

USING THE DELTA METHOD WITH PROC MIXED TO GENERATE MEANS AND CONFIDENCE INTERVALS FROM A LINEAR MIXED MODEL ON THE ORIGINAL SCALE, WHEN THE ANALYSIS IS DONE ON THE LOG SCALE

Brandy R. Sinco, MS, University of Michigan, Ann Arbor, MI
 Edith Kieffer, MPH, PhD, University of Michigan, Ann Arbor, MI
 Michael S. Spencer, MSW, PhD, University of Michigan, Ann Arbor, MI

ABSTRACT

Background. SAS® Proc Mixed does not provide an option to report log-transformed analyses on the original scale.

Objective. Implement the delta method to report the means and confidence intervals of difference scores from log-transformed longitudinal data, so that results can be reported on the original scale of the outcome.

Demonstrate how this methodology can be implemented on output from lmm (linear mixed model) from SAS Proc Mixed.

Compare results from the delta method to reporting percent change, another common approach to difference scores from log-transformed data.

Methods. We focused on a lmm for data with two treatment groups and two time points. By using the delta method and the exponential function as the inverse, a formula was derived to obtain estimates for the differences between the means between two time points, along with confidence intervals.

As a demonstration, the delta method was applied to nutrition outcome data from the Healthy Moms study, a randomized clinical trial that showed improved dietary outcomes among 275 Latina women.

Results. Using the delta method, the total fat outcome can be reported as decreasing between baseline and follow-up by -12.4g (-17.8g, -7.1g) within the treatment group, with an intervention effect of -9.1g (-17.0g, -1.2g) relative to the control group. Without the delta method, the results would be -16.7% (-23.0%, -10.0%) within the treatment group, with an intervention effect of -12.9% (-22.0%, -2.7%).

Conclusion. The delta method provides a vehicle for reporting intervention effects in units that are more meaningful in public health and clinical practice.

Outline

Background	Pages 1 - 2
Introduction – Linear Mixed Model with Log-Transformed Data	Page 2
Why Reporting Outcomes as % Change	Page 2
Univariate Delta Method.	Pages 2 - 3
Multivariate Delta Method	Page 3
Application of the Multivariate Delta Method to Mean Estimates from a Linear Mixed Model	Page 3 - 4
Important Findings – Simple Formula for Variances When the Only Covariates are Binary Indicators for Treatment Group and Time Point	Pages 4 - 5
Delta Method With Adjusted Estimates	Page 5
Confidence Intervals for Difference Scores	Pages 5 - 6
SAS Code – Procs Mixed and IML	Pages 6 - 7
Views of LSMEANS and COVB Before And After Processing for Proc IML	Page 7 – 9
Examples Demonstrating the Advantage to Using the Delta Method Over Reporting % Change	Pages 9 - 10
Conclusions	Page 10
References	Page 10

Background

Reporting intervention effects in units that are meaningful in social work, public health, and clinical practice is important for improving the understanding and use of study results. Common statistical software, such as SAS Proc Mixed, does not provide an option to report log transformed analyses on the original, and more interpretable, scale.

For example, consider a weight loss study, in which the data was log transformed and the results were reported with percent change. If the treatment group lost an average of 10% of their body weight and the control group lost 1%, the intervention effect would not necessarily sum to 9%. However, by using the delta method, if the treatment group lost 20 pounds and the control group lost 2 pounds, the intervention effect would be 18 pounds, which would be easier to interpret. In the section below on reporting outcomes as percent change, there is a proof and an example of percent changes in the treatment and intervention groups not always adding to the percent change of the intervention effect. This demonstrates the advantage of the delta method in making the intervention effects easier to interpret in physical units.

Introduction to the Linear Mixed Model¹ With Log-Transformed Data

- Let X_{ijk} = original variable; (i, j, k) = (randomization, time point, subject).
- $Y_{ijk} = \ln(X_{ijk})$.
- $i = 0$ for control and 1 for treatment.
- $j = 1$ for pre-intervention and 2 for post-intervention.
- $k = k^{\text{th}}$ subject.
- Let $R = 0$ for control and 1 for treatment.
- Let $T = 0$ for pre-intervention and 1 for post-intervention.
- Linear Mixed Model (LMM) for $X_{ijk} = \alpha_0 + \alpha_1R + \alpha_2T + \alpha_3RT + \varepsilon_{ijk}$, where ε_{ijk} = error term and $\varepsilon_{ijk} \sim N(0, \Sigma)$.
- Estimated Mean: $E(X_{ij}) = \alpha_0 + \alpha_1R + \alpha_2T + \alpha_3RT$.
- The means for the control group are α_0 at pre-intervention and $(\alpha_0 + \alpha_2)$ at post-intervention.
- The means for the treatment group are $(\alpha_0 + \alpha_1)$ at pre-intervention and $(\alpha_0 + \alpha_1 + \alpha_2 + \alpha_3)$ at post-intervention.
- The change scores are α_2 for the control group and $(\alpha_2 + \alpha_3)$ for the treatment group.
- The intervention effect is the difference in change scores for the treatment and control groups = α_3 .

