QUESTION 1

Refer to the article "Universal hepatitis B vaccination and the decreased mortality from fulminant hepatitis in infants in Taiwan", and to the SAS input, programs and output in the APPENDIX below.

a. The supplied plot (next page) already shows the ***observed*** rates. For each model:
   - calculate and plot the ***fitted*** rates for the years 1975(0) 1985(10) and 1995(20).
   - show your calculations.
   - connect the 3 fitted rates using a curve.

b. Model 2 was successfuly fit with 22 observations (i.e., with data from 1997 and 1998 excluded). With all 24 observations, "the specified model did not converge". Why?

c. From which model did the authors get the 1.10 mentioned in the statement "the ratio of yearly mortality from 1975 to 1998 was 1.10 (95% CI: 1.07-1.11, ***P*** < .001), representing a progressive decrease" ?

   Show how they obtained the reported 1.10.

d. Why is the deviance/df high for all 4 fitted models? What implication does it have?

e. From which model did the authors get the 0.32 mentioned in the statement "The ratio of the average mortality in the period from 1985 to 1998 to that in the period from 1975 to 1984 was 0.32" ? ***(Don't worry if your calculations do not agree exactly, to 2 decimal places, with theirs)***

f. If you ran the following LOGISTIC or GENMOD procedure,

```
proc logistic descending data = hbv ;     proc genmod data = hbv ;
 model deaths/births = per85_98 ;          model deaths/births = per85_98 /
                                           link = logit dist = binomial;
```

what output would you get ? Fill in the blanks ... (and explain your reasoning)

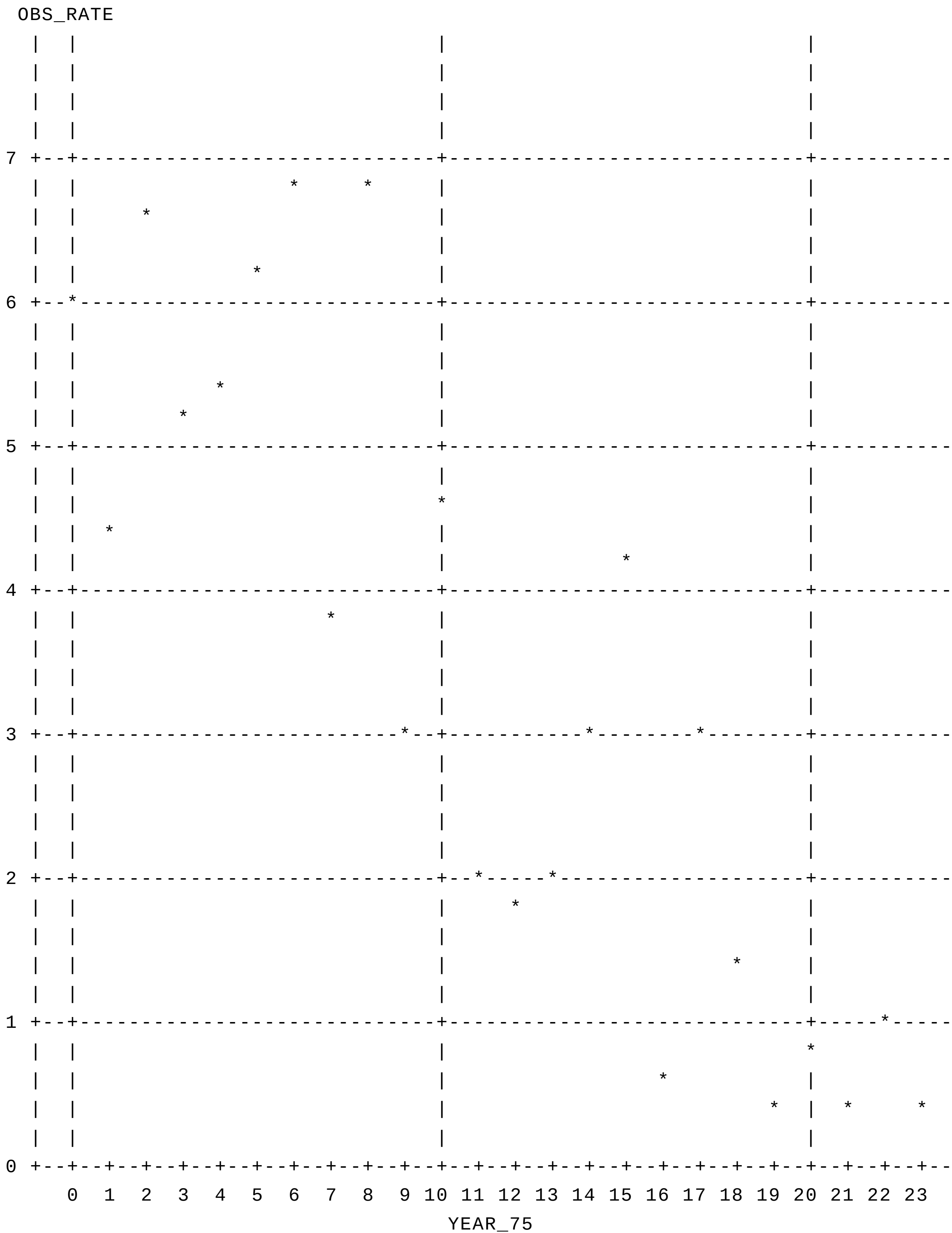
```
Analysis of Maximum Likelihood Estimates

             Parameter        Odds
Variable      Estimate        Ratio


INTERCPT     ________


PER85_98     ________         _____
```

