Econometric analysis of the decision to increase the sales volume of the organization

Gabrielyan Anna A.
ASUE, Faculty of Computer Science and Statistics, aspirant (Yerevan, RA)
gabrielianna5@gmail.com

UDC 330.43:339.13
Keywords: Indicators (benchmarks), Amazon, KPI (key performance indicator), online business, econometric analysis

Կազմակերպության վաճառքի ծավալների ավելացման որոշումների կայացման էկոնոմետրիկ վերլուծությունը

Գաբրիելյան Աննա Ա.
ՀՊՏՀ տնտ. ինֆ և տեղ. համ. ամբիոնի ասպիրանտ (Երևան, ՀՀ)
gabrielianna5@gmail.com

Ամփոփագիր. Յուրաքանչյուրը, ով լուրջ է վերաբերվում Ամազոնում օնլայն առևտրի բիզնեսին պետք է ուշադրություն դարձնի գործունեության հիմնական ցուցանիշներին (KPI): Լավ վաճառք ապահովելու համար անհրաժեշտ է մշտապես հետևել մի շարք ցուցանիշների:Սակայն մեկից ավել չափանիշների դեպքում դժվար է որոշել, թե որ տվյալները կարող են օգտակար լինել: Ժամանակի ընթացքում մշակվել և ստեղծվել են առաջադեմ ալգորիթմներ և ավտոմատ գործընթացներ, որոնք իրենց վրա են վերցնում այս գործառույթը: Այսպիսով, Ամազոնում հաջողություն գրացելու համար անհրաժշետ է իմանալ, թե որոնք են Ամազոնում գործունեության հիմնական չափանիշները, ինչու են դրանք կարևոր և ինչպես կարող եք դրանք վերահսկել և ստուգել: Այս հոդվածում նկարագրվել է Ամազոնում գործունեության հիմնական ցուցանիշներից մի քանիսը և կատարվել է էկոնոմետրիկ վերլուծություն դրանց հիման վրա:
Հանգուցաբառեր՝ ցուցանիշներ(չափանիշներ), Ամազոն, KPI (կատարման հիմնական ցուցանիշին), օնլայն բիզնես, էկոնոմետրիկ վերլուծություն

Эконометрический анализ решения об увеличении объема продаж организации

Габриелян Анна А.
аспурант, АГЭУ, Факультет информатики и статистики (Ереван, РА)
gabrielianna5@gmail.com

Аннотация: Всем, кто серьезно относится к онлайн-торговле на Amazon, стоит обратить внимание на ключевые показатели эффективности (KPI). Чтобы обеспечить хорошие продажи, вам нужно всегда следить за несколькими показателями, но при использовании нескольких показателей трудно определить, какие данные могут быть полезны. Со временем были разработаны усовершенствованные алгоритмы, и автоматизированные процессы взяли на себя эту функцию. Итак, чтобы добиться успеха в Amazon, вам нужно знать, какие ключевые показатели есть на Amazon, почему их можно проверять и как ими можно управлять. В этой статье описаны некоторые ключевые показатели эффективности Amazon и на их основе проведен эконометрический анализ.
Ключевые слова: Индикаторы (бенчмарки), Amazon, KPI (ключевой показатель эффективности), онлайн-бизнес, эконометрический анализ.

Anyone who is serious about doing business online at Amazon should pay attention to key performance indicators (KPIs). For doing real and healthy sales, it is necessary to constantly follow these indicators. But with so many criteria, it is difficult to determine which data may be useful. Over time, advanced algorithms have been developed, automated processes that perform these operations. So to be successful on Amazon, you need to know what the key performance metrics are on Amazon, why they are important, how can be checked, and how you can control them. This article describes some of the key performance indicators on Amazon, econometric analysis has been performed.

