THE STATISTICS OF EARTHQUAKES WITH RESPECT TO THE MOON

This is a companion document to the Exposing PseudoAstronomy podcast Episode #50 on "Lunatic Tides," what the evidence shows for whether Earth's moon triggers earthquakes. The episode is available online at http://podcast.sjrdesign.net/shownotes_050.php .

Claim: Statistical evidence shows that earthquakes occur, statistically, much more frequently near times of lunar syzygy (new or full moon, when the moon is aligned with the sun relative to Earth), and perigee (when the moon is closest to Earth).

"Oh, people can come up with statistics to prove anything, Kent. 14% of people know that." – Homer Simpson

Needed Data:

- Complete (as in, not picked for most costly, etc.) list of earthquakes. Need at least date/time and magnitude of the earthquake.
- List of dates and times of new, full, perigee, and – for completeness – apogee moons.

Data Sources:

- Southern California Data Center (http://www.data.scec.org).
- USGS (http://earthquake.usgs.gov/earthquakes/eqarchives/epic/epic_global.php).
- Moon Perigee/Apogee/New/Full Calculator (http://www.fourmilab.ch/earthview/pacalc.html).

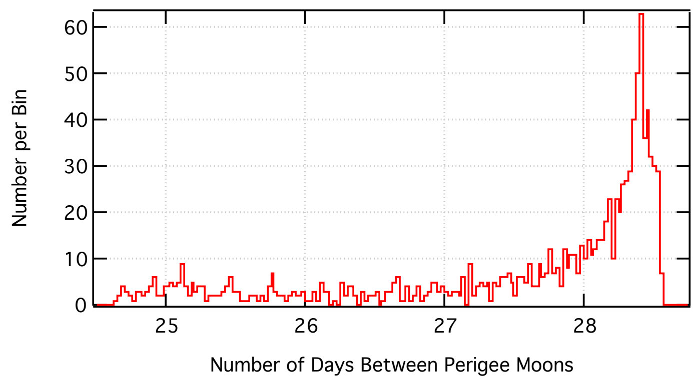
➢ I used the Southern California Data Center's data from 1932 through August 2012. It included data down to magnitude 3.0 earthquakes for the greater California area and magnitude 6.0 quakes for the globe. $N_{\text{quakes}} = 43{,}048$.

Lunar Data, and Why Analytic Statistics Won't Work

The time between new and full moon averages to 29.531 days, but it is not constant from one lunation to the next: There is a minimum of 29.28 and a maximum of 29.83 days. Fortunately, the distribution is roughly flat.

However, the number of days between apogee and perigee moons varies less predictably. Apogee moons occur at an average of 27.554 days apart, with a range of 26.98-27.90 days and a mode of 27.77 days – they are not normally distributed.

The range for perigee moons is highly variable, spanning 24.65-28.56 days with μ=27.555 days but a mode 28.4 days. It is roughly a Lorentzian distribution with μ=28.4.



Because these are variable, one cannot use analytic statistical techniques to estimate probabilities when combining moons (*i.e.*, probability of being within *x* days of both a new moon and perigee moon). Numerical methods are required.

Numerical Testing of Random Data

A 10^6-point Monte Carlo simulation was run for times between January 1932 and September 2012. The times of the simulated earthquakes were binned in intervals of 0.5 days for when they occurred relative to a perigee, apogee, new, or full moon. The distribution, as expected was random with respect to each moon type.

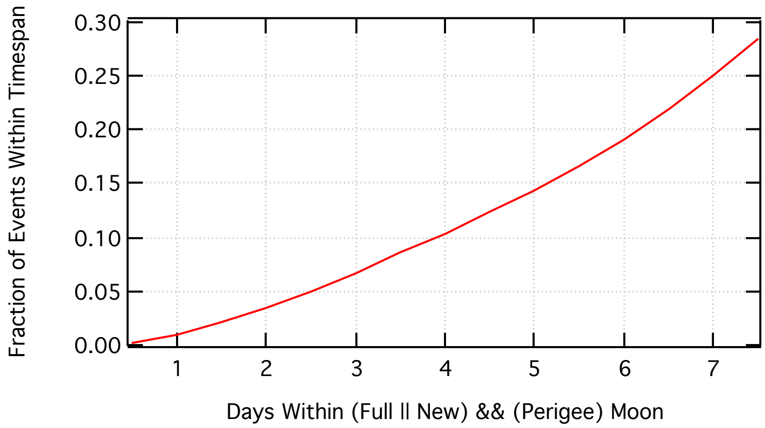
The data were then tested for when they occurred within both a new or full and perigee moon – in Boolean logic, (perigee) && (new || full).

This is where analytic methods would be faster but not possible. Ideally, if the interval between new moons and full moons were the same, say D_1, and the interval between perigee moons were D_2, then the probability that an earthquake would occur by chance within *x* days of both perigee and syzygy is:

$$p = \frac{2x}{D_1} \cdot \frac{2x}{D_2},$$

the factor of 2 being multiplied in because the earthquake could occur within *x* days on either side of the event.

