Shonubi Case

2024-03-27

Shonubi's case started a long discussion about the proper use of statistical methods in the courtroom. Shonubi was a drug smuggler that was caught red handed on JFK airport on 1991-12-10 when smuggling heroin in his stomach. He was found to have swallowed 103 ballons filled with an aggregate of 427.4 grams of heroin. Investigators find out that it was not his only trip from Nigeria to the US, on a popular drug smuggling route. Shonubi had in total 8 trips, but only the data about the last one, mainly the amount of heroin he was carrying, is available. However, the amount of drugs he smuggled in total is significant, because the sentence he would receive for smuggling more than 3000 grams will be much higher than for smuggling less than 3000 grams. The whole discussion is how to asses the amount of drugs he smuggled in total, and how to use this information in the court, as the only information possible to obtain are projections with some degree of unccertainty.

To create estimations of drugs that Shonubi could smuggle a dataset of drug smugglers was obtained. It's not perfect unfortunetally ass there are missing values when it comes to net and gross amount of drugs that was smuggled in the stomachs of those people. We have a divsion on net and gross values because drugs were smuggled in the ballons, and the weight that counts is the pure narcotics.

The most roughly and initial estimate of the narcotics he could smuggle was $427.4 \times 8 = 3,419.2$, so just a multiplication of his last trip smuggle.

> Boyum simulated 100,000 possible series of seven previous trips by Shonubi using repeated sampling with replacement of the 117 net weights. The simulations produced a Gaussian-looking histogram of seven-trip totals. Boyum concluded from these simulations that "there is a 99

But this estimation didn't not account for so called "trip effect", where the smugglers are expect to swallow more with each trip, as they are getting more experienced. Also an epistemological worry, are we justified to project Shonubis perforamnce as a drug trafficar given the information only about different people? It it a valid worry, especially when we look at the variety of amounts of drugs that were smuggled by different people.

Main problems of this case and the estimations:

> They point out that (1) outliers in the data tend to distort implicit distributional assumptions and hence any resulting inferences; (2) the 427.4 grams Shonubi carried on his last trip was suspiciously smaller than would be expected based upon the data; (3) if a learning curve were used to impute missing data, the sensitivity of the resulting estimate of the total amount of heroin carried by Shonubi in his eight trips would depend upon the manner in which the learning curve is defined; and (4) regardless of how the learning curve is defined, it would still be very difficult to obtain an estimate of that total that would be less than 1,000 grams.

```
shonubi <- read.csv("../ShonubiCaseDataset.csv")

head(shonubi, n=3)
```

```
  obs dataset balloons gross_wt net_wt purity age_yrs gender
1   1       1       79    742.4  503.2   0.51      NA     NA
2   2       1       90    901.7  576.9   0.32      NA     NA
3   3      12       90    800.2  573.3   0.85      38      1
```

The dataset has 248 observations, many values are missing what is especially problematic for a dataset of this size, importantly only **PERCENT** of observations have both net and gross weight of the drugs.