Proceedings of the World Congress on Engineering and Computer Science 2017 Vol I
WCECS 2017, October 25-27, 2017, San Francisco, USA

Model of Study Results in Master Degree at Faculty of Management in Jindřichův Hradec

Vladislav Bína, *Member, IAENG,* and Jiří Přibil, *Member, IAENG*

***Abstract*—A recent increase of university capacities in the Czech Republic together with present demographical decrease in the cohort of applicants results in a significant change of behavior of students and raises the need to change a strategy of university marketing. In last years, the management of many Czech universities faced a serious issue of significantly lower numbers of applicants and substantially higher rate of unsuccessful finish of studies or prolongation of the study period. At Faculty of Management in Jindřichův Hradec the above mentioned problem appeared to be significant in the Bachelor's Degree, particularly in the blended learning form.**

The numbers of students comprises a basis of a significant part of universities funding in the Czech Republic. The goal of this paper is to search for the most important factors influencing study results of Master's Degree of Faculty of Management. Since the variables describing study data are categorical, for the analysis of important factors we can employ means of probabilistic analysis of dependency structure, particularly so called compositional models.

The compositional models rank among the means of probabilistic modeling and present an alternative capable of representing and modeling dependency structures without the necessity to employ graphical apparatus of directed acyclic graphs (unlike the Bayesian networks rather confusingly hinting at non-existing causal relations).

***Index Terms*—Faculty of Management, study results, master's degree, probabilistic dependency structure, compositional model.**

I. Introduction and Motivation

THE primary goal of this paper is a contribution to the analysis of data concerning Master's Degree students studying (from 2011 to 2015) in an ECTS-labeled study program at Faculty of Management in Jindřichův Hradec (University of Economics, Prague). The issues of successful completion and study problems is important world-wide (see, e.g., DeAngelo et al. [8] or Vossensteyn et al. [16]). But since the numbers of students comprises a basis of a significant part of universities funding it is an acute and ubiquitous problem of the Czech universities (the factors influencing study results and overall success of students are analyzed, e.g., in Maryška et al. [12], Doucek and Maryška [9] or in Míková et al. [13]).

Since the issues of study problems, factors affecting probability of successful graduation and results in particular subjects are complex, we significantly simplify the analysis and focus only on determinants of study results in compulsory subjects and subjects in students' minor specialization. Thus, in order to simplify our considerations, we exclude all optional subjects, which do not appear to be problematic and are usually chosen in connection to the student's interests. Such subjects are in general successfully finished or substituted by another (non-problematic) subject by almost every active student.

The presented results can be considered as stand-alone. But in a way it extends the research performed in case of study results in Bachelor's Degree study at Faculty of Management in Jindřichův Hradec published by Bína and Přibil [7].

A. Grading and Study System

Study results in the particular Master's Degree subjects are represented by categorical grades on a standard grading scale used by the University of Economics, Prague. The student receives a grade of '1' for outstanding performance with 90–100 points, a grade of '2' for a very good result of 75–89 points, a grade of '3' for good results between 61–74, a grade of '4+' for insufficient results with possibility of repetition between 50 and 59 points, and '4' for students failing with less than 50 points. In the case of state exams, students cannot obtain a '4+' grade (but in the case of failure they can repeat twice). Ungraded courses are finished with 'Z' in the case of successful completion and by 'NZ' in the case of failure. Special category '–' is reserved for students who did not show at exams and 'O' for excused unsuccessful ending.

For the sake of simplicity and in order to represent previous unsuccessful attempts in particular subjects we started with a preprocessing step converting current results and previous unsuccessful endings to the simplified scale. In this new scale '1', '2', '3' remain unchanged. Successful ending of ungraded courses is denoted by 'S'. All subjects not finished in a successful way are put together to the category 'F' and finally any successful completion of subject after previous failure is labeled by '!'.

As we mentioned above and proposed already in [7], the studied topic is rather complex and very often concerns unobserved (or even unmeasurable) factors. Therefore, we study a simplified variant of the problem and, thanks to the categorical nature of the study data, it is natural to employ means of description used for multidimensional problems of uncertain character. These are usually handled using multivariate probabilistic tools.

