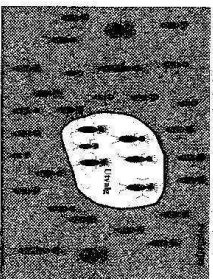


Kapittel 1: Introduksjon til statistikk og dataanalyse

TMA4240 Statistikk (F2 og E7)

Foreleses mandag 16.august 2004



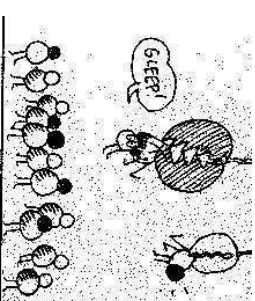
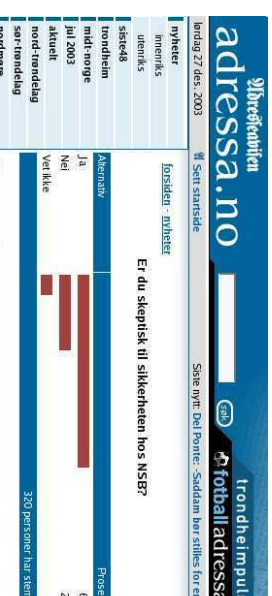
Olav Petter Løden@math.ntnu.no – p.1/16

Matematikk 1: spørreundersøkelse

- Mål med undersøkelsen: hva påvirker karakter i Matematikk 1?
- Våren 2003: 571 studenter (av 1500) svarte på 40 spørsmål (6 kategorier).
 - Når Matematikk 1: 475 studenter i H2002.
 - 17 linjer; deriblant E5 (50), O2 (5), O3 (37), K3 (11), SDK (11), F2 (59), E7 (42).
- Er dette et tilfeldig utvalg?
- ”ethvert utvalg av størrelse n har lik sjanse for å bli trukket ut (svare på undersøkelsen)”.
 - Representativitet: kan vi trekke konklusjoner for alle som har tatt Matematikk 1?

TMA4240 (F2 og E7): Kapittel 1 – p.2/16

Er dette et tilfeldig utvalg?



Figur fra Cartoon Guide to Statistics – p.3/16

Spørreundersøkelse for Matematikk 1.

Våren 2003

1: Innledning

1. Kjønn
Mannen
Kvinnen
2. Hvilken linje gikk du på høsten 2001
A. E5, E6, E7, F1, F2, G, K1, K3, N, NAN, O2, O3, SDK, SEM, SPP
B. E5, E6, E7, F1, F2, G, K1, K3, N, NAN, O2, O3, SDK, SEM, SPP
3. Når begynte du studier ved NTNU siv.-ing?
Høsten 1998, Høsten 1999, Høsten 2000, Høsten 2001, Høsten 2002
4. Studenterkarakter i 3MX33AN
6, 5, 4, 3, 2, annen bakgrunn
5. Hvilket år er karakteren i 3MX33AN fra?
2002, 2001, 2000, 1996 – 1999, Før 1996, Annen bakgrunn
6. Har du tatt 3FY? (Ja/Nei)
7. Studenterkarakter i ditt avsluttende fysikkfag
6, 5, 4, 3, 2
8. Hvilken karakter fikk du til eksamen i Matematikk 1?
A, B, C, D, E, F, Ikke avlagt eksamen
9. I hvilken grad mener du den karakteren du fikk var rettferdig i forhold til prestasjonen din til eksamen i Matematikk 1.
For dårlig (1) For god (9) Ingen karakter (10)
10. I hvilken grad mener du karakteren din i Matematikk 1 gir et riktig bilde av den kompetensen du selv mener å ha i faget?
For dårlig (1) For god (9) Ingen karakter (10)
11. I hvilken grad er du enig i følgende utsagn: Jeg var bare interessert i å nå, og lyvde meg ikke om hvilken skår karakter jeg eventuelt fikk i Matematikk 1.
Uenig (1) Enig (9)
12. Delok du på opptakingskurset i august? Ja/Nei
13. Delok du på midsemesterprøven? Ja/Nei
14. Har du (bealt eller ikke bealt) jobb ved siden av studiet i semesteret? Antall i så fall: antall timer per uke i gjennomsnitt som går med til arbeid utenom studiene.
0 – 5 timer

