Practical exercise

Maria Beatriz Walter Costa
Viral Ecology and Omics (VEO)

Foto: Anne Günther

FRIEDRICH-SCHILLER-
UNIVERSITÄT
JENA

INSTITUTO FEDERAL
Goiano
Campus Rio Verde

Question: are the cultivation temperatures of two Bacterial groups different?

- **Null hypothesis:** cultivation temperatures of the two groups are <u>not</u> different
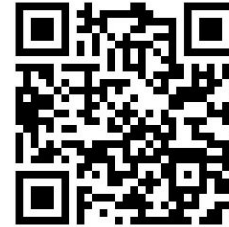
- **Samples: independent and unpaired**
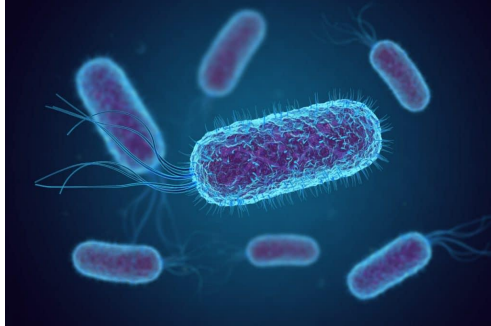
T-TEST
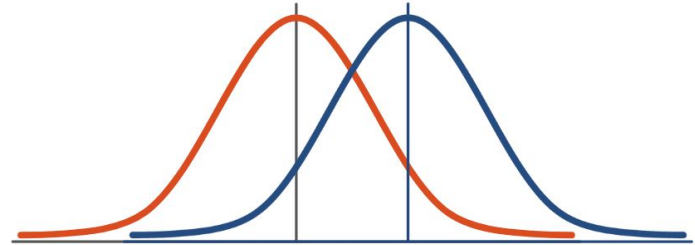
Outline of the practical exercise

- **Collect data**
  - BacDive: https://bacdive.dsmz.de/

- **Prepare data for t-test**
  - calculate: mean, standard deviation, sample size

- **Calculate t-value**

- **Compare t-value to critical value**

- **Interpret results**

- **Slides at:**
  https://github.com/waltercostamb/statistics

BacDive: cultivation metadata of temperature

BacDive

Statistics   API ® test finder

| Strain identifier | BacDive ID: 132485 | Type strain: ✔ | Species: Lysobacter tolerans | Strain Designation: UM1 | Culture col. no.: DSM 28473, KCTC 42936, MCC 2572 | Strain history: <- R. Lal, Univ. Delhi, Dept. Zoology; UM |

« Browse strain by BacDive ID »

Lysobacter tolerans UM1 is an aerobe, Gram-negative, rod-shaped bacterium that forms circular colonies and was isolated from soil from hexachlorocyclohexane-contaminated dumpsite.

colony-forming   Gram-negative   rod-shaped   aerobe   16S sequence   Bacteria   genome sequence

▼ ☰ **Name and taxonomic classification**

▼ 📈 **Morphology**

▼ ⚗ **Culture and growth conditions**

[Ref.: #43676]   **Culture medium**   LB (Luria-Bertani) MEDIUM

[Ref.: #43676]   **Culture medium growth**   ✔

Temperatures

Kind of temperature ⇕	Temperature ⇕
growth	25-40 °C
optimum	28 °C

[Ref.: #43676]
[Ref.: #43676]

pH

Kind of pH ⇕	pH ⇕
optimum	8
growth	3-10

[Ref.: #43676]
[Ref.: #43676]

▼ 📋 **Physiology and metabolism**

[Ref.: #43676]   **Oxygen tolerance**   aerobe

[Ref.: #43676]   **Ability of spore formation**   no

Halophily

Salt ⇕	Tested relation ⇕	Salt conc. ⇕

FRIEDRICH-SCHILLER-UNIVERSITÄT JENA

INSTITUTO FEDERAL Goiano Campus Rio Verde

Collect data

- **Go to BacDive**: https://bacdive.dsmz.de/

- **Collect the first reported temperature for the isolates:**

- **Group 1**
  - *Abyssibacter profundi* OUC007
  - *Abiotrophia defectiva* DSM 9849
  - *Acetivibrio clariflavus* EBR-02E-0045
  - *Acetivibrio thermocellus* DSM 1237

- **Group 2**
  - *Achromobacter denitrificans* 12
  - *Achromobacter insolitus* DSM 23807
  - *Acidovorax defluvii* DSM 12644

Collect data: solution

- **Group 1**
    - *Abyssibacter profundi* OUC007: 37 C
    - *Abiotrophia defectiva* DSM 9849: 37 C
    - *Acetivibrio clariflavus* EBR-02E-0045: 55 C
    - *Acetivibrio thermocellus* DSM 1237: 55 C

- **Group 2**
    - *Achromobacter denitrificans* 12: 30 C
    - *Achromobacter insolitus* DSM 23807: 28 C
    - *Acidovorax defluvii* DSM 12644: 30 C

Prepare data for t-test

- **Write down sample size for group 1 (n1)**
- **Write down sample size for group 2 (n2)**

- **Calculate means for group 1**
- **Calculate means for group 2**

- **Calculate standard deviation for group 1**
- **Calculate standard deviation for group 2**

