Innføring i statistikk og dataanalyse for studenter i idretts- og helsefag

Jon Ingulf Medbø

Innføring i statistikk og dataanalyse for studenter i idretts- og helsefag

CAPPELEN DAMM AKADEMISK

© CAPPELEN DAMM AS, Oslo, 2021

ISBN 978-82-02-73011-6
1. utgave, 1. versjon 2021

E-utgaven baserer seg på 1. trykte utgave, 2. opplag,
978-82-02-51688-8, Cappelen Damm Akademisk, 2018

Illustrasjoner: Røy Søbstad
Omslagsdesign: Affair.no
Sats: Bøk Oslo AS
Elektronisk tilrettelegging: Bøk Oslo AS

Boka er utgitt med støtte fra Lærebokutvalget for høyere utdanning.

Materialet i denne publikasjonen er omfattet av åndsverklovens bestemmelser. Uten særskilt avtale med forlaget er enhver eksemplarfremstilling og tilgjengeliggjøring bare tillatt i den utstrekning det er hjemlet i lov eller tillatt gjennom avtale med Kopinor, interesseorgan for rettighetshavere til åndsverk.

www.cda.no
akademisk@cappelendamm.no

Cappelen Damm AS er sertifisert iht. FSC®, PEFC og Miljøfyrtårn.

Innhold

Forord 15

Kapittel 1
Innledning – hva er statistikk? 17
Innledende eksempler 17
Hva er statistikk? 20
«Lov og tilfeldighet» 21
«Do you speak statistics?» 21
Hvordan lærer en statistikk? 22
Statistikk er et middel i data-analyse 23
Misbruk av statistikk 23
Oversikt over bokas innhold og oppbygninga av kapitlene 25
Sammenfatning av data – deskriptiv statistikk 27
Praktiske eksempler og oppgaver 27
Oppgaver og kontrollspørsmål 28
Merknader 29

Kapittel 2
Innføring i kvantitative metoder 32
Hva er kvantitative metoder? 32
Forskningsprosessen 32
Første trinn: problemstilling, hypotese og utvalg 33
Andre trinn: valg av framgangsmåte eller metode for datainnsamlinga 35
Tredje trinn: innsamling av resultater, bearbeiding og framstilling av dem 42
Fjerde trinn: tolking av resultatene, å sette dem inn i en større sammenheng 42
Klassifisering av data – typer måleskalaer 43
Nominelle skalaer, kategoriske variabler 43
Ordinalskalaer, rangskalaer 44

Intervallskalaer 44
Forholdsskalaer, ratioskalaer 45
Sammenfatning om måleskalaer 46
Diskrete og kontinuerlige målinger 46
Mål på lokalisering og spredning 49
Sentralmål – mål på lokalisering 49
Spredningsmål 49
Standardiserte størrelser 51
Normalskårer, *z*-skårer 51
Stanineskårer 51
Handtering av tall 52
Tall med siffer eller bokstaver? 53
Måltall 54
Skriving av tall 54
Brøk, prosent og desimaltall 56
Bruk og misbruk av prosenter 57
Prosenter og prosentpoeng 57
Bruk av forkortinger i faglige sammenhenger 58
Handtering av (fysiske) størrelser med målenheter 60
Skriving av statistiske størrelser 63
Tolking av intervaller. Overganger mellom diskrete og kontinuerlige skalaer 63
Diskrete og kontinuerlige størrelser og skalaer 63
Lengden av diskrete intervall 64
Halvkorreksjon 64
Hvor mange siffer i svaret? Om å unngå desimalsjuken 65
Regler for avrunding av tall 65
Desimalsjuken 66
Størrelsesorden 69
Bruk av måleskalaer 70
Null er et viktig tall 70
Titallsystemet 72
VAS-skalaer 73
Likertskalaer 73
Gagges subjektive temperaturskala 74
Borg-skalaer 75
Klassifisering av kurver 78
Oppgaver og kontrollspørsmål 80
Vedlegg 81
Mer om mål på lokalisering 81
Mer om spredningsmål 84