Unlike the previously studied case of Bachelor's Degree (Bína and Přibil [7]), in the Master's Degree at Faculty of Management after block of subjects called a common basis students choose one of the subsidiary specializations, i.e. *Business Management*, *Management of Public Services*, *Information Management* or *Health-care Management*. Since the study problems usually appears after the choice of specialization, it is natural to divide the problem into 4 subproblems and analyze students of all four specializations separately.

Manuscript received July 23, 2017; revised August 07, 2017. The research was supported by GA ČR under Grant No. 15-00215S.

V. Bína and J. Přibil are with the Department of Exact Methods, Faculty of Management, University of Economics, Prague, Jarošovská 1117/II, 37701 Jindřichův Hradec, Czech Republic, email: {bina,pribil}@fm.vse.cz.

ISBN: 978-988-14047-5-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

B. Study Data and Methodology

The data file concerns 1,319 Master's Degree students and we study (according to specialization and respective counts of obligatory subjects) from 21 to 23 important factors and search for their probabilistic dependence structure. It is obvious that it is impossible to easily sample a multivariate distribution with more than 20 categorical variables. In the case of dichotomic variables it would mean to estimate more than 1 million table cells; but our situation is even worse since most of the variables are multinomial, usually with 4 or 5 possible values. Therefore, it is important to construct a multidimensional model which can be composed from marginals of lower dimensions where it is feasible to estimate probability tables. This can be done using a compositional model, which is a probabilistic model capable of representing and modeling dependency structures without the necessity to employ graphical apparatus of directed acyclic graphs (unlike the Bayesian networks confusingly hinting at non-existing causal relations). For basics of the theory of compositional models see, e.g., Jiroušek [10] and Bína and Jiroušek [5]. For the construction of compositional models we use an adaptation of statistical structure learning principles based on the likelihood ratio test of independence (see Agresti [1]) and its important property of decomposability (in context of compositional models see, e.g., Bína [3] or [4]).

II. Notation, Basic Notions and Methodology

As we already sketched above, the structure learning can be based on the so-called neighborhood structures of decomposable compositional models (for details see Bína [4]) and on the decomposability of a likelihood criterion usable for local computation in the process of (sub)optimal search in the space of decomposable compositional models (see Bína [3], where the notion of decomposability in probabilistic models is well known and dates back more than 40 years ago, see, e.g., Wermuth [17]). Let us summarize the most important properties essential for the formulation of the structure learning algorithm.

For $K \subset N$, symbol $\kappa(x_K)$ denotes a $|K|$-dimensional distribution of variables from the system $X_K = \{X_i\}_{i \in K}$, which is defined on all subsets of a Cartesian product $\mathbf{X}_K = \times_{i \in K} \mathbf{X}_i$. In order to keep the notation simple, symbol $\kappa(x_K)$ is also used to denote a value of probability distribution κ at the point x_K. For $L \subset K$ the symbol $\kappa(x_L)$ denotes the corresponding marginal distribution.

A. Composition

Two multidimensional distributions can be composed in the following manner (embedding a relation of conditional independence between groups of variables).

Definition 1 (Operator of composition): For two discrete probability distributions $\kappa \in \Pi^{(K)}$ and $\lambda \in \Pi^{(L)}$ such that $\kappa(x_{K \cap L}) \ll \lambda(x_{K \cap L})$, their *composition* is defined by the formula

$$\kappa(x_K) \triangleright \lambda(x_L) = \kappa(x_K)\lambda(x_{L \setminus K} \mid x_{K \cap L}).$$

Here the symbol $\ll$ stands for the relation of *dominance* (also referred to as *absolute continuity*).

The operator of composition can be iterated and the result of the repeated application to the sequence of low-dimensional distributions is (if defined) a multidimensional distribution. The resulting multidimensional distribution $\hat{\kappa}$ is so-called *compositional model*

$$\hat{\kappa} = (\ldots((\kappa_1 \triangleright \kappa_2) \triangleright \kappa_3) \triangleright \ldots) \triangleright \kappa_n.$$

Because of its properties, this model can be written as a plain sequence of low-dimensional distributions (a *generating sequence*), where only sets of variable indices are noted $\hat{\kappa} = (K_1 \bullet K_2 \bullet \ldots \bullet K_k)_\kappa$ for the sake of simplicity.

