Case Study 1: Maternal Smoking and Infant Death

Group members:
Luying Jiang A13513110
Siyao Mi A13806003
Nuo Chen A13709502
Jining Jin A13688539
Zhijian Hu A92070159

Introduction

Despite the fact that it is commonly acknowledged that smoking brings harmful effects to multiple aspects in people's life, the influence of maternal smoking is especially apparent in raising infants' risks towards various health problems. As in ways to evaluate infants, babies' maturity, measured by birth weight and gestational age, is considered a significant factor since babies both born early and small have lower survival rates.[1]

The objective of this study is to evaluate the relationship between maternal smoking and babies maturity through investigations in the influence of maternal smoking on birth weight and gestational age separately, including the consideration of connection between these two factors.

Data

The data collected for our study is part of the Child Health and Development Studies (CHDS), which included all pregnancies that occurred between 1960 and 1967 among women in Kaised Health Plan in Oakland, California. The subjects involved in this study are the women who enrolled in Kaiser Health Plan, acquired prenatal care in San Francisco area and delivered in any of the Kaiser hospitals in Northern California. The length, weight and head circumstance of 1236 babies in this data are obtained at birth measurements. All of the babies are boys who lived at least 28 days and are in single births.

In order to demonstrate our hypothesis, our group utilize the specific data about birth weight, gestation period and whether or not the mother smoked. Among these variables, weight and gestation period in our case are numerical and discrete variables; smoking status is categorical variable.Besides, in the data of babies.text, 0 means no, 1 means yes and 9 means unknown in the column of smoke. In the data of babies23.text, 0 means never, 1 means yes now, 2 means smoking until pregnancy, 3 means smoking once before and 9 means unknown in the column of smoke.

In further support our investigation and avoid some possible items that can mask the effect of smoking on the low weight and gestational age, our group eliminates some confounders from our primary data sources.

For the part of babies' weight,

1)we exclude those babies whose mother's age are under 20 and over 40 since mother's age will influence the babies' weight[2];

2)We exclude those babies whose mother's Body Mass Index (BMI) is under 18.5, since those mothers can be the confounder of low birth weight[3].

---

[1] Lecture note page5.

[2] Explaining differences in birth outcomes in relation to maternal age: the Generation R Study, 2011

[3] Rick Scott, Pre-Pregnancy Body Mass Index and Related Maternal Health and Infant Outcomes Among Mothers in Florida, 2009-2011

The formula of BMI= $\frac{Weight*0.45}{(height*0.025)^2}$

3)We also exclude the unknown smoking condition from the data set, since the data includes some unknown values, such as 9 in smoke column of babies23.text

After eliminating above confounders in gestational age, our group initiated a new scope of data for analyzing the effect of smoking upon gestational age. There are 811 samples in our new scope.

For the part of gestational age,

1)We exclude those babies whose mother's Body Mass Index (BMI) is over 30 since compared with normal weight mothers, overweight mothers can be the confounder of our project[4].

2)We also exclude some inaccuracy outliners, such as 9 in smoke column and 999 in gestation days.

After eliminating above confounders in gestational age, our group initiated a new scope of data for analysing the effect of smoking upon gestational age. There are 1205 samples in our new scope.

Background

According to the lecture notes, since babies born early and small have lower survival rates[5], both birth weight and gestational age are indispensable to infant mortality. Based on the *National Vital Statistics Reports*, baby born before 37 weeks are defined as preterm delivery; babies born at less than 2500 grams are defined as low birth weight[6]. And maternal smoking will lead to the result of "substances such as nicotine, hydrogen cyanide, and carbon monoxide from the placenta into the fetal blood supply"[7], which in turn can impede babies' healthy growth. In light of this, to further analyze the problem of fetal growth in babies, our group decides to investigate whether smoke could affect fetal growth. In other words, we want to investigate the significance of difference between birth weight and gestational age born to smokers and born to nonsmokers.

Investigations (weight)

Numerical analysis

Table 1 contains the numerical comparison (mean, median, standard deviation and etc.) the two distributions: the weights of the babies born to non-smoking mothers and the weights of the babies born to mothers who smoke during their pregnancy period. The data contains more samples of the non-smoking mothers. From the table, we can see that the median, mean, 1st quantile and the 3rd quantile of the birth weight from the babies born to smoking mothers are lower than those from non-smoking mothers.

---

4 Rick Scott, Pre-Pregnancy Body Mass Index and Related Maternal Health and Infant Outcomes Among Mothers in Florida, 2009-2011

5 Lecture note page 5

6 Lecture note page 10

7 Infant Mortality Statistics from the 2001 Period Linked Birth/Infant Death Data Set

	Babies' weights (Non-smoking mother)	Babies' weights (Smoking mother)
Count	638	400
Minimum	95.0	96.0
1st quantile	117.0	114.0
Median	126.0	121.0
Mean	127.0	122.9
3rd quantile	135.2	131.0
Maximum	174.0	163.0
Standard deviation	14.05024	14.10361
Skewness	-0.16624	-0.02389
kurtosis	1.00673	-0.0169

Table 1: numerical statistics of birth weights for two baby groups

The box plot(Figure 1) below shows the baby birth weights of mothers who have different smoking habits during the pregnancy period. The 1st quantile and the 3rd quantile for the birth weights of the babies born to smoking mothers are lower than that of the babies born to non-smoking mothers. In addition, we can clearly see that the left box also indicates that there are many outliers for the birth weights of babies born to non-smoking mothers.