Lokaliseringsmål, spredningsmål og måleskalanivå 88
Skeivhet 88
Kurtose 90

Kapittel 3
Innføring i tilfeldige prosesser og sannsynlighetsfordelinger 94
Tilfeldige (stokastiske) prosesser 94
Den diskrete uniforme fordelinga 94
Normalfordelinga 96
Kumulative fordelinger 98
Eksempel på normalfordelte størrelser: kroppshøyden hos norske menn 99
Eksempel med tilfeldige fordelinger 101
Utfallsrom – definisjon 104
Oppgaver og kontrollspørsmål 105

Kapittel 4
Innføring i sannsynlighetsregning 106
Bakgrunn 106
Gunstige og mulige utfall 107
Kombinatorikk og permutasjoner 109
Utvelging med eller uten tilbakelegging 109
Ordna og ikke-ordna utvalg 109
Ikke-ordna utvalg uten tilbakelegging 111
Avsluttende merknader om kombinatorikk og sannsynligheter 115
von Mises' og Kolmogorovs sannsynlighetsbegrep 115
von Mises' sannsynlighetsbegrep 115
Kolmogorovs sannsynlighetsbegrep 117
Subjektiv sannsynlighet 119
Punktsannsynligheter og sannsynlighetstettheter 120
Punktsannsynligheter 121
Kontinuerlige sannsynlighetsfunksjoner. Sannsynlighetstettheter 121
Kumulative fordelinger 123
Kontinuerlige og diskrete fordelinger – hvor går skillet? 127
Betinga sannsynligheter 128
Marginale sannsynlighetsfordelinger 131
Uavhengighet 132
Statistiske fordelinger 133
Forventning og varians 134
Viktige diskrete statistiske fordelinger 137
Den diskrete uniforme fordelinga 137

Binomialfordelinga ... 137
Poissonfordelinga ... 140
Viktige kontinuerlige statistiske fordelinger ... 143
Den uniforme fordelinga (rektangulærfordelinga) ... 143
Normalfordelinga ... 144
Gossets («Students») *t*-fordeling ... 145
Sentralgrenseteoremet ... 149
Oppgaver og kontrollspørsmål ... 155
Vedlegg ... 156
Viktige regler utleda fra Kolmogorovs aksiom ... 156
Formler for betinga sannsynligheter. Bayes teorem ... 158
Stirlings formel i enkel form ... 158
Den multinomiske fordelinga ... 160
Den hypergeometriske fordelinga ... 161
Kjikvadratfordelinga ... 162
Den eksponensielle fordelinga ... 163
Fisher-fordelinga ... 164
Student-Fishers setning ... 165

Kapittel 5
Estimering og hypotesetesting ... 166
Innledning ... 166
Estimering ... 168
Estimering av forventning og varians for (tilnærma) normalfordelte størrelser ... 168
Medianen ... 170
Hodges-Lehman-Sen-estimatoren (HLS-estimatoren) ... 171
Estimering av den binomiske suksess-sannsynligheten ... 173
Avsluttende merknader om estimering ... 175
Hypotesetesting ... 175
Formelle prinsipp for hypotesetesting ... 176
Hypotesetesting i praksis ... 181
Teststyrke – styrkefunksjonen ... 181
Avsluttende merknader om hypotesetesting ... 183
Konfidensintervall ... 184
Teststyrke og konfidensintervall ... 188
Konfidensintervall og hypotesetesting ... 191
Suffisiens ... 191
Suffisiens og sentralgrenseteoremet ... 192
Oppgaver og kontrollspørsmål ... 193

Vedlegg 194
Konfidenskurver 194

Kapittel 6
Toutvalgstester 196
Bakgrunn med et eksempel 196
Gossets (Students) toutvalgs ***t***-test 199
Estimering 200
Hypotesetesting 202
Forutsetninger for Gossets toutvalgs ***t***-test 204
Wilcoxon og Mann-Whitneys ikke-parametriske toutvalgstest 204
Wilcoxons tilnærming 205
Mann-Whitneys tilnærming 205
Sammenfallende verdier 208
Estimering 210
Forutsetninger for Wilcoxon Mann-Whitneys toutvalgs test. Hva måler egentlig testen? 211
Regnetekniske krav til Wilcoxon Mann-Whitney-testen 212
Behrens-Fishers problem. Heteroskedastiske data 213
Welchs tilnærming 215
Estimering 216
Avsluttende merknad om antall frihetsgrader 217
Råd for valg av test og testopplegg – effisiens 217
Effisiens 218
Tester på lik varians 219
Fishers F-test 219
Sammenfatning 221
Oppgaver og kontrollspørsmål 222
Vedlegg. Andre foreslåtte tester 226
Kolmogorov-Smirnovs test 227
Wald-Wolfowitz' følgetest 229
Siegel-Tukeys varianstest 230
Mer om Behrens-Fishers problem og Welchs tilnærming 232

