Metody statystyczne

kacper.topolnicki@uj.edu.pl

12 XI. 2017

Literatura

„Numerical Recepies in C"

W. H. Press, S. A. Teukolsky

W. T. Wetterling, B. P. Flannery

Cambridge University Press

rozdział 7.2

Funkcja Gęstości Prawdopodobieństwa

$f_x(x)$ zmienna losowa $\quad a \leq x \leq b$

FGP

Prawdopodobieństwo znalezienia x w przedziale $x_0 \leq x < x_1$:

$$P_x(x_0 \leq x \leq x_1) = \int_{x_0}^{x_1} f_x(x)\, dx$$

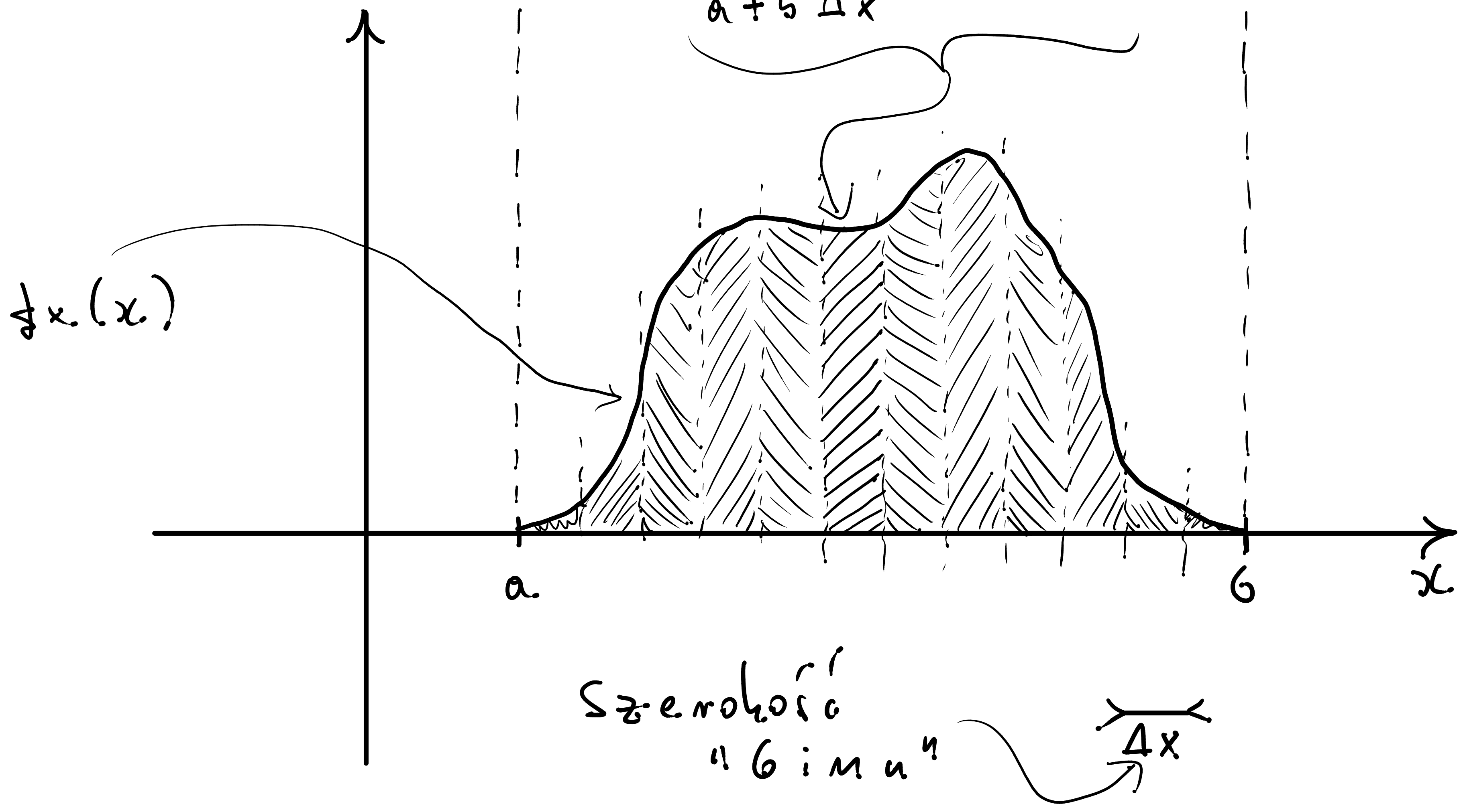
Normalizacja:

$$P_x(a \leq x \leq b) = \int_a^b f(x)\, dx = 1$$

$P_x(x = x_0)$ – nie ma sensu!

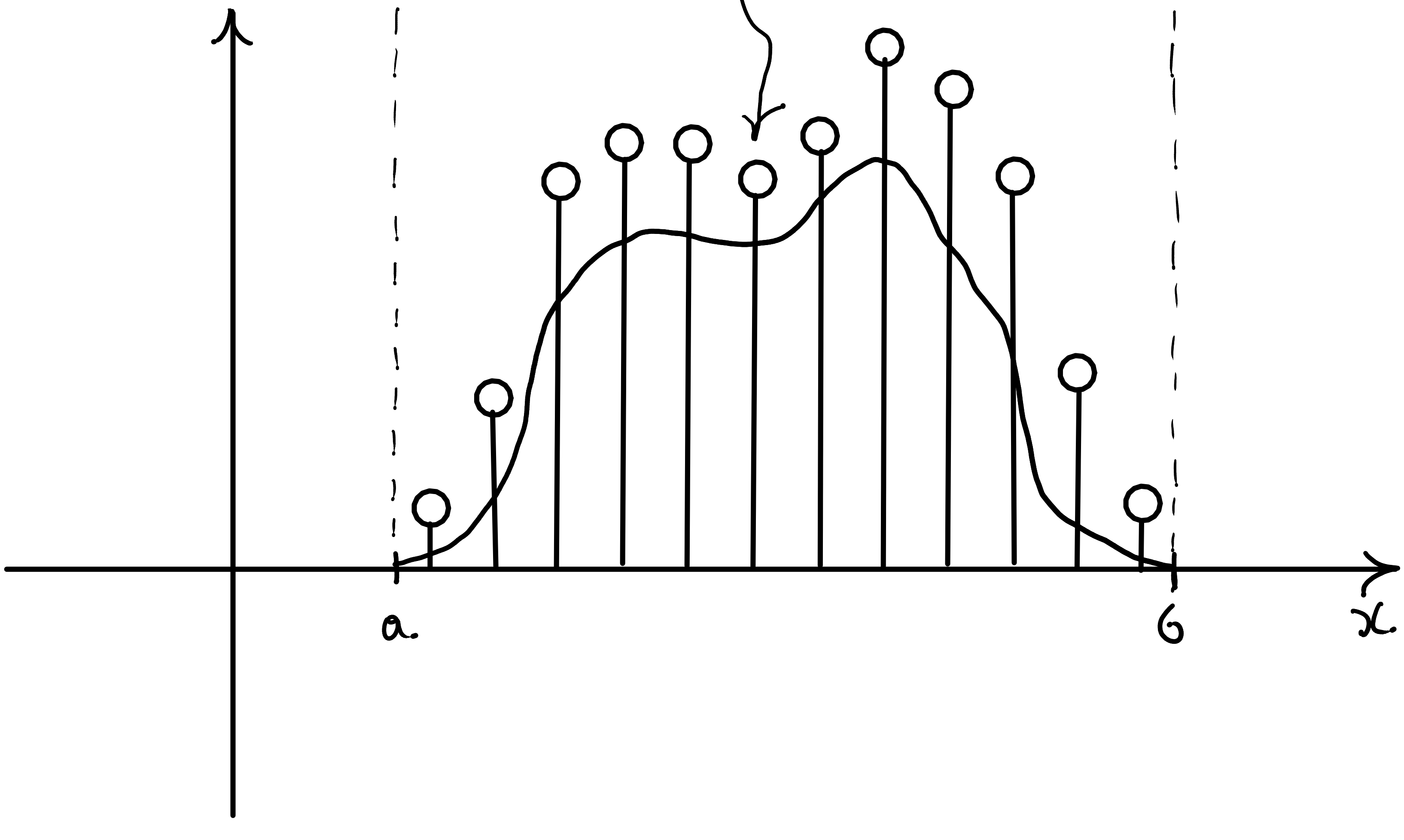
Histogram:

$$\int_{a+5\Delta x}^{a+6\Delta x} f_x(x)\,dx = \text{wysokość słupka}$$



$$\int_{a+5\Delta x}^{a+6\Delta x} f_X(x)\,dx$$

Porównując histogram z FGP należy pamiętać o Δx!



Box - Muller

x_1, x_2 - losowane z rozkładu jednorodnego na przedziale $(0, 1)$

$$y_1 = \sqrt{-2\ln(x_1)}\cos(2\pi x_2) \qquad\qquad x_1 = \exp\left(-\frac{1}{2}(y_1^2 + y_2^2)\right)$$

$$y_2 = \sqrt{-2\ln(x_1)}\sin(2\pi x_2) \quad \longrightarrow \quad x_2 = \frac{1}{2\pi}\arctan\left(\frac{y_2}{y_1}\right)$$

$$f_Y(y_1, y_2)\,dy_1\,dy_2 = f_X(x_1, x_2)\left|\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)}\right| dy_1\,dy_2$$

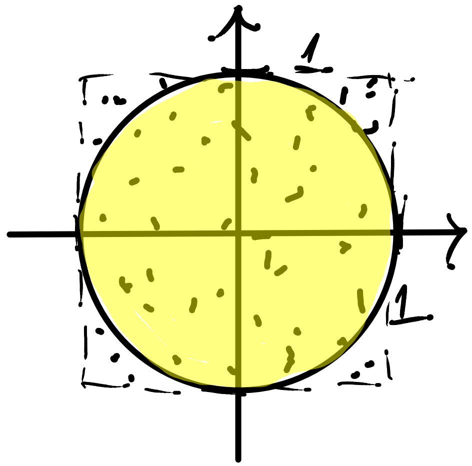
$$\left|\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)}\right| = \left|\begin{bmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{bmatrix}\right| = (-)\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y_1^2}{2}\right) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y_2^2}{2}\right)$$

$$f_Y(y_1, y_2)\, dy_1 dy_2 = \underbrace{f_X(x_1, x_2)}_{\substack{= \\ 1}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y_1^2}{2}\right) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y_2^2}{2}\right) dy_1 dy_2$$

rozkład normalny
niezależne zmienne losowe

Metoda polarna

v_1, v_2 najpierw losowane w okręgu:



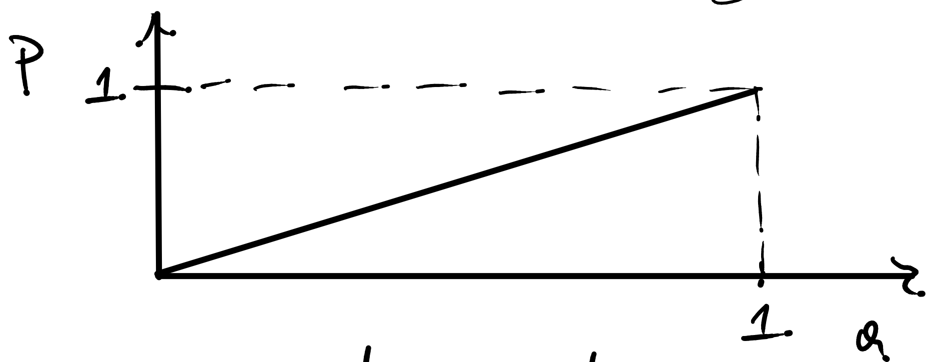
Nowe zmienne losowe:

$R^2 = v_1^2 + v_2^2$ rozkład jednorodny $(0,1)$

$\Theta = \arctan\left(\frac{v_1}{v_2}\right)$ rozkład jednorodny $(0, 2\pi)$

$T = R^2$ jakie jest prawdopodobieństwo wylosowania $0 \le T \le a$

$$P(0 \le T \le a) = \frac{\pi a}{\pi 1^2} = a$$



rozkład jednorodny

$$y_1 = \sqrt{-2\ln(x_1)}\ \cos(2\pi x_2)$$

$$y_2 = \sqrt{-2\ln(x_1)}\ \sin(2\pi x_2)$$

$$y_1 = \sqrt{-2\ln(R^2)}\ \overbrace{\cos(\ \theta\ )}^{\frac{v_1}{R}}$$

$$y_2 = \sqrt{-2\ln(R^2)}\ \underbrace{\sin(\ \theta\ )}_{\frac{V_2}{R}}$$

y_1 – rozkład normalny

y_2 – rozkład normalny

Problem ruiny gracza
(wykład R. S. Z.)

Gracz A – kapitał początkowy $a \in \mathbb{Z}$

Gracz B – kapitał początkowy $b \in \mathbb{Z}$

$$z = a + b$$

A wygrywa 1 – prawdopodobieństwo p

B wygrywa 1 – prawdopodobieństwo $q = 1 - p$

Q_i – zdarzenie ruiny A przy kapitale początkowym i

M - wygrana w pierwszej kolejce

$$P(Q_i) = P(Q_i|M)\,P(M) + P(Q_i|\overline{M})\,P(\overline{M})$$

$$P(Q_i|M) = P(Q_{i+1}) = r_{i+1}$$

$$P(Q_i|\overline{M}) = P(Q_{i-1}) = r_{i-1}$$

$$r_i = P(Q_i)$$

$$r_i = r_{i+1}\cdot p + r_{i-1}\,q$$

$$r_i\overbrace{(p+q)}^{1} = r_{i+1}\,p + r_{i-1}\,q$$

$$q(r_i - r_{i-1}) = p(r_{i+1} - r_i)$$

$$\frac{r_{i+1} - r_i}{r_i - r_{i-1}} = \frac{q}{p} = \text{const}$$

$$\frac{\Delta_{i+1}}{\Delta_i} = \frac{q}{p}$$ Δ_i – ciąg geometryczny

$$\underbrace{\Delta_1}_{r_1 - r_0} + \underbrace{\Delta_2}_{r_2 - r_1} + \underbrace{\Delta_3}_{r_3 - r_2} = r_3 - r_0$$

$$r_n - r_0 = \Delta_1 \frac{1-\left(\frac{q}{p}\right)^{n+1}}{1-\left(\frac{q}{p}\right)}$$

suma pierwszych "n" wyrazów ciągu geometrycznego

$$\frac{r_n - \overset{1}{\overset{\shortparallel}{r_0}}}{\underset{\underset{0}{\shortparallel}}{r_z} - \underset{\underset{1}{\shortparallel}}{r_0}} = \frac{\Delta_n}{\Delta_z} = \frac{1-\left(\frac{p}{q}\right)^{n}}{1-\left(\frac{p}{q}\right)^{z}} = (-1)(r_n - 1)$$

$$\boxed{r_n = \frac{\left(\frac{q}{p}\right)^{n} - \left(\frac{q}{p}\right)^{z}}{1-\left(\frac{q}{p}\right)^{z}}}$$

$p \neq q \neq \frac{1}{2}$

$$p = q = \frac{1}{2}$$

$$r_i = r_{i+1} \cdot p + r_{i-1}\, q$$

$$r_i = r_{i+1} \cdot \frac{1}{2} + r_{i-1} \cdot \frac{1}{2} = \frac{1}{2}\left(r_{i+1} - r_{i-1}\right)$$

$$\frac{1}{} = r_0, \quad r_1 = \frac{r_0 + r_2}{2} = \frac{1}{2}, \quad 0 = r_2 = z$$

$$r_n = \frac{1}{} - \frac{n}{z}$$

$$r_2 = \frac{1}{} - \frac{1}{2} = \frac{1}{2}$$

Jakie jest prawdopodobieństwo, że gra się zakończy?

