**How to Read Math**

Today’s lab covers basic information that you have almost certainly had before, in statistics or even as far back as high school algebra. However, if it’s been a while since you’ve taken a math course or dealt with equations on a regular basis, chances are that it will take you a while to adjust to reading the equations in the textbook. These equations are key to much of the course material – measurement theory is generally expressed in mathematical terms – so you **cannot** simply skim over them. In general, you will have an easier time with this course if you:

* Allow extra time to read the material. Reading once straight through is not likely to be sufficient.
* “Translate” equations into words as you read, to make sure you understand.
* When the text provides a numerical example, work it out yourself and make sure your answer matches that of the text at every step. This is invaluable for understanding the more complex equations.
* Remember that I will never ask you to memorize equations or to do proofs – many people are anxious about these things. I do, however, expect you to be able to interpret the *meaning* of equations as well as to calculate them correctly.

For more specific guidance, this “cheat sheet” contains some useful tips, based on common mistakes made by previous students. If you find additional tips or reminders that are helpful for you, by all means share them with the class via the Canvas discussion boards.

**Commonly Used Variables**

Some variables tend to have *fairly* consistent accepted meanings. Some are consistent across measurement theory; others are specific to this textbook but are at least consistent within the text. Of course, you will sometimes see exceptions to these usages, so be sure to pay attention to the context.

Within the text:

*x* usually means a participant’s (quantitative) response to a particular item.

*n* usually means the number of participants.

*m* usually means the number of items.

*F*  usually means a participant’s score on a common factor, *not* the F statistic found in ANOVA.

Remember that Latin letters (*a, b,* etc.) usually refer to sample statistics, while Greek letters (α, β, etc.) usually refer to the corresponding population parameters. The difference is that sample statistics describe your particular sample, while parameters are *estimates* of the value you expect for the population as a whole. Properly, parameter estimates are indicated by a “hat”: , but the text often omits the “hat” for simplicity.

Common statistics/parameters:

Statistics Parameters

*p =* sample proportion *π* = population proportion

or *=* sample mean *μ* = population mean

*s =* samplestandard deviation *σ* = population standard deviation

*s2 =* samplevariance *σ2* = population variance

*Note: this notation emphasizes that the standard deviation is the square root of the variance (and vice versa). If you see the quantity s2 or σ2 in an equation and you know the variance, just insert the variance –* ***don’t square it again.***

*sjk =* sample covariance between two variables *σjk* = pop. covariance between two variables

*Sometimes you will see sjj or σjj. Extending the notation, this is the covariance of a variable with itself, which is the same thing as the variance. Note that sjj = (the variance), NOT sj (the standard deviation).*

*r* = sample correlation *ρ* = population correlation

*This is sometimes written as:*

*rxy =* sample correlation between *x* and *y ρxy =* population correlation between *x* and *y*

*As with the covariances, in the context of reliability, you will sometimes see or . This can be thought of as the correlation of the test with itself (e.g., at another point in time or in an alternate form – less than 1.0 due to measurement error).This is called the reliability coefficient.*

Specific to measurement (you will learn more about these as the course progresses):

λ = factor loading

ψ*2* = unique variance – note that this is like the variance, the *2* is part of the variable name. You do not need to square these again.

α = usually indicates the Guttman-Cronbach reliability coefficient alpha, *not* the significance level for hypothesis testing.

ω = McDonald’s omega reliability coefficient

Other common notation:

*P*{ *x* } = probability of *x*

*Z* = standard score – equal to

β = regression weight

**Order of Operations**

This is basic algebra that I imagine most of you are familiar with, but you’d be surprised how many errors occur as a result of performing calculations in the wrong order.

* Remember that anything within parentheses gets done first:
  + Example: *x* = *y*(10 – *z*)
  + Subtract *z* from 10 first, *then* multiply the result by *y*.
  + If *y* = 5 and *z* = 4, then:
    - 10 – 4 = 6
    - 5(6) = 30 = *x*
* If you have one set of parentheses nested inside another one, work from the inside parentheses out.
  + Example: *x* = (*y +* (10 – *z*))2
  + If *y* = 5 and *z* = 4, then:
    - 1 0 – 4 = 6
    - 5 + 6 = 11
    - 112 = 121 = *x*
  + Example 2: *x* = (*y +* (10 – *z*) 2)
    - Note that the *only* difference is the position of the exponent!
  + If *y* = 5 and *z* = 4, then:
    - 1 0 – 4 = 6
    - 62 = 36
    - 5 + 36 = 41 = *x*
* If you have two sets of parentheses that are not nested, *first* calculate the inside of each set of parentheses separately, *then* combine them.
  + Example: *x* = (*y + 1) /* (10 – *z*)
  + If *y* = 5 and *z* = 4, then:
    - 5 + 1 = 6
    - 10 – 4 = 6
    - 6 / 6 = 1

**Subscripts**

Subscripts indicate that there is more than one of a particular type of variable. For example, if you have a 4-item test, each of those items will have its own mean and variance (among other parameters). It would get awfully confusing in a hurry if we used a different letter for each parameter (e.g., item 1 mean = *m*, item 2 mean = *n,* item 1 variance = *o*… get the picture?). So instead, we use a subscript to indicate that is the mean for item 1, is the mean for item 2, and so on. This way we know that all of the values we are talking about are item means, but for different items.

