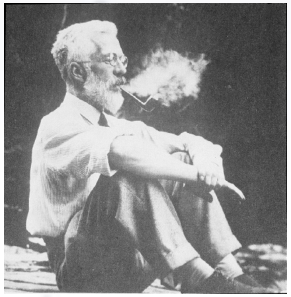
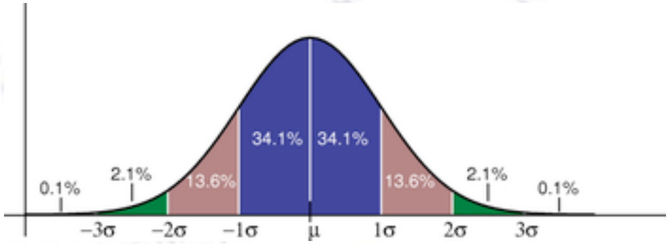
Applied Statistics

Problem Set - Solutions & Comments



Troels C. Petersen (NBI)



"Statistics is merely a quantisation of common sense"

Quantify!!!!

Problem 1.1

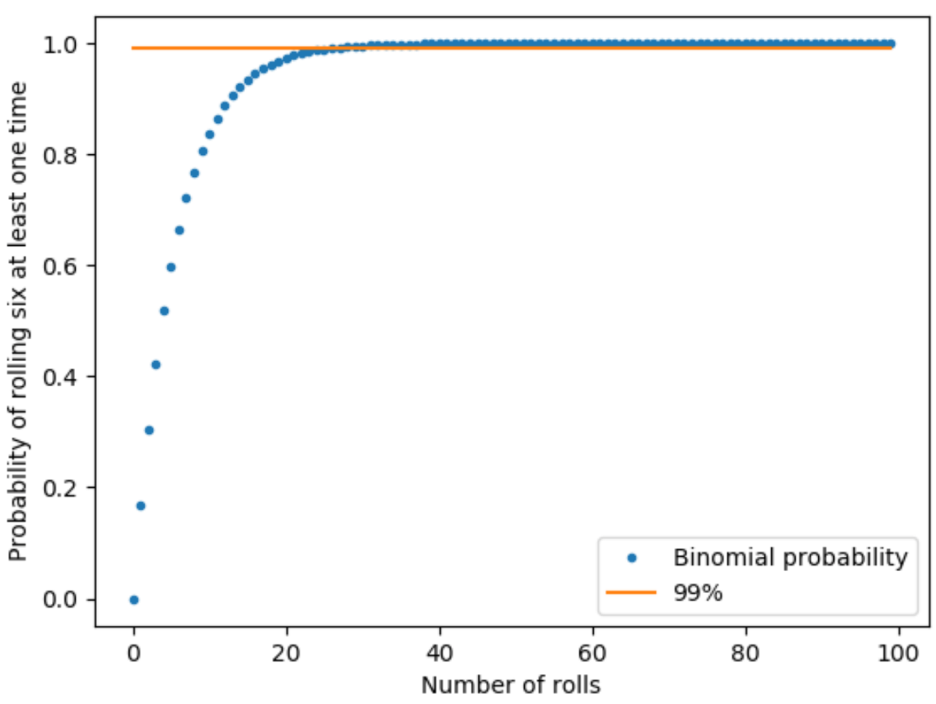
This is a probability calculation, which can be solved in two ways:

Using the formula:

$$0.01 > (1 - 1/6)^N \implies N ln(5/6) < ln(0.01) \implies N > ln(0.01)/ln(5/6) = 25.2585$$

Using the Binomial distribution and choosing N such that p(Nhit = 0) < 0.01.

$$p_{\text{Binomial}}(N = 26, p = 1/6, N_{\text{succes}} = 0)$$



$$= 0.008735$$

Problem 1.1

This is a probability calculation, which can be solved in two ways:

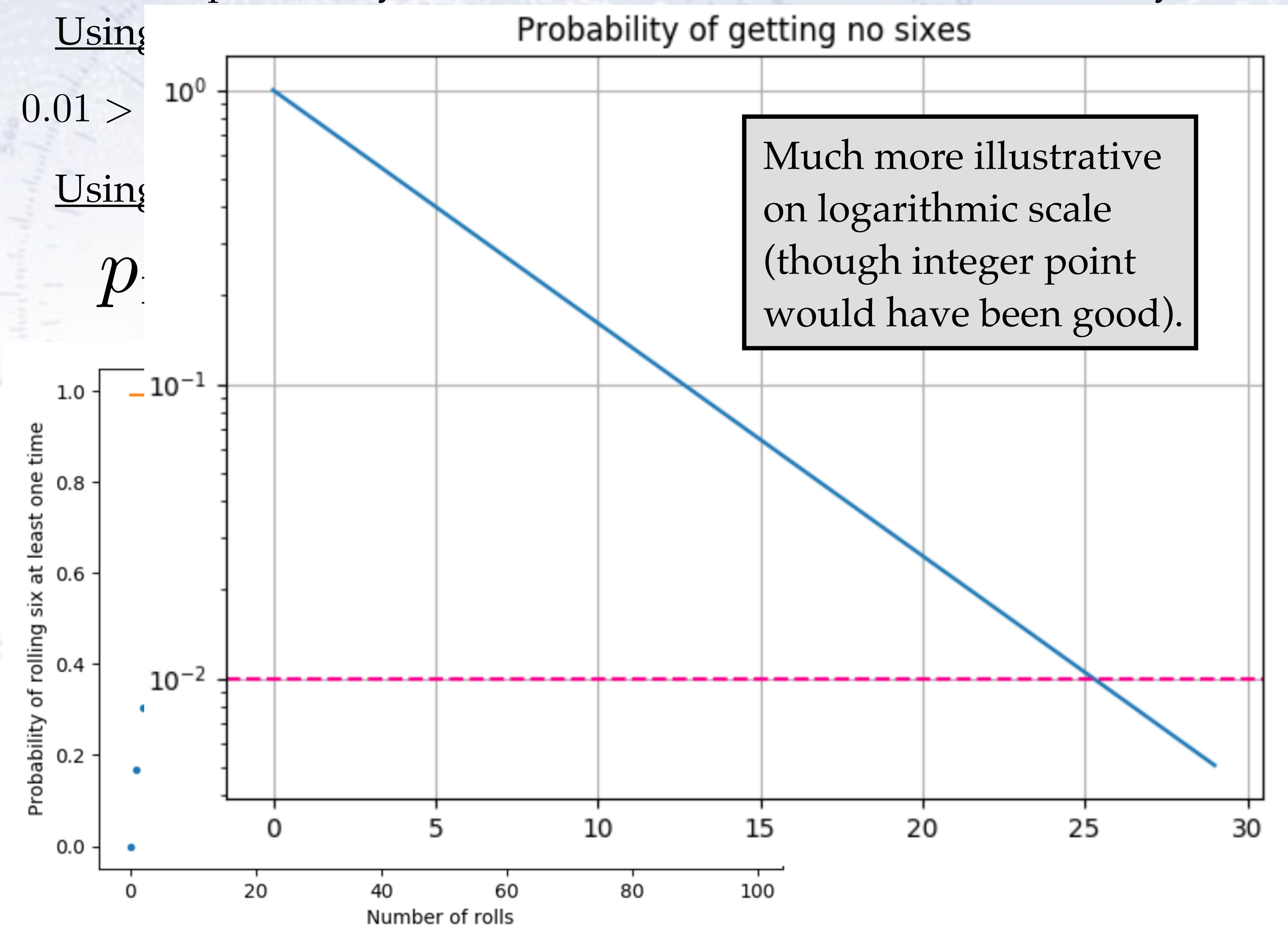
Using

$0.01 >$ … 25.2585

Using … $0.01.$

p … $))$

… 35



Problem 1.2

The mean should have been clear from just reading the problem, while the width is harder to obtain. Below is the (beautifully laid out) "full" solution:

$$z = \frac{X - \mu}{\sigma}$$

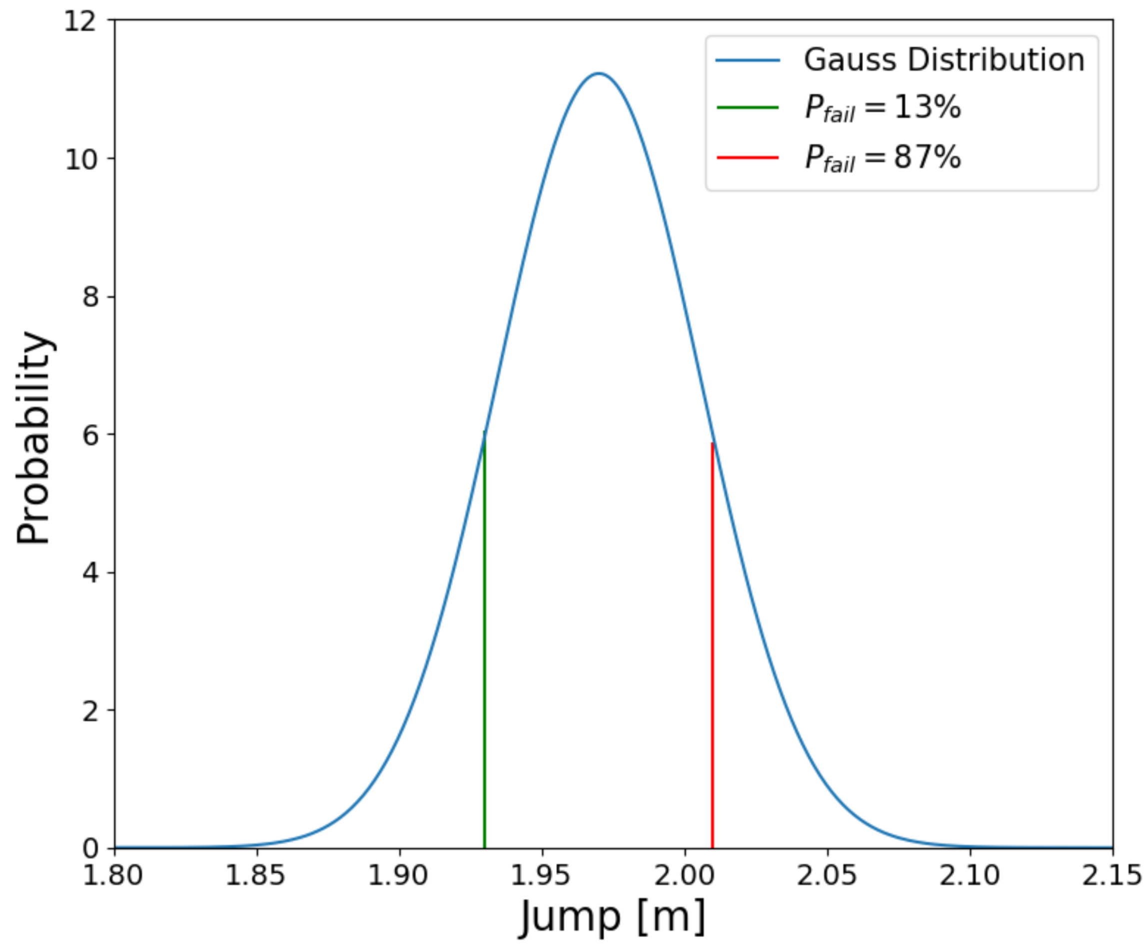
$$-1.1 = \frac{1.93m - \mu}{\sigma}$$

$$1.1 = \frac{2.01m - \mu}{\sigma}$$

$$\sigma = \frac{1.93m - \mu}{-1.1}$$

$$\mu = 1.97m$$

$$\sigma = 0.04m$$



Problem 1.3

Bayes' Theorem

$$P_{rain} = 14/365$$

$$P_{Forecast}(Rain) = 0.8$$

$$P_{Forecast}(Sun) = 0.1$$

And adapting Bayes' Theorem such that

$$P_{Rain}(Forecast) = \frac{P_{Forecast}(Rain) * P_{Rain}}{P_{Forecast}}$$

We only miss

$$P_{Forecast} = P_{Forecast}(Sun) * PSun + P_{Forecast}(Rain) * P_{Rain} = 0.1 * (1 - \frac{14}{365}) + 0.8 * (\frac{14}{365})$$

$$= 0.1268 \approx 0.13$$

Before we conclude with:

$$P_{Rain}(Forecast) = \frac{0.8 * \frac{14}{365}}{0.1268}$$

$$0.2420 \approx 0.24$$

NOTE: Here I do not take into account days where it isn't raining, but it isn't sunny (overcast days). I looked it up, ([1]) and found that there are actually only 269 sunny days. This would introduce some guesswork into the dataset as well, as overcast days are neither one nor the other. I would go on, but I have to submit this before noon.

Problem 1.4

- The distribution should follow a Poisson distribution. There can be a lot of plants in one km^2, so N is high, but there are only a few (in this case $7.1/\text{km}^2$), so p is low. The mean for $0.3\ \text{km}^2$ is $2.13/\text{km}^2$. Then the probability to find four or more plants is given by:

$$P = 1 - \sum_{k=0}^{3} poisson(2.13, k) = 0.167 \tag{5}$$

- The mean for species A is scaled down to $1.42/\text{km}^2$, and for species B to $2.52/\text{km}^2$. The probability to find exactly two of each species is given by:

$$P = poisson(1.42, 2) \cdot poisson(2.52, 2) = 0.062 \tag{6}$$

The probability to find in total more than four is given by:

$$\begin{aligned} P &= P_A(X=0) \cdot P_B(X \geq 5) + P_A(X=1) \cdot P_B(X \geq 4) + P_A(X=2) \cdot P_B(X \geq 3) \quad (7) \\ &+ P_A(X=3) \cdot P_B(X \geq 2) + P_A(X=4) \cdot P_B(X \geq 1) + P_A(X \geq 5) \cdot P_B(X \geq 0) \quad (8) \\ &= 0.359 \end{aligned}$$

with $P_i(X \geq a) =$

Case:	$P_A(n_A)$	$P_B(n_B \text{ or less})$	P_{case}
0 A plants, and 4 B plants or less	0.242	0.888	0.215
1 A plant, and 3 B plants or less	0.343	0.753	0.259
2 A plants, and 2 B plants or less	0.244	0.539	0.131
3 A plants, and 1 B plant or less	0.115	0.283	0.033
4 A plants, and 0 B plants	0.04	0.08	0.003
Total probability of less than four plants:			**0.641**

Table 1: The 5 different cases of finding less than four plants in total, with the probability P_{case} of each scenario, and the total probability being $P(4 \text{ plants or less}) = \sum_{i=1}^{5} P_{case,i}$.

Problem 1.4

- The distribution should follow a Poisson distribution. There can be a lot of plants in one km^2, so N is high, but there are only a few (in this case $7.1/km^2$), so p is low. The mean for $0.3\ km^2$ is $2.13/km^2$. Then the probability to find four or more plants is given by:

$$P = 1 - \sum_{k=0}^{3} poisson(2.13, k) = 0.167 \quad (5)$$

However, since adding Poisson distributions (A and B) give a new Poisson distribution with $\lambda = \lambda(A) + \lambda(B)$, it can be solved in a one-liner:
pAB_MoreThan4_Total2 = 1.0-poisson.cdf(4, (lambdaA+lambdaB) * area2)

The probability to find in total more than four is given by:

$$P = P_A(X=0) \cdot P_B(X \geq 5) + P_A(X=1) \cdot P_B(X \geq 4) + P_A(X=2) \cdot P_B(X \geq 3) \quad (7)$$
$$+ P_A(X=3) \cdot P_B(X \geq 2) + P_A(X=4) \cdot P_B(X \geq 1) + P_A(X \geq 5) \cdot P_B(X \geq 0) \quad (8)$$
$$= 0.359$$

with $P_i(X \geq a) =$

Case:	$P_A(n_A)$	$P_B(n_B$ or less)	P_{case}
0 A plants, and 4 B plants or less	0.242	0.888	0.215
1 A plant, and 3 B plants or less	0.343	0.753	0.259
2 A plants, and 2 B plants or less	0.244	0.539	0.131
3 A plants, and 1 B plant or less	0.115	0.283	0.033
4 A plants, and 0 B plants	0.04	0.08	0.003
Total probability of less than four plants:			**0.641**

Table 1: The 5 different cases of finding less than four plants in total, with the probability P_{case} of each scenario, and the total probability being $P(4 \text{ plants or less}) = \sum_{i=1}^{5} P_{case,i}$.

Problem 1 problems

General:
Bad formatting of text and code.

1.1
Rounding n to 25.

1.2
- confusing the probability with the probability density or not understanding what was meant with 'clearing a height' : mu=0.87*1.94+0.13*2.01=1.94
- reading from the table for the two tailed integral: P=87% -> z=1.5 -> sigma=0.027 m.

1.3
Some rare cases of not using bayes theorem

1.4
- thinking that p(x=3) == p(x>3) and just calculating one value from the distribution
- not multiplying area to the mean density (taking lamba=7.1). In this case some people just multiply the probability and the area of the field or take the factorial of n/area.
- confusion between p(x>4) and p(x>=4) lead some people to 55% instead than 36 % (including myself!)
- not knowing that the sum of 2 Poissonian variables follows a Poissonian.

Problem 2.1

I assume no correlations between any of these measurements. The error on the sine factors are:

$$\Delta(\sin(\theta)) = \Delta\theta \cdot \cos(\theta) \tag{10}$$

This results in $\sin(\theta_1) = 0.959 \pm 0.002$ and $\sin(\theta_2) = 0.622 \pm 0.003$. The index of refraction n_2 is:

$$n_2 = n_1 \cdot \frac{\sin(\theta_1)}{\sin(\theta_2)} \tag{11}$$

$$\Delta(n_2) = n_2 \cdot \sqrt{\left(\frac{\Delta(\sin(\theta_1))}{\sin(\theta_1)}\right)^2 + \left(\frac{\Delta(\sin(\theta21))}{\sin(\theta_2)}\right)^2 \left(\frac{\Delta(n_1)}{na_1)}\right)^2} \tag{12}$$

The refraction index is $n_2 = 1.542 \pm 0.008$. In Table 1 the total error on n_2 as well as the single contributions are given.

