9th BILETA Conference
The Changing Legal Information Environment

11th & 12th April 1994
Scarman House
University of Warwcick
Coventry

Statistically Analysing Court Decisions on Custody Disputes

P.A.W. Piepers and C.J.M. Combrink-Kuiters

Keywords: Jurimetrics - Legal Decision-making - Statistical Prediction Models - Custody Disputes

Abstract: An examination of custody disputes subject to Dutch family law. The use of statistical analysis via linear regression and cross validation in order to determine the correlation between facts presented during a custody dispute and the decision reached by the judge or judges presiding.

This paper describes the progress of a jurimetrical research project carried out by the authors and presents the results of same showing that the above-noted techniques can give a high percentage of correct predictions.

Introduction

This paper aims to determine whether there is a significant relationship between the presence or absence of facts adduced in custody disputes, and the decision made by a judge or panel of judges at the conclusion of the dispute.

The statistical techniques of linear regression and cross validation have been used to determine these relationships.

This paper will examine the following areas:
(i) Dutch law relating to custody disputes;
(ii) The methods used to obtain and prepare data to be used in the statistical analysis;
(iii) An explanation of the technique of linear regression together with a listing of the most significant correlations between all presented facts and decisions;
(iv) Results of the Analysis;
(v) Constraints on the analysis.

Custody disputes according to Dutch family law

The subject of this paper is custody disputes according to Dutch family law. According to Dutch law (Dutch Civil Code, first book, title 9 and 15 (BW, boek 1, titels 9 en 15)) all minors, must be subject to authority. During a marriage, both parents are jointly responsible for their minor children; together they hold parental authority. Divorce procedure usually results in the judge granting custody to one of the parents. The parent who has been granted custody of the child is responsible for bringing up and caring for that child. He/she is the legal representative of the child and may have to manage the assets of the child. The other parent is normally awarded a passive guardianship and rights of access to the child. The parent in the role of passive guardian will monitor this financial management in addition to the child's general up-bringing. The non-custodial parent can request the Court to revoke

its original decision and grant custody in his/her favour. He has to state and prove "changed conditions" since the last decision. When the custodial parent does not want to give up custody the judge has to make a choice between the two parents. According to Dutch law a parent can appeal to a higher court and after that even to the Supreme Court when he disagrees with the verdict. For more information on custody according to Dutch law see Poek, 1984] and [Pitlo, 1989] (both in Dutch), [Fokkema, 1978] and [Freeman, 1984].

Collection and preparation of data to be used in the analyses

Following the choice of subject area it is necessary to collect a number of cases to be used in the analysis. For practical reasons it was decided to start with verdicts that had been published in a legal journal called "de Nederlandse Jurisprudentie", a collection of case law publications which are, for one reason or another, of interest to a larger public. The cases - usually verdicts of the Supreme Court - are stored on the Kluwer CD-ROM Databank and therefore easily obtainable. Dutch case law has been stored in this databank from 1965 onwards. By using the incorporated search method it is possible to retrieve relevant documents using key words. The subject "parental custody disputes after divorce according to Dutch Family Law" provided a sufficient number for analysis. Each of the thirty-five cases selected could be analysed using a variety of perspectives. However, the perspective chosen in this instance is whether custody is granted to gather or mother.

This perspective is fixed by defining the legal item in terms of *the outcome for the father* in custody disputes:

To what extent in a final hearing was the father able to quash a custody order in which the mother had been granted custody or in which the mother's custody was affirmed, or to what extent was he, the father, able to prevent the overturnmg of an order in a final hearing in which he had been granted custody or in which his custody had been affirmed.

In order to establish which facts caused the judge or panel of judges to reach the decisions, the texts of all cases had to be read and analysed. Whilst reading and re-reading these verdicts as many facts as possible were listed. This was a time-consuming task because after introducing a new fact all previous verdicts had to be re-read in order to establish whether this fact was present or not. In the future an attempt will be made to use the computerised format in the coding of the facts. This would be a special application of "conceptual data retrieval" [Wildemast, 1992].

In total 337 possibly important facts were determined divided in two types of facts: dichotomous and numeric facts. Most facts were of the dichotomous type and were allotted a code of "+1" in case the fact was present, a code "-1" if the fact was absent and a code of "0" if it could not be determined whether the fact was present or absent. The numeric facts kept their numeric value. All dichotomous facts which had a code of "+1" or "-1" less than three times, or over 32 times out of 35 cases, were removed because they were regarded as too rare or too common. (For the statistical analysis, too rare and too common are similar obstacles.) After the removal, 148 facts remained on the list.

