Planning Experimental Studies of Vehicle Movement Patterns under Various Road Conditions

Konstantine Mchedlishvili, Giorgi Kakhadze
Georgian Technical University, M.Kostava st 77, 0175, Tbilisi, Georgia
kakhadze.g@gtu.ge
DOI: https://doi.org/10.52340/building.2025.72.02.11

Abstract The article is dedicated to the experimental study of the regularities of vehicle movement under various road conditions, the determination of the size of adjustable and controllable variables acting on the Driver–Vehicle–Road–Environment (DVRE) system, the definition of the sequence among variable values to ensure the compactness of the experiment and the effectiveness of the results, and the calculation of the required number of trials and measurement accuracy to obtain reliable and convincing outcomes. The article also discusses conducting the experiment using a so-called “randomized design.” For planning the experiments and analyzing the obtained data, the normal distribution law and the so-called “Student’s t-test” are applied.

Key words: experiment, number of trials, normal distribution, data analysis

Introduction

Applied science is based on research conducted through theoretical and, predominantly, experimental methods derived from the achievements of fundamental sciences, resulting in solutions to problems that are directly necessary for practical use. Among these is the main goal of our study — a critical analysis of the existing methods, standards, and regulations for studying the regularities of road traffic, improving outdated approaches, and creating new ones. The application of these improved or newly developed methods in practice will ensure the achievement of maximally effective results with minimal time, material, and financial expenditures.

In the early stages of civilization, engineering relied on the results of experimental research, but the analysis and further development of these results belonged more to the realm of art. Instead of the optimal synthesis of mutually conflicting requirements, the main focus was placed on ensuring the style and technologies appropriate to the era, as well as on excessive strength and durability. As an example, we can mention the multilayer pavement structures of ancient Rome’s main roads during the classical era, whose strength and durability exceed even the requirements needed for today’s automotive loads.

Main Part

Rational planning of an experiment allows us to minimize the duration of the research, avoid major errors, and obtain the maximum amount of useful information with minimal time and labor expenditure. For the effective execution of an experiment, it is necessary to determine in advance the size of the adjustable and controllable variables. To do this, we rely on intuition, experience, and the following two conditions: [1,2,3]

The variation step must be sufficient to reveal a noticeable change in the studied object. For this purpose, we take some baseline value of X, denoted as $X\delta$, and define a step ΔX, which provides the lower and upper bounds of the variation interval — X_{low} and X_{high}.

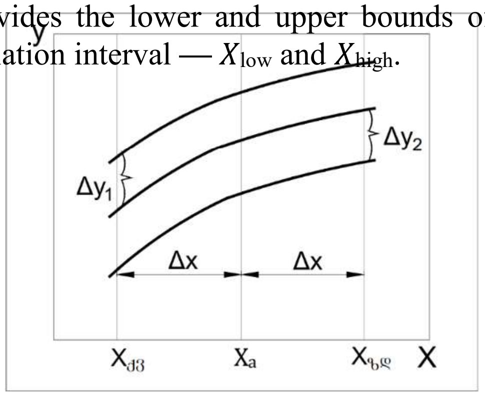


Fig. 1. Scheme for Determining the Variation Step

$$X\delta-\Delta X=X_{\text{ქვ}},\ X\delta+\Delta X=X_{\text{ზდ}}$$

The variation step must be such that the increments of the studied object at the lower and upper bounds cancel each other out:

$$\Delta Y_1 - \Delta Y_2 = 0$$

The variation step must also be small enough to allow the description of the regularities of the studied object's changes within this interval using a linear function. In Fig. 2, it is evident that for linear approximation, the step ΔX is better than ΔX_1.

Such analysis is performed across the entire range of possible values of the variables, and based on this, taking into account the results of preliminary experiments, the required step size is determined.

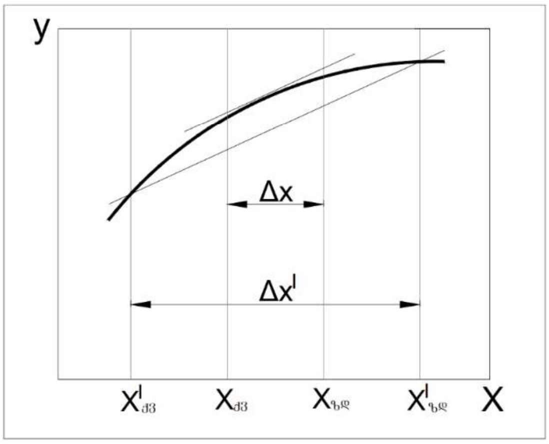


Fig. 2. Selection of the Variation Step Based on the Condition of Linear Approximation

For the experiment to yield effective results and for the experiment itself to remain compact, we must determine in advance the sequence of the variable’s values. The theory of experimentation [3] considers two cases here: The first case is when we start from one extreme value of the variable and gradually move toward the other extreme value — for example, from $X_{\min}$ to $X_{\max}$.

