

ISTx1001: Industriell statistikk

6. november 2023

Vi starter klokka 15:15

Timen blir tatt opp :)

Thea Bjørnland, Institutt for matematiske fag, NTNU

Prosjektarbeid i grupper

Uke ~~43~~, ~~44~~, 45, 46. Innlevering mandag uke 47

Undervisning

~~Del 1: Uke 43 og 44~~

~~Del 2/3: Uke 45 og 46~~

Veiledning

Alle uker

Innlevering

FØR
20. november
12:00

Plan

Se Blackboard for
utfyllende info

Forelesninger (små endringer kan forekomme)

- Mandag 23. oktober 15:15-16:00: Introduksjon og informasjon om prosjektet
- Tirsdag 24. oktober 14:15-16:00: Forsøksplanlegging (del 1)
- Mandag 30. oktober 15:15-16:00: Forts. forsøksplanlegging (del 1)
- Tirsdag 31. oktober 14:15-16:00 Forts. forsøksplanlegging (del 1),
- **Mandag 6. november 15:15-16:00: Statistisk kvalitetsstyring (del 2)**
- Mandag 13. november 15:15-16:00: Statistisk kvalitetsstyring (del 2)
- Forelesningene vil i all hovedsak utdype prosjektoppgavene og svare på spørsmål. Pensum dekkes av kompendium, videoer og notatbøker.

Plan

Se Blackboard for
utfyllende info

Veiledning

- Digitalt:
 - Forumet er alltid åpent og vi følger med. Spørsmål fra forum kan også bli tatt opp i zoom-timer
 - Digitale veiledningsmøter ved behov
- Fysisk:
 - Trondheim: Torsdag ~~26/10, 2/11~~, 9/11, 16/11 14:15-16:00 i EL5
Tirsdag 7/11, 14/11 i F1
 - Gjøvik: Torsdag ~~26/10, 2/11~~, 9/11, 16/11 14:15-16:00 i A-160
Tirsdag 7/11, 14/11 i A-160
 - Ålesund: Avtal veiledning med Siebe

Del 2 av prosjektet i ISTx1001

Ressurser:

- ✧ Læreboka kapittel 11
- ✧ Temavideoer (laget av Ketil Arnesen)
- ✧ Forelesninger med slides



inkludere mye
Python

Spesielt mandag 13-nov

Statistisk kvalitetskontroll

- 
- [Shewharts kontrollidiagram](#)
 - [s-diagram og p-diagram](#)
 - [Karakteristikk](#)
 - [Kapabilitet](#)

Del 2 av prosjektet i ISTx1001

Oppgave 2 - Statistisk prosesskontroll

Målet med oppgaven er å **bestemme kontrollgrenser** i en prosesskontroll, samt visualisere og forklare hva dette innebærer ved hjelp av stokastiske simuleringer.

Tenk dere at dere skal lage et produkt basert på forsøket dere gjennomførte i Oppgave 1, f.eks muffins, skråplan med ball eller papirhelikopter. Dere må selv velge nivåer for forklaringsvariablene som produktet skal produseres under. En stor produksjon skal settes i gang, og responsvariabelen fra Oppgave 1 skal brukes for å kontrollere (med stikkprøver) at prosessen er under kontroll.

Eks: Papirbroa

Hvitt papir ($x_1 = +1$)

8cm ($x_2 = -1$)

2 brett ($x_3 = +1$)

Husk:

responsen er beregnet (gram)

Husk også

$$Y \sim N(\mu, \sigma)$$

↑↑

må hente estimater for disse fra Del 1.

Del 2 av prosjektet i ISTx1001

2a) Kontrollgrenser

Finn en estimert forventningsverdi og et standardavvik på Y fra Oppgave 1 for den valgte faktorkombinasjonen. Dette setter dere som verdier på μ og σ for en prosess som er under kontroll. Bestem hvor mange (n) produkter dere vil kontrollere i hver stikkprøve.

