The SAS System

The FREQ Procedure

Frequency	Table 1 of Embarked by Survived		
	Controlling for Pclass=1		
		Survived	
Embarked	0	1	Total
C	48	80	128
Q	1	2	3
S	65	86	151
Total	114	168	282
	Frequency Missing = 2		

Frequency	Table 2 of Embarked by Survived		
	Controlling for Pclass=2		
		Survived	
Embarked	0	1	Total
C	12	12	24
Q	4	1	5
S	133	99	232
Total	149	112	261

Frequency	Table 3 of Embarked by Survived		
	Controlling for Pclass=3		
		Survived	
Embarked	0	1	Total
C	36	24	60
Q	25	17	42
S	305	94	399
Total	366	135	501

The SAS System

The FREQ Procedure

Frequency	Table of FamilySize by Survived		
	Survived		
FamilySize	0	1	Total
1	409	181	590
2	96	110	206
3	63	81	144
4	9	30	39
5	13	4	17
6	20	5	25
7	11	5	16
8	7	1	8
11	1	0	1
Total	629	417	1046

The SAS System

The FREQ Procedure

Frequency	Table of IsAlone by Survived		
	Survived		
IsAlone	0	1	Total
0	220	236	456
1	409	181	590
Total	629	417	1046

The SAS System

The LOGISTIC Procedure

Model Information	
Data Set	WORK.CLEAN_ALL
Response Variable	Survived
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	1046
Number of Observations Used	1044

Response Profile		
Ordered Value	Survived	Total Frequency
1	1	415
2	0	629

Probability modeled is Survived='1'.

Note: 2 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information			
Class	Value	Design Variables	
Sex	female	1	
	male	0	
Embarked	C	1	0
	Q	0	1
	S	0	0
Pclass	1	1	0
	2	0	1
	3	0	0

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	1405.113	799.658
SC	1410.064	844.215
-2 Log L	1403.113	781.658

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	621.4553	8	<.0001
Score	541.4771	8	<.0001
Wald	337.8160	8	<.0001

Type 3 Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
Sex	1	303.0576	<.0001
Age	1	22.1326	<.0001
Pclass	2	58.0281	<.0001
Embarked	2	0.4578	0.7954
FamilySize	1	12.1349	0.0005
IsAlone	1	4.4145	0.0356

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-0.7465	0.4284	3.0368	0.0814
Sex	female	1	3.6515	0.2098	303.0576	<.0001
Age		1	-0.0350	0.00744	22.1326	<.0001
Pclass	1	1	2.0907	0.2752	57.7287	<.0001
Pclass	2	1	0.8985	0.2381	14.2357	0.0002
Embarked	C	1	0.1646	0.2442	0.4543	0.5003
Embarked	Q	1	-0.00394	0.4503	0.0001	0.9930
FamilySize		1	-0.3396	0.0975	12.1349	0.0005
IsAlone		1	-0.5806	0.2763	4.4145	0.0356

Odds Ratio Estimates		

Effect	Point Estimate	95% Wald Confidence Limits	
Sex female vs male	38.533	25.544	58.127
Age	0.966	0.952	0.980
Pclass 1 vs 3	8.090	4.718	13.874
Pclass 2 vs 3	2.456	1.540	3.916
Embarked C vs S	1.179	0.731	1.902
Embarked Q vs S	0.996	0.412	2.407
FamilySize	0.712	0.588	0.862
IsAlone	0.560	0.326	0.962

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	89.8	Somers' D	0.796
Percent Discordant	10.2	Gamma	0.797
Percent Tied	0.1	Tau-a	0.382
Pairs	261035	c	0.898

The SAS System

The HPFOREST Procedure

Performance Information	
Execution Mode	Single-Machine
Number of Threads	4

Data Access Information			
Data	Engine	Role	Path
WORK.CLEAN_ALL	V9	Input	On Client

Model Information		
Parameter	Value	
Variables to Try	2	(Default)
Maximum Trees	100	
Actual Trees	100	
Inbag Fraction	0.6	(Default)
Prune Fraction	0	(Default)
Prune Threshold	0.1	(Default)
Leaf Fraction	0.00001	(Default)
Leaf Size Setting	1	(Default)
Leaf Size Used	1	
Category Bins	30	(Default)
Interval Bins	100	
Minimum Category Size	5	(Default)
Node Size	100000	(Default)
Maximum Depth	20	(Default)
Alpha	1	(Default)
Exhaustive	5000	(Default)
Rows of Sequence to Skip	5	(Default)
Split Criterion	.	Variance
Preselection Method	.	Loh
Missing Value Handling	.	Valid value

