



NJÅL FOLDNES | STEFFEN GRØNNEBERG | GUDMUND HORN HERMANSEN

STATISTIKK OG DATAANALYSE

En moderne innføring

**Njål Foldnes,
Steffen Grønneberg og
Gudmund Horn Hermansen**

Statistikk og dataanalyse

En moderne innføring

Kapitteloversikt

DEL 1 INTRODUKSJON TIL STATISTIKK

- Kapittel 1 Populasjon og utvalg 19
- Kapittel 2 Utvalgsmetoder 33
- Kapittel 3 Observere eller eksperimentere? 47
- Kapittel 4 Datamaskinens rolle i statistisk analyse 57

DEL 2 BESKRIVENDE STATISTIKK: OPPSUMMERING OG VISUALISERING AV DATA

- Kapittel 5 Variabler 67
- Kapittel 6 Bruk av grafer for å beskrive data 79
- Kapittel 7 Bruk av tall til å oppsummere data 104
- Kapittel 8 Samvariasjon mellom to variabler 132

DEL 3 SANNSYNLIGHET

- Kapittel 9 Grunnleggende sannsynlighetsteori 157
- Kapittel 10 Generell sannsynlighetsregning 186
- Kapittel 11 Mer om sannsynlighetsregler (*) 208
- Kapittel 12 Diskrete tilfeldige variabler 216
- Kapittel 13 Kontinuerlige tilfeldige variabler 237
- Kapittel 14 Sentraltendens, spredning og samvariasjon for tilfeldige variabler 265
- Kapittel 15 Utvalgsfordelinger og sentralgrenseteoremet 292
- Kapittel 16 Flere egenskaper ved normalfordelte variabler (*) 317

DEL 4 STATISTISK INFERENS: TEORI

- Kapittel 17 Introduksjon til statistisk inferens 327
- Kapittel 18 Oversikt over konfidensintervaller 338
- Kapittel 19 Oversikt over hypotesetesting 352
- Kapittel 20 Mer om hypotesetester 382

DEL 5 STATISTISK INFERENS: PRAKSIS

- Kapittel 21 Inferens for et gjennomsnitt i små utvalg 411
- Kapittel 22 Inferens for en andel 432
- Kapittel 23 Inferens for å sammenlikne to grupper 447
- Kapittel 24 Khikvadrattest for sannsynligheter 468

DEL 6 INFERENS MED FLERE VARIABLER

- Kapittel 25 Inferens for enkel regresjon 483
- Kapittel 26 Samvariasjon for to kategoriske variabler 518

Vedlegg 531

Stikkordregister 540

Innhold

DEL 1 INTRODUKSJON TIL STATISTIKK

KAPITTEL 1

Populasjon og utvalg 19

- 1.1 Om statistikk 20
- 1.2 Et reelt eksempel: Spørreundersøkelse om bankbransjen 21
- 1.3 Populasjon og utvalg 23
- 1.4 Elementene i en statistisk analyse 24
- 1.5 Beskrivende statistikk og statistisk inferens 25
- 1.6 Parametre og observatorer 26
- 1.7 Hva har vi lært? 29
- 1.8 Oppgaver 30
- 1.9 Oppgaveløsninger 32

KAPITTEL 2

Utvalgsmetoder 33

- 2.1 En analogi 34
- 2.2 Om gode og dårlige utvalgsmetoder 34
- 2.3 Tilfeldig utvalg 37
- 2.4 Klyngeutvalg (*) 38
- 2.5 Stratifisert utvalg (*) 40
- 2.6 Hva har vi lært? 43
- 2.7 Oppgaver 44
- 2.8 Oppgaveløsninger 46

KAPITTEL 3

Observere eller eksperimentere? 47

- 3.1 Årsakssammenheng eller skjulte variabler? 49
- 3.2 Randomiserte eksperimenter 51
- 3.3 Observasjonelle studier 52
- 3.4 Hva har vi lært? 54
- 3.5 Oppgaver 55
- 3.6 Oppgaveløsninger 56

KAPITTEL 4

Datamaskinens rolle i statistisk analyse 57

- 4.1 Statistisk programvare 58
- 4.2 Datafila 59
- 4.3 Hva har vi lært? 61
- 4.4 Oppgaver 62
- 4.5 Oppgaveløsninger 63

