Akademie věd České republiky
Ústav teorie informace a automatizace, v.v.i.

Academy of Sciences of the Czech Republic
Institute of Information Theory and Automation

RESEARCH REPORT

Evženie Uglickich and Ivan Nagy

Recursive mixture estimation with univariate multimodal Poisson variable

2394 May 2022

Project Arrowhead Tools
ECSEL 826452 and MSMT 8A19009

ÚTIA AV ČR, P.O.Box 18, 182 08 Prague, Czech Republic
Tel: (+420)266052422, Fax: (+420)286890378, Url: http://www.utia.cas.cz, E-mail: utia@utia.cas.cz

Contents

1 Introduction

Analysis of count variables described by the Poisson distribution is required in many application fields [1]. Examples of the count variables observed per a time unit can be, e.g., number of customers [2, 3], passengers [4], road accidents [5], Internet traffic packet arrivals [6], bankruptcies [7], virus attacks [8], etc.

If the behavior of such a variable exhibits a multimodal character, the problem of clustering and classification of incoming count data arises. This issue can touch, for instance, detecting clusters of the different behavior of drivers in traffic flow analysis as well as cyclists or pedestrians.

This work focuses on the model-based clustering of Poisson-distributed count data with the help of the recursive Bayesian estimation of the mixture of Poisson components. The parameters of mixtures of Poisson distributions and weights are often estimated using the expectation-maximization (EM) algorithm [9] and other primarily iterative techniques, see, e.g. [10, 11, 12, 13], etc. Here, the Bayesian methodology [14, 15] will be used for the Poisson mixture estimation. The work applies the general approach to various components started in [16] up to Poisson components. The aim of the work is to explain the methodology in details with an illustrative simple example, so that the work is limited to the univariate case and static pointer [14]. The Poisson proximity function (see [17, 16, 18]) will be used directly unlike to the paper [4].

The layout of the work is as follows: Section 2 provides the recursive estimation of a single Poisson model, which serves as the preliminaries to the multimodal case presented in Section 3. Section 3 gives the theoretical scheme of the recursive estimation, the algorithm and program. Section 4 demonstrates results of experiments of clustering simulated and real data. The discussion can be found in Section 4.3. Conclusions in Section 5 close the work.

2 Basic facts on single Poisson model

Let the random variable y be described by the Poisson distribution

$$f\left(y=y_t|\lambda\right)=e^{-\lambda}\frac{\lambda^{y_t}}{y_t!} \tag{1}$$

with the unknown parameter λ, and let $y_t \in \{0,1,\ldots\}$ be a realization of y observed at discrete time instants $t=1,2,\ldots,T$. The realizations y_t of the variable y create the count data set $y(t)=\{y_0,y_1,\ldots,y_t\}$, including the prior knowledge y_0.

For analysis of the count variable it is necessary to estimate the unknown parameter λ. As it is well known, the maximum likelihood estimate of λ is the average of the measured realizations, see, e.g., [19]

$$\hat{\lambda}=\frac{1}{T}\sum_{t=1}^{T}y_t. \tag{2}$$

The recursive estimation of this parameter is based on the use of the Bayes rule, see, e.g., [20],

$$\underbrace{f(\lambda|y(t))}_{\text{posterior}}\propto\underbrace{f\left(y=y_t|\lambda\right)}_{\text{model}}\underbrace{f(\lambda|y(t-1))}_{\text{prior}}, \tag{3}$$

where the probability density function (pdf) $f(\lambda|y(t-1))$ in the right side of the formula (3) is the Gamma prior pdf conjugate to the Poisson model (1), see, for instance, [21]. According to [21], due to the independent realizations of y, the likelihood function for the data measured for time $t=1,2,\ldots,T$ can be written just as a product of the Poisson distributions, i.e.,

$$\mathcal{L}(\lambda)=\prod_{t=1}^{T}f\left(y=y_t|\lambda\right)=\prod_{t=1}^{T}e^{-\lambda}\frac{\lambda^{y_t}}{y_t!}=e^{-T\lambda}\lambda^{\sum_{t=1}^{T}y_t}\frac{1}{y_t!}\propto e^{-T\lambda}\lambda^{\sum_{t=1}^{T}y_t}. \tag{4}$$

The Gamma prior pdf conjugate to the Poisson distribution has the following form [21]:

$$f(\lambda|y(t-1))=\frac{b^a}{\Gamma(a)}\lambda^{a-1}e^{-b\lambda}, \tag{5}$$

which means that substituting (5) and (1) into the Bayes rule (3), the posterior Gamma pdf takes the form

$$f(\lambda|y(t))\propto e^{-T\lambda}\lambda^{\sum_{t=1}^{T}y_t}\lambda^{a-1}e^{-b\lambda}=e^{-\lambda(T+b)}\lambda^{\sum_{t=1}^{T}y_t+a-1}, \tag{6}$$

where $\frac{b^a}{\Gamma(\alpha)}$ and $\frac{1}{y_t!}$ are "hidden" under the denotation $\propto$.

