Teknisk Måling

Statistik

Poul S. Eriksen

Steffen L. Lauritzen

Institut for Matematiske Fag
Aalborg Universitet

Forord

Disse noter udgør en stærkt omarbejdet og udvidet udgave af noter, som er udarbejdet over en årrække i forbindelse med undervisningen i statistik og teknisk måling på landinspektørstudiet i Aalborg. Det har specielt været hensigten at gøre notesættet uafhængigt af de statistikbøger, som eventuelt har været anvendt på tidligere semestre, således at det giver mening at læse dem selvstændigt.

På den anden side er det en ubetinget fordel, hvis læseren tidligere har haft et elementært statistikkursus, og iøvrigt kender til udjævning efter mindste kvadraters metode.

Aalborg, januar 2004

Poul S. Eriksen og Steffen L. Lauritzen

Indhold

1 Grundlæggende begreber

I dette afsnit skal vi repetere elementære begreber fra sandsynlighedsregning og statistik.

1.1 Stokastiske variable og deres fordelinger

I det følgende vil X betegne en *kontinuert stokastisk variabel*. X har en *tæthedsfunktion*, som vi vil betegne f_X. Da gælder

$$f_X(x) \geq 0$$

samt for ethvert interval $[a, b]$

$$P(X \in [a, b]) = \int_a^b f_X(x)\,dx.$$

Da $P(-\infty < X < \infty) = 1$ fås at

$$\int_{-\infty}^{\infty} f_X(x)\,dx = 1.$$

Middelværdien af X er givet ved

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)\,dx.$$

Lad h være en reel funktion. Da er $Y = h(X)$ en stokastisk variabel, og der gælder

$$E(Y) = E\{h(X)\} = \int_{-\infty}^{\infty} h(x) f_X(x)\,dx. \tag{1}$$

Hvis a og b er reelle tal og $h(x) = ax + b$, indses let at

$$E(aX + b) = aE(X) + b.$$

Variansen på X er givet ved

$$\mathrm{Var}(X) = E\left[\{X - E(X)\}^2\right].$$

Idet $h(x) = \{x - E(X)\}^2$ fås af (1) at

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} \{x - E(X)\}^2 f_X(x)\,dx.$$

Heraf indses forholdsvis let at

$$\mathrm{Var}(X) = E(X^2) - \{E(X)\}^2.$$

Hvis a og b er reelle tal gælder endvidere

$$\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X). \tag{2}$$

Spredningen på X defineres som

$$\mathrm{Spr}(X) = \sqrt{\mathrm{Var}(X)}.$$

Af (2) følger så at

$$\mathrm{Spr}(aX + b) = |a|\mathrm{Spr}(X).$$

Ved måling af positive størrelser, som for eksempel afstande, benytter man ofte *variationskoefficienten*, der er bestemt som $\mathrm{Spr}(X)/E(X)$. Den angives normalt i % og er dimensionsløs, idet vi for $a > 0$ har, at

$$\frac{\mathrm{Spr}(aX)}{E(aX)} = \frac{a\mathrm{Spr}(X)}{aE(X)} = \frac{\mathrm{Spr}(X)}{E(X)}.$$

Fordelingsfunktionen for X defineres som

$$F_X(x) = P(X \leq x).$$

Da

$$P(X \leq x) = P(X \in]-\infty, x]) = \int_{-\infty}^{x} f_X(t)dt,$$

gælder med andre ord

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt,$$

hvoraf det så følger at

$$f_X(x) = \frac{d}{dx}F_X(x) = F_X'(x).$$

Det er ofte nyttigt at arbejde med værdierne af den inverse fordelingsfunktion. Disse betegnes fraktiler (eng. quantiles). Konkret er α-*fraktilen* x_α for X er givet ved ligningen

$$F_X(x_\alpha) = \alpha,$$

eller ækvivalent hermed

$$x_\alpha = F^{-1}(\alpha).$$

Lad nu $a > 0$ og b være reelle tal og sæt $Y = aX + b$. Så gælder om fordelingsfunktionen for Y at

$$F_Y(y) = P(aX + b \leq y) = P\left(X \leq \frac{y-b}{a}\right) = F_X\left(\frac{y-b}{a}\right),$$

hvoraf

$$\begin{aligned} f_Y(y) &= \frac{d}{dy}\left\{F_X\left(\frac{y-b}{a}\right)\right\} \\ &= F_X'\left(\frac{y-b}{a}\right)\frac{d}{dy}\left(\frac{y-b}{a}\right) = f_X\left(\frac{y-b}{a}\right)\frac{1}{a}. \end{aligned}$$

Det vil sige at

$$F_{aX+b}(y) = F_X\left(\frac{y-b}{a}\right) \tag{3}$$

og endvidere

$$f_{aX+b}(y) = \frac{1}{a} f_X\left(\frac{y-b}{a}\right).$$

Om fraktilerne gælder den simple relation

$$y_\alpha = a x_\alpha + b.$$

1.2 Normalfordelingen

Antag at X har tæthedsfunktionen givet ved

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right).$$

Da $\phi(x) = \phi(-x)$ indses let at

$$E(X) = \int_{-\infty}^{\infty} x\phi(x)\,dx = 0.$$

Endvidere kan det vises at

$$E(X^2) = \int_{-\infty}^{\infty} x^2\phi(x)\,dx = 1.$$

Heraf følger at $\mathrm{Var}(X) = 1$.

Fordelingen af X kaldes *normalfordelingen med middelværdi 0 og varians* 1, eller undertiden *den standardiserede normalfordeling* eller endda *U-fordelingen.* Fordelingsfunktionen for X betegnes ofte med Φ og er altså givet ved

$$\Phi(a) = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx.$$

Denne funktion er tabelleret mange forskellige steder. Mange lommeregnere indeholder Φ som standardfunktion, og den er også tilgængelig i stort set enhver form for software med statistiske faciliteter.

Lad $\sigma > 0$ og $\mu \in R$ og betragt den stokastiske variabel $Y = \sigma X + \mu$. Da fås middelværdien til

$$E(Y) = \sigma E(X) + \mu = \mu.$$

Tilsvarende variansen

$$\mathrm{Var}(Y) = \sigma^2 \mathrm{Var}(X) = \sigma^2$$

og tæthedsfunktionen

$$f_Y(y) = \frac{1}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(y-\mu)^2}{2\sigma^2}\right\}.$$

Fordelingen af Y kaldes *normalfordelingen med middelværdi μ og varians* σ^2, og man skriver ofte $Y \sim \mathcal{N}(\mu, \sigma^2)$. Fordelingsfunktionen for Y er i henhold til (3) givet ved

$$F_Y(y) = \Phi\left(\frac{y-\mu}{\sigma}\right).$$

Herved kan man finde sandsynligheder i normalfordelingen med middelværdi μ og varians σ^2 ud fra den standardiserede normalfordeling.

1.3 Uafhængighed

Et sæt $X_1, \ldots, X_n$ af stokastiske variable siges at være *uafhængige*, såfremt der for alle intervaller $A_1, \ldots, A_n$ på R gælder

$$P(X_1 \in A_1, X_2 \in A_2, \ldots, X_n \in A_n) = P(X_1 \in A_1)P(X_2 \in A_2) \cdots P(X_n \in A_n).$$

Lad h være en funktion af $k < n$ variable. Hvis $X_1, X_2, \ldots, X_n$ er uafhængige stokastiske variable, da gælder at

$$h(X_1, \ldots, X_k), X_{k+1}, \ldots, X_n \text{ er uafhængige.} \tag{4}$$

Lad X og Y være stokastiske variable. Da er $X + Y$ igen en stokastisk variabel, hvorom det gælder, at

$$E(X + Y) = E(X) + E(Y). \tag{5}$$

Med hensyn til produktet af X og Y gælder, at

$$E(XY) = E(X)E(Y), \quad \text{hvis } X \text{ og } Y \text{ er uafhængige.} \tag{6}$$

Af (5) og (6) følger, at

$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y), \text{ hvis } X \text{ og } Y \text{ er uafhængige.} \tag{7}$$

Man har således, at hvis $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$ og $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$ er uafhængige, så er $E(X+Y) = \mu_1 + \mu_2$ og $\mathrm{Var}(X+Y) = \sigma_1^2 + \sigma_2^2$. Da man endvidere kan vise, at $X+Y$ er normalfordelt, har vi med andre ord, at hvis $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$ og $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$ er uafhængige, så gælder at

$$X + Y \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2). \tag{8}$$

1.4 Fordelinger afledt af normalfordelingen

I det følgende beskriver vi tre fundamentale fordelinger. De fremkommer alle udfra normalfordelingen ved forskellige regneoperationer, som typisk foretages i forbindelse med udjævning. Fordelingerne beskrives efter voksende kompleksitet, og historisk tog det omkring 60 år at finde dem alle tre, i den angivne rækkefølge og med omtrent 20 års mellemrum imellem opdagelserne.

1.4.1 Helmerts χ^2

Hvis $U_1, \ldots, U_d$ er uafhængige og standard normalfordelte kan man vise, at

$$Y = U_1^2 + \cdots + U_d^2 \tag{9}$$

har tæthedsfunktion

$$f_Y(y;d) = k_d y^{d/2-1} e^{-y/2}, \text{ for } y > 0,$$

hvor k_d er en konstant, som sikrer, at tæthedsfunktionen har integral 1

$$k_d^{-1} = \int_0^\infty y^{d/2-1} e^{-y/2}\, dy.$$

Normaliseringskonstanten kan beregnes til $k_d^{-1} = \Gamma(d/2)2^{d/2}$, hvor Γ er gammafunktionen. Denne fordeling kaldes χ^2-*fordelingen med d frihedsgrader* (*eng.* degrees of freedom).

Residualkvadratsummen efter en udjævning med d overbestemmelser er netop χ^2-fordelt med d frihedsgrader, hvis målingerne er angivet i enheder, så a priori spredningen er lig med 1.

Dette generelle resultat blev første gang vist af den tyske geodæt og landmåler F. R. Helmert i 1876. Udtrykket 'frihedsgrader' henviser til frihedsgraderne i det ligningssystem, som hører til udjævningen.

Der gælder endvidere, at

$$E(Y) = d, \quad \text{Var}(Y) = 2d, \tag{10}$$

altså at både middelværdi og varians vokser proportionalt med antallet af frihedsgrader.

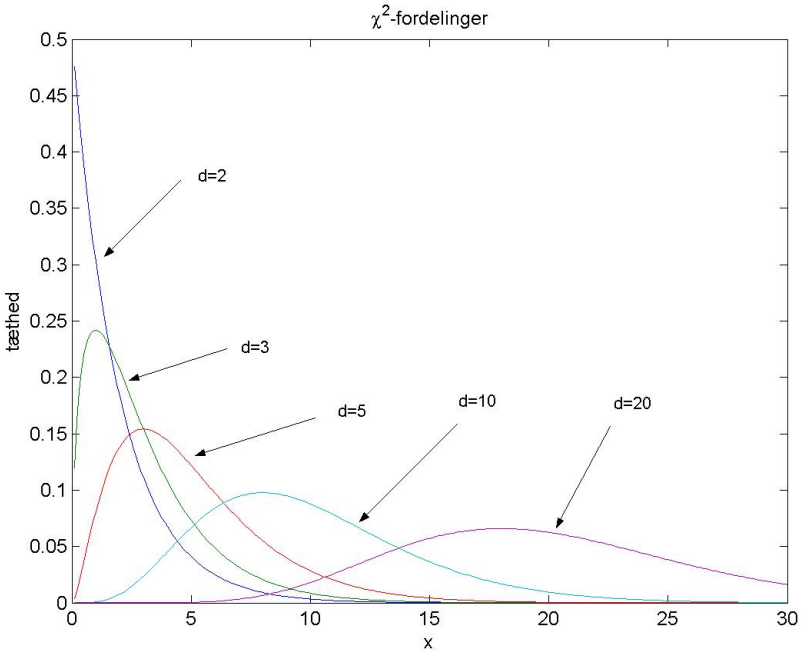
Det følger direkte af (9), at hvis Y_1 og Y_2 er uafhængige og χ^2-fordelte med hhv. d_1 og d_2 frihedsgrader, så bliver summen $Y = Y_1 + Y_2$ χ^2-fordelt med $d_1 + d_2$ frihedsgrader. For vi har, at

$$Y = Y_1 + Y_2 = U_1^2 + \cdots + U_{d_1}^2 + U_{d_1+1}^2 + \cdots + U_{d_1+d_2}^2.$$

Dette resultat er vigtigt, når residualkvadratsummer fra forskellige udjævninger foretaget med samme a priori spredning kombineres ved simpel addition.

A posteriori variansen efter udjævning bestemmes normalt som Y/d, hvor Y er residualkvadratsummen, se nærmere i afsnit 2.2.

χ^2-fordelingen har top i nulpunktet, medmindre $d \geq 3$, altså vil man meget ofte se ekstremt små a posteriori varianser, hvis man kun har 1 eller 2 overbestemmelser, for eksempel hvis man baserer en udjævning udelukkende på dobbelt- eller triple-målinger.



Figur 1: Forskellige χ^2-fordelinger. Når antallet d af frihedsgrader vokser, bliver fordelingen mere symmetrisk og kommer til at ligne en normalfordeling.

1.4.2 Students t

Endnu en vigtig fordeling er *t-fordelingen med d frihedsgrader.* Den fremkommer som

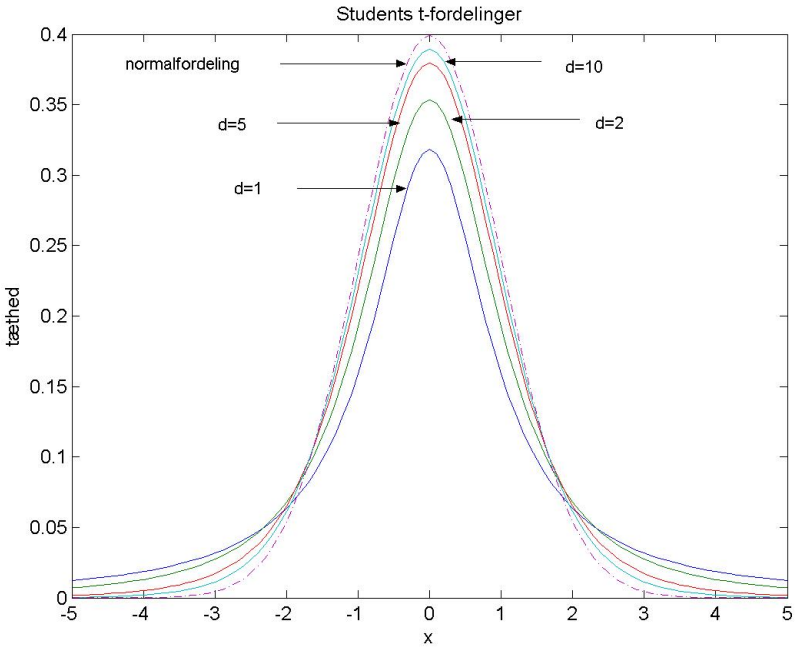
$$T = \frac{U}{\sqrt{Y/d}},$$

hvor U og Y er uafhængige, U er U-fordelt, og Y er χ^2-fordelt med d frihedsgrader. Tæthedsfunktionen er

$$f_T(t;d) = c_d \left(1 + t^2/d\right)^{-(d+1)/2},$$

hvor c_d igen er en normaliseringskonstant

$$c_d^{-1} = \int_{-\infty}^{\infty} \left(1 + t^2/d\right)^{-(d+1)/2} dt.$$



Figur 2: Eksempler på Students t-fordelinger. Når antallet d af frihedsgrader vokser, nærmer fordelingen sig til en standard normalfordeling.

Fordelingen har meget tunge haler for d lille. For $d \to \infty$ nærmer t-fordelingen sig til en U-fordeling. Det er svært at se forskel på disse to fordelinger når $d > 30$. Nogle eksempler på t-fordelinger er vist på Figur 2.

"Student" er et pseudonym for den engelske statistiker W. S. Gossett, som fandt denne fordeling omkring år 1900.

I landmålingen optræder t-fordelingen typisk, når et residual U efter udjævningen skaleres med a posteriori spredningen $\sqrt{Y/d}$. Hvis antallet af overbestemmelser er mindre end 30, kan det være misvisende at ignorere denne skalering og bruge normalfordelingen i stedet for.

1.4.3 Fishers F

Fordelingen af

$$V = \frac{Y_1/d_1}{Y_2/d_2},$$

hvor Y_1 og Y_2 er uafhængige og χ^2-fordelte med hhv. d_1 og d_2 frihedsgrader kaldes *F-fordelingen med (d_1, d_2) frihedsgrader.* Tæthedsfunktionen er

$$f_V(v; d_1, d_2) = k_{d_1,d_2} \frac{v^{(d_2/2)-1}}{(d_1 v + d_2)^{(d_1+d_2)/2}}, \text{ for } v > 0,$$

hvor k_{d_1,d_2} som sædvanlig er en normaliseringskonstant, der sikrer, at tæthedsfunktionen har integral 1.

F-fordelingen er relevant, når man skal sammenligne to a posteriori varianser, for eksempel beregnet ud fra to forskellige udjævninger.

Fordelingen er første gang fundet af den engelske statistiker R. A. Fisher omkring 1920 og kendes også under navnet *v^2-fordelingen.* Der gælder

$$E(V) = \frac{d_2}{d_2 - 2}.$$

Også denne fordeling har meget tunge haler, især når antallet af frihedsgrader for nævneren er lille, altså når nævnerens residualkvadratsum er baseret på meget få overbestemmelser. Formlen ovenfor fortæller endda, at middelværdien ikke er endelig medmindre $d_2 \geq 3$. Eksempler på F-fordelinger er vist i

Det er oplagt, at hvis V er F-fordelt med (d_1, d_2) frihedsgrader, så er $1/V$ F-fordelt med (d_2, d_1) frihedsgrader.

Endvidere er der en simpel forbindelse mellem t-fordelingen og F-fordelingen med 1 frihedsgrad i tælleren. Idet vi har

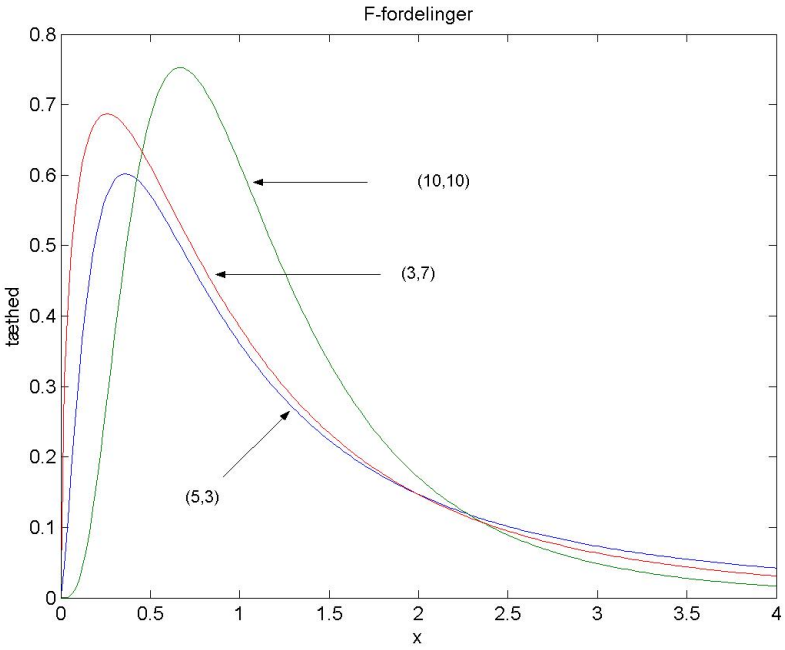
$$T^2 = \left(\frac{U}{\sqrt{Y/d}} \right)^2 = \frac{U^2}{Y/d},$$

ses let, at hvis T er t-fordelt med d frihedsgrader, så vil T^2 være F-fordelt med $(1, d)$ frihedsgrader.