Why Reporting Outcomes as % Change²

Next, consider the LMM for Y, the log transform of X.

- Linear Mixed Model (LMM) $Y_{ijk} = \beta_0 + \beta_1R + \beta_2T + \beta_3RT + \varepsilon'_{ijk}$, where ε'_{ijk} is the error term.
- Estimated Mean: $E(Y_{ij}) = \beta_0 + \beta_1R + \beta_2T + \beta_3RT$.
- Post – Pre Mean for Control: $\ln(X_{02}) - \ln(X_{01}) = \beta_2$; $X_{02}/X_{01} = \exp(\beta_2)$.
- % Change Post – Pre Mean for Control = $100\%(\exp(\beta_2) - 1)$.
- Post – Pre Mean Treatment: $\ln(X_{12}) - \ln(X_{11}) = \beta_2 + \beta_3$; $X_{12}/X_{11} = \exp(\beta_2 + \beta_3)$.
- % Change Post – Pre Mean for Treatment = $100\%(\exp(\beta_2 + \beta_3) - 1)$.
- Intervention Effect = $\ln(X_{12}) - \ln(X_{11}) - [\ln(X_{02}) - \ln(X_{01})] = \beta_3$.
- Intervention Effect = $\ln(X_{12}/X_{11}) - \ln(X_{02}/X_{01}) = \beta_3$.
- $100\%[(X_{12}/X_{11}) - (X_{02}/X_{01})]/(X_{02}/X_{01}) = 100\%(\exp(\beta_3) - 1)$.

Important Note: $100\% \times (\exp(\beta_3) - 1)$ is ratio of percent changes, not the difference in percent changes. That's why (percent change in the treatment group) - (percent change in the control group) is not always equal to the intervention effect, as we see from Table 1 below. While the differences in mean changes will sum to the intervention effect when using the delta method, this property does not always hold for percent change, although the percent changes; $-12.9\% \neq -16.7\% - (-4.4\%)$.

Table 1: Change in Fat Consumption Using Percent Change and Delta Method

Group	Total Fat Change (%)	Total Fat Change (g), Delta Method
Control	-4.40%	-3.38 g
Treatment	-16.70%	-12.45 g
Intervention Effect	-12.90%	-9.07 g

The Univariate Delta Method³

- Let $Y \sim \text{Normal}(\mu, \sigma^2)$, $\mu, \sigma^2 \neq 0$; n = sample size.
- Let $g(Y)$ be a differentiable function of Y with non-zero first derivative.
- Then, a first-order Taylor series for $g(Y) = g(\mu) + g'(\mu)(Y - \mu)$.
- Mean of $g(Y) \approx g(\mu)$ and Variance of $g(Y) \approx (g'(\mu))^2 \sigma^2$.

- Delta Method Theorem: $\lim_{n \rightarrow \infty} \sqrt{n} (g(Y) - g(\mu)) \xrightarrow{D} \text{Normal}(0, \sigma^2 (g'(\mu))^2)$.
- I.E., asymptotic distribution of $g(Y) = \text{Normal}(g(\mu), (g'(\mu))^2 \sigma^2)$.
- **Example: Let $Y \sim \text{Normal}(\mu, \sigma^2)$. Let $W = g(Y) = e^Y$.**
- $g'(\mu) = e^\mu$.
- Using the delta method, mean of $W = g(\mu) = e^\mu$ and
- Variance of $W = (g'(\mu))^2 \sigma^2 = e^{2\mu} \sigma^2$; Standard deviation of $W = e^\mu \sigma$.
- Asymptotic distribution of $W = \text{Normal}(e^\mu, e^{2\mu} \sigma^2)$.

