

LINEAR MODELS FOR LONGITUDINAL DATA

Estimation of the Mean Model (β -coefficients)

Weighted/Generalized Least Squares

- **General linear model** for longitudinal data is:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

where

$$\boldsymbol{\epsilon} \sim MVN(0, V)$$

where V is the **variance-covariance** matrix of $\boldsymbol{\epsilon}$

- How are the regression parameters $\boldsymbol{\beta}$ estimated?
- Recall that the log-likelihood is

$$L(\boldsymbol{\beta}) = -\frac{1}{2}N \log(2\pi) - \frac{1}{2} \log |V| - \frac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'V^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \quad (1)$$

- Assuming for the moment that V is known, the **maximum likelihood estimator** $\hat{\beta}$ is obtained by **minimizing**

$$\text{RSS} = (\mathbf{y} - X\boldsymbol{\beta})'V^{-1}(\mathbf{y} - X\boldsymbol{\beta}) \quad (2)$$

which is the **weighted sum of squared errors**

- $\hat{\beta}$ that minimizes RSS is called the **weighted least squares** estimator. It is also maximum likelihood
- In practice, v-c-c parameters γ are estimated via ReML, and \hat{V} are plugged in RSS to estimate β by minimizing RSS
- With **correlated data**, V is of **block diagonal** form:

$$\begin{pmatrix} V_1 & 0 & \dots & 0 \\ 0 & V_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & V_m \end{pmatrix}$$

so that RSS is

$$\text{RSS} = \sum_i (\mathbf{y}_i - X_i \boldsymbol{\beta})' \hat{V}_i^{-1} (\mathbf{y}_i - X_i \boldsymbol{\beta}) \quad (3)$$

(check at home)

- **Special case:** If the data are **independent** with constant variance, then \hat{V}_i (variance-covariance matrix for the i th subject) takes the form

$$\begin{pmatrix} \hat{\sigma}^2 & 0 & \dots & 0 \\ 0 & \hat{\sigma}^2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \hat{\sigma}^2 \end{pmatrix}$$

and RSS takes the form

$$\text{RSS} = \sum_{ij} \frac{1}{\hat{\sigma}^2} (y_{ij} - \mathbf{x}'_{ij} \boldsymbol{\beta})^2, \quad (4)$$

and minimizing RSS gives rise to the OLS estimator (check at home)

- The OLS estimator equally “weights” each observation by $1/\hat{\sigma}^2$.
- Comparing (3) and (4), we see that WLS generalizes OLS by “weighting” i th subject’s data by \hat{V}_i^{-1}
- RSS (3) sums over subjects i , “weighting” the “square” of each subjects’ vector of residuals $(\mathbf{y}_i - X_i\boldsymbol{\beta})$ by the inverse of the estimated variance-covariance matrix \hat{V}_i of \mathbf{Y}_i

- Minimizing RSS:

Use the fact of matrix derivation (\mathbf{x} is $n \times 1$, A is $n \times n$):

$$\frac{\partial \mathbf{x}' A \mathbf{x}}{\partial \mathbf{x}} = \mathbf{x}' (A + A'),$$

we have

$$\begin{aligned} \frac{\partial \text{RSS}}{\partial \boldsymbol{\beta}} &= 2(\mathbf{y} - X\boldsymbol{\beta})' V^{-1} \frac{\partial (\mathbf{y} - X\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \\ &= -2(\mathbf{y} - X\boldsymbol{\beta})' V^{-1} X \end{aligned}$$

Solve $\frac{\partial \text{RSS}}{\partial \boldsymbol{\beta}} = 0$ and we can obtain **WLS/GLS estimator**:

$$\hat{\boldsymbol{\beta}} = (X' \hat{V}^{-1} X)^{-1} X' \hat{V}^{-1} \mathbf{y}$$

- In the case of independent and homoscedastic data, this reduces to OLS:

$$\hat{\boldsymbol{\beta}}_I = (X' X)^{-1} X' \mathbf{y}$$

- **Special case:**

WLS estimator for independent observations with Heteroskedasticity

- ϵ_i ($i = 1, \dots, N$) are independent but have different variance $\text{var}(\epsilon_i) = \sigma_i^2$ in the model

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

where $\boldsymbol{\epsilon} \sim MVN(0, V)$, and V is

$$\begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_N^2 \end{pmatrix}$$

- WLS is a solution to Heteroskedasticity:
 - Model the variance $\sigma_i^2 = f(x_i)$ as a function of x_i to estimate $\hat{\sigma}_i$
 - obtain WLS estimator by plugging \hat{V} in $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\hat{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\hat{V}^{-1}\mathbf{y}$

- Variance-covariance matrix of $\hat{\beta}$ is

$$\begin{aligned}\text{var}(\hat{\beta}|X) &= (X'V^{-1}X)^{-1}X'V^{-1}\text{var}(\mathbf{y}|X)V^{-1}X(X'V^{-1}X)^{-1} \\ &= (X'V^{-1}X)^{-1}X'V^{-1}VV^{-1}X(X'V^{-1}X)^{-1} \\ &= (X'V^{-1}X)^{-1}\end{aligned}$$

Plug in estimated \hat{V} to have **estimated variance-covariance matrix** of $\hat{\beta}$

$$\widehat{\text{var}}(\hat{\beta}|X) = (X'\hat{V}^{-1}X)^{-1}$$

- Note: This shows that the estimated $\widehat{\text{var}}(\hat{\beta}|X)$ rely on correctly specified v-c-c model for V
- This $(\hat{\beta}$ and $\widehat{\text{var}}(\hat{\beta}))$ is what is computed by computer packages implementing WLS regression, including SAS proc mixed.

