Przykład (obrona okrętów USA przed atakami lotnictwa japońskiego)

Możliwe dwie wykluczające się taktyki:

M = manewr

A = artyleria przeciwlotnicza

Departament Marynarki Wojennej na podstawie danych z wojny na Pacyfiku przeprowadził analizę statystyczną i wypracował zalecenia taktyczne (były tajne do 1960 roku).

Przyjęto, że danych jest tak dużo, że prawdopodobieństwa można zastąpić częstościami.

Okręty podzielono na dwie kategorie:

L = duże,

S = małe

Niech T oznacza trafienie okrętu.

	L		S		Razem	
TAKTYKA	Ataki	T	Ataki	T	Ataki	T
Manewr M	36	8	144	52	180	60
Art. p-lot A	61	30	124	32	185	62
Razem	97	38	268	84	365	122

Obliczono prawdopodobieństwa warunkowe:

$P(T) =$

$P(T|M) =$

$P(T|A) =$

Oraz

$P(T|L) =$

$P(T|S) =$

$P(T|L \cap M) =$

$P(T|L \cap A) =$

$P(T|S \cap M) =$

$P(T|S \cap A) =$

$P(T) = 122/365 = 0{,}334$

$P(T|M) = (60/365)/(180/365) = 0{,}333$

$P(T|A) = 62/185 = 0{,}335$

Oraz

$P(T|L) = 38/97 = 0{,}392$

$P(T|S) = 84/268 = 0{,}313$

$P(T|L\cap M) = 8/36 = 0{,}222$

$P(T|L\cap A) = 30/61 = 0{,}492$

$P(T|S\cap M) = 52/144 = 0{,}361$

$P(T|S\cap A) = 32/124 = 0{,}258$

Wniosek.
Duże okręty powinny manewrować, małe powinny kontratakować.

Literatura: B.A. Ogunnaike,"*Random Phenomena*".

Problem.

Gra polega na skreśleniu 6 liczb spośród 49. Uczestnik gry wygrywa gdy wśród wylosowanych przez organizatora 6 liczb są co najmniej 3 wytypowane przez gracza. Obliczyć prawdopodobieństwo, że gracz wygra?.

Problemy tego typu (losowanie bez zwracania gdy wśród elementów jest pewien wyróżniony podzbiór) modelujemy rozkładem **hipergeometrycznym**.

Rozkład hipergeometryczny

Dla danej liczby obiektów N z których M ma określoną własność losujemy n elementów bez zwracania. X - liczba wylosowanych obiektów o określonej własności określamy funkcję prawdopodobieństwa

$$P(X=k)=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$$

gdzie n, N, M to liczby całkowite nieujemne, $M \leq N$, $n \leq N$,
$\max(0, M + n - N) \leq k \leq \min(M, n)$

W MS EXSCEL:

ROZKŁAD.HIPERGEOM(k;n;M;N)

Rozwiązanie Problemu:

W tym przypadku n = 6, N = 49, M =6.

X – liczba odgadniętych liczb,

wtedy

$$P(X=k)=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$$

Prawdopodobieństwa dla k od 3 do 6 można przedstawić w tabelce:

k	3	4	5	6
p_k	0,017650403867	0,000968619724	0,000018449900	0,000000071511

Suma tych prawdopodobieństw wynosi około 0,018637545002 i jest to szukane prawdopodobieństwo.

Wniosek. Średnio co 54 zakład powinien wygrywać.

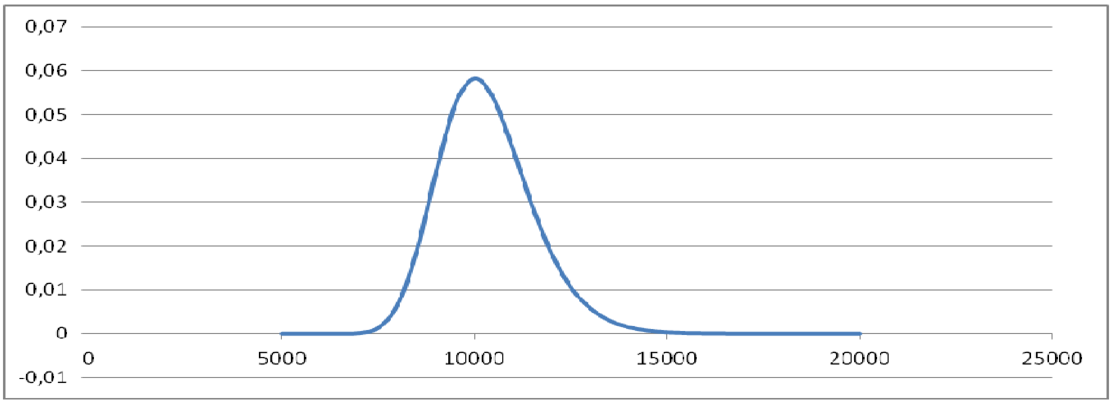
Przykład. (zastosowanie rozkładu hipergeometrycznego)

