431 Class 07

https://thomaselove.github.io/431-2024/

2024-09-17

## Today’s Agenda

* Building an Uncertainty Interval when we have Two Independent Samples
  + Finishing the Example from Last Class
* One-Factor Analysis of Variance: Comparing More than Two Independent Samples
  + Using Regression to Develop an ANOVA model

## Load packages and set theme

library(janitor)  
library(knitr)  
library(readxl) # to read in an .xlsx file  
library(car)  
library(ggdist) # for raincloud plots  
library(MKinfer)  
library(patchwork)  
library(rstanarm)  
library(easystats)  
library(tidyverse)  
  
theme\_set(theme\_bw())  
knitr::opts\_chunk$set(comment = NA)  
  
source("c07/data/Love-431.R") # for the lovedist() function

## Returning to the DM-464 data

dm7 <- read\_csv("c07/data/dm464\_class07.csv", show\_col\_types = FALSE) |>  
 janitor::clean\_names() |>  
 mutate(across(where(is.character), as\_factor)) |>  
 mutate(statin\_f = as\_factor(statin),  
 statin\_f = fct\_recode(statin\_f,   
 "Statin" = "1", "No Statin" = "0")) |>  
 mutate(id\_code = as.character(id\_code)) |>  
 mutate(a1c\_diff = a1c\_end - a1c\_base)  
  
dim(dm7)

[1] 464 10

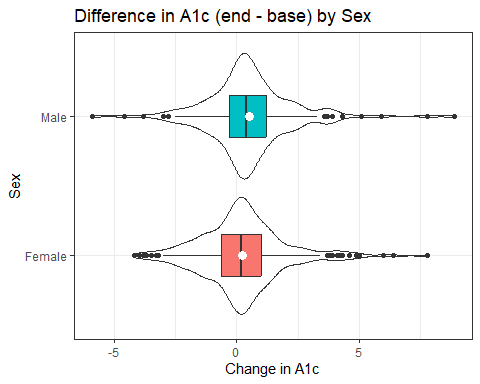
## A Strategy for Independent Samples

Suppose we want to estimate an uncertainty interval for the difference in means across two groups. To begin, plot the data from each sample.

* If the sample data in each group are well described as “Normal”, then use the OLS / pooled t procedure to obtain a confidence interval, or a Bayesian model to obtain a credible interval.
  + If the sample sizes are not the same, and the sample variances aren’t close to each other, consider using a Welch t procedure.
* If either sample’s data are “symmetric but with outliers”, then consider using the bootstrap to obtain a confidence interval, although an OLS or Bayesian result may also be reasonable.
* If either sample’s data are best described as “substantially skewed”, then consider whether you actually want to summarize with the difference in means, and consider whether a transformation might be helpful.

## Difference in A1c by Sex

ggplot(dm7, aes(x = a1c\_diff, y = sex)) +  
 geom\_violin() + geom\_boxplot(aes(fill = sex), width = 0.3) +  
 stat\_summary(fun = mean, geom = "point", shape = 16, size = 3, col = "white") +  
 labs(x = "Change in A1c", y = "Sex",  
 title = "Difference in A1c (end - base) by Sex") +  
 guides(fill = "none")



## Numerical Summary

dm7 |> group\_by(sex) |>   
 reframe(lovedist(a1c\_diff)) |>   
 print\_md(digits = 3)

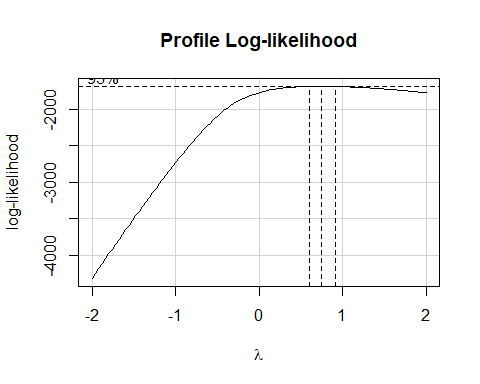
| sex | n | miss | mean | sd | med | mad | min | q25 | q75 | max |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Female | 288 | 0 | 0.227 | 1.747 | 0.200 | 1.186 | -4.200 | -0.625 | 1.000 | 7.800 |
| Male | 176 | 0 | 0.503 | 1.838 | 0.400 | 1.186 | -5.900 | -0.300 | 1.200 | 8.900 |

* Observed difference in sample means is 0.503 - 0.227 = 0.276
* Do we have a balanced design?
* Do we feel comfortable pooling these standard deviations?
* Do we feel comfortable assuming Normality for each sex?

