

Hap 823

Instructor: Farrokh Alemi

By: Akhil Anto

Team Mate: Elina Bajracharaya

Question 5: The objective of this analysis is to find response to antidepressants. You can select one of the antidepressants.

1. These data come from STAR*D experiment conducted by National Institute of Medicine. Read about the study protocol. [Protocol](#)
2. Download data. Use instructor's last name as password. [Data](#)
3. The data are report bi-weekly or monthly. There are 22254 records for about 4,000 patients. Organize the data so there is one row for each patient. Exclude patients who drop out of the study or have remission prior to the antidepressant you have focused on.
 - **Focus:** The data report on citalopram, bupropion, mirzapine, buspirone, lithium, nortriptyline, sertraline, thyroid, tranylcypromine, and venlafaxine. Please focus the analysis on only one of the antidepressants or a combination of two antidepressants taken simultaneously. For the time being ignore the dose of the medication.
 - **Exclusions:** Patients received treatment at 4 levels, depending on whether the initial antidepressant was effective. Exclude all patients whose depression went into remission prior to taking the antidepressant you have selected.
 - **Treatment:** If the patient has taken the antidepressant at any time during the study period, then mark it as 1, otherwise 0. Notice that some patients have taken the medication and others have not.
 - **Covariates:** For the covariates, include gender, risk of suicide, heart, vascular, haematopoietic, eyes ears nose throat larynx, gastrointestinal, renal, genitourinary, musculoskeletal Integument, neurological, psychiatric illness, respiratory, liver, endocrine, alcohol, amphetamine, cannabis use, opioid use, panic, specific phobia, social phobia, OCD, PTSD, anxiety, borderline personality, dependent personality, antisocial personality, paranoid personality, personality disorder, anorexia, bulimia, and cocaine use. If the variable is ever present assume that it is present. Exclude any variable that is not present for any of the patients.
 - **Outcome:** The medication is considered to have caused the remission if (a) the patient has had remission and (b) the patient has not switched from the antidepressant to other medications before remission is observed.
4. Balance the data to remove the effects of covariates. Show visually that you have successfully balanced the data. Use the following steps to accomplish this:
 - **Calculate Propensity Score:** Calculate the propensity of taking the antidepressant. Regress taking of the antidepressant on the covariates.
 - **Weights:** Calculate inverse propensity weights
 - **Verify Balance:** Verify that weighted regression removes the effects of all covariates. Regress the antidepressants on the covariates and verify that none have a statistically significant effect on selection of the antidepressant.
 - **Estimate Impact on Response:** Regress response to the antidepressant on the covariates and taking the antidepressant.
5. Describe how well the model was balanced and how well the impact of antidepressant was estimated

Answer

Linear Regression

Call:

```
lm(formula = Remission ~ Buspirone + Gender + RiskOfSuicide +  
Heart + Vascular + Haematopoietic + Eyes_Ears_Nose_Throat_Larynx +  
Gastrointestinal + Renal + Genitourinary + Musculoskeletal_Integument +  
Neurological + Psychiatric_Illness + Respiratory + Liver +  
Endocrine + Alcohol + Amphetamine + Cannabis + Opioid + Panic +  
Specific_Phobia + Social_Phobia + OCD + PTSD + Anxiety +  
Borderline_Personality + Dependent_Personality + Antisocial_Personality +
```

Paranoid_Personality + Personality_Disorder + Anorexia +
Bulimia + Cocaine, data = mydata)

Residuals:

Min	1Q	Median	3Q	Max
-0.4782	-0.2832	-0.2324	0.6122	1.0759

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	0.318277	0.006529	48.746	< 2e-16	***
Buspirone	0.069795	0.013859	5.036	4.79e-07	***
Gender2	0.005076	0.006310	0.804	0.421127	
RiskOfSuicide	-0.199693	0.026860	-7.435	1.09e-13	***
Heart	-0.022545	0.008407	-2.682	0.007327	**
Vascular	-0.036969	0.007324	-5.048	4.51e-07	***
Haematopoietic	-0.012666	0.010539	-1.202	0.229445	
Eyes_Ears_Nose_Throat_Larynx	-0.010549	0.006217	-1.697	0.089722	.
Gastrointestinal	-0.022976	0.006329	-3.630	0.000284	***
Renal	0.011311	0.012762	0.886	0.375482	
Genitourinary	0.001817	0.007364	0.247	0.805083	
Musculoskeletal_Integument	-0.026309	0.006315	-4.166	3.11e-05	***
Neurological	-0.030382	0.006881	-4.416	1.01e-05	***
Psychiatric_Illness	-0.001942	0.009218	-0.211	0.833134	
Respiratory	-0.030418	0.006466	-4.705	2.56e-06	***
Liver	-0.041923	0.010435	-4.018	5.90e-05	***
Endocrine	0.013218	0.007246	1.824	0.068130	.
Alcohol	-0.021368	0.014493	-1.474	0.140414	
Amphetamine	-0.012912	0.041175	-0.314	0.753831	
Cannibis	-0.024647	0.027501	-0.896	0.370148	
Opioid	-0.003280	0.046844	-0.070	0.944174	
Panic	-0.018373	0.014876	-1.235	0.216831	
Specific_Phobia	0.051098	0.040892	1.250	0.211469	
Social_Phobia	-0.039746	0.016070	-2.473	0.013393	*
OCD	0.064444	0.039116	1.648	0.099469	.
PTSD	-0.110407	0.013326	-8.285	< 2e-16	***
Anxiety	-0.027753	0.012612	-2.201	0.027779	*
Borderline_Personality	-0.096774	0.040835	-2.370	0.017802	*
Dependent_Personality	-0.099289	0.052345	-1.897	0.057865	.
Antisocial_Personality	-0.071566	0.065823	-1.087	0.276941	
Paranoid_Personality	-0.220690	0.177723	-1.242	0.214337	
Personality_Disorder	-0.040304	0.051232	-0.787	0.431460	
Anorexia	-0.051252	0.143554	-0.357	0.721077	
Bulimia	0.139776	0.060426	2.313	0.020722	*
Cocaine	0.034463	0.021001	1.641	0.100800	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4295 on 22178 degrees of freedom
Multiple R-squared: 0.02362, Adjusted R-squared: 0.02212
F-statistic: 15.78 on 34 and 22178 DF, p-value: < 2.2e-16

Propensity Score Using Logistic Regression

Call:

```
glm(formula = Buspirone ~ Gender + RiskOfSuicide + Heart + Vascular +  
Haematopoietic + Eyes_Ears_Nose_Throat_Larynx + Gastrointestinal +  
Renal + Genitourinary + Musculoskeletal_Integument + Neurological +  
Psychiatric_Illness + Respiratory + Liver + Endocrine + Alcohol +  
Amphetamine + Cannibis + Opioid + Panic + Specific_Phobia +  
Social_Phobia + OCD + PTSD + Anxiety + Borderline_Personality +  
Dependent_Personality + Antisocial_Personality + Paranoid_Personality +  
Personality_Disorder + Anorexia + Bulimia + Cocaine, family = "binomial",  
data = mydata)
```

Deviance Residuals:
 Min 1Q Median 3Q Max
 -0.5723 -0.3286 -0.2985 -0.2699 2.9474

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-2.98993	0.07157	-41.775	< 2e-16	***
Gender2	-0.25602	0.06990	-3.663	0.00025	***
RiskOfSuicide	-0.73781	0.41598	-1.774	0.07612	.
Heart	-0.19296	0.09709	-1.987	0.04687	*
Vascular	0.11197	0.08031	1.394	0.16325	
Haematopoietic	0.04184	0.11610	0.360	0.71858	
Eyes_Ears_Nose_Throat_Larynx	0.17215	0.06925	2.486	0.01292	*
Gastrointestinal	-0.04829	0.07069	-0.683	0.49448	
Renal	-0.02022	0.13873	-0.146	0.88412	
Genitourinary	-0.14076	0.08451	-1.666	0.09580	.
Musculoskeletal_Integument	-0.19963	0.07127	-2.801	0.00510	**
Neurological	0.15715	0.07458	2.107	0.03512	*
Psychiatric_Illness	0.02985	0.10419	0.287	0.77449	
Respiratory	0.29625	0.07004	4.230	2.34e-05	***
Liver	0.02673	0.11493	0.233	0.81606	
Endocrine	0.12087	0.07898	1.530	0.12592	
Alcohol	-0.41872	0.18291	-2.289	0.02206	*
Amphetamine	0.91232	0.40340	2.262	0.02372	*
Cannibis	-0.41167	0.38906	-1.058	0.29000	
Opioid	-13.32932	326.57719	-0.041	0.96744	
Panic	-0.10713	0.17758	-0.603	0.54634	
Specific_Phobia	-0.03045	0.44841	-0.068	0.94586	
Social_Phobia	0.22176	0.17367	1.277	0.20165	
OCD	0.45623	0.36371	1.254	0.20970	
PTSD	0.15575	0.14179	1.098	0.27201	
Anxiety	-0.31096	0.15951	-1.949	0.05124	.
Borderline_Personality	0.31052	0.40077	0.775	0.43845	
Dependent_Personality	-14.29934	446.52547	-0.032	0.97445	
Antisocial_Personality	-14.47118	593.82834	-0.024	0.98056	
Paranoid_Personality	-14.57253	1604.64909	-0.009	0.99275	
Personality_Disorder	-14.41464	467.78133	-0.031	0.97542	
Anorexia	-14.43478	1317.92657	-0.011	0.99126	
Bulimia	-14.42094	549.75084	-0.026	0.97907	
Cocaine	-12.30392	180.83433	-0.068	0.94575	

