The Golden Ratio and Facial Beauty. A statistical investigation

Diego Ramos Crespo

December 1, 2024

Introduction

The golden ratio, often referred as the golden number, divine proportion or golden proportion, approximately 1.618, is an irrational number usually represented by the Greek letter φ (phi) which has been regarded as a representation of beauty, balance, order and proportions that are naturally pleasing to the human eye (Carlson, 2024).

In a recent study an anthropologist registered 25 female faces from a tribe in Africa, a community with beauty standards that differ from those of Western cultures. The researcher measured the length (w) and width (h) of the faces, calculating the length-to-width ratios (w/h). The purpose of this analysis is to assess whether the tribe has similar beauty standards as Western societies or their preferences differ. In order to achieve this goal, there will be used statistical tools such as confidence intervals by two mainstreams: a classical approach and a computer-based approach which is the *Bootstrap* method, along with some visualization techniques.

1. Classical approach

The data for this analysis consists of 25 length-to-width ratios (w/h), as depicted in **Figure 1**, representing the facial measurements of facial geometry of female individuals from an African tribe.

Figure 1

$$[1.443,\ 1.335,\ 1.529,\ 1.617,\ 1.493,\ 1.511,\ 1.488,\ 1.626,\ 1.650,\ 1.619,\ 1.449,\ 1.592,\ 1.497,\ 1.618,\ 1.637,\ 1.650,\ 1.642,\ 1.664,\ 1.808,\ 1.754,\ 1.185,\ 1.736,\ 1.072,\ 1.615,\ 1.620]$$

Note: Twenty-five length-to-width ratios of female faces in an African tribe.

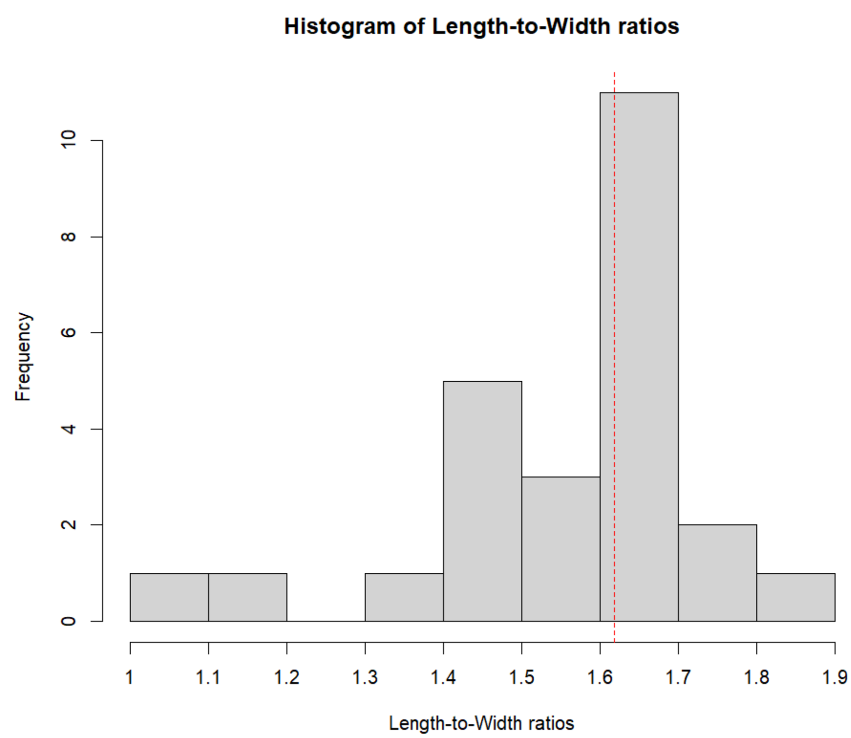
For constructing a confidence interval (CI) through the classical approach it is important to assume that the population follows a normal distribution. Besides the observations should be independent from one another and they must be randomly sampled from the population of interest to ensure avoiding bias in the data.

