AUTOMATYKA/ AUTOMATICS • 2015 • Vol. 19 • No. 1

http://dx.doi.org/10.7494/automat.2015.19.1.17

Waldemar Bauer*

New approach in modelling of the patients behavior in Primary Health Care (PHC)

1. Introduction

The purpose of this paper is to show a method of modelling behavior of patients in Primary Health Care (PHC). The reason for that is caused by the fact, that nowadays the methods of modelling PHC systems don't exist. In the paper the author described primary health care systems, stochastic process modelling and presents result of simulation patient number in year for different distribution functions.

2. Primary Health Care system

Primary Health Care (PHC) is a multidimensional and country-to-country diverse part of the health care system, thus it is a real challenge for modelling these systems. The Primary Health Care Activity Monitor in Europe (PHAMEU) study (held in 2008 through 2010) in 31 European countries aimed to develop a universal matrix for international comparisons of PHC systems. The classical quality model by Donabedian (structure-process-outcome) has been filled with PHC features (accessibility, coordination, comprehensiveness, continuity) and quality, effectiveness, governance, financing and workforce measures (see [2]). The international research team used existing datasets, reports, publications as well as grey data and opinions of key-informants to provide innovative international comparisons of PHC systems in Europe. Also it proved the usefulness of the PHC quality assessment model. In 2011 another study – Quality and Costs of Primary Care in Europe (QUALICOPC) has been initiated, using the PHAMEU-based matrix for PHC quality evaluation (see [3]). In Poland, as in other Central and Eastern European countries the methods for reliable monitoring of care provided to patients has not been introduced in primary health care.

* AGH University of Science and Technology, Department of Automatics and Biomedical Engineering, Krakow, Poland

Moreover, the data collected by the National Health Found are not be published. The above situation makes a difficult take steps in family physician practices to improve the quality of care (see [2, 3]).

The world currently uses the following methods of financing PHC (see [13]):

Fee for service – is the pay gap between doctor (service providers, or any other professional employee) on the basis of the number and type of individual benefits which have been awarded to patients. For any service shall be determined a specific price usually including reimbursement of costs and margin. The number of granted benefits of a specific price determines the height of the achieved revenue.

Capitation – a system in which a physician (or other medical worker or trader) receives a fixed amount for each person covered by the care, in order to be able to provide within the time period a specified level of benefits and offer a defined package of medical services. If the operation of this system is to create a list of names of patients under care. Consideration on capitation may cover only the cost of consulting a doctor himself, but often includes other related costs. Even more often in the way benefits are financed through a specific package of services including personnel costs.

Fixed fee (salary) – a typical solution for full-time employees, the dominant employment especially in public structures. Doctors and other medical personnel receive a fixed sum for a predetermined time. The remuneration shall be determined due to a qualifications, a years of work and often in relation to salaries in other sectors of the public. Fixed remuneration (salary) in principle does not depend on the work, as measured, for example. the number of visits, the number of patients treated, the severity of the case, etc. These items may be taken into account in the framework of the so-called. Incentive bonus, which use in the health system, however, is significantly limited.

Fee for the case – the way in which the payment service provider receives a fixed salary for a comprehensive investigation in a particular case, or disease entity. The sum paid for each case is calculated on the basis of the expected procedures established and confirmed in the Protocol of treatment. Salary requires the development of special rank order, uniquely identifying the case and called for him to pay.

The International Classification of Diseases (ICD) is the standard diagnostic tool for epidemiology, health management and clinical purposes. ICD-10 is a medical classification list by the World Health Organization (WHO). It contains codes for diseases, signs and symptoms, abnormal findings, complaints, social circumstances, and external causes of injury or diseases. It is used to monitor the incidence and prevalence of diseases and other health problems, proving a picture of the general health situation of countries and populations. All the described methods of finance primary health care based on the number of patients served at the facility. It follows that an efficient way simulation of PHC will be a simulation number of patients in a time range.