Aby oszacować liczbę N ryb w stawie wyłowiono M = 2000 ryb oznakowano je i wpuszczono do stawu. Następnie wyłowiono n = 300 ryb, okazało się, że wśród nich jest k = 60 ryb oznakowanych.

Szukamy N aby prawdopodobieństwo $\dfrac{\binom{2000}{60}\binom{N-2000}{300-60}}{\binom{N}{300}}$ **maksymalne.**

Obliczmy to prawdopodobieństwo dla różnych N:

N	p(N)
7000	0,000139477
7500	0,001357092
8000	0,006618486
8500	0,01914646
9000	0,037089475
9500	0,052598482
10000	0,058380748
10500	0,053360835
11000	0,041770769
11500	0,028875407
12000	0,01805906
12500	0,010417044
13000	0,005628622
13500	0,002884777
14000	0,001416808
14500	0,00067241
15000	0,000310511



Zatem należy sądzić, że w całym stawie jest ok. 10000 ryb.

Drugi sposób to badanie jak zmienia się

$$P(N)/P(N-1) = (N-M)(N-n)/[(N-M-n+k)N]$$

co daje N~Mn/k = 10000.

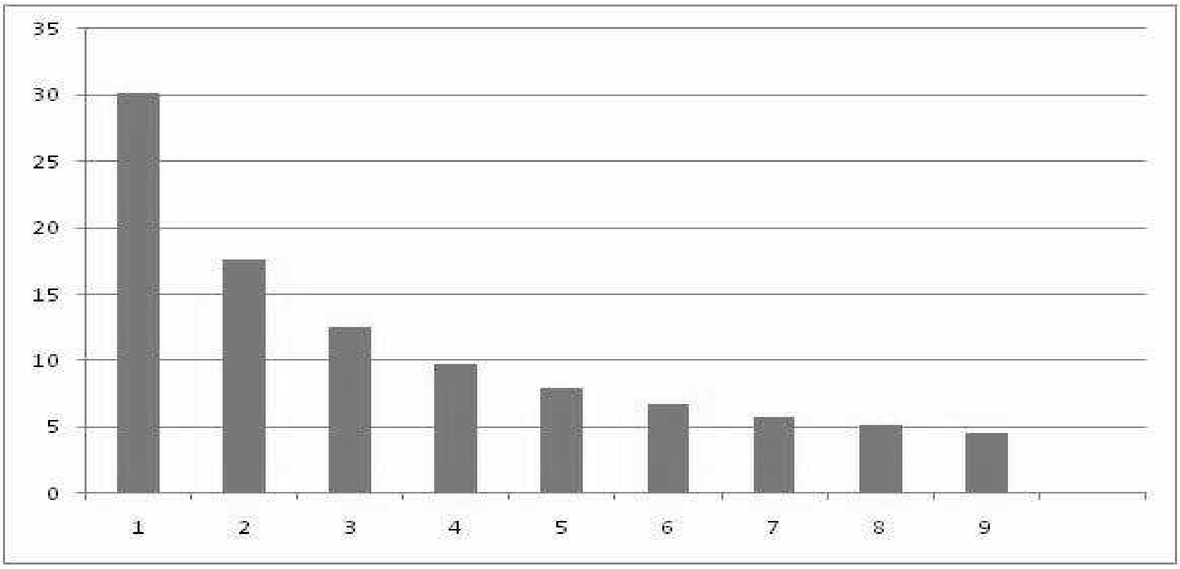
Probabilistyczny błąd mafii

Nowojorska mafia w 1920 roku organizuje **nielegalną** loterię. Jako pięciocyfrowy wynik losowania biorą pięciocyfrową końcówkę długu publicznego publikowanego przez Departament Stanu.

