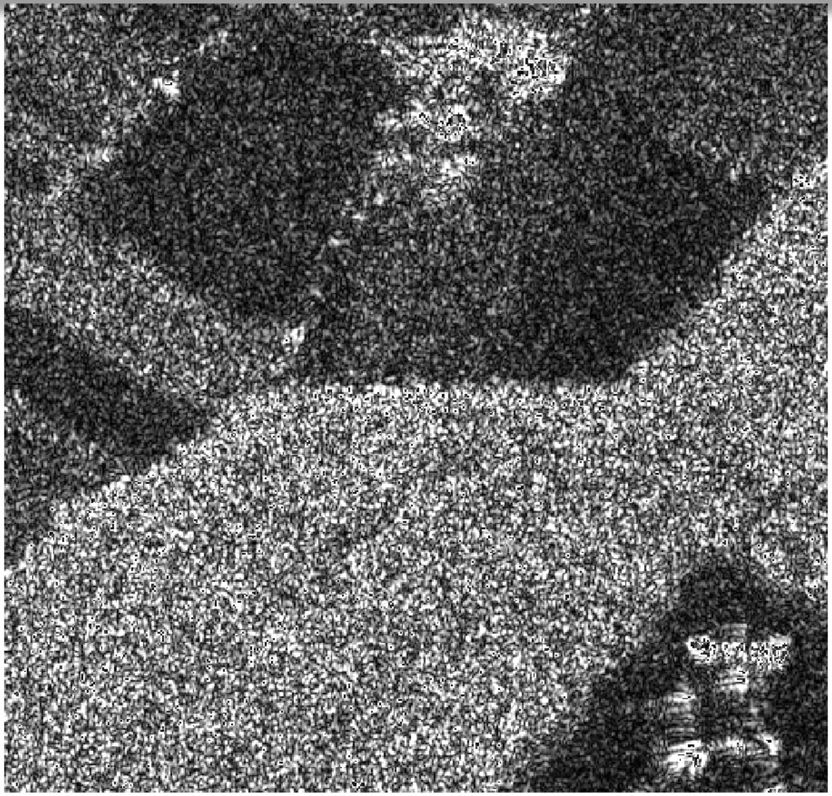
Statistics and Remote Sensing: An Educational Approach

(with emphasis on SAR image analysis)

Alejandro C. Frery
School of Mathematics and Statistics
Te Herenga Waka – Victoria University of Wellington
New Zealand



My first encounter with SAR

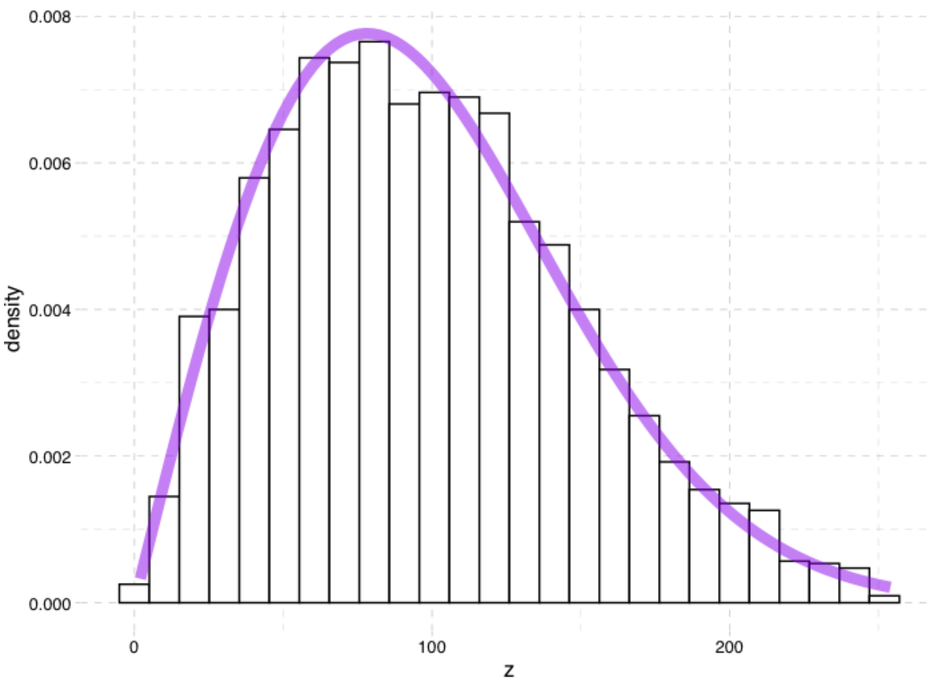
During my PhD, in 1990, my advisor told me to look at this image

It was obtained by the SAR-580 campaign

He also told me that the theory stated that the data should follow a Rayleigh distribution, so I selected an area, devised and implemented a maximum likelihood estimator, and…

Voilà, I was hooked with this kind of images!

