



Universitetet  
i Stavanger

DET TEKNISK-NATURVITENSKAPELIGE FAKULTET

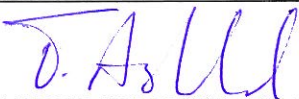
## MASTEROPPGAVE

Studieprogram/spesialisering:  
Master i teknologi, (siving)/ risikostyring

Vårsemesteret, 2012

Åpen

Forfatter: Tore Askeland

  
.....  
(signatur forfatter)

Fagansvarlig: Lasse Berg Andersen

Tittel på masteroppgaven: Subjektive sannsynligheter – overføring av kunnskap til sannsynligheter for bruk i risikoanalyser

Engelsk tittel: Subjective probabilities – converting knowledge into probabilities in Risk Analysis.

Studiepoeng: 30

Emneord:

Subjektiv, Sannsynlighet, Risikoanalyse,  
Ekspert

Eliciting, Probabilities, Risk analysis, Expert

Sidetall: 71

+ vedlegg/annet: 3

Stavanger, 15. juni 2012

## SAMMENDRAG

En risikovurdering er en systematisk kartlegging og evaluering av risiko. Vurdering av sannsynlighet inngår som en aktivitet i den delen av risikovurderingen som kalles risikoanalyse, og er en av de aktivitetene som i stor grad påvirker vurderingens resultat. Dagens praksis er ofte slik at det gjøres beregninger hvor man skal regne seg frem til det man tror er en "riktig" sannsynlighet, hvilket gir et skinn av objektivitet som ikke er riktig. Det tas heller ikke tilstrekkelig hensyn til de kognitive og motivasjonsmessige feilkilder som inntreffer i en slik prosess.

Med utgangspunkt i behov for sannsynligheter i en risikoanalyse var problemstillingen for denne oppgaven følgende; hvordan bør en prosess for å konvertere kunnskap til sannsynlighet være for best mulig å skape tillit hos en beslutningstager? Målene med denne oppgaven var følgende:

- Kartlegge forskningsfronten innenfor konvertering av kunnskap til tall til bruk i risikoanalyser, herunder kartlegge ulike prosesser som benyttes i dag for å konvertere kunnskap til sannsynlighet ved gjennomføring av risikoanalyser.
- Utarbeide kriterier for evaluering av tilgjengelige teknikker for å overføre kunnskap til sannsynligheter.
- Evaluere de kartlagte teknikker henhold til foregående punkt.
- Etablere en (forbedret) systematikk/ teknikk for å konvertere kunnskap til sannsynlighet.

Vi kartla forskningsfronten, og fant ut at det er utgitt en mengde bøker og artikler om konvertering av kunnskap til sannsynligheter, men at dette neppe er det mest utforskede området i statistikken. Et utvalg av den aktuelle litteraturen ble studert i detalj.

For å evaluere et sett med metoder som er representative for forskningsfronten ble det utarbeidet to evalueringskriterier; tillit og ressursbruk. For å ha tillit så vi at det må legges en subjektiv fortolkning av sannsynlighetsbegrepet til grunn, hvor sannsynlighet er ekspertens grad av tro på at hendelsen inntreffer eller ikke, og at bruk av en frekventistisk tilnærming gir en falsk illusjon om objektiv sannsynlighet som gjør at prosessen ikke inngir tillit. Videre må prosessen forhindre heuristikker og bias, og vi må være sikre på at eksperten faktisk er ekspert innen det område vi trenger ekspertkompetanse på.

Det ble avdekket at de fleste refererte prosessene har fokus på å håndtere heuristikker og bias, og på å skaffe reelle eksperter i vurderingene. Flere av forfatterne har imidlertid en beskrivelse av konverteringsprosessen hvor man legger en tilnærming til sannsynlighet som er frekventistisk, og hvor det skapes en illusjon av objektivitet rundt konverteringene.

Videre fant vi at det er hensiktsmessig å tilpasse metoden etter situasjonen og analyseobjektet. Dette vil sammen med å forebygge skjevfordelinger være med på å skape tillit til resultatet fra ekspertvurderingene. Vi fant at en mekanisk aggregering av både sannsynligheter og sannsynlighetsfordeling er enkelt, og kan benyttes dersom ikke aggregering skjer i en gruppe.

Denne oppgaven skulle bidra til mer viten omkring hvordan en prosess for å konvertere kunnskap til sannsynlighet for best mulig å skape tillit hos en beslutningstager skulle være. I stedet for å rangere metodene og kåre en "vinner", ble funnene benyttet til å lage et forslag til en ny "best practice". Denne er forholdsvis lik flere av de viste prosesser, men hensikten er at den på en tydeligere måte skal sikre at ekspertene er eksperter, at det tas hensyn analyseobjektet, og at heuristikker og bias håndteres proaktivt. En subjektivistisk forståelse av sannsynlighetsbegrepet, og hvor resultatet representerer ekspertens grad av tro, basert på den tilgjengelige informasjonen, er en forutsetning som ligger til grunn for prosessen.

## INNHOOLD

<b>SAMMENDRAG .....</b>	<b>2</b>
<b>1. INTRODUKSJON .....</b>	<b>5</b>
1.1 Introduksjon til temaet .....	5
1.2 Hensikt og mål med oppgaven .....	7
1.3 Metode og gjennomføring .....	7
1.4 Forkortelser.....	8
<b>2. KUNNSKAP OG SANNSYNLIGHETER I RISIKOANALYSER .....</b>	<b>9</b>
2.1 Norsk standard .....	9
2.2 ISO 31000 (Risk Management – principles and guidelines) .....	9
2.3 COSO .....	9
2.4 Sannsynligheter og sannsynlighetsfordelinger.....	10
2.5 Definisjoner.....	11
2.5.1 Kunnskap.....	11
2.5.2 Ekspert og ekspertvurdering.....	11
2.5.3 Eliciting.....	12
2.5.4 Heuristikker og bias.....	13
2.5.5 Risikoanalyser .....	14
2.5.6 Operasjonell risiko .....	15
<b>3. KRITERIER FOR EVALUERING AV KONVERTERINGSPROSESSER</b>	<b>17</b>
3.1 Ressursbruk .....	17
3.2 Tillit.....	17
3.3 Oppsummering kriterier .....	18
<b>4. KONVERTERINGS-PROSESS – ULIKE MODELLER .....</b>	<b>19</b>
4.1 Introduksjon .....	19
4.2 O’Hagan med flere .....	20
4.3 Meyer & Booker .....	23
4.4 Merkhofer .....	27
4.5 Clemen & Reilly .....	31
4.6 Vose.....	33
4.7 Encyclopedia of quantitative risk analysis and assessment .....	36
4.8 Terje Aven .....	39
4.9 Rausand og Utne .....	41
4.10 Oppsummering av funn.....	41
<b>5. DISKUSJON .....</b>	<b>44</b>
5.1 Sammenligning av metoder.....	44
5.1.1 Forutsetninger for prosess.....	44
5.1.2 Innhold i prosessen .....	45

5.1.3	Roller .....	48
5.1.4	Heuristikker og Bias .....	49
5.1.5	Aggregering .....	49
5.1.6	Om svært små sannsynligheter.....	52
5.2	Evaluering av ressursbruk og tillit .....	54
5.3	Oppsummering diskusjon .....	58
<b>6.</b>	<b>REVIDERT PROSESS TIL BRUK I RISIKOANALYSER.....</b>	<b>60</b>
6.1	Overordnet modell .....	60
6.2	Initiering.....	62
6.3	Design .....	63
6.4	Data innsamling/konvertering.....	64
6.5	Data analyse .....	66
<b>7.</b>	<b>KONKLUSJON .....</b>	<b>68</b>
<b>8.</b>	<b>REFERANSER.....</b>	<b>70</b>

## VEDLEGG

### VEDLEGG 1 – LINDLEYS PARADOKS

# 1. INTRODUKSJON

## 1.1 Introduksjon til temaet

En risikovurdering er en systematisk kartlegging og evaluering av risiko. Vurdering av sannsynlighet inngår som en aktivitet i den delen av risikovurderingen som kalles risikoanalyse, og er en av de aktivitetene som i stor grad påvirker vurderingens resultat. For å gjennomføre sannsynlighetsvurderingen benyttes sannsynlighetsregning, en ferdighet som må beherskes av risikoanalytikeren.

Sannsynlighetsregning defineres i Store Norske leksikon som "*den delen av matematikken som omhandler regning med statistiske sannsynligheter*". Sannsynlighetsregningen har sitt opphav i Europa på 15- og 1600-tallet hvor man forsøkte å finne gode regneregler til bruk i pengespill, og hvordan en gevinst skulle fordeles når man avsluttet et pengespill.

Gunnar Løvås (1999) beskriver i sin lærebok i statistikk og sannsynlighetsregning at det i hovedsak finnes 3 ulike måter å tallfeste sannsynlighet på:

1. Hvis alle utfall har like stor sannsynlighet for å inntreffe, kan vi resonnerer oss frem til en uniform sannsynlighetsmodell

$$P(A) = \frac{\text{Antall gunstige utfall for hendelsen } A}{\text{Antall mulige utfall for } A}$$

2. Hvis det er mulig, kan vi utføre mange forsøk og registrere relativ frekvens av hendelsen. Dette medfører en *frekvensbasert* sannsynlighet

$$P(A) = \frac{\text{Antall ganger hendelsen har inntruffet}}{\text{totalt antall forsøk}}$$

3. Hvis vi ikke kan gjøre som beskrevet i punkt 1 eller 2, må vi gi uttrykk for "*vår grad av tro*" om hvorvidt en hendelse inntreffer eller ikke. Dette kalles en *subjektiv* sannsynlighet.

Han beskriver at disse tre måtene gir opphav til ulike måter å fortolke sannsynlighet på. Såkalte "*frekventister*" vil tolke sannsynlighet som den relative frekvensen av en hendelse etter uendelig antall forsøk (hypotetiske eller praktiske). Såkalte "*subjektivist*" vil i følge Løvås alltid tolke en sannsynlighet som graden av tro om hvorvidt en hendelse vil inntreffe eller ikke.

I sannsynlighetsteoriens barndom var det den frekventistiske tankemåten som ble utviklet. Fremtredende bidragsytere i denne første perioden var spesielt italieneren G. Cardano som skrev sin bok "*De ludo aleae*" i 1564. Her gir han gode råd vedrørende spill og gambling, og forsøker å redegjøre for sannsynligheter for ulike utfall i spill. Han ga flere gode resonnementer innen spillteori, men lyktes imidlertid ikke med å bevise regneregler som understøttet teorien på en god måte (Hacking 2006, s 51).

Pierre de Fermat (1601-1665) og Blaise Pascal (1623-1662), begge fra Frankrike, brevvekslet og utvekslet ideer om sannsynlighetsregning (de brukte aldri ordene sannsynlighet i sin korrespondanse eller senere verk, men tekniske spillteoretiske uttrykk (Hacking, 2006, s 69)). Fermat er mest kjent for "*Fermats siste teorem*" som først lot seg bevise i 1995, mens Pascal er kjent for sin "*Pascals trekant*", som gir binomialkoeffisientene som er kjent fra kombinatorikken. Pascal regnes av flere som den viktigste av disse første sannsynlighetsteoretikere, da hans resonnementer var logiske, og han innså muligheten av at usikkerhet ikke nødvendigvis har med Guds påvirkning å gjøre, noe som til da var en fremtredende tanke (Hacking 2006, s 69-72).

Mens både Cardano, Fermat og Pascal beskrev sine løsninger ved bruk av brøker, bidro Jacob Bernoulli (1654-1705) fra Sveits med sitt arbeid til flere forandringer. Han begynte å angi sannsynligheter som tall mellom 0 og 1, og hans verk "Ars conjectandi" ble utgitt etter hans død, men regnes som et vesentlig bidrag til sannsynlighetsregningen. I dette verket beskriver han den til da kjente sannsynlighetsregningen med flere nye/ alternative bevis. Her redegjør han også for "store talls lov". "Store talls lov" sier at "dersom et forsøk gjentas mange ganger, vil den frekvensbaserte sannsynligheten nærme seg sin teoretisk riktige verdi", og er en begrunnelse for en frekvensbasert tilnærming hvor man altså mener at en gitt hendelse har en egenskap med en teoretisk riktig verdi. Uttrykket "*Bernoulli-forsøk*" stammer også fra denne boka, og brukes om forsøk som kan deles opp i  $n$  delforsøk, hvor hvert delforsøk kun har to utfall (A eller ikke A), hvor

$$p = P(A)$$

er lik i alle forsøk, og hvor alle delforsøkene er statistisk uavhengige av hverandre. Resultatet vil da ha en binomisk fordeling (Walpole, Myers, Myers, Ye (2007) s143-145).

Thomas Bayes (1702-1761), britisk statistiker, utviklet det som senere skulle bli kjent som "Bayes teorem", og ga også navnet til den del av statistikken som benevnes Bayesiansk statistikk. Både frekventister og subjektivistisk anerkjenner de samme bayesianske regnereglene, hvor en "*a priori*" sannsynlighet oppdateres i lys av ny kunnskap, nye erfaringer til en "*a posteriori*" sannsynlighet. Bayesiansk statistikk har flere definisjoner, men av flere (Løvås, Walpole, Myers) kobles den til subjektiv sannsynlighet, det vil si den sistnevnte måten å tallfeste sannsynligheter (av de 3 nevnt over).

Det var imidlertid Pierre Simon Laplace (1749-1827), fransk matematiker som i 1812 utga "Théorie analytique des probabilités, som først introduserte det som senere er blitt kjent som Bayesiansk statistikk (Myers et al, s 726), og som først benyttet dette i anvendt regning. Det skulle imidlertid vare helt til 1990 før gjennombruddet for Bayesianske anvendelsesmetoder ble oppnådd med økt forståelse og betydelig bruk av datamaskiner for å løse beregninger.

Italieneren Bruno de Finetti (1906-1985) var en av de viktigste predikanter for å se på sannsynlighet som en persons "grad av tro", og var i følge en annen statistiker i samme periode, Lindley, opptatt av aforismen "*Probability does not exist*" (Lindley 1986), da dette på en god måte viste hvordan en "riktig" eller "sann" sannsynlighet ikke finnes, men kun er uttrykket for en persons synspunkt - grad av tro.

Dennis Victor Lindley (1923-) er en britisk statistiker og kjent som en ledende tilhenger av bayesiansk statistikk. Han har gitt opphavet til Lindley's paradoks, et uttrykk for den situasjon man kan komme i ved en beregning som gjennomføres både med en frekventistisk og en bayesiansk tilnærming, men som gir ulikt resultat. Det er flere eksempler på at man kan komme i denne situasjonen som Lindley kaller et paradoks, og et eksempel er gjengitt i vedlegg 1.

Løsningene til de to fremgangsmåtene ser ut til å være i konflikt med hverandre, og det er dette som er "paradokset". Når statistikere ikke er enig om metode og resultat, hvordan kan en beslutningstaker ha tillit til risikoanalysen vi benytter sannsynlighetsberegningene i?

