BST 222 : Analysis of Factors Affecting Observed Breastfeeding Durations

Elsie Basa, Shirley Lin, Laura Wang

1 Introduction

1.1 Background

For newborns, current recommendations by the CDC are to primarily breastfeed a child for the first 6 months of their life and then move to gradually introduce solid foods into their diet. Previous studies have reported the mean and median breastfeeding times to be about 5.7 and 6 months respectively (Wasie Kasahun et al.). However, breastfeeding can continue for as long as the mother and child feel like they can continue for. In general, there are few drawbacks to breastfeeding for longer times, with some recommendations suggesting continuing breastfeeding for as long as 2 years ("Breastfeeding FAQs: How Much and How Often (for Parents) - Nemours KidsHealth"). Breastfeeding is considered beneficial for children, protecting them from certain diseases and past studies have shown that breastfed babies have lower risks of asthma, obesity, and sudden infant death syndrome (SIDS) ("Breastfeeding Benefits Both Baby and Mom | DNPAO | CDC"). While there are certainly multiple benefits to breastfeeding newborns, there's a lot of variability in the time until a mother chooses to wean her child. In some cases, mothers will only breastfeed their children for approximately a week before weaning them. The primary purpose of this analysis is to identify trends in the duration of the breastfeeding period as well as identify variables that appear to be associated with these trends.

1.2 Data Description

The data was sourced from the `KMsurv` package in R. The original data is from the National Longitudinal Survey of Youth (a survey which began in 1979 and ended in 1988). The package data is a cleaned and subsetted version of the original data.

Beginning in 1983, the women in the study population were asked about any pregnancies they had and various data about them were collected. For this dataset, only first-born children born after 1978 at a gestational age of 20 to 45 weeks were included. Then, the data was further narrowed down to only include responses from mothers who chose to breastfeed their children. After taking all of these factors into account, we end up with 927 total observations. For this analysis, the response (survival) is generated using the breastfeeding duration (weeks) and an indicator for whether or not the child was weaned.

The variables included in this dataset are shown in Table 1.

Table 1: Factor Levels for Categorical Variables

race	Race of mother (1=white, 2=black, 3=other)
poverty	Mother in poverty (1=yes, 0=no)
smoke	Mother smoked at birth of child (1=yes, 0=no)
alcohol	Mother used alcohol at birth of child (1=yes, 0=no)
agemth	Age of mother at birth of child
ybirth	Year of birth
yschool	Education level of mother (years of school)
pc3mth	Prenatal care after 3rd month (1=yes, 0=no)

It is important to note that, for this dataset in particular, formula feeding is not considered breastfeeding despite formula being a form of liquid nourishment for infants. The indicator for poverty only indicates whether or not the mother was in poverty at the time of birth of her child. Table 2 shows some very basic summary statistics for the outcome variables (`duration` and `delta`). When `delta` takes on a value of 1, it means that the child completed breastfeeding (was weaned) at the end of the reported period of time. Otherwise, the value of 0 meant that the information was either censored (unobserved) or there was loss to follow-up. All of the children who were reported to not have finished breastfeeding yet were born during or after 1984.

Table 2: Summary of Data Outcome

bfeed Outcome Summary	**N = 927**
duration (weeks)	
Mean	16
Median	10
Minimum-Maximum	1-192
delta (=1)	892 (96%)

[1] n (%)

Table 3: Summary of Data by Variable

bfeed Variable Summary	**0**, N = 35	**1**, N = 892
race		
white	28 (80%)	634 (71%)
black	4 (11%)	113 (13%)
other	3 (8.6%)	145 (16%)
yschool		
noHS	1 (2.9%)	219 (25%)
HSgrad	13 (37%)	425 (48%)
someCollege	21 (60%)	248 (28%)
poverty	3 (8.6%)	168 (19%)
alcohol	3 (8.6%)	76 (8.5%)
smoke	7 (20%)	263 (29%)
agemth		
Mean (SD)	25 (2)	21 (3)
pc3mth	8 (23%)	156 (17%)

