
Lectio Magistralis: **Science, Data and Statistics**

Anthony Davison

Galileo Galilei (1564–1642)



Galileo's telescope



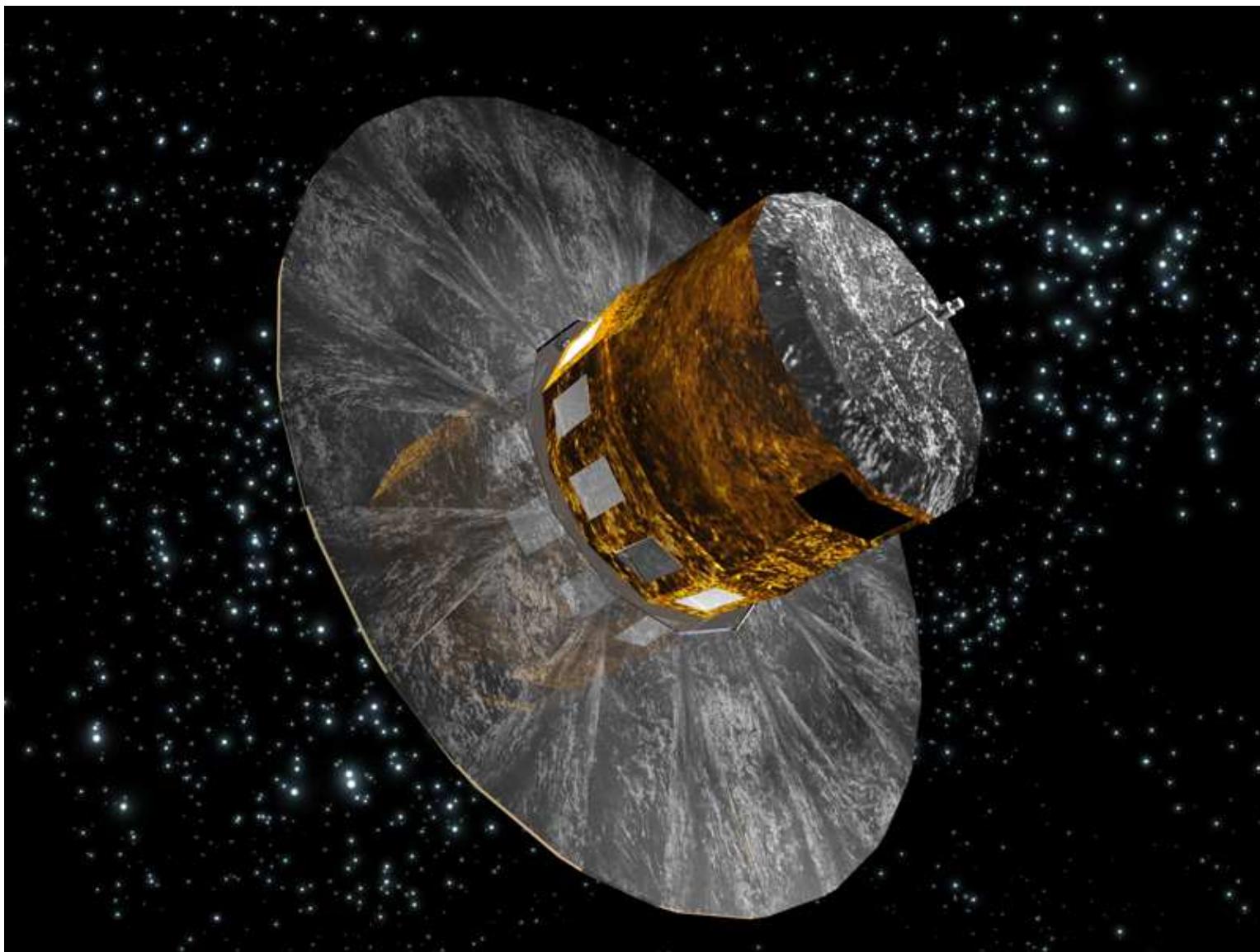
Il Saggiatore, 1623



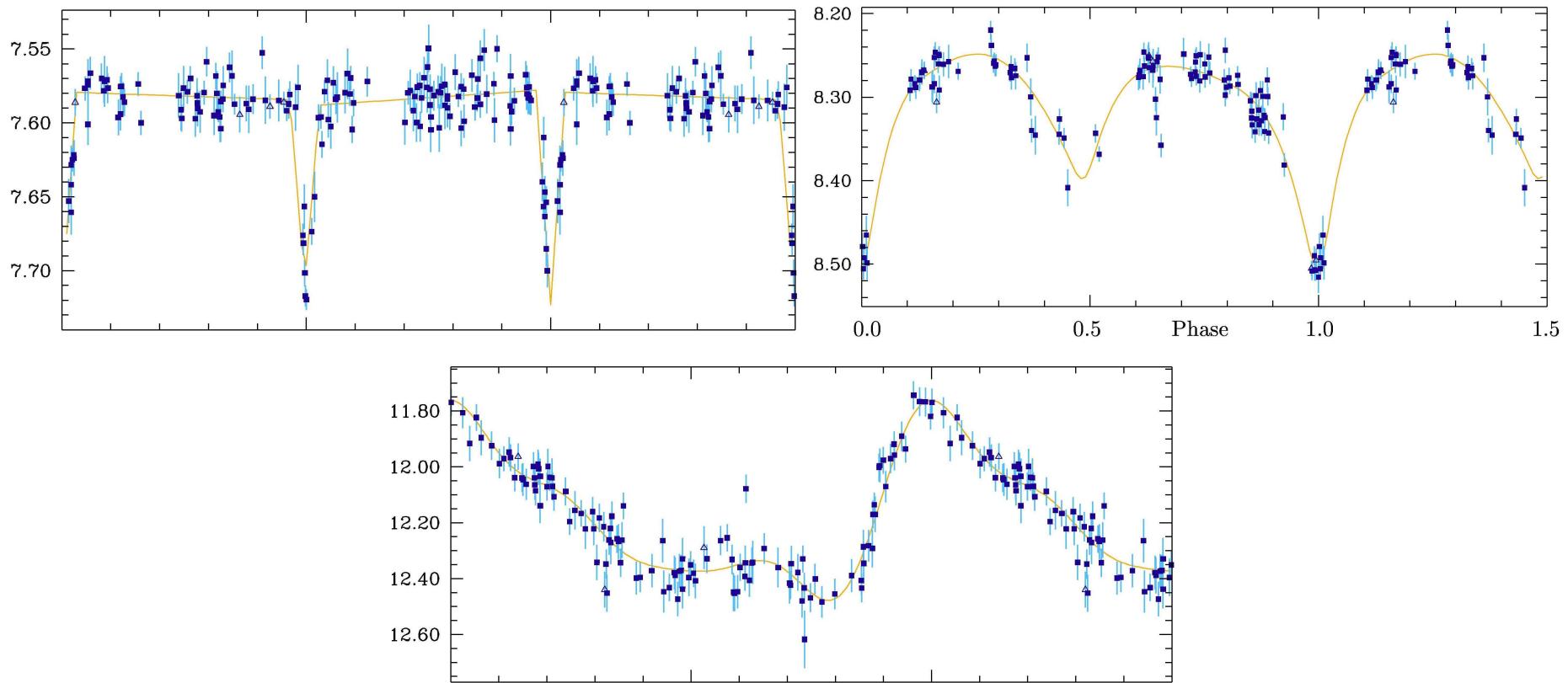
Galileo Galilei, Il Saggiatore, 1623

La filosofia è scritta in questo grandissimo libro che continuamente ci sta aperto innanzi a gli occhi (io dico l'universo), ma non si può intendere se prima non s'impara a intender la lingua, e conoscer i caratteri, ne' quali è scritto. Egli è scritto in lingua matematica, e i caratteri son triangoli, cerchi, ed altre