

# **FFI RAPPORT**

## **DRIFTSKOSTNADSVEKST ESTIMERT MED BAYESIANSKE METODER**

DAHL Fredrik A.

**FFI/RAPPORT-2005/01676**



FFI-Analyse

Kjeller 14. april 2005

**DRIFTSKOSTNADSVEKST ESTIMERT MED  
BAYESIANSKE METODER**

DAHL Fredrik A.

FFI/RAPPORT-2005/01676

**FORSVARETS FORSKNING SINSTITUTT**  
**Norwegian Defence Research Establishment**  
Postboks 25, 2027 Kjeller, Norge







**INNHOOLD**

	<b>Side</b>	
1	INNLEDNING	7
2	DATA	7
3	MODELL	8
3.1	Tilfeldig støy i observasjonene	8
3.2	Tilfeldig støy i prosessen	9
3.3	Parametre	9
4	BAYESIANSK ANALYSE	10
5	UTFORMING AV APRIORIFORDELING	11
5.1	Prosesstøy	11
5.2	Observasjonsstøy	11
5.3	Driftskostnadsvekst	11
6	APOSTERIORIFORDELING	12
7	MODELLEVALUERING	13
8	VIDERE ARBEID	14
9	KONKLUSJON	14
	Litteratur	16





## DRIFTSKOSTNADSVEKST ESTIMERT MED BAYESIANSKE METODER

### 1 INNLEDNING

Forsvaret har i de senere år vært i mediens søkelys på grunn av økende kostnader. En kilde til de økte kostnadene antas å være personellrelatert drift, ved at kostnadene for hvert årsverk i Forsvaret øker. Den prosentvise årlige økningen i kostnadene for hvert årsverk, utover konsumprisindeksen, kalles personellrelatert *driftskostnadsvekst* (DKV).

Hensikten med arbeidet som dokumenteres i denne rapporten, er å kombinere innsamlede data om personellkostnader fra perioden 1994–2002 med ekspertvurderinger, ut fra en statistisk analysemetode. Analysen gir en sannsynlighetsfordeling for DKV, gitt data og vurderinger. Den bidrar dermed til å tallfeste hvilke DKV-verdier som har høy sannsynlighet, og hva som med rimelighet kan oppfattes som en øvre og nedre grense for DKV.

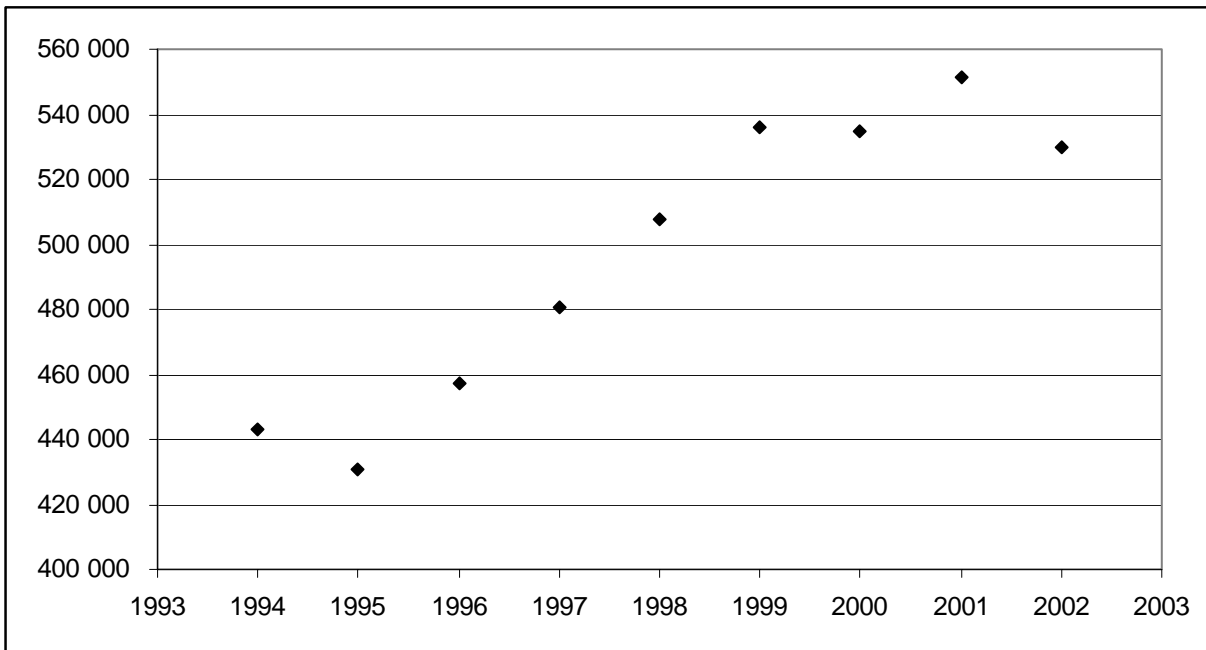
Hvorvidt analysen har relevans for framtidig kostnadsutvikling er bl. a. avhengig av om man lykkes med å omstrukturere Forsvaret, og dette sier analysen ingenting om. Vår vurdering er at analysen har klar relevans for de neste 5–10 årene, fordi en rekke tiltak for å redusere driftskostnadsveksten får virkning først på sikt. Man må også huske på at det har vært mye fokus på reduksjon av kostnader også i den perioden som analyseres (1994–2002), og det er ikke uten videre gitt at man vil lykkes vesentlig bedre i fremtiden.

