



Международно Висше Бизнес Училище
International Business School

Международно висше бизнес училище Ботевград

Дистанционно обучение

Учебна дисциплина: Бизнес статистика

КАЗУС

Преподавател: доц. д-р Недка Гатева

Разработил:

Фак.№.....

Специалност:.....

2018г.

1. Каква е познавателната цел на дисперсионния анализ?

Дисперсионният анализ се използва за проверка на хипотези. Той намира приложение тогава, когато се резултатната променлива може да се изрази в интервална скала, а факторите могат да се изразят в ординална, номинална или друга. Съвременните приложения на дисперсионния анализ обхващат широк спектър от проблеми в икономиката, социологията, биологията и технологиите; прилага се за откриване на систематични разлики между резултатите от преките измервания, извършвани при специфични условия.

2. Как се дефинират нулевата и алтернативната хипотези при еднофакторния анализ?

$H_0 : \bar{x}_1 = \bar{x}_2 = \bar{x}_3 = \dots = \bar{x}_k$ т.е.
ако съществуват разлики между тях то те са случайни;

$H_1 : \bar{x}_1 \neq \bar{x}_2 \neq \bar{x}_3 \neq \dots \neq \bar{x}_k$ т.е. разликата не е случайна, а се дължи на систематичното действие на фактора, който се изследва.

3. Какъв е смисълът на разлагането на общата девиация на съставни компоненти?

За представяне на разлагането се използва следната формула: $SS_i = SS_m + SS_b$, където:

SS_m - междугруповата оценка на дисперсията,

SS_b - вътрешногруповата оценка на дисперсията.

Разлагането е възможно по силата на адитивното свойство. Като то се прилага се дефиниране на влиянието на крайно действащите и останалите фактори.

По представеният начин биха могли да се сравнят две независими и самостоятелни оценки на общата девиация.

4. Какво се разбира под независими оценки на дисперсията?

Разбира се междугрупова и вътрешногрупова девиация, които са независими оценки на общата вариация на зависимата променлива.

5. Защо са различни степените на свобода при двете независими оценки на дисперсията?

Разликата се дължи на начина на тяхното установяване: $\varphi_1 = k - 1$ – За междугрупова дисперсия ; $\varphi_2 = n - k$ – За вътрешногруповата дисперсия.

6. Защо емпиричната характеристика на F трябва да се сравнява с теоретичната?

По същество това е сравнение на емпирична с теоретичната характеристика. Целта е вземане на решение за приемане и отхвърляне на хипотеза/и.

7. Как се намира теоретичната (критичната) стойност на характеристиката?

За целта има разработени таблици при дадени α ; φ_1 ; φ_2 .

8. Как се формулира заключението при еднофакторния дисперсионен анализ?

Ако $F_{emp.} \geq F_T$, то нулевата хипотеза (H_0) се отхвърля и се приема алтернативната (H_1).

Ако $F_{emp.} \leq F_T$, то алтернативната хипотеза (H_1) се отхвърля и се приема нулевата (H_0). Приема се, че липсва статистически значима разлика при съответното равнище на значимост $\alpha = 0,01$; $0,05$ и др..

Представете решението на следната задача

В таблицата по-долу са дадени данни за оборота на запасите от независими случайни извадки за три групи търговски фирми с еднакъв предмет на дейност. (Данните са примерни). Необходимо е да се провери чрез дисперсионен анализ има ли основание да се твърди, че мащабите на фирмите оказват влияние върху скоростта (времето) на оборота на запасите им, т.е. дали разликите между средните обороти за отделните извадки са статистически значими (неслучайни). Заключението да се направи при равнище на значимост $\alpha = 0,05$. Представете подробно всички стъпки на анализа.

Групи фирми по размер на месечния оборот	Брой на фирмите в извадките (n_i)	Продължителност на оборота на отделните фирми в дни (x)	Средна продължителност на оборота (средни на извадките)
--	---------------------------------------	---	---

Малки	12	15,14,11,18,12,16,12,14,15,13,12,18	14,2
Средни	10	14,11,12,14,12,14,13,12,11,12	12,5
Големи	8	16,10,11,11,12,11,10,14	11,9
Общо:	30		(13,0)

Решение:

Може да се използва дисперсионният анализ:

1/ Дефиниране на хипотезите – нулева хипотеза и алтернативната хипотеза:

Нулевата хипотеза => H₀: Дали разликите между средните обороти за отделните извадки са статистически значими (неслучайни).

Алтернативната хипотеза: H₁: Дали разликите между средните обороти за отделните извадки са статистически незначими.

2/ Равнище на значимост $\alpha = 0,05$.