For the recursive estimation of the parameter λ of the single Poisson model (1), the key points are the statistics $\sum_{t=1}^{T} y_t$ and T obtained during the computation of the likelihood function (4). Setting the prior statistics S_0 and κ_0, their update is as follows:

$$\text{sum} \quad S_t = S_{t-1} + y_t, \quad \text{counter} \quad \kappa_t = \kappa_{t-1} + 1, \tag{7}$$

and then the point estimate of the parameter λ at time t is

$$\hat{\lambda}_t = \frac{S_t}{\kappa_t}, \tag{8}$$

which is the result identical to (2).

3 Mixture estimation with Poisson components

In this section, the count Poisson variable y is observed on a multimodal system, whose behavior changes among N_c working regimes. Thus, the task is *to capture the clusters existing in the count data space and classify the data into the detected clusters.*

In this case, the generated variable y should be described by a mixture of N_c Poisson *components* describing each of the observed system regimes. The component is hence the Poisson distribution (1)

$$f(y = y_t|\lambda, c_t = i) = e^{-\lambda_i} \frac{\lambda_i^{y_t}}{y_t!}, \quad i = \{1, 2, \ldots, N_c\}, \tag{9}$$

where the parameter λ is now the collection of all component parameters, i.e., $\lambda \equiv \{\lambda_i\}_{i=1}^{N_c}$, and $\lambda = \lambda_i$, if $c_t = i$. The random discrete variable c_t is the so called *pointer* [14], whose value indicates the active component that generates data at the time instant t.

Switching the active components is described by the following static pointer model [14] $f(c_t = i|\alpha)$ with the categorical distribution

c_t	1	2	$\cdots$	N_c
$f(c_t = i\|\alpha)$	α_1	α_2	$\ldots$	α_{N_c}

(10)

where α_i are the probabilities of the value i of the variable c_t and $\alpha = \{\alpha_i\}_{i=1}^{N_c}$. According to [15], the parameter α is recursively estimated using the Bayes rule with the conjugate prior Dirichlet distribution $f(\alpha|y(t-1))$. Its statistics denoted by $\nu_t = \{\nu_{i;t}\}_{i=1}^{N_c}$ is updated as follows [15]:

$$\nu_{i;t} = \nu_{i;t-1} + 1, \tag{11}$$

starting from random or uniform initial statistics. The normalization of the updated statistics gives the point estimates of α at time t

$$\hat{\alpha}_{i;t} = \frac{\nu_{i;t}}{\sum_{l=1}^{N_c} \nu_{l;t}}. \tag{12}$$

The formulated task is specified as *the recursive estimation of the component parameters λ, the pointer model parameters α and the value of the pointer c_t to denote the active component at time t.* Using the Bayes and the chain rules [20, 22] and according to the Bayesian methodology [14, 15], the mixture estimation algorithm is derived as follows:

$$f(\lambda, \alpha, c_t = i|y(t)) \propto f(y = y_t, \lambda, \alpha, c_t = i|y(t-1))$$

$$= \underbrace{f(y = y_t|\lambda, c_t = i)}_{(9)} \underbrace{f(\lambda|y(t-1))}_{\text{Gamma prior}} \underbrace{f(c_t = i|\alpha)}_{(10)} \underbrace{f(\alpha|y(t-1))}_{\text{Dirichlet prior}}, \tag{13}$$

where, as usual under the adopted methodology, the joint pdf of the unknown variables is decomposed to obtain the product of the used models and mentioned prior pdfs [15]. The component parameters λ and the pointer model parameters α are assumed to be mutually independent as well as the pointer c_t and λ.

To obtain the posterior distribution of the pointer for clustering and classification of the data, the factorized joint distribution (13) is marginalized with the help of the integrals over the entire definition spaces λ^* and α^* as follows [14, 15]:

$$f(c_t = i|y(t)) = \int_{\alpha^*}\int_{\lambda^*} \underbrace{f(y = y_t|\lambda, c_t = i)}_{(9)} \underbrace{f(\lambda|y(t-1))}_{\text{Gamma prior}} \underbrace{f(c_t = i|\alpha)}_{(10)} \underbrace{f(\alpha|y(t-1))}_{\text{Dirichlet prior}} \, d\lambda \, d\alpha. \quad (14)$$

Here, the integral over λ^* provides the approximation, which is called the *proximity* of the actual realization y_t to each of the Poisson components at time t. The definition of the proximity is explained in the details in [17, 16, 18].

The proximity is the value, obtained by substituting the point estimates of the component parameters λ available for the last time instant $t-1$, the current realization y_t into the Poisson component (9), i.e.,

$$m_i = \mathcal{POI}\left(y_t, \hat{\lambda}_{i;t-1}\right), \quad (15)$$

which represents the approximated "closeness" of the data item to the i-th component [17, 16, 18], and where $\hat{\lambda}_{i;t-1}$ is obtained according to (8).

The integral over α^* in (14) serves to obtain the stationary prediction of the pointer [15] via the point estimate of the pointer model parameter α.

