

Orthogonal Matrices

1. Estimates

The median is around 70, but the distribution tends lower, so I guess $mean = 69$.

My guess : $\sigma = 4$.

Units on both are whatever height was measured in, but the axis isn't labeled, so actually knowing is hard.

Seriously? Label your axes.

To be honest, it doesn't matter that much, jost indicates a lack of proofreading on your part.

2. Histogram estimates

The median is around \$30,000, but the distribution tends substantially higher, so I guess $mean = \$60,000$. This estimate is really hard with the "over 100,000" bucket being so big.

My guess : $\sigma = \$35,000$.

3. 1's and 3's

$mean = 2, \sigma = 1$

4. Manual histograms

Bin size = 5 (round number)

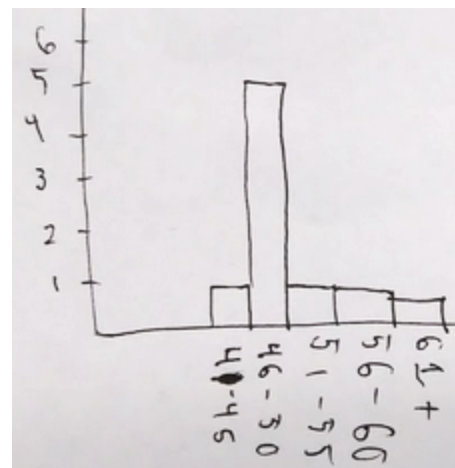
Centering around ...1 - ...5

Results in 5 bins for 9 points, which is a little low, but tolerable

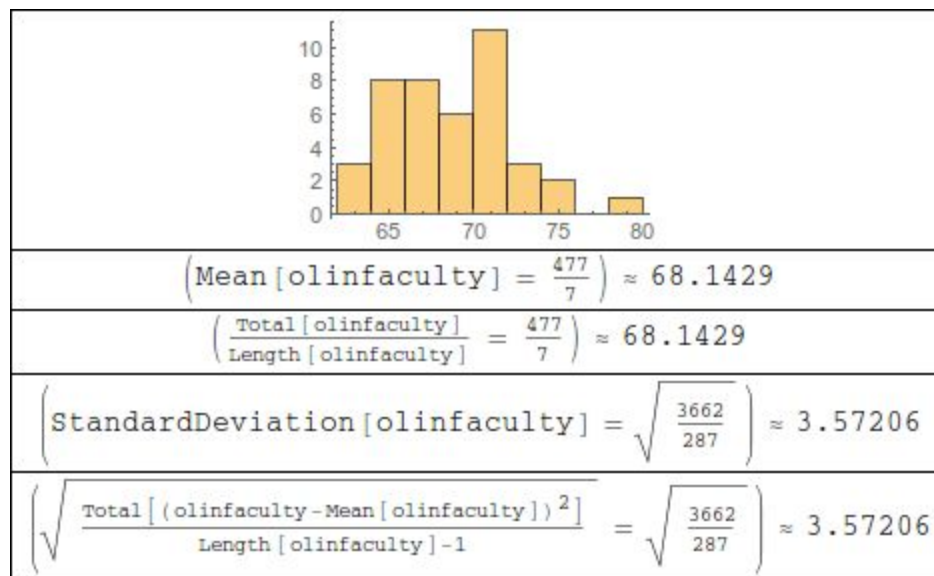
$$mean = \frac{446}{9} \approx 49.6$$

$$variance = \frac{725}{18} \approx 40.3$$

$$\sigma \approx 6.35$$



5. Automatic Histograms



Mathematica's definition of variance divides by $N - 1$, which seems to agree with the internet but not with your description.
Both values look reasonable.

Edit: after working further, I realized that your definitions all agree with “population” statistics, while Mathematica’s definitions match “sample” statistics. I’m curious where that difference is coming from, and while I can see the merit of introducing the simpler math first, it is probably worth a footnote explaining what is going on.

Correlation

6. States

University \Leftrightarrow Income
 Infant Mortality \Leftrightarrow -Income
 Unemployment \nRightarrow Doctors

7. Hand calculations

Working out these calculations by hand isn’t a useful use of my time. I’ll come back if I get less busy.