The court decisions on the legal item were encoded in a similar way. When the outcome for the father was positive a code of "+1" was given, otherwise a code of "-1" was given.

This procedure resulted in a data matrix suitable for statistical analyses. Figure 1 shows the matrix with case nos. shown in columns and the facts in rows. For each case the code for the court's decision is shown as 'Y'.

	Y	X1	X2	X3	..	X146	X147	X148
Case1								
Case2								

Case3								
.....								
Case33								
Case34								
Case35								

Figure 1

Linear Regression

A linear statistical method that has been used in the past by De Mulder [De Mulder, 1984] which uses all the information available and has, therefore, been used as a first statistical technique herein. In short, the model is based on the following idea: a line of regression is calculated for the verdict (Y) and each fact (X1 .. X148). The formula for such a line of regression is:

$$Y=a+bX$$

In this formula [Runyon, 1980] *a* represents a constant and *b* the slope of the line relating values of X to values of Y. The number of lines of regression equals the number of facts in the data set, leading to 148 lines of regression in the used data set. Besides the constant *a* and the slope *b* the Pearson correlation coefficient was calculated for each fact and decision, leading to 148 correlation coefficients. A list holding the number of the fact, constant *(a),* slope *(b)* and the correlation coefficient was subsequently ordered on the absolute value of the correlation coefficient. A strong positive correlation is statistically as significant as a strong negative correlation, the sign only points to the direction of the relationship.

Figure 2 shows all facts in which the correlation is statistically significant. To be statistically significant in a data set of 35 cases the correlation must exceed the value of 0.28, [Ferguson, 1971] and [Wijvekate, 1979] (in Dutch). Out of the 148 correlation coefficients, 20 proved to be statistically significant:

Factno.	slope (b)	constant (a)	correlation
65	0.6515	-0.0152	0.6260
56	0.5189	-0.0644	0.4986
113	0.7097	0.2903	0.4673
64	0.4216	-0.2451	0.4361
87	0.6875	0.3125	0.3983
50	-0.3791	-0.2680	-0.3921
49	-0.3791	-0.2680	-0.3921
61	0.500	0.1000	0.3621
67	0.3939	-0.0996	0.3540
27	-0.4483	-0.5517	-0.3497
117	0.4634	0.0209	0.3283
148	-0.0303	0.2874	-0.3162
59	-0.4333	-0.5667	-0.3138
18	-0.4333	-0.5667	-0.3138

109	-0.4333	-0.5667	-0.3138
4	0.3000	-0.300	0.3073
3	-0.3000	-0.3000	-0.3073
2	0.0316	-0.6833	0.2906
20	0.4274	0.0726	0.2814
107	0.4274	0.0726	0.2814

Figure2

To give an impression of what these facts stand for some descriptions are given. For example, facts 64 and 65 are positive decisions for the father in the court and appeal court respectively. Fact 61 stands for the father referring to the uncertain future of the mother. Fact 87 stands for the willingness of the father to cooperate in a scheme for the mother visiting the child. The facts 18 and 27 stand for specific judges participating in the decision-making. In Supreme Court cases there are 5 judges on the panel.

Exactly what it means for these facts to have a statistically significant correlation coefficient is not clear at this point in the analysis. A closer look must be taken at the interaction between the facts themselves. It is likely that some facts have a strong correlation between themselves, which could lead to a further reduction of the number of relevant facts, since facts correlating to a high degree usually represent the same information. Also the statistical technique of factor analysis could be applied, but this causes problems in interpreting the new constructed facts.

For the technique of lines of regression all 148 constants and slopes were used. Although the correlation is not statistically significant, the slopes and constants still have influence on predicting or extrapolating a case. The verdict of a court case is calculated by filling in the ,values of each X for each case (1..35). In formula:

$$Y_1=(a_1 +b_1X_{11})+ (a_2+b_2X_{12})+ .. + (a_{m-1} + b_{m-1}X_{im-1}) + (a_m + b_mX_{im}) \text{ for } i = 1, 2,.., 35$$

If the calculated values of Y, here called the interpolations, are ordered, it is possible to determine the point in the list where the positive court decisions can be distinguished from the negative court decisions. The idea behind this statistical method is that the high interpolations correspond with the positive decisions while the low interpolations correspond with the negative decisions. An example of such a list is shown in Figure 3. The optimal distinction (in the list represented by a line) between the positive and negative decisions is the one that leads to the minimum number of incorrectly extrapolated decisions. The column 'Decision' gives the actual decisions made:

Case	Decision	Interpolation	Correct
1	1	-0.1186	1
4	1	-0.1285	1
15	1	-0.1399	1
6	1	-0.1419	1
14	1	-0.1552	1
35	1	-0.1581	1
23	1	-0.1594	1
18	1	-0.1632	1
11	1	-0.1664	1
16	1	-0.1684	1

5	1	-0.1743	1
19	-1	-0.1849	-1
26	1	-0.1998	1
13	1	-0.2008	1
---	---	---	---
22	-1	-0.2037	1
32	-1	-0.2062	1
7	-1	-0.2075	1
33	-1	-0.2098	1
2	-1	-0.2145	1
31	-1	-0.2167	1
17	-1	-0.2168	1
10	-1	-0.2224	1
3	-1	-0.2225	1
8	-1	-0.2244	1
20	-1	-0.2252	1
25	-1	-0.2317	1
9	-1	-0.2336	1
12	-1	-0.2402	1
24	-1	-0.2413	1
30	-1	-0.2445	1
21	-1	-0.2545	1
34	-1	-0.2552	1
28	-1	-0.2629	1
29	-1	-0.2646	1
27	-1	-0.2661	1

Figure3

Cases above the distinction line are extrapolated positively, whilst decisions below the distinction line are extrapolated negatively. In the column 'Correct', a correctly extrapolated court decision is coded with a "1" and an incorrectly extrapolated court decision with a "-1". In the given table, only case 19 is predicted incorrectly. Thus, in 34 out of 35 cases (97%) the verdicts are predicted correctly.

Case 19 is a peculiar case. The father went to court and asked the judge to change custody in his favour. His request was granted and the child went to live with the father and his family. After a period of approximately 5 years the mother asked for custody to be changed in her favour. During these 5 years, however, the child became attached to her stepmother. The Council for Child Care and Protection recommended that custody should not be changed again. The judge followed this advice. Upon appeal this verdict was overruled and this higher court's finding upheld by the Supreme Court. The father was accused of frustrating the mother's right of access to the child, but he had only started to do so at the time the mother asked for the change of custody. In the author's opinion the verdict could have been reversed: in favour of the father.

Cross validation algorithm

A possible objection to the method described above is that each case contributes to its own extrapolation. It might even be expected that this contribution is substantial, because of the ratio between the number of cases and the number of facts. To find out whether the results found are valid a cross validation method has been developed and applied.

Each case can only be decided by the judge or panel of judges in two ways: positive or negative. This phenomenon was used for the cross validation method. For each case it was assumed that the actual decision was unknown. Then the statistical method of regression lines was applied to a data matrix with the case positively decided and to a data matrix in which the case was negatively decided. Both calculations give a list of extrapolations as above. To determine which decision fits best, both extrapolation lists are compared by applying the cross validation algorithm. The algorithm for selecting the most likely decision uses three criteria. The first criterion is the overall consistency of the list of extrapolations. In other words, how many cases are predicted correctly? The second criterion is whether the case itself is predicted correctly. The third criterion calculates how evident the prediction of the case itself is. Cases closer to the distinction line are not as strong as cases further away from the distinction line. The stronger a case the better. A value of "1" was given to the distance of the two cases the furthest away above and below the distinction line, while other cases got a value for the distance related to the actual position in the list and the number of cases above or below the distinction line. If these three criteria are not decisive, this will be reported.

The above idea has been applied in a slightly different manner to the data set. In the left column the actual decision made is always shown, while in the right column the situation is shown if the decision of the case would have been the reverse. Then the cross validation algorithm is applied and a conclusion is drawn which situation would be preferred according to the above mentioned three criteria. If the situation is preferred in which the reversed decision is used this is counted as a wrongly predicted case. In Figure 4 an example of the outcome of the cross validation algorithm for case 9, (marked in both columns with a *), is shown. In this case the decision would be predicted correctly according to the first criterion:

Constraints and further research

The general idea is that the constructed model correctly predicts as many court decisions as possible. The best possible result (a hundred percent correct predictions), however, is not likely on theoretical grounds. A model which always predicts correctly would imply that judges do not make mistakes and that all relevant facts have been included in the model. It is always possible to have such a unique fact in a case that a deviant decision is justified. It is equally difficult to assess an absolute percentage of correct predictions which a model should be able to provide. Only a lower boundary can be given: the a priori probability for a positive or negative decision. The terms "positive" and "negative" should be interpreted in relation to the legal item that is being analysed: "Does the judge decide in favour of the father?" From the above table it becomes clear that the number of correct predictions is 17 per cent higher than the a priori probability.

