Jednoczynnikowa analiza wariancji

Zmienna zależna – ilościowa, numeryczna

Zmienna niezależna – grupująca (dzieli próbę na więcej niż dwie grupy), nominalna – zmienną wyrażoną tekstem należy w SPSS przerekodować Przekształcenia >>> Automatyczne rekodowanie

Hipoteza zerowa

Średnie we wszystkich grupach są równe

H_0: $\mu_1 = \mu_2 = ... = \mu_k$

Hipoteza alternatywna

Średnie pomiędzy jakimiś grupami różnią się istotnie

Dla trzech grup (k=3)

H_1 : $\mu_1 \neq \mu_2$ lub $\mu_2 \neq \mu_3$ lub $\mu_1 \neq \mu_3$

ANOVA –założenia

- Rozkład wyników w każdej grupie ma **rozkład zbliżony do normalnego**
- Wariancje w grupach są jednorodne (**homogeniczność wariancji**)
- **Równoliczność obserwacji** w podgrupach

Minimalna wielkość próby

- Minimum absolutne, aby program wykonał test: N-k > 0
- Minimum zalecane w naukach przyrodniczych **5-7** w podgrupie
- Najlepiej, jak najwięcej, co najmniej **30** w podgrupie

Weryfikacja założeń

- Sprawdzenie zgodności z rozkładem normalnym i jednorodność wariancji Analiza >>> Opis statystyczny >>> Eksploracja
- Wykresy normalności z testami + histogramy + rozrzut-poziom z testem Levene'a (Nie przekształcone, w wynikach patrzymy na *Bazując na średniej*)
  Lub SPSS: Analiza >>> porównywanie średnich >>> Jednoczynnikowa ANOVA >>> opcje >>> jednorodność wariancji
- Równoliczność grup – gdy obserwowane są różnice w liczebnościach - test zgodności z rozkładem teoretycznym Chi-kwadrat Analiza >>> Testy nieparametryczne >>> Testy tradycyjne >>> Chi-kwadrat Weryfikujemy czy rozkład obserwacji w poszczególnych grupach różni się istotnie od założonego (1:1:1:…:1) . H_0: Rozkład jest zgodny z rozkładem teoretycznym.

Przy równolicznych grupach ANOVA jest odporna na brak równości wariancji i normalności rozkładu.

Gdy mamy nierównoliczne grupy i brak równości wariancji:

- Statystyka F będzie przeszacowana, jeśli wariancja jest większa w grupie mniej licznej – większe ryzyko fałszywego odrzucenia H0
- Statystyka F będzie niedoszacowana, jeśli wariancja jest większa w grupie liczniejszej – większe ryzyko nieodrzucenia fałszywej H0

Brak normalności rozkładu w niewielkim stopniu wpływa na wyniki ANOVY. Duża liczba obserwacji (>30 w każdej podgrupie) zapewnia działanie centralnego twierdzenia granicznego gdy nie ma rozkładu normalnego – szczególnie przy rozkładach skośnych

Brak jednorodności wariancji

- Sprawdzamy czy nie ma obserwacji odstających w grupach
- Dokonujemy transformacji zmiennej zależnej
- Wykluczamy grupę, w której wariancja różni się od pozostałych
- Stosujemy test nieparametryczny
- **Używamy korekty Welch'a lub Brown-Forsythe'a – poprawki do statystyki F uwzględniającej nierówne wariancji. Stosowane wymiennie, jednak test Welch'a jest bardziej konserwatywny i ma większą moc.**

SPSS: Analiza >>> porównywanie średnich >>> Jednoczynnikowa ANOVA >>> opcje >>> Welch