I'd never seen such a good agreement between theory and evidence.

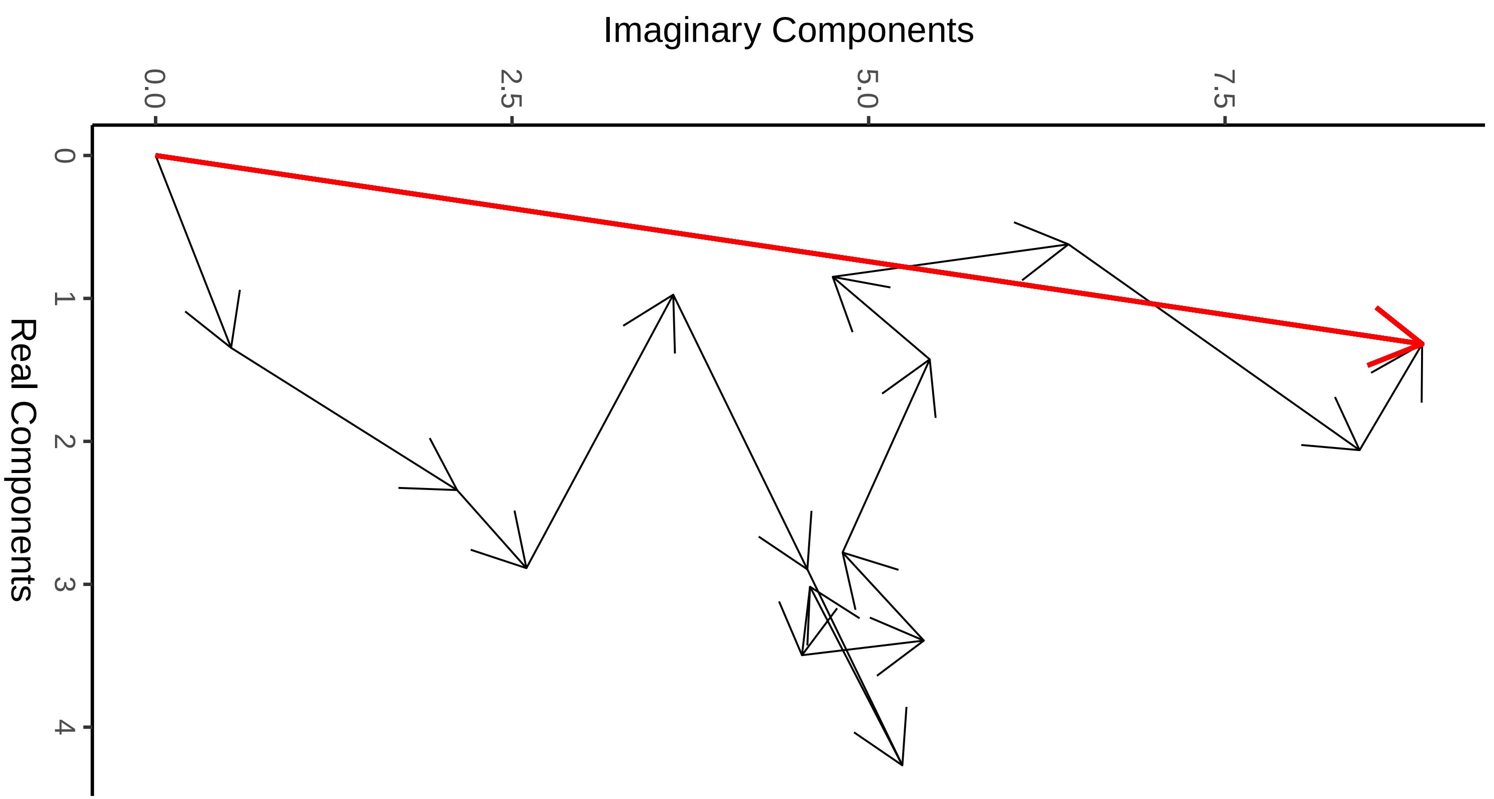


The theory of SAR

- What was the underlying theory leading to the Rayleigh distribution?
- The observation at each pixel is the result of the sum of
  - Infinitely many elementary backscatterers
  - Each backscatterer contributes with random complex signal (amplitude and phase)
  - Amplitudes and phases are collectively independent
  - Amplitudes are comparable
  - Phases are not informative

