

# 1 General Principles

## Within Cases

- A case contributes a DV value for every value of a categorical IV
- It is natural to expect data from the same case to be correlated - NOT independent
- For example, the same subject appears in several treatment conditions
- Hearing study: How does pitch affect our ability to hear faint sounds? The same subjects will hear a variety of different pitch and volume levels (in a random order). They press a key when they think they hear something.

## Student's Sleep Study (*Biometrika*, 1908)

- Patients take two sleeping medicines several days apart.
- Half get *A* first, half get *B* first.
- Reported hours of sleep are recorded.
- It's natural to subtract, and test whether the mean *difference* equals zero.
- That's what Gossett did.
- But some might do an independent *t*-test with  $n_1 = n_2$ .
- It's wrong, but is it harmful?

## Matched pairs, testing $H_0 : \mu_1 = \mu_2$

- If population covariance between the two measurements is positive, matched *t*-test has better power
  - Each case serves as its own control.
  - Many unknown influences are removed by subtraction.
  - This makes the analysis more precise.

- If population covariance between measurements is *negative*, independent *t*-test has Type I error rate greater than 0.05.
  - Matched *t*-test still has the correct Type I error rate.
  - Negative covariance is unlikely to happen in most real situations.

### Within-cases Terminology

- **Longitudinal:** The same variables are measured repeatedly over time. Usually there are lots of variables, including categorical ones, and large samples. If there's an experimental treatment, it's usually once at the beginning, like a surgery. Longitudinal studies basically track what happens over time.
- **Repeated measures:** Usually, the same subjects experience two or more experimental treatments. Usually quantitative DVs, and often small samples.

### Advantages of Within-cases Designs

- Convenience (sometimes)
- Each case serves as its own control. A huge number of extraneous variables are automatically held constant. The result can be a very sensitive analysis.
- For some models, you can have lots of measurements on just a few subjects — if you are willing to make some assumptions.

### Three Main Approaches

- Classical Univariate
- Multivariate
- Covariance Structure

## Classical Univariate Approach

- “Case” (or Subject) is one of the factors.
- Case is a *random effects* factor that is *nested* within combinations of the between-cases factors, and *crosses* the within-cases factors.
- Uses a mixed model ANOVA.
- $F$ -tests are possible only for balanced experimental designs.
- Can also do it with the covariance structure approach, and don’t need balance.

## Multivariate Approach

- When a case (subject) provides data under more than one set of conditions, it is natural to think of the measurements as multiple dependent variables.
- Start with the humble matched  $t$ -test

## 2 Within-Cases Factors Only

### Why the matched $t$ -test makes sense

- Let  $Y_1 =$  HS English mark
- Let  $Y_2 =$  HS Calculus mark
- Let  $D = Y_1 - Y_2$
- $E(Y_1) = \mu_1$  and  $E(Y_2) = \mu_2$
- $H_0 : \mu_1 = \mu_2$

$$E(D) = E(Y_1 - Y_2) = E(Y_1) - E(Y_2) = \mu_1 - \mu_2$$

$$E(D) = E(Y_1 - Y_2) = E(Y_1) - E(Y_2) = \mu_1 - \mu_2$$

- Mean difference is the difference between means.
- So test whether the population mean difference equals zero.
- Use a one-sample t-test of  $H_0 : \mu_D = 0$
- It's called a matched t-test

### **Multivariate version of the *t*-test**

- It's called Hotelling's *T*-squared
- Can be one-sample or two-sample
- Multivariate matched *t*: can test whether several mean differences are zero, simultaneously.
- Optimal for small samples if the data are multivariate normal.
- Robust against normality for moderately large *n*.
- No equal variance assumption.

### **Example: Grapefruit Sales**

- Cases are stores
- At each store, sales are measured at three different price levels, yielding  $Y_1$ ,  $Y_2$  and  $Y_3$
- Stores are randomly assigned to one of the 6 different orders
- Want to test  $H_0 : \mu_1 = \mu_2 = \mu_3$
- $D_1 = Y_1 - Y_2$  and  $D_2 = Y_2 - Y_3$
- Test  $H_0 : E(D_1) = E(D_2) = 0$
- Same as testing  $H_0 : \mu_1 - \mu_2 = \mu_2 - \mu_3 = 0$
- That is,  $H_0 : \mu_1 = \mu_2 = \mu_3$

### Mean of a difference is a difference between means

- Calculate a collection of difference variables. These are linear combinations (contrasts) of the data values for a case.
- Data values  $Y_1, \dots, Y_k$  are multivariate normal.
- So the collection of difference variables  $D_1, \dots, D_r$  is multivariate normal too (theorem).
- Expected value of a linear combination is a linear combination of expected values
- Test whether the expected values (population means) of the difference variables all equal zero.
- If there is more than one within-cases factor, no problem: Collections of contrasts represent main effects and interactions as usual. Note that these contrasts are all differences.

### Archery Example: Bow and Arrow

- Cases are archers. There are  $n$  archers.
- Test two bows, three arrow types.
- Warmup, then each archer takes 10 shots with each Bow-Arrow combination — 60 shots.
- In a different random order for each archer, of course.
- $Y_{i,1}, \dots, Y_{i,6}$  are mean distances from arrow tip to centre of target, for  $i = 1, \dots, n$ .
- Each  $Y_{i,j}$  is based on 10 shots.
- $E(Y_{i,j}) = \mu_j$  for  $j = 1, \dots, 6$ .