2 Estimationsteori

2.1 Det grundlæggende problem

Teorien for statistisk estimation beskæftiger sig med den generelle situation, hvor man har en række målinger, der kan opfattes som observationer



Figur 3: Eksempler på Fishers F-fordelinger. Når antallet d_2 af frihedsgrader i nævneren vokser, nærmer fordelingen sig til en χ^2/d_1-fordeling med d_1 frihedsgrader, en χ^2-fordeling som er skaleret til at have middelværdi 1.

$x_1, \ldots, x_n$ af stokastiske variable $X_1, \ldots, X_n$ med tæthedsfunktioner, som afhænger af en ukendt *parameter* θ. Parameteren er typisk et element i en udjævning eller lignende.

Nu ønsker man at bestemme værdien af θ udfra $x_1, \ldots, x_n$. Vi siger at θ ønskes *estimeret*. Hertil anvendes typisk en *estimator* $\hat{\theta} = g(x_1, \ldots, x_n)$, altså en passende valgt funktion af observationerne.

Men umiddelbart er der vilkårligt mange muligheder for at vælge sådan en estimator $\hat{\theta}$. Der er derfor behov for at diskutere principper, så man kan identificere visse estimatorer som gode og rigtige og andre som dårlige og forkerte.

Eksempel 1 Vi forestiller os, at vi har målt en højdeforskel mellem punkterne A og B 3 gange og fået følgende resultater (målt i mm)

$$119, 112, 114.$$

De parametre, vi er interesseret i at estimere kunne for eksempel være

- μ—den sande højdeforskel,
- σ—målemetodens sande spredning.

Man kunne nu foreslå flere måder at gøre dette på:

1. $\hat{\mu}_1 = \bar{x} = 115$, $\hat{\sigma}_1 = s = 3.61$;
2. $\hat{\mu}_2 = x_{(2)} = 114$, $\hat{\sigma}_2 = (x_{(3)} - x_{(1)})/2 = 3.5$.

Den første af disse metoder er den, som vi normalt anvender for normalfordelte observationer (se nedenfor), og som vi kender fra almindelig udjævning, altså hvor gennemsnittet bruges til at estimere den teoretiske middelværdi og den empiriske spredning til at estimere den teoretiske.

Den anden metode ser måske usædvanlig ud, men på forhånd behøvede den ikke at være ufornuftig. Her estimerer vi μ med den midterste observation og spredningen bedømmes udfra forskellen på den største og mindste observation. Vi vil senere diskutere dette eksempel i detaljer. □

2.2 Estimation i normalfordelingen

Før vi mere generelt diskuterer principper for estimation, vil vi repetere, hvorledes vi normalt estimerer parametre i normalfordelingen i det simpleste tilfælde med gentagne målinger af samme størrelse.

Lad derfor $X_1, X_2, \ldots, X_n$ være uafhængige således at $X_i \sim \mathcal{N}(\mu, \sigma^2)$ $i = 1, \ldots, n$. Det almindeligt brugte estimat for μ — efter mindste kvadraters metode — er givet ved gennemsnittet

$$\hat{\mu} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n}(X_1 + \cdots + X_n),$$

mens estimatet for σ^2 tilsvarende er givet ved

$$\hat{\sigma}^2 = S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

De stokastiske variable $\bar{X}$ og S^2 kan vises at være stokastisk uafhængige og

$$\bar{X} \sim \mathcal{N}(\mu, \sigma^2/n), \quad (n-1)S^2/\sigma^2 \sim \chi^2(n-1)$$

det vil sige, $(n-1)S^2/\sigma^2$ er χ^2-fordelt med $(n-1)$ frihedsgrader.

Mere generelt giver mindste kvadraters udjævning anledning til et estimat for observationernes middelværdivektor, og variansen estimeres ved hjælp af *a posteriori variansen*, som er den normerede residualkvadratsum $\hat{r}^\top C\hat{r}/d$, hvor d er antallet af overbestemmelser.

2.3 Egenskaber for estimatorer

Vi vil nu se på forskellige egenskaber, som estimatorer kan have.

2.3.1 Centralitet

En estimator $\hat{\theta} = g(x_1, \ldots, x_n)$ kaldes *central* (*eng.* unbiased) for θ hvis

$$Eg(X_1, \ldots, X_n) = \theta.$$

Man bruger også betegnelsen 'middelværdiret' idet ligningen ovenfor kan tolkes således, at man 'i middel' estimerer parameteren korrekt og uden en systematisk skævhed (*eng.* bias).

I normalfordelingstilfældet gælder, at $\hat{\mu} = \bar{x}$ er en central estimator for middelværdien μ, og $\hat{\sigma}^2 = s^2$ er en central estimator for σ^2.

Generelt er det rigtigt, at mindste kvadraters udjævning er central, både for elementerne i en elementudjævning og også for middelværdivektoren, og endvidere er a posteriori variansen en central estimator for σ^2.

Det er naturligvis en udmærket egenskab for disse estimatorer. Men det bør dog bemærkes, at $\hat{\sigma}$ ikke er en central estimator for σ, idet der gælder

$$E(\hat{\sigma}) \leq \sqrt{E(\hat{\sigma}^2)} = \sigma,$$

så spredningen i den forstand altid underestimeres. Så undertiden bruger vi åbenbart gerne estimatorer, der ikke er centrale.

Centralitet er især vigtigt, når man ønsker at kombinere estimatorer fra flere forskellige måleserier. Hvis vi har centrale estimatorer $\hat{\theta}_1, \ldots, \hat{\theta}_n$ for en parameter θ målt med præcisioner (vægte) $p_1, \ldots, p_n$, så kan vi kombinere dem i et gennemsnit til

$$\hat{\theta} = \frac{p_1\hat{\theta}_1 + \cdots + p_n\hat{\theta}_n}{p_1 + \cdots + p_n},$$

som så igen vil være en central estimator for θ som har en større præcision end hver af de enkelte estimatorer. Hvis derimod $\hat{\theta}_i$ hver især har en skævhed β, så

$$E(\hat{\theta}_i) = \theta + \beta,$$

vil der også gælde $E(\hat{\theta}) = \theta + \beta$, så skævheden ikke reduceres ved kombinationen. På den måde vil skævheden blive den totalt dominerende fejlkilde, hvis n er meget stor.

Det er (blandt andet) derfor vigtigt, at man ved kombination af resultater fra forskellige udjævninger beregner vægtede gennemsnit af a posteriori varianser (som er centrale estimatorer) og *ikke* gennemsnit af a posteriori spredninger.

Eksempel 2 Vi betragter igen situationen fra før, hvor vi har observeret højdeforskelle $119, 112, 114$, alt målt i mm.

Her er *medianen* (den midterste observation) $x_{(2)} = 114$ også en central estimator for højdeforskellen og derfor et muligt alternativ til $\bar{x}$.

Hvis observationerne er normalfordelte $\mathcal{N}(\mu, \sigma^2)$ kan det vises, at

$$E\hat{\sigma}_1 = \frac{\sqrt{\pi}}{2}\sigma = 0.89\sigma$$

og

$$E\hat{\sigma}_2 = 0.85\sigma.$$

Ingen af disse estimatorer er derfor centrale for spredningen. Så måske var det bedre at korrigere disse estimatorer og bruge

$$\hat{\sigma}_3 = s/0.89 = 4.15 \text{ eller } \hat{\sigma}_4 = \hat{\sigma}_2/0.85 = 4.118.$$

Vi skal vende tilbage til denne problemstilling. □

2.3.2 Minimal varians

Hvis man ønsker at anvende centrale estimatorer — og det kan ofte være hensigtsmæssigt — vælger man naturligvis helst en estimator med så lille varians som muligt, altså med så stor præcision som muligt.

Hvis målingerne er normalfordelte, giver mindste kvadraters udjævning den bedst mulige centrale estimator (altså den med minimal varians) for

elementerne. Man siger, at estimatoren er en *Minimum Variance Unbiased Estimator* (MVUE).

Endvidere er a posteriori variansen også MVUE for variansen på målefejlen i det normalfordelte tilfælde. Dette er et hovedargument for at anvende denne i stedet for de alternativer, der er nævnt i eksemplet ovenfor.

Nu er det ikke altid rimeligt at antage, at målefejlene er eksakt normalfordelte. Derfor er det vigtigt at vide, at mindste kvadraters udjævning *altid* giver den bedst mulige centrale estimator blandt de, som kan udtrykkes som en *lineær* funktion af observationerne. Her bruges udtrykket *Best Linear Unbiased Estimator* (BLUE). Dette beviste Gauss omkring 1823 og brugte det som et hovedargument for at anvende mindste kvadraters metode til udjævning.

Udenfor normalfordelingen er s^2 stadig en central estimator for σ^2. Den er også *konsistent*, d.v.s. at hvis antallet af frihedsgrader går mod uendelig, vil a posteriori variansen nærme sig den sande varians. Men s^2 har ikke nødvendigvis andre optimalitetsegenskaber udenfor normalfordelingen.

Eksempel 3 Vi vender atter tilbage til situationen med de tre højdemålinger $119, 112, 114$.

Hvis observationerne kan antages normalfordelte, er gennemsnittet $\bar{x} = 115$ den bedste centrale estimator. I eksemplet er *medianen* $x_{(2)}$ ikke en lineær funktion af målingerne og under visse antagelser om fejlfordelingen, er den mere præcis end gennemsnittet.

Estimatorer af typen $x_{(2)}$ er generelt mere *robuste* end $\bar{x}$ mod grove fejl, men sværere at beregne. De opstår ved at minimere summen af absolutte afvigelser $\sum |\hat{r}_i|$ af residualerne i stedet for kvadratsummen, som anvendes i mindste kvadraters metode.

Estimatoren $(x_{(3)} - x_{(1)})/2$ for spredningen er til gengæld meget dårlig. En enkelt grov fejl vil slå fuldstændigt igennem på sådan en estimator. I eksemplet har $\hat{\sigma}_3$ spredning 0.521σ mens $\hat{\sigma}_4$ har spredning 0.725σ.

Havde vi haft en grov fejl, så 119 var blevet til 129, ville vores robuste estimator $x_{(2)}$ være uændret, mens $\bar{x}$ ville være steget fra 115 til 118.3.

De grove fejl har alvorlige konsekvenser for estimation af spredningen uanset hvilken estimator vi vælger; i eksemplet ville $(x_{(3)} - x_{(1)})/2$ være steget fra 3.5 til 8.5 og s fra 3.61 til 9.29. □

2.4 Konfidensintervaller

Man er ikke altid tilfreds med at få et enkelt gæt på værdien af θ. I stedet anvendes *konfidensmængder*, for eksempel konfidensintervaller, eller konfidenssellipser (i to dimensioner). Man taler undertiden også om *intervalestimater*.

Et konfidensinterval beregnes ved to grænser: en nedre grænse $C_1 = g_1(X_1, \ldots, X_n)$ og en øvre grænse $C_2 = g_2(X_1, \ldots, X_n)$. Intervallet $[C_1, C_2]$ siges at have *konfidensgrad* γ, hvis

$$P(C_1 \leq \theta \leq C_2) = \gamma.$$

Konfidensgraden angiver altså sandsynligheden for, at intervallet omfatter den sande værdi af parameteren. Vi vil derfor helst have små intervaller med stor konfidensgrad.

Konfidensintervaller kan ofte dannes udfra en estimator $\hat{\theta}$. Hvis denne har fordelingsfunktion F_θ, så vi har

$$P(\hat{\theta} \leq x) = F_\theta(x),$$

kan konfidensgrænserne i princippet vælges som fraktiler $c_1 = x_{\alpha/2}$ og $c_2 = x_{1-\alpha/2}$, hvor $\alpha = 1 - \gamma$. Problemet er, at disse fraktiler i almindelighed afhænger af den ukendte værdi θ. Men i en række interessante specialtilfælde kan man komme uden om det problem ved et simpelt kneb, som vi skal se eksempler på i det følgende.

2.4.1 Konfidensintervaller i normalfordelingen

Kendt spredning Hvis målevariansen er *kendt* og lig med a priori variansen σ_0, er det bedste konfidensinterval for μ af formen

$$c_1 = \bar{x} - u_{1-\alpha/2}\sigma_0/\sqrt{n}, \quad c_2 = \bar{x} + u_{1-\alpha/2}\sigma_0/\sqrt{n},$$

hvor $u_{1-\alpha/2}$ er $1 - \alpha/2$-fraktilen i normalfordelingen, altså $\Phi(u_{1-\alpha/2}) = 1 - \alpha/2$.

Konfidensgraden for dette interval er $\gamma = 1 - \alpha$, hvilket beregnes som nedenstående

$$\begin{aligned} P(C_1 < \mu < C_2) &= P(\bar{X} - u_{1-\alpha/2}\sigma_0/\sqrt{n} < \mu < \bar{x} + u_{1-\alpha/2}\sigma_0/\sqrt{n}) \\ &= P\left(\sqrt{n}\frac{\bar{X}-\mu}{\sigma_0} - u_{1-\alpha/2} < 0 < \sqrt{n}\frac{\bar{X}-\mu}{\sigma_0} + u_{1-\alpha/2}\right) \\ &= P\left(-u_{1-\alpha/2} < -\sqrt{n}\frac{\bar{X}-\mu}{\sigma_0} < u_{1-\alpha/2}\right) \\ &= \Phi(u_{1-\alpha/2}) - \Phi(-u_{1-\alpha/2}) \\ &= 1 - \alpha/2 - (1 - (1 - \alpha/2)) = 1 - \alpha \end{aligned}$$

hvor vi i sidste linie har udnyttet, at U-fordelingen er symmetrisk, så $-U$ har samme fordeling som U og $\Phi(-x) = 1 - \Phi(-x)$.

Det kneb, vi har anvendt i konstruktionen af konfidensintervallet er, at størrelsen $U = \sqrt{n}(\bar{X} - \mu)/\sigma_0$ er U-fordelt. Altså har vi fra vores estimator $\bar{X}$ kunnet konstruere en størrelse U, hvis fordeling *ikke afhænger af parameteren,* og derefter har vi kunnet 'regne baglæns'. En sådan kaldes en *pivotal* størrelse (*eng.* pivot).

Ukendt spredning Hvis vi i stedet ønsker at anvende a posteriori variansen s estimeret med d overbestemmelser, kan vi bruge den pivotale størrelse $T = \sqrt{n}(\bar{X} - \mu)/s$. Denne er nemlig t-fordelt med $d = n - 1$ frihedsgrader, så fordelingen afhænger altså heller ikke af μ og σ. Grænserne bliver nu

$$c_1 = \bar{x} - t_{1-\alpha/2}(d)s/\sqrt{n}, \quad c_2 = \bar{x} + t_{1-\alpha/2}(d)s/\sqrt{n}.$$

Her verificeres konfidensgraden på nøjagtig samme måde:

$$\begin{aligned} P(C_1 < \mu < C_2) &= P(\bar{X} - t_{1-\alpha/2}(d)s/\sqrt{n} < \mu < \bar{x} + t_{1-\alpha/2}(d)s/\sqrt{n}) \\ &= P\left(\sqrt{n}\frac{\bar{X}-\mu}{s} - t_{1-\alpha/2}(d) < 0 < \sqrt{n}\frac{\bar{X}-\mu}{s} + t_{1-\alpha/2}(d)\right) \\ &= P\left(-t_{1-\alpha/2}(d) < -\sqrt{n}\frac{\bar{X}-\mu}{s} < t_{1-\alpha/2}(d)\right) \\ &= F_T(t_{1-\alpha/2}; d) - F_T(-t_{1-\alpha/2}; d) \\ &= 1 - \alpha/2 - (1 - (1 - \alpha/2)) = 1 - \alpha, \end{aligned}$$

hvor vi også her har udnyttet, at T-fordelingen er symmetrisk, se Figur 2.

Eksempel 4 Vi vender atter tilbage til vores tre højdemålinger $119, 112, 114$. Hvis vi antager, at spredningen σ_0 er kendt og lig med 4, fås et 95% konfidensinterval til

$$115 \pm u_{0.975} \times \sigma_0/\sqrt{n} = 115 \pm 1.96 \times 4/\sqrt{3} = 115 \pm 4.53.$$

Med ukendt spredning fås

$$115 \pm t(2)_{0.975} \times s/\sqrt{n} = 115 \pm 4.30 \times 3.61/\sqrt{3} = 115 \pm 8.97,$$

altså et væsentligt bredere interval på grund af, at der også er usikkerhed på bestemmelse af spredningen. □

2.4.2 Konfidensintervaller for spredningen

Fra en a posteriori varians s^2 baseret på d overbestemmelser bestemmes et konfidensinterval for σ ved at udnytte den pivotale størrelse $Y = ds^2/\sigma^2$, som jo er χ^2-fordelt med d frihedsgrader. Konfidensintervallet bliver nu

$$c_1 = \sqrt{\frac{d}{\chi^2(d)_{1-\alpha/2}}}s, \quad c_2 = \sqrt{\frac{d}{\chi^2(d)_{\alpha/2}}}s,$$

hvor $\chi^2(d)_\lambda$ er λ-fraktilen i χ^2-fordelingen med r frihedsgrader. At konfidensgraden er netop bliver $\gamma = 1 - \alpha$ ses ved følgende regninger

$$\begin{aligned}
P(C_1 < \sigma < C_2) &= P\left(\sqrt{\frac{d}{\chi^2(d)_{1-\alpha/2}}}s < \sigma < \sqrt{\frac{d}{\chi^2(d)_{\alpha/2}}}s\right) \\
&= P\left(\frac{ds^2}{\chi^2(d)_{1-\alpha/2}} < \sigma^2 < \frac{ds^2}{\chi^2(d)_{\alpha/2}}\right) \\
&= P\left(\frac{1}{\chi^2(d)_{1-\alpha/2}} < \frac{\sigma^2}{ds^2} < \frac{1}{\chi^2(d)_{\alpha/2}}\right) \\
&= P\left(\chi^2(d)_{\alpha/2} < \frac{ds^2}{\sigma^2} < \chi^2(d)_{1-\alpha/2}\right) \\
&= F_Y(\chi^2_{1-\alpha/2}; d) - F_Y(\chi^2_{\alpha/2}; d) = 1 - \alpha/2 - \alpha/2 = 1 - \alpha.
\end{aligned}$$

Bemærk, at vi her har brugt forskellige fraktiler i de to af fordelingens haler, da χ^2-fordelingen ikke er symmetrisk.