- **Calculate the degrees of freedom (df)**
  - **df = n1 + n2 -2**

$$\text{Mean} = \frac{\text{Sum of All Data Points}}{\text{Number of Data Points}}$$

Fórmulas

$$\sigma = \sqrt{\frac{\sum(x_i - \mu)^2}{N}}$$

σ = desvio padrão da população

N = o tamanho da população

x_i = cada valor da população

μ = a média da população

FRIEDRICH-SCHILLER-
UNIVERSITÄT
JENA

INSTITUTO FEDERAL
Goiano
Campus Rio Verde

Prepare data for t-test: solution

- Write down sample size for group 1 (n1): **4**
- Write down sample size for group 2 (n2): **3**

- Calculate means for group 1: **46**
- Calculate means for group 2: **29.33**

- Calculate standard deviation for group 1: **9**
- Calculate standard deviation for group 2: **0.94**

- Calculate the degrees of freedom (df): **5**

Calculate t-value

$$\frac{|\overline{x}_1 - \overline{x}_2|}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

- **x1 = mean of group 1**
- **x2 = mean of group 2**

- **s1 = standard deviation of group 1**
- **s2 = standard deviation of group 2**

- **n1 = sample size of group 1**
- **n2 = sample size of group 2**

- **We will see the balance between <u>signal</u> and <u>noise</u>**

FRIEDRICH-SCHILLER-
UNIVERSITÄT
JENA

INSTITUTO FEDERAL
Goiano
Campus Rio Verde

YouTube Channel "Bozeman Science", Student's t-test,

9

Calculate t-value: solution: **3.7**

- x1 = mean of group 1 = **46**
- x2 = mean of group 2 = **29.33**

- s1 = standard deviation of group 1 = **9**
- s2 = standard deviation of group 2 = **0.94**

- n1 = sample size of group 1 = **4**
- n2 = sample size of group 2 = **3**

- **Conclusion: signal is larger than noise (t-value > 1)**

$$\frac{|\overline{x}_1 - \overline{x}_2|}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

FRIEDRICH-SCHILLER-
UNIVERSITÄT
JENA

INSTITUTO FEDERAL
Goiano
Campus Rio Verde

Compare t-value to critical value on t-table

- **We know, our signal is larger than noise**

- **Now we compare t-value to critical value to know if this is due to chance**

- **t-value < critical value**: we do not reject the null hypothesis
- **t-value > critical value**: we reject the null hypothesis

YouTube Channel "Bozeman Science", Student's t-test,

Critical value for df=5 and p-value = 0.05 on t-table: 2.571

df/α	0.9	0.5	0.3	0.2	0.1	0.05	0.02	0.01	0.001
1	0.158	1	2	3.078	6.314	12.706	31.821	64	637
2	0.142	0.816	1.386	1.886	2.92	4.303	6.965	10	31.598
3	0.137	0.765	1.25	1.638	2.353	3.182	4.541	5.841	12.929
4	0.134	0.741	1.19	1.533	2.132	2.776	3.747	4.604	8.61
5	0.132	0.727	1.156	1.476	2.015	2.571	3.365	4.032	6.869
6	0.131	0.718	1.134	1.44	1.943	2.447	3.143	3.707	5.959
7	0.13	0.711	1.119	1.415	1.895	2.365	2.998	3.499	5.408
8	0.13	0.706	1.108	1.397	1.86	2.306	2.896	3.355	5.041
9	0.129	0.703	1.1	1.383	1.833	2.263	2.821	3.25	4.781
10	0.129	0.7	1.093	1.372	1.812	2.228	2.764	3.169	4.587

FRIEDRICH-SCHILLER-UNIVERSITÄT JENA

INSTITUTO FEDERAL
Goiano
Campus Rio Verde

Compare t-value to critical value on t-table

- **We compared t-value = 3.7 to critical value = 2.571**

- **t-value > critical value**: we reject the null hypothesis

- **Conclusion: there is a statistical difference between temperatures of groups 1 and 2!**

- **Remember that: t-tests assume normal distributions of samples and similar variance**
  - We do not meet these criteria, so for a scientific experiment, we should do a non-parametric test instead!
  - We used the t-test today for learning purposes

df/α	0.9	0.5	0.3	0.2	0.1	0.05	0.02	0.01	0.001
1	0.158	1	2	3.078	6.314	12.706	31.821	64	637
2	0.142	0.816	1.386	1.886	2.92	4.303	6.965	10	31.598
3	0.137	0.765	1.25	1.638	2.353	3.182	4.541	5.841	12.929
4	0.134	0.741	1.19	1.533	2.132	2.776	3.747	4.604	8.61
5	0.132	0.727	1.156	1.476	2.015	2.571	3.365	4.032	6.869
6	0.131	0.718	1.134	1.44	1.943	2.447	3.143	3.707	5.959
7	0.13	0.711	1.119	1.415	1.895	2.365	2.998	3.499	5.408
8	0.13	0.706	1.108	1.397	1.86	2.306	2.896	3.355	5.041
9	0.129	0.703	1.1	1.383	1.833	2.263	2.821	3.25	4.781
10	0.129	0.7	1.093	1.372	1.812	2.228	2.764	3.169	4.587

Thank you!