\omega\in\Omega \mid (\omega_{t_1};\omega_{t_2};\dots;\omega_{t_d})\in B_{t_1;\dots;t_d}\}$

for some $d\in\mathbb{N}$, $t_1;\dots;t_d\in T$ and $B_{t_1;\dots;t_d}\in\hat{S}^{\otimes d}=\underbrace{\hat{S}\otimes\hat{S}\otimes\cdots\otimes\hat{S}}_{d\text{ times}}$, the d-fold product of σ-algebrae. In other words, whether a function ω belongs to a cylinder C or not it is determined by its values on a finite number of coordinates.

Note that the cylinder representation (1) is not unique, for example the same cylinder C could be expressed as

$C=\{\omega\in\Omega \mid (\omega_{t_1};\omega_{t_2};\dots;\omega_{t_d};\omega_{t_{d+1}})\in B_{t_1;\dots;t_d}\times S\}$

1.

Claim: the cylinders $\hat{C}=\{C\subseteq\Omega \mid C \text{ is a cylinder}\}$ form an algebra of Ω-events.

Proof:

Direct proof:

Obviously $\hat{C}\subseteq 2^\Omega$.

Obviously $\Omega\in\hat{C}$ ($d=1$, $t_1\in T$; $B_{t_1}=S$).

Let $C\in\hat{C}$. Let $C=\{\omega\in\Omega \mid (\omega_{t_1};\omega_{t_2};\dots;\omega_{t_d})\in B_{t_1;\dots;t_d}\}$. Then

$\Omega\setminus C=:C^c=\{\omega\in\Omega \mid (\omega_{t_1};\dots;\omega_{t_d})\in S^d\setminus B_{t_1;\dots;t_d}\}$ where $S^d\setminus B_{t_1;\dots;t_d}\in\hat{S}^{\otimes d}$

Thus $C^c \in \hat{C}$.

Let $A, B \in \hat{C}$. Let $A = \{\omega \in \Omega \mid (\omega_{t_1}; \dots; \omega_{t_d}) \in A_{t_1;\dots;t_d}\}$ and $B = \{\omega \in \Omega \mid (\omega_{s_1}; \dots; \omega_{s_{d'}}) \in B_{s_1;\dots;s_{d'}}\}$ with $d, d' \in \mathbb{N}$ and $t_1;\dots;t_d;$ $s_1;\dots;s_{d'} \in T$, $A_{t_1;\dots;t_d} \in \hat{S}^{\otimes d}$, $B_{s_1;\dots;s_{d'}} \in \hat{S}^{\otimes d'}$. Now let

$C_{t_1;\dots;t_d;s_1;\dots;s_{d'}} = A_{t_1;\dots;t_d} \times S^{d'} \cup S^d \times B_{s_1;\dots;s_{d'}} \in \hat{S}^{\otimes(d+d')}$

Then $A \cup B = \{\omega \in \Omega \mid (\omega_{t_1}; \dots; \omega_{t_d}; \omega_{s_1}; \dots; \omega_{s_{d'}}) \in C_{t_1;\dots;t_d;s_1;\dots;s_{d'}}\}$.

Thus $A \cup B \in \hat{C}$.

Hence $\hat{C}$ is an algebra. Whether $\hat{C}$ is an algebra of Ω-events depends on what σ-algebra of events we put on Ω. If we put the standard σ-algebra generated by cylinders (also $\sigma(\hat{C})$) then $\hat{C}$ is an algebra of Ω-events.

Hence the claim is true. $\square$

2. However, the cylinders do not form a σ-algebra when T is infinite. Let us find an example where the countable intersection of cylinders is not a cylinder.

Let us choose $S = \mathbb{R}$, $\hat{S} = B(\mathbb{R})$, $T = \mathbb{N}$ and $\Omega = S^T = \mathbb{R}^{\mathbb{N}}$. Let us define the cylinders

$C_n := \{\omega \in \Omega \mid (\omega_n) \in \{0\}\}$ for $n \in \mathbb{N}$ (here $d = 1$, $t_1 = t_d = n \in T = \mathbb{N}$,

$B_{t_1} = \{0\} \in \hat{S}^{\otimes d} = \hat{S}^{\otimes 1} = \hat{S} = B(\mathbb{R})$)

Now $\bigcap_{n=1}^{\infty} C_n = \{(0,0,0,\dots)\} \subseteq \Omega$, but $\{(0,0,0,\dots)\}$ is not a cylinder as the values of a finite number of coordinates of $\omega \in \Omega$ do not determine whether $\omega \in \{(0,0,0,\dots)\}$, id est, whether $\omega = (0,0,0,\dots)$.

A consistent family P of finite dimensional distributions is a collection of probability measures $P_{t_1;\dots;t_d}$ on the respective product σ-algebrae $\hat{S}^{\otimes d}$ indexed by $t_1;\dots;t_d \in T$, where d varies in $\mathbb{N}$, satisfying the properties:

(*) $P_{t_1;\dots;t_d}(B_{t_1} \times \dots \times B_{t_d}) = P_{t_{\pi(1)};\dots;t_{\pi(d)}}(B_{t_{\pi(1)}} \times \dots \times B_{t_{\pi(d)}})$

for every d, $t_1;\dots;t_d \in T$ and π permutation of $\{1,2,\dots,d\}$, and $B_{t_i} \in \hat{S}$,

$P_{t_1;\dots;t_d}(B_{t_1;\dots;t_d}) = P_{t_1;\dots;t_d;t_{d+1}}(B_{t_1;\dots;t_d} \times S)$

$\forall d \in \mathbb{N}\ \forall t_1;\dots;t_d;t_{d+1} \in T$ and $B_{t_1;\dots;t_d} \in \hat{S}^{\otimes d}$.

3.

Claim: the map $P_0 : \hat{C} \to [0,1]$ with $P_0(C) = P_{t_1;\dots;t_d}(B_{t_1;\dots;t_d})$ for cylinders C with representation (1) is well defined, meaning that it does not depend on the particular representation of the cylinder C, and that P_0 is finitely additive on the algebra $\hat{C}$.

Proof: As $C = \Omega\ (=\emptyset)$ can only be represented by $B_{t_1;\dots;t_d} = S^d\ (=\emptyset)$, we have $P_0(C) = P_{t_1;\dots;t_d}(B_{t_1;\dots;t_d}) = 1\ (=0)$ well-defined, so we ignore the special case $C = \Omega\ (=\emptyset)$

Direct proof:

Let $C \in \hat{C}$, and let d' be the minimal $d \in \mathbb{N}_0$ such that C has a representation (1) with $B_{t_1;\dots;t_d} \in \hat{S}^{\otimes d}$ (a non-empty set of natural numbers has a least element). We agree that the case $d=0$ means $C = \Omega$ or $C = \emptyset$.

Now, let $B_{t'_1;\dots;t'_{d'}} \in \hat{S}^{\otimes d'}$ be such that it can be used in representation $(1)^d$ of C; $t'_1;\dots;t'_{d'} \in T$.

Now, all the alternative representations of type (1) of C are such that

$C = \{\omega \in \Omega \mid (\omega_{t_1};\dots;\omega_{t_d}) \in B_{t_1;\dots;t_d}\}$

where $d \geq d'$ and $\{t'_1;\dots;t'_{d'}\} \subseteq \{t_1;\dots;t_d\}$, for if say $t'_j \notin \{t_1;\dots;t_d\}$, then for $\omega \in \Omega$:

(2) $(\omega_{t'_1};\dots;\omega_{t'_{d'}}) \in B_{t'_1;\dots;t'_{d'}} \Leftrightarrow \omega \in C \Leftrightarrow (\omega_{t_1};\dots;\omega_{t_d}) \in B_{t_1;\dots;t_d}$

Now let $\omega \in \Omega$ be such that $(\omega_{t'_1};\dots;\omega_{t'_{d'}}) \in B_{t'_1;\dots;t'_{d'}}$.

Then by the above, $(\omega_{t_1};\dots;\omega_{t_d}) \in B_{t_1;\dots;t_d}$. Then let $a \in S$ and let $\omega^a \in \Omega$, $\omega^a_t := \begin{cases} \omega_t; t \neq t'_j \\ a; t = t'_j \end{cases}$. Now as $t'_j \notin \{t_1;\dots;t_d\}$,

$(\omega^a_{t_1};\dots;\omega^a_{t_d}) = (\omega_{t_1};\dots;\omega_{t_d}) \in B_{t_1;\dots;t_d}$. Thus by (2) $(\omega^a_{t'_1};\dots;$ $\omega^a_{t'_{d'}}) \in B_{t'_1;\dots;t'_{d'}}$. This holds for all $a \in S$, so that

$B_{t'_1;\dots;t'_{d'}} \cong Pr_{t'_1;\dots;t'_{j-1};t'_{j+1};\dots;t'_{d'}}(B_{t'_1;\dots;t'_{d'}}) \times S$ ($\cong$ means bijective correspondence)

where Pr are coordinate projections. As the projections are measurable sets, we find a representation of type (1) for C with set $Pr_{t'_1;\dots;t'_{j-1};t'_{j+1};\dots;t'_{d'}}(B_{t'_1;\dots;t'_{d'}})$

which shows that d' is not minimal $\lightning$.

Thus we have $\{t'_1;\dots;t'_{d'}\} \subseteq \{t_1;\dots;t_d\}$.

Now obviously we have

$Pr_{t'_1;\dots;t'_{d'}}(B_{t_1;\dots;t_d}) = B_{t'_1;\dots;t'_{d'}}$

because of (2). Also, because of (2), since $B_{t'_1;\dots;t'_{d'}}$ and $B_{t_1;\dots;t_d}$ characterize the same set C, there exist a permutation π of $\{1,\dots,d\}$ such that

$\pi(B_{t_1;\dots;t_d}) = B_{t'_1;\dots;t'_{d'}} \times S^{d-d'}$ and $(t_1;\dots;t_d) \overset{\pi}{\mapsto} (t'_1;\dots;t'_{d'};\tilde{t}_1;\dots;\tilde{t}_{d-d'})$

Now, from the property $(*)$ it follows that

$P_{\tilde{t}_1;\dots;\tilde{t}_{\tilde{d}}}(B) = P_{\tilde{t}_{\tilde{\pi}(1)};\dots;\tilde{t}_{\tilde{\pi}(\tilde{d})}}(\tilde{\pi}(B))$

for all $\tilde{d} \in \mathbb{N}$, $\tilde{\pi} \in S_{\tilde{d}}$, $B \in \mathcal{S}^{\otimes \tilde{d}}$ (as the products $B_1 \times \dots \times B_d$ form a π-system spanning $\mathcal{S}^{\otimes d}$).

Thus finally

$P_0(C) = P_{t_1;\dots;t_d}(B_{t_1;\dots;t_d}) = P_{t_{\pi(1)};\dots;t_{\pi(d)}}(\pi(B_{t_1;\dots;t_d})) = P_{t'_1;\dots;t'_{d'}}$

$= P_{t'_1,\dots,t'_{d'},\tilde{t}_1,\dots,\tilde{t}_{d-d'}}(B_{t'_1,\dots,t'_{d'}} \times S^{d-d'}) = P_{t'_1,\dots,t'_{d'}}(B_{t'_1,\dots,t'_{d'}}) \in [0,1]$

As this holds for all $B_{t_1,\dots,t_d}$ we see that $P_0(C)$ is well-defined.

So $P_0 : \hat{C} \to [0,1]$ is a well-defined mapping.

Let $A, B \in \hat{C}$, and let $A_{t_1,\dots,t_d}$ and $B_{s_1,\dots,s_{d'}}$ be as in exercise 1. Let $A \cap B = \emptyset$.

Now $A \cup B = \{\omega \in \Omega \mid (\omega_{t_1},\dots,\omega_{t_d},\omega_{s_1},\dots,\omega_{s_{d'}}) \in C_{t_1,\dots,t_d,s_1,\dots,s_{d'}}\}$

where $C_{t_1,\dots,t_d,s_1,\dots,s_{d'}} = A_{t_1,\dots,t_d} \times S^{d'} \sqcup S^d \times B_{s_1,\dots,s_{d'}}$ and the

union is disjoint because $A \cap B = \emptyset$. Thus

$P_0(A \cup B) = P_{t_1,\dots,t_d,s_1,\dots,s_{d'}}(C_{t_1,\dots,t_d,s_1,\dots,s_{d'}}) = P_{t_1,\dots,t_d,s_1,\dots,s_{d'}}$
$(A_{t_1,\dots,t_d} \times S^{d'}) + P_{t_1,\dots,t_d,s_1,\dots,s_{d'}}(S^d \times B_{s_1,\dots,s_{d'}}) = P_{t_1,\dots,t_d}(A_{t_1,\dots,t_d})$
$+ P_{s_1,\dots,s_{d'}}(B_{s_1,\dots,s_{d'}}) = P_0(A) + P_0(B)$

as $P_{t_1,\dots,t_d,s_1,\dots,s_{d'}}$ is a probability measure.

Thus P_0 is finitely additive in $\hat{C}$.

For each t, let Q_t be a probability on $(S, \hat{S})$.

