Teoria miary
WPPT IIr. semestr zimowy 2009
Wykłady 8, 9 i 10.
Dystrybuanty i miary borelowskie na prostej
Funkcje mierzalne
Zbieżność prawie wszędzie i zbieżność według miary

5 i 12/11/09

DYSTRYBUANTY i MIARY BORELOWSKIE NA PROSTEJ

Definicja 1: Funkcję $F : \mathbb{R} \to \mathbb{R}$ nazwiemy *dystrybuantą* jeśli spełnia ona 3 warunki:
1. F jest niemalejąca,
2. F jest prawostronnie ciągła (tzn. $\forall_{t\in\mathbb{R}} \lim_{s\to t^+} F(s) = F(t)$)
3. $\lim_{t\to 0^-} F(t) = 0$

Definicja 2: Niech μ będzie miarą borelowską (czyli określoną na zbiorach borelowskich) na prostej rzeczywistej $\mathbb{R}$, która **na przedziałach właściwych przyjmuje wartości skończone**. *Dystrybuantą miary* μ nazywamy funkcję $F_\mu : \mathbb{R} \to \mathbb{R}$ określoną wzorem

$$F_\mu(t) = \begin{cases} \mu([0,t]); & t \geq 0 \\ -\mu((0,t)); & t < 0 \end{cases}$$

Łatwo pokazuje się, że F_μ jest dystrybuantą sensie definicji 1. W obliczeniach korztysta się z ciągłości miary z dołu i z góry. Miara co prawda nie musi być skończona, ale w potrzebnych tu obliczeniach wszystkie zbiory można zawrzeć w jednym przedziale właściwym, który z założenia ma miarę skończoną, zatem można korzystać z ciągłości miary z góry.

Twierdzenie 1. *Każda dystrybuanta (w sensie definicji 1) jest dystrybuantą (w sensie definicji 2) pewnej miary μ skończonej na przedziałach właściwych. Odwzorowanie $\mu \mapsto F_\mu$ jest bijekcją (funkcją różnowartościową i ,,na") ze zbioru wszystkich miar borelowskich na prostej skończonych na przedziałach właściwych na zbiór wszystkich dystrybuant (w sensie definicji 1).*

Szkic dowodu: Dana jest dystrybuanta F. Chcemy skonstuować miarę μ taką, że $F_\mu = F$ (w każdym punkcie t). To znaczy, chcemy aby $\mu([0,t]) = F(t)$ dla t nieujemnych oraz $\mu((t,0) = -F(t))$ dla t ujemnych. Z tego wynika łatwo, że ta miara musi spełniać $\mu((a,b]) = F(b) - F(a)$ dla dowolnych par liczb $a < b$. Konstrukcja takiej miary μ przebiega dokładnie tak samo jak konstrukcja miary Lebesgue'a, z zastąpieniem warunku $\lambda([a,b)) = b - a$ warunkiem $\mu((a,b]) = F(b) - F(a)$, co praktycznie niczego nie zmienia (zmiana rodzajów końców przedziału jest również nieistotna w konstrukcji). Przypomnijmy kroki konstrukcji (z uwzględnieniem tych zmian):
Krok I:
Na rodzinie $\mathcal{A}$ przedziałów postaci $(a,b]$ $(a,b \in \mathbb{R},\ a < b)$ określamy funkcję ,,masy" ν wzorem

$$\nu((a,b]) = F(b) - F(a).$$

Krok II:
Określamy miarę zewnętrzną

$$\bar{\mu}(A) = \inf\{\sum_{n=1}^{\infty} \nu((a_n, b_n]) : A \subset \bigcup_{n=1}^{\infty} (a_n, b_n]\}.$$

Krok III:
Niech Σ oznacza sigma-ciało zbiorów mierzalnych w sensie Carathéodory'ego względem miary zewnętrznej $\bar{\mu}$. Obcięcie $\bar{\mu}$ do Σ jest miarą. Oznaczamy je przez μ.

To jest koniec konstrukcji. Trzeba sprawdzić, że dowolna półprosta postaci $(-\infty, a]$ ($a \in \mathbb{R}$) jest mierzalna względem miary zewnętrznej $\bar{\mu}$, oraz że dla $a < b$ $\bar{\mu}((a, b]) = F(b) - F(a)$. Dowody są prawie identyczne jak dla miary Lebesgue'a. Jedyne różnice, to takie, że wszędzie na liczby a, b, a_n, itp. w dowodzie trzeba nałożyć funkcję F (czyli zastąpić przez $F(a), F(b), F(a_n)$, itp) oraz w ostatniej części dowodu, ze względu na zmianę kierunków nawiasów, trzeba zmienić kierunek ,,indukcji", to znaczy zacząć od punktu b i własność (*) określić wzorem:

$$(*) \qquad x < b \text{ oraz } \exists_{\mathbb{N}_x \subset \mathbb{N}} \Big(\inf_{n \in \mathbb{N}_x} a_n = x \;\&\; \sum_{n \in \mathbb{N}_x} (F(b_n) - F(a_n)) = F(x) - F(a) \Big)$$

Następnie pokazać, że zbiór X punktów spełniających (*) osiąga minimum (a nie, jak poprzednio, maximum), i że minimum to jest mniejsze bądź równe a.

W ten sposób udowodnimy istnienie miary μ takiej, że $F_\mu = F$. Wzajemna jednoznaczność odwzorowania $\Phi : \mu \mapsto F_\mu$ wynika teraz z prostej obserwacji, że powyżej właśnie określiliśmy odwzorowanie odwrotne do Φ przyporządkowujące każdej dystrybuancie F miarę μ taką, że $\Phi(\mu) = F$. Wiadomo, że jeśli istnieje odwzorowanie odwrotne do danego, to odwzorowanie to jest bijekcją. To kończy dowód twierdzenia 1. $\square$

Przykłady miar borelowskich (skończonych na przedziałach) i ich dystrybuant.
1. Miara Lebesgue'a: Oczywiście na przedziałach właściwych przyjmuje ona wartości skończone. Policzmy jej dystryuantę:

$$F_\lambda(t) = \begin{cases} \lambda([0, t]) = t & t \geq 0 \\ -\lambda((t, 0)) = -(0 - t) = t & t < 0. \end{cases}$$

Czyli w obu przypadkach po prostu $F_\lambda(t) = t$.