Risikovurderinger benyttes hyppig innenfor stadig flere bransjer og områder i samfunnet. Gjennom Basel II er det eksempelvis kommet krav om å gjøre risikoanalyser for finansielle institusjoner for å beregne operasjonell risiko og tilhørende kapitalkrav. Risikoanalyser benyttes stadig oftere som beslutningsgrunnlag ved prosjektering av infrastruktur i samfunnet, og det er krav om risikoanalyser som del av arealplanlegging for å evaluere om den nye planen medfører et økt eller redusert risikonivå for planområdet, og sikre at et eventuelt økt risikonivå håndteres i den videre planperioden.

I enkelte sammenhenger finnes databaser som kan gi opphav til generiske overordnede risikoverdier, men det er lite data innefor operasjonelle risikoer, slik som hvor ofte et underslag finner sted i en gitt bank, hvor ofte en ansatt har feil i sine regneark etc., og det er umulig å slå opp noe sted hvor ofte et ras skal gå i en fjellskrent i et gitt reguleringsplanområde.

Virksomheter og personer i/ nær analyseobjektet har ofte mye ekspertkunnskap, og god kjennskap til lokale forhold som kultur, holdninger, kunnskapsnivå og lignende som påvirker risikoen. Det er svært varierende i hvilken grad denne informasjonen systematiseres og gjøres tilgjengelig, om den er dokumentert eller om den kun finnes som taus kunnskap, samt i hvilken grad informasjonen overhodet benyttes i risikoanalyser. Ofte brukes ikke kompetansen, det gjøres beregninger hvor man skal regne seg frem til det man tror er en "riktig" sannsynlighet, hvilket gir et skinn av objektivitet som ikke er riktig, og det tas ikke tilstrekkelig hensyn til de kognitive og motivasjonsmessige feilkilder som inntreffer i en slik prosess.

Denne ekspertkunnskapen bør på et mer systematisk vis identifiseres, og benyttes til å overføre kunnskap til en sannsynlighet som uttrykk for usikkerheten knyttet til hva som vil skje i fremtiden, samtidig som man mer systematisk reduserer effekten av de kognitive og motivasjonsmessige feilkilder. Dette vil sannsynligvis øke troverdigheten og øke verdien av risikoanalysene for oppdragsgiveren/ beslutningstakeren.

Med utgangspunkt i behov for sannsynligheter i en risikoanalyse er problemstillingen for denne oppgaven følgende; hvordan bør en prosess for å konvertere kunnskap til sannsynlighet være for best mulig å skape tillit hos en beslutningstager?

## 1.2 Hensikt og mål med oppgaven

Hensikten med denne oppgaven er todelt:

- For det første skal eksisterende prosesser/ teknikker for å omgjøre kunnskap til tall kartlegges og evalueres.
- For det andre er det ønskelig å bidra til forbedret innsikt i temaet, og om mulig å utarbeide en (forbedret) systematikk for konvertering av kunnskap til sannsynligheter i kommende risikoanalyser.

For å nå denne hensikten ønskes det å oppnå følgende mål:

- Kartlegge forskningsfronten innenfor konvertering av kunnskap til tall til bruk i risikoanalyser, herunder kartlegge ulike prosesser som benyttes i dag for å konvertere kunnskap til sannsynlighet ved gjennomføring av risikoanalyser.
- Utarbeide kriterier for evaluering av tilgjengelige teknikker for å overføre kunnskap til sannsynligheter.
- Evaluere de kartlagte teknikker henhold til foregående punkt.
- Etablere en (forbedret) systematikk/ teknikk for å konvertere kunnskap til sannsynlighet.

## 1.3 Metode og gjennomføring

Arbeidet er gjennomført som en selvstendig studie våren 2012. Kunnskap er hovedsakelig blitt tilegnet ved hjelp av studie av faglitteratur, oppslagsverk i bokform og på Internett, samt uformelle intervjuer med personer som enten kjøper eller gjennomfører risikoanalyser.

Litteraturen som er gjennomgått er gjengitt i kildehenvisningen. Det er i hovedsak faglitteratur og artikler innen risikoanalyse og statistikk som er benyttet, men også et filosofisk verk omkring sannsynlighetsbegrepet.

#### **1.4 Forkortelser**

COSO – The committee of sponsoring Organizations of the Treadway Commission

DSB – Direktoratet for Samfunnsikkerhet og beredskap

ISO - International Organization for Standardization

NRC – Nuclear Regulations Commission

NS – Norsk standard

NTNU – Norges teknisk-naturvitenskapelige universitet

ROS – Risiko- og sårbarhetsanalyse

SIL – Safety Integrity Level

UiS – Universitetet i Stavanger



## 2. KUNNSKAP OG SANNSYNLIGHETER I RISIKOANALYSER

For å avdekke krav til en prosess som skal brukes i risikoanalyser er det ønskelig å gå gjennom relevante standarder/ veileder for å kartlegge hva de eventuelt beskriver om bruk og konvertering av kunnskap til sannsynligheter, og eventuelle føringer for håndtering av subjektive vurderinger. Videre gjennomgås definisjoner som er relevante for resten av oppgaven.

### 2.1 Norsk standard

NS 5814 ("Krav til risikovurderinger (2008)") har en forholdsvis spesifikk presisering av hva slags kompetanse som skal være til stede i en risikoanalyseprosess. NS 5814 beskriver i punkt 3.2.2. at den / de som deltar i risikovurderingen samlet skal ha:

- Kunnskap om og erfaring med bruk av risikoanalytiske metoder
- Kunnskap om analyseobjektet og aktuelle farer
- Kunnskap om samspillet mellom analyseobjektet og andre forhold, internt og eksternt
- Nødvendig kjennskap til alle relevante fag.

NS 5815 (Risikovurdering av anleggsarbeid (2006)") beskriver i pkt 4.3 "*Organisering av arbeidet*" at arbeidet skal ledes av en person med god kunnskap om og erfaring i risikovurdering. Videre sies det at arbeidet skal bestå av personer med nødvendig kunnskap og erfaring innenfor det aktuelle anleggsarbeid. Videre skal personer med kunnskap om de nærmeste omgivelser delta. Det er presisert at lederen av arbeidsgruppen skal påse at arbeidsgruppen har nødvendig kompetanse. Kvalitetssikring av risikovurderingen skal tillegges en kompetent person som ikke deltar i selve arbeidet.

Pkt 5.4 "*Vurdering av sannsynligheten for en uønsket hendelse*" beskriver grunnlag for fastsettelse av sannsynligheter. Subjektive sannsynligheter/ ekspertvurdering er ikke nevnt.

### 2.2 ISO 31000 (Risk Management – principles and guidelines)

I ISO 31000 punkt 5.4.2 "*Risk identification*" står det beskrevet at "*People with appropriate knowledge should be involved in identifying risks*". Det er ikke beskrevet hva slags kunnskap, eller på hvilket nivå kunnskapen skal være, så dette forutsettes identifisert fra analyse til analyse.

I punkt 2.19 "*Likelihood*" er det beskrevet at sannsynligheter skal være objektive eller subjektive, kvalitativ eller kvantitativ, og skal være beskrevet ved hjelp av kjente begrep som sannsynlighet eller frekvens i en gitt tidsperiode.

### 2.3 COSO

Kompetanse inngår som en del av "Internt miljø" i Cosos rammeverk. Kompetanse beskrives som "*den kunnskap og de ferdigheter som er nødvendige for å utføre oppgaver. Ledelsen beslutter hvor godt disse oppgavene trenger å utføres, og veier virksomhetens strategi og målsettinger mot planer for å sette dem ut i livet*". Videre sies det at "*blant de faktorer som bør vurderes når man skal fastsette nivå for kunnskap og ferdigheter er graden av skjønnsmessige vurderinger som kreves i den aktuelle jobben*".

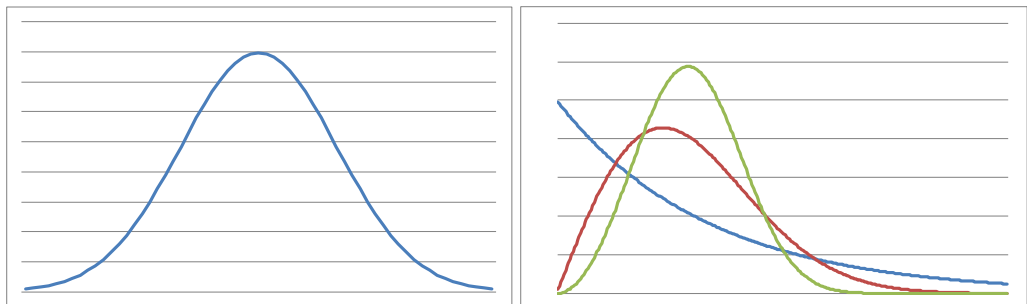
Coso beskriver at det er ønskelig med objektive data, og at objektive data er å foretrekke fremfor subjektive data. Som eksempel sies det at man for å finne en sannsynlighet undersøker frekvens av en hendelse internt, og supplerer denne med standardverdier i bransjen, og at dette muliggjør et mer presist estimat. Coso kommer ikke inn på ekspertvurderinger.

## 2.4 Sannsynligheter og sannsynlighetsfordelinger

Tittelen på oppgaven er "overføring av kunnskap til sannsynligheter for bruk i risikoanalyser". Ofte bruker vi eksperter for å komme ikke bare med separate sannsynligheter, men også for å forstå hvilke sannsynlighetsfordelinger som kan være aktuelt å bruke, siden ulike fordelinger på ulik måte er mer eller mindre gode til å beskrive hvordan en variabel opptrer. Kildene i boken skiller ikke konsekvent på ordene sannsynligheter og sannsynlighetsfordelinger (eller kun fordelinger), men disse brukes om hverandre ut fra hva som er hensiktsmessig i hver enkelt situasjon. Når det refereres til heuristikker hos en kilde, og deretter henvises til sannsynlighetsfordelinger, behøver ikke det å bety at dette gjelder kun når fordelinger identifiseres, men også enkeltsannsynligheter, og leseren bør ha i mente at disse ordene i noen grad brukes om hverandre.

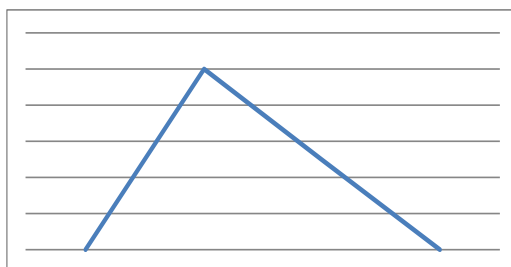
Innenfor fordelinger skiller vi mellom to kategorier sannsynlighetsfordelinger, parametriske og ikke parametriske fordelinger (Vose s 166).

- Parametriske fordelinger er basert på en matematisk funksjon hvis form og rekkevidde er bestemt av en eller flere fordelingsparametre. Disse fordelingsparametrene har liten åpenbar eller intuitiv sammenheng med den form på fordelingen de angir. Eksempler på parametriske funksjoner er lognormal, normal, beta, weibull, pareto, hypergeometrisk med flere.



Figur 1: Parametriske fordelinger; Normalfordeling (venstre) og Weibullfordelinger (høyre)

- Ikke- parametriske fordelinger har i motsetning til parametriske fordelinger en form og rekkevidde som er bestemt av sine parametre direkte, og i en intuitiv og selvfølgelig form. Fordelingen er forenklet sett en matematisk beskrivelse av deres form, slik som uniform, trekant, kumulativ og diskret.



Figur 2: Ikke-parametriske fordelinger; Trekantfordeling

## 2.5 Definisjoner

I det følgende er det lagt vekt på å definere de viktigste begrepene som omhandles i oppgaven, foruten ordet sannsynlighet som ble gjennomgått i introduksjonen.

- Kunnskap
- Ekspert
- Eliciting
- Heuristikker og bias
- Risikoanalyser
- Operasjonell risiko

### 2.5.1 Kunnskap

I pedagogisk ordbok (2006) defineres kunnskap som "den mengde informasjon og viten en person, gruppe, institusjon eller kultur rår over". Det skilles mellom kunnskap ervervet gjennom dagliglivets erfaringer, forskning, og gjennom studier av litterære eller andre kilder. Videre deles kunnskap inn i nivåer, der forståelse, evne til å analysere og kombinere, evne til å trekke slutninger og se sammenhenger regnes som kunnskap av høyere nivå, mens kjennskap til, og evne til å reprodusere fakta er på et lavere nivå. Den delen av kunnskap som ikke kan beskrives eller forklares gjennom instruksjoner kalles gjerne praksiskunnskap og inngår i begrepet kompetanse, men skilles gjerne fra kunnskapsbegrepet og kalles vanligvis ferdigheter.

Kunnskap inngår således i det utvidede begrepet kompetanse, noe som vi også finner hos Edgar Schein (2006) som kompetanse kan deles i kunnskap, ferdigheter og holdninger. Ferdighet defineres i pedagogisk ordbok som "en organisert rekkefølge av handlinger som er korrekt og/eller godt gjennomført, og som vanligvis ytrer seg i fleksible mønstre".

I "Managing technology in higher education" (2011) skilles det ikke like kategorisk mellom kunnskap og ferdigheter ("skills"), men ferdigheter inngår som en del av kunnskap, da det i stadig større grad er nødvendig med ferdigheter for å finne, og utnytte andres kunnskap. Ferdigheter sees på som en viktig del av kunnskap, eksempelvis ferdigheter i analyseteknikker, og eksempelvis dataverktøy for å finne frem i et stadig større kunnskapsvolum, da det ikke er gitt noen å kunne alt i fagområder som utvikler seg raskt. Kunnskapsområdet er under endring, og det hevdes at kunnskap ikke lenger genereres og valideres på utdanningsinstitusjoner alene, men at kunnskap i stadig større grad genereres av nettbaserte strømmer og diskusjoner mellom millioner av internettbrukere, en teori som kalles "connectivism" (Siemens 2004).

Det skilles i samme bok mellom to læreretninger innen pedagogikken, "objektivist" og "konstruktivist". "Objektivist" argumenterer med at det eksisterer en pålitelig mengde faktakunnskap, prinsipper, konsepter og teorier, og at disse kan bli bevist i den dag i dag. "Konstruktivist" definerer derimot kunnskap som subjektiv, etablert med bakgrunn i vår forståelse og felles forståelse (constructed from our perceptions and mutually agreed on conventions).

### 2.5.2 Ekspert og ekspertvurdering

Ekspertrollen står sentralt i denne oppgaven, og det finnes en rekke definisjoner på ordet ekspert.

- Spesielt dyktig eller sakkyndig person/ person med grundig kjennskap til et emne (Store Norske Leksikon).
- En person som har bakgrunn i det aktuelle området, og som er anerkjent blant sine likemenn eller de som gjennomfører en studie som kvalifisert til å besvare spørsmål (Meyer & Booker 1991).

- En som har eller påstås å ha overlegen kunnskap om/ erfaring med data, modeller og regelverk innenfor et spesielt tema (Encyclopedia of quantitative risk analysis and assessment, 2008). Det er i samme verk listet opp indikatorer som kan være til hjelp for å identifisere eksperter:
  - Vedkommende har publisert forskning og/ eller har mottatt stipend innenfor det aktuelle temaet.
  - Personen henvises til i andres arbeid/ forskning.
  - Har mottatt akademiske titler, priser eller andre typer anerkjennelse.
  - Person som anbefales av andre eksperter.
  - Har en relevant stilling.

