STK1110 Høsten 2024

Enveis variansanalyse

Tilsvarer Avsnitt 11.1 og 11.3 (s. 662-663)

Enveis variansanalyse

- Kap 8-9: Ett utvalg
- Kap 10: To utvalg
- Kap 11: Flere (> 2) utvalg
  - Ønsker å test om det er forskjell mellom utvalgene

Eksempel

- For å undersøke opptak av fett fra ulike meieriprodukter, ble 30 rotter fôret med ulike produkter.

Eksempel

- For å undersøke opptak av fett fra ulike meieriprodukter, ble 30 rotter fôret med ulike produkter.
- Rottene ble delt inn i fem like store grupper på seks.

Eksempel

- For å undersøke opptak av fett fra ulike meieriprodukter, ble 30 rotter fôret med ulike produkter.
- Rottene ble delt inn i fem like store grupper på seks.
- Hver gruppe ble fôret med ett av følgende meieriprodukter:

Eksempel

- For å undersøke opptak av fett fra ulike meieriprodukter, ble 30 rotter fôret med ulike produkter.
- Rottene ble delt inn i fem like store grupper på seks.
- Hver gruppe ble fôret med ett av følgende meieriprodukter:
  1. kremost

Eksempel

- For å undersøke opptak av fett fra ulike meieriprodukter, ble 30 rotter fôret med ulike produkter.
- Rottene ble delt inn i fem like store grupper på seks.
- Hver gruppe ble fôret med ett av følgende meieriprodukter:
  1. kremost
  2. rømme

Eksempel

- For å undersøke opptak av fett fra ulike meieriprodukter, ble 30 rotter fôret med ulike produkter.
- Rottene ble delt inn i fem like store grupper på seks.
- Hver gruppe ble fôret med ett av følgende meieriprodukter:
  1. kremost
  2. rømme
  3. fløte

Eksempel

- For å undersøke opptak av fett fra ulike meieriprodukter, ble 30 rotter fôret med ulike produkter.
- Rottene ble delt inn i fem like store grupper på seks.
- Hver gruppe ble fôret med ett av følgende meieriprodukter:
  1. kremost
  2. rømme
  3. fløte
  4. smørbart smør (smør innblandet 25% vegetabilsk olje)

Eksempel

- For å undersøke opptak av fett fra ulike meieriprodukter, ble 30 rotter fôret med ulike produkter.
- Rottene ble delt inn i fem like store grupper på seks.
- Hver gruppe ble fôret med ett av følgende meieriprodukter:
  1. kremost
  2. rømme
  3. fløte
  4. smørbart smør (smør innblandet 25% vegetabilsk olje)
  5. smør

Eksempel

- For å undersøke opptak av fett fra ulike meieriprodukter, ble 30 rotter fôret med ulike produkter.
- Rottene ble delt inn i fem like store grupper på seks.
- Hver gruppe ble fôret med ett av følgende meieriprodukter:
  1. kremost
  2. rømme
  3. fløte
  4. smørbart smør (smør innblandet 25% vegetabilsk olje)
  5. smør
- Hver rotte fikk en mengde av meieriproduktet tilsvarende 300 mg fett, og totalt opptak av fett (i mg) i løpet av 8 timer ble målt (via mengden av fettsyrer i lymfeårene).

Eksempel

- For å undersøke opptak av fett fra ulike meieriprodukter, ble 30 rotter fôret med ulike produkter.
- Rottene ble delt inn i fem like store grupper på seks.
- Hver gruppe ble fôret med ett av følgende meieriprodukter:
  1. kremost
  2. rømme
  3. fløte
  4. smørbart smør (smør innblandet 25% vegetabilsk olje)
  5. smør
- Hver rotte fikk en mengde av meieriproduktet tilsvarende 300 mg fett, og totalt opptak av fett (i mg) i løpet av 8 timer ble målt (via mengden av fettsyrer i lymfeårene).
- Er opptaket av fett det samme for alle typer meieriprodukter?

Enveis variansanalyse (ANOVA)

- Enveis variansanalyse: Sammenligning av forventningsverdien i flere populasjoner,
  - populasjonene er definert ut fra én bestemt faktor.

Enveis variansanalyse (ANOVA)

- Enveis variansanalyse: Sammenligning av forventningsverdien i flere populasjoner,
  - populasjonene er definert ut fra én bestemt faktor.
- Anta
  - I grupper (behandlinger)
  - J observasjoner i hver (for enkelhets skyld er det like mange observasjoner i hver gruppe).

Enveis variansanalyse (ANOVA)

- Enveis variansanalyse: Sammenligning av forventningsverdien i flere populasjoner,
  - populasjonene er definert ut fra én bestemt faktor.
- Anta
  - I grupper (behandlinger)
  - J observasjoner i hver (for enkelhets skyld er det like mange observasjoner i hver gruppe).
- Anta
  - $X_{i1}, \ldots, X_{iJ} \overset{uif}{\sim} N(\mu_i, \sigma^2)$, der X_{ij} er j-te måling i gruppe i,
    - Felles σ^2 for alle grupper!
  - X_{ij}-ene er uavhengige på tvers av grupper.

Enveis variansanalyse (ANOVA)

- Enveis variansanalyse: Sammenligning av forventningsverdien i flere populasjoner,
  - populasjonene er definert ut fra én bestemt faktor.
- Anta
  - I grupper (behandlinger)
  - J observasjoner i hver (for enkelhets skyld er det like mange observasjoner i hver gruppe).
- Anta
  - $X_{i1}, \ldots, X_{iJ} \overset{uif}{\sim} N(\mu_i, \sigma^2)$, der X_{ij} er j-te måling i gruppe i,
    - Felles σ^2 for alle grupper!
  - X_{ij}-ene er uavhengige på tvers av grupper.
- Vi ønsker å teste

$H_0 : \mu_1 = \ldots = \mu_I = \mu$ mot H_a : minst to μ_i-er er forskjellige.

Eksempel 11.1

- Flere forskjellige typer bokser ble sammenlignet med hensyn til trykkstyrke

Table 11.1 The data and summary quantities for Example 11.1

Type of box	Compression strength (lb)			Sample mean	Sample SD
1	655.5	788.3	734.3	713.00	46.55
	721.4	679.1	699.4		
2	789.2	772.5	786.9	756.93	40.34
	686.1	732.1	774.8		
3	737.1	639.0	696.3	698.07	37.20
	671.7	717.2	727.1		
4	535.1	628.7	542.4	562.02	39.87
	559.0	586.9	520.0		
			Grand mean =	682.50	

Eksempel 11.1

- Flere forskjellige typer bokser ble sammenlignet med hensyn til trykkstyrke

Table 11.1 The data and summary quantities for Example 11.1

Type of box	Compression strength (lb)			Sample mean	Sample SD
1	655.5	788.3	734.3	713.00	46.55
	721.4	679.1	699.4		
2	789.2	772.5	786.9	756.93	40.34
	686.1	732.1	774.8		
3	737.1	639.0	696.3	698.07	37.20
	671.7	717.2	727.1		
4	535.1	628.7	542.4	562.02	39.87
	559.0	586.9	520.0		
			Grand mean =	682.50	

Her er det altså $I = 4$ grupper med $J = 6$ observasjoner i hver gruppe.

