

Three-way Anova with R

stats package - No install required

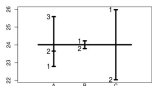
Y : numeric continuous variable
A, B, C, ... : factor (categorical) variables

Goal: Find which factors influence a quantitative continuous variable, taking into account their possible interactions

Graphical exploration

Plot the mean of Y for the different factors levels

```
plot.design(Y ~ ., data = data)
```



Plot the mean of Y for two-way combinations of factors

```
interaction.plot(data$A, data$B, data$Y)
```



Model building

Generate the full model

```
m1 <- aov(Y ~ A * B * C, data = data)
```

Update an Anova model

```
m2 <- update(m1, . ~ . - A:B:C) #Remove interaction
```

Analysis of variance table

```
summary(m1)
```

Comparison between nested models

```
anova(m1, m2)
```

```
> summary(m1)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
A	2	32.75	16.38	7.256	0.0086 **
B	1	1.14	1.14	0.505	0.4911

[These are just the first two rows!]

Write a formula in R

Full model

```
Y ~ A * B * C
```

```
Y ~ A + B + C + A:B + A:C + B:C + A:B:C
```

} same as

Update a formula

```
update(oldformula, newformula)
```

```
ex update(Y ~ A + B, . ~ . + C)
```

```
#> Y ~ A + B + C
```

```
update(Y ~ A + B + C, . ~ . - C)
```

```
#> Y ~ A + B
```

Diagnostics & Prediction

Residual plots

```
plot(m2)
```

Predict new values

```
predict(object = m2, newdata = newdata)
```

```
> anova(m1, m2)
```

Analysis of Variance Table

Model 1: Y ~ A * B * C

Model 2: Y ~ A + B + C + A:B + A:C + B:C

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	12	27.084				
2	14	31.533	-2	-4.4491	0.9856	0.4015