INTERNATIONAL JOURNAL ON HUMAN COMPUTING STUDIES

www.journalsresearchparks.org/index.php/IJHCS e-ISSN: 2615-8159|p-ISSN: 2615-1898

Volume: III Issue: 3 May-June 2021

OTHER WAYS TO BUILD CORRELATION MODELS

Sh. B. Ochilov, A. N. Tagaev, O.K.Khudayberdieva

Bukhara, Uzbekistan

Bukhara engineering technological institute

---***---

Abstract: *This article has developed a new method for building production functions. On the basis of accurate data, the methodology for calculating the coefficients of any mathematical models is shown.*

Key words: *production function, correlation analysis, product function, approximation, forecasting.*

Introduction

Reflecting the main connections in nature and society in mathematical language has always been one of the main problems of science. Everyone knows that this problem has been solved in general cases in science. The most complex methods of this process are the collection of statistical figures about an existing object, the preliminary hypothesis that the function available on the basis of these figures is reflective of this process, proving it in mathematical ways. By this time, all specialists will effectively use this path.

Of course, there are shortcomings in this method, as well as in certain shortcomings of any method. Predictions derived from the construction of mathematical models in economics do not always provide clarity. Everyone knows that seismology has a very low accuracy level of forecasts based on models under construction.

Therefore, the results obtained from mathematical models in many areas are used not as the main tool for compiling conclusions, but as auxiliary information. This thing has the same drawbacks as in medicine, physics and biology.

For instance, let y be the cotton yield and x be the amount of local fertilizer applied to the soil.

As a result of statistical analysis, it is known that the $y = 18{,}0 + 1{,}6x$ link is established. In that case, when 10 tons of local fertilizer is applied to 1 hectare of land, the employee who works in practice will never accept the assertion that the yield will be $y = 34\ s/ha$ as 100% correct.

So are there any ways to further increase the accuracy level of the $y = f\left(x_1; x_2; \ldots x_n\right)$ correlation relationships that are being built, or their role in their production, or are the methods used to date the last resort?

This article has tried to prove that there are such ways.

First of all, let's think about the fact that the most mistakes can come out within the hypotheses that are made in tradition methods:

- The first shortcoming is that in our opinion is y_i - the results of the experiment directly - correspond directly to a certain class of functions.

- We think that the second shortcoming is that we do not take into consideration the change nature in the experiments results obtained from point to point.

INTERNATIONAL JOURNAL ON HUMAN COMPUTING STUDIES
www.journalsresearchparks.org/index.php/IJHCS e-ISSN: 2615-8159|p-ISSN: 2615-1898
Volume: III Issue: 3 May-June 2021

If it were possible to avoid the above drawbacks, then we would increase the level of correlation that will be detected to real realities. Research conducted in this way shows that in order to realistically express the processes, we need to look for the product of this function, not the relationship $y=f(x_1;x_2;....x_n)$, based on the given statistical numbers.

With a fully constructed derivative function, it will not be difficult to find a single integrated primary connection. Based on the results of the given experiment, the structure of the derivative function makes it possible to take into account the characteristics of the decrease in the growth of the function, flexure– bending, which characterizes this process.

So on the basis of y_i - we deal with the problem of finding the product of the function, not the function itself.

It is known from the course of mathematical analysis that $\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}\approx f'(x_0)$ or $\frac{y_{i+1}-y_i}{\Delta x}\approx f'(x_0)$ for statistical numbers.

Hence, if y_i - are statistical numbers and their corresponding x_i are given by $i=1,n$, then at y $n-1$ points we can also find the values of the product function corresponding to the point corresponding to the points x_i. This allows us to find $\varphi=f'(x)$. Let's take a brief example.

