02402 Statistik (Polyteknisk grundlag)

Uge 11: Ensidet variansanalyse - ANOVA

Nicolai Siim Larsen
DTU Compute
Danmarks Tekniske Universitet
2800 Kgs. Lyngby

Dagsorden

Variansanalyse - ANOVA

"ANalysis Of VAriance"(ANOVA) blev introduceret af R.A. Fisher for ca. 100 år siden som en systematisk måde at analysere grupper på og har siden da været helt centralt i statistik og anvendelser deraf.

- I dag: Et inddelingskriterium (ensidet ANOVA)
- Næste uge: To inddelingskriterier (tosidet ANOVA)
- Inddelingskriterium = faktor
- Første faktor kaldes typisk *treatment*, anden faktor *block*

Dagsorden

Ensidet variansanalyse - Eksempel

Gruppe A	Gruppe B	Gruppe C
2.8	5.5	5.8
3.6	6.3	8.3
3.4	6.1	6.9
2.3	5.7	6.1

Er der forskel (i middelværdien) på grupperne A, B og C?

Variansanalyse (ANOVA) kan anvendes til analysen, såfremt observationerne i hver gruppe kan antages at være normalfordelte.

Envejs variansanalyse – eksempel i R

```
# Indlæs data
y <- c(2.8, 3.6, 3.4, 2.3,
       5.5, 6.3, 6.1, 5.7,
       5.8, 8.3, 6.9, 6.1)

# Definer (treatment) grupper
treatm <- factor(c(1, 1, 1, 1,
                   2, 2, 2, 2,
                   3, 3, 3, 3))

# Plot data mod grupperne
par(mfrow = c(1,2))
plot(y ~ as.numeric(treatm), xlab = "Gruppe (Treatment)", ylab = "Værdi")
boxplot(y ~ treatm, xlab = "Gruppe (Treatment)", ylab = "Værdi")
```


Dagsorden

Ensidet variansanalyse - Model

- Modellen kan opskrives som

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}\,,$$

  hvor det antages ε_{ij} er i.i.d. med

$$\varepsilon_{ij} \sim N(0, \sigma^2)\,.$$

- μ er den samlede middelværdi
- α_i angiver effekten af gruppe (treatment) i
- Y_{ij} er måling j i gruppe i (j går fra 1 til n_i)

Ensidet variansanalyse - Hypotesetest

- Vi vil nu sammenligne (flere end to) middelværdier $(\mu + \alpha_i)$ i modellen

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}, \quad \varepsilon_{ij} \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2).$$

- Nulhypotesen er givet ved:

$$H_0: \quad \alpha_i = 0 \quad \text{for alle } i.$$

- Modhypotesen (alternativhypotesen) er givet ved:

$$H_1: \quad \alpha_i \neq 0 \quad \text{for mindst et } i.$$

Dagsorden

Ensidet variansanalyse - Dekomposition og ANOVA-tabellen

- Med modellen

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}, \quad \varepsilon_{ij} \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2)$$

kan den totale variation i data opspaltes:

$$SST = SS(Tr) + SSE.$$

- 'Ensidet' hentyder til, at der kun er én faktor i forsøget (med k niveauer).
- Metoden kaldes variansanalyse, fordi testningen foregår ved at sammenligne varianser.

Formler for kvadratafvigelsessummer

- Den samlede variation

$$SST = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2$$

- Variation inden for grupperne (Variation tilbage efter model, dvs. af residualerne)

$$SSE = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$$

- Variation mellem grupperne (Variantion forklaret af modellen)

$$SS(Tr) = \sum_{i=1}^{k} n_i (\bar{y}_i - \bar{y})^2$$

Ensidet variansanalyse - Parameterestimater

- $\hat{\mu} = \bar{y}$
- $\hat{\alpha}_i = \bar{y}_i - \bar{y}$
- $\hat{\sigma}^2 = MSE = \frac{SSE}{n-k}$

```
# Samlet gennemsnit
mean(y)

## [1] 5.233

# Gruppegennemsnit
tapply(y, treatm, mean)

##     1     2     3
## 3.025 5.900 6.775

# SSE: Brug anova(..)
```