[1] n (%)

General guidelines for pregnant mothers recommend seeking prenatal care after the third month of pregnancy. However, as seen in Table 3, only approximately 18% of the mothers chose to seek out prenatal care after the third month of their pregnancies. Also, while a fairly low percentage of the mothers consumed alcohol around the time of the birth of their child, approximately 30% of them reported smoking at the time. However, the degree of their smoking, such as how much or how often they smoked, wasn't included in the dataset. Visual representations of these summary tables can be found in the exploratory data analysis portion of the paper.

2 Exploratory Data Analysis

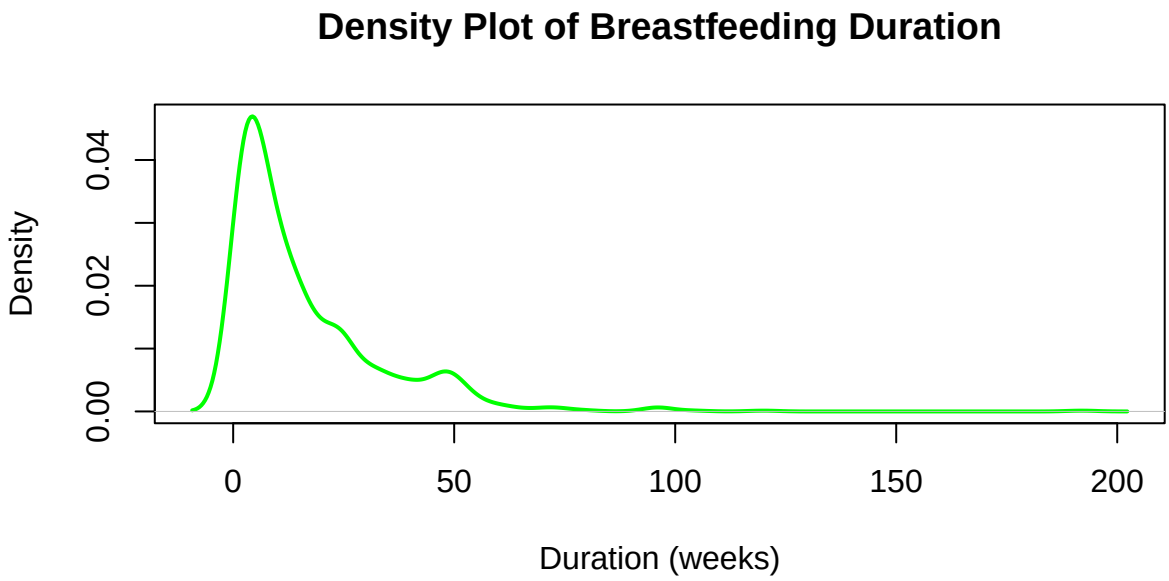

Figure 1: Density Plot of Breastfeeding Duration

Figure 1 is a density plot shows that almost all of the mothers in the study only breastfeed for no more than about 50 weeks, which is about a year. On the graph there is a spike at around 10-12 weeks indicating that the majority of mothers only breastfeed for about a few months.