 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 8222.7 on 22212 degrees of freedom
 Residual deviance: 8089.8 on 22179 degrees of freedom
 AIC: 8157.8

Number of Fisher Scoring iterations: 16

Attaching prop score to the data set

```
mydata$psvalue<-predict(propscore,type="response")
```

Calculating weights & attaching it to Data Set

```
mydata$weight<-ifelse(mydata$Remission==1,1/mydata$psvalue,1/(1-  
mydata$psvalue))
```

This is not correct; the weights must be assigned based on treatment variable and not the outcome variable. Remission is the outcome variable. Weights must be assigned based on the antidepressant.

Weighted Regression Analysis

Call:

```
lm(formula = Remission ~ Buspirone + Gender + RiskOfSuicide +  
Heart + Vascular + Haematopoietic + Eyes_Ears_Nose_Throat_Larynx +  
Gastrointestinal + Renal + Genitourinary + Musculoskeletal_Integument +  
Neurological + Psychiatric_Illness + Respiratory + Liver +  
Endocrine + Alcohol + Amphetamine + Cannibis + Opioid + Panic +  
Specific_Phobia + Social_Phobia + OCD + PTSD + Anxiety +  
Borderline_Personality + Dependent_Personality + Antisocial_Personality +  
Paranoid_Personality + Personality_Disorder + Anorexia +  
Bulimia + Cocaine, data = mydata, weights = weight)
```

Weighted Residuals:

```
      Min      1Q   Median      3Q      Max  
-1.08488 -1.02510 -1.01905 -0.00015  1.63261
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	9.998e-01	9.369e-05	10670.851	< 2e-16	***
Buspirone	-9.161e-02	9.997e-03	-9.164	< 2e-16	***
Gender2	-2.624e-07	3.681e-05	-0.007	0.994312	
RiskOfSuicide	4.695e-04	2.404e-04	1.953	0.050870	.
Heart	-5.019e-05	4.746e-05	-1.058	0.290266	
Vascular	-2.440e-05	7.122e-05	-0.343	0.731891	
Haematopoietic	1.473e-04	1.282e-04	1.149	0.250574	
Eyes_Ears_Nose_Throat_Larynx	-8.665e-05	5.998e-05	-1.445	0.148570	
Gastrointestinal	-1.004e-04	5.612e-05	-1.789	0.073626	.
Renal	1.487e-04	1.060e-04	1.403	0.160606	
Genitourinary	2.920e-05	5.971e-05	0.489	0.624902	
Musculoskeletal_Integument	-8.524e-05	5.924e-05	-1.439	0.150196	
Neurological	6.855e-05	3.846e-05	1.782	0.074714	.
Psychiatric_Illness	-1.658e-05	3.453e-05	-0.480	0.631175	
Respiratory	4.869e-05	4.829e-05	1.008	0.313369	
Liver	-7.085e-05	3.823e-05	-1.853	0.063891	.
Endocrine	-1.390e-04	6.303e-05	-2.205	0.027461	*
Alcohol	-5.170e-05	4.220e-05	-1.225	0.220522	
Amphetamine	-8.204e-05	6.456e-05	-1.271	0.203828	
Cannibis	-7.096e-05	5.418e-05	-1.310	0.190352	
Opioid	3.789e-04	1.577e-04	2.403	0.016285	*
Panic	-5.625e-05	1.104e-04	-0.510	0.610341	
Specific_Phobia	-1.131e-01	3.152e-02	-3.588	0.000333	***
Social_Phobia	-1.017e-05	1.149e-04	-0.088	0.929518	
OCD	-3.651e-04	1.663e-04	-2.196	0.028089	*
PTSD	1.799e-04	9.275e-05	1.940	0.052416	.
Anxiety	-1.224e-04	6.743e-05	-1.815	0.069491	.
Borderline_Personality	-3.398e-01	4.987e-02	-6.813	9.81e-12	***
Dependent_Personality	3.595e-04	1.580e-04	2.275	0.022893	*
Antisocial_Personality	4.228e-04	1.894e-04	2.232	0.025623	*
Paranoid_Personality	-9.240e-01	3.610e-01	-2.559	0.010494	*
Personality_Disorder	2.629e-04	1.123e-04	2.340	0.019270	*
Anorexia	2.455e-04	1.403e-04	1.750	0.080191	.
Bulimia	3.138e-04	1.240e-04	2.531	0.011391	*
Cocaine	2.098e-04	8.149e-05	2.575	0.010034	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8828 on 22178 degrees of freedom
Multiple R-squared: 0.007241, Adjusted R-squared: 0.005719
F-statistic: 4.758 on 34 and 22178 DF, p-value: < 2.2e-16