	θ_1	θ_2	n_1	total
$\Delta(n_2)$	0.003	0.008	0.00015	0.008

Table 1: Error on n_2 and the single contributions to this error

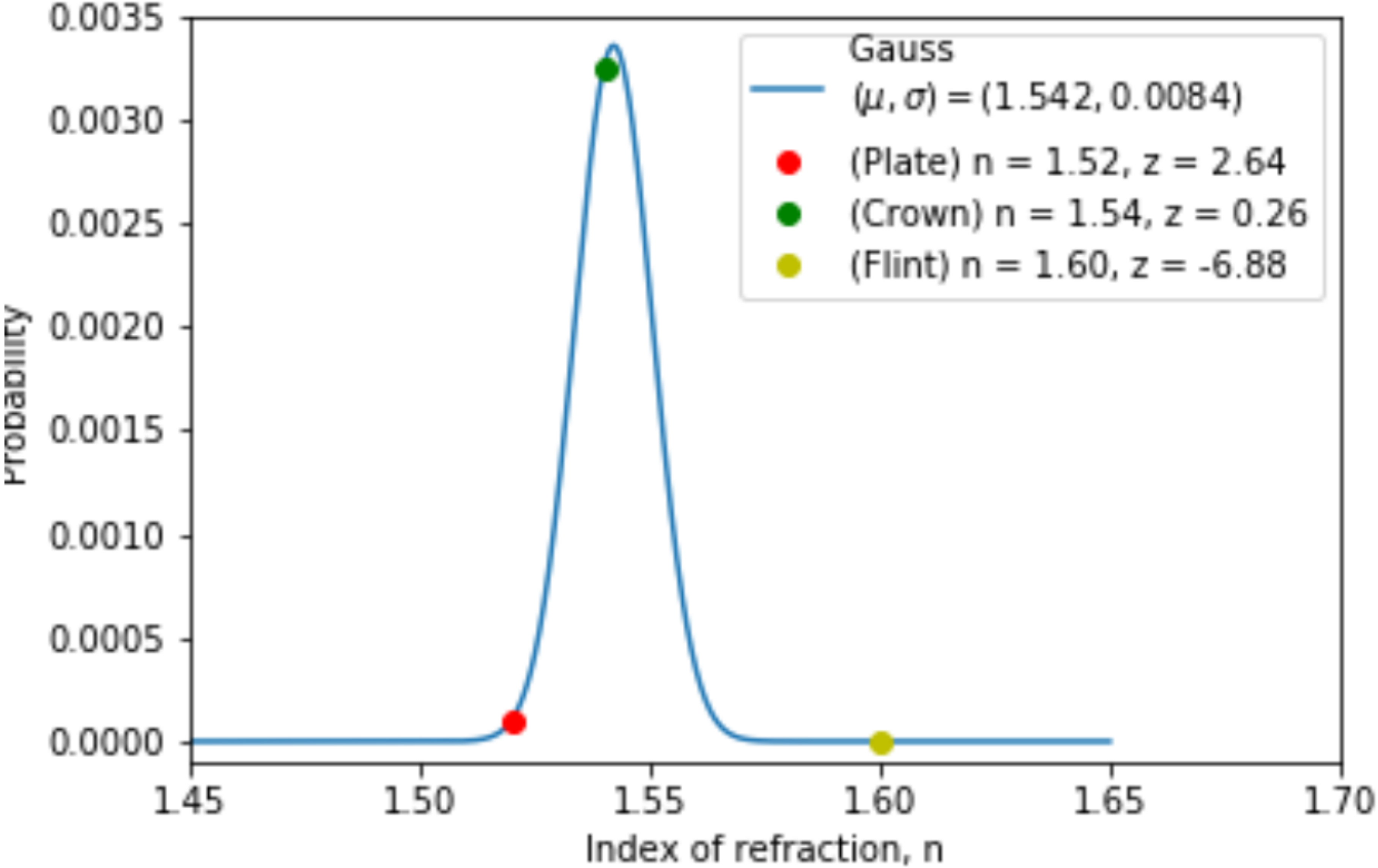
In Table 2 the difference of this value to the refraction index of Plate, Crown, and Flint glass is given in number of sigmas, as well as the p-value (the p-value is always the probability to get a worse result if one assumes a Gaussian distribution of the errors).

	Plate ($n = 1.52$)	Crown ($n = 1.54$)	Flint glass ($n = 1.60$)
σ	-2.6	-0.26	6.9
p-value	0.0047	0.397	$2.6 \cdot 10^{-12}$

Table 2: Comparison of result to refraction index of three materials

The measurement is most likely from Crown, but it could also be from Plate. Here and in the following I choose the limit to be statistically significant to be 3σ.

Problem 2.1



Gauss
$(\mu, \sigma) = (1.542, 0.0084)$
(Plate) n = 1.52, z = 2.64
(Crown) n = 1.54, z = 0.26
(Flint) n = 1.60, z = -6.88
Probability
Index of refraction, n
0.0035
0.0030
0.0025
0.0020
0.0015
0.0010
0.0005
0.0000
1.45
1.50
1.55
1.60
1.65
1.70

Problem 2.2

To find the W boson mass we'll use a weighted mean and the weighted uncertainty

$$\bar{x} = \frac{\sum x_i/\sigma_i^2}{\sum 1/\sigma_i^2} \quad , \quad \sigma(\bar{x}) = \sqrt{\frac{1}{\sum 1/\sigma_i^2}}$$

Using these equations we get the mass of the W bosson:

$$W\,mass = \underline{80.375 \pm 0.011\ GeV}$$

For comparing the two values of the W mass with/without the Higgs bosson included, we'll use a *Two-sample test.*

$$z_{two} = \frac{\bar{x_1} - \bar{x_2}}{\sqrt{\sigma_1^2 + \sigma_2^2}}$$

Giving us

$$(With) \quad z_{two} = \frac{80.375 - 80.358}{0.011^2 + 0.008^2} = \underline{1.3}$$

$$(Without) \quad z_{two} = \frac{80.375 - 80.249}{0.011^2 + 0.008^2} = 9.3$$

Measurement	ALEPH	Delphi	Opal	L3	CDF	D0	ATLAS
Uncertainty (GeV)	0.050	0.066	0.050	0.054	0.015	0.024	0.015

$$\sigma^2 = \sigma_{True}^2 + s^2 \quad \Rightarrow \quad \sigma_{True} = \sqrt{\sigma^2 - s^2}$$

$$W\ mass_{true} = \underline{80.376 \pm 0.009 \pm 0.011\ GeV}$$

Problem 2.2

To find the W boson mass we'll use a weighted mean and the weighted uncertainty

$$\bar{x} = \frac{\sum x_i/\sigma_i^2}{\sum 1/\sigma_i^2} \quad , \quad \sigma(\bar{x}) = \sqrt{\frac{1}{\sum 1/\sigma_i^2}}$$

Using these equations we get the mass of the W bosson:

$$W\,mass = \underline{80.375 \pm 0.011\ GeV}$$

For comparing the two values of the W mass with/without the Higg *test.*

Should you have tested, if the measurements were consistent? How about a ChiSquare test?

$$z_{two} = \frac{\bar{x_1} - \bar{x_2}}{\sqrt{\sigma_1^2 + \sigma_2^2}}$$

Giving us

$$(With) \quad z_{two} = \frac{80.375 - 80.358}{0.011^2 + 0.008^2} = \underline{1.3}$$

$$(Without) \quad z_{two} = \frac{80.375 - 80.249}{0.011^2 + 0.008^2} = 9.3$$

Measurement	ALEPH	Delphi	Opal	L3	CDF	D0	ATLAS
Uncertainty (GeV)	0.050	0.066	0.050	0.054	0.015	0.024	0.015

$$\sigma^2 = \sigma_{True}^2 + s^2 \quad \Rightarrow \quad \sigma_{True} = \sqrt{\sigma^2 - s^2}$$

$$W\ mass_{true} = \underline{80.376 \pm 0.009 \pm 0.011\ GeV}$$

Measurement situation

There are four possible situations in experimental measurements of a quantity:

One measurement, no error:

X = 3.14

Situation: You are f***ed!

You have no clue about uncertainty, and you can not obtain it!

Several measurements, no errors:

X1 = 3.14
X2 = 3.21
X3 = …

Situation: You are OK

You can combine the measurements, and from RMS get error on mean.

One measurement, with error:

X = 3.14 ± 0.13

Situation: You are OK

You have a number with error, which you can continue with.

Several measurements, with errors:

X1 = 3.14 ± 0.13
X2 = 3.21 ± 0.09
X3 = …

Situation: You are on top of things!

You can both combine to a weighted, average and check with a chi-square.

Problem 2 problems

<u>General:</u>
- different number of significant digits between value and error. Too many/few digits.

P2.1:
- deriving correctly error propagation of Snell's law,
- rejecting Plate glass, even though the one-sample test gives <3 sigma,
- testing glass type inside arbitrary confidence limit (1-2 sigmas), without proper test.

P2.2:
- no chi2 and probability when using the weighted mean,
- using One sample instead of two sample test, when reference has an uncertainty,
- adding the systematic error to the uncertainty calculated in the previous question,
- adding systematic error before calculating the weighted mean (syst. / sqrt(N)!!!),
- adding or subtracting uncertainties NOT in quadrature,

Several students used a chi2 test for p2.2.2, for which I considered correct (when the execution of the chi2 was correct though), even if the two sample test has not been used.

Problem 3.1

This problem didn't give too many troubles.

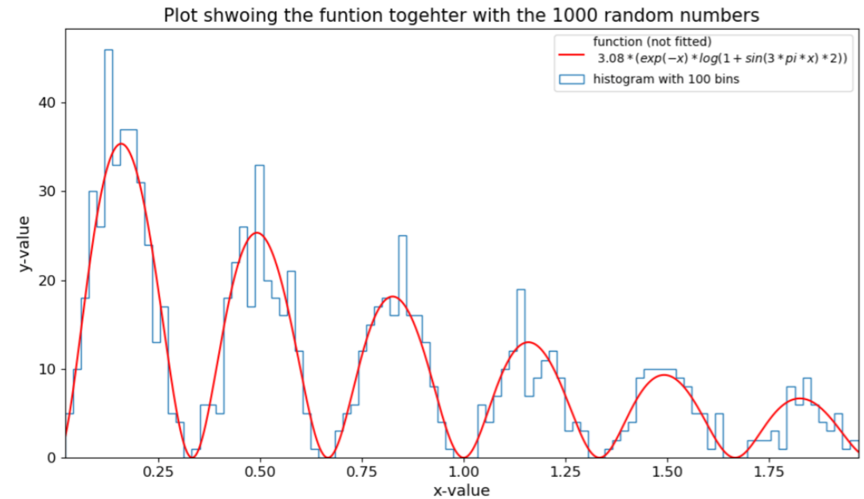
One has to use the Accept-Reject method for f(x), and a combination for g(x), though several used "extended Accept-Reject" approx.

The fit probably required a **note about the low statistics** in several of the bins.

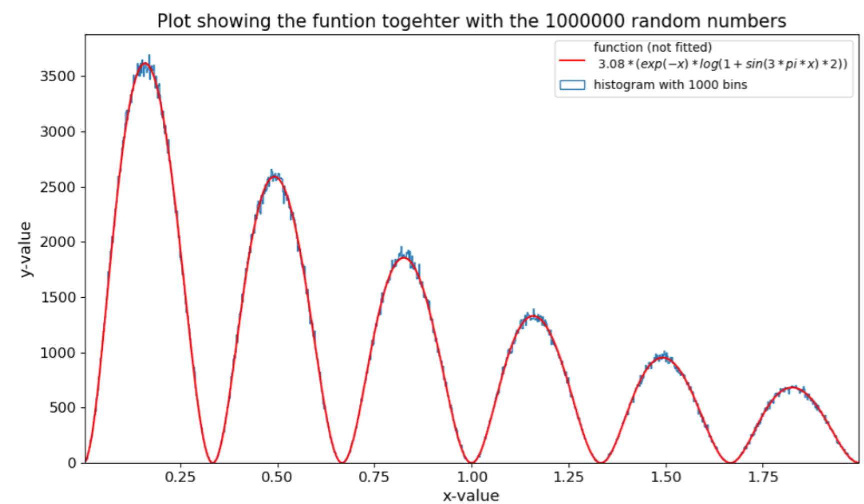
The first comparison could be done with fit (or Shapiro-Wilks test).

The second comparison should be done by Kolmogorov-Smirnov (or Anderson-Darling).

BNV226



Due to the binning the generated numbers doesn't seems to be located spot on, so to clarify that we made another plot with 1000 bins instead of 100 and with 1,000,000 generated numbers instead of just 1,000. This result is:



Now it is much more clear that the generated numbers are match the function.

Problem 3.1

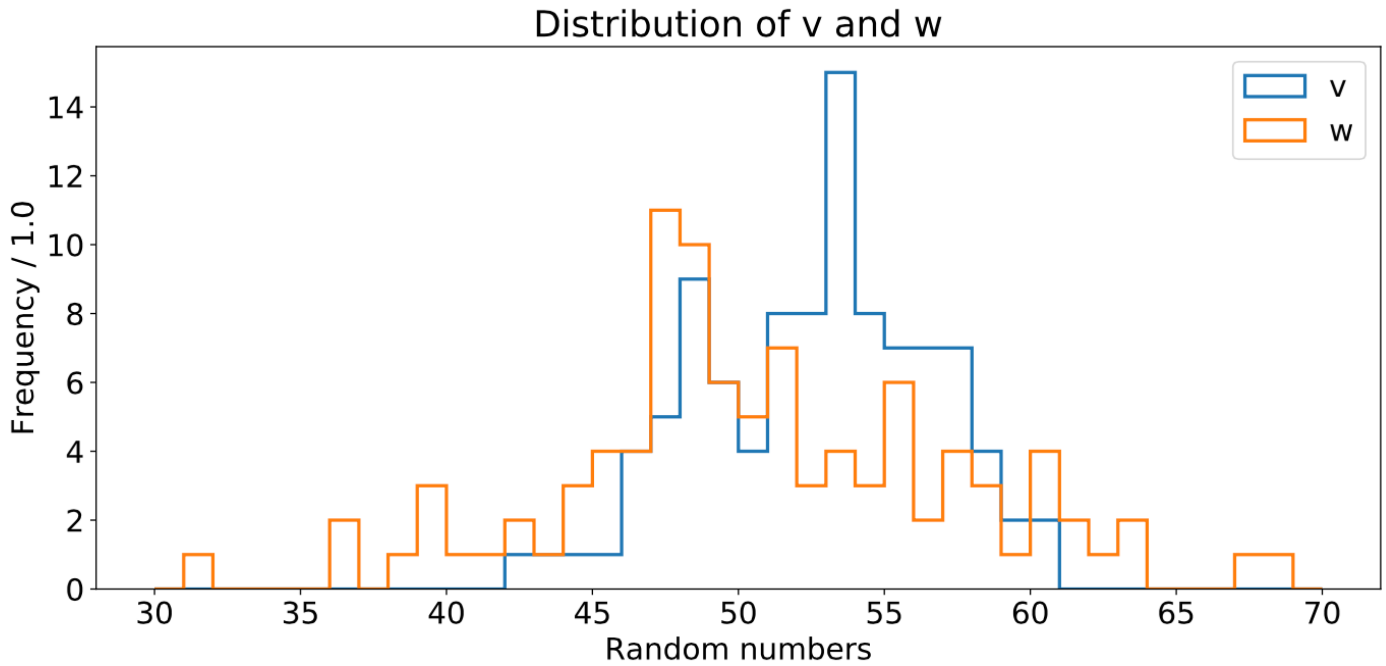


Figure 4: 100 numbers of v and w generated as a sum of 75 and 50 random numbers from $f(x)$ and $g(x)$ respectively.

Test		Probability
Difference in mean	Difference: 1.7 ± 0.8, 2.1 sigma away	0.01949
Binned χ^2		0.13774
Binned Kolomogrov-Smirnov		0.00646
Unbinned Kolomorov-Smirnov		0.00303

Problem 3.1

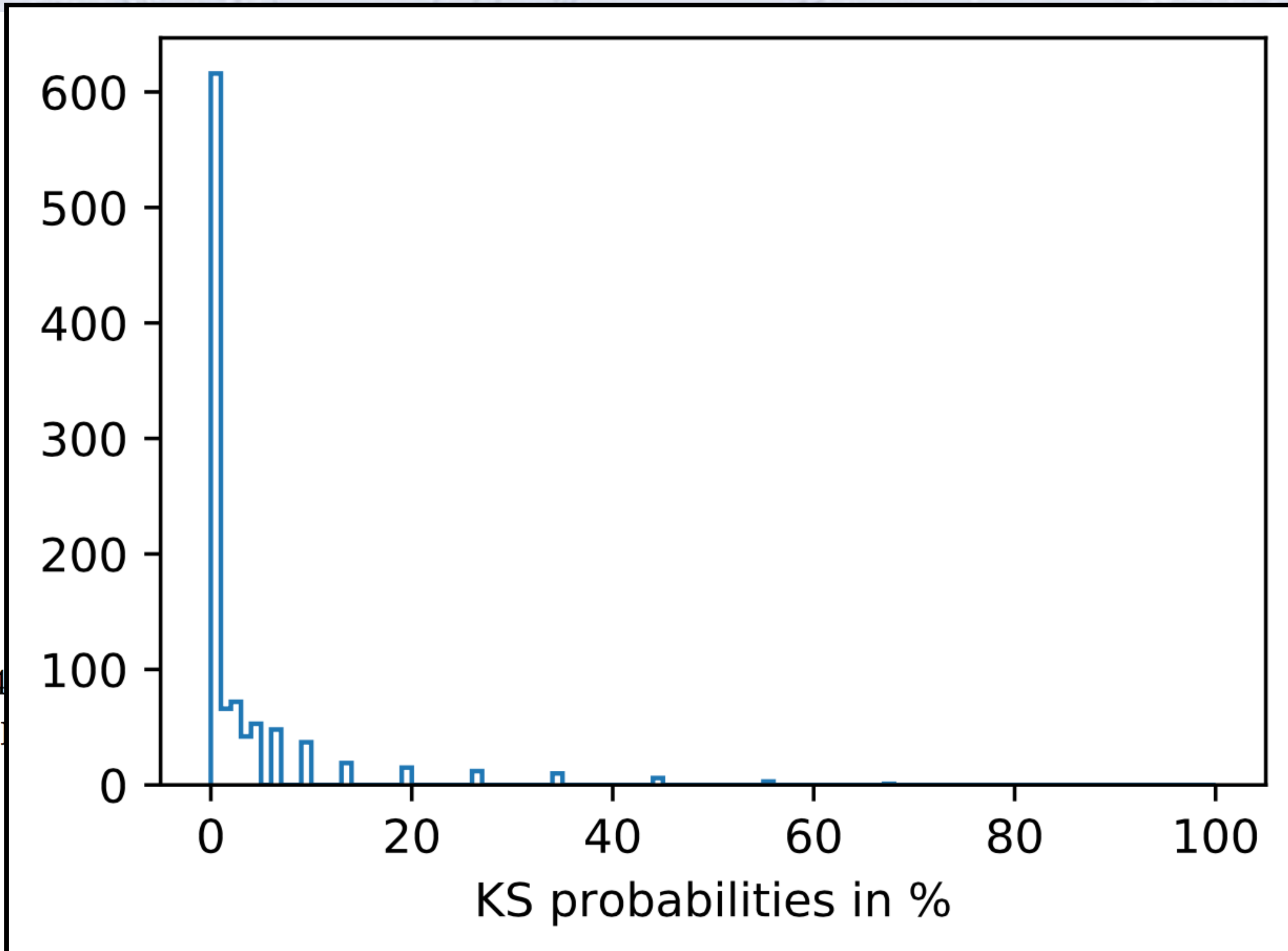


Figure 4 … $f(x)$ and $g(x)$ res…

Binned Kolomogrov-Smirnov	0.00646
Unbinned Kolomorov-Smirnov	0.00303

Problem 3 problems

General:

- Confusion about how to interpret a TS. (ie rejecting the TS when you should accept it)

3.1 Most people only wrote 1/2 reasons why using the hit-and-miss method.

3.2

- Forgetting to normalize the function prior to generating the random number (4)
- Buggy / failed number production or wrong frequency of the function

3.3

- Fitting -p0 instead of p0 or no labels/traces of the fit procedure on the plots.
- Failling to quantify the deviation between fitted and expected parameters.
- Annoying: three word sentences like "fit looks nice".

