MATH 189 Case Study 1

Maternal Smoking and Infant Weight

Yuxin Zou, A13996888
Yifei Li, A92082060
Bo Hu, A13805750
Bo Zhang, A13616488
Xuanyu Wu, A13569778

Introduction

It is well known that smoking is harmful to one's health, but the adverse effects of maternal smoking on infant health are not as notable. These effects may increase the risk of perinatal morbidity, mortality, newborn low birth weight, preterm delivery, and so on. While we cannot fully comprehend each individual substance's effects on infant health, studies have shown that nicotine and carbon monoxide are among the major causes. Nicotine residues from the smoke circulate around the body as well as the fetus. Apart from causing the newborn to weigh less, it can also has behavioural and neural consequences due to nicotine exposure. Carbon monoxide also poses great threat to the growth of fetus. Since carbon monoxide dissolves far better than oxygen in water, blood in maternal body does not supply enough oxygen to the baby, causing growth to be significantly slower, thus a reduced infant weight.

Hypothesis

Different time stage of maternal smoking relates to babies' health, which is measured by a number of metrics, such as baby weight, length, etc. In particular, infants born to mothers who smoke during pregnancy have lower weights than infants whose mothers smoke in other stages.

Data Description & Data Cleaning

The data for this report is from babies23.txt, which is part of the Child Health and Development Studies (CHDS). The data is collected of all pregnancies that occurred between 1960 and 1967 among women in Kaiser Health Plan of the San Francisco region. In this dataset, these women were all enrolled in Kaiser Health Plan and had obtained prenatal care in San Francisco area and were delivered in any of the Kaiser hospitals in Northern California, which means that they received medical care at a very early stage of their pregnancies. This dataset is comprised of 1236 babies with all the same gender (boys) who have lived at least 28 days and were all single births (no twins). The birth weights of these babies were all measured in ounces, which is a discrete numerical value. Also, maternal smoking status, which is a categorical variable, is represented by an integer indicator to differentiate different smoking time stages, of which '0' indicates that the Mom has never smoked before, 1 indicates that she is still smoking even now, '2' indicates that she smoked until pregnancy and '3' indicates that she once smoked but quit now. Furthermore, the age of Mom is also provided in years as an ordinal categorical variable.

The data babies23.txt also consists of missing values, indicated as unknown. In order to facilitate the future analysis on this dataset, observations of unknown value in the fields were dropped. E.g. observations of unknown maternal smoking state (indicated by '9'), observations of unknown Mom's race (indicated by '99') and observations of unknown education of Mom (indicated by '9').

Also, the age of Mom is considered as a confounding factor with the possibility of masking the effect of smoking on the infant death rate. "Confounding variables and confounders are often defined as the variable correlate (positively or negatively) with both the dependent variable and the independent variable" (Pourhoseingholi, Baghestani, Vahedi), where in this case with dependent variable being infant weights and independent variable being maternal smoking smoking status. In order to eliminate the confounding effects by age of Mom, we use the method stratification to fix the level of the confounder (the range of age) and produce groups within which the confounder does not vary. From the SSMHealth, "if you're 35 years old or older pregnant, you're considered to be of Advanced Maternal Age (AMA)" and "Teenage Pregnancy" is considered for pregnant women below 17 years old. The justification for this classification is associated with the risk of pregnancy. "Women tend to have a window of time when it's easier on their body to grow a baby and give birth" (ssmhealth.com). Thus, based on this classification, we only select Mom's of ages ranging from 17 to 35 in order to erase the effect the confounding factor. After stratification, the size of data for each smoking status is: 416 observations for smoking status 0, 376 observations for smoking status 1, 75 observations for smoking status 2, 67 observations for smoking status 3.

Background

Previous studies have shown that "infants born to mothers who smoke are a few hundred grams smaller, on average, than the infants of nonsmokers." Considered as evidence of toxicity of smoking cigarettes, it coincides with Michael's finding that "maternal smoking is associated with post-neonatal mortality. (Michael, et al.)" Besides the effect of smoking on infant death, Friede's study demonstrated the "association between young maternal age and high infant mortality and low birth weight. (Allen)" Furthermore, the study conducted by Alam illustrated the relationship between parity and infant mortality: "first-born children were at higher odds of dying in infancy than second births if mothers were in their twenties."

Schoendorf and Kiely's study shifts the focus to "the relationship between tobacco exposure during infancy and Sudden Infant Death Syndrome(SIDS)" (Kennith, John). This case-control analysis assigns infants into three exposure groups: "maternal smoking during both pregnancy and infancy (combined exposure)", "maternal smoking only during infancy (passive exposure)", and "no maternal smoking." This study also illustrates that race may function as a confounding factor by comparing the odds ratio between infants of different race and tobacco exposure. After adjustment for demographic risk factors, the study concludes that "both intrauterine and passive tobacco exposure are associated with an increased risk of SIDS" (Kennith, John).