Benford zupełnie przypadkiem zauważył (1938), że strony tablic logarytmicznych są znacznie bardziej zabrudzone na początkowych, niż na końcowych stronach. (Wcześniej (1881) przez S. Newcomba). T. P. Hill (1995) analiza teoretyczna.

Rozkład Benforda - prawdopodobienstwo wystapienia cyfry *n* na pierwszej pozycji

$$P(n) = \lg\left(1 + \frac{1}{n}\right)$$



rozkład pierwszej cyfry .

1	2	3	4	5	6	7	8	9
30,1	17,6	12,5	9,7	7,9	6,7	5,8	5,1	4,6

Prawo Benforda = zjawisko częstego występowania powyższego rozkładu w danych statystycznych.

Przykłady:
Dane geograficzne,
Dane fizyczne,
Dane ekonomiczne.

H.Tijms, "Understanding Probability"

Przykład (wnioskowanie)

Ocenić wielkość produkcji czołgów w III rzeszy na podstawie numerów seryjnych zdobytych czołgów.

Założenie: czołgi są numerowane kolejnymi liczbami naturalnymi od 1 do N.

Zagadnienie: ocenić N.

Rozwiązanie 1. Metoda momentów.

Wartość oczekiwaną $\frac{N+1}{2}$ przyrównujemy do wartości średniej *m*. (Uzasadnić wartość oczekiwaną)

$$m = \frac{N+1}{2}$$

Stąd $N = 2m - 1$

Przykład 1.

dane: 10, 70, 90, 120

N = 2(290/4) – 1 = 144

Przykład 2.

dane: 10, 20, 100, 170

N = 2(300/4) – 1 = 149

Wynik bez sensu.

Wniosek należy ocenę uzależnić od $k = \max\{n_i\}$ i liczebności próby *n*.

Propozycja

$$N = k + \frac{k-n}{n}$$

Obliczyć N dla danych z Przykładu 1 i 2.

Rzut niesymetryczną monetą.
p -prawdopodobieństwo wyrzucenia orła,

Rzucamy 10 razy

ROOOOROOOO

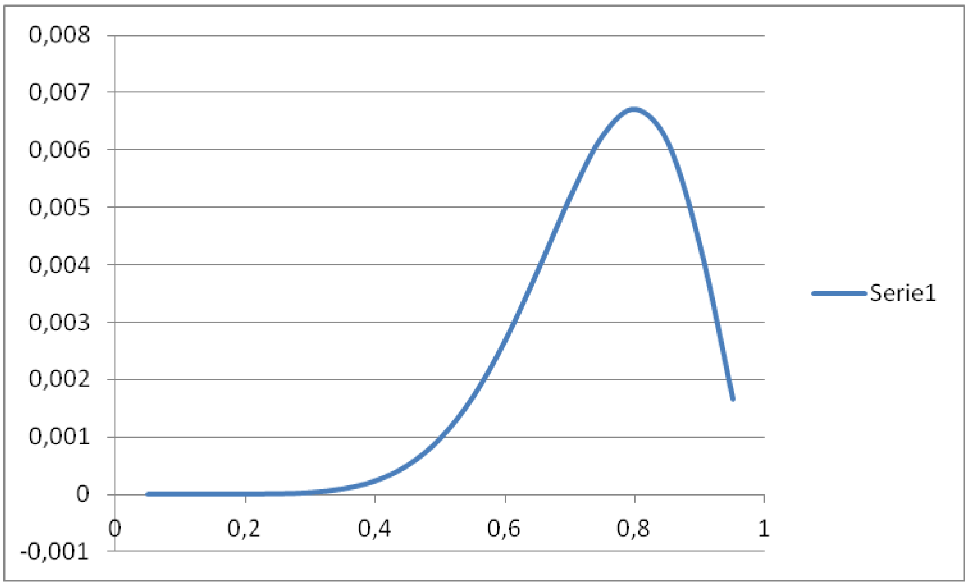
Ocenić wartość p.

