Unit 4 HW Solution

**Note that you have already completed a test using the data below… the goal in this homework will be to conduct a power analysis of this test and suggest modifications to this test using additional power analysis.**

1. In the United States, it is illegal to discriminate against people based on various attributes. One example is age. An active lawsuit, filed August 30, 2011, in the Los Angeles District Office is a case against the American Samoa Government for systematic age discrimination by preferentially firing older workers. Though the data and details are currently sealed, suppose that a random sample of the ages of fired and not fired people in the American Samoa Government are listed below:

**Fired**

34 37 37 38 41 42 43 44 44 45 45 45 46 48 49 53 53 54 54 55 56

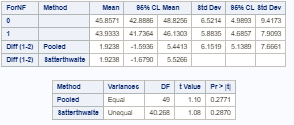
**Not fired**

27 33 36 37 38 38 39 42 42 43 43 44 44 44 45 45 45 45 46 46 47 47 48 48 49 49 51 51 52 54

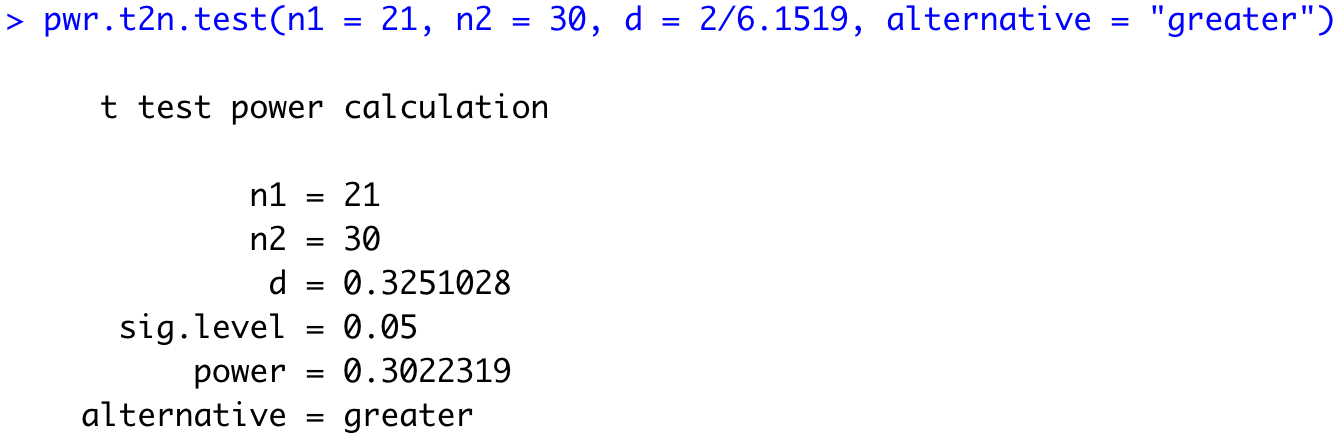
1. Assume the assumptions for the pooled t-test are met. The difference of sample means  ) was found to be 1.92 years although the result was not statistically significant. If there really is a difference in mean years of age between these two groups, what type of error was made here?

**Type II Error**

1. Again, assume the assumptions for the pooled t-test are met and remember the difference of sample means () was found to be 1.92 years although the result was not statistically significant. Let’s say the researchers conclude that they still believe that the Fired group has a higher mean age and they believe that it failed to reject based on the randomness of the observations in their sample. If they used a one-sided test with the data above, what is the power of their test to detect a difference of 2 points? For the standard deviation, since we are assuming the standard deviation of each group to be equal, we will use the pooled standard deviation (you will need to find this.).