Flere definisjoner kunne vært ramset opp, men de sier alle noe om god kompetanse innen et fagområde, tema eller objekt, og det er en lite kontroversiell definisjon.

De opplistede kriterier ovenfor bærer preg av at man er i et forum hvor fokus er på akademiske prestasjoner. Eksperter som benyttes for å overføre kunnskap til sannsynligheter trenger ikke være akademikere. Det vil i risikoanalyser ofte være svært nyttig å se på en persons kompetanse utover eventuelle akademiske prestasjoner, og uskolerte kan også være eksperter (Fischhoff, 1989). Dersom man gjør en risikoanalyse av en maskin, vil en person som har vært operatør på maskinen i 20 år være en nyttig ekspert på svært mange spørsmål knyttet til maskinen og bruk av denne. I det følgende må derfor ekspert ikke sees på som akademiker, men betraktes som en sakkyndig person med grundig kjennskap til et bestemt emne.

En ekspertvurdering er som navnet tilsier – en vurdering utført av en ekspert. Det kan være en geologs vurdering av sannsynlighet for ras i en fjellskråning i forbindelse med en ROS-analyse i en reguleringsplan, eller en mekaniker som kjenner driften rundt en ventil som man ønsker data på i forbindelse med en risikoanalyse på en oljeplattform.

Vurderingene det i denne oppgaven er snakk om er sannsynlighetsvurderinger som del av en risikoanalyse. Slike vurderinger inngår som en viktig del av de fleste typer analyser hvor andre målemetoder er umulig eller for kostbart (Meyer & Booker 1991).

Ekspertvurderingene kan foregå med en ekspert, eller det kan være flere eksperter som skal komme frem til en samlet vurdering, enten i fellesskap, eller hver for seg med påfølgende aggregering av resultatet.

### 2.5.3 Eliciting

Det engelske verbet "*elicit*" blir av clue 7.1, (2006) oversatt til å "*lokke, kalle fram, bringe fram*". Når det i artikler om ekspertvurderinger står "*eliciting process*" kan det derfor oversettes med "*frembringingsprosess*" eller "*fremlokkingsprosess*". Dette låter ikke så godt på norsk, da det kan høre ut som man lokker og lurer noen, noe som selvsagt ikke er tilfelle.

Den prosessen som handler om å få en ekspert til å gi sin vurdering i form av sannsynlighet vil kalles her en *konverteringsprosess*, da det handler om å konvertere kunnskap til et tall mellom 0 og 1. Denne forståelsen er også i tråd med O'Hagan et al (2006) som beskriver at respondenten har en form for kunnskap eller overbevisning i "sitt hode", og det er vår oppgave å konstruere hensiktsmessige spørsmål for å utvinne/ trekke ut denne informasjonen i form av sannsynligheter. For å gjøre språket mindre monotont vil det også benyttes ord som *kvantifisering* og *fastsettelse av sannsynlighet*.

#### 2.5.4 Heuristikker og bias

Heuristikker er definert i Store norske leksikon som følger: *“enkel fremgangsmåte eller strategi som en problemløser kan ta i bruk for å øke sjansen til å løse en oppgave. Eksempler er å se etter analogier, arbeide baklengs, eller tenke i middel-mål-relasjoner. Beslutninger og bedømmelser i dagliglivet baserer seg for en stor grad på heuristikker som f.eks. det å velge det kjente fremfor det ukjente, eller å foreta sammenligninger med en “forankrings-verdi”. Slike bedømmelser er effektive, men garanterer ikke riktig resultat og kan dermed føre til utfall som har systematiske skjevheter (bias)”*.

Heuristikker kalles ofte for tommelfingerregel for beslutninger - for rask beslutningsgjennomføring, men personer som bruker disse vil ofte være eksponert for systematiske feil. Disse systematiske feilene kalles bias, og fører til at resultatene av vurderinger får en helning eller slagside som sannsynligvis ikke er der uten heuristikken, og som anses som ikke-optimalt, ikke-rasjonelt eller ikke-konsistent (Mengel, 2011).

Den mest refererte og kjente beskrivelsen av heuristikker og bias er gjort av Amos Tversky, Paul Slovic og Daniel Kahneman (Judgment under uncertainty: Heuristics and biases, 1982). Her redegjør de for de ulike forhold som kan påvirke eksperter under en konverteringsprosess, slik de fant det under sitt forskningsprogram hvor de kartla ulike typer instinktive prosesser som mennesker tar i bruk under vurderinger hvor det er usikkerhet til stede. Forskningen var banebrytende, og har blitt stående som en vesentlig lærebok i mulige feilkilder i konverteringsprosesser, og hvor alle refererer til Tversky, Slovic og Kahneman.

I deres verk gjennomgås heuristikkene, og det er tre hovedgrupper heuristikker som oftest gjengis i litteratur omkring konverteringsprinsippene (de ble også presentert i en etter hvert kjent artikkel i tidsskriftet “Science” så tidlig som i 1974), og de gjennomgås derfor i det følgende:

- Representativitet (representativeness)
- Tilgjengelighet (availability)
- Ankringspunkt (referansepunkt) og justering (anchor and adjustment)

##### Representativitet

Med representativitetsheuristikken menes vår tilbøyelighet til å gi høyere sannsynlighet for at en person skal tilhøre en gitt gruppe, dersom personen ligner svært på den stereotypen vi har av personer i den gitte gruppen (Tversky, Slovic og Kahneman, 2006). Dersom man får vite at en person er stille, beskjeden, setter pris på orden og struktur og liker å lese, er det da mest sannsynlig at han har et yrke som bibliotekar, flyver, bonde eller butikkansatt? Mange vil gi bibliotekar høyest sannsynlighet, da beskrivelsen av ham passer godt med det bilde mange har av bibliotekarer.

Representativitetsheuristikken kan føre til alvorlige feil, da likhet eller representativitet ikke er påvirket av flere faktorer som bør påvirke bedømmelser av sannsynligheter. Denne heuristikken er også svært aktuell i bayesiansk regning, da ikke-sensitivitet i forhold til a priori sannsynlighet er en typisk feil som kommer innunder representativitetsheuristikken. Det fakta at andelen bibliotekarer er lavt i en befolkning bør tas hensyn til i den tidligere nevnte vurderingen om yrke.

Ikke-sensitivitet i forhold til utvalgsstørrelsen er en annen faktor som kan medføre feil, og som går innunder denne heuristikken. Mange unnlater å ta hensyn til at store utvalg vanskelig fraviker middelveien enn små utvalg (O’Hagan et al 2006). Denne feilen har fått navnet “små talls lov” som et ordspill på store talls lov (Tversky, Slovic og Kahneman, 2006).

Sammenblanding av betingede sannsynligheter kommer også innunder representativitet. Dersom man sammenblander  $P(A|B)$  og  $P(B|A)$  vil dette selvfølgelig få uønskede konsekvenser, da regnestykket blir feil.

#### Tilgjengelighet

Med tilgjengeighetsheuristikk menes vår tilbøyelighet til å vurdere sannsynlighet for en hendelse ulikt ut fra hvor lett vi klarer å få frem en tilsvarende hendelse i hukommelsen (Aven 2003). Hendelser hvor eksperten svært lett husker tilsvarende hendelser gis en høyere sannsynlighet enn dersom man ikke hadde hørt om en hendelse (O'Hagan et al 2006).

To risikoanalyser for Rørosbanen utført henholdsvis rett før og rett etter Åstaulykken i januar 2000, hvor den samme hendelsen ble vurdert, vil i følge denne teorien kunne fått to forskjellige sannsynligheter, da man i desember 1999 ikke hadde hatt andre vesentlige togulykker siden Trettenulykken i 1975, mens man i februar 2000 hadde Åstaulykken friskt i minne.

#### Ankringspunkt og justering

Dersom en analytiker angir et første forslag til en ekspert som skal vurdere sannsynlighet, viser forskningen til Tversky, Slovic og Kahneman at eksperten har en tilbøyelighet til å justere sitt tenkte anslag til en verdi nærmere det første forslaget som angis. Det samme er tilfelle i grupper der flere eksperter samles, og sannsynligheter diskuteres åpent. Det første forslaget kan virke som et anker, hvorfra de andre justerer sine forslag. Man kan anta at ankerverdien vil variere ut fra hvem som gir den første verdien, noe som påvirker resultatet. Som en følge av ankring og justering blir det færre ekstremverdier, og mindre variasjon enn man ellers ville hatt.

### **2.5.5 Risikoanalyser**

En risikovurdering er en systematisk kartlegging og evaluering av risiko. Ofte er hensikten med risikovurderinger å identifisere risiko, rangere denne, og legge grunnlag for en proaktiv håndtering av risiko i en virksomhet. På denne måten reduseres risikoen, og kostnader og andre uønskede konsekvenser av uønskede hendelser blir mindre.

En risikovurdering består av tre hovedelementer (Standard Norge 2008):

- Initiering/ planlegging
- Risikoanalyse
- Risikoevaluering

Risikoanalysens innhold består av:

- Årsaksanalyse
- Vurdering av sannsynlighet
- Vurdering av konsekvens

I årsaksanalysen blir årsaker til en gitt hendelse kartlagt, gjennom dette arbeidet fremkommer informasjon som kan benyttes for å vurdere hendelsens sannsynlighet, det vil si hvor ofte (sjelden) hendelsen kan tenkes å inntreffe.

Sannsynlighet for en hendelse uttrykkes med et tall mellom 0 og 1. Tallet gir uttrykk for usikkerheten knyttet til hvorvidt hendelsen kan inntreffe eller ikke. Ofte benyttes returperioder når svært små sannsynligheter skal kommuniseres.

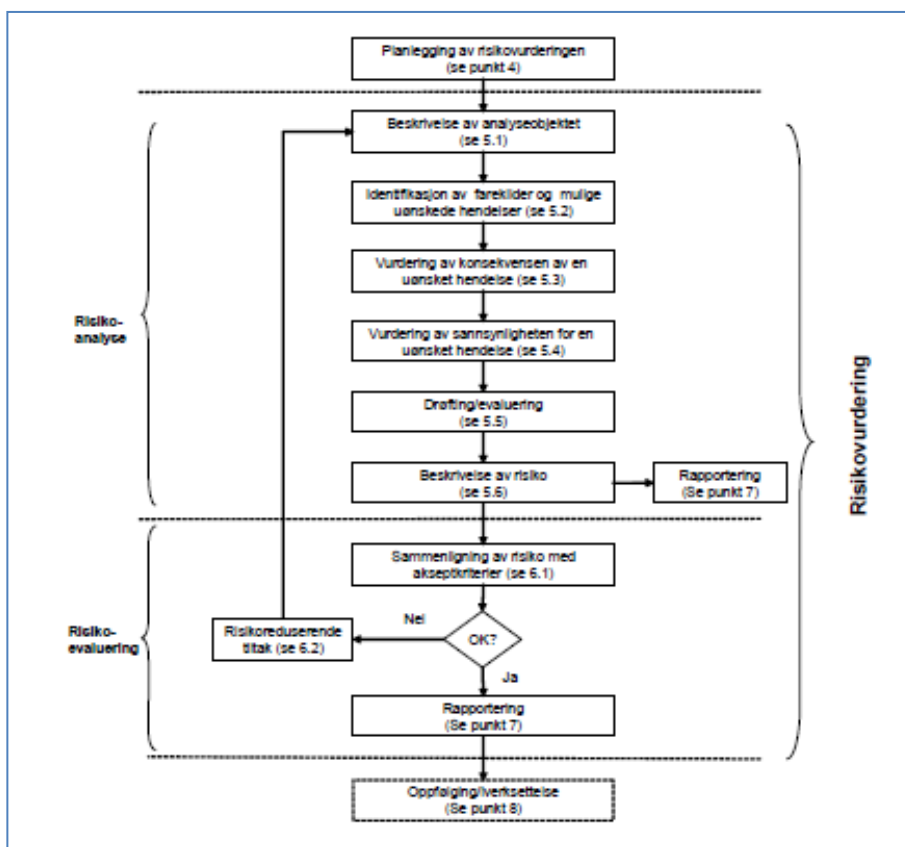
I vurdering av konsekvens vurderes det hvilken konsekvens hendelsen vil få, gitt at den inntreffer. Konsekvens kan oppgis i form av kostnader, men man kan også se på direkte

konsekvenser for liv & helse, miljø, eller andre parametre som er hensiktsmessig i forhold til risikovurderingens beslutningskriterier.

Det er flere standarder og veiledninger for utførelse av risikovurderinger, eksempelvis:

- ISO 31000 (2009) Risk management - Principles and guidelines
- ISO 17799 (2000) Information Technology - Code of practice for information security management
- NS 5814 (2008), krav til risikovurderinger
- NS 5815 (2006), risikovurderinger av anleggsarbeid
- NASA (2002) Probabilistic Risk Assessment Procedures Guide for NASA Managers and Practitioners
- NORSOK Z-013 (2001) Risk and Emergency Preparedness Analysis

I stor grad sammenfaller prosessene i risikoanalysene i de ulike risikoanalysene, og vurdering av sannsynlighet inngår som en aktivitet i risikoanalysen, se eksempel i figuren under.



Figur 3: Risikoanalyseprosessen slik den er gjengitt i NS5815 (Risikovurdering av anleggsarbeid)

### 2.5.6 Operasjonell risiko

En virksomhets risiko deles ofte inn i ulike typer, avhengig hvor enten hvor kilden til risikoen er, eller hvor alvorlig, og hvilken type konsekvenser en risiko medfører. Inndelingen varierer noe fra bransje til bransje, men det er vanlig å dele inn risiko minimum i følgende kategorier:

- Strategisk
- Finansiell
- Operasjonell

I finansielle virksomheter ser man innenfor finansiell risiko på kreditrisiko, markedsrisiko og likviditetsrisiko.

Det finnes flere definisjoner av inndelingen av risiko. I følge Aven, Røed, Wiencke (2008) kan de deles inn som nevnt i tabell under, men er nok da mest vanlig blant ikke finansielle virksomheter.

**Tabell 1: Definisjon risikokategorier (Aven, Røed og Wiencke)**

Strategisk risiko	Finansiell risiko	Operasjonell risiko
<p>Omfatter forhold som er viktige for virksomhetens langsiktige strategi og planer, eksempelvis:</p> <ul style="list-style-type: none"> <li>• Oppkjøp</li> <li>• Teknologi</li> <li>• Konkurrenter</li> <li>• Politiske forhold</li> <li>• Lover og regler</li> <li>• Arbeidsmarkedet</li> </ul>	<p>Omfatter virksomhetens finansielle situasjon og inkluderer blant annet:</p> <ul style="list-style-type: none"> <li>• Markedsrisiko, som er knyttet til priser på varer og tjenester, valutakurser, og verdipapirer (aksjer og obligasjoner etc.)</li> <li>• Kreditrisiko, som er knyttet til debitorers betalingsproblemer</li> <li>• Likviditetsrisiko, som er knyttet til virksomhetens tilgang på kapital</li> </ul>	<p>Omfatter forhold som påvirker den normale driftssituasjonen:</p> <ul style="list-style-type: none"> <li>• Ulykkeshendelser</li> <li>• Villedede handlinger: sabotasje, utro tjenere osv.</li> <li>• Tap av kompetanse/nøkkelpersonell</li> <li>• Juridiske forhold, for eksempel forhold knyttet til mangelfulle kontrakter og ansvarsforsikringer.</li> </ul>

Norges Bank definerer operasjonell risiko som følger:

”Risiko knyttet til faren for forstyrrelser og avbrudd av driftsmessig art, for eksempel brudd på prosedyrer, feil i IT-systemer eller maskinvarer, regelbrudd, bedragerier, brann og terrorangrep”.

