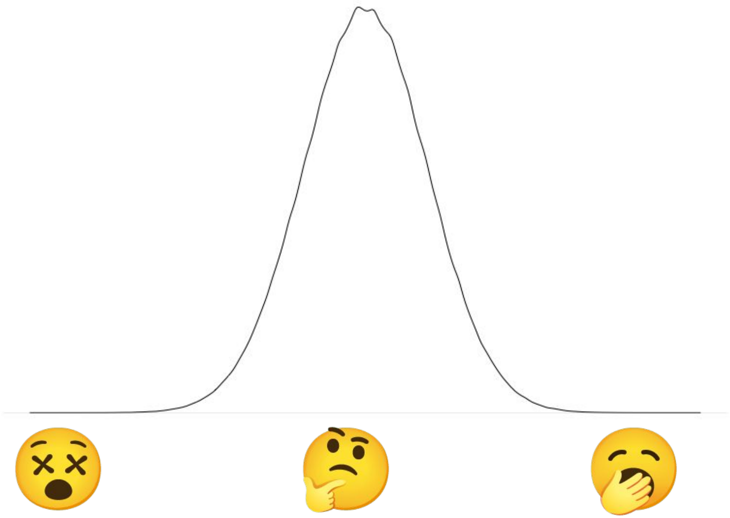
Philosophy of Science and Statistical Reasoning

F-distribution & Nonparametric Inference

But first, …

📢



Student Puzzlement Scale

📰 News

> ***The Harvard Professor and the Bloggers***
>
> When Francesca Gino, a rising academic star, was accused of falsifying data — about how to stop dishonesty — it didn't just torch her career. It inflamed a crisis in behavioral science.

— The New York Times (Sep. 30, 2023)

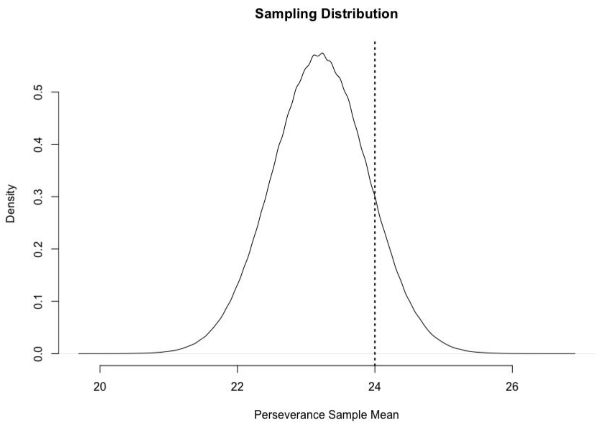
Wayback Machine

Data Colada (response post)

The New York Times (free subscription)

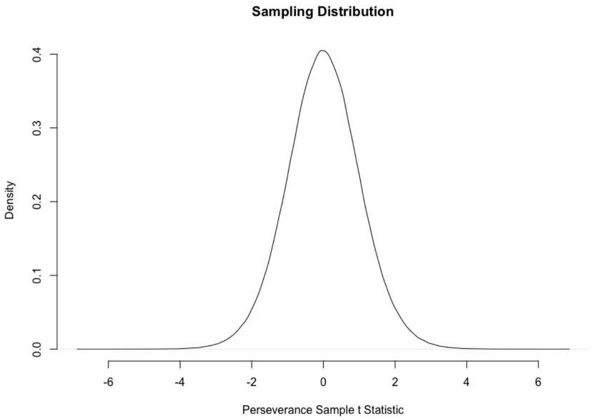
🍺 Pub quiz

Previously, on statistical reasoning



Normal distribution
mu = population mean
sd = standard error of sample mean

Test-statistic: sample mean



Student's *t*-distribution
df = n − 1

Test-statistic: *t*-statistic

What will we learn today?

Topics

Statistical reasoning
Empirical cycle
Probability distributions
Frequentist inference
Sample / sampling distribution
Central limit theorem
Normal distribution
P-value
Type I/II errors
Effect size
Confidence interval
Power
Test statistics
Linear regression
t-Test
Moderation
F-distribution
Nonparametric inference
ANOVA
Bayesian inference

Questions

How can we use frequentist statistics to test more complex models? (*F*-distribution)

What can we do if we can't meet a test's assumptions? (nonparametric inference)

🎁 Two new probability distributions and two new test statistics!

