

Zadania na estymację i weryfikację hipotez

1. Estymacja, minimalna liczebność próby.

Zad. 1. W próbie 600 telewizorów typu A wykryto 48 usterek. Oszacować metodą punktową i przedziałową procent wadliwych telewizorów ($1-\alpha=0,95$).

Zad. 2. Waga tabliczek czekolady jest zmienną losową o rozkładzie $N(100g; D(x))$. Na podstawie próby losowej 120 tabliczek czekolady otrzymano $\bar{x} = 9g$ oraz $S(x)=10g$. Przy współczynniku ufności 0,90 oszacować przedziałowo wariancję wagi wszystkich produkowanych tabliczek czekolady.

Zad. 3. W badaniu wysokości wynagrodzeń w 2011 r. w przemyśle odzieżowym wylosowano 500 pracowników. Z próby tej otrzymano dane: średnie miesięczne wynagrodzenie wyniosło 2778 tys.zł. a odchylenie standardowe 1155 tys.zł. Przyjmując współczynnik ufności 0.90 metodą przedziałową oszacować przeciętne zróżnicowanie (dyspersję) wynagrodzeń w przemyśle odzieżowym.

Zad. 4. Dane są dwa wektory zmiennej losowej:

- wektor dolnych granic 4 przedziałów klasowych [0; 10; 20; 30]
- wektor wartości dystrybuanty empirycznej dla górnych granic przedziałów [0.1; 0.5; 0.9; 1.0] przy czym próba liczyła 26 jednostek. Dystrybuanta w punkcie 40 wynosi 1.0. Czy liczebność próby jest wystarczająca jeżeli założyć, że maksymalny błąd szacunku wartości oczekiwanej tej zmiennej przy współczynniku ufności 0.99 powinien wynieść 0.4?

Zad. 5. Wytrzymałość pewnego materiału budowlanego jest zmienną losową o rozkładzie normalnym $N(m,\sigma)$. W celu oszacowania nieznanej średniej m wytrzymałości tego materiału dokonano pomiaru wytrzymałości na $n=5$ wylosowanych niezależnie sztukach tego materiału. Wyniki pomiarów były następujące: 20,4; 19,6; 22,1; 20,8; 21,1. Przyjmując współczynniki ufności 0,90 i 0,99 zbuduj przedziały ufności dla średniej wytrzymałości tego materiału i porównaj długość przedziałów.

Zad. 6. Oblicz z jakim prawdopodobieństwem określono przedziały ufności dla średnich miesięcznych wydatków na artykuły spożywcze w przeliczeniu na 1 osobę w gospodarstwie domowym, jeśli otrzymano przedział o rozpiętości 90 tys. zł. a wśród losowo dobranej próby 65 gospodarstw domowych przeciętnie zróżnicowanie tych wydatków równało się 350 tys. zł. Przeciętna liczba osób w gospodarstwie domowym wynosiła 4.

Zad. 7. Obserwowano 5 świetlówek ze względu na długość czasu świecenia (tys. godz.). Otrzymano dane: 2,4; 2,5; 2,0; 2,1; 2,0. Przyjmując 95% współczynnik ufności oszacować przedziałowo średni czas świecenia świetlówek. Podać wzory, wartości i interpretację przeciętnego i maksymalnego błędu szacunku średniego czasu świecenia świetlówek.

Zad. 8. W celu oszacowania rozrzutu wagi opakowań herbaty wylosowano niezależnie $n=10$ opakowań i uzyskano : średnia $=9,9$ dkg. oraz $S^2(x)=0,25$ (dkg.)². Przyjmując współczynnik ufności 0,90 oszacuj metodą przedziałową poziom odpowiedniego parametru zróżnicowania wagi opakowań herbaty.

Zad. 9. Dla 26-elementowej próby losowej pochodzącej z populacji o rozkładzie normalnym stwierdzono, że $\bar{x}=60$ a $S(x)=40$. Jaki współczynnik ufności przyjęto przy estymacji przedziałowej m jeżeli długość przedziału ufności wyniosła 27,36.