The update of statistics of the components as well as the pointer model is performed using the result of (14), which is the posterior distribution of the pointer. It is the *weighting vector* $w_t = [w_{1;t}, \ldots, w_{N_c;t}]$, which contains the probabilities of the activity of the components at time t. Each weight is obtained [15] by

$$\tilde{w}_{i;t} = m_i \, \hat{\alpha}_{i;t-1}, \quad i = \{1, \ldots, N_c\} \quad (16)$$

and then is normalized

$$w_{i;t} = \frac{\tilde{w}_{i;t}}{\sum_{k=1}^{N_c} \tilde{w}_{k;t}} \quad (17)$$

to enter the vector w_t $\forall i = \{1, \ldots, N_c\}$. The computed weights are used in the update of statistics (7) of each component based on the Bayesian mixture estimation theory [14, 15]

$$S_{i;t} = S_{i;t-1} + w_{i;t} x_t, \quad \kappa_{i;t} = \kappa_{i;t-1} + w_{i;t}, \quad (18)$$

starting with the initial statistics of the components, which can be chosen, for instance, with the help of the histogram analysis. The pointer model statistics update has the form [14, 15]

$$\nu_{i;t} = \nu_{i;t-1} + w_{i;t}. \quad (19)$$

The point estimates of all the component parameters λ are obtained using (18) in (8), while the point estimates of α are calculated with the help of (19) substituted into (12).

At this stage, at each time instant t, the *clusters* are identified by the estimated Poisson components with their means and variances $\hat{\lambda}_{i;t}$ $\forall i = \{1, \ldots, N_c\}$.

To *classify* the data at time t, the point estimate of the pointer c_t is determined as the index i of the maximum weight $w_{i;t}$ in the weighting vector w_t, i.e.,

$$\hat{c}_t = \arg\max_i [w_{1;t}, \ldots, w_{N_c;t}], \quad i \in \{1, \ldots, N_c\}, \quad (20)$$

which points to the active component where the current realization y_t belongs. The clustering and classification algorithm is summarized in the following algorithm.

3.1 Clustering algorithm

Algorithm

{Mixture initialization (for $t = 1$)}

Set the number of components N_c.

for all $i \in \{1, \ldots, N_c\}$ **do**

Set the initial statistics $S_{i;t}$, $\kappa_{i;t}$ and $\nu_{i;t}$.

Compute the point estimates of components $\hat{\lambda}_{i;t}$ and of the pointer model $\hat{\alpha}_{i;t}$ according to (8) and (12) respectively.

end for
{Recursive clustering)}
for $t = 2, 3, \ldots, T$ **do**
Measure the current realization y_t.
for all $i \in \{1, \ldots, N_c\}$ **do**
Compute the proximity m_i (15).
Obtain the weight $w_{i;t}$ [15] according to (16) and (17).
Update the statistics $S_{i;t}$, $\kappa_{i;t}$ and $\nu_{i;t}$ according to (18) and (19).
Re-compute the point estimates of parameters $\hat{\lambda}_{i;t}$ and $\hat{\alpha}_{i;t}$ according to (8) and (12).
end for
{Classification)}
Obtain the point estimate of c_t according to (20) and declare the active component.
end for

An example of the program implementation of the algorithm for the univariate variable y is given in the next section. The algorithm was implemented in a free and open source programming environment Scilab (www.scilab.org).

