Analysis of Road Fatalities

Fiona McLean

Introduction

Understanding road fatalities is important as legislators and health officials look to limit fatalities and make roads safer. Men are involved in accidents more than women, and the proportion of accidents which are fatal is higher for men than for women. This may be due in part to women being more reluctant than men to walk outdoors late at night or in poor weather, and could also reflect men being on average more likely to engage in more risky behavior than women. In this report, we assess whether the UK road accident data are consistent with the hypothesis that women tend to be, on average, safer as pedestrians than men, particularly as teenagers and in early adulthood.

Methods

The data set used in the analysis is from the UK government, which contains all road traffic accidents in the UK. There are 1,159,453 observations spanning from 1979 to 2015.

As an exploratory tool, tables of slight and fatal casualties and sex were created. Fatal accidents as treated as cases and slight injuries are treated as controls, moderate injuries are ignored.

In order to determine if women tend to be, on average, safer as pedestrians than men, particularly as teenagers and in early adulthood data is stratified so that observation of accidents from the same light and weather conditions in the same year, month, day and hour were grouped together. Strata with no cases or no controls were removed. Since we strata on time and weather conditions, the fact that more men may be out at night or in bad weather conditions is controlled for, and only risky behavior is left in the analysis. We set men aged 25-36 as the baseline group. In order to test our hypothesis, a conditional logistic regression model was constructed, where we condition on being in an accident. The fixed effects in the model are age of the person in the accident and an interaction term age and sex. We include the interaction between sex and age because we believe that men and women propensity to be in fatal accidents may be different between different ages.

The model is:

$Y_{ij} \sim Bernoulli(P_{ij})$
$logit(P_{ij}) = \alpha_i + X_{ij}\beta$, where:

- Y_{ij} is the probability person i from strata j had a fatal accident, given they were in an accident
- α_i is the intercept of each strata
- $X_{ij}\beta$ are the covariates age and age*sex

Results

To gain some intuition as to how different the number of fatal accidents are between sexes, given the person was in an accident, a table was created to show the number of fatal and slight accidents for each sex.

Table 1: The table below shows that the proportion of fatal accidents are higher for men than for women in the data set.

	Male	Female
Slight	637919	481811
Fatal	24429	15212

To better understand how the number of fatal accidents change for each sex and age group, a chart showing the number of fatal accidents for each sex in each age group was created. The graph shows that more accidents are fatal for men in every age category except for over 75.

