Lecture 03a – t-Tests

ENVX2001 Applied Statistical Methods

Januar Harianto

Feb 2026

History

William Gosset (1908)

Figure 1: William Gosset

- Worked as a statistician at Guinness Brewery in Dublin
- Published under the pseudonym **"Student"** (company policy prevented use of real name)
- Developed the t-distribution and t-tests for **small sample sizes**
- Needed to compare barley varieties with limited data
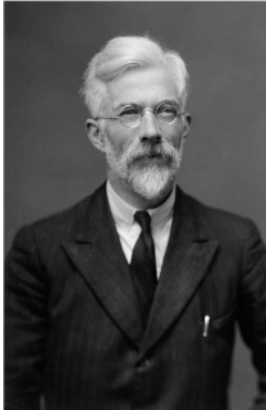
Ronald Fisher (1925)

Figure 2: Ronald Fisher

- Extended and formalised Gosset's work
- Introduced the concept of **degrees of freedom**
- Developed **Analysis of Variance (ANOVA)** – next lecture!
- Established many foundations of modern statistical inference

Cattle data

The problem

- Weights of two breeds of cattle are to be compared
- 12 samples were randomly taken from Breed 1, and 15 from Breed 2
- **Are there differences in the mean weight between the two breeds?**

```
cattle ← read.csv("data/cattle.csv")
cattle
```

```
   Breed1 Breed2
1   187.6  148.1
2   180.3  146.2
3   198.6  152.8
4   190.7  135.3
5   196.3  151.2
6   203.8  146.3
7   190.2  163.5
8   201.0  146.6
9   194.7  162.4
10  221.1  140.2
11  186.7  159.4
```

```
12  203.1  181.8
13     NA  165.1
14     NA  165.0
15     NA  141.6
```

Reshaping the data

The data is in **wide** format – we need to convert it to **long** format for analysis.

```r
cattle ← read.csv("data/cattle.csv") |>
  pivot_longer(
    cols = everything(),
    names_to = "breed",
    values_to = "weight"
  ) |>
  mutate(breed = as.factor(breed)) |>
  drop_na()

cattle
```
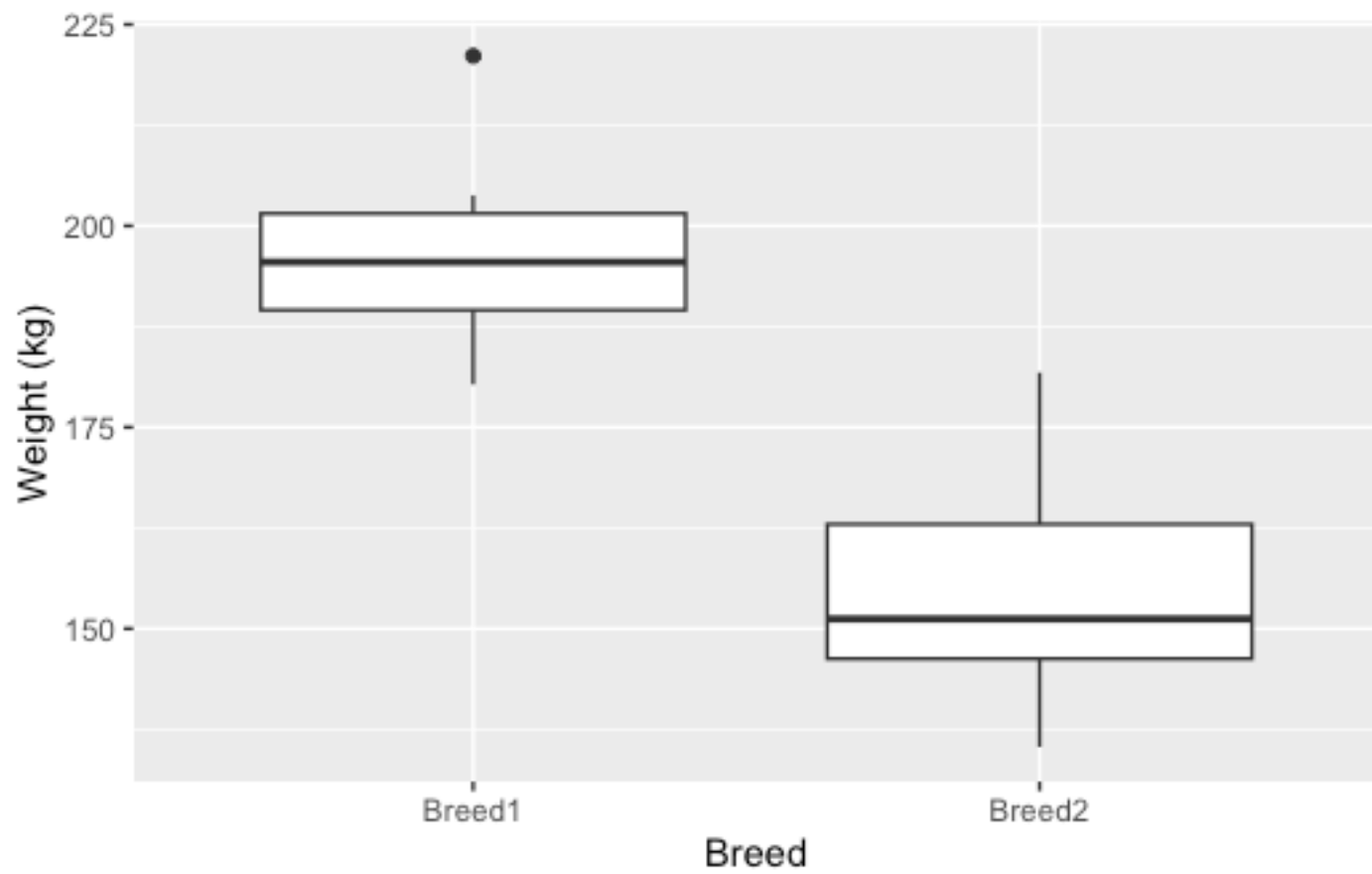
```
# A tibble: 27 × 2
   breed  weight
   <fct>   <dbl>
 1 Breed1   188.
 2 Breed2   148.
 3 Breed1   180.
```

```
 4 Breed2    146.
 5 Breed1    199.
 6 Breed2    153.
 7 Breed1    191.
 8 Breed2    135.
 9 Breed1    196.
10 Breed2    151.
# i 17 more rows
```

Visualising the data

```r
ggplot(cattle, aes(breed, weight)) +
  geom_boxplot() +
  labs(x = "Breed", y = "Weight (kg)")
```

Model and hypotheses

Two-sample t-test model

Observed data = Group Mean + Random Error (residuals)

$$y_{ij} = \mu_i + \varepsilon_{ij}$$

- $i = 1, 2$ (group); $j = 1, 2, ..., n_i$ (replicate)

In the cattle example:

- μ_1 = mean body weight (kg) for Breed 1
- μ_2 = mean body weight (kg) for Breed 2

Hypotheses

- **Null hypothesis**: $H_0 : \mu_1 = \mu_2$
- **Alternative hypothesis**: $H_1 : \mu_1 \neq \mu_2$

Test statistic:

$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{s^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} = \frac{\bar{y}_1 - \bar{y}_2}{SE(\bar{y}_1 - \bar{y}_2)} = \frac{\Delta \text{ in mean}}{SE \text{ of the } \Delta \text{ in mean}}$$

Degrees of freedom: $n_1 + n_2 - 2$

t-distribution
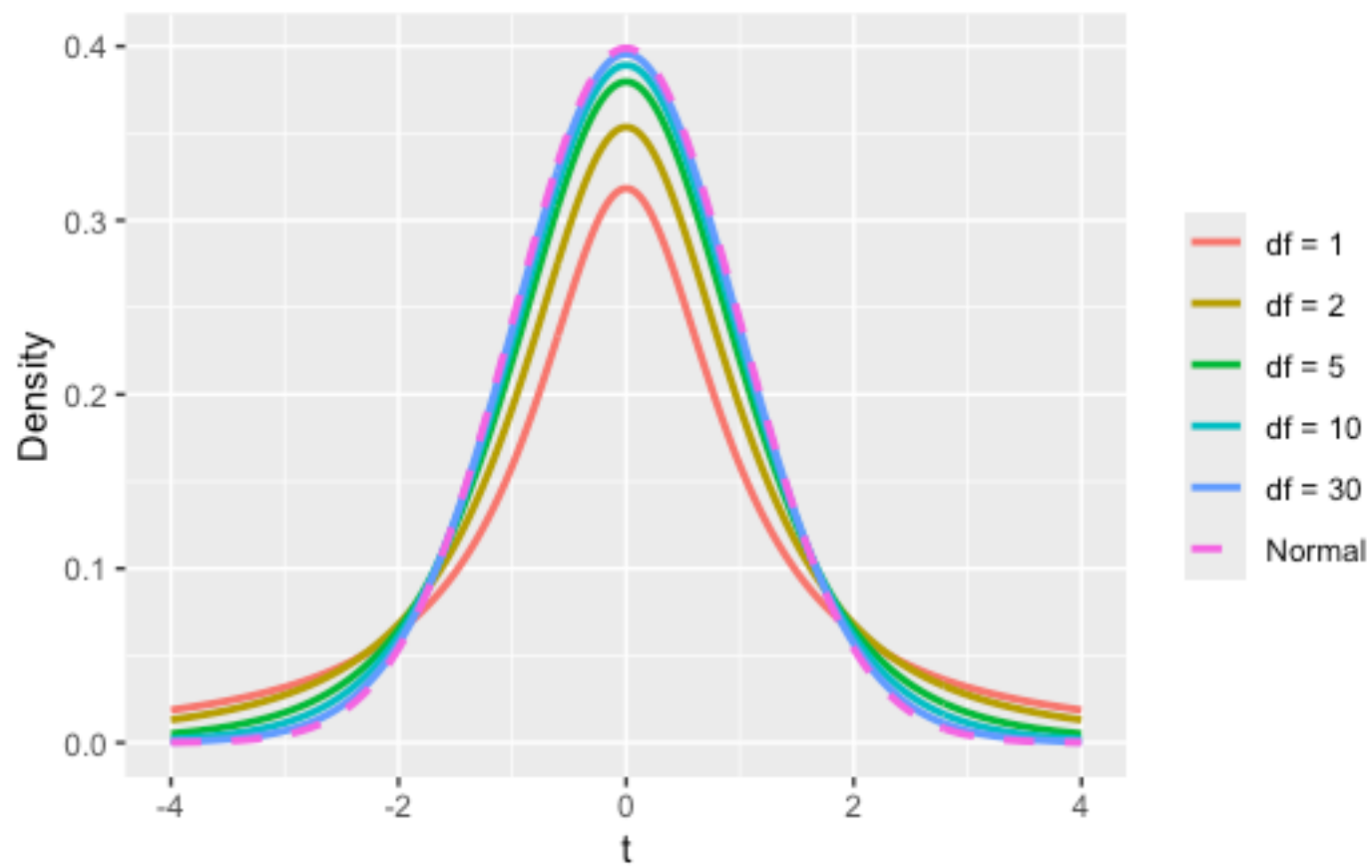
Shape changes with degrees of freedom

```r
x ← seq(-4, 4, length.out = 300)

t_data ← map_dfr(c(1, 2, 5, 10, 30), function(df) {
  tibble(x = x, density = dt(x, df), df = paste("df =", df))
})

normal_data ← tibble(x = x, density = dnorm(x), df = "Normal")

bind_rows(t_data, normal_data) |>
  mutate(df = fct_relevel(df, "df = 1", "df = 2", "df = 5", "df = 10", "df = 30", "Normal")) |>
  ggplot(aes(x, density, colour = df, linetype = df)) +
  geom_line(linewidth = 1) +
  scale_linetype_manual(values = c(rep("solid", 5), "dashed")) +
  labs(x = "t", y = "Density", colour = NULL, linetype = NULL) +
  theme(legend.position = "right")
```
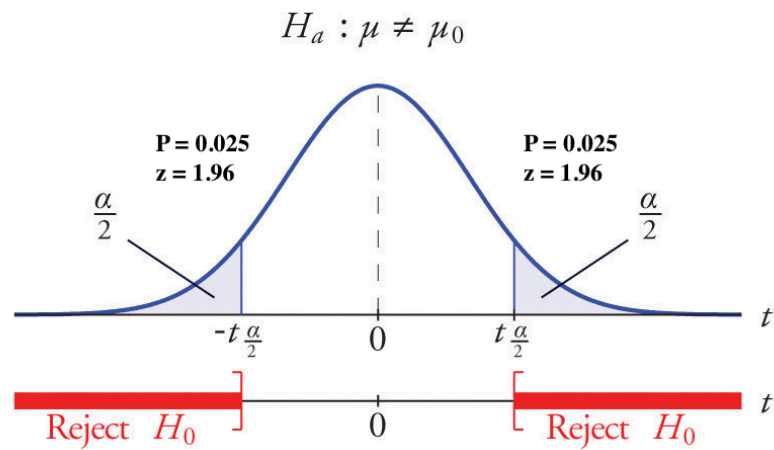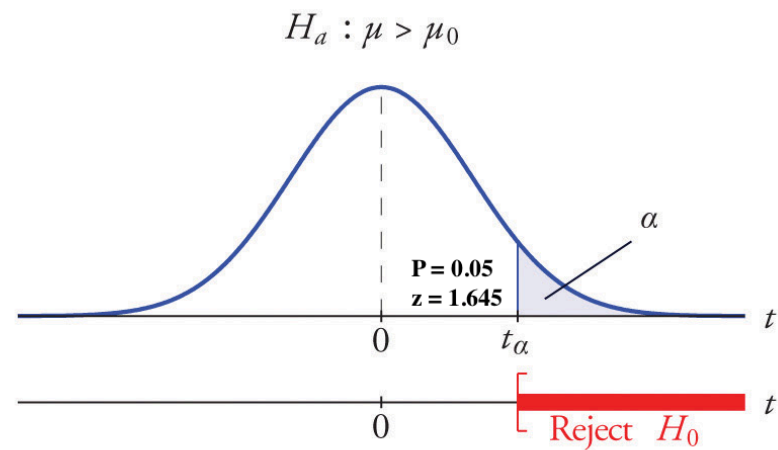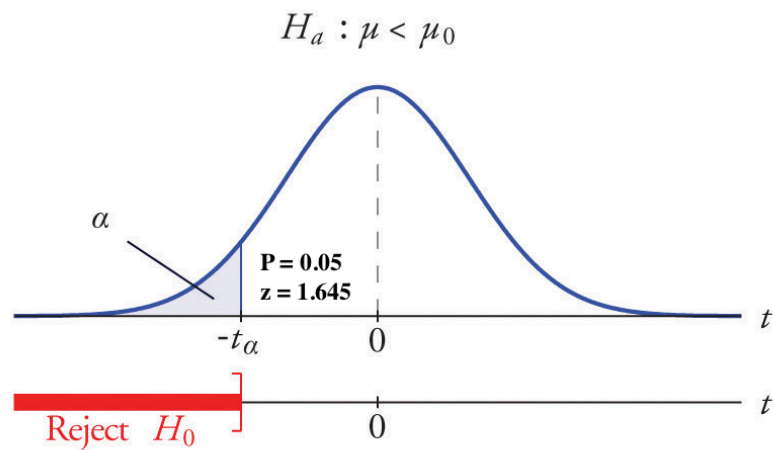
> **i Note**
>
> As degrees of freedom increase, the t-distribution approaches the standard normal distribution.

One-tailed vs two-tailed tests

$H_a : \mu < \mu_0$

α

P = 0.05
z = 1.645

$-t_\alpha$ 0 t

Reject H_0 0 t

$H_a : \mu > \mu_0$

α

P = 0.05
z = 1.645

0   t_α t

0   Reject H_0 t

$H_a : \mu \neq \mu_0$

P = 0.025
z = 1.96

P = 0.025
z = 1.96

$\frac{\alpha}{2}$

$\frac{\alpha}{2}$

$-t\frac{\alpha}{2}$ 0 $t\frac{\alpha}{2}$ t

Reject H_0 0   Reject H_0 t

- **Two-tailed**: tests for a difference in *either* direction ($\mu_1 \neq \mu_2$)
- **One-tailed**: tests for a difference in a *specific* direction ($\mu_1 > \mu_2$ or $\mu_1 < \mu_2$)
- In most cases, we use **two-tailed** tests unless we have a strong *a priori* reason for one direction.

Manual calculations

Group means and standard deviations

```r
cattle_summary ← cattle |>
  group_by(breed) |>
  summarise(
    n = n(),
    mean_wt = mean(weight, na.rm = TRUE),
    sd_wt = sd(weight, na.rm = TRUE)
  )

cattle_summary
```

```
# A tibble: 2 × 4
  breed       n mean_wt sd_wt
  <fct>   <int>   <dbl> <dbl>
1 Breed1     12    196.  10.6
2 Breed2     15    154.  12.3
```

Computing the t-statistic

```r
# Extract values
n1 <- cattle_summary$n[1]
n2 <- cattle_summary$n[2]
m1 <- cattle_summary$mean_wt[1]
m2 <- cattle_summary$mean_wt[2]
s1 <- cattle_summary$sd_wt[1]
s2 <- cattle_summary$sd_wt[2]

# Pooled variance
sp2 <- ((n1 - 1) * s1^2 + (n2 - 1) * s2^2) / (n1 + n2 - 2)

# t-statistic
t_stat <- (m1 - m2) / sqrt(sp2 * (1/n1 + 1/n2))
t_stat
```

```
[1] 9.462409
```

```r
# Degrees of freedom
df <- n1 + n2 - 2

# p-value (two-tailed)
p_val <- 2 * pt(-abs(t_stat), df)
p_val
```

```
[1] 9.663383e-10
```

Running `t.test()` in R

The easy way

```r
t.test(weight ~ breed, data = cattle, var.equal = TRUE)
```

```
	Two Sample t-test

data:  weight by breed
t = 9.4624, df = 25, p-value = 9.663e-10
alternative hypothesis: true difference in means between group Breed1 and group Breed2 is not equal
to 0
95 percent confidence interval:
 33.23011 51.71989
sample estimates:
mean in group Breed1 mean in group Breed2
            196.175              153.700
```

> 💡 **Tip**
>
> `var.equal = TRUE` performs the **pooled** (Student's) t-test, which assumes equal variances. This matches our manual calculation.

Assumptions

Three key assumptions

1. **Equal variances**: $\sigma_1^2 \approx \sigma_2^2$
2. **Normality**: $\varepsilon_{ij} \sim N(0, \sigma^2)$
3. **Independence** of observations

Checking equal variances

The only difference between the two groups should be the **location of the mean** – all else should remain the same.

General guide: $\frac{\text{larger SD}}{\text{smaller SD}} < 2.0$

```
max(cattle_summary$sd_wt) / min(cattle_summary$sd_wt)
```
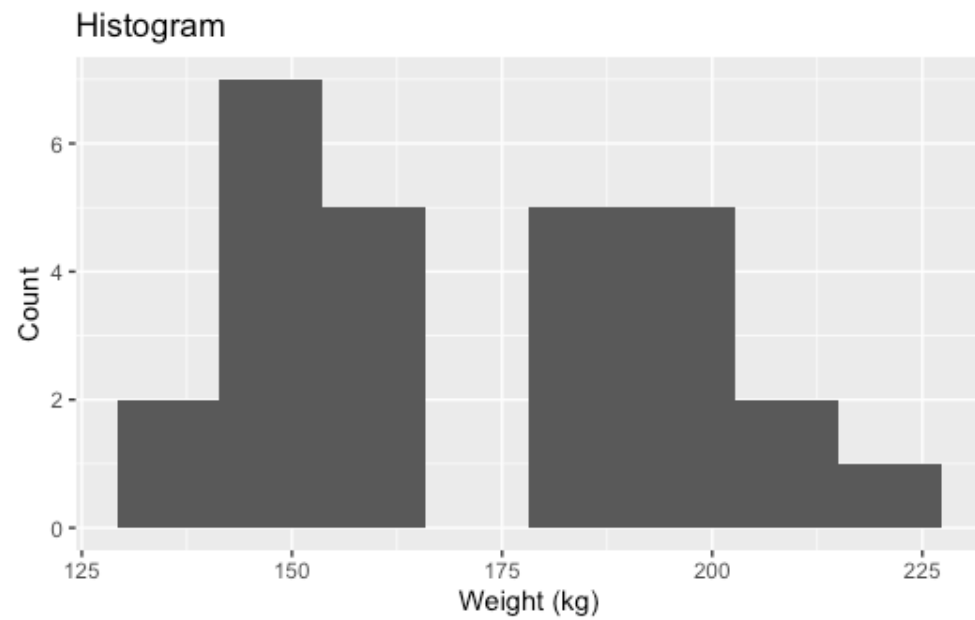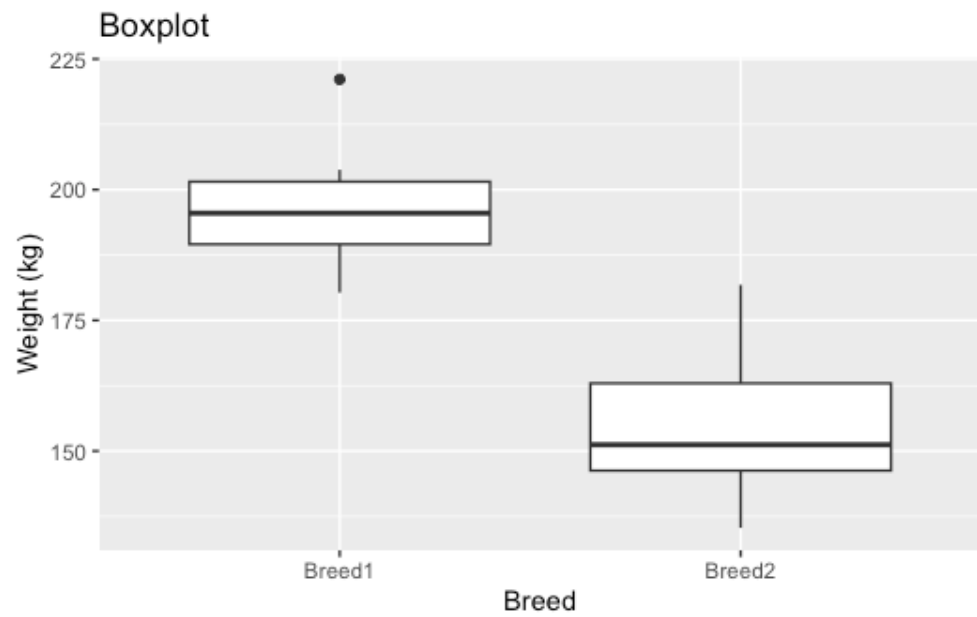
```
[1] 1.158756
```

The ratio is well under 2, so we can assume equal variances.

Checking normality: visual

```r
ggplot(cattle, aes(breed, weight)) +
  geom_boxplot() +
  labs(title = "Boxplot", x = "Breed", y = "Weight (kg)")
ggplot(cattle, aes(weight)) +
  geom_histogram(bins = 8) +
  labs(title = "Histogram", x = "Weight (kg)", y = "Count")
```

Checking normality: Shapiro-Wilk test

```
shapiro.test(cattle$weight)
```

```
	Shapiro-Wilk normality test

data:  cattle$weight
W = 0.93704, p-value = 0.103
```

- If $p > 0.05$, the data is **not** significantly different from a normal distribution, *i.e.* we can assume normality.

What if assumptions are not met?

- **Equal variances** not met – use Welch's t-test (the default in R's `t.test()` when `var.equal = FALSE`)
- **Normality** not met – if $N > 30$, we can often assume normality anyway (Central Limit Theorem)
- **Independence** not met – consider a **paired t-test** (e.g. before/after measurements on the same subjects)

Summary

Key points

- The **two-sample t-test** compares means of two independent groups
- The test statistic follows a t-distribution with $n_1 + n_2 - 2$ degrees of freedom
- Always check assumptions: **equal variances**, **normality**, **independence**
- Use `t.test()` in R for quick analysis

Next up

- What if we have **more than two groups**? We need **ANOVA**!

Thanks!

Questions?

This presentation is based on the SOLES Quarto reveal.js template and is licensed under a Creative Commons Attribution 4.0 International License.