

# STA 441s18 Formulas and Printouts

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_{p-1} x_{p-1} + \epsilon$$

$$SST = SSR + SSE$$

$$R^2 = \frac{SSR}{SST}$$

$$a = \frac{R_F^2 - R_R^2}{1 - R_R^2}$$

$$a = \frac{sF}{n-p+sF}$$

$$F = \left(\frac{n-p}{s}\right) \left(\frac{a}{1-a}\right)$$

If an overall test has  $s$  numerator degrees of freedom and critical value  $c$ , the critical value of a Scheffé follow-up test with  $r$  degrees of freedom is  $\left(\frac{s}{r}\right) \cdot c$ .

$$\ln\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 x_1 + \cdots + \beta_{p-1} x_{p-1} = L$$

$$\pi = \frac{e^L}{1+e^L}$$

$$\ln\left(\frac{\pi_1}{\pi_3}\right) = \beta_{0,1} + \beta_{1,1} x_1 + \cdots + \beta_{p-1,1} x_{p-1} = L_1$$

$$\pi_1 = \frac{e^{L_1}}{1+e^{L_1}+e^{L_2}}$$

$$\ln\left(\frac{\pi_2}{\pi_3}\right) = \beta_{0,2} + \beta_{1,2} x_1 + \cdots + \beta_{p-1,2} x_{p-1} = L_2$$

$$\pi_2 = \frac{e^{L_2}}{1+e^{L_1}+e^{L_2}}$$

$$\pi_3 = \frac{1}{1+e^{L_1}+e^{L_2}}$$

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_k \end{pmatrix} = \begin{pmatrix} E[y_1|\mathbf{X}=\mathbf{x}] \\ E[y_2|\mathbf{X}=\mathbf{x}] \\ \vdots \\ E[y_k|\mathbf{X}=\mathbf{x}] \end{pmatrix} = \begin{pmatrix} \beta_{0,1} + \beta_{1,1}x_1 + \cdots + \beta_{p-1,1}x_{p-1} \\ \beta_{0,2} + \beta_{1,2}x_1 + \cdots + \beta_{p-1,2}x_{p-1} \\ \vdots \\ \beta_{0,k} + \beta_{1,k}x_1 + \cdots + \beta_{p-1,k}x_{p-1} \end{pmatrix}$$

Unknown  
(type=un)

Compound Symmetry  
(type=cs)

Autoregressive  
(type=ar(1))

$$\begin{pmatrix} \sigma_1^2 & \sigma_{1,2} & \sigma_{1,3} & \sigma_{1,4} \\ \sigma_{1,2} & \sigma_2^2 & \sigma_{2,3} & \sigma_{2,4} \\ \sigma_{1,3} & \sigma_{2,3} & \sigma_3^2 & \sigma_{3,4} \\ \sigma_{1,4} & \sigma_{2,4} & \sigma_{3,4} & \sigma_4^2 \end{pmatrix} \quad \begin{pmatrix} \sigma^2 + \sigma_1^2 & \sigma_1^2 & \sigma_1^2 & \sigma_1^2 \\ \sigma_1^2 & \sigma^2 + \sigma_1^2 & \sigma_1^2 & \sigma_1^2 \\ \sigma_1^2 & \sigma_1^2 & \sigma^2 + \sigma_1^2 & \sigma_1^2 \\ \sigma_1^2 & \sigma_1^2 & \sigma_1^2 & \sigma^2 + \sigma_1^2 \end{pmatrix} \quad \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{pmatrix}$$

The following pages contain

- SAS program for the Birthweight data
- Results for the Birthweight data
- SAS program for the  $CO_2$  data
- Results for the  $CO_2$  data

```

/***** FinalBirthWeight.sas *****/
title 'STA441s18 Final Exam: Low Birth Weight Data';

proc format; /* Value labels used in data step below */
  value lowfmt 0 = '2500 g +' 1 = 'Under 2500 g';
  value racefmt 1 = 'White'
                2 = 'Black'
                3 = 'Other';
  value ynfmt 0 = 'No' 1 = 'Yes';

data bigbaby;
  infile '/folders/myfolders/441s18/HW/bweight.data.txt' firstobs=2;
  input id low age lwt race smoke ptl ht ui ftv bwt;
  if race = 2 then r2 = 1; else r2=0;
  if race = 3 then r3 = 1; else r3=0;
  label id = 'Identification Code'
        low = 'Low Birth Weight'
        lwt = 'Weight at Last Period'
        smoke = 'Smoke during Pregnancy'
        ptl = 'History of Premature Labour'
        ht = 'History of Hypertension'
        ui = 'Presence of Uterine Irritability'
        ftv = 'Visits to Doctor During 1st trimester'
        bwt = 'Birth Weight in Grams'
        r2 = 'Black vs White'
        r3 = 'Other vs White';
  if ptl > 1 then ptl = 1; /* Recode ptl from # of times to Yes-No */
/***** Value labels defined above in proc format *****/
format low lowfmt.;
format race racefmt.;
format ht ui smoke ptl ynfmt.;

ods exclude Association;
proc logistic;
  model low (event='Under 2500 g') = age lwt smoke ptl ht ui ftv r2 r3;
  Test1: test r2=r3=0;
  Test2: test r2=r3;
  Test3: test age=0, ui=0, ftv=0;
run;