Provide information on cotton yields and organic fertilizers used (figures are conditional). Based on them, we will find a productive function and then integrate the main function. We draw conclusions by comparing the numbers found using the new path with the numbers found based on the connection found in the inaccessible way.

	y	x	Δy	Δx	$\frac{\Delta y}{\Delta x}$	$\frac{\Delta y}{\Delta x}\cdot x$	x^2	x_{-1}
	24	3,5	-	-	-	-	-	-
	24,3	3,8	0,3	0,3	1	3,8	14,44	3,8
	24,7	4,1	0,4	0,3	1,33	5,47	16,81	4,1
	25,2	4,3	0,5	0,2	2,5	10,75	18,49	4,3
	25,3	4,5	0,1	0,2	0,5	2,25	20,25	4,5
	24,8	4,7	-0,5	0,2	-2,5	-11,75	22,09	4,7
	25,7	4,75	0,9	0,05	18	85,5	22,5625	4,75
	25,9	4,8	0,2	0,05	4	19,2	23,04	4,8
	26,1	4,85	0,2	0,05	4	19,4	23,5225	4,85
	27	5	0,9	0,15	6	30	25	5
Σ	253	44,3			34,83	164,62	186,21	40,8

Here y - is the cotton yield and x - is the amount of local fertilizer applied to 1 ha of land.

Suppose $\frac{\Delta y}{\Delta x}$ - grows linearly, i.e let $\frac{dy}{dx}=a_0+bx$. Let's find this connection. In the small squares method, $\sum\frac{\Delta y_i}{\Delta x_i}\cdot x_i$ is used instead of $\sum\frac{\Delta y}{\Delta x}\sum y_i\cdot x_i$ instead of $\sum y_i$ and we have the following system of equations.

$$\begin{cases}\sum_{i=1}^{9}\frac{\Delta y}{\Delta x}=9a_0+b\cdot\sum_{i=1}^{9}x_i\\ \sum\frac{\Delta y}{\Delta x}\cdot x=a_0\cdot\sum_{i=1}^{9}x_i+\sum_{i=1}^{9}x\end{cases}$$

Based on the numbers above, we get the following and solve this system to find the unknowns of the necessary coefficient.

$$\begin{cases}34{,}83=9a_0+40{,}8b\\ 164{,}6=40{,}8a_0+186{,}17b\end{cases}$$

$$a_0=-23{,}645$$

$$b=6{,}07$$

That is, by integrating $\frac{dy}{dx} = -23{,}645 + 6{,}07x$, c is determined based on the initial condition $y = \int(-23{,}645 + 6{,}07x)dx$

$y = -23{,}645x + 3{,}035x^2 + c$ and

$y = 3{,}035x^2 - 23{,}645x + 69{,}5795$ 2y-3 is determined.

But using traditional methods it is possible to solve this problem and find the linear function $y = 18{,}212 + 1{,}6x$. Now the question arises as to what if the process is in the form of a multivariable function. It should be noted that in this case a lot of mathematical problems arise. In this case, the problem cannot be solved in one attempt.

We propose to define this problem for each variable separately as $y_i = f(x_i)$ for each variable, and then to define the resulting function as the average of the sum of the functions. That is $y = f(x_1; x_2; \ldots x_n) = \frac{f_1(x_1) + f_2(x_2) + \ldots + f_n(x_n)}{n}$

Let's look at this solution in a concrete example.For convenience, we consider the above example as a function of two variables rather than one variable. Hence y - is the cotton yield; let x_1 - be the amount of local fertilizer (T) per hectare and x_2 - be the amount of mineral fertilizer per hectare (kg). Below are the results of several years of observations. Let $y = f(x_1; x_2)$ be required to determine the relationship by constructing the product function.

