Parametric Tests using R

Dr. Kapil Kumar
Associate Professor, Department of Statistics
Chaudhary Charan Singh University, Meerut, India
kapilstats@gmail.com

Five-day Training Programme on Data Analysis and Modeling using R (23–27 March 2026)

Department of Statistics, University of Lucknow, with the National Statistics Office (CQCD), Zonal Office, Lucknow

25 March 2026 · Lecture 1 of 2

Plan of the session

1. Parametric tests and their assumptions
2. Checking normality
3. Checking equal variances
4. The three t-tests
5. Case study: two golf balls
6. Reporting, and common mistakes

By the end you can state the assumptions of a parametric test, check normality with plots and formal tests, test equality of variances, choose and run the right t-test, and read and report the R output correctly.

PART 1

Parametric tests

What is a parametric test?

- A test whose validity rests on an assumed **form** for the population, described by a few **parameters**, for example $N(\mu, \sigma^2)$
- The hypotheses are statements about those parameters: $\mu = 50$, $\mu_1 = \mu_2$, $\sigma_1^2 = \sigma_2^2$
- Familiar examples: z-test, t-tests, F-test, ANOVA, tests in regression
- Today: normality tests, variance tests and the three t-tests

Why use them When their assumptions hold they use the data efficiently, are more powerful than rank-based tests, and give confidence intervals for quantities such as a mean difference.

But Their p-values are only as good as the assumptions behind them.

The assumptions, and how to check them

Assumption	What it means	How to check
Independence	one value does not affect another	study design, not a test
Scale	numeric, interval or ratio data	definition of the variable
Normality	each population roughly normal	Q–Q plot, Shapiro–Wilk
Equal variances	same spread in the groups (some tests)	F, Bartlett, Levene
Design	one sample, two groups, or pairs	how the data were collected

Ask before any test Independent? Numeric? Roughly normal? Similar variances? One sample, two groups, or pairs?

If one fails Unequal variances: Welch t. Strong skew, small n: Mann–Whitney or Wilcoxon. Dependence: paired methods.

Hypothesis testing: a quick refresher

- H_0: the claim tested (no difference); H_1: the alternative, two-sided or one-sided, fixed *before* seeing the data
- α: the accepted risk of rejecting a true H_0, usually 0.05
- p-value: the probability, *if* H_0 *is true*, of a result at least as extreme as the one observed
- Reject H_0 when $p \leq \alpha$; otherwise "do not reject" (never "accept" or "prove")

Decision	H_0 true	H_0 false
Reject	Type I (α)	correct
Do not reject	correct	Type II (β)

Power $= 1 - \beta$: the chance of detecting a real effect. It grows with sample size and with the size of the effect.

Packages and settings used today

```r
# install once:
# install.packages(c("nortest", "car"))
library(nortest)   # Anderson-Darling, Lilliefors
library(car)       # leveneTest()
options(digits = 5)          # digits in output
col_main <- "#1F5FA8"        # plot colours
col_alt  <- "#B45309"
```

- Base R (**stats**) has `shapiro.test()`, `ks.test()`, `var.test()`, `bartlett.test()` and `t.test()`
- **nortest** adds more normality tests; **car** adds Levene's test
- The case study reads `data/Golf.csv` with a relative path

PART 2

Checking normality

The normal distribution and why we check it

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \quad x \in \mathbb{R}$$

- Symmetric and bell-shaped about μ; mean = median = mode
- σ sets the spread: about 68%, 95% and 99.7% of values lie within 1, 2 and 3 σ of μ

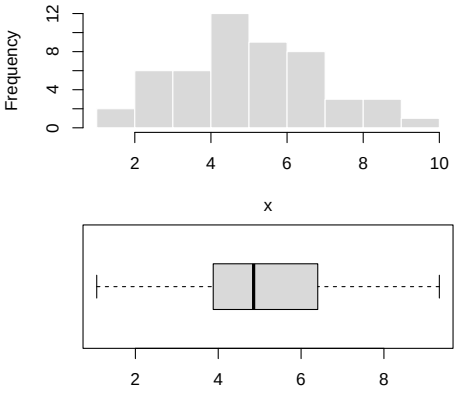
Why check With strongly non-normal data in small samples, t-test p-values and confidence intervals can mislead.

Two routes, used together
Graphs: histogram, box plot, Q–Q plot. **Tests**: Shapiro–Wilk, Anderson–Darling, Lilliefors.

A sample to work with: histogram and box plot

```r
set.seed(123)
x <- rnorm(50, mean = 5, sd = 2)
c(mean = mean(x), median = median(x), sd = sd(x))
par(mfrow = c(2, 1), mar = c(4, 4, 0.5, 1))
hist(x, col = "grey85", border = "white", main = "")
boxplot(x, horizontal = TRUE, col = "grey85")

##   mean median     sd
## 5.0688 4.8547 1.8517
```

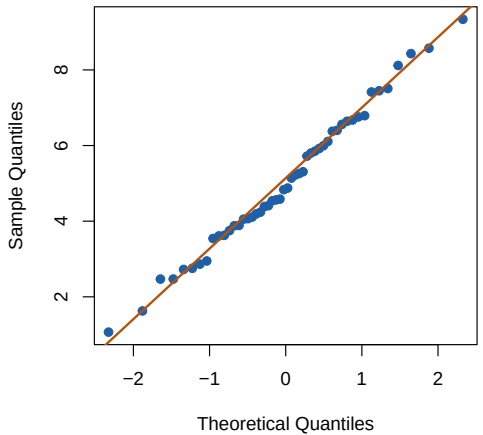


Reading it Roughly symmetric, one central peak, mean close to median, no outlying points.

The normal Q–Q plot

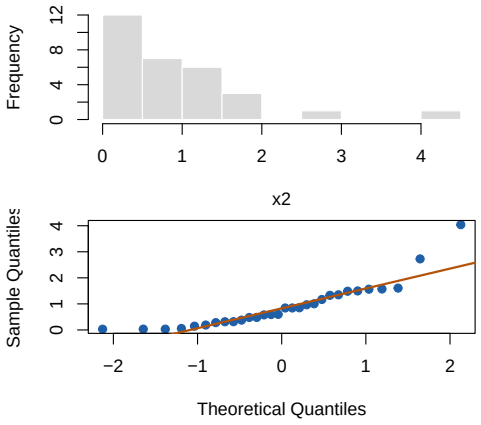
```
par(mar = c(4, 4, 1, 1))
qqnorm(x, pch = 19, col = col_main, main = "")
qqline(x, col = col_alt, lwd = 2)
```

- Sample quantiles against the quantiles a normal distribution would give
- Points close to the line: normality plausible
- A curved, banana-shaped pattern: skewness
- Points peeling away at both ends: heavy tails or outliers



A skewed sample for contrast

```r
set.seed(123)
x2 <- rexp(30, rate = 1)      # exponential
par(mfrow = c(2, 1), mar = c(4, 4, 0.5, 1))
hist(x2, col = "grey85", border = "white",
     main = "", xlab = "x2")
qqnorm(x2, pch = 19, col = col_main, main = "")
qqline(x2, col = col_alt, lwd = 2)
```



Reading it A long right tail in the histogram, and a Q–Q plot that bends away from the line: the typical picture of right skewness.

The Shapiro–Wilk test

```
shapiro.test(x)
shapiro.test(x2)$p.value

##
##  Shapiro-Wilk normality test
##
## data:  x
## W = 0.989, p-value = 0.93
##
## [1] 0.00019256
```

H_0: the data come from a normal population

$$W = \frac{\left(\sum_i a_i \, x_{(i)}\right)^2}{\sum_i (x_i - \bar{x})^2}, \quad x_{(i)} \text{ ordered}$$

Reading it x: $W = 0.989$, $p = 0.93$, no evidence against normality. x2: $p = 0.00019$, clear non-normality. A large p does not *prove* normality.

Kolmogorov–Smirnov and the Lilliefors correction

```r
ks <- ks.test(x, "pnorm", mean = mean(x), sd = sd(x))
lf <- lillie.test(x)                    # nortest
c(D_ks = unname(ks$statistic),
  D_lillie = unname(lf$statistic))
c(p_ks = ks$p.value, p_lillie = lf$p.value)

##     D_ks D_lillie
## 0.083232 0.083232
##     p_ks p_lillie
##  0.85064  0.52335
```

$$D = \max_x |F_n(x) - F_0(x)|$$

Reading it The same distance D, but different p-values (0.85 vs 0.52).

Why `ks.test()` assumes μ and σ are known in advance. Estimating them from the same data makes its p-value too large; Lilliefors corrects this.

A battery of normality tests

```
tests <- c("Shapiro-Wilk", "Anderson-Darling", "Cramer-von Mises",
           "Lilliefors", "Shapiro-Francia")
p_normal <- c(shapiro.test(x)$p.value, ad.test(x)$p.value, cvm.test(x)$p.value,
              lillie.test(x)$p.value, sf.test(x)$p.value)
p_skewed <- c(shapiro.test(x2)$p.value, ad.test(x2)$p.value, cvm.test(x2)$p.value,
              lillie.test(x2)$p.value, sf.test(x2)$p.value)
data.frame(test = tests, x = round(p_normal, 3), x2 = round(p_skewed, 4))

##              test     x     x2
## 1     Shapiro-Wilk 0.928 0.0002
## 2 Anderson-Darling 0.853 0.0025
## 3 Cramer-von Mises 0.741 0.0111
## 4       Lilliefors 0.523 0.0655
## 5  Shapiro-Francia 0.940 0.0003
```

`x2`: four tests reject at 5%; Lilliefors ($p = 0.065$) does not. Tests differ in power.

Normality tests and sample size

```r
set.seed(2026)
n_values <- c(20, 100, 500)
reject <- numeric(3)
for (i in 1:3) {          # t(5): mildly heavy tails
  p <- replicate(1000,
       shapiro.test(rt(n_values[i], df = 5))$p.value)
  reject[i] <- mean(p < 0.05)
}
data.frame(n = n_values, rejected = reject)
```

```
##     n rejected
## 1  20    0.184
## 2 100    0.576
## 3 500    0.993
```

Reading it The same mild departure is detected in 18% of samples of 20 but in 99% of samples of 500.

Practical rule Small n: tests miss real problems; look at the plots. Large n: tests flag harmless departures, and the CLT protects the t-test.

PART 3

Equal variances

Two groups: the F-test

```r
group1 <- c(12, 15, 14, 10, 18, 16, 13, 17)
group2 <- c(22, 25, 24, 20, 28, 26, 23, 27)
var.test(group1, group2)

##
##  F test to compare two variances
##
## data:  group1 and group2
## F = 1, num df = 7, denom df = 7, p-value = 1
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.2002 4.9949
## sample estimates:
## ratio of variances
##                  1
```

$$H_0 : \sigma_1^2 = \sigma_2^2$$
$$F = S_1^2 / S_2^2 \sim F_{n_1 - 1,\, n_2 - 1}$$

Reading it `group2` is `group1` plus 10: the centre moves, the spread does not, so $F = 1$, $p = 1$. Beware: the F-test is very sensitive to non-normality.

Three or more groups: Bartlett's test

```r
values <- c(12, 15, 14, 10, 18, 16,
            22, 25, 24, 20, 28, 26,
            32, 35, 30, 31, 36, 34)
group <- factor(rep(c("A", "B", "C"), each = 6))
tapply(values, group, var)
bartlett.test(values ~ group)
```

```
##      A      B      C
## 8.1667 8.1667 5.6000
##
##  Bartlett test of homogeneity of variances
##
## data:  values by group
## Bartlett's K-squared = 0.208, df = 2, p-value = 0.9
```

$H_0 : \sigma_1^2 = \cdots = \sigma_k^2$; under H_0 Bartlett's statistic is approximately χ^2_{k-1}

Reading it Variances 8.17, 8.17 and 5.6; $p = 0.9$: equal variances are plausible.

Like the F-test, Bartlett's test assumes normal data.

Levene's test, and which test to use

```
leveneTest(values ~ group)        # car: ANOVA on |deviations from the group median|

## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value Pr(>F)
## group  2    0.03   0.97
##       15
```

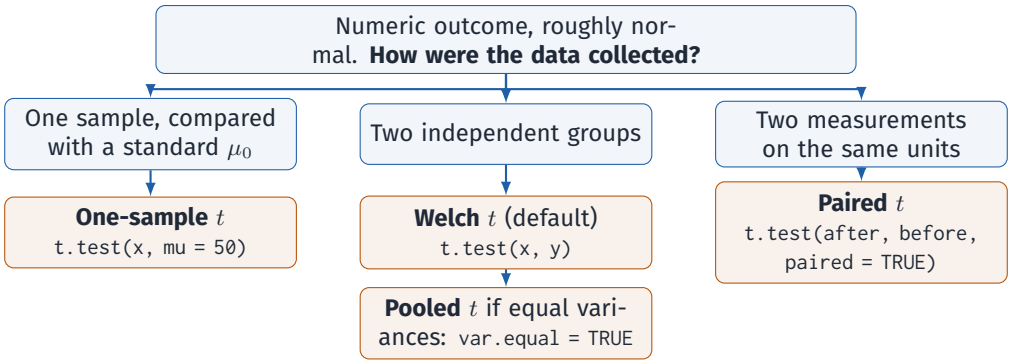
Levene's test compares the groups' average distance from their centre. **car** uses the median (the Brown–Forsythe form), which is robust to non-normality. Here $p = 0.97$, in line with Bartlett.

Situation	Test
2 groups, normal data	`var.test()`
k groups, normal data	`bartlett.test()`
normality doubtful	`leveneTest()`
unequal variances	Welch t, `oneway.test()`

PART 4

The t-tests

Which t-test? Start from the design



The first question is not “which command?” but “how were my data collected?”

One-sample t-test: the idea, by hand in R

$H_0 : \mu = \mu_0$ vs $H_1 : \mu \neq \mu_0$ (or $>$, $<$)

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \sim t_{n-1} \text{ under } H_0$$

$$s^2 = \frac{1}{n-1}\sum (x_i - \bar{x})^2$$

Example: ten marks; is the mean mark 50? The steps on the right are the "manual method", done with R as a calculator.

```
marks <- c(48, 52, 49, 55, 51, 47, 53, 50, 46, 54)
n    <- length(marks)
xbar <- mean(marks)
s    <- sd(marks)
t0   <- (xbar - 50) / (s / sqrt(n))
c(n = n, mean = xbar, sd = s, t = t0, df = n - 1)
2 * (1 - pt(abs(t0), df = n - 1))  # p-value

##        n     mean       sd        t       df
## 10.00000 50.50000  3.02765  0.52223  9.00000
## [1] 0.61412
```

One-sample t-test with `t.test()`

```
t1 <- t.test(marks, mu = 50)
t1

##
##  One Sample t-test
##
## data:  marks
## t = 0.522, df = 9, p-value = 0.61
## alternative hypothesis: true mean is not equal to 50
## 95 percent confidence interval:
##  48.334 52.666
## sample estimates:
## mean of x
##      50.5
```

Reading the output $t = 0.522$ on 9 df, the same as by hand. $p = 0.614$: no evidence that the mean differs from 50. The 95% CI for μ (48.33 to 52.67) contains 50.

A two-sided test at 5% rejects exactly when the 95% CI excludes μ_0.

One-sided alternatives and p-values from `pt()`

```r
t.test(marks, mu = 50,
       alternative = "greater")$p.value
nu <- n - 1                    # degrees of freedom
c(left  = pt(t0, nu),
  right = 1 - pt(t0, nu),
  two   = 2 * (1 - pt(abs(t0), nu)))

## [1] 0.30706
##    left   right     two
## 0.69294 0.30706 0.61412
```

Alternative	p-value
$H_1 : \mu < \mu_0$	$P(T_\nu \le t_0)$
$H_1 : \mu > \mu_0$	$P(T_\nu \ge t_0)$
$H_1 : \mu \neq \mu_0$	$2\,P(T_\nu \ge \lvert t_0 \rvert)$

Reading it With $t_0 = 0.522$ the right-tail p (0.307) is half the two-sided p. Choose the direction from the question, never from the data.

Two independent samples: pooled or Welch?

Pooled (equal variances assumed)

$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_p\sqrt{1/n_1 + 1/n_2}} \sim t_{n_1+n_2-2}$$

Welch (variances may differ)

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S_1^2/n_1 + S_2^2/n_2}},$$ with approximate df

$$\nu = \frac{(S_1^2/n_1 + S_2^2/n_2)^2}{\frac{(S_1^2/n_1)^2}{n_1-1} + \frac{(S_2^2/n_2)^2}{n_2-1}}$$

Practice $H_0 : \mu_1 = \mu_2$. `t.test(x, y)` gives Welch by default; add `var.equal = TRUE` for the pooled test. Welch loses little when variances are equal and protects you when they are not, so it is the safer default.

Two independent samples in R

```r
groupA <- c(45, 48, 50, 52, 47, 49)
groupB <- c(55, 58, 54, 57, 56, 53)
t2 <- t.test(groupA, groupB, var.equal = TRUE)
t2
```

```
##
##  Two Sample t-test
##
## data:  groupA and groupB
## t = -5.59, df = 10, p-value = 0.00023
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -9.7889 -4.2111
## sample estimates:
## mean of x mean of y
##      48.5      55.5
```

```r
sp2 <- (5 * var(groupA) +
        5 * var(groupB)) / 10
(mean(groupA) - mean(groupB)) /
  sqrt(sp2 * (1/6 + 1/6))  # t
```

```
## [1] -5.5925
```

Reading it By hand and by `t.test()`: $t = -5.593$ on 10 df, $p = 0.00023$. The CI for $\mu_A - \mu_B$ lies well below 0.

Pooled and Welch side by side

```r
var.test(groupA, groupB)$p.value   # F-test
t2w <- t.test(groupA, groupB)      # Welch
cmp <- rbind(
  pooled = c(t2$statistic, t2$parameter, t2$p.value),
  welch  = c(t2w$statistic, t2w$parameter, t2w$p.value))
colnames(cmp) <- c("t", "df", "p")
round(cmp, 5)

## [1] 0.58055
##              t     df       p
## pooled -5.5926 10.000 0.00023
## welch  -5.5926  9.388 0.00029
```

Reading it The F-test gives no reason to doubt equal variances ($p = 0.58$). Welch uses fewer degrees of freedom (9.4 instead of 10), but both tests give the same t (equal group sizes) and the same conclusion.

The paired t-test

```r
before <- c(60, 62, 58, 64, 57, 59)
after  <- c(65, 66, 60, 68, 61, 62)
d <- after - before
(tp <- t.test(after, before, paired = TRUE))  # ( ) assigns and prints

##
##  Paired t-test
##
## data:  after and before
## t = 8.7, df = 5, p-value = 0.00033
## alternative hypothesis: true mean difference is not equal to 0
## 95 percent confidence interval:
##  2.5828 4.7505
## sample estimates:
## mean difference
##          3.6667
```

$d_i = \text{after}_i - \text{before}_i$
$H_0 : \mu_d = 0$
$t = \bar{d} / (s_d / \sqrt{n}) \sim t_{n-1}$

Reading it Scores rose by 3.67 points on average ($t = 8.696$, $p = 0.00033$). Same as `t.test(d, mu = 0)`.

What if we ignore the pairing?

```
c(paired      = t.test(after, before,
                       paired = TRUE)$p.value,
  independent = t.test(after, before)$p.value)
c(sd_scores = sd(before), sd_change = sd(d))

##      paired independent
##  0.00033266  0.05332220
## sd_scores sd_change
##    2.6077    1.0328
```

A costly mistake Treated as two independent groups, the same data give $p = 0.053$: not significant at 5%.

Why pairing helps Scores vary a lot between people (sd 2.61), changes much less (sd 1.03). Pairing removes the between-person variation.

The three t-tests at a glance

Test	Question	R call	Example
One-sample	mean $= \mu_0$?	`t.test(x, mu = 50)`	mean mark 50?
Two-sample	two means differ?	`t.test(x, y)`	two golf balls
pooled	… equal variances	`t.test(x, y, var.equal = TRUE)`	
Paired	mean change zero?	`t.test(a, b, paired = TRUE)`	before / after

Same output, every time R prints the t statistic, degrees of freedom, p-value, a confidence interval and the estimate(s). Read all five, not only the p-value.

PART 5

Case study: two golf balls

The question and the data

- A manufacturer (“Par, Inc.”) has a new golf ball and compares it with its current ball
- Driving distance in yards for 40 drives with each ball: `data/Golf.csv`, columns `Current` and `New`
- Public teaching data set (Kaggle: *Golf ball testing data set from Par Inc*)

Research question Does the mean driving distance of the new ball differ from that of the current ball?

μ_1: mean distance, current ball
μ_2: mean distance, new ball
$H_0 : \mu_1 = \mu_2$ vs $H_1 : \mu_1 \neq \mu_2$

The two sets of 40 drives are treated as **independent samples**: no drive with one ball is matched to a drive with the other. So this is a two-sample t-test.

Reading and summarising the data

```r
golf <- read.csv("data/Golf.csv")
head(golf, 2)
sum(is.na(golf))                    # missing values?
round(sapply(golf, function(v) c(n = length(v),
        mean = mean(v), sd = sd(v))), 2)
```

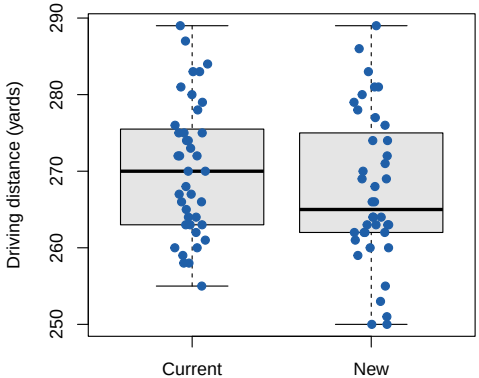
```
##   Current New
## 1     264 277
## 2     261 269
## [1] 0
##      Current   New
## n      40.00  40.0
## mean  270.28 267.5
## sd      8.75   9.9
```

Reading it 40 drives per ball, no missing values. The current ball averages 270.3 yards (sd 8.8), the new ball 267.5 yards (sd 9.9): a difference of 2.77 yards in favour of the *current* ball.

Look at the data first

```r
par(mar = c(4, 4, 1, 1))
boxplot(golf, col = "grey90",
        ylab = "Driving distance (yards)")
stripchart(golf, vertical = TRUE,
           method = "jitter", pch = 19,
           col = col_main, add = TRUE)
```

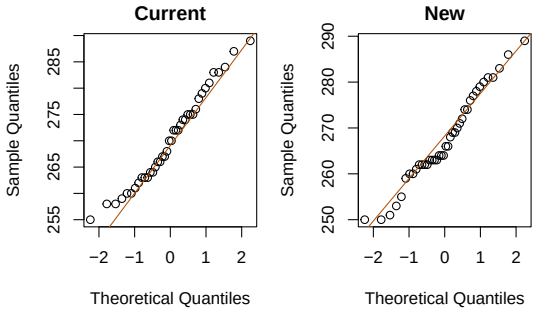
Reading it The two distributions overlap almost completely and have similar spread; neither shows strong skewness or extreme outliers.



Assumption 1: normality in each group

```r
par(mfrow = c(1, 2), mar = c(4, 4, 2, 1))
qqnorm(golf$Current, main = "Current")
qqline(golf$Current, col = col_alt)
qqnorm(golf$New, main = "New")
qqline(golf$New, col = col_alt)
p_cur <- shapiro.test(golf$Current)$p.value
p_new <- shapiro.test(golf$New)$p.value
c(current = p_cur, new = p_new)

## current     new
## 0.37880 0.30647
```



Reading it Points on the lines, both $p > 0.05$: normality is plausible.

Assumption 2: equal variances

```r
vt <- var.test(golf$Current, golf$New)
c(F = unname(vt$statistic), p = vt$p.value)
golf_long <- data.frame(         # "long" format: values and groups
  distance = c(golf$Current, golf$New),
  ball = factor(rep(c("Current", "New"), each = 40)))
leveneTest(distance ~ ball, data = golf_long)

##       F       p
## 0.78219 0.44653
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value Pr(>F)
## group  1    0.26   0.61
##       78
```

Reading it Variance ratio 0.78 (F-test $p = 0.45$); Levene agrees ($p = 0.61$). Equal variances are plausible, so the pooled test is justified.

The two-sample t-test

```
golf_t <- t.test(golf$Current, golf$New, var.equal = TRUE)
golf_t

##
##  Two Sample t-test
##
## data:  golf$Current and golf$New
## t = 1.33, df = 78, p-value = 0.19
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -1.384  6.934
## sample estimates:
## mean of x mean of y
##    270.27    267.50
```

Reading it $t = 1.33$ on 78 df, $p = 0.188$. The 95% CI for $\mu_1 - \mu_2$ runs from -1.38 to 6.93 yards and contains 0: no significant difference.

Welch's test gives $p = 0.188$: the same conclusion.

Interpretation: size, not just significance

```
m1 <- mean(golf$Current); m2 <- mean(golf$New)
sp <- sqrt((39 * var(golf$Current) +
            39 * var(golf$New)) / 78)
c(difference = m1 - m2,
  cohen_d = (m1 - m2) / sp)

## difference    cohen_d
##    2.77500    0.29703
```

Cohen's $d = (\bar{x}_1 - \bar{x}_2)/S_p$: the difference in standard deviations

Conclusion for the manufacturer
No significant difference in mean driving distance. The new ball averaged 2.8 yards less, a small effect ($d = 0.3$); anything from 6.9 yards shorter to 1.4 yards longer is consistent with the data.

"Not significant" does not prove the balls are identical: a sharper answer needs more drives.

A trap: the one-sided question

```r
# H1: the new ball goes farther
t.test(golf$New, golf$Current,
       alternative = "greater",
       var.equal = TRUE)$p.value
```

```
## [1] 0.90603
```

Note the order: `t.test(New, Current)` with `"greater"` tests $\mu_2 > \mu_1$.

Reading it If the question is "does the new ball travel farther?", $p = 0.906$: the sample mean actually went the other way.

Rule Fix the direction of H_1 from the business question before looking at the data. Switching to the side the data favour doubles the real Type I error rate.

PART 6

Reporting and pitfalls

Reading and reporting the output

- **Normality / variance tests**, large p: "no evidence against normality (equal variances)"; small p: "evidence of a departure"
- **t-tests**: "we reject H_0" or "we do not reject H_0", never "H_0 is proved"
- For non-statisticians: "the data provide (do not provide) evidence of a difference"

A complete report of the golf test A pooled two-sample t-test found no significant difference in mean driving distance between the current ball ($M = 270.3$, $SD = 8.8$) and the new ball ($M = 267.5$, $SD = 9.9$), $t(78) = 1.33$, $p = 0.188$; 95% CI for the difference -1.38 to 6.93 yards.

Estimate, interval, test statistic with df, and p: all four belong in a report.

Common mistakes in practice

Assumptions

- Skipping the normality check with small, skewed samples
- Using the pooled test without looking at the variances (prefer Welch)
- Analysing paired data as two independent groups
- Applying a t-test to ranks or categories

Interpretation

- Reporting only p, without the estimate and its CI
- Reading “significant” as “important”, or “not significant” as “no difference”
- Choosing a one-sided H_1 after seeing the data
- Trusting one normality test alone in very small or very large samples

Key take-aways

1. **Design first**: one sample, two independent groups or pairs decides the test.
2. **Check normality** with plots and a test together; mind the sample size.
3. **Check variances** with F, Bartlett or (more robust) Levene; prefer Welch when in doubt.
4. **Read all of the output**: t, df, p, confidence interval and estimates.
5. **Pairing matters**: ignoring it can hide a real effect.
6. **Report size and uncertainty**, in plain language for decision makers.

Next session: reporting these results with R Markdown, turning the golf-ball analysis into a reproducible report.

References

- Dalgaard, P. (2008). *Introductory Statistics with R*, 2nd Ed., Springer, New York.
- Verzani, J. (2014). *Using R for Introductory Statistics*, 2nd Ed., CRC Press, Boca Raton.
- Fox, J., and Weisberg, S. (2019). *An R Companion to Applied Regression*, 3rd Ed., Sage, Thousand Oaks.
- Shapiro, S. S., and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4), 591–611.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna. `https://www.R-project.org`
- Data: *Golf ball testing data set from Par Inc*, Kaggle. `https://www.kaggle.com/datasets/ipravin/golf-ball-testing-data-set-from-par-inc`

Thank you

Questions and discussion