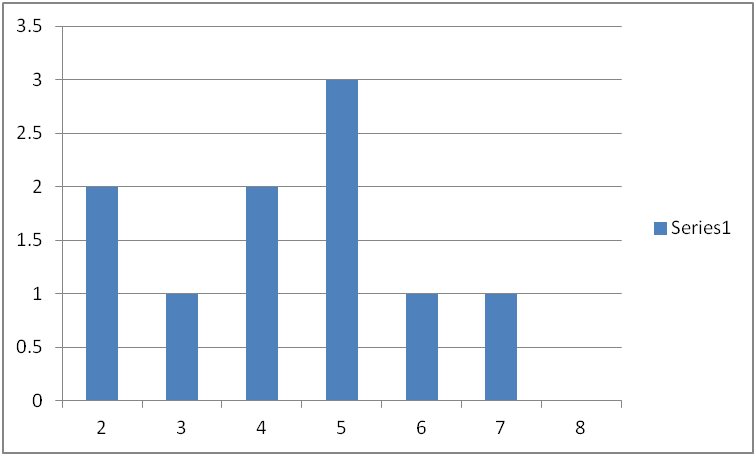
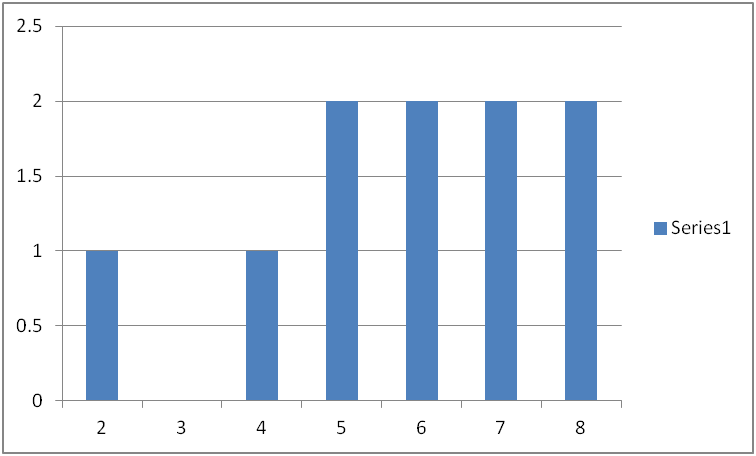
Single Factor ANOVAs

1. Single factor experiments
   1. Science used to be single manipulations of control and experimental groups because it was easy to tell what was going on (binary decisions, no post hocs).
   2. Now experiments are much more complicated, often with multiple levels and independent variables. It’s not any easier, but more economical to test them all in one experiment (also complicated theories require complicated experiments).
   3. Definitions
      1. Factor – the independent variable, the different types of treatments
      2. Levels (treatments/treatment levels) – are the individual groups or things the IV is broken into
   4. Example
      1. You could test 5 different drugs – then compare each drug to each other drug (how many tests would that be?).
      2. That would be tedious. Instead put them all in one experiment.
      3. Instead we ANOVA!
2. Variability and the Sum of Squares
   1. Examples (not going to do the example in the book because you can read…other examples to help you figure it out).
      1. Two conditions: Right handed words, left handed words
      2. Remember all the rules from the previous chapter, random assignment, controlling for nuisance variables, etc.
      3. Here’s the data (attached excel).
   2. Histograms – visual representations of the frequency of scores in the data.
      1. X axis = the event/score (the numbers that could be represented in the data)
      2. Y axis = number of events (how many people rated right handed words at a 5)
      3. Histograms show you the middle, tallest, and variability of a distribution, which is helpful to see if your data is the same for both groups.
         1. HOW TO SPSS HERE





* 1. Distributional statistics
     1. Median – center of a distribution (middle score)
     2. Mean – average of the scores
        1. HOW TO SPSS HERE.
        2. R = 4.3, L = 5.8
     3. So can I say that people prefer left handed words?
        1. This difference is affected by the actual difference between scores and the accidental factors (error, people weirdness) between scores.
        2. So basically is that difference real or due to chance?

