Young American Tobacco Usage

Vladyslav Shuvalov

September 24, 2019

Introduction

Tobacco is consumed in several forms sometimes with the help different tools. Using the 2014 American National Youth Tobacco Survey this report will examine whether regular use of chewing tobacco, snuff or dip is no more common amongst Americans of European ancestry than for Hispanic-Americans and African-Americans, once one accounts for the fact that white Americans more likely to live in rural areas and chewing tobacco is a rural phenomenon. In this report a regular user of chewing tobacco is someone that has consumed chewing tobbacco 6 or more times in the past month. The report will also examine whether the likelihood of having used a hookah or waterpipe on at least one occasion is the same for two individuals of the different sexes, provided their age, ethnicity, and other demographic characteristics are similar.

Methods

Since the hypothesis is testing whether the likelihood of having used a hookah or waterpipe on at least one occasion is the same for two individuals of the different sexes the binomial regression with the count of 1 was chosen to fit the categorical data.

From the smoking dataset the covariates: use of chewing tobacco or snuff, age, race, sex, and demographic were chosen to test the hypothesis.

```
##        Race            Age        Sex       RuralUrban
##  white   :9651   Min.   : 9.00   M:10114   Urban:10631
##  black   :3260   1st Qu.:13.00   F: 9874   Rural: 9357
##  hispanic:5726   Median :14.00
##  asian   : 949   Mean   :14.54
##  native  : 325   3rd Qu.:16.00
##  pacific :  77   Max.   :19.00
##  ever_tobacco_hookah_or_wa
##  Mode :logical
##  FALSE:17177
##  TRUE :2811
##
##
##
```

Fitting a Binomial generalized linear model to the data with a logit link $log(\frac{p}{1-p}) = \beta_0 + \beta_1 X_1 + I_{Urban}\beta_2 X_2 + I_{Sex}\beta_3 X_3 + I_{Hispanic}\beta 4 X_4 + I_{Black}\beta_5 X_5$

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-7.845	0.184	-42.707	0.000
Age	0.409	0.011	35.748	0.000
RuralUrbanRural	-0.389	0.044	-8.813	0.000
SexF	0.039	0.043	0.917	0.359
Raceblack	-0.629	0.070	-8.945	0.000
Racehispanic	0.351	0.048	7.285	0.000

	Estimate	Std. Error	z value	Pr(>|z|)
Raceasian	-0.629	0.118	-5.352	0.000
Racenative	0.181	0.188	0.962	0.336
Racepacific	0.957	0.270	3.551	0.000

In the figure a 95% CI for the influence of being female on the chance of a person trying a hookah over a male is (-0.0469, 0.125) or on the natural scale ($\frac{e^(-0.0469)}{1+e^(-0.0469)} = 0.488, \frac{e^(0.125)}{1+e^(0.125)} = 0.5312094$). 0 is in the confidence interval of β_3 therefore there is not enough evidence to suggest that there is a difference in likelihood of having used a hookah or waterpipe on at least one occasion between two induviduals of the different sexes at the 95% confidence level.

Since the hypothesis is testing Regular use of chewing tobacco, snuff or dip is no more common amongst Americans of European ancestry than for Hispanic-Americans and African-Americans the binomial regression with the count of 1 was chosen to fit the categorical data.

From the smoking dataset the covariates: Age, Race, Rural or Urban, and number of days used chewing tobacco out of the past 30. The dataset then filtered out those living in a rural community since there is a known bias of white Americans more likely to live in rural areas and chewing tobacco is a rural phenomenon. A column designating regular users was added.

```
##        Race           Age        RuralUrban    chewing_tobacco_snuff_or
##  white   :4416   Min.   : 9.00   Urban:10937   Mode :logical
##  black   :1927   1st Qu.:13.00   Rural:    0   FALSE:10724
##  hispanic:3653   Median :14.00                 TRUE :213
##  asian   : 751   Mean   :14.49
##  native  : 148   3rd Qu.:16.00
##  pacific :  42   Max.   :19.00
##  days_use_chew_tobacco_snu   Regular
##  Min.   :1.000             Length:10937
##  1st Qu.:1.000             Class :character
##  Median :1.000             Mode  :character
##  Mean   :1.066
##  3rd Qu.:1.000
##  Max.   :7.000
```

