Mixed models for the analysis of repeated measurements[1]

Jerry Brunner, Department of Statistical Sciences

[1]This slide show is an open-source document. See last slide for copyright information.

The Role of Statistics in the Discourse of Science

- The Hidden Advisor study (Garfinkel, 1967).
- Skeptic in a box.

The skeptic says

- Maybe this is coincidence.
- Under the following reasonable circumstances, what are the chances of getting results like these (or even stronger) if the treatment actually had no effect?

For example

- You claim that it takes longer for Spanish-speaking children to learn English.
- You have data in which the mean length of time required for a set of Russian-speaking children to reach a certain level of proficiency was shorter than the mean for a set of Spanish-speaking children.
- Well, what if these were independent random samples from normal populations with exactly the same means, but possibly different variances?
- What are the chances of observing a difference as large as or larger than the one you observed?
- If the probability is small enough, perhaps coincidence can be ruled out.
- And we can discuss this.

They always say

- It's good to think about the statistical analysis as you are designing a study.
- They're right.
- But it can be taken too far.

Statistics should not be a Procrustean Bed

Mixed models for the analysis of repeated measurements

- First, repeated measurements.
- Are we really shorter in the evening?
- Measure a set of people in the morning and the evening.
- Well, what if these measurements were independent random samples from normal populations with exactly the same means ...
- Wait, hold it!

The main benefits of repeated measurement analysis

From a statistical point of view

- Increased statistical power for a given sample size.
- Each participant serves as her own control.
- Person-to-person variation cancels out.
- Signal is more detectable against a lower level of background noise.

Random effects vs. fixed effects

A random factor is one in which the values of the factor are a random sample from a populations of values.

- Randomly select 20 fast food outlets, survey customers in each about quality of the fries. Outlet is a random effects factor with 20 values. Amount of salt could be a fixed effects factor.
- Randomly select 10 schools, test students at each school. School is a random effects factor with 10 values.
- Randomly select 15 naturopathic medicines for arthritis (there are quite a few), and then randomly assign arthritis patients to try them. Drug is a random effects factor.
- Randomly select 15 lakes. In each lake, measure how clear the water is at 20 randomly chosen points. Lake is a random effects factor.

Mixed models

Mixed models have both fixed and random effects.

Random shocks

A nice simple example

- Randomly select 5 farms.
- Randomly select 10 cows from each farm.
- Randomly select 5 to have free access to pasture, and 5 to be kept in the barn.
- Record the amount of milk from each cow.
- Farm is a random factor, access to pasture is fixed.

The idea is that "Farm" is a kind of random shock that pushes all the amounts of milk in a particular farm up or down by the same amount. It's considered random because the farms were randomly selected.