3.2 Program example

```
/// recursive mixture estimate of univariate multimodal Poisson variable

exec("ScIntro.sce",-1), mode(0)

///// SIMULATION
nd=1200;

al=fnorm(rand(3,1,'u'));        // alfa for c1
nc=size(al,1);                    // number of components

// parameters od y1
La=abs([rand(1,1,'n') 0.5*rand(1,1,'n')+15 rand(1,1,'n')+45]);


for t=1:nd
    c(t)=sum(rand(1,1,'uniform')>cumsum(al))+1;           // pointer
    y(t)=grand(1, 1, "poi", La(c(t)));// output
end
////FIGURES of DATA
scf(1);
plot(y);
title('Simulated data','Fontsize',5);
xlabel('Time','Fontsize',4);ylabel('y','Fontsize',4);

scf(2);
histc(y);
title('Histogram','Fontsize',5);
xlabel('y','Fontsize',4);ylabel('Frequencies','Fontsize',4);


///// MIXTURE INITIALIZATION
S=[2 16 47]'; /// statistics of components from histogram
ka=ones(nc,1);/// counters of components
nu=rand(nc,1,'uniform'); /// statistics of pointer
LaE=S./ka; /// point estimates of components
alE=nu./sum(nu);/// point estimates of pointer


///// RECURSIVE CLUSTERING AND CLASSIFICATION
```

```
for t=1:nd
    for i=1:nc
          /// poisson proximity for components
          m(i)=(LaE(i))^(y(t))*exp(-LaE(i))*(1/factorial(y(t)));
    end
    tildew=m.*alE;
    w=tildew./sum(tildew);  /// normalized weighting vector
    [nic,cE(t)]=max(w);   /// pointer estimation
    wt(:,t)=w;
    ///// statistics update
    for i=1:nc
       S(i)=S(i)+w(i)*y(t); /// of components
       ka(i)=ka(i)+w(i);/// counters of components
    end
    nu=nu+w; /// statistics update of pointer

    //// point estimates
    LaE=S./ka; /// point estimates of components
    alE=nu./sum(nu);/// point estimates of pointer

    //// prediction
    pred=2;
 select pred
    case 1
    bb=grand(1, 1, "poi", LaE(cE(t)));// from active component

    case 2
    bb=0;
    for i=1:nc
     bb=bb+w(i)*LaE(i);   /// weighted average of component point estimates
    end
end
    yp(t)=bb;


end

////RESULTS
scf(3);
ss=nd-200:nd;
plot(ss,c(ss),'bs');
plot(ss,cE(ss),'rx');
title('Pointer estimation','Fontsize',5);
set(gca(),"data_bounds",[ss(1) ss($) min(cE)-0.1 max(cE)+0.1]);
legend('simulated', 'estimated');
xlabel('Time','Fontsize',4); ylabel('Values of pointer','Fontsize',4);


////Classification accuracy
ep=sum(c(:)~=cE(:));
PE=(ep*100)/nd;
disp('Persentage of wrong classification',PE)


///// prediction error
RMSE = sqrt(mean((y(:) - yp(:)).^2));
// relative prediction error
RPE=variance(y(:)-yp(:))/variance(y(:));
```

```
/// filtered relative prediction error
RPEF=rpef(y,yp);
disp(['RMSE RPE RPEF'],[RMSE RPE RPEF]);


scf(9);
plot(ss,y(ss),'b');
plot(ss,yp(ss),'r');
title('Prediction','Fontsize',5);
xlabel('time','Fontsize',4);ylabel('y','Fontsize',4);
legend('simulated','prediction');



scf(10);
plot(y,cE,'.')
title('Detected clusters','Fontsize',5);
set(gca(),"data_bounds",[0 max(y)+1 min(cE)-0.1 max(cE)+0.1]);
xlabel('y','Fontsize',4); ylabel('Values of pointer','Fontsize',4);


//// which clusters the data belong to
[cE y (1:nd)']; // in time

/// descriptive statistics of data in clusters
kv=zeros(3,3);
for j=1:nc
    tc=find(cE==j);
    yc=y(tc);
    ym(j)=mean(yc);
    yv(j)=variance(yc);
    kvv=quart(yc);
    ma(j)=max(yc);
    mi(j)=min(yc);
    kv(:,j)=kvv;
   yclust{j}=yc;
   si(j)=size(yclust{j},1);
  end

disp(['mean variance'],[ym yv]);
disp(['min  max'],[mi ma]);
disp(['Q1,Q2,Q3'],[kv]);
disp('Size of clusters',si' )



scf(11);
bplot(yclust);
title('Boxplot','Fontsize',5);
ylabel('y','Fontsize',4);
```

4 Experiments

To validate the functionality of the algorithm, the experiments with simulated data sets were conducted. The quality of the classification was evaluated using the following criteria:

- Number of incorrect classifications (i.e., comparing the simulations and the point estimates of the pointer);

- The error of the data prediction from the components.

4.1 Simulated data

Using the program from Section 3.2, 1200 realizations of the count variable y were generated within each data set. Approximately 100 data sets were used for these experiments. An example of the histogram and the data plot of one of the data sets are given in Figure 1.

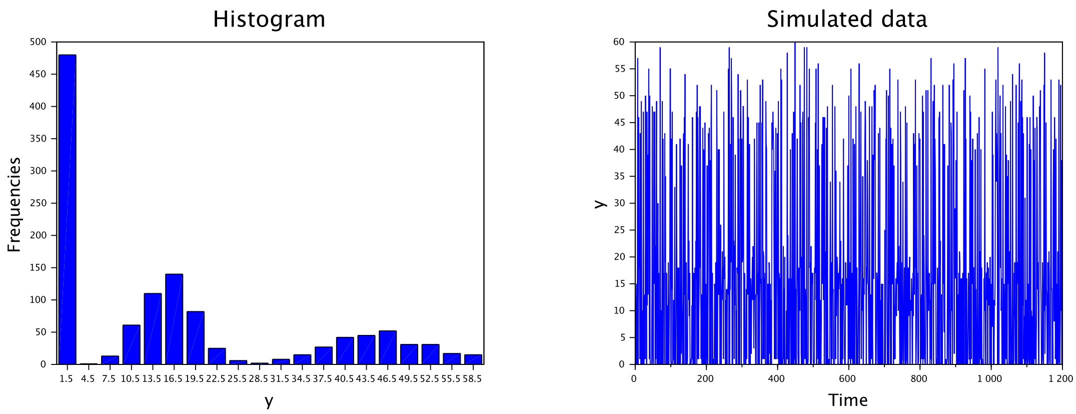

Figure 1: The example of the histogram (left) and values (right) of the simulated data set

4.1.1 Results

The average percentage of the incorrect classification during performing the experiments with simulated data sets was 0.5%, which was the value calculated as a result of comparison of the simulated and estimated pointer values. The example of the evolution of the pointer point estimates is given in Figure 2. For better illustration, a fragment of the estimation is shown.

