Wydział Matematyki i Nauk Informacyjnych
Politechnika Warszawska

STATYSTYKA MATEMATYCZNA

z pakietem R

V. Weryfikacja hipotez, cz. I

Przemysław Grzegorzewski
Konstancja Bobecka-Wesołowska
Marek Gągolewski

Spis treści

Spis treści 1

1 Wprowadzenie 2

1.1 Testy normalności 2

1.2 Testy parametryczne dla jednej i dwóch prób 2

1.2.1 Test dla wartości oczekiwanej dla jednej próby 2

1.2.2 Test dla wartości oczekiwanej dla dwóch prób niezależnych 3

1.2.3 Test dla wariancji dla dwóch prób niezależnych 3

1.2.4 Test dla wskaźnika struktury (proporcji) dla jednej próby 3

1.2.5 Test dla wskaźników struktury dla dwóch prób niezależnych 4

2 Zadania rozwiązane 4

3 Zadania do rozwiązania 12

4 Wskazówki i odpowiedzi 15

1 Wprowadzenie

1.1 Testy normalności

Do weryfikacji hipotezy o normalności rozkładu badanej cechy można użyć w programie R np. funkcji `shapiro.test()`, za pomocą której przeprowadzany jest test Shapiro-Wilka. Wektor wejściowy `x`, zawierający wartości próbki, dla której chcemy przeprowadzić test zgodności z rozkładem normalnym, podajemy jako argument tej funkcji: `shapiro.test(x)`. Dopuszczalna liczność próby wynosi między 3 a 5000 elementów. Test ten jest narzędziem z wyboru w wielu zastosowaniach.

Inne dostępne w R testy normalności są zamieszczone w pakiecie `nortest`[1]. Są to: test Cramera-von Misesa (`cvm.test()`), test Andersona-Darlinga (`ad.test()`), test Lillieforsa (`lillie.test()`), test chi-kwadrat Pearsona (`pearson.test()`), test Shapiro-Francii (`sf.test()`).

Nadto, przydatną, poglądową metodą jest wykres kwantylowy dla rozkładu normalnego (ang. *normal Q-Q plot* bądź *quantile-quantile plot*), zaimplementowany jako funkcja `qqnorm()`.

Więcej informacji na temat testów zgodności poznamy przy okazji omawiania testów nieparametrycznych w VI części materiałów.

1.2 Testy parametryczne dla jednej i dwóch prób

W R mamy do dyspozycji funkcje pozwalające przeprowadzać testy dla wartości oczekiwanej (dla jednej i dwóch prób) w modelu normalnym, test równości dwóch wariancji w modelu normalnym oraz testy dla wskaźnika struktury (proporcji) w modelu dwupunktowym (dla jednej i dwóch prób).

1.2.1 Test dla wartości oczekiwanej dla jednej próby

Niech $(X_1, \ldots, X_n)$ będzie próbą z rozkładu normalnego $\mathrm{N}(\mu, \sigma)$ o nieznanych parametrach μ i σ. Do weryfikacji hipotez dotyczących wartości oczekiwanej μ postaci:

$$\begin{aligned} &H : \mu = \mu_0 \\ &K : \mu \neq \mu_0; \quad K' : \mu < \mu_0; \quad K'' : \mu > \mu_0 \end{aligned}$$

można użyć funkcji `t.test()`. Pierwszym argumentem tej funkcji jest wektor zawierający wartości próbki, na podstawie których szacujemy μ. Wartość μ_0 podajemy jako drugi argument pisząc: `mu` $= \mu_0$, zaś postać hipotezy alernatywnej jako trzeci, pisząc: `alternative="two.sided"` (dla K), `"less"` (dla K') lub `"greater"` (dla K''). Domyślnie przeprowadzany jest test dwustronny.

Np. dla wektora `x` zawierającego wartości próbki:
`t.test(x, mu=1)` — testujemy hipotezę $H : \mu = 1$ przeciwko $K : \mu \neq 1$.
`t.test(x, mu=1, alternative="less")` — testujemy hipotezę $H : \mu = 1$ przeciwko $K : \mu < 1$.

[1] Który może nie być automatycznie ładowany przy starcie R-a. Jeżeli odpowiednie funkcje nie są dostępne, należy wykonać polecenie: `library(nortest);`.

1.2.2 Test dla wartości oczekiwanej dla dwóch prób niezależnych i obserwacji parami zależnych

Dla dwóch prób niezależnych: $(X_1, \ldots, X_{n_1})$ z rozkładu $\mathrm{N}(\mu_1, \sigma_1)$ oraz $(Y_1, \ldots, Y_{n_2})$ z rozkładu $\mathrm{N}(\mu_2, \sigma_2)$ możemy za pomocą funkcji `t.test()` weryfikować hipotezy o równości dwóch wartości średnich:

$$H : \mu_1 = \mu_2$$
$$K : \mu_1 \neq \mu_2; \quad K' : \mu_1 < \mu_2; \quad K'' : \mu_1 > \mu_2$$

Jako pierwszy i drugi argument funkcji `t.test()` należy podać w tym przypadku wektory `x` i `y` zawierające wartości próbek. Jako trzeci argument podajemy postać hipotezy alternatywnej (domyślnie przeprowadzany jest test dwustronny). Jako kolejny argument funkcji `t.test()` podajemy informację, czy wariancje σ_1^2 i σ_2^2 są sobie równe, czy też nic o tym nie wiadomo. Jeśli wariancje są równe to piszemy: `var.equal=TRUE` (domyślnie równość wariancji nie jest zakładana).