DEL 2 BESKRIVENDE STATISTIKK: OPPSUMMERING OG VISUALISERING AV DATA

KAPITTEL 5

Variabler 67

- 5.1 Variabler 68
- 5.2 Kategoriske og kvantitative variabler 69
- 5.3 Variabler i gråsonen mellom kategorisk og kvantitativ 70
- 5.4 Målenivå 71
- 5.5 Respons- og prediktorvariabler 74
- 5.6 Hva har vi lært? 76
- 5.7 Oppgaver 77
- 5.8 Oppgaveløsninger 78

KAPITTEL 6

Bruk av grafer for å beskrive data 79

- 6.1 Et første blick på dataene 80
- 6.2 Grafer for kategoriske data 80
- 6.3 Grafer for kvantitative data 83
- 6.4 Litt om skjeve og symmetriske fordelinger 88
- 6.5 Fra histogram til tetthetskurve 90
- 6.6 Grafer som viser samvariasjon for to variabler 91
- 6.7 Hva har vi lært? 98
- 6.8 Oppgaver 99
- 6.9 Oppgaveløsninger 102

KAPITTEL 7

Bruk av tall til å oppsummere data 104

- 7.1** Oppsummere kategoriske variabler 105
- 7.2** Sentraltendens: Gjennomsnitt og median 106
- 7.3** Spredning: Standardavvik og varians 109
- 7.4** Tolkning av s når histogrammet er nær normalfordelt 113
- 7.5** Standardisere en observasjon: z-verdien 115
- 7.6** Standardisering av normalfordelingskurver 116
- 7.7** Å sjekke normalfordelingen med QQ-plott 119
- 7.8** En alternativ måte å måle spredning på: Interkvartilbredde 123
- 7.9** Kalkulatortrening 124
- 7.10** Hva har vi lært? 126
- 7.11** Oppgaver 127
- 7.12** Oppgaveløsninger 130

KAPITTEL 8

Samvariasjon mellom to variabler 132

- 8.1** Samvariasjon mellom to kategoriske variabler 133
- 8.2** Samvariasjon mellom to kvantitative variabler 135
- 8.3** Korrelasjon 136
- 8.4** Om rette linjer 142
- 8.5** Minste kvadraters metode og regresjonslinja 144
- 8.6** Tolkning og prognose 146
- 8.7** Hva har vi lært? 149
- 8.8** Oppgaver 150
- 8.9** Oppgaveløsninger 153

DEL 3 SANNSYNLIGHET

KAPITTEL 9

Grunnleggende sannsynlighetsteori 157

- 9.1** Hvorfor trenger vi sannsynlighetsregning i statistikk? 158
- 9.2** Hva er en sannsynlighet? 158
- 9.3** Tilfeldige eksperimenter og utfallsrommet 163

- 9.4** Hendelser og mengdelære 166
- 9.5** De første sannsynlighetsmodellene 168
- 9.6** Telling, permutasjoner og kombinatorikk 172
- 9.7** Hva har vi lært? 181
- 9.8** Oppgaver 182
- 9.9** Oppgaveløsninger 184

KAPITTEL 10

Generell sannsynlighetsregning 186

- 10.1** Betinget sannsynlighet 186
- 10.2** Uavhengighet 193
- 10.3** Loven om total sannsynlighet 197
- 10.4** Bayes' formel 199
- 10.5** Anvendelse i DNA-testing 202
- 10.6** Hva har vi lært? 205
- 10.7** Oppgaver 206
- 10.8** Oppgaveløsninger 207

KAPITTEL 11

Mer om sannsynlighetsregler (*) 208

- 11.1** Forklaring av hvorfor «gunstige delt på mulige» holder 208
- 11.2** Mer om addisjonsregelen for disjunkte hendelser 209
- 11.3** Begrunnelse av formelen for betinget sannsynlighet 210
- 11.4** Utvidelse av addisjonsloven til ikke-disjunkte mengder 211
- 11.5** Hva har vi lært? 213
- 11.6** Oppgaver 214
- 11.7** Oppgaveløsninger 215

KAPITTEL 12

Diskrete tilfeldige variabler 216

- 12.1** Tilfeldige variabler og hendelser 217
- 12.2** Sannsynlighetsfordelinger for diskrete variabler 218
- 12.3** Binomisk fordeling 222
- 12.4** Hypergeometrisk fordeling 226
- 12.5** Poisson-fordeling 230
- 12.6** Hva har vi lært? 232
- 12.7** Oppgaver 233
- 12.8** Oppgaveløsninger 235

