Bootstrapping:

When the two-sample t-test is performed, it is assumed that the distributions are normally distributed. However, according to the result of chi-square test of goodness of fit above (shown in *Table 3*), the distribution of birth weights of babies born to mothers who avoid smoking during pregnancy is not a normal distribution. Thus, bootstrapping must be applied here to get a large enough sample. Then, according to the Central Limit Theorem, we can conclude that distributions of samples with sufficient sample sizes are normally distributed.

After bootstrapping, normality is again checked for both distributions by looking at the kurtosis values, the histograms, the QQ-plots, and performing the chi-square GOF test with significance 0.001. The numerical values are shown in *Table 5*, while histograms and QQ-plots are each shown in *Figure 6, 7,* and *8*.

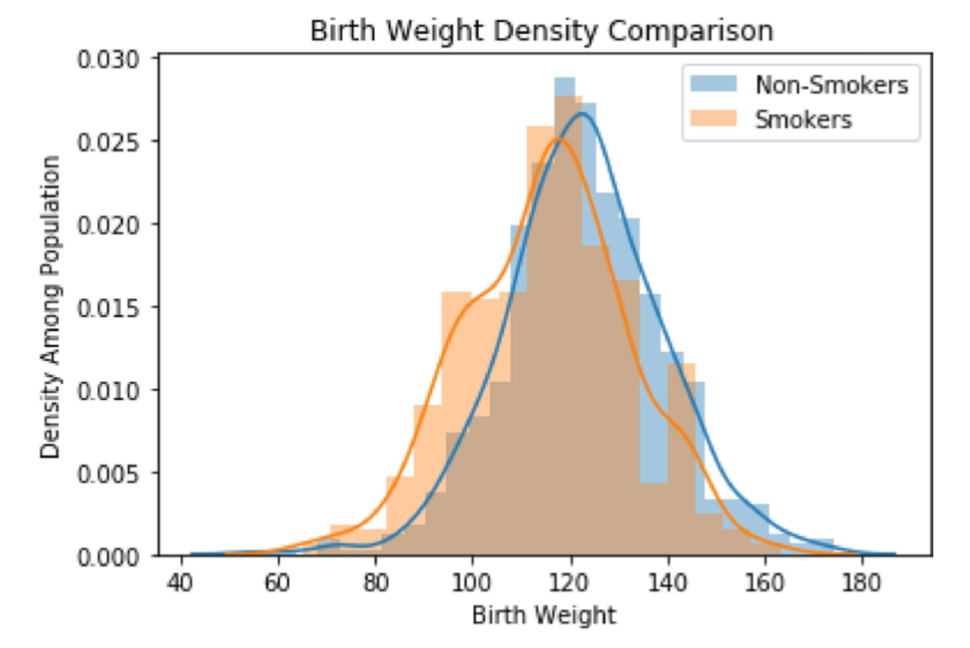
From observing *Table 5* below, both distributions have kurtosis values around 3, which is the kurtosis value for a standard normal distribution. The chi-square GOF results also show that we cannot conclude that the distributions are not normally distributed.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Kurtosis | P-value for Chi-Square GOF Test | Conclusion |
| Distribution of Birth Weights of Babies Born to Non-Smoking Mothers During Pregnancy | 3.765430 | 0.003008 | The null hypothesis cannot be rejected and that we cannot say that this distribution in not normally distributed. |
| Distribution of Birth Weights of Babies Born to Smoking Mothers During Pregnancy | 2.974981 | 0.774119 | The null hypothesis cannot be rejected and that we cannot say that this distribution in not normally distributed. |

