

MIT OpenCourseWare
<http://ocw.mit.edu>

11.220 Quantitative Reasoning & Statistical Methods for Planners I Spring 2009

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.

Quantitative Reasoning and Statistical Methods

Ezra Glenn

April 2, 2009

	"Planning Conservatism"			Total
	Low	Medium	High	
Sample (100 towns)	26	38	36	100

Table: "Planning Conservatism" in Sample Towns

	"Planning Conservatism"			Total
	Low	Medium	High	
Sample (100 towns)	26	38	36	100

Table: "Planning Conservatism" in Sample Towns

Best guess of level of "Planning Conservatism" would be the mode ("Medium"). We'd be right 38% of the time (and wrong 62%).

	"Planning Conservatism"			
	Low	Medium	High	Total
Sample (100 towns)	26	38	36	100

Table: "Planning Conservatism" in Sample Towns

	"Planning Conservatism"			Total
	Low	Medium	High	
Sample (100 towns)	26	38	36	100

Table: "Planning Conservatism" in Sample Towns

	"Planning Conservatism"			Total
	Low	Medium	High	
Small Towns	6	15	20	41
Mid-size Towns	10	15	11	36
Large Towns	10	8	5	23
Total	26	38	36	100

Table: "Planning Conservatism" in Sample, by Town Size

	"Planning Conservatism"			Total
	Low	Medium	High	
Small Towns	6	15	20	41
Mid-size Towns	10	15	11	36
Large Towns	10	8	5	23
Total	26	38	36	100

Table: "Planning Conservatism" in Sample, by Town Size

Predictions and Errors

- For Small Towns: predict "High" (still wrong $\frac{21}{41} = .51$)
- For Mid-size Towns: predict "Medium" (still wrong $\frac{21}{36} = .58$)
- For Large Towns: predict "Low" (still wrong $\frac{13}{23} = .56$)

Predictions and Errors

- For Small Towns: predict “High” (still wrong $\frac{21}{41} = .51$)
- For Mid-size Towns: predict “Medium” (still wrong $\frac{21}{36} = .58$)
- For Large Towns: predict “Low” (still wrong $\frac{13}{23} = .56$)

$$\begin{cases} .51 \text{ error} \times .41 \text{ of cases} = .21 \\ .58 \text{ error} \times .36 \text{ of cases} = .21 \\ .56 \text{ error} \times .23 \text{ of cases} = .13 \end{cases} \quad (1)$$

An improvement

$$.21 + .21 + .13 = .55.$$

(Compare this to the previous error of .62)

The “Proportional Reduction in Error” is also called λ :

$$\lambda = \frac{(\text{Error w/o conditional info}) - (\text{Error w/conditional info})}{\text{Error w/o conditional info}} \quad (2)$$

Here:

$$\lambda = \frac{.62 - .55}{.62} = .113 \quad (3)$$

	"Planning Conservatism"			
	Low	Medium	High	Total
Small Towns	6	15	20	41
Mid-size Towns	10	15	11	36
Large Towns	10	8	5	23
Total	26	38	36	100

Table: χ^2 test (Observed)

	"Planning Conservatism"			Total
	Low	Medium	High	
Small Towns	6	15	20	41
Mid-size Towns	10	15	11	36
Large Towns	10	8	5	23
Total	26	38	36	100

Table: χ^2 test (Observed)

	"Planning Conservatism"			Total
	Low	Medium	High	
Small Towns	10.66	15.58	14.76	41
Mid-size Towns	9.36	13.68	12.96	36
Large Towns	5.98	8.74	8.28	23
Total	26	38	36	100

Table: χ^2 test (Expected)

	"Planning Conservatism"		
	Low	Medium	High
Small Towns	2.04	0.02	1.86
Mid-size Towns	0.04	0.13	0.30
Large Towns	2.70	0.06	1.30

Table: χ^2 test (cell contributions)

$$\chi^2 = 8.45, df = 4, p \text{ value} = 0.076 > .05$$

	"Planning Conservatism"		
	Low	Medium	High
Small Towns	2.04	0.02	1.86
Mid-size Towns	0.04	0.13	0.30
Large Towns	2.70	0.06	1.30

Table: χ^2 test (cell contributions)

Not Significant!

$\chi^2 = 8.45$, $df = 4$, $p \text{ value} = 0.076 > .05$

	"Planning Conservatism"			
	Low	Medium	High	Total
Small Towns	6	15	20	41
Mid-size Towns	10	15	11	36
Large Towns	10	8	5	23
Total	26	38	36	100

Table: Old sample

	"Planning Conservatism"			Total
	Low	Medium	High	
Small Towns	6	15	20	41
Mid-size Towns	10	15	11	36
Large Towns	10	8	5	23
Total	26	38	36	100

Table: Old sample

	"Planning Conservatism"			Total
	Low	Medium	High	
Small Towns	12	30	40	82
Mid-size Towns	20	30	22	72
Large Towns	20	16	10	46
Total	52	76	72	200

Table: New Sample (surprisingly similar proportions...)

