

Velkommen til TMA4240

- Hva er statistikk
- Hva skal dere lære i dette kurset
- Praktisk info
- Litt om Deskriptiv statistikk
- Sannsynlighetsregning

Hva er Statistikk??

Hva er Statistikk??

Statistikk er en gren av matematikk som tar for seg analyse av tallfestede dataobservasjoner med den hensikt å kunne trekke konklusjoner basert på disse dataene

Statistikk er en verktøy for å ta beslutnings under usikkerhet og variasjon.

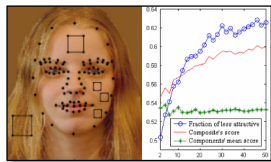
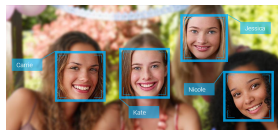
Statistikk handler om å tallfeste vår usikkerhet

Recommender systems



- Basert på hvordan andre har vurdert en product vil amazon vite hvordan *du* ville vurdere en produkt du ikke har vurdert
- I 2006 ga Netflix tilgang til 100480507 vurdering gitt av 480189 brukere til 17770 filmer

Face recognition



- Finn ansiktet i en bilde
- finn ut om ansiktet er for eksempel i dårlig belysning eller i profil
- velge ut unike egenskaper i ansiktet
- sammenlign de unike egenskapene til det ansiktet til alle menneskene du allerede kjenner

30 av 100 ble friske ved bruk av medisin A og 20 av 50 ble friske ved bruk av medisin B. Er medisin A bedre enn medisin B?

30 av 100 ble friske ved bruk av medisin A og 20 av 50 ble friske ved bruk av medisin B. Er medisin A bedre enn medisin B?

- 40% er høyere enn 30%, men er det nok?
- Det kan være svært kostbart å trekke feil konklusjon
- Statistikk svarer på spørsmålet: hvor sikker er vi på at medisin B er bedre enn medisin A?

Statistikk handler om å tallfeste vår usikkerhet

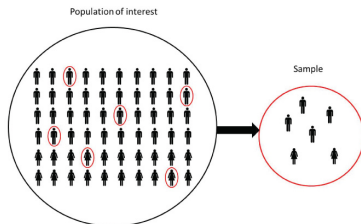
Vi vil vite hvor høyt nordmenn er i gjennomsnit.

- Veldig lett å svare: mål alle og ta gjennomsnit!!!
- ...men det er veldig komplisert å måle alle!!

Statistikk handler om å tallfeste vår usikkerhet

Vi vil vite hvor høyt nordmenn er i gjennomsnit.

- Veldig lett å svare: mål alle og ta gjennomsnit!!!
- ...men det er veldig komplisert å måle alle!!
- Vi må måle noe (et utvalg) og bruke vår resultat til å gi et svar for hele populasjon
- Resultate blir da usikkert
- Det beste vi kan gjør er å tallfeste hvor mye usikkerhet er det rundt vår svar



De finnes mange forskjellige data...

Kontinuerlig data Høyde, vekt, temperatur

Diskret data Antall studenter som tar TMA4240, antall søkere til NTNU, antall øyne på en terning

Kategorisk data Syk/ikke-syk, øyefarge, stryk/bestått

..og mange forskjellige problemer

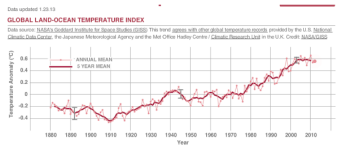
- Hva er årlig tilsig i en vassdrag hvor vi ikke har målinger?
- Er det en trend i temperatur?
- Lever røykere kortere enn ikke-røykere?
- Hva blir været i morra?