Define the family $\hat{Q}$ of finite dimensional distributions

$Q_{t_1,\dots,t_d} = Q_{t_1} \otimes Q_{t_2} \otimes \cdots \otimes Q_{t_d}$

as the product measure on the product space S^d equipped with product σ-algebra $\hat{S}^{\otimes d}$.

4.

Claim: $\hat{Q}$ is a consistent family of finite dimensional distributions.

Proof:

Direct proof:

$\hat{Q}$ is obviously a suitably indexed collection of probability measures.

Let $d \in \mathbb{N}$, $t_1, \dots, t_d, t_{d+1} \in T$.

Let $\pi \in S_d$ and let $B_{t_1}, \dots, B_{t_d} \in \hat{S}$. Let $B_{t_1, \dots, t_d} \in \hat{S}^{\otimes d}$

Now

$$Q_{t_1, \dots, t_d}(B_{t_1} \times \dots \times B_{t_d}) = Q_{t_1} \otimes \dots \otimes Q_{t_d}(B_{t_1} \times \dots \times B_{t_d}) = Q_{t_1}(B_{t_1}) \dots Q_{t_d}(B_{t_d}) = Q_{t_{\pi(1)}}(B_{t_{\pi(1)}}) \dots Q_{t_{\pi(d)}}(B_{t_{\pi(d)}}) = Q_{t_{\pi(1)}} \otimes \dots \otimes Q_{t_{\pi(d)}}(B_{t_{\pi(1)}} \times \dots \times B_{t_{\pi(d)}}) = Q_{t_{\pi(1)}, \dots, t_{\pi(d)}}(B_{t_{\pi(1)}} \times \dots \times B_{t_{\pi(d)}})$$

and

$$Q_{t_1, \dots, t_d, t_{d+1}}(B_{t_1, \dots, t_d} \times S) = (Q_{t_1} \otimes \dots \otimes Q_{t_d}) \otimes Q_{t_{d+1}}(B_{t_1, \dots, t_d} \times S) = (Q_{t_1} \otimes \dots \otimes Q_{t_d}(B_{t_1, \dots, t_d}))\, Q_{t_{d+1}}(S) = Q_{t_1, \dots, t_d}(B_{t_1, \dots, t_d})$$

Thus $\hat{Q}$ is a consistent family of finite dimensional distributions. Hence the claim is true. □

Remark: the next question which will be addressed in the lectures is: can we extend uniquely P_0 to a σ-additive probability defined on the σ-algebra $\sigma(\hat{C})$ generated by the cylinders? By Caratheodory theorem, it is enough to show that P_0 is σ-additive on the cylinder algebra, namely if $(C_n : n \in \mathbb{N}) \subseteq \hat{C}$ is a cylinder sequence with $C_n \downarrow \emptyset$, necessarily $P_0(C_n) \downarrow 0$. This is the content of Kolmogorov extension theorem, which requires an additional assumption on the probability space $(S, \mathcal{S})$.

5.

In general, let Ω be an abstract space and $\mathcal{E} \subseteq 2^{\Omega}$ a collection of Ω-subsets. Let $\mathcal{F} = \sigma(\mathcal{E})$ be the σ-algebra generated by $\mathcal{E}$.

Claim: $\Lambda \in \mathcal{F}$ if and only if $\Lambda \in \sigma(\hat{\mathcal{E}})$ for some countable collection $\hat{\mathcal{E}} \subseteq \mathcal{E}$, which may depend on Λ.

Proof:

Direct proof:

As for all $\hat{C} \subseteq \mathcal{E}$, $\sigma(\hat{C}) \subseteq \mathcal{F} = \sigma(\mathcal{E})$, the "if"-part ($\Leftarrow$) is clear.

$\Rightarrow$ Let us define

$\cdot \mathcal{B} = \{A \in \mathcal{F} \mid A \in \sigma(\hat{C})$ for some countable $\hat{C} \subseteq \mathcal{E}\}$

We show that $\mathcal{B}$ is both a π-class and Dynkin class and it contains $\mathcal{E}$. By Dynkin's π-λ-theorem $\mathcal{B}$ is a σ-algebra, and as $\mathcal{E} \subseteq \mathcal{B}$, $\mathcal{F} = \sigma(\mathcal{E}) \subseteq \mathcal{B}$, thus proving the "only if" ($\Rightarrow$) part.

First we note: $A \in \mathcal{E} \Rightarrow (A \in \mathcal{F} = \sigma(\mathcal{E}) \wedge A \in \sigma(\{A\})) \Rightarrow A \in \mathcal{B}$. Thus $\mathcal{E} \subseteq \mathcal{B}$.

Then let us show $\mathcal{B}$ is a π-system.

As $\Omega \in \mathcal{F}$ and $\Omega \in \sigma(\emptyset)$, $\emptyset \subseteq \mathcal{E}$ countable, $\Omega \in \mathcal{B}$, so $\mathcal{B} \neq \emptyset$.

Let $A_1; A_2 \in \mathcal{B}$. Then $A_1 \cap A_2 \in \mathcal{F}$ and $A_1 \cap A_2 \in \sigma(\hat{C}_{A_1} \cup \hat{C}_{A_2})$ where $\hat{C}_{A_1}; \hat{C}_{A_2} \subseteq \mathcal{E}$ are countable collections as in the definition of $\mathcal{B}$. As $\hat{C}_{A_1} \cup \hat{C}_{A_2}$ is countable, $A_1 \cap A_2 \in \mathcal{B}$.

So $\mathcal{B}$ is a π-system

Let us assume $\Omega \neq \emptyset$ (for otherwise "Dynkin class on Ω" is not defined). Let us show $\mathcal{B}$ is a Dynkin class.

As above, $\Omega \in \mathcal{B}$.

Let $A \in \mathcal{B}$. Then $A^c \in \mathcal{F}$ and $A^c \in \sigma(\hat{C}_A)$, where $\hat{C}_A \subseteq \mathcal{E}$ is a countable collection such that $A \in \sigma(\hat{C}_A)$. Thus $A^c \in \mathcal{B}$.

Let $A_1; A_2; \ldots \in \mathcal{B}$ be such that $A_i \cap A_j = \emptyset \ \forall i \neq j$. Then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$ and $\bigcup_{n=1}^{\infty} A_n \in \sigma(\bigcup_{n=1}^{\infty} \hat{C}_{A_n})$ where $\hat{C}_{A_n} \subseteq \mathcal{E}$ is a countable collection such that $A_n \in \sigma(\hat{C}_{A_n})$, $n \in \mathbb{N}$. Now as $\bigcup_{n=1}^{\infty} \hat{C}_{A_n} \subseteq \mathcal{E}$ is a countable collection, $\bigcup_{n=1}^{\infty} A_n \in B$.

So B is a Dynkin system.

Hence the claim is true. $\square$

6\. We come back to the construction of the σ-algebra generated by the cylinders on $\Omega = S^T$. Using the previous exercise, let us show that a set A in the σ-algebra $\sigma(\hat{C})$ generated by the cylinders is determined by at most countably many T-coordinates.

Let $A \in \sigma(\hat{C})$. By the previous exercise 5 $A \in \sigma(B)$ for some countable collection $B \subseteq \hat{C}$.

Now, as each cylinder $C \in B$ is determined by finitely many T-coordinates, performing countable set-operations to these produces sets determined by countably many T-coordinates. Taking the closure with respect to countable set-operations of these sets determined by countable many T-coordinates produces only sets determined by countably many T-coordinates as we can enumerate the sets by ordinals according to which state of the process they appear (sets in B are 0-level sets, sets produced by countable set operations applied to 0-level sets are 1-level sets, et cetera) and notice that

$$\sigma(B) \subseteq \{A \subseteq \Omega \mid A \text{ appears before level } \aleph_1\} \text{ (aleph-1)}$$

as clearly the right-hand side is closed with respect to countable set operations, and thus a σ-algebra, and it contains B.

To be more precise, let us define

$$K := \{A \in \sigma(\hat{C}) \mid \exists I \subseteq T \text{ countable}: \forall \omega \in \Omega\, \forall a \in A: (\forall i \in I: \omega_i = a_i$$

$$\Rightarrow \omega \in A)\}$$

Obviously all cylinder sets $C \in K$, as they are defined by a finite number of T-coordinates.

$\Omega \in K$ as we can choose $I = \emptyset$ is countable. Thus $K \neq \emptyset$.

For $A \in K$ we may choose $I_{A^c} = I_A$ to show $A^c \in K$ (as if we have $\omega \in \Omega$, $a \in A^c$ and $\forall i \in I_A: \omega_i = a_i$, then as $a \in A^c$ it must be the case that $\forall b \in A \exists i_b \in I: \omega_{i_b} \neq b_{i_b}$ as otherwise we have $\exists b \in A: \forall i \in I: a_i = \omega_i = b_i \Rightarrow a \in A$ ↯. So $\forall b \in A \exists i_b \in I: \omega_{i_b} \neq b_{i_b}$ $\Rightarrow \omega \in A^c$).

Then, for $A_1; A_2; \ldots \in K$, let $I_{A_1}; I_{A_2}; \ldots \subseteq T$ countable as in the definition of K; then

$$\forall \omega \in \Omega \; \forall a \in \bigcap_{n=1}^{\infty} A_n: \forall i \in \bigcup_{n=1}^{\infty} I_{A_n}: \omega_i = a_i \Rightarrow \omega \in \bigcap_{n=1}^{\infty} A_n$$

so that $\bigcap_{n=1}^{\infty} A_n \in K$.

Thus K is a σ-algebra.

Now $\sigma(\hat{C}) \subseteq K$ (as $\hat{C} \subseteq K$). This proves the claim: $A \in \sigma(\hat{C})$ is determined by countably many T-coordinates.

In particular let $T = \mathbb{R}^m$ and $S = \mathbb{R}^d$.

Claim: the space of continuous functions

$$C(\mathbb{R}^m, \mathbb{R}^d) = \{\omega: \mathbb{R}^m \to \mathbb{R}^d \text{ continuous function}\} \subseteq (\mathbb{R}^d)^{\mathbb{R}^m}$$

is not in the σ-algebra $\sigma(\hat{C})$ generated by cylinders.

Proof:

Proof by contradiction:

Counterassumption: $C(\mathbb{R}^m, \mathbb{R}^d) \in \sigma(\hat{C})$.

By above, there exists $I \subseteq \mathbb{R}^m$ countable: $\forall f \in (\mathbb{R}^d)^{\mathbb{R}^m} \; \forall g \in C(\mathbb{R}^m, \mathbb{R}^d): f \upharpoonright I = g \upharpoonright I \Rightarrow f \in C(\mathbb{R}^m; \mathbb{R}^d)$.

Now, choose $g \in C(\mathbb{R}^m, \mathbb{R}^d)$ and $f = g + \mathbb{1}_x$; $x \in \mathbb{R}^m \setminus I$. Now f is not continuous, but $f \upharpoonright I = g \upharpoonright I$, so $f \in C(\mathbb{R}^m; \mathbb{R}^d)$ ↯.

Thus the counterassumption leads to a contradiction.

Hence the claim is true. $\square$

PROBLEMS 4.

1.

Claim: if P is a probability measure on $\Omega = \mathbb{R}^d$ equipped with the Borel σ-algebra $\mathcal{B}(\mathbb{R}^d)$, every Borel set $B \in \mathcal{B}(\mathbb{R}^d)$ satisfies the following Approximation Property: for every $\varepsilon > 0$ there is an open set $U \subseteq \mathbb{R}^d$ and a closed set $C \subseteq \mathbb{R}^d$ such that $U \supseteq B \supseteq C$ and $P(U \setminus C) \leq \varepsilon$.

Proof:

Direct proof: Let P be a probability measure on $\Omega = \mathbb{R}^d$ equipped with the Borel σ-algebra $\mathcal{B}(\mathbb{R}^d)$.
Let us consider the class of events

$D = \{ B \in \mathcal{B}(\mathbb{R}^d) \mid B \text{ has the Approximation Property} \} \subseteq \mathcal{B}(\mathbb{R}^d)$

Let us first show that for the class

$\hat{C} = \{ C \subseteq \mathbb{R}^d \mid C \text{ is closed} \}$

we have $\hat{C} \subseteq D$ and $\hat{C}$ is a π-class (closed under intersections).

First, let $C \in \hat{C}$. Then C is closed. Let us define, for $\varepsilon > 0$

$C^\varepsilon := \{ y \in \mathbb{R}^d \mid \exists x \in C : |x-y| < \varepsilon \} \supseteq C$

Now C^ε is open (for $y \in C^\varepsilon$, let $x \in C$ be such that $|x-y| < \varepsilon$. Now $B(y, \varepsilon - |x-y|) \subseteq C^\varepsilon$, so C^ε is open), and

$$C = \bigcap_{n \in \mathbb{N}} C^{\frac{1}{n}}$$

(as $C \subseteq C^{\frac{1}{n}}$ $\forall n \in \mathbb{N}$ and $z \in \mathbb{R}^d \setminus C \Rightarrow \exists r > 0 : B(z,r) \subseteq \mathbb{R}^d \setminus C$ as $\mathbb{R}^d \setminus C$ is open (C is closed) and thus $z \notin C^r$, so $\exists n' \in \mathbb{N} : \frac{1}{n'} < r$ so that $z \notin C^{\frac{1}{n'}}$).

