Rachunek prawdopodobieństwa i statystyka

Kurs dla kierunku
Informatyki stosowanej
Uniwersytet Jagielloński
Kraków, 2026/2027

Dr hab. Roman Skibiński, prof. UJ

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-0-03
(jeżeli możliwe, to proszę się zapowiedzieć mailowo)
lub w innym terminie w B-2-34 po umówieniu email:
roman.skibinski@uj.edu.pl

2

Warunki zaliczenia (1)

- Ćwiczenia: ocena będzie średnią ocen cząstkowych (w skali 2.0-5.0) - z dwóch kolokwiów 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 kolokwiów 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 ustalane są w grupach ćwiczeniowych.
- Warunkiem zaliczenia ćwiczeń jest zaliczenie (najpóźniej w wyniku poprawy) obu kolokwiów co najmniej na ocenę dostateczną.

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 niezbędnej liczby zadanych programów.**

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 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.

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"

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 całej grupie na podstawie informacji zebranej na części grupy, np. sondaże przedwyborcze), jak oceniać wiarygodność takiego wnioskowania (hipotezy statystyczne)

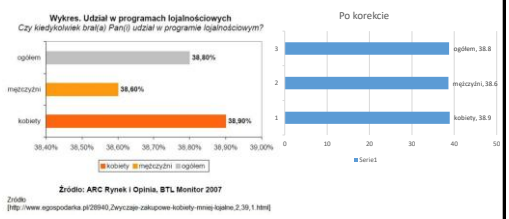
7

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)

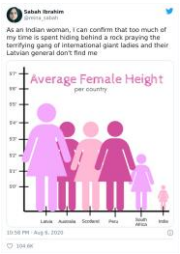
8

Rozgrzewka – dobór osi



9

Rozgrzewka – dobór osi



https://betaandbit.github.io/Wykresy/mutant.html

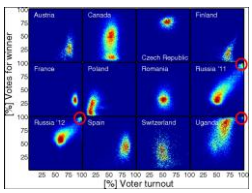


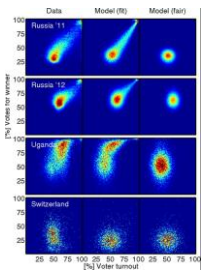
https://www.boredpanda.com/average-women-height-data-chart-latvian-indian-women/

10

Rozgrzewka – wiarygodność wyników wyborczych

- P.Klimek et al., Proceedings of the National Academy of Sciences (PNAS) 109 (41) 16469-16473 (2012)



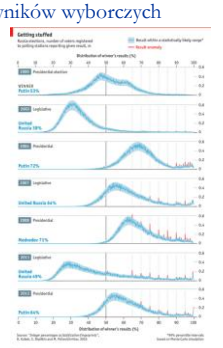


11

Rozgrzewka – wiarygodność wyników wyborczych

Statystyka jako test uczciwości wyborów

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



12

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.

13

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.Kołmogorow (1903-1987): nowoczesna tzw. aksjomatyczna teoria prawdopodobieństwa
- Wykorzystanie komputerów – nowe możliwości i nowe zadania

14

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)

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

15

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. $0 \le f_k(n) \le 1$

$$\sum_k f_k(n) = 1$$

Częstościowa definicja prawdopodobieństwa:

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

-trudne w praktyce (nieskończona liczba eksperymentów, powtarzalność doświadczeń, definicja eksperymentu (prawdopodobieństwo urodzenia dziewczynki/chłopca),

-wynika z aksjomatycznej teorii prawdopodobieństwa,

-przykład: aplet „Fałszywa kostka"

16

Pojęcia wstępne

- **Przestrzeń próbek eksperymentu przypadkowego** to zbiór Ω wszystkich możliwych wyników tego eksperymentu (dyskretna lub ciągła lub mieszana, skończona lub nieskończona)
- **Zdarzenie elementarne** – każdy możliwy wynik eksperymentu przypadkowego.
  Powtarzając eksperyment przypadkowy jako wynik otrzymujemy jedno i tylko jedno zdarzenie elementarne; zdarzenia elementarne wykluczają się wzajemnie
- **Zdarzenie** to podzbiór przestrzeni próbek.
- Szczególne zdarzenia:
  - **zdarzenie niemożliwe** – pusty podzbiór przestrzeni Ω
  - **zdarzenie pewne** – cała przestrzeń Ω

17

Działania na zbiorach (przypomnienie)

- Operacje na zbiorach
  (= podzbiorach przestrzeni próbek = zdarzeniach)

Suma $A \cup B: \quad x \in A \vee x \in B$

Iloczyn $A \cap B: \quad x \in A \wedge x \in B$

Dopełnienie $\overline{A}: \quad A + \overline{A} = \Omega$

Rozłączność $A \cap B = \phi$

Zawartość $A \subset B: \quad x \in A \Rightarrow x \in B$

Równość $A = B: \quad A \subset B \wedge B \subset A$

Własności działań na zdarzeniach:

Przemienność $A \cup B = B \cup A \qquad A \cap B = B \cap A$

Łączność $A \cup (B \cup C) = (A \cup B) \cup C \qquad A \cap (B \cap C) = (A \cap B) \cap C$