Np. dla wektora `x` zawierającego wartości pierwszej próbki i wektora `y` zawierającego wartości drugiej próbki:
`t.test(x,y)` — testujemy hipotezę $H : \mu_1 = \mu_2$ przeciwko $K : \mu_1 \neq \mu_2$, bez zakładania równości wariancji.
`t.test(x,y,alternative="less",var.equal=TRUE)` — testujemy hipotezę $H : \mu_1 = \mu_2$ przeciwko $K' : \mu_1 < \mu_2$, przy założeniu $\sigma_1^2 = \sigma_2^2$.

Jeśli zamiast prób niezależnych $(X_1, \ldots, X_n)$ z rozkładu $\mathrm{N}(\mu_1, \sigma_1)$ oraz $(Y_1, \ldots, Y_n)$ z rozkładu $\mathrm{N}(\mu_2, \sigma_2)$, mamy obserwacje zależne w parach, to jako argument funkcji `t.test()` podajemy dodatkowo: `paired=TRUE`.

1.2.3 Test dla wariancji dla dwóch prób niezależnych

Dla dwóch prób niezależnych: $(X_1, \ldots, X_{n_1})$ z rozkładu $\mathrm{N}(\mu_1, \sigma_1)$ oraz $(Y_1, \ldots, Y_{n_2})$ z rozkładu $\mathrm{N}(\mu_2, \sigma_2)$, hipotezę o równości wariancji

$$H : \sigma_1^2 = \sigma_2^2$$
$$K : \sigma_1^2 \neq \sigma_2^2; \quad K' : \sigma_1^2 < \sigma_2^2; \quad K'' : \sigma_1^2 > \sigma_2^2$$

można zweryfikować testem F, za pomocą funkcji `var.test()`. Jako pierwszy i drugi argument tej funkcji podajemy wektory `x` i `y`, zawierające wartości próbek. Jako trzeci argument podajemy postać hipotezy alternatywnej (domyślnie przeprowadzany jest test dwustronny).

Np. dla wektora `x` zawierającego wartości pierwszej próbki i wektora `y` zawierającego wartości drugiej próbki:

`var.test(x,y)` — testujemy hipotezę $H : \sigma_1^2 = \sigma_2^2$ przeciwko $K : \sigma_1^2 \neq \sigma_2^2$.
`var.test(x,y,alternative="less")` — testujemy hipotezę $H : \sigma_1^2 = \sigma_2^2$ przeciwko $K' : \sigma_1^2 < \sigma_2^2$.

1.2.4 Test dla wskaźnika struktury (proporcji) dla jednej próby

Niech $(X_1, \ldots, X_n)$ będzie próbą z rozkładu dwupunktowego Bern(p). Do weryfikacji hipotez dotyczących wartości wskaźnika struktury (proporcji) p postaci:

$$H : p = p_0$$
$$K : p \neq p_0; \quad K' : p < p_0; \quad K'' : p > p_0$$

można użyć funkcji `binom.test()` lub `prop.test()`, przy czym w drugim przypadku przeprowadzany jest test asymptotyczny.

Pierwszym argumentem obu tych funkcji jest liczba jedynek w naszej próbie (odpowiadająca liczbie elementów posiadających interesującą nas cechę), a drugim — liczność próby n. Wartość p_0 podajemy jako trzeci argument pisząc: `p` $= p_0$, zaś postać hipotezy alernatywnej jako czwarty (domyślnie przeprowadzany jest test dwustronny).

Na przykład:
`prop.test(4,100,p=0.05)` — testujemy hipotezę $H : p = 0{,}05$ przeciwko $K : p \neq 0{,}05$, dla próby 100 elementowej, w której 4 elementy mają interesującą nas cechę.
`prop.test(4,100,p=0.05,alternative="less")` — testujemy hipotezę $H : p = 0{,}05$ przeciwko $K' : p < 0{,}05$.

