Insulin Data

Arden Miller

Executive Summary

Two preparations of insulin, the standard preparation and a test preparation, were tested on laboratory mice. The preparations were applied for doses that ranged from 3.4 to 29.0. For both preparations the probability of convulsion increases as the dose increases. There is very strong evidence that for a fixed level of dose the probability of convulsions is less for the test preparation than for the standard preparation. The difference between the two preparations in the probability of convulsions depends on the level of dose but the probability for the test solution is always smaller than the probability for the standard solution. The estimated difference is over 0.2 for dose levels from 10 to 17.5.

Insulin Data

The data (see Table 1) for this assignment came from a study that investigated the effect of insulin on laboratory mice. Two types of insulin preparation were investigated: the standard preparation and a new "test" preparation. Each preparation was tested for different levels of dose. These ranged from 3.4 to 28.0 for the standard preparation and from 6.5 to 29.0 in the test preparation. For each dose levels for each preparation from 30 to 40 mice were injected and the number that suffered convulsions were recorded.

Standard Preparation			Test Preparation		
dose	with convulsions	total	dose	with convulsions	total
3.4	0	33	6.5	2	40
5.2	5	32	10.0	10	30
7.0	11	38	14.0	18	40
8.5	14	37	21.5	21	35
10.5	18	40	29.0	27	37
13.0	21	37			
18.0	23	31			
21.0	30	37			
28.0	27	30			

Table 1: Convulsions in mice data.

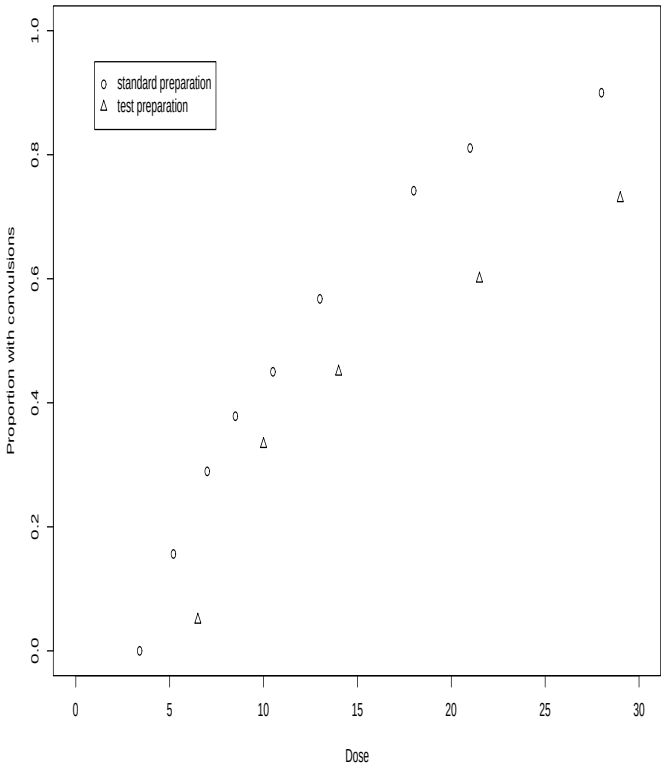


Figure 1: Observed proportions of mice that suffered convulsions.

We are interested in modelling how the proportion of mice with convulsions varies with the type of preparation and with the dose applied. Figure 1 contains plot that shows the observed proportions of mice for each combination of preparation and dose investigated. It is clear that for both preparation the proportion of mice that had convulsions increased as the dose level increases. It is also appears that for a fixed level of dose mice are less apt to have convulsions with the test preparation than with the standard preparation.

