Hypothesetoets voor meer dan twee populatiegemiddeldes

Sandra Van Aert

1 december 2011

Enkelvoudige variantie-analyse

- one-way ANOVA (analysis of variance)
- uitbreiding van de tweezijdige toets voor de gelijkheid van twee gemiddelden:
$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_a : \mu_1 \neq \mu_2 \end{cases}$$
- algemeen: toets voor de gelijkheid van g gemiddeldes
$$\begin{cases} H_0 : \mu_1 = \mu_2 = \cdots = \mu_g = \mu \\ H_a : \text{niet } H_0 \end{cases}$$

Voorbeelden

- vergelijken vochtabsorptie tussen betonsamenstellingen
- vergelijken mortaliteit in functie van socio-economische factoren
- vergelijken vervuilingsgraad voor meerdere rivieren
- vergelijken werking meerdere medicijnen
- vergelijken gewicht in verschillende nesten
- vergelijken smelttemperatuur voor verschillende legeringen

Hoe?

- van elke populatie nemen we een steekproef, zodat we g steekproeven bekomen
- de steekproeven zijn meestal even groot, maar dat hoeft niet
- noteer de steekproefgroottes als $n_1, n_2, n_3, \ldots, n_g$

g steekproeven

steekproef 1	steekproef 2	…	steekproef g
X_{11}	X_{21}		X_{g1}
X_{12}	X_{22}		X_{g2}
$\vdots$	$\vdots$		$\vdots$
X_{1n_1}			X_{gn_g}
	X_{2n_2}		
$\downarrow$	$\downarrow$		$\downarrow$
$\overline{X}_1$	$\overline{X}_2$		$\overline{X}_g$

$$\underbrace{\overline{X}_1 \qquad \overline{X}_2 \qquad \overline{X}_g}_{\overline{X}}$$

Voorbeeld

- plastic folie wordt gefabriceerd op 3 extruders
- om na te gaan of de gemiddelde dikte voor elk van de 3 extruders dezelfde is, wordt op 6 lukraak bepaalde plaatsen in de folies, geproduceerd door elke machine, een meting uitgevoerd:

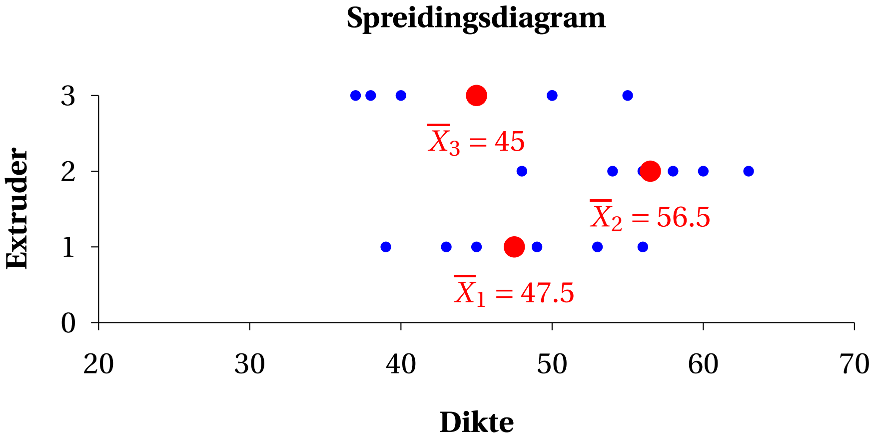
	Meting (in μm)					
	1	2	3	4	5	6
Extruder 1	39	45	43	53	49	56
Extruder 2	56	48	54	58	60	63
Extruder 3	38	37	40	55	50	50