When we are using equations to express theory, we don’t care *which* particular item mean (for example) we are talking about. We just want to indicate that we are talking about *an* item mean, so we use a variable subscript:

This tells us that there are lots of possible s out there, but that our equation applies to any one of them. We use the capital version of this letter to represent the number of possible *j*s (items), *J =* 4 .

Sometimes we will want to use two (or even more) subscripts. For example, we might have a test with *J* items that we have given to *I* participants. In this case, represents the response of participant *i* to item *j*. This means that there are *I x J* potential responses *x.* If we want to talk about two items, the first is often item *j* and the second is item *k*.

When we aggregate across all of the elements in a subscript, that subscript goes away. For example, if we calculated the mean score for item *j* across all of our *I* participants, the resulting value would be . If, on the other hand, we calculated the mean score for participant *i* across all *J* items, the resulting value would be .

**Summation Notation**

Subscripts become particularly important in *summation notation*. Basically, summation notation tells you that you are going to take several values and add them up. Formally, summations are written like this:

In this example, *xj* represents the response of a participant to one particular item, item *j*. We have *m* items, so we might want to add up this participant’s responses to each of the *m* items. So we would start with item 1 (*j* = 1), add the response to item 2 (*j =* 2), and keep going until we hit item *m.* The text sometimes simplifies this notation to:

It means exactly the same thing. Since you only have one subscript, *j*, you know that you should add up all of the *xj* you have across all of the values of *j*. Sometimes, however, you have more than one subscript, and more than one summation sign:

In this case, we might have responses *xij* to *n* items from *m* participants, and we might want to add ALL of them. In other words, we would add up all of participant 1’s responses across all *j* items, then add those to all of participant 2’s responses, and so on until you reach participant *i*'s response to item *j*. Note that this means you are adding *i x j* numbers.

Often there are calculations within the summation sign, like this:

This is where order of operations becomes *extremely* important! In this example, you will:

* Take each value of *xj* and square it
* Add one to each of the squared values you just obtained
* *Then* add up all of the values you obtained in the previous step.

This is quite different from:

In this example, you would take each value of *xj* and square it, then add up all the squared values, and only *then* add 1.

You will, of course, come across a few other variables and mathematical operations over the course of the semester, but these should be enough to get you started. If you run across something that confuses you, carefully doublecheck the context (the text does usually explain new variables), then ask me or a classmate, or Google.

**Lab Exercise:**

This exercise is to be done **in lab**. Usually, you will turn in lab exercises in electronic form, but you may do this one in hard copy if you prefer. You are welcome to discuss the items with one another, but each person needs to turn in one worksheet.

1. Work the following problems, showing your work:

(8-2)^2 / 4- 1

6^2 / 3

36/3

12

* 1. where *x* = 3

[1/3^2)(5-3)] / (3+1)

(1/9)(2) / 4

0.111(2) /4

0.22222 /4

0.0555556

(3\*1 +2) + (3\*2 +2) + (3\*3 +2) + (3\*4 +2) + (3\*5 +2)

5 + 8 + 11 + 14 + 17

55

Use the data in the following table to calculate the next two problems:

|  |  |
| --- | --- |
| Item | Participant’s response: |
| 1 | 4 |
| 2 | 5 |
| 3 | 4 |
| 4 | 3 |
| 5 | 5 |

(4^2) + (5^2) + (4^2) + (3^2) + (5^2)

16 + 25 + 16 + 9 +25

91

(4 + 5 + 4 + 3 + 5)^2

21^2

441

1. Translate the following equations into *words* (NO mathematical symbols). Tell me what you should do *step by step*, in order. Be detailed and specific.

Step 1: Identify the values that need to be plugged in.

Step 1a: is the individual response to an item. Each participant will have a different value here. This will change multiple times.

Step1b: is the mean of all responses to an item.

Step1c: n is the number of particpants in the sample

Step 2: plug in a particpants value for xji and subtract it from the mean of the items.

Step 3: Square the value obtained in step 2. Keep this value, we will refere back to it.

Step 4: subtract 1 from the sample size.

Step 5: divide the result in step 3 by the result in step 4.

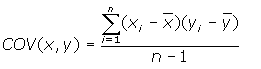
Step 6: Retain this value.

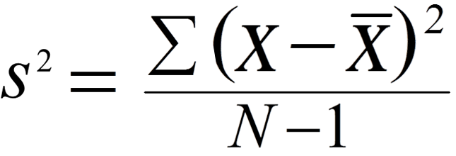
Step 7: Repeat stesps 1-6 until every item response has been inputted into the equation.

Step 8: Once step 7 is complete, sum all of the values obtained in step 6.

Step 9: Observe answer and ensure it makes sense with the data given.

Is this equivalent to formula 2.2 in the R&M text? Explain why or why not.

Covariance Equation:

Variance Equation: 

where *p*  is the number of items, and *Xi* and *Xj* are items.

(R & M Equation 6.9)

No, these are not equivalent equarions. To start, I have listed the Variance and covariance equations above. When You plug these in, You find that you can immediately you can cancel out the “n-1” and

Making the equation

Then We can combine the equivalent summation operators. This is the most I can get the equation reduced down to. Because there is the incorporation of the covariance equation ip top, we can not properly eliminate the divisible portion on the bottom. Therefore, they are not equivalent.