B. Neighborhood Structure

The class of decomposable models corresponds to an analogous subclass of undirected graphs and, in the case of compositional models, we can also introduce its definition using the running intersection property (RIP); see, e.g., Koller and Friedman [11].

Within the class of decomposable compositional models, we can introduce a neighborhood structure given by the following pair of assertions (Theorems 1 and 2).

Theorem 1: For any decomposable model $\hat{\kappa}$ such that $\hat{\kappa} = (K_1 \bullet K_2 \bullet \ldots \bullet K_k)_\kappa$ (with the exception of the independent model as a product of one-dimensional marginals) there exists a decomposable model $\hat{\kappa}' = (K'_1 \bullet K'_2 \bullet \ldots \bullet K'_{k'})_\kappa$ where one additional conditional independence relation is introduced between a pair of variables which appear (as a whole) in only one set of indices K_i ($i \in \{1, \ldots, k\}$).

Theorem 2: For any decomposable model $\hat{\kappa}$ such that $\hat{\kappa} = (K_1 \bullet K_2 \bullet \ldots \bullet K_k)_\kappa$ (non-trivial, i.e., embedding at least one conditional independence relation), there exists a decomposable model $\hat{\kappa}' = (K'_1 \bullet K'_2 \bullet \ldots \bullet K'_{k'})_\kappa$, such that there exists a pair of variables which are conditionally independent given the rest of variables in the model in the case of model $\hat{\kappa}$, but not in the case of model $\hat{\kappa}'$.

For further clarification of notions, proofs and simple examples, see again Bína [4].

C. Likelihood-ratio Statistics

For testing whether the compositional model $\hat{\kappa}$ sufficiently faithfully approximates the original data distribution κ (both with variables from $\mathbf{X}_K$), the likelihood-ratio test statistic is defined by the formula

$$G^2 = 2 \sum_{x_K \in \mathbf{X}_K} \kappa(x_K) \log \frac{\kappa(x_K)}{\hat{\kappa}(x_K)}$$

and, under certain conditions, has χ^2 distribution with the appropriate number of degrees of freedom (see above). In the case of a likelihood-ratio statistic, the sample large enough for the approximation of χ^2 distribution is usually considered when the sample is at least five times larger than the number of cells in a contingency table (see Agresti [1]).

ISBN: 978-988-14047-5-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

D. Decomposition of Likelihood-ratio Statistic

Now we shall take an advantage of the decomposability of models in order to decompose[1] the G^2 test statistics. This method employs the neighborhood structure of decomposable models as described in Theorems 1 and 2.

Using the properties of logarithm, we can take advantage of neighboring models and arrive at formula

$$G^2_{\hat{\kappa}'} = G^2_{\hat{\kappa}} + \\ + 2 \sum_{x \in \mathbf{X}_{K_i}} \kappa(x_{K_i}) \log \frac{\kappa(x_{K_i \setminus \{\ell,m\}})\kappa(x_{K_i})}{\kappa(x_{K_i \setminus \{\ell\}})\kappa(x_{K_i \setminus \{m\}})} \quad (1)$$

which allows the enumeration of the likelihood-ratio statistic using a pre-computed value of the neighboring model and employing only local computations.

E. Degrees of Freedom for Likelihood-Ratio Decomposition

The number of degrees of freedom for the likelihood-ratio statistics in the model $\hat{\kappa}$ are given by formula

$$df = \prod_{k \in K} r_k - 1 - \sum_{i=1}^{n} \left(\prod_{j \in K_i \setminus U_i} r_j - 1 \right) \cdot \prod_{j \in K_i \cap U_i} r_j$$

where symbol r_k denotes the number of categories for the corresponding variable. If we introduce a new conditional independence relation among the pair of variables with indices $\ell, m \in K_i$, the change in the number of degrees of freedom Δdf can be computed from the previous values as follows

$$\Delta df = (r_\ell - 1)(r_m - 1) \prod_{j \in K_i \setminus \{\ell,m\}} r_j. \quad (2)$$

This change in the number of degrees of freedom for neighboring models (after introduction of one additional conditional independence relation) is in agreement with the results obtained in the case of hierarchical log-linear models and Bayesian networks (see, e.g., Agresti [1] or Neapolitan [14].