In the second case, the value of the variable is chosen arbitrarily across its entire range of variation. The first case represents an experiment with a so-called “sequential design,” while the second corresponds to an experiment with a “randomized design.” Such a design offers several advantages: often, other factors influencing the process under study are not reliably stabilized — this mainly concerns controllable but non-adjustable factors. In our case, these include ambient temperature, rain, and moderate or strong wind, all of which may vary during the course of the experiment. If, during the same period, we continuously vary the studied variable X in a fixed sequence, then the studied output Y may change not only due to variations in factor X, but also under the influence of other, poorly stabilized factors. However, if we vary X randomly, we eliminate the possibility of mixing the effects of meteorological factors and the influence of X on Y. Given the specifics of our experiments, which are conducted outdoors, we always perform the trials in sunny, windless, and dry weather.

Selecting random values for X also eliminates the influence of the operator’s changing performance due to fatigue or learning, as well as the limitations of the available time for using the measuring instrument. It is evident that randomization allows us to minimize the “noise” caused by uncontrollable factors. However, in our study, where observations are conducted on pedestrian behavior or the time intervals between vehicles, as well as on the intensity and composition of traffic flow, the experiment is carried out using video recording. If video recording is not used, observers are rotated every 15 minutes when traffic intensity is high, and when traffic is low, one observer can reliably work for 45–60 minutes. Based on the above considerations, and given the characteristics of the studied object, for which obtaining objective information is fully possible through a strict sequential arrangement of variables, we therefore use experiments with a “sequential design.”

In our studies, there are often objects that depend on several adjustable and controllable variables. The same object may also be influenced by several uncontrollable factors. In such cases, depending on the number of adjustable and controllable variables, we can conduct two-factor, three-factor, and higher-order experiments. For example, the waiting time of pedestrians at a regulated intersection depends on the duration of the red signal, the width of the roadway, the presence of a median strip, and so on. These are adjustable and controllable variables. In addition, there are uncontrollable factors, such as air temperature, rain intensity, wind strength, and the mentality of pedestrians and drivers.

In cases of such multi-factorial influences, we can conduct experiments using either a classical or a factorial design. In our studies [5,6], we mainly use multi-factorial experiments with a classical design. For example, consider an object R that is a function of the adjustable and controllable variables X, Y, and Z: $R = f\ (X,Y,Z)$. In the classical design, all these variables except one are kept constant, i.e., stabilized. The remaining variable is varied across its entire possible range. Then we vary the second variable while keeping the others stabilized at constant levels, and so on. The influence of uncontrollable factors is taken into account through randomization. For instance, consider a two-factor experiment where each factor is varied at five levels. We investigate the maximum speed achievable under stability conditions on a regulated intersection during left turns with a small radius and its dependence on the turning angle $\alpha°$ and the trajectory radius R. The sequence scheme of changes for both factors, each varied at five levels, can be represented as shown in Fig. 3.

	50	40	70	60	80
200			+		
175			+		
150	+	+	+	+	+
100			+		
50			+		

Fig. 3. Randomization matrix of variables at 5 levels

To randomize factors such as vehicle type, driver characteristics, pavement parameters, curve superelevation, etc., it is advisable to conduct the experiment according to the following variable pairings: 200 m – 70°, 75 m – 70°, 150 m – maximum, and so on. Based on our experience, during a classical experiment it is not necessary for the experimental plan to be fully balanced. That is, variables can be varied at different levels. For example, we know that the turning radius is more important than the turning angle of the trajectory. Therefore, we can vary the radius at 8 levels and the turning angle at 5 levels. If a more complex type of mathematical relationship is expected, such as trigonometric, logarithmic, etc., the matrix can be filled more fully, and the experiment can be conducted with a greater nu mber of combinations of x and y variables. To obtain reliable results from the trials, it is necessary to determine in advance the required number of trials and the accuracy of measurements. Based on our practical experience [5,6], it is advisable to evaluate the reliability of the results using the so-called reliability probability (α). The limits within which the studied quantity is contained with the corresponding reliability probability are called reliable limits, denoted as a_1 and a_2 (Figure 4). The interval between the reliable limits is called the reliable interval $\mu\alpha$. Using preliminary data from observations of vehicle movement modes, we determine the magnitude of the studied parameter’s dispersion R.