To be fair, this problem is far from useless. I was just in a bad mood, and had enough other things on my daily schedule. If I’d planned ahead farther, I would have done it.

8. Proper calculations

University	Income	Mortality
-1.08552	-1.21398	1.57928
-0.00509633	1.34336	-0.0652596
-0.453573	-0.391869	-0.456817
-1.73785	-1.59578	1.18772
0.463766	0.605823	-1.55318
1.68688	0.206469	-1.005
1.68688	1.35674	-0.61344
0.0356743	0.305217	1.0311
-0.310876	-0.707149	0.247986
0.0356743	-0.401486	0.874478
0.361839	1.21983	-1.08331
-0.677812	-0.727176	-0.143571

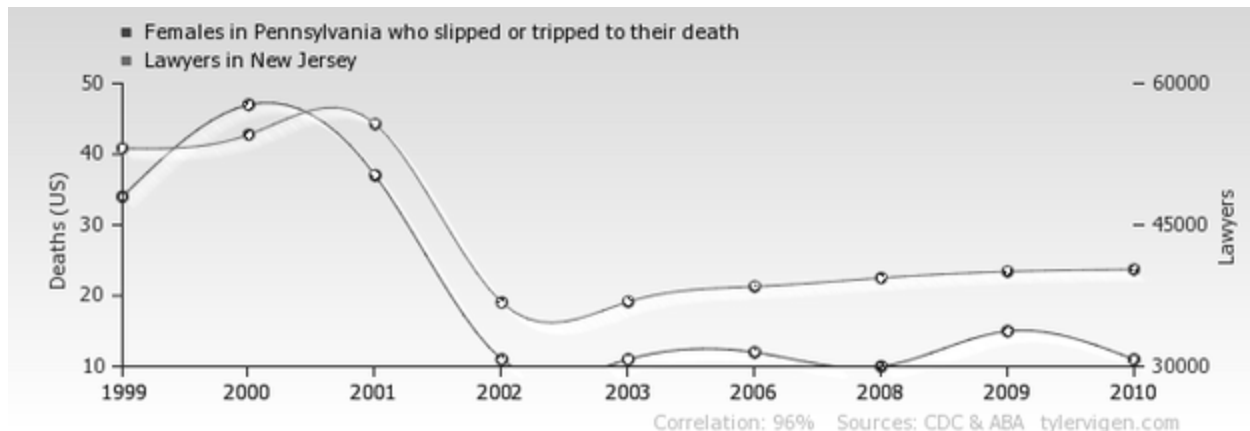
Normalized vectors

Correlation coefficients come from the dot products of the two vectors

Matrix multiplication (with the horizontal matrix in front) then division by the length of the list produces the desired result

$$\begin{pmatrix} 1. & 0.757318 & -0.665918 \\ 0.757318 & 1. & -0.648249 \\ -0.665918 & -0.648249 & 1. \end{pmatrix}$$

9. Spurious Correlations



Linear Regression

Oh man, I did a bad job of analyzing my work here. I should get better at that, it will probably improve retention and comprehension substantially.

10. Equations

$$\left\{ \left\{ a \rightarrow \frac{1}{120}, b \rightarrow \frac{43}{2} \right\} \right\}$$

11. Matrix equations 1

$$\begin{pmatrix} 60 & 1 \\ 300 & 1 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 60a + b \\ 300a + b \end{pmatrix}$$

12. Matrix equations 2

$$A^{-1} = \begin{pmatrix} -\frac{1}{240} & \frac{1}{240} \\ \frac{5}{4} & -\frac{1}{4} \end{pmatrix}$$

$$A^{-1}b = \begin{pmatrix} \frac{1}{120} \\ \frac{43}{2} \end{pmatrix}$$

13. Cricket sums

In order: {285, 595, 9589, 19441}

14. Equations:

$$\{9589a + 285b = 19441, 285a + 9b = 595\}$$

15. Solving the Regression

$$\begin{pmatrix} 9589 & 285 \\ 285 & 9 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 19441 \\ 595 \end{pmatrix}$$

$$b = \frac{82385}{2538} \wedge a = \frac{899}{846}$$