F. Akaike Information Criterion

The test statistic itself does not contain information about the number of parameters used for the representation of the model. This information is, in the case of hypothesis testing, employed in the form of degrees of freedom. But in the 1970s, the Akaike information criterion was formulated on the basis of the parsimony principle (see [2]). It can be expressed in a form using likelihood–ratio statistics

$$AIC_{G^2} = G^2 - 2 \cdot df$$

where G^2 is the likelihood–ratio and df is the number of degrees of freedom.

Now, in the case of neighboring decomposable compositional models, thanks to the above-expressed Formulae (1) and (2) we can again locally compute the value of Akaike information criterion and use it for search among models.

[1]The word decomposable (decompose) is used here in an ambiguous manner; it has two different meanings. The decomposability of a model is a structural property (characterizable by the RIP property). But the decomposability of a test statistic means the possibility to perform only local computations and hence take advantage of previous computations in the process of searching for a suitable model.

TABLE I
THE MOST IMPORTANT FACTORS DESCRIBING STUDENTS.

Variable	Meaning	Values
Gender	Gender of the student	female, male
State	Citizenship of the student	Czech Rep., Slovakia, etc.
Form	Study form	face-to-f., blended learn.
Bc School	Bachelor's Degree from	FM, VŠPJ, VŠTE, Other

TABLE II
OBLIGATORY SUBJECTS COMMON FOR ALL ANALYZED STUDENTS.

Variable	Meaning	Values
6MI401	Managerial Decision-Making	1,2,3,!,F
6MI405	Managerial Informatics	1,2,3,!,F
6MP401	Strategic Management and Analysis	1,2,3,!,F
6MP415	Marketing Communication	1,2,3,!,F
6MP455	Managerial Economics	1,2,3,!,F
6MP486	Management of Organizations	1,2,3,!,F
6MP580	International Management	1,2,3,!,F
6SE410	Economics II	1,2,3,!,F
6HV433	Psychology and Sociology for Managers	1,2,3,!,F
6MV410	Regional Management	1,2,3,!,F

G. Suboptimal Search among Decomposable Models

The search algorithm uses a simple idea of limiting the number of decomposable models tested in each iteration to a certain limit k. The algorithm is based on the use of neighborhood Theorems 1 and 2.

Suboptimal search using the test criterion algorithm starts with the saturated model (with no conditional independence relation introduced) and then the three steps described below follow.

1) *Generate all possible decomposable models with one additional conditional independence relation between a pair of variables.*
2) *Choose k models with the lowest values of the criterion.*
3) *Repeat steps 1 and 2 as long as the values of the criterion decrease.*

Obviously, we obtain only a suboptimal solution due to the greedy character of the algorithm, since it does not search the entire space of decomposable models.

III. STUDY DATA 2011–2015

As we already mentioned, in the presented paper we analyze anonymized data from the university information system InSIS concerning 1,319 Master's Degree students of a management study program on Faculty of Management in Jindřichův Hradec in the years 2011–2015. For the sake of clarity, in Table I we present a selection of the most important factors describing students and in Table II we cover only obligatory subjects.

The following four Tables III, IV, V and VI cover subjects of the four minor specializations. Namely Table III shows three profiling subjects, diploma seminar, thesis defense and state exams of Business Management specialization, Table IV presents situation in specialization of Management of Public Services, Table V describes situation in specialization of Information Management and Table VI presents subjects in specialization of Health-care Management. Thus, we will

TABLE III
OBLIGATORY SUBJECTS IN SPECIALIZATION OF BUSINESS MANAGEMENT.