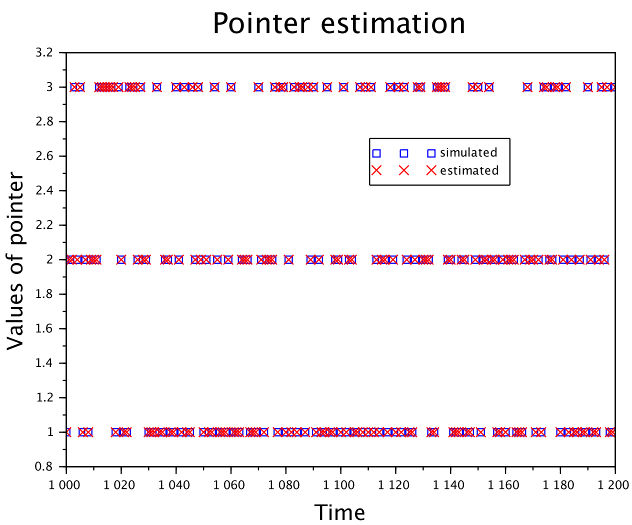

Figure 2: The fragment of the pointer estimation

The fragment of the data prediction can be seen in Figure 3.

In this work, the prediction accuracy was evaluated with the help of the relative prediction error (RPE)

$$\text{RPE} = \frac{D[y_t - \hat{y}_t]}{D[y_t]}, \tag{21}$$

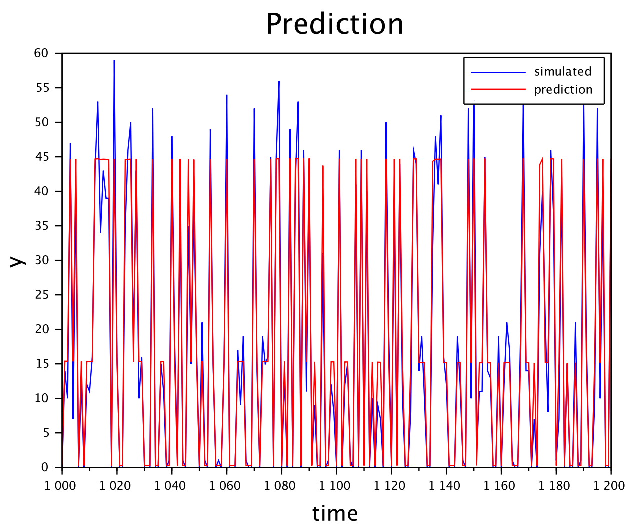

Figure 3: The fragment of the data prediction

where D denotes variance. RPE is defined on the interval from 0 to 1 and clearly shows whether the prediction corresponds to the data evolution. RPE with filtered data computed for averages from 3 previous values (RPE_F) was calculated as well. The average results obtained with the simulated data sets were: RPE = 0.0474102 and $\text{RPE}_F = 0.0444386$.

The clusters detected in the data space are demonstrated in Figure 4. As this work is limited to the univariate variable clustering, the clusters can be visualized only by plotting the data against the pointer estimates.

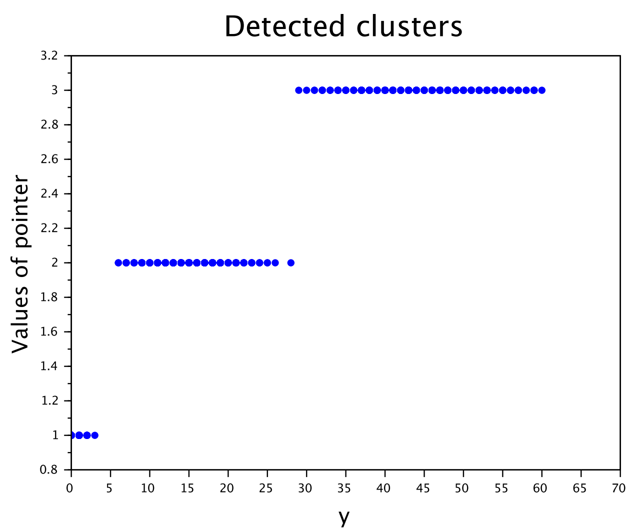

Figure 4: The detected clusters

4.2 Real data

To test the clustering algorithm on real data, the data set dealing with residential electric vehicle (EV) charging in apartment buildings from [23] described in [24] and related to [25] was used.

The data set was preprocessed to obtain the number of EV plugins in the residential garages observed per hour, i.e., the modeled variable y is the number of EV plugins per hour. The data set contains 9758 realizations y_t measured in the period of from December 2018 to January 2020.

