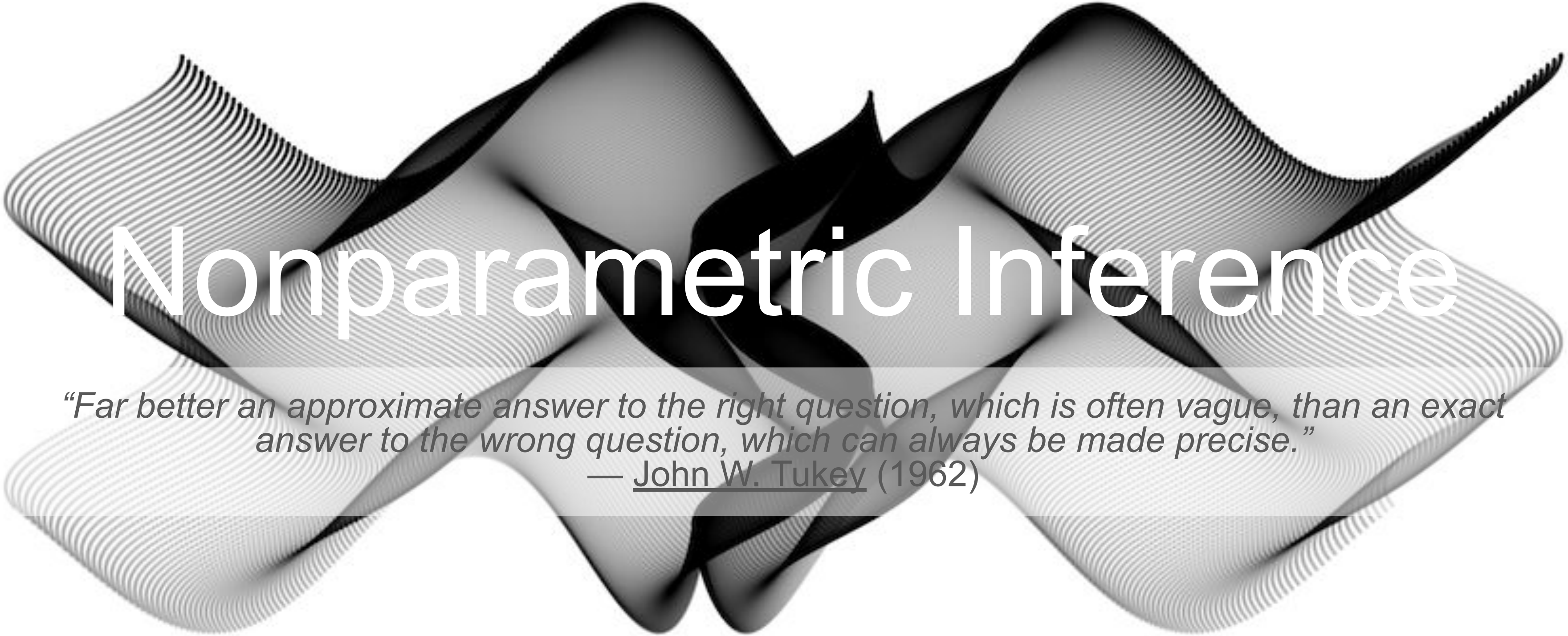
Nonparametric Inference



“Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise.”
— John W. Tukey (1962)

👨‍🏫 Statistical Reasoning Lecture #5
Alexander Savi, 2024
📚 Whitlock/Schluter, Ch. 13, 15.2, 16.3, 17.5, 18.5

Untitled by Katharina Brunner (generativeart package) 🎨

📢 Announcements

- Lecture 7?

Warming Up

Untitled by Katharina Brunner (generativeart package) 🎨

📰 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)
(free subscription)

Geen zin om een scriptie te schrijven? In Indonesië huur je dan een joki in

Fraude In Indonesië wordt volop gefraudeerd aan universiteiten. Status is belangrijker dan kennis. Ghostwriters schijven scripties en proefschriften. „Hier leest niemand een boek."

— NRC (Sep. 18, 2024)

🔥 Data Colada (response post)

🔁 Recap

Come up with a question.

Construct your hypotheses.

Operationalize your hypotheses (for instance into an experimental design).

Formulate your expectations.

Collect data.

Run the analyses.