## Might a transformation help here?

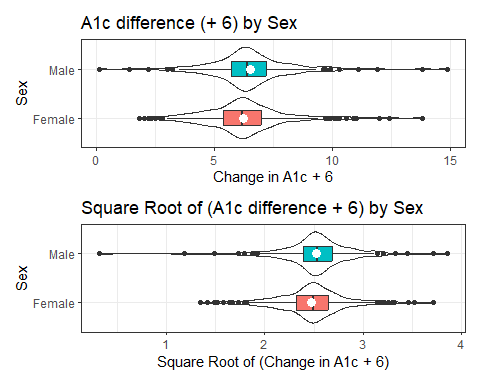
Some of the changes in A1c are negative (minimum was -5.9)

dm7 <- dm7 |> mutate(a1c\_diff\_p6 = a1c\_diff + 6)  
fit4 <- lm(a1c\_diff\_p6 ~ sex, data = dm7)  
boxCox(fit4)



## Does using a square root help us here?

p1 <- ggplot(dm7, aes(x = a1c\_diff\_p6, y = sex)) +  
 geom\_violin() + geom\_boxplot(aes(fill = sex), width = 0.3) +  
 stat\_summary(fun = mean, geom = "point", shape = 16, size = 3, col = "white") +  
 labs(x = "Change in A1c + 6", y = "Sex",  
 title = "A1c difference (+ 6) by Sex") +  
 guides(fill = "none")  
  
p2 <- ggplot(dm7, aes(x = sqrt(a1c\_diff\_p6), y = sex)) +  
 geom\_violin() + geom\_boxplot(aes(fill = sex), width = 0.3) +  
 stat\_summary(fun = mean, geom = "point", shape = 16, size = 3, col = "white") +  
 labs(x = "Square Root of (Change in A1c + 6)", y = "Sex",  
 title = "Square Root of (A1c difference + 6) by Sex") +  
 guides(fill = "none")  
  
p1 / p2



## Addressing outliers?

### OLS model

fit5 <- lm(a1c\_diff ~ sex, data = dm7)  
  
model\_parameters(fit5, ci = 0.90)

Parameter | Coefficient | SE | 90% CI | t(462) | p  
-----------------------------------------------------------------  
(Intercept) | 0.23 | 0.11 | [ 0.05, 0.40] | 2.16 | 0.031  
sex [Male] | 0.28 | 0.17 | [ 0.00, 0.56] | 1.62 | 0.105

Uncertainty intervals (equal-tailed) and p-values (two-tailed) computed  
 using a Wald t-distribution approximation.

* Male subjects had A1c changes that were 0.28 larger than females, on average, with 90% uncertainty interval (0, 0.56) according to our OLS model (pooled t test.)

## Addressing outliers?

### Bayesian model

set.seed(20240912)  
fit6 <- stan\_glm(a1c\_diff ~ sex, data = dm7, refresh = 0)  
  
model\_parameters(fit6, ci = 0.90)

Parameter | Median | 90% CI | pd | Rhat | ESS | Prior  
---------------------------------------------------------------------------------------  
(Intercept) | 0.22 | [ 0.05, 0.39] | 98.42% | 1.000 | 3666.00 | Normal (0.33 +- 4.46)  
sexMale | 0.28 | [ 0.00, 0.57] | 94.80% | 1.000 | 3649.00 | Normal (0.00 +- 9.19)

Uncertainty intervals (equal-tailed) and p-values (two-tailed) computed  
 using a MCMC distribution approximation.

* Male subjects had A1c changes that were 0.28 larger than females, on average, with 90% uncertainty interval (0, 0.57) according to our Bayesian model.

