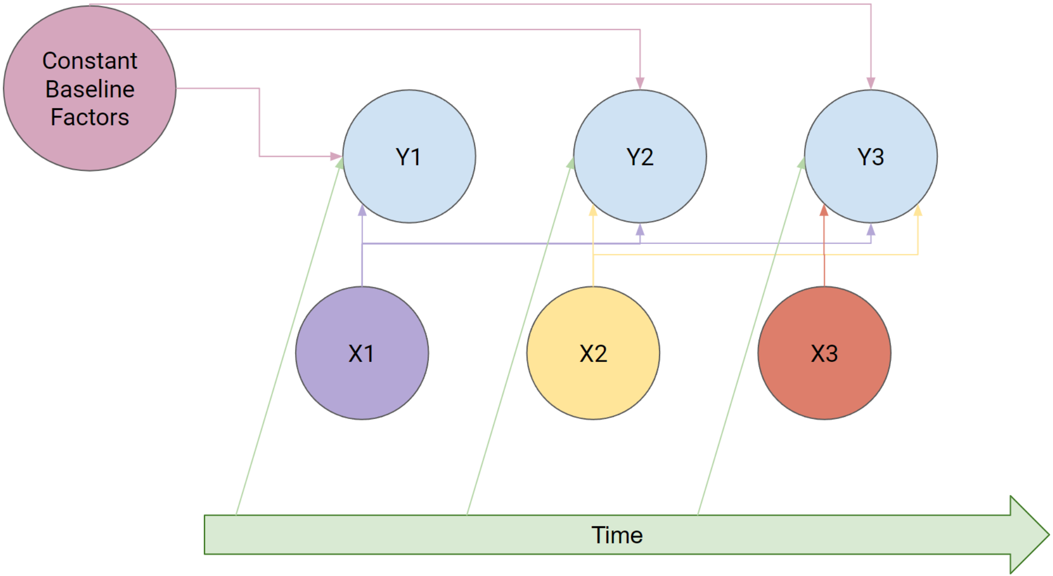
How do we analyze continuous longitudinal data?

What is our goal?



Stated Mathematically

We want to **fit a model** that gives

$$E[Y_{ij}|X_{ij}, t_{ij}],$$

in terms of **interpretable parameters**.

Let's use an example!

ID	Trt	W0	W1	W4	W6	ID	Trt	time	W
1	P	30.8	26.9	25.8	23.8	1	P	1	30.8
2	A	26.5	14.8	19.5	21	2	A	1	26.5
3	A	25.8	23	19.1	23.2	3	A	1	25.8
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
98	A	29.4	22.1	25.3	4.1	98	A	4	4.1
99	A	21.9	7.6	10.8	13	99	A	4	13
100	A	20.7	8.1	25.7	12.3	100	A	4	12.3

- Consider the TLC trial data, in **wide format** (left-hand side) and then in **long format** (right-hand side).

Let's use an example!

ID	Trt	W0	W1	W4	W6	ID	Trt	time	W
1	P	30.8	26.9	25.8	23.8	1	P	1	30.8
2	A	26.5	14.8	19.5	21	2	A	1	26.5
3	A	25.8	23	19.1	23.2	3	A	1	25.8
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
98	A	29.4	22.1	25.3	4.1	98	A	4	4.1
99	A	21.9	7.6	10.8	13	99	A	4	13
100	A	20.7	8.1	25.7	12.3	100	A	4	12.3

- Consider the TLC trial data, in **wide format** (left-hand side) and then in **long format** (right-hand side).
- In the right-hand side we have an outcome (W), with two explanatory factors ($\{\text{Trt}, \text{time}\}$).

Let's use an example!

ID	Trt	W0	W1	W4	W6	ID	Trt	time	W
1	P	30.8	26.9	25.8	23.8	1	P	1	30.8
2	A	26.5	14.8	19.5	21	2	A	1	26.5
3	A	25.8	23	19.1	23.2	3	A	1	25.8
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
98	A	29.4	22.1	25.3	4.1	98	A	4	4.1
99	A	21.9	7.6	10.8	13	99	A	4	13
100	A	20.7	8.1	25.7	12.3	100	A	4	12.3

- Consider the TLC trial data, in **wide format** (left-hand side) and then in **long format** (right-hand side).
- In the right-hand side we have an outcome (W), with two explanatory factors ($\{\text{Trt}, \text{time}\}$).
  - We want $E[W|\text{Trt}, \text{time}]$. **Is this familiar?**

Why can't we just use linear regression?