Map of the Vansjø-Hobøl Catchment, Norway. Image: Skarbøvik and Bechmann (2010). Bioforsk Report vol. 5

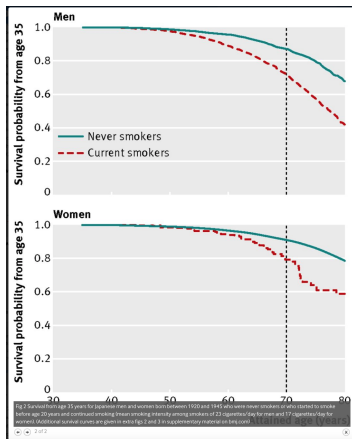
..og mange forskjellige problemer

- Hva er årlig tilsig i en vassdrag hvor vi ikke har målinger?
- Er det en trend i temperatur?
- Lever røykere kortere enn ikke-røykere?
- Hva blir været i morra?



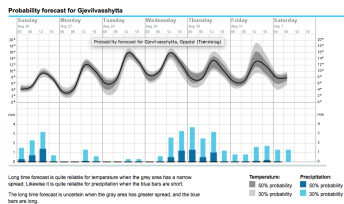
..og mange forskjellige problemer

- Hva er årlig tilsig i en vassdrag hvor vi ikke har målinger?
- Er det en trend i temperatur?
- **Lever røykere kortere enn ikke-røykere?**
- Hva blir været i morra?

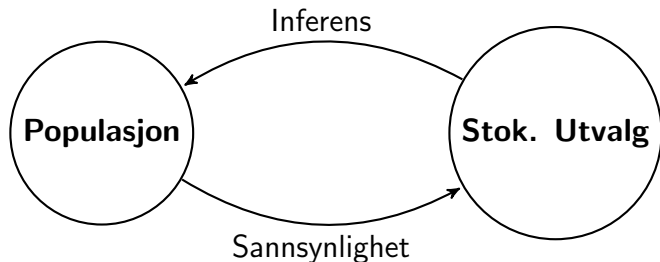


..og mange forskjellige problemer

- Hva er årlig tilsig i en vassdrag hvor vi ikke har målinger?
- Er det en trend i temperatur?
- Lever røykere kortere enn ikke-røykere?
- Hva blir været i morra?



Hva skal vi lære i dette kurset?



- Deskriptiv Statistikk, Kap 1 (Les selv)
 - Presentere og beskrive data i tall, tabeller, figurer
- Sannsynlighet, Kap 2-7 (Fra idag)
 - regne men sannsynlighet
- Statistikk inferens, Kap 8-11 (Senere)
 - trekke slutninger om en hel populasjon fra observasjoner på et utvalg

Sannsynlighetseksempel Du har en rettferdig mynt (like sannsynlighet for hoder eller haler). Hva er sannsynligheten for 60 eller flere hoder hvis du kaster mynt 100 ganger?

Statistikk eksempel Du har en ukjent mynt. Du kaster du det 100 ganger og teller 60 hoder. Er mynten rettferdig?

Sannsynlighetseksempel Du har en rettferdig mynt (like sannsynlighet for hoder eller haler). Hva er sannsynligheten for 60 eller flere hoder hvis du kaster mynt 100 ganger?

Det er bare ett svar (ca. 0.028444), og vi skal lære å beregne det.

Statistikk eksempel Du har en ukjent mynt. Du kaster du det 100 ganger og teller 60 hoder. Er mynten rettferdig?

Sannsynlighetseksempel Du har en rettferdig mynt (like sannsynlighet for hoder eller haler). Hva er sannsynligheten for 60 eller flere hoder hvis du kaster mynt 100 ganger?

Det er bare ett svar (ca. 0.028444), og vi skal lære å beregne det.

Statistikk eksempel Du har en ukjent mynt. Du kaster du det 100 ganger og teller 60 hoder. Er mynten rettferdig? Det er mange måter å gå videre. Faktisk kan forskjellige statistikere trekke forskjellige konklusjoner.

