Homework 4

Your student ID #

Homework Guidance

To ensure the integrity and educational value of your work, please adhere to the following guidelines as you complete Homework 1:

1) Originality and Collaboration: You are encouraged to work independently and ensure that the submissions are your own. While collaboration with your peers for understanding concepts and discussing problems is allowed, directly copying work from your colleagues is strictly prohibited. Your submission should reflect your understanding and your ability to apply what you have learned.

2) Use of GPT-like Platforms: You are permitted to use GPT-like platforms for assistance with your homework. However, this tool should only be used when you fully understand the answers it provides. The purpose of using such platforms is to enhance your learning, not to bypass the learning process. ***Keep in mind that midterm and final exams will be conducted without internet access***. If there is a significant discrepancy between the code you submit for homework and your ability to write similar code during an exam, ***it will be considered cheating***. Such instances will result in a score of zero for the involved exam component.

3) Submission Quality: Your focus should be on submitting code that you comprehend thoroughly. Fancy or complex code that goes beyond your level of understanding is not the goal. We value honesty and genuine effort. Make sure that you can explain and justify every line of code you submit. This approach will not only help you in your homework but also prepare you for the no-internet exams.

Your coding HWs (check week03 pdf files) from 1 to 4 should be answered here:

1.

2.

3.

4.

5. Calculate both the nonparametric and parametric 95-percent confidence intervals for the median of the following ***Granodiorite*** data. Which of these interval estimates is more suited to this data, and why?

Granodiorite (mg/L) data:
[6, 0.5, 0.4, 0.7, 0.8, 6.0, 5.0, 0.6, 1.2, 0.3, 0.2, 0.5, 0.5, 10, 0.2, 0.2, 1.7, 3.0]

6. A yield of 0.70 gallons per minute per foot was recorded for a well in Virginia. Does this yield fit within the same distribution as the data from the ***VA wells*** dataset, or is it indicative of a larger trend? Determine this by calculating the relevant 90-percent parametric and non-parametric confidence intervals. Assess which type of interval is better suited for this data.

VAwells (gallons/min/foot) data:
[0.001, 0.03, 0.1, 0.003, 0.04, 0.454, 0.007, 0.041, 0.49, 0.02, 0.077, 1.02]

7. Create the most suitable 95-percent confidence interval estimates for both the mean and median annual streamflows (cubic feet per second) of the Conecuh River at Brantley, Alabama,

as given below. Employ a bootstrap technique for estimating the median, in addition to the conventional method that utilizes the binomial distribution. For the mean, apply the parametric method and also explore a bootstrap method for estimation.

Year	Flow (cfs)
1941	369
1942	683
1943	923
1944	1193
1945	413
1946	1025
1947	894
1948	859
1949	1157
1950	524
1951	327
1952	574
1953	762
1954	578
1955	379
1956	374
1957	581
1958	581
1959	530
1960	929

8. (Bonus point: +5, which is equivalent to the value of 5 homework problems.)

A water intake is located on the Potomac River at Little Falls, just above Washington, D.C. We want to select a design discharge for this water intake so that in only 1 year out of 10 does the 1-day minimum discharge of the river go so low that the intake becomes inoperable because it is not submerged in the river.

(Imagine there is a spot on the Potomac River, near Little Falls just before it flows through Washington, D.C., where we've installed a device to take water from the river. This device needs to be underwater to work properly.
Now, we are trying to figure out the least amount of water (design discharge) that needs to be flowing in the river for our device to keep working. We want to be really cautious and plan for a scenario where, in a bad year, the river's flow drops to its lowest level for just one day. But, we want to make sure this kind of low flow only happens once every ten years. In other words, we want our water intake to stay operational 9 out of 10 years, even on the day when the river flow is at its minimum.)

But, we want to be conservative in our selection of this design discharge because we know that an estimate of the 0.1 frequency annual minimum flow is somewhat uncertain. We want to set the design discharge such that there is only a 5-percent probability that the true 0.1 frequency annual minimum flow is below the design discharge. Our dataset (see below) consists of the annual 1-day minimum discharge (Q) at this site for the 84-year period 1932–2015. Use the concept of a nonparametric one-sided confidence interval for the 0.1 quantile on the distribution of annual minimum discharges.

year	discharge
1932	21.38
1933	16.57
1934	52.1
1935	33.7
1936	46.44
1937	34.55
1938	56.92
1939	35.96
1940	30.02
1941	60.03
1942	18.18
1943	65.13
1944	24.18
1945	24.18
1946	62.86
1947	28.32
1948	30.58
1949	61.45
1950	59.18
1951	44.46
1952	34.83
1953	49.84
1954	31.15
1955	24.92
1956	37.38
1957	60.6

1958	23.45
1959	36.53
1960	20.5
1961	31.15
1962	40.78
1963	29.73
1964	19.99
1965	10.59
1966	12.29
1967	3.43
1968	47.01
1969	22.68
1970	16.93
1971	33.13
1972	52.39
1973	61.45
1974	61.45
1975	41.63
1976	86.08
1977	31.71
1978	22.54
1979	47.01
1980	88.35
1981	33.98
1982	23.9
1983	37.66
1984	33.98
1985	64.56
1986	28.12
1987	16.34
1988	21.44
1989	21.01
1990	63.71
1991	42.48
1992	14.55
1993	45.31
1994	25.17
1995	46.16
1996	26.96
1997	155.74
1998	35.11
1999	27.95

2000	4.93
2001	48.7
2002	23.22
2003	7.31
2004	119.5
2005	52.39
2006	26.33
2007	27.41
2008	17.16
2009	28.2
2010	33.7
2011	11.47
2012	39.93
2013	48.7
2014	44.74
2015	44.46