Paper 192-29

An Animated Guide©: The Logic of Hypothesis Testing and One Way ANOVA

Russell Lavery, Contractor, About-Consulting LLC, Chadds Ford, PA

ABSTRACT

This presentation will show how easy it is to understand statistical reasoning when the examples are familiar to the listener. In the first section of this paper illustrates the logic, and subtleties, of hypothesis testing. The second section of this paper shows the logic, and many subtleties, of ANOVA.

INTRODUCTION

The first section of this paper illustrates the logic, and subtleties, of hypothesis testing through the story of a failed romance (Names have been changed to protect the embarrassed). Topics covered are the logical process of hypothesis testing, Ho, Ha, the burden of proof, alpha error, beta error and the meaning of the P value.

The second section of this paper shows the logic, and many subtleties, of ANOVA via the story of a first grade playground contest. Topics covered are: Variability, the hypothesis of ANOVA, SSE, SST, SSModel, "within group variability", between group variability", MSModel, MSE and R-squared. The logic of the F test is also explained in simple terms.

While not discussed in any depth, the SAS code for ANOVA is very simple, flexible and powerful. An example follows.

```
proc ANOVA data=kickball;
class name;
model kick=name;
means name;
run;
```

IT'S EASY

Often people are surprised to hear that they knew the logic of hypothesis testing and ANOVA before they ever took a statistics class. Sadly, after finishing a statistics course, students often do not recognize their use of the logic of Hypothesis Testing and ANOVA in their everyday thought processes. This paper will display connections between these statistics and common thought processes/experiences. It is thought that familiar stories will make the process more understandable and everyone is familiar with romance.

HYPOTHESIS TESTING:

Anyone who has watched a soap opera, or had a teenage romance should relate to the example below. It is not presented in a serious manner, but does contain all the complexity of a formal hypothesis test. If you know all the answers to the questions, and understand the process, well…the opening paragraph asserted that you did. This is the story of Harry Heart-Throb and his girlfriend-of-record, Sue Sweetheart.

This paper attempts to summarize an animated presentation in a few pages. As a result, concepts and graphics that are on several different slides, in the presentation, must be compressed into one complex figure in the paper.

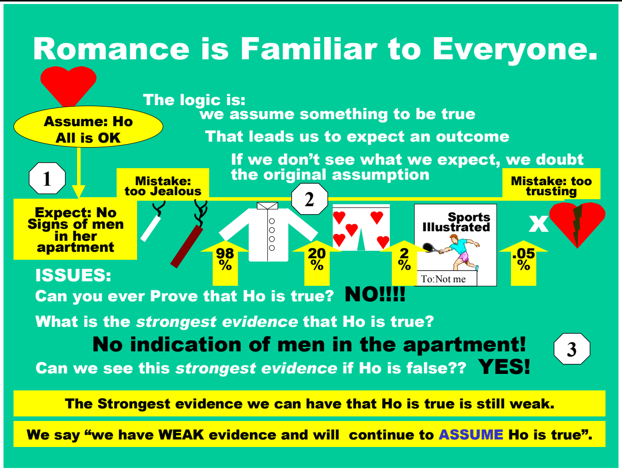

Figure 1

HO AND HA

The starting point for hypothesis testing is stating the two hypotheses, or "describing the two states of nature/reality being tested". The hypotheses are typically called Ho and Ha. The hypothesis testing process is: we assume Ho to be true. That assumption leads us to expect to see certain things when we observe nature/reality. If we do not see what we expect to see, we stop assuming Ho and conclude Ha.

There are only two hypotheses in a hypothesis test and they must be mutually exclusive and exhaustive. Mutually exclusive means that nature/reality can not exist in both Ho and Ha. Exhaustive means that all possible states of nature/reality are described by either Ho or Ha.

In this example (Figure 1) Harry is in a relationship and assumes that all is fine. Ho is that "things are fine". Ha is that Harry has romantic problems. Note that nature/reality (A.K.A. Harry's love life) is in either one or the other situations, not both. Also note that Ho, and Ha, taken together have described all the possible states of Harry's love life (all the states of nature/reality).

Believing Ho to be true leads Harry to make certain predictions about what he will see in nature/reality. Ho leads Harry to expect reality to be [Label (1) in Figure 1] "no signs of men" in Sue's apartment. When Harry observes things in nature/reality (or informally, her apartment), that differ from his expectation Harry, will doubt the truth of Ho. Under the horizontal line [Label (2) in Figure 1] we see a list of incidents that are (left to right) increasingly different from Harry's expectation (under Ho being true).

We use Figure 1 to look at Harry's romantic life over time. Please look to the details of Figure 1 for the issues Poor Harry must handle. Figure 1 shows Harry went over one day and found a cigarette in the ashtray, and Sue does not smoke. This is not what Harry expected, but not *too far* from what he expected. Harry is trusting. He continued to assume Ho and did not start a fight.

On his next visit, he found a cigar in the ashtray and felt less confident in Ho. The following week Harry found a man's shirt (buttons on the right) in the closet…and the shirt was not his. Harry's belief Ho was pretty shaky, but there were possible explanations for the shirt (state of nature). Sue might have a brother who left the shirt, or Sue might be collecting clothes for Goodwill. Harry is very trusting. He did not want to start an argument and then find out that the shirt was going to Goodwill. Harry said nothing.