Kapittel 7
Ettutvalgstester, para tester 234
Bakgrunn 234
Et innledende eksempel 235
Annen bruk av partester – grunnprinsippet bak testene 237
Gossets («Students») ettutvalgs ***t***-test 237

Estimering ... 239
Enkel hypotesetesting ... 239
Forutsetninger for Gossets ettutvalgs *t*-test ... 240
Wilcoxons ettutvalgs rangtest ... 241
Forutsetninger for Wilcoxons ettutvalgstest ... 242
Hva måler egentlig Wilcoxons ettutvalgstest? ... 242
Estimering ... 243
Sammenlikning av testene ... 243
Toutvalgstest på para data ... 244
Toutvalgs *t*-test ... 245
Wilcoxon Mann-Whitney-test ... 245
*Kolmogorov-Smirnov-test ... 246
*Wald-Wolfowitz' følgetest ... 246
Sammenfatning om bruk av toutvalgstester på para data ... 246
Sammenlikning av endringer for to utvalg ... 247
Vanlige feil ... 247
Oppgaver og kontrollspørsmål ... 248
Vedlegg. Fortegnstesten ... 249
Utregninger ... 249
Matematiske sider ved partester ... 250

Kapittel 8
Sammenlikning av mer enn to grupper. Innføring i variansanalyse ... 252
Innledning ... 252
Faktorer og nivåer ... 257
Forventningskomponenter og varianskomponenter. Blanda modeller ... 257
Variansanalysens grunnprinsipp ... 259
Variansanalysens to trinn ... 260
Sammenfatning ... 260
Enveis variansanalyse ... 261
Trinn to i en variansanalyse ... 262
Scheffés metode ... 262
Fishers minste signifikante skilnad (LSD) ... 263
Studentiserte variasjonsbredder ... 265
Dunnetts test ... 266
Pre hoc-analyser ... 266
Mer om variansanalyse ... 266
Randomiserte blokker ... 266
Varianskomponentmodeller ... 267
Toveis og flerveis variansanalyse ... 270

Avsluttende merknader om variansanalyse 275
Variansanalyse er en omnibustest 276
Oppgaver og kontrollspørsmål 278
Vedlegg 280
Kruskal-Wallis' test. En ikke-parametrisk enveis variansanalyse 280
Friedmans test – en ikke-parametrisk test for blokkdelte forsøk 285

Kapittel 9
Sammenhenger mellom to eller flere variabler 286
Regresjon 286
Ei presisering 289
Viktige egenskaper ved en regresjonsmodell 291
Viktige parametere i lineærregresjon 292
Estimering og statistisk testing i lineærregresjon 294
Lineærregresjon når *X* blir *Y* og *Y* blir *X* 295
Geometrisk middel 297
Interpolering. Ekstrapolering 298
Kalibrering 299
Bland-Altmans metode og B-A-plottet 302
Multippel regresjon 305
Ikke-lineær regresjon 306
Ikke-parametriske metoder. Medianregresjon og kvantilregresjon 309
Kovariansanalyse 310
Korrelasjon 312
Pearsons produktmoment koeffisient 314
Spearmans rangkorrelasjonskoeffisient 317
Kendalls korrelasjonskoeffisient 318
Multippel korrelasjon 319
Partiell korrelasjon 320
Avsluttende merknader om korrelasjon 321
Oppgaver og kontrollspørsmål 321

Kapittel 10
Binomiske forsøk. Poissonprosesser 324
Forsøk med binomisk utfall 324
Usikkerhet for ekstreme binomiske utfall. Treerregelen 327
Sammenlikning av punktsannsynligheter for to binomiske forsøksrekker 328
Utfall av poissonprosesser 333
Enkle poissonprosesser 333
Sammenlikning av to poissonprosesser 337

