Intelligent Design Found in the Sky with $p < 0.001$

David Mumford

Looking at the sky from my hot tub in Tenants Harbor, as night falls earlier and earlier in the fall, I wait for the first sighting of Orion. One evening, there it is, a warrior resplendent against the southeastern sky. Its seven principal stars all carry names - Rigel, Betelgeuse, Bellatrix, Saiph, Mintaka, Alnitak and Alnilam - and are among the 67 brightest stars in the whole sky[1]. The constellation is unmistakable not only as a cluster of so many very bright stars but also by its striking humanoid shape: Betelgeuse and Bellatrix form the shoulders, Saiph and Rigel the knees and Alnitak, Alnilam and Mintaka the belt. In addition, below the belt are the three stars, one the great nebula of Orion, forming Orion's sword. Every culture has recognized this striking cluster of stars: it was the god Osiris in Egypt, the Vedic creator of the universe, Prajapati, in India, one of the mansions of the White Tiger in China and the great father Hunhunahpo in Mayan Mexico. It is even conjectured to be the carving in a tusk dating from 32,500 BCE[2].

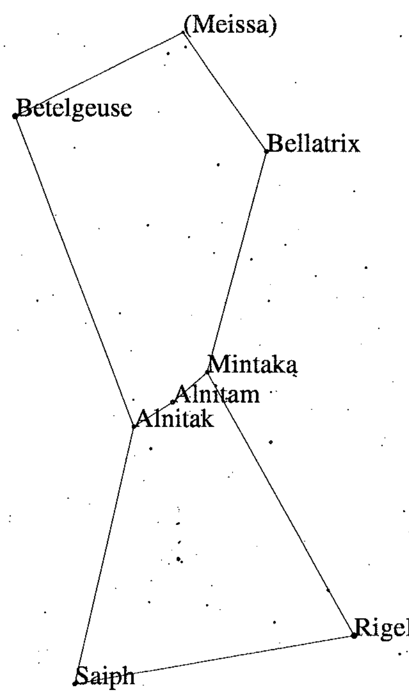


Figure 1. The constellation Orion and its seven principal stars.

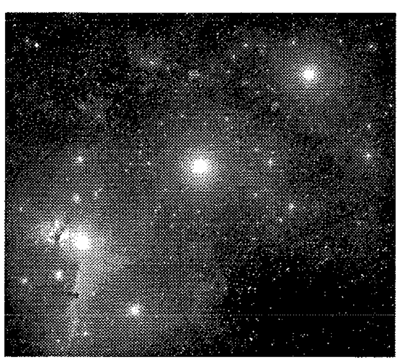


Figure 2 The three stars of Orion's belt blurred in a long exposure in order to show the nebulosity next to Alnitak

This year the thought crossed my mind: is it not very improbable, if 67 stars were scattered at random in the celestial sphere, that such a pattern would be present? Having worked in computer vision, it is conceivable that the statistical models used in object recognition could quantify this. However, full human body models are not really ready for 'prime time'. But at least we can ask whether it is probable or not that 7 out of the 67 brightest stars should wind up so close to each other? Moreover,

[1]Because of variable and binary stars, there is some ambiguity in ordering stars by brightness, but using the listing in http://www.astro.uiuc.edu/~kaler/sow/bright.html the seven principal stars in Orion have ranks 7,11,26,29,30,52 and 67

[2]This is an analysis of Michael Rappenglueck, as reported in the BBC News of 21 January, 2003

the key component in what is sometimes called 'early vision' - that is the first steps in the analysis of the patterns of an image - is the identification of straight lines and extended curves in images. Psychophysics, esp. the experiments of the gestalt school, has confirmed that human perception recognizes these patterns in the midst of clutter with amazing sensitivity. Such curves can be contours of objects or parts of objects (such as limbs of trees). The three stars in the belt of Orion are striking not only because they are very close but because they are almost exactly regularly spaced in a line. Now the occurrence of such a linear pattern is easy to quantify.

Firstly, in the table below, we give the key facts about the seven main stars of Orion.