Classical F-tests

For mixed models and pure random effects models

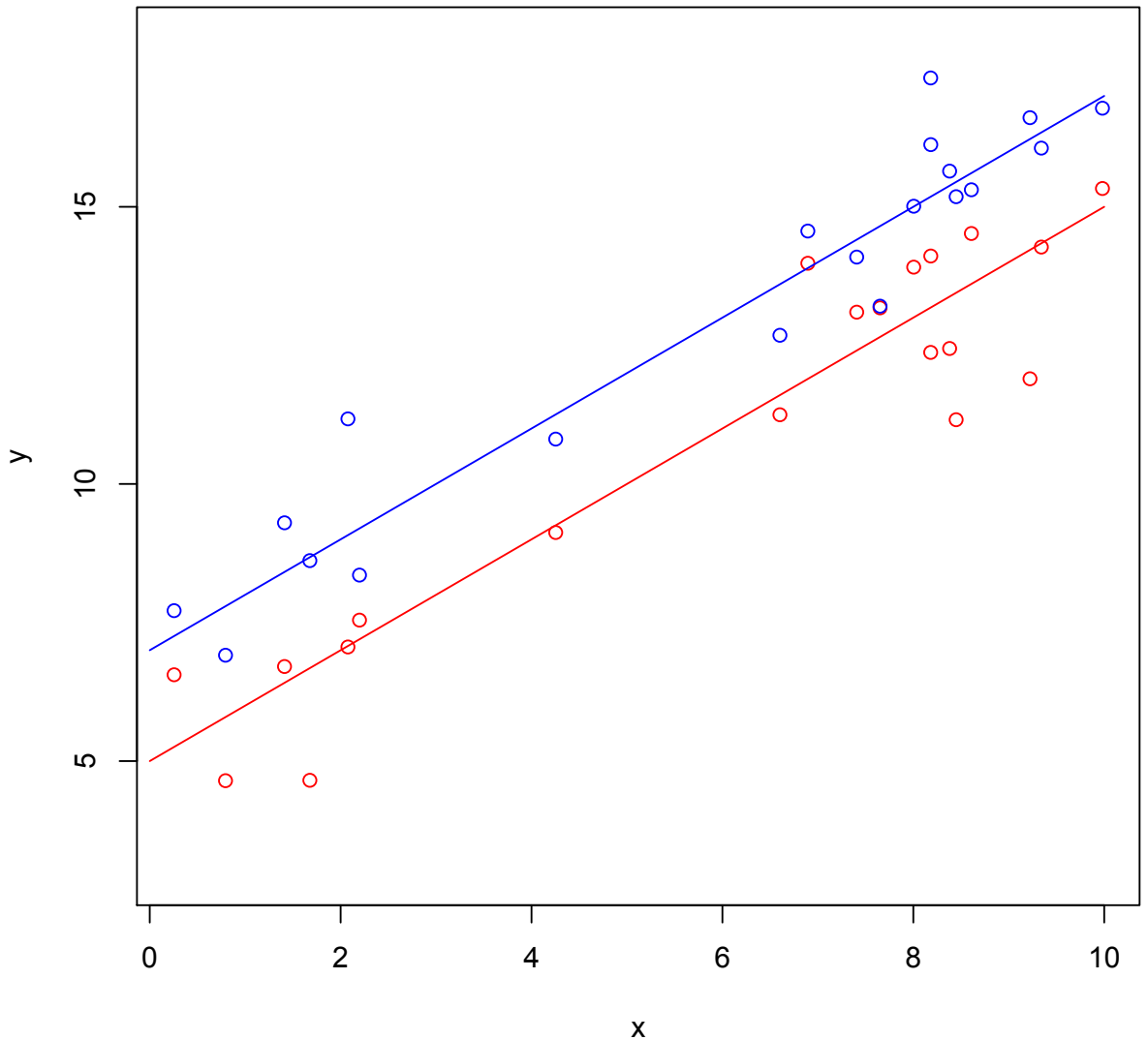
- Data are normal, including the unobservable random shocks.
- Design is balanced, meaning equal or proportional sample sizes.
- Then exact F-tests are usually possible.
- Tests for random effects are asking whether the variance of the (normal) random shock is zero.
- Balance matters.
- Normality matters.

Repeated measures

- An individual is tested under more than one condition, and contributes a response in each treatment condition.
- **One can view "subject" as just another random effects factor, because subjects supposedly were randomly sampled.**
- Subject would be nested within sex, but might cross stimulus intensity.
- Sex and simulus intensity are fixed effects.
- There might be other random effects in the model, like item.
- This is the classical way to analyze repeated measures.
- It's not the only way.
- It implies equal correlations between measurements from the same individuals. This is unrealistic for some data sets.
- But it generalizes naturally to models for categorical responses.

Random intercepts

This is all multiple regression under the surface



Dichotic listening study

- Left-handed and right-handed subjects push a key when they hear their names over background noise.
- They are wearing stereo headphones.
- Signal comes in the left ear, the right ear, or both.
- 50 trials in each condition, in random order.
- Dependent variable is median reaction time in milliseconds.
- Conceptually, there are two factors: handedness and ear.
- Repeated measures on ear.
- Technically there are three factors: handedness, ear and subject.
- Subject is a random effect that crosses ear and is nested within handedness.