- **Note:** For any (full rank) symmetric weight matrix W , WLS estimator

$$\hat{\beta}_W = (X'WX)^{-1}X'W\mathbf{y}$$

is unbiased, since

$$E(\hat{\beta}_W) = (X'WX)^{-1}X'WX\beta = \beta.$$

- Variance of $\hat{\beta}_W$ is then

$$\text{var}(\hat{\beta}|X) = (X'WX)^{-1}X'WVWX(X'WX)^{-1}$$

- If ignore correlation and use OLS estimator:
 - $\hat{\beta}_{OLS} = (X'X)^{-1}X'\mathbf{y}$ is still unbiased
 - True variance $\text{var}(\hat{\beta}_{OLS}|X) = (X'X)^{-1}X'VX(X'X)^{-1}$
 - If ignore correlation, variance estimator $\widehat{\text{var}}(\hat{\beta}_{OLS}) = \sigma^2(X'X)^{-1}$ is biased
- We will talk about robust estimator for $\text{var}(\hat{\beta})$ to “fix up” misspecified V model in future

- WLS estimator with $W = V^{-1}$ is **best linear unbiased estimator (BLUE)**:

To see this, recall:

– $\hat{\beta}_{OLS} = (X'X)^{-1}X'y$ is BLUE when $\epsilon \sim MVN(0, \sigma^2 I)$,

– V can be orthogonally diagonalized as $C'VC = \Lambda$,

where C is matrix of eigenvectors corresponding to matrix of eigenvalues Λ .

Let $V^{-1/2} = C\Lambda^{-1/2}C'$, we have $V^{-1/2}VV^{-1/2'} = I$

Hence, for linear transformed response $V^{-1/2}Y$, we have

$$V^{-1/2}Y = V^{-1/2}X\beta + V^{-1/2}\epsilon$$

where $V^{-1/2}\epsilon \sim MVN(0, I)$.

OLS estimator for this linear transformed response is our original WLS estimator, which is BLUE.

Inferences in the Mean Model

- We have now seen how models are fitted using WLS
- Now how about hypothesis tests and confidence intervals for β ?
- Two options:
 - LRT for hypothesis tests about β (with χ^2 -based P -values), but ML (not ReML) is required for these tests
 - \mathcal{F} -tests (and related confidence intervals)

\mathcal{F} -test

- Eg, with $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4)'$ we would test

$$H_0 : \beta_1 = \beta_4 = 0 \quad \text{vs.} \quad H_A : \beta_1 \neq 0 \text{ or } \beta_4 \neq 0$$

- This can be set up with a **contrast matrix**, L :

We wish to test $L\boldsymbol{\beta} = 0$, for this example,

$$L = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

- In general, we would test

$$H_0 : L\boldsymbol{\beta} = 0 \quad \text{vs.} \quad H_A : L\boldsymbol{\beta} \neq 0$$

- Then we compute $L\hat{\beta}$ and

$$\widehat{\text{var}}(L\hat{\beta}) = L\widehat{\text{var}}(\hat{\beta})L' = L\hat{C}L'$$

and finally

$$F = (L\hat{\beta})'(L\hat{C}L')^{-1}(L\hat{\beta})$$

where $\hat{C} = \widehat{\text{var}}(\hat{\beta}|X) = (X'\hat{V}^{-1}X)^{-1}$

- Statistic F has an approximate \mathcal{F} -distribution
 - numerator DF=2 (rank of L)
- What about the denominator DF?
 - There are several methods for estimating the denominator DF.
 - One that SAS uses is called the “between-within” method

Note on Computation of Degrees of Freedom

- SAS `proc mixed` has several ways of handling degrees of freedom and provides some explanation / references for each of these approaches.
- Note SAS does not use the “between-within” method with the unstructured variance-covariance model.
DDFM=KENWARDROGER may be a better option for this case.
- Some relevant material from the `proc mixed` manual for DDFM=BETWITHIN option:

It is computed by dividing the residual degrees of freedom into between-subject and within-subject portions. PROC MIXED then checks whether a fixed effect changes within any subject. If so, it assigns within-subject degrees of freedom to the effect; otherwise, it assigns the between-subject degrees of freedom to the effect.

- There are other ways to estimate the denominator DF. See, e.g., McCulloch and Searle (Generalized, Linear, and Mixed Models. 2001. New York: Wiley), p.167
- If the correlation structure and parameter values are **known**, and only the total variance $\text{var}(\epsilon_{ij}) = \sigma^2$ is being estimated, then the denominator DF is

N – total number of covariates (incl. intercept)

Example: \mathcal{F} -test and Confidence Interval for One Parameter

- **Example:** Nepalese children: Suppose we wanted to test for an interaction between age and weight in predicting arm circumference
- We fit the following mean model assuming “exchangeable plus exponential” correlation structure (assume we have already decided to eliminate the age^2 term)

$$E(\text{arm}_{ij}) = \beta_0 + \beta_1 \text{wt}_{ij} + \beta_2 \text{age}_{ij} + \beta_3 \text{sex}_i + \beta_4 \text{wt}_{ij} * \text{age}_{ij}$$

and test

$$H_0 : \beta_4 = 0 \quad \text{vs.} \quad H_A : \beta_4 \neq 0$$

- We wish to test

$$H_0 : L\boldsymbol{\beta} = 0 \quad \text{vs.} \quad H_A : L\boldsymbol{\beta} \neq 0$$

where

$$L = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

- Compute F :

$$F = \frac{(\hat{\beta}_4 - 0)^2}{\widehat{\text{var}}(\hat{\beta}_4)},$$

- Approximately, $F \sim \mathcal{F}$ distribution with numerator DF=1 (Just like in OLS regression)

- Here is the SAS code

```
proc mixed data=temp dfbw ;
  title2 "1st run" ;
  class id;
  model arm = wt age sex wtage / s cl ;
  random intercept / subject=id ;
  repeated / subject=id type=sp(pow)(age4) ;
  contrast 'F test for wt'
    wt      1 ,
    wtage   1 ;
  estimate 'wt estimate for 2 yrs'
    wt      1  wtage -12
    / cl ;
run;
```

- CI's are given by PROC MIXED with cl option on model statement
- contrast and estimate will be used for tests later.