The histogram of the data is demonstrated in Figure 5. In this work, it is described by the mixture of

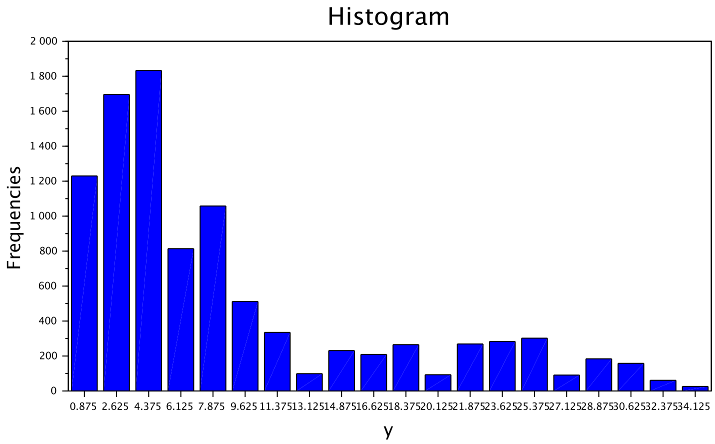

Figure 5: The histogram of the number of EV plugins

three Poisson components, whose hills can be guessed in the histogram as 3, 8, 23 plugins. These initial statistics were used for the mixture initialization.

4.2.1 Results

The pointer estimation for the whole period of learning is shown in Figure 6. Three values of the pointer were estimated, which means that the number of components was correctly initialized.

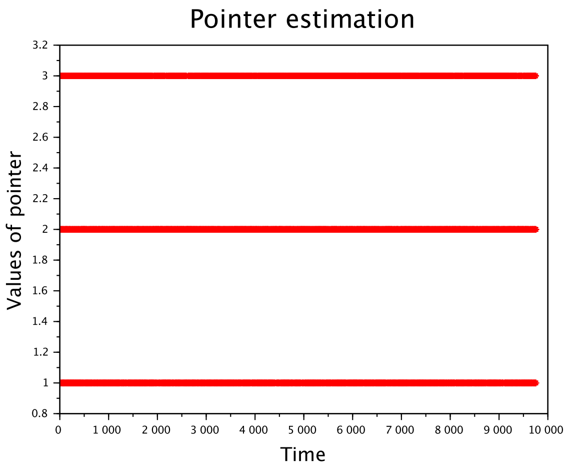

Figure 6: The pointer estimation

A fragment of the data prediction is shown in Figure 7. Three versions of predictions were compared: (i) with the weighted average of the point estimates of the component parameters, (ii) with the point estimates of the active component at each time instant directly, and (iii) by generating the values from the Poisson distribution of the active component. The prediction based on the weighted averages of the point estimates of the component parameters gave the most accurate results, which can be also seen in Table 1, where it has the lowest values of RPE and RPE_F.

The number of data captured into three clusters is 3855, 3839 and 2064 values respectively. The descriptive statistics of the data captured into three clusters are given in Table 2. It can be seen that the real data in the detected clusters meet the Poisson assumption of the equidispersion. The very slight deviations are insignificant. The boxplots of the data in the detected clusters can be seen in Figure 8. The figure and the table demonstrate that the range, interquartile range and location of the data values in the clusters are different.

Table 1: Comparison of RPE and RPE_F

Predictions	RPE	RPE_F
The weighted average	0.082544	0.0843112
The point estimates	0.1181072	0.115001
Generating from active distributions	0.2351426	0.2290116