```

STA441s18 Final Exam: Low Birth Weight Data

The LOGISTIC Procedure

| Model Information         |                  |                  |
|---------------------------|------------------|------------------|
| Data Set                  | WORK.BIGBABY     |                  |
| Response Variable         | low              | Low Birth Weight |
| Number of Response Levels | 2                |                  |
| Model                     | binary logit     |                  |
| Optimization Technique    | Fisher's scoring |                  |

|                             |     |
|-----------------------------|-----|
| Number of Observations Read | 189 |
| Number of Observations Used | 189 |

| Response Profile |              |                 |
|------------------|--------------|-----------------|
| Ordered Value    | low          | Total Frequency |
| 1                | 2500 g +     | 130             |
| 2                | Under 2500 g | 59              |

Probability modeled is low='Under 2500 g'.

| Model Convergence Status                      |  |
|-----------------------------------------------|--|
| Convergence criterion (GCONV=1E-8) satisfied. |  |

| Model Fit Statistics |                |                          |
|----------------------|----------------|--------------------------|
| Criterion            | Intercept Only | Intercept and Covariates |
| AIC                  | 236.672        | 216.750                  |
| SC                   | 239.914        | 249.168                  |
| -2 Log L             | 234.672        | 196.750                  |

| Testing Global Null Hypothesis: BETA=0 |            |    |            |
|----------------------------------------|------------|----|------------|
| Test                                   | Chi-Square | DF | Pr > ChiSq |
| Likelihood Ratio                       | 37.9219    | 9  | <.0001     |
| Score                                  | 35.4005    | 9  | <.0001     |
| Wald                                   | 28.6172    | 9  | 0.0008     |

| Analysis of Maximum Likelihood Estimates |    |          |                |                 |            |
|------------------------------------------|----|----------|----------------|-----------------|------------|
| Parameter                                | DF | Estimate | Standard Error | Wald Chi-Square | Pr > ChiSq |
| Intercept                                | 1  | 0.6445   | 1.2239         | 0.2773          | 0.5985     |
| age                                      | 1  | -0.0395  | 0.0383         | 1.0659          | 0.3019     |
| lwt                                      | 1  | -0.0151  | 0.00703        | 4.5942          | 0.0321     |
| smoke                                    | 1  | 0.8595   | 0.4098         | 4.3975          | 0.0360     |
| ptl                                      | 1  | 1.2185   | 0.4630         | 6.9256          | 0.0085     |
| ht                                       | 1  | 1.8604   | 0.7082         | 6.9016          | 0.0086     |
| ui                                       | 1  | 0.7193   | 0.4634         | 2.4091          | 0.1206     |

|     |   |        |        |        |        |
|-----|---|--------|--------|--------|--------|
| ftv | 1 | 0.0509 | 0.1755 | 0.0842 | 0.7717 |
| r2  | 1 | 1.2188 | 0.5332 | 5.2253 | 0.0223 |
| r3  | 1 | 0.8194 | 0.4505 | 3.3089 | 0.0689 |

| Odds Ratio Estimates |                |                            |        |
|----------------------|----------------|----------------------------|--------|
| Effect               | Point Estimate | 95% Wald Confidence Limits |        |
| age                  | 0.961          | 0.892                      | 1.036  |
| lwt                  | 0.985          | 0.972                      | 0.999  |
| smoke                | 2.362          | 1.058                      | 5.274  |
| ptl                  | 3.382          | 1.365                      | 8.381  |
| ht                   | 6.426          | 1.604                      | 25.750 |
| ui                   | 2.053          | 0.828                      | 5.092  |
| ftv                  | 1.052          | 0.746                      | 1.484  |
| r2                   | 3.383          | 1.190                      | 9.620  |
| r3                   | 2.269          | 0.938                      | 5.487  |