Stosujemy metodę największej wiarygodności

Wyznacz p aby funkcja $L(p) = (1-p)p^4(1-p)p^4 = (1-p)^2p^8$ osiągała maksimum

p	L(p)
0,05	3,53E-11
0,1	8,1E-09
0,15	1,85E-07
0,2	1,64E-06
0,25	8,58E-06
0,3	3,21E-05
0,35	9,51E-05
0,4	0,000236
0,45	0,000509
0,5	0,000977
0,55	0,001696
0,6	0,002687
0,65	0,003903
0,7	0,005188
0,75	0,006257
0,8	0,006711
0,85	0,006131
0,9	0,004305
0,95	0,001659

Zatem p = 0,8



Przykład zastosowania estymacji

Chcemy w dyskretny sposób (obawa karalności) ocenić odsetek *k* osób dających łapówki.

Można to zrobić następująco.

Pytana osoba rzuca monetą i wynik rzutu zachowuje do swojej wiadomości.

Przygotowujemy dużą liczbę kart na połowie których jest pytanie: "czy wypadł orzeł?" a na drugiej połowie kart jest pytanie "czy dajesz łapówki?". Karty losujemy. Pytany losuje kartę i odpowiada TAK (T) lub NIE na wylosowane pytanie.

Rozpatrywane doświadczenie ma rozkład zerojedynkowy z nieznanym parametrem *p*.

Niech K1 wylosowanie karty z pytaniem nr 1.

Niech K2 wylosowanie karty z pytaniem nr 2.

Wtedy

$p = P(T) = P(K1)\, P(T|K1) + P(K2)\, P(T|K2) =$

$= 0,5 \bullet 0,5 + 0,5k$

Estymatorem dla p jest średnia w.

Stąd estymatorem k jest $k \cong 2w - 0,5$.

Paradoks zmiany decyzji w "Idź na całość"

Wpływ dodatkowej informacji na prawdopodobieństwo. Oto opis najczęściej spotykanej wersji: Teleturniej "Idź na całość"

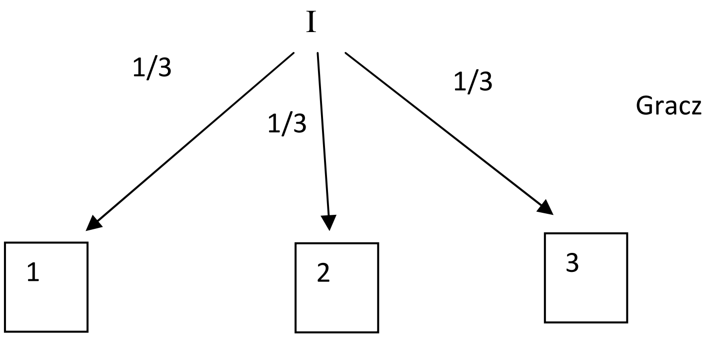
Gracz ma wybrać jedną z trzech skrzynek: w jednej jest cenna nagroda. Gracz wskazuje skrzynkę - nic nie wiedząc wybiera ją losowo. Przed otwarciem skrzynki prowadzący teleturniej (który wie, gdzie jest nagroda) otwiera jedną z pozostałych dwóch skrzynek, pustą. Prowadzący pyta gracza: "Pozostajesz przy swoim wyborze, czy zmieniasz?". Co ma zrobić gracz? Pozostać przy swoim początkowym wyborze, czy go zmienić?

Rozwiązanie:

Ponumerujmy skrzynki: 1 - pusta, 2 - pusta, 3 - pełna. Najpierw gracz (G) losuje jedną z trzech możliwości, potem losuje prowadzący (P):