Tidligere FFI-analyser av driftskostnadsvekst har vært basert på regresjonskurver (1) som implisitt bygger på en matematisk modellantagelse om tilfeldig støy kun i observasjonen av de virkelige kostnadene. I denne rapporten modelleres også eventuell tilfeldighet i den faktiske kostnadsutviklingen, i tillegg til at ekspertkunnskap inkluderes i analysen.

I kapittel 2 beskrives datamaterialet. Kapittel 3 beskriver og begrunner den valgte matematiske modellen for DKV. I kapittel 4 forklares prinsippene bak Bayesiansk analyse. I kapittel 5 beskrives arbeidet med å produsere apriorifordelinger fra ekspertkunnskap. Kapittel 6 gir aposteriorifordelingen for DKV, mens kapittel 7 presenterer en evaluering av modellen, og kapittel 8 oppsummerer konklusjonene fra analysen.

### 2 DATA

Tallmaterialet vi analyserer i denne rapporten er en tidsserie med ni datapunkter fra årene 1994–2002. Verdiene, vist i figur 2.1, er beregnede gjennomsnittlige kostnader for årsverk i Forsvaret, indeksregulert til 2002-kroner ved konsumprisindeksen (KPI).



Figur 2.1 Utviklingen i personellrelaterte driftskostnader per årsverk, i 2002-kroner

Det ligger et betydelig arbeid bak beregningen av verdiene (1), bl. a. fordi identifisering av personellrelaterte driftskostnader fra Forsvarets regnskaper ikke er triviell. Verdier før 1994 er ikke tilgjengelige pga vesentlige endringer i regnskapsføringen fra 1993 til 1994. Dataene i figur 2.1 ble benyttet i analysen i (1), med unntak av verdien for 2002 som ikke var tilgjengelig på det tidspunktet.

### 3 MODELL

Vi antar at kostnadsutviklingen har en underliggende eksponentiell trend. Dette betyr at kostnadene på lang sikt har en prosentvis årlig økning. Den grunnleggende sammenhengen vi leter etter er derfor på formen:

$$kost_{i+1} \approx (1 + DKV)kost_i \quad (3.1)$$

hvor  $DKV$  er rentesatsen på årsverkskostnadene, utover prisstigningen, og  $i \in \{1, 2, \dots, 9\}$  er gitt ved  $i = \text{årstall} - 1993$ . Vi betrakter logaritmen til kostnadene  $y_i = \log(kost_i)$ , som gir:

$$y_{i+1} \approx y_i + d \quad (3.2)$$

hvor  $d = \log(1 + DKV)$ . Transformasjonen fra  $kost_i$  til  $y_i$  forenkler de matematiske beregningene, og man kan naturligvis regne seg tilbake til kostnader ved å invertere transformasjonen:  $kost_i = \exp(y_i)$ .

#### 3.1 Tilfeldig støy i observasjonene

Gjennom arbeidet med grunnlagsdataene i figur 2.1 er det tydelig at det er usikkerhet rundt de faktiske årsverkskostnadene. Dette modellerer vi som observasjonsstøy, ved å la variabelen  $x_i$  representere logaritmen til den faktiske årsverkskostnaden for år  $1993 + i$ :

$$y_i = x_i + \varepsilon_i \quad (3.3)$$

Støyvariablene  $\varepsilon_i$  er uavhengige og identisk fordelte med forventning 0. Støyen er en sum av flere feilkilder, slik som variasjoner i postering av forskjellige kostnader, bruk av overtid og variasjon i øvingsmønster. Det er derfor rimelig å modellere observasjonsstøyen som normalfordelt med et visst (ukjent) standardavvik  $\sigma_y$ :  $\varepsilon_i \square N(0, \sigma_y^2)$ .

