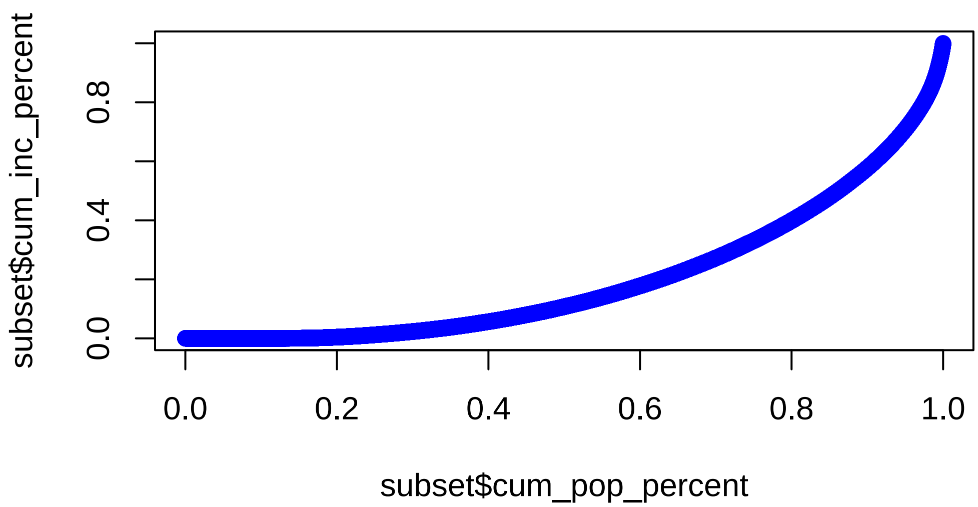
Lorenz Curve and Gini

Rowan, Mamoru, Ari, Gwen

To calculate the Gini index we first needed to estimate the coefficients of the Lorenz curve. To do this we downloaded a data set from Ipums USA which contained individual, household level, census data about the total yearly income of each household. We cleaned and ordered the data, and then calculated a cumulative population percentage, and a cumulative income percentage variable. When we graph them against each other we can see that classically shaped Lorenz curve emerge:



Next we needed to estimate the coefficients of the model. As we know, the Gini coefficient is given by:

$$G = 1 - 2\int_0^1 l(x)dx$$

Where:

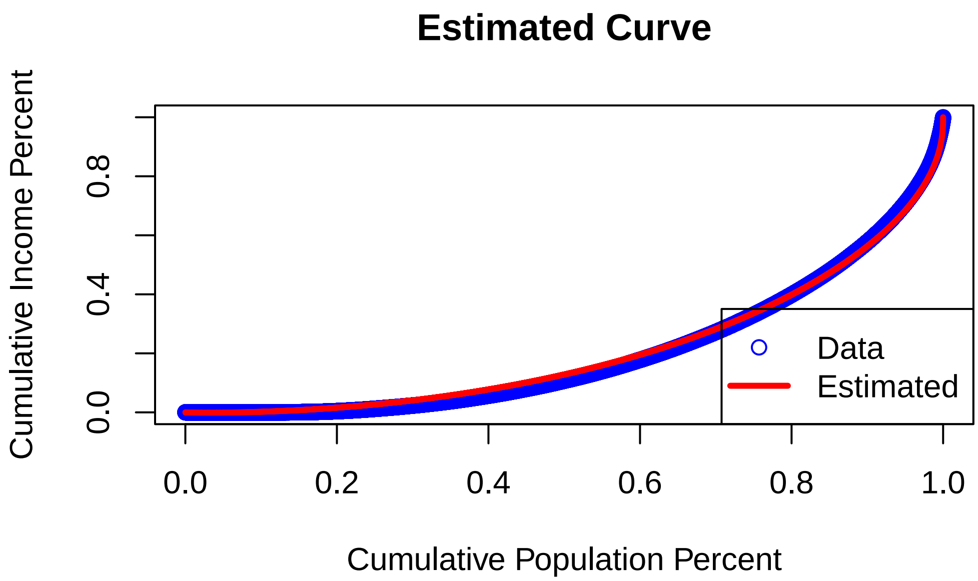
$$l(x) = x - (ax^a(1-x)^b)$$

Therefore we need to use a nonlinear least squares estimation method to estimate the coefficients for a and b. Here you can see the code we ran, and the results for the estimated coefficients for the model. Part of the equation is cut off the edge of the screen due to the pdf rendering method we are using, I couldn't figure it out in time.

```
lorenz_model<- nls(cum_inc_percent ~ cum_pop_percent - (a * cum_pop_percent^a * (1 -
                                                                            cum_pop_per
                        data  = subset,
                        start = list(a = 0.5, beta = 0.5),  # Starting values for paramete
                        lower = c(a = 0, beta = 0),  # Lower bounds to ensure positivity
                        upper = c(a = 10, beta = 10), # Upper bounds
                        algorithm = "port")  # Port algorithm for bounded fitting

subset$fitted_values_new <- predict(lorenz_model)

print(lorenz_model)
```

```
Nonlinear regression model
  model: cum_inc_percent ~ cum_pop_percent - (a * cum_pop_percent^a *     (1 - cum_pop_percen
   data: subset
     a   beta
0.9816 0.4195
 residual sum-of-squares: 6.69

Algorithm "port", convergence message: relative convergence (4)
```

After estimating the coefficients, we can graph the estimated curve against the actual observed data to check if it visually appears correct.



Finally we can use R to calculate the Gini index by taking the integral of the model, and plugging it into the equation for the Gini index.

```
# Define the fitted Lorenz curve function based on the estimated parameters a and beta
lorenz_function <- function(x, a, beta) {
  x - a * x^a * (1 - x)^beta
}

# Get the estimated parameters from your model
a_estimate <- coef(lorenz_model)["a"]
beta_estimate <- coef(lorenz_model)["beta"]

# Define the function to integrate
gini_integrand <- function(x) {
  x - lorenz_function(x, a_estimate, beta_estimate)
}

# Numerically integrate the function from 0 to 1
gini_result <- integrate(gini_integrand, lower = 0, upper = 1)

# Calculate the Gini coefficient
Gini_coefficient <- 1 - 2 * gini_result$value
```

```
# Print the Gini coefficient
print(Gini_coefficient)
```

```
[1] 0.4214416
```

But this is a math class so we are going to show all the math for calculating the integral here:

Derivation and Computation of the Gini Coefficient

We are given a Lorenz curve model defined as:

$$l(x) = x - ax^{a}(1-x)^{\beta},$$

where

$$(\alpha = 0.978) and (\beta = 0.413)$$

(a = = 0.978) and (= 0.413).

The Gini coefficient (G) is defined as:

$$G = 1 - 2\int_0^1 [x - l(x)]\, dx.$$

First, we simplify the integrand:

$$x - l(x) = x - [x - ax^{a}(1-x)^{\beta}] = ax^{a}(1-x)^{\beta}.$$

Thus, the integral we need to evaluate is:

$$\int_0^1 [x - l(x)]\, dx = \int_0^1 ax^{a}(1-x)^{\beta}\, dx.$$

Factor out (a):

$$= a\int_0^1 x^{a}(1-x)^{\beta}\, dx.$$

Recognize that this integral is related to the Beta function. The Beta function (B(p,q)) is defined as:

$$B(p,q) = \int_0^1 t^{p-1}(1-t)^{q-1}\,dt.$$

Let (p = a + 1) and (q = + 1). Then:

$$\int_0^1 x^a(1-x)^\beta\,dx = B(a+1,\beta+1).$$

Using the relationship between the Beta and Gamma functions:

$$B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)},$$

we have:

$$B(a+1,\beta+1) = \frac{\Gamma(a+1)\Gamma(\beta+1)}{\Gamma(a+\beta+2)}.$$

Substitute this back into our integral:

$$\int_0^1 [x - l(x)]\,dx = a\frac{\Gamma(a+1)\Gamma(\beta+1)}{\Gamma(a+\beta+2)}.$$

With (a = 0.978) and (= 0.413), evaluating the Gamma functions numerically (using a computational tool) gives approximately:

$$\int_0^1 [x - l(x)]\,dx \approx 0.29.$$

Finally, we compute the Gini coefficient:

$$G = 1 - 2\int_0^1 [x - l(x)]\,dx = 1 - 2(0.29) = 1 - 0.58 = 0.42.$$

Conclusion:

The estimated Gini coefficient is approximately:

$$G = 0.42.$$