1. Hypothesis testing
   1. The main statistics we just reported are *descriptive* because they describe what the data looks like.
   2. We want to do *inferential* statistics to tell what the real difference is.
   3. Definitions:
      1. Samples (group of people tested).
      2. Population (everyone you are interested in studying).
      3. Statistics – summary descriptions from the sample (roman letters)
      4. Parameters – measures calculated from the observations in a population (greek letters)
   4. Statistical Hypotheses
      1. Research hypothesis is a general statement about what you expect to happen in a study (i.e. there will be a difference in scores).
      2. Statistical hypotheses are more precise about the parameters involved in the study (i.e. M1 =/ M2).
      3. Null hypothesis – H0 = the hypothesis we are going to test
         1. H0 = u1 = u2 = u3, etc.
         2. Basically this says no effect.
      4. Alternative hypothesis – the opposite of the null hypothesis
         1. Ha = means are not equal
      5. Why would we test the null and not what we actually want to find?
         1. You are setting yourself up here to either reject the null or retain the null.
         2. You are assuming that most of the time, there are no differences, but this one time when you did something to the groups, you should find a difference.
         3. You want the probability of the differences to be very small (hence p<.05) so that when you reject the null you are saying that the likelihood of these differences is very small, so that H0 can’t be probable under these conditions
   5. Experimental error
      1. Experimental error/error variability/error variance – all the variance due to the fact that people are people
      2. Potential causes of scores
         1. Permanent or stable abilities – how good people are at the experiment
         2. Treatment effect – what you did to them (here right/left) that changes their scores; would be nothing if the null were true
            1. Systematic source contributing to the differences between means.
         3. Internal variability – temporary subject changes, such as mood, tiredness, etc.
         4. External variability – changes from one subject to the next
            1. Measurement error – our measurement tools are not perfect, so it’s the variability in the machine (windows start clock)
      3. Differences between means can be both systematic (treatment) and random (experimental error).
   6. Evaluation of the null
      1. Basically we want to compare the treatment effect to the error and see if you factor out the error that the treatment effect is sufficiently large to reject the null hypothesis.
      2. Null is true = error / error = 1 or treatment / error = < 1
      3. Null not true = treatment + error / error > 1
      4. So we can use this ratio to determine if we should reject the null or not by checking the chance probability of that ratio value under the curve
2. Component deviations
   1. How do we actually measure treatment effects (between groups) and error (within groups) variance?
   2. We are going to add “equal words” to our experiment, so that you can see the calculations with three groups.
   3. Steps:
      1. Figure the means for all three groups: R = 4.3 L = 5.8 E = 7.4
      2. Figure the grand mean: 5.833
      3. DRAW THE PICTURE.
      4. MAKE THE CHART

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Score | Value | Total | Between | Within |
| Each persons score |  | Individual score minus grade mean | Group average minus grand mean | Individual score minus group mean |
| RIGHT 1 | 5 | -.83 | -1.53 | 0.7 |
| RIGHT 2 | 4 | -1.83 | -1.53 | -.3 |
| Etc. |  |  |  |  |

* 1. Between – the deviation of the group means from the middle
  2. Within – the deviation of each person from the middle of that group

1. Sum of squares – defining formulas
   1. Variances – this IS an Analysis of Variance so we have to change these deviations to the more useful variances.
   2. Variance is the average squared deviations from the mean
      1. Why squared?
      2. Variance = sum of squared deviations / degrees of freedom
      3. Sum of squared deviations = SS
      4. DF = the number of ways that deviations are able to vary from each other
         1. Go over more later…always less than the actual Number of people
   3. Deviations = total = between + within, same here SStotal = SSbetween + SSwithin
   4. SS total
      1. Sum of each individual score minus the grand mean squared
   5. SS between
      1. Since this is the group average minus the grand mean squared, we can just multiply by the number of people to get each person’s between group variance (this helps even out unequal groups)
      2. n(Group mean – Grand mean)^2 – add all them up
   6. SS within
      1. Each person’s score minus the mean, squared, add them all up, across all groups (but be sure you are subtracting the right group mean).
2. Computational formulas
   1. Much easier to hand calculate, but aren’t always necessarily clear what you are measuring (these should be clear between versus within).
   2. Some special notation in this book…
      1. Yij = a person’s score, i = subject number, j = group number
      2. I goes from 1 to n (sample size)
      3. J goes from 1 to a (number of groups)
   3. Treatment sums (group totals)
      1. Sum up all the scores in each group
   4. Treatment means (group averages)
      1. Divide each sum by the n (sample size)
      2. Ybarj = average for each group
   5. Grand sum
      1. Sum of all the groups = T in formula
   6. Grand average
      1. Average of all the groups = YbarT = T / an
   7. Bracket terms
      1. [] – steps
         1. Square all the quantities
         2. Sum all the squares
         3. Divide by the number of scores that you were working with
      2. [Y] = sum of all the scores squared
      3. [A] = sum of each group sum squared / n
      4. [T] = T2 / an
   8. Therefore
      1. SST = [Y] – [T] (individual minus grand mean)
      2. SSbetween = [A] – [T] – group minus grand mean
      3. SSwithin = [Y] – [A] – individual minus group mean