

*****Zadanka do poćwiczenia vol. :)*****
 ****agregacja****funkcje dwóch zmiennych****kowariancja****standaryzacja****

Zadanie 1. Poniżej mamy niekompletny opis statystyczny dotyczący liczby wypożyczonych książek w roku 2001 przez czytelników pewnej biblioteki. Opis ten został oddzielnie wykonany dla kobiet i mężczyzn, mamy również wartości niektórych parametrów dotyczące wszystkich czytelników. Uzupełnij poniższą tabelę, podając – najdokładniej jak to możliwe – wartości brakujących parametrów.

Jeśli dla któregoś z nich nie da się wyznaczyć nawet dopuszczalnego przedziału, wstaw N.

	Kobiety	Mężczyźni	Wszyscy czytelnicy
a) liczba uczniów	144	96	240
b) średnia liczba przeczytanych książek	6		5,2
c) odsetek osób które nie wypożyczyły żadnej książki	0,1	0,15	
d) odsetek osób które wypożyczyły więcej niż 10 książek	0,05	0,01	
e) mediana	7	5	
f) modalna	6		6
g) odchylenie przeciętne od mediany wzrostu	3,2	4,2	
h) błąd modalnej	0,75	0,75	

Zadanie 2. Mamy podane informacje o parametrach zmiennej X w całej zbiorowości. Podzieliliśmy naszą zbiorowość na dwie podzbiorowości A i B. Podzbiorowość A jest czterokrotnie mniejsza od podzbiorowości B. Mamy informacje o parametrach tej zmiennej w podzbiorowości A. Co możemy powiedzieć o wartościach tych parametrów w podzbiorowości B. Odpowiedzi uzasadnij

	cała zbiorowość	A	B
średnia	31,6	30	
mediana	29	27	
modalna	28	28	
błąd modalnej	0,8	0,8	
wariancja	10,44	9	
odchylenie przeciętne od mediany	5,8	5	
entropia	3,4	3,4	
minimum	15	18	

Zadanie 3. W pewnej bibliotece wydziałowej sprawdzono, ile książek w ciągu roku wypożyczyli studenci. 160 osób studiuje na tym wydziale w trybie dziennym, 40 – w trybie wieczorowym, 200- w trybie zaocznym. W poniższej tabeli znajdują się wybrane parametry zmiennej „liczba wypożyczonych książek”. Wyznacz, najdokładniej jak to możliwe, wartości parametrów dla całej zbiorowości.

Jeżeli dla któregoś z nich nie da się wyznaczyć nawet dopuszczalnego przedziału wstaw literę N.

	studenci			RAZEM
	dzienni	wieczorowi	zaoczni	
N	160	40	200	400
średnia	10	10	15	
wariancja	8	9	9	
modalna	10	{10;11}	10	
mediana	12	14	<13,15>	
odchylenie przeciętne od mediany	4	4	4	
błąd modalnej	0,7	0,7	0,9	
entropia	2,1	2,1	2,1	
minimum	0	0	2	
rozstęp	24	20	20	

Zadanie 4. Po zakończeniu wyborów na prezydenta miasta komisja wyborcza zaczęła przeliczać głosy. Spośród „pierwszych” 85% przeliczonych głosów 60% było oddanych na kandydata M . Czy po przeliczeniu pozostałych 15% może się okazać, że kandydat M uzyska mniej 50% wszystkich głosów?

Zadanie 5. Jan Pretensjonalny napisał fascynującą powieść. Po roku od wydania tej powieści zaniepokoił go umiarkowany rozgłos jego dzieła. W związku z tym do członków ze „Związku Krytyków Profesjonalnych” wysłał list z zapytaniem o to czy czytał(a) jego powieść. 80% krytyków odpisało na jego list a spośród nich 90% odpowiedziało, że przeczytało jego powieść (zakładamy, że napisali prawdę). Nic jednak nie wiadomo o pozostałych 20%, którzy nie odpisali na list. Określ najdokładniej jak potrafisz, jaki odsetek wszystkich krytyków przeczytał powieść Jana Pretensjonalnego. Odpowiedź uzasadnij!
[Nie zakładamy przy tym, że jeśli ktoś nie odpisał na list to jest mniej prawdopodobne, że przeczytał tę powieść].

Zadanie 6. Studenci uczą się statystyki w dwóch grupach. Pierwsza liczy 24 studentów, druga 16 studentów. Obydwie grupy napisały identyczny test. Średnia uzyskanych punktów z testu w grupie I była większa o 10 niż średnia w grupie II. Wariancja uzyskanych punktów z testu w grupie I wyniosła 100 a w grupie II 64. Jaka była wariancja zmiennej „liczba uzyskanych punktów” w połączonej zbiorowości obydwu grup?

