Applied Statistics

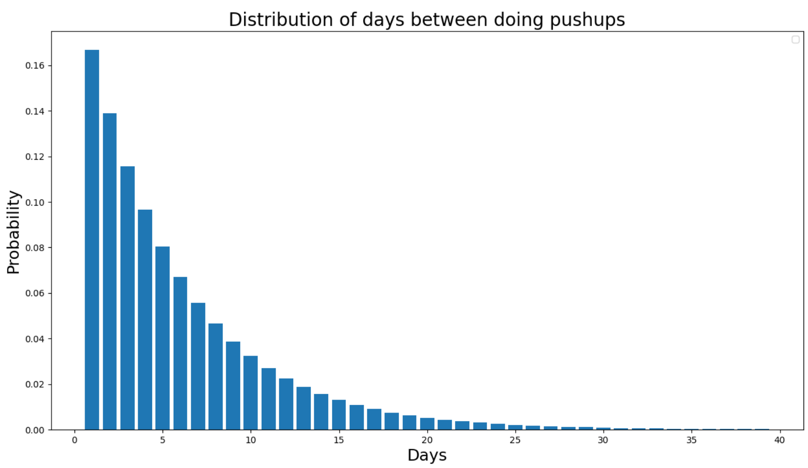
Exam in Applied Statistics 2024/25 by vfd881

1 Distributions and probabilities

1.1 1.1 (6 points) Every day, you roll a normal die, and if you get a six, you do roll the die 100 times and do as many pushups as you get sixes. Otherwise you don't do any pushups

1.1.1 What is the distribution of days between doing pushups?

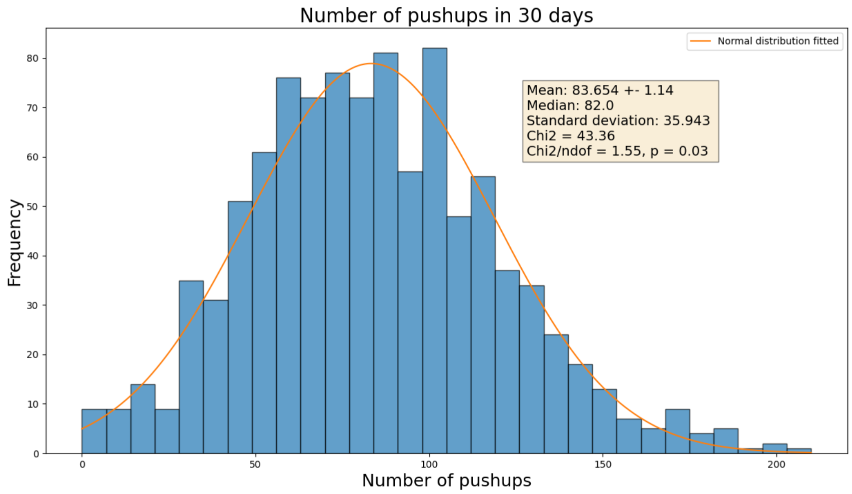
Here I will use a geometric distribution, as it will give me the probability of the first success happening on a given day. This I can use to plot the distribution of the days it will take to get a six and afterwards plot it.



Figur 1: Geometric distribution of day between doing pushups

1.1.2 What is the mean, median, and standard deviation of number of pushups in 30 days?

While the exact distribution of total pushups comes from a binomial distribution, the Central Limit Theorem (CLT) says that the sum of random variables (here, total pushups) approaches a normal distribution. Therefore a normal distribution provides a good approximation for the histogram (assuming poisson error on the counts), seen here:



Figur 2: Number of pushups over 30 days

I get a mean of 83.7 ± 1.1 with a standard deviation of 35.9. My chi2/ndof looks relative fine, but with a low p-value. A higher sample of days, would (as a result of CLT) make the histogram approach a better normal distribution.

1.2 1.2 (4 points) The Djoser pyramide in Egypt is North-South aligned to 3 degrees

1.2.1 Estimate the probability that the pyramid is North-South aligned by coincidence.

If it's by coincidence then I can assume a uniform probability distribution of angles, meaning it could be alligned to any angle within 360 degrees. I will run a simulation (1000000 simulations to be exact) to see how many fall in the range of ± 3 degrees:
Probability of being aligned by coincidence: $1.67\% \pm 0.01\%$

2 Error propagation

2.1 (5 points) Water on Earth ($\oplus$) has a Deuterium to Hydrogen ratio of $r_\oplus = (149 \pm 3) \times 10^{-6}$. The hydrogen of the proto-solar system ($\odot$) has a ratio of $r_\odot = (25 \pm 5) \times 10^{-6}$, while that of comets ($C$) has been measured to be $r_C = (309 \pm 20) \times 10^{-6}$

2.1.1 From these numbers, what fraction of water on Earth do you estimate come from the original proto-solar system, and what fraction do you attribute to comets?

To do this I assume a linear correlation between the ratio of water on earth and the ratios of water coming from the protosolar system and the comet.

$$r_E = f_p \cdot r_p + f_c \cdot r_c \tag{1}$$

where r_E is the ratio of water on earth, f_p is the fraction of water coming from the protosolar system, r_p is the ratio of water in the protosolar system, f_c is the fraction of water coming from the comet, and r_c is the ratio of water in the comet. This also means:

$$f_p + f_c = 1 \tag{2}$$

$$r_E = f_p \cdot r_p + (1 - f_p) \cdot r_c \tag{3}$$

$$f_p = \frac{r_c - r_E}{r_c - r_p} \tag{4}$$

$$f_c = 1 - f_p \tag{5}$$

To calculate the error I used error propagation

$$f_{p,\text{error}} = \sqrt{\left(\frac{r_{E,\text{err}}}{r_C - r_P}\right)^2 + \left(\frac{r_{C,\text{err}}(r_E - r_P)}{(r_C - r_P)^2}\right)^2 + \left(\frac{r_{P,\text{err}}(r_C - r_E)}{(r_C - r_P)^2}\right)^2}$$

$$f_{C,\text{error}} = \sqrt{\left(\frac{r_{E,\text{err}}}{r_C - r_P}\right)^2 + \left(\frac{r_{C,\text{err}}(r_P - r_E)}{(r_C - r_P)^2}\right)^2 + \left(\frac{r_{P,\text{err}}(r_C - r_P)}{(r_C - r_P)^2}\right)^2}$$

The results yielded

- **Fraction of proto-solar material:** 0.563 ± 0.105
- **Fraction of comet material:** 0.437 ± 0.137

2.2 (8 points) You run a detector for a time interval of $\Delta t = 98.4\,\text{s}$, during which the detector yields $N = 1971$ counts. The time interval uncertainty is $\sigma_{\Delta t} = 3.7\,\text{s}$, independent of Δt.

As I have an error on the time interval, I need the uncertainty on the counts. These I assume to be poisson errors meaning the erros is the square root of counts.

$$r = \frac{N}{t}$$

$$r_{\text{err}} = \sqrt{\left(\frac{N_{\text{err}}}{t}\right)^2 + \left(\frac{N \cdot t_{\text{err}}}{t^2}\right)^2}$$

Results yielded (I added the relative error to compare with the next question):

- **Rate of decay:** 20.030 ± 0.878 decays/s
- **Relative error of rate of decay:** 4.38

2.2.1 How long should you measure to get a relative uncertainty on the rate r below 2.5%?

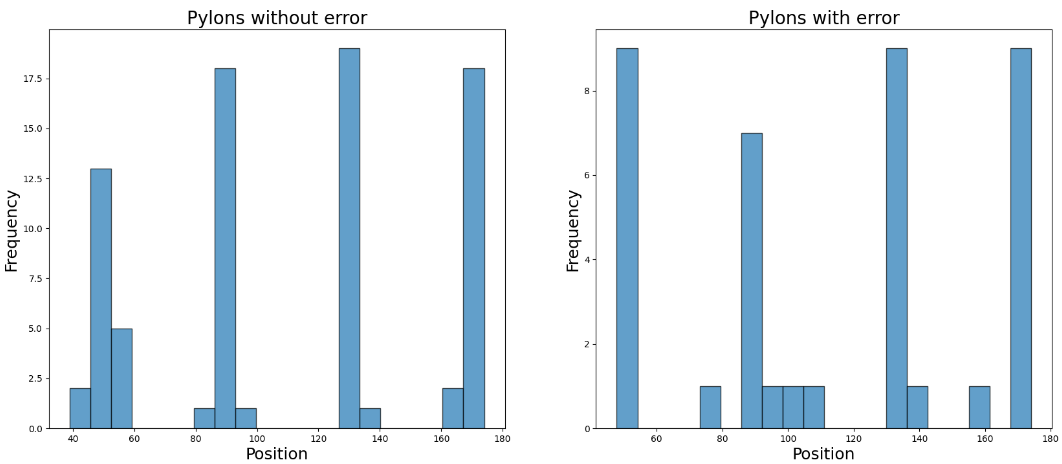
I will run a simulation over a range of t's to see how the error changes with the counts and find when the relative error is less than 2.5%. I will start at t = 98.4, since I know that gives a relative rate of 4.4%. I will do this in stepsizes of 0.1 s. Looking for: $\frac{\sigma_r}{r} < 0.025$

The time needed for a relative error of 2.5% is: 341.2 s

2.3 (14 points) The file PylonPositions.csv contains position measurements (both with and without uncertainties) of four pylons for a bridge.