### 3.2 Tilfeldig støy i prosessen

Det er mange forhold i samfunnet som påvirker kostnadsutviklingen i Forsvaret, og siden vi ikke kan ta alle disse med i analysen, modellerer vi dem også som tilfeldig støy:

$$x_{i+1} = x_i + d + \delta \quad (3.4)$$

hvor støyvariablene  $\delta_i$  er uavhengige og identisk fordelte med forventning 0. Siden støyen er en sum av mange forskjellige forhold i samfunnet (f eks oljepris, terrorfrykt, rentenivå, boligmarked, etc) er det rimelig å modellere også denne som normalfordelt med et visst (ukjent) standardavvik  $\sigma_x$ :  $\delta_i \square N(0, \sigma_x^2)$ .

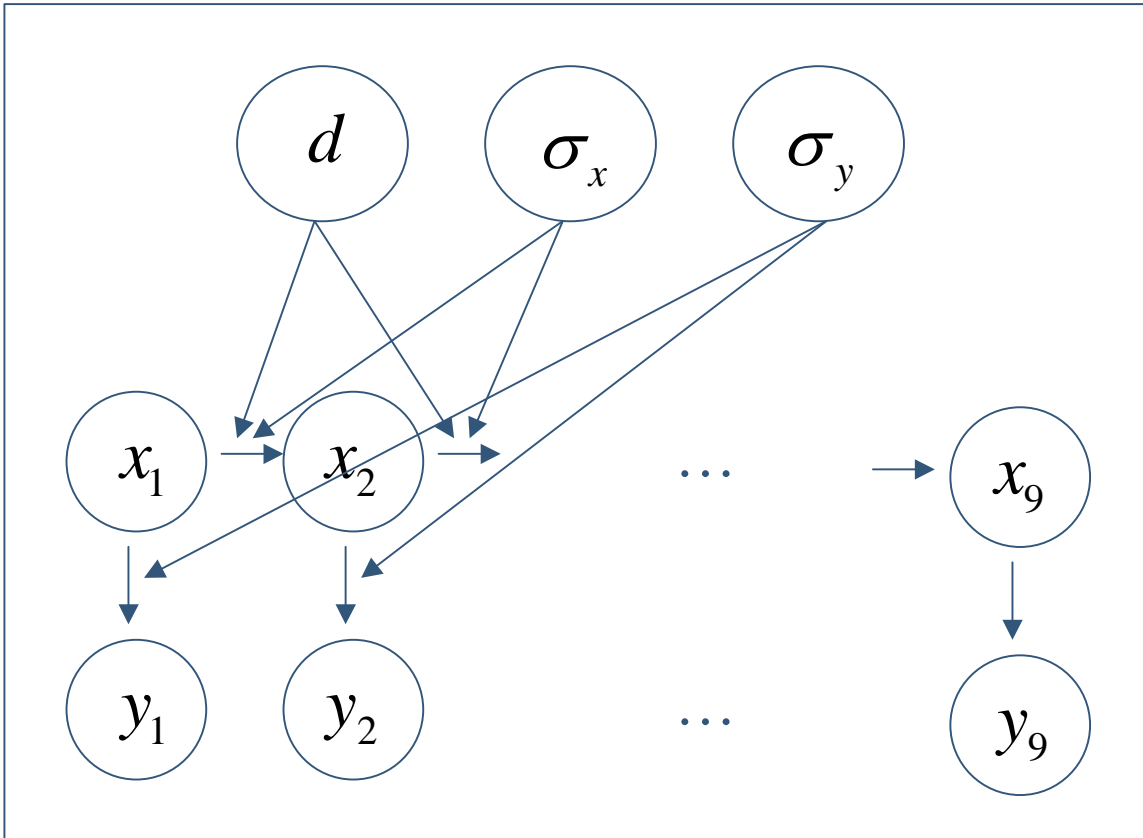
Legg merke til at denne modelleringen ikke ville vært rimelig uten logaritmetransformasjonen av dataene, fordi enhver normalfordelt variabel har positiv sannsynlighet for verdier under  $-(1+d)$ , som ville gitt negative kostnader. Negative verdier for den faktiske veksten i kostnadene er derimot ikke noen prinsipiell umulighet (dvs  $\delta_i < -d$  for et år  $i$ ), så lenge kostnadene er positive.

