

# Convolutions and Their Tails

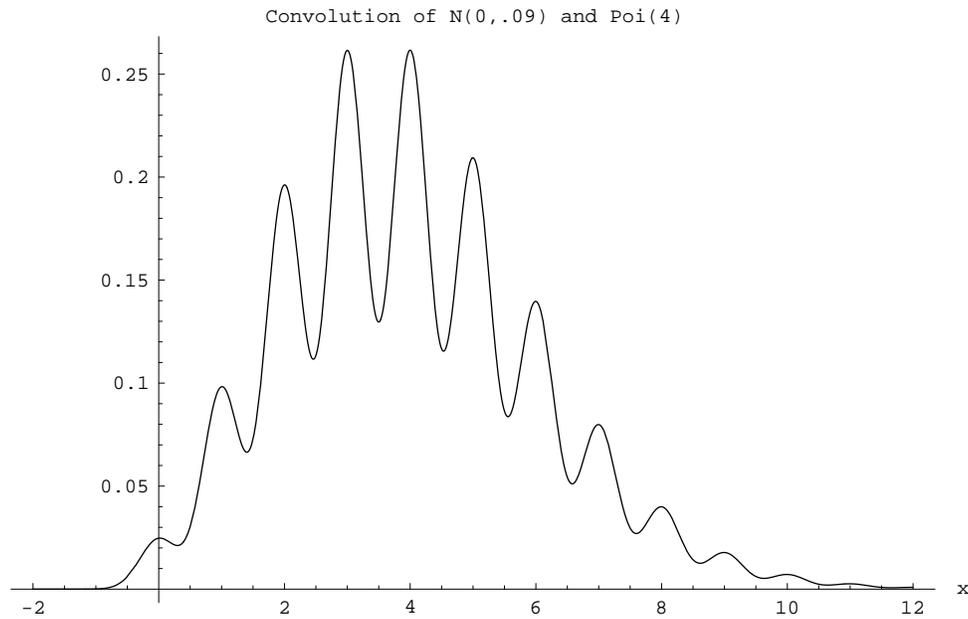
Anirban DasGupta

$$Y, Z : (\Omega, \mathcal{A}, P) \longrightarrow (\mathcal{X}, \mathcal{B}),$$

$Y, Z$  independent.

Distribution of  $Y + Z =$  Convolution of the distributions of  $Y, Z$ .

Colloquial: Convolution of  $Y$  and  $Z$ .

