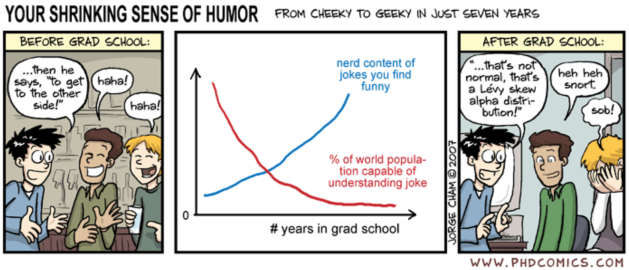
Wrapping up

Goals for the semester

In addition the the learning objectives listed in your syllabus our overarching goals for the semester are to develop:

- your ability to critically assess statistical information presented to you in scientific and non-scientific fora
- your sense of how to approach answering real world questions with data
- your ability to concisely and accurately describe statistical methods and results

Goals for the semester



***footnote:** *Thanks to Daragh at George Mason U. for this comic idea.*

Day 1 argument: This is a relevant class

I hoped to convince everyone here that statistics is relevant to everyone

You make many decisions during your day that are influenced by statistics

Statistics is not just relevant for public health, but also for other professions, including: policy, journalism and law

As we have tried to illustrate via the recurring "statistics is everywhere" segments, statistics is useful for understanding the news and the world around us - certainly during this pandemic we have seen a lot of public health and statistics in the news.

Science in the time of COVID

Lancet. 2021 19 December 2020-1 January; 396(10267): 1941.
Published online 2020 Dec 17. doi: 10.1016/S0140-6736(20)32709-4

PMCID: PMC7833527
PMID: 33341126

Science during COVID-19: where do we go from here?

The Lancet

Statistics is Everywhere!!

CORONAVIRUS | News

Experts on the one-and-done advantage offered by Johnson & Johnson's COVID-19 vaccine

Melissa Couto Zuber
The Canadian Press Staff
Contact

Published Tuesday, April 27, 2021 4:48AM EDT

MEDICINE

Blood Clots and the Johnson & Johnson Vaccine: What We Know So Far

Infectious disease physician-scientist Wilbur Chen discusses the rare cases of blood clots linked to the immunization

By Jim Daley on April 23, 2021

Statistics is Everywhere!!

What comparisons would you want to make?

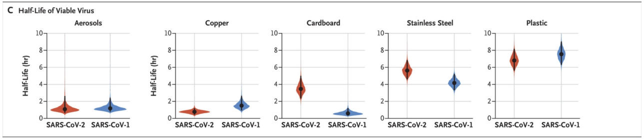
Statistics is Everywhere!!

Why are these numbers for clotting not presented with a margin of error?

How would you describe the risks and benefits to a family member?

Statistics is Everywhere!!

From a March 2020 New England Journal of Medicine Article



Statistics is Everywhere!!

What kind of a statistical test would you use to compare these data?

Have you changed your behavior based on these types of studies?

Statistics is Everywhere!!

Interesting opinion piece in the Lancet

Non-statistical considerations

Remembering that in our interpretation process we want to think about not just the statistical results. . . .

Evidence based suggestions: longhand notes

The Pen Is Mightier Than the Keyboard: Advantages of Longhand Over Laptop Note Taking

Pam A. Mueller[1] and Daniel M. Oppenheimer[2]
[1]Princeton University and [2]University of California, Los Angeles

Evidence based suggestions: longhand notes

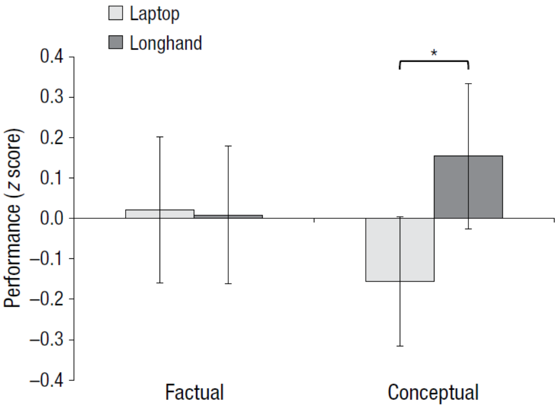


Fig. 1. Mean z-scored performance on factual-recall and conceptual-application questions as a function of note-taking condition (Study 1). The asterisk indicates a significant difference between conditions ($p < .05$). Error bars indicate standard errors of the mean.