figure geometriche, senza i quali mezzi è impossibile a intenderne umanamente parola; senza questi è un aggirarsi vanamente per un oscuro laberinto.

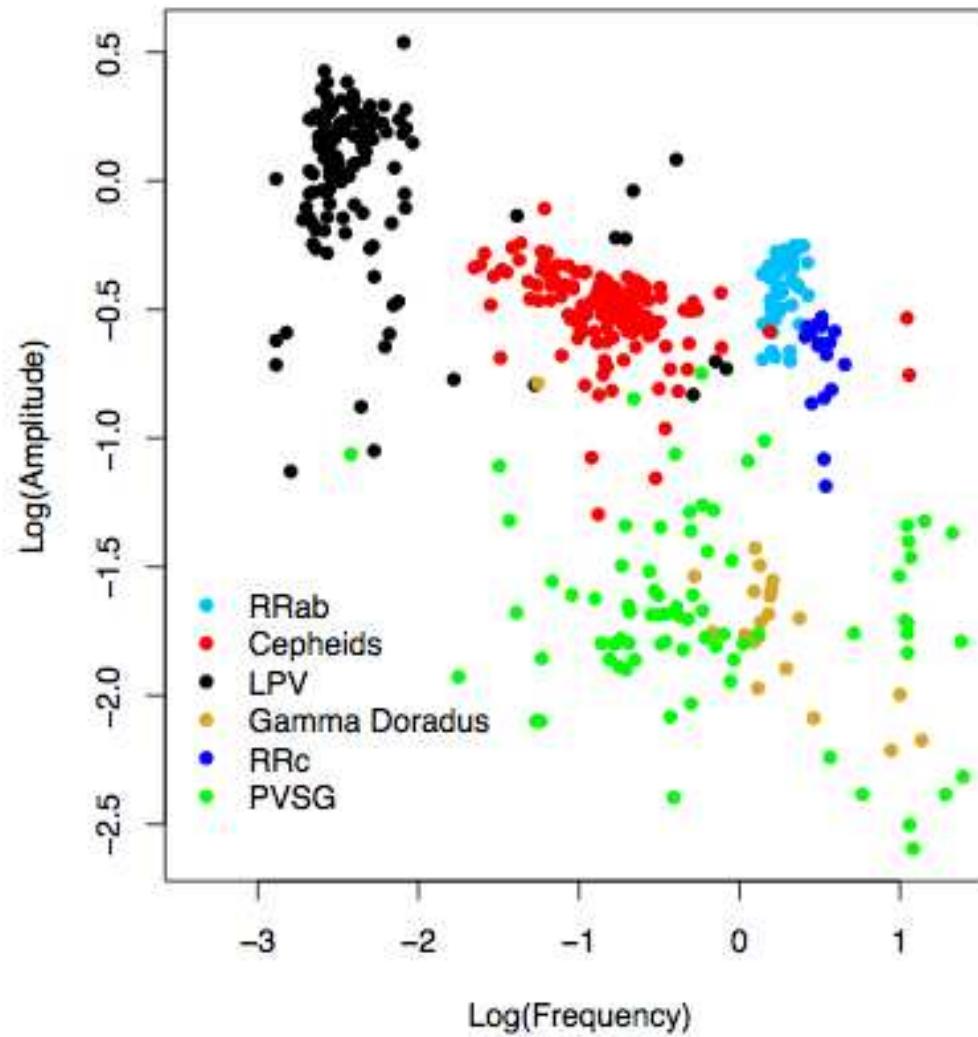
The book of the Universe cannot be understood unless one first learns to comprehend the language and to understand the alphabet in which it is composed. It is written in the language of mathematics, and its characters are triangles, circles, and other geometric figures, without which it is humanly impossible to understand a single word of it; without these, one wanders about in a dark labyrinth.