↓

3 niveaus van de factor extruder,

behandelingen of treatments

Voorbeeld

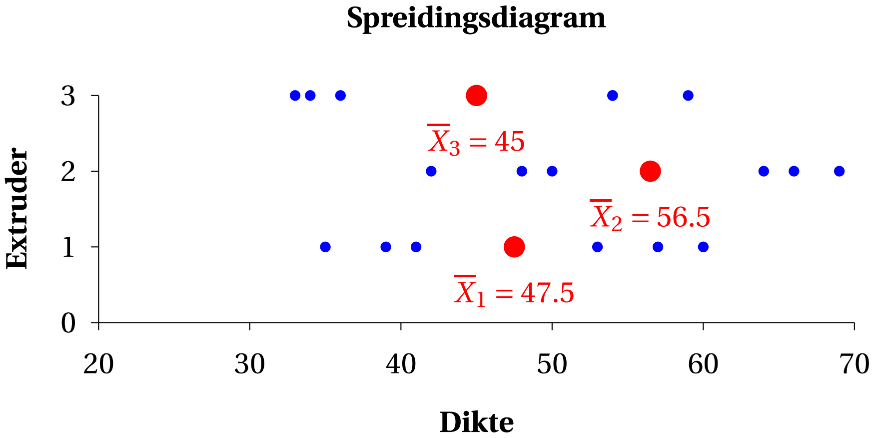

Spreidingsdiagram
Extruder
Dikte
3
2
1
0
20
30
40
50
60
70
$\overline{X}_3 = 45$
$\overline{X}_2 = 56.5$
$\overline{X}_1 = 47.5$

Variant

- plastic folie wordt gefabriceerd op 3 extruders
- om na te gaan of de gemiddelde dikte voor elk van de 3 extruders dezelfde is, wordt op 6 lukraak bepaalde plaatsen in de folies, geproduceerd door elke machine, een meting uitgevoerd:

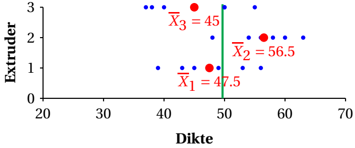
	Meting (in μm)					
	1	2	3	4	5	6
Extruder 1	35	41	39	57	53	60
Extruder 2	50	42	48	64	66	69
Extruder 3	34	33	36	59	54	54

Variant



Spreidingsdiagram
Extruder
3
2
1
0
$\overline{X}_3 = 45$
$\overline{X}_2 = 56.5$
$\overline{X}_1 = 47.5$
20
30
40
50
60
70
Dikte

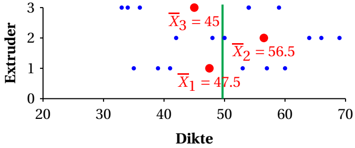
voorbeeld:



gemiddeldes liggen even ver van elkaar in beide gevallen.
algemeen gemiddelde is in beide gevallen identiek:

$$\overline{X} = 49.67$$

variant:

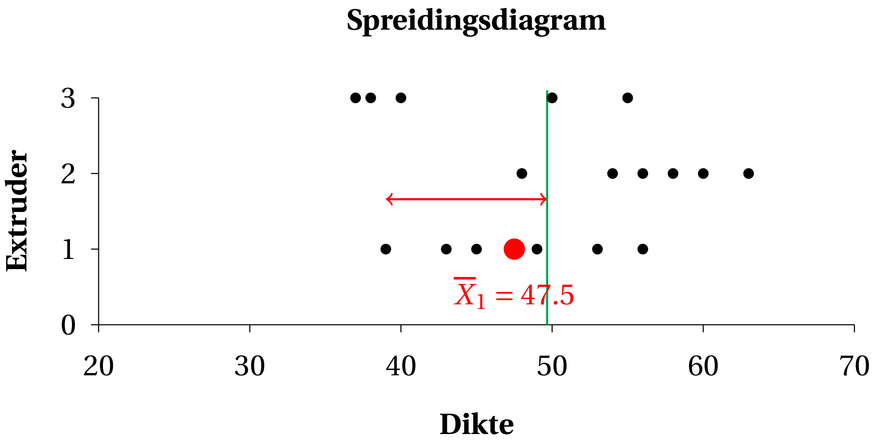


in de eerste figuur is er minder variabiliteit in de gemeten diktes bij elke machine.
daarom geeft de eerste figuur een sterkere indicatie dat er een verschil is in gemiddelde.

