Answers to some of the problems in chapter 13:

Your text also has answers to some of them in the back.

[13.3, p. 607] *{13.2.3}*

Note that the 4th edition actually gives you a distinct clue not found in the other editions.

In any case, note that there are two continuous variables. The data is not arranged into samples (say, men vs. women, or medicine vs. control). It consists of a series of heights and a series of respiratory volumes (PEF). For each height, there is one respiratory volume.

The analysis that comes to mind is to try to see if PEF is somehow related to height (note that testing if PEF is different from height is silly - they are to completely different things). So if we are trying see if there is a relationship between two continuous variables we want to use either correlation or regression.

Which one? If we're just trying to see if PEF goes up as height goes up (the logical thing to think), then one could use correlation. However, it probably makes more sense here to see what PEF goes with a specific height - i.e., we're trying to predict PEF based on height, ***so regression is probably a bit better.*** It also makes sense that if a person is taller, her respiratory volume should increase, so we're interested in a **one sided alternative** here.

[13.4, p. 608] *{13.2.5, p. 563}*

Here we're looking at colors in snapdragons (a type of flower). The researcher looked at 97 progeny, and found 22 red flowers, 52 pink flowers, and 23 white flowers. Genetic theory says the offspring here should occur in a 1:2:1 ratio (problem says .25,.50,.25, which is the same thing). (Aside: if you know genetics, this is an example of incomplete dominance).

The data that are given are simply counts for each type of flower. There are no measurements here. We are simply interested in how many flowers there are of each color. The "factor" we're investigating is flower color (and only flower color).

Since the data are all in counts, we should be thinking "categorical" analysis. Since there is only one factor (flower color), this is a χ^2 **goodness of fit test**. With three groups, this can not be one sided.

[13.7, p. 608] *{13.2.8, p. 563}*

This one might seem a bit confusing as what the biologists did was count the number of trees, so you might be tempted to think categorical data again.

However, note that what we're really interested in is the number of trees per plot. Since we have 13 samples for each plot, we should be thinking along the lines of "is there a difference in the average number of trees/plot between the three plot types". In other words, the counts are really measurements here.

(If we had just one count of trees per plot type, we might think about a goodness of fit test, but that's obviously not the case here).

So, we have three samples (square, round, and rectangular), each with 13 measurements. Since we have three samples, we should be thinking ANOVA.

(***214 students note:*** *we did not learn how to do this, so in your case, you* ***should*** *be able to figure out that you don't know how to do this analysis*).

However, before you decide to do an **ANOVA**, you really ought to see if the data are approximately normal. If not, you can do a **KW** test here. Since we have three samples, this can't be one sided.

[13.16, p. 609] *{13.2.17, p. 564}*

Here we are interested in how tamoxifen affects microvessel density in 18 patients with cervical cancer. As the book points out, if this density goes down, it's a good thing (less blood flows to the tumor). Each patient was measured twice, once before treatment was given, once after treatment was given.

What we are given here are two sets of continuous measurements. So we have two "samples", one before treatment, one after treatment (it ought to be reasonably obvious that the data here are not categorical - we didn't, for example, count patients with cancer, or patients with a positive response - we just measured their microvessel density.

This is a classic example of a **paired t-test** (assuming the differences are normal). It's also **one sided** because we're interested in microvessel density decreasing after treatment.