

ISTx1001: Industriell statistikk

4. november 2024

Vi starter klokka 14:15

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
18. november
12:00 i INSPERA

Plan

Forelesninger (små endringer kan forekomme)

- Mandag 21. oktober 15:15-16:00: Introduksjon
- Tirsdag 22. oktober 14:15-16:00: Forsøksplanlegging (del 1)
- Mandag 28. oktober 15:15-16:00: Forts. forsøksplanlegging (del 1)
- Tirsdag 29. oktober 14:15-16:00 Forts. forsøksplanlegging (del 1)
- **Mandag 4. november 14:15-15:00: Statistisk kvalitetsstyring (del 2)**
- Mandag 11. november 14:15-15:00: Statistisk kvalitetsstyring (del 2)
- Forelesningene vil i all hovedsak utdype prosjektoppgavene og svare på spørsmål. Pensum dekkes av kompendium, videoer og notatbøker.

Tips til data-analyse i del 1

/ istx100y2024 /

9/17/24, 3:08 PM ▲	Last Modified
1001	18 seconds ago
1002	4 days ago
1003	16 days ago
forprosjekt	last month
uke1	last month
uke2	2 months ago
uke4	last month
uke5	last month
uke6	last month
uke7	27 days ago
uke8	last month
uke9	16 days ago
README.md	3 months ago
• velkommenIS...	2 months ago

/ istx100y2024 / 1001 /

Name ▲	Last Modified
•  DOE_EKS1.ipy...	6 days ago
•  DOE_EKS2.ip...	6 days ago
•  DOE_EKS3.ip...	10 days ago
•  DOE_introduk...	6 days ago
 DOE_multipp...	17 days ago
 DOEfunksjone...	30 seconds ago
 papirbro.ipynb	17 days ago
 produktunifor...	17 days ago

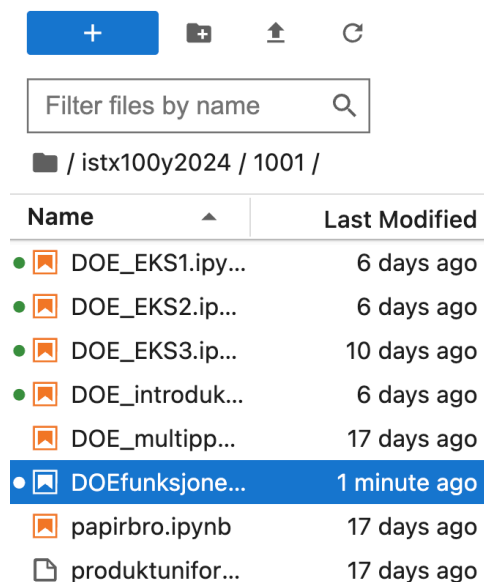
Tips til data-analyse i del 1

DOE eksempel 2: Produktuniformitet

Eksempelet og dataene er hentet fra National Institute of Standards and Technology (USA)

<https://www.itl.nist.gov/div898/handbook/pri/section3/pri3331.htm>

```
samspillsplott(data = df, k = 3)
```



Name	Last Modified
DOE_EKS1.ipynb	6 days ago
DOE_EKS2.ipynb	6 days ago
DOE_EKS3.ipynb	10 days ago
DOE_introduk...	6 days ago
DOE_multipp...	17 days ago
DOEfunksjone...	1 minute ago
papirbro.ipynb	17 days ago
produktunifor...	17 days ago



```
plt.show()

[39]: def samspillsplott(data, k,
                        faktornavn = ["A", "B", "C", "D", "E"],
                        variabelnavn = ["x" + str(i) for i in range(1, 6)]):

    # Sjekk om input er gyldig for ISTx1001
    if k <= 1 or k > 5:
        raise ValueError("Ugyldig antall faktorer k")

    farger = ['black', '#e41a1c']

    ymax = max(data['y'])
    ymin = min(data['y'])

    kolonnenavn = list(data.columns)

    if(k == 2):
        fig, axes = plt.subplots(1, k, figsize=(8, 4))
        i = 0
        for j in range(k):
            if(i==j):
                axes[j].axis('off')
                axes[j].annotate(variabelnavn[i]+" = -1", (0.5, 0.5), color = farger[i])
```

