9.3 Models with Random Slopes

1. Idea: Remember the model where dummy variable interacted with another continuous variable? Similarly, what if partial effect of a variable depends on unobserved factors that vary by population unit as follows:

   $$y_i = a_i + b_i x_i$$

   In our usual model we have $E(b_i) = \beta_1$ and $a_i = \beta_0 + u_i$. This model is called the **random coefficient model** or random slope model.

   The unobserved slope coefficient b_i is considered a random draw from the population along with observed data (x_i, y_i). So in some sense, for every individual i, you know his/her x_i, y_i and b_i.

2. *Estimation*: In such cases we estimate the average response across individuals denoted by $E(b_i) = \beta_1$ known as the *average partial effect* (APE) or average marginal effect.

3. To understand these models as versions of our usual model we write them as follows:

$$y_i = a_i + b_i x_i = \beta_0 + c_i + \beta_1 x_i + d_i x_i = \beta_0 + \beta_1 x_i + u_i$$

This tells us that we could potentially write the random part of the model into its two components (a) the average partial effect and (b) part of the random effect. For instance $a_i = \beta_0 + c_i$.

(a) Conditional independence assumption implies $E(u_i|x_i) = 0$. Both the models produce unbiased estimates.

(b) The error term i.e. u_i contains heteroskedasticity. If $Var(c_i|x_i) = \sigma_c^2$, $Var(d_i|x_i) = \sigma_d^2$ and $Cov(c_i, d_i|x_i) = 0$, then

$$Var(u_i|x_i) = \sigma_c^2 + \sigma_d^2 x_i^2$$

This could be addressed using methods learnt in chapter 8. Notice that random slope models could be sources of heteroskedasticity. Caution: this does not allow for heteroskedasticity in a_i or b_i. This implies that we cannot distinguish between a random slope model, where the intercept and slope are independent of x_i, and a constant slope model with hereoskedasticity in a_i

4. *Summary*: Estimating random slope models as used in chapter 6 is easy if slopes are independent of the explanatory variables.