

Tilastollisten mallien peruskurssi – H3

2.

```
> shapiro.test(saasto2013)

      Shapiro-Wilk normality test

data:  saasto2013
W = 0.86229, p-value = 1.628e-12

> shapiro.test(saasto2014)

      Shapiro-Wilk normality test

data:  saasto2014
W = 0.90003, p-value = 2.343e-10

> shapiro.test(saasto2015)

      Shapiro-Wilk normality test

data:  saasto2015
W = 0.93678, p-value = 1.143e-07
```

Kaikissa p-arvo<0.001, joten poiketaan merkittävästi normaalijakaumaoletuksesta.

3.

```
> # Toistettujen mittausten varianssianalyysi
> fit1 = aov(euroa ~vuosi + Error(id/vuosi), data=dat2)
> summary(fit1)

Error: id
      Df Sum Sq Mean Sq F value
Residuals 1 108562328 108562328
Pr(>F)
Residuals

Error: id:vuosi
      Df Sum Sq Mean Sq
vuosi 2 4.762e+09 2.381e+09

Error: within
      Df Sum Sq Mean Sq F value
vuosi 2 1.028e+09 514157711 13.86
Residuals 597 2.215e+10 37104739
Pr(>F)
vuosi 1.31e-06 ***
Residuals
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

F-testin p-arvo<0.001, joten vuosien välillä on tilastollisesti merkitseviä eroja.

4.

```
> t.test(saasto2013,saasto2015, paired=T)

    Paired t-test

data:  saasto2013 and saasto2015
t = 10.373, df = 200, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 5973.927 8778.366
sample estimates:
mean of the differences
      7376.147

> t.test(saasto2014,saasto2015, paired=T)

    Paired t-test

data:  saasto2014 and saasto2015
t = 9.5842, df = 200, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 1783.661 2707.740
sample estimates:
mean of the differences
      2245.7
```

Säästöt 2013 vs säästöt 2015: $p < 0.002$

Säästöt 2014 vs säästöt 2015: $p < 0.002$

5, 6 ja 7.

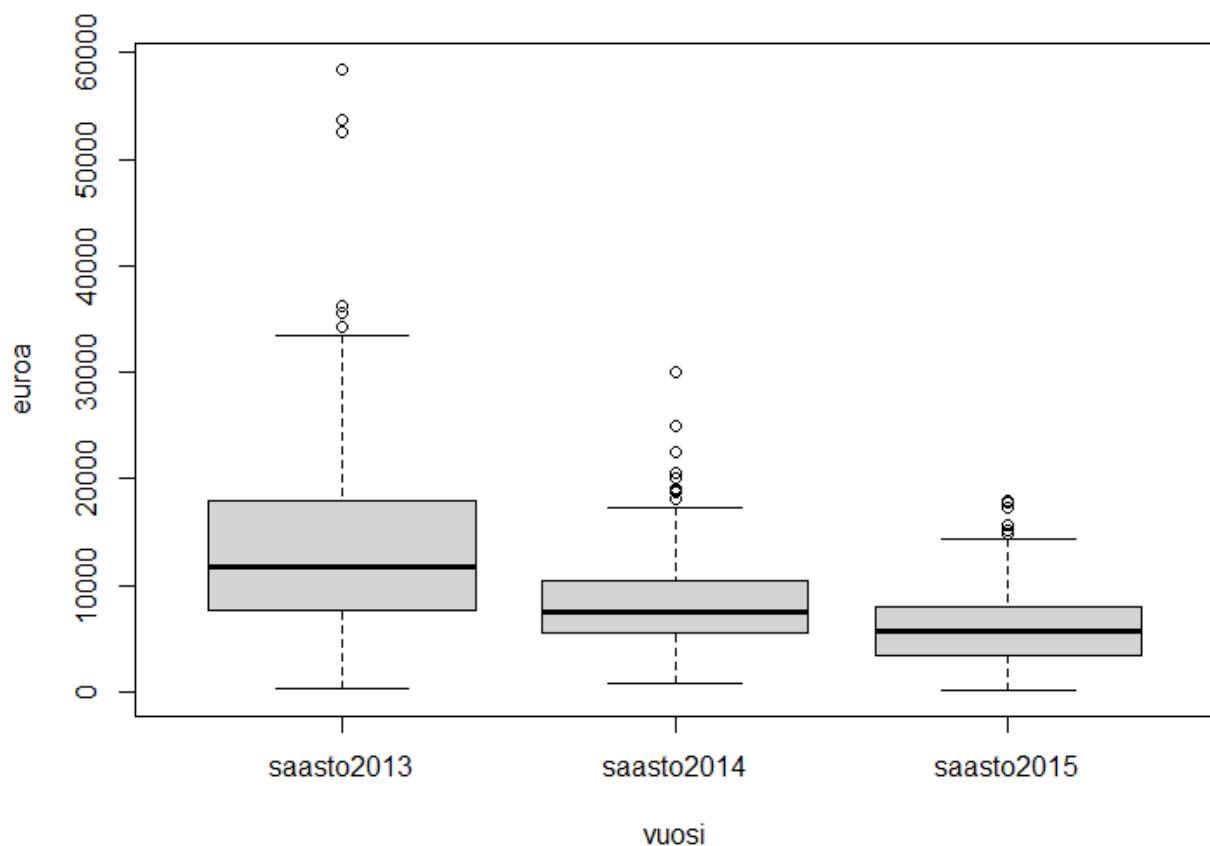
```
> friedman.test(euroa ~vuosi | id, data=dat2)

