

dr hab. Barbara Więckowska

Katedra i Zakład Informatyki i Statystyki

Zalecenia dotyczące analizy wariancji (ANOVA) dla danych niezależnych
wraz z testami post-hoc.

data przygotowania: 2024-03-16

ANOVA pomaga sprawdzić czy istnieje istotna różnica w pomiarach
między kilkoma kategoriami.

Istnieje kilka klasycznych, parametrycznych analiz typu ANOVA oraz ich nieparametryczne odpowiedniki. Dla danych ciągłych zaleca się wykonać analizy i opisać ich wyniki zgodnie ze schematem zawartym w przykładzie.

Załóżmy, że badacz chce sprawdzić, czy istnieje różnica w średnim tętnie spoczynkowym (HR) u biegaczy z kilku różnych kategorii wiekowych.

Co wiemy o danych?

- (1) Tętno to cecha dająca się uporządkować i można obliczyć o ile jedna osoba ma wyższe tętno niż inna – jest to skala interwałowa a mierzony parametr ma charakter ciągły.
- (2) Kategorie wiekowe można uporządkować, ale nie da się sprawdzić na ich podstawie o ile ktoś jest starszy, jedynie do której kategorii wiekowej należy – jest to skala porządkowa

Jak przygotować dane?

Dane przygotuj w dwóch kolumnach

Wartość α najczęściej wynosi 0.05 i jest to poziom błędu którego nie chcemy przekroczyć prowadząc badanie.

Trendy możemy sprawdzać jeśli cecha, która tworzy kategorie daje się uporządkować i jeśli interesuje nas sprawdzenie czy ze wzrostem tej cechy (u nas wieku) wzrasta/spada druga z badanych cech (w naszym przykładzie tętno spoczynkowe).

Jak opisać wyniki?

- Gdy dane miały rozkład normalny wyniki opisz podając średnią z odchyleniem standardowym (SD – standard deviation) lub błędem (SE – standard error) lub przedziałem ufności np. w zakresie 95% (95% CI – confidence interval)
- Gdy dane nie miały rozkładu normalnego wyniki opisz podając medianę i kwartyle. Mediana to 50. percentyl, wartość pierwszego kwartyła to 25. Percentyl, a trzeciego kwartyła to 75. percentyl. Czasami podaje się wartość najmniejszą (minimum) /największą (maksimum).

W dalszej analizie można utworzyć grupy jednorodnych pomiarów (nieróżniących się w sposób istotny statystycznie) i wskazać je na wykresie lub w tabeli. Jeśli wartość p dla ANOVY jest > 0.05 , czyli nieistotna, nie ma potrzeby analizować wyników testów post-hoc ani na ich podstawie budować grup jednorodnych.

Przykładowy wynik

Wartość $p = 0.0018$ testu ANOVA dla danych niezależnych wskazuje, że kategorie wiekowe różnią się wartością HR w spoczynku. Innymi słowy, kategoria wiekowa jest czynnikiem, który istotnie różnicuje wartość tętna spoczynkowego. Najwyższe wartości osiąga kategoria najmłodsza, najniższe kategoria najstarsza. Jednak wskazane $p = 0.0018$ nie wskazuje między którymi kategoriami istnieje taka różnica! Ta wartość mówi jedynie, że kategorie wiekowe istotnie

różnią się spoczynkową wartością HR. Dopiero post-hoc test, np. NIR Fishera, może wyodrębnić kategorie, które różnią się między sobą – przykład podano w tabeli.

Wartość p	21-30 lat	31-40 lat	41-50 lat	51-60 lat
21-30 lat		0.02881219	0.0360886	0.00013804
31-40 lat	0.02881219		0.53887178	0.04609668
41-50 lat	0.0360886	0.53887178		0.07148519
51-60 lat	0.00013804	0.04609668	0.07148519	

Jak widać, częstość pracy serca w najmłodszej kategorii wiekowej jest istotnie różna od wszystkich osób w pozostałych kategoriach. Podobnie, osoby z najwyższej kategorii wiekowej mają istotnie różną częstość pracy serca od pozostałych osób. Z kolei nie ma istotnych różnic między osobami z kategorii wiekowych 31-40 lat oraz 41-50 lat. Podsumowując można wyniki opisać, zaznaczając że osoby najmłodsze mają istotnie wyższe tętno od wszystkich pozostałych osób, a najstarsze istotnie niższe od pozostałych.