Eksempel 11.1

- Flere forskjellige typer bokser ble sammenlignet med hensyn til trykkstyrke

Table 11.1 The data and summary quantities for Example 11.1

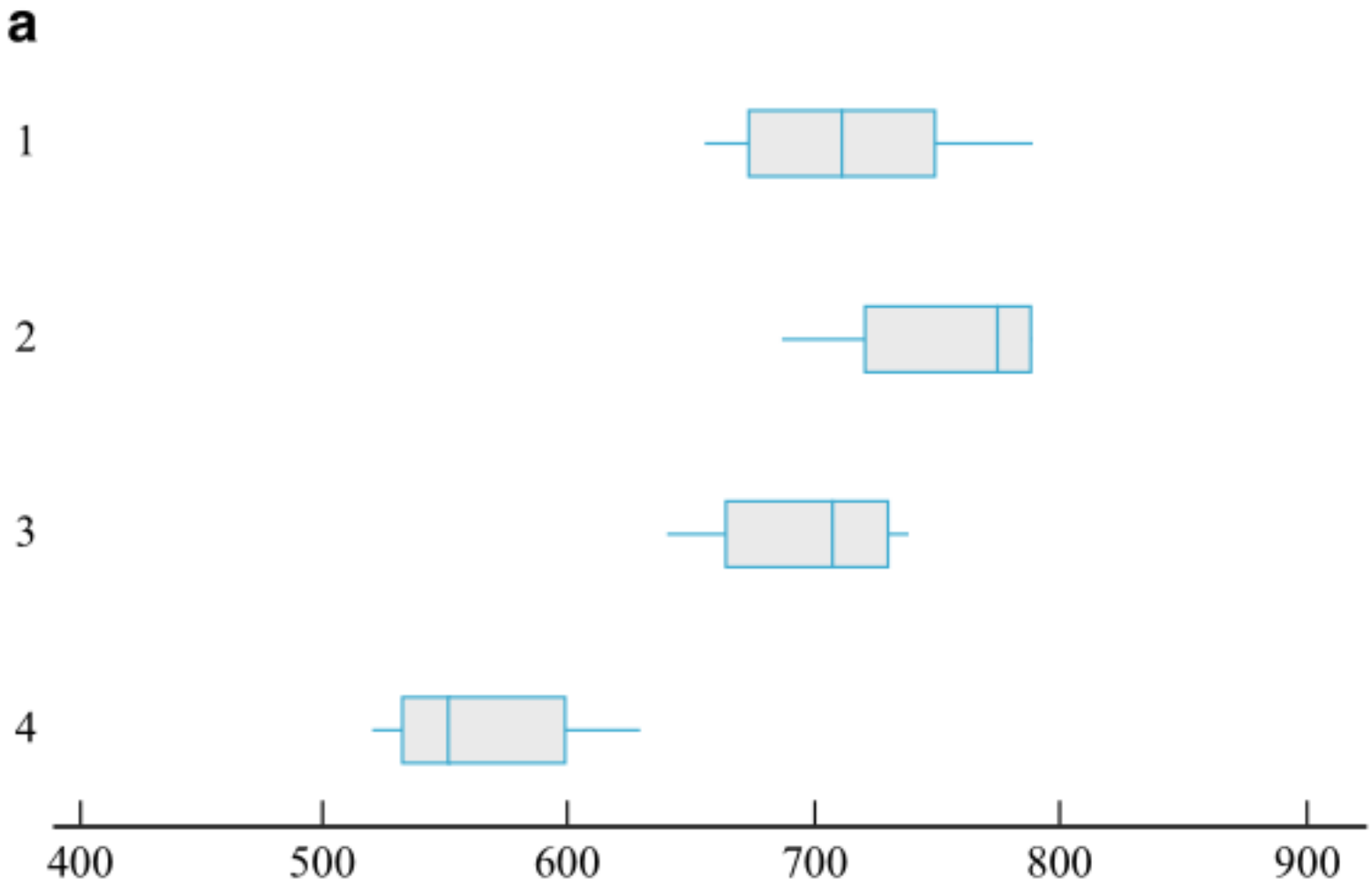
Type of box	Compression strength (lb)			Sample mean	Sample SD
1	655.5	788.3	734.3	713.00	46.55
	721.4	679.1	699.4		
2	789.2	772.5	786.9	756.93	40.34
	686.1	732.1	774.8		
3	737.1	639.0	696.3	698.07	37.20
	671.7	717.2	727.1		
4	535.1	628.7	542.4	562.02	39.87
	559.0	586.9	520.0		
			Grand mean =	682.50	

Her er det altså $I = 4$ grupper med $J = 6$ observasjoner i hver gruppe.

Er forventet trykkstyrke ulikt i de forskjellige gruppene?