## Addressing outliers?

### Bootstrap with boot.t.test from MKinfer

set.seed(20240912)  
boot.t.test(a1c\_diff ~ sex, data = dm7, conf.level = 0.90)

Bootstrap Welch Two Sample t-test  
  
data: a1c\_diff by sex  
number of bootstrap samples: 9999  
bootstrap p-value = 0.1096   
bootstrap difference of means (SE) = -0.2787276 (0.1715815)   
90 percent bootstrap percentile confidence interval:  
 -0.566107955 0.002070707  
  
Results without bootstrap:  
t = -1.6028, df = 355.56, p-value = 0.1099  
alternative hypothesis: true difference in means is not equal to 0  
90 percent confidence interval:  
 -0.561349458 0.008003498  
sample estimates:  
mean in group Female mean in group Male   
 0.2267361 0.5034091

## Comparing Our Results

These comparisons describe the Female - Male differences. The point estimate is -0.28.

| Method | 90% Uncertainty Interval |
| --- | --- |
| OLS / pooled t | (-0.56, 0.00) |
| Welch t | (-0.56, 0.01) |
| Bayesian model | (-0.57, 0.00) |
| Bootstrap[[1]](#footnote-43) | (-0.57, 0.00) |

## 2 Independent Samples Strategy

Suppose we want to estimate an uncertainty interval for the difference in means across two groups. To begin, plot the data from each sample.

* If the sample data in each group are well described as “Normal”, then use the OLS / pooled t procedure to obtain a confidence interval, or a Bayesian model to obtain a credible interval.
  + If the sample sizes are not the same, and the sample variances aren’t close to each other, consider using a Welch t procedure.
* If either sample’s data are “symmetric but with outliers”, then consider using the bootstrap to obtain a confidence interval, although an OLS or Bayesian result may also be reasonable.
* If either sample’s data are best described as “substantially skewed”, then consider whether you actually want to summarize with the difference in means, and consider whether a transformation might be helpful.

## Paired Samples Strategy

Suppose we want to estimate an uncertainty interval for the mean of a set of paired differences.

* Calculate the paired differences, then plot them.
* If the sample data are well described as “Normal”, then use the OLS / paired t procedure to obtain a confidence interval, or a Bayesian model to obtain a credible interval.
* If the sample data are best described as “symmetric but with outliers”, then use the bootstrap to obtain a confidence interval, although an OLS or Bayesian result may be quite similar.
* If the sample data are best described as “substantially skewed”, then consider whether you actually want to summarize with the mean, and consider whether a transformation might be helpful.

# What if we have more than two independent samples?

## The ohio\_2020 data

ohio\_2020.xlsx rows describe Ohio’s 88 counties:

* FIPS code (identifier for mapping), state and county name
* health outcomes (standardized: more positive means **better** outcomes, because we’ve taken the negative of the Z score CHR provides)
* health behavior ranking (1-88, we’ll divide into 4 groups)
* clinical care ranking (1-88, we’ll split into 3 groups)
* and some other variables

### Sources (these bullets are links)

* [County Health Rankings](https://www.countyhealthrankings.org/app/ohio/2020/downloads) (2020 Ohio Data)
* [Wikipedia for 2016 Election Results](https://en.wikipedia.org/wiki/2016_United_States_presidential_election_in_Ohio#By_county)

## Importing Data / Creating Factors

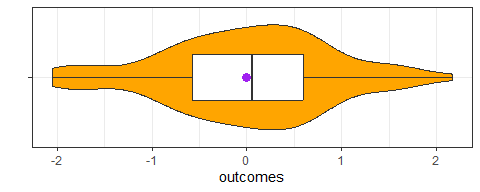
ohio20 <- read\_xlsx("c07/data/ohio\_2020.xlsx") |>  
 mutate(behavior = Hmisc::cut2(rk\_behavior, g = 4),  
 clin\_care = Hmisc::cut2(rk\_clin\_care, g = 3)) |>  
 mutate(behavior = fct\_recode(behavior,  
 "Best" = "[ 1,23)", "High" = "[23,45)",  
 "Low" = "[45,67)", "Worst" = "[67,88]")) |>  
 mutate(clin\_care = fct\_recode(clin\_care,  
 "Strong" = "[ 1,31)", "Middle" = "[31,60)",  
 "Weak" = "[60,88]")) |>  
 select(FIPS, state, county, outcomes, behavior, clin\_care,   
 everything())

