Prawo Serii w ujęciu matematycznym

Tomasz Downarowicz
Instytut Matematyki i Informatyki
Politechnika Wrocławska

12 stycznia, 2011

Każdy słyszał zapewne o prawie serii w rozumieniu potocznym. Najczęściej mówi się o nim żartobliwie, gdy w naszym życiu natkniemy się na pozornie niezależne powtórzenia zdarzeń identycznych lub do siebie podobnych, albo przynajmniej posiadających wspólne cechy (jak na przykład to, że są dla nas korzystne - wtedy mówimy o dobrej passie, lub niekorzystne – wtedy mówimy, że *nieszczęścia chodzą parami*). Prawo to jednak manifestuje się czasem tak wyraźnie, że trudno już myśleć o nim w kategoriach żartobliwych i zaczynamy doszukiwać się w nim przejawu jakiejś niewyjaśnionej siły występującej w przyrodzie, sprawa zaczyna pachnieć zjawiskami nadprzyrodzonymi, a w skrajnych przypadkach (zwłaszcza serii nieszczęść) – może budzić pierwotny zabobonny lęk i wzmagać wiarę w niektóre przesądy.

Wśród środowisk naukowych zainteresowanie zjawiskami typu serii i nadzwyczajnymi zbiegami okoliczności (tzw. koincydencjami) pojawiło się na przełomie XIX i XX wieku i trwało mniej więcej do końca lat 60-tych. Później nastąpiło pewne nasycenie możliwości gromadzenia nowych argumentów i dyskusje na ten temat, mające zresztą głównie charakter filozoficzno-akademicki, jak i zainteresowanie całą tą tematyką zdecydowanie osłabło. Powracała ona jedynie na łamach książek i czasopism popularno-naukowych lub wręcz paranaukowych, co spowodowało jeszcze wyraźniejsze odcięcie się od niej szanujących się środowisk naukowych.

W ostatnich latach, w ramach teorii ergodycznej, badając wraz z moim francuskim kolegą Yves'em Lacroix tzw. asymptotyczne rozkłady czasów powrotu dla długich cylindrów, udowodniliśmy pewne twierdzenia związane z pojęciem entropii, które opisują zjawiska do złudzenia przypominające te interpretowane potocznie jako prawo serii. Przy pewnej dozie swobody w interpretacji rzeczywistości można nawet uważać, że wyniki te w jakiś sposób przekładają się na występowanie prawa serii w niektórych typach zjawisk. Dlatego, pomimo pewnej rezerwy i początkowo chłodnego przyjęcia ze strony autorytetów nauki, postanowiliśmy w prezentacji naszych wyników nawiązywać do potocznego prawa serii i nie stronić od kojarzenia tych dwóch sposobów rozumienia tego „prawa". Staramy się jednak przy tym wyraźnie zaznaczać, że istnieją dwa odrębne światy: świat zjawisk rzeczywistych i świat modeli matematycznych. Nasze *ergodyczne* *prawo serii* stosuje się wyłącznie do drugiego z nich. Jego stosowalność do rzeczywistości pozostawiamy na ogół ocenie naszych czytelników i słuchaczy.

1 Potoczne prawo serii

Zacznijmy więc może od krótkiego wprowadzenia do potocznego prawa serii, składającego się z kilku prostych przykładów z życia oraz rysu historycznego debaty naukowej, po których przejdziemy ergodycznego prawa serii, podając niezbędne definicje pojęć matematycznych z teorii ergodycznej i procesów stochastycznych, sformułowania twierdzeń i szkic jednego z dowodów.

Zjawisko prawa serii w pierwszej kolejności zanotowali hazardziści uprawiający gry losowe. Obserwowali oni tak zwane dobre i złe passy. Sztandarowym przykładem jest tu przypadek Charlesa Wellsa, który w roku 1891 jednej nocy kilkakrotnie rozbił bank (100 tys. franków – w tamtych czasach – majątek!), grając w ruletkę w jednym ze znanych kasyn w Monte Carlo. Rozbicie banku jest zdarzeniem niezwykle rzadkim, a Wellsowi przydarzało się ono raz za razem. Charakterystyczne jest przy tym to, że kolejne liczby i kolory wypadające w grze w ruletkę uważane są za zdarzenia kompletnie od siebie niezależne, tak więc nie można upatrywać wyjaśnienia wygranych w stosowaniu jakiegoś systemu. Na temat tego zdarzenia napisano piosenkę "The man who broke the Bank at Monte Carlo", a nawet nakręcono film (pod tym samym tytułem).

Prawo serii występuje często w historii badań naukowych. Nierzadko bywa tak, że jakiś problem naukowy pozostaje otwarty przez dziesiątki lat, a kiedy zostaje rozwiązany, okazuje się, że dokonały tego mniej więcej w tym samym czasie, niezależnie od siebie i nie wiedząc o sobie, dwa lub więcej zespoły badawcze w różnych zakątkach świata. Tu jednak można mieć pewne wątpliwości co do niezależności tych odkryć. Nieraz bowiem jakieś wcześniejsze odkrycie naukowe otwiera drogę do rozwiązania i wtedy zbieżności w czasie pojawienia się kilku rozwiązań nie można uznać za przypadkową.

Istnieją inne, zupełnie niewiarygodne przykłady serii identycznych lub podobnych zdarzeń, gdzie trudno dopatrzeć się związku przyczynowego. Wiele z nich znaleźć można w literaturze dotyczącej tzw. zjawisk niewyjaśnionych. Są to najczęściej zbieżności życiorysów nie spokrewnionych ze sobą osób (często noszących te same nazwiska), niezwykłe podobieństwo okoliczności śmierci bliźniąt, powtórzenia niemal identycznych sekwencji zdarzeń w życiu jednej osoby (*deja vu*), itp. Pomijając kwestię wiarygodności tych relacji, wskazują one na to, że zdarzenia tak niesamowite, że już prawdopodobieństwo jednego wystąpienia wydaje się znikome, mimo to zdarzają się, i to w seriach po dwa lub kilka naraz, w stosunkowo niewielkich odstępach czasu.

