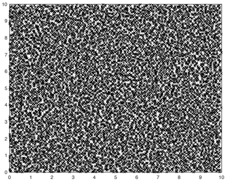
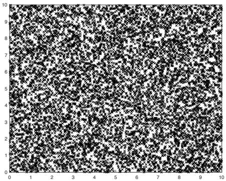
Enjoy learning data analysis, least squares, statistics, and so on.
Births and Moon cycle

Frédéric Chambat, Anne-Laure Fougères, Alexandre Elyildirim, 2021,
Comptes Rendus Mathématique (Académie des Sciences), to appear soon.

École Normale Supérieure de Lyon
Laboratoire de Géologie, Terre Planètes Environnement

Les Houches - doctoral school - 2021





Why talking about that during a doctoral school on geology ?

- This is fun
- A lot of people are interested in this question
- It's about the Moon (and tides a little bit)
- A nice example of signal analysis
- An example of generalized least squares
- An introduction to statistic tests.

Why talking about that during a doctoral school on geology ?

- This is fun
- A lot of people are interested in this question
- It's about the Moon (and tides a little bit)
- A nice example of signal analysis
- An example of generalized least squares
- An introduction to statistic tests.

Why talking about that during a doctoral school on geology ?

- This is fun
- A lot of people are interested in this question
- It's about the Moon (and tides a little bit)
- A nice example of signal analysis
- An example of generalized least squares
- An introduction to statistic tests.

Why talking about that during a doctoral school on geology ?

- This is fun
- A lot of people are interested in this question
- It's about the Moon (and tides a little bit)
- A nice example of signal analysis
- An example of generalized least squares
- An introduction to statistic tests.

Why talking about that during a doctoral school on geology ?

- This is fun
- A lot of people are interested in this question
- It's about the Moon (and tides a little bit)
- A nice example of signal analysis
- An example of generalized least squares
- An introduction to statistic tests.

Why talking about that during a doctoral school on geology ?

- This is fun
- A lot of people are interested in this question
- It's about the Moon (and tides a little bit)
- A nice example of signal analysis
- An example of generalized least squares
- An introduction to statistic tests.

A common idea is that
there is more births on full moon days.

Is this idea well founded ?

Yes... and no

Part of the room is right, but with the wrong reasons.
Part of the room is not right, but with good reasons.

May be we don't all ask the same question

- Does the moon affect births ?
- Are maternity hospitals crowded on full moon days ?
- Are there more births on full moon days ?
- Are there many more births on full moon days ?
- Are there significantly more births on full moon days ?
- Is the popular belief, that there are more births on full moon days, based on observations ?
- If there are more births on full moon days, can we see it in a maternity ward ?

May be we don't all ask the same question

- Does the moon affect births ?
- Are maternity hospitals crowded on full moon days ?
- Are there more births on full moon days ?
- Are there many more births on full moon days ?
- Are there significantly more births on full moon days ?
- Is the popular belief, that there are more births on full moon days, based on observations ?
- If there are more births on full moon days, can we see it in a maternity ward ?

May be we don't all ask the same question

- Does the moon affect births ?
- Are maternity hospitals crowded on full moon days ?
- Are there more births on full moon days ?
- Are there many more births on full moon days ?
- Are there significantly more births on full moon days ?
- Is the popular belief, that there are more births on full moon days, based on observations ?
- If there are more births on full moon days, can we see it in a maternity ward ?

May be we don't all ask the same question

- Does the moon affect births ?
- Are maternity hospitals crowded on full moon days ?
- Are there more births on full moon days ?
- Are there many more births on full moon days ?
- Are there significantly more births on full moon days ?
- Is the popular belief, that there are more births on full moon days, based on observations ?
- If there are more births on full moon days, can we see it in a maternity ward ?

May be we don't all ask the same question

- Does the moon affect births ?
- Are maternity hospitals crowded on full moon days ?
- Are there more births on full moon days ?
- Are there many more births on full moon days ?
- Are there significantly more births on full moon days ?
- Is the popular belief, that there are more births on full moon days, based on observations ?
- If there are more births on full moon days, can we see it in a maternity ward ?

May be we don't all ask the same question

- Does the moon affect births ?
- Are maternity hospitals crowded on full moon days ?
- Are there more births on full moon days ?
- Are there many more births on full moon days ?
- Are there significantly more births on full moon days ?
- Is the popular belief, that there are more births on full moon days, based on observations ?
- If there are more births on full moon days, can we see it in a maternity ward ?

May be we don't all ask the same question

- Does the moon affect births ?
- Are maternity hospitals crowded on full moon days ?
- Are there more births on full moon days ?
- Are there many more births on full moon days ?
- Are there significantly more births on full moon days ?
- Is the popular belief, that there are more births on full moon days, based on observations ?
- If there are more births on full moon days, can we see it in a maternity ward ?

Plan

- 1. How, and why, do we come to study these kinds of questions ?
- 2. Existing analyses about 'Lunar effects'
- 3. Birth data in France over 50 years
- 4. Least squares and statistical analysis

Plan

- 1. How, and why, do we come to study these kinds of questions ?
- 2. Existing analyses about 'Lunar effects'
- 3. Birth data in France over 50 years
- 4. Least squares and statistical analysis

Plan

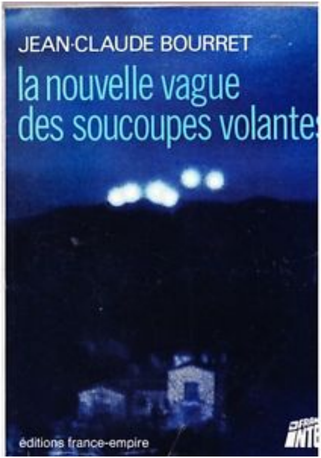
- 1. How, and why, do we come to study these kinds of questions ?
- 2. Existing analyses about 'Lunar effects'
- 3. Birth data in France over 50 years
- 4. Least squares and statistical analysis

Plan

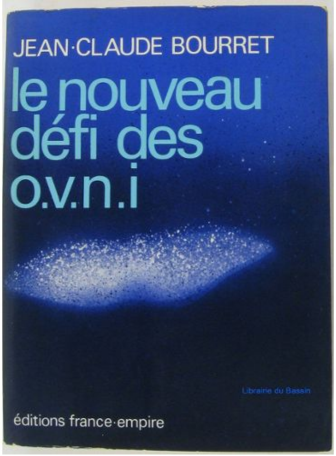
- 1. How, and why, do we come to study these kinds of questions ?
- 2. Existing analyses about 'Lunar effects'
- 3. Birth data in France over 50 years
- 4. Least squares and statistical analysis

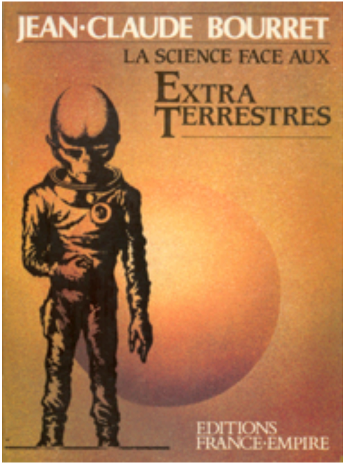
1. How, and why, do we come to study these kinds of questions ?

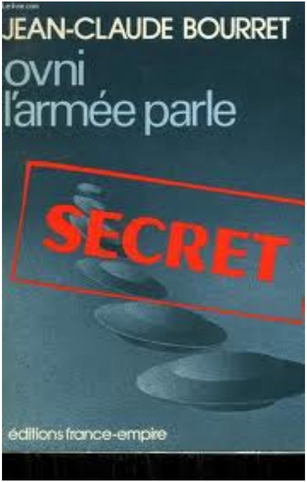
Do french people here recognize him ?

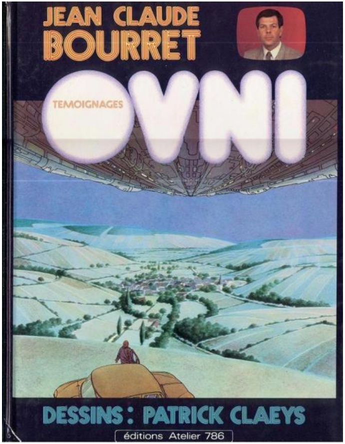


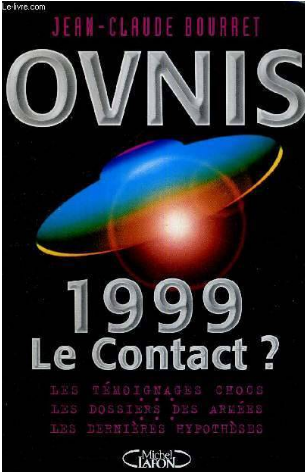
The new wave of flying saucers (Unidentified Flying Objects)

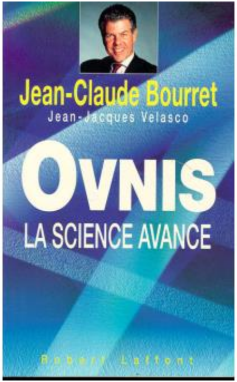

JEAN·CLAUDE BOURRET
le nouveau défi des o.v.n.i
Librairie du Bassin
éditions france·empire


JEAN-CLAUDE BOURRET
LA SCIENCE FACE AUX
EXTRA TERRESTRES
EDITIONS
FRANCE-EMPIRE

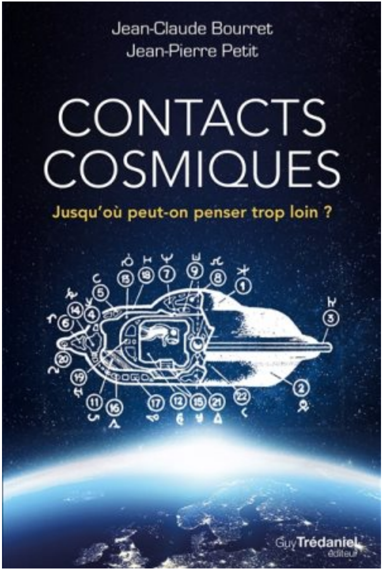

JEAN-CLAUDE BOURRET
ovni
l'armée parle
SECRET
éditions france-empire


JEAN CLAUDE
BOURRET
TEMOIGNAGES
OVNI
DESSINS : PATRICK CLAEYS
éditions Atelier 786


Le-livre.com
JEAN-CLAUDE BOURRET
OVNIS
1999
Le Contact ?
LES TÉMOIGNAGES CHOCS
LES DOSSIERS DES ARMÉES
LES DERNIÈRES HYPOTHÈSES
Michel LAFON


Jean-Claude Bourret
Jean-Jacques Velasco
OVNIS
LA SCIENCE AVANCE
Robert Laffont

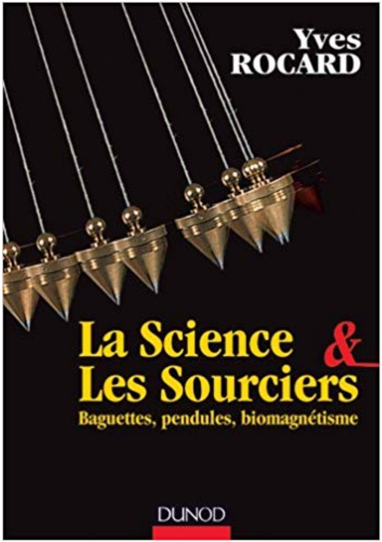

JEAN-CLAUDE BOURRET
JEAN-PIERRE PETIT
OVNI
L'EXTRAORDINAIRE
DÉCOUVERTE
GuyTrédaniel
éditeur


Jean-Claude Bourret
Jean-Pierre Petit
CONTACTS COSMIQUES
Jusqu'où peut-on penser trop loin ?
Guy Trédaniel éditeur

UFOs came on Earth mainly in the 50-70's.
In 2021, nobody seems to see them anymore...