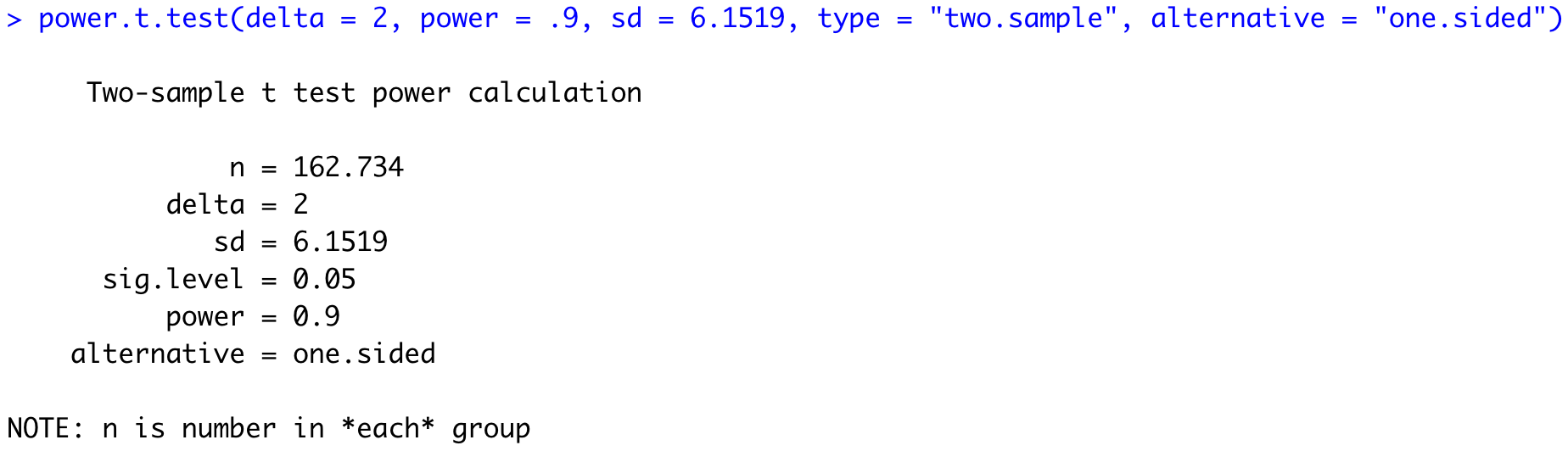
The pooled SD is 6.1519.

The power of the test to detect a 2-point actual difference in means is .3022 as seen below:



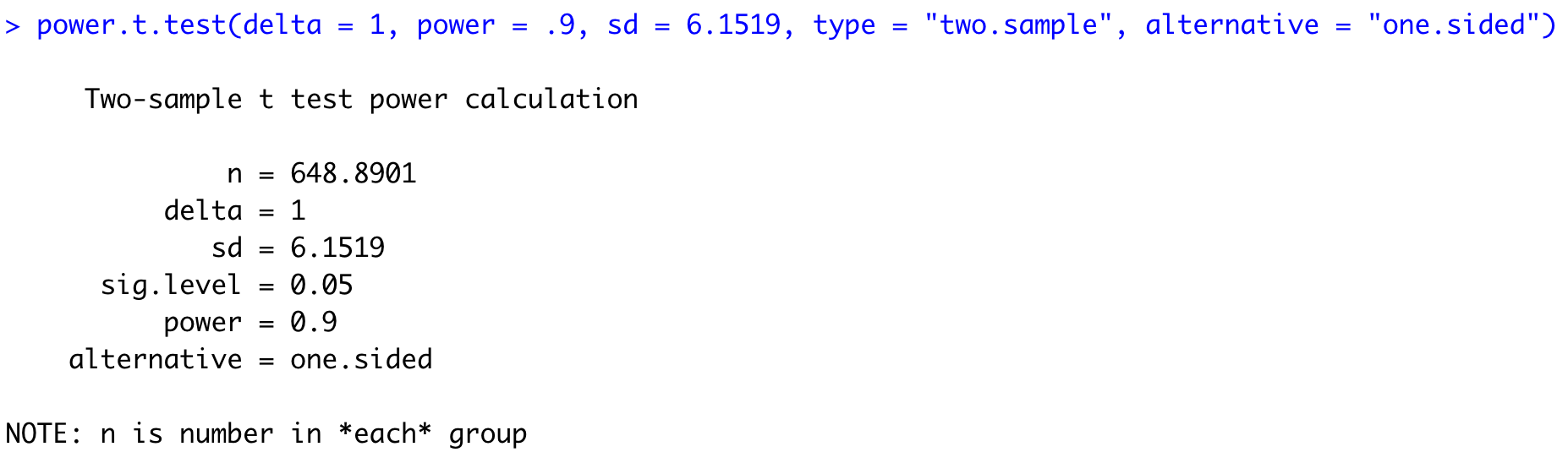
1. Next, they have found the money to conduct a new study and would like to calculate the sample size required to have 90% power to detect if the mean of the Fired group is 2 points higher than the Non-Fired group. (using a one-sided test.) Calculate this sample size.

Since we are not given any information about the if the groups are to have equal sample sizes, we will have to assume that they will.