2.3.1 Using measurements without uncertainty, determine the four pylon positions.

First I plotted the data in histograms to see visually if I could guess where the pylons where located:

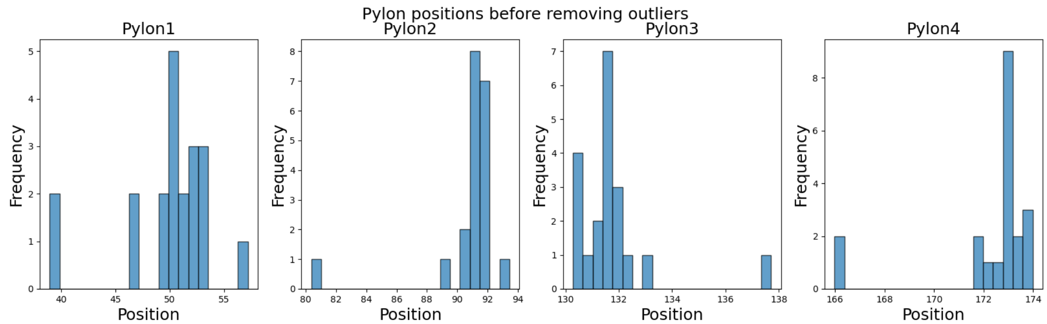


Figur 3: Distribution of pylons positions

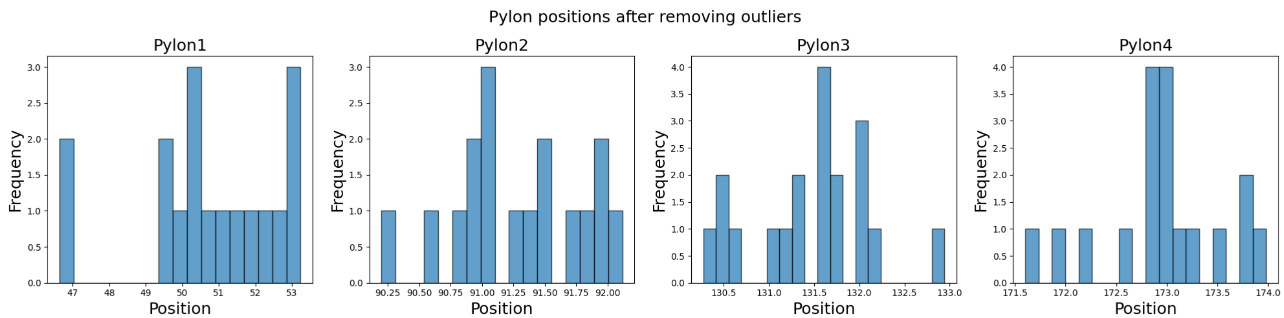
I will start with the pylons without error, where it is very easy to see the 4 pylons as seen on the histogram. These are the masks for the position I used to identify the pylons:

- Pylon 1 < 70
- Pylon 2 > 70 and < 110
- Pylon 3 > 110 and < 150
- Pylon 4 > 150

As they don't have any uncertainty, I will simply take the mean of the 4 peaks as the position of the pylons. Furthermore I will use chauvenet's criterion to remove the outliers, if there is any. To do this I will use student's t-distribution to test for outliers as the varience is unknown and N is small. I set my criterion for the t-test to have to be less than 0.25 for me to remove the data point. Here I show the data before and after removing outliers:



Figur 4: Positions of pylons before using chauvenet



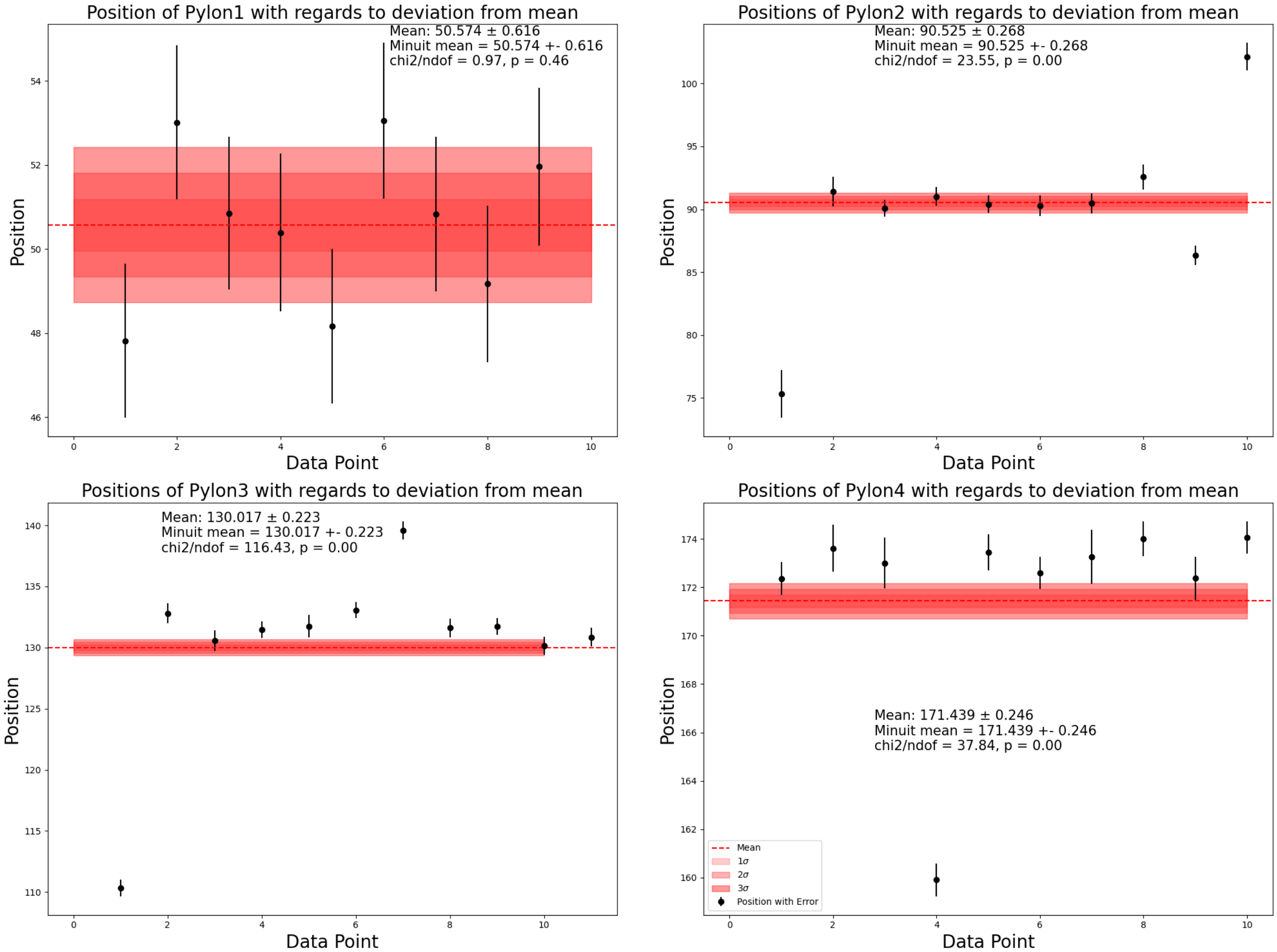
Figur 5: Positions of pylons after using chauvenet

Using the "good data" meaning the data with the outliers removed I get these positions:

$$\begin{aligned} \text{Pylon1 position:} \quad & 50.75 \pm 0.47 \\ \text{Pylon2 position:} \quad & 91.30 \pm 0.12 \\ \text{Pylon3 position:} \quad & 131.47 \pm 0.16 \\ \text{Pylon4 position:} \quad & 172.95 \pm 0.14 \end{aligned}$$

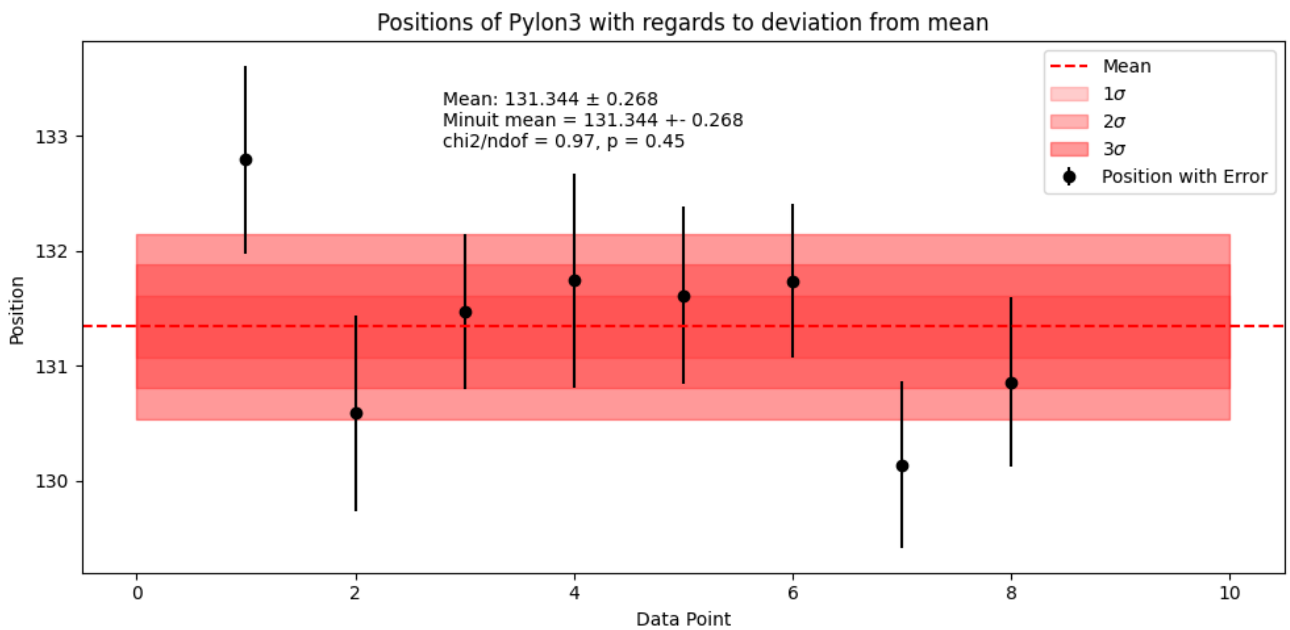
2.3.2 Using measurements with uncertainty, determine the four pylon positions