| Linear Hypotheses Testing Results |                 |    |            |
|-----------------------------------|-----------------|----|------------|
| Label                             | Wald Chi-Square | DF | Pr > ChiSq |
| Test1                             | 6.2284          | 2  | 0.0444     |
| Test2                             | 0.5343          | 1  | 0.4648     |
| Test3                             | 3.6958          | 3  | 0.2962     |

```
/* FinalCO2.sas */
title 'STA441s18 Final Exam: CO2 Data';

data gas;
  infile '/folders/myfolders/441s18/HW/CO2.data.txt' firstobs=2;
  input lineno Plant $ Type $ Treatment $ conc uptake;
  label conc = 'Ambient CO2 Concentration';

proc mixed data=gas;
  title2 'Repeated Measures';
  class Plant Type Treatment conc;
  model uptake = Type|Treatment|conc;
  repeated / type=cs subject=Plant;
  lsmeans Type Treatment conc;
run;

/* Get interaction plots from proc glm, separately by Type */

proc sort; by type; /* Need to sort first */
quit; ods select IntPlot;
proc glm;
  title3 'Interaction plots';
  class conc Treatment;
  model uptake = Treatment|conc;
  by Type;
run;
ods select all; /* Turn selection off */
```

STA441s18 Final Exam: CO2 Data  
Repeated Measures

The Mixed Procedure

| Model Information         |                   |
|---------------------------|-------------------|
| Data Set                  | WORK.GAS          |
| Dependent Variable        | uptake            |
| Covariance Structure      | Compound Symmetry |
| Subject Effect            | Plant             |
| Estimation Method         | REML              |
| Residual Variance Method  | Profile           |
| Fixed Effects SE Method   | Model-Based       |
| Degrees of Freedom Method | Between-Within    |

| Class Level Information |        |                                                 |
|-------------------------|--------|-------------------------------------------------|
| Class                   | Levels | Values                                          |
| Plant                   | 12     | Mc1 Mc2 Mc3 Mn1 Mn2 Mn3 Qc1 Qc2 Qc3 Qn1 Qn2 Qn3 |
| Type                    | 2      | Mississi Quebec                                 |
| Treatment               | 2      | chilled nonchill                                |
| conc                    | 7      | 95 175 250 350 500 675 1000                     |

| Dimensions            |    |
|-----------------------|----|
| Covariance Parameters | 2  |
| Columns in X          | 72 |
| Columns in Z          | 0  |
| Subjects              | 12 |
| Max Obs per Subject   | 7  |

| Number of Observations          |    |
|---------------------------------|----|
| Number of Observations Read     | 84 |
| Number of Observations Used     | 84 |
| Number of Observations Not Used | 0  |

| Iteration History |             |                 |            |
|-------------------|-------------|-----------------|------------|
| Iteration         | Evaluations | -2 Res Log Like | Criterion  |
| 0                 | 1           | 308.98928384    |            |
| 1                 | 1           | 283.89732028    | 0.00000000 |

Convergence criteria met.

| Covariance Parameter Estimates |         |          |
|--------------------------------|---------|----------|
| Cov Parm                       | Subject | Estimate |
| CS                             | Plant   | 4.4892   |
| Residual                       |         | 3.9298   |

| Fit Statistics          |       |
|-------------------------|-------|
| -2 Res Log Likelihood   | 283.9 |
| AIC (Smaller is Better) | 287.9 |

|                          |       |
|--------------------------|-------|
| AICC (Smaller is Better) | 288.1 |
| BIC (Smaller is Better)  | 288.9 |

| Null Model Likelihood Ratio Test |            |            |
|----------------------------------|------------|------------|
| DF                               | Chi-Square | Pr > ChiSq |
| 1                                | 25.09      | <.0001     |

| Type 3 Tests of Fixed Effects |        |        |         |        |
|-------------------------------|--------|--------|---------|--------|
| Effect                        | Num DF | Den DF | F Value | Pr > F |
| Type                          | 1      | 8      | 95.20   | <.0001 |
| Treatment                     | 1      | 8      | 27.95   | 0.0007 |
| Type*Treatment                | 1      | 8      | 6.38    | 0.0354 |
| conc                          | 6      | 48     | 172.56  | <.0001 |
| Type*conc                     | 6      | 48     | 15.88   | <.0001 |
| Treatment*conc                | 6      | 48     | 4.28    | 0.0016 |
| Type*Treatment*conc           | 6      | 48     | 4.75    | 0.0007 |