Let us now use the σ-additivity of P to show that C has the Approximation Property. Namely, as P is σ-additive we have (Gasbarra: Lecture notes in probability theory fall semester 2015: Lemma 1.1.1.):

By the Dynkin lemma (Yasbarra: Lecture notes in probability theory Fall semester 2015: Lemma 1.1.3.2): $d(\hat{C}) = \sigma(\hat{C})$, and as we have

$$B(\mathbb{R}^d) = \sigma(\hat{C}) = d(\hat{C}) \subseteq D,$$

we can conclude that all Borel sets have the Approximation Property.

Hence the claim is true. $\square$

Claim: when $B \in B(\mathbb{R}^d)$ is a Borel set, $\forall \varepsilon > 0$ one can find an open set U and a compact set K with $K \subseteq B \subseteq U$ and $P(U \setminus K) < \varepsilon$.

Proof:

Direct proof:

Let $B \in B(\mathbb{R}^d)$. Let $\varepsilon > 0$. Let $U \subseteq \mathbb{R}^d$ be open, $C \subseteq \mathbb{R}^d$ be closed, such that $C \subseteq B \subseteq U$ and $P(U \setminus C) < \frac{\varepsilon}{2}$.

Let us have $K_n = \overline{B(0,n)} \cap C$ $\forall n \in \mathbb{N}$. K_n is closed and bounded $\forall n \in \mathbb{N}$, so K_n is compact $\forall n \in \mathbb{N}$. Furthermore $K_n \subseteq K_{n+1}$ $\forall n \in \mathbb{N}$ and $C = \bigcup_{n=1}^{\infty} K_n$. Thus $P(C) = \lim_{n \to \infty} P(K_n)$. Now

$$0 \leq P(U \setminus C) = P(U) - P(C) = P(U) - \lim_{n\to\infty} P(K_n) < \frac{\varepsilon}{2}$$

and thus, for some $k \in \mathbb{N}$, $K_k \subseteq C \subseteq B \subseteq U$, K_k is compact and $P(U \setminus K_k) = P(U) - P(K_k) < \varepsilon$.

Hence the claim is true. $\square$

Remark: we have used the Approximation Property of the Borel sets in the proof of Kolmogorov extension theorem.

2.

Let $f: \mathbb{R} \to \mathbb{R}$ be a non-decreasing function, $f(s) \leq f(t)$ when $s \leq t$.

Claim: f is Borel measurable, which means that for every Borel set $B \in B(\mathbb{R})$, the counterimage $f^{-1}[B] := \{t \in \mathbb{R} \mid f(t) \in B\}$ is a Borel set.

Proof:

Direct proof:

Let us define

$D := \{B \in \mathcal{B}(\mathbb{R}) \mid f^{-1}[B] \in \mathcal{B}(\mathbb{R})\}$.

Now we note that for $t \in \mathbb{R}$

$f^{-1}[\,]-\infty, t]\,] = \,]-\infty, \sup\{x \in \mathbb{R} \mid f(x) \le t\}\, I$

is an interval (open or closed in the right; here we write I on the right because we do not know whether the interval should be open or closed at right boundary); here we agree $\sup \emptyset = -\infty$.

Thus $]-\infty, t] \in D \ \forall t \in \mathbb{R}$.

Furthermore, $f^{-1}[\mathbb{R}] = \mathbb{R} \in \mathcal{B}(\mathbb{R})$, so $\mathbb{R} \in D$; also, by Gasbarra: Lecture notes in probability theory fall semester 2015: Lemma 2.0.1 we see that

$A \in D \Rightarrow A^c \in D$,

$(A_n)_{n \in \mathbb{N}} \subseteq D \Rightarrow \bigcup_{n=1}^{\infty} A_n \in D$

Hence D is a σ-algebra that contains all intervals $]-\infty, t]$, $t \in \mathbb{R}$. By Gasbarra: Lecture notes in probability theory fall semester 2015: Example 1.1.2 we have

$\sigma(\{]-\infty, t] \mid t \in \mathbb{R}\}) = \mathcal{B}(\mathbb{R}) \subseteq D$.

Hence f is Borel-measurable.

Thus the claim is true. $\square$

3.

On a probability space $(\Omega, \mathcal{F}, P)$, let $(A_n : n \in \mathbb{N})$ be any sequence of pairwise disjoint events, which means $A_i \cap A_j = \emptyset$ when $i \neq j$.

Claim: $\lim_{n\to\infty} P(A_n)=0$

Proof:

Direct proof:

$$1=P(\Omega)\geq P\left(\bigcup_{n=1}^{\infty} A_n\right)=\sum_{n=1}^{\infty} P(A_n)\geq 0$$

Also, $\sum_{n=1}^{\infty} P(A_n)$ converges, so $\lim_{n\to\infty} P(A_n)=0$.

Hence the claim is true. $\square$

5.

On a probability space $(\Omega, \mathcal{F}, P)$, let $(A_\alpha : \alpha\in I)$ be a family of pairwise disjoint events, indexed by an index set I.

Claim: if $P(A_\alpha)>0\ \forall \alpha\in I$, then I must be countable.

Proof:

Direct proof:

Let us show that $\forall n\in\mathbb{N}$ the set $\{\alpha\in I \mid P(A_\alpha)>\frac{1}{n}\}$ is finite.

From this it follows that $\{\alpha\in I \mid P(A_\alpha)>0\}=\bigcup_{n\in\mathbb{N}}\{\alpha\in I \mid P(A_\alpha)>\frac{1}{n}\}$ is countable.

Now, if $\{\alpha\in I \mid P(A_\alpha)>\frac{1}{n}\}$ would have more than n members, we would have $A_{\alpha_1},\dots,A_{\alpha_{n+1}}\in\{A_\alpha \mid P(A_\alpha)>\frac{1}{n}\}$ and thus

$$1=P(\Omega)\geq P(A_{\alpha_1}\cup\dots\cup A_{\alpha_{n+1}})=P(A_{\alpha_1})+\dots+P(A_{\alpha_{n+1}})\geq \underbrace{\frac{1}{n}+\dots+\frac{1}{n}}_{n+1 \text{ times}}$$

$$=\frac{n+1}{n}>1 \ \lightning.$$

Thus $\{\alpha\in I \mid P(A_\alpha)>\frac{1}{n}\}$ has at most n members, $n\in\mathbb{N}$.

Thus $\{\alpha\in I \mid P(A_\alpha)>0\}$ is countable.

Thus the claim is true. $\square$

5.

Let P be a probability on $\Omega = \mathbb{R}$ equipped with the Borel σ-algebra $\mathcal{B}(\mathbb{R})$. We have shown the cumulative distribution function $F(x) = P(]-\infty, x])$ is right-continuous, which means

$$F(x+) = \lim_{u \downarrow x} F(u) = F(x) \ \forall x \in \mathbb{R}$$

Denote the jump size of F at x by $\Delta F(x) = F(x) - F(x-)$ where $F(x-) = \lim_{s \uparrow x} F(s)$ is the limit from the left.

a) Claim: $P(\{x\}) = \Delta F(x)$

Proof:

Direct proof:

$$P(\{x\}) = P(]-\infty, x] \setminus]-\infty, x[) = P(]-\infty, x]) - P(]-\infty, x[) = F(x)$$

$$- P\left(\bigcup_{n=1}^{\infty}]-\infty, x - \tfrac{1}{n}]\right) = F(x) - \lim_{n \to \infty} P(]-\infty, x - \tfrac{1}{n}]) = F(x) - \lim_{n \to \infty} F(]-\infty, x - \tfrac{1}{n}])$$

$$= F(x) - \lim_{s \uparrow x} F(s) = F(x) - F(x-) = \Delta F(x).$$

Hence the claim is true. □

b) Claim: the set $\{x \in \mathbb{R} \mid \Delta F(x) > 0\}$ is at most a countable set.

Proof:

Direct proof:

Using part a), we can write $\{x \in \mathbb{R} \mid \Delta F(x) = P(\{x\}) > 0\}$. The claim then follows from exercise 4 applied to family $\{A_\alpha = \{\alpha\} : \alpha \in \{x \in \mathbb{R} \mid \Delta F(x) > 0\}\}$ of pairwise disjoint events indexed by index set $I = \{x \in \mathbb{R} \mid \Delta F(x) > 0\}$.

Hence the claim is true. □

6.

Suppose a function $F: \mathbb{R} \to [0,1]$ is given by

$$F(x) = \sum_{n=1}^{\infty} 2^{-n} \mathbb{1}(x \geq \tfrac{1}{n})$$

a) Claim: $F(t)$ is the cumulative distribution function of a probability P on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Proof:

Direct proof:

We see that

$$\lim_{t\to-\infty} F(t) = \lim_{t\to-\infty} \sum_{n=1}^{\infty} 2^{-n} \mathbb{1}(t \geq \tfrac{1}{n}) = \sum_{n=1}^{\infty} 2^{-n} \lim_{t\to-\infty} \mathbb{1}(t \geq \tfrac{1}{n}) = \sum_{n=1}^{\infty} 2^{-n} \cdot 0$$

$$= 0$$

$$\lim_{t\to\infty} F(t) = \lim_{t\to\infty} \sum_{n=1}^{\infty} 2^{-n} \mathbb{1}(t \geq \tfrac{1}{n}) = \sum_{n=1}^{\infty} 2^{-n} \lim_{t\to\infty} \mathbb{1}(t \geq \tfrac{1}{n}) = \sum_{n=1}^{\infty} 2^{-n} \cdot 1 = 1$$

Also we see that for $s \leq t$

$$F(s) = \sum_{n=1}^{\infty} 2^{-n} \mathbb{1}(s \geq \tfrac{1}{n}) \leq \sum_{n=1}^{\infty} 2^{-n} \mathbb{1}(t \geq \tfrac{1}{n}) = F(t)$$

so F is non-decreasing.

Finally, for $t \in \mathbb{R}$,

$$F(t+) = \lim_{u \downarrow t} F(u) = \lim_{u\downarrow t} \sum_{n=1}^{\infty} 2^{-n} \mathbb{1}(u \geq \tfrac{1}{n}) = \sum_{n=1}^{\infty} 2^{-n} \lim_{u\downarrow t} \mathbb{1}(u \geq \tfrac{1}{n})$$

$$= \sum_{n=1}^{\infty} 2^{-n} \mathbb{1}(t \geq \tfrac{1}{n}) = F(t)$$

so F is right-continuous.

Thus F is a cumulative distribution function of a probability P on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.
Hence the claim is true. $\square$

For such P, let us compute the probabilities of the following events:

$A = [1, \infty[$
$B = [\frac{1}{10}, \infty[$
$C = \{0\}$
$D = [0, \frac{1}{2}[$
$E =]-\infty, 0[$
$G =]0, \infty[$

$P(A)=P([1,\infty[)=1-P(]-\infty,1[)=1-\lim_{s\uparrow 1}F(s)=1-\lim_{s\uparrow 1}\sum_{n=1}^{\infty}2^{-n}\mathbb{1}(s\geq\frac{1}{n})$

$=1-\sum_{n=1}^{\infty}2^{-n}\lim_{s\uparrow 1}\mathbb{1}(s\geq\frac{1}{n})=1-\sum_{n=2}^{\infty}2^{-n}\cdot 1-2^{-1}\cdot 0=\frac{1}{2}$

$P(B)=P([\frac{1}{10},\infty[)=1-P(]-\infty,\frac{1}{10}[)=1-\lim_{s\uparrow\frac{1}{10}}F(s)=1-\lim_{s\uparrow\frac{1}{10}}\sum_{n=1}^{\infty}2^{-n}\mathbb{1}(s\geq\frac{1}{n})$

$=1-\sum_{n=1}^{\infty}2^{-n}\lim_{s\uparrow\frac{1}{10}}\mathbb{1}(s\geq\frac{1}{n})=1-\sum_{n=11}^{\infty}2^{-n}\cdot 1-\sum_{n=1}^{10}2^{-n}\cdot 0=1-2^{-10}$

$P(C)=P(\{0\})=\Delta F(0)=F(0)-\lim_{s\uparrow 0}F(s)=\sum_{n=1}^{\infty}2^{-n}\mathbb{1}(0\geq\frac{1}{n})-\sum_{n=1}^{\infty}2^{-n}\lim_{s\uparrow 0}$

$\mathbb{1}(s\geq\frac{1}{n})=\sum_{n=1}^{\infty}2^{-n}\cdot 0-\sum_{n=1}^{\infty}2^{-n}\cdot 0=0$

$P(D)=P([0,\frac{1}{2}[)=P(]-\infty,\frac{1}{2}])-P(]-\infty,0[)-P(\{\frac{1}{2}\})=F(\frac{1}{2})-\lim_{s\uparrow 0}F(s)$

$-\Delta F(\frac{1}{2})=F(\frac{1}{2})-\sum_{n=1}^{\infty}2^{-n}\lim_{s\uparrow 0}\mathbb{1}(s\geq\frac{1}{n})-F(\frac{1}{2})+\lim_{s\uparrow\frac{1}{2}}F(s)=-\sum_{n=1}^{\infty}2^{-n}\cdot 0$

$+\sum_{n=1}^{\infty}2^{-n}\lim_{s\uparrow\frac{1}{2}}\mathbb{1}(s\geq\frac{1}{n})=\sum_{n=3}^{\infty}2^{-n}\cdot 1+2^{-2}\cdot 0+2^{-1}\cdot 0=2^{-2}=\frac{1}{4}$

$P(E)=P(]-\infty,0[)=P(]-\infty,0])-P(\{0\})=F(0)-\Delta F(0)=\lim_{s\uparrow 0}F(s)$

$=\sum_{n=1}^{\infty}2^{-n}\lim_{s\uparrow 0}\mathbb{1}(s\geq\frac{1}{n})=\sum_{n=1}^{\infty}2^{-n}\cdot 0=0$

$P(G)=P(]0,\infty[)=1-P(]-\infty,0[)-P(\{0\})=1-P(E)-P(C)=1-0-0=0$

b) Let us define a random variable X on a probability space $(\Omega,\mathcal{F},P)$ of our choice, with a probability P of our choice, such that the distribution $P(\{\omega\in\Omega\,|\,X(\omega)\leq t\})=F(t)$.