The next week Harry noticed that there were some neatly folded boxer shorts, with hearts on them, in the dresser. He knew he had not left boxers with hearts on them in Sue's apartment. This was pretty far from what a true Ho would lead Harry to expect. Maybe, Harry hoped, the boxers are also going to Goodwill. Harry is very, very trusting and again said nothing.

The following week Harry noticed that some guy was having a subscription to Sports Illustrated mailed to Sue's apartment. Well, Harry thought, SI can make mistakes. (but, he did think Ho was looking pretty weak) . Finally, there was some event (here represented as an X) that was so far from "no signs of men in Sue's apartment" that even Harry rejected Ho and concluded Ha.

PROOF AND ASSUMPTION

Let's examine this process. We want to consider what we can <u>prove</u> to be true and what we must <u>assume</u> to be true.

Can Harry ever prove that Ho is true? The answer is no. The strongest evidence that Ho is true (no signs of men in the apartment) *can* easily occur when Ho is false. Sue could be "cheating" on Harry and be good at cleaning her apartment. Or Sue could be going to "his" apartment. *Importantly; the strongest evidence that Ho is true, can* ***EASILY*** *occur when Ho is false.* As a general statement, the strongest evidence we could ever have supporting Ho being true is simply a lack of evidence against it. Therefore, the strongest evidence supporting Ho is, when you stop to think about it, is pretty weak evidence. The fact that the best evidence in support of Ho is weak, is the reason why we say we can NOT prove Ho. We are only willing to say that we do not have enough (or any) evidence to reject Ho. The common expression is "We *continue to assume* Ho".

Can Harry ever prove Ho is false? Yes he can! There is a picture (modestly hidden under the X in Figure 1) that is so outrageous that it will "not be seen in nature if Ho is true". We can imagine (and I leave that task to the reader) the existence of fairly strong evidence that Ho is false. Evidence going counter to Ho can be very strong. We can, when we see that strong counter evidence, reject Ho and be pretty certain we are making the correct decision. Evidence against Ho can be strong so we can say "we *conclude* Ha is true" or "we can *prove* Ha to be true".

What does statistics do for us, since we have all known this logic since we were adolescents? Statistics can tell us the *chance* of our being wrong when we make a decision. When Harry rejects Ho and concludes Ha, Harry is going to start an argument. He's going to make accusations and get a bit loud. If he's wrong he's going to have to apologize, buy roses, do laundry and several other things that he doesn't really enjoy. Statistics (and a situation that has numbers and uses statistical formulas) would let Harry reject Ho, and *have an idea of how likely it is that he is wrong*. Figure 1 has several gold arrows. These arrows show (very roughly) how often Harry might be "wrong" if Harry were to see the pictured event and start to make accusations.

TWO TYPES OF MISTAKES

There are two types of mistakes a boyfriend can make. He could be "Too Jealous" and start arguments over little things. This is a Type One Error (A.K.A. an Alpha error). An Alpha Error is rejecting a true Ho, or starting a fight over nothing. Or one could be "Too Trusting" and make a type Two Error (A.K.A. a Beta error). A Type Two Error is continuing to believe a false Ho.

A MORE NUMERIC EXAMPLE MIGHT BE HELPFUL:

Assume Harry is a manager in a food factory (Figure 2) and Harry's department has a machine that fills jars with tomato paste. If the machine is in adjustment, jars should contain 32 ounces of tomato paste. One day Harry had the machine adjusted and immediately weighed 11 jars and plotted the weights with white x's. He then had data on how adjusted machines behave. That plot is shown in Figure 2. We often approximate these plots of x's with a smooth curve (see black line in Figure 2). The height of the curve is an indication of how often that weight occurs.

Harry was now in a position to do some simple hypothesis tests with his data. If the machine were in adjustment (Ho=True) he could expect to see "jar weights" of about 32 ounces (historically 31 to 33 oz.). Imagine, one day Harry weighed a jar at 36 ounces. This is not what he expected, if Ho is true. He reasoned it unlikely that the 36 ounce jar came from the distribution centered at 32oz and more likely came from a "shifted distribution", like one of the distributions in blue. We call them shifted distributions because the averages of these distributions are "shifted" away from the 32 ounce Ho value, and would come from a machine out of adjustment.