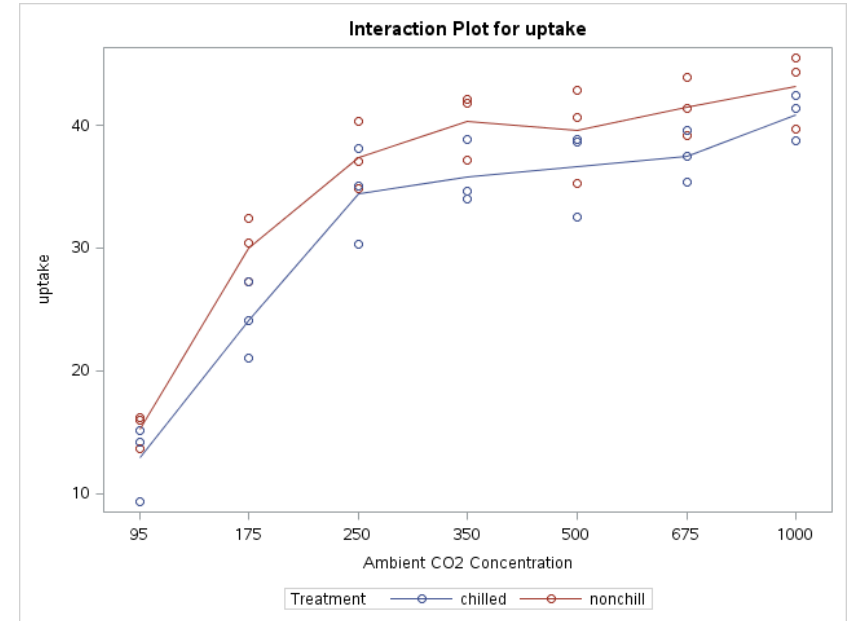
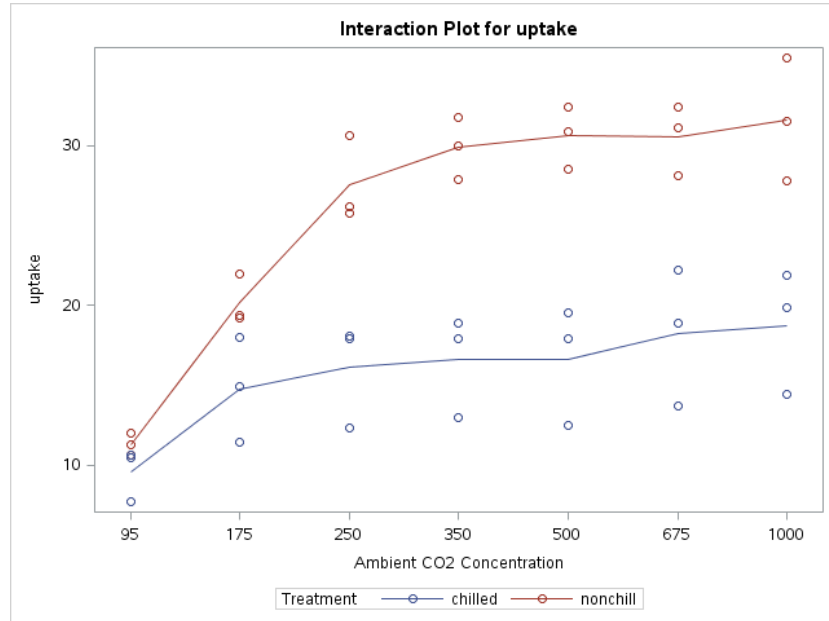
| Least Squares Means |          |           |                           |          |                |    |         |         |
|---------------------|----------|-----------|---------------------------|----------|----------------|----|---------|---------|
| Effect              | Type     | Treatment | Ambient CO2 Concentration | Estimate | Standard Error | DF | t Value | Pr >  t |
| Type                | Mississi |           |                           | 20.8833  | 0.9175         | 8  | 22.76   | <.0001  |
| Type                | Quebec   |           |                           | 33.5429  | 0.9175         | 8  | 36.56   | <.0001  |
| Treatment           |          | chilled   |                           | 23.7833  | 0.9175         | 8  | 25.92   | <.0001  |
| Treatment           |          | nonchill  |                           | 30.6429  | 0.9175         | 8  | 33.40   | <.0001  |
| conc                |          |           | 95                        | 12.2583  | 0.8376         | 48 | 14.64   | <.0001  |
| conc                |          |           | 175                       | 22.2833  | 0.8376         | 48 | 26.60   | <.0001  |
| conc                |          |           | 250                       | 28.8750  | 0.8376         | 48 | 34.47   | <.0001  |
| conc                |          |           | 350                       | 30.6667  | 0.8376         | 48 | 36.61   | <.0001  |
| conc                |          |           | 500                       | 30.8750  | 0.8376         | 48 | 36.86   | <.0001  |
| conc                |          |           | 675                       | 31.9500  | 0.8376         | 48 | 38.14   | <.0001  |
| conc                |          |           | 1000                      | 33.5833  | 0.8376         | 48 | 40.09   | <.0001  |

STA441s18 Final Exam: CO2 Data  
Repeated Measures  
Interaction plots

The GLM Procedure

Dependent Variable: uptake

Type=Mississi



**STA441s18 Final Exam: CO2 Data**

**Repeated Measures  
Interaction plots**

The GLM Procedure

Dependent Variable: uptake

Type=Quebec