3.4

- Missing the link between the exercice and the Central Limit Theorem (4)
- Annoying: plots without error bars, fit curve without fit result on the plot

3.5 When choosing to extend f(x) to a large bound, failing to estimate the precision level

3.6

- For many people, the histograms of u and v were largely separated (i.e. wrong)
- For some of the above, both f(x) and g(x) were sampled over the wrong x interval
- Lack of information about whether students have renormalized g(x)
- Stopping the comparison after looking at distribution width (no KS test)

Problem 4.1

- I use the `mean()` and `std()` functions from the numpy library in Python to find the mean and standard deviation of the heights of the persons measured (removing 1 degree of freedom when taking the standard deviation, as it has been used to calculate the mean). The uncertainty on the mean is then the standard deviation divided by the square root of the number of persons measured. This yields

$$\underline{\underline{\bar{h} = (1.7317 \pm 0.0016)\text{m}.}}$$

- I calculate the mean for men and women separately, in the same fashion as described above. This yields

$$\bar{h}_m = (1.800 \pm 0.002)\text{m}$$
$$\bar{h}_w = (1.6828 \pm 0.0016)\text{m}$$

where $\bar{h}_m$ is the mean of heights for men and $\bar{h}_w$ is the mean of heights for women. Combining this in a weighted mean yields

$$\underline{\underline{\bar{h} = (1.7237 \pm 0.0013)\text{m}.}}$$

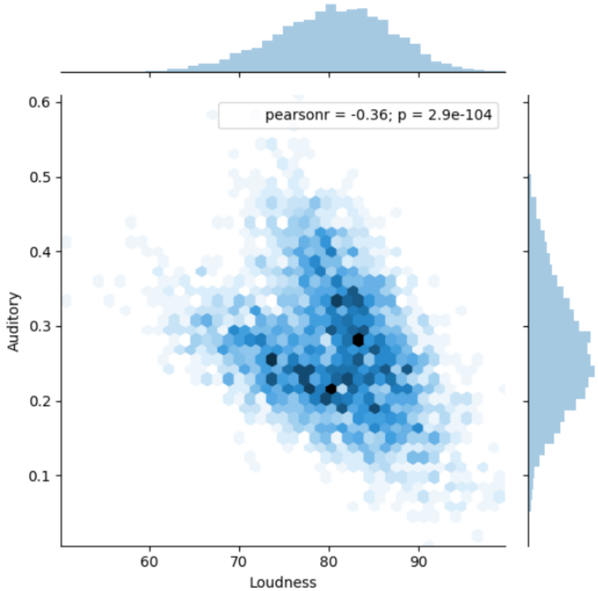
This value lies a distance of 3.8σ from the value found above, and thus they can not be said to be in agreement.

This is an example of STRATIFICATION, where the uncertainty is minimised by first dividing into similar samples, and afterwards combining them.

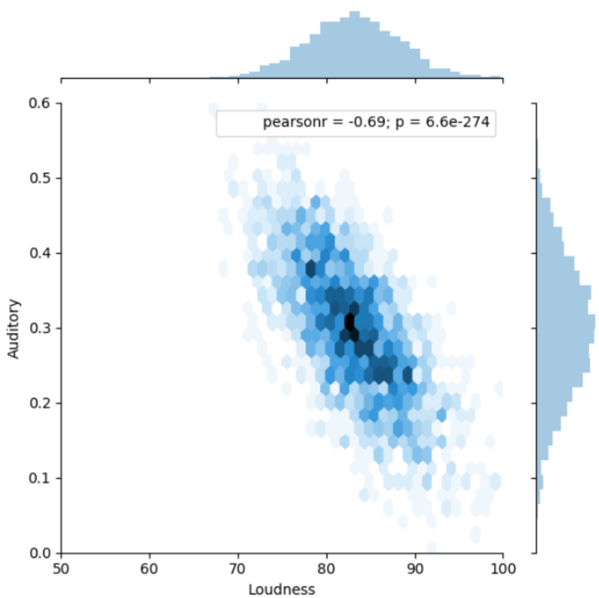
Problem 4.1

Very nice figure, which shows both the distributions and their correlation.

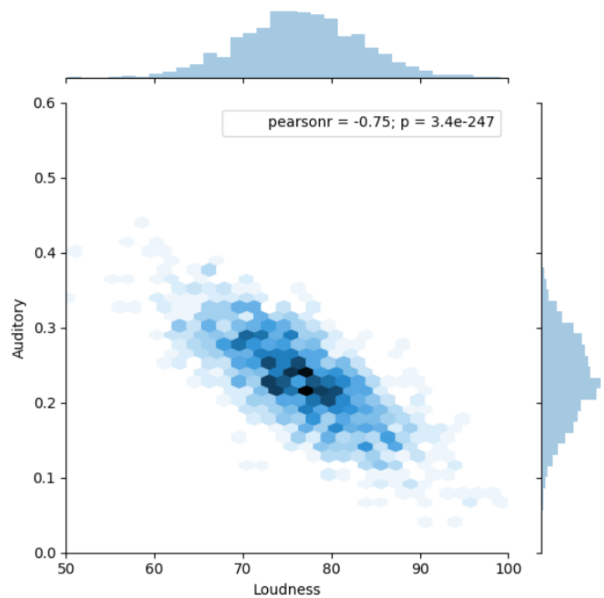
Some persons even gave p-values for the two variables not being correlated.



(a) all



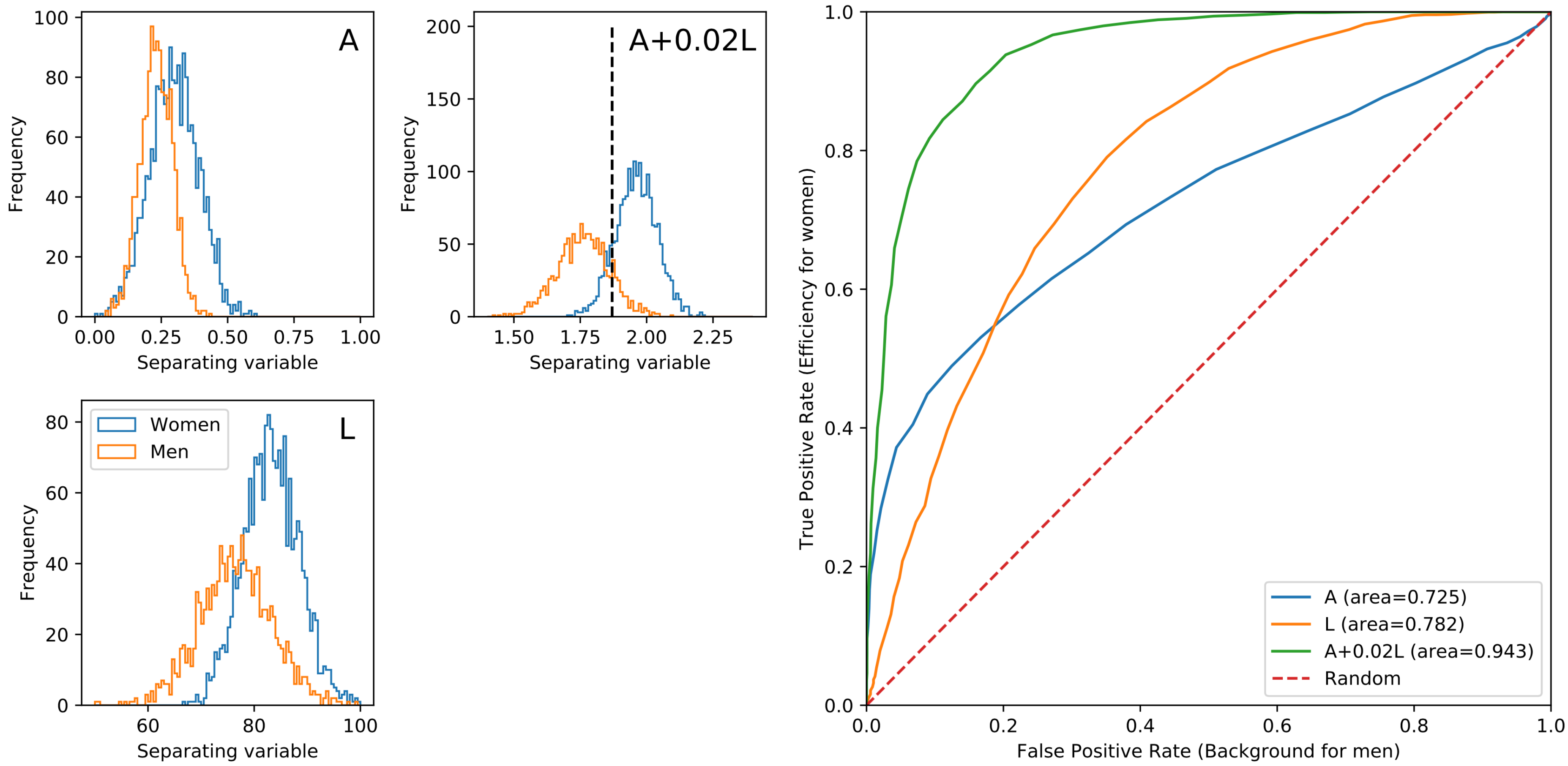
(b) women



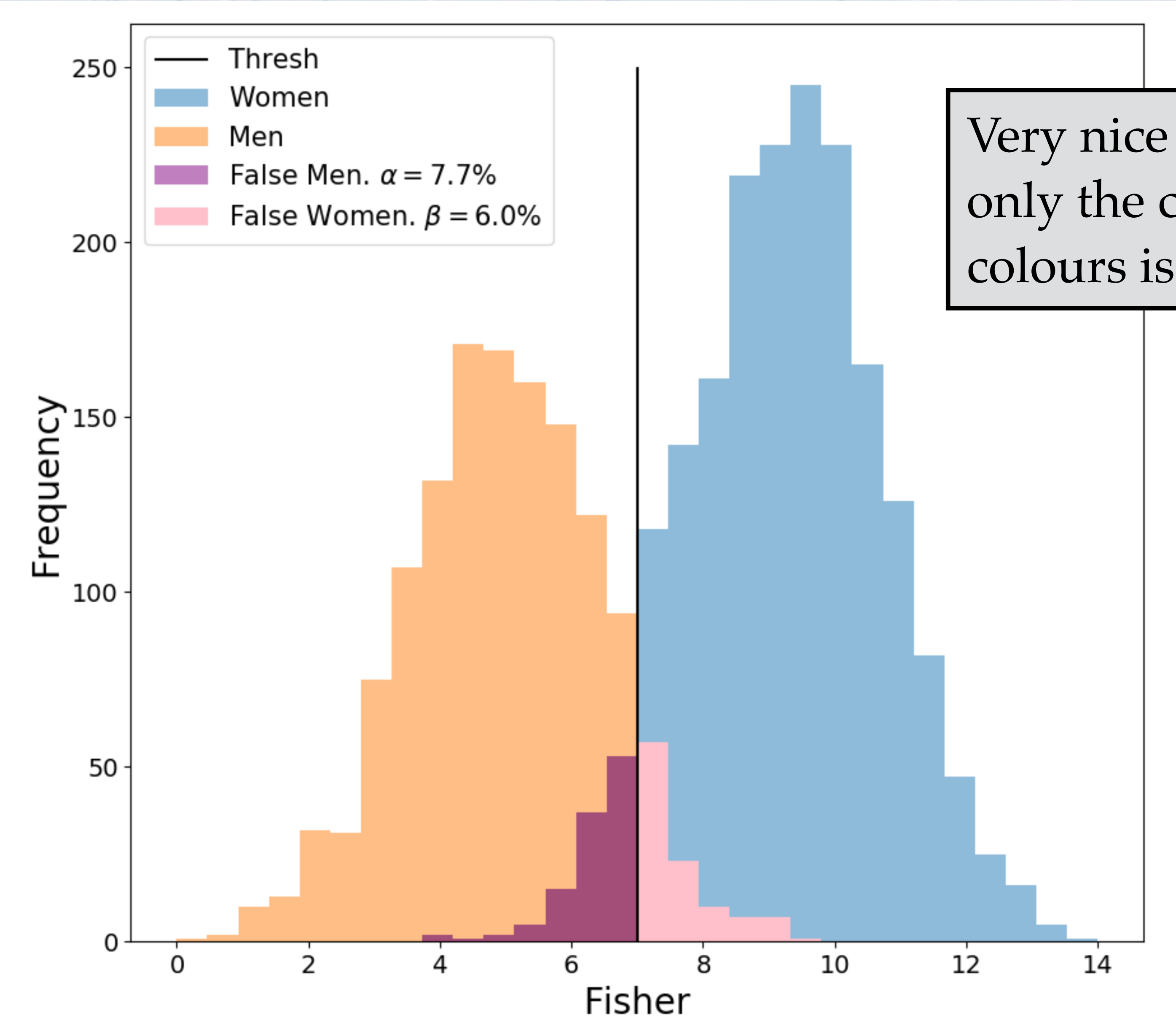
(c) men

Problem 4.1

The problem in 4.1 didn't require anything "fancy", only a good treatment of what was done. Here is an "ad hoc" Fisher = A + 0.02*L



Problem 4.1



Very nice figure, only the choice of colours is a bit odd!

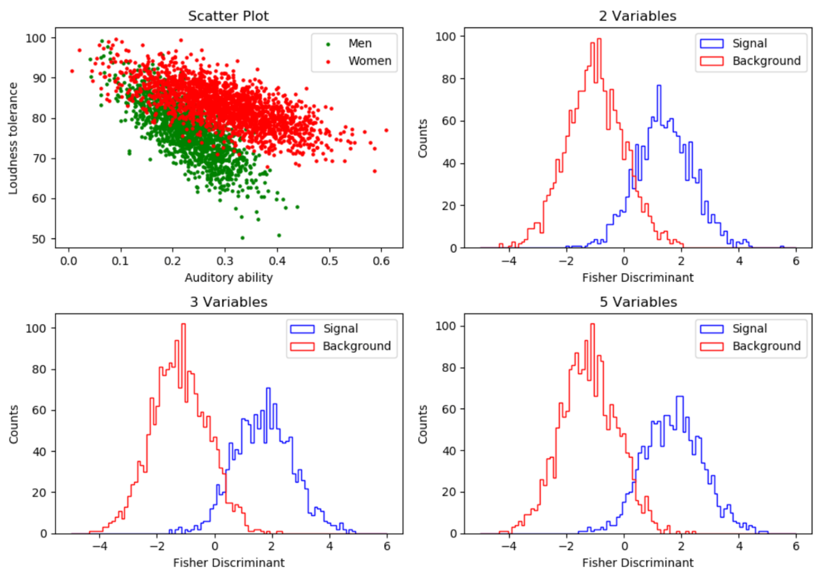
Problem 4.1

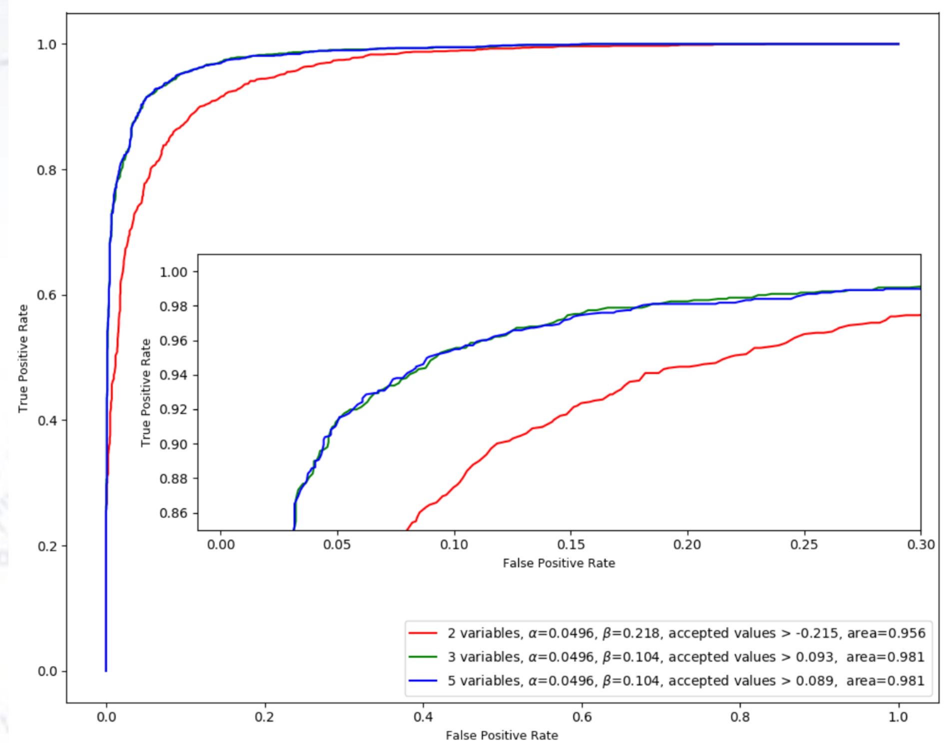
Here is a comparison of Fishers with a range of variables included. Clearly, three variables are enough and hardly anything is gained by including the last two variables… shown very nicely.

