Solution:

(a) Let $\mathbf{X} = (X_1, X_2, \dots, X_n)$,

$$\begin{aligned} P(\mathbf{X};\lambda) &= \prod_{i=1}^{n} e^{-\lambda}\frac{\lambda^{X_i}}{X_i!} \\ &= (e^{-n\lambda}) \times (\lambda^{\sum_i X_i}) \times (\prod_{i=1}^{n} \frac{1}{X_i!}) \\ &= \exp\{\eta(\lambda)T(\mathbf{X}) - B(\lambda)\}h(\mathbf{X}), \end{aligned}$$

where $T(\mathbf{X}) = \sum_i X_i$, $\eta(\lambda) = \log(\lambda)$, $B(\lambda) = n\lambda$ and $h(\mathbf{X}) = \prod_{i=1}^{n} \frac{1}{X_i!}$. $P(\mathbf{X};\lambda)$ is exponential family distribution with a single parameter. Therefore, $T(\mathbf{X}) = \sum_i X_i$ is sufficient and complete statistics.

(b) Let $T(\mathbf{X}) = \sum_i X_i$, the moment generation function of $T(\mathbf{X})$

$$M_{T(\mathbf{X})}(s) = \prod_{i=1}^{n} M_{X_i}(s) = \prod_{i=1}^{n} e^{\lambda(e^t-1)} = e^{n\lambda(e^t-1)}$$

$$\Rightarrow \qquad T(\mathbf{X}) \sim Poisson(n\lambda).$$

Therefore,

$$\begin{aligned} E(B^{T(\mathbf{X})}) &= \sum_{k=0}^{\infty} B^k e^{-n\lambda}\frac{(n\lambda)^k}{k!} \\ &= e^{-n\lambda}\sum_{k=0}^{\infty}\frac{(Bn\lambda)^k}{k!} \\ &= e^{-n\lambda}e^{Bn\lambda}\sum_{k=0}^{\infty} e^{-Bn\lambda}\frac{(Bn\lambda)^k}{k!} \\ &= e^{n\lambda(B-1)}. \end{aligned}$$

(c) Let $B = \frac{n-1}{n}$, using part (b),

$$E(B^{T(\mathbf{X})}) = e^{n\lambda(\frac{n-1}{n}-1)} = e^{-\lambda}.$$

Therefore, $\left(\frac{n-1}{n}\right)^{T(\mathbf{X})}$ is an unbiased estimator of $e^{-\lambda}$. $\left(\frac{n-1}{n}\right)^{T(\mathbf{X})}$ is UMVUE, since $\left(\frac{n-1}{n}\right)^{T(\mathbf{X})}$ is a function of a complete and sufficient statistics, and is unbiased.

(d) Based on SLLN, $\bar{X}_n \xrightarrow{p} \lambda$, e^t is a continues function, thus

$$e^{-\bar{X}_n} \xrightarrow{p} e^{-\lambda}.$$

(e) Based on CLT,

$$\sqrt{n}(\bar{X}_n - \lambda) \xrightarrow{d} N(0, \lambda).$$

Following Delta method, $\sqrt{n}(h(\bar{X}_n) - h(\lambda)) \xrightarrow{d} N(0, [h'(\lambda)]^2\lambda)$,

$$\sqrt{n}(e^{-\bar{X}_n} - e^{-\lambda}) \xrightarrow{d} N(0, \left[\frac{\partial e^{-\lambda}}{\partial \lambda}\right]^2 \lambda)$$

$$\Rightarrow$$

$$\sqrt{n}(e^{-\bar{X}_n} - e^{-\lambda}) \xrightarrow{d} N(0, e^{-2\lambda}\lambda)$$

$$\Rightarrow$$

$$\sigma^2 = e^{-2\lambda}\lambda.$$

Solution to Problem 2:

There is some leeway in the type of regression model that could be fit, but given the fact that the responses are counts that are relatively low, a Poisson regression model seems the best choice. The normal linear model does not give a horrible fit, but the nature of the response data leads on toward a Poisson model.

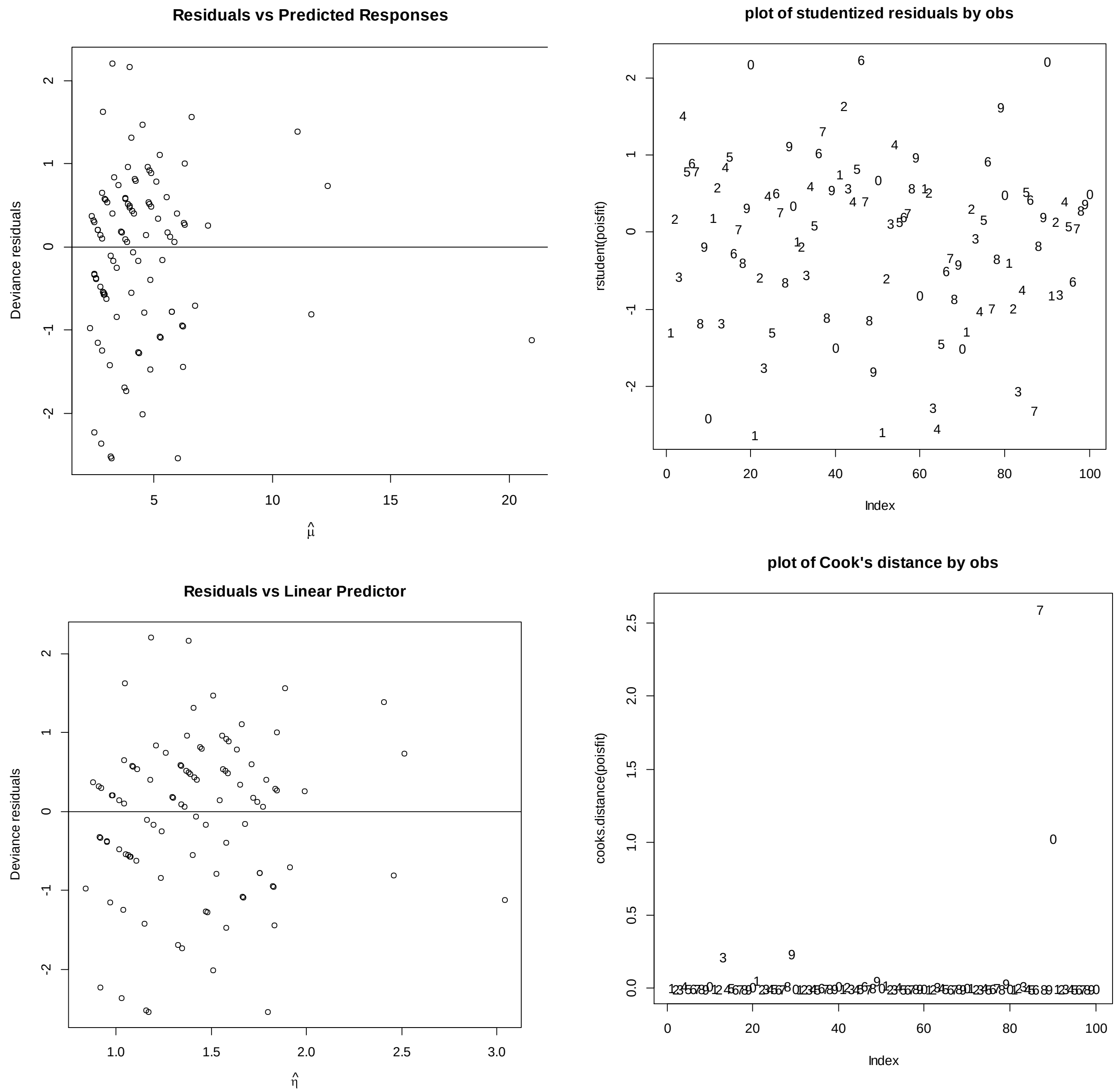
Of paramount importance is that the correct terms be included in the model. A 'weekend' indicator variable should be created and included in the model. The 'holiday' variable should be in the model as well, and there should be a 'weekend × holiday' interaction term in the model, based on the president's suspicions about those factors' joint effects. The 'high temperature' variable should be in the model, and some type of non-linear form should be explored: Below I include it with a quadratic effect, but other explorations of form are possible. The 'date' variable should be included in the model, and the relevant hypothesis test is whether its associated marginal effect is positive.

