Item and Test Statistics Algebra of Expectations

Measurement Methods

Lecture 5

2/9/06

Today's Class

- Recap of covariance/correlation matrices.
- Statistics for quantitatively scored items.
- Test score statistics.
- More "expectations."

Carrying on From Last Class…

- If one has multiple items, a matrix of correlations or covariances can be constructed.
- This matrix is symmetric – everything in lower triangle is repeated exactly in the upper triangle:

TABLE 3.6
Binary Covariance and Correlation Matrices

(a) Covariance Matrix

	1	2	3	4	5
1	.25	.10	.15	.10	−.05
2	.10	.24	.00	.08	−.06
3	.15	.00	.25	.10	−.05
4	.10	.08	.10	.16	−.02
5	−.05	−.06	−.05	−.02	.09

(b) Correlation Matrix

	1	2	3	4	5
1	1	.408	.600	.500	−.333
2	.408	1	.000	.408	−.408
3	.600	.000	1	.500	−.333
4	.500	.408	.500	1	−.167
5	−.333	−.408	−.333	−.167	1

Correlation Matrix Inspection

- As we progress through this class, we will come to understand that a lot of information can be culled from a correlation matrix.
- Often times, groups of items are highly intercorrelated within, but relatively uncorrelated between.
- These ideas will serve as a base to answer one of our fundamental questions about a test:
  - Does one score suffice or should two scores be used?

Expected Values

- Imagine you have a random variable X.
- The expected value of X is then the mean of the variable:
  - $E(X) = \mu$
- The variance is also phrased as an expected value:
  - $\mathrm{Var}(X) = E[\,(X-\mu)^2\,]$

Covariance as an Expected Value

- As you could probably guess, we can phrase the covariance as an expected value, too:
  - $\mathrm{Cov}(X,Y) = E[\ (X\text{-}\mu_X)\ (Y\text{-}\mu_Y)\ ]$

- To demonstrate, consider the following example.
  - We have two binary items
  - We go out and find proportion of responses to each.
  - We put our responses in a 2 x 2 contingency table.

Counts of Responses to Our Two Items

X_k	0	1	Sum
0	35	5	40
1	35	25	60
Sum	70	30	100

$\mu_{Xj} = \pi_{Xj} = 0.60$ $\mu_{Xk} = \pi_{Xk} = 0.30$

Constructing Covariances from Expectations

	X_j			
X_k		0	1	Sum
	0	30	10	40
	1	20	40	60
Sum		50	50	100

Xj	Xk	n	P(Xj,Xk)	Xj - E(Xj)	Xk - E(Xk)	[Xj - E(Xj)][Xk - E(Xk)]	P(Xj,Xk)*[Xj - E(Xj)][Xk - E(Xk)]
0	0	30	0.3	-0.6	-0.3	0.18	0.054
0	1	20	0.2	-0.6	0.7	-0.42	-0.084
1	0	10	0.1	0.4	-0.3	-0.12	-0.012
1	1	40	0.4	0.4	0.7	0.28	0.112
						Covariance	0.07

Algebra of Expected Values

Algebra of Expected Values

- Expected values really don't do us a whole lot of good by themselves.

- Really, for our purposes, we must consider what the expected values are for functions of variables.

- For instance, imagine if you have a constant, c, what would the expected value of $X + c$ be?

Rules of Expected Values

- Expected value of a constant:
  - $E(c) = c$

- Expected value of a sum of a constant and a random variable:
  - $E(X + c) = \mu + c$
  - This is the same as saying the mean of variable that has added

More Rules

- The variance of a sum of a constant and a random variable:
  - $\mathrm{Var}(X + c) = \sigma^2$
  - Adding a constant to a variable does not change the variance of the variable.

More Rules

- Multiplication of a random variable by a constant:
  - $E(cX) = c\mu$
  - $Var(cX) = c^2\ Var(X) = c^2\sigma^2$

- Consider a linear combination:
  - $Y = a + bX$
  - $E(Y) =$
  - $Var(Y) =$

Even More Rules

- The expected value of a sum of random variables is the sum of their expected values:
  - $E(X+Y) = E(X) + E(Y)$
- The variance of a sum of random variables is given by:
  - $Var(X+Y) = Var(X) + Var(Y) + Cov(X,Y)$

What About Covariances?