### 3.3 Parametre

Likningene (3.3) og (3.4) er vår matematiske modell med parameterne  $d$ ,  $\sigma_x$  og  $\sigma_y$ . Implisitt har vi også definert en parameter gjennom startverdien  $x_1$ . Imidlertid er modellen vår translasjonsinvariant, på den måten at  $(\tilde{x}, \tilde{y})$  gitt ved  $\tilde{x}_i = x_i + c$ ,  $\tilde{y}_i = y_i + c$ , hvor  $c$  er en konstant, har samme sannsynlighetstetthet som  $(x, y)$ . Derfor kan de tre primære parametrene analyseres uavhengig av  $x_1$ .

Tolkningen av parameteren  $d$  er den underliggende driftskostnadsveksten (eller strengt tatt logaritmen til  $1+DKV$ , men for aktuelle verdier av  $DKV$  er denne forskjellen liten).

Figur 3.1 gir en grafisk fremstilling av modellen. De horisontale pilene viser at årsverkskostnadene endrer seg fra ett år til den neste. Pilene fra parameterne  $d$  og  $\sigma_x$  indikerer at endringene i  $x$  påvirkes av disse parameterne. De vertikale pilene viser at (de logaritmetransformerte) dataene  $y$  er observasjoner av de tilsvarende  $x$ -verdiene. Pilene fra  $\sigma_y$  viser at størrelsen på den tilfeldige feilen i observasjonen er bestemt av denne parameteren.



Figur 3.1 Grafisk fremstilling av modellen

#### 4 BAYESIANSK ANALYSE

I vår anvendelse har vi få datapunkter i forhold til antall parametere som må estimeres. Samtidig har vi tilgang til betydelig domenekunnskap i form av generell økonomisk teori, kostnadsdata fra andre tidsperioder/samfunnsområder/land samt innsikt i støykildene i modellen gjennom arbeid med datainnsamling.

Bayesiansk analyse er en gren av matematisk statistikk som muliggjør integrasjon av harde data, slik som kostnadstall, med mer kvalitativ og vurderingsbasert kunnskap, og er derfor ideell for denne anvendelsen (2). I Bayesiansk analyse opererer man med en subjektiv sannsynlighetsfordeling, kalt *apriorifordeling*, på parameterne i modellen, og beregner den betingede sannsynlighetfordelingen til parameterne, kalt *aposteriorfordelingen*, gitt de tilgjengelige dataene. Dette betyr at man kan beregne sannsynligheten for at en parameter, f. eks.  $d$ , ligger innenfor et gitt intervall.

Vi tolker apriorifordelingen som en representasjon av vår samlede kunnskap om parameterverdiene før vi har sett på dataene. Fra denne tankegangen gir vi høy vekt til parameterverdier som vi har stor tro på, liten vekt til verdier som ville overraske oss, og null i vekt på verdiområder som vi oppfatter som helt uaktuelle. Apriorifordelingen er derfor en subjektiv sannsynlighetsfordeling fordi den tallfester vår oppfatning av hvor sannsynlige forskjellige parameterverdier er.

I mange praktiske anvendelser av Bayesianske metoder legger man liten vekt på utformingen av

apriorifordelingen, med argumentasjon om at data vil "vise vei" ved å styre aposteriorifordelingen inn mot de riktige parameterverdiene. Man legger da gjerne mest vekt på at man ikke utelukker noen muligheter, og utstyres parameterne med mer eller mindre flate (såkalt "ikkeinformative") apriorifordelinger. Vår anvendelse er imidlertid det man kan kalle "ren" Bayesiansk, ved at vi ønsker å bruke apriorifordelingen til å representere den kunnskapen vi faktisk har om domenet. Dette gjør vi fordi datamaterialet er svært begrenset.

## 5 UTFORMING AV APRIORIFORDELING

Arbeidet med å utforme apriorifordelingen ble utført i samarbeid mellom FFIs gruppe for forsvarsøkonomi (GFØ), forsker Else Helene Feet og forfatteren av denne rapporten. Gruppen diskuterte hver parameter for seg og antok uavhengighet mellom dem.

### 5.1 Prosesstøy

Parameteren for prosesstøy er  $\sigma_x$ . Størrelsen på denne gir standardavviket til den tilfeldige variasjonen i de faktiske gjennomsnittlige årsverkskostnadene. Hvis vi f. eks. har  $\sigma_x = 0.01$ , betyr det at den tilfeldige endringen av årsverkskostnadene fra ett år til det neste er av størrelsesorden ett prosentpoeng. Gruppen mente 0.05 kunne være en realistisk øvre grense for denne parameteren. Man hadde vanskelig for å tallfeste noen nedre grense, og endte opp med en uniform apriorisannsynlighet i området  $[0, 0.05]$ .