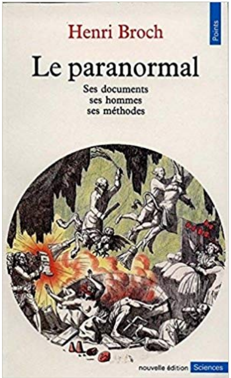
Pendulum


AMPHITHEATRE YVES ROCARD
ULSORY
ONE WAY DIRECTION



"Sourcier" = Dowser


AMPHITHEATRE Y
R
I
CARD
ONE WAY DIRECTION



Cercle zététique

(skeptical association that studies paranormal phenomena)

The prevalence of beliefs is high

- Three out of four Americans believe in at least one unfounded concept (ghosts, witches, astrology, etc.).
- Four out of ten Americans think that humans have been created in the last 10,000 years.
- 6 percent of Americans think Apollo landings on the Moon are fake. **(Margot, 2015)**
- 80 % of nurses, and 60 % of doctors, in emergency units, think the moon has an effect on patients.

(Daniel F. et col, 1986)

So we should like to find a way to study paranormal phenomena and extraordinary beliefs.

There is one belief that can be tested.
For which there is data, accurate and in large numbers.
The civil registry office and the national statistics institute (INSEE) collect birth data.

Why should we be interested in beliefs and test them ?

For brain hygiene.

« Allowing our brain to develop beliefs that contradict indisputable facts is bad brain hygiene » (Margot, 2015)

Brush your teeth for oral hygiene.
Think properly to avoid smelling bad brain.

Misconceptions have a huge cost

Health professionals have difficulty containing a measles outbreak due to questionable beliefs about vaccine safety. Yet vaccines are widely considered to be one of the greatest public health achievements, but vaccine-preventable diseases kill people because of beliefs that are not in line with scientific evidence. In 2018 in France : 2,800 cases of measles (90 % not vaccinated), including 3 deaths.

Imminent extinction of the northern white rhino (horns), mutilation of reindeer (antlers), sharks (fins), seals (penises) etc... due to superstition of their alleged aphrodisiac character
(thanks to the inventors of molecules such as Viagra, Tadalafil...)

Many species are becoming extinct or decreasing sharply, huge population displacements will occur, due to climate sceptical beliefs.

In 2021, coronavirus antivax people have a cost : some of them will spend some time in intensive care units at hospital, other will miss work, etc.

Probabilities and statistics

All this has to do with statistics :

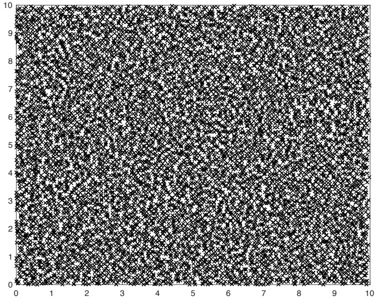
- statistics are a way to thwart illusions

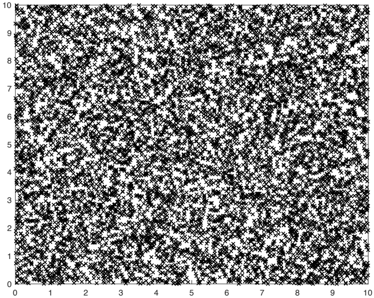
- statistics make it possible to analyse observations rationally.

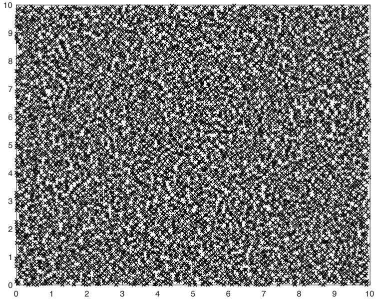
Statistics is not always intuitive

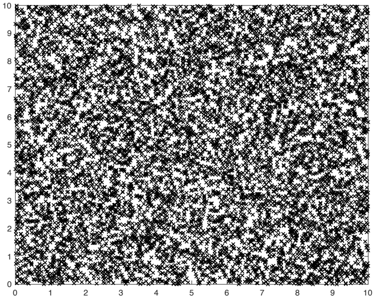
Which square is filled the most randomly ?

Left or right ?









Uniform $\neq$ Random

Apparent Miracles

« An abandoned dog 100 km away finds his home ! »

100,000 abandoned dogs and cats every year.

Let's say they are abandoned *averages* 100 km from home. And that they know the environment of their home about 100 m away, which means that if they get within 100 m of home, they are sure to find their master.
Suppose, as a skeptic, that an abandoned dog goes in a totally random direction ! And that 1 in 10 dogs will wander at least 100 km.
So, the expected number of dogs returning home each year is

$$(100,000/10) \times (2 \times 0.1/100)/(2\pi) = 3.$$

This return makes the headlines in the local newspapers.

However, we can see that they are not so improbable (in France and over a year).

Correlation and causality

60 % of deaths occur in hospitals or clinics,
25 % at home,
13 % in a retirement home.
2 % elsewhere.

Thus, hospital kills.

Correlation and causality

Suppose we find that :
heavy coffee drinkers have more cancers than the average population.
What should we conclude ?

Many studies have found a positive association between coffee and cigarette consumption.

An epidemiological study found an extremely strong correlation ($r = 0.99$) between number of cigarettes and the number of cups of coffee consumed. This correlation is not found with the tea.

38 % of smokers of both sexes are heavy consumers of coffee, compared to 12 % and 16 % for non-smokers, depending on gender.

« It must therefore be concluded that there are factors in coffee that stimulate smoking consumption, and that it is not caffeine, which would rather be tendency to limit it. »

http ://www.tabac-humain.com/wp-content/uploads/2011/07/Tabac-et-Cafe.pdf

2. Existing studies about 'Lunar effects'

Meta-analysis and reviews

The Moon Was Full and Nothing Happened:
A Review of Studies on the Moon and Human Behavior and Human Belief

by I.W. KELLY, JAMES ROTTON, and ROGER CULVER

Psychological Bulletin
1985, Vol. 97, No. 2, 286–306

0033-2909/85/$00.75

Much Ado About the Full Moon:
A Meta-Analysis of Lunar-Lunacy Research

James Rotton
Florida International University

I. W. Kelly
University of Saskatchewan
Saskatoon, Saskatchewan, Canada

Data from 37 published and unpublished studies were combined in a meta-analysis that examined relations between phase of the moon, type of lunar cycle, sex, publication practices, geographical features (latitude, population density), and several types of lunacy, including mental hospital admissions, psychiatric disturbances, crisis calls, homicides, and other criminal offenses. Although a few statistically significant relations emerged, effect size estimates indicated that phases of the moon accounted for no more than 1% of the variance in activities usually termed lunacy. Alleged relations between phases of the moon and behavior can be traced to inappropriate analyses, a failure to take other (e.g., weekly) cycles into account, and a willingness to accept any departure from chance as evidence for a *lunar effect*.

Current Biology *18*, R784–R794, September 9, 2008 ©2008 Elsevier Ltd All rights reserved DOI 10.1016/j.cub.2008.07.003

Human Responses to the Geophysical Daily, Annual and Lunar Cycles — **Review**

Russell G. Foster[1] and Till Roenneberg[2]

an internal ~24 hour circadian timer in anticipation of the

Table 1. Pathologies and some physiological and behavioural phenomena in humans which have been linked to the lunar cycle but which in reality occur independently of the phase of the moon.

Condition/event with no consistent lunar influence	Reference
Psychosis, depression, anxiety	[67–75]
Violent behaviour/aggression	[76–80]
Seizures	[81]
Suicide	[79,82–87]
Absenteeism rates	[88]
Coronary failure	[89]
Conception (*in vitro* fertilization)	[90]
Birth	[91–93]
Menstruation	[94]
Surgery and survival of breast cancer	[95]
Postoperative outcome (general)	[96–98]
Renal colic	[99]
Outpatient admissions (general)	[100]
Automobile accidents	[101,102]

For example, human copulation frequencies and lunar rhythms

Diurnal and Weekly, But No Lunar Rhythms in Human Copulation

By John D. Palmer,[1] J. Richard Udry[2] and Naomi M. Morris[3]

ABSTRACT

Using the copulation times from 78 White married women, representing an average of 72 reporting days per woman over a 12-month period, two expected prominent copulatory rhythms were found and described quantitatively. The daily rhythm is characterized by a major evening peak, which encompasses 58% of the daily copulations, and a minor peak in the morning. The weekly rhythm describes a rather constant copulatory rate during weekdays, with a large increase on weekends. The probability of orgasm was found to be significantly greater in copulations taking place after midday. Because of the long-time speculation that the moon plays some role in the timing of human love-making, the data were examined for such a relationship. None was found.

Human Biology, 1982.