Case	Decision	Interpolation	Correct	Case	Decision	Interpolation	Correct
1	1	-0.1186	1	1	1	-0.0807	1
4	1	-0.1285	1	4	1	-0.0899	1
15	1	-0.1399	1	15	1	-0.1035	1
6	1	-0.1419	1	6	1	-0.1046	1
14	1	-0.1552	1	14	1	-0.1115	1
35	1	-0.1581	1	23	1	-0.1201	1
23	1	-0.1594	1	18	1	-0.1229	1
18	1	-0.1632	1	35	1	-0.1241	1

11	1	-0.1664	1	11	1	-0.1279	1
16	1	-0.1684	1	5	1	-0.1326	1
5	1	-0.1743	1	16	1	-0.1391	1
19	-1	-0.1849	-1	19	-1	-0.1546	-1
26	1	-0.1998	1	13	1	-0.1581	1
13	1	-0.2008	1	9	1	-0.1589	1
---	---	---	---	---	---	---	---
22	-1	-0.2037	1	3	-1	-0.1601	1
32	-1	-0.2062	1	2	-1	-0.1605	1
7	-1	-0.2075	1	26	1	-0.1636	-1
33	-1	-0.2098	1	33	-1	-0.1677	1
2	-1	-0.2098	1	22	-1	-0.1680	1
31	-1	-0.2167	1	32	-1	-0.1694	1
17	-1	-0.2167	1	20	-1	-0.1762	1
10	-1	-0.2224	1	7	-1	-0.1767	1
3	-1	-0.2225	1	17	-1	-0.1796	1
8	-1	-0.2244	1	31	-1	-0.1808	1
20	-1	-0.2252	1	10	-1	-0.1830	1
25	-1	-0.2317	1	8	-1	-0.1833	1
9	-1	-0.2336	1 *	25	-1	-0.1904	1
12	-1	-0.2402	1	24	-1	-0.1939	1
24	-1	-0.2413	1	30	-1	-0.2032	1
30	-1	-0.2445	1	12	-1	-0.2062	1
21	-1	-0.2545	1	34	-1	-0.2114	1
34	-1	-0.2552	1	21	-1	-0.2156	1
28	-1	-0.2629	1	28	-1	-0.2261	1
29	-1	-0.2646	1	29	-1	-0.2298	1
27	-1	-0.2661	1	27	-1	-0.2334	1

Figure 4

Cases predicted incorrectly: 1
Case itself predicted correctly
Distance to distinction line: 0.62

Cases predicted incorrectly: 1
Case itself predicted correctly
Distance to distinction line: 0.07

Conclusion: The original case 9 is preferred

Once this was done for all cases the following overall results were achieved:

Correct predictions : 29 82.86%
Incorrect predictions : 6 17.14%

Not predicted : 0 0.00%

Total number of cases : 35
Total number of facts : 148

A priori percentage positively decided cases . 34.29%
A priori percentage negatively decided cases . 65.71%
Percentage gain of correct predictions . 17.14%

Nonetheless, a significantly higher percentage of correct predictions than the a priori probability of the most likely decision does not say everything about the reliability of the model as it is. Not all relevant facts may have been stated by the courts in all their verdicts. The coding of the cases and the facts into the data matrices is done by humans, which implies subjective decisions and possible mistakes, which may mean that the data used is inaccurate and possibly unsuitable for the statistical calculations. The "significantly higher percentage of correct predictions" would then be based on unreliable data.

A theory must be developed which explains why the facts found are important for the court decision in those types of cases and why to that extent. This explanatory theory does not have to be strictly legal. From a jurimetrical point of view [Loevinger, 1949; De Mulder, 1984 and De Mulder, 1988] it is possible that courts also weigh legal or non-legal facts in their decision, which are not mentioned in the text of any verdict. In this kind of jurimetrical research it is important to find out which facts judges have actually used in reaching a verdict [Bell, 1988]; [Hogarth, 1971]; [Kate, 1984]; [Leith, 1991]; [Lindley, 1991]; [Malsch, 1989]; [Winterfeldt, 1986]; and [Wright, 1987]. If it is possible to find an explanatory theory for the model, the conclusion that the model is valid has more meaning than when the model is just the best possible mathematical way of predicting the court decision. This is the direction in which the author's research will move in future. More research into the facts and their influence will be undertaken.

