ON THE SECURITY OF CELL SUPPRESSION IN CONTINGENCY TABLES WITH QUANTITATIVE FACTORS

Josep Domingo i Ferrer Josep M. Mateo Sanz

Universitat Rovira i Virgili*†
Tarragona, Catalonia

Abstract

We show in this paper how to break the security of cell suppression methods when these are used to anonymize contingency tables with at least one quantitative factor. The procedure combines distribution fitting and linear programming and may obtain interval estimates for the suppressed cells which are narrower than the disclosure interval chosen in the suppression method.

1. INTRODUCTION

This work analyzes the security of cell suppression methods when they are used to guarantee anonymity in contingency tables having at least one quantitative factor. Given a table with suppressed cells, we try to determine interval estimates for such cells by adjusting a probability distribution for the quantitative factor. For each cell, if the interval estimate obtained in this way is narrower than the interval that results from a linear programming strategy, then we add the new interval as a supplementary constraint. When this procedure has been completed for all cells, we start again using the linear programming strategy with the marginal sum constraints and the new interval constraints for the suppressed cells. This procedure may break the security of the suppression method by providing interval estimates for the suppressions that are narrower than the disclosure interval (see a definition of this concept in Cox, 1993 or in section 2 below) used in the suppression method.

In section 2 we recall some background on cell suppresion methods. In section 3 we sketch our attack. Section 4 illustrates our approach on a practical example. Section 5 contains some conclusions.

2. BACKGROUND ON CELL SUPPRESSION

In this section we recall the basic principles of cell suppression methods, following Cox (1993).

*Autovia de Salou, s/n, E-43006 Tarragona, Catalonia, Spain. E-mail {jdomingo,jmateo}@etse.urv.es
†This work is partly supported by the Spanish CICYT under grant no. TIC95-0903-C02-02.

Cell suppression belongs to the family of disclosure control methods relying on query set size control (see Schackis, 1993). The basic principle is to eliminate the so-called *disclosure cells*, which are cells that could lead to disclosure of individual data. In contingency tables, a common criterion is to consider as disclosure cells those containing statistics computed over less than t individuals, where usually $t = 3$ or $t = 4$. In the special case where cells contain frequencies, disclosure cells are those containing frequencies lower than t. Once a disclosure cell X has been identified, a *disclosure interval* $I(X)$ is determined. The security target of the cell suppression method is not only to prevent exact disclosure of the value of X, but to ensure that $I(X)$ is a sub-interval of all interval estimates of the value of X that are obtainable from the contingency table.

The first step in the suppression method is to suppress disclosure cells. Suppressions at this step are called primary suppressions. Due to linear relations between cell values (basically the marginal sums), additional non-disclosure cells must be usually suppressed to provide the required interval protection to X. Thus the second step consists of determining and suppressing these additional cells. Suppressions at this step are called secondary suppressions. Given a set of disclosure cells each with its disclosure interval, the way to determine the secondary suppressions is not unique. There are several criteria which aim at minimizing the loss of data quality. The most popular criteria are

- Choose secondary suppressions such that the number of suppressed cells is minimal.
- Choose secondary suppressions such that the overall suppressed value is minimal.

Checking that a secondary suppression is appropriate is not trivial, but requires a linear programming approach. Each suppressed cell is viewed as a variable. For each disclosure cell X primarily suppressed, $\min(X)$ and $\max(X)$ are computed over the linear system defined by the aggregation equations (marginal sums) and the variables representing suppressed cells. Then the secondary suppression is appropriate if and only if for each disclosure cell the interval $I(X)$ is contained in the interval $[\min(X), \max(X)]$. The following is a two-dimensional example from Cox (1993) that clearly illustrates this issue.

