

Statistical Literacy

Problem Set 8

Rémi Viné

Due November 20th, 2023

1. Choose the right distribution.

- (a) Constructing a 99% confidence interval on the average accidents at work in a given sector (number of accidents per year), you have a sample of 81 employees in the sector. What distribution should you use to find the critical value?
- (b) What is the critical value?
- (c) Constructing a 95% confidence interval on the average time to commute, you use a sample of 200 people and you use the standard deviation taken from last year's census (where all the population of interest was investigated). Assume that the population variance has not changed since then. What distribution is required to find the critical value?
- (d) What is the critical value?
- (e) If a statistician made a mistake and used the Standard Normal distribution rather than the Student distribution (T-distribution), what would be the problem in the critical value's calculation? And how would this affect the confidence interval in general?

2. Confidence intervals for means.

Three researchers are interested in three different variables in three different contexts. The first researcher investigated seats availability in buses in Dakar during rush hours and obtained a representative sample of 61 buses. The second researcher focused on the price of 1L water bottles in Geneva and obtained a representative sample of 101 bottles of water (with their respective prices) in the city. The third researcher looked at fossil fuel consumption per capita in India and gathered district level data for 500 districts, assumed to be representative of all 803 districts. Sample statistics are given in table (1).

Variable	Sample size	Average	Variance (sample)
Bus seats available in rush hours (Dakar)	61	2	3
Price of 1L of bottled water in Geneva, CHF	101	1.5	0.81
Fossil fuel consumption per capita in Indian districts (kWh)	500	6319	4900

Table 1: Some variables

- (a) Using the three variables in table 1 and constructing 95% confidence intervals for each, what variable would have the larger critical value?
- (b) Constructing a 99% confidence interval for each variable, for what variable would the interval have a larger rejection area (α)?
- (c) Construct and interpret a 80% confidence interval for the first variable of table 1.
- (d) Construct and interpret a 99.8% confidence interval for the second variable of table 1.
- (e) Construct and interpret a 95% confidence interval for the last variable of table 1.