

ULA Period Tracker - Logic Math.

ULA Period Tracker's Predictions are made via the following mathematical formulations:

Prediction Engine:

The prediction engine is based upon the BAvC, the Biological Average Calculation, i.e. the average length of a Regular Cycle. As obtained from Mayo Clinic information the BAvC is equal to 28 days.

However, clockwork regularity is rare, so the Prediction Engine utilizes User Data to create the Personalized Average Calculation, PAvC, an adaptive calculation where current cycle information allows the Engine to predict further period cycles.

Given a set of recent cycles $C = c_1, c_2, \dots, c_n$ where $n \leq 6$

$$PAvC = \frac{1}{n} \sum_{i=1}^n L_i$$

Where L_i is the length of cycle c_i in days

But, even this set of information must account for standard deviation, as such, using only 4 Cycles, the Engine can deviate In the following manner.

Where $\bar{L} = PAvC$, or if missing, BAvC as an Edge Case backup and $n \geq 2$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (L_i - \bar{L})^2$$

then Standard Deviation = $s = \sqrt{s^2}$

Where a sample standard deviation is used ($n-1$) because a small sample gives better statistical behaviour with the unbiased estimator.

However, because the equations deal with statistics between biological limits, as such a Clamped Standard Deviation, present for Code Logic purposes, must be introduced.

$$\begin{aligned} \text{lowerBound} &= 2 \\ \text{upperBound} &= \min(10, \frac{1}{2}PAvC) \end{aligned}$$

Then:

$$s_{\text{clamped}} = \min(\max(s, 2), \text{upperBound})$$

Which prevents the prediction window from becoming either too narrow (<2 days) or impossibly large (>10 days, i.e. about half a standard cycle)

The Prediction Engine also utilizes Deterministic math to predict the expected period date and/or window, ovulation, and fertility windows.

Period Date:

Where D_0 = Date of Last Period Start and \bar{L} = PAvC, or BAvC if need be

$$D_{\text{expected}} = D_0 + \text{round}(\bar{L})$$

Period Prediction Window:

$$\begin{aligned} \text{Using } s_{\text{clamped}} &\Rightarrow D_{\min} = D_0 + \text{round}(\bar{L} - s_{\text{clamped}}) \\ D_{\max} &= D_0 + \text{round}(\bar{L} + s_{\text{clamped}}) \end{aligned}$$

Ovulation Prediction:

Assuming the Luteal Phase to be an average of 14 days:

$$\begin{aligned} \text{Given } D_{\text{next}} &= \text{expected next period start} \\ O &= D_{\text{next}} - 14 \end{aligned}$$

Fertile Window:

Given ovulation date O: $D_{\text{fertile start}} = O - 5$ and $D_{\text{fertile end}} = O + 1$

therefore

$$[O - 5, O + 1] = \text{Fertile Window}$$

All of this composes the Prediction Engine, that calculates the main predictive functions of ULA

Analytics Engine:

However, what are predictions without analysis. The Analytics Engine computes cycle and period statistics, irregular cycle allowances and summaries, the Prediction Confidence Score, and miscellaneous descriptive metrics.

It does this via standard Statistical Metrics, such as standard Mean and Median Averages, Min Max, and standard Deviation utilizing an unbiased estimator.

Cycle and Period Length:

In code the Cycle and Period Lengths are calculated using dateComponents, making the calculation a simple date difference. For Cycle, given a start date S, and an end date E, the cycle is equal to days(S, E). Similarly for Period Length, using Bleeding Episodes as a measurement, it's days(S, E) + 1 as it counts both the first and last day as a full bleeding event.

Irregular Cycles:

Due to the nature of their irregularity, the Analytics Engine only follows an irregularity rate, where it calculates the last 180 days of regularity/irregularity, where (using actual in code labels):

$$\begin{aligned} \text{tooShortCount} &= \#\{i : \text{reason}_i = \text{tooShort}\} \\ \text{tooLongCount} &= \#\{i : \text{reason}_i = \text{tooLong}\} \\ \text{total} &= n \end{aligned}$$

Measured as such:

$$\begin{aligned} \text{toStart}_i &\geq \text{today} - \text{recentDays} \\ &\text{Using default of 180 days} \end{aligned}$$

Then:

$$\text{recentIrregularCount} = \#\{i : \text{toStart}_i \geq \text{today} - 180\}$$

Prediction Confidence Score:

Markedly the most 'algorithmic' part of the Analytics Engine:

Given:

$$\begin{aligned} \text{Cycle Lengths: } &L_1, \dots, L_n \\ \text{Mean: } &\bar{L} \\ \text{Standard Deviation: } &s \end{aligned}$$

And assuming that:

$$\text{For } n < 2 \text{ confidence} = 0.2$$

Then, we can compute the variability ratio:

$$v = \frac{s}{L}$$

Then normalize to a confidence score:

$$\text{raw} = 1 - \min\left(\frac{v}{0.3}, 1\right)$$

And clamp to [0, 1]:

$$\text{confidence} = \min(\max(\text{raw}, 0), 1)$$

Which results in cycles with low variability where the Standard Deviation nears 0, it gives a high confidence, but cycles with a larger variability lower our confidence... resulting in a heuristic stability score.

For the proper functioning of the system other minor engines exist, such as the Cycle Analyzer which gives computations on period episodes relating to ordered dates, computing as well the output from the Analytics Engine, resulting necessary for proper functioning of the Logic System. Or even the MoonPhaseEngine, a bit of fun that calculates the Moon Phases on-device via the knowledge of the New Moon of the 6th of January 2000, and correlates it to our MonthBuilder, and through it, to the UI, and the User.