Table 2

	y	x_1	x_2	Δy	Δx_1	Δx_2	$\frac{\Delta y}{\Delta x_1}$	$\frac{\Delta y}{\Delta x_2}$	$\frac{\Delta y}{\Delta x_2} \cdot x_2$	$\frac{\Delta y}{\Delta x_1} \cdot x_1$	x_2^2
	24	3,5	80	-	-	-	-	-	-	-	-
	24,3	3,8	87	0,3	0,3	7	1	0,043	3,729	3,8	7569
	24,7	4,1	95	0,4	0,3	8	1,33	0,050	4,750	5,467	9025
	25,2	4,3	98	0,5	0,2	3	2,5	0,167	16,333	10,75	9604
	25,3	4,5	105	0,1	0,2	7	0,5	0,014	1,500	2,25	11025
	24,8	4,7	110	-0,5	0,2	5	-2,5	-0,100	-11,000	-11,75	12100
	25,7	4,75	117	0,9	0,05	7	18	0,129	15,043	85,5	13689
	25,9	4,8	121	0,2	0,05	4	4	0,050	6,050	19,2	14641
	26,1	4,85	127	0,2	0,05	6	4	0,033	4,233	19,4	16129
	27	5	135	0,9	0,15	8	6	0,113	15,188	30	18225
Σ	253	44,3	1075				34,833	0,498	55,826	164,617	112007

The first of the functions sought, $y_1 = f_1'(x_1)$ had already been calculated. So now we only calculate $y_2 = f_2(x_2)$. Taking the relevant data from the table №2 above, we construct an equations system for $y_2 = f_2(x_2)$

$$\begin{cases} 0{,}469 = 9a_0 + 995b \\ 54{,}98 = 995a_0 + 112011b \end{cases}$$

Solve the equation and find $y = -0{,}11989 + 0{,}001556x_2$. But we found $\frac{dy_2}{dx_2} = -0{,}11989 + 0{,}001556x_2$. Integrating it on the basis of the condition $y_2(80) = 24$, we get $y_2 = 28{,}62 - 0{,}11989x_2 + 0{,}000776x_2^2$.

So we also created $f_1(x_1) = 3{,}035x^2 - 23{,}645x + 69{,}5795$ and $f_2(x_2) = 28{,}62 - 0{,}11989x_2 + 0{,}000776x_2^2$. Now the resulting relationship $y = f(x_1; x_2) = 49 - 11{,}8225x_1 + 1{,}5175x_1^2 - -0{,}0599x_2 + 0{,}000388x_2^2$ since $f(x_1; x_2) = \frac{f_1(x_1) + f_2(x_2)}{2}$.

This function can be used to estimate how close the calculated relationship is to the actual values by calculating the values at the corresponding points.

Now, based on the above numbers, let us construct a model of $f(x_1; x_2)$ in the form $y = a_0 + a_1x_1 + a_2x_2$ using traditional methods and compare the results with the results of the connection constructed on the basis of the method we propose.

Table № 3

$$\begin{cases} 252,7 = 10a_0 + 44,3a_1 + 1075a_2 \\ 1126,95 = 44,3a_0 + 198,455a_1 + 4838,5a_2 \\ 27302 = 1075a_0 + 4838,5a_1 + 118407a_2 \end{cases}$$

Solving the equations system, we obtain $y = 21,8344 - 1,12383x_1 + 0,07827x_2$.

Now we have the function $f(x_1; x_2) = 49 - 11,8225x_1 + 1,5175x_1^2 - 0,0599x_2 + 0,000388x_2^2$ generated by the new method and $y = 21,8344 - 1,12383x_1 + 0,07827x_2$ found in the traditional method. Comparing the results obtained on the basis of these two functions, it is possible to be sure that the proposed method does not lag behind the traditional methods.

The following table shows y_x is an actual experimental results, y_T are results calculated by the traditional method, y_φ is a sequence of numbers calculated by the proposed method.

But there are many questions that require mathematical validation.

1) is the proposed method always effective?

2) is it possible to compare these two methods mathematically.

3) how the reliability of predictions can be assessed mathematically.

These are questions that need to be addressed in the future.

Reference

1. Kremer N. Sh. Econometrics: textbook . – M.: Publishing House, 2018 Yurayt. – p 354.

2. State statistics committee of the Republic of Uzbekistan [electronic resource] - access procedure:http://www.stat.uz