The Multivariate Delta Method³

- Let Y be a multivariate vector of m normal variables, $Y = [Y_1 Y_2 \dots Y_m]$.
- $Y \sim N(\mu, \Sigma)$, where Σ is a $m \times m$ covariance matrix.
- Let $g(Y)$ be a differentiable function of Y with non-zero first derivative.
- The multivariate delta method states that if $\lim_{n \rightarrow \infty} \sqrt{n} (Y - \mu) \xrightarrow{D} N(0, \Sigma)$, then

$$\lim_{n \rightarrow \infty} \sqrt{n} (g(Y) - g(\mu)) \xrightarrow{D} N(0, J_g(\mu) \Sigma J_g^T(\mu)).$$

- Where $J_g(\mu)$ is the Jacobian matrix, evaluated at $Y = \mu$.

$$J_g(\mu) = \begin{bmatrix} \frac{\partial g_1(Y)}{\partial Y_1} & \frac{\partial g_1(Y)}{\partial Y_2} & \dots & \frac{\partial g_1(Y)}{\partial Y_m} \\ \frac{\partial g_2(Y)}{\partial Y_1} & \frac{\partial g_2(Y)}{\partial Y_2} & \dots & \frac{\partial g_2(Y)}{\partial Y_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_m(Y)}{\partial Y_1} & \frac{\partial g_m(Y)}{\partial Y_2} & \dots & \frac{\partial g_m(Y)}{\partial Y_m} \end{bmatrix} \text{ evaluated at } Y = \mu.$$

Application of the Multivariate Delta Method to Mean Estimates from a Linear Mixed Model

- Recall Estimated Mean from LMM: $Y_{ij} = \beta_0 + \beta_1 R + \beta_2 T + \beta_3 RT$.
- The β 's are assumed to be multivariate normal, with covariance matrix Σ_β .
- Let Y_{01} = mean of control group at time 1 on log scale = β_0 .
- Let $W_1 = Y_{01}$ on the original scale; Let $g_1(\beta) = \exp(\beta_0)$.
- $W_1 = g_1(\beta) = \exp(\beta_0)$. Note that $\partial g_1(\beta) / \partial \beta_0 = g_1(\beta) = \exp(\beta_0)$ and $\partial g_1(\beta) / \partial \beta_i = 0$ if $i > 0$.
-
- Let Y_{02} = mean of control group at time 2 on log scale = $\beta_0 + \beta_2$;
- Let $W_2 = g_2(\beta) = \exp(\beta_0 + \beta_2)$. Note that $\partial g_2(\beta) / \partial \beta_0 = g_2(\beta)$, $\partial g_2(\beta) / \partial \beta_2 = g_2(\beta)$, and $\partial g_2(\beta) / \partial \beta_i = 0$ for $i \neq 0$ and $i \neq 2$.
-
- Same property holds for g_3 and g_4 derivatives with respect to the β 's.
- Let Y_{11} = mean of treatment group at time 1 on log scale = $\beta_0 + \beta_1$;
- Let $W_3 = g_3(\beta) = \exp(\beta_0 + \beta_1)$.
- $\partial g_3(\beta) / \partial \beta_0 = g_3(\beta)$, $\partial g_3(\beta) / \partial \beta_1 = g_3(\beta)$, and $\partial g_3(\beta) / \partial \beta_i = 0$ for $i > 1$.
-
- Let Y_{12} = mean of treatment group at time 2 on log scale = $\beta_0 + \beta_1 + \beta_2 + \beta_3$;
- Let $W_4 = g_4(\beta) = \exp(\beta_0 + \beta_1 + \beta_2 + \beta_3)$.
- $\partial g_4(\beta) / \partial \beta_i = g_4(\beta)$, $0 \leq i \leq 3$.

Jacobian Matrix for $g(\beta)$.

$$J_g(\beta) = \begin{bmatrix} \exp(\beta_0) & 0 & 0 & 0 \\ \exp(\beta_0 + \beta_2) & 0 & \exp(\beta_0 + \beta_2) & 0 \\ \exp(\beta_0 + \beta_1) & \exp(\beta_0 + \beta_1) & 0 & 0 \\ \exp(\beta_0 + \beta_1 + \beta_2 + \beta_3) & \exp(\beta_0 + \beta_1 + \beta_2 + \beta_3) & \exp(\beta_0 + \beta_1 + \beta_2 + \beta_3) & \exp(\beta_0 + \beta_1 + \beta_2 + \beta_3) \end{bmatrix}$$

$$\text{In simpler form, } J_g(\beta) = \begin{bmatrix} W_1 & 0 & 0 & 0 \\ W_2 & 0 & W_2 & 0 \\ W_3 & W_3 & 0 & 0 \\ W_4 & W_4 & W_4 & W_4 \end{bmatrix}.$$

$$\text{Define the covariance matrix for the } \beta\text{'s as } \Sigma_\beta = \begin{bmatrix} \sigma_{\beta 00}^2 & \sigma_{\beta 01} & \sigma_{\beta 02} & \sigma_{\beta 03} \\ \sigma_{\beta 01} & \sigma_{\beta 11}^2 & \sigma_{\beta 12} & \sigma_{\beta 13} \\ \sigma_{\beta 02} & \sigma_{\beta 12} & \sigma_{\beta 22}^2 & \sigma_{\beta 23} \\ \sigma_{\beta 03} & \sigma_{\beta 13} & \sigma_{\beta 23} & \sigma_{\beta 33}^2 \end{bmatrix}.$$

The covariance matrix for W (estimates for each group at each time point) will be $\Sigma_W = J_g(\beta) \times \Sigma_\beta \times (J_g(\beta))^T$.