In order to make effective brand decisions on Amazon, you need to use key performance metrics to help you oversee your business at Amazon. One of these indicators is the number of page views, which shows how successful paid and organic marketing strategies are on Amazon. One of these indicators is the session, which allows you to determine, for example, how many out of 100 visits will lead to sales. Important indicators are also the number of clicks and Click-Through Rate (CTR). No less important are CPA (Cost per Acquisition), Average Sales per Order Item, Ads - Impressions, Cost Per Click (CPC), etc.

With the help of the above indicators it is possible to identify the weak points of the business, which can be improved and to have more sales և a large number of page views. Thus, in online trading organizations, a systematic approach is needed to study problems, offer solutions, economical-mathematical modeling to maximize sales and / or page views, and to identify all the key phenomena within that system.

An econometric analysis was performed based on research-based factors to maximize product sales on Amazon and / or page views. In the first stage, the factors characterizing the field were selected. Then, an econometric analysis was performed to determine the relationship between the relationships between them, a correlation matrix was drawn to determine whether the observed indicators were the factors influencing the output, and a time series stationary check was performed to make the estimates reliable and effective. The Fisher test tested the statistical value of the regression equation, the strength of the bond, and so on.

Thus, the indicators characterizing the system are the following for the number of sales of the product (Y1): Sessions (X1), Impressions (X2), Click-Thru Rate (X3), Cost Per Click (X4), Average Sales per Order Item (X5), Clicks (X6), and for page count (Y2): Average Sales per Order Item (X1), Sessions (X2), Clicks (X3), Cost Per Click (X4), Click-Through Rate (X5) Impressions (X6).

The choice of these indicators is due to the fact that the company for which the analysis is performed sells on the Amazon platform, the latter provides statistics on the type of limited indicators. However, these indicators are the main factors influencing the number of sales and views.

As a result of the correlation, it turned out that only the following factors meet the set conditions:

	Y_1		Y_2
X_1	0.9255	X_2	0.9997
X_2	0.5882	X_3	0.5416
X_3	-0.549	X_6	0.5524

T**able 1**. *From the correlation matrix of indicators characterizing the organization*

Since each set of system characteristics is closely correlated with the corresponding Y, econometric analysis can be continued to make a regression equation based on the data obtained. They will look like this:

$$Y_1 = 4.23X_1 - 5.23X_2 - 2780.7X_3 + 1825.76$$
$$Y_2 = 1.52X_2 + 17.76X_3 - 0.13X_6 - 29.86$$

When making time series predictions, several statistical models are usually built, and there is a need to choose the one that best suits the current situation and is more reasonable. In this regard, the better the model explains the past of the event under study, the more realistic it will predict the future. The value of the models is then assessed using actual and predictive indicators that show how applicable the chosen model is. One such method of estimation is the calculation of the Tale coefficient. It shows the degree of compliance of the time series. The closer the Tale coefficient is to 0, the more comparable the time series.

$$v = \frac{\sqrt{\frac{1}{T}\sum_{t=1}^{T}(Y_t^s - Y_t^a)^2}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T}(Y_t^s)^2} + \sqrt{\frac{1}{T}\sum_{t=1}^{T}(Y_t^a)^2}}$$

The calculations for the obtained model show that the Tale coefficient is $Y_1 = 0.35206$ and

$Y_2 = 0.50296$ (*calculated by the author*), which indicates the practical significance of the obtained model.

Because the observations of the studied indices were made for 24 months, it is necessary to check the stationarity of time series, otherwise the construction of models based on time series with different types of stationarity may lead to inadequate models. For the latter, the provisions of the least squares method will not apply, which in turn will lead to instability and inefficiency of the received estimates.

To check the stationarity of dynamic series, we used the method of dividing the series into successive groups. That is, each time series was divided into four six groups and calculated for each groups average arithmetic and dispersion. Then, the student values of the Student's t-criterion are compared with the tabular values, since all these quantities were smaller than the tabular values, the stationarity of the series is confirmed.