KAPITTEL 13

Kontinuerlige tilfeldige variabler 237

- 13.1 Tetthetsfordelingen til en kontinuertlig tilfeldig variabel 238
- 13.2 Normalfordelingen 243
- 13.3 Å finne sannsynligheter for normalfordelingen 247
- 13.4 Den uniforme sannsynlighetsfordelingen 251
- 13.5 Eksponentialfordelingen 254
- 13.6 Kvantiler 256
- 13.7 Hva har vi lært? 260
- 13.8 Oppgaver 261
- 13.9 Oppgaveløsninger 263

KAPITTEL 14

Sentraltendens, spredning og samvariasjon for tilfeldige variabler 265

- 14.1 De store talls lov 266
- 14.2 Forventning til en tilfeldig variabel 270
- 14.3 Varians og standardavvik til en tilfeldig variabel 275
- 14.4 Forventning og varians til noen kjente fordelinger 279
- 14.5 Samvariasjon mellom to tilfeldige variabler 281
- 14.6 Hva har vi lært? 286
- 14.7 Oppgaver 287
- 14.8 Oppgaveløsninger 290

KAPITTEL 15

Utvalgsfordelinger og sentralgrenseteoremet 292

- 15.1 Å tilpasse til en sannsynlighetsfordeling ved å trekke fra en tilfeldig variabel 293
- 15.2 Om utvalgsfordelingen til gjennomsnittet 299
- 15.3 Mer om de store talls lov 302
- 15.4 Forventning og varians til en sum 303
- 15.5 Forventning og varians til gjennomsnittet 305

- 15.6 Sentralgrenseteoremet 307

- 15.7 Hva har vi lært? 312

- 15.8 Oppgaver 313

- 15.9 Oppgaveløsninger 315

KAPITTEL 16

Flere egenskaper ved normalfordelte variabler (*) 317

- 16.1 Affine transformasjoner og standardisering av en normalfordelt variabel 318
- 16.2 Sum av uavhengige normalfordelte variabler 320
- 16.3 Fordelingen til gjennomsnittet når vi trekker fra en normalfordelt variabel 321
- 16.4 Hva har vi lært? 323

DEL 4 STATISTISK INFERENS: TEORI

KAPITTEL 17

Introduksjon til statistisk inferens 327

- 17.1 Estimatorer og deres usikkerhet 328
- 17.2 Hvordan estimere en populasjonsparameter? 330
- 17.3 En eksplorativ analyse 332
- 17.4 Estimator for populasjonsandel 334
- 17.5 Estimatorer for varians og standardavvik 335
- 17.6 Notasjon for estimatorer 336
- 17.7 Hva har vi lært? 337

KAPITTEL 18

Oversikt over konfidensintervaller 338

- 18.1 Konfidensintervaller for populasjonsgjennomsnitt 342
- 18.2 Hva påvirker konfidensintervallet? 345
- 18.3 Mer om antagelsene som ligger bak konfidensintervallet (*) 347
- 18.4 Hva har vi lært? 349
- 18.5 Oppgaver 350
- 18.6 Oppgaveløsninger 351

KAPITTEL 19

Oversikt over hypotesetesting 352

- 19.1** Introduksjon til hypotesetesting 352
- 19.2** En generell beskrivelse av hypotesetester 353
- 19.3** Tosidige hypotesetester for populasjonsgjennomsnitt 356
- 19.4** p-verdier 360
- 19.5** Utregning av p-verdien for tosidige hypotesetester 362
- 19.6** Testobservatorer: Et annet perspektiv på testing 366
- 19.7** Hva hvis H_A er sann? 368
- 19.8** Signifikansnivået. Hva må til for å kalle H_0 urimelig? 369
- 19.9** Forholdet mellom konfidensintervaller og tosidige hypotesetester 373
- 19.10** En matematisk forklaring (*) 375
- 19.11** Beskrivelse av en tosidig hypotesetest med mer matematikk (*) 376
- 19.12** Hva har vi lært? 379
- 19.13** Oppgaver 380
- 19.14** Oppgaveløsninger 381

KAPITTEL 20

Mer om hypotesetester 382

- 20.1** Praktisk signifikans og statistisk signifikans 382
- 20.2** To vanlige misforståelser om p-verdier 383
- 20.3** Ensidige hypotesetester for populasjonsgjennomsnitt 384
- 20.4** Hvorfor nullhypotesen kan settes som en likhet (*) 389
- 20.5** Beslutninger, formelle hypotesetester og type I- og type II-feil 392
- 20.6** Type I- og type II-feil for store utvalg: Konsistente tester 395
- 20.7** Fordelingen til p-verdien 396
- 20.8** Konsistens og definisjonen av p-verdier for ensidige tester (*) 397
- 20.9** Hva skal vi velge som nullhypotese og alternativ hypotese? 399

