Midterm Project

Jiun Lee

1. Introduction

This report aims to use multilevel modeling and evaluate it. The dataset in this report is from [European Social Survey(ESS)](https://ess-search.nsd.no/en/study/172ac431-2a06-41df-9dab-c1fd8f3877e7), which is the research infrastructure consortium in Europe. The behavior and attitudes survey of people over 14 is conducted every two years with 38 countries’ participation. In ESS Data Portal, you can download variable datasets provided by ESS. The dataset in this report is the latest dataset among them. The variables of the dataset consist of respondents’ different attitudes such as nationality, media, politics, understanding of democracy, social relationships, values, and so on. This project is conducted to see the effect of different attributes on happiness and health within countries and between countries.

1. EDA Process

The variables ‘cntry’, ‘happy’, ‘health’, ‘ipmodst’, ‘impsafe’, ‘ipfrule’, and ‘stfmjob’ are chosen for analysis. NA values (Refusal, No answer, Don’t know) are all deleted and the variables are sorted in ascending order during the cleaning process.

This is the description of each variable.

* cntry: respondents’ countries
  + BG(Bulgaria), CZ(Czechia), EE(Estonia), FI(Finland), FR(France), HR(Croatia), HU(Hungary), LT(Lithuania), SI(Slovenia), SK(Slovakia))
* happy: How happy are you
  + 0(extremely unhappy) – 3
* health: How is your health in general?
  + 1(Very bad) – 5(Very good)
* ipmodst: important to be humble and modest, not draw attention
  + 1(not like me at all) - 6(Very much like me)
* impsafe: Important to live in secure and safe surroundings
  + 1(not like me at all) - 6(Very much like me)
* ipfrule: Important to do what is told and follow rules
  + 1(not like me at all) - 6(Very much like me)
* stfmjob: How satisfied are you in your main job

To approximately scan the different variables’ effects on health and happiness, I plotted the association of different variables with health and health within each country.

Calendar

Description automatically generated

* SK(Slovakia) seems to have the biggest negative relationship of ‘ipmodst’ and ‘health’.

Calendar

Description automatically generated

* SK(Slovakia) seems to have biggest positive relationship of ‘impsafe’ and ‘health’.

Chart

Description automatically generated

* All countries generally have a positive effect of ‘stfmjob’ on ‘health’. The satisfaction of a job seems strongly associated with health.

Chart, scatter chart

Description automatically generated

* All countries generally have a significant positive effect of ‘happy’ on ‘health’. Happiness seems strongly associated with health.

Chart

Description automatically generated with medium confidence

* The effects of ‘ipmodst’ on ‘happy’ in BG(Bulgaria) and SI(Slovenia) are different from other countries.

Chart, scatter chart

Description automatically generated

* The effects of ‘impsafe’ on ‘happy’ are significant in HU(Hungary) and SK(Slovakia).

Chart

Description automatically generated

* The effect of ‘ipfrule’ on happy is significant in SK(Slovakia).

Chart, scatter chart

Description automatically generated

* All countries generally have a significant positive effect of ‘stfmjob’ on ‘happy’. The satisfaction of a job seems strongly associated with happiness.

Chart, scatter chart

Description automatically generated

* All countries generally have a significant positive effect of ‘health’ on ‘happy’. The health seems strongly associated with happiness.