Pierwszym i od razu fanatycznym badaczem prawa serii był działający na przełomie XIX i XX w. austriacki biolog Paul Kammerer (1880-1926). W swojej książce *Das Gesetz der Serie* [K] (czyli właśnie *Prawo serii*) zawarł wiele przykładów z życia swojego i swoich bliskich. Są one podobne do przytoczonych powyżej, choć może mniej spektakularne: kilkakrotne natknięcie się w ciągu jednego dnia na nazwę tej samej, odległej, mało znanej miejscowości, spotkanie tego samego dnia kilku osób o tym samym nazwisku, itp. Kammerer przeprowadzał też eksperymenty polegające na obserwacji przechodniów i notowaniu czasów pojawienia się osób o jakiejś wspólnej cesze (np. w okularach, z laską, itp.), albo na notowaniu dokładnych czasów wejść klientów do jakiegoś sklepu. Przykładowo, przy średniej 60 klientów na godzinę, tylko w stosunkowo małym procencie przedziałów jednominutowych odnotował on dokładnie jednego klienta. W większości takich przedziałów było albo zero, albo więcej niż jeden klient. To, zdaniem Kammerera, dowodziło, że klienci chodzą „seriami". Kammerer, jako biolog, nie miał pojęcia o procesach stochastycznych i w swoim rozumowaniu popełniał oczywisty błąd. Jednak błąd Kammerera naprowadzi nas na to, jak poprawnie zdefiniować matematyczne prawo serii.

Jak już wspomnieliśmy, przykłady serii mieszają się w literaturze z przykładami innych niewiarygodnych zbiegów okoliczności. Pionierską teorię o nieobjętych prawami fizyki siłach prowokujących m. in. serie zdarzeń podobnych i inne koincydencje wysnuli (posiłkując się między innymi przykładami z książki Kammerera), dwaj znani naukowcy, Karol Gustaw Jung (1875-1961, szwajcarski profesor filozofii i psychologii) oraz zdobywca nagrody Nobla w dziedzinie fizyki, Austriak, Wolfgang Pauli (1900-1958). Postulowali oni istnienie w naturze swoistego „przyciągania" w przestrzeni i czasie zjawisk lub obiektów posiadających wspólne cechy (tzw. teoria synchroniczności) [J-P, J].

Pogląd przeciwstawny do teorii synchroniczności głosi, że wszelkie serie, koincydencje i temu podobne, są wynikiem czystego przypadku i nie kryje się za nimi żadna nadprzyrodzona lub niewyjaśniona siła. Pogląd taki wyraził m. in. amerykański matematyk, Warren Weaver (1894-1978), bliski współpracownik Claude Shannona – twórcy teorii informacji i pojęcia entropii. Rzeczywistość przeprowadza w każdej minucie miliony prób losowych polegających na przypadkowym zestawianiu w przestrzeni i czasie różnych liczb, nazwisk, wydarzeń, itp., nie ma zatem niczego nadzwyczajnego w tym, że od czasu do czasu pojawi się seria elementów identycznych lub podobnych, albo koincydencja odbierana przez nas jako niewiarygodna. Każda z nich ma bowiem prawdopodobieństwo niezerowe, więc przy odpowiednio dużej liczbie prób ma prawo, a nawet „obowiązek" kiedyś się pojawić. Nasz problem polega na ignorowaniu sekwencji zdarzeń nie noszących znamion nadzwyczajności, a przez to nie dostrzeganiu globalnej liczby „porażek" jakie towarzyszą wystąpieniu jednego „sukcesu" w postaci niezwykłego zbiegu okoliczności.

W odniesieniu do prawa serii, argumentacja ta wygląda następująco: wystąpienia powtórzeń określonego zjawiska uważane za niezależne od siebie są modelowane przez proces Poissona. Odstępy między sygnałami w takim procesie nie są oczywiście jednakowe – czasem są one większe, czasem mniejsze. Miejsca na osi czasu, gdzie zbiegnie się kilka małych odstępów (a takowe występują z odpowiednią dodatnią frekwencją w czasie) mogą być interpretowane jako „serie". Richard von Mises w swojej książce [vM] wyraźnie sugeruje, że to właśnie takie naturalne serie obserwował Kammerer w większości swoich eksperymentów.

A oto inny argument przeciwników teorii synchroniczności i prawa serii. Wiele zjawisk w przyrodzie występuje seriami z zupełnie racjonalnych powodów. Może to być związane z fizycznym prowokowaniem powtórzeń. Za przykład może tu służyć występowanie zachorowań na jakoś chorobę zakaźną w określonym regionie. Erupcje wulkanów, powodzie czy inne zjawiska w przyrodzie również występują seryjnie w okresach występowania tzw. warunków sprzyjających, które pojawiają się i znikają zgodnie z innym procesem o bardzo wolnym przebiegu. Oczywiście, nikt nie doszukuje się magicznego prawa serii tam, gdzie jego przyczyny są oczywiste. Prawo serii dotyczyć ma zdarzeń, których kolejne pojawienia się uważane są za wzajemnie niezależne. Angielski matematyk i popularyzator nauki Robert Matthews w jednym ze swych esejów wyraża pogląd, że w przypadku wielu takich „niewyjaśnionych" serii ta niezależność jest tylko pozorna. Przy wnikliwszym zbadaniu mogłoby się okazać, że powtórzenia są ze sobą fizycznie powiązane i de facto prowokują się nawzajem. Tylko, że pobieżni obserwatorzy nie zadają sobie trudu, aby tych powiązań dociekać. Przecież zawsze to bardziej podniecające odkrywać zjawiska paranormalne!

