

Rachunek prawdopodobieństwa i statystyka

Kurs dla kierunku
Informatyki stosowanej
Uniwersytet Jagielloński
Kraków, 2024/2025

Dr hab. Roman Skibiński, prof. UJ

RPIS 2024/2025 1

1

UWAGA:

Slajdy zawierają (prawie) całość materiału przedstawianego na wykładach;

Powinny być uzupełnione wiadomościami z ćwiczeń

<http://koza.if.uj.edu.pl/~rpis/>

<http://koza.if.uj.edu.pl/~aplety/>

Konsultacje: czwartki 9-10 pokój H-0-05
(jeżeli możliwe proszę się zapowiedzieć mailowo)
lub w innym terminie w B-2-34 po umówieniu email

RPIS 2024/2025 2

2

Warunki zaliczenia (1)

- Ćwiczenia: ocena będzie średnią ocen cząstkowych (w skali 2.0-5.0) - z dwóch kolokwium z zadań i aktywności. Niezaliczone kolokwia będzie można poprawiać na kolokwium poprawkowym, poprawa dotyczy tylko niezaliczonej części materiału. W przypadku poprawiania jednego kolokwium średnia będzie liczona jako średnia z czterech ocen (2 kolokwia + aktywność + kolokwium poprawkowe) przeliczona jak wyżej. Analogicznie w przypadku poprawiania obu kolokwium średnia będzie liczona z pięciu ocen.
- Średnia ta zostanie przeliczona na ocenę z ćwiczeń, wpisywaną do indeksu, w następujący sposób:
[2.66-3.20] – ocena dostateczna,
[3.20-3.70] – ocena dostateczna plus,
[3.70-4.20] – ocena dobra,
[4.20-4.50] – ocena dobra plus,
[4.51-5.0] – ocena bardzo dobra.
- Zasady dla oceny z aktywności na ćwiczeniach ustalone są w grupach ćwiczeniowych.
- Warunkiem zaliczenia ćwiczeń jest zaliczenie (najpóźniej w wyniku poprawy) obu kolokwium co najmniej na ocenę dostateczną.

RPIS 2024/2025 3

3

Warunki zaliczenia (2)

- Wykład: cztery kartkówki na ćwiczeniach, oceniane w skali 0-5 pkt. Suma punktów prowadzi do oceny:
[12-13.5] – ocena dostateczna,
[13.5-15.0] – ocena dostateczna plus,
[15.0-16.5] – ocena dobra,
[16.5-18.0] – ocena dobra plus,
[18.0-20.0] – ocena bardzo dobra.
 - Przed kartkówkami (niezapowiedzianymi) dostępne będą zagadnienia do przygotowania
 - W trakcie semestru można poprawić jedną, najgorzej napisaną kartkówkę + te z usprawiedliwionymi nieobecnościami
 - Osoby, które nie uzyskają 12 punktów piszą kolokwium poprawkowe z całości wykładu
- Aby uzyskać zaliczenie z wykładu ocena z ćwiczeń musi być nie mniejsza niż 3.0.
- Ocena końcowa (wpisywana do indeksu, z wykładu):**
(2/3*ocena z ćwiczeń + 1/3*ocena z wykładu)*0.9
Powyższy algorytm prowadzi do oceny wpisywanej do indeksu:
do 2.70 – ocena niedostateczna,
[2.70-3.15] – ocena dostateczna,
[3.15-3.60] – ocena dostateczna plus,
[3.60-4.05] – ocena dobra,
[4.05-4.5] – ocena dobra plus.
Osoby, które rozwiążą zadania programistyczne (co najmniej wszystkie -1) będą miały ocenę podwyższoną o pół stopnia np. z dobrej plus na bardzo dobrą. Dotyczy to podwyższania od oceny 3.0. **Zatem warunkiem otrzymania oceny bardzo dobrej w indeksie jest napisanie niezbednej liczby zadanych programów.**

RPIS 2024/2025 4

4

Literatura

- W.Krysicki, J.Bartos i inni, „Rachunek prawdopodobieństwa i statystyka matematyczna w zadaniach” tomy 1 i 2, PWN 2005

Literatura dodatkowa:

- J.Jakubowski, R.Sztencel „Rachunek prawdopodobieństwa i statystyka matematyczna dla (prawie) każdego”,SCRIPT, W-wa 2006.
- S.Brandt „Analiza danych”, PWN (od 1999)
- R.Nowak „Statystyka dla fizyków” PWN 2002.
- V.Rohatgi, „Statistical inference”,J.Wiley&Sons, Inc, 1984.

RPIS 2024/2025 5

5

Definiujemy sztukę przewidywania, inaczej sztukę stochastyki, jako sztukę oceniania z największą możliwą dokładnością prawdopodobieństwa zdarzeń, tak żebyśmy w naszych osądach i działaniach zawsze mogli opierać się na tym, co okazało się najlepsze, najodpowiedniejsze, najpewniejsze, najsensowniejsze; jest to jedyny cel mądrości filozofa i roztropności męża stanu.

J.Bernoulli „Ars Conjectandi”
(„Sztuka przewidywania”) 1713

Za I.Stewart „Oswajanie nieskończoności.
Historia matematyki”

RPIS 2024/2025 6

6

Zakres wykładu

- **Rachunek prawdopodobieństwa** – jak liczyć prawdopodobieństwa zdarzeń i jak je globalnie opisywać.
- **Statystyka matematyczna** – jak wnioskować w sytuacjach, gdy mamy niepełną informację (wniosek o całość grupy na podstawie informacji zebranej na części grupy, np. sondaże przedwyborcze), jak oceniać wiarygodność takiego wnioskowania (hipotezy statystyczne)

RPiS 2024/2025 7

7

Dlaczego ?