Tips til data-analyse i del 1

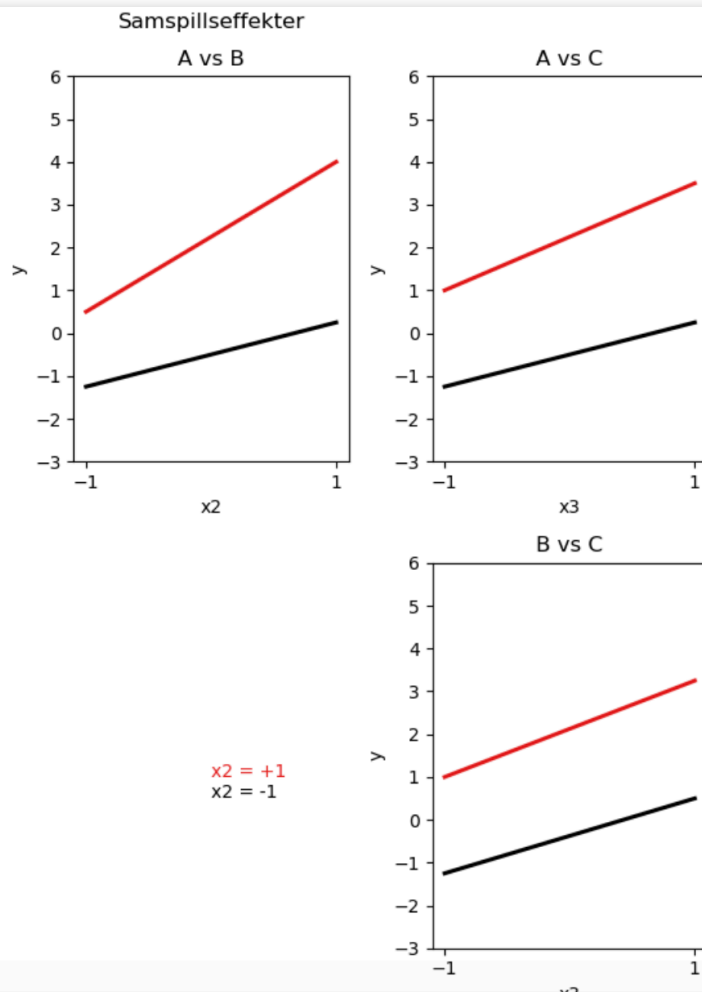
DOE eksempel 2: Produktuniformitet

Eksempelet og dataene er hentet fra National Institute of Standards and Technology (USA)

<https://www.itl.nist.gov/div898/handbook/pri/section3/pri3331.htm>

```
samspillsplott(data = df, k = 3)
```

x1 = +1
x1 = -1



Digital veiledning i morgen

<https://wiki.math.ntnu.no/istx100y/2024h/1001#veiledning>

Veiledning

Digitalt: Bruk [mattelab-forumet vårt](#) og underforumet 1001.

Zoom-veiledning Tirsdag 5. nov klokka 13:00 - 15:00, registrer tidspunkt [i dette skjemaet](#), zoom-lenke i skjemaet.

	A	B	C	D	E	F
1	Digital veiledning 5. november 2024 ISTx1001 prosjekt					
2	TID	NAVN (Dette navnet må dere også ha i zoom venterom)				
3	13:00					
4	13:10					
5	13:20					
6	13:30					
7	13:40					
8	13:50					
9	Pause					
10	14:10					
11	14:20					
12	14:30					
13	14:40					
14	14:50					
15						
16						
17	Zoom:	https://NTNU.zoom.us/j/95727089069?pwd=EbbcQvSNw4nvqAFveM9r1gFXY4941K.1				
18						

Del 2 av prosjektet i ISTx1001

Ressurser:

Læreboka kapittel 11

Temavideoer (laget av Ketil Arnesen)

Forelesninger med slides

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. 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.

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])
```

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
```