1.2.5 Test dla wskaźników struktury dla dwóch prób niezależnych

Dla dwóch prób niezależnych: $(X_1, \ldots, X_{n_1})$ z rozkładu Bern(p_1) oraz $(Y_1, \ldots, Y_{n_2})$ z rozkładu Bern(p_2), możemy za pomocą funkcji `prop.test()` weryfikować hipotezy o równości dwóch wskaźników struktury (test asymptotyczny):

$$H : p_1 = p_2$$
$$K : p_1 \neq p_2; \quad K' : p_1 < p_2; \quad K'' : p_1 > p_2$$

Jako pierwszy argument funkcji `prop.test()` należy podać wektor (k_1, k_2) a jako drugi wektor (n_1, n_2), gdzie n_i są licznościami próbek, a k_i — liczbami jedynek w każdej próbce (odpowiadającymi liczbom elementów posiadających interesującą nas cechę), $i = 1, 2$. Jako trzeci argument podajemy postać hipotezy alternatywnej (domyślnie przeprowadzany jest test dwustronny).

Na przykład:
`prop.test(c(4,6),c(100,120))` — testujemy hipotezę $H : p_1 = p_2$ przeciwko $K : p_1 \neq p_2$, dla próbek 100 i 120 elementowych, w których jest, odpowiednio, 4 i 6 elementów wyróżnionych.
`prop.test(c(4,6),c(100,120),alternative="less")` — testujemy hipotezę $H : p_1 = p_2$ przeciwko $K' : p_1 < p_2$.

2 Zadania rozwiązane

Zadanie 5.1. Wytrzymałość na ciśnienie wewnętrzne jest ważną charakterystyką jakościową szkła butelek. Pewna rozlewnia chce stosować butelki, których średnia wytrzymałość przewyższa 1,20 N/mm^2. Na podstawie dotychczasowych doświadczeń wiadomo, że rozkład ciśnienia jest normalny z odchyleniem standardowym 0,07 N/mm^2. Pobrano próbę losową 20 butelek, które następnie umieszczono w maszynie hydrostatycznej, zwiększając ciśnienie aż do zniszczenia butelki. Otrzymano następujące wyniki (w N/mm^2):

1.36, 1.14, 1.27, 1.15, 1.20, 1.29, 1.27, 1.18, 1.23, 1.36, 1.38, 1.37, 1.30, 1.21, 1.33, 1.28, 1.32, 1.29, 1.33, 1.25.

Na poziomie istotności 0,04 stwierdzić, czy dana partia butelek spełnia postawione wymagania jakościowe.

Rozwiązanie.
Mamy tutaj do czyniania z próbą $(X_1, \ldots, X_n)$ z rozkładu normalnego $N(\mu, \sigma)$ o nieznanym parametrze μ i znanym $\sigma = 0{,}07$.

Do weryfikacji hipotezy zerowej $H_0 : \mu = \mu_0$ przeciw alternatywie $K : \mu > \mu_0$ użyjemy tzw. testu z, ponieważ ma on większą moc niż test t. Wyznaczymy go ręcznie, gdyż nie został on zamieszczony w standardowej bibliotece R-a.

Test ten jest następującej postaci. Odrzucamy hipotezę zerową, gdy statystyka testowa

$$z = \frac{\overline{\mathbf{X}} - \mu_0}{\sigma}\sqrt{n}. \tag{1}$$

przyjmie wartość z przedziału krytycznego: $W_\alpha = [z_{1-\alpha}, +\infty)$, gdzie $z_{1-\alpha}$ jest kwantylem rzędu $1 - \alpha$ rozkładu normalnego standardowego.

W naszym zadaniu: $n = 20$, $\alpha = 0{,}04$, $\mu_0 = 1{,}2$.

Wyznaczamy wartość statystyki testowej:

```
> x <- c(1.36, 1.14, 1.27, 1.15, 1.20, 1.29, 1.27, 1.18, 1.23, 1.36,
  1.38, 1.37, 1.30, 1.21, 1.33, 1.28, 1.32, 1.29, 1.33, 1.25);
> mu0 <- 1.2
> sigma <- 0.07
> (z <- (mean(x)-mu0)/sigma*sqrt(length(x)))

[1] 4.823518
```

Teraz wartość odpowiedniego kwantyla wynosi

```
> qnorm(0.96)

[1] 1.750686
```

Jako że wartość statystyki testowej z należy do przedziału krytycznego $W_{0,04} = [1{,}75069, +\infty)$, to odrzucamy hipotezę zerową. Zatem możemy stwierdzić, że wytrzymałość na ciśnienie badanych butelek istotnie ($\alpha = 0{,}04$) przewyższa 1,2 N/mm^2.

$\boxdot$

Zadanie 5.2. Nominalna waga netto kawy sprzedawanej w opakowaniu szklanym winna wynosić 150 g. Występuje jednakże duża zmienność wagi. Istotnie, próba losowa siedmiu słoiczków kawy konkretnej marki sprzedawanej w sieci handlowej Żuczek wykazała następujące wagi netto (w gramach): 142, 151, 148, 151, 145, 150, 141. Zakładając normalność rozkładu wagi, przetestuj hipotezę głoszącą, że waga netto tej marki kawy wynosi faktycznie 150 g. Przyjmij poziom istotności $\alpha = 0{,}05$.

