

# Applied Statistics Exam

Riz Fernando Noronha

January 20, 2023

## I

### 1.1

- We need to integrate a normalized Gaussian in the range 55 to 65. I numerically integrated it (as an analytical closed form solution for the gaussian integral doesn't exist) to obtain

$$\text{Fraction in } [55, 65] = 0.1747$$

- After sampling 120 random samples from the given distribution, I get

$$\begin{aligned}\mu &= 49.4 \\ \sigma_\mu &= 1.9\end{aligned}$$

- Had there been no correlation (independent test scores), the fraction who got over 60 in both tests would simply be

$$(\text{SF}_{\text{gaussian}}(60))^2 \quad \rho = 0$$

Where SF represents the survival function, or  $1 - \text{CDF}(x)$ . With perfect positive correlation, the people who get above 60 in one test are likely to get the same grade in the other test, and thus the fraction would be

$$\text{SF}_{\text{gaussian}}(60) \quad \rho = 1$$

Thus, with a correlation of  $0 \leq \rho \leq 1$ , it seems reasonable to expect that the fraction is given by

$$(\text{SF}_{\text{gaussian}}(60))^{2-\rho}$$

Which, for  $\rho = 0.6$ , gives us **19.28%**.

### 1.2

- The number of successes will follow a binomial distribution,

$$P_{\text{binomial}}(r, N, p) = \binom{N}{r} p^r (1-p)^{N-r}$$

where  $r$  represents the number of successes,  $N$  the number of trials, and  $p$  the probability of success.

$$\begin{aligned}P(11 \text{ or more wins}) &= \sum_{r=11}^{20} P_{\text{binomial}}(r, 20, 0.4) \\ &= 0.12752\end{aligned}$$

## II

### 2.1

As the partial derivatives themselves are differentiable,  $\frac{\partial z}{\partial x} \frac{\partial z}{\partial y} = \frac{\partial z}{\partial y} \frac{\partial z}{\partial x}$  and since the correlation matrix is symmetric ( $V_{x,y} = V_{y,x}$ ) we can combine the cross terms.

The error propagation formula is then

$$\sigma_z^2 = \left( \frac{\partial z}{\partial x} \right)_{\bar{x}, \bar{y}}^2 \sigma_x^2 + \left( \frac{\partial z}{\partial y} \right)_{\bar{x}, \bar{y}}^2 \sigma_y^2 + 2 \left( \frac{\partial z}{\partial x} \frac{\partial z}{\partial y} \right)_{\bar{x}, \bar{y}} V_{x,y}$$

•

$$\begin{aligned} \sigma_{z_1 \text{ from } x} &= \frac{\partial z_1}{\partial x} \sigma_x \\ &= (y e^{-y})_{\bar{y}} \sigma_x \\ &= \mathbf{0.00091} \\ \sigma_{z_1 \text{ from } y} &= \frac{\partial z_1}{\partial y} \sigma_y \\ &= (x e^{-y} - x y e^{-y})_{\bar{x}, \bar{y}} \sigma_y \\ &= \mathbf{0.20602} \end{aligned}$$

As  $|\sigma_{z_1 \text{ from } y}| > |\sigma_{z_1 \text{ from } x}|$ ,  $y$  is the major contributor to error on  $z_1$

•

$$\begin{aligned} \sigma_{z_1}^2 &= (y e^{-y})_{\bar{y}}^2 \sigma_x^2 + (x e^{-y} - x y e^{-y})_{\bar{x}, \bar{y}}^2 \sigma_y^2 + 2 [y e^{-y} (x e^{-y} - x y e^{-y})]_{\bar{x}, \bar{y}} \rho \sigma_x \sigma_y \\ &= 0.04260 \\ \sigma_{z_1} &= \mathbf{0.2064} \\ \sigma_{z_2}^2 &= \left( -\frac{(y+1)^3}{(x-1)^2} \right)_{\bar{x}, \bar{y}}^2 \sigma_x^2 + \left( \frac{3(y+1)^2}{x-1} \right)_{\bar{x}, \bar{y}}^2 \sigma_y^2 + 2 \left[ -\frac{(y+1)^3}{(x-1)^2} \cdot \frac{3(y+1)^2}{x-1} \right]_{\bar{x}, \bar{y}} \rho \sigma_x \sigma_y \\ &= 519.4874 \\ \sigma_{z_2} &= \mathbf{22.7923} \end{aligned}$$

- Assuming the errors on  $z_1$  and  $z_2$  are the correlated errors calculated above, we can calculate

$$\begin{aligned} \rho_{z_1, z_2} &= \frac{\mathbb{E}[(z_1 - \mu_{z_1})(z_2 - \mu_{z_2})]}{\sigma_{z_1} \sigma_{z_2}} \\ &= \frac{\int_{-1}^1 \int_{-1}^1 (z_1 - \mu_{z_1})(z_2 - \mu_{z_2}) dz_1 dz_2}{\sigma_{z_1} \sigma_{z_2}} \\ &= \mathbf{2.128} \end{aligned}$$

Something seems to have gone wrong, as the Pearson's correlation coefficient should be in the range  $[-1, 1]$ . I believe we should replace  $z_1$  and  $z_2$  in the integration step with  $x$  and  $y$ , and that would perhaps get the right result.

## 2.2

When we have multiple measurements with their own uncertainties, we can combine them through a weighted mean:

$$\hat{\mu} = \frac{\sum x_i / \sigma_i^2}{\sum 1 / \sigma_i^2}$$

$$\hat{\sigma}_\mu = \sqrt{\frac{1}{\sum 1 / \sigma_i^2}}$$

- Naively combining the measurements (discussed further below), we get

$$g = 5.279 \pm 0.060 \text{ g/cm}^3$$

- If we actually look at our data, we notice that the results are inconsistent with each other: there are only a few data points, and yet several points lie outside of  $1\sigma$  confidence intervals with each other. As a weighted mean is a chi-square fit of a constant function, we can calculate a ‘goodness of fit’ measure.