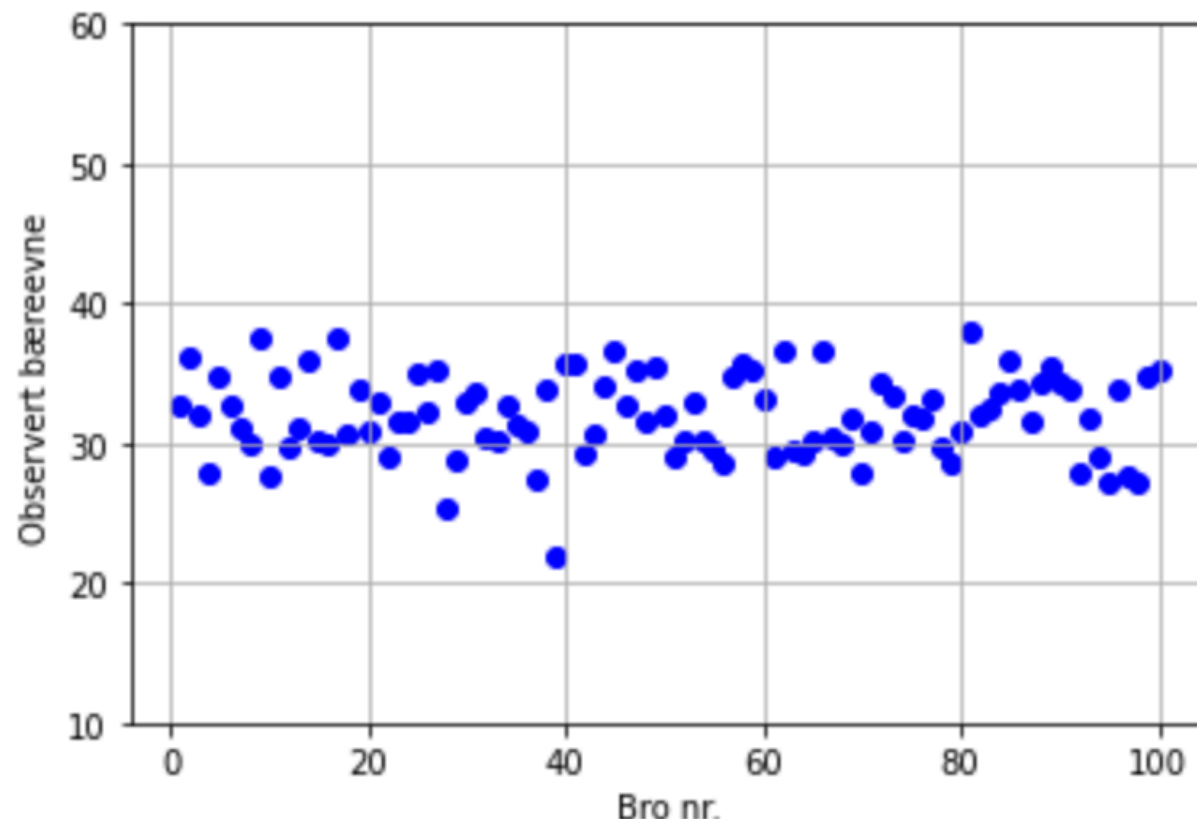
Stokastisk simulering: normal prosess

```
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()
```



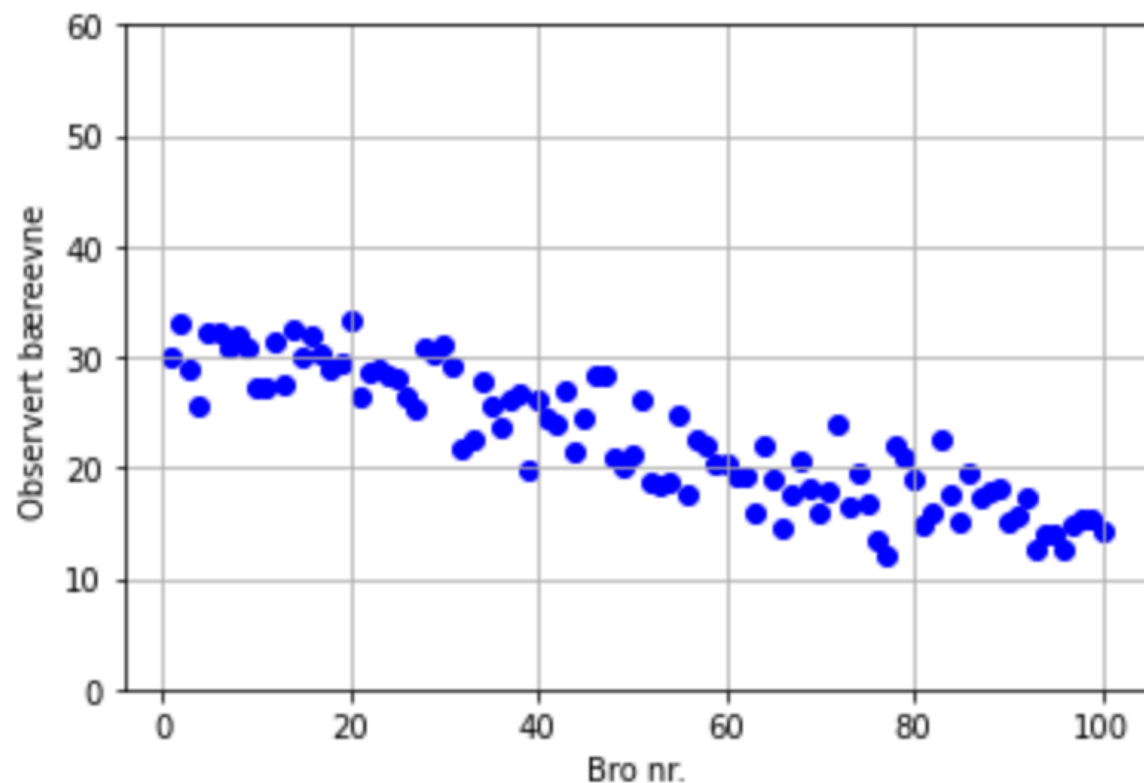
Stokastisk simulering: hva kan gå galt?

```
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()
```