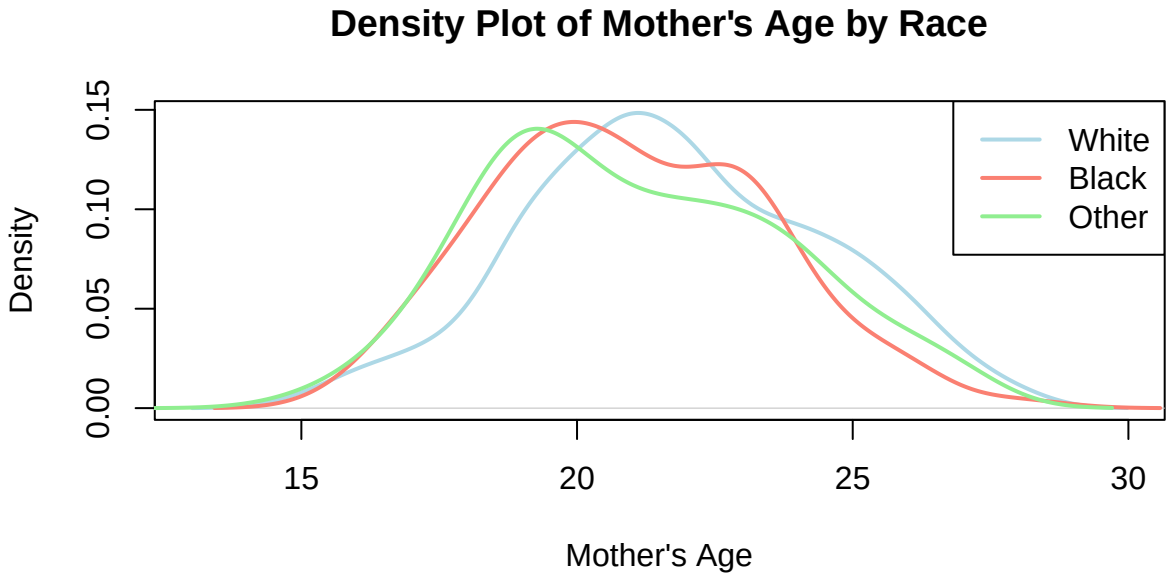

Figure 2: Density Plots Mother's Age by Race

Figure 2 shows the density of the mothers' age by race demonstrates that black mothers and mothers of other races tend to have their first kids younger than white mothers.

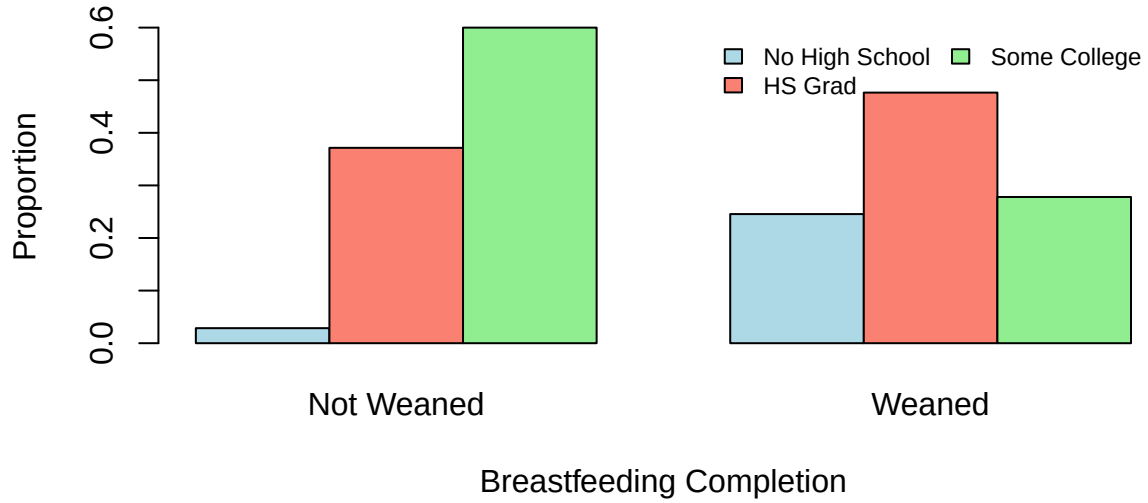

Figure 3: Bar Plot on Education proportion By Breastfeeding Status

Figure 3 compares the group of mothers who weaned their babies versus those who did not wean their babies. Those who didnn't wean their children had higher proportion of mothers who went some college in their education background. Mothers who weaned their babies had a higher proportion of those who did not finish high school.

