Homework due Oct. 20

Assigned exercises: Ch.5, OpenIntro Stats, ex. 5.16, 5.17, 5.18, 5.19, 5.23.
Supplemental exercises linked via homework web page:
Pg.469: 11, 15, 16, 22. Pg.498: 21, 24, 33. (12 exercises total)
Graded exercises: 5.16, 5.23; supp-pg.469: 16, 22(a-c); supp-pg.498: 24(a-c).
Total (maximum) possible points = 20.
3 pt for each of 5 graded problems, plus 5 for completion of the rest.

Exercises from Ch.5, OpenIntro Stats

(5.16) (a) Let μ = true mean calorie intake of diners at that restaurant. The hypotheses are

$H_0 : \mu = 1100$
$H_A : \mu \neq 1100$

H_A is 2-tailed, since we want to test whether there is a difference in calorie intake.

(b) Let p = true proportion of Wisconsin adults who consumed alcohol last year.

$H_0 : p = 0.7$ (i.e., same as the national rate)
$H_A : p \neq 0.7$

Again, we have a 2-tailed H_A, since we want to test whether there is a difference.

(5.23) Given information:

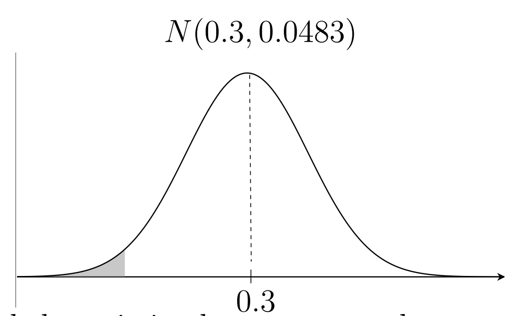
$H_0 : p = 0.3$, $H_A : p \neq 0.3$, $n = 90$, P-value=0.05.

We want to find $\hat{p}$, assuming all the conditions for inference are met.

Since $p = 0.3$, the sampling distribution model is $N(0.3, \sqrt{\frac{0.3(1-0.3)}{90}}) = N(0.3, 0.0483)$.

Since the P-value=0.05, and H_A is 2-tailed, each tail contains 0.025 of the probability.

Using z-tables, or software, and doing an inverse lookup for 0.025, we get $z = \pm 1.96$.



Thus, there are two possible values for the sampled statistic that correspond to a P-value of 0.05: $\hat{p} = 0.3 \pm 1.96 \times 0.0483$ $\boxed{= 0.3947 \text{ or } 0.2053.}$

Exercises from supplement linked via homework web page

(pg.469:16) Assuming the conditions for inference are met, a P-value of 0.017 is quite small, since typical significance levels for this type of study tend to be around 0.05 or more. It means there is only a 1.7% chance of observing the sampled statistic due to sampling variability, if the true proportion of high schoolers who have cars is the same as it was 10 years ago. Thus, it is reasonable to conclude that more high schoolers have cars.

(pg.470:22) (a) Let p = current proportion of nation's children who have congenital abnormalities.

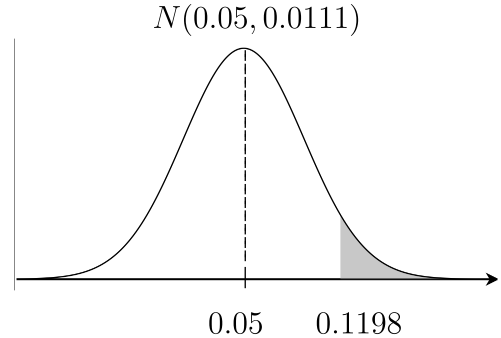
$H_0 : p = 0.05$ (i.e., same as in the 1980s)
$H_A : p > 0.05$

H_A is 1-tailed, since we want to know whether the rate has increased.

(b) Check for independence: There is no clear indication about how the 384 children in the study were selected. Unless the sample was random or, at the very least, representative of the nation's children, the independence condition may not be met.

Check for sample size: $n \cdot p = 384 \cdot 0.05 = 19.2 > 10$; and $n \cdot (1-p) = 364.8 > 10$. Thus the sample is large enough.

(c) Since $p = 0.05$, the sampling distribution model is $N(0.05, \sqrt{\frac{0.05(1-0.05)}{384}})$ $= N(0.05, 0.0111)$.



The sampled statistic is $\hat{p} = \frac{46}{384} = 0.1198$.

$$z = \frac{0.1198 - 0.05}{0.0111} = 6.29$$

Since we're more than 6 SD away from the mean, the P-value ≈ 0.

(d)-(f) are not graded, but here are the answers:

(d) If the true proportion of children with congenital abnormalities is 5%, then the chance of observing our sampled statistic (or worse) is essentially 0.

(e) Since the P-value is so low, we would ordinarily reject the null hypothesis and conclude that the proportion of children with congenital abnormalities is currently more than 5%. However, we would advise caution, since our sample may not have met the independence condition.

(f) No, we cannot infer a causal relationship between environmental chemicals and congenital abnormalities.

(pg.499:24) (a) In order to use these statistics for inference we must assume: (1) the observations in the sample are independent, or at least representative; and (2) the data are approximately normally distributed.

(b) Computations:

Confidence interval $= \bar{y} \pm t^*_{df} \cdot \frac{s}{\sqrt{n}}$

Sampled statistics: $n = 44, \quad \bar{y} = \$126, \quad s = \$15$.

$df = n - 1 = 43$, and for 90% confidence, we lookup tables, or software, and get: $t^*_{43} \approx t^*_{40} = 1.684$.

Confidence interval $= 126 \pm 1.684 \cdot \frac{15}{\sqrt{44}}$ $\boxed{= [\$122.2, \$129.8]}$

(c) Interpretation: With 90% confidence, the true mean daily income of the parking garage is between \$122.2 and \$129.8.

(d)-(e) are not graded, but here are the answers:

(d) "90% confidence" means: 90% of all random samples of size 44 will produce confidence intervals that contain the true mean.

(e) Since \$130 is outside the upper end of the confidence interval, we would conclude that the mean daily revenue of the parking garage is below \$130.