Variansanalyseskema

Source of variation	Deg. of freedom	Sums of squares	Mean sum of squares
Treatment	$k-1$	$SS(Tr)$	$MS(Tr) = \frac{SS(Tr)}{k-1}$
Residual	$n-k$	SSE	$MSE = \frac{SSE}{n-k}$
Total	$n-1$	SST	

```
# Ensidet ANOVA med anova() og lm()
anova(lm(y ~ treatm))

## Analysis of Variance Table
##
## Response: y
##           Df Sum Sq Mean Sq F value  Pr(>F)
## treatm     2   30.8   15.40    26.7 0.00017 ***
## Residuals  9    5.2    0.58
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```


Dagsorden

Envejs variansanalyse - F-test

- Vi har (Sætning 8.2)

$$SST = SS(Tr) + SSE$$

- Herfra kan udlede teststørrelsen:

$$F = \frac{SS(Tr)/(k-1)}{SSE/(n-k)} = \frac{MS(Tr)}{MSE},$$

  hvor
  - k er antal nivauer af faktoren,
  - n er antal observationer.
- Vælg et signifikansniveau α og beregn teststørrelsen F.
- Sammenlign teststørrelsen med $(1-\alpha)$-fraktilen i F-fordelingen:

$$F \sim F(k-1, n-k) \text{ (Sætning 8.6)}$$

F-fordelingen og F-testen

```
# Under H0:

# Antal grupper
k <- 3

# Antal observationer
n <- 12

# Talrække plot
xseq <- seq(0, 10, by = 0.1)

# Plot tætheden for F-fordelingen
plot(xseq, df(xseq, df1 = k-1, df2 = n-k), type = "l", xlab = "x", ylab = "f(x)")

# Plot kritiske værdier for 5%-signifikansniveauet
cr <- qf(0.95, df1 = k-1, df2 = n-k)
abline(v = cr, col = "red")
```


Variansanalyseskema

Source of variation	Deg. of freedom	Sums of squares	Mean sum of squares	Test-statistic F	p-value
treatment	$k-1$	$SS(Tr)$	$MS(Tr) = \frac{SS(Tr)}{k-1}$	$F_{\text{obs}} = \frac{MS(Tr)}{MSE}$	$P(F > F_{\text{obs}})$
Residual	$n-k$	SSE	$MSE = \frac{SSE}{n-k}$		
Total	$n-1$	SST			

```
anova(lm(y ~ treatm))

## Analysis of Variance Table
##
## Response: y
##           Df Sum Sq Mean Sq F value  Pr(>F)
## treatm     2   30.8   15.40    26.7 0.00017 ***
## Residuals  9    5.2    0.58
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```


Ensidet ANOVA F-test “i hånden”

```
k <- 3; n <- 12  # Antal grupper og observationer

# Samlet variation: SST
SST <- sum((y - mean(y))^2)

# Variation af residualerne (inden for grupperne): SSE
y1 <- y[1:4]; y2 <- y[5:8]; y3 <- y[9:12]

SSE <- sum((y1 - mean(y1))^2) +
       sum((y2 - mean(y2))^2) +
       sum((y3 - mean(y3))^2)

# Variation forklaret af modellen/grupperingen (mellem grupperne): SS(Tr)
SSTr <- SST - SSE

# Teststørrelsen
Fobs <- (SSTr/(k-1)) / (SSE/(n-k))

# P-værdien
1 - pf(Fobs, df1 = k-1, df2 = n-k)
```


Dagsorden

Variabilitet og sammenhæng med t-testen for to stikprøver (Sætning 8.4)

Residualkvadratafvigelsessummen, SSE, divideret med $n-k$, også kaldet middelkvadratafvigelsen $MSE = SSE/(n-k)$, er et vægtet gennemsnit af stikprøvevarianserne for grupperne:

$$MSE = \frac{SSE}{n-k} = \frac{(n_1-1)s_1^2 + \cdots + (n_k-1)s_k^2}{n-k},$$

$$s_i^2 = \frac{1}{n_i-1}\sum_{i=1}^{n_i}(y_{ij}-\bar{y}_i)^2.$$

KUN når $k=2$: (jf. Metode 3.52)

$$MSE = s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n-2},$$

$$F_{\text{obs}} = t_{\text{obs}}^2,$$

hvor t_{obs} er den sammenvejede t-teststørrelse fra Metode 3.52 og 3.53.

Dagsorden

Post hoc konfidensinterval – Metode 8.9

- En enkelt *forudplanlagt* sammenligning af forskellen på behandling i og j findes ved:

$$\bar{y}_i - \bar{y}_j \pm t_{1-\alpha/2}\sqrt{\frac{SSE}{n-k}\left(\frac{1}{n_i}+\frac{1}{n_j}\right)},$$

  hvor $t_{1-\alpha/2}$ er fra t-fordelingen med $n-k$ frihedsgrader.

- Bemærk de færre frihedsgrader, da der estimeres flere parametre i beregningen af $MSE = SSE/(n-k) = s_p^2$ (det sammenvejede variansestimat)
- Hvis alle $M = k(k-1)/2$ kombinationer af parvise konfidensintervaller udregnes, så brug formlen M gange, men hver gang med $\alpha_{\text{Bonferroni}} = \alpha/M$.

