Osiem i pół

Andrzej DĄBROWSKI, Wrocław

Chodzi tu o artykuły „Osiem" oraz „Osiem i pół".

Znowu wraca problem sensownego określenia prawdopodobieństwa w zbiorze liczb. Mój wspólnik aby uniknąć kłopotliwej sytuacji i aby nie straszyć czytelnika, używa słowa-wytrychu ***szansa,*** podobnie jak w teorii liczb używa się słowa ***gęstość*** ([2]). W artykule [3] pisałem o możliwości wybrnięcia z kłopotliwej sytuacji przez wprowadzenie bardzo intuicyjnego pojęcia warunkowej przestrzeni probabilistycznej Rényi'ego [6].

W przypadku zadania o prawdopodobieństwie (przepraszam, szansie) wylosowania liczby naturalnej parzystej intuicja podpowiada, że wynosi ono (ona?) 1/2. W pytaniu, jaka jest szansa wylosowania liczby naturalnej zaczynającej się ósemką, intuicja, jak pokazał mój wspólnik, prowadzi na manowce. Co w takim wypadku należałoby zrobić?

Rozwiązanie, podane w artykule „Osiem" jest matematycznie pociągające. Każdej liczbie naturalnej przypisuje się *arbitralnie* pewną wagę $n^{-q} - (n+1)^{-q}$. Suma tych wag jest równa 1, są one nieujemne, więc mogą posłużyć jako jedna z możliwych propozycji na określenie prawdopodobieństwa na zbiorze liczb naturalnych. Propozycji jest zresztą dość dużo, bo tyle, ile można wybrać wartości $q > 0$. Autor zauważa, że dla dużych q małe liczby naturalne są uprzywilejowane. Więc, po wyliczeniu stosownych prawdopodobieństw przy ustalonym q, autor przechodzi z q do granicy równej 0. I wtedy otrzymuje się wynik, zgodny ze spostrzeżeniem Benforda (zwanym ***prawem Benforda***):

Frank Benford, fizyk zatrudniony w General Electric Company uważany jest za autora tego prawa. Jak to często w nauce bywa, Benford swoim artykułem [1] przyczynił się do spopularyzowania tego problemu. Tymczasem, 57 lat przed Benfordem, Simon Newcomb, astronom z wykształcenia, edytor *The American Journal of Mathematics*, rozpoczął pracę [5] od następującego stwierdzenia: „To, że dziesięć cyfr nie pojawia się z równą częstością musi być oczywiste dla każdego, kto korzysta z tablic logarytmicznych i widzi, że pierwsze strony są bardziej zużyte, niż ostatnie".

Prawdopodobieństwo, że liczba naturalna zaczyna się cyfrą p wynosi $\log(p+1) - \log(p)$.

Mnie takie rozwiązanie problemu nie w pełni cieszy. Martwi mnie dowolność wyboru sposobu liczenia prawdopodobieństwa. Zaniepokojony może być też nauczyciel, który musi przekonać ucznia, że wybrany przez niego sposób obliczania prawdopodobieństwa nie może być dowolny, że musi wynikać z opisu doświadczenia (modelu). Udamy się zatem na

Poszukiwanie modelu

Nie utrudnimy sobie życia, jeśli będziemy szukać zadowalającego modelu **na zbiorze wszystkich liczb rzeczywistych dodatnich**. Będzie nas bowiem interesować odpowiedź na pytanie:

Jakie jest prawdopodobieństwo, że losowo wybrana liczba rzeczywista dodatnia zaczyna się cyfrą p.

Dla liczb mniejszych od 1 chodzi o pierwszą cyfrę znaczącą (czyli różną od 0).

Wybrać losowo – to znaczy wybrać każdą liczbę rzeczywistą dodatnią z jednakowym prawdopodobieństwem.

Naszym celem jest skonstruowanie modelu losowego wyboru liczby rzeczywistej dodatniej. Wiadomo, że taki model w klasycznym rachunku prawdopodobieństwa nie istnieje. Zaproponuję model prawdopodobieństwa warunkowego.

Osobom, nie zainteresowanym szczegółami, proponuję teraz rozpocząć czytanie artykułu od definicji modelu jednorodnego.