Investigation

Numerical Analysis

Baby Weight	count	mean	std	min	25%	50%	75%	max	skewness	kurtosis
Smoke Type										
Never	493.0	122.452333	16.945880	55.0	113.00	123.0	132.00	174.0	-0.151657	4.426977
During Pregnancy	426.0	113.650235	17.922374	58.0	101.00	114.5	125.00	161.0	-0.012740	3.050064
Until Pregnancy	84.0	122.345238	18.200023	62.0	110.75	121.5	136.00	163.0	-0.388596	3.457181
Used To	85.0	123.941176	18.650516	65.0	112.00	124.0	138.00	160.0	-0.450883	3.359515
All	1088.0	119.113971	18.086470	55.0	108.00	119.0	130.25	174.0	-0.154868	3.497915

Table 1: Descriptive statistics of baby weight grouped by mother smoke type

There are 493 babies born from mothers who have never smoked in their life, while 426 mothers have been continuously smoking through their entire pregnancy. There are also two other occasions where 84 mothers have been smoking until they got pregnant and where 85 mothers used to smoke at some point in the past.

Just by looking at the average, which gives the center of the distribution of babies' weights, we can see that mothers who did not smoke during pregnancy have similar mean birth weights for their babies, around 123 ounces. 112.452 for mothers who never smoke, 122.345 for mothers who smoke until pregnancy, 123.941 for mothers who once smoked. For mothers who continue to smoke, their babies have a lower mean weight of 113.650 ounces. We plan to investigate into this correlation further in the following sections.

By looking at the standard deviation, which gives the degree of variability around the center of the distribution, we see that mothers who never smoke have the lowest standard deviation for their babies' weights, 16.946. For mothers who continue to smoke during their pregnancy, the standard deviation is slightly higher, of 17.922. For mothers who stop smoking for their babies and ones who once smoked, the standard deviations are respectively 18.200 and 18.650. We can see that smoking in any point in a mother's like can suggest more variability in their babies weights, but the variabilities are not so significantly different in these different cases.

By looking at the lower quartile, which indicates at least 25% falls below this number, mothers who continue to smoke during pregnancy has the lowest lower quartile of 101.00. In other cases, the lower quartiles are around 112, with mothers who stopped smoking right before pregnancy slight lower at 110.75. This seems to suggest that you can expect a higher chance for universally low birth weight if mothers are smoking during pregnancy.

For skewness, which is a measure of how unsymmetric is the dataset, we can see that mothers who smoke during pregnancy have a nearly symmetrical distribution for their babies' weights. Babies whose mothers once smoked have a weight distribution that is moderately skewed to the left, at -0.389 and -0.451, while those whose mothers who never smoke have a slight skew toward left. For kurtosis, which is a measure of how tail-heavy is the distribution from that of normal distribution, we can see that mothers who smoke during pregnancy have lesser kurtosis of 3.050, meaning the babies' weights are more located around the center when compared to other groups. Moms who never smoke have a relatively higher kurtosis, suggesting that the distribution is heavy-tailed, or that the presence of outliers is higher than other groups.

Next, we group the data by a potential confounding factor, mother's age at her baby's birth, and compare these descriptive statistics from the following two tables.

Baby Weight	count	mean	std	min	25%	50%	75%	max	skewness	kurtosis
Pregnancy Type										
Teenage	7.0	125.285714	8.360907	114.0	121.5	124.0	127.50	141.0	0.920741	3.106651
Advanced	147.0	120.197279	21.205336	55.0	108.0	120.0	134.00	174.0	-0.136939	3.263140
Common	934.0	118.897216	17.600611	58.0	108.0	119.0	130.00	174.0	-0.166154	3.467351
All	1088.0	119.113971	18.086470	55.0	108.0	119.0	130.25	174.0	-0.154868	3.497915

Table 2: Descriptive statistics of baby weight grouped by mother age

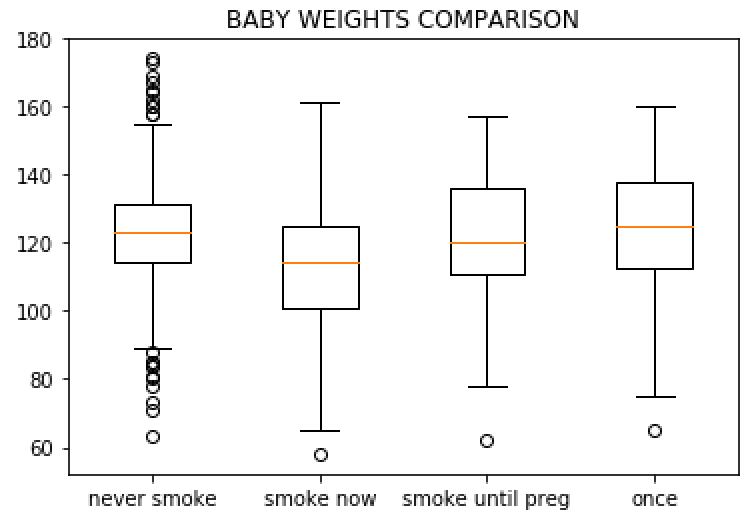
Pregnancy Type	Smoke Type	count	mean	std	min	25%	50%	75%	max	skewness	kurtosis
Teenage	During Pregnancy	3.0	120.333333	6.506407	114.0	117.00	120.0	123.50	127.0	0.229937	1.500000
	Never	3.0	125.000000	2.645751	123.0	123.50	124.0	126.00	128.0	1.457863	1.500000
	Until Pregnancy	1.0	141.000000	NaN	141.0	141.00	141.0	141.00	141.0	NaN	0.000000
Common	During Pregnancy	387.0	113.754522	17.776365	58.0	101.00	115.0	125.00	161.0	-0.021245	3.068256
	Never	430.0	122.479070	16.460400	55.0	114.00	123.0	132.00	174.0	-0.149223	4.479057
	Until Pregnancy	77.0	121.740260	18.064506	62.0	111.00	121.0	136.00	157.0	-0.488126	3.449403
	Used To	69.0	123.739130	18.289866	65.0	112.00	125.0	137.00	160.0	-0.579566	3.709809
Advanced	During Pregnancy	36.0	111.972222	20.152967	75.0	100.00	110.0	122.75	154.0	0.169034	2.764725
	Never	60.0	122.133333	20.588352	71.0	110.75	123.5	135.25	174.0	-0.126586	3.688163
	Until Pregnancy	6.0	127.000000	20.909328	106.0	111.25	126.0	132.50	163.0	1.004205	2.520321
	Used To	16.0	124.812500	20.746787	84.0	112.00	122.5	142.25	158.0	-0.075825	2.198977

