

Sensorveiledning og løsningsforslag

Emnet bruker karakterskala bestått/ikke bestått. Nedenfor følger løsning på alle oppgavene med poeng angitt i rødt. Sensorveiledningen er gitt til slutt i dokumentet.

Oppgave 1

1. Hvis vi setter styrken til 80% og signifikansnivået til 5%, hvor mange pasienter bør vi inkludere i studien? (6)

I STATA bruker man: “Statistics” -> “Power, precision and sample size”; og deretter: “Proportions” -> “Chi-squared test comparing two independent proportions”, og får resultatet nedenfor:

```
. power twoproportions 0.4 0.2, test(chi2)

Performing iteration ...

Estimated sample sizes for a two-sample proportions test
Pearson's chi-squared test
H0: p2 = p1 versus Ha: p2 != p1

Study parameters:

      alpha =    0.0500
      power =    0.8000
      delta =   -0.2000 (difference)
      p1 =     0.4000
      p2 =     0.2000

Estimated sample sizes:

      N =          164
      N per group =          82
```

Vi trenger 164 pasientene totalt, 82 i hver gruppe, til å få den bestemte teststyrke.

2. Dersom styrken øker til 90%, hvor mange pasienter trenger vi da i hver gruppe? (2)
Som forventet trenger man flere pasienter i studien når styrken øker: utvalgsstørrelsen må da være 218 (109 i hver gruppe).

```
. power twoproportions 0.4 0.2, test(chi2) power(0.9)

Performing iteration ...

Estimated sample sizes for a two-sample proportions test
Pearson's chi-squared test
H0: p2 = p1 versus Ha: p2 != p1

Study parameters:

      alpha =    0.0500
      power =    0.9000
      delta =   -0.2000 (difference)
      p1 =     0.4000
      p2 =     0.2000

Estimated sample sizes:

      N =          218
      N per group =          109
```

Studien ble gjennomført med 70 pasienter i hver gruppe.

3. Gitt denne utvalgsstørrelsen, hva var styrken til studien for å oppdage den samme reduksjonen i andelen pasienter med alvorlig tilstand? Kommentarer? (4)

ISTATA: “Statistics” -> “Power, precision and sample size”; og deretter: “Proportions” -> “Chi-squared test comparing two independent proportions”, men nå trenger man å velge «Power» i «Compute:». Når man setter utvalgsstørrelsen til 140 (70 i hver gruppe), får man en styrke på 73.8% (se på bilde nedenfor). Dette er, som forventet, lavere enn 80% fordi det ikke ble rekruttert nok pasienter til studien. Nedgangen i styrken er derimot ikke betydelig stor, så det kan være nok pasienter til å avdekke en betydelig effekt.

```
. power twoproportions 0.4 0.2, test(chi2) n(140)

Estimated power for a two-sample proportions test
Pearson's chi-squared test
H0: p2 = p1 versus Ha: p2 != p1

Study parameters:

      alpha =    0.0500
      N      =    140
N per group =    70
      delta =   -0.2000 (difference)
      p1     =    0.4000
      p2     =    0.2000

Estimated power:

      power =    0.7381
```

Av de pasientene som fikk morfin (gruppe 1), ble bare 16 av pasientenes tilstand evaluert som alvorlig 10 dager etter kirurgi, mens av de som fikk halotan ble 39 klassifisert som alvorlig (gruppe 2).

4. Estimer andelen pasienter med alvorlig tilstand i de to gruppene. Beregn 95% konfidensintervall for andelen. (6)

$$\hat{p}_1 = \frac{16}{70} = 0.23 \text{ med 95\% KI (0.13, 0.33)}$$

$$\hat{p}_2 = \frac{39}{70} = 0.56 \text{ med 95\% KI (0.44, 0.67)}$$

Til å beregne KI brukte vi formel: $\hat{p} \pm 1.96 \sqrt{\frac{\hat{p}*(1-\hat{p})}{n}}$

5. Undersøk om andelen pasienter med alvorlig tilstand er signifikant forskjellig i de to gruppene. Beregn RR med 95% KI, og gi en kort sammenfatning av resultatene du har funnet. (8)

Siden utvalgsstørrelsen er 140, kan man bruke en Kjikvadrat test til å sjekke sammenheng mellom tilstand og grupper. Det er nyttig å lage en frekvenstabell til å representere resultatene.