X_{11}	X_{12}	X_{13}	10	20	80
10	X_{22}	X_{23}	5	15	60
X_{31}	10	10	X_{34}	10	90
X_{41}	5	15	X_{44}	5	40
75	35	65	45	50	270

In the original table cells containing statistics computed over 1 or 2 individuals were considered disclosure cells. Imagine that X_{11} is a disclosure cell whose disclosure interval is $I(X_{11}) = [15, 25]$. One might think that X_{11} is well protected because its row and column contain three suppressions each. Using the row and column equations only, it turns out that the best interval estimate for the cell is $0 < X_{11} < 50$, which contains $I(X_{11})$. However, using linear programming we get $\min(X_{11}) = \max(X_{11}) = 20$, so $X_{11} = 20$ is revealed, which means exact disclosure.

Several algorithms have been proposed to determine the optimal set of secondary suppressions. Besides the security limitations that will be discussed in this paper, all known approaches share computational limitations which make them impractical for k-dimensional tables when k is large. In Cox (1993) a solution is described which relies on network optimization. In Gopal *et al.* (1995) several heuristic algorithms are given to accomplish the suppression task not only for contingency tables, but also for statistical databases. A branch-and-cut algorithm is proposed in Salazar and Fischetti (1995); this algorithm has been used in the example of section 4. Although computationally different, all the above solutions are functionally equivalent and can be used to minimize the number of suppressed cells or the overall value suppressed.

3. DESCRIPTION OF THE ATTACK

Let us assume that we have an anonymized contingency table with at least one quantitative factor. More precisely, let us assume that there are k factors $F_1, \cdots, F_k$, of which the first q factors are quantitative. Some cells in the table have been suppressed after using a cell suppression method, and there is no information on which suppressions are primary or secondary. Then, the following algorithm is proposed to determine interval estimates for suppressed cells that, in the case of disclosure cells, may be narrower than the disclosure interval chosen at the time of anonymization.

Algorithm 1

1. *Considering each suppressed cell as a variable X, compute by linear programming* $\min(X)$ *and* $\max(X)$ *subject to the aggregation equations (marginal sums) in terms of the other suppressed cells.*

2. *For each quantitative factor F_i, from $i = 1$ to q*

   (a) *For each combination c of levels of factors F_j, for $j \neq i$, consider the row that is obtained when F_i varies. If the proportion of suppressed cells in this row is higher than a preselected threshold then discard the row; otherwise, try to fit a distribution to data in the row. We assume at this step that marginal sums are known, which means that the* total *value of suppressed cells in the row is known and used as a single cell at the time of distribution fitting. Several distributions can be tried whose parameters are determined in such a way that the χ^2 goodness-of-fit statistic is minimized. This guarantees the interesting property that the χ^2 statistic approximately follows a χ^2 distribution with $n - s - m$ d. f. where n is the total number of cells in the row, s is the number of suppressed cells in the row and m is the number of parameters of the distribution (see Moore, 1986 and Fisher, 1924). The best-fitting distribution is considered for subsequent computation.*

   (b) *For each fitted row, take the significance level α_c of the χ^2 statistic under the χ^2_{n-s-m} distribution.* Assume *that if the same distribution model were fitted to the original row (without suppressions), the significance level would have as expected value α_c, but under χ^2_{n-m-1}. This assumption provides a critical value $\chi^2_{\alpha_c,n-m-1}$. Now, since the original row is unknown, the value of the χ^2 goodness-of-fit statistic for that row can be viewed as a random variable having a χ^2-like distribution with expected value the critical value $\chi^2_{\alpha_c,n-m-1}$. If this expected value is an integer, then the* *distribution of the goodness-of-fit statistic is a* χ^2 *with* $\chi^2_{\alpha_c,n-m-1}$ *degrees of freedom. In the general case, the distribution the goodness-of-fit statistic is a gamma with parameters* $\alpha = \frac{\chi^2_{\alpha_c,n-m-1}}{2}$ *and* $\beta = 2$*. The distribution of the goodness-of-fit statistic for the original row allows to determine a confidence interval for this statistic.*