Eksempel 5 Vi vender tilbage til de tre højdemålinger $119, 112, 114$, hvor vi beregnede a posteriori spredningen til $s = 3.61$. Et 95% konfidensinterval for spredningen kan nu beregnes til

$$\left[\sqrt{\frac{2}{\chi^2(2)_{.975}}}3.61, \sqrt{\frac{2}{\chi^2(2)_{.025}}}3.61\right] = [1.88, 22.6],$$

da $\chi^2(2)_{.975} = 7.38$ og $\chi^2(2)_{.025} = 0.051$.

Intervallet er uhyggeligt bredt, og det minder endnu engang om, at man skal være meget forsigtig med at fæste lid til a posteriori spredninger, som er baseret på få overbestemmelser!

Bemærk, at χ^2-fordelingens asymmetri betyder, at $\hat{\sigma} = s = 3.61$ ikke ligger midt i konfidensintervallet, som det var tilfældet med konfidensintervaller for middelværdien. □

Man kan nu passende spørge, hvor mange overbestemmelser man skal have for at konfidensintervallerne bliver passende snævre.

For at give et simpelt svar på det spørgsmål, benytter vi os af at både χ^2-fordelingen og fordelingen af $\hat{\sigma} = s$ ligner normalfordelinger, når antallet af overbestemmelser er stort. Det er derfor korrekt, at et interval af formen $\hat{\sigma} \pm 2\mathrm{Spr}(\hat{\sigma})$ approximativt har 95% konfidens.

Vi mangler blot at bestemme "spredningen på spredningen": $\mathrm{Spr}(\hat{\sigma})$. Det fremgår af (10), at middelværdien i en χ^2-fordeling er $E(Y) = d$ og variansen $\mathrm{Var}(Y) = 2d$. Vi får derfor for variansen på variansestimatet at

$$\mathrm{Var}(\hat{\sigma}^2) = \mathrm{Var}(\sigma^2 Y/d) = 2d\sigma^4/d^2 = 2\sigma^4/d, \tag{11}$$

og $E(\hat{\sigma}^2) = \sigma^2$ da estimatoren er central. Altså er

$$\mathrm{Spr}(\hat{\sigma}^2) = \sqrt{\mathrm{Var}(\hat{\sigma}^2)} = \sigma^2\sqrt{2/d}.$$

Men nu var det $\mathrm{Spr}(\hat{\sigma})$, som vi var interesserede i. Da $\hat{\sigma} = \sqrt{\hat{\sigma}^2}$ kan vi bruge fejlforplantlingsloven på funktionen $f(x) = \sqrt{x}$ som har differentialkvotient $f'(x) = 1/(2\sqrt{x})$. Fra (11) får vi derfor at der approximativt gælder at

$$\mathrm{Var}(\hat{\sigma}) \approx \mathrm{Var}(\hat{\sigma}^2)/(2\sqrt{\sigma^2})^2 = \sigma^2/(2d)$$

og derfor at *spredningen på spredningen er approximativt givet ved*

$$\mathrm{Spr}(\hat{\sigma}) = \sqrt{\mathrm{Var}(\hat{\sigma})} \approx \sigma/\sqrt{2d}.$$

Den *relative spredning* eller variationskoefficienten er derfor givet som

$$\mathrm{Spr}(\hat{\sigma})/\sigma \approx 1/\sqrt{2d}.$$

Hvis vi ønsker at sikre os, at den relative spredning på spredningen ikke overstiger en given størrelse $0 < \rho < 1$ skal vi med andre ord løse ligningen

$$1/\sqrt{2d} = \rho$$

hvilket fører til at vi skal mindst skal have

$$d = 1/(2\rho^2)$$

overbestemmelser. For $\rho = 10\% = 1/10$ fås for eksempel at vi mindst skal have

$$d = 1/(2 \times 0.01) = 50$$

overbestemmelser og tilsvarende for $\rho = 5\% = 1/20$ skal vi bruge

$$d = 1/(2 \times 0.0025) = 200$$

overbestemmelser.

2.4.3 Ensidige konfidensintervaller

I alle eksemplerne ovenfor har vi valgt at konstruere intervallerne tosidigt, altså med begge konfidensgrænser C_1 og C_2 tilfældige, og vi har endda gjort det således, at sandsynlighederne for at intervallet rammer for lavt eller rammer for højt er lige store, nemlig $\alpha/2$:

$$P(C_2 < \theta) = P(\theta < C_1) = \alpha/2.$$

Somme tider kan det være mere relevant at lave et ensidigt interval, altså et interval, hvor enten c_1 (eller c_2) er så lille (eller stor) som overhovedet muligt.

Det kan være fordi vi af den ene eller anden grund primært vil sikre os ved at angive en sandsynlig øvre (eller nedre) grænse for vores vores estimat for den sande værdi af θ.

Hvis vi lader c_1 være så lille som mulig, skal vi altså bestemme en øvre konfidensgrænse, som så er givet ved x_γ-fraktilen, hvor $\gamma = 1 - \alpha$ som før. Omvendt, hvis vi vil have c_2 til at være så stor som muligt, beregnes c_1 som x_α-fraktilen. Igen vil disse fraktiler normalt afhænge af den ukendte parameter θ, men anvender vi vores pivotale kunstgreb ovenfor, fås i de tre tilfælde

Øvre konfidensinterval med kendt spredning

$$c_1 = -\infty, \quad c_2 = \bar{x} + u_{1-\alpha}\sigma/\sqrt{n}.$$

Øvre konfidensinterval med ukendt spredning

$$c_1 = -\infty, \quad c_2 = \bar{x} + t_{1-\alpha}(d)s/\sqrt{n}.$$

Øvre konfidensinterval for spredningen

$$c_1 = 0, \quad c_2 = \sqrt{\frac{d}{\chi^2(d)_\alpha}}s.$$

Bemærk, at vi skal bruge en nedre fraktil i χ^2 fordelingen for at få et øvre konfidensinterval for spredningen. Det skyldes, at uligheden i χ^2-fordelingen bliver vendt om, når vi tager reciprokværdier.

For at vise, at intervallerne har den rigtige konfidensgrad $\gamma = 1 - \alpha$ gennemføres regninger, som er helt analoge til de tidligere udførte.

Eksempel 6 Hvis vi i tilfældet med de tre højdemålinger $119, 112, 114$ ønsker at lave øvre 95% konfidensgrænser fås i de tre tilfælde

1. $\sigma = 4$ og kendt:

   $$[-\infty, 115 + 1.645 \times 4/\sqrt{3}] = [-\infty, 115 + 3.80],$$

   som kan sammenlignes med det tosidige interval 115 ± 4.53.

2. Med ukendt spredning fås

$$[-\infty, 115 + 2.92 \times 3.61/\sqrt{3}] = [-\infty, 115 + 6.09],$$

som kan sammenlignes med det ensidige 115 ± 8.97.

3. Som øvre konfidensinterval for spredningen fås tilsvarende

$$\left[0, \sqrt{\frac{2}{\chi^2(2)_{.05}}} 3.61\right] = [0, 15.94],$$

idet $\chi^2(2)_{.05} = 0.1026$. Dette kan igen sammenlignes med det tosidige interval $[1.88, 22.6]$. Selv om vi laver intervallet ensidigt, får vi en meget stor øvre konfidensgrænse for spredningen, da estimatet for denne kun er baseret på 2 overbestemmelser.

Det overlades til læseren at beregne nedre konfidensintervaller. □

2.4.4 Alternative konfidensintervaller

Alle de beregnede konfidensintervaller ovenfor er baseret på en antagelse om normalfordelte observationer. Det kan i en konkret situation være tvivlsomt.

C. F. Gauss beviste i 1823 en ulighed, som alternativt kan anvendes til at konstruere konfidensintervaller. Nærmere betegnet viste han, at hvis fordelingen af målefejlen X er symmetrisk og har en tæthedsfunktion $f_X(x)$ som er aftagende med $|x|$, så vil

$$P(|X| \leq t\sigma) \leq 1 - \frac{4}{9t^2} \text{ for } t \geq 2/\sqrt{3}.$$

I den situation har et 3σ-interval derfor mindst en konfidensgrad på

$$1 - \frac{4}{9 \times 9} = 77/81 = 95.1\%.$$

Derfor er det en god ide med 3σ-intervaller, og det er muligvis baggrunden for, at man i dag i landmålingen typisk anvender 3σ- og ikke 2σ-intervaller, som ellers vil give en konfidensgrad på 95% for normalfordelte observationer. Men 3σ-intervallernes konfidensgrad når næppe op på de ca. 99%, som man

ofte regner med. Denne konfidensgrad opnås kun, når fejlene er normalfordelte.

Hvis man ikke engang kan regne med, at fejlfordelingen er symmetrisk og har aftagende tæthedsfunktion, kan man lave konfidensintervaller ved hjælp af Chebyshevs ulighed

$$P(|X| \leq t\sigma) > 1 - t^{-2},$$

som gælder i alle tilfælde, hvor fordelingen har en middelværdi og varians. Det følger, at et 5σ-interval mindst har en konfidensgrad på

$$1 - 1/25 = 24/25 = 96\%,$$

men så er det også blevet noget bredt efterhånden...

Til sidst vil vi som et kuriosum nævne, at det for en symmetrisk fejlfordeling gælder, at intervallet $[x_{(1)}, x_{(n)}]$ fra den mindste til den største observation vil have en konfidensgrad på $1 - 2^{-n+1}$, så man allerede ved $n = 6$ observationer har en konfidensgrad på

$$1 - 1/32 = 31/32 = 96.8\%.$$

Ved flere observationer stiger konfidensgraden naturligvis, men det gør intervallets bredde også!

3 Testteori

3.1 Idéen bag det statistiske test

I et statistisk test konfronteres en bestemt *hypotese* med virkeligheden i form af observationer $x_1, \ldots, x_n$ af stokastiske variable.

Hypotesen betegnes traditionelt med H_0, og derfor bruger man også udtrykket "nulhypotesen".

Den undersøgte hypotese kan være af videnskabelig natur eller af kontrollerende natur.

Det er vigtigt at gøre sig klart, at hypoteser kun kan *falsificeres* (forkastes) ved statistiske test, aldrig bekræftes. Derfor må man somme tider formulere hypoteserne lidt 'bagvendt' i forhold til normal sprogbrug.

Som eksempler på videnskabelige hypoteser kan nævnes:

- Himmellegemerne bevæger sig med Jorden i centrum (Galileo Galilei).
- Lyset bevæger sig uendelig hurtigt (Ole Rømer).
- Jorden er kugleformet (P. S. Laplace).

Videnskabelige hypoteser er ofte formuleret med henblik på falsificering, og det videnskabelige fremskridt fremkommer netop ved at hypotesen falsificeres.

Galilei udførte verdenshistoriens første statistiske test, idet han måtte forkaste den herskende hypotese, da den blev konfronteret med hans astronomiske observationer.

Laplace havde af en eller anden grund fået den ide, at Jorden var pæreformet. Det lykkedes ham aldrig at falsificere hypotesen om Jordens kugleform, selv om satellit- og andre målinger senere har bekræftet, at han havde ret!

Nedenfor angives en række eksempler på kontrollerende hypoteser:

- En måleserie er foretaget med forventet præcision (global test).
- En bestemt måling er ikke behæftet med grov fejl (data snooping).
- Et objekt har ikke flyttet sig.
- Et objekt er ikke deformeret.
- Et objekt tilfredsstiller bestemte specifikationer.

Kontrollerende hypoteser er mest almindelige i landmålingen. Derfor vælger vi her at fremstille testteorien på en sådan måde, at det kontrollerende synspunkt er i centrum.

Ved en kontrollerende hypotese går man egentlig ud fra, at den formulerede hypotese er korrekt, men ønsker at den skal underkastes en kontrol, så man bedre kan stole på den.

Eksempel 7 Vi betragter endnu engang vores lille eksempel med tre højdemålinger $119, 112, 114$.

Her er en række hypoteser, som kunne være relevante i forbindelse med sådan en måleserie.

- Spredningen på måleinstrumentet er $2mm$.
- Højdeforskellen er $111mm$ som specificeret i landskabsplanen.
- For et år siden måltes samme højdeforskel til $110, 112, 109$. Har noget flyttet sig? Hypotese: Intet har flyttet sig.
- Er 119 en grov fejl? Hypotese: den første måling er ikke en grov fejl.

Bemærk, at alle hypoteserne er kontrollerende. Vi skal udføre det statistiske test for at sikre os, at hypoteserne kan stå for en nærmere prøvelse.

Bemærk også, at i de to sidste tilfælde formulerer vi hypotesen 'bagvendt' i forhold til den måde, vi normalt udtrykker os på. Hypotesen skal formuleres, så den er præcis og egnet til falsificering. □

3.2 Konstruktion og udførelse af et statistisk test

Første trin ved konstruktion af et statistisk test er, at man vælger en *testvariabel* $W = g(X_1, \ldots, X_n)$ (*eng.* test statistic), som skal afsløre afvigelser fra hypotesen. Man kan i nogen grad opstille formelle principper for dette valg, men det fører ikke altid til praktisk tilfredsstillende løsninger. Her nøjes vi derfor med at fastslå, at testvariablen skal være velegnet til at afsløre afvigelser fra hypotesen.

I næste trin afgøres, om store eller små værdier af W (eller både store og små værdier) bør betragtes som *kritiske* for hypotesen, altså som indikation på, at hypotesen ikke kan opretholdes.

For at afgøre, hvilke værdier, der skal betragtes som kritiske, formuleres ofte en *alternativ hypotese* H_A, og den alternative hypotese kan også bruges til vejledning, når testvariablen skal vælges, idet den angiver, hvilke afvigelser, man især er på vagt overfor.

Der vælges eventuelt på forhånd et *signifikansniveau* α. Eksempelvis bruges $\alpha = 5\%$, men ofte andre, alt efter situationen.

Udfra testvariablen og signifikansniveauet findes nu et *kritisk område,* K_α. Hypotesen H_0 betragtes derefter som (falsificeret) *forkastet på niveau* α*,* hvis den observerde værdi $W = w_{\text{obs}}$ ligger i det kritiske område. I modsat fald siges hypotesen H_0 at være *accepteret.*

Hvis store værdier af W er kritiske for hypotesen bestemmes det kritiske område som

$$K_\alpha = \{w \mid w > w_c\},$$

hvor w_c kaldes *den kritiske værdi.* Den kritiske værdi skal bestemmes således, at der gælder

$$P(W \in K_\alpha \mid H_0) = \alpha. \tag{12}$$

Dermed angiver signifikansniveuaet α sandsynligheden for at forkaste en korrekt nulhypotese. Man taler også om, at α angiver sandsynligheden for at begå en *fejl af type I.* Naturligvis ønsker man derfor, at α skal være lille.

Komplementærmængden til det kritiske område K_α kaldes også *accept-området* for det givne test.

Ved videnskabelige hypoteser kan det virke unaturligt at vælge signifikansniveau på forhånd (hvilket niveau skulle Galilei have valgt?), men især ved kontrollerende hypoteser kan det have en vigtig, standardiserende funktion.

Bemærk, at vi kun kan bestemme det kritiske område efter (12) og dermed konstruere et test med veldefineret signifikansniveau, hvis sandsynligheden for, at W ligger i det kritiske område, ikke afhænger af en ukendt parameter θ under antagelse af nulhypotesen. Derfor vil testvariablen W ofte blive konstrueret udfra en pivotal størrelse, som det var tilfældet, når vi skulle konstruere konfidensintervaller.

Eksempel 8 Vi vender tilbage til vores tre højdemålinger $119, 112, 114$ og ønsker at efterprøve, om hypotesen $H_0 : \sigma_0 = 2$ kan opretholdes i lyset af disse observationer.

Som testvariabel vælger vi $Y = dS^2/\sigma_0^2 = S^2/2$, med den observerede værdi $y_{\text{obs}} = 3.61^2/2 = 13/2 = 6.5$.

Vi betragter store værdier som kritiske, idet vi kun er interesseret i at afsløre, om målingernes kvalitet er ringere end vi forventer. Vores alternative hypotese er altså $H_A : \sigma > \sigma_0 = 2$.

Hvis nulhypotesen er rigtig, følger Y en χ^2-fordeling med $d = 2$ frihedsgrader. Den kritiske værdi skal derfor bestemmes som $(1-\alpha)$-fraktilen i denne fordeling. Ved et signifikansniveau 5% fås

$$y_c = \chi^2(2)_{0.95} = 5.99.$$

Hypotesen forkastes altså på niveau 5%, idet den observerede testvariabel er større end den kritiske værdi. Vi konkluderer derfor, at målingernes kvalitet er ringere end forventet. Formelt konkluderer vi måske $H_A : \sigma > \sigma_0 = 2$, men reelt er værdien 119 måske fremkommet ved en grov fejl, og så er det måske et andet alternativ, der er gældende? □

Hvis små værdier af W skal betragtes som kritiske, skal det kritiske område konstrueres som et område af formen

$$K_\alpha = \{w \,|\, w < w_c\},$$

men ellers er fremgangsmåden analog.

I mange situationer er det naturligt at betragte både store og små værdier af W som kritiske. Det kritiske område skal derfor have formen

$$K_\alpha = \{w \,|\, \underline{w}_c < w\} \cup \{w \,|\, w < \overline{w}_c\},$$

så både en *nedre kritisk værdi* $\underline{w}_c$ og en *øvre kritisk værdi* $\overline{w}_c$ skal bestemmes. Det er nu et problem, at der bliver to ubekendte i ligningen (12), så den ikke længere har en entydig løsning. Konventionelt vælger man ofte grænserne symmetrisk, så de to dele af det kritiske område hver har en sandsynlighed på $\alpha/2$:

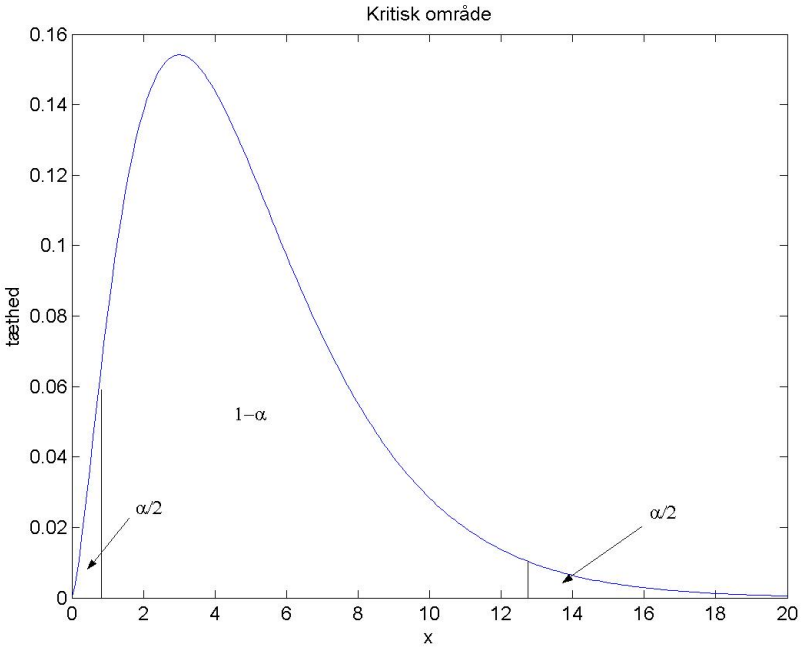
$$P(\underline{w}_c < W) = P(W < \overline{w}_c\} = \alpha/2,$$

se illustrationen i Figur 4. I det følgende vil vi se, hvorledes testvariable og kritiske områder konstrueres i en række typiske situationer.

