

Examining Records (Task 2 continued)

1. The objective of this exercise is to see how many records one might expect from certain types of data. To get started, import 99 years (1895-1993) of temperature data into Splus by following the menu options File > Import Data > From File. After the Import Data dialog box opens, Look in folder G, subfolder st192, change the Files of type to Microsoft Excel Files [*.xl*], and select the Excel file named tundra. A new data structure called tundra will be created. This data structure acts like a matrix with the rows corresponding to years and the columns to months.

```
> tundra[,1] [displays average maximum temperature for Jan for the 99 years]
> tundra[,7] [displays average maximum temperature for July for the 99 years]
> tundra[2,] [displays the temperature for the 12 months in 1896]
> tundra[3,7] [displays the temperature for July of 1897]
> tsplot(tundra) [plots the temperature data for each of the 12 months]
```

a) For each month, count the number of record highs and record lows. Display the results in a table. You might think about an efficient way to "count" these records in Splus.

b) Construct table for reverse highs and lows obtained by reading the data in reverse order. (The Splus command `rev(tundra[,1])` reverses the order of the January data.)

c) Euler's constant. Plot the function $f(n) = \sum_{j=1}^n \frac{1}{j^{2.1}}$, as a function of $n = 1000, \dots, 100,000$. Does $f(n)$ appear to asymptote?

2. What do you think the probability is of observing 10 or more records in a sample of size 100? Approximate this probability via simulation. That is, in 1000 simulations of a sample of size 100, calculate the proportion of simulations for which the number of records is 10 or more.