Instructions: Always show how you set up the problem, in order to receive partial credit for a problem. On most problems, you don’t need to get the numerical answer, so long as you show how you would proceed to get the answer by inserting the appropriate numbers into the correct formula. You may use the book only for the tables in the back.

1) (18 points) The annual numbers of death from tornadoes in the US from 1990 through 2000 are (as given by NOAA):

\[
\begin{array}{cccccc}
53 & 39 & 39 & 33 & 69 & 30 \\
25 & 67 & 130 & 94 & 42 \\
\end{array}
\]

a) Give a stem and leaf plot for the data.
b) Some summary statistics are:
\[ M = 42, \quad Q_1 = 33, \quad Q_3 = 69, \quad IQR = 36 \]
Sketch a boxplot for these data. (Give a few immediate calculations so one can follow how you got the boxplot).

c) Suppose that a journal article does not give the individual observations, but states that the mean is 56.5 and the standard deviation is 31.9 for these 11 observations. Use the Chebyshev Inequality to make some statement, with some given probability, on the maximum number of tornadoes you would expect to find in a given year.
2) (25 points) For the population of females between the ages of 3 and 74, the distribution of hemoglobin levels has mean 13.3 (g/100 ml) and standard deviation 1.12 (g/100 ml). We assume this distribution is approximately normal.

a) What is the probability that a random female has a hemoglobin level of 12.0 or less, which is considered to be low? (If you can’t find it, you may assume henceforth that the probability is 0.25.)

b) This class has 20 females. How many would you expect to have a level of 12 or less?

c) Find the probability that only one of the 20 females has a level of 12 or less. (You just need to show the equation to find this.)

d) What is the probability that the mean hemoglobin level for the 20 females in this class is 12 or less?
3) (30 points) The winglength of the common adult housefly is a variable that is known to be normally distributed. The mean of this population is 45.5 (in units of 0.1 mm) with a population variance of 15.2. A scientist believes that a new mutant species of houseflies has been found, which is characterized by a shorter winglength. Suppose $n = 5$ of these new houseflies are sampled, and that they have a sample mean of 41. We assume they have the same population variance of 15.2. Assuming $\alpha = .05$, do you have statistical evidence to claim that this mutant population has a reduced mean winglength, i.e. a mean less than 45.5?

$H_0 : $ ____________________________

$H_a : $ ____________________________

Test Statistic: ____________________________

Rejection Region: (in terms of $t, z$ or $X$):

Calculations:

Decision: ____________________________
4) (25 points) The scientist wonders whether a sample size of 5 is sufficient.

   a) Find the power of the hypothesis test in Problem 3 to reject the null hypothesis and thus detect the new species if the true winglength for mutants is 43.2 (which is a 5% reduction in the previous mean.)

   b) How large should the sample size be in order to detect a 5% change in mean winglength (i.e. from 45.5 to 43.2) with a power of 90%?
5) (12 points) Short Answer.

a) In problem 3, a colleague takes another sample of \( n = 5 \) houseflies, and finds a sample mean of \( \bar{X} = 47 \). What is the \( p \)-value for this result with the hypothesis test in problem 3?

b) The confidence interval for \( \mu \) is often misinterpreted. Suppose a journal gives a 95% confidence interval of (38,44) for \( \mu \) in Problem 3, and states that there is “a 95% chance that the interval contains \( \mu \).” Is this a correct interpretation? If so, state why. If not, give a correct interpretation for the given 95% confidence interval.