Eksempel 9 I eksemplet fra før kan det være naturligt at udføre testet tosidigt, svarende til en mere diffus alternativ hypotese $H_A : \sigma \neq \sigma_0 = 2$. Vi skal så finde den nedre kritiske værdi som $\alpha/2$-fraktilen og den øvre kritiske værdi som $(1 - \alpha/2)$-fraktilen. Det giver

$$\underline{w}_c = \chi^2(2)_{0.025} = 0.0506, \quad \overline{w}_c = \chi^2(2)_{0.975} = 7.378.$$

Da vi observerede $w_{obs} = 13/2 = 6.5$ og denne værdi ligger i acceptområdet $[0.0506, 7.378]$, accepterer vi altså i dette tilfælde nulhypotesen på niveau $\alpha = 5\%$. □



Figur 4: Kritisk område for et tosidigt test. Hver af de to haler i fordelingen har sandsynlighed $\alpha/2$.

3.3 Test i en enkelt normalfordelt stikprøve

3.3.1 Test for bestemt varians

Antag at $X_1, \ldots, X_n$ er en stikprøve fra $\mathcal{N}(\mu, \sigma^2)$, hvor μ og σ er ukendte. Hvis vi ønsker at teste $H_0 : \sigma^2 = \sigma_0^2$, hvor σ_0^2 er kendt, bruges testvariablen

$$Y = \frac{(n-1)S^2}{\sigma_0^2}.$$

I landmålingens normale sprogbrug benyttes ofte udtrykket *a priori spredning* om σ_0 og *a posteriori spredning* om S. Under H_0 haves altså at $Y \sim \chi^2(n-1)$.

Hvis $H_A : \sigma^2 \neq \sigma_0^2$ er alternativet, vil både store og små værdier af Y være kritiske, det vil sige, at testet skal være tosidigt.

Hvis y_{obs} er den observerede værdi, og F_{n-1} er fordelingsfunktionen for $\chi^2(n-1)$-fordelingen, er testets acceptområde altså givet ved

$$[\chi^2(n-1)_{\alpha/2}, \chi^2(n-1)_{1-\alpha/2}].$$

Dette test omtales ofte som χ^2-*test* på grund af testvariablens fordeling. Vi har allerede set på eksempler på dette test ovenfor i eksempel 8 og 9.

I forbindelse med en mere generel udjævning, anvendes χ^2-testet også til et såkaldt "globalt test" (Cederholm 2000), hvor testvariablen er givet som

$$Y = \frac{\hat{d}^\top C \hat{r}}{\sigma_0^2},$$

den *skalerede residualkvadratsum.* Her er d antallet af overbestemmelser og testvariablen er også her χ^2-fordelt med d frihedsgrader.

I Cederholm (2000) foreslås et tosidigt alternativ, men hvis formålet primært er at sikre sig imod grove målefejl, kan det synes mere hensigtsmæssigt med et ensidigt test, så kun store værdier anses for kritiske. Et argumentet for også at anse små værdier for kritiske kunne være, at en særlig lille værdi af testvariablen tyder på, at målingerne er udført på en måde, så deres fejl er korrelerede, og man også ønsker at kunne afsløre den slags fejl ved hjælp af testet.

3.3.2 Test for bestemt middelværdi. Kendt varians

Antag $X_1, \ldots, X_n$ er en stikprøve fra $\mathcal{N}(\mu, \sigma_0^2)$, hvor μ er ukendt og σ_0^2 er kendt. Vi minder da om, at $\bar{X}$ er estimatet for μ og $\bar{X} \sim \mathcal{N}(\mu, \sigma_0^2/n)$. Hvis vi ønsker at teste $H_0 : \mu = \mu_0$, hvor μ_0 er kendt, bruges testvariablen

$$U = \frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma_0}$$

Under H_0 ses let at $U \sim \mathcal{N}(0, 1)$ (heraf navnet U-fordeling), og hvis $H_A : \mu \neq \mu_0$ er alternativet, vil både store og små værdier af U være kritiske, det vil sige, at testet skal være tosidigt. Med andre ord bliver testets acceptområde:

$$[u_{\alpha/2}, u_{1-\alpha/2}].$$

Testet omtales normalt som *u-test.*

Eksempel 10 I vores standardeksempel med målinger $119, 112, 114$ kunne vi være interesseret i at undersøge, om højden overholder en specifikation på

111 mm, altså i nulhypotesen $H_0 : \mu = 111$, og vi har et tosidigt alternativ $H_A : \mu \neq 111$.

Hvis man kan stole på a priori spredningen $\sigma_0 = 2$, anvendes u-test. Den tilhørende testvariabel fås til

$$U = \sqrt{n}\frac{\bar{X} - 111}{\sigma_0}; \quad u_{\text{obs}} = \sqrt{3}\frac{115 - 111}{2} = 3.464.$$

Ved et signifikansniveau på 5% findes den nedre og øvre kritiske værdi til at være hhv. -1.96 og 1.96. Da den observerede værdi er klart større end den øvre kritiske grænse, kan hypotesen ikke opretholdes med det givne signifikansniveau. □

3.3.3 Test for bestemt middelværdi. Ukendt varians

Vi betragter stadig situationen, hvor vi har en stikprøve $x_1, \ldots, x_n$ af normalfordelte observationer og ønsker at teste hypotesen $H_0 : \mu = \mu_0$ mod alternativet $H_A : \mu \neq \mu_0$.

Hvis vi ikke kan stole på a priori spredningen anvendes i stedet a posteriori spredningen og testvariablen

$$T = \frac{\sqrt{n}(\bar{X} - \mu_0)}{s}.$$

Da denne er t-fordelt med $d = n-1$ frihedsgrader bliver testets acceptområde nu i stedet

$$[t(n-1)_{\alpha/2}, t(n-1)_{1-\alpha/2}].$$

Dette test omtales også som et t-test.

Eksempel 11 Hvis vi i vores eksempel ikke tror på a priori spredningen $\sigma_0 = 2$, anvender vi a posteriori spredningen og t-fordelingen: Vi får så værdien af vores testvariabel til

$$T = \sqrt{n}\frac{\bar{X} - 111}{s}; \quad t_{\text{obs}} = \sqrt{3}\frac{115 - 111}{3.606} = 1.922.$$

Testets acceptområde er givet som

$$[t(2)_{0.025}, t(2)_{0.975}] = [-4.3027, 4.3027].$$

Da den observerede værdi af testvariablen ligger i acceptområdet, godkender vi her nulhypotesen på niveau 5%. Bemærk, at konklusionen er den stik modsatte af, hvad vi kom frem til i det tilfælde, hvor vi brugte a priori variansen i eksempel 10. □

3.4 Test i to eller flere normalfordelte stikprøver

3.4.1 Sammenligning af to varianser

Det kan ofte være relevant at sammenligne to uafhængige måleserier for at undersøge, om de to måleserier er udført med samme præcision.

Vi antager, at vi fra udjævning af to uafhængige måleserier har fundet a posteriori spredninger s_1 og s_2 som er baseret på hhv. d_1 og d_2 overbestemmelser. Idet vi lader σ_1 og σ_2 betegne de teoretiske spredninger hørende til de to måleserier, ønsker vi nu at undersøge hypotesen $H_0 : \sigma_1 = \sigma_2$ om, at de to måleserier er udført med samme nøjagtighed.

Som testvariabel anvender vi forholdet mellem a posteriori varianserne

$$V = S_1^2/S_2^2.$$

Under antagelse af nulhypotesen er dette forhold F-fordelt med (d_1, d_2) frihedsgrader, se afsnit 1.4.3.

Hvis alternativet blot er $H_A : \sigma_1 \neq \sigma_2$, anvendes et tosidigt test, og med et signifikansniveau α bliver testets acceptområde

$$[\underline{v}_c, \overline{v}_c] = [v(d_1, d_2)_{\alpha/2}, v(d_1, d_2)_{1-\alpha/2}],$$

hvor $v(d_1, d_2)_\beta$ betegner β-fraktilen i F-fordelingen med (d_1, d_2) frihedsgrader.

Dette test omtales normalt som et F-test, på grund af testvariablens fordeling.

Eksempel 12 Vi ser igen på vores måleserie $119, 112, 114$, men ønsker nu at sammenligne den med en måleserie lavet på et tidligere tidspunkt, hvor resultaterne blev $110, 112, 109, 114$. Er disse to måleserier foretaget med samme præcision?

Vores testvariabel $V = S_1^2/S_2^2$ beregnes til

$$v_{\text{obs}} = 13/4.9167 = 2.6441,$$

og med et signifikansniveau på 5% bliver testets acceptområde

$$[v(2,3)_{0.025}, v(2,3)_{0.0975}] = [0.0255, 16.04],$$

så det kan ikke dokumenteres, at spredningen er ændret. Bemærk igen, at med så få overbestemmelser er acceptområdet meget bredt. Vi skal observere en forøgelse af spredningen på mere end 4 gange (16 gange for variansen) før ændringen bliver signifikant. □

3.4.2 Sammenligning af middelværdier. Kendte varianser

Vi antager, at vi har to normalfordelte stikprøver $X_1, \ldots, X_m$ og $Y_1, \ldots, Y_n$, hvor $X_i \sim \mathcal{N}(\mu_1, \sigma_1^2)$ og $Y_j \sim \mathcal{N}(\mu_2, \sigma_2^2)$, hvor spredningerne σ_1 og σ_2 antages at være kendte. Vi ønsker at undersøge, om de to middelværdier er ens og vil derfor teste hypotesen $H_0 : \mu_1 = \mu_2$.

Det er naturligt at forsøge at konstruere en testvariabel på baggrund af forskellen $\bar{X} - \bar{Y}$ på de to gennemsnit. Under antagelse af, at H_0 er korrekt, er $\bar{X} - \bar{Y} \sim \mathcal{N}(0, \sigma_1^2/m + \sigma_2^2/n)$, så testvariablen

$$U = \frac{\bar{X} - \bar{Y}}{\sqrt{\sigma_1^2/m + \sigma_2^2/n}} \tag{13}$$

er U-fordelt, og hvis alternativet er tosidigt $H_A : \mu_1 \neq \mu_2$, er acceptområdet for et test med signifikansniveau α derfor givet ved $[u_{\alpha/2}, u_{1-\alpha/2}]$.

Eksempel 13 Vi ser igen på de to måleserier fra før, hvor vi i den ene serie fik observationerne $119, 112, 114$ og i den anden $110, 112, 109, 114$. Nu ønsker vi at undersøge, om middelværdien af de to serier er ens, enten for at sikre os, at objektets højde ikke er ændret, eller for at sikre, at en af serierne ikke er behæftet med systematiske fejl.

Hvis vi antager, at de to spredninger er kendte med hhv. $\sigma_1 = 3$ og $\sigma_2 = 2$, fås testvariablen til

$$u_{\text{obs}} = \frac{\bar{x} - \bar{y}}{\sqrt{\sigma_1^2/m + \sigma_2^2/n}} = \frac{115 - 111.25}{\sqrt{9/3 + 4/4}} = 1.875.$$

Ved et signifikansniveau på $\alpha = 5\%$ falder denne værdi indenfor acceptområdet $[-1.96, 1.96]$, så vi godkender hypotesen om, at middelværdien ikke har ændret sig. □

3.4.3 Sammenligning af middelværdier. Ukendt fælles varians

Vi antager, at vi har to normalfordelte stikprøver $X_1, \ldots, X_m$ og $Y_1, \ldots, Y_n$, hvor $X_i \sim \mathcal{N}(\mu_1, \sigma_1^2)$ og $Y_j \sim \mathcal{N}(\mu_2, \sigma_2^2)$, men antager nu, at spredningerne σ_1 og σ_2 er ukendte, men at vi til gengæld kan antage at de er ens. Vores residualkvadratsummer $(m-1)S_1^2$ og $(n-1)S_2^2$ fra de to stikprøver er χ^2-fordelte med hhv. $m-1$ og $n-1$ frihedsgrader og kan derfor kombineres til et fælles variansestimat

$$S^2 = \frac{(m-1)S_1^2 + (n-1)S_2^2}{m+n-2},$$

hvor så $(m+n-2)S^2$ er χ^2-fordelt med $m+n-2$ frihedsgrader. Testvariablen for hypotesen $H_0 : \mu_1 = \mu_2$ kan nu konstrueres som

$$T = \frac{\bar{X} - \bar{Y}}{S\sqrt{1/m + 1/n}},$$

som er t-fordelt med $m+n-2$ frihedsgrader. Hvis alternativet er tosidigt $H_A : \mu_1 \neq \mu_2$, er acceptområdet for et test med signifikansniveau α derfor givet ved $[t(m+n-2)_{\alpha/2}, t(m+n-2)_{1-\alpha/2}]$.

Eksempel 14 Vi ser igen på de to måleserier fra før, hvor vi i den ene serie fik observationerne $119, 112, 114$ og i den anden $110, 112, 109, 114$, men antager nu at to spredninger er ukendte men ens. Det fælles variansestimat bliver

$$s^2 = (2 \times 13 + 3 \times 4.9167)/5 = 8.15,$$

så $s = \sqrt{8.15} = 2.855$ og testvariablen derfor bliver

$$t_{\text{obs}} = \frac{\bar{x} - \bar{y}}{s\sqrt{1/m + 1/n}} = \frac{115 - 111.25}{2.855\sqrt{1/3 + 1/4}} = 1.72.$$

Ved et signifikansniveau på $\alpha = 5\%$ falder denne værdi indenfor acceptområdet $[-2.57, 2.57]$, så vi godkender hypotesen om, at middelværdien ikke har ændret sig. □

3.4.4 Sammenligning af middelværdier. Ukendte varianser

Hvis varianserne ikke er ens, optræder der et såkaldt Behrens–Fisher problem, idet man ikke kan lave et test med et fast, veldefineret signifikansniveau. I praksis kan man derimod godt anvende testvariablen

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{S_1^2/m + S_2^2/n}},$$

som er konstrueret udfra (13) ved at erstatte de teoretiske varianser med de empiriske.

Fordelingen af Z kan ikke beregnes explicit, men Z kan dog approximativt antages at være t-fordelt med q frihedsgrader, hvor q skal bestemmes som

$$q = 1/\{a^2/(m-1) + (1-a)^2/(n-1)\},$$

hvor

$$a = \frac{s_1^2/m}{s_1^2/m + s_2^2/n}.$$

Eksempel 15 Vi ser på måleserierne $119, 112, 114$ og $110, 112, 109, 114$; men antager nu at to spredninger er helt ukendte. Testvariablen bliver

$$z_{\text{obs}} = \frac{\bar{x} - \bar{y}}{\sqrt{s_1^2/m + s_2^2/n}} = \frac{115 - 111.25}{\sqrt{13/3 + 4.9167/4}} = 1.59.$$

Frihedsgraderne for den t-fordeling, vi skal anvende, beregnes som

$$a = (13/3)(13/3{+}4.9167/4) = 0.779, \quad q = 1/(0.779^2/2{+}0.221^2/3) = 3.1279.$$

Bemærk, at frihedsgraderne q for t-fordelingen ikke er heltallige. Ved et signifikansniveau på $\alpha = 5\%$ falder den observerede værdi af testvariablen $z_{\text{obs}} = 1.59$ indenfor acceptområdet, som er givet ved

$$[t(3.1279)_{0.025}, t(3.1279)_{0.975}] = [-3.11, 3.11],$$

så vi godkender også her hypotesen om, at middelværdien ikke har ændret sig. □

3.4.5 Sammenligning af flere varianser

Vi betrager nu den situation, hvor varianser fra flere forskellige måleserier skal sammenlignes på en gang.

Vi antager derfor, at $s_1^2, \ldots, s_k^2$ er uafhængige a posteriori varianser, det vil sige, at $d_i S_i^2/\sigma_i^2 \sim \chi^2(d_i)$ for $i = 1, \ldots, k$. Hvis vi ønsker at teste

$$H_0 : \sigma_1 = \cdots = \sigma_k = \sigma,$$

hvor σ er ukendt, mod det alternativ at mindst to af disse spredninger er forskellige, bruges testvariablen

$$B = c^{-1}(d \ \log S^2 - \sum_{i=1}^{k} d_i \ \log S_i^2)$$

hvor log betegner logaritmen med grundtal e,

$$d = \sum_{i=1}^{k} d_i, \quad c = 1 + \frac{1}{3(k-1)}(\sum_{i=1}^{k} d_i^{-1} - d^{-1}),$$

og

$$S^2 = d^{-1} \sum_{i=1}^{k} d_i S_i^2$$

er et estimat for den fælles varians σ^2.

Fordelingen af B under H_0 er svær at bestemme, men den kan approksimeres med en $\chi^2(k-1)$-fordeling. Approksimationen er god, hvis $d_i \geq 2$ for alle $i = 1, \ldots, k$. Her vil man kun anse store værdier af B for at være kritiske, og derfor bestemmes den øvre kritiske grænse for et test med signifikansniveau α som

$$\overline{b}_c = \chi^2(k-1)_{1-\alpha}.$$

Approksimationen til B's fordeling er fundet af den engelske statistiker M. S. Bartlett i 1937, og testet omtales derfor normalt som *Bartlett's test.*

Eksempel 16 Vi ser igen på vores måleserier 119, 112, 114 og 110, 112, 109, 114, men tænker os nu, at vi yderligere har en måleserie 108, 110, 112, 109, 108. Er disse tre måleserier udført med samme præcision?

Vi anvender Bartlett's test og beregner derfor

$$d = 2 + 3 + 4 = 9, \quad c = 1 + \frac{1}{6}(1/2 + 1/3 + 1/4 - 1/9) = 1.1620,$$

og videre

$$s^2 = (2 \times 13 + 3 \times 4.9167 + 4 \times 2.8)/9 = 5.7722,$$

så den observerede værdi af testvariablen B bliver

$$b_{\text{obs}} = (9 \log 5.7722 - 2 \log 13 - 3 \log 4.9167 - 4 \log 2.8)/1.162 = 1.507,$$

og da den øvre kritiske grænse ved et signifikansniveau på 5% er

$$\chi^2(2)_{0.95} = 5.99,$$

accepteres hypotesen klart på niveau 5%. Der er altså ingen grund til at tro, at måleserierne er udført med forskellig præcision. Men igen må vi tage det forbehold, at der er meget få overbestemmelser, så forskellene i præcision skal være meget markante for at vi kan identificere dem. □

3.5 Test og konfidensintervaller

Der er en snæver forbindelse mellem konstruktion af test og konstruktion af konfidensintervaller. Faktisk gælder det som regel, at man udfra et test med signifikansniveau α på naturlig måde kan konstruere et konfidensområde med konfidensgrad $1-\alpha$ for den ukendte parameter som indgår i testet. Omvendt kan man udfra et konfidensområde med konfidensgrad γ konstruere et test med signifikansniveau $1 - \alpha$.

