# 250C Assignment 5

*Team Awesome (aka Steph and Shelley)*

*4/30/2019*

## Bootstrap confidence interval for RAP

### 1. Write out the statistical model that you fit to calculate the RAP in terms of the parameters (e.g. $\beta$-coefficients). What is the expression for the RAP in terms of the model parameters? (15 points)

The model that we fit was a cox proportional hazards model for time to incident CVD as a function of BMI, age, sex, education and current smoking status. BMI and education were each categorized into 4 categories (with appropriate indicators.)

$$\begin{aligned} log[h(t|\boldsymbol{x})] \quad &= log[h_0(t)] + BMI_{underweight} \times \beta_1 + BMI_{overweight} \times \beta_2 + BMI_{obese} \times \beta_3 + \\ &\quad age \times \beta_4 + sex \times \beta_5 + highschool \times \beta_6 + \\ &\quad somecollege \times \beta_7 + morethancollege \times \beta_8 + currentsmoking \times \beta_9 \end{aligned}$$

The RAP (rate advancement period) of how much obesity increases the risk of CVD with age is the ratio of the log hazard ratio of obesity on CVD, divided by the log hazard ratio of age on CVD. Thus,

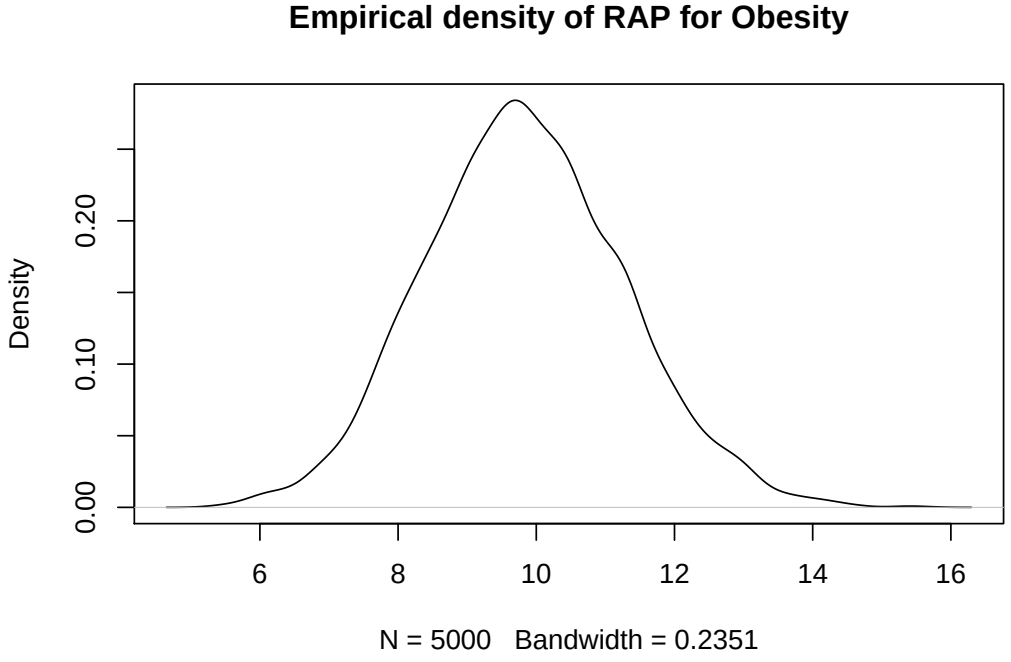
$$RAP_{\text{obesity by age}} = \frac{logHR_{\text{obesity}}}{logHR_{\text{age}}} = \frac{\beta_{\text{obesity}}}{\beta_{\text{age}}} = \frac{\beta_3}{\beta_4}$$

### 2. What is the value for the RAP for the obese vs. normal weight exposure level and its 95% CI from the delta method? Offer an interpretation of this effect measure. (15 points)

Based on the delta method analysis, the RAP for obese vs. normal weight participants is 9.83 (7, 12.67) . This implies that for a one year increase in age, the hazard of CVD based on obesity increases by 9.83. If we were to take many samples of participants from the same underlying population, and calculate a 95% confidence interval from each sample using these methods, 95% of those confidence intervals should contain the true population value for the RAP. (7, 12.67) is one such interval.

**3. What is the bootstrapped mean and 95% confidence interval estimate for the RAP for the obesity effect. How does this bootstrapped estimate of the 95% confidence interval compare to the estimate from the delta method (remembering that they are both approximations!). Turn in the density plot for the bootstrapped RAP. (10 points)**

Based on our bootstrap, the RAP is estimated to be 9.83 ( 7.1, 12.81). This estimate is very similar to the estimate from the delta method. See density plot for the bootstrapped RAP below.




## Linear regression model of BMI

**4. Using the R code provided, complete Table 1. (20 points)**

Table 1. Posterior means and 95% credible intervals for slope coefficient from linear regression model of BMI on age, sex, and education level.

| Variable | Vague Prior | Informative Prior 1* | Informative Prior 2† |
|---|---|---|---|
| Age(per year increase) | | | |
| Female sex (vs. male) | | | |
| High school education (vs. <HS) | | | |
| Some college (vs. <HS) | | | |
| College+ (vs. <HS) | | | |
| Current smoker (vs. non) | | | |

*Prior mean for effect of current smoking=100, prior variance = 1000

† Prior mean for effect of current smoking=100, prior variance = 0.1225

**5. What seems to be more influential, varying the prior mean, or the prior variance? In one sentence, briefly explain what you think is happening? (10 points)**

**6. Using the trace plots, density plots and autocorrelation plots (focus on 1st chain) from the diagnostics for the first model ("Vague prior"), briefly describe any evidence of convergence (or lack of convergence) that you see. Attach these plots (2 pages for trace/density plots; 1 page for autocorrelation plots). (20 points)**

**7. From the results of the Geweke test, is there evidence for lack of convergence? Justify your answer. (10 points)**