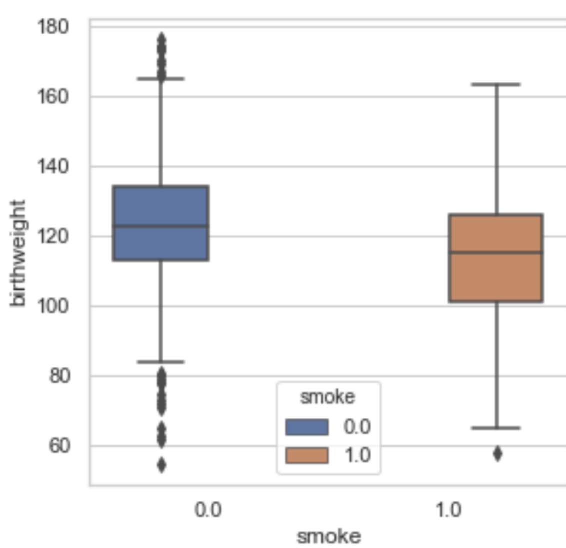


Figure 1: Boxplot for two distribution

The Figure 2 shows the normal distribution of these two categories. As we can see from the graph, the distribution of the baby weights born to smoking mothers are unimodal and generally symmetric.This distribution has one mode about 115. The shape of the histogram to the left of the peak looks roughly like the miser image of the part of the histogram to the right of the peak. This histogram indicates a few outliers. Similarly, the distribution of the baby weights from non-smoking mothers are also unimodal and generally symmetric. But there is a slight difference between these two distributions. The density of the weights of babies born to non-smoking mother left skewed (more babies have higher weights) and that of babies born to smoking mothers right skewed comparatively. Moreover, the mode for nonsmoking group is 125, which is higher than that of smoking group in general.

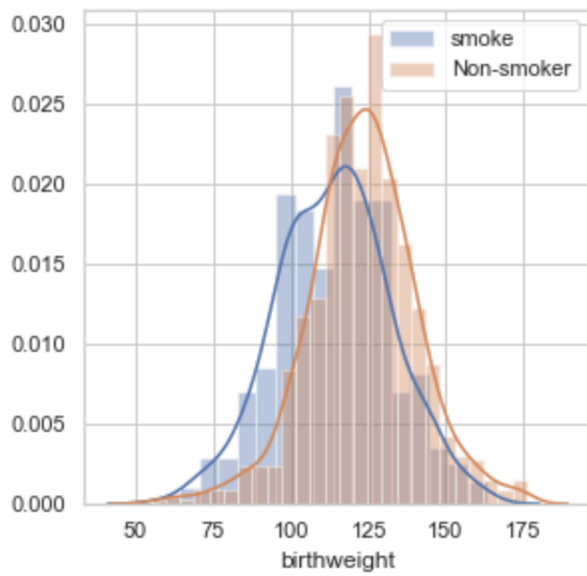
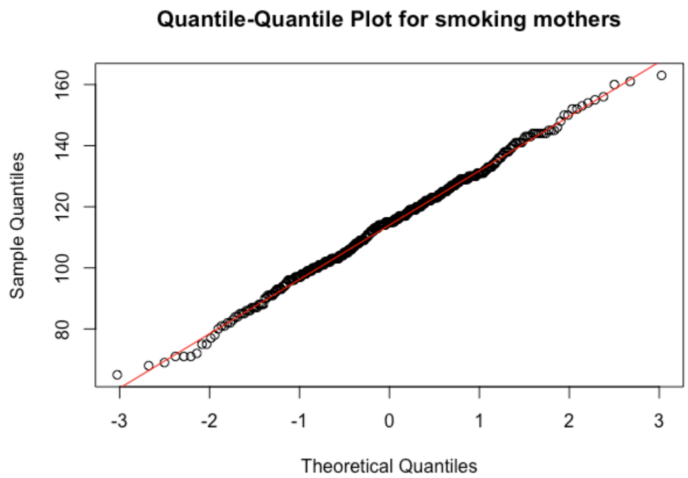


Figure 2: Histograms and density curve comparison between two different groups

In order to check whether the two data sets come from a common distributions, we draw a Quantile-Quantile Plot comparison for the weights of the babies born to smoking mothers and non-smoking mothers during their pregnancy. The Quantile-Quantile Plots for smoking mothers below suggest that while there are some points that are specifically on the comparison line at the end, most of the points do lie on the line. So this tells us that the weight of babies born to non-smoking mothers during their pregnancy are most likely normally distributed. (Figure 3) Similarly, as we can see from the QQ plot for smoking mothers (Figure 4), there is also a normal distribution with majority of plots in the graph lie on the y=x straight 45 degree line with few exceptions.



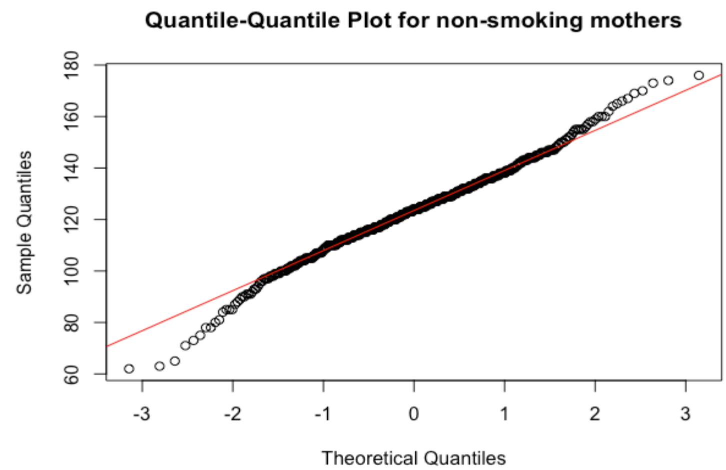


Figure 3: Quantile-Quantile Plot for Smoking mothers   Figure 4: Quantile-Quantile Plot for Non-smoking mothers

Then in order to clearly demonstrate the difference between smoking and non-smoking groups, we did another Q-Q plot to further investigate their relationship. From the figure 5, the plots lies roughly above the 45 degree straight line, which indicates that the two distributions are slightly different but roughly linear related. The mean is roughly equal to the median.

