On Reporting the Results of a Statistical Analysis

Demetri Pananos

Introduction

Writing up an analysis is just as important as doing the analysis correctly. This document is intended to demonstrate effective ways of communicating statistics to statisticians and non-statisticians alike.

One day, you are all going to get jobs. Assuming you will be working with numbers in some capacity[1], your ability to effectively communicate results will far outweigh your ability to fit complex models. Everyone can fit a regression, but few people can fit a regression and explain it to someone who hasn't done statistics in 20 years.

Example

Suppose you are working with a clinician who asks the question *"Are there gender differences in BMI after controlling for weight?"*. I'll walk you through this small example and show you three ways you could possibly answer this question. I'll include my own thoughts on their effectiveness and quality, but leave you to be the judge.

This class will have you answer questions of a similar calibre and of a similar complexity. The art of writing a scientific paper, or full statistical report, is much more involved. If you like, I can point you to a very useful document that I still reference when writing.

Answer 1

One possible answer is just to show your collaborator what you see. Shown below is a summary of a linear regression in R.

```
##
## Call:
## lm(formula = BMI ~ weight + gender, data = diabetes)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.7494  -2.8942  -0.1467   3.3668  10.5693
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   9.22033    1.99724   4.617 1.21e-05 ***
## weight        0.33472    0.01153  29.040  < 2e-16 ***
## gendermale  -10.62100    0.92575 -11.473  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.497 on 96 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.9009, Adjusted R-squared:  0.8988
## F-statistic: 436.2 on 2 and 96 DF,  p-value: < 2.2e-16
```

---

[1] Data skills are the most indemand skills at the moment, so I think this is a reasonable assumption

Put yourself in a clinician's shoes. You've not done statistics in a very long time, and now your statistician has dumped this mess on your desk when all you really wanted was a short answer to this question.

There are 4 p-values in this table. Would you trust someone who isn't so statistically savy to pick the right p-value? What if your table included many more covariates? What if, like SAS does, the output reported several tests of the null hypothesis using different estimators? My rule of thumb is to assume that if left to their own devices, my collaborators will pick the wrong summary, not because they are ignorant, but perhaps out of confusion.

To me, this answer is lazy and suggests that whoever made this doesn't really understand what the output means. You should be able to create your own summaries of the results and tell your collaborator exactly what they want to know. Including code or code output can be informative, but it is usually best to keep it in an appednix, in the supplementary material, or preferably in a public github repository.

Answer 2

A lot of people use statistics as a black box for inference. Here is an example of such a person's potential answer:

> We regressed gender and weight onto BMI. The p value for gender is less than 0.05, so the effect of gender is statistically signficiant.

I will put my philosophical thoughts about Null Hypothesis Significance Testing aside for this answer. What does this answer tell us? Sure, the result is "statistically significant", but does it tell us the size of the effect? Does it tell us if males have on average higher or lower BMI than females of the same weight? Most importantly, does it provide us with any measures of uncertainty in our result? The answer to all of these questions is "No".

This answer means well, but falls short of effectively communicating the result. People want to hear p-values, but they need to hear uncertainty estimates. Which brings me to my next answer.

Answer 3

A really good answer should have, in my opinion, the following:

1. The answer should succintly address the question being asked. What are we intersting in determining?

2. The answer should include a plot if it aids in the understanding of the results. Tables are nice, but plots are better. Plots should always be captioned with an informative title or description. Try to be concious of color here. Many people are color blind, and choosing a poor color pallete means they would not be able to understand your plot. Consider using different shapes, or a color pallete which is able to be read by a color blind person.

3. The answer should address the method of estimation by name. This aids in reproducibility and transparency, both of which are cornerstones of science.

4. The answer should include an estimate and an accompanying confidence interval. Confidence intervals give a measure of uncertainty, which is much more valuable than a p-value. In fact, you can derive a p-value from a confidence interval, so I usually leave them out of the analysis all together.

