Lecture 25: Blocking

Dewayne E Perry
ENS 623
perry@ece.utexas.edu

Nature of Blocking

- Remember one of the ways to increase power:
  - Increase the size of the effect
- One way to increase effect size
  - Decrease the size of the within-group or error variation
- *Blocking* does this - increases precision
  - Stratifying or subdividing of subjects/samples
  - In such a way that those within a common block are more similar to each other on the dependent variable than they are to subjects/samples in another block/group

Nature of Blocking

- Example
  - Block according to anxiety level in study of new type of treatment
  - Measure anxiety level prior to treatment
  - Take top scores, randomly assign to T and C, iteratively
  - Anxiety

Anxiety	hst	h	m	l	lst	sum
T	8	6	3	1	1	19
C	9	7	5	3	1	25
sum	17	13	8	4	2	44

  - Mean: T=3.8, C=5.0, GM=4.4
  - Summary of sources

	SS	*df*	*MS*	*F*	*p*	*eta*
T	3.60	1	3.60	10.29	.04	.85
AB	77.40	4	19.35	55.29	.002	.99
R	1.40	4	0.35			

  - Comparison
    - Treatment effect is large and significant at $p < .05$
    - blocking variable effect even more so

Nature of Blocking

- Omitting blocking

	SS	*df*	*MS*	*F*	*p*	*eta*
T	3.60	1	3.60	0.37	.56	.21
R	78.80	4	9.85			

- Summary
  - Effect of treatment not very significant
  - But same mean squares
  - Residual variance decomposed into a large between blocks component compared with the small one in the blocked
  - Removes the large sources of variation known to be associated with the systematic pre-experimental differences among subjects

Benefits of Blocking

- Consider the size of the sample needed to achieve the same *F* ratio for blocked and unblocked analyses

$$reps = \frac{MS_{error_{unblocked}}(\text{no. of blocks})}{MS_{error_{blocked}}}$$

$$reps = (9.85 \times 5)/0.35 = 140.7$$

  - Would need 140 pairs in the unblocked experiment to reach the same *F* we have in our blocked with 5 pairs
  - But a difference in *df* and *p*: 4 and .04 vs 279.4 and .002
- To achieve the same significance level, would need 60.3 subjects/samples - a ratio of about 12 to 1 against blocked
- Example designed to show dramatic effects of blocking
- the larger the correlation between the blocking variable and the dependent variable
  - the greater the benefits
  - the greater the precision

Blocking and Covariance

- Analysis of covariance - a special case of ANOVA
  - observed scores adjusted for individual differences within conditions of
    - some predictor variable, or
    - some covariate known to correlate with the dependent variable
  - typical covariate is the pretest administration of the same (or similar) test that is to be employed as the dependent variable
- Increasing precision
  - Might have used ANOCVA instead of blocking
    - Sometimes better at increasing precision
    - Especially when pre and post test scores are highly correlated
    - Special case of ANOVA when perfectly correlated
  - Useful rules of thumb (Cox 57)
    - Blocking better when correlation is .6 or less
    - ANOCVA better when correlation .8 or more
    - neither clearly better when correlation is between .6 and .8

Blocking and Covariance

- Blocking equally efficient for both linear and curvilinear; ANCOVA only when linear
- Blocking also useful when the blocks differ in qualitative rather than quantitative ways
- Blocking always imposes some cost in terms of loss of *df* for error.
  - Cost usually small in relation to decreased MS error
  - if little reduction in MS error, can always unblock and recapture the lost *df*

Blocking and Covariance

- Detecting interactions
  - another benefit besides increasing precision
    - detection of interactions between experimental and blocking variable
    - usually in designs where each block has a number of replications for each treatment condition
  - Example (Treatment, Control, Mean, Residual)
    - above 60: TM=6, CM=7 TR=.33, CR= -.33
    - 40-59: TM=3, CM=6 TR= -.67, CR=.67
    - below 40: TM=6, CM=7 TR=.33, CR= -.33
    - Middle age tend to benefit more than younger or older

Note on r^2 or eta^2

- Note that in these analyses we have disregarded all other sources of variance except the treatment *SS* and the error *SS*

$$eta = \sqrt{\frac{SS_{treatment}}{SS_{treatment} + SS_{error}}}$$

  - generally more useful than

$$eta = \sqrt{\frac{SS_{treatment}}{SS_{total}}}$$

  - reflects growing ability to predict behavior
    - by decreasing the error term
      - ✓ by learning to specify other important sources of variance
      - ✓ and subtract them from the error term by means of blocking

Blocking and Repeated Measures

- Remember matched pairs *t* test?
  - Example of blocking: each pair of observations is a block
  - simplest form of repeated measures design
- All repeated measures designs are examples of blocked designs
  - the more positively correlated the successive observations on the same sampling units, the more we benefit from increased precision
    - versus between subjects design
    - typically get greatest precision when block on sampling unit itself

Blocking within Blocks

- Example: determining accuracy of decoding non-verbal clues from face, body, voice
  - Repeated measures design
  - 30 students, 60 clips, 20 of each
    - might have face items over-represented in last half, eg
  - Alternative: divide into 20 blocks of 3 each
    - randomly present one of each
    - doesn't necessarily increase precision
    - does eliminate possible confounding effects of the order of presentation
    - does allow us to learn from blocks x channels interaction
      - ✓ extent to which differences change over time

Blocking within Blocks

- ANOVA of 30 x 60 - *df* for each source
  - decoders=29, channels=2, channels-decoders=58, items=1710, total=1799
- ANOVA of 30 x 20 x 3
  - decoders=29, channels=2, channels-decoders=58, blocks=19, blocks-decoders=551, blocks-channels=38, blocks-channels-decoders=1102, total=1799