We note that we can always define the random variable as the identity map in the space where it takes values, in this case $\mathbb{R}$ equipped with the Borel σ-algebra $\mathcal{B}(\mathbb{R})$.

So, let us choose $\Omega=\mathbb{R}$, $\mathcal{F}=\mathcal{B}(\mathbb{R})$ and $X:\Omega\to\mathbb{R}$ (also $X:(\Omega,\mathcal{F})\to(\mathbb{R},\mathcal{B}(\mathbb{R}))$) as $X=id$ (measurability of X is trivial).

Let us choose P on $(\Omega,\mathcal{F})=(\mathbb{R},\mathcal{B}(\mathbb{R}))$ as the measure whose cumulative distribution function is F; also, the measure that gives $P(]a,b])=F(b)-F(a)$ $\forall a<b, a,b\in\mathbb{R}$. This measure exists (and is unique) by Gasbarra: Lecture notes in probability theory fall semester 2015; Example on page 22. Now we have $P(\{\omega\in\Omega\,|\,X(\omega)\leq t\})=P(\{s\in\mathbb{R}\,|\,id(s)\leq t\})=P(]-\infty,t])=F(t)$ so $(\Omega,\mathcal{F},P)$ is as wanted (P has a discrete distribution with $P(\{\frac{1}{n}\})=2^{-n}$ $\forall n\in\mathbb{N}$, $P(\{x\})=0$ $\forall x\in\mathbb{R}\setminus\{\frac{1}{n}\,|\,n\in\mathbb{N}\}$).

PROBLEMS 5.

1.

Let us consider the probability space $\Omega = [0,1]$ equipped with the Borel σ-algebra $\mathcal{F} = \mathcal{B}([0,1])$ and the uniform probability measure P such that $P(]a,b]) = b-a$ for $0 \le a \le b \le 1$, which is also called Lebesgue measure.

Claim: the identity map $U : \Omega \to [0,1]$ with $U(\omega) = \omega$ is an uniformly distributed random variable, which means

$$P(\{\omega \in \Omega \mid U(\omega) \in]a,b]\}) = b-a.$$

Proof:

Direct proof:

For any Borel-set $B \in \mathcal{B}([0,1])$, $U^{-1}[B] = B \in \mathcal{B}([0,1]) = \mathcal{F}$, so U is a measurable map, thus U is a random variable.

Also we see that

$$P(\{\omega \in \Omega \mid U(\omega) \in]a,b]\}) = P(]a,b]) = b-a.$$

Thus U is an uniformly distributed random variable.

Hence the claim is true. $\square$

Let now $(\Omega, \mathcal{F}, P)$ be an abstract probability space and $U : \Omega \to [0,1]$ a random variable with uniform distribution on $[0,1]$, which means

$$P(\{\omega \in \Omega \mid U(\omega) \in]a,b]\}) = b-a.$$

Let $F : \mathbb{R} \to [0,1]$ be a cumulative probability distribution function (c.d.f.), which is right continuous and non-decreasing with $F(+\infty) = 1$ and $F(-\infty) = 0$.

We shall construct a random variable on $(\Omega, \mathcal{F})$ with values in $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that

$P(\{\omega \in \Omega \mid X(\omega) \le t\}) = F(t)$

Let us assume for simplicity that $F(t)$ is continuous and strictly increasing, with $F(s) < F(t)\ \forall s<t$.

In this case there is an unique inverse $F^{-1}: [0,1] \to \mathbb{R}$ such that $F(F^{-1}(u)) = u\ \forall u \in [0,1]$ and $F^{-1}(F(t)) = t\ \forall t \in \mathbb{R}$.

Claim: $X(\omega) = F^{-1}(U(\omega))$ is a random variable with

$P(\{X(\omega) \le t\}) = F(t)$.

Proof:

Direct proof:

Set $X(\omega) := F^{-1}(U(\omega))$.

For any $t \in \mathbb{R}$, $(F^{-1})^{-1}[]-\infty, t]] = F[]-\infty, t]] =]-\infty, F(t)] \cap [0,1] \in \mathcal{B}([0,1])$. Thus, similarly as in exercise 2 of problem set 4, we see that $F^{-1}: [0,1] \to \mathbb{R}$ is a measurable map.

Hence $X = F^{-1} \circ U$ is a measurable map, thus X is a random variable.

We also have

$P(\{X(\omega) \le t\}) = P(\{U(\omega) \le F(t)\}) = P(\{\omega \in \Omega \mid U(\omega) \in [0, F(t)]\})$

$= P(\{\omega \in \Omega \mid U(\omega) \in]0, F(t)]\}) = F(t) - 0 = F(t)$

as $P(\{\omega \in \Omega \mid U(\omega) = 0\}) = P(\{\omega \in \Omega \mid U(\omega) \in [0,1]\}) - P(\{\omega \in \Omega \mid U(\omega)$

$\in]0,1]\}) = P(\Omega) - (1-0) = 1 - 1 = 0.$

Hence the claim is true. $\square$

Using a generalized inverse, this construction extends also to the general cumulative distribution function, which does not need to be continuous from the left nor strictly increasing.

2.

On an abstract probability space $(\Omega, \mathcal{F}, P)$, let $X(\omega) \geq 0 \; \forall \omega \in \Omega$ be a non-negative random variable.

We have defined the expectation of X as

$$E_P(X) = \sup_{\substack{0 \leq Y \leq X \\ Y \in SF^+}} E_P(Y)$$

where the supremum is taken over the simple random variables Y (taking finitely many values) such that $0 \leq Y(\omega) \leq X(\omega) \; \forall \omega \in \Omega$.

Let us assume that $X(\omega) \in \mathbb{N} \; \forall \omega \in \Omega$.

a) Claim: $E_P(X) = \sum_{n=1}^{\infty} n P(\{\omega \in \Omega \mid X(\omega) = n\}) = \sum_{n=1}^{\infty} n P_X(\{n\})$

where $P_X(\{n\}) = P(\{\omega \in \Omega \mid X(\omega) = n\})$ is the distribution of X with $E_P(X) \in [0, +\infty]$ (the series may also diverge).

Proof:

Direct proof:

Let us note that, for $k \in \mathbb{N}$

$$Y_k = \sum_{n=1}^{k-1} n \mathbb{1}(X = n) + k \mathbb{1}(X \geq k)$$

is a simple random variable taking finitely many values such that $0 \leq Y_k \leq X$. Thus

$$E_P(X) \geq E_P(Y_k) = \sum_{n=1}^{k-1} n P(\{\omega \in \Omega \mid X(\omega) = n\}) + k P(\{\omega \in \Omega \mid X(\omega) \geq k\}) \geq \sum_{n=1}^{k-1} n P_X(\{n\})$$

for all $k \in \mathbb{N}$. Thus $E_P(X) \geq \sum_{n=1}^{\infty} n P(\{\omega \in \Omega \mid X(\omega) = n\})$.

On the other hand, let Y be a simple random variable taking only finitely many values such that $0 \leq Y \leq X$. Let $k' \in \mathbb{N}$ be such that $Y \leq k'$ (as Y takes only finitely many values, such k' exists). Now as $X \in \mathbb{N}$ we see that $Y \leq Y_{k'}$ as for any $\omega \in \Omega$, if $X(\omega) \leq k'$, $Y(\omega) \leq X(\omega) = Y_{k'}(\omega)$, and if $X(\omega) > k'$, $Y(\omega) \leq k' = Y_{k'}(\omega)$.

Hence $E_P(Y) \leq E_P(Y_{k+1})$. Thus

$$E_P(X) = \sup_{k \in \mathbb{N}} E_P(Y_k).$$

Now we note

$$E_P(Y_k) = \sum_{n=1}^{k-1} n P_X(\{n\}) + k P(\{\omega \in \Omega \mid X(\omega) \geq k\}) = \sum_{n=1}^{k-1} n P_X(\{n\})$$
$$+ \sum_{n=k}^{\infty} k P_X(\{n\}) \leq \sum_{n=1}^{\infty} n P_X(\{n\})$$

so that $E_P(X) \leq \sum_{n=1}^{\infty} n P(\{\omega \in \Omega \mid X(\omega) = n\})$.

Hence $E_P(X) = \sum_{n=1}^{\infty} n P(\{\omega \in \Omega \mid X(\omega) = n\})$.

Hence the claim is true. □

b) Let us show one non-trivial example with $X(\omega)$ taking countably many values in $\mathbb{N}$ and choosing the distribution of X, $P_X(\{n\})$ such that $E_P(X) < \infty$, and another example where $E_P(X) = +\infty$.

First, let $X : \Omega \to \mathbb{N}$ be a random variable such that $P_X(\{n\}) = \frac{1}{2^n}$ $\forall n \in \mathbb{N}$. Then

$$\sum_{n \in \mathbb{N}} P_X(\{n\}) = 1$$

$$E_P(X) = \sum_{n \in \mathbb{N}} n P_X(\{n\}) = \sum_{n=1}^{\infty} n \frac{1}{2^n} < \infty.$$

Then, let $X : \Omega \to \mathbb{N}$ be a random variable such that $P_X(\{2^n\}) = \frac{1}{2^n}$ $\forall n \in \mathbb{N}$ (and $P_X(\{m\}) = 0$ if $m \neq 2^n$ for some $n \in \mathbb{N}$). Then

$$\sum_{n \in \mathbb{N}} P_X(\{n\}) = 1$$

$$E_P(X) = \sum_{n \in \mathbb{N}} n P_X(\{n\}) = \sum_{n=1}^{\infty} 2^n \frac{1}{2^n} = \infty$$

3.

On an abstract probability space $(\Omega, \mathcal{F}, P)$, let $N(\omega)$ be a Poisson distributed random variable with parameter $\lambda > 0$, such that

$P(\{\omega \in \Omega \mid N(\omega) = k\}) = P_\lambda(\{k\}) = \exp(-\lambda)\frac{\lambda^k}{k!}$

a) Let us check that $(P_\lambda(\{k\}))_{k\in\mathbb{N}}$ defines a probability distribution on $\mathbb{N}_0 = \{0,1,2,\dots\}$, in particular that $P_\lambda(\mathbb{N}_0)=1$.

Obviously $P_\lambda(\{k\}) = \exp(-\lambda)\frac{\lambda^k}{k!} > 0 \quad \forall k \in \mathbb{N}_0$, and

$$P_\lambda(\mathbb{N}_0) = \sum_{k=0}^{\infty} P_\lambda(\{k\}) = \sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}e^{\lambda} = 1$$

so $(P_\lambda(\{k\}))_{k\in\mathbb{N}}$ defines a probability distribution on $\mathbb{N}_0$.

b) Let us compute the moment generating function $m: \mathbb{R} \to [0,\infty]$

$$m(\theta) = E_P(\exp(\theta N)), \ \theta \in \mathbb{R}$$

We have

$$m(\theta) = E_P(e^{\theta N}) = \sum_{k=0}^{\infty} e^{\theta k} e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{(e^\theta\lambda)^k}{k!} = e^{-\lambda}e^{e^\theta\lambda} = e^{\lambda(e^\theta-1)}$$

c) Claim: Stein equation for the Poisson distribution:

$$E_P(\lambda g_{N+1}) = E_P(N g_N)$$

for every bounded sequence $(g_k)_{k\in\mathbb{N}} \subseteq \mathbb{R}$.

Proof:

Direct proof:

Let $(g_k)_{k\in\mathbb{N}} \subseteq \mathbb{R}$ be bounded. Thus the, expectations below exist

$$E_P(\lambda g_{N+1}) = \sum_{k=0}^{\infty} \lambda g_{k+1} e^{-\lambda}\frac{\lambda^k}{k!} = \sum_{k=0}^{\infty}(k+1)g_{k+1}e^{-\lambda}\frac{\lambda^{k+1}}{(k+1)!} = \sum_{k'=1}^{\infty} k' g_{k'} e^{-\lambda}\frac{\lambda^{k'}}{k'!} = \sum_{k'=0}^{\infty} k' g_{k'} e^{-\lambda}\frac{\lambda^{k'}}{k'!} = E_P(N g_N)$$

Hence the claim is true. □

d) Let us compute the expectations (moments) $E_P(N^q)$ for $q \in \mathbb{N}$.

