Department of Statistics
COURSE STATS 330
Assignment 5, 2004

Instructions: Hand in your completed assignment to the Student Resource Centre by 4pm on Thursday 21st October.

In a survey of American college students, the following variables were measured:

alcohol:	Has the student tried alcoholic drinks? (Yes/No)
cigarettes:	Has the student tried smoking (tobacco) cigarettes? (Yes/No)
marijuana:	Has the student tried marijuana (cannabis)? (Yes/No)
race:	Student's race (White/Other)
sex:	Student's gender (F/M)

The data are shown overleaf. Note that unlike previous assignments, you must type these data in yourself, it is not available on the Web page. See Tutorial 9 for ways of doing this.

1. Create a suitable data frame for the data. Check for errors. Note that the data overleaf are correct, so any errors will be due to your typing or programming.

2. Collapse the table over the factors Race and Sex. Using the variables alcohol, cigarettes, and marijuana only, find a suitable conditional independence model for these data. Use your fitted model to decide if trying alcohol and trying cigarettes are related. If they are, do you think the degree of the relationship depends on trying marijuana?

3. Now fit a saturated model using all the variables. Are there any relationships between trying alcohol, cigarettes, and marijuana on the one hand, and gender and race on the other? Fit suitable additional models to explore these issues.

Useful R functions

The function **`expand.grid`** is useful for creating factors, so you need only type in the counts. To collapse an existing table, turn the table into a data frame, then you can use the functions **`tapply, as.vector`** and **`data.frame`** to create a data frame containing the collapsed table.

See Tutorial 9 for examples.

Data for Assignment 5.

						Race			
						White		Other	
						Alcohol		Alcohol	
						Yes	No	Yes	No
Sex	Male	Marijuana	Yes	Cigarettes	Yes	405	1	23	0
					No	13	1	2	0
			No	Cigarettes	Yes	268	17	23	1
					No	218	117	19	12
	Female	Marijuana	Yes	Cigarettes	Yes	453	1	30	1
					No	28	1	1	0
			No	Cigarettes	Yes	228	17	19	8
					No	201	133	12	17

Data for Assignment 4

crab	colour	spine	width	weight	satelite	crab	colour	spine	width	weight	satelite
1	2	3	28.3	3050	1	88	4	1	25.5	2750	0
2	3	3	22.5	1550	0	89	4	3	23.5	1900	0
3	1	1	26.0	2300	1	90	2	2	24.0	1700	0
4	3	3	24.8	2100	0	91	2	1	29.7	3850	1
5	3	3	26.0	2600	1	92	2	1	26.8	2550	0
6	2	3	23.8	2100	0	93	4	3	26.7	2450	0
7	1	1	26.5	2350	0	94	2	1	28.7	3200	0
8	3	2	24.7	1900	0	95	3	3	23.1	1550	0
9	2	1	23.7	1950	0	96	2	1	29.0	2800	1
10	3	3	25.6	2150	0	97	3	3	25.5	2250	0
11	3	3	24.3	2150	0	98	3	3	26.5	1967	1
12	2	3	25.8	2650	0	99	3	3	24.5	2200	1
13	2	3	28.2	3050	1	100	3	3	28.5	3000	1
14	4	2	21.0	1850	0	101	2	3	28.2	2867	1
15	2	1	26.0	2300	1	102	2	3	24.5	1600	1
16	1	1	27.1	2950	1	103	2	3	27.5	2550	1
17	2	3	25.2	2000	1	104	2	2	24.7	2550	1
18	2	3	29.0	3000	1	105	2	1	25.2	2000	1
19	4	3	24.7	2200	0	106	3	3	27.3	2900	1
20	2	3	27.4	2700	1	107	2	3	26.3	2400	1
21	2	2	23.2	1950	1	108	2	3	29.0	3100	1
22	1	2	25.0	2300	1	109	2	3	25.3	1900	1
23	2	1	22.5	1600	1	110	2	3	26.5	2300	1
24	3	3	26.7	2600	1	111	2	3	27.8	3250	1
25	4	3	25.8	2000	1	112	2	3	27.0	2500	1
26	4	3	26.2	1300	0	113	3	3	25.7	2100	0
27	2	3	28.7	3150	1	114	2	3	25.0	2100	1
28	2	1	26.8	2700	1	115	2	3	31.9	3325	1
29	4	3	27.5	2600	0	116	4	3	23.7	1800	0
30	2	3	24.9	2100	0	117	4	3	29.3	3225	1
31	1	1	29.3	3200	1	118	3	3	22.0	1400	0
32	1	3	25.8	2600	0	119	2	3	25.0	2400	1
33	2	2	25.7	2000	0	120	3	3	27.0	2500	1
34	2	1	25.7	2000	1	121	3	3	23.8	1800	1
35	2	1	26.7	2700	1	122	1	1	30.2	3275	1
36	4	3	23.7	1850	0	123	3	3	26.2	2225	0
37	2	3	26.8	2650	0	124	2	3	24.2	1650	1
38	2	3	27.5	3150	1	125	2	3	27.4	2900	1
39	4	3	23.4	1900	0	126	2	2	25.4	2300	0
40	2	3	27.9	2800	1	127	3	3	28.4	3200	1
41	3	3	27.5	3100	1	128	4	3	22.5	1475	1
42	1	1	26.1	2800	1	129	2	3	26.2	2025	1
43	1	1	27.7	2500	1	130	2	1	24.9	2300	1
44	2	1	30.0	3300	1	131	1	2	24.5	1950	1
45	3	1	28.5	3250	1	132	2	3	25.1	1800	0
46	3	3	28.9	2800	1	133	2	1	28.0	2900	1
47	2	3	28.2	2600	1	134	4	3	25.8	2250	1