3. Stochastic process modelling

Many stochastic processes used for the modeling of financial marks, biological systems, social systems and other systems in engineering are Markovian. To simulate the process can be used Markov Chain Monte Carlo (MCMC). In statistics MCMC method is a class of algorithms for sampling from a probability distribution. Probability distributions can be found the context of Bayesian data analysis. The goal will be to find parameter values in a probabilistic model that the best explain the data. It is based on known information (a priori). The created mathematical models are then created a posteriori. Such approach guarantees that solution is influenced by the known data, therefore the models are often more accurate than obtained with other methods (see [10, 11]).

In this section, the author will describe basic definitions from stochastic theory, the Markov Chain process, Markov Chain Monte Carlo Simulation and the Metropolis Hasting algorithm.

4. Stochastic process

A stochastic process X^n is a family of random variables indexed by parameter n (this parameter can by associated with time). Formally, a stochastic process for probability space (Ω, F, P) and measurable space (S, Σ). The sample space Ω is a set of outcomes, where an outcome is the result of a single execution of a stochastic model. F is set of all events in the model and P is the probability measure. The probability measure is the function returning an event's probability (P: $F \to [0, 1]$) this can by. The S-valued stochastic process is a collection of S-valued random variables on Ω. The stochastic process is indexed by a totally ordered set T. That is, a stochastic process X is the collection:

$$\left\{X^n\!: n \in T\right\}$$

5. The Markov Chain process

Markov Chain is a stochastic process where we transition from one state to another using a sequential procedure. We start Markov Chain in state $x^{(0)}$, and use a transition function $p\left(x^{(n)} \mid x^{(n-1)}\right)$, to determine the next state, conditional on the last state. We can say a stochastic process $\left\{n^{(n)}\!: n \geq 0\right\}$ is a Markov Chain if for all times $n \geq 0$ and all state $i_0, i_1, \ldots, i_n, j \in S$ (state space):

$$P\left(x^{\{n\}} = j \mid x^{(n-1)} = i_{(n-1)}, x^{(n-2)} = i_{(n-2)}, \ldots, x^{(0)} = i_0\right) =$$

$$= P\left(x^{(n)} = j \mid x^{(n-1)} = i_{(n-1)}\right) = P_{i,j}$$

$P_{i,j}$ denotes the probability that the chain moves from state $x^{(n-1)}$ to state $x^{(n)}$. This value is referred to as a one-step transition probability. The square matrix $\boldsymbol{P} = P(i, j) \in S$, is called the one-step transition matrix and each row sum to one (see [8, 9]).

6. Probability density functions

The probability density function (pdf or p.d.f) of continuous random variable X with support S is an integrable function $f(x)$ satisfying the following:

- $f(x)$ is positive everywhere in the support S, that is $f(x) > 0$ for all x in S,
- the area under the curve $f(x)$ in the support S is 1, that is:

  $$\int_S f(x)dx = 1,$$

- if $f(x)$ is the probability density function of x, then the probability that x belongs to A, true is equality:

  $$P(X \in A) = \int_A f(x)dx.$$

A continuous random variable takes on an uncountably infinite number of possible values. For a discrete random variable X that takes on a finite or countably infinite number of possible values, we determined $P(X = x)$ for all of the possible values of X, and called it the probability mass function. For continuous random variables, as we shall soon see, the probability that X takes on any particular value x is 0. That is, finding $P(X = x)$ for a continuous random variable X is not going to work. Instead, we'll need to find the probability that X falls in some interval (a, b), that is, we'll need to find $P(a < X < b)$. We'll do that using a probability density function.

Example

Let X be a continues random variable whose probability density function is:

$$f(x) = 2x^2$$

For $x \in [0, 1]$. Let's respond to questions:

1. Do $f(x)$ is a valid probability density function?
   - $f(x)$ is positive for all $x \in [0, 1]$,
   - $\int_0^1 f(x)dx = x^4\Big|_0^1 = 1.$

   Answer: Because $f(x)$ is positive for all $x \in [0, 1]$ and $\int_0^1 f(x)dx = 1$ that is valid probability density function.

2. What is the probability that X falls between $\frac{1}{2}$ and 1? That is, what is $P\left(\frac{1}{2} < X < 1\right)$?

$$P\left(\frac{1}{2} < X < 1\right) = \int_{1/2}^{1} 4x^3 dx = x^4 \Big|_{1/2}^{1} = 1 - \frac{1}{16} = \frac{15}{16}$$

Answer: Probability that X falls between $\frac{1}{2}$ and 1 is equal $\frac{7}{12}$.