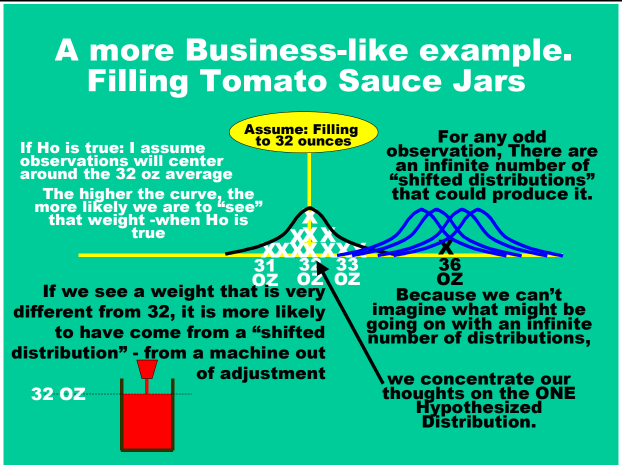

Figure 2

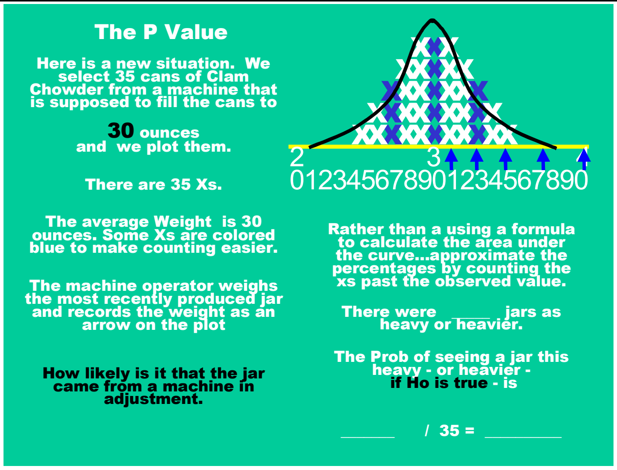

Figure 3

There are an infinite number of shifted distributions that are more likely, than a machine in adjustment, to produce "fill weights" of 36 oz. – not just the four shown. When Harry rejects Ho, and concludes Ha, he is concluding that the 36 ounce jar came from one of the infinite number of blue, shifted distributions and not from the black one. He'd also conclude that the machine needs adjustment.

P-VALUE

There is another concept deserving discussion, the P value. Lets use Figure 3 to investigate the P-value. Figure 3 is cluttered, but it will allow us to work through several examples. The machine can be out of adjustment Heavy, or Light, and so we will concentrate on how far an observation is from the mean in both directions.

Assume that Harry also has a machine that fills Clam Chowder cans to 30 ounces. He weighed 35 jars from a recently adjusted machine and plotted the values. He colored some of the X's Blue to make counting easier. If the machine is in adjustment, he expects to observe fill weights of about 30 ounces. Every half-hour Harry has the machine operator weigh a jar and plot the value as a blue arrow on the plot shown in Figure 3. Figure 3 shows the operator has made five observations and plotted five blue arrows.

A jar weighing 41 ounces (the far right arrow) would be a signal for immediate action. A weight that heavy is so unusual (*if* the machine is in adjustment) that the decision is easy. Harry would immediately send someone to adjust the machine if the sample jar weighed 41 ounces. Harry has recorded 35 observations from a machine in adjustment and NONE of them are close to 41 ounces.

If a jar weighed 33 ounces, the situation would be less clear. When the machine is in adjustment, it produces jars 3 (or more) ounces away from 30 ounces 12 times out of 35. We can figure this percentage by counting the x's that are 3 (or more) ounces from 30. A machine in adjustment produces a jar of 3 ounces "off" from 30 ounces 12/35 (34.2%) of the time. *That 34.2% is the P value for a weight of 33 ounces.* The P value is the chance of "seeing" a deviation that large, or larger, from an adjusted machine.

You can think of the P value as the chance of Harry seeing a jar that far from the mean of 30, when the machine in adjustment. You can also think of it as the chance of sending out the repairman – *and being wrong*. The repairman does not like Harry and makes rude comments when Harry sends him out to adjust a machine that is not really "out of whack". Harry does not want to be ridiculed by the shop crew and wants a small chance of error when he sends out the repairman.

Looking back; the P value for a 41 ounce jar is zero. We can count the x's past 41 but by looking at the picture we KNOW, without statistics, we should send out the repairman a sample jar weighs 41 oz. There are no X's GE 41 (or LE 19) ounces from a machine in adjustment. The probability of getting a 41 oz. jar from a machine that is in adjustment is (close to) zero. Linking our understanding of the picture to the P-value percentage we will make a rule: *"When the P value is small, we can reject Ho". For small P values, we reject Ho and send out the repairman.*

The P value for a 35 ounce jar is 2/35 = .0572. Harry might feel fairly safe from ridicule if a sample jar weighed 35 ounces and Harry sent out the repairman. In that case, Harry has a 6% (about 1 in 15) chance of being wrong. The P value for a 31-ounce fill is .856.

Looking at the picture, we see that a 31-ounce jar is not a rare event when the machine is in adjustment. Just looking at the plot of the data, Harry would not send out the repairman if he observed a 31-ounce jar. Linking our understanding of the picture to the P-value percentage we will make a rule: *If the P value is large, do not reject Ho. If P is large, do not send out the repairman.*

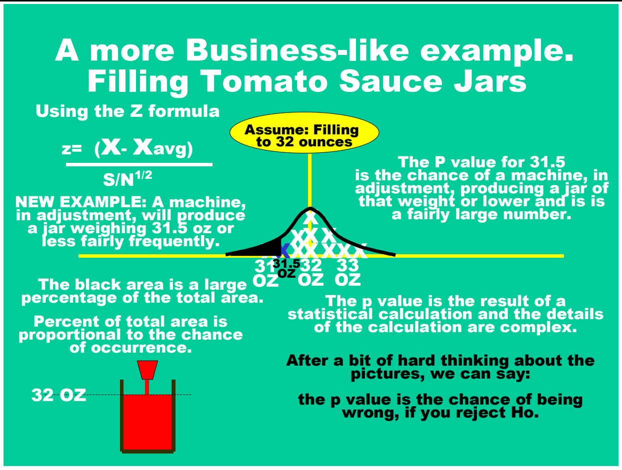

Figure 4

In the real world we would not plot a distribution from a machine in adjustment and take sample sizes of one. There are better procedures than what we have used, but the examples above illustrate the logic of a P value. Counting the number of x's was a simple way of avoiding some formulas that are shown in Figure 4. The formulas in Figure 4 are more precise, and complicated, than the calculations just performed but they have the same logic as the examples above.

