Investigation of Long-Term Change in Web-Based Data Traffic in Terms of Performance Evaluation

Web Tabanlı Veri Trafiği Uzun Vadeli Değişiminin Performans Değerlendirme Açısından İncelenmesi

Gözde BOZKURT[1], Ahmet Mete ÇİLİNGİRTÜRK[2]

ÖZET

Çalışma, web sitelerinin tıklama verilerinden elde edilen sayfa merkezli verilerin dağılımlarını modellemeyi amaçlamaktadır. Bu modelleme sayesinde simülasyon ve planlama modelleri için kaynak oluşturulması amaçlanmaktadır. Çevrimiçi pazarlama davranışlarının anlaşılmasında web sitelerinin kullanımı ve göstergelerinin analizi oldukça önemlidir. İş hedefleri doğrultusunda hazır giyim ve aksesuarların pazarlandığı e-ticaret ve bilgi edinme ve yayma amaçlı haber siteleri hedef olarak belirlenmiş web siteleridir. İlgili web sitelerinin hemen çıkma oranı, sayfa görüntüleme ve kalış süresi göstergelerinin dağılımları incelenmiştir. Elde edilen dağılım bilgileri ile özel sektörde kullanılan satış simülasyonları başta olmak üzere web siteleri arası performans değerlendirme gibi çeşitli alanlarda kullanılabileceği düşünülmektedir.

ABSTRACT

In this study, the development levels of Developing 8 countries were tried to be concretized in line with the growth-centered, human-centered and environment-centered development approach and criteria. From 1997 to the present, Bangladesh, Nigeria and Pakistan have arisen from the low-income group to the lower-middle-income group, Indonesia and Iran form the lower-middle-income group to the upper-middle income group. Lower-middle income Egypt, Malaysia and Turkey have continued to stay in the upper-middle income groups. Between 1997-2018 the share of industrial production in total value-added Indonesia, Iran, Malaysia, Nigeria and Turkey have decreased. It remained stable in Egypt and increased in Bangladesh and Pakistan. The share of manufacturing production in total value added follows a similar course to industrial productions. Between 1997 and 2018, D-8 countries except Nigeria managed to raise their human development levels to the next level. Bangladesh and Pakistan have been promoted from low human development level to medium human development. Egypt, Indonesia and Iran were able to move from medium human development level to high human development level. Malaysia has risen from a high level of human development to a very high level of human development. Meanwhile, Turkey was able to jump very high human development levels from medium human development level. It observed that the performance of D-8 countries in realizing the United Nations Sustainable Development Objectives varies between 50% and 70%. Iran, Malaysia and Turkey, for the moment, is located above the middle of sustainable development league. Egypt is right in the middle, Indonesia, Bangladesh, Pakistan and Nigeria are located below in the middle. It seen that Malaysia takes the first place among D-8 countries in terms of development of financial institutions and financial development course between 1997 and 2018, it is observed that the financial development index value shows a steady increase trends in countries other than Egypt, Indonesia and Pakistan.

[1] *University of Beykent, Department of Economics, İstanbul/Türkiye, gozdebozkurt@beykent.edu.tr,* **ORCID-ID:** *0000- 0001-8413-1099*
[2] *Prof. Dr., University of Marmara, Department of Econometrics, İstanbul/Türkiye, acilingi@marmara.edu.tr,* **ORCID-ID:** *0000-0001-8677- 7969*

INTRODUCTION

Along with the high acceleration of information technologies, the impact of many different disciplines, especially communication channels, is felt intensely. The new platforms, which have been formed as a result of the traditional communication channels changing shape with the developing technology, have also changed the communication of the institutions with their customers. Social Software, which is described as interactive and internet-based applications, has increased the speed of this dynamic structure. Due to the transfer of communication to new channels, an extensive communication network has been formed, and the necessity of providing quality network service has occurred. This requirement brought about planning and modeling of web-based data traffic. These models are named as planning and simulation models and allow them to produce experimental results based on real life.

In this study, the aim is to estimate the distribution of the page-centered data consisting of user-centered click data of the websites and to create a source for simulation and planning models. The websites examined for the purpose are websites with news sources for information sharing and fashion websites evaluated within the scope of e-commerce. Due to the content of these differently constructed websites, there are two different views that the website is differentiated in terms of the duration of the website and the number of pages the users visit, or on the contrary, it does not differ. The main objective of the study is to determine whether the distribution patterns of website statistics have changed in the long run, and if there is constancy, it is recommended to focus on the distributions in which planning is determined.