Number of Observations	

Type	N
Number of Observations Read	1046
Number of Observations Used	1046

Baseline Fit Statistics	
Statistic	Value
Average Square Error	0.240

Fit Statistics			
Number of Trees	Number of Leaves	Average Square Error (Train)	Average Square Error (OOB)
1	12	0.13525	0.16342
2	35	0.11709	0.14350
3	57	0.11237	0.13848
4	78	0.11219	0.13709
5	97	0.11319	0.13485
6	127	0.11048	0.12972
7	156	0.10906	0.12751
8	178	0.10827	0.12542
9	203	0.10742	0.12398
10	227	0.10647	0.12162
11	241	0.10691	0.12067
12	267	0.10679	0.11950
13	286	0.10636	0.11911
14	306	0.10629	0.11862
15	322	0.10668	0.11901
16	347	0.10676	0.11924
17	363	0.10740	0.11894
18	384	0.10729	0.11836
19	402	0.10768	0.11847
20	429	0.10716	0.11803
21	448	0.10759	0.11875
22	473	0.10748	0.11865
23	490	0.10784	0.11862
24	509	0.10766	0.11826

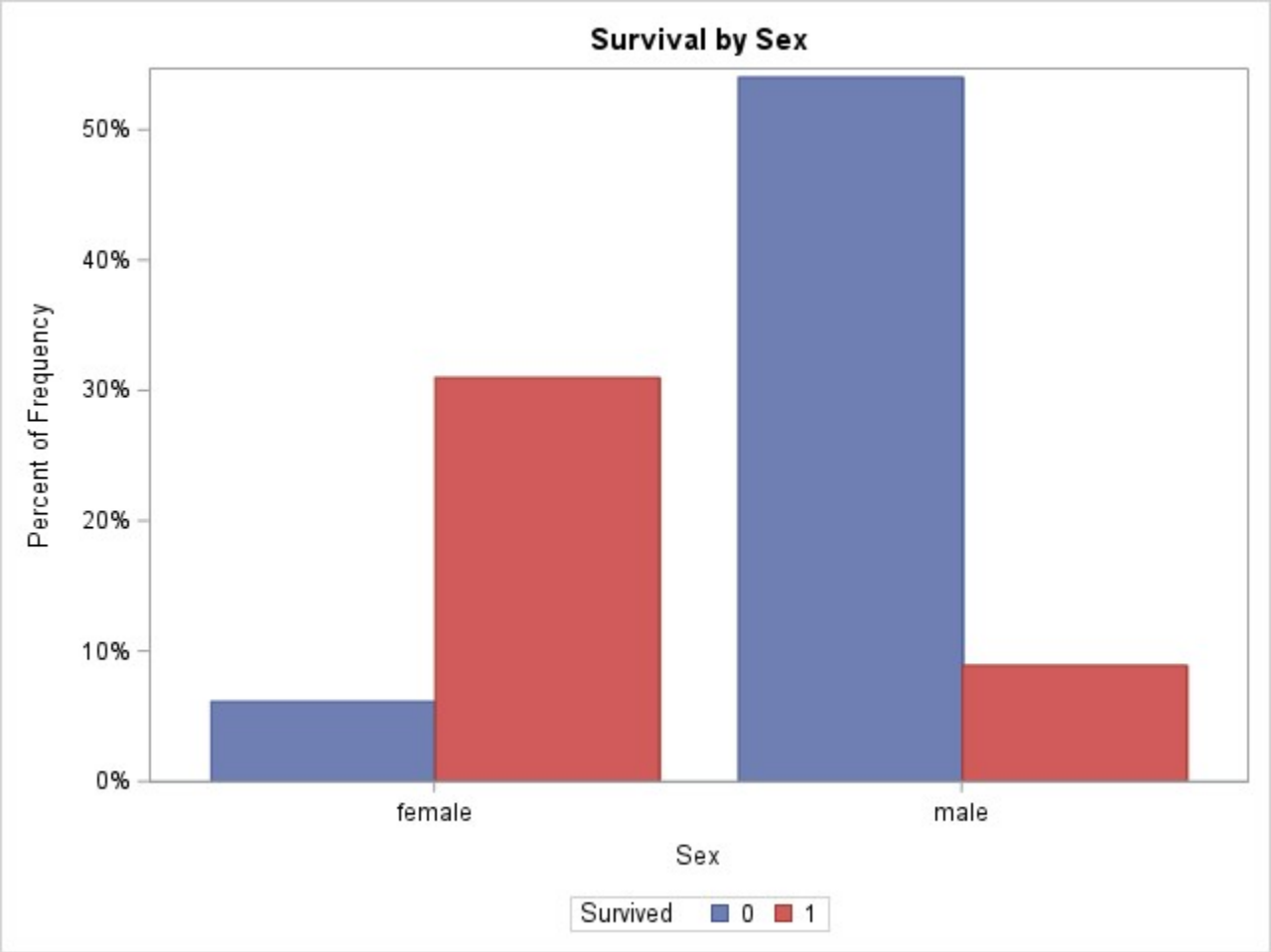
25	**538**	0.10767	0.11829
26	**564**	0.10748	0.11819
27	**581**	0.10750	0.11789
28	**599**	0.10722	0.11772
29	**631**	0.10703	0.11779
30	**657**	0.10701	0.11766
31	**684**	0.10714	0.11773
32	**709**	0.10699	0.11762
33	**725**	0.10742	0.11799
34	**747**	0.10743	0.11788
35	**774**	0.10746	0.11783
36	**793**	0.10736	0.11782
37	**807**	0.10725	0.11754
38	**825**	0.10727	0.11744
39	**849**	0.10726	0.11730
40	**878**	0.10715	0.11717
41	**901**	0.10719	0.11719
42	**935**	0.10696	0.11685
43	**961**	0.10692	0.11685
44	**978**	0.10691	0.11673
45	**999**	0.10697	0.11676
46	**1034**	0.10686	0.11679
47	**1056**	0.10697	0.11666
48	**1076**	0.10684	0.11657
49	**1100**	0.10687	0.11685
50	**1122**	0.10690	0.11689
51	**1140**	0.10704	0.11721
52	**1160**	0.10704	0.11703
53	**1171**	0.10726	0.11692
54	**1195**	0.10718	0.11692
55	**1215**	0.10728	0.11700
56	**1241**	0.10724	0.11691
57	**1266**	0.10722	0.11687
58	**1287**	0.10714	0.11669
59	**1312**	0.10714	0.11670