Interpret the results.

```
model <- dep_var ~ indep_vars # specify your model
fit_regression <- lm(formula = model, data = ...) # fit a regression model
fit_anova <- aov(fit_regression) # or an ANOVA
?plot.lm # check assumptions
```

🔮 Overview

Topics

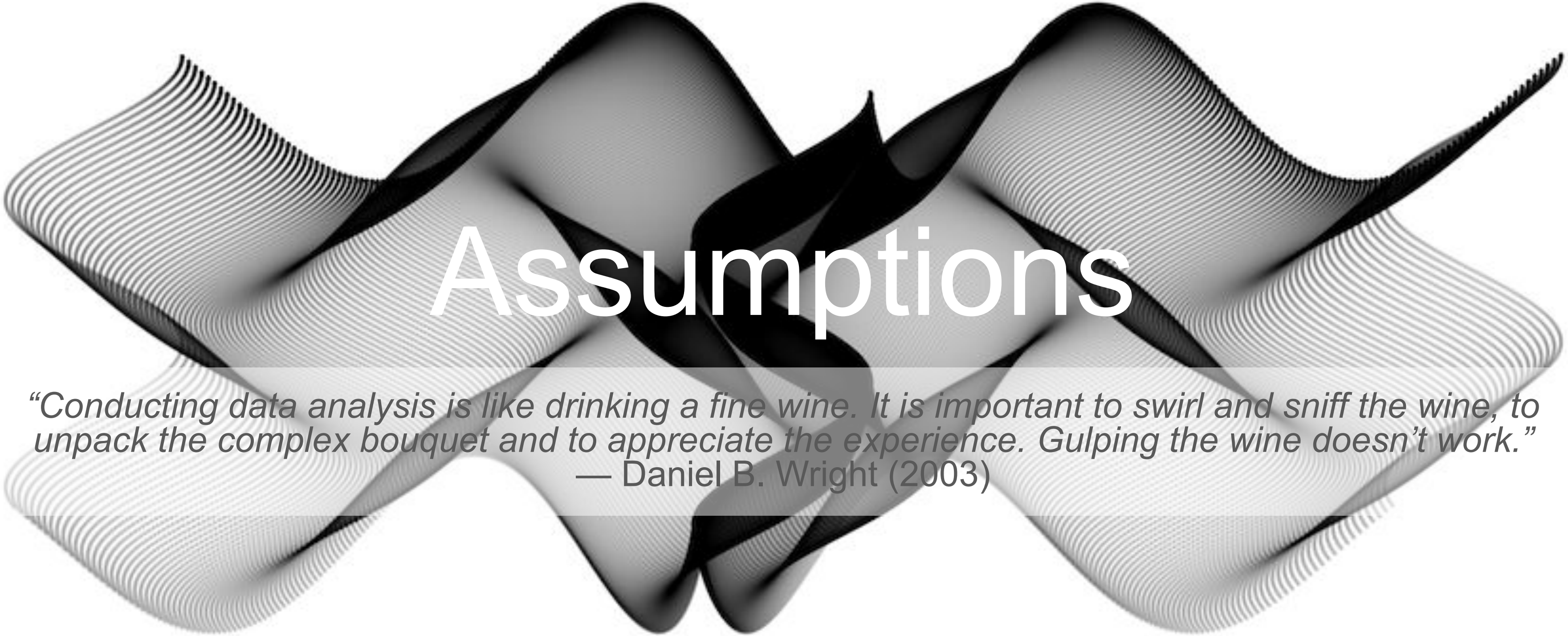
Probabilities & distributions
Frequentist inference
Multiple linear regression
Factorial ANOVA
| Assumptions
| Nonparametric inference
Bayesian inference

Learning goals

Check assumptions of statistical models.

Perfom nonparametric tests if assumptions do not hold.

🎁 Brand new probability distribution and brand new test statistic!



Assumptions

"Conducting data analysis is like drinking a fine wine. It is important to swirl and sniff the wine, to unpack the complex bouquet and to appreciate the experience. Gulping the wine doesn't work."
— Daniel B. Wright (2003)

Untitled by Katharina Brunner (generativeart package) 🎨

Validity

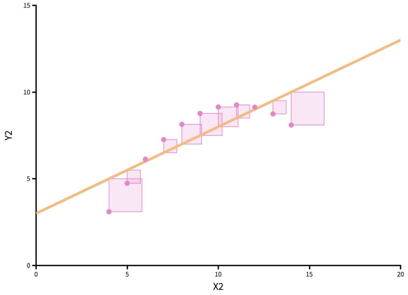
- Outcome reflects measure of interest?
- Outcome is representative?
- Inputs are relevant and necessary?
- Sample represents population of interest?

`predict(fit)` # external validity

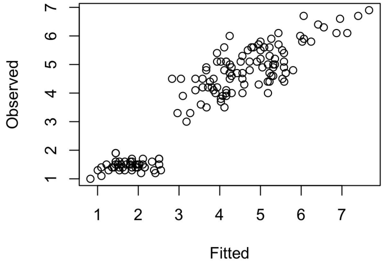
⚠️ Important

💡 Prioritization taken from Andrew Gelman.