	morfin (gruppe 1)	halotan (gruppe 2)
alvorlig	16	39
ikke alvorlig	54	31

I STATA: “Statistics” -> “Epidemiology and related” -> “Tables for Epidemiologists” -> “Cohort study risk ratio etc. Calculator”, som gir resultatet nedenfor: det er en signifikant forskjell i andelen pasienter med alvorlig tilstand i de to gruppene ($p=0.001$). Endringen i anestesimidlet fører til en signifikant redusert risiko av alvorlig tilstand etter kirurgi, med RR = 0.41 og 95% KI til RR (0.25, 0.66).

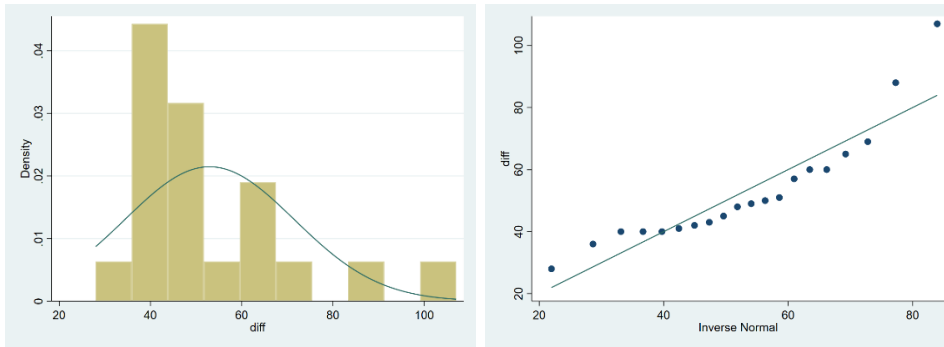
```
. csi 16 39 54 31
```

	Exposed	Unexposed	Total
Cases	16	39	55
Noncases	54	31	85
Total	70	70	140
Risk	.2285714	.5571429	.3928571
	Point estimate		[95% conf. interval]
Risk difference	-.3285714		-.4809419 -.176201
Risk ratio	.4102564		.254275 .6619225
Prev. frac. ex.	.5897436		.3380775 .745725
Prev. frac. pop	.2948718		

chi2(1) = 15.84 Pr>chi2 = 0.0001

Oppgave 2

1. *Hva slags data er dette? (2)*
 Pardata, analyse baseres på differansene.
2. *Har immunisering en signifikant innvirkning på konsentrasjonen av antipneumokokkantistoffer? Utfør en test og begrunn hvorfor du valgte denne. (10)*
 Siden analysen baseres på differansene, må man første beregne dem:
`generate diff = pneumo_etter - pneumo_for`
 «diff» er en ny variabel som inkluderer differansene i konsentrasjoner av antipneumokokkantistoffer (her, positive differansene betyr økende konsentrasjoner). Man kan gi en beskrivelse av differansene med flere grafiske måter (histogram, normalitetsplott), og man kan også bruke en Shapiro-Wilks test til å sjekke normalitetsfordeling av differansene. Både histogrammet og normalitetsplottet (nedenfor) viser ganske skjevfordelte differanser, og p-verdien til Shapiro-Wilks testen er $p = 0.008$. Derfor viser dette at man ikke kan anta normalfordeling, og siden utvalgsstørrelsen er ganske lav (20 observasjoner) bruker man en ikke-parametriske metode å analysere differansene.



. swilk diff

Shapiro-Wilk W test for normal data

Variable	Obs	W	V	z	Prob>z
diff	20	0.86157	3.277	2.392	0.00838

Siden differansene er ganske skjevfordelte, er det bedre å bruke Sign test (og ikke Wilcoxon signed rank test) for paradata. Det gir en veldig lav p-verdien ($p < 0.001$), som betyr at immunisering har en signifikant innvirkning på konsentrasjonen av antipneumokokkantistoffer.

Sign test

Sign	Observed	Expected
Positive	20	10
Negative	0	10
Zero	0	0
All	20	20

One-sided tests:

H0: median of diff = 0 vs.

Ha: median of diff > 0

Pr(#positive >= 20) =

Binomial(n = 20, x >= 20, p = 0.5) = 0.0000

H0: median of diff = 0 vs.

