**An Examination of the Relationship Between Various Human Bone Lengths and Gender**

**Abstract:**

On March 8, 2018, CNN news reported that Richard Jantz, professor emeritus of anthropology at the University of Tennessee, attempted to use bone length data collected in 1940 to determine if bones found on the Pacific Island of Nikumaroro could be those of Amelia Earhart.

In 1940, Dr. D. W. Hoodless studied and measured these bones first hand and concluded that they belonged to a human male. The bones were lost shortly afterwards.

Richard Jantz used Dr. Hoodless’s bone measurements from 1940 to perform a modern forensic analysis using information and technology not available at that time. In his opinion, what he found “strongly supports the conclusion that the Nikumaroro bones belonged to Amelia Earhart”. However, other reporting indicates that this conclusion is suspect due to other circumstantial evidence surrounding the case.

Nonetheless, the idea of determining gender from bone measurements is an interesting one. This research project will look at raw data from various human bones and attempt to create models to accurately determine gender from bone length. This project will examine the mean and standard deviation between the various bones of each gender. It will analyze the correlations between height, humerus length, and femur length. Several logistic regressions will be run to model the relationship between bones and gender.

The results indicate that it is possible to predict gender using bone length. Predictions can be made using individual logistic regressions or multi-variable logistic regressions. Predictions when bone length height proportion data is used, but the predictions are less accurate.

**Introduction:**

Male and female bone length and height raw data was obtained from the Forensic Anthropology Data Bank for the following human bones: humerus, radius, ulna, femur, tibia, and fibula. Using this data, we will look for relationships between different bones lengths and gender. The goal is to see if we can find a reliable model or reliable models to determine gender from bone length.

The main goals of this project are to determine:

* Which single bone is the best predictor of gender?
* How can we best tell the difference between male and female bones?
* Does looking at multiple bones give more predictive power than looking at just one bone?
* Are men and women different in bone proportions after accounting for height?

**Methods, Models, and Calculations:**

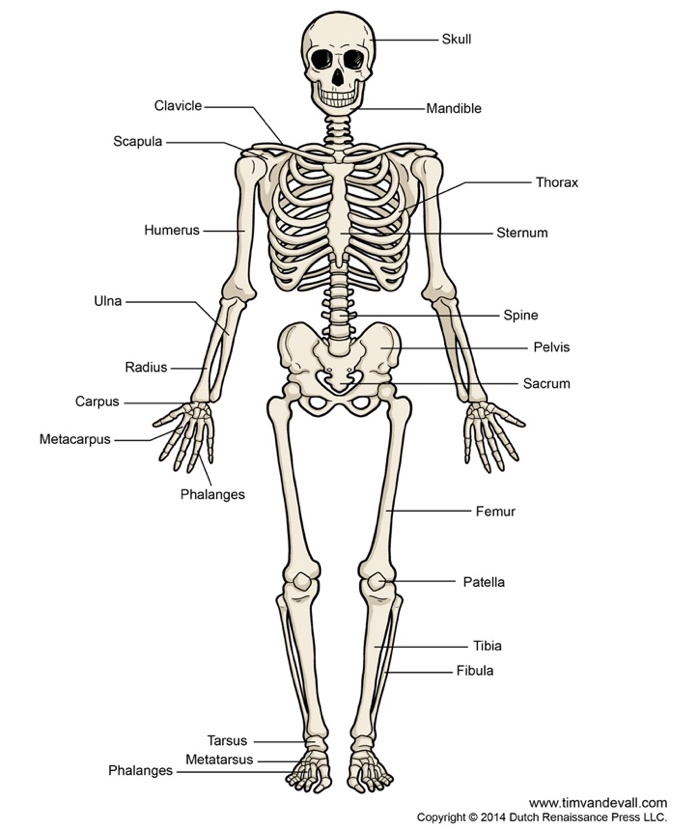
All calculations and graphing will be done in Excel except for the summary output data for logistic regressions which will be done using XLMiner in Google Sheets.

We will look at the correlation(s) between the largest bones (humerus and femur) and height.

