Example of inference for a percentile

In a recent year the first quartile of SAT verbal test scores nationally was $\theta_{.25} = 430$. A particular school district claims that their new instructional methods will lead to an increase in the first quartile of SAT verbal scores. In a random sample of 100 students from the district, 19 students have SAT verbal scores below 430. Does this data support their claim?

Let $H_0 : \theta_{.25} = 430$ versus $H_a : \theta_{.25} > 430$. If X_i is a student's SAT verbal score, then under H_0 we have $P(X_i < 430) = .25$, and $B =$ the number of students having scores below 430, is binomially distributed with $p = .25$ and $n = 100$. Under H_a we expect to have fewer scores under 430, so our hypotheses about the quartile $\theta_{.25}$, $H_0 : \theta_{.25} = 430$ versus $H_a : \theta_{.25} > 430$ translate into hypotheses about $p = P(X_i < 430)$ as $H_0 : p = .25$ versus $H_a : p < .25$. To calculate the P value for this test, we calculate the probability of observing $B \leq 19$ under H_0, which is:

$$P \text{ value} = P(B \leq 19 | n = 100, p = .25) = \sum_{x=0}^{19} \binom{100}{x} (.25)^x (.75)^{100-x} \ .$$

We can calculate this probability directly or use the normal approximation to obtain a Z value:

$$Z = \frac{19 - np}{\sqrt{np(1-p)}} = \frac{19 - (100)(.25)}{\sqrt{(100)(.25)(.75)}} = \frac{-6}{4.33} = -1.38.$$

The approximation can be improved by using a continuity correction:

$$Z \text{ (with continuity correction)} = \frac{19.5 - np}{\sqrt{np(1-p)}} = \frac{19.5 - (100)(.25)}{\sqrt{(100)(.25)(.75)}} = \frac{-5.5}{4.33} = -1.27.$$

It turns out that $P(B \leq 19 | n = 100, p = .25) = .10$, $P(Z < -1.38) = .084$, and $P(Z < -1.27) = .102$, so at the $\alpha = .05$ significance level we would not reject H_0 in favor of H_a.

To calculate a confidence interval for $\theta_{.25}$, we can use the same reasoning for the median and use:

$$P(X_{(a)} < \theta_{.25} < X_{(b)}) = \sum_{x=a}^{b-1} \binom{n}{x} (.25)^x (.75)^{n-x}$$

or its normal approximation to obtain an (approximate) confidence interval. For the SAT issue above we can use a normal approximation to solve for a and b for a 95% confidence interval for $\theta_{.25}$ by using:

$$\frac{(b-1)-np}{\sqrt{np(1-p)}} = \frac{(b-1)-25}{4.33} = 1.96$$

and

$$\frac{a-np}{\sqrt{np(1-p)}} = \frac{a-25}{4.33} = -1.96$$

to obtain $a = 16.51$ and $b = 34.49$. Using the book's rounding rule (which can lead to less than 95% coverage) we would use $(X_{(17)}, X_{(34)})$ as our 95% confidence interval for $\theta_{.25}$.