(c) *For each confidence interval obtained at substep 2b, compute the set of value assignments to suppressed cells in the corresponding fitted row that yield a goodness-of-fit statistic value in the confidence interval and satisfy the marginal sum of the row. If there are* s *suppressed cells, it follows from its defining conditions that the set of feasible assignments can represented in an* s*-dimensional Euclidean space as the intersection of an ellipsoidal* s*-annulus (space between two concentric* s*-ellipsoids) and a hyperplane. For each suppressed cell, this set of feasible assignments projects a feasible value set which is usually a single interval, but in rare cases may consist of two intervals.*

(d) *For each suppressed cell, compare the feasible interval obtained at substep 2c with the interval obtained at step 1 for the same cell. Discard the former if it contains the latter. Otherwise take the intersection of both intervals as a supplementary constraint for the suppressed cell variable and concatenate the new constraint to the marginal sum constraints.*

3. *If no new constraints have been found during step 2, then exit the algorithm. Otherwise consider each suppressed cell as a variable* X *and compute by linear programming* $\min(X)$ *and* $\max(X)$ *subject to the aggregation equations (marginal sums) and the constraints found at step 2. Output* $[\min(X), \max(X)]$ *(if the feasible value set for* X *found at substep 2d consists of two intervals, then the intersection of the previous interval and that value set is output).*

The above algorithm *may* output a feasible interval for a primarily suppressed cell X that is contained in the disclosure interval $I(X)$ for that cell; this may happen as a result of step 2 (distribution fitting) or as a result of step 3 (distribution fitting plus linear programming). If a feasible interval contained in a disclosure interval is obtained, then the security of the cell suppression algorithm is *probably* broken. We say "probably" because, unlike disclosure intervals, the feasible intervals output by the algorithm are based on confidence intervals and are thus probabilistic.

Central to algorithm 1 is the following natural assumption, which is supported by extensive tests on real data rows

Assumption 1 *Consider a row where the proportion of suppressed cells is small. If the* χ^2 *goodness-of-fit statistic results in an* α_c *significance level when fitting a distribution model to the non-suppressed cells in a row, then the goodness-of-fit statistic if the same distribution model were fitted to the complete original row is expected to yield about the same significance level* α_c*. Plainly speaking, the goodness-of-fit measure for the whole row is expected to be similar as the goodness-of-fit measure for the non-suppressed part of the row.*

Of course, the above assumption is no longer reasonable when the proportion of suppressed cells in a row is too high. Given a row size, the threshold used in substep 2a of algorithm 1 is a parameter to be chosen.

The width of the feasible interval for a cell obtained at substep 2c depends on two things, of which the first is more influential:

- The expected value of the suppressed cell according to the fitted distribution. The smaller the expected value, the wider the feasible interval.

- The significance level α_c determined at substep 2b. The larger α_c, *i. e.* the better the fit, the narrower the feasible interval.

4. A PRACTICAL EXAMPLE

Table 1 (see end of paper) is a contingency table containing the count of farms in Catalonia by cultivated area (in hectares) and district (see IEC, 1995, p. 227). This table is two-dimensional and the cultivated area is a quantitative factor; thus it is suitable to illustrate our attack. Imagine that table 1 is not disclosure-protected and is to be protected using a cell suppression method. So, we anonymize table 1 by using a cell suppression software by the authors of Salazar and Fischetti (1995), with the following options:

- Primary disclosure cells are those containing a count less than four and greater than zero. They are circled in table 1. Zero counts are not considered as disclosure cells.

- The disclosure interval for a disclosure cell with a count x is $I(x) = [1, 2x]$.

- Secondary suppressions are determined so as to minimize the total value of suppressed cells. Like for primary suppressions, zero counts have not been considered for secondary suppression.

Table 2 is the output of the cell suppression procedure. At the place of each suppressed cell, we show the interval computed by linear programming at step 1 of algorithm 1. Since the statistics in the original table are frequency counts, throughout this example "interval" is equivalent to the set of integers comprised between both ends of the interval.