Dystrybutywność (rozłączność sumy i iloczynu)

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \qquad A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

Prawa de'Morgana

$$\overline{A \cap B} = \overline{A} \cup \overline{B} \qquad \overline{A \cup B} = \overline{A} \cap \overline{B}$$

18

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 $A_1, A_2, \ldots$ jest ciągiem rozłącznych zdarzeń to $P(\bigcup_k A_k) = \sum_k P(A_k)$

19

Podsumowanie: przestrzeń probabilistyczna

Formalnie eksperymenty losowe opisujemy zatem za pomocą trzech elementów:

I: Przestrzeń próbek Ω (zbiór zdarzeń elementarnych)

II: Przestrzeń zdarzeń losowych: zbiór podzbiorów przestrzeni Ω próbek z dołączonym zbiorem pustym (σ-ciało)

III: funkcja prawdopodobieństwa P(A)

np. obliczane poprzez (klasyczna definicja prawdopodobieństwa) ze wzoru (ćwiczenia)

20

Przypisywanie prawdopodobieństwa zdarzeniom:

Klasyczna definicja prawdopodobieństwa $P(A) \equiv \frac{\overline{\overline{A}}}{\overline{\overline{\Omega}}}$

gdzie $\overline{\overline{A}}$ to ilość równoprawdopodobnych zdarzeń elementarnych sprzyjających zdarzeniu A,

zaś $\overline{\overline{\Omega}}$ to ilość wszystkich zdarzeń elementarnych w przestrzeni Ω.

Obliczamy np. kombinatorycznie lub drzewkiem

21

Pojęcia kombinatoryczne

- Permutacja bez powtórzeń (ustawiamy n elementów) $P_n = n!$
- Wariacja bez powtórzeń (wybieramy k spośród n elementów) $V_{n,k} = \frac{n!}{(n-k)!}$
- Kombinacja bez powtórzeń (wybieramy k spośród n elementów, kolejność nieistotna) $C_{n,k} = \binom{n}{k} = \frac{n!}{k!(n-k)!}$
- Wariacja z powtórzeniami (wybieramy k spośród n elementów, dopuszczamy powtórzenia) $\overline{V}_{n,k} = n^k$
- Kombinacja z powtórzeniami (wybieramy k spośród n elementów, kolejność nieistotna, dopuszczamy powtórzenia) $\overline{C}_{n,k} = \binom{n+k-1}{k}$
- Permutacja z powtórzeniami (rozdzielamy n elementów na m grup, z których każda ma określoną wielkość) $\overline{P}_{n,n_1,n_2,\ldots,n_m} = \frac{n!}{n_1!\,n_2!\ldots n_m!}$

22

Pojęcia kombinatoryczne

A={1,2,3}, n=3, k=2

- Permutacja bez powtórzeń (ustawiamy n elementów) — (1,2,3),(2,1,3),(3,2,1),(1,3,2),(2,3,1),(3,1,2)
- Wariacja bez powtórzeń (wybieramy k spośród n elementów) — (1,2),(2,1),(1,3),(3,1),(2,3),(3,2)
- Kombinacja bez powtórzeń (wybieramy k spośród n elementów, kolejność nieistotna) — (1,2),(1,3),(3,2)
- Wariacja z powtórzeniami (wybieramy k spośród n elementów, dopuszczamy powtórzenia) — (1,2),(2,1),(1,3),(3,1),(2,3),(3,2),(1,1),(2,2),(3,3)
- Kombinacja z powtórzeniami (wybieramy k spośród n elementów, kolejność nieistotna, dopuszczamy powtórzenia) — (1,2),(1,3),(3,2),(1,1),(2,2),(3,3)
- Permutacja z powtórzeniami (rozdzielamy n elementów na m grup, z których każda ma określoną wielkość)

23

Przypisywanie prawdopodobieństwa zdarzeniom - przykład

Przykład: W urnie znajdują się dwie kule białe i trzy czarne. Wybieramy losowo dwie kule bez zwracania. Jakie jest prawdopodobieństwo tego, że obie wyciągnięte kule będą tego samego koloru?

B – zdarzenie: wyciągnięto kulę białą (w jednym losowaniu)

C – zdarzenie: wyciągnięto kulę czarną (w jednym losowaniu)

BB – zdarzenie: wyciągnięto dwie kule białe

CC - zdarzenie: wyciągnięto dwie kule czarne

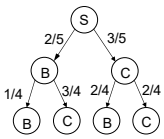
KK - zdarzenie: wyciągnięto dwie kule tego samego koloru

I sposób:

$$P(KK) = P(BB \cup CC) = P(BB) + P(CC) = \frac{\binom{2}{2}}{\binom{5}{2}} + \frac{\binom{3}{2}}{\binom{5}{2}} = \frac{\frac{2!}{2!(2-2)!} + \frac{3!}{2!(3-2)!}}{\frac{5!}{2!(5-2)!}} = \frac{1+3}{10} = 0.4$$

II sposób (metoda drzewka):

$$P(KK) = P(BB \cup CC) = P(BB) + P(CC) = \frac{2}{5}\cdot\frac{1}{4} + \frac{3}{5}\cdot\frac{2}{4} = \frac{2}{20} + \frac{6}{20} = 0.4$$



24