Operasjonell risiko defineres i Basel II som:

”Risikoen for tap som følge av utilstrekkelige eller sviktende interne prosesser eller systemer, menneskelige feil, eller eksterne hendelser.” Basels II s definisjon legges til grunn for denne oppgaven.



### 3. KRITERIER FOR EVALUERING AV KONVERTERINGSPROSESSER

For å evaluere de ulike konverteringsprosessene senere i oppgaven, er det ønskelig å definere egnede kriterier. Disse kriteriene skal være sider ved en prosess som det i ettertid er enkelt å si seg enig i at enten er en til stede eller ikke i prosessenene.

For å identifisere aktuelle kriterier er dette benyttet litteraturgjennomgang, egen erfaring, samt at kollegaer som arbeider med risikoanalyser, samt beslutningstagere som kjøper tjenester innen risikostyring uformelt er spurt. Det er liten grad av uenighet eller usikkerhet omkring hvilke kvaliteter som er ønsket ved konverteringsprosesser, og dermed hvilke kriterier som kan egne seg til evaluering. Disse kan deles i to hovedgrupper:

- Ressursbruk
- Tillit til resultatet

#### 3.1 Ressursbruk

Med ressursbruk menes de ressurser som kreves for å gjennomføre konverteringsprosessen. Ressurser kan være i form av:

- Timeforbruk som vanligvis gjenspeiles i honorar til eksperter, analytikere etc.
- Tidsbruk, det antas en rask leveranse er mer verdt enn en leveranse på et senere tidspunkt.
- Andre fasiliteter, som laboratorier, utstyr, respondenter, møtelokaler, reiser etc. som påløper under gjennomføringen, og som medfører en kostnad for oppdragsgiver.

Det er ingen definert grense på hvor store ressurser som kan benyttes i en gitt prosess. Dette må styres av det faktiske behovet ut fra situasjonen, og en lønnsomhetsbetraktning hos en beslutningstaker. Beslutningstakeren kan gi føringer for hvor mye prosessen kan koste, og når resultatet må foreligge. Ofte vil slike praktiske begrensninger også gi føringer for designen på konverteringsprosessen. Man kan eksempelvis tenke seg ulike muligheter og to svært forskjellige måter å gjennomføre en prosess på avhengig av om resultatet må foreligge om 5 dager, med en økonomisk ramme på NOK 40 000,- eller om et halvt år, med en økonomisk ramme på en NOK 1 000 000,-.

#### 3.2 Tillit

Grad av tillit kan defineres som det omfang en person frivillig setter seg i et avhengighetsforhold til en annen person med en følelse av relativ sikkerhet, selv om negative konsekvenser er mulig (*"Trusting Behavior is the extent to which one person voluntarily depends on another person in a specific situation with a feeling of relative security, even though negative consequences are possible"* (McKnight og Chervany 1996)).

En beslutningstager investerer ressurser i en konverteringsprosess for å få støtte i form av ekspertvurdering(er) til en beslutning. Dersom han ikke har tillit til resultatet, vil han heller ikke bruke det som beslutningsgrunnlag, og nytten av arbeidet er således borte for oppdragsgiveren. Med tillit til resultatet menes således at oppdragsgiver må ha tro på at den gjennomføre prosessen har vært hensiktsmessig, og at resultatet som foreligger etter prosessen kan brukes som grunnlag for hans beslutninger.

Det er flere forhold som påvirker hvilken grad av tillit en oppdragsgiver har til resultatet:

- At det benyttes eksperter som faktisk er eksperter, og som vet hva de svarer på (Thomsen, 2012)
- At ekspertene fri for forutsigbare heuristikker og bias.
- At metodene som er benyttet anerkjent og etter beste praksis (Mengel 2011)

Marius Mengel (2011) beskriver at tillit vanskelig kan skapes til ekspertvurderinger hvor det legges en frekventistisk tanke til grunn, hvor man går ut fra at det finnes "sanne" verdier for sannsynlighet, og hvor usikkerhet er differansen på et estimat og den sanne verdien. Han oppsummerer med at "det skapes antatt tillit ved å presentere en illusjon om objektivitet i målingene og verktøyene som brukes, og ved å gi inntrykk av at strukturelle forsikringer i form av trening gir garantier for optimale vurderinger". En forutsetning for tillit er derfor at vi legger en subjektivistisk tilnærming til sannsynlighet til grunn, hvor sannsynlighet sees på som ekspertens grad av tro, og usikkerhet knyttet til hvorvidt en hendelse inntreffer eller ikke.

### 3.3 Oppsummering kriterier

Som en oppsummering kan vi således si at det er en fordel dersom en prosess kan gjennomføres med lite enn med store ressurser. Det forutsettes dog at det benyttes ressurser som er nødvendig for å få et hensiktsmessig resultat, men det er ingen øvre eller nedre cut-off verdi, da dette vil avhenge av situasjonen (kost-nyttebetraktning hos beslutningstager) og analyseobjektet.

I litteraturen er det lite fokus på ressursbruk, men det er selvfølgelig i oppdragsgivers interesse, og etisk riktig, at et arbeid gjøres så økonomisk fordelaktig for ham som mulig, selv om dette ikke alltid er en reell begrensning.

Når det gjelder tillit er det et absolutt krav om at beslutningstager har tillit til resultatet. I den grad man kan se på tillit som en binær tilstand, hvor tillit enten er til stede eller ikke er til stede, vil prosesser hvor beslutningstager ikke får tillit til resultatet være uegnede prosesser. For at det skal være tillit til en prosess må det legges en subjektivistisk sannsynlighetstolkning til grunn, prosessen må forhindre heuristikker og bias, og ekspertene må være faktiske eksperter.

## 4. KONVERTERINGS-PROSESS – ULIKE MODELLER

### 4.1 Introduksjon

Arbeidet med å konvertere kunnskap fra en eller flere personer (eksperter), og gjøre dette om til sannsynligheter, er en kvalitativ praksis, hvor resultatet kan benyttes i kvantitative beregninger (Encyclopedia 2008). Ekspertvurderinger inngår som en vesentlig forutsetning i bayesiansk statistikk, og ekspertvurderinger anbefales benyttet når det finnes lite relevant data, og når en analytikers vurderinger sannsynligvis kommer til å bli studert nærmere (Aven 2003). For å lykkes med bayesianske tilnærminger i risikoanalyser, og særlig innen operasjonell risiko der sannsynligheter kan være lave, er det derfor avgjørende at vi klarer å fasilitere og benytte ekspertvurderinger i vår konverteringsprosess på en god måte, slik at beslutningstager har tillit til det resultat han har fått.

Det er beskrevet flere teknikker opp gjennom årene. Frem til 80-tallet ble ekspertkunnskap benyttet for å gi input til feiltrær og hendelsestrær etc. Etter hvert som datamaskiner og software har utviklet seg videre med eksempelvis lettvinne Monte Carlo-løsninger som @Risk og Crystal ball, har ekspertvurderinger i større grad blitt benyttet til å vurdere spredning og resultater av fordelinger som resultat av simuleringer med ulike inputparametre.

Flere av prosessene er blitt evaluert i kontrollerte tester, men disse testene har ikke til nå avdekket en prosess som er mer "riktig" eller "brukbar" enn de andre (Encyclopedia of Risk). Det testene derimot har avdekket, er at ekspertene som deltar er tilbøyelige til å gi svar som kan være skjevfordelt og med mer eller mindre godhet. Godheten til en prosess vil derfor i stor grad være knyttet til hvor vellykket man klarer å håndtere disse heuristikkene. En god prosess vil derfor være avhengig av god kunnskap ikke bare om det emnet selve risikoanalysen dreier seg om, men også kunnskap om disse heuristikkene, om kommunikasjon og intervjuetknikker, foruten å forstå ekspertens behov og begrensninger i en slik prosess. I det følgende gjennomgås derfor teori knyttet til heuristikker og bias, før et knippe etablerte prosesser presenteres.

Det er skrevet en del om dette temaet de 3-4 siste tiår, men det er ikke det best dokumenterte område innenfor verken statistikk eller psykologi. Etter å ha søkt på Google, Amazon og i bibliotekbasen er det ca 30 bøker som er skrevet om emnet og tilstøtende områder. Noen er lite relevante da de har fokus på bruk av ekspertvurderinger i andre fagfelt, og ikke selve metoden i seg selv. Tilsvarende søk i artikler i bibliotekdatabasen viser omkring 60 relevante artikler. Mengden er overkommelig, men utfordringen initielt er å identifisere hvilke av disse som har størst relevans, og som i størst grad er benyttet i henvisninger i artikler. Det er nødvendig og gjøre et utvalg for å studere enkelte metoder i detalj.

Foruten gjennomgang av Tversky, Slovic og Kahneman (Judgment under uncertainty) for gjennomgang av heuristikker og bias, er de utvalgte kilder delt inn som følger:

- De mest kjente/ mest henviste (Uncertain Judgments – eliciting experts' Probabilities, O'Hagan et al, og Eliciting and analyzing expert Judgment – a practical guide, Meyer & Booker)
- Noen som ser ut til å vise en litt annerledes forståelse (Quantifying Judgemental Uncertainty: Methodology, Merkhofer, og Quantative Risk Analysis, Vose)
- Oppslagsverk og lærebøker (Encyclopedia of quantitative risk analysis and assessment, og Making Hard Decisions, Clemen & Reilly).
- Betydelige miljøer i Norge innenfor risikoanalyser (for å se hva som står om konverteringsprosesser hos disse) (Aven ved UiS og Rausand ved NTNU).

I kapittelet benyttes de uttrykk som forfatterne har benyttet, for eksempel usikkerhet, estimering, pålitelighet osv slik de har beskrevet disse. Det er altså ikke denne forfatterens ordbruk som benyttes dersom ikke det skrives, noe leseren bes ha i mente når han leser ulik bruk av ord, og ordbruk som ikke passer godt med en subjektivistisk metode. I kapittel 5 diskuteres det hvordan disse uttrykkene kan være med på å si noe om forfatterens syn på sannsynlighet, og om de angir en frekventistisk eller subjektivistisk holdning.

## 4.2 O'Hagan med flere

O'Hagan har sammen med sju andre personer skrevet "Uncertain Judgments – eliciting experts' probabilities". De definerer en ekspert i sin bok som rollen hvis mening skal utledes, uansett om han er en ekspert eller ikke på det faktiske analyseobjektet. Han påpeker så viktigheten av de som velges ut som eksperter faktisk er de riktige menneskene som skal utspørres, og at utvelgelsen av ekspert er et av de viktigste valg i en konverteringsprosess.

Kapittel 2 i O'Hagans bok omhandler hele prosessen med å forberede og gjennomføre en konvertering med en eller flere eksperter. Generelt er det to måter å samle informasjon på, og bruke denne informasjonen (de forutsetter flere eksperter i prosessen);

- Samle flere eksperter samtidig, og få gruppens mening, og behandle gruppens mening som en ekspertvurdering, en såkalt "behavioural aggregation", (ikke eget norsk ord, men kan kanskje oversette med "aggregering basert på adferd", siden den enkeltes adferd i en gruppeprosess vil bety mye for resultatet).
- Gjennomføre prosess med en og en ekspert, og deretter aggregere resultatet matematisk

Dersom antall eksperter blir svært høyt, kan det være hensiktsmessig å gjennomføre eksempelvis nettbaserte spørreundersøkelser. O'Hagan påpeker imidlertid at svakheten ved disse er at det er vanskelig å lage disse så gode at alle ekspertene tolker alle spørsmålene likt, slik at de svarer på de nøyaktig samme tingene. Et face-to-face intervju vil være best med en ekspert, og dersom face-to-face intervjuer gjennomføres ved bruk av flere eksperter gir det mulighet til å rette opp misforståelser umiddelbart. Ressursbruken ved bruk av slike metoder for mange eksperter vil imidlertid være svært store. O'Hagan mener imidlertid at en face-to-face prosess mellom fasilitator og ekspert er den beste.

O'Hagan viser til Wood og Ford som i 1993 beskrev fire ulike måter som skiller en eksperts tilnærming fra en amatørs tilnærming;

- Ekspertens kompetanse er basert på reelle tilfeller (ikke kun teori).
- Eksperten uttrykker seg ved hjelp av anerkjente uttrykk/ prinsipper
- Eksperten løser problemer med kjente strategier
- Eksperten stoler mindre på forklarende kunnskap (fakta) og mer på prosesskunnskap (forbindelser).

Videre påpeker O'Hagan at en ekspert er i stand til å fastsette og uttrykke usikkerhet omkring det aktuelle temaet på en god måte.

O'Hagan har klart definert 4 roller i en konverteringsprosess.

1. Beslutningstakeren, eller personen eller gruppen som etterspør resultatet av prosessen (kunden).
2. Ekspert på analyseobjektet, det vil si den som kjenner godt til usikkerhetene ved det objektet eller den prosessen som analyseres.
3. Statistikeren, en veiledende ekspert på statistiske spørsmål, som kan gi opplæring og veiledning innen sannsynligheter, og som kan gi validere resultater og gi tilbakemelding.

4. Fasilitatoren, en ekspert i å lede konverteringsprosesser, og som leder dialogen med eksperten.

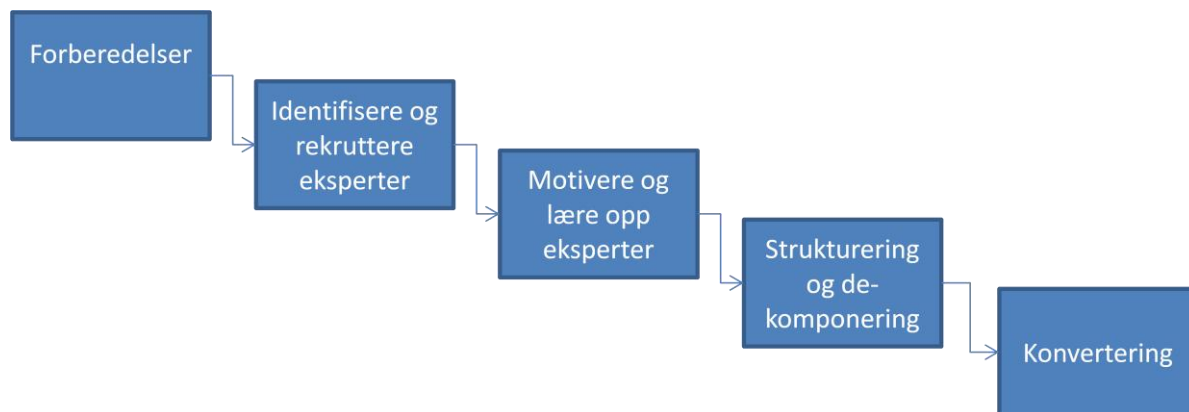
Han presiserer at en person kan ha flere roller, og at spesielt rollene som statistiker og fasilitator kan være samme person. Likeledes kan kunden være eksperten som kjenner analyseobjektet godt. Således kan to personer til sammen fylle de 4 definerte rollene.