Variable	Meaning	Values
6MP485	Managerial Decision-Making	1,2,3,!,F
6MP509	Managerial Informatics	1,2,3,!,F
6MP413	Strategic Management and Analysis	1,2,3,!,F
6MP001	Diploma Seminar I	S,F
6MP002	Diploma Seminar II	S,F
6DP_PS	Diploma Thesis Defense	1,2,3,!,F
6SZVPS	Specialization State Exam	1,2,3,!,F
6SZ_MN	Major Field State Exam	1,2,3,!,F

TABLE IV
OBLIGATORY SUBJECTS IN SPECIALIZATION OF MANAGEMENT OF PUBLIC SERVICES.

Variable	Meaning	Values
6MV420	Public Economics	1,2,3,!,F
6MV441	Management of Public Services	1,2,3,!,F
6MV510	Theory and praxis of public administration	1,2,3,!,F
6MV511	Diploma Seminar	S,F
6DP_VS	Diploma Thesis Defense	1,2,3,!,F
6SZVVS	Specialization State Exam	1,2,3,!,F
6SZ_MN	Major Field State Exam	1,2,3,!,F

TABLE V
OBLIGATORY SUBJECTS IN SPECIALIZATION OF INFORMATION MANAGEMENT.

Variable	Meaning	Values
6MI404	Analysis and design of information systems	1,2,3,!,F
6MI413	Knowledge Management	1,2,3,!,F
6MI420	Modeling in Economy	1,2,3,!,F
6MI510	Diploma Seminar	S,F
6DP_MI	Diploma Thesis Defense	1,2,3,!,F
6SZVMI	Specialization State Exam	1,2,3,!,F
6SZ_MN	Major Field State Exam	1,2,3,!,F

TABLE VI
OBLIGATORY SUBJECTS IN SPECIALIZATION OF HEALTH-CARE MANAGEMENT.

Variable	Meaning	Values
6MZ401	Health Care Management	1,2,3,!,F
6MZ402	Public Health and Epidemiology	1,2,3,!,F
6MZ503	Health Economics and Health Care Financing	1,2,3,!,F
6MZ504	Information Technology in Health Care	1,2,3,!,F
6MZ505	Legal and Ethical Aspects of Health Care	1,2,3,!,F
6MZ600	Diploma Seminar	S,F
6DP_MZ	Diploma Thesis Defense	1,2,3,!,F
6SZVMZ	Specialization State Exam	1,2,3,!,F
6SZ_MN	Major Field State Exam	1,2,3,!,F

divide the analysis into four models for each study specialization in such a way that it concerns the results of obligatory subjects together with the subjects of considered specialization.

IV. RESULTS OF STRUCTURE LEARNING

The procedure of the structure learning taking advantage of neighborhood structure of decomposable model was starts with the full model (without any conditional independence relation). It then continues with as long as any model simplification results in an accepted model.

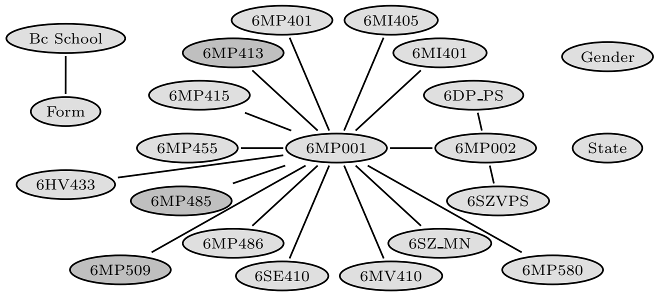

Fig. 1. A dependency structure of study results data in a specialization of Business Management.

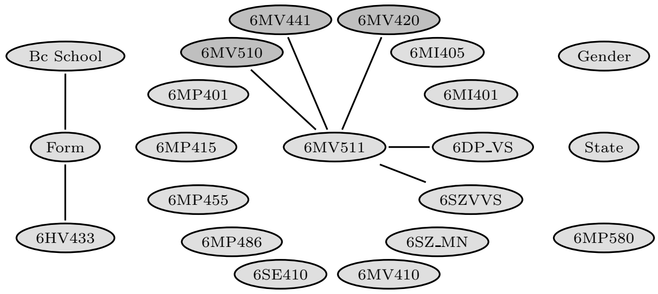

Fig. 2. A dependency structure of study results data in a specialization of Management of Public Services.