Rozwiązanie.
Mamy próbę $(X_1, \ldots, X_n)$ z rozkładu normalnego $N(\mu, \sigma)$ o dwu nieznanych parametrach μ oraz σ. Do weryfikacji hipotezy zerowej $H : \mu = 150$ przeciw alternatywie $K : \mu \neq 150$ użyjemy testu t. Jest on w naszym ulubionym programie zaimplementowany jako funkcja `t.test()`.

```
> kawa <- c(142, 151, 148, 151, 145, 150, 141);
> mean(kawa) # zobaczmy, jaka jest średnia waga kawy w pobranej próbce

[1] 146.8571
```

```
> t.test(kawa, mu=150)

        One Sample t-test

data:  kawa
t = -1.9704, df = 6, p-value = 0.0963
alternative hypothesis: true mean is not equal to 150
95 percent confidence interval:
 142.9542 150.7601
sample estimates:
mean of x
 146.8571
```

Wartość statystyki testowej wynosi $t = -1{,}9704$. Porównujemy p-value $= 0{,}0963$ z żądanym poziomem istotności testu $\alpha = 0{,}05$. Ponieważ p-value $> \alpha$, to nie odrzucamy hipotezy zerowej. Oznacza to, że nie mamy podstaw, by powiedzieć, że średnia waga kawy istotnie ($\alpha = 0{,}05$) różni się od 150 g.

Ponadto, przypomnijmy, funkcja `t.test()` zwraca nam m.in. granice (domyślnie 95%) przedziału ufności dla średniej wagi kawy.

⊡

Zadanie 5.3. Wylosowana niezależnie z partii żarówek 12-elementowa próba dała następujące wyniki pomiarów czasu świecenia (w godzinach): 2852, 3060, 2631, 2819, 2805, 2835, 2955, 2595, 2690, 2723, 2815, 2914.

a) Wyznacz 97% przedział ufności dla maksymalnego średniego czasu świecenia żarówek.

b) Czy średni czas świecenia żarówek jest istotnie krótszy od 2900 godzin? Przyjmij poziom istotności $\alpha = 0{,}05$.

Rozwiązanie.
Ponieważ w zadaniu nie mamy informacji na temat rozkładu, z którego pochodzi próba, sprawdzimy najpierw testem Shapiro-Wilka, czy można przyjąć założenie o jego normalności. W teście tym hipotezy są postaci: H_0 : rozkład próby jest normalny *versus* K : rozkład ten nie jest normalny.

```
> zar <- c(2852, 3060, 2631, 2819, 2805, 2835, 2955, 2595, 2690, 2723, 2815, 2914)
> shapiro.test(zar)

        Shapiro-Wilk normality test

data:  zar
W = 0.9747, p-value = 0.9532
```

Ponieważ $W = 0{,}9747$, p-value $= 0{,}9532 > \alpha = 0{,}05$, to możemy przyjąć, że nasza próba pochodzi z rozkładu normalnego $\mathrm{N}(\mu, \sigma)$ (o nieznanych μ i σ), jako że nie mamy podstaw do odrzucenia hipotezy zerowej na zadanym poziomie istotności.

Do wyznaczenia jednostronnego przedziału ufności dla μ oraz zweryfikowania hipotezy $H : \mu = 2900$ przeciw $K : \mu < 2900$ możemy teraz użyć funkcji `t.test()`. Gdyby założenie o normalności nie było spełnione, powinniśmy raczej w tym miejscu użyć np. jakiegoś testu nieparametrycznego. Narzędzia te poznamy w następnej części materiałów dydaktycznych.

```
> t.test(zar, conf.level=0.97, mu=2900, alternative="less")

        One Sample t-test

data:  zar
t = -2.3855, df = 11, p-value = 0.01807
alternative hypothesis: true mean is less than 2900
97 percent confidence interval:
     -Inf 2888.819
sample estimates:
mean of x 
 2807.833 
```

Uwaga

Zauważmy, że podanie parametru `conf.level` nie wpływa na wynik testu. Służy on tylko do określania poziomu ufności dla przedziału ufności. Parametr `alternative` zarówno determinuje rodzaj przedziału, jak i postać hipotezy alternatywnej.

97%-przedział ufności dla maksymalnego średniego czasu świecenia żarówek wynosi $[0, 2888{,}819]$ godzin (zwróćmy uwagę, że czas nie może być ujemny).

Wartość statystyki testowej testu t dla wartości oczekiwanej wynosi $t = -2{,}3855$. Jako że p-value $= 0{,}0181 < \alpha = 0{,}05$, odrzucamy więc hipotezę zerową na rzecz alternatywnej; średni czas świecenia żarówek jest istotnie krótszy niż 2900 godzin.

$\boxdot$

Zadanie 5.4. W stopie metalicznym pewnego typu zastosowano dwa różne pierwiastki utwardzające. Wyniki pomiarów twardości stopów utwardzanych obiema metodami wyglądają następująco:

Metoda I	$145, 150, 153, 148, 141, 152, 146, 154, 139, 148$
Metoda II	$152, 150, 147, 155, 140, 146, 158, 152, 151, 143, 153$

Przyjmuje się, że twardość ma rozkład normalny oraz że odchylenia standardowe dla obu metod są równe. Czy na podstawie przeprowadzonych pomiarów można stwierdzić, że średnia twardość stopu utwardzanego drugą metodą przewyższa średnią twardość stopu utwardzanego pierwszą metodą? Przyjmij poziom istotności $\alpha = 0{,}05$.