TMA4240 (F2 og E7): Kapittel 1 – p.4/16

Diskrete data

- Diskrete data: endelig (telbart) antall mulige utfall.
- Standpunktkarakterer i 3MX/3MN:
 - 6, 5, 4, 3, 2, ”annen bakgrunn”
- Karakter til eksamen i Matematikk 1:
 - A, B, C, D, E, F, ”ikke avlagt eksamen”
- I hvilken grad mener du karakteren din i Matematikk 1 *gir et riktig bilde* av den kompetansen du selv mener å ha i faget:
 - For dårlig (1), ..., For god(9), ”Ingen karakter”(10)
- Jeg klarte å holde meg faglig ajour gjennom semesteret: Aldri (1), ..., Alltid(9)

TMA4240 (F2 og E7): Kapittel 1 – p.5/16

Relativ frekvens

- Karakter til eksamen i Matematikk 1

Intervall	Midtpunkt	Frekvens	Relativ frekvens
86-100	A	35	0.06
76-85	B	61	0.11
66-75	C	90	0.16
56-65	D	74	0.13
36-55	E	144	0.26
0-35	F	156	0.28
0-100	A-F	560	1.00

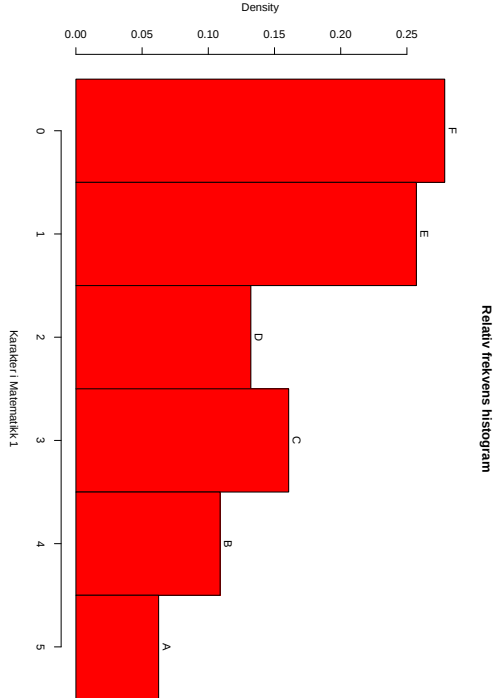
TMA4240 (F2 og E7): Kapittel 1 – p.7/16

Kontinuerlige data

- Kontinuerlige data: uendelig antall mulige utfall.
- Ingen kontinuerlige data i denne undersøkelsen.
- Spørsmål om timeforbruk og prosentandel tilstede på ulike aktiviteter kunne vært angitt som kontinuerlige tall.
- Målinger (mengde, volum, andel) ofte kontinuerlig.
- Spørsmål om ”høyde” i [www-spørreundersøkelsen](#) i TMA4240.
- Sannynlighetsteori ulik for diskrete og kontinuerlige data – spesielt kapittel 3, 5, 6.