- All of the things previous to this were geared to describe what happens to the covariance…
- Imagine you have two variables X_j and X_k.
- Imagine further that you want to create two new variables:
  - $Y_j = a + b\,X_j$
  - $Y_k = c + d\,X_k$
- What is $\mathrm{Cov}(Y_j, Y_k)$?

Covariances

- Cov(Y_j, Y_k) =

Cov($a + b\,X_j$, $c + d\,X_k$) =

E[($a + b\,X_j - a + b\,\mu_{Xj}$) ($c + d\,X_k - c + d\mu_{Xk}$)] =

E[b ($X_j - \mu_{Xj}$) d ($X_k - \mu_{Xk}$)] =

$b\,d$ E[($X_j - \mu_{Xj}$) ($X_k - \mu_{Xk}$)] =

$b\,d$ Cov(X_j, X_k)

- We will come to appreciate how this will function for our models.

Quantitative Item Scores

Quantitative Item Scores

- Recall from our first lecture the item:
  - Assault weapons do not belong in private hands.

- Imagine you are interested in assessing attitudes toward gun control.

- You could collect data on this type of item using a "Likert" scale.

Typical “Likert” Scale Responses

- Imagine you use the following scale to assess a person’s opinion:
  - Strongly agree
  - Agree
  - Neither agree nor disagree
  - Disagree
  - Strongly Disagree

Assigning Integers to Responses

- Likert (1932) showed that complicated scaling techniques provided little improvement over the assignment of numbers to the responses.

- So, to make your life simple, you then code each response with an integer:
  - 5 for strongly agree.
  - 4 for agree.
  - 3 for neither agree nor disagree.
  - 2 for disagree.
  - 1 for strongly disagree.

Now What?

- Recall from our last class the random sampling process we did to ask a single question to a bunch of people.

- We could do this same process, just with our new question.

Our Sample

TABLE 3.7
Constructed Quantitative Variable

13433	24333	333213	22544
33533	33223	13243	43331
43315	33234	12322	32521
31222	43232	24433	44412
23453	35243	33342	22315
. . .			

- Above is a sample of 100 responses to the gun control attitude item.
- These were generated by a random process where each response option had a certain probability (called π) of being selected.

Our Sample, Revisited

TABLE 3.8
Distribution of Quantitative Variable

X	n_c	p	π	πx	πx^2
5	7	.07	.1	0.5	2.5
4	17	.17	.2	0.8	3.2
3	41	.41	.4	1.2	3.6
2	25	.25	.2	0.4	0.8
1	10	.10	.1	0.1	0.1
				$\mu = \Sigma\pi x = 3.0$	$\Sigma\pi x^2 = 10.2$

$$\sigma^2 = 10.2 - 3.0^2 = 1.2$$

- Here, we find that using the population values, we can determine what we expect our mean and variance to be.
- Our sample values for the mean and variance will be slightly different because we have a finite sample.
- The law of large numbers still applies, however.
  - The sample mean will converge to the population mean.
  - The sample variance will converge to the population variance.

Item Difficulty

- As with binary items, we regard the difficulty of an item to be indicated by its mean.
- Again, we hesitate to call non-cognitive items “difficult” or “easy”

Variance of Integer-valued Items

- One property of Likert scale or integer valued items is that the maximum variance is known.
- For an item with *k* responses, the maximum variance is:

$$\sigma^2_{\mathrm{max}} = \frac{k-1}{2}$$

Variance of Integer-valued Items

- Additionally, the variance of these items is *not* determined by the mean.
  – But the variance and mean are related.

- Variance is a measure of the *diversity* of the response.
  – In our item, variance could be regarded as diversity of opinion.

Satisfaction With Life Scale (SWLS) Example

- As an example, consider our familiar SWLS items:
  1. I am satisfied with my life.
  2. The conditions of my life are excellent.
  3. In most ways my life is close to the ideal.
  4. So far I have gotten the important things I want from life.
  5. If I could live my life over, I would change almost nothing.