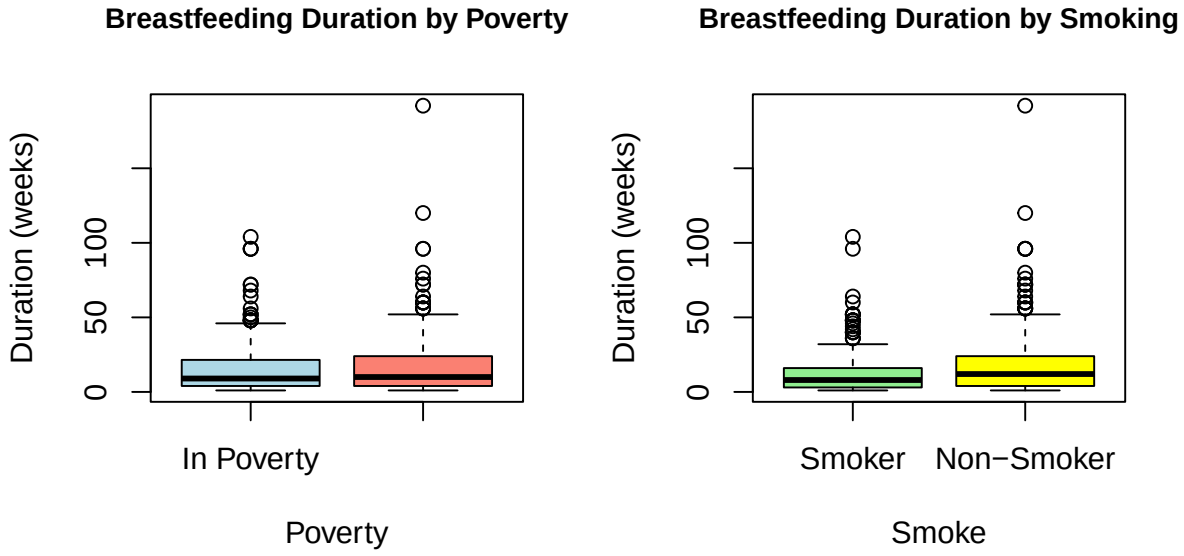

Figure 4: Box Plot for Categorical Variables

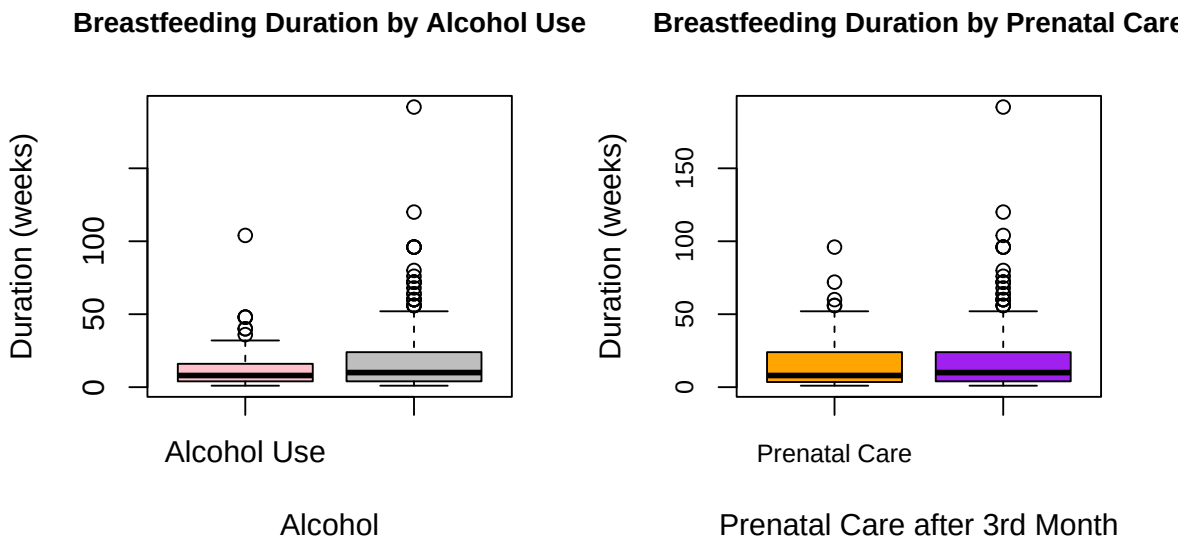

Figure 5: Box Plot for Categorical Variables

The first box plot shows that the breastfeeding duration for those in poverty compared to those who are not in poverty are similar but those in poverty have a slightly lower and narrower breastfeeding duration range. The second box plot shows that mothers who smoked had a sligtly shorter breastfeeding duration and that the range for breastfeeding duration is also narrower compared to mothers who did not smoke.

The third box plot, demonstrates that mothers who use alcohol have a narrower breastfeeding distribution with most mothers breastfeeding betwween 0 and under 