EPIB-621 Solutions 2009 Final Exam

1. (a) A χ^2 or a Z test can be used. In either case, the hypotheses are:

$$H_0: \quad \text{No effect from distance to highway}$$

$$H_A: \quad \text{There is an effect from distance to highway}$$

Using usual methods, one finds either $\chi^2 = 37.1$ or a $Z = 6.1$, both of which lead to p-values very close to zero (approximately 1×10^{-9}). Therefore the null hypothesis can be rejected, and we conclude that there is at least some effect from distance to the highway.

(b) Using the usual formula, the CI is (0.00378, 0.00872). Therefore while the null hypothesis can be rejected, any effects likely very small (although they can add up over thousands of exposed individuals).

2. (a) This is the interaction term for spring compared to summer for the effect of age, so: For any fixed value for x_1, $\beta_7 x_1 + \beta_4$ gives the effect of spring.

(b) Plugging into the regression equation one calculates 10,900 steps.

(c) Plugging into the regression equation twice and subtracting, one finds $14000 - 7700 = 6300$ is the difference.

3. (a) $OR = \exp(0.6) = 1.822$, with 95% CI of $(\exp(0.5), \exp(0.7)) = (1.648, 2.014)$.

4. (a) $OR = \exp(0.3501) = 1.42$, with 95% calculated from $exp(0.3501 \pm 1.96 \times 0.06956) = (1.23, 1.63)$.

(b) Almost every variable is confounded with every other, coefficient values change almost every time a variable enters or exits the model, se's also unstable.

5. (a) Using the model averaged coefficients, we have:

$$logit(Y) = X = -3.919 + 0.727 \times A1 + 0.006 \times A2 + 0.345 \times T + 1.401 \times S$$

The probabilities are then given by

$$\frac{\exp(X)}{1 + \exp(X)}$$

(b) We plug in to

$$\frac{\exp(X)}{1+\exp(X)}$$

where $X = -3.919 + 1.401$ to find the probability is 0.0746.

(c) Same as (b), but now $X = -3.919$ to find the probability is 0.01947.

6. (a) $\exp(0.11276) = 1.1194$.

(b) $\exp(0.11276 + 0.08634) = 1.22$.

(c) $\exp(3 \times (0.11276 + 0.08634)) = 1.817$.

7. Various answers coud be acceptable, but main points are:

(i) Provides a simple unbiased estimate, but loses precision. Might be acceptable for this simple situation.

(ii) Probably the worst choice, since final inferences will be *too precise* quite possibly biased, and weight might actually be increasing over time rather than stable.

(iii) Best method, as it is unbiased with minimum loss in precision, but may not be worth the trouble in this simple example.

8. (a) We plug in to

$$\frac{\exp(X)}{1+\exp(X)}$$

where $X = 1.377$ to calculate 0.7985.

(b) Yes, region 4 is larger by about 8.3% (95% credible interval of (7.4%, 9.2%).