Zadanie 7. W miasteczku A liczącym 1000 osób błąd modalnej zmiennej „liczba posiadanych filiżanek kawy” wynosi 0,4. Wiadomo, że w miasteczku B błąd modalnej tej zmiennej wynosi 0,7 a połączonej zbiorowości mieszkańców obydwu miast wynosi 0,52. Co można powiedzieć o liczbie mieszkańców miasteczka B?

Zadanie 8. W pewnym instytucie naukowym zrobiono badanie dotyczące tego ile razy w 2004, poszczególni pracownicy zakupili kwartalnik „Tylko statystyka” (zmienna X). Niestety, nie zdołano dotrzeć do 10% pracowników tego instytutu (w momencie badania nie można było zastać ich w pracy; z jakich powodów? Długo by o tym opowiadać...). Na podstawie informacji o przebadanych 90% pracowników określ możliwie najdokładniej parametry zmiennej X w całym instytucie.

dane o 90% pracowników	
średnia	2,2
mediana	2
modalna	2
błąd modalnej	0,4

Zadanie 9. Mamy podane wybrane informacje o parametrach dwuwartościowej zmiennej X . Nie znamy liczebności podzbiorowości A i B. Uzupełnij poniższą tabelkę tam gdzie jest to możliwe. Odpowiedzi uzasadnij

	A	B	cała zbiorowość
Minimum	0	0	
Maximum			1
Wariancja	0,21	0,24	
Średnia			
Mediana		1	
Modalna	1		
Błąd modalnej			0,36
Odchylenie przeciętne od mediany			

Zadanie 10. Mamy podany rozkład łączny dwóch zmiennych

X\Y	1	2	3
0	0,2	0	0,1
1	0,1	0,2	0,1
2	0	0,1	0,2

Wyznacz rozkład zmiennej W - „warunkowy błąd modalnej X względem Y ”, oblicz jej medianę i minimum
Wyznacz rozkład zmiennej Z - „warunkowa mediana Y względem X ”, oblicz jej maximum i błąd modalnej

Zadanie 11. Mamy podane informacje o błędzie modalnej zmiennej 0-1-kowej w podzbiorowościach A i B. Podzbiorowość A jest trzykrotnie większa od podzbiorowości B. Uzupełnij w miarę możliwości poniższą tabelkę. Odpowiedzi uzasadnij

	A	B	cała zbiorowość
błąd modalnej	0,4	0,3	
odchylenie przeciętne od mediany			

Zadanie 12. Mamy zbiorowość czterech osób

nr	X	Y
1	1	1
2	2	2
3	1	4
4	2	3

- a) oblicz kowariancję zmiennych X i Y
 b) Wyznacz zmienną $Z = E[(X|Y)]$ i oblicz medianę, wariancję i rozstęp tej zmiennej
 c) Wyznacz zmienną $W = 3X + 2Y$ i oblicz medianę tej zmiennej

Zadanie 13. Trzech chłopców Maksencjusz, Gotfryd i Euzebiusz za rozwiązanie zadania z arytmetyki dostali w nagrodę od Pani: pyszne czekoladowe ciasteczka!!! Wiadomo, że każdy z nich dostał po tyle samo. Po pierwszej przerwie lekcyjnej Maksencjuszowi i Euzebiuszowi zostało po 3 ciasteczka, Gotfrydowi nie zostało żadne. Średnia liczba ciasteczek zjedzonych w czasie przerwy, wyniosła 3 (żaden z chłopców nie podzielił się z nikim innym). Po ile ciasteczek dostał każdy z chłopców od Pani.

(Wiadomo też, że po powrocie do domu Gotfryd strasznie się rozchorował choć ta informacja w żaden sposób nie pomaga rozwiązać zadania)

Zadanie 14. Rozkład łączny częstości zmiennych X i Y przedstawia się następująco:

X\Y	0	1	2
1	0,1	0	0,05
2	0,2	0,2	0,05
3	0,1	0,2	0,1

- a) Wyznacz rozkład zmiennej $Z = 2X + Y$ i oblicz jej średnią
 b) Wyznacz rozkład zmiennej $W = (X + Y)^2$ i oblicz jej medianę
 c) Oblicz kowariancję zmiennych X i Y
 d) Wyznacz rozkład warunkowych błędów modalnych Y względem X i oblicz medianę tej zmiennej
 e) Oblicz miernik zależności stochastycznej zmiennej X względem Y