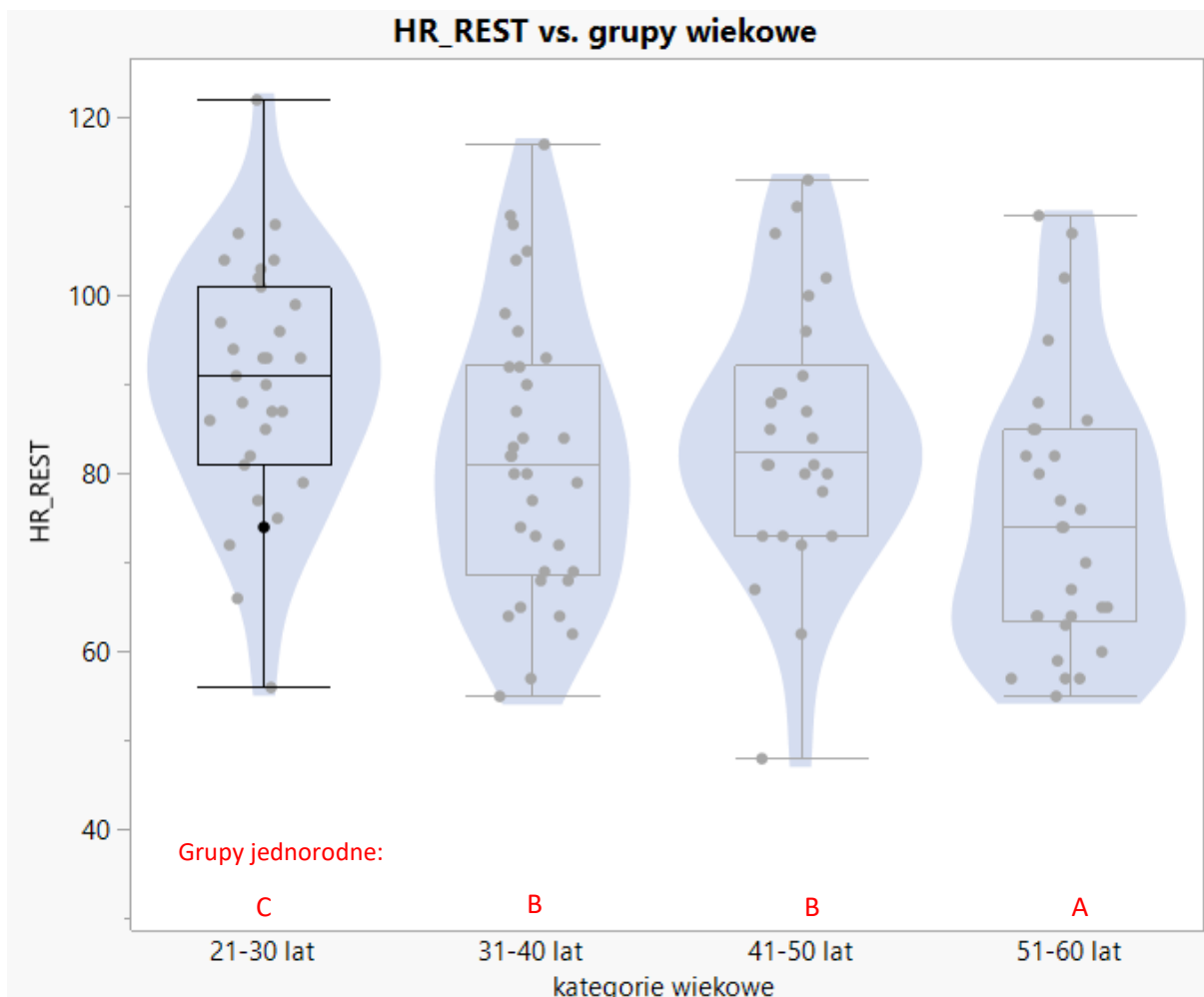
Grupy jednorodne

Zdarza się, że porównuje się więcej niż 3 lub 4 kategorie, np. 10 lub więcej różnych. Wówczas w celu uniknięcia analizowania złożonych tabel z dziesiątkami lub setkami wartości p tworzy się wspólne grupy jednorodne. Ponieważ na podstawie testu post-hoc ustaliliśmy, że osoby w najmłodszej kategorii wiekowej różniły się wartością HR od osób w pozostałych kategoriach wiekowych znalazły się one w odrębnej grupie jednorodnej C (o najwyższej średniej wartości HR). Podobnie najstarsi różnili się od wszystkich pozostałych kategorii wiekowych i zostali przydzieleni do odrębnej grupy jednorodnej A (o najniższej średniej wartości HR). Pośrednie kategorie wiekowe tzn. 31-40 lat oraz 41-50 lat nie różniły się między sobą ale różniły się od kategorii najmłodszej i najstarszej, dlatego znalazły się we wspólnej grupie jednorodnej B (o przeciętnej wartości HR). Czasami grupy jednorodne zazębiają się wówczas kategorie wieku mogą mieć przypisaną więcej niż jedną literę w opisie. Prezentacja wyników zamiast wartości p dla poszczególnych porównań między kategoriami może zostać uproszczona to podania liter

oznaczających grupy jednorodne. Tylko różne litery wskazują, że dana grupa (która może mieć więcej niż jedną kategorię) różni się od pozostałych grup z inną literą. Wszystkie kategorie z tą samą literą nie różnią się między sobą.

Kategorie wiekowe	21-30 lat	31-40 lat	41-50 lat	51-60 lat
HR REST Średnia ± SD	90.1 [85 - 95.1]	81.8 [76.3 - 87.3]	84.2 [78.2 - 90.3]	74.8 [69 - 80.6]
Grupy jednorodne	C	B	B	A

Na wykresie przedstawiono rozkład spoczynkowej częstości pracy serca w poszczególnych kategoriach wieku z przypisanymi etykietami dla grup jednorodnych.



Kiedy badać istnienie trendu?

Jeśli kategorie (oś X na wykresie) są cechą porządkową i uzasadnione jest oczekiwać (a priori) pojawienia się trendu zgodnego ze wzrostem tej cechy, wówczas możemy badać trendy tworzone przez takie kategorie. Trendy mogą mieć charakter liniowy lub inny.

W przedstawionym przykładzie kategorie wiekowe stanowią cechę porządkową. Możemy oczekiwać, a nawet widzimy na wykresie, że wraz ze wzrostem wieku (zmianą kategorii wiekowej na kolejną) uzyskujemy niższe wartości HR. Taka obserwacja pozwala na rezygnację z ANOVAY wraz z testami post-hoc i zastosowanie w zamian prostszej hipotezy: hipotezy o trendzie. W podanym przykładzie ustawiając kontrast liniowy Fishera uzyskaliśmy wartość $p=0.0003$.

Możemy więc potwierdzić istnienie malejącego trendu liniowego i napisać w sposób prosty, że osoby należące do kolejnych kategorii wiekowych mają coraz niższą wartość HR w spoczynku.