Friedman rank sum test

data: euroa and vuosi and id
Friedman chi-squared = 100.17, df = 2, p-value < 2.2e-16
```

$p < 0.001$, on tilastollisesti merkitseviä eroja

Friedman boxplot



Wilcoxonin testit

```
> wilcox.test(saasto2013,saasto2015, paired=T)

wilcoxon signed rank test with continuity correction

data: saasto2013 and saasto2015
V = 17927, p-value < 2.2e-16
alternative hypothesis: true location shift is not equal to 0

> wilcox.test(saasto2014,saasto2015, paired=T)

wilcoxon signed rank test with continuity correction

data: saasto2014 and saasto2015
V = 17397, p-value < 2.2e-16
alternative hypothesis: true location shift is not equal to 0
```

Säästöt 2013 vs säästöt 2015 $p < 0.001$

Säästöt 2014 vs säästöt 2015 $p < 0.001$

Normaalijakautumaan perustuva ja epäparametrinen testaus tuottivat saman tuloksen.

8.

```
> with(dat, tapply(saasto2013, taltyyt, shapiro.test))
$`erittäin tyytyväinen`

      shapiro-wilk normality test

data:  X[[i]]
W = 0.95513, p-value = 0.6085

$`melko tyytyväinen`

      shapiro-wilk normality test

data:  X[[i]]
W = 0.88821, p-value = 2.494e-06

$`melko tyytymaton`

      shapiro-wilk normality test

data:  X[[i]]
W = 0.84099, p-value = 6.303e-06

$`erittäin tyytymaton`

      shapiro-wilk normality test

data:  X[[i]]
W = 0.80164, p-value = 9.565e-07
```

Ainoastaan erittäin tyytyväisten joukko noudattaa normaalijakaumaoletusta.

```
> with(dat, tapply(saasto2014, taltyyt, shapiro.test))
$`erittäin tyytyväinen`

      shapiro-wilk normality test

data:  X[[i]]
W = 0.92025, p-value = 0.1943

$`melko tyytyväinen`

      shapiro-wilk normality test

data:  X[[i]]
W = 0.90409, p-value = 1.175e-05

$`melko tyytymaton`

      shapiro-wilk normality test

data:  X[[i]]
W = 0.93562, p-value = 0.007466

$`erittäin tyytymaton`

      shapiro-wilk normality test

data:  X[[i]]
W = 0.89405, p-value = 0.0003075
```

Ainoastaan erittäin tyytyväiset noudattaa normaalijakaumaoletusta.

```

> with(dat, tapply(saasto2015, taltyyt, shapiro.test))
$`erittäin tyytyväinen`

      shapiro-wilk normality test

data:  x[[i]]
W = 0.91341, p-value = 0.1527

$`melko tyytyväinen`

      shapiro-wilk normality test

data:  x[[i]]
W = 0.91606, p-value = 4.149e-05

$`melko tyytymaton`

      shapiro-wilk normality test

data:  x[[i]]
W = 0.92405, p-value = 0.002656

$`erittäin tyytymaton`

      shapiro-wilk normality test

data:  x[[i]]
W = 0.90164, p-value = 0.0005447

```

Ainoastaan erittäin tyytyväiset noudattaa normaalijakaumaoletusta.

9.

