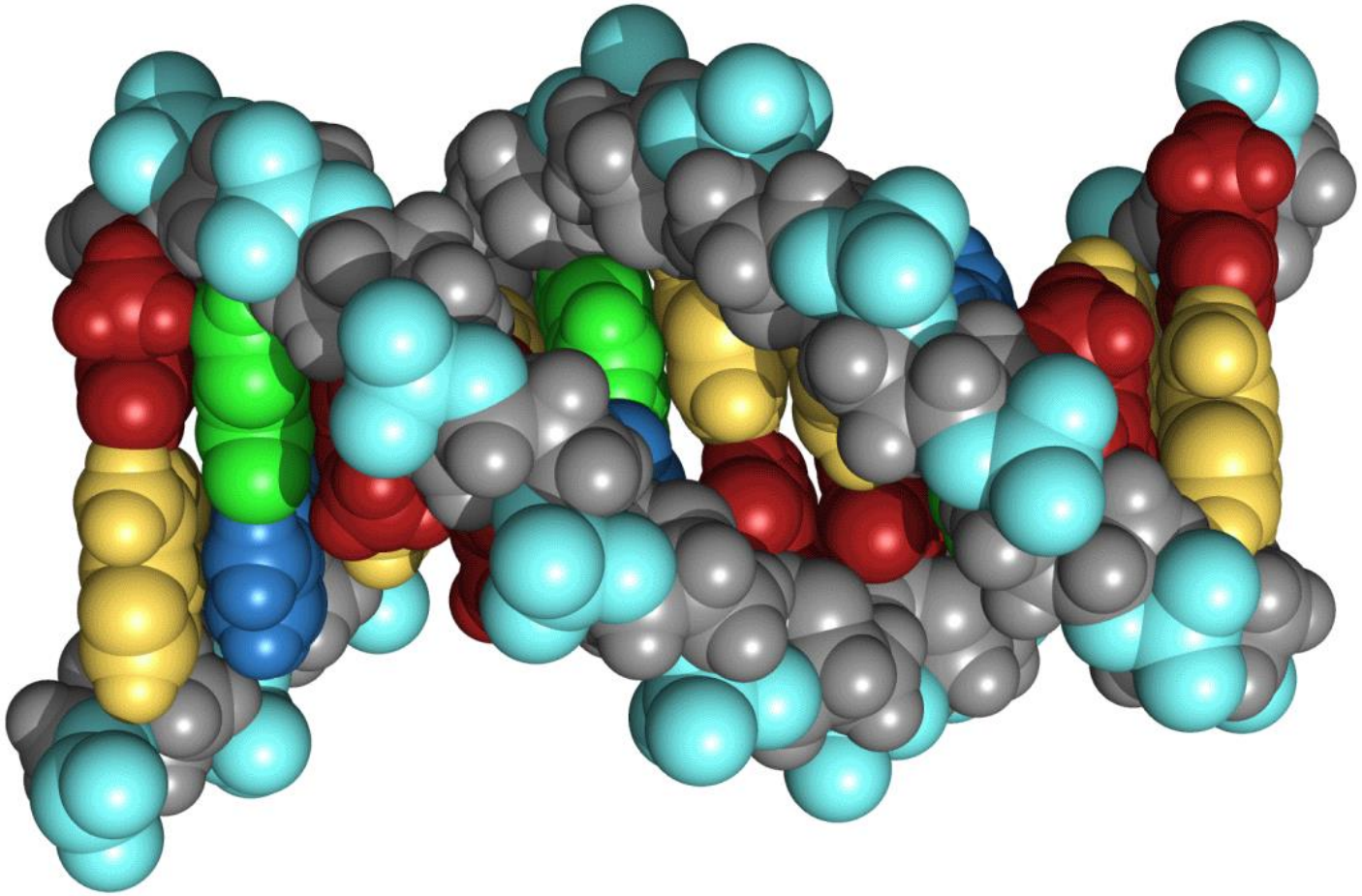