Table 3: Descriptive statistics of baby weight grouped by mother age and smoke type

Since there are only 7 cases of teenage pregnancies, we only compare the case of common pregnancy and advanced pregnancy. We find that there is a universal increase (17.601 to 21.205) in standard deviation for advanced pregnancy when compared to common ones. This accounts for the nature of more variability in factors for older women pregnancies. For the average, smoking during pregnancy has a lower mean of 113.754, which is consistent with our finding earlier before we stripped out this potential confounding factor. Overall, the results are similar. As for the shape of the distribution, the skewness of baby weights for women who never smoke in both conditions are nearly the same. We will investigate the normality of the various distributions in later chapters.

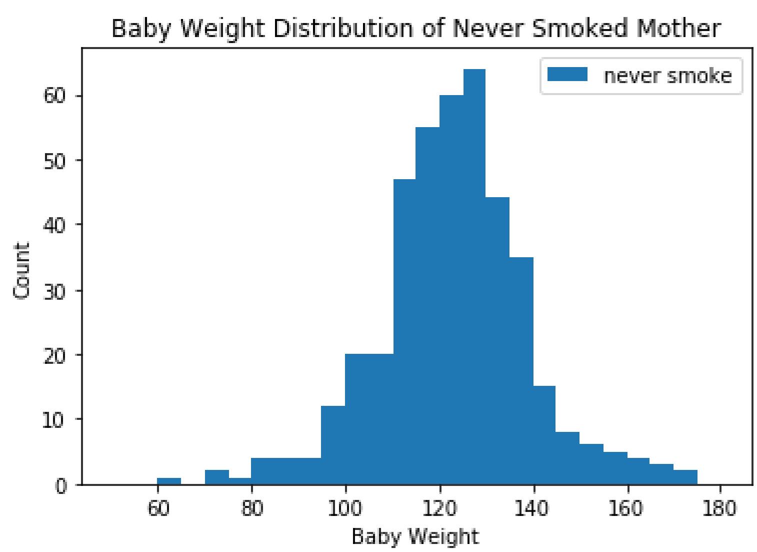
Graphical Analysis

Figure 1: Boxplot for baby weights comparison



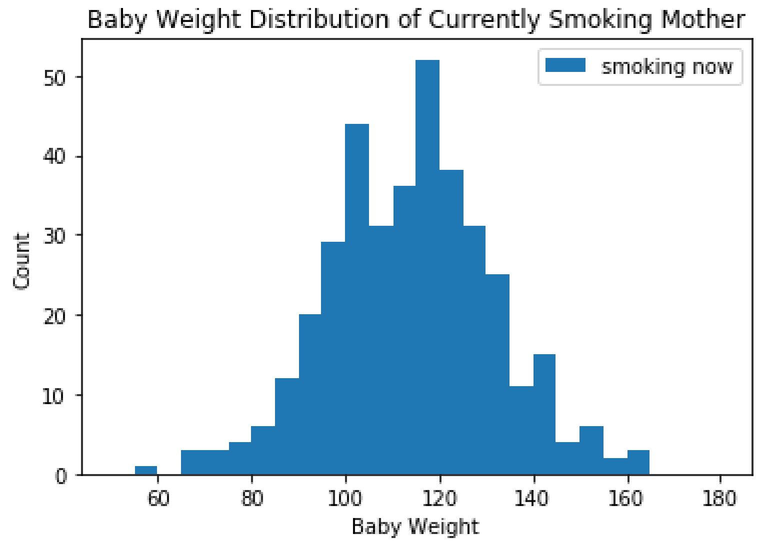
The boxplot(Figure 1) displays the distribution of babies weights of different mothers(never smoked, currently smoking, smoked until pregnant, once smoked) based on a five number summary(“minimum”, first quartile, median, third quartile, and “maximum”). From the plot, the median, first quartile, and third quartile of baby weights of currently smoking mother are lower than the other three. The baby weights of never smoked, smoked until pregnant, and once smoked mother are close. Note that there are many outliers for the baby weights of never smoked mother.