Przypomnę definicję Rényi'ego ([6], [3]):

Układ $(\Omega, \mathscr{A}, \mathscr{B}, P(.\mid .))$, gdzie
- Ω jest niepustym zbiorem – *przestrzenią zdarzeń elementarnych*,
- $\mathscr{A}$ jest σ-algebrą Boole'a podzbiorów Ω,
- $\mathscr{B} \subset \mathscr{A}$ jest niepustą rodziną zbiorów – *rodziną warunków*,
- $P(A \mid B)$ jest dwuargumentową funkcją zbioru dla $A \in \mathscr{A}$, $B \in \mathscr{B}$,

nazywamy ***warunkową przestrzenią probabilistyczną***, gdy spełnione są aksjomaty:

1. $P(A \mid B)$ jest dla każdego ustalonego $B \in \mathscr{B}$ prawdopodobieństwem w przestrzeni $(\Omega, \mathscr{A}, \mathscr{B}, P(. \mid B))$,
2. Dla każdej trójki $A \in \mathscr{A}$, $B \in \mathscr{B}$, $B \subset C$, $P(B \mid C) > 0$ zachodzi

$$P(A \mid B) = \frac{P(A \cap B \mid C)}{P(B \mid C)}.$$

W naszym przypadku przestrzenią Ω jest zbiór liczb rzeczywistych dodatnich z rodziną zbiorów borelowskich $\mathscr{A}$ jako σ-algebrą Boole'a, ze zbiorem $\mathscr{B}$ wszystkich odcinków na półprostej dodatniej jako rodziną warunków.

Funkcja P, która będzie opisywać prawdopodobieństwo warunkowe powinna odzwierciedlać nasze założenie o losowym wyborze liczb naturalnych. Proponuję następujący model:

Model jednorodny

Dla każdego $\alpha \geq 1$ i dla każdego $u > 0$ zachodzi relacja

$$P([\alpha u, \infty) \mid [u, \infty)) = G(\alpha)$$

dla pewnej *funkcji skalującej* G.

Warunek ten jest bardzo naturalny. Wyobraźmy sobie, że stoimy w punkcie u na osi liczbowej i patrzymy w prawo (czyli w stronę coraz większych liczb dodatnich, lub jak kto woli, patrzymy w stronę plus nieskończoności). Wskazujemy losowo liczbę na prawo od tej, przy której stoimy. Chcemy oszacować prawdopodobieństwo, że wskażemy liczbę co najmniej α razy od niej większą. W modelu jednorodnym zakładamy, że prawdopodobieństwo to zależy jedynie od wzajemnej proporcji współrzędnych tych punktów. Jest to więc model ***jednorodności względem skali***.

Zadziwiające jest, że przy takim skromnym założeniu można wyznaczyć w pełni szukane prawdopodobieństwo warunkowe.

Model znaleziony!

Ustalmy na pewien czas liczbę $u_0 > 0$. Niech u będzie dowolną liczbą rzeczywistą, nie mniejszą niż u_0 (czyli spełniającą nierówność $u \geq u_0$), natomiast niech $\alpha \geq 1$. Z warunku 2. definicji warunkowej przestrzeni probabilistycznej uzyskamy ciąg równości:

$$(*) \qquad \begin{aligned} G(\alpha) &= P([\alpha u, \infty) \mid [u, \infty)) = \\ &= \frac{P([\alpha u, \infty) \cap [u, \infty) \mid [u_0, \infty))}{P([u, \infty) \mid [u_0, \infty))} = \frac{P([\alpha u, \infty) \mid [u_0, \infty))}{P([u, \infty) \mid [u_0, \infty))}. \end{aligned}$$

Wprowadźmy oznaczenie $H(u) = P([\alpha u, \infty) \mid [u_0, \infty))$. O funkcji tej, określonej na półprostej $[u_0, \infty)$ można powiedzieć, że

(1) $\qquad H(u_0) = 1 \quad \text{i} \quad H(u) \leq 1,$

(2) $\qquad H(u) \xrightarrow{u \to u_1 - 0} H(u_1),$

(3) $\qquad H(u) \xrightarrow{u \to \infty} 0,$

(4) $\qquad H(\alpha u) = G(\alpha) H(u).$

Własność (1) wynika bezpośrednio z warunku 1, definiującego prawdopodobieństwo warunkowe. Podobnie, własności (2) i (3) wynikają z warunku 2, gdyż, jak nietrudno zauważyć, funkcja $1 - H(u)$ jest dystrybuantą (a **każda** dystrybuanta jest lewostronnie ciągła). Własność (3) wynika z pierwszej i ostatniej równości ciągu $(*)$.