Now step 2 of algorithm 1 is run for the only quantitative factor in the table. Rows with more than *two* suppressed cells are discarded and the rest of rows are fitted to some distribution. Distributions tried include Weibull, beta, gamma, exponential, Snedecor's F, lognormal, normal, Poisson and binomial. The best-fitting distribution for each row and the significance level of the goodness-of-fit statistic can be summarized as follows:

Row	Distribution	G-of-fit sig. level
Alt Camp	Weibull	10^{-6}
Alta Ribagorça	lognormal	0.07
Baix Llobregat	lognormal	0.01
Baix Penedès	lognormal	0.43
Garraf	lognormal	0.99
Priorat	lognormal	10^{-6}
Tarragonès	lognormal	0.12

The only row with a really high goodness-of-fit is the one corresponding to the Garraf district, but this does not prevent from using all rows. Assumption 1 is independent of whether the goodness-of-fit is high or low. Now table 3 shows the feasible intervals obtained at the end of step 2 for each suppressed cell in the fitted rows (95% confidence intervals have been used in substep 2b). In all cases, a *single* interval per cell has been found. The circled intervals improve on the intervals found at step 1 and shown in table 2. Therefore, step 3 is run. The final intervals for all suppressed cells are shown in table 4. By comparing this table and table 1, it can be seen that, in the Baix Llobregat row, the interval $[1, 3]$ is contained in the disclosure interval $[1, 4]$ corresponding to the disclosure cell with value 2. Also, in the Baix Penedès row, the interval $[1, 5]$ is contained in the disclosure interval $[1, 6]$, corresponding to the disclosure cell with value 3. The security of the cell suppresion method is probabilistically broken in both cases. Besides, all circled intervals in table 4 are contained in the corresponding intervals found at step 1 by pure linear programming.

5. CONCLUSION

A methodology of attack to the security of cell suppression methods has been presented that can be used on contingency tables with at least one quantitative factor. The idea is to combine distribution fitting and linear programming, which may yield probabilistic intervals for primarily suppressed cells that are contained in their disclosure interval. The attack has reasonable chances of success, and it almost surely improves on the intervals for suppressed cells that can be obtained by straightforward pure linear programming. Therefore, if a contingency table has one or more quantitative factors, then it is wise to use a disclosure control method other than cell suppression.

REFERENCES

Cox, L. H. (1993) "Solving confidentiality protection problems: a model for cell suppresion in U.S. economic censuses", in D. Lievesley (Ed.), *Proceedings of the First International Seminar on Statistical Confidentiality*, Luxembourg: Eurostat, pp. 229-245.

Fisher, R. A. (1924) "The conditions under which χ^2 measures the discrepancy between observation and hypothesis", *Journal of the Royal Statistical Society* **87**, pp. 442-450.

Gopal, R. D, Goes, P. B., and Garfinkel, R. S. (1995) "Interval protection of confidential information in a database", *Management Science* (submitted).

IEC (1995) *Anuari Estadístic de Catalunya*, Barcelona: Institut d'Estadística de Catalunya.

Moore, D. S. (1986) "Tests of chi-squared type", in R. B. D'Agostino and M. A. Stephens (Eds.), *Goodness-of-Fit Techniques*, New York and Basel: Dekker, pp. 63-95.

Salazar, J. J., and Fischetti, M. (1995) "Técnicas poliédricas para garantizar la privacidad en tablas estadísticas públicas", Univ. de la Laguna-Univ. di Padova (working paper).

Schackis, D. (1993) *Manual on Disclosure Control Methods*, Luxembourg: Eurostat.