1. Multilevel analysis

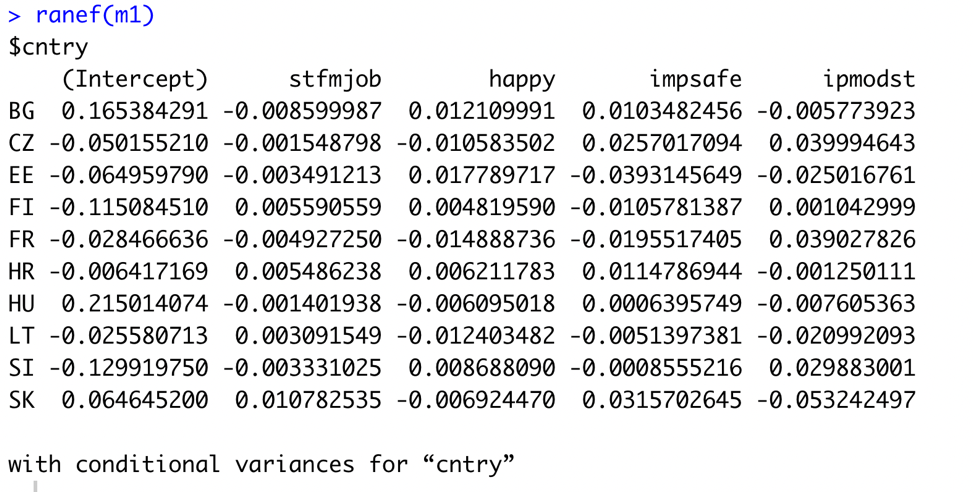
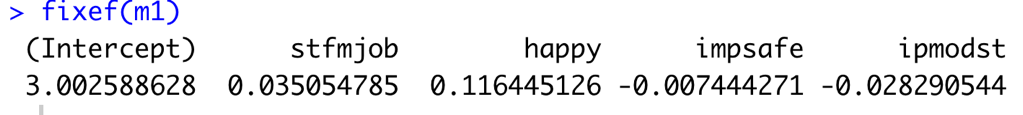
From the EDA process, I verified that there were differences in the effects between countries. To see the random intercept and random slope, we can set the model like this.

Firstly, the response to health.

m1 <- stan\_lmer(health ~ stfmjob + happy + impsafe + ipmodst + (1|cntry) +(0+stfmjob|cntry)+(0+happy|cntry)+(0+impsafe|cntry)+(0+ipmodst|cntry),data=data2)

Table

Description automatically generated with medium confidence



Text

Description automatically generated

* Interpretation of fixed effects (fixef(m1))
  + Intercept: The average value of ‘health’ for every country is 3
  + Stfmjob: The average effect of stfmjob on health for every country is 0.04
  + Happy: The average effect of happy on health for every country is 0.12
  + Impsafe: The average effect of impsafe on health for every country is -0.01
  + Ipmodst: The average effect of ipmodst on health for every country is -0.03.
* Interpretation of random effects (ranef(m1),VarCorr(m1))
  + Intercept: The difference between the average health of all countries and the average health of each country
  + Stfmjob: The difference between all countries’ effect of stfmjob on health and each country’s effect of stfmjob
  + Happy: The difference between all countries’ effect of Happy on health and each country’s effect of Happy
  + Impsafe: The difference between all countries’ effect of Impsafe on health and each country’s effect of Impsafe
  + Ipmodst: The difference between all countries’ effect of Ipmodst on health and each country’s effect of Ipmodst
  + Std.Dev.(Intercept) : There are solid between-country differences in average value of health.(0.20)
  + Std.Dev(stfmjob,happy,impsafe,ipmodst): There are solid between-country differences in association between ‘stfmjob,happy,impsafe,ipmodst’ and health.
* Interpretation of coefficients (coef(m1))
  + Stfmjob: slopes are generally similar between countries.
  + Happy: slopes are generally similar between countries.
  + Impsafe: EE has the biggest negative impact and SK has the biggest positive impact on health among countries.
  + Ipmodst: SK has the biggest negative impact on health among countries.

To check the model’s plausibility, we can use ICC(Intraclass correlation coefficient). If the ICC of each variable is over 0.1, we can consider the use of the multilevel model.

Text

Description automatically generatedText

Description automatically generated

The ICC of each variable are all over 0.1. The multilevel model ‘m1’ is plausible.

Also, we can use MCMC(Markov Chain Monte Carlo) for testing single parameters.