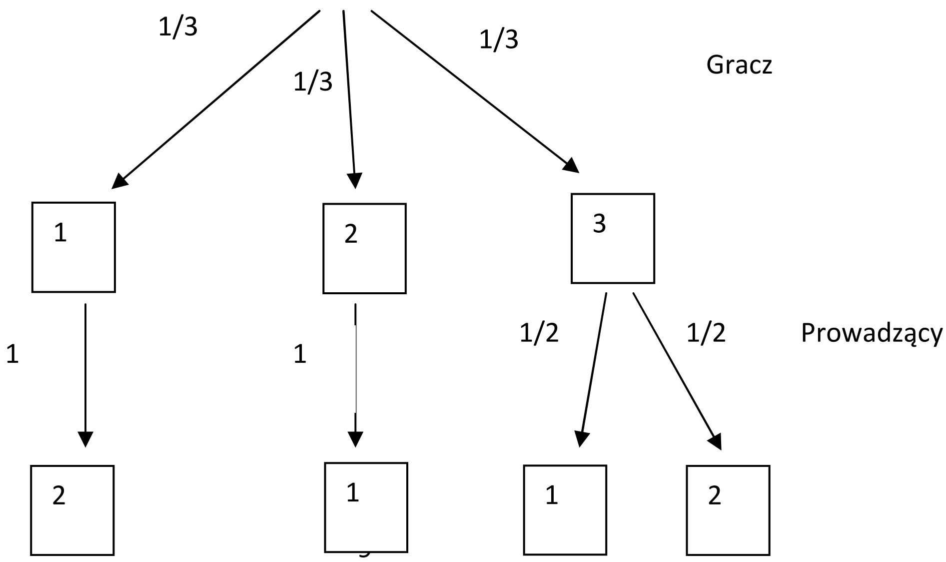
wybór gracza 1 2 3

Prawdopodobieństwo: 1/3 1/3 1/3

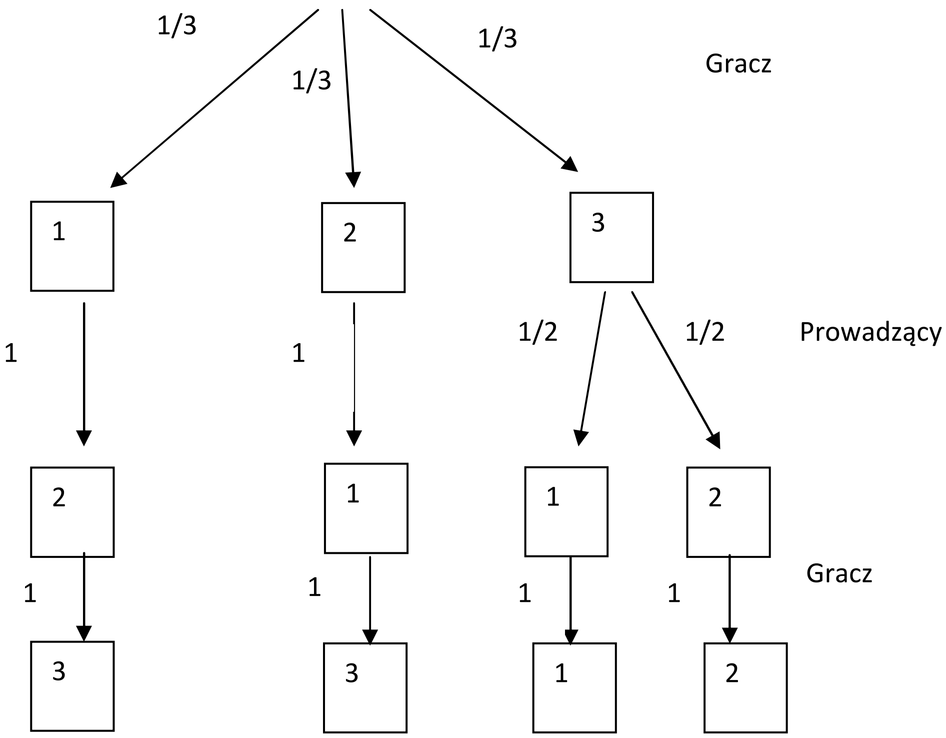


Szansa na wygraną wynosi 1/3.

Wybór prowadzącego. Niewiadome prawdopodobieństwo losowania przez prowadzącego, gdy gracz trafił, przyjmujemy równe 1/2.



Zmiana pierwotnego wyboru



Szansa na wygraną wynosi teraz 1/3 + 1/3 = 2/3.

Zatem zmiana pierwotnego wyboru podwaja szanse.

Fałszywa moneta.

Gdy chcemy coś rozstrzygnąć bezstronnie, rzucamy monetą.

Przypuśćmy, że mamy tylko niesymetryczną monetę (przesunięty środek ciężkości). Czy można taką monetą uzyskać dwa równie prawdopodobne wyniki?

Tak, należy rzucać dwukrotnie i pomijać wyniki identyczne (OO), (RR).

Wtedy wyniki (OR), (RO) są jednakowo prawdopodobne.

Jest to pomysł matematyka von Neumanna.