Gdy w analizie wariancji odrzucimy H_0 >>> Testy post-hoc

<u>Równoliczne grupy, równe wariancje:</u>

- **Tukey**, nazwy alternatywne:
    - Po polsku: UIR – test Uczciwie Istotnych Różnic
    - Po angielsku: HSD – Honestly Significant Differences
- **Test pierwszego wyboru przy spełnionych założeniach**. Podobny do poprawki Bonferroniego, ale ma większą moc przy dużej liczbie porównań.
- Inne popularne testy:
    - Dunnett, Duncan

<u>Nierównoliczne grupy, równe wariancje:</u>

- **LSD Fishera**, nazwy alternatywne:
    - Po polsku: NIR – Najmniejsza Istotna Różnica
    - Po polsku: LSD – Least Squared Difference
    - Polega na wykonaniu k(k-1)/2 testów t-studenta bez poprawki
    - Bardzo liberalny test, stosowany pomocniczo
- **Bonferroni**
    - Polega na wykonaniu k(k-1)/2 testów t-studenta z poprawką na liczbę porównań
    - Poprawka polega na podzieleniu błędu pierwszego rodzaju (α) na liczbę porównań

    - Nie zakłada się równoliczności grup
    - Konserwatywny, większa moc przy małej liczbie porównań
- **Sheffe** – jak wyżej, ale dzieli α na 2 * k(k-1)/2. Najbardziej konserwatywny.

Nierównoliczne grupy, nierówne wariancje :

- C Dunnetta
    - Porównania stosowany przy niespełnionych założeniach dla równości wariancji
- T2 Tamhane'a
    - Konserwatywny test stosowany przy niespełnionych założeniach dla równości wariancji

W testach post-hoc >>> istotność <0,05 odrzucamy H_0, są istotne różnice między grupami

Hipotezy kierunkowe – Kontrasty

- Umożliwiają planowane porównania przed przystąpieniem do analizy, w przeciwieństwie do testów post hoc, które wykonujemy dopiero po analizie
- Kontrasty umożliwiają testowanie tylko wybranych porównań pomiędzy grupami, a nie wszystkich jak ma to miejsce przy analizie wariancji
- Testy kontrastów polegają na przypisaniu wag każdej grupie

Przyznawanie wag:

- Wagi są wartościami dodatnimi i ujemnymi, które wskazują które grupy będziemy porównywać
- Grupy, które chcemy porównać powinny mieć przypisaną taką samą wartość wag, ale o przeciwnym znaku np. -1 i +1
    - wartość nie jest ważna, ale dla uproszczenia przypisujemy liczby całkowite
    - Grupie, którą wyłączamy z porównań przypisujemy 0
- Suma wag musi być równa 0
    - Porównując grupę 1 z grupą 2 i 3, przypisujemy np. -2 grupie pierwszej oraz +1 grupie 2 i grupie 3.

Przykładowe wagi:

- Porównanie metody CF AVG z metodami GOR i PHD

Metoda	CF AVG	GOR	PHD
Waga	-2	+1	+1

- Przykładowe wagi dla hipotezy kierunkowej porównującej metodę CF AVG z metodą GOR, z pominięciem PHD