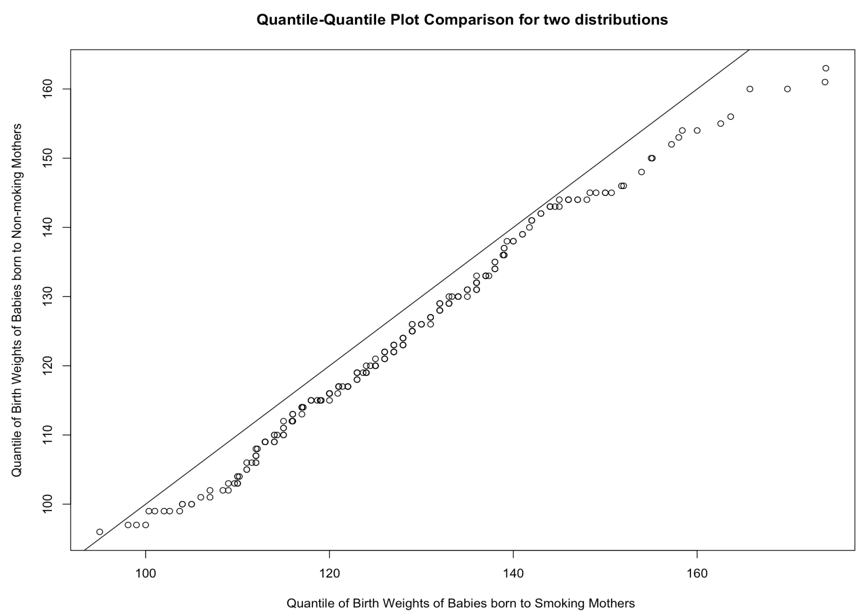


Figure 5: Quantile-Quantile Plot Comparison for the two distributions

<u>Frequency and incidence</u>

The bar plot (Figure 6) shows the frequency of low birth weights in the two groups. The non-smoking-mother group has a higher frequency of normal weight and a lower frequency of low birth weight at 88.2 ounces threshold. We can know that the proportion of low birth weight babies in the whole group is higher in the smoking-mother group. After calculation, according to

Table 2, mothers who smoke during pregnancy indeed have a higher frequency of low birth weight rate (0.08264) compared with mothers who do not smoke during the pregnancy (0.03100).

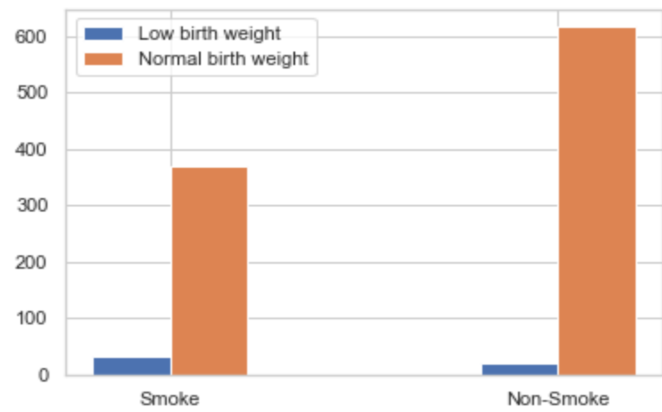


Figure 6: Frequency of low birth weight babies in each group with cutoff: 88.2 ounces

	Non-smoking mother	Smoking mother
Proportion	0.03100	0.08264

Table 2: percentage of low birth weight babies in each group

To test the reliability of the estimates, we changed the threshold to have a fewer more or fewer babies of the weight that is classified as low birth weight. In this case, we can see the changes in incidence as the classification standard changes. The thresholds range from 86 ounces to 91 ounces. However, in table 3, the changes are minimal. The mothers who smoke during pregnancy still have a higher frequency of low birth weight compared with the mothers who do not smoke during the pregnancy.

	Non-smoking mother	Smoking mother
Proportion (Threshold 86 ounces)	0.02830	0.05372
Proportion (Threshold 87 ounces)	0.02830	0.06198
Proportion (Threshold 89 ounces)	0.03100	0.08264
Proportion (Threshold 90 ounces)	0.03369	0.08264
Proportion (Threshold 91 ounces)	0.03639	0.08678

Table 3: percentage of low birth weight babies in each group

Importance of the Difference

Despite the fact that the minimum value for the babies, the numerical statistics suggest that babies born to smoking mothers during the pregnancy have relatively lower birth weight in 1st quantile, 3rd quantile, population mean and median.

The graphical methods also support this argument. As we can see in the boxplot and the histogram, there is a significant difference in median and mean. We can know from the Quantile-Quantile plot that the distributions for the two populations have the same shape. It also suggests that each distribution from the two population follows the normal distribution. With the standard of low weight at 88.2 ounces, the bar plot shows that the proportion of underweight babies born to smoking mothers is higher than that to non-smoking mothers. Also, after changing the thresholds, the incidence graphs suggests that the difference is significant and robust.