Ha: median of diff < 0

Pr(#negative >= 0) =

Binomial(n = 20, x >= 0, p = 0.5) = 1.0000

Two-sided test:

H0: median of diff = 0 vs.

Ha: median of diff != 0

Pr(#positive >= 20 or #negative >= 20) =

min(1, 2*Binomial(n = 20, x >= 20, p = 0.5)) = 0.0000

3. *Estimer differansen i antipneumokokkantistoffer konsentrasjon fra før til etter immunisering. Gi også 95% konfidensintervall for differansen. (6)*

Siden differansene i antipneumokokkantistoffer konsentrasjon ikke er normalfordelte, må man bruke median og 95% KI for medianen. Medianen til differansene er 48.5, med 95% KI (41.1, 59.6) – se STATA nedenfor.

. summarize diff, detail

diff				
Percentiles		Smallest		
1%	28	28		
5%	32	36		
10%	38	40	Obs	20
25%	40.5	40	Sum of wgt.	20
50%	48.5		Mean	52.95
		Largest	Std. dev.	18.57977
75%	60	65		
90%	78.5	69	Variance	345.2079
95%	97.5	88	Skewness	1.481526
99%	107	107	Kurtosis	4.993227

. centile diff

Variable	Obs	Percentile	Centile	Binom. interp. [95% conf. interval]	
diff	20	50	48.5	41.11647	59.65059

Oppgave 3

1. Lag en tabell som viser antall barn, gjennomsnitt og standardavvik for variablene indeks og timer, for jenter og gutter. (4)

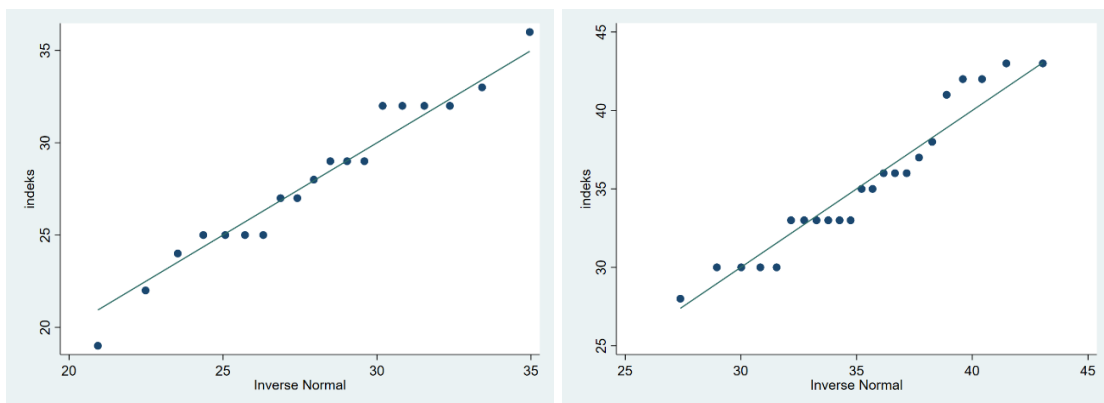
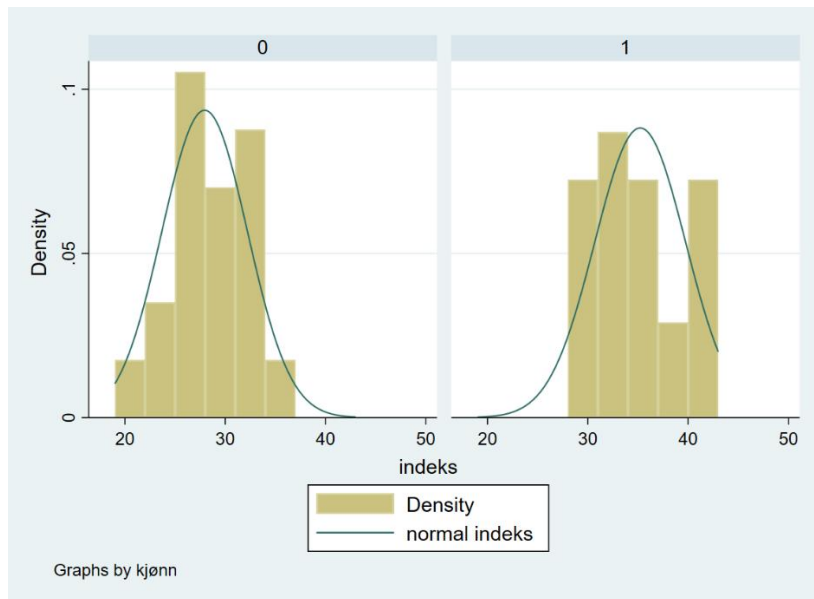
	Antall	Gjennomsnitt	Standardavvik
Indeks			
Jenter	19	27.9	4.26
Gutter	23	35.2	4.52
Timer			
Jenter	19	6.32	0.96
Gutter	23	8.13	0.86

2. Undersøk om det er en signifikant forskjellig indeks hos jenter og gutter. Utfør en test og begrunn hvorfor du valgte denne. Angi også effektmål med konfidensintervall. (6)
Vi skal sammenligne to uavhengige grupper, derfor trenger vi først å undersøke om normalitetsforutsetningen er oppfylt i hver gruppe: histogrammet, normalitetsplottet og resultatet av Shapiro-Wilks testen er nedenfor. Det ser ut som normalitetsforutsetningen er oppfylt. Derfor kan vi bruke to-utvalgs t-test med kommandoen: «ttest indeks, by(kjønn)». P-verdien fra t-testen er < 0.001 , så det er signifikant forskjellig indeks av selvtillit hos jenter og gutter.

Estimert differanse i gjennomsnitt mellom de to gruppene: -7.27*

*jenter i forhold til gutter

95% KI for differanse i gjennomsnitt: (-10.03, -4.51)



Normalitetsplottet for jenter (venstre side) og gutter (høyre side).

```
. swilk indeks if kjønn == 1
```

Shapiro-Wilk W test for normal data

Variable	Obs	W	V	z	Prob>z
indeks	23	0.95235	1.246	0.448	0.32719

```
. swilk indeks if kjønn == 0
```

Shapiro-Wilk W test for normal data

Variable	Obs	W	V	z	Prob>z
indeks	19	0.98727	0.291	-2.482	0.99347

Shapiro-Wilks normalitetsfordelings testen for jenter gir p-verdien 0.99, og p-verdien for gutter er 0.33. Normalitetsfordeling kan antas i begge grupper.

3. *Undersøk ved hjelp av lineær regresjon om det er en signifikant sammenheng mellom selvtillit hos barnet («indeks») og friluftsliv («timer»). Angi effektmål med konfidensintervall. Hvordan tolkes regresjonskoeffisienten? (8)*

Svarene finnes fra STATA output nedenfor. Regresjonskoeffisienten er 3.75 med 95% KI (2.98, 4.51). Effekten er veldig signifikant ($p < 0.001$). Regresjonskoeffisienten betyr at barna blir 3.75 ganger mer selvstendig (på «indeks» poengskalaen) for hver ekstra time de tilbringer i naturen.

. regress indeks timer

Source	SS	df	MS	Number of obs	=	42
Model	941.180628	1	941.180628	F(1, 40)	=	97.63
Residual	385.605086	40	9.64012715	Prob > F	=	0.0000
				R-squared	=	0.7094
				Adj R-squared	=	0.7021
Total	1326.78571	41	32.3606272	Root MSE	=	3.1049

indeks	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
timer	3.746206	.3791373	9.88	0.000	2.979941	4.512471
_cons	4.536672	2.813309	1.61	0.115	-1.149237	10.22258

4. Angi også R og R^2 for regresjonsmodellen i 3. Hvordan tolker du disse? (2)

R er Pearsons korrelasjonskoeffisient. Det er et mål for lineær sammenheng mellom de to variablene indeks og timer, og finnes via kommandoen «pwwcorr», som gir $R = 0.842$. Det indikerer en ganske sterk lineær sammenheng.

$R^2 = 0.709$ (fra STATA output). R^2 er forklart varians. Denne regresjonsmodellen med timer forklarer 70.9% av variasjonen i indeks.