2. Miara z gęstością: Niech f będzie funkcją nieujemną całkowalną (po każdym przedziale właściwym) w sensie Riemanna. Określamy

$$F(t) = \begin{cases} \int_0^t f(s)\, ds & t \geq 0 \\ -\int_t^0 f(s)\, ds & t < 0 \end{cases}$$

Łatwo sprawdza się, że F jest dystrubuantą oraz że odpowiadająca jej miara μ spełnia

$$\mu(\langle a, b \rangle) = \int_a^b f(s)\, ds$$

niezależnie od tego jakie są końce przedziału od a do b; pod znaczek $\langle$ można podstawić dowolnie (lub [i podobnie pod $\rangle$ dowolnie) lub].

3. Miara czysto atomowa: Niech $\{x_n : n \in \mathbb{N}\}$ będzie podzbiorem przeliczalnym (lub skończonym) prostej. Załóżmy dla uproszczenia, że jest to zbiór *dyskretny*, tzn., że odległości między punktami x_n są co najmniej pewne $r > 0$. Niech $\{p_n : n \in \mathbb{N}\}$ będzie ciągiem liczb dodatnich. Określamy miarę μ (na wszystkich podzbiorach prostej – w szczególności na zbiorach borelowskich) wzorem

$$\mu(A) = \sum_{n: x_n \in A} p_n.$$

Ponieważ w każdym przedziale właściwym jest co najwyżej skończenie wiele punktów x_n (bo odległości między nimi są co najmniej $r > 0$), to miara ta jest skończona na takich przedziałach. Jest to *miara czysto atomowa* a każdy punkt x_n nazywamy jej *atomem* o *masie* p_n. Dystrybuanta tej miary jest funkcją schodkową o następującym przebiegu: pomiędzy punktami x_n jest to funkcja stała, a w każdym punkcie x_n ma ona skok (w górę) o p_n. Trzeba tę funkcję tak ustawić aby spełniała warunek 3 w definicji 1 (granica lewostronna w zerze równa zero). Jeśli w zerze jest ona ciągła, to po prostu ustawiamy ją tak aby $F(0) = 0$. Jeśli w zerze jest skok (co oznacza, że zero jest atomem), to $F(0)$ będzie równe masie atomu w zerze.

Ogólnie, jeśli dla dowolnej miary (na dowolnej przestrzeni) punkt x spełnia $\mu(\{x\}) > 0$, to x nazywamy *atomem* tej miary. Dla miar borelowskich na prostej, skończonej na przedziałach właściwych atomy ZAWSZE odpowiadają punktom nieciągłości dystrybuanty. Ponieważ dystrybuanta jest niemalejąca i prawostronnie ciągła, to wszelkie nieciągłości są typu ,,skok w górę" i wartość w takim punkcie ,,przyklejona jest" do granicy prawostronnej. Masa atomu odpowiada wielkości skoku. Innymi słowy mamy ogólny wzór dla dowolnej dystrybuanty:

$$\mu(\{x\}) = F(x) - \lim_{t \to x^-} F(t)$$

(Wzór ten zgadza się nawet jeśli x nie jest atomem – po obu stronach jest wtedy zero, gdyż dystrybuanta jest tam ciągła.) Wzór ten wynika natychmiast z ciągłości z góry:

$$\{x\} = \bigcap_{t_n \to x^-}\!\!\downarrow (t_n, x] \quad \text{zatem} \quad \mu(\{x\}) = \lim_{t_n \to x^-} (F(x) - F(t_n)) = F(x) - \lim_{t \to x^-} F(t).$$

Dowolna miara, której dystrybuanta jest ciągła w każdym punkcie nazywa się albo *bezatomowa* albo po prostu *ciągła*. Na przykład miary z gęstością (przykład 1) są ciągłe, gdyż funkcja określona jako całka z funkcji całkowalnej jest zawsze ciągła. Okazuje się, że NIE KAŻDA MIARA CIĄGŁA JEST MIARĄ Z GĘSTOŚCIĄ. Dystrybuanty miar z gęstością są nie tylko ciągłe, ale *absolutnie ciągłe* (definicję tego pojęcia pomijamy). W każdym razie jest tak, że wśród funkcji ciągłych funkcje absolutnie ciągłe stanowią zdecydowaną mniejszość. Miary odpowiadające dystrybuantom ciągłym lecz nie absolutnie ciągłym (a takich jest najwięcej) nazywamy *miarami singularnymi* (to nie jest całkiem ścisłe, miara taka może rozkładać sięna sumę miary absolutnie ciągłej i singularnej – poznamy to w przyszłości). Przykład miary singularnej jest poniżej.

4. Miara na zbiorze Cantora:
Przedstawimy teraz konstrukcję pewnej dystrybuanty ciągłej związanej ze zbiorem Cantora $\mathfrak{C}$. Kładziemy

$$F(t) = 0 \;\text{ dla } t \leq 0 \quad \text{oraz} \quad f(t) = 1 \;\text{ dla } t \geq 1.$$

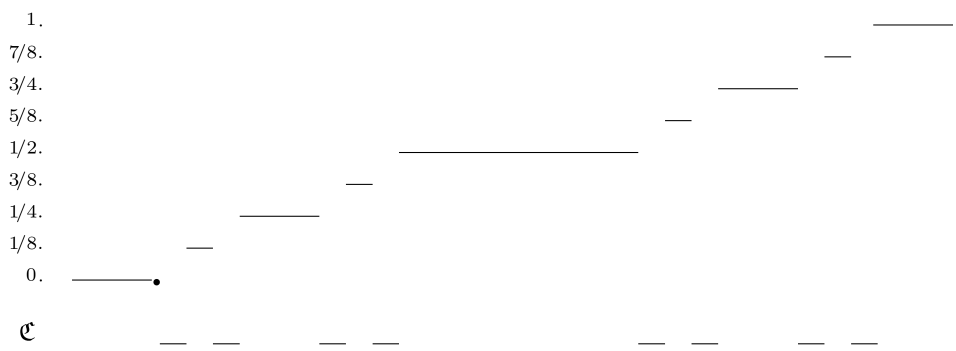
Następnie

$$F(t) = \frac{1}{2} \;\text{ dla } \tfrac{1}{3} < t < \tfrac{2}{3}$$

Dalej

$$F(t) = \frac{1}{4} \;\text{ dla } \tfrac{1}{9} < t < \tfrac{2}{9} \quad \text{oraz} \quad F(t) = \frac{3}{4} \;\text{ dla } \tfrac{7}{9} < t < \tfrac{8}{9}$$

i.t.d. (pomijamy szczegółowy opis indukcyjny). W ten sposób funkcję określimy na DOPEŁNIENIU zbiory Cantora. Widać to na rysunku (kropka lokalizuje punkt $(0,0)$, a kreski na samym dole to zbiór Cantora na osi (reszty osi nie narysowano):