	"Planning Conservatism"			Total
	Low	Medium	High	
Small Towns	12	30	40	82
Mid-size Towns	20	30	22	72
Large Towns	20	16	10	46
Total	52	76	72	200

Table: New Sample

Predictions and Errors (this should look similar...)

- For Small Towns: predict "High" (still wrong $\frac{42}{82} = .51$)
- For Mid-size Towns: predict "Medium" (still wrong $\frac{42}{72} = .58$)
- For Large Towns: predict "Low" (still wrong $\frac{26}{46} = .56$)

Predictions are no "better" (errors, λ , etc., all still the same)

	"Planning Conservatism"			
	Low	Medium	High	Total
Small Towns	12	30	40	82
Mid-size Towns	20	30	22	72
Large Towns	20	16	10	46
Total	52	76	72	200

Table: Observed

	"Planning Conservatism"			Total
	Low	Medium	High	
Small Towns	12	30	40	82
Mid-size Towns	20	30	22	72
Large Towns	20	16	10	46
Total	52	76	72	200

Table: Observed

	"Planning Conservatism"			Total
	Low	Medium	High	
Small Towns	21.32	31.16	29.52	82
Mid-size Towns	18.72	27.36	25.92	72
Large Towns	11.96	17.48	16.56	46
Total	52	76	72	200

Table: Expected

	"Planning Conservatism"		
	Low	Medium	High
Small Towns	4.07	0.04	3.72
Mid-size Towns	0.09	0.25	0.59
Large Towns	5.40	0.13	2.60

Table: χ^2 test (cell contributions)

$$\chi^2 = 16.9, df = 4, p \text{ value} = 0.002$$

	"Planning Conservatism"		
	Low	Medium	High
Small Towns	4.07	0.04	3.72
Mid-size Towns	0.09	0.25	0.59
Large Towns	5.40	0.13	2.60

Table: χ^2 test (cell contributions)

Significant!

$\chi^2 = 16.9$, $df = 4$, $p \text{ value} = 0.002$

	"Planning Conservatism"			Total
	Low	Medium	High	
Small Towns	12	30	40	82
Mid-size Towns	20	30	22	72
Large Towns	20	16	10	46
Total	52	76	72	200

	"Planning Conservatism"			Total
	Low	Medium	High	
Low Income Towns	36	30	24	90
High Income Towns	16	46	48	110
Total	52	76	72	200

Table: "Planning Conservatism" of Towns by Income

	"Planning Conservatism"			Total
	Low	Medium	High	
Low Income Towns	36	30	24	90
High Income Towns	16	46	48	110
Total	52	76	72	200

Table: "Planning Conservatism" of Towns by Income

$$\chi^2 = 17.23, df = 2, p \text{ value} = 0.0001 \text{ significant!}$$

$$\lambda = \frac{.62 - .58}{.62} = .065 \leftarrow \text{Not as good as town size}$$

	"Planning Conservatism"			Total
	Low	Medium	High	
Small Towns	2	3	4	9
Mid-size Towns	16	16	12	44
Large Towns	18	11	8	37
Total	36	30	24	90

Table: Low Income Towns

	"Planning Conservatism"			Total
	Low	Medium	High	
Small Towns	10	27	36	73
Mid-size Towns	4	14	10	28
Large Towns	2	5	2	9
Total	16	46	48	110

Table: High Income Towns

	"Planning Conservatism"			Total
	Low	Medium	High	
Small Towns	2	3	4	9
Mid-size Towns	16	16	12	44
Large Towns	18	11	8	37
Total	36	30	24	90

Table: Low Income Towns

$\chi^2 = 3.24$, $df = 4$, p value = 0.52 Not significant

	"Planning Conservatism"			Total
	Low	Medium	High	
Small Towns	10	27	36	73
Mid-size Towns	4	14	10	28
Large Towns	2	5	2	9
Total	16	46	48	110

Table: High Income Towns

$\chi^2 = 3.55$, $df = 4$, p value = 0.47 Not significant

	Town Size			Total
	Small	Medium	Large	
Low Income Towns	9	44	37	90
High Income Towns	73	28	9	110
Total	82	72	46	200

Table: Town Size by Income

$$\chi^2 = 69.2427, df = 2, p \text{ value} < .001$$

Multicollinearity often occurs when dealing with income, education, class, race/ethnicity, geography—many of the things we care about. . . .

- think about making predictions and “the smooth and the rough”
- think about missing/lurking/confounding variables
- draw causal diagrams
- beware the “ecological fallacy”