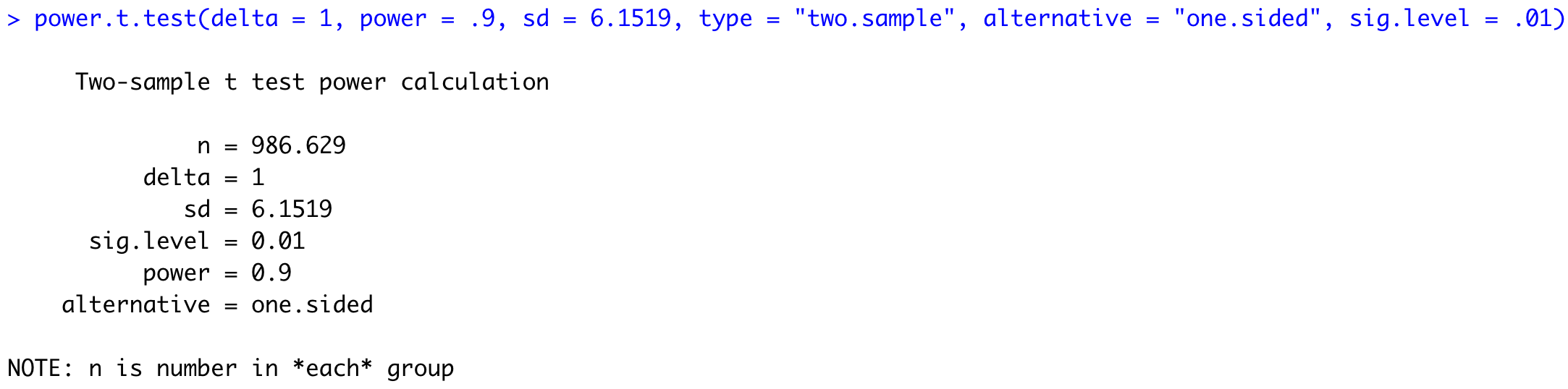
Therefore, the study would need to have 163 members per group in order to achieve at least 90% power.

1. Next assume that a difference of 1 year is the legal threshold to define “age discrimination”. Next, calculate the sample size needed per group to again achieve 90% power to detect this minimum legal threshold.



This study would need to be much larger if the effect size was reduced to 1. The study would need to have 649 members per group to achieve at least 90% power.

1. While it is important to be able to detect this minimum difference if it actually exists, the researchers also do not want to detect a significant result if there really is not a difference in the mean ages. Assume we would like to reduce the probably of making this type of error to 1%. How does that affect the sample size calculation in the last problem?



This study would need to be much larger still if the significance level was reduced to .01. The study would need to have 987 members per group to achieve at least 90% power.

1. Someone in the research group proposed to invest in a sample size that hits the point of diminishing returns with respect to power. Create a power curve to estimate the sample size, per group (assuming equal sample sizes per group), at which increasing the sample size would not provide “significantly” greater power to justify the increase in cost for detecting this minimum legal threshold. This answer will vary per student! Please provide your power curve as well as your justification for your choice of sample size.



I used alpha = .01 here although I could see using alpha = .05 as well. It looks to me that the slope of the power curve is steepest at about 800 members per group which would yield about 82% power. The study still receives a significant boost in power by adding samples after 800, but the gait seems to dissipate at this value. Again, answers will vary.

*#Sample Size*

*powerholder = c()*

*samplesizes = seq(100,1500,length = 20)*

*for(i in 1:20)*

*{*

*powerholder[i] = power.t.test(n = samplesizes[i], delta = 1, sd = 6.1519, sig.level = .01, type = "two.sample",alternative = "one.sided")$power*

*}*

*plot(samplesizes,powerholder,type = "l", col = "blue", main = "POWER CURVE", ylab = "POWER", lwd = 3)*

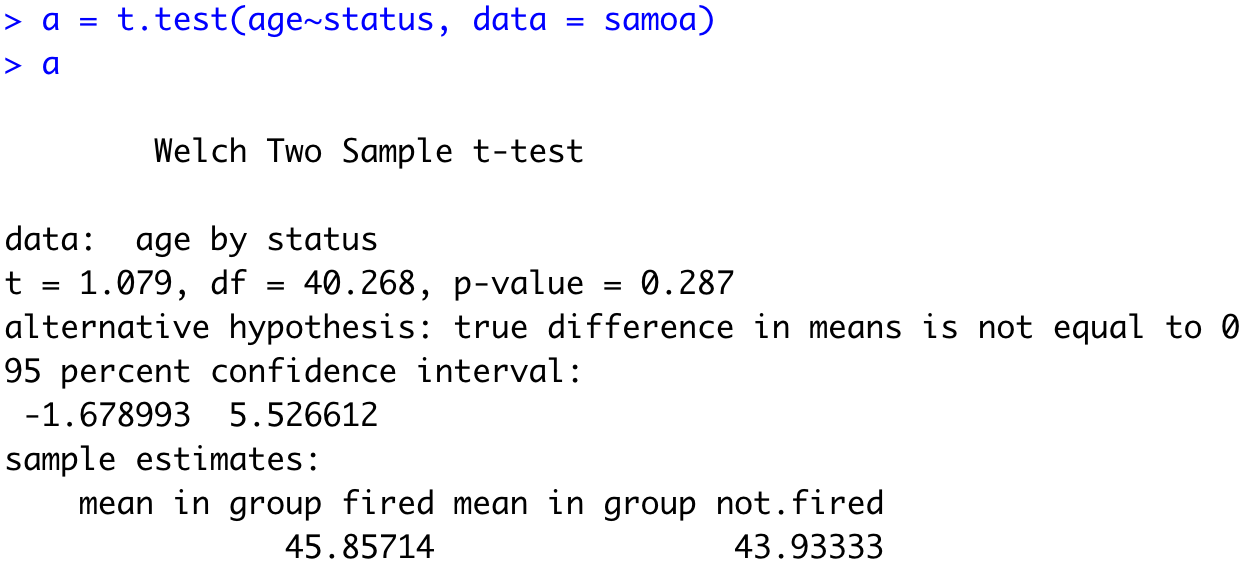
*abline(h = .820, col = "red", lwd = 3)*

*abline(v = 800, col = "red", lwd = 3)*

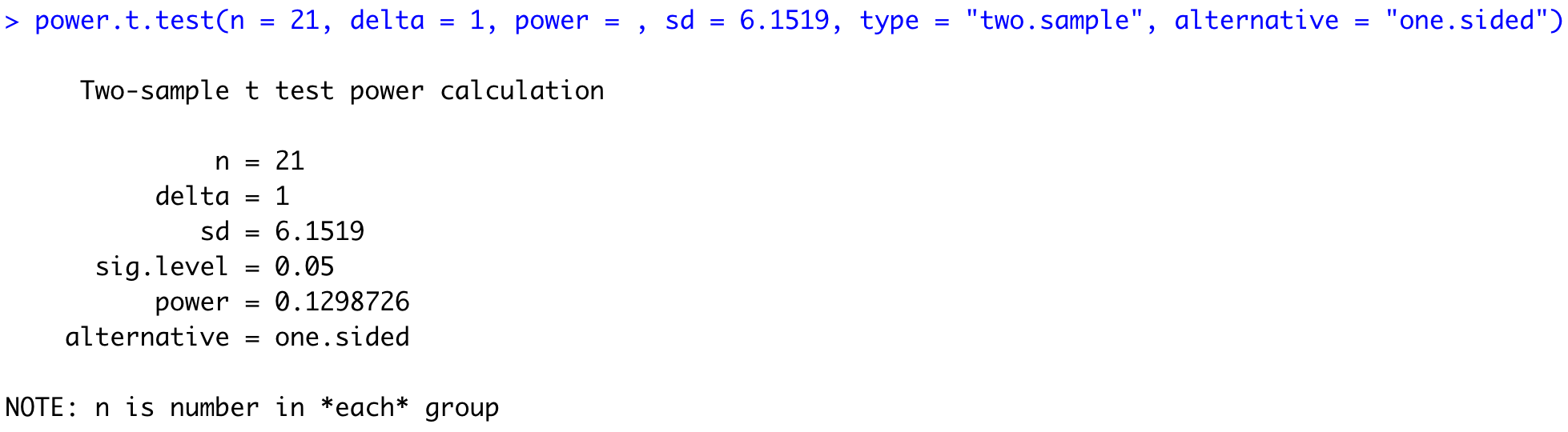
1. Finally, let’s say that we are not comfortable saying that the standard deviation assumption is met. What will happen to the power if we use Welch’s test with the Satterthwaite adjustment to the degrees of freedom?

The power will decrease since the degrees of freedom of the test, a proxy for the sample size, decreases.

An estimate of the decrease in power can be found by calculating the Sattherwaite adjustment to the degrees of freedom and then finding the power for a test with the corresponding sample size.



The degrees of freedom for the Welch’s test are about 40 which would be close to the degrees of freedom for the pooled test which were 49. This is a rough estimate as I will assume equal sample sizes in each group to calculate the power.



Also assuming that the effect size was 1, the power of this test is .1299 or 12.99%.

Solution using function that will use different standard deviation, but it still forces us to use the same sample size…. :/

Note that we figured the power if both sample sizes were 21 and if both sample sizes were 30. The power varied between 7.3% and 8.9%.

Graphical user interface, text, application, email

Description automatically generated