Variable star data

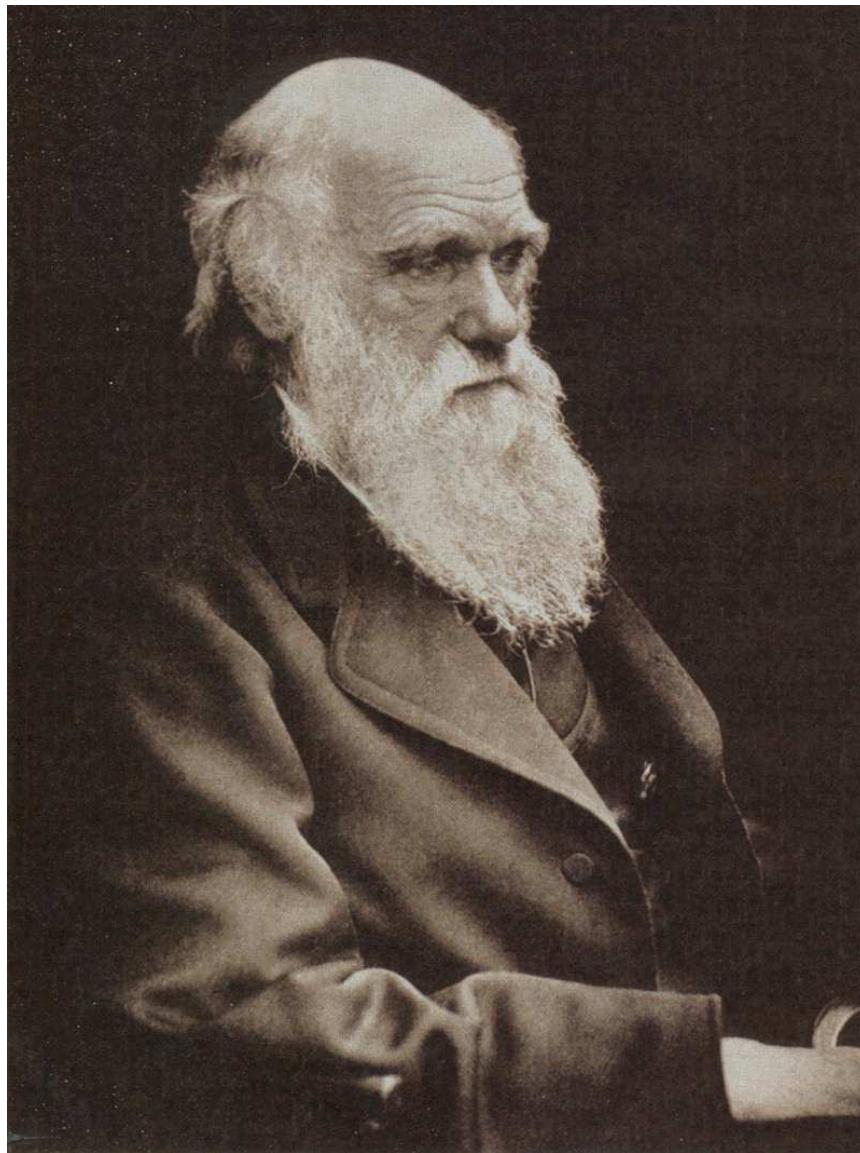


Classification

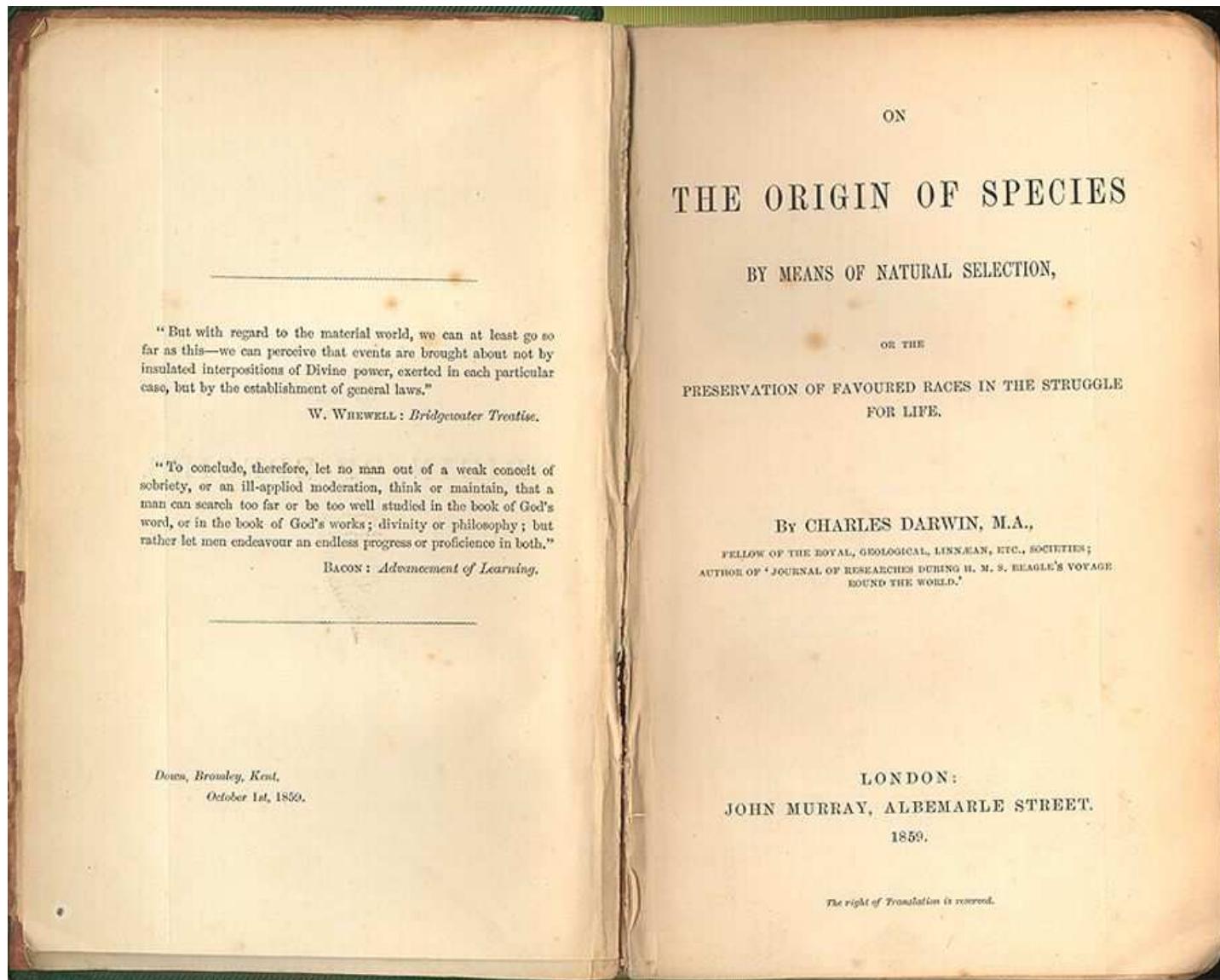


Thanks: Maria Süveges

Charles Robert Darwin (1809–1882)