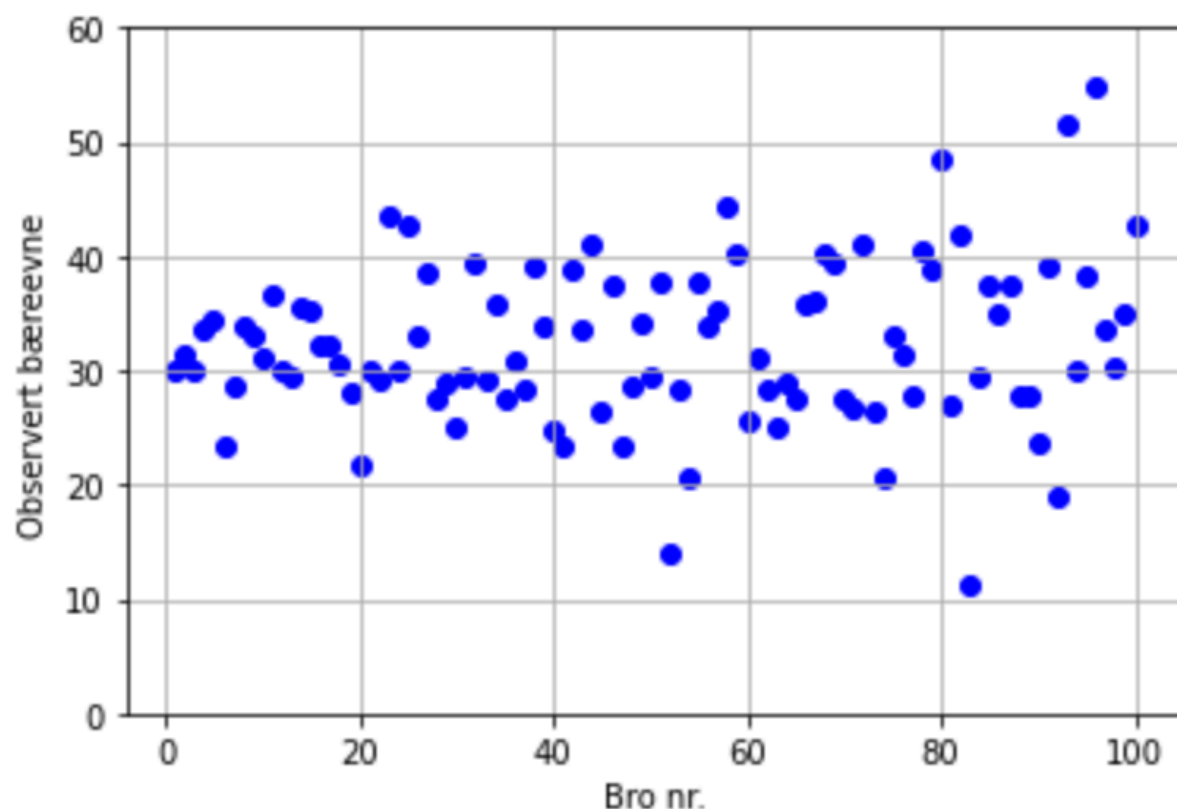
Stokastisk simulering: hva kan gå galt?

```
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!

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 $\alpha/2 = 0.002$.

Kontrollgrenser for gjennomsnittet

Kontrollgrenser for gjennomsnittet

Kontrollgrenser for standardavviket

Kontrollgrenser for standardavviket

```
stats.chi2.ppf(0.002,4)
```

0.12923771486623198

```
stats.chi2.ppf(1-0.002,4)
```

16.923758195804652

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 $\alpha/2 = 0.002$.