The program is used to study physics, geography, medicine, etc. In addition to being used by researchers in different disciplines, the usability of different distributions in the field of statistics has also been demonstrated. The correlation of the variables examined obtained marginal distributions, and from this point of view, it seems possible to obtain conditional distributions in other studies to be carried out. For example, with a conditional distribution that can be obtained, if the bounce rate of the users on the websites is 1, the information about the number of reads will be reached. In addition to this, it is possible to use Bayesian distribution methods as a method. It is thought that the distribution information obtained from the variables can be used in many areas, such as sales simulations used in the private sector, the self-performance of multimedia salespeople, and the performance evaluation between websites.

According to the study, it is possible to use the distribution results of the variables in the performance evaluation between the websites, especially in the sales simulations in the private sector and in the evaluation of the performances of the multimedia sales representatives. Due to the changing behavior of internet users, it would be appropriate for researchers to make evaluations for other website indicators using distributions based on constantly updated data.

1. LITERATURE REVIEW

There are two different theories in the related literature on simulation and planning models, namely "social impact" and "social presence." In the social impact theory, the effect of individuals on each other in the face of any social situation is examined. Social impact occurs as the instant announcement of the incident, perceived influence power of the resources, and the number of individuals responding to the resource (Adeyemi vd., 2018). In the theory of social presence, the

media is acceptable to the extent that visual, auditory, or physical communication between the communication sickles can be achieved.The behavioral interaction between communication partners also acts in direct proportion with social presence. Some sample studies conducted in the literature for planning and simulation models within the scope of related theories are given in **Table 1**.

Author(s)	Publication Year	Publication Name	Result
Huberman B.A., Pirolli P., Pitkow J.E. and Lukose R.	1998	Strong Regularities in World Wide Web Surfing	It has been determined that the number of pages that users move from one web page to another fits the Reverse Gaussian distribution. Thus, it has been suggested that web pages can be used as useful information in the modeling of their costs and designs. Also, some firms spread information to more pages to enable users to visit more pages. For this reason, it was emphasized that the total time spent on the page should be taken into consideration instead of the number of pages.
Johnson E.J., Bellman S. and Lohse G.L.	2003	Cognitive Lock-In and the Power Law of Practice	It is determined that the time spent on the website decreases more frequently.
Arora A., Krishnan R., Nandkumar A., Telang R. and Yang Y.	2004	Impact of Vulnerability Disclosure and Patch Availability - An Empirical Analysis	The logarithm of this variable is used in models, assuming that the time spent in the session comes from the exponential distribution family. It has been stated that the intervals between transactions made over the internet comply with heterogeneous hazard mixture models that cannot be observed.
Dalessandro B., Hook R., Perlich C., and Provost F.J.	2012	Evaluating and Optimizing Online Advertising: Forget the Click, but there are Good Proxies	Click, turnover, and hits are used in the performance measurements of e-commerce sites, which is one of the channels developed by the web servers and where communication between communication partners is provided. It has been demonstrated that the number of hits in the current research turned into buying behavior with rates of 2% in 1995 and 0.08% in 2010. For this reason, it was thought by managers that the number of clicks and the rate of the purchase were insufficient and misleading information. Therefore, it has been argued that access to websites and time spent is more suitable indicators.
Nottorf F.	2014	Modeling the Clickstream Across Multiple Online Advertising Channels Using a Binary Logit with Bayesian Mixture of Normals	After entering the relevant web page, the time spent in the session, the number of previous entries, and the total time spent was found to be significant in these models. It is assumed that the model parameters comply with the Gaussian distributions, and the unit totals are Multinomial distributed.
Heuer D., Brettel M. and Kemper J.	2015	Brand Competition in Fashion ECommerce	In the study, which deals with companies that carry out e-commerce in the field of ready-to-wear, it is tried to explain the competitive situations of the companies based on product addiction by using the transactions of four products in this field. As a result, it was seen that ten brands represent this sector, national and private, in total.
Adeyemi O.J., Popoola S.I., Atayero, A.A., Afolayan D.G., Ariyo M. and Adetiba E.	2018	Exploration of Daily Internet Data Traffic Generated in a Smart University Campus	In the daily-based analysis of the volume or download of user-based internet data and the distribution of loading amounts, model estimates have been made in line with the prediction that these distributions will generally comply with the Weibull, Logistics, Rician and Nakagami distributions. In the obtained result, it was determined that it shows the best fit for nonparametric distributions as well as the lowest log-similarity value in Nakagami distribution.

