From: Hayden Smith
To: Professor Neal E. Devins and Professor Allison Orr Larsen
Date: December 17 2023
Subject: DDR statistical significance

Statistical tests were run in Microsoft Excel® version 16.81. Under this test, the alpha value is 0.05 and the p-value (for a result to be found statistically significant) is less than or equal to 0.05. Here, the p-value is 0.015 for the one-tailed test and 0.030 for the two-tailed test.

1. Formulated a research hypothesis

- Trump-appointed judges author more DDRs than Other Republican judges.
- Considering the alternative hypothesis:
    - In order to calculate the value of t, I formulated this hypothesis to determine whether the differences between means (the volume of DDRs) resulted from random error or real effect. This hypothesis is designed to show that the relationship is not the result of chance. Because this research hypothesis states a clear direction—that Trump judges author more DDRs—we considered whether to use a "one-tail" or "two-tail test" to find the value of t of the distribution.

2. Formulated a null hypothesis

- Trump-appointed judges do not author more DDRs than Other Republican judges.
    - The null hypothesis provides that no relationship exists between two variables to exclude the possibility of chance. Here, the null hypothesis states that my research hypothesis that Trump-appointed judges author more DDRs than Other Republican judges is not true.

3. Determined the Significance Level

- I selected the probability of error (alpha level) was 0.05. In other words, I specified a 5% chance of making a Type I error that Trump judges do not author more DDRs than Other Republicans. This is a standard value.

4. Selected the test

- I selected a Paired t-Test to determine whether the difference between the means of the volume of DDRs by Trump judges and Other Republicans is statistically significant.
- The t-Test measures statistical significance between two groups using the same variable based on the mean for each group. Here, I pose the question: Do Trump judges author DDRs at a higher rate than Other Republican judges?
    - I considered using the Two-Sample t-Test Assuming Equal Variances; however, under that independent test, the DDR data samples would need to be pulled from the same group of people or judges. Under the Paired t-Test, I can assess two groups of observations (two separate columns measuring the volume of DDRs by Trump and non-Trump judges) since the dependent samples are grouped by year and not randomly. I pulled the volume of DDRs by Republican judges (either

Trump Judges or non-Trump judges) from 2019-2023.[1] Once Trump judges began authoring DDRs in 2018, and as soon as they matched the volume of Other Republican judges in 2019, they began authoring more DDRs every year.

- Typically a Paired t-Test is used for the same subject, but it applies here since each year is a specific timeframe while the DDR authors are unique.
- In my previous research, I tested the same stimulus using a Two-Sample t-Test Assuming Equal Variances and found that the one-tailed t-Test yielded a p-value of **0.029**, which would establish statistical significance by rejecting our null hypothesis ($p < 0.05$).
- Since a statistical significance measurement is more accurate when based on larger sample sizes, and paired a t-Test measures two variables from the same group, it would make sense to conduct a separate test for each Trump Judge and every Other Republican judge. However, the time frame segmented by year here serves the purpose of creating the same "subject" by which I measured the volume of DDRs.

- The table in Figure 1 is a useful visual. Data for the Paired t-Test was pulled from the columns highlighted in grey.

Figure 1:

Republican Judges Authoring DDRs (2017-2023)[2]							
	Appointing President						
YEAR	**Reagan**	**HW Bush**	**GW Bush**	**Trump**	**Total Non-Trump Republicans**	**Combined Total**	**% of DDRs by Trump Judges**
2017	12	3	14	0	29	29	0%
2018	4	1	8	5	13	18	28%
2019	4	2	11	17	17	34	50%
2020	5	1	8	28	14	42	67%
2021	4	2	8	21	14	35	60%
2022	3	0	8	23	11	34	68%
2023	0	2	2	11	4	15	73%
TOTAL	32	11	59	105	102	207	51%

4. Considered a power analysis for the sample size

- n = 5 (2019 to 2023).

---

[1] This data was provided by Paul Hellyer. The database search inquiry applied the following parameters: *DISSEN! /10 (DEN! REFUS! DECLIN! FAIL! /S ("EN BANC" "IN BANC")) & DA ([4-digit year])* in Westlaw's U.S. Courts of Appeals database. Results were then reviewed to exclude false positives and Federal Circuit cases. *See* US Court of Appeals, *Additional Authorized Judgeships*, https://www.uscourts.gov/sites/default/files/appealsauth.pdf.). See also Jeremy Horowitz, *Not Taking "No" for an Answer: An Empirical Assessment of Dissents from Denial of Rehearing En Banc*, 102 Geo. L. J. 59, 59 (2013).

