УДК 338.534: 338.516.44: 338.516.54

O.T. Devinyak[1,2], T.I. Khudan[1], M.I. Kusiy[3]

CLUSTER ANALYSIS AS A RATIONAL METHOD FOR PHARMACEUTICAL MARKET SEGMENTATION ON PRICE BASIS

State University «Uzhgorod National University»[1],
Uzhgorod, Ukraine
Danylo Halytsky Lviv National Medical University[2],
Lviv, Ukraine
Lviv State University of Life Safety[3],
Lviv, Ukraine

e-mail: *o.devinyak@gmail.com*

Summary: Three possible methods for post-hoc market segmentation on price basis are compared in the paper. They are k-means clustering, hierarchical clustering and exhaustive search. The results have showed that exhaustive search is preferred for price-based segmentation in inventory management where fast and reliable method is needed and a comprehensive picture of prices distribution is not required. However, hierarchical clustering is more valuable for product positioning due to its inner capability to represent prices distribution via the specific plot – dendrogram. Both strong and weak sides of the methods are discussed and practical recommendations are given.

Keywords: market segmentation, cluster analysis, pharmaceutical products segmentation, pricing, k-means, hierarchical clustering.

Introduction. Market segmentation is a broadly used marketing strategy that involves the dividing of market elements (products or customers) into corresponding groups. The product segmentation is required during product positioning. Besides different qualitative and technical characteristics of a product, the price can also be used as a criterion for product segmentation. Price-based product segmentation results in the dividing of products into several (usually three) different price segments or niches. Commonly they are low price segment, mid-segment and premium segment. Each of these segments has its own demand that should be satisfied with proper supply. Thus price-based segmentation is useful not only for product positioning and pricing by manufacturing companies, but for inventory management and forecasting of retail companies. Actually, when products are easy interchangeable, a demand concerns not specific product names but any product from a single specific price segment. Since generic drugs take the leading position on the pharmaceutical market of Ukraine, there is a vast amount of interchangeable OTC-drugs proposed in a pharmacy. Additionally, according to paragraph 78-1 of Ukrainian law № 2801-XII (19.11.1992) «Basic Law of Ukraine on Health Care» a pharmacist cannot refuse to give or give wrong information about the presence of drugs with the same active substance, especially with lower price. Thus price-based segmentation is a crucial part of pharmaceutical companies' activity.

The price-based segmentation of commercial products is usually carried out as a part of a specific market segment analysis and has different aims. *T. Bashuk* used such segmentation in the pricing process of a new book[2]. The price-based segmentation of drugs into three groups has been carried out by *V. Gonchar* in order to show the overall price dynamics of pharmaceutical market during 2008-2010[3]. *O. Gudzenko* reported the distribution of antiallergic medicines over three price segments pointing that high proportion of low-price segment is a positive social-economic characteristics of antiallergic agents market[4]. *Z. Mnushko* and N. *Sotnikova* show the relationship between price segments and a country of origin in the Ukrainan market of dietary supplements[8]. *L. Kulyba* and *T. Dibrova* discuss the dependence of retail margin on the price segment of pharmaceutical market[5]. It is worth to note, that the choose of

price intervals that were used to create segments seems to be manual and arbitrary in all pharmacy-related papers[3-5,8]. For example, the price intervals to divide the whole pharmaceutical market were taken to be 0-10 UAH, 10-50 UAH and higher than 50 UAH[5]. Almost at the same time price intervals were fixed at 0-15, 15-35 and higher than 35 UAH for dietary supplements market[8]; at 0-20, 20-40 and higher than 40 conventional units (actually, they are also UAH) for antiallergic medicines[4]; and price intervals were not disclosed by *V. Gonchar*[3]. Such assigning of price intervals can be called *ad-hoc* or *apriori* since methodologically it has no close relationship with price data and can be carried out before data collection. On the other hand, the choose of intervals can be based on the collected data so they obtained segments would fit prices more accurately – that is *post-hoc* approach. For example, *T. Bashuk* defined price intervals after the inspection of price distribution plot[2]. Although, that method is subjective as well. The methods of segmentation are objective when the choice of intervals depends only on data. The problem of partitioning objects into maximally different groups is covered by cluster analysis. The idea of market segmentation via cluster analysis is not novel. However, literature sources do not cover price as segmentation criterion for clustering[1,6,7,10]. So the efficiency and special features of market segmentation via cluster analysis with price criterion are still unknown. Moreover, the results of market segmentation via clustering have been never compared with exhaustive search method.

The aim of research are: 1) to compare the efficiency of three post-hoc methods of market segmentations: k-means clustering, hierarchical clustering and exhaustive search using price as grouping criterion; 2) to describe the peculiarities of price as a criterion for product segmentation and to report the guideline for price-based product segmentation; 3) to present the distribution of antihistamine drugs, hepatoprotectors, cough syrups, sprays and lozenges for treating of throat pain, remedies for acne and vitamins over three price segments as a case product segmentation.

