

Stat601 Final Problem 1: Schizophrenia

James Young

December 16, 2019

The Problem

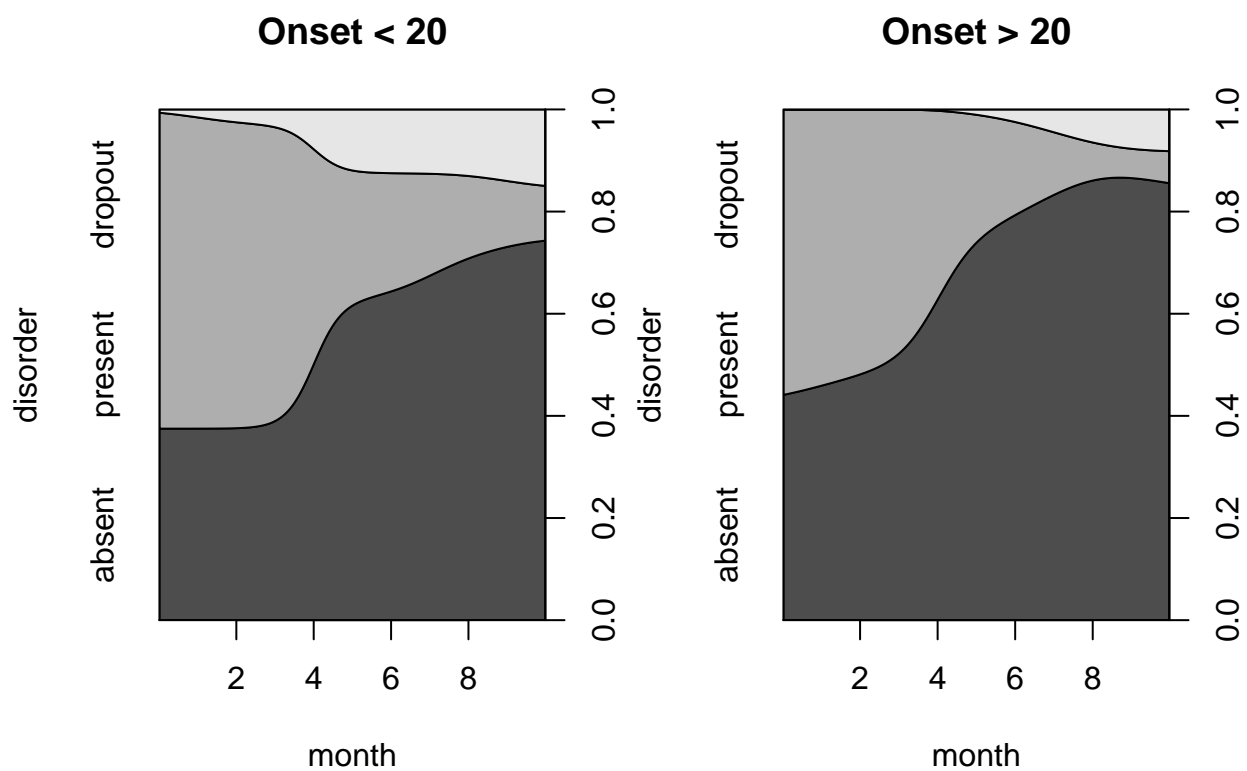
I am given the task of determining if the time of onset of schizophrenia has a significant effect on the progression of the disease. The onset in this dataset is measured as occurring either before or after 20 years of age.

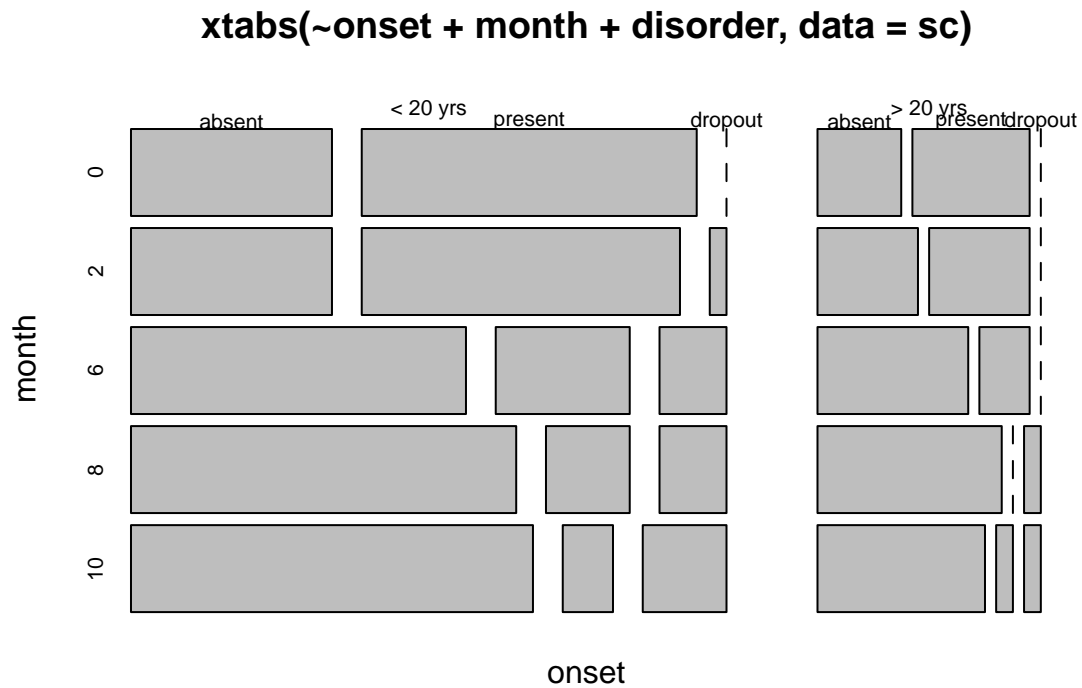
The Data

The data is composed of 4 variables including *subject*, *onset*, *disorder*, and *month*. The response is the factor variable, *disorder*, with options present or absent. There are a few NA's in *disorder* which will be considered dropouts, but we cannot say for sure what reasons they dropped out for. I notice that there are about 3 times as many observations in the early onset group as the late onset group, as well.