Star	Magnitude	Right Ascension	Declination	Distance
Alnitak	1.74	05 40 45.5	-01 56 34	815 ly
Alnilam	1.70	05 36 12.8	-01 12 07	1340 ly
Mintaka	2.23	05 32 00.4	-00 17 57	915 ly
Betelgeuse	0.70	05 55 10.3	+07 24 25	425 ly
Bellatrix	1.64	05 25 07.9	+06 20 59	245 ly
Rigel	0.12	05 14 32.3	-08 12 06	775 ly
Saiph	2.06	05 47 45.4	-09 40 11	720 ly

The data is from the Yale Bright Star Catalog (available via `ftp://cdsarc.u-strasbg.fr/cats/V/50/catalog.gz`), with recent distances from the Hipparcos satellite data, (found in `http://www.astro.uiuc.edu/~kaler/sow/bright.html`). One checks that all seven stars are within 9.82º of Alnilam, the central belt star. Within the belt, Alnitak and Alnilam are 1.356º apart, Alnilam and Mintaka 1.386º apart, a difference of only 2.2%. And that the exterior angle in the polygon joining Alnitak, Alnilam and Mintaka is only 7.5º.

To quantify the improbability of this, we turn to hypothesis testing. Hypothesis testing is the gold standard, for instance, of medical tests. Does some treatment improve a patient's chances of getting better? Well, suppose you know from past history that p_U is the probability of recovery in untreated patients. Now you take 1,000 patients and give them the treatment. Suppose p_T is the proportion of the treated patients who get better. Of course p_T had better be bigger than p_U or you can stop there. Then you imagine a game in which p_U is the chance of winning and you calculate the probability p of winning this game 1000* p_T or more times if you play it 1000 times. In other words, we consider the null hypothesis that the treatment had no effect and then ask, if we assume the null hypothesis, what is the chance of seeing a proportion p_T or larger of patients being cured in a population of 1000. If $p < .01$, it is customary to give the treatment a seal of approval. In other words, when your health is at stake, if there is 1% or less chance of the medical test results coming out the way they did under the assumption that the treatment is worthless, you declare to the world at large that the treatment is worth taking.

We want to apply hypothesis testing to Orion. We use the null hypothesis that the stars are scattered at random in the sky and we ask: what is the probability that the circle of radius 9.82° around one of them should contain 6 others. This is trivial to compute:

$$\text{Prob} \leq 67 \times \binom{66}{6} \times \left(\frac{\text{area(spherical disk),r=}9.82^{o}}{4\pi}\right)^{6} \approx .001$$

BUT we are now committing the cardinal sin of hypothesis testing: we are choosing our test after we have the data, not before. This is the standard problem with people noticing "coincidences". Some striking thing occurs (Barlow used to talk of seeing five yellow VW bugs on the street one morning) and you say - "the probability of this happening by accident is tiny, so there must be some reason". What you don't do is try to imagine how many million other odd things might have happened but didn't. You picked the one test for which your reality had a low probability. What you need to do is apply the Bonferonni correction: if there are N possible remarkable events of which one actually occurred, you should take the p-value of that event, its probability under the assumption that everything is normal, and multiply it by N and ask if this probability is small, e.g. less than .05.[3]

In the case of Orion, we chose to test for a tight cluster of 7 stars from the brightest 67. But there are many other possibilities, e.g. the Pleiades, a much tighter cluster but not all as bright. This was considered by John Mitchell in 1767 as we shall discuss later. If we put ourselves in the shoes of a person who has not seen the stars and ask what tests they might make to see if there are remarkable clusters, one approach, for example, would be to use the classification of stars by magnitude. Visible stars range in magnitude from -1 (the brightest) to 5 (maybe 6 but this requires very clear dry air which is in short supply these days). The seven major stars of Orion are all of magnitudes 0, 1 or 2. The six brightest stars of Pleiades are of magnitudes 3 and 4. There are, by one count, 2, 6, 14, 69, 192, 610 and 1929 stars of magnitudes respectively -1, 0, 1, 2, 3, 4 and 5. We might assign the significance level p and form a test for each magnitude n and cluster size m. If there are $N(n)$ stars s_k of magnitude at most n, we take as our test statistic:

$$f_{m,n}(s_1,\ldots,s_{N(n)}) = \min_{i(1),\ldots,i(m)\in(1,\ldots,N(n)} \max(\text{dist}(s_{i(2)}-s_{i(1)}),\ldots,\text{dist}(s_{i(m)}-s_{i(1)}))$$

[3]If all the tests are made at the same level p of significance, then the probability of one occuring under the null hypothesis is $1-(1-p)^{N}$ which is about Np