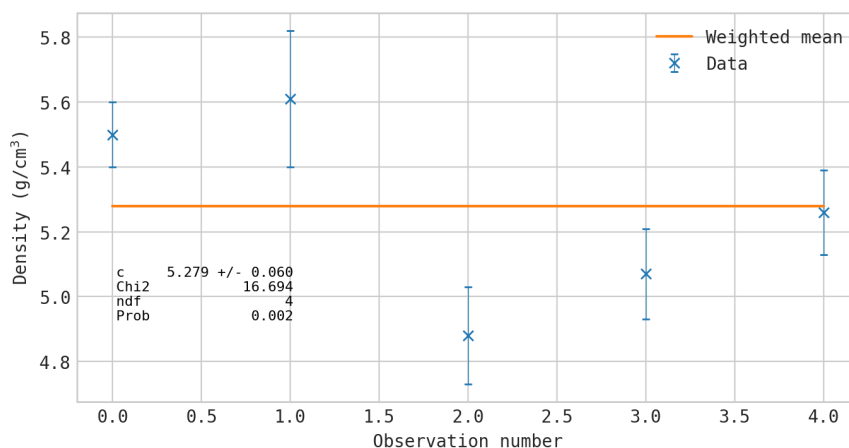


Figure 1: A chi-square fit to find the weighted mean of the experimental data. Notice that a majority of points do not contain the weighted mean value, within their error bars.

Figure 1 contains a plot of the data to see how it compares to the weighted mean. We get a  $\chi^2$  of 16.7, which (together with 4 degrees of freedom) gives us a chi-square probability of 0.002. The value should be  $\sim 0.5$ , and the low probability means that we have underestimated our errors. In order to make the data more consistent, I assume that the errors are misreported, and scale them all by a factor of 2, shown in Figure 2 below:

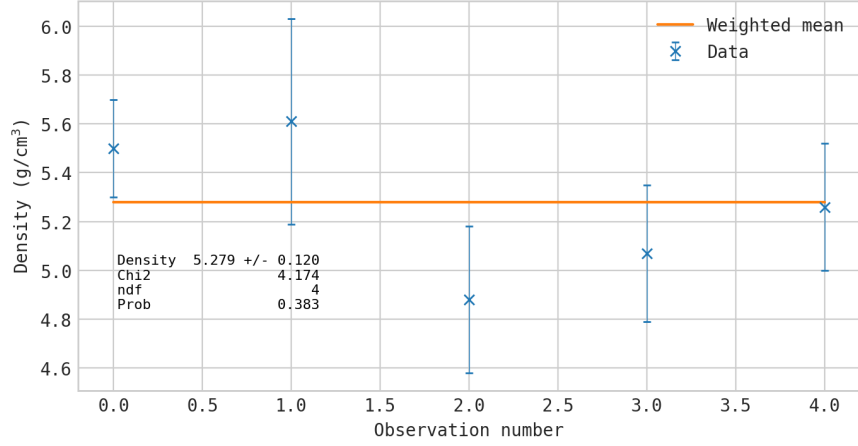


Figure 2: A chi-square fit to find the weighted mean of the experimental data, after doubling all the errors.

We get a much more realistic probability of 0.383 ( $\chi^2=4.2$ ,  $N_{dof}=4$ ), which inspires much more confidence than before. The best estimate we can make is thus  $\rho=5.28\pm0.12$ .

- To quantify how close our observation is, we can calculate the  $z$ -value of it, using the formula

$$z = \frac{\rho_{\text{true}} - \hat{\mu}_\rho}{\hat{\sigma}_\rho}$$

Which gives us a  $z$ -value of  $\sim 1.95$ . This tells us that our observation is around  $2\sigma$  away from the true value, and we should expect a value as bad as this or worse around 5% of the time.

## 2.3

- 

$$A = \pi a^2 \sqrt{1 - e^2}$$

$$= 2.3928$$

$$\sigma_A^2 = \left(2\pi a \sqrt{1 - e^2}\right)_{\bar{a}, \bar{e}}^2 \sigma_a^2 + \left(\frac{-\pi a^2 e}{\sqrt{1 - e^2}}\right)_{\bar{a}, \bar{e}}^2 \sigma_e^2$$

$$= 1.7127$$

$$\sigma_A = 1.3$$

The area of ellipse  $E$  is thus  $2.4 \pm 1.3$

- 

$$(\text{Lower bound error})^2 = \left(4\sqrt{2 - e^2}\right)_{\bar{e}}^2 \sigma_a^2 + \left(\frac{-4ae}{\sqrt{2 - e^2}}\right)_{\bar{a}, \bar{e}}^2 \sigma_e^2$$

$$= 1.8288$$

$$\text{Lower bound error} = 1.3523$$

$$(\text{Upper bound error})^2 = \left(\pi\sqrt{4 - 2e^2}\right)_{\bar{e}}^2 \sigma_a^2 + \left(\frac{-2\pi ae}{\sqrt{4 - 2e^2}}\right)_{\bar{a}, \bar{e}}^2 \sigma_e^2$$

$$= 2.2562$$

$$\text{Upper bound error} = 1.5021$$

Thus, the circumference of the ellipse and it's boundaries turns out to be

$$5.09 \pm 1.4 < C < 5.65 \pm 1.5$$

However, the boundaries overlap: because the mathematical bound is not the same as an error, it cannot be combined in the same way. Thus, we should report the boundary on the circumference with the minimum and maximum values (within the respective uncertainties). That turns out to be

$$5.09 - 1.4 < C < 5.65 + 1.5$$

## III

### 3.1

I assume that the arrival and departure times are normally distributed, with the standard deviation of the distributions being the reported uncertainty. While the question may be able to be solved analytically (perhaps by using convolutions?) I've chosen to simulate it.