A. Resulting Model for Specialization of Business Management

The greedy search for a (sub)optimal model of dependence structure in case of specialization of Business Management resulted in a model depicted in Figure 1. Let us mention that profiling specialization subjects are denoted by darker shade of grade filling of the graph nodes.

The resulting model is characterized by

$$G^2 = 39,356.25; \quad df \doteq 8.79 \cdot 10^{13};$$
$$AIC = 44,140.21; \quad \Delta_{AIC} = 23.64.$$

The resulting suboptimal solution shows interestingly that no other variable is probabilistically dependent on Gender and State. But even more surprisingly, study results are neither dependent on previous Bachelor's Degree School and study Form. Less surprising, but still interesting appears the relation of Diploma Seminar I (6MP001) with all previous study subjects and its relation to the subsequent Diploma seminar II (6MP002) and Major State Exam. The second Diploma Seminar is then in relation with Diploma Thesis Defense and Specialization State Exam. (Business Management is the only specialization which organizes diploma seminar in two semesters.)

B. Resulting Model for Specialization of Management of Public Services

Now, the greedy search for a (sub)optimal model of dependence structure in case of specialization of Management of Public Services resulted in a model depicted in Figure 2. Again, the profiling specialization subjects are denoted by darker shade of grade filling of the graph nodes.

ISBN: 978-988-14047-5-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

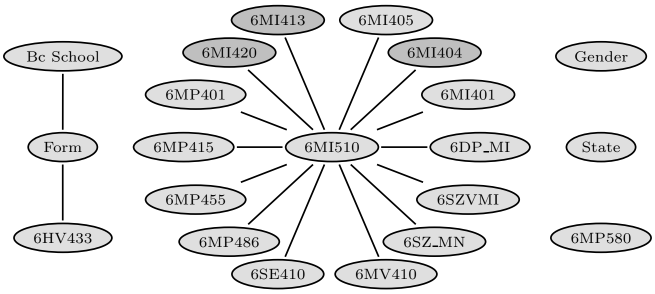

Fig. 3. A dependency structure of study results data in a specialization of Information Management.

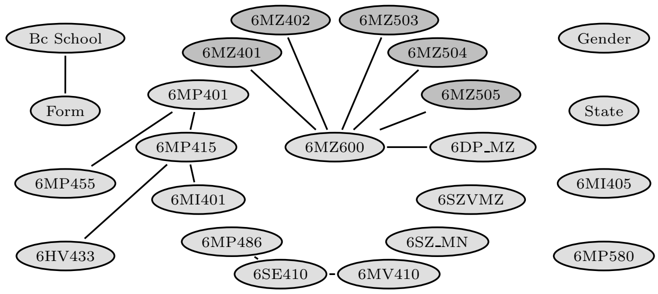

Fig. 4. A dependency structure of study results data in a specialization of Health-care Management.

The resulting model is characterized by

$$G^2 = 39,631.54; \quad df \doteq 3.52 \cdot 10^{13};$$
$$AIC = 42,875.38; \quad \Delta_{AIC} = 17.98.$$

The resulting suboptimal solution shows again that no other variable is probabilistically dependent on Gender and State but there appears ten compulsory subjects that are surprisingly independent with any other considered variable including the Major State Exam. Again, the previous Bachelor's Degree School and study Form are interconnected, but now the study Form is related to the subject of Psychology and Sociology (6HV433). And even more interestingly than in previous case, study results in the specialization appear to be in relation only with other subjects of the Management of Public Services specialization including the Diploma Thesis Defense and Specialization State Exam (the crutial point is again the Diploma Seminar (6MV511).

C. Resulting Model for Specialization of Information Management

In this case, the greedy search for a (sub)optimal model of dependence structure in case of specialization of Information Management resulted in a model depicted in Figure 3. Again, the profiling specialization subjects are denoted by darker shade of grade filling of the graph nodes.