Evidence based suggestions: anxiety reduction

A Comparison of Two In-Class Anxiety Reduction Exercises Before a Final Exam

Teaching of Psychology
2019, Vol. 46(1) 92-95

DOI: 10.1177/0098628318816182
journals.sagepub.com/home/top
SAGE

Virginia Clinton[1] and Stacy Meester[2]

Evidence based suggestions: anxiety reduction

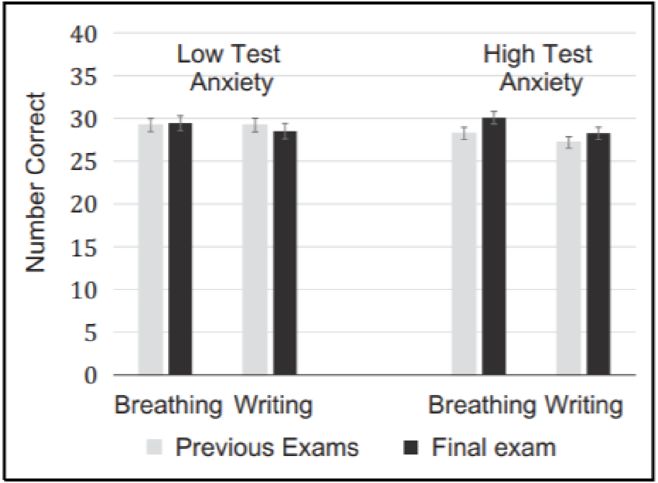


Figure 1. Previous exam and final exam performance by condition and level of trait test anxiety (means and $\pm$ 1 *SE*).

Some examples - what tests?

Example 1

August 8, 2019 NEJM

Vitamin D Supplementation and Prevention of Type 2 Diabetes Anastassios G. Pittas, M.D., Bess Dawson-Hughes, M.D., Patricia Sheehan, R.N., M.P.H., M.S., James H. Ware, Ph.D., William C. Knowler, M.D., Dr.P.H., Vanita R. Aroda, M.D., Irwin Brodsky, M.D., Lisa Ceglia, M.D., Chhavi Chadha, M.D., Ranee Chatterjee, M.D., M.P.H., Cyrus Desouza, M.B., B.S., Rowena Dolor, M.D., et al., for the D2d Research Group*

BACKGROUND Observational studies support an association between a low blood 25-hydroxyvitamin D level and the risk of type 2 diabetes. However, whether vitamin D supplementation lowers the risk of diabetes is unknown.

Example 1 cont.

METHODS We randomly assigned adults who met at least two of three glycemic criteria for prediabetes (fasting plasma glucose level, 100 to 125 mg per deciliter; plasma glucose level 2 hours after a 75-g oral glucose load, 140 to 199 mg per deciliter; and glycated hemoglobin level, 5.7 to 6.4%) and no diagnostic criteria for diabetes to receive 4000 IU per day of vitamin D3 or placebo, regardless of the baseline serum 25-hydroxyvitamin D level. The primary outcome in this time-to-event analysis was new-onset diabetes, and the trial design was event-driven, with a target number of diabetes events of 508.