Real managers would be using the logic above but formulas like those in Figure 4. Formulas like the one in Figure 4 are taught in statistics courses. Since our goal was the Logic of Hypothesis Testing, not the Statistics of Hypothesis Testing, we will just show the formula and consider the goal met.

ANOVA

ANOVA is another statistical concept that is more easily understood if the example used is familiar. ANOVA is a common statistical technique that has hypothesis testing as a component. ANOVA compares the averages of three, or more, people/process to determine if the averages are likely to be the same, or likely to be different.

Lets look at an incident that happened to the author in First Grade. I challenged a Second Grader named Ajoy and a Third Grader named Beth to a kick-ball contest. Kick ball is played on a field like a baseball field. A pitcher rolls a 12 inch diameter rubber ball to the kicker who kicks the ball and tries to run around four bases. If the ball is caught in the air, or the runner is tagged with the ball, the runner is out. There are three outs to an inning and nine innings to a game. My Kickball field had lines painted on it so that players could see how far they kicked the ball. I challenged Ajoy and Beth to a distance kicking contest. (Figure 5)

Kicking is a random process and we need to digress and explore random processes. A random process has a mean and some variability. While I had an average kicking distance for First Grade, not all my kicks were to my overall average. Some were shorter and some were longer. On a "lucky" days most of the kicks were longer than my overall average. On bad days, most were shorter than my overall average. This is a "good day", "bad day" effect and it happens to everyone.

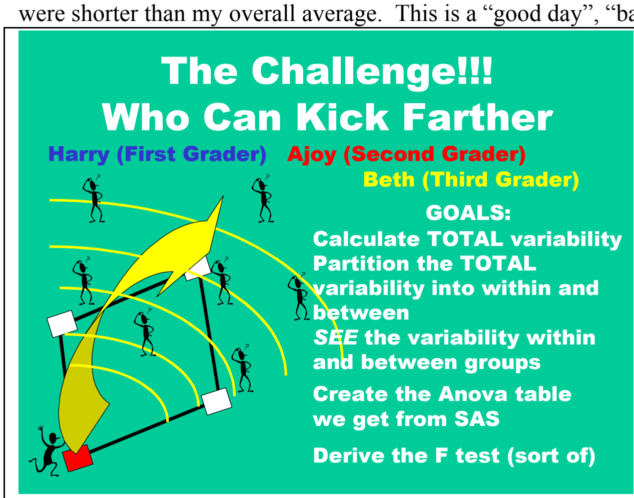

Figure 5

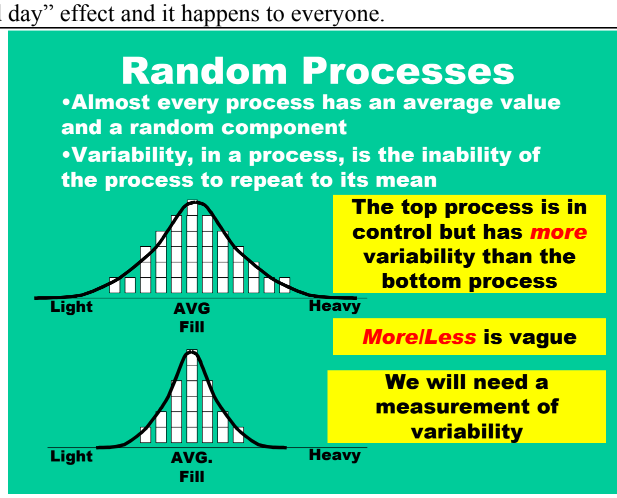

Figure 6

A random process has an average and variability about that average. **Variability, in a process, is the inability of a process to repeat to its mean.**

Processes that are "in control" have predictable distributions without unusual observations. The 36-oz. weight that we saw in Figure 2 came from a process that was "out of control". There was some special cause for the 36 oz. weight – for the excess variability. Adult Harry sent the repairman out to fix the machine when it filled to 41 ounces (Figure 3)- he wanted the repairman to find and eliminate the "special causes" of the excess variability. The distributions in Figure 6 are both in control (they have no unusual observations) but they do have different variability. The bottom distribution has less variability. It is more consistent.