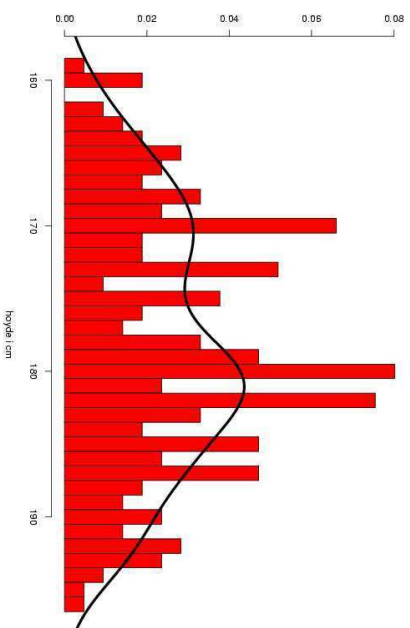
TMA4240 (F2 og E7): Kapittel 1 – p.6/16

Relativ frekvens: karakter



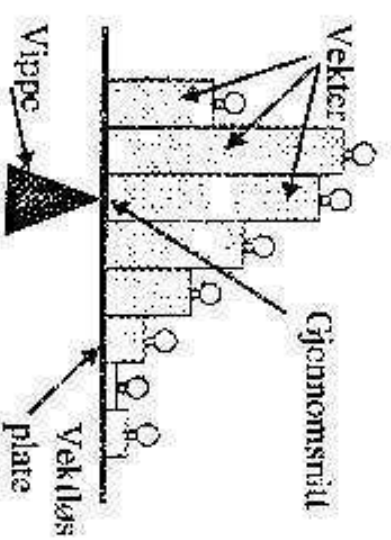
TMA4240 (F2 og E7): Kapittel 1 – p.8/16

Relativ frekvens: høyde



TMA4240 (F2 og E7): Kapittel 1 – p.9/16

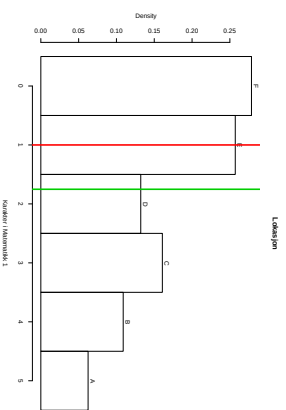
Gjennomsnittet



Figur fra G. Løvdal: "Statistikk" – p.11/16

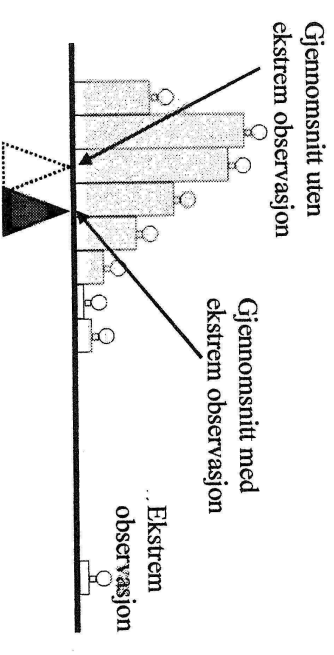
Mål for beliggenhet

- Gjennomsnitt (mean): $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$
- Median
- n odde $\tilde{x} = x_{(n+1)/2}$
- n like $\tilde{x} = \frac{1}{2}(x_{n/2} + x_{n/2+1})$



TMA4240 (F2 og E7): Kapittel 1 – p.10/16

Gjennomsnittet



Figur fra G. Løvdal: "Statistikk" – p.12/16

Mål for variabilitet

- Empirisk varians (DEF 1.1):
 - $s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$
- Gjennomsnittlig kvadratavvik til gjennomsnittet.
- Empirisk standardavvik (DEF 1.1):
 - $s = \sqrt{s^2}$
- Samme måleenhet som originalobservasjonene.
- Inter-kvaril rekkevidde(IQR): Q3-Q1
 - Q1=verdien som 25% av data er mindre enn
 - Q3=verdien som 25% av data er større enn

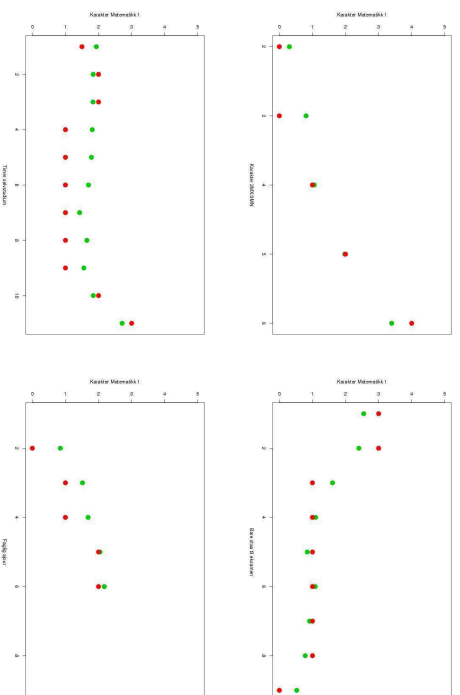
TMA4240 (F2 og E7: Kapittel 1 – p.13/16

Oppsummering Matematikk 1

- Generelt: god karakter i 3MX og 3FY “gir” god karakter i Matematikk 1.
- Linjevisе forskjeller.
- Registrering av kvantitet av tidsforbruk ikke nok – må også registrere kvalitet?
- Hva mener studentene de lærer mest av?
 - Forberedelse til eksamen
 - Samarbeid med medstudenter
 - Regne øvinger selv!
- Egen forklaring på F:
 - Jobba for lite
 - Jobba på feil måte

TMA4240 (F2 og E7: Kapittel 1 – p.15/16

Sammenheng



TMA4240 (F2 og E7: Kapittel 1 – p.14/16

TMA4240 Statistikk

Beskrivende statistikk

- Kapittel 1
- Presentere og beskrive data i tall, tabeller og figurer

Sannsynlighetsteori

- Kapittel 2-7
- Matematisk regning med sannsynlighet og sannsynlighetsfordelinger

Statistisk inferens

- Kapittel 8-11
- Estimere, teste, prediktere, basert på data under usikkerhet.

TMA4240 (F2 og E7: Kapittel 1 – p.16/16