## A Quick Look at the Data

ohio20 |> filter(county == "Cuyahoga") |>  
 select(FIPS, county, outcomes, behavior, clin\_care)

# A tibble: 1 × 5  
 FIPS county outcomes behavior clin\_care  
 <chr> <chr> <dbl> <fct> <fct>   
1 39035 Cuyahoga -0.807 Worst Strong

ggplot(ohio20, aes(x = "", y = outcomes)) + geom\_violin(fill = "orange") +  
 geom\_boxplot(width = 0.4) +  
 stat\_summary(fun = mean, geom = "point", shape = 16, size = 3, col = "purple") +  
 coord\_flip() + labs(x = "")



## Key Measure Details

* **outcomes** = quantity that describes the county’s premature death and quality of life results, weighted equally and standardized (z scores).
  + Higher (more positive) values indicate better outcomes in this county.

## Key Measure Details

* **behavior** = (Best/High/Low/Worst) reflecting adult smoking, obesity, food environment, inactivity, exercise, drinking, alcohol-related driving deaths, sexually transmitted infections and teen births.
  + Counties in the Best group had the best behavior results.

## Key Measure Details

* **clin\_care** = (Strong/Middle/Weak) reflects rates of uninsured, care providers, preventable hospital stays, diabetes monitoring and mammography screening.
  + Strong means that clinical care is strong in this county.

### Today’s Question

1. How do average health outcomes vary across groups of counties defined by health behavior?

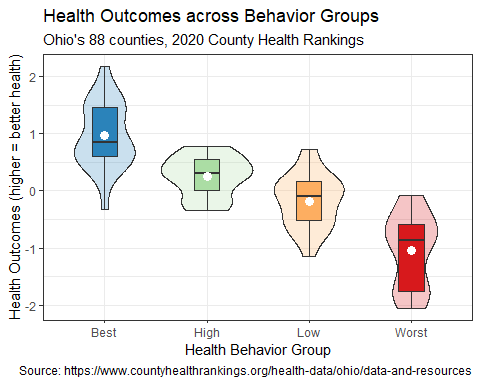
## Samples: Comparing Means

1. What is the outcome under study?
2. What are the (in this case, ) treatment/exposure groups?
3. Were the data in fact collected using independent samples?
4. Are the data random samples from the population(s) of interest? Or is there at least a reasonable argument for generalizing from the samples to the population(s)?
5. What is the uncertainty level we require?
6. Are we doing one-sided or two-sided testing? (usually 2-sided)
7. What does the distribution of each individual sample tell us about which inferential procedure to use?
8. Are there meaningful differences between population means?
9. Can we identify pairwise comparisons of means that show meaningful differences using an appropriate procedure that protects against Type I error expansion due to multiple comparisons? (to be discussed next time)

## Our Question

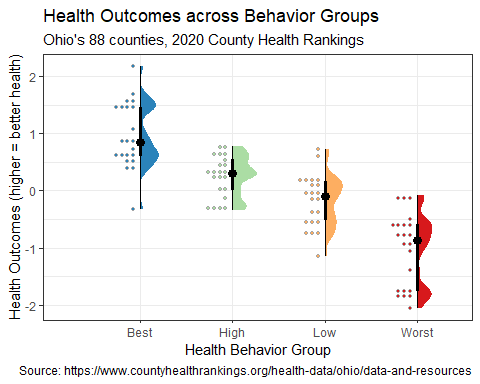
Do average health outcomes differ by health behavior?

ggplot(ohio20, aes(x = behavior, y = outcomes,   
 fill = behavior)) +  
 geom\_violin(alpha = 0.25) +  
 geom\_boxplot(width = 0.25) +  
 stat\_summary(fun = mean, geom = "point", shape = 16, size = 3, col = "white") +  
 guides(fill = "none") +   
 scale\_fill\_brewer(palette = "Spectral", direction = -1) +  
 labs(x = "Health Behavior Group",   
 y = "Health Outcomes (higher = better health)",  
 title = "Health Outcomes across Behavior Groups",  
 subtitle = "Ohio's 88 counties, 2020 County Health Rankings",  
 caption = "Source: https://www.countyhealthrankings.org/health-data/ohio/data-and-resources")



