[1] Thirty nine kids, who had corrective spinal surgery, were followed-up for a period of 12 months after surgery. Of them, 17 developed Kyphosis and 22 did not. Their ages (in months) at the time of surgery are given in Data T2.1.

a) Fit a logistic regression model to study how age affects the probability of developing Kyphosis within one year following spinal surgery. Compute a 95% Confidence Interval for the odds ratio when age of the kid increases one.

b) Do a residual analysis of the model in part a). Please include the residual plot against age with a smooth line fitted to check for curvilinear shape.

c) Does the above residual plot suggest that the probability of developing Kyphosis within one year following spinal surgery depends on age quadratically? Fit this suggested model. Is the squared term of age significant? Please report the Wald test and Likelihood ratio test results. How would you go about testing whether the quadratic term would improve the model fit?

d) Repeat part b) for the suggested model in part c). What do you conclude?

e) Using the model involving linear and quadratic terms in age, if age increases by one month, how does the probability of developing Kyphosis change? Will the change depend on age?

[2] In a designed experiment, 12 formulations of a capsule were made based on different amount of three medications A, B and C in the capsule. Each formulation was prescribed to 100 male and 100 female patients. The number of patients whose blood chemical level dropped by more than 5% within 24 hours is recorded in Table T2.2. (**Hint:** The response variable has binomial distribution with n=100.)

Let π be the proportion of patients whose blood chemical level drops by more than 5% within 24 hours.

a) Propose a model to describe the dependence of π on gender and the amounts of three medications (do not include the interaction terms). Fit the model to the data. How good is the fit? (**Remark:** 1. Consider the amounts of three medications as continuous variables. 2. You need to modify the data for fitting the model by adding a gender variable and then you should have 12*2=24 observations instead of 12.)

b) Test the statistical significance of the gender effect on π. Interpreter 95% confidence interval for odds ration conditioned on any given formulation of the capsule.

c) Based on the following SAS output, what is the optimal capsule formulation for each gender, if it is desired to maximize π. Report the estimated optimal π's.

d) **<u>Based on the following SAS output</u>**, compute the 95% C.I for π, when gender is male and formulation of the capsule is 2 (A=15, B=20, C=20 corresponding to the second observation). (**Remark:** If you do not know how to compute it by hand, you could run the SAS or R procedure to obtain the result to receive partial points.)

The GENMOD Procedure

Estimated Covariance Matrix

	Prm1	Prm2	Prm3	Prm4	Prm5
Prm1	0.09255	-0.000508	-0.002228	-0.001358	-0.003963
Prm2	-0.000508	0.0000152	2.7916E-6	1.5639E-6	-4.776E-6
Prm3	-0.002228	2.7916E-6	0.0000868	2.4071E-6	-7.343E-6
Prm4	-0.001358	1.5639E-6	2.4071E-6	0.0000856	-4.095E-6
Prm5	-0.003963	-4.776E-6	-7.343E-6	-4.095E-6	0.008512

Analysis Of Parameter Estimates

Parameter		DF	Estimate	Standard Error	Likelihood Ratio 95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept		1	-2.7730	0.3042	-3.3732	-2.1802	83.08	<.0001
a		1	0.0457	0.0039	0.0381	0.0533	137.40	<.0001
b		1	0.0699	0.0093	0.0517	0.0882	56.28	<.0001
c		1	0.0390	0.0093	0.0209	0.0572	17.77	<.0001
gender	fema	1	-0.1191	0.0923	-0.3001	0.0617	1.67	0.1969
gender	male	0	0.0000	0.0000	0.0000	0.0000	.	.
Scale		0	1.0000	0.0000	1.0000	1.0000		

NOTE: The scale parameter was held fixed.

e) Would it make any difference if gender is considered a nominal variable (declared by a class statement), or a numerical variable? Justify your answer.

f) Fit a simpler model, if possible, with fewer medications as predictors, without compromising the goodness of fit substantially. (Hint: This is an open question. You could use different criteria to find your optimal model. You could try AIC or BIC (SC) criteria. You could also use backward selection procedure to get a simpler model with reasonable model fit through goodness of fit test.)