Shortcomings of Marginal Linear Models

The New York Times

Scientists Predict Omicron Will Peak in the U.S. in Mid-January But Still May Overwhelm Hospitals

Azeen Ghorayshi

January 1, 2022 · 4 min read

Can linear marginal models make these forecasts?

Generalized Marginal Models for Longitudinal Data

Where are we right now?

So far we have **expanded** linear regression models into **linear marginal models**, allowing for an analysis of **continuous, longitudinal data**.

How do we handle **other outcomes**?

Linear Marginal Models ARE TO Linear Regression Models

WHAT

__________?__________ ARE TO Generalized Linear Models

Generalized Marginal Models

Generalized Marginal Models

- Recall that for **linear marginal models** we specified a conditional mean, a variance structure, and pairwise correlations.

Generalized Marginal Models

- Recall that for **linear marginal models** we specified a conditional mean, a variance structure, and pairwise correlations.
- Recall that for a **generalized linear model** we need to specify a linear predictor, a link function, and a distribution†.

Generalized Marginal Models

- Recall that for **linear marginal models** we specified a conditional mean, a variance structure, and pairwise correlations.
- Recall that for a **generalized linear model** we need to specify a linear predictor, a link function, and a distribution†.
- What if we specify:

Generalized Marginal Models

- Recall that for **linear marginal models** we specified a conditional mean, a variance structure, and pairwise correlations.
- Recall that for a **generalized linear model** we need to specify a linear predictor, a link function, and a distribution†.
- What if we specify:
  1. A linear predictor, $\eta_{ij} = X_{ij}\beta$, with a link function, $g(E[Y_{ij}|X_i]) = \eta_{ij}$.

Generalized Marginal Models

- Recall that for **linear marginal models** we specified a conditional mean, a variance structure, and pairwise correlations.
- Recall that for a **generalized linear model** we need to specify a linear predictor, a link function, and a distribution†.
- What if we specify:
  1. A linear predictor, $\eta_{ij} = X_{ij}\beta$, with a link function, $g(E[Y_{ij}|X_i]) = \eta_{ij}$.
  2. A variance structure, as $\text{var}(Y_{ij}) = \phi V(\mu_{ij})$, for some function $V(\cdot)$.

Generalized Marginal Models

- Recall that for **linear marginal models** we specified a conditional mean, a variance structure, and pairwise correlations.
- Recall that for a **generalized linear model** we need to specify a linear predictor, a link function, and a distribution†.
- What if we specify:
  1. A linear predictor, $\eta_{ij} = X_{ij}\beta$, with a link function, $g(E[Y_{ij}|X_i]) = \eta_{ij}$.
  2. A variance structure, as $\text{var}(Y_{ij}) = \phi V(\mu_{ij})$, for some function $V(\cdot)$.
  3. A pairwise correlation matrix, as $\mathbf{R}_i(\rho)$.

Generalized Marginal Models

- Recall that for **linear marginal models** we specified a conditional mean, a variance structure, and pairwise correlations.
- Recall that for a **generalized linear model** we need to specify a linear predictor, a link function, and a distribution†.
- What if we specify:
  1. A linear predictor, $\eta_{ij} = X_{ij}\beta$, with a link function, $g(E[Y_{ij}|X_i]) = \eta_{ij}$.
  2. A variance structure, as $\text{var}(Y_{ij}) = \phi V(\mu_{ij})$, for some function $V(\cdot)$.
  3. A pairwise correlation matrix, as $\mathbf{R}_i(\rho)$.
- This provides the *natural extension* of a GLM to longitudinal data!

Example 1: Continuous Data

If Y_{ij} are continuous, then we can take

$$\begin{aligned} E[Y_{ij}|X_{ij}] &= X_{ij}\beta \\ \text{var}(Y_{ij}) &= \phi V(\mu_{ij}) = \phi \\ \text{cor}(Y_{ij}, Y_{i\ell}) &= \rho_{j\ell}, j \neq \ell. \end{aligned}$$

This provides us with *effectively* the same as the linear marginal model, except **without assuming normality**.

Example 2: Binary Data

If Y_{ij} are binary indicators, then we can take

$$\begin{aligned} P[Y_{ij} = 1|X_{ij}] = E[Y_{ij}|X_{ij}] = \text{expit}(X_{ij}\beta) &= \pi_{ij} \\ \text{var}(Y_{ij}) = \phi V(\mu_{ij}) &= \phi\pi_{ij}(1 - \pi_{ij}) \\ \text{cor}(Y_{ij}, Y_{i\ell}) &= \rho_{j\ell}, j \neq \ell. \end{aligned}$$

This is a **natural generalization** of logistic regression, where arbitrary correlations are permitted!

Example 3: Count Data

If Y_{ij} are counts, then we can take

$$\begin{aligned} E[Y_{ij}|X_{ij}] &= \exp(X_{ij}\beta) = \lambda_{ij} \\ \text{var}(Y_{ij}) &= \phi V(\mu_{ij}) = \phi\lambda_{ij} \\ \text{cor}(Y_{ij}, Y_{i\ell}) &= \rho_{j\ell}, j \neq \ell. \end{aligned}$$

This is a **natural generalization** of log-linear regression, where arbitrary correlations are permitted!

Do these models *actually* work?

While this is *compelling* in terms of model specification, we have not discussed how to **estimate** the relevant parameters.

We **can** estimate the parameters, and they will be **nicely behaved** and useful for inference.

This requires a slightly deeper dive into *M-estimation*.

Summary

- Generalized Marginal Models extend generalized linear models to longitudinal data.

Summary

- Generalized Marginal Models extend generalized linear models to longitudinal data.
- We specify a **linear predictor**, and a **link function**, in addition to a **variance function**, and **pairwise associations**.

Summary

- Generalized Marginal Models extend generalized linear models to longitudinal data.
- We specify a **linear predictor**, and a **link function**, in addition to a **variance function**, and **pairwise associations**.
- Any GLM extends *naturally* to the case of longitudinal data, with whatever correlation patterns we want to accommodate.

How can we estimate these models?