Podobne poglądy znaleźć można w pracach matematyka amerykańskiego, Williama Kruskala (1919–2005). Mówiąc w dużym skrócie, poddaje on gruntownej krytyce wszechobecne w modelowaniu matematycznym zjawisk zachodzących w rzeczywistości założenie niezależności pewnych zdarzeń (zob. np. [Kr]). Inny współczesny matematyk, Piercy Diakonis, znany jest z badań nad własnościami realnego procesu rzutów monetą i wskazaniu, że nie jest to, wbrew powszechnemu przekonaniu, proces Bernoulli'ego z wagami 1/2, 1/2 (zob. n.p. [D-H-M]).

Współcześnie o istnieniu prawa serii i synchroniczności, które są czymś więcej niż naturalnym rozmieszczeniem sygnałów w procesie Poissona czy skutkiem wytłumaczalnych zależności, przekonani są jeszcze tylko badacze zjawisk paranormalnych. Francuz, Jean Moisset jest samoukiem i zajmuje się parapsychologią. Jego dorobek liczy kilkanaście książek. Moisset łączy prawo serii ze zjawiskiem psychokinezy i sugeruje nawet, że przy pewnym treningu można sterować prawem serii czy zjawiskiem synchroniczności w korzystny dla siebie sposób (zob. n.p. [M]).

2 Ergodyczne prawo serii

W końcówce XX wieku w ramach teorii ergodycznej powstało kilkanaście prac poświęconych rozkładom czasu powrotu do zbioru miary dodatniej (zob. n.p. [C, A-G]). W szczególności badano rozkład graniczny, gdy średnica lub miara tego zbioru maleje do zera. Jak już wspomniałem na wstępie, wraz z moim francuskim kolegą zajmowaliśmy podobnymi zagadnieniami, w szczególności pytaniem, w jaki sposób dodatnia entropia wpływa na dystrybuantę czasów powrotu długich cylindrów w procesie ergodycznym o skończonej liczbie stanów. Wiadomo było, że w układach specjalnego typu, np. w procesach Bernoulli'ego, albo w procesach o dostatecznie szybkim tempie mieszania, rozkłady te po unormowaniu zbiegają wraz z długością bloku do rozkładu wykładniczego (zob. n.p. [A]). Zastanawialiśmy się, czy wystarczy do tego dodatnia entropia. Z pomocą nieżyjącego już dziś Dana Rudolpha znaleźliśmy na to kontrprzykłady. Ale co nas zastanowiło – we wszystkich tych kontrprzykładach dystrybuanta graniczna unormowanego rozkładu czasu oczekiwania leżała poniżej dystrybuanty rozkładu wykładniczego. Po kilku tygodniach wytężonej pracy udało nam się udowodnić, że istotnie tak musi być. Dodatnia entropia powoduje, że dystrybuanty czasów oczekiwania na długie bloki leżą, z dokładnością do epsilona, poniżej dystrybuanty rozkładu wykładniczego. Dowód jest skomplikowany i zaawansowany, i uważaliśmy go (i do dziś uważamy) za duże osiągnięcie.

W trakcie spisywania szczegółów zaczęliśmy też zastanawiać się, jaka jest interpretacja tego zjawiska. Co oznacza takie położenie dystrybuanty czasu oczekiwania. Bardzo szybko odkryliśmy, że oznacza to po prostu tendencję wystąpień danego bloku wzdłuż osi czasu w realizacji procesu do skupiania się w grupy. Tak jakby pierwsze jego wystąpienie prowokowało kilka następnych, a potem następowała dłuższa kompensacyjna przerwa. Takie zachowanie bloków nasunęło nam skojarzenie z potocznym prawem serii. Zrozumieliśmy, że paradoksalnie, teoria ergodyczna dostarcza argumentów za, a nie przeciw istnieniu prawa serii. Odkryta przez nas interpretacja tak nas zafascynowała, że we wstępie do powstałej pracy naukowej bardzo śmiało do niej nawiązaliśmy. Napisaliśmy wręcz, że nasze wyniki rzucają nowe światło na zjawisko, o które toczy się odwieczny spór, że odkryliśmy matematyczne podstawy Jungowskiej teorii synchroniczności, itp. itd.

W efekcie, żadne szanujące się czasopismo matematyczne nie chciało opublikować naszej pracy i była ona systematycznie odrzucana z adnotacją, że jest ona „ bardzo interesująca matematycznie, jednak autorzy wyciągają zbyt śmiałe i zbyt daleko idące wnioski". Dopiero po kilku latach zmieniliśmy taktykę. Złagodziliśmy nasz wstęp i napisaliśmy wyraźnie jakich serii nasze wyniki dotyczą, a jakich nie. W nowej wersji praca jest od niedawna dostępna na łamach *Ergodic Theory and Dynamical Systems* [D-L].

Wprowadzimy teraz kilka podstawowych definicji, sformułowania naszych głównych twierdzeń oraz naszkicujemy metody użyte w dowodzie. Sposób w jaki to zrobimy jest wynikiem kilkuletnich przeróbek i odbiega od tego, jak dochodziliśmy do tych wyników. W czasach prowadzenia naszych pierwszych badań, byliśmy zupełnie nieświadomi interpretacji; widzieliśmy jedynie suchą asymptotykę pewnych zmiennych losowych stowarzyszonych z określoną klasą procesów stochastycznych.