Model Information

Dependent Variable	arm
Covariance Structures	Variance Components, Spatial Power
Subject Effects	id, id
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Between-Within

<snip>

Dimensions

Covariance Parameters	3
Columns in X	5
Columns in Z Per Subject	1
Subjects	197
Max Obs Per Subject	5
Number of Observations Read	1000
Number of Observations Used	877
Number of Observations Not Used	123

<snip>

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > t	Alpha
Intercept	8.1660	0.2552	195	32.00	<.0001	0.05
wt	0.6593	0.02367	677	27.85	<.0001	0.05
age	-0.05142	0.006181	677	-8.32	<.0001	0.05
sex	0.2895	0.09502	195	3.05	0.0026	0.05
wtage	-0.00070	0.000408	677	-1.72	0.0866	0.05

Solution for Fixed Effects

Effect	Lower	Upper
Intercept	7.6627	8.6694
wt	0.6128	0.7057
age	-0.06355	-0.03928
sex	0.1021	0.4769
wtage	-0.00150	0.000101

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
wt	1	677	775.44	<.0001
age	1	677	69.19	<.0001
sex	1	195	9.28	0.0026
wtage	1	677	2.94	0.0866

- The F statistic for testing $\beta_4 = 0$ is computed as

$$F = \frac{(-0.00070 - 0)^2}{(0.000408)^2} = 2.94 = (-1.72)^2 ,$$

and the P -value (0.0866) is obtained using an \mathcal{F} -distribution with 677 denominator DF

- Conclusion: At $\alpha = 0.05$, these data do not strongly suggest a weight \times age interaction

- Note: $\sqrt{F} = |t|$ with the same (denominator) degrees of freedom, and this can be used as in OLS to **construct confidence** intervals:
- A 95% CI for β_4 is

$$\hat{\beta}_4 \pm t_{\text{DF},.975} \hat{\text{se}}(\hat{\beta}_4)$$

which here is very close to a CI based on a Z (because DF=677)

$$-0.00070 \pm t_{677,.975} \times 0.000408 = [.000150, 0.000101]$$

- Computation of DF for nepal data:
 - Between DF:
Total number of subjects: 197
Number of between covariates: 2
Residual between DF: $197 - 2 = 195$
 - Then, the residual between DF is applied to all “between effects”
 - Within DF:
Total number of observations: 877
Total number of subjects: 197
Total within DF: $877 - 197 = 680$
Number of within covariates: 3
Residual within DF: $680 - 3 = 677$
 - Then, residual within DF is applied to all “within effects”

Example: \mathcal{F} -test for Vector-Valued Parameter

- Example: Suppose we wished to test: whether weight had any effect on arm circumference, allowing for possible interactions with age
- In this model

$$E(\text{arm}_{ij}) = \beta_0 + \beta_1 \text{wt}_{ij} + \beta_2 \text{age}_{ij} + \beta_3 \text{sex}_i + \beta_4 \text{wt}_{ij} * \text{age}_{ij}$$

we would test

$$H_0 : \beta_1 = \beta_4 = 0 \quad \text{vs.} \quad H_A : \beta_1 \neq 0 \text{ or } \beta_4 \neq 0$$

- We wish to test

$$H_0 : L\beta = 0 \quad \text{vs.} \quad H_A : L\beta \neq 0$$

where

$$L = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

- Numerator DF = 2, Denominator DF = within DF
- SAS does this for you with the contrast statement in PROC MIXED:

```
contrast 'F test for wt'
      wt      1 ,
      wtage   1 ;
```

The contrast statement sets up an L matrix, wherein:

- 1st row: all 0's except 1 for the coefficient of wt
- 2nd row: all 0's except 1 for the coefficient of wtage

- This results in the output

Contrasts				
Label	Num DF	Den DF	F Value	Pr > F
F test for wt	2	677	401.42	<.0001

- Conclusion: Weight is a significant predictor of arm circumference, adjusting for sex and allowing for differential effects of weight by age
- Note: The contrast statement is like the test statement in stata

Example: Estimating Linear Combinations of Mean Parameters

- Example: We wanted to estimate the effect of 1-unit change in wt for two-year olds (24 months)
- Note: age (in Month) is centerized at 36 before fitting model
→
 - effect of 1-unit change in wt for three-year olds (36 months)
 $= \beta_1$
 - effect of 1-unit change in wt for two-year olds (24 months)
 $= \beta_1 + (24 - 36)\beta_4$
- L matrix (with one row):

$$L = (0, 1, 0, 0, -12)$$

so that we are estimating $L\boldsymbol{\beta} = \beta_1 - 12\beta_4$ for this example

- Estimate $L\beta$ by $L\hat{\beta}$
- For variance of $L\hat{\beta}$, recall:

$$\widehat{\text{var}}(L\hat{\beta}) = L\widehat{\text{var}}(\hat{\beta})L' = L(X'\hat{V}^{-1}X)^{-1}L'$$

- This is obtained via the estimate statement in PROC MIXED

```
estimate 'wt estimate for 2 yrs'
      wt      1  wtage -12
      / cl ;
```

- Output

Estimates					
Label	Estimate	Standard Error	DF	t Value	Pr > t
wt estimate for 2 yrs	0.6677	0.02357	677	28.33	<.0001

Estimates

Label	Alpha	Lower	Upper
wt estimate for 2 yrs	0.05	0.6214	0.7139

- Interpretation: Among 2-year olds, on average, for each difference in one unit of weight, arm circumference differs by 0.6677 cm, adjusting for sex
- Note: The estimate statement is like `lincom` in stata

Alternative Test for Vector-Valued Parameter Likelihood Ratio Test

- If you are not able to construct contrasts for \mathcal{F} -tests, you could also use a LRT with a chi-square test
- If you do use an LRT, you must be **sure** you are using the full likelihood (not ReML) and are doing joint estimation of the mean (β) and variance parameters (γ)
- The \mathcal{F} -test with ReML: does not involve re-estimation of the variance-covariance parameters under the simpler (null) model
- If you use a LRT, you need to fit two models.