Forbindelsen er, at en hypotese af formen $H_0 : \theta = \theta_0$ accepteres hvis og kun hvis værdien θ_0 ligger i det observerede konfidensområde for θ, og konfidensområdet for θ består af alle de værdier θ_0, hvor hypotesen $H_0 : \theta = \theta_0$ kan accepteres.

Hvis θ er en endimensional parameter, vil der typisk være tale om et konfidensinterval, men hvis θ er flerdimensional, kan konfidensområdet få en anden form. For todimensionale parametre fås for eksempel ofte konfidensellipser, se afsnit 4.2, som netop er konstrueret på denne måde udfra et tilsvarende test.

Denne sammenhæng beror på, at udgangspunktet i begge tilfælde er en pivotal størrelse

$$W = w(X_1, \ldots, X_n, \theta),$$

hvis fordeling ikke afhænger af den ukendte parameter. Ofte vil denne ligning (14) kunne løses med hensyn til θ, så vi kan skrive

$$W = w(x_1, \ldots, x_n, \theta) \iff \theta = h(x_1, \ldots, x_n, w), \tag{14}$$

hvor h er en voksende funktion af w.

Vi danner nu et test for hypotesen $H_0 : \theta = \theta_0$ med signifikansniveau α ved at bruge $W_0 = w(x_1, \ldots, x_n, \theta_0)$ som testvariabel og området

$$A_\alpha = [w_1, w_2]$$

som acceptområde, hvor grænserne (w_1, w_2) er valgt så

$$P(w_1 < w(X_1, \ldots, X_n, \theta) < w_2) = \gamma = 1 - \alpha. \tag{15}$$

Tilsvarende bliver konfidensintervallet for θ dannet ved at anvende (14) til at isolere θ i uligheden (15), så vi får

$$P(h(X_1, \ldots, X_n, w_1) < \theta < h(X_1, \ldots, X_n, w_1)) = \gamma,$$

og dermed får at intervallet

$$[C_1, C_2] = [h(X_1, \ldots, X_n, w_1), < h(X_1, \ldots, X_n, w_1)]$$

netop har konfidensgrad $\gamma = 1 - \alpha$.

Eksempel 17 Betragt igen observationerne $119, 112, 114$ og hypotesen $H_0 : \mu = \mu_0 = 111$ i tilfældet, hvor spredningen σ antages ukendt.

Den pivotale størrelse, som bruges både til konfidensinterval og test er her

$$T = \sqrt{n}\frac{\mu - \bar{X}}{s} = \sqrt{3}\frac{\mu - 115}{3.61}. \tag{16}$$

Denne er t-fordelt med 2 frihedsgrader, altså afhænger fordelingen ikke af de ukendte parametre μ og σ. Vi får så grænserne (t_1, t_2) som

$$(t(2)_{0.025}, t(2) =_{0.975}) = (-4.30, 4.30).$$

Dette bliver så acceptområde for testet med testvariabel

$$T = \sqrt{3}\frac{115 - \mu_0}{3.61} = \sqrt{3}\frac{4}{3.61} = 1.92$$

Løser vi ligningen (16) med hensyn til μ fås

$$\mu = \bar{x} + ts/\sqrt{n} = 115 + t \times 3.61/\sqrt{3},$$

så det tilsvarende konfidensinterval bliver

$$115 \pm 4.30 \times 3.61 = 115 \pm 8.97.$$

Vi ser, at testvariablen ligger i acceptområdet og tilsvarende, at den værdi af μ_0, som vi tester — 111 — ligger i konfidensintervallet. □

3.6 Styrkefunktion

Vi har i en række eksempler ovenfor bemærket, at en hypotese, som ikke kunne forkastes ved et statistisk test, måske alligevel kunne være tvivlsom, men at vi ikke havde tilstrækkeligt mange eller tilstrækkeligt nøjagtige observationer til at forkaste hypotesen på det givne signifikansniveau.

Dette problem kan belyses yderligere ved at undersøge testets *styrke,* et begreb, som vi skal forklare nærmere nedenfor. Vi har tidligere nævnt, at hvis man forkaster en hypotese H_0, som er sand, begår man en *fejl af type I.* Men man kan naturligvis også lave en anden slags fejl. Hvis man godkender en forkert hypotese, siger vi, at man begår en *fejl af type II.*

Ved et test på signifikansniveau α er sandsynligheden for at begå fejl af type I lig med α. Det kritiske område K_α er jo bestemt så der gælder

$$P(W \in K_\alpha \,|\, H_0) = \alpha,$$

hvor W er den anvendte testvariabel, og man begår netop en fejl af type I hvis H_0 er sand og $W \in K_\alpha$.

Tilsvarende begår vi en fejl at type II hvis H_0 er forkert og $W \notin K_\alpha$. Et test har stor styrke, hvis sandsynligheden for at begå fejl af type II er lille. Styrken betegnes ofte med β:

$$\beta = P(W \in K_\alpha \,|\, H_A),$$

og sandsynligheden for at begå type II fejl er så $1-\beta$.

Typisk vil β *afhænge af størrelsen* på afvigelsen fra H_0 og derfor er udtrykket ovenfor ikke umiddelbart veldefineret.

Ofte findes en parameter δ, som angiver afvigelsens størrelse og derved bestemmer styrken. Funktionen $\beta(\delta)$

$$\beta(\delta) = P(W \in K_\alpha \,|\, H_A(\delta)),$$

hvor $H_A(\delta)$ præciserer alternativet H_A nøjere, kaldes da *styrkefunktionen.*

Idet vi som regel vælger δ således, at $H_A(0) = H_0$, har vi nu, at styrke og signifikansniveau er relateret ved

$$\alpha = \beta(0) = P(W \in K_\alpha \,|\, H_A(0)) = P(W \in K_\alpha \,|\, H_0).$$

Eksempel 18 Vi har højdemålinger $119, 112, 114$ og antager nu, at spredningen vides at være $\sigma = 3$ og ønsker at teste, om det kan antages at $\mu = 111$.

Et test på signifikansniveau 1% fås ved at forkaste, når $\bar{x}(= 115)$ ikke ligger i intervallet

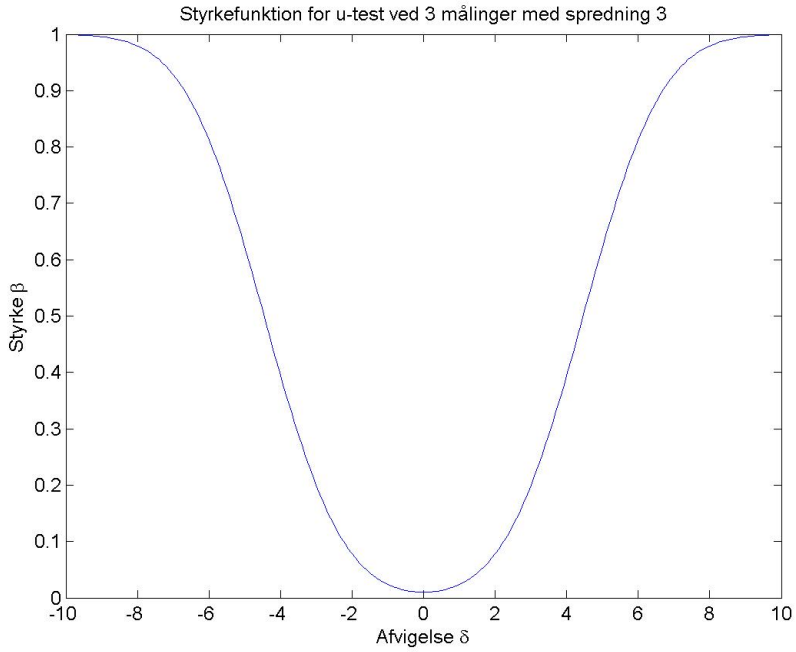
$$\begin{aligned}&[111 - u_{0.995} * 3/\sqrt{3}, 111 + u_{0.995} * 3/\sqrt{3}] = \\ &\quad [111 - 2.576 * 3/\sqrt{3}, 111 + 2.57632/\sqrt{3}] = [106.54, 115.46].\end{aligned}$$

Vi accepterer altså hypotesen på niveau 1%.

Men hvor stor en afvigelse fra $\mu = 111$ skal der egentlig til, for at vi kan regne med at afsløre den på dette signifikansniveau med kun tre observationer og $\sigma = 3$? Vi beregner styrken for $\mu = 111 + \delta$ til

$$\begin{aligned}\beta(\delta) &= 1 - P(106.54 \leq \bar{X} \leq 115.46 \,|\, h = 111 + \delta) \\ &= 1 - \Phi\left(\frac{115.46 - 111 - \delta}{3\sqrt{1/3}}\right) + \Phi\left(\frac{106.54 - 111 - \delta}{3\sqrt{1/3}}\right) \\ &= 1 - \Phi\left((4.46 - \delta)/\sqrt{3}\right) + \Phi\left((-4.46 - \delta)/\sqrt{3}\right).\end{aligned}$$

Denne styrkefunktion er afbildet i Figur 5. Vi ser på figuren, at med kun tre målinger og den angivne spredning på $\sigma = 3$, skal afvigelsen være omkring $7mm$ for at man kan opdage den med 95% sandsynlighed. □



Figur 5: Styrkefunktionen ved tre højdemålinger med spredning $\sigma = 3$ og et signifikansniveau på 1%.

3.6.1 Styrkefunktionen for χ^2-fordelte testvariable

Vi tænker os en testvariabel af typen

$$Y = \frac{r^\top C r}{\sigma_0^2},$$

for eksempel anvendt i forbindelse med et globalt test. Med d overbestemmelser er denne χ^2-fordelt med d frihedsgrader. I et test for hypotesen $H_0 : E(R) = 0$ er det kritiske område af formen $y_{obs} > y_0 = \chi^2(d)_{1-\alpha}$, og testets styrkefunktion er givet ved sandsynligheden for at forkaste hypotesen

$$\beta(\xi) = P\{Y > y_0 \,|\, E(R) = \xi\}.$$

Som vi ser, afhænger styrken af testet umiddelbart af hele middelværdivektoren ξ for den alternative hypotese.

I dette tilfælde kan det dog vises, at χ^2-testets styrke $\beta(\xi)$ kun afhænger af afvigelsens længde δ, målt med vægtmatricen:

$$\delta = \xi^\top C \xi.$$

idet Y, når $\xi \neq 0$, følger en *ikke-central χ^2-fordeling* med d frihedsgrader og *ikke-centralitetsparameter* δ.

Tæthedsfunktionen for denne fordelingen er givet som

$$f(y \mid d; \delta) = \sum_{j=0}^{\infty} \frac{e^{-\delta/2}\delta^j}{2^j j!} \frac{e^{-y/2} y^{d/2+j-1}}{2^{d/2+j}\Gamma(\frac{d}{2}+j)}.$$

Fordelingen er nærmere beskrevet side 53-57 i Pearson and Hartley (1972) og tabelleret side 232–250 i samme reference. Den findes også som standard-funktion i en del statistisk software, men den er ikke så almindelig som de andre funktioner, vi ellers har stødt på.

Eksempel 19 Vi vender tilbage til vores eksempel, hvor vi har målinger $119, 112, 114$ målt med en spredning på $\sigma_0 = 3$ og mistænker den første observation for at være en grov fejl. I dette eksempel er altså $\xi = (\lambda, 0, 0)$ og derfor er $\delta = \lambda^2/9$. Den tilsvarende styrkefunktion er afbildet i Figur 6. □

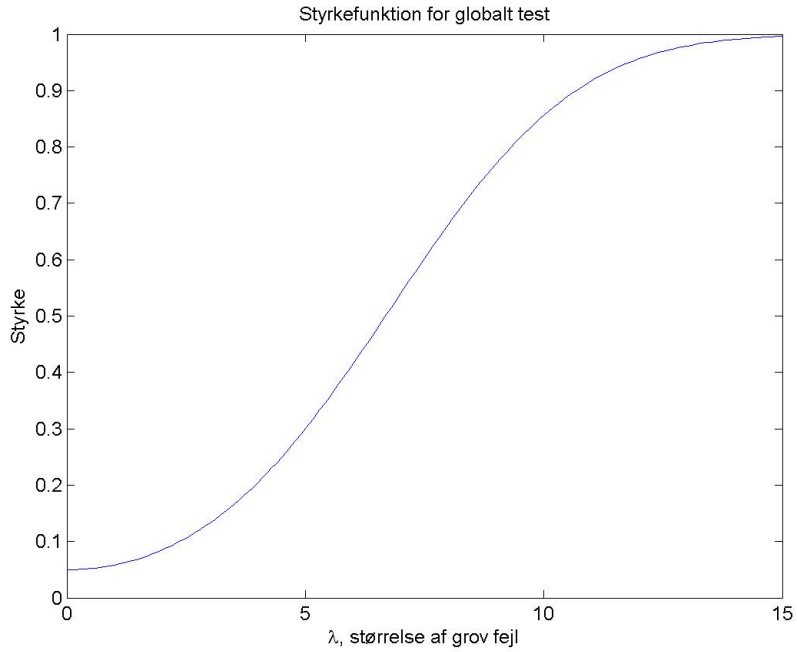
Man kan i almindelighed spørge, hvor stor ξ skal være, før at man kan opdage $\xi \neq 0$ med en given sandsynlighed? Svaret kommer an på ξ's retning. Lad $e_{\min}$ være egenvektoren (normeret) for vægtmatricen C svarende til C' s mindste egenværdi $\gamma_{\min}$. Da er $\delta = \xi^\top C \xi$ minimal, hvis δ har samme retning som $e_{\min}$, altså hvis $\xi = \rho e_{\min}$. Så fås nemlig

$$\delta = \rho e_{\min}^\top C \rho e_{\min} = \rho^2 ||e_{\min}||^2 \gamma_{\min} = \rho^2 \gamma_{\min}.$$

Ved at finde δ så

$$P\{Y \leq y_0 \mid \delta\} = 1 - \beta_0,$$

fås at det svarer til $\rho = \delta/\sqrt{\gamma_{\min}}$. Altså, jo mindre $\gamma_{\min}$ er, jo større skal afvigelsen ρ være, for at blive opdaget.



Figur 6: Styrkefunktionen for et globalt test med tre frihedsgrader. Styrken angiver sandsynligheden for at opdage en enkelt grov fejl af størrelse λ.

3.6.2 Tolerance

Styrkefunktioner er især anvendelige i forbindelse med planlægning af måleprocedurer.

Som et eksempel på dette vil nu se på, hvorledes man kan sikre, at et givet objekt overholder en bestemt teknisk specifikation i et simpelt eksempel. Nærmere betegnet ønsker vi at sikre, at objektet har en længde på μ_0, med en specificeret tolerance på τ. Med andre ord, skal vi garantere, at længden af objektet ligger i intervallet $[\mu_0 - \tau, \mu_0 + \tau]$.

Vi antager, at vi kan foretage n målinger med en individuel spredning på σ_0. Vi afpræver om objektet overholder specifikationerne ved at udføre et u-test for hypotesen $H_0 : \mu = \mu_0$. Da vi ikke ønsker at kassere for mange objekter uden grund, ønsker vi at holde sandsynligheden for fejl af type I nede, og anvender et lavt signifikansniveau α. Men hvor mange målinger skal

vi så udføre for at sikre, at vores objekter overholder specifikationerne med den givne tolerance?

Ved kendt σ_0 skal vi med andre ord sikre, at vi ved n målinger har en tilstrækkelig stor styrke $\beta_n(\delta)$ for $\delta = \tau$. Vi har

$$\begin{aligned}\beta_n(\delta) &= 1 - \Phi\left(\frac{\mu_0 + u_{1-\alpha/2}\sigma_0\sqrt{1/n} - \mu_0 - \delta}{\sigma_0\sqrt{1/n}}\right) \\ &\quad + \Phi\left(\frac{\mu_0 - u_{1-\alpha/2}\sigma_0\sqrt{1/n} - \mu_0 - \delta}{\sigma_0\sqrt{1/n}}\right) \\ &= 1 - \Phi\left(u_{1-\alpha/2} - \frac{\delta\sqrt{n}}{\sigma_0}\right) + \Phi\left(u_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma_0}\right).\end{aligned}$$

Hvis vi fastlægger, at styrken $\beta(\tau)$ skal være mindst β_0, skal vi altså løse ligningen

$$\beta_0 = 1 - \Phi\left(u_{1-\alpha/2} - \frac{\tau\sqrt{n}}{\sigma_0}\right) + \Phi\left(u_{\alpha/2} - \frac{\tau\sqrt{n}}{\sigma_0}\right), \tag{17}$$

med hensyn til n. Dette kan nemt gøres numerisk.

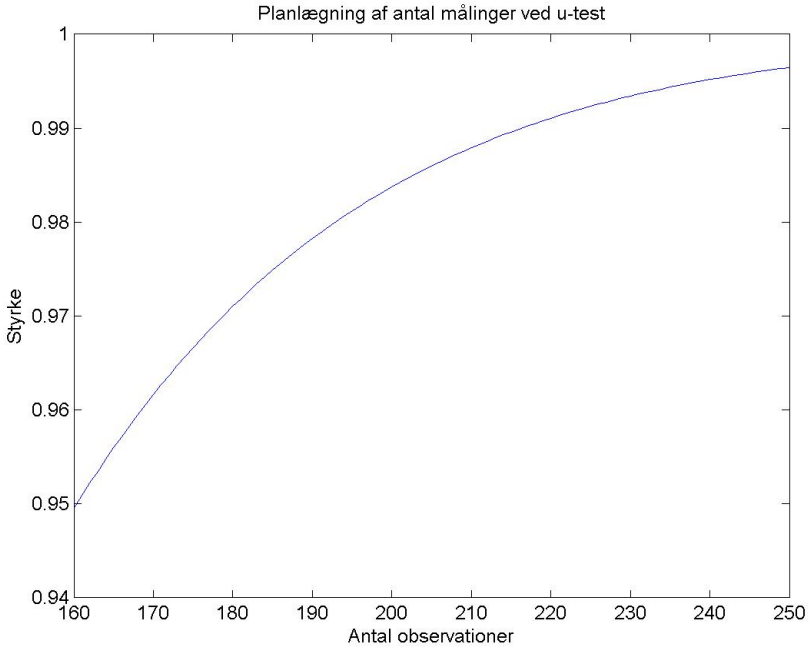
Eksempel 20 Vi ser igen på vores højdemålinger, som er foretaget med kendt varians $\sigma = 3$. Vi ønsker at garantere at $110 \leq \mu \leq 112$ på baggrund af en kontrolprocedure, hvor vi tester $H_0 : \mu = 111$ med signifikansniveau α=1%.