- I sampled a million random numbers from a gaussian with mean 0 and standard deviation 120, to represent the truck arrival time (in minutes). I then did the same for a Gaussian with mean  $\Delta t = 130$  and a standard deviation of 50. I then pair them up and compare if the second random number is smaller than the first, that means the ship departed before the truck arrived, and thus the container needs to wait an additional day.

I ran the above simulation 25 times, and got an average fraction of **15.87%  $\pm$  0.04% containers** which needed to wait an additional day.

- I run the above simulation, but with a tweak: this time, I find all the days when the ship departs before the truck arrives, and calculate the waiting time by using the ship departure of the next day. Although it would ultimately not make a major difference to just add 1440 (24 hours, in minutes) to any negative time, it was fairly simple to implement the 'next day departure' by rolling the array, and it didn't increase the execution time. I first tried values of  $\Delta t$  from 0 to 1440 (Figure 3), and after noticing that there's only one minimum, reran my simulations for values in the range [180, 280].

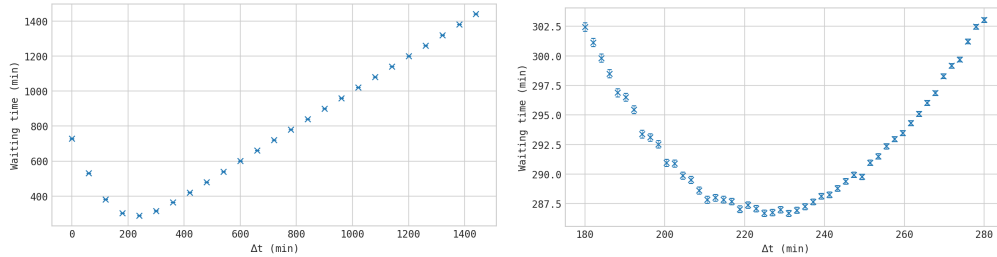


Figure 3: The waiting times, calculated from a numerical simulation. Error bars represent the error on the mean ( $\sigma/\sqrt{N}$ ). Left: domain = [0, 1440], Right: domain = [180, 280]

Close to a minimum, a function can be approximated as a parabola (which can be seen by a Taylor expansion). As such, I squeeze the interval once more, this time to [200, 250], and use chi-square to fit the following parabola to it:

$$f(x, a, b, c) = ax^2 + bx + c$$

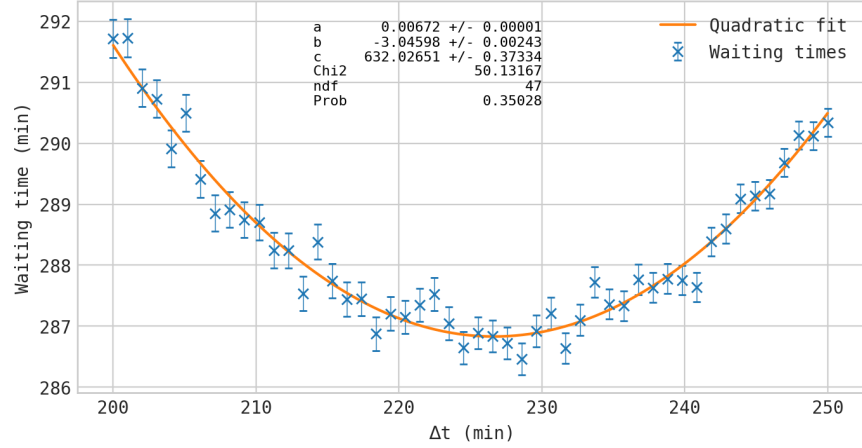


Figure 4: The waiting times calculated from a numerical simulation around the minimum, along with a quadratic chi-squared fit.

The quadratic fit in Figure 4 above has  $\chi^2 = 50.1$ ,  $N_{dof} = 47$  which gives us a great probability of 0.35. By fitting a parabolic equation in the standard form, we can differentiate it and set the derivative to 0 to calculate the  $x$ -coordinate of the minima. We get the formula

$$x_{\text{minima}} = -\frac{b}{2a}$$

Which tells us that the minimal waiting time is obtained at **226.64 ± 0.16 minutes**.

### 3.2

- I would determine random numbers by **inverse transform sampling**: The function can be integrated within the given domain, so we can find a CDF. While the CDF is not analytically invertible, we can use numerical methods to find a numerical inversion. If we have a function value  $y$ , we want to find  $x$  such that

$$\text{CDF}(x) = y$$

We can use a root finding algorithm such as the bisection method or Newton's method to find an  $x$  where

$$\text{CDF}(x) - y = 0$$

We then sample uniform random numbers, and by plugging them in to the (approximate) inverse CDF, we get random numbers from the required distribution.