Figure 2: Histogram for baby weights of never smoked mother



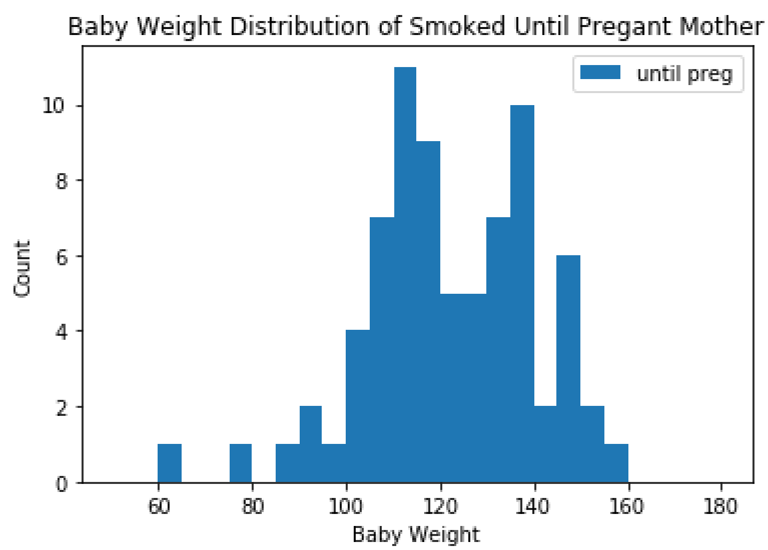
The histogram of baby weights of never smoked mother(Figure 2) is unimodal and symmetric, which suggests that the baby weights of never smoked mother is likely to be normally distributed. From the plot, most of the baby weights are between 110-140 ounce.

Figure 3: Histogram for baby weights of currently smoking mother



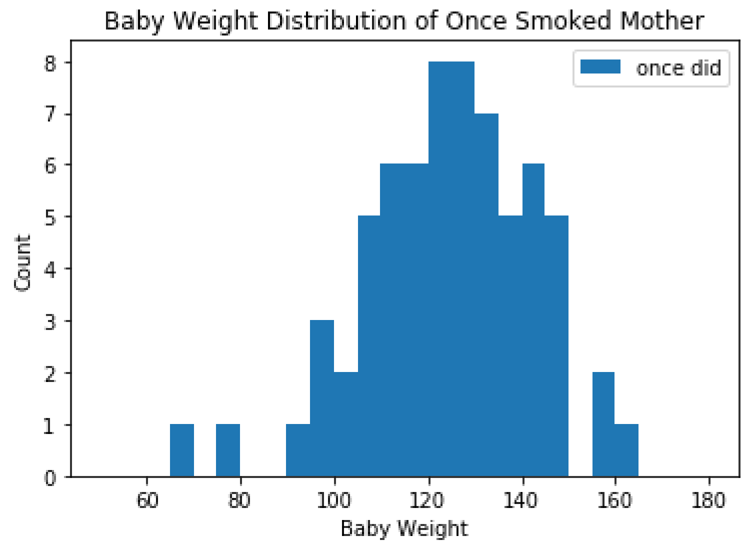
The histogram of baby weights of currently smoking mother(Figure 3) is bimodal and symmetric. From the plot, two peaks of the baby weights are around 100 ounce and 115 ounce.

Figure 4: Histogram for baby weights of smoked until pregnant mother



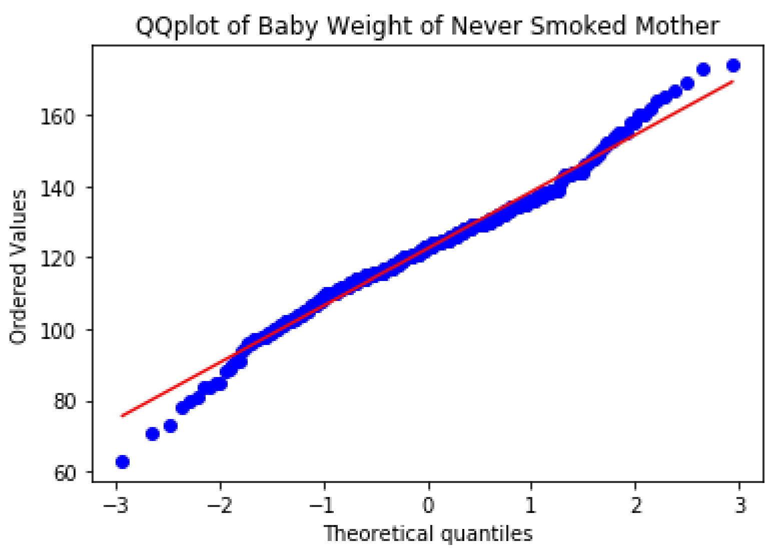
The histogram of baby weights of smoked until pregnant mother(Figure 4) is multimodal and symmetric, which suggests that the baby weights of smoked until pregnant mother has a large variety. From the plot, most of the baby weights are between 110-140 ounce.

Figure 5: Histogram for baby weights of once smoked mother



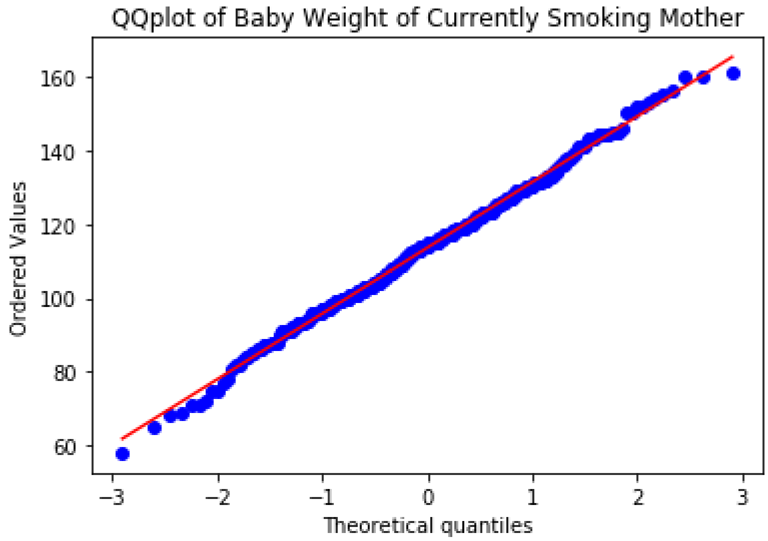
The histogram of baby weights of once smoked mother(Figure 5) is unimodal and symmetric, which suggests that the baby weights of once smoked mother is likely to be normally distributed. From the plot, most of the baby weights are between 105-150 ounce.

