Testy statystyczne w R.
Testy t-Studenta

Agnieszka Goroncy

UNIWERSYTET
MIKOŁAJA KOPERNIKA
W TORUNIU
Wydział Matematyki
i Informatyki

Testy statystyczne

Informacje dotyczące testów statystycznych można znaleźć w plikach:

Testy statystyczne - podstawowe informacje

Testy statystyczne - teoria rozszerzona

p-wartość

Testy statystyczne przeprowadzone w środowisku R nie dadzą jednoznacznej odpowiedzi na to, czy hipotezę zerową należy przyjąć, czy też odrzucić. W zamian otrzymujemy coś bardziej użytecznego - tzw. *p*-wartość (`p-value`).

Czym jest *p*-wartość? Jest to **graniczny poziom istotności** - najmniejszy, przy którym zaobserwowana wartość statystyki testowej prowadzi do odrzucenia hipotezy zerowej. Jest to wiec taki poziom istotności, przy którym zmienia się decyzja testu (zwiększając poziom istotności testu α od zera: najpierw dla najmniejszych poziomów istotności nie mamy podstaw do odrzucenia hipotezy zerowej, a później, po przekroczeniu *p*-wartości, zaczynamy odrzucać hipotezę zerową).

p-wartość pozwala bezpośrednio ocenić wiarygodność hipotezy: im *p*-wartość jest większa, tym bardziej hipoteza zerowa jest prawdziwa. Mała *p*-wartość świadczy przeciwko hipotezie zerowej.

p-wartość: podsumowanie

Znajomość p-wartości pozwala przeprowadzić testowanie dla dowolnego ustalonego poziomu istotności testu $\alpha \in (0, 1)$:

- odrzucamy hipotezę zerową, gdy p-wartość $\leqslant \alpha$,
- nie mamy podstaw do odrzucenia hipotezy zerowej, gdy p-wartość $> \alpha$.

Kiedy p-wartość uznamy za „małą”, a kiedy za „dużą”?

Zarówno poziom istotności testu α, jak i p-wartość, są liczbami z przedziału (0,1).

Poziom istotności, czyli maksymalne dopuszczalne prawdopodobieństwo popełnienia błędu I rodzaju (odrzucenia hipotezy zerowej, gdy jest ona prawdziwa), z reguły przyjmuje się bardzo mały, np. 0,05 (standardowo), 0,01.

W związku z tym p-wartość równa np. 0,2, mimo, że w przedziale (0,1) wydaje się „mała”, dla nas już będzie „duża” i nie będzie przemawiała za odrzuceniem hipotezy zerowej. Z kolei p-wartość równa np. 2,2e-16 (=0,00000000000000022) jest dla nas „mała” i skłoni do odrzucenia hipotezy zerowej.

Uwaga: Jeśli w zadaniu poziom istotności testu α nie jest zadany, to bierzemy poziom standardowy, czyli $\alpha = 0{,}05$.

Testy t-Studenta

Testy t-Studenta to parametryczne testy dla jednej lub dwóch prób, polegające na testowaniu hipotezy dotyczącej wartości oczekiwanej (nazywane też testami istotności). **Założenia stosowania:** pomiary podlegają rozkładowi normalnemu bądź rozmiar próbki jest duży (z reguły wystarczy $n > 30$).

Funkcja **t.test()** pozwala przeprowadzić test t-Studenta.

Pierwszym argumentem funkcji jest wektor z danymi. Kolejne ważne (opcjonalne) argumenty są następujące:

- `y` - drugi wektor z danymi,
- `alternative`: two.sided / less / greater - określa hipotezę alternatywną (domyślnie: dwustronna),
- `mu` - liczba określająca hipotetyczną wartość oczekiwaną rozkładu (bądź różnicę wartości oczekiwanych w przypadku dwóch prób), domyślnie równa 0,
- `paired` - wektor logiczny określający, czy test ma być przeprowadzony dla danych powiązanych (zależnych), domyślnie FALSE. Jeżeli paired=TRUE, wówczas wektory z danymi muszą być równej długości.

Uwaga: W przypadku, gdy nie są spełnione założenia testu, alternatywą dla testu t-Studenta jest test Wilcoxona.

Test t-Studenta: jedna próba

Przykład 1. Według normy technicznej wykonanie obróbki mechanicznej jednego pierścienia stalowego powinno zajmować szlifierzowi 22 minuty. Wylosowano 16 stanowisk roboczych, dla których czas obróbki wynosił odpowiednio: 23, 22, 24, 23, 26, 25, 26, 24, 23, 23, 27, 23, 22, 24, 25, 24 minuty. Zakładając, że czas obróbki ma rozkład normalny, zweryfikuj na poziomie istotności $\alpha = 0,05$ hipotezę $H_0 : \mu = 22$ wobec hipotezy alternatywnej

(a) $H_1 : \mu \neq 22$,

(b) $H_1 : \mu < 22$,

(c) $H_1 : \mu > 22$.