For the pylons with uncertainties, I assumed the same mask for the positions as I used in the previous question. To determine any outliers, I calculated a weighted average of each position and then plotted their respective position as a function of the data-point index. This was used to compare it to the average with a three sigma deviation, to easy identify any points outside of a three sigma range. These points was then removed. Futhermore I calculated the weighted mean, as well a fitted using minuit, as seen in the following graphs.



Figur 6: Pylon position as a function of index with their respective errors

For Pylon 1 it seems that all points are withing 3 sigma of the mean, so I will not remove any points. For the others I will start removing the furthest points and see if the change to the mean will affect the other points.I will do this iteratively until I have removed all the outliers. This can also be seen since we have a low p-value and a high reduced chi-squared for all pylons except pylon 1.



Figur 7: Positions of pylon 1 after I removed outliers

In figure 7 is the remaining positions after removing outliers. Same process was done for the 3 other pylons, seen in Appendix 6. We see a much better value for both the reduced chi-square and the p-value for all pylons. The positions for the pylons with uncertainties:

$$\begin{aligned} \text{Pylon1 position:} &\quad 50.57 \pm 0.62 \\ \text{Pylon2 position:} &\quad 90.7 \pm 0.3 \\ \text{Pylon3 position:} &\quad 131.34 \pm 0.27 \\ \text{Pylon4 position:} &\quad 173.19 \pm 0.26 \end{aligned}$$

2.3.3 Combine the two measurements. Do they match each other?

To check whether or not the values agree with one another, I will peform a two sided t-test to see if the means are the same. And if they are, I will combine them in a weighted average to get the final position of the pylon:

$$\begin{aligned} \text{Test for Pylon1:} &\quad 0.23,\ p = 0.41 \\ \text{Test for Pylon2:} &\quad 1.84,\ p = 0.06 \\ \text{Test for Pylon3:} &\quad 0.41,\ p = 0.35 \\ \text{Test for Pylon4:} &\quad -0.82,\ p = 0.22 \end{aligned}$$

$$\begin{aligned} \text{Pylon1 combined position:} &\quad 50.69 \pm 0.37 \\ \text{Pylon2 combined position:} &\quad 91.21 \pm 0.12 \\ \text{Pylon3 combined position:} &\quad 131.44 \pm 0.13 \\ \text{Pylon4 combined position:} &\quad 173.00 \pm 0.12 \end{aligned}$$

As seen they all yield an acceptable p-value, which mean I cannot reject the Null-Hypothesis that they are from the same distribution, I therefore combined as seen in the above table.

2.3.4 Test if the four measured pylon positions are equidistant.

Here I simply calculated the distance between each pylon and used errorpropagation to also have an error on the distance.

- Distance between pylons: [40.51939184, 40.23158347, 41.56352332] $\pm$ [0.3890923, 0.17684916, 0.18285625]
- Average distance between pylons: 40.84 $\pm$ 0.12
- Test for distance between Pylon 1 & 2 agreeing with the average distance: -0.79, $p = 0.26$
- Test for distance between Pylon 2 & 3 agreeing with the average distance: -2.85, $p = 0.05$
- Test for distance between Pylon 3 & 4 agreeing with the average distance: 3.30, $p = 0.04$

To check if the pylons are equidistant, I performed a 1 sided t-test to see if the how much the differ from the mean, as if they were truly equidistant they should all be within a few sigma of the mean, with a p-value above 0.05. I see that for the last Pylons (3 to 4) they do not seem to agree with the mean, and therefore, from this data I can reject the Null hypothesis that they are equidistant.

2.4 The pylon distance should be 40.83m with a tolerance (i.e. maximally allowed deviation) of 1.05m. Do the pylon positions live up to this requirement?

I perform the same test, as in the above question, just with the given pylon distance with its tolerance of 1m

- Pylons distance should be: 40.84 $\pm$ 1.05
- Test for distance between Pylon 1 & 2 agreeing with the average distance: -0.29, $p = 0.40$
- Test for distance between Pylon 2 & 3 agreeing with the average distance: -0.57, $p = 0.31$
- Test for distance between Pylon 3 & 4 agreeing with the average distance: 0.68, $p = 0.28$

Here I clearly see all p-values are above a threshold of 0.05, and I therefore I cannot reject the Null-hypothesis that they live up to the requirement set.

3 Simulation / Monte Carlo

3.1 (8 points) Circles A and B are centered at (0,0) and (3,7) and have radii of 6 and 4, respectively.

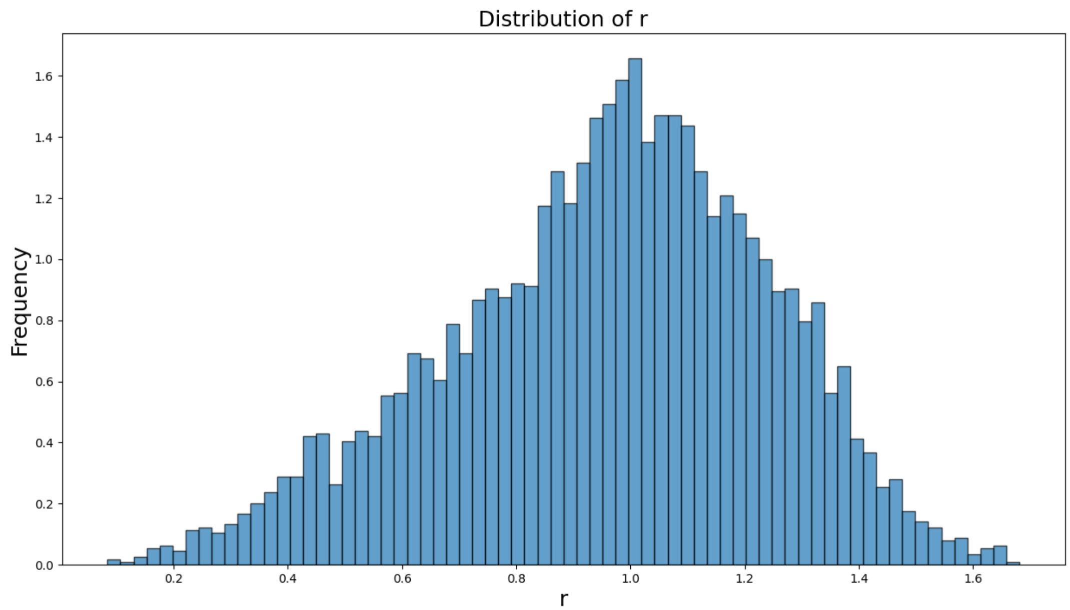
3.1.1 What fraction of A overlaps with B? And conversely, what fraction of B overlaps with A? and If the circles were 4D 'hyperballs' centered at (0,0,0,0) and (3,7,-1,2), respectively, and with the same radii, what would the answers then be?

I will make a box around both circles with dimensions (13, 17), as that will precisely cover both boxes. Then I will run a monte carlo simulation and generate 1000 points for each simulation that is uniformly distributed in the box. I will then check if the point is within the each circle and how many points are within both circles. Dividing this the sum of points inside each cirlce, will give me the fraction A overlaps B and vice verca. Same method can be used in 4 dimensions, of course just with a box of dimension (13,17,12,12). Below I calculated the overlap from A on B and vice versa with the standard deviation on the mean.

$$\begin{aligned} &\text{In 2D:} \\ \text{Mean overlap of A on B:} \quad & 9.085\% \pm 0.006\% \\ \text{Mean overlap of B on A:} \quad & 20.449\% \pm 0.012\% \\ &\text{In 4D:} \\ \text{Mean overlap of A on B:} \quad & 1.461\% \pm 0.004\% \\ \text{Mean overlap of B on A:} \quad & 7.39\% \pm 0.02\% \end{aligned}$$

3.2 (15 points) You want to simulate the radial material distribution m(r) from a uniform explosion