Below the table we can see cdplots and mosaic plots visualizing the progression of disease. We see that the early onset group has proportionally more dropouts and that both groups tend to not have schizophrenia by month 10. It appears that the late onset group may get better slightly quicker.

subject	onset	disorder	month
1 : 5	< 20 yrs:160	absent :132	Min. : 0.0
2 : 5	> 20 yrs: 60	present: 72	1st Qu.: 2.0
3 : 5	NA	NA's : 16	Median : 6.0
4 : 5	NA	NA	Mean : 5.2
5 : 5	NA	NA	3rd Qu.: 8.0
6 : 5	NA	NA	Max. :10.0
(Other):190	NA	NA	NA

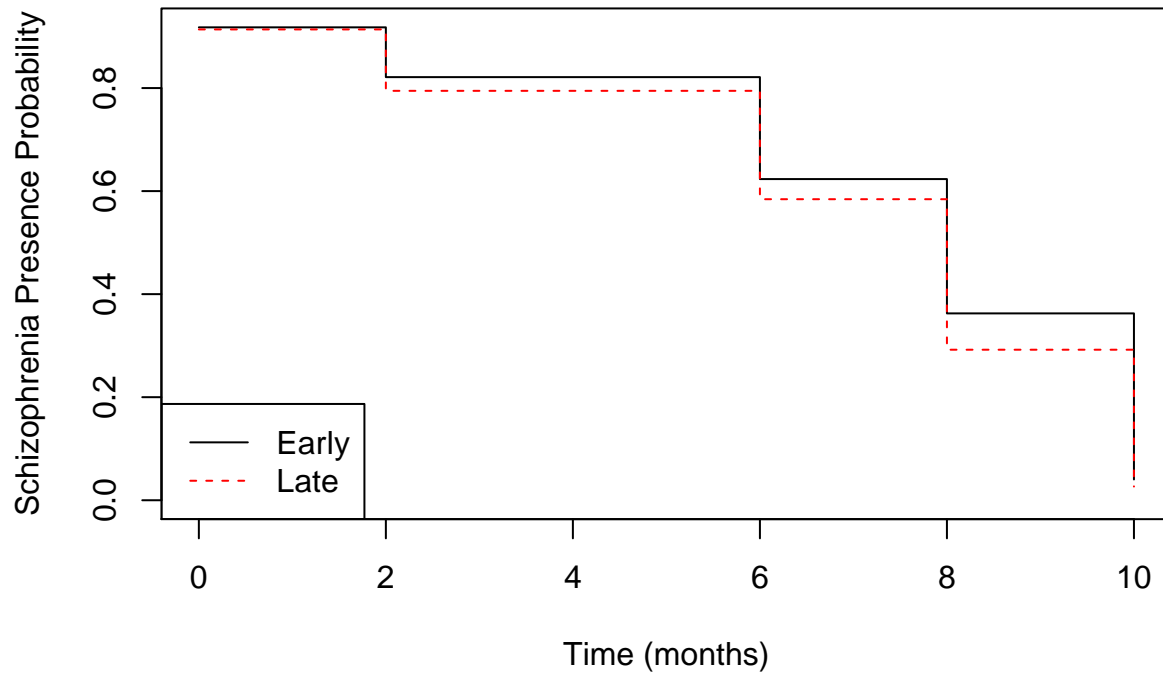




The Model

I used a survival model to visualize and, further below, quantify any difference between the progress of schizophrenia conditional on time of onset. The survival plot tracks time in months on the x-axis and probability of schizophrenia being present on the y-axis. The red dashed line tracks the probability of schizophrenia being present for the late onset group and the black solid line tracks it for the early onset group. Visually, it looks like there is a small difference, with schizophrenic symptoms being absent slightly sooner for the late onset group. This is consistent with the plots above. We will see further below if there is a true statistically significant difference.

Effect of Onset on Schizophrenia Persistence



Does Onset Have a Significant Effect Progression of Schizophrenia?

To determine whether there is statistically significant difference in disease progression conditional on disease onset, I will perform a chi-square test on the survival plots and evaluate the p-value with $\alpha = 0.05$ as the threshold for significance. We can see below that the p-value is greater than 0.05 and can conclude, based on this data, that the onset of schizophrenia does not effect the progression of the disease in a statistically significant way.

```
## Call:
## survdiff(formula = Surv(month, disorder == "present") ~ onset,
##   data = schizophrenia2)
##
## n=204, 16 observations deleted due to missingness.
##
##               N Observed Expected (O-E)^2/E (O-E)^2/V
## onset=< 20 yrs 146      55     51.4    0.259    1.03
## onset=> 20 yrs  58      17     20.6    0.644    1.03
##
## Chisq= 1  on 1 degrees of freedom, p= 0.3
```

Conclusions

Through data exploration I noticed a slight difference between the progression of schizophrenia between early and late onset group. It seemed visually that the late onset group stopped having symptoms of schizophrenia

faster than the early onset group. However, using a more analytical approach where I made survival plots for each group and did a chi-square test to look for a difference, it was apparent that there was no statistically significant difference in disease progression based on time of onset. Based on the data table presented at the beginning, we know there were almost 3 times more observations in the early onset group than the late onset group. Collecting more data on this issue may reveal a different result.