O'Hagan kommer inn på hvorvidt en person kan fylle alle rollene, men mener dette er vanskelig. Han henviser til Savage (1971) som diskuterer utfordringene knyttet til å gjøre konverteringsprosessen alene, da man teoretisk kan tenke seg at en person har innsikt og mandat til å fylle alle roller. Han påpeker ulempene ved dette, og at uklarhet raskt blir et resultat av dette, og påpeker i de tilfeller et egnet sett med scoringsregler for å dokumentere resultatene og at man ikke "lurer" seg selv.

I kapittel 3 kommer O'Hagan inn på heuristikker og bias, og han henviser til Tversky og Kahneman og deres forskningsprogram vedrørende hva slags instinktive prosesser mennesker tar i bruk når de skal ta bedømmelser under usikkerhet. De definerer heuristikker som en strategi som benyttes for å løse slike konverteringsutfordringer som omtales i denne oppgaven. Se pkt 2.5.4.

O'Hagan definerer konverteringsprosessen i 5 aktiviteter:

1. Forberedelser (background & preparation)
2. Identifisere og rekruttere eksperter (identify and recruit expert(s)).
3. Motivere og lære opp eksperter (motivation and training the expert(s)).
4. Strukturering og dekomponering (structuring and decomposition).
5. Utlokking/konvertering (the elicitation).



**Figur 4:Prosesstegene i O'Hagans konverteringsprosess**

#### Forberedelser:

O'Hagan legger vekt på at den første fasen er helt vesentlig for at hele prosessen skal være nyttig, og for at resultatet av de enkelte konverteringsprosesser skal være egnet til den del av analysen hvor de inngår. Det er helt avgjørende at man klarer å identifisere hvilke variabler man trenger sannsynligheter for, og dermed hvilke områder man trenger eksperter på. Han påpeker at dette høres enkelt ut, men at det kan være en svært omfattende oppgave, da det også krever at man har kommet så langt at man har en viss formening om sannsynlighetsfordelinger, og om man eksempelvis ønsker a priori fordelinger i en bayesiansk analyse etc.

Han påpeker at dette er fasen hvor fasilitatoren/ statistikeren er i dialog med kunden, og antagelig sammen avklarer hvilke data som er nødvendig, og hvor fasilitatoren/ statistikeren har mulighet til å tilegne seg såpass med kompetanse om analyseobjektet at de kan kommunisere fornuftig med en ekspert på området, slik at de blir i stand til å lede prosessen, og veilede eksperten på sannsynlighetsvurderingene senere i prosessen.

#### Identifisere og rekruttere eksperter:

O'Hagan beskriver i liten grad hvordan eksperter skal identifiseres og rekrutteres. Han påpeker at kunden i mange tilfeller vil peke på hvem som er ekspert på området (ofte en som arbeider hos kunden). Videre viser de til Hora og von Winterfeldt (1997) som har arbeidet med temaet innenfor atomavfall. Der påpekes viktigheten av at prosessen med å velge eksperter gjennomføres slik at den kan være gjenstad for senere granskning, siden temaet er kontroversielt, og har stor offentlig interesse. De kom frem til seks kriterier for eksperter:

1. Konkret bevis på ekspertise
2. Omdømme
3. Tilgjengelighet og villighet til å delta
4. Forståelse for problemstilling
5. Upartiskhet/ objektivitet
6. Mangel på økonomisk eller personlig interesse i saken

Det påpekes at dersom det ikke er mulig å rekruttere eksperter som tilfredsstillende pkt 5 og 6, er det svært viktig å journalføre/ registrere enhver potensiell interessekonflikt som en ekspert måtte ha/ kan komme i.

#### Motivere og lære opp eksperter:

Det påpekes at det vil være fornuftig med å starte en opplæring/ motivasjon med å forklare viktigheten av å informere om hensikten med hvorfor ekspertene deltar, og hvordan resultatet (ekspertenes vurdering) skal brukes.

For å øke effektiviteten i konverteringen, anbefales det at eksperten bør ha en generell forståelse av tallord, og helst en viss kjennskap til sannsynlighetsregning. Det er trolig ubehagelig ikke å kunne uttrykke seg godt nok innenfor sannsynligheter for en som er ekspert på et område, og det anbefales derfor at eksperter som ikke har god kunnskap innen sannsynlighetsregning gis dette i forkant av konverteringsprosessen. Det anbefales derfor en opplæringsseanse i tre hoveddeler:

1. Sannsynlighetsregninger og sannsynlighetsfordelinger
2. Informasjon om de vanligste heuristikker og bias (forutinntatthet)
3. Øvelse i konvertering av kunnskap til sannsynligheter. Dette bør omfatte eksempler hvor det finnes "sanne" sannsynligheter, men hvor disse ikke er kjent for eksperten.

Vi legger merke til at O'Hagan legger til grunn at "sanne" sannsynligheter eksisterer, han gir ikke uttrykk for en problematisering rundt uttrykket "sanne" sannsynligheter når han bruker det.

#### Strukturering og dekomponering:

Enhver ekspert vil ha en egen oppfatning av avhengigheter og forholdet mellom disse, som kan påvirke sannsynligheten. Det er avgjørende for resultatet, mener O'Hagan, at man avdekker ekspertens syn på disse avhengighetene, og hvordan de ulike forholdene (årsakene) henger sammen, slik at dette kan tas høyde for dersom det gjennomføres konvertering med flere eksperter hver for seg, da disse sannsynligvis ikke har et ensartet syn på avhengigheter.

Kunden og statistikeren kan i forberedelsene forberede noe av dette, men det påpekes at eksperten ganske sikkert kan ha et avvikende syn på dette, og at han finner et foreslått årsaksbilde unaturlig eller direkte feil, og at ulike syn må identifiseres og håndteres på en hensiktsmessig måte.

#### Konvertering:

Selve konverteringen av kunnskap til en sannsynlighet, eller en fordeling, er siste aktivitet i prosessen. O'Hagan påpeker at det er en forutsetning at de foregående aktiviteter er gjennomført på en god måte for at selve konverteringen skal være tilstrekkelig god.

O'Hagan beskriver to ulike formater på konverteringen:

- Indirekte metoder
- Direkte metoder

Direkte metoder innebærer at eksperten selv svarer direkte, enten med å angi en sannsynlighet, eller med å angi svar på en for anledningen etablert skala (response scales). En responskala kan for eksempel være en visuell skala på et ark hvor  $p=0$  er den ene ytterkanten, mens  $P=1$  er den andre ytterkanten. Svaret kan angis direkte på linjen, hvoretter man måler distansen fra ytterpunktene for å finne sannsynligheten.

Skalaen kan også være oppdelt i klasser, hvor eksperten markerer sitt svar i en klasse på skalaen. Det er gjort flere studier omkring slike skalaer, og det kan tyde på at eksperten kan bli påvirket av en eventuell klasseinndeling, og at man får bias knyttet til at eksperten tror det er verdier som er mer forventet enn andre, og dermed påvirkes resultatet.

Såkalte sannsynlighetshjul (probability wheels) med skala er en annen måte å angi sannsynlighet på. En sirkel deles inn med flere skiver, ofte med ulik farge, slik at man kan angi en samlet sirkel inndelt med to farger som eksempelvis representerer henholdsvis tro på at hendelse inntreffer eller at den ikke inntreffer. Igjen måles sannsynligheten ut fra hvor stor del av hjulet som har hvilken farge ved hjelp av areal, omkrets etc.

Indirekte metoder er metoder der man ikke angir sannsynligheten direkte, men i stedet går omveien om andre metoder som eksempelvis veddemål. Et eksempel på veddemål kan være å spørre en ekspert om han hvilken av følgende to situasjoner han ville satset penger på; at hendelse A inntreffer, eller at han får en ener på et enkelt terningkast. Dersom han sier at han ville satset penger på hendelse A, har han indirekte sagt at han vurderer sannsynligheten for hendelse A til å være høyere enn  $P=0,167$ . Prosessen beskrives som en iterativ prosess der eksperten holder på til han er likegyldig med hvilken av løsningene som velges.

Konverteringsprosessene kan skje med en ekspert, eller ved flere eksperter. Ved flere eksperter kan disse fasiliteres hver for seg, eller i en gruppeprosess. O'Hagan beskriver også at man kan aggregere sannsynligheter på to ulike måter:

- Matematisk aggregering (ved separate eksperter)
- "Gruppeaggregering" (behavioral aggregation)

### **4.3 Meyer & Booker**

Mary A. Meyer og Jane M. Booker er forfatterne av "*Eliziting and analyzing expert judgment – a practical guide*". Den benyttes hyppig som henvisning i andre artikler og bøker, men var relativt vanskelig å få tak i i Norge. Boken er utgitt i 1991, og det er den utgaven som her er beskrevet. Den er imidlertid utgitt også i 2001. Meyer and Booker gjennomgår i detalj introduksjon til ekspertvurderinger med avklaringer og mulige fallgruver, bias etc., ulike prosesser for ekspertvurderinger, og analysemetoder.

Ekspert(er), analytiker, beslutningstaker og prosjektpersonell er rollene som benyttes hos Meyer og Booker. Datainnsamler er eksempel på en rolle som inngår i prosjektpersonell.

Meyer og Booker tar utgangspunkt i de aktiviteter en ekspert gjennomfører når han/ hun løser problemer. Mange tror at eksperter løser problemer hver dag, og at et intervju med en ekspert for å få tak i ekspertens syn på en sannsynlighet ikke er annet en problemløsning lik den eksperten gjør hver dag. Problemløsning er imidlertid ikke lett i seg selv, og en konverteringsprosess som her medfører ofte ytterligere byrder til den naturlige prosessen det er å løse oppgaver. Dette påpekes i det følgende.

I boken skisseres 4 aktiviteter som eksperten faktisk utfører når de sitter i en intervjusituasjon, for eksempel i en konverteringsprosess:

1. Forstå spørsmålet
2. Huske relevant informasjon
3. Gjøre seg opp en mening
4. Formulere og rapportere et svar

Når eksperten skal forstå spørsmålet må han/ hun forholde seg til en intervjuer som man ikke nødvendigvis har erfaring med og kjenner. Intervjueren kan ikke nødvendigvis så mye om det faglige som eksperten, og kan uttrykke seg annerledes enn det eksperten ville gjort dersom han/ hun skulle stilt de samme spørsmålene. Kanskje benyttes det andre ord enn det eksperten normalt bruker, og et annet språk kan benyttes.

Når eksperten skal huske informasjon, må eksperten tidligere ha mottatt og lagret relevant informasjon. Det er ikke slik at all relevant ligger ferdig lagret, klar til å prosesseres i en intervjusituasjon, sier forfatterne, men bilder, situasjoner etc. fremkaller en viss informasjonsmengde, og spørsmålene og situasjonen kan påvirke hva som hentes frem. Mennesker er også slik at de ofte husker informasjonen som understøtter en viss hypotese, mens man glemmer informasjon som er i konflikt med den samme hypotesen.

For å gjøre seg opp en mening, aktivitet 3, benytter mennesker heuristikker, de mentale snarveiene. Forfatterne nevner her Tversky og Kahneman og deres arbeid, ref pkt 2.5.4.

I aktivitet 4, å formulere og rapportere et svar, krever det at eksperten utarbeider et svar på en bestemt måte. Dersom man er bedt om å svare på et bestemt format, eksempelvis med ett tall mellom 0 og 1, blir eksperten nødt til å "oversette" sitt svar internt til det formatet.

Forfatteren går dypere for å forklare leseren om bias - skjevfordeling. De hevder at det er vanskelig å si objektivt at noe er skjevfordelt. Dersom noe er skjevfordelt, må det være skjevfordelt i forhold til noe. Dette noe må da sees på som riktig, eller som virkeligheten eller sannheten. Hva som er riktig er imidlertid ikke alltid objektivt, siden vi alltid omformer inntrykk fra syn, hørsel, følelse etc. til forståelse i hjernen vår. Disse oppfatningene kan imidlertid variere fra person til person.

Vårt minne av en hendelse, vår oppfatning av det som skjedde i fortiden, og vår forventning til fremtiden kan aldri være objektiv, da det uunngåelig skapes av et subjektivt og personlig bilde av hendelser og erfaringer (Meyer og Booker, 1991).

Meyer og Booker viser til to ulike syn på bias:

- Motiv bias – når eksperten bevisst eller ubevisst forsøker å tilfredsstillte intervjueren

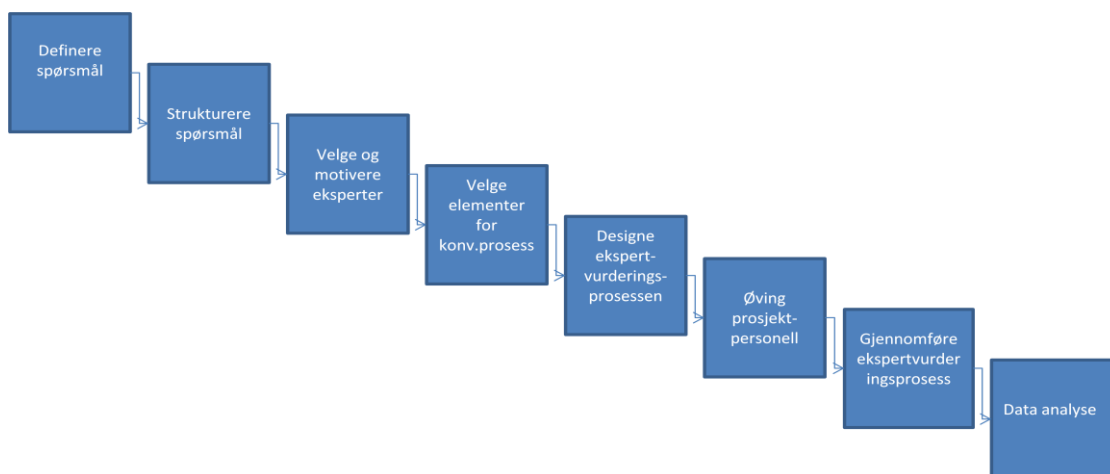


- Kognitiv bias - når eksperten ikke følger objektive regler og standarder/ ekspertens svar passer ikke med det som kan bli målt eller observert.

Forfatterne presiseres at det er hensiktsmessig med en plan for å identifisere og håndtere underveis i prosessen, og et program foreslås, se pkt 5.1.4.

