Lab 5

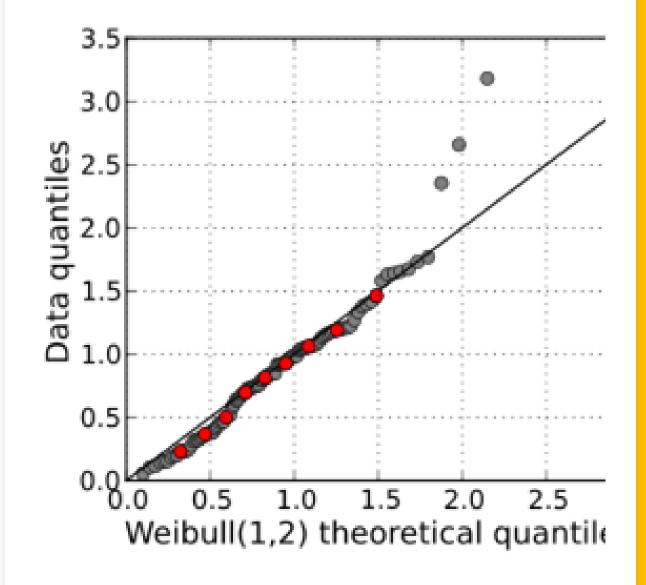
Lab G5

Cindy J. Pang

Wednesday, 05/14/2024, 1:00-1:50pm

Q-Q Plots (Quantile-Quantile)

- Graphical tool to help assess if a set of data comes from some theoretical distribution
- Plot Quantiles against each other
- Relies on Rank Statistics (or ordering the values from lowest to highest first)

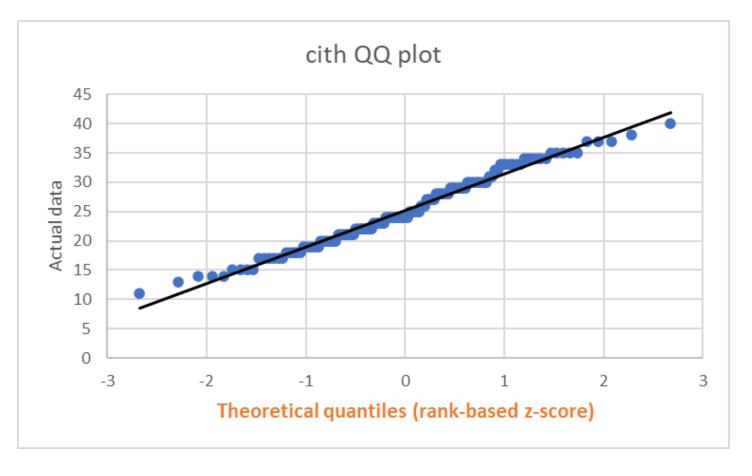


Constructing a Q-Q Plot (X-axis values)

cith		rank	percentile	rank-based z-score	
	11	1	0.0037313	-2.675459606	11
	13	2	0.011194	-2.283718847	1/3
	14	3	0.0186567	-2.082318	14
	14	4	0.0261194	-1.941160555	14
	14	5	0.0335821	-1.83057382	14
	15	6	0.0410448	-1.738688601	15
	15	7	0.0485075	-1.659500771	15
	15	8	0.0559701	-1.589532164	15
	15	9	0.0634328	-1.526579244	15
	17	10	0.0708955	-1.469153923	17
	17	11	0.0783582	-1.416201838	17
	17	12	0.0858209	-1.366947407	17
	17	13	0.0932836	-1.320802685	17
	17	14	0.1007463	-1.277310814	17
	17	15	0.108209	-1.236109375	17
	18	16	0.1156716	-1.196905773	18
	18	17	0.1231343	-1.159460185	18
	18	18	0.130597	-1.123573438	18
	18	19	0.1380597	-1.089078199	18
	18	20	0.1455224	-1.055832437	18
	19	21	0.1529851	-1.02371449	19
	19	22	0.1604478	-0.992619286	19

Come from a normal distribution (theoretical distribution)