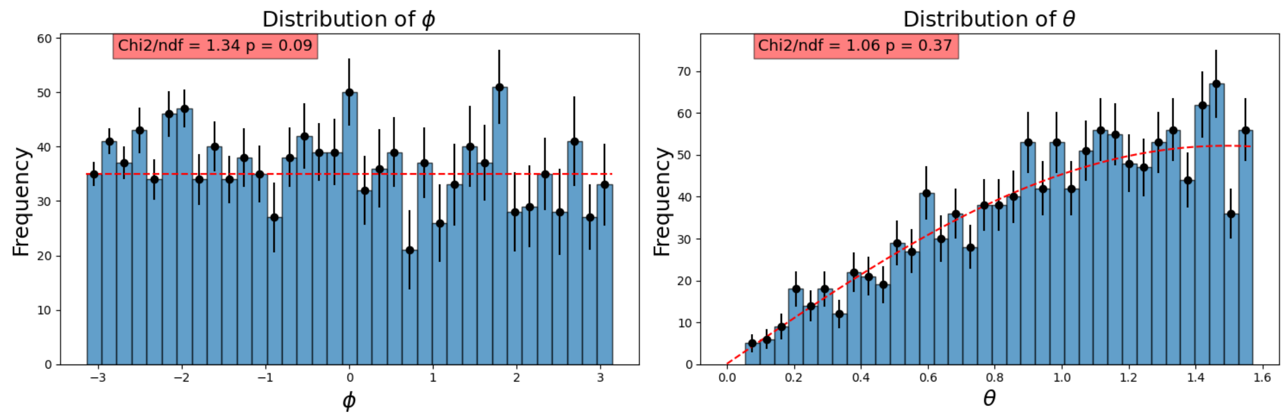
I will here use Monte-Carlo simulations to simulate the radial maetrial distribution. Letting $r = \sqrt{(x^2 + y^2 + z^2)}$. I assume that it is uniform distributed within a cube of 2x2x2.



Figur 8: Distribution of r, showing the density of points at different radial distances

3.2.1 Selecting only points with z > 0 and r < 1, what distributions in θ and ϕ do you obtain?

Here I use that $r = \sqrt{(x^2 + y^2 + z^2)}$ and $\theta = \arccos(\frac{z}{r})$ and $\phi = \arctan 2(y, x)$ to convert the cartesian coordinates to spherical coordinates.



Figur 9: Distribution of ϕ and θ, fitted to a uniform and sinusoidal distribution respectively

The distribution of ϕ is uniform, as expected as it is calculated from x and y and both are uniform. The distribution of θ is not uniform, as it is calculated from z and r, and r is not uniform and seems to follow a sinusoidal, furthermore it's limited to the range of $[0, \pi/2]$, since it's only z > 0. Both of these are distributions have given a reasonable p-value and therefore I can say that they are uniform and sinusoidal respectively.

3.2.2 How would you produce random velocities v according to $f(v) = \left(\frac{v}{v_0}\right)^2 \exp\left(-\frac{v}{v_0}\right)$?

Since I can assume that v_0 is a constant and let $x = \frac{v}{v_0}$ I can rewrite the equation:

$$f(x) = x^2 e^{-x} \tag{6}$$

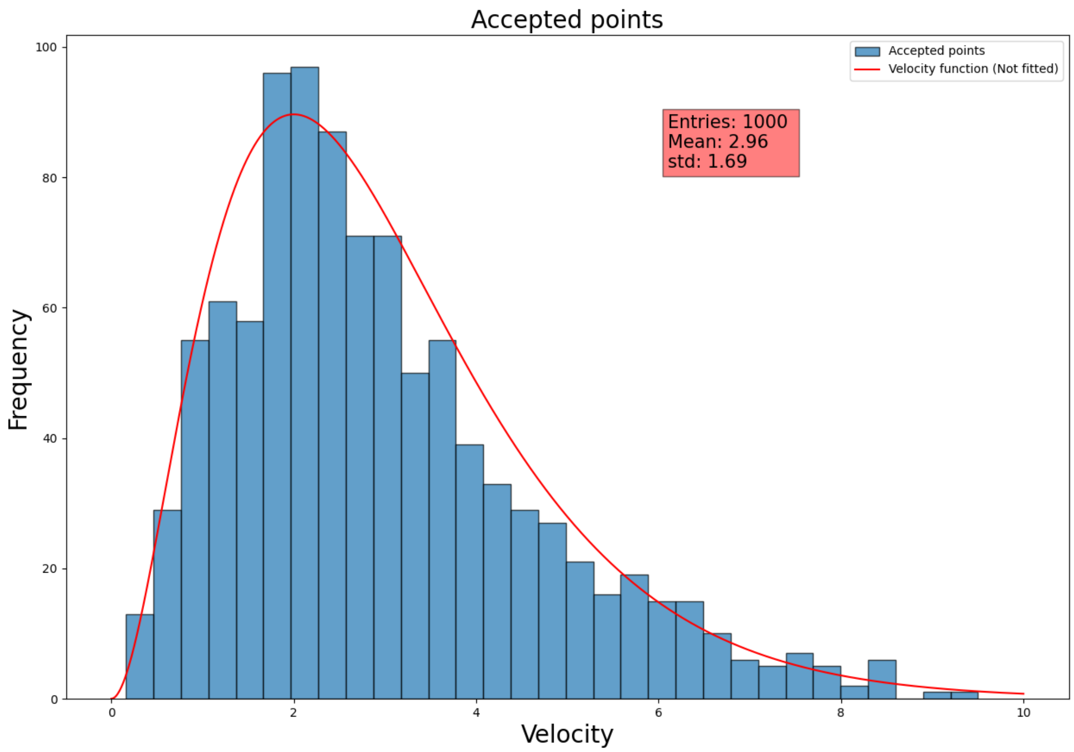
I can then find the CDF by integrating the function from 0 to x and inverting it, to use the transformation method to generate random velocites from the given PDF.

$$F(x) = \int_0^x f(x)dx \tag{7}$$

Solved using maple I get:

$$F(x) = 2 + (-x^2 - 2x - 2)e^{-x} \tag{8}$$

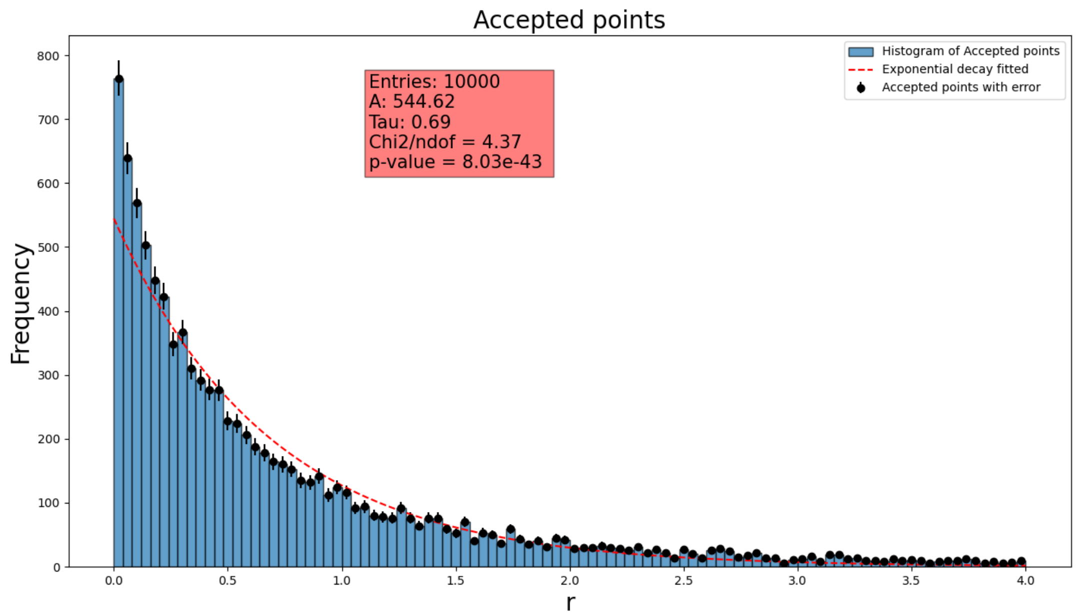
Looks like it's not possible to use the transformation method to generate random velocities from the given PDF, as the inverse of the CDF is not possible to solve for, I will therefore use the accept/reject method instead. And from plotting I see the function approaches 0 at x=10, so I will use that as my range. And use the range for y from 0 to 1, as I can clearly see 1 is way above the maxium of the function! I will use 1000 points to try the accept/reject method.



Figur 10: Histogram of accepted points using the accept/reject method - Efficiency: $19.47\% \pm 0.55\%$

3.2.3 Given $v_0 = 100\,\mathbf{m/s}$ and that the radial distance of material r as a function of velocity is $r(v) = \frac{\sin(\theta)v^2}{g}$ ($g = 9.82\,\mathbf{m/s}^2$), simulate 10000 values of θ and v. Combine these to obtain values of r, and plot the resulting distribution $m(r)$.

