Quiz 1 Hints and Comments

Stat 134, Spring 2020, Kolesnik

2/15/2020

These hints and comments are meant to help you fix your work if you missed any questions. They contain common sources of error and hints to push you in the right direction.

Question 1:

The key here is to recognize from the setup of the problem that since $\Phi(-z, z)$ has exactly the same value as $P(|p - \hat{p}| \leq k)$, we have that $z\sigma_{\hat{p}} \leq k$. To see this, rewrite $P(|p - \hat{p}| \leq k) = P(\frac{|\hat{p}-p|}{\sigma_{\hat{p}}} \leq \frac{k}{\sigma_{\hat{p}}}) = \Phi(\frac{k}{\sigma_{\hat{p}}}) - \Phi(-\frac{k}{\sigma_{\hat{p}}}) = \Phi(-\frac{k}{\sigma_{\hat{p}}}, \frac{k}{\sigma_{\hat{p}}})$.
Note that $\mathbb{E}(\hat{p}) = p$.
Common Mistakes:

1. Proportion of heads vs number of heads

   It is true that the proportion of heads $\hat{p} = \frac{X}{n}$, where X is the number of heads. Most of you converted the probability given into a statement about the number of heads, which is a perfectly valid approach. However, some of you said that $P(|p - \hat{p}| \leq k) = P(|n - n\hat{p}| \leq nk)$ which is not true, because $np \neq n$ unless $p = 1$, and since $pq \leq \frac{1}{4}$, you know $np \neq n$.

   Another common mistake in this process is to not recognize that $\mathbb{E}(\hat{p}) = p$, or in other words, that $\mathbb{E}(p - \hat{p}) = 0$.

   Lastly, there's continuity correction on $\hat{p}$, the proportion of heads. $\hat{p}$ is a continuous random variable. This means that when you do normal approximation, you cannot use continuity correction because, as the name suggests, it's a correction value applied to discrete random variable to make up for its lack of continuity.

Question 2:

A lot of you seemed to have struggled a little with understanding the setup of this question, so we'll explain the setup here. The lot has 5 defective items, so that is the size of the bad population. And the rule for passing a lot is that if there is no defective item in the first sample then it's good. However, if there is one defective item, we will take another larger sample from the remaining items, and pass the lot if and only if there is no defective item in the new sample.
Common Mistakes:

1. Independence

The samples are drawn without replacement. This means that the trials are not independent, so for every item you draw the probability that it is defective is different depending on your previous draws. Instead of a binomial distribution which requires independence, this problem should be solved using hypergeometric distribution.

2. Hypergeometric distribution parameters and formula

   The parameters of a hypergeometric distribution are: G which is the size of the good population, N which is the size of the total population, and n which is the sample size. The formula of $P(X = g)$, where X is the number of good items in a sample of size n drawn from a population of size N with G good items is given by $\frac{\binom{G}{g}\binom{N-G}{n-g}}{\binom{N}{n}}$. As can be seen from this formula, you need to make sure the top numbers in the combination factors in the numerator add up to the top number in the combination factor in the denominator. The same goes for bottom number as well.

3. Complement

   Some of you tried to use the complement rule to solve this problem. Note that the complement to the lot passing is actually not very easy to write down. You have to be comprehensive if you chose to do it this way. For a lot to be rejected, it has to

   $$\begin{cases} \text{has at least 2 defective items in the first draw, OR} \\ \text{has 1 defective item in the first draw AND at least 1 defective item in the second draw} \end{cases}$$

   Most of you missed the second case above, leading to an answer that is not complete.