Meyer & Booker skisserer en prosess med 8 aktiviteter:

- Definere av spørsmålsområder og konkrete spørsmål
- Strukturere spørsmål
- Velge og motivere eksperter
- Velge elementer for konverteringsprosess
- Designe selve ekspertvurderingsprosessen
- Øving prosjektpersonell
- Gjennomføre ekspertvurderingsprosess
- Data analyse



**Figur 5: Prosesstrinnene i Meyer & Bookers prosess**

#### Definere av spørsmålsområder og konkrete spørsmål

For det første må prosjektets hensikt og mål klarlegges. Dersom man ikke er sikre på hensikten med prosjektet må det klargjøres før man går videre. De tema som skal være gjenstand for konverteringsprosessen må kartlegges, og deretter selve spørsmålene innenfor hvert område.

#### Strukturere spørsmål

For å sikre god kvalitet på spørsmålene bør spørsmålene forbedres gjennom strukturering, det vil si å stille spørsmålene på en organisert og kontrollert måte. Å strukturere spørsmålene gir flere fordeler:

- Det fokuserer ekspertens oppmerksomhet på hva han faktisk skal levere
- Det letter den kognitive byrden av å løse problemene ved å presentere de på en mer fordøyet/ assimilerbar form
- Det avgrenser spørsmålet slik at eksperten ikke tolker spørsmålet annerledes enn tenkt
- Det gjør spørsmålet lettere å akseptere for eksperten fordi det er strukturert til å passe deres mening og bruke deres terminologi.

Måten å strukturere på kan variere fra sak til sak. Forfatterne presenterer flere teknikker for hvordan dette kan gjøres.

### Velge og motivere eksperter

Forfatterne skisserer to ulike anvendelsesområder for eksperter, de som kun skal brukes til å avgi et svar, og begrunnelsen for svaret, og de som skal brukes for å dokumentere en større problemløsningsprosess. Det påpekes at risikoanalyser typisk kommer innunder den første kategorien, hvor fokus er å få tak i ekspertens svar.

En ekspert anses her som en som har god kunnskap på det område som skal analyseres (substansiell kunnskap). Forfatterne påpeker imidlertid at eksperten også må ha kunnskap om sannsynlighetsregning (normativ kunnskap), for å kunne gi gode svar slik at det ikke svekker prosessen og dermed påvirker resultatet. Deres studier har vist at eksperter som bare kjenner et fagfelt (substansiell eller normativ), kan påvirke resultatet negativt, og det konkluderes med at en ekspert innenfor begge fagfelt er en bedre ekspert enn en som kun har substansiell ekspertise.

Om man bør bruke en eller flere eksperter innen et spørsmålsområde blir også diskutert. Det konkluderes med at flere eksperter er det optimale, da disse vil ha ulik erfaring, og man har muligheten til å se ting fra ulike synsvinkler. Det vises også til at resultatet etter aggregering fra flere eksperter gir et bedre resultat enn ved bruk av enkelt-eksperter, når resultatene blir sammenlignet med den "korrekte" verdien, når denne er kjent. Det sies at 9 eksperter er det maksimale av hva som kan håndteres i en gruppe hvor man ønsker at alle eksperter skal gi sin dybdevurdering, uten å komme i konflikt med potensielle effekter som kan komme av gruppedynamikk.

Det beskrives også at betaling av eksperter må være siste utveg for å gjennomføre et prosjekt hvor man søker eksperter til å delta. Det fremkommer ikke her om man mener utover normal lønn, eller utover honorar for rådgivere/ konsulenter etc. Forfatterne mener at man må søke å presentere prosjektet som så interessant og nyttig at ekspertene ønsker å være med, slik at man unngår å betale da dette blir dyrt. Det problematiseres ikke at man ikke får tak i de beste ekspertene dersom man ikke betaler, eller at markedet kan være slik at de ikke har tid til å stille på grunn av betalte oppdrag.

### Velge elementer for konverteringsprosess

Det listes flere elementer som de ansvarlige må ta stilling til for å etablere en god prosjektplan:

- Hvordan skal man kommunisere mellom eksperter og prosjekt, face-to-face, e-post, telefon
- Hvordan skal konverteringsprosessene skje, i grupper, en og en, eller med hjelp av delphiteknikk<sup>1</sup>?
- Hvordan skal svar avgis fra ekspert, muntlig eller skriftlig
- Hvordan skal prosjektet aggregere svar, matematisk eller i gruppe?
- Hvordan skal prosess dokumenteres, lydopptak, video, rapport?

Noe av dette vil være situasjonsavhengig, mens det for andre punkter kan være mer åpent.

### Design selv ekspertvurderingsprosessen

Denne aktiviteten er selv prosjektplanleggingen, hvor logistikk og kostnader må håndteres. Det påpekes at eksperter ofte er travle mennesker, opptatt med sitt daglige virke. Det vil være utfordringer knyttet til et design der man forsøker å samle alle samtidig, men hvor vanskelig og også hvor kostbart det er vil jo helt avhenge av antall personer som møtes.

---

<sup>1</sup> Delphiteknikken går ut på ekspertene hver for seg avgir hvert sitt svar på et gitt spørsmål. Svarene samles inn, anonymiseres, og deles ut til alle deltagerne. Deltagerne får mulighet til å justere sitt svar på bakgrunn av de svar man ser de andre har gitt. Teknikken kan brukes for å oppnå konsensus i en gruppe, eller som input i en iterativ prosess for å komme frem til et endelig svar fra hver ekspert.

Metodebruk og en detaljering av de områder man kom frem til i forrige aktivitet detaljeres. Utarbeidelse av de variabler man skal ha svar på, konkrete spørsmålsstillinger og svarskjema utarbeides.

#### Øving prosjektpersonell

Med ordet "øving" mener Meyer and Booker å holde prøve på/ gjengi, slik at man blir mer forberedt og mer profesjonell på det som skal gjøres med eksterne til stede.

Det presenteres flere forhold som bør øves av prosjektpersonellet:

- Orienteringer og briefers som avholdes til eksperter
- Prosedyrer for konverteringsprosessen, spesielt når det er andre enn de som har planlagt aktiviteten som skal gjennomføre den
- Aggregering
- Dokumentering med input av data i datamodeller

Det vises også til at det kan være behov for å øve gjennomføre tester med en liten gruppe eksperter til stede, for å få erfaring med det planlagte oppsett.

Videre kan det være behov for å gjennomføre opplæring/ øving med eksperter:

- Introdusere ekspertene til den planlagte prosessen og de praktiske forhold.
- Opplæring i heuristikker og bias.
- Hvordan eksperter skal benyttes eventuelle dokumenter/ programmer for å avgi svar.
- Hvordan aggregering skal skje dersom ekspertene involveres i dette.

#### Gjennomføre ekspertvurderingsprosess

Her går forfatterne veldig detaljert til verks, og beskriver eksempelvis selvfølgeligheter som hva man skal si når man ringer en ekspert for å avtale et møte, at man bør vite navnet på eksperten før et intervju etc.

Det gis mange råd og vink, og det skinner gjennom at forfatterne har fokus på prosjekter med mange eksperter, og en stor stab i sving for å gjennomføre konverteringsprosessen. Videre gis det et inntrykk av at eksperten er (bør være) en akademiker, men dette står ikke direkte.

#### Data analyse

I siste del av boken gjennomgås vanlige statistiske metoder for å analysere data.

En relevant del av denne delen av boken omhandler hvordan man kan detektere bias i en data analyse fase. Det påpekes at man kan gjennomgå data, og identifisere personer som har en motivasjons bias, ved å avdekke om enkelte alltid er mer positive eller negative enn andre. Dette kan gjøres ved hjelp av korrelasjonsanalyser.

## **4.4 Merkhofer**

Miley W. Merkhofer gir i sin artikkel "*Quantifying Judgmental Uncertainty: Methodology, experiences and insights*" en oppsummering av hans virksomhets prosess for å konvertere kunnskap til tallverdier, herunder sannsynligheter, eller "encoding probabilities" som er deres uttrykk, men også andre kvantitative verdier. Han begrunner behovet for en formell prosess rundt konverteringen med omfattende feil/ feiloppfattelser(bias) hos de enkelte eksperter har større mulighet for å bli oppdaget og korrigert en strukturert prosess, og at en systematisk aggregering er nødvendig for å få et resultat andre kan ha tillit til.

Merkhofer definerer klart to roller i sin prosess, men påpeker at det er gjennom felles (joint) aktiviteter fra følgende aktører:

- Eksperten (the subject) – person med omfattende kunnskap i det relevante fagfelt
- Analytikeren – personen som intervjuer og fasiliterer prosessen

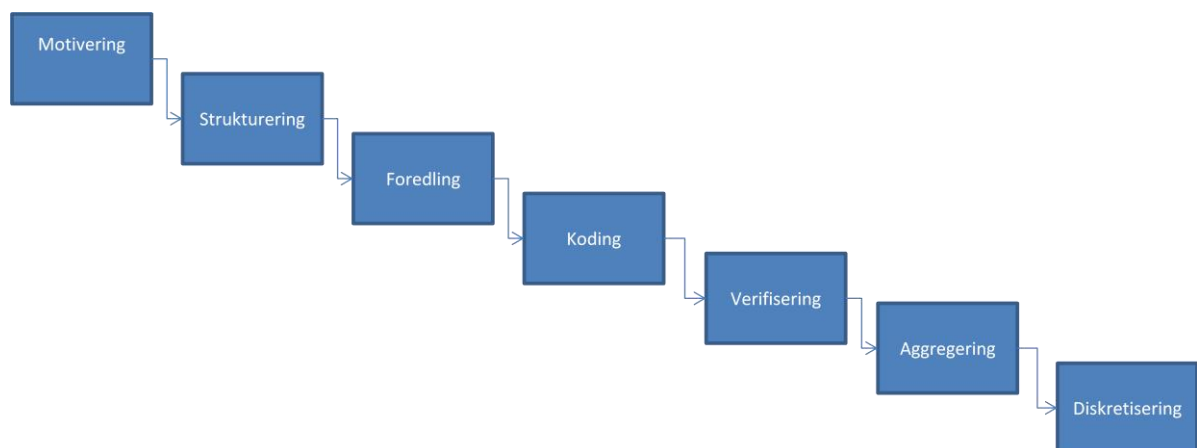
Han påpeker at det er ugunstig, selv om det skulle være teoretisk mulig, at en person fasiliterer seg selv. I stedet foreslås en strategi der analytikeren søker å forstå metoden eksperten bruker for å utlede sannsynlighetene, og identifisere eventuelle heuristikker/ bias som eksperten måtte ta i bruk, slik at disse kan håndteres/ tas høyde for i den videre prosessen.

Merkhofer kommer i sin artikkel inn på flere forhold som er interessante, og som kommer i tillegg til de tre tidligere nevnte heuristikker og bias.

- Merkhofer skiller i management-bias og ekspert-bias, hvor management-bias er det som inntreffer når en ekspert tror ledelsen vil høre et spesielt tall, i stedet for å identifisere usikkerheten omkring tallet og hva årsakene til usikkerhetene er.
- Noen eksperter underestimerer usikkerhet, fordi man har en forståelse av at andre forventer at en ekspert skal være "sikker", ikke gi uttrykk for usikkerhet.
- Dersom en ekspert har økonomisk interesse i den saken hvor informasjonen skal benyttes, vil eksperten være uegnethet på grunn av økonomisk påvirkning. Dette kan påvirke resultatet i svært stor grad, og det er viktig at slike interesser avdekkes under sleve identifiseringen av eksperter, og slik at disse ikke benyttes i prosessen, eller som minimum, at man er klar over ekspertens interesse, og kan lese resultatene i lys av dette.

Merkhofers prosess er delt inn i 7 aktiviteter:

- Motivering
- Strukturering
- Foredling
- Koding
- Verifisering
- Aggregering
- Diskretisering



**Figur 6: Merkhofer's prosess for konvertering av kunnskap til sannsynligheter**

#### Motivering

Hensikten med motivasjonsaktiviteten er i følge Merkhofer å etablere et tilstrekkelig nært forhold til eksperten, og undersøke om det er potensial for feil/ avvik (bias) på bakgrunn av ekspertens motivasjon, holdning, økonomiske situasjon etc. Merkhofer legger stor vekt på at eksperter med interesse i resultatet må identifiseres, og eventuelt fratras rollen.

Før selve konverteringen starter må eksperten få forklart modellen som benyttes, hvor sannsynlighetene skal inngå som inputverdier, og at det er viktig at eksperten uttrykker sin usikkerhet omkring de faktorer som påvirker sannsynligheten, og at poenget ikke er å komme frem til ent forhåndsdefinert tall som er ønsket av andre aktører.

Merkhofer anbefaler at man oppfordrer til at eksperten gir sine vurderinger skriftlig, da dette forhindrer sleivete/ unøyaktige vurderinger, og kan være med på å sikre at eksperten gjør mer overveide vurderinger.

### Strukturering

Struktureringsfasen har to hensikter:

- Sikre at de usikre forholdene som man ønsker sannsynlighetsvurderinger på er organisert på en hensiktsmessig måte som passer til ekspertens syn på årsaker etc.
- Undersøke hva eksperten tenker om de usikre forholdene mer generelt for å avdekke om han er under påvirkning av kognitive bias.

Merkhofer påpeker at fasen er viktig for analytikeren som skal søke å forstå hvordan eksperten tenker og uttrykker seg når han prater om de enkelte usikkerhetene, slik at kommunikasjonen mellom de to går lettere under konverteringen, og slik at man øker sannsynlighetene for at analytikeren forstår tankegangen til eksperten under de senere konverteringene.

Strukturering kan skje i 4 faser:

1. Klart presisere/ definere de variablene som skal bli vurdert/ konvertert. Et triks som Merkhofer anbefaler er å ta "synskhetstesten" før man godtar definisjonen av den enkelte variabel. Eksempelvis er forbruk pr mil på et kjøretøy i år 2020 ikke en variabel som går gjennom testen. Man må vite hva slags kjøretøy det er snakk om, hva slags drivstoff kjøretøyet bruker, hvor det skal kjøre etc. for å kunne gi et godt svar, slik at variabelen i dette tilfellet må presiseres bedre.
2. Undersøke muligheten for og nytten av å dele opp variablene i delvariabler, og angi verdi på disse. Dette vil ofte forebygge feilvurderinger, og er ofte nyttig for sannsynligheter, da man ved å tallfeste årsaker ofte reduserer en gitt verdi på sannsynligheten for en hendelse som kan ha flere årsaker (dette virker veldig basic når det gjelder sannsynligheter, men Merkhofer har som tidligere sagt også strukturert metoden for å angi usikkerheter i verdier, ikke kun sannsynligheter)
3. Liste alle forutsetninger/ antagelser som er tatt ved vurderingen av de enkelte variablene
4. Fastsette skala på de vurderinger som skal gis verdi. Dersom eksperten ønsker å bruke benevninger som fraviker analytikerens ønske, bør man velge den eksperten foretrekker. Denne kan så senere omgjøres etter behov.

### Foredling

Hensikten med fasen er i følge Merkhofer i hovedtrekk å klargjøre og umiddelbart å dra ut all tilgjengelig bevisst kunnskap som er relatert til variabelen fra eksperten, og undersøke om han baserer den på case-spesifikk eller generell kunnskap. Fasen benyttes også til å forberede eksperten på de feilkilder som kan påvirke resultatene i prosessen.