SWLS Response Options

- The response options for the SWLS were integer-valued:

7 - Strongly agree

6 - Agree

5 - Slightly agree

4 - Neither agree nor disagree

3 - Slightly disagree

2 - Disagree

1 - Strongly disagree

SWLS Sample

TABLE 3.10
Satisfaction With Life Data

	Items						
Respondent	*1*	*2*	*3*	*4*	*5*	*Total*	*Mean*
1	2	4	3	5	2	16	3.2
2	5	7	7	7	6	32	6.4
3	3	5	5	4	1	18	3.6
4	6	6	6	6	5	29	5.8
5	7	7	6	2	2	24	4.8
6	5	2	6	7	2	22	4.4
7	2	3	3	3	1	12	2.4
8	4	3	6	3	3	19	3.8
9	3	5	5	5	1	19	3.8
10	4	4	5	6	4	23	4.6
Total	41	46	52	48	27		
Mean	4.1	4.6	5.2	4.8	2.7		

SWLS Covariance and Correlation

TABLE 3.11
Satisfaction With Life Covariance/Correlations

(a) Covariance Matrix

	1	2	3	4	5
1	2.78	1.49	1.76	0.24	1.48
2	1.49	2.93	0.98	−0.09	1.31
3	1.76	0.98	1.73	0.71	1.40
4	0.24	−0.09	0.71	3.07	1.71
5	1.48	1.31	1.40	1.71	3.12

(b) Correlation Matrix

	1	2	3	4	5
1	1	.53	.80	.08	.50
2	.52	1	.43	−.03	.43
3	.80	.43	1	.31	.60
4	.08	−.03	.31	1	.55
5	.50	.43	.60	.55	1

Test Scores

Test Scores

- We will consider (until otherwise noted) that a test score will be a simple sum of the scores of all the items of the test.
  - Later we will discuss different weighting of each item.
- The total test score for examinee *i*, then (for *m* items) is:

$$y_i = \sum_{j=1}^{m} x_{ji}$$

Other Versions of Test Scores

- The **relative test score** or **mean test score** is given by:

$$m_i = \left(\sum_{j=1}^{m} x_{ji} \right) / m$$

- For binary items:
  - The total test score is the number right.
  - The relative test score is the proportion right.

Total Test Score Mean

- The mean of the total test score can be computed by either:
  - Summing each examinee's total test score and dividing by the sample size.
  - Summing the mean of each item.
- Summing the mean of each item comes from the algebra of expectations:
  - $Y = X_1 + X_2 + \ldots + X_m$
  - $E(Y) = E(X_1 + X_2 + \ldots + X_m) =$

    $E(X_1) + E(X_2) + \ldots + E(X_m)$

Total Test Score Variance

- The variance of the total test score can be computed by either:
  - Taking the variance of each examinee's total test score.
  - Summing all of the elements of the covariance matrix of the items.

- For some purposes, we make a distinction of the elements of the covariance matrix:
  - We sum the diagonal values – call them ***D***
  - We sum the off-diagonal values (all of them) – call them ***O***

Total Test Score Variance

- From our SWLS Example:
  - $D = 13.63$
  - $O = 21.98$
- Then the total test score is:
  - $s_y^2 = 35.6$

TABLE 3.11
Satisfaction With Life Covariance/Correlations

(a) Covariance Matrix

	1	2	3	4	5
1	2.78	1.49	1.76	0.24	1.48
2	1.49	2.93	0.98	−0.09	1.31
3	1.76	0.98	1.73	0.71	1.40
4	0.24	−0.09	0.71	3.07	1.71
5	1.48	1.31	1.40	1.71	3.12

(b) Correlation Matrix

	1	2	3	4	5
1	1	.53	.80	.08	.50
2	.52	1	.43	−.03	.43
3	.80	.43	1	.31	.60
4	.08	−.03	.31	1	.55
5	.50	.43	.60	.55	1

Covariance Considerations

- Imagine we have a large pool of potential items.
  - We want to create a test of a smaller subset of the items.
- It can be argued that we should choose the items that give the largest ratio of test variance to the sum of item variances:

$$s_y^2 / \left(\sum_{j=1}^{m} s_j^2 \right) = (D + O_d)/D$$

Covariance Considerations

- To maximize the ratio on the previous slide, it makes sense to choose items with large positive covariances.
  - A test whose items have large positive *correlations* is called **internally consistent**.
    - We will revisit this topic later in the class.
- What we are doing, in effect, is making the test longer…
  - Longer tests are generally better measures of a trait.

Covariance Considerations

- Additionally, adding items with high covariances typically means we are adding items that measure the same content area.
- For binary items, this means that we may end up choosing items with variances near 0.25…
  - As we will see, we will end up with a test that measures the middle of a scale well.

Next Time

- Chapter 4 – the concept of a scale.

- Discussion about class projects.