Iwona Mróz,

Instytut Fizyki Doświadczalnej,

Uniwersytet Wrocławski

**„Podstawy statystyki i analizy danych” – materiały do wykładu nr 11 z dnia 20.12.2019 r**

Niniejsze materiały mają charakter roboczy. Bardzo proszę o zgłaszanie zauważonych błędów, braków, niedociągnięć i niejasności. Prośba dotyczy też przypisów. Z góry dziękuję za pomoc☺.

**Wybrane testy statystyczne służące do porównywania dwóch grup (populacji)**

**Testy parametryczne**

***Porównanie dwóch grup niezależnych***

*Opracowano na podstawie podręcznika Andrzeja Stanisza „Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny”, StatSoft Polska Sp. z o.o., Kraków 2006, tom 1, str. 221-225.*

Uwaga ogólna: testy parametryczne, czyli takie, które porównują parametry rozkładów badanych cech są lepsze od testów nieparametrycznych. Dlatego zawsze warto sprawdzić, czy dla danego zadania możemy zastosować test parametryczny, a dopiero jeśli jest to niemożliwe, zastosować test nieparametryczny.

Bardzo często chcemy porównać średnie w dwóch grupach niezależnych. Służy do tego zestaw testów parametrycznych. Niezwykle ważne jest, aby sprawdzić czy są spełnione założenia danego testu. Ogólnie, założenia dla omawianych tutaj testów parametrycznych sprowadzają się do:

1. Przestrzegania zasad randomizacji – o tym założeniu, bardzo ważnym dla doświadczalnictwa, jeszcze nie mówiliśmy, ale powiemy.
2. Właściwej ocenie, czy badane grupy są zależne czy niezależne,
3. Ocenie, czy próba jest duża, czy mała. A Stanisz przyjmuje, ze próby są duże jeśli każda z nich liczy więcej niż 50 obserwacji. Wielkość próby wpływa na wybór właściwego testu. Generalnie, próby powinny mieć zbliżoną liczebność, choć przy testach dla grup niezależnych ścisła równoliczność prób nie jest wymagana. Oczywiście dla grup zależnych obie próby są równoliczne. O liczebności też jeszcze więcej sobie powiemy.
4. Założenie o normalności rozkładu (w każdej z badanych grup).
5. Założenie, czy wariancje są jednorodne. Do oceny jednorodności wariancji służą, między innymi, testy: F, Bartletta, Levene’a czy Browna-Forsythe’a

Przywołując A. Stanisza mamy następujący schemat postępowania:

Hipotezy statystyczne:

Niech *μ1* i *μ2* oznaczają średnie w porównywanych populacjach.

H0:

H1: lub lub (zob. materiały dodatkowe do wykładu, część 1)

1. Jeżeli obie próby są duże (*n1* > 50i *n2* > 50, niektórzy autorzy przy porównywaniu dwóch grup uznają próby za duże jeśli ich liczebność jest większa od 30),stosujemy test z. Statystyka testu z jest następująca:

 (11.1)

*mi, si2, ni, i = 1,2* to, odpowiednio, średnia, wariancja i liczebność i-tej próby.

Statystyka z ma standardowy rozkład normalny N(0,1).

2. Jeżeli przynajmniej jedna próba nie jest duża, to pytamy o normalność rozkładu w obu grupach (populacjach). Jeżeli przynajmniej jeden rozkład nie jest normalny, to wybieramy testy nieparametryczne.

3. Jeżeli rozkłady w obu grupach są normalne, pytamy, czy znamy wariancje obu populacji. Jeżeli wariancje te są znane, stosujemy test U, którego statystyka ma postać:

 (11.2)

Gdzie dla *i = 1,2*: *mi, ni,* to, odpowiednio, średnia i liczebność próby, a *i2* to znana wariancja w i-tej populacji. Statystyka U ma standardowy rozkład normalny N(0,1).

3. Jeżeli wariancje populacji nie są znane, pytamy, czy wariancje te są równe. Równość wariancji możemy przetestować przy pomocy wspomnianych wyżej testów na jednorodność wariancji.

