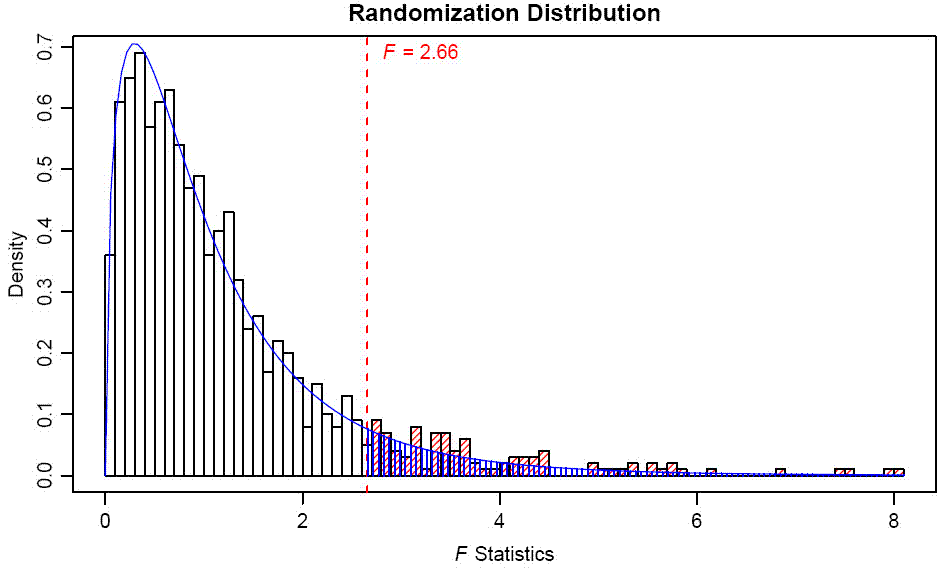
Chapter 3 Variance and the F Ratio

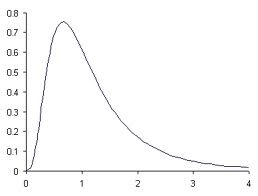
1. Completing the analysis
   1. Summary table – ANOVA calculations

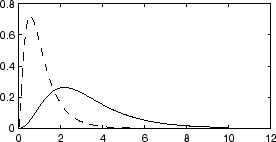
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Source | Term | SS | DF | MS | F |
| A (between) | [A] | [A]-[T] | a – 1 | SSa/dfa | MSa/MSs/a |
| S/A (within) | [Y] | [Y]-[A] | a(n-1) | SSs/a/dfs/a |  |
| Total | [T] | [Y]-[T] | an-1 |  |  |

* + 1. 1st – calculations to get SS
    2. 2nd – SS calculations already discussed in C2
    3. 3rd – df = degrees of freedom – number of independent pieces of information that enter into the SS calculation.
       1. So if you estimate the population mean (u) with the sample mean (M), you have to subtract 1 for that estimation
       2. Number of things free to vary to keep score same
    4. 4th – mean squares – variance associated with treatment effect and error effect
       1. Variance formula demonstrations – SS = top, df = bottom
  1. Confidence intervals for treatment means
     1. Range that should contain the population mean a certain percentage of the time
     2. You set the percent = confidence = (100)(1-alpha)
     3. Formula
        1. Ybar – t(sm) < um <= Ybar + t(sm)
        2. Sm = standard error mean
        3. t from Table a.2 – t-test number
     4. Estimating the Sm – standard error of the mean
        1. Group = SE = SD / sqrt(n) – for individual group, can get this from SPSS
        2. Pooled = SE = sqrt (MSs/a / n(for group))
     5. To get t – use the df (n-1), need alpha
  2. F-ratio – test statistic for ANOVA that allows you to assess the null hypothesis
     1. This value reflects error + treatment on the top, error on the bottom
     2. Will be close or less than 1 when the null hypothesis is true
     3. Much greater than 1 when the null hypothesis is unlikely

1. Evaluating the F-ratio
   1. Null hypothesis significance testing – the comparison of a null result to alternative hypothesis to determine if the null hypothesis is likely or not.
   2. Sampling distribution – when you run an experiment many times and figure the test statistic for each one (such as F here). As you plot the statistics, you will see a known distribution with known probabilities.
   3. Monte Carlos – repeated random draws from a known population.
   4. Show F-distribution here





* 1. Enough of these monte carlos have been run to know the probabilities with different DFs for our studies (or close enough to estimate).
  2. The F-table.
     1. You will need the df numerator (a or between)
     2. Df denominator (s/a or within)
     3. Appendix A.1
     4. Alpha (p-value)
     5. Reporting here
  3. The f-distribution when the null is false
     1. Central F distribution – the normal F distribution based on the true null idea (means are equal)
     2. Noncentral F distribution – the F distribution when means are not equal F’
     3. 
  4. Testing the Null Hypothesis
     1. Null hypothesis states that the means are equal because it is the easiest thing to state clearly
     2. Alternative hypothesis has to be mutually exclusive, says that at least one of the means is not equal (inexact because we haven’t actually said WHAT will happen)
     3. So we evaluate the Null, so we use the F distribution instead of the F’ distribution (because we know what the F distribution looks like)
     4. As a discipline we’ve picked the magic 5% marker (p<.05) where a F score is very unlikely, therefore making the null hypothesis very unlikely and we are going to reject it.
  5. Decision Rules
     1. Reject the null hypothesis when the observed F falls within the incompatible region
     2. We commonly use alpha = .05, generally called the significance level
     3. Critical value – where alpha and the degrees of freedom mark a point for compatible and incompatible hypotheses
     4. You reject or retain the null (not accept/prove/etc.). You support the research hypothesis or not. See how the wording is odd.
     5. Some people use .01 and .10 as well.
     6. However, just because p<.01 does not make the result “better stronger” = this is where effect size becomes important
     7. We tend to look at SIG in SPSS (talk about how you are looking for less than .05).
     8. P values DO NOT equal the probability of getting that effect if you repeat the study.
  6. Comments on NHST (yeah ok it’s not the best)
     1. Plot your data. Visualize it.
     2. Not everything can be tested with NHST.
     3. Significance levels indicate sampling effects – not replication.
     4. What does that significant effect mean?
     5. Effect sizes.
     6. Context of other studies.

1. Errors in hypothesis testing

|  |  |  |
| --- | --- | --- |
|  | Reject Ho | Retain Ho |
| Means equal Ho true | Type 1 error  Alpha | YAY |
| Means unequal Ho false | YAY  Power | Type 2 error  Beta |

* 1. Type 1 – probability of rejecting when you shouldn’t
  2. Type 2 – probability of not rejecting when you should
  3. Alpha – probability IF Ho true
  4. Beta – probability IF Ho false
  5. Power – probability of rejecting null when alternative hypothesis is true
     1. Power = 1-beta
  6. These are on a seesaw – as one increased the other decreases – but you do not add them up (a + b)
  7. DEMO.
     1. Have students create a random variable
     2. Sort by that random variable
     3. Choose the first 15 lines
     4. Run an ANOVA
     5. See how many significant

1. Complete example here (go back to the qwerty example)
   1. Show SS through ANOVA table
   2. SPSS
   3. Graphs with error bars
2. Unequal sample sizes
   1. How often do you get equal groups? (only if you force it right?)
   2. Basically the formulas change where you had n (for sample size for all groups)
      1. [A] = sigma A2 / n
      2. dfs/a = N – a (or n-1 + n-1 + n-1)
      3. Grand mean = T/N instead of averaging treatment means