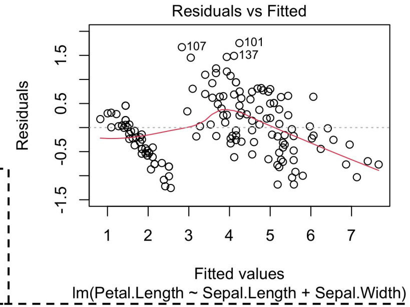
Additivity & linearity




```
plot(x = fitted(fit), y =
iris$Petal.Length)
```




```
plot(x = fit, which = 1)
```



`car::residualPlots(model = fit)` # for each predictor

👍 Approximately linear; horizontal at 0

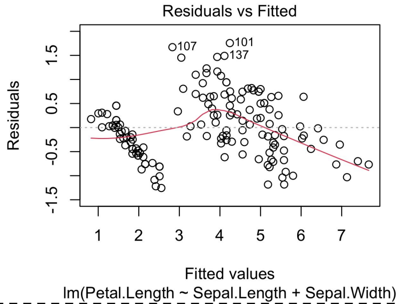
⚠️ Important

Independence of errors

Multilevel models, time-series models, …

👍 Horizontal
⚠️ Important

```
plot(x = fit, which = 1)
```

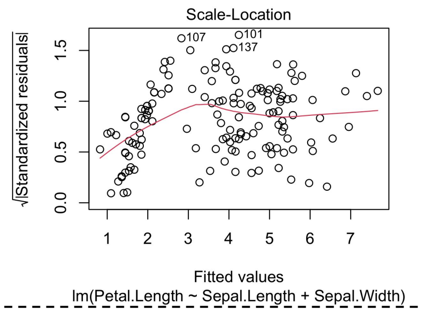


Equal variance of errors

Homogeneity of variance, homoscedasticity.
Similar to sphericity in repeated measures ANOVA.

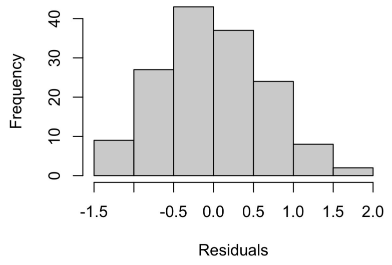
👍 Horizontal
⚠️ Issue with prediction, otherwise minor

```
plot(x = fit, which = 3)
```

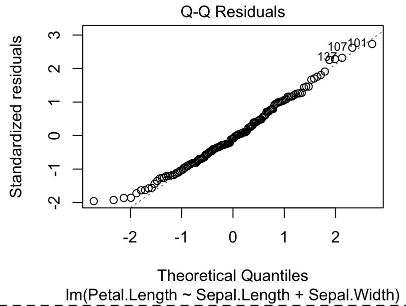


Normality of errors

```
hist(x = resid(fit))
```




```
plot(x = fit, which = 2)
```



👍 Approximately normal; linear
⚠️ Issue with predicting individual data points, otherwise not an issue

(Multi)collinearity

```
cor(iris[, c("Sepal.Length",
"Sepal.Width")])
```

```
car::vif(mod = fit)
```

👍 Low correlations between predictors; low VIF

⚠️ For explanation, less/not for prediction

Influential observations

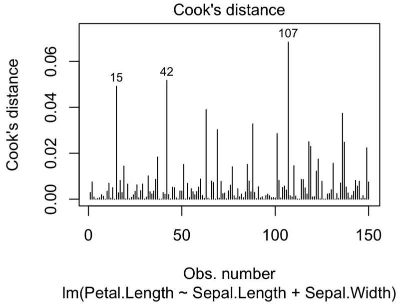
Outlier, leverage, influential

Error, interesting, random ([Leys et al., 2019](#))

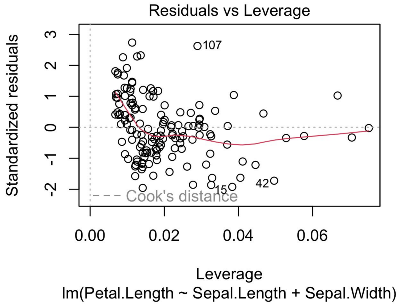
👍 Cook's distance < 1 or < 4/N; horizontal at 0