## Question 1 Raincloud Plots?

ggplot(ohio20, aes(x = behavior, y = outcomes,   
 fill = behavior)) +  
 ggdist::stat\_halfeye(adjust = 0.5, width = 0.3, .width = c(0.5, 1)) +  
 ggdist::stat\_dots(side = "left", dotsize = 1, justification = 1.05, binwidth = 0.1) +  
 guides(fill = "none") +   
 scale\_fill\_brewer(palette = "Spectral", direction = -1) +  
 labs(x = "Health Behavior Group",   
 y = "Health Outcomes (higher = better health)",  
 title = "Health Outcomes across Behavior Groups",  
 subtitle = "Ohio's 88 counties, 2020 County Health Rankings",  
 caption = "Source: https://www.countyhealthrankings.org/health-data/ohio/data-and-resources")



## Numerical Summaries by Group

How do average health outcomes vary across groups of counties defined by health behavior?

ohio20 |>  
 group\_by(behavior) |>  
 reframe(lovedist(outcomes)) |>  
 kable(digits = 2)

| behavior | n | miss | mean | sd | med | mad | min | q25 | q75 | max |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Best | 22 | 0 | 0.96 | 0.57 | 0.86 | 0.60 | -0.33 | 0.60 | 1.46 | 2.17 |
| High | 22 | 0 | 0.25 | 0.35 | 0.30 | 0.41 | -0.35 | 0.00 | 0.55 | 0.77 |
| Low | 22 | 0 | -0.18 | 0.47 | -0.09 | 0.43 | -1.15 | -0.52 | 0.16 | 0.73 |
| Worst | 22 | 0 | -1.04 | 0.63 | -0.87 | 0.65 | -2.05 | -1.75 | -0.59 | -0.08 |

Note that there is no missing data here.

## Analysis of Variance: Question 1

Does the mean outcomes result differ meaningfully across the behavior groups?

$$
H\_0: \mu\_{Best} = \mu\_{High} = \mu\_{Low} = \mu\_{Worst} \mbox{ vs. } \\
H\_A: \mbox{At least one } \mu \mbox{ is different.}
$$

To test this set of hypotheses, we will build a linear model to predict each county’s outcome based on what behavior group the county is in.

## Building the Linear Model

Do we see meaningful differences in population means of outcomes across behavior groups?

model\_one <- lm(outcomes ~ behavior, data = ohio20)  
  
model\_parameters(model\_one, ci = 0.90)

Parameter | Coefficient | SE | 90% CI | t(84) | p  
------------------------------------------------------------------------  
(Intercept) | 0.96 | 0.11 | [ 0.78, 1.15] | 8.74 | < .001  
behavior [High] | -0.71 | 0.16 | [-0.97, -0.45] | -4.55 | < .001  
behavior [Low] | -1.14 | 0.16 | [-1.40, -0.88] | -7.32 | < .001  
behavior [Worst] | -2.01 | 0.16 | [-2.27, -1.75] | -12.85 | < .001

Uncertainty intervals (equal-tailed) and p-values (two-tailed) computed  
 using a Wald t-distribution approximation.

How do we interpret this result?

## Meaning of indicator variables?

outcomes = 0.96 - 0.71 behaviorHigh   
 - 1.14 behaviorLow - 2.01 behaviorWorst

| group | behaviorHigh | behaviorLow | behaviorWorst |
| --- | --- | --- | --- |
| Best | 0 | 0 | 0 |
| High | 1 | 0 | 0 |
| Low | 0 | 1 | 0 |
| Worst | 0 | 0 | 1 |

* So what is the predicted outcomes score for a county in the High behavior group, according to this model?