To multiply the above three matrices, the authors suggest using a symbolic matrix calculator, such as

http://wims.unice.fr/~wims/en_tool-linear-matrix.html

Because SAS Proc IML will multiply numerical matrices, but not matrices with characters, the above symbolic matrix calculator is recommended.

Important Findings – Simple Formula for Variances When the Only Covariates are Binary Indicators for Treatment Group and Time Point.

First, when the estimator for the mean has only covariates for treatment group and time point that are either 1 or 0, such as $Y_{ij} = \beta_0 + \beta_1 R + \beta_2 T + \beta_3 RT$, the variances of the point estimates have a simple form.

- $\text{Var}(W_1) = \text{Var}(\text{control group at baseline}) = \exp(2\beta_0)\sigma_{\beta 00}^2$;
- $\text{standard error}(W_1) = \exp(\beta_0)\sigma_{\beta 00} = W_1 \times \text{Std error}(\text{estimate of } Y_{01} \text{ on log scale}).$

- Standard errors of W_2, W_3, W_4 have also the same format as in the univariate example with the delta method earlier in this paper.

- $\text{Var}(W_2)$ (control group at follow-up) = $\exp(2(\beta_0 + \beta_2))[\sigma_{\beta 00}^2 + \sigma_{\beta 22}^2 + 2\sigma_{\beta 02}]$.
- $\text{Standard error}(W_2) = \exp(\beta_0 + \beta_2)[\sigma_{\beta 00}^2 + \sigma_{\beta 22}^2 + 2\sigma_{\beta 02}]^{1/2} =$
- $W_2 \times \text{Std err}(\text{estimate of } Y_{02} \text{ log scale}).$

- $\text{Var}(W_3)$ (treatment group at baseline) = $\exp(2(\beta_0 + \beta_1))[\sigma_{\beta 00}^2 + \sigma_{\beta 11}^2 + 2\sigma_{\beta 01}]$.
- $\text{Standard error}(W_3) = \exp(\beta_0 + \beta_1)[\sigma_{\beta 00}^2 + \sigma_{\beta 11}^2 + 2\sigma_{\beta 01}]^{1/2} =$
- $W_3 \times \text{Std err}(\text{estimate of } Y_{11} \text{ log scale}).$

- $\text{Var}(W_4)$ (treatment group at follow-up) =
- $\exp(2(\beta_0 + \beta_1 + \beta_2 + \beta_3))[\sigma_{\beta 00}^2 + \sigma_{\beta 11}^2 + \sigma_{\beta 22}^2 + \sigma_{\beta 33}^2 + 2\sigma_{\beta 01} + 2\sigma_{\beta 02} + 2\sigma_{\beta 03} + 2\sigma_{\beta 12} + 2\sigma_{\beta 13} + 2\sigma_{\beta 23}]$.
- $\text{Standard error}(W_4) = \exp(\beta_0 + \beta_1 + \beta_2 + \beta_3)[\text{Var}(\beta_0 + \beta_1 + \beta_2 + \beta_3)]^{1/2} =$
- $W_4 \times \text{Std err}(\text{estimate of } Y_{12} \text{ log scale}).$

Second, if randomization to the two treatment groups was successful, then (W_1, W_2) will be independent from (W_3, W_4) and therefore $\text{Cov}(W_1, W_3) = \text{Cov}(W_1, W_4) = \text{Cov}(W_2, W_3) = \text{Cov}(W_2, W_4) = 0$. The Σ_W matrix will have the following form:

$$\Sigma_W = \begin{bmatrix} \sigma_{W11}^2 & \sigma_{W12} & 0 & 0 \\ \sigma_{W12} & \sigma_{W22}^2 & 0 & 0 \\ 0 & 0 & \sigma_{W33}^2 & \sigma_{W34} \\ 0 & 0 & \sigma_{W34} & \sigma_{W44}^2 \end{bmatrix}$$

Below is a picture of the Σ_W matrix from nutrient analysis from the “Healthy Mothers on the Move” project⁴.