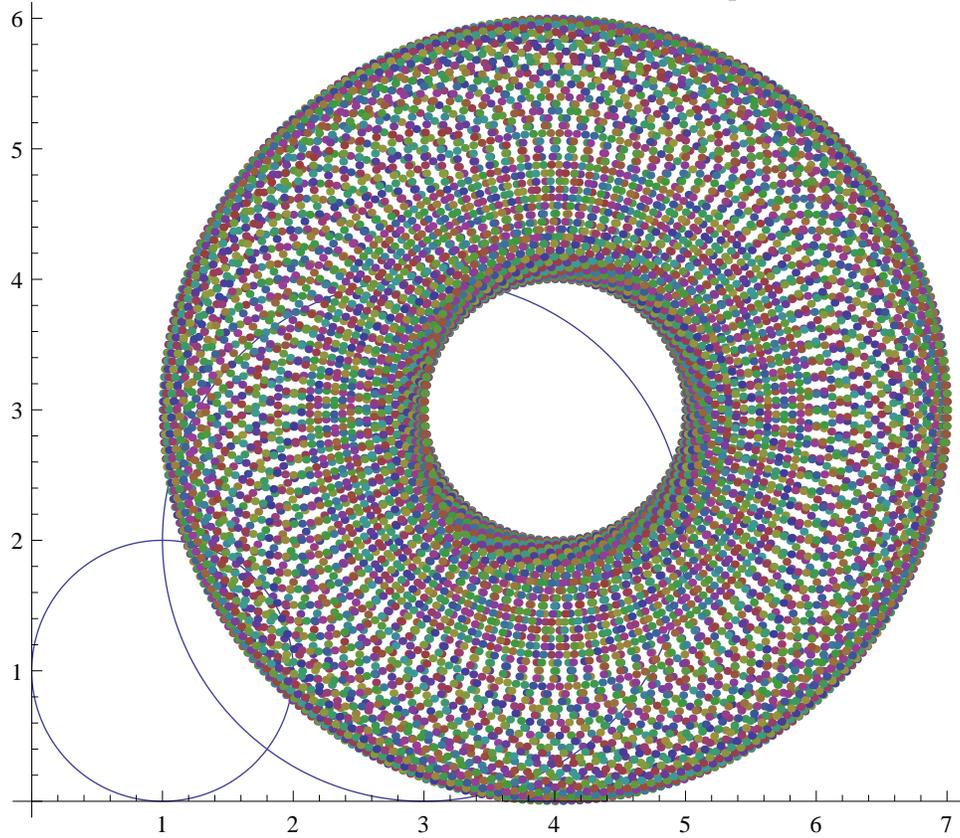


## Examples

1.  $Y \sim Poi(4), Z \sim N(0, .05)$ .

Then the convolution has a density, and here is the density.

Realization of the Convolution of Two Random Spheres



2.  $Y, Z$  are two *random spheres*.

Then, here is how one realization from the convolution of  $Y$  and  $Z$  looks like.

## Dilations

$Z$  a general random variable;  $Y$  an independent nonnegative real valued random variable.

The distribution of  $YZ$  is a dilation of the distribution of  $Z$ .

### Example

$$Z \sim N(0, 1), \quad \frac{1}{Y^2} \sim \chi_1^2 \implies YZ \sim C(0, 1).$$

**Fact** If the random variables concerned take values in some finite dimensional Euclidean space, and we are allowed to both dilate and convolve, then we can fold in a lot of distributions.

**Theorem 1**

Let  $Z$  be an arbitrary  $\mathcal{R}^d$ -valued random variable with an a.c. distribution. Let  $Y \geq 0, W \mathcal{R}^d$ -valued. Everything independent. Then, the dilation-convolution class of random variables  $YZ + W$  is dense in the class of all a.c. distributions on  $\mathcal{R}^d$ .

**Corollary** The class of Gaussian dilation-convolutions is dense.

## Theorem 2 (When are Dilations also Convolutions)

Let  $Z \sim N(0, 1)$  and consider a Gaussian dilation  $YZ$ . The dilation is also a Gaussian convolution  $Z + W$  if and only if  $P(Y < 1) = 0$ , in which case  $W$  is also a Gaussian dilation,  $W = UZ$ , and  $P(U \leq u) = P(1 \leq Y \leq u + 1)$ .

Corollary The Gaussian convolution class does not contain familiar distributions.

For example,

$$C(0, 1) \neq N(0, 1) + \text{Anything.}$$

But, curiously,

$$C(0, 1) = t(3) + \text{Something,}$$

etc.

## Message

- (a) The Gaussian dilation-convolution class is very rich.
  - (b) The Gaussian dilation class is quite rich.
  - (c) The Gaussian convolution class is not practically rich.
  - (d) Gaussian dilations are automatically unimodal.
  - (e) Arbitrary Gaussian convolutions are not necessarily unimodal. But Gaussians convolved with unimodals are always unimodal.
  - (f) Which Gaussian convolutions are unimodal can be characterized.
- Two theorems on property (e) and property (f) are the following:

**Theorem 3** Consider a Gaussian convolution  $Z + Y$ .

(a) If  $Y$  is unimodal, so is  $Z + Y$  (Ibragimov)

(b) If  $Y$  has at most  $k$  local maximas, then  $Z + Y$  has at most  $k + 1$  local maximas (Follows from a joint thm. of Samuel Karlin and Herman Rubin).

The next theorem is:

## **Theorem 4**

Consider an arbitrary Gaussian convolution. It is unimodal if and only if conditions † and ‡ hold.

A useful corollary of this if and only if result is this:

**Corollary** Consider a general symmetric unimodal random variable  $X$ . Write  $X$  in the form  $X = U[-V, V]$ .  $X$  cannot be a Gaussian convolution if  $V$  is infinitely divisible.

## Example

Take  $X = N(0, 2) = N(0, 1) + N(0, 1)$ , by construction a Gaussian convolution. Now write  $N(0, 2) = U[-V, V]$ . The associated  $V$  has density  $cv^2e^{-v^2/4}$ ,  $v > 0$ .  $V$  is not infinitely divisible.