Source: Compiled by authors.

It is possible to obtain "user" and "site" centered data of the websites targeted to be examined according to the indicators. However, the need for user-centered data, such as individual recording and shared computer use, constitutes the disadvantage of this data collection technique (Bucklin, 2009). The role and effect of the internet, which is used for advertising and persuasion purposes, enables the understanding of online marketing behaviors as the usage of the websites can be analyzed as a result of obtaining the data of the indicators (Nottorf, 2014).

2. METHODOLOGY

The data analyzed by the purpose of the study were provided manually via the "Alexa Traffic Rank" site. Through this site, where the measurements of the websites are made and published, the indicators determined by the theory of simulation models for a total of 127 websites, 65 of which are e-commerce websites and 62 of them are news websites. These indicators are the rate of leaving the homepage, the duration of the site, and the number of page views. The number of page views refers to the average number of pages opened within the same site, and the duration of the site refers to the average number of pages opened by visitors on the site. It is aimed to make comparisons by estimating the distribution of the statistics of the websites: August-2018, January-2019, and March 2020. Related websites are grouped as e-commerce and news. Since there was a statistically significant difference between the groups during the periods examined, the distribution of the data from these two different sources was handled separately.

In the examination of whether there is a significant difference between the old and new distributions of the variables of the related websites, a separation analysis based on the Minimum Kullback-Leibler Distance Ratio was applied. The primary purpose of this analysis is to determine the fixed distribution by comparing the parameters of the theoretical distributions determined for the indicators of news and e-commerce sites. Determining the fixed distribution is essential in terms of examining the long-term changes of the relevant websites according to the indicators. It is an unsymmetrical measure of the distance between the two density functions, such as $f_1(x)$ and $f_2(x)$ first defined by Kullback and Leibler in 1951 for an X random variable in statistics. Bromideh and Valizadeh (2013) proposed the Minimum Kullback-Leibler Distance Ratio (RMKLD) method to determine which possible distribution can be modeled from any data set. This method determines the correct selection possibilities for distributions. The method's hypothesis, test statistics, and decision phase are given below.

$$H_0: X \sim f_1(x,.); \; H_a: X \sim f_2(x,.)$$

$$RMKDL = ln\left[\frac{D_{KL}(f_1 || f_2)}{D_{KL}(f_2 || f_1)}\right] \qquad (2.1)$$

If the test statistic is *RMKDL <0*, the H_0 hypothesis cannot be rejected. For the RMKDL method, log-similarity values were obtained from the probability density functions of the old (f_2) and new (f_1) distributions. Gauss formula, which expresses how likely the distribution observed in obtaining log-like values, was used.

$$L = \prod_i \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\hat{y})^2}{2\sigma^2}} \propto e^{-\chi^2} \text{ and generally}$$ *(2.2)*

$$L_{distr.}(\theta_1, \theta_2, ..., \theta_k; x_1, x_2, ..., x_n) = \prod_{i=1}^{n} f(x_i; \theta_1, \theta_2, ..., \theta_k)$$ it is written as.

Minimizing the chi-square value is equivalent to maximizing the likelihood of log-similarity. While the formula gives the full result in the case of the normal distribution, it provides the closest possible result when used in different distributions.

3. FINDINGS

This study aims to estimate the distributions of the page-centered data consisting of user-centered click data of the websites and to create a source for simulation and planning models. For this purpose, firstly, it has been investigated whether there are significant differences in the variables examined according to e-commerce and news sites, and the results are summarized in **Table.2**. Statistical analyzes were carried out via SPSS-18 version. In order to repeat the study regularly from 2018 and compare the results, the results of the relevant dates are given together in the tables.