⚠️ It depends

```
plot(x = fit, which = 4)
```

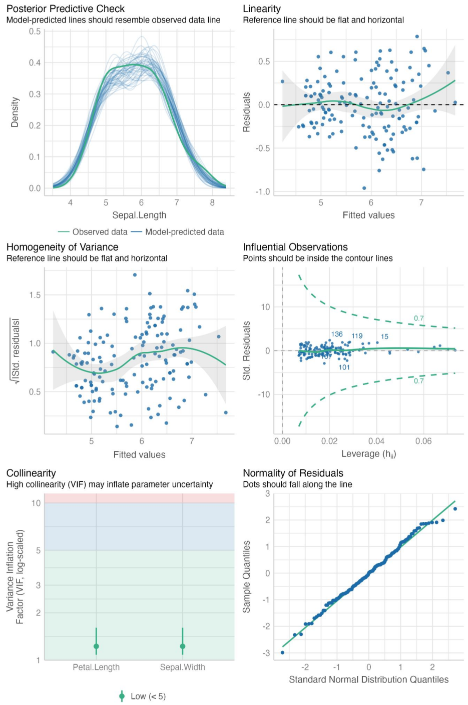



```
plot(x = fit, which = 5)
```



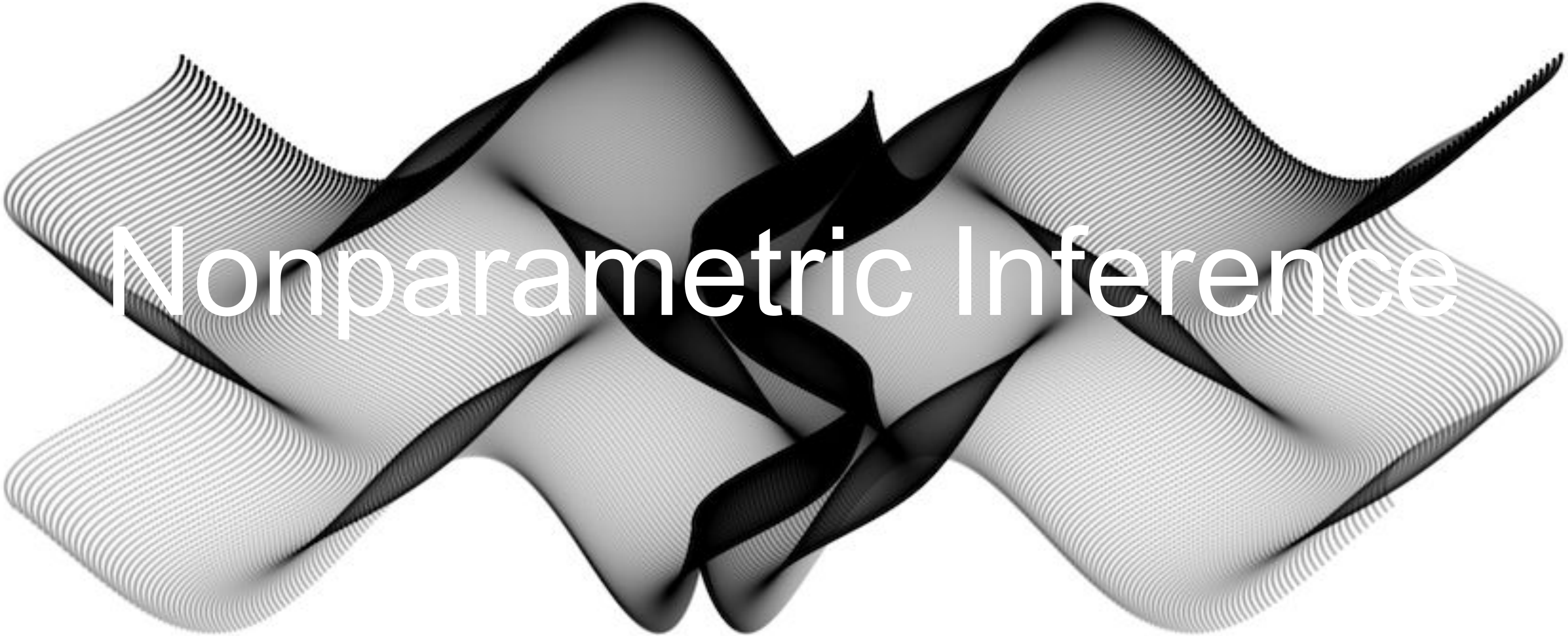
Take it easy

```
library("easystats")
performance::check_model(fit)
```



💡 [Interpretations and solutions]()

Nonparametric Inference



Untitled by Katharina Brunner (generativeart package) 🎨

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

But… we forgot to measure their actual length!

Let me handle that. Do you have a picture?

I do!

Great, *Nonparametric Ninja* at your service.

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: 🕔📏⚖️ 🌡️ (K)