Post hoc parvis hypotesetest – Metode 8.10

- For en enkelt *forudplanlagt* hypotesetest

$$H_0:\ \mu_i = \mu_j,\ \ H_1:\ \mu_i \neq \mu_j$$

  på niveau α, benyttes teststørrelsen

$$t_{\text{obs}} = \frac{\bar{y}_i - \bar{y}_j}{\sqrt{MSE\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}}$$

  og p-værdien

$$p = 2P(T > |t_{\text{obs}}|),$$

  hvor t-fordelingen med $n-k$ frihedsgrader anvendes.

- Hvis alle $M = k(k-1)/2$ kombinationer af parvise hypotesetest udføres, så bruges det korrigerede signifikansniveau $\alpha_{\text{Bonferroni}} = \alpha/M$.

Dagsorden

1. Introduktion
2. Model og hypoteser
3. Beregning: Variansdekomposition og ANOVA-tabellen
4. Hypotesetest (F-test)
5. Variabilitet og sammenhæng med t-testen for to stikprøver
6. Post hoc sammenligninger
7. **Modelkontrol**
8. Et gennemregnet eksempel – fra bogen

Varianshomogenitet

Se på box-plottet om spredningen ser (meget) forskellig ud for hver gruppe

```
# Tjek antagelsen om varianshomogenitet
plot(treatm, y)
```


Normalfordelingsantagelsen

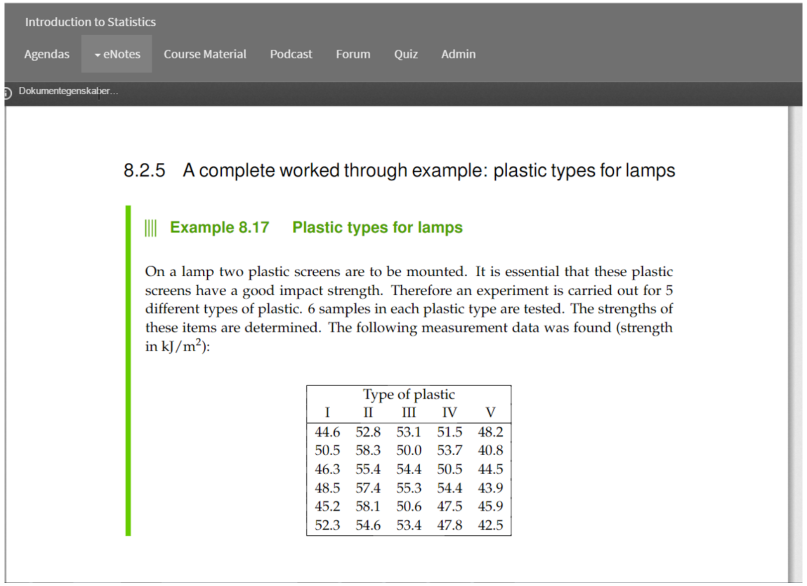
Se normalfordelings-QQ-plottet af residualerne:

```
# Tjek antagelsen om normalfordeling
fit1 <- lm(y ~ treatm)
qqnorm(fit1$residuals)
qqline(fit1$residuals)
```


Dagsorden

1. Introduktion
2. Model og hypoteser
3. Beregning: Variansdekomposition og ANOVA-tabellen
4. Hypotesetest (F-test)
5. Variabilitet og sammenhæng med t-testen for to stikprøver
6. Post hoc sammenligninger
7. Modelkontrol
8. **Et gennemregnet eksempel – fra bogen**

Et gennemregnet eksempel – fra bogen



8.2.5 A complete worked through example: plastic types for lamps

Example 8.17 Plastic types for lamps

On a lamp two plastic screens are to be mounted. It is essential that these plastic screens have a good impact strength. Therefore an experiment is carried out for 5 different types of plastic. 6 samples in each plastic type are tested. The strengths of these items are determined. The following measurement data was found (strength in kJ/m^2):

Type of plastic				
I	II	III	IV	V
44.6	52.8	53.1	51.5	48.2
50.5	58.3	50.0	53.7	40.8
46.3	55.4	54.4	50.5	44.5
48.5	57.4	55.3	54.4	43.9
45.2	58.1	50.6	47.5	45.9
52.3	54.6	53.4	47.8	42.5

Dagsorden

1. Introduktion
2. Model og hypoteser
3. Beregning: Variansdekomposition og ANOVA-tabellen
4. Hypotesetest (F-test)
5. Variabilitet og sammenhæng med t-testen for to stikprøver
6. Post hoc sammenligninger
7. Modelkontrol
8. Et gennemregnet eksempel – fra bogen