The usually obtained regression equation is statistically evaluated on the basis of the F-test. In our case $F_{account} = 2.04$. For the analysis of the equation, do the inverse: 1 / 2.04 = 0.49 and find the critical value F in the corresponding table, provided that the degree of freedom for the numerator is f_1=k, where k is the number of factors, 3, and for the denominator, f_2=n-k-1, where n is the number of observations, is f_2=24-3-1=20. We will have $F_{crit} = 3.10$. It is obvious that the account $F_{account} < F_{crit}$, then it can be argued that our regression equation has a fairly high degree of equivalence.

The number of observations is 24, the number of variables is 3, then $R^2_{կրհun}$= 0.446. With this indicator, we are sure that the obtained regression

equation is quite reliable, since the R^2account $>R^2$ crit.

For the completeness of the analysis, it is necessary to analyze the individual values of the regression coefficients using the Student's t-criterion. Student distribution is widely used in experimental data processing. for example, when it is necessary to construct reliability intervals Hypothesis of mean dispersion of unknown dispersion [**1**, pp. 51-52]. To determine t_{crit} , use the Student Distribution Chart when v=n-k-1=24-3-1=20: ($t_{crit} = 2{,}06$).

Comparing this data with the data of our t-test, we see that in Y_1 the factors x_1 and x_3, and in Y_2 x_2 and x_3 do not satisfy the $t_{account} \geq t_{criterion}$. Then we subtract the unsatisfactory factors from our regression equation in Y1 x_1 and x_3, and we subtract x_2 and x_3 from the equation Y2.

After multi-step regression analysis, the regression equation looks like this:

$$Y_1 = 6.63x_2 + 914.88$$
$$Y_2 = 3.68_{X6} + 48.57$$

The coefficient of determination for Y_1 became 0.645048 and for Y_2 0.4988355, so the equation value hypothesis is not rejected with 95% probability.

Residue graphs show that residues (which should be random, not correlated with each other, not having the same dispersion) are inadequate:

a) there are segments with strongly positive residues segments with strongly negative residues, i.e. the residues are related to each other. For example, if the remainder in a passage is positive, then most likely the remainder next to it is positive. At first glance, this is a sign of autocorrelation, so you need to check for it.

b) the level of dispersion of residues is not stable. There are residues with large modulus (high dispersion) and residues with medium modulus (low dispersion). At first glance, this is a sign of heteroscedasticity.

c)

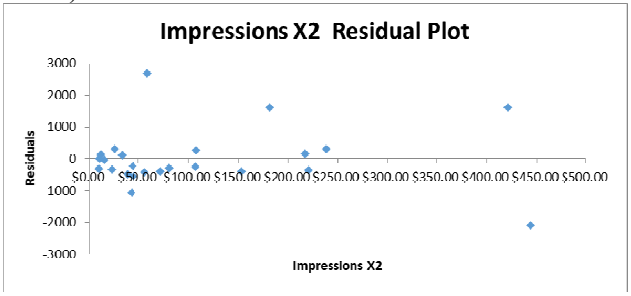

Figure 1. Dynamics of balances (deviations) for Y_1 (*The chart was edited by the author*)

Figure 2. Dynamics of balances (deviations) for Y_2 (*The chart was edited by the author*)

The independence of the residues is checked using the Darbin-Watson criterion:

$$d_1 = \frac{\sum_{i=2}^{n}(\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{i=2}^{n}\varepsilon_t^2} = \frac{143{,}210{,}464.92}{110{,}079{,}142.63} = 1.3$$

$d_1 = 1.3$

$$d_2 = \frac{\sum_{i=2}^{n}(\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{i=2}^{n}\varepsilon_t^2} = \frac{16{,}646{,}923.12}{13{,}406{,}767.63} = 1.24$$

$d_2 = 1.24$

Since $d<d_1$ and $d<d_2$ then we can say that there is autocorrelation. This means that our information on product sales and page views is incomplete. To eliminate autocorrelation, it is necessary to include other indicators in the model that will clarify the model.

To find out if there are different dispersions, use the Goldfeld-Quantt method. Arrange the observations as the x variable increases, divide them into two groups of 12, and construct the corresponding tables.