r_a^A — poprzednio

r_b^B — $p \to q$, $q \to p$

$p = q = \frac{1}{2}$

$$r_a^A + r_b^B = \left(1 - \frac{a}{a+b}\right) + \left(1 - \frac{b}{a+b}\right) =$$

$$= \left(\frac{\cancel{a} + b - \cancel{a}}{a+b}\right) + \left(\frac{a + \cancel{b} - \cancel{b}}{a+b}\right) = \frac{a+b}{a+b} = 1$$

$p \neq q \neq \frac{1}{2}$

$$r_a^A + r_b^B = \frac{\left(\frac{q}{p}\right)^a - \left(\frac{q}{p}\right)^z}{1 - \left(\frac{q}{p}\right)^z} + \frac{\left(\frac{p}{q}\right)^b - \left(\frac{p}{q}\right)^z}{1 - \left(\frac{p}{q}\right)^z} =$$

$$= \ldots = 1$$

gra się zawsze zakończy

Problem: A ma kwotę a skończy grę gdy

1) zbankrutuje
2) wygra δa

$$P(\text{A wygra } \delta a) = ?$$

$$P(\underbrace{\text{A wygra } \delta a}_{\text{B bankrutuje}}) + P(\underbrace{\text{A przegra } a}_{\text{A bankrutuje}}) = \underbrace{1}_{\text{gra się skończy}}$$

$$P(\text{A wygrywa } \delta A) = 1 - P(\text{A przegrywa } a)$$

Problem: A ma ∞ kapitał, B ma b

$p = q = \frac{1}{2}$

$$P(A \text{ wygrywa}) = \lim_{a \to \infty} r_b^B = \lim_{a \to \infty} \frac{a}{a+b} = 1$$

$p \neq q \neq \frac{1}{2}$

$$P(A \text{ wygrywa}) = \lim_{a \to \infty} r_b^B = \lim_{a \to \infty} \frac{1 - \left(\frac{q}{p}\right)^a}{1 - \left(\frac{q}{p}\right)^{a+b}} =$$

$$= \begin{cases} 1 & q < p \\ \left(\frac{p}{q}\right)^b & p < q \end{cases}$$

Problemy

A

- Implementacja generatora liczb losowych z rozkładu $N(0, 1)$

wartość oczekiwana μ

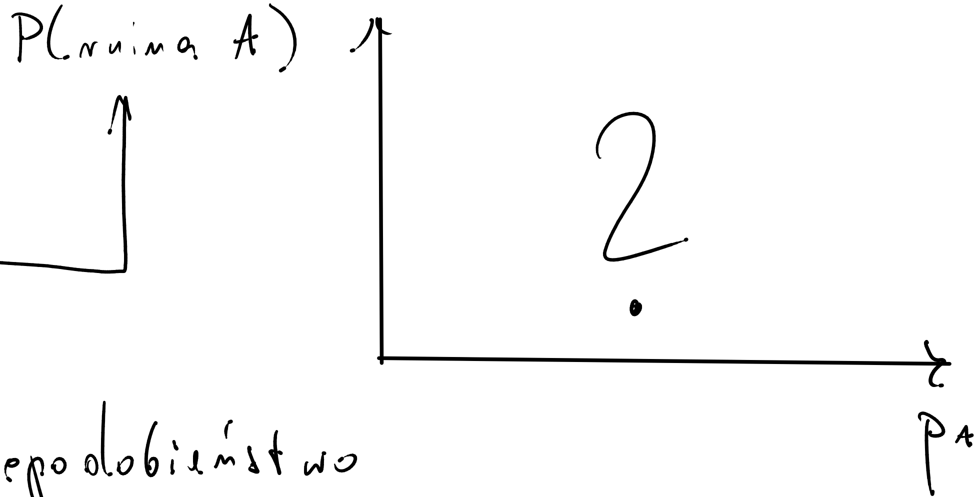
wariancja σ^2

- Narysowanie histogramu i porównanie ze wzorem analitycznym

$$f(x) = \frac{1}{\sqrt{2\sigma^2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