Metoda	CF AVG	GOR	PHD
Waga	0	-1	+1

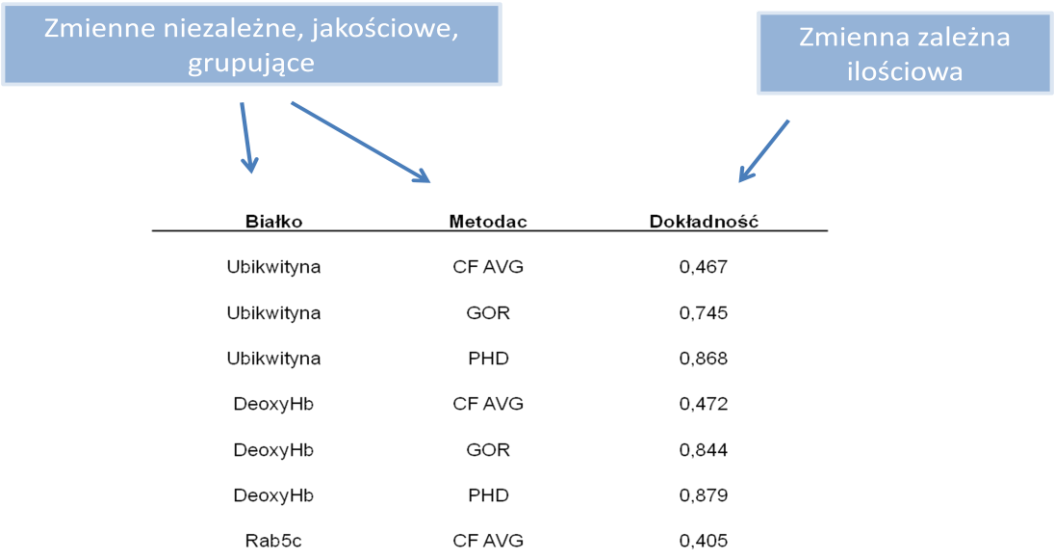
- Ważne – przy porównywaniu dwóch grup średnia grupy z negatywną wagą jest odejmowana od średniej grupy z wagą pozytywną

H_0: Brak istotnych różnic pomiędzy grupami

Dwuczynnikowa analiza wariancji

Jedna zmienna zależna – ilościowa, numeryczna

Dwie (lub więcej) zmiennych niezależnych (grupujących)



Przy układzie doświadczenia 2X3 (dwie zmienne niezależne, jedna dzieli próbę na dwie a druga na trzy grupy) uzyskujemy sześć podgrup. Normalność rozkładu, równoliczność, jednorodność wariancji i wymagania dotyczące minimalnej liczebności dotyczy każdej podgrupy.

Weryfikacja założeń w każdej podgrupie – należy najpierw narzucić podział na grupy według jednej zmiennej (wygodniej gdy jest to ta z jak najmniejszą liczba grup)

Dane >>> Podziel na podzbiory >>> Przedstaw wyniki w podziale na grupy

Przy podzielonej próbie wykonujemy Analiza >>> Opis statystyczny >>> Eksploracja >>> Wprowadzamy zmienna zależną i tą zmienną niezależną (jako czynnik) według której **nie** dzieliliśmy próby.

Weryfikujemy zgodność z rozkładem normalnym.

Równoliczność grup – tak jak przy jednoczynnikowej + wyżej opisany podział danych

Jednorodność wariancji – pomiędzy wszystkimi podgrupami Analiza >>> Ogólny model liniowy >>> Jednej zmiennej >>> Opcje >>> Testowanie jednorodności

Hipoteza zerowa i alternatywna – jak w jednoczynnikowej

<u>Efekty główne i efekty interakcyjne</u>

Efekt główny:

- Różnice między średnimi zmiennej zależnej, wyróżnionymi ze względu na poziomy jednej zmiennej niezależnej
- **Interpretowany jak przy jednoczynnikowej analizie wariancji**
- Istotność statystyczna oznacza że co najmniej jedna średnia różni się od pozostałych

Efekt interakcji:

- Wzór różnic pomiędzy średnimi zmiennej zależnej na różnych poziomach czynnika pierwszego jest różny na poszczególnych poziomach czynnika drugiego
- współdziałanie czynników ze sobą przynosi efekt odmienny niż każdego z osobna
- Wpływ czynników nie musi być istotny statystycznie żeby istotny był wpływ interakcji

Moderator - zmienna niezależna która wprowadzona do analizy zmienia istotność i kierunek różnic zmiennej zależnej, ze względu na drugą zmienną niezależną

- To która zmienna jest moderatorem ustalamy na podstawie wcześniejszej wiedzy, celu analizy, itp. tak aby moderator i jego działanie było interpretowalne w świetle teorii w ramach której formułowane są hipotezy