Formeel

- totale deviatie: $X_{ij} - \overline{X}$

Formeel

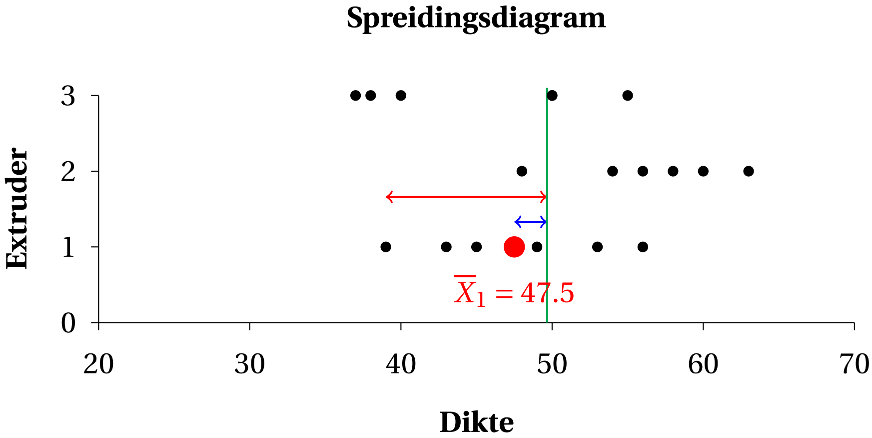


Spreidingsdiagram
Extruder
3
2
1
0
$\overline{X}_1 = 47.5$
20
30
40
50
60
70
Dikte

Formeel

- totale deviatie: $X_{ij} - \overline{X}$
- verschil tussen de machines: $\overline{X}_i - \overline{X}$

Formeel

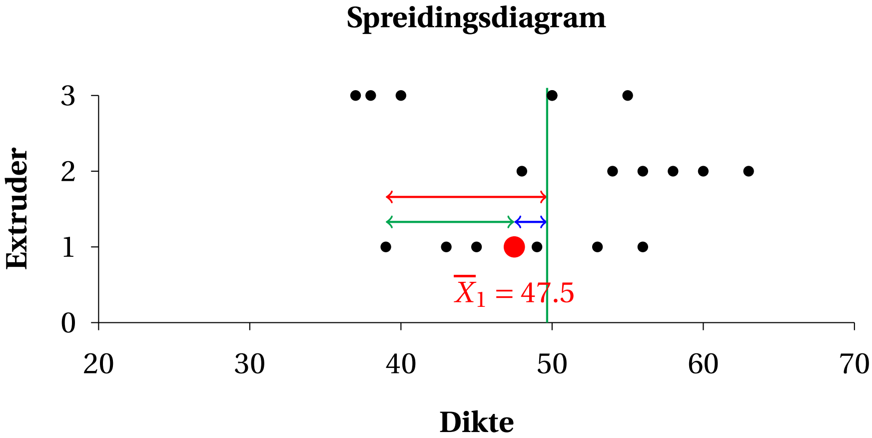


Spreidingsdiagram
Extruder
3
2
1
0
$\overline{X}_1 = 47.5$
20
30
40
50
60
70
Dikte

Formeel

- totale deviatie: $X_{ij} - \overline{X}$
- verschil tussen de machines: $\overline{X}_i - \overline{X}$
- deviaties binnen een machine: $X_{ij} - \overline{X}_i$

Formeel


Spreidingsdiagram
Extruder
Dikte
3
2
1
0
20
30
40
50
60
70
$\overline{X}_1 = 47.5$

Formeel

- totale deviatie: $X_{ij} - \overline{X}$
- verschil tussen de machines: $\overline{X}_i - \overline{X}$
- deviaties binnen een machine: $X_{ij} - \overline{X}_i$
- totale deviatie: $X_{ij} - \overline{X} = \left(\overline{X}_i - \overline{X}\right) + \left(X_{ij} - \overline{X}_i\right)$
- totale kwadraatsom:

$$\sum_{i=1}^{3}\sum_{j=1}^{6}\left(X_{ij}-\overline{X}\right)^2 = \sum_{i=1}^{3}\sum_{j=1}^{6}\left(\overline{X}_i-\overline{X}\right)^2 + \sum_{i=1}^{3}\sum_{j=1}^{6}\left(X_{ij}-\overline{X}_i\right)^2$$

$$\underbrace{\sum_{i=1}^{3}\sum_{j=1}^{6}\left(X_{ij}-\overline{X}\right)^2}_{\text{SST}} = \underbrace{\sum_{i=1}^{3}6\left(\overline{X}_i-\overline{X}\right)^2}_{\text{SSA}} + \underbrace{\sum_{i=1}^{3}\sum_{j=1}^{6}\left(X_{ij}-\overline{X}_i\right)^2}_{\text{SSE}}$$