Som tidligere beregnet, ser det ikke godt ud med kun 3 målinger. Testet er alt for svagt og der må flere målinger til. Vi ønsker at sikre en styrke på mindst $\beta_0 = 99\%$ i endepunkterne af toleranceintervallet.

Hvis $\alpha = 1\%$ kan ligningen (17) omskrives til

$$\beta_n(1) = 0.99 \quad = \quad 1 - \Phi\left(2.576 - \sqrt{n}/3\right) + \Phi\left(-2.576 - \sqrt{n}/3\right).$$

På Figur 7 er højresiden afbildet som funktion af antal observationer n. Vi kan derefter aflæse grafisk, at vi skal bruge omkring $n = 220$ målinger foretaget med den givne spredning. Det er mange, men spredningen er jo også tre gange så stor som tolerancen. □



Figur 7: Styrken for et u-test med spredning 3 som funktion af antal observationer.

4 Den todimensionale normalfordeling

Et par af stokastiske variable (X, Y) siges at have en todimensional normalfordeling med parametre $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \rho)$, hvis tæthedsfunktionen har følgende udseende

$$f_{(X,Y)} = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-Q(x,y)/2\right\}, \tag{18}$$

hvor $Q(x, y)$ er det kvadratiske udtryk

$$Q(x,y) = \frac{1}{1-\rho^2}\left\{\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{x-\mu_2}{\sigma_2}\right) + \left(\frac{x-\mu_2}{\sigma_2}\right)^2\right\}.$$

Man kan så vise, at der gælder

$$EX = \mu_1, EY = \mu_2, \mathrm{Var}(X) = \sigma_1^2, V(Y) = \sigma_2^2, \mathrm{Cov}(X,Y) = \rho\sigma_1\sigma_2,$$

så ρ angiver *korrelationen* mellem X og Y

$$\rho = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}.$$

Matricen

$$\Sigma = \left\{ \begin{array}{cc} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{array} \right\}$$

kaldes *kovariansmatricen* for (X,Y), mens

$$\mu = \left(\begin{array}{c} \mu_1 \\ \mu_2 \end{array} \right)$$

kaldes *middelværdivektoren* for (X,Y). I det følgende vil $(X,Y) \sim \mathcal{N}_2(\mu,\Sigma)$ betegne, at (X,Y) har en todimensional normalfordeling med middelværdivektor μ og kovariansmatrix Σ.

4.1 Konturellipser

Niveaukurverne for tæthedsfunktionen for den todimensionale normalfordeling er ellipseformede og betegnes derfor *konturellipser*. De er givet ved ligningen

$$\frac{1}{1-\rho^2}\left\{\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right\} = c^2. \quad (19)$$

Hvis vi skriver denne ligning på matrixform, får vi

$$(x-\mu)^\top \Sigma^{-1}(x-\mu)$$

hvor $M^\top$ er den transponerede til matricen M, $x = (x_1,x_2)^\top$, $\mu = (\mu_1,\mu_2)^\top$ og

$$\Sigma = \left\{ \begin{array}{cc} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{array} \right\}.$$

Ellipserne har alle centrum i μ og de er større, jo større c er.

Som det kan beregnes ud fra formlerne i kap. 8.6 i Thomas and Finney (1987) og angivet i Hald (1948) formel (17.57), angiver (19) en ellipse, hvor den ene akses hældning α med 1.-aksens retning er givet ved ligningen

$$\tan 2\alpha = \frac{2\rho\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2},$$

hvis $\sigma_1^2 \neq \sigma_2^2$ og ellers er $\alpha = \pi/4$. Excentriciteten af ellipsen er uafhængig af c^2 og bestemt ved $e = \sqrt{a^2 - b^2}/a$, hvor a og b er længden af de to hovedakser, se Thomas and Finney (1987), kap. 8.4. Disse er ifølge Hald (1948) (17.62) og (17.63), givet ved ligningerne

$$\begin{aligned}(a+b)^2 &= \sigma_1^2 + \sigma_2^2 + 2\sigma_1\sigma_2\sqrt{1-\rho^2} \\ (a-b)^2 &= \sigma_1^2 + \sigma_2^2 - 2\sigma_1\sigma_2\sqrt{1-\rho^2}.\end{aligned}$$

Hvis alle parametre undtagen ρ fastholdes, medens denne varieres, fås en skare ellipser, som alle har samme omskrevne rektangel, men hvor store numeriske værdier af ρ svarer til meget smalle ellipser, se Fig. 17.4 i Hald (1948). Forholdet mellem siderne i dette rektangel er σ_1/σ_2.

Det er let at beregne, hvor stor en del af sandsynlighedsmassen, der ligger inden i denne ellipse, for venstresiden af (19) følger en χ^2-fordeling med 2 frihedsgrader. En sådan har tæthed

$$f(y) = e^{-y/2}/2, \quad y > 0,$$

så vi får

$$p_c = F_2(c^2) = \int_0^{c^2} f(y)\,dy = \int_0^{c^2} e^{-y/2}/2\,dy = 1 - e^{-c^2/2}.$$

Ønsker vi en konturellipse, som omfatter 95% sandsynlighedsmassen, skal c^2 derfor bestemmes som

$$c^2 = -2\log .05 = 5.991.$$

Dette kunne naturligvis også have været fundet ved at slå op i en tabel over χ^2-fordelingen.

4.2 Konfidensellipser og test

En 95% konturellipse indeholder 95% af sandsynlighedsmassen. I modsætning hertil er en 95% konfidensellipse et "ellipseformet estimat" af den ukendte middelværdivektor μ, som har 95% sandsynlighed for at omfatte den sande værdi. Konturellipsen har centrum i den 'sande' værdi μ, medens konfidensellipserne har centrum i den estimerede værdi $\hat{\mu}$.

Antag nu at $(X_i, Y_i) \sim \mathcal{N}_2(\mu, \Sigma)$ for $i = 1, \ldots, n$ er uafhængige. Estimatet for μ er givet ved

$$\hat{\mu} = \begin{pmatrix} \bar{X} \\ \bar{Y} \end{pmatrix}.$$

Man kan vise at $\hat{\mu} \sim \mathcal{N}_2(\mu, \Sigma/n)$.

Konfidensellipser konstrueres som regel ud fra statistiske test på en måde, som er analog til konfidensintervaller. Vi ser således på test for hypotesen $H_0 : \mu = \mu_0$. Vi har nu en række forskellige tilfælde.

4.2.1 Kovariansmatrix kendt

Antag at vi har en stikprøve som ovenfor af størrelse n fra en todimensional normalfordeling med middelværdivektor $\mu = (\mu_1, \mu_2)^\top$ og kendt kovariansmatrix $\Sigma = \Sigma_0$. Betragt så hypotesen

$$H_0 : \mu = \mu_0,$$

hvor μ_0 er en kendt vektor. En rimelig testvariabel for H_0 mod $H_A : \mu \neq \mu_0$ er givet ved

$$\begin{aligned} X^2 &= n(\hat{\mu} - \mu_0)^\top \Sigma_0^{-1} (\hat{\mu} - \mu_0) \\ &= p_{11}(\hat{\mu}_1 - \mu_{01})^2 + p_{22}(\hat{\mu}_2 - \mu_{02})^2 + 2p_{12}(\hat{\mu}_1 - \mu_{01})(\hat{\mu}_2 - \mu_{02}) \end{aligned}$$

hvor $P = \Sigma_0^{-1}$ er vægtmatricen og store værdier af X^2 er kritiske. Endvidere kan man vise at

$$X^2 \sim \chi^2(2).$$

Det vil sige, at det kritiske område er givet ved

$$K_\alpha = \{x \,|\, x^2 > \chi^2(2)_{1-\alpha}\}.$$

Vil man fastlægge $(1-\alpha)$-konfidensellipsen, er denne således givet ved de μ_0'er, som opfylder uligheden

$$n(\hat{\mu}-\mu_0)^\top \Sigma_0^{-1}(\hat{\mu}-\mu_0) \leq \chi^2(2)_{1-\alpha}$$

hvor $\chi^2(2)_{1-\alpha}$ er $1-\alpha$ fraktilen i χ^2-fordelingen med 2 frihedsgrader, eller simplere

$$n(\hat{\mu}-\mu_0)^\top \Sigma_0^{-1}(\hat{\mu}-\mu_0) \leq -2\log\alpha.$$

Her har vi benyttet, at χ^2-fordelingen med 2 frihedsgrader er en eksponentialfordeling og fordelingsfunktionen derfor kan udtrykkes explicit ved eksponentialfunktionen.

4.2.2 Kovariansmatrix ukendt

Hvis kovariansmatricen Σ er ukendt, er estimatet for denne givet ved

$$S = \left\{ \begin{array}{cc} s_x^2 & s_{xy} \\ s_{xy} & s_y^2 \end{array} \right\}.$$

Betragt som ovenfor hypotesen

$$H_0 : \mu = \mu_0,$$

hvor μ_0 er en kendt vektor. Idet S^{-1} betegner den inverse matrix til S, er en rimelig testvariabel givet ved

$$T^2 = \frac{n(n-2)}{2(n-1)}(\hat{\mu}-\mu_0)^\top S^{-1}(\hat{\mu}-\mu_0).$$

Testvariablen måler den vægtede kvadratiske afvigelse mellem $\hat{\mu}$ og μ_0, og derfor er store værdier af T^2 kritiske. Man kan vise, at $T^2 \sim F(2, n-2)$, det vil sige, at det kritiske område er er givet ved

$$\{t^2 > v^2(2, n-2)_{1-\alpha}\} \tag{20}$$

hvor $v^2(d_1, d_2)_\gamma$ er γ-fraktilen i $F(2, n-2)$-fordelingen.

Dette test er indført af den amerikanske statistiker Harold Hotelling og er kendt som *Hotelling's* T^2. Testet er en generalisering af det tilsvarende test i den endimensionale normalfordeling, se afsnit 3.3.3, hvor testvariablen var

$$t = \frac{\sqrt{n}(\bar{x} - \mu_0)}{s} \sim t(n-1).$$

Kvadreres denne størrelse, ses at store værdier af $t^2 = n(\bar{x}-\mu_0)^2/s^2$ er kritiske og $t^2 \sim F(1, n-1)$, jvf. afsnit 1.4.3.

For at finde konfidensellipsen, skal vi bestemme de μ_0-er, for hvilke hypotesen $\mu = \mu_0$ accepteres på niveau α. På denne måde defineres et $(1-\alpha)$-konfidensområde. Med andre ord bestemmes ellipsen som

$$(\hat{\mu} - \mu_0)^\top S^{-1}(\hat{\mu} - \hat{\mu}_0) \leq \frac{2(n-1)}{n(n-2)} v^2(2, n-2)_{1-\alpha}.$$

4.3 Relative konfidensellipser

Antag at et punkt i planen koordineres til 2 forskellige tidspunkter, og der haves estimater for punktkoordinater til tidspunkt i:

$$\hat{\mu}_i \ \sim \ \mathcal{N}_2(\mu_i, \Sigma_{0i}/n_i), \quad i = 1, 2$$

hvor disse er uafhængige. Kovariansmatricen Σ_{0i} for $\hat{\mu}_i$ antages for eksempel bestemt ved udjævning af et net af målinger, hvori punktet er indgået.

Test for $\mu_1 = \mu_2$ baseres på

$$\hat{d} = \hat{\mu}_1 - \hat{\mu}_2 \ \sim \ \mathcal{N}_2(\delta, \Sigma_{01}/n_1 + \Sigma_{02}/n_2)$$

d.v.s. vi får testvariablen

$$X^2 = \hat{d}(\Sigma_{01}/n_1 + \Sigma_{02}/n_2)^{-1}\hat{d}^\top$$

som skal vurderes i en $\chi^2(2)$-fordeling. Derfor er $(1-\alpha)$-konfidensellipsen for differensvektoren δ bestemt ved

$$(\delta - \hat{d})(\Sigma_{01}/n_1 + \Sigma_{02}/n_2)^{-1}(\delta - \hat{d})^\top = \chi^2(2)_{1-\alpha} \leq -2\log\alpha$$

og kaldes den *relative fejlellipse*[1]. Den angiver altså et elliptisk estimat for deformationen af det pågældende punkt.

Det er noget mere indviklet at diskutere relative konfidensellipser, hvis kovariansmatricerne ikke kan antages kendte.

5 Den flerdimensionale normalfordeling

Lad $X_1, \ldots, X_n$ være stokastiske variable og betragt den *stokastiske vektor* $X = (X_1, \ldots, X_n)^\top$. Vi regner med søjlevektorer og $^\top$ angiver som tidligere transponering, det vil sige at

$$(X_1, \ldots, X_n)^\top = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}.$$

Lad $a = (a_1, \ldots, a_n)^\top \in R^n$ betegne en vektor i R^n. Vi siger da, at X har en *n-dimensional normalfordeling,* hvis der for alle vektorer $a \in R^n$ gælder at

$$a \cdot X = a_1 X_1 + \cdots + a_n X_n$$

er normalfordelt.

Vektoren $E(X) = \mu = (\mu_1, \ldots, \mu_n)^\top$, hvor $\mu_i = E(X_i)$ for $i = 1, \ldots, n$, kaldes for *middelværdivektoren* og matricen $\Sigma = \{\sigma_{ij}\}_{i,j=1}^n$ med elementer $\sigma_{ij} = \mathrm{Cov}(X_i, X_j)$, kaldes *kovariansmatricen.* Hvis Σ har en invers, det vil sige at $\det(\Sigma) \neq 0$, kaldes denne *vægtmatricen* og betegnes ofte med $P = \Sigma^{-1}$. I så tilfælde har fordelingen tætthedsfunktion

$$f(x_1, \ldots, x_n) = \det(P)(2\pi)^{-n/2} \exp\left\{-(x-\mu)^\top P(x-\mu)/2\right\}.$$

Alt i alt siges X at have en n-dimensional normalfordeling med middelværdivektor μ og kovariansmatrix Σ, hvilket også betegnes

$$X \sim \mathcal{N}_n(\mu, \Sigma).$$

[1]Dette stemmer ifølge Kai Borre ikke overens med den i landmålingen gængse betydning af udtrykket.

Eksempel 21 Antag at $X_1, \ldots, X_n$ er uafhængige og $X_i \sim \mathcal{N}(\mu_i, p_i^{-1})$. Da gælder, som vi tidligere har omtalt, at $a_1X_1 + \cdots + a_nX_n$ er normalfordelt. Endvidere er $\sigma_{ij} = \mathrm{Cov}(X_i, X_j) = 0$ for $i \neq j$, og $\sigma_{ii} = \mathrm{Cov}(X_i, X_i) = \mathrm{Var}(X_i) = p_i^{-1}$.

Det vil altså sige, at $X \sim \mathcal{N}_n(\mu, \Sigma)$, hvor $\Sigma = P^{-1}$ er en diagonalmatrix, hvor det i'te element i diagonalen er lig med p_i^{-1}. □

I det følgende vil vi antage at $X \sim \mathcal{N}_n(\mu, \Sigma)$. Vi ved fra definitionen, at $a \cdot X$ er normalfordelt for alle $a \in R^n$. Vi ønsker at bestemme middelværdien og variansen af $a \cdot X$. Først beregnes middelværdien til

$$E(a \cdot X) = E\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i E(X_i) = \sum_{i=1}^{n} a_i \mu_i = a \cdot \mu.$$

For variansen fås, idet $a \cdot \mu$ blot er en konstant,

$$\begin{aligned}
\mathrm{Var}(a \cdot X) &= \mathrm{Var}(a \cdot X - a \cdot \mu) = \mathrm{Var}\{a \cdot (X - \mu)\} \\
&= E\left\{a \cdot (X - \mu)^2\right\} - \left[E\left\{a \cdot (X - \mu)\right\}\right]^2 \\
&= E\left\{\sum_{i=1}^{n} a_i (X_i - \mu_i) \sum_{j=1}^{n} a_j (X_j - \mu_j)\right\} - 0 \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j E\left\{(X_i - \mu_i)(X_j - \mu_j)\right\} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \sigma_{ij} = a^\top \Sigma a.
\end{aligned}$$

Overvejelserne ovenfor kan formelt sammenfattes som

$$X \sim \mathcal{N}_n(\mu, \Sigma) \iff a \cdot X \sim \mathcal{N}(a \cdot \mu, a^\top \Sigma a) \text{ for alle } a \in R^n. \tag{21}$$

Den flerdimensionale normalfordeling har følgende pæne egenskab. Lad A være en $m \times n$ matrix og $\eta \in R^m$ en m-vektor. Hvis $X \sim \mathcal{N}_n(\mu, \Sigma)$ så er

$$Y = AX + \eta \sim \mathcal{N}_m(A\mu + \eta, A\Sigma A^\top). \tag{22}$$

Man viser (22) ved at indse, at Y opfylder (21), det vil sige, at man for ethvert $b \in R^m$ viser, at

$$b \cdot Y \sim \mathcal{N}(b \cdot (A\mu + \eta), b^\top A\Sigma A^\top b),$$

idet man udnytter, at $Y = AX + \eta$, og at X opfylder (21).

Man kan vise, at hvis Σ er en kovariansmatrix, så findes der altid en $n \times n$ matrix A således at $\Sigma = AA^\top$. For eksempel kan A findes ved Cholesky's metode, og i så tilfælde er A endda en trekantsmatrix. Hvis vi endvidere antager, at Σ har en invers, det vil sige at $\det(\Sigma) \neq 0$, så har også A en invers, og det følger så af (22) at

$$Y = A^{-1}(X - \mu) \sim \mathcal{N}_n(0, I_n) \tag{23}$$

hvor I_n er diagonalmatricen, hvis diagonalelementer alle er lig med 1. Dette betyder med andre ord at

$$Y_i \sim \mathcal{N}(0,1), \quad i = 1, \ldots, n$$

er uafhængige. Heraf kan så sluttes, at

$$Y^\top Y = \sum_{i=1}^{n} Y_i^2 \sim \chi^2(n),$$

altså at $Y^\top Y$ er χ^2-fordelt med n frihedsgrader. Da nu $Y^\top = (X-\mu)^\top (A^{-1})^\top$, ses at

$$\begin{aligned} Y^\top Y &= (X-\mu)^\top \left(A^{-1}\right)^\top A^{-1}(X-\mu) \\ &= (X-\mu)^\top (AA^\top)^{-1}(X-\mu) \\ &= (X-\mu)^\top \Sigma^{-1}(X-\mu). \end{aligned}$$

Vi har med andre ord vist, at hvis $X \sim \mathcal{N}_n(\mu, \Sigma)$, så er

$$(X-\mu)^\top \Sigma^{-1}(X-\mu) \sim \chi^2(n). \tag{24}$$

6 Den generelle udjævningsmodel

I det følgende antager vi, at kovariansmatricen $\Sigma = \Sigma_0$ er kendt. Dette er normalt i forbindelse med en udjævning, hvor $\Sigma_0^{-1} = P$ er vægtmatricen (der som oftest er diagonal).