The CDF can be calculated by integrating the PDF from  $\infty$  to  $x$ :

$$\begin{aligned} \text{CDF}(x) &= \int_0^x \frac{x'}{\sigma^2} e^{-x'^2/(2\sigma^2)} dx' \\ &= 1 - e^{-x^2/(2\sigma^2)} \end{aligned}$$

- The random numbers generated can be seen (according with the PDF) in Figure 5 below:

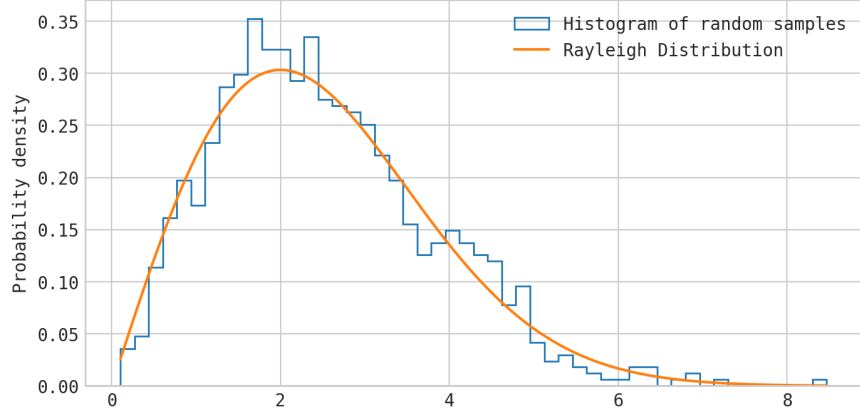


Figure 5: A plot of 1000 random points generated from a Rayleigh Distribution (with  $\sigma=2$ ) with inverse transform sampling, binned into 50 bins.

- I perform a chi-squared fit to the distribution. While it might be worse than an unbinned likelihood fit in cases where you have several empty bins, I feel that the benefit of the chi-square p-value being able to quantify the quality of the fit outweighs the drawback. I fit the function

$$f(x, \sigma, C) = C \cdot \frac{x}{\sigma^2} \cdot e^{-x'^2/(2\sigma^2)}$$

to obtain values of  $C$  and  $\sigma$ . I assume Poisson errors on the bins: the error is equal to the square root of values within the bin. The fit is presented in Figure 6:

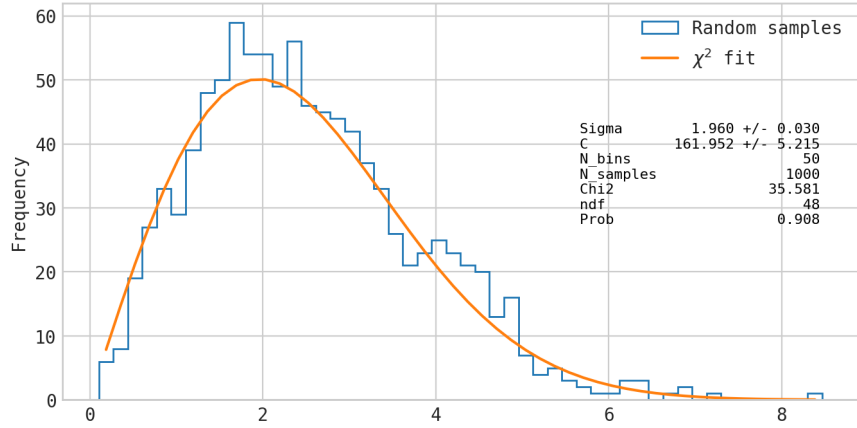


Figure 6: A plot of 1000 random points generated from a Rayleigh Distribution (with  $\sigma = 2$ ) with inverse transform sampling, binned into 50 bins, along with a  $\chi^2$  fit to estimate  $\sigma$

The fit looks very good by eye, and we get a decent probability of  $p=0.936$  ( $\chi^2 = 144$ ,  $N_{dof} = 48$ ).  $\sigma$  is determined with  $\sim 3\%$  precision.

- I chose 100 values of  $N$  from a log-spaced scale in the interval  $[50, 5000]$ , generated a histogram and did a chi-square fit to evaluate  $\sigma$ . However, the number of bins must be modulated according to the number of samples. I used the following equation to set the number of bins:

$$N_{bins} = \max\left(10, \left\lfloor \frac{N}{20} \right\rfloor\right)$$

$N/20$  lends itself fairly well for large values of  $N$ , but for a value like  $N=50$ , we get too few bins (in our case, just 2), and hence we set a minimum value of 10 bins. I repeat the process 10 times for every  $N$ , in order to get error bars on the uncertainty in  $\sigma$ .

I then perform a power law chi-square fit to the uncertainty in sigma data, using the function

$$f(N, a, \alpha) = aN^\alpha$$

We would expect a decent  $\chi^2$  value at  $\alpha \approx -0.5$ , as errors typically scale with  $1/\sqrt{N}$ .

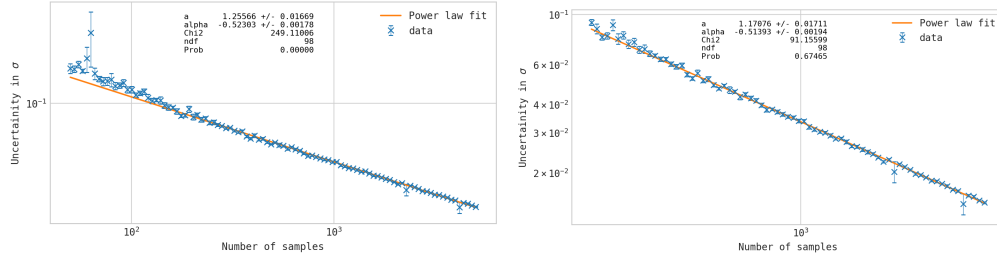


Figure 7: The uncertainty in  $\sigma$  from Rayleigh Distribution as a function of the number of samples, with a chi-square power law fit. Left: domain [50, 5000], Right: domain [159, 5000]

The left plot in Figure 7 shows the fit on the entire domain [50, 5000]. The initial points deviate from the trend, having a higher than usual uncertainty: I believe this is a combination of the small sample size and the number of bins used. Due to this fact, the chi-square value obtained is very bad ( $\chi^2=249$ ,  $N_{\text{dof}}=98$ ,  $p=0$ ). If we remove the first 25 elements and redo the fit (corresponding to the domain [159,5000]), we get a much better chi-square probability of 0.67 ( $\chi^2=91$ ,  $N_{\text{dof}}=98$ ). The fit also gives an  $\alpha$  of  $-0.5139 \pm 0.0019$ , similar to the -0.5 we would expect for  $1/\sqrt{N}$  scaling.