The Origin of Species, 1859



Ronald Alymer Fisher (1890–1962)



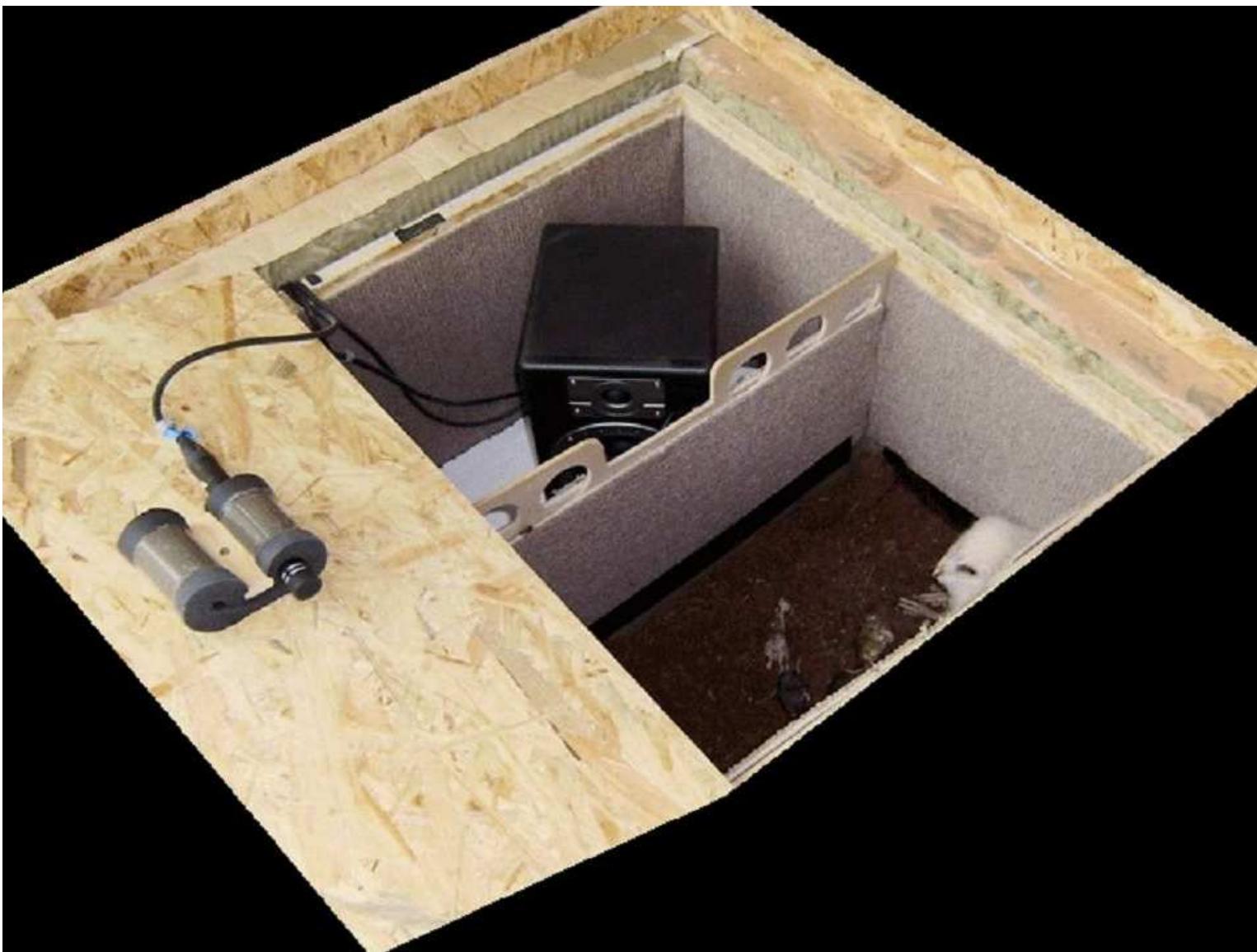
Ronald Alymer Fisher, 1938

To call in the statistician after the experiment is done may be no more than asking him to perform a postmortem examination: he may be able to say what the experiment died of.

