

ISTx1001: Industriell statistikk

24. oktober 2023

Vi starter klokka 14:15

Timen blir tatt opp :)

Thea Bjørnland, Institutt for matematiske fag, NTNU

I dag: Steg 1, 2, og 3 i Del 1 av prosjektet

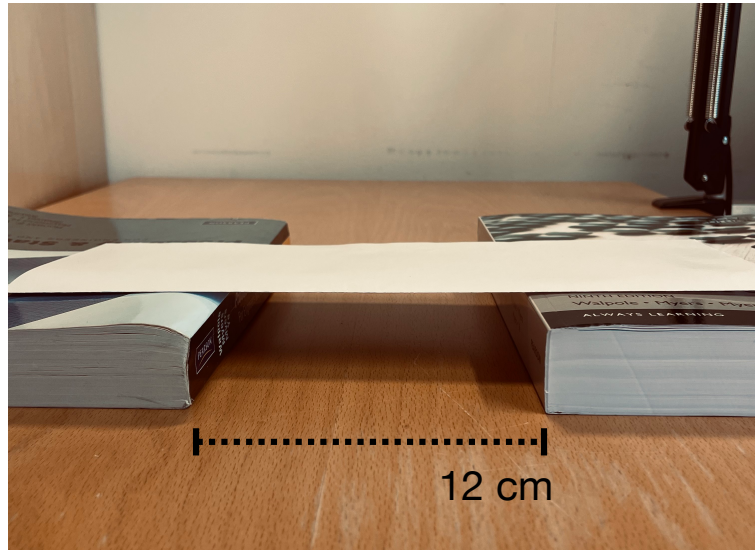
+ sammenhengen med lineær regresjon!

Oppgavebeskrivelse

Målet med prosjektet er å planlegge, gjennomføre og diskutere et 2-nivå faktorielt forsøk. Dere står selv fritt til å velge problemstilling og eksperiment. Det kan være et laboratorieeksperiment eller et problem fra dagliglivet. Her er stegene gruppa skal gjennomføre

1. Bestem dere for en problemstilling som dere kan studere med et 2^k faktorielt forsøk. Hva er responsvariabelen og hva er mulige forklaringsvariabler? Hva er egnede nivåer for forklaringsvariablene? Merk at responsvariabelen Y må være en kontinuerlig stokastisk variabel.
2. Design et forsøk med 16, 24 eller 32 målinger. Bruk enten 3 eller 4 faktorer. Dersom dere bruker 3 faktorer, må dere ha minst ett gjentak.
3. Planlegg innsamling av data.
4. Samle inn data.
5. Analyser data.
6. Skriv en rapport.

Papirbro - hvor mye vekt tåler den?



Responsvariabel: bæreevne (g)

Faktor: papirtype
↑

2 nivåer

lavt nivå: grått ark

høyt nivå: hvitt ark



4.35 gram

bæreevne (g)

8.7

0

grått

hvitt

Estimert effekt av å bytte
Papirtype = 8.7 g

Estimert effekt av å endre papirtype

Matematisk notasjon

Y : bæreevne (stokastisk variabel)

$$E(Y) = \beta_0 + \beta_1 x$$

↑
papirtype

$$\begin{aligned} x &= -1 \\ x &= +1 \end{aligned}$$

grått papir
hvitt papir

(lavt nivå)
(høyt nivå)

y : observasjon av Y

lavt nivå

$$E(Y | x = -1) = \beta_0 - \beta_1$$

høyt nivå

$$E(Y | x = +1) = \beta_0 + \beta_1$$

Effekt:

$$(\beta_0 + \beta_1) - (\beta_0 - \beta_1) =$$

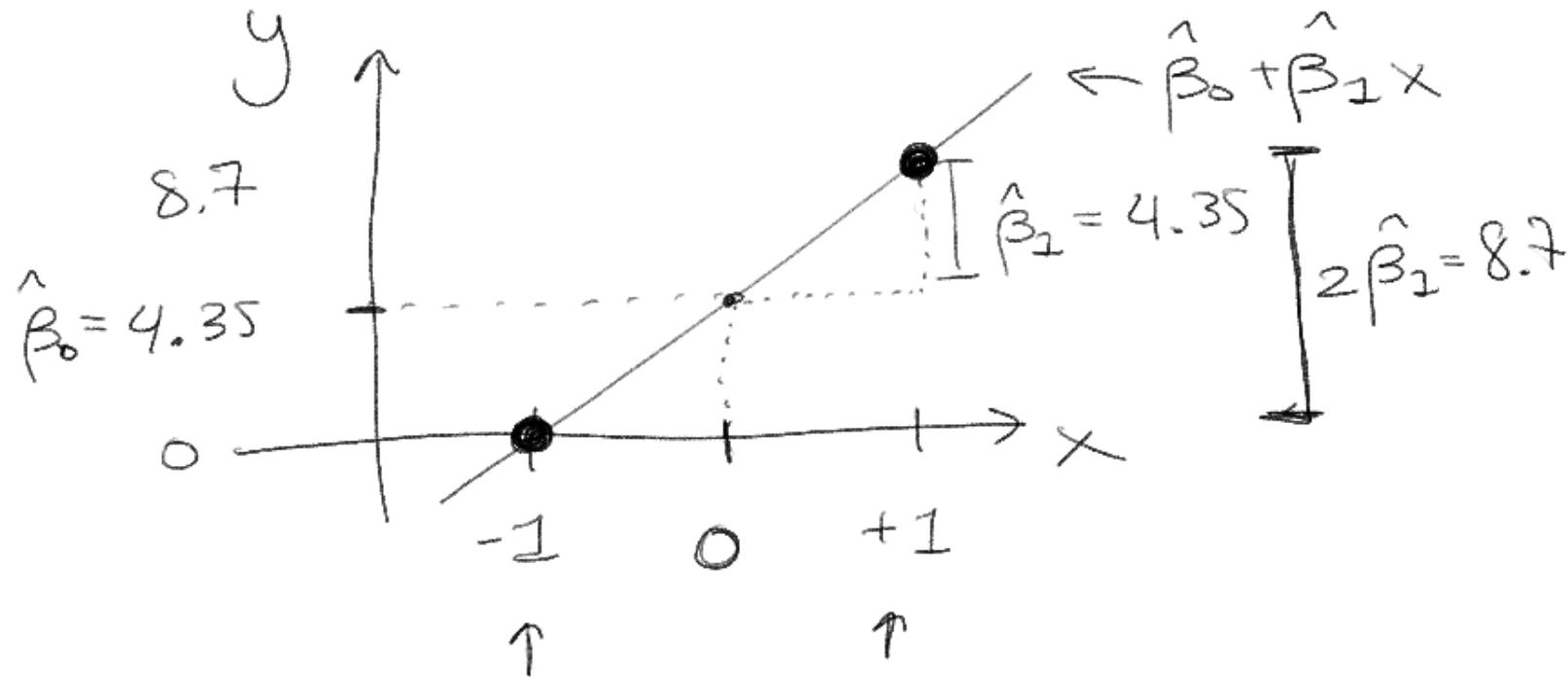
$$\beta_0 + \beta_1 - \beta_0 + \beta_1 =$$

$$\boxed{2\beta_1}$$

Estimert effekt av å endre papirtype

Datasett

x	y
-1	0
+1	8.7



Husk uke 9:

$$H_0: \beta_1 = 0 \quad \text{mot} \quad H_1: \beta_1 \neq 0$$

(+-test)

$$Y \sim \text{Normal}(\beta_0 + \beta_1 x, \sigma)$$

Estimert effekt av å endre papirtype

Signifikant effekt eller tilfeldig variasjon?