$$F_2 = -14.92 \cdot \text{aud} - 0.215 \cdot \text{loud} + 21.30 \quad (17)$$

$$F_3 = 7.64 \cdot \text{height} - 12.41 \cdot \text{aud} - 0.181 \cdot \text{loud} + 4.67 \quad (18)$$

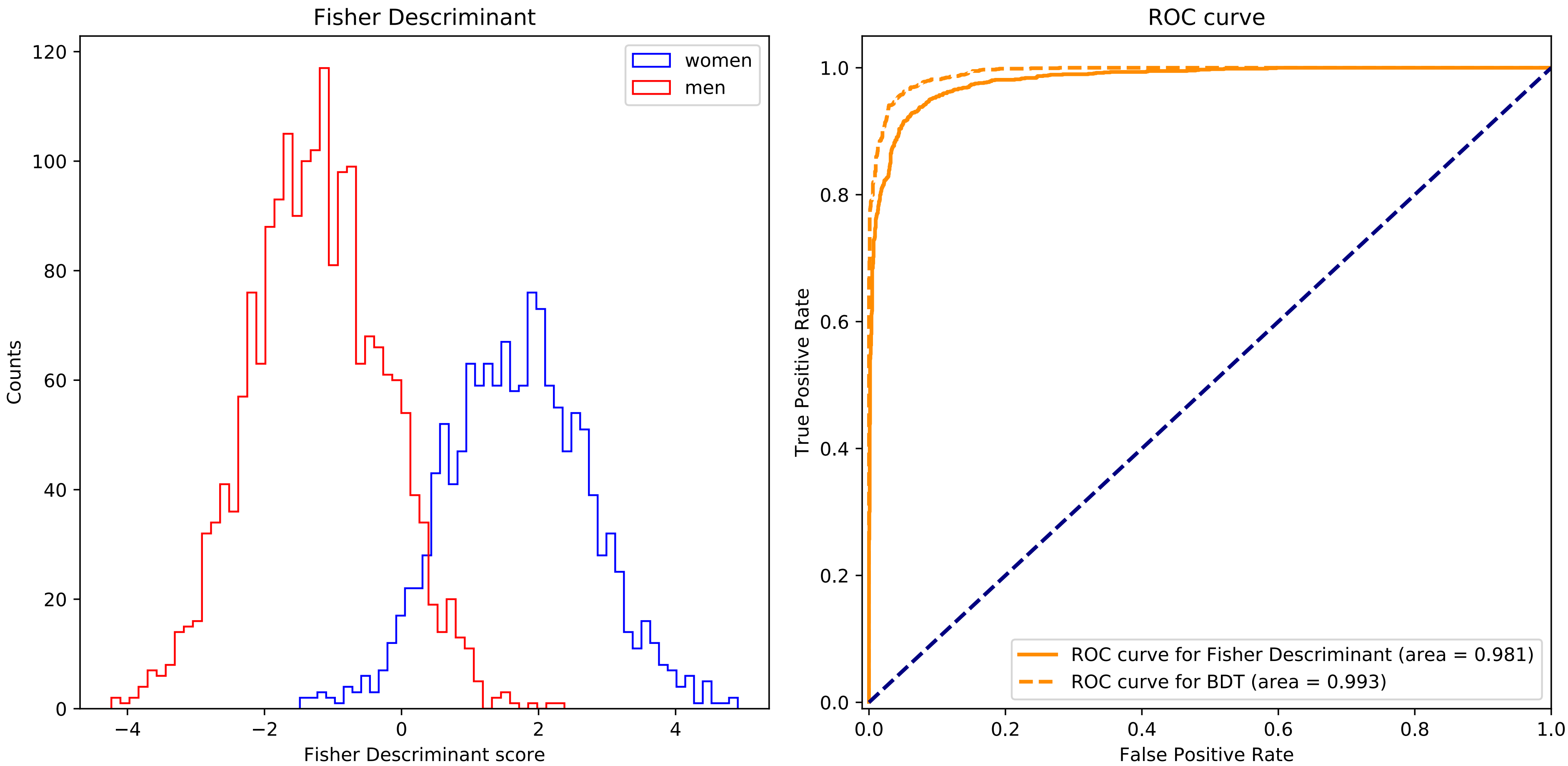
$$F_5 = 7.64 \cdot \text{height} - 12.37 \cdot \text{aud} - 0.181 \cdot \text{loud} - 0.00358 \cdot \text{eye} + 0.000489 \cdot \text{IQ} + 4.93 \quad (19)$$





Problem 4.1

The machine learning solution improves on top of the Fisher approach, especially since the IQ was symmetric and hence not usable by the Fisher!



Problem 4.2

So the averages are $\hat{\mu}_{away} = 1.2 \pm 0.5$ and $\hat{\mu}_{home} = 1.5 \pm 0.5$. So from this there is no significant average being on a home field (though i bet you both fans and players would tell you differently!)

Problem 4.2

So the averages are $\hat{\mu}_{away} = 1.2 \pm 0.5$ and $\hat{\mu}_{home} = 1.5 \pm 0.5$. So from this there is no significant average being on a home field (though i bet you both fans and players would tell you differently!)

"I bet you, both fans and players disagree with your uncertainties!!!"
[My thoughts (cocky moment late night)]

Problem 4.2

So the averages are $\hat{\mu}_{away} = 1.2 \pm 0.5$ and $\hat{\mu}_{home} = 1.5 \pm 0.5$. So from this there is no significant average being on a home field (though i bet you both fans and players would tell you differently!)

"I bet you, both fans and players disagree with your uncertainties!!!"
[My thoughts (cocky moment late night)]

- Given the amount of goals scored in a game, $n_{goals} = 0, 1, ..., 9$, and the amount of matches in which that many goals have been scored, $n_{matches}$ I calculate the average number of goals scored as

$$\mu = \frac{1}{N} \sum_{i=1}^{10} (n_{matches,i} \cdot n_{goals,i}),$$

where N is the total amount of matches played, $N = 3420$. The variance is found correspondingly as

$$V(n_{goals}) = \frac{1}{N} \sum_{i=1}^{10} (n_{matches,i} \cdot (n_{goals,i} - \mu)^2).$$

The uncertainty on the mean can then be found as $\sigma_\mu = \sqrt{V}/\sqrt{N}$. This yields the average number of goals scored for home and away:

$$\mu_{home} = 1.56 \pm 0.02$$
$$\mu_{away} = 1.16 \pm 0.02$$

Problem 4.2

So the averages are $\hat{\mu}_{away} = 1.2 \pm 0.5$ and $\hat{\mu}_{home} = 1.5 \pm 0.5$. So from this there is no significant average being on a home field (though i bet you both fans and players would tell you differently!)

"I bet you, both fans and players disagree with your uncertainties!!!"
[My thoughts (cocky moment late night)]

- Given the amount of goals scored in a game, $n_{goals} = 0, 1, ..., 9$, and the amount of matches in which that many goals have been scored, $n_{matches}$ I calculate the average number of goals scored as

$$\mu = \frac{1}{N} \sum_{i=1}^{10} (n_{matches,i} \cdot n_{goals,i}),$$

where N is the total amount of matches played, $N = 3420$. The variance is found correspondingly as

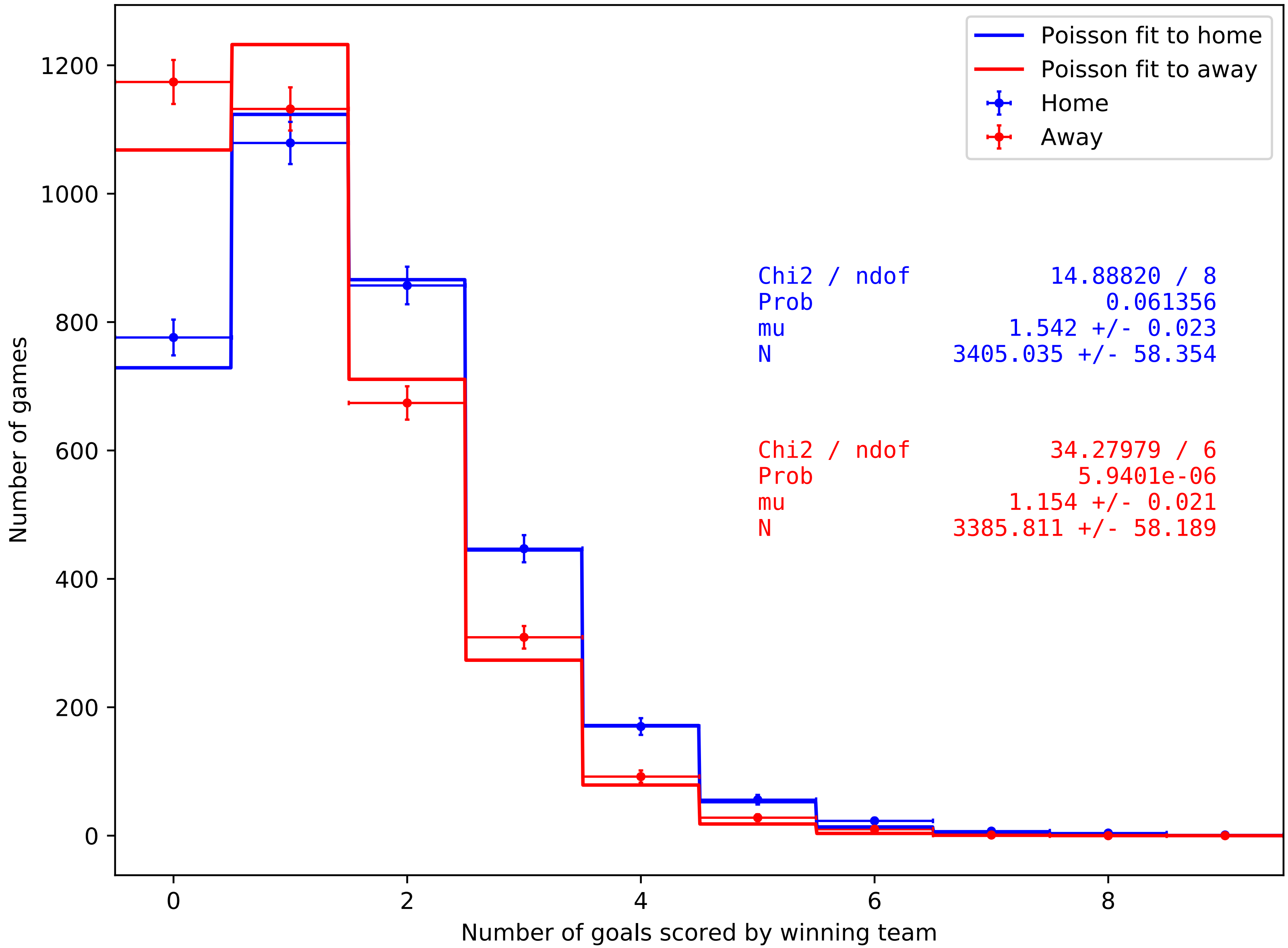
$$V(n_{goals}) = \frac{1}{N} \sum_{i=1}^{10} (n_{matches,i} \cdot (n_{goals,i} - \mu)^2).$$

Note that since these distributions are close to being Poisson, one can get the uncertainty from the RMS = sqrt(mean)… at least as a check.

Th … age
nu

$$\mu_{home} = 1.56 \pm 0.02$$
$$\mu_{away} = 1.16 \pm 0.02$$

Problem 4.2



Poisson fit to home
Poisson fit to away
Home
Away
Chi2 / ndof 14.88820 / 8
Prob 0.061356
mu 1.542 +/- 0.023
N 3405.035 +/- 58.354
Chi2 / ndof 34.27979 / 6
Prob 5.9401e-06
mu 1.154 +/- 0.021
N 3385.811 +/- 58.189
Number of games
Number of goals scored by winning team
0
200
400
600
800
1000
1200
0
2
4
6
8

Problem 4.2

The 2x2 contingency table allowed for the use of Fishers exact test, but the Chi2 gives more or less the same result. If one includes more detailed information, then the chance of these being truly independent drops.

		Away	
		0	1+
Home	0	278	498
	1+	896	1748

χ^2 contingency test for 3x3, prob = **2.690%**
χ^2 contingency test for 4x4, prob = **0.007%**
χ^2 contingency test for 5x5, prob = **0.033%**
χ^2 contingency test for 6x6, prob = **0.706%**
χ^2 contingency test for 7x7, prob = **4.431%**
χ^2 contingency test for 8x8, prob = **31.823%**

Fisher's exact test, prob = **32.3%.**
χ^2 contingency test, prob = **33.9%.**

This last problem caused some troubles, and many got a very high correlation, despite hardly any change in the A or H distribution when changing H or A.

Problem 4.2

The 2x2 contingency table allowed for the use of Fishers exact test, but the Chi2 gives more or less the same result. If one includes more detailed information, then the chance of these being truly independent drops.

		Away	
		0	1+
Home	0	278	498
	1+	896	1748

χ^2 contingency test for 3x3, prob = **2.690%**
χ^2 contingency test for 4x4, prob = **0.007%**
χ^2 contingency test for 5x5, prob = **0.033%**
χ^2 contingency test for 6x6, prob = **0.706%**
χ^2 contingency test for 7x7, prob = **4.431%**
χ^2 contingency test for 8x8, prob = **31.823%**

Fisher's exact test, prob = **32.3%.**
χ^2 contingency test, prob = **33.9%.**

This last problem caused some troubles, and many got a very high correlation, despite hardly any change in the A or H distribution when changing H or A.

Giving us a correlation of 2.13, which seems quite high.

From a solution

Problem 4 problems

General:

Flushing full Notebook output with a bunch of not-explained/not formatted numbers in cell output with answers. IMHO a poor solution.

4.1

RMS in 4.1.1, 4.1.2

"premature" rounding, leading to wrong z-score.

When using Fisher in 4.1.4, usually only 2 variables are used (why not all available?)

4.1.4 Stopping by producing (Fisher) distributions in better case including z-test separation. Not mentioning value to cut/distinguish two sample, missing alpha, beta errors.

4.2

No uncertainty for average away/home goals

Uncertainty on average as sqrt of number (works for Poisson number not average of them)

Judging "Poissonity" by only checking if mean == variance

Judging "Poissonity" by fitting Poison and only comparing lambdas

Contingency table is not covariance table: Argument about 0 on off-diagonal terms for uncorrelated observables.

Problem 4 problems

General:
Flushing full Notebook output with a bunch of not-explained/not formatted numbers in cell output with answers. IMHO a poor solution.

4.1
RMS in 4.1.1, 4.1.2
"premature" rounding, leading to wrong z-score.
When using Fisher in 4.1.4, usually only 2 variables are used (why not all available?)
4.1.4 Stopping by producing (Fisher) distributions in better case including z-test separation.
Not mentioning value to cut/distinguish two sample, missing alpha, beta errors.

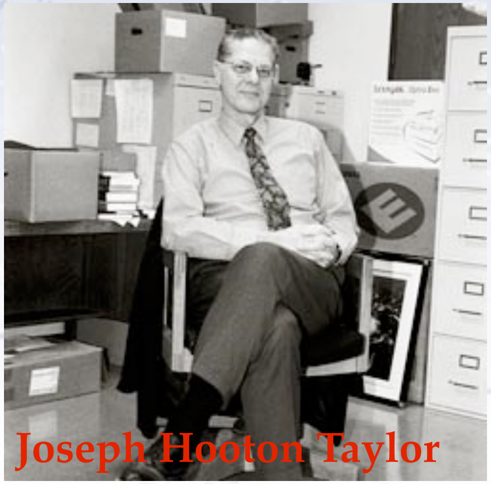
4.2
No uncertainty for average away/home goals
Uncertainty on average as sqrt of number (works for Poisson number not average of them)
Judging "Poissonity" by only checking if mean == variance
Judging "Poissonity" by fitting Poison and only comparing lambdas
Contingency table is not covariance table: Argument about 0 on off-diagonal terms for uncorrelated observables.

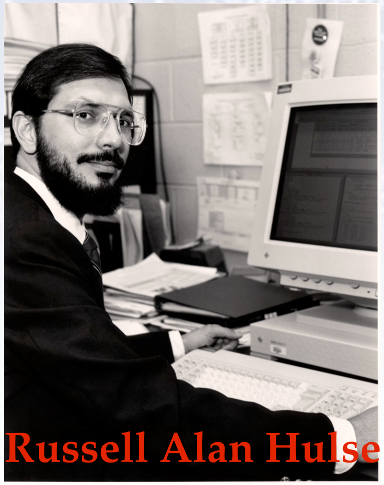
I also learned something new:
p-value of Fisher's exact test is **only for this exact configuration** (~2%), should be a sum of this value and worse (~32%). Not reducing points for that!