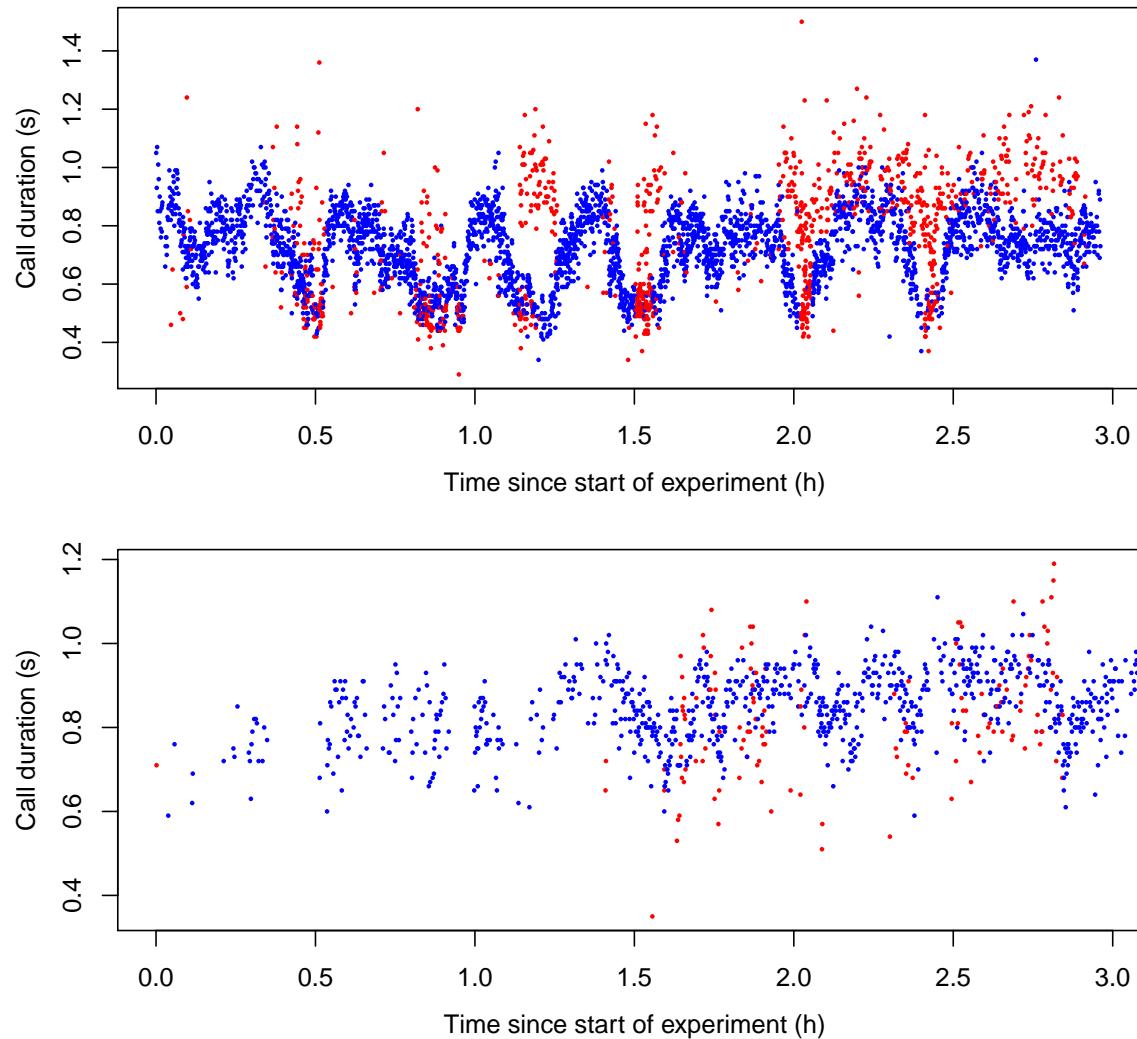
Owl nestling experiments



Listening in ...



Owl interaction data



Thanks: Amélie Dreiss, Charlène Ruppli

Chief Economist at Google:

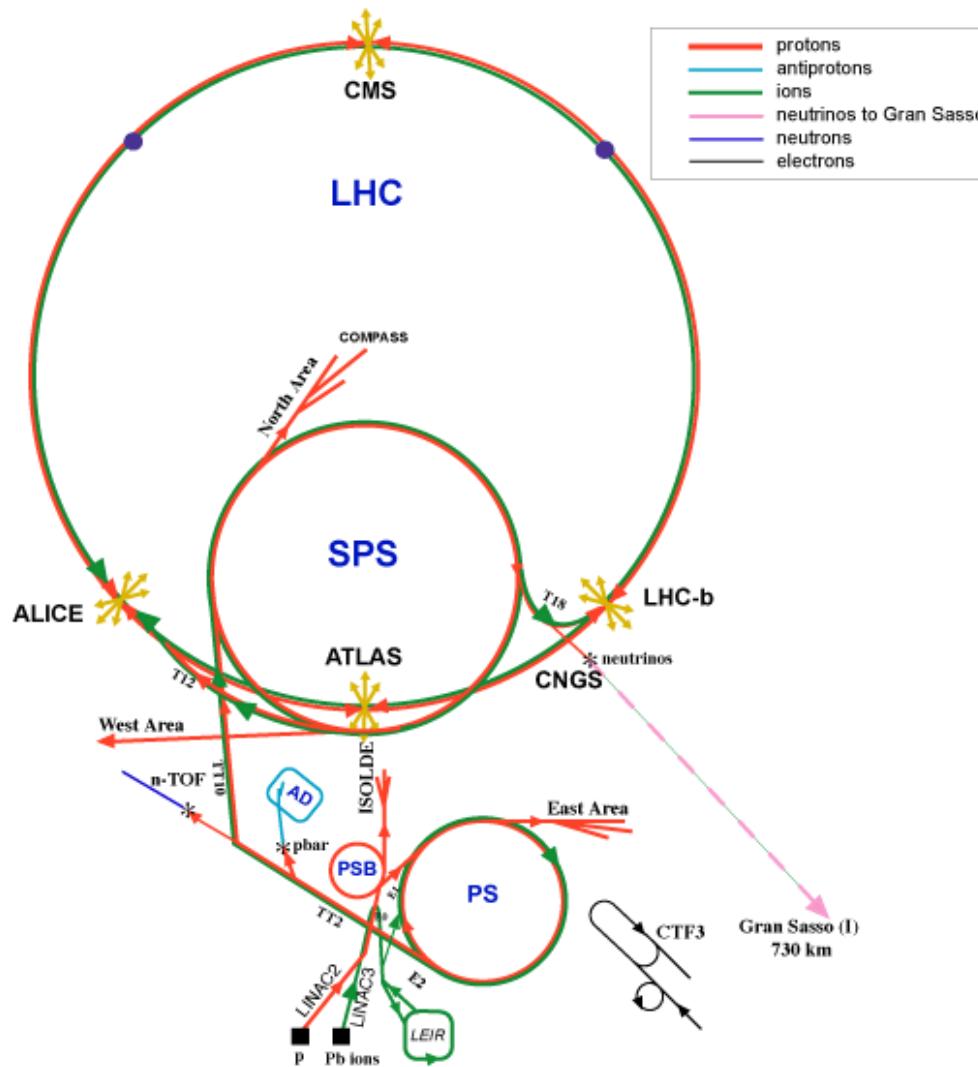
I keep saying the sexy job in the next ten years will be statisticians. People think I'm joking, but ... the ability to take data—to be able to understand it, to process it, to extract value from it, to visualize it, to communicate it—that's going to be a hugely important skill in the next decades.