B

- Ruina gracza dla 2 graczy A, B



prawdopodobieństwo sukcesu w jednej rozgrywce

prawdopodobieństwo ruiny gracza A

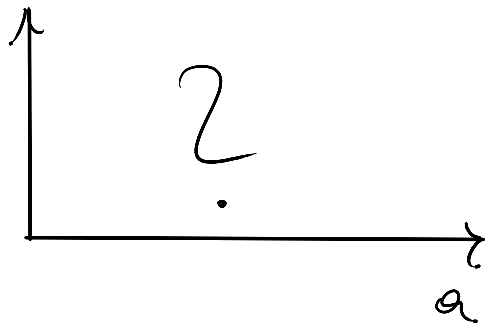
$a = 50$
$b = 50$
kapitały początkowe A, B

- Porównanie z wynikiem analitycznym dla różnych a, b

C

- Ruina gracza dla 2 graczy A, B

P(ruina A)



$a + b = 100$

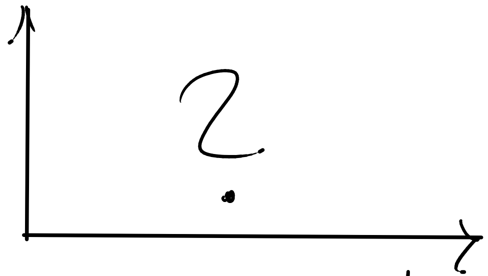
$p_A = \frac{1}{2}$

- Porównanie wyniku z teorią

D

- Liczba rozgrywek do ukończenia gry – L

prawdopodobieństwo $\rightarrow$ P(L)



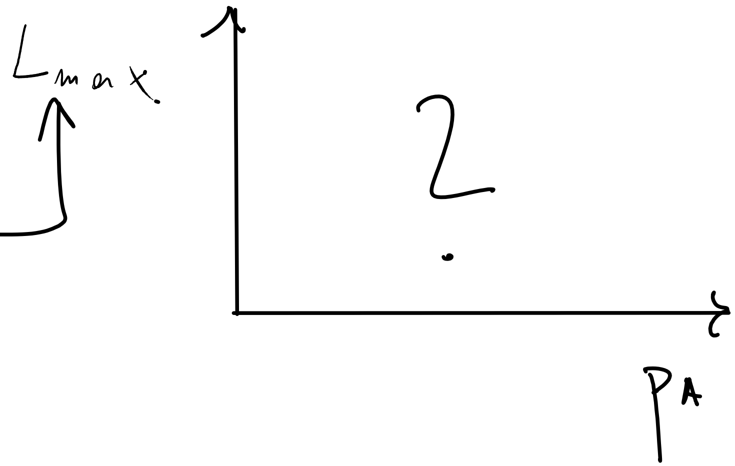
długość rozgrywki $\rightarrow$ L

$p_A = \frac{1}{5}, \frac{1}{2}, \frac{4}{5}$ $a = b = 50$

- Wyliczyć średnią długość rozgrywki

E

• Rozkład



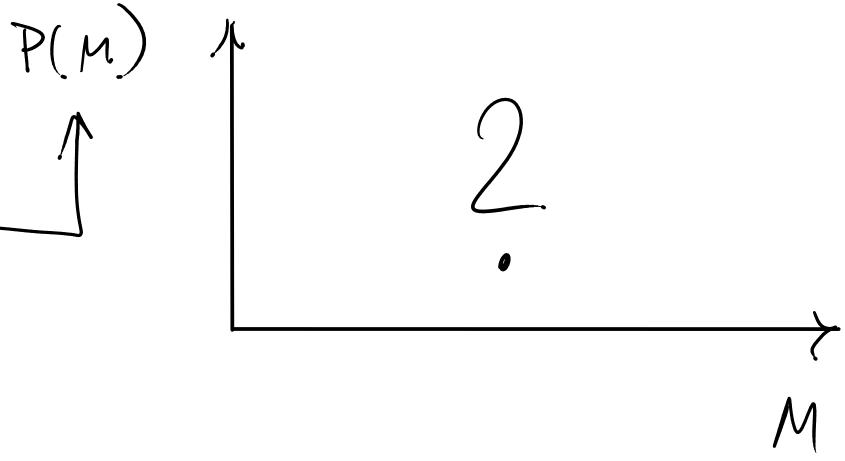
prawdopodobieństwo wygrania kolejki przez A