How would you do it?

	Arrow			
Bow	1	2	3	
1	$\mu_1$	$\mu_2$	$\mu_3$	$\frac{\mu_1 + \mu_2 + \mu_3}{3}$
2	$\mu_4$	$\mu_5$	$\mu_6$	$\frac{\mu_4 + \mu_5 + \mu_6}{3}$
	$\frac{\mu_1 + \mu_4}{2}$	$\frac{\mu_2 + \mu_5}{2}$	$\frac{\mu_3 + \mu_6}{2}$	

To test main effects for arrow type:

- For each archer  $i = 1, \dots, n$ , calculate

$$D_{i,1} = (Y_{i,1} + Y_{i,4}) - (Y_{i,2} + Y_{i,5})$$

$$D_{i,2} = (Y_{i,2} + Y_{i,5}) - (Y_{i,3} + Y_{i,6})$$

- Notice that the population mean differences are

$$E(D_{i,1}) = (\mu_1 + \mu_4) - (\mu_2 + \mu_5)$$

$$E(D_{i,2}) = (\mu_2 + \mu_5) - (\mu_3 + \mu_6)$$

- Test the multivariate hypothesis  $\boldsymbol{\mu} = \mathbf{0}$  on the difference variables, using Hotelling's  $T$ -squared.
- The null hypothesis is  $H_0 : \mu_1 + \mu_4 = \mu_2 + \mu_5 = \mu_3 + \mu_6$ .

### Hotelling's $T$ -squared

- The only trick is persuading SAS to calculate  $T^2$ .
- Official SAS documentation claims that SAS won't do it, but ...
- $T^2 = (n - 1) \left( \frac{1}{\lambda} - 1 \right)$ , so just get Wilks' Lambda from the `mtest` statement of `proc reg`. The  $p$ -value will be correct.
- In a regression model with *no independent variables*,  $E(\mathbf{D}_i) = \boldsymbol{\beta}_0$ , so test  $H_0 : \boldsymbol{\beta}_0 = \mathbf{0}$ .

```
proc reg;
  model D1 D2 = ;
  Arrow: mtest intercept=0;
```

### 3 Within and Between

#### One Between, One Within

- Grapefruit study
- Within stores factor: Three price levels
- Between-stores factor: Incentive program for produce managers (Yes-No)

#### Monkey Study

- Train monkeys on discrimination tasks, at 16, 12, 8, 4 and 2 weeks prior to treatment. Different task each time, equally difficult (randomize order).
- Treatment is to block function of the hippocampus (with drug, not surgery), re-tested. Get 5 scores for each monkey.

---

Train	Train	Train	Train	Train	Inject	Test
-16	-12	-8	-4	-2	0	

- 11 randomly assigned to treatment, 7 to control
- Treatment is between, time elapsed since training is within

#### Time by Treatment

	Time				
Treatment	2	4	8	12	16
Drug					
Saline					

Form linear combinations

- Four difference variables for the five time levels.
- Also, calculate average score across the 5 time levels.
- Now we have 5 potential DVs.

### Main Effect of Treatment

	Time				
Treatment	2	4	8	12	16
Drug					
Saline					

- DV is average discrimination score, averaging across time.
- Means of this variable for the two treatment groups are the marginal means, averaging across time.
- So just do an ordinary univariate  $F$ -test for difference between means.

### Main effect of Time

- DVs are the 4 difference variables.
- With effect coding on Treatment, intercepts are grand means: Averages of the between-cases cell means
- Multivariate test of the intercepts = zero is a test of whether, averaging across the between-subjects factor, the marginal means of the within-subjects factor are equal.
- This is the test for main effect of time

	$E(Y_1 - Y_2 x)$	$E(Y_2 - Y_3 x)$	$E(Y_3 - Y_4 x)$	$E(Y_4 - Y_5 x)$
Drug	$\beta_{01} + \beta_{11}$	$\beta_{02} + \beta_{12}$	$\beta_{03} + \beta_{13}$	$\beta_{04} + \beta_{14}$
Saline	$\beta_{01} - \beta_{11}$	$\beta_{02} - \beta_{12}$	$\beta_{03} - \beta_{13}$	$\beta_{04} - \beta_{14}$

$$H_0 : \beta_{01} = \beta_{02} = \beta_{03} = \beta_{04} = 0$$

### Time by Treatment Interaction

- DV is the set of difference variables.
- IV is Treatment.
- If the average set of mean differences (trend over time) *depends* on treatment, then there is an interaction.



## In General

- Suppose there are  $k$  measurement for each case
- Calculate difference (contrast) variables to represent main effects and interactions of the within-cases factors. There will be  $k - 1$  of them.
- Calculate one more variable: The mean or sum of the  $k$  measurements
- Do a multivariate regression, with effect coding on the between-cases factors.

## Tests

- Tests of within-cases main effects and interactions are tests on the intercepts.
- Tests of between-cases main effects and interactions are tests on the sum or average variable.
- Interactions between between and within-cases factors: Test the between-cases IVs on the difference variables.
- For example, if  $A$  and  $B$  are between and  $C$  and  $D$  are within, test of  $A \times B$  on the set of differences representing the  $C \times D$  interaction is a test of the  $A \times B \times C \times D$  interaction, because the  $C \times D$  interaction *depends* on the  $A \times B$  combination.