**Problem 1:**This problem is designed to have you think through what a confidence level means, and reinforce why a confidence interval is more informative than a hypothesis test. There are a total of 5 tables; one for each of the distributions you considered in problem 2 of the previous assignment.

The stated confidence level is 95% (i.e. α=0.05).

The samples drawn for this simulation study are the same ones used for the sampling distribution assignment. The population distributions are provided below.

Discrete Cases: Symmetric Binomial (p = 0.5) andSkewed Binomial (p = 0.04), top and bottom rows respectively:



Continuous Cases: Uniform (1, 9) {symmetric}, Chi-Square (3) {right skewed; truncated at 0}, and Mixture distribution {right skewed, two peaks}:



|  |  |  |
| --- | --- | --- |
| Chi-Square |  | Skewed Binomial |
|   | n = 5 | n = 15 | n = 45 | n = 135 |  |   | n = 5 | n = 15 | n = 45 | n = 135 |
| Reps = 10 | 80 | 100 | 100 | 100 |  | Reps = 10 | 30 | 40 | 60 | 100 |
| Reps = 100 | 91 | 94 | 93 | 93 |  | Reps = 100 | 25 | 42 | 80 | 87 |
| Reps = 1000 | 89.7 | 91.9 | 94.2 | 94.4 |  | Reps = 1000 | 20.2 | 47.7 | 83.7 | 88.9 |
| Reps = 10,000 | 90.26 | 92.55 | 94.11 | 94.8 |  | Reps = 10,000 | 18.57 | 45.63 | 84.3 | 89.75 |
|  |  |  |  |  |  |  |  |  |  |  |
| Uniform Distribution |  | Symmetric Binomial |
|   | n = 5 | n = 15 | n = 45 | n = 135 |  |   | n = 5 | n = 15 | n = 45 | n = 135 |
| Reps = 10 | 80 | 70 | 90 | 100 |  | Reps = 10 | 100 | 90 | 90 | 80 |
| Reps = 100 | 94 | 88 | 92 | 96 |  | Reps = 100 | 95 | 89 | 94 | 93 |
| Reps = 1000 | 93.4 | 95.5 | 95.5 | 95.7 |  | Reps = 1000 | 91.6 | 87.6 | 94.2 | 94 |
| Reps = 10,000 | 93.77 | 95.03 | 94.93 | 95.14 |  | Reps = 10,000 | 93.33 | 88.35 | 92.98 | 93.96 |
|  |  |  |  |  |  |  |  |  |  |  |
| Mixture Distribution |  |  |  |  |  |  |
|   | n = 5 | n = 15 | n = 45 | n = 135 |  |  |  |  |  |  |
| Reps = 10 | 80 | 100 | 100 | 80 |  |  |  |  |  |  |
| Reps = 100 | 89 | 94 | 97 | 92 |  |  |  |  |  |  |
| Reps = 1000 | 90.8 | 93.8 | 95.3 | 94.2 |  |  |  |  |  |  |
| Reps = 10,000 | 90.8 | 93.95 | 94.72 | 94.59 |  |  |  |  |  |  |

**1.1)** While holding sample size fixed, comment on the empirical coverage probability versus the stated confidence level for each level of replication. If you observe a difference between coverage probability and stated confidence level, give your hypothesis as to why this may be occurring.

**1.2)**While holding the number of replications fixed, comment on the empirical coverage probability versus the stated confidence level for each level of sample size. If there you observe a difference between coverage probability and stated confidence level, give your hypothesis as to why this may be occurring.

**1.3)**If a confidence interval sits wholly above the hypothesized value, what can you say about the p-value relative to the significance level (α). Explain.

**1.4)**Putting it all together: Relate, what you have observed via simulation to the concepts of type I and type II error from the lecture notes. Why do you think I stated that confidence intervals provide more information than hypothesis tests?