Bestem kontrollgrenser for gjennomsnittet til en stikkprøve. Bruk Shewarts' anbefaling om 3-sigma-grenser.

Bestem kontrollgrenser for det empiriske standardavviket til en stikkprøve basert på kritiske verdier fra kjikvadratfordelingen med $\alpha/2 = 0.002$.

$Y \sim N(\mu, \sigma)$

↑
broas
bæreevne
ved valgte innstillinger

hva er estimator på
 μ og σ fra Del 1?

Python

```
print(df)
```

	x1	x2	x3	y
0	-1.0	-1.0	-1.0	8.70
1	1.0	-1.0	-1.0	17.40
2	-1.0	1.0	-1.0	0.00
3	1.0	1.0	-1.0	8.70
4	-1.0	-1.0	1.0	8.70
5	1.0	-1.0	1.0	30.45
6	-1.0	1.0	1.0	4.35
7	1.0	1.0	1.0	13.05
0	-1.0	-1.0	-1.0	4.35
1	1.0	-1.0	-1.0	17.40
2	-1.0	1.0	-1.0	4.35
3	1.0	1.0	-1.0	8.70
4	-1.0	-1.0	1.0	8.70
5	1.0	-1.0	1.0	34.80
6	-1.0	1.0	1.0	4.35
7	1.0	1.0	1.0	21.75

```
modell = smf.ols('y~x1+x2+x3 + x1*x2 + x2*x3 + x1*x3', data=df).fit()
```

```
modell.predict()
```

```
array([ 5.709375, 18.215625,  2.990625,  7.884375,  9.515625, 31.809375,  
        3.534375, 18.215625,  5.709375, 18.215625,  2.990625,  7.884375,  
        9.515625, 31.809375,  3.534375, 18.215625])
```

observert bærceve $y = 30,45$ gram

$y = 34,80$ gram

$\hat{y} = 31,81$ gram

estimert forventningsverdi
for $\bar{Y}$ ved valgt faktorkombinasjon

$\mu = 31,81$ gram
(estimert fra Del 1)

Python

```
print(df)
```

	x1	x2	x3	y
0	-1.0	-1.0	-1.0	8.70
1	1.0	-1.0	-1.0	17.40
2	-1.0	1.0	-1.0	0.00
3	1.0	1.0	-1.0	8.70
4	-1.0	-1.0	1.0	8.70
5	1.0	-1.0	1.0	30.45
6	-1.0	1.0	1.0	4.35
7	1.0	1.0	1.0	13.05
0	-1.0	-1.0	-1.0	4.35
1	1.0	-1.0	-1.0	17.40
2	-1.0	1.0	-1.0	4.35
3	1.0	1.0	-1.0	8.70
4	-1.0	-1.0	1.0	8.70
5	1.0	-1.0	1.0	34.80
6	-1.0	1.0	1.0	4.35
7	1.0	1.0	1.0	21.75

```
modell = smf.ols('y~x1+x2+x3 + x1*x2 + x2*x3 + x1*x3', data=df).fit()
```

```
modell.mse_resid
```

8.541406249999998

← estimat på σ^2

estimat på $\sigma = \sqrt{8.54, \dots}$
 $= 2.92 \text{ gram}$

Fra Del 1:

$\bar{Y} \sim N(31.81, 2.92)$

↑ beregne ved hvitt pipi, 8 cm, 2 Brett.