Funkcję $G(\alpha)$ można też wyrazić przy pomocy H:

$$G(\alpha) = G(\alpha) H(u_0) = H(\alpha u_0).$$

Ponieważ argumenty i wartości funkcji H są dodatnie, więc możemy wprowadzić dalsze oznaczenia (ln oznacza logarytm o podstawie e):

$$\alpha = \mathrm{e}^{\beta}, \quad \text{gdzie } \beta \geq 0,$$
$$u = \mathrm{e}^{v}, \quad u_0 = \mathrm{e}^{v_0},$$
$$\ln(H(\mathrm{e}^{v})) = V(v),$$

W ten sposób nasze podstawowe równanie (4) przyjmie postać:

(5) $V(\beta + v) = V(\beta + v_0) + V(v), \quad \text{gdzie } \beta \geq 0,$

(6) $V(v_0) = 0.$

Równanie funkcyjne (5) należy rozwiązać dla funkcji V lewostronnie ciągłych, spełniających warunek (6). Łatwo zauważyć, że

$$V(2\beta + v) = V(\beta + (\beta + v)) = V(\beta + v_0) + V(\beta + v) = 2V(\beta + v_0) + V(v).$$

Podstawiając do powyższego wzoru $v = v_0$ i powtarzając to postępowanie, otrzymamy wzór

$$V(k\beta + v_0) = kV(\beta + v_0), \quad \text{gdzie } k \in \mathbb{N}.$$

Z tego wzoru, po podstawieniu zamiast β wartość β/k otrzymamy

$$V(\beta + v_0) = V(k(\beta/k) + v_0) = kV(\beta/k + v_0),$$

$$V\left(\frac{\beta}{k} + v_0\right) = \frac{1}{k}V(\beta + v_0).$$

Niech teraz $w = \frac{k}{l}$ będzie dowolną, nieujemną liczbą wymierną. Z dopiero co udowodnionych wzorów wynika

$$V(w\beta + v_0) = V\left(k\left(\frac{\beta}{l}\right) + v_0\right) = kV\left(\frac{\beta}{l} + v_0\right) = \frac{k}{l}V(\beta + v_0) = wV(\beta + v_0).$$

Niech teraz $\lambda > 0$ będzie dowolną liczbą rzeczywistą, $\{w_n\}$ ciągiem liczb wymiernych, zbieżnych lewostronnie do λ. Korzystając z lewostronnej ciągłości V otrzymamy z poprzedniej tożsamości równanie

$$V(\lambda\beta + v_0) = \lambda V(\beta + v_0).$$

Z drugiej strony

$$V(v) = V(\beta + (-\beta + v)) = V(\beta + v_0) + V(-\beta + v)$$

$$V(-\beta + v) = -V(\beta + v_0) + V(v).$$

Weźmy teraz liczbą rzeczywistą $\lambda < 0$. Z poprzednich równań otrzymamy tożsamość:

$$V(\lambda\beta + v_0) = V(-|\lambda|\beta + v_0) = -V(|\lambda|\beta + v_0) = -|\lambda|V(\beta + v_0) = \lambda V(\beta + v_0)$$

Ostatecznie, V musi spełniać równanie funkcyjne:

$$\boxed{V(\lambda\beta + v_0) = \lambda V(\beta + v_0), \quad \lambda \in \mathbb{R}, \quad \text{gdzie } \beta \geq 0.}$$

Łatwo sprawdzić, że funkcja $V_0(v) = q_0(v - v_0)$ spełnia to równanie. Oznaczmy przez V_1 taką funkcję, że $V(v) = V_0(v)V_1(v)$. Funkcja V_1 jest funkcją stałą.

Mamy bowiem dla lewej strony równania definiującego V

$$V(\lambda\beta + v_0) = V_0(\lambda\beta + v_0) = \lambda V_0(\beta + v_0)V_1(\lambda\beta + v_0).$$

Podobnie, dla prawej strony tego równania:

$$\lambda V(\beta + v_0) = \lambda V_0(\beta + v_0)V_1(\beta + v_0).$$

Z porównania tych dwóch stron otrzymamy równanie

$$V_1(\lambda\beta + v_0) = V_1(\beta + v_0), \quad \lambda \in \mathbb{R}, \quad \beta \geq 0.$$

Kładąc $\beta = 1$ otrzymamy natychmiast naszą tezę, że V_1 jest stała. Ostatecznie,

$$V(v) = -q(v - v_0), \quad \text{gdzie } q > 0$$

(V jako logarytm z $H \leq 1$ ma wartości ujemne).