We find the value $f(n, m)$ such that:

$$\text{Prob}(f_{m,n}(s_1, \dots, s_{N(n)}) \leq f(m,n)|\text{stars random}) = p$$

Then we check the values of this test statistic on the actual stars. The seven major stars of Orion are of 2^{nd} magnitude at most and there are 91 of these on Kaler's web site referred to above (counting double stars as one). Then

$$\text{Prob}(f_{7,2} \leq 9.82^o) \leq 91 \times \binom{90}{6} \times \left(\frac{\text{area(spherical disk),r=9.82}^o}{4\pi}\right)^6 \approx .009$$

Aha: this means that if we chose p equal to the standard level 0.01 of statistical significance, we would find Orion causes us to reject the null hypothesis and conclude that the stars were not randomly distributed. But we have still committed the sin of fitting our statistic to the data by choosing the numbers $n = 2$ and $m = 7$. We can apply the same criterion to the belt, where 3 stars are within 1.386º of the center star:

$$\text{Prob}(f_{3,2} \leq 1.386^o) \leq 91 \times \binom{90}{2} \times \left(\frac{\text{area(spherical disk),r=1.386}^o}{4\pi}\right)^2 \approx .008$$

This is similarly 'statistically significant' - but not with a truly tiny p-value. I have not systematically examined for which m and n such significant clusters exist. This would be necessary to go on to ask whether, if the stars were random, this collection of clusters was unlikely. Instead, I want to turn to a more unlikely situation which appears to be present in Orion.

Let's examine the belt more closely. Its amazingly symmetric configuration - three almost equally spaced stars very nearly on a line - is highly unusual. Such a configuration is called a 'linelet' in computer vision. If you consider clusters of three stars, there are only two striking special geometric configurations: equally spaced on a line or the vertices of an equilateral triangle. The Gestalt school of psychophysicists[4] investigated at great length what patterns in an image caused its points, lines and other parts to be grouped, to be seen as part of one object. The belief is that, for ecological reasons, what humans see is determined by what 2D patterns are most helpful in working out the 3D world around us. *Proximity* and *alignment* turn out to be the two strongest factors leading to visual grouping. An equilateral triangle is not a configuration found by the Gestalt school to be highly salient to the human visual system. This is presumably because equilateral triangles are not common in our visual experience whereas straight lines, whole or partially occluded, repetitive texture patterns and linear motion are very common.

[4]See, for example, G.Kaniza's book *La Grammaire du Voir*,Diderot, 1997 (origionally *Grammatica del Vedere*,1980

We need to develop a specific statistic to measure the linearity of the belt. The most natural is the discrete second derivative, the angular distance from the middle star to the midpoint of the first and third star:

$$c = \mathrm{dist}(s_2, \mathrm{midpt}(s_1, s_3))$$

If $b = \mathrm{dist}(s_1, s_3)$ is the overall size of the 'linelet', then we are associating to every triple of stars the simple, elementary and natural pair (b, c) which measures how closely it is a small 'linelet' (to use the terminology of computer vision). To develop a test, we need to combine b and c. It is easy to see that for three random stars, they are independent and have a distribution with density $\sin 9b)\sin(c)/4)dbdc$. Since they are independent, we take as our test statistic $T = bc$. But this being small is not surprising unless both b and c are reasonably small, e.g. b should be less than the expected diameter of the smallest triple among the 91 randomly placed stars. A Monte Carlo simulation shows this to be about 6.7^o or 0.117 radians. For the triple to look remotely like a linelet, we ask $c \leq b/8$ which means the spacing at worst 3:5 and the exterior angle at the middle star is less than 29^o. Then if we observe $T = T_0$, the p-value of this event among all stars of magnitude at most two is:

$$91 \times (91 \times 89/2) \times \iint_R \frac{\sin(b)\sin(c)}{4} dbdc, R = \{b, c | bc \leq T_0, c \leq b/4, b \leq 0.117\}$$

In the case of Orion's belt, b $\approx$ 0.048 radians, c is merely 5.5 arc minutes or about 0.0016 radians, thus $T_0 \approx 0.000076$. To evaluate the integral, we approximate $\sin(b)$ by b and $\sin(c)$ by c and find easily that p $\approx$ .00034.