Figure 6: QQplot for baby weights of never smoked mother



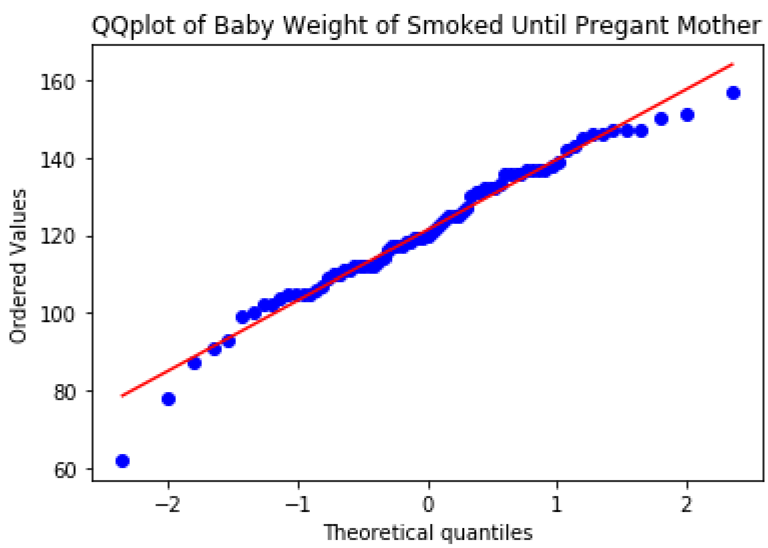
Quantile-Quantile Plot of baby weight of never smoked mother(Figure 6) is drawn to see if it came from normal distribution as expected. From the graph, we can see that the points forming a line is roughly straight, which suggests that it is likely to be a normal distribution. Performing KS-test(in the python code right below the qqplot), the p-value of this data is 0.1629928333642301, which indicates that it is normally distributed, though it has some outliers at both ends of the line.

Figure 7: QQplot for baby weights of currently smoking mother



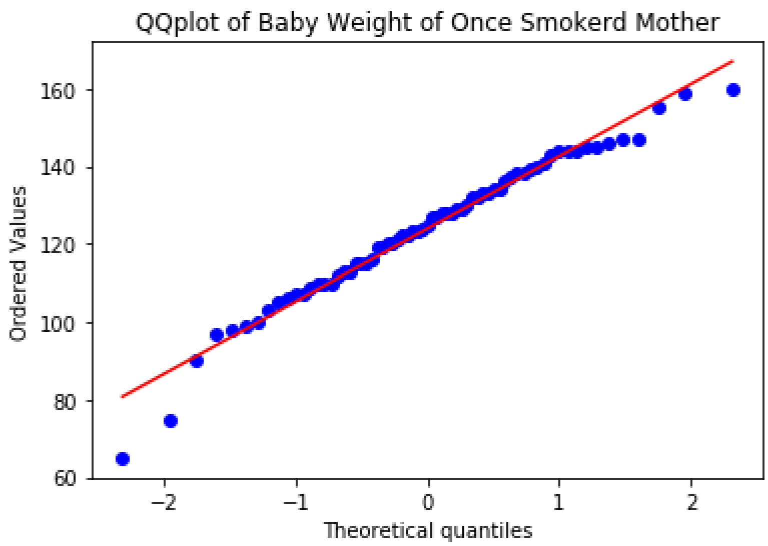
Quantile-Quantile Plot of baby weight of currently smoking mother(Figure 7) is drawn to see if it came from normal distribution as expected. From the graph, we can see that the points forming a line is roughly straight, which suggests that it is likely to be a normal distribution. Performing KS-test(in the python code right below the qqplot), the p-value of this data is 0.6383697462168512, which indicates that it is normally distributed.

Figure 8: QQplot for baby weights of smoked until pregnant mother



Quantile-Quantile Plot of baby weight of smoked until pregnant mother(Figure 8) is drawn to see if it came from normal distribution as expected. From the graph, we can see that the points forming a line is roughly straight, which suggests that it is likely to be a normal distribution. Performing KS-test(in the python code right below the qqplot), the p-value of this data is0.8032849422874576, which indicates that it is normally distributed.