Przechodząc do wyjściowych oznaczeń, otrzymamy

$$\ln(H(u)) = \ln(H(e^v)) = V(v) = -q(v - v_0),$$

$$H(u) = e^{-q(v - v_0)} = \left(\frac{u_0}{u}\right)^q,$$

czyli

$$\boxed{P([u, \infty) \mid [u_0, \infty)) = \left(\frac{u_0}{u}\right)^q, \quad \text{gdzie } u \geq u_0, \quad q > 0.}$$

Funkcja po prawej stronie jest ciągła, z czego wynika, że $P(\{u\} \mid [u_0, \infty)) = 0$. Oznacza to, że w dalszych rachunkach możemy nie zwracać uwagi na końce przedziału, którego prawdopodobieństwo chcemy obliczyć.

Od razu warto obliczyć

$$P([a,b] \mid [u_0,\infty)) = \begin{cases} \left(\frac{u_0}{a}\right)^q - \left(\frac{u_0}{b}\right)^q, & \text{dla } u_0 \le u, \\ 1 - \left(\frac{u_0}{b}\right)^q, & \text{dla } u_0 \in (a,b), \\ 0, & \text{dla } u_0 \ge b. \end{cases}$$

Przy okazji otrzymaliśmy postać funkcji skalującej $G(\alpha)$

$$G(\alpha) = H(\alpha u_0) = \alpha^{-q}.$$

Zauważmy, że im q jest bliższe 0, tym mniejszy jest wpływ funkcji skalującej na obliczenie prawdopodobieństwa $P([\alpha u, \infty) \mid [u_0, \infty)) = G(\alpha) = \alpha^{-q}$. A więc, dla q bliskich 0 prawdopodobieństwo to jest *niezmiennicze względem skali*.

Używając naszego porównania z patrzeniem w prawo na oś liczbową, q jest współczynnikiem odpowiedzialnym za zasięg naszego patrzenia – im q jest bliższe 0, tym łatwiej zauważyć daleką liczbę. Na przykład, prawdopodobieństwo zauważenia liczby 100 razy większej niż najmniejsza z dostępnego nam zakresu wynosi: 0,1 dla $q = 0{,}5$, 0,40 dla $q = 0{,}2$, 0,63 dla $q = 0{,}1$ i 0,96 dla $q = 0{,}01$.

Pora więc na odkrycie kart i

Obliczenie prawdopodobieństwa, że liczba zaczyna ósemką

Oznaczmy przez D_p zdarzenie, polegające na tym, że pierwszą cyfrą znaczącą (czyli pierwszą cyfrą różną od 0) losowo wybranej liczby dodatniej jest p ($p = 1, 2, \ldots, 9$). Nietrudno zauważyć, że $D_p = \bigcup_{m=-\infty}^{\infty} [10^m p, 10^m (p+1))$.

Niech p_0 będzie pierwszą cyfrą znaczącą u_0. Wtedy $u_0 = 10^t z$, gdzie $1 \le z < 10$, $p_0 = [z]$, $t \in \mathbb{Z}$ (gdzie $\mathbb{Z}$ jest zbiorem liczb całkowitych).

Rozważymy dwa przypadki

Przypadek 1. *z jest liczbą całkowitą.*
Wtedy

$$u_0 \le 10^m p \Leftrightarrow 10^t p_0 \le 10^m p \Leftrightarrow \begin{cases} m = t, & \text{gdy } p \ge p_0, \\ m > t, & \text{gdy } p < p_0. \end{cases}$$

Natomiast nierówność

$$10^m p < u_0 < 10^m (p+1) \Leftrightarrow 10^m p < 10^t p_0 < 10^m (p+1)$$

nigdy nie zachodzi.