Table 2. Statistical Analysis Results for Variables

Variable	**Daily Time on Site**			**Bounce Rate**			**Daily Pageviews per Visitor**		
Analysis	*August 2018*	*January 2019*	*March 2020*	*August 2018*	*January 2019*	*March 2020*	*August 2018*	*January 2019*	*March 2020*
Clothing-E-commerce	05:37 (03:42)	06:22 (04:37)	04:41 (02:48)	32,80 (10,22)	43,12 (20,34)	36,67 (12,46)	5,30 (2,66)	4,59 (2,83)	4,79 (2,51)
Skewness	2,170	1,218	1,538	1,005	1,123	1,663	2,377	1,862	3,103
Kurtosis	5,779	2,052	3,307	1,150	0,474	4,505	10,309	6,527	16,193
News-Source	06:14 (04:01)	08:23 (06:13)	04:01 (01:55)	54,79 (14,25)	64,01 (17,54)	55,70 (14,48)	3,11 (1,85)	2,47 (1,45)	3,15 (2,35)
Skewness	1,61	1,120	0,857	-0,022	-0,131	-0,079	2,334	1,452	2,484
Kurtosis	2,903	1,599	-0,152	0,784	0,490	-0,592	7,642	1,732	6,864
t-Test p-value	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Mann-Whitney-U	0,399	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Levene Homogeneity	0,086	0,121	0,385	0,003	0,006	0,000	0,065	0,072	0,50
K-S Normality	0,000	0,064	0,000	0,062	0,334	0,001	0,045	0,005	0,000
Chi-Square Normality	0,398	1,000	1,000	0,346	1,000	1,000	0,063	0,435	1,000

In the examination of the significant difference between the two independent groups, e-commerce and news sites, a t-test was applied, and it was observed that there was a significant difference between the groups. In order to check the consistency of this result, the Mann-Whitney U test, which does not seek normality conditions, was used, and it was observed that the same result was achieved. It was determined that the meaningful difference between the groups gave the same result in all the periods examined. Kolmogorov-Smirnov and Chi-square fitness tests were used to

examine the normal distribution of the groups. As a result of the tests, it was concluded that the skewness and kurtosis coefficients deviate from the normal distribution. The relationships between the variables examined within the scope of the groups are examined and given in **Table.3**.

Table 3. Correlation Results for Variables by Category

Correlation	Bounce Rate – Daily Pageviews per Visitor		Bounce Rate – Daily Time on Site		Daily Pageviews per Visitor – Daily Time on Site	
	News	E-Commerce	News	E-Commerce	News	E-Commerce
Pearson	-0,712	-0,674	-0,846	-0,716	0,738	0,916
Spearman	-0,834	-0,833	-0,840	-0,862	0,811	0,905

The purpose of analyzing the relationship between variables is that multiple related variables are needed in models used in simulation studies. As can be seen in **Table.3**, it is possible to say that there is a nonlinear relationship between the variables since the Spearman correlation coefficient between the variables gives stronger results. It is seen that there is a high positive correlation between the number of page views and the duration of the page in news and e-commerce sites. There is a negative relationship between the other variables examined for both groups. It was determined that the relationship structure between the variables examined in August 2018 and January 2019 was the same as stated above. Therefore, it is possible to say that the continuity of the relationship structure between the variables is consistent. EasyFit 5.6 program was used to examine the distribution of relevant variables examined within the scope of e-commerce and news site groups. The properties of the distributions determined by conformity tests and whose parameters are estimated are given in **Table.4**.

Table 4. Statistical Distribution Results of the Variables for the Categories

Category	E-Commerce Sites			News Sites		
Variable	August 2018	January 2019	March 2020	August 2018	January 2019	March 2020
Daily Time on Site	**Distribution of Dagum** k=0,80507 α=3,4082 β=0,21975	**Distribution of Dagum** k=0,29954 α=4,2096 β=0,29371	**Distribution of Dagum** k=0.70108 α=3.5011 β=0.19926	**Distribution of Johnson SB** γ=0.04165 δ=0.97989 λ=64.239 ξ=24.335	**Distribution of Inverse GaussTers Gauss (3P)** λ=0,98638 μ=0,26655 γ=-0,07814	**Distribution of Johnson SB** γ=0.86176 δ=0.71202 λ=0.35261 ξ=0.06528
Bounce Rate	**Distribution of Burr** k=0,80234 α=6,5053 β=29,559	**Distribution of Dagum** k=4,4936 α=2,9956 β=20,36	**Distribution of Burr** k=0.58786 α=7.5227 β=30.553	**Distribution of Johnson SB** γ=0.04165 δ=0.97989 λ=64.239 ξ=24.335	**Distribution of Johnson SB** γ=-0.31689 δ=1.5719 λ=121.18 ξ=-2.1521	**Distribution of Johnson SB** γ=-0.16275 δ=1.3946 λ=90.069 ξ=8.3141
Daily Pageviews per Visitor	**Distribution of Dagum** k=0,53738 α=5,2908 β=5,9679	**Distribution of Dagum** k=0,44352 α=4,2354 β=5,6958	**Distribution of Dagum** k=0.64766 α=5.0276 β=5.0463	**Distribution of Dagum** k=0,53738 α=5,2908 β=5,9679	**Distribution of Log-Pearson (3P)** α=22.813 β=0.1101 γ=-1.747	**Distribution of Burr** k=0.25491 α=7.0774 β=1.5894