By relating the sums of the squares of the balances, we get the required calculation value:

$$F_{Ñ^3ßí} = \frac{S_{2\hat{y}}}{S_{1\hat{y}}} = \frac{149{,}487{,}444.22}{160{,}186{,}731.21} = 0.93$$

And since $F_{crit}(0.05,12,12)=2.69$, according to the Fisher scale, it turns out that $F_{Ñ^3ßí}<F_{Ïñçí}$, then the hypothesis of heteroskedasticity is refuted, we see that the order of equality of the dispersions of the residual quantities is not violated.

More detailed studies allow us to derive the coefficient of elasticity, Ey, which shows how sales volume and / or page views change with a one percent change in the x index.

$$E_{y1} = \frac{6.63 \cdot x}{914.88 + 6.63 \cdot x}$$

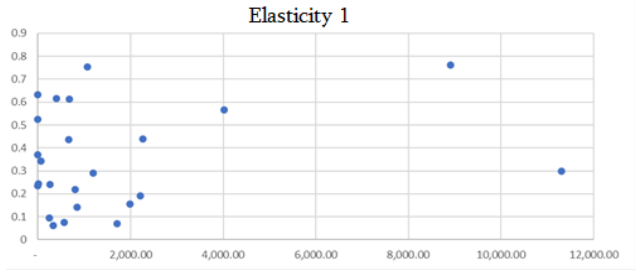

Figure 3. Elasticity of the relationship between sales volume and x (impressions) (*The chart was edited by the author*)

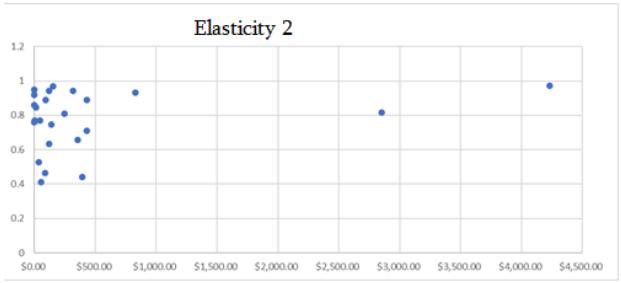

Figure 4. Elasticity of the relationship between page views and x (impressions) (*The chart was edited by the author*)

The coefficient of elasticity shows that a 1% increase in the x index does not lead to an increase in sales even in the long run.

$$E_{y2} = \frac{3{,}68 \cdot x}{48{,}57 + 3{,}68 \cdot x}$$

The elasticity coefficient shows that a 1% increase in the x index does not lead to an increase in page views even in the long run.

From the four evaluated functions, the connection of the appearance of a polynomial was chosen as the calculation.

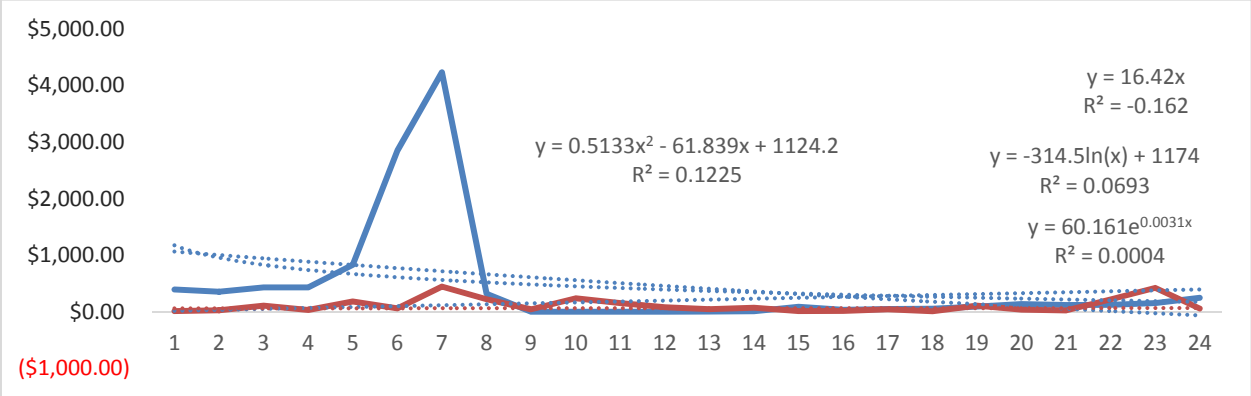

Figure 5. Types of sales volume functions (*The chart was edited by the author*)