3a. Jeżeli wariancje są równe, stosujemy test t (to jest najbardziej popularny test Studenta) o statystyce:

 (11.3)

*mi, si2, ni, i = 1,2* to, odpowiednio, średnia, wariancja i liczebność i-tej próby.

Jeżeli H0 jest prawdziwa statystyka t ma rozkład t-Studenta o stopniach swobody.

3b. Jeżeli wariancje nie są równe, stosujemy inny test, np. test Cochrana-Coxa, którego statystyka C ma postać: 

 (11.4)

*mi, si2, ni, i = 1,2* to, odpowiednio, średnia, wariancja i liczebność i-tej próby.

Statystyka testu Cochrana-Coxa na rozkład t-Studenta, liczbę stopni swobody *df* można wyliczyć.

***Porównanie dwóch grup zależnych***

***Opracowano na podstawie podręcznika Andrzeja Stanisza „Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny”, StatSoft Polska Sp. z o.o., Kraków 2006, tom 1, str. 225.***

W przypadku porównywania dwóch grup zależnych możemy zastosować test parametryczny należący do grupy testów Studenta i nazywany testem t-Studenta dla grup zależnych lub testem różnic.

W przypadku grup zależnych liczebność obu prób jest zawsze taka sama, a pomiarów dokonuje się dla tych samych obiektów, np. przed i po zastosowaniu procedury badawczej. Badana cecha statystyczna musi być ilościowa i mierzalna.

Niech liczebność każdej próby wynosi *n*. Oznaczymy *xi* oraz *yi*, *i = 1,…,n* odpowiadające sobie obserwacje, odpowiednio przed i po przeprowadzeniu procedury. Tworzymy zbiór różnic:  .

Założenia testu:

1. Obowiązuje przestrzeganie zasad randomizacji – powiemy o nich za jakiś czas.
2. Rozkład różnic powinien być rozkładem normalnym. UWAGA: Założenie o normalności rozkładu różnic znajdą Państwo m.in. w podręczniku A. Stanisza, na który się powołuję. W literaturze spotykałam też założenie o normalności rozkładów badanej cechy w obu próbach. Nie sprawdzałam, czy oba założenia są równoważne, ale można to zrobić.

Hipotezy:

H0: , gdzie jest średnią różnicą miedzy wynikami „przed” i „po” dla populacji generalnej. Innymi słowy, hipoteza zerowa mówi, że w populacji średnie „przed” i „po” są takie same. Mówiąc kolokwialnie, dla populacji procedura nie działa.

H1: . Można też sformułować hipotezy jednostronne.

Statystyka testu ma postać:

 (11.5)

Gdzie jest średnią arytmetyczną różnic z próby, - odchyleniem standardowym różnic z próby, a *n* oznacza liczbę obserwacji w każdej z prób.

Statystyka *t* ma rozkład t-Studenta o .

Jeżeli nie możemy przeprowadzić testu różnic, np. z powodu braku spełnienia założeń, wykonujemy testy nieparametryczne: test Wilcoxona lub test znaków (zob. poniżej).

**Testy nieparametryczne**

***Porównanie dwóch grup niezależnych***

***Opracowano na podstawie podręcznika Andrzeja Stanisza „Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny”, StatSoft Polska Sp. z o.o., Kraków 2006, tom 1, str. 371-373.***

Pamiętajmy, że testy nieparametryczne porównują rozkłady, nie parametry rozkładów. Testów takich jest wiele. Jeżeli naszym zadaniem jest porównanie dwóch grup niezależnych, testem często stosowanym jest test omówiony poniżej U Manna-Whitneya. Innym wartym poznania testem jest test Kołmogorova-Smirnova, którego w na naszym wykładzie raczej nie damy rady omówić.

Test U Manna-Whitneya opiera się na procedurze nadawania rang (rangowania).