Kolmogorov-Smirnov, Chi-square Suitability, and Anderson Darling test results were examined to determine the suitability among the distribution models of the series examined within the scope of e-commerce and news sites. In line with the results given in **Table.4**, it has been seen that the distribution type does not change in the duration of stay in e-commerce sites, and it shows the Dagum distribution feature, especially used in finance and economy fields. At the same time, it was observed that the distribution type of the page view number variable in e-commerce sites was Dagum again and did not change in the changing date ranges. In the news sites, it was observed that only the distribution type of the Bounce Rate variable did not change during the relevant periods and the Johnson SB distribution. This distribution is a four-parameter distribution unlike other distribution types, but it is a successful type in different modeling types, mainly because its structure is quite flexible. The distributions obtained by comparing the theoretical distributions of the indicators with the RMKDL method and their parameters are given in **Table 5**.

Table 5. Analyzed Statistical Distribution Results of Variables

Category / *Variable*	E-Commerce Sites			News Sites		
	August 2018	January 2019	March 2020	August 2018	January 2019	March 2020
Daily Time on Site	**Distribution of Dagum** k=0.80507 α=3.4082 β=0.21975	**Distribution of Dagum** k=0.29954 α=4.2096 β=0.29371	**Distribution of Dagum** k=0.70108 α=3.5011 β=0.19926	**Distribution of Johnson SB Dağılımı** γ=1.8858 δ=0.94562 λ=1.2767 ξ=0.06057	**Distribution of Johnson SB Dağılımı** γ=2.8448 δ=1.6011 λ=1.6116 ξ=-0.07089	**Distribution of Johnson SB Dağılımı** γ=0.86514 δ=0.71378 λ=0.35358 ξ=0.0652
Bounce Rate	**Distribution of Dagum** k=1.2927 α=5.5117 β=29.179	**Distribution of Dagum** k=4.4936 α=2.9956 β=20.36	**Distribution of Dagum** k=1.6206 α=5.2076 β=30.352	**Distribution of Johnson SB Dağılımı** γ=0.04165 δ=0.97989 λ=64.239 ξ=24.335	**Distribution of Johnson SB Dağılımı** γ=-0.31689 δ=1.5719 λ=121.18 ξ=-2.1521	**Distribution of Johnson SB Dağılımı** γ=-0.16275 δ=1.3946 λ=90.069 ξ=8.3141
Daily Pageviews per Visitor	**Distribution of Dagum** k=0.53738 α=5.2908 β=5.9679	**Distribution of Dagum** k=0.44352 α=4.2354 β=5.6958	**Distribution of Dagum** k=0.64766 α=5.0276 β=5.0463	**Distribution of Johnson SB Dağılımı** γ=2.8728 δ=0.9391 λ=24.458 ξ=1.3198	**Distribution of Johnson SB** γ=1.3938 δ=0.71964 λ=7.9776 ξ=0.96657	**Distribution of Johnson SB Dağılımı** γ=1.8231 δ=0.58808 λ=16.194 ξ=1.4997

Information about the properties of the distributions whose parameters are given in **Table 5** is given in **Table 6**.

