A Technical Report under $AMA\hat{}-1$ perspective *

Álvaro Novillo *Universidad Carlos III*
Polo *Universidad Carlos III*

This article discusses the implementation of AMA^{-1}, a methodology that emphasizes the idea that statistics is used to solve real-world problems. The acronym AMA^{-1}, represents the three fundamental steps in this methodology: understanding the real-world problem and related data (A), translating the real-world problem into a statistical problem and solving it (M), and translating the statistical results back into terms of the original problem (A-1). The article presents basic aspects of each step along with an example of its implementation

Keywords: AMA^{-1}, Statistics

Introduction (A)

In this case, we will implement the AMA-1 methodology in R using the dataset "bioage.csv". To do so, we will follow the three fundamental steps of the methodology: understanding the real-world problem and related data (A), translating the real-world problem into a statistical problem and solving it (M), and translating the statistical results back into terms of the original problem (A-1).

The dataset used includes data on a specific animal species, detailing their weight, illness status (1= sick), and chronological age. Our article aims to investigate if there exists a method of determining the biological age of these animals and how it compares to their chronological age.

Table 1 present the preliminary analysis of the dataset. The results indicate that over half of the animals in our sample are affected by illness, with an illness status average of 0.6. Additionally, the age variable is left-skewed, suggesting a higher frequency of middle-aged animals.

Method (M)

Observations 10 R2 / R2 adjusted 0.815 / 0.763

*Replication files are available on the author's Github account (https://github.com/AlvaroNovillo). **Current version**: noviembre 10, 2023; **Corresponding author**: alvanovi@ucm.es.

Table 1: Summary of the dataset

weight	seek	age
Min. :131.4	Min. :0.0	Min. :20.09
1st Qu.:170.0	1st Qu.:0.0	1st Qu.:38.29
Median :235.2	Median :1.0	Median :44.88
Mean :229.0	Mean :0.6	Mean :44.15
3rd Qu.:290.6	3rd Qu.:1.0	3rd Qu.:53.34
Max. :304.1	Max. :1.0	Max. :58.28

Parameter	Coefficient	SE	95% CI	t(7)	p
(Intercept)	7.83	9.34	(-14.25, 29.92)	0.84	0.429
weight	0.15	0.05	(0.03, 0.28)	2.85	0.025
seek	2.09	6.99	(-14.44, 18.62)	0.30	0.774

Results (A^{-1})

```
# Step A-1: Translating the statistical results back into terms of the original problem
# Create a scatter plot of weight versus age with a regression line
library(ggplot2)
ggplot(data, aes(x=weight, y=age)) + geom_point() + geom_smooth(method="lm")
```