← estimert

Stokastisk simulering: normal prosess

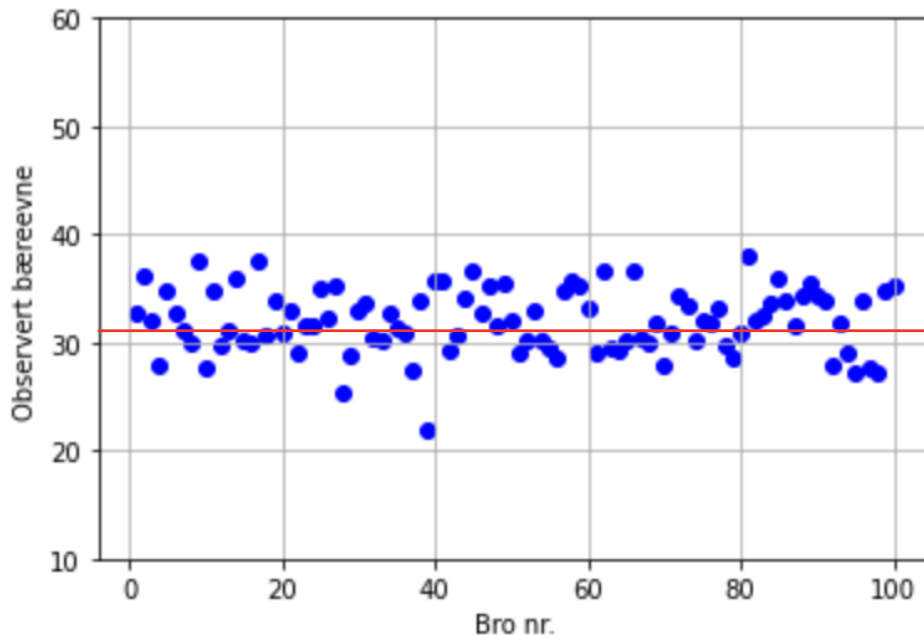
$\mu =$
 $\sigma =$
→ mean = 31.81
std_dev = 2.92
antall_broer = 100
bæreevne = np.zeros(antall_broer)

```
for i in range(antall_broer):  
    bæreevne[i] = np.random.normal(mean, std_dev)  
  
bro_nr = np.arange(1, antall_broer + 1)  
  
plt.scatter(bro_nr, bæreevne, marker='o', color='b')  
plt.ylim(10, 60); plt.xlabel("Bro nr."); plt.ylabel("Observert bæreevne")  
plt.grid()  
plt.show()
```

trekker et tall
fra $N(31.81, 2.92)$ fordeling

Antagelsen:

$$Y \sim N(31.81, 2.92)$$



$$\mu = 31.81$$

Prosess som
er under
kontroll

Stokastisk simulering: hva kan gå galt?

mål med prosesskontroll:
oppdage systematiske kvalitetsavvik

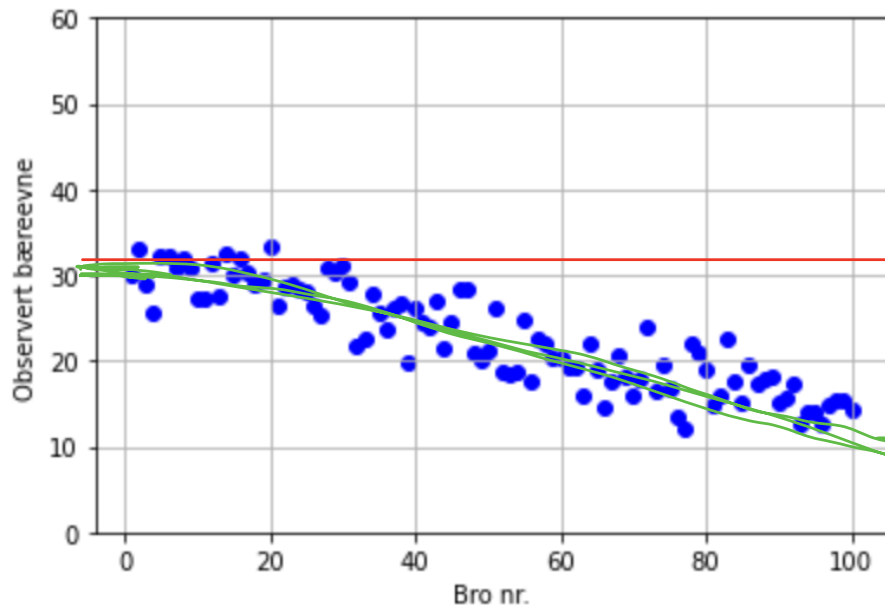
Stokastisk simulering: hva kan gå galt?

```
mean = 31.81
std_dev = 2.92
antall_broer = 100
bæreevne = np.zeros(antall_broer)

for i in range(antall_broer):
    if (i + 1) % 5 == 0:
        mean -= 0.3 * std_dev
        bæreevne[i] = np.random.normal(mean, std_dev)

bro_nr = np.arange(1, antall_broer + 1)

plt.scatter(bro_nr, bæreevne, marker='o', color='b')
plt.ylim(0,60); plt.xlabel("Bro nr."); plt.ylabel("Observert bæreevne")
plt.grid()
plt.show()
```



