Practice questions

- 1. A medical journal publishes analyses about the efficacy θ of treatments that can be proven to exceed some threshold at the 95% confidence level.
 - (a) If a journal issue publishes 30 analyses what is the probability that at least one of them does not meet the threshold? Assume that each analysis barely meets the 95% confidence level requirement (it does not exceed it) and that the analyses are mutually independent.
 - (b) If the journal editors want to be 95% confident that all 30 analyses are correct, how should they modify the requirement for each treatment?
- 2. The measurements of ten random athlete heights in centimeters are

152, 163, 188, 201, 192, 176, 194, 166, 215, 184.

- (a) Assuming the heights are independent normal random variables with known standard deviation $\sigma = 20$, give a 95% confidence interval for the mean height.
- (b) How many samples do you need for a 95% confidence interval of width 5cm?
- 3. A large company conducts a job satisfaction survey among its 6250 employees. Out of 250 employees that are sampled (with repetition), 142 are satisfied with their jobs.
 - (a) Calculate a 99% confidence interval for the number of employees that are satisfied with their job.
 - (b) Find a confidence interval of width 100 for the number of satisfied employees and estimate the confidence level for it.
- 4. Let MAX be the sample maximum for 10 independent samples of a Uniform $(0, \theta)$ random variable.
 - (a) What is the CDF of the random variable MAX/θ ?
 - (b) Find values z_{-} and z_{+} for which $P(MAX/\theta < z_{-}) = P(MAX/\theta > z_{+}) = 2.5\%$.
 - (c) Find a 95% confidence interval for θ based on your answers in part (b).
 - (d) **(Optional)** There are multiple 95% confidence intervals for θ . What is the narrowest one that you can find?

Additional ESTR 2020 questions

- 5. A continuous PowerLaw(α) random variable has PDF of the form $f(x) \propto x^{-\alpha}$ for x > 0. Here α is a parameter that can take values greater than 1 (for otherwise $\int_0^\infty f(x) dx$ diverges).
 - (a) Show that the sample mean of n independent PowerLaw(α) random variables has "heavy tails", for example $P(\overline{X} \ge \mu + \sigma) \ge cn^{-d}$ for some constants c and d that depend on α . (μ and σ are the mean and standard deviation of a PowerLaw(α) random variable.)
 - (b) From Chebyshev's inequality we know that $d \ge 1$. Can you come up with an α for which d = 1?
- 6. A discrete PowerLaw(α) random variable has PMF $f(x) \propto x^{-\alpha}$ over the positive integers. Such random variables have been considered to model connectedness in social networks, e.g., number of friends of a random facebook profile. How would you go about estimating the value of α from (independent) PowerLaw(α) samples? This is a possible project topic.