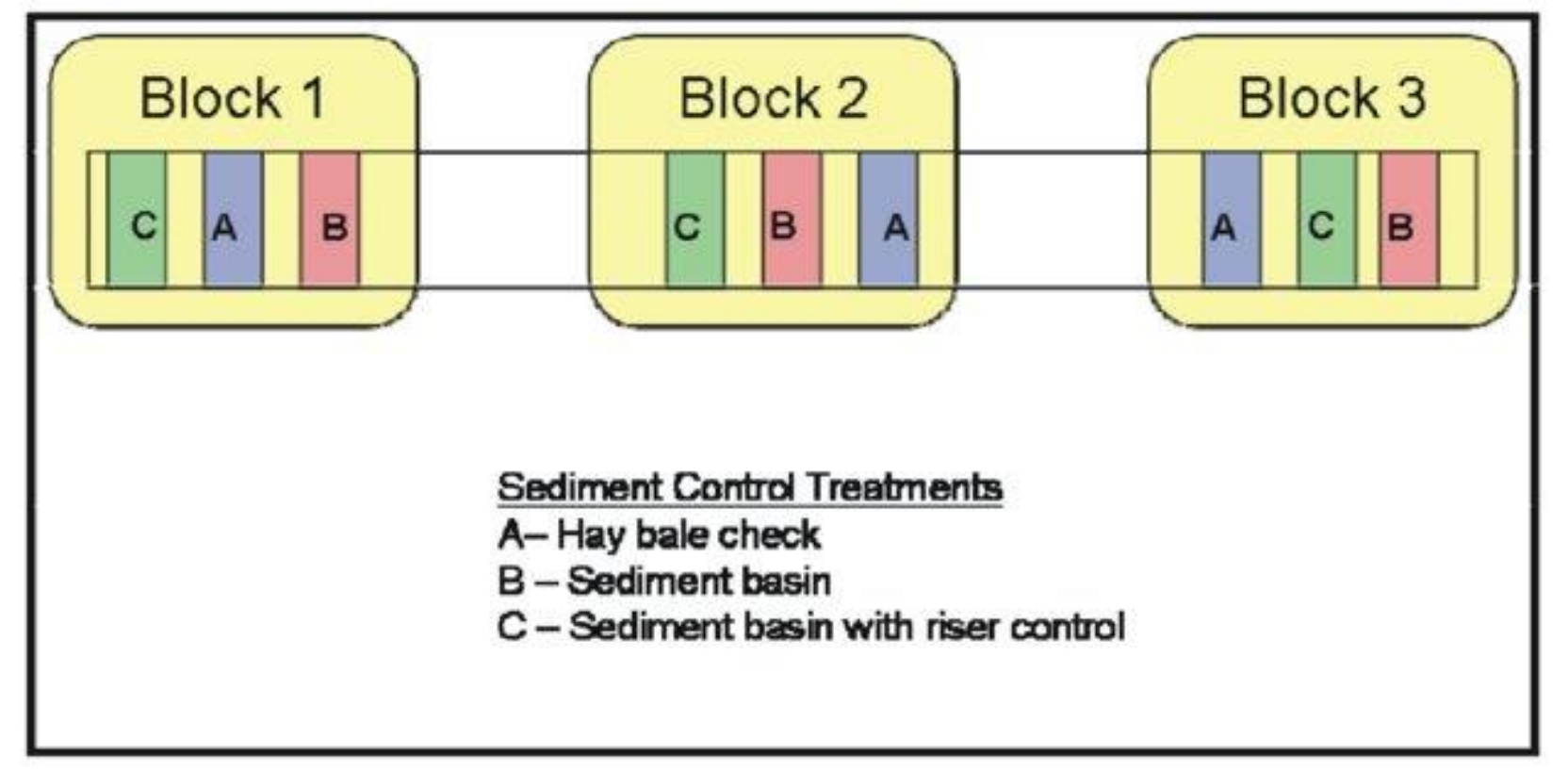
Statistiek 3 BIN