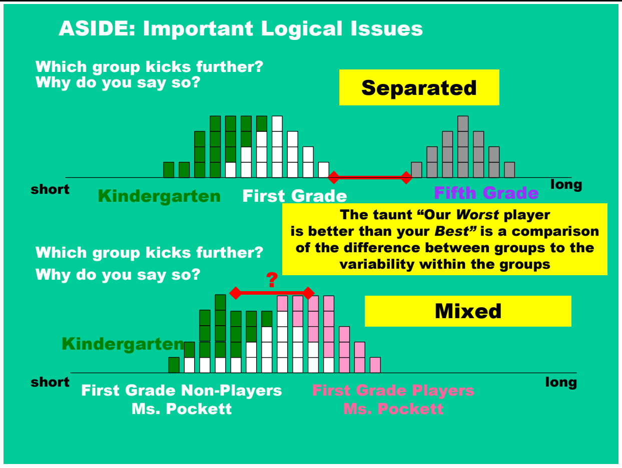

Figure 7

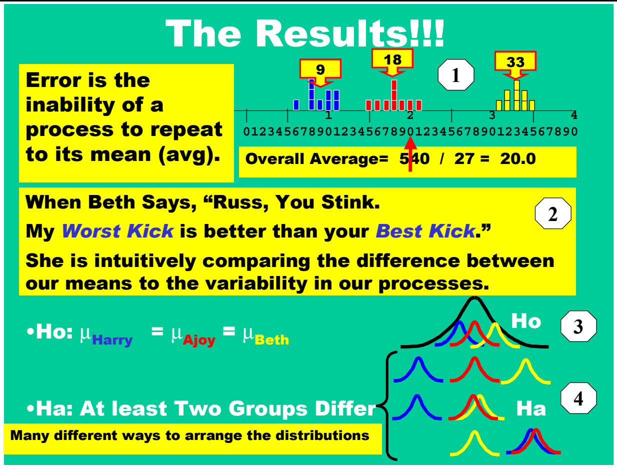

Figure 8

Figure 7 shows some distributions of kicks from my elementary school. In the upper picture, it is pretty plain that the fifth graders kick longer than the kindergartners or first graders. Their taunt, shown in Figure 7, is not only rude but powerful logic. We would all be comfortable predicting that if the contest were repeated tomorrow, fifth graders would again be the winners. Looking at the bottom set of distributions, we would be less willing to predict that the first grade non-players will beat the kindergartners in the next contest.

Figure 8 allows us the first peek at the kicking contest data. The results are plotted in the upper right part, of Figure 8, and I lost pretty badly. My average (blue) was 9 feet. Ajoy (red) averaged 18 feet and Beth (gold) averaged 33 feet. That day Beth yelled out a cruel schoolyard taunt. She said to the author [Label (2)] "Russ, you Stink! My worst kick is better than your best kick".

In one sentence, she showed her nasty temper as well as an innate understanding of statistical logic. She compared the difference between her mean and mine (33-9) to the variability within our kicks (all kicks were within a few feet of our means). Even on one of my lucky days, I would lose. Comparing the difference **between** means" to the "variability **within** contestant distributions" is the basis for ANOVA.

ANOVA is investigated via hypothesis testing and we need to understand the form of the hypothesis used in the ANOVA technique. Ho is that all contestants are all equally good kickers [see the picture labeled (3) in Figure 8]. If Ho is true, the three observed distribution of kicks come from one parent distribution (the black distribution). If Ho is true and all kicks really came from one parent distribution and the differences observed between the three kicking distributions are the result of a "good day"- "bad day" effect. If Ho is true, Ajoy or myself would be just as likely as Beth to win the next contest.

Ha [see the pictures labeled (4) in Figure 8] is that at least two groups differ. There are many situations where Ha could be true and three of them are pictured. All contestants could have different averages. Or I could be bad, while Beth and Ajoy are equally good. Or Ajoy and myself could be equally as good, and both of us better than Beth. The three pictures shown in Figure 8 all satisfy, though in different ways, the requirement that "at least two differ"

Please step back for a second and look at the plot of the kicks in Figure 8. Without using any statistics, do you think the three children have equal averages? Decide now and see, at the end, see if the SAS output supports your opinion.

ADD SOME MATH TO THE LOGIC- TOTAL VARIABILITY

In Figure 9 we calculate the total variability in the data set. The TOTAL variability in the data set is the summed squared distance of each observation from the overall mean of the data. Total Variability measures the players' inability to kick to the *overall average of the data.* The basis for ANOVA (ANalysis Of VAriability or ANalysis Of VAriance) is to take the total variability and split into two parts – "within competitor variability" and "between competitor variability".

We start off by thinking/assuming that "between competitor variability" is caused by a difference in kicking ability (different average distances) and we will return to this to do some subtle thinking later.

We want to do some calculations that will split the total variability into:

1) variability caused by the inability of the contestants to kick the overall average distance (called variability between contestants) and

2) variability caused by the contestants inability to always kick their own individual average distance (called variability within contestants).