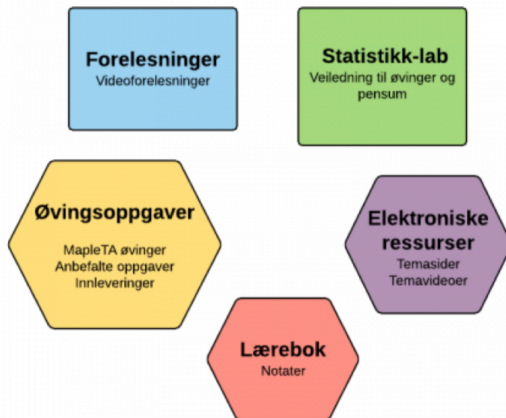
Sannsynlighetseksempel Tilfeldige prosessen fullt kjent (sannsynlighet for hodene = .5).

Målet er å finne sannsynligheten for et bestemt utfall (minst 60 hoder).

Statistikk eksempel Resultatet kjent (60 hoder) og målet er å belyse den ukjente tilfeldige prosessen (sannsynligheten for hoder).

TMA4240: Tilgjengelig resurser

- Faglærer: Sara Martino og Håkon Tjelmeland
- Hjemmeside:
<https://wiki.math.ntnu.no/tma4240/2019h/start>



- To Parallell:
 - Mandag-Tirsdag 8:15-10:00
 - Mandag-Tirsdag 12:15-14:00 (video opptak)
- Du kan velge hvilken forelesning du ønsker
- Begge parallel har samme pensum og de er delt ”verticalt” mellom to foreleser
 - 4 uker: Sara Martino
 - 4 uker: Håkon Tjelmeland
 - 3 uker: Sara Martino
 - 3 uker: Håkon Tjelmeland

Elektronisk
øving

Anbefalt
øving

Skriftlig
øving

Statistikklab

Krav til godkjent øvingsopplegg

- Elektronisk (Maple TA) øving
 - **Minst 4** godkjente øvinger fra Blokk 1
 - **Minst 4** godkjente øvinger fra Blokk 2
- Skriftlige innleveringer
 - **3 godkjente innleveringer**, hvorav **minst 1** må være fra Blokk 1 og **minst 1** fra Blokk 2
- Anbefalte øvinger
- Anvarling: Ingeborg Gullikstad Hem (ingeborg.hem@ntnu.no)

Her kan du samarbeide med andre og få hjelp av studentassistenter

Dag	Tid	Rom
Mandag	14:15-18:00	S5
Torsdag	12:15-16:00	A2
Fredag	8:15-12:00	S5

- **Walpole, Myers, Myers and Ye (2012): Probability and Statistics for engineers (9. utgave), Pearson / Prentice Hall.**
 - Det finnes en egen utgave for NTNU (ISBN 978-1-78726-081-8) som selges på Akademika.
 - Det finnes en NTNU-versjon av 9. utgave med ISBN 978-1-78448-014-1 som også kan brukes. Originalversjonen (internasjonal versjon) av 9. utgave kan også brukes.
 - Det finnes en utgave med **blå skrift** (ISBN 10:1-292-02392-9. ISBN 13: 978-1-292-02392-2), som **vi fraråder** da flere har gitt tilbakemelding om at den er vanskelig å lese og å finne frem i.
- **Tabeller of formler i statistikk. Akademika**

- Temasider
- Temavideo

- Vi trenger 2/3 studenter som kan være i referansegruppe
 - 2/3 møter i løpet av semester
 - skrive en kort rapport til slutt

Descriptive Statistikk (Kap. 1)

- Anta at man ha samlet et datamateriale : x_1, x_2, \dots, x_n
- Hvordan forstå dette datamateriale?
 - Hvis $n = 10000$ er det ikke så informativ bare å lese alle tallene!!
- Hvordan visualizere datane?
 - histogram
 - det finnes flere grafiske plottemetoder
- Hvis man skal gi et (eller 2,3) tall som beskriver datamaterialet, hva skal du velge?
 - gjennomsnitt
 - median
 - empirisk varians
- Skal ikke undervise dette kapittel. Les selv med spesielt fokus på:
 - mål for lokasjon (eng: location)
 - mål for variasjon (eng: variability)
 - histogram
 - boxplott
 - scatterplott