Algemeen

- totale deviatie: $X_{ij} - \overline{X}$
- verschil tussen de machines: $\overline{X}_i - \overline{X}$
- deviaties binnen een machine: $X_{ij} - \overline{X}_i$
- totale deviatie: $X_{ij} - \overline{X} = \left(\overline{X}_i - \overline{X}\right) + \left(X_{ij} - \overline{X}_i\right)$
- totale kwadraatsom:

$$\sum_{i=1}^{g}\sum_{j=1}^{n_i}\left(X_{ij}-\overline{X}\right)^2 = \sum_{i=1}^{g}\sum_{j=1}^{n_i}\left(\overline{X}_i-\overline{X}\right)^2 + \sum_{i=1}^{g}\sum_{j=1}^{n_i}\left(X_{ij}-\overline{X}_i\right)^2$$

$$\underset{\text{SST}}{\sum_{i=1}^{g}\sum_{j=1}^{n_i}\left(X_{ij}-\overline{X}\right)^2} = \underset{\text{SSA}}{\sum_{i=1}^{g} n_i\left(\overline{X}_i-\overline{X}\right)^2} + \underset{\text{SSE}}{\sum_{i=1}^{g}\sum_{j=1}^{n_i}\left(X_{ij}-\overline{X}_i\right)^2}$$

Terminologie

- SST:
    - totale kwadraatsom
    - total sum of squares
- SSE:
    - binnenkwadraatsom
    - “error” sum of squares
- SSA:
    - tussenkwadraatsom
    - “treatment” sum of squares

E(SSE)

$$
\begin{aligned}
E\left[\sum_{j=1}^{n_i}\left(X_{ij}-\overline{X}_i\right)^2\right] &= E\left[\sum_{j=1}^{n_i}\frac{n_i-1}{n_i-1}\left(X_{ij}-\overline{X}_i\right)^2\right] \\
&= (n_i-1)E\left[\frac{\sum_{j=1}^{n_i}\left(X_{ij}-\overline{X}_i\right)^2}{n_i-1}\right] \\
&= (n_i-1)E[S_i^2] \\
&= (n_i-1)\sigma_i^2 \qquad \text{we veronderstellen dat} \\
&\qquad\qquad\qquad\qquad \sigma_1^2=\sigma_2^2=\cdots=\sigma_g^2=\sigma^2 \\
&= (n_i-1)\sigma^2
\end{aligned}
$$

$$\Rightarrow E[\text{SSE}] = \sum_{i=1}^{g}(n_i-1)\sigma^2 = (n-g)\sigma^2$$

Binnenvariantie MSE

- binnenvariantie

  = gemiddelde binnenkwadraatsom

$$\mathrm{MSE} = \frac{\mathrm{SSE}}{n-g} = \frac{\sum_{i=1}^{g}\sum_{j=1}^{n_i}\left(X_{ij}-\overline{X}_i\right)^2}{n-g}$$

- verwachte waarde:

$$E\left[\mathrm{MSE}\right] = E\left[\frac{\mathrm{SSE}}{n-g}\right] = \sigma^2$$

- indien $X_{ij} \sim N(\mu, \sigma^2)$:

$$\frac{(n-g)\mathrm{MSE}}{\sigma^2} = \frac{\mathrm{SSE}}{\sigma^2} \sim \chi^2_{n-g}$$

Tussenvariantie MSA

- $E[\text{SSA}] = (g-1)\sigma^2 + \sum_{i=1}^{g} n_i(\mu_i - \mu)^2$
- $E[\text{MSA}] = \sigma^2 + \frac{\sum_{i=1}^{g} n_i(\mu_i - \mu)^2}{g-1}$
- als H_0 waar is:

$$E\,[\text{MSA}] = \sigma^2$$

$$\frac{(g-1)\text{MSA}}{\sigma^2} = \frac{\text{SSA}}{\sigma^2} \sim \chi^2_{g-1}$$