W1	W2	W3	W4
14.17	7.22	0	0
7.22	13.65	0	0
0	0	6.28	5.07
0	0	5.07	7.02

Delta Method With Adjusted Estimates

- If adjusting for covariates, such as age (assuming average age of 50), $Y_{ij} = \beta_0 + \beta_1 R + \beta_2 T + \beta_3 RT + 50\beta_{Age}$.
- Use same procedure, check whether Σ_W matrix has a form with simple formulas, such as $\text{standard error}(W_1) = \exp(\beta_0)\sigma_{\beta_{00}}$.
- For the above example, I found the same pattern of formulas for the estimates by group and time point.
- Let Y_{01} = mean of control group at time 1 on log scale = $\beta_0 + 50\beta_{Age}$.
- Let $W_1 = Y_{01}$ on the original scale; Let $g_1(\beta) = \exp(\beta_0 + 50\beta_{Age})$.
- $W_1 = g_1(\beta) = \exp(\beta_0 + 50\beta_{Age})$.
- $\partial g_1(\beta)/\partial \beta_0 = g_1(\beta) = \exp(\beta_0 + 50\beta_{Age}) = W_1$, $\partial g_1(\beta)/\partial \beta_i = 0$ if $i > 0$ and $i \neq \text{Age}$.
- $\partial g_1(\beta)/\partial \beta_{Age} = 50g_1(\beta) = 50\exp(\beta_0 + 50\beta_{Age}) = 50W_1$.
- Let Y_{02} = mean of control group at time 2 on log scale = $\beta_0 + \beta_2 + 50\beta_{Age}$;
- Let $W_2 = g_2(\beta) = \exp(\beta_0 + \beta_2 + 50\beta_{Age})$.
- $\partial g_2(\beta)/\partial \beta_0 = g_2(\beta)$, $\partial g_2(\beta)/\partial \beta_2 = g_2(\beta)$, and $\partial g_2(\beta)/\partial \beta_i = 0$ for i not in $(0, 2, \text{Age})$.
- $\partial g_2(\beta)/\partial \beta_{Age} = 50g_2(\beta) = 50\exp(\beta_0 + \beta_2 + 50\beta_{Age}) = 50W_2$.
- Same properties holds for g_3 and g_4 derivatives with respect to the β 's.
- Let Y_{11} = mean of treatment group at time 1 on log scale = $\beta_0 + \beta_1 + \beta_{Age}$;
- Let $W_3 = g_3(\beta) = \exp(\beta_0 + \beta_1 + \beta_{Age})$.
- $\partial g_3(\beta)/\partial \beta_0 = g_3(\beta)$, $\partial g_3(\beta)/\partial \beta_1 = g_3(\beta)$, and $\partial g_3(\beta)/\partial \beta_i = 0$ for $i > 1$ and $i \neq \text{Age}$.
- $\partial g_3(\beta)/\partial \beta_{Age} = 50g_3(\beta) = 50\exp(\beta_0 + \beta_1 + 50\beta_{Age}) = 50W_3$.
-
- Let Y_{12} = mean of treatment group at time 2 on log scale = $\beta_0 + \beta_1 + \beta_2 + \beta_3 + \beta_{Age}$;
- Let $W_4 = g_4(\beta) = \exp(\beta_0 + \beta_1 + \beta_2 + \beta_3 + \beta_{Age})$.
- $\partial g_4(\beta)/\partial \beta_i = g_4(\beta)$, $0 \leq i \leq 3$ and $\partial g_4(\beta)/\partial \beta_{Age} = 50g_4(\beta) = 50\exp(\beta_0 + \beta_1 + \beta_2 + \beta_3 + 50\beta_{Age}) = 50W_4$.

- $\Sigma_W = J_g(\beta) \times \Sigma_\beta \times (J_g(\beta))^T$, where $J_g(\beta) = \begin{bmatrix} W_1 & 0 & 0 & 0 & 50W_1 \\ W_2 & 0 & W_2 & 0 & 50W_2 \\ W_3 & W_3 & 0 & 0 & 50W_3 \\ W_4 & W_4 & W_4 & W_4 & 50W_4 \end{bmatrix}$.

- Σ_W will always be a 4 x 4 matrix. However, when the model is adjusted for covariates, such as age, Σ_W will usually not have zeroes in the upper and lower quadrants, as in the unadjusted model.
- For this example, $J_g(\beta)$ has dimensions (4×5) , Σ_β is (5×5) , and $(J_g(\beta))^T$ is (5×4) .

Confidence Intervals (CI) for Difference Scores

- Run the linear mixed model. Output the estimates on the log scale for each group and each time point to Excel. In SAS, the best estimates for the means by time point and randomization are produced by the LSMEANS statement.