Decomposition of variability

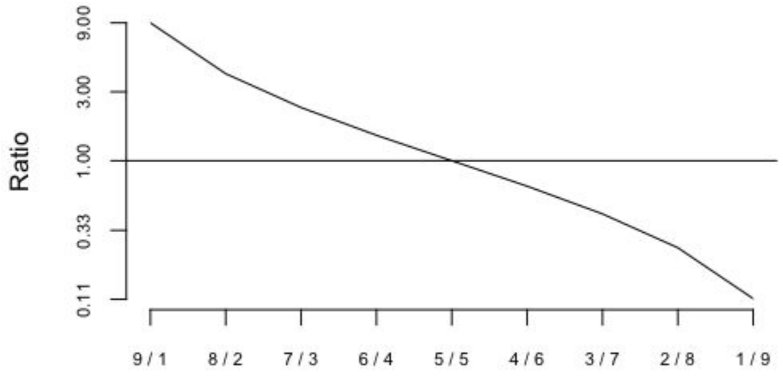
F

=

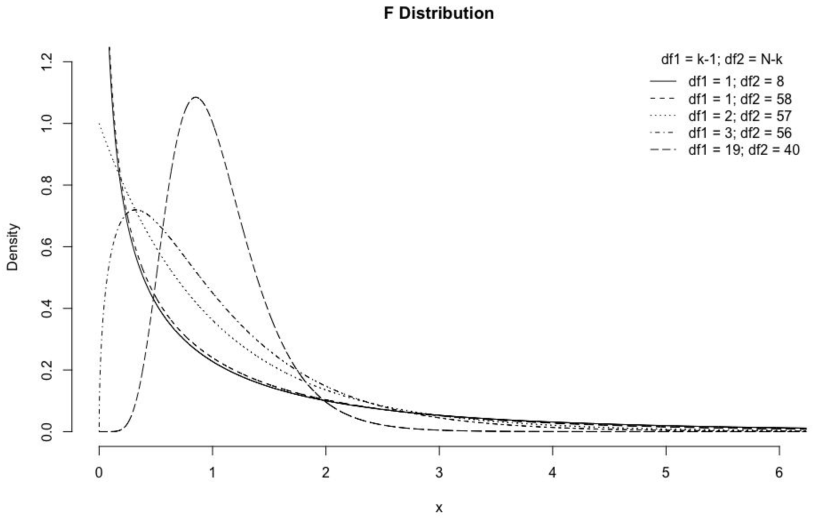
explained variance / unexplained variance

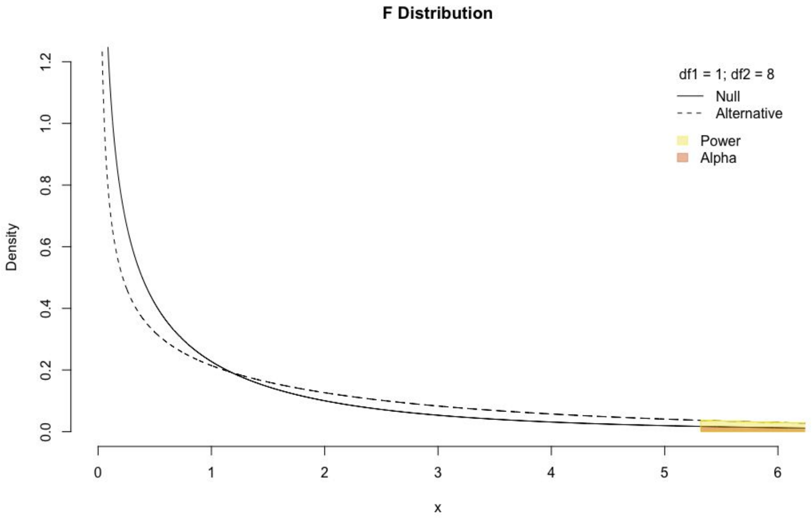
=

model / error



F-distribution





$df_{model} = k - 1$
$df_{error} = N - k$

How t F R^2 all of this related?