20.10 En introduksjon til teststyrke og utvalgsstørrelse (*) 401

20.11 Hva har vi lært? 405

20.12 Oppgaver 406

20.13 Oppgaveløsninger 407

DEL 5 STATISTISK INFERENS: PRAKSIS

KAPITTEL 21

Inferens for et gjennomsnitt i små utvalg 411

21.1 t-fordelingen 412

21.2 Er t-metodene robuste? 416

21.3 Konfidensintervall for populasjonsgjennomsnitt 417

21.4 Hypotesetest for et gjennomsnitt 421

21.5 Hva har vi lært? 427

21.6 Oppgaver 428

21.6 Oppgaveløsninger 430

KAPITTEL 22

Inferens for en andel 432

22.1 Utvalgsfordelingen til utvalgsandelen 433

22.2 Konfidensintervall for en andel 435

22.3 Hypotesetest for en andel 438

22.4 Hva har vi lært? 442

22.5 Oppgaver 443

22.6 Oppgaveløsninger 445

KAPITTEL 23

Inferens for å sammenlikne to grupper 447

23.1 Relaterte og uavhengige utvalg 448

23.2 Sammenlikne to gjennomsnitt 449

23.3 Sammenlikning av andeler i to grupper 457

23.4 Hva har vi lært? 463

23.5 Oppgaver 464

23.6 Oppgaveløsninger 466

KAPITTEL 24

Khikvadrattest for sannsynligheter 468

- 24.1** Observerte og forventete verdier 469
- 24.2** Test for sannsynligheter 471
- 24.3** Hva har vi lært? 476
- 24.4** Oppgaver 477
- 24.5** Oppgaveløsninger 479

DEL 6 INFERENS MED FLERE VARIABLER

KAPITTEL 25

Inferens for enkel regresjon 483

- 25.1** Regresjonsmodellen: Gjennomsnittet til y avhenger av x 484
- 25.2** Når er den enkle lineære regresjonsmodellen rimelig å bruke? 486
- 25.3** Inferens for β_0 og β_1 488
- 25.4** Regresjonstabell fra statistikkprogram 490
- 25.5** En nærmere kikk på residualene 492
- 25.6** Konfidensintervall for regresjonsmodellen 494
- 25.7** Mer om hypotesetester for regresjonsmodellen 496
- 25.8** Regresjonens standardfeil og standardfeilen til b_1 501
- 25.9** Prediksjonsintervall (*) 503

- 25.10** Mer om normalitetsantagelsen. En simulering (*) 506
- 25.11** Veien videre: Multippel lineær regresjon 508
- 25.12** Hva har vi lært? 510
- 25.13** Oppgaver 511
- 25.14** Oppgaveløsninger 516

KAPITTEL 26

Samvariasjon for to kategoriske variabler 518

- 26.1** Er variablene uavhengige, eller samvarierer de? 519
- 26.2** Observerte og forventete verdier i krysstabellen 520
- 26.3** Khikvadrattesten for samvariasjon mellom kategoriske variabler 521
- 26.4** Hva har vi lært? 525
- 26.5** Oppgaver 526
- 26.6** Oppgaveløsninger 528