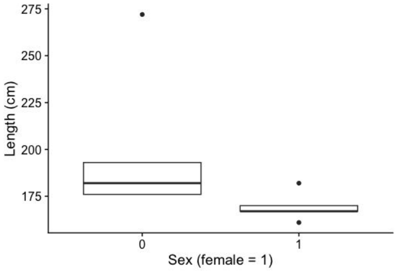
Ranking

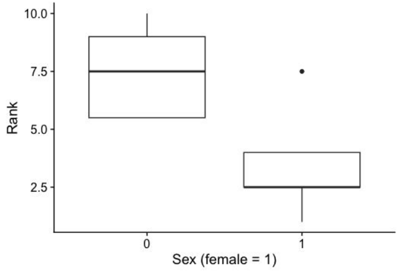
interval

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





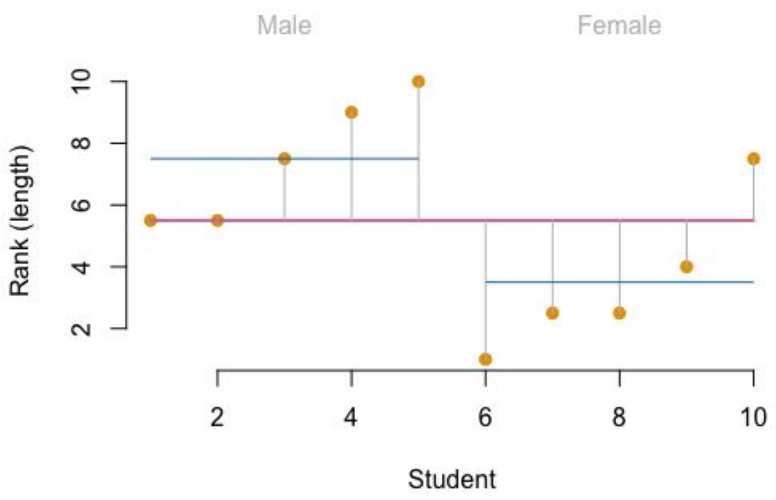

```
dat <- data.frame(
  sex = c(1, 1, 1, 1, 0, 0, 1, 0, 0, 0),
  rank = c(1, 2.5, 2.5, 4, 5.5, 5.5, 7.5,
7.5, 9, 10))

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

Ranking

interval 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

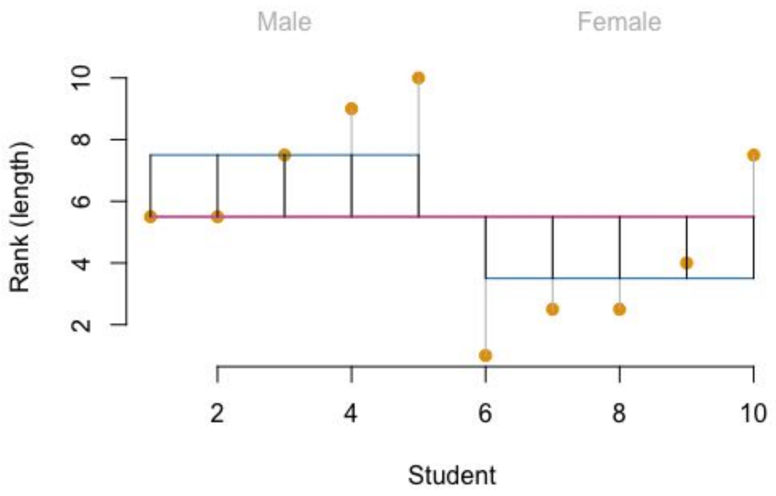
Ordered by length	Ranked length	Ranked length /w ties
1	1	1
1	2	(2+3)/2 = 2.5
1	3	(2+3)/2 = 2.5
1	4	4
0	5	(5+6)/2 = 5.5
0	6	(5+6)/2 = 5.5
1	7	(7+8)/2 = 7.5
0	8	(7+8)/2 = 7.5
0	9	9
0	10	10



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

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

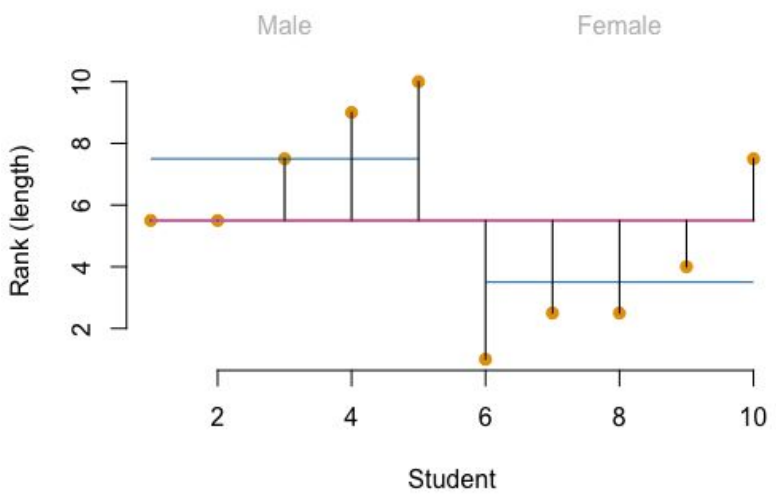