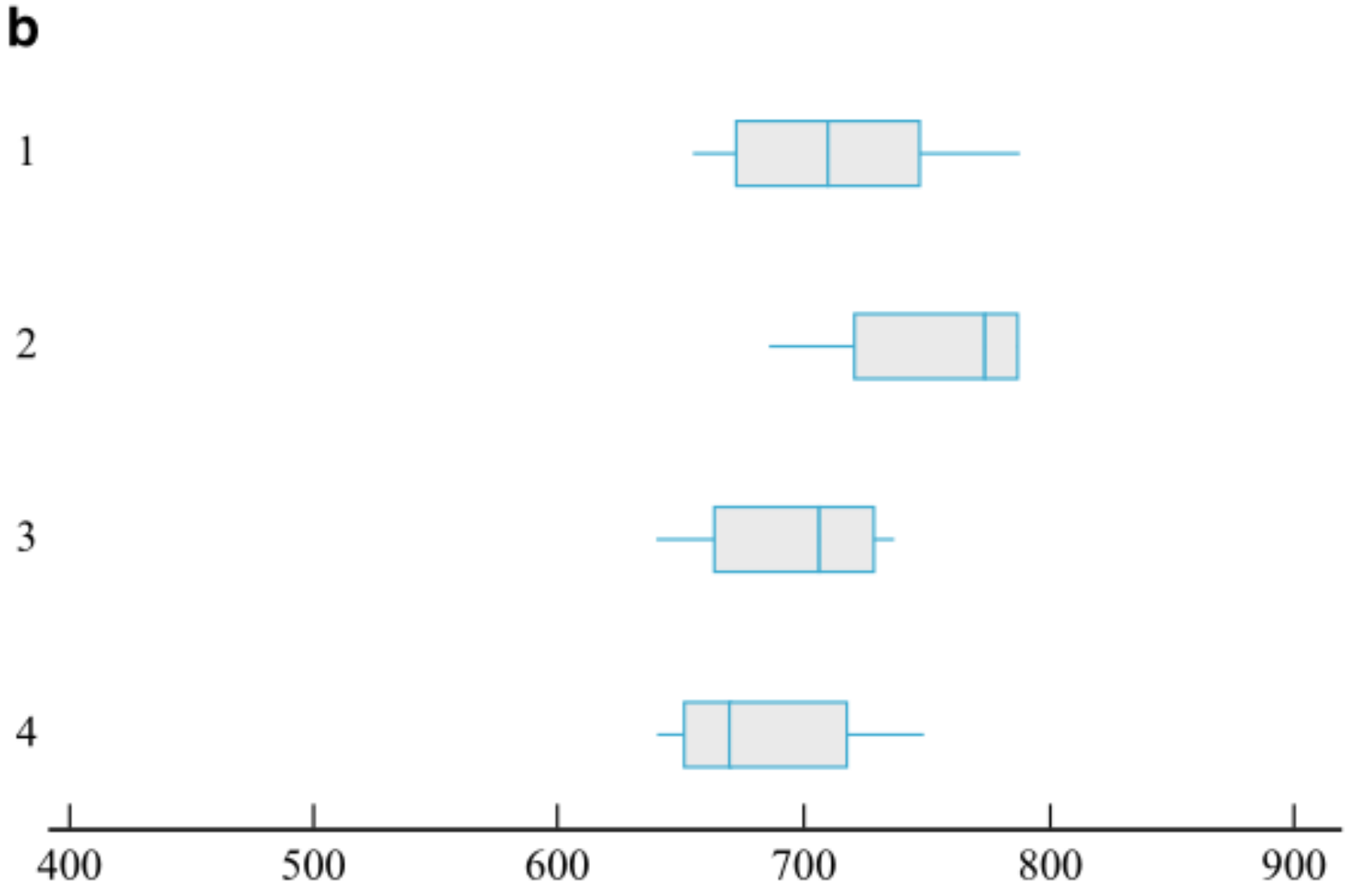
Eksempel 11.1 - forts

Orginale data



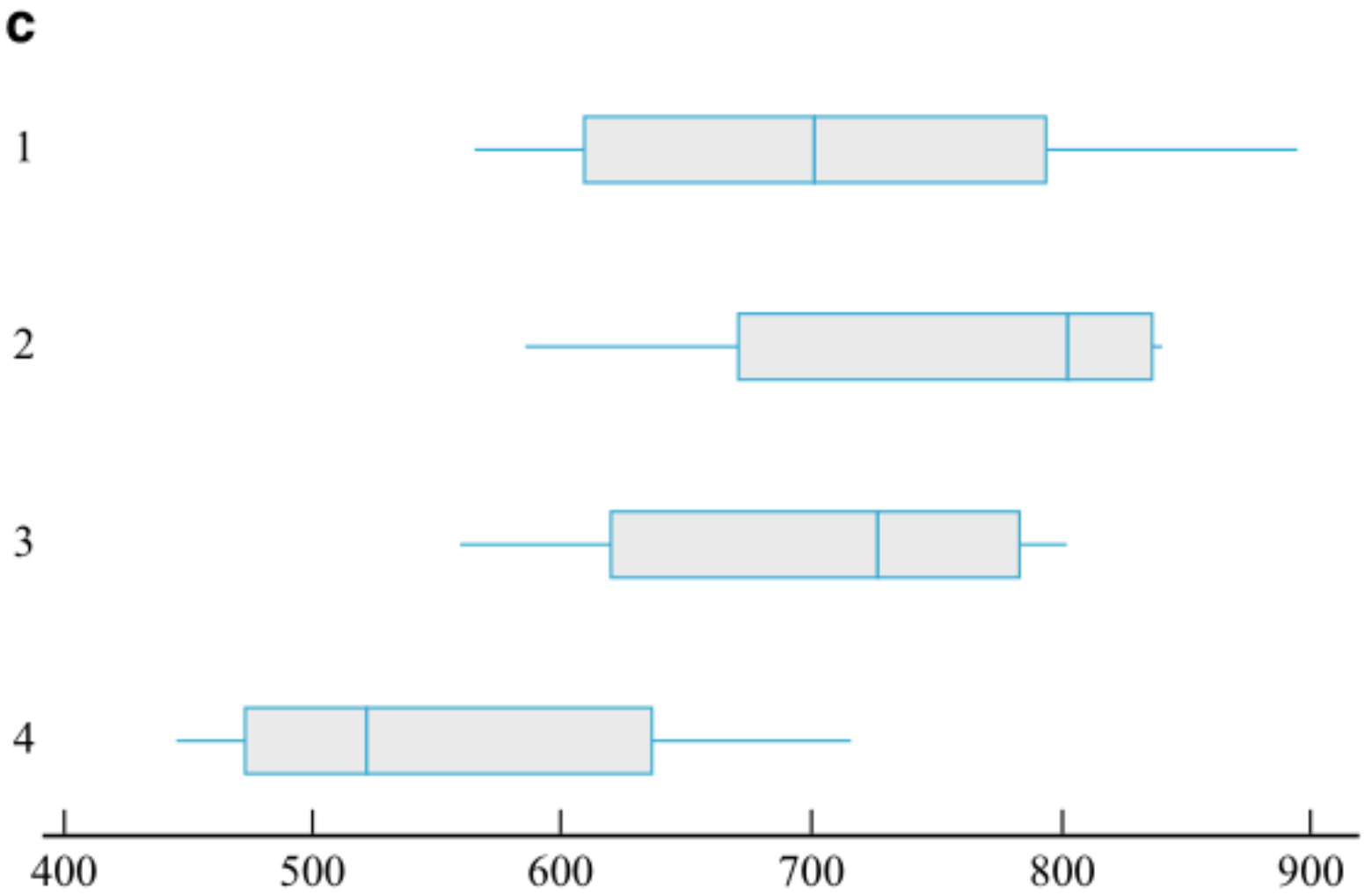
Eksempel 11.1 - forts

Gruppe 4 manipulert: lagt til 120 til hver observasjon i gruppa



Eksempel 11.1 - forts

Variabilitet manipulert i alle grupper, mens gruppegjennomsnittet er beholdt



Utrykket variansanalyse?

- Vi er faktisk interessert i eventuelle forskjeller mellom forventingene i de ulike gruppene

Utrykket variansanalyse?

- Vi er faktisk interessert i eventuelle forskjeller mellom forventingene i de ulike gruppene
- For å oppnå dette vil vi

Utrykket variansanalyse?

- Vi er faktisk interessert i eventuelle forskjeller mellom forventingene i de ulike gruppene
- For å oppnå dette vil vi
- sammenligne variasjonen innad i gruppene

Utrykket variansanalyse?

- Vi er faktisk interessert i eventuelle forskjeller mellom forventingene i de ulike gruppene
- For å oppnå dette vil vi
- sammenligne variasjonen innad i gruppene
- med variasjonen mellom gruppene

Utrykket variansanalyse?

- Vi er faktisk interessert i eventuelle forskjeller mellom forventingene i de ulike gruppene
- For å oppnå dette vil vi
- sammenligne variasjonen innad i gruppene
- med variasjonen mellom gruppene
- Hvis denne er stor nok indikerer dette at nullhypotesen ikke holder

Utrykket variansanalyse?

- Vi er faktisk interessert i eventuelle forskjeller mellom forventingene i de ulike gruppene
- For å oppnå dette vil vi
- sammenligne variasjonen innad i gruppene
- med variasjonen mellom gruppene
- Hvis denne er stor nok indikerer dette at nullhypotesen ikke holder

Må spesifisere næremere hva dette betyr.

Utrykket variansanalyse?

- Vi er faktisk interessert i eventuelle forskjeller mellom forventingene i de ulike gruppene
- For å oppnå dette vil vi
- sammenligne variasjonen innad i gruppene
- med variasjonen mellom gruppene
- Hvis denne er stor nok indikerer dette at nullhypotesen ikke holder

Må spesifisere næremere hva dette betyr.

Muligens ville variasjonsanalyse vært et bedre uttrykk enn variansanalyse.

Variasjon innad i gruppene

- La, for $i = 1, \ldots, I$,

$$\bar{X}_{i.} = \frac{1}{J} \sum_{j=1}^{J} X_{ij}$$

Variasjon innad i gruppene

- La, for $i = 1, \ldots, I$,

$$\bar{X}_{i.} = \frac{1}{J} \sum_{j=1}^{J} X_{ij}$$

$$S_i^2 = \frac{1}{J-1} \sum_{j=1}^{J} (X_{ij} - \bar{X}_{i.})^2, \quad i = 1, \ldots, I.$$

Variasjon innad i gruppene

- La, for $i = 1, \ldots, I$,

$$\bar{X}_{i.} = \frac{1}{J} \sum_{j=1}^{J} X_{ij}$$

$$S_i^2 = \frac{1}{J-1} \sum_{j=1}^{J} (X_{ij} - \bar{X}_{i.})^2, \quad i = 1, \ldots, I.$$

være gjennomsnitt og empirisk standardavvik for gruppe.

