1.1 a) In NFIP, the control group is simply not given anything whereas in the double-blind the control group is given a placebo. The nonblindness of the first experiment can lead to some biases. In addition, the amount of people in each respective group is quite different. (i.e. 725k vs 200k and 125k vs. 350k).
Also, the NFIP study split patients up by grade which did not happen with the double blind experiment.

b) NFIP: 25 vs. 54 and 25 vs. 44
Double blind: 28 vs. 71 and 28 vs. 46

c) Each of these groups had their own prior. The patients in the control group might've known they could be at risk, so they were more eager to join the study. Similarly, the non-consent group could have some idea that they are most likely ok, so they did not consent to avoid the hassle of the experiment. This can lead to a discrepancy in results.

d) Yes, when both the patient and the doctor performing the experiment know if the patient is in the treatment or control group this can lead to biases from the doctor preferring a certain result or the participant behaving a certain way they believe they are supposed to behave.

e) The parents are incorrect because the consenting group also contains kids who received the placebo and not the actual treatment. This greatly inflates the polio rate, but should be a cause for concern for parents.

1.2 a) we are looking at the # of events in a time interval, so each interval can be represented as a poisson distribution.

b) Say each interval has t_i has rate λ_i.
Define Θ as the space of 100 dim vectors representing the Rates for every interval. Define $\Theta_0 \subset \Theta$ as the subset where all 100 entries are the same i.e., the rate is constant.
If Λ is the 100 dim vector of all our rates, we get:

$$\boxed{H_0: \Lambda \in \Theta_0, \; H_A: \Lambda \in \Theta \setminus \Theta_0}$$

c) $\mathcal{L}_0(D) = \prod_{i=1}^{100} \frac{(\lambda t_i)^{x_i} e^{-\lambda t_i}}{x_i!}$ Note in this case we have $t_i = \lambda_j \; \forall (i,j)$, so let $\lambda = \lambda_i \; \forall i$

$$\ell_0(D) = \sum_{i=1}^{100} x_i \ln(\lambda t_i) - \lambda t_i - \ln(x_i!)$$

$$\frac{\partial \ell_0(D)}{\partial \lambda} = \sum_{i=1}^{100} \frac{x_i}{\lambda} - t_i = 0$$

$$\Rightarrow \boxed{MLE: \lambda = \frac{\sum x_i}{\sum t_i} \approx 0.00388}$$

$\Lambda(t_i)$

d) $\mathcal{L}(D) = \prod_{i=1}^{100} \frac{(\lambda_i t_i)^{x_i} e^{-\lambda_i t_i}}{x_i!}$

$$\ell(D) = \sum_{i=1}^{100} x_i \ln(\lambda_i t_i) - \lambda_i t_i - \ln(x_i!)$$

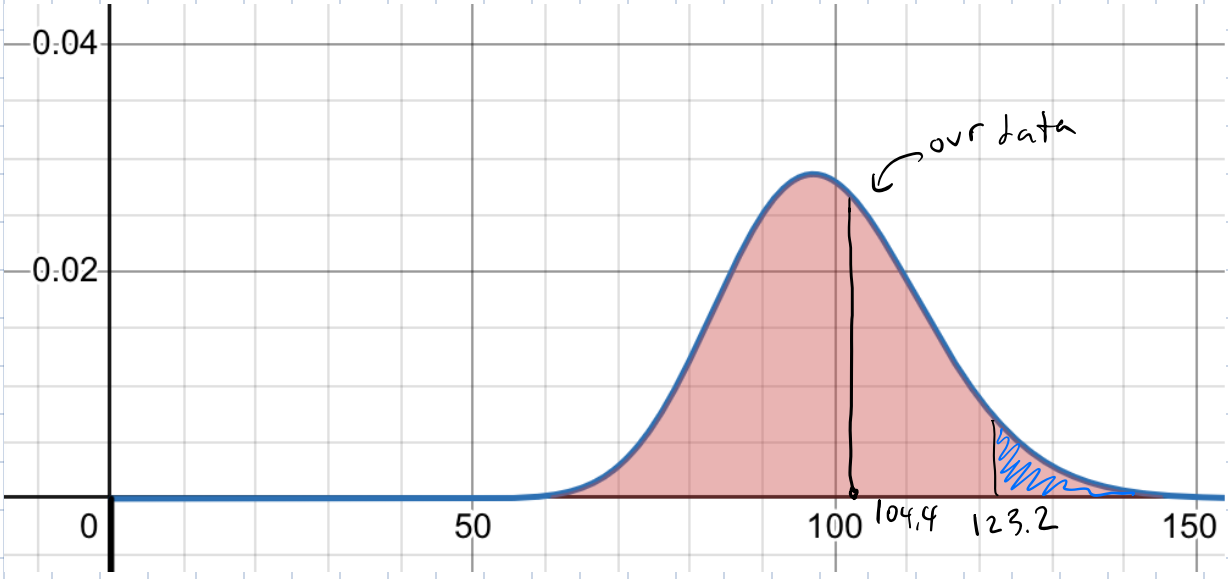
$$\frac{\partial \ell(D)}{\partial \lambda_i} = \frac{x_i}{\lambda_i} - t_i = 0$$

$$\Rightarrow \boxed{MLEs: \lambda_i = \frac{x_i}{t_i}}$$

e) Our test statistic is our likelihood ratio statistic.

$\Lambda(D) = -2\ln(L(D)) = -2(\ell_0(D) - \ell(D))$ where $\ell_0(D)$ and $\ell(D)$ are defined as in parts c) and d) and we take the MLE parameters that we solved for. By Wilkes Thm, under H_0, $\Lambda(x) \to \chi^2_d$ where $d = 100 - 1 = 99$.

$d = 99$ χ^2 distribution:



: rejection region

f) we reject H_0 if $\Lambda(x)$ is too big.

Using an online calculator, we see that $\mathbb{P}(X \geq 123.2) = 0.05$ if X is drawn from 99 df χ^2 distribution. Therefore, our rejection region is just if our likelihood ratio statistic is greater than 123.2.

I've shaded this region on the plot in the previous part.

g) We get that $\ell_0(D) = 5.25 \times 10^{-40}$

and that $\ell(D) = 2.46 \times 10^{-17}$

$\Rightarrow \Lambda(D) = 104.4$

This gives us a p-value of 0.336.

This is not statistically significant signaling that we fail to reject the null hypothesis.

1.3a) No, it means that under the assumption our null hypothesis is true, death rate is same in control and treatment group, the probability we see the observed data is 0.0012.

The assumption used was that the deaths of patients followed a binomial distribution where each patient independently has some probability of dying.

b) Probably not. This seems like a case of correlation not being the same as causation. More likely, there could be a confounding variable such as how developed the economy of a country is which can yield more chocolate consumption and more education ⇒ Nobel prizes.

c) No, thinking about the definition of the p-value, it says the observed data has a 5% chance of happening under the null. Therefore, with 100 features, most likely one can find some pair which is "statistically significant" purely by chance.

d) No, what the lab has done is accept the alternate hypothesis, but in reality they should be rejecting the null hypothesis. This is because the data shows that the change in performance is unlikely to be from random chance.

e) For: interpretable measure of an experiment that can give some idea as to if some factors are influential or not.

Against: Easily "hackable" as we saw in part c. While the real meaning is interpretable, it is also easy believe alternative incorrect meanings as seen in part a.