Fitting a Binomial generalized linear model to the data with a logit link $log(\frac{p}{1-p}) = \beta_0 + \beta_1 X_1 + I_{Hispanic}\beta 2 X_2 + I_{Black}\beta_3 X_3$

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-7.747	0.660	-11.731	0.000
Raceblack	-1.002	0.291	-3.437	0.001
Racehispanic	-0.393	0.186	-2.110	0.035
Raceasian	-1.595	0.590	-2.704	0.007
Racenative	0.438	0.598	0.732	0.464
Racepacific	1.268	0.614	2.064	0.039
Age	0.252	0.042	5.973	0.000

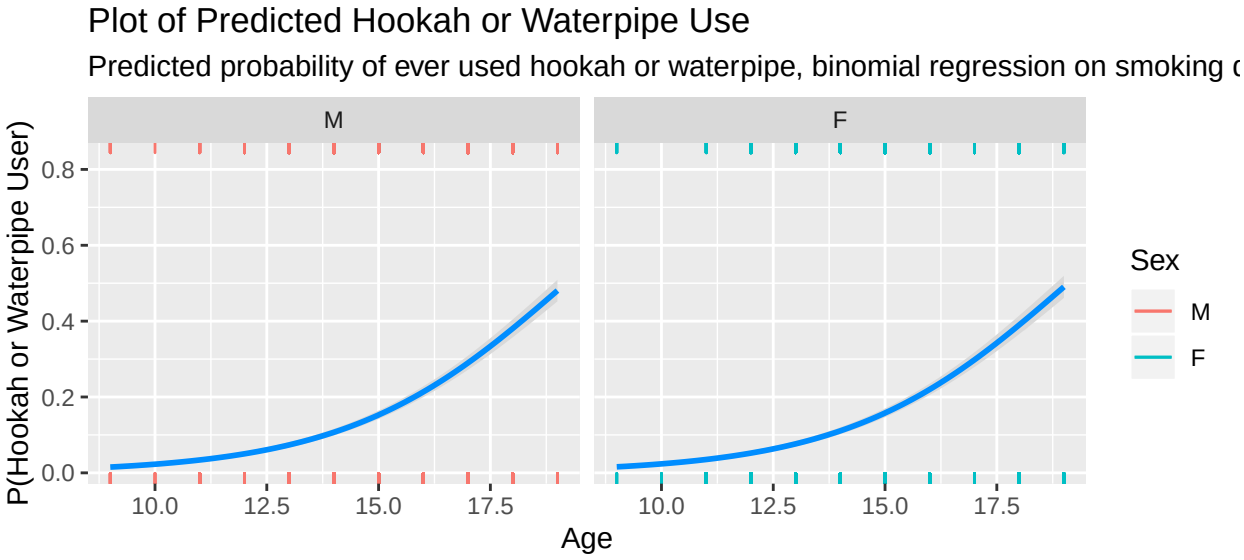
In the figure a 95% CI for the influence of being African-American person on the chance of being a regular chewing tobacco user over a European-American person is (-1.584, -0.420) or on the natural scale ($\frac{e^(-1.584)}{1+e^(-1.584)} = 0.170, \frac{e^(-0.420)}{1+e^(-0.420)} = 0.396$). While the influence of being Hispanic-American person on the chance of being a regular chewing tobacco user over a European-American person is (-0.765, -0.021)

or on the natural scale ($\frac{e^(-0.765)}{1+e^(-0.765)} = 0.317, \frac{e^(-0.021)}{1+e^(-0.021)} = 0.494$). There is evidence to suggest that European-Americans are more likely to regularily consume chewing tobbacco than both African-Americans and Hispanic-Americans.

Summary

In conclusion when accounting for age, ethnicity, and other demographic charactaristics there is not enough statistical evidence to conclude that there is a difference in likelihood of having used a hookah or waterpipe on at least one occasion is the same for two individuals of the different sexes.

The regression plot below shows a very similar trend among both sexes across all recorded ages.



While European-Americans are more likely to regularily consume chewing tobbacco than both African-Americans and Hispanic-Americans when accounting for demographic bias.

The regression plot below shows different trends for the three cultural groups. The difference becomes more prominent with age.