Key for HW #4

1. 
   **#11**
   a. Ha  b. Ho  c. Ha  d. Ho  e. Ha

   **#17**
   The 1% figure refers to a p-value of 0.01. The researchers’ conclusion can be restated correctly like this: There is strong evidence (p-value=0.01) that the sex ratio of children born to families in the native community in Ontario is different from the continental average sex ratio. A correct interpretation of the 1% figure is like this: Had the sex ratio in the native community been equal to the continental sex ratio, there would have been a 1% chance of observing the results they had in their experiment, or results even more extreme.

   **#18**
   a. **Yes,** if we already know some biology about snakes, i.e. that snakes are cold-blooded animals who use external factors to warm up. The problem description suggests that the one-sided alternative hypothesis was formulated before the data was gathered, which is appropriate.

   b. Ho: Snakes have no preference (probability p=0.50 of choosing either site).

   Ha: Snakes have a preference for one or the other sites (probability p not equal to 0.50)

   c. 2* one-sided p-value=2*0.03=**0.06**

2. 
   **#23.**
   Y=# of observing the wrinkled pea among 12 pea plants

   Y~B(12, 1/4)

   a. E(Y)=n*p=12*0.25=3

   b. 
      \[
      \begin{align*}
      \Var(\hat{p}) & = \frac{p(1-p)}{n} = \frac{0.25*0.75}{12} = 0.015625 \\
      \SD(\hat{p}) & = \sqrt{0.015625} = 0.125
      \end{align*}
      \]
c. 
\[ Var(\hat{p}) = \frac{p \cdot (1 - p)}{n} = \frac{0.25 \cdot 0.75}{12} = 0.015625 \]

d. 
\[ P(Y = 2) = \frac{12!}{10! \cdot 2!} \cdot 0.25^2 \cdot 0.75^{10} = 0.23229 \]

3. 
#20.

a. 
Success rate = \( \hat{p} = \frac{4}{10} = 0.4 \)
\[ Y \sim B(10, \frac{1}{5}) \]
\[ E(\hat{p}) = E\left( \frac{Y}{10} \right) = \frac{10 \cdot 0.2}{10} = 0.2 \]

b. 
Ho : p=0.2 vs Ha : p>0.2  
(It is a one-sided test.)
\[ Y \sim B(10, \frac{1}{5}) \), under H_0 \]
mean = 10 \cdot 0.2 = 2, under H_0  
Observed data: 4, which goes in the direction of Ha. 
\[ p - value = P(Y \geq 4) = 1 - P(Y \leq 3) \]
\[ = 1 - \{ P(Y = 0) + P(Y = 1) + P(Y = 2) + P(Y = 3) \} = 0.1208739 \]
Since the p-value is larger than 0.05, we can't reject Ho.

c. 
The success rate is 0.35 which is lower than that of (a). But, after we do the hypothesis test, we will obtain a p-value. 
Ho : p = 0.2 vs Ha : p > 0.2  (It is a one-sided test.)
\[ Y \sim B(1000, \frac{1}{5}) \), under H_0 \]
\[ n \cdot p = 1000 \cdot 0.2 = 200 > 5 \]
\[ n(1 - p) = 1000 \cdot 0.8 = 800 > 5 \]
 \( (p \) comes from H_0 (in this case, p = 0.2) \)
\[
\frac{Y}{n} = \hat{\rho} \sim N(0.2, 0.01264911), \text{ approximately}
\]

Observed data: \( \frac{350}{1000} = 0.35 \), which goes in the direction of \( H_a \)

\[ p \text{- value} = P(\hat{\rho} \geq 0.35) \]
\[ = P\left( \frac{\hat{\rho} - 0.2}{0.01264911} \geq \frac{0.35 - 0.2}{0.01264911} \right) \]
\[ = P(Z \geq 11.85854) < 0.00002, \quad Z \sim N(0,1) \]

Since the \( p \)-value is smaller than 0.001, we can reject \( H_0 \) strongly.

So, for (a), we can't reject \( H_0 \). But, we can reject \( H_0 \) for (c).

This result shows us that the biological importance in (c) is smaller than that in (b), but the statistical significance in (c) is stronger than that in (b).

4.

#21.

a.

95% Confidence interval (Agresti-Coull method)

\[ \tilde{p} = \frac{6101 + 2}{9821 + 4} = 0.6211705, \]

\[ [\tilde{p} - 1.96\sqrt{\frac{\tilde{p}(1-\tilde{p})}{n + 4}}, \tilde{p} + 1.96\sqrt{\frac{\tilde{p}(1-\tilde{p})}{n + 4}}] \]
\[ = [0.6115783, 0.6307627] \]

b.

\( H_0 : p = 0.5 \) \text{ vs } \( H_a : p \neq 0.5 \)

\( Y \sim B(9821, 0.5), \text{ under } H_0 \)

\( n \cdot p = 9821 \cdot 0.5 = 4910.5 > 5 \)

\( n \cdot (1-p) = 9821 \cdot 0.5 = 4910.5 > 5 \)

\( (p \text{ comes from } H_0 \text{ (in this case, } p = 0.5 \text{) } ) \)

\( \hat{\rho} \sim N(0, 0.00504536), \text{ approximately} \)

Observed data: \( \frac{6101}{9821} = 0.6212198 \)

\[ p \text{- value} = 2 \cdot P(\hat{\rho} \geq 0.6212198) \]
\[ = 2 \cdot P\left( \frac{\hat{\rho} - 0.5}{0.00504536} \geq \frac{0.6212198 - 0.5}{0.00504536} \right) \]
\[ = 2 \cdot P(Z \geq 24.026) \]
\[ < 0.00004 \]

Since the \( p \)-value is smaller than 0.001, we reject \( H_0 \) strongly.