PLOT of rates from article on <u>Universal hepatitis B vaccination & mortality</u>



APPENDIX for question on

Universal hepatitis B vaccination & mortality from fulminant hepatitis

```
DATA HBV; INPUT YEAR births deaths;
           b_100k = births/100000; ln_b100k = log(b_100k);
           year_75 = year - 1975;  per85_98 = (year >= 1985);
           obs_rate = round( deaths / b_100k , 0.1);

 OBS  YEAR  BIRTHS  DEATHS   B_100K  LN_B100K  YEAR_75  PER85_98 OBS_RATE

   1  1975  331760    20    3.31760   1.19924      0        0      6.0
   2  1976  359611    16    3.59611   1.27985      1        0      4.4
...
  23  1997  305622     3    3.05622   1.11718     22        1      1.0
  24  1998  281091     1    2.81091   1.03351     23        1      0.4
```

```
MODEL 1.    proc genmod data = hbv ;
              model deaths = year_75  /
              link = log dist = poisson offset = ln_b100k;
```

Model Information

Description	Value
Data Set	WORK.HBV
Distribution	POISSON
Link Function	LOG
Dependent Variable	DEATHS
Offset Variable	LN_B100K
Observations Used	24

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	22	41.9260	1.9057
Scaled Deviance	22	41.9260	1.9057
Pearson Chi-Square	22	42.4069	1.9276
Scaled Pearson X2	22	42.4069	1.9276
Log Likelihood	.	446.9566	.

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Std Err	ChiSquare	Pr>Chi
INTERCEPT	1	2.0290	0.0942	463.6761	0.0001
YEAR_75	1	-0.0881	0.0101	76.4395	0.0001
SCALE	0	1.0000	0.0000	.	.

NOTE: The scale parameter was held fixed.

```
MODEL 2.     proc genmod data = hbv ;
                model deaths = b_100k b_100k*year_75 /
                link = identity dist = poisson noint;
                where year <= 1996;  <-- NB
```

Model Information

Description	Value
Data Set	WORK.HBV
Distribution	POISSON
Link Function	IDENTITY
Dependent Variable	DEATHS
Observations Used	22

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	20	29.9617	1.4981
Scaled Deviance	20	29.9617	1.4981
Pearson Chi-Square	20	30.5626	1.5281
Scaled Pearson X2	20	30.5626	1.5281
Log Likelihood	.	453.6429	.

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Std Err	ChiSquare	Pr>Chi
INTERCEPT	0	0.0000	0.0000	.	.
B_100K	1	6.6909	0.4643	207.7045	0.0001
B_100K*YEAR_75	1	-0.3009	0.0286	110.3250	0.0001
SCALE	0	1.0000	0.0000	.	.

APPENDIX continued...

MODEL 2-failed

```
proc genmod data = hbv ;
model deaths = b_100k b_100k*year_75 /
link = identity dist = poisson noint;
```

(same as model 2a but with all 24 datapoints)

ERROR: THE MEAN PARAMETER IS EITHER INVALID OR AT A LIMIT OF ITS RANGE FOR SOME OBSERVATIONS.

WARNING: THE SPECIFIED MODEL DID NOT CONVERGE.

MODEL 3.

```
proc genmod data = hbv ;
   model deaths = per85_98  /
   link = log dist = poisson offset = ln_b100k;
```

Model Information

Description	Value
Data Set	WORK.HBV
Distribution	POISSON
Link Function	LOG
Dependent Variable	DEATHS
Offset Variable	LN_B100K
Observations Used	24

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	22	55.4183	2.5190
Scaled Deviance	22	55.4183	2.5190
Pearson Chi-Square	22	55.0601	2.5027
Scaled Pearson X2	22	55.0601	2.5027
Log Likelihood	.	440.2104	.

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Std Err	ChiSquare	Pr>Chi
INTERCEPT	1	1.6812	0.0716	551.1612	0.0001
PER85_98	1	-1.0866	0.1346	65.1717	0.0001
SCALE	0	1.0000	0.0000	.	.

NOTE: The scale parameter was held fixed.

MODEL 4.

```
proc genmod data = hbv ;
   model deaths = b_100k b_100k*per85_98 /
   link = identity dist = poisson noint;
```

Model Information

Description	Value
Data Set	WORK.HBV
Distribution	POISSON
Link Function	IDENTITY
Dependent Variable	DEATHS
Observations Used	24

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	22	55.4183	2.5190
Scaled Deviance	22	55.4183	2.5190
Pearson Chi-Square	22	55.0601	2.5027
Scaled Pearson X2	22	55.0601	2.5027
Log Likelihood	.	440.2104	.

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Std Err	ChiSquare	Pr>Chi
INTERCEPT	0	0.0000	0.0000	.	.
B_100K	1	5.3721	0.3847	195.0000	0.0001
B_100K*PER85_98	1	-3.5596	0.4366	66.4609	0.0001
SCALE	0	1.0000	0.0000	.	.

NOTE: The scale parameter was held fixed.