## IV

### 4.1

- In order to quantify which measure gives the highest separation between healthy and sick individuals, I used a **ROC curve**. Figure 8 below shows both the histograms (divided into healthy and sick people) and ROC curves for every test statistic:



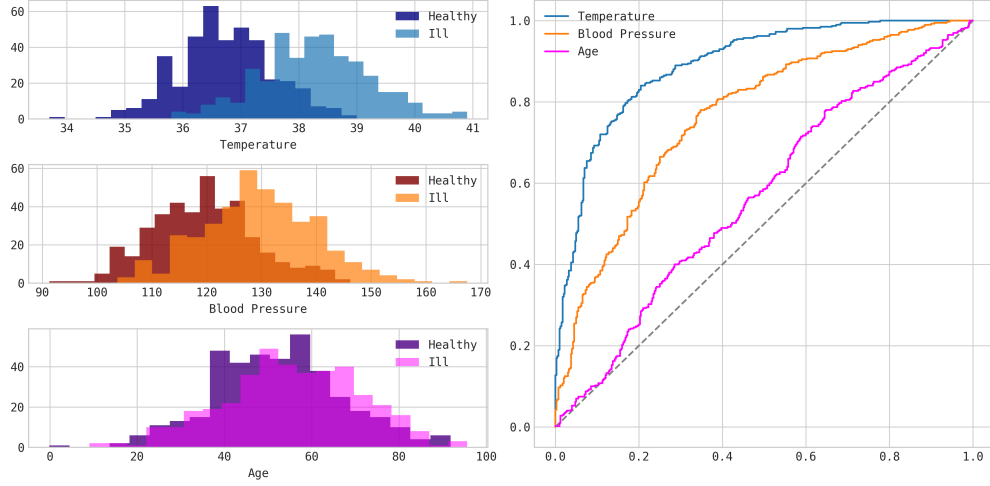


Figure 8: Left: Histograms of the characteristics of the control sample, segregated by healthy and ill people. Right: ROC curves of the different characteristics.

By visually inspecting the histograms in the figure, we can see that temperature is very good at separating the two populations. Blood pressure is also a decent check, and age has a minor difference. The behaviour is visible on the ROC curve: the further the curve is to the top left corner, the closer it is to an ideal classifier. We can see that age is only slightly better than a random classifier.

- In order to compare the distributions, I used the **Kolmogorov Smirnov test**. The test compares whether two given sets of samples originate from the same distribution, which is exactly what we need to do. If we set our significant threshold to 0.05, we find a **statistically significant difference** between the two distributions (KS statistic=0.1325,  $p=0.0018$ ).
- For simplicity, I do not consider any combination of the three characteristics, but only every linear combination of them. The optimal linear combination (which separates the two populations as best as it can) can be calculated with the **Fisher Discriminant**. After implementing it, we get the following weights:

$$\begin{aligned} w_{\text{temperature}} &= 0.93810394 \\ w_{\text{blood pressure}} &= 0.02806819 \\ w_{\text{age}} &= 0.01381178 \end{aligned}$$

Which aligns with what we see in Figure 8: temperature is such a great classifier that majority of the Fisher test characteristic is based on it.

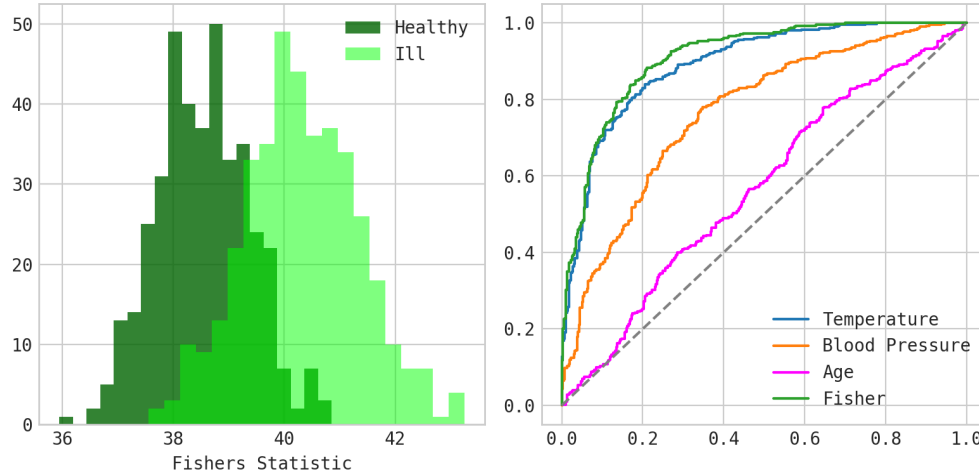


Figure 9: Left: Histograms of the control sample under the fisher test-statistic, segregated by healthy and ill people. Right: ROC curves of the different characteristics.

Figure 9 above shows the histograms of the populations under the newly created characteristic. As the ROC curve shows, it proves to be a better predictor than any of the previous characteristics.