Vanligvis baserer eksperten seg i følge Merkhofer på to typer informasjon:

- Case-spesifikk informasjon, det vil si informasjon knyttet til den konkrete variabelen som vurderes
- Informasjon om fordelingen av verdiene til tilsvarende variabler

Forskning (Kahneman og Tversky, 1973 og 1982) har vist at eksperter ofte har en tendens til å undervekte informasjonen fra sistnevnte, den informasjonen som er tilgjengelig fra tilsvarende variabler, mens informasjon knyttet til den aktuelle variabelen overvektes. Dersom analytikeren mener at dette kan være tilfelle når informasjon skal innhentes, vil det være nyttig å informere eksperten om denne forskningen, slik at bevissthet omkring dette kan motvirke en eventuell ubevisst skjevvekting.

Videre henviser Merkhofer til psykologiforskning (Kahneman og Tversky, 1973 og Ross, 1977) og til fenomenet "regression toward the mean" hvor han beskriver at ting sjelden er så gode eller så dårlig som de først ser ut, og at tidligere nevnte feilen, overvekting av case-spesifikk informasjon, kan gi mer ekstreme resultater enn de burde. Dersom eksperten opprettholder ikke-regressive svar, anbefaler Merkhofer følgende prosess:

- Eksperten utspørres om hvilken sannsynlighet en ekspert vil gi dersom han/hun ikke har tilgang til case-spesifikk informasjon. Svaret gir a priori sannsynlighet til bruk i Bayes formel.
- Sannsynligheten som eksperten gir med case-spesifikk informasjon benyttes så for å beregne a posteriori sannsynlighet som i passende grad vektet både case-spesifikk informasjon og informasjon om fordelingen på tilsvarende variabler.

#### Koding (konvertering)

Denne fasen er selve kvantifiseringen av sannsynligheten, og Merkhofer deler i tre måter å gjøre kvantifiseringen på:

- Sannsynlighetsmetode – eksperten svarer med å oppgi konkrete sannsynlighetsverdier direkte
- Verdimetode – eksperten svarer ut fra verdier på en skala
- Sannsynlighets/ verdiskala – eksperten svarer med å kombinere sannsynligheter og verdier

Sannsynligheter kan angis direkte eller ved hjelp av odds. Forskning (Philips & Edwards, 1966, og Lusted et al, 1980) viser at eksperter med liten kvantitativ erfaring ofte foretrekker å bruke ulike oddsteknikker, mens eksperter med mer teknisk bakgrunn ofte foretrekker å bruke sannsynligheter.

#### Verifisering

Denne fasen har til hensikt å verifisere at eksperten virkelig tror på de sannsynlighetene han har angitt. Merkhofer beskriver en veddemålsmetode, der analytikeren lager kumulative sannsynlighetsfordelinger basert på ekspertens svar, og deretter gjør disse om til sannsynlighetsfordelinger. Totoppede fordelinger og eventuelle ekstremverdeier skal eksperten kommentere, og til slutt kontrolleres det at eksperten i et tenkt veddemål satser sine penger tilsvarende det resultat han/hun har kommet frem til. Dersom uoverensstemmelser eller feil identifiseres under verifiseringen, må man returnere til kodingsfasen, og gjenta denne.

#### Aggregering

Aggregering av resultater kan være hensiktsmessig når flere enn en ekspert gir sine resultater, og disse skal sammenstilles til en beslutningstager. Merkhofer angir to hovedtyper på aggregering:

- Gruppearbeid
- Mekanisk (matematisk)

I gruppearbeid kan flere eksperter komme frem til konsensus, men dette er en metode som Merkhofer fraråder. Resultatet av en interaksjon ansikt til ansikt i gruppesammenheng kan ha

flere uheldige sider, hvor noen er svært dominerende, mens andre er underdanige, og villige til å gi seg raskt. Gruppepress kan påvirke en eller flere til å endre syn, uten at påvirkning står i forhold til deres egenskap som ekspert på det aktuelle emnet.

Gruppearbeid kan imidlertid forbedre resultatet også, (Merkhofer henviser til Ashton & Ashton, 1985), og det kan føre til at deltagerne får innsyn i hva andre mener, slik at gruppen får et mer felles ståsted enn de har hver for seg (Merkhofer henviser til Seaver, 1976 og Rohrbaugh, 1979). Det siste kan være nyttig når gruppen skal arbeide videre sammen i flere problemstillinger.

Med en mekanisk aggregering benyttes enkle matematiske regler for å regne sammen gjennomsnitt av de enkeltes sannsynligheter. Flere studier (Merkhofer henviser til Brown, 1973 og von Holstein, 1972) viser at gjennomsnittsverdier er bedre enn de individuelle svarene. Enkle gjennomsnittsbetraktninger er det som er oftest benyttet, men det åpnes også for en vektingsmodell, der ekspertenes svar vektet i forhold til beslutningstakerens syn på deres kunnskap, eller basert på ekspertenes egenervering.

Merkhofer kommenterer også at en mekanisk beregning kan skje gjennom bruk av bayesiansk tilnærming, hvor sannsynligheter som er utledet fra eksperter er å anse som nyinnsamlede data, og kan benyttes til å beregne en a posteriori sannsynlighet. En utfordring med nesten all mekanisk aggregering er at de forlanger at ekspertene har lagt like forutsetninger til grunn, noe som er vanskelig å få til, og som sjelden er tilfelle.

#### Diskretisering

Siste steg i Merkhofers metodikk er diskretisering av data, hvor de innsamlede sannsynlighetene er dataene, som ofte kan være kontinuerlige fordelinger, diskretiseres, slik at de uendelig mange tallene som er identifisert i kontinuerlige fordelinger reduseres til et endelig antall sannsynligheter som kan analyseres og håndteres på en mer lettvinn måte.

## **4.5 Clemen & Reilly**

Clemen & Reilly har skrevet boken "Making hard decisions" som omhandler beslutninger og relevante verktøy (Decision Tools), som også UiS tilbyr undervisning i gjennom faget "Beslutningsanalyse". Forfatterne beskriver at et frekventistisk syn ikke alltid er tilstrekkelig til bruk i risikoanalyser, da ikke alle ønskede sannsynligheter kan gjennomføres i forsøk slik at man oppnår en frekvens av en hendelse. Til det trenger man en subjektiv forståelse av sannsynlighet, som kan beskrives som grad av tro av en hendelse skal inntreffe. Forfatterne legger således en såkalt subjektivistisk fortokning til grunn i sitt arbeid.

Clemen & Reilly har definert følgende roller:

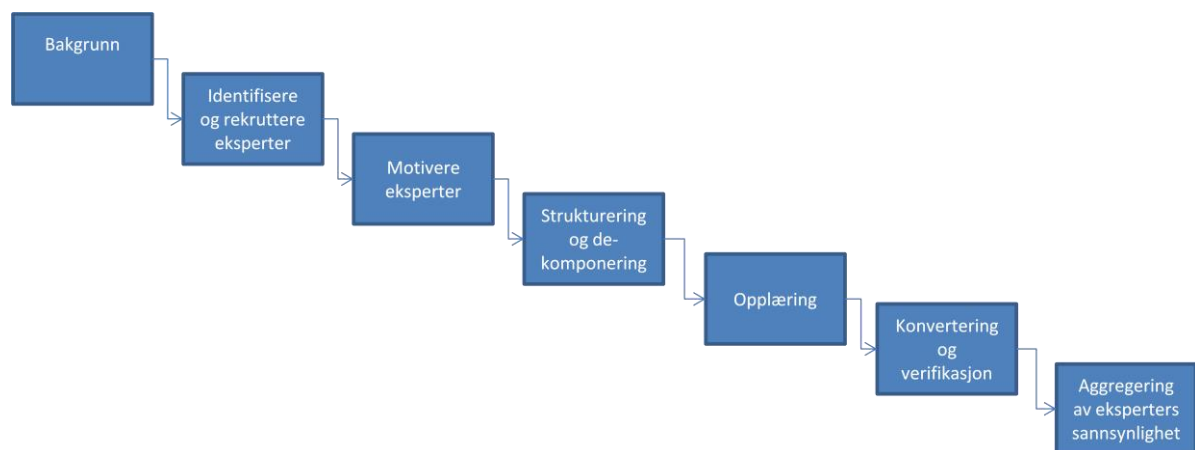
1. Beslutningstakeren, den som har behov for resultatet av vurderingen, og som ses på som kunden.
2. Fasilitator/ Policy maker, den som syr ekspertvurderingene sammen til et produkt som beslutningstakeren har nytte av.
3. Ekspert, den som i egenskap av å besitte kompetanse om et bestemt tema skal utspørres.

Det presiseres at det å velge en ekspert kan være kontroversielt, og at det ikke alltid er enkelt å velge en ekspert ut fra kjente/ synlige egenskaper, og selv om man er ekspert på et tema er det ikke sikkert at man vet så mye om sannsynligheter.

Clemen og Reilly viser også til Tversky og Kahnemans forskning, og beskriver de samme typer heuristikker og bias som tidligere er beskrevet.

Generelt kan prosessen ses på som enhver vitenskapelig datainnsamlingsprosess mener forfatterne, med de krav til reproduksjon og validitet som er allment kjent og som forventes. Clemen og Reilley har definert 7 aktiviteter i sin konverteringsprosess. Disse er:

- Bakgrunn
- Identifisere og rekruttere eksperter
- Motivere eksperter
- Strukturering og dekomponering
- Trening på konvertering
- Konvertering og verifikasjon
- Aggregering av ekspertenes sannsynligheter/ distribusjoner



**Figur 7: Clemen & Reillys konverteringsprosess**

#### Bakgrunn (identifisere variabler)

Første steg er å definere de variabler hvor man har behov for ekspertvurderinger i. Det påpekes at dette høres selvfølgelig ut, men at dette er en viktig oppgave. Videre påpekes det at det er hensiktsmessig å gå grundig til verks, og avdekke vitenskapelige undersøkelser som er foretatt på det aktuelle området, slik at man klarer å ta stilling til hvorvidt det er behov for ekspertvurderinger, og på hvilke områder det er behov for eksperter, og hvordan det er hensiktsmessig å legge opp konverteringen.

#### Identifisere og rekruttere eksperter

Clemen & Reilly påpeker at relevante eksperter kan være vanskelig å finne, men at de ofte finnes i relevante organisasjoner som kundens, eller faglige organisasjoner, og at andre fagpersoner er de mest relevante å spørre om hvem som kan være aktuelle.

#### Motivere eksperter

Clemen & Reilly har erfart at eksperter kan være kritiske overfor konverteringsprosesser, siden de ofte er vitenskapelig ansatte, og foretrekker å stole på mer tradisjonelle vitenskapsprosesser for å utlede svar på spørsmål. Når det da er behov for slike eksperter påpeker Clemen & Reilly at man ofte er nødt til å ta beslutninger, og dermed gjennomføre konverteringer selv med liten informasjon tilgjengelig, noe som mange misliker. Det blir da helt vesentlig å etablere et nært forhold med disse ekspertene, og skape entusiasme for det prosjektet man ønsker deres deltagelse i.

#### Strukturering og dekomponering

Kunnskapsundersøkelse kunne aktiviteten også hete i følge forfatterne. I denne fasen mener man at det vesentlige er å få eksperten til å lage en modell som på en hensiktsmessig måte viser



hvordan eksperten mener de ulike variablene henger sammen, og hva årsaksbildet til de ulike variablene er. Resultatet vil gi god innsikt i hvilke sannsynligheter som skal angis separat, og hvilke som bør utledes med årsak i andre forhold, og dermed kan en hensiktsmessig rekkefølge på variablene som skal tallfestes utledes.

#### Opplæring

Mange eksperter har ikke mye erfaring i verken bruk av konverteringsprosesser, eller de forholdene under prosessen som gjør at vi bruker vår forutinntatthet (heuristikker og bias). Det påpekes derfor at samtlige eksperter må få opplæring i de vanligste heuristikker og bias, og hvordan man kan håndtere slike under prosessen.

#### Konvertering og verifikasjon

Selve konverteringen, med tallfesting av sannsynligheter skjer under veiledning av en fasilitator i denne fasen. Det foreslås at ekspertens sannsynligheter verifiseres (at sannsynligheter summeres til 1, at betingede sannsynligheter er konsistente med grenseverdier og felles sannsynlighetsvurderinger etc.). Det anbefales videre at man forsøker å dokumentere ikke bare sannsynligheter og fordelinger, men også rasjonale bak vurderingene som gjennomføres, slik at dette kan gjennomgås ved behov.

#### Aggregering av ekspertenes sannsynligheter/ distribusjoner

Dersom flere eksperter har vært involvert i en vurdering kan det være ønskelig å aggregere sannsynlighetene/ fordelingene til en felles sannsynlighet/ fordeling, slik at det blir mer anvendelig for en beslutningstaker. Clemen & Reilly påpeker at dette kan gjøres på to måter:

- Enten kan man be ekspertene bli enige i fellesskap. Dette krever at ekspertene deler kunnskap og erfaring, og det kan være omfattende i tids- og ressursbruk. Det påpekes også at en slik prosess kan innebære at resultatet påvirkes på en måte som ikke er ønskelig, da man gjør seg sårbar for de nevnte heuristikker og bias.
- Dersom man ikke blir enige, eller dersom det av praktiske årsaker er vanskelig å la ekspertene møtes, kan en aggregering skje matematisk ved å lage gjennomsnitt av sannsynligheter eller fordelinger.

## **4.6 Vose**

David Vose er forfatter av boka "Quantitative Risk analysis", som nettopp gir en innføring i teknikker for bruk i kvantitative modeller og analyser. Kapittel 8 heter "Defining distributions from expert opinion", og omhandler ekspertvurderinger. Han argumenter for at det i risikoanalyser så å si alltid vil være et behov for å gjøre subjektive vurderinger, og at det ellers ikke er mulig å tilegne seg tilstrekkelig gode data som på en god måte ivaretar usikkerheten i de variablene som en risikoanalyse er avhengig av, av flere årsaker:

- Aktuelle data er ikke samlet inn tidligere
- Dataene er for kostbare å holde oppdatert
- Tidligere data er irrelevante (ny teknologi, endringer i politisk eller kommersielt miljø etc.)
- Tilgjengelige data er sparsomme, og ekspertvurderinger må gjøres for å tette "hull" i data.
- Objektet som analyseres og modelleres er nytt.

Vose sier at usikkerhet i subjektive estimater har årsak i to forhold:

- Usikkerhet i selve variabelen
- Usikkerhet som kommer fra ekspertens manglende kunnskap om variabelen

Vose definerer usikkerhet som differansen mellom estimatet og en tenkt "sann" sannsynlighet, noe som antyder en frekventistisk tenkemåte.

Vose skisserer to roller i sin modell:

- Analytiker – opptreter også som møteleder ved møter (chairman).
- Ekspert – den personen som har kunnskap som skal omgjøres til sannsynlighet.

Vose har som mange andre beskrevet Kahnemans og Tverskys momenter som nevnt i 2.5.4.

