Minitab Handout 2

This handout illustrates some Minitab commands and interpretations of the outputs useful for doing Homework 2. Minitab menu sequences are in bold. First, read your data set into Minitab as you did for graphical summaries.

- Numerical measures:
  See **Minitab Handout 1** for the menu sequence. For five-number summary, you will need `Minimum`, `Q1`, `Median`, `Q3`, and `Maximum` from the numerical summary. `Mean` and `StDev` represent the mean $\bar{x}$ and the standard deviation s of the variable specified. If your data set is reasonably small, it may be a good idea to compare your hand-calculated results with the Minitab output. You may notice that Minitab has automated procedures of finding the quartiles, which are slightly different from those described in the textbook. However, differences would be negligible, so you can go with either Minitab answers or your own calculations. Limited precisions in Minitab output may give slightly different numerical summaries as well.

- Boxplot:
  In the textbook, it is explained that the lines from the central box extend to the minimum and the maximum. Minitab produces a modified box plot by default. This means that the lines (or whiskers) stretch to the greatest value less than or equal to (the third quartile plus 1.5× IQR) and the smallest value greater than or equal to (the first quartile minus 1.5× IQR). Any observations beyond the upper or the lower limit are plotted individually.

- How to eliminate outliers from Minitab Worksheet for recalculation of $\bar{x}$ and s?
  There are more than one way to do this. After eliminating outliers by one of the following options, use the same menu sequence as before to get numerical summaries.

  (i) Identify the value(s) to eliminate in Worksheet and replace them by the asterisk `*`, which is the symbol of a missing value in Minitab. In this case, if you save the modified data set under the same name as the original data, you will lose the value(s) you have eliminated.

  (ii) You may create a new variable by copying. **Manip; Copy Columns...** and select the variable you want to copy from and type a new name in the `To columns` box, then click **OK**. Apply (i) to the new variable to eliminate value(s). This allows you to easily refer to the original values.

  (iii) Identify the row(s) containing the value(s) you want to eliminate. **Manip; Delete Rows...** and specify the row(s) to delete in the box of `Delete rows` and the variable(s) in the `From columns` box, and then click **OK**. Again, do not overwrite the original data set.

- Looking up standard normal distribution table by Minitab
  Note: It may be handy to use Minitab to get the area under the standard normal curve to the left of a given value z. Still, you need to know how to read the standard normal distribution table!
  Calc; Probability Distributions; Normal... will pop up a dialog box. Let's consider the standard normal distribution N(0,1) only here.

  (i) $x \rightarrow$ the proportion (or probability) of $X \leq x$:
  Check `Cumulative probability`, which is indeed default in Minitab. If you have created a column with x values, then check `Input column`, select the column name, and click **OK**. For a single value, check `Input constant` and type x in the box, then click **OK**.

  (ii) the proportion (probability) of $X \leq x \rightarrow x$:
  All the same as (i) except that you select `Inverse cumulative probability` instead of `Cumulative probability`.

- Normal Quantile plot
  Graph; Probability Plot... and select a variable, and then click **OK**. By default, the normal distributions are selected in the box, `Distribution`. Notice that observations are on x-axis and the z-scores are used for scaling y-axis in the plot. The plot also contains other information and further output is printed in the session window, but you may ignore them for a moment. Just check if the plot looks like a straight line and whether it stays inside the band drawn around the line which represents ideal match between the data and the normal distribution assumption.