When classifying data, there are several different metrics that we can use. There is typically a tradeoff between the false positive rate (i.e, type I errors) and the false negative rate (type II errors). Where to draw the line is heavily scenario dependent: for a hospital with vast resources, it's probably beneficial to reduce the false negative rate as much as possible, so that everyone with the disease is caught (although there could be several false positives, repeated testing could potentially help identify them).

Unfortunately, I'm a computational physicist and not an epidemiologist. As such, I've decided to arbitrarily choose a cutoff threshold which seems reasonable: simply reduce the total number of people who are wrongly classified (i.e, the sum of false positives and false negatives). This approach gives us a value of 39.515 as the cutoff fisher statistic, which intuitively lines up with what we'd expect (given the histograms in Figure 9). Using this threshold, I find out that **54 individuals** in the unknown group are ill.

- Applying Bayes' Theorem (note: I have only used the control sample for calculations here),

$$\begin{aligned}
 P(\text{Ill} | T = 38.6^\circ) &= \frac{P(T = 38.6^\circ | \text{Ill}) \cdot P(\text{Ill})}{P(T = 38.6^\circ)} \\
 &= \frac{(14/400) \cdot (0.01)}{15/800} \\
 &= \frac{7}{375} \\
 &= \mathbf{1.867\%}
 \end{aligned}$$

## 4.2

- 

Mean = 17362.045

Median = 6677.0

25% quantile = 2047.0

75% quantile = 23762.0

- In order to test whether  $\log_{10}(\text{PopSize})$  is Gaussian, I used the **Shapiro-Wilk test**. While there are alternate tests one could use (the Kolmogorov-Smirnov test, the Anderson-Darling test or the Lilliefors test), Shapiro-Wilk has the most predictive power (though Anderson-Darling isn't far behind). With a significance level of  $p = 0.05$ , the test showed **no significant results** (Test statistic=0.987,  $p=0.25$ ) and thus we cannot reject the hypothesis that the value follows a Gaussian distribution. As a cross-check, the Anderson-Darling test obtained the same result (statistic=0.52, which wouldn't even be rejected with a 15% confidence level).
- Pearson's correlation coefficient was fairly large ( $\rho_{\text{pearson}}=0.765$ ,  $p \approx 10^{-26}$ ), which shows a **strong correlation** between the happiness and education indexes. Spearman's correlation coefficient, which tests the rank correlation (and thus can find monotonically correlated data instead of just linearly correlated data) was slightly higher ( $\rho_{\text{spearman}}=0.804$ ,  $p \approx 10^{-31}$ ).
- A preliminary plot of the happiness index as a function of GDP shows that most of the data points are clustered together at low GDPs, while a couple of outliers have a large GDP, and thus distort the scale of the graph (which can also be seen from the means and medians calculated in the previous part of the question). As such, I decided to plot it on a log-log scale, and a roughly linear trend is visible (Figure 10).

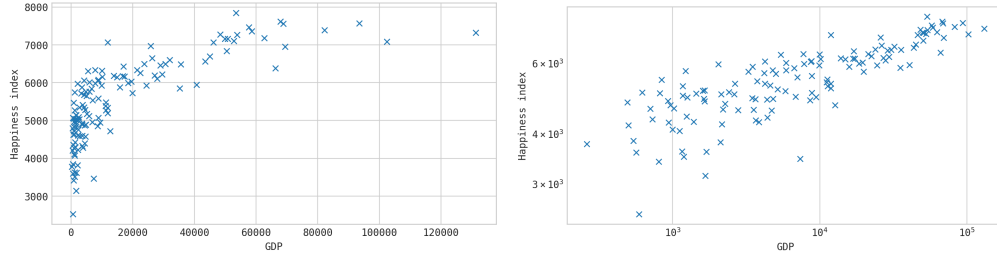


Figure 10: The happiness index as a function of GDP. Left: linear scale, Right: log-log scale.

If a trend appears linear on a log-log plot, that means it's a *power law*. Thus, I fitted the following power law to the data:

$$f(x, \lambda, a, c) = Ax^\lambda + C$$

As I wanted to estimate the error on the happiness index, I made the assumptions that the error on a value will be proportional to the value itself. As such, I only need to find the scale  $k$  such that

$$\sigma_{\text{Happiness}} = k \cdot (\text{Happiness Index})$$

gives realistic errors. For this reason, I chose to use a chi-squared fit, as the chi-square probability tells you whether your errors are underestimated ( $p \sim 0$ ) or overestimated ( $p \sim 1$ ). Figure 11 shows an example of one such fit for  $k = 0.13$ :

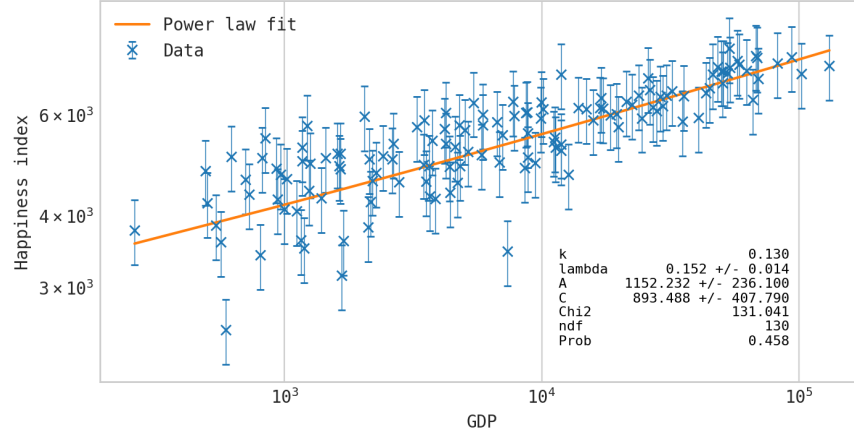


Figure 11: The happiness index as a function of GDP, fitted with a  $\chi^2$  fit. The errors are assumed to be relative with a scale factor  $k = 0.13$