Rozwiązanie.
Mamy tutaj do czynienia z dwoma niezależnymi próbami: $(X_1, \ldots, X_{n_1})$ z rozkładu $\mathrm{N}(\mu_1, \sigma_1)$ oraz $(Y_1, \ldots, Y_{n_2})$ z rozkładu $\mathrm{N}(\mu_2, \sigma_2)$, przy czym wiadomo *a priori*, że $\sigma_1^2 = \sigma_2^2$.

Weryfikujemy $H : \mu_1 = \mu_2$ przeciw $K : \mu_1 < \mu_2$. Użyjemy testu t dla dwóch prób o takich samych wariancjach:

```
> m1 <- c(145,150,153,148,141,152,146,154,139,148)
> m2 <- c(152,150,147,155,140,146,158,152,151,143,153)
> t.test(m1, m2, alternative="less", var.equal=T)
```

```
        Two Sample t-test

data:  m1 and m2
t = -0.9466, df = 19, p-value = 0.1779
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
     -Inf 1.758418
sample estimates:
mean of x mean of y
 147.6000  149.7273
```

Wnioski pozostawiamy Czytelnikowi.

Ponadto, sprawdźmy, jakie uzyskalibyśmy wyniki, gdybyśmy nie wiedzieli o równości wariancji.

```
> t.test(m1, m2, alternative="less") # domyślnie: var.equal=FALSE

        Welch Two Sample t-test

data:  m1 and m2
t = -0.9496, df = 18.973, p-value = 0.1771
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
     -Inf 1.746503
sample estimates:
mean of x mean of y
 147.6000  149.7273
```

Uwaga

Jeżeli $\sigma_1^2 = \sigma_2^2$, to drugi z zastosowanych testów ma mniejszą moc (jest bardziej zachowawczy). W przypadku przeciwnym użycie pierwszego jest nieuzasadnione.

⊡

Zadanie 5.5. Spośród pracowników pewnego przedsiębiorstwa wylosowano niezależnie 15 pracowników fizycznych i 9 pracowników umysłowych. Otrzymano następujące dane dotyczące stażu pracy (w latach):

Umysłowi	$14, 17, 7, 33, 2, 24, 26, 22, 12$
Fizyczni	$13, 15, 3, 2, 25, 4, 1, 18, 6, 9, 20, 11, 5, 1, 7$

Wiadomo, że rozkład stażu pracy w przedsiębiorstwie jest normalny. Zweryfikuj hipotezę głoszącą, że średni staż pracy pracowników fizycznych jest istotnie krótszy niż staż pracy pracowników umysłowych. Przyjmij poziom istotności $\alpha = 0{,}05$.

Rozwiązanie.

Mamy dwie próby niezależne: $(X_1, \ldots, X_{n_1})$ z rozkładu $\mathrm{N}(\mu_1, \sigma_1)$ oraz $(Y_1, \ldots, Y_{n_2})$ z rozkładu $\mathrm{N}(\mu_2, \sigma_2)$.

Weryfikujemy $H : \mu_1 = \mu_2$ przeciw $K : \mu_1 > \mu_2$. Użyjemy do tego testu t-Studenta dla dwóch prób.

Najpierw sprawdzamy, czy można przyjąć założenie o równości wariancji. Do weryfikacji hipotezy $H : \sigma_1^2 = \sigma_2^2$ przeciw $K : \sigma_1^2 \neq \sigma_2^2$ użyjemy funkcji `var.test()`, przeprowadzającej test F:

```
> um <-  c(14,17,7,33,2,24,26,22,12)
> fiz <- c(13,15,3,2,25,4,1,18,6,9,20,11,5,1,7)
> var.test(um, fiz)

        F test to compare two variances

data:  um and fiz
F = 1.725, num df = 8, denom df = 14, p-value = 0.3557
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.5250833 7.1238787
sample estimates:
ratio of variances
           1.72505
```

Zatem można przyjąć, że $\sigma_1^2 = \sigma_2^2$. Teraz test t:

```
> t.test(um, fiz, alternative="greater", var.equal=T)

        Two Sample t-test

data:  um and fiz
t = 2.2937, df = 22, p-value = 0.01587
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 2.038754      Inf
sample estimates:
mean of x mean of y
17.444444  9.333333
```

Więc średni staż pracy pracowników fizycznych jest istotnie krótszy niż staż pracy pracowników umysłowych.

$\boxdot$

Zadanie 5.6. Grupę 10 dzieci poddano pewnemu testowi pamięci. Po pewnym czasie, w którym dzieci wykonywały w domu ćwiczenia usprawniające pamięć, poddano je ponownie równoważnemu testowi. Na podstawie wyników zamieszczonych w tabeli stwierdź, czy zaproponowane ćwiczenia w istotny sposób usprawniają pamięć. Przyjmij poziom istotności równy 5%.