Punktem wyjścia będzie pewien typ procesu stochastycznego, tzw. proces sygnałowy. Jest to proces $\mathsf{X}_t(\omega)$ (t reprezentuje czas i przebiega wszystkie liczby rzeczywiste bądź też tylko nieujemne, a ω jest elementem abstrakcyjnej przestrzeni probabilistycznej i reprezentuje realizację procesu) o wartościach naturalnych i o trajektoriach startujących z zera, niemalejących i powiedzmy prawostronnie ciągłych. Sygnały odpowiadają skokom trajektorii. Zakładamy, że nasz proces jest jednorodny, tzn. rozkłady wielowymiarowe przyrostów $\mathsf{X}_t - \mathsf{X}_s$ są niezmiennicze na przesunięcie w czasie. Do celów interpretacji, za sygnały można uważać właśnie wystąpienia jakiegoś obserwowanego przez nas zdarzenia losowego. Z jednorodnym procesem sygnałowym związana jest pewna zmienna losowa zwana czasem oczekiwania na pierwszy sygnał i określona wzorem

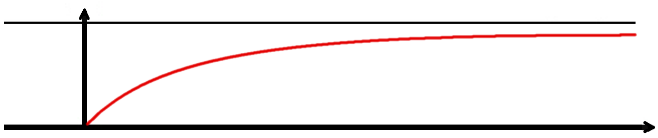
$$W(\omega) = \min\{t : \mathsf{X}_t(\omega) > 0\}.$$

Inną pomocniczą wartością takiego procesu jest jego intensywność zdefiniowana jako średnia ilość sygnałów w jednostce czasu, czyli jest to $\lambda = \mathsf{E}(\mathsf{X}_1)$. Teoretycznie wartość ta może być nieskończona, jednak procesy sygnałowe o nieskończonej intensywności, z dość oczywistych względów, nie będą nas interesować.

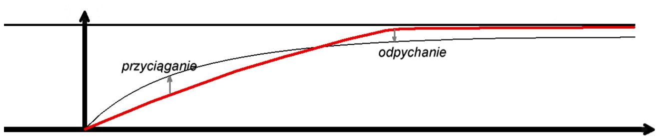
Dobrze wiadomo, że sygnały nadchodzące z określoną intensywnością λ ale w sposób niezależny od siebie (proces sygnałowy bez pamięci) opisywane są przez tzw. proces Poissona. Gdy obserwujemy powtórzenia jakiegoś losowego zjawiska w przyrodzie, co do którego mamy przekonanie, że są one od siebie wzajemnie niezależne, nie prowokują się ani nie wygaszają się nawzajem, to będziemy modelować je właśnie przy pomocy procesu Poissona. Tak powinny zachowywać się przykładowe czasy wejść klientów do sklepu. Rozkład czasu oczekiwania w procesie Poissona jest wykładniczy, w związku z tym odstępy pomiędzy sygnałami nie są stałe, tylko losowe, czasem większe czasem mniejsze. Nic więc dziwnego, że Kammerer obserwował przewagę jednominutowych przedziałów, gdzie liczba sygnałów była różna od wartości oczekiwanej. Aby uniknąć „błędu Kammerera" musimy zatem przyjąć proces Poissona za punkt odniesienia. Uznać, że w procesie Poissona nie ma żadnego prawa serii. Prawo serii zdefiniujemy jako odpowiednie odchylenie procesu sygnałowego od procesu Poissona. Oczywiście dany proces możemy porównywać tylko z procesem Poissona o takiej samej intensywności.

Okazuje się, że poszukiwana definicja jest niezwykle prosta i odwołuje się wyłącznie do rozkładu czasu oczekiwania. Nie trzeba analizować skomplikowanych rozkładów wielowymiarowych procesu. Przypomnijmy, że w procesie Poissona rozkład czasu oczekiwania jest wykładniczy z parametrem λ, gdzie λ jest intensywnością procesu i jej dystrybuanta zapisuje się wzorem

$$F_\lambda(t) = 1 - e^{-\lambda t}.$$

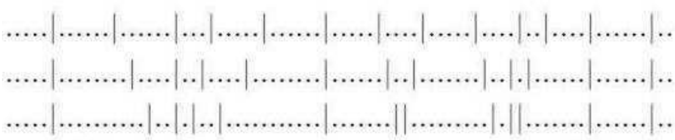


Definicja 2.1. *Powiemy, że sygnały w procesie sygnałowym o intensywności λ wzajemnie się przyciągają z odległości t, jeśli $F_W(t) < F_\lambda(t)$, a odpychają się, jeśli zachodzi nierówność przeciwna, gdzie F_W jest dystrybuantą czasu oczekiwania w danym procesie. Moduł różnicy między prawą a lewą stroną nierówności nazwiemy siłą przyciągania (lub odpowiednio odpychania).*



Napomknijmy, że z dość oczywistych powodów (o których nie będziemy tu pisać) dystrybuanta czasów powrotu jest zawsze funkcją wklęsłą, o czym trzeba pamiętać rysując jej przykładowe przebiegi.

Wyjaśnimy teraz, dlaczego te proste nierówności określamy właśnie mianem przyciągania i odpychania. Przypuśćmy, że obserwujemy nasz proces przez czas t. Wartość $F_W(t)$ to prawdopodobieństwo, że czas oczekiwania na pierwszy sygnał w naszym procesie jest mniejszy od t, czyli że zobaczymy co najmniej jeden sygnał. Jeśli $F_W(t) < F_\lambda(t)$, to mamy podwyższone (w porównaniu z procesem Poissona) prawdopodobieństwo braku sygnałów w przedziale czasu długości t. Czyli w typowej realizacji procesu będziemy mieli więcej pustych przedziałów długości t. Ponieważ intensywność jest ustalona, ta sama średnia ilość sygnałów musi zatem bardziej koncentrować się w niepustych przedziałach o długości t. Będzie to więc (obserwowalna w skali czasu t) podwyższona skłonność sygnałów do skupiania się. Odpychanie to efekt przeciwny: obserwować będziemy mniej pustych przedziałów, co oznacza bardziej wyrównane wzdłuż osi czasu rozmieszczenie sygnałów z typowej realizacji. Poniższy rysunek ukazuje przykładowe rozmieszczenie sygnałów na osi czasu w procesie z odpychaniem, Poissona i z przyciąganiem.