Table 6. Statistical Distributions

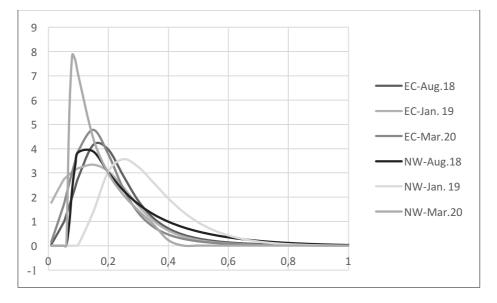| Characteristic Distribution	Parameter	Probability Function
Dagum	k: continuous form parameter **(k>0)** α: continuous form parameter **(α>0)** β: continuous scale parameter **(β>0)** γ: continuous local parameter **($\gamma \equiv 0$ Returns the Dagum distribution with three parameters)**	$(\gamma \leq x < +\infty)$ $(0 \leq x < +\infty)$ $$f(x) = \frac{\alpha k \left(\frac{x}{\beta}\right)^{\alpha k - 1}}{\beta \left(1 + \left(\frac{x}{\beta}\right)^{\alpha}\right)^{k+1}} \rightarrow 3P$$ $$f(x) = \frac{\alpha k \left(\frac{x-\gamma}{\beta}\right)^{\alpha k - 1}}{\beta \left(1 + \left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)^{k+1}} \rightarrow 4P$$
Johnson SB	γ: continuous form parameter δ: continuous form parameter **(δ>0)** λ: continuous scale parameter **(λ>0)** ξ: continuous local parameter	$(\xi \leq x \leq \xi + \lambda)$ $$f(x) = \frac{\delta}{\lambda \sqrt{2\pi} z (1-z)} exp\left(-\frac{1}{2}\left(+\delta ln\left(\frac{z}{1-z}\right)\right)^2\right)$$

Source: Compiled by authors.

In Table 6, **"3P"** refers to the distribution with three parameters and **"4P"** with four parameters. The change in the number of parameters is adapted by changing the number of distribution parameters so that the lower and upper boundaries of the distribution do not approach -∞ because the variables used in this study have no negative value.

In accordance with the results obtained, distributions of variables according to the periods examined for e-commerce and news sites were obtained by obtaining probability density functions, and their graphics were presented. The graphs created to examine the similarities and differences between the distributions of variables that have the same distribution over time are given, respectively.

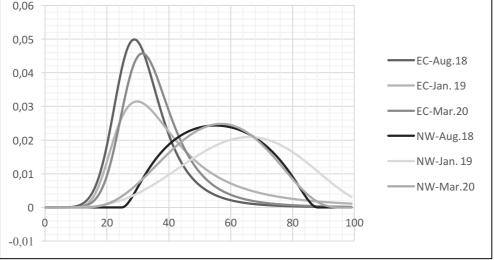**Figure 1.** Distribution of Time on Site on E-Commerce and News Sites

(***Note:*** "EC" means of E-Commerce sites and "NW" News sites.)

As can be seen in **Figure 1**, the homogeneity in the duration of stay in news sites deteriorated over time and showed more difference. After August 2018, it was observed that the weight shifted towards lower periods in January 2019. By March 2020, the weight shifted towards high times. The deterioration in the homogeneity of e-commerce residence time is similar to that in news sites. When it came to January 2019, after August 2018, it was observed that the weight shifted mostly to low periods. In March 2020, there is a shift towards long periods. The primary hypothesis is rejected because the calculated RMKDL = 3.35> 0 belonging to news sites. A meaningful change is observed in the distribution of residence time in news sites. In e-commerce sites, the primary hypothesis cannot be rejected because RMKDL = -5.11 <0. No significant change was observed in the distribution of residence time in e-commerce sites.

Figure 2: Bounce Rate Distributions from E-Commerce and News Sites

As seen in **Figure 2**, Bounce Ratenda homogeneity status has deteriorated over time by showing more differences from news sites. When it comes from August 2018 to January 2019, it is seen that Bounce Ratenda is rising from related sites. In March 2020, it was determined that this rate decreased. As the value of the "ξ" position parameter, which is one of the parameters of the Johnson SB distribution, has increased compared to the old dates, the bounce rates also increase. When looking at e-commerce sites, there is a situation similar to the change of homogeneity in news sites. As the studied date range progresses, there is an increase in bounce rates from related sites. The basic hypothesis is rejected because the calculated RMKDL = 2.22> 0 belonging to news sites. There is a significant change in the distribution of the number of page views on news sites. In e-commerce sites, the basic hypothesis is rejected because RMKDL = 1.62> 0. A significant change is observed in the distribution of the number of page views in e-commerce sites.
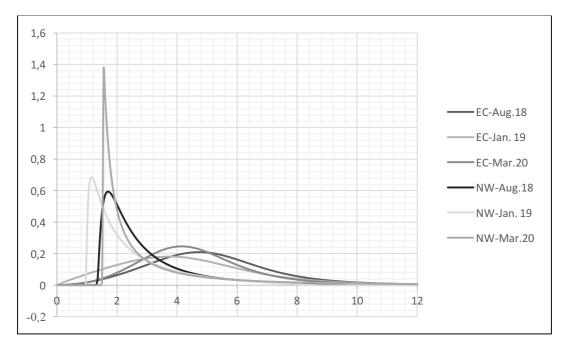
Figure 3. Distribution of Page Views on E-Commerce and News Sites