Nadawanie danym rang (rangowanie, rank procedure):

Przeprowadzenie testu nieparametrycznego często wymaga uporządkowania danych. Porządkowanie danych jest możliwe, gdy dane można rozpatrywać przynajmniej w skali porządkowej. Uporządkowanie danych pozwala na nadanie im tzw. rang.

Procedura nadawania rang przebiega następująco:

- porządkujemy rosnąco wszystkie obserwacje (z obu rozpatrywanych prób),

- zaczynając od wartości najmniejszej lub największej przyporządkowujemy obserwacjom kolejne liczny naturalne,

- w przypadku, gdy mamy dane o wartościach jednakowych, przypisujemy im tzw. rangi wiązane, czyli średnie arytmetyczne rang, które byłyby im nadane.

Test U Manna-Whitneya (porówanie dwóch grup niezależnych)

Założenia:

- Badana cecha statystyczna ma w populacji rozkład typu ciągłego,

- dane można rozpatrywać w skali porządkowej (można je uporządkować),

- pobrano dwie próbki o liczebnościach n1 i n2.

Hipotezy:

A Stanisz podaje hipotezy H0 i H1 w oparciu o dystrybuanty. W innych źródłach, np. w podręczniku Aczela, mogą Państwo spotkać hipotezy sformułowane inaczej. Na przykład Aczel określa H0 jako „oba rozkłady są identyczne” , Według H1 jako „oba rozkłady nie są identyczne”.

Według A. Stanisza:

Niech F(x) i G(x) oznaczają dystrybuanty w rozpatrywanych populacjach.

Weryfikacja hipotezy:

H0: F(x) = G(x)

wobec alternatywnej:

H1: F(x)  G(x).

Weryfikacja hipotezy opiera się na statystyce U:

U = n1n2 + n1(n1+1)/2 – R1 (11.5)

gdzie R1 jest sumą rang przypisanych danym z pierwszej próby. Rozkład statystyki U sprawdza się w tablicach lub komputerowo, gdyż nie jest ona prosta. Ilustrację poglądową statystyki U przedstawiają J.E. Freud i B.M. Perles Perle podręczniku „Modern Elementary Statistics”.

Jeżeli próby są duże, co dla testu U Manna-Whitneya oznacza, że n1 >20 i n2 >20, zamiast statystyki U posługujemy się statystyką Z:

Z = (R1-R2-(n1-n2)(n+1)/2)/(sqrt(n1n2(n+1)/3) (11.6)

gdzie n jest całkowitą liczbą obserwacji (n = n1+n2), a R2 jest sumą rang przypisanych danym z drugiej próby. Statystyka Z ma w przybliżeniu rozkład normalny.

Uwaga! Dla podanej postaci testu ze względu na rozkład statystyki U hipotezę zerową odrzucamy, gdy obliczona wartość statystyki jest mniejsza od wartości krytycznej.

***Porównanie dwóch grup zależnych***

***Opracowano na podstawie podręcznika Andrzeja Stanisza „Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny”, StatSoft Polska Sp. z o.o., Kraków 2006, tom 1, str. 382.***

Porównanie dwóch grup zależnych wymaga pamiętania wyników obserwacji w parach (przed i po procedurze). Tak jak w przypadku testu (parametrycznego) Studenta dla dwóch grup zależnych, pracujemy na różnicach. W najprostszym ujęciu, filozofia postępowania jest taka: nadajemy różnicom znaki (+) lub (-) i patrzymy, jak układa się sekwencja znaków. Znaki różnic (+) i (-) mogą opisywać zmiany cech jakościowych, czyli np. wystąpienie zjawiska.

Hipoteza zerowa mówi o braku systematycznego efektu wywołanego procedurą, zatem prawdopodobieństwa: P(+) = P(-). Ten wariant testu nazywa się testem znaków.

Jego bardziej rozwinięta forma to test Wilcoxona, w którym patrzymy nie tylko na znaki różnic, ale też ich wartości bezwzględne oraz kolejność. Porządkujemy różnice i nadajemy im rangi, a następnie, sumując rangi różnic dodatnich i ujemnych obliczamy odpowiednią statystykę.