36-402/608 Homework #10 Solutions 4/1

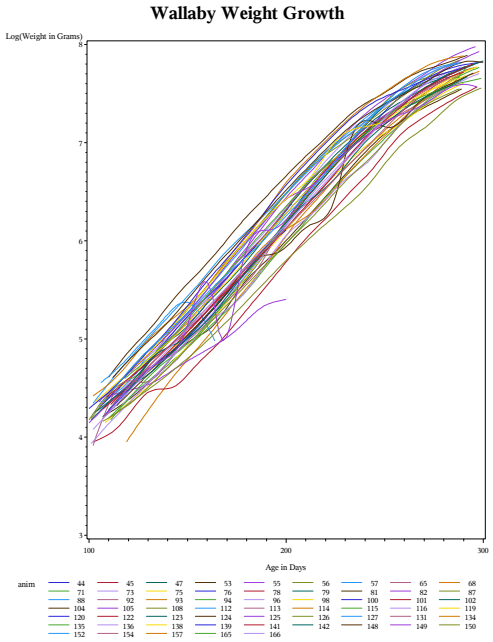
1. Fixing Breakout 17 (60 points)

   You must use SAS for this problem!

   Modify the code in wallaby.sas to load the wallaby data and to create a new outcome in the form of the log of the grams. Follow the model selection ideas of Handout 17 and Breakout 17 to select an appropriate random effects plus possible serial correlation model for the rich fixed effects model (using the default REML fitting method). Different from the handout, as your first model use pure fixed effects (no random or repeated statement). Be sure to use "TYPE=UN" on the RANDOM statement if you include more than one random effect. If you get a "0" standard error for a random effect, drop it regardless of what the BIC says.

   Then use backward selection with method=ML to pare down the fixed effects model. Finish with a REML fit of the final model, including a residual vs. fit plot. Turn in the BIC values (labeled with a model description and the "METHOD") for each model you fit, the final "Covariance Parameter Estimates" and "Solution for Fixed Effects", the residual plot, and a paragraph or two describing what the final model says about wallaby growth.

   The EDA shown here (lowess curves per subject0 may help in model choices.



   The plot show one bend, so at least a quadratic term will be needed. It shows parallel curves, so at least a random intercept will be needed (or maybe just strong serial correlation).

Method	BIC	Random Effect	Fixed Effects	Comment
REML	-240.6	none	loca M*(days+days2+days3)	
REML	-888.5	Intercept	loca M*(days+days2+days3)	
REML	-1004.8	Int. + days slope	loca M*(days+days2+days3)	RS SE=0
REML	-1314.4	Int. + AR1/Sp.	loca M*(days+days2+days3)	RI Est.=0
REML	-1314.4	days slope + AR1/Sp.	loca M*(days+days2+days3)	RS Est.=0
REML	-1314.4	AR1/Sp.	loca M*(days+days2+days3)	BIC not comp. to next
ML	-1418.6	AR1/Sp.	loca M*(days+days2+days3)	days3*M p=0.94
ML	-1422.7	AR1/Sp.	loca M*(days+days2) +days3	loca p=0.47 (M in IA)
ML	-1457.7	AR1/Sp.	M*(days+days2) +days3	remainder needed
REML	-1365.2	AR1/Sp.	M*(days+days2) +days3	BIC not comp. to prev.

```
ods graphics on / imagename="MyResPlot" imagefmt = pdf;
proc mixed covtest method=REML plots=studentpanel(conditional);
  class male;
  model  logGrams = male|daysC male|daysC2 daysC3 / solution;
  repeated / subject=anim type=sp(pow)(daysC);
run;
ods graphics off;
```

Model Information

Data Set	HERE.WALLABY
Dependent Variable	logGrams
Covariance Structure	Spatial Power
Subject Effect	anim
Estimation Method	REML

Dimensions

Covariance Parameters	2
Columns in X	10
Columns in Z	0
Subjects	61
Max Obs Per Subject	16

Convergence criteria met.