F-test

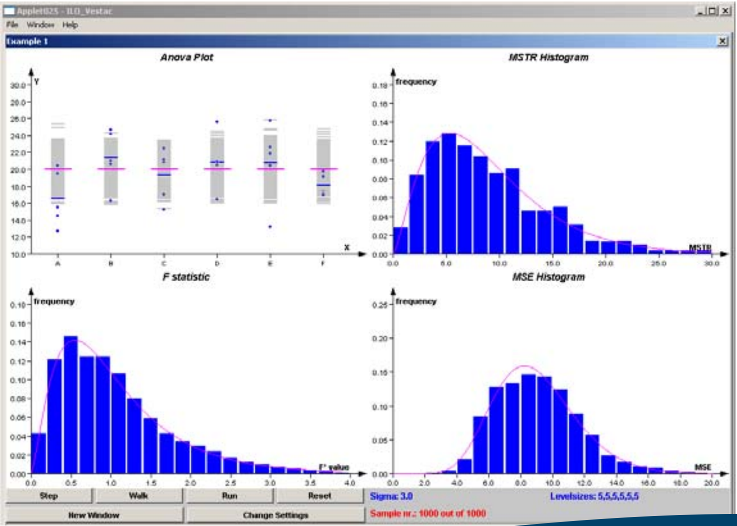
- toetsingsgrootheid:

$$F = \frac{\mathrm{MSA}}{\mathrm{MSE}}$$

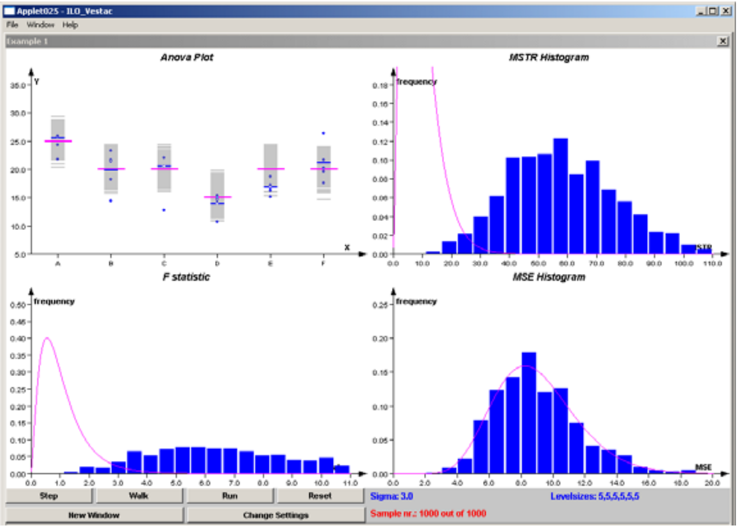
- veronderstelling: alle X_{ij} zijn onafh. en $N(\mu, \sigma^2) \quad (\text{dus } \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_g^2 = \sigma^2)$
- $\dfrac{\mathrm{SSE}}{\sigma^2}$ en $\dfrac{\mathrm{SSA}}{\sigma^2}$ zijn onafh. χ^2 variabelen met resp. $n-g$ en $g-1$ vrijheidsgraden
- $\dfrac{\frac{\mathrm{SSA}/\sigma^2}{g-1}}{\frac{\mathrm{SSE}/\sigma^2}{n-g}} = \dfrac{\mathrm{MSA}}{\mathrm{MSE}}$ is $F_{g-1,n-g}$-verdeeld als H_0 waar

Illustratie: H_0 is waar

http://lstat.kuleuven.be/java/ (Anova/Histograms)



Illustratie: H_0 is onwaar



Beslissingsregel hypothesetoets

- H_0 verwerpen indien

$$\text{MSA} \gg \text{MSE}$$

$$\frac{\text{MSA}}{\text{MSE}} \gg 1$$

toetsingsgrootheid f groot

$$f > \text{ kritieke waarde } F_{\alpha,g-1,n-g}$$

- H_0 verwerpen indien $p < \alpha$
waarbij $p = P(F_{g-1,n-g} \geq f)$

Variantie-analyse tabel: voorbeeld

Source	SS	df	MS	F	p-value
Treatment	439	2	219.500	5.251	0.0187
Error	627	15	41.800		
Total	1066	17			

Variantie-analyse tabel: variant

Source	SS	df	MS	F	p-value
Treatment	439.00	2	219.500	1.752	0.2071
Error	1879.00	15	125.267		
Total	2318.00	17			

Wat als H_0 verworpen wordt?