Testobservator for $H_0 : \beta_1 = 0$ mot $H_1 : \beta_1 \neq 0$:

$$T = \frac{\hat{\beta}_1 - 0}{\text{SE}(\hat{\beta}_1)} \underset{H_0}{\sim} t_{n-2}$$

$$\hat{\beta}_1 = 4.35$$

$$\text{SE}(\hat{\beta}_1) = \frac{S}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}}$$

S er estimat
på σ

$$S^2 = \frac{1}{n-2} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

antall obs
 $n=2$

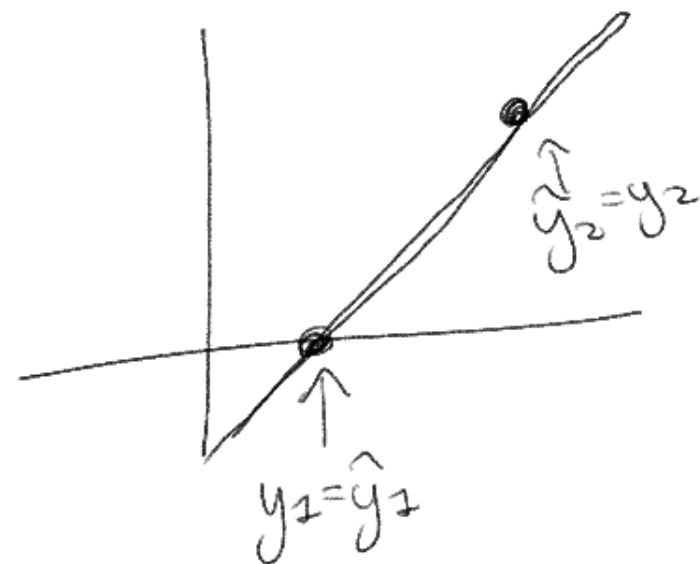
y_i $\hat{y}_i$

$$S^2 = \frac{0}{0}$$

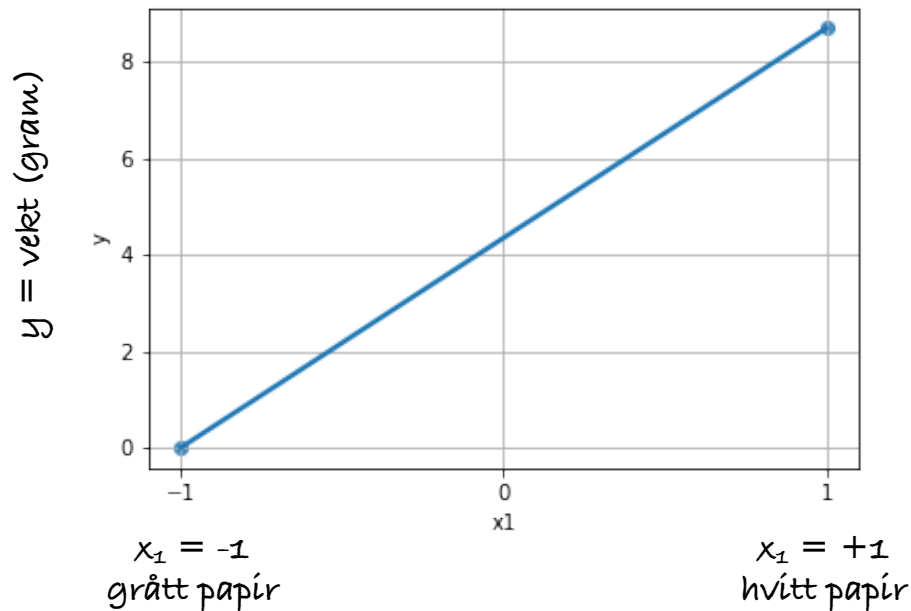
$n=2$

$\beta_2, \beta_0 \leftarrow 2$ parametre

estimere σ ?
Nei



Signifikant effekt eller tilfeldig variasjon?



