Propagating Monte Carlo Error

Team A7: Dylan Arrabito, Nicole Holden, Daniel Monteagudo, Giovana Puccini

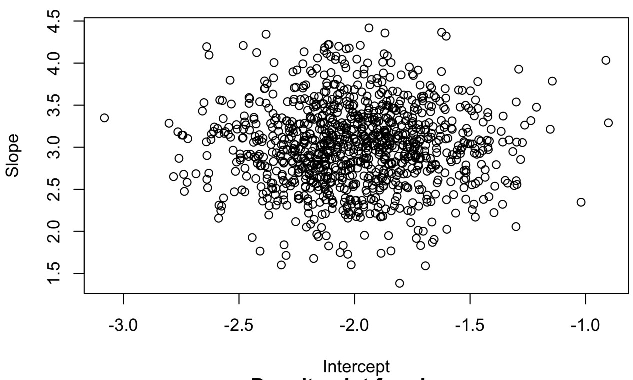
12/06/2018

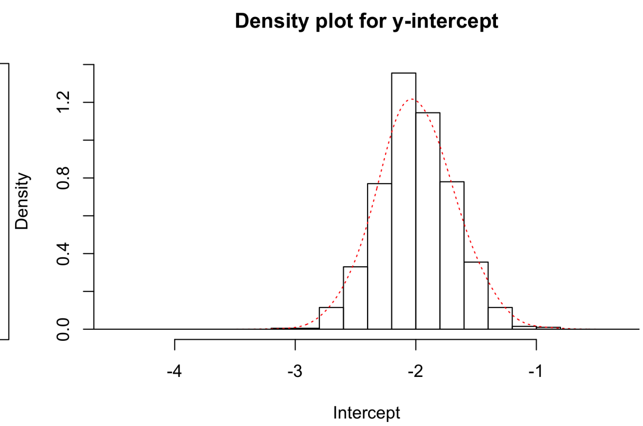
The functions we are using to generate the fake data are:

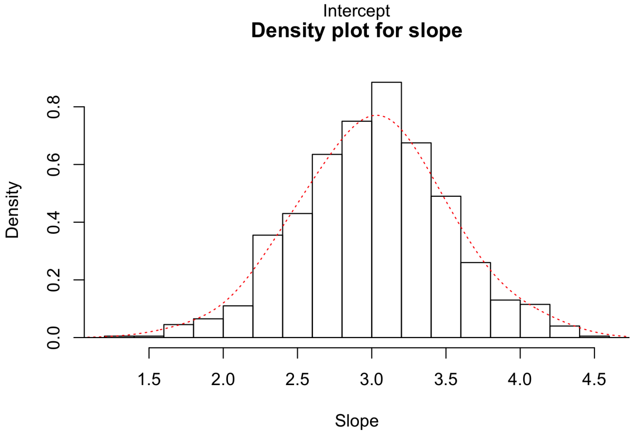
1. $f_1(x) = -2 + 3x$
2. $f_2(x) = 3$
3. $f_3(x) = 6x^2 + 3x + 3$
4. $f_4(x) = 10x + 3$
5. $f_5(x) = -4x - 6$

Generation and plotting of the 1000 $\hat{c}_0$ and $\hat{c}_1$ values from fake data