- Ma wpływ na nasze życie (gry hazardowe, ubezpieczenia, handel, kryminalistyka, medycyna, polityka, manipulacje)
- Zastosowania w informatyce:
 - symulacje komputerowe
 - metody obliczeniowe
 - modelowanie rzeczywistości (grafika)
 - probabilistyczna (statystyczna) analiza algorytmów
 - algorytmy probabilistyczne (w tym komputer kwantowy)
 - systemy kolejkowe
 - eksploracja danych (Big Data)
 - uczenie maszynowe
 - układy z szumem (np. rozpoznawanie mowy)

RPiS 2024/2025 8

8

Rozgrzewka I

Statystyka jako test uczuciowości wyborów

- Zał: kandydat wygrywa z 70% poparciem w skali kraju. Jaki będzie rozkład głosów w poszczególnych komisjach ?

RPiS 2024/2025 9

9

Rozgrzewka II - matury

Matematyka 2020

Język niemiecki 2020

Język polski p 2021

2020

2013

RPiS 2024/2025 10

10

Rozgrzewka II - Egzamin ósmoklasisty 2024

1. Wynik w skali 0-100

2. Wynik w skali 0-1000

3. Wynik w skali 0-1000

RPiS 2024/2025 11

11

Rozgrzewka III – dobór osi

Wykres. Udział w programach tożamościowych Czy kiedykolwiek brałaś (Pan!) udział w programie tożamościowym?

RPiS 2024/2025 12

12

Rozgrzewka III – dobór osi

RPIS 2024/2025 13

13

Historia

- Starożytność, Średniowiecze – gry losowe
- XVI w G.Cardano (1501-1576), „Księga o grach losowych” – podstawy prawdopodobieństwa (gry w kości i w karty, dodatkowo rozdział o skutecznym oszukiwaniu)
- A.Gombaud (Chevalier de Méré, 1607-1684) – korespondencja pomiędzy B.Pascalem (1623-1662; 1654, 1655 – trójkąt Pascala) a P.de Fermat (1601-1665), problem podziału puli przy przerwaniu gry („problem of points”)

Rozwiązanie biorące pod uwagę tylko dotychczasowe wyniki jest błędne, należy uwzględnić możliwe zdarzenia do zaplanowanego końca gry.

RPIS 2024/2025 14

14

Historia

- Ch.Huygens(1629-1695), J.Bernoulli (1654-1705; 1713 – „Sztuka przewidywania”, białe i czarne kamyki w urnie),
- problemy typu „rzut uczciwą monetą”. Ale co to znaczy „uczciwa moneta”?
- T.Bayes (1701-1761): analiza bayesowska
- P.Laplace(1749-1827), K.Gauss(1777-1855)
- teoria miary
- A.Quetelet (1796-1874); 1835 statystyka społeczeństwa
- F.Galton (1822-1911); 1865 dziedziczenie, regresja
- K.Pearson (1857-1936); używane do dziś pojęcia i metodologia, zastosowania w naukach przyrodniczych
- XX w: A.N.Kolmogorow (1903-1987): nowoczesna tzw. aksjomatyczna teoria prawdopodobieństwa
- Wykorzystanie komputerów – nowe możliwości i nowe zadania

RPIS 2024/2025 15

15

Pojęcia wstępne

- **Eksperyment deterministyczny** – warunki wyznaczają wynik (np. tylko białe kule w urnie)
- **Eksperyment przypadkowy (zdarzenie losowe)** to taki eksperyment, którego wyniku nie potrafimy przewidzieć, mimo, że powtarzamy go w takich samych warunkach. (np. białe i czarne kule w urnie)

Jedynę co możemy zrobić to zebrać możliwe wyniki i określić ich prawdopodobieństwo.

RPIS 2024/2025 16

16

Definicja częstościowa

Powtarzamy eksperyment n razy
 $N_k(n)$ – liczba wystąpienia wyniku k w n eksperymentach
 $f_k(n)$ – względna częstość wyniku k $f_k(n) = \frac{N_k(n)}{n}$
 spełnia z def. $0 \leq f_k(n) \leq 1$
 $\sum_k f_k(n) = 1$

Częstościowa definicja prawdopodobieństwa:

$$P(k) \equiv \lim_{n \rightarrow \infty} f_k(n)$$

-trudne w praktyce (nieskończona liczba eksperymentów, powtarzalność doświadczeń, definicja eksperymentu (prawdopodobieństwo urodzenia dziewczynki/chłopca),
 -wynika z aksjomatycznej teorii prawdopodobieństwa,
 -przykład: aplet „Falszywa kostka”

RPIS 2024/2025 17

17

Aksjomatyczna definicja prawdopodobieństwa

Każdemu zdarzeniu A w przestrzeni próbek Ω przyporządkowujemy liczbę rzeczywistą $P(A)$ zwaną prawdopodobieństwem, tak by miała ona następujące własności:

I: $\forall A \subset \Omega \quad P(A) \geq 0$
 II: $P(\Omega) = 1$
 III: Jeżeli A_1, A_2, \dots jest ciągiem rozłącznych zdarzeń to $P(\bigcup_k A_k) = \sum_k P(A_k)$

RPIS 2024/2025 18

18