1230 copulations, 48 couples, 3rd sept. - 2nd dec. 1973
711 copulations, 30 couples, 8th juin - 5th sept. 1974

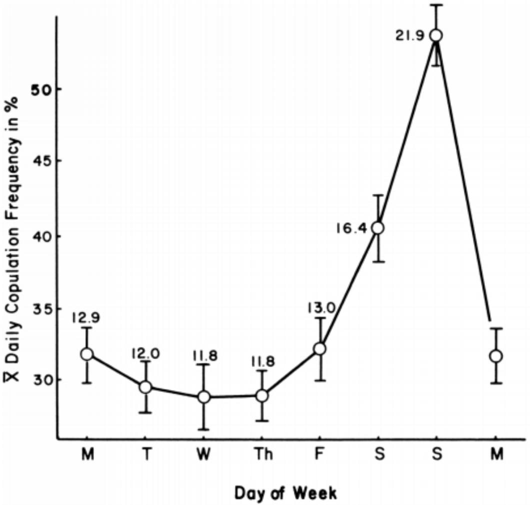


FIG. 1. The mean weekly copulation rhythm of 78 women. The curve is based on 1,941 copulatory events. Standard errors indicate the relative significance of the daily means. The Sunday rate differs significantly from that of Saturday (P = <0.001), and the Saturday rate from any weekday (P = <0.01). The numbers displayed above each mean represent the day's percentage of the week's total number of copulations.

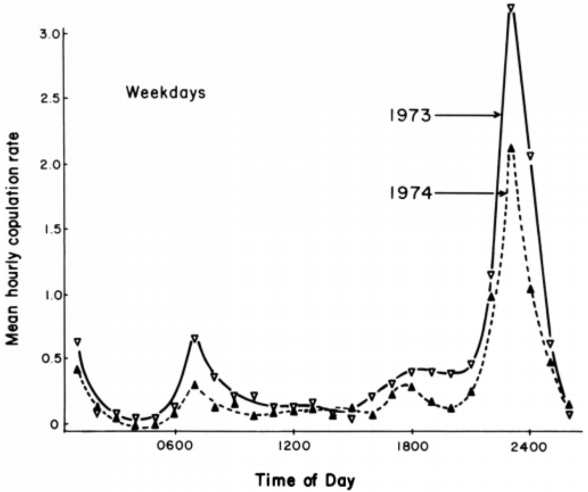
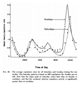


FIG. 2A. The average solar-day pattern for weekdays during three-month studies in 1973 and 1974. The 1973 curve is based on 744 copulations; the 1974 one on 451. The major peaks in both years encompass slightly over 60 percent of the daily copulations. A second, minor peak occurs at 7 AM.



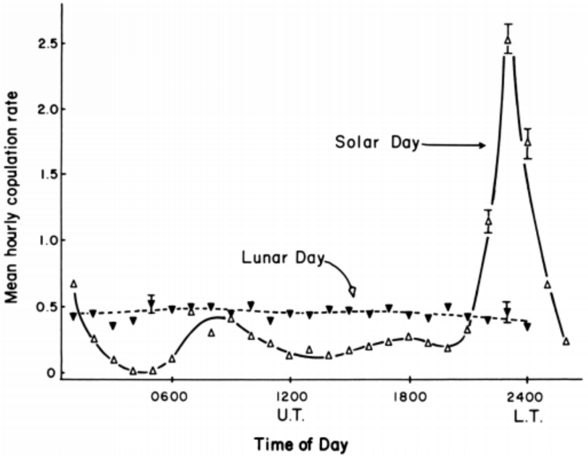


FIG. 3. The mean solar- and lunar-day patterns of copulation based on all the data collected in both studies. A total of 5,584 "copulating-couple days," and 1,941 copulations were used in the construction of each curve. Representative standard errors are given on some points to indicate relative significance, or lack thereof, between points. UT = upper transit of the moon; LT = lower transit. The solar-day pattern emphasizes the presence of a population rhythm, characterized by a prominent peak between 10 PM and 1 AM which encompasses 56 per cent of the daily copulations. There is no indication of a lunar-day rhythm.

(1 'lunar day' = 24,8 h)

In a further analysis, the number of copulations which took place while the moon was above the horizon (i.e., the intervals between moonrise and moonset) were compared to the number when the moon was below the horizon (the times between moonset and moonrise). The copulation numbers were found to be 7.6% greater under the former condition, but the difference was not statistically significant ($p = 0.2$).

Next, copulation numbers as a function of the four major phases of the moon were examined. The average rate was found to be 14.6% greater for the 5-day intervals centered on full moon, than for the same intervals centered on the nights of first quarter of the moon (the two extremes in the average-monthly pattern). The difference was not significant ($p = \sim 0.1$). It should be stressed that the large-amplitude weekly rhythm present in the data, would be expected to override any relatively small correlation that might exist between copulation rates and the moon's phases.

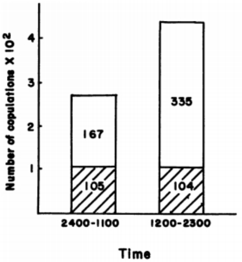


FIG. 4. AM and PM comparison of orgasmic success during copulation for 30 women over 711 copulations. The fraction of the copulations in which orgasm was not attained is indicated by the shaded portion of the base of each bar.

In the 1974 study, the women recorded whether they had experienced orgasm during intercourse. As a group, they failed to reach orgasm in 29.4% of their encounters. The mean-daily curve of orgasm achievement was, unsurprisingly, quite similar in appearance to that for copulation, but showed important non-uniform differences in amplitude: greater orgasmic success was attained after midday. As seen in Fig. 4, between the hours of midnight and 11 AM, 38.6% of all copulation failed to result in orgasm for the women; but between noon and midnight, the rate improved significantly ($X^2 = 17.9$; $p = <0.001$), the failure rate fell to 23.7%.

Studies about births and lunar cycle

First method : Fourier spectrum

Naissances, fertilité, rythmes et cycle lunaire

Etude statistique sur 5 927 978 naissances

P. Guillon, D. Guillon, J. Lansac *, J.H. Soutoul *

Avec la collaboration de P. Bertrand et J.P. Hornecker **

* *Département de Clinique Gynécologique, Obstétricale et de la Reproduction, Clinique Universitaire, 23, avenue de Roubaix, F 37100 Tours.*
** *Département de Mathématiques, Faculté des Sciences d'Orléans.*

RÉSUMÉ

Existe-t-il une liaison entre la survenue des naissances et le cycle lunaire synodique ? Certains travaux ont déjà établi l'existence d'une telle liaison. Nous avons procédé à l'étude de 5 927 978 naissances françaises comprises entre le 1er janvier 1968 et le 31 décembre 1974. Par analyse spectrale de Fourier, il est mis en évidence deux rythmes dans la survenue des naissances :
— un rythme hebdomadaire caractérisé par un nombre minimum de nouveau-nés le dimanche et maximum le mardi;
— un rythme annuel, avec un maximum de naissances en mai et un minimum en septembre-octobre.

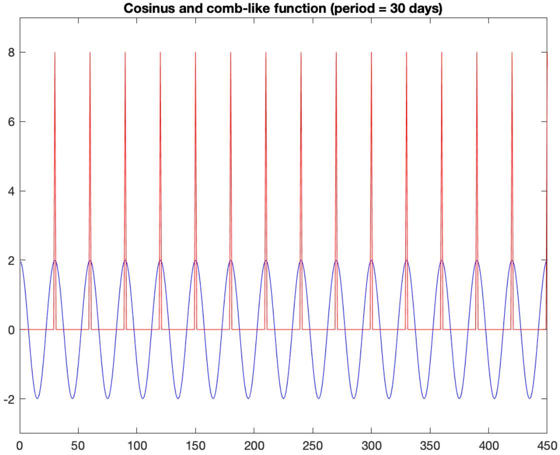
L'étude statistique de la répartition des nouveau-nés dans le cycle lunaire révèle un plus grand nombre de naissances entre le Dernier Quartier et la Nouvelle Lune et un moins grand nombre autour du Premier Quartier. Les écarts entre la distribution observée dans la lunaison et la distribution théorique sont significatifs.

No pic at 29.53 days in the Fourier spectrum.

(J. Gynecol. Obstet. Biol. Reprod., 1986)

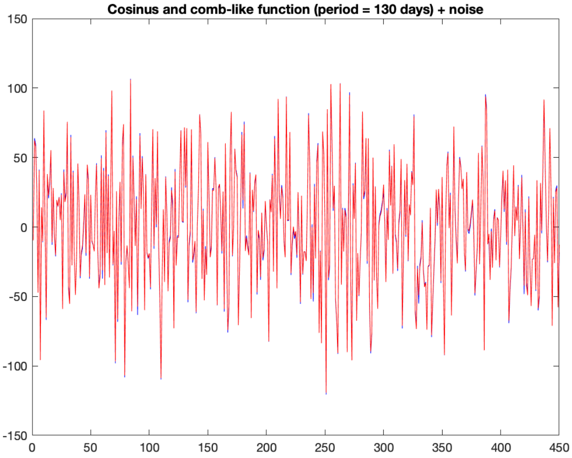
Moon synodic month = 29.53 days
(time delay between two Full Moons)

The Fourier spectrum is not always the best way to detect a periodicity



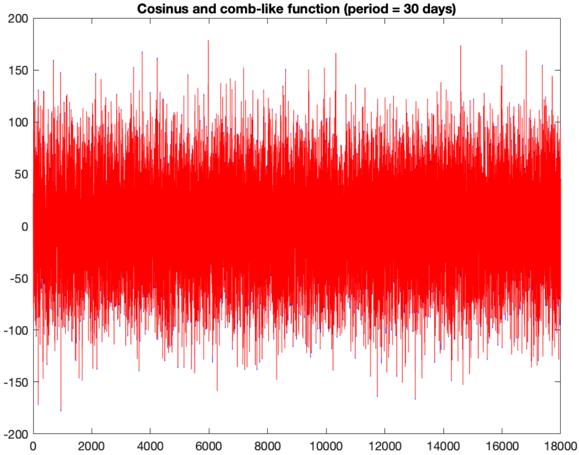
A sinusoid and a comb-like signal, with a period of one month (30 days)

I add a gaussian (and independent) noise :



Sinusoid + noise (blue) and a comb-like signal + noise (red)
(the same noise on both)

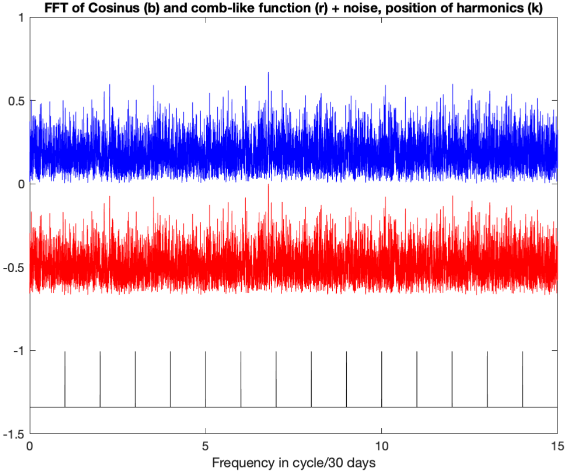
I will not show it entirely every time, but I consider a long signal :



Sinusoid + noise (blue) and a comb-like signal + noise (red)
(the same noise on both)

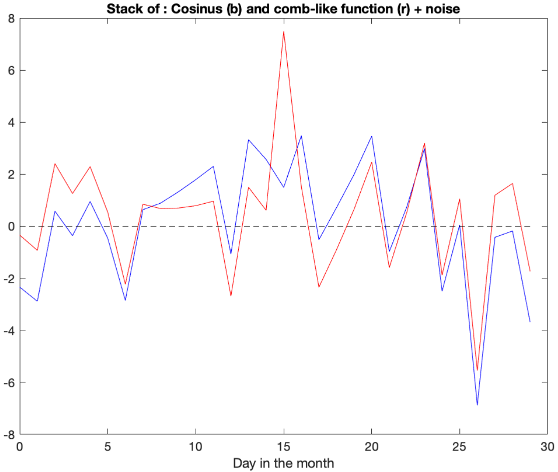
Which signal appears the best on the Fourier transform ?