- Output the covariance matrix for β from SAS Proc Mixed.
- Compute Σ_W in a matrix routine, such as Proc IML.
- Output the estimate for Σ_W to Excel.
- $W_2 - W_1 =$ Mean Follow-Up – Baseline Score for the Control Group.
- $95\% \text{ CI} = (W_2 - W_1) \pm 1.96\sqrt{\sigma_{W11}^2 + \sigma_{W22}^2 - 2\sigma_{W12}}$.
- Similarly, $W_4 - W_3 =$ Mean Follow-Up – Baseline Score for the Treatment Group.
- $95\% \text{ CI} = (W_4 - W_3) \pm 1.96\sqrt{\sigma_{W33}^2 + \sigma_{W44}^2 - 2\sigma_{W34}}$.
- Intervention Effect = Difference Between Change Scores, Treatment – Control = $(W_4 - W_3) - (W_2 - W_1)$.
- $95\% \text{ CI} = (W_4 - W_3 - W_2 + W_1) \pm 1.96\sqrt{\sigma_{W11}^2 + \sigma_{W22}^2 + \sigma_{W33}^2 + \sigma_{W44}^2 + 2(-\sigma_{W12} - \sigma_{W13} + \sigma_{W14} + \sigma_{W23} - \sigma_{W24} - \sigma_{W34})}$.
- At first, the formula for the variance of the intervention might look like a nightmare to compute, but here is a shortcut.
- First, sum the diagonal entries on the Σ_W matrix. This can be done with the trace() function in SAS Proc IML or in Excel.
- Next, sum 2 × entries in the upper triangle of Σ_W , where each entry is multiplied by 1 or -1. For example, multiply σ_{12} by -1 and σ_{14} by 1.

If randomization worked, then the treatment and control group should be independent and

Intervention Effect = Difference Between Change Scores, Treatment – Control = $(W_4 - W_3) - (W_2 - W_1) =$

$$(W_4 - W_3 - W_2 + W_1)$$

$95\% \text{ CI} = (W_4 - W_3 - W_2 + W_1) \pm 1.96\sqrt{\sigma_{W11}^2 + \sigma_{W22}^2 + \sigma_{W33}^2 + \sigma_{W44}^2 + 2(-\sigma_{W12} - \sigma_{W34})}$. So, the variance for the intervention

effect, $(W_4 - W_3) - (W_2 - W_1)$, will be the sum of the variances for the control group, $(W_2 - W_1)$, and the treatment group, $(W_4 - W_3)$.

SAS Code – Procs Mixed and IML

```

/* Add ODS output statements to Proc Mixed to create datasets for the LSMEANS */
/* covariance matrix for the betas, so that these datasets can be processed */
/* in Proc IML. */
%Macro MixTime(DepVar, VarType, DSet);
Proc datasets library=work memtype=data;
delete LSMeans CovB; /* delete output from previous run */ run;
title "Outcome=&DepVar, Dataset=&DSet";
ods html; ods graphics on;
proc mixed data=&DSet method=REML NOCLPRINT;
class id TimepointN RandomizationN;
model &DepVar = randomizationN timepointN timepointN*randomizationN
/ solution influence (iter = 5 effect=id est) residual
ddfm = kr CovB;
Repeated timepointN / type = &VarType subject = id;
LSMEANS TimepointN*RandomizationN;
ods output LSMeans=LSMeans CovB=CovB;
run;
ods graphics off; ods html close;
%MEnd MixTime;

%MixTime(log_calories, CS,, MOMs_nutr);
%MixTime(log_calories, UN,, MOMs_nutr);

```

```

/* Manipulate LSMeans dataset for Proc IML */
/* Create new variable, Effect, to represent Y1, Y2, Y3, Y4 */
data LSMeansIML;
set LSMeans(keep=Effect timepointN randomizationN Estimate);
if (timepointN=0 and randomizationN=0) then Effect='Y1';
if (timepointN=-1 and randomizationN=0) then Effect='Y2';
if (timepointN=0 and randomizationN=-1) then Effect='Y3';
if (timepointN=-1 and randomizationN=-1) then Effect='Y4';
run;

proc sort data=LSMeansIML;
by Effect; run;

ods html path="c:\temp";
proc print data=LSMeansIML;
run;
ods html close;

/* Manipulate CovB for Proc IML */
/* Rename the columns to correspond to B0, B1, B2, B3 */
/* Due to the class statements, SAS created extra columns that we don't need. */
Data CovBIML;
set CovB(keep=Col1 Col2 Col4 Col6);
Rename Col1=B0 Col2=B1 Col4=B2 Col6=B3;
If _N_ in (1,2,4,6); Run;

*** Use Proc IML to get covariance matrix of W ***;
ods html path="c:\temp";
Proc IML;
Use LSMeansIML;
read all var {'Estimate'} into Y;
Close LSMeansIML;

expY=exp(Y);
print Y, expY;

Use CovBIML;
Read all var {'B0' 'B1' 'B2' 'B3'} into CovB;
Close CovBIML;
print CovB;

/* Construct W, 4x4 matrix */
print (expY[1]);
W=j(4,4,0);
W[1,1]=expY[1];
z={1 3};
W[2,z]=expY[2];
W[3,1:2]=expY[3];
W[4,]=expY[4];
print W;

/* V=W*CovB*T(W) = cov matrix of expY */
/* reset fuzz statement will clean up values in scientific notation, */
/* such a 7E-10, that are actually zero, but have tiny values due to rounding. */
reset fuzz;
V=W*CovB*T(W);
print V;
Quit; ods html close;