Zadanie 15. Uzupełnij niekompletne dane o zmiennych X oraz Y wiedząc, że:

- $E(X+Y) = E(X-Y) = 0$
- Wariancja zmiennej X nie zmienia się po standaryzacji zmiennej.
- $D^2(X+Y) = 5$

nr	X	Y
1	1	-4
2	-1	0
3	1	
4	-1	0
5		2
6		0

Zadanie 16..Zmienna X została zestandaryzowana. Zmienną zestandaryzowaną nazwano Z. Wiemy, że rozstęp zmiennej Z wynosi 2 a rozkład zmiennej $W = Z^2$ jest jednopunktowy. Ponadto wiemy, że $E(X) = 8 + E(Z)$ a $D^2(X) = 25$. Odtwórz rozkład zmiennej X

Zadanie 17. Opisujemy narciarzy biegowych należących do pewnego klubu za pomocą następujących zmiennych:

W-kategoria wiekowa (1-młodszy juniorzy; 2-starsi juniorzy 3-seniorzy)

X- preferowany styl (1-klasyczny; 2-łyżwa)

Y- czy brał udział w Biegu Piastów (0-nie, 1-tak)

Z – liczba godzin treningu w styczniu 2005

<i>Wśród biegaczy, którzy brali udział w biegu Piastów większość stanowią seniorzy. Które z poniższych zapisów formalnych tego zdania są poprawne (tzn. są mu logicznie równoważne):</i>	
$Mo(W Y=1)=3$	N
$Me(W Y=1)=3$	T
$P(Y=1 \wedge W=3) > P(Y=1 \wedge W<3)$	T
$P(W=3 Y=1) > P(W<3 Y=1)$	T

$H(Y W=2) = 1.$ Czy z tego zapisu wynika, że	
Połowa starszych juniorów brała udział w Biegu Piastów	T
Wszyscy starsi juniorzy brali udział w Biegu Piastów	N
Wśród zawodników, którzy brali udział w Biegu Piastów 1/3 stanowią starsi juniorzy, 1/3 stanowią młodsi juniorzy a 1/3 seniorzy	N
Połowa starszych juniorów nie brała udziału w Biegu Piastów	T

<i>Wśród młodszych juniorów odsetek biegaczy preferujących „łyżwę” jest większy niż odsetek biegaczy preferujących styl klasyczny wśród starszych juniorów. Które z poniższych zapisów formalnych tego zdania są poprawne</i>	
$P(X=2 W=1) > P(X=1 W=2)$	T
$P(X=2 \wedge W=1) > P(X=1 \wedge W=2)$	N
$P(W=1 X=2) > P(W=2 X=1)$	N
$Mo(X W=1) = 2$	N

<i>Wśród biegaczy, którzy preferują „łyżwę” i trenowali poniżej 60 godzin odsetek juniorów młodszych jest większy od odsetka juniorów starszych. Które z poniższych zapisów formalnych tego zdania są poprawne</i>	
$P(W=1 Z<60 \wedge X=2) > P(W=2 Z<60 \wedge X=2)$	T
$N(W=1 \wedge Z<60 \wedge X=2) > N(W=2 \wedge Z<60 \wedge X=2)$	T
$P(W=1 \wedge X=2 Z<60) > P(W=2 \wedge X=2 Z<60)$	T
$P(Z<60 \wedge X=2 W=1) > P(Z<60 \wedge X=2 W=2)$	N

<i>Wśród biegaczy, którzy trenowali w styczniu 2005 więcej niż 90 godzin przeważają biegacze preferujący styl klasyczny. Które z poniższych zapisów formalnych tego zdania są poprawne</i>	
$Mo(X Z>90)=1$ i jest to jedyna modalna	T
$E(X Z>90)<1,5$	T
$N(Z>90 \wedge X=1) > N(Z>90 \wedge X=2)$	T
$P(X=1 Z>90) > P(X=1 Z<90)$	N

$Me(Z Y=1) = 120.$ Czy z tego wynika, że	
Wśród biegaczy, którzy brali udział w Biegu Piastów co najmniej połowa trenowała 120 godzin lub więcej	T
Wśród biegaczy, którzy brali udział w Biegu Piastów co najmniej połowa trenowała 120 godzin lub mniej	T
Wśród biegaczy, którzy trenowali więcej niż 120 godzin co najmniej połowa brała udział w biegu Piastów	N
Wśród biegaczy, którzy brali udział w Biegu Piastów co najmniej połowa trenowała 120 godzin i co najwyżej połowa trenowała 120 godzin	N