The resulting model is characterized by

$$G^2 = 35,931.86; \quad df \doteq 2.11 \cdot 10^{13};$$
$$AIC = 40,104.34; \quad \Delta_{AIC} = 9.43.$$

The resulting suboptimal solution shows again that no other variable is probabilistically dependent on Gender and State and results of International Management (6MP580). Once more, the previous Bachelor's Degree school and study form are interconnected and study Form is related to the subject of Psychology and Sociology (6HV433). Similarly to the case of Business Management specialization, study results in the specialization of Information Management shows the relation of the Diploma Seminar (6MI510) with nearly all previously studied subjects and relations to the subsequent Specialization and Major State Exam.

D. Resulting Model for Specialization of Health-care Management

Finally, the greedy search for a (sub)optimal model of dependence structure in case of specialization of Health-care Management resulted in a model depicted in Figure 4. Again, the profiling specialization subjects are denoted by darker shade of grade filling of the graph nodes.

The resulting model is characterized by

$$G^2 = 28,970.95; \quad df \doteq 1.10 \cdot 10^{14};$$
$$AIC = 36,382.38; \quad \Delta_{AIC} = 21.55.$$

The resulting suboptimal solution shows again that no other variable is probabilistically dependent on Gender and State and results of several compulsory subjects. Even in this case, the previous Bachelor's Degree school and study form are interconnected. Similarly to the case of Management of Public Services specialization, study results in the specialization of Health-care Management shows the relation of the Diploma Seminar (6MZ600) with all subjects previously studied in specialization and relations to the subsequent Diploma Thesis defense. There also appears an interesting dependence structure between pairs of variables among five compulsory subjects.

TABLE VII
A CONDITIONAL DISTRIBUTION OF STUDY FORM GIVEN BACHELOR'S DEGREE EDUCATION FOR BUSINESS MANAGEMENT SPECIALIZATION.

Form	FM UEP	VŠPJ	VŠTE	Other
Blended learning	0.244	0.353	0.432	0.771
Face-to-face	0.756	0.647	0.568	0.229

E. Closer Look at Interesting Results

Within the limited space of the presented paper, let us show several interesting results which can be inferred from the resulting models. Table VII represents one of the building stones of resulting compositional models, particularly, conditional distribution for the relation between the previous Bachelor's Degree education and study Form of students.

From a conditional distribution (Table VII) of study Form given Bachelor's Degree education in case of Business Management specialization we can observe that the lowest frequency of blended learning choice (less than a quarter) appear in the case of students from our faculty (Faculty of Management, University of Economics, Prague). Less than one half of students chooses the blended learning in case of two regional universities providing only Bachelor's Degree education. Students from other universities are motivated to choose our faculty in more the three quarters of cases because of the possibility to study in blended learning form. Let us mention that the observed proportions differ statistically significantly (Pearson's Chi-squared test: $\chi^2 = 314.69$, $df = 3$, $p < 2.2 \cdot 10^{-16}$).

ISBN: 978-988-14047-5-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

TABLE VIII
A CONDITIONAL DISTRIBUTION OF STUDY FORM AND RESULTS OF 6HV433 FOR MANAGEMENT OF PUBLIC SERVICES SPECIALIZATION.

Form	1	2	3	!	F
Blended learning	0.132	0.461	0.697	0.658	0.840
Face-to-face	0.868	0.539	0.303	0.342	0.160

This situation appears to be similar also in the cases of other specializations.

Another interesting result can be observed Table VIII which represents a building stone of resulting compositional model in case of specialization of Management of Public Services. Namely, it is the conditional distribution describing the relation between the study Form of students and results in Psychology and Sociology (6HV433).

From a conditional distribution (Table VIII) of proportion of study Forms for different results of Psychology and Sociology (6HV433) in case of Management of Public Services specialization we can observe that higher proportions of blended learning form appear in case of worse results (complete failure to finish the subject or finalization after unsuccessful previous result. Let us mention that the observed proportions differ statistically significantly (Pearson's Chi-squared test: $\chi^2 = 168.75$, $df = 4$, $p < 2.2 \cdot 10^{-16}$).