Now this is *much more* significant from a statistical viewpoint. But we still ought to allow for alternate tests for events that might have occurred but did not. While looking for unusual alignments, perhaps our cutoff at 2nd magnitude is arbitrary and perhaps 4 aligned stars should be considered too. This part of the argument really cannot be made precise. A common procedure is to allow some factor for this: I suggest 3, making the conservative p-value for the alignment of Orion's belt 0.001.

Now if the null hypothesis is rejected, what can be the cause of this alignment? In Gestalt psychology, alignment of some points in an image leads the perceiver to assume the world points projecting to these points on the image are aligned in three dimensions, unless there is strong evidence to the contrary. Aligned points in the world will be seen as aligned on the retina no matter what the viewpoint. Likewise, a cluster of salient points in an image is assumed to be caused by a cluster of points in the world.

As we mentioned, John Michell in 1767[5] applied statistics to the Pleiades. Using the null hypothesis that the stars are scattered at random over the full

[5]Michell J.(1767). An inquiry into the probable Parallax, and Magnitude, of the Fixed Stars, from the Quantity of Light which they afford us, and the particular Circumstances of their Situation,*Philosophical Transactions*, v.57, p. 234-264

celestial sphere and neglecting the caveats we have discussed, he asked how likely was it to find six stars as close together as they are in the Pleiades, among all the stars at least as bright. He found $p = .000002$ for the Pleiades occurring by random chance. He deduced from this that the null hypothesis was wrong and proposed that the Pleiades must be clustered in 3-space so that their positions in the sky were correlated, not independent. He actually went a bit farther and for this he was greatly criticized: he proposed assigning *prior* probabilities to the possibilities that these stars were close in 3-space vs. being distant in 3-space and merely close from the earth's vantage point. He could then apply Bayes's rule to deduce that .000002 was also the probability that the Pleiades were a cluster in space. In fact, his conclusion was right: the Pleiades is indeed a cluster designated M45 in Messier's catalog.

How are the 3 stars of Orion's belt aligned in space? Fortunately, the Hipparcos satellite has provided excellent data on stellar distances. The result for Alnitak, Alnilam and Mintaka is shown in figure 3. It is clear that if the sun were positioned a little bit above or below the plane of the belt, the three stars would fall out of alignment immediately, and the central star, Alnilam would move away from the other two. So Mitchell's alternate hypothesis does not explain Orion's belt.

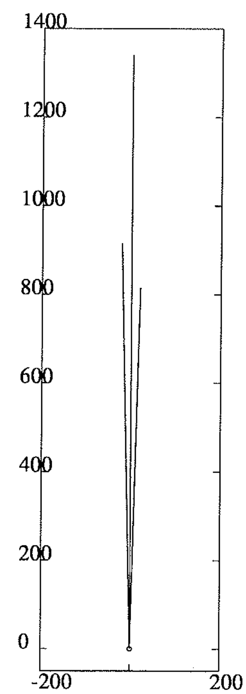


Figure 3 Alnitak, Alnilam and Mintaka in space. The small circle is the sun

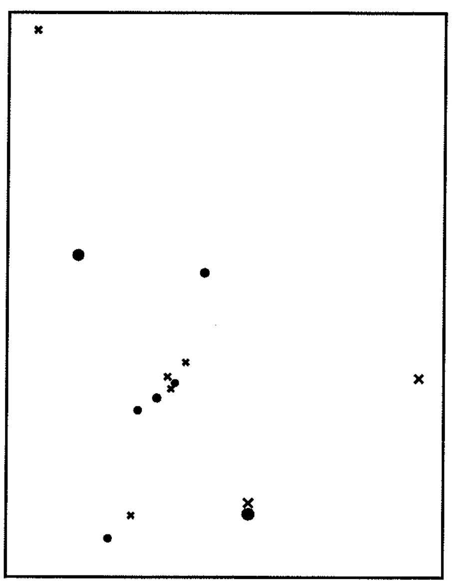


Figure 4 The seven stars of Orion now (dot) and 2 million years ago (cross)

As the sun and the seven stars of Orion move around our galaxy, the shape and the very existence of the grouping we call Orion will not remain. Betelgeuse is moving the fastest relative to the rest of Orion, flying left and up (north). Bellatrix is moving to the right and down nearly as fast. The rates are roughly a degree every 200,000 years or so. Alnitak is leaving the other 2 belt stars by a degree every 1-2 million years: enough to break its symmetry.