$\mu = 31,81$

forventningsverdien synker

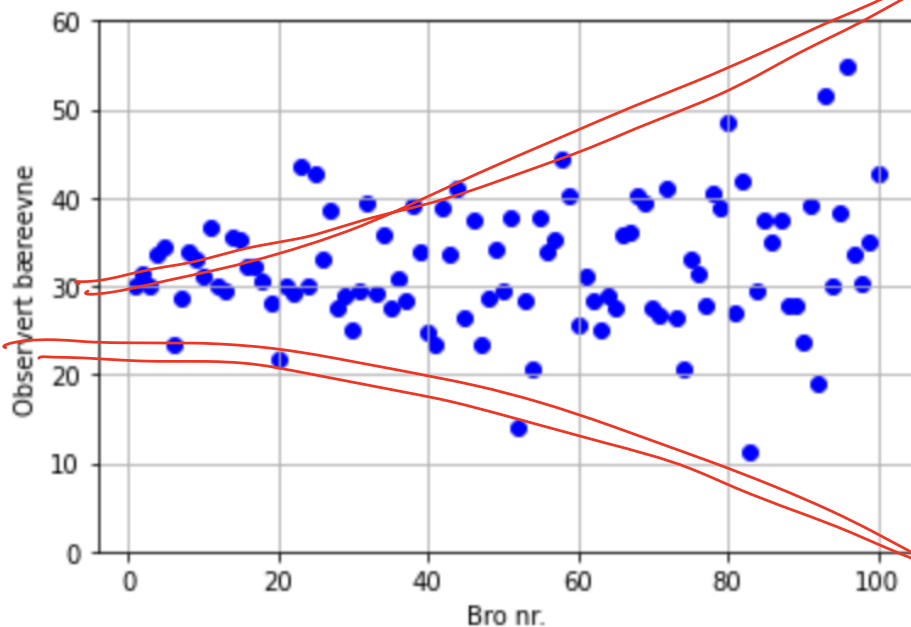
Stokastisk simulering: hva kan gå galt?

```
mean = 31.81
std_dev = 2.92
antall_broer = 100
bæreevne = np.zeros(antall_broer)

for i in range(antall_broer):
    if (i + 1) % 5 == 0:
        std_dev += 0.5
        bæreevne[i] = np.random.normal(mean, std_dev)

bro_nr = np.arange(1, antall_broer + 1)

plt.scatter(bro_nr, bæreevne, marker='o', color='b')
plt.ylim(0,60); plt.xlabel("Bro nr."); plt.ylabel("Observert bæreevne")
plt.grid()
plt.show()
```



Stokastisk simulering: hva kan gå galt?

Vi må følge med på forventningsverdi og standardavvik!

- Vi vil ikke teste alle brøer som produseres
(veldig kostbar)

- Teste et lite utvalg med jevne mellomrom,
f.eks $n=5$ brøer annenhver dag

⇓
"stikkprøve"

↙
gjennomsnitt
som et estimat
på forventningsverdi

→ empirisk
standardavvik (s)
som et estimat
på standardavviket (σ)

Del 2 av prosjektet i ISTx1001

2a) Kontrollgrenser

Finn en estimert forventningsverdi og et standardavvik på Y fra Oppgave 1 for den valgte faktorkombinasjonen. Dette setter dere som verdier på μ og σ for en prosess som er under kontroll. Bestem hvor mange (n) produkter dere vil kontrollere i hver stikkprøve.