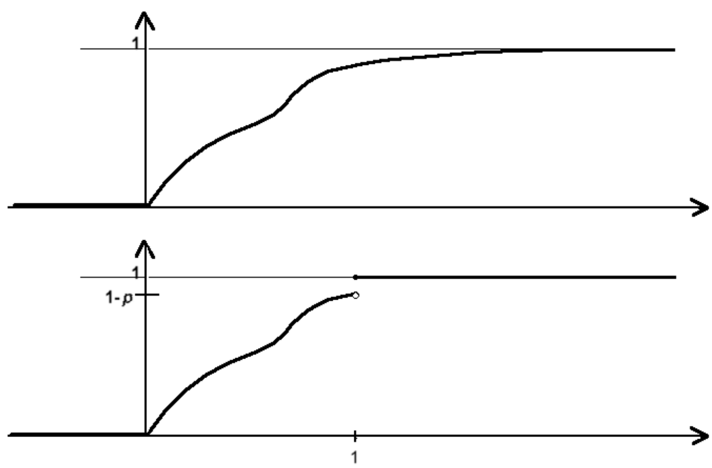
Teraz określamy F w pozostałych punktach $x \in \mathfrak{C}$ wzorem

$$F(x) = \sup\{F(y) : y \notin \mathfrak{C}, y < x\}.$$

Tak zdefiniowana funkcja jest monotoniczna i ciągła na całej prostej i $F(0) = 0$, czyli jest to dystrybuanta pewnej miary bezatomowej μ. Widać, że na każdym przedziale ,,wyrzuconym" ze zbioru Cantora (na przykład $(\frac{1}{9}, \frac{2}{9})$) funkcja F jest stała (a nawet na jego domknięciu), czyli jego miara jest zerowa (bo jest to różnica F na końcach przedziału). Stąd miara dopełnienia zbioru Cantora jest zerowa (jako przeliczalna suma takich przedziałów), czyli miara jest ,,skupiona" na $\mathfrak{C}$. Z kolei miara Lebesgue'a zbioru $\mathfrak{C}$ jest zero, bowiem z odcinka $[0,1]$ wyrzuciliśmy przedziały o łącznej długości 1 (jest to elementarne do sprawdzenia, że $\frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \cdots = 1$). Gdyby μ była miarą z gęstością, to gęstość ta f musiałaby być zero λ-prawie wszędzie na dopełnieniu zbioru Cantora (bo miara tego dopełninia jest zero), zatem niezerowa mogłaby być tylko na $\mathfrak{C}$. Ale ponieważ $\mathfrak{C}$ ma miarę λ zerową, to f byłaby równa zero λ-prawie wszędzie, a całka a takiej funkcji jest zero po każdym przedziale. Z tego wynikałoby, że dystrybuanta F tej miary jest wszędzie równa zero, a przecież tak nie jest! Dokładne uzasadnienie wykorzystanych tu faktów dotyczących całkowania funkcji poznamy na kolejnych wykładach.

5. Miary mieszane. Oprócz miar absolutnie ciągłych, czysto atomowych i singularnych jest jeszcze cała masa miar mieszanych. Są to sumy miar powyższcyh trzech typów. Dystrybuanty takich miar są dowolne: w co najwyżej przeliczalnie wielu punktach występują nieciągłości (skoki w górę) – to są atomy tej miary, w pozostałych punktach dystrybuanta nie musi być stała (schodkowa), może tam być ściśle rosnąca. O ile oddzielenie części atomowej od ciągłej jest stosunkowo proste, o tyle rozkład części ciągłej na część absolutnie ciągłą i singularną jest trudny. W przyszłości poznamy odpowiednie twierdzenia (w ogólniejszym kontekście).

Przykład praktyczny: Wyobraźmy sobie, że mierzymy jakaś zmienną losową nieujemną o rozkładzie ciągłym przyrządem, którego skala kończy się na pewnej wartości (np. na 1) i wszystkie wartości powyżej 1 są pokazywane jako 1. Załóżmy, że wartości powyżej 1 występują z dodatnim prawdopodobieństwem, powiedzmy p. Wtedy zmienna losowa ,,odczyt pomiaru" jest mieszany: dla wartości poniżej 1 jest to ten sam rozkład co mierzonej zmiennej losowej (czyli ciągły) natomiast w 1 występuje atom o masie p, który zastępuje cały przedzial $[1, \infty)$ dla zmiennej mierzonej (1-ka ,,zbiera" miarę z tego przedziału). Rysunek przedstawia dystrybuanty rozkładów mierzonej zmiennej i zmiennej ,,odczyt pomiaru".



FUNKCJE MIERZALNE

Rozważamy dwie przestrzenie mierzalne $(X, \mathcal{F})$ i $(Y, \mathcal{G})$. Później skoncentrujemy się na przypadku, gdy $(Y, \mathcal{G})$ jest prostą rzeczywistą z sigma-ciałem zbiorów borelowskich. Ilekroć w roli Y występuje $\mathbb{R}$ tylekroć za $\mathcal{G}$ bierzemy domyślnie (bez przypominania) sigma-ciało zbiorów borelowskich Σ_B.

Definicja: Odzworowanie $f : X \to Y$ nazywa się *mierzalne* jeśli $f^{-1}(\mathcal{G}) \subset \mathcal{F}$ (tzn. $\forall_{B\in\mathcal{G}}\ f^{-1}(B) \in \mathcal{F}$). Jeśli $Y = \mathbb{R}$ to f nazywać będziemy po prostu funkcją (rzeczywistą) mierzalną.

Przykład: Niech f będzie odwzorowaniem stałym, to znaczy $f(x) = y_0$, gdzie $y_0 \in Y$ jest ustalony. Takie odwzorowanie (lub funkcja) jest zawsze mierzalne, bo $f^{-1}(B)$ dla dowolnego zbioru $B \subset Y$ jest albo zbiorem pustym (jeśli $y_0 \notin B$) albo całym X (jeśli $y_0 \in B$).