- Availability
- Representativeness
- Adjustment and anchoring

Vose har imidlertid også flere andre interessante feilkilder:

- Ekspert er feil ekspert – det hender at en ekspert ikke kan nok om det aktuelle området, og i stedet for å innrømme dette, og henvise til en annen, er det uheldig om vedkommende prøver å være hjelpsom, og gi sitt beste estimat uten å si fra. Dersom en analytiker opplever dette, bør vedkommende finne andre, og konferer om de resultater man har fått så langt.
- Organisasjonskultur – arbeidsmiljøet i noen organisasjoner er slik at man ønsker å være lojal gjennom å vise at man tror på bedriftens ideer. Eksempelvis<sup>2</sup> kan man tenke seg at det kan være vanskelig for ansatte i Vegvesenet dersom de skulle være med i en risikoanalyse for den nye tunnelen på E6 i Gudbrandsdalen som mange mener Vegvesenet bygger i strid med Tunnelsikkerhetsforskriften (ettløpstunnel med årsdøgnsstrafikk over 8000, uten nødutganger, slik Aftenposten beskrev i 2011). Når en løsning, som det er svært diskutabelt om ivaretar sikkerheten, blir valgt av Vegdirektoratet, kan det være vanskelig å være ansatt og si hva man mener om sannsynligheten for ulykker, siden man forventes å være lojal mot organisasjonens vedtak, selv om man er helt uenig i vedtaket.
- Skjult/ motstridende agenda – innebærer enkeltpersoner som har en økonomisk eller personlig interesse av et spesielt resultat av beslutning. Analytikeren må være observant på slike interessekonflikter sier Vose, og må hente inn supplerende ekspertvurderinger dersom han blir klar over at slike konflikter er gjeldende.
- Uvillighet til å vurdere ekstremverdier – eksperter kan av og til være uvillige til å vurdere ekstremverdier, men bør oppmuntres til dette for å få frem det reelle spennet i sannsynligheter sier Vose, uten egentlig å grunngi eller eksemplifisere problemet ytterligere.
- Ønsker å si "de rette tingene" – enkelte eksperter ønsker i stor grad å tilfredsstille en analytiker eller en beslutningstager ved å si det man tror analytikeren/ beslutningstageren vil høre, fremfor det han/ hun egentlig mener. Vose fremhever viktigheten av åpne spørsmål i en intervjusituasjon, og forhindre at man i eksemplet bruker tallverdier som kan brukes til ledetråder i svaret.
- Ekspert for travel – en analytiker som kommer til en virksomhet for å hente ut data fra virksomhetens eksperter på et spesielt tema er sjelden særlig velkomne. Ekspertene har det travelt, og ser ikke alltid nytten av egen deltagelse. Disse kan være bryske mot analytikeren, eller avfeie alle problemstillinger sier Vose. Symptomer på dette er at eksperten overforenkler estimatene, og setter X+/- y %, eller minimum, forventet og maksimumsestimater som er likt angitt for alle variabler. Løsningen i slike situasjoner,

---

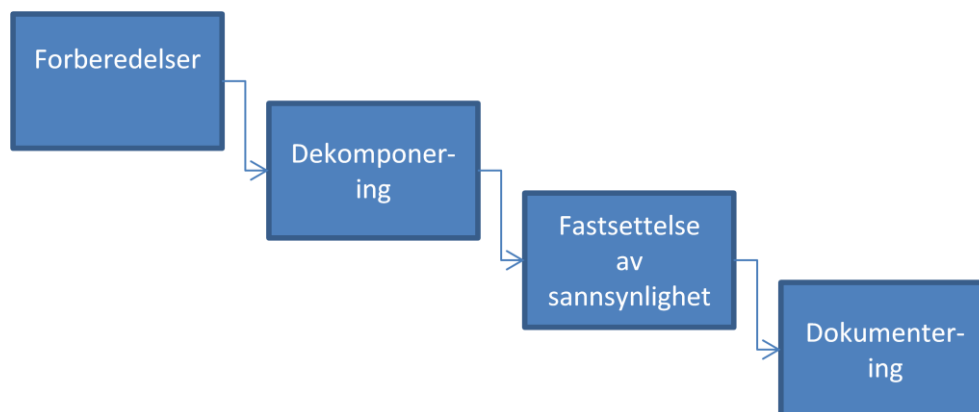
<sup>2</sup> Eksempel er ikke hentet fra boken, men fra denne forfatterens tolkning av Vose.

sier Vose, er å gå til ledelsen i virksomheten, og forankre prosessen der, slik at ekspertene får melding om å prioritere dette arbeidet.

- Enhetsforvirring – Vose oppgir forvirring i enheter som problem, og henviser til engelske enheter som er forandret i løpet av siste levetid, hvor noen bruker gallons, og andre pints.
- Ekspertene skal være helt sikre – enkelte kan tro at eksperter skal gi svar med liten usikkerhet rundt en middelvei, og tror at stor spredning er det samme som mangel på kompetanse, noe enkelte kan tro vil medføre et dårligere omdømme som ekspert.

Vose har ikke en strukturert beskrevet en prosess, men har gjennom innholdet beskrevet aktiviteter som skal gjennomføres. Han beskriver også hvordan man bør gjennomføre intervjuer og idédugnader. Prosessen har følgende aktiviteter:

- Forberedelser
- Dekomponering
- Fastsettelse av sannsynlighet
- Dokumentering



**Figur 8: Konverteringsprosessen til Vose**

#### Forberedelser

Alle deltagere bør i god tid få tilsendt informasjon som man skal sette seg inn i før idédugnaden/intervjuet.

#### Dekomponering

Vose påpeker viktigheten av at problemstillinger dekomponeres slik at de variablene ekspertene skal estimere er konkrete og håndterbare. Dette kan gjøre at det blir lettere å fastsette sannsynligheter. Dette gjør også at årsaksforhold kommer til syne, og hva som skal være inputnoder for sannsynligheter, og hva som på bakgrunn av disse blir mellomregninger, blir avklart.

#### Fastsettelse av sannsynlighet/ sannsynlighetsfordeling

Selve fastsettelsen av sannsynligheter skjer i en idédugnad og/eller i et intervju med en eller flere eksperter til stede. Ved bruk av flere eksperter anbefaler Vose en teknikk der han først gjennomfører en idédugnad i et fellesmøte, hvor analytikeren sørger for at arbeidet struktureres godt, og at det skrives referat fra møtet.

Etter møtet anbefaler Vose at det sendes referat til den enkelte deltager, og det gjennomføres deretter intervjuer med hver enkelt, en og en, for å få den enkeltes meninger om gruppens resultat. Denne fremgangsmåten har flere fordeler mener Vose:

- Hver ekspert får tid og mulighet til å tenke gjennom problemene før møtet og intervjuet
- Hver ekspert utfordres til å si sin mening etter først å ha hørt andres og gruppens mening
- Den fåmælt og forsiktige personen får like stor oppmerksomhet som den dominerende
- Meningsforskjeller mellom ekspertene blir lettere å identifisere
- Hele prosessen kan bli gjennomført ryddig og effektivt.

Vose mener denne teknikken er svært lik Delphiteknikken, som her beskrives som at ekspertene får tilgang til andres mening (anonymt). Vose beskriver også matematiske teknikker for aggregering etc. dersom deltagerne ikke blir enige.

#### Dokumentering

Det skrives referat fra møtet som møtedeltagerne får tilgang til før intervju. I tillegg dokumenteres arbeidet i notat/ rapport.

### **4.7 Encyclopedia of quantitative risk analysis and assessment**

Encyclopedia of quantitative risk analysis and assessment er et anerkjent firebinds oppslagsverk for risikoanalyse, og er tilgjengelig ved UiS. To av kapitlene heter henholdsvis "Expert elicitation for risk assessment" og "Expert Judgment". Her påstås det at det å dra kunnskap ut av en ekspert til syvende og sist dreier seg om å ha en strukturert spørsmålsstrategi, og at selve prosessen kan skje i "face to face"-situasjoner, i gruppearbeid, ved bruk av instrumenter(eksempelvis for avstemming etc.), simuleringer og rollespill.

En vellykket prosess forutsetter at man har en detaljert forståelse av det tema hvor eksperten skal utale seg om dette videre og aktuelle risikoer her, samt kjennskap til spesialistens behov og måte å kommunisere og uttrykke seg på.

Før man starter arbeidet med å samle eksperter for å diskutere et tema bør følgende forhold sikres:

- Er det planlagte opplegg for å få kunnskap fra en ekspert etisk riktig, og er man sikre på at velferden til eksperten er sikret under prosessen?
- Hvilke kulturelle eller sosiale forhold kan være til stede som fører til at eksperten kan misforstå, eller svare misvisende på spørsmål?
- Er det riktige eksperter som skal delta? Hvor godt undersøker spørsmålene de tema som eksperten besitter, og hvor godt egner de seg til faktisk å identifisere risikoen (sannsynligheten)?
- Er spørsmålene formulert og fremstilt på en måte som sikrer gode svar som er fri for uønsket påvirkning?

Forfatterne mener at strategien for spørsmålene må være mer betinget ut fra situasjonen, enn fra en bestemt sett med regler for hvordan prosessen skal gjennomføres.

Forfatterne legger vekt på at det ikke finnes en standard konverteringsprosess som er best i enhver situasjon, da det er hensiktsmessig med forskjellig tilnærming avhengig av hva analyseobjektet er, inndelt som følger (her følger noen situasjoner der ekspertuttalelser er relevant, men ikke nødvendigvis sannsynligheter):

- Hendelse
- System/prosess

- Konkurransen/ spill (med flere aktører med motsatt vilje)
- Samfunnsforhold

### Hendelser

De fleste konverteringsprosesser fokuserer på sannsynligheten for hendelser. Hensikten med en prosess for å konvertere kunnskap til sannsynlighet er å skaffe en best tilgjengelig verdi for den potensielle hendelsen basert på den kjennskap man har til hendelsen, i følge forfatterne. Prosessen fokuserer på å kvantifisere verdien.

Det er gjennomført undersøkelser av slike prosesser, og man vet forholdsvis mye om når eksperter bidrar godt og mindre godt til slike prosesser. Av slike forsøk har man etablert mye av den kunnskap man har knyttet til tommelfingerregler og feil (heuristics and bias).

Eksempler på bruk av ekspertvurderinger for å finne sannsynligheter på hendelser finnes i eksempler fra installasjoner i atomkraftverk, i sikkerhetsstudier, i vurderinger av organtransplantasjon, IT-systemer og søk- og redningsaksjoner, i eiendomsmarkedet, og i modeller for klimaendring med mer.

### System/ prosess

Ekspertvurderinger knyttet til risiko fra prosesser nødvendigvis gjør en litt annerledes tilnærming enn for hendelser. Risikoer fra hendelser forutsettes å være uavhengige, noe som tillater at dekomponerte sannsynligheter til å bli vurdert separat, for senere å bli satt sammen i et feiltre, et bayesiansk nettverk etc.

Risikoer knyttet til prosesser medfører imidlertid at det er korrelasjon knyttet til disse, og en konverteringsprosess må fokusere på å identifisere sammenhenger og dermed korrelerte underprosesser. Dette medfører komplikasjoner, da det kanskje ikke finnes eksperter som har kontroll på alle forhold i hele prosessen. Å sikre en prosess som utlokker alle relevante data vil således være krevende. Å verifisere en modell som har slike data vil være enda krevende, kanskje med motstridende ekspertuttalelser avhengig av ståsted i prosessen.

Prosesser knyttet til å utlokke sannsynligheter fra systemrisiko er mye mindre studert enn sannsynligheter fra hendelser. Det finnes mye mindre etablert kunnskap om hvilke feil og skjevheter og tommelfingerregler som vil bli benyttet i slike prosesser enn ved hendelser. Det å innhente kunnskap gjennom konverteringsprosesser for systemer er en ganske ny tilnærming (i forhold til hendelser), men er blitt mer aktuelt ved bruk av Markovkjede Monte Carlo-tilnærminger. Det anbefales å bruke mange av de samme prinsipper som for hendelser, og forutsette de samme feil/ skjevheter, gitt det nivå man per i dag er på med tilnærming av prosesser.

### Konkurransen/ spill

I denne kategorien kommer de forhold hvor to parter er i et konkurranseforhold hvor partene har motsatt rettet vilje. Eksempel kan være kriminalitet, terrorisme, krig etc. For å forstå hvordan en god utlokking av kunnskap kan skje i denne sammenheng, er det viktig å ha god kjennskap til spillteori (game theory).

Risiko i denne typen forhold er ikke basert på hendelser som enten inntreffer eller ikke, men blir aktivt etablert av en motpart med motsatt vilje. Forfatterne presiserer at risikoer av denne type ikke må ses på som en tilfeldig hendelse som enten inntreffer eller ikke, og derfor må en prosess knyttet til utlokking av sannsynligheter ikke ses på om sannsynligheter fra binære hendelser. Generelt vil vurderingen av slike risikoer avhenge av følgende:

- Hvor godt kjenner ekspertene strategien, den tilgjengelige informasjon og kost-nytteanalyser på hver side av spillet
- Egenskapen ved spillet på begge sider, er det en engangshendelse eller hendelser som skjer suksessivt?
- Evnen til å utlokke informasjon fra de ulike partene

For slike risikoer er det vanskelig å få tak i eksperter som vet tilstrekkelig om situasjonen på begge sider av forholdet. Kriminelle og terrorister etc. svarer neppe på anmodninger om å delta på konverteringsprosesser for å få innsikt i sannsynligheter dersom man skulle be om slike. Løsningen blir da ofte å etablere to lag, og etablere rollespill hvor partene går inn i hver sin rolle, og gjennomføre øvelser, pc-baserte spill, simuleringer etc. Resultatet av slike prosesser bli ra ofte en liste med de foretrukne strategier gitt strategien til det andre laget for begge sider i konkurransen.

#### Samfunnsmessige forhold

Utlokking av kunnskap i forbindelse med samfunnsmessige forhold har den forskjellen fra de andre typene av at en kvantifisering av risikoen (sannsynligheten) ikke er hensikten, men heller som et virkemiddel for å forhandle, og etablere kompromisser ved uenigheter som involverer eksempelvis fordeling av økonomiske midler, lokasjonsdiskusjoner etc. Ofte vil de ulike partene delta i en prosess med hvert sitt bilde av den tilstedeværende risiko, og hensikten er ikke å bli enige om hva risikonivået er, men å bruke risikonivået til å forhandle frem en løsning, for eksempel en annen plassering på en installasjon etc.

Suksesskriteriet for utlokking av ekspertkunnskap i slike situasjoner er å sikre at begge (alle) parter opplever at prosessen er god og ryddig. Kvantifisering av risiko (sannsynligheter) er neppe nødvendig utover å skape et konsensus.

Encyclopedia omtaler ikke de ulike rollene som inngår i prosessen. Det er kun fokus på eksperten, men administrator og intervjuer nevnes også kort.

Encyclopedia beskriver flere heuristikker, blant annet de kognitive som er nevnt i kapittel 2.5.4, og i tillegg legger forfatterne vekt på heuristikker ut fra motivasjon, og ut fