Australian National University

Applied Statistics

Ameer Sultan
ameer.sultan@anu.edu.au

Tutorial 1

Consultation Timings

- About me.
- Consultations Hours:
  - – Wednesdays 12 noon -1 PM
    - → Room 3.35 CBE

Example Problem:
Yellowness of Tail Feathers of the Northern Flicker Birds

- "Odd" feathers are presumed to be regrown feathers after being lost
- In the population of all northern flicker birds, is there sufficient evidence to say that "odd" feathers are significantly different from the "typical" feathers in terms of yellowness?
- In the population of all northern flicker birds, how odd are the odd feathers compared to typical feathers in terms of yellowness?

Data Set for Yellowness Index in 16 Birds

Bird	Typical	Odd
A	-0.255	-0.324
B	-0.213	-0.185
C	-0.19	-0.299
D	-0.185	-0.144
E	-0.045	-0.027
F	-0.025	-0.039
G	-0.015	-0.264
H	0.003	-0.077
I	0.015	-0.017
J	0.02	-0.169
K	0.023	-0.096
L	0.04	-0.33
M	0.04	-0.346
N	0.05	-0.191
O	0.055	-0.128
P	0.058	-0.182

Describing the Data

Sample Mean: $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

For the example, let us solve for the mean difference of yellowness:

$$\bar{X} = \frac{1}{16}\sum_{i=1}^{16} X_i = \frac{1}{16}(2.194) = 0.137125$$

By this statistic, it can be said that the typical feathers are on the average 0.137125 units more yellow than odd feathers.

Describing the Data

Sample Variance:

$$S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right)$$

Sample Standard Deviation:

$$S_X = \sqrt{S_X^2} = \sqrt{\frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right)}$$

Describing the Data

For the example, let us solve for the variance and standard deviation of the difference of yellowness:

$$S_X^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right) = \frac{1}{16-1}\left(0.574 - 16 \times 0.137125^2\right)$$

$$= 0.01820985$$

$$S_X = \sqrt{0.01820985} \approx 0.134943877$$

By this statistic, it can be said that the difference in yellowness between typical and odd feathers might fluctuate on average by about 0.1349 units above or below the stated average difference of 0.137125 units.

Making Inferences

On answering the question "how odd...?"

Confidence Interval Estimation for θ, "Typical" Method:

Estimator of $\theta \mp$ (Distribution Quantile) $\times$ (SE of Estimator)

95% Interval Estimation for Population Mean, assuming Normality:

$$
\begin{aligned}
\bar{X} \mp t_{0.05/2,n-1}\frac{S_X}{\sqrt{n}} &= 0.137125 \mp t_{0.05/2,16-1} \times 0.03373596931 \\
&= 0.137125 \mp 2.131 \times 0.03373596931 \\
&= 0.137125 \mp 0.0718913506 \\
&= (0.0652336496, 0.2090163506)
\end{aligned}
$$

Making Inferences

95% Interval Estimation for Population Mean, assuming Normality:

$$\bar{X} \mp t_{0.05/2,n-1}\frac{S_X}{\sqrt{n}} = (0.0652336496, 0.2090163506)$$

There is 95% that the true population mean difference of yellowness between typical and odd feathers is between *0.0652 units and 0.2090 units*. Since zero is not inside the interval, there is evidence to support that the typical feathers are different from odd feathers in terms of yellowness. The typical feathers tend to be more yellow than odd feathers.

Making Inferences

On answering the question "is there significant difference...?"

Hypothesis Testing Procedure, "Typical" Method:

Testing $H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$

$$\text{Test Statistic} = \frac{\text{Estimator } - \theta_0}{(\text{SE of Estimator})}$$

Given a certain level of significance α, we reject H_0 if test statistic value $\in$ {rejection region}.

Steps For Hypothesis Testing

1. Stating the Alternative and Null Hypotheses and Selecting Alpha. The Alternate or research hypothesis (H1) states the relationship in which we are really interested and null hypothesis (H0) always contradicts the research hypothesis, usually stating that there is no difference between the population mean and some specified value.

2. Selecting the Sampling Distribution and Specifying the Test Statistic – Student's *t* distribution.

3. Making a Decision and Interpreting the Results – if $|t| > |t_c|$ reject null hypothesis.

Making Inferences

On answering the question "is there significant difference...?"

Hypothesis Testing Procedure: Paired T Test

Testing $H_0 : \mu_D = 0$ vs. $H_1 : \mu_D \neq 0$

Test Statistic $T = \dfrac{\bar{X} - 0}{\frac{s_D}{\sqrt{n}}} = \dfrac{0.137125}{0.03373596931} = 4.064652737$

at level of significance $\alpha = 0.05$, we reject H_0 if $T > 2.131$ or $T < -2.131$.

Since $4.0647 > 2.131$, we reject H_0. There is sufficient evidence to conclude that there is a difference in yellowness between typical and odd feathers for the population of northern flicker birds.

R Studio

- We will now replicate these results using R software. Let’s calculate the following on R:
  - Mean
  - Std. Dev
  - Variance
  - Confidence Intervals
  - Hypothesis Testing