Tutorial 9.

Q1.

| | Before | | After | diff B-A |
|---|---|---|---|---|
| 1 | 10 | → | 12 | |
| ⋮ | ⋮ | | | |
| 10 | ⋮ | | | |

2 samples → paired / matched

t. test for paired or
t. test for difference.

if samples are normal or have large size.

if not large size, not normal → use non parametric, that is W. S. R test

Q2. what test to use?

comparing GMP(y) for 2 types of transmission (0, 1)

→ 2 samples. → independent.

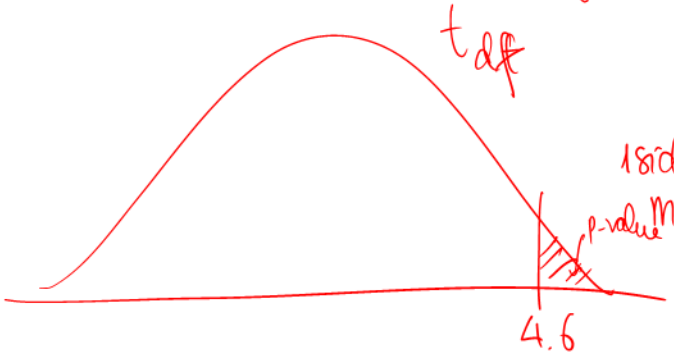
→ to compare mean, the suitable tests:

2 indpd sample t-test ← if assumptions are met: normal or large size.

OR M. W. U test ← is used if the assumptions is not met

⇒ should use 2 indpd sample t-test (equal or unequal var)

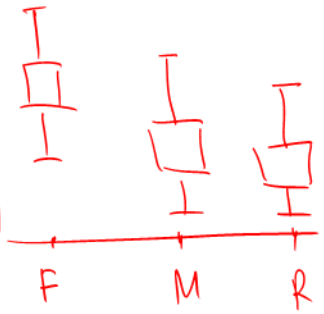
→ check equality of var using Bartlett test




manual - auto > 0.

Q3. what test should we use?

plot boxplot of sales for 3 locations




→ How many samples? → 3 samples.

→ 6 stores (obs) in each samples.

| 1K | Front | M | Rear |
|---|---|---|---|
| → 1st | 10. | | |
| → | 12 | | |
| → | ⋮ | | |
| 6th → | | | |

⇒ indpd.

⇒ compare mean sale for 3 locations

⇒ Anova ← assumption is met.

or Kruskall Wallis test ←

Assumptions for Anova
- equal var for 3 groups. ← Bartlett
- normality for each of 3 groups
  - hist + normal curve (*) ←
  - shapiro ✓
  - QQ plot ✓

⇒ can use Anova ⇒ still need to check if residuals $\sim N(0,1)$

$H_0$ : Mean sales of stores in 3 locations are the same.

$H_1$ : - - - - not the same.

p-value from Anova : $0.000524 < 0.001$

⇒ Very strong evidence against $H_0$.

⇒ Check if S. Residuals of Anova is normal:
- hist of SR + normal density curve —
- QQplot —
- shapiro test. —

---

(b) which location is different ?

⇒ pairwise test → Bonf. correction ✓
or Tukey correction ✓
Tukey is better if samples satisfying Anova.

⇒ From Tukey, at FWER = 0.05 then

$F \neq M$

$F \neq R$

$M \approx R$.