Figure 9: QQplot for baby weights of once smoked mother



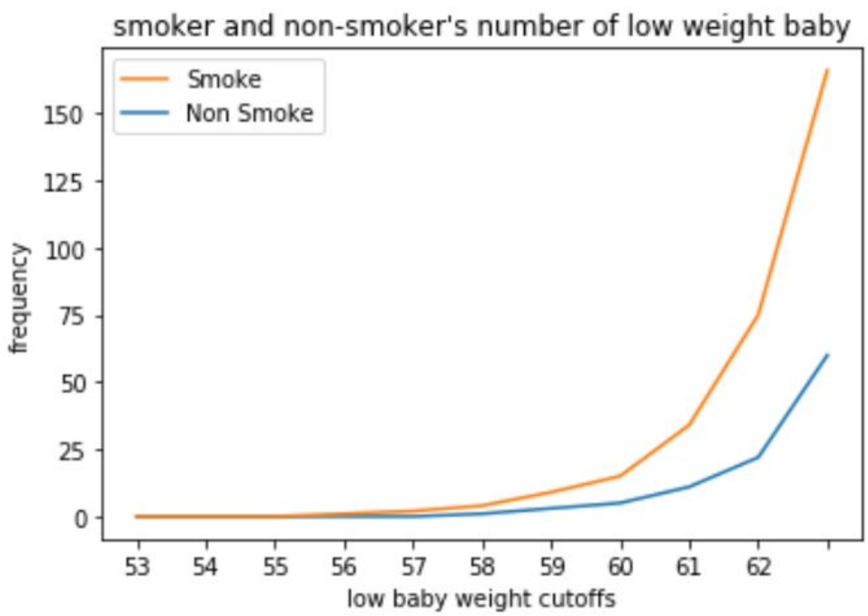
Quantile-Quantile Plot of baby weight of once smoked mother(Figure 9) is drawn to see if it came from normal distribution as expected. From the graph, we can see that the points forming a line is roughly straight, which suggests that it is likely to be a normal distribution. Performing KS-test(in the python code right below the qqplot), the p-value of this data is 0.9690488557137379, which indicates that it is normally distributed.

Frequency and Incidence:

Babies' weight might have an association with babies health, or even babies death rate. Maureen Hack, et al, suggest from their study that 1500g which is 53 ounces would counted as low weight birth, and that would affect babies survival rate (pediatrics). For this part, our main focus would be how the number of babies belong to low weight range changes if we alternate the definition of a low weight baby given the mother's smoking habit. Gathered from previous researchers' works, we find out we could use the range [53,63] in ounces to test on the frequency's changes. First, we check on the number of low weight babies whose mother smokes and with those whose mother does not smoke. We filter out one of the confounding factors from the data, which is the mother's age that might potentially affect the babies' weight.

	frequency	total	incidence	low_weight_cut
non_smoker	0	416	0.000000	53
smoker	0	518	0.000000	53
non_smoker	0	416	0.000000	63
smoker	2	518	0.003861	63

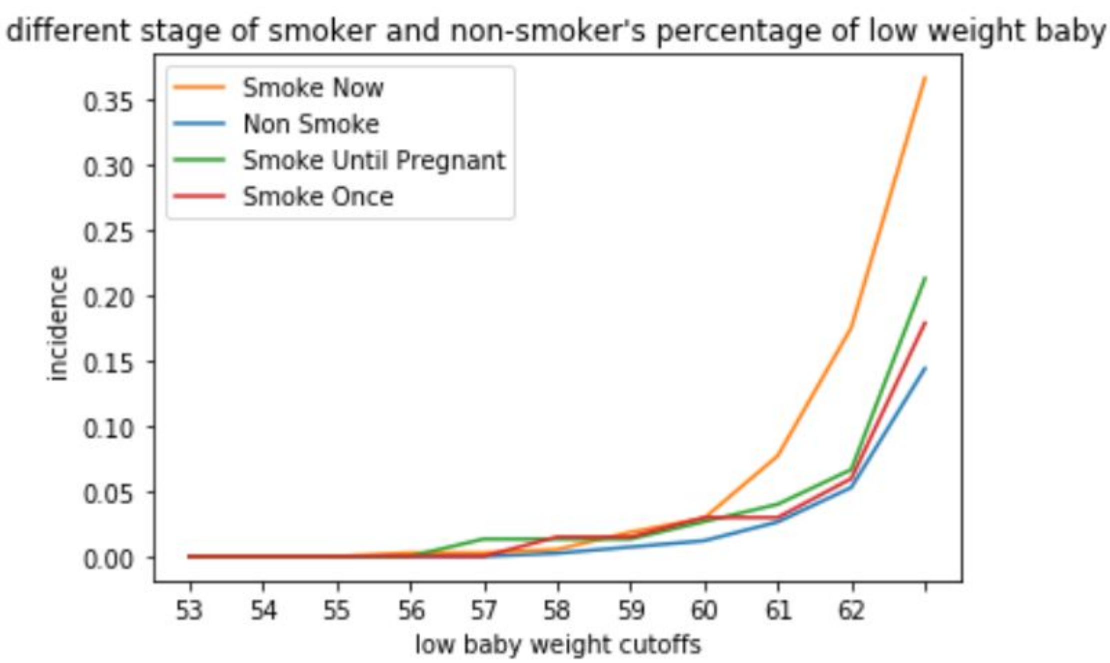
Usually, mother who gives birth at a younger age or at an older age might have a higher chance to have a baby with poor health condition. Institute has identified that if the mothers' age are less that 17 or above 35, then they would be considered at high-risk to give birth to babies (ssmhealth.com). If that is the case, we decided to only select the group of mothers who are under age that is normal to be pregnant in order to subside the effect from the confounding factor. We choose the range of mother's age that is within 17 and 35 for several reasons. First, this group has the greatest amount of data for us to analyze. Second, this group would considerably to be the common group so that the babies' weight could be less affected by the age of the mother. After cleaning and filtering the data, we create a table which have the number of the low weight babies for both non-smoker mother, and smoker mother, setting the cutoff initially at 53oz then to 63oz. From the table, we could not find enough information to see how the cutoffs would influence these two groups. Then, we decided to increase the cutoffs gradually in order to see the different increasing rates for the two groups.



As what we expected, the smoker group would have more and more babies counted as low weight babies, while the non smoker groups increasing rate is much slower than the smoker group. Continuing on this process, we further group the smoker mothers to different smoking stages. Some of them only smoked once in their life, some of them smoke until they are pregnant, and some of them keep smoking no matter if they are pregnant or not.

	frequency	total	incidence	low_weight_cut
non_smoker	0	416	0.000000	53
smoker_smoke_now	0	376	0.000000	53
smoker_untill_prg	0	75	0.000000	53
smoker_once	0	67	0.000000	53
non_smoker	0	416	0.000000	63
smoker_smoke_now	1	376	0.002660	63
smoker_untill_prg	1	75	0.013333	63
smoker_once	0	67	0.000000	63

The result is that mothers who keeps smoking would have more low weight babies compared to other groups. Surprisingly, the frequency of low weight baby for non smoking mothers is higher than the other two groups(smoking once, and smoking until pregnant). One possible explanation for that is the total number of data points for these groups are different. In that case, the frequency would not be a good indicator. We then, plot the incidence rather than the frequency to see the ratios of low weight babies for these groups.



This time, the result shows that there is not that much difference between “never smoke mother”, “smoke once”, and “smoke until pregnant mothers” incidence of low weight baby. Mothers who smoking until now have the highest incidence increasing rate compared to the rest of the three groups.

Advanced Analysis

From numerical analysis, we see that in cases of common pregnancies, the average birth weight of babies born to mothers who smoke during pregnancy is almost 10 ounces lower than those of babies born to mothers with other smoking stage. To investigate the statistical significance of difference between 4 birth weights, it makes sense to do hypothesis testing to compare their means.

In order to compare the means of two populations, it makes sense to do two-sample t test. We need to check that the assumptions of two-sample t test are met. First, from the ks-test of normality, we fail to reject the null hypothesis that the distribution is normal for birth weights of babies born to mothers with all four smoking stages. Therefore, we can assume that the distributions of birth weights are all approximately normal. Second, it is reasonable to assume that smokers and nonsmokers samples are independent. Infants birth weights of smoker mothers should not affect those of nonsmoker mothers. Third, as these mothers were all enrolled in Kaiser Health Plan, samples might not be simple random samples from the San Francisco. However, common pregnancy samples can be seen as simple random samples from all mothers of age ranging from 17 to 30 in Kaiser Health Plan. Thus, the results of hypothesis testing can only be generalized to mothers of common pregnancy in Kaiser Health Plan but not the entire San Francisco.

With assumptions all met, we can now perform two-sample t test. The first mission is to investigate the variances of each two samples. As distributions are all normal, it is safe to perform F test of equality of two variances. Here, let X = the population in which the first sample is drawn, and Y = the population in which the second sample is drawn. The null hypothesis is $H_0 : var(X) = var(Y)$, and the alternative hypothesis is $H_1 : var(X) \neq var(Y)$. We set the significance level to be 0.05. The table below summarizes the p-values and decisions of each F test.

Table 1: F test of equal variances results of different samples

Sample 1	Sample 2	P-value	Decision
Never Smoke	Yes Now	**0.04069**	Reject
Never Smoke	Until Preg	0.16213	Fail to reject
Never Smoke	Once	0.10370	Fail to reject

Yes Now	Until Preg	0.83611	Fail to reject
Yes Now	Once	0.63419	Fail to reject
Until Preg	Once	0.82848	Fail to reject

From the results of F test, we only reject the null hypothesis that the variances are equal when comparing variances of birth weights of babies born to mothers who never smoke to those born to mothers who always smoke. Therefore, we should do Welch's t-test with unequal population variances to test means of these two groups and do student t-test with equal variances for other groups. Again, we let X = the population in which the first sample is drawn, and Y = the population in which the second sample is drawn. This time the null hypothesis is $H_0 : \mu_X = \mu_Y$, and the alternative hypothesis is $H_1 : \mu_X \neq \mu_Y$, where μ_X is the mean of the first population, and μ_Y is the mean of the second population. We set the significance level to be 0.05. The table below summarizes the p-values and decisions of each two-sample t test with equal or unequal mean.

Table 2: Two-sample t test results of different samples

Sample 1	Sample 2	P-value	Decision
Never Smoke	Yes Now	**4.87831 * 10^(-13)**	Reject
Never Smoke	Until Preg	0.57397	Fail to reject
Never Smoke	Once	0.50984	Fail to reject
Yes Now	Until Preg	**0.00066**	Reject
Yes Now	Once	**1.72146 * 10^(-5)**	Reject
Until Preg	Once	0.40253	Fail to reject

Therefore, we reject the null hypothesis that two population means are equal in 3 out of 6 cases. We can conclude that there are statistically significant difference between the means of birth weights of babies born to mothers who always smoke and those born to mothers of other smoking stages. However, there is no significant difference between the means of birth weights related to mothers who never smoke, smoke until pregnancy, and have smoked once.

Furthermore, in order to know whether infants born to mothers who smoke all the time have lower or higher birth weights, we now do one-sided instead of two-sided hypothesis testing. Making our alternative hypothesis $H_1 : \mu_X > \mu_Y$, we find that babies born to consistently smoking mothers have significantly lower birth weights than babies born to non-smoker mothers,

mothers who have smoked once, and mothers who smoke until pregnancy, with p-values being 2.43916*10^(-13), 8.60731*10(-6), and 0.00032 respectively.