Variasjon innad i gruppene

- La, for $i = 1, \ldots, I$,

$$\bar{X}_{i.} = \frac{1}{J} \sum_{j=1}^{J} X_{ij}$$

$$S_i^2 = \frac{1}{J-1} \sum_{j=1}^{J} (X_{ij} - \bar{X}_{i.})^2, \quad i = 1, \ldots, I.$$

  være gjennomsnitt og empirisk standardavvik for gruppe.
- Videre defineres kvadratsum innad for alle gruppene ved

Variasjon innad i gruppene

- La, for $i = 1, \ldots, I$,

$$\bar{X}_{i.} = \frac{1}{J} \sum_{j=1}^{J} X_{ij}$$

$$S_i^2 = \frac{1}{J-1} \sum_{j=1}^{J} (X_{ij} - \bar{X}_{i.})^2, \quad i = 1, \ldots, I.$$

  være gjennomsnitt og empirisk standardavvik for gruppe.
- Videre defineres kvadratsum innad for alle gruppene ved

$$SSE = \sum_{i=1}^{I} \sum_{j=1}^{J} (X_{ij} - \bar{X}_{i.})^2$$

Variasjon innad i gruppene

- La, for $i = 1, \ldots, I$,

$$\bar{X}_{i.} = \frac{1}{J} \sum_{j=1}^{J} X_{ij}$$

$$S_i^2 = \frac{1}{J-1} \sum_{j=1}^{J} (X_{ij} - \bar{X}_{i.})^2, \quad i = 1, \ldots, I.$$

  være gjennomsnitt og empirisk standardavvik for gruppe.
- Videre defineres kvadratsum innad for alle gruppene ved

$$SSE = \sum_{i=1}^{I} \sum_{j=1}^{J} (X_{ij} - \bar{X}_{i.})^2$$

- samt "mean square error" innad i alle gruppene ved

Variasjon innad i gruppene

- La, for $i = 1, \ldots, I$,

$$\bar{X}_{i.} = \frac{1}{J} \sum_{j=1}^{J} X_{ij}$$

$$S_i^2 = \frac{1}{J-1} \sum_{j=1}^{J} (X_{ij} - \bar{X}_{i.})^2, \quad i = 1, \ldots, I.$$

  være gjennomsnitt og empirisk standardavvik for gruppe.
- Videre defineres kvadratsum innad for alle gruppene ved

$$SSE = \sum_{i=1}^{I} \sum_{j=1}^{J} (X_{ij} - \bar{X}_{i.})^2$$

- samt "mean square error" innad i alle gruppene ved

$$MSE = \frac{1}{I} \sum_{i=1}^{I} S_i^2 = \frac{1}{I(J-1)} \sum_{i=1}^{I} \sum_{j=1}^{J} (X_{ij} - \bar{X}_{i.})^2 = \frac{SSE}{I(J-1)},$$

Variasjon innad i gruppene

- La, for $i = 1, \ldots, I$,

$$\bar{X}_{i.} = \frac{1}{J}\sum_{j=1}^{J} X_{ij}$$

$$S_i^2 = \frac{1}{J-1}\sum_{j=1}^{J}(X_{ij} - \bar{X}_{i.})^2, \quad i = 1, \ldots, I.$$

  være gjennomsnitt og empirisk standardavvik for gruppe.
- Videre defineres kvadratsum innad for alle gruppene ved

$$SSE = \sum_{i=1}^{I}\sum_{j=1}^{J}(X_{ij} - \bar{X}_{i.})^2$$

- samt "mean square error" innad i alle gruppene ved

$$MSE = \frac{1}{I}\sum_{i=1}^{I} S_i^2 = \frac{1}{I(J-1)}\sum_{i=1}^{I}\sum_{j=1}^{J}(X_{ij} - \bar{X}_{i.})^2 = \frac{SSE}{I(J-1)},$$

- Vi har da
$\mathrm{E}(MSE) = \frac{1}{I}\sum_{i=1}^{I}\underbrace{\mathrm{E}(S_i^2)}_{\sigma^2} = \sigma^2$, dvs. forventningsrett for σ^2

Variasjon mellom gruppene

- La $\bar{X}_{..} = \frac{1}{IJ} \sum_{i=1}^{I} \sum_{j=1}^{J} X_{ij}$ være det totale gjennomsnittet ("grand mean")

Variasjon mellom gruppene

- La $\bar{X}_{..} = \frac{1}{IJ} \sum_{i=1}^{I} \sum_{j=1}^{J} X_{ij}$ være det totale gjennomsnittet ("grand mean")
- Definer kvadratsummen mellom gruppene ved

Variasjon mellom gruppene

- La $\bar{X}_{..} = \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J} X_{ij}$ være det totale gjennomsnittet ("grand mean")
- Definer kvadratsummen mellom gruppene ved

$$SSTr = J\sum_{i=1}^{I}(\bar{X}_{i.} - \bar{X}_{..})^2$$

Variasjon mellom gruppene

- La $\bar{X}_{..} = \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J} X_{ij}$ være det totale gjennomsnittet ("grand mean")
- Definer kvadratsummen mellom gruppene ved

$$SSTr = J\sum_{i=1}^{I}(\bar{X}_{i\cdot} - \bar{X}_{..})^2$$

- Størrelsen "mean square error" defineres deretter ved

$$MSTr = \frac{1}{I-1}SSTr$$

Variasjon mellom gruppene

- La $\bar{X}_{..} = \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J} X_{ij}$ være det totale gjennomsnittet ("grand mean")
- Definer kvadratsummen mellom gruppene ved

$$SSTr = J\sum_{i=1}^{I}(\bar{X}_{i\cdot} - \bar{X}_{..})^2$$

- Størrelsen "mean square error" defineres deretter ved

$$MSTr = \frac{1}{I-1}SSTr$$

- Vi skal vise at også $\mathrm{E}(MSTr) = \sigma^2$

Variasjon mellom gruppene

- La $\bar{X}_{..} = \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J} X_{ij}$ være det totale gjennomsnittet ("grand mean")
- Definer kvadratsummen mellom gruppene ved

$$SSTr = J\sum_{i=1}^{I}(\bar{X}_{i\cdot} - \bar{X}_{..})^2$$

- Størrelsen "mean square error" defineres deretter ved

$$MSTr = \frac{1}{I-1}SSTr$$

- Vi skal vise at også $\mathrm{E}(MSTr) = \sigma^2$
- men bare under $\mathrm{H}_0 : \mu_1 = \mu_2 = \cdots = \mu_I$
- Under H_a vil $\mathrm{E}(MSTr) > \sigma^2$

Variasjon mellom gruppene

- La $\bar{X}_{..} = \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J} X_{ij}$ være det totale gjennomsnittet ("grand mean")
- Definer kvadratsummen mellom gruppene ved

$$SSTr = J\sum_{i=1}^{I}(\bar{X}_{i\cdot} - \bar{X}_{..})^2$$

- Størrelsen "mean square error" defineres deretter ved

$$MSTr = \frac{1}{I-1}SSTr$$

- Vi skal vise at også $\mathrm{E}(MSTr) = \sigma^2$
- men bare under $\mathrm{H}_0 : \mu_1 = \mu_2 = \cdots = \mu_I$
- Under H_a vil $\mathrm{E}(MSTr) > \sigma^2$

- Merk at MSE er forventningsrett for σ^2 både under H_0 og H_a

MSTr under H_0

Anta nå at $H_0 : \mu_1 = \mu_2 = \cdots = \mu_I = \mu$ holder.

MSTr under H_0

Anta nå at $H_0 : \mu_1 = \mu_2 = \cdots = \mu_I = \mu$ holder.

Da gjelder

- $\bar{X}_{i\cdot} \sim N\left(\mu, \frac{\sigma^2}{J}\right)$, $i = 1, \ldots, I$.

MSTr under H_0

Anta nå at $H_0 : \mu_1 = \mu_2 = \cdots = \mu_I = \mu$ holder.