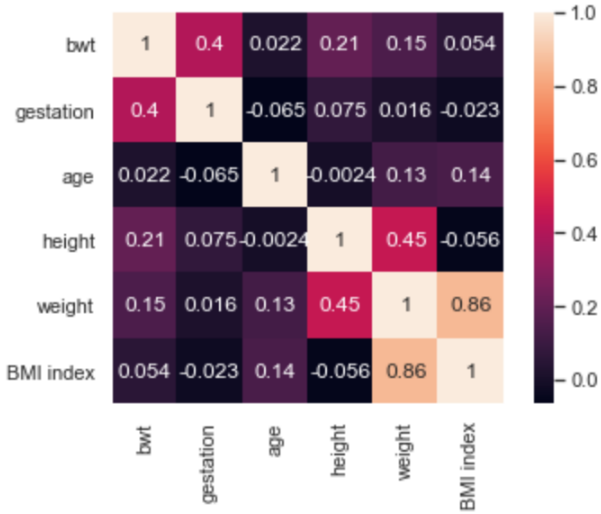
Hence, the results from the above method and graphs show a strong association between the smoking status of mothers and the birth weight of the babies. If mother smoked during the pregnancy period, there is a higher chance to give birth to a lower weight baby. According to NVSR, after excluding the effect of confounders, they get the result that “Infant mortality rates were much higher for low-birthweight infants than for infants with birthweights of 2,500 grams or more for all race and ethnic groups studied. Overall, the infant mortality rate for very-low-birth-weight infants (those with birthweights of less than 1,500 grams) was 244.4, more than 100 times the rate for infants with birthweights of 2,500 grams or more.”[8] In other words, NVSS thought that the low birth weight exert a negative effect on fetal growth and has a greater chance of high mortality among babies. Related with our group project, through investigating the relationship between smoking status and birth weight, the differences of three types of comparisons is of greater importance because all of them cogently build the relationship between low birth weight and maternal smoking, which in turn further strengthen our hypothesis that the smoking status can influence the birth weight and then babies’ survival rate.

Additional Analysis:

The above analysis gives the general relationship between birthweight and smoke, but further investigation needs the research to eliminate the confounder since gestation period may naturally contribute to larger birth weight. The correlation is also suggested in journal “Hirve , Siddhi S., and Bela R. Ganatra. *DETERMINANTS OF LOW BIRTH WEIGHT: A COMMUNITY BASED STUDY*.” The correlation between gestation and birth weight is significant enough to be addressed according to the “correlation heat map” below:

[8] Reexamining the effects of gestational age, fetal growth, and maternal smoking on neonatal mortality

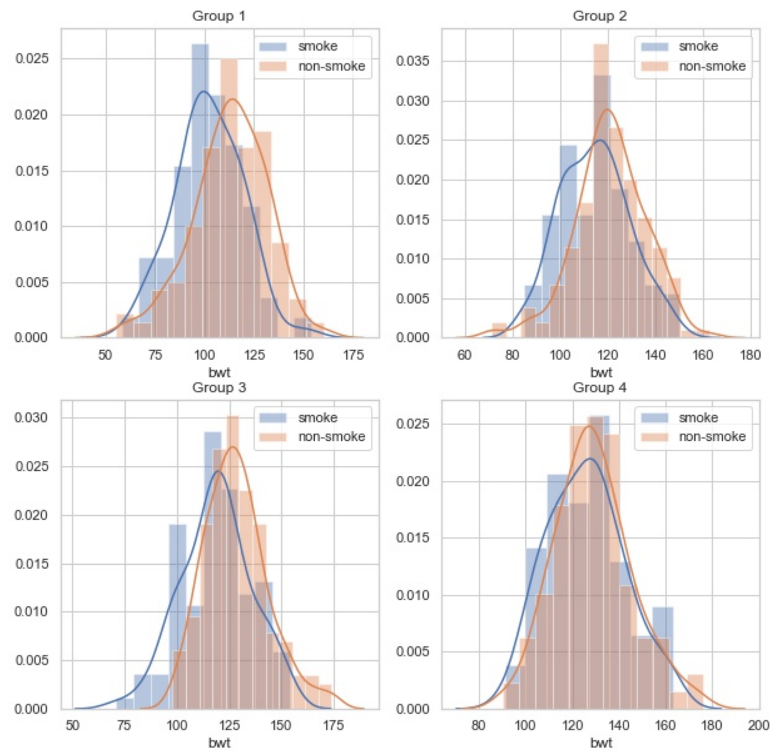
```
<matplotlib.axes._subplots.AxesSubplot at 0x1b55b301470>
```



Note: From the figure, mother's weight and height are closely related, but this correlation will not be addressed in the project.

The first step is to divide the overall population according to quantile: group 1is minimum to 1st quantile(25%), group 2 is data from 1st quantile to mean(25%-50%), group3 is data from mean to 3rd quantile and group3 from 3rd quantile to maximum(75%-100%).

The second step is to further divide the data for each group to two partitions smoke and non-smoke.The project implemented statistics method including kurtosis, skewness and Kolmogorov Smirnov Test to determine whether the distribution is a normal distribution. Figure below presents the general distributions of partitions:



Group 1:
Distribution Simulation:

	0	1	2	3
0	kurtosis for smoke:	0.021703	skewness for smoke:	0.024476
1	kurtosis for Non-smoke:	0.336489	skewness for Non-smoke:	-0.420069

```
stats.kstest(normalize(s1['bwt']), 'norm')
```

```
KstestResult(statistic=0.05195875619612467, pvalue=0.885671756921O855)
```

```
stats.kstest(normalize(ns1['bwt']), 'norm')
```

```
KstestResult(statistic=0.06914297887844378, pvalue=0.4138208394658598)
```

Given that kurtosis and skewness is equal to 0 in normal distribution and both distribution p-value is larger than 0.05 significant level, we can simulate the distributions to a normal distribution.

Bootstrap step:

Implement resample method to perform 1000 times of resampling Implement resample method 1000 times for both Smoke group and Non-smoke partitions in group 1 and calculate p-value through two sample T-test. Then calculate the mean of 1000 statistics.

Since the two distributions are close to normal distribution, we perform two sample T-test under the assumption that variance is unknown and variances of two groups are not equal.

Set up the Null Hypothesis H_0: $\mu_1=\mu_2$ and Alternative Hypothesis H_1: $\mu_1>\mu_2$

```
print('The mean p-value is: ', np.mean(bt))
```

```
The mean p-value is:  0.0006985700770165726
```

Because the p-value is far less than significant level $\alpha=0.05$, we reject H_0 and conclude that mean of birth weight for non-smoking mothers is significantly larger than smoking mothers.

Group 2

Distribution Simulation:

	0	1	2	3
0	kurtosis for smoke:	-0.456983	skewness for smoke:	0.151926
1	kurtosis for Non-smoke:	0.696749	skewness for Non-smoke:	-0.306987

```
stats.kstest(normalize(s2['bwt']), 'norm')
```

```
KstestResult(statistic=0.07160879598278153, pvalue=0.529534536559876)
```

```
stats.kstest(normalize(ns2['bwt']), 'norm')
```

```
KstestResult(statistic=0.05850244374565028, pvalue=0.5970564375272418)
```

Given that kurtosis and skewness is equal to 0 in normal distribution and both distribution p-value is larger than 0.05 significant level, we can simulate the distributions to a normal distribution.

Since the two distributions are close to normal distribution, we perform two sample T-test under the assumption that variance is unknown and variances of two groups are not equal.

Set up the Null Hypothesis H_0: $\mu_1=\mu_2$ and Alternative Hypothesis H_1: $\mu_1>\mu_2$

```
print('The mean p-value is: ', np.mean(bt2))
```

```
The mean p-value is:  0.0059093854720566955
```

Because the p-value is less than significant level $\alpha=0.05$, we reject H_0 and conclude that mean of birth weight for non-smoking mothers is significantly larger than that of smoking mothers.

Group 3

Distribution Simulation

	0	1	2	3
0	kurtosis for smoke:	-0.176585	skewness for smoke:	-0.124142
1	kurtosis for Non-smoke:	0.424900	skewness for Non-smoke:	0.569729

```
print(stats.kstest(normalize(s3['bwt']), 'norm'))
print(stats.kstest(normalize(ns3['bwt']), 'norm'))
```

```
KstestResult(statistic=0.060510232897182326, pvalue=0.8533512474725926)
KstestResult(statistic=0.06625317604580283, pvalue=0.38734800406728986)
```

Given that kurtosis and skewness is equal to 0 in normal distribution and both distribution p-value is larger than 0.05 significant level, we can simulate the distributions to a normal distribution.

Since the two distributions are close to normal distribution, we perform two sample T-test under the assumption that variance is unknown and variances of two groups are not equal.

Set up the Null Hypothesis H_0: $\mu_1=\mu_2$ and Alternative Hypothesis H_1: $\mu_1>\mu_2$

```
print('The mean p-value is: ', np.mean(bt3))
```

```
The mean p-value is:  0.0006449392814525881
```

Because the p-value is far less than significant level, we reject H_0 and conclude that mean of birth weight for non-smoking mothers is significantly larger than smoking mothers.

Group 4

Distribution simulation

	0	1	2	3
0	kurtosis for smoke:	-0.438254	skewness for smoke:	0.255058
1	kurtosis for Non-smoke:	0.122453	skewness for Non-smoke:	0.349934

```
print(stats.kstest(normalize(s4['bwt']), 'norm'))
print(stats.kstest(normalize(ns4['bwt']), 'norm'))
```

```
KstestResult(statistic=0.05580891386345749, pvalue=0.9515974991524998)
KstestResult(statistic=0.05198208287427053, pvalue=0.7347264326097929)
```

Given that kurtosis and skewness is equal to 0 in normal distribution and both distribution p-value is larger than 0.05 significant level, we can simulate the distributions to a normal distribution.

Since the two distributions are close to normal distribution, we perform two sample T-test under the assumption that variance is unknown and variances of two groups are not equal.

Set up the Null Hypothesis H_0: $\mu_1=\mu_2$ and Alternative Hypothesis H_1: $\mu_1>\mu_2$

```
print('The mean p-value is: ', np.mean(bt4))
```

```
The mean p-value is:  0.261822952679899
```

The p-value is larger than significant level, we fail to reject H_0. Therefore, there is no significant difference between mean of birth weight for Non-smoking mothers and Smoking mothers based on the sample.

In the investigation above, it can be proved that differences of means of baby birth weight for smoke and non-smoke group are significant enough in Group 1, Group 2 and Group 3. However, we cannot prove significant difference in Group 4 (75%~Maximum). Therefore, we can conclude that smoke has critical influence on baby birth weight when gestation is less than 75% of sample (<288 days), but it has minor impact on baby birth weight when gestation is above 75 percentile.

Investigations (gestation)

Numerical analysis

Table 4 below contains numerical distributions of gestational age for babies born to women who smoked during their pregnancy and for babies born to women who did not smoke during their pregnancy. Larger size of samples are contained in babies born to non-smoking mothers. As the comparison of median and mean from two sets of data shown, maternal smoking tends to shorten babies gestational age. Also the 1st quantile and 3rd quantile of gestational age distribution are both lower in the smoking group with differences of 2 days and 3 days.

	Babies gestational age (days) (Non-smoking mother)	Babies gestational age (days) (Smoking mother)
Count	602	404
Minimum	148.0	223.0
1st quantile	273.0	271.0
Median	281.0	279.0
Mean	288.9	283.9
3rd quantile	289.0	286.0
Standard deviation	80.72386	67.24155
Skewness	-1.071203	-0.2293753
Kurtosis	11.73105	5.083699

Table 4: numerical statistics of gestational age for two baby groups

Graphical Methods

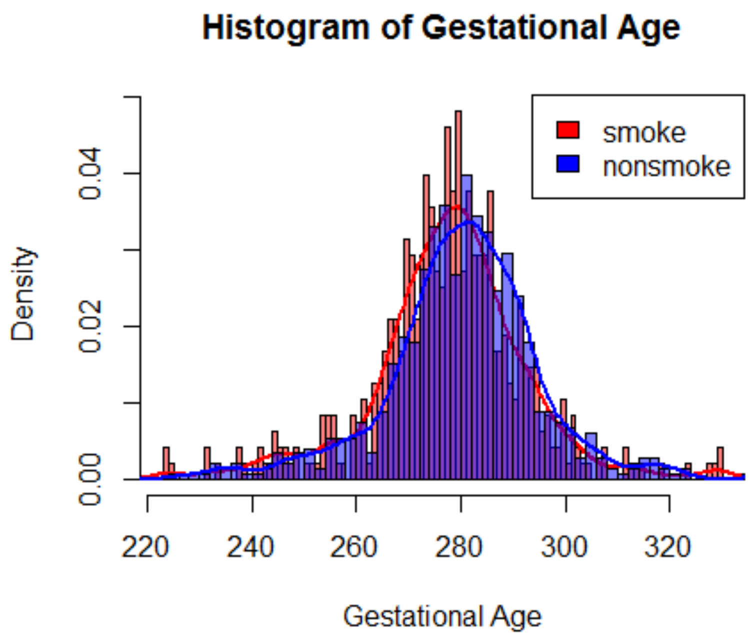


Figure 7: Histograms and density curve comparison between two different groups

From the figure 7. above, we observed that histograms of each category are unimodal and symmetric. The shapes of both histograms to the left of the peak looks roughly like the miser image of the part of that histogram to the right of the peak. The bell shape of symmetric graphs of both groups indicate a great possibility of normal distribution. However, the density of the group of babies born to non-smoking mothers skewed to the right compared to the distribution of smoking mothers group, as in babies born by non-smoking mothers have relatively higher gestational age. Also, the non-smoking group has data with larger domain in gestational age and less dense in the range from 260 days to 300 days, possibly due to the fact that it has a larger sample size. Since there are extreme outliers (gestational age = 999) in both distributions, we removed them in the graph in order to improve the clarity of histograms.

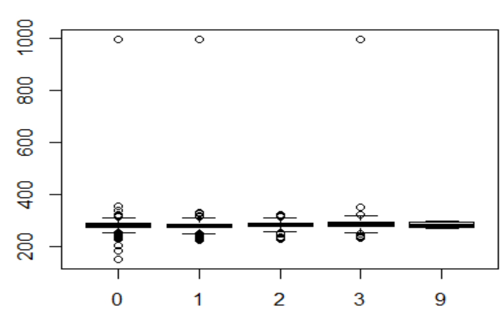


Figure 8: Boxplot for two distribution with outliers

Similarly, the boxplot shown below (figure8) is not clear when we have extreme outliers, and thus we construct a new graph which does not contain those outliers.

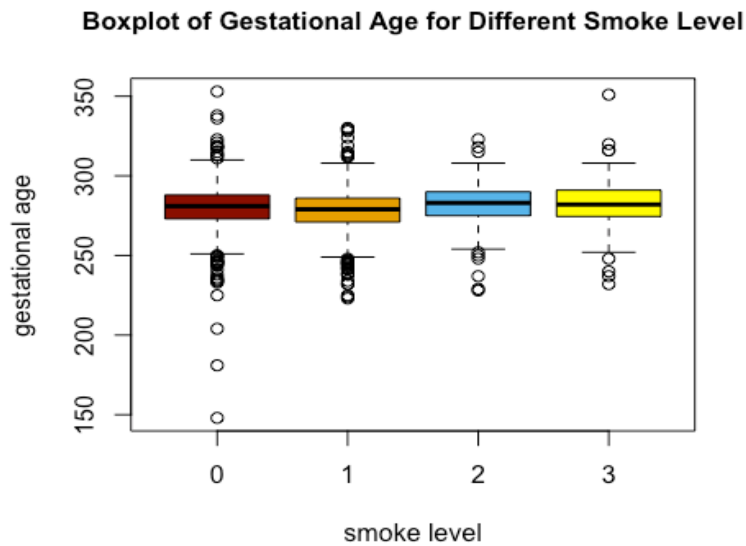
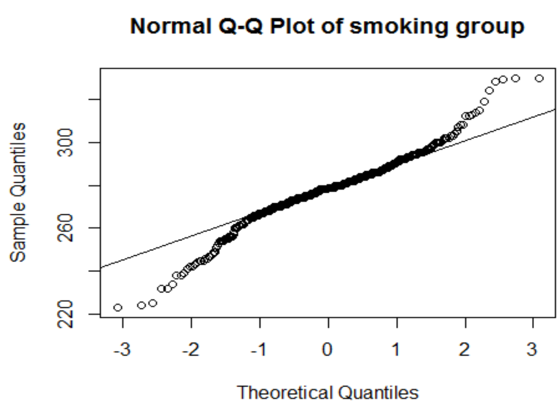