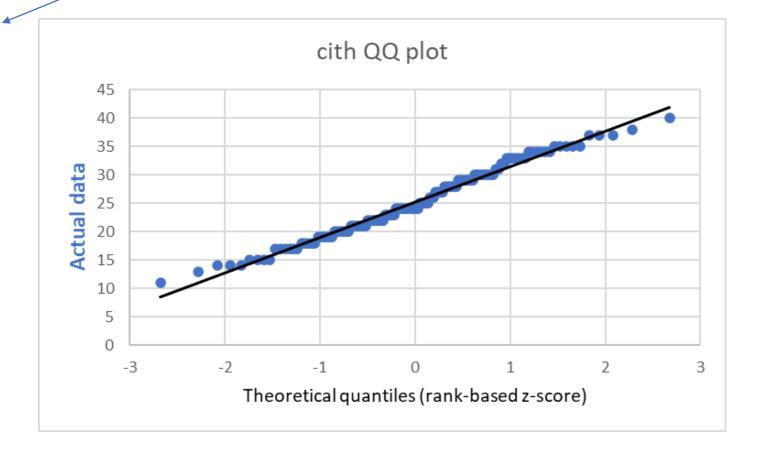
Calculated based on the percentiles (click on the values in Excel to see formula)



Constructing a Q-Q Plot (Y-axis values)

		. • •		1.1
	rank	•		cith
11	1	0.0037313	-2.675459606	11
13	2	0.011194	-2.283718847	13
14	3	0.0186567	-2.082318	14
14	4	0.0261194	-1.941160555	14
14	5	0.0335821	-1.83057382	14
15	6	0.0410448	-1.738688601	15
15	7	0.0485075	-1.659500771	15
15	8	0.0559701	-1.589532164	15
15	9	0.0634328	-1.526579244	15
17	10	0.0708955	-1.469153923	17
17	11	0.0783582	-1.416201838	17
17	12	0.0858209	-1.366947407	17
17	13	0.0932836	-1.320802685	17
17	14	0.1007463	-1.277310814	17
17	15	0.108209	-1.236109375	17
18	16	0.1156716	-1.196905773	18
				18
				18
	_			18
	_			18
	_			19
				19
	11 13 14 14 15 15 15 17 17 17 17 17 18 18 18 18	11 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	11 1 0.0037313 13 2 0.011194 14 3 0.0186567 14 4 0.0261194 14 5 0.0335821 15 6 0.0410448 15 7 0.0485075 15 8 0.0559701 15 9 0.0634328 17 10 0.0708955 17 11 0.0783582 17 12 0.0858209 17 13 0.0932836 17 14 0.1007463 17 15 0.108209 18 16 0.1156716 18 17 0.1231343 18 18 0.130597 18 20 0.1455224 19 21 0.1529851	11 1 0.0037313 -2.675459606 13 2 0.011194 -2.283718847 14 3 0.0186567 -2.082318 14 4 0.0261194 -1.941160555 14 5 0.0335821 -1.83057382 15 6 0.0410448 -1.738688601 15 7 0.0485075 -1.659500771 15 8 0.0559701 -1.589532164 15 9 0.0634328 -1.526579244 17 10 0.0708955 -1.469153923 17 11 0.0783582 -1.416201838 17 12 0.0858209 -1.366947407 17 13 0.0932836 -1.320802685 17 14 0.1007463 -1.277310814 17 15 0.108209 -1.236109375 18 16 0.1156716 -1.196905773 18 17 0.1231343 -1.159460185 18 19 0.1380597 -1.089078199 18 20 0.1455224 -1.055832437