Dziecko	1	2	3	4	5	6	7	8	9	10
Wynik przed	27,	21,	34,	24,	30,	27,	33,	31,	22,	27
Wynik po	29,	32,	29,	27,	31,	26,	35,	30,	29,	28

Rozwiązanie.
Tym razem nasze zmienne losowe, odpowiadające wynikom poszczególnych dzieci przed i po ćwiczeniach, tzn. $X_1, \ldots, X_{10}$ oraz $Y_1, \ldots, Y_{10}$, są zależne parami. Wynik bowiem, dajmy na to, Jasia „przed” nie jest oderwany od jasiowego wyniku „po” — wpływają na niego m.in. różne predyspozycje dziecka, jego sposób funkcjonowania w szkole itp.

Ponieważ nie mamy informacji na temat rozkładów, z których pochodzą próby, sprawdzimy najpierw testem Shapiro-Wilka, czy można przyjąć założenie o normalności danych z obydwu próbek.

```
> przed <- c(27,21,34,24,30,27,33,31,22,27)
> po    <- c(29,32,29,27,31,26,35,30,29,28)
> shapiro.test(przed)

        Shapiro-Wilk normality test

data:  przed
W = 0.9501, p-value = 0.6694


> shapiro.test(po)

        Shapiro-Wilk normality test

data:  po
W = 0.9498, p-value = 0.666
```

bądź:

```
> shapiro.test(przed-po) # test dla różnic w wynikach

        Shapiro-Wilk normality test

data:  przed - po
W = 0.9355, p-value = 0.504
```

Zatem możemy przyjąć, że $(X_1, \ldots, X_n)$ jest próbą z rozkładu $\mathrm{N}(\mu_1, \sigma_1)$ (test Shapiro-Wilka, $W = 0{,}9501$, p-value $= 0{,}669$) oraz że $(Y_1, \ldots, Y_n)$ jest próbą z rozkładu $\mathrm{N}(\mu_2, \sigma_2)$ (test Shapiro-Wilka, $W = 0{,}9498$, p-value $= 0{,}666$).

Do weryfikacji hipotezy zerowej $H_0 : \mu_1 = \mu_2$ względem $K : \mu_1 < \mu_2$ (badamy, czy nastąpiła poprawa) użyjemy testu t dla par.

```
> t.test(przed, po, alternative="less", paired=T)

        Paired t-test

data:  przed and po
t = -1.4302, df = 9, p-value = 0.09322
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
      -Inf 0.5634468
sample estimates:
mean of the differences
                     -2
```

Toteż nie mamy podstaw by odrzucić H_0 na poziomie istotności $\alpha = 0{,}05$. Ćwiczenia nie wydają się istotnie usprawniać pamięci dzieci.

Uwaga

Zauważmy, że w celu weryfikacji naszej hipotezy możemy również użyć równoważnego powyższemu testu t dla jednej próby:

```
> t.test(przed-po, mu=0, alternative="less");
```

```
        One Sample t-test

data:  przed - po
t = -1.4302, df = 9, p-value = 0.09322
alternative hypothesis: true mean is less than 0
95 percent confidence interval:
      -Inf 0.5634468
sample estimates:
mean of x
       -2
```

⊡

Zadanie 5.7. W czasie poprawnej pracy maszyny frakcja wytwarzanych przez nią elementów wadliwych nie powinna przekraczać 4%. Jeżeli liczba ta będzie większa, wówczas należy podjąć czynności mające na celu wyregulowanie procesu produkcji. Pracownik zajmujący się kontrolą jakości pobrał próbkę losową 200 elementów i znalazł w niej 14 elementów wadliwych. Czy zaistniała sytuacja wymaga wyregulowania procesu produkcji? Zweryfikować odpowiednią hipotezę na poziomie istotności 0,05.

Rozwiązanie.
Mamy próbę $(X_1, \ldots, X_n)$ z rozkładu dwupunktowego Bern(p). Do weryfikacji hipotezy $H_0 : p = 0{,}04$ przeciw $K > 0{,}04$, możemy użyć funkcji `prop.test()` bądź `binom.test()`.

```
> prop.test (14, 200, p=0.04, alternative="greater")

        1-sample proportions test with continuity correction

data:  14 out of 200, null probability 0.04
X-squared = 3.9388, df = 1, p-value = 0.02359
alternative hypothesis: true p is greater than 0.04
95 percent confidence interval:
 0.04371856 1.00000000
sample estimates:
   p
0.07


> prop.test (14, 200, p=0.04, alternative="greater", correct=F) # korekta ciągłości

        1-sample proportions test without continuity correction

data:  14 out of 200, null probability 0.04
X-squared = 4.6875, df = 1, p-value = 0.01519
alternative hypothesis: true p is greater than 0.04
95 percent confidence interval:
 0.04570864 1.00000000
sample estimates:
   p
0.07


> binom.test(14, 200, p=0.04, alternative="greater") # test dokładny
```

```
        Exact binomial test

data:  14 and 200
number of successes = 14, number of trials = 200, p-value = 0.03121
alternative hypothesis: true probability of success is greater than 0.04
95 percent confidence interval:
 0.04281265 1.00000000
sample estimates:
probability of success
                  0.07
```

Wyniki wszystkich testów wskazują, że należy odrzucić hipotezę zerową na poziomie istotności 0,05. Zatem proces produkcji został najprawdopodobniej rozregulowany.