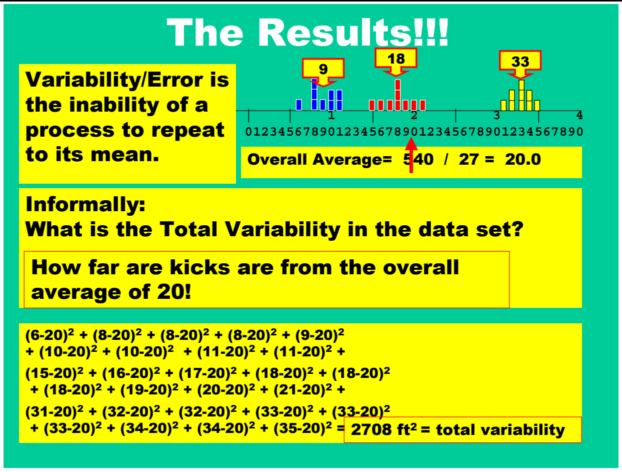

Figure 9

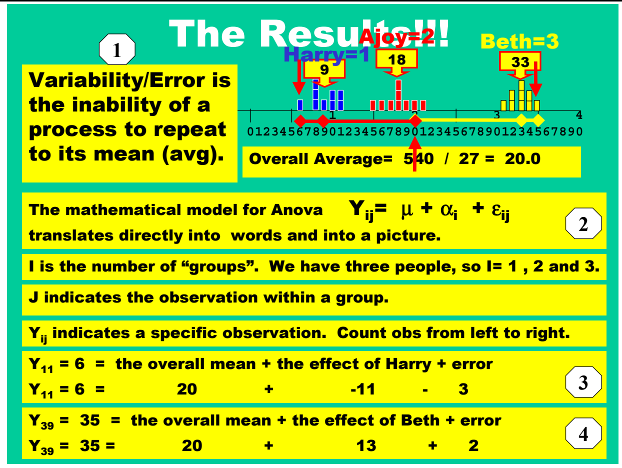

Figure 10

Figure 10 shows the mathematical model for ANOVA [Label (2)], translates the model into a picture (red & yellow lines below the plot) and shows actual numbers [Labels (3) & (4)] associated with the formula. Note that the long red line in Figure 10 is 11 units long and the short red line in Figure 10 is three units long. This is the physical meaning of the numbers in the formula shown in (3).

Figure 10 (1) restates the important concept that error/ variability is the inability of a process to repeat to its mean. The mathematical model for ANOVA (2) builds on that idea. Lets look at a few observations.

Consider my worst kick of six feet and the formula labeled (3). My worst kick is 14 feet below the overall average of 20. We can split that 14 feet into distance caused by my being less skilled *(you know, I really did stink)* than average (20-9 feet from the overall mean) and distance from the overall average caused by this being a poor kick for me (9-6 feet from my mean).

The mathematical model $Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$ can be translated into numbers [Labels (3) (4)] in Figure 9] and into pictures [the red and yellow lines in Figure 9). We can apply the mathematical model for ANOVA to my worst kick and get 6 = 20 -11 –3. The 14 foot difference from the mean (total variability) has been partitioned into between [-11] and within [-3] error.

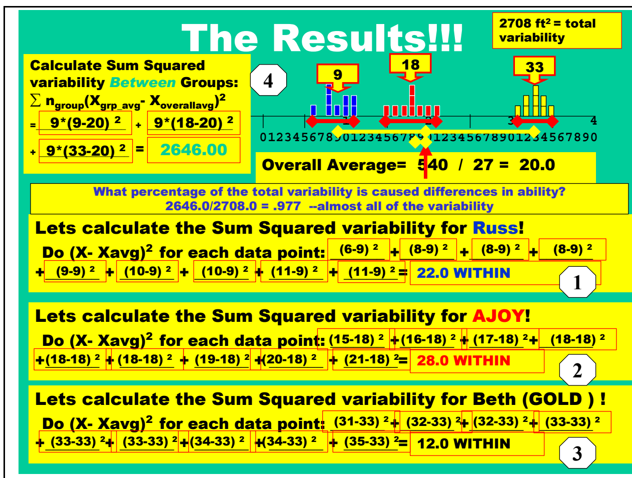

Figure 11

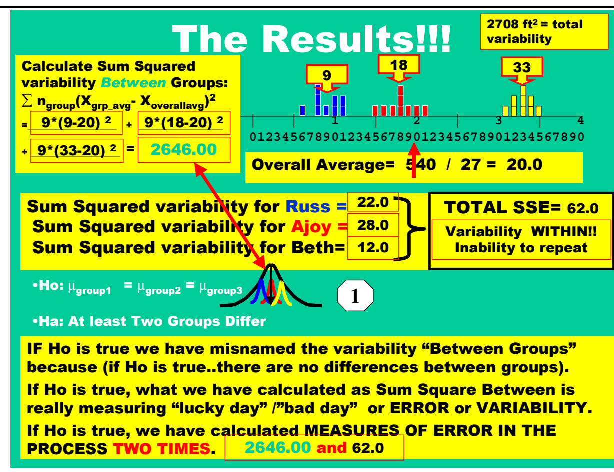

Figure 12

Figure 11 shows the calculations [labels (1) (2) (3)] for variability within the contestants (also called variability within groups). These calculations involve all the observations in the data set. The calculations are similar, in structure, to the calculation for total variability.

The calculation for variability caused by between contestant differences (4) is a bit more complex but has a similar structure. If the calculation in figure 11 does not make sense, just remember that we are splitting total variability into two parts- within and between. We have already calculated total (2708 in Figure 9) variability and within (22+28+12=62 in Figure 11) variability. Since we are splitting total into two parts we could subtract and back into the number by calculating:

total variability – within variability = between group variability
2708 - 62 = 2646

We know how to calculate total variability and we know how to calculate within variability so can subtract to get the between variability. This is what was done in the formula above.

