HW1 problem 5 clarification (problem 4 in 2020/21)

5. Exact and approximate small-sample CIs for the mean.

For each of the cases 1-3 below, complete the steps (a)-(e) below.

For j=1,...,1000,

(a) set the random seed to j,

(b) generate a random sample of size 4 from Normal(0,1)

(c) compute a 2-sided 95% CI for the mean.

(d) record whether the CI containts the true mean (=0).

(e) for cases 2 and 3, also store the length of the CI

standard normal quantile

Case 1: sig2 is known and is equal to 1. CI : $\bar{X}_n \pm \Phi^{-1}(0.975) \cdot \frac{\sigma}{\sqrt{n}}$

Case 2: sig2 is unknown and exact small sample CI is computed in (c) [using t-distribution quantiles] : $\bar{X}_n \pm Q(0.975) \cdot S/\sqrt{n}$; S - sample st. dev.

Case 3: sig2 is unknown and an approximate large-sample CI is computed in (c) [using Normal(0,1) quantiles] $\bar{X}_n \pm \Phi^{-1}(0.975) \cdot S/\sqrt{n}$

Report

(1) the "empirical frequency of coverage" (average of (d))

(2) histogram of interval widths (when appropriate).

Discuss your findings (particularly, try to relate (1) and (2)).

A suggestion to stats/biostats folks: think LLN and CLT + exact vs approximate CIs.