Zad. 10. a) Cecha X ma w populacji generalnej rozkład $N(m;\sigma)$. Jaki błąd popełniamy stosując do estymacji punktowej wartości średniej w populacji generalnej średnią arytmetyczną z 200-elementowej próby losowej.

b) Podaj w jaki sposób można zwiększyć dokładność szacunku w przypadku estymacji przedziałowej

c) Do estymacji parametru θ wykorzystać można estymator Z_1 lub Z_2 . Wiadomo, że $D^2(Z_1)=D^2(Z_2)$ oraz, że $E(Z_1)=\theta$ a $E(Z_2)=\theta + \frac{1}{n}$. Który z estymatorów należy wybrać? Odpowiedź uzasadnić.

Zad. 11. W celu zbadania struktury wieku lekarzy wiejskich i miejskich pobrano losowo dwie próby: 9 elementową lekarzy miejskich i 8 elementową lekarzy wiejskich uzyskując między innymi następujące wyniki (w latach):

lekarze:	miejscy	wiejscy
przeciętny wiek	46	42

przeciętne zróżnicowanie wśród wszystkich losowo wybranych lekarzy wyniosło 9,7 lat. Zbuduj przedział ufności dla przeciętnego wieku całej populacji lekarzy (miejskich i wiejskich) przyjmując współczynnik ufności 0,97.

Zad.12.a./Oceń (punktowo i przedziałowo) przeciętny czas trwania operacji międzybankowych na terenie m.Warszawy w marcu b.r. w zakresie kredytów krótkoterminowych , jeśli średni czas dla 200 losowo wybranych przesylek międzybankowych wynosił 20 minut , przy 20% zmienności.

Przyjąć $P = 1 - \alpha = 0,90$.

Podaj wartość liczbowa i zinterpretuj: średni, maksymalny i względny błąd szacunku.

Wyjaśnij pojęcia: precyzji i dokładności szacunku. Jakie elementy mogą wpływać na ich wielkość?.

Zad.13.Wartość przeprowadzonych emisji obligacji komunalnych w losowo wybranych dużych miastach w Polsce (w mln. zł) w 2011 roku była następująca:

EMITENT	WARTOŚĆ EMISJI W MLN ZŁ
1.Gdynia	28
2.Łódź	29
3.Bielsko-Biała	100
4.Gdańsk	99,3
5.Kraków	15
6.Ostrów Wielkopolski	7,5
7.Lublin	90
8.Wałbrzych	20

Oceń punktowo i przedziałowo przeciętną wartość emisji obligacji komunalnych przyjmując $P=0,90$ i $P=0,99$. Określ wartość średniego , maksymalnego i względnego błędu szacunku.

Otrzymane wyniki skomentuj słownie.

Zad.14.Z jakim prawdopodobieństwem oszacowano frakcję (odsetek) dorosłych Polaków , którzy uważają, że polityka gospodarcza rządzącej koalicji jest właściwa, jeśli otrzymany przedział ma końce: 0,777 - 0,822, a oszacowany został na podstawie 1200 elementowej próby losowej, wśród której 960 respondentów miało zdanie na tak.

Czy liczebność 1200 osób w próbie byłaby konieczna do oszacowania przedziału ufności, gdyby przyjąć maksymalny błąd szacunku równy 3%.

Zad.15. Jak liczna powinna być próba dokumentów księgowych pobranych do kontroli przez biegłego w czasie badania bilansu firmy , aby przy współczynniku ufności 0,95 oszacować % dokumentów wadliwych , jeśli nie chcemy pomylić się o więcej niż o 4%?.

Zad.16. W losowej próbie 500 telewizorów typu „LX-A5” wykryto w czasie kontroli technicznej 35 usterek. Oszacuj punktowo i przedziałowo % wadliwych telewizorów przyjmując $P=0,95$.