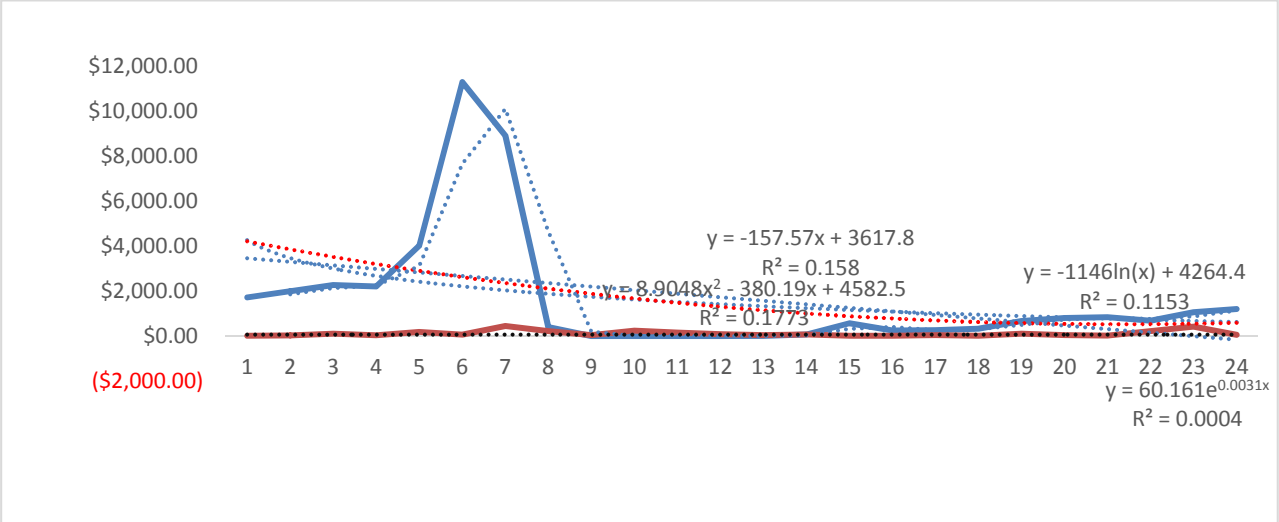

Figure 6. Page view function counts (*The chart was edited by the author*)

Y_1=8.9048x_2-380.19x+4582.5
Y_2=0.5133x_2-61.839x+1124.2

As a result of econometric analysis, we see that the same indicator has the greatest impact on the two results, that is, the display of ads / page - Impressions. n addition, we have obtained functions for two output indicators that describe the correlation between the input variable ы output indicator. We will consider these two functions as target functions for solving the STEM decision-making problem.

List of used literature

1. **Кобзарь А. И.**, Прикладная математическая статистика. Для инженеров и научных работников, ФИЗМАТЛИТ, 2006.
2. https://www.roirevolution.com/blog/2020/02/top-9-business-level-kpis/?fbclid=IwAR2kzC-Vah4D_TGdzObiCOoMwJBHEKXOrp86AJOi0rZ1bo5PTE_LGaotBPA
3. https://salesfunnelhq.com/amazon-kpi/?fbclid=IwAR2pqfIaTv4ChK7ogSSfD3jq4fTudbSHUI9cSojW45pOak2ldm-1gMYENBY

Сдана/Հանձնվել է՝ 17.11.2021
Рецензирована/Գրախոսվել է՝ 30.11.2021
Принята/Ընդունվել է՝ 08.12.2021