- Les 11

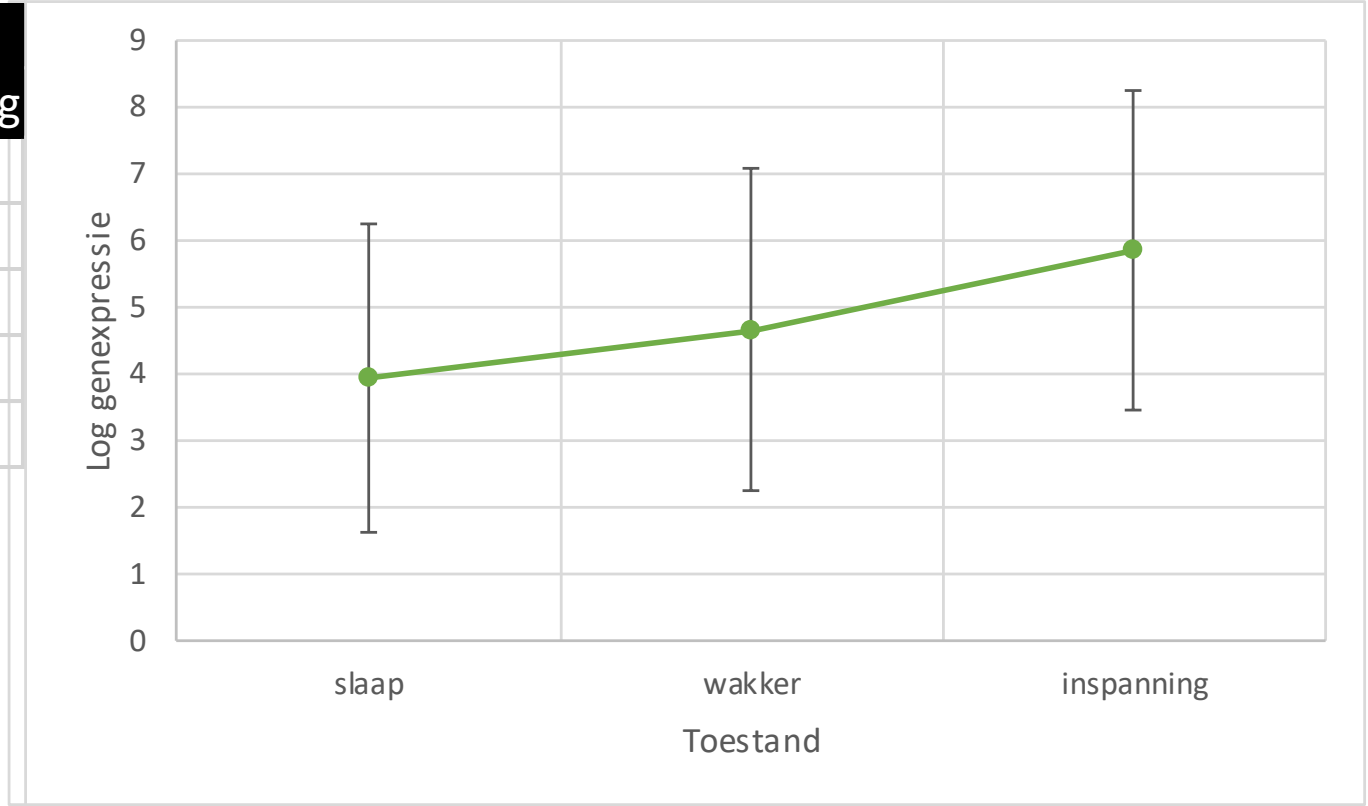


Block ANOVA


Block 1
C
A
B
Block 2
C
B
A
Block 3
A
C
B
Sediment Control Treatments
A– Hay bale check
B – Sediment basin
C – Sediment basin with riser control

1-way ANOVA

log genexpressie		toestand		
		slaap	wakker	inspanning
diersoort	koe	2.1	2.8	3.8
	geit	3.0	3.7	4.9
	hond	4.1	4.7	6.0
	paard	3.5	4.2	5.6
	mens	7.0	7.9	9.0



Statistisch model:

$$y_{ik} = \mu + \alpha_i + \varepsilon_{ik}$$

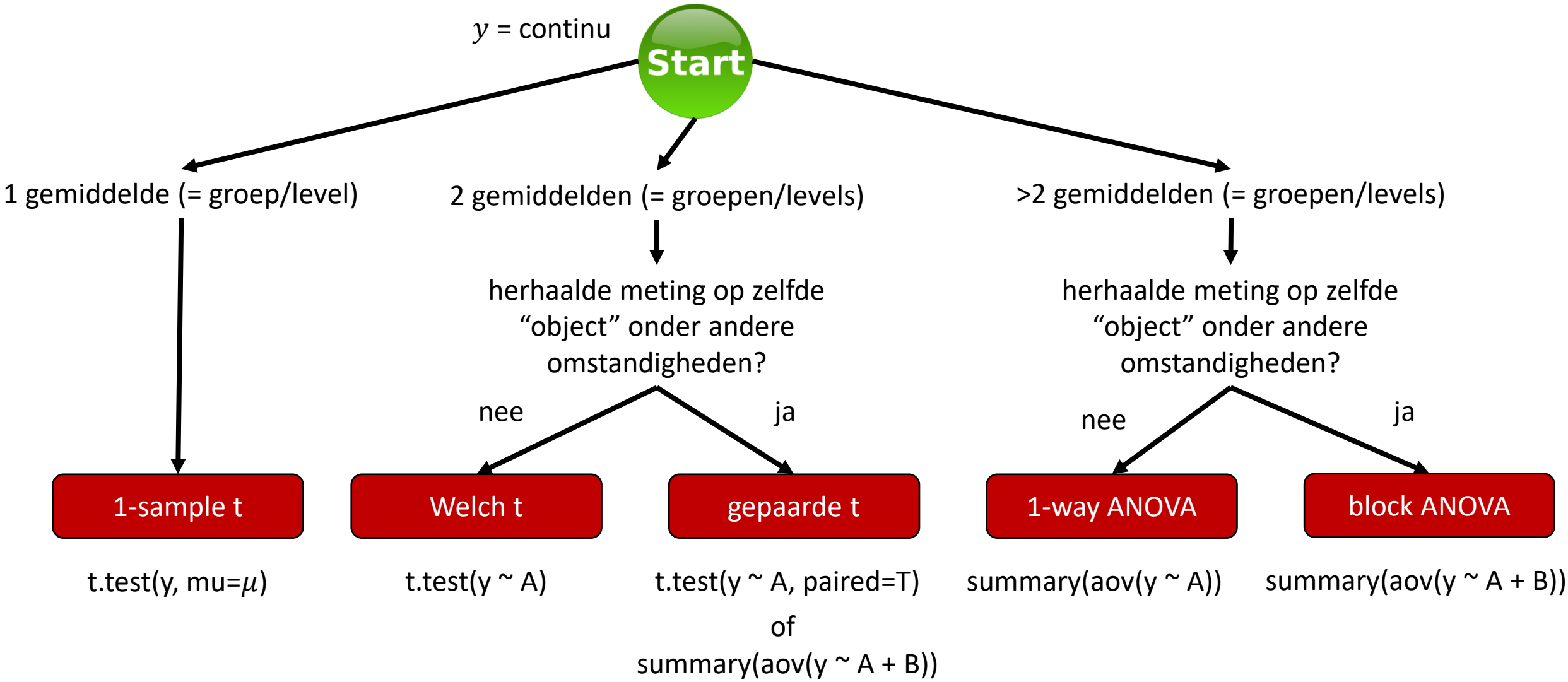
Factor A met $a = 3$ levels

1-way ANOVA

```{r}
( fit.oneway <- summary(aov(E ~ A)) )
```

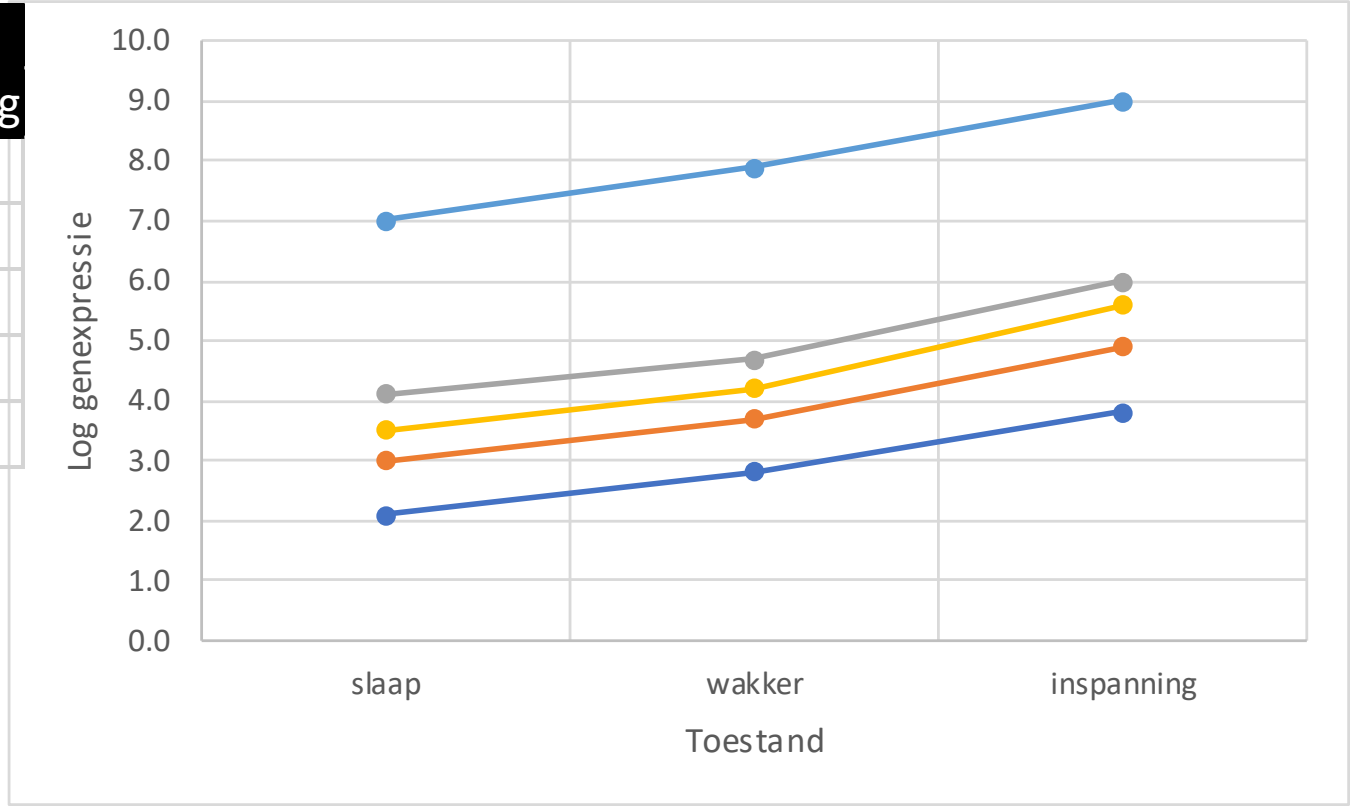
```
            Df Sum Sq Mean Sq F value Pr(>F)
A            2   9.41   4.704   1.281  0.313
Residuals   12  44.06   3.671
```

