**************************zadanka do poćwiczenia vol. 9 :) ***************************

***** estymacja i nie tylko!!!*********** estymacja i nie tylko!!!*****

We wszystkich poniższych zadaniach z wyjątkiem 4b) mamy do czynienia z doborem prostym niezależnym

1)Spośród kobiet powyżej 18 roku życia mieszkających w Polsce wylosowano próbę liczącą 1000 osób. Średnia wzrostu w tej próbie wyniosła 165 cm a wariancja 81.

a)Oszacuj przedział ufności dla średniej wzrostu godząc się na ryzyko błędnej oceny wynoszące 0,01.

b)Dane takie jak w poprzednim zadaniu. Oszacuj minimalną liczebność próby jeśli dokładność oszacowania ma być taka, by długość przedziału ufności nie przekroczyła 3 cm.

2)Spośród mieszkańców Warszawy wylosowano próbę 300-osobową. 100 osób w tej próbie pali papierosy.

a)Oszacuj przedziałowo odsetek palaczy w populacji na poziomie ufności 0,95.

b)Dla danych powyżej określ minimalną liczebność próby, taką by długość przedziału ufności nie przekroczyła 1% na poziomie ufności równym 0,95

3) Chcemy oszacować odsetek osób znających język angielki w pewnej zbiorowości. Określ minimalną liczebność próby, taką by długość przedziału ufności nie przekroczyłą 3% dla poziomu ryzyka równego 0,05

4)Chcemy oszacować średnią *ilość czasu przeznaczonego na rozmowy telefoniczne w ciągu miesiąca* w populacji. Wylosowano do badania próbę 1000 osób. Średnia *czasu rozmów* w tej próbie wyniosła 3 godziny 15 minut, odchylenie standardowe 40 minut.

a) Oszacuj przedział ufności dla średniej *tej zmiennej* godząc się na ryzyko błędnej oceny wynoszące 0,01.

b) Gdyby powyższe dane zostały uzyskane dla próby bezzwrotnej jak mogłaby się zmienić dokładność oszacowania?

b) Jeżeli chcemy by długość przedziału ufności nie przekroczyła 20 minut na takim samym poziomie ufności to ile powinna wynosić liczebność próby?

5)Spośród mieszkańców Warszawy wylosowano próbę 400-osobową. 160 osób w tej próbie posiada samochód.

a)Oszacuj przedziałowo odsetek posiadaczy samochodu na poziomie ufności 0,95.

b)Dla danych powyżej określ minimalną liczebność próby, taką by długość przedziału ufności wyniosła 1% dla poziomu ryzyka równego 0,01.

6) Posługując się próba 900 osobową badacz oszacował przedziałowo odsetek osób w zbiorowości słuchających jazzu <7,95; 12,05>. Do oszacowania wykorzystał wariancję z próby. Jaki musiał przyjąć poziom ufności (w przybliżeniu)

7) Średnia zmiennej X w zbiorowości wynosi 90 a odchylenie standardowe 5. Z populacji losujemy 200 elementowe próby w sposób zwrotny. Ile będzie równy siódmy decyl zmiennej „średnia z 200 –elementowej próby"

8) Zmienna X ma rozkład normalny o wartości oczekiwanej równej 7 i odchyleniu standardowym równym 1.

a) Ile wynosi mediana i 6 decyl zmiennej X
b) Ile wyniesie mediana i 6 decyl zmiennej Z „średnia z 1-elementowej próby dobranej <u>w sposób zależny"</u>
c) Ile wyniesie mediana i 6 decyl zmiennej W „średnia z 4-elementowej próby dobranej <u>w sposób niezależny"</u>

9) W pewnej zbiorowości wartość oczekiwana zarobków mężów (X_M) mają rozkład normalny o wartości oczekiwanej równej 2000 i odchyleniu standardowym równym 500 a zarobki żon (X_Z) mają rozkład normalny o wartości oczekiwanej równej 1800 i odchyleniu standardowym równym 300. Współczynnik korelacji liniowej pomiędzy zarobkami żon i mężów jest równy 0,3.

a) Ile wynosi wartości oczekiwana i odchyleniu standardowe łącznych zarobków męża i żony (zmienna X_S).

b) Ile wynosi wartości oczekiwana i odchylenie standardowe różnicy pomiędzy zarobkami męża i żony (zmienna X_R)

c) Uzupełnij poniżej wstawiając znak „=", „>", „<" (nie trzeba dużo liczyć)

$P(X_M > 2100)$_______$P(X_M < 1900)$
$P(X_M > 2500)$_______$P(X_Z > 1800)$
$P(X_M < 2250)$_______$P(X_Z > 1600)$
$P(X_M > 2500)$_______$P(X_Z < 1500)$
$P(X_M > 3000)$_______$P(X_Z < 1200)$
$P(X_M < 2000)$_______$P(X_Z > 1800)$

10) Wiadomo, że badacz oszacował na podstawie 144-elementowej próby średnią liczbę posianych płyt z muzyka klasyczną (zmienna X) jako przedział <19,02; 20, 98> a średnią liczbę posianych płyt z muzyka rozrywkową (zmienna Y) jako przedział <37,04; 40,96>. W obydwu przypadkach przyjął przedział ufności równy 0,95. Zakładamy, ze korelacja pomiędzy zmiennymi wynosi 0,3. Na tym samym poziomie ufności oszacuj średnią liczbę posiadanych płyt łącznie (zarówno z muzyką rozrywkową jak i klasyczną), wykorzystując – tam gdzie jest to konieczne - informacje, na podstawie podanych wyżej wyników z próby.