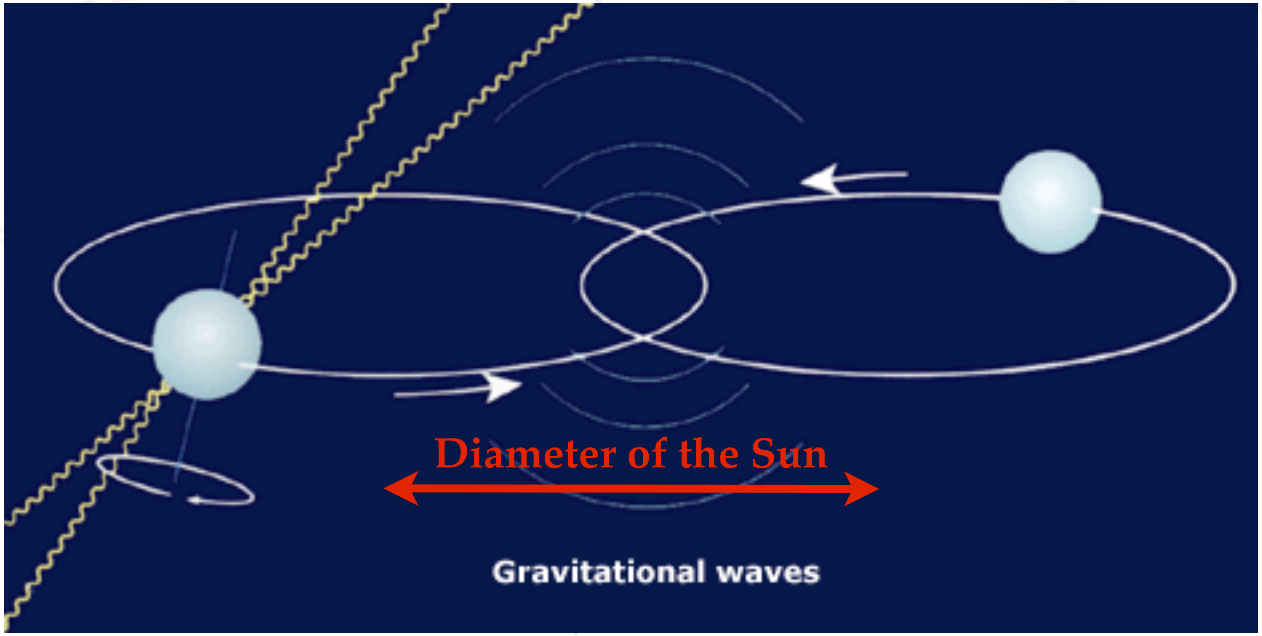

A small intermezzo…
…the Hulse-Taylor-pulsar

Hulse-Taylor-pulsaren

In 1973 Hulse & Taylor discovered a very special pulsar… the period for its signal was NOT constant, but had a variation with a period of 8 hours! As it turns out, this was to become a "jewel" in the test of Einstein's theory of relativity.





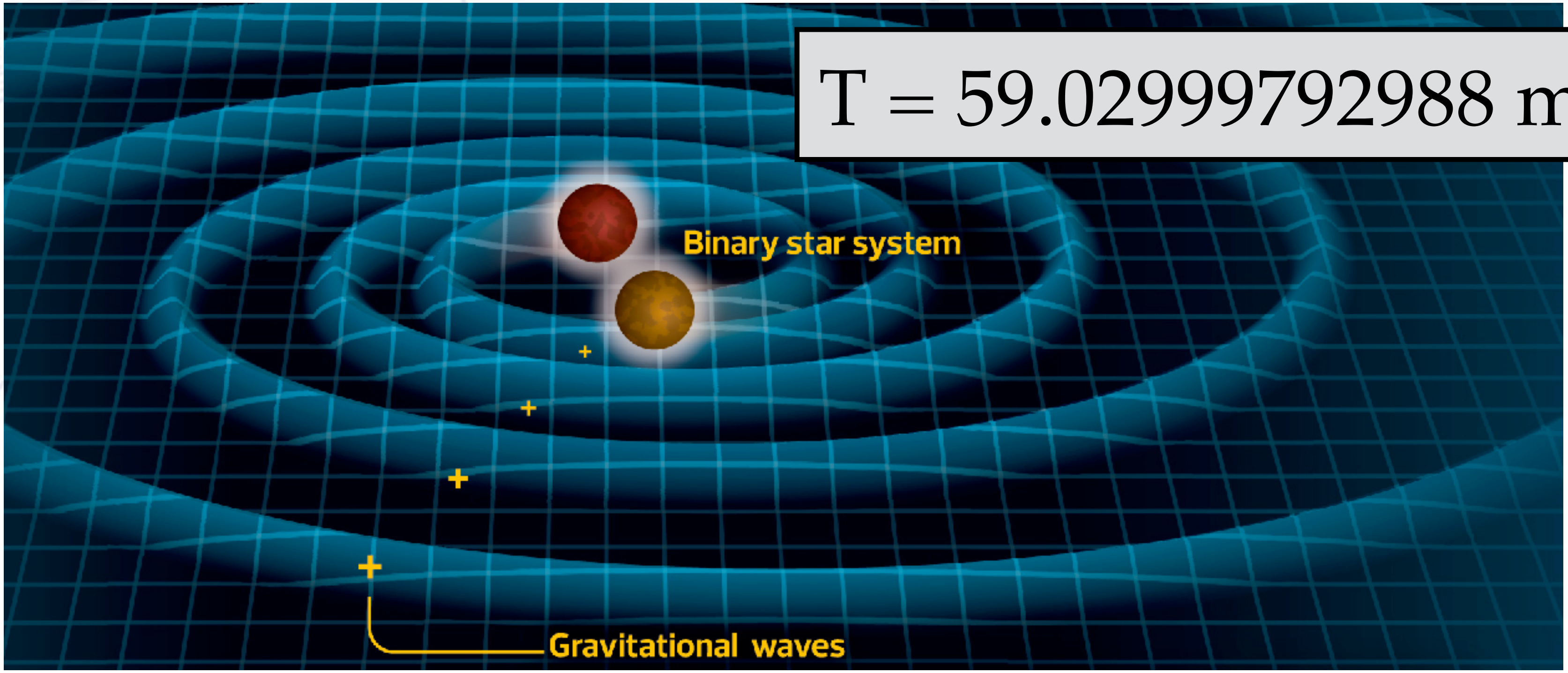


Præcisionsmålinger

In the following years, they measured the pulsar parameters with great precision:

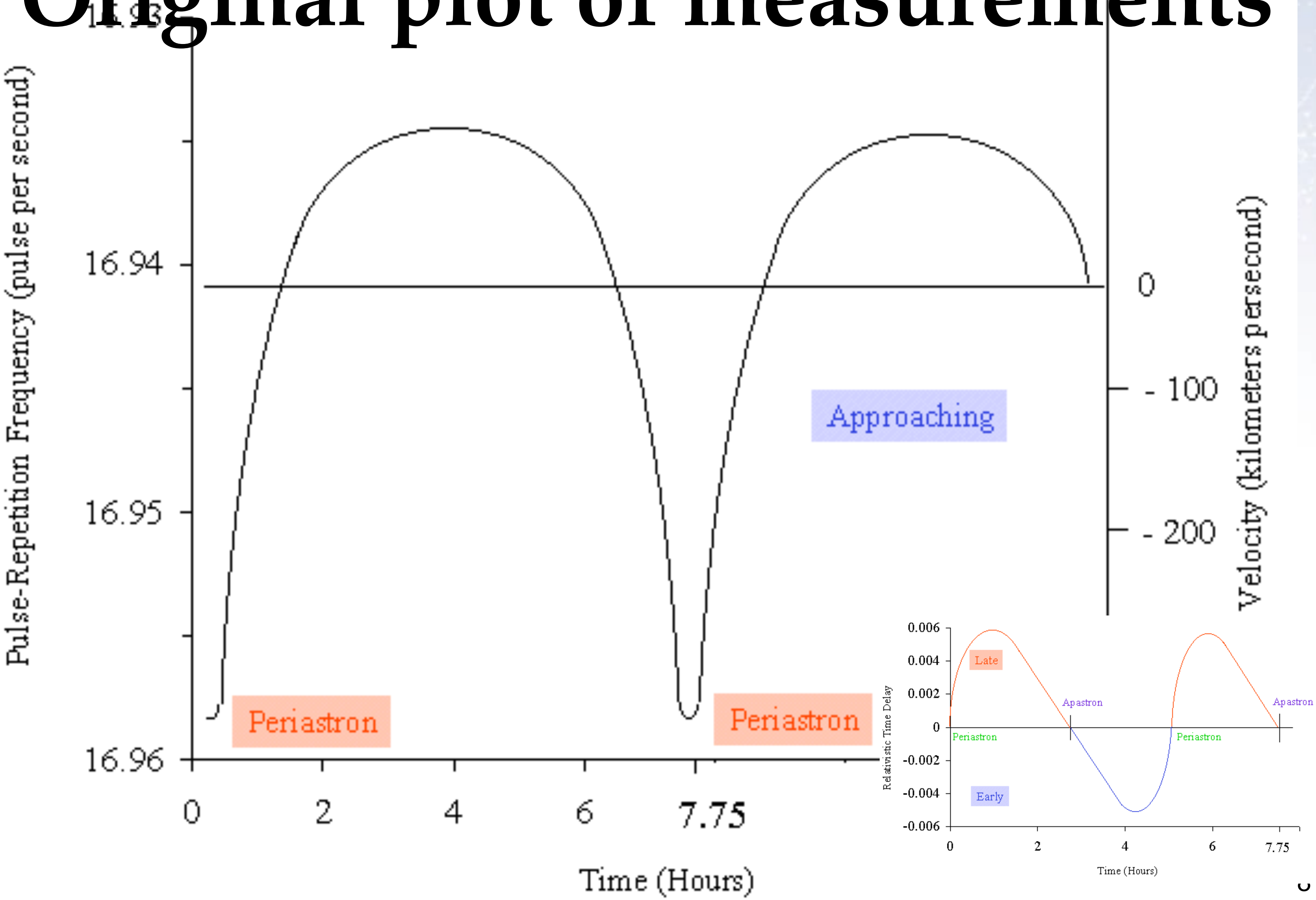
- Mass of companion: 1.387 MSun
- Total mass of the system: 2.828378(7) MSun
- Orbital period: 7.751938773864 hr
- Eccentricity: 0.6171334
- Semimajor axis: 1,950,100 km
- Periastron separation: 746,600 km
- Apastron separation: 3,153,600 km
- Orbital velocity of stars at periastron (relative to center of mass): 450 km/s
- Orbital velocity of stars at apastron (relative to center of mass): 110 km/s

The measurements were possible, partly because of the large relativistic effects. What takes **a century** for Mercury, takes **a day** for the Hulse-Taylor-pulsar!



$T = 59.02999792988$ ms

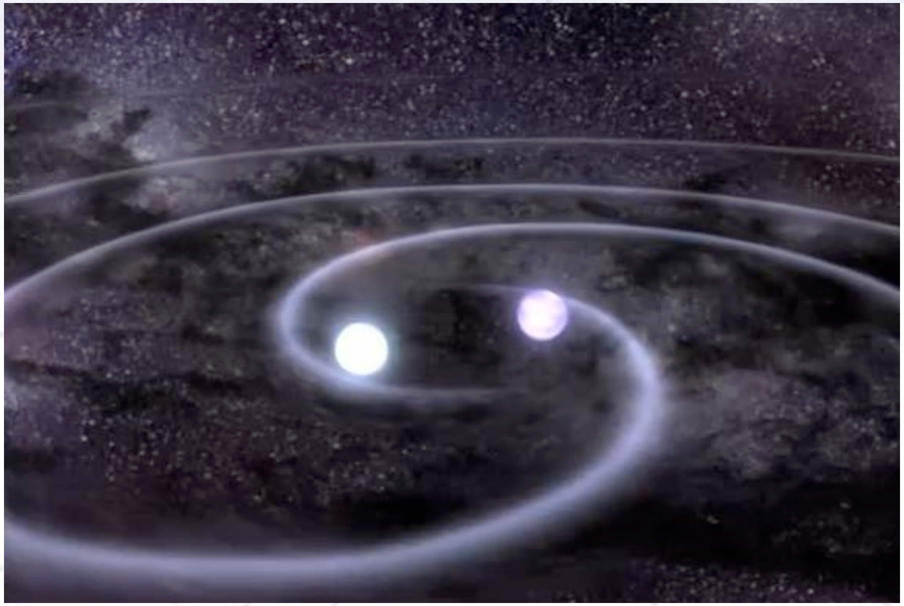
Original plot of measurements

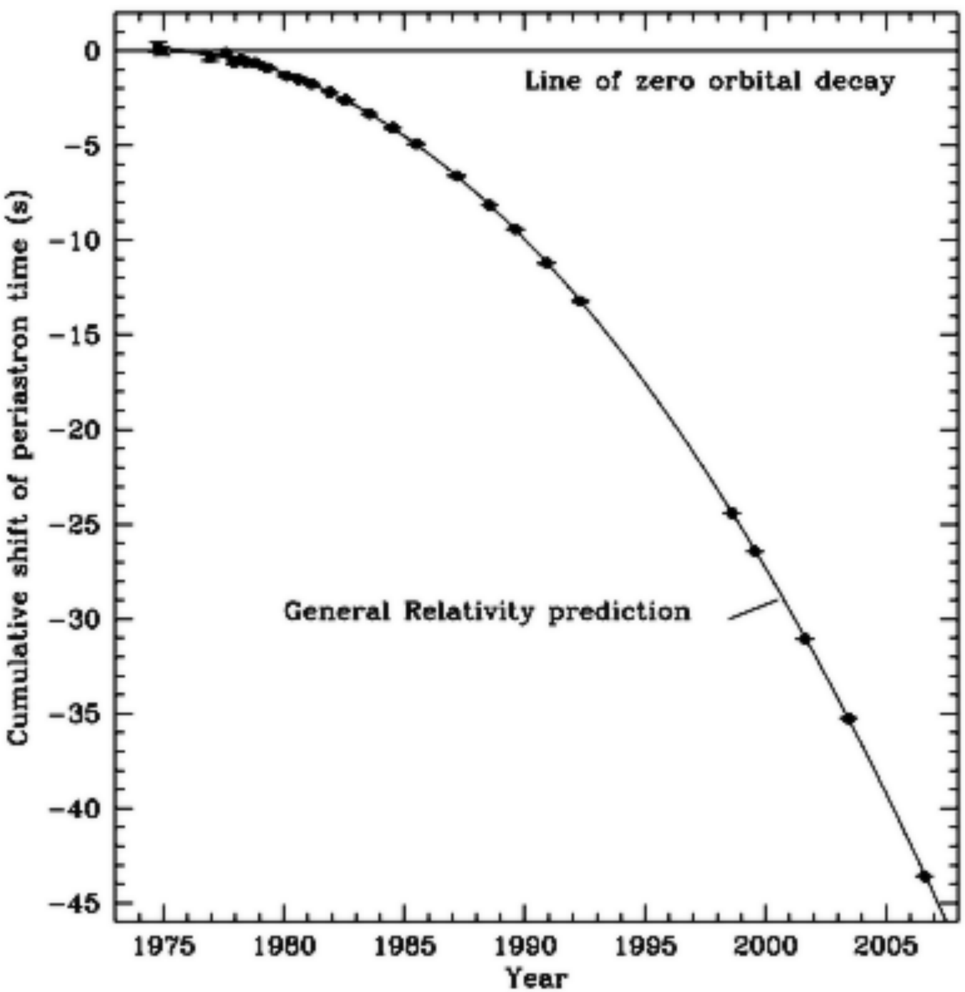


The discovery

After years of observation it became clear, that the pulsars spiral towards each other.

Conclusion: They loose energy (fast). Immediate question: How?!?





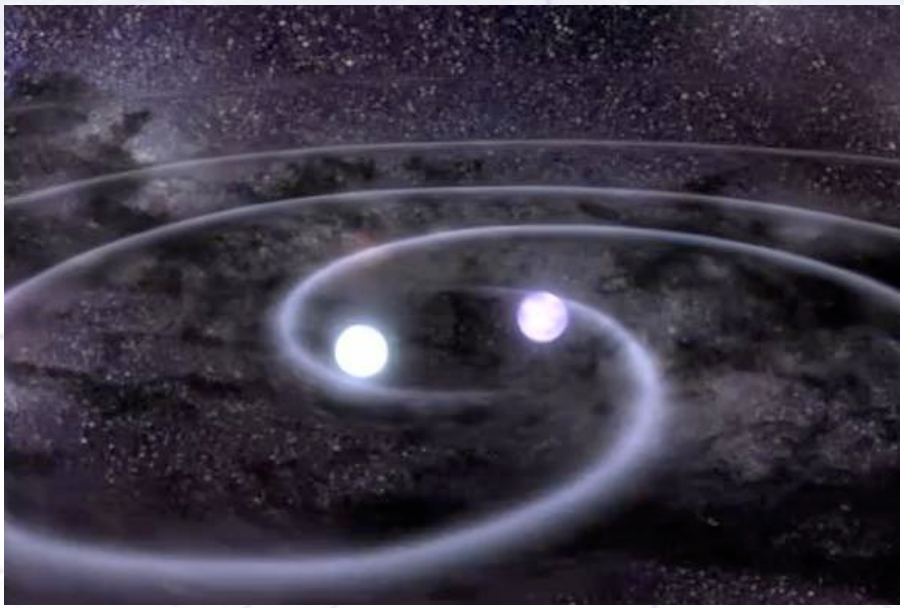
The answer is gravitational waves. What is normally a tiny effect is here so large (100x suns output!), that we (with a natural high precision clock) can see it.

(Sun+Earth: 200 W, 10^{-15} m radius “lost” pr. day)

The discovery

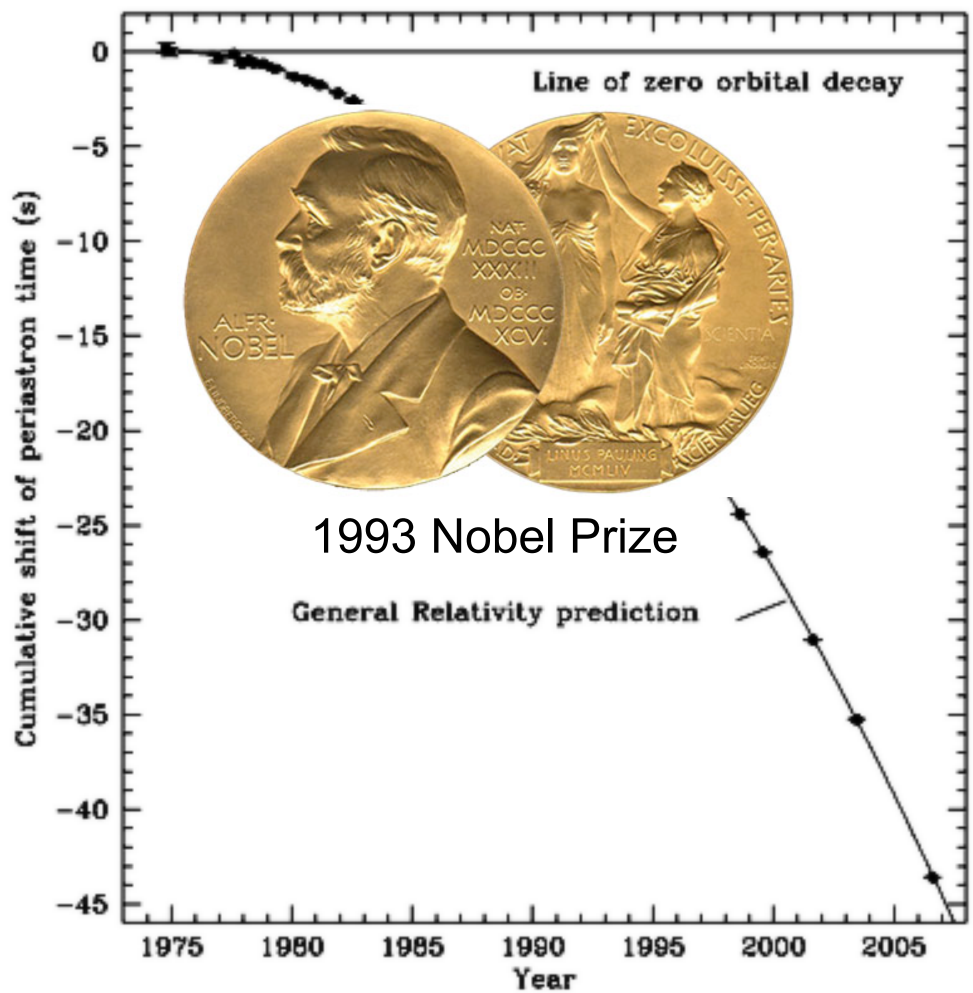
After years of observation it became clear, that the pulsars spiral towards each other.