R Output for Poisson regression model:

```
Call:
glm(formula = sales ~ weekend + holiday + weekend:holiday + hightemp +
    hightempsq + date, family = poisson)


Coefficients:
                       Estimate Std. Error z value Pr(>|z|)
(Intercept)          -0.5536997  1.4619218  -0.379    0.705
weekendyes            0.0779313  0.1073342   0.726    0.468
holidayyes            1.5276885  0.2266158   6.741 1.57e-11 ***
hightemp              0.0435623  0.0442320   0.985    0.325
hightempsq           -0.0003298  0.0003295  -1.001    0.317
date                  0.0330950  0.0057679   5.738 9.59e-09 ***
weekendyes:holidayyes -0.3071655  0.3056174  -1.005    0.315
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 194.46  on 99  degrees of freedom
Residual deviance: 100.57  on 93  degrees of freedom
AIC: 426.52

Number of Fisher Scoring iterations: 4
```

Before looking at model diagnostics, we immediately see that our preliminary model sheds some light on some of the president's suspicions: (1) Expected car sales are higher on holidays (given the other predictors in the model). This effect does not seems to depend on whether the holiday is on a weekend, since the interaction term is non-significant. In fact, the 'weekend' factor is not significant at all.
(2) As suspected, the later dates in a month do yield higher expected car sales, as the coefficient of date is significantly positive. (3) Based on this output, 'high temperature' does not seem to have an effect on car sales (in fact, even if we remove the quadratic term, 'high temperature' is non-significant). But see below…

Some diagnostic residual plots are given here. These exact diagnostic plots need not be done, but some type of diagnostics should be attempted.



Based on these plots, the model seems to fit well overall. There are not any serious outliers, although Observation 87 may be an influential case based on its large Cook's distance. (This is one of the days with a sales count of 16.) We fit the model without Observation 87 to see whether any substantive conclusions change:

```
glm(formula = sales ~ weekend + holiday + weekend:holiday + hightemp +
    hightempsq + date, family = poisson, subset = -87)

Coefficients:
                       Estimate Std. Error z value Pr(>|z|)
(Intercept)          -2.9221447  1.8725814  -1.560   0.1186
weekendyes            0.0775512  0.1072520   0.723   0.4696
holidayyes            1.6057996  0.2301503   6.977 3.01e-12 ***
hightemp              0.1104070  0.0553623   1.994   0.0461 *
hightempsq           -0.0007987  0.0004054  -1.970   0.0488 *
date                  0.0351601  0.0058331   6.028 1.66e-09 ***
weekendyes:holidayyes 0.3186657  0.4026778   0.791   0.4287
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 176.622  on 98  degrees of freedom
Residual deviance:  95.452  on 92  degrees of freedom
AIC: 416.78

Number of Fisher Scoring iterations: 5
```

Most conclusions remain the same, but there is one important difference: We see a marginally significant quadratic effect of high temperature on car sales. The major conclusions about the effects of ‘date, ‘holiday’ and ‘weekend’ remain the same.

Solution to Problem 3:

(a) The parameter space is $\{\underset{\sim}{p}_A, \underset{\sim}{p}_B, \underset{\sim}{p}_C\}$

where $\underset{\sim}{p}_A = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{6}, \frac{1}{6}\right)$

$\underset{\sim}{p}_B = \left(\frac{1}{6}, \frac{1}{3}, \frac{1}{3}, \frac{1}{6}\right)$

$\underset{\sim}{p}_C = \left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{2}\right)$

(b) We can evaluate the likelihood at each point of the parameter space:

$$L(\underset{\sim}{p}_A) \propto \left(\frac{1}{3}\right)^5 \left(\frac{1}{3}\right)^{11} \left(\frac{1}{6}\right)^6 \left(\frac{1}{6}\right)^7 = 1.778663 \times 10^{-18}$$

$$L(\underset{\sim}{p}_B) \propto \left(\frac{1}{6}\right)^5 \left(\frac{1}{3}\right)^{11} \left(\frac{1}{3}\right)^6 \left(\frac{1}{6}\right)^7 = 3.557326 \times 10^{-18}$$

$$L(\underset{\sim}{p}_C) \propto \left(\frac{1}{6}\right)^5 \left(\frac{1}{6}\right)^{11} \left(\frac{1}{6}\right)^6 \left(\frac{1}{2}\right)^7 = 5.935571 \times 10^{-20}.$$

Since $\underset{\sim}{p}_B$ produces the highest likelihood, the MLE is

$$\hat{\underset{\sim}{p}}_{ML} = \underset{\sim}{p}_B = \left(\frac{1}{6}, \frac{1}{3}, \frac{1}{3}, \frac{1}{6}\right)$$

c) By Bayes' Theorem,

$$P(A \mid \text{data}) = \frac{P(\text{data} \mid A)\, P(A)}{P(\text{data} \mid A)\, P(A) + P(\text{data} \mid B)\, P(B) + P(\text{data} \mid C)\, P(C)}$$

$$= \frac{(1.778663 \times 10^{-18})\left(\frac{1}{3}\right)}{(1.778663 \times 10^{-18})\left(\frac{1}{3}\right) + (3.557326 \times 10^{-18})\left(\frac{1}{3}\right) + (5.935571 \times 10^{-20})\left(\frac{1}{3}\right)}$$

$$= \frac{1.778663 \times 10^{-18}}{5.3953447 \times 10^{-18}} \approx \boxed{0.3297}$$

Solution to Problem 3 (continued)

(d) There is probably more than one way to do this, but perhaps the simplest in to use a likelihood ratio test involving

$$\lambda = \frac{\text{likelihood with } \underset{\sim}{p}_c \text{ plugged in}}{\text{likelihood with MLE plugged in}}$$

Then for large samples, $-2 \ln \lambda$ is approximately χ^2 with $3-0=3$ d.f.* So we would reject H_0 if $-2 \ln \lambda > \chi^2_{.10,3} = 6.25$

* = Note there are 3 free parameters that are fixed by H_0 and there are 0 free parameters that are fixed by $\underset{\sim}{p}$ being in the general parameter space.

For these data,

$$\lambda = \frac{\left(\frac{1}{6}\right)^5 \left(\frac{1}{6}\right)^{11} \left(\frac{1}{6}\right)^6 \left(\frac{1}{2}\right)^7}{\left(\frac{1}{6}\right)^5 \left(\frac{1}{3}\right)^{11} \left(\frac{1}{3}\right)^7 \left(\frac{1}{6}\right)^7} = \frac{3^7}{(2^{11})(2^6)} = \frac{3^7}{2^{17}}$$

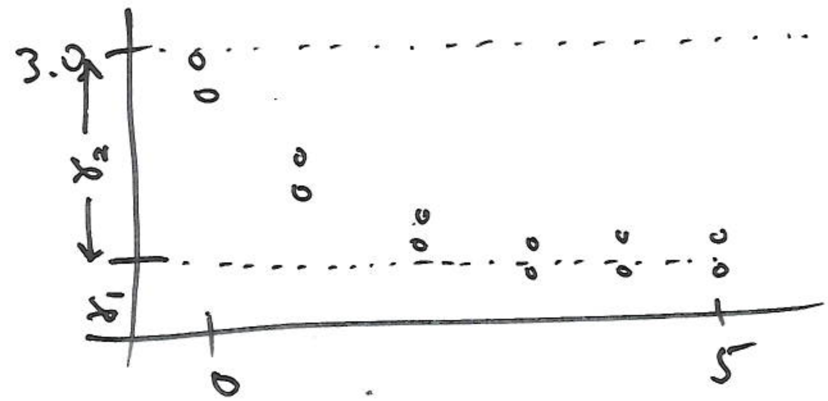
$$\Rightarrow \lambda = 0.0166855 \Rightarrow -2 \ln(\lambda) = 8.19$$

Since $8.19 > 6.25$, we reject H_0 and conclude $\underset{\sim}{p} \neq \underset{\sim}{p}_c$. The approximate P-value is 0.042.

This is an <u>approximate</u> test, and the approximation may not be great, since the sample size is not huge.

Problem 4

(a)



$\Rightarrow \gamma_1 \approx 0.5, \quad \gamma_2 \approx 2.5, \quad \gamma_3 \approx -1$

$\sigma \approx 0.5$. See attached output.

(b) Fitted reg. curve & scatterplot show model fits fine. Residual plot confirms this [random scatter & constant var. okay]. See attached output.

(c) γ_1 is asymptotic mean of transmitted light as conc. increases. 95% CI for γ_1 is $(-0.3, 0.35)$ so we do not reject

$\mu(x) = \gamma_1 + \gamma_2 e^{\gamma_3 x} \longrightarrow$ ⊙ at 5% level.

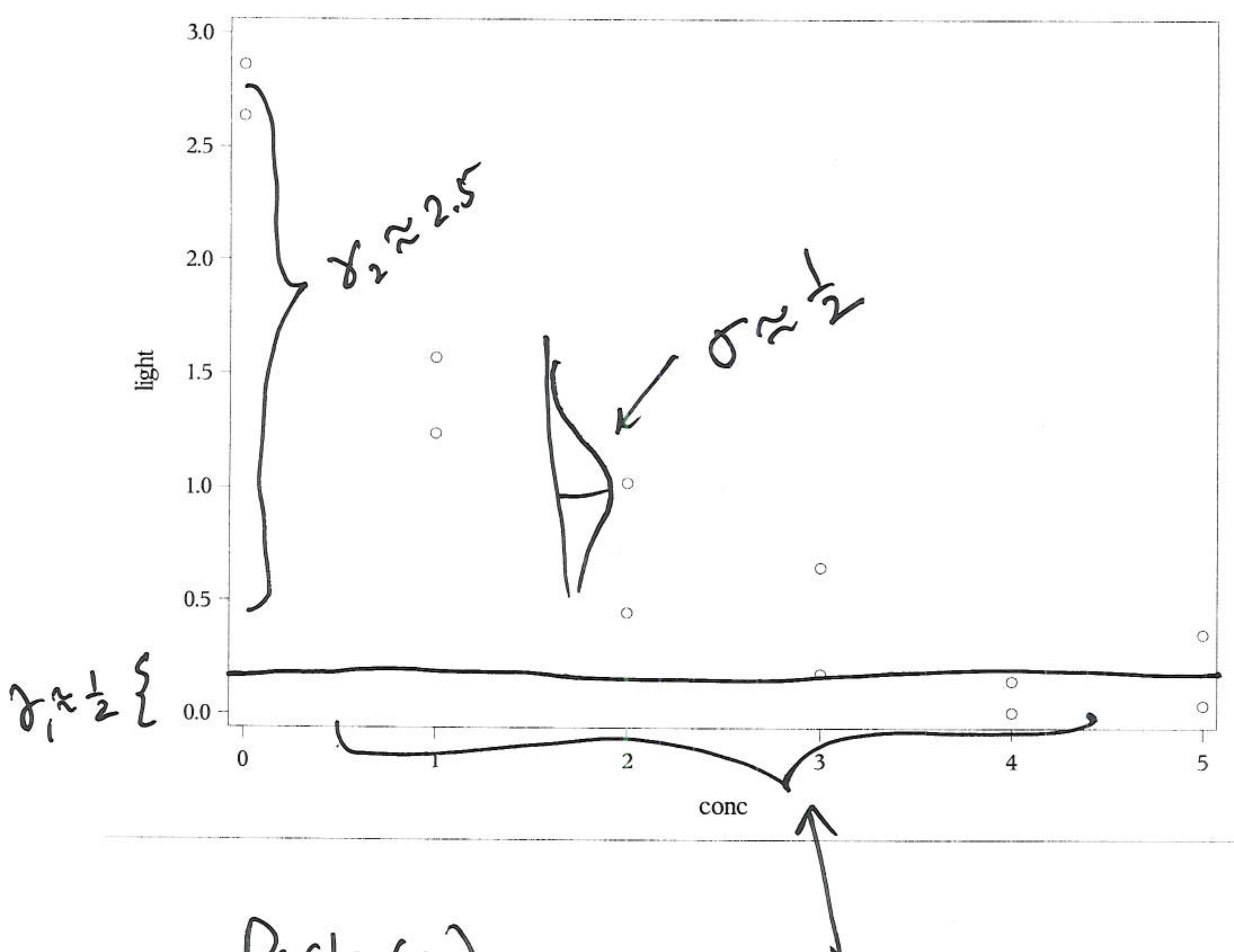
(d) From SAS, $\gamma_1 + \gamma_2$ estimated to be 2.75 w/ 95% CI $(2.45, 3.05)$.

(e) Testing $H_0: \tau_1 = 0$ gives p-value $= 0.825$, we accept constant variance.

(f) AIC for proposed model $= 2.9 \leftarrow$ lower!
AIC for quadratic $= 4.3$

Nonlinear exponential model preferred!

Also quadratic mean <u>increases</u> transmitted light after conc ≈ 4
$\Rightarrow$ non-intuitive!



Part (a)

exponential dies down @ about 5, so $\gamma_3 \approx -1$.

The NLMIXED Procedure

Fit Statistics	
-2 Log Likelihood	-5.1
AIC (smaller is better)	2.9
AICC (smaller is better)	8.6
BIC (smaller is better)	4.9

← (f)

Parameter Estimates								
Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	95% Confidence Limits		Gradient
g1	0.02875	0.1472	12	0.20	0.8484	-0.2919	0.3494	0.000046
g2	2.7233	0.1820	12	14.97	<.0001	2.3268	3.1198	-0.00004
g3	-0.6828	0.1211	12	-5.64	0.0001	-0.9467	-0.4189	0.000129
sigma	0.1959	0.03999	12	4.90	0.0004	0.1088	0.2830	-0.00048

(a)

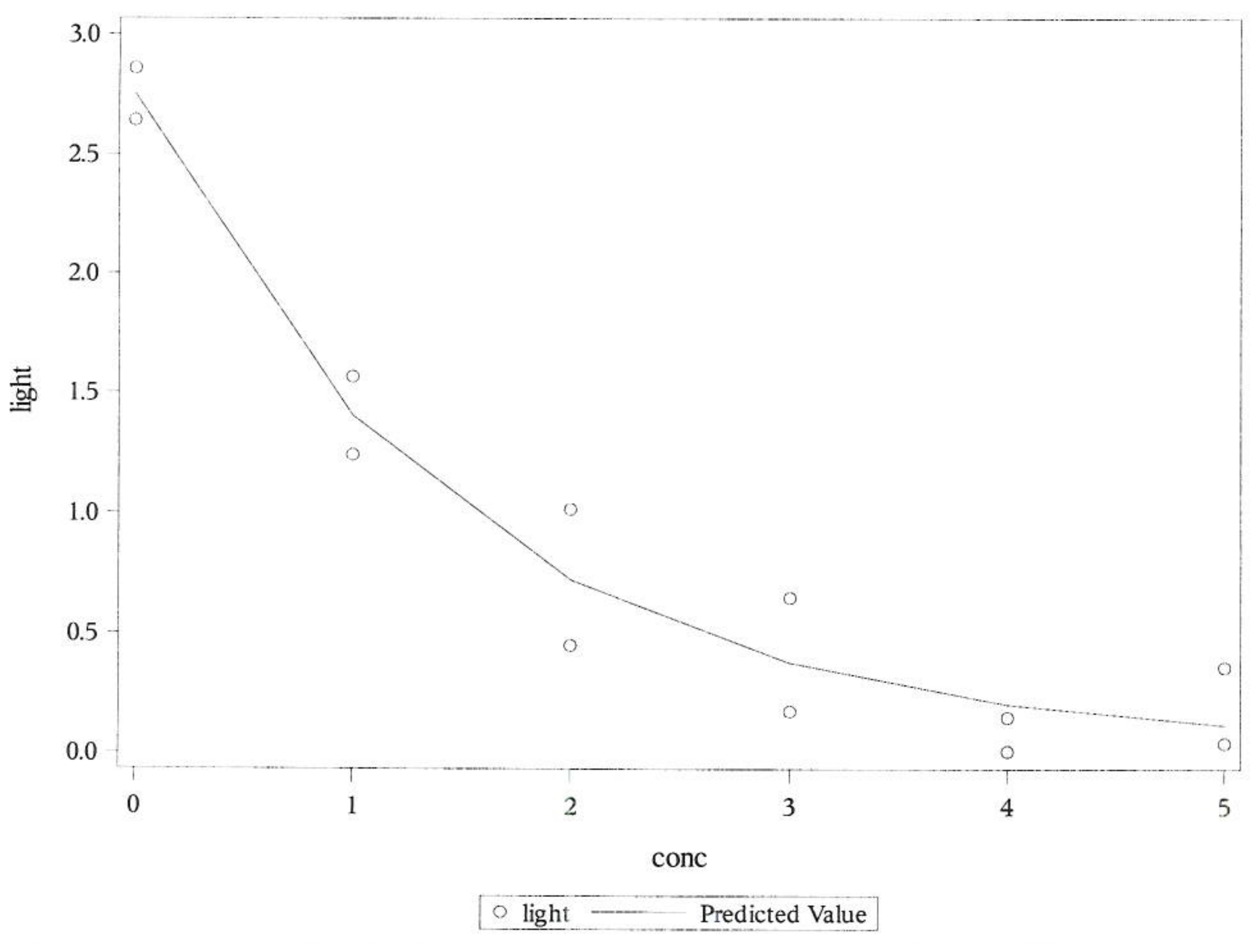
Additional Estimates								
Label	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
mean at zero conc.	2.7520	0.1362	12	20.21	<.0001	0.05	2.4554	3.0487

(d)