Istotny efekt interakcji zmienia interpretacje

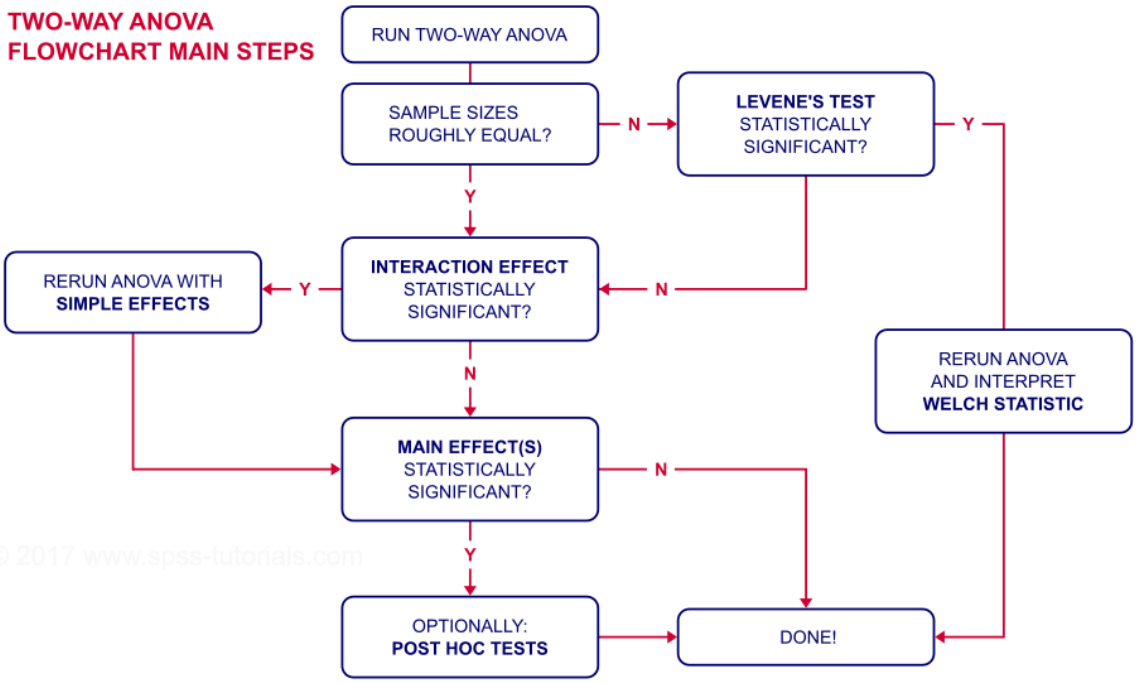
Główne efekty proste

- Liczenie jednoczynnikowych analiz wariancji dla jednego czynnika oddzielnie dla każdego poziomu drugiego czynnika
    - Obliczamy efekt główny jednej zmiennej utrzymując drugą zmienną na stałym poziomie
- Przy istotnej interakcji istotne może być porównanie tylko na wybranych poziomach zmiennej niezależnej
- Aby sprawdzić pomiędzy którymi grupami są różnice oszacujemy **efekty proste** (ang. simple effects**)** = porównania parami

Efekty proste

- To samo co porównania post-hoc w jednoczynnikowej analizie wariancji – te same testy statystyczne np. Bonferroni, Tukey
- Wykonujemy tylko, jeżeli zaobserwowaliśmy istotny prosty efekt główny na danym poziomie zmiennej niezależnej
- Liczba poziomów czynnika który będziemy porównywać wynosi co najmniej 3

Analiza >>> Ogólny model liniowy >>> jednej zmiennej >>> Opcje: Test jednorodności wariancji + Ocena wielkości efekty >>> Wykresy: wrzucamy zmienne niezależne >>> Zapisz: standaryzowane reszty

<u>Wybór i interpretacja testów Post-hoc – jak przy jednoczynnikowej</u>

Dwuczynnikowa ANOVA schemat