Function number 1







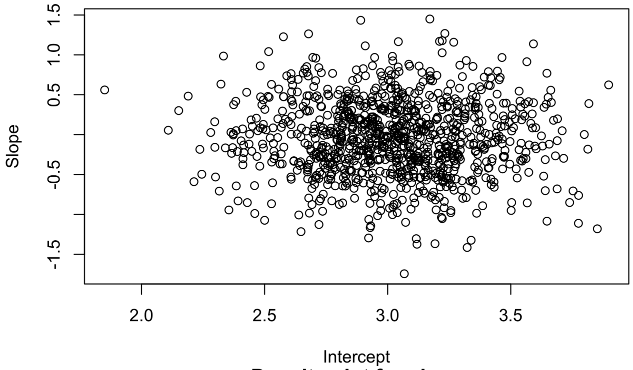
Mean of the intercept: -1.9966

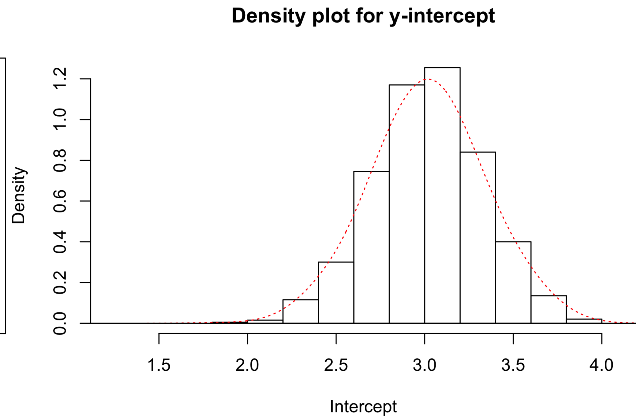
Variance of the intercept: 0.0946

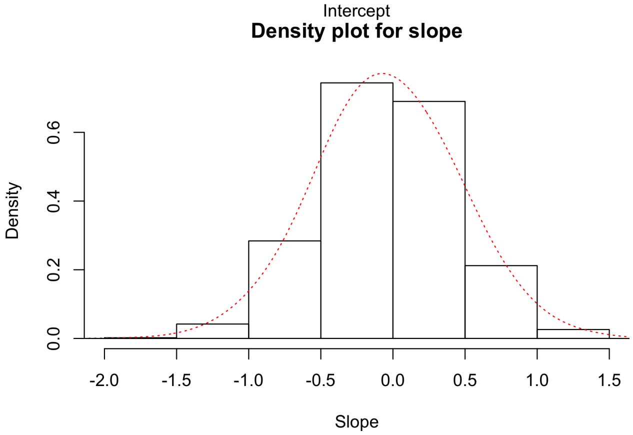
Mean of the slope: 3.0098

Variance of the slope: 0.24616

Function number 2







Mean of the intercept: 3.0159

Variance of the intercept: 0.095975

Mean of the slope: -0.044604

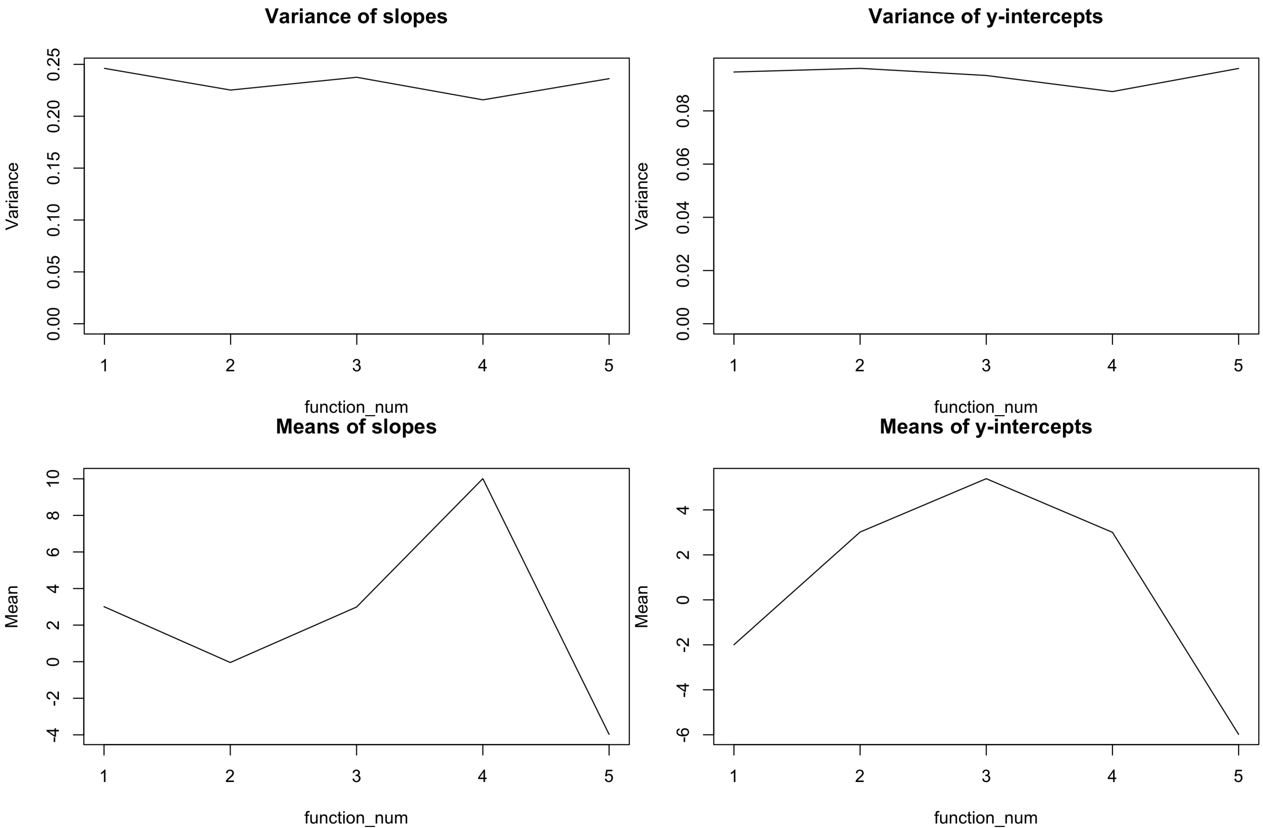
Variance of the slope: 0.22523

Conclusion:

The bell shaped histogram indicated that the values of $\hat{c}_0$ and $\hat{c}_1$ are normally distributed.

Calculation of the variance and mean value of $\hat{c}_0$ and $\hat{c}_1$ for every function.