Using Linear Regression

	Estimate	Std. Error	Pr(>|t|)
(Intercept)	26.540	0.9370175	0.0000000
time2	-13.018	1.3251428	0.0000000
time3	-11.026	1.3251428	0.0000000
time4	-5.778	1.3251428	0.0000166
TreatmentP	-0.268	1.3251428	0.8398322
time2:TreatmentP	11.406	1.8740349	0.0000000
time3:TreatmentP	8.824	1.8740349	0.0000035
time4:TreatmentP	3.152	1.8740349	0.0933783

We can fit the model in R, using `lm`. **Is this valid?**

What does this `lm` imply about our data?

- There is a **linear conditional mean** structure:

$$
\begin{aligned}
E[W_{ij}|\text{Trt}_i, t_j] = \beta_0 &+ \beta_1 \text{Trt}_i + \beta_2 I(t_j = 2) + \beta_3 I(t_j = 3) \\
&+ \beta_4 I(t_j = 4) + \beta_5 \text{Trt}_i I(t_j = 2) + \beta_6 \text{Trt}_i I(t_j = 3) \\
&+ \beta_7 \text{Trt}_i I(t_j = 4).
\end{aligned}
$$

What does this `lm` imply about our data?

- There is a **linear conditional mean** structure:

$$\begin{aligned} E[W_{ij}|\text{Trt}_i, t_j] = \beta_0 &+ \beta_1 \text{Trt}_i + \beta_2 I(t_j = 2) + \beta_3 I(t_j = 3) \\ &+ \beta_4 I(t_j = 4) + \beta_5 \text{Trt}_i I(t_j = 2) + \beta_6 \text{Trt}_i I(t_j = 3) \\ &+ \beta_7 \text{Trt}_i I(t_j = 4). \end{aligned}$$

- There is **constant variance** such that $\text{var}(W_{ij}) = \sigma^2$ for all i, j.

What does this `lm` imply about our data?

- There is a **linear conditional mean** structure:

$$
\begin{aligned}
E[W_{ij}|\text{Trt}_i, t_j] = \beta_0 + \beta_1 \text{Trt}_i + \beta_2 I(t_j = 2) + \beta_3 I(t_j = 3) \\
+ \beta_4 I(t_j = 4) + \beta_5 \text{Trt}_i I(t_j = 2) + \beta_6 \text{Trt}_i I(t_j = 3) \\
+ \beta_7 \text{Trt}_i I(t_j = 4).
\end{aligned}
$$

- There is **constant variance** such that $\text{var}(W_{ij}) = \sigma^2$ for all i, j.
- The values of W_{ij} are **independent**.

What does this `lm` imply about our data?

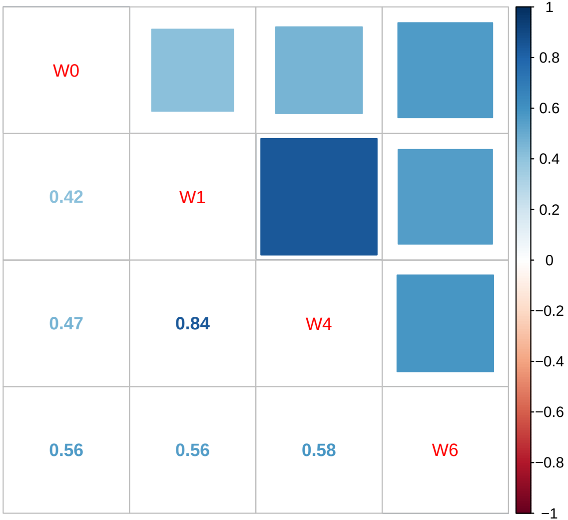
- There is a **linear conditional mean** structure:

$$
\begin{aligned}
E[W_{ij}|\text{Trt}_i, t_j] = \beta_0 &+ \beta_1 \text{Trt}_i + \beta_2 I(t_j = 2) + \beta_3 I(t_j = 3) \\
&+ \beta_4 I(t_j = 4) + \beta_5 \text{Trt}_i I(t_j = 2) + \beta_6 \text{Trt}_i I(t_j = 3) \\
&+ \beta_7 \text{Trt}_i I(t_j = 4).
\end{aligned}
$$

- There is **constant variance** such that $\text{var}(W_{ij}) = \sigma^2$ for all i, j.
- The values of W_{ij} are **independent**.
  - Uh oh...