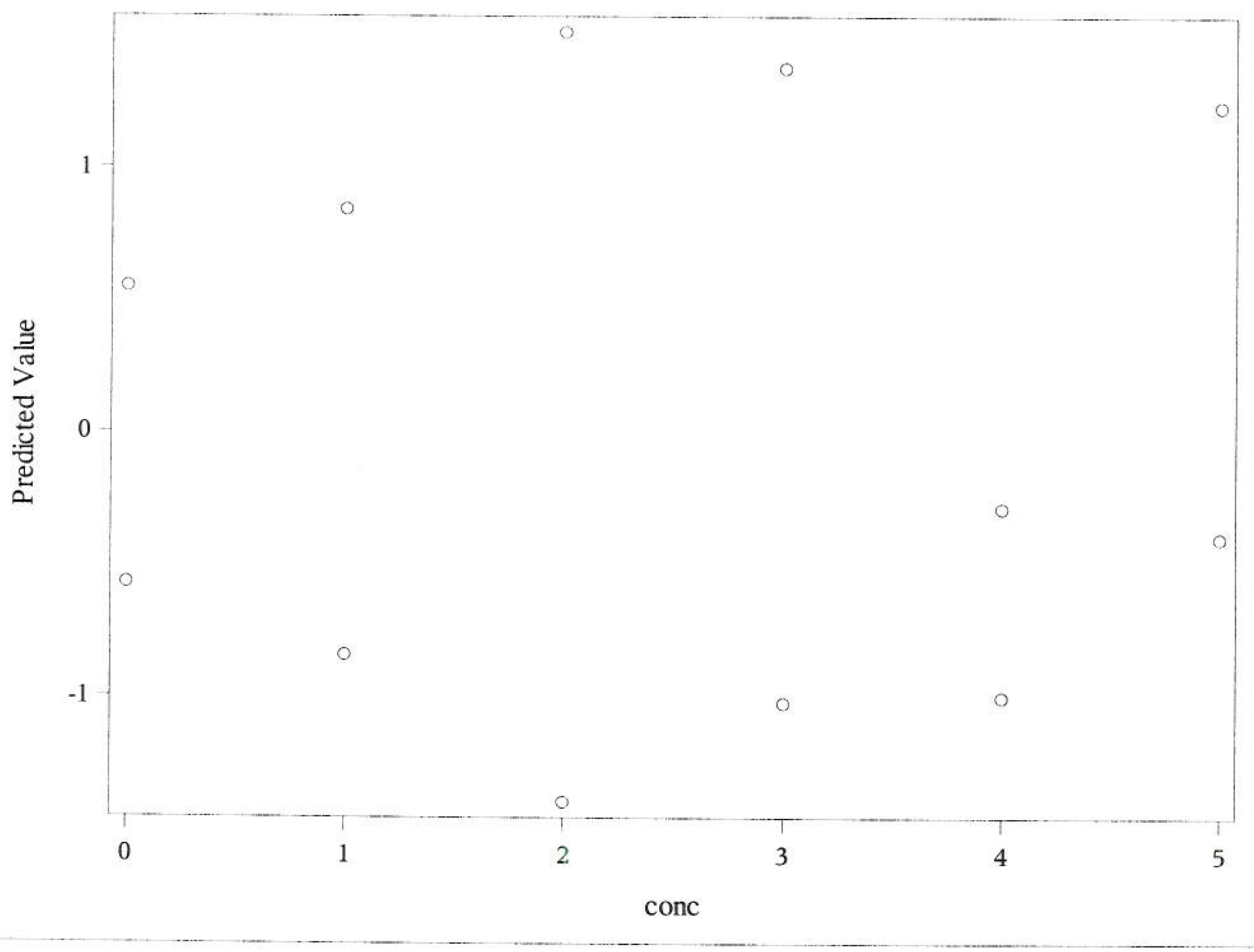
(c) $\hat{\gamma}_1 = 0.0288$, 95% CI = (−0.292, 0.349)

Accept $H_0: \gamma_1 = 0$ @ 5% level

[p-val = 0.85]



(5) Model fits well!



(b) Random scatter & constant. var. Okay!

The NLMIXED Procedure

Iteration History					
Iteration	Calls	Negative Log Likelihood	Difference	Maximum Gradient	Slope
14	46	-2.5614761	6.026E-6	0.000950	-0.00001
15	49	-2.5614761	7.796E-9	0.000069	-1.73E-8

NOTE: GCONV convergence criterion satisfied.

Fit Statistics	
-2 Log Likelihood	-5.1
AIC (smaller is better)	4.9
AICC (smaller is better)	14.9
BIC (smaller is better)	7.3

Parameter Estimates								
Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	95% Confidence Limits		Gradient
g1	0.02474	0.1555	12	0.16	0.8762	-0.3140	0.3635	-0.00005
g2	2.7263	0.1820	12	14.98	<.0001	2.3297	3.1229	-4.13E-6
g3	-0.6796	0.1204	12	-5.65	0.0001	-0.9419	-0.4173	0.000014
tau0	-3.4392	0.8732	12	-3.94	0.0020	-5.3418	-1.5367	-0.00002
tau1	0.06978	0.3088	12	0.23	0.8250	-0.6030	0.7425	-0.00007

(e) Accept $H_0: \tau_1 = 0 \Rightarrow$ Constant variance okay!

The NLMIXED Procedure

Iteration History

Iteration	Calls	Negative Log Likelihood	Difference	Maximum Gradient	Slope
14	53	-1.8369175	2.779E-8	0.003230	-6.61E-8
15	56	-1.8369175	7.19E-10	9.374E-6	-1.44E-9

NOTE: GCONV convergence criterion satisfied.

Fit Statistics

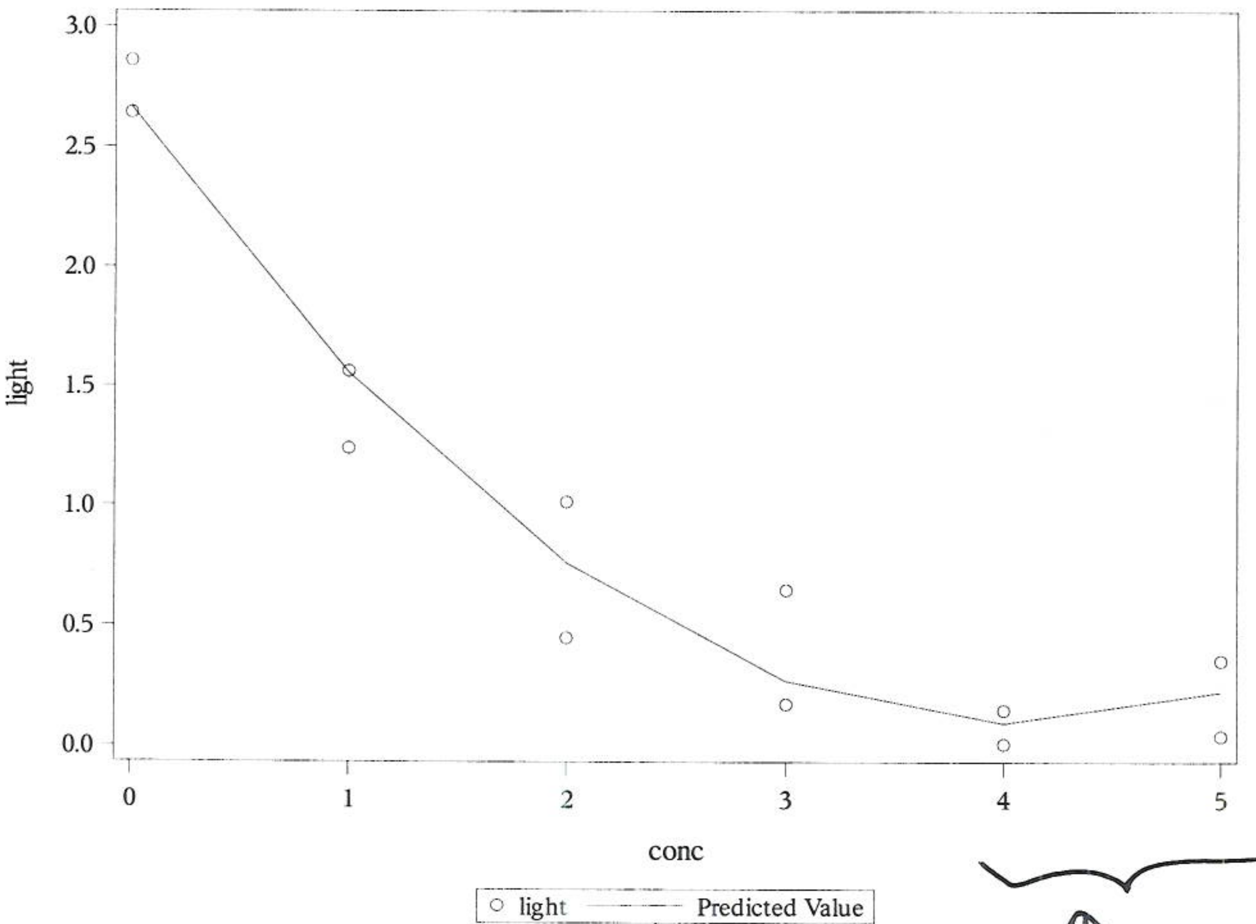
-2 Log Likelihood	-3.7
AIC (smaller is better)	4.3
AICC (smaller is better)	10.0
BIC (smaller is better)	6.3

← (f)

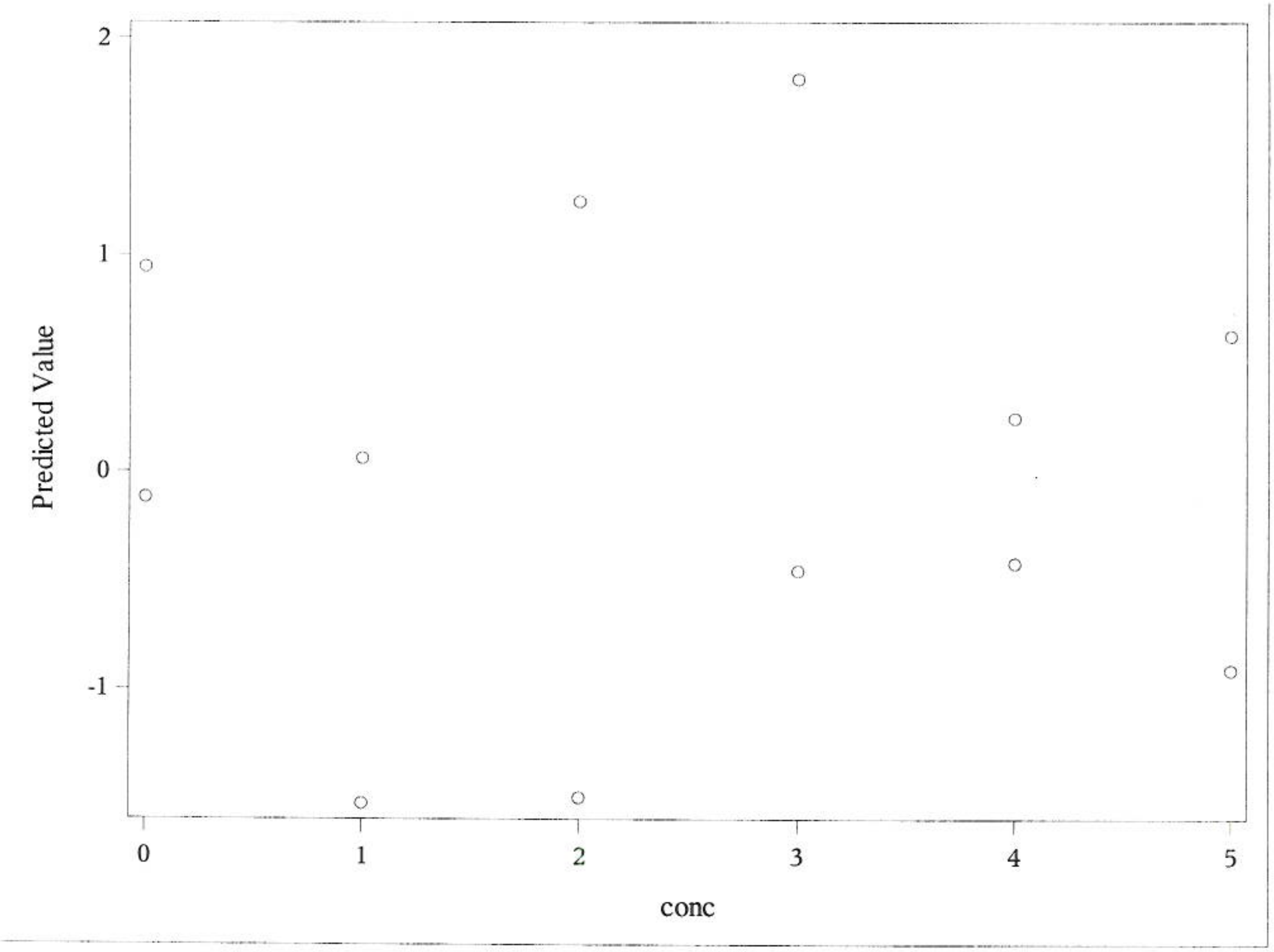
Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	95% Confidence Limits		Gradient
g1	2.6641	0.1331	12	20.02	<.0001	2.3742	2.9540	-2.73E-9
g2	-1.2607	0.1252	12	-10.07	<.0001	-1.5334	-0.9880	-1.6E-6
g3	0.1547	0.02403	12	6.44	<.0001	0.1024	0.2071	-9.37E-6
sigma	0.2076	0.04238	12	4.90	0.0004	0.1153	0.3000	-4.83E-7

(f)



theoretically should <u>not</u> increase!



(f) Residual ~~plot~~ not bad though

SAS code:

```
data light;
input light conc @@;
datalines;
2.86 0.0 2.64 0.0
1.57 1.0 1.24 1.0
0.45 2.0 1.02 2.0
0.65 3.0 0.18 3.0
0.15 4.0 0.01 4.0
0.04 5.0 0.36 5.0
;

* initial scatterplot to eyeball initial values;
proc sgscatter data=light; plot light*conc; run;

* main fit of model;
proc nlmixed data=light;
parms g1=0.5 g2=2.5 g3=-1 sigma=0.5;
mu=g1+g2*exp(g3*conc);
model light ~ normal(mu,sigma*sigma);
predict g1+g2*exp(g3*conc) out=fit;
predict (light-mu)/sigma out=res;
estimate "mean at zero conc." g1+g2;
```

```
* fitted values & raw data;
proc sgplot data=fit;
scatter x=conc y=light;
series x=conc y=pred;

* Pearson resdual plot;
proc sgplot data=res;
scatter x=conc y=pred;

* model with non-constant variance;
proc nlmixed data=light;
parms g1=0.5 g2=2.5 g3=-1 tau0=-1 tau1=0;
mu=g1+g2*exp(g3*conc);
sigma=exp(0.5*tau0+0.5*tau1*conc);
model light ~ normal(mu,sigma*sigma);

* quadratic model;
proc nlmixed data=light;
parms g1=2.5 g2=-1 g3=0 sigma=0.5;
mu=g1+g2*conc+g3*conc*conc;
model light ~ normal(mu,sigma*sigma);
predict g1+g2*conc+g3*conc*conc out=fit;
predict (light-mu)/sigma out=res;

* quadratic model fit;
proc sgplot data=fit;
scatter x=conc y=light;
series x=conc y=pred;

* quadratic model residuals;
proc sgplot data=res;
scatter x=conc y=pred;
run;
```