As seen in **Figure 3**, the homogeneity in the number of page views on news sites deteriorated over time and showed more difference. However, there is a difference in this indicator compared to other indicators. The homogeneity condition deteriorated over time as it shifted to low values, not too high values. It is seen that there is a decrease in the number of pages displayed when coming from August 2018 to January 2019. In March 2020, it was determined that there was much more agglomeration in low periods. E-commerce sites also have a similar situation to news sites. By 2018, by 2020, it was observed that the weight shifted to lower values , and the aggregation was high. The basic hypothesis is rejected because the calculated RMKDL = 1.85>0 belonging to news sites. There is a significant change in the distribution of the number of page views on news sites. In e-commerce sites, the basic hypothesis is rejected because RMKDL = 0.71>0. There is a significant change in the distribution of page views on e-commerce sites.

RESULT

In line with the findings obtained from the websites examined in the study, it is thought that the use of the web environment is more purpose-oriented and increased. The main goal of targeted usage is to observe the increase of Bounce Rate and decrease of page views from related websites. Therefore, it can be concluded that the Bounce Rate and page view count indicators can be analyzed with Dagum and Johnson SB distributions in the simulation and planning models to be made. At the same time, since there is a statistically significant difference in the number of page views on news sites, it can be concluded that it can be examined with the Johnson SB distribution. However, since it is determined that there is no statistically significant difference in the page display indicator in e-commerce sites, this comment does not seem possible. This situation is consistent with the results obtained by examining the distributions of August 2018 and January 2019.

It is possible to say that the distribution results of the relevant website indicators show consistency as a result of following the websites at varying date intervals. In line with the results obtained, it was observed that the website statistics are not stable. In this study, where the changing status in the distributions on the changing dates is examined, the trend structure of the Bounce Rate indicator cannot be determined exactly. However, it was seen that the page view count indicator also had a decreasing trend on changing dates. In addition, there is no seasonal impact on e-commerce and news site bounce rates within the context of relevant dates. However, the number of page views on e-commerce sites shifted to lower numbers in the winter period, including January 2019 and March 2020. As of August 2018, it was determined that more page views were reached in the summer period compared to other winter periods. From this point of view, there is a seasonal effect on the number of page views on the relevant websites, and users search for more fashion-related products in the summer season, etc. it seems possible to make comments. Therefore, it is believed that companies make comparisons according to seasonal period sales figures of e-commerce sites and user indicators in web-based systems. When the change in the number of page views was evaluated in terms of news sites, there was no seasonal effect. However, when the difference between distributions is analyzed, although the decreasing trend continues in March 2020, it has been observed that there is a significant increase in the areas with low page views compared to other dates. In addition, the structure of the Bounce Rate, which had an increasing trend, changed, and the rate decreased in March 2020. This situation is thought to be due to the fact that the websites examined the news sources, and the virus outbreak named COVID-19 was brought to the agenda in our country on this date. The basis of this idea is that the increase in the duration of the page is high, and the Bounce Ratenda decrease is observed from the page. The fact that the users do not display too many pages about the relevant news sites and that they stay on the access page for a longer time supports this idea. The amount of hits and clicks of web pages also depends on the ads and the theory of social availability and social impact. According to the findings obtained as a result of the study, comments about the increase in the duration of the page stay in March 2020 and the decrease in the bounce rate are within the scope of the social impact theory. The distribution type is an indicator of social impact, and the number of pages visited is considered to be a reaction.

