Separations under Sunshine: An Analysis of Divorce in California

Emily Hou, Ashlyn Jew, Kirsten Landsiedel, Hana Lim, Aidan O'Sullivan, and Natasha Vuong
Statistics 140SL

UCLA Department of Statistics

Introduction & Objective

The **divorce rate in the United States has been declining overall** for the past 40 years, reaching 14.9% in 2019, which was the lowest seen in 50 years (Wang 2020). Depending on who is asking, this is probably great news, but divorce is still significant — it may have an effect on socioeconomic and gender inequality, and parents' marital changes are associated with reduced levels of their children's well being (Raley & Sweeney 2020). Because divorce varies in occurrence and consequences across different demographics, we want to **determine which factors affect whether a person is married or divorced**. We also had a hunch that there was a difference in women's divorce rates based on whether they had children or not, so we test that as well.

Methods

Using IPUMS USA, we retrieved data from the U.S. Census Bureau's American Community Survey (ACS). Since the ACS samples 3.5 million addresses of mostly categorical information, we made our analysis more manageable by restricting it to California, selecting variables we found interesting, and combining factors where possible.

We used a z-test on the following hypotheses:

- H_0: **There is no difference in divorce rates between women with children and women without children.**
- H_a: **There is a difference in divorce rates between women with children and women without children.**

We subsetted our data to only include those who were married or divorced and built a binary logistic regression model with a variety of predictors. We also ran a random forest model to further examine model accuracy and significant variables.

Results

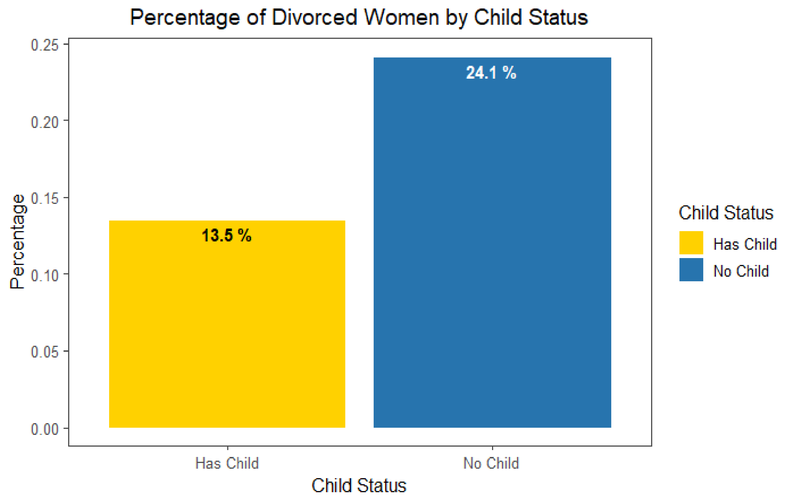
Our z-test had a p-value of less than 2.2×10^{-16}, meaning we **reject the null hypothesis** that there is no difference in divorce rates between women with children and women without children.

Our logistic model had 84% accuracy, and the **most significant variables included household income, sex = female, age, race = African American, and race = Asian/Pacific Islander**.

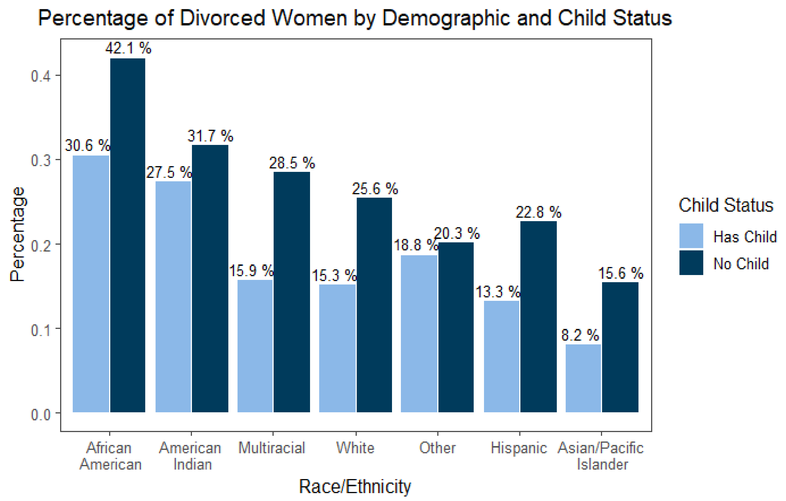
To further explore the importance of variables, we ran two random forest models: one with all variables, which had 88% accuracy, and a reduced model that excluded the more obvious variables, which had slightly less accuracy of 82.9%. The variable importance plot of the reduced model showed that race had the biggest effect on accuracy when removed.

Visualizations

Based on all of these results, we decided to ignore household income, sex, and age in favor of exploring race further.



Our graph shows that there are 10.6% more divorced women without children than women with children.



This graph shows the percentage of divorced women for each race/ethnicity, also divided by whether they have children or not. Notably, African American women have the highest percentages of divorce, while Asian/Pacific Islander women have the lowest.

Discussion

The results of our z-test and our first visualization suggest that **more women without children tend to be divorced, when compared to women with children**. We theorize that this could be due to women wanting to stay married for the sake of their children.

Our second visualization supported the findings of our logistic regression model. Race = African American had a relatively large positive coefficient, which was reflected in the **high percentages of divorced African American women**. Similarly, race = Asian/Pacific Islander was the only significant race variable with a negative coefficient, which was reflected in the **low percentages of divorced Asian American women**. It is possible that Asian American women tend to stay married due to the emphasis on the collective and the family within many Asian cultures.

It is generally difficult to make definitive conclusions about race and its consequences because race is socially constructed (Omi and Winant 2014). The ACS relies on people to identify as one of the racial categories predetermined by the government, even though these categories typically have different definitions across individuals. Hispanic is also a separate category for sociohistorical reasons, but we treated it as another factor within the race/ethnicity variable for visualization purposes. There is overlap between those who identify as Hispanic and those who identify as a different race/ethnicity.

References

1. Wang, W. (2020). The U.S. Divorce Rate Has Hit a 50-Year Low. Institute for Family Studies. Retrieved March 9, 2021, https://ifstudies.org/blog/the-us-divorce-rate-has-hit-a-50-year-low
2. Raley, R. K., Sweeney, M. M. (2020). Divorce, Repartnering, and Stepfamilies: A Decade in Review. Journal of Marriage and Family, 82(1), 8199. https://doi.org/10.1111/jomf.12651
3. Minnesota Population Center, University of Minnesota. U.S. CENSUS DATA FOR SOCIAL, ECONOMIC, AND HEALTH RESEARCH. https://usa.ipums.org/usa/
4. Raley, R. K., Sweeney, M. M., Wondra, D. (2015). The Growing Racial and Ethnic Divide in U.S. Marriage Patterns. The Future of Children / Center for the Future of Children, the David and Lucile Packard Foundation, 25(2), 89109.
5. Amato, P. R. (2010). Research on Divorce: Continuing Trends and New Developments. Journal of Marriage and Family, 72(3), 650666. https://doi.org/10.1111/j.1741-3737.2010.00723.x
6. Omi, M., Winant, H. (2014). Racial formation in the United States. Routledge.