Let us use the moment generating function:

$$m(\theta)=E_p(e^{\theta N})=E_p\Big(\sum_{k=0}^{\infty}\frac{\theta^k N^k}{k!}\Big)=\sum_{k=0}^{\infty}E_p(N^k)\frac{\theta^k}{k!}=e^{\lambda(e^{\theta}-1)}=\sum_{k=0}^{\infty}\frac{\lambda^k(e^{\theta}-1)^k}{k!}$$

$$=\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\sum_{n=0}^{k}\binom{k}{n}e^{\theta n}(-1)^{k-n}=\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\sum_{n=0}^{k}\binom{k}{n}(-1)^{k-n}\sum_{p=0}^{\infty}\frac{\theta^p n^p}{p!}$$

$$=\sum_{p=0}^{\infty}\frac{\theta^p}{p!}\sum_{k=0}^{\infty}\lambda^k\frac{1}{k!}\sum_{n=0}^{k}(-1)^{k-n}\binom{k}{n}n^p=\sum_{p=0}^{\infty}\frac{\theta^p}{p!}\sum_{k=0}^{\infty}\lambda^k{p \brace k}$$

where ${p \brace k}$ is the Stirling number of the second kind,

$${p \brace k}:=\frac{1}{k!}\sum_{n=0}^{k}(-1)^{k-n}\binom{k}{n}n^p$$

Now, because ${p \brace k}=0$ for $k>p$, we have

$$m(\theta)=\sum_{p=0}^{\infty}E_p(N^p)\frac{\theta^p}{p!}=\sum_{p=0}^{\infty}\frac{\theta^p}{p!}\sum_{k=0}^{p}\lambda^k{p \brace k}$$

from which we can read

$$E_p(N^q)=\sum_{k=0}^{q}\lambda^k{q \brace k}$$

$\Rightarrow$ TO BE CONTINUED.

e) Let us compute the expectations $E_p(N^q e^{\theta N})$ for $\theta\in\mathbb{R}$ and $q\in\mathbb{N}$.

Using the moment generating function, we find that

$$\frac{d^q}{d\theta^q}m(\theta)=\frac{d^q}{d\theta^q}E_p(e^{\theta N})=E_p\Big(\frac{d^q}{d\theta^q}e^{\theta N}\Big)=E_p(N^q e^{\theta N})=\frac{d^q}{d\theta^q}\sum_{k=0}^{\infty}E_p(N^k)\frac{\theta^k}{k!}=\sum_{k=0}^{\infty}E_p(N^k)\frac{d^q}{d\theta^q}\frac{\theta^k}{k!}=\sum_{k=q}^{\infty}E_p(N^k)\frac{\theta^{k-q}}{(k-q)!}=\sum_{k=0}^{\infty}E_p(N^{k+q})\frac{\theta^k}{k!}$$

$$=\sum_{k=0}^{\infty}\frac{\theta^k}{k!}\sum_{n=0}^{k+q}\lambda^n{k+q \brace n}$$

$\Leftarrow$ CONTINUE 3 d)

Let us prove ${p \brace k}=\frac{1}{k!}\sum_{n=0}^{k}(-1)^{k-n}\binom{k}{n}n^p=0$ for $k>p$.

We note

$$\sum_{n=0}^{k}(-1)^{k-n}\binom{k}{n}n^p=\sum_{j=0}^{k}(-1)^j\binom{k}{j}(k-j)^p$$

($\binom{k}{n}=\binom{k}{k-n}$) and we give the following combinatorial interpretation for the formula: let A,B be sets, $|A|=p$, $|B|=k$. Let $X=\{f:A\to B\}$. We shall employ inclusion-exclusion principle to calculate $|Y|$ for $Y := \{f : A \xrightarrow{\text{ONTO}} B\} \subseteq X$.

Let for $b \in B$ $X_b := \{f : A \to B,\ b \notin \operatorname{ran} f\}$. Then we note

$|X_b| = (k-1)^p$ (each member of A has $k-1$ choices, $|A| = p$)

and for $b_1, \dots, b_j \in B$ distinct

$|X_{b_1} \cap \dots \cap X_{b_j}| = (k-j)^p$.

Now we have ($|X| = k^p$)

($\binom{k}{j}$ different way to choose $b_1, \dots, b_j \in B$ distinct)

$$|Y| = |X \setminus (\bigcup_{b \in B} X_b)| = |X| - |\bigcup_{b \in B} X_b| = k^p - \left(\sum_{j=1}^{k} (-1)^{j-1} \binom{k}{j} (k-j)^p\right)$$

$$= \sum_{j=0}^{k} (-1)^j \binom{k}{j} (k-j)^p$$

But, for $k > p$, $Y = \emptyset$ and thus

$$\sum_{j=0}^{k} (-1)^j \binom{k}{j} (k-j)^p = 0.$$

PROBLEMS 6.

1.

When the cumulative distribution function $F_X(t) = P(X \le t)$ of a $\mathbb{R}$-valued random variable X has derivative $f_X(t)$ almost everywhere (for all t outside a set of zero Lebesgue measure) $f_X(t)$ is called probability density function. In such case, for every non-negative and Borel measurable test function $g(x) \ge 0$ we have

$$(1)\ E_P(g(X)) = \int_\Omega g(X(\omega))\, P(d\omega) = \int_{\mathbb{R}} g(t)\, P_X(dt) = \int_{\mathbb{R}} g(t) F(dt) = \int_{\mathbb{R}} g(t) f_X(t) dt$$

where $P_X(B) = \mathbb{P}(\{\omega \in \Omega \mid X(\omega) \in B\})$ is the pushforward measure of $\mathbb{P}$ by the random variable X. The integral with respect to P_X on $\mathbb{R}$ is the same as the Lebesgue-Stieltjes integral with respect to dF, meaning P_X coincides with the measure induced by the cumulative distribution function $F(t)$ on $\mathbb{R}$.

Let us study this problem by applying the monotone class theorem. Let us define

$$\mathcal{C} := \{ g : \mathbb{R} \longrightarrow \mathbb{R},\ g \text{ bounded and Borel measurable such that (1) holds} \}.$$

We shall prove that $\mathcal{C}$ is a monotone class.

$\mathcal{C}$ is a vector space with $\mathbb{R}$-coefficients, as a $\mathbb{R}$-linear combination of bounded and Borel measurable functions is a bounded Borel measurable function, and the expectation and all the integrals in (1) are linear operations; for $f, g \in \mathcal{C}$, $a, b \in \mathbb{R}$:

$$E_P(af + bg) = a E_P(f) + b E_P(g) = a \int_\Omega f(X(\omega)) P(d\omega) + b \int_\Omega g(X(\omega)) P(d\omega)$$

$$= \int (af + bg)(X(\omega)) P(d\omega) = a \int_{\mathbb{R}} f(t) P_X(dt) + b \int_{\mathbb{R}} g(t) P_X(dt) = \int_{\mathbb{R}} (af$$

$$+ bg)(t) P_X(dt) = a \int_{\mathbb{R}} f(t) F(dt) + b \int_{\mathbb{R}} g(t) F(dt) = \int_{\mathbb{R}} (af + bg)(t) F(dt)$$

$= a\int_{\mathbb{R}} f(t) f_X(t) dt + b\int_{\mathbb{R}} g(t) f_X(t) dt = \int_{\mathbb{R}} (af + bg)(t) f_X(t) dt$

Note that if we restrict to positive functions $g: \mathbb{R} \to [0, \infty[$ in $\hat{C}$, then $\hat{C}$ is not a $\mathbb{R}$-vector space (and as $\mathbb{R}_{\geq 0}$ does not form a ring, it is not $\mathbb{R}_{\geq 0}$-vector space or anything like that).

Obviously constant function $\mathbb{1} \in \hat{C}$, for

$$E_P(\mathbb{1}) = 1 =: \int_\Omega \mathbb{1}\, P(d(\omega))$$

$$\int_{\mathbb{R}} \mathbb{1}\, P_X(dt) = P_X(\mathbb{R}) = 1$$

$$\int_{\mathbb{R}} 1\, F(dt) = \lim_{t \to +\infty} F(t) - \lim_{t \to -\infty} F(t) = 1 - 0 = 1$$

$$(2)\ \int_{\mathbb{R}} 1\, f_X(t) dt = \lim_{t \to +\infty} F(t) - \lim_{t \to -\infty} F(t) = 1 - 0 = 1,$$

where we note that equality (2) holds if F is absolutely continuous. If F is not required to be absolutely continuous we can choose F to be the Cantor function (devil's staircase), whence $f_X \equiv 0$ (almost everywhere), and $\int_{\mathbb{R}} 1 \cdot f_X(t) dt = 0$, thus forming a counterexample to the claim.

Remember that absolute continuity is equivalent to

$$F(b) - F(a) = \int_a^b f_X(t) dt \quad \forall a < b,\ a, b \in \mathbb{R}$$

and we know $\int_{\mathbb{R}} f_X(t) dt = \lim_{k \to \infty} \int_{-k}^{k} f_X(t) dt$ as $f_X(t) \geq 0\ \forall t \in \mathbb{R}$

(for those $t \in \mathbb{R}$ for which $f_X(t)$ is defined), for F is a non-decreasing function (having $F'(x) < 0$ for some point $x \in \mathbb{R}$ would mean $F(y) < F(x)$ for $x < y < \varepsilon > 0$ for some $\varepsilon > 0$).

Let us then show $\hat{C}$ is closed with respect to monotone limits. Let $(g_n)_{n \in \mathbb{N}} \subseteq \hat{C}$ be such that

$$0 \leq g_n \leq g_{n+1}\ \forall n \in \mathbb{N}$$

$$g := \lim_{n \to \infty} g_n \text{ is a bounded function}$$

Now by Monotone convergence theorem (Gasbarra: Lecture notes in probability theory fall semester 2015 (moniste): Theorem 4.1.1) for expectation and for general measures (remember that Lebesgue-Stieltjes integration is actually integration with respect to Lebesgue-Stieltjes measure)

$$(3)\ E_P(g(X)) = \lim_{n\to\infty} E_P(g_n(X)) = \lim_{n\to\infty} \int_\Omega g_n(X(\omega))\,P(d\omega) = \int_\Omega g(X(\omega))\,P(d\omega)$$

$$= \lim_{n\to\infty} \int_{\mathbb{R}} g_n(t)\,P_X(dt) = \int_{\mathbb{R}} g(t)\,P_X(dt) = \lim_{n\to\infty} \int_{\mathbb{R}} g_n(t)\,F(dt) = \int_{\mathbb{R}} g(t)\,F(dt)$$

$$= \lim_{n\to\infty} \int_{\mathbb{R}} g_n(t) f_X(t)\,dt = \int_{\mathbb{R}} g(t) f_X(t)\,dt$$

Note that in the final step we need again that $f_X \geq 0$.

Thus $g \in \hat{C}$ (obviously conditions $g: \mathbb{R} \to \mathbb{R}$, g Borel-measurable are satisfied).

So $\hat{C}$ is a monotone class. Furthermore, for $\forall a \leq b \in \mathbb{R}$, $\mathbb{1}_{]a,b]} \in \hat{C}$ (obviously real valued bounded Borel-measurable function with

$$E_P(\mathbb{1}_{]a,b]}(X)) = P(X \in]a,b]) =: \int_\Omega \mathbb{1}_{]a,b]}(X(\omega))\,P(d\omega)$$

$$\int_{\mathbb{R}} \mathbb{1}_{]a,b]}(t)\,P_X(dt) = P_X(]a,b]) = P(X \in]a,b])$$

$$\int_{\mathbb{R}} \mathbb{1}_{]a,b]}(t)\,F(dt) = F(b) - F(a) = P(X \in]a,b])$$

$$\int_{\mathbb{R}} \mathbb{1}_{]a,b]}(t) f_X(t)\,dt = F(b) - F(a);\ F \text{ absolutely continuous}$$

(F absolutely continuous $\Rightarrow \int_{[a+\frac{1}{n},b]} f_X(t)\,dt = F(b) - F(a+\frac{1}{n})\ \forall n \in \mathbb{N}$,

$f_X \geq 0 \Rightarrow \int_{]a,b]} f_X(t)\,dt = \lim_{n\to\infty} \int_{[a+\frac{1}{n},b]} f_X(t)\,dt = F(b) - F(a+) = F(b) - F(a)$,

F right-continuous)

We know that the intervals $]a,b]$ $\forall a \leq b \in \mathbb{R}$ is a π-system generating the Borel σ-algebra. Thus we get, by Gasbarra: Lecture Notes in Probability theory Fall semester 2015 (moniste): Theorem 2.0.1 that

$$\{g: \mathbb{R} \to \mathbb{R} \text{ bounded, Borel-measurable}\} \subseteq \hat{C}$$

To generalize this to all non-negative Borel-measurable functions g, we note that we can find a sequence of simple functions $0 \le g_n \uparrow g$; as simple functions necessary are bounded (and Borel-measurable), we have $(g_n)_{n\in\mathbb{N}} \subseteq \mathcal{C}$; by using Monotone convergence theorem as in (3) above, we can finally deduce the claim for g.