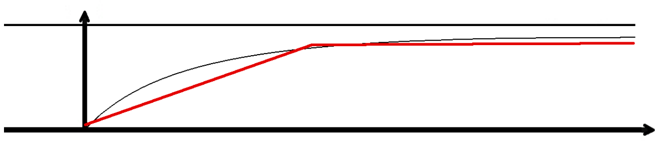
Jeśli teraz w danym procesie występuje przyciąganie z jednych odległości i odpychanie z innych, to skłonność do skupiania się sygnałów (powstawania serii) będzie niejednoznaczna, zależna od perspektywy czasu obserwacji. Ale jeśli dany proces wykazuje wyłącznie przyciąganie, to w dowolnej skali czasu zobaczymy jedynie tendencję do tworzenia się serii. To jest właśnie to, co chcemy nazwać mianem prawa serii:

Definicja 2.2. *Powiemy, że w procesie sygnałowym o intensywności λ występuje prawo serii jeśli spełnione są jednocześnie dwa postulaty:*

Postulat 1. Z żadnej odległości nie ma w nim odpychania;

Postulat 2. Z przynajmniej jednej odległości jest przyciąganie.

Innymi słowy $F_W(t) \leqslant F_\lambda(t)$, i funkcje te nie są sobie równe. W praktyce będziemy tolerować marginalne odpychanie, z siłą o wiele mniejszą od siły występującego przyciągania, jak to pokazano na rysunku.



W ten sposób sprecyzowaliśmy pojęcie, o które (bez jego definiowania) spierano się w pierwszej połowie ubiegłego stulecia. Widać teraz, że nie można rozstrzygać o występowaniu prawa serii w procesach, w których dysponujemy niewielką liczbą obserwacji, nie pozwalającą nawet na ustalenie jego intensywności. Obserwacja dwóch lub trzech powtórzeń jest całkowicie niewystarczająca. Tym bardziej, jeśli o tym, co uznamy za powtórzenie będziemy decydować dopiero w trakcie trwania obserwacji i w ten sposób tendencyjnie zmieniać przebieg procesu (a tak właśnie wygląda większość obserwacji potocznego prawa serii).

Jednakże, w pewnym sensie paradoksalnie, jak wynika z naszych twierdzeń, to właśnie zachowanie dystrybuanty czasu oczekiwania zgodne z naszą formalną definicją prawa serii dominuje dla określonych zdarzeń w procesach ergodycznych o skończonej liczbie stanów. Dlatego też uznaliśmy, że w układach dynamicznych odkryliśmy ową mistyczną siłę synchroniczności Pauli'ego i Junga. Jest nią mianowicie entropia.

Rozważamy klasyczny teorio-miarowy układ dynamiczny $(X, \mathfrak{A}, \mu, T)$, gdzie μ jest miarą probabilistyczną, a $T : X \to X$ jest transformacją mierzalną zachowującą miarę μ. Prostym sposobem otrzymania procesu sygnałowego jest ustalenie jakiegoś zdarzenia losowego (zbioru mierzalnego $B \subset X$) i przyjęcie za sygnał pojawienie się zdarzenia B w realizacji procesu. Zachowywanie miary implikuje, że taki proces sygnałowy jest jednorodny, a Twierdzenie Ergodyczne mówi nam, że jego intensywność jest równa prawdopodobieństwu zbioru B. Jest to co prawda proces sygnałowy z czasem dyskretnym a nie ciągłym, ale jeśli B jest zdarzeniem o bardzo małej mierze, to przyjmując za nową jednostkę czasu wartość oczekiwaną czasu powtórzenia (czyli odwrotność miary zbioru B – ten zabieg nazywa się *normalizacją*) czas będzie miał skoki tak drobne, że będziemy mogli z pomijalnym błędem traktować go jako ciągły. Możemy zatem zupełnie sensownie zapytywać o zjawiska przyciągania i odpychania w tym procesie, w szczególności o prawo serii.

Nas interesować będą specjalne zdarzenia zwane cylindrami. Ustalamy skończone rozbicie mierzalne $\mathcal{P}$ przestrzeni X. Rozbicie takie można interpretować jako *rozdzielczość* z jaką obserwujemy ewolucję naszego układu dynamicznego. Jeśli każdemu elementowi rozbicia przyporządkujemy jeden symbol, to trajektoria każdego punktu opisana będzie przez nieskończony ciąg symboli. Często tak opisany układ dynamiczny nazywamy procesem o skończonej liczbie stanów lub układem symbolicznym. W takim procesie cylindrem nazwiemy zdarzenie polegające na pojawieniu się określonego skończonego łańcucha znaków, czyli tzw. bloku

$$B = (b_1, b_2, \ldots, b_n).$$

Inaczej mówiąc, jest to klatka rozbicia połączonego

$$\mathcal{P}^n = \mathcal{P} \vee T^{-1}(\mathcal{P}) \vee \cdots \vee T^{-n+1}(\mathcal{P}).$$

Otóż właśnie dla takich cylindrów udało nam się udowodnić prawo serii przy jednym jedynym (i jakże naturalnym) założeniu, że nasz układ dynamiczny ma dodatnią entropię, czyli innymi słowy jest niedeterministyczny. Precyzyjne sformułowanie obejmuje dwa twierdzenia, każde z nich odpowiada jednemu z postulatów w definicji prawa serii:

Twierdzenie 2.3. *[D-L] Rozważamy ergodyczny proces o skończonej liczbie stanów i dodatniej entropii. Wtedy maksymalna siła odpychania cylindra długości n zawierającego x (jako funkcja na X) zbiega według miary do zera.*

Oznacza to, że dla niemal każdego (oprócz bloków o łącznej mierze ϵ) dostatecznie długiego bloku B wykres dystrybuanty czasu oczekiwania na B leży poniżej funkcji $1 - e^{-\lambda t} + \epsilon$. Oznacza to, że wszystkie te cylindry prawie nie wykazują odpychania (pomijając możliwe odpychanie z pewnych odległości z marginalnie małą siłą). Przypomnijmy, że brak odpychania stanowi pierwszy postulat w definicji prawa serii.