Dokładność estymacji przedziałowej średniej zmiennej X w populacji na podstawie n-elementowej próby w koncepcji Neymana jest tym większa:	
im liczniejsza jest próba	T
im większa jest średnia z próby	N
im większy jest poziom ufności	N
im mniejsza jest wariancja zmiennej „średnia X z *n*-elementowej próby”	T

Na podstawie próby wyznaczono przedział ufności średniej zmiennej X jako ⟨a;b⟩ na poziomie ufności równym 0,95. Czy wynika z tego, że:.	
$\mid \bar{x} - a \mid = \mid \bar{x} - b \mid$ gdzie $\bar{x}$ oznacza średnią w badanej próbie	**T**
|m-a|=|m-b| gdzie m oznacza średnią zmiennej X w badanej populacji	N
Procedura, którą stosujemy z prawdopodobieństwem 95 % pozwala na wyznaczenie przedziału ufności, którego końce obejmą średnią w populacji	T
Jeśli dla tej samej próby przy poziomie ufności równym 0,99 średnią oszacowanoby jako przedział ⟨c;d⟩ to |d-c|>|b-a|.	T

Która z poniższych statystyk z n-elementowej próby dobranej w sposób prosty – niezależny jest nieobciążonym estymatorem średniej zmiennej X w populacji	
$\bar{X}$ czyli średnia X z *n*-elementowej próby	**T**
Wartość zmiennej X dla pierwszego elementu *n*-elementowej próby	T
$2\bar{X}$	N
$\bar{X} \cdot \frac{n-1}{n}$	N

Mamy zmienną Z – „średnia zmiennej X z n-elementowej próby dobranej w sposób prosty niezależny”. Czy jest prawdą, że	
E(X)-E(Z)=0	T
$D^2(X)= D^2(Z)$	N
Zmienna ta ma rozkład normalny	N
$P(\|Z-E(X)\|>a) \leq D^2(X) / (na^2)$	T

Mamy populację 5-cio elementową (ω_1, ω_2, ω_3, ω_4, ω_5), i losujemy z niej próby dwu-elementowe w sposób prosty niezależny. Wartości zmiennej X dla poszczególnych obiektów są następujące $\omega_1(X)=1$ $\omega_2(X)=3$ $\omega_3(X)=2$ $\omega_4(X)=3$ $\omega_5(X)=3$. Czy prawdą jest, że	
Możliwych wyników losowania jest 20	N
Prawdopodobieństwo wylosowania próby {ω_1, ω_2} jest takie samo jak próby {ω_1, ω_3}	T
Jest bardziej prawdopodobne zdarzenie, że wylosujemy próbę, w której suma wartości X dla elementów wylosowanych wyniesie 4 niż próbę dla której ta suma wyniesie 3	T
Najbardziej prawdopodobne jest, że średnia w wylosowanej próbie wyniesie 12/5	N

Czy jest prawdą, że:	
$E(S^2)=D^2(X)$	N
Średnia z próby jest nieobciążonym estymatorem średniej w populacji	T
Efektywny estymator średniej z próby to taki, który ma największą wariancję	N
$E(S^2) / D^2(X)=(n-1)/n$	T

O dowolnej próbie losowej można powiedzieć, że:	
Prawdopodobieństwo wylosowania każdego elementu z populacji jest znane	T
Prawdopodobieństwo wylosowania każdego elementu jest takie samo	n
Element populacji ω_1, może zostać wylosowany do danej próby tylko raz	n
Średnia zmiennej z próby nie rożni się znacząco od średniej z populacji	n

Szacujemy przedziałowo odsetek czytelników pewnego czasopisma. Na poziomie ufności równym 0,97 odsetek ten oszacowano jako <18%, 24%>. Czy oznacza to, że	
Prawdopodobieństwo, że odsetek czytelników tej gazety nie należy do przedziału <18%, 24%> jest równe 0,03	**n**
Nasza metoda daje poprawne oszacowanie z prawdopodobieństwem równym 0,97	t
Odsetek czytelników w badanej próbie wyniósł 21%	t
Gdybyśmy przyjęli poziom ufności równy 0,95 wówczas przedział ufności wyniósłby <21% - 5,0 %; 21% + 5,0 %>	n
Gdybyśmy przyjęli poziom ufności równy 0,95 wówczas przedział ufności wyniósłby <21% - 1,96% ; 21% + 1,96%>	t