### 5.2 Observasjonsstøy

Parameteren for observasjonsstøy er  $\sigma_y$ . Størrelsen på denne gir standardavviket til den tilfeldige målingen vi gjør av årsverkskostnadene. På samme måte som for prosesstøy betyr en verdi  $\sigma_y = 0.01$  at den tilfeldige feilen vi gjør i målingen av kostnadene har standardavvik på ett prosentpoeng. I denne støyen inngår bl. a. usikkerhet i gruppering av kostnader som personell-relatert eller ikke. Gruppen mente denne støyen antakeligvis var større en prosesstøyen, med en maksimumsverdi på 0.10. Ut fra eget arbeid med dataene var gruppen også sikker på at observasjonsstøyen er på minst tre prosentpoeng. Det resulterende intervallet  $[0.03, 0.10]$  ble i likhet med prosesstøyen, også gitt uniform apriorisannsynlighet.

### 5.3 Driftskostnadsvekst

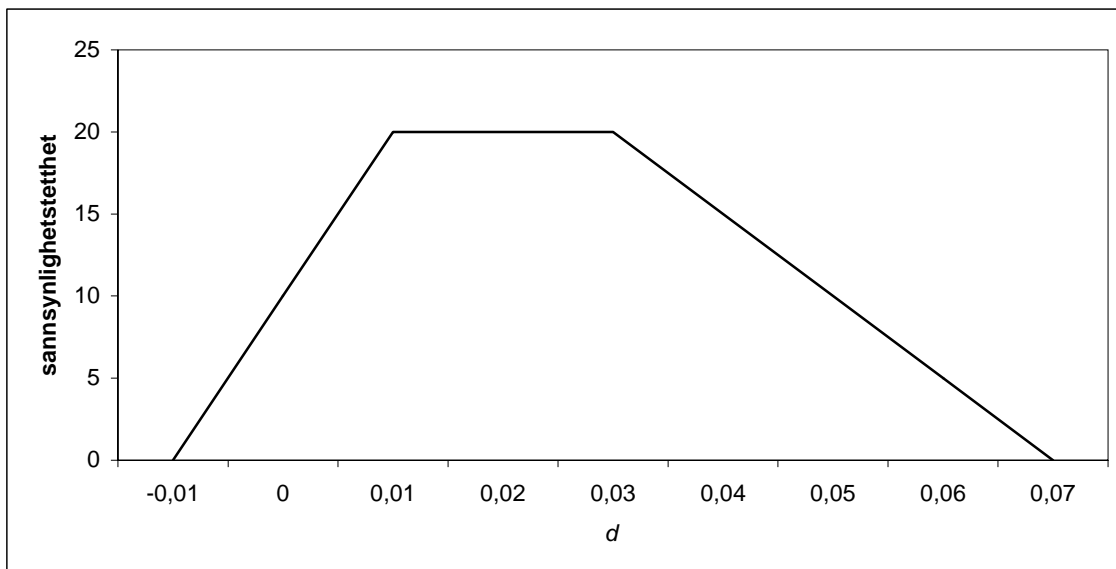
Driftskostnadsveksten  $d$  er naturligvis den viktigste parameteren i analysen, og aprioriantagelser om dennes fordeling vil normalt påvirke sluttresultatet mer enn apriorifordelingen på støyvariablene (som ofte omtales som *nuisance parameters* i statistisk litteratur). Man la derfor mer arbeid i modellering av apriorifordelingen til  $d$ .

Generelt i samfunnet er det en positiv velstandsutvikling som medfører at ansatte i de fleste bransjer får økt sin kjøpekraft selv om arbeidsinnsatsen holdes konstant, se Statistisk sentralbyrås estimater for økning i inflasjonsjustert brutto nasjonalprodukt (3). Dette er et utslag av at arbeid generelt rasjonaliseres og effektiviseres, bl. a. gjennom innføring av ny teknologi. Enten slik rasjonalisering lykkes i Forsvaret eller ikke, er det rimelig å tro at ansatte i denne sektoren vil forlange, og få, en viss vekst i reallønningene, som gir utslag i kostnadsvekst. Det er

derfor god grunn til å tro at parameteren  $d$  er positiv. Etter en lengre diskusjon kom man likevel fram til å tillate en svakt negativ verdi, ned til  $-0.01$ , med liten sannsynlighetsvekt.

Som øvre grense for fordelingen til parameteren endte man på tallet  $0.07$ . Gruppens oppfatning var at dette var en like ekstrem verdi som den nedre grensen på  $-0.01$ . Hvis spennet  $[-0.01, 0.07]$  virker smalt, må man huske på at dette er verdien til den underliggende trenden. Den faktiske kostnadsendringen fra et år til neste inneholder også et tilfeldig ledd med spredning gitt av  $\sigma_x$ , og den observerte kostnadsverdien har enda større spredning gjennom  $\sigma_y$ .