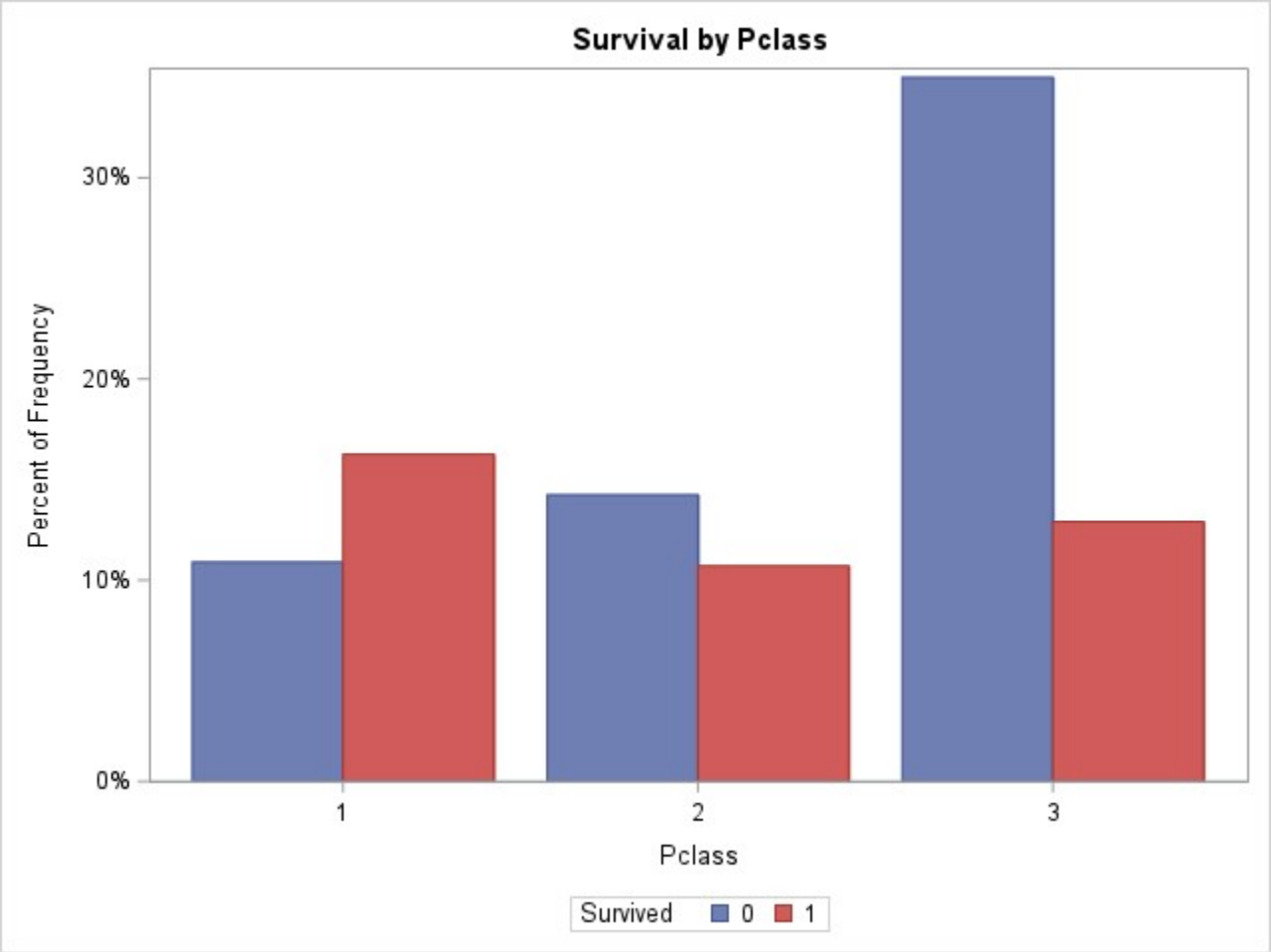
	1338	0.10697	0.11661
61	1361	0.10701	0.11658
62	1378	0.10706	0.11660
63	1394	0.10716	0.11666
64	1412	0.10716	0.11666
65	1437	0.10719	0.11669
66	1460	0.10704	0.11653
67	1483	0.10689	0.11628
68	1505	0.10685	0.11619
69	1533	0.10672	0.11601
70	1553	0.10668	0.11581
71	1569	0.10670	0.11579
72	1597	0.10661	0.11572
73	1618	0.10659	0.11573
74	1637	0.10650	0.11566
75	1665	0.10659	0.11573
76	1690	0.10666	0.11585
77	1716	0.10668	0.11580
78	1740	0.10666	0.11583
79	1757	0.10671	0.11593
80	1778	0.10665	0.11595
81	1795	0.10668	0.11611
82	1812	0.10670	0.11613
83	1833	0.10668	0.11599
84	1855	0.10662	0.11603
85	1874	0.10662	0.11595
86	1898	0.10660	0.11593
87	1926	0.10665	0.11595
88	1941	0.10661	0.11586
89	1970	0.10658	0.11586
90	1992	0.10658	0.11581
91	2017	0.10652	0.11578
92	2042	0.10650	0.11575
93	2064	0.10659	0.11585
94	2092	0.10655	0.11595