RESULTS A total of 2423 participants underwent randomization (1211 to the vitamin D group and 1212 to the placebo group). By month 24, the mean serum 25-hydroxyvitamin D level in the vitamin D group was 54.3 ng per milliliter (from 27.7 ng per milliliter at baseline), as compared with 28.8 ng per milliliter in the placebo group (from 28.2 ng per milliliter at baseline). After a median follow-up of 2.5 years, the primary outcome of diabetes occurred in 293 participants in the vitamin D group and 323 in the placebo group (9.39 and 10.66 events per 100 person-years, respectively).

Example 1

Per the discussion in the article:

“Because vitamin D supplements are used increasingly in the U.S. adult population,29 approximately 8 of 10 participants had a baseline serum 25-hydroxyvitamin D level that was considered to be sufficient according to current recommendations ($\geq$ 20 ng per milliliter) to reduce the risk of many outcomes,23,30 including diabetes.6 The high percentage of participants with adequate levels of vitamin D may have limited the ability of the trial to detect a significant effect.”"

Example 2: Staph infections

Researchers recruited 917 patients who had tested positive for staphylococcus Aureus and randomly assignmed them to a staph-killing nasal ointment or placebo. They were interested in testing weather this drug was associated with a reduction in post-surgical infections. In the active treatment group17 of 504 patients developed infections, in the placebo group 32 of 413 patients developed infections.

- What are the exposure and outcome variables?
- What kind of a test would you use for these data?
- What is the null hypothesis of this test?

Common Issues

Odds v risks

from: Perspect Clin Res. 2015 Oct-Dec; 6(4): 222–224. doi: 10.4103/2229-3485.167092

Table 3

RR and OR for different event rates

	Deaths	Survivors	Odds of death	OR	Risk of death	RR
(a)						
Intervention 1	10	90	10/90=0.11	0.11/0.11=1.0	10/100=0.10	0.10/0.10=1.0
Control	10	90	10/90=0.11		10/100=0.10	
(b)						
Intervention 2	1	99	1/99=0.01	0.01/0.11=0.09	1/100=0.01	0.01/0.1=0.10
Control	10	90	10/90=0.11		10/100=0.10	
(c)						
Intervention 3	3	97	3/97=0.0309	0.0309/0.0101=3.06	3/100=0.03	0.03/0.01=3.0
Control	1	99	1/99=0.0101		1/100=0.01	
(d)						
Intervention 4	30	70	30/70=0.43	0.43/0/11=3.9	30/100=0.30	0.30/0.10=3.0
Control	10	90	10/90=0.11		10/100=0.10	
(e)						
Intervention 5	45	55	45/55=0.82	0.82/0.11=7.45	45/100=0.45	0.45/0.1=4.5
Control	10	90	10/90=0.11		10/100=0.1	

OR=Odds ratio, RR=Relative risk

Common Issues

method used vs P-value and CI vs. conclusion

Appropriate interpretation of a P-value and CI

Common Issues

Relationship between a confidence interval and a p-value

Common Issues

Confusion about the ANOVA null hypothesis and interpretation of results

Power example

Example

Lecture 21 has a worked example from the Pagano text that is worth reviewing Here we will go through another example from the Baldi and Moore textbook. This example assumes you are planning a quality control study to look at whether storage impacts the percieved sweetness of a beverage. The manufacturers are concerned that storage will decrease the sweetness. Ten professional tasters will rate the sweetness on a 10 point scale before and after storage. We know that the standard deviation of sweetness ratings is $= 1$. We also know that a mean sweetness change of 0.8 on this scale is noticed by consumers. We want 90% power and an alpha of 0.05 for our study. We have a set of 10 values representing the difference in sweetness caused by storage.

What is the null hypothesis here?

What is our alternative?

Is our hypothesis one or two sided?

Example

We will start by finding the Z alpha: This is the value on a standard normal distribution that corresponds to an α of 0.05 Here we are looking at the lower tail because we are interested in sweetness decreases only.

```
qnorm(.05)
```

```
## [1] -1.644854
```

Example

$$Z_\alpha = \frac{\overline{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

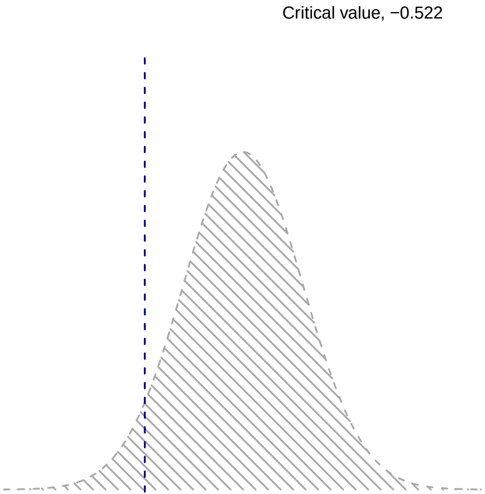
$$-1.645 = \frac{\overline{x} - 0}{\frac{1}{\sqrt{10}}}$$

Solve this for $\overline{x}$

$$\bar{x} = -1.645 \times \frac{1}{\sqrt{10}} = -0.522$$

null distribution

So now we have our null distribution in terms of our variable (not the Z score) with the value at which we reject the null



Example

We must choose a value at which to evaluate β. Here we will choose an alternate hypothesis that the mean sweetness difference is -0.8. Since we know a sample mean greater than -0.522 causes us to fail to reject H_0 we need to calculate the proportion of a distribution centered at -0.8 that would be below this value.

Using R to calculate the probability,

```
pnorm(-.522, mean=-0.8)
```

```
## [1] 0.6094938
```

Thus β P(do not reject null(0)|Null is false (true sweetness change is -0.8)) is $\sim$ 0.609

Remember that Power is 1-β = P(reject null | null is false)

In this example, Power is 1-0.609 or $\sim$ 0.391

more examples

dangerous skiiers?

We have learned that The typical skier death in CO is a 37-year-old experienced male skier wearing a helmet who loses control on an intermediate, groomed run and hits a tree. There is a suspicion that wearing a helmet when skiing increases the sense of safety and influences skiiers to take more risks than they would if they were not wearing a helmet. We want to test this by examining the proportion of skiers at a resort who get non-head related inuries while wearing a helmet. We know that in general the injury rate for skiers is 3 per 1000 person days. We collect a random sample of 100 skiers who consistently wear helmets and who skied 10 days in the most recent ski season. In our sample 6 individuals had an injury during the season.

How would we get a confidence interval for this?

dangerous skiiers?

Let's work this out by hand. . . .

dangerous skiiers?

- Let $\tilde{p} = \frac{\text{number of successes} + 2}{n+4}$
- Let $SE = \sqrt{\frac{\tilde{p}(1-\tilde{p})}{n+4}}$
- Then the CI is:

$$\tilde{p} \pm z^* \sqrt{\frac{\tilde{p}(1-\tilde{p})}{n+4}}$$

```
p.tilde <- (6 + 2)/(100 + 4)
se <- sqrt(p.tilde*(1-p.tilde)/104) # standard error
c(p.tilde - 1.96*se, p.tilde + 1.96*se) # CI
```

```
## [1] 0.02570932 0.12813684
```

what would a hypothesis test look like?

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

what would a hypothesis test look like?

The null value would be based on the underlying rate of 3 per 1000 person days, so for 100 skiers who each skied 10 days, we would have expected 3 injuries or 3%.

$$z = \frac{0.06 - 0.03}{\sqrt{\frac{0.03(0.07)}{100}}} = 6.55$$

what would a hypothesis test look like?


```r
# use r to get the p value
2*(pnorm(6.55, lower.tail=FALSE))
```

```
## [1] 5.753708e-11
```

Why would we be concerned about interpreting this result?

dangerous skiiers?