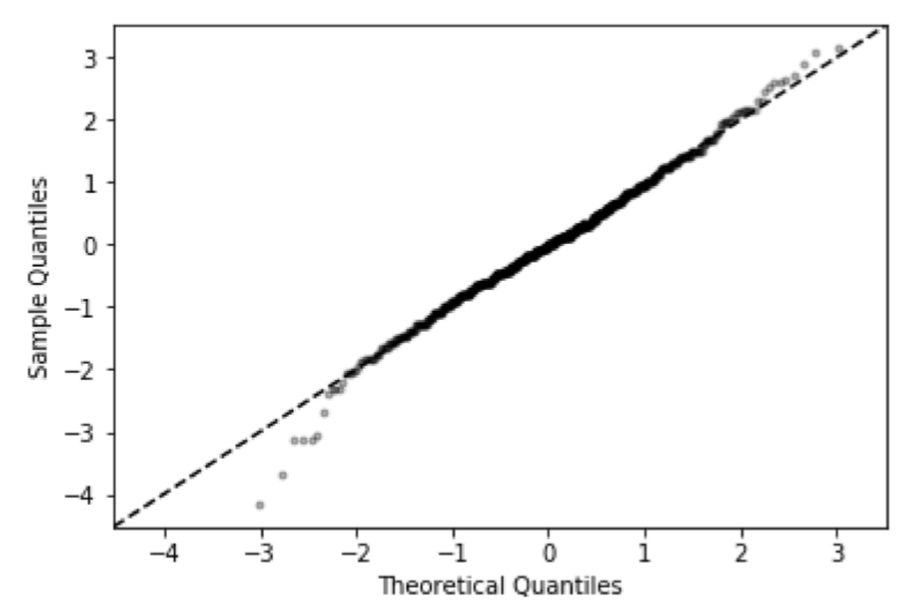
*Table 5: Kurtosis Values and Chi-Square Goodness of Fit Test Results*

*for Both Distributions after Bootstrapping*

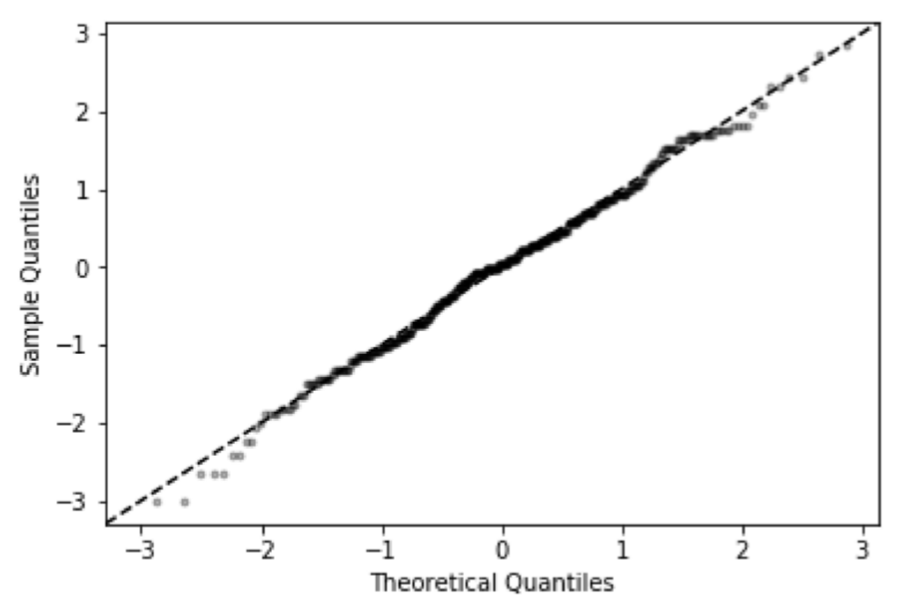
Histograms shown in *Figure 6* below show that both distributions, after bootstrapping, are now roughly bell-shaped like standard normal curves. Moreover, QQ-plots for both distributions after bootstrapping shown in *Figure 7* and *8* below tell that both distributions have lied closer to the theoretical normal distributions shown as dashed lines compare to those before bootstrapping (*Figure 3* and *4*).



*Figure 6: Histograms and Density Curve Comparison between the Babies’ Birth Weight Born to Non-smoking Pregnant Mothers and Born to Pregnant Smoking Mothers after Bootstrapping*



*Figure 7: Quantile-Quantile Plot for Non-smoking distribution after Bootstrapping*



*Figure 8: Quantile-Quantile Plot for Smoking distribution after Bootstrapping*