1 Modelling the Probability of Convulsions

Logistic regression analysis was used to create a model that relates the probability of convulsions to the type of preparation and the level of dose. This model can be summarised as:

$$\text{Standard preparation:} \quad \hat{\pi} = \frac{\exp\left(4.837 - 15.733 \times \mathsf{dose}^{-1/2}\right)}{1 + \exp\left(4.837 - 15.733 \times \mathsf{dose}^{-1/2}\right)}$$

$$\text{Test preparation:} \quad \hat{\pi} = \frac{\exp\left(3.927 - 15.733 \times \mathsf{dose}^{-1/2}\right)}{1 + \exp\left(3.927 - 15.733 \times \mathsf{dose}^{-1/2}\right)}$$

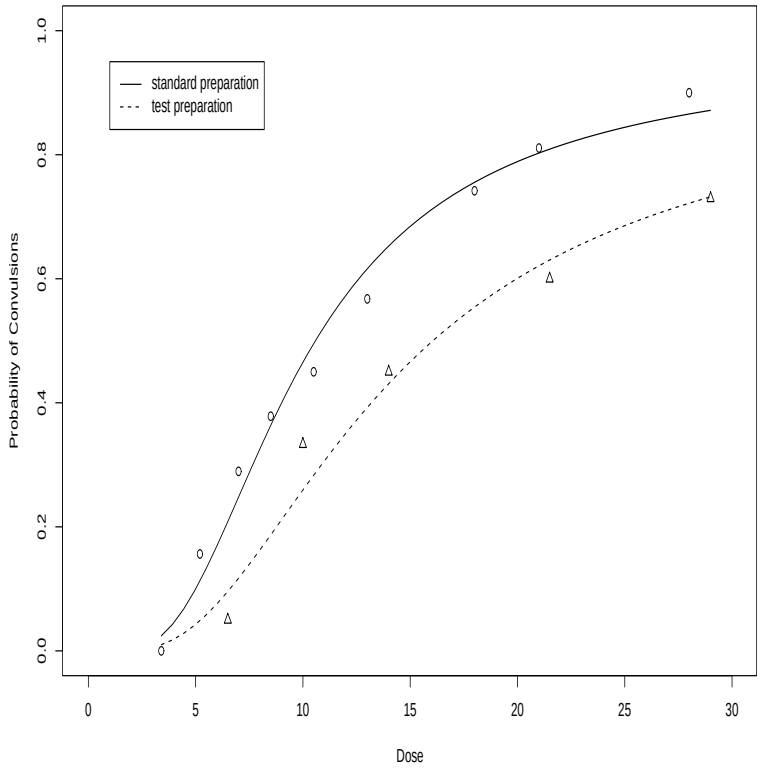


Figure 2: Fitted logistic regression lines.

A graph of the fitted relationships between the probability of convulsion and dose for each preparation are given in Figure 2. The observed proportions of mice with convulsions are superimposed on these plots. The fitted lines appear to track the observed proportions quite well. This plot confirms that the probability of convulsions is consistently less for the test preparation than for the standard preparation. Strictly, speaking we should only apply this conclusion to the range of dose for which we have data for both the test preparation and the standard preparation. In this case that would be for doses from 6.5 to 28.0.

To quantify how big the difference in probabilities is between the two preparations estimates were made for each preparation for levels of dose ranging from 7.5 to 27.5 – see Table 2. The table also reports the 95% confidence intervals for each of the estimates - these are intervals that should contain the true values of the probabilities 95% of the time. Note that for each dose, the two confidence intervals do not overlap. Thus we can be very confident that the true probability for the test solution is lower than that for the standard solution in each case. The difference in the estimated probabilities is given in the last column of Table 2. We can see that this difference is highest when the level of dose is approximately 12.5. Over the range of doses covered by Table 2 (7.5 to 27.5) the difference in estimated probabilities is always at least .147.

	estimated probabilities		
dose	standard	test	difference
7.5	.287 (.225, .359)	.140 (.092, .206)	.147
10.0	.465 (.399, .533)	.259 (.194, .338)	.206
12.5	.596 (.528, .660)	.372 (.296, .455)	.224
15.0	.685 (.616, .746)	.466 (.385, .549)	.219
17.5	.746 (.678, .804)	.541 (.457, .624)	.205
20.0	.789 (.722, .843)	.601 (.514, .681)	.188
22.5	.821 (.756, .871)	.648 (.561, .726)	.173
25.0	.844 (.782, .891)	.686 (.599, .761)	.158
27.5	.863 (.802, .907)	.716 (.630, .789)	.147

Table 2: Estimated probabilities of convulsions

Statistical Appendix

I started by fitting the model that just uses prep and dose as regressors.