Conclusion: They loose energy (fast). Immediate question: How?!?



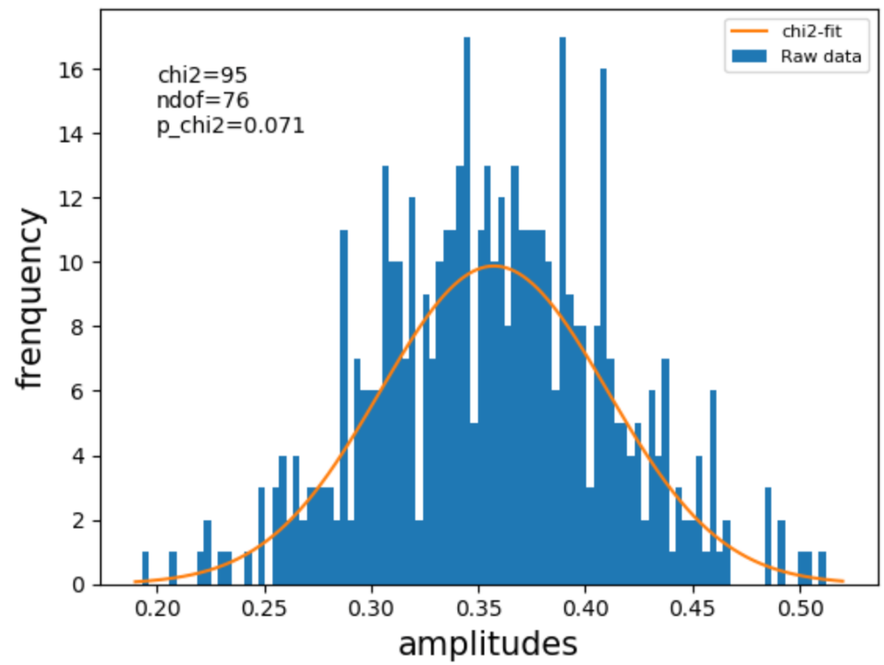
The answer is gravitational waves. What is normally a tiny effect is here so large (100x suns output!), that we (with a natural high precision clock) can see it.

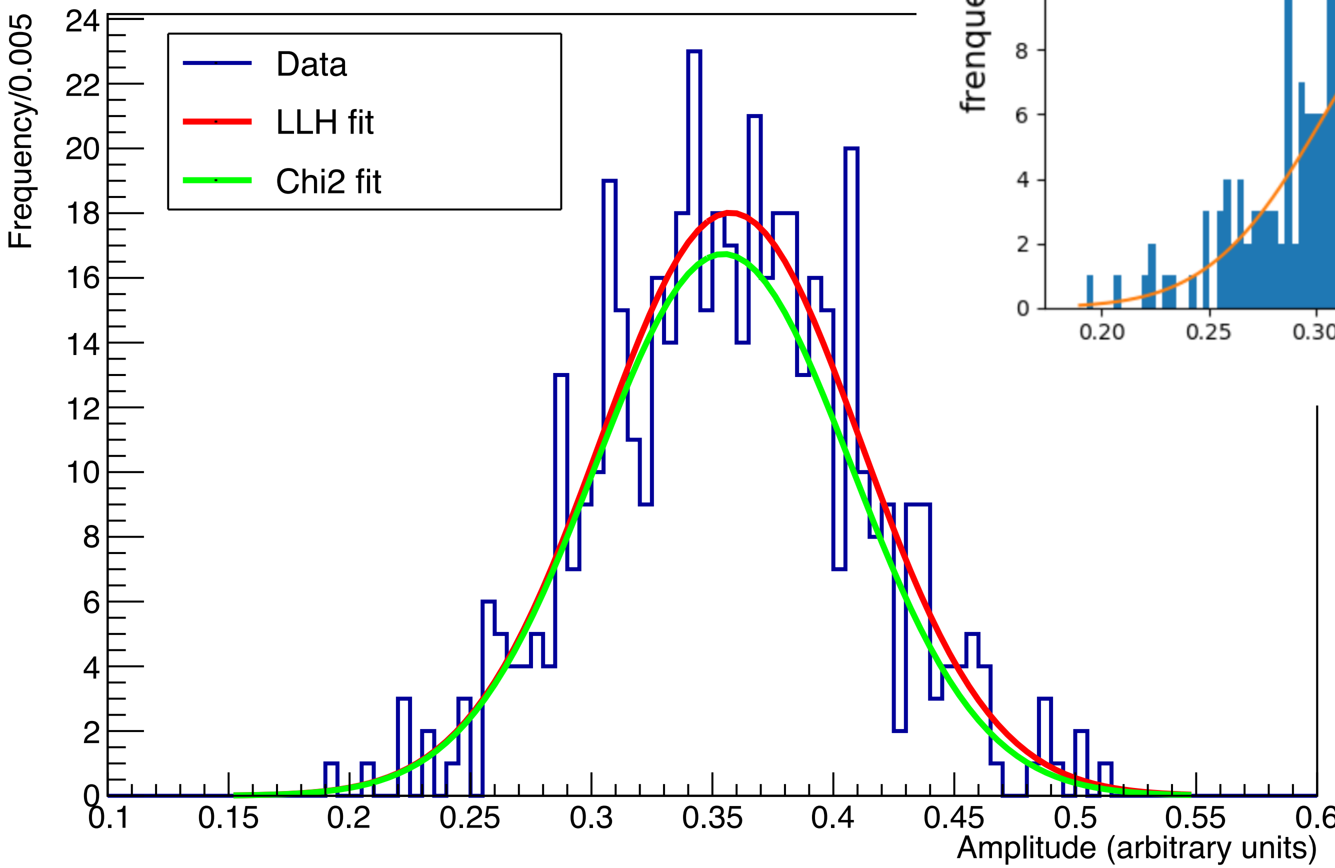
(Sun+Earth: 200 W, 10^{-15} m radius "lost" pr. day)



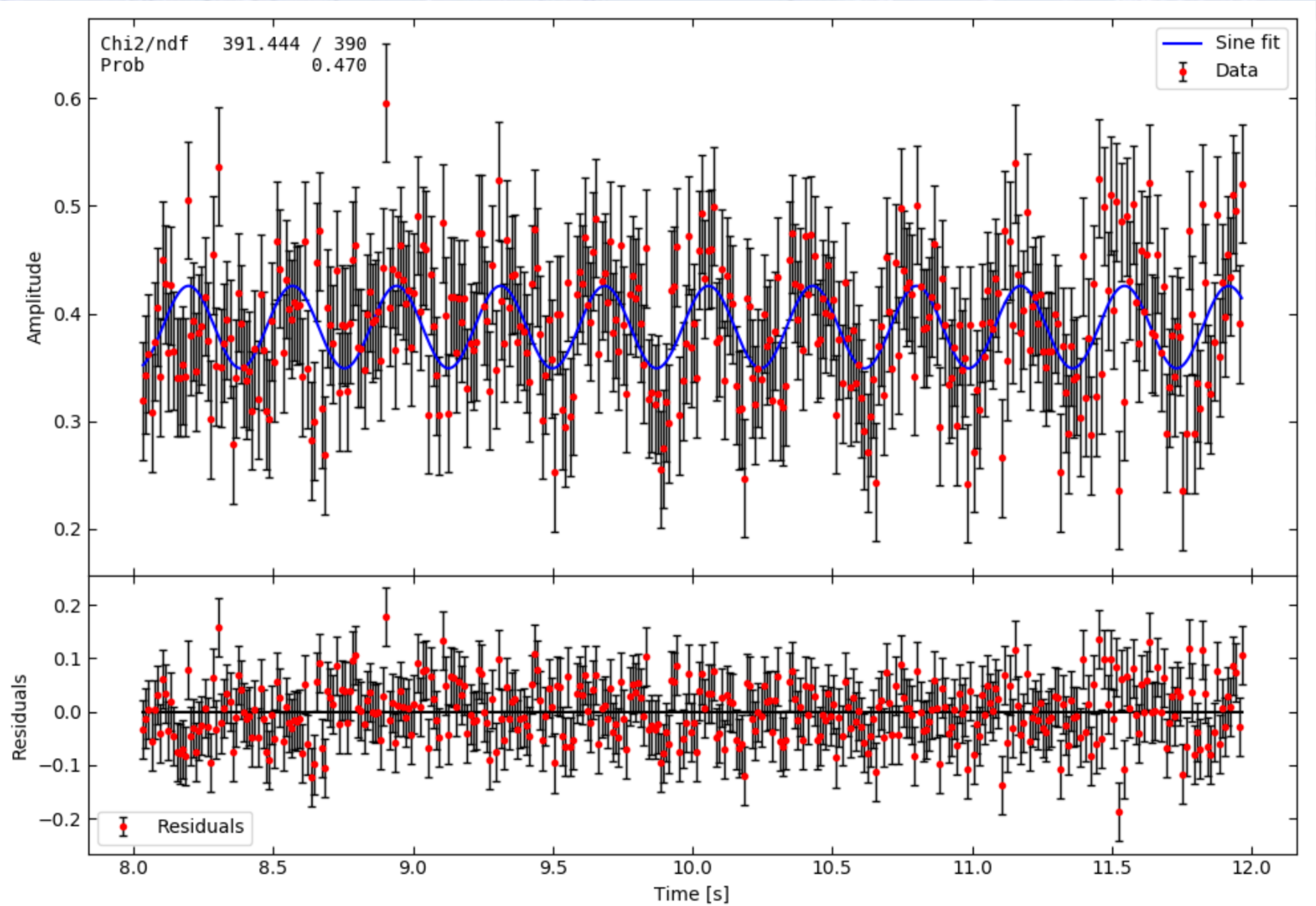
Problem 5.1

The distribution is Gaussian, and the RMS (in amplitude) is 0.054.
Binning had a (slight) effect, so careful!!!





Problem 5.1



Chi2/ndf 391.444 / 390
Prob 0.470
Sine fit
Data
Amplitude
Residuals
Residuals
Time [s]
0.6
0.5
0.4
0.3
0.2
0.2
0.1
0.0
−0.1
−0.2
8.0
8.5
9.0
9.5
10.0
10.5
11.0
11.5
12.0

Problem 5.1

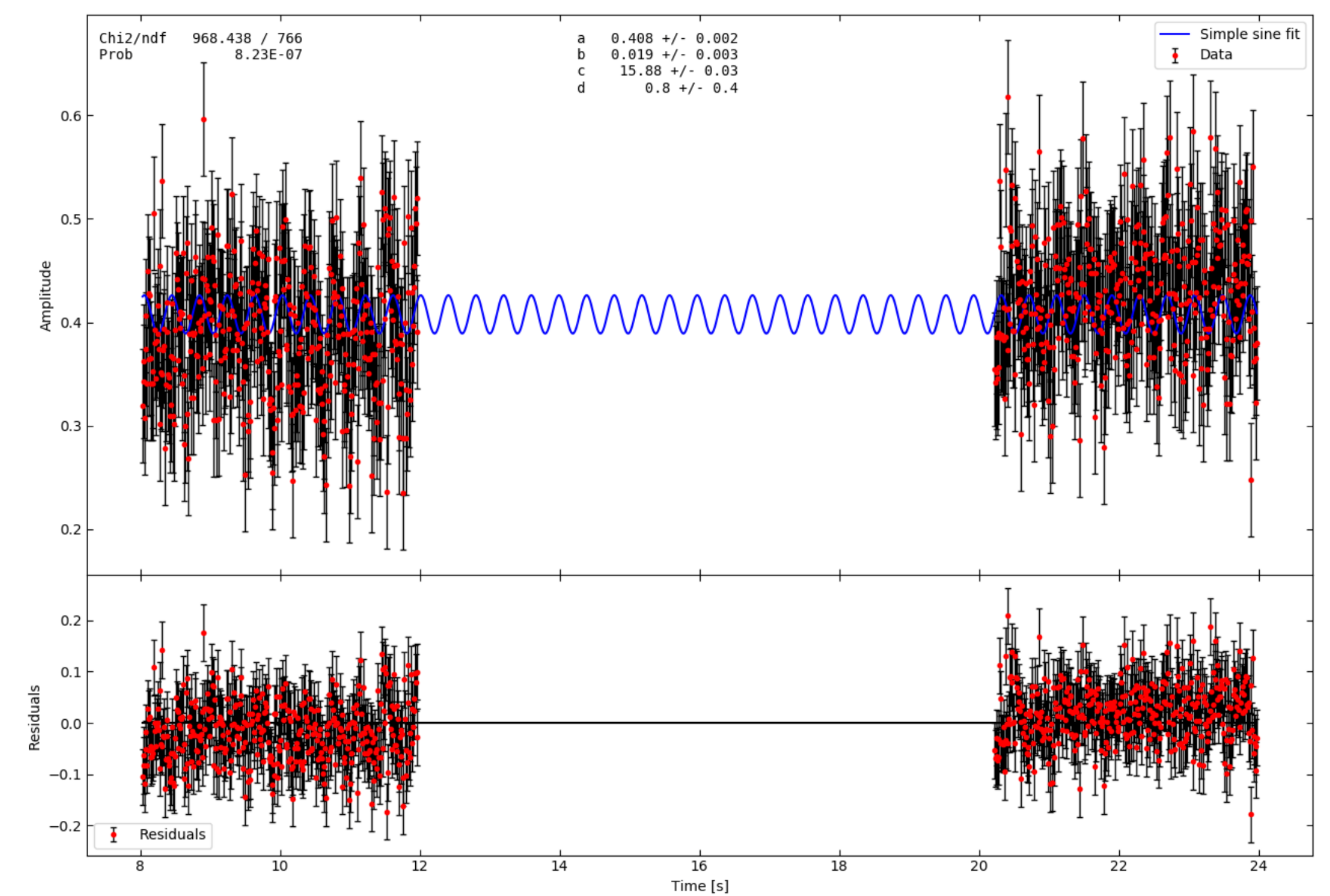
The two observation periods had the same period and amplitude, but the background noise level had changed.

One could either change this with a step function, but a better and more reasonable solution is to include a linear term. This is also suggested by the calibration run, which was (purposely) put into the same timeline.

	a	b	c	d
first run	0.387 ± 0.003	0.038 ± 0.004	16.89 ± 0.10	11.2 ± 1.0
second run	0.429 ± 0.003	0.034 ± 0.004	16.95 ± 0.11	25 ± 2
σ	-11	0.84	-0.44	-5.2
p-value	$1.9 \cdot 10^{-28}$	0.200	0.330	$1.0 \cdot 10^{-7}$

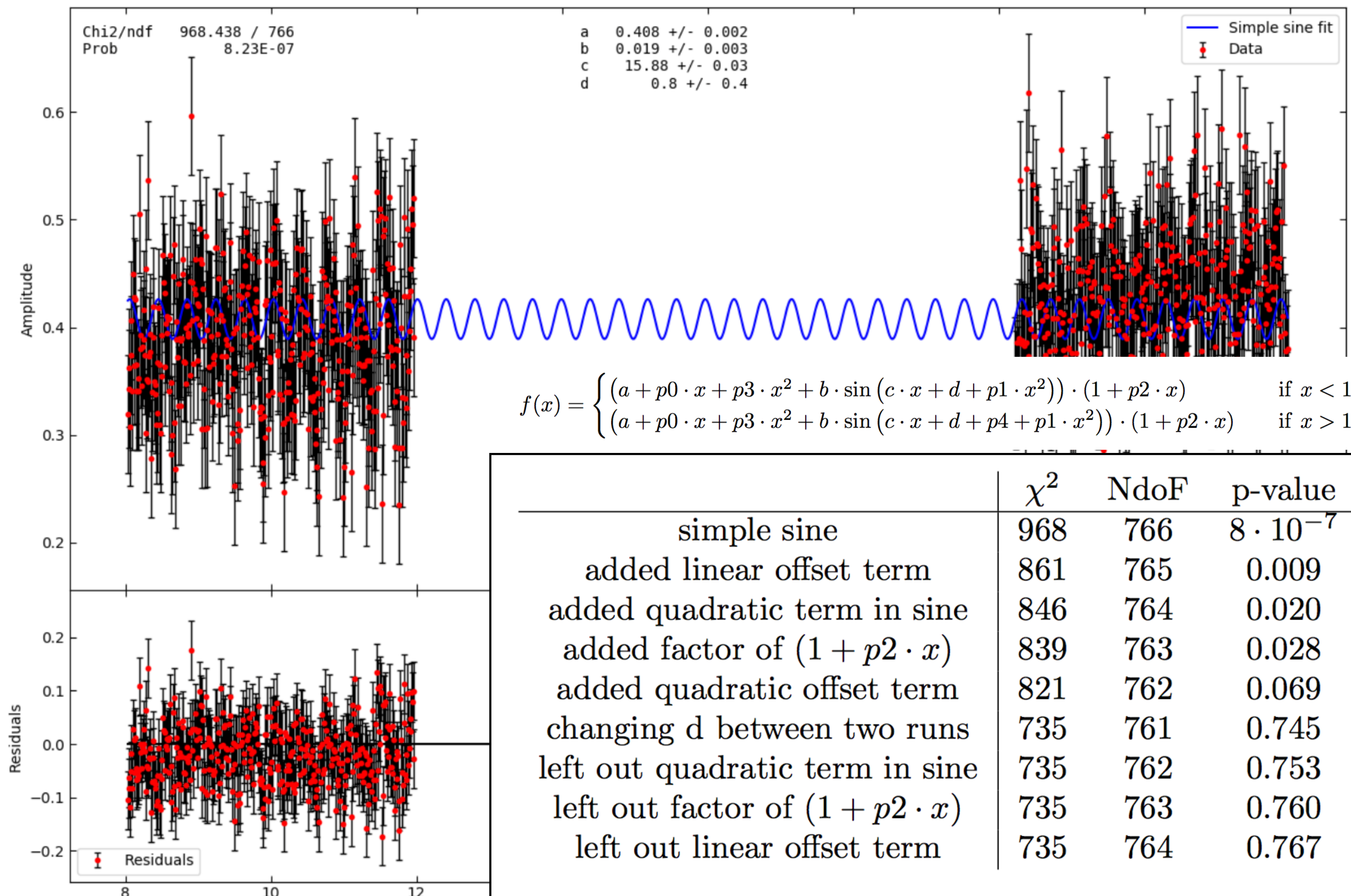
Fit value	First run	Second run	Deviation	Comment
A	0.038 ± 0.004	0.034 ± 0.004	0.85σ	Not significant
k	0.387 ± 0.003	0.429 ± 0.003	10.7σ	Highly significant
ω	$16.89 \pm 0.10\ \text{s}^{-1}$	$16.95 \pm 0.10\ \text{s}^{-1}$	0.44σ	Not significant
t_0	9.685 ± 0.007 s	9.33 ± 0.08 s	4.5σ	Significant, but unimportant