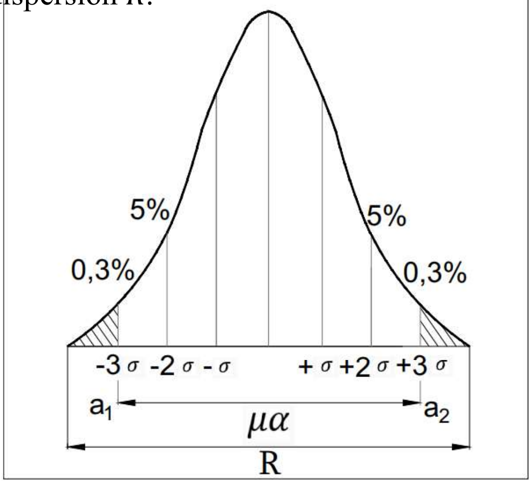
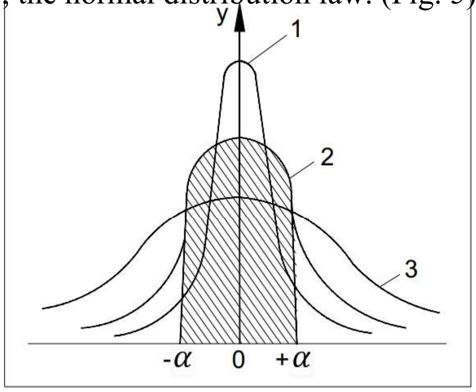


Figure 4. Reliable interval and its limits according to a predetermined reliability probability

The distribution of the studied parameter can follow different regularities. To analyze it, histograms and variation series of the studied quantity are constructed, and then, by fitting the appropriate theoretical distribution curve, statistical investigation is carried out using well-known methods. The reliability of the obtained hypothesis regarding the type of distribution is tested using the Pearson criterion, $P(x)^2$ P=0,05, corresponding to a 95% reliability probability. Analysis of the obtained data and a review of specialized literature indicate that the distribution of speeds predominantly follows the Gaussian, i.e., the normal distribution law. (Fig. 5)



1) σ=0.5; 2) σ=1.0; 3) σ=2.0

Fig. 5. Shape of the normal distribution curve

The distribution of time intervals, on the other hand, more often follows the Poisson or Pearson Type III distribution (Figure 6).

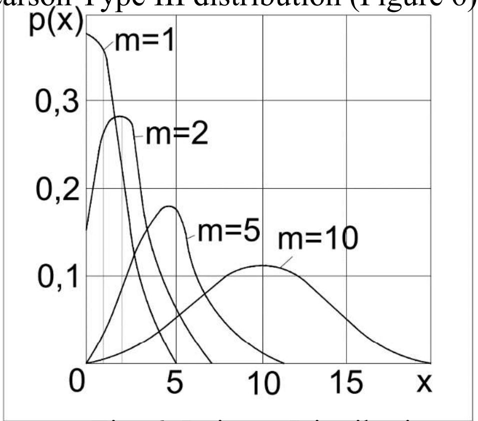


Fig. 6. Poisson Distribution

m-mathematical expectation, i.e., the weighted average value. Under conditions of relatively low intensity, with traffic load levels in the range Z=0.3–0.6, according to our data, the Poisson distribution, in terms of its spread characteristics, approximates the normal distribution. In high-intensity flows, when Z= 0.7–0.85, it corresponds to Pearson’s Type III distribution. Since most of our experiments were conducted under traffic loads with Z≤0.6, for the design of experiments and analysis of data, we use the “normal” distribution law. In such cases, to determine reliable limits, we use the number of standard deviations that should be measured to the right and left of the distribution center so that the probability of the distribution within the resulting interval is α [2, 3].

Experience from studies on speed distributions under certain road conditions shows that most values are located within the range X±3σ, which corresponds to a confidence probability of 99.7%. This indicates that we can take the spread magnitude as R=6σ. Accordingly, if the spread magnitude is known from preliminary experiments, the probable value of the standard deviation is σ = R/6.

In road studies, results are traditionally accepted with 85% and 95% confidence, which shows that the existence of the studied quantity within any predetermined interval can be confirmed only with a 15% or 5% allowable error of the total number of cases. To determine a reliable and convincing value of the studied quantity, and subsequently the nature of its distribution, it is necessary to establish the required number of experiments using the following expression:

$$n=\frac{t^2 * \sigma^2}{\Delta^2}$$

Where n is the number of trials, Δ is the measurement accuracy, and t is the confidence coefficient, i.e., the function of the reliability probability according to the so-called “Student’s criterion” [2, 3]. Its value is chosen according to the required confidence level based on the values presented in the table below.

Table N1

Confidence α%	85	90	95	98	99	99,7
t	1,25	1,7	2,0	2,4	2,6	3,3

Below is the calculation of the minimum number of trials required for analyzing speeds recorded on a specific road section.

Variant I. The speed range on the given road section is V=40–76km/h, i.e., the spread R=36 km/h. Then the standard deviation is $\sigma = \frac{R}{6}$=6km/h σ=6 R=6km/h. We select a confidence level of 85%. From the table, α=85% corresponds to t=1.25. If the measurement accuracy is Δ = 1 km/h, then the minimum number of trials required is:

$$n=\frac{1{,}25^2 * 5^2}{1^2} = 1{,}56*36=56 \text{ trials}$$

For analyzing the distribution of time spent on the section, we perform the following calculation.

Variant I. The time range is t=5–35 min, so the spread is R=30 min. Choosing a confidence level of 85% gives t=1.25, and measurement accuracy Δ=1 min. Then the minimum number of trials required is: $n=\frac{1{,}25^2*5^2}{1^2}$=1,56*25=39 trials.

If we choose a confidence level of 95%, the required minimum number of trials increases because the Student's criterion becomes t=2. In this case, for speed distribution analysis: $n_v=\frac{2{,}0^2*6^2}{1^2}$ =144 and for time distribution analysis: $n_t=\frac{2^2*5^2}{1^2}$ =100 trials.

If we reduce the measurement accuracy to Δ=2km/h and Δ=2min, while keeping other parameters the same, the minimum required number of trials decreases accordingly: n_v=36 trials, n_t=25 trials. It should also be noted that for our conducted experiments, the calculation values were taken according to Variant I, i.e., α=85%, t=1.25, Δ=1km/h, and Δ=1 min.

Conclusions

When experimentally studying the patterns of vehicle movement under various road conditions, it is essential to determine the size of the set of controllable and measurable variables acting on the driver–vehicle–road–environment (DVRE) system. The range of variation of the variables should be sufficient to reveal noticeable changes in the studied object. At the same time, the variation should be small enough that the regularities of the changes in the studied object can be described using a linear function. To ensure the compactness of the experiment and the efficiency of the results, the sequence of variable values should be determined in advance. Two approaches are possible here: Sequential Plan Experiment – In this approach, the experiment starts from one extreme value of the variable and proceeds to the other extreme. This method is suitable for objects that depend on several controllable and measurable variables. Depending on the number of such variables, two-factor, three-factor, or higher-order experiments can be conducted, in which one factor is varied while the others are reliably stabilized. Subsequently, the second factor is varied while the remaining factors remain stabilized, and so on. This is the classical sequential plan for multi-factor experiments. Randomized Plan Experiment – When it is impossible to reliably stabilize some of the factors acting on the studied object because they are uncontrollable, it is advisable to select variable values randomly across the entire range. Random variation of the variable protects the experiment from the influence of poorly stabilized factors. For reliable and convincing results, it is necessary to determine in advance the required number of trials and measurement accuracy. Conducting experimental studies on road factors is advisable when the traffic load level does not exceed 0.6, i.e., when a capacity reserve of at least 40% is ensured. Otherwise, high traffic density prevents the influence of road factors on vehicle movement modes from being observed. When the load level is below 0.6, it is possible to use the so-called "normal" distribution law for planning experiments and analyzing the obtained data. The confidence level of the trials is chosen using the Student's criterion. A level of 85% or 95% ensures that errors occur in only 15% or 5% of cases, respectively. Using this criterion, the required number of trials is determined. If necessary, confidence levels of 99% or 99.7% can also be ensured, but this significantly increases the number of trials. Based on our many years of practical experience in conducting experiments, it is advisable in road studies to use confidence levels of 85% and 95%.

References

1. M. Tsikarishvili. Fundamentals of Engineering and Scientific Research. Publishing House "Technical University", Tbilisi, 2021, 152 pp.
2. M. Okroshvili. Fundamentals of Scientific Research. Tbilisi, 2009, 131 pp.
3. V.V. Nalimov. Theory of Experiment. "Nauka" Publishing, Moscow, 1991, 205 pp.
4. V. Sidenko, I. Grushko. Fundamentals of Scientific Research. "Vyshcha Shkola" Publishing, Kharkiv, 1997, 198 pp.
5. K. Mchedlishvili, M. Siradze. Some Features of Engineering Calculations for Road Traffic Regulation. Collection of Works of the Georgian Automobile and Road Institute, No. 3, Tbilisi, 2007, pp. 146–149.
6. K. Mchedlishvili, Z. Bogvelashvili, V. Gabiashvili. Investigation of Traffic Flow Density. Collection of Works of the Georgian Automobile and Road Institute, No. 3, Tbilisi, 2007, pp. 61–65.