Zad.17. W badaniu wynagrodzeń miesięcznych w przemyśle rolno-spożywczym wylosowano 200 pracowników. Dla wyników tej próby otrzymano: średnie miesięczne wynagrodzenie równe 850 zł, odchylenie standardowe 200 zł. Przyjmując współczynnik ufności równy 0,95 oszacuj metodą przedziałową poziom przeciętnego wynagrodzenia w przemyśle rolno-spożywczym.

Zad.18. Zakłada się, że miesięczne wydatki na żywność w rodzinach czteroosobowych posiadają rozkład $N(m; \sigma)$.

a/Oszacuj metodą przedziałową przeciętną wartość tych wydatków, jeśli dla losowej próby 10 gospodarstw domowych otrzymano średnią = 546 zł i $S(x) = 100$ zł, przyjmując $P=0,98$,

b/Jaki przedział otrzyma się , jeśli zostanie przyjęte założenie , że wydatki na żywność mają rozkład normalny $N(m; 100)$,

c/Jaki przedział ufności otrzyma się przy $P=0,98$ i $P=0,90$, jeśli założymy, że $x = 546$, $S(x) = 100$ a wyniki

pochoǳą z próby 200 losowo dobranych gospodarstw domowych,

d/Dla zbudowanych przedziałów ufności w punktach: a, b, c określ wielkość średniego, maksymalnego i wzglęǳnego błęǳu szacunku,

e/Jaki zapis posiada estymacja punktowa w przypadku posiadanych informacji w punktach: a, b, c.

Zad.19. Zakłada się, że czas pracy żarówek „AVOX” ma rozkład normalny $N(800; \sigma)$. Na podstawie 16 elementowej próby otrzymano: $S^2(x)=2700$. Przyjmując $P=0,95$ określ przedział ufności dla wariancji i odchylenia standardowego czasu pracy żarówek.

Zad.20. Wydatki półroczne (w tys. zł) na reklamę można uznać za cechę ciągłą o rozkładzie normalnym $N(m; \sigma)$. Do badania wylosowano próbę prostą 200 zakładów usługowych i otrzymano:

Półroczne wydatki na reklamę	0 - 5	5 - 10	10 - 15	15 - 20	Σ
Liczba zakładów	20	40	80	60	200

Przy współczynniku ufności $P = 0,96$ wyznacz przedział ufności dla:

a/ przeciętnych półrocznych wydatków na reklamę,

b/ odchylenia standardowego wydatków na reklamę,

c/ jak zmieni się precyzja szacunku i pewność , jeśli przyjąć współczynnik ufności : $P = 0,90$, $P = 0,99$.

Zad.21. a/ Oblicz jaka powinna być minimalna liczebność próby niezbędna do oszacowania odsetka zakładów , które wydają na reklamę półrocznie mniej niż 10 tys. zł z maksymalnym błędem szacunku równym 2% i z $P=0,99$.

b/Oszacuj drugi raz minimalną liczebność próby przy wykorzystaniu informacji zawartych w treści zadania nr 8.

Zad.22. Zakłada się, że waga produkowanych detali ma rozkład normalny . Na podstawie 10 losowo wybranych detali wyznaczono odchylenie standardowe wagi tych detali równe 0,5 kg. Należy:

a/ na poziomie $P=0,90$ oszacować przedział ufności dla wariancji wagi detali,

b/ jak zmieni się przedział ufności , gdy $P=1-\alpha = 0,90$,

c/ ile elementów należałoby dołosoować do próby , aby oszacować średnią wagę detali z dokładnością do 0,1 kg na poziomie $P=0,98$.

1. Hipotezy – parametryczne i nieparametryczne.

Zad.1. Na 800 zbadanych wypadków drogowych okazało się, że 320 zostało spowodowanych nadmierną szybkością. Na poziomie $\alpha= 0,05$ zweryfikuj hipotezę, że % wypadków spowodowanych nadmierną szybkością jest równy 35%.