Well known Gaussian dilations are usually not Gaussian convolutions.

This is a generalizable phenomenon, in some sense. Here is one theorem.

**Theorem 5** Consider a random variable  $Z$  whose characteristic function is of regular variation (at infinity). Consider dilations of such a random variable  $Z$ . They cannot be obtained by convolving  $Z$  with any absolutely continuous distributions.

Author proves it by using the regular variation, and by proving that the cf of the convolver must have a strictly positive liminf, which'd violate Riemann-Lebesgue lemma.

Interestingly, the phrase *cannot be obtained by convolving  $Z$  with any absolutely continuous distribution* is not relaxable. Here is a counterexample.

## Example

Consider  $Z$  a standard double exponential; its cf is  $\frac{1}{1+t^2}$ , which is of regular variation.

Consider  $X$ , also a double exponential, but with a scale parameter  $\sigma > 1$ . So,  $X$  is a trivial dilation of  $Z$ !

Curiously,

$$X = Z + Y,$$

where

$$Y \sim \left(1 - \frac{1}{\sigma^2}\right)DE(\sigma) + \frac{1}{\sigma^2}\delta_0.$$

In words, if  $\sigma > 1$ ,

$$DE(\sigma) = DE(1) + \left[\left(1 - \frac{1}{\sigma^2}\right)DE(\sigma) + \frac{1}{\sigma^2}\delta_0\right].$$

So, here is a DE dilation which is also a DE convolution, but the convolver  $Y$  isn't a.c.

## Tails

Convolutions increase uncertainty. So, tails of Gaussian convolutions would be appropriately heavier than the standard normal tail.

## Question

What is a precise quantification?

## Theorem 6

Suppose  $p(z)$  is a density function on the real line, and

(a)  $q(z)$  is ultimately nonnegative;

(b)  $w(z) = -\frac{d}{dz} \log q(z)$ ,  $v(z) = -\frac{d}{dz} \log p(z)$  exist and are functions of regular oscillation;

(c)

$$\begin{aligned} & \liminf_{z \rightarrow \infty} \frac{d}{dz} \log q(z) \\ & > \liminf_{z \rightarrow \infty} \frac{d}{dz} \log p(z). \end{aligned}$$

Then,

$$\begin{aligned} p * q(\gamma) &= \int_{-\infty}^{\infty} q(z)p(\gamma - z)dz \\ &= q(\gamma) \int_{-\infty}^{\infty} e^{-zw(\gamma)} p(z)dz (1 + o(1)), \end{aligned}$$

as  $\gamma \rightarrow \infty$ .

## A Statistical Application

### Example (Regularized Estimate of Normal Mean)

$X_1, \dots, X_n \stackrel{iid}{\sim} N(\theta, 1)$ . Let

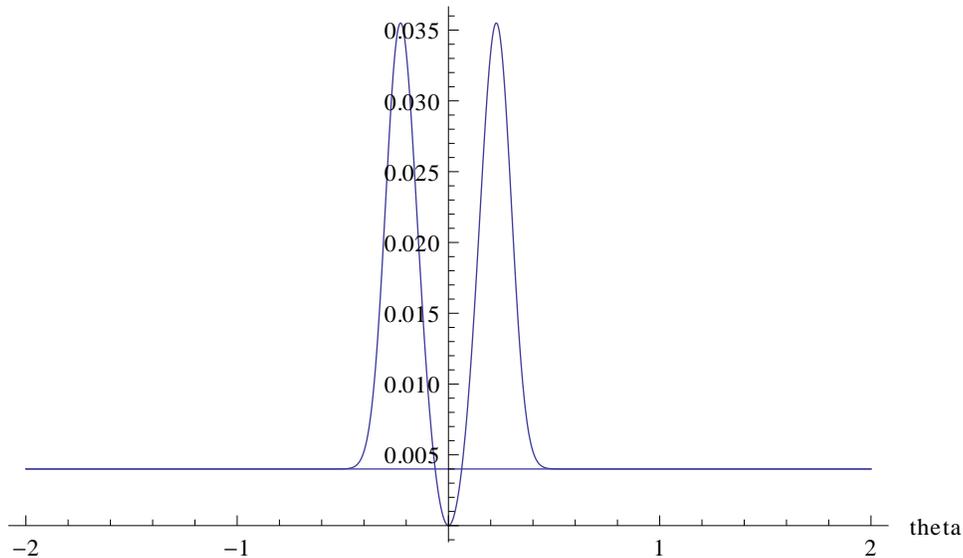
$$S_n(X_1, \dots, X_n) = \begin{cases} \bar{X}_n & \text{if } |\bar{X}_n| > c_n \\ 0 & \text{if } |\bar{X}_n| \leq c_n \end{cases}$$