- uitvissen waar de belangrijke verschillen zitten
  $\rightarrow$ significante verschillen
- alle gemiddeldes twee per twee toetsen
  $$\begin{cases} H_0 : \mu_i = \mu_j \\ H_a : \mu_i \neq \mu_j \end{cases}$$
  met behulp van t-toets

Probleem

- stel 3 paarsgewijze toetsen:

$$\mu_1 \text{ versus } \mu_2 \rightarrow 1-\alpha \text{ betrouwbaarheid}$$
$$\mu_1 \text{ versus } \mu_3 \rightarrow 1-\alpha$$
$$\mu_2 \text{ versus } \mu_3 \rightarrow 1-\alpha$$

- gezamenlijke betrouwbaarheid: $(1-\alpha)^3$
- voorbeeld:
  $\alpha = 0.05 \Rightarrow 1-\alpha = 0.95 \Rightarrow (1-\alpha)^3 = 0.857 \neq 0.95$
- je wil gezamenlijke betrouwbaarheid van bijvoorbeeld 0.95

Oplossing

- kies kleinere α
- gemiddeldes twee per twee toetsen
- Tukey methode
- simultane constructie van betrouwbaarheidsintervallen zodat de gezamenlijke betrouwbaarheid gelijk is aan 95%

Tukey methode (R)

```
TukeyHSD(aov(Meting~Extruder))
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Meting ~ Extruder)

$Extruder
     diff        lwr       upr     p adj
2-1   9.0  -0.695676 18.695676 0.0707885
3-1  -2.5 -12.195676  7.195676 0.7842171
3-2 -11.5 -21.195676 -1.804324 0.0196076
```

Tukey methode (Matlab)

```
multcompare(stats)

ans =

    1.0000    2.0000  -18.6957   -9.0000    0.6957
    1.0000    3.0000   -7.1957    2.5000   12.1957
    2.0000    3.0000    1.8043   11.5000   21.1957
```

Testen van voorwaarden voor ANOVA

- normaliteit van de residuele waarden
  kwantieldiagram, Shapiro Wilk toets, Lilliefors toets
- gelijkheid van varianties
  Bartlett test

Normaliteit - Shapiro Wilk toets (R)

```
qqnorm(lm1$residuals)

shapiro.test(lm1$residuals)

        Shapiro-Wilk normality test

data:  lm1$residuals
W = 0.9388, p-value = 0.2760
```

normaliteit aanvaarden indien p-waarde groter dan α (=0.05)

Normaliteit - Lilliefors toets (Matlab)

```
qqplot(residuals(:))

 [h,p]=lillietest(residuals(:))

h =

     0


p =

    0.5000
```

normaliteit aanvaarden indien p-waarde groter dan α (=0.05)

Gelijkheid varianties (R)

```
bartlett.test(Meting~Extruder)

        Bartlett test of homogeneity of variances

data:  Meting by Extruder
Bartlett's K-squared = 0.6401, df = 2, p-value = 0.7261
```

gelijkheid van varianties aanvaarden indien p-waarde groter dan α (=0.05)

Gelijkheid varianties (Matlab)

```
vartestn(X)

Group                    Count       Mean      Std Dev
----------------------------------------------------
1                        6          47.5       6.37966
2                        6          56.5       5.20577
3                        6          45         7.58947
Pooled                  18          49.6667    6.46529

Bartlett's statistic     0.64012
Degrees of freedom       2
p-value                  0.7261
```

gelijkheid van varianties aanvaarden indien p-waarde groter dan α (=0.05)