Jeśli przeprowadzamy testowanie hipotez „ręcznie", czyli bez pomocy komputera, to zazwyczaj to robimy za pomocą tzw. obszaru krytycznego (patrz plik *Testy statystyczne - podstawowe informacje*). Przedstawimy najpierw tę metodę.

Przykład 1. Obliczenia

Mamy:

$$\begin{aligned} n &= 16, \\ \bar{x} &= \frac{1}{16}(23 + \ldots + 24) = 24, \\ s &= \sqrt{\frac{1}{15}((23-24)^2 + \ldots + (24-24)^2)} \simeq 1,46. \\ \mu_0 &= 22, \\ \alpha &= 0,05. \end{aligned}$$

Hipoteza zerowa testu ma postać

$$H_0 : \mu = 22.$$

Statystyka testowa wynosi

$$T_n = \sqrt{n}\frac{\bar{x} - \mu_0}{s} = \sqrt{16}\frac{24-22}{1,46} \simeq 5,48.$$

Przykład 1. Obliczenia, c.d.

(a) Hipoteza alternatywna $H_1 : \mu \neq 22$
Korzystając z *Tablice kwantyli rozkładu t-Studenta*, mamy: $F^{-1}_{t_{n-1}}(1-\frac{\alpha}{2}) = F^{-1}_{t_{15}}(0{,}975) = 2{,}131$ (skoro $n-1=15$, to należy patrzeć 15-ty wiersz i kolumnę 0,975). Zatem obszar krytyczny ma postać

$$\begin{aligned} K &= \left(-\infty, -F^{-1}_{t_{n-1}}(1-\frac{\alpha}{2})\right) \cup \left(F^{-1}_{t_{n-1}}(1-\frac{\alpha}{2}), +\infty\right) \\ &= (-\infty, -2{,}131) \cup (2{,}131, \infty). \end{aligned}$$

Decyzja: $T_n \in K$, zatem **odrzucamy** $H_0 : \mu = 22$ **na rzecz** $H_1 : \mu \neq 22$.

(b) Hipoteza alternatywna $H_1 : \mu < 22$
Z *Tablice kwantyli rozkładu t-Studenta* otrzymujemy: $F^{-1}_{t_{n-1}}(1-\alpha) = F^{-1}_{t_{15}}(0{,}95) = 1{,}753$.
Zatem obszar krytyczny ma postać

Przykład 1. Obliczenia, c.d.

$$K = \left(-\infty, -F^{-1}_{t_{n-1}}(1-\alpha)\right) = (-\infty, -1{,}753).$$

Decyzja: $T_n \notin K$, zatem **nie mamy podstaw do odrzucenia** $H_0 : \mu = 22$**.**

(c) Hipoteza alternatywna $H_1 : \mu > 22$
Mamy z *Tablice kwantyli rozkładu t-Studenta*:
$F^{-1}_{t_{n-1}}(1-\alpha) = F^{-1}_{t_{15}}(0{,}95) = 1{,}753$.
Zatem obszar krytyczny ma postać

$$K = (F^{-1}_{t_{n-1}}(1-\alpha), \infty) = (1{,}753, \infty).$$

Decyzja: $T_n \in K$, zatem **odrzucamy** $H_0 : \mu = 22$ **na rzecz** $H_1 : \mu > 22$.

A teraz jak to robimy za pomocą programu `R`.

```
> x=c(23, 22, 24, 23, 26, 25, 26, 24, 23, 23, 27, 23,
+ 22, 24, 25, 24)
```

Przykład 1.

Dla sprawdzenia wyniku ze str. 8 możemy obliczyć wartość statystyki testowej (nie jest to konieczne):

```
> n=length(x)
> T=sqrt(n)*(mean(x)-22)/sd(x)
```

Jeśli hipoteza alternatywna H_1 brzmi: $\mu \neq 22$ (przypadek (a)):

```
> t.test(x, mu=22)
```

Skoro p-wartość$=6.372e{-}05 < 0.05$, odrzucamy H_0 na rzecz H_1.

Jeśli hipoteza alternatywna H_1 brzmi: $\mu < 22$ (przypadek (b)):

```
> t.test(x, mu=22, alternative="l")
```

Skoro p-wartość$=1 > 0.05$, nie mamy podstaw do odrzucenia hipotezy H_0.

Jeśli hipoteza alternatywna H_1 brzmi: $\mu > 22$ (przypadek (c)):

```
> t.test(x, mu=22, alternative="g")
```

Skoro p-wartość$=3.186e{-}05 < 0.05$, odrzucamy H_0 na rzecz H_1.

Test t-Studenta: jedna próba

Przykład 2 Wygenerujmy próbkę losową pochodzącą z rozkładu jednostajnego na przedziale $(-1, 1)$. Przetestujmy najpierw, czy możemy przyjąć, że wartość oczekiwana rozkładu, z którego pochodzi próbka, wynosi -2:

```
> x<-runif(40,-1,1)
> t.test(x, mu=-2)
```

Skoro p-wartość jest znacznie mniejsza od standardowej 0,05, odrzucamy hipotezę o średniej równej -2.