Here I again used the accept/reject method, for both v and θ



Figur 11: Distribution of r, from random generated v and θ. An exponential was fitted

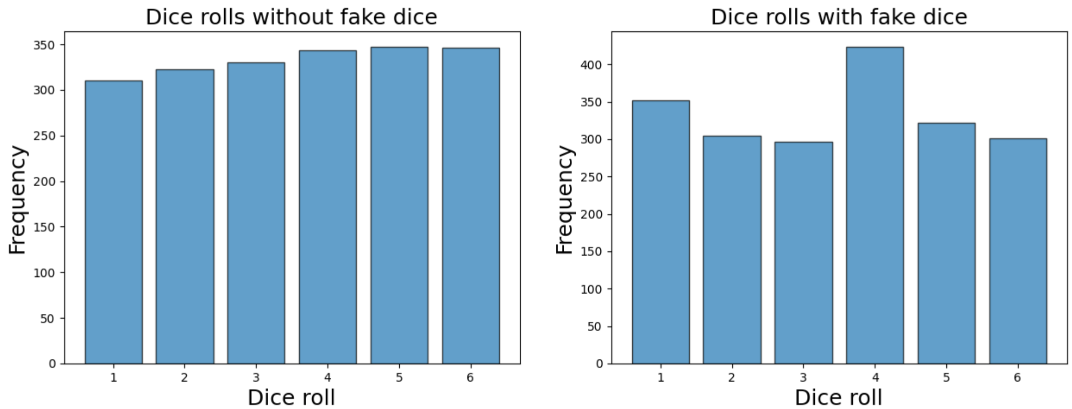
On the figure 11 we see a relatively bad p-value and I didn't not succeed in adjusting my fit perfectly. I suspect that I should have used another model or adjusted my function as I see a high correlation between my parameters A and tau of 0.8.

4 Statistical tests

4.1 (12 points) You get a permanently closed box with 20 normal (i.e. six-sided) dices in. One of the dice is potentially fake, with all the sides having the same (unknown) value. You can shake the box and see the resulting 20 dice roll as many times as you like

4.1.1 Simulate 100 box rolls and plot the die frequencies, both with and without a fake die in.

To do this I simulated of 100 box rolls, one without a fake dice, and one with a fake dice, that was chosen randomely before the simulation began and then added on for every roll.



Figur 12: Distribution of dice rolls over 100 box rolls

For 100 box rolls, you can more or less clearly tell that dice face 4, has been rolled more times, for the one with a fake dice in it.

4.1.2 For both of your simulated datasets, test if there is a fake die or not.

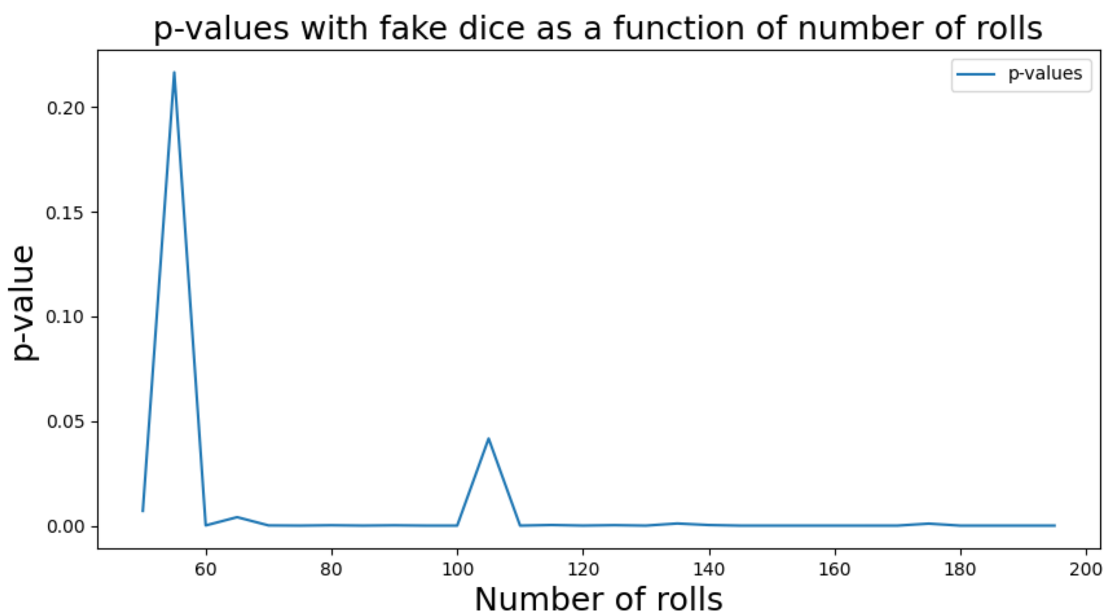
To test this I will use a chi2-squared test to see whether the dice rolls are consistent with being uniform, as they should if the dice is fair. I will use a significance level of 0.05. Since the chi2-squared test, test for the whole distribution compared to an expected distribution it didn't tell exactly which dice was the outlier. Therefore I also used a binomial test to test each dice number individually to see if they matched the expected distribution. I used a significance level of 0.05 for this test as well.

p_value for dice 1:	0.17
p_value for dice 2:	0.55
p_value for dice 3:	0.86
p_value for dice 4:	0.53
p_value for dice 5:	0.42
p_value for dice 6:	0.45
p_value for fake dice 1:	0.27
p_value for fake dice 2:	0.09
p_value for fake dice 3:	0.03
p_value for fake dice 4:	0.00
p_value for fake dice 5:	0.51
p_value for fake dice 6:	0.05
Chi2 for dice rolls:	3.37, $p = 0.64$
Chi2 for dice rolls with fake dice:	35.06, $p = 0.00$

Both the binomial test and chi-squared test, resulted more or less as expected, where for the normal dice, there are more or less no outliers or deviants, whereas for the one with a fake dice, we clearly see a very low p-value for dice 4. The Chi-squared test also shows a high chi-square value and low p-value for for the fake dice. Therefore, for the distribution with fake dice, I can reject the Null-hypothesis that the distribution of dice numbers is uniform.

4.1.3 How many rolls would you require before you would argue, that you could tell the difference?

I perform a simulation that will run over a number of rolls, starting from 50 and increasing by a stepsize of 5. For each number of rolls, I will calculate a p-value using the chi2-squared test and see when the p-value is below 0.05 for the one with the fake dice. I will then use the number of rolls that gave a p-value below 0.05 as the number of rolls needed to reject the null hypothesis that the dice is fair. Furthermore I also plot the p-value as a function of the number of rolls.



Figur 13: P-values as a function of Number of Rolls

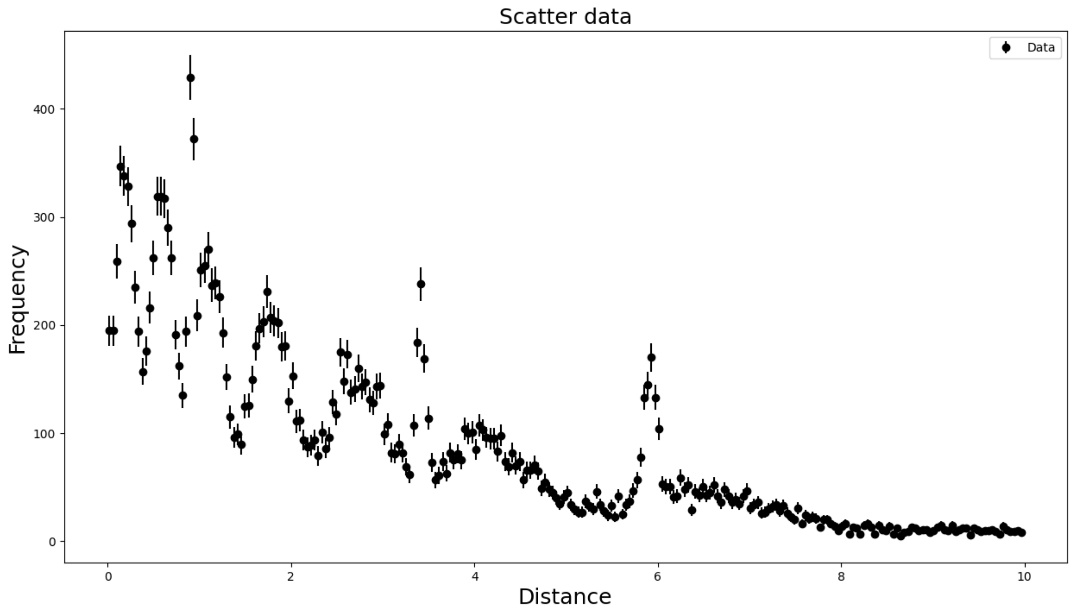
Strictly from look at figure 13, I see that around 60 Box rolls, the p-value flattens out. However seems like something went wrong with how I coded the problem. I would suspect that the number of rolls needed should be between 50 and 100.

5 Fitting data

5.1 The file InconstantBackground.csv contains molecular interspacing measurements d (in nm) from a scattering experiment

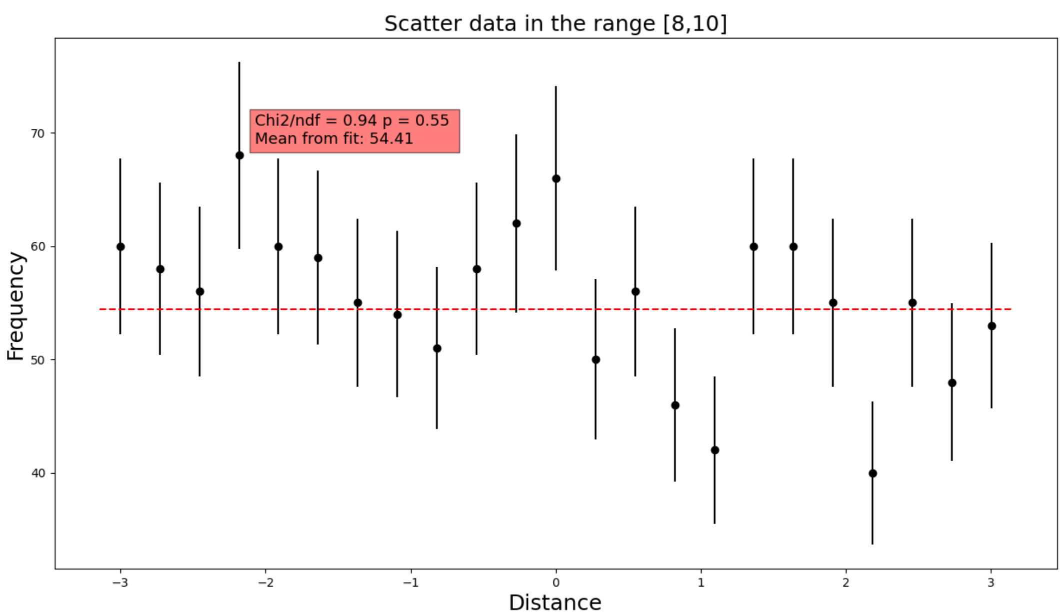
5.1.1 Plot the data and test to what extend the background in the range [8,10] is uniform.