## Interpreting the Indicator Variables

outcomes = 0.96 - 0.71 behaviorHigh   
 - 1.14 behaviorLow - 2.01 behaviorWorst

What predictions does the model make? Do these make sense?

| group | High | Low | Worst | Prediction |
| --- | --- | --- | --- | --- |
| Best | 0 | 0 | 0 | 0.96 |
| High | 1 | 0 | 0 | 0.96 - 0.71 = 0.25 |
| Low | 0 | 1 | 0 | 0.96 - 1.14 = -0.18 |
| Worst | 0 | 0 | 1 | 0.96 - 2.01 = -1.05 |

## Interpreting the Indicator Variables

outcomes = 0.96 - 0.71 behaviorHigh   
 - 1.14 behaviorLow - 2.01 behaviorWorst

ohio20 |> group\_by(behavior) |>  
 summarise(n = n(), mean = round\_half\_up(mean(outcomes),2)) |>   
 kable(digits = 2)

| behavior | n | mean |
| --- | --- | --- |
| Best | 22 | 0.96 |
| High | 22 | 0.25 |
| Low | 22 | -0.18 |
| Worst | 22 | -1.04 |

## ANOVA for Linear Model

Do we see meaningful differences in mean outcome across the behavior groups?

$$
H\_0: \mu\_{Best} = \mu\_{High} = \mu\_{Low} = \mu\_{Worst} \mbox{ vs. } \\
H\_A: \mbox{At least one } \mu \mbox{ is different.}
$$

anova(model\_one)

Analysis of Variance Table  
  
Response: outcomes  
 Df Sum Sq Mean Sq F value Pr(>F)   
behavior 3 46.421 15.4736 57.718 < 2.2e-16 \*\*\*  
Residuals 84 22.519 0.2681   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## So, what’s in the ANOVA table? (df)

The ANOVA table reports here on a single **factor** (behavior group) with 4 levels, and on the residual variation in health **outcomes**.

anova(model\_one)[1:2]

Df Sum Sq  
behavior 3 46.421  
Residuals 84 22.519

**Degrees of Freedom** (df) is an index of sample size…

* df for our factor (behavior) is one less than the number of categories. We have four behavior groups, so 3 degrees of freedom.
* Adding df(behavior) + df(Residuals) = 3 + 84 = 87 = df(Total), one less than the number of observations (counties) in Ohio.
* *n* observations and *g* groups yield residual df in a one-factor ANOVA table.

## ANOVA table: Sum of Squares

anova(model\_one)[1:3]

Df Sum Sq Mean Sq  
behavior 3 46.421 15.4736  
Residuals 84 22.519 0.2681

**Sum of Squares** (Sum Sq, or SS) is an index of variation…

* SS(factor), here SS(behavior) measures the amount of variation accounted for by the behavior groups in our model\_one.
* The total variation in outcomes to be explained by the model is SS(factor) + SS(Residuals) = SS(Total) in a one-factor ANOVA table.
* We describe the proportion of variation explained by a one-factor ANOVA model with (“eta-squared”: same as in regression)

## ANOVA table: (Mean Square, F ratio)

anova(model\_one)[1:4]

Df Sum Sq Mean Sq F value  
behavior 3 46.421 15.4736 57.718  
Residuals 84 22.519 0.2681

**Mean Square** (Mean Sq, or MS) = Sum of Squares / df

* MS(Residuals) estimates the **residual variance**, the square of the residual standard deviation (residual standard error in earlier work).
* The ratio of MS values is the ANOVA **F value**.

## ANOVA Table p value

anova(model\_one)

Analysis of Variance Table  
  
Response: outcomes  
 Df Sum Sq Mean Sq F value Pr(>F)   
behavior 3 46.421 15.4736 57.718 < 2.2e-16 \*\*\*  
Residuals 84 22.519 0.2681   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

* The *p* value is derived from the ANOVA F statistic, as compared to the F distribution.
* Which F distribution is specified by the two degrees of freedom values…

pf(57.718, df1 = 3, df2 = 84, lower.tail = FALSE)

[1] 2.377323e-20

## Alternative ANOVA display

summary(aov(model\_one))

Df Sum Sq Mean Sq F value Pr(>F)   
behavior 3 46.42 15.474 57.72 <2e-16 \*\*\*  
Residuals 84 22.52 0.268   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

estimate\_means(model\_one, ci = 0.90)

We selected `by = c("behavior")`.