What makes longitudinal data special?

Longitudinal data are characterized by **correlation** *within* individuals.



The previous `lm` will work **only if** we are willing to assume that the observations are **independent**.

How can we adapt linear regression to allow for this association?

Longitudinal Data as Multivariate Data

- When the data are in **long format**, it appears that the outcomes are univariate.

Longitudinal Data as Multivariate Data

- When the data are in **long format**, it appears that the outcomes are univariate.
- When the data are in **wide format**, we can view the outcome as a vector of outcomes, (e.g., $\mathbf{W} = (W_0, W_1, W_4, W_6)$).

Longitudinal Data as Multivariate Data

- When the data are in **long format**, it appears that the outcomes are univariate.
- When the data are in **wide format**, we can view the outcome as a vector of outcomes, (e.g., $\mathbf{W} = (W_0, W_1, W_4, W_6)$).
- The analysis of longitudinal data is **multivariate analysis**.

Longitudinal Data as Multivariate Data

- When the data are in **long format**, it appears that the outcomes are univariate.
- When the data are in **wide format**, we can view the outcome as a vector of outcomes, (e.g., $\mathbf{W} = (W_0, W_1, W_4, W_6)$).
- The analysis of longitudinal data is **multivariate analysis**.
    - This accounts for the **lack of independence** in the outcomes!

Multivariate Normal

Instead of assuming that $Y_{ij} \sim N(X_{ij}\beta, \sigma^2)$, what if took

$$Y_i \sim \text{MVN}(X_i\beta, \Sigma_i)?$$

Recall that the multivariate normal (MVN) has a density given by

$$f(y;\mu,\Sigma) = \left([2\pi]^k|\Sigma|\right)^{-1/2} \exp\left\{\frac{1}{2}(y-\mu)'\Sigma^{-1}(y-\mu)\right\}.$$

Linear Marginal Models

- In this proposal, we specify a **linear form** for the conditional mean.

Linear Marginal Models

- In this proposal, we specify a **linear form** for the conditional mean.
  - That is, $E[Y_i|X_i] = X_i\beta$, where X_i is a matrix and Y_i is a vector!

Linear Marginal Models

- In this proposal, we specify a **linear form** for the conditional mean.
  - That is, $E[Y_i|X_i] = X_i\beta$, where X_i is a matrix and Y_i is a vector!
- We allow for **correlation** through the individual covariance matrix, Σ_i.

Linear Marginal Models

- In this proposal, we specify a **linear form** for the conditional mean.
  - That is, $E[Y_i|X_i] = X_i\beta$, where X_i is a matrix and Y_i is a vector!
- We allow for **correlation** through the individual covariance matrix, Σ_i.
- We could (theoretically) find the **MLE** under the assumption of multivariate normality.

Covariance Matrix

Recall that $\text{cor}(X, Y) = \frac{\text{cov}(X,Y)}{\sqrt{\text{var}(X)\text{var}(Y)}}$, and so, re-arranging,

$$\text{cov}(X, Y) = \text{cor}(X, Y)\sqrt{\text{var}(X)\text{var}(Y)}.$$

Moreover, recall that a variance/covariance matrix is

$$\text{cov}(Y_i) = \Sigma_i = \begin{bmatrix} \text{var}(Y_{i1}) & \text{cov}(Y_{i1}, Y_{i2}) & \cdots & \text{cov}(Y_{i1}, Y_{ip}) \\ \text{cov}(Y_{i2}, Y_{i1}) & \text{var}(Y_{i2}) & \cdots & \text{cov}(Y_{i2}, Y_{ip}) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(Y_{ip}, Y_{i1}) & \text{cov}(Y_{ip}, Y_{i1}) & \cdots & \text{var}(Y_{ip}) \end{bmatrix}.$$

Covariance Matrix Simplification

If **we assume** that $\text{var}(Y_{ij}) = \sigma^2$ for all i, j, *and* we denote $\text{cor}(Y_{ij}, Y_{i\ell}) = \rho_{j\ell}$ for all i, then note that

$$\text{cov}(Y_{ij}, Y_{i\ell}) = \text{cor}(Y_{ij}, Y_{i\ell})\sqrt{\text{var}(Y_{ij})\text{var}(Y_{i\ell})} = \sigma^2 \rho_{j\ell}.$$

