Ekonometria

Robert Pietrzykowski

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

WNE Zarządzanie Studia stacjonarne 2013 - 2014

Zadanie problemowe

W pewnej firmie zajmującą się sprzedażą gier planszowych chciano stwierdzić czy metoda sprzedaży ma wpływ na wielkość sprzedaży [w szt]. Firma prowadzi trzy metody sprzedaży: A - realizowaną w supermarketach, B - realizowaną przez internet, C - realizowaną w sieci sklepów zajmującą się tylko sprzedażą gier planszowych. Poza tym firma chce stwierdzić która metoda sprzedaży pozwala na uzyskanie największej sprzedaży gier plaszowych.

Wprowadzenie

POPULACJA: dni sprzedaży gier planszowych w pewnym okresie czasu. Czynnik różnicujący sprzedaż to różny sposób sprzedaży. Mamy trzy poziomy czynnika.
CECHA ZALEŻNA: wielkość sprzedaży [szt], która jest różnicowana ze względu na sposób sprzedaży. Możemy jednak uznać, że jest to cecha ilościowa, ciągła.
CECHA NIEZALEŻNA: rodzaj sprzedaży, cecha opisowa, jakościowa.

Zagadnienie badania wpływu cechy o charakterze jakościowym na cechę ciągłą.

Model analizy wariancji

$$Y_{ij} = \mu_i + \epsilon_{ij}$$

$$Y_{ij} = \mu + a_i + \epsilon_{ij}$$

$\epsilon_{ij} \backsim N(0, \sigma^2)$
a_i - efekt i-tego poziomu czynnika

ZAŁOŻENIA:
1) $Y_i \backsim N(\mu_i, \sigma_i^2) \quad i = 1, \ldots, k$
2) $Y_1, Y_2, \ldots, Y_k$ – są niezależne
3) $\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2$

Hipoteza

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

$$H_0 : a_1 = a_2 = \ldots = a_k$$

$$H_0 : \sum_{i=1}^{k} a_i^2 = 0$$

Źródło zmienności	Stopnie swobody	Sumy kwadratów	Średnie kwadraty	Femp
Czynnik	k-1	varA	$S_a^2 = \frac{varA}{k-1}$	$\frac{S_a^2}{S_e^2}$
Błąd losowy	N-k	varE	$S_e^2 = \frac{varE}{N-k}$	
Ogółem	N-1	varT		

Hipoteza

Źródło zmienności	Stopnie swobody	Sumy kwadratów	Średnie kwadraty	Femp
Czynnik	k-1	varA	$S_a^2 = \frac{varA}{k-1}$	$\frac{S_a^2}{S_e^2}$
Błąd losowy	N-k	varE	$S_e^2 = \frac{varE}{N-k}$	
Ogółem	N-1	varT		

$$varT = varA + varE$$

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij} - \bar{Y})^2 = \sum_{i=1}^{k} n_i(\bar{Y}_i - \bar{Y})^2 + \sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij} - \bar{Y}_i)^2$$

Porównania szczegółowe

Grupy jednorodne, homogeniczne – podzbiory średnich, które można uznać za takie same.
Procedury porównań wielokrotnych – postępowanie statystyczne mające na celu uzyskanie podziału zbioru średnich na grupy jednorodne.
Procedury: Tukeya, Scheffego, Bonfferoniego, Duncana i inne.

Porównania szczegółowe

Algorytm działania:

1. wyznaczamy pewną miarę która będzie służyła do określenia czy porównywane średnie różnią się między sobą. Określamy ją jako NIR (Najmniejsza Istotna Różnica).
2. Porównujemy wszystkie pary średnich
3. Jeżeli $|\bar{Y}_i - \bar{Y}_j| < NIR$ to uznajemy, że badane średnie są takie same $(\mu_i = \mu_j)$ czyli możemy uznać, że tworzą grupę jednorodną.
4. Jeżeli $|\bar{Y}_i - \bar{Y}_j| > NIR$ to uznajemy, że badane średnie nie są takie same $(\mu_i\mu_j)$ czyli nie możemy uznać, że tworzą grupę jednorodną.

Sprawdzenie założeń

Badanie równości wariancji:

$$H_0 : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2$$

Cecha Y_i ma rozkład normalny $N(\mu_i, \sigma_i^2)$ o nieznanej średniej μ_i i wariancji σ_i^2
Test statystyczny: Hartleya, Cochrana, Bartletta
Badanie normalności rozkładu cech:

$$H_0 : Y_i \text{ ma rozkład normalny}$$

Test statystyczny: Shapiro-Wilka, Kołomogorowa, Lilieforsa, χ^2 zgodności