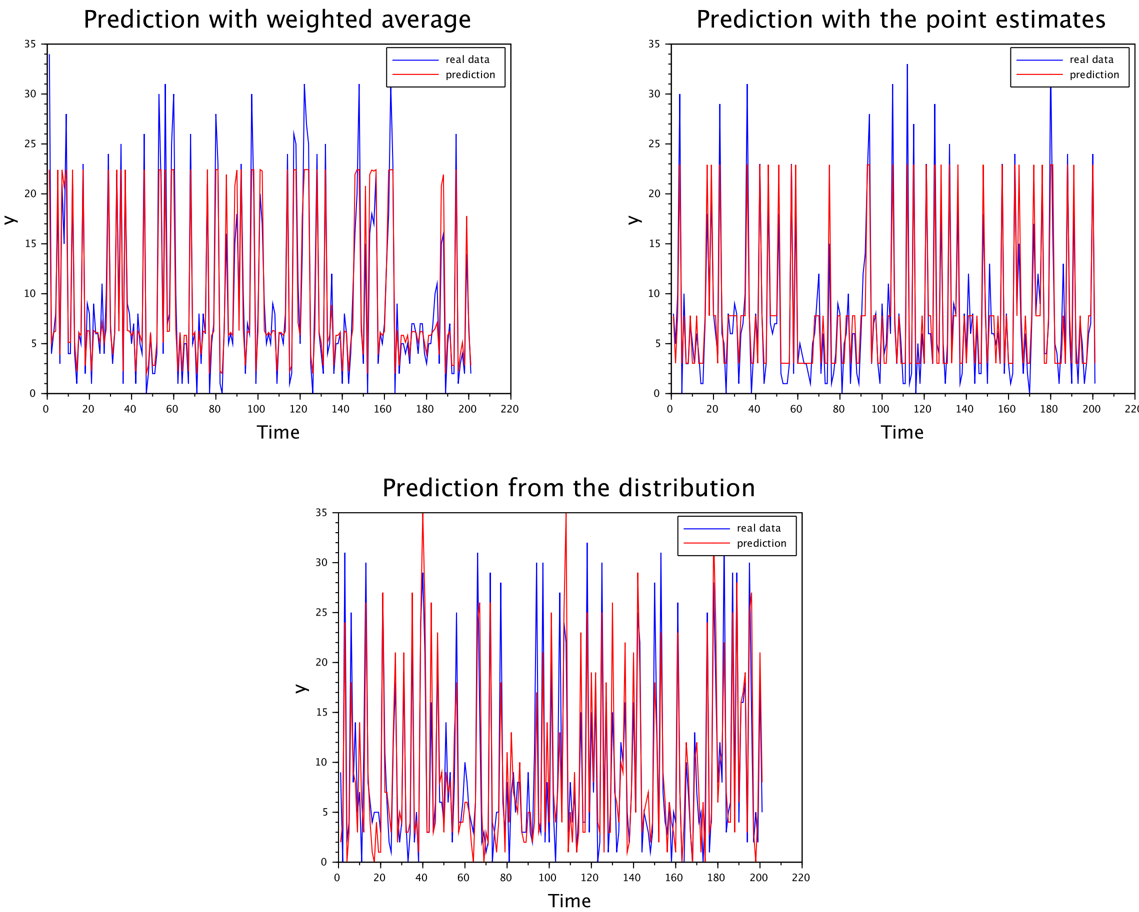

Figure 7: The comparison of the prediction of the number of EV plugins (fragment)

Table 2: Descriptive statistics of the data in clusters

	cluster 1	cluster 2	cluster 3
Mean	2.2941634	7.5274811	23.145833
Variance	1.7525736	6.0096015	24.610327
Minimum	0	5	13
Maximum	5	14	35
Q1	1	6	19
Q2	2	7	23
Q3	3	9	27

4.3 Discussion

The main aim of this work was to explain the clustering algorithm and to verify it using simulations, where the number of clusters is known in advance and thus, results of the classification can be compared with the simulated values of the pointer. Then, the testing of the algorithm on real data should have been

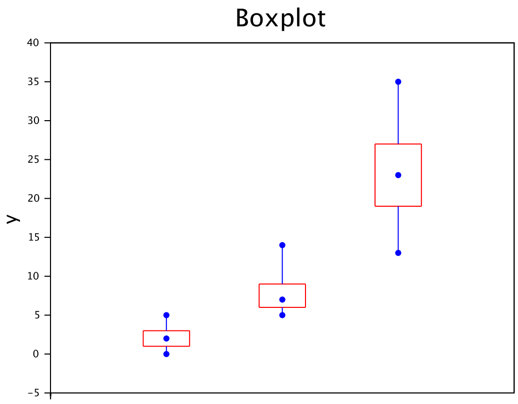

Figure 8: The boxplots of the data in the detected clusters

performed. It should be stated that the aim of the work has been successfully achieved: the classification accuracy is high and gives only 0.5% of wrong values of the pointer. With real data, the accuracy can be judged by means of the prediction errors from the estimated Poisson components, which were low. The lowest prediction errors was obtained for the case of using weighted averages of the point estimates of the component parameters.

It should be noted that the adopted methodology [14, 15] with a dynamic pointer [26] has been already applied to mixtures of the Poisson components in [4], however, with the Gaussian approximation of the proximity function. The presented experiments chosen to demonstrate a simple univariate case show that the Poisson probability function can be used as the proximity directly.

However, it should not be forgotten that the presented case study is a trivial recursive clustering of a univariate Poisson variable. It can serve as a preparation for the clustering and classification of a data vector composed from several variables. Some of them can explain the behavior of a target variable, which can be then used for predicting clusters in the data space.

5 Conclusion

The work deals with the task of clustering and classification of count data described by the univariate multimodal Poisson variable. The problem is solved with the help of the recursive Bayesian estimation of a mixture of Poisson distributions. The recursive statistics updates of the Poisson components and the static pointer model were used. The detailed clustering algorithm and a simple example of its program implementation were demonstrated along with illustrative experiments. Statistical analysis of the data captured in the detected clusters was performed.

Using the obtained results, the clustering of multidimensional Poisson variables will be the problem to be discussed in the future work.

Acknowledgements

The research was supported by project Arrowhead Tools, the project number ECSEL 826452 and MSMT 8A19009.