Zad.2.Przy kontroli pracy dwóch central telefonicznych w losowo dobranym dniu stwierdzono, że na 200 połączeń w centrali A 16 było pomyłkowych, natomiast na 100 połączeń w centrali B złych połączeń było 10. Na poziomie $\alpha=0,05$ zweryfikuj hipotezę , że % złych połączeń jest jednakowy w obu centralach.

Zad.3. W hurtowni owoców cytrusowych wylosowano niezależnie 120 skrzynek z cytrusami. Okazało się , że w 16 skrzynkach znaleziono zepsute owoce. Na poziomie $\alpha=0,10$ oceń, czy partia badanych owoców zawiera więcej niż 5% skrzynek z zepsutymi owocami?.

Zad.4. W dwóch grupach firm: krajowych i zagranicznych przeprowadzono badanie zadłużeń z tytułu kredytów na działalność eksportową . Dla losowych prób : $n_1 = 150$ firm krajowych i $n_2 = 100$ firm zagranicznych otrzymano następujące informacje o wysokości zadłużenia:

firmy krajowe

średnia = 15,8 tys. zł

$S_1(x) = 3,5$ tys. zł

firmy zagraniczne

średnia = 35,0 tys. zł.

$S_2(x) = 15,0$ tys. zł

Czy na poziomie $\alpha= 0,01$ można uznać, że średni poziom zadłużenia w obu typach firm jest taki sam?

Przy jakim innym poziomie istotności decyzja weryfikacyjna może ulec zmianie?

Zad.5. Zmierzono czas reakcji na bodziec u 8 sportowców badanych w pracowni psychometrycznej przed i w 15 minut po wypiciu testowanego leku. Wyniki przed wypiciem leku były następujące(w sekundach):

0,22 0,18 0,16 0,19 0,20 0,23 0,17 0,25. Czasy reakcji po podaniu leku były następujące(w sekundach): 0,28 0,25 0,20 0,30 0,19 0,26 0,28 0,24 . Na poziomie istotności równym 0,05 zweryfikować hipotezę, że podawany do testowania lek wydłuża czas reakcji na bodziec. Uwaga: w procedurze weryfikacji hipotez zastosować dwa testy: dla dwu prób i dla jednej różnic wyników u tego samego pacjenta

Zad.6. Miesięczne wydatki na żywność w gospodarstwach domowych mają rozkład $N(m, \sigma)$.

Uważa się, że wartość przeciętna tych wydatków jest wyższa niż 400 zł.

Należy:

a/ zweryfikować prawdziwość tego sądu na poziomie istotności równym 0,01 , jeśli dla losowo wybranej grupy 10 rodzin otrzymano: przeciętną kwotę wydatków równą 480 zł i odchylenie standardowe 108 zł.

b / czy na poziomie istotności równym 5% można uważać, że odchylenie standardowe wydatków wynosi 90 zł?.

Zad.7. Czas pracy pewnego rodzaju baterii ma rozkład $N(m; 70)$. Na poziomie istotności $\alpha=0,05$ zweryfikuj hipotezę , że przeciętny czas pracy tego typu baterii wynosi ponad 500 godzin, jeśli dla losowo wybranych baterii otrzymano średnią równą 560 godzin.

Zad.8. Obserwując liczbę kilometrów jaką przebywają prywatne samochody w Polsce , otrzymano na losowej próbie 100 samochodów średnią równą 15.500 km i odchylenie standardowe 3.500 km Na poziomie istotności równym 0,01 należy zweryfikować hipotezę, że przeciętna liczba kilometrów przebytych przez prywatne samochody w Polsce wynosi 15.000.

Zad.9. Wyprodukowana partia detali jest uważana za dobrą wówczas, gdy odchylenie standardowe wymiaru ogółu wyprodukowanych detali nie przekracza 0,017 mm. W celu zbadania jakości kolejnej partii pobrano z niej losowo 6 detali i otrzymano następujące wymiary (w mm): 6,3 5,9 6,2 5,8 5,7 6,1.

a/ Czy otrzymane wyniki w próbie dają podstawę do uznania badanej partii za dobrą?; przyjąć $\alpha=0,10$.

b/ Przy jakim poziomie istotności decyzja weryfikacyjna może ulec zmianie?

Zad.10. Analiza rozkładu czasu pracy w godzinach nadliczbowych w dwóch losowo dobranych grupach pracowników z każdej zmiany dostarczyła następujących informacji:

- w grupie 100 pracowników pierwszej zmiany odnotowano średnią godzin nadliczbowych 4 godz. i odchylenie standardowe 1,2 godz.

- w grupie 120 pracowników drugiej zmiany odnotowano odpowiednio: średnią równą 6 godzin i odchylenie standardowe 2,4 godz.

Podać z jakim prawdopodobieństwem błędu I-go rodzaju można twierdzić, że średnio biorąc pracownicy pierwszej zmiany pracują krócej w czasie nadliczbowym od pracowników drugiej zmiany.

Zad. 11 . Dla próby 17 gospodarstw rolnych zaciągających kredyty w BGŻ zbadano poziom kwartalnych spłat , których rozkład jest normalny. Średnia arytmetyczna spłat w próbie wynosiła 3000 zł, odchylenie standardowe 420 zł. Bank twierdzi, że średnia wysokość spłat w rzeczywistości wynosi 3200 zł. Z jakim najniższym ryzykiem błędu I-go rodzaju można uznać różnicę pomiędzy wynikami badań a opinią banku za statystycznie istotną.

Zad.12. Większość dorosłych Polaków uważa, że średnie dochody osobiste w 1996 nie zmieniły się w porównaniu do 1995, kiedy wynosiły one przeciętnie 550 zł miesięcznie na osobę. W 1996 roku zbadano 81 osób, których średni dochód (w cenach stałych) wyniósł 580 zł, przy wariancji 10000 zł². Zakładając normalny rozkład dochodów należy sprawdzić z ryzykiem błędu I-go rodzaju równym 0,05 , czy powyższa większość ma rację.

Zad.13. Postawiono przypuszczenie, że empiryczny rozkład zysków brutto wtys. zł 1000 losowo dobranych firm w Polsce jest zgodny z rozkładem normalnym $N(150; 60)$. Utworzono 6 klas zysków i dla pięciu kolejnych z nich obliczono $\chi^2 = 12,3$. Klasa szósta liczyła 7 firm o zyskach 300 tys zł i więcej. Oblicz ostateczną wartość statystyki chi-kwadrat dla badanego rozkładu 1000 firm, przyjmując poziom istotności 0,05 sprawdź czy przypuszczenie było słuszne. Przy jakim poziomie istotności decyzja ulegnie zmianie. Odpowiedź uzasadnij.

Zad.14. W dwóch firmach przewozowych badano odległości przejazdów i otrzymano:

- dla firmy A : $n_1 = 15$; $S_1(x) = 158$ km

- dla firmy B : $n_2 = 10$; $S_2(x) = 283$ km

Czy na poziomie istotności $\alpha=0,01$ można uważać , że wariancje odległości przewozów w obu firmach są takie same.

Zad.15. W ciągu 100 dni obserwowano liczbę awarii w sieci wodno-kanalizacyjnej w mieście wojewódzkim i otrzymano:

Liczba awarii	0	1	2	3	4	5	Σ
Liczba dni	13	30	25	15	10	7	100

Czy można sądzić, że przeciętna ilość awarii w tym mieście jest równa 2, przyjmując poziom istotności równy 0,02.

W innym mieście wojewódzkim na podstawie obserwacji w ciągu 50 dni otrzymano średnią liczbę awarii

równą 3 i wariancję z próby równą 4. Czy na poziomie istotności równym 0,02 można sądzić, że przeciętna liczba awarii w obu miastach wojewódzkich jest taka sama.

Korzystając z szeregu empirycznego , przyjmując $\alpha=0,05$ oceń zgodność badanego rozkładu z rozkładem Poissona.