Fakt: Jeśli dodatkowo $(Z, \mathcal{H})$ jest przestrzenią mierzalną i $g : Y \to Z$ jest mierzalna, to złożenie $g \circ f : X \to Z$ jest mierzalne.

Dowód: Niech $A \in \mathcal{H}$. Wtedy $(g\circ f)^{-1}(A) = f^{-1}(g^{-1}(A))$. Oznaczając $B = g^{-1}(A)$ mamy $B \in \mathcal{G}$ z mierzalności g, a dalej $(g \circ f)^{-1}(A) = f^{-1}(B) \in \mathcal{F}$ z mierzalności f. $\square$

Uwaga: Do mierzalności odzworowania (funkcji) wystarczy, że $f^{-1}(\mathcal{A}) \subset \mathcal{F}$, gdzie $\mathcal{A}$ jest rodziną zbiorów w Y taką, że $\sigma(\mathcal{A}) = \mathcal{G}$. Dowód tego faktu wymaga indukcji pozaskończonej. Pokazuje się, że dla każdej liczby porządkowej $\alpha < \omega_1$ zachodzi $f^{-1}(\mathcal{A}_\alpha) \subset \mathcal{F}$, gdzie $\mathcal{A}_\alpha$ są rodzinami występującymi w pozsakończonej konstrukcji sigma-ciała $\sigma(\mathcal{A})$.

Przykład: Niech X i Y będą przestrzeniami topologicznymi i rozważajmy w nich sigma-ciała zbiorów borelowskich (generowane przez *topologie*, czyli rodziny zbiorów otwartych). Wtedy mamy następujący fakt:

Jeśli odwzorowanie $f : X \to Y$ jest ciągłe, to jest ono mierzalne.

Dlaczego tak jest? Otóż ciągłość jest równoważna z warunkiem: $f^{-1}(U)$ jest otwarty (w X) dla dowolnego U otwartego w Y. Ale to oznacza, że $f^{-1}(U)$ jest borelowski w X. Ponieważ zbiory otwarte U generują zbiory borelowskie w Y, to z powyższej uwagi wynika mierzalność f.

Wniosek: $f : X \to \mathbb{R}$ jest mierzalna wtedy i tylko wtedy gdy zbiory postaci $\{x : f(x) < t\}$ są mierzalne dla wszystkich $t \in \mathbb{R}$ (wystarczy nawet $t \in Q$ lub w innym zbiorze gęstym). Wynika to z faktu, że $\{x : f(x) < t\} = f^{-1}((-\infty, t))$ oraz z tego, że rodzina półprostych $\{(-\infty, t) : t \in \mathbb{R}\}$ (albo $t \in Q$) generuje sigma-ciało zbiorów borelowskich.

Funkcje rzeczywiste mierzalne można dodawać, mnożyć, brać granice, itp:

Twierdzenie 2: Jeśli f, g, f_n oznaczają rzeczywiste funkcje mierzalne, to funkcje $f + g$, $f - g$, fg, $\frac{f}{g}$, f^g, $\sup_n f_n$, $\inf_n f_n$, $\limsup_n f_n$, $\liminf_n f_n$, $\lim_n f_n$, $\sum_n f_n$ (za każdym razem przy założeniach gwarantujących sens danej funkcji) są funkcjami mierzalnymi. Ponadto funkcje postaci $\sin(f)$, $\exp(f)$, $\ln(f)$, $\log_f(g)$, itp., są mierzalne (też przy odpowiednich założeniach gwarantujących sens).

Dowód: Zaczniemy od funkcji $\ln(f)$. Oczywiście trzeba założyć, że f przyjmuje wartości nieujemne. Wtedy jako przeciwdziedzinę można wziąć przestrzeń $Y = (0, \infty)$. Teraz funkcja $\ln : Y \to \mathbb{R}$ jest określona wszędzie i ciągła, a zatem mierzalna. Zatem $\ln(f)$ jako złożenie funkcji mierzalnych jest mierzalna. Podobnie uporamy się z $\sin(f)$, $\exp(f)$, itp. Natomiast $\log_f(g)$ trzeba zapisać jako $\frac{\ln(f)}{\ln(g)}$ i skorzystać jeszcze z mierzalności ilorazu funkcji mierzalnych (którą wykażemy za chwilę).
Udowodnimy teraz mierzalność sumy $f + g$. Wystraczy pokazać, że zbiory postaci $\{x : f(x) + g(x) < t\}$ są mierzalne. Mamy

$$\{x : f(x)+g(x) < t\} = \{x : f(x) < t-g(x)\} = \{x : \exists_{q\in Q} : f(x) < q < t-g(x)\} = \sum_{q\in Q}\{x : (f(x) < q) \wedge (g(x) < t - q)\} = \sum_{q\in Q}\{x : f(x) < q\} \cap \{x : g(x) < t - q\},$$

a to jest zbiór mierzalny, jako przeliczalna suma przekrojów zbiorów mierzalnych.
Zajmiemy się teraz iloczynem fg. Otóż $fg = \exp(\log(f) + \log(g))$. No i dalej wszystko powinno być jasne. Pozostałe funkcje zapiszmy podobnie, w sposób, który wyjaśnia ich mierzalność:

$$f - g = f + (-1)g, \quad f^g = \exp(g \cdot \log(f)), \quad \frac{f}{g} = f \cdot g^{-1}.$$

Ostatnie rozumowanie dotyczyć będzie $\inf_n f_n$. Funkcję $\sup_n f_n$ zamienimy na $-\inf_n(-f_n)$. Funkcję $\limsup_n f_n$ zapiszemy jako $\inf_n(\sup_{m\geq n} f_m)$, a liminf otrzymamy przez zmianę znaku z limsup. Granica (o ile istnieje) jest równa granicy górnej, a szereg jest granicą sum skończonych.
A więc rozważmy $g = \inf_n f_n$. Dla $t \in \mathbb{R}$ mamy

$$\{x : g(x) < t\} = \{x : \inf_n f_n(x) < t\} = \{x : \exists_n f_n(x) < t\} = \bigcup_n \{x : f(x) < t\},$$

a to jest suma przeliczalna zbiorów z założenia mierzalnych. $\square$

Funkcje proste.
Definicja: *Funkcją charakterystyczną* zbioru $A \subset X$ nazywamy funkcję $\mathbf{1}_A : X \to \mathbb{R}$ określoną wzorem

$$\mathbf{1}_A(x) = \begin{cases} 1; & x \in A \\ 0; & x \notin A \end{cases}.$$

Funkcją prostą (w postaci ogólnej) nazywamy każdą funkcję o postaci

$$f = \sum_{i=1}^{n} a_n \mathbf{1}_{A_n},$$

gdzie $A_1, A_2, \ldots, A_n$ są jest dowolnym skończonym ($n \in \mathbb{N}$ jest również dowolne) układem podzbiorów X oraz $a_1, a_2, \ldots, a_n$ jest dowolnym układem liczby rzeczywistych. (Czyli funkcja prosta = kombinacja liniowa funkcji charakterystycznych.)