Markov Chain Monte Carlo simulation

Markov Chains are relatively easy to simulate from, they can be used to sample from an a priori unknown and probability distribution. Monte Carlo sampling allows one to estimate various characteristics of a distribution such as the mean, variance, kurtosis, or any other statistic of interest to a researcher. Markov Chains involve a stochastic sequential process where we can sample states from some stationary distribution.

The Markov Chain Monte Carlo (MCMC) method is a general simulation method for sampling from posterior distributions and computing posterior quantities of interest. MCMC methods sample successively from a target distribution. Each sample depends on the previous one, hence the notion of the Markov Chain. A Markov Chain is a sequence of random variables, θ^1, θ^2, … for which the random variable θ^t depends on all previous θ s only through its immediate predecessor.

Monte Carlo, as in Monte Carlo integration, is mainly used to approximate an expectation by using the Markov Chain samples. In the simplest version

$$\int_s g(\theta) p(\theta) d\theta \cong \frac{1}{n} \sum_{t=1}^{n} g\left(\theta^t\right)$$

Where g is a function of interest and θ^t are samples from $p(\theta)$ on its support S. This approximates the expected value of $g(\theta)$. The Markov Chain method has been quite successful in modern Bayesian computing. The simplest Bayesian models recognize the analytical forms of the posterior distributions and summarize inferences directly. In moderately complex models, posterior densities are too difficult to work with directly. With the MCMC method, it is possible to generate samples from an arbitrary posterior density $p(\theta|y)$ and to use these samples to approximate expectations of quantities of interest (see [4, 5]).

7. Metropolis-Hasting algorithm

To illustrate the work of all MCMC methods the Metropolis-Hastings method has been described.

Suppose our goal is to sample from the target density $p(\theta)$. The Metropolis-Hastings method creates a Markov Chain that produces a sequences of state:

$$\theta^{(0)} \to \theta^{(1)} \to \cdots \theta^{(i)} \to \cdots$$

where $\theta^{(t)}$ is a state at iteration i. The samples from the chain, after burning, reflect samples from the target distribution $p(\theta)$. In this algorithm, we initialize the first state from random value. We then use a proposal distribution $p\left(\theta^{(n)} \mid \theta^{(n-1)}\right)$ to generate a new candidate state θ^*, that is conditional on the previous state. The proposal distribution is chosen by the research and good choices for the distribution depend on the problem. To the choose proposed distribution we can use e.g.: maximum entropy, nuclear estimators and transformation groups.

The next step is to either accepted or reject proposal state. The probability of accepting the sate θ^* is:

$$\alpha = \min\left(1, \frac{p\left(\theta^*\right) q\left(\theta^{(i-1)} \middle| \theta^*\right)}{p\left(\theta^{(i-1)}\right) q\left(\theta^* \middle| \theta^{(i-1)}\right)}\right).$$

To make a decision on whether to actually accept or reject the proposed state, we generate a uniform deviate u. If $u \geq \alpha$, the proposal is accepted and the next state value is equal θ^*, else we reject the proposal and next state value is equal to the old state value. We continue generating new proposals conditional on the current state of the method, and either accept or reject the proposals. This procedure continues until the sample reaches convergence. At this point, samples $\theta^{(i)}$ the samples from the target distribution $p(\theta)$.

This can be converted into an algorithm as follows:

1) generate initial value of u, and set $\theta^{(0)}$ and $i = 0$;
2) set max iteration number N;
3) repeat:
   a) $i+ = 1$,
   b) generate proposal θ^* from $p\left(\theta^{(i)} \mid \theta^{(i-1)}\right)$,
   c) calculate the accept probability: $\alpha = \min\left(1, \frac{p\left(\theta^*\right) q\left(\theta^{(i-1)} \middle| \theta^*\right)}{p\left(\theta^{(i-1)}\right) q\left(\theta^* \middle| \theta^{(i-1)}\right)}\right)$,
   d) generate u from a uniform (0, 1) distribution,
   e) if $u \geq \alpha$, accept new state and $\theta^{(i)} = \theta^*$, else set $\theta^{(i)} = \theta^{(i-1)}$;
4) until $i = N$.