Stąd

$$P(D_p \mid [u_0,\infty)) = \begin{cases} \sum_{m=t}^{\infty} \left(\left(\frac{10^t p_0}{10^m p}\right)^q - \left(\frac{10^t p_0}{10^m (p+1)}\right)^q \right), & \text{gdy } p \ge p_0, \\ \sum_{m=t+1}^{\infty} \left(\left(\frac{10^t p_0}{10^m p}\right)^q - \left(\frac{10^t p_0}{10^m (p+1)}\right)^q \right), & \text{gdy } p < p_0, \end{cases}$$

co upraszcza się do wyrażenia:

$$P(D_p \mid [u_0,\infty)) = \begin{cases} p_0^q \frac{p^{-q} - (p+1)^{-q}}{1 - 10^{-q}}, & \text{gdy } p \ge p_0, \\ \left(\frac{p_0}{10}\right)^q p_0^q \frac{p^{-q} - (p+1)^{-q}}{1 - 10^{-q}}, & \text{gdy } p < p_0. \end{cases}$$

Przypadek 2. *z nie jest liczbą całkowitą.*
Wtedy

$$u_0 \le 10^m p \Leftrightarrow 10^t z \le 10^m p \Leftrightarrow \begin{cases} m = t, & \text{gdy } p \ge p_0, \\ m > t, & \text{gdy } p < p_0 \end{cases}$$

oraz

$$10^m p < u_0 < 10^m (p+1) \Leftrightarrow 10^m p < 10^t z < 10^m (p+1) \Leftrightarrow m = t \text{ i } p = p_0.$$

Analogicznie, jak w przypadku 1,

$$P(D_p \mid [u_0,\infty)) = \begin{cases} \sum_{m=t+1}^{\infty}\left(\left(\frac{u_0}{10^m p}\right)^q - \left(\frac{u_0}{10^m(p+1)}\right)^q\right), & gdy p < p_0, \\ \sum_{m=t+1}^{\infty}\left(\left(\frac{u_0}{10^m p}\right)^q - \left(\frac{u_0}{10^m(p+1)}\right)^q\right)+ & \\ \qquad +1-\left(\frac{u_0}{10^t(p+1)}\right)^q, & \text{gdy } p = p_0, \\ \sum_{m=t}^{\infty}\left(\left(\frac{u_0}{10^m p}\right)^q - \left(\frac{u_0}{10^m(p+1)}\right)^q\right), & \text{gdy } p > p_0. \end{cases}$$

Podsumowując, otrzymamy ostateczny wynik:

$$P(D_p \mid [u_0,\infty)) = \begin{cases} \left(\frac{z}{10}\right)^q \frac{p^{-q}-(p+1)^{-q}}{1-10^{-q}}, & \text{gdy } p < p_0, \\ \left(\frac{z}{10}\right)^q \frac{p^{-q}-(p+1)^{-q}}{1-10^{-q}} + 1 - \left(\frac{z}{p+1}\right)^q, & \text{gdy } p = p_0, \\ z^q \frac{p^{-q}-(p+1)^{-q}}{1-10^{-q}}, & \text{gdy } p > p_0. \end{cases}$$

Ostateczna wartość prawdopodobieństwa, że liczba rzeczywista zaczyna się cyfrą p, gdy ograniczymy się do liczb nie mniejszych od u_0 wynosi:

$$\boxed{P(D_p \mid [u_0,\infty)) = \begin{cases} \left(\frac{z}{10}\right)^q \frac{p^{-q}-(p+1)^{-q}}{1-10^{-q}}, & \text{gdy } p < p_0, \\ \left(\frac{z}{10}\right)^q \frac{p^{-q}-(p+1)^{-q}}{1-10^{-q}} & \\ \qquad +1 - \left(\frac{z}{p+1}\right)^q, & \text{gdy } p = p_0,\ z > p_0, \\ z^q \frac{p^{-q}-(p+1)^{-q}}{1-10^{-q}}, & \text{gdy } p = p_0,\ z = p_0, \\ z^q \frac{p^{-q}-(p+1)^{-q}}{1-10^{-q}}, & \text{gdy } p > p_0, \end{cases}}$$

gdzie p_0 jest pierwszą cyfrą znaczącą $u_0 = 10^t z$ i gdzie $1 \leq z < 10$, $t \in C$.

Z tego olbrzymiego wzoru na prawdopodobieństwo widać, że nie zależy ono od cechy t liczby u_0. Ciekawe jest też to, że na prawdopodobieństwo wpływa fakt, czy w u_0 jest jedna, czy więcej cyfr znaczących.