	coef	std err	t	P> t	[0.025	0.975]
Intercept	4.3500 ← β_0	inf	0	nan	nan	nan
x1	4.3500 ← β_1	inf	0	nan	nan	nan

↑
summary()

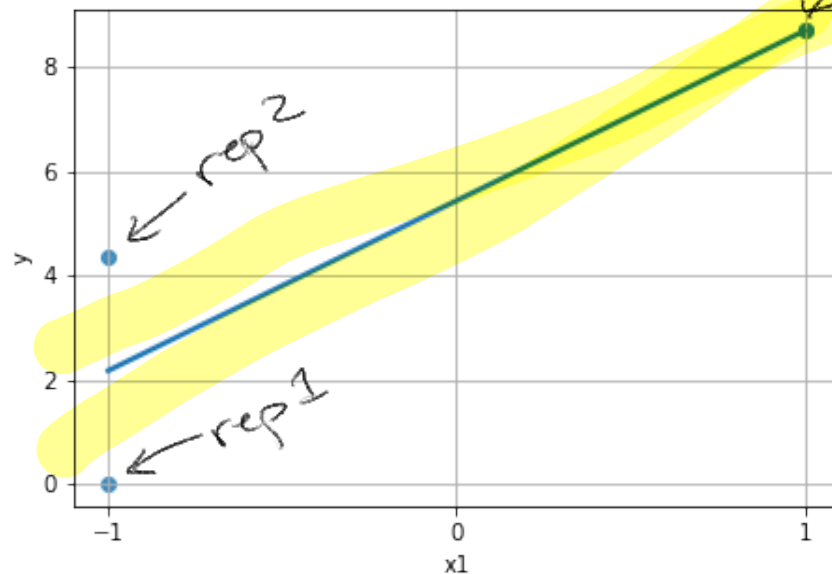
Papirbro - flere målinger

rep 1

x1	y
-1.0	0.00
1.0	8.70

rep 2

-1.0	4.35
1.0	8.70



$$y = 5.44 + 3.2625 x_1$$

signifikansnivå
 $\alpha = 0.05$

	coef	std err	t	P> t	[0.025	0.975]
Intercept	5.4375	1.087	5.000	0.038	0.758	10.117
x1	3.2625	1.087	3.000	0.095	-1.417	7.942

$\hat{\beta}_0 = 5.44$
 $\hat{\beta}_1 = 3.26$

$H_0: \beta_1 = 0$
 $p = 0.095$

$H_1: \beta_1 \neq 0$
 Kan ikke forkaste H_0 .

Oppsummering av notasjon ++

Y : responsvariabel

x : forklaringsvariabel / kovariat

"numerisk koding av faktornivåer"

$x = -1$ lavt nivå

$x = +1$ høyt nivå

Anta: $Y|x \sim N(\beta_0 + \beta_1 x, \sigma)$

Effekt av faktor "hovedeffekt"

$$2\beta_1$$

dersom t-test fra linear regresjon gir $\beta_1 \neq 0$ (H_1 sann)
så er effekten av faktoren statistisk signifikant.

Oppsummering av notasjon ++

PAUSE TIL

15.10

$$E(Y|x) = \beta_0 + \beta_1 x$$

↑
 $2\beta_1$

$$p = 0.095$$

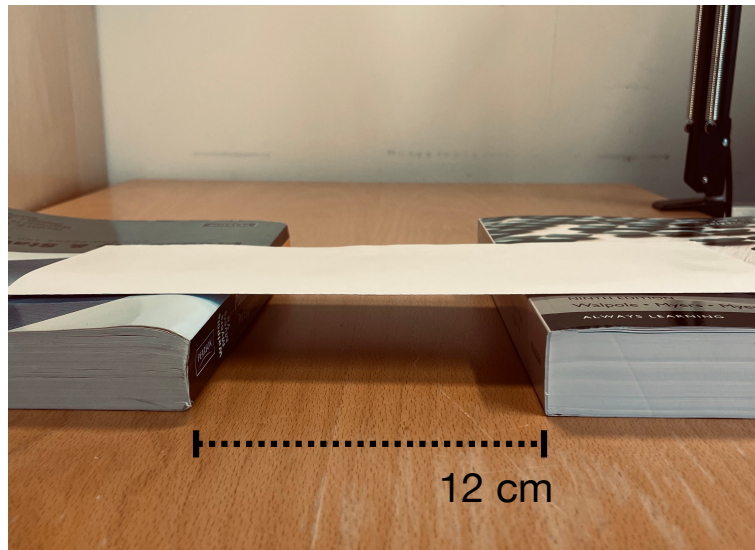
når faktoren endres fra
lavt til høyt nivå så endres