We write

$$\mathbf{R}(\rho) = \begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \cdots & \rho_{pp} \end{bmatrix} = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{12} & 1 & \cdots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1p} & \rho_{2p} & \cdots & 1 \end{bmatrix}.$$

With this notation,

$$\Sigma_i = \sigma^2 \mathbf{R}(\rho).$$

Linear Marginal Models

Under the previous specification we can find the MLE to be

$$\widehat{\beta} = \left(\sum_{i=1}^{n} X_i' \mathbf{R}_i^{-1} X_i\right)^{-1} \sum_{i=1}^{n} X_i' \mathbf{R}_i^{-1} Y_i.$$

For the variance parameter, we get

$$\widehat{\sigma}^2 = \frac{1}{nk} \sum_{i=1}^{n} (Y_i - X_i\beta)' \mathbf{R}_i^{-1} (Y_i - X_i\beta).$$

And we can solve numerically for $\mathbf{R}_i$.

We want to model $E[Y_{ij}|X_i]$ (for some purpose) and so we specify a **multivariate linear model**. By assuming that the variance **is constant across different times**, and we can accommodate the correlation expected within each individual.

The **multivariate normality** assumption gives us a process for computing the MLE, which can produce estimates for the parameters of interest, denoted $\widehat{\beta}$.

What comes next?

Next Steps

- How can we conduct **inference** on the estimated parameters? (Why do we want to?)

Next Steps

- How can we conduct **inference** on the estimated parameters? (Why do we want to?)
- How can we specify **time trends** in the model for the mean?

Next Steps

- How can we conduct **inference** on the estimated parameters? (Why do we want to?)
- How can we specify **time trends** in the model for the mean?
- How can we use this model to answer **scientific questions of interest**?

Next Steps

- How can we conduct **inference** on the estimated parameters? (Why do we want to?)
- How can we specify **time trends** in the model for the mean?
- How can we use this model to answer **scientific questions of interest**?
- What can we do about the **correlation matrix**? (Are there any shortcomings with our assumptions?)

Inference

Asymptotic Normality

It can be shown that, asymptotically,

$$\widehat{\beta} \sim MVN(\beta, \text{var}(\widehat{\beta})),$$

where

$$\text{var}(\widehat{\beta}) = \left[\frac{1}{\sigma^2} \sum_{i=1}^{n} X_i' \mathbf{R}_i^{-1} X_i \right]^{-1},$$

which can be estimated by plugging-in $\widehat{\sigma}^2$ and $\widehat{\rho}$.

This gives us s.e.$(\widehat{\beta}_j) = \left[\widehat{\text{var}}(\widehat{\beta}) \right]^{1/2}_{(j,j)}$.

Inference based on Wald Statistics

As a result,

$$\frac{(\widehat{\beta}_j - \beta_j)}{\text{s.e.}(\widehat{\beta}_j)} \dot{\sim} N(0,1).$$

This can be used to test $H_0 : \beta_j = \beta^*$, or for confidence intervals, **just like with linear regression**!

Note this is equivalent to $\frac{(\widehat{\beta}_j - \beta_j)^2}{var(\widehat{\beta}_j)} \sim \chi^2_1$.

Time Trends

Time as a Covariate

- Generally speaking, we can simply include **time** as a **covariate** in the model.

Time as a Covariate

- Generally speaking, we can simply include **time** as a **covariate** in the model.
- If the data are **balanced** and there are *relatively few* time points, we can include it as a factor.

Time as a Covariate

- Generally speaking, we can simply include **time** as a **covariate** in the model.
- If the data are **balanced** and there are *relatively few* time points, we can include it as a factor.
- If the data are **not balanced** or there are *too many* time points, we can include it as a continuous variable.

Time as a Covariate

- Generally speaking, we can simply include **time** as a **covariate** in the model.
- If the data are **balanced** and there are *relatively few* time points, we can include it as a factor.
- If the data are **not balanced** or there are *too many* time points, we can include it as a continuous variable.
  - We can also include **quadratic** time trends, or **logarithmic** time trends, or any other functional form.