Estimated Marginal Means  
  
behavior | Mean | SE | 90% CI  
----------------------------------------  
Best | 0.96 | 0.11 | [ 0.78, 1.15]  
High | 0.25 | 0.11 | [ 0.07, 0.44]  
Low | -0.18 | 0.11 | [-0.36, 0.01]  
Worst | -1.04 | 0.11 | [-1.22, -0.86]  
  
Marginal means estimated at behavior

So, what might we conclude? Is this a surprise?

## Session Information

xfun::session\_info()

R version 4.4.1 (2024-06-14 ucrt)  
Platform: x86\_64-w64-mingw32/x64  
Running under: Windows 11 x64 (build 22631)  
  
Locale:  
 LC\_COLLATE=English\_United States.utf8   
 LC\_CTYPE=English\_United States.utf8   
 LC\_MONETARY=English\_United States.utf8  
 LC\_NUMERIC=C   
 LC\_TIME=English\_United States.utf8   
  
Package version:  
 abind\_1.4-5 arrangements\_1.1.9 askpass\_1.2.0   
 backports\_1.5.0 base64enc\_0.1-3 bayesplot\_1.11.1   
 bayestestR\_0.14.0 BH\_1.84.0.0 bit\_4.0.5   
 bit64\_4.0.5 blob\_1.2.4 boot\_1.3-31   
 broom\_1.0.6 bslib\_0.8.0 cachem\_1.1.0   
 callr\_3.7.6 car\_3.1-2 carData\_3.0-5   
 cellranger\_1.1.0 checkmate\_2.3.2 cli\_3.6.3   
 clipr\_0.8.0 cluster\_2.1.6 coda\_0.19-4.1   
 codetools\_0.2-20 colorspace\_2.1-1 colourpicker\_1.3.0   
 commonmark\_1.9.1 compiler\_4.4.1 conflicted\_1.2.0   
 correlation\_0.8.5 cowplot\_1.1.3 cpp11\_0.5.0   
 crayon\_1.5.3 crosstalk\_1.2.1 curl\_5.2.2   
 data.table\_1.16.0 datasets\_4.4.1 datawizard\_0.12.3   
 DBI\_1.2.3 dbplyr\_2.5.0 Deriv\_4.1.3   
 desc\_1.4.3 digest\_0.6.37 distributional\_0.4.0   
 doBy\_4.6.22 dplyr\_1.1.4 DT\_0.33   
 dtplyr\_1.3.1 dygraphs\_1.1.1.6 easystats\_0.7.3   
 effectsize\_0.8.9 emmeans\_1.10.4 estimability\_1.5.1   
 evaluate\_0.24.0 exactRankTests\_0.8-35 fansi\_1.0.6   
 farver\_2.1.2 fastmap\_1.2.0 fontawesome\_0.5.2   
 forcats\_1.0.0 foreach\_1.5.2 foreign\_0.8-87   
 Formula\_1.2-5 fs\_1.6.4 gargle\_1.5.2   
 generics\_0.1.3 ggdist\_3.3.2 ggplot2\_3.5.1   
 ggridges\_0.5.6 glmnet\_4.1-8 glue\_1.7.0   
 gmp\_0.7-5 googledrive\_2.1.1 googlesheets4\_1.1.1   
 graphics\_4.4.1 grDevices\_4.4.1 grid\_4.4.1   
 gridExtra\_2.3 gtable\_0.3.5 gtools\_3.9.5   
 haven\_2.5.4 highr\_0.11 Hmisc\_5.1-3   
 hms\_1.1.3 htmlTable\_2.4.3 htmltools\_0.5.8.1   
 htmlwidgets\_1.6.4 httpuv\_1.6.15 httr\_1.4.7   
 ids\_1.0.1 igraph\_2.0.3 inline\_0.3.19   
 insight\_0.20.4 isoband\_0.2.7 iterators\_1.0.14   
