Practice questions

1. Let N and N' be independent Normal$(0,1)$ random variables. Find the probabilities that (a) $2N > N' + 1$ and (b) $(N - N')^2 + (N + N')^2 > 4$.

2. The midterm test scores of six random students are 81, 84 ,83, 73, 76, 83.

   (a) What is the sample variance?

   (b) Assuming their scores are independent Normal(μ, σ) random variables. Give as large a value for $\hat{\Theta}_-$ as you can so that $(\hat{\Theta}_-, 10)$ is a 95% confidence interval for σ.

3. A food processing company packages honey in glass jars. The volume of honey in millilitre in a random jar is a Normal(μ, σ) random variable. 5 random jars are picked and the volume of honey inside them in millilitre are 108, 101, 103, 109 and 104.

   (a) Suppose μ is unknown and σ is known to be 5. Give a 95% confidence interval for μ.

   (b) Suppose μ and σ are both unknown. Give a 95% confidence interval for μ.

   (c) Suppose μ and σ are both unknown. Give a 95% prediction interval for the next sample.

4. In this question you will study prediction intervals for normal random variables with known and unknown standard deviation σ.

   (a) If $X_1, \ldots, X_n, X_{n+1}$ are independent Normal(μ, σ) random variables, what kind of random variable is $X_{n+1} - (X_1 + \cdots + X_n)/n$?

   (b) Based on part (a) come up with a prediction interval for the next Normal(μ, σ) sample for unknown μ and known σ.

   (c) Assuming $\sigma = 5$, give a 95% prediction interval for the next sample from the following data: $37.1, 44.0, 49.3, 49.7, 37.9, 65.5, 63.7, 44.1, 58.3, 60.9$.

   (d) Repeat part (c) for unknown σ. How do the confidence intervals in parts (c) and (d) compare?

Additional ESTR 2020 questions

5. The Cauchy random variable is the ratio N/N' of two independent Normal$(0,1)$ random variables.

   (a) Show that the Cauchy random variable has a PDF that is symmetric around the origin and has infinite variance.

   (b) Let $X_1, X_2, \ldots$ be independent Cauchy samples. What does the sequence $(\overline{X}_1, \overline{X}_2, \ldots)$ of sample means $\overline{X}_n = (X_1 + \cdots + X_n)/n$ look like? Does it typically converge?

   (c) What happens if you replace $\overline{X}_n$ with the sample median MM_n of $\sqrt{n}$ means of $\sqrt{n}$ samples each (restricting n to perfect squares only?)

6. The median-of-means estimator for the mean μ requires knowledge of the variance σ^2 of the data. If the variance is not known, one possibility is to estimate the variance using

$$V = \text{median}\left\{\tfrac{1}{2}(X_1 - X_1')^2, \ldots, \tfrac{1}{2}(X_n - X_n')^2\right\},$$

   where $X_1, X_1', \ldots, X_n, X_n'$ are independent samples.

(a) Assuming that σ is finite, show that V rarely overestimates the variance by much, namely $\mathrm{P}(V > 3\sigma^2) < \exp(-cn)$ for some constant $c > 0$. What is the smallest c you can get? (**Hint:** Markov's inequality)

(b) Can you also be sure that $\mathrm{P}(V < \frac{1}{3}\sigma^2) < \exp(-cn)$?

(c) As a potential project, you can experimentally evaluate the confidence error of the resulting estimator for the mean of data with unknown variance on suitable data sets (and compare it with the sample mean estimator).