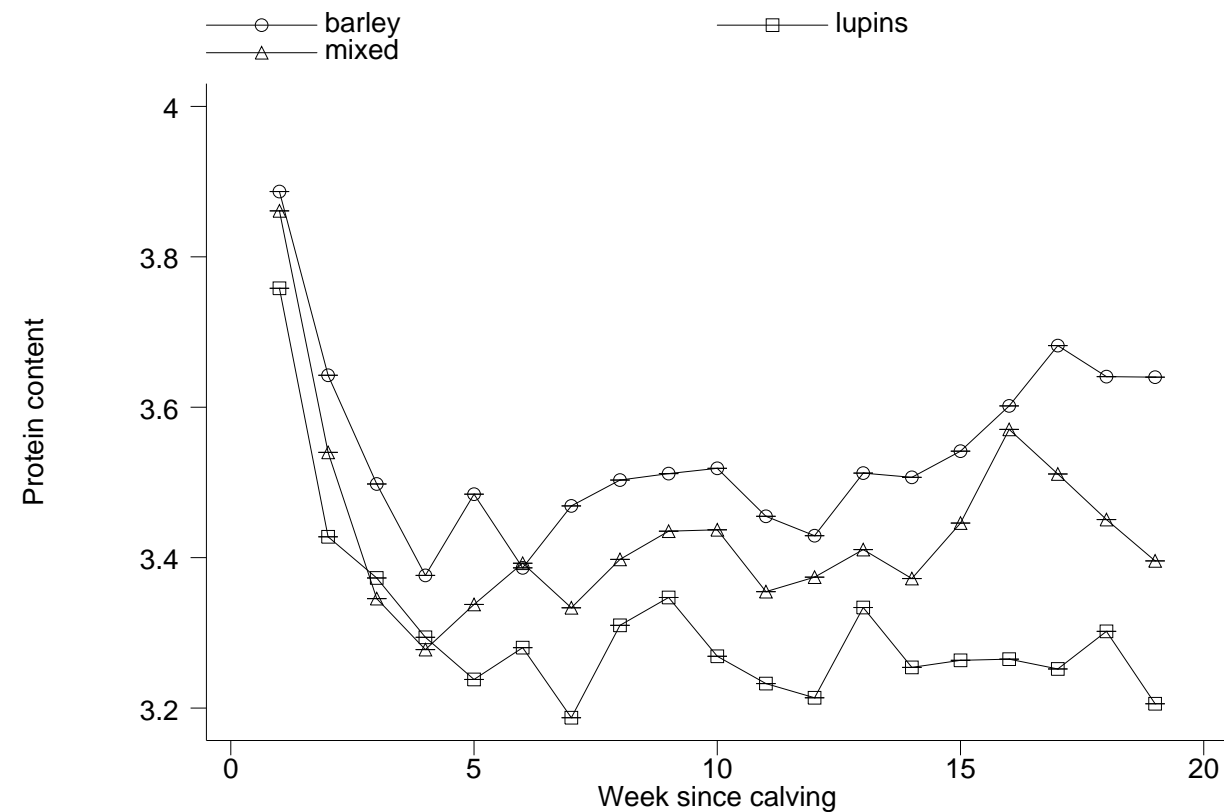
Some Notes on Tests and CIs for Mean Parameters

- The \mathcal{F} -tests are more convenient and more powerful than an LRT using a χ^2 -distribution
- The \mathcal{F} -tests are more accurate especially for small-to-moderate sample sizes
- All inferences (standard errors, tests, confidence intervals) for mean model parameters β rely on the v-c-c model for V being **correctly-specified**, at least to a reasonable approximation.
 - So, some effort in modeling V is warranted.
 - Will discuss ways around this problem (for large sample sizes) later
- Use contrast for omnibus tests. This will be especially important for sets of dummy variables corresponding to a categorical covariate

Mean Model Building and Inferences

Case Study in Time Profile Analysis

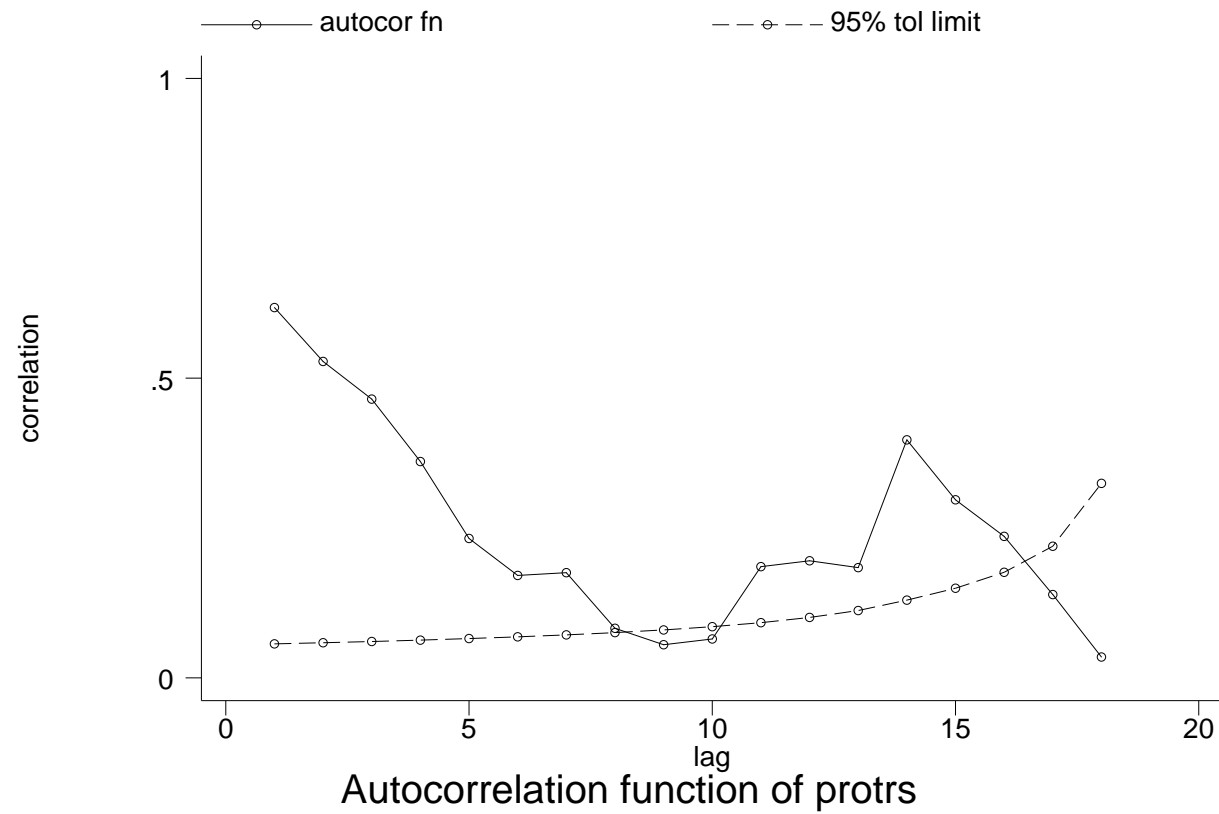
- **Example:** Protein content in cows' milk depends on diet
- Recall mean response profile for each diet group