We plotted the means and variances of $\hat{c}_0$ and $\hat{c}_1$ for five different functions to see if the choice of $f(x)$ affected them.

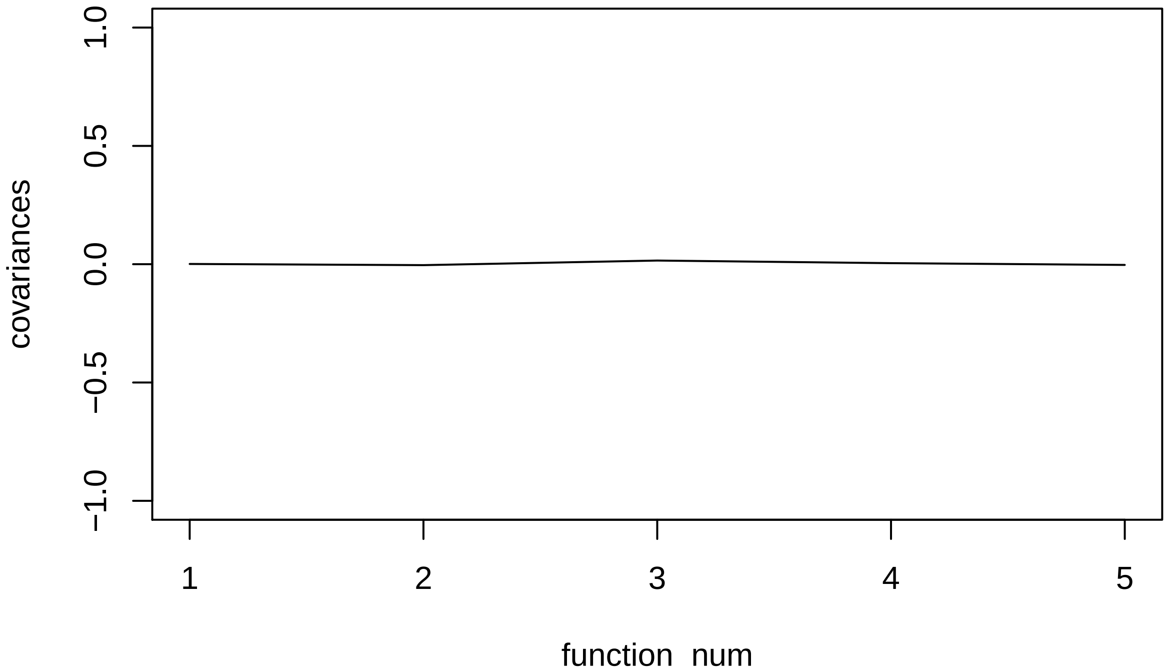


Conclusion:

Because the variance is roughly a straight line between all 5 functions, we can see that it does not depend on the choice of $f(x)$. Conversely, looking at the means, we can see that they vary wildly depending on your choice of function.

Covariance Calculation

First we want to plot the covariances for all five different functions to determine whether covariance depends on the choice of $f(x)$.