Bestem kontrollgrenser for gjennomsnittet til en stikkprøve. Bruk Shewarts' anbefaling om 3-sigma-grenser.

Bestem kontrollgrenser for det empiriske standardavviket til en stikkprøve basert på kritiske verdier fra kjikvadratfordelingen med $\alpha/2 = 0.002$.

Hvis $\bar{y}$ eller s fra en stikkprøve er utenfor kontrollgrensene $\rightarrow$ Alarm!

Men, vi vil ikke at alarmen skal gå for ofte dersom det ikke er noe feil.

3-sigma-grenser $\rightarrow$ sannsynlighet på 0,0027 for falsk alarm.

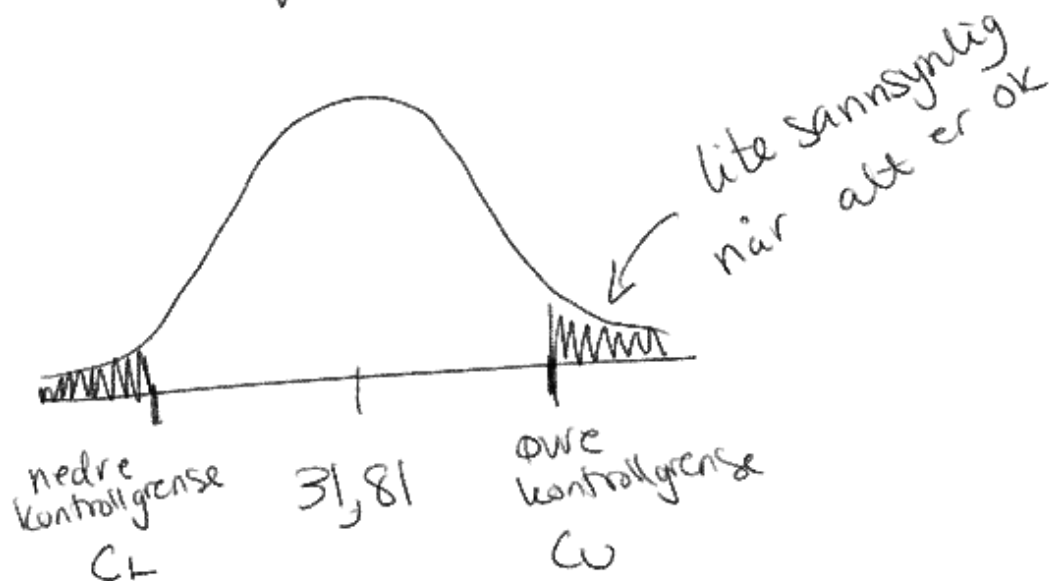
Kontrollgrenser for gjennomsnittet

Husk: $Y \sim N(\mu, \sigma)$ $\mu = 31,81$ } når alt er OK
 $\sigma = 2,92$ } "under kontroll"

Stikprøve $Y_1, Y_2, \dots, Y_n$ $Y_i \sim N(\mu, \sigma)$

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

↓ fordeling



$n=5$

Shewarts 3-sigma-regel

$$C_L = \mu - 3 \frac{\sigma}{\sqrt{n}}$$

$$C_U = \mu + 3 \frac{\sigma}{\sqrt{n}}$$

$$C_L = 31,81 - 3 \frac{2,92}{\sqrt{5}} = 27,89 \text{ g}$$

$$C_U = 31,81 + 3 \cdot \frac{2,92}{\sqrt{5}} = 35,72 \text{ g}$$

Y : representerer bæreevne til en papirbro med
produksjonsinnstillinger
hvitt papir, 8cm, 2 brettinger

$$\mu = 31,81 \text{ gram} \quad \sigma = 2,92 \text{ gram}$$

estimerer fra
Del 1 av
prosjektet,
som vi nå
tenker på som
sanne verdier

I produksjonsprosess tar vi en stikkprøve
med jevne mellomrom, eks: $n=5$ papirbroer
og regner ut gjennomsnittlig bæreevne $\bar{y}$ og
empirisk standardavvik s .