Dla modelu niezmienniczego względem skali zachodzi, niezależnie od u_0, prawo Benforda

$$\boxed{P(D_p \mid [u_0,\infty)) \xrightarrow{q\to 0+} \log\left(\frac{p+1}{p}\right).}$$

W niezwykle ciekawej pracy [5], gdzie podana jest obszerna historia prawa Benforda, oraz gdzie omówione są różne, matematyczne metody podejścia do tego problemu, przedstawiona jest tablica liczb, przedstawiająca częstość pojawiania się pierwszej cyfry znaczącej w różnych zbiorach danych.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Prawo Benforda	0,301	0,176	0,125	0,097	0,079	0,067	0,058	0,051	0,046
Tablice matematyczne	0,306	0,185	0,124	0,094	0,080	0,064	0,051	0,049	0,047
Elektryczność	0,316	0,167	0,116	0,087	0,085	0,064	0,057	0,050	0,057
Miejscowości	0,190	0,200	0,185	0,168	0,098	0,065	0,043	0,037	0,013

W powyższej tablicy w pierwszym wierszu występuje prawdopodobieństwo według prawa Benforda, w drugim – dane z tablic matematycznych o masach atomowych 91 pierwiastków, w trzecim – są dane od 1243 mieszkańców Wysp Salomona, których zużycie elektryczności w październiku 1969 roku przekraczało 1 kilowatogodzinę, wreszcie w ostatnim wierszu – dane, zawierające liczbę mieszkańców miejscowości w pięciu stanach USA w latach 1960–1970, przekraczających 2500.

W ostatnim wierszu tablicy widać wyraźne odstępstwo od prawa Benforda. Moim zdaniem wynika to z dwóch faktów:

(i) dane nie zaczynają się od potęgi 10 (jak w innych danych, przytoczonych w tablicy), ani nie są większe od liczby o jednej cyfrze znaczącej ($u_0 = 2500$),

(ii) współczynnik q jest odległy od 0.

W tablicy, którą podaję poniżej (za [5]), w wierszach są podane prawdopodobieństwa przy różnych q, dla u_0, będącego potęgą 10 z wyjątkiem ostatniego wiersza, gdzie u_0 jest równe 2500. Widać, że dane z tablic Raimiego można dobrze przybliżyć przez obliczone przez nas prawdopodobieństwa: dla danych z tablic matematycznych z $q = 0{,}02$ i dla danych o elektryczności z $q = 0{,}06$. Stanowi to pewien przyczynek do oszacowania, jak daleko są te dane od liczb spełniających prawo Benforda, gdzie $q = 0$. W ostatnim wierszu starałem się uzyskać prawdopodobieństwa, zbliżone do danych o miejscowościach w USA. Współczynnik q w tym przypadku ma wartość 0,84.

q	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0,02	0,306	0,177	0,125	0,096	0,078	0,066	0,057	0,050	0,045
0,06	0,316	0,179	0,124	0,095	0,077	0,064	0,055	0,048	0,043
2	0,758	0,140	0,049	0,023	0,012	0,007	0,005	0,003	0,002
0,5	0,428	0,190	0,113	0,077	0,057	0,044	0,036	0,030	0,025
0,2	0,351	0,184	0,122	0,090	0,070	0,057	0,048	0,042	0,036
0,05	0,313	0,178	0,124	0,095	0,077	0,065	0,056	0,049	0,043
0,84	0,161	0,201	0,215	0,135	0,093	0,068	0,052	0,041	0,034

A dla miłośników ósemki i czytelników artykułu K. Omiljanowskiego wyróżniłem w obu tablicach prawdopodobieństwo otrzymania ósemki na początku losowo wybranej liczby.

Literatura

[1] F. Benford, *The law of anomalous numbers*, Proc. Amer. Phil. Soc. 78 (1938), 551–572.

[2] W. Kuczyński, *O pewnym zadaniu maturalnym*, Matematyka 1 (1994), 10–11.

[3] A. Dąbrowski, *O losowym wyborze liczb naturalnych*, Matematyka 1 (1994), 12–14.

[4] S. Newcomb, *Note on the frequency of use of the different digits in natural numbers*, Amer. J. Math., 4 (1881), 39–40.

[5] R. A. Raimi, *The first digit problem*, Amer. Math. Monthly, 83 (1976), 521–538.

[6] A. Rényi, *Probability theory*, Budapest 1970, Akadémiai Kiadó.