A logistic regression will be done for each bone. A multi-variable logistic regression will be done for all bones. A multi-variable logistic regression using bone/height proportions will be done for all bones. A bone/height proportion logistic regression will be done for the humerus to see how it compares to the raw data (non-proportion) humerus logistic regression.

An example graph, calculated probabilities, and summary data will be shown for the humerus logistic regression. Only summary data will be shown for the remaining regressions (the graphs and summaries for all regressions can be found in the attached Excel workbook).

**Skeletal Diagram for Reference:**

****

**Mean Heights and Mean Bone Lengths:**

We will first compare the mean and overall height standard deviation for all of the bones for both males and females.





**Correlation of Height, Humerus, and Femur:**







In the above graphs, looking at R^2, we see moderate correlations between height, femur length, and humerus length. We also see that male bones tend to be longer than female bones.

**Graphed Logistic Regression, Probabilities, and Summary Data for the Humerus Bone:**



****

****

**Results:**

**Logistic Regression of the Humerus:**

****

**Logistic Regression of the Radius:**

****

**Logistic Regression of the Ulna:**

****

**Logistic Regression of the Femur:**

****

**Logistic Regression of the Tibia:**

****

**Logistic Regression of the Fibula:**

****

**Multi-Variable Logistic Regression for All Bones:**

****

**Multi-Variable Logistic Regression for All Bones Using Bone/Height Proportions:**

****

**Logistic Regression of the Humerus Using Bone/Height Proportions:**

****

**Conclusions:**

**Which single bone is the best predictor of gender?**

When looking at the summary data for single bone logistic regressions we can see that the humerus has the highest Chi-Squared result of 54.9093259. Therefore, the most useful bones for gender prediction when looking at logistic regressions for each individual bone are (from best to worst predictor):



**How can we best tell the difference between male and female bones?**

The multi-variable linear regression for all bones summary data we can see that only the humerus, tibia, and fibula are statistically significant. This can be determined by looking at the p-values for the bones against alpha = 0.05. The femur, radius, and ulna are not statistically significant. Therefore, the best way to tell the difference between a set of male and female bones would be to examine the humerus, tibia, and fibula together.



**Does looking at multiple bones give more predictive power than looking at just one bone?**

Comparing the chi-squared values for the individual logistic regression equations with the multi-variable linear regression, we can see that the individual equations summarize to:



And the multi-variable linear regression summarizes to:



Thus, the multi-variable linear regression using multiple bones is a better predictor of gender.

**Are men and women different in bone proportions after accounting for height?**

Examining the results of the multi-variable linear regression for all bones using bone/height proportions we have a chi-squared result of:



If we compare this to the multi-variable linear regression using multiple bones (no proportions) chi-squared value of:



This indicates that we will get better results using the multi-variable linear regression and that it will be more difficult to predict gender with the bone/height proportion summary data.

Examining logistic regression of the humerus using bone/height proportions we have a chi-squared of:

****

If we compare this to the logistic regression of the humerus bone (no proportions) chi-squared value of:



This indicates that we will get better results using the non-proportion humerus bone logistic regression and that it will be more difficult to predict gender with the bone/height proportion humerus logistic regression summary data.

Men and women are still different in bone proportions after accounting for height. Predictions are more accurate when we do not use bone/height proportion data for the model.

**Final Thoughts:**

Cleary, there is some overlap in size between the bones of men and women. Looking at the mean and standard deviation for all bones and looking at the logistic regression summary data and probabilities, we do have some ability to predict gender from bone length, but there will be some bone lengths that will leave gender ambiguous such as:





**References:**

Data from Forensic Anthropology Data Bank (FDB)

<http://www.utdanacenter.org/mathtoolkit/mmacd/curric/pr/pdf/pr-tg.pdf>

<https://edition.cnn.com/2018/03/08/health/amelia-earhart-bones-island-intl/index.html>

<https://www.timvandevall.com/templates/human-skeleton-diagram-printable/>

*Introduction to Statistics and Data Analysis, 4th* Edition - by Rosy Peck (Author),‎ Chris Olsen (Author),‎ Jay L. Devore (Author)

*What is a p-value anyway? 34 Stories to Help You Actually Understand Statistics 1st Edition* - by Andrew J. Vickers (Author) (used for information on logistic regression)