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
SP(POW)	anim	0.9957	0.000704	1413.3	<.0001
Residual		0.03465	0.005182	6.69	<.0001

Fit Statistics

BIC (smaller is better) -1365.2

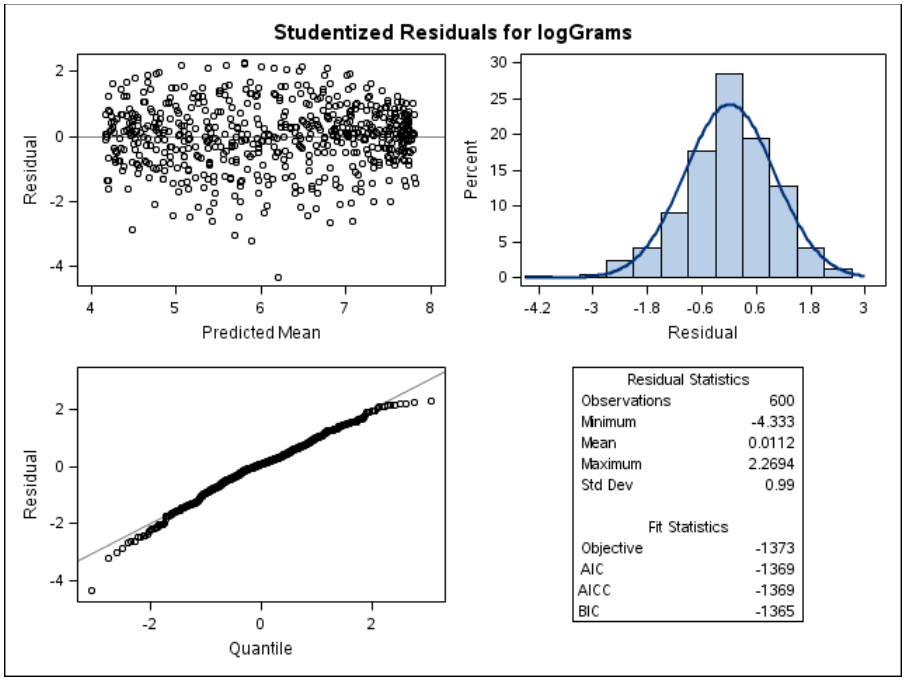
Solution for Fixed Effects

Effect	male	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		4.1652	0.03722	59	111.91	<.0001
male	0	0.001849	0.05535	59	0.03	0.9735
male	1	0	.	.	.	.
daysC		0.01485	0.000746	534	19.91	<.0001
daysC*male	0	0.002074	0.000770	534	2.69	0.0073
daysC*male	1	0	.	.	.	.
daysC2		0.000092	7.925E-6	534	11.65	<.0001
daysC2*male	0	-0.00001	3.559E-6	534	-3.47	0.0006
daysC2*male	1	0	.	.	.	.
daysC3		-3.76E-7	0	534	-Infty	<.0001

Note: The zero Estimates represent the (non-meaningful) redundant X columns listed under "Columns in X".

Note: This is not really a "mixed" model because Z and G are unused.

We interpret this as separate cubic male and female average curves (fixed effects), not random effects, and an AR(1) serial correlation structure with a very high (0.996) correlation between measurements that are one day apart (and $0.996^{d_2-d_1}$ for measuements on days d_1 and d_2. The residual error is sd=$\sqrt{0.03465}$=0.19.



The residual vs. fit plot is quite good (perhaps with very slight nonlinearity). The quantile-normal plot is not perfect, but the non-normality is not to a degree likely to exceed the robustness of the method. This is great, and far better than the model without the log.

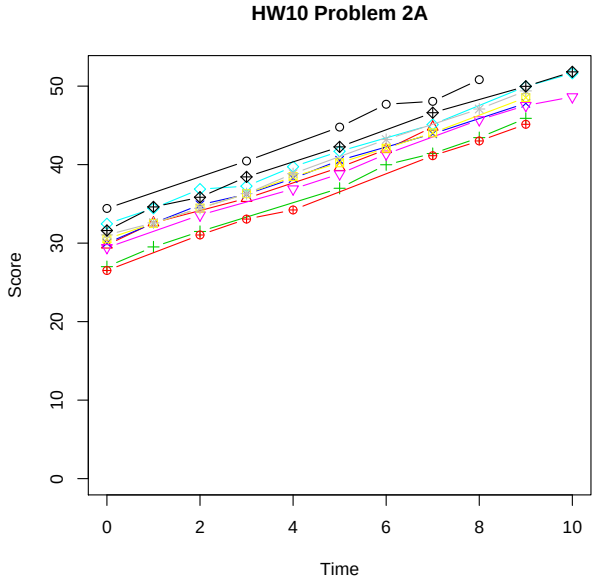
From 100 to 300 days wallabies grow dramatically, along a cubic shape. The males and females follow different curves with females having a steeper slope (by 0.00207+/-0.00077SE) and a slightly smaller curvature (0.00008 vs. 0.00009).