Materials and methods. Retail price data has been collected from a pharmacy of Beregovo town (Transcarpathian region) and weighted average prices have been obtained from website *compendium.com.ua*. These prices have been collected in the December, 2012 and thus may differ from current prices. Totally 147 drug names were included into analysis. The choice of pharmacotherapeutic groups was not arbitrary: these groups consist of interchangeable OTC-drugs that can be proposed to a consumer by pharmacist. The drugs that were not available at the Beregovo pharmacy have been excluded from the analysis. Since there are different dosages and drug intake regimens in each pharmacotherapeutic group, the prices were adjusted to fixed time of therapy. Briefly, the number of studied antihistamine drugs is 31 and price is adjusted for 10 days of therapy. The prices of 15 hepatoprotectors are adjusted for 10-day therapy also. Cough syrups (28) are adjusted for 4 days, sprays for treating throat pain (13) – for 10 days, lozenges for treating of throat pain (23) – for 5 days, vitamins (28) – for 30 days and remedies for acne (9) are not adjusted since they have topical application.

The three methods used to divide drugs into price segments are k-means clustering, hierarchical clustering and exhaustive search. The procedure of clustering has been carried out in *R 3.0.1 software*[9]. K-means belongs to stochastic algorithm that is iteratively searching for optimal solution. So possible drawbacks of k-means method are the requirement the number of clusters as an input parameter and convergence to a local optimum rather than to global. Hierarchical clustering is a determined algorithm that is looking for locally optimal step in each iteration. Such algorithms are called «greedy» and the obtained results often are non-optimal (though still close to the optimal solution). The Ward's link function has been used as an objective for minimization during hierarchical clustering[11]. Like k-means algorithm, the Ward's method minimizes the sum of squared Euclidean distances within clusters. The exhaustive search tests all possible variants of partitioning into defined number of clusters. We have written the function for exhaustive search using R as programming language and sum of squared Euclidean distances within clusters as a criterion to choose the best combination. This allows us to compare all three methods with the same ruler. Actually, exhaustive search will always give the best possible solution. The major drawback of this approach is merely its computational time.

Results and discussion. First of all, we would like to note that since market segmentation has customer attraction as its final aim, the created segments should be consisted with customer view on the market. Thus using price as segmentation criterion has its own specificity. Actually, the correlation between customers' point of view on prices and the prices themselves is not linear. For example, in a customer subjective opinion the difference between 2

UAH and 20 UAH is much larger than between 100 UAH and 120 UAH, while arithmetically the relationship is inverse. That is the one of the fundamental features of human mind and it is described by the psychophysical law of *Weber and Fechner*. *Weber-Fechner* law states that subjective sensation is proportional to the logarithm of the stimulus intensity, and this law covers various fields of human perception: vision, sound perception, weight perception and numerical cognition as well. Hence the segmentation based on price logarithms is much more close to consumer's point of view than the segmentation using raw price values. As far as we know, this is the first proposal to make logarithmic transform on prices in order to perform price-based market segmentation.

This way, the price logarithms splitted into pharmacotherapeutic groups were subjected to cluster analysis via the three different methods. It has been observed that k-means algorithm finds the best solution in 86% of cases. That is why at least two independent runs of k-means with the selection of best one are needed to obtain best solution with small risk to make a wrong decision. The accuracy of hierarchical cluster analysis is even worse with 57% of best solution cases. However, the hierarchical analysis gives a precious information about mutual distribution of prices by means of special plot – dendrogram. For example, looking at the dendrogram of antihistamine drugs (fig. 1.) it is apparent that there are exactly three well-defined price segments.
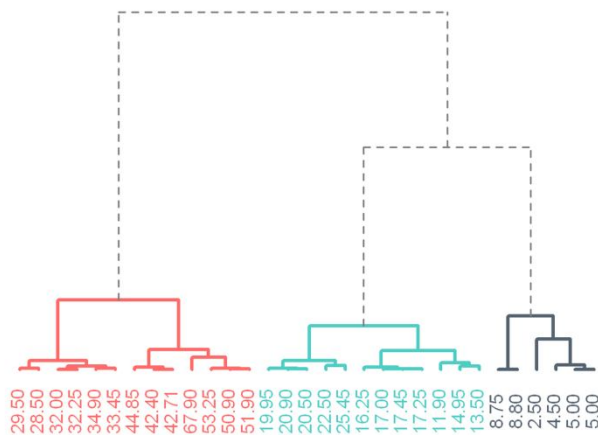

