A San Diego reproductive clinic reported 49 live births to 207 women under the age of 40 who had been previously unable to conceive.

1. What is the population here?

2. A 95% confidence interval for \( \hat{p} \pm 1.96 \text{SE}(\hat{p}) \)

   Conditions: Randomization Condition, the 10% Condition, the Success/Failure Condition

   ```r
   > phat = 49/207
   > qhat = 1 - phat
   > se = sqrt ( phat*qhat/207)
   > zcrit = qnorm(.975)
   > zcrit
   [1] 1.959964
   > c( phat - zcrit * se , phat + zcrit * se)
   [1] 0.1788096 0.2946204
   ```

   (Homework problem: 90% confidence interval, change the critical value!)

3. Three approximations:
   (a) The binomial — 10% condition
   (b) The normal as approximation to the binomial — Success/Failure Condition
   (c) Standard error substituted for standard deviation
4. Eliminating the third approximation. The probability is (approximately) 95% that

\[ p - 1.96 \sqrt{\frac{pq}{n}} < \hat{p} < p + 1.96 \sqrt{\frac{pq}{n}} \]

After much algebra and the quadratic equation, this is equivalent to

\[ \hat{p} + \frac{1.96^2}{2n} - 1.96 \sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{1.96^2}{4n^2}} < p < \hat{p} + \frac{1.96^2}{2n} + 1.96 \sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{1.96^2}{4n^2}} \]

\[ \hat{p} + \frac{1.96^2}{2n} - 1.96 \sqrt{\hat{p}(1-\hat{p}) + \frac{1.96^2}{4\hat{n}}} < p < \hat{p} + \frac{1.96^2}{2n} + 1.96 \sqrt{\hat{p}(1-\hat{p}) + \frac{1.96^2}{4\hat{n}}} \]

\[ \hat{p} + \frac{1.96^2}{2n} - 1.96 \sqrt{\hat{p}(1-\hat{p}) + \frac{1.96^2}{4\hat{n}}} < p < \hat{p} + \frac{1.96^2}{2n} + 1.96 \sqrt{\hat{p}(1-\hat{p}) + \frac{1.96^2}{4\hat{n}}} \]

The last formula was messy. A simple substitute: The Agresti-Coull interval. Pretend that there are two more successes and two more failures in the sample! And then do what we did before. Bonus: Success/Failure condition reduces to 5.

\[ \hat{p} = \frac{y + 2}{n + 4} \quad \hat{n} = n + 4 \quad \text{confidence interval:} \quad \hat{p} \pm 1.96 \sqrt{\frac{\hat{p}(1-\hat{p})}{\hat{n}}} \]

5. The last formula was messy. A simple substitute: The Agresti-Coull interval. Pretend that there are two more successes and two more failures in the sample! And then do what we did before. Bonus: Success/Failure condition reduces to 5.

\[ n_{\text{tilde}} = 211 \]
\[ p_{\text{tilde}} = 51/211 \]
\[ \hat{p}_{\text{tilde}} = \text{sqrt} ( p_{\text{tilde}} * (1-p_{\text{tilde}})/n_{\text{tilde}}) \]
\[ c( p_{\text{tilde}} - z_{\text{crit}} * \hat{p}_{\text{tilde}}, p_{\text{tilde}} + z_{\text{crit}} * \hat{p}_{\text{tilde}}) \]
\[ [0.1839363, 0.2990876] \]

6. What’s the problem? For \( p \) close to 0 or 1, the binomial is noticeably not symmetric.

\[ \text{for } p \text{ close to 0 or 1, the binomial is noticeably not symmetric.} \]

\[ \text{prop.test(49,207,correct=F)} \]
\[ 1\text{-sample proportions test without continuity correction} \]
\[ \text{data: 49 out of 207, null probability 0.5} \]
\[ X^2 = 57.3961, df = 1, \text{ p-value = 3.563e-14} \]
\[ \text{alternative hypothesis: true } p \text{ is not equal to 0.5} \]
\[ 95 \text{ percent confidence interval:} \]
\[ 0.1839363 \text{ 0.2990876} \]
\[ \text{sample estimates:} \]
\[ p \]
\[ 0.2367150 \]

\[ \text{Homework} \]

(a) Read pages 470–473.

(b) No further problems assigned at this time.