Weight measurements on similar days are highly correlated. Location was not a significant effect.

2. Mixed Model EDA (40 points)

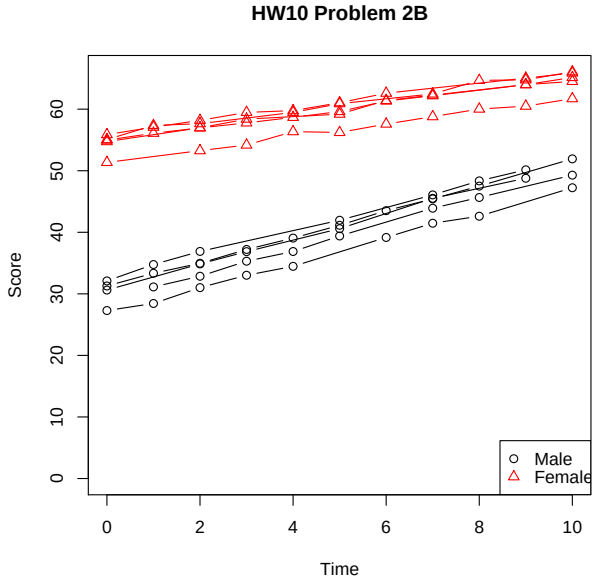
   For each EDA plot, give the most likely fixed and random effect models.

   (a)



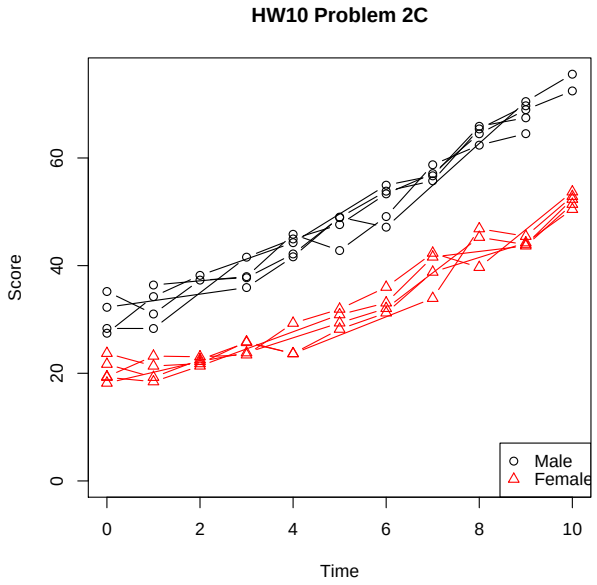
   This plot suggests a linear trend over time with a random intercept (parallel curves for the different subjects).

   (b)



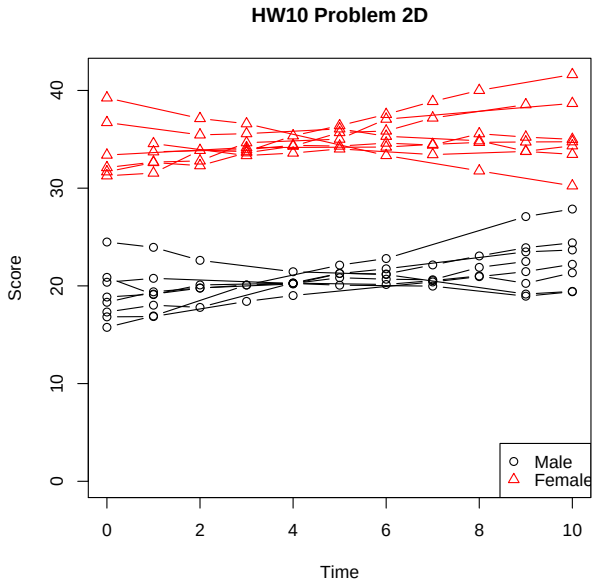
   This plot suggests a linear trend over time with separate intecepts and slopes for males vs. females (fixed termsL intercept, male, time, male:time interaction) and a random intercept per subject.

(c)



This plot suggests quadratic curves with different male vs. female intercept and different slopes and/or curvatures. There is no evidence in the plot of a random effect (all curves share the overall gender curve).

(d)



Here we see random intercept and random slope, with a gender difference in (fixed) intercept.