Fourier spectrum :



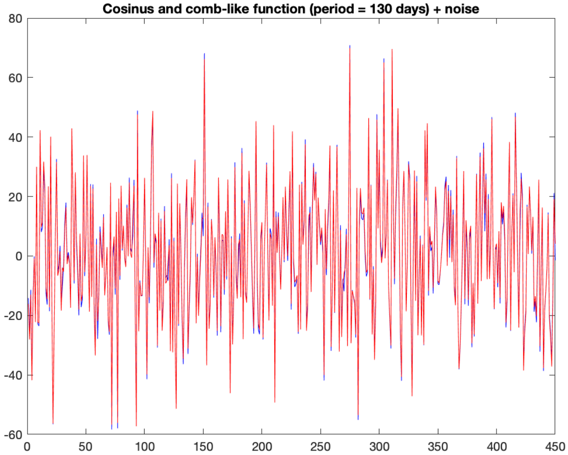
Fourier transform of sinusoid + noise (blue) and a comb-like + noise (red)

In that case, it is better to stack all the months :



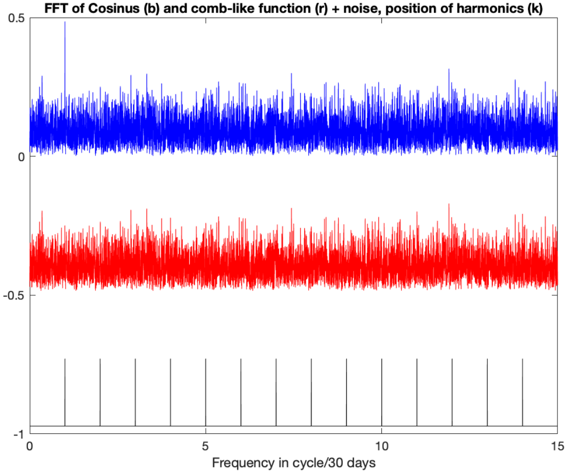
Stack of sinusoid (blue) and a comb-like signal + noise (red)

With a smallest noise :



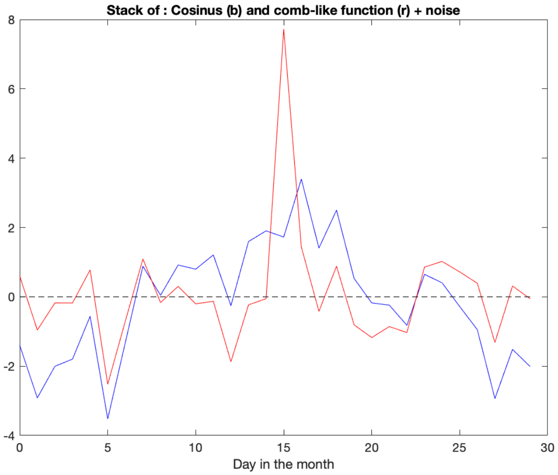
Sinusoid + noise (blue) and a comb-like + noise (red)

The sinus is visible in the Fourier spectrum, but not the comb-like signal (Why ?)



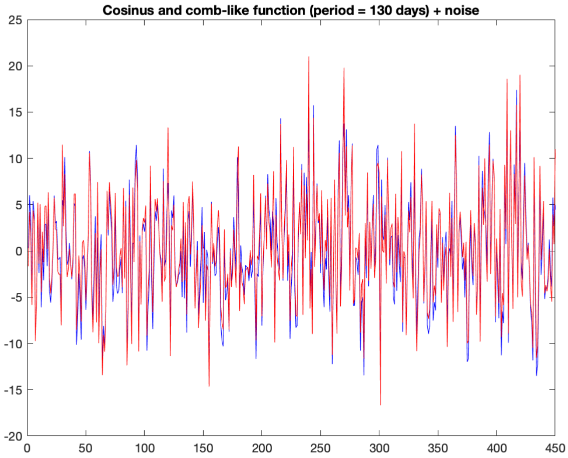
In case of moderate signal/noise ratio, we should do a statistical test to say what is the probability that a given pic could be in the spectrum by chance (i.e. with a random signal).

Stack :

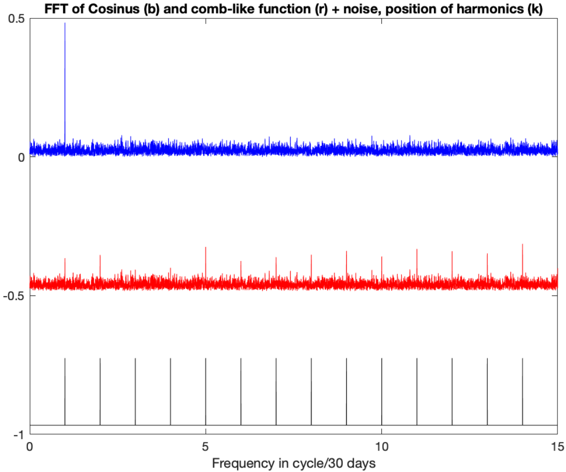


Stack of : Cosinus (b) and comb-like function (r) + noise
8
6
4
2
0
-2
-4
0
5
10
15
20
25
30
Day in the month

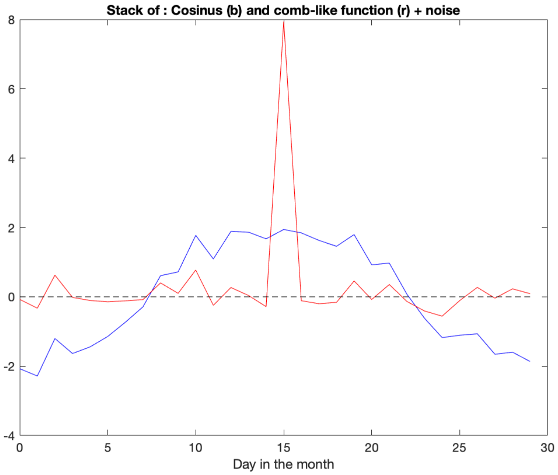
With an even smallest noise :



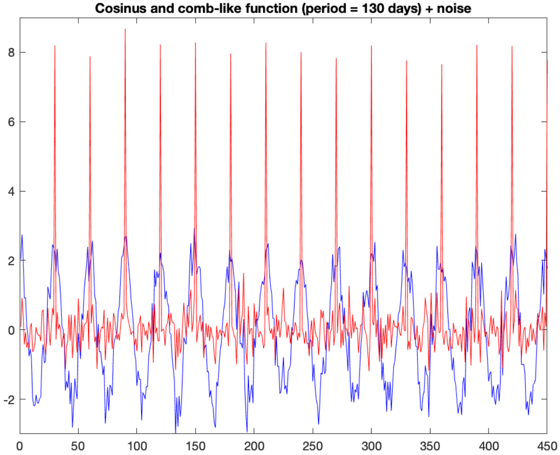
Fourier spectrum :



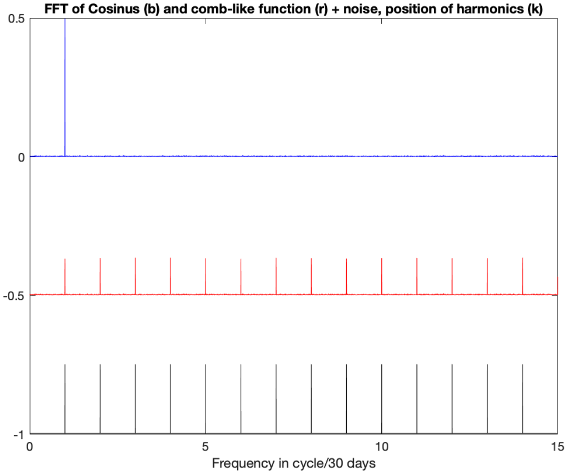
Stack :


Stack of : Cosinus (b) and comb-like function (r) + noise
8
6
4
2
0
-2
-4
0
5
10
15
20
25
30
Day in the month

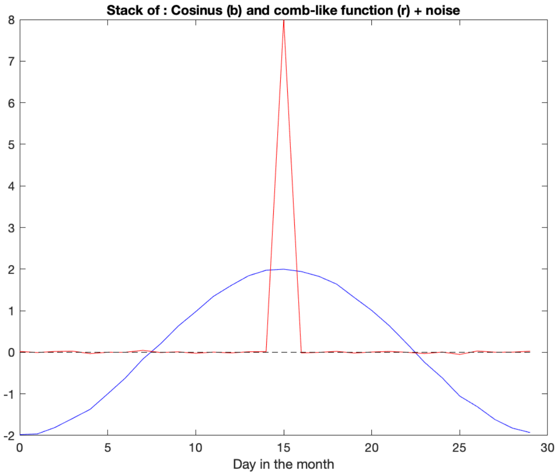
With a very small noise



Fourier spectrum :



Stack



Stack of : Cosinus (b) and comb-like function (r) + noise
8
7
6
5
4
3
2
1
0
-1
-2
0
5
10
15
20
25
30
Day in the month

Second method : stack (on the supposed period)

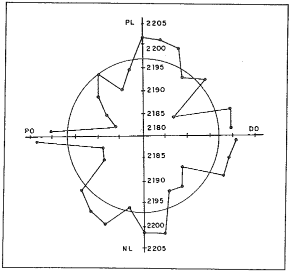


Fig. 4. — Répartition des naissances dans le cycle lunaire synodique. Chaque point de la courbe indique le nombre moyen d'enfants nés au cours d'une journée du cycle. La moyenne, qui est de 2 196,7 enfants par jour, est représentée par le cercle.