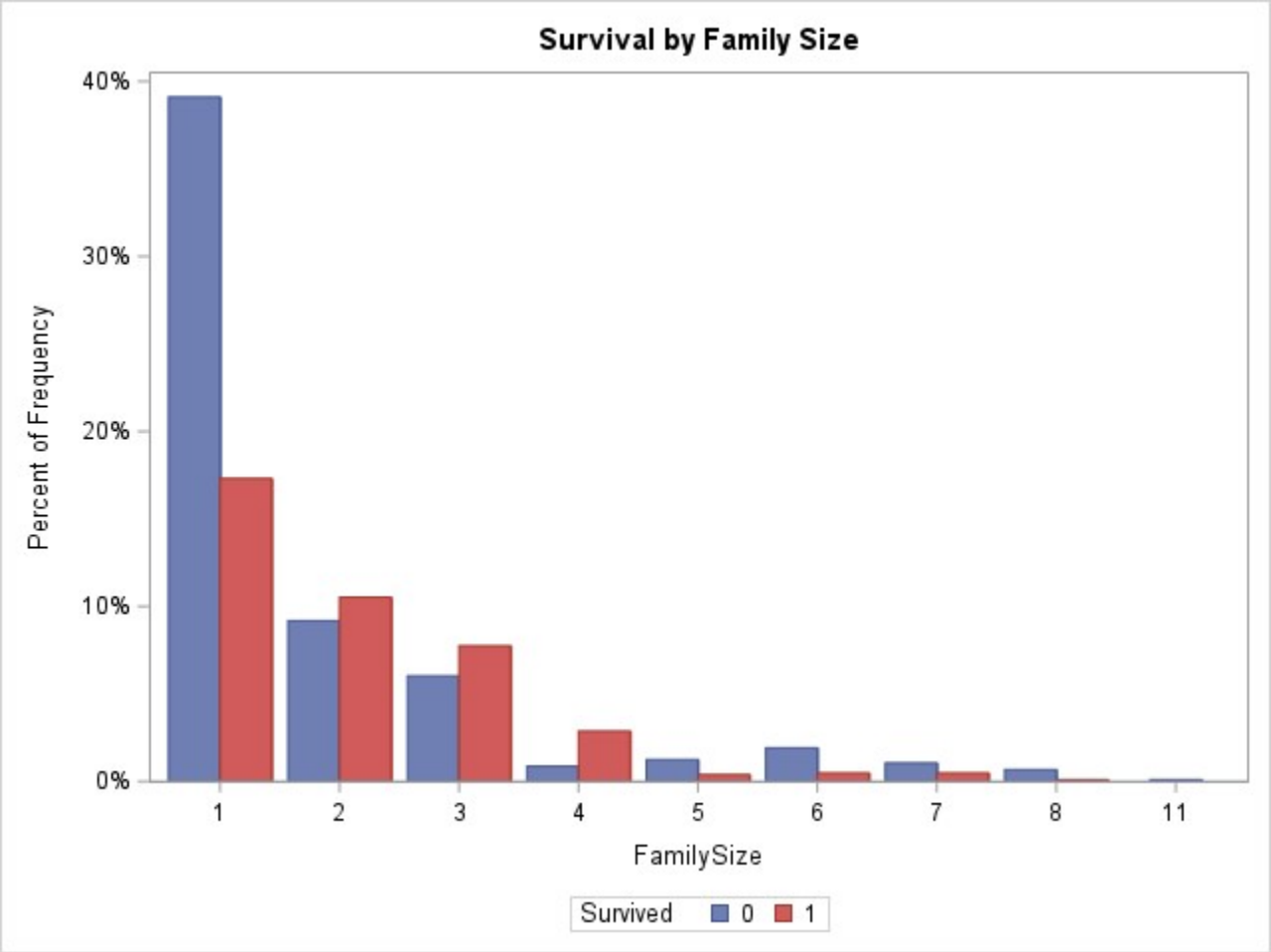
	2116	0.10661	0.11605
96	2146	0.10650	0.11594
97	2164	0.10646	0.11592
98	2186	0.10641	0.11583
99	2209	0.10629	0.11567
100	2229	0.10628	0.11569