- total number of observations across all groups
- the number of groups
- the number of observations in group *i*
- the rank (among all observations) of observation *j* from group *i*
- the average rank of all observations in group *i*
- the average of all the r_{ij}



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}



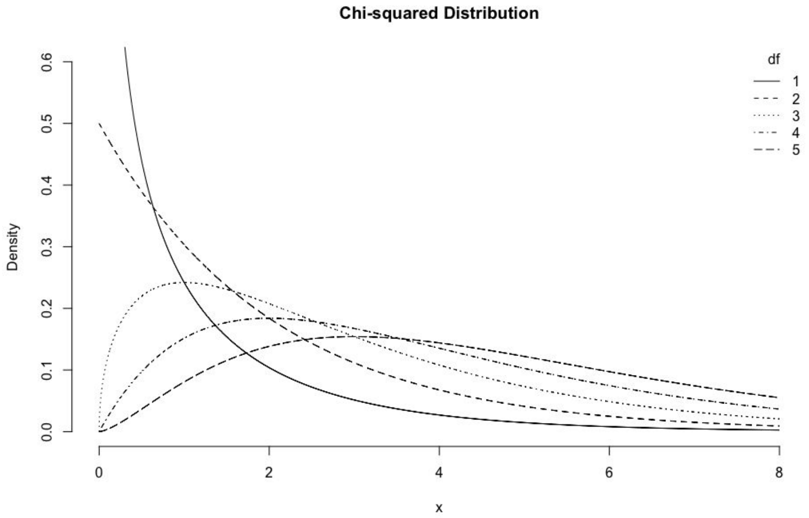
Kruskall–Wallis test *(one-way ANOVA on ranks)*

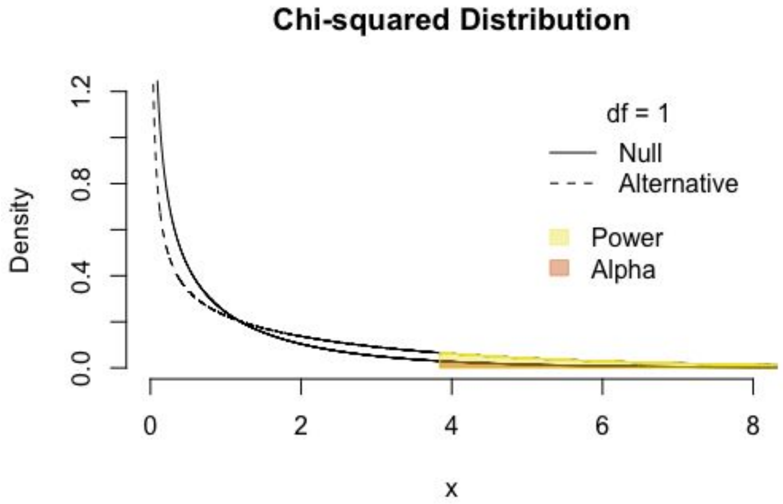
$$H = (N-1)\frac{\sum_{i=1}^{g} n_i(\bar{r}_{i\cdot} - \bar{r})^2}{\sum_{i=1}^{g}\sum_{j=1}^{n_i}(r_{ij} - \bar{r})^2}$$

- 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.)

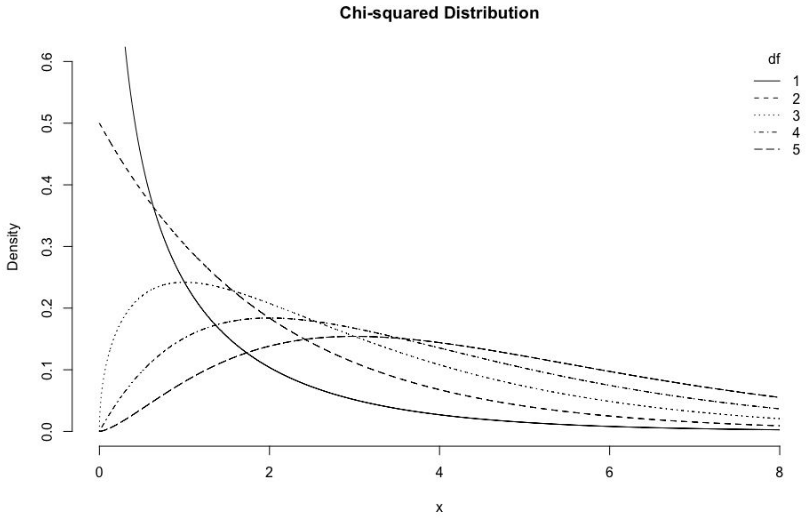




H = 4.44

df = 1

Chi-squared distributed (approx.)

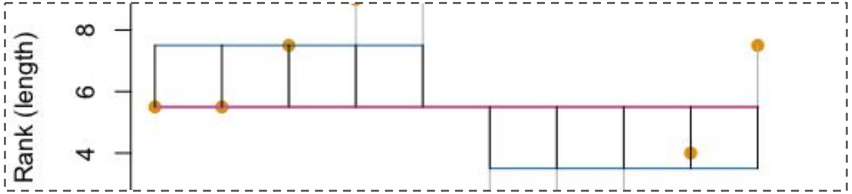


H = 4.44

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
```

