Hypothesis Testing for Sampling

Anastassia Kornilova

Let $x_1 \cdots x_N$ be a population with an unknown mean μ. Suppose we pick a sample without replacement of $y_1 \cdots y_n$. We want to test if the sample sum will have the same sign as the population sum, by using a hypothesis test. Instead of testing for signs directly, sssume WLOG that $S_n > 0$. Let the null hypothesis be $S < 0$. Let:

$$S = \sum_{i=1}^{N} x_i = n\mu \quad S_n = \sum_{j=1}^{n} y_j$$

By the bound defined in Serfling in *Probability Inequalities for Sum In Sampling Without Replacement*:

$$P(S_n - S > t) \leq \exp(-2t^2 * N/(n(N-n-1)(b-a)^2))$$

where a and b are bounds on the population. Because $S < 0$ iff $S_n - S < S_n$, and because the above bound holds for any $t \geq 0$:

$$P(S_n - S < S_n) \leq \exp(-2S_n^2 * N/(n(N-n-1)(b-a)^2))$$

It follows that if:

$$\exp(-2S_n^2 * N/(n(N-n-1)(b-a)^2)) < \alpha$$

then we can reject H_0 with p-value of α, because in that case $P(H_0) < \alpha$.