```
Error: id
      Df    Sum Sq   Mean Sq
taltyyt 1 108562328 108562328

Error: id:vuosi
      Df    Sum Sq   Mean Sq
vuosi  2 4.762e+09 2.381e+09

Error: within
      Df    Sum Sq   Mean Sq F value    Pr(>F)
taltyyt  3 2.762e+07   9206627    0.255 0.858118
vuosi    2 1.028e+09  514157711   14.215 9.36e-07 ***
taltyyt:vuosi 6 8.553e+08 142553961    3.941 0.000713 ***
Residuals 588 2.127e+10  36171064
---
signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Onko taloudellisen tyytyväisyyden luokkien välillä mittauskerrat yhdistettynä tilastollisesti merkitseviä eroja? $p=0.858$ eli ei ole tilastollisesti merkitseviä eroja

Onko vuosien välillä tilastollisesti merkitseviä eroja? $p<0.001$ eli on tilastollisesti merkitseviä eroja

Onko vuosien ja taloudellisen tyytyväisyyden välillä yhdysvaikutusta? $p<0.001$ eli on yhdysvaikutusta

10.

Erittäin tyytyväiset:

```
> t.test(saasto2013,saasto2015, paired=T)

Paired t-test

data: saasto2013 and saasto2015
t = 1.4059, df = 14, p-value = 0.1816
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1209.494  5811.827
sample estimates:
mean of the differences
      2301.167

> t.test(saasto2014,saasto2015, paired=T)

Paired t-test

data: saasto2014 and saasto2015
t = -3.236, df = 14, p-value = 0.005977
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1877.0139  -380.6527
sample estimates:
mean of the differences
     -1128.833
```

Säästöt 2013 vs säästöt 2015 bonferroni-korjattu $p=1$, ei merkitsevä

Säästöt 2014 vs säästöt 2015 bonferroni-korjattu $p=0.048$, merkitsevä

Melko tyytyväiset:

```
> t.test(saasto2013,saasto2015, paired=T)

Paired t-test

data: saasto2013 and saasto2015
t = 7.0672, df = 83, p-value = 4.516e-10
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  5319.01  9485.49
sample estimates:
mean of the differences
      7402.25

> t.test(saasto2014,saasto2015, paired=T)

Paired t-test

data: saasto2014 and saasto2015
t = 0.8151, df = 83, p-value = 0.4174
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -228.7018  546.3090
sample estimates:
mean of the differences
      158.8036
```

Säästöt 2013 vs säästöt 2015 bonferroni-korjattu $p<0.008$, merkitsevä

Säästöt 2014 vs säästöt 2015 bonferroni-korjattu $p=1$, ei merkitsevä

Melko tyytymättömät:

```
> t.test(saasto2013,saasto2015, paired=T)

Paired t-test

data: saasto2013 and saasto2015
t = 6.4785, df = 51, p-value = 3.627e-08
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 6358.381 12068.648
sample estimates:
mean of the differences
      9213.514

> t.test(saasto2014,saasto2015, paired=T)

Paired t-test

data: saasto2014 and saasto2015
t = 11.086, df = 51, p-value = 3.421e-15
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 3283.440 4735.705
sample estimates:
mean of the differences
      4009.572
```

Säästöt 2013 vs säästöt 2015 bonferroni-korjattu $p < 0.008$, merkitsevä

Säästöt 2014 vs säästöt 2015 bonferroni-korjattu $p < 0.008$, merkitsevä

Erittäin tyytymättömät:

```
> t.test(saasto2013,saasto2015, paired=T)

Paired t-test

data: saasto2013 and saasto2015
t = 4.3717, df = 49, p-value = 6.41e-05
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 3751.926 10135.924
sample estimates:
mean of the differences
      6943.925

> t.test(saasto2014,saasto2015, paired=T)

Paired t-test

data: saasto2014 and saasto2015
t = 10.791, df = 49, p-value = 1.516e-14
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 4011.557 5847.683
sample estimates:
mean of the differences
      4929.62
```

Säästöt 2013 vs säästöt 2015 bonferroni-korjattu $p < 0.008$, merkitsevä

Säästöt 2014 vs säästöt 2015 bonferroni-korjattu $p < 0.008$, merkitsevä