Dodatkwo powiemy, że funkcja prosta jest zapisana *w postaci rozłącznej*, jeśli zbiory $A_1, \dots, A_n$ są parami rozłączne, a jeśli dodatkowo liczby a_i parami różne, oraz zbiory A_i dają w sumie całe X, to powiemy, że funkcja jest w postaci *kanonicznej.* Postać kanoniczna jest jednoznaczna (z dokładnością do przestawienia składników sumy).

Fakt: *Funkcja f jest prosta wtedy i tylko wtedy gdy przyjmuje ona tylko skończenie wiele wartości. Każda funkcja prosta daje się zapisać w postaci kanonicznej.*

Dowód: Niech f będzie funkcją prostą zadaną w postaci ogólnej wzorem

$$f = \sum_{i=1}^{n} a_n \mathbf{1}_{A_n},$$

Oczywiście funkcja taka przyjmuje tylko skończenie wiele wartości (są to sumy skończonych układów wybranych spośród liczb a_i). Weźmy więc teraz dowolną funkcję f przyjmującą tylko skończenie wiele wartości, np. $b_1, b_2, \dots, b_m$ (zapisujemy je bez powtórzeń, czyli są to liczby parami różne). Pokażemy, że jest to funkcja prosta od razu w postaci kanonicznej. Po prostu zdefiniujmy B_i jako $f^{-1}(\{b_i\})$ (czyli $\{x : f(x) = b_i\}$). Zbiorów tych jest skończenie wiele (m), są one parami rozłączne (jako przeciwobrazy zbiorów rozłącznych) oraz dają w sumie całe X (bo wzięliśmy wszystkie wartości funkcji). Zatem funkcja

$$g = \sum_{i=1}^{m} b_i \mathbf{1}_{B_i}$$

jest prosta w postaci kanonicznej. Pokażemy, że $f = g$ w każdym punkcie. Niech $x \in X$. Istnieje jedyne i takie, że $f(x) = b_i$, czyli $x \in B_i$. Wtedy $g(x) = b_i$ bo w sumie definiującej g tylko ten jeden składnik jest niezerowy. Czyli $f(x) = b_i = g(x)$. Aby wykazać jednoznaczność przedstawienia kanonicznego wystarczy zauważyć, że w tej formie zawsze liczby b_i reprezentują wszystkie (różne) wartości funkcji prostej, a zbiory A_i muszą być przeciwobrazami tych liczb. □

Po zapisaniu funkcji prostej w postaci kanonicznej bardzo łatwo stwierdzić kiedy jest ona mierzalna:

Fakt: *Funkcja prosta w postaci kanonicznej $f = \sum_{i=1}^{n} a_n \mathbf{1}_{A_n}$ jest mierzalna wtedy i tylko wtedy, gdy wszystkie zbiory A_i $(i = 1, \dots, n)$ są mierzalne.*

Dowód: Jeśli f jest mierzalna, to zbiory A_i jako przeciwobrazy zbiorów jednopunktowych (a więc domkniętych, zatem borelowskich) $\{b_i\}$ muszą być mierzalne. Na odwrót. Załóżmy, że te zbiory są mierzalne. Do mierzalności f wystarczy sprawdzić, że przeciwobraz dowolnej półprostej $(-\infty, t)$ jest mierzalny. Ale

$$f^{-1}((-\infty, t)) = \{x : f(x) < t\} = \{x : f(x) = b_i \text{ oraz } b_i < t\} = \\ = \bigcup_{i : b_i < t} f^{-1}(\{b_i\}) = \bigcup_{i : b_i < t} A_i.$$

Jako suma skończona zbiorów (z założenia) mierzalnych jest to zbiór mierzalny. □

UWAGA: Dla f w postaci ogólnej mamy tylko implikację w jedną stronę: jeśli zbiory A_i są mierzalne, to f jest mierzalna. W drugą stronę implikacja nie zachodzi. Na przykład funkcja stała 0 może być zapisana jako $a_1\mathbf{1}_{A_1} + a_2\mathbf{1}_{A_2}$, gdzie $a_1 = -a_2$ oraz $A_1 = A_2$. Zbiór A_1 może tu być zupełnie dowolny, na przykład niemierzalny.

Funkcje proste stanowią istotny element w teorii funkcji mierzalnych. Oto powód:

Twierdzenie 3: *Każda funkcja mierzalna f jest granicą (w każdym punkcie) ciągu funkcji prostych mierzalnych. W każdym punkcie ciąg ten jest od pewnego numeru niemalejący. Jeśli f jest ograniczona z dołu, to ciąg ten jest niemalejący (od początku). Jeśli f jest ograniczona (i z dołu i z góry) to zbieżność ta jest jednostajna.*

Dowód: Ustalmy funkcję mierzalną $f : X \to \mathbb{R}$. Ustalmy $n \in \mathbb{N}$. Podzielmy przedział $[-n, n)$ na $2n2^n$ przedziałów długości 2^{-n} postaci $[\frac{k}{2^n}, \frac{k+1}{2^n})$. Nazwijmy je (i przenumerujmy) jako I_i $(i = 1, \ldots, n^2)$. Dorzućmy do tego dwie półproste $I_0 = (-\infty, -n)$ i $I_{2n2^n+1} = [n, \infty)$. W ten sposób $\mathbb{R}$ jest rozłączną sumą $2n2^n + 2$ rozłącznych przedziałów (a więc zbiorów borelowskich). Określmy $A_i = f^{-1}(I_i)$ $(i = 0, 1, \ldots, 2n2^n + 1)$. Są to zbiory mierzalne, rozłączne, dają w sumie całe X, a z mierzalności f wynika, że są one mierzalne. Niech $a_i = \min I_i$ (oprócz $i = 0$) oraz $a_0 = -n - 1$. Zdefiniujmy n-tą funkcję prostą wzorem

$$f_n = \sum_{i=0}^{2n2^n+1} a_i \mathbf{1}_{A_i}$$

(oczywiście jest to od razu postać kanoniczna). Sprawdzimy, czy $f_n(x)$ dąży (i czy monotonicznie) do $f(x)$ w dowolnym punkcie $x \in X$. Dla dostatecznie dużych n zachodzi $-n < f(x) < n$. Dla takich n, $f(x)$ należy do jednego z przedziałów I_i o numerze od 1 do $2n2^n$ (przedziały te mają długość $\frac{1}{2^n}$). Wtedy $x \in f^{-1}(I_i) = A_i$. Czyli $f_n(x) = \min I_i$. Ponieważ $f(x) \in I_i$, to po pierwsze $f_n(x) \leq f(x)$, po drugie różnica między tymi liczbami $f(x) - f_n(x)$ nie przekracza $\frac{1}{2^n}$. Ponieważ $\frac{1}{2^n}$ maleje do zera, to widać, że $f_n(x) \to f(x)$. Jeśli teraz zwiększymy n, to ,,nowy" przedział I_i zawierający x będzie podzbiorem ,,starego" (to jest powód, dla którego dzielimy $[-n, n)$ na $2n2^n$ przedziałów, a nie na $2n^2$; przy podziale na $2n^2$ (czyli każdego przedziału długości 1 na n kawałków) ,,nowe" przedziały nie muszą być zawarte w ,,starych"), zatem wartość $\min I_i$ przypisana punktowi x przez funkcję f_n może co najwyżej wzrosnąć (infimum po podzbiorze jest nie mniejsze). Tak będzie nawet dla x należących do ostatniego przedziału $[n, \infty)$, gdyż tu również stosujemy wzór $f_n(x) = \min I_i$ i tu również ,,nowy" przedział zawierający x jest zawarty w ,,starym". (Tyle, że tu nie będzie jeszcze oszacowania $f(x) - f_n(x) < \frac{1}{2^n}$.) To oznacza, że rozważany ciąg funkcji f_n jest w każdym punkcie x niemalejący od pewnego miejsca (od tak dużego n, że $-n < f(x)$).

Jeśli teraz f jest ograniczona z dołu, to warunek $-n < f(x)$ jest spełniony dla dostatecznie dużego n jednocześnie dla wszystkich x i można ciąg funkcji zacząć od takiego numeru n. Taki ciąg będzie niemalejący od początku w każdym punkcie.

Wreszcie jeśli f jest również ograniczona z góry, to dla dostatecznie dużego n spełniony będzie warunek $-n < f(x) < n$ jednocześnie dla wszystkch x. Wtedy będziemy mieli spełniony warunek $f(x) - f_n(x) < \frac{1}{2^n}$ jednocześnie dla wszystkich x, a to daje zbieżność jednostajną. □

ZBIEŻNOŚĆ PRAWIE WSZĘDZIE I ZBIEŻNOŚĆ WEDŁUG MIARY

W naszej przestrzeni $(X, \mathcal{F})$ wprowadzimy teraz miarę μ. Będziemy mówić wyłącznie o funkcjach rzeczywistych (niekoniecznie mierzalnych).

Definicja: Powiemy, że funkcje f i g są *równe μ-prawie wszędzie* jeśli istnieje zbiór A miary zero taki, że $f(x) = g(x)$ dla wszystkich $x \in A^c$ (funkcje są sobie równe poza pewnym zbiorem miary zero).

Przykład: Równość prawie wszędzie, w zależności od miary, może faktycznie oznaczać równość poza ,,pomijalnie małym" zbiorem, ale nie zawsze. Weźmy miarę $\mu = \delta_x$ (atomową o jednym atomie w x o masie 1). Wtedy do równości μ-prawie wszędzie dwóch funkcji wystarcza ich równość w jednym punkcie x.
Inny przykład: W przestrzeni X z miarą μ ustalmy pewien podzbiór mierzalny B miary dodatniej skończonej. Teraz zdefiniujemy tzw. *miarę warunkową (unormowaną)* na B. Mianowicie, dla dowolnego mierzalnego zbioru A kładziemy

$$\mu_B(A) = \frac{\mu(A \cap B)}{\mu(B)}.$$

Nietrudno sprawdza się, że jest to dobrze określona miara (na tym samym sigmaciele $\mathcal{F}$ co miara μ) oraz, że $\mu_B(X) = \mu_B(B) = 1$ (a więc jest to miara *probabilistyczna*). Ponadto to miara μ_B dopełnienia zbioru B (i każdego jego podzbioru) jest zero. Widać więc, że do równości μ_B-prawie wszędzie jakiejś pary funkcji wystarczy aby zgadzały się one na B, a poza B mogą się one od siebie dowolnie różnić.

Zbieżność prawie wszędzie.
Wracamy do sytuacji, gdy na $(X, \mathcal{F})$ ustalona jest jedna miara μ. Od teraz nie będziemy rozważać innych miar. W tej sytuacji w zwrocie ,,μ-prawie wszędzie" będziemy pomijać μ i mówić po prostu ,,prawie wszędzie" (i skracać ,,p.w."). Aby usprawnić wypowiadanie warunku ,,prawie wszędzie" umawiamy się, że zbiór $X' \subset X$ jest *pełny* jeśli jego dopełnienie jest miary zero. W przestrzeniach z miarą skończoną zbiór X' jest pełny wtedy i tylko wtedy, gdy $\mu(X') = \mu(X)$, ale gdy miara jest nieskończona, to w X może być dużo zbiorów o mierze nieskończonej (a więc równej $\mu(X)$), które nie są pełne. Wtedy musimy niestety popatrzeć na dopełnienie (czy jest ono miary zero).