$F = t^2$

$F = (R^2 / (1 - R^2)) \times (df_{error} / df_{model})$

🧑‍💻

```
?sleep
mod <- extra ~ group
summary(aov(mod, sleep))[[1]]["F value"][1, ]
t.test(mod, sleep)$statistic^2
```

🧑‍💻

```
?iris
mod <- Petal.Length ~ Sepal.Length + Sepal.Width
fit <- summary(lm(formula = mod, data = iris))
fit$fstatistic["value"]
(fit$r.squared / (1 - fit$r.squared)) *
(fit$fstatistic["dendf"] / fit$fstatistic["numdf"])
```

F-ratio for our length data (one-way ANOVA)

But… we forgot to measure their actual length!

Let me handle that. You have a picture?

🤦 🥷

☕

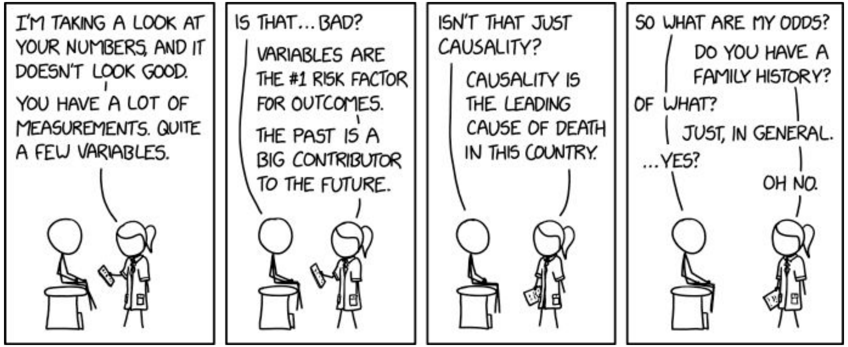


Illustration by [Randall Munroe](wtf) (wtf)

15:00

Nonparametric inference

🥷

🤯 data does not have the precision of an interval scale
🤯 serious concerns about (extreme) deviations from normal distribution
🤯considerable difference in the number of subjects for each group

Advantages: ordinal data, more robust (not sensitive to outliers), any distribution of the data

Disadvantages: less power

[Level of measurement]
Nominal: 🍏🍓🍌
Ordinal: ☹️😐😀
Interval: 📅 🌡️ (°C)
Ratio: 🕔📏⚖️

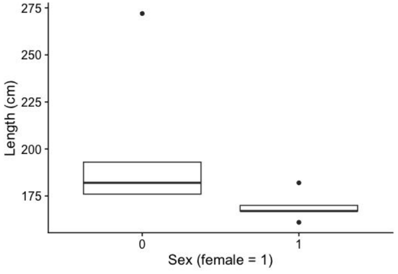
Ranking

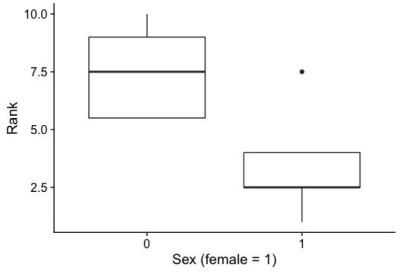
interval

ordinal

Sex (f=1)	Length (cm)
1	161
1	167
0	272
1	170
0	176
1	182
0	182
1	167
0	176
0	193

Ordered by length	Ranked length	Ranked length /w ties
1	1	1
1	2	(2+3)/2 = 2.5
1	3	(2+3)/2 = 2.5
1	4	4
0	5	(5+6)/2 = 5.5
0	6	(5+6)/2 = 5.5
1	7	(7+8)/2 = 7.5
0	8	(7+8)/2 = 7.5
0	9	9
0	10	10






```
dat <- data.frame(
  sex = c(1, 1, 1, 1, 0, 0, 1, 0, 0, 0),
  rank = c(1, 2.5, 2.5, 4, 5.5, 5.5, 7.5, 7.5, 9, 10))

rank(c(161, 167, 272, 170, 176, 182, 182, 167,
176, 193), ties.method = "average")
```