As said previously, the first step is to take a view of the distribution of the ratios by generating a histogram, which is shown in **Figure 2**.

The distribution is somewhat skewed. With this small sample size it might seem an unknown distribution, however, there is a remarkable peak around 1.6 and 1.7, which is close to the golden ratio (approximately 1.618) indicated by the red vertical dashed line.

Nevertheless, there is a noticeable variability in the data with ratios lower than 1.5 and some others approaching 1.8, this is why the sample mean equals to **1.554** which is far from the value of the golden ratio.

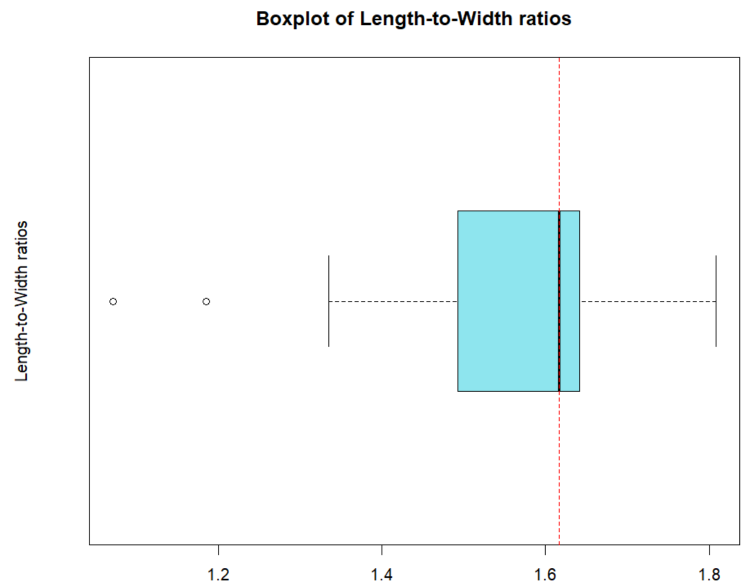
Figure 2



Note: Histogram of length-to-width ratios from the sample where the vertical red dashed line indicates the value of the golden ratio.

In order to verify the skewness of the distribution, it is helpful to observe the boxplot, which indeed also depicts some outliers. In **Figure 3**, the dashed line shows that the median lies very close to the golden ratio; in this case, the median of the sample is 1.617; this suggests that the median might be a more robust indicator than the mean as it is less sensitive to outliers.

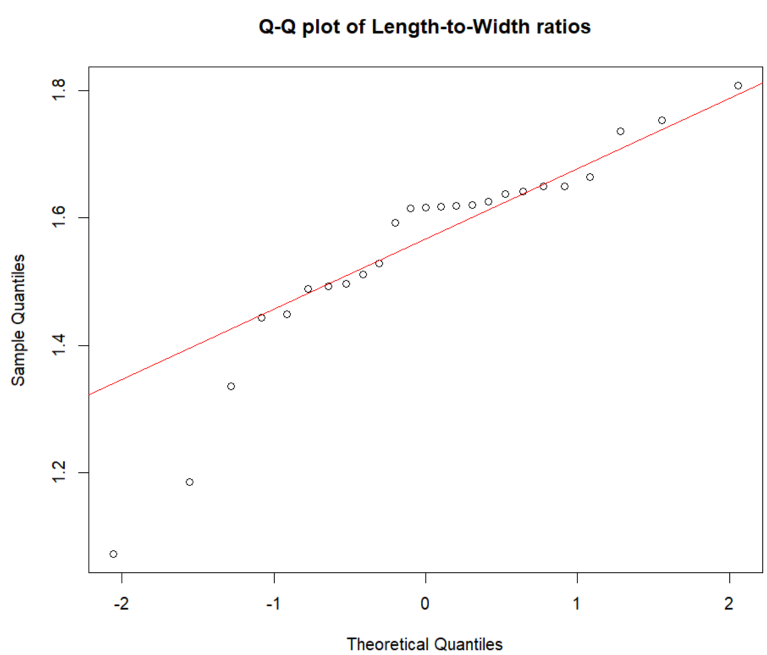
Figure 3



Note: Boxplot of the length-to-width ratio where the vertical red dashed line represents the value of the golden ratio.