[2] The timeframe for DDR data is January 19, 2017 to July 30, 2023. Republican-appointed judges authored 207 of the 303 total DDRs during that period. Data taken from the Federal Judicial Center's *Biographical Directory of Article III Federal Judges, 1789-present*, available at https://www.fjc.gov/history/judges (downloaded December 10, 2023).

- A large sample size may yield smaller results because the sample size is actually too large; however, we run the opposite risk in this test.
- I considered possible performance errors in my calculation. First, this sample does not pull from a population (like a typical paired t-Test). Second, the sample size is constrained to a relatively small time period when compared to a larger sample size.

5. Conduct the t-Test

- By using Excel's Data Analysis ToolPak, I ran the t-Test by inputting the variables from Figure 1 above. Variable 1 column contained the volume of DDRs by Trump judges segmented by year. Variable 2 column contained the volume of DDRs by Other Republicans segmented by year. I entered the hypothesized mean difference as 0. This is the value of the null hypothesis (representing no effect).
- Excel produced the following output in Figure 2.

Figure 2:

t-Test: Paired Two Sample for Means

	Variable 1	*Variable 2*
Mean	20	12
Variance	41	24.5
Observations	5	5
Pearson Correlation	0.56793356	
Hypothesized Mean Difference	0	
df	4	
t Stat	3.29354788	
P(T<=t) one-tail	**0.01505753**	
t Critical one-tail	2.13184679	
P(T<=t) two-tail	**0.03011507**	
t Critical two-tail	2.77644511	

6. Interpreted the results

- The t-Test yielded a one-tailed p-value of 0.015 and a two-tailed p-value of 0.030. Here, the p-value determines whether the research hypothesis test yielded a statistically significant outcome.
  - Since my hypothesis stated a direction (that Trump judges author more DDRs), a one-tailed t-Test might suffice, but in general, the two-tailed is better because it measures both extreme ends of normal distribution instead of one. A two-tailed t-test typically requires a smaller value to prove statistical significance. Here, both p-values fall below 0.05.
- Here, since the p-values are less than our significance level, we can reject the Null Hypothesis. The difference between the two means is statistically significant. Thus, Trump judges author more DDRs than Other Republicans to a statistically significant degree.

The following internet sources helped guide me through the scientific method to pressure-test my hypotheses.

- https://statisticsbyjim.com/hypothesis-testing/t-tests-excel/
- https://home.csulb.edu/~msaintg/ppa696/696stsig.htm

I ran the calculations using Excel's Data Analysis ToolPak.

From: Hayden Smith
To: Professor Neal E. Devins and Professor Allison Orr Larsen
Date: January 30, 2024
Subject: **Attempt to Calculate the Increase in Republican Judges Authoring DDRs**

Research Hypothesis: There is a statistically significant relationship between political party and DDR authorship between the time periods 1943-2012 and 2014-2022.

The p value of 0.01 (one-tailed) is below the 0.05 alpha (95% confidence interval), rejecting our *null* hypothesis and finding that there is a statistically significant relationship between political party and DDR authorship from 1943-2012 compared to 2014-2022.

Year	**Total**		**Republicans**			**Democrats**		
	# of DDRs	**# of Active Judges**	**# of DDRs**	**% of DDRs**	**# of Active Judges**	**# of DDRs**	**% of DDRs**	**# of Active Judges**
1943-2012[3]	1276	569	662	**52%**	292 (51%)	602	**47%**	277 (49%)
2014-2022	382	242	253	**66%**	127 (52%)	129	**34%**	115 (48%)

t-Test: Paired Two Sample for Means

	Variable 1	*Variable 2*
Mean	0.428252556	0.56704525
Variance	0.016400858	0.01814833
Observations	2	2
Pearson Correlation	1	
Hypothesized Mean Difference	0	
df	1	
t Stat	-29.51657941	
P(T<=t) one-tail	**0.010779982**	
t Critical one-tail	6.313751515	
P(T<=t) two-tail	**0.021559963**	
t Critical two-tail	12.70620474	

[3] Data from 1943 to 2012 taken from Jeremy Horowitz, *Not Taking "No" for an Answer: An Empirical Assessment of Dissents from Denial of Rehearing En Banc*, 102 Geo. L. J. 59, 59 (2013).