I udjævningen skal vi fastlægge m ukendte parametre $\alpha = (\alpha_1, \ldots, \alpha_m)^\top$, (eksempelvis koordinater) på basis af observationen $X = (X_1, \ldots, X_n)^\top$ (eksempelvis længde- og vinkelmålinger), hvor vi må antage at $n \geq m$, det vil sige, at vi har $(n-m)$ overbestemmelser.

Generelt vil nettets struktur give anledning til ikke-lineære relationer

$$EX_i = f_i(\alpha_1, \ldots, \alpha_m), \quad i = 1, \ldots, n,$$

men i praksis erstattes disse af lineære approximationer

$$EX_i = x_{0i} + b_{i1}\alpha_1 + \cdots + b_{im}\alpha_m, \quad i = 1, \ldots n,$$

hvor

$$b_{ij} = \left.\frac{\partial f_i(\alpha_0, \ldots, \alpha_m)}{\partial \alpha_j}\right|_{\alpha=\alpha_0},$$

$\alpha_0 = (\alpha_{01}, \ldots, \alpha_{0n})$ er en kendt vektor og $x_0 = f(\alpha_0) = (x_{01}, \ldots, x_{0n})$. Idet B betegner $n \times m$ matricen $\{b_{ij}\}$ (som altså er kendt), kan dette også udtrykkes

$$EX = x_0 + B\alpha,$$

hvor x_0 er en kendt vektor. Vi vælger nu at foretage en nulpunktskorrektion, så vi med andre ord antager, at

$$X \leftarrow X - x_0, \quad X \sim \mathcal{N}_n(B\alpha, \Sigma_0).$$

6.1 Estimation

Hvordan skal vi estimere α? Man kunne for eksempel vælge det α, som minimerer afstanden mellem X og $B\alpha$, hvilket indebærer, at alle X_i'er indgår med samme vægt. Dette virker urimeligt, hvis X_i'erne har forskellige varianser, idet en observation med lille varians må indeholde mere information om parametrene end en observation med stor varians.

Lad derfor A_0 være bestemt ved at $\Sigma_0 = A_0 A_0^\top$. Da vi så har at $A_0^{-1}\Sigma_0(A_0^{-1})^\top = I_p$ følger det at

$$Y = A_0^{-1}X \sim \mathcal{N}_n(A_0^{-1}B\alpha, I_n),$$

det vil sige, at komponenterne i $Y = (Y_1, \ldots, Y_n)^\top$ er uafhængige, og de har alle varians 1.

Vi kan nu konstruere et rimeligt estimat for α ved at minimere afstanden mellem $Y = A_0^{-1}X$ og $A_0^{-1}B\alpha$. Denne afstands kvadrat er givet ved

$$\begin{aligned} ||A_0^{-1}X - A_0^{-1}B\alpha||^2 &= \left\{A_0^{-1}(X - B\alpha)\right\}^\top \left\{A_0^{-1}(X - B\alpha)\right\} \\ &= (X - B\alpha)^\top (A_0^{-1})^\top A_0^{-1}(X - B\alpha) \\ &= (X - B\alpha)^\top \Sigma_0^{-1}(X - B\alpha). \end{aligned}$$

Dette udtryk minimeres (vægtede mindste kvadrater) ved at løse *normalligningerne*

$$B^\top \Sigma_0^{-1} X = (B^\top \Sigma_0^{-1} B)\alpha,$$

som har løsningen

$$\hat{\alpha} = (B^\top \Sigma_0^{-1} B)^{-1} B^\top \Sigma_0^{-1} X.$$

Dernæst ønsker vi at finde fordelingen af $\hat{\alpha}$. Dette gøres let ved at observere, at $\hat{\alpha}$ blot er X multipliceret med $m \times n$ matricen $(B^\top \Sigma_0^{-1} B)^{-1} B^\top \Sigma_0^{-1}$. Det følger så fra (22) at

$$\hat{\alpha} \sim \mathcal{N}_m \left\{ \alpha, \left(B^\top \Sigma_0^{-1} B \right)^{-1} \right\},$$

idet

$$\left\{ (B^\top \Sigma_0^{-1} B)^{-1} B^\top \Sigma_0^{-1} \right\} \Sigma_0 \left\{ (B^\top \Sigma_0^{-1} B)^{-1} B^\top \Sigma_0^{-1} \right\}^\top = \left(B^\top \Sigma_0^{-1} B \right)^{-1}.$$

6.2 Modelkontrol

I det følgende vil vi betragte hypotesen

$$H_0 : E(X) = B\alpha \text{ mod } H_1 : E(X) \neq B\alpha.$$

Dette er relevant for at kontrollere, at X ikke er behæftet med systematiske eller grove fejl.

Som testvariabel vil vi betragte "afstanden" fra X til $B\hat{\alpha}$. På samme måde som tidligere kan man argumentere for, at det er rimeligt at betragte den vægtede afstand, det vil sige at vores testvariabel er givet ved

$$W = (X - B\hat{\alpha})^\top \Sigma_0^{-1} (X - B\hat{\alpha}).$$

Det er klart, at store værdier af W er kritiske, og man kan vise, at hvis H_0 er sand, så gælder at

$$W \sim \chi^2(n - m).$$

Vi får således et kritisk område af formen

$$K_\alpha = \{ w \,|\, w > \chi^2(n - m)_{1-\alpha} \}.$$

Dette test er også kendt som et *globalt test* (Cederholm 2000).

Størrelsen $\sqrt{w_{obs}/(n-m)}$ kaldes også *spredningen på vægtenheden* og er standardoutput for udjævningsprogrammer (forudsat at $\sigma_0 = 1$), så det er lige ud ad landevejen at bestemme w_{obs}.

Det er måske på sin plads at bemærke, at størrelsen $r = x - B\hat{\alpha}$ ofte kaldes *residualvektoren,* og at $[Prr] = r^\top Pr$ er en ofte anvendt betegnelse for w, hvor $P = \Sigma_0^{-1}$ er vægtmatricen. I landmålingslitteraturen har man traditionelt modsat fortegn på residualerne i forhold til disse noter.

Hvis man forkaster sin model, har det naturligvis interesse at undersøge, hvor de signifikante afvigelser er opstået. En mulig forklaring er systematiske eller grove fejl på en eller flere målinger. Dette vil afspejles i, at de tilhørende residualer bliver "for store". Størrelsen af residualerne vurderes som regel i forhold til deres spredning.

Hvis modellen er rigtig, kan vi umiddelbart bestemme fordelingen af residualvektoren R ved at observere at

$$R = X - B\hat{\alpha} = (I_n - B(B^\top \Sigma_0^{-1} B)^{-1} B^\top \Sigma_0^{-1})X.$$

Det følger så fra (22) og en del matrixmanipulationer, at

$$R \sim \mathcal{N}_n(0, \Sigma_0 - B(B^\top \Sigma_0^{-1} B)B^\top),$$

og hermed at

$$\mathrm{Var}(R_i) = \sigma_{ii} - b_i^\top (B^\top \Sigma_0^{-1} B)^{-1} b_i,$$

hvor

$$\sigma_{ii} = \mathrm{Var}(X_i) = (P^{-1})_{ii}$$

og $b_1, \ldots, b_n$ er de transponerede til rækkerne i B, altså

$$B^\top = \{b_1 : b_2 : \ldots : b_n\}.$$

Vi kan nu vurdere om et observeret residual r_i tilhører acceptområdet for et u-test med signifikansniveau α ved at undersøge om

$$\frac{|r_i|}{\sqrt{\mathrm{Var}(r_i)}} \leq u_{1-\alpha/2}.$$

Det skal understreges, at residualerne er stokastiske afhængige, hvorfor ovenstående kun kan betragtes som en overfladisk indikator for eventuelle fejlkilder. Eksempelvis kan en grov fejl på een måling give anledning til at flere residualer bliver "for store", og ofte andre residualer end det, hvor fejlen er placeret.

6.3 Test for $\alpha = \alpha_0$

Antag at vi har estimeret α og gerne vil undersøge om $\alpha = \alpha_0$, hvor α_0 er en kendt vektor. Dette er eksempelvis problemet ved kontrol af en afsætning. Vi har altså at

$$\hat{\alpha} \sim \mathcal{N}_m(\alpha, (B^\top \Sigma_0^{-1} B)^{-1})$$

og vil teste

$$H_0 : \alpha = \alpha_0 \text{ mod } H_1 : \alpha \neq \alpha_0.$$

Ved at gennemføre overvejelser (omkring vægtning) af samme karakter som i det foregående, kan man indse, at en fornuftig testvariabel er givet ved

$$W = (\hat{\alpha} - \alpha_0)^\top B^\top \Sigma_0^{-1} B(\hat{\alpha} - \alpha_0).$$

Det følger af (24) at $W \sim \chi^2(m)$, og da store værdier er kritiske, er det kritiske område bestemt som

$$K_\alpha = \{w \,|\, w > \chi^2(m)_{1-\alpha}\}.$$

I praksis kan w beregnes på følgende måde. Ved udjævning bestemmes størrelsen

$$w_1 = (x - B\hat{\alpha})^\top \Sigma_0^{-1} (x - B\hat{\alpha})$$

det vil sige at "afstanden" mellem observationen og $B\hat{\alpha}$. Hvis vi endvidere kender afstanden mellem observationen og $B\hat{\alpha}_0$, det vil sige

$$w_2 = (x - B\alpha_0)^\top \Sigma_0^{-1} (x - B\alpha_0)$$

så er "afstanden" mellem $\hat{\alpha}$ og α_0 (eller $B\hat{\alpha}$ og $B\alpha_0$) givet ved

$$w = (\hat{\alpha} - \alpha_0)^\top B^\top \Sigma_0^{-1} B(\hat{\alpha} - \alpha_0) = w_2 - w_1.$$

Aktuelt kan w_2 beregnes direkte eller ved at lave en udjævning, hvor alle parametre på forhånd er fikserede, det vil sige at at $\alpha = \alpha_0$. Som output fås spredningen på vægtenheden, der vil være lig med $\sqrt{w_2/n}$.

6.4 Test for $\alpha_1 = \alpha_2$

I mange situationer er det relevant at undersøge om to middelværdivektorer er ens. Det er for eksempel typisk spørgsmålet i en deformationsanalyse.

Antag at vi til tidspunkt 1 har estimeret middelværdivektoren α_1, eksempelvis koordinater til punkter på en bro, og at vi har

$$\hat{\alpha}_1 \sim \mathcal{N}_m(\alpha_1, \Sigma_1).$$

Formodentlig er Σ_1 givet ved $\Sigma_1 = (B_1^\top P_1 B_1)^{-1}$ hvor B_1 og P_1 (jvf. tidligere betragtninger) er størrelser, som er karakteristiske for nettet etableret til tidspunkt 1.

Antag endvidere, at vi til tidspunkt 2 har estimeret koordinaterne til de samme punkter, det vil sige at

$$\hat{\alpha}_2 \sim \mathcal{N}_m(\alpha_2, \Sigma_2),$$

hvor Σ_2 er bestemt via nettet til tidspunkt 2. Det er klart at vi kan antage, at $\hat{\alpha}_1$ og $\hat{\alpha}_2$ er uafhængige stokastiske vektorer.

Vi er interesseret i at undersøge om "broen" har flyttet sig, hvilket naturligvis kontrolleres ved at teste hypotesen

$$H_0 : \alpha_1 = \alpha_2 \text{ mod } H_1 : \alpha_1 \neq \alpha_2.$$

Som testvariabel vil vi betragte "længden" af $\hat{\alpha}_1 - \hat{\alpha}_2$. Af det foregående følger, at hvis H_0 er sand, da gælder, at

$$\hat{\alpha}_1 - \hat{\alpha}_2 \sim \mathcal{N}_m(0, \Sigma_1 + \Sigma_2)$$

og via (24) fås så, at

$$W = (\hat{\alpha}_1 - \hat{\alpha}_2)^\top (\Sigma_1 + \Sigma_2)^{-1} (\hat{\alpha}_1 - \hat{\alpha}_2) \sim \chi^2(m)$$

det vil sige at er det kritiske område bestemt som

$$K_\alpha = \{w \,|\, w > \chi^2(m)_{1-\alpha}\}.$$

Ved hjælp af et standard udjævningsprogram, kan man bestemme w på følgende måde: Ved udjævningerne til tidspunkt i, $i = 1, 2$, bestemmes blandt andet

$$w_i = (n_i - m)\hat{\sigma}_i^2,$$

hvor $n_i - m$ er antallet af overbestemmelser ved den $i$te udjævning og $\hat{\sigma}_i$ er spredningerne på vægtenheden. Størrelsen $w_1 + w_2$ repræsenterer da afstanden mellem vore målinger og modellen, hvor punkternes koordinater ikke er antaget ens.

Dernæst laves en udjævning, hvor de to net "lægges oven i hinanden". Hvis $\hat{\sigma}_{1+2}$ er spredningen på vægtenheden ved denne udjævning, så vil

$$w_{1+2} = (n_1 + n_2 - m)\hat{\sigma}_{1+2}^2$$

repræsentere afstanden mellem målingerne og modellen, hvor punkternes koordinater er antaget ens.

Afstanden mellem modellen med ens punkter og modellen, hvor punkternes koordinater antages forskellige, er nu givet ved

$$w = w_{1+2} - w_1 - w_2.$$

Hvis vi accepterer $\alpha_1 = \alpha_2 = \alpha$, er man ofte interesseret i et estimat for den fælles værdi af α_1 og α_2. Dette opnås ved at finde det α, som minimerer den totale, vægtede kvadratsum, givet ved udtrykket

$$(\hat{\alpha}_1 - \alpha)^\top \Sigma_1^{-1}(\hat{\alpha}_1 - \alpha) + (\hat{\alpha}_2 - \alpha)^\top \Sigma_2^{-1}(\hat{\alpha}_2 - \alpha).$$

Dette giver estimatet

$$\hat{\alpha} = \left(\Sigma_1^{-1} + \Sigma_2^{-1}\right)^{-1} \left(\Sigma_1^{-1}\hat{\alpha}_1 + \Sigma_2^{-1}\hat{\alpha}_2\right),$$

og fordelingen af $\hat{\alpha}$ bestemmes til at være

$$\hat{\alpha} \sim \mathcal{N}_m \left\{\alpha, \left(\Sigma_1^{-1} + \Sigma_2^{-1}\right)^{-1}\right\}.$$

6.5 Test for flytning

I det følgende vil vi se på, hvordan man kan undersøge om et net bestående af k punkter har flyttet sig i forhold til nogle givne fikspunkter, uden at nettet som sådant er blevet deformeret.

Lad $\hat{\alpha}_i \in R^2$, $i = 1, 2, \ldots, k$ være de estimerede koordinatpar for de k punkter efter en udjævning, det vil sige, at idet $\hat{\alpha}^\top = (\hat{\alpha}_1, \ldots, \hat{\alpha}_k)^\top$, har vi, at

$$\hat{\alpha} \sim \mathcal{N}_{2k} \left\{\alpha, (B^\top P B)^{-1}\right\},$$

hvor

$$E\hat{\alpha} = \alpha = (\alpha_1, \ldots, \alpha_k)^\top.$$

Lad $\alpha_0 = (\alpha_{10}, \ldots, \alpha_{k0})^\top$ være de k punkters koordinatpar før den eventuelle flytning og antag, at disse er kendte.

En flytning består som bekendt af en rotation (1 frihedsgrad) samt en translation (2 frihedsgrader), hvorfor vi med andre ord har $2k - 3$ overbestemmelser. Lad

$$U(\theta) = \left\{ \begin{array}{cc} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{array} \right\}, \quad \theta \in [0; 2\pi)$$

være matricen for rotationen og lad $m \in R^2$ være translationsvektoren. Vores hypotese er da, at der findes et par (θ, m), så

$$H_0 : E\hat{\alpha} = \left(\begin{array}{c} U(\theta)\alpha_{10} + m \\ \vdots \\ U(\theta)\alpha_{k0} + m \end{array} \right) = \alpha(\theta, m).$$

Vi tester mod alternativet, at dette ikke er tilfældet, altså at vi for alle par (θ, m) har

$$H_A : E\hat{\alpha} \neq \alpha(\theta, m).$$

Estimatet $(\hat{\theta}, \hat{m})$ for (θ, m) bestemmes under hypotesen ved at minimere den vægtede kvadratiske afstand mellem $\hat{\alpha}$ og $\alpha(\theta, m)$ eller — mere præcist — ved at minimere

$$\{\hat{\alpha} - \alpha(\theta, m)\}^\top B^\top PB \{\hat{\alpha} - \alpha(\theta, m)\}.$$

Testet for H_0 udføres dernæst ved at vurdere om den minimale afstand er for stor, hvilket indebærer, at det er kritisk med store værdier af testvariablen

$$W = \left\{\hat{\alpha} - \alpha(\hat{\theta}, \hat{m})\right\}^\top B^\top PB \left\{\hat{\alpha} - \alpha(\hat{\theta}, \hat{m})\right\}.$$

Fordelingen af W er (approximativt) $\chi^2(2k-3)$, det vil sige at det kritiske område er approximativt givet ved

$$K_\alpha = \{w \,|\, w > \chi^2(2k-3)_{1-\alpha}\}.$$

6.5.1 Bestemmelse af W og $(\hat{\theta}, \hat{m})$

Ved udjævningen til bestemmelse af $\hat{\alpha}$ beregnes spredningen $\hat{\sigma}_{01}$ på vægtenheden, som repræsenterer "afstanden" mellem vore observationer og $\hat{\alpha}$, idet

$$(n-2k)\hat{\sigma}_{01}^2 = (x - B\hat{\alpha})^\top P(x - B\hat{\alpha}).$$

Da en flytning af nettet (inklusive aksespejling, hvad der næppe kan være aktuelt) netop er karakteriseret ved, at længder og vinkler bevares, kan vi estimere (θ, m) ved at lave en udjævning, hvor vi fastholder en passende del af nettets vinkler og afstande. Hvis α_{i0} for $i = 1, 2, \ldots, k$ repræsenterer de k punkter, kan man eksempelvis fastholde afstandene mellem α_{i0} og $\alpha_{i+1,0}$ for $i = 1, 2, \ldots, k-1$ samt vinklerne mellem $\alpha_{i0} - \alpha_{i-1,0}$ og $\alpha_{i0} - \alpha_{i+1,0}$ for $i = 2, \ldots, k-1$.