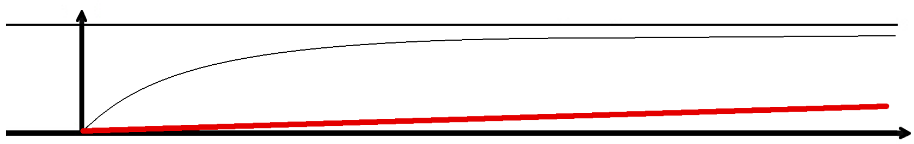
Doszedłszy do takich wniosków zaczęliśmy wraz z Yves'em Lacroix zastanawiać się (już teraz celowo) nad postulatem drugim. Oczywiście nie można oczekiwać, że wszystkie procesy o dodatniej entropii będą wykazywać przyciąganie dla jakichś cylindrów. Jak już wspomniałem, istnieje wiele procesów, w których czas oczekiwania na wszystkie długie bloki jest niemal dokładnie wykładniczy. Udało nam się jednak udowodnić, że procesy takie są wyjątkowe w (sensie kategorii Baire'a), i że w typowym procesie dla wielu cylindrów będziemy obserwować przyciąganie i to bardzo silne.

W standardowej przestrzeni probabilistycznej rodzina wszystkich rozbić mierzalnych (modulo miara) o ustalonej liczności m stanowi przestrzeń metryzowalną w sposób zupełny, zatem jest sens mówić o własnościach typowych w sensie kategorii.

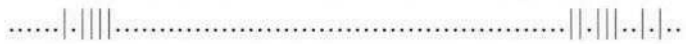
Twierdzenie 2.4. *[D-L-L] Rozważamy dowolny ergodyczny i nieokresowy teoriomiarowy układ dynamiczny $(X, \mathfrak{A}, \mu, T)$. Wtedy istnieje zbiór rezydualny (gęsty zbiór typu G_δ) rozbić o ustalonej liczebności m, które generują procesy o następującej własności:*

- *Istnieje zbiór $\mathbb{S} \subset \mathbb{N}$ o górnej gęstości 1 taki, że dla dowolnego $\epsilon > 0$ i dostatecznie dużego $n \in \mathbb{S}$, dla dowolnego cylindra B długości n, dystrybuanta F_B czasu oczekiwania na B leży poniżej prostej $y = \epsilon t$.*

Innymi słowy, w typowym procesie o m stanach, wszystkie cylindry wybranych długości n wykazują silne przyciąganie, jak to widać na poniższym rysunku dystrybuanty:



Na osi czasu daje taki oto przebieg realizacji:



Znaczenie tych twierdzeń zilustrujemy następującą, nieco zabawną anegdotką. Identycznym przykładem (jednak nie w kontekście prawa serii, tylko w celu zilustrowania co oznacza dodatnie prawdopodobieństwo) posłużył się w swym eseju z roku 1909 Émil Borel. Wyobraźmy sobie małpę siedzącą przy maszynie do pisania i bezmyślne stukającą w klawisze. Jest jasne, że jeśli by dać jej nieskończenie wiele czasu, to kiedyś napisze ona *Hamleta* (jak również wszelkie inne dzieła Shakespeare'a i wszelkich innych autorów). Mało tego, na mocy Twierdzenia Poincaré, *Hamlet* zostanie napisany nieskończenie wiele razy. Myśl ta znana jest pod hasłem *Paradoxe du singe savant* czyli „paradoks wykształconej małpy" (po angielsku przetłumaczono to, niewiadomo dlaczego, jako *Infinite Monkey Theorem*). Rzecz jasna, małpa symbolizuje tu generator losowy, a nie żywą małpę. (Sama idea jest zresztą znacznie starsza i wywodzi się jeszcze od Arystotelesa, oczywiście w jakiejś innej wersji bez maszyny do pisania). My możemy uzupełnić myśl Borela o dodatkowy szczegół: ponieważ proces generowany przez małpę nie będzie dokładnie procesem Bernoulli'ego lecz jego wersją zaburzoną, mamy prawo przypuszczać, że wpadnie on do zbioru rezydualnego, o którym mówi Twierdzenie 2.4. Istnieją też duże szanse na to, że długość *Hamleta* jest liczbą z wyróżnionego w tym twierdzeniu zbioru o górnej gęstości 1. W takim przypadku *Hamlet* będzie występować w tej pisaninie bardzo wyraźnymi seriami, wbrew intuicji, według której jego wystąpienia powinny tworzyć proces Poissona.

Inna sprawa to skala czasu, w której cokolwiek będzie można zaobserwować. Odstępy czasu między wystąpieniami *Hamleta* nawet wewnątrz takiej serii mogą wynosić po kilka milionów lat...

Widzimy więc, że nasze Ergodyczne Prawo Serii może wyjaśniać powstawanie serii nawet w procesach, które uważamy za niezależne. W rzeczywistości nie ma bowiem procesów idealnie Bernoulli'ego czy idealnie Poissona. Wszelkie procesy są zaburzone i dlatego mogą podpadać pod kategorię procesów *typowych*, czyli należących do ustalonego zbioru rezydualnego. Łatwo wyobrazić sobie zastosowania tych twierdzeń w teorii informacji, transmisji danych, genetyce i w wielu innych dziedzinach, gdzie mamy do czynienia z naprawdę długimi ciągami znaków. W oczywisty sposób nie stosują się one jednak do zjawisk takich, jak powtarzanie się jakiegoś nazwiska czy też wygrywającej kombinacji w totolotku (są to bardzo krótkie ciągi znaków). Tym bardziej nie stosują się one do zdarzeń, które w ogóle nie mają formy łańcucha znaków. Ponadto nasze twierdzenia dotyczą powtórzeń zdarzeń identycznych, a nie jedynie podobnych, jak to ma miejsce w przypadku przysłowiowych nieszczęść chodzących parami. W ostatnich latach uzyskaliśmy co prawda dodatkowe wyniki w przypadku powtórzeń zdarzeń podobnych, ale nadal podobieństwo to dotyczy cylindrów nad bardzo długimi blokami. Jest to część doktoratu pani Pauliny Grzegorek (zob. [D-G-L]), tutaj jednak nie będziemy tego tematu rozwijać.

3 Szkic dowodu

Postaramy się teraz bardzo szkicowo przedstawić ideę dowodu Twierdzenia 2.3. Ustalamy blok B dużej długości n i podzielimy go na dwie części: $B = AC$, gdzie A ma długość n_0 , a C ma długość r, przy czym $n_0 \gg r$. Obserwujemy najpierw proces wystąpień bloku A:

$$.......A....................A....A...............................A....A.................A.......A......$$

Następnie interesujemy się blokami długości r występującymi bezpośrednio po bloku A:

$$.......AX.................AX.AX............................AX.AX..............AX....AX...$$

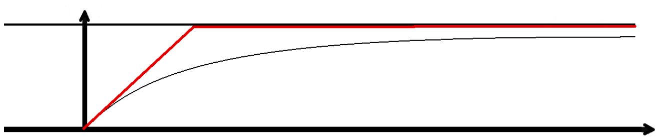
Bloki X (X, rzecz jasna, oznacza tu blok zmienny) tworzą tzw. proces *indukowany* na zbiorze A przez rozbicie $\mathcal{P}^r$. Pierwszym, najistotniejszym i najtrudniejszym krokiem dowodu jest wykazanie, że proces ten jest ϵ-niezależny (wewnętrznie) jak również ϵ-niezależny od procesu wystąpień bloku A, gdzie ϵ można uczynić dowolnie małym poprzez dobór dostatecznie dużej długości n_0 bloku A (w tej części dowodu długość r nie jest istotna i można przyjąć po prostu $r = 1$). Wrócimy do tego za chwilę.

Druga część rozumowania jest znacznie łatwiejsza. Przypuśćmy, że mamy udowodnione powyższe ϵ-niezależności. Dla uproszczenia przyjmijmy, że są to niezależności pełne. Na początek rozmieścimy bloki A w równych odstępach (zachowując intensywność):

$$............AX............AX............AX............AX............AX............AX..$$

i teraz wśród bloków X poszukajmy bloków C. Ponieważ proces ukrywający się pod X-ami jest niezależny, więc blok C wystąpi tam tak, jak w procesie Bernoullie'go, tyle tylko, że jednostką czasu będzie teraz odstęp między blokami A. Czas oczekiwania na sukces w procesie Bernoulli'ego ma rozkład geometryczny, i jeśli prawdopodobieństwo sukcesu jest bardzo małe, to możemy z niewielkim błędem przyjąć, że jest to rozkład wykładniczy. To, że prawdopodobieństwo wystąpienia C bezpośrednio po bloku A jest małe, uzyskujemy przez wcześniejszy dobór dostatecznie dużego parametru r. W tym miejscu musimy skorzystać z tego, że entropia procesu jest dodatnia; te szczegóły są dość techniczne i pominiemy je. Zatem w tej sytuacji czas oczekiwania na blok $B = AC$ ma rozkład (niemal) wykładniczy.

Jednakże w rzeczywistości bloki A rozmieszczone są inaczej wzdłuż osi czasu (przy zachowaniu intensywności). Rozmieszczenie w równych odstępach czasu modeluje maksymalne możliwe odpychanie, czyli maksymalnie wysoko przebiegającą dystrybuantę czasu oczekiwania na A (jest to funkcja na rysunku):



Każde inne rozmieszczenie może więc jedynie *obniżyć* dystrybuantę czasu oczekiwania na A. Teraz następuje niezwykle elementarny rachuneczek na dystrybuantach, którego wynik jest intuicyjnie tak oczywisty, że nie będziemy go przytaczać: obniżenie dystrybuanty dla bloku A może mieć wyłącznie jeden skutek: obniżenie dystrybuanty dla bloku $B = AC$. (W tym rachunku istotna jest niezależność procesu X-ów od procesu pojawień się bloku A). Tak więc skoro w przypadku maksymalnym dystrybuanta czasu oczekiwania na B była mniej więcej wykładnicza, to w każdym innym przypadku będzie ona leżeć poniżej dystrybuanty wykładniczej (plus ϵ). To kończy dowód Twierdzenia 2.3 modulo fakty o ϵ-niezależności.

W dowodach ϵ-niezależności podstawowym narzędziem jest entropia. Entropia skończonego rozbicia $\mathcal{P}$ jest zadana wzorem:

$$H(\mathcal{P}) = -\sum_{A\in\mathcal{P}} \mu(A)\log(\mu(A)).$$

Mając dwa rozbicia $\mathcal{P}$ i $\mathcal{Q}$ entropię warunkową zdefiniujemy wzorem

$$H(\mathcal{P}|\mathcal{Q}) = H(\mathcal{P}\vee\mathcal{Q}) - H(\mathcal{Q}).$$

Mając pod-sigma-ciało $\mathcal{F} \subset \mathfrak{A}$ entropię warunkową $H(\mathcal{P}|\mathfrak{F})$ definiujemy jako granicę $H(\mathcal{P}|\mathcal{Q}_n)$, gdzie $\mathcal{Q}_n$ jest rosnącym (w sensie rozdrabniania) ciągiem rozbić $\mathfrak{F}$-mierzalnych generujących $\mathfrak{F}$.

Dla nas istotne będą następujące fakty dotyczące entropii:

1. Dla dowolnej pary rozbić $\mathcal{P},\mathcal{Q}$ i sigma-ciała $\mathfrak{F}$ zachodzi wzór

   $$H(\mathcal{P}|\mathcal{Q}\vee\mathfrak{F}) = \sum_{A\in\mathcal{Q}} \mu(A)H_A(\mathcal{P}|\vee\mathfrak{F}),$$

   gdzie H_A oznacza entropię liczoną względem unormowanej miary warunkowej na A.

2. Rozbicie $\mathcal{P}$ jest ϵ-niezależne od sigma-ciała $\mathfrak{F}$ jeśli

   $$H(\mathcal{P}|\mathfrak{F}) \geqslant H(\mathcal{P}) - \epsilon,$$

   (a niezależne, gdy $H(\mathcal{P}|\mathfrak{F}) = H(\mathcal{P})$).