The Q-Q Plot illustrated in **Figure 4** reaffirms the fact that the sampling distribution differs from a normal distribution. In an ideal case, where the data is normally distributed, the points would align closely onto the red line that represent the theoretical quantiles. However, in this case some deviations from the line are observed, mainly at the tails of the distribution, indicating skewness or the presence of outliers.

Figure 4



Note: Q-Q plot of length-to-width ratios.

With this data set, the construction of a confidence interval using the classical approach may not be the most appropriate method because it assumes that the sample is normally distributed. Nonetheless, the previous analysis demonstrates that this sample distribution differs from a normal distribution, as a result, the confidence interval might not accurately represent the population's true mean of the length-to-width ratios.

For the purpose of the CI computation, the confidence level will be 96%. Besides, as the true variance of the population is unknown, there will be considered the t-distribution with 24 degrees of freedom to calculate the critical value.

The confidence interval is calculated below:

Statistical concepts	**Formula**	**Values**
$Sample\ size$	n	25
$Sample\ mean$	$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$	1.554
$Sample\ variance$	$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{X})^2$	0.0276
$Standard\ deviation$	$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{X})^2}$	0.1663
$Significance\ level$	$\propto$	0.04
$Critical\ value$	$t_{\frac{\alpha}{2},(n-1)}$	$t_{\frac{0.04}{2},(24)} = 2.1715$
$Standard\ error\ (SE)$	$\frac{S}{\sqrt{n}}$	$\frac{0.1663}{\sqrt{25}} = 0.0333$
$Margin\ of\ error\ (MOE)$	$t_{\frac{\alpha}{2},(n-1)} \times \frac{S}{\sqrt{n}}$	0.0722
$Lower\ limit\ of\ the\ inteval$	$\bar{X} - t_{\frac{\alpha}{2},(n-1)}\left(\frac{S}{\sqrt{n}}\right)$	1.4818
$Upper\ limit\ of\ the\ interval$	$\bar{X} + t_{\frac{\alpha}{2},(n-1)}\left(\frac{S}{\sqrt{n}}\right)$	1.6262

Therefore, the confidence interval for the population length-to-width ratio mean is $[1.4818, 1.6262]$.

We can be 96% confident that the previous interval contains the true population mean; this suggests that the mean length-to-width ratio for the faces in this African tribe is likely to be close 1.618 (the golden ratio). This supports the idea that the facial proportions in females may align with the golden ratio, a concept often associated with aesthetic harmony and symmetry in Western societies.

2. Computer-based approach: Bootstrapping

When the sample distribution does not fulfill the conditions needed to apply classical statistical techniques, computer-based approaches provide a powerful and flexible alternative, in particular, methods like bootstrap, which will be used in this second part of the analysis.

In the current context, bootstrapping is particularly useful in this because the sample size is relatively small. By resampling the data, a large number of samples gets generated, allowing to have more information for statistical inference.