VEDLEGG 531

Tabell A 532

Tabell B 534

Tabell C 536

Tabell D 538

STIKKORDREGISTER 540

Eksempler

Eksempel 1	21	Eksempel 39	121	Eksempel 77	189
Eksempel 2	35	Eksempel 40	133	Eksempel 78	189
Eksempel 3	35	Eksempel 41	134	Eksempel 79	190
Eksempel 4	36	Eksempel 42	135	Eksempel 80	191
Eksempel 5	36	Eksempel 43	135	Eksempel 81	192
Eksempel 6	39	Eksempel 44	138	Eksempel 82	194
Eksempel 7	39	Eksempel 45	139	Eksempel 83	195
Eksempel 8	41	Eksempel 46	140	Eksempel 84	198
Eksempel 9	42	Eksempel 47	142	Eksempel 85	199
Eksempel 10	48	Eksempel 48	143	Eksempel 86	200
Eksempel 11	50	Eksempel 49	146	Eksempel 87	200
Eksempel 12	51	Eksempel 50	147	Eksempel 88	201
Eksempel 13	52	Eksempel 51	148	Eksempel 89	204
Eksempel 14	53	Eksempel 52	148	Eksempel 90	212
Eksempel 15	68	Eksempel 53	159	Eksempel 91	219
Eksempel 16	69	Eksempel 54	161	Eksempel 92	219
Eksempel 17	70	Eksempel 55	162	Eksempel 93	221
Eksempel 18	71	Eksempel 56	163	Eksempel 94	222
Eksempel 19	72	Eksempel 57	164	Eksempel 95	222
Eksempel 20	73	Eksempel 58	164	Eksempel 96	223
Eksempel 21	74	Eksempel 59	165	Eksempel 97	225
Eksempel 22	74	Eksempel 60	166	Eksempel 98	226
Eksempel 23	86	Eksempel 61	169	Eksempel 99	227
Eksempel 24	87	Eksempel 62	171	Eksempel 100	228
Eksempel 25	90	Eksempel 63	172	Eksempel 101	228
Eksempel 26	105	Eksempel 64	173	Eksempel 102	229
Eksempel 27	109	Eksempel 65	174	Eksempel 103	230
Eksempel 28	110	Eksempel 66	175	Eksempel 104	231
Eksempel 29	110	Eksempel 67	175	Eksempel 105	231
Eksempel 30	111	Eksempel 68	176	Eksempel 106	238
Eksempel 31	113	Eksempel 69	177	Eksempel 107	240
Eksempel 32	114	Eksempel 70	177	Eksempel 108	242
Eksempel 33	115	Eksempel 71	178	Eksempel 109	245
Eksempel 34	117	Eksempel 72	178	Eksempel 110	246
Eksempel 35	118	Eksempel 73	179	Eksempel 111	247
Eksempel 36	120	Eksempel 74	186	Eksempel 112	248
Eksempel 37	120	Eksempel 75	187	Eksempel 113	248
Eksempel 38	121	Eksempel 76	187	Eksempel 114	249

Eksempel 115	250	Eksempel 152	329	Eksempel 189	450
Eksempel 116	251	Eksempel 153	331	Eksempel 190	451
Eksempel 117	253	Eksempel 154	338	Eksempel 191	453
Eksempel 118	255	Eksempel 155	339	Eksempel 192	456
Eksempel 119	257	Eksempel 156	344	Eksempel 193	456
Eksempel 120	258	Eksempel 157	346	Eksempel 194	457
Eksempel 121	258	Eksempel 158	352	Eksempel 195	458
Eksempel 122	266	Eksempel 159	356	Eksempel 196	458
Eksempel 123	267	Eksempel 160	359	Eksempel 197	460
Eksempel 124	268	Eksempel 161	364	Eksempel 198	461
Eksempel 125	271	Eksempel 162	365	Eksempel 199	469
Eksempel 126	271	Eksempel 163	369	Eksempel 200	470
Eksempel 127	272	Eksempel 164	373	Eksempel 201	470
Eksempel 128	273	Eksempel 165	374	Eksempel 202	471
Eksempel 129	274	Eksempel 166	385	Eksempel 203	473
Eksempel 130	274	Eksempel 167	386	Eksempel 204	474
Eksempel 131	275	Eksempel 168	389	Eksempel 205	484
Eksempel 132	276	Eksempel 169	392	Eksempel 206	490
Eksempel 133	276	Eksempel 170	393	Eksempel 207	493
Eksempel 134	277	Eksempel 171	403	Eksempel 208	494
Eksempel 135	278	Eksempel 172	404	Eksempel 209	495
Eksempel 136	280	Eksempel 173	414	Eksempel 210	497
Eksempel 137	281	Eksempel 174	414	Eksempel 211	498
Eksempel 138	282	Eksempel 175	415	Eksempel 212	499
Eksempel 139	283	Eksempel 176	419	Eksempel 213	502
Eksempel 140	284	Eksempel 177	419	Eksempel 214	503
Eksempel 141	285	Eksempel 178	420	Eksempel 215	504
Eksempel 142	293	Eksempel 179	421	Eksempel 216	505
Eksempel 143	294	Eksempel 180	425	Eksempel 217	519
Eksempel 144	296	Eksempel 181	426	Eksempel 218	520
Eksempel 145	297	Eksempel 182	433	Eksempel 219	521
Eksempel 146	300	Eksempel 183	434	Eksempel 220	522
Eksempel 147	301	Eksempel 184	435	Eksempel 221	522
Eksempel 148	302	Eksempel 185	437	Eksempel 222	522
Eksempel 149	308	Eksempel 186	441	Eksempel 223	523
Eksempel 150	310	Eksempel 187	441		
Eksempel 151	311	Eksempel 188	449		