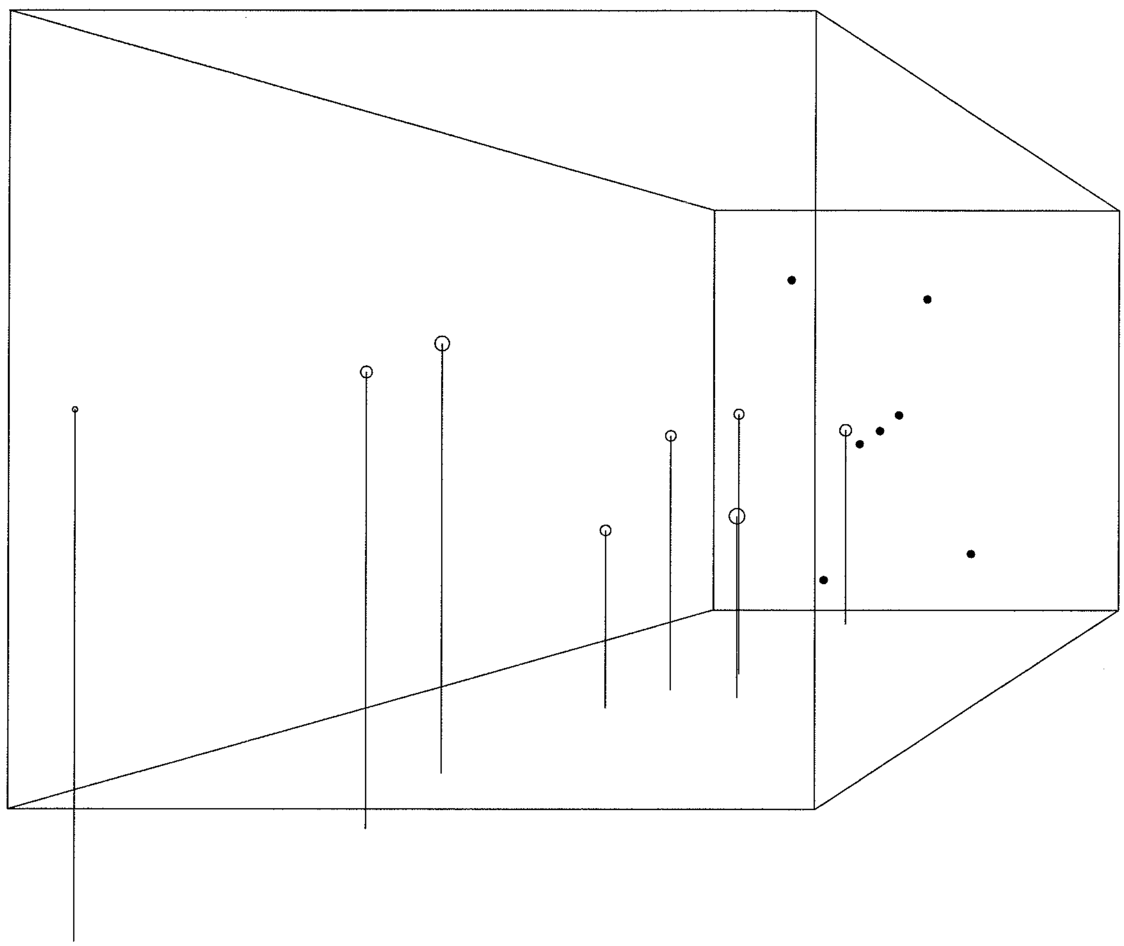
According to Rappenglueck (op.cit.), the shape of Orion has altered enough since Neolithic times that this can be detected in the prehistoric carving he analyzed. Figure 4 above shows a reconstruction of how Orion looked 2 million years ago with Betelgeuse off to the left, Bellatrix at the top.

What alternate hypotheses are we left with? Some might indeed infer from this evidence for intelligent design: that the creator has caused these 7 stars to assemble themselves as a great warrior just as homo sapiens emerged on earth[6]. Frequentist statistics is a wonderful tool. Bayesians, on the other hand, put priors on alternate hypotheses such as intelligent design and, depending on your personal prior, this can radically alter your conclusions.

◇ ◇ ◇ ◇

Orion in space

(Black spot closest to observer is position of our sun)



[6]One of my sons suggested this could be described as God 'micro-mangaging' the world