5.a) $n\int_a^b \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy$

f) Firstly, we compare observed and expected frequencies in part (b). We observe that observed and expected values are not close and so the Normal model is not a suitable model for these data. To be specific, for instance, expected value is 40.0 while observed value is 2 when the interval is $[30, 40)$. Also, expected value is 7.16 while observed value is 5 when the interval is $[40, 50)$. Moreover, expected value is 5.52 while observed value is 10 when the interval is $[50, 60)$. The limitation of the method is that different selected intervals would cause difference values, which means that if we change the interval from 10 to 5, it is possible the observed and expected values would change.

Secondly, by observing the Histogram of gas in part (c), we find that there exists an obvious systematic departure between the

histogram and the normal density line on it. To be specific, there is a long left tail of the histogram, but the Normal model is symmetric. Thus, the Normal model is not a suitable model for these data. The limitation of this method is that the intervals for the histogram must be chosen because the data are grouped into intervals to form the histogram. Different intervals might cause different shapes of the histogram. Also, it requires us to estimate μ and σ.

Thirdly, by part (d), we find that there is good agreement between the curve and points and so Normal model seems suitable for these data. The limitation of ecdf plot is that it does not show the shape of distribution clearly. It also need us to estimate mean and standard deviation. The merit of it is that it shows the exact height in the sample.

Fourthly, by part (e), we find that the set of points lie reasonably along the straight line. Although the points at both ends of the line lies further from the line, it is still reasonable since the quantiles of the

Normal distribution change in value more rapidly in the tails. Thus, it suggests that the Normal model is good. The merit of the Qqplots is that μ and σ do not need to be estimated. However, it does not show the shape of distribution.

From the above four comparisons, we conclude that the Normal model is not suitable for these data.