```

Views of LSMEANS and COVB Before And After Processing for Proc IML

Below are pictures of the original and processed ODS output from LSMEANS and COVB, indicating the need for the above SAS code to process the output:

Table 2a: Raw LSMEANS Output

Obs	Effect	timepointN	randomizationN	Estimate	StdErr	DF	tValue	Probt
1	timepoint*randomizat	-1	-1	3.7762	0.06568	432	57.50	<.0001
2	timepoint*randomizat	-1	0	4.0301	0.06566	425	61.38	<.0001
3	timepoint*randomizat	0	-1	4.0231	0.06144	389	65.48	<.0001
4	timepoint*randomizat	0	0	4.1016	0.06229	391	65.84	<.0001

Table 2b: LSMEANS Output, Processed for Proc IML

Obs	Effect	timepointN	randomizationN	Estimate
1	Y1	0	0	4.1016
2	Y2	-1	0	4.0301
3	Y3	0	-1	4.0231
4	Y4	-1	-1	3.7762

Table 3a: Raw COVB Output

			B0	B1		B2		B3
Effect	timepointN	randomizationN	Col1	Col2	Col3	Col4	Col5	Col6
Intercept			0.00193	-0.0019		-0.001		0.00102
randomizat ionN		-1	-0.0019	0.00381		0.00102		-0.002
randomizat ionN		0						
timepointN	-1		-0.001	0.00102		0.00228		-0.0023
timepointN	0							
timepoint*r andomizat	-1	-1	0.00102	-0.002		-0.0023		0.00454
timepoint*r andomizat	-1	0						
timepoint*r andomizat	0	-1						
timepoint*r andomizat	0	0						

Table 3b: COVB Output, Processed for Proc IML

Obs	B0	B1	B2	B3
1	0.003880	-0.00388	-0.00176	0.001758
2	-0.00388	0.007655	0.001758	-0.00346
3	-0.00176	0.001758	0.003948	-0.00395
4	0.001758	-0.00346	-0.00395	0.007882

Examples Demonstrating the Advantage to Using the Delta Method Over Reporting % Change

The tables below are from published articles and display the outcomes as percent change and in the actual units, via use of the delta method. Table 4A displays nutrition data from the “Healthy Mothers on the Move” study, a randomized controlled Diabetes prevention intervention trial with pregnant Latina women. Table 4B is an additional example from the REACH Detroit study, a randomized controlled community health worker intervention among African American and Latino adults with Type 2 Diabetes.

Table 4A^{*4}: Unadjusted Intervention Effects (95% CI) for Nutrition Estimates, Comparison of % Change and Delta Method (Only covariates are time and treatment group, N = 275)

Outcome	Percent Change	Delta Method
Calories, kcal	-7.3% (-16.5%, 2.9%)	-136.0 (-341.1, 69.1) kcal
Fruit, servings	3.3% (-11.5%, 20.5%)	0.1 (-0.5, 0.7) servings
Vegetables, servings	41.9% (19.2%, 68.8%)	0.7 (0.4, 1.1) servings
Fiber, g	15.9% (3.1%, 30.3%)	3.1 (0.7, 5.6) g
Added sugar, g	-16.1% (-29.6%, -0.1%)	-8.1 (-16.0, -0.1) g
Percent calories from added sugar	-9.7 (-21.0, 3.1)	-1.0 (-2.5, 0.4) %
Total fat, g	-12.9% (-22.0%, -2.7%)	-9.1 (-17.0, -1.2) g
Total saturated fat, g	-15.7% (-25.2%, -5.0%)	-4.1 (-7.1, -1.0) g

*Data Source: Kieffer, E., Welmerink, D., Welch, K., Clayton, E., Schumann, C., Sinco, B., Uhley, V. Dietary Outcomes of Healthy MOMs/Madres Saludables: A Randomized Controlled Diabetes Prevention Intervention Trial with Pregnant Latina Women. American Journal of Public Health: March 2014, Vol. 104, No. 3, pp. 526-533.

