Normal Data Sets Report Form

Section: ______ Team number: _____

We affirm in accordance with the Duke Community Standard that this data was gathered by the members of our team and that all the work in this lab is our original work. [Sign below.]

Team members: ____________________________ ____________________________

____________________________ ____________________________

Analyzing the Data

1. What did you measure? Be sure you describe your measurements precisely enough for someone else to be able to read your report and to duplicate your experiment. How many measurements did you make?

2. How did you gather your data? Describe briefly the process and the source(s).

The Mean and Standard Deviation of Discrete Data

1. Suppose that $\{70, 65, 72, 69, 71, 68, 75\}$ are heights of seven men. Compute $\overline{x}$ and s both by hand and also by using the statistical features of your calculator. All members of the team who completed this step should initial below:

_____________ _____________ _____________ _____________

2. The value, s, is supposed to measure how spread out the data points are; thus, computing $(x_k - \overline{x})$ may seem like a good idea, but can you explain why that measure is NOT used and why it's a good idea to add the *squares* of these terms rather than simply the terms $(x_k - \overline{x})$?

The Mean and Standard Deviation of Your Data

1. Attach a list of your data to this report. All team members who helped to gather the data should sign the list.

2. Compute the mean and standard deviation of your data. Record this information below.

Sample mean, $\overline{x}$: ______________ Sample standard deviation, s: _____________

3. If the data were normally distributed, what fraction of the data would you expect to fall within one standard deviation (*sd*) of the mean (*m*)? Answer: ________________

Find the fraction of your data values that actually fall in the interval $[m - sd, m + sd]$. Is it consistent with normal data? Display and explain your calculations.

Approximating the Distribution and Density Functions for Your Data

1. Let *a* and *b* represent numbers a little below and a little above the low and high values, respectively, in your data. Divide the interval [*a*,*b*] into n subintervals of equal width, $\Delta x = \frac{b-a}{n}$. We will designate these endpoints by $x_0,\ x_1,\ x_2,\ \ldots, x_{n-1}, x_n$. Note that $x_0 = a$ and $x_n = b$. Fill in the blanks below.

$a =$ ________ $b =$ ________ $n =$ ________ $\Delta x =$ _________

$x_0 =$ _______ $x_1 =$ _______ $x_2 =$ _______ $x_{n-1} =$ _______ $x_n =$ _______

2. For each x_k, compute by hand the corresponding values of the distribution function $F(x) = P(X \leq x)$. For example, $F(a) = 0$ and $F(b) = 1$. Display two sample computations below.

3. Show a plot of the values of $F(x_k)$ in an x-y plane below. Do these points approximate a typical distribution curve? Explain your answer.

4. Approximate values of the density function for your data. Show two sample computations below.

5. Use your calculator to make a plot in the x-y plane of the approximate values you just computed for F'. Now, <u>superimpose</u> the graph of $f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ over this previous plot. For σ and μ you should use, of course, the values of the *standard deviation* and *mean* that you computed for your data. Attach (or carefully draw) a copy of the plot and graph together.

6. Do you think your data is normally distributed? Justify your conclusion. Comment on any unusual or inconsistent parts of your data, and explain how someone could improve upon the study that you have done.