The fact that asymmetric proposal distributions can be used allows the Metropolis--Hastings procedure to sample from target distributions that are defined on a limited range (see [6, 7]). With bounded variables, care should be taken in constructing a suitable proposal distribution. Therefore, the sampler will move towards the regions of the state space where the target function has high density. However, note that if the new proposal is less likely than the current state, it is still possible to accept this "worse" proposal and move toward it. This process of always accepting a "good" proposal, and occasionally accepting a "bad proposal insures that the sampler explores the whole state space, and samples from all parts of a distribution (including the tails). In the numerical experiments all algorithms have been prepared by the author.

8. Numeric experiments

Experiment 1

In order to depict the described simulation methods and demonstrate the actual data, it will be discussed how to simulate the amount of visits to the Cabinet at the PHC for the age group of children 1 year based on data from 2009. In 2009, in the age group was 622 visits within 215 days. The average value of the day is advice about 2.44, while standard deviation is about 1.7. The maximum number of visits in a single day to 18 and the minimum 1. As you can see in Figure 1 the probability distribution for this sample can be estimated using the normal distribution, given by the formula:

$$p(x,\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Where μ is the expected value and standard deviation σ.

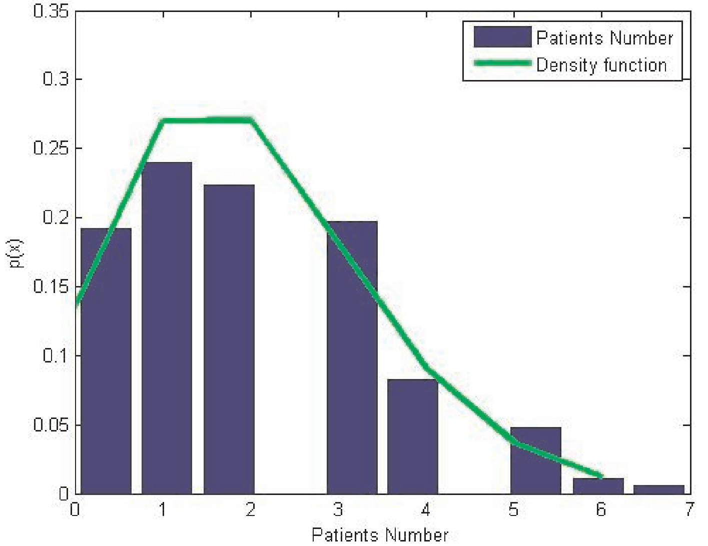

Fig. 1. Real data histogram and normal density function

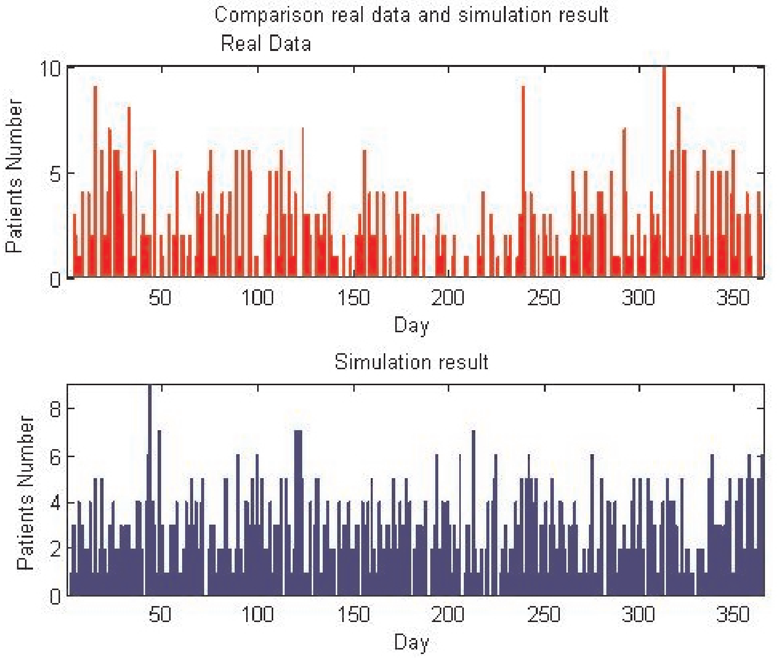

Fig. 2. Real data and simulation results for the normal density function

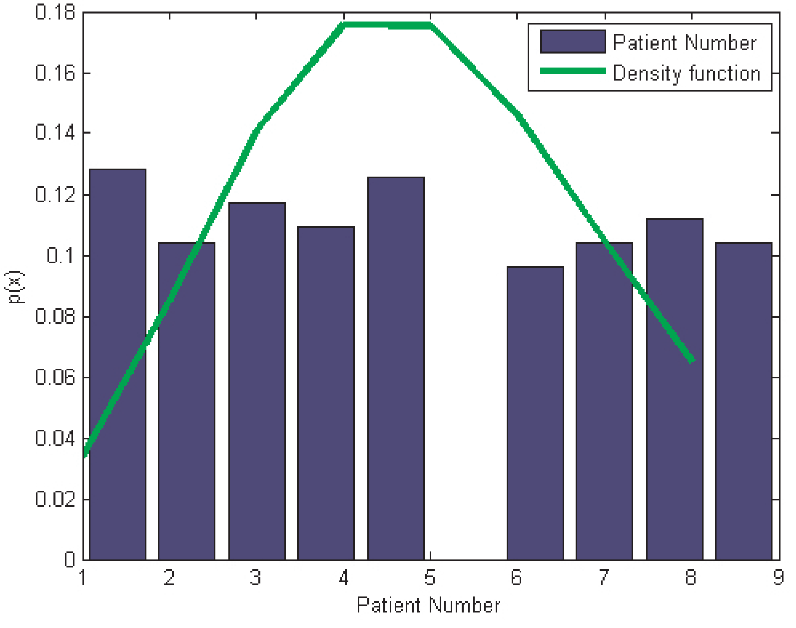

Fig. 3. Simulation data histogram and normal density function

Let's say that on the basis of this information, we want to carry out simulations for the same clinic for the year again, except that it will be open 365 days a year. Therefore, the function of acceptance State the Metropolias-Hastings will take the form of:

$$\alpha = \min\left(1, \frac{p\left(\theta^{*}, \mu, \sigma\right)}{p\left(\theta^{(i-1)}, \mu, \sigma\right)}\right).$$

Simulations with the given conditions, gave the results as shown in Figure 2. In addition, it should be noted that for a specified period simulation of 365 days, the algorithm generated in the United States where the number of visits per day does not exceed 18. In total, in the simulation have been 893 visits or expected value was approximately 2.44 and the standard deviation of about 1.7. This means that have been preserved all the properties of the simulated what confirms the Figure 3.

Experiment 2

For data from experiment 1 suppose the probability distribution can be described using the Poisson distribution:

$$p(x, \mu) = \frac{\mu^{x} e^{-\mu}}{x!}.$$

Where μ is the expected value. This example illustrates the Figure 4.

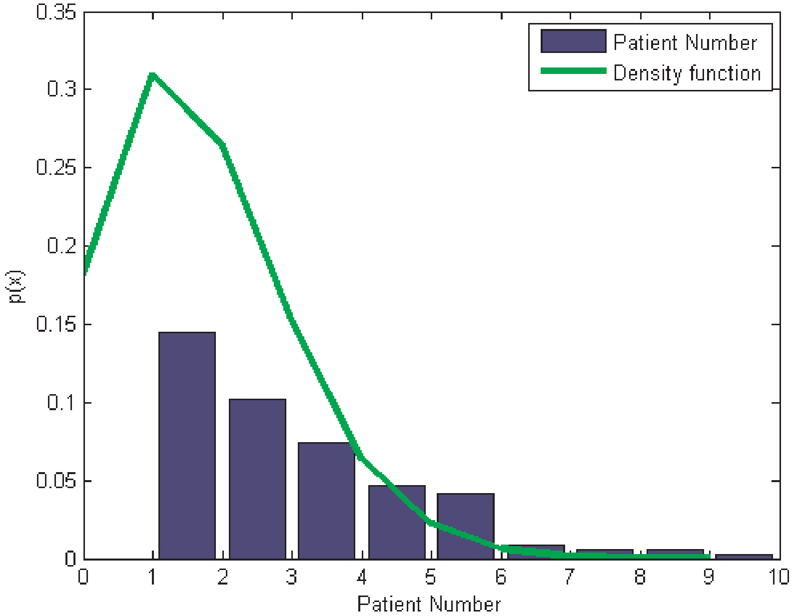

Fig. 4. Real data histogram and Poisson density function

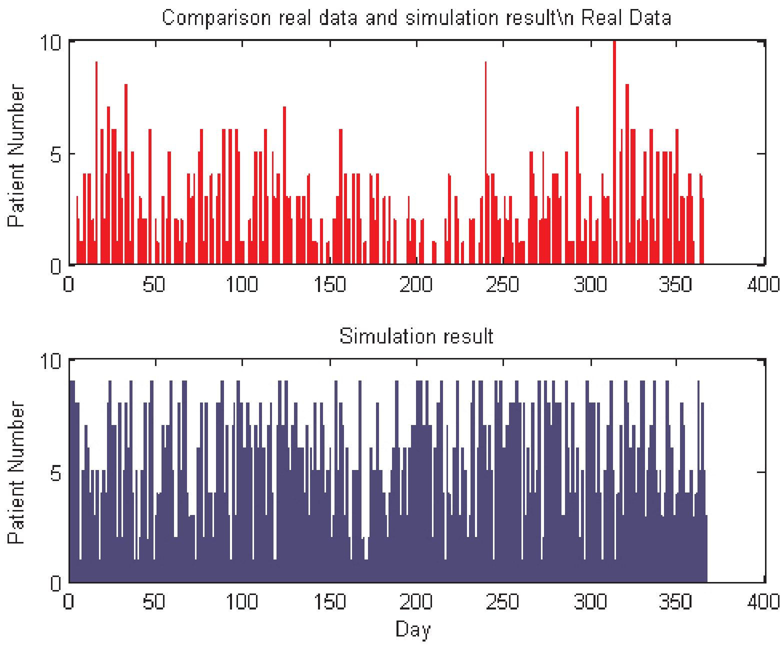

Fig. 5. Real data and simulation result for the Poisson density function

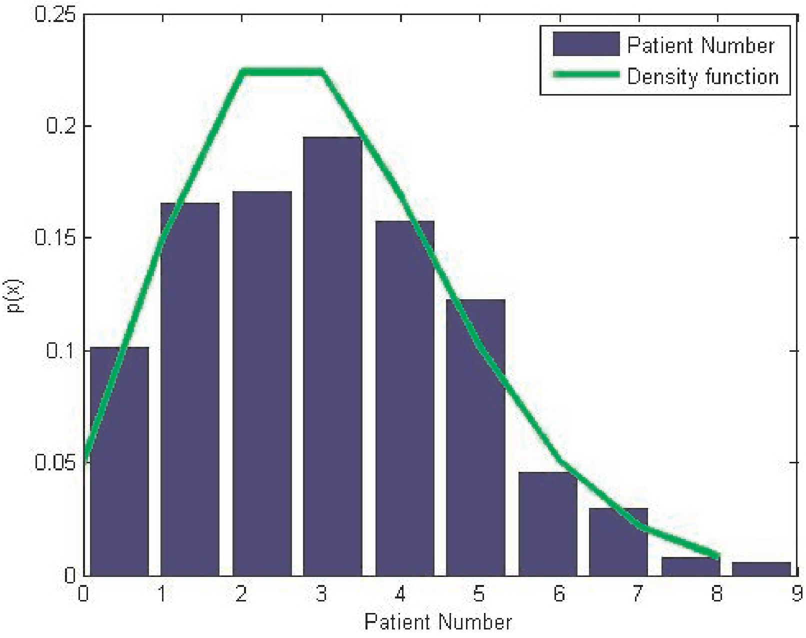

Fig. 6. Simulation data histogram and the Poisson density function

As it's easy to see that the results of the simulation for the Poisson distribution (Fig. 5 and 6) differ significantly from results for the normal distribution (Fig. 2). It should be noted, however, that also in this case the parameter describing the distribution (the expected value) has not changed as a result of the simulation. This example illustrates the behavior of the simulation in the case of proper selection of the distribution for this phenomenon.

9. Conclusion and future work

The main objective of future work is to create a model of the patients behavior for Primary Health Care (PHC) unit. This issue can be divided into subtasks: quantitative analysis of acquired data, determination of the patients behavior in PHC model, design and creation of the simulator. Quantitative analysis based on provided data. Statistical analysis of the acquired data will be made using Bayesian statistics methods, in particular the methods of estimating probability distributions, e.g.: maximum entropy, nuclear estimators and transformation groups. As a result of these analyzes, it is planned to draw up a mathematical model describing the work of PHC units based on the Monte Carlo method using the Markov Chain.

References

[1] Brochu E., Brochu T., de Freitas N., *A Bayesian Interactive Optimization Approach to Procedural Animation Design*, Symposium on Computer Animation, 2010, pp. 103–12.

[2] Kringos D.S., Boerma W.G.W., Bourgueil Y., Cartier T., Hasvold T., Hutchinson A., Lember M., Oleszczyk M., Pavlic D.R., Svab I., Tedeschi P., Wilson A., Windak A., Dedeu T., Wilm S., *The European primary care monitor: structure, process and outcome indicators*, BMC Family Practice 2010, Vol. 11, pp. 81.

[3] Schäfer W.L., Boerma W.G.W., Kringos D.S., De Ryck E., Greß S., Heinemann S., Murante A.M., Rotar-Pavlic D., Schellevis F.G., Seghieri C., Van den Berg M.J., Westert G.P., Willems S., Groenewegen P.P., *Measures of quality, costs and equity in primary health care instruments developed to analyse and compare primary care in 35 countries*, Quality in Primary Care 2013, Vol. 21, pp. 67–79.

[4] Jaynes E.T., *Probability Theory: The Logic of Science*, Cambridge University Press 2003.

[5] Sivia D., Skilling J., *Data Analysis: A Bayesian Tutorial*, Oxford University Press, Oxford 2006.

[6] Lizotte D., Wang T., Bowling M., Schuurmans D., *Automatic Gait Optimization with Gaussian Process Regression*, 2007, pp. 944–949.

[7] Martinez-Cantin R., de Freitas N., Brochu E., Castellanos J., Doucet A., *A Bayesian Exploration-Exploitation Approach for Optimal Online Sensing and Planning with a Visually Guided Mobile Robot*, Autonomous Robots 2009, Vol. 27, No. 3, pp. 93–103.

[8] Srinivas N., Krause A., Kakade S.M., Seeger M.W., *Information-theoretic regret bounds for Gaussian process optimization in the bandit setting*, IEEE Transactions on Information Theory 2012, Vol. 58, No. 5, pp. 3250–3265.

[9] Garnett R., Osborne M.A., Roberts S.J., *Bayesian optimization for sensor set selection*, International Conference on Information Processing in Sensor Networks – IPSN 2010, pp. 209–219.

[10] Bergstra J., Yamins D., Cox D.D., *Hyperopt: A Python Library for optimizing the hyperparameters of machine learning algorithms*, Python in Science Conference, 2013, pp. 13–20.
[11] Thornton C., Hutter F., Hoos H.H., Leyton-Brown K., *Auto-WEKA: Combined Selection and Hyperparameter Optimization of Classification Algorithms*, in: *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, New York 2013, pp. 847–855.
[12] Snoek J., Larochelle H., Adams R., *Practical Bayesian Optimization of Machine Learning Algorithms*, Neural Information Processing Systems Conference, 2012, pp. 1–12.
[13] Baranowski J., Windak A., *Optymalizacja polskiego systemu finansowania podstawowej opieki zdrowotnej*, Technical Report, Ernst & Young Usługi Finansowe Audyt sp. z o.o. Polska sp.k., Warszawa 2012.