Time as a Covariate

- Generally speaking, we can simply include **time** as a **covariate** in the model.
- If the data are **balanced** and there are *relatively few* time points, we can include it as a factor.
- If the data are **not balanced** or there are *too many* time points, we can include it as a continuous variable.
  - We can also include **quadratic** time trends, or **logarithmic** time trends, or any other functional form.
- We can include time as **calendar time**, **time since baseline**, **index of time point**, **age**, etc.

Time as a Covariate

- Generally speaking, we can simply include **time** as a **covariate** in the model.
- If the data are **balanced** and there are *relatively few* time points, we can include it as a factor.
- If the data are **not balanced** or there are *too many* time points, we can include it as a continuous variable.
  - We can also include **quadratic** time trends, or **logarithmic** time trends, or any other functional form.
- We can include time as **calendar time**, **time since baseline**, **index of time point**, **age**, etc.
  - This will depend on what we have **measured** and what we are **interested** in.

The **choice** of how we include time will be dictated **both** by the *available* data, and by **the scientific questions of inquiry**.

This goes for the **form** it takes in the model, and the **time scale** that we choose to use.

Scientific Questions

Linear Transformations of Parameters

If we want to test a **joint hypothesis** or make a **prediction of outcome**, we are interested in $L\widehat{\beta}$ for some matrix L.

Standard results give that

$$L\widehat{\beta} \dot{\sim} N(L\beta, L\widehat{\text{var}}(\widehat{\beta})L').$$

As a result we get that

$$\left[L\widehat{\beta} - L\beta\right]' \left[L\widehat{\text{var}}(\widehat{\beta})L'\right]^{-1} \left[L\widehat{\beta} - L\beta\right] \dot{\sim} \chi^2_r,$$

where r is the rank of L.

Why do we care?

Many scientific questions of interest can be posed as $H_0 : L\beta = C$.

- Is the effect of placebo any different from the effect of experimental treatment?

Why do we care?

Many scientific questions of interest can be posed as $H_0 : L\beta = C$.

- Is the effect of placebo any different from the effect of experimental treatment?
  - Two equivalent effects, take $L = (0, \cdots, 1, \cdots, -1, \cdots, 0)$ and $C = 0$ ($r = 1$).

Why do we care?

Many scientific questions of interest can be posed as $H_0 : L\beta = C$.

- Is the effect of placebo any different from the effect of experimental treatment?
  - Two equivalent effects, take $L = (0, \cdots, 1, \cdots, -1, \cdots, 0)$ and $C = 0$ ($r = 1$).
- Is the time trend the same between individuals in the placebo and control group?

Why do we care?

Many scientific questions of interest can be posed as $H_0 : L\beta = C$.

- Is the effect of placebo any different from the effect of experimental treatment?
  - Two equivalent effects, take $L = (0, \cdots, 1, \cdots, -1, \cdots, 0)$ and $C = 0$ ($r = 1$).
- Is the time trend the same between individuals in the placebo and control group?
  - Simultaneous effects, take $L = \begin{pmatrix} 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \end{pmatrix}$ and $C = 0$ ($r = 2$).

Why do we care?

Many scientific questions of interest can be posed as $H_0 : L\beta = C$.

- Is the effect of placebo any different from the effect of experimental treatment?
  - Two equivalent effects, take $L = (0, \cdots, 1, \cdots, -1, \cdots, 0)$ and $C = 0$ ($r = 1$).
- Is the time trend the same between individuals in the placebo and control group?
  - Simultaneous effects, take $L = \begin{pmatrix} 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \end{pmatrix}$ and $C = 0$ ($r = 2$).
- What is the expected outcome for a patient with a specific set of variates, X?

Why do we care?

Many scientific questions of interest can be posed as $H_0 : L\beta = C$.

- Is the effect of placebo any different from the effect of experimental treatment?
  - Two equivalent effects, take $L = (0, \cdots, 1, \cdots, -1, \cdots, 0)$ and $C = 0$ ($r = 1$).
- Is the time trend the same between individuals in the placebo and control group?
  - Simultaneous effects, take $L = \begin{pmatrix} 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \end{pmatrix}$ and $C = 0$ ($r = 2$).
- What is the expected outcome for a patient with a specific set of variates, X?
  - Take $L = X$, and do not run it as a hypothesis, but instead build a confidence interval (for instance).

Correlation

Possible Shortcomings of our Method

- It assumes constant variance across observations.

Possible Shortcomings of our Method

- It assumes constant variance across observations.
- It requires $\frac{k(k-1)}{2}$ different correlation paramaters.

Possible Shortcomings of our Method

- It assumes constant variance across observations.
- It requires $\frac{k(k-1)}{2}$ different correlation paramaters.
- It requires **balanced** data.

Correlation Pattern Matrices

- Our current assumption is **unstructured**. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = \rho_{j\ell}$ without restriction.

Correlation Pattern Matrices

- Our current assumption is **unstructured**. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = \rho_{j\ell}$ without restriction.
- We could assume **independence**. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = 0$ for all $j \neq \ell$.

Correlation Pattern Matrices

- Our current assumption is **unstructured**. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = \rho_{j\ell}$ without restriction.
- We could assume **independence**. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = 0$ for all $j \neq \ell$.
  - This would become our previously fit linear model!

Correlation Pattern Matrices

- Our current assumption is **unstructured**. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = \rho_{j\ell}$ without restriction.
- We could assume **independence**. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = 0$ for all $j \neq \ell$.
  - This would become our previously fit linear model!
- We could assume that the observations are **exchangeable**. That is $\text{cor}(Y_{ij}, Y_{i\ell}) = \rho$, for $j \neq \ell$.

Correlation Pattern Matrices

- Our current assumption is **unstructured**. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = \rho_{j\ell}$ without restriction.
- We could assume **independence**. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = 0$ for all $j \neq \ell$.
  - This would become our previously fit linear model!
- We could assume that the observations are **exchangeable**. That is $\text{cor}(Y_{ij}, Y_{i\ell}) = \rho$, for $j \neq \ell$.
  - This is also referred to as **uniform** or **compound symmetry**.

Correlation Pattern Matrices

- Our current assumption is **unstructured**. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = \rho_{j\ell}$ without restriction.
- We could assume **independence**. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = 0$ for all $j \neq \ell$.
  - This would become our previously fit linear model!
- We could assume that the observations are **exchangeable**. That is $\text{cor}(Y_{ij}, Y_{i\ell}) = \rho$, for $j \neq \ell$.
  - This is also referred to as **uniform** or **compound symmetry**.
- We could assume an **autoregressive** structure. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = \rho^{|j-\ell|}$.

Correlation Pattern Matrices

- Our current assumption is **unstructured**. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = \rho_{j\ell}$ without restriction.
- We could assume **independence**. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = 0$ for all $j \neq \ell$.
  - This would become our previously fit linear model!
- We could assume that the observations are **exchangeable**. That is $\text{cor}(Y_{ij}, Y_{i\ell}) = \rho$, for $j \neq \ell$.
  - This is also referred to as **uniform** or **compound symmetry**.
- We could assume an **autoregressive** structure. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = \rho^{|j-\ell|}$.
- We could assume an **exponential** structure. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = \exp(-\rho|t_j - t_\ell|)$.

Correlation Pattern Matrices

- Our current assumption is **unstructured**. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = \rho_{j\ell}$ without restriction.
- We could assume **independence**. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = 0$ for all $j \neq \ell$.
  - This would become our previously fit linear model!
- We could assume that the observations are **exchangeable**. That is $\text{cor}(Y_{ij}, Y_{i\ell}) = \rho$, for $j \neq \ell$.
  - This is also referred to as **uniform** or **compound symmetry**.
- We could assume an **autoregressive** structure. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = \rho^{|j-\ell|}$.
- We could assume an **exponential** structure. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = \exp(-\rho|t_j - t_\ell|)$.
  - This pattern can handle **unequal spaced times!**.

Correlation Pattern Matrices

- Our current assumption is **unstructured**. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = \rho_{j\ell}$ without restriction.
- We could assume **independence**. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = 0$ for all $j \neq \ell$.
  - This would become our previously fit linear model!
- We could assume that the observations are **exchangeable**. That is $\text{cor}(Y_{ij}, Y_{i\ell}) = \rho$, for $j \neq \ell$.
  - This is also referred to as **uniform** or **compound symmetry**.
- We could assume an **autoregressive** structure. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = \rho^{|j-\ell|}$.
- We could assume an **exponential** structure. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = \exp(-\rho|t_j - t_\ell|)$.
  - This pattern can handle **unequal spaced times!**.
- We can expand the exponential to the **Gaussian**. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = \exp(-\rho|t_j - t_\ell|^2)$.