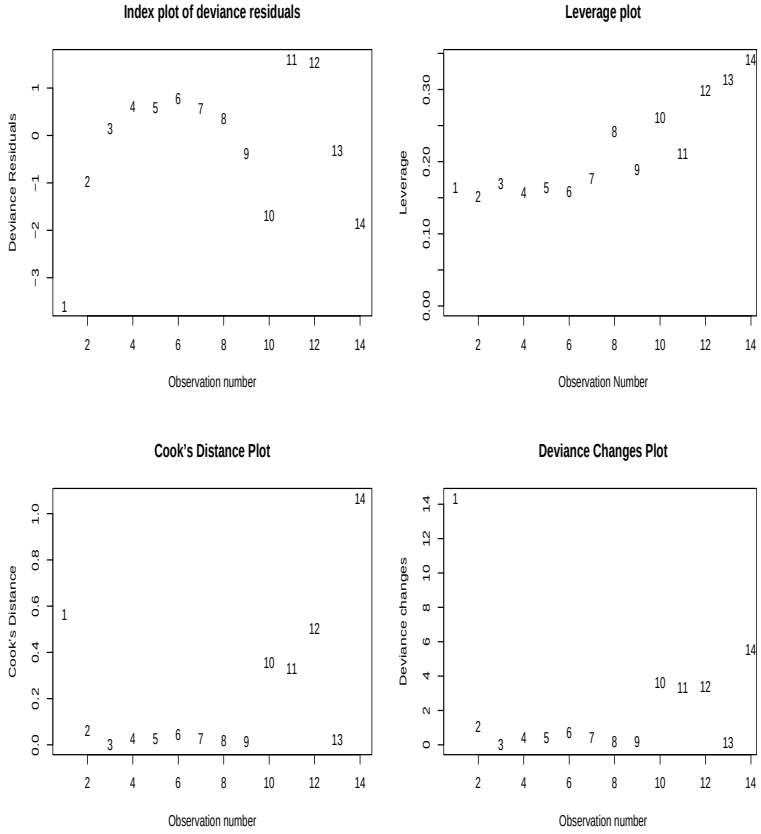
```
Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -2.07827    0.22688  -9.160  < 2e-16 ***
preptst     -0.87525    0.23393  -3.742 0.000183 ***
dose         0.16126    0.01601  10.069  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 166.834  on 13  degrees of freedom
Residual deviance:  27.098  on 11  degrees of freedom
```

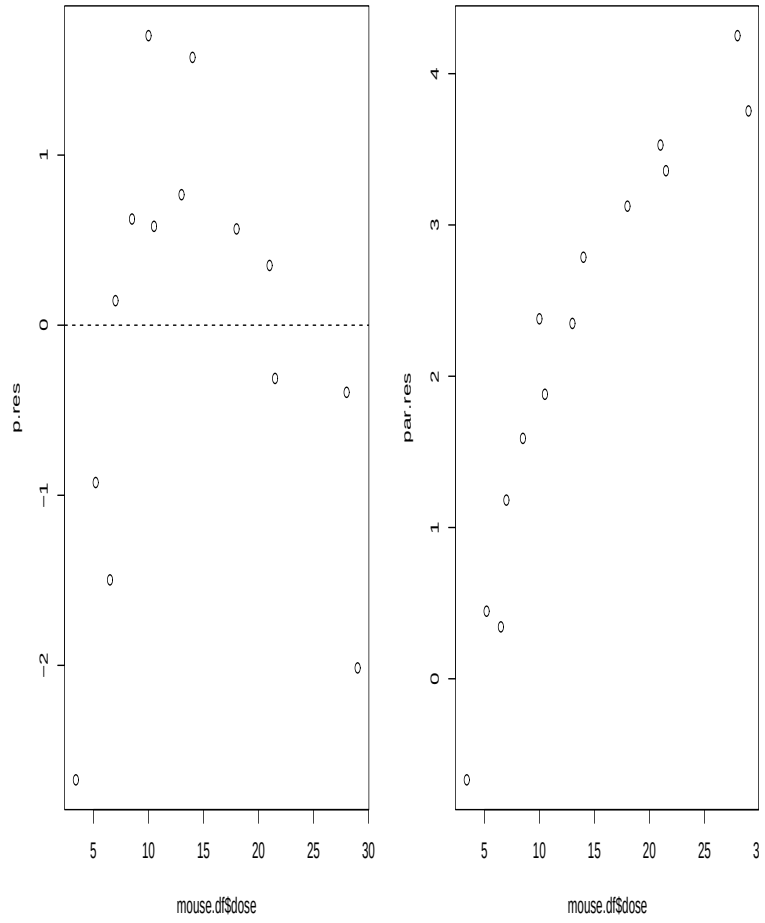
Since we have grouped data with 30 to 40 observations per covariate pattern a goodness-of-fit test can be conducted using the residual deviance as the test statistic. For this model the P-value is 0.0044 which represents very strong evidence against the model being adequate.