V. CONCLUSION

In the above presented paper we briefly introduced methodology of compositional models capable to represent efficiently multidimensional probabilistic distributions. Moreover, the methodology serves as a data mining approach revealing dependence structure among the considered variables and searching for the factors influencing strongly some important variable.

But compositional models can be used not only to present interesting relations and hidden dependencies. They can be also used for the computation of results of hypothetical interventions (inspired by Judea Pearl's do-calculus, see, [15], for compositional models see Bína and Jiroušek [6]).

The limited scope of paper allowed us to present the general methodology on an interesting real-life example of Master's Degree study data in four study specializations at the Faculty of Management, University of Economics, Prague. The presented dependency structures appear to be to some extent surprising but still plausible; we were able to present only brief look at two conditional probability tables showing interesting dependencies.

ACKNOWLEDGMENT

The authors would like to thank to the members of university management who generously extracted and provided study data concerning the students of Faculty of Management in Jindřichův Hradec.

REFERENCES

[1] A. Agresti, *Categorial Data Analysis*. John Wiley & Sons Inc., New York, 2002.
[2] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716-723, 1974.
[3] V. Bína, "Greedy Search among Decomposable Models," in *Proceedings of The Fourteenth IASTED International Conference on Artificial Intelligence and Soft Computing, Greece, 2011*, pp. 280-287.
[4] V. Bína, *Multidimensional Probability Distributions: Structure and Learning: Disertation Thesis*. Faculty of Management, VŠE Praha, 2011.
[5] V. Bína and R. Jiroušek, "Marginalization in Multidimensional Compositional Models," *Kybernetika*, vol. 42, no. 4, pp. 405-422, 2006.
[6] V. Bína and R. Jiroušek, "On Computations with Causal Compositional Models," *Kybernetika*, vol. 51, no. 1, pp. 525-539, 2015.
[7] V. Bína, J. Přibil, "Study Results at Faculty of Management in Jindřichův Hradec," in *Proceedings of the 35th International Conference on Mathematical Methods in Economics, MME 2017, Hradec Králové, Czech Republic*, accepted for publication.
[8] L. DeAngelo, R. Franke, S. Hurtado, J. H. Pryor and S. Tran, *Completing college: Assessing graduation rates at four-year institutions*. Higher Education Research Institute, UCLA, Los Angeles, 2011.
[9] P. Doucek, M. Maryška, "Are Applicants for Programs in Economics Good Both in Mathematics and in English?" in *Turkish Online Journal of Educational Technology, INTE 2016*, pp. 1428-1435.
[10] R. Jiroušek, "Foundations of Compositional Model Theory," *International Journal of General Systems*, vol. 40, no. 6, pp. 623-678, 2011.
[11] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*. Massachusetts Institute of Technology, 2009.
[12] M. Maryška, P. Doucek and H. Mikovcová, "Mathematics knowledge analysis of registered students at the Universite of Economics, Prague," in *Efficiency and Responsibility in Education, Prague, Czech Republic, 2013*, pp. 405-411.
[13] I. Míková, P. Pudil and L. Komárková, "Adopting Proper Methods For Studying Management In Post-Transformation Economies And Their Effect On The Employment Rate," in *Turkish Online Journal of Educational Technology, November 2016*, pp. 218-226.
[14] R. E. Neapolitan, *Learning Bayesian Networks*. Prentice Hall, 2003.
[15] J. Pearl, *Causality: Models, Reasoning, and Inference*. Cambridge University Press, New York, 2009.
[16] J. J. Vossensteyn, A. Kottmann, B. W. A. Jongbloed, F. Kaiser, L. Cremonini, B. Stensaker, E. Hovdhaugen, S. Wollscheid, *Dropout and completion in higher education in Europe: main report*. European Union, 2015.
[17] N. Wermuth, "Model search among multiplicative models," *Biometrics*, vol. 32, pp. 253-263, 1976.

ISBN: 978-988-14047-5-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)