$\qquad$
May 3, 2011

Write, compile, and test a $\mathrm{C}++$ program that determines the slope and $y$-intercept of the line that best fits (in the least-squares sense) a collection of data points of the form $\left(x_{i}, y_{i}\right)$. The work-horse for your program should be a function named LinReg with a function header line similar to

```
short LinReg( int n, double x[], double y[], double &slope,
    double &yint, double &r )
```

This function should be called by using the syntax $\operatorname{LinReg}(\mathrm{n}, \mathrm{x}, \mathrm{y}$, slope, yint, r ), where

- n is the number of data points;
- x is a one-dimensional array containing the $x$-values;
- y is a one-dimensional array containing the corresponding $y$-values;
- slope is the computed slope of the best fit line;
- yint is the computed $y$-intercept of the best fit line;
- $r$ is the linear correlation coefficient that measures the strength of the linear relationship between the paired $x$ - and $y$-values; and
- the return value is 1 or -1 according to whether the function is successful or not.

Use the following formula for the linear correlation coefficient:

$$
r=\frac{\sum_{i=1}^{n}\left(x_{i}-\bar{x}\right)\left(y_{i}-\bar{y}\right)}{\sqrt{\sum_{i=1}^{n}\left(x_{i}-\bar{x}\right)^{2} \sum_{i=1}^{n}\left(y_{i}-\bar{y}\right)^{2}}}
$$

The main function should read the collection of data points from a data file until end-of-file is reached. Your data file should be formatted so that there is exactly one $(x, y)$ pair on each line. For example...

$$
\begin{array}{ll}
53.0 & 80 \\
67.5 & 344 \\
72.0 & 416 \\
72.0 & 348 \\
73.5 & 262 \\
68.5 & 360 \\
73.0 & 332 \\
37.0 & 34
\end{array}
$$

Program output should include number of data points in file, slope of best fit line, $y$-intercept of best fit line, and linear correlation coefficient.