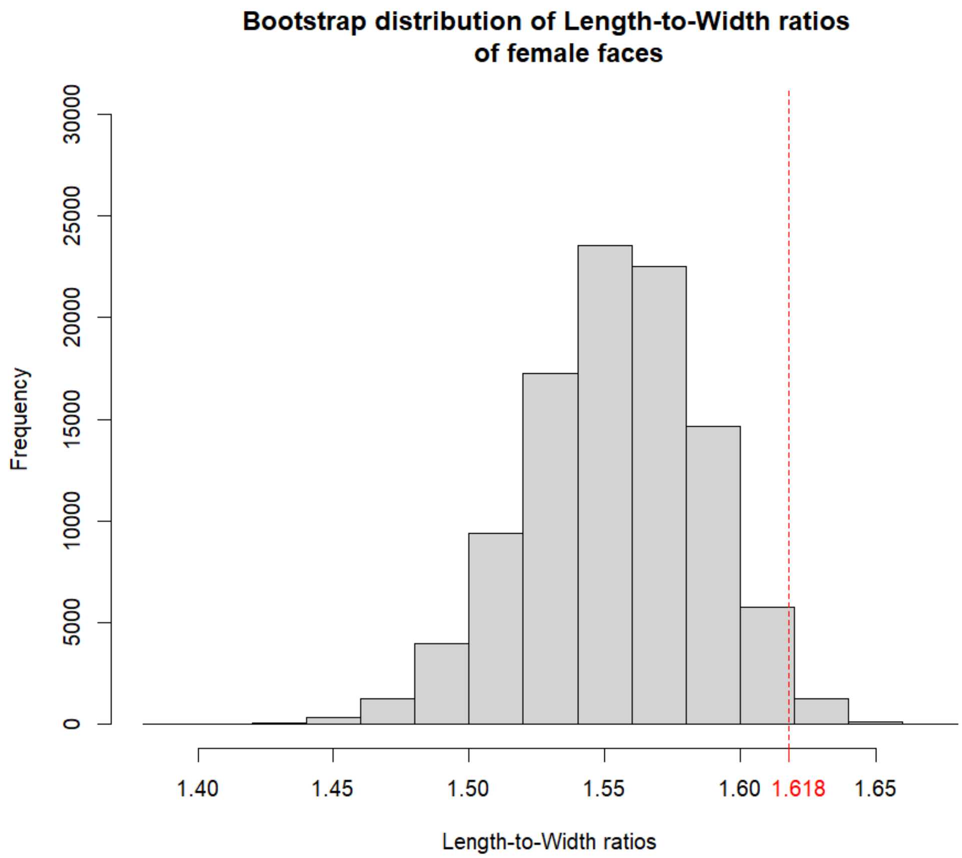
In R, the function used to make the resampling process was *sample()* and the arguments given were:

- The original sample so that the function could pick among all those ratios.
- The size of the new resamples, which were set to be the same as the original sample size (25).
- The *replace = TRUE* argument, so that all the values can be taken again, this means that the possible number of resamples is $25^{25} \approx 8.88178 \times 10^{34}$.

The process was repeated 100,000 times and for each resample, the mean was computed and stored in a vector.

Then, the histogram is plotted in **Figure 5** and now it does seem to resemble a normal distribution, although there is still “a bit” of asymmetry because it looks left-skewed. Furthermore, the majority of ratios appear to be around 1.52 and 1.58, moving away from the golden ratio which is represented with the dashed line in red.

Figure 5



Note: Histogram using Bootstrap method to produce 100,000 resamples of the original sample. The red dashed vertical line shows the value of the golden ratio.

Due to Central Limit Theorem (CLT) we can state that as the number of resampled means increases, the distribution of these means will converge to a normal distribution, regardless of the shape of the original population distribution. For that reason, we can use a normal distribution to calculate the critical values for the CI even though the variance is still unknown because now the sample size is considerably large (10^5).

Moreover, one can observe that both critical values are quite similar. For a normal distribution is 2.053749, while for a t-distribution with the corresponding degrees of freedom, it is approximately 2.053776. The similarity is due to the fact that the sample size is relatively large, making the t-distribution converge closely to the normal distribution.

In the current analysis, we will stick to the quantiles of the normal distribution, as it simplifies computations without compromising accuracy.

The next table shows the CI computations.