It was shown that the program used in the study of statistical distributions in the study could be used in distribution modeling in the field of statistics as well as its use in different disciplines such as medicine, physics, and geography. Since the variables examined within the scope of the study were related, marginal distributions were obtained. From this point, it is possible to obtain conditional distributions in different studies. In addition, it seems possible to use bayesian distribution methods. However, working constraints should not be ignored. The news and e-commerce sites examined are quantitatively limited due to insufficient access to data on the websites targeted to be examined. In addition, the fact that the number of logins per person accessed to the related websites is not prevented from examining the distributions of this data. Besides these limitations, it should be remembered that there may be cultural and social differences due to the study and inference of the websites published in Turkish only. When comparisons between countries are desired, news and shopping follow-up behaviors of societies may change.

Interactive websites are an essential resource for identifying the needs of target consumer groups and developing attitudes in line with these needs. Especially today, when the technological development momentum is high, websites are critical to the brand effectiveness of companies. It

seems possible to get feedback from consumers with website statistics. Thus, mutual communication provides a bond between consumers and the brand. In particular, the developments experienced or experiences in website design and function increase the value of the relevant brand for the consumer. Marketing strategy advantages can be provided through web pages. At the same time, it can be determined how the presentation should be made to the consumers who want to get information about the products and services. According to the study results obtained without ignoring the constraints, it is possible to use the distribution results of the variables in the performance evaluation between the websites, especially in the sales simulations in the private sector and in the evaluation of the performances of the multimedia sales representatives. Due to the changing behavior of internet users, it would be appropriate for researchers to make evaluations for other website indicators using distributions based on constantly updated data.

As a result, the purpose of creating resources with the distributions determined in the planning and simulation models was realized by examining the page-centered data.

REFERENCES

Adeyemi, O. J., Poola, S. I., Atayero, A. A., Afolayan, D. G., Ariyo, M., & Adetiba, E., (2018), "Exploration of Daily Internet Data Traffic Generated In A Smart University Campus". Data in Brief, 20. https://doi.org/10.1016/j.dib.2018.07.039

Arora, A., Krishnan, R., Nandkumar, A., Telang, R., & Yang, Y., (2004), "Impact of Vulnerability Disclosure and Patch Availability - An Empirical Analysis. H. John Heinz III School of Public Policy and Management, 32.

Bromideh, A.A., & Valizadeh, R., (2013), "Discrimination Between Gamma and Log-Normal Distributions By Ratio of Minimized Kullback-Leibler Divergence", Pakistan Journal of Statistics and Operation Research, 9(4). https://doi.org/10.18187/pjsor.v9i4.487

Bucklin, R. E., & Sismeiro, C., (2009), "Click Here for Internet Insight: Advances in Clickstream Data Analysis in Marketing", Journal of Interactive Marketing, 23(1). https://doi.org/10.1016/j.intmar.2008.10.004

Dalessandro, B., Hook, R., Perlich, C., & Provost, F.J., (2012), "Evaluating And Optimizing Online Advertising: Forget The Click, But There Are Good Proxies", Big Data, 3(2). https://doi.org/10.1089/big.2015.0006

Heuer, D., Brettel, M., & Kemper, J., (2015)., "Brand Competition In Fashion Ecommerce", Electronic Commerce Research and Applications, 14(6). https://doi.org/10.1016/j.elerap.2015.07.007

Huberman, B. A., Pirolli, P., Pitkow, J. E., & Lukose, R., (1998), "Strong Regularities in World Wide Web Surfing", Science, 280(5360). https://doi.org/10.1126/science.280.5360.95

Johnson, E. J., Bellman, S., & Lohse, G.L., (2003), "Cognitive Lock-In And The Power Law of Practice", Journal of Marketing, 67(2). https://doi.org/10.1509/jmkg.67.2.62.18615

Levina, O., & vilnai-Yavetz, I., (2015), "E-visibility Maturity Model: A Tool for Assessment And Comparison of Individual Firms And Sets of Firms In E-Business", Electronic Commerce Research and Applications, 14(6), https://doi.org/10.1016/j.elerap.2015.07.004

Nottorf, F., (2014), "Modeling The Clickstream Across Multiple Online Advertising Channels Using A Binary Logit With Bayesian Mixture of Normals", Electronic Commerce Research and Applications, 13(1), https://doi.org/10.1016/j.elerap.2013.07.004