3/ Избор на метод

Това ще бъде еднофакторния дисперсионен анализ, защото се изследва корелационна връзка между две явления

$$F_{em} = \frac{\sigma_m^2}{\sigma_b^2},$$

където:

$$\sigma_m^2 = \frac{\sum(x_i - \bar{x})^2 n_i}{k-1}$$

$$\sigma_b^2 = \frac{\sum(x - \bar{x}_i)^2}{n-k}$$

4/. Изчисляване на емпиричната стойност:

Групи фирми по размер на месечния оборот	Брой на фирмите в извадките (n _i)	Продължителност на оборота на отделните фирми в дни (x)	Σ(x _j i)	средни на извадките \bar{x}
--	---	---	---------------------	-------------------------------

Малки	12	15,14,11,18,12,16,12,14,15,13,12,18	170	14,2
Средни	10	14,11,12,14,12,14,13,12,11,12	125	12,5
Големи	8	16,10,11,11,12,11,10,14	95	11,9
Общо:	30		390	(13,0)

Оценка на средните значения на x_{ji} за трите типа фирми (малки; средни и големи) според размера на техният месечен оборот с помощта на средна аритметична величина:

$$\bar{x}_1 = \frac{\sum_{j=1}^{n_i} x_{ij}}{n_i}$$

$$\text{За малките фирми: } \bar{x}_1 = \frac{\sum_{j=1}^{n_i} x_{ij}}{n_i} = \frac{170}{12} = 14,2$$

$$\text{За средните фирми: } \bar{x}_2 = \frac{\sum_{j=1}^{n_i} x_{ij}}{n_i} = \frac{125}{10} = 12,5$$

$$\text{За големите фирми: } \bar{x}_3 = \frac{\sum_{j=1}^{n_i} x_{ij}}{n_i} = \frac{95}{8} = 11,9$$

След това се изчислява обща средна величина по формулата:

$$\bar{x}_0 = \frac{\sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij}}{n} = \frac{390}{30} = 13,0$$

Прави се оценка на междугруповата девиация

$$\sigma_m^2 = \frac{\sum (x_i - \bar{x})^2 n_i}{k-1}$$

$$\sum (\bar{x}_i - \bar{x})^2 n_i = (\bar{x}_1 - \bar{x})^2 \cdot n_1 + (\bar{x}_2 - \bar{x})^2 \cdot n_2 + (\bar{x}_3 - \bar{x})^2 \cdot n_3 = (14,2 - 13)^2 \cdot 12 + (12,5 - 13)^2 \cdot 10 + (11,9 - 13)^2 \cdot 8 = 17,29 + 2,5 + 9,68 = 29,47$$

$$\sigma_m^2 = \frac{\sum (x_i - \bar{x})^2 n_i}{k-1} = \frac{29,47}{3-1} = \frac{29,47}{2} = 14,74$$

Прави се оценка на вътрешногруповата девиация:

$$\sigma_b^2 = \frac{\sum(x - \bar{x}_i)^2}{n - k}$$

$$\sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 = (15-14,2)^2 + (14-14,2)^2 + (11-14,2)^2 + (18-14,2)^2 + (12-14,2)^2 + (16-14,2)^2 + (12-14,2)^2 + (14-14,2)^2 + (15-14,2)^2 + (13-14,2)^2 + (12-14,2)^2 + (18-14,2)^2 + (14-12,5)^2 + (11-12,5)^2 + (12-12,5)^2 + (14-12,5)^2 + (12-12,5)^2 + (14-12,5)^2 + (13-12,5)^2 + (12-12,5)^2 + (11-12,5)^2 + (12-12,5)^2 + (16-11,9)^2 + (10-11,9)^2 + (11-11,9)^2 + (11-11,9)^2 + (12-11,9)^2 + (11-11,9)^2 + (10-11,9)^2 + (14-11,9)^2 = 103,06$$

$$\sigma_b^2 = \frac{\sum(x - \bar{x}_i)^2}{n - k} = \frac{103,06}{30 - 3} = \frac{103,06}{27} = 3,82$$

$$F = \frac{\sigma_m^2}{\sigma_e^2} = \frac{14,74}{3,82} = 3,86$$

5/. Определяне степените на свобода:

$$\varphi_1 = k - 1 = 3 - 1 = 2$$

$$\varphi_2 = n - k = 30 - 3 = 27$$

6/. Определяне на критичната област –двустранна критична област, защото не е установена посоката на проблема.

7/. Оценява се стойността на теоретичната характеристика на хипотезата **Ft.**

То се намира по зададен риск за грешка $\alpha = 0,05$, $\varphi_1 = 2$ и $\varphi_2 = 27$ от таблица за критичните стойности на F-разпределението.

Таблица с критичните стойности на F-разпределението, $\alpha = 0,05$

$\varphi_1 \backslash \varphi_2$	1	2	3	4	5
...
27	4,21	3,35	2,96	2,73	2,57

$$F_T=3,35$$

$$F_{\text{emp.}} \geq F_T,$$

$$3,86 > 3,35$$

8/. Вземане на решение за приемане или отхвърляне на съответната хипотеза

Както посочихме по-горе, ако $F_{\text{emp.}} \geq F_T$, то се приема алтернативната хипотеза и се отхвърля нулевата, т.е. разликите между средните обороти за отделните извадки са статистически незначими.