EPIB-621 Solutions 2009 Midterm Exam

1. (a) Intercept is approximately 810, the number of DMF teeth expected per 100 children in towns with a fluoride concentration of 0.

(b) Need to calculate the slope of the tangent line drawn on the graph, easiest done by taking two points, and calculating the change in the y-axis divided by the change in the x-axis between these points. Points chose further apart will give more accurate estimates, in general. Exact value of the slope is -467.4, but I accepted anything reasonably close to that.

2. (a) The value of $\alpha = 200$ provides the expected number of visits to this hospital's emergency room on a day with no precipitation and an average temperature of 0 $°C$.

(b) X_3 is entered as an interaction term in the model, and basically indicates whether any precipitation most likely fell as rain or snow. Thus, the way the model is set up, snow will contribute to the number of emergency room visits, but not rain.

(c) Plugging into the regression equation, we calculate $Y = 200+25+2\times 10 = 245$. Note that it is 2 and not 20 times 10 because the snowfall must be converted to rainfall.

3. (a) We calculate

```
> -0.1892 +c(-1,1) * 1.96*0.0514
[1] -0.289944 -0.088456
```

So an approximate 95% CI is (-0.29, -0.09). Very roughly, this means that for every extra year of age of the donor, survival time likely decreases by somewhere between .1 and .3 months, or between 3 to 9 days, not a very large change. For a ten year difference in donor age one can expect between one to three months survival change, which may be more clinically interesting.

(b) Model seems reasonably linear and residuals are close to normal (they are at least symmetric if not exactly normal), but variance seems to decrease for larger donor ages.

4. (a) Variables X_1 and X_2 seem highly confounded with each other, as witnessed by the large changes in both estimated coefficients and standard errors of both variables, when the other is in or out of the model. X_3 has remarkably stable coefficient and SE estimates, so appears not to be affected by any confounding by either X_1 or X_2.

(b) Because there is no confounding, effect is remarkably stable across models. from model 2, we calculate:

```
> 4.933475 +c(-1,1) * 1.96*2.135339
[1] 0.7482106 9.1187394
```

So an approximate CI is (0.75, 9.12). For each additional week gestation, birth weight of the baby increases, on average, by somewhere between 1 and 9 grams, which is very small, almost negligible. Therefore, while there is some effect, it is not likely to be clinically important (no claim to realism for this example!).

5. (a) From the first study, using the formula for the difference in binomial proportions, 95% CI is:

```
> prop.test(c(.3*50, .25*50), c(50, 50), correct=F)
95 percent confidence interval:
 -0.1247558  0.2247558
```

I disagree with the first researcher, he stated that there was no difference, but his small study does not rule out differences potentially as large as 22% in one direction and 12% in the other, so his study is really inconclusive.

(b) From the second study, the 95% CI is:

```
> prop.test(c(.3*1000, .25*1000), c(1000,1000), correct=F)
95 percent confidence interval:
 0.01092341 0.08907659
```

I agree with the second research, it does seem likely that Drug A is more effectiv e than Drug B, but it is not clear if this effect is of clinical interest, as it could be as small as 1%.

(c) Overall, it seems that Drug A is at least as effective as Drug B, and could be as different as 9% more effective, although further research will need to be carried out to further narrow down the CI so a more definite conclusion can be made.