I will quickly plot the data, assuming poission errors on the counts of my bins. And then I will fit a line to the data, using a least square fit in the range [8,10] to check if the data is uniform.



Figur 14: Data of interspacing measurments in nm

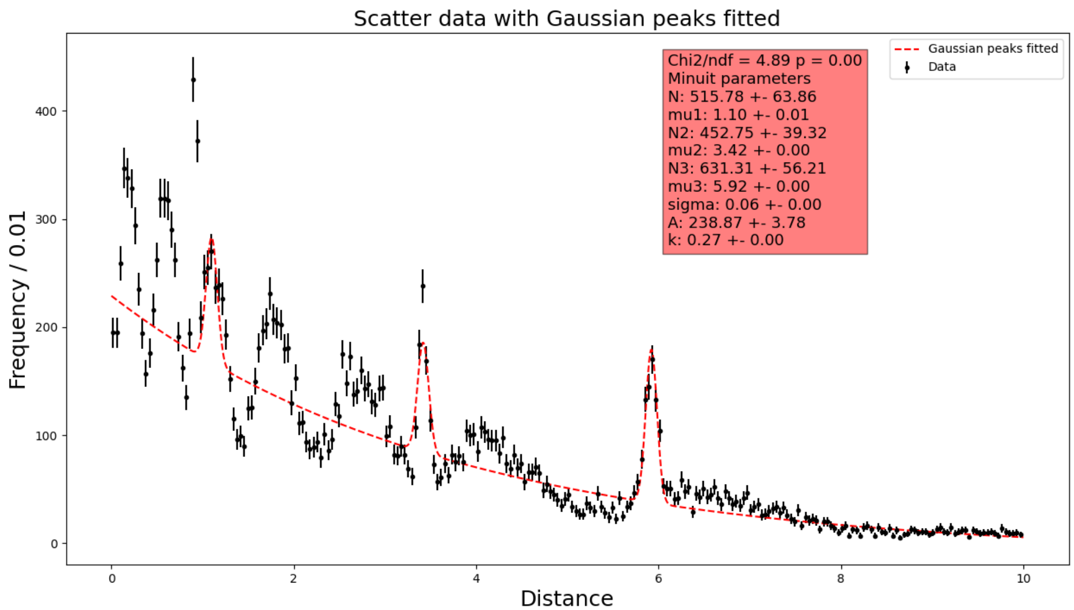
I now only look in the range [8,10] and try and fit a linear function to it.



Figur 15: Data of interspacing measurments in nm with a straight line fitted

With a p-value of 0.55 and a good chi2/ndof I can say that the data is consistent with being uniform, since I cannot reject the Null Hypothesis that the data in the range [8,10] is uniform, and the mean of the freqeuency is seems consistent with the plot.

5.1.2 Fit the three Gaussian peaks at around d = 0.9, 3.4, and 5.9 nm, including local background.



Figur 16: three Gaussian peaks fitted at around d = 0.9, 3.4, and 5.9 nm

From figure 16 I see that there is some deviation around the peak at 1, which makes it hard to fit a guassian peak. It looks like from the data, that two peaks are very close to each other, with one peak deviating a lot from the rest of the data. I first tried to calculate some z-values for the range [0,2] and remove any outliers that way, but that didn't seem to work. There is also some high correlation between the parameters A and k, but below 0.85, so I will keep them both in the fit. I ended up masking the range of the fit, to ensure I fitted to the peak visible in the data. All the parameters from Minuit looks reasonable and as expected.

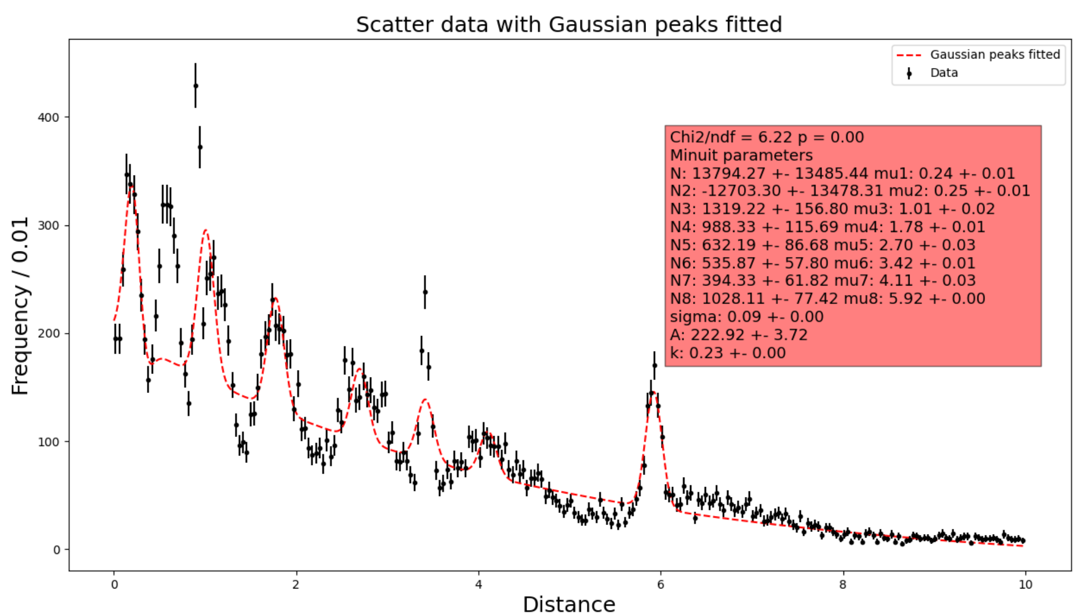
5.1.3 Test if the peaks have the same intensity, i.e. number of measurements in them.

To do this I will simply set some bounds for the peaks using the parameters from the fit (The mean and the standard deviation), to give me a number of points in the peaks and using the errors for the peaks given from minuit. Afterwards I will perform a z-test to see if the peaks are significantly different from each other. I will use a significance level of 0.05. Furthermore I will do a chi2-squared test to see if the data is consistent with having equal amount of entries in each, compared to the total amount of entries in the 3 peaks. I will use a significance level of 0.05.

Total entries:	21933
Entries in peak 1:	785 ± 64
Entries in peak 2:	602 ± 39
Entries in peak 3:	465 ± 56
Chi2 for the entries in the peaks:	14.39, $p = 0.00$
z_{12} :	2.44, $p = 0.01$
z_{13} :	3.76, $p = 0.00$
z_{23} :	2.00, $p = 0.02$

As seen both they z-test and the chi2-squared test gives a p-value below 0.05, so I can reject the null hypothesis that the peaks are the same and that the data is consistent with having equal amount of entries in each peak.

5.1.4 Try to fit the entire spectrum or parts of it best possible and comment on your results.



Figur 17: Fitting of all peaks

The reduced chi2 of 6.22 most likely indicates a significant deviation from the model used, which is also seen in figure 17, as one of the peaks was not fitted. Some improvements that could have been made, is to revaluate some of the outliers, as they seem to impact the fit for the second peak seen in the figure. Maybe even try to look at overlapping gaussians, as the peaks are very close to each other, and therefore their parameters are correlated, as I can see that μ_1 and μ_2 are highly correlated (0.994) and the same for N and N1 (first 2 peaks) with a correlation of -0.993. For future reference I would have rewritten the function for the highly correlated peaks, dividing each peak into a fraction of the total gaussian, ensuring less correlation between parameters. I also assumed equal standard deviations, which is unlikely looking at figure 17

5.2 5.2 (14 points) The table below lists the North-South alignment of Egyptian pyramids (in arc minutes).

Pyramid	1.Meidum	2.Bent	3.Red	4.Khufu	5.Khafre	6.Menk.	7.Sahure	8.Nefer.
Align. year	2600 BC	2583 BC	2572 BC	2554 BC	2522 BC	2489 BC	2446 BC	2433 BC
East Align.	-20.6±1.0	-17.3±0.2	-8.7±0.2	-3.4±0.2	6.0±0.2	12.4±1.0	23±10	30±10
West Align.	-18.1±1.0	-11.8±0.2	–	-2.8±0.2	6.0±0.2	14.1±1.8	–	–

5.2.1 Test to what extend the East (E) and West (W) alignment values agree.

For the ones where there is no alignment in the west alignment, I have replaced the NaN with 0 and skip them in the test coming up. I will then use the t-test, as we have a low number statistics, with the amount of data points, to see if the East and West alignment values agree with each other. I will use a significance level of 0.05.

$$\begin{aligned} \text{Meidum:} &\quad z = -1.77,\ p = 0.04 \\ \text{Bent:} &\quad z = -19.45,\ p = 0.00 \\ \text{Red:} &\quad \text{Has no data for West alignment} \\ \text{Khufu:} &\quad z = -2.12,\ p = 0.02 \\ \text{Khafre:} &\quad z = 0.00,\ p = 0.50 \\ \text{Menk:} &\quad z = -0.83,\ p = 0.20 \\ \text{Sahure:} &\quad \text{Has no data for West alignment} \\ \text{Nefer:} &\quad \text{Has no data for West alignment} \end{aligned}$$

The only two pyramids, where I cannot reject the Null-Hypothesis that the two alignments are in agreement is for Khafre and Menk. The rest I reject the Null-hypothesis that the two alignments are in agreement. A case could be made for Meidum and Khufu, as their are both within three sigma.