What is the R code we would use to generate an appropriate test?

```
prop.test(x=6,n=100, p=.03, correct=FALSE)
```

```
## Warning in prop.test(x = 6, n = 100, p = 0.03, correct = FALSE): Chi-squar
## approximation may be incorrect

##
##  1-sample proportions test without continuity correction
##
## data:  6 out of 100, null probability 0.03
## X-squared = 3.0928, df = 1, p-value = 0.07864
## alternative hypothesis: true p is not equal to 0.03
## 95 percent confidence interval:
##  0.02778612 0.12476815
## sample estimates:
##    p
## 0.06
```

Today's fun fact

The length of your hand from your wrist to your elbow is the same as the length of your foot from your heel to your big toe. If I want to show that these two measures are almost perfectly correlated how might I do that?

You could do a correlation test, a linear regression, a paired t, you could show a scatterplot. . .

Today's fun fact

What kind of plot would I expect to see for these data?

Perfectly correlated scatterplot. Observations falling on a straight line with increasing slope.

Staph infections

Researchers recruited 917 patients who had tested positive for staphylococcus Aureus and randomly assignmed them to a staph-killing nasal ointment or placebo. They were interested in testing weather this drug was associated with a reduction in post-surgical infections. In the active treatment group 17 of 504 patients developed infections, in the placebo group 32 of 413 patients developed infections.

- What are the exposure and outcome variables? exposure is drug vs placebo, outcome is development of infection vs not
- What kind of a test would you use for these data? 2 sample test of proportions, you could also do a chi-squared test of independence
- What is the null hypothesis of this test? There is no difference in proportions, There is no association between drug exposure and development of infections

Staph infections

Let's do this by hand

Staph infections

we can create the z-test for the difference between two proportions as:

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

remember

- If the null hypothesis is true, then p_1 is truly equal to p_2. In this case, our best estimate of the underlying proportion that they are both equal to is

$$\hat{p} = \frac{\text{no. successes in both samples}}{\text{no. individuals in both samples}}$$

- Also, our best guess at the SE for $\hat{p}$ is:

$\sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}}$

$\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_1} + \frac{1}{n_2})}$

This is the formula for the SE for the difference between two proportions but we have substituted $\hat{p}$ for p_1 and p_2.

Staph infections

$$z = \frac{.037 - .07748}{\sqrt{.0534 * 0.9466\left(\frac{1}{504} + \frac{1}{413}\right)}}$$

Staph infections


```
## [1] 0.001641061
```

In this case we are only interested in a reduction in infections - so we will only look at the left tail of the distribution.

Staph infections

In R?

```
##
##  2-sample test for equality of proportions without continuity
##  correction
##
## data:  c(17, 32) out of c(504, 413)
## X-squared = 8.5906, df = 1, p-value = 0.00169
## alternative hypothesis: less
## 95 percent confidence interval:
##  -1.00000000 -0.01839005
## sample estimates:
##     prop 1     prop 2
## 0.03373016 0.07748184
```

```
##
##  Exact binomial test
##
```

Coffee and race speed

I am interested in testing whether drinking coffee 20 minutes prior to a race increases sprinting speed. I recruit 11 runners from a running club and have them sprint 200 meters. I then give them each a cup of coffee and 20 minutes later ask them to sprint 200 meters again.

Coffee and race speed

Here are the data I collect:

Time 1	Time 2
32	31
40	44
26	24
28	23
34	36
56	48
24	21
28	30
26	24
36	40
30	25
____	____

coffee and race speed

In r?

```
2*pnorm(-.975)

time1<-c(32, 40,26,28,34,56,24,28,26,36,30)

time2<-c(31,44,24,23,36,48,21,30,24,40,25)

wilcox.test(time1,time2, paired=TRUE)
```

coffee and race speed

```
## [1] 0.3295603

## Warning in wilcox.test.default(time1, time2, paired = TRUE): cannot comput
## exact p-value with ties

##
##  Wilcoxon signed rank test with continuity correction
##
## data:  time1 and time2
## V = 44, p-value = 0.3477
## alternative hypothesis: true location shift is not equal to 0
```