- Scientific goal: Relationship of time profile (since calving) of protein content of milk to diet (lupins, barley, mixed)
- Preliminary observations:
 - Protein content drops strongly over first several weeks after calving and then levels off
 - Cows on diets with higher barley content have higher protein content (on average)
 - Mean response profiles for each group follow a somewhat similar pattern over time
- Selection of variance-covariance-correlation model (sketch)

Recall from earlier exploratory analysis that

 - Variance was approximately constant over time
 - The autocorrelation function (via correlogram) was:



- Correlation does not approach one for short lags
- Correlation drops off strongly over time
- Some positive autocorrelation at long lags

- Suggests starting with “exponential plus measurement error” model and comparing to “exchangeable plus exponential” model
 - Use flexible mean model with different mean for each week \times diet
 - “Exponential plus measurement error”:

```
proc mixed data=temp dfbw covtest ;
  title '1st run -- expon + mst error';
  class id diet week;
  model prot = diet*week / ;
  repeated / subject=id type=sp(pow)(week) local ;
run;
```

This yields:

Model Information

Data Set	WORK.TEMP
Dependent Variable	prot
Covariance Structure	Spatial Power
Subject Effect	id
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based

Degrees of Freedom Method Between-Within

Dimensions

Covariance Parameters	3
Columns in X	58
Columns in Z	0
Subjects	79
Max Obs Per Subject	19

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
Variance	id	0.07502	0.007697	9.75	<.0001
SP(POW)	id	0.8591	0.02004	42.88	<.0001
Residual		0.02246	0.002074	10.83	<.0001

Fit Statistics

-2 Res Log Likelihood	69.7
-----------------------	------

- “Exchangeable plus exponential”:

```
proc mixed data=temp dfbw ;
  title '1st run -- exch + expon';
  class id diet week;
  model prot = diet*week / ;
  random intercept / subject=id ;
  repeated / subject=id type=sp(pow)(week) ;
run;
```

This yields

Model Information

Data Set	WORK.TEMP
Dependent Variable	prot
Covariance Structures	Variance Components, Spatial Power

Covariance Parameter Estimates

Cov Parm	Subject	Estimate
Intercept	id	0.01838

SP(POW)	id	0.5646
Residual		0.07662

	Fit Statistics	
-2 Res Log Likelihood		126.8

- The two models have the same number of parameters.
- Comparing (ReML) log-likelihood values, we would conclude that the exponential plus measurement error model is a better fit.
- We proceed with mean modeling using the exponential plus measurement error v-c-c model

- **Model 1: Linear model for each time group.**

$$\begin{aligned} E(Y_{ij}|X_i) = & \beta_0 + \beta_1(\text{week}_{ij} - 1) + \beta_2 I(\text{diet}_i = 2) + \beta_3 I(\text{diet}_i = 3) \\ & + \beta_4 I(\text{diet}_i = 2) \times (\text{week}_{ij} - 1) \\ & + \beta_5 I(\text{diet}_i = 3) \times (\text{week}_{ij} - 1) , \end{aligned}$$

where Y_{ij} is protein content for i th cow on j th time point

Here, recall

- β_1 = difference in mean response for each unit difference in time (time slope) among cows on diet 1
- β_2 = difference in mean response comparing cows on diet 2 to those in diet 1, **at** week 1 (baseline treatment effect)
- β_4 = difference in time slope of mean response comparing diet 2 to diet 1 (response drops more quickly on diet 2 than diet 1)

or

estimated difference in treatment effect comparing diet 2 to diet 1
for each unit difference in time (slope of treatment effect; treatment
effect greater with increasing time)

SAS proc mixed code:

```
data temp;
  set cows;
  mixed = 1*(diet eq "mixed");
  barley = 1*(diet eq "barley");
  wk1=week-1 ;
    mixedwk1 = 1*(diet eq "mixed")*wk1;
    barleywk1 = 1*(diet eq "barley")*wk1;
  wk3=1*(week ge 3)*(week-3) ;
    mixedwk3 = 1*(diet eq "mixed")*wk3;
    barleywk3 = 1*(diet eq "barley")*wk3;
run;
proc mixed data=temp dfbw ;
  title '2nd run -- linear interaction';
  class id diet ;
  model prot = mixed barley wk1 mixedwk1 barleywk1 / s outpm=run2 ;
  repeated / subject=id type=sp(pow)(week) local ;
run;
```