 janitor\_2.2.0 jomo\_2.7-6 jquerylib\_0.1.4   
 jsonlite\_1.8.8 knitr\_1.48 labeling\_0.4.3   
 later\_1.3.2 lattice\_0.22-6 lazyeval\_0.2.2   
 lifecycle\_1.0.4 lme4\_1.1-35.5 loo\_2.8.0   
 lubridate\_1.9.3 magrittr\_2.0.3 markdown\_1.13   
 MASS\_7.3-61 Matrix\_1.7-0 MatrixModels\_0.5.3   
 matrixStats\_1.4.0 memoise\_2.0.1 methods\_4.4.1   
 mgcv\_1.9.1 mice\_3.16.0 miceadds\_3.17-44   
 microbenchmark\_1.5.0 mime\_0.12 miniUI\_0.1.1.1   
 minqa\_1.2.8 mitml\_0.4-5 mitools\_2.4   
 MKdescr\_0.8 MKinfer\_1.2 modelbased\_0.8.8   
 modelr\_0.1.11 multcomp\_1.4-26 munsell\_0.5.1   
 mvtnorm\_1.3-1 nlme\_3.1-164 nloptr\_2.1.1   
 nnet\_7.3-19 numDeriv\_2016.8.1.1 openssl\_2.2.1   
 ordinal\_2023.12.4.1 pan\_1.9 parallel\_4.4.1   
 parameters\_0.22.2 patchwork\_1.2.0 pbkrtest\_0.5.3   
 performance\_0.12.3 pillar\_1.9.0 pkgbuild\_1.4.4   
 pkgconfig\_2.0.3 plyr\_1.8.9 posterior\_1.6.0   
 prettyunits\_1.2.0 processx\_3.8.4 progress\_1.2.3   
 promises\_1.3.0 ps\_1.7.7 purrr\_1.0.2   
 quadprog\_1.5-8 quantreg\_5.98 QuickJSR\_1.3.1   
 R6\_2.5.1 ragg\_1.3.2 rappdirs\_0.3.3   
 RColorBrewer\_1.1-3 Rcpp\_1.0.13 RcppArmadillo\_14.0.0.1  
 RcppEigen\_0.3.4.0.2 RcppParallel\_5.1.9 readr\_2.1.5   
 readxl\_1.4.3 rematch\_2.0.0 rematch2\_2.1.2   
 report\_0.5.9 reprex\_2.1.1 reshape2\_1.4.4   
 rlang\_1.1.4 rmarkdown\_2.28 rpart\_4.1.23   
 rstan\_2.32.6 rstanarm\_2.32.1 rstantools\_2.4.0   
 rstudioapi\_0.16.0 rvest\_1.0.4 sandwich\_3.1-0   
 sass\_0.4.9 scales\_1.3.0 see\_0.9.0   
 selectr\_0.4.2 shape\_1.4.6.1 shiny\_1.9.1   
 shinyjs\_2.1.0 shinystan\_2.6.0 shinythemes\_1.2.0   
 snakecase\_0.11.1 sourcetools\_0.1.7.1 SparseM\_1.84.2   
 splines\_4.4.1 StanHeaders\_2.32.10 stats\_4.4.1   
 stats4\_4.4.1 stringi\_1.8.4 stringr\_1.5.1   
 survival\_3.7-0 sys\_3.4.2 systemfonts\_1.1.0   
 tensorA\_0.36.2.1 textshaping\_0.4.0 TH.data\_1.1-2   
 threejs\_0.3.3 tibble\_3.2.1 tidyr\_1.3.1   
 tidyselect\_1.2.1 tidyverse\_2.0.0 timechange\_0.3.0   
 tinytex\_0.52 tools\_4.4.1 tzdb\_0.4.0   
 ucminf\_1.2.2 utf8\_1.2.4 utils\_4.4.1   
 uuid\_1.2.1 V8\_5.0.0 vctrs\_0.6.5   
 viridis\_0.6.5 viridisLite\_0.4.2 vroom\_1.6.5   
 withr\_3.0.1 xfun\_0.47 xml2\_1.3.6   
 xtable\_1.8-4 xts\_0.14.0 yaml\_2.3.10   
 zoo\_1.8-12

1. without pooling the standard deviations [↑](#footnote-ref-43)