3.4 Average of Cauchy variables

The probability distribution for a Cauchy variable is

$$p(x) = \frac{1}{\pi(1+x^2)}$$

and so we can see that the variance is indeterminate, because

$$\int x^2 \,\mathrm{prob}(x)\, dx$$

diverges.

We have seen that the sum of variables is a convolution: therefore the product of the Fourier transforms of the distribution should give us the distribution of the mean. The transform of this Cauchy distribution is

$$P(k) = e^{-|k|}$$

where k is the Fourier variable, and there will be some constant in front, depending on the particular definition of the transform.

If we add N identically-distributed Cauchy variables to get a sum s, then the transform of the distribution of s will be

$$P(k') = e^{-|Nk'|}$$

and, inverting this transform, we find

$$p(s) \propto \frac{N}{\pi(N^2 + s^2)}.$$

Remembering that the average $a = s/N$, change variables to get

$$p(a) \propto \frac{N^2}{\pi(N^2 + (Na)^2)}$$

which shows that the mean a has the same distribution as a single observation x!

To find a better estimator: if our N data are X_i, and the distribution has an unknown location μ, then the probability of getting this particular dataset is

$$\mathcal{L} = \prod_i \frac{1}{1 + (X_i - \mu)^2}.$$

We can find the value of μ that maximizes $\mathcal{L}$ – this is intuitive. We discuss this estimator much more in Chapter 6. In the present case, maximization gives a high-order polynomial equation in μ. There is no neat analytical solution, but numerical maximization is fast. Try this with some test data drawn from a Cauchy distribution. The estimator finds the mean, and the scatter on the estimator drops off nicely as $\sqrt{N}$, just as you might hope.

Jaynes (1983) gives an insightful discussion of this problem (in the paper "Confidence Intervals vs Bayesian Intervals") where he shows how, with just two data, the estimate $X_1 + X_2$ can be greatly improved by making it conditional on the (also known!) value $X_1 - X_2$.