

This is not the level of work I was hoping for on the final project. You have not researched this topic very thoroughly, and what you wrote contains some pretty bad misunderstandings. The paper itself is quite a bit under the required length. You have also not followed any standard method for citing your sources and documenting where the ideas came from.

First, to say how the  $\chi^2$  goodness of fit test works in general, the idea of “goodness of fit” needs to be formulated in terms of a specific *probability model* for how the data would be generated. So the general description of goodness of fit as considering “how well the data fits what it should be” is vague and somewhat inaccurate. But you do something like the correct approach in describing how the Mendelian inheritance table would generate the 9 : 3 : 3 : 1 ratios for the frequencies of the round/wrinkled and yellow/green pea traits depending on the traits of the parent plants. So this is not the most serious problem. (But note that the model is really the hypothesis here and the question is whether the model fits that data.)

Second, the example computation is good up until the final interpretation of what the test is telling you. (By the way, did you generate this computation yourself, or did you take it from one of your sources? There’s no penalty for using results like this as long as you say where they came from. But this should be made clear to the reader!) The numbers generated here are supposed to be uniformly distributed on the interval 0 - 100, so the bin sizes correspond to the 9 : 3 : 3 : 1 ratios correctly. So this computation is simulating the kind of data that Mendel *should have reported* (see below). The computation of the  $\chi^2$  statistic is correct. But I think you misunderstood how you should interpret the results.

The null hypothesis for this goodness of fit test is that the 9 : 3 : 3 : 1 ratios *do describe the data* reasonably well. The alternative is that they *do not*. This is actually a bit different from the tests we discussed in class and you needed to appreciate that to understand what was going on here. The larger the  $\chi^2$  statistic is, the more evidence you have to say that the 9 : 3 : 3 : 1 ratios do not fit. But with 3 d.f., the value  $\chi^2 = 3.509$  would give a *p*-value between 0.5 and 0.1. (Your table gives the probability of observing a certain value of  $\chi^2$ , or something larger, if the null hypothesis is true – that is, if the 9 : 3 : 3 : 1 model does describe how the data was generated.) In other words, there *is not enough evidence to say that the 9 : 3 : 3 : 1 model does not fit the data* (and of course there shouldn’t be, because that is how the data *was* generated(!)) We do not reject the null hypothesis in this case.

Third, and more seriously (given the length of what you did submit), you have not included any of the more detailed descriptions of where the  $\chi^2$  distributions come from, how the  $\chi^2$  tables would be generated, or how R.A. Fisher used the  $\chi^2$  distributions to analyze Mendel’s reported data. Recall that the project topic description asked for all of that. That would have been interesting, because the thing that Fisher found was that Mendel’s data was, in a sense, “too good to be true.” The agreement with the predicted

9 : 3 : 3 : 1 ratios was *so close* that  $\chi^2$  came out very small. If you look at the  $\chi^2$  table, you can see that equivalently the  $p$ -value for the test was extremely large, so the null hypothesis looked like a sure bet. The conclusion that Fisher drew (and that has been debated ever since) was that Mendel had actually “*fudged*” or “*cherry-picked*” his data to fit the pattern he thought should be true. Of course we now know that Mendel was right. But if he did that, then he was at the very least not following what we think of now as sound scientific methods(!)

*Paper:* 15/20

*Presentation:* 0/5

*Bibliography:* 5/5

*Total:* 20/30 (67%)