50 weeks. Mothers who did not consume alcohol tend to breastfeed between 0 and 52 weeks. Those mothers who consumed alcohol tend breastfeed for a shorter amount of time compared to those who did not. The fourth box plot showed that the breastfeeding duration distribution of mothers who had prenatal care compared to those are very similar, with majority of both groups mainly brestfeeding for 0 weeks to 50 weeks.

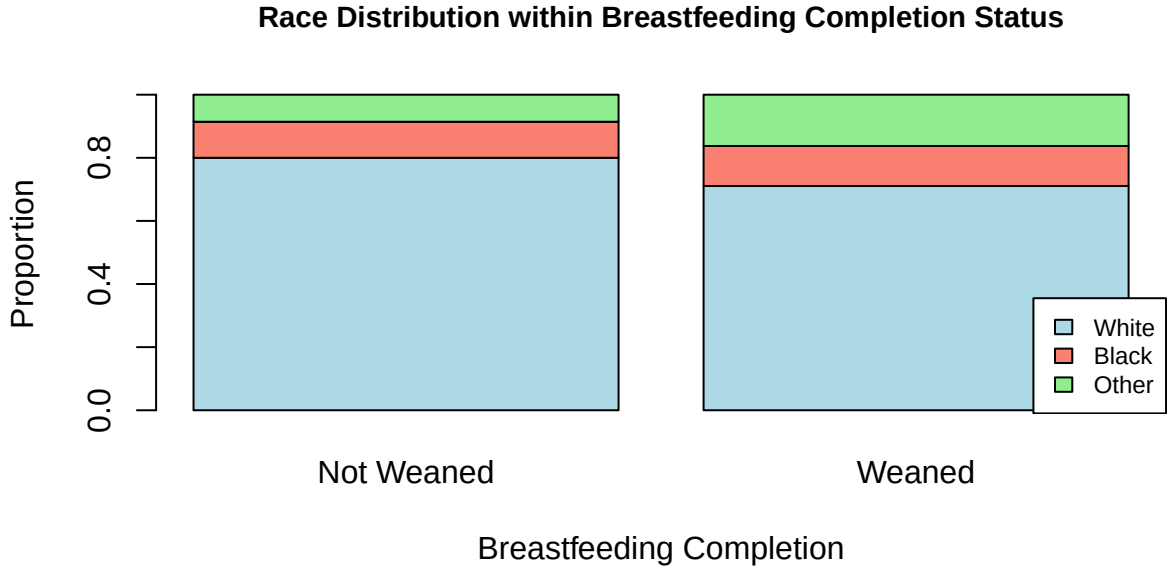

Figure 6: Box Plot of Race Distribution by Breastfeeding Status

Figure 6 compares the proportion of each race by breastfeeding status. The group of mothers who were able to wean have a high proportion of mothers of other race and black mothers compared to the group that did not wean.

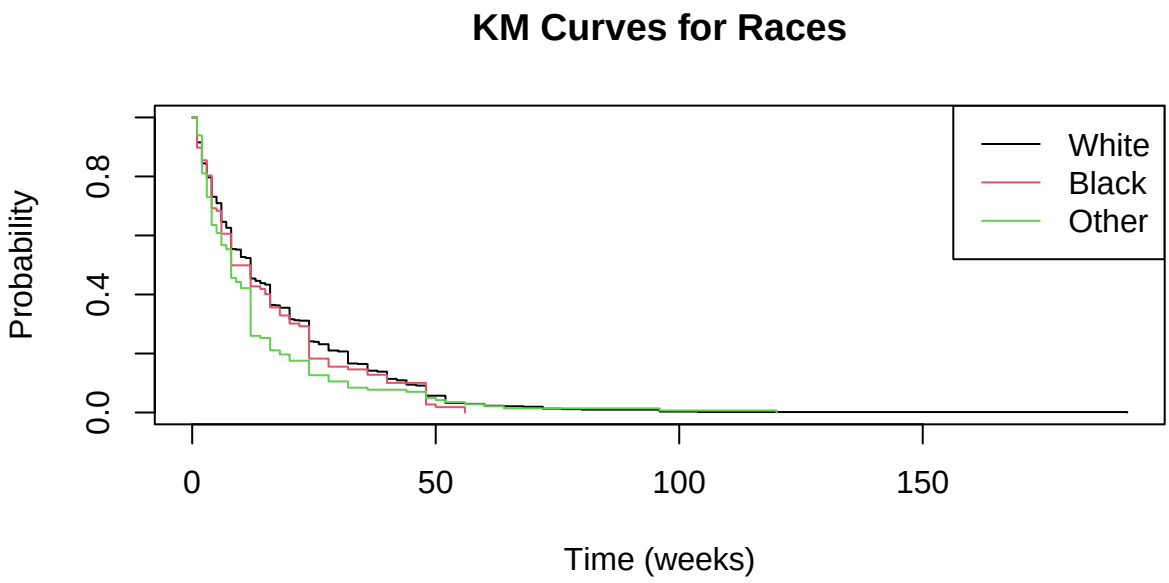

Figure 7: KM Curve - Race

Figure 7 compares the probability a mother doesn't wean for each of the different race categories. For each race, there is a steep decline around the 10-20 week mark indicating many mothers stop breastfeeding around that time. Additionally, mothers of other race tend to wean sooner then black mothers and white mothers. Black mothers also weaned faster than white mothers.