5. På bakgrunn av analysene i 2. og 3. utførte forskerne en multippel regresjonsanalyse (med «indeks» som avhengig variabel). Hvilke variabler vil du ha med i den multiple regresjonsmodellen, og hva betyr det i forhold til konklusjonen i 2.? Begrunn svaret. Hva er R^2 for denne modellen? (8)

. regress indeks timer kjønn

Source	SS	df	MS	Number of obs	=	42
Model	946.292906	2	473.146453	F(2, 39)	=	48.50
Residual	380.492809	39	9.75622586	Prob > F	=	0.0000
				R-squared	=	0.7132
				Adj R-squared	=	0.6985
Total	1326.78571	41	32.3606272	Root MSE	=	3.1235

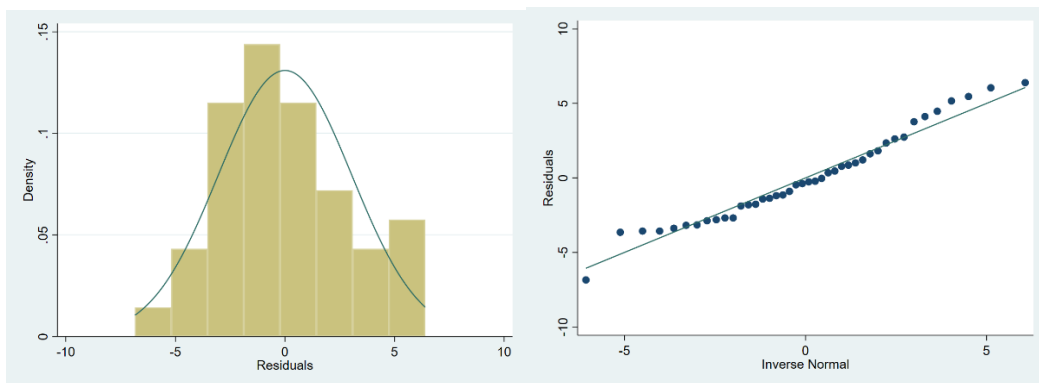
indeks	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
timer	3.465695	.5437286	6.37	0.000	2.3659	4.56549
kjønn	.9992571	1.380419	0.72	0.473	-1.792904	3.791418
_cons	6.04053	3.510843	1.72	0.093	-1.060821	13.14188

STATA viser at variabelen «timer» er fortsatt signifikant i multippel regresjonsmodellen for «indeks» (p -verdien < 0.001), men «kjønn» er ikke det (p -verdien = 0.47). Dette betyr at effekten av kjønn på indeks ikke er signifikant da vi

justerer for variabel «timer», og derfor kan effekten som vi estimerte i 2. forklares med at gutter bruker i gjennomsnitt mer tid enn jenter i naturen.
 $R^2=0.713$ for den multiple regresjonsmodellen.

6. Er forutsetningene tilfredsstillt for å gjøre multippel regresjon? Begrunn svaret. (6)
 Residual analyse viser at forutsetningene er oppfylt:

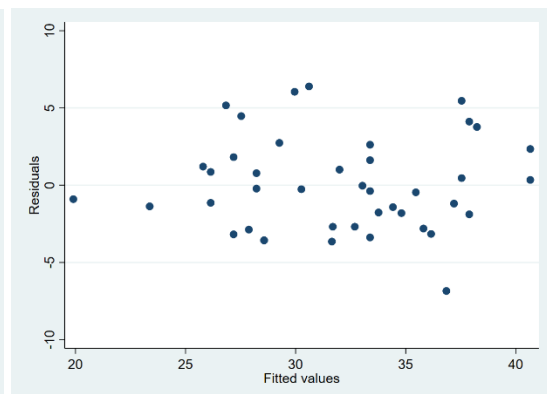
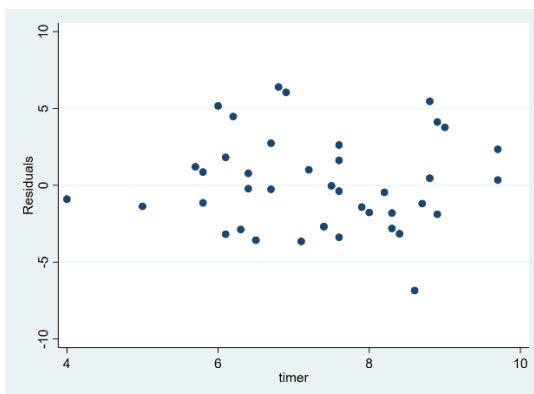
- Uavhengighet;
- Residualene er tilnærmet normalfordelte (histogram og q-q plott);
- Alle standardiserte residualer bortsett fra en er innenfor intervallet fra -3 til 3;
- Plott med standardiserte residualer på y-aksen og kovariat «timer» på x-aksen viser at residualene fordeler seg 'tilfeldig' rundt verdien 0.
- Plott med standardiserte residualer på y-aksen og standardiserte predikerte verdier på x-aksen viser at residualene fordeler seg 'tilfeldig' rundt verdien 0.



```
. predict rstd, rstandard
```

```
. summarize rstd, detail
```