Da gjelder

- $\bar{X}_{i\cdot} \sim N\left(\mu, \frac{\sigma^2}{J}\right)$, $i = 1, \ldots, I$.
- Dessuten er $\bar{X}_{i\cdot}$-ene er uavhengige

MSTr under H_0

Anta nå at $H_0 : \mu_1 = \mu_2 = \cdots = \mu_I = \mu$ holder.

Da gjelder

- $\bar{X}_{i.} \sim N\left(\mu, \frac{\sigma^2}{J}\right)$, $i = 1, \ldots, I$.
- Dessuten er $\bar{X}_{i.}$-ene er uavhengige
- Dermed blir $\frac{1}{I-1}\sum_{i=1}^{I}(\bar{X}_{i.} - \bar{X}_{..})^2$ empirisk varians for $\bar{X}_{i.}$-ene.

MSTr under H_0

Anta nå at $H_0 : \mu_1 = \mu_2 = \cdots = \mu_I = \mu$ holder.

Da gjelder

- $\bar{X}_{i.} \sim N\left(\mu, \frac{\sigma^2}{J}\right)$, $i = 1, \ldots, I$.
- Dessuten er $\bar{X}_{i.}$-ene er uavhengige
- Dermed blir $\frac{1}{I-1}\sum_{i=1}^{I}(\bar{X}_{i.} - \bar{X}_{..})^2$ empirisk varians for $\bar{X}_{i.}$-ene.
- og altså forventningsrett for σ^2/J

MSTr under H_0

Anta nå at $H_0 : \mu_1 = \mu_2 = \cdots = \mu_I = \mu$ holder.

Da gjelder

- $\bar{X}_{i.} \sim N\left(\mu, \frac{\sigma^2}{J}\right)$, $i = 1, \ldots, I$.
- Dessuten er $\bar{X}_{i.}$-ene er uavhengige
- Dermed blir $\frac{1}{I-1}\sum_{i=1}^{I}(\bar{X}_{i.} - \bar{X}_{..})^2$ empirisk varians for $\bar{X}_{i.}$-ene.
- og altså forventningsrett for σ^2/J

MSTr under H_0

Anta nå at $H_0 : \mu_1 = \mu_2 = \cdots = \mu_I = \mu$ holder.

Da gjelder

- $\bar{X}_{i\cdot} \sim N\left(\mu, \frac{\sigma^2}{J}\right)$, $i = 1, \ldots, I$.
- Dessuten er $\bar{X}_{i\cdot}$-ene er uavhengige
- Dermed blir $\frac{1}{I-1}\sum_{i=1}^{I}(\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2$ empirisk varians for $\bar{X}_{i\cdot}$-ene.
- og altså forventningsrett for σ^2/J
- Dermed er $MSTr = \frac{1}{I-1}SSTr = \frac{J}{I-1}\sum_{i=1}^{I}(\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2$ forventningsrett for σ^2.

MSTr under H_0

Anta nå at $H_0 : \mu_1 = \mu_2 = \cdots = \mu_I = \mu$ holder.

Da gjelder

- $\bar{X}_{i\cdot} \sim N\left(\mu, \frac{\sigma^2}{J}\right)$, $i = 1, \ldots, I$.
- Dessuten er $\bar{X}_{i\cdot}$-ene er uavhengige
- Dermed blir $\frac{1}{I-1}\sum_{i=1}^{I}(\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2$ empirisk varians for $\bar{X}_{i\cdot}$-ene.
- og altså forventningsrett for σ^2/J
- Dermed er $MSTr = \frac{1}{I-1}SSTr = \frac{J}{I-1}\sum_{i=1}^{I}(\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2$ forventningsrett for σ^2.

altså alt under H_0.

F-testen for $H_0 : \mu_1 = \mu_2 = \cdots = \mu_I = \mu$

Vi har dermed etablert at $MSE = \frac{1}{I(J-1)} SSE$ og $MSTr = \frac{1}{I-1} SSTr$ er forventningsrette for σ^2 under nullhypotesen.

F-testen for $H_0 : \mu_1 = \mu_2 = \cdots = \mu_I = \mu$

Vi har dermed etablert at $MSE = \frac{1}{I(J-1)} SSE$ og $MSTr = \frac{1}{I-1} SSTr$ er forventningsrette for σ^2 under nullhypotesen.

Vi skal teste nullhypotesen ved testobservator

F-testen for $H_0 : \mu_1 = \mu_2 = \cdots = \mu_I = \mu$

Vi har dermed etablert at $MSE = \frac{1}{I(J-1)} SSE$ og $MSTr = \frac{1}{I-1} SSTr$ er forventningsrette for σ^2 under nullhypotesen.

Vi skal teste nullhypotesen ved testobservator

$$F = \frac{MSTr}{MSE}$$

F-testen for $H_0 : \mu_1 = \mu_2 = \cdots = \mu_I = \mu$

Vi har dermed etablert at $MSE = \frac{1}{I(J-1)} SSE$ og $MSTr = \frac{1}{I-1} SSTr$ er forventningsrette for σ^2 under nullhypotesen.

Vi skal teste nullhypotesen ved testobservator

$$F = \frac{MSTr}{MSE}$$

Det vil faktisk vise seg at

$$F \sim F_{I-1, I(J-1)},$$

F-testen for $H_0 : \mu_1 = \mu_2 = \cdots = \mu_I = \mu$

Vi har dermed etablert at $MSE = \frac{1}{I(J-1)} SSE$ og $MSTr = \frac{1}{I-1} SSTr$ er forventningsrette for σ^2 under nullhypotesen.

Vi skal teste nullhypotesen ved testobservator

$$F = \frac{MSTr}{MSE}$$

Det vil faktisk vise seg at

$$F \sim F_{I-1, I(J-1)},$$

altså F-fordelt med $I - 1$ og $I(J - 1)$ frihetsgrader.

F-testen for $H_0 : \mu_1 = \mu_2 = \cdots = \mu_I = \mu$

Vi har dermed etablert at $MSE = \frac{1}{I(J-1)} SSE$ og $MSTr = \frac{1}{I-1} SSTr$ er forventningsrette for σ^2 under nullhypotesen.

Vi skal teste nullhypotesen ved testobservator

$$F = \frac{MSTr}{MSE}$$

Det vil faktisk vise seg at

$$F \sim F_{I-1, I(J-1)},$$

altså F-fordelt med $I - 1$ og $I(J - 1)$ frihetsgrader.

Store verdier av F indikere at nullhypotesen ikke holder.

F-testen for $H_0 : \mu_1 = \mu_2 = \cdots = \mu_I = \mu$

Vi har dermed etablert at $MSE = \frac{1}{I(J-1)} SSE$ og $MSTr = \frac{1}{I-1} SSTr$ er forventningsrette for σ^2 under nullhypotesen.

Vi skal teste nullhypotesen ved testobservator

$$F = \frac{MSTr}{MSE}$$

Det vil faktisk vise seg at

$$F \sim F_{I-1, I(J-1)},$$

altså F-fordelt med $I - 1$ og $I(J - 1)$ frihetsgrader.

Store verdier av F indikere at nullhypotesen ikke holder.
Dermed forkastes H_0 med nivå α dersom $F > F_{\alpha, I-1, I(J-1)}$

F-testen for $H_0 : \mu_1 = \mu_2 = \cdots = \mu_I = \mu$

Vi har dermed etablert at $MSE = \frac{1}{I(J-1)} SSE$ og $MSTr = \frac{1}{I-1} SSTr$ er forventningsrette for σ^2 under nullhypotesen.

Vi skal teste nullhypotesen ved testobservator

$$F = \frac{MSTr}{MSE}$$

Det vil faktisk vise seg at

$$F \sim F_{I-1, I(J-1)},$$

altså F-fordelt med $I - 1$ og $I(J - 1)$ frihetsgrader.

Store verdier av F indikere at nullhypotesen ikke holder. Dermed forkastes H_0 med nivå α dersom $F > F_{\alpha, I-1, I(J-1)}$ $= (1 - \alpha)$ 100 persentilen i $F_{I-1, I(J-1)}$-fordelingen.