⊡

Zadanie 5.8. 455 spośród 700 absolwentów techników i 517 spośród 1320 absolwentów liceów nie zdało egzaminu wstępnego z matematyki na politechnikę. Czy na podstawie powyższych wyników można stwierdzić, że absolwenci techników są słabiej przygotowani do egzaminu z matematyki niż absolwenci liceów?

Rozwiązanie.
Mamy tutaj dwie próby niezależne: $(X_1, \ldots, X_{n_1})$ z rozkładu $\mathrm{Bern}(p_1)$ oraz $(Y_1, \ldots, Y_{n_2})$ z rozkładu $\mathrm{Bern}(p_2)$. Do weryfikacji hipotezy $H_0 : p_1 = p_2$ przeciw $K : p_1 > p_2$ użyjemy testu dla dwu proporcji zaimplementowanego w funkcji `prop.test()`

```
> prop.test(c(455,517), c(700,1320), alternative="greater")

        2-sample test for equality of proportions with continuity correction

data:  c(455, 517) out of c(700, 1320)
X-squared = 121.2477, df = 1, p-value < 2.2e-16
alternative hypothesis: greater
95 percent confidence interval:
 0.2202584 1.0000000
sample estimates:
   prop 1    prop 2
0.6500000 0.3916667
```

Wnioski pozostawiamy Czytelnikowi.

⊡

3 Zadania do rozwiązania

Zadanie 5.9. Biolog, badający pewien gatunek ryb, pobrał losową próbę 15 ryb i zmierzył ich długość. Otrzymał następujące wyniki (w mm):

92, 88, 85, 82, 89, 86, 81, 66, 75, 61, 78, 76, 91, 82, 82.

Zakładając, że rozkład długości ryb badanego gatunku jest normalny, zweryfikuj hipotezę, że ich średnia długość przekracza 78 mm. Przyjmij poziom istotności $\alpha = 0{,}01$.

Zadanie 5.10. Pewien księgowy przypuszcza, że przeciętne saldo na kontach klientów jego firmy jest mniejsze niż 31 tys. €. Żeby to sprawdzić, pobrał losową próbę kont, otrzymując następujące wyniki dotyczące przeciętnego salda (w tys. €):

30.0, 30.0, 29.9, 31.3, 32.0, 32.0, 32.1, 30.5, 32.3, 29.5, 27.8, 27.3, 31.1, 30.7, 24.5, 28.3, 31.3, 32.7, 33.3, 26.8.

Czy prawdziwe jest przypuszczenie księgowego? Zweryfikuj stosowną hipotezę na poziomie istotności 0,01.

Zadanie 5.11. Na podstawie danych zawartych w pliku `samochody.csv` zweryfikuj przypuszczenie, że średnia moc silnika samochodów wyprodukowanych w latach 1979–1981 wynosi 84 KM (wykorzystaj zmienne `moc` i `rok`). Przyjmij poziom istotności 0,01.

Zadanie 5.12. Oszacowano przecietną długość życia w wybranych losowo 18 krajach. Wyniki przedstawia poniższa tabela:

Kraj	Długość życia	Kraj	Długość życia	Kraj	Długość życia
Argentyna	70,5	Japonia	79	Sudan	53
Etiopia	51,5	Kenia	61	Tajwan	75
Niemcy	76	Meksyk	72	Tajlandia	68,5
Indie	57,5	Maroko	64,5	Turcja	70
Iran	64,5	RPA	64	Ukraina	70,5
Włochy	78,5	Hiszpania	78,5	USA	75,5

Czy na podstawie tych danych możemy twierdzić, że średnia długość życia przekracza 62 lata? Przyjmij poziom istotności 0,05.

Zadanie 5.13. Na podstawie danych dotyczących parametrów kilku wybranych marek samochodów (plik `samochody.csv`) stwierdź, czy występuje statystycznie istotna różnica w przyspieszeniu samochodów produkowanych w USA i w Japonii. Przyjmij poziom istotności $\alpha = 0{,}05$.