maksymalne długości rozgrywki przy 1000 rozgrywkach

Nie Obowiązkowe

F

- 2 graczy



kapitał gracza 1 po $n = 10$, $n = 0.5\,\bar{n}$, $n = 0.9\,\bar{n}$ rozgrywkach

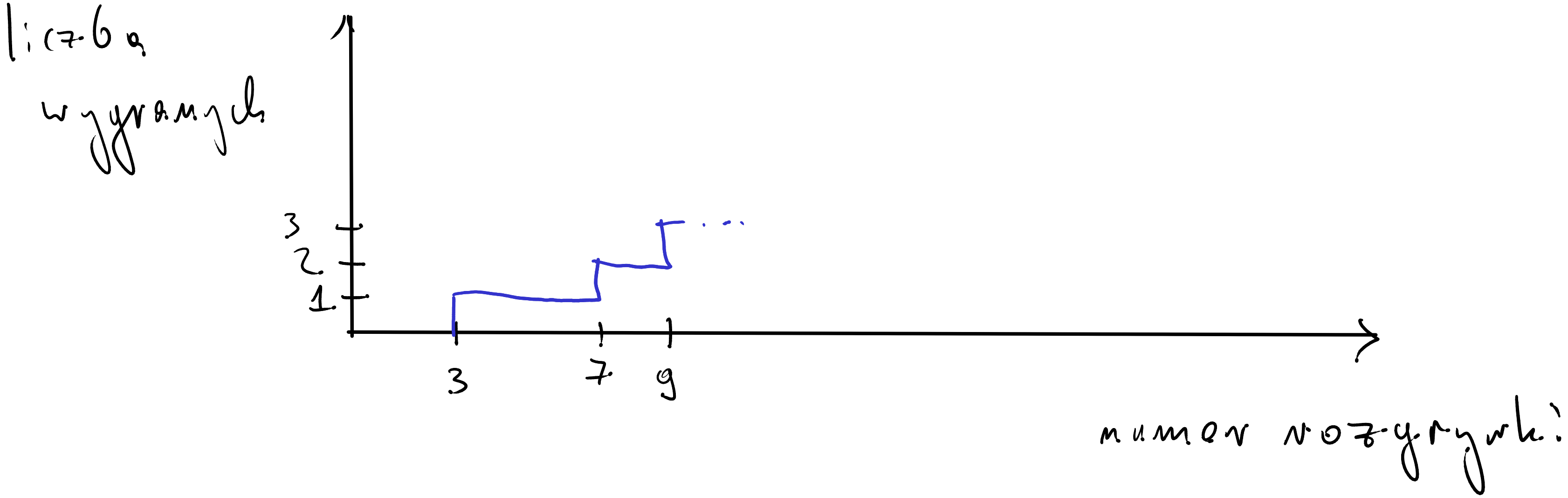
prawdopodobieństwo kapitału M

$a = b = 50$ $\qquad p_A = \frac{1}{5}, \frac{1}{2}$

nie obowiązkowe

G

- Trajektoria liczby wygranych dla 1 z 2 graczy



dla kilku gier (3, 4, ...)

H

- B, C, D, G dla kilku graczy
  (w G trajektorie naraz dla wszystkich graczy)

różne kombinacje p_i (prawdopodobieństwo wygrania gracza "i" w kolejce)

Nie Obowiązkowe