Large Hadron Collider

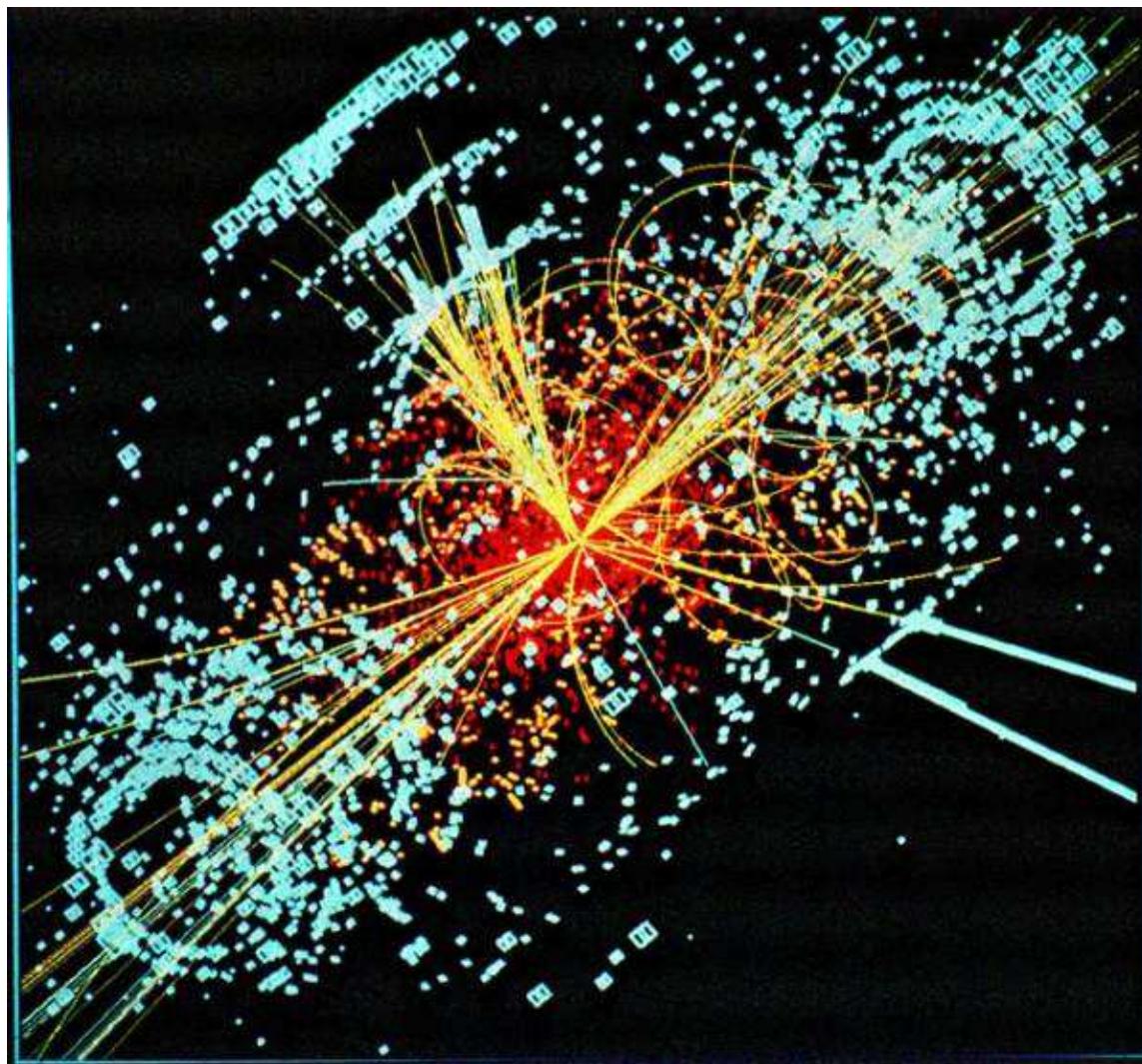


Thanks: Nicola Sartori

Experiments at CERN



Simulated Higgs event



$$\Pr[T_A < 1, T_B < 1] = \Phi_2(\Phi^{-1}(F_A(1)), \Phi^{-1}(F_B(1)), \gamma)$$

Here's what killed your 401(k) *David X. Li's Gaussian copula function as first published in 2000. Investors exploited it as a quick—and fatally flawed—way to assess risk. A shorter version appears on this month's cover of Wired.*

Probability

Specifically, this is a joint default probability—the likelihood that any two members of the pool (A and B) will both default. It's what investors are looking for, and the rest of the formula provides the answer.

Copula

This couples (hence the Latinate term copula) the individual probabilities associated with A and B to come up with a single number. Errors here massively increase the risk of the whole equation blowing up.

Survival times

The amount of time between now and when A and B can be expected to default. Li took the idea from a concept in actuarial science that charts what happens to someone's life expectancy when their spouse dies.

Distribution functions

The probabilities of how long A and B are likely to survive. Since these are not certainties, they can be dangerous: Small miscalculations may leave you facing much more risk than the formula indicates.

Equality

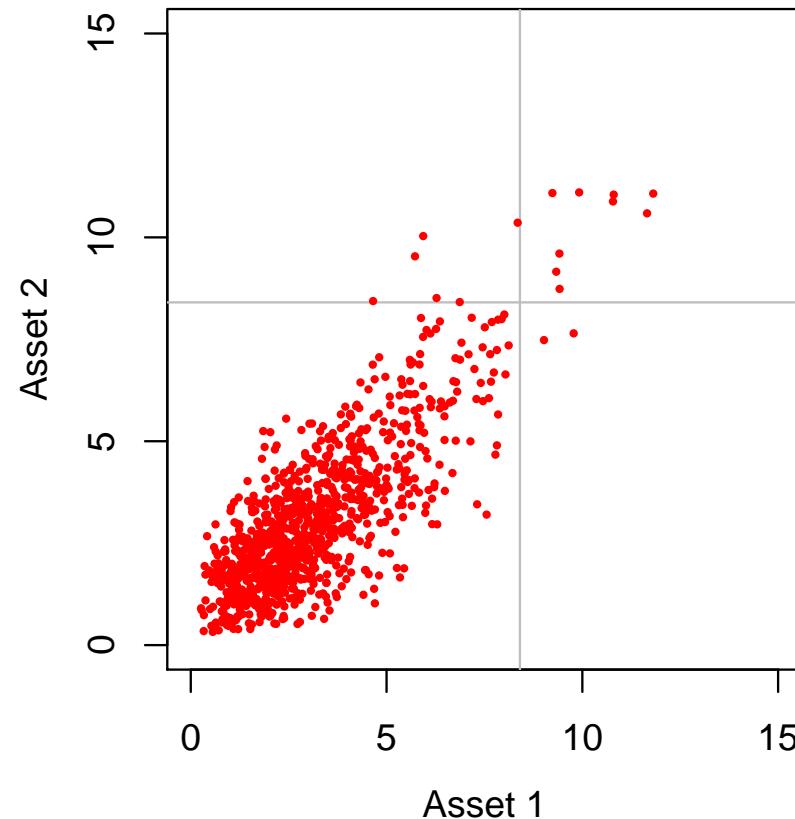
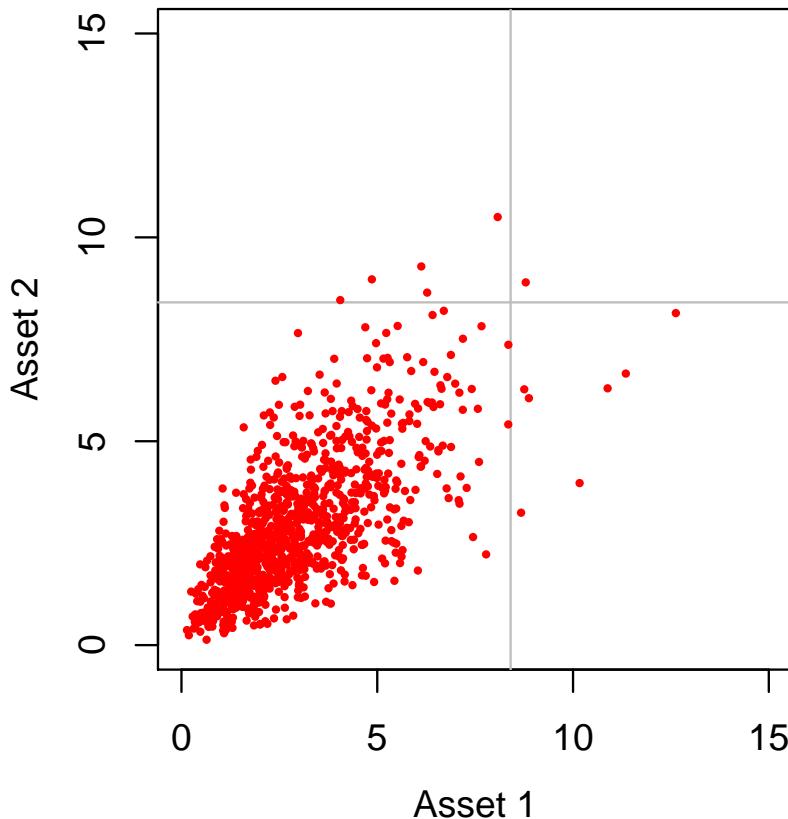
A dangerously precise concept, since it leaves no room for error. Clean equations help both quants and their managers forget that the real world contains a surprising amount of uncertainty, fuzziness, and precariousness.

Gamma

The all-powerful correlation parameter, which reduces correlation to a single constant—something that should be highly improbable, if not impossible. This is the magic number that made Li's copula function irresistible.

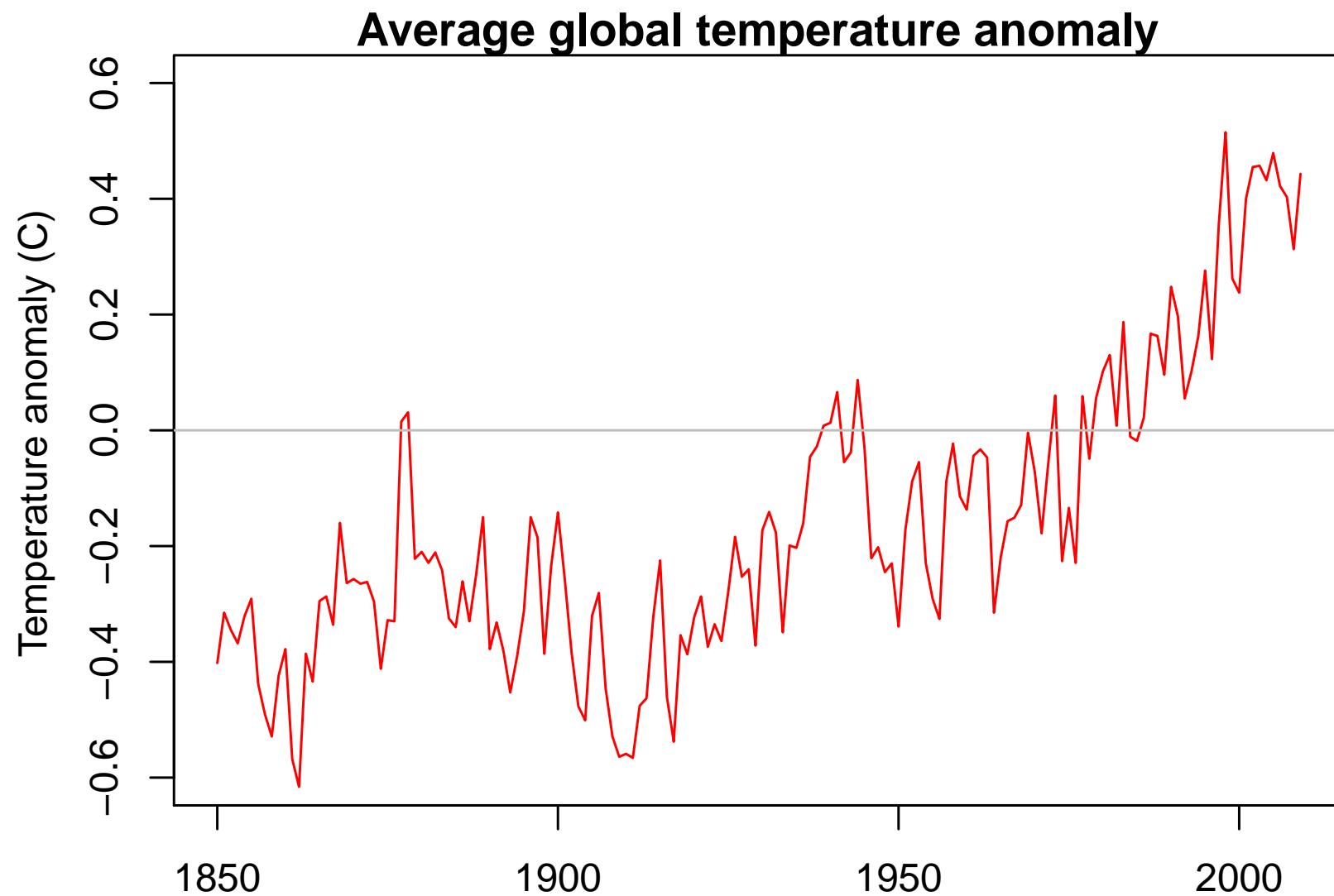
Effect of dependence model, $\rho = 0.7$