Correlation Pattern Matrices

- Our current assumption is **unstructured**. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = \rho_{j\ell}$ without restriction.
- We could assume **independence**. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = 0$ for all $j \neq \ell$.
  - This would become our previously fit linear model!
- We could assume that the observations are **exchangeable**. That is $\text{cor}(Y_{ij}, Y_{i\ell}) = \rho$, for $j \neq \ell$.
  - This is also referred to as **uniform** or **compound symmetry**.
- We could assume an **autoregressive** structure. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = \rho^{|j-\ell|}$.
- We could assume an **exponential** structure. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = \exp(-\rho|t_j - t_\ell|)$.
  - This pattern can handle **unequal spaced times!**.
- We can expand the exponential to the **Gaussian**. That is, $\text{cor}(Y_{ij}, Y_{i\ell}) = \exp(-\rho|t_j - t_\ell|^2)$.
  - Produces more rapid decay!

Relaxing Constant Variance

All of the previous pattern matrices replace $\mathbf{R}_i(\rho)$ with a parsimonious model. What about the **constant variance?**

If we define

$$A_i = \begin{bmatrix} \mathrm{var}(Y_{i1}) & 0 & 0 & \cdots & 0 \\ 0 & \mathrm{var}(Y_{i2}) & 0 & \cdots & 0 \\ 0 & 0 & \mathrm{Var}(Y_{i3}) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \mathrm{var}(Y_{ik_i}) \end{bmatrix},$$

then we can write

$$\Sigma_i = A_i^{1/2}\mathbf{R}_i(\rho)A_i^{1/2}.$$

Notation

Since it is **unlikely** that we will *correctly* specify $\mathbf{R}_i(\rho)$, we typically denote

$$V_i = A_i^{1/2} \mathbf{R}_i(\rho) A_i^{1/2},$$

and refer to this as **the working covariance matrix**, to distinguish it from the true, Σ_i.

The Punchline

Any correlation structure you can imagine, can be used in theory. **In practice** the previously listed ones will work quite well!

These structures can be made to vary based on time, or on other covariates as well! *(We will see this in R code!)*

Is that all that we need?

Limitations

- We require the complete specification of the **mean model** and the **covariance structure**.

Limitations

- We require the complete specification of the **mean model** and the **covariance structure**.
- If the covariance structure is *misspecified* then inference regarding $\widehat{\beta}$ is **invalid**.

Limitations

- We require the complete specification of the **mean model** and the **covariance structure**.
- If the covariance structure is *misspecified* then inference regarding $\widehat{\beta}$ is **invalid**.
- Our results so far have relied on the assumption of **normality**, which is not ideal!

Limitations

- We require the complete specification of the **mean model** and the **covariance structure**.
- If the covariance structure is *misspecified* then inference regarding $\widehat{\beta}$ is **invalid**.
- Our results so far have relied on the assumption of **normality**, which is not ideal!
- We have not handled **non-linear** outcomes.

What comes next?

Despite these shortcomings, this serves as an incredibly useful basis for **analyzing continuous longitudinal data**.

Next we will explore these **theoretical properties**, deriving the results that were explored here, and providing **an application in R** of these concepts. The application will further explore how we can ask scientific questions of these models, test hypothesis, and interpret the parameters!

From there, we will begin to **tackle these limitations**.

Summary

- We cannot use linear regression to analyze continuous longitudinal data **unless** we assume independence.

Summary

- We cannot use linear regression to analyze continuous longitudinal data **unless** we assume independence.
- Instead, we can generalize this by assuming **multivariate normality** and imposing a specific **correlation structure**.

Summary

- We cannot use linear regression to analyze continuous longitudinal data **unless** we assume independence.
- Instead, we can generalize this by assuming **multivariate normality** and imposing a specific **correlation structure**.
- This gives us asymptotic results for **inference** and hypothesis testing.

Summary

- We cannot use linear regression to analyze continuous longitudinal data **unless** we assume independence.
- Instead, we can generalize this by assuming **multivariate normality** and imposing a specific **correlation structure**.
- This gives us asymptotic results for **inference** and hypothesis testing.
- Imposing **specific correlation structures** allow for us to explore more parsimonious models.