SS2864B, 2021

Assignment #2 due to February 22, 11:55pm, 2021

Instructions Submit an electronic version (pdf, words, etc) of your solutions (appropriately annotated with comments, plots, and explanations) to owl. Save all your R codes in one script file with proper comments and submit it as well to owl.

1. Write an R function to calculate both

   (a) $\sum_{i=1}^{n} \min(2^i, i^3)$,

   (b) $\sum_{i=1}^{n} \max(2^i, i^3)$,

   i.e., both sums need to be calculated in one function and should return two values. Make sure to have two error checkings on input n and test your function for a few bad/wrong n's. Then execute your function for $n = seq(200, 5000, by = 600)$.
   Note: **for**, **while**, or **repeat** looping is not allowed. Please check the usages of R functions **pmin** and **pmax**. You can use R function **sapply** to carry out the last step calculation.

2. For iid sample $X_1, \ldots, X_n$ from population distribution $F(x)$ with mean μ and variance $\sigma^2 < \infty$, the 95% confidence bands for μ can be computed as

   $$\text{low.band} = \bar{X} - 1.96 * \sigma/\sqrt{n}; \ \text{high.band} = \bar{X} + 1.96 * \sigma/\sqrt{n},$$

   provided that σ is known, where $\bar{X}$ is the sample mean.

   (a) For given a vector x and $\sigma = 3$ as inputs, write an R function called **my.conf** to compute those low and high bands and choose a proper return value(s). Please do one error checking on x.

   (b) If σ is unknown, σ can be replaced by the sample standard deviation of x (to compute confidence bands). With only input x, write an R function called **my.conf2** to compute those two bands. You can reuse the function implemented in (a).

3. Implement an R function, say **IQR.outliers**, to compute the inter-quartile-range (IQR), find outliers if they exist. The function should have one argument x. In the function body,

   (a) Do at least two error checkings on x before any computation.

   (b) Compute the IQR as Q_3 (3rd quartile) - Q_1 (1st quartile).

   (c) Use 1.5*IQR rule to detect suspected outlier(s) if they exist. First check if there is(are) value(s) of x that is(are) less than Q_1-1.5*IQR. If so, this (those) is(are) suspected outlier(s) on left tail. Similarly, checks any value(s) of x over Q_3+1.5*IQR on right tail.

   (d) Please choose a proper output object to represent required information. The return values must contain IQR and outlier(s) from left and/or right tail.

   Test your function with the variables **weight** and **Time** in the data.frame **ChickWeight**. Also test your function with wrong inputs (checking if testing procedures are working or not).

4. NASA's GISS Surface Temperature Analysis (GISTEMP) is an estimate of global surface temperature change recorded monthly. The detail information can be found in https://data.giss.nasa.gov/gistemp. A file, GLB.Ts_dSST.csv, has uploaded to owl in Data sets folder. It contains monthly temperatures from 1880 to 2020. Please do the following steps to extract some basic information. Any looping such as for, while, repeat is not allowed.

   (a) Import the dataset into R as a data frame. Create another data frame to keep only **Years**, **Jan**, ..., **Dec** 13 columns.

   (b) Write an R function with input x to calculate the mean of x without the first element $x[1]$. Do one error checking on x. Then with the help of R function **apply** to generate a vector of yearly average temperatures from 1880 to 2020 and use **plot.ts** to plot it as a time series. Comment your findings.

   (c) Find the vector (**temp.1900**) for the monthly temperatures from Jan to Dec for the year 1900. Also find two vectors (**temp.1960** and **temp.2020**) for the monthly temperatures from Jan to Dec for the years 1960 and 2020 respectively. Run the following codes

```
plot.ts(temp.1900, ylim=c(-0.35, 1.25))
lines(temp.1960, col = "blue")
lines(temp.2020, col = "red")
```

   and comment your findings.

5. Suppose Ms. Wu wishes to take out a mortgage on a house. She wants to know what her periodic payment will be. If P is the initial amount mortgage, $i.r$ is the effective interest rate, and n the length of the mortgage, then the periodic payment R is given by

$$R = \frac{P\, i.r}{1 - (1 + i.r)^{-n}}.$$

   (a) Construct a function called **mortgage.payment** which employs this formula.

   (b) Calculate Ms. Wu's monthly payments, if the initial amount is \$50,000, the interest rate is 2% and the number of interest conversion periods is 300.