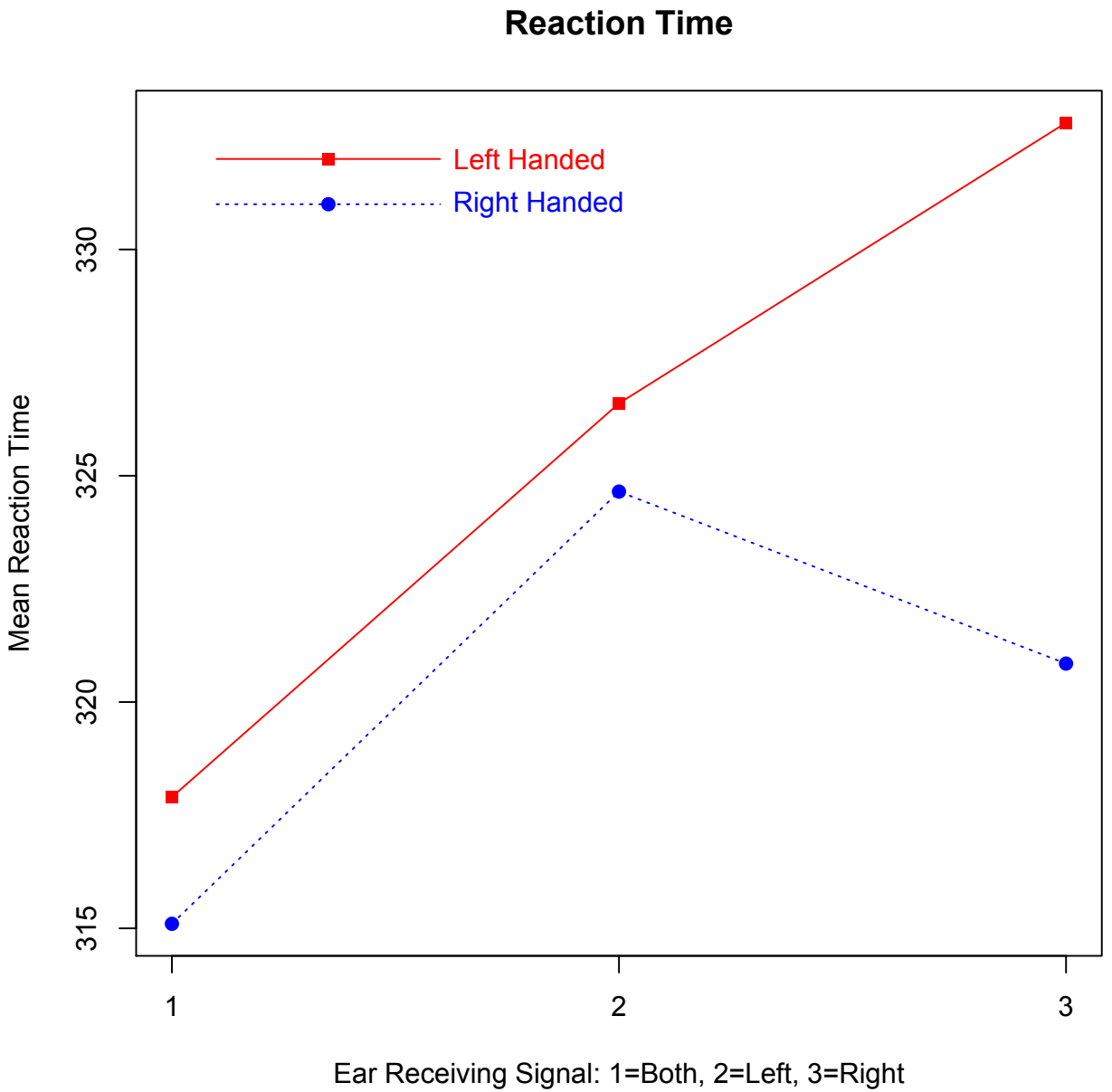
Treatment Means

	Presentation			
Handedness	Both ears	Left ear	Right ear	
Left	317.90	326.60	332.80	325.77
Right	315.10	324.65	320.85	320.20
	316.50	325.62	326.82	322.98

Vocabulary

- Marginal means
- Main effects
- Interaction

Interaction Plot



Another example (Baayen, Davidson and Bates , 2008)

Treatment is SOA

Subject	Item	Treatment	ReactionTime
s1	w1	Long	466
s1	w2	Long	520
s1	w3	Long	502
s1	w1	Short	475
s1	w2	Short	494
s1	w3	Short	490
s2	w1	Long	516
s2	w2	Long	566
s2	w3	Long	577
s2	w1	Short	491
s2	w2	Short	544
s2	w3	Short	526
s3	w1	Long	484
s3	w2	Long	529
s3	w3	Long	539
s3	w1	Short	470
s3	w2	Short	511
s3	w3	Short	528

The modern approach

For unbalanced designs

- Estimate the parameters with Restricted Maximum Likelihood (REML)
- Exact F-tests are still out of reach for unbalanced designs.
- There are good approximations for testing the fixed effects.
- Satterthwaite adjustment to df.
- So-called F statistics reduce to the classical Fs when the design is balanced.
- Inference for the random effects is a challenge.
- Not an issue for repeated measures analysis.

References

Garfinkel, H. (1967). *Studies in Ethnomethodology*. Englewood Cliffs, New Jersey: Prentice-Hall. See pp. 79-94

Baayen, R.H., Davidson, D.J. and Bates, D.M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390-412.

Copyright Information