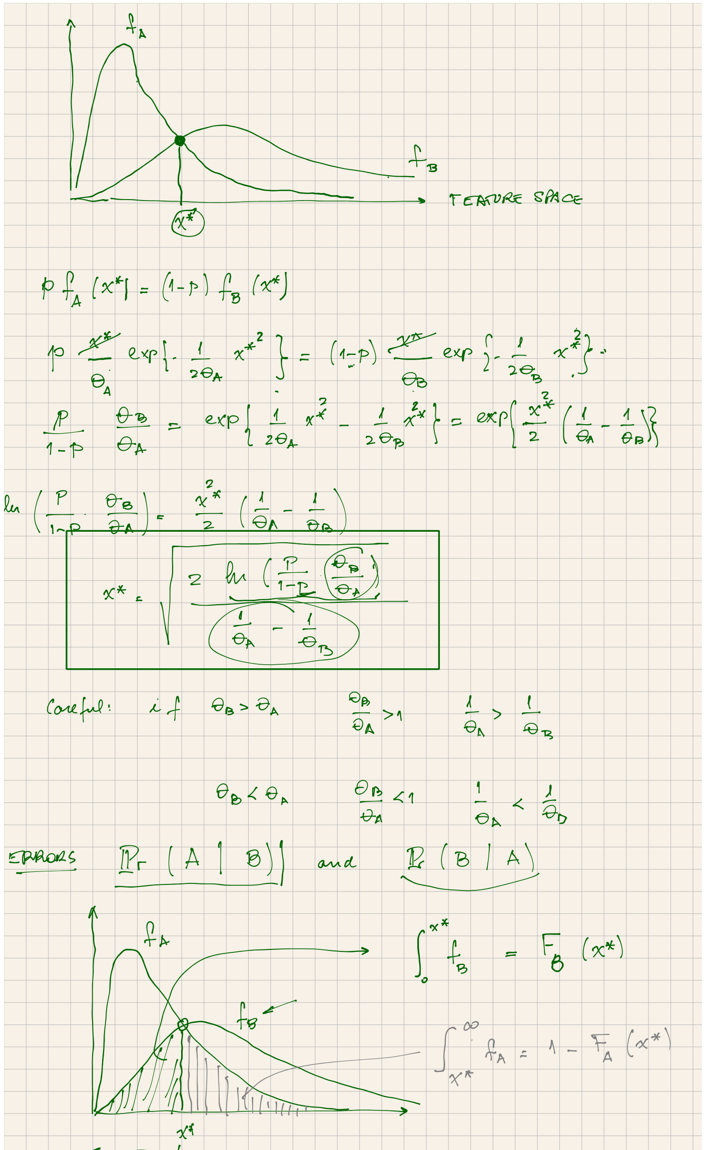
Geoscience and Remote Sensing Society

The theory of SAR



The theory of SAR

- We can then apply the Central Limit Theorem
- The intensity follows an Exponential distribution, and the amplitude obeys a Rayleigh law
- All the information lies in the parameter that indexes the Rayleigh model.
- This is a phenomenological bottom-up physically-based way of describing the data
- This is a very successful model:
  - Expressive
  - Tractable



I spent a good deal of my PhD working with the Rayleigh distribution

The Theory of SAR

- But technology evolved
- The amplitude format was no longer mandatory
- Multilook images became easier to obtain

The Equivalent Number of Looks

- SAR images are very noisy
- The noise can be mitigated while processing the data by *multilooking*
- It consists of averaging L ideally independent intensity observations: the *looks*
- The resulting distributions for the data are
  - Gamma law, if in intensity format
  - Square Root of Gamma (Nakagami) model, if in amplitude format

The Equivalent Number of Looks

- Images are provided with a *nominal* value of L
- The number of looks is a measure of the signal-to-noise ratio, i.e., of the image quality: more looks, less noise, more value
- The nominal looks seldom reflect the actual properties of the data
- Estimating L was the way to measure the image quality