Statistical concepts	**Formula**	**Values**
$Sample\ size$	n	100,000
$Sample\ mean$	$\bar{X}_{boot} = \frac{1}{n}\sum_{i=1}^{n} x_i$	1.553873
$Sample\ variance$	$S^2{}_{boot} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{X})^2$	0.001059
$Standard\ deviation$	$S_{boot} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{X})^2}$	0.032541
$Significance\ level$	$\propto$	0.04
$Critical\ value$	$Z_{\frac{\alpha}{2}}$	$\mathrm{t}_{\frac{0.04}{2}} = 2.053749$
$Standard\ error\ (SE_{boot})$	$\frac{S_{boot}}{\sqrt{n}}$	$\frac{0.032541}{\sqrt{100{,}000}} = 0.000103$
$Margin\ of\ error\ (MOE_{Boot})$	$Z_{\frac{\alpha}{2}} \times \frac{S_{boot}}{\sqrt{n}}$	0.000211
$Lower\ limit\ of\ the\ inteval$	$\bar{X}_{boot} - Z_{\frac{\alpha}{2}}\left(\frac{S_{boot}}{\sqrt{n}}\right)$	1.553661
$Upper\ limit\ of\ the\ interval$	$\bar{X}_{boot} + Z_{\frac{\alpha}{2}}\left(\frac{S_{boot}}{\sqrt{n}}\right)$	1.554084

According to the previous calculations, the confidence interval for the population length-to-width ratio mean is $[1.553661, 1.554084]$.

We can say that we are 96% confident that the interval above contains the true mean population length-to-width ratio. As one can notice, the CI does not contain the golden ratio, actually is far from it. This finding suggests that the beauty standards of this African tribe appear not to align with the golden ratio, a concept historically associated with Western ideals of aesthetic symmetry and harmony. This highlights how cultural perceptions of beauty can differ widely across societies.

The Bootstrap method has several advantages, specially because it does not require assumptions about the distribution of data, unlike the classical approach it does not require that data come from a normal distribution, so it can be applied to any data set no matter if the distribution is skewed or non-normal.

It is a flexible method that allows researchers to obtain more data in order to make some statistical inferences when available data set is limited. Besides, it is a pretty intuitive technique that everyone could comprehend.

Conclusion

After careful analysis of the data set of height-to-width ratios of women's facial symmetry in an African tribe we were able to create two confidence intervals for the mean of those ratios, each one of them using different approaches: On one side, using the *classical method* in which the researcher must assume that the sampling distribution is normally distributed, what causes that the CI does not provide reliable results. Within this context, it was observed through different plots that the original sample did not resemble a normal distribution and in consequence the confidence interval could be somehow inaccurate. Since there was few data (25 observations) and the population variance was unknown, it was decided to make the CI calculations by using a t-distribution with its respective degrees of freedom; the resulting interval was very wide ($[\mathbf{1.4818}, \mathbf{1.6262}]$), including values that seem to be far from the real value of the population mean and CI actually contains the value of the golden ratio (approximately 1.618).

On the other hand, the *Bootstrap* method was also used, this computer-based approach allowed to have a more robust data set due to the resampling process, in particular, there were taken the mean of 100,000 samples. After plotting the histogram with the data obtained it was possible to see that the data resembled a distribution very similar to the normal one, thus, in this case it was used the normal distribution to compute the critical values since the sample size was considerably large, which led to a narrower interval ($[\mathbf{1.553661}, \mathbf{1.554084}]$) that did not contained the value of the golden ration, indeed, the bounds are far from it. The bootstrapping can be particularly beneficial when the traditional assumptions are not met.

Finally, due to the previous research it is possible to conclude that the African tribe's beauty standards seem not to align with the Western concept of harmony, aesthetic and symmetry often related with the golden ratio. This suggests that beauty standards are not universal and may vary significantly depending on each society perspective.

Bibliography

Carlson, S. C. (2024, October 21). Golden ratio. Britannica. https://www.britannica.com/science/golden-ratio