Gruppens oppfatning var at det mest sannsynlige området for driftskostnadsveksten er intervallet  $[0.01, 0.03]$ . Til sammenlikning forutsatte Forsvarsstudien 2000 (4) en  $DKV$  på  $1.5\%$ . Gruppen ble enig om å legge konstant apriorisannsynlighet på intervallet  $[0.01, 0.03]$ , og interpolere lineært ned til yttergrensene  $-0.01$  og  $0.07$ . Dette gir trapesfordelingen vist i figur 5.1.



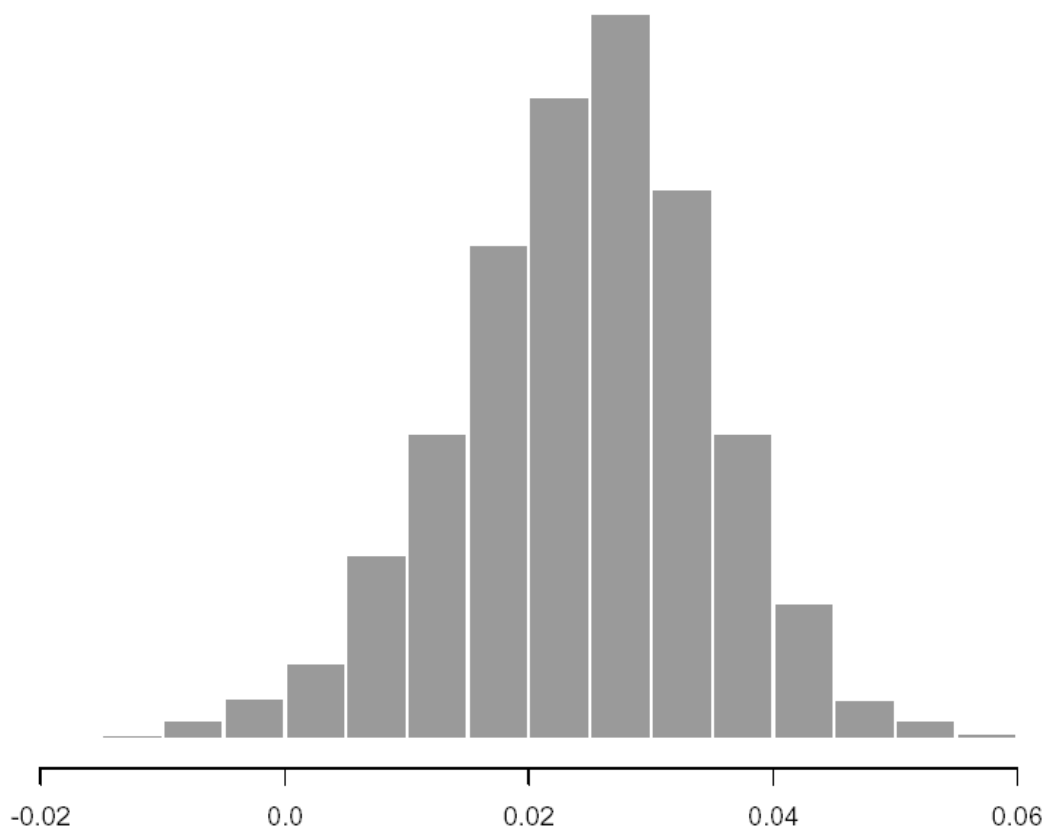
Figur 5.1: Sannsynlighetstetthet for apriorifordelingen til parameteren  $d$

Apriorisannsynligheter fremkommer som arealer under denne kurven. Maksimumsverdien er satt til 20 for at det samlede arealet under kurven skal bli 1.

## 6 APOSTERIORIFORDELING

Når apriorifordelingen og data er gitt, er aposteriorifordelingen på parameterne entydig bestemt, men beregningen av den er ikke triviell. Ved hjelp av Kalmanfilter (5) kan vi beregne likelihood-funksjonen for parameterne  $(d, \sigma_x, \sigma_y)$ , gitt  $y$ -verdiene. Med Markovkjede Monte Carlo simulering, forkortet MCMC i litteraturen, (6), kan vi så trekke et tilfeldig utvalg fra aposteriorifordelingen. Det er utenfor rammen av denne rapporten å beskrive den matematiske framgangsmåten i detalj, og det henvises til Metropolis-Hastings-algoritmen (6).