Logistisk regresjon 338
Oppgaver og kontrollspørsmål 340
Vedlegg 340
Mer nøyaktig formel for konfidensintervall for binomisk punktsannsynlighet 340
Utledning av treerregelen 341
Mer nøyaktig formel for konfidensintervall for intensiteten til poissonske punktprosesser 343

Kapittel 11
Multinomiske forsøksrekker. Analyse av kategoriske data 344
Bakgrunn 345
Tre ulike modeller 347
Kjikvadrattilnærminga 348
Yates' korreksjon for kontinuitet 351
Irwin-Fishers eksakte test for 2×2-tabeller 352
Konfidensintervall for 2×2-modeller 354
Nyere tester for 2×2-modeller 356
Para sammenlikninger: McNemars test 357
Estimering med konfidensintervall 359
Nyere forskning på para sammenlikninger 360
Analyse av utfallet av multinomiske forsøk mot fastsatte fordelinger 360
Analyse av tabeller med rangerte verdier 362
Enveis rangstruktur 363
Toveis rangstruktur 364
Hva taper en ved å kategorisere data? 368
Videre lesing 369
Oppgaver og kontrollspørsmål 370
Vedlegg 371
Matematikken bak analyser av multinomiske forsøksrekker med kjikvadrattesten 371

Kapittel 12
Data-analyse. Avsluttende merknader 375
Følger av variasjoner i tyngdekrafta 375
Metningstrykket for vanndamptrykket avhenger av temperaturen 377
Skøyteløp 378
Holde svingen 378
Kutte svingen 379
Tidsforsinking på grunn av lydfarten 380

Varmere klima – følger for skiføret 380
Norsk toppfotball gjennom femti år 381
Analyse av en sammensatt idrett – skiskyting 384
Presentasjon av resultat 388
Standardavvik eller standardfeil. Konfidensintervall og yttergrenser for samsvar 388
Bruk og misbruk av P-verdier 391
P-fisking 391
Publiseringsskeivheit 392
Hypotesetesting eller estimering? P-verdier eller konfidensintervall? 392
Matematiske snarveier, fallgruver og vrakgods 393
«Do you speak statistics?» 396
Statistiske moteretninger 397
Statistikkens tre? 398
Data-analyse uten statistikk 399

Vedlegg 400
Videregående matematikk 400
Integrering 400
Logaritmer og eksponentialfunksjoner 406

Stikkordregister 409

Forord

Innføring i statistikk og dataanalyse for studenter i idretts- og helsefag er ei brukerretta lærebok for studenter som jobber med praktisk statistikk. Tidligere måtte krevende regnearbeid av denne typen gjøres for hand, og å lære seg å bruke de rette formlene sto derfor sentralt i opplæringa i praktisk statistikk. I dag har vi dataprogram som utfører regnearbeidet for oss. Da er det viktigere at brukeren veit hvilket verktøy som skal brukes, enn å kjenne til de regnetekniske detaljene.

Boka er skrevet som ei vanlig lærebok med tekst, figurer og tabeller, og etter hvert kapittel gis det oppgaver (noen kapitler har også oppgaver underveis). Det er utarbeida et hefte med løsningsforslag til alle oppgavene. Dette heftet gir også en kort innføring i bruk av regnearket Excel og statistikkprogrammet SPSS. Studentene kan også søke opp hovedbokens ISBN 9788202516888 på www.cappelendammundervisning.no, gå til fanen Tilleggsmateriell og laste ned inngangsdata og andre talldata tilhørende bokens oppgaver, dette for å slippe tidkrevende inntastinger.

Regelmessig fysisk aktivitet er viktig for helsa, og her kommer tilknytninga mellom idrettsfag og helsefag til syne. Med den nye femårige mastergraden ved lærerutdanninga bør boka også være aktuell i masterarbeidet for mange av disse studentene.