Stacking on a month shows 'significant variations'.

(Guillon et al., Rev. Gynecol. Obstét., 1988)

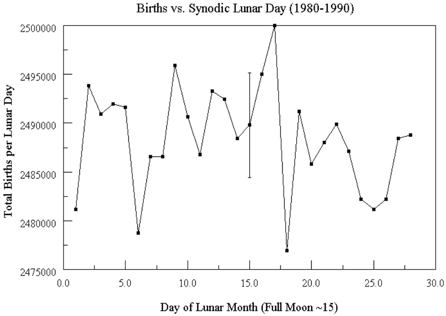
But ! with a χ^2 square test :

$$\chi^2 \propto \frac{\text{Variance(Fluctuations in the lunar month)}}{\text{Variance(Multinomial Law)}}$$

Hypothesis = the random variables (= each birth) are independent and identically distributed (iid) (= 'strong white noise'). Wrong ! χ^2 square test inadequate here.

Third method : stack and do Student, Fischer, ANOVA, and related tests

The plot below shows the result for all 20 years, about 70 million births. The standard deviation for this set of 28 points is about 5360 and is shown as an error bar at the data point for day 15.



Most of the points lie with one standard deviation of the mean, and all but one lie within two sigmas.

Caton D.B., 2002, Personal communication .

$$\text{Test statistic} \propto \frac{\text{Variance(Fluctuations in the lunar month)}}{\text{Variance(Fluctuations in the lunar days)}}$$

Hypothesis = the random variables (= number of births/days) are independent and Gaussian. A little bit wrong on raw data, might be true after having 'cleaned' the data.

Economics and Human Biology 11 (2013) 545–550

Contents lists available at SciVerse ScienceDirect

Economics and Human Biology

journal homepage: http://www.elsevier.com/locate/ehb

The lunar cycle, sunspots and the frequency of births in Germany, 1920–1989

Thomas K. Bauer [a,b,c], Stefan Bender [d], Jörg Heining [d,*], Christoph M. Schmidt [a,b,c]

[a] Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI Essen), Hohenzollernstr. 1-3, 45128 Essen, Germany
[b] Ruhr-Universität Bochum, Germany
[c] IZA Bonn, Germany
[d] Institut für Arbeitsmarkt- und Berufsforschung (IAB) der Bundesagentur für Arbeit (BA), Forschungsdatenzentrum, Regensburger Str. 104, 90478 Nürnberg, Germany

ARTICLE INFO

Article history:
Received 10 April 2012
Received in revised form 22 November 2012
Accepted 24 November 2012
Available online 3 December 2012

JEL classification:
I12
J13

Keywords:
Daily frequency of births
Lunar cycle
Sunspots

ABSTRACT

Based on multivariate linear regression models, we analyze the effect of the lunar cycle and the number of sunspots occurring on a particular day on the number of births using social security data and controlling for a number of other potential confounders. The daily numbers of births between 1920 and 1989 have been calculated from the full sample of individuals who have been registered at least once in the German social security system. While the lunar cycle does not affect the number of births, the number of sunspots is positively correlated to the number of births. The empirical results may be explained by medical technological progress making natural influences on births less important over time. This interpretation is supported by the results on the intertemporal influence of weekends and holidays on the frequency of daily births.

61.4 million births in Germany, period 1920-1989.
No link between the number of births and the lunar phases.

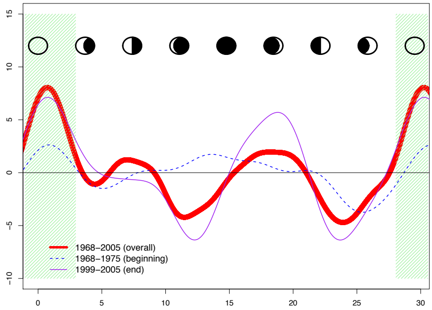


Figure 6: Average value for the mean detrended daily number of births along the moon cycle (in days).

Charpentier, 2008, Charpentier et Causeur, 2009 <hal-00482743>.

30 M births, from 1968 to 2005.
Article in 2008 : more birth between FM-1 et FM+3 : p-value $= 0,02$.
Article in 2009 : more birth (+9), between FM-3 et FM+2, on 8 % about 469 lunar cycles.

3. Birth data in France during 50 years

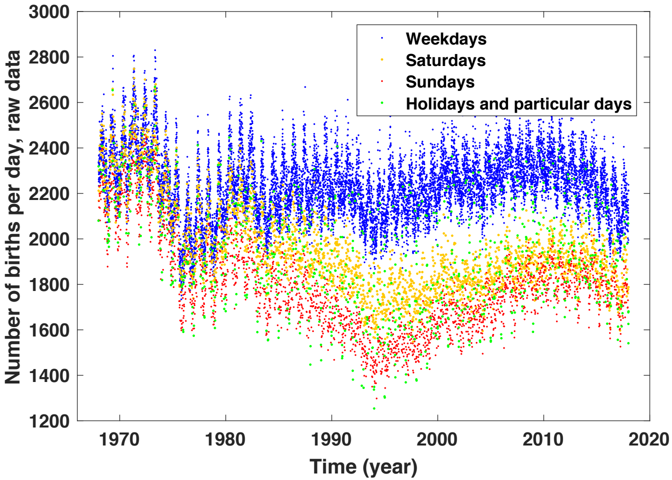
Data

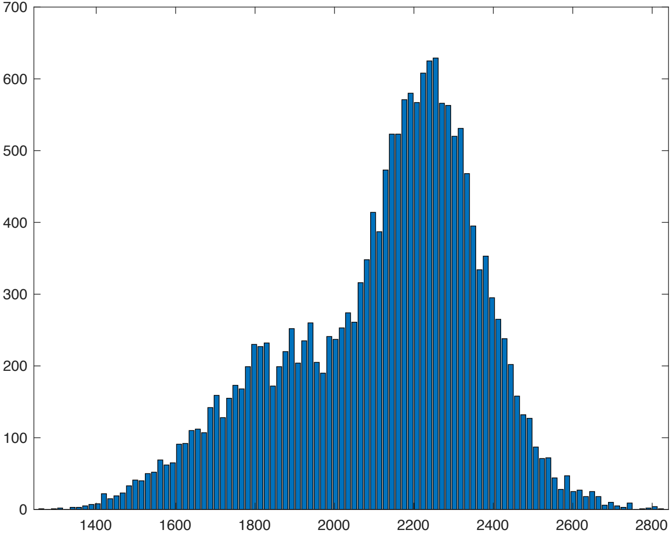
Number of births per day in metropolitan France

Total number without distinctions :
- natural triggers/activated triggers (20 %) /without trigger ('déclenchement')

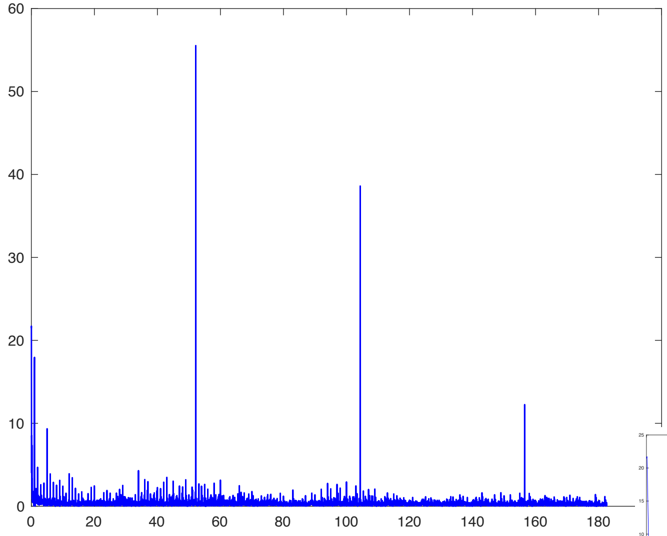
nor
- natural/medicalized/caesarean delivery.

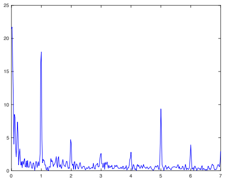
1968-2017
total = 38 741 555 births,
average = 2121 births per day,
18263 days



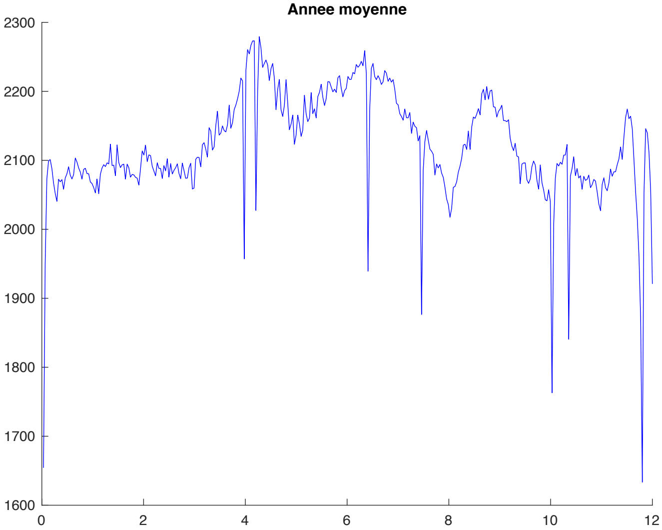


Histogram of data

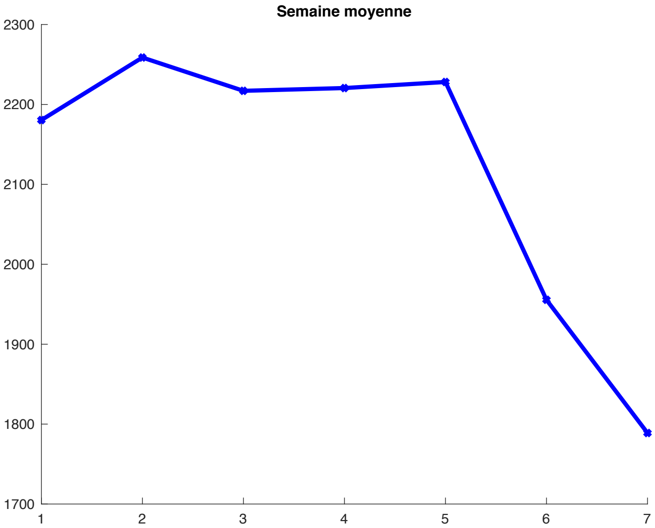




Fourier transform
Frequency in cycle/year

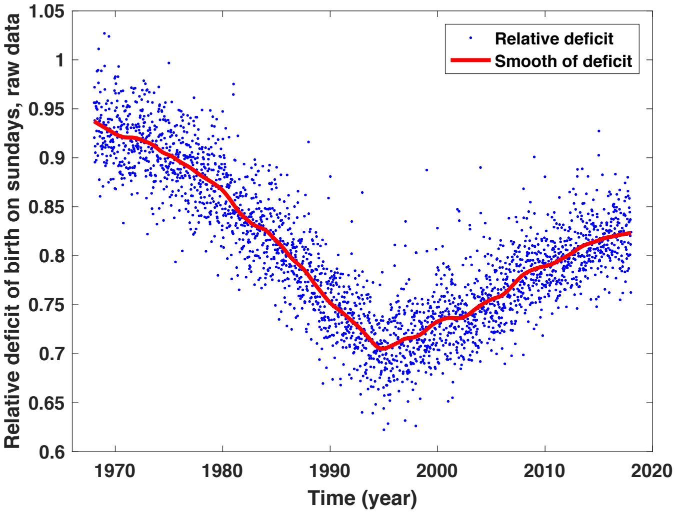

Annee moyenne
2300
2200
2100
2000
1900
1800
1700
1600
0
2
4
6
8
10
12

Mean year



Mean week

Ratio of births on Sundays to births on weekdays



And the Moon ?

