《国际关系定量分析基础》2020 秋季

第四次作业 (共计 100 分)

学生姓名　　　　学生学号

截止时间：**2020 年 11 月 30 日 11：59 am**

注意事项：

- 作业在网络学堂提交
- 请将 `Chunk` 中的 `eval=FALSE` 改为 `eval=TRUE` 再 `knit`
- 请将文件解压缩后，直接在 R Markdown 文件中完成本次作业
- 学生可以互相讨论作业，但作业必须是自己本人独立完成
- 提交作业的文件名需以 `HW-4-YourName.Rmd`,`HW-4-YourName.pdf` 或者 `HW-4-YourName.html`，请将 `YourName` 替换为你的姓名。(若 R Markdown 出现无法 `knit` 为 `pdf` 情况，则使用 `bookdown::html_document2`：会生成为 `html`)
- 请显示每道题的 R Code 于 pdf 中，注重 Code 的整洁性和可读性，可参考Google's R Style Guide

本次作业需要的数据已经提供，请将数据与 `HW-4-YourName.Rmd` 放在同一工作路径的文件夹内

```
load("HW-4.RData")
```

本次作业的 `HW-4.RData` 数据包括 `broz_et_al` 和 `map` 两个数据，其中数据 `broz_et_al` 来自于 J. Lawrence Broz, Zhiwen Zhang 以及 Gaoyang Wang 发表于《国际组织》2020 年第 2 期的复制数据（见 J. Lawrence Broz, Zhiwen Zhang, and Gaoyang Wang, "Explaining Foreign Support for China's Global Economic Leadership," *International Organization*, Vol. 74, Summer 2020, pp. 417–52）。该文章检验了世界各国对中国一带一路峰会态度的影响因素。

其中部分变量如下：

- `countryname`: The Correlates of War (COW) country name
- `attendance`: "DV=Attendance"（1= Yes; 0 otherwise）
- `obor_nations`: "One Belt, One Road Position"（1= Yes; 0 otherwise）
- `ftas`: "FTA with China"（1= Yes; 0 otherwise）

- `bits`: “BIT with China” （1= Yes; 0 otherwise）
- `fc_dummy_cumcount_bank_s1_9016`: “Financial Crises” （1= Yes; 0 otherwise）
- `ka_open_sd9016`: “Variability of Capital Account Policy”
- `mean_portfolio_vol`: “Volatility of Portfolio Outflows”
- `imf_dummy_unrest_index_sum_9017`: “Social Unrest During IMF Programs”
- `wto_cases_cumulcount9516`: “WTO Complaints Against The U.S.”
- `imf_governance_deficit_usd_2015`: “IMF Governance Deficit”

表-1 统计了部分变量的统计分布特征。请利用 `broz_et_al` 数据完成以下各题。

表 1: 变量的描述性统计

Statistic	N	Mean	Median	Max	Min	St. Dev.
attendance	192	0.151	0	1	0	0.359
obor_nations	192	0.349	0	1	0	0.478
ftas	192	0.125	0	1	0	0.332
bits	192	0.547	1	1	0	0.499
fc_dummy_cumcount_bank_s1_9016	162	4.358	0.000	27.000	0.000	6.475
ka_open_sd9016	178	0.140	0.116	0.431	0.000	0.105
mean_portfolio_vol	93	0.162	0.011	7.324	0.00002	0.797
imf_dummy_unrest_index_sum_9017	192	13.625	2	188	0	26.984
wto_cases_cumulcount9516	192	0.651	0	17	0	2.246
imf_governance_deficit_usd_2015	184	0.066	0.052	1.924	−7.653	0.630

数据可视化

1.(20 分) 利用 `map` 数据和 `ggplot` 绘制一幅各国参加一带一路峰会的地图（提示：参与国家的变量为 `attendance`)。根据地图，你发现参与国家的地理分布有何模式和特征?

```
ggplot(data= map)+
  geom_polygon(aes(x = long, y = lat, group = group, fill = attendance)) +
  coord_fixed() +
  scale_fill_manual(values = c("blue", "red"), na.value = "gray")+
   theme(line = element_blank(),
    legend.position = "right",
    legend.title=element_blank(),
```

```
    panel.border=element_blank(),
    panel.grid=element_blank(),
    axis.ticks=element_blank(),
    axis.text=element_blank())
```

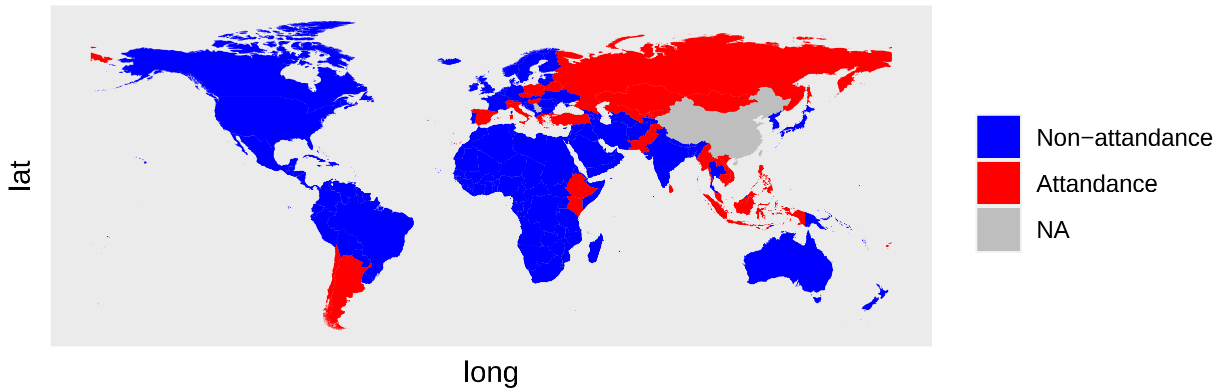

图 1: Coutries attended BRI Submit

估计二分类因变量模型

2.（15 分）利用线性概率模型（linear probability model）估计以下模型，将其命名为 m1, 并制作一个回归表格。根据回归表格，解读对应的回归系数 β_1。

LPM：

$$\begin{aligned} attendance = & \beta_0 + \beta_1 * \text{One Belt, One Road Position}+ \\ & \beta_2 * \text{FTA with China}+ \\ & \beta_3 * \text{BIT with China} + \epsilon \end{aligned}$$

```
m1 <- lm(attendance ~ obor_nations + ftas + bits, data = broz_et_al)
library(stargazer)
stargazer(m1, type = "latex", header = FALSE)
```

- 需要主要 LPM 的 β 可以直接解释为概率的变化：在控制其他变量不变的情况下，一带一路沿线国家参加峰会的概率比非一带一路沿线国家的概率高 0.13.

3.（15 分）利用 Logit 回归估计以下模型，将其命名为 m2, 并制作一个回归表格。根据回归表格，解读对应的回归系数 β_1 和 β_2。

Logit Model：

表 2:

	Dependent variable:
	attendance
obor_nations	0.129^{**}
	(0.055)
ftas	0.304^{***}
	(0.074)
bits	0.126^{**}
	(0.051)
Constant	-0.001
	(0.035)
Observations	192
R^2	0.212
Adjusted R^2	0.200
Residual Std. Error	0.321 (df = 188)
F Statistic	16.906^{***} (df = 3; 188)
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

$$\begin{aligned}
P(attendance = 1) = &\beta_0 + \beta_1 * \text{Financial Crises}+ \\
&\beta_2 * \text{Variability of Capital Account Policy}+ \\
&\beta_3 * \text{Volatility of Portfolio Outflows}+ \\
&\beta_4 * \text{Social Unrest During IMF Programs}+ \\
&\beta_5 * \text{WTO Complaints Against The U.S.}+ \\
&\beta_6 * \text{IMF Governance Deficit} + \epsilon
\end{aligned}$$

```r
m2 <- glm(attendance ~  fc_dummy_cumcount_bank_s1_9016 +
            ka_open_sd9016 + mean_portfolio_vol + imf_dummy_unrest_index_sum_9017 +
            wto_cases_cumulcount9516 + imf_governance_deficit_usd_2015,
          family = binomial(link = "logit"), data = broz_et_al)
stargazer(m2, type = "latex",  header = FALSE)
```

- 对于 Logit 模型，直接对回归系数的解释需要在 log-odds 层面，更好的回答是将其转换为概率：其他条件不变的情况下，发生金融危机的国家出席峰会的概率为：$\frac{exp(0.037)}{1+exp(0.037)} = 0.5092489$

4.（20 分）利用线性概率模型（linear probability model）、Logit 和 Probit 分别估计以下模型，将其命名为 `m3`, `m4`, `m5`，同时制作一个回归系数图。

$$\begin{aligned}
P(attendance = 1) = &\beta_0 + \beta_1 * \text{One Belt, One Road Position}+ \\
&\beta_2 * \text{FTA with China}+ \\
&\beta_3 * \text{BIT with China}+ \\
&\beta_4 * \text{Financial Crises}+ \\
&\beta_5 * \text{Variability of Capital Account Policy}+ \\
&\beta_6 * \text{Volatility of Portfolio Outflows}+ \\
&\beta_7 * \text{Social Unrest During IMF Programs}+ \\
&\beta_8 * \text{WTO Complaints Against The U.S.}+ \\
&\beta_9 * \text{IMF Governance Deficit} + \epsilon
\end{aligned}$$

```r
#LPM
m3 <- lm(attendance ~ obor_nations + ftas + bits + fc_dummy_cumcount_bank_s1_9016 +
            ka_open_sd9016 + mean_portfolio_vol + imf_dummy_unrest_index_sum_9017 +
            wto_cases_cumulcount9516 + imf_governance_deficit_usd_2015,
           data = broz_et_al)
#probit
m4 <- glm(attendance ~ obor_nations + ftas + bits + fc_dummy_cumcount_bank_s1_9016 +
```

表 3:

	Dependent variable:
	attendance
fc_dummy_cumcount_bank_s1_9016	0.037
	(0.048)
ka_open_sd9016	2.044
	(2.584)
mean_portfolio_vol	0.517
	(0.424)
imf_dummy_unrest_index_sum_9017	0.014
	(0.010)
wto_cases_cumulcount9516	−0.109
	(0.132)
imf_governance_deficit_usd_2015	0.144
	(0.495)
Constant	-2.023^{***}
	(0.596)
Observations	85
Log Likelihood	−41.836
Akaike Inf. Crit.	97.671
Note:	*p<0.1; **p<0.05; ***p<0.01

```
            ka_open_sd9016 + mean_portfolio_vol + imf_dummy_unrest_index_sum_9017 +
            wto_cases_cumulcount9516 + imf_governance_deficit_usd_2015,
           family = binomial(link = "probit"), data = broz_et_al)
#logit
m5 <- glm(attendance ~ obor_nations + ftas + bits + fc_dummy_cumcount_bank_s1_9016 +
            ka_open_sd9016 + mean_portfolio_vol + imf_dummy_unrest_index_sum_9017 +
            wto_cases_cumulcount9516 + imf_governance_deficit_usd_2015,
           family = binomial(link = "logit"), data = broz_et_al)

library(dotwhisker)
dwplot(list(m3, m4, m5), conf.level = .95, show_intercept = TRUE) +
        theme_bw() +
        ggtitle("Coefficient Plot ")+
        scale_color_discrete(name="Model Name",
                             labels=c("Model 3","Model 4","Model 5"))
```