Fig. 1   **Dendrogram of antihistamine drugs**

But in the case of lozenges for treating of throat pain the dendrogram (fig. 2.) shows that three is not the best number of clusters. There is «septefril» in low-price segment with extremely low price (as being compared with its neighbors). Thus the distribution of lozenges prices is better represented as «three segments + septefril».
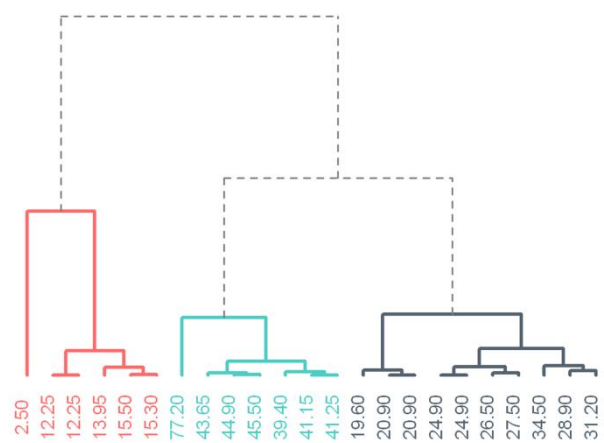

Fig. 2   **Dendrogram of lozenges for treating of throat pain**

Such thorough analysis of prices distribution is useful when the market positioning of new drug is carried out. While setting the price for new drug, the closest competitors can be identified this way. This information may be helpful in the development of drug promotion strategy.
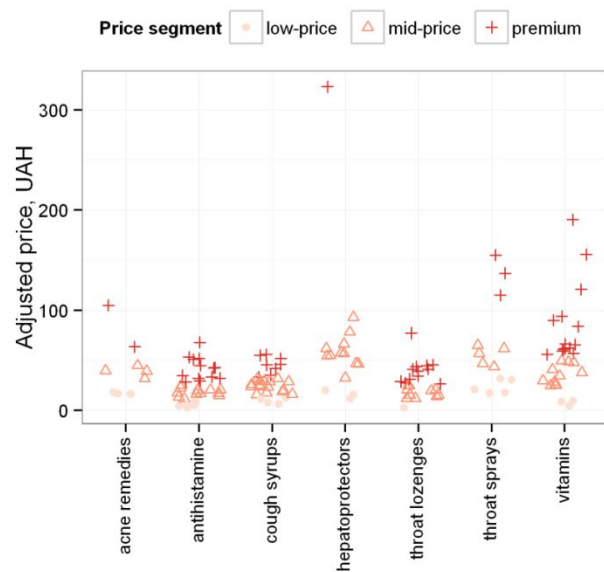

Fig. 3   **The distribution of prices for different pharmacotherapeutic groups**

Exhaustive search always results in the best solution. That is the way how we found out what segmentation is the optimal one and, consequently, what is its frequency in k-means or hierarchical clustering results. Though the main limitation of exhaustive search is significant computational time, the application of this method to splitting about 30 objects took less than one second on *Intel Pentium Dual-Core* processor with core clock rate 1.86 GHz. That is why, when extended price information is not required, the exhaustive search is preferred

over k-means and hierarchic clustering. This situation takes place in inventory management. Particularly in the retail sector, the integration of automatic price-based segmentation into usual workflow allows to swap the object of inventory management from various trade names to unified price segments. Once again, this claim is fair for mutually interchangeable drugs.

Finally, the price-based segmentation has been carried out using both Beregovo pharmacy prices and weighted average prices from *compendium.com.ua*. The results have shown that though two sets have some differences, the assignment of each drugs to one of the price segments remains the same. The distribution of studied pharmacotherapeutic groups over three price segments is reported in fig. 3.

The estimated highest adjusted prices in low- and mid-price segments are the following (Beregovo pharmacy data): acne remedies – 17.9 and 44.9; antihistamine agents – 8.8 and 25.45; cough syrups – 12.5 and 31.9; hepatoprotectors – 19.9 and 92.9; throat lozenges – 2.5 and 24.9; throat sprays – 31.4 and 65; vitamins – 9.7 and 49.5;

The listed highest adjusted prices are boundaries between different price segments. Here we can see, that one set of boundaries should not be used for all groups. Moreover, these boundaries move with time since drug prices are constantly changing. That is why the given boundaries should be treated as a case report. While management decisions should be based on the results of up-to-date segmentation with current prices. To reach that goal the price-based segmentation should be integrated into accounting and managerial software. Fortunately, the exhaustive search is a rather simple procedure, whereas free implementations of k-means and hierarchical clustering are available in most of the programming languages. Being used for inventory management, price-based clustering can also be combined with ABC analysis, strengthening the last in such way that despite of products' importance all price segments will be supplied enough (indeed, excessive deleting of C-items can lead to price gaps in product offers).