	< 1	[1, 2)	[2, 5)	[5, 10)	[10, 20)	[20, 50)	[50, 100)	≥ 100	Total
Alt Camp	727	452	635	435	379	247	27	3	2905
Alt Empordà	865	382	548	548	542	439	78	25	3427
Alt Penedès	395	486	866	606	381	159	39	8	2940
Alt Urgell	162	106	181	233	232	148	41	26	1129
Alta Ribagorça	16	21	50	59	39	39	13	21	258
Anoia	212	196	324	253	270	323	142	28	1748
Bages	651	309	460	338	291	329	86	25	2489
Baix Camp	1243	1095	1658	945	438	111	9	5	5504
Baix Ebre	2248	1843	2803	1445	573	162	20	20	9114
Baix Empordà	452	249	415	380	347	264	31	9	2147
Baix Llobregat	736	521	668	248	88	35	7	2	2305
Baix Penedès	156	203	319	208	138	61	5	3	1093
Barcelonès	32	21	23	10	1	1	0	1	89
Berguedà	251	189	247	207	206	204	40	27	1371
Cerdanya	25	18	32	62	116	192	43	29	517
Conca de Barberà	301	263	479	372	354	339	74	22	2204
Garraf	145	123	174	107	60	34	8	1	652
Garrigues	236	298	745	852	975	540	52	14	3712
Garrotxa	288	231	423	339	214	111	25	12	1643
Gironès	492	239	358	348	265	142	12	7	1863
Maresme	884	662	633	152	49	12	1	2	2395
Montsià	1419	949	1404	932	489	189	20	5	5407
Noguera	342	256	608	661	807	858	217	66	3815
Osona	335	273	512	521	589	335	48	29	2642
Pallars Jussà	64	69	204	253	313	353	76	36	1368
Pallars Sobirà	74	53	183	198	127	110	19	52	816
Pla d'Urgell	299	267	624	690	493	260	27	11	2671
Pla de l'Estany	130	77	167	185	161	112	13	5	850
Priorat	396	390	670	488	263	88	6	1	2302
Ribera d'Ebre	399	507	1007	786	525	192	19	5	3440
Ripollès	113	70	124	153	137	179	97	80	953
Segarra	43	42	116	163	254	584	214	45	1461
Segrià	878	902	1859	1679	1500	806	130	45	7799
Selva	428	281	414	247	197	106	14	4	1691
Solsonès	26	40	81	79	119	265	107	34	751
Tarragonès	753	530	732	421	216	80	22	3	2757
Terra Alta	174	244	609	699	807	338	19	4	2894
Urgell	175	233	527	487	518	567	139	20	2666
Val d'Aran	39	45	151	78	33	1	1	19	367
Vallès Occidental	393	250	267	140	92	76	21	5	1244
Vallès Oriental	656	472	682	373	224	113	18	4	2542
Catalonia	17653	13857	22982	17380	13822	9504	1980	763	97941

Table 1: **Farms in Catalonia by cultivated area (hectares) and district. Source: Institut d'Estadística de Catalunya (1995).**