Figure 9: Boxplot for two distribution without outliers

As we can see from the above boxplot (figure 9), group 1, which represents moms currently smoking, have babies with the lowest median of gestational age than other groups who do not smoke during pregnancy. Although there are no major differences in upper and lower bounds between these four categories of mothers, 1st quantile and 3rd quantile of group 1 are the lowest among all groups.

To further discuss the data distribution of smoking mothers and non-smoking mothers, we plot three quantile-quantile plots. According to the histograms we have above, gestational age of both groups is both very likely to be normal distribution, and two normal q-q plots below will help us to confirm.



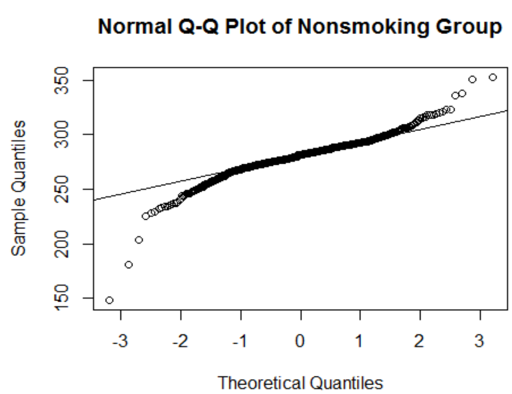


Figure 10: Quantile-Quantile Plot for Smoking mothers Figure 11: Quantile-Quantile Plot for Non-smoking mothers

In the two figures above (figure 10 and figure 11), the plotted points point roughly on the line while a few points lie above or under the normal line. Therefore, it is indicated that data for both groups have approximate normal distributions.

In order to compare data of these two groups, we plot a new Q-Q plot with smoking group on the x-axis and non-smoking group on the y-axis. As the points lie much roughly on the Q-Q line, we can say that samples from smoking groups and non-smoking groups are extremely similarly. Since there still exist some points departure from the straight line, a small difference between the two distributions is indicated. Also, we know that data of these two groups have different means or standard deviation since the plot is roughly linear, but slope and intercept is not 1 and 0.

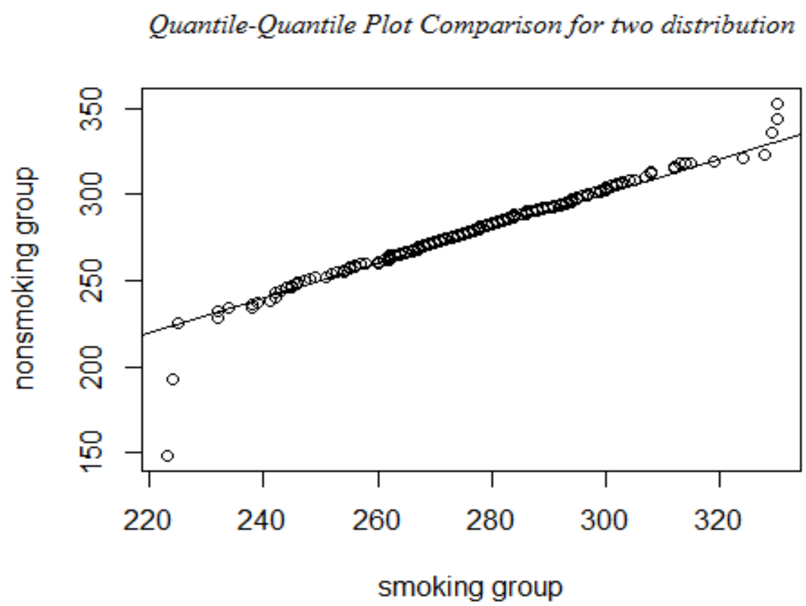


Figure 12: Quantile-Quantile Plot Comparison for the two distributions

Frequency and incidence

	Non-smoking mother	Smoking mother
Proportion (Threshold 35 weeks)	0.0260989	0.0293501
Proportion (Threshold 36 weeks)	0.0467033	0.0524109
Proportion (Threshold 37 weeks)	0.07692308	0.08595388
Proportion (Threshold 38 weeks)	0.1181319	0.1383648
Proportion (Threshold 39 weeks)	0.2307692	0.2955975

Table 5: percentage of small gestational age babies in each group

The sample size of smoking group without outliers is 477 people and of nonsmoking group without outliers is 728 people. In order to estimate whether our estimation is effective, we adjust the standard of small gestational age babies (less than 37 weeks originally) by increasing and decreasing the weeks of gestational age which considered as classification of preterm period. Then, we divide the number of smoke/nonsmoker mothers in each classification by the total number of its group to get the percentage of small gestational age, shown in Table 5. From the table above, it is obvious that the proportion of smoking mothers is greater than non-smoking mothers in every classification of small gestational age.