Histogrammet til den estimerte aposteriorifordelingen for  $d$  er vist i figur 6.1. Vi ser at aposteriorifordelingen er smalere enn apriorifordelingen, med et standardavvik på  $0.011$ , mot  $0.017$  for apriorifordelingen.



Figur 6.1 Histogram for aposteriorfordelingen til parameteren  $d$

Som punktestimater benyttes gjerne aposteriorforventningen, dvs. gjennomsnittet fra histogrammet i figur 6.1, og denne ligger på 0.024. Et 95 % troverdighetsintervall for  $d$  er gitt ved  $[0.002, 0.044]$ . Det betyr at det er 95 % sannsynlighet for at driftskostnadsveksten ligger mellom 0.2 % og 4.5 %. (Transformasjonen fra  $d$  til  $DKV$  betyr altså et tidels prosentpoeng for et øverste endepunkt i intervallet, mens forskjellen blir mindre enn avrundingsenheten for gjennomsnittsverdien og den nedre intervallgrensen.)

## 7 MODELLEVALUERING

Analysen og konklusjonene i foregående kapittel bygger naturligvis på forutsetningene som er gjort om modell og apriorifordeling, og det er derfor viktig å evaluere modellen. Vi benytter såkalt *posterior predictive p-value* (7), forkortet *ppp*. Dette baseres på et avviksmål  $D(y, \theta)$ , hvor  $y = [y_1, y_2, \dots, y_9]$  og  $\theta = [d, \sigma_x, \sigma_y]$ , som tallfester hvor overraskende dataene  $y$  er, gitt parameterverdiene  $\theta$ . Vi definerer følgende:

$$D(y, \theta) = \sum_{i=1}^8 \frac{(y_{i+1} - y_i - d)^2}{\sigma_x^2 + 2\sigma_y^2} \quad (7.1)$$

Hvert ledd i summen adresserer økningen av observerte kostnadsverdier fra ett år til det neste. Den forventede verdien for hoppet  $y_{i+1} - y_i$  er parameteren  $d$ . Telleren  $(y_{i+1} - y_i - d)^2$  gir derfor kvadratet av dette hoppets avvik fra sin forventningsverdi. Nevneren  $\sigma_x^2 + 2\sigma_y^2$  er variansen til

$y_{i+1} - y_i$ . Hvert ledd i summen gir derfor et mål på hvor ekstremt hoppet  $y_{i+1} - y_i$  er, gitt verdien på parameterne, og summen gir et mål på hvor ekstrem hele datavektoren  $y$  er. Observatoren  $ppp(y)$  tar verdier i intervallet  $[0,1]$ , hvor verdier nær 0 indikerer at dataene strider mot antagelser i modellen.  $Ppp$ -verdien kan sees som en generalisering av  $P$ -verdier brukt for hypotesetesting innen klassisk statistikk. For definisjonen av  $ppp$ , henvises det til (7).

I vårt tilfelle kom  $ppp$ -verdien på 0.61, som indikerer godt samsvar mellom modell og data. Generelt har  $ppp$ -verdier en tendens til å samle seg nær 0.5. For å korrigere for denne effekten er det i (8) definert en kalibrert versjon av  $ppp$ , kalt  $cppp$ , som har uniform sannsynlighet i intervallet  $[0,1]$ , gitt at dataene kommer fra den forutsatte modellen med den aktuelle apriorifordelingen. I vårt tilfelle fikk vi en  $cppp$ -verdi på 0.75, som også indikerer god overensstemmelse mellom modell og data.

Målsettingen med en slik analyse er å teste om dataene motsier modellen. Gjennom definisjonen av avviksmålet  $D$  fokuserer vi på om hoppene i de observerte  $y$ -verdiene avviker urimelig mye fra sin forventning. Man kan aldri bevise at en statistisk modell er korrekt, men når modellen i vårt tilfelle er så vidt godt faglig begrunnet og består testen vi har gjennomført, føler vi oss trygge på at den gir en brukbar beskrivelse av virkeligheten<sup>1</sup>.

## 8 VIDERE ARBEID

Ekspertkunnskapen i form at apriorifordelingen på parametrene tilfører analysen kunnskap, som påvirker resultatene. Dette er nettopp hensikten med å velge en Bayesiansk tilnærming. Når man har blitt enige om en slik fordeling, er det prinsipielt viktig å ikke endre på den for å få mer troverdige resultater. Likevel kunne det være interessant å kartlegge følsomheten av resultatene med hensyn på apriorifordelingen, ved å prøve ut forskjellige alternativer. På den måten kunne man tallfeste hvor stor påvirkning de subjektive sidene av analysen har. Med samme formål kan det også være aktuelt å sammenlikne resultatene våre med en frekventistisk (dvs ikke-Bayesiansk) analyse av modellen. I en slik analyse vil man estimere parametrene f. eks. gjennom maksimum likelihood-beregninger, og beregne konfidensintervaller<sup>2</sup> for disse.