We get a nice chi-square fit, with  $\chi^2 \approx 131$  and  $N_{dof} \approx 130$ , which gives us a very acceptable probability of 0.458. By running multiple such fits for different values of  $k$ , we can calculate a  $\chi^2$  value for each, and (in particular) check where the fit becomes “acceptable” with a significance level of  $1\sigma$  ( $0.3173 \leq p \leq 0.6827$ ). The range appears to be  $0.127 \leq k \leq 0.135$ , with  $p = 0.5$  being obtained at  $k \sim 0.131$ . Thus, we can quantify the error on the happiness index to be:

$$\sigma_{\text{Happiness}} = (0.131 \pm 0.004) \cdot (\text{Happiness Index})$$

## V

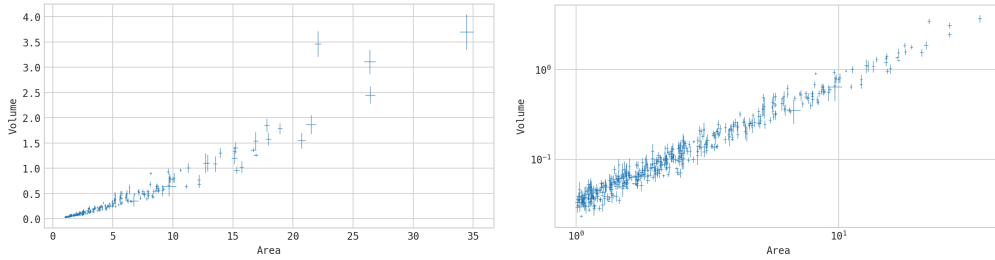


Figure 12: A plot of glacier volume as a function of area. Left: Linear scale, Right: log-log scale

- Figure 12 above shows the data along with the uncertainties on both area and volume. From the plot, we can see that the relative uncertainty on the volume appears to be bigger than that of the area. In order to quantify it, I calculated the relative errors to be

$$E(\sigma_A/A) = 0.01328 \pm 0.00019$$

$$E(\sigma_V/V) = 0.1007 \pm 0.0024$$

It's clear to see that the relative error in volume is an order of magnitude larger than the error in the area.

- I performed a chi-square fit with the following function

$$V(A, k) = k A^{3/2}$$

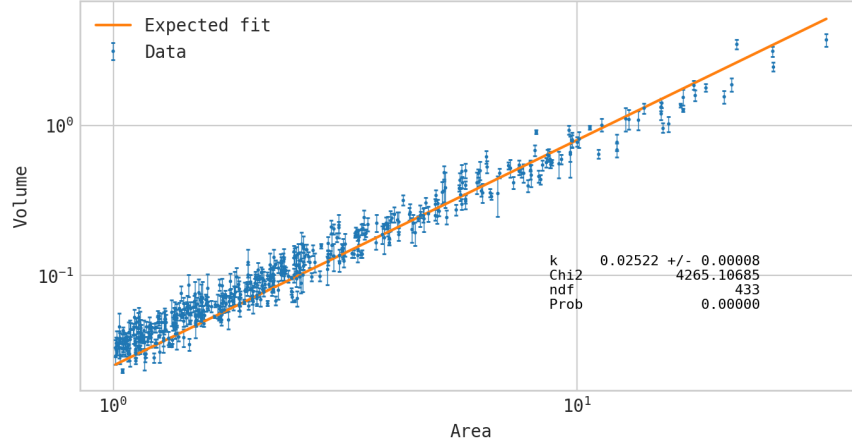


Figure 13: Glacier volume as a function of area, along with a chi-square fit of the form  $V = k A^{2/3}$

- The fit in Figure 13 above passes the eye test: the line appears to go through majority of the points, and fit the data pretty well. However, we get an abnormally low  $\chi^2$  probability of 0 ( $\chi^2=4265$ ,  $N_{dof}=433$ ), and hence the fit isn't satisfactory. A Wald Wolfowitz Runs test yielded no significant results. However, the residuals tell another story - a histogram is plotted below (Figure 14).

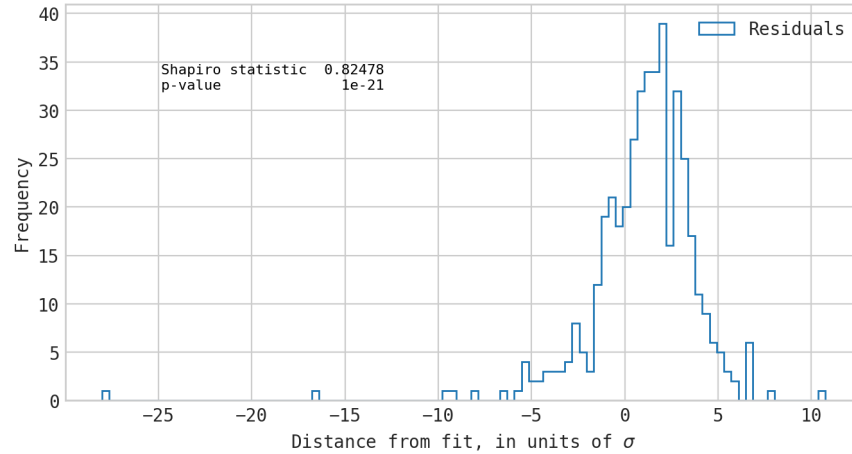


Figure 14: Residuals of the fit performed in Figure 13

The chi-squared test assumes that the residuals are normally distributed, however they did not appear to be Gaussian. This was tested with a Shapiro-Wilk test, which resulted in a statistic of 0.825, with a  $p$ -value of  $10^{-21}$ . The main problem lies in a couple of outlying points, whose errors are small and thus their residuals are large.