Loss Reduction Variable Importance

Variable	Number of Rules	MSE	OOB MSE	Absolute Error	OOB Absolute Error
Sex	270	0.093728	0.09385	0.186284	0.186222
Pclass	498	0.012125	0.01001	0.020757	0.018023
IsAlone	255	0.004482	0.00357	0.006503	0.005570
FamilySize	414	0.006714	0.00309	0.013138	0.009702
Embarked	388	0.002946	0.00052	0.005041	0.002466
Age	304	0.001114	-0.00665	0.013540	0.006311


Survival by Sex
50%
40%
30%
20%
10%
0%
Percent of Frequency
female
male
Sex
Survived 0 1


Survival by Pclass
30%
20%
10%
0%
Percent of Frequency
1
2
3
Pclass
Survived 0 1


Survival by Family Size
40%
30%
20%
10%
0%
Percent of Frequency
1
2
3
4
5
6
7
8
11
FamilySize
Survived
0
1

Survival by Family Size

The FREQ Procedure

Frequency	Table of FamilySize by Pclass			
	Pclass			
FamilySize	1	2	3	Total
1	128	142	320	590
2	97	52	57	206
3	39	45	60	144
4	9	20	10	39
5	5	1	11	17
6	6	1	18	25
7	0	0	16	16
8	0	0	8	8
11	0	0	1	1
Total	284	261	501	1046