Table 4B^{5}: Adjusted Intervention Effects (95% CI) for Outcomes from a Diabetes Study, Comparison of % Change and Delta Method (Covariates are time, treatment group, gender, race/ethnicity, clinic site, age, N = 164)**

Outcome	Percent Change*	Delta Method
Hemoglobin A1c	-9.7% (-15.9%, -3.0%)	-0.7 (-1.3, -0.1)
LDL Cholesterol, (mg/dl)	-5.8% (-15.6%, 5.1%)	-6.6 (-18.3, 5.1) mg/dl
Systolic Blood Pressure, (mm Hg)	1.0% (-3.1%, 5.1%)	1.9 (-3.3, 7.1) mg Hg
Diastolic Blood Pressure, (mm Hg)	2.2% (-3.1%, 7.9%)	1.9 (-2.2, 5.9) mg Hg
Body Mass Index, kg/m ^{2k}	2.1% (-1.0%, 5.3%)	0.6 (-0.5, 1.6) kg/m ^{2k}
PAID (Problem Areas In Diabetes), 0 – 100 scale	-21.9% (-49.0%, 19.4%)	-2.2 (-7.5, 3.2)

**Data Source: Spencer, M., Rosland, A.R., Kieffer, E., Sinco, B., Valerio, M., Palmisano, G., Anderson, M., Guzman, J.R., Heisler, M., Effectiveness of a Community Health Worker Intervention Among African American and Latino Adults With Type 2 Diabetes: A Randomized Controlled Trial. *American Journal of Public Health*: June 2011, No. 12, Vol. 101, pp. 2253-2260.

Conclusions

- When the analysis is done on the log scale to reduce skewness, the delta method is a useful tool to convert the results from percent change to the original scale of the outcome. The delta method provides a vehicle for reporting intervention effects in units that are more meaningful in social work, public health, and clinical practice.
- The equations from the delta method can be implemented with SAS. All that is needed is the ability to output the covariance matrix of the mixed model coefficients and then manipulate the matrix in a program, such as Proc IML.
- After outputting the estimates by treatment group and time point, along with the covariance matrix of the estimates (W) on the original scale to Excel, confidence intervals can be obtained by programming cell formulas into Excel.
- Limitations. First, implementation of the delta method does require Proc IML, in addition to Proc Mixed. Second, since physical data is seldomly a perfect normal distribution, the delta method may not always produce the desired intervals, although it has worked fine on samples above $N \geq 164$, based on my experience.

REFERENCES

1. West BT, Welch KB, Galecki AT. *Linear mixed models: A practical guide using statistical software*. 2nd ed. New York, NY: Chapman & Hall/CRC; 2014.
2. Vittinghoff E, Shiboski SC, Glidden DV, McCulloch CE. *Regression methods in biostatistics: Linear, logistic, survival, and repeated measures models*. New York, NY: Springer; 2005.
3. Casella G, Berger RL. *Statistical inference*. 2nd ed. Pacific Grove, CA: Duxbury; 2002.
4. Kieffer E, Welmerink D, Welch K, et al. Dietary outcomes of healthy MOMs/madres saludables: A randomized controlled diabetes prevention intervention trial with pregnant latina women. *AJPH*. 2014;104(3):526-533.
5. Spencer M, Rosland AM, Kieffer E, et al. Effectiveness of a community health worker intervention among african american and latino adults with type 2 diabetes: A randomized controlled trial. *AJPH*. 2011;101(12):2253-2260.

ACKNOWLEDGEMENTS

I would like to acknowledge the REACH Detroit Partnership project team and Family Health Advocates, and the Healthy Mothers on the Move project team and Women's Health Advocates, and clients for their innovative work.

GRANTS:

- REACH Detroit Partnership: NIH R18DK0785501A1,
- Michigan Diabetes Research and Training Center (NIH grant 5P60-DK20572).
- Healthy Mothers on the Move. NIH/NIDDK (R18 DK062344).

On a personal level, I would like to thank my parents for raising their children to value education.

CONTACT INFORMATION

Your comments and questions are welcome and encouraged. Contact the author at:

Ms. Brandy R. Sinco
University of Michigan
School of Social Work
1080 S. University St.
Box 183
Ann Arbor, MI 48109-1106

Phone: 734-763-7784
Fax: 734-763-3372
E-Mail: brsinco@umich.edu

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.