References

[1] Guenni L.B.: Poisson Distribution and Its Application in Statistics. In: Lovric M. (eds) International Encyclopedia of Statistical Science. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-04898-2_448 (2011)

[2] Donnelly Jr., R.: Business Statistics. 3rd edn. Pearson (2019)

[3] Anderson, D. R., Sweeney, D. J., Williams, T. A., Camm, J. D., Cochran, J. J.: Essentials of Modern Business Statistics with Microsoft Office Excel (Book Only). 7th edn. Cengage Learning, (2017)

[4] Petrouš, M., Suzdaleva, E., Nagy, I.: Modeling of passenger demand using mixture of Poisson components. In: Gusikhin, O., Madani K., Zaytoon J. (eds.) Proceedings of the 16th INTERNATIONAL CONFERENCE ON INFORMATICS IN CONTROL, AUTOMATION AND ROBOTICS (ICINCO 2019), p. 617–624 (2019)

[5] Lukusa, M. T., Phoa, F. K. H. A Horvitz-type estimation on incomplete traffic accident data analyzed via a zero-inflated Poisson model, Accident Analysis & Prevention, vol.134, 105235, (2020)

[6] González-Arévalo, B. and Roy, J. Simulating a Poisson cluster process for Internet traffic packet arrivals, Computer Communications, 33 (5): 612-618, (2010)

[7] Jaggia, S., Kelly, A.: Business Statistics: Communicating with Numbers. 3rd edn. McGraw-Hill Education (2018)

[8] Doane, D., Seward, L.: Applied Statistics in Business and Economics. 3rd edn. Mcgraw-Hill (2010)

[9] Gupta, M.R. and Chen, Y. Theory and use of the EM algorithm. Foundations and Trends in Signal Processing, 4(3), 223-296, (2011)

[10] Jayasekare, R. R., Gill, R., Lee, K. Modeling discrete stock price changes using a mixture of Poisson distributions, Journal of the Korean Statistical Society, 45(3): 409–421, (2016)

[11] Ma, J., Liu, J., Ren, Z. Parameter estimation of Poisson mixture with automated model selection through BYY harmony learning, Pattern Recognition, 42(11): 2659–2670, (2009)

[12] Harris, I. R., Shen, S. The minimum L2 distance estimator for Poisson mixture models, Journal of Statistical Planning and Inference, 141(3): 1088–1101, (2011)

[13] Karlis, D., Xekalaki, E. Minimum Hellinger distance estimation for Poisson mixtures, Computational Statistics & Data Analysis, 29(1): 81-103, (1998)

[14] Kárný, M., Kadlec, J., Sutanto, E.L.: Quasi-Bayes estimation applied to normal mixture. In: Rojíček, J., Valečková, M., Kárný, M., Warwick, K. (eds.). Preprints of the 3rd European IEEE WORKSHOP ON COMPUTER-INTENSIVE METHODS IN CONTROL AND DATA PROCESSING, Prague, CZ, pp. 77–82 (1998)

[15] Kárný, M., Böhm, J., Guy, T.V., Jirsa, L., Nagy, I., Nedoma, P., Tesař, L.: Optimized Bayesian Dynamic Advising: Theory and Algorithms. Springer, London (2006)

[16] Nagy, I., Suzdaleva, E.: Algorithms and Programs of Dynamic Mixture Estimation. Unified Approach to Different Types of Components, SpringerBriefs in Statistics. Springer International Publishing, Heidelberg (2017)

[17] Nagy, I.; Suzdaleva, E., Pecherková, P. Comparison of various definitions of proximity in mixture estimation. In: Proceedings of the 13th International Conference on Informatics in Control, Automation and Robotics - Volume 1: ICINCO, ISBN 978-989-758-198-4, pp. 527–534, (2016)

[18] Jozová, Š.; Uglickich, E. and Nagy, I. Bayesian mixture estimation without tears. In: Proceedings of the 18th International Conference on Informatics in Control, Automation and Robotics - ICINCO, ISBN 978-989-758-522-7, pp. 641–648, (2021)

[19] Sinharay, S. Discrete Probability Distributions. In: International Encyclopedia of Education. 3rd Edition), Editor(s): P. Peterson, E. Baker, B. McGaw, Elsevier, pp.132-134, (2010)

[20] Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., Rubin, D. B.: Bayesian Data Analysis (Chapman & Hall/CRC Texts in Statistical Science). 3rd edn. Chapman and Hall/CRC (2013)

[21] HyvPönen, V. and Tolonen, T. Bayesian Inference 2019. https://vioshyvo.github.io/Bayesian_inference/. Last accessed March 2019

[22] Peterka, V. Bayesian system identification, in Eykhoff, P. (Ed.), Trends and Progress in System Identification. Oxford, Pergamon Press, pp. 239-304, (1981)

[23] Sørensen, Å. L. Residential electric vehicle charging datasets from apartment buildings, Mendeley Data, V1, doi: 10.17632/jbks2rcwyj.1, (2021)

[24] Sørensen, Å. L., Lindberg, K. B., Sartori, I., Andresen, I. Residential electric vehicle charging datasets from apartment buildings, Data in Brief, volume 36, 107105, ISSN 2352-3409. doi: 10.1016/j.dib.2021.107105, (2021)

[25] Sørensen, Å. L., Lindberg, K. B., Sartori, I., Andresen, I. Analysis of residential EV energy flexibility potential based on real-world charging reports and smart meter data, Energy and Buildings, volume 241, 110923, (2021)

[26] Nagy, I., Suzdaleva, E., Kárný, M., Mlynářová, T. Bayesian estimation of dynamic finite mixtures. International Journal of Adaptive Control and Signal Processing. 25(9): 765-787, (2011)