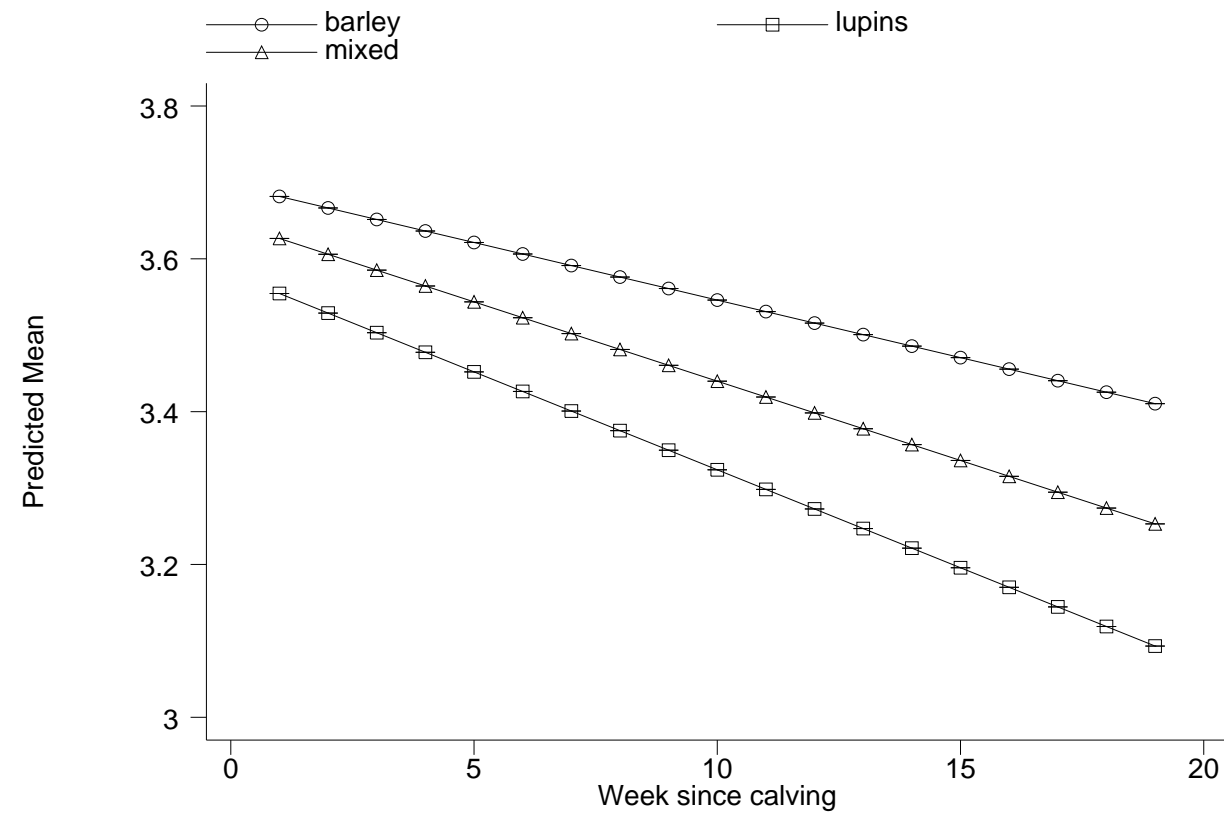
Some output:

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > t
Intercept	3.5546	0.05626	76	63.18	<.0001
mixed	0.07209	0.07953	76	0.91	0.3676
barley	0.1270	0.08110	76	1.57	0.1214
wk1	-0.02563	0.005183	1255	-4.95	<.0001
mixedwk1	0.004878	0.007308	1255	0.67	0.5045
barleywk1	0.01057	0.007437	1255	1.42	0.1556

A plot of fitted response profiles:

```
proc sort data=run2; by week; run;
proc sgplot data=run2;
    series y=pred x=week/group=diet;
run;
```

- Model does not fit the data (not flexible enough), **but** can see how time trends differ across groups
- Obtain very easily interpretable parameters capturing the effect of diet on protein content at week 1 **and** on the slope over time of

protein content

- In many longitudinal studies comparing groups over time, treatment by time interactions are the most important parameters
 - How do time trajectories vary with treatment?

- **Model 2: Linear spline model with a “knot” at 3 weeks.**

In this model, we assume that each group has a linear trend over weeks 1–3, and then a possibly different linear trend over weeks 3–19, where the two lines connect at week 3

$$\begin{aligned}
 E(Y_{ij}|X_i) = & \beta_0 + \beta_1(\text{week}_{ij} - 1) + \beta_2 I(\text{diet}_i = 2) + \beta_3 I(\text{diet}_i = 3) \\
 & + \beta_4 I(\text{diet}_i = 2) \times (\text{week}_{ij} - 1) \\
 & + \beta_5 I(\text{diet}_i = 3) \times (\text{week}_{ij} - 1) \\
 & + \beta_6 (\text{week}_{ij} - 3)_+ \\
 & + \beta_7 I(\text{diet}_i = 2) \times (\text{week}_{ij} - 3)_+ \\
 & + \beta_8 I(\text{diet}_i = 3) \times (\text{week}_{ij} - 3)_+ ,
 \end{aligned}$$

where

$$(\text{week}_{ij} - 3)_+ = \begin{cases} 0 , & \text{week} \leq 3 \\ (\text{week} - 3) , & \text{week} > 3 \end{cases}$$

This is called a **linear spline model** with a knot at 3 weeks

Interpretation of parameters: β_1 (for lupins), $(\beta_1 + \beta_6)$ (for lupins), β_6 (for lupins), β_4 , $(\beta_4 + \beta_7)$:

SAS proc mixed code:

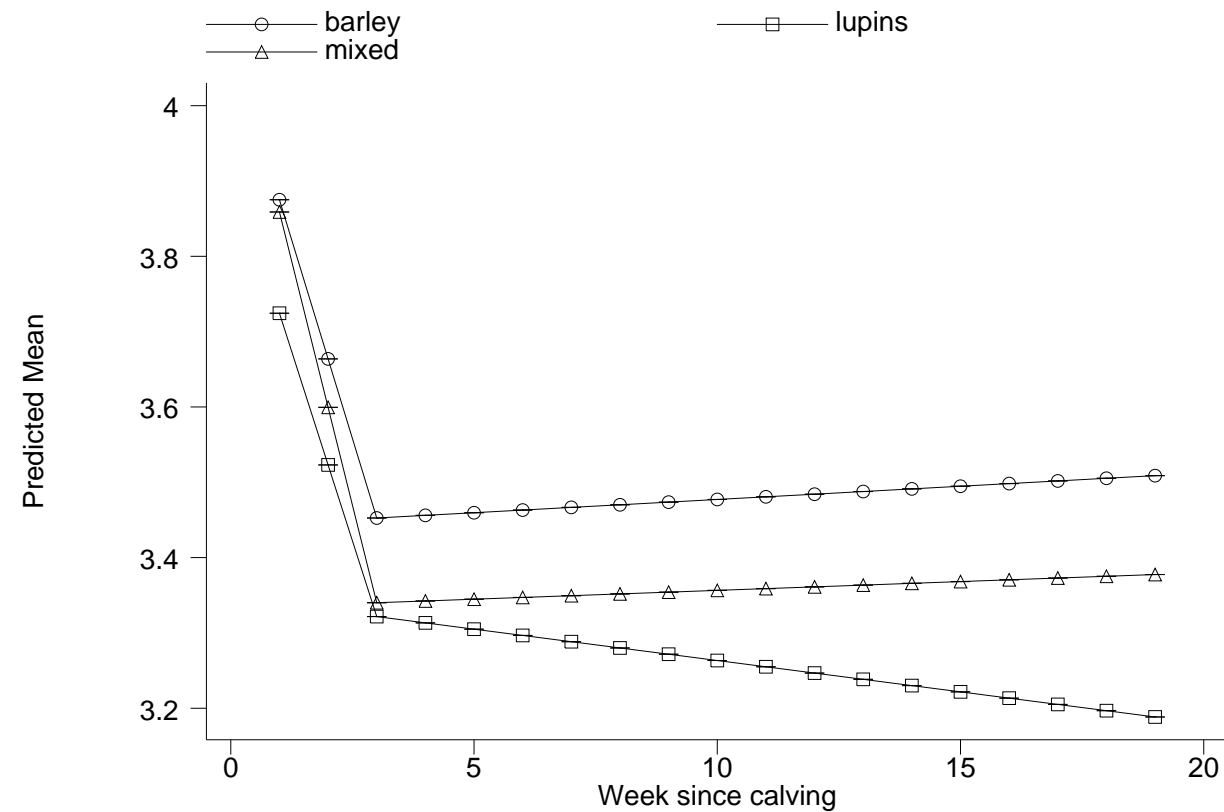
```
proc mixed data=temp dfbw ;  
  title '3rd run -- linear splines w interaction';  
  class id diet ;  
  model prot = mixed barley wk1 mixedwk1 barleywk1 wk3 mixedwk3 barleywk3  
    / s outpm=run3 ;  
  repeated / subject=id type=sp(pow)(week) local ;  
  contrast 'spline interactions'  
    mixedwk3 1 ,  
    barleywk3 1 ;  
run;
```

Some output:

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > t
Intercept	3.7244	0.05826	76	63.93	<.0001
mixed	0.1345	0.08239	76	1.63	0.1067
barley	0.1508	0.08404	76	1.79	0.0768
wk1	-0.2014	0.02615	1252	-7.70	<.0001
mixedwk1	-0.05804	0.03698	1252	-1.57	0.1168
barleywk1	-0.00993	0.03771	1252	-0.26	0.7924
wk3	0.1930	0.02815	1252	6.86	<.0001
mixedwk3	0.06871	0.03979	1252	1.73	0.0845
barleywk3	0.02177	0.04057	1252	0.54	0.5916

A plot of fitted response profiles



- This model allows for a different time trend for each of the three groups, each linear with a knot at 3 weeks.

– Can test

$$H_0 : \beta_7 = \beta_8 = 0 \text{ vs. } H_A : \beta_7 \neq 0 \text{ or } \beta_8 \neq 0$$

The contrast statement tests H_0 and yields:

Contrasts

Label	Num DF	Den DF	F Value	Pr > F
spline interactions	2	1252	1.55	0.2117

we could conclude that the spline interaction terms are not needed

- **Model 3: “Semiparametric” model with linear interactions.**

A model that is very flexible with respect to time, but that constrains how the groups differ from one another

$$\begin{aligned} E(Y_{ij}|X_i) = & \alpha_0 + \alpha_2 I(\text{week}_{ij} = 2) + \dots + \alpha_{19} I(\text{week}_{ij} = 19) \\ & + \beta_2 I(\text{diet}_i = 2) + \beta_3 I(\text{diet}_i = 3) \\ & + \beta_4 I(\text{diet}_i = 2) \times (\text{week}_{ij} - 1) \\ & + \beta_5 I(\text{diet}_i = 3) \times (\text{week}_{ij} - 1) \end{aligned}$$

This model contains a flexible baseline **profile analysis**, a set of “main effects” (treatment comparison at 1st week), and a parsimonious treatment-by-time effect:

- $\alpha_0, \alpha_2, \dots, \alpha_{19}$ flexibly model the time trend for diet 1 (using information from all three diets, owing to constraints in how trajectories vary from group to group)

- β_2 = difference in mean response comparing cows on diet 2 to those in diet 1, **at** week 1 (baseline treatment effect)
 - β_4 = difference in time trend of mean response comparing diet 2 to diet 1 (response drops more quickly on diet 2 than diet 1)
- or
- estimated difference in treatment effect comparing diet 2 to diet 1 for each unit difference in time (slope of treatment effect)
- **Note:** This is exactly the same interpretation as in the linear model, but in a more flexible and better fitting model
 - This model does quite a nice job of capturing the time trends in the data (flexibility with respect to trends “shared” across all groups)
 - But, also yields very interpretable parameters about effects of diet on:
 - overall level at baseline: β_2, β_3
 - time trend: β_4, β_5

SAS proc mixed code:

```
proc mixed data=temp dfbw ;  
  title '4th run -- flexible w linear interaction';  
  class id diet week ;  
  model prot = mixed barley week mixedwk1 barleywk1 / s outpm=run4 ;  
  repeated / subject=id type=sp(pow)(week) local ;  
  contrast 'linear interactions'  
    mixedwk1 1 ,  
    barleywk1 1 ;  
  contrast 'all diet effects'  
    mixed      1 ,  
    barley     1 ,  
    mixedwk1   1 ,  
    barleywk1  1 ;  
run;
```

Some output:

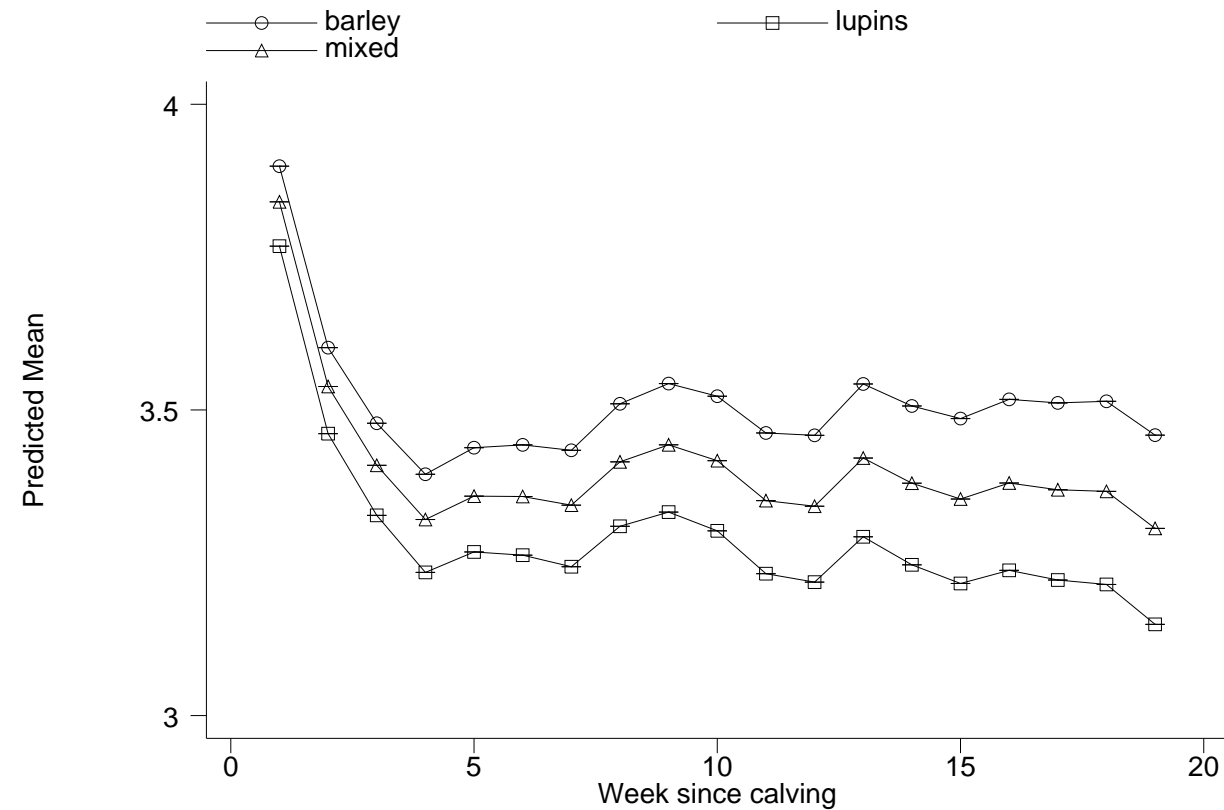
Solution for Fixed Effects

Effect	Week since calving	Estimate	Standard Error	DF	t Value	Pr > t
Intercept		3.1492	0.06653	76	47.33	<.0001
mixed		0.07251	0.07452	76	0.97	0.3336
barley		0.1308	0.07600	76	1.72	0.0892
week	1	0.6187	0.08884	1238	6.96	<.0001
<snip>						
week	19	0
mixedwk1		0.004693	0.006602	1238	0.71	0.4773
barleywk1		0.009928	0.006720	1238	1.48	0.1398

Contrasts

Label	Num DF	Den DF	F Value	Pr > F
linear interactions	2	1238	1.09	0.3360
all diet effects	4	1238	4.55	0.0012

A plot of fitted response profiles shows that the model fits the data pretty well:



This model allows for a linear difference in mean responses among the three groups as a function of (continuous) time.

- Test whether the difference among the three groups **changes** over time:
 - The first contrast statement tests for different **time slopes** across the three diet groups, specifically

$$H_0 : \beta_4 = \beta_5 = 0 \text{ vs. } H_A : \beta_4 \neq 0 \text{ or } \beta_5 \neq 0$$

This yields:

Contrasts				
Label	Num DF	Den DF	F Value	Pr > F
linear interactions	2	1238	1.09	0.3360

- Test whether there are **any** differences among the three groups over time
 - The second contrast statement tests for **any differences** among the three diet groups across time, i.e.,

$$H_0 : \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0 \quad \text{vs.} \quad H_A : \text{at least one of } \beta_2, \dots, \beta_5 \neq 0$$

This yields

all diet effects	4	1238	4.55	0.0012
------------------	---	------	------	--------

- Conclusions:
 - The fitted model suggests that a barley diet yields the highest protein content, followed by a mixed diet, relative to a lupins diet.
 - The hypothesis test for differences among the three groups, **allowing for** (but not requiring) those differences to vary with time, is significant ($F = 4.55$ on 4 df, $P\text{-value} = 0.0012$).

- However, a follow-up hypothesis test does not indicate strong support for the group differences varying with time ($F = 1.09$ on 2 df, P -value = 0.34).
- The above analysis of mean models illustrates an important aspect: **substantive** parameters and **nuisance** parameters
 - The baseline trend is the **non-parametric** or **nuisance** part of the model
 - The difference in time trends form the **parametric** or **substantive** part of the model
- Recommended reading: Fitzmaurice's book (Applied Longitudinal Analysis) §§ 3.4–3.6

Remark: Analysis of Response Profiles

Strengths:

- Allows arbitrary patterns in the mean response over time (no time trend assumed) and arbitrary patterns in the covariance (by unstructured).
- Analysis has a certain robustness since potential risks of bias due to mis-specification of models for mean and covariance are minimal.

Drawbacks:

- Require balanced longitudinal design
- The number of estimated parameters, grows rapidly with the number of measurement occasions.