Hvis udjævningsprogrammet ikke levner muligheder for at fastholde sådanne størrelser, kan man i stedet indføre dem som målinger med uendelig (meget stor) vægt. Man skal dog være opmærksom på, at dette kan give anledning til numeriske problemer. Spredningen $\hat{\sigma}_{01+2}$ på vægtenheden i forbindelse med denne udjævning repræsenterer da "afstanden" mellem vore observationer og $\alpha(\hat{\theta}, \hat{m})$, idet

$$(n-3)\hat{\sigma}_{01+2}^2 = \left\{x - B\alpha(\hat{\theta}, \hat{m})\right\}^\top P \left\{x - B\alpha(\hat{\theta}, \hat{m})\right\}.$$

Den vægtede, kvadratiske afstand mellem $\hat{\alpha}$ og $\alpha(\hat{\theta}, \hat{m})$ er da

$$\begin{aligned} w &= \left\{\hat{\alpha} - \alpha(\hat{\theta}, \hat{m})\right\}^\top B^\top PB \left\{\hat{\alpha} - \alpha(\hat{\theta}, \hat{m})\right\} \\ &= (n-3)\hat{\sigma}_{01+2}^2 - (n-2k)\hat{\sigma}_{01}^2. \end{aligned}$$

Estimatet $(\hat{\theta}, \hat{m})$ kan bestemmes via de estimerede koordinatpar $\alpha(\hat{\theta}, \hat{m})$. Betragtes for eksempel estimaterne $\alpha_i(\hat{\theta}, \hat{m}) = U(\hat{\theta})\alpha_{i0} + \hat{m}$ for $i = 1, 2$ for de to første koordinatpar, ses at

$$\alpha_1(\hat{\theta}, \hat{m}) - \alpha_2(\hat{\theta}, \hat{m}) = U(\hat{\theta})(\alpha_{10} - \alpha_{20}).$$

Drejningen kan således bestemmes ved at sammenligne denne vektor med $\alpha_{10} - \alpha_{20}$. Når drejningen er bestemt, ses det umiddelbart, at

$$\hat{m} = \alpha_1(\hat{\theta}, \hat{m}) - U(\hat{\theta})\alpha_{10}.$$

6.5.2 Styrkefunktionen for flytning af et punkt

Vi betragter to uafhængige bestemmelser af koordinaterne til et punkt i et 3-dimensionalt system $\hat{\alpha}_1$ og $\hat{\alpha}_2$ med

$$\hat{\alpha}_i \sim \mathcal{N}_3(\alpha_i, \Sigma_i),$$

hvor kovariansmatricerne Σ_i antages kendte (bestemt fra GPS-måling, for eksempel). Som i afsnit 6.4 tester vi for flytning baseret på teststørrelsen

$$W = (\hat{\alpha}_2 - \hat{\alpha}_1)^\top (\Sigma_1 + \Sigma_2)^{-1} (\hat{\alpha}_2 - \hat{\alpha}_1),$$

som følger en χ^2-fordeling med 3 frihedsgrader, dersom punktet ikke har flyttet sig.

Testet udføres ved at forkaste, hvis

$$P(W \geq w_{\text{obs}}) \leq \alpha,$$

svarende til, at for $\alpha = 0.05; 0.01$ forkastes hypotesen, hvis $w_{\text{obs}} \geq 7.81; 11.3$ hhv.

For at bestemme, hvor store deformationer, der kan opdages ved et sådant test, skal vi beregne, ved en flytning δ (altså $\alpha_2 = \alpha_1 + \delta$), hvad er sandsynligheden

$$\beta(\delta) = P(W \geq w_0 \,|\, \delta), \tag{25}$$

hvor $w_0 = 7.81$ hhv. 11.3. Hvis $\delta = 0$, haves at denne er 5% hhv. 1%, men hvis $\delta \neq 0$ skal den ikke-centrale χ^2-fordeling tages i brug som beskrevet i afsnit 3.6.1. Idet vi som der sætter $\lambda = \delta^\top (\Sigma_1 + \Sigma_2)^{-1} \delta$, er W i almindelighed ikke-centralt χ^2-fordelt med 3 frihedsgrader og ikke-centralitetsparameter λ.

Idet $\lambda = \delta^\top (\Sigma_1 + \Sigma_2)^{-1} \delta$, afhænger styrkefunktionen β af, i hvilken retning deformationen er sket. I værste fald er deformationen sket i den retning, som angives af egenvektoren svarende til den mindste egenværdi $\gamma_{\min}$ for vægtmatricen $(\Sigma_1 + \Sigma_2)^{-1}$.

Hvis $\delta = \rho e_{\min}$, hvor $e_{\min}$ er en normeret egenvektor, er $\lambda = \rho^2/\gamma_{\min}$. For eksempel kan man da ved $\alpha = 0.05$ opdage deformationer af størrelse

$$\gamma = \sqrt{\rho^2/\gamma_{\min}} = \sqrt{19.262/\gamma_{\min}}$$

med 97% sandsynlighed.

Litteratur

Cederholm, P.: 2000, Udjævning, Forelæsningsnoter, Aalborg Universitet.

Hald, A.: 1948, *Statistiske Metoder*, Akademisk Forlag, København.

Pearson, E. S. and Hartley, H. O. (eds): 1972, *Biometrika Tables for Statisticians*, Vol. II, Cambridge University Press, Cambridge, UK.

Thomas, G. B. and Finney, R. L.: 1987, *Calculus and Analytic Geometry*, 7th edn, Addison-Wesley, Reading, Mass.

7 Opgaver

Opgave 1 Resultatet X af en afstandsmåling antages $\mathcal{N}(\mu, \sigma^2)$, hvor $\mu = 30m$ og $\sigma = 5cm$. Find

(a) $P(X \leq 30.05)$;

(b) $P(X \geq 29)$;

(c) $P(29.9 \leq X \leq 30.1)$

Opgave 2 Antag at vi har 4 uafhængige målinger af afstanden i opgave 1.

(a) Hvad er sandsynligheden for, at alle målingerne er over $30m$?

(b) Hvad er sandsynligheden for, at alle målingerne er under $29.9m$?

(c) Hvad er sandsynligheden for, at mindst en måling er over $30.15m$?

Opgave 3 Simuler en datamatrix med 1000 måleserier på hver 4 målinger med fordeling som angivet i opgave 1.

(a) I hvor mange af disse serier er alle målingerne over $30m$?

(b) I hvor mange af disse serier er alle målingerne under $29.9m$?

(c) I hvor mange af disse serier er mindst en måling over $30.15m$?

(d) Sammenlign disse resultater med svarene i opgave 2.

Opgave 4 Antag at vi har foretaget 1000 længdemålinger som er normalfordelt med hver sin middelværdi og samme spredning $\sigma = 5cm$.

Hvad er det forventede antal målinger, hvis afvigelse fra middelværdien overstiger

(a) $7.5cm$?

(b) $15cm$?

Opgave 5 En mand skyder til måls mod en cirkelformet skydeskive med radius $1m$. Antag, at det ramte punkts afvigelse fra centrum i vandret og lodret retning er uafhængige og normalfordelte med middelværdi 0 og spredning 1. Hvad er sandsynligheden for, at manden rammer skydeskiven?

Opgave 6 Antag, at a priori variansen i en udjævning er sat til 1.

(a) Hvis der er 20 frihedsgrader i udjævningen, hvad er så sandsynligheden for, at a posteriori variansen ligger mellem 0.8 og 1.2? Vink: 20 gange a posteriori variansen er χ^2-fordelt med 20 frihedsgrader.

(b) Samme spørgsmål som ovenfor, men med kun 10 frihedsgrader i udjævningen.

Opgave 7 Lad T være t-fordelt med 2 frihedsgrader. Hvad er sandsynligheden for, at T er større end 2? Større end 3? Hvad ville svarene være, hvis T i stedet var U-fordelt?

Opgave 8 Vis resultaterne (2) og (7).

Opgave 9 Lad X være en kontinuert stokastisk variabel, $a < 0$ og $b \in R$.

(a) Vis at

$$F_{aX+b}(y) = 1 - F_X(a^{-1}(y-b)). \tag{26}$$

(b) Slut heraf, at $f_{aX+b}(y) = a^{-1} f_X(a^{-1}(y-b))$.

(c) Vis udfra dette, at hvis $X \sim \mathcal{N}(\mu, \sigma^2)$, da gælder

$$aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2). \tag{27}$$

Opgave 10 Lad $\mu = 0, \sigma^2 = 1$ og $a = -1$. Vis ved hjælp af (26) og (27), at

$$\Phi(x) = 1 - \Phi(-x)$$

det vil sige, vi behøver kun tabellen $\Phi(x)$ for $x \geq 0$.

Opgave 11 Antag at $X_1, \ldots, X_n$ er uafhængige og $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$.

(a) Vis ved hjælp af (4) og (8) at

$$X_1 + \cdots + X_n \sim \mathcal{N}(\mu_1 + \cdots + \mu_n, \sigma_1^2 + \cdots + \sigma_n^2).$$

(b) Lad $\mu_i = \mu$, $\sigma_i^2 = \sigma^2$ for $i = 1, \ldots, n$. Slut ved hjælp af (27), at gennemsnittet $\bar{X}$ er normalfordelt $\sim \mathcal{N}(\mu, \sigma^2/n)$.

Opgave 12 Fra to forskellige udjævninger med henholdsvis 8 og 17 overbestemmelser fås a posteriori varianser $s_1^2 = 16.4$ og $s_2^2 = 4.1$. Derfor vil vi gerne vide:

(a) Hvad er sandsynligheden for, at s_1^2 er mere end 4 gange så stor som s_2^2?

(b) Hvad er sandsynligheden for, at den største af sådanne to a posteriori varianser er mere end 4 gange så stor som den mindste?

Opgave 13 En udjævning har givet en a posteriori spredning på $s = 2.7$. Beregn 95% konfidensintervaller for den sande værdi af spredningen i tilfælde af at udjævningen er udført med

(a) 10 overbestemmelser

(b) 20 overbestemmelser

(c) 100 overbestemmelser.

Opgave 14 En højdeforskel mellem to punkter er målt 9 gange med følgende resultater, angivet i mm:

$$110, 109, 110, 108, 111, 112, 109, 111, 112.$$

(a) Bestem et 95% konfidensinterval for højdeforskellen under antagelse af, at variansen er kendt og lig med a priori variansen 1.

(b) Bestem et 95% konfidensinterval for højdeforskellen baseret på a posteriori variansen.

(c) Bestem et 95% konfidensinterval for den ukendte spredning.

Opgave 15 Der er foretaget 15 aflæsninger på et stadie med et præcisionsnivellerinstrument. Aflæsningerne, der formodes at være en tilfældig stikprøve fra en normal population, er, angivet i mm:

1412.80	1412.85	1412.87
1413.09	1412.50	1412.80
1412.86	1412.84	1412.66
1412.80	1412.84	1412.84
1412.78	1413.02	1412.72

(a) Angiv 95% og 99% konfidensintervaller for målingens sande værdi.

(b) Angiv 50% og 95% konfidensintervaller for måleusikkerheden.

Undersøg følgende hypoteser:

(c) $H_0 : \mu = 1413.00mm$ mod $H_1 : \mu \neq 1413.00mm$.

(d) $H_0 : \mu = 1412.75mm$ mod $H_1 : \mu \neq 1412.75mm$.

(e) $H_0 : \sigma = 0.08mm$ mod $H_1 : \sigma \neq 0.08mm$.

(f) $H_0 : \sigma = 0.20mm$ mod $H_1 : \sigma \neq 0.20mm$.

Opgave 16 De nedenfor gengivne 2 måleræk̊ker angiver de afvigelser, der fremkom ved 25 gentagne målinger af vinklen 0 gon. En teodolit blev gang på gang indstillet i en retning, der var bestemt af en kollimator. Horisontalkredsen er aflæst for hver indstilling. Forskellen mellem to på hinanden følgende aflæsninger er da den observerede fejl på vinkelmålingen.

række 1	række 2	række 1	række 2
−13.5	+13.5	+0.5	+8.0
−2.0	−12.5	+14.5	+5.0
−6.0	+5.0	−1.0	−5.0
−3.0	−9.0	−3.5	+8.0
−12.5	−8.0	−1.0	−3.0
+2.0	+6.0	+4.0	−9.5
+13.5	−4.5	+10.0	+11.5
+2.5	−12.0	+7.5	+12.5
+5.0	+6.0	−7.5	−5.5
+3.5	−4.5	+1.0	−1.5
−12.0	−7.0	−4.0	+2.5
+5.0	−4.0	+1.5	−7.0
+2.5	+3.5		

(a) Udregn gennemsnit og spredning for hver måleræk ke.

(b) Er middelværdien nul eller er målingerne behæftet med systematiske fejl?

(c) Har de to måleræk ker samme varians?

(d) Har de to måleræk ker samme middelværdi?

Hvis der er svaret bekræftende på de to sidste spørgsmål, kan vi betragte tallene som hidrørende fra en og samme måleræk ke. Vi kan derfor igen spørge:

(e) Er middelværdien nul eller er målingerne behæftet med systematiske fejl?

(f) Beregn et 90% konfidensinterval for den eventuelle systematiske fejl.

Opgave 17 Vi betragter situationen i opgave 16, men antager iøvrigt at spredningen σ på afvigelsen mellem to på hinanden følgende vinkelmålinger er kendt og lig med 7.5 gon.

(a) Bestem teststørrelse og acceptområde ved test på niveau 5% for, om målingerne er behæftet med systematiske fejl. Måleækken har 50 observationer.

(b) Beregn styrkefunktionen for ovennævnte test.

(c) Hvor mange målinger skal man udføre for at kunne afsløre en systematisk fejl af størrelse 1 gon med en sikkerhed på 90%?

Opgave 18 En afstand er målt 10 gange. Alle målinger er uafhængige og har samme varians. A priori spredningen på hver måling vides at være $0.025m$. Man fik følgende værdier, angivet i meter:

307.532	307.500	307.474	307.549	307.490
307.527	307.556	307.502	307.489	307.514

(a) Find gennemsnit for denne måleække.

(b) Konstruer 50% og 95% konfidensintervaller for afstanden.

(c) Beregn a posteriori spredningen for måleækken.

(d) Er der overensstemmelse mellem a posteriori spredningen og a priori spredningen?

(d) Hvis spredningen på de enkelte målinger *ikke* kan antages kendt, bestem da 50% og 95% konfidensintervaller for den ukendte spredning under den forudsætning, at målingerne er normalfordelte.

Opgave 19 I Aalborg-området er der målt en del afstande til kontrol af de fra foto- og landmålingen afledte afstande. På grundlag af 36 dobbeltbestemmelser af afstande blev spredningen for de fotomålte afstande beregnet til $s = 4.7cm$. Landmålingsresultaterne er opdelt i perioden indtil 1966 og efter 1966, idet EDM blev almindeligt anvendt fra dette tidspunkt.

Periode	Målegruppe nr.	Frihedsgrader	Spredning (cm)
før 1966	1	145	11.0
før 1966	2	142	10.0
efter 1966	3	71	6.1
efter 1966	4	69	4.7

Giver disse resultater grundlag for at formode, at fotomåling er mere nøjagtig end landmåling?

Opgave 20 Resultaterne af satsudjævninger foretaget i 7 forskellige opstillingspunkter gengives nedenfor

St. nr.	# Retninger n_1	# Satser k_1	Frihedsgrader $f_i = (n_i - 1)(k_i - 1)$	Retningsspredning τ_i (mgon)
1	4	3	6	2.02
2	4	3	6	2.57
3	3	4	6	1.83
4	5	3	8	2.45
5	6	3	10	4.10
6	4	3	6	1.50
7	4	3	6	1.64

Undersøg ved hjælp af Bartlett's test om de enkelte τ_i^2 kun adskiller sig ved fejl af tilfældig natur eller om der virkelig foreligger signifikante forskelle. Nulhypotesen for dette test er:

$$H_0 : \tau_1^2 = \tau_2^2 = \cdots = \tau_7^2 = \tau^2.$$

Hvorledes ændres testet, hvis resultaterne fra punkt 5 udelades?

Opgave 21 De tilfældige fejl i den plane beliggenhed (x, y) af et punkt antages todimensionalt normalfordelt med følgende kovariansmatrix:

$$\begin{Bmatrix} 0.090\text{m}^2 & - & 0.096\text{m}^2 \\ -0.096\text{m}^2 & & 0.160\text{m}^2 \end{Bmatrix}.$$

Find halvakserne og orienteringen af punktets fejlellipse. Skitser fejlellipsen.

Opgave 22 Beregningen af en lukket polygon resulterer i følgende gab i x og y koordinaterne:

$$\begin{aligned} x_b - x_0 &= -3.3\text{cm} \\ y_b - y_0 &= 6.9\text{cm} \end{aligned}$$

hvor x_0 og y_0 er de givne koordinater i polygonens begyndelsespunkt. De beregnede koordinaters kovariansmatrix er:

$$\left\{ \begin{array}{cc} 4.63\text{cm}^2 & -0.87\text{cm}^2 \\ -0.87\text{cm}^2 & 8.83\text{cm}^2 \end{array} \right\}.$$

Er gabet acceptabelt?

Opgave 23 Lidt øvelse i normalfordelingens teori:

(a) Prøv at bevise (22).

(b) Overbevis dig om at (23) er rigtig.

Opgave 24 Lad A være en $n \times m$ matrix og B en $m \times q$ matrix.

(a) Vis at $(AB)^\top = B^\top A^\top$.

Antag at $n = m = q$ og at både A og B har en invers.

(b) Vis at $(AB)^{-1} = B^{-1}A^{-1}$

(c) Slut ved hjælp af (a) at $(A^\top)^{-1} = (A^{-1})^\top$.

Opgave 25 Antag at $X = (X_1, X_2, X_3)^\top \sim \mathcal{N}_3(\mu, \Sigma)$, hvor

$$\mu = \left(\begin{array}{c} 0 \\ -1 \\ 2 \end{array} \right) \text{ og } \Sigma = \left\{ \begin{array}{ccc} 2 & 1 & -1 \\ 1 & 4 & 2 \\ -1 & 2 & 5 \end{array} \right\}.$$

(a) Beregn middelværdi og varians af størrelserne

$$\begin{aligned} Y_1 &= \frac{1}{3}(X_1 + X_2 + X_3) \\ Y_2 &= \frac{1}{3}(X_1 + X_2 - X_3). \end{aligned}$$

(b) Hvad er korrelationen mellem Y_1 og Y_2?

Opgave 26 Antag at $X = (X_1, X_2, X_3)^\top \sim \mathcal{N}_3(\mu, \Sigma)$, hvor vægtmatricen P er givet ved

$$P = \Sigma^{-1} = \left\{ \begin{array}{ccc} 5 & 10 & 1 \\ 10 & 40 & -11 \\ 1 & -11 & 12 \end{array} \right\}$$

Antag endvidere, at vi har observeret værdien $(1.5, 0.9, 1.4)$ af $X^\top$. Er dette i overensstemmelse med hypotesen $\mu^\top = (1, 1, 1)$?

Opgave 27 Lad $\hat{\alpha} = (\hat{\alpha}_1, \ldots, \hat{\alpha}_m)^\top$ være estimatet for α, dvs.

$$\hat{\alpha} \sim \mathcal{N}_m \left\{ \alpha, (B^\top \Sigma_0^{-1} B)^{-1} \right\}.$$

Antag at $(\hat{\alpha}_1, \hat{\alpha}_2)$ er koordinaterne til et punkt i planen. Hvordan er fordelingen af $(\hat{\alpha}_1, \hat{\alpha}_2)^\top$? Dette er aktuelt ved konstruktion af fejlellipser.