Thus, the results of hypothesis testing matches our hypothesis. Infants born to mothers who smoke all the time indeed have lower weights than infants whose mothers of other smoking stages. Besides, for mothers who do not smoke all the time, the means of their babies' birth weights are not significantly different.

Theory

Numerical Statistics:

Mean: Mean is a measure of the central tendency of the distribution.

Standard Deviation: Standard deviation is a measure of the variation of a set of data values.

Quantile: Quantiles are cut points dividing the range of a probability distribution into intervals with equal probabilities.

Skewness: Skewness is a measure of the asymmetry of the probability distribution. It is calculated as the third standardized moment. Skewness can indicate which direction and how far a distribution deviates from normal.

Kurtosis: Kurtosis is a measure of "tailedness" of the probability distribution and is calculated as the fourth standardized moment. Like skewness, Kurtosis is another way of indicating how far a distribution deviates from normal by measuring the thickness of the tail.

Graphs:

Boxplot: A boxplot is a standardized way of displaying the distribution of data based on a five number summary ("minimum", first quartile (Q1), median, third quartile (Q3), and "maximum"). It can tell you about your outliers and what their values are. It can also tell you if your data is symmetrical, how tightly your data is grouped, and if and how your data is skewed. By comparing boxplots of different groups of data, we can receive a direct visual relations among different distributions.

Histogram: Histogram is an accurate representation of the distribution of numerical data by displaying with bars of different heights. It groups numbers into bins of ranges and plot the height of each bar by the frequency of data points within each bin. With histogram, we can have a direct visualization of the distribution of the data.

Q-Q plot: Q-Q plot stands for quantile-quantile plot. It is a graphical method to compare two probability distribution by plotting their quantiles against each other. By setting one of probability distributions to normal distribution, we can visualize the deviation of the target distribution from normal.

Testing Method:

Kolmogorov–Smirnov test: KS test is a nonparametric test of equality of probability distributions of two samples . The KS statistic quantifies the distance between the empirical distribution of the sample and the cumulative distribution function of the reference distribution (Normal Distribution in this case). Thus, it can be used as a more formal way to determine the normality of the probability distribution by setting one of the distributions to be normal.

Two Sample t-test: The two sample T-test is often used for evaluating the means of two variables or distinct groups, providing information as to whether the means between the two populations differs. In this case, t-test is preferred over z-test because the variance of the data is known.

Conclusion and Further discussion:

From the above analysis, we can infer that there is a strong relationship between (mothers who) currently smoking and the low babies weight. And according to studies, low babies weights is a strong indicator of babies' health. Note that since the data is not collected in a controlled experiment, we cannot establish a causal relationship between the smoking status of the mothers and the low baby weights. Furthermore, (The mother who) once smoked does not show a strong relationship with low baby weight, however, we cannot conclude that smoking prior to pregnancy has no effect on baby weights. Therefore, further experiments are needed in order to establish a causal relationship between smoking and low babies weights, but maternal smoking is something that we need to be cautious about when considering the health of the offspring.

Through this case, we find out the relationship between mother's smoking habit with their babies' weight. Alghouth babies' weight is one of the good indicators for babies health. There are lots of other variables that we could pick as indicators for babies' health, such as gestation, babies length, and etc. For future study, we could change different variables to see mother's smoking habit relates to those indicators. In that way, we could get a full picture of how mother's smoking habit could affect babies' health in multiple ways. In addition, there are still some confounding factors in the study. We only eliminate mother's age as one of them. In future studies, we could try to filter the data in more detail in order to get a more homogeneous group. In that way, the result would be more towards the relationship between the desired variables with smoking habit, and less influence from other factors that we do not want to include.

Works Cited:

1."Age and Pregnancy." SSM Health Saint Louis University Hospital | SSM Health, www.ssmhealth.com/maternity/high-risk-pregnancy/age-pregnancy.

2.Hack, Maureen, et al. "Very Low Birth Weight Outcomes of the National Institute of Child Health and Human Development Neonatal Network." Pediatrics, American Academy of Pediatrics, 1 May 1991, pediatrics.aappublications.org/content/87/5/587.short.

3. Pourhoseingholi MA, Baghestani AR, Vahedi M. How to control confounding effects by statistical analysis. Gastroenterol Hepatol Bed Bench 2012;5(2):79-83.

4. Kenneth C. Schoendorf and John L. Kiely. "Relationship of Sudden Infant Death Syndrome to Maternal Smoking During and After Pregnancy." American Academy of Pediatrics, 1992.

5. Michael H. Malloy, et al. , "The Association of Maternal Smoking with Age and Cause of Infant Death." *American Journal of Epidemiology*, Volume 128, Issue 1, 1 July 1988, Pages 46–55, https://doi.org/10.1093/oxfordjournals.aje.a114957.

6. Allen J. Wilcox, "Birth Weight and Perinatal Mortality: The Effect of Maternal Smoking", *American Journal of Epidemiology*, Volume 137, Issue 10, 15 May 1993, Pages 1098-1104, https://doi.org/10.1093/oxfordjournals.aje.a116613.

7. Alam, N., "Teenage Motherhood and Infant Mortality in Bangladesh: Maternal Age_dependent Effect of Parity One". *Journal of Biosocial Science*, 32(2), 229-236.