3. Rozbicia $\mathcal{P}^{[-n,-1]}$ (zadane przez odpowiednie współrzędne ujemne) generują sigma-ciało $\mathcal{P}^- = \mathcal{P}^{(-\infty,-1]}$ zwane przeszłością procesu, zatem entropie warunkowe $H(\mathcal{P}|\mathcal{P}^{[-n,-1]})$ zbiegają po n do entropii warunkowej $H(\mathcal{P}|\mathcal{P}^-)$ zwaną entropią procesu.

4. Proces generowany przez rozbicie $\mathcal{P}$ nazywamy procesem (wewnętrznie) ϵ-niezależnym jeśli rozbicie $\mathcal{P}$ jest ϵ-niezależne od przeszłości procesu $\mathcal{P}^-$, czyli gdy $H(\mathcal{P}|\mathcal{P}^-) \geqslant H(\mathcal{P}) - \epsilon$. Interpretujemy to tak, że symbol na współrzędnej zerowej mało zależy od poprzedzających symboli.

My chcemy wykazać wewnętrzną ϵ-niezależność procesu generowanego przez $\mathcal{P}$ (w tym rozumowaniu przyjmujemy $r = 1$) *indukowanego* na długim cylindrze A. Musimy więc posługiwać się marą warunkową na A. Wystarczy wykazać, że

$$H_A(\mathcal{P}|\mathcal{P}^-) \geqslant H_A(\mathcal{P}) - \epsilon.$$

(Przeszłość procesu indukowanego jest sigma-ciałem zawartym w przeszłości pełnego procesu, dlatego odpowiednia entropia warunkowa będzie pomiędzy prawą a lewą w stroną powyższej nierówności.)

Znajdujemy n_0 tak duże, by entropia warunkowa $H(\mathcal{P}|\mathcal{P}^{[-n_0,-1]})$ różniła się od granicznej $H(\mathcal{P}|\mathcal{P}^-)$ o mniej niż ϵ^2. Wtedy piszemy

$$H(\mathcal{P}|\mathcal{P}^-) = H(\mathcal{P}|\mathcal{P}^{[-n_0,-1]} \vee \mathcal{P}^-) = \sum_{A\in\mathcal{P}^{[-n_0,-1]}} \mu(A) H_A(\mathcal{P}|\mathcal{P}^-),$$

i to jest tylko o ϵ^2 mniej niż

$$H(\mathcal{P}|\mathcal{P}^{[-n_0,-1]}) = \sum_{A\in\mathcal{P}^{[-n_0,-1]}} \mu(A) H_A(\mathcal{P}).$$

Z drugiej strony, każdy wyraz pierwszej sumy jest mniejszy równy od odpowiedniego wyrazu sumy dolnej. Z tego w oczywisty sposób wynika, że nierówność

$$H_A(\mathcal{P}|\mathcal{P}^-) < H_A(\mathcal{P}) - \epsilon$$

może być spełniona na zbiorach A o łącznej mierze co najwyżej ϵ. Na pozostałych zbiorach A (czyli na większości cylindrów długości n_0) mamy

$$H_A(\mathcal{P}|\mathcal{P}^-) \geqslant H_A(\mathcal{P}) - \epsilon,$$

a to jest żądana ϵ-niezależność.

Najtrudniejsza część dowodu dotyczy ϵ-niezależności powyższego procesu indukowanego od procesu pojawień się bloku A. Jest to rozumowanie zbyt zawiłe by przytaczać tutaj jakiekolwiek szczegóły. Na tym więc pozwolimy sobie zakończyć ten szkic.

Na bazie wykładu na Forum Matematyków Polskich, Olsztyn 2010

Literatura

[A] Abadi, M., *Exponential approximation for hitting times in mixing processes*, Math. Phys. Electron. J. **7** (2001) 19 pp. (electronic)

[A-G] Abadi, M. and Galves, A., *Inequalities for the occurrence times of rare events in mixing processes. The state of the art*, Inhomogeneous random systems (Cergy-Pontoise, 2000), Markov Process. Related Fields **7** no. 1 (2001) pp 97–112

[C] Coelho, Z., *Asymptotic laws for symbolic dynamical systems*, Topics in symbolic dynamics and applications (Temuco, 1997), London Math. Soc. Lecture Note Ser. **279** (2000) pp 123–165

[D-H-M] Diaconis, P., Holmes, S. and Montgomery, R., *Dynamical bias in the coin toss*, SIAM Rev. **49**, no. 2 (2007) pp 211–235

[D-G-L] Downarowicz, T., Grzegorek, P. and Lacroix, Y., *Attracting and repelling in homogeneous signal processes*, Nonlinearity **23** (2010) pp 2793–2813

[D-L] Downarowicz, T. and Lacroix, Y., *The law of series*, Ergodic Theory Dynam. Systems **31** (2010) pp 351–357

[D-L-L] Downarowicz, T., Lacroix, Y. and Leandri, D., *Spontaneous clustering in theoretical and some empirical stochastic processes*, ESAIM. Probab. Statist., **14** (2010) pp 256–262

[J] Jung, C. G., *Jung on Synchronicity and the Paranormal: Key Readings*, Routledge, 1977

[J-P] Jung, C. G. and Pauli, W., *The Interpretation of Nature and Psyche*, Pantheon Books, New York, 1955

[K] Kammerer, P., *Das Gesetz der Serie. Eine Lehre von den Wiederholungen im Lebens- und Weltgeschehen*, Deutsche Verlags-Anstalt, Stuttgart/Berlin, 1919

[Kr] Kruskal, W., *Miracles and statistics: the casual assumption of independence*, J. Amer. Statist. Assoc. **83** no. 404 (1988) pp 929–940

[M] Moisset, J., *La loi des séries*, Ed. JMG Editions, 2000

[vM] von Mises, R., *Probability, statistics and truth*, Dover Publications Inc., New York, 1981

Instytut Matematyki i Informatyki, Politechnika Wrocławska
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław
e-mail: downar@pwr.wroc.pl