Doctor, is this normal?
ULAB, Division of Sincerely Serious Science

Gavin Tamasi

November 15, 2024

1   Introduction

According to my sources, around May 2021 an image began circulating around the internet about a broken sidewalk. The image, which can be seen in Figure 1, claims that the sidewalk crack resembles the form of a normal distribution. This lab report will test that claim by fitting this curve with a normal standard distribution (also called a Gaussian distribution).

Figure 1: Sidewalk image from the internet
Figure 1

2   Methods

2.1   Preprocessing

First, we used the Perspective Wrap tool in Photoshop to correct the titled perspective of the image. Then, we worked with the `Python Imaging Library` and `NumPy` and converted our image to grayscale.

A threshold mask was applied to the image so that only dark pixels would be retained. Note that some

minimal manual processing was required to remove the vertical sidewalk crack as well as outlier pixels.

Finally, the location of the dark pixels are recorded as coordinates in a Cartesian plane.

2.2 Curve Fitting

We worked with `SciPy` to fit a normal distribution, which is given by:

$$f(x) = a \cdot \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) \tag{1}$$

where a is a scaling constant, μ is the mean, and σ is the standard deviation.

3 Results

After our data analysis, Figure 2 reveals that a normal distribution fits our data well.
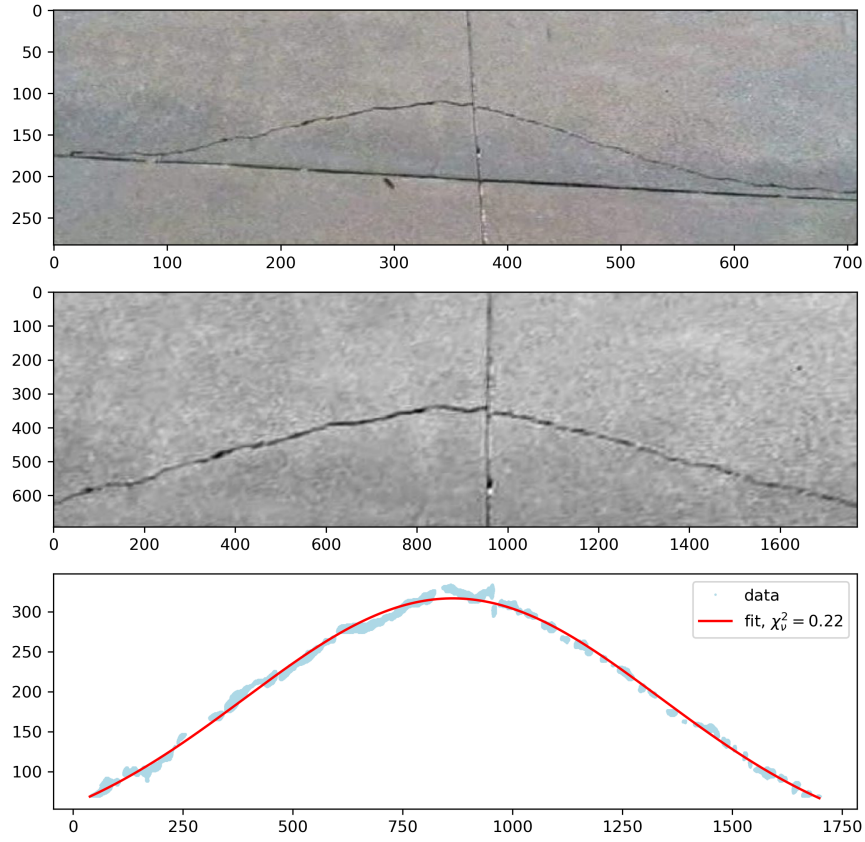

Figure 2: Top: original image, Middle: processed image with perspective correction and greyscale transformation, Bottom: curve fit of the sidewalk crack points.

Furthermore, we determined the fit to have a reduced chi-squared value, defined as:

$$\chi_\nu^2 = \frac{1}{\nu} \sum \frac{(O - E)^2}{E} = 0.22 \tag{2}$$

where ν is the number of degrees of freedom. Of note, we used the Pearson chi-squared definition because it is difficult to estimate the uncertainty of our points.

The reduced chi-squared value is less than 1 and quite small. Therefore, our model is likely correct, but the data may be slightly overfitted. Regardless, we have shown that the crack is well-described by a normal curve!

References

[1] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, vol. 17, pp. 261–272, 2020.

[2] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. Fernández del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant, "Array programming with NumPy," *Nature*, vol. 585, p. 357–362, 2020.