The output from `glm.diag.plots` for this model are:



From these plots it seems that observation has a highish value of Cook's Distance and that observation 1 has a high value of deviance changes.

To check the linearity assumption I looked at a plot of the Pearson's residuals versus dose and a plot of the partial residuals versus dose:



Curvature is clearly evident in both these plots which indicates that dose should be transformed (or that a polynomial model in dose should be used).

The next model that I tried was to include the `prep:dose` interaction:

```
             Estimate Std. Error z value Pr(>|z|)
(Intercept)  -2.44388    0.30018  -8.141 3.91e-16 ***
preptst       0.06098    0.50282   0.121   0.9035
dose          0.19295    0.02351   8.208 2.25e-16 ***
preptst:dose -0.06612    0.03224  -2.051   0.0403 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 166.834  on 13  degrees of freedom
Residual deviance:  22.867  on 10  degrees of freedom
AIC: 78.719
```

The interaction is significant and thus improves the model. However a goodness of fit test provides strong evidence that this model is not adequate (P-value = 0.011).

Next, I tried transforming dose. First I tried log(dose). This resulted in the residual deviance dropping to 8.7912 which corresponds to a P-value of 0.64 for the goodness-of-fit test. Thus this model provides an adequate description of the data. I tried adding the `prep:log(dose)` interaction. However there was no evidence that this improves the model - the added variable Chi-square test gives a P-value of 0.55.

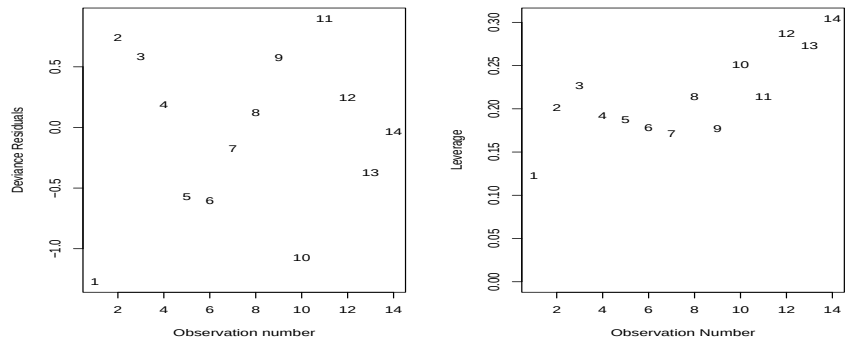
I decided to try a few other power transformations for dose that were close to $p = 0$ (i.e. log). The following table summarises the Residual deviances found for these models and the P-values for the corresponding goodness-of-fit tests:

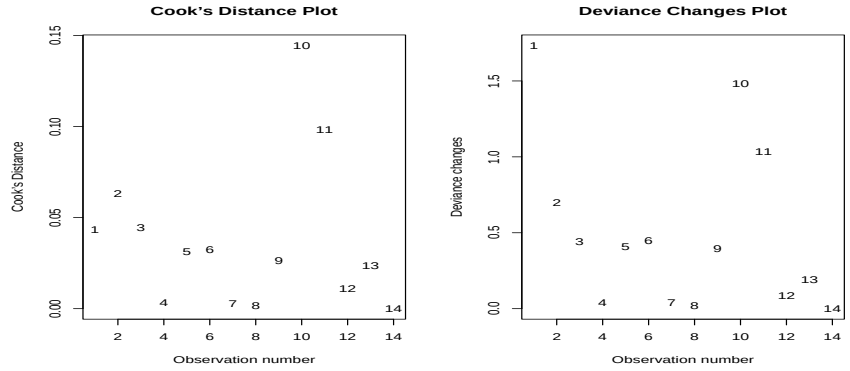
transformation	Resid. Dev.	P-value
dose	27.098	0.004
$\text{dose}^{1/2}$	16.318	0.130
log(dose)	8.791	0.641
$\text{dose}^{-1/2}$	5.769	0.888
dose^{-1}	7.125	0.789

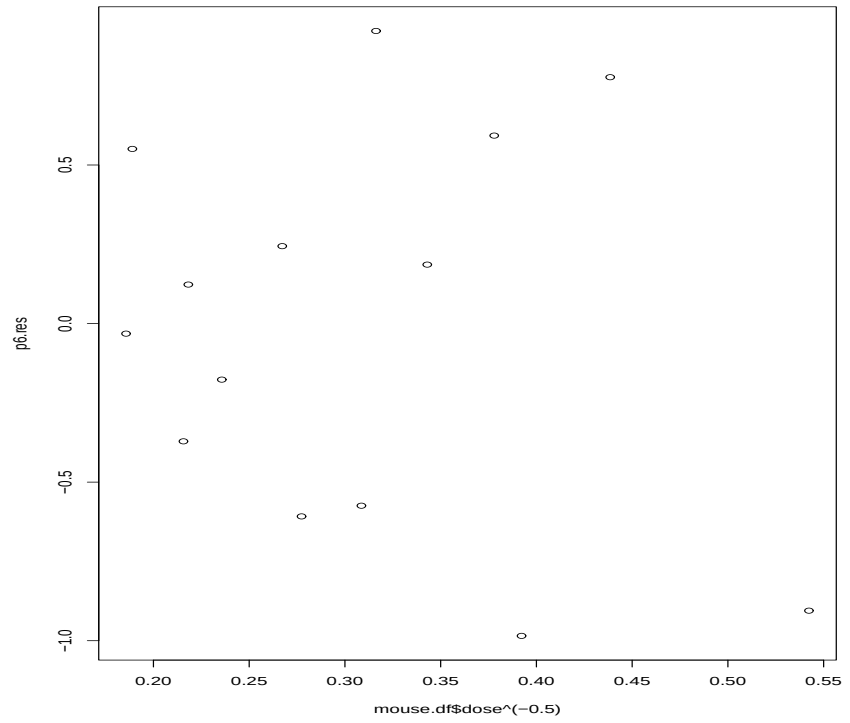
This table indicates that the residual deviance is smallest for $p \approx -1/2$ but p anywhere in the range of (0, -1) would certainly be reasonable.

I also tried using polynomial models in dose. Using the quadratic model for dose gives a Residual deviance of 11.784 whereas using a cubic model gives 7.125. These models require more terms and still have a higher residual deviance than some of the models that use a power transformation of dose. Thus I decided against using one of these.

Overall, the model that uses $\text{dose}^{-1/2}$ give the lowest residual deviance and thus I decided to use this model. Diagnostic plots for this model are:







These plots do not indicate any serious problems with the fitted model. Points 10 and 11 have a slightly high values of Cook's distance but given the small number of points I decided not to try deleting these. Note that observations 10 and 11 represent the two smallest doses (6.5 and 10) for the test solution so deleting these would only leave 3 observations for the test preparation for doses from 14.0 to 29.0. This would limit the useful range of our model to this range of doses.

The estimated confidence intervals in Table 2 were created by first creating confidence intervals for $\text{logit}\,\hat{\pi}$ and then using the logistic function to transform these to confidence intervals for π.