5. a) U_1, U_2 i.i.d $Uniform(0,1)$. $f(u_1,u_2) = 1; 0 < u_1, u_2 < 1$.

$$Z_1 = \cos(2\pi U_1)\sqrt{-2\log U_2},\ Z_1 = \sin(2\pi U_1)\sqrt{-2\log U_2}$$

Then since $-1 \le \cos(\cdot) \le 1$ on $[0, 2\pi]$ and $-\infty < \log(\cdot) < 0$ on $(0,1)$ we have $-\infty < Z_1 < \infty$ and $-\infty < Z_1 < \infty$.

$$\frac{Z_1}{Z_2} = \tan(2\pi U_1) \ \Rightarrow\ U_1 = \frac{1}{2\pi}\arctan\left(\frac{Z_1}{Z_2}\right)$$

$$Z_1^2 + Z_2^2 = -2\log U_2 \ \Rightarrow\ U_2 = e^{-\frac{1}{2}(Z_1^2+Z_2^2)}$$

Jacobian:

$$\begin{vmatrix} \frac{du_1}{dz_1} & \frac{du_1}{dz_2} \\ \frac{du_2}{dz_1} & \frac{du_2}{dz_2} \end{vmatrix} = \begin{vmatrix} \frac{1}{2\pi}\cdot\frac{1}{1+\left(\frac{z_1}{z_2}\right)^2}\cdot\frac{1}{z_2} & \frac{1}{2\pi}\cdot\frac{1}{1+\left(\frac{z_1}{z_2}\right)^2}\cdot\frac{-z_1}{z_2^2} \\ e^{-\frac{1}{2}(z_1^2+z_2^2)}(-z_1) & e^{-\frac{1}{2}(z_1^2+z_2^2)}(-z_2) \end{vmatrix}$$

$$= \left(-\frac{1}{2\pi}\cdot\frac{e^{-\frac{1}{2}(z_1^2+z_2^2)}}{1+\left(\frac{z_1}{z_2}\right)^2}\right) - \left(\frac{1}{2\pi}\cdot\frac{e^{-\frac{1}{2}(z_1^2+z_2^2)}}{1+\left(\frac{z_1}{z_2}\right)^2}\cdot\left(\frac{z_1}{z_2}\right)^2\right)$$

$$= -\frac{1}{2\pi}e^{-\frac{1}{2}(z_1^2+z_2^2)}$$

Thus

$$f(z_1,z_2) = \frac{1}{2\pi}e^{-\frac{1}{2}(z_1^2+z_2^2)}\ ,\ -\infty < z_1, z_2 < \infty$$

i.e Z_1, Z_2 are iid $N(0,1)$.

b) We know that If $X = \mathbf{m} + \mathbf{BY}$ where $\mathbf{Y} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ then $\mathbf{X} \sim Normal$ with $E(X) = \mathbf{m} + \mathbf{B}\boldsymbol{\mu}$ and $Var(X) = \mathbf{B\Sigma B'}$.

Thus for the present problem we have $Y \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$.
Since $\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-1}\mathbf{X'Y}$, we have $\hat{\boldsymbol{\beta}} \sim Normal$ with

$$E[\hat{\boldsymbol{\beta}}] = (\mathbf{X'X})^{-1}\mathbf{X'E(Y)} = (\mathbf{X'X})^{-1}\mathbf{X'X}\boldsymbol{\beta} = \boldsymbol{\beta}$$

$$Var[\hat{\boldsymbol{\beta}}] = (\mathbf{X'X})^{-1}\mathbf{X'} * \boldsymbol{\sigma^2}\mathbf{I} * \mathbf{X}(\mathbf{X'X})^{-1} = \sigma^2(\mathbf{X'X})^{-1}$$

c) $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I})$ and $\mathbf{V} = \boldsymbol{\mu} + \mathbf{A'Z}$. Then $\mathbf{V} \sim Normal$ with

$$E[\mathbf{V}] = \boldsymbol{\mu} + \mathbf{A'}E[\mathbf{Z}] = \boldsymbol{\mu}$$

$$Var[\mathbf{V}] = \mathbf{A'}Var[\mathbf{Z}]\mathbf{A} = \mathbf{A'IA} = \boldsymbol{\Sigma}$$

Thus if we have $\mathbf{Z}$ then we can get $\mathbf{V}$ by linear transformation.(See .doc file for R code and plot)

d) Let us write $\mathbf{S} = (\mathbf{X'X})^{-1}$ (here $\sigma = 1$), also suppose $\mathbf{A}$ is the Choleski composition of $\mathbf{S}$, then we can generate $\hat{\boldsymbol{\beta}}$ as follows:

- Generate 1000 copies of uniform U_1 and U_2.
- Get 1000 copies of standard normal iid Z_1 and Z_2 by

$$Z_1 = \cos(2\pi U_1)\sqrt{-2\log U_2},\ Z_1 = \sin(2\pi U_1)\sqrt{-2\log U_2}$$

  Let $\mathbf{Z} = \begin{pmatrix} \mathbf{Z_1} \\ \mathbf{Z_2} \end{pmatrix}$.

- Get 1000 copies of $\hat{\boldsymbol{\beta}}$ by linear transformation $\boldsymbol{\beta} + \mathbf{A}' * \mathbf{Z}$.

e) The covariance matrix for $\hat{\boldsymbol{\beta}}$ is $\mathbf{S} = (\mathbf{X}'\mathbf{X})^{-1}$.
From the R calculation $S = \begin{pmatrix} 1.50 & -0.25 \\ -0.25 & 0.05 \end{pmatrix}$.

- From the matrix the correlation coefficient is **-0.9128709**.

- In general, this correlation is free of both $\boldsymbol{\beta}$ and σ^2.

Problem 6

(a) We expect no treatment differences at baseline; an interaction allows treatment differences to manifest gradually over time.

(b) $$\frac{OR(t=1, m, u)}{OR(t=0, m, u)} = \frac{e^{\beta_0 + \beta_1 + \beta_2 m + \beta_3 m + u}}{e^{\beta_0 + \beta_2 m + u}}$$

$$= e^{\beta_1 + \beta_3 m}$$

(c) I would recommend proc glimmix here as it is more stable than proc nlmixed. Also, we primarily used glimmix in class. If you include "class treatment id;" then SAS takes baseline to be treatment = 1

See attached output...

Parameter	est.	CI
β_0	-0.72	$(-1.19, -0.25)$
β_1	-0.026	$(-0.685, 0.633)$
β_2	-0.278	$(-0.341, -0.215)$
β_3	-0.096	$(-0.196, 0.004)$
σ^2	4.709	$(3.53, 5.89)$

(d) See attached output. The 95% CI does <u>not</u> include one at <u>9 months</u>. Treatment 1 takes <u>at least</u> 6 months to significantly reduce odds of severe separation relative to Treatment 0.