	<1	$[1,2)$	$[2,5)$	$[5,10)$	$[10,20)$	$[20,50)$	$[50,100)$	≥ 100	Total
Alt Camp	727	452	635	435	379	247	[21,29]	[1,9]	2905
Alt Empordà	865	382	548	548	542	439	78	25	3427
Alt Penedès	395	486	866	606	381	159	39	8	2940
Alt Urgell	162	106	181	233	232	148	41	26	1129
Alta Ribagorça	16	[21,33]	50	59	39	39	[1,13]	21	258
Anoia	212	196	324	253	270	323	142	28	1748
Bages	651	309	460	338	291	329	86	25	2489
Baix Camp	1243	1095	1658	945	438	111	9	5	5504
Baix Ebre	2248	1843	2803	1445	573	162	20	20	9114
Baix Empordà	452	249	415	380	347	264	31	9	2147
Baix Llobregat	736	521	668	248	88	35	[1,8]	[1,8]	2305
Baix Penedès	156	203	319	208	138	61	[1,7]	[1,7]	1093
Barcelonès	32	[9,21]	23	10	[1,13]	[1,12]	0	[1,9]	89
Berguedà	251	189	247	207	206	204	40	27	1371
Cerdanya	25	18	32	62	116	192	43	29	517
Conca de Barberà	301	263	479	372	354	339	74	22	2204
Garraf	145	123	174	107	60	34	[1,8]	[1,8]	652
Garrigues	236	298	745	852	975	540	52	14	3712
Garrotxa	288	231	423	339	214	111	25	12	1643
Gironès	492	239	358	348	265	142	12	7	1863
Maresme	884	662	633	152	49	[1,12]	[1,13]	[1,9]	2395
Montsià	1419	949	1404	932	489	189	20	5	5407
Noguera	342	256	608	661	807	858	217	66	3815
Osona	335	273	512	521	589	335	48	29	2642
Pallars Jussà	64	69	204	253	313	353	76	36	1368
Pallars Sobirà	74	53	183	198	127	110	19	52	816
Pla d'Urgell	299	267	624	690	493	260	27	11	2671
Pla de l'Estany	130	77	167	185	161	112	13	5	850
Priorat	396	390	670	488	263	88	[1,6]	[1,6]	2302
Ribera d'Ebre	399	507	1007	786	525	192	19	5	3440
Ripollès	113	70	124	153	137	179	97	80	953
Segarra	43	42	116	163	254	584	214	45	1461
Segrià	878	902	1859	1679	1500	806	130	45	7799
Selva	428	281	414	247	197	106	14	4	1691
Solsonès	26	40	81	79	119	265	107	34	751
Tarragonès	753	530	732	421	216	80	[16,24]	[1,9]	2757
Terra Alta	174	244	609	699	807	338	19	4	2894
Urgell	175	233	527	487	518	567	139	20	2666
Val d'Aran	39	45	151	78	[21,33]	[1,12]	[1,13]	19	367
Vallès Occidental	393	250	267	140	92	76	21	5	1244
Vallès Oriental	656	472	682	373	224	113	18	4	2542
Catalonia	17653	13857	22982	17380	13822	9504	1980	763	97941

Table 2: **Intervals for suppressed cells found by LP.**

	< 1	[1, 2)	[2, 5)	[5, 10)	[10, 20)	[20, 50)	[50, 100)	≥ 100	Total
Alt Camp	727	452	635	435	379	247	[23,29]	[1,7]	2905
Alt Empordà	865	382	548	548	542	439	78	25	3427
Alt Penedès	395	486	866	606	381	159	39	8	2940
Alt Urgell	162	106	181	233	232	148	41	26	1129
Alta Ribagorça	16	[21,28]	50	59	39	39	[6,13]	21	258
Anoia	212	196	324	253	270	323	142	28	1748
Bages	651	309	460	338	291	329	86	25	2489
Baix Camp	1243	1095	1658	945	438	111	9	5	5504
Baix Ebre	2248	1843	2803	1445	573	162	20	20	9114
Baix Empordà	452	249	415	380	347	264	31	9	2147
Baix Llobregat	736	521	668	248	88	35	[6,8]	[1,3]	2305
Baix Penedès	156	203	319	208	138	61	[3,7]	[1,5]	1093
Barcelonès	32		23	10			0		89
Berguedà	251	189	247	207	206	204	40	27	1371
Cerdanya	25	18	32	62	116	192	43	29	517
Conca de Barberà	301	263	479	372	354	339	74	22	2204
Garraf	145	123	174	107	60	34	[5,8]	[1,4]	652
Garrigues	236	298	745	852	975	540	52	14	3712
Garrotxa	288	231	423	339	214	111	25	12	1643
Gironès	492	239	358	348	265	142	12	7	1863
Maresme	884	662	633	152	49				2395
Montsià	1419	949	1404	932	489	189	20	5	5407
Noguera	342	256	608	661	807	858	217	66	3815
Osona	335	273	512	521	589	335	48	29	2642
Pallars Jussà	64	69	204	253	313	353	76	36	1368
Pallars Sobirà	74	53	183	198	127	110	19	52	816
Pla d'Urgell	299	267	624	690	493	260	27	11	2671
Pla de l'Estany	130	77	167	185	161	112	13	5	850
Priorat	396	390	670	488	263	88	[1,6]	[1,6]	2302
Ribera d'Ebre	399	507	1007	786	525	192	19	5	3440
Ripollès	113	70	124	153	137	179	97	80	953
Segarra	43	42	116	163	254	584	214	45	1461
Segrià	878	902	1859	1679	1500	806	130	45	7799
Selva	428	281	414	247	197	106	14	4	1691
Solsonès	26	40	81	79	119	265	107	34	751
Tarragonès	753	530	732	421	216	80	[16,24]	[1,9]	2757
Terra Alta	174	244	609	699	807	338	19	4	2894
Urgell	175	233	527	487	518	567	139	20	2666
Val d'Aran	39	45	151	78				19	367
Vallès Occidental	393	250	267	140	92	76	21	5	1244
Vallès Oriental	656	472	682	373	224	113	18	4	2542
Catalonia	17653	13857	22982	17380	13822	9504	1980	763	97941