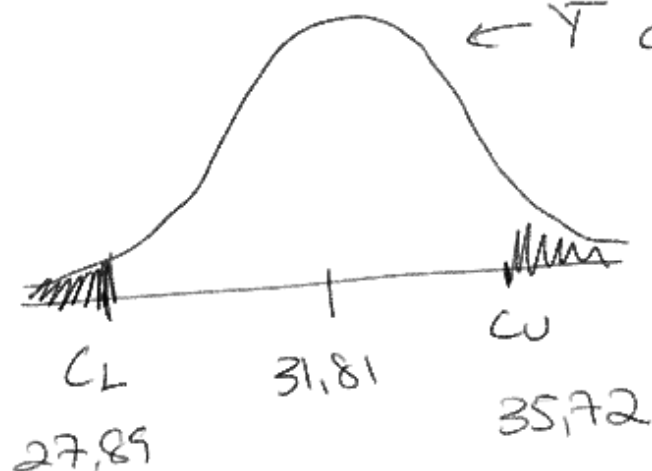
Kontrollgrenser for gjennomsnittet

Shewarts 3-sigma-regel

$$C_L = \mu - 3\frac{\sigma}{\sqrt{n}}$$

$$C_U = \mu + 3\frac{\sigma}{\sqrt{n}}$$

← $\bar{Y}$ over n brøer i stikkprøve



$$P(\text{falsh alarm}) \approx 0,0027$$

Kontrollgrenser for standardavviket

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$$



estimator for variance

$$\frac{(n-1)}{\sigma^2} s^2 \sim \chi^2 \quad \begin{array}{l} \text{kjikkvadrat} \\ \text{chi-squared} \end{array}$$



$$C_L = \sigma \cdot \sqrt{\frac{\chi^2_{\alpha/2}}{n-1}}$$

$$C_U = \sigma \sqrt{\frac{\chi^2_{1-\alpha/2}}{n-1}}$$

merk: $\chi^2_{\alpha/2}$ og $\chi^2_{1-\alpha/2}$ er kritiske verdier fra χ^2 -fordelingen med $n-1$ frihetsgrader.

0.002 OBS:
vanlig å velge
 $\frac{\alpha}{2} = 0.002$

Kontrollgrenser for standardavviket

C_L : `stats.chi2.ppf(0.002,4)`
0.12923771486623198 $\leftarrow \chi^2_{\alpha/2}$

$n=5$
 $n-1=4$

C_U : `stats.chi2.ppf(1-0.002,4)`
16.923758195804652 $\leftarrow \chi^2_{1-\alpha/2}$

$\mu = 31,81$ $\sigma = 2,92$ $n=5$

$$C_L: \sigma \sqrt{\frac{\chi^2_{\alpha/2}}{n-1}} = 2,92 \cdot \sqrt{\frac{0.13}{4}} = 0.53$$

$$C_U: \sigma \sqrt{\frac{\chi^2_{1-\alpha/2}}{n-1}} = 2,92 \cdot \sqrt{\frac{16.92}{4}} = 6.00$$

Del 2 av prosjektet i ISTx1001

2a) Kontrollgrenser

Finn en estimert forventningsverdi og et standardavvik på Y fra Oppgave 1 for den valgte faktorkombinasjonen. Dette setter dere som verdier på μ og σ for en prosess som er under kontroll. Bestem hvor mange (n) produkter dere vil kontrollere i hver stikkprøve.

Bestem kontrollgrenser for gjennomsnittet til en stikkprøve. Bruk Shewarts' anbefaling om 3-sigma-grenser.

Bestem kontrollgrenser for det empiriske standardavviket til en stikkprøve basert på kritiske verdier fra kjikvadratfordelingen med $\alpha/2 = 0.002$.

↑
Video i morgen.