Keuzeschema toetsen



Block ANOVA

log genexpressie		toestand		
		slaap	wakker	inspanning
diersoort	koe	2.1	2.8	3.8
	geit	3.0	3.7	4.9
	hond	4.1	4.7	6.0
	paard	3.5	4.2	5.6
	mens	7.0	7.9	9.0



Statistisch model:

$$y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

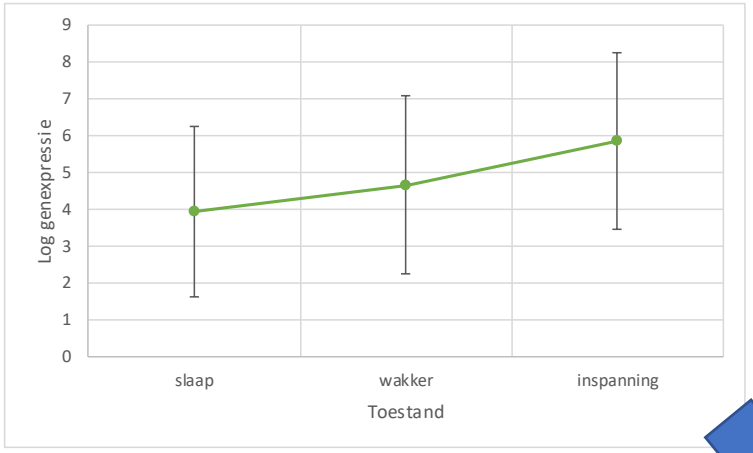
Factor A met $a = 3$ levels

Factor B met $b = 5$ levels

Block ANOVA

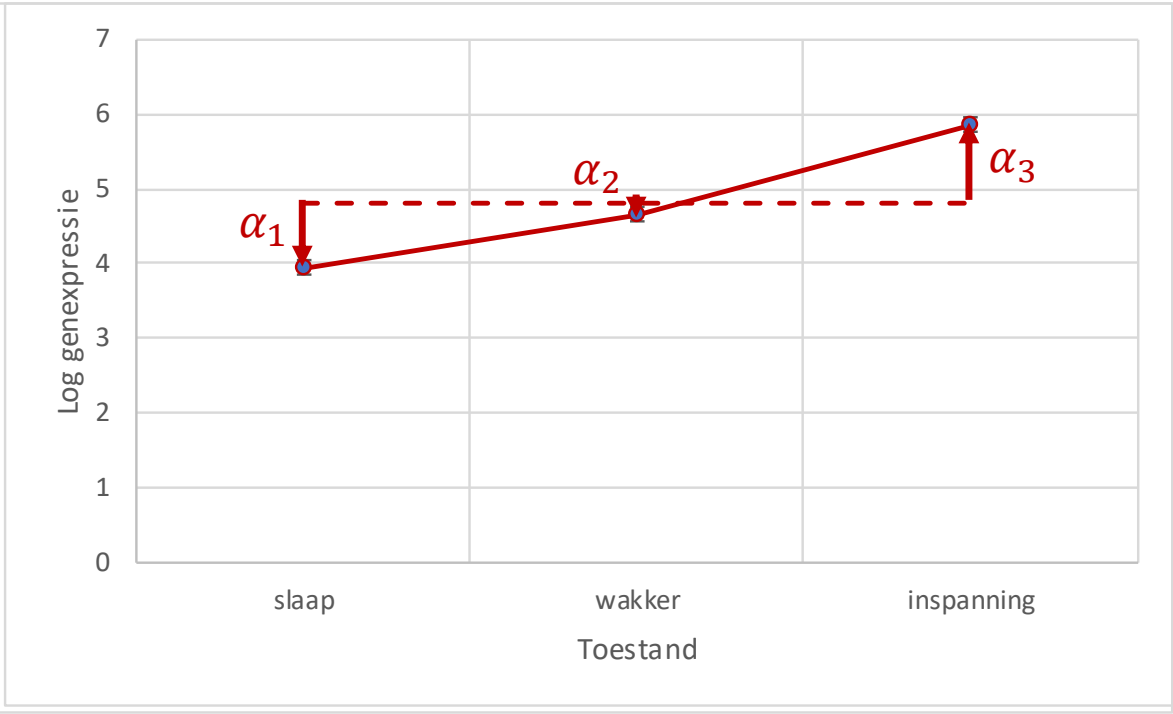
```{r}
( fit.block <- summary(aov(E ~ A + B)) )
```