Conclusion

The statistical analysis described in this paper provides a strong tool for predicting a court decision on the basis of the analysis of the texts of verdicts on a chosen legal subject. This statistical technique combined with the cross validation method makes it possible to predict a higher percentage than the a priori probability of the most likely court decision correctly. The algorithm could be used in legal knowledge systems. It would not only provide a "yes" or "no" prediction, but it would also give an estimation of the relative strength of a case. A case is stronger if its prediction is high or low in the order of extrapolations, i.e. further away from the distinction line. Future research will focus on the relationship between facts themselves and on why the influence of facts on the decisions.

References

Bell, DE., Ralifa, H. and Tversky, A. (eds), (1988). Decision Making, Cambridge University Press, Cambridge.

Combrink-Kuiters, C.J.M. and P.A.W. Plepers, (1993). 'The implementation of Predictive Capabilities into Legal Computer Advice Systems', paper for the 8th BILETA Conference, Building Systems, 1-2 April, John Moores University, Liverpool, 1993. Published in: Pre-Proceedings of the Conference, pp.63-72.

De Mulder, R.V., (1984). Een model voor juridische informatica (A model for the application of

computer science to law), Vermande, Lelystad.

De Mulder, R.V., (1988). Juridische Informatica: in de eerste plaats een hulpmiddel bij de jurimetrie (Legal Informatics: in the first place an aid to jurimetrics), in: Berg, P. van den, et al (Eds), RI-paradigmata: Toekomst van Nederlands onderzoek in de rechtsinformatica (Paradigms in legal informatics: Exploring the future of Dutch research programmes in legal informatics), Vermande, Lelystad.

Doek, J.E. and Drewe:, H.M.A., (1984). Jeugdrecht en jeugdbeschermingsrecht in Nederland, Loghum Slaterus, Deventer.

Ferguson, GA., (1971). Statistical Analysis in Psychology and Education, 3th Edition, McGraw-Hill, Tokyo.

Fokkema, D.C. (ed), (1978). Introduction to Dutch Law for Foreign Lawyers, Kluwer, Deventer.

Freeman, M.D.A. (ed), The State, the Law, the Family, Travistock Publishers, London, New York.

Nogaith, J., (1971). Sentencing as a Human Process, Joronto Press, Toronto.

Kate, J. ten and P.J. van Koppen, (1984). Determinanten van privaatrechtelijke beslissingen

(Determinants of private law decision-making), Gouda Quint, Arnhem.

Leith, P., (1991). The Computerised Lawyer: A Guide to the Use of Computers in the Legal Profession, Springer Verlag, London.

Lindley, D.V., Making decisions, second edition, John Wiley and Sons, London.

Loevinger, L, (1949). tjurimetrics: the next step forward', Minnesota Law Review, Volume 33.

Malsch, M., (1989). Lawyers' predictions of judicial decisions, PhD thesis, University of Leiden, the Netherlands.

Piepers, P.A.W., C.J.M. Combrink-Kulters and R.V. De Mulder, (1993). 'Towards a Model for Knowledge Representation; Predicting Court Decisions on Custody Disputes', paper for the Conference in Celebration of the 25th Anniversary of the Istituto per la documentazione giuridica of the Consigho Nazionale delle Ricerche, 1-3 December, Florence, Italy. Published in: Pre-Proceedings of the Conference, pp.159-160.

Pitlo, A., (1989). Het Nederlands B.W. deel 1, Het personen- en familierecht, Gouda Quint, Arnhem.

Runyon. R.P., and Haber, A., (1980). Fundamentals of behavioral statistics, fourth edition, Addison-Wesley, Reading, Massachussets.

Wijvekate, M.L., (1979). Verklarende statistiek, UtrechtlAntwerpen.

Wildemast. C.A.M., and De Mulder, R.V., (1992). Some Design Considerations for a Conceptual Legal Information Retrieval System. In: Grütter, C.A.F.M., Breuker, J.A.P.J., Herik, H.J. van den, Schmidt, A.H.J., and Vey Mestdagh, C.N.J. de (eds), Legal Knowledge Based Systems: Information Technology and Law, JURIX, Vermande, Lelystad,.

Winterfeldt, D. von and Edward, W., (1986). Decision Analysis and Behavioral Research, Wright,

G. and Ayton, P. (eds), (1987). Judgmental Forecasting.