Conclusion:

The graph shows that the covariance does not depend on the choice of $f(x)$ and we can see that the **covariance for any function is approximately zero** so we can say that the two coefficients $\hat{c}_0$ and $\hat{c}_1$ are uncorrelated for any $f(x)$.

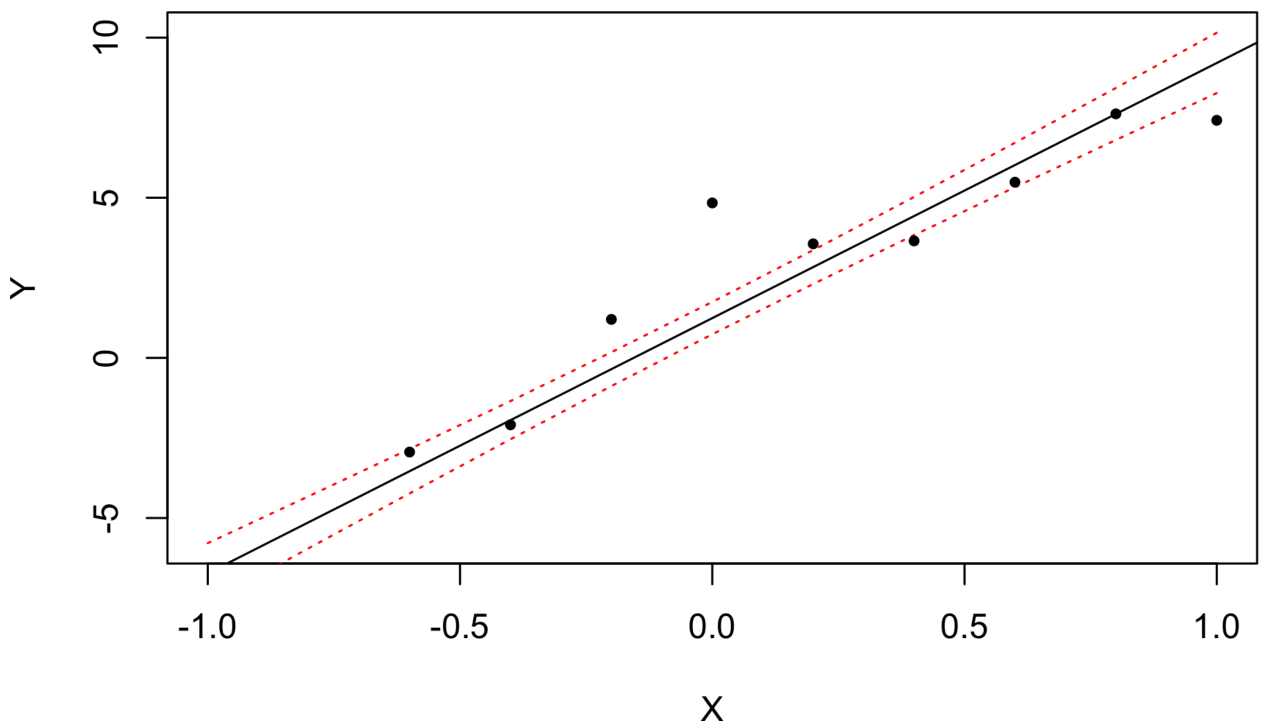
Variance of a linear fit

From the covariance and variance values found, we used the formula for the variance of a linear combination $(V(f(x)) = V(\hat{c}_0) + V(\hat{c}_1)x^2 + 2xCov(\hat{c}_0, \hat{c}_1))$ to find an expression for the variance of our linear model.

$V(f(x)) = 0.093383 + 0.23208x^2 + 0$

Plotting a confidence interval for a linear fit

From the previous variance formula, we calculated the 90% confidence interval for the linear model. We plotted the interval together with the best fit linear model and the original data.



Solving the integral $\hat{I} = \int(\hat{c}_0 + \hat{c}_1 x)dx$ for $\hat{I}$, we found that $\hat{I} = 2\hat{c}_0$

Therefore $V(\hat{I}) = 4V(\hat{c}_0)$

CI $= \hat{I} \pm z_{0.05}V(\hat{I})$

CI $= 2\hat{c}_0 \pm 1.0023$

Final confidence interval: $[1.4812, 3.4859]$