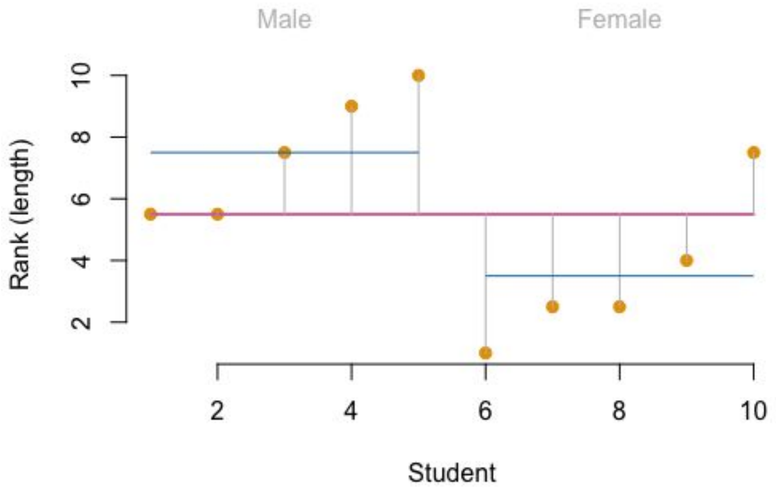
Ranking

interval → Length (cm)

ordinal → Ordered by length

Sex (f=1)	Length (cm)
1	161
1	167
0	272
1	170
0	176
1	182
0	182
1	167
0	176
0	193

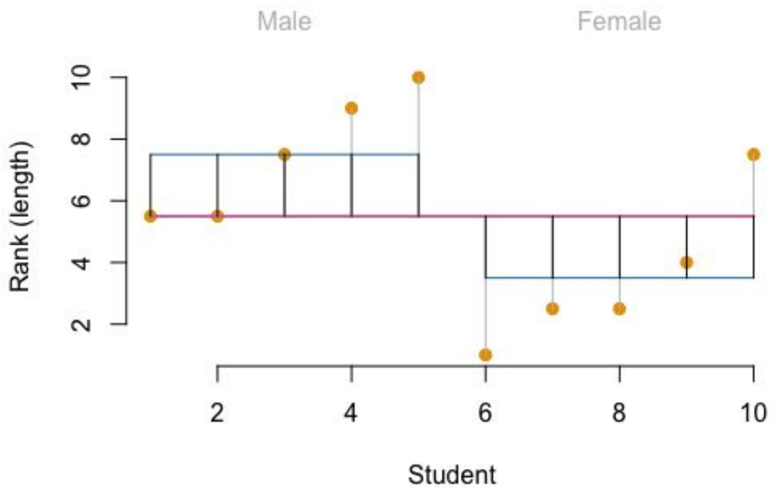
Ordered by length	Ranked length	Ranked length /w ties
1	1	1
1	2	(2+3)/2 = 2.5
1	3	(2+3)/2 = 2.5
1	4	4
0	5	(5+6)/2 = 5.5
0	6	(5+6)/2 = 5.5
1	7	(7+8)/2 = 7.5
0	8	(7+8)/2 = 7.5
0	9	9
0	10	10



Kruskall–Wallis test *(one-way ANOVA on ranks)*

$$H = (N-1)\frac{\sum_{i=1}^{g} n_i(\bar{r}_{i\cdot} - \bar{r})^2}{\sum_{i=1}^{g}\sum_{j=1}^{n_i}(r_{ij} - \bar{r})^2}$$

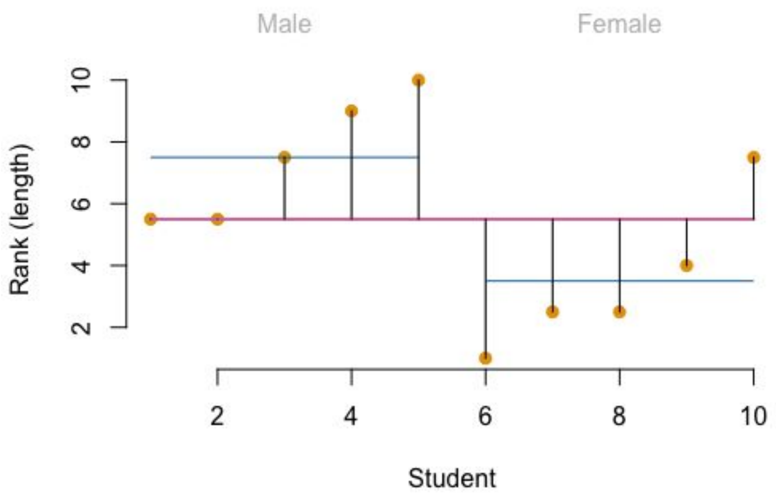
- total number of observations across all groups
- the number of groups
- the number of observations in group *i*
- the rank (among all observations) of observation *j* from group *i*
- the average rank of all observations in group *i*
- the average of all the r_{ij}



