Stat 140: Cat Ownership

Evan Ray (adapted from Rossman and Chance)

December 4, 2017

Cat Ownership

In a sample survey of 50,347 households in December 2011, approximately 15305 of the households, or about 30.4%, owned a pet cat. (2012 U.S. Pet Ownership & Demographics Sourcebook, American Veterinary Medical Association, see also http://news.vin.com/vinnews.aspx?articleId=31369). Although there are problems with the way this survey was conducted, let's assume for the purpose of this example that this was a representative sample of independent households in the US.

(a) Is the number 0.304 a population parameter or a sample statistic? What symbol would you use for the population parameter and what symbol would you use for the sample statistic?

(b) The proportion of US households who own cats is something I think about a lot, and before I read about this survey I always believed that 31% of US households owned cats. Conduct a hypothesis test of whether the sample data provide evidence that the population proportion who own a pet cat differs from 31%. State the hypotheses, report the p-value, and draw a conclusion in the context of this study using an $\alpha = 0.01$ significance level. You may use the R output below:

```
binom.test(15305, 50347, p = .31, ci.method = "wald")
```

```
##
##  Exact binomial test (with Wald CI)
##
## data:  15305 out of 50347
## number of successes = 15000, number of trials = 50000, p-value =
## 0.004
## alternative hypothesis: true probability of success is not equal to 0.31
## 95 percent confidence interval:
##  0.300 0.308
## sample estimates:
## probability of success
##                  0.304
```

(c) Interpret the 95% confidence interval for the proportion of households in the US who own cats in context. Could you recreate the confidence interval in the R output "by hand"?