However, because the lunar intervals are not even, the Monte Carlo data are used as a substitute to simulate what the probabilities should be. The Monte Carlo data were run for earthquakes within intervals of ±0.5 days for between 0.5 and 7.5 days of perigee and syzygy, the new and full moon results being averaged together:

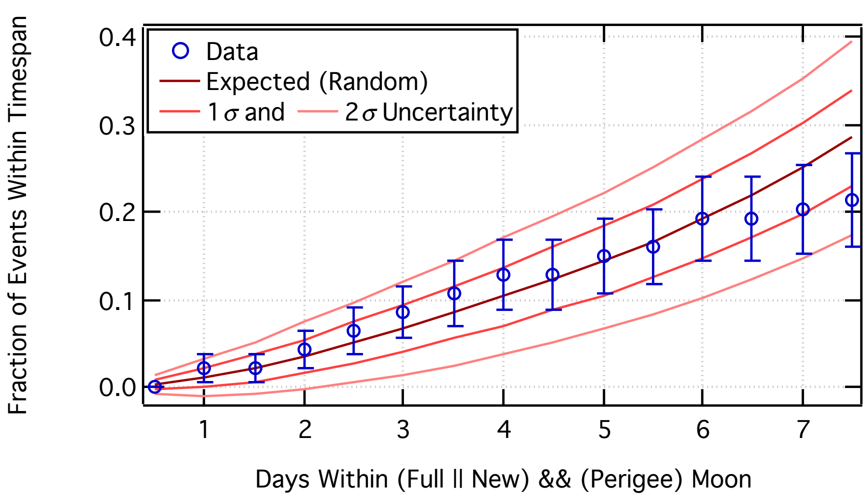


The uncertainty of each data point is dependent upon the number *N* of earthquakes in the sample size, for I used Poisson statistics to estimate the expected variation.

Note that by approx. 7 days from syzygy and perigee, the Sun-Earth-Moon system forms a right angle and tides would be at their lowest. Ergo, any significant earthquakes triggered by the hypothesis would need to be within just a "few" days of syzygy and perigee – I would argue anything beyond ±3 days is meaningless for this hypothesis.

Examination of Major California Earthquakes (≥6.0 magnitude)

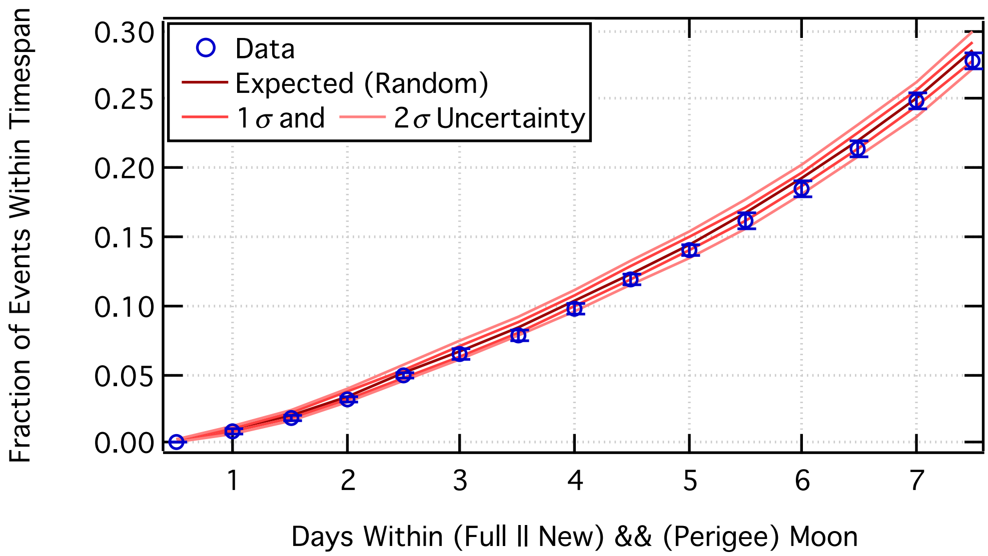
The data for 42 California earthquakes show no significant trend of when they occur relative to perigee, apogee, new, or full moons. Perigee and syzygy moon data were also examined:



Because there are only 42 earthquakes in the sample size, the expected range of the fraction of earthquakes within both perigee and syzygy is somewhat large. However, the data are fully within the expected confidence bands, illustrating the California earthquake data show **no** statistical significance beyond expected random chance for earthquakes happening within *x* days of perigee and syzygy. This is despite David Nabhan's claim that 33% of major California earthquakes occurred within ±0.5 days of both perigee and syzygy.