Importance of difference:

In order to assess the importance of the difference in gestational age we found between smoking and non-smoking groups, a two-sided year can be performed. We set the null hypothesis as there's no influence on babies gestational age from maternal smoking and the alternative hypothesis is that the smoking of mothers actually affects gestational age of babies. That is, the percentage of small gestational age of babies of smoking mothers is greater than those of non-smoking mothers. Set up the Null Hypothesis H_0: μ_1-μ_2 =0 and the Alternative Hypothesis H_1: μ_1-μ_2 >0. Assume that our significant level is $\alpha = 0.05$. We know that the formula of t test is t-test = $\frac{(x1-x2)}{\sqrt{\frac{s1^2}{n1}+\frac{s2^2}{n2}}}$. By calculating the mean, standard deviation, and sample size of two groups, we get the t = 2.348905. Since t > 1.64, we reject the null hypothesis and in favor of H_1. Thus, we can conclude that we are 95% confident babies gestational age of smoking mothers is shorter than those of non-smoking mothers.

According to NVSR, "Similarly, the infant mortality rate for very preterm infants (those born at less than 32 weeks of gestation) was 181.0, more than 72 times the rate for infants born at term (2.5) (37–41 weeks of gestation) (tables 1 and 2)." That is, the gestational age is a factor that affects infant mortality directly. Thus, since our study shows the relationship between smoking and babies' gestational age, we consider our findings as important and necessary.

Theory:

Numerical Summaries:

Through finding the mean, standard deviation, minimum, maximum, and quantiles of babies' weight and gestational age, our group not only gets a straightforward understanding of the data but also gain the numerical summaries of the location and the spread of the data. These numerical data facilitate us to observe the influence upon babies' weight and gestational age by smoking and nonsmoking mothers.

Graphical Summaries:

After the basic analysis of numerical analysis, our group draw histograms which enable us to clearly see the shape of the distribution and then investigate the relationship between smoking and nonsmoking mothers towards babies' weight and gestational age.

Since boxplot provides visualization of comparing a column of multiple categories and shows the relation between two variables, we use it to see how data spread out and compare the distribution of baby's weight and gestational age among different smoking categories.

In order to further test the normality of the data set, our group then utilize the Quantile-Quantile plots to further support our assumption that the dependent variables in our data set are normally distributed. Moreover, we did bar plots to show the frequencies of birth weight and gestational age in each group with different cutoffs, which further demonstrate the relationship between numerical variables and categorical variable in this project.

Central Limit Theorem:

Based on above histograms shown, variables in our case are independent, identically distributed. The central limit theorem properly applies to our analysis. The individual observation does not depend on another observation. Also, since the distribution of the sum of variables looks normal, the rule properly applies to each standardized data.

Skewness Coefficient

Our group utilizes the skewness coefficient to check for normality, which is a way to test the average of the third power of the standardized data

$$\text{Skewness} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{xi-\bar{x}}{s}\right)^3$$

Kurtosis Coefficient

Our group also utilizes the Kurtosis Coefficient to check for normality, which is a way to test the average of the fourth power of the standardized data

$$\text{Kurtosis} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{xi-\bar{x}}{s}\right)^4$$

Conclusion

Based on the above investigation and specific analysis, our group insists that there is a significant difference between smoking mothers and nonsmoking mothers upon fetal growth. By excluding the influence of several confounders, owing to the numerical summaries and graphical summaries, we get the result that smoking status exerts a negative effect on babies' weight and gestational age. According to the definition in lecture note, birth weight and gestational age are the two determinants of baby's maturity[9]. In light of this, from the result we got from numerical summaries and graphical summaries, we can see that compared with non-smoking mothers, smoking mothers have a higher risk of born babies early and small, which is troublesome for the baby's maturity and then have a greater chance of lower survival rate.

[9] Lecture note page 3

Reference

Lecture note, Chapter 2 Maternal smoking and infant death
https://math189.edublogs.org/files/2019/01/chp2-2gfykid-2fdsk4v.pdf

Scott, R. (2011). *Pre-Pregnancy Body Mass Index and Related Maternal Health and Infant Outcomes Among Mothers in Florida, 2009-2011*. [online] Floridahealth.gov.

Available at: http://www.floridahealth.gov/statistics-and-data/survey-data/pregnancy-risk-assessment-monitoring-system/reports/_documents/bmireport.pdf [Accessed 3 Feb. 2019].

R Bakker, EAP Steegers, AA Biharie, JP Mackenbach, A Hofman and VWV Jaddoe (2011). *Explaining differences in birth outcomes in relation to maternal age: the Generation R Study*. [online] Obstetrics & Gynecology. Available at: https://obgyn.onlinelibrary.wiley.com/doi/full/10.1111/j.1471-0528.2010.02823.x [Accessed 3 Feb. 2019].

T. J. Mathews, M. S. Fay Menacker, Dr. P. H, Marian F. Macdorman and Ph. D (2003). *Infant Mortality Statistics from the 2001 Period Linked Birth/Infant Death Data Set*. [online] Citeseerx.ist.psu.edu.

Available at: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.207.3631 [Accessed 3 Feb. 2019].

Hirve , Siddhi S., and Bela R. Ganatra. *DETERMINANTS OF LOW BIRTH WEIGHT: A COMMUNITY BASED STUDY*. www.indianpediatrics.net/oct1994/1221.pdf.