Kruskall–Wallis test *(one-way ANOVA on ranks)*

$$H = (N - 1)\frac{\sum_{i=1}^{g} n_i(\bar{r}_{i\cdot} - \bar{r})^2}{\sum_{i=1}^{g}\sum_{j=1}^{n_i}(r_{ij} - \bar{r})^2}$$

- total number of observations across all groups
- the number of groups
- the number of observations in group *i*
- the rank (among all observations) of observation *j* from group *i*
- the average rank of all observations in group *i*
- the average of all the r_{ij}



Kruskall–Wallis test *(one-way ANOVA on ranks)*

$$H = (N-1)\frac{\sum_{i=1}^{g} n_i(\bar{r}_{i\cdot} - \bar{r})^2}{\sum_{i=1}^{g}\sum_{j=1}^{n_i}(r_{ij} - \bar{r})^2}$$

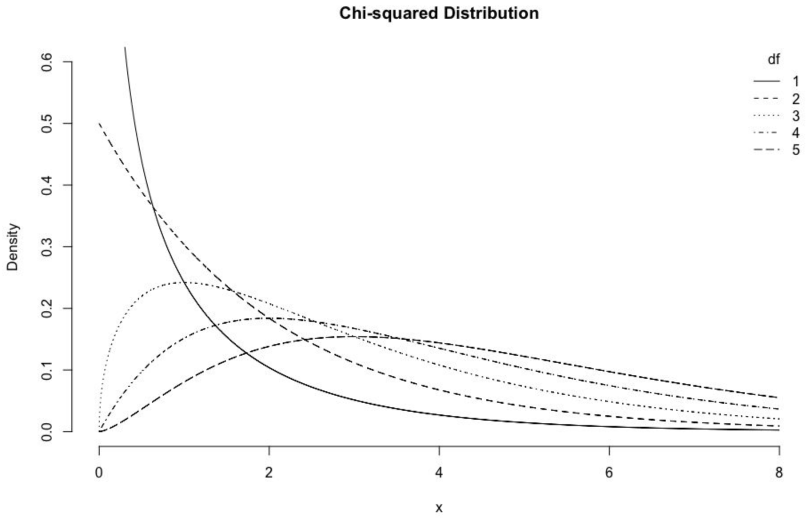
- total number of observations across all groups
- the number of groups
- the number of observations in group *i*
- the rank (among all observations) of observation *j* from group *i*
- the average rank of all observations in group *i*
- the average of all the r_{ij}

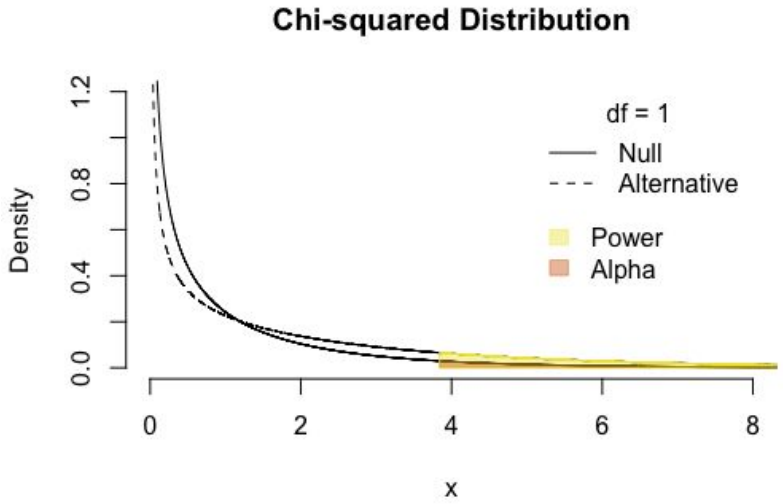
🧑‍💻

```
N <- nrow(dat)
g <- length(unique(dat$sex))
n_i <- aggregate(rank ~ sex,
                 data = dat, length)$rank
r_ij <- dat$rank
r_mean_i <- aggregate(rank ~ sex,
                      data = dat, mean)$rank
r_mean <- mean(dat$rank)

H <- (N - 1) *
  (sum(n_i * (r_mean_i - r_mean)^2) /
    sum((r_ij - r_mean)^2))
df <- g - 1
```