$c_n$  is a sequence  $\rightarrow 0$ , *BUT NOT TOO FAST*. In fact, you'd like  $\gamma_n = \sqrt{n}c_n \rightarrow \infty$ .

Such regularized estimates of the mean  $\theta$  provide large gains in estimation accuracy (risk) if  $\theta \approx 0$ , but have larger risks than unregularized estimates otherwise.

Here is a typical picture of risk of regularized vs. unregularized estimates.

Risk of Hodges' Estimate for  $n = 250$



Therefore, you'd expect that if  $\theta$  has a *sufficiently peaked prior distribution*, then the regularized estimate has smaller Bayes risk than the unregularized estimate.

**What Can be Proved?** Suppose

$$\pi(\theta) = \pi_n(\theta) = \sqrt{nh}(\theta\sqrt{n}), \quad -\infty < \theta < \infty.$$

Define the signed function

$$q(z) = \int_0^z (t^2 - 1)h(t)dt - h'(z), \quad -\infty < z < \infty;$$

$$w(z) = -\frac{d}{dz} \log q(z).$$

Let  $B_n(S_n, \pi)$  denote Bayes risk of the regularized estimate  $S_n$  under  $\pi$ .

Then, we have

## Theorem 7

$$B_n(S_n, \pi) = \frac{1}{n} + \frac{2}{n} (q * \phi)(\gamma_n),$$

where  $\gamma_n = \sqrt{nc_n}$ .

This is how convolution tails enter into this problem.

This leads to the general result on Bayes risks of regularized estimates:

### **Theorem 8**

Suppose  $-\log q(z)$  is a function of regular variation at  $z = \infty$ . Then,

$$B_n(S_n, \pi) = \frac{1}{n} + \frac{2q(\gamma_n)e^{\frac{1}{2}[w(\gamma_n)]^2}}{n} (1+o(1)),$$

as  $n \rightarrow \infty$ .

## Example

Suppose the prior was normal!

$$\theta \sim N\left(0, \frac{\tau^2}{n}\right).$$

Then,

- 1)  $B_n(S_n, \pi) \equiv \frac{1}{n} \forall n$ , if  $\tau^2 = 1$ ;
- 2)  $B_n(S_n, \pi) > \frac{1}{n}$  for all large  $n$  if  $\tau^2 > 1$ ;
- 3)  $B_n(S_n, \pi) < \frac{1}{n}$  for all large  $n$  if  $\tau^2 < 1$ .

## A Small Open Problem

Recall the second theorem of this talk:  
If  $Z \sim N(0, 1)$ ,  $X$  is a dilation of  $Z$ ,  $X = YZ$ , then  $X$  is a Gaussian convolution,  
 $X = Z + W$  iff  $P(Y < 1) = 0$ .

A generalization of the only if part of this theorem is as follows.

## **Theorem 9**

Let  $X_1 = Y_1Z$ ,  $X_2 = Y_2Z$  be Gaussian dilations. Let  $\nu(A) = P(Y_1 \in A) - P(Y_2 \in A)$ , and let  $\nu = \nu^+ - \nu^-$  be the Jordan decomposition of  $\nu$ .

Then,  $X_1 = X_2 + W$  only if  $\nu^-[0, t) = 0 \Rightarrow \nu^+[0, t) = 0$ .

In Theorem 2, we had  $Y_2 \equiv 1$ , and in this case, we can easily show

$$\nu^+[0, 1) = P(Y_1 < 1); \quad \nu^-[0, 1) = 0.$$

So, by Theorem 9, if  $P(Y_1 < 1) > 0$ , then we cannot write  $X_1$  as  $X_2 + W$ , which was one part of Theorem 2. In Theorem 2, the result was if and only if; in Theorem 9, only the only if part has been proved.

Merci beaucoup