

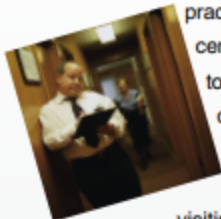
## Curbstoning

The glossary for the Census defines *curbstoning* as “the

practice by which a census enumerator fabricates a questionnaire for a residence without actually visiting it.” Curbston-

ing occurs when a census enumerator sits on a curbstone (or anywhere else) and fills out survey forms by making up responses. Because data from curbstoning are not real, they can affect the validity of the Census. The extent of curbstoning has been investigated in several studies, and one study showed that about 4% of Census enumerators practiced curbstoning at least some of the time.

The methods of Section 7-2 assume that the sample data have been collected in an appropriate way, so if much of the sample data have been obtained through curbstoning, then the resulting confidence interval estimates might be very flawed.



A confidence interval is associated with a specific confidence level, such as 0.95 (or 95%). The confidence level gives us the *success rate of the procedure* used to construct the confidence interval. See the following and notice the relationship between the confidence level and the corresponding value of  $\alpha$ .

Most Common Confidence Levels	Corresponding Values of $\alpha$
90% (or 0.90) confidence level:	$\alpha = 0.10$
95% (or 0.95) confidence level:	$\alpha = 0.05$
99% (or 0.99) confidence level:	$\alpha = 0.01$

The confidence level of 95% is most common because it provides a good balance between precision (as reflected in the width of the confidence interval) and reliability (as expressed by the confidence level).

Here's an example of a confidence interval found later (in Example 3), which is based on the sample data of 1007 adults polled, with 85% of them knowing what Twitter is:

**The 0.95 (or 95%) confidence interval estimate of the population proportion  $p$  is  $0.828 < p < 0.872$ .**

## Interpreting a Confidence Interval

We must be careful to interpret confidence intervals correctly. There is a correct interpretation and many different and creative incorrect interpretations of the confidence interval  $0.828 < p < 0.872$ .

**Correct:** “We are 95% confident that the interval from 0.828 to 0.872 actually does contain the true value of the population proportion  $p$ .” This means that if we were to select many different samples of size 1007 and construct the corresponding confidence intervals, 95% of them would actually contain the value of the population proportion  $p$ . (In this correct interpretation, the confidence level of 95% refers to the *success rate of the process* used to estimate the population proportion.)

**Incorrect:** “There is a 95% chance that the true value of  $p$  will fall between 0.828 and 0.872.”

**Incorrect:** “95% of sample proportions will fall between 0.828 and 0.872.”

**CAUTION** Know the correct interpretation of a confidence interval, as given above.

At any specific time, a population has a fixed and constant value of the proportion  $p$ , and a confidence interval constructed from a sample either includes  $p$  or does not. Similarly, if a baby has just been born and the doctor is about to announce its gender, it's incorrect to say that there is a probability of 0.5 that the baby is a girl; the