

# PROPOSAL FOR FITTING LINEAR MODELS WITH GENERALISED HYPERBOLIC NOISE

ZULFIKAR MOINUDDIN AHMED

I was the pioneer in fitting social science data with Barndorff-Nielsen's Generalised Hyperbolic models. In this note I will propose that all noise in Social Science is Generalised Hyperbolic. In fact my hypothesis is that all noise in Nature is Generalised Hyperbolic.

## 1. THE EXAMPLE THAT JUSTIFIES PROPOSAL

This is Happiness predicted by three factors, Trust of all people, Autonomy, and Respect. Respect is proxied by importance of teaching young children tolerance and respect for all people. The ordinary least squares fits the data for  $N = 79$  countries with  $R^2 = 0.67$ .

```
> summary(mod.haptrust)
```

Call:

```
lm(formula = trustMost ~ Happy + childResp  
+ Autonomy, data = haptrust)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-29.775	-6.516	0.719	5.106	33.856

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-62.3829	13.4673	-4.632	1.49e-05 ***

---

*Date:* April 9, 2021.

Happy	0.5853	0.1496	3.912	0.000199	***
childResp	0.3409	0.0911	3.742	0.000355	***
Autonomy	1.3159	0.1489	8.836	3.06e-13	***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 10.52 on 75 degrees of freedom  
 Multiple R-squared: 0.6714, Adjusted R-squared: 0.6583  
 F-statistic: 51.08 on 3 and 75 DF, p-value: < 2.2e-16

## 2. BENEFITS OF NEW NOISE MODELS

A little bit of experimentation shows that it is quite difficult to improve on the  $R^2$  in the above model with linear models with Gaussian assumptions, i.e. the standard linear model. I use R which has an extremely high quality fitting and analysis code for fitting these.

Now I will show you the residual distribution of model in the last section.

Asymmetric Generalised Hyperbolic Distribution:

Parameters:

lambda	alpha.bar	mu
-2.2650043	0.9937496	-1.1687394
sigma	gamma	
10.2361129	1.1861608	

Call:

fit.ghypuv(data = hapres)

Optimization information:

log-Likelihood:	-293.4147
AIC:	596.8294

### 3. CONCLUSION

We see that we have nontrivial  $(\lambda, \bar{\alpha}, \gamma)$  in our linear model residuals proving they are fit much better with Generalised Hyperbolic Distribution. This is not an exceptional situation. I claim that every single residual in every single social science linear model will have better fit by GHD and therefore that linear models with Gaussians are *never* scientifically appropriate for Social Science. I propose that for Social Science and for all of Science least square linear fits be permanently replaced by linear models with assumed GHD fits that can simultaneously estimate the linear parameters *and* the noise parameters  $(\mu, \sigma, \lambda, \bar{\alpha}, \gamma)$ . The optimal linear parameters will differ but *be much more accurate to Nature*, i.e. we will have scientifically superior models every single time than linear models and Social Science will be advanced because the noise is so high in social science data, no one will see any better models at all without addressing the noise in this way.

### 4. THE OPTIMISATION PROBLEM

Suppose we consider the model

$$(1) \quad y = Ax + b + \varepsilon$$

a multivariate linear model with noise from a parametric distribution

$$(2) \quad \varepsilon \sim G(\theta)$$

Let

$$(3) \quad \ell(\theta)$$

be the log-likelihood function for  $G(\theta)$ . Then the optimisation problem for fitting linear model with  $G(\theta)$  noise is

$$(4) \quad (A^*, b^*, \theta^*) = \arg \min_{(A, b, \theta)} \|y - Ax - b\|_2 - \ell(\theta)$$

And the solution to this optimisation problem will give us both the linear parameters  $(A, b)$  and the noise parameter  $\theta$ .

There are many ways to solve these problems efficiently in the literature. Any of these will resolve the problem. I will examine some natural methods to implement these later, but I am not interested in novelty in fitting techniques. The important *scientific* issue is that for all Natural Sciences and Social Science, the particular choice of  $G(\theta)$  being Ole Barndorff-Nielsen's Generalised Hyperbolic Distribution will produce a *Scientific Advance* for all of Social Science. Note that (4) is not the same at all as the least-square problem, so we expect that the optimal values will produce different  $(A, b)$  estimates than least squares, and they might differ significantly from least square linear model estimates.