Conclusions:

Price-based segmentation is a valuable and informative marketing procedure. Among the three studied methods, the exhaustive search has absolute accuracy and takes less than one second to proceed. Thus it is preferred for price-based segmentation in inventory management where fast and reliable method is needed and a comprehensive picture of prices distribution is not required. However, hierarchical clustering is more valuable for product positioning due to its inner capability to represent prices distribution via the specific plot – dendrogram. In the last case the low accuracy of given solution is overshadowed by information gain. In both cas-es the price should be logarithm transformed prior to clustering. The rationale of this transformation is explained by the psychophysical law of Weber and Fechner. The distribution of antihistamine drugs, hepatoprotectors, cough syrups, sprays and lozenges for treating of throat pain, remedies for acne and vitamins over three price segments is computed as a case product segmentation. Proposed in the paper *post-hoc* methods of price-based segmentations can be integrated into accounting and managerial software providing the information support for management decisions.

References:

1. *Antonyuk A.A.* (2009), Countries and regions. Series: Economics and Business [in ukr.], №3, P.23–27.
2. *Bashuk T.O., Haydabrus N.V.* (2011), Marketing and management of innovation [in ukr.], №3, P.135-142.
3. *Hontchar V.I.* (2011), Sustainable economic development [in ukr.], №2, P.267-271.
4. *Hudzenko O.P., Kozytska K.I.* (2009), Ukrainian Medical Almanac [in ukr.], Vol.12, №5, P.63-66.
5. *Kulyba L.R., Dibrova T.H.* (2010), Actual problems and economics and manamement [in ukr.], №4, available at: http://probl-economy.kpi.ua/ru/node/167
6. *Kuo R.J., Ho L.M., Hu C.M.* (2002), Computers & Operations Research, №11, P.1475-1493.
7. *Lockshin L.S., Spawton A.L., Macintosh G.* (1997), Journal of Retailing and Consumer Services, №3, P.171-183.
8. *Mnushko Z.M., Sotnikova N.V.* (2006), News of Pharmacy [in ukr.], №3, P.C.57-62.
9. R: a language and environment for statistical computing (2013), available at: http://www.lsw.uni-heidelberg.de/users/christlieb/teaching/UKStaSS10/R-refman.pdf
10. *Sokolova L.V., Veryasova H.M., Sokolov O.Ye.* (2011), Journal of National University «Lviv Politechnic». Series «Management and Entrepreneurship in Ukraine: the stages of formation

and development problems» [in ukr.], №720, P.240-248.

11. *Ward J.H.* (1963), Journal of the American Statistical Association, №301, P.236-244.

КЛАСТЕРНЫЙ АНАЛИЗ КАК РАЦИОНАЛЬНЫЙ СПОСОБ СЕГМЕНТАЦИИ ФАРМАЦЕВТИ-ЧЕСКОГО РЫНКА ПО ЦЕНОВОМУ ПРИЗНАКУ

О.Т. Девиняк[1,2], Т.И. Худан[1], М.И. Кусий [3]
ГВУЗ «Ужгородский национальный университет»[1], г. Ужгород, Украина
Львовский национальный медицинский университет имени Данила Галицкого[2], г. Львов, Украина
Львовский государственный университет безопасности жизнедеятельности[3], г. Львов, Украина

Резюме: В данной статье проведено сравнение трех возможных методов *post-hoc* сегментации рынка по ценовому признаку: метода k-средних, иерархического кластерного анализа и метода исчерпывающего поиска. Раскрыты сильные и слабые стороны указанных подходов и даны рекомендации для практического применения.

Ключевые слова: сегментация рынка, кластерный анализ, сегментация фармацевтической продукции, ценообразование, метод k-средних, иерархический кластерный анализ.

КЛАСТЕРНИЙ АНАЛІЗ ЯК РАЦІОНАЛЬНИЙ СПОСІБ СЕГМЕНТАЦІЇ ФАРМАЦЕВТИЧНОГО РИНКУ ЗА ЦІНОВОЮ ОЗНАКОЮ

О.Т. Девіняк[1,2], Т.І. Худан[1], М.І. Кусій[3]
ДВНЗ «Ужгородський національний університет»[1], м. Ужгород, Україна
Львівський національний медичний університет імені Данила Галицького[2], м. Львів, Україна
Львівський державний університет безпеки життєдіяльності[3], м. Львів, Україна

Резюме: У даній статті проведено порівняння трьох можливих методів *post-hoc* сегментації ринку за ціновою ознакою: методу k-середніх, ієрархічного кластерного аналізу та методу вичерпного пошуку. Розкрито сильні та слабкі сторони вказаних підходів та надані рекомендації для їх практичного застосування.

Ключові слова: сегментація ринку, кластерний аналіз, сегментація фармацевтичної продукції, ціноутворення, метод k-середніх, ієрархічний кластерний аналіз.