Problem 5.1



Chi2/ndf 968.438 / 766
Prob 8.23E-07
a 0.408 +/- 0.002
b 0.019 +/- 0.003
c 15.88 +/- 0.03
d 0.8 +/- 0.4
Simple sine fit
Data
Amplitude
Residuals
Residuals
Time [s]

Problem 5.1

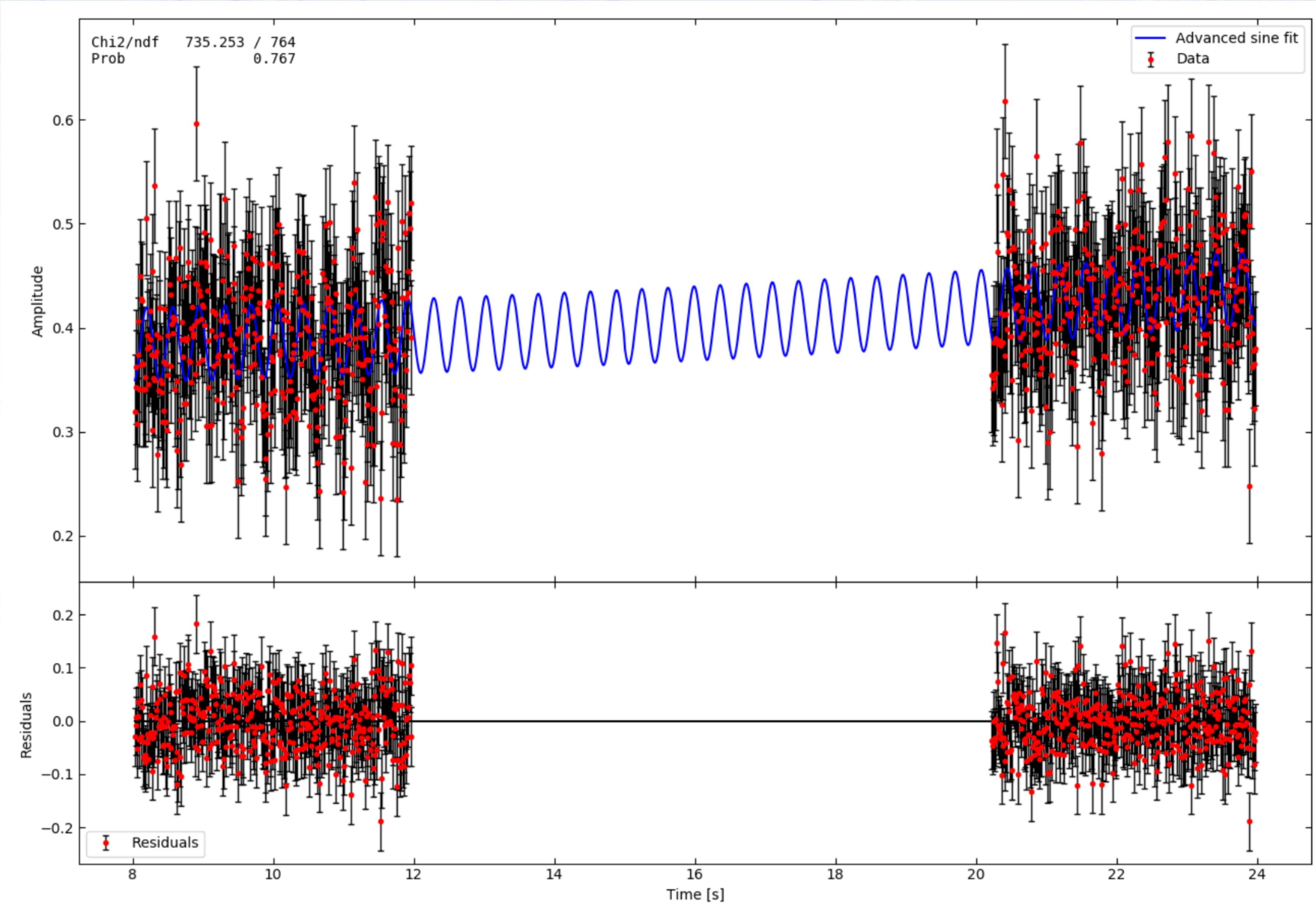


$$f(x) = \begin{cases} (a + p0 \cdot x + p3 \cdot x^2 + b \cdot \sin(c \cdot x + d + p1 \cdot x^2)) \cdot (1 + p2 \cdot x) & \text{if } x < 15 \text{ s} \\ (a + p0 \cdot x + p3 \cdot x^2 + b \cdot \sin(c \cdot x + d + p4 + p1 \cdot x^2)) \cdot (1 + p2 \cdot x) & \text{if } x > 15 \text{ s} \end{cases}$$

	χ^2	NdoF	p-value
simple sine	968	766	$8 \cdot 10^{-7}$
added linear offset term	861	765	0.009
added quadratic term in sine	846	764	0.020
added factor of $(1 + p2 \cdot x)$	839	763	0.028
added quadratic offset term	821	762	0.069
changing d between two runs	735	761	0.745
left out quadratic term in sine	735	762	0.753
left out factor of $(1 + p2 \cdot x)$	735	763	0.760
left out linear offset term	735	764	0.767

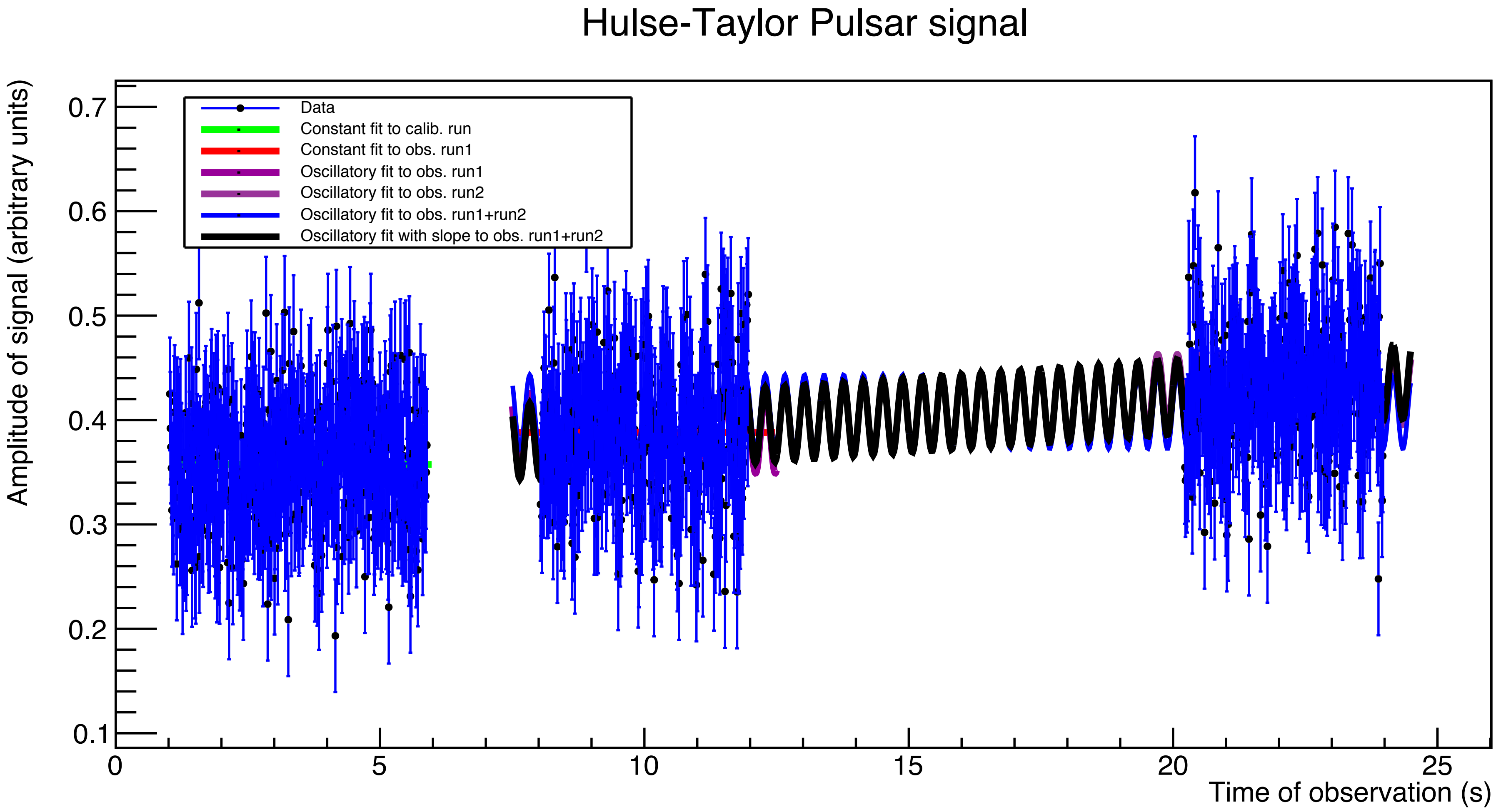
Table 6: Performance changes for the sine fit

Problem 5.1



Chi2/ndf 735.253 / 764
Prob 0.767
Advanced sine fit
Data
Amplitude
0.2
0.3
0.4
0.5
0.6
Residuals
−0.2
−0.1
0.0
0.1
0.2
Residuals
8
10
12
14
16
18
20
22
24
Time [s]

Problem 5.1



Hulse-Taylor Pulsar signal
Data
Constant fit to calib. run
Constant fit to obs. run1
Oscillatory fit to obs. run1
Oscillatory fit to obs. run2
Oscillatory fit to obs. run1+run2
Oscillatory fit with slope to obs. run1+run2
Amplitude of signal (arbitrary units)
Time of observation (s)
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0
5
10
15
20
25

Problem 5 problems

- Plotting histograms as lines
- Plotting error bar plots as (connected) lines (or even without errors)
- Plotting histograms of amplitude (in observation runs) and not amplitude vs time
- Fitting a constant != fitting a straight line
- Saying that slope == 0 means that data is consistent with being constant (only approx.)
- **Only providing the reduced chi^2**
- Forgetting to include the actual graphs and figures
- Shifting the second observation run closer in time to the first one (i.e. removing the space/time between the 2 runs and thus altering the data)
- Using vague/non-statistical formulations "almost similar" "seems like" "close enough to"
- **Proving hypotheses** (in contrary to disproving wrong hypotheses)
- Not defining which functions they are fitting with (or what the fit parameters are)
- Not remembering the phase in the oscillating function
- "N_dof hacking" i.e. fitting first, then removing background, and then refitting again - only now with 1 dof less (=> better P value).

<u>Annoying observations:</u>

- Not making the final result clear (which one of the intermediate results is the final one)
- Not stating which problem this is the solution to
- Writing the results/answers only in the figure captions
- Writing the results/answers only in the figure itself (TP: Or not at all!)

Typical mistakes

While some mistakes were individual, there we clear patterns in mistakes:

- stats.poisson.sf(4, lambda) is summing ABOVE four, not including 4...
- Not doing weighted mean, when errors are available - or even when asked to!!!
- Hit 'n Miss: a lot of people not sampling the correct function.
  Try to plot it in e.g. Maple/WolframAlpha to test
- People showing mean +/- RMS!! Sigma_mean != RMS.
- People not quantifying, e.g. "yes the data looks Gaussian"
- Always add uncertainties, when possible; for 4.2.1 and 5.1.3 you need them to calculate the distances (many did not calculate the distances at all!)
- Do not use the KS test on data you believe come from a discrete distribution.
  scipy's ks_2samp() cannot be used on binned data, either!
- Often, people would fit something and get extremely high or low chi2 values without commenting on the obviously wrong values! The error might often be that people forgot to add the uncertainty to the fitting code.

Comment on code sharing!

Out of interest, we ran Moss (Measure Of Software Similarity) on last years code, which is an automatic system for determining the similarity of programs (e.g. detecting plagiarism in programs).
Nothing suspicious was found at that time. Thank you!

Moss Results

Wed Jan 24 16:21:52 PST 2018

Options -l python -m 10

Applied statistics 2017

[How to Read the Results | Tips | FAQ | Contact | Submission Scripts | Credits]

File 1	File 2	Lines Matched
[redacted] (15%)	[redacted].py (12%)	178
[redacted] (13%)	[redacted]py (10%)	147
[redacted]py (9%)	[redacted]y (9%)	158
[redacted]py (14%)	[redacted].py (10%)	325
[redacted].py (6%)	[redacted].py (4%)	148
[redacted]py (8%)	[redacted]py (11%)	117
[redacted]py (5%)	[redacted].py (4%)	321
[redacted]py (11%)	[redacted]py (11%)	186
[redacted]py (4%)	[redacted].py (6%)	130
[redacted].py (9%)	[redacted]py (6%)	129
[redacted].py (5%)	[redacted]py (8%)	132
[redacted]py (6%)	[redacted]py (8%)	120

Your results…. (100% done)

Your final results

	1.1	1.2	1.3	1.4.1	1.4.2	2.1	2.2.1	2.2.2	2.2.3	3.1.1	3.1.2	3.1.3	3.1.4	3.1.5	3.1.6	4.1.1	4.1.2	4.1.3	4.1.4	4.2.1	4.2.2	4.2.3	5.1.1	5.1.2	5.1.3	5.1.4	Total
	5	**5**	**5**	**4**	**4**	**6**	**3**	**3**	**3**	**3**	**3**	**3**	**3**	**3**	**3**	**2**	**3**	**3**	**6**	**4**	**4**	**4**	**4**	**4**	**4**	**6**	**100**
pmc938	5	5	5	4	4	5	3	3	0	3	3	0	1	0	0	2	3	3	6	4	3	3	4	2.5	3	3.5	**78**
dlw792	5	5	5	4	4	4.5	3	3	0	2	2	0	2	0	0	0	2	3	5	2	1	0	3.5	2	0	0	**58**
mvp544	5	0	5	2	2.5	4.5	3	2	2	2	1	1	0	0	0	0	2	3	1	2	2	0	3	2	1	2	**48**
dbz341	5	5	5	4	4	5	3	2	2	3	3	2	1	0	0	2	2.5	1	2	2	2.5	1	2.5	1	0	0	**60.5**
AE	4	5	2	4	4	5	3	3	1	2	3	1	0	0	0	2	3	3	0	4	4	3	3	2	2.5	2	**65.5**
bpg317	5	5	5	4	5	5	3	3	2	3	3	3	2	3	3	1	3	3.5	8	4	6	4	3	3	4	4	**97.5**
rjp530	5	0	5	4	2.5	6	3	3	0	3	3	3	2	2	1	0	1	3	7	2	4	4	1	0	0	0	**64.5**
wzx358	5	5	5	3.5	3.5	6	3	3	2	3	3	3	3	3	2	0	2	4	6	2	4	4	3	2.5	3	4	**87.5**
ntd132	5	5	2	4	4	4	3	2	2	3	3	1	1	2	0	2	3	3	2	4	3	2	1	1	0.5	0	**62.5**
msd988	5	5	5	4.5	4	6	2.5	3	1	3	3	3	3	3	3	2	3	4	5	3	5	3.5	4	3.5	2	5.5	**94.5**
rcn776	5	3	5	4	3	5	3	3	3	3	3	3	2	3	2	0	2	4	6	3	4	4	3.5	3	4	6	**89.5**
AK	5	5	5	4	4	5	3	3	1	3	3	1	2	1	3	2	3	3.5	6	4	4	3	3	2	2	3	**83.5**
xvm402	5	5	5	4	4	5	3	3	3	3	3	3	2	3	1	0	3	4	6	3	4	4	2.5	4	3	6	**91.5**
fsj413	5	5	5	4	4	5	3	3	0	2	1	1	1	2	3	2	3	3	3	2	3.5	3	3.5	1.5	2	2	**72.5**
flv917	5	0	5	2	0	5	3	1.5	1	3	3	1	1	1	0	0	2	3	1	3	2	1	3.5	2	0	0	**49**
wgx965	5	5	5	4	4	5	2	2	1	3	2	2	2	3	2	3	2.5	3	5	3	4	3.5	4	2	2	5	**84**
wpl223	5	5	5	4	5	5	3	3	3	3	3	2	2	3	2	2	3	3	7	2	4	3	4	4	4	6	**95**
psh270	**5**	**0**	**5**	**1**	**0**	**6**	**1**	**1**	**1**	**3**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**3**	**1**	**2**	**1**	**0**	**3.5**	**0**	**0**	**0**	**34.5**
krn362	5	5	5	4	4	5	3	3	3	3	3	2	2	3	2	2	2	3.5	7	4	4	0	4	1	0	0	**79.5**
bpq912	5	4.5	5	4	0	4.5	3	1.5	0	3	3	0	1	0	0	0	1	0	0	2	2	0	0	0	0	0	**39.5**
hwl395	5	5	5	4	4	6	3	3	3	3	3	3	3	3	3	2	4	4	6	3	1	4	4	3	2	4	**93**
pjt158	5	5	5	4	3	5	3	1	2	3	3	3	3	0	1	0	2	3	7	2	4	4	4	3.5	3	5.5	**84**
kvc457	5	1	5	4	2	4	3	3	4	3	3	3	2	2	0	0	2	3	7	3	5	4	4	3.5	4	5.5	**85**
rmb528	5	5	5	4	4	5	3	3	1	2	2	2	1	2	2	2	2	4	2	2	2	2	3	2	2.5	2.5	**72**
vpr254	5	5	5	4	4	5	3	3	2	3	3	1	2	3	0	2	3	3	6	4	4	2.5	3	2.5	1	2.5	**81.5**
fhr821	5	0	3	4	4	6	3	3	3	3	3	2	3	3	3	0	2	3.5	7	3	4	4	3	4	2.5	5	**86**
xls818	5	5	5	4	4	5	3	2	0	2	1	2	2	0	0	0	2	3.5	1	2	4	3	0	0	0	0	**55.5**
lnj387	5	0	5	4	4	5	3	2	3	2	3	3	2	3	3	0	2	3	7	3	4	3	3	3.5	3.5	5	**84**
grd219	5	5	5	4	4	4	3	3	2	3	3	3	2	0	1	0	2	3	6	2	1	0	3.5	3	3.5	4	**75**
ctl564	0	0	2	4	2	4	3	2	0	2	0	1	2	0	0	2	2	3	2	4	3	3	3.5	1	0.5	0	**46**
lxd483																											**0**
mfz821	5	5	5	0	0	5	1	3	1	3	2	0	1	0	0	0	0	0	0	0	0	0	3.5	1	0	0	**35.5**
ksb688	5	5	5	2	4	5	3	3	2	3	3	3	2	2	1	0	2	3	7	4	2	1.5	2.5	4	2	5	**81**
fbg128	5	0	2	4	4	4	2.5	2	2.5	3	1	2	1	1	0	2	2.5	3	2	2	1	4	3	0.5	0	0	**54**
tpl245	5	5	5	4	4	5	3	3	2	3	3	3	3	3	1	2	2.5	4	6	3	4	0	3.5	4	4	6	**91**
fsd776	5	5	5	4	4	5	3	3	2	3	3	2	3	0	0	2	3	3	5	2	2	3	4	4	4	3	**82**