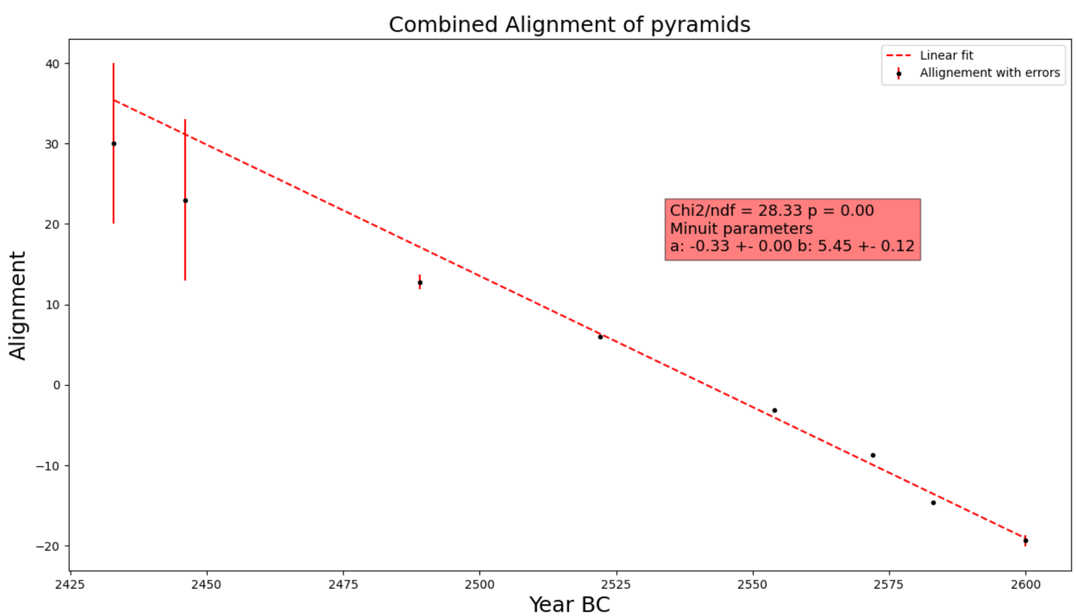
5.2.2 Combine East and West values. Include systematic uncertainties to ensure agreement.

I will use a weighted average for each of the pyramids to include the systematic uncertainties and find the combined alignment with errors. Noting here that some of the pyramids only have an alignment to the East and it therefore counts as its combined alignment.

$$\begin{aligned}
\text{Meidum:} &\quad -19.35 \pm 0.71 \\
\text{Bent:} &\quad -14.55 \pm 0.14 \\
\text{Red:} &\quad -8.70 \pm 0.20 \\
\text{Khufu:} &\quad -3.10 \pm 0.14 \\
\text{Khafre:} &\quad 6.00 \pm 0.14 \\
\text{Menk:} &\quad 12.80 \pm 0.87 \\
\text{Sahure:} &\quad 23.00 \pm 10.00 \\
\text{Nefer:} &\quad 30.00 \pm 10.00
\end{aligned}$$

5.2.3 If the alignments were done using circumpolar stars, these drift with Earth's precession. Test if the alignment of the pyramids shifts linearly as a function of time.

Here it is important to note that since my values of x (year) is far from 0, it will most likely cause a lot of correlation between the a and b (parameters), as a small change in a, will cause a big change in b. To counter this I will substract the average year from x.



Figur 18: Combined alignment of pyramids as a function of time, fitted with a linear model. Noting here that the x-axis is years BC, therefore it goes from latest years to earliest

The model seems to be able to fit to the model, though with a very high reduced chi-square and low p-value. In Appendix 6 I also fitted just using the West and East alignments, where the West alignment gave a good p-value of 0.24, seeming to tell me, for at least the west alignment, I cannot reject the Null-hypothesis that the alignments of the pyramids shifts linearly as a function of time.
A lot of the points lie outside of a 3 sigma range, as to why I also get a very low p-value (Could be a

underestimation of uncertainties, if it is truly linear). I could look into removing some outliers, but as I deem all points relevant, with the low number of data points, I will keep them in the fit.

5.2.4 The astronomically predicted shift as a function of time is 2467 BC + 0.274 arc min./year. Does the slope of the linear fit match the predicted value?

The slope of my fits should be equal to the shift as a function of time, so I will compare the 3 slopes to the real value and evaluate using a z-test and chi2-test to see if they are consistent with each other. I will use a significance level of 0.05.

$$\begin{aligned} &\text{z-test} \\ \text{Combined shift:}\quad & z = 17.32,\ p = 0.00 \\ \text{East shift:}\quad & z = 18.56,\ p = 0.00 \\ \text{West shift:}\quad & z = 4.09,\ p = 0.00 \end{aligned}$$

$$\begin{aligned} &\text{chi2-test} \\ \text{Combined shift:}\quad & \chi^2 = 299.87,\ p = 0.00 \\ \text{East shift:}\quad & \chi^2 = 344.34,\ p = 0.00 \\ \text{West shift:}\quad & \chi^2 = 16.70,\ p = 0.00 \end{aligned}$$

None of them are within 3 sigma of the real value, but it is clearly to see the deviation comes from the east shift, as the west shift by itself is almost within 4 sigma of the real value, so an argument could be made that it is consistent to some degree with the real value.

5.2.5 What alignment date of Khufu (historically 2554 ± 100 BC) is the astronomical prediction?

I will test the value of Khufu's allignment both for the combined shift and the west and east seperatly to see which one is the most consistent with the real value. I will use a z-test and chi2-test to see if they are consistent with each other. I will use a significance level of 0.05. Furthermore I will use both the value of my fit and the given value in the assignemnt .

$$\begin{aligned} \text{Khufu combined:}\quad & 2551.09 \pm 0.44 \\ \text{Khufu east:}\quad & 2492.59 \pm 0.56 \\ \text{Khufu west:}\quad & 2494.78 \pm 0.72 \\ \text{Khufu real:}\quad & 2493.69 \pm 0.55 \end{aligned}$$

It seems I'm getting really low uncertainties on the values, which is a bit odd, as I would expect them to be higher. This can make it a bit harder to compare to the historical value. Instead I will use the uncertainty on the historical value as the "observed" value, to use for the test's and see if they are consistent with each other. I realize this is not optimal, as I suspect my uncertainties should have been much higher on the year of alligmenent. I will use a significance level of 0.05.

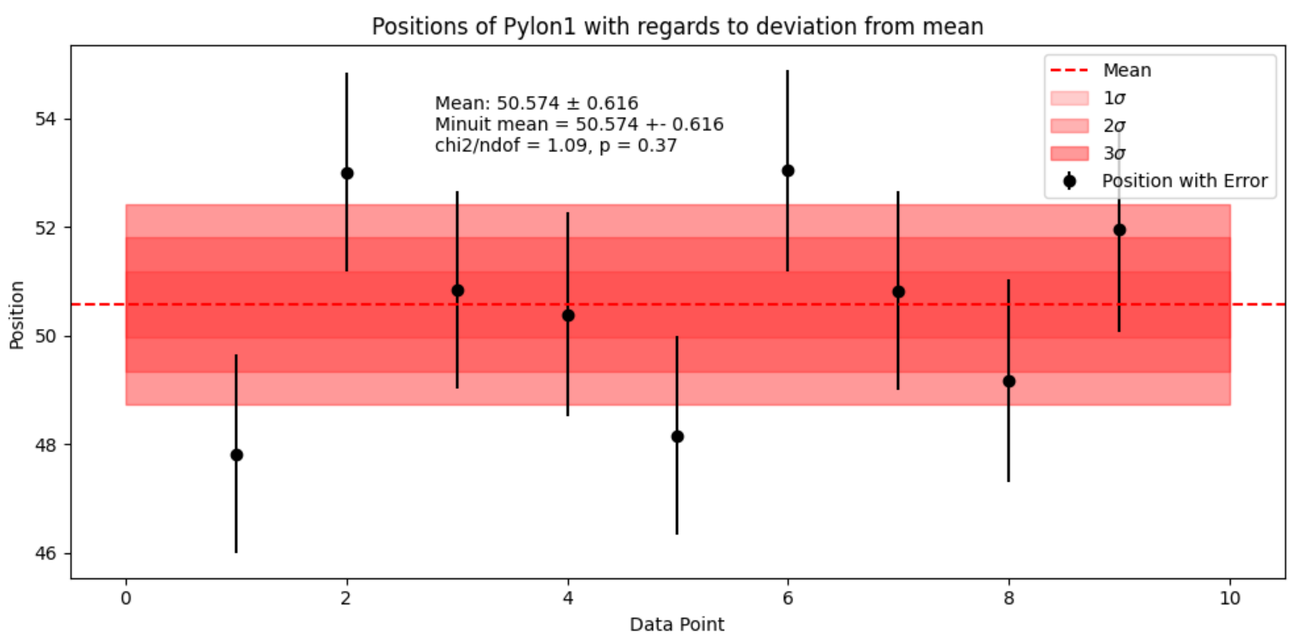
$$
\begin{aligned}
&\text{Chi2 test} \\
\text{Combined:}\quad & \chi^2 = 0.00,\ p = 0.98 \\
\text{East:}\quad & \chi^2 = 0.38,\ p = 0.54 \\
\text{West:}\quad & \chi^2 = 0.35,\ p = 0.55 \\
\text{Real:}\quad & \chi^2 = 0.36,\ p = 0.55
\end{aligned}
$$

$$
\begin{aligned}
&\text{Z test} \\
\text{Combined:}\quad & z = 0.03,\ p = 0.49 \\
\text{East:}\quad & z = 0.61,\ p = 0.27 \\
\text{West:}\quad & z = 0.59,\ p = 0.28 \\
\text{Real:}\quad & z = 0.60,\ p = 0.27
\end{aligned}
$$