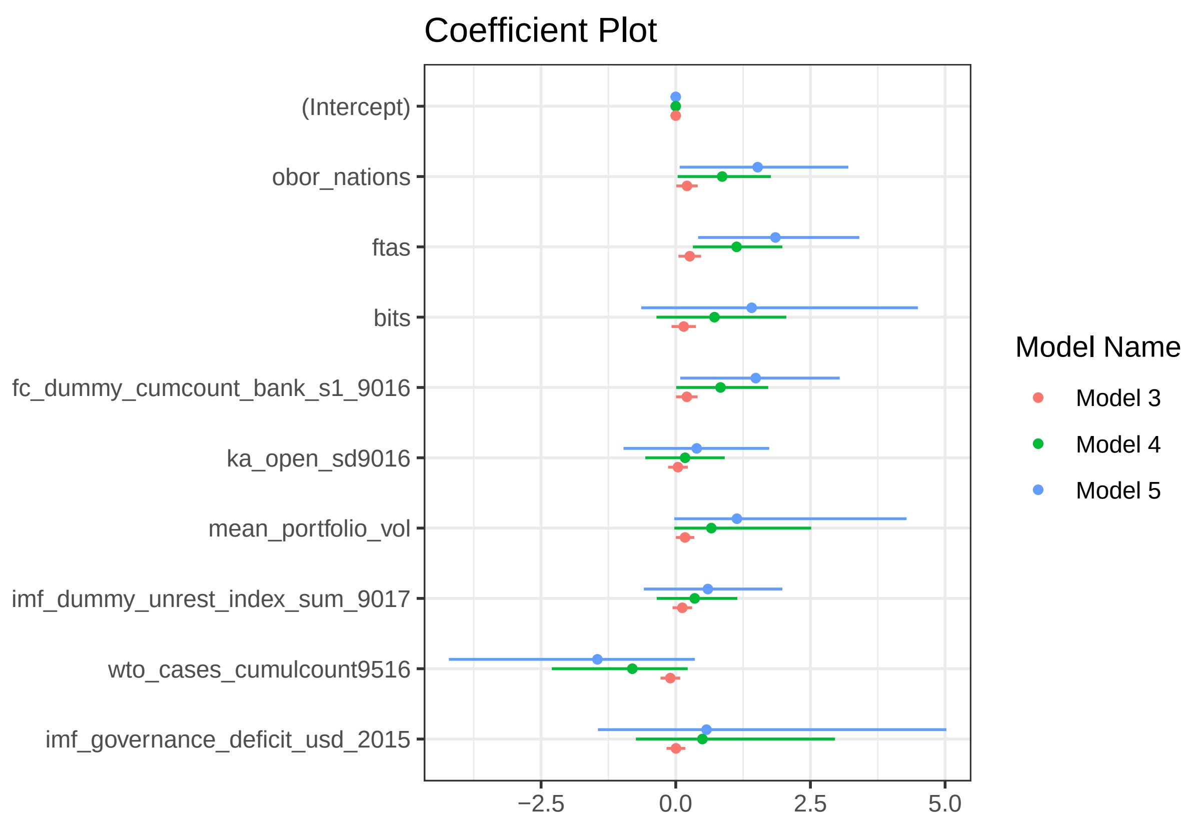

图 2: Coefficient Plots

5.（30 分）根据 `broz_et_al` 数据和 Logit 模型 5，计算澳大利亚 (Australia)、法国（France）和

印度（India）支持中国经济领导的概率（包括 95% 置信区间）。（提示：根据模型 5 结果以及这三国对应变量取值，计算预测概率及其置信区间）

```
au_newdata <- broz_et_al %>%
          filter(countryname == "Australia") %>%
          data.frame()

fr_newdata <- broz_et_al %>%
          filter(countryname == "France") %>%
          data.frame()

ind_newdata <- broz_et_al %>%
          filter(countryname == "India") %>%
          data.frame()

au_prob <- predict(m5, newdata = au_newdata, se.fit = TRUE)
fr_prob <- predict(m5, newdata = fr_newdata, se.fit = TRUE)
ind_prob <- predict(m5, newdata = ind_newdata, se.fit = TRUE)

au_upr <- au_prob$fit + (1.96*au_prob$se.fit)
au_lwr <- au_prob$fit - (1.96*au_prob$se.fit)
au_fit <- au_prob$fit

fr_upr <- fr_prob$fit + (1.96*fr_prob$se.fit)
fr_lwr <- fr_prob$fit - (1.96*fr_prob$se.fit)
fr_fit <- fr_prob$fit

ind_upr <- ind_prob$fit + (1.96*ind_prob$se.fit)
ind_lwr <- ind_prob$fit - (1.96*ind_prob$se.fit)
ind_fit <- ind_prob$fit

# function log-odds -> probability
getpro <- function(x) { 1/(1 + exp(-x))}

probs <- data.frame(country = c("Australia", "France", "India"),
                    fitted = c(getpro(au_fit),getpro(fr_fit), getpro(ind_fit)),
                    lwr = c(getpro(au_lwr),getpro(fr_lwr), getpro(ind_lwr)),
```

```
                    upr = c(getpro(au_upr),getpro(fr_upr), getpro(ind_upr)))


kable(probs)
```

country	fitted	lwr	upr
Australia	0.2149431	0.0386330	0.6510101
France	0.0281778	0.0006595	0.5602455
India	0.3333695	0.0387331	0.8612353