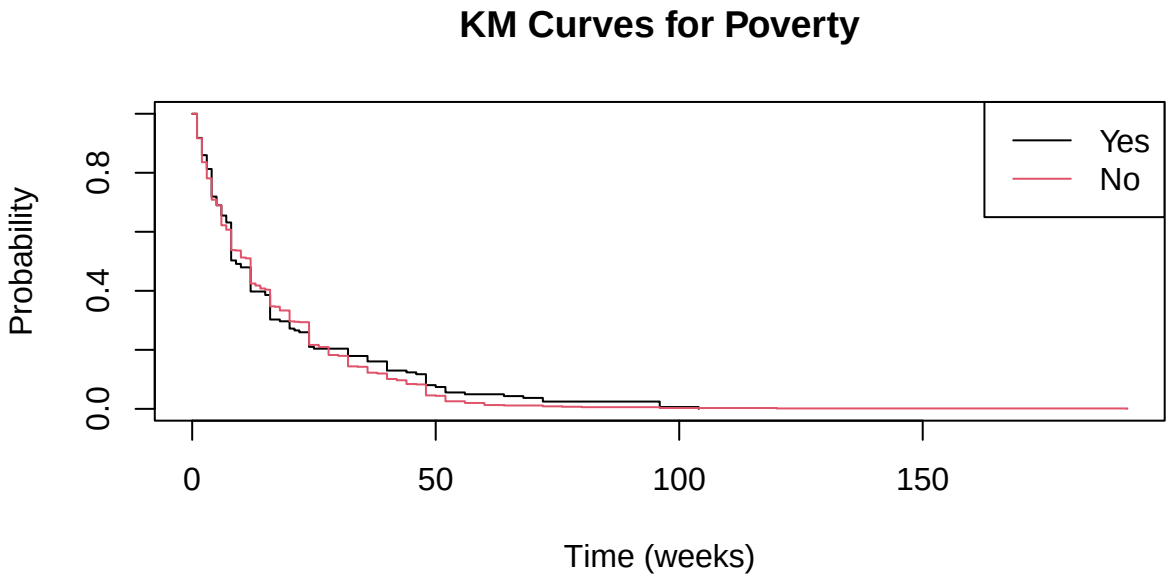

Figure 8: KM Curve - Poverty

Figure 8 compares the probability a mother doesn't wean depending if they are in poverty or not. For both groups, there is a steep decline around the 10-20 week mark indicating many mothers stop breastfeeding around that time. The Kaplen Meier curves are similar and do overlap til around week 30 where it can be seen that mother's who are not in poverty have a higher probability of weaning their child.

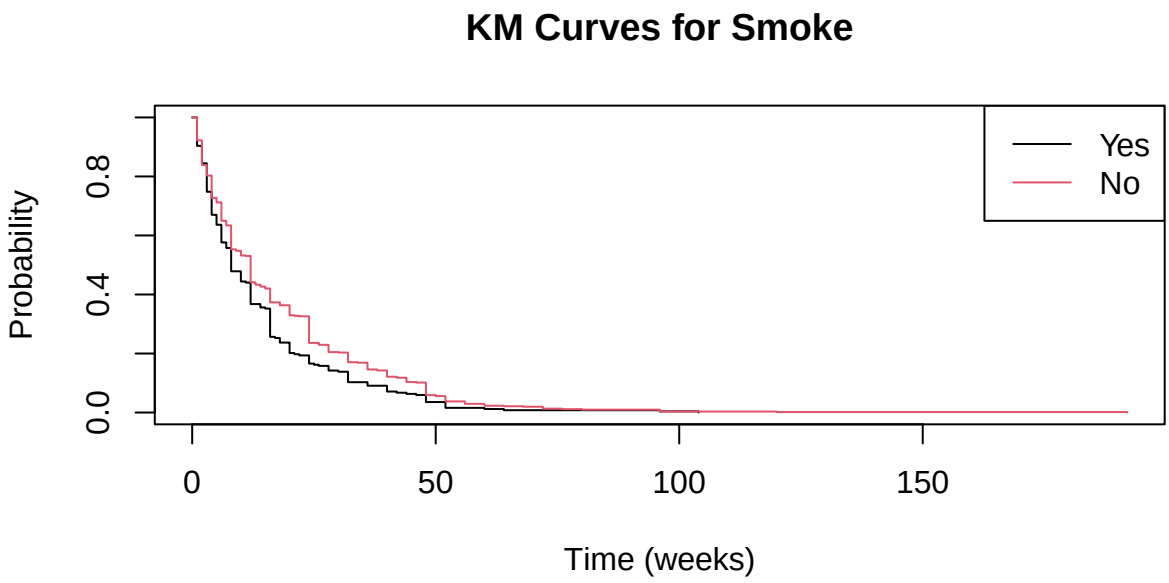

Figure 9: KM Curve - Smoke

Figure 9, compares the probability a mother does not wean for mothers who smoke and mothers who do not smoke. The curves have a steep decline around the 10-20 week mark indicating that many mothers stop breastfeeding around that time period. The Kaplen Meier curves indicate that there is a higher probability that mothers who smoke will wean compare to those who do not smoke.