Sprawdźmy, czy test zachowa się prawidłowo przy wartości oczekiwanej równej 0 (mu$=$0 to wartość domyślna):

```
> t.test(x)
```

Zgodnie z oczekiwaniami, p-wartość jest większa od 0,05, co nie pozwala na odrzucenie hipotezy zerowej. Możemy zatem przyjąć, że wartość oczekiwana rozkładu, z którego pochodzą dane, wynosi 0.

Uwaga: Próbka nie pochodzi z rozkładu normalnego, ale ma liczebność większą od 30.

Test t-Studenta: dwie próby niezależne

Przykład 3 Sprawdzamy równość wartości oczekiwanych rozkładów dwóch prób:

```
> x1=c(50, 48, 39, 53, 51, 49, 50, 55, 47, 53)
> x2=c(46, 54, 58, 49, 52, 51, 58, 46)
```

Skoro próbki mają różne rozmiary, na pewno są to próbki **niezależne**.

Standardowy test t-Studenta zakłada, że wariancje rozkładów obu próbek są zbliżone. Sprawdzimy więc najpierw, czy to ma miejsce. Stosujemy **test na równość wariancji**:

```
> var.test(x1,x2)
```

Test ten testuje hipotezę H_0: wariancje rozkładów próbek są takie same przeciw hipotezie H_1: wariancje rozkładów próbek są różne. Skoro p-wartość wynosi $0{,}8278 > 0{,}05$, nie mamy podstaw do odrzucenia H_0.

Test t-Studenta: dwie próby niezależne

Zatem wykonujemy test t-Studenta dla próbek o jednakowej wariancji:

```
> t.test(x1,x2,var.equal=TRUE)
```

Otrzymana p-wartość wynosi $0{,}3146 > 0{,}05$, więc nie mamy podstaw sądzić, że wartości oczekiwane obydwu rozkładów są różne.

Gdyby się okazało, ze w teście na równość wariancji odrzucilibyśmy hipotezę zerową, to wykonalibyśmy test t-Studenta dla dwóch prób za pomocą polecenia:

```
> t.test(x1,x2)
```

(czyli opcja var.equal=FALSE jest domyślna). W tym przypadku braku równości wariancji do wartości statystyki testowej wprowadza się specjalną korektę.

Uwaga: Przy małej liczebności prób w tym przykładzie niezbędne jest założenie o normalności rozkładów.

Test t-Studenta: dwie próby zależne

Przykład 4 Mamy dwa zestawy danych równej wielkości, które reprezentują wagi osób stosujących specyfik odchudzający przed i po kuracji. Chcemy sprawdzić, czy ów specyfik jest skuteczny:

```
> y1=c(65, 79, 103, 64, 78, 80, 66, 97, 104, 67)
> y2=c(64, 76, 100, 66, 77, 84, 64, 96, 110, 60)
```

Mamy dwie próbki jednakowego rozmiaru. Powiedziane jest, że pierwsza próbka reprezentuje wagi osób przed kuracją, a druga wagi *tych samych osób* po kuracji. Więc, są to próbki powiązane, czy innymi słowy zależne. Sprawdzamy, czy próbki mają w przybliżeniu takie same wariancje:

```
> var.test(y1,y2)
```

Skoro p-wartość wynosi $0{,}7857 > 0{,}05$, nie mamy podstaw sądzić, że wariancje próbek są istotnie różne. Zatem wykonujemy test t-Studenta dla próbek zależnych o jednakowej wariancji. Specyfik jest skuteczny, gdy waga osób po kuracji się zmniejszyła.

Test t-Studenta: dwie próby zależne

Więc,

```
> t.test(y1,y2,var.equal=TRUE,paired=TRUE,
+ alternative="g")
```

lub

```
> t.test(y2,y1,var.equal=TRUE,paired=TRUE,
+ alternative="l")
```

Skoro p-wartość wynosi $0{,}3124 > 0{,}05$, nie mamy podstaw do odrzucenia H_0, czyli niestety stwierdzamy, że specyfik nie działa na tyle, aby średnia waga przed kuracją była większa niż średnia waga po kuracji.

Uwaga: Test t-Studenta dla dwóch prób zależnych jest tożsamy z testem t-Studenta dla jednej próby, będącej różnicą wartości z obu próbek.

Uwaga: Przy małej liczebności prób w tym przykładzie niezbędne jest założenie o normalności rozkładów.

Odwoływanie się do wyników testu

Każdy test w R zwraca nam pewne informacje, do których możemy się odwoływać poprzez odpowiednią nazwę zmiennej.

Przykład:

```
> wynik=t.test(x1,x2,var.equal=TRUE)
> # sprawdzamy jakie informacje mamy dostępne:
> wynik[]
> wynik$p.value
> wynik$alternative
> wynik$data.name
```

Literatura

- Jacek Koronacki, Jan Mielniczuk, **Statystyka dla studentów kierunków technicznych i przyrodniczych**, WNT, Warszawa, 2001
- Przemysław Biecek, **Przewodnik po pakiecie** R, Oficyna Wydawnicza GiS, Wrocław, 2011
- Joseph Adler, **R in a Nutshell**, O'Reilly Media, 2009