Why *H*?



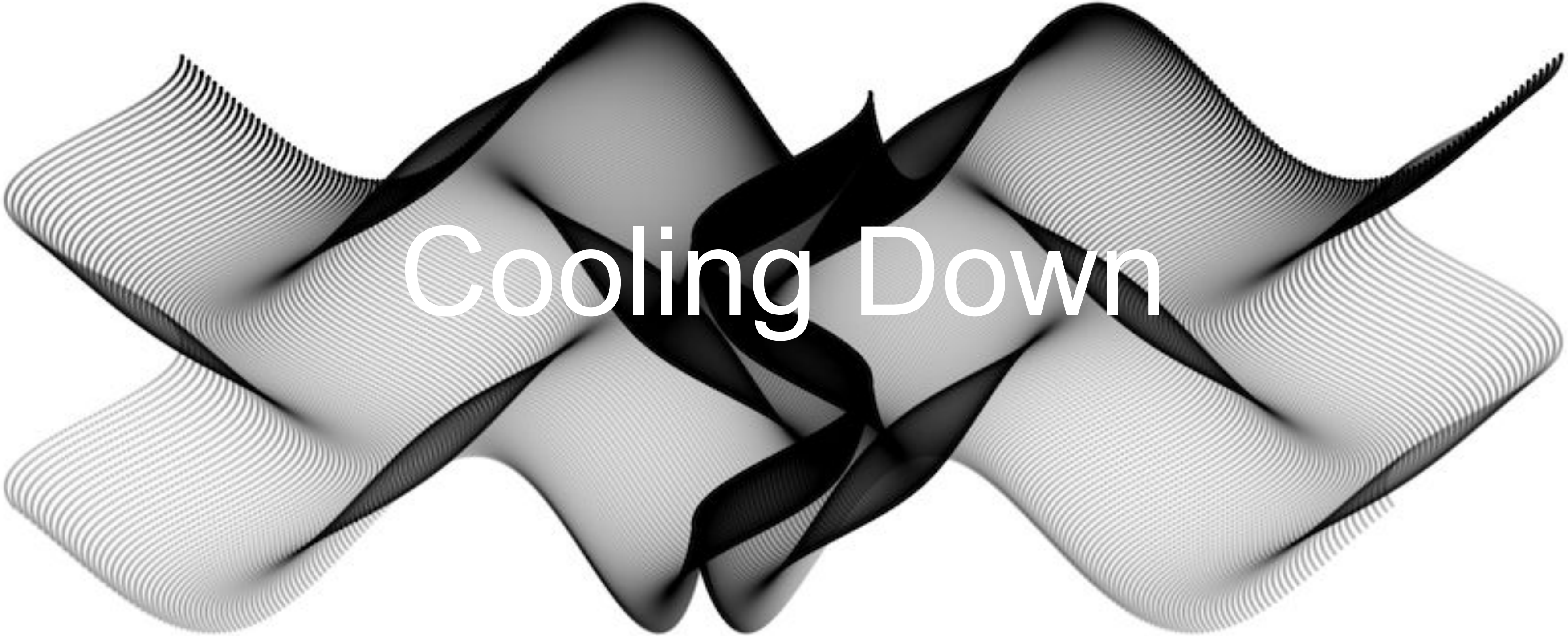
$$H = \text{[redacted]} \frac{\sum_{i=1}^{g} n_i(\bar{r}_{i\cdot} - \bar{r})^2}{\sum_{i=1}^{g}\sum_{j=1}^{n_i}(r_{ij} - \bar{r})^2}$$

"Effect size"

$$H = (N-1)\,\text{[redacted]}$$

"Power"

Cooling Down



Untitled by Katharina Brunner (generativeart package) 🎨

Takeaways

Illustration by [Amii Illustrates]

Takeaways

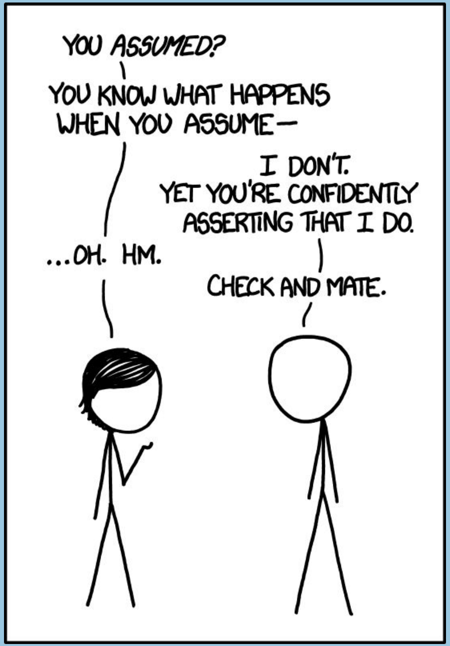


Illustration by Randall Munroe (wtf)