## 9 KONKLUSJON

Vi har i denne rapporten beskrevet en Bayesiansk analyse av driftskostnadsvekst for personell i Forsvaret. Analysen kombinerer subjektive – men begrunnede – sannsynlighetfordelinger for modellparametere med kostnadsdata fra Forsvarets regnskaper. En matematisk evaluering viser godt samsvar mellom dataene og den forutsatte statistiske modellen, noe som øker vår tillit til analysen.

<sup>1</sup> Modellen er Bayesiansk og hierarkisk, og evaluering av denne typen modeller er arbeidsområdet til forskningsprosjektet "Evaluation of Bayesian Hierarchical Models" ved Avdeling for statistikk og forsikringsmatematikk under Matematisk institutt ved Universitetet i Oslo. Undertegnede utførte analysen og modelleringen som en del av dette prosjektet, i et fagmiljø som er i forskningsfronten innen Bayesiansk analyse generelt, og evaluering av Bayesianske hierarkiske modeller spesielt.

<sup>2</sup> Innen frekventistisk statistikk kan man ikke definere sannsynligheten for at en parameter ligger innenfor et intervall, fordi parameteren selv ikke er noen tilfeldig størrelse. Konfidensintervaller er funksjoner av data, og kan oppfattes som tilfeldige før dataene er gitt. Konfidensintervaller har derfor en annen tolkning enn Bayesianske troverdighetintervaller, men en tallmessig sammenlikning kan likevel være interessant.



Konklusjonen blir at den mest sannsynlige verdien for den underliggende trenden er en årlig personellkostnadsvekst pr årsverk utover KPI på 2.4 % i tidsperioden 1994-2002, og at denne vekstraten med høy grad av sikkerhet ligger mellom 0.2 % og 4.5 %. Til sammenlikning ga beregninger dokumentert i (1) verdien 3.8 %. Disse beregningene inkluderte imidlertid ikke kostnadene for 2002 hvor kostnadene hadde en klart negativ utvikling, samtidig som de inkluderte en liten korreksjon i forhold til kostnader for vernepliktige.

Det er et åpent spørsmål hvor langt inn i fremtiden en slik trend vil vare. På lang sikt vil antakelig endringer i den generelle utviklingen i norsk økonomi være avgjørende. På kortere sikt må man imidlertid ta hensyn til spesifikke forhold når det gjelder utviklingen av Forsvaret sammenliknet med andre samfunnssektorer. Man kan imidlertid benytte våre estimater som referanse, og justere opp eller ned avhengig av hvordan man mener framtidens utvikling vil skille seg fra den vi ser i dag. Vår vurdering er at analysen i alle tilfeller har klar relevans for de neste 4–8 årene, fordi en rekke tiltak for å redusere driftskostnadsveksten får virkning først på sikt. Man må også huske på at det har vært mye fokus på reduksjon av kostnader også i den perioden som er analysert (1994–2002), og det er ikke uten videre gitt at man vil lykkes vesentlig bedre i fremtiden.

**Litteratur**

- (1) GULICHSEN, Steinar (2002): *Driftskostnadsvekst i Forsvaret*, FFI/RAPPORT-2002/02999
- (2) GELMAN, Andrew, et. al. (2004): *Bayesian data analysis*, Chapman & Hall/CRC
- (3) Statistisk sentralbyrås hjemmesider: <http://ssb.no/>
- (4) Forsvarssjefens Forsvarsstudie 2000, sluttrapport
- (5) BØLVIKEN, Erik, CHRISTOPHERSEN, Nils, STORVIK, Geir (1998): *Linear dynamic models, Kalman filtering and statistics*, University of Oslo
- (6) GILKS, W.R., Richardson, S., Spiegelhalter, D.J. (1996): *Markov chain Monte Carlo in practice*, Chapman & Hall
- (7) GELMAN, A., MENG, X.L., STERN, H. (1996): Posterior predictive assessment of model fitness via realized discrepancies, *Statistica sinica*, 6 (4) 733-760
- (8) HJORT, Nils, DAHL, Fredrik, STEINBAKK, Gunnhildur H. (2005): Post-processing posterior predictive p-values, under vurdering av *Journal of the American Statistical Association (JASA)*