Please note, in Figure 11 [calculations Labeled (1) (2) and (3)], that each observation (unless it is exactly on the mean) makes a contribution to variability. The contribution of an observation to variability is either 0 or positive. So, the more observations a data set has, the larger we expect the total variability to be. This difference in number of observations is a problem with the using the results of the calculations in Figure 11.

We want to compare "between variability" (three observations) with "within observations" (27 observations). The different number of observations makes comparisons of variability difficult. Note that every observation makes (if it not exactly on the mean) a contribution to summed variability. Data sets with more observations will, generally, have more variability.

There is a solution. If you want to compare the two data sets, you often divide the total by the number of observations (N) you made (or by N-1). We often take averages to facilitate comparisons. Average variability per observations, better than total variability, allows comparisons of variability between groups with different numbers of observations.

Figure 12 [Label (1)] shows a picture that illustrates a very important logical condition. If Ho ***is*** true the data we have collected on kick distances all come from one parent distribution – the black distribution in figure 12. If Ho is true, the three students have equal long-term, overall averages. If Ho is true we are all equally good and our long-term, overall averages are the average of the black parent distribution.

Very importantly, if Ho is true, the differences between the three averages that we saw on the day of the contest are just a measure of "good day, bad day" and not measures of "real differences in ability among the children".

The following subtle logic is the key to ANOVA. If Ho is true, variability between groups has been mis-named. It is really a measure of the "good-day, bad-day" effect and the "good-day, bad day" effect is error!!! "good-day, bad day" is the inability of the contestants to kick to their true overall mean. If Ho is true, we have calculated a measure of error **twice** – once when we calculated within group variability and once when we calculated between group variability. (If Ho is true or false, within group variability is a true measure of the error in the process.)

Next come two important pieces of logic and a bit of fuzzy explanation about the number of observations. We are now going to adjust the variability for the fact that the two measures of variability were calculated with different numbers of observations.

We had 27 observations – giving 24 *independent* observations – in the data set that we used to calculate Sum Squared Error **Within** groups. Divide Sum Square Variability Within by 24 to get a measure of average variability per independent observation 62/24 = 2.58. (The full name is Mean Squared Error Caused by the inability of contestants to repeat to their mean – any wonder we call it MSE).

We calculated Sum Square Variability caused by **between** contestant differences by using three observations, giving 2 *independent observations*. We should divide the Sum Square Variability caused by **between** group differences by 2 to get the Mean Squared Variability per independent observation caused by between contestant differences: 2646 /2 = 1323. (Any wonder statisticians call it MSTr or MST - though we might argue for MSC where c is for contestants).

If Ho is true, we have two independent measures of the same thing - error. They are two measures of the inability of the process to repeat to its average. Think a bit; since they are estimating the same thing, the estimates should be equal! Just as a convent way of seeing how equal the estimates are, we typically divide one estimate by the other estimate. If the estimates are truly equal, the result of the division should be about 1.

By convention, we calculate MSTr/ MSE. If Ho is true, the result should equal 1. In Figure 13, we see it is 512. Look at the picture again. Do you think that the three children are all equally good (Ho= $\mu_{Russ} = \mu_{Ajoy} = \mu_{Beth}$ IsTrue) or does the picture suggest that at least two children differ (Ha=True) in their ability?

If Ho is true, we expect MSTr/MSE to be approximately 1. Instead it equals 512.13. That is very far from the 1.00 that Ho led us to expect. According to the Logic of Hypothesis Testing, we doubt that Ho is true because what we observed in nature is very different from what we expected if Ho were true. Ho was that the three children had the same kicking average. We reject that Hypothesis and conclude that at least two children differ.

Is that statistical conclusion in agreement with your interpretation of the picture?