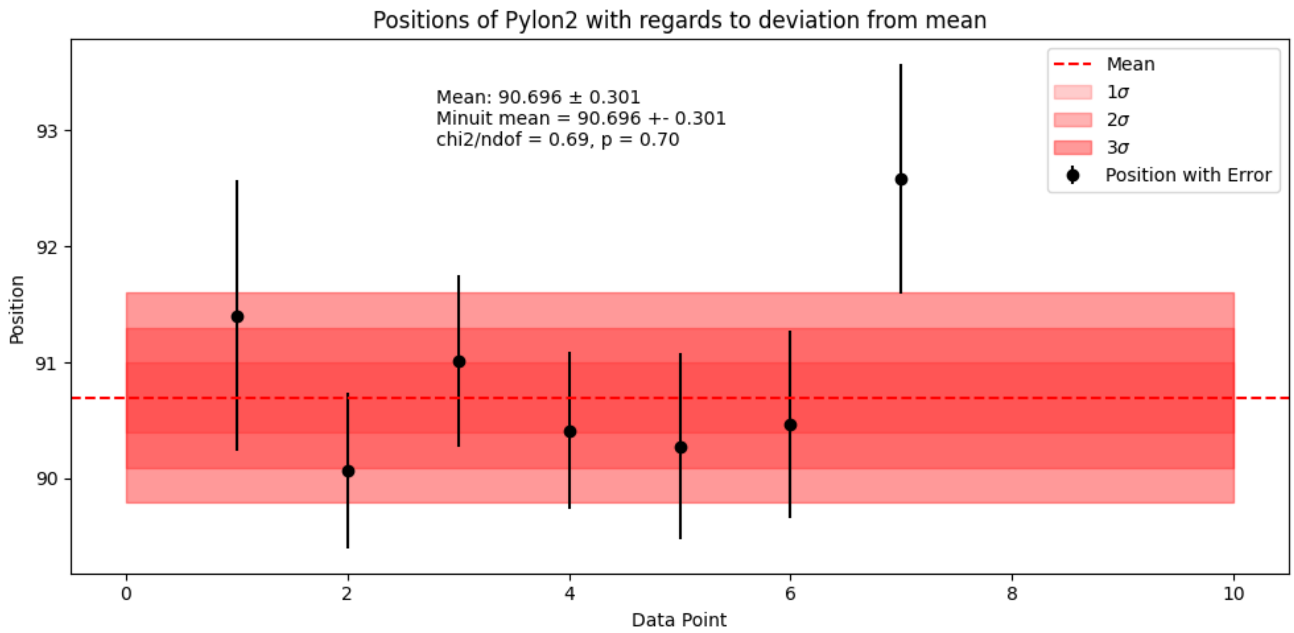
Doing it this way will of course give me some good p-values, but still can be used to see how the different values compare, as here it looks like the combined value is the most consistent with the historical value. For chi2 test I even get a very high p-value which could indicate overestimated errors on the historical value, if you were to solely look at the combined aligment compared to the historical value.

6 Appendix

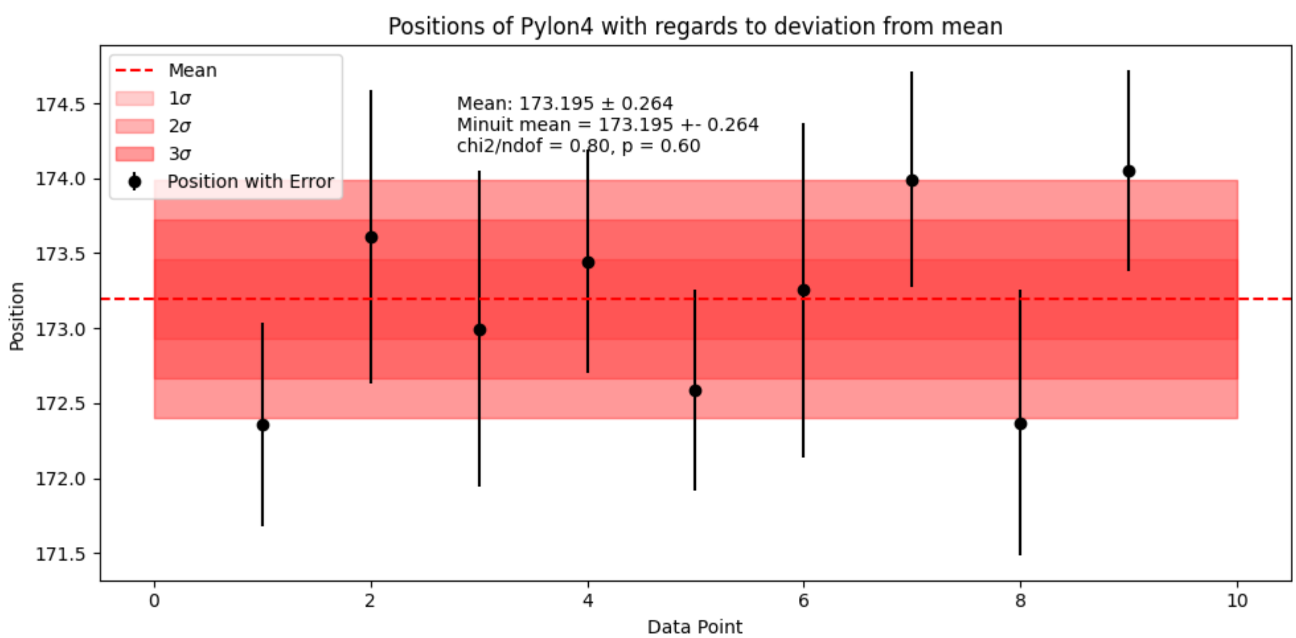
2.3



Figur 19: Positions of pylon 1 after I removed outliers



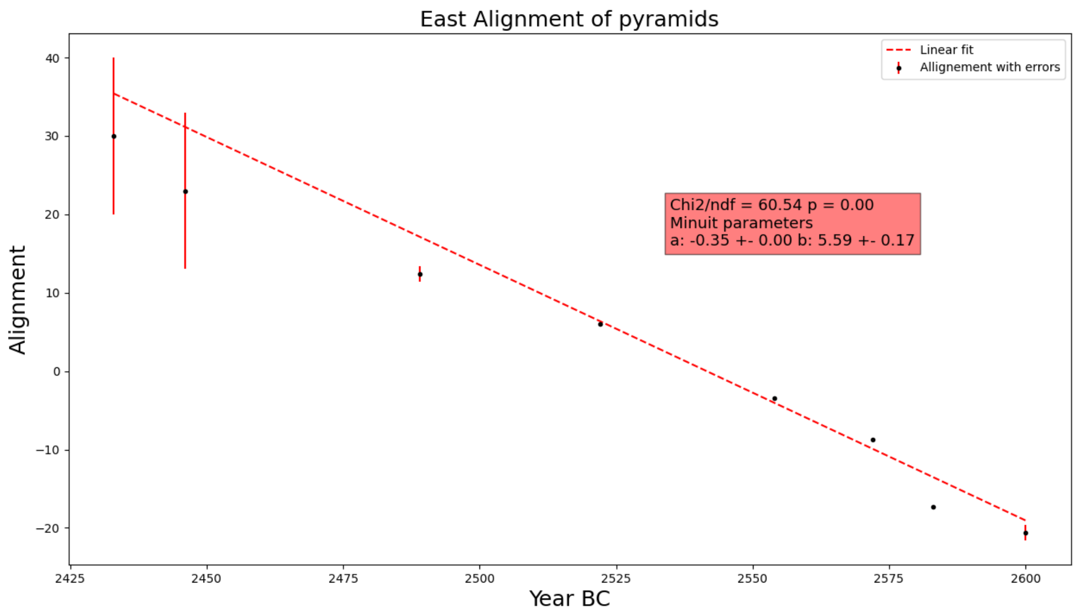
Figur 20: Positions of pylon 2 after I removed outliers



Figur 21: Positions of pylon 4 after I removed outliers

5.2

The outcome of a repeated process follows not chance but statistics.



Figur 22: East alignment of pyramids as a function of time