Your final results

	1.1	1.2	1.3	1.4.1	1.4.2	2.1	2.2.1	2.2.2	2.2.3	3.1.1	3.1.2	3.1.3	3.1.4	3.1.5	3.1.6	4.1.1	4.1.2	4.1.3	4.1.4	4.2.1	4.2.2	4.2.3	5.1.1	5.1.2	5.1.3	5.1.4	Total
	5	**5**	**5**	**4**	**4**	**6**	**3**	**3**	**3**	**3**	**3**	**3**	**3**	**3**	**3**	**2**	**3**	**3**	**6**	**4**	**4**	**4**	**4**	**4**	**4**	**6**	**100**
ndx115	5	0	5	4	2	4	3	2	0	2	3	1	1	3	0	2	3	3	0	2	2	2	4	3.5	3.5	4	**64**
wtq360	5	0	5	2	2	5	3	2	2	2	1	0	0	1	0	0	2	3	1	2	3	0	3.5	3	2.5	4	**54**
wsl764	5	0	5	3.5	4	4	3	3	2	2	3	2	2	1	3	2	3.5	3	8	4	3.5	3	3.5	3	3	4	**83**
rfd312	5	5	5	4	4	5	3	3	2	3	3	3	2	3	2	2	2.5	3	7	3	6	3.5	4	4	4	6	**97**
nqj779	4	0	5	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**13**
txb814	5	5	5	4	3	6	3	3	2	3	3	2	2	2	3	0	2	4	6	3	4	4	4	4	3.5	5.5	**91**
dhg491	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	**2**
kzl542	5	5	5	4	3	6	3	3	3	3	3	3	1	3	3	2	4	3	6	4	4	0	4	4	4	6	**94**
mdv971	5	5	5	4	4	6	3	3	3	3	3	3	2	3	2	2	3	3	7	4	3	2	3.5	3	2.5	5	**92**
kxz925	5	5	5	2	0	5	3	3	2	3	3	1	3	3	3	0	2	3.5	6	3	4	4	3	4	2.5	4	**82**
xjg423	4	0	5	4	4	5	3	2	2	0	0	0	0	0	0	0	2	2	0	0	0	0	2.5	0.5	0	0	**36**
svq398	5	0	0	3	2	2	3	3	2.5	0	0	0	0	0	0	0	2	3	5	2	2	0	0	0	0	0	**34.5**
gtf274	5	2	5	4	4	3	1	2	3	3	2	0	1	1	0	2	3	3	1	4	0	0	2	0	0	0	**51**
kjf816	5	0	5	4	3	5	3	3	3	2	3	2	3	3	2	0	1	3	6	2	4	0	3.5	4	2.5	5.5	**77.5**
gfs647	5	0	5	4	4	5	3	3	3	3	3	3	2	3	3	2	3	3	8	2	4	4	4	4	3.5	6	**92.5**
klq995																											**0**
dxw176	5	5	5	4	3	6	3	3	1	2	1	2	2	2	3	0	2	3	7	2	4	2.5	2	2.5	3	4.5	**79.5**
chp915	5	5	5	4	4	6	3	1	3	1	2	0	1	2	2	2	3	3	2	3.5	3	0	0	0	0	0	**60.5**
wjz988	5	0	5	4	4	5	3	3	2	2	3	2	2	0	0	2	3	3	7	3	2	4	3	3	2	3	**75**
mqt845	5	5	5	4	4	6	3	3	2	2	3	2	1	3	2	0	2	3	7	2	4	4	4	3.5	3	5	**87.5**
mfq130	5	5	5	4	0	4	3	0	0	2	1	0	1	1	0	0	2	1.5	0	2	0	0	2.5	2.5	0	0.5	**42**
bcp320	5	5	5	4	3	5	2	3	2	2	3	0	0	0	0	0	2	3	5	2	2	0	3	2	1.5	2	**61.5**
kmz569	5	0	5	3.5	3	4	3	3	2	2	3	2	2	1	0	0	2	3.5	1	2	2	0	3	3	3.5	0	**58.5**
xhn538	**5**	**5**	**5**	**4**	**3**	**6**	**3**	**3**	**2**	**3**	**3**	**3**	**3**	**3**	**0**	**1**	**1**	**2**	**6**	**4**	**4**	**4**	**4**	**4**	**4**	**6**	**91**
rqg746	5	4.5	5	4	4	5	3	3	2	3	3	2	2	2	3	2	3.5	3	5	4	4	0	4	2	0	0	**78**
vlf201	**5**	**3**	**5**	**4**	**4**	5	2	3	2	2	3	0	1	0	0	0	2	3	0	2	2	0	1.5	1.5	1	0	**52**
JW	5	5	5	4	2	6	3	3	1	2	3	0	0	3	3	0	2	3	2	2	3	2	3.5	2	2	1	**67.5**
fvb662	5	4.5	5	4	4	6	3	4	3	1	3	3	3	3	3	2	3	3	6	3.5	4	4	3.5	3	3.5	6	**96**
pgz731	5	4.5	5	4	0	5	3	2	0	2	2	0	1	1	0	0	2	3	5	2	1	0	2	0.5	0	0	**50**
kmx917	5	0	5	4	3	5	3	2	2	2	3	1	0	1	0	2	3	3	1.5	2	2	0	0.5	0	0	0	**50**
pjl124	5	5	5	4	4	5	3	3	2	2	3	2	2	2	1	2	3	3	6	4	4	2	2	3.5	2.5	3	**83**
twg529	5	5	5	4	4	4	3	1	2.5	2	3	2	2	3	0	1	2	3	1	2	1	2	2.5	1.5	0	0	**61.5**
qbp511	0	0	3	2	1	5	2	3	2	2	2	1	2	2	1	2	3	3	0	2	4	1	2.5	1.5	0.5	0	**47.5**
fjm796	5	5	5	4	4	6	3	3	3	3	3	3	2	3	3	2	3	3.5	7	4	2	4	3	3	2.5	4	**93**
zkp986	5	5	5	4	2	4	3	1	3	1	1	0	0	0	0	0	1	3	0.5	2	1	1	3	1.5	0.5	0	**47.5**
kpr426	5	5	5	4	4	4.5	3	3	2.5	2	3	0	1	2	2	0	1	3	1	2	2	0	3.5	3	2.5	2.5	**66.5**

Your final results

	1.1	1.2	1.3	1.4.1	1.4.2	2.1	2.2.1	2.2.2	2.2.3	3.1.1	3.1.2	3.1.3	3.1.4	3.1.5	3.1.6	4.1.1	4.1.2	4.1.3	4.1.4	4.2.1	4.2.2	4.2.3	5.1.1	5.1.2	5.1.3	5.1.4	Total
	5	**5**	**5**	**4**	**4**	**6**	**3**	**3**	**3**	**3**	**3**	**3**	**3**	**3**	**3**	**2**	**3**	**3**	**6**	**4**	**4**	**4**	**4**	**4**	**4**	**6**	**100**
nlb108	**5**	**1**	**5**	**4**	**4**	4.5	3	2	2	**3**	**3**	**2**	**2**	**3**	**2**	**2**	**3**	**3**	**0**	**4**	**2**	**2**	3	2	1	2	**69.5**
LL	5	5	5	4	1	1	3	2.5	0	2	3	2	1	1	0	1	2	2	1	2	2	0	1.5	1.5	1	0	**49.5**
lkg182	5	3	5	2	0	5	3	2.5	2	3	2	0	2	1	2	2	3	3	6	4	3.5	2	3.5	3.5	2.5	5.5	**76**
lrc229	5	3	5	4	4	4	2	1	2	2	1	2	1	2	1	1.5	2	3	4	2	3.5	0	3.5	3.5	2	4	**68**
jnc582	4	0	5	4	4	6	3	1	1	1	0	2	2	0	0	1.5	2.5	3	5.5	3	3	0	3	1.5	0	0	**56**
pqc735	5	0	5	4	4	4	3	3	3	2	1	1	2	3	2	0	1	3	6	4	2	0	4	3.5	2	0	**67.5**
jwt903	5	5	5	4	4	5	3	3	1	3	3	2	3	3	2	2	2	3	6	4	4	0	4	1	0	0	**77**
bvf334	4	0	5	4	0	6	3	2	2	3	3	3	3	3	1	2	3	3	7	2	4	2	4	3	3.5	5	**80.5**
pbv615	5	5	5	4	4	5	3	4	3	2	3	2	2	2	3	2	3	3	6	3	4	4	3	3	2	4	**89**
zvn474	5	0	5	4	4	5	3	3	1	3	3	2	1	3	0	2	3	3	0.5	3.5	1	0	3.5	2	3	4	**67.5**
qls370	5	5	5	4	4	5	3	3	2	3	3	3	3	3	1	2	3	3	6	3.5	4	0	4	3.5	2.5	4.5	**88**
xrl936	5	5	5	2	1	5	3	3	3							2	3	3	6	4	2	2	3	3.5	3.5	5.5	**69.5**
wxj562	4	5	5	4	3	5	3	3	2	3	1	1	2	2	1	0	2	3	1	0	0	0	3	3	2.5	3	**61.5**
jpw919	5	0	5	4	3	5	2	2	2	2	3	2	2	2	2	2	2	3	5	2	2	2	3	3.5	2.5	4	**72**
xgl132																											**0**
bqh863																											**0**
dhm812	5	1	5	4	4	4	3	2	2	2	3	1	2	1	0	2	3	3	0	2	2	0	3	1	0	0	**55**
rqg460	4	0	2	3	2	4	3	2	2.5	2	3	2	2	1	3	2	3	3	6	2	4	3	3.5	2.5	1	3	**68.5**
hfk576	5	0.5	5	2	2	5	3	2	2	3	1	0	0	0	0	0	2	3	1	2	1	0	3	0	0	0	**42.5**
wvd868	5	0	4	4	3	6	3	2	2	2	3	2	0	0	0	0	2	3	4	2	3.5	2	1.5	2	2	3	**61**
flr522	5	5	5	4	4	6	3	3	2	3	3	2	1	0	0	0	2	3	7	3	4	4	3	3	2.5	6	**83.5**
jzg995	5	2	5	4	2	4	2.5	2	0	3	3	2	0	2	0	2	3	3	0	0	0	0	0	0	0	0	**44.5**
gcf945	**5**	**2**	**5**	**4**	**3.5**	4	1	2	2	**2**	**2.5**	**3**	**3**	**2**	**1**	**1.5**	**1.5**	**3**	**5**	**2**	**3.5**	**1**	2.5	2	2.5	3	**69.5**
lpk331	5	5	5	2	0	6	3	3	3	3	3	3	2	3	2	0	2	3	7	2	4	2	3.5	3.5	3	0.5	**78.5**
jvf427	5	0	5	4	3	5	3	2	1	2	3	2	2	1	2	2	3	3	6	2	2	0	3.5	2	2	2	**67.5**
jrt134	5	5	5	4	4	4	3	3	3	3	3	2	3	3	0	0	2	4	6	3	4	3.5	3.5	4	3	5.5	**88.5**
mbv396	5	0	5	3.5	4	4.5	3	3	2	2	3	0	0	0	0	2	2.5	3	1	2	2	0	1.5	0.5	0	0	**49.5**
tmv646	5	0	0	4	1	4.5	3	2	2	3	0	0	0	0	0	2	3	3	1	2	1	0	0	0	0	0	**36.5**
PH	5	5	0	2	0	3	3	1	3	2	3	0	2	0	0	0	1	2.5	1	3	4	4	3.5	3.5	3	5.5	**60**
fbg450	5	5	4	4	4	5	3	3	3	1	3	1	3	3	1	0	2	3	5	3	4	0	3.5	3	2	0	**73.5**
dxj847	5	4	5	4	4	4	1	2	2	2	2	0	1	3	0	1.5	3	3	5	2	3	1	3.5	3	2.5	3.5	**70**
ctq599	5	5	5	4	3	5	3	3	1	1	3	3	3	3	1	0	2	3	2.5	2	4	0	3.5	4	3.5	5.5	**78**
mnb838	5	4	5	3.5	4	5	3	2	1	3	3	3	3	2	2	2	3	3	1	2	3	2	2	0	0	0	**66.5**
lnk170	5	3	5	4	4	5	3	2	1	2	3	3	2	2	0	2	3	3	1	2	3	1	2	0.5	0	0	**61.5**
sgq951	5	5	5	4	3	5	3	1	1	3	2	2	1	3	2	2	1.5	3	2	2	2	2	3.5	3	2.5	4	**72.5**
knb392	5	5	5	3	3	6	3	3	2	2	3	1	2	3	3	0	2	3	6	2	2	0	3.5	1	0	0	**68.5**

Your final results

	1.1	1.2	1.3	1.4.1	1.4.2	2.1	2.2.1	2.2.2	2.2.3	3.1.1	3.1.2	3.1.3	3.1.4	3.1.5	3.1.6	4.1.1	4.1.2	4.1.3	4.1.4	4.2.1	4.2.2	4.2.3	5.1.1	5.1.2	5.1.3	5.1.4	Total
	5	**5**	**5**	**4**	**4**	**6**	**3**	**3**	**3**	**3**	**3**	**3**	**3**	**3**	**3**	**2**	**3**	**3**	**6**	**4**	**4**	**4**	**4**	**4**	**4**	**6**	**100**
SD	5	5	5	4	3	5	3	2.5	2.5	2	2	0	2	1	2	0	2	3.5	0	4	5	2	4	3.5	3	6	77
zxp220	5	5	5	2	2	5	3	2.5	2	3	3	2	2	3	3	0	1	3	6	3.5	4	2	3.5	2	2.5	4	79
fkg212	5	5	5	4	3	6	3	3	2	2	1	2	1	1	0	0	2	3	4	2	4	2	2.5	3	2	2.5	70
dsc459	5	5	5	4	4	6	1	3	3	0	3	1	2	3	3	2	4	4	6	3.5	4	5	4	4	3	6	93.5
kjd405	5	0	5	4	4	5	3	2	2	2	3	3	2	2	2	0	1	3	6	2	4	1	3.5	1	0	0	65.5
dqz113	5	5	5	4	4	6	3	2	1	3	3	2	2	0	0	2	3.5	3	1	4	3.5	2	2.5	0	0	0	66.5
mkw755	5	5	5	4	3	6	3	3	2.5	3	1	3	2	3	3	0	1.5	3	7	2	4	4	3.5	3	3	4.5	87
szl778	5	5	5	4	4	6	3	3	2	3	2	2	3	3	2	0	2	3	1	2	4	4	4	3	3	4	82
qxp321	5	0	5	4	3	5	3	1	2	2	3	3	3	0	0	0	3	3	6	2	4	3	4	3	1.5	3	71.5
bfw491	5	5	5	4	4	6	3	3	2	0	2	2	1	2	2	0	2	1	5	4	3	2	3.5	3	1	2	72.5
bkm394	5	5	5	4	4	6	3	3	2	2	3	2	2	0	0	0	2	3	1	0	0	0	0.5	0	0	0	52.5
txd908	5	5	5	4	4	5	3	3	3	3	3	2	3	3	0	0	2	4	6	2	4	4	3.5	3	2.5	4	86
nth404	5	5	4	4	3	4	3	3	2	2	3	3	3	2	2	0	2.5	3.5	4	2	4	4	3.5	4	3	5	83.5
ksr966	5	5	5	4	3	5	3	2	2	2	3	2	2	2	0	0	2.5	2.5	7	2	4	4	4	4	2.5	5.5	83
cfr686	5	0	5	4	4	5	3	2	2	2	3	2	3	2	3	2	3.5	3.5	1	2.5	4	0	3.5	4	4	5	78
ntv598	5	5	5	4	4	5	3	3	2	3	3	3	2	2	3	2	3	3	6	4	4	2	4	3.5	3	6	92.5
rls244	4	5	5	3	0	5	3	3	0	2	3	2	2	2	1	0	2.5	3	7	3	4	0	3.5	3.5	3	6	75.5
mqb610	5	5	5	4	4	6	3	3	3	2	1	2	2	1	3	2	4	3	6	4	1	4	3.5	4	3	6	89.5
vcg135	4	5	4	4	2	2	0	2	0	2	3	0	0	0	0	1	2	3	0	2	1	1	1.5	2.5	2	3	47
dvc678	5	5	4	4	4	4	3	3	0	2	3	1	0	0	0	0	2.5	3	0	0	1	1	1.5	1	0.5	0	48.5