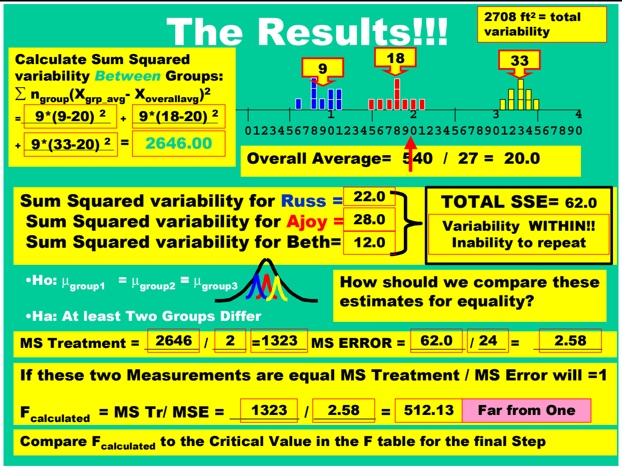

Figure 13

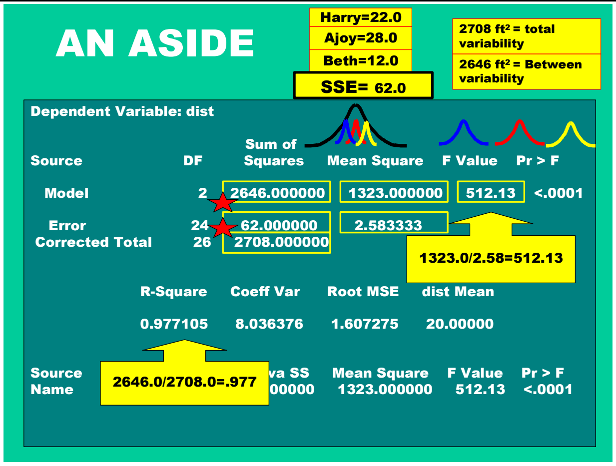

Figure 14

THINKING MORE IN DEPTH ABOUT WITHIN AND BETWEEN VARIABILITY

Now please do a "thought experiment". Imagine the three distributions, in Figure 13, sliding closer to the overall average of 20 without changing their shape. As the three means get closer together, the "total variability" will drop and the "variability between" will drop, but the "variability within" will not decrease. The ratio MSTr/MSE will approach 1 as the means of the distributions "slide" in towards the overall average. If the individual means are all **on** the overall average, MSTr/MSE will equal 1 and SSModel will =0.

To be statistically justified in rejecting Ho, we would need to compare the 512.12 to a number in a statistical table of the F distribution. However, that is the Statistics of ANOVA and this paper is about the Logic of ANOVA. An interested reader can consult almost any statistical textbook for instruction on the use of the F table in ANOVA. This paper hopes to demonstrate why we reject Ho if F is a large number.

A major deliverable of this paper is Figure 13 above. While a complex slide, it is very rich in information. It shows: a picture of the data, Ho and Ha (in symbols and Ho in a picture), the calculations that need to be done for ANOVA and the resulting F test.

Hopefully, this linking of pictures, formulas and F test will take some of the mystery out of ANOVA.

BACK TO SAS

Finally, Figure 14 shows the output from SAS when one uses Proc ANOVA on the kickball data. Note that the numbers that SAS prints are the same numbers that were calculated manually in the paper (and shown in Figure 13). R-squared is the percent of total variability caused by between group variability. When R-squared gets large the picture of the distributions shows separation between the individual distributions. As F gets large, so does R-squared. Since R-squared is SSmodel/SST, it is limited between 0 and 1.

Hopefully, the SAS output is now less intimidating and more understandable.

CONCLUSION

The logical underpinnings of hypothesis testing and ANOVA are easy to understand and we have done them much of our lives. It is hoped that this discussion of the logic behind these techniques will make them less daunting.

ACKNOWLEDGMENTS

Thanks to Dr. Steve Bajgier

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Russell Lavery – Contractor for About-Consulting LLC – Chadds Ford PA
About-Consulting.com
Or Russ.Lavery @verizon.net

```
*********example 1 one way ANOVA**************************;
data kickball;
INFILE DATALINES;
input name  $ kick @@;
DATALINES;
russ   6   ajoy   15    beth   31
russ   8   ajoy   16    beth   32
russ   8   ajoy   17    beth   32
russ   8   ajoy   18    beth   33
russ   9   ajoy   18    beth   33
russ   10  ajoy   18    beth   33
russ   10  ajoy   19    beth   34
russ   11  ajoy   20    beth   34
russ   11  ajoy   21    beth   35
;
run;

OPTIONS LS=100 NOCENTER;
proc chart data=kickball ;
Vbar kick / DISCRETE subgroup=name
            AXIS=0 TO 10 BY 1
            MIDPOINTS=0 TO 40 BY 1;
run;

%LET I_WANT=RUSS;
%PUT _USER_;
proc univariate data=kickball;
where upcase(name)="&I_want";
run;
**NOTE: Cut and paste the above 5 lines and change the VLAUE OF I_WANT;

proc ANOVA data=kickball;
class name;
model kick=name;
means name;
run;
```

```
***********example 2  one way ANOVA*********************************;
data BOOKSREAD;
INFILE DATALINES;
```

```
*a recode of books read over the summer, by local schoolchildren;
input grade $  gender $ books @@;
DATALINES;
first  F 0   second  M 5    third  M 1
first  M 0   second  M 6    third  M 2
first  F 2   second  M 7    third  F 2
first  M 4   second  M 8    third  F 3
first  F 4   second  F 8    third  M 3
first  M 3   second  F 8    third  F 3
first  F 0   second  M 9    third  M 4
first  M 2   second  F 0    third  F 4
first  F 0   second  F 1    third  M 5
;
run;


%MACRO EXMPL2(THE_VAR=);
%PUT _USER_;
proc chart data=BOOKSREAD ;
TITLE "A QUICK PICTURE OF THE VARIABLE &THE_VAR";
Vbar BOOKS / DISCRETE subgroup=&THE_VAR
            AXIS=0 TO 10 BY 1
            MIDPOINTS=0 TO 20 BY 1;
run;

proc ANOVA data=BOOKSREAD;
TITLE 'A QUICK PICTURE OF THE VARIABLE &THE_VAR'; *THIS HAS PROBLEMS;
class &THE_VAR;
model BOOKS=&THE_VAR;
means &THE_VAR;
run;
%MEND EXMPL2;



%EXMPL2(THE_VAR=GRADE);

%EXMPL2(THE_VAR=GENDER);
```