Zad.16. Struktura wieku losowej próby inwestorów giełdowych w pewnej grupie zawodowej w Polsce jest następująca:

Wiek w latach	15 - 25	25 - 35	35 - 45	45 - 55	55 - 65
Liczba osób	20	30	60	50	10

Oceń na poziomie istotności 0,01 zgodność badanego rozkładu z rozkładem normalnym.

Zad. 17. W mieście wojewódzkim wylosowano 200 kobiet i 150 mężczyzn pytając ich, czy byli by skłonni, przy wszystkich warunkach sprzyjających, mieć o jedno dziecko więcej. Odpowiedzi pozytywnej udzieliło 48 kobiet i 45 mężczyzn. Na poziomie istotności $\alpha=0.01$ zweryfikować hipotezę, że istnieje zgodność opinii kobiet i mężczyzn.

DODATKOWE ZADANIA

Zad 1

Zmienna losowa X ma rozkład normalny. Zakładając znajomość parametru σ należy oszacować metodą największej wiarygodności parametr m . Czy jest to estymator nieobciążony?

Zad 2

Własność nieobciążoności estymatora:

- a) Oznacza, że stosując estymator unikamy błędów losowych.
- b) Oznacza, że uzyskana ocena jest przeciętnie równa wartości szacowanego parametru.
- c) Oznacza, że warto (w miarę możliwości pokrycia kosztów) zwiększać liczebność próby.
- d) Oznacza, że stosując estymator unikamy błędów nielosowych.

Zad 3

Zmienna X ma rozkład normalny $N(m; \sigma)$. Poniższa tabela przedstawia parametr Θ , proponowany estymator T_n oraz wartość oczekiwaną i wariancję tego estymatora.

Lp.	Θ	T_n	$E(T_n)$	$D^2(T_n)$
1	σ^2	$s^2 = \frac{1}{n} \sum_i (X_i - \bar{X})^2$	$\frac{n-1}{n} \sigma^2$	$\frac{2\sigma^4(n-1)}{n^2}$
2	σ	$d = \frac{1}{n} \sum_i X_i - m $	$\frac{2}{\pi} \sigma$	$\frac{\sigma^2(\pi-2)}{n\pi}$
3	p	$w = \frac{x}{n}$	p	$\frac{p(1-p)}{n}$
4	m	mediana	m	$\frac{\pi\sigma^2}{2n}$
5	m	\bar{X}	m	$\frac{\sigma^2}{n}$
6	σ^2	$\hat{S}^2 = \frac{1}{n-1} \sum_i (X_i - \bar{X})^2$	σ^2	$\frac{2\sigma^4}{n-1}$

Które z proponowanych estymatorów są estymatorami zgodnymi?

Zad 4

Zbadano staż pracy 100 losowo wybranych urzędników uzyskując poniższe informacje.

Staż pracy w latach	Do 2	Do 6	Do 10	Do 14
Skumulowana częstość pracowników	0,1	0,4	0,8	0,9

Oszacuj punktowo odsetek urzędników, którzy pracują dłużej niż 10 lat.

Zad 5

Wylosowano próbę 100 kartonów jogurtów (zawierających po 48 opakowań) i zbadano je ze względu na liczbę opakowań w kartonie, które zostały uszkodzone podczas transportu do sklepu. Rozkład liczby uszkodzonych opakowań w kartonie przedstawiony jest poniżej.

Liczba uszkodzonych opakowań w kartonie	0	1	2	3
Liczba kartonów	40	30	20	10

Oszacować punktowo średnią liczbę uszkodzonych opakowań w kartonie.

Zad 6

Z populacji generalnej o rozkładzie normalnym wylosowano próbę prostą: 4, 2, 6, 5, 6, 3, 3, 5, 6, 2, 2. Zakładając, że wariancja w populacji generalnej jest równa wariancji z próby, należy

zbudować i zinterpretować 90% przedział ufności dla wartości oczekiwanej. Czy rozwiązanie się zmieni, jeśli owo założenie (o równości wariancji z próby i populacji) nie zostanie przyjęte?