YOUR DATA

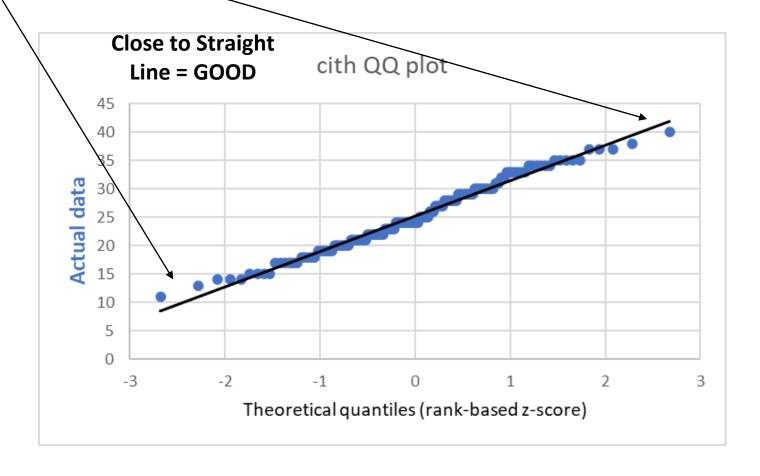


Is it any GOOD?

rank	percentile	rank-based z-score	cith	
1	0.0037313	-2.675459606		11
2	0.011194	-2.283718847		13
3	0.0186567	-2.082318		14
4	0.0261194	-1.941160555		14
5	0.0335821	-1.83057382		14
6	0.0410448	-1.738688601		15
7	0.0485075	-1.659500771		15
8	0.0559701	-1.589532164		15
9	0.0634328	-1.526579244		15
10	0.0708955	-1.469153923		17
11	0.0783582	-1.416201838		17
12	0.0858209	-1.366947407		17
13	0.0932836	-1.320802685		17
14	0.1007463	-1.277310814		17
15	0.108209	-1.236109375		17
16	0.1156716	-1.196905773		18
17	0.1231343	-1.159460185		18
18	0.130597	-1.123573438		18
19	0.1380597	-1.089078199		18
20	0.1455224	-1.055832437		18
21	0.1529851	-1.02371449		19
22	0.1604478	-0.992619286		19
	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21	1 0.0037313 2 0.011194 3 0.0186567 4 0.0261194 5 0.0335821 6 0.0410448 7 0.0485075 8 0.0559701 9 0.0634328 10 0.0708955 11 0.0783582 12 0.0858209 13 0.0932836 14 0.1007463 15 0.108209 16 0.1156716 17 0.1231343 18 0.130597 19 0.1380597 20 0.1455224 21 0.1529851	1 0.0037313 -2.675459606 2 0.011194 -2.283718847 3 0.0186567 -2.082318 4 0.0261194 -1.941160555 5 0.0335821 -1.83057382 6 0.0410448 -1.738688601 7 0.0485075 -1.659500771 8 0.0559701 -1.589532164 9 0.0634328 -1.526579244 10 0.0708955 -1.469153923 11 0.0783582 -1.416201838 12 0.0858209 -1.366947407 13 0.0932836 -1.320802685 14 0.1007463 -1.277310814 15 0.108209 -1.236109375 16 0.1156716 -1.196905773 17 0.1231343 -1.159460185 18 0.130597 -1.089078199 20 0.1455224 -1.055832437 21 0.1529851 -1.02371449	1 0.0037313 -2.675459606 2 0.011194 -2.283718847 3 0.0186567 -2.082318 4 0.0261194 -1.941160555 5 0.0335821 -1.83057382 6 0.0410448 -1.7386888601 7 0.0485075 -1.659500771 8 0.0559701 -1.589532164 9 0.0634328 -1.526579244 10 0.0708955 -1.469153923 11 0.0783582 -1.416201838 12 0.0858209 -1.366947407 13 0.0932836 -1.320802685 14 0.1007463 -1.277310814 15 0.108209 -1.236109375 16 0.1156716 -1.196905773 17 0.1231343 -1.159460185 18 0.130597 -1.089078199 20 0.1455224 -1.055832437 21 0.1529851 -1.02371449

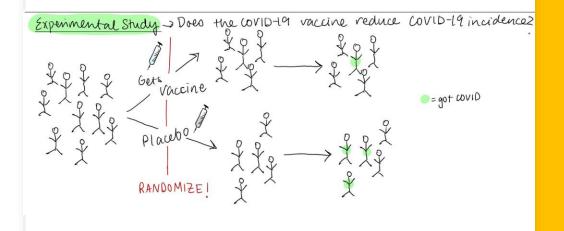
$R^2 \approx 1$ Means there is a STRONG Association

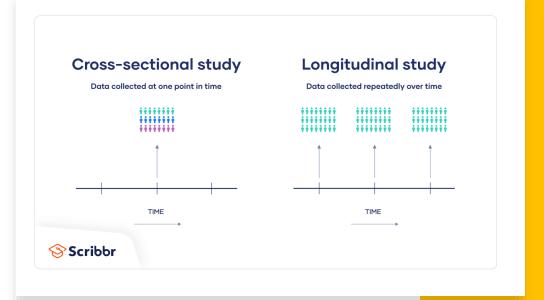
We will circle back to *r* later...



Study Types

- Experimental IS AN INTERVENTION with the subjects
 - Example: Randomized Control Trials (RCT)
- Longitudinal measures an outcome at multiple points in time to measure changes over time
- Cross-Sectional measures an outcome at one point in time





Meta-analysis

 "Meta-analysis is a research process used to systematically synthesize or merge the findings of single, independent studies, using statistical methods to calculate an overall or 'absolute' effect. Meta-analysis does not simply pool data from smaller studies to achieve a larger sample size."

Shorten A, Shorten B What is meta-analysis? Evidence-Based Nursing 2013; 16:3-4

Effect Size

• "an effect size is defined as a metric quantifying the relationship between two entities. It captures the direction and magnitude of this relationship. If relationships are expressed as the same effect size, it is possible to compare them."

• Examples:

- Correlation (Pearson's r)
- Mean Difference
- Risk Ratio
- Odds Ratio

Harrer, M., Cuijpers, P., Furukawa, T.A., & Ebert, D.D. (2021). *Doing Meta-Analysis with R: A Hands-On Guide*. Boca Raton, FL and London: Chapmann & Hall/CRC Press. ISBN 978-0-367-61007-4.

Pearson's Correlation Coefficient (r)

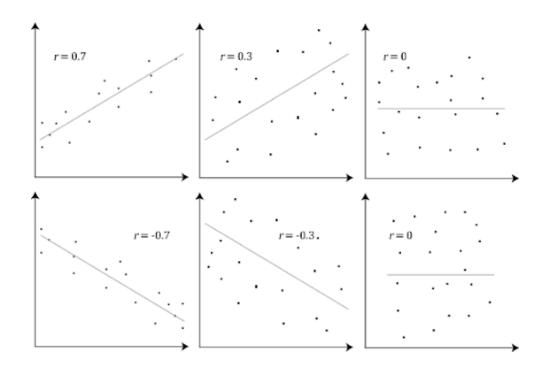
Measures the correlation between two variables:

$$r = Corr(x, y) = \frac{Cov(x, y)}{\sigma_x \sigma_y} = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum (x_i - \overline{x})^2 \sum (y_i - \overline{y})^2}}$$

Sometimes correlation can be abbreviated as $\rho(x,y)$ as well

Interpretation:

- r = 1: perfect positive correlation
- r = 0: no correlation
- r = -1: perfect negative correlation



^{*} There is also Spearman's Rank Coefficient Correlation

Cohen's d

"Cohen's d is used to describe the standardized mean difference of an effect. This value can be used to compare effects across studies, even when the dependent variables are measured in different ways"

$$d_S = \frac{\overline{X_1} - \overline{X_2}}{SD_{pooled}}$$

$$SD_{pooled} = \sqrt{\frac{(n_1 - 1)SD_1^2 + (n_2 - 1)SD_2^2}{n_1 + n_2 - 2}}$$

 n_i = sample size of group i where $i = \{1,2\}$

 $\overline{X_i}$ = mean of group *i*

Interpretation of Cohen's d:

 $d = 0.2 \rightarrow \text{Small Effect Size}$

 $d = 0.5 \rightarrow Medium Effect Size$

 $d = 0.8 \rightarrow \text{Large Effect Size}$

Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front Psychol. 2013 Nov 26;4:863. doi: 10.3389/fpsyg.2013.00863. PMID: 24324449; PMCID: PMC3840331.