5. The answer should round any estimates to a reasonable amount of significant figures. Your collaborator probably doesn't care about the *exact* value of the estimate. Even if they did, unless the uncertaintys is vanishly small, anything after 2, or maybe 3, digits is almost certainly a result of random/sampling variation.

6. The answer should interpret any results.

7. The answer should be concise. No one wants to read a two page answer to a two line question.
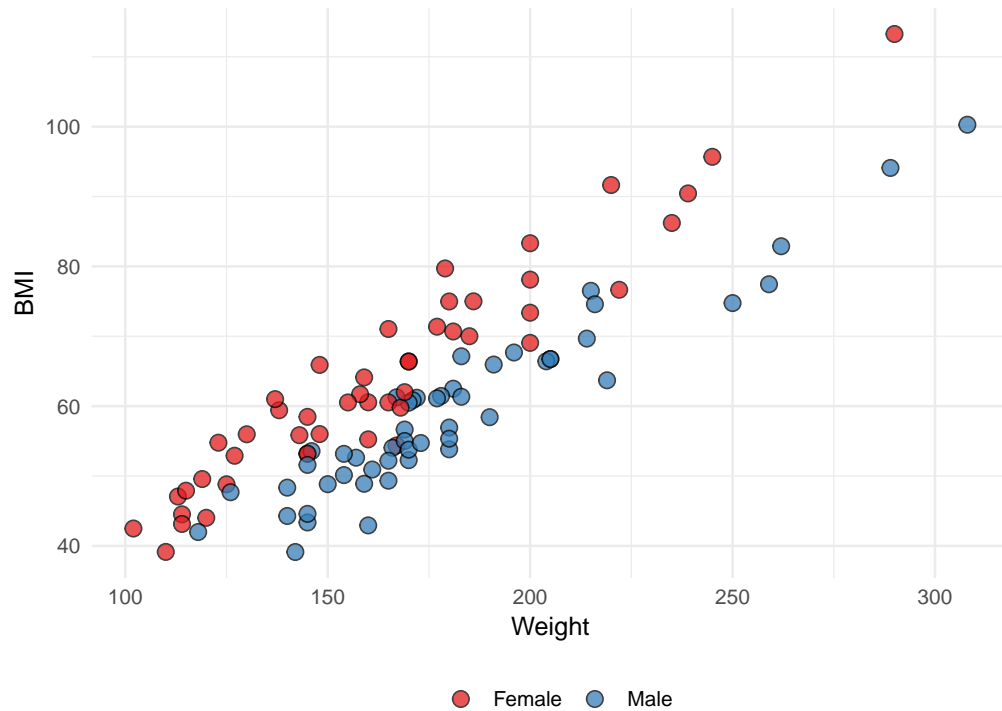
Figure 1: Plots of weight vs. BMI. The data suggest that males (blue) may have a lower BMI than females (red) of the same weight

Here is how I would answer this question:

> *I sought to determine if there was an effect of gender on BMI after controlling for weight. Figure 1 shows a plot of weight against BMI, with each data point colored by gender. From this plot, it is apparent that males have a lower BMI than females of the same weight.*
>
> *I used linear regression to perform the estimation. I estimate that males have, on average, 10.6 fewer BMI points (95% CI: (8.8, 12.4)) than women of the same weight.*

This answer provides a thorough and informative answer to the clinician's question. Now the clinican not only knows if there is an effect, but in which direction the effect operates. We are also given a 95% confidence interval, which is quite narrow, meaning we can me reasonably certain in our estimate.

Note that this answer assumes that all the assumptions of the regression have been checked. I've done this, but have not explicitly communicated it. If you go to check your assumptions, and something is wrong, you should report that as a defficiency and possible limitation of the model.

Conclusion

Good answers to statistical answers are dense in information while being easy to read and short. It takes practice to become good at this sort of thing, but it pays in spades.