Definicja: Powiemy, że ciąg f_n funkcji (rzeczywistych na X) *jest zbieżny prawie wszędzie* do funkcji f (co zapiszemy $\lim_n f_n = f$ μ-p.w., lub $\lim_n f_n = f$ p.w. lub jeszcze krócej $f_n \to f$ p.w.) jeśli istnieje zbiór pełny X' taki, że dla każdego $x \in X'$ zachodzi zbieżność

$$\lim_n f_n(x) = f(x).$$

UWAGA: Funkcja graniczna jest wtedy jednoznaczna z dokładnością do równości prawie wszędzie, tzn., jeśli g jest również granicą p.w. tego samego ciągu f_n, to $f = g$ p.w. (to jest elementarne ćwiczenie i wynika z jednoznaczności granicy ciągu liczbowego).

Podamy teraz definicję zbieżności według miary:

Definicja: Powiemy, że ciąg f_n *jest zbieżny według miary* do funkcji f (co zapiszemy $f_n \xrightarrow{\mu} f$, albo μ-$\lim_n f_n = f$) jeśli zachodzi następujący warunek:

$$\forall_{\epsilon>0} \quad \lim_{n\to\infty} \mu\{x : |f_n(x) - f(x)| \geq \epsilon\} = 0$$

(implicite zakłada się tu, że zbiory pod znakiem miary są mierzalne). Dla usprawnienia dalszych wywodów wprowadzimy oznaczenie

$$A_{n,\epsilon} = \{x : |f_n(x) - f(x)| \geq \epsilon\}.$$

Przykłady: Niech $X = \mathbb{R}$ z miarą Lebesgue'a. Niech $f_n = \mathbf{1}_{[n,\infty)}$. Oczysiście $\lim_n f_n(x) = 0$ wszędzie (nie tylko prawie wszędzie), gdyż dla każdego $x \in \mathbb{R}$ od pewnego miejsca mamy $x \notin [n, \infty)$. Z drugiej strony dla $\epsilon = 1$ mamy równość $A_{n,\epsilon} = \{x : |f_n(x) - 0| \geq \epsilon\} = [n, \infty)$ i miara Lebesgue'a tego zbioru jest nieskończona, a więc miary te nie dążą do zera. Zatem w tym przykładzie jest zbieżność p.w., ale nie według miary.

Teraz niech $X = [0, 1]$ (też z miarą Lebesgue'a obciętą do odcinka (patrz powyżej pojęcie miary warunkowej). Niech I_n będzie ciągiem odcinków wypisanym poniżej:

$$\begin{aligned}
&[0, \tfrac{1}{2}), [\tfrac{1}{2}, 1], \\
&[0, \tfrac{1}{4}), [\tfrac{1}{4}, \tfrac{2}{4}), [\tfrac{2}{4}, \tfrac{3}{4}), [\tfrac{3}{4}, 1], \\
&[0, \tfrac{1}{8}), [\tfrac{1}{8}, \tfrac{2}{8}), [\tfrac{2}{8}, \tfrac{3}{8}), [\tfrac{3}{8}, \tfrac{4}{8}), [\tfrac{4}{8}, \tfrac{5}{8}), [\tfrac{5}{8}, \tfrac{6}{8}), [\tfrac{6}{8}, \tfrac{7}{8}), [\tfrac{7}{8}, 1], \\
&\vdots
\end{aligned}$$

Niech $f_n = \mathbf{1}_{I_n}$. Widać, że każdy punkt wpada do nieskończenie wielu zbiorów I_n i do nieskończenie wielu z nich nie wpada. Zatem ciąg $f_n(x)$ nie jest zbieżny w żadnym punkcie. Czyli żadna funkcja f nie jest granicą p.w. tego ciągu. Z drugiej strony, jeśli za f przyjmiemy funkcję stale równą zero, to dla dowolnego $\epsilon > 0$ (mniejszego lub równego 1) jest równość: $A_{n,\epsilon} = \{x : |f_n(x) - f(x)| \geq \epsilon\} = I_n$. Ponieważ miary tych zbiorów dążą do zera, zachodzi zbieżność według miary (do $f \equiv 0$).

Podamy teraz twierdzenia, które jednak dają wynikanie jedego rodzaju zbieżności z drugiego (ale przy dodatkowych założeniach):

Twierdzenie 4: *Jeśli miara μ jest skończona, a funkcje f_n i f są mierzalne, to ze zbieżności $f_n \to f$ prawie wszędzie wynika zbieżność według miary.*

Dowód: Niech $f_n \to f$ p.w. Ustalmy $\epsilon > 0$. Dla każdego $x \in X'$ (to jest ten zbiór pełny na którym jest zbieżność) istnieje n_x takie, że $\forall_{n\geq n_x} |f_n(x) - f(x)| < \epsilon$. Dla $n \in \mathbb{N}$ niech $X_n = \{x : n_x \leq n\}$. Oczywiście zbiory X_n tworzą ciąg rosnący i jego suma zawiera X': $\biguplus_n X_n \supset X'$. Zbiory X_n są mierzalne, co widać, gdy się je zapisze tak:

$$X_n = \{x : \forall_{m\geq n} |f_n(x) - f(x)| < \epsilon\} = \bigcap_{m\geq n} \{x : |f_n(x) - f(x)| < \epsilon\}.$$

Trzeba teraz zauważyć, że funkcje $|f_n - f|$ są mierzalne. Wynika to z tego, że $f_n - f$ jest różnicą funkcji mierzalnch a moduł (jako funkcja rzeczywista zmiennej rzeczywistej) jest funkcją ciągłą, a więc mierzalną.
Z ciągłości miary z dołu mamy zatem

$$\lim_n \mu(X_n) = \mu\Big(\bigcup_n^{\uparrow} X_n\Big) \geq \mu(X') = \mu(X).$$

Z tego wynika, że miary dopełnień zbiorów X_n dążą do zera (tu korzystamy ze skończoności miary). Z drugiej strony warunek $|f_n(x) - f(x)| \geq \epsilon$ implikuje, że $n_x > n$ czyli, że $x \notin X_n$. Zatem zbiory $A_{n,\epsilon} = \{x : |f_n(x) - f(x)| \geq \epsilon\}$ w definicji zbieżności według miary są podzbiorami dopełnień zbiorów X_n, a więc ich miary tym bardziej dążą do zera, a to jest właśnie zbieżność weług miary. $\square$

Teraz twierdzenie w pewnym sensie przeciwne. Nie zakładmy w nim ani skończoności miary, ani mierzalności występujących tu funkcji (tylko, jak zwykle, mierzalność zbiorów pod znakiem miary).