<u>Examination of Major Worldwide Earthquakes (≥6.0 magnitude)</u>

3073 earthquakes from around the world since 1932 of magnitude 6.0 or larger are in the database used. As with the major California earthquakes, no significance was seen with respect to when earthquakes occur relative to only apogee, perigee, new, or full moons. The same analysis was then run for perigee and syzygy:

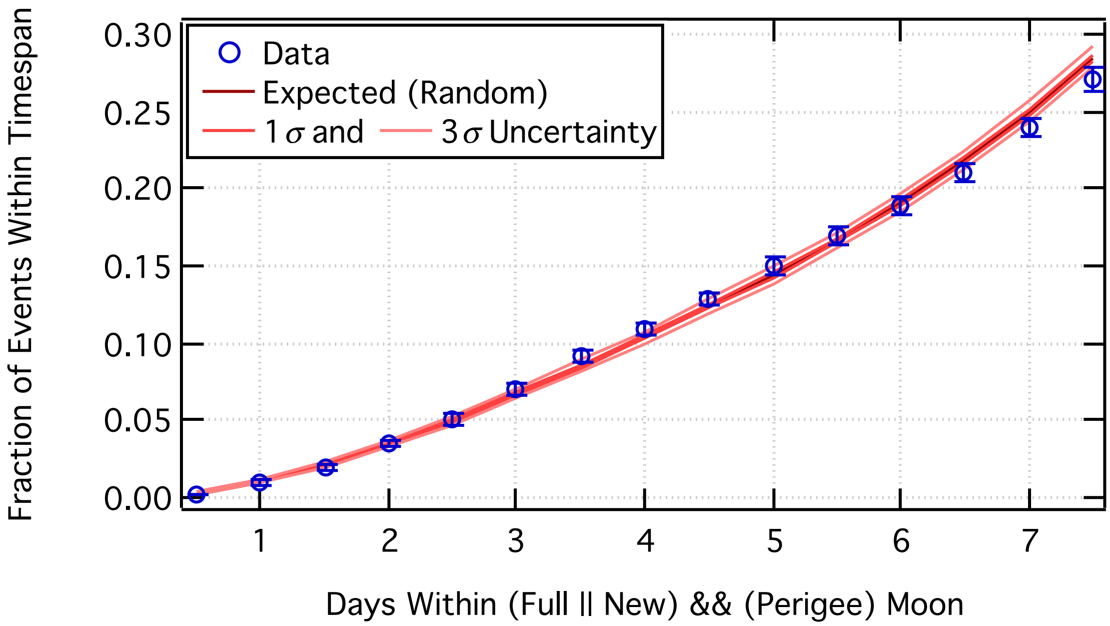


The confidence bands this time, as well as the individual uncertainty on each data point, are significantly smaller than before. That's because of the much larger sample size. In Poisson statistics, the standard expected uncertainty is the square-root of the counts. While the raw numbers with 3073 earthquakes will be larger than those from a sample size of 42, the *relative* uncertainty is significantly smaller. Since the y axis on these graphs is the <u>fraction</u> of earthquakes, both the actual number and the uncertainty are divided by the total number of earthquakes, hence the smaller uncertainties.

And, as can be seen, <u>the worldwide earthquake data show no significant deviation from random chance.</u>

Examination of All (≥3.0 magnitude) California Earthquakes

These are the only data that show a statistically significant deviation from random chance. N in this sample was 26,227 earthquakes. Of being within individual moons, the data were statistically off from random chance, but they were completely random as to the time relative to perigee, apogee, full, or new moons. For example, earthquakes were statistically more likely to occur 11 days after perigee, within ±4 days of apogee, 10 and 3 days before new, and 1, 6, and 13 days after full.



In the perigee and syzygy test, note the change in display – out to 3σ bands are displayed, and the error bars on each individual data point this time are also $\pm 3\sigma$.

The data are fully within the expected $\pm 1\sigma$ range for being within ±3 days of perigee and syzygy. However, beyond those days, the data show a statistically significant – albeit 3σ value – enhancement of earthquakes. Beyond approx. 5 days, there is a statistically significant (though again within $\pm 3\sigma$) deficit of earthquakes.

What does each σ level mean? A 1σ band means that, roughly 68.3% of the time, the data should fall in that range. A 2σ band means that 95.45% of the time, the data should fall within that range. So if it falls outside of that range, there's still a 4.55% chance (if it's Gaussian-distributed data) that it's due to purely random chance. A 3σ band means that 99.73% of the data should fall within that range, so a 3σ result means that it only had a 0.27% chance of happening by pure, random chance. In physics, the "gold standard" is a 5σ result, meaning that the odds of it occurring by random chance are only 0.000573% – roughly 1 in 1.7 million.

So is a 3σ result significant here? Possibly, but unlikely. The intervals of significance relative to moons are in multiples of about 1 week indicating possible bias in the data, and the data for days within ±2 of perigee and syzygy are still well within the 1σ range. A 3σ result centered at 3.5 days out is far from the most significant tides, and so I would argue that even if this is a significant result, it's (a) still only for California, and (b) does not actually support the initial hypothesis.

After this analysis, the null hypothesis, that tides do not trigger earthquakes, is upheld.