1. W biochemicznym doświadczeniu badano czas życia komórek w pewnym środowisku. Dokonano ośmiu pomiarów uzyskując wyniki (w godzinach): 4.7, 5.3, 4.0, 3.8, 6.2, 5.5, 4.5, 6.0 ($\bar{x} = 5,\ s = 0,891227083$). Czy można uznać, że średni czas życia komórek w badanym środowisku wynosi 4 godziny?

Rozwiązanie.
Populacja: Komórki w pewnym środowisku.
Cecha: Czas życia komórki.
Niech X oznacza czas życia wylosowanej komórki.
Założenie: $X \sim N(\mu, \sigma^2)$, $\mu-$średni czas życia komórek w pewnym środowisku.

Cel $\to$ Zweryfikować hipotezę, że średni czas życia komórek wynosi 4:

$$H_0: \ \mu = 4 \text{ (średni czas wynosi 4)}$$
$$H_1: \ \mu \neq 4$$

Zadaję poziom istotności $\alpha = 0.05$

Rachunki: Obliczenia w pakiecie R
z1 $<-$c(4.7,5.3,4,3.8,6.2,5.5,4.5,6)
t.test(z1,m=4)

Fragment wydruku:
t = 3.1736, df = 7, p-value = 0.01563

Weryfikacja hipotezy: Ponieważ p-value $<$ 0.05, hipotezę odrzucam.
Wniosek: Średni czas życia komórek w pewnym środowisku nie wynosi 4.

2. W próbie dwustu warszawskich dzieci w wieku od sześciu do siedmiu lat rozpoznano osiemnaście przypadków astmy. Oszacować odsetek chorych na astmę w populacji wszystkich warszawskich dzieci w wieku od sześciu do siedmiu lat.

Rozwiązanie.
Populacja: dzieci warszawskie w wieku od sześciu do siedmiu lat.
Niech X oznacza liczbę przypadków astmy wśród wylosowanych dzieci.
Ponieważ rozmiar wylosowanej próby wynosi 200, a populacja jest bardzo liczna, możemy założyć, że $X \sim B(200, p)$, gdzie p oznacza prawdopodobieństwo wylosowania dziecka chorego na astmę.

Cel $\to$ Oszacować p

Zadaję poziom ufności $1 - \alpha = 0.95$.

Rachunki: Obliczenia w pakiecie R.
prop.test(18,200)$conf.int

Fragment wydruku:
[1] 0.0557122 0.1406878

Wniosek: Odsetek dzieci chorych na astmę wynosi co najmniej 5.6%, ale nie więcej niż 14.1%. Zaufanie do wniosku wynosi 95%.

3. Obserwujemy uczucie suchości skóry wśród ludzi z atopowym zapaleniem skóry (w skrócie z AZS) i bez AZS. Pobrana próba dała wyniki:

	uczucie suchości skóry	bez uczucia suchości skóry
bez AZS	2025	9392
z AZS	2591	4512

Interesuje nas różnica pomiędzy populacjami osób z AZS oraz bez AZS, ze względu na odsetek osób, które mają uczucie suchości skóry. Wyznaczyć odpowiedni do zagadnienia przedział ufności oraz zinterpretować uzyskany wynik.

Rozwiązanie.

Populacja 1: zbiorowowść osób z AZS; Populacja 2: zbiorowość osób bez AZS.
Niech X oznacza liczbę osób z uczuciem suchości skóry wśród osób, które stanowią próbę pobraną z populacji 1.
Niech Y oznacza liczbę osób z uczuciem suchości skóry wśród osób, które stanowią próbę pobraną z populacji 2.
Można założyć, że $X \sim B(2025 + 9392, p_1)$ oraz $Y \sim B(2591 + 4512, p_2)$.

Cel $\rightarrow$ wyznaczyć przedział ufności dla $p_1 - p_2$

Obliczenia w R:
z7< −matrix(c(2025,9392,2591,4512),ncol=2,byrow=T)
colnames(z7)< −c("uczucie suchości","brak uczucia suchości") # ta linia nie jest konieczna do obliczeń
rownames(z7)< −c("bez AZS","z AZS") # ta linia nie jest konieczna do obliczeń
prop.test(z7)$conf.int

Fragment wydruku z R:
[1] -0.2007290 -0.1740878

Wniosek: Różnica między badanymi odsetkami wynosi co najmniej 17%, ale nie więcej niż 20%. Uczucie suchości skóry częściej odczuwają osoby z AZS (bo końce przedziału są ujemne).

Uwaga: Na różnicę między odsetkami możemy popatrzeć przez pryzmat ryzyka względnego, albo ilorazu szans. W środowisku medycznym „najbardziej popularny" jest iloraz szans. I jeszcze jedna rzecz. Wyznaczony w zadaniu przedział ufności jest przedziałem warunkowym (przy warunku na rozmiary prób).

4. W pewnym doświadczeniu farmakologicznym z podawaniem dwu preparatów badano potęgowanie narkozy. Dla preparatu A otrzymano następujące przedłużenia narkozy: 4, 3, 5, 2, 4, 6, 4, 5 ($\bar{x} = 4.125,\ s = 1.246423455$), a dla preparatu B: 6, 10, 8, 9, 9, 10, 8, 7 ($\bar{x} = 8.375,\ s = 1.407885953$). Czy można uznać, że preparaty w różnym stopniu przedłużają czas narkozy, jeżeli doświadczenie przeprowadzono na dobranych losowo osobach?

Rozwiązanie.

Populacja: pacjenci
Czynnik: Rodzaj preparatu używanego w narkozie.
Niech X oznacza czas, o który jest przedłużona narkoza przy preparacie A.
Niech Y oznacza czas, o który jest przedłużona narkoza przy preparacie B.
Ponieważ cechy X oraz Y realizują się na wylosowanych pacjentach, możemy przyjąć, że są to cechy losowe.
Założenie: $X \sim N(\mu_1, \sigma^2)$, $Y \sim N(\mu_2, \sigma^2)$

Zadanie możemy rozwiązać na przykład za pomocą przedziału ufności dla $\mu_1 - \mu_2$.
W tym celu zadajemy poziom ufności $1 - \alpha = 0.95$ i wykonujemy rachunki w R:
z8.A< −c(4,3,5,2,4,6,4,5)
z8.B< −c(6,10,8,9,9,10,8,7)
t.test(z8.A,z8.B,m=0,equal.var=T)

Wydruk z R:
95 percent confidence interval:
-5.677831 -2.822169
t = -6.3929, df = 13.797, p-value = 1.793e-05

Przedział ufności dla $\mu_1 - \mu_2$ wynosi $(-5.677831, -2.822169)$. Ponieważ przedział ufności nie zwiera zera, możemy uznać, że średnie μ_1, μ_2 różnią się między sobą. Wniosek jest taki, że preparaty mają inny wpływ na przedłużenie narkozy.

Uwaga. Żeby zmierzyć czas przedłużenia narkozy, nie potrzebujemy każdego pacjenta poddawać dwukrotnej narkozie. Zmierzony czas narkozy wystarczy porównać z przeciętnym czasem trwania narkozy standardowej. W przypadku szacowanie różnicy (takiej, jak w zadaniu) nawet i tego nie trzeba robić. Dlaczego?