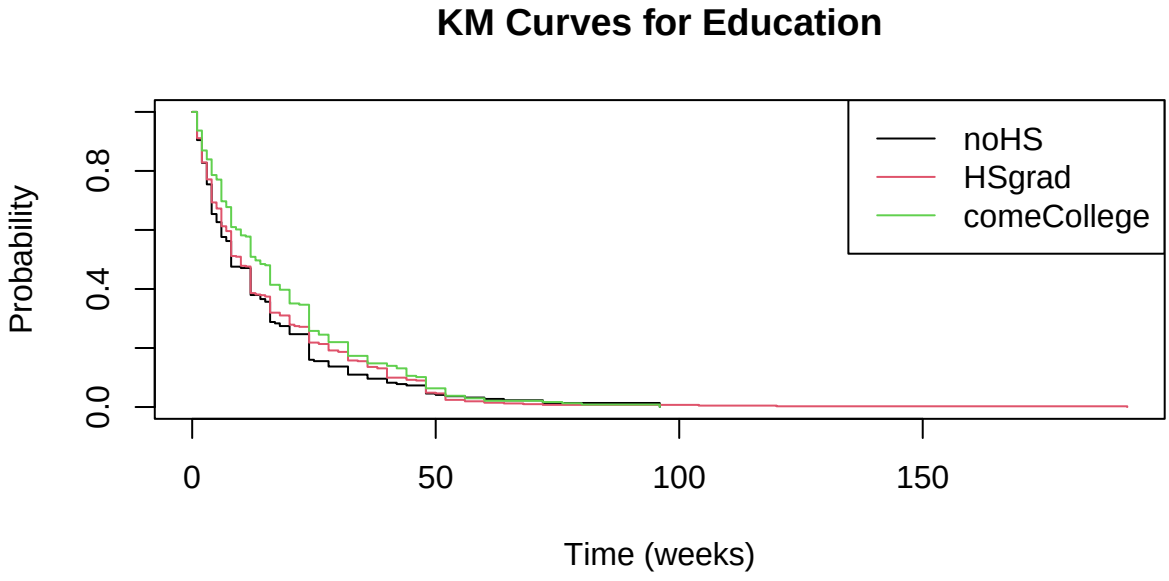

Figure 10: KM Curve - Education

Figure 10, compares the probability a mother does not wean for mothers in different education levels. The curves have a steep decline around the 10-20 week mark indicating that there is a low probability that mothers continue breastfeeding around that time period. Mothers with some college education are more likely to not wean compare to mothers with high school education and mothers with no high school education. For the KM curves for mothers with no high school education and those with only a high school degree have similar likelihoods of not weaning their child till around week 20. Around that time those without a high school education are more likely to wean than those without one.

3 Model Building

4 Conclusion

4.1 Discussion

5 Bibliography

"Breastfeeding Benefits Both Baby and Mom | DNPAO | CDC." Centers for Disease Control and Prevention, 7 Sept. 2023, https://www.cdc.gov/nccdphp/dnpao/features/breastfeeding-benefits/index.html.

"Breastfeeding FAQs: How Much and How Often (for Parents) - Nemours KidsHealth." Nemours KidsHealth - the Web's Most Visited Site about Children's Health, https://kidshealth.org/en/parents/breastfeed-often.html. Accessed 7 Dec. 2023.

"Frequently Asked Questions (FAQs) | Breastfeeding | CDC." Centers for Disease Control and Prevention, 18 Apr. 2023, https://www.cdc.gov/breastfeeding/faq/index.htm.

"Kaplan Meier Curve • Simply Explained - DATAtab." Online Statistics Calculator: Hypothesis Testing, t-Test, Chi-Square, Regression, Correlation, Analysis of Variance, Cluster Analysis, https://datatab.net/tutorial/kaplan-meier-curve. Accessed 4 Dec. 2023.

Wasie Kasahun, Abebaw, et al. "Predictors of Exclusive Breastfeeding Duration among 6–12 Month Aged Children in Gurage Zone, South Ethiopia: A Survival Analysis | International Breastfeeding Journal | Full Text." BioMed Central, 21 Apr. 2017, https://internationalbreastfeedingjournal.biomedcentral.com/articles/10.1186/s13006-017-0107-z.

6 Appendix