

Calculating confidence intervals for crosstabulations

Note that these confidence intervals are based on a simple random survey (ie excluding design effects from stratification and surveyor variability) and therefore *underestimate* the size of the interval.

Example

1. SPSS syntax and output to produce required tables

```
GET
  FILE='D:\DATA\EHCS2001\Database\Provisional\coretgo3.sav'.
MATCH FILES /FILE=*
  /TABLE='D:\DATA\EHCS2001\Database\Derived\physical summary.sav'
  /BY aacod01.
EXECUTE.
MATCH FILES /FILE=*
  /TABLE='D:\DATA\EHCS2001\Database\Derived\tenure.sav'
  /BY aacod01.
EXECUTE.
MATCH FILES /FILE=*
  /TABLE='D:\DATA\EHCS2001\Database\Derived\unfit01x.sav'
  /BY aacod01.
EXECUTE.
```

*** The first table calculates 'p' (the probability of unfitness occurring in specified populations - in this example tenure groups and the whole population of dwellings) expressed as a percentage .

```
WEIGHT
  BY grdwtotg .
CROSSTABS
  /TABLES=ten401x BY unfit01x
  /FORMAT= AVALUE TABLES
  /CELLS= ROW .
```

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Tenure 2001 - 4 categories * Fitness - final 2001	21355925	100.0%	0	.0%	21355925	100.0%

Tenure 2001 - 4 categories * Fitness - final 2001 Crosstabulation

% within Tenure 2001 - 4 categories

		Fitness - final 2001		
		unfit	fit	Total
Tenure 2001 - 4 categories	owner occupied	3.2%	96.8%	100.0%
	private rented	11.0%	89.0%	100.0%
	LA	4.7%	95.3%	100.0%
	RSL	3.4%	96.6%	100.0%
Total		4.2%	95.8%	100.0%

*** The second table indicates 'n' (the number of raw cases in the sample comprising the specified populations) and which are given by the marginal totals of the crosstabulation .

```
WEIGHT
OFF.
CROSSTABS
/TABLES=ten401x BY unfit01x
/FORMAT= AVALUE TABLES
/CELLS= COUNT .
```

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Tenure 2001 - 4 categories * Fitness - final 2001	17532	100.0%	0	.0%	17532	100.0%

Tenure 2001 - 4 categories * Fitness - final 2001 Crosstabulation

Count		Fitness - final 2001		
		unfit	fit	Total
Tenure 2001 - 4 categories	owner occupied	295	8413	8708
	private rented	187	1358	1545
	LA	223	4324	4547
	RSL	85	2647	2732
Total		790	16742	17532

2) Use of standard tables to calculate approximate confidence intervals for each tenure and for all dwellings, and a comparison with the actual confidence intervals produced by applying standard formula.

The standard formula to derive the 95% confidence interval (expressed as a percentage) is given by:

$$CI = (1.96 \times \text{SQRT}((p \times (1 - p/n)) \times 100.$$

Unfitness rate by tenure, 2001

tenure 2001:	unfitness p(%)	marginal count n(raw cases)	approx CI from tables +/-	calculated CI (simple random sample) +/-
owner occupied	3.2	8,708	0.36	0.37
private rented	11.0	1,545	1.58	1.56
LA	4.7	4,547	0.64	0.62
RSL	3.4	2,732	0.67	0.68
all tenures	4.2	17,532	0.29	0.30