The tidal force is extremely low

The inclination and intensity of this force are very low.

— o —

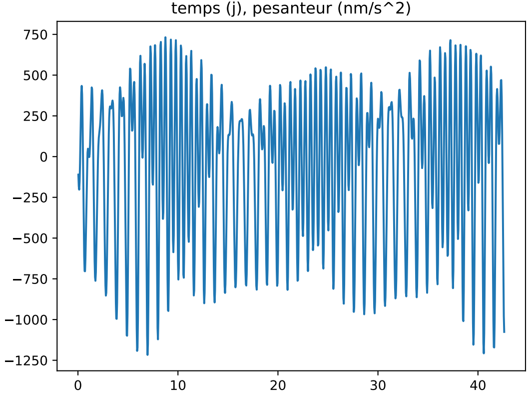
The difference in water level is
≈ 1 m over 2000 km = angle of 5.10^{-7} rad = 0.00003
(and 10 times less on the continents).
Whether the inclination of the tidal force had an effect on the initiation of childbirth, so to have a baby, all you have to do is tilt your head.

— o —

The acceleration we experience due to the lunar tidal force is $= 10^{-7}g$.
The acceleration of an elevator is calibrated to $10^{-1}g$, sometimes more.

Whether the intensity of the tidal force had an effect on the initiation of childbirth, so to have a baby, all you have to do is take the elevator.

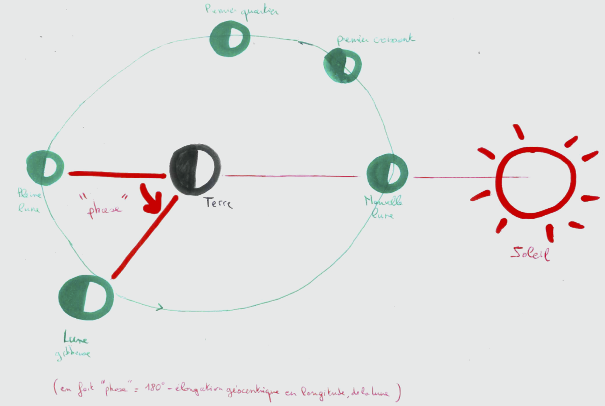
If the attraction of the Moon and the Sun influenced the number of births, then that number should have a period of 12 h 15



Gravity measurements over time

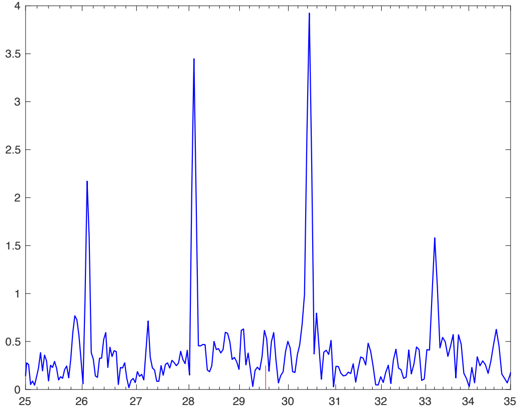
Main tide period : 12 h 15 (and 24 h 30). + Modulation of 0.5 $\times$ 29.53 days.

And the Moon phases ?



Synodic lunar month = 29.53 days (variations $< \pm$ 6h) .

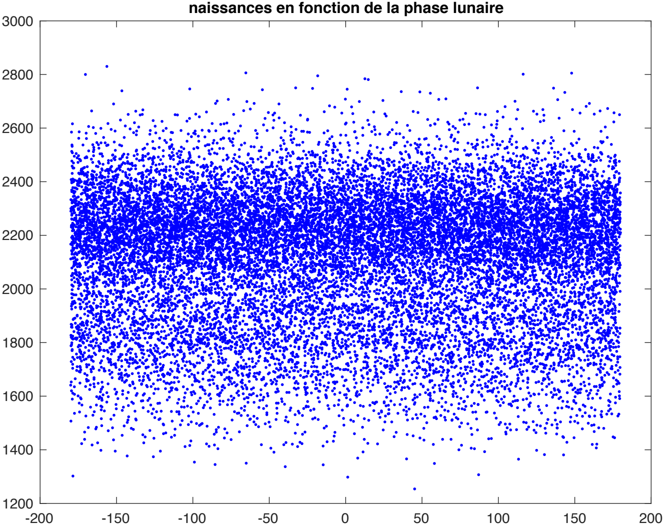
Fourier transform



Fourier Transform around 29.53 days

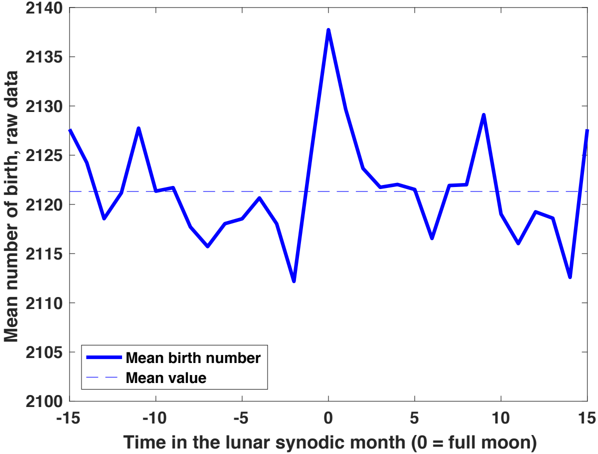
Main problem : low signal/noise, and FFT not the most efficient in detecting noisy non-sinusoidal signals

If there is an effect it is not obvious at all



Birth number as a function of the moon phase angle

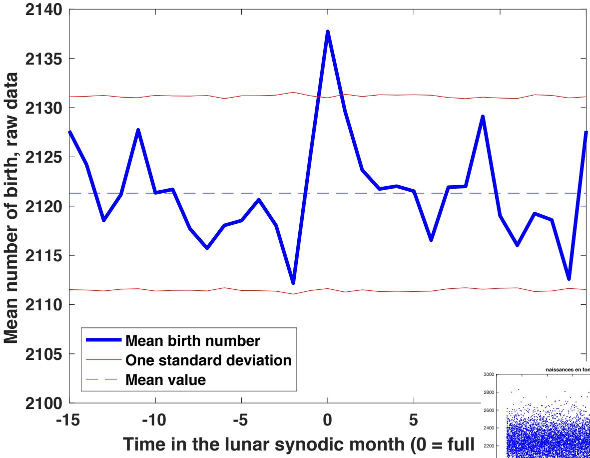
Average birth number every 12° (≈ 1 day)

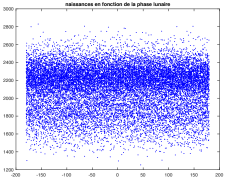


There are a little more births on full moon days, but :

- it's very low (16 births out of 2121, or +0.8 %).
- can this be due to chance ?
- or an interaction with weekly and annual fluctuations ?
- no effect of moonlight (sinus + excess in 0)

1st problem. It is necessary to compare with statistical fluctuations (= standard deviation with a statistical test)





What is the probability that the observed statistics of +16 births on FM could have been produced by chance, given the fluctuations of ± 240 births/day in the data ?

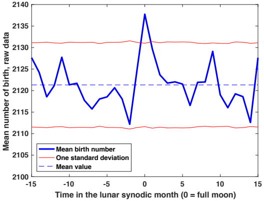
This probability is called a p-value.

The smallest is the p-value, the smallest is the probability of being wrong if you say 'my data have not been produced by the random process'.

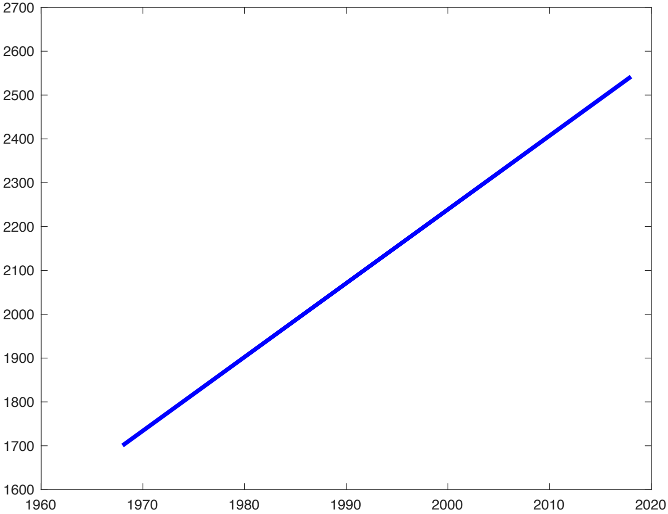
Calculus : take numbers fluctuating with a standard deviation of ± 240 births/day, distribute them in 30 days randomly, do it a large number of times, and determine the proportion of times the full moon is $> +16$ births. More simple : you know the law of your random variable, you calculate its cumulative distribution function, and then its value for the computed statistic (here 16 -> 0,07).