Fra tidligere

- Anta $Z_1, ..., Z_n \overset{uif}{\sim} N(\mu, \sigma^2)$

Fra tidligere

- Anta $Z_1, ..., Z_n \overset{uif}{\sim} N(\mu, \sigma^2)$
- Da er

$$\frac{\sum_{j=1}^{n}(Z_j - \bar{Z})^2}{\sigma^2} \sim \chi^2_{n-1}$$

Fra tidligere

- Anta $Z_1, ..., Z_n \overset{uif}{\sim} N(\mu, \sigma^2)$
- Da er

$$\frac{\sum_{j=1}^{n}(Z_j - \bar{Z})^2}{\sigma^2} \sim \chi^2_{n-1}$$

- Nå: $X_{ij} \overset{uif}{\sim} N(\mu_i, \sigma^2)$
- Gir

$$\frac{\sum_{j=1}^{J}(X_{ij} - \bar{X}_{i\cdot})^2}{\sigma^2} \sim \chi^2_{J-1}$$

Fra tidligere

- Anta $Z_1, ..., Z_n \overset{uif}{\sim} N(\mu, \sigma^2)$
- Da er

$$\frac{\sum_{j=1}^{n}(Z_j - \bar{Z})^2}{\sigma^2} \sim \chi^2_{n-1}$$

- Nå: $X_{ij} \overset{uif}{\sim} N(\mu_i, \sigma^2)$
- Gir

$$\frac{\sum_{j=1}^{J}(X_{ij} - \bar{X}_{i\cdot})^2}{\sigma^2} \sim \chi^2_{J-1}$$

- Får da også at

$$\sum_{i=1}^{I} \frac{\sum_{j=1}^{J}(X_{ij} - \bar{X}_{i\cdot})^2}{\sigma^2} \sim \sum_{i=1}^{I} \underbrace{\chi^2_{J-1}}_{uavh} = \chi^2_{I(J-1)}$$

Fra tidligere

- Anta $Z_1, ..., Z_n \overset{uif}{\sim} N(\mu, \sigma^2)$
- Da er

$$\frac{\sum_{j=1}^{n}(Z_j - \bar{Z})^2}{\sigma^2} \sim \chi^2_{n-1}$$

- Nå: $X_{ij} \overset{uif}{\sim} N(\mu_i, \sigma^2)$
- Gir

$$\frac{\sum_{j=1}^{J}(X_{ij} - \bar{X}_{i\cdot})^2}{\sigma^2} \sim \chi^2_{J-1}$$

- Får da også at

$$\sum_{i=1}^{I} \frac{\sum_{j=1}^{J}(X_{ij} - \bar{X}_{i\cdot})^2}{\sigma^2} \sim \sum_{i=1}^{I} \underbrace{\chi^2_{J-1}}_{uavh} = \chi^2_{I(J-1)}$$

- Under H_0: $\bar{X}_{i\cdot} \overset{uif}{\sim} N(0, \sigma^2/J)$
- Gir

$$\frac{J\sum_{j=1}^{J} X_{i\cdot} - \bar{X}_{\cdot\cdot})^2}{\sigma^2} \sim \chi^2_{I-1}$$

Test og variansanalysetabell

- Vi har da at $\frac{SSTr}{\sigma^2} = \frac{J\sum_{i=1}^{I}(\bar{X}_{i\cdot}-\bar{X}_{\cdot\cdot})^2}{\sigma^2} \sim \chi^2_{I-1}$ under H_0

Test og variansanalysetabell

- Vi har da at $\frac{SSTr}{\sigma^2} = \frac{J\sum_{i=1}^{I}(\bar{X}_{i\cdot}-\bar{X}_{\cdot\cdot})^2}{\sigma^2} \sim \chi^2_{I-1}$ under H_0
- og $\frac{SSE}{\sigma^2} = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}(X_{ij}-\bar{X}_{i\cdot})^2}{\sigma^2} \sim \chi^2_{I(J-1)}$

Test og variansanalysetabell

- Vi har da at $\frac{SSTr}{\sigma^2} = \frac{J\sum_{i=1}^{I}(\bar{X}_{i\cdot}-\bar{X}_{\cdot\cdot})^2}{\sigma^2} \sim \chi^2_{I-1}$ under H_0
- og $\frac{SSE}{\sigma^2} = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}(X_{ij}-\bar{X}_{i\cdot})^2}{\sigma^2} \sim \chi^2_{I(J-1)}$
- og at disse størrelsene er uavhengige.

Test og variansanalysetabell

- Vi har da at $\frac{SSTr}{\sigma^2} = \frac{J\sum_{i=1}^{I}(\bar{X}_{i\cdot}-\bar{X}_{\cdot\cdot})^2}{\sigma^2} \sim \chi^2_{I-1}$ under H_0
- og $\frac{SSE}{\sigma^2} = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}(X_{ij}-\bar{X}_{i\cdot})^2}{\sigma^2} \sim \chi^2_{I(J-1)}$
- og at disse størrelsene er uavhengige.

Test og variansanalysetabell

- Vi har da at $\frac{SSTr}{\sigma^2} = \frac{J\sum_{i=1}^{I}(\bar{X}_{i\cdot}-\bar{X}_{\cdot\cdot})^2}{\sigma^2} \sim \chi^2_{I-1}$ under H_0
- og $\frac{SSE}{\sigma^2} = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}(X_{ij}-\bar{X}_{i\cdot})^2}{\sigma^2} \sim \chi^2_{I(J-1)}$
- og at disse størrelsene er uavhengige.
- Da er $F = \frac{SSTr/(I-1)}{SSE/(I(J-1))} = \frac{MSTr}{MSE} \sim F_{I-1,I(J-1)}$.

Test og variansanalysetabell

- Vi har da at $\frac{SSTr}{\sigma^2} = \frac{J\sum_{i=1}^{I}(\bar{X}_{i\cdot}-\bar{X}_{\cdot\cdot})^2}{\sigma^2} \sim \chi^2_{I-1}$ under H_0
- og $\frac{SSE}{\sigma^2} = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}(X_{ij}-\bar{X}_{i\cdot})^2}{\sigma^2} \sim \chi^2_{I(J-1)}$
- og at disse størrelsene er uavhengige.
- Da er $F = \frac{SSTr/(I-1)}{SSE/(I(J-1))} = \frac{MSTr}{MSE} \sim F_{I-1,I(J-1)}$.

Test og variansanalysetabell

- Vi har da at $\frac{SSTr}{\sigma^2} = \frac{J\sum_{i=1}^{I}(\bar{X}_{i\cdot}-\bar{X}_{\cdot\cdot})^2}{\sigma^2} \sim \chi^2_{I-1}$ under H_0
- og $\frac{SSE}{\sigma^2} = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}(X_{ij}-\bar{X}_{i\cdot})^2}{\sigma^2} \sim \chi^2_{I(J-1)}$
- og at disse størrelsene er uavhengige.
- Da er $F = \frac{SSTr/(I-1)}{SSE/(I(J-1))} = \frac{MSTr}{MSE} \sim F_{I-1,I(J-1)}$.
- En test for
  $H_0 : \mu_1 = \ldots = \mu_I$ mot H_a : minst to μ_i-er er forskjellige
  med signifikansnivå α er dermed gitt ved å forkaste H_0
  dersom $F > F_{\alpha,I-1,I(J-1)}$.

Test og variansanalysetabell

- Vi har da at $\frac{SSTr}{\sigma^2} = \frac{J\sum_{i=1}^{I}(\bar{X}_{i\cdot}-\bar{X}_{\cdot\cdot})^2}{\sigma^2} \sim \chi^2_{I-1}$ under H_0
- og $\frac{SSE}{\sigma^2} = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}(X_{ij}-\bar{X}_{i\cdot})^2}{\sigma^2} \sim \chi^2_{I(J-1)}$
- og at disse størrelsene er uavhengige.
- Da er $F = \frac{SSTr/(I-1)}{SSE/(I(J-1))} = \frac{MSTr}{MSE} \sim F_{I-1,I(J-1)}$.
- En test for
  $H_0 : \mu_1 = \ldots = \mu_I$ mot H_a : minst to μ_i-er er forskjellige
  med signifikansnivå α er dermed gitt ved å forkaste H_0
  dersom $F > F_{\alpha,I-1,I(J-1)}$.

- Vi sammenfatter ofte resultatene i en såkalt

  Variansanalysetabell (ANOVA tabell):

Kilde til var.	**Df**	**SS**	**MS**	**f**
Behandlinger	$I-1$	$SSTr$	$MSTr = SSTr/(I-1)$	$MSTr/MSE$
Feil	$I(J-1)$	SSE	$MSE = SSE/(I(J-1))$	
Total	$IJ-1$	SST		

Forventning MSTr under H_a (og under H_0)

- Definerte

$$MSTr = \frac{SSTr}{I-1} = \frac{J}{I-1}\sum_{i=1}^{I}(\bar{X}_{i.} - \bar{X}_{..})^2,$$

Forventning MSTr under H_a (og under H_0)

- Definerte

$$MSTr = \frac{SSTr}{I-1} = \frac{J}{I-1}\sum_{i=1}^{I}(\bar{X}_{i.} - \bar{X}_{..})^2,$$

- Generelt kan det vises at

$$\mathrm{E}(MSTr) = \mathrm{E}\left(\frac{SSTr}{I-1}\right) = \sigma^2 + \frac{J}{I-1}\sum_{i=1}^{I}(\mu_i - \bar{\mu})^2\,,$$

,

Forventning MSTr under H_a (og under H_0)

- Definerte

$$MSTr = \frac{SSTr}{I-1} = \frac{J}{I-1}\sum_{i=1}^{I}(\bar{X}_{i.} - \bar{X}_{..})^2,$$

- Generelt kan det vises at

$$\mathrm{E}(MSTr) = \mathrm{E}\left(\frac{SSTr}{I-1}\right) = \sigma^2 + \frac{J}{I-1}\sum_{i=1}^{I}(\mu_i - \bar{\mu})^2,$$

, der $\bar{\mu} = \frac{1}{I}\sum_{i=1}^{I}\mu_i$.

- Under H_0: $\mu_i = \bar{\mu}$ forsvinner siste ledd og vi får altså

$$\mathrm{E}(MSTr) = \sigma^2$$

Forventning MSTr under H_a (og under H_0)

- Definerte

$$MSTr = \frac{SSTr}{I-1} = \frac{J}{I-1}\sum_{i=1}^{I}(\bar{X}_{i.} - \bar{X}_{..})^2,$$

- Generelt kan det vises at

$$\mathrm{E}(MSTr) = \mathrm{E}\left(\frac{SSTr}{I-1}\right) = \sigma^2 + \frac{J}{I-1}\sum_{i=1}^{I}(\mu_i - \bar{\mu})^2,$$

, der $\bar{\mu} = \frac{1}{I}\sum_{i=1}^{I}\mu_i$.

- Under H_0: $\mu_i = \bar{\mu}$ forsvinner siste ledd og vi får altså

$$\mathrm{E}(MSTr) = \sigma^2$$

- Men hvis H_0 er gal: $\mathrm{E}(MSTr) > \sigma^2$

Dekomponering av variasjon

- Den totale kvadratsummen er gitt ved

$$SST = \sum_{i=1}^{I} \sum_{j=1}^{J} (X_{ij} - X_{..})^2$$

Dekomponering av variasjon

- Den totale kvadratsummen er gitt ved

$$SST = \sum_{i=1}^{I} \sum_{j=1}^{J} (X_{ij} - X_{..})^2$$

- Vi kan da dekomponere $SST = SSTr + SSE$

Dekomponering av variasjon

- Den totale kvadratsummen er gitt ved

$$SST = \sum_{i=1}^{I} \sum_{j=1}^{J} (X_{ij} - X_{..})^2$$

- Vi kan da dekomponere $SST = SSTr + SSE$
- Tilsvarende har vi for antall frihetsgrader er $df_{Tot} = I - 1 + I(J-1) = IJ - 1$.

Dekomponering av variasjon

- Den totale kvadratsummen er gitt ved

$$SST = \sum_{i=1}^{I} \sum_{j=1}^{J} (X_{ij} - X_{..})^2$$

- Vi kan da dekomponere $SST = SSTr + SSE$
- Tilsvarende har vi for antall frihetsgrader er $df_{Tot} = I - 1 + I(J-1) = IJ - 1$.
- Den totale variasjonen kan altså dekomponeres variasjonen som skyldes forskjellige behandlinger ($SSTr$) og variasjonen som skyldes tilfeldige feil (SSE).

Dekomponering av variasjon

- Den totale kvadratsummen er gitt ved

$$SST = \sum_{i=1}^{I} \sum_{j=1}^{J} (X_{ij} - X_{..})^2$$

- Vi kan da dekomponere $SST = SSTr + SSE$
- Tilsvarende har vi for antall frihetsgrader er $df_{Tot} = I - 1 + I(J-1) = IJ - 1$.
- Den totale variasjonen kan altså dekomponeres variasjonen som skyldes forskjellige behandlinger ($SSTr$) og variasjonen som skyldes tilfeldige feil (SSE).
- Hvis $SSTr$ er liten sammenlignet med SSE, er det lite som motsier H_0.

Dekomponering av variasjon

- Den totale kvadratsummen er gitt ved

$$SST = \sum_{i=1}^{I} \sum_{j=1}^{J} (X_{ij} - X_{..})^2$$

- Vi kan da dekomponere $SST = SSTr + SSE$
- Tilsvarende har vi for antall frihetsgrader er $df_{Tot} = I - 1 + I(J-1) = IJ - 1$.
- Den totale variasjonen kan altså dekomponeres variasjonen som skyldes forskjellige behandlinger ($SSTr$) og variasjonen som skyldes tilfeldige feil (SSE).
- Hvis $SSTr$ er liten sammenlignet med SSE, er det lite som motsier H_0.
- Derimot vil en stor $SSTr$ sammenlignet med SSE være en indikasjon på at H_0 ikke stemmer.

Dekomponering av variasjon

- Den totale kvadratsummen er gitt ved

$$SST = \sum_{i=1}^{I}\sum_{j=1}^{J}(X_{ij} - X_{..})^2$$

- Vi kan da dekomponere $SST = SSTr + SSE$
- Tilsvarende har vi for antall frihetsgrader er $df_{Tot} = I - 1 + I(J-1) = IJ - 1$.
- Den totale variasjonen kan altså dekomponeres variasjonen som skyldes forskjellige behandlinger ($SSTr$) og variasjonen som skyldes tilfeldige feil (SSE).
- Hvis $SSTr$ er liten sammenlignet med SSE, er det lite som motsier H_0.
- Derimot vil en stor $SSTr$ sammenlignet med SSE være en indikasjon på at H_0 ikke stemmer.

Eksempel

Fettopptak i rotter.

Sjekk av antakelser - normalfordeling

- For å kunne teste antakelsene vi har gjort, trenger vi å definere feilleddene ved $\varepsilon_{ij} = X_{ij} - \mu_i$.

Sjekk av antakelser - normalfordeling

- For å kunne teste antakelsene vi har gjort, trenger vi å definere feilleddene ved $\varepsilon_{ij} = X_{ij} - \mu_i$.
- Vi har antatt at $\varepsilon_{ij} \sim N(0, \sigma^2)$ og at ε_{ij}-ene er uavhengige.