Jeg takker Asgeir Mamen og Boye Welde som har kommet med nyttige innspill og gode kritiske tilbakemeldinger underveis i arbeidet. Jeg retter en like stor takk til forlagets anonyme fagkonsulenter for de mange gode forslaga og innspillene de har kommet med. Det har resultert i ei bedre bok. Takk til Nils Lid Hjort, Terje Myklebust og Futoshi Ogita for data og innspill. Jeg takker videre redaktør Bjørn Olav Aas Hansen hos Cappelen Damm Akademisk for svært godt samarbeid og Marius Hjeldnes for godt redaksjonelt arbeid. En stor takk også til språkvasker Siv Rekve for tålmodig arbeid og for tallrike gode innspill som har forbedra den språklige framstillinga.

Jon Ingulf Medbø,
Sogndal, oktober 2017

KAPITTEL 1

Innledning – hva er statistikk?

Innhold og læringsmål

Målet med dette kapitlet er å gi noe av grunntanken bak statistiske analyser av data. Det sier også noe om hvordan en bør gå fram for å lære å bruke og tolke statistiske data. Leseren skal også få litt kjennskap til mulig misbruk av statistikk. Kapitlet gir til slutt en oversikt over innholdet i resten av boka.

Innledende eksempler

Statistikk blir stadig mer brukt, ikke minst i forskning og når en tolker data i forskningsliknende oppgaver. Det gjelder også for idretts- og helsefag. Noen eksempler kan vise det:

- Ei gruppe amerikanske forskere under ledelse av Jonas E. Salk utvikla i 1950-åra en vaksine mot sjukdommen poliomyelitt som ubehandla fører til muskellammelser og i noen tilfeller død. Denne sjukdommen ramma mange, blant annet den amerikanske presidenten Franklin D. Roosevelt (president 1933–1945) som var lamma i underkroppen etter poliosjukdom i 1921.[1]

  For å undersøke om vaksinen virkelig hadde noen virkning, blei nesten én million barn vaksinert i 1954, mens mer enn én million barn ikke blei vaksinert og tjente som kontroller. Så fulgte Salk og medarbeidere alle barna og registrerte forekomster av poliomyelitt i hver av de to gruppene de følgende månedene. De fant langt færre tilfeller av poliomyelitt blant de vak-

1 https://en.wikipedia.org/wiki/Franklin_D._Roosevelt#Paralytic_illness_.281921.29. Lest 26.9.2017.

sinerte barna sammenlikna med i den ubehandla gruppa. Forskerne viste i 1955 at skilnaden var så stor at den ikke kunne forklares med statistiske tilfeldigheter. Det er slikt arbeid som har ført til at sjukdommen poliomyelitt for lengst er så godt som utrydda i industrialiserte land.[2]

- Redaksjonen i det medisinske tidsskriftet *New England Journal of Medicine* valgte i januar 2000 å starte det nye tusenåret med å sammenfatte den medisinske utviklinga i det tusenåret som nettopp var avslutta.[3] Den pekte ut elleve hovedemner som den mente hadde vært særlig viktige for utviklinga. Det at legevitenskapen bygger på kunnskap i anatomi og fysiologi, på kunnskap om celler og oppbygninga av dem, og på de biokjemiske prosessene i dem, er åpenbart. Det samme gjelder at kunnskap om bakterier og det å kunne bekjempe dem, enten gjennom kroppens eget immunsystem eller ved hjelp av antibiotika, er viktig for å unngå infeksjonssjukdommer. Ett av de elleve fagfeltene som blei framheva, var statistisk bearbeiding av innsamla materiale. Bruk av statistikk i medisinsk sammenheng blei altså framheva på linje med bruk av kunnskaper i anatomi og fysiologi, kunnskap om celler og bakterier, om anestesi, farmakologi og medisiner og så videre.
- Hurtigløp på skøyter er en tradisjonell norsk idrett. I mesterskap på skøytesprint går en i dag to 500-meterløp, den ene gangen med start i indre bane, og den andre gangen med start i ytre bane. Det har lenge vært diskutert om det er en fordel med «første indre». Nils L. Hjort ved Universitetet i Oslo undersøkte spørsmålet statistisk ved å gå gjennom resultatlister fra ti verdensmesterskap i skøytesprint. Ut fra de beregningene viste han at en gjennomsnittlig sprinter tjente 0,05 s på å starte i første indre sammenlikna med start i første ytre. Videre viste han at selv om det var variasjoner mellom løpere og fra stevne til stevne, var det en gjennomgående skilnad: Den var statistisk utsagnskraftig («signifikant») og kunne ikke bare forklares med tilfeldigheter.[4]
- Ordninga med én førstedivisjon/toppdivisjon i norsk fotball (nå tippeligaen) blei oppretta forut for sesongen 1963. Etter at sesongen 2016 er fullført, er det i løpet av 54 sesonger[5] gjennomført 8314 kamper der det er registrert 3935 hjemmeseiere (47,3) %, 2007 uavgjorte kamper (24,1 %) og