Takeaways

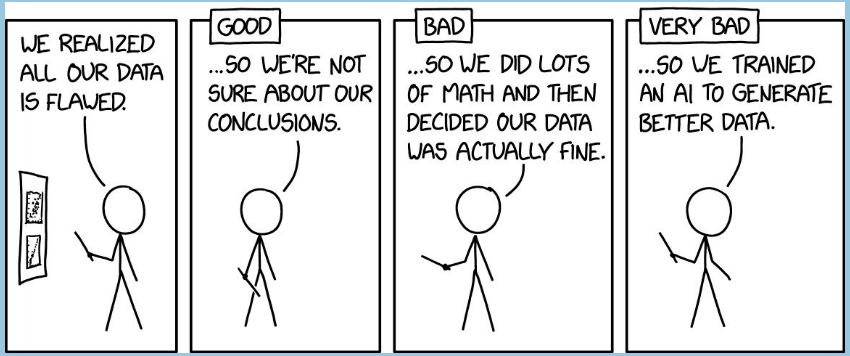


Illustration by [Randall Munroe](#) ([wtf](#))

Takeaways

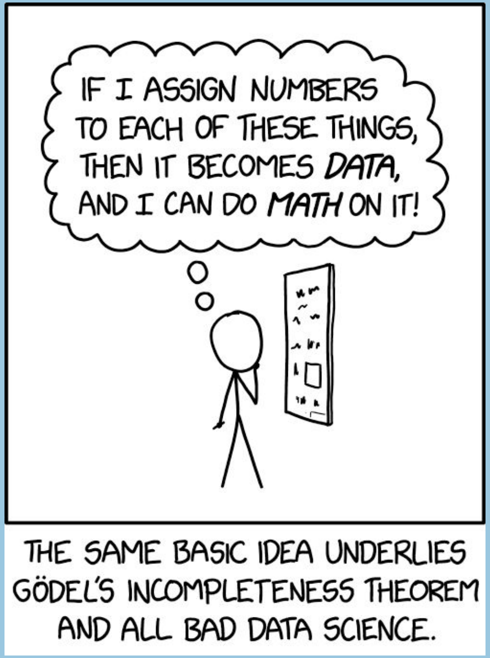


Illustration by [Randall Munroe](#) ([wtf](#))

Takeaways

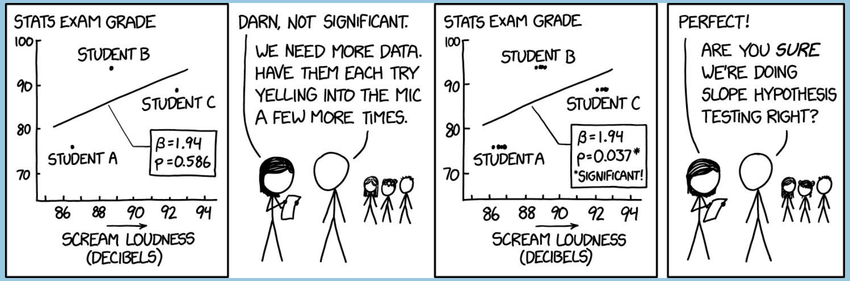


Illustration by [Randall Munroe](#) ([wtf](#))

Takeaways

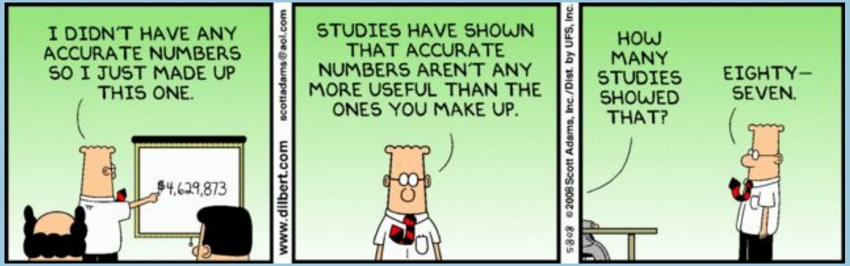


Illustration by [Scott Adams]

🥇 Exam(ple) question

In welke van de onderstaande gevallen is het verstandig om een nonparametrisch alternatief voor je toets te kiezen?

☐ Er zijn grote verschillen in het aantal deelnemers in de verschillende condities die je met elkaar wilt vergelijken.

☐ De data lijken op het oog normaal verdeeld, maar uit de Shapiro–Wilk toets blijkt dat toch niet het geval te zijn.

☐ Je hebt weinig data en zoekt een toets die veel power heeft om de verwachte relatie aan te tonen.

☐ De afhankelijke variabele is gemeten op een ordinale schaal.

🧑‍💻 Take-home assignments

📆 Weekly assignment

🍻 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

🔮 Overview

Topics

Probabilities & distributions
Frequentist inference
Multiple linear regression
Factorial ANOVA
Nonparametric inference
| Bayesian inference

Illustration by [Jennifer Cheuk]

🗃️ Look here!

Assumptions

Normal variation of the QQ plot (Yihui Xie).

⚠️ Don't look here!

Replicate Yihui Xie's animation of the normal variation of the QQ plot (see the 'Look here!' section). Just simulate a small collection of plots, it doesn't have to be animated (but it can be).

Additional challenge: pick another assumption and examine its normal variation.

Hints (select and copy/paste the invisible text below to reveal it)

0.

1.

2.

3.

✍️ Colophon

Slides

alexandersavi.nl/teaching/

License