Sjekk av antakelser - normalfordeling

- For å kunne teste antakelsene vi har gjort, trenger vi å definere feilleddene ved $\varepsilon_{ij} = X_{ij} - \mu_i$.
- Vi har antatt at $\varepsilon_{ij} \sim N(0, \sigma^2)$ og at ε_{ij}-ene er uavhengige.
- Disse kan approksimeres med residualene $e_{ij} = X_{ij} - \bar{X}_{i.}$.

Sjekk av antakelser - normalfordeling

- For å kunne teste antakelsene vi har gjort, trenger vi å definere feilleddene ved $\varepsilon_{ij} = X_{ij} - \mu_i$.
- Vi har antatt at $\varepsilon_{ij} \sim N(0, \sigma^2)$ og at ε_{ij}-ene er uavhengige.
- Disse kan approksimeres med residualene $e_{ij} = X_{ij} - \bar{X}_{i.}$.
- Hvis antakelsene våre stemmer, vil $e_{ij} \overset{tiln.}{\sim} N(0, \sigma^2)$.

Sjekk av antakelser - normalfordeling

- For å kunne teste antakelsene vi har gjort, trenger vi å definere feilleddene ved $\varepsilon_{ij} = X_{ij} - \mu_i$.
- Vi har antatt at $\varepsilon_{ij} \sim N(0, \sigma^2)$ og at ε_{ij}-ene er uavhengige.
- Disse kan approksimeres med residualene $e_{ij} = X_{ij} - \bar{X}_{i.}$.
- Hvis antakelsene våre stemmer, vil $e_{ij} \overset{tiln.}{\sim} N(0, \sigma^2)$.
- Antakelsen om normalfordeling kan vi sjekke med et normalt kvantilplott av e_{ij}-ene.

Sjekk av antakelser - normalfordeling

- For å kunne teste antakelsene vi har gjort, trenger vi å definere feilleddene ved $\varepsilon_{ij} = X_{ij} - \mu_i$.
- Vi har antatt at $\varepsilon_{ij} \sim N(0, \sigma^2)$ og at ε_{ij}-ene er uavhengige.
- Disse kan approksimeres med residualene $e_{ij} = X_{ij} - \bar{X}_{i.}$.
- Hvis antakelsene våre stemmer, vil $e_{ij} \overset{tiln.}{\sim} N(0, \sigma^2)$.
- Antakelsen om normalfordeling kan vi sjekke med et normalt kvantilplott av e_{ij}-ene.

Sjekk av antakelser - normalfordeling

- For å kunne teste antakelsene vi har gjort, trenger vi å definere feilleddene ved $\varepsilon_{ij} = X_{ij} - \mu_i$.
- Vi har antatt at $\varepsilon_{ij} \sim N(0, \sigma^2)$ og at ε_{ij}-ene er uavhengige.
- Disse kan approksimeres med residualene $e_{ij} = X_{ij} - \bar{X}_{i.}$.
- Hvis antakelsene våre stemmer, vil $e_{ij} \overset{tiln.}{\sim} N(0, \sigma^2)$.
- Antakelsen om normalfordeling kan vi sjekke med et normalt kvantilplott av e_{ij}-ene.

Eksempel

Fettopptak i rotter.

Sjekk av antagelser - like varianser

- For å sjekke antakelsen om lik varians kan vi enten bruke log-likelihood ratio-testen (kalles Bartletts test i denne sammenhengen) eller Levenes test, som er en ANOVA F-test basert på $|x_{ij} - \bar{x}_{i.}|$ i stedet for x_{ij}.

Sjekk av antagelser - like varianser

- For å sjekke antakelsen om lik varians kan vi enten bruke log-likelihood ratio-testen (kalles Bartletts test i denne sammenhengen) eller Levenes test, som er en ANOVA F-test basert på $|x_{ij} - \bar{x}_{i.}|$ i stedet for x_{ij}.

Sjekk av antagelser - like varianser

- For å sjekke antakelsen om lik varians kan vi enten bruke log-likelihood ratio-testen (kalles Bartletts test i denne sammenhengen) eller Levenes test, som er en ANOVA F-test basert på $|x_{ij} - \bar{x}_{i.}|$ i stedet for x_{ij}.
- Sistnevnte test er mer robust overfor antakelsen om normalfordeling.

Sjekk av antagelser - like varianser

- For å sjekke antakelsen om lik varians kan vi enten bruke log-likelihood ratio-testen (kalles Bartletts test i denne sammenhengen) eller Levenes test, som er en ANOVA F-test basert på $|x_{ij} - \bar{x}_{i.}|$ i stedet for x_{ij}.
- Sistnevnte test er mer robust overfor antakelsen om normalfordeling.

Eksempel

Fettopptak i rotter.
Eksempel 11.1

Alternativ formulering

- Enveis variansanalyse kan reformuleres til en form som er lettere å utvide til flerveis variansanalyse og også lettere ses i sammenheng med lineær regresjon.

Alternativ formulering

- Enveis variansanalyse kan reformuleres til en form som er lettere å utvide til flerveis variansanalyse og også lettere ses i sammenheng med lineær regresjon.
- La

$$\bar{\mu} = \frac{1}{I}\sum_{i=1}^{I} \mu_i \quad \text{og} \quad \alpha_i = \mu_i - \bar{\mu}.$$

Alternativ formulering

- Enveis variansanalyse kan reformuleres til en form som er lettere å utvide til flerveis variansanalyse og også lettere ses i sammenheng med lineær regresjon.
- La

$$\bar{\mu} = \frac{1}{I}\sum_{i=1}^{I} \mu_i \quad \text{og} \quad \alpha_i = \mu_i - \bar{\mu}.$$

- Da er

$$X_{ij} = \bar{\mu} + \alpha_i + \varepsilon_{ij}, \quad \text{der} \sum_{i=1}^{I} \alpha_i = 0.$$

Alternativ formulering

- Enveis variansanalyse kan reformuleres til en form som er lettere å utvide til flerveis variansanalyse og også lettere ses i sammenheng med lineær regresjon.
- La

$$\bar{\mu} = \frac{1}{I}\sum_{i=1}^{I} \mu_i \quad \text{og} \quad \alpha_i = \mu_i - \bar{\mu}.$$

- Da er

$$X_{ij} = \bar{\mu} + \alpha_i + \varepsilon_{ij}, \quad \text{der} \sum_{i=1}^{I} \alpha_i = 0.$$

- Hypotesetesten kan da reformuleres til

$$H_0 : \alpha_1 = \ldots = \alpha_I = 0 \quad \text{mot} \quad H_a : \alpha_i \neq 0 \text{ for minst én } i.$$

Alternativ formulering

- Enveis variansanalyse kan reformuleres til en form som er lettere å utvide til flerveis variansanalyse og også lettere ses i sammenheng med lineær regresjon.
- La

$$\bar{\mu} = \frac{1}{I}\sum_{i=1}^{I} \mu_i \quad \text{og} \quad \alpha_i = \mu_i - \bar{\mu}.$$

- Da er

$$X_{ij} = \bar{\mu} + \alpha_i + \varepsilon_{ij}, \quad \text{der} \sum_{i=1}^{I} \alpha_i = 0.$$

- Hypotesetesten kan da reformuleres til

$$H_0 : \alpha_1 = \ldots = \alpha_I = 0 \quad \text{mot} \quad H_a : \alpha_i \neq 0 \text{ for minst én } i.$$

- Videre er
$\mathrm{E}(MSTr) = \sigma^2 + \frac{J}{I-1}\sum_{i=1}^{I}(\mu_i - \bar{\mu})^2 = \sigma^2 + \frac{J}{I-1}\sum_{i=1}^{I}\alpha_i^2.$