

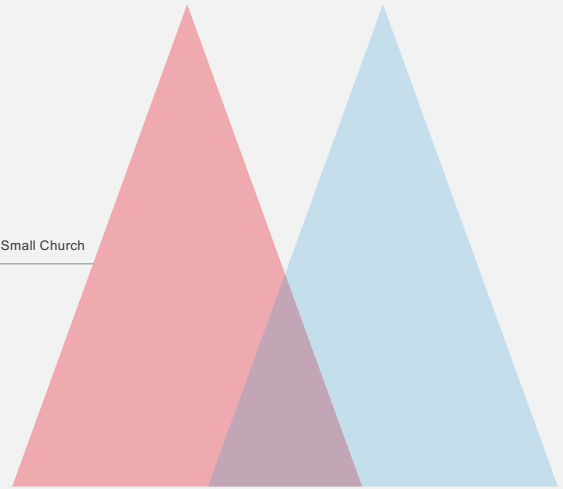
KIOSC

Korean Incubating Organization for Small Church

Various Statistical Analysis Method I

Noha Lim, Ph.D.

#KIOSC #TRAINING



1



Statistics in clinical trial

- For each clinical trial contributing to a marketing application, all important details of its design and conduct and the principal features of its proposed statistical analysis should be clearly specified in a protocol written before the trial begins. (mandate for late phase confirmatory trial)
 - Although the early phases of drug development consist mainly of clinical trials that are exploratory in nature, statistical principles are also relevant to these clinical trials.
- ICH E9 guidance

2



p-value < 0.05

- **Statistically significant**



Anthony S.
Fauci

National Institute of Allergy and Infectious Diseases Director **Anthony S. Fauci, MD**, said today that data from a multinational randomized control trial showed that Gilead's investigational antiviral remdesivir "has a clear-cut significant positive effect in diminishing time to recovery" for patients with COVID-19.

3



p-value < 0.05

- **Magic word in biomedical researches**
- **Statisticians warning against too much emphasis on this concept (ASA statement, 2016)**
- **More meaningful for well designed and controlled studies**
- **Still FDA standard for demonstrating "substantial evidence of effectiveness"**

4



What is p-value?

- p means “Probability”
- Inferential statistics
- Hypothesis ($H_0: A=B$; $H_a: A \neq B$) for population
- Statistical testing using random samples

	Reject H_0	Fail to reject H_0
H_0 is True	Type I error (α)	correct
H_a is True	correct	Type II error (β)

5



$H_0: A=B$

- Determine variable to be tested (i.e. blood pressure, achieve remission, etc.)
- Parameter for population A and B (i.e. mean, proportion, etc.)
- Different statistical tests depending on variable/parameter types

6



t-test

- Most commonly used tool to compare means
- Introduced by William Sealy Gosset in early 1900
- Also known as Student's t-test
- Gosset used pen name, Student in his first publication
- In this class, two sample t-test

7



t-test

- Continuous variable (i.e. height, blood pressure, etc.)
- Parameter for population is mean (i.e. mean of height, mean of blood pressure, etc.)
- Null hypothesis
 - Mean of blood pressure in group A (experimental arm) = Mean of blood pressure in group B (control arm)
 - $\mu_A = \mu_B \Rightarrow \mu_A - \mu_B = 0$

8



t-test

- Underlying assumption
 - Random sample
 - Independence between group A and B
 - Reasonably large sample size
 - Data follow normal distribution
 - Equal variance of A and B

9



t-test

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_p \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

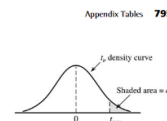
where

$$s_p = \sqrt{\frac{(n_1 - 1)s_{X_1}^2 + (n_2 - 1)s_{X_2}^2}{n_1 + n_2 - 2}}$$

and

Table A.5 Critical Values for t Distributions

df	α						
	.10	.05	.025	.01	.005	.001	.0005
1	3.078	6.314	12.706	31.821	63.657	318.31	636.62
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
32	1.309	1.694	2.037	2.449	2.738	3.365	3.622
34	1.307	1.691	2.032	2.441	2.728	3.348	3.601
36	1.306	1.688	2.028	2.434	2.719	3.333	3.582
38	1.304	1.686	2.024	2.429	2.712	3.319	3.566
40	1.303	1.684	2.021	2.425	2.704	3.307	3.551
50	1.299	1.676	2.009	2.403	2.678	3.262	3.496
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291