Chi-squared distributed (approx.)

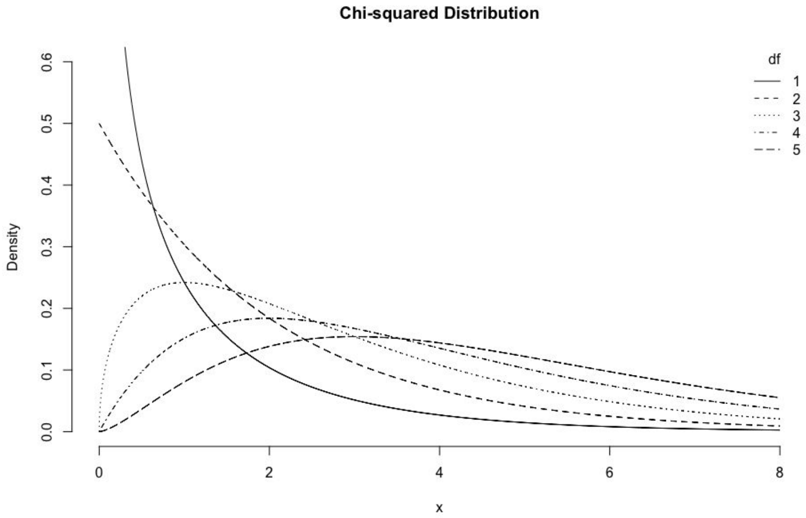




H = 4.44

df = 1

Chi-squared distributed (approx.)



H = 4.44

df = 1

🧑‍💻

pchisq(q = H, df = df, lower.tail = FALSE)

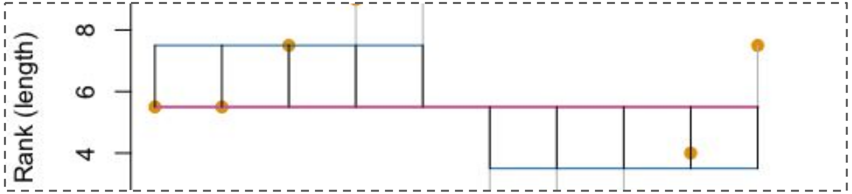
🦥

kruskal.test(rank ~ sex, data = dat)

kruskal.test(length ~ sex, data = dat) # if you have the original length data

lm(rank ~ sex, data = dat) # or model it as a linear regression

Why *H*?



$H =$ [redacted] $\dfrac{\sum_{i=1}^{g} n_i(\bar{r}_{i\cdot} - \bar{r})^2}{\sum_{i=1}^{g} \sum_{j=1}^{n_i}(r_{ij} - \bar{r})^2}$

"Effect size"

$H = (N - 1)$ [redacted]

"Power"

Cooling down

What did we learn?

Topics

Statistical reasoning
Empirical cycle
Probability distributions
Frequentist inference
Sample / sampling distribution
Central limit theorem
Normal distribution
P-value
Type I/II errors
Effect size
Confidence interval
Power
Test statistics
Linear regression
t-Test
Moderation
F-distribution
Nonparametric inference
ANOVA
Bayesian inference

Illustration by [Jennifer Cheuk]

Take-home assignments

📆 Weekly assignment

Q3 what's the *most likely option* given the CI's?
Q4/5/6 Shapiro–Wilk not previously discussed

🍻 Pub quiz

Create an *informative* four-choice question about the content of today's lecture.

An informative question has a large spread in responses across answer options.

Clarify answer options (which are (in)correct and why).

Illustration adapted from Snippets.com

Colophon

Slides
alexandersavi.nl/teaching/

License