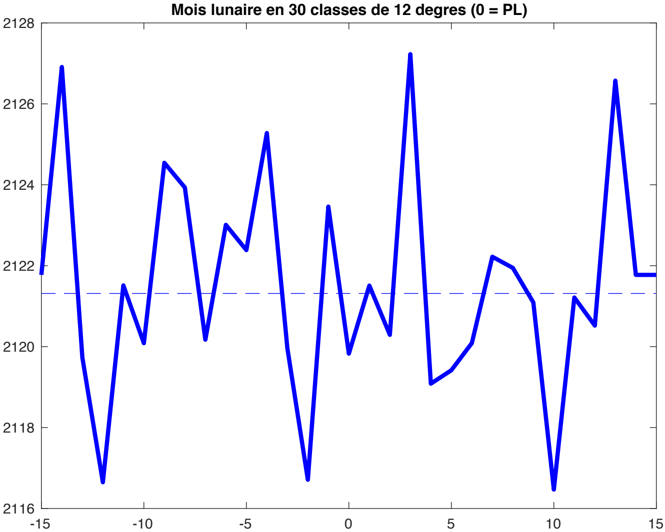
You find : p-value(FM) $= 0.07$

You can also determine the probability that one of the 30 days has $> +16$ births : global p-value $= 1 - (1 - 0.07)^{30} = 0.9$



2nd problem. An ultra-simple deterministic model of births :





± 7 births

Deterministic fluctuations (non random fluctuations) do not respect the conditions of statistical tests.

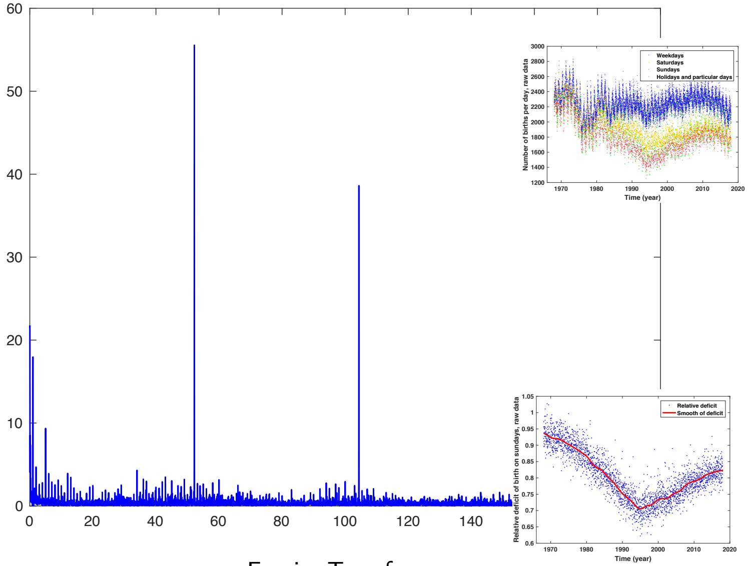
It is therefore necessary to remove deterministic fluctuations.

4. Statistical analysis

It is therefore necessary to :

- remove deterministic fluctuations : raw data $\longrightarrow$ residuals

- perform a statistical analysis of the residuals : residuals $\longrightarrow$ statistical tests



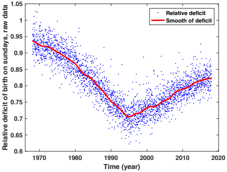
Fourier Transform
Frequency in cycle/year

Least squares

$$n(t) = \sum_{\omega_k} \left(a_k(t)\cos(\omega_k t) + b_k(t)\sin(\omega_k t) \right) + T(t) + (1+K(t)) \sum_{p=1}^{P} \alpha_p I_p(t) + \epsilon(t)$$

with :

- $\omega_k = k \times \frac{2\pi}{7 \text{ days}} \quad k \in \mathbb{N}$
- $\omega_{k'} = k' \times \frac{2\pi}{365{,}25 \text{ days}} \quad k' \in \mathbb{N}$
- $\omega_{k''} = \omega_k \pm \omega_{k'}$
- $T(t)$ is the long-term tendency ,
- $\begin{cases} I_p(t) = 1 & \text{if } t \text{ is the particular day } p, \\ I_p(t) = 0 & \text{otherwise,} \end{cases}$
- α_p is the corresponding coefficient.
- $K(t)$ long-term modulation of particular days.
- $r(t)$ are the (Gaussian) residuals.



Least squares

$$n(t) = \sum_{\omega_k} \left(a_k(t) \cos(\omega_k t) + b_k(t) \sin(\omega_k t) \right) + T(t) + (1 + K(t)) \sum_{p=1}^{P} \alpha_p \, l_p(t) + r(t)$$

$$n = g(m) + r$$

$$n \in \mathbb{R}^{18,263} = \text{data (observations)}$$

$$m \in \mathbb{R}^{12.3\ \text{million}} = \text{model (parameters)}$$

$$r \in \mathbb{R}^{18,263} = \text{residuals supposed random gaussian}$$

Tarantola, A., Valette, B., Rev. Geophys. Space Phys. 20 (1982).

$$n = g(m) + r$$

Find m that minimises :

$$||r||^2 = r^T C_n^{-1} r = (n - g(m))^T C_n^{-1} (n - g(m))?$$

Underdetermined !

Better find the smallest m too :

$$||r||^2 + ||m||^2 = r^T C_n^{-1} r + m^T C_m^{-1} m = (n - g(m))^T C_n^{-1} (n - g(m)) + m^T C_m^{-1} m$$

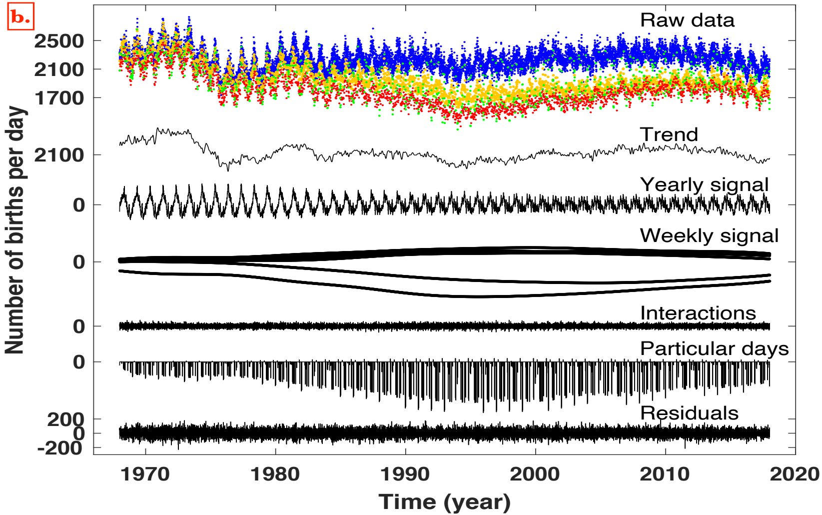
We know how to find a minimum. For example with Newton iterative scheme on the derivative :

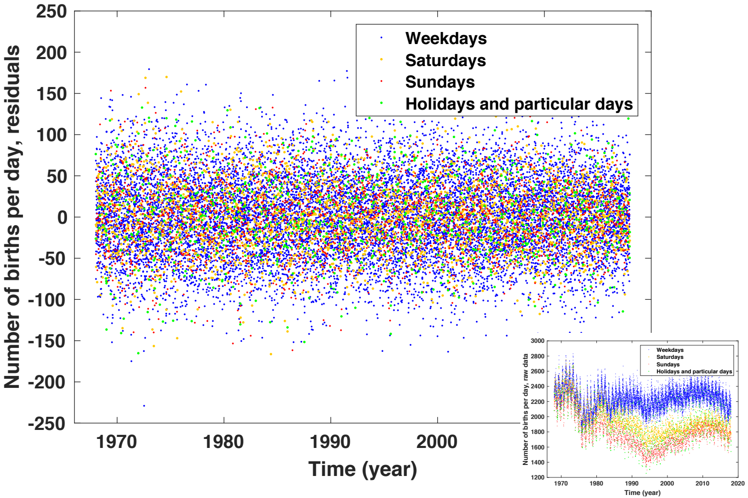
$$\hat{m}_{k+1} = C_m \, G_k^T \, \left(C_n + G_k \, C_m \, G_k^T \right)^{-1} \left(n - g(\hat{m}_k) + G_k \, \hat{m}_k \right)$$

with $G_k = \frac{\partial g}{\partial m}(\hat{m}_k)$.

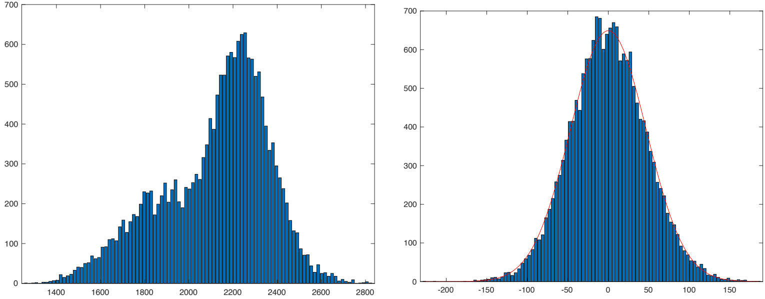
Decomposition of the signal

$$n(t) = \sum_{\omega_k} \left(a_k(t) \cos(\omega_k t) + b_k(t) \sin(\omega_k t) \right) + T(t) + (1+K(t)) \sum_{p=1}^{P} \alpha_p I_p(t) + r(t)$$

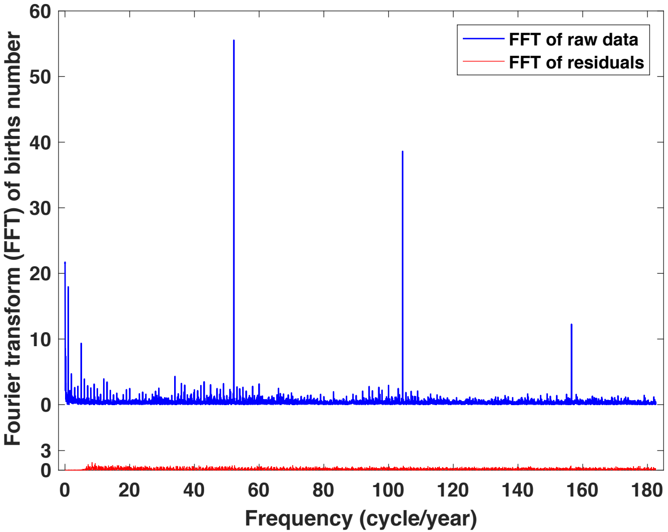




Weekdays
Saturdays
Sundays
Holidays and particular days
Number of births per day, residuals
Time (year)
Number of births per day, raw data



