STAT 22000: Homework 9

Zixi Li

Problem 1 Textbook review

A is (1)Histogram of the ratings given by all potential readers. Because the population distribution is exactly based on all potential readers, and the ratings are: 50% of the readers give 5 stars, 25% give 4 stars, 25% of the readers give 2 stars.

B is (4) Histogram of 5000 sample means of the ratings from random samples of each size 100. Because the histogram looks very normal and the mean is 4 stars.

C is (2) Histogram of the ratings given by 100 randomly sampled readers. Because there are only 3 kinds of ratings and the distribution is not completely same with the population distribution, apparently it's a sample distribution.

D is (3) Histogram of 5000 sample means of the ratings from random samples of each size 10. Because the histogram has a mean equals to 4, but the distribution is not quite normal as B.

Problem 2 Ideal number of children

(a)

The point estimate of the population mean is simply the sample mean 3.22.

(b)

$SE = SD(samplemean) = SD(sample)/\sqrt{n} = 1.99/\sqrt{678} = 0.0764$

(c)

The 99% confidence interval can be calculated as

$mean \pm 2.58SE = 3.22 \pm 2.58 \times 0.0764 = (3.02, 3.42)$

(d)

It's not plausible, because for 99% of the time, the population mean will be within (3.02, 3.42). But it's still possible.

Problem 3 True or faulse of problem 2

(a) is true, because the mean is larger than median, the distribution is right-skewed.

(b) is false, because provided the sample size is large, the distribution of sample mean will be normal, even if the population distribution or sample distribution is not normal.

(c) is false, because the average of the ideal number of children in a family answered by the 678 females is 3.22, we are 100% confident.

(d) is false, because the confident interval is for covering the population mean, not for covering 99% of the entire population.

(e) is false, because the new sample size is undecided, the new sample mean can be very random can cover a large interval.

(f) is true, because for the confidence interval $samplemean \pm z * SE$, the margin of error is $z * SE$, which is 0.2 here.

(g) is false, because in order to decrease the margin of error by half, we need to decrease the standard error of sample mean by half, then we need to increase the sample size by 4 times as large.

(h) is true, we don't need to be as sure about our estimate, the z value for 95% CI is 1.96, for 99% CI is 2.58 instead.