The Equivalent Number of Looks

The gamma distribution and ENL

$$Z \sim \Gamma(\mu, L) \qquad f_Z(z; \mu, L) = \frac{L^L}{\mu^L \Gamma(L)} z^{L-1} \exp\{-Lz/\mu\}$$

Moments $\quad \mathbb{E}(Z) = \mu \qquad \mathrm{Var}(Z) = \dfrac{\mu^2}{L}$

$$CV(Z) = \frac{\sqrt{\mathrm{Var}(Z)}}{\mathbb{E}(Z)} = \frac{1}{\sqrt{L}}$$

Therefore $\quad L = \dfrac{1}{CV^2(Z)}$ and $\boxed{\hat{L} = \dfrac{\bar{\mu}^2}{s^2}}$

We may take samples, compute the ratio of the square sample mean to the sample variance to estimate L, and then their (weighted?) average.

The Equivalent Number of Looks

Also

$$\sqrt{L} = \frac{E(z)}{\sqrt{Var(z)}}$$

$$\sqrt{Var(z)} = \frac{1}{\sqrt{L}} E(z)$$

We may take samples, and form a regression model

There are at least five different approaches in the literature.

What does a practitioner need to know?

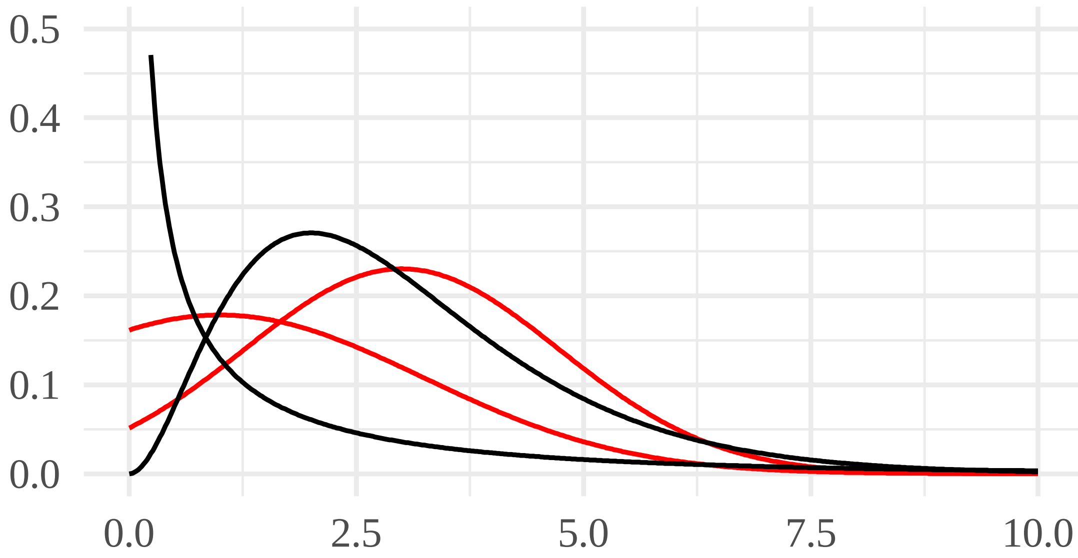
The required knowledge up to this point is:

- The Central Limit Theorem
- The Gamma model
  - Where it comes from
  - Its properties
- Estimation
  - Moments
  - Maximum likelihood
  - Robust methods
- Regression

At which level?

- Practitioners **may** have a superficial knowledge, but
  - They will miss the **fun**
  - They will lack a deep **understanding**
  - They may use techniques when they are not **valid**

Is this *Statistical Puritanism?*



Not really

Missing good models may have terrible consequences when, for instance, making classifications

Is this *Statistical Puritanism?*

Not really

Missing good models may have terrible consequences when, for instance, designing speckle filters

What does a practitioner need to know?

- Limit Theorems (CLT, LLN)
- Gamma and related models
- Operations with random variables
- Estimation techniques
- Regression (linear and generalised)
- Hypotheses tests, *p*-values
- Multivariate statistics
- ➢ All at undergraduate level

Contact

Alejandro C. Frery
School of Mathematics and Statistics
Te Herenga Waka – Victoria University of Wellington
`alejandro.frery@vuw.ac.nz`
`https://people.wgtn.ac.nz/alejandro.freryorgambide`