Standardized residuals				
	Percentiles	Smallest		
1%	-2.248496	-2.248496		
5%	-1.174064	-1.214184		
10%	-1.111099	-1.174064	Obs	42
25%	-.8867633	-1.174064	Sum of wgt.	42
50%	-.1063175		Mean	-.0003868
		Largest	Std. dev.	1.008638
75%	.5986986	1.701815	Variance	1.017351
90%	1.471361	1.80085	Skewness	.353898
95%	1.80085	1.999498	Kurtosis	2.598798
99%	2.154175	2.154175		



7. Lag en tabell som oppsummerer resultatene av studien. Gi estimerte regresjonskoeffisienter, konfidensintervaller og p-verdier. Gi både de ujusterte og de justerte resultatene. (6)

Variabel	Ujustert effekt			Justert effekt		
	effekt	95% KI	p-verdi	effekt	95% KI	p-verdi
Timer	3.75	(2.98, 4.51)	<0,001	3.47	(2.37, 4.57)	<0,001
Kjønn*	-7.27*	(-10.03,-4.51)	<0,001	.99	(-1.79, 3.79)	0.47

*jenter i forhold til gutter

Oppgave 4

1. Forklar hva denne testens andel med falske positive og falske negative betyr i denne sammenheng. (4)

Andel med falske positive: for 3% av personer som ikke har COVID-19, vil denne hjemmetesten indikere at de har sykdommen.

Andel med falske negative: for 5% av personer som har COVID-19, vil denne hjemmetesten indikere at de ikke har sykdommen.

2. Hva tror du er mest alvorlig: et falskt positivt eller et falskt negativt utslag? Begrunn svaret ditt. (2)

Et falsk negativt utslag er mer alvorlig. Folk som får et falskt negativt utslag tror de er friske, og dermed vil de kunne spre sykdommen. Folk som får et falsk positivt utslag kan (og gjør det vanligvis) testes på nytt for å bekrefte, eller bare sette i karantene for å unngå å spre viruset til andre mennesker.

Godkjenning av testen ble gjort etter en studie med 4000 symptomatiske personer som brukte hjemmetesten for COVID-19 først, og deretter ble testet med PCR for å bekrefte diagnosen. Tabellen nedenfor viser resultatene fra denne studien.

		PCR	
		Positiv	Negativ
Hjemmetest	Positiv	190	114
	Negativ	10	3686

3. Beregn andelen med positiv hjemmetest som er korrekt klassifisert. Hva kaller vi denne sannsynligheten? (5)

Denne sannsynligheten kalles PPV (Positiv Prediktiv Verdi).

$$PPV = 190/(190+114) = 0.625$$

4. Beregn andelen med negativ hjemmetest som er korrekt klassifisert. Hva kaller vi denne sannsynligheten? (5)

Denne sannsynligheten kalles NPV (Negative Prediktiv Verdi).

$$NPV = 3686/(3686+10) = 0.997$$

Sensorveiledning:

Oppgave	Maksimal poengsum
1	26
2	18
3	40
4	16
Sum	100

Krav til bestått:

- Grensen for bestått er 65 av 100 mulige poeng.
- Besvarelsen skal vise at kandidaten har bred faglig kunnskap innen hele emnet og god kunnskap innen de mest sentrale områdene. Ingen av disse oppgavene er utenfor de mest sentrale områdene. Dette betyr at manglende/totalt utilfredsstillende besvarelse av enkelte oppgaver ikke kan kompenseres ved svært god besvarelse av andre.
- Bruken av fagterminologi er tilfredsstillende og besvarelsen er klar og presis.
- Kandidaten viser god faglig vurderingsevne og selvstendighet.