Twierdzenie 5: *Jeśli zachodzi zbieżność według miary $f_n \xrightarrow{\mu} f$ z dodatkowym silniejszym warunkiem*

$$\forall_{\epsilon>0}\ \exists_{n_\epsilon} \quad \sum_{n \geq n_\epsilon} \mu(A_{n,\epsilon}) < \infty,$$

to ciąg f_n jest zbieżny do f prawie wszędzie.

UWAGA: Oczywiście warunek ten jest silniejszy niż zwykła zbieżność według miary, bowiem sumowalność ciągu implikuje zbieżność jego wyrazów do zera (ale nie odwrotnie!).
Do dowodu potrzebny jest lemat zwany Lematem Borela-Cantelli'ego:

Lemat Borela-Cantelli'ego: *Dane są zbiory mierzalne A_n takie, że $\sum_n \mu(A_n) < \infty$. Wtedy $\mu(\limsup A_n) = 0$*

Dowód lematu: Przypomnijmy, że $\limsup A_n = \bigcap_n \bigcup_{m \geq n} A_m$ jest zbiorem tych x, które należą do nieskończenie wielu spośród zbiorów A_n. Dla każdego n_0 mamy $\limsup A_n \subset \bigcup_{m \geq n_0} A_m$, zatem z przeliczalnej podaddytywności miary,

$$\mu(\limsup A_n) \leq \mu\Big(\bigcup_{m \geq n_0} A_m\Big) \leq \sum_{m \geq n_0} \mu(A_m).$$

Ostatnia suma jest ogonem szeregu zbieżnego, zatem jej wartość zbiega do zera gdy $n_0 \to \infty$. Stąd $\mu(\limsup A_n) \leq 0$ (czyli oczywiście $= 0$). $\square$

Dowód twierdzenia 5: Zastanówmy się na jakim zbiorze NIE zachodzi zbieżność $f_n(x) \to f(x)$. Otóż aby tak było, musi istnieć $\epsilon > 0$ (który można zastąpić liczbą $\frac{1}{k}$, gdzie $k \in \mathbb{N}$) taki, że warunek $|f_n(x) - f(x)| \geq \epsilon$ zachodzi dla nieskończenie wielu n (to jest wprost zaprzeczenie definicji granicy ciągu). Czyli x należy wtedy do granicy górnej zbiorów $A_{n,\epsilon} = \{x : |f_n(x) - f(x)| \geq \epsilon\}$. Zatem zbieżność nie zachodzi na zbiorze

$$\bigcup_\epsilon \limsup_n A_{n,\epsilon},$$

gdzie suma po ϵ jest *de facto* przeliczalna (gdyż można ją zastąpić sumą po k, gdzie w roli ϵ wystąpi $\frac{1}{k}$). Zatem wystarczy pokazać, że dla każdego $\epsilon > 0$ zbiór $\limsup A_{n,\epsilon}$ jest miary zero. Wtedy suma przeliczalna też będzie miary zero, a poza tą sumą zachodzi już zbieżność $f_n(x) \to f(x)$, a więc jest to zbieżność prawie wszędzie.
Ale założyliśmy, że dla każdego $\epsilon > 0$ miary zbiorów $A_{n,\epsilon}$ tworzą ciąg od pewnego miejsca sumowalny. Zatem z Lematu Borela-Cantelli'ego wynika, że (dla każdego ϵ) $\mu(\limsup_n A_{n,\epsilon}) = 0$. (W lemacie zakłada się sumowalność całego szeregu, ale to nie ma znaczenia, gdyż granica górna ciągu zbiorów nie zmieni się, gdy z ciągu tego usuniemy skończenie wiele początkowych wyrazów.) No i o to nam chodziło. Dowód jest skończony. $\square$

Powyższe twierdzenie ma bardzo ważny wniosek, który stosuje się już bez żadnych dodatkowych założeń:

Wniosek: Jeśli $f_n \xrightarrow{\mu} f$, to istnieje podciąg n_k taki, że $f_{n_k} \to f$ μ-p.w.

Dowód: Wystraczy skonstruować podciąg, który spełni założenia twierdzenia 5. Oto jak go skonstruować: Wybierzmy ciąg ϵ_k malejący do zera. Dla ϵ_1 szukamy n_1 takiego, aby zbiór A_{n_1,ϵ_1} miał miarę mniejszą od 2^{-1} (takie n_1 istnieje, gdyż z założenia o zbieżności według miary, miary zbiorów A_{n,ϵ_1} dążą po n do zera). Następnie szukamy n_2 takiego, żeby $\mu(A_{n_2,\epsilon_2}) < 2^{-2}$. I tak dalej. Uzyskujemy podciąg n_k o własności $\mu(A_{n_k,\epsilon_k}) < 2^{-k}$. Teraz ustalmy dowolny $\epsilon > 0$. Istnieje k_ϵ takie że $\epsilon_k < \epsilon$ dla wszystkich $k \geq k_\epsilon$. Wtedy $A_{n,\epsilon_k} \supset A_{n,\epsilon}$ dla każdego n (wystarczy popatrzeć na definicję zbioru $A_{n,\epsilon}$ i zauważyć, że zależy on malejąco od ϵ). W szczególności $A_{n_k,\epsilon_k} \supset A_{n_k,\epsilon}$. Zatem

$$\sum_{k=k_\epsilon}^{\infty} \mu(A_{n_k,\epsilon}) \leq \sum_{k=k_\epsilon}^{\infty} \mu(A_{n_k,\epsilon_k}) \leq \sum_{k=k_\epsilon}^{\infty} 2^{-k}.$$

Po prawej jest szereg sumowalny, a więc i po lewej też. To kończy dowód. $\square$

UWAGA: Dowód wniosku jest nieco prostszy w przypadku przestrzeni skończonej. Założenie twierdzenia 5 jest wtedy równoważne z warunkiem, że CAŁY szereg (od $n = 1$ a nie tylko od $n = n_\epsilon$) jest sumowalny, bowiem ze skończoności miary, suma skończenie wielu początkowych wyrazów szeregu jest skończona i nie wpływa na jego sumowalność. Jednak w przestrzeni o mierze nieskończonej, brakujące wyrazy mogą być nieskończone i wtedy suma całego szerego (mimo sumowalności od pewnego miejsca) może nie zachodzić. To uproszczenie przenosi się do dowodu wniosku. Pomijam szczegóły, gdyż nie zmienia to sformułowania wniosku.