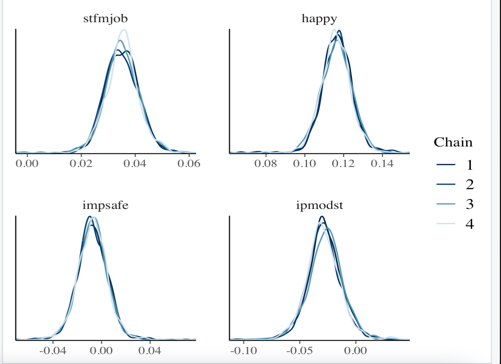
* mcmc\_intervals: Confidence Interval of each parameter

Chart

Description automatically generated

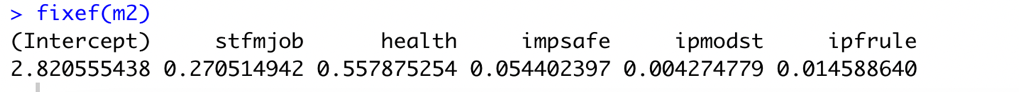
* mcmc areas, mcmc\_dens\_overlay

A picture containing chart

Description automatically generated

Now, try the response to happiness.

Text

Description automatically generated

Text

Description automatically generated

Text

Description automatically generated

* Interpretation of fixed effects (fixef(m2))
  + Intercept: The average value of happiness for every country is 2.82
  + Stfmjob: The average effect of stfmjob on happiness for every country is 0.27
  + health: The average effect of health on happiness for every country is 0.56
  + Impsafe: The average effect of impsafe on happiness for every country is 0.05
  + Ipmodst: The average effect of ipmodst on happiness for every country is 0.004
  + ipfrule: The average effect of ipfrule on happiness for every country is 0.01
* Interpretation of random effects (ranef(m2),VarCorr(m2))
  + Intercept: The difference between the average happiness of all countries and the average happiness of each country
  + Stfmjob: The difference between all countries’ effect of stfmjob on happiness and each country’s effect of stfmjob
  + health: The difference between all countries’ effect of health on happiness and each country’s effect of health
  + Impsafe: The difference between all countries’ effect of Impsafe on happiness and each country’s effect of Impsafe
  + Ipmodst: The difference between all countries’ effect of Ipmodst on happiness and each country’s effect of Ipmodst
  + Ipfrule: The difference between all countries’ effect of Ipmodst on happiness and each country’s effect of ipfrule
  + Std.Dev.(Intercept) : There are solid between-country differences in average value of happiness.(1.20)
  + Std.Dev(stfmjob,happy,impsafe,ipmodst,ipfrule): There are solid between-country differences in the association between ‘stfmjob, health, impsafe, ipmodst’ and happiness.
* Interpretation of coefficients (coef(m2))
  + Intercept: FR has the biggest average happiness among countries.
  + Stfmjob: slopes are generally similar between countries.
  + health: slopes are generally similar between countries.
  + Impsafe: SK has the biggest positive impact on happiness among countries.
  + Ipmodst: BG has the biggest negative impact on happiness among countries.
  + Ipfrule: HU has the biggest negative impact and SK has the biggest positive impact on happiness among countries.
* ICC

Text

Description automatically generatedText

Description automatically generated

The ICC of each variables are all over 0.1. The multilevel model ‘m2’ is plausible.

Also, we can use MCMC(Markov Chain Monte Carlo) for testing single parameters.

* mcmc\_intervals: Confidence Interval of each parameter

Chart, scatter chart

Description automatically generated

* mcmc areas, mcmc\_dens\_overlay

A picture containing chart

Description automatically generated

Diagram

Description automatically generated

1. Result

In model “m1”, we could find out that satisfaction from a job and happiness are significantly related to health in every country with a similar amount of effect. The countries’ average health is similar as we can see from each intercept. However, it seemed that people who think it’s important to live in secure and safe surroundings are less healthy in Estonia but are healthier in Slovakia. Also, people who think it’s important to be humble and modest tend to live less healthy in Slovakia

In model “m2”, we could also find out that satisfaction from a job and health are significantly related to happiness in every country with a similar amount of effect. But the average happiness in France is the biggest among other countries. Also, people in Slovakia who think it’s important to live in secure and safe surroundings are happier. In Bulgaria, people who find it important to be humble are less happy. People who think it’s important to follow rules are less happy in Hungary but are happier in Slovakia.

In conclusion, this report has handled the process of multilevel regression. We confirm that there are differences among countries with multiple parameters.