Histogramms of raw data and of residuals


FFT of raw data
FFT of residuals
Fourier transform (FFT) of births number
Frequency (cycle/year)
60
50
40
30
20
10
0
3
0
0
20
40
60
80
100
120
140
160
180

Particular days in the year ?

$$n(t) = \sum_{\omega_k} \big(a_k(t) \cos(\omega_k t) + b_k(t) \sin(\omega_k t) \big) + T(t) + (1 + K(t)) \sum_{p=1}^{P} \alpha_p \, I_P(t) + r(t)$$

Particular days and holidays in the year

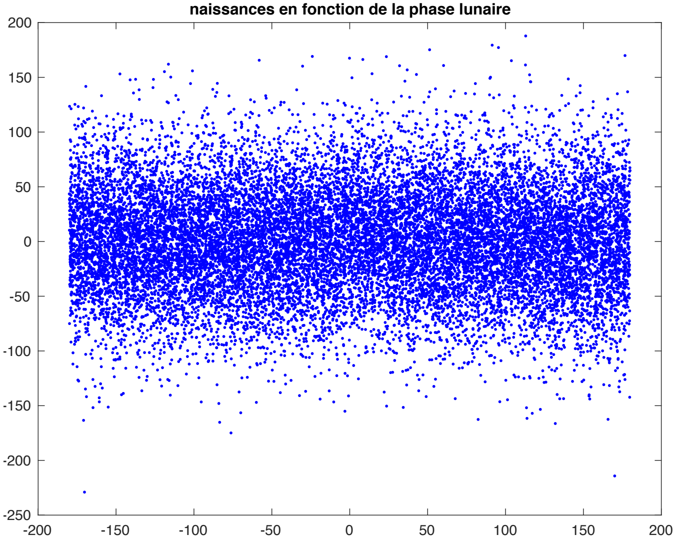
N°	Day	Event	Average deficit of birth
Fixed date holidays			
1	January 1	New year	-466
2	May 1	Labour day	-389
3	May 8	Surrender of Germany	-457
4	July 14	National Day	-393
5	August 15	Assumption	-340
6	November 1	All Saints	-356
7	November 11	Armistice	-372
8	December 25	Christmas	-491 (-23 %)
Mobile holidays			
9		Easter Monday	-364
10		Ascension Thursday	-391
11		Whit Monday	-388
Other particular days			
12	January 2	New Year's Next day	-115
13	December 22	J-3 before Christmas	-64
14	December 23	J-2 before Christmas	-107
15	December 24	Christmas Eve	-214
16	December 31	New Year's Eve	-149
17	Extra long weekend	Monday or Friday	-71
18	Friday 13	Superstition	-43
19	February 29	Leap day	-117
20	February 14	Valentine's Day	+23
21	Last Sunday in March	Switch to summer time	-59 (-3 %)
22	Last Sunday in October	Switch to winter time	+38 (+2 %)

p-values : all $< 1.5 \times 10^{-6}$ except Switch to winter time $= 6 \times 10^{-4}$ and Valentine's day = 2 %.
bigskip Superstition may affect the number of births since a deficit of birth appears in some days when the medical staff is a priori as numerous as the other days (Friday 13, Leap day)

Paraskevidékatriaphobie (*)

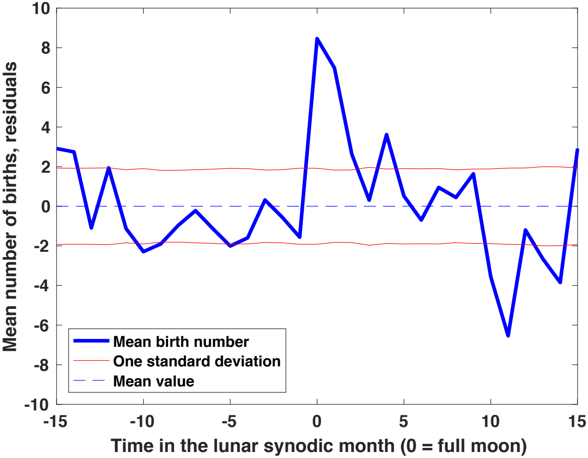
Operating rooms, Médipole, Villeurbanne, 2019

(*) Phobia for the friday the 13th.

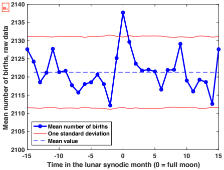


Birth residulas as a function of the lunar phase

Significant excess of births at FM and FM+1



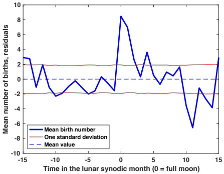
We understand why Fourier transform
is not the best
We have to decide wether
these values are different from 0 or not
= do a statistical test.



Global test (test if all means = 0)

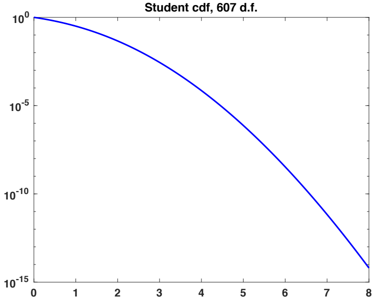
$$\text{Test statistic : } S = \frac{\text{variance between lunar days}}{\text{variance inside lunar days}} \approx 2.46$$

$$\text{p-value(global)} =$$
$$\text{Fischer cum. distrib. function } (x = 2.46, \text{ deg. freedom} = 30 \text{ and } 18233) =$$
$$= 1.5 \times 10^{-5}$$



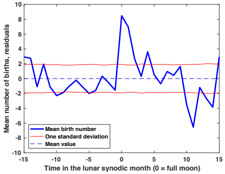
Full moon day (test if mean(FM)=0)

$$\text{Test statistic}: S = \frac{\text{mean}(FM)}{\text{standard deviation}(FM)} \approx 4.42$$



p-value(FM) =
Student cum. distrib. funct.
(x = 4.42, deg. freedom = 607)
$= 10^{-5}$

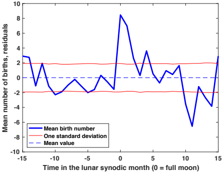
(and 1.5×10^{-6} with the exact FM day)



Correction for multiple testing
also called Family-Wise Error Rate

Probability than one mean over 30 has a statistic ≥ 4.42 :

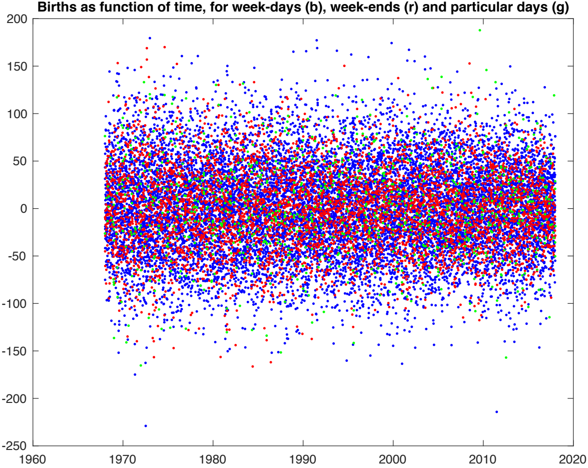
$$\text{Corrected p-value(FM)} = 1 - (1 - \text{pvalue})^{30}$$
$$= 4 \times 10^{-4}$$



Corrected p-values are all greater than 80 % except three : p-value(FM) $= 4 \times 10^{-4}$, p-value(FM+1) $= 4 \times 10^{-3}$, and p-value(FM+11) $= 2 \times 10^{-2}$.

It is very unlikely that the excess at FM and FM+1 is due to the random fluctuations of the number of births throughout days.

Isolated special days ? (outliers)



Isolated special days ('outliers')

Day	Event	Diff. of birth
Wednesday, July 16, 1969	Launch of Apollo 11	+216
Sunday, May 4, 1975	Extra long week-end	-205
Friday, July 15, 1977	Extra long week-end	-222
	between July 14th and the weekend	
Tuesday June 30, 1987	First day of summer holidays	+288
	the next day	
Friday, January 6, 1989	Weekend vigil	+254
	(return from vacation on the 5th)	
Friday, December 27, 1991	Weekend vigil in Christmas time	+185
Thursday, October 17, 1996	Public Service Strike	-199
Wednesday, August 11, 1999	Solar Eclipse	-237
Friday, December 31, 1999	Millennium New Year's Eve	-213
Monday, March 19, 2001	Midwives' Strike the next day	+203
Tuesday, March 20, 2001	Start of midwives strike movement	-339
Saturday, July 14, 2007	National holiday after a Friday 13	+263

Table – Days for which the residuals were greater than 3 standard deviations after a first application of the least squares program and for which we found an event that could influence the number of births on that day. In a second least squares step, we corrected the 12 data corresponding to these dates by the values indicated in the second column of the table. The residuals corresponding to the sum of the two stages of the least squares.

Individual p-values all $< 6.10^{-5}$.

Corrected, for multiple testing, p-values are smaller than 1 % only for the five days when Diff. > 237.

Observation time and size required

- With the raw data, the probability that the peak of births on the FM day is due to chance is 7 %. It's too big to conclude that it's significant.
- To fall to a probability of 0.001, always at the rate of 2000 births per day, it would take 200 years to observe !
- By processing the data well (residuals), it only takes 40 years.
- This increase in the number of births at the FM (+8 births for an average of 2121) is therefore **totally undetectable in a maternity ward** (10 births/day), or in a city, even over a lifetime !
- A belief that is not based on any observation.

Conclusions

- Significant variations in the number of births during the year ($\pm$ 20 %)
- Significant decrease during week-ends and holidays (-20 %)
- Slight increase on FM and FM+1 days (+0.4 %)
- No significant variation on the other days of the lunar month ; so lunar light doesn't seem to be a triggering factor.
- Neither is the attraction, otherwise the period would be 12 h 15.
- This increase is totally undetectable unless you have a lot of data and do a lot of statistical work.
- The number of medical staff affects the number of births
- Medical staff believes in the lunar effect
- Some maternity hospitals have (might have) more staff on full moon days ! ?

- The excess of births at the FM might be a self-fulfilling prophecy.
- Staff belief $\Rightarrow$ more staff at the FM $\Rightarrow$ more births.
- Self-fulfilling prophecies are everywhere :

  . Teachers towards the students
  (information received $\rightarrow$ differentiated expectations and attention),
  . Economic agents (announce a price increase, stock market...),
  . War or peace (hostility or supposed kindness of the other),
  . Placebo/nocebo effect...

Thanks !