Both models have the same correlation and the same marginal distributions, but the tail dependence is much stronger on the right. The left model was used by financial institutions to model risk.

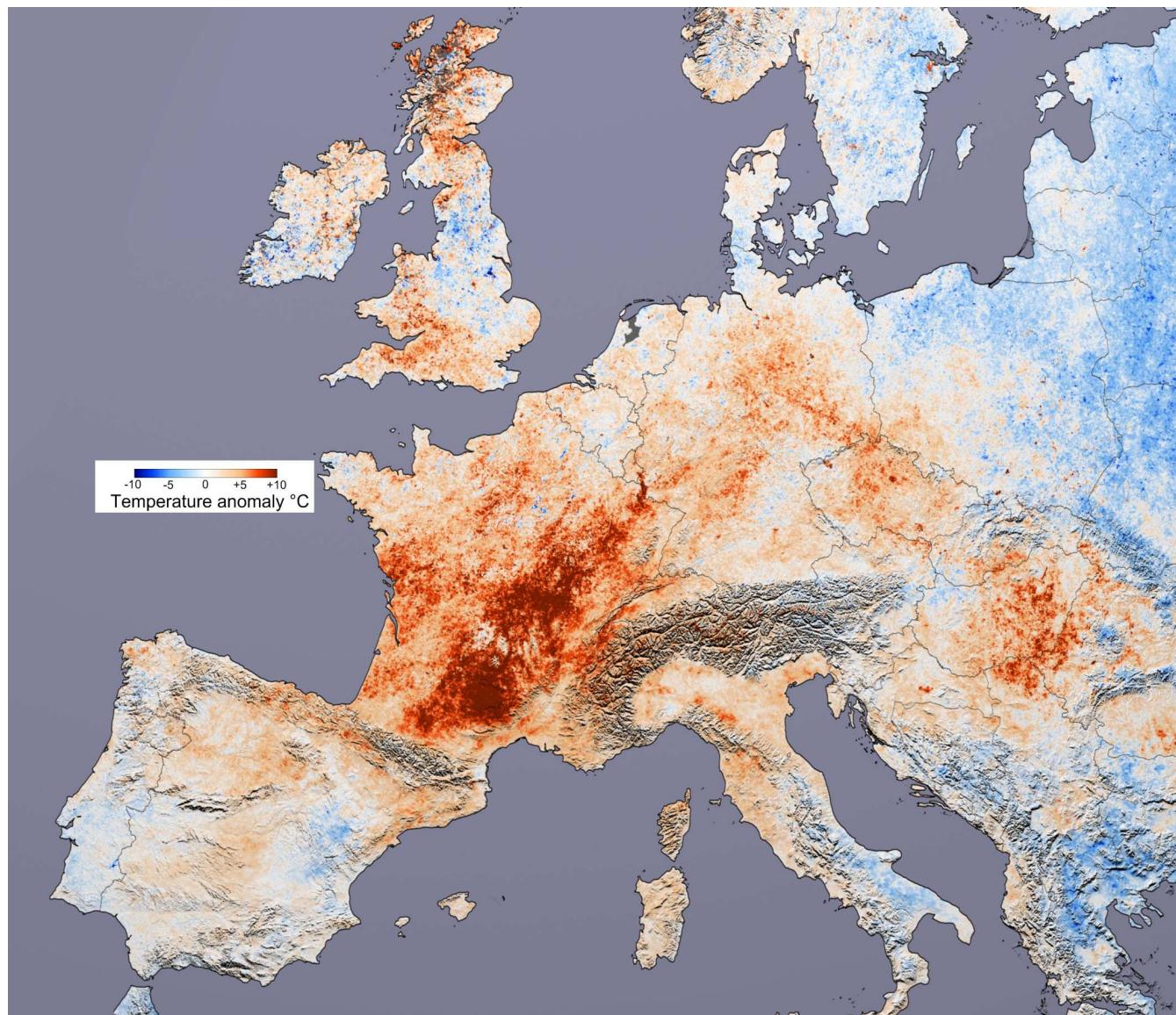


Thanks: Paul Embrechts

Global warming, 1850–2009



Europe 2003



Le ciel est tombé sur la tête des Tessinois



Malgré les routes coupées et le trafic perturbé, le Tessin a échappé au pire en fin de semaine.

vait devant sa maison à côté d'une fontaine.

Bien que moins touchée, la Suisse romande a eu droit à son lot de précipitations. A Genève, il est tombé plus de 155 mm d'eau en l'espace de trois jours. Dans le canton de Vaud, 50 mm de pluie pour la seule journée de samedi ont arrosé la région de la Dôle. Entre Saint-Prex et Villars-sous-Yens, le trafic a été perturbé par de l'eau sur la chaussée. – ATS

MALHEUREUX GOTHARD

Une coulée de boue a recouvert la ligne CFF, hier vers 12 h 40, entre Lavorgo et Pianotondo, interrompant le trafic durant tout l'après-midi. C'est la deuxième perturbation, coup sur coup, qui frappe le trafic au Gothard. La ligne avait en effet déjà été coupée durant 15 heures jeudi dernier, après la collision entre une machine de chantier de 72 tonnes et quatre wagons remplis de ballast. Cet accident était survenu dans un tunnel entre Giornico et Lavorgo. – AP

Brig, 1993



Thanks: Christophe Ancey

Thank you!