(e) A test of $H_0: \sigma^2 = 0$ yields p-value < 0.0001.

(f) Yes, odds of "greater separation" decrease over time for both treatment "1" and "0."

However treatment "1" has smaller odds (lower probability) of greater separation according to the plot.

The GLIMMIX Procedure

Model Information	
Data Set	WORK.TOENAIL
Response Variable	Response
Response Distribution	Binomial
Link Function	Logit
Variance Function	Default
Variance Matrix Blocked By	ID
Estimation Technique	Residual PL
Degrees of Freedom Method	Containment

Class Level Information		
Class	**Levels**	**Values**
ID	294	1 2 3 4 6 7 9 10 11 12 13 15 16 17 18 19 20 21 22 23 24 25 28 29 30 31 33 35 37 38 39 40 41 45 48 49 50 51 52 53 54 55 56 58 59 60 61 63 64 65 66 68 69 70 72 73 75 76 78 79 80 81 82 83 84 85 86 87 88 89 90 93 94 95 96 97 99 101 102 104 105 106 107 108 109 110 111 114 116 117 118 119 120 123 124 125 126 127 129 131 132 133 134 136 137 138 139 140 141 142 143 144 145 146 149 150 151 152 154 156 157 158 160 161 162 163 164 165 166 168 169 170 172 173 174 175 176 177 178 180 181 182 185 186 188 189 190 191 192 193 194 195 197 198 199 200 201 202 203 204 205 206 207 209 210 211 212 213 214 215 216 217 218 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 237 239 240 241 242 243 245 246 247 248 249 250 251 252 254 255 256 258 259 260 261 262 263 264 266 269 270 271 273 275 276 277 278 279 283 284 287 288 289 290 292 293 294 295 297 298 300 301 302 305 306 307 308 309 310 311 312 313 314 316 319 321 324 325 327 328 330 331 332 333 334 335 336 337 338 340 341 343 346 350 351 352 353 354 355 356 357 358 359 360 361 363 364 365 366 367 368 369 372 373 374 377 381 382 383

Number of Observations Read	1908
Number of Observations Used	1908

Dimensions	
G-side Cov. Parameters	1
Columns in X	4
Columns in Z per Subject	1
Subjects (Blocks in V)	294
Max Obs per Subject	7

Optimization Information	
Optimization Technique	Dual Quasi-Newton
Parameters in Optimization	1
Lower Boundaries	1
Upper Boundaries	0

The GLIMMIX Procedure

Optimization Information	
Fixed Effects	Profiled
Starting From	Data

Iteration History					
Iteration	**Restarts**	**Subiterations**	**Objective Function**	**Change**	**Max Gradient**
0	0	5	8519.1940045	0.95204430	0.00027
1	0	4	9475.4880329	0.47981014	0.000654
2	0	5	10396.037151	0.20592568	9.789E-7
3	0	4	10932.841553	0.07454117	0.000022
4	0	3	11106.553815	0.02771427	1.836E-6
5	0	2	11147.350274	0.00667517	2.223E-6
6	0	2	11156.554733	0.00146360	1.998E-7
7	0	1	11158.606044	0.00032309	6.36E-6
8	0	1	11159.060897	0.00007112	3.102E-7
9	0	1	11159.161118	0.00001571	1.792E-8
10	0	1	11159.183261	0.00000347	2.117E-8
11	0	1	11159.188155	0.00000077	5.59E-10
12	0	0	11159.189239	0.00000000	3.635E-6

Convergence criterion (PCONV=1.11022E-8) satisfied.

Fit Statistics	
-2 Res Log Pseudo-Likelihood	11159.19
Generalized Chi-Square	1489.85
Gener. Chi-Square / DF	0.78

Covariance Parameter Estimates			
Cov Parm	**Subject**	**Estimate**	**Standard Error**
Intercept	ID	4.7095	0.6024

← (c)

The GLIMMIX Procedure

Solutions for Fixed Effects								
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
Intercept	-0.7204	0.2370	292	-3.04	0.0026	0.05	-1.1868	-0.2540
Treatment	-0.02594	0.3360	1612	-0.08	0.9385	0.05	-0.6850	0.6331
Month	-0.2782	0.03222	1612	-8.64	<.0001	0.05	-0.3414	-0.2150
Treatment*Month	-0.09583	0.05105	1612	-1.88	0.0607	0.05	-0.1960	0.004307

(c)

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Treatment	1	1612	0.01	0.9385
Month	1	1612	74.57	<.0001
Treatment*Month	1	1612	3.52	0.0607

Estimates									
Label	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper	Exponentiated Estimate
m= 0	-0.02594	0.3360	1612	-0.08	0.9385	0.05	-0.6850	0.6331	0.9744
m= 3	-0.3134	0.3093	1612	-1.01	0.3110	0.05	-0.9201	0.2932	0.7309
m= 6	-0.6009	0.3540	1612	-1.70	0.0898	0.05	-1.2953	0.09348	0.5483
m= 9	-0.8884	0.4493	1612	-1.98	0.0482	0.05	-1.7698	-0.00702	0.4113
m=12	-1.1759	0.5704	1612	-2.06	0.0394	0.05	-2.2948	-0.05698	0.3086

(d)

Estimates		
Label	Exponentiated Lower	Exponentiated Upper
m= 0	0.5041	1.8835
m= 3	0.3985	1.3408
m= 6	0.2738	1.0980
m= 9	0.1704	0.9930
m=12	0.1008	0.9446

The GLIMMIX Procedure

Tests of Covariance Parameters Based on the Residual Pseudo-Likelihood					
Label	**DF**	**-2 Res Log P-Like**	**ChiSq**	**Pr > ChiSq**	**Note**
No G-side effects	1	11398	238.83	<.0001	MI

(e)

MI: P-value based on a mixture of chi-squares.

SAS code:

```
proc glimmix data=toenail;   class id;
model Response=Treatment|Month / dist=bin link=logit s cl;
random intercept / subject=id;
covtest zerog;
estimate "m= 0" treatment 1 treatment*month 0 / exp cl;
estimate "m= 3" treatment 1 treatment*month 3 / exp cl;
estimate "m= 6" treatment 1 treatment*month 6 / exp cl;
estimate "m= 9" treatment 1 treatment*month 9 / exp cl;
estimate "m=12" treatment 1 treatment*month 12 / exp cl;
```

Alternatively, one can use “random id;” to get the same output.

R code to get plot:

```
> m=seq(0,12,0.1)
> odds1=exp(-0.7204-0.02594*1-0.2782*m-0.09583*m*1)
> odds0=exp(-0.7204-0.02594*0-0.2782*m-0.09583*m*0)
> plot(m,odds1,type="l",ylim=c(0,0.5),xlab="months",ylab="odds")
> lines(m,odds0,lty=3)
```

(f)