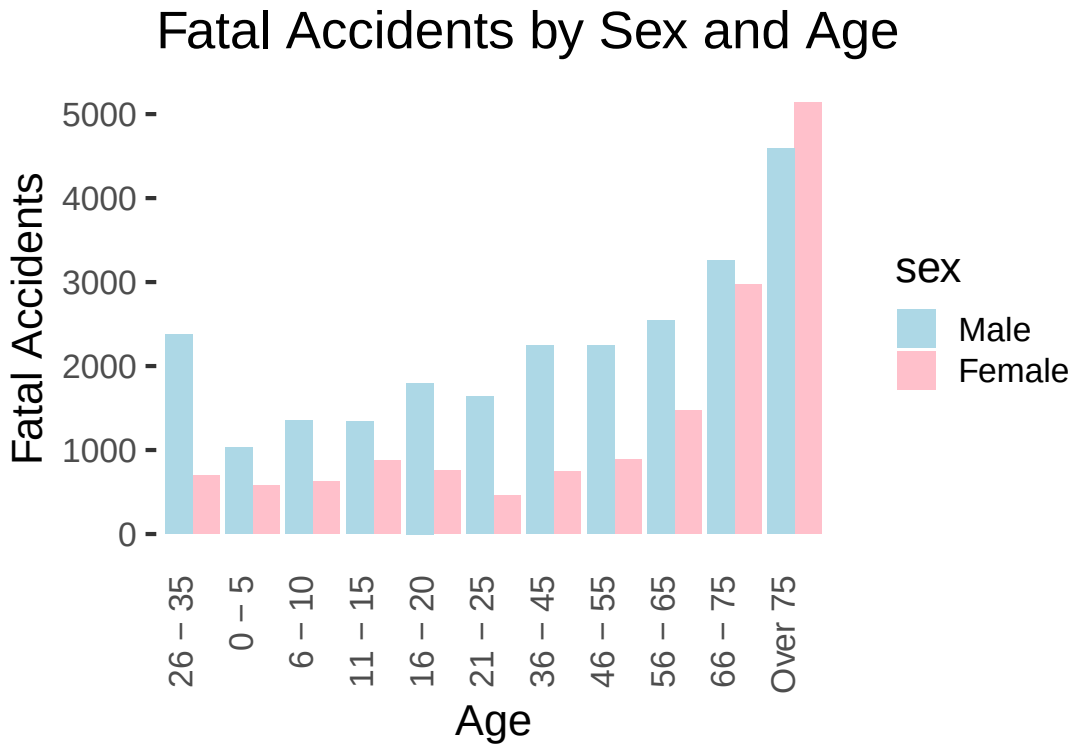

Figure 1: The figure above shows that in the vast majority of age groups, men have more fatal accidents. As it is a scientific fact that women live longer than men, perhaps women have more fatal accidents in the over 75 group simply because there are more women alive and able to drive than men at this age.

We now run the model described above to determine whether the UK road accident data are consistent with the hypothesis that women tend to be, on average, safer as pedestrians than men, particularly as teenagers and in early adulthood. We plot the results of the model for better interpretability.

Figure 2: The left plot shows the odds that an accident is fatal, given the man is in an accident compared to the baseline of 25-36 year old men. The right plot show the odds of a fatal accident for a woman, given she is in an accident, compared to a man of the same age.

We first start by discussing the results of the model displayed for men in the left plot. For each age group, we compare the odds of being in a fatal accident given the person is in an accident compared to the baseline age of 25-36. For example, for men under the age group of 25-36, it appears the odds of being in a fatal accident given the person is in an accident is lower than for men of the age 25-36. However, after the age of 25-36, the odds of a man being in a fatal accident given he is in an accident increases compared to the baseline age of 25-36. At the age of 70, the odds a man is in a fatal accident given he is in an accident is 10 times the odds of a 25-36 year old man. Therefore, we can conclude that the odds of being in a fatal accident for men given they are in an accident increase steadily after the age of 30, compared to men aged 25-36. Now we analyze the graph on the right, which displays the odds that a woman is in a fatal accident, given she is in an accident compared to a man of the same age. For example, at the age of 10, the odds of woman is in a fatal accident, given she is in an accident, is about .8 the odds of a 10 year old man compared to the baseline of men ages 25-36. We can see that for almost all ages, the odds of a woman being in a fatal accident given she is in an accident compared to a man of the same age is under 1, meaning that the women is less likely to be in a fatal accident.

To answer our main research question as to if women tend to be, on average, safer as pedestrians than men, particularly as teenagers and in early adulthood, we compare the odds of a woman being in a fatal accident given that she is in an accident to a man of the same age between 10 and 25. We can see that the odds a woman is in a fatal accident compared to men is much lower, about .8 the odds of man of the same age at 10, decreasing all the way to about .65 the odds of a man the same age at 25. As noted before, at almost every age the odds of a woman being in a fatal accident given she is in an accident compared to a man is lower than 1. Therefore, the model suggests that women are safer pedestrians than men at almost every age, but especially teenage to young adult ages.

Conclusion

To determine how fatality rates of accidents differ between men and women at different ages, a conditional logistic regression model was constructed. In order to remove the possibility that more men drive in the dark or bad weather conditions, the model considers time down to the hour, as well as weather and lighting conditions. The model suggests that at almost every age group, the odds a woman is in a fatal accident given she is in an accident compared to a man the same age is lower. At the teenage and young adult ages, the odds of a fatal accident compared to a man the same age are particularly lower, between .8 to .65 the odds of man, suggesting that women are especially safer pedestrians in the young adult ages.