Table 3: **Feasible intervals for suppressed cells in fitted rows.**

	< 1	$[1, 2)$	$[2, 5)$	$[5, 10)$	$[10, 20)$	$[20, 50)$	$[50, 100)$	≥ 100	Total
Alt Camp	727	452	635	435	379	247	[23,29]	[1,7]	2905
Alt Empordà	865	382	548	548	542	439	78	25	3427
Alt Penedès	395	486	866	606	381	159	39	8	2940
Alt Urgell	162	106	181	233	232	148	41	26	1129
Alta Ribagorça	16	[21,28]	50	59	39	39	[6,13]	21	258
Anoia	212	196	324	253	270	323	142	28	1748
Bages	651	309	460	338	291	329	86	25	2489
Baix Camp	1243	1095	1658	945	438	111	9	5	5504
Baix Ebre	2248	1843	2803	1445	573	162	20	20	9114
Baix Empordà	452	249	415	380	347	264	31	9	2147
Baix Llobregat	736	521	668	248	88	35	[6,8]	**[1,3]**	2305
Baix Penedès	156	203	319	208	138	61	[3,7]	**[1,5]**	1093
Barcelonès	32	[14,21]	23	10	[1,8]	[1,8]	0	[1,8]	89
Berguedà	251	189	247	207	206	204	40	27	1371
Cerdanya	25	18	32	62	116	192	43	29	517
Conca de Barberà	301	263	479	372	354	339	74	22	2204
Garraf	145	123	174	107	60	34	[5,8]	[1,4]	652
Garrigues	236	298	745	852	975	540	52	14	3712
Garrotxa	288	231	423	339	214	111	25	12	1643
Gironès	492	239	358	348	265	142	12	7	1863
Maresme	884	662	633	152	49	[5,12]	[1,9]	[1,9]	2395
Montsià	1419	949	1404	932	489	189	20	5	5407
Noguera	342	256	608	661	807	858	217	66	3815
Osona	335	273	512	521	589	335	48	29	2642
Pallars Jussà	64	69	204	253	313	353	76	36	1368
Pallars Sobirà	74	53	183	198	127	110	19	52	816
Pla d'Urgell	299	267	624	690	493	260	27	11	2671
Pla de l'Estany	130	77	167	185	161	112	13	5	850
Priorat	396	390	670	488	263	88	[1,6]	[1,6]	2302
Ribera d'Ebre	399	507	1007	786	525	192	19	5	3440
Ripollès	113	70	124	153	137	179	97	80	953
Segarra	43	42	116	163	254	584	214	45	1461
Segrià	878	902	1859	1679	1500	806	130	45	7799
Selva	428	281	414	247	197	106	14	4	1691
Solsonès	26	40	81	79	119	265	107	34	751
Tarragonès	753	530	732	421	216	80	[16,24]	[1,9]	2757
Terra Alta	174	244	609	699	807	338	19	4	2894
Urgell	175	233	527	487	518	567	139	20	2666
Val d'Aran	39	45	151	78	[26,33]	[1,8]	[1,8]	19	367
Vallès Occidental	393	250	267	140	92	76	21	5	1244
Vallès Oriental	656	472	682	373	224	113	18	4	2542
Catalonia	17653	13857	22982	17380	13822	9504	1980	763	97941

Table 4: **Feasible intervals for all suppressed cells.**