Zadanie 5.14. Badano wytrzymałość 20 losowo wybranych wsporników betonowych, przy czym 10 z nich wykonano metodą tradycyjną, a pozostałe niedawno opatentowaną, nową metodą. Wyniki pomiarów (w MPa) podano w poniższej tabeli:

Metoda tradycyjna	$53, 51, 62, 55, 59, 56, 61, 54, 47, 57$
Nowa metoda	$62, 55, 61, 58, 54, 49, 56, 60, 52, 63$

Czy na podstawie tych danych można stwierdzić, że wytrzymałość wsporników wykonanych nową metodą przewyższa istotnie wytrzymałość wsporników wykonanych metodą tradycyjną? Przyjmij poziom istotności 0,04.

Zadanie 5.15. Badano liczbę recept wypisywanych w ciągu 14 losowo wybranych dni przez pewnych dwóch lekarzy. Otrzymano następujące wyniki:

Lekarz I	$19, 21, 15, 17, 24, 12, 19, 14, 20, 18, 23, 21, 17, 12$
Lekarz II	$17, 15, 12, 12, 16, 15, 11, 13, 14, 21, 19, 15, 11, 10$

Zakładając, że badana cecha ma rozkład normalny, zweryfikuj przypuszczenie, że lekarz I wypisuje średnio więcej recept niż lekarz II. Przyjmij poziom istotności 0,05.

Zadanie 5.16. Badano przeciętną długość filmów produkowanych przez dwie konkurujące ze sobą firmy. W tym celu wylosowano do badania kilka filmów i otrzymano następujące dane (w minutach):

Długości filmów produkcji A	$102, 86, 98, 109, 92, 102, 95, 120$
Długości filmów produkcji B	$81, 165, 97, 134, 92, 87, 114, 120, 95, 136, 170$

Czy można twierdzić, że przeciętna długość filmów produkcji A przewyższa przeciętną długość filmów produkcji B? Zweryfikuj stosowną hipotezę na poziomie istotności 0,01.

Zadanie 5.17. Badano wpływ nowego leku na zmianę poziomu pewnej substancji we krwi (w mg/ml). W tym celu zmierzono poziom tej substancji u 8 losowo wybranych osób, a następnie, po upływie 30 minut od podania owego leku, powtórzono badanie na tej samej grupie osób. Otrzymano następujące wyniki:

Pacjent	1	2	3	4	5	6	7	8
Poziom przed	2.76,	5.18,	2.68,	3.05,	4.10,	7.05,	6.60,	4.79
Poziom po	7.02,	3.10,	5.44,	3.99,	5.21,	10.26,	13.91,	14.53

Czy na podstawie powyższych danych można stwierdzić, że nowy lek powoduje istotne podwyższenie poziomu owej substancji we krwi? Przyjmij poziom istotności $\alpha = 0{,}05$ oraz założenie o normalności rozkładu badanej cechy.

Zadanie 5.18. Badano wagę (w kilogramach) losowo wybranych palaczek przed i 5 tygodni po rzuceniu przez nich palenia papierosów. Otrzymano następujące wyniki:

Palaczka	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Waga przed	67,	65,	62,	62,	66,	65,	61,	63,	64,	71,	69,	65,	61,	60,
Waga po	69,	71,	65,	67,	74,	62,	69,	64,	70,	68,	73,	71,	67,	62,

Czy na podstawie powyższych danych można stwierdzić, że rzucenie palenia wpływa na wzrost wagi palącej papierosy kobiety? Przyjmij poziom istotności $\alpha = 0{,}05$.

Zadanie 5.19. Wśród pracowników naukowych pewnej uczelni przeprowadzono ankietę dotyczącą stażu pracy i stanu cywilnego. Otrzymano następujące wyniki liczby osób wedle stanu i stażu (w latach):

Staż pracy	Panna/kawaler	Mężatka/żonaty
0–5	6	20
5–10	8	20
10–15	3	60
15–20	2	25
20–25	1	15

Zweryfikuj hipotezę, że w grupie mężatek i żonatych, odsetek osób pracujących na owej uczelni dłużej niż 15 lat wynosi 0,3. Przyjmij poziom istotności $\alpha = 0{,}05$.

Zadanie 5.20. Na podstawie danych zawartych w pliku `samochody.csv`:

a) Podaj przedział ufności dla odsetka samochodów mających moc większą niż 80 KM (wykorzystaj zmienną `moc`). Przyjmij poziom ufności 0,95.

b) Zweryfikuj hipotezę, że ponad 50% samochodów ma moc większą niż 80 KM. Przyjmij poziom istotności 0,06.

c) Rozpatrz te problemy ponownie, tym razem ograniczając się do samochodów produkowanych tylko w Ameryce i Japonii (wykorzystaj nadto zmienne `producent` i `legenda`).

4 Wskazówki i odpowiedzi

Wskazówka do zadania 5.10. Tu i dalej, jeśli zachodzi potrzeba, zweryfikuj hipotezę o normalności rozkładu danych.

Wskazówka do zadania 5.13. Tu i dalej, jeśli zachodzi potrzeba, zweryfikuj hipotezę o równości wariancji w obydwu próbach.

Wskazówka do zadania 5.17. Obserwacje parami zależne.

Wskazówka do zadania 5.18. Obserwacje parami zależne.