$E(Y)$ med $2\beta_1$

$$\hat{\beta}_1 = 3.26$$

↑

Flere faktorer



A : papirtype

B : avstand

C : brette

$k=3$

2^3 ← 3 faktorer
forsøk
↑

to-niva
faktorielt
forsøk

2 nivåer for hver faktor

A lavt nivå : grøtt ark høyt nivå : hvitt ark

B lavt nivå : 8cm høyt nivå : 12cm

C lavt nivå : 1 Brett høyt nivå : 2 brett

A : $x_1 = -1$, $x_1 = +1$

B : $x_2 = -1$, $x_2 = +1$

C : $x_3 = -1$, $x_3 = +1$

Men Samspill :

Det kan hende at brekkingen (faktor C) er viktigere når avstanden mellom bøkene er lang (faktor B, høyt nivå) enn når avstanden er kort (faktor B, lavt nivå)

Regresjonsmodell med samspill (interaksjoner)

$$E(Y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \beta_{123} x_1 x_2 x_3$$

8 koeffisienter

Designmatrise

2^3 -forsøk

→ måle Y for alle mulige faktor kombinasjoner

$x_1(A)$	$x_2(B)$	$x_3(C)$
-1	-1	-1
+1	-1	-1
-1	+1	-1
+1	+1	-1
-1	-1	+1
+1	-1	+1
-1	+1	+1
+1	+1	+1
$\vec{x}_1$	$\vec{x}_2$	$\vec{x}_3$

8 faktorkombinasjoner

$$2^3 = 8$$

må måle Y for
hver av de 8
faktor kombinasjonene.

NB!

randomisert
rekkefølge

* 8 faktorkomb. $\rightarrow$ måle y i alle 8 tilfeller

* randomisert rekkefølge

* like forutsetninger

$n=8$
observasjoner

8 parameter

Husk 8 koeffisienter (8 "betaer")

$$\begin{aligned} E(y) = & \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \\ & + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 \\ & + \beta_{123} x_1 x_2 x_3 \end{aligned}$$

må gjøre
gjentak

8 + 8
forsøk

Oppgavebeskrivelse

Målet med prosjektet er å planlegge, gjennomføre og diskutere et 2-nivå faktorielt forsøk. Dere står selv fritt til å velge problemstilling og eksperiment. Det kan være et laboratorieeksperiment eller et problem fra dagliglivet. Her er stegene gruppa skal gjennomføre

1. Bestem dere for en problemstilling som dere kan studere med et 2^k faktorielt forsøk. Hva er responsvariabelen og hva er mulige forklaringsvariabler? Hva er egnede nivåer for forklaringsvariablene? Merk at responsvariabelen Y må være en kontinuerlig stokastisk variabel.
2. Design et forsøk med 16, 24 eller 32 målinger. Bruk enten 3 eller 4 faktorer. Dersom dere bruker 3 faktorer, må dere ha minst ett gjentak.
3. Planlegg innsamling av data.
4. Samle inn data.
5. Analyser data.
6. Skriv en rapport.

2^3 , minst ett gjentak

2^4

$\frac{x_1 x_2 x_3}{x_1 x_2 x_4}$ osv.

$\frac{x_1 x_2 x_3 x_4}{x_1 x_2 x_3 x_4}$