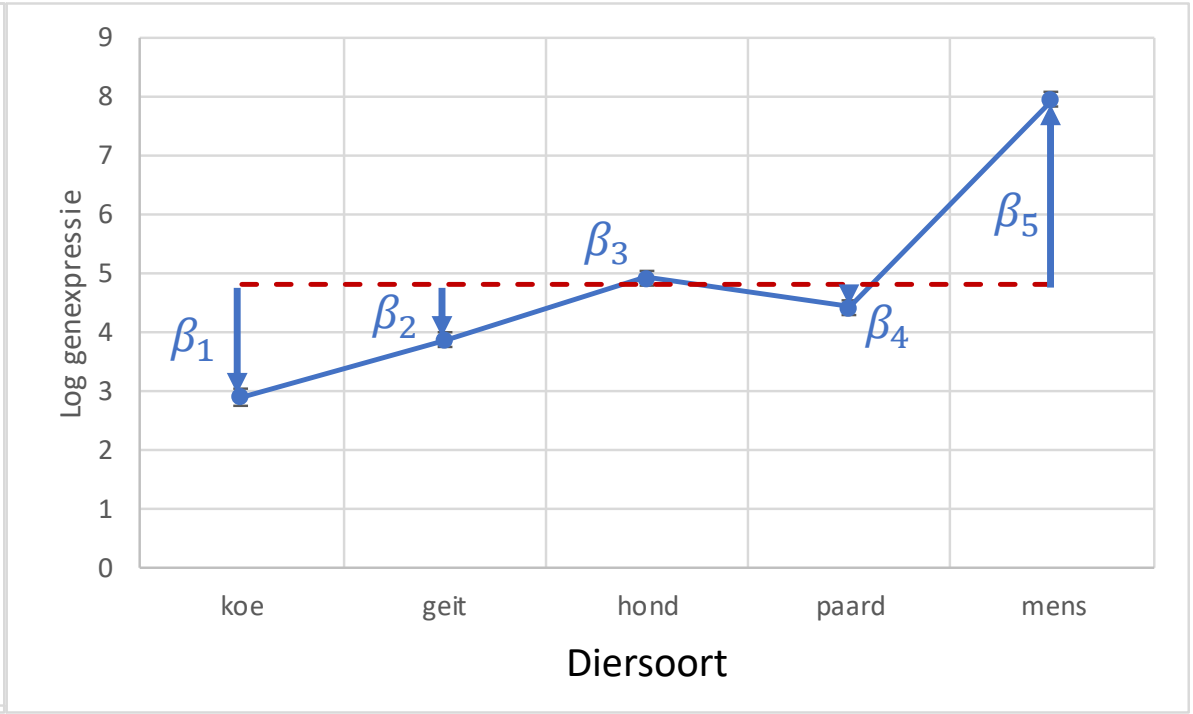
```
            Df Sum Sq Mean Sq F value   Pr(>F)
A            2   9.41   4.704   478.4 4.73e-09 ***
B            4  43.98  10.994  1118.1 5.08e-11 ***
Residuals    8   0.08   0.010
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

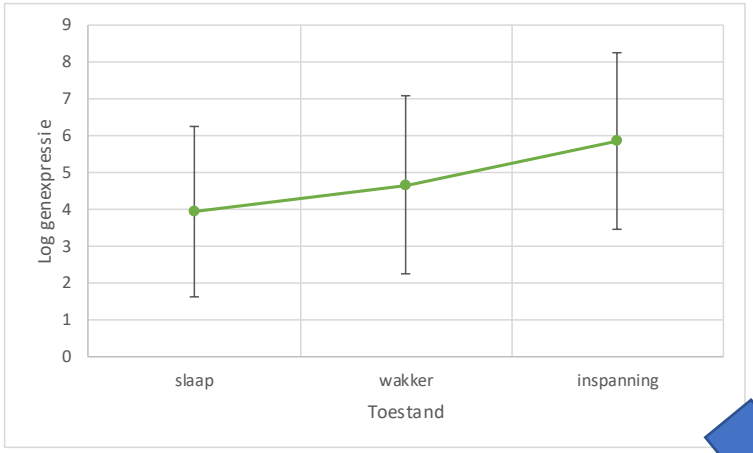


Block vs 1-way ANOVA

$$y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$



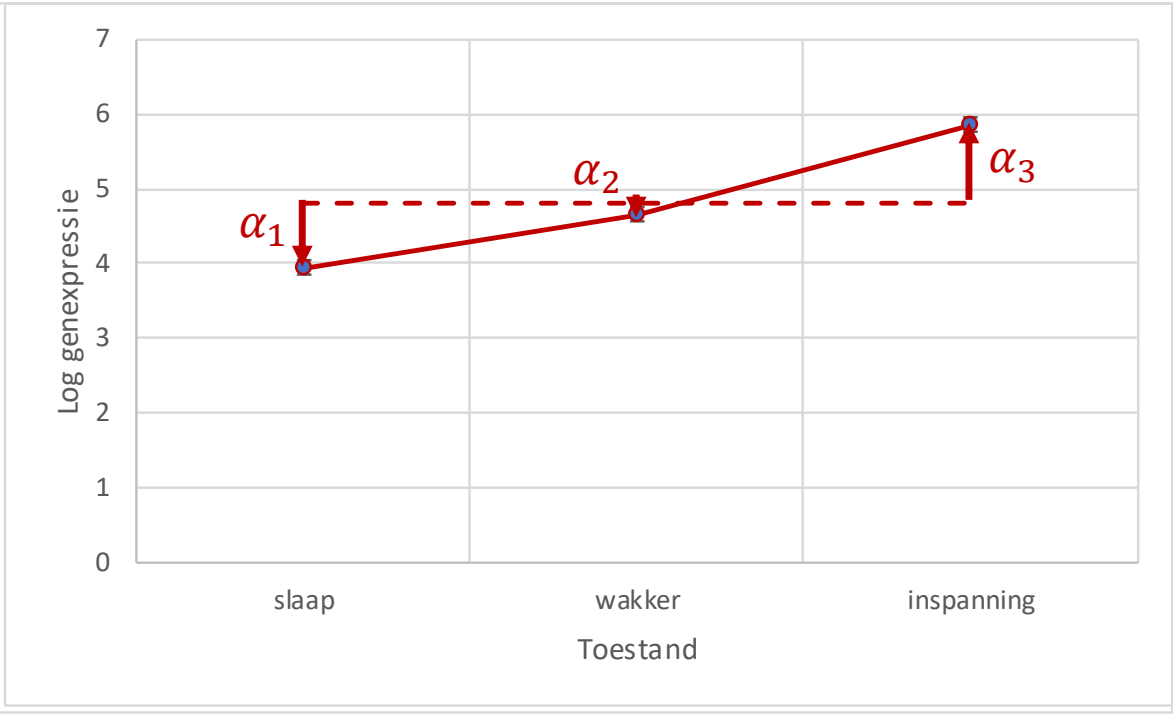


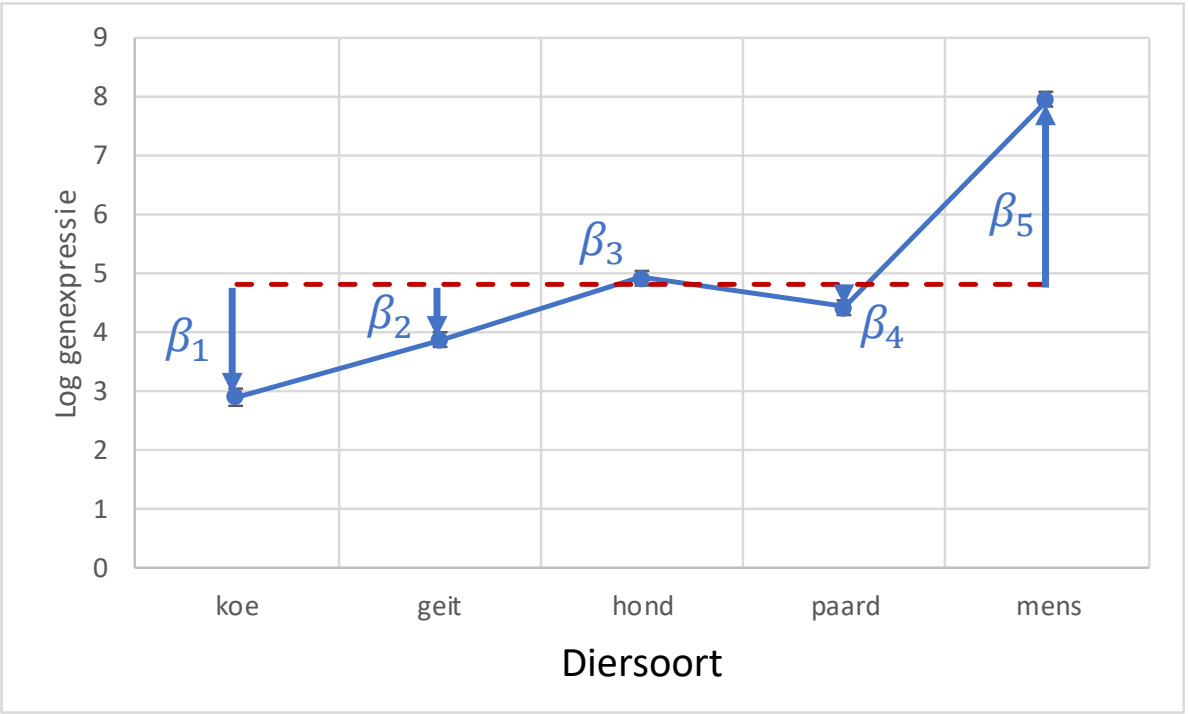


Block vs 1-way ANOVA

1-way: $SS_{\text{tot}} = SS_{\text{A}} + SS_{\text{err}}$

Block: $SS_{\text{tot}} = SS_{\text{A}} + SS_{\text{B}} + SS_{\text{err}}$





Omzetten dataframes van “wide” naar “long”

ID	T	P.1	P.2	P.3
1	24.3	10.2	5.5	2.1
2	23.4	10.4	5.7	2.8
3	22.1	10.5	5.9	3.1
4	19.9	10.2	5.2	2.4

ID	Channel	T	P
1	1	24.3	10.2
2	1	23.4	10.4
3	1	22.1	10.5
4	1	19.9	10.2
1	2	24.3	5.5
2	2	23.4	5.7
3	2	22.1	5.9
4	2	19.9	5.2
1	3	24.3	2.1
2	3	23.4	2.8
3	3	22.1	3.1
4	3	19.9	2.4

Data omzetten van “wide” naar “long” (1)

- Data y met 1 factor A:

Data in "wide" format:

```r
WideData <- data.frame(level1 = c(1, 2, 3),
                       level2 = c(4, 5, 6),
                       level3 = c(7, 8, 9))
WideData
```

level1 <dbl>	level2 <dbl>	level3 <dbl>
1	4	7
2	5	8
3	6	9

Data omzetten van “wide” naar “long” (2)

Data van "wide" naar "long" format:

```{r}
LongData <- stack(WideData)
colnames(LongData) <- c("y", "A")
LongData
```

y <dbl>	A <fctr>
1	level1
2	level1
3	level1
4	level2
5	level2
6	level2
7	level3
8	level3
9	level3

Data omzetten van “wide” naar “long” (3)

Data van "long" naar "wide" format:

```{r}
WideData.2 <- unstack(LongData)
WideData.2
```

level1 <dbl>	level2 <dbl>	level3 <dbl>
1	4	7
2	5	8
3	6	9

Data omzetten van “wide” naar “long” (4)

- Data y met 1 factor A en herhaalde metingen (factor B):

Data in "wide" format:

```{r}
WideData <- data.frame(level1 = c(1, 2, 3),
                       level2 = c(4, 5, 6),
                       level3 = c(7, 8, 9))
rownames(WideData) <- paste0("Sample.",1:3)

WideData
```

	level1 <dbl>	level2 <dbl>	level3 <dbl>
Sample.1	1	4	7
Sample.2	2	5	8
Sample.3	3	6	9

Data omzetten van “short” naar “long” (5)

Data van "wide" naar "long" format:

```{r}
LongData <- stack(WideData)
colnames(LongData) <- c("y", "A")
a <- ncol(WideData)
b <- nrow(WideData)
B <- factor(rep(rownames(WideData), a))
LongData <- cbind(LongData, "B" = B)
LongData
```

y <dbl>	A <fctr>	B <fctr>
1	level1	Sample.1
2	level1	Sample.2
3	level1	Sample.3
4	level2	Sample.1
5	level2	Sample.2
6	level2	Sample.3
7	level3	Sample.1
8	level3	Sample.2
9	level3	Sample.3

Data omzetten van “short” naar “long” (6)

Data van "long" naar "wide" format (naar levels A):

```{r}
# Eerst juist sorteren zodat factor B weer rownames kan worden:
LongData <- LongData[order(LongData$A, LongData$B), ]
# Unstack naar "wide":
WideData.2 <- unstack(LongData, y ~ A)
rownames(WideData.2) <- levels(LongData$B)
WideData.2
```

	level1 <dbl>	level2 <dbl>	level3 <dbl>
Sample.1	1	4	7
Sample.2	2	5	8
Sample.3	3	6	9

Data omzetten van “short” naar “long” (7)

Data van "long" naar "wide" format (naar levels B):

```r
# Eerst juist sorteren zodat factor A weer rownames kan worden:
LongData <- LongData[order(LongData$B, LongData$A), ]
# Unstack naar "wide":
WideData.3 <- unstack(LongData, y ~ B)
rownames(WideData.3) <- levels(LongData$A)
WideData.3
```

	Sample.1 <dbl>	Sample.2 <dbl>	Sample.3 <dbl>
level1	1	2	3
level2	4	5	6
level3	7	8	9