Changing the functional form will not be able to improve the  $\chi^2$  value a whole lot: these outlier points will only be able to be “fitted” if we use several degrees of freedom and thus overfit the system. Alternatively, one could choose to discard them (say, by Chauvenet's Criterion) but I'm personally against the idea of discarding data for not matching your a-priori assumptions, and thus I choose to keep it in.

- I now change the function to have a few more free parameters: I let the exponent (which was previously  $3/2$ ) be free, and add in a  $y$ -intercept. The function to be fitted becomes:

$$V(A, k, \alpha, c) = kA^\alpha + c$$

A fit with the modified function is shown below (Figure 15).

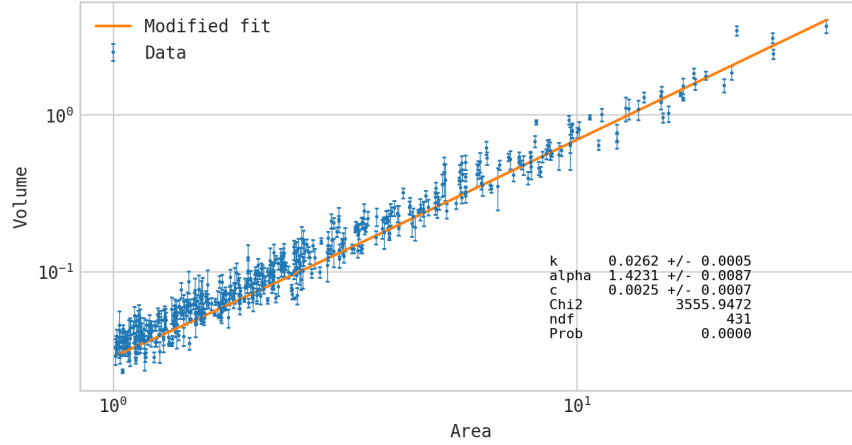


Figure 15: Glacier volume as a function of area, along with a chi-square fit of the form  $V = k A^{2/3}$

The chi-squared probability doesn't increase, for the same reason as before (Shapiro-Wilk  $p \approx 10^{-23}$ ). We can clearly see that the  $\chi^2$  value has decreased significantly, from 4265 to 3555, being scaled by a factor of 0.834. That said, the value of  $N_{dof}$  has gone down, and we can consider the ratio of  $\chi^2/N_{dof}$  for both the plots, which shows us that we get nearly a 20% improvement.

- The earlier fits were both chi-square fits, assuming no area uncertainty. I was not able to find a generalization of the chi-square quantity for when you have errors on both the dependent and independent variables, and thus I couldn't use a chi-square fit.

Instead, I choose to implement a non-linear total least squares algorithm, in particular, *the Levenberg-Marquardt algorithm* to minimize the total residuals. When the errors on  $x$  and  $y$  are identical, it can be thought of as minimizing the orthogonal distance of points to the curve, and thus is also called Orthogonal Distance Regression (ODR). The plot can be seen in Figure 16.

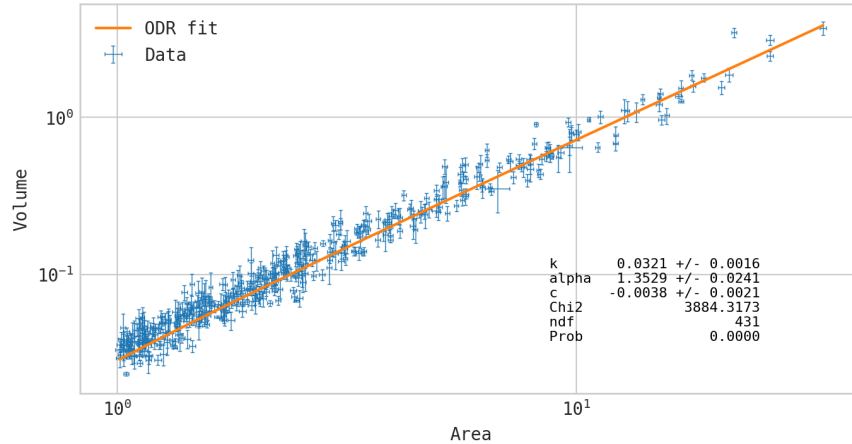


Figure 16: Glacier volume as a function of area, along with a total least-squares fit of the form  $V = k A^{2/3}$

While  $\chi^2$  is not used here, I choose to calculate it and display the value (3884) to illustrate how it increases from the previous plot: this is because the previous plot minimized it, whereas now it's minimizing something more general in 2 dimensions. Another important observation is that the errors on the parameters  $k$ ,  $\alpha$ , and  $c$  are all approximately tripled: previously they had only a  $y$ -error, but

by adding in  $x$  errors not only do we increase the total overall error, but it typically also increases the size of the trust region (i.e, there are many more solutions that are “almost as good” as the optimum).

•

$$\begin{aligned}\sigma_V^2 &= (A^\alpha)_{\bar{A},\bar{\alpha}}^2 \sigma_k^2 + (k \ln(A) A^\alpha)_{\bar{A},\bar{k},\bar{\alpha}}^2 \sigma_\alpha^2 + \sigma_c^2 \\ &= 0.00000484646 \\ \sigma_V &= 0.0022\end{aligned}$$

The volume will thus be **0.0088  $\pm$  0.0022 km<sup>3</sup>**.