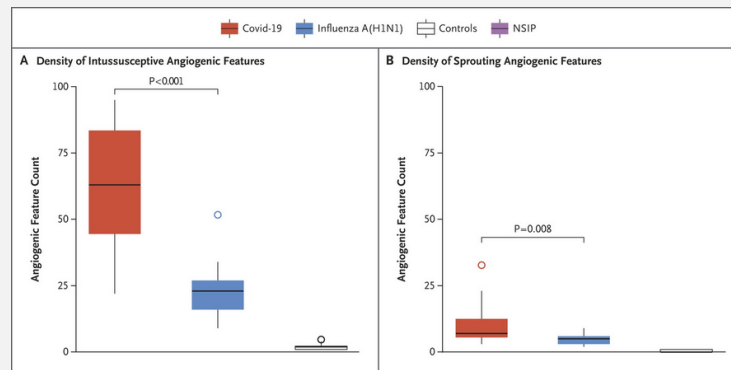
=> p-value

10



t-test

Pulmonary Vascular Endothelialitis, Thrombosis, and Angiogenesis in Covid-19 (Ackermann et al. 2020 in NEJM)



11



ANOVA

- Analysis of Variance
- Extension of t-test
- Compare means among three or more groups
- Developed by Ronald Fisher
- Use F-test
- In this class, one-way ANOVA

12



ANOVA

- Like t-test, continuous variable
- Parameter of interest is mean
- Null hypothesis
 - $\mu_A = \mu_B = \mu_C = \dots$
- Alternative hypothesis: Not all means are equal
- For example, different dose groups vs placebo; different combination of therapies vs placebo; etc.

13



ANOVA

- Underlying assumption
 - Random sample
 - Independence between group A and B and C
 - Reasonably large sample size
 - Data follow normal distribution
 - Equal variance of A and B and C

14



ANOVA

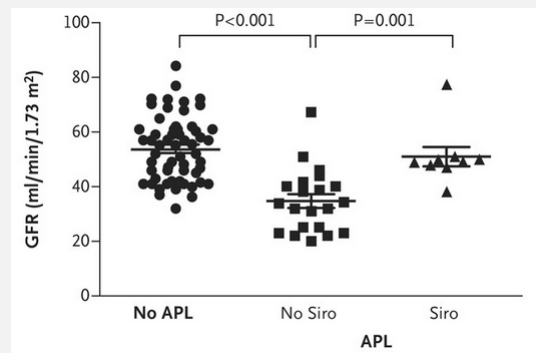
- F value: ratio of sum of squares (variance)
- Square of t distribution is same as F distribution
- Hence two sample t-test is basically equal to two group ANOVA

15



ANOVA

Inhibition of the mTORC Pathway in the Antiphospholipid Syndrome (Canaud et al. 2016 in NEJM)



16



ANCOVA

- Analysis of Covariance
- Combination of ANOVA and regression
- In comparison to one-way ANOVA
 - Still same continuous variable
 - But comparing adjusted mean
 - Adjust variable of interest by covariate using regression technique

17



ANCOVA

- “covariate” : variable expected to have an important influence on the primary variable
 - Baseline values
 - Age, Gender, etc.
- Pre-specification required
- Improve the power of significance tests and the precision of estimates of treatment effect

18

KIOSC

Thank You

Biostatistics for Non-statistician

- ☐ Basic I - Various Statistical Analysis Methods I
- ☐ Basic II - Various Statistical Analysis Methods II
- ☐ Advanced I - Study Design and Power & Sample Size
- ☐ Advanced II - Adaptive Trial Design