Solutions, Biostat III exam, December 15, 2010

1. Question 1

   (a) Yes, age is the underlying time scale and all effects of the covariates in the model are adjusted for the time scale.

   (b) For each unit increase in the level of calcium the risk for hip fracture is increased with 52%. This is true within all levels of the remaining covariates and the effect is adjusted for all covariates including the time scale.

   (c) HR = 0.52 OR The risk is 48% lower OR The effect of receiving the higher dose is 0.52 times the effect of receiving the lower dose. (irrespective of calcium level)

2. Question 2

   (a) The HR comparing women with high dose and zero calcium level to women with low dose and zero calcium level is 0.0016.

   (b) $HR = \frac{\exp(\beta_1 * protect + \beta_2 * 1 + \beta_3 * 8 + \beta_4 * 1 * 8)}{\exp(\beta_1 * protect + \beta_2 * 0 + \beta_3 * 8 + \beta_4 * 0 * 8)} = \exp(\beta_2 + \beta_4 * 8) = \exp(\beta_2) * \exp(\beta_4)^8 = 0.0016 * 1.786^8 = 0.166$

   (c) The formal approach to test whether there is need to include an interaction term between dose and calcium level in the model is to conduct a likelihood ratio test.

   The null hypothesis to be tested is:

   $$H_0 : \text{The interaction effect is 0.}$$

   against

   $$H_A : \text{The interaction effect is not 0.}$$

   The test statistic is defined as

   $$-2 \times (\text{log likelihood (model A)} - \text{log likelihood (model B)})$$

   and has a χ^2 distribution under H_0. Note that the degrees of freedom is determined by how many 'extra' parameters we need to estimate in the model containing the interaction compared to the main effects model.

   We thus calculate the difference in log likelihoods between model A (log likelihood: -27.278096) and model B (log likelihood: -26.572158) and get -0.705938. This is multiplied by -2 to get 1.41 which is our test statistic.

   The critical value of a χ^2 with 1 degrees of freedom is 3.84 at the 95 % significance level. Since our test statistic of 1.41< 3.84 the LR test is not statistically significant and we cannot reject the null hypothesis that the interaction effect is 0.

3. Question 3

   (a) There should be a random scatter of the residuals and the smoother through the dots should be horizontal.

   (b) The null hypothesis to be tested is:

   H_0 : A straight line through the residuals (for the covariate calcium) as a function of time has slope 0.

   against

   H_A : A straight line through the residuals (for the covariate calcium) as a function of time has not slope 0.

   (c) The assumption of proportional hazards is not rejected since the p-value is 0.7438.

4. Question 4

   (a) $HR = \frac{\exp(-0.2755*12)}{\exp(-0.2755*8)} = \exp(-0.2755 * 4) = 0.33$. This is not significant since the effect of calcium is not significant (p-value 0.227).

   (b) Hazard rate $= \exp(3.3 - 1.8 - 10 * 0.28 + 0.14) = 0.31$ person-years.

5. Question 5

   (a) Competing risks are events that prevent an event of interest from occurring, rather than just prevent you from seeing it happen (censoring).

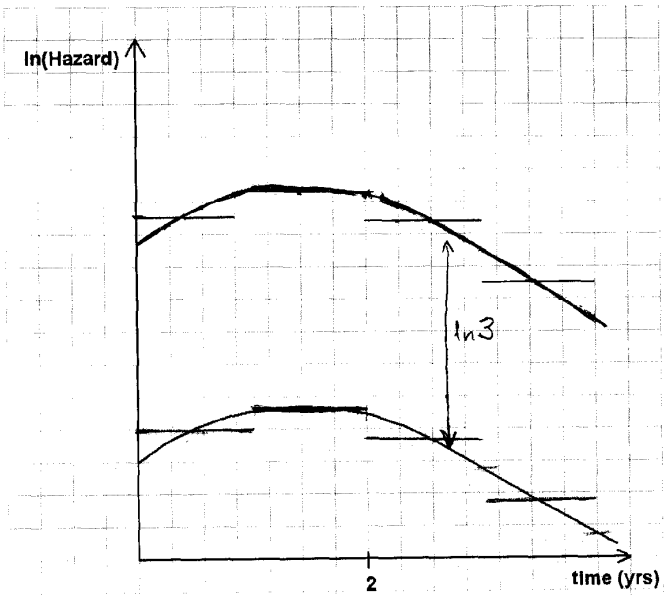
   (b) If you have a well-defined cohort and would like to collect additional information that would be expensive or time-consuming to collect for the whole cohort, such as blood samples. But it requires that the outcome is not too common, or else almost the whole cohort would be used anyway.

   (c) For each cohort member who experience the event, everyone who is AT RISK at that event time are eligible controls. This means that controls can serve as controls for many cases and controls might also became a case (experience the event later on).

6. Question 6

   (a) See graph

   (b) See graph



   (c) If time is splitted into very small time bands in the Poisson model. (It would give the exact same estimate if time was split on every event time.)

   (d) If you have two time-scales.

   (e) Include interaction between the time bands and the variable, than test if the interaction is statistically significant.