In the 1980s it was generally believed that congenital abnormalities affected about 5% of the nation’s children. Some people believe that, in recent years, the increase in the number of chemicals in the environment has led to an increase in the incidence of abnormalities. A recent study with 384 children found that 46 of them showed signs of an abnormality. Can we infer that the risk has increased? Carry out a hypothesis test at a 5% significance level. Solution must show all steps including hypotheses, check of conditions, sampling distribution model (with sketch), all calculations, and conclusion.

Solution
Let $p$ = proportion of children with congenital abnormalities.

* Null hypothesis $H_0 : p = 0.05$
* Alt. hypothesis $H_A : p > 0.05$

Test is 1-tailed, since we want to know if there is an increase in the incidence of abnormalities.

* Check conditions/ assumptions that the sample must satisfy:

1. Sample random, or representative? Problem doesn’t clarify. Assume yes, but interpret results with caution.
2. Is $n < 10\%$? It is true that 384 $< 10\%$ of the population (i.e., nation’s children).
3. Sample large enough? Check whether at least 10 successes/failures:
   
   $np_0 = 384 \times 0.05 = 19.2 > 10 \checkmark$ 
   
   $n(1 - p_0) = 364.8 > 10 \checkmark$

* Sampling distribution model (based on null hypothesis) is normal $N(0.05, 0.01112)$:

   \[
   SD = \sqrt{\frac{0.05(1 - 0.05)}{384}} = 0.01112
   \]

   \[
   \hat{p} = \frac{46}{384} = 0.1198
   \]

Want to find the area shown above, for $\hat{p} > 0.1198$.

   \[
   z = \frac{0.1198 - 0.05}{0.01112} = 6.276
   \]

   From calculator: $P$-value = $\text{normalcdf}(6.276, 10) = 1.7 \times 10^{-10}$ (or z-table).

   Therefore, $P$-value is, pretty much, 0.

* Conclusion: This $P$-value is well below the 5% significance level, indicating the probability of sampling variability causing this much deviation from $H_0$ is extremely small. We reject null hypothesis, and conclude the rate of abnormalities is higher than 5%.

Grading: Total points possible = 10.

1 pt - Any reasonable effort.
2 pt = correct hypotheses.
2 pt = know & check the right assumptions.
1 pt = correct mean & SD of sampling distribution model
1 pt = sketch of sampling distr. model showing what we want to find.
1 pt = correct $\hat{p}$ and $z$-score.
1 pt = find correct $P$-value.
1 pt = conclusion & interpretation.