48	2	3	25.0	2100	1	135	2	3	27.9	3050	1
49	2	3	28.5	3000	1	136	2	3	24.9	2200	0
50	2	1	30.3	3600	1	137	2	1	28.4	3100	1
51	4	3	24.7	2100	1	138	3	3	27.2	2400	1
52	2	3	27.7	2900	1	139	2	2	25.0	2250	1
53	1	1	27.4	2700	1	140	2	3	27.5	2625	1
54	2	3	22.9	1600	1	141	2	1	33.5	5200	1
55	2	1	25.7	2000	1	142	2	3	30.5	3325	1
56	2	3	28.3	3000	1	143	3	3	29.0	2925	1
57	2	3	27.2	2700	1	144	2	1	24.3	2000	0
58	3	3	26.2	2300	1	145	2	3	25.8	2400	0
59	2	1	27.8	2750	0	146	4	3	25.0	2100	1
60	4	3	25.5	2250	0	147	2	1	31.7	3725	1
61	3	3	27.1	2550	0	148	2	3	29.5	3025	1
62	3	3	24.5	2050	1	149	3	3	24.0	1900	1
63	3	1	27.0	2450	1	150	2	3	30.0	3000	1
64	2	3	26.0	2150	1	151	2	3	27.6	2850	1
65	2	3	28.0	2800	1	152	2	3	26.2	2300	0
66	2	3	30.0	3050	1	153	2	1	23.1	2000	0
67	2	3	29.0	3200	1	154	2	1	22.9	1600	0
68	2	3	26.2	2400	0	155	4	3	24.5	1900	0
69	2	1	26.5	1300	0	156	2	3	24.7	1950	1
70	2	3	26.2	2400	1	157	2	3	28.3	3200	0
71	3	3	25.6	2800	1	158	2	3	23.9	1850	1
72	3	3	23.0	1650	1	159	3	3	23.8	1800	0
73	3	3	23.0	1800	0	160	3	2	29.8	3500	1
74	2	3	25.4	2250	1	161	2	3	26.5	2350	1
75	3	3	24.2	1900	0	162	2	3	26.0	2275	1
76	2	2	22.9	1600	0	163	2	3	28.2	3050	1
77	3	2	26.0	2200	1	164	4	3	25.7	2150	0
78	2	3	25.4	2250	1	165	2	3	26.5	2750	1
79	3	3	25.7	1200	0	166	2	3	25.8	2200	0
80	2	3	25.1	2100	1	167	3	3	24.1	1800	0
81	3	2	24.5	2250	0	168	3	3	26.2	2175	1
82	4	3	27.5	2900	0	169	3	3	26.1	2750	1
83	3	3	23.1	1650	0	170	3	3	29.0	3275	1
84	3	1	25.9	2550	1	171	1	1	28.0	2625	0
85	2	3	25.8	2300	0	172	4	3	27.0	2625	0
86	4	3	27.0	2250	1	173	2	2	24.5	2000	0
87	2	3	28.5	3050	0						