2 https://en.wikipedia.org/wiki/Jonas_Salk. Lest 26.9.2017.

3 Editorial. *New England J. Med*. 2000; 342(1): 42–49.

4 http://www.uio.no/studier/emner/matnat/math/STK4160/v11/olympic_partone.pdf, http://www.uio.no/studier/emner/matnat/math/STK4160/v11/olympic_parttwo.pdf

5 Regelverket for sesongen 1987 var avvikende med straffesparkkonkurranse etter en uavgjort kamp ved full tid. Det er i statistikken her brukt resultatet ved full tid, uten å ta med resultatet av straffesparkkonkurransen.

2372 (28,5 %) borteseiere.[6] Denne statistikken for norske toppseriekamper gjennom mer enn femti år sett under ett viser at nesten annenhver kamp ender med hjemmeseier, i underkant av hver fjerde kamp ender uavgjort, mens det blir borteseier i opp mot 30 % av alle kampene. Disse talla understreker blant annet den store fordelen det er med kamp på hjemmebane.

- Ifølge en artikkel på Wikipedia er antall mål i britiske fotballkamper poissonfordelt med intensitetsparameter 2,5 mål per kamp.[7] Hva har det å si for fordelinga av antall mål per kamp? Hvor sannsynlig er det med en målløs kamp? Hvor sannsynlig er det med mer enn fire mål i en kamp? Slike spørsmål kan en belyse med statistiske metoder.
- I norsk toppfotball for menn er det fra og med 1963 til og med 2016 skåra til sammen 25 093 mål på 8314 kamper.[8] Det gir en midlere skåringsintensitet på 3,02 mål/kamp. Hva slags videre statistiske vurderinger kan en bruke slike tall til?
- Mange land, også Norge, har egne kreftregistre. Her blir alle nye forekomster av kreft registrert sammen med andre opplysninger om pasientene. Så kan en seinere hente fram opplysninger og undersøke hva slags systematiske skilnader det er mellom de som har fått en kreftsjukdom, og de som ikke er ramma. Det var slike tilnærminger som gjorde at en i 1950 kunne slå fast at røyking var en svært viktig årsak til lungekreft.[9] Allerede før det hadde en klare, statistiske holdepunkter for at det var en klar sammenheng mellom røyking og lungekreft. Seinere har en påvist at røyking også er en viktig årsak til mange andre former for kreft og videre til hjerte-karlidelser og alvorlige luftveislidelser. Det er videre vist at mer enn annenhver vanerøyker (en som røyker daglig over lengre tid) dør for tidlig og av følger av røykinga. Disse slutningene bygger i første rekke på statistisk analyse av livslengde og dødsårsaker.
- Idretten har mange uløste spørsmål av statistisk art. En del idretter består av sammensatte øvelser, der kombinert skirenn kan være et eksempel. Det består av en hoppkonkurranse og et langrenn. Denne idrettens idé er at begge øvelsene skal være (omtrent) like utslagsgivende. Det skal ikke være slik at bare en er god til å hoppe, er det ikke så farlig med egenskapene som langrennsløper, eller omvendt. Skihopping består igjen av det å hoppe

---

6 Resultatene er tatt fra nettsida Alt om fotball: http://www.altomfotball.no/element.do?useFullUrl=false høsten 2016 og med løyve fra TV 2.

7 https://en.wikipedia.org/wiki/Poisson_distribution. Lest 26.9.2017.

8 Mål i straffesparkkonkurranser som avslutta og avgjorde uavgjorte kamper i 1987-sesongen er ikke tatt med.

9 https://en.wikipedia.org/wiki/Health_effects_of_tobacco. Lest 26.9.2017.