Now, as the integral with respect to P_X on $\mathbb{R}$ of non-negative (Borel-measurable) functions g is the same as the Lebesgue-Stieltjes integral of these g with respect to dF, we can deduce that actually for all integrable functions g (integrable with respect to $P_X \Leftrightarrow$ Lebesgue-Stieltjes integrable with respect to dF), the integral over $\mathbb{R}$ with respect to P_X is the same as the Lebesgue-Stieltjes integral over $\mathbb{R}$ with respect to dF.

In particular, for $B \in \mathcal{B}(\mathbb{R})$ a Borel set,

$$P_X(B) = \int_{\mathbb{R}} \mathbb{1}_B(x) P_X(dx) = \int_{\mathbb{R}} \mathbb{1}_B(x) F(dx) = m_F(B)$$

where m_F is the measure induced by the cumulative distribution function F on $\mathbb{R}$.

Also, P_X coincides with the measure induced by the cumulative distribution function F on $\mathbb{R}$.

2. Linearity of the expectation.

The expectation of a random variable $X(\omega)$ is defined as

$$E_P(X) = E_P(X^+) - E_P(X^-)$$

where $X^+ = \max\{X, 0\} \ge 0$, $X^- = \max\{-X, 0\} \ge 0$ are non-negative random variables, and we have defined first for non-negative random variables

$$E_P(X) = \sup_{Y \in SF : 0 \le Y \le X} E_P(Y)$$

In this way the expectation is well defined unless

$E_P(X^+) = E_P(X^-) = +\infty$.

In the lectures we have shown (first for simple random variables and then by the monotone convergence theorem) that when $X(\omega) \geq 0$, $Y(\omega) \geq 0$ P-almost surely (outside a P-null set), and $a, b \geq 0$

$$(4)\ E_P(aX + bY) = a E_P(X) + b E_P(Y)$$

Claim: Linearity holds for any random variables X, Y and $a, b \in \mathbb{R}$ when the expectations on both left and right sides in (4) are finite.

Proof:

Direct proof:

Let X, Y be random variables, $a, b \in \mathbb{R}$. Let the expectations in (4) be finite.

Let us use the representations $X = X^+ - X^-$, $Y = Y^+ - Y^-$, $a = a^+ - a^-$, $b = b^+ - b^-$, and write

$$aX + bY = a^+X^+ - a^+X^- - a^-X^+ + a^-X^- + b^+Y^+ - b^+Y^- - b^-Y^+ + b^-Y^-$$

$$= (a^+X^+ + a^-X^- + b^+Y^+ + b^-Y^-) - (a^+X^- + a^-X^+ + b^+Y^- + b^-Y^+)$$

Hence, since

$$aX + bY = (aX + bY)^+ - (aX + bY)^-$$

we have

$$(aX+bY)^+ + a^+X^- + a^-X^+ + b^+Y^- + b^-Y^+ = (aX+bY)^- + a^+X^+ + a^-X^- + b^+Y^+ + b^-Y^-$$

and thus (all terms positive, all coefficients positive)

$$E_P((aX+bY)^+) + a^+E_P(X^-) + a^-E_P(X^+) + b^+E_P(Y^-) + b^-E_P(Y^+) = E_P((aX+bY)^-)$$

$$+ a^+E_P(X^+) + a^-E_P(X^-) + b^+E_P(Y^+) + b^-E_P(Y^-)$$

$\Leftrightarrow E_P((aX+bY)^+) - E_P((aX+bY)^-) =: E_P(aX+bY) = (a^+ - a^-)(E_P(X^+)$

$- E_P(X^-)) + (b^+ - b^-)(E_P(Y^+) - E_P(Y^-)) = a E_P(X) + b E_P(Y)$

In manipulating the equations above we use the fact that the expectations on both sides of (4) are finite, also

$E_P(aX+bY) < \infty \Leftrightarrow 0 \leq E_P((aX+bY)^+) < \infty > E_P((aX+bY)^-) \geq 0$

$E_P(X) < \infty \Leftrightarrow 0 \leq E_P(X^+) < \infty > E_P(X^-) \geq 0$

$E_P(Y) < \infty \Leftrightarrow 0 \leq E_P(Y^+) < \infty > E_P(Y^-) \geq 0$

Hence the claim is true. $\Box$

3.

Let $U(\omega)$ be an uniformly distributed random variable with values in $[0,1]$ such that $P(\{U \in]a,b]\}) = b-a$ for $0 \leq a \leq b \leq 1$.

a) Claim: the powers $U(\omega)^z$, with $z \in \mathbb{Z}$ (the integers) are random variables

Proof:

Direct proof:

As $U(\omega) \in [0,1]$, and for $z \in \mathbb{Z}$, $z<0$ we have not defined 0^z, we have a problem which we can solve in 2 ways: first, as $P(U=0) = 1 - P(U \in]0,1]) = 1-(1-0) = 0$, we can ignore the problem and say that $U(\omega)^z$ is defined only P-almost everywhere, where $U(\omega) \in]0,1]$; because $U \upharpoonright U^{-1}[]0,1]] : U^{-1}[]0,1]] \to]0,1]$ is a measurable map (Borel-measure in $]0,1]$, restriction of original σ-algebra on $U^{-1}[]0,1]]$) and $x \mapsto x^z :]0,1] \to \mathbb{R}$ is continuous, $\omega \mapsto U(\omega)^z : U^{-1}[]0,1]] \to \mathbb{R}$ is a random variable. On the other hand, we can define $0^z = \infty$ for $z<0$, then $\omega \to U(\omega)^z : \Omega \to \mathbb{R} \cup \{\infty\}$ and since $x \to x^z : [0,1] \to \mathbb{R} \cup \{\infty\}$ is continuous, $U(\omega)^z$ is a random variable.
Either way, $U(\omega)^z$ is a random variable and the claim is true. $\Box$

b) Let us compute the moments $E_P(U^z) \in [0,\infty]$ for $z \in \mathbb{Z}$.

We have

$$E_P(U^z) = \int_{[0,1]} x^z P_U(dx) = \int_0^1 x^z dx = \begin{cases} \frac{1}{z+1}(1^{z+1} - 0^{z+1}) = \frac{1}{z+1}; z > -1 \\ \ln(1) - \ln(0) = \infty, \ z = -1 \\ \frac{1}{z+1}(1^{z+1} - 0^{z+1}) = \infty, \ z < -1 \end{cases}$$

c) Let us compute the exponential moments $E_P(e^{tU})$ for $t \in \mathbb{R}$.

We have

$$E_P(e^{tU}) = \int_{[0,1]} e^{tx} P_U(dx) = \int_0^1 e^{tx} dx = \begin{cases} \frac{1}{t}(e^{t\cdot 1} - e^{t \cdot 0}) = \frac{1}{t}(e^t - 1), \ t \neq 0 \\ 1, \ t = 0 \end{cases}$$

d) Let us compute the trigonometric moments $E_P(\cos(2\pi t U))$ and $E_P(\sin(2\pi t U))$ for $t \in \mathbb{R}$.

We have

$$E_P(\cos(2\pi t U)) = \int_{[0,1]} \cos(2\pi t x) P_U(dx) = \int_0^1 \cos(2\pi t x) dx$$

$$= \begin{cases} \frac{1}{2\pi t}(\sin(2\pi t \cdot 1) - \sin(2\pi t \cdot 0)) = \frac{1}{2\pi t}\sin(2\pi t), \ t \neq 0 \\ 1, \ t = 0 \end{cases}$$

$$E_P(\sin(2\pi t U)) = \int_{[0,1]} \sin(2\pi t x) P_U(dx) = \int_0^1 \sin(2\pi t x) dx$$

$$= \begin{cases} \frac{1}{2\pi t}(\cos(2\pi t \cdot 0) - \cos(2\pi t \cdot 1)) = \frac{1}{2\pi t}(1 - \cos(2\pi t)), \ t \neq 0 \\ 0, \ t = 0 \end{cases}$$

5.

Let $f : [0,T] \to \mathbb{R}^+$ be a non-negative and bounded measurable function.

We define its upper and lower Riemann-integrals as follows:

$\int^+(f) := \inf\{ I(g) \mid g \geq f, \ g$ takes finitely many values and is piecewise continuous$\}$

$J^-(f) = \sup\{I(g) \mid g \leq f,\ g$ takes finitely many values and is piecewise continuous$\}$

where the integral $I(g)$ of a piecewise continuous function g taking finitely many values is the usual finite sum.

Note that on the real line, a piecewise continuous simple function taking finitely many values is piecewise constant, with representation

$$g(x) = \sum_{k=1}^{n} a_k \mathbb{1}_{E_k}(x), \text{ with } I(g) = \sum_{k=1}^{n} a_k \text{ length}(E_k)$$

where E_k are intervals. In the construction of Lebesgue integral, the general definition uses Borel sets instead of intervals.

We say that f is Riemann integrable when $J^+(f) = J^-(f)$ which defines the Riemann integral $J(f)$ (it is possible that $J(f) = +\infty$).

a) Claim: When f is Riemann integrable the Riemann integral $J(f)$ coincides with the Lebesgue integral $I(f)$ defined in the lectures.

Proof:

Direct proof:

Let f be Riemann integrable.

We define the Lebesgue integral $I(f)$ of a Borel measurable non-negative function with respect to Lebesgue measure as

$I(f) = \sup\{I(g) \mid g \leq f,\ g$ is measurable and takes finitely many values$\}$

Let us denote $\{g \geq f,\ g$ takes finitely many values and is piecewise continuous$\} = A$, $\{g \le f, g$ takes finitely many values and is piecewise continuous$\} = B$, $\{g \le f, g$ is measurable and takes finitely many values$\} = C$.

As $B \subseteq C$, we have $J^-(f) \le I(f)$.

As for every $g \in A, g' \in C$: $g \ge f \ge g' \Rightarrow I(g) \ge I(g')$, so that

$J^+(f) \ge I(f)$.

For f Riemann integrable,

$J(f) = J^+(f) \ge I(f) \ge J^-(f) = J(f)$

$\Rightarrow J(f) = I(f)$.

Hence the claim is true. $\square$

b) Claim: a non-negative continuous function f is Riemann integrable on the compact set $[0,T]$.

Proof:

Direct proof:

A continuous function is uniformly continuous on compact sets. So let $\varepsilon > 0$ and $\delta > 0$ such that

$\forall x, y \in [0,T]: |x-y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$.

Choose a partition of the interval $[0,T]$ with $\frac{\delta}{2}$-mesh: $t_0 = 0 < t_1 = \frac{\delta}{2} < t_2 = \delta < \dots < t_{n-1} < t_n = T$ (let us agree that the distance $t_n - t_{n-1} \le \frac{\delta}{2}$, so it need not be $t_n - t_{n-1} = \frac{\delta}{2}$), for some $n \in \mathbb{N}$. Define functions

$g: [0,T] \to \mathbb{R}$, $g(x) = \sum_{k=1}^{n} \inf_{t_{k-1} \le y \le t_k} f(y) \, \mathbb{1}_{[t_{k-1}; t_k[}(x)$ with the last term $\propto \mathbb{1}_{[t_{n-1}; t_n]}(x)$;

$\tilde{g}: [0,T] \to \mathbb{R}$, $\tilde{g}(x) = \sum_{k=1}^{n} \sup_{t_{k-1} \le y \le t_k} f(y) \, \mathbb{1}_{[t_{k-1}; t_k[}(x)$ with the last term

$\alpha \mathbb{1}_{[t_{n-1};t_n]}(x)$.

Now (same denotation as in part a)) $\tilde{g} \in A, g \in B$, so

$I(g) \leq J^-(g)$

$I(\tilde{g}) \geq J^+(g)$

and as for every $h \in B, h' \in A$: $h \leq g \leq h' \Rightarrow I(h) \leq I(h')$
and thus $J^-(g) \leq J^+(g)$, so we have

$I(g) \leq J^-(g) \leq J^+(g) \leq I(\tilde{g})$

and we can calculate

$$I(\tilde{g}) - I(g) = \sum_{k=1}^{n} \left(\sup_{t_{k-1} \leq y \leq t_k} g(y) - \inf_{t_{k-1} \leq y \leq t_k} g(y) \right) \text{length}([t_{k-1}; t_k[)$$

$$\leq \sum_{k=1}^{n} \varepsilon (t_k - t_{k-1}) = \varepsilon (T - 0) = \varepsilon T$$

So, letting $\varepsilon \to 0$ we see that we must have

$J^-(g) = J^+(g)$

also g is Riemann integrable.

Hence the claim is true. $\square$

c) Let $g(x) = \mathbb{1}_{\mathbb{Q}}(x)$ where $\mathbb{Q}$ are the rationals.

Claim: g is Borel measurable, but not Riemann integrable on $[0,T]$.

Proof:

Direct proof:

As $\mathbb{Q}$ is countable, $\mathbb{Q} \in \mathcal{B}(\mathbb{R})$, so $\mathbb{1}_{\mathbb{Q}}$ is Borel-measurable

Also we notice that, on a compact interval $[0,T]$, if $g: [0,T] \to \mathbb{R}$ takes finitely many values and is piecewise continuous, then if

$g \le \mathbb{1}_{\mathbb{Q}} \Rightarrow g \le 0$, and if

$g \ge \mathbb{1}_{\mathbb{Q}} \Rightarrow g \ge 1$

as $\mathbb{Q}$ is dense. So $J^-(g) \le 0 = I(0)$ and $J^+(g) \ge T = I(1)$.

Thus $J^-(g) \ne J^+(g)$ and g is not Riemann integrable.

Hence the claim is true. $\square$

d) Claim: For the Lebesgue integral we have

$$I(g) = \int_0^T g(x)\,dx = 0$$

Proof:

Direct proof:

$$I(g) = \int_0^T g(x)\,dx = \int_0^T \mathbb{1}_{\mathbb{Q}}(x)\,dx = m(\mathbb{Q}) = 0$$

where m is Lebesgue measure, and $m(\mathbb{Q}) = 0$ as $\mathbb{Q}$ is countable.

Hence the claim is true. $\square$

5.

a) Claim: (Chebyshev inequality): For a random variable X with $X(\omega) \ge 0$ P-almost surely,

$$P(X > t) \le \frac{E_P(X)}{t} \quad \forall t > 0$$

Proof:

Direct proof:

Note that $0 \le t\,\mathbb{1}_{\{X(\omega) > t\}} \le X$, so

$$E_P(X) \ge E_P(t\,\mathbb{1}_{\{X(\omega) > t\}}) = t\,P(X > t)$$

$\Rightarrow P(X>t) < \frac{E_P(X)}{t} \quad \forall t>0$

Hence the claim is true. $\square$

b) Claim: (Chentsov inequality):

$$P(X>t) \leq \inf_{\theta>0} \{\exp(-\theta t) E_P(\exp(\theta X))\}$$

Proof:

Direct proof:

Note that for any $\theta>0$, $X>t \Leftrightarrow e^{\theta X} > e^{\theta t}$, also for $\theta>0$,

$$X>t \Leftrightarrow e^{\theta X} > e^{\theta t} \Leftrightarrow e^{-\theta t} e^{\theta X} \geq 1 = \mathbb{1}_{\{X>t\}}$$

- So we have

$$e^{-\theta t} e^{\theta X} \geq \mathbb{1}_{\{X>t\}} \geq 0$$

$$\Rightarrow E_P(e^{-\theta t} e^{\theta X}) = \exp(-\theta t) E_P(\exp(\theta X)) \geq E_P(\mathbb{1}_{\{X>t\}}) = P(X>t)$$

for all $\theta>0$, so

$$P(X>t) \leq \inf_{\theta>0} \{\exp(-\theta t) E_P(\exp(\theta X))\}$$

Hence the claim is true. $\square$

c) Consider a random variable $N(\omega)$ with Poisson(λ) distribution, where $\lambda>0$ is the parameter and

$$P_\lambda(N=k) = e^{-\lambda} \frac{\lambda^k}{k!}, \; k \in \mathbb{N} = \{0,1,2,\dots\}$$

Knowing that $E(\exp(\theta N)) = \exp(\lambda(e^\theta - 1))$ (computed in the exercise sheet number 5), let us use the Chentsov inequality to bound from above the probability $P_\lambda(N>t)$, for $t>0$

Set $t>0$. We have (b)-part)

$$P_\lambda(N>t) \leq \inf_{\theta>0} \{\exp(-\theta t) E_P(\exp(\theta N))\} = \inf_{\theta>0} \{\exp(-\theta t)$$

$\exp(\lambda(e^{\theta}-1))\} = \min\{\exp(-0\cdot t)\exp(\lambda(e^{0}-1)), \exp(-(\ln \frac{t}{\lambda})t)$

$\exp(\lambda(e^{\ln \frac{t}{\lambda}}-1))\} = \min\{1, \exp(t - t\ln t + t\ln\lambda - \lambda)\}$

$= \min\{1, \exp(-t\ln t + (1+\ln\lambda)t - \lambda)\}$

where we note that $\theta \mapsto \exp(-\theta t)\exp(\lambda(e^{\theta}-1))$ is a C^{∞}-function in whole $\mathbb{R}$, and

$$\lim_{\theta\to\infty} \exp(-\theta t + \lambda e^{\theta} - \lambda) = \infty$$

so in the interval $]0,\infty[$, the infimum of $\exp(-\theta t)\exp(\lambda(e^{\theta}-1))$ is either its value at $\theta=0$ (which is 1) or its value at a zero of derivative:

$$\frac{d}{d\theta}\exp(-\theta t + \lambda e^{\theta} - \lambda) = (-t + \lambda e^{\theta})\exp(-\theta t + \lambda e^{\theta} - \lambda) = 0$$

$\Leftrightarrow t = \lambda e^{\theta}$

$\Leftrightarrow \theta = \ln \frac{t}{\lambda}$

We note that $\exp(-t\ln t + (1+\ln\lambda)t - \lambda) \xrightarrow[t\to\infty]{} 0$.

PROBLEMS 7.

In the problems all random variables live in a probability space $(\Omega, \mathcal{F}, P)$.

1.

On a probability space $(\Omega, \mathcal{F}, P)$, let $(X_n)_{n\in\mathbb{N}}$ be a sequence of exponential random variables such that

$$P(X_1 > t_1; \ldots; X_n > t_n) = \exp(-\lambda \sum_{i=1}^{n} t_i) \quad \forall n\in\mathbb{N},\ t_1; \ldots; t_n \geq 0$$

where $\lambda > 0$ is a parameter.

a) Claim: the random variables $(X_n)_{n\in\mathbb{N}}$ are independent under P.

Proof:

Direct proof:

Let $I = \{i_1; \ldots; i_n\} \subseteq \mathbb{N}$ be a finite subset, $n\in\mathbb{N}$. Let $i_1 < i_2 < \cdots < i_n$. Now let us note that

$A := \{]t, \infty[\mid t \in \mathbb{R}\}$ (Obviously a π-system, obviously $\sigma(A) = \sigma(\{]-\infty, t] \mid t\in\mathbb{R}\})$)

is a π-system spanning $\mathcal{B}(\mathbb{R})$ (Gasbarra: Lecture notes in probability theory fall semester 2015 (moniste): Example 1.1.2).

Then the set

$$\tilde{B} := \{A_1 \times A_2 \times \cdots \times A_n \mid A_i \in A\ \forall i = 1, \ldots, n\}$$

is a π-system spanning $\mathcal{B}(\mathbb{R}^n) = \mathcal{B}(\mathbb{R})^{\otimes n}$. Now we use Gasbarra: Lecture notes in probability theory fall semester 2015 (moniste): Proposition 1.1.1 to establish that from

(1) $P(X_{i_1} \in A_1; \ldots; X_{i_n} \in A_n) = P(X_{i_1} \in A_1) \cdots P(X_{i_n} \in A_n) \quad \forall A_1; \ldots; A_n \in A$

it follows that

(2) $P(X_{i_1} \in B_1; \ldots; X_{i_n} \in B_n) = P(X_{i_1} \in B_1) \cdots P(X_{i_n} \in B_n) \quad \forall B_1; \ldots; B_n \in \mathcal{B}(\mathbb{R})$

Here on the left-hand side of (1),(2) we have the probability measure P, and on the right hand side we have the product probability measure $P_{X_{i_1}} \otimes P_{X_{i_2}} \otimes \cdots \otimes P_{X_{i_n}}$.

So, let $t_1; \dots; t_n \geq 0$ and let us calculate

$$P(X_{i_1} > t_1; \dots; X_{i_n} > t_n) = P(X_1 > 0; X_2 > 0; \dots; X_{i_1} > t_1; \dots; X_{i_n} > t_n)$$

$$= \exp(-\lambda(0 + 0 + \dots + t_1 + \dots + t_n)) = \exp(-\lambda t_1) \cdots \exp(-\lambda t_n)$$

$$= P(X_{i_1} > t_1) \cdots P(X_{i_n} > t_n)$$

where we note $P(X > 0) = 1$ for exponential random variable X (and we can ignore 0-probability events) and $P(X > t) = e^{-\lambda t}$ for exponential random variable X with parameter λ (the cumulative distribution function of X is $F(t) = 1 - e^{-\lambda t}$).

Hence, by (1),(2) above, we have that $(X_n)_{n \in \mathbb{N}} \overset{P}{\perp\!\!\!\perp}$.

Hence the claim is true. □

b) Let

$$Y_n(\omega) := \min\{X_1(\omega); X_2(\omega); \dots; X_n(\omega)\}, \omega \in \Omega$$

Let us compute $P(Y_n > t)$, and let us compute also the probability density function of Y_n.

$$P(Y_n > t) = P(X_1 > t; X_2 > t; \dots; X_n > t) = e^{-\lambda \sum_{i=1}^{n} t} = e^{-\lambda n t}$$

Thus the cumulative distribution function F_Y for Y is

$$F_Y(t) = P(Y \leq t) = 1 - P(Y > t) = 1 - e^{-\lambda n t}$$

Thus the probability density function for Y is

$$f_Y(t) = \frac{dF_Y(t)}{dt} = \lambda n e^{-\lambda n t}$$

c) Let $X_n^*(\omega) := \max\{X_1(\omega); X_2(\omega); \dots; X_n(\omega)\}, \omega \in \Omega$.

Let us compute $P(X_n^* \leq t)$, Let us compute also the probability density function of X_n^*.

$$P(X_n^* \le t) = P(X_1 \le t; X_2 \le t; \dots; X_n \le t) = P(X_1 \le t) \cdots P(X_n \le t)$$

$$= (1 - P(X_1 > t)) \cdots (1 - P(X_n > t)) = (1 - e^{-\lambda t}) \cdots (1 - e^{-\lambda t}) = (1 - e^{-\lambda t})^n$$

The cumulative distribution function for X_n^* is $F_{X_n^*}(t) = P(X_n^* \le t) = (1 - e^{-\lambda t})^n$; thus the probability density function is

$$f_{X_n^*}(t) = \frac{dF_{X_n^*}(t)}{dt} = n(1 - e^{-\lambda t})^{n-1} \cdot \lambda e^{-\lambda t} = \lambda n e^{-\lambda t}(1 - e^{-\lambda t})^{n-1}$$

d) Let us compute $\lim_{n \to \infty} P(\lambda X_n^* \le t + \log(n))$.

$$\lim_{n \to \infty} P(\lambda X_n^* \le t + \log(n)) = \lim_{n \to \infty} P(X_n^* \le \frac{1}{\lambda}(t + \log(n))) = \lim_{n \to \infty} (1 - \exp(-\lambda \cdot \frac{1}{\lambda}(t + \log(n))))^n = \lim_{n \to \infty} (1 - \frac{e^{-t}}{n})^n = \exp(-e^{-t})$$

where we note $(1 + \frac{x}{n})^n \underset{n \to \infty}{\to} e^x$.

2.

Let us consider a sequence of random variables $(U_k)_{k \in \mathbb{N}}$ such that for $\forall t_1, \dots, t_n \in [0,1]$,

$$P(U_1 \le t_1; \dots; U_n \le t_n) = \prod_{k=1}^{n} t_k.$$

a) Claim: $(U_k)_{k \in \mathbb{N}}$ are independent and uniformly distributed on $[0,1]$.

Proof:

Direct proof:

Uniformly distributed: $F_{U_k}(t) = P(U_k \le t) = P(U_1 \le 1; U_2 \le 1; \dots; U_k \le t) = 1 \cdot 1 \cdot \dots \cdot t = t \quad \forall t \in [0,1] \Rightarrow U_k$ is uniformly distributed, $k \in \mathbb{N}$.

Independent: For $I = \{i_1 < i_2 < \dots < i_n\} \subseteq \mathbb{N}$ we have

$$P(U_{i_1} \le t_1; \dots; U_{i_n} \le t_n) = P(U_1 \le 1; U_2 \le 1; \dots; U_{i_1} \le t_1; \dots; U_{i_n} \le t_n)$$

$= 1 \cdot 1 \cdot \ldots \cdot t_1 \cdot \ldots \cdot t_n = P(U_{i_1} \leq t_1) \cdots P(U_{i_n} \leq t_n)$

for all $t_1, \ldots, t_n \in [0,1]$; as $\{]-\infty, t] \cap [0,1] \mid t \in \mathbb{R}\}$ is a π-system spanning $\mathcal{B}([0,1])$, the claim follows as in exercise 1 part a).

Hence the claim is true. □

b) Let us consider $\bar{U}_n(\omega) := \max\{U_1(\omega); \ldots; U_n(\omega)\}$, $\omega \in \Omega$.

Let us compute the cumulative distribution function of $\bar{U}_n$; $F_{\bar{U}_n}(t) = P(\bar{U}_n \leq t)$.

$$F_{\bar{U}_n}(t) = P(\bar{U}_n \leq t) = P(U_1 \leq t; \ldots; U_n \leq t) = t^n.$$

c) Claim: $\lim_{n \to \infty} \bar{U}_n(\omega) = 1$ P-almost surely.

Proof:

Direct proof:

Let us pick $\omega \in \Omega$. As $(\bar{U}_n(\omega))_{n \in \mathbb{N}}$ is a monotonously increasing, almost surely bounded sequence ($P(U_k \in [0,1]) = 1$ $\forall k \in \mathbb{N} \Rightarrow P(\bar{U}_n \in [0,1]) = 1$ $\forall n \in \mathbb{N} \Rightarrow P(\forall n \in \mathbb{N}: \bar{U}_n \in [0,1]) = 1$ by monotone converge of measure; $\lim_{n \to \infty} 1 = 1$), the limit exists almost surely. As $P(\forall n \in \mathbb{N}: \bar{U}_n \in [0,1]) = 1$, the limit is a number $q \in [0,1]$. If $q < 1$, we have $\exists n' \in \mathbb{N}: \bar{U}_n(\omega) \leq q + \frac{1}{2}(1-q)$ $\forall n \geq n'$. We have

$$P(\forall n \geq n': \bar{U}_n \leq q + \tfrac{1}{2}(1-q)) = \lim_{n \to \infty} P(\bar{U}_{n'} \leq q + \tfrac{1}{2}(1-q); \ldots; \bar{U}_n \leq q + \tfrac{1}{2}(1-q))$$

$$= \lim_{n \to \infty} P(U_1 \leq 1; U_2 \leq 1; \ldots; U_{n'} \leq q + \tfrac{1}{2}(1-q); \ldots; U_n \leq q + \tfrac{1}{2}(1-q)) = \lim_{n \to \infty} 1 \cdot 1 \cdot \ldots$$

$$\cdot (q + \tfrac{1}{2}(1-q)) \cdot \ldots \cdot (q + \tfrac{1}{2}(1-q)) = \lim_{n \to \infty} (q + \tfrac{1}{2}(1-q))^{n-n'} = 0$$

So (by monotone convergence of measure), $P(\exists n' \in \mathbb{N}\ \forall n \geq n': \bar{U}_n \leq q + \frac{1}{2}(1-q)) = 0$. Thus $P(\lim_{n \to \infty} \bar{U}_n < 1) = P(\bigcup_{k=1}^{\infty} \{\lim_{n \to \infty} \bar{U}_n < 1 - \frac{1}{k}\}) \leq \sum_{k=1}^{\infty} 0 = 0$ (as $P(\lim_{n \to \infty} \bar{U}_n = q < 1) = 0$ $\forall q < 1$).

Thus $\lim_{n \to \infty} \bar{U}_n(\omega) = 1$ P-almost surely. Hence the claim is true. □

d) Let $\underline{U}_n(\omega) := \min\{U_1(\omega); \ldots; U_n(\omega)\}$, $\omega \in \Omega$.

Let us compute the cumulative distribution function of $\underline{U}_n$, $F_{\underline{U}_n}(t) = P(\underline{U}_n \le t)$.

$F_{\underline{U}_n}(t) = P(\underline{U}_n \le t) = 1 - P(\underline{U}_n > t) = 1 - P(U_1 > t; \ldots; U_n > t) = 1 - P(U_1 > t)$

$\cdots P(U_n > t) = 1 - (1-t)^n$

e) Claim: $\lim_{n\to\infty} \underline{U}_n(\omega) = 0$ P-almost surely.

Proof:

Direct proof:

Let us note that $V_n = 1 - U_n$ has the same distribution as U_n:

$P(V_n \le t) = P(1 - U_n \le t) = P(U_n \ge 1-t) = 1 - P(U_n < 1-t) = 1 - \lim_{k\to\infty} P(U_n \le 1 - t - \frac{1}{k})$

$= 1 - \lim_{k\to\infty}(1 - t - \frac{1}{k}) = t = P(U_n \le t) \quad \forall t \in [0,1]$

(from which it follows that $P(V_n \in B) = P(U_n \in B)$ $\forall B \in \mathcal{B}([0,1])$ as $\{]-\infty, t] \cap [0,1] \mid t \in \mathbb{R}\}$ is a π-system generating $\mathcal{B}([0,1])$). Thus $\underline{U}_n = \min\{U_1; \ldots; U_n\}$ has the same distribution as $\min\{V_1; \ldots; V_n\} = \min\{1-U_1; \ldots; 1-U_n\} = 1 - \max\{U_1; \ldots; U_n\} = 1 - \overline{U}_n$ (this follows easily from Gasbarra: Lecture notes in probability theory fall semester 2015 (moniste): Definition 2.1.1; given a set of similarly distributed random variables, any Borel function of them is similarly distributed).

Now as $(\underline{U}_n)_{n\in\mathbb{N}}$ is a monotonously decreasing, almost surely bounded sequence ($\underline{U}_n \in [0,1]$ almost surely), we have

$\lim_{n\to\infty} \underline{U}_n = \inf_{n\in\mathbb{N}} \underline{U}_n$ (which is a well-defined random variable)

Now $P(\inf_{n\in\mathbb{N}} \underline{U}_n \ge t) = P(\bigcap_{n\in\mathbb{N}}\{\underline{U}_n \ge t\}) \overset{(*)}{=} \lim_{n\to\infty} P(\underline{U}_n \ge t) = \lim_{n\to\infty} P(1 - \overline{U}_n \ge t)$

$\overset{(*)}{=} P(\bigcap_{n\in\mathbb{N}}\{1 - \overline{U}_n \ge t\}) = P(\inf_{n\in\mathbb{N}}(1 - \overline{U}_n) \ge t) = P(1 - \sup_{n\in\mathbb{N}} \overline{U}_n \ge t)$; $t \in [0,1]$

which implies that $\inf_{n\in\mathbb{N}} \underline{U}_n$ and $1 - \sup_{n\in\mathbb{N}} \overline{U}_n$ are similarly distributed, now $\sup_{n\in\mathbb{N}} \bar{U}_n = \lim_{n\to\infty} \bar{U}_n$ and thus

$P(\lim_{n\to\infty} \underline{U}_n = 0) = P(1 - \lim_{n\to\infty} \bar{U}_n = 0) = 1.$

Hence $\lim_{n\to\infty} \underline{U}_n = 0$ P almost surely.

Thus the claim is true.

In $(*)$ above we use monotone convergence of measure, $\{\underline{U}_{n+1} \geq \varepsilon\} \subseteq \{\underline{U}_n \geq \varepsilon\}$ et cetera. □

3.
a) Let $X, X_n, n\in\mathbb{N}$ be random variables such that $X_n(\omega) \xrightarrow[n\to\infty]{} X(\omega)$ P-almost surely.

Claim: the Cesaro mean converges P-almost surely to X

$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i(\omega) = X(\omega)$ P-almost surely

Proof:

Direct proof:

Let $\omega\in\Omega$ be such that $X_n(\omega) \xrightarrow[n\to\infty]{} X(\omega)$. Let $\varepsilon>0$, and let

$n'\in\mathbb{N}$ be such that $|X_n(\omega) - X(\omega)| < \varepsilon$ $\forall n \geq n'$. We have for $n\geq n'$

$$\left|\frac{1}{n}\sum_{i=1}^{n} X_i(\omega) - X(\omega)\right| \leq \left|\frac{1}{n}\sum_{i=1}^{n'-1} X_i(\omega)\right| + \left|\frac{1}{n}\sum_{i=n'}^{n}(X_i(\omega) - X(\omega))\right| + \left|\frac{1}{n}(n'-1)X(\omega)\right|$$

$$\leq \frac{1}{n}\left|\sum_{i=1}^{n'-1} X_i(\omega)\right| + \frac{1}{n}\sum_{i=n'}^{n}|X_i(\omega) - X(\omega)| + \frac{n'-1}{n}|X(\omega)| \leq \frac{1}{n}\left|\sum_{i=1}^{n'-1} X_i(\omega)\right|$$

$$+ \frac{1}{n}(n-n'+1)\varepsilon + \frac{n'-1}{n}|X(\omega)| \xrightarrow[n\to\infty]{} \varepsilon.$$

Thus $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i(\omega) = X(\omega)$ P-almost surely.

Hence the claim is true. □

b) Let us assume now that $E_P(|X_n - X|) \to 0$ as $n\to\infty$ (without assuming P-almost sure convergence).

Claim: the Cesaro mean is converging in $L^1(P)$, that is,

$\lim_{n\to\infty} E_P(|\frac{1}{n}\sum_{i=1}^{n} X_i - X|) = 0$

Proof:

Direct proof:

Let $\varepsilon > 0$ and let $M \in \mathbb{N}$ be such that

$E_P(|X_n - X|) < \varepsilon \quad \forall n \geq M$

Then by triangle inequality

$$|\frac{1}{n}\sum_{i=1}^{n} X_i - X| \leq \frac{1}{n}\sum_{i=1}^{n} |X_i - X| = \frac{1}{n}\sum_{i=1}^{M} |X_i - X| + \frac{1}{n}\sum_{i=M+1}^{n} |X_i - X|$$

$\forall n \geq M$, where the inequalities are preserved after taking the expectation, so that

$$E_P(|\frac{1}{n}\sum_{i=1}^{n} X_i - X|) \leq \frac{1}{n}\sum_{i=1}^{M} E_P(|X_i - X|) + \frac{1}{n}\sum_{i=M+1}^{n} E_P(|X_i - X|)$$

$$\leq \frac{1}{n}\sum_{i=1}^{M} E_P(|X_i - X|) + \frac{1}{n}(n-M)\varepsilon \xrightarrow[n\to\infty]{} \varepsilon.$$

Thus $\lim_{n\to\infty} E_P(|\frac{1}{n}\sum_{i=1}^{n} X_i - X|) = 0.$

Hence the claim is true. $\square$

5.

Let $X, (X_n)_{n\in\mathbb{N}}$ be random variables on a probability space $(\Omega, \mathcal{F}, P)$.

Claim: if $\forall \varepsilon > 0: \sum_{n=1}^{\infty} P(|X_n - X| > \varepsilon) < \infty$, it follows that $\lim_{n\to\infty} X_n(\omega)$

$= X(\omega)$ P-almost surely.

Proof:

Direct proof:

Let $\forall \varepsilon > 0: \sum_{n=1}^{\infty} P(|X_n - X| > \varepsilon) < \infty.$

Let us note that

$\{\omega \in \Omega \mid X_n(\omega) \not\xrightarrow[n\to\infty]{} X(\omega)\} = \bigcup_{k\in\mathbb{N}} \{\omega\in\Omega \mid |X_n(\omega) - X(\omega)| > \frac{1}{k}$ infinitely often for $n\}$

which follows from the definition of a limit: $X_n(\omega) \not\xrightarrow[n\to\infty]{} X(\omega)$ $\Leftrightarrow \exists \varepsilon>0\ \forall n'\in\mathbb{N}\ \exists n \geq n' : |X_n(\omega) - X(\omega)| > \varepsilon \Leftrightarrow \exists \varepsilon > 0 : |X_n(\omega) - X(\omega)| > \varepsilon$ infinitely often in n.

Now, defining

$A_n^k := \{\omega\in\Omega \mid |X_n(\omega) - X(\omega)| > \frac{1}{k}\}; k, n \in \mathbb{N}$

we have by assumption that $\forall k\in\mathbb{N}: \sum_{n=1}^{\infty} P(A_n^k) < \infty$. Thus by Borel-Cantelli lemma (Gasbarra: Lecture notes in probability theory fall semester 2015 (moniste): Lemma 5.1.1)

$P(\{\omega\in\Omega \mid \omega\in A_n^k$ for infinitely many $n\}) = 0 \quad \forall k \in \mathbb{N}$

Thus

$P(\{\omega\in\Omega \mid X_n(\omega) \not\xrightarrow[n\to\infty]{} X(\omega)\}) = P(\bigcup_{k\in\mathbb{N}} \{\omega\in\Omega \mid \omega\in A_n^k$ for infinitely many $n\}) \leq \sum_{k\in\mathbb{N}} 0 = 0$

Thus $\lim_{n\to\infty} X_n(\omega) = X(\omega)$ P-almost surely.

Hence the claim is true. □

5.

Let us consider a random variable X with $E_P(|X|) < \infty$.

Claim: $E_P(|X| \mathbb{1}(|X| > n)) = \int_\Omega |X(\omega)| \mathbb{1}(|X(\omega)| > n) P(d\omega) \to 0$ as $n\to\infty$.

Proof:

Direct proof:

$E_P(|X|) = E_P(|X|(\mathbb{1}(|X| \leq n) + \mathbb{1}(|X| > n))) = E_P(|X| \mathbb{1}(|X| \leq n)) + E_P(|X| \mathbb{1}(|X| > n))$

$$\overset{E_p(|X|)<\infty}{\Longleftrightarrow} E_p(|X|\mathbb{1}(|X|>n)) = E_p(|X|) - E_p(|X|\mathbb{1}(|X|\leq n)) \quad \forall n \in \mathbb{N}:$$

$$\lim_{n\to\infty} E_p(|X|\mathbb{1}(|X|>n)) = E_p(|X|) - \lim_{n\to\infty} E_p(|X|\mathbb{1}(|X|\leq n))$$

$$= E_p(|X|) - E_p(|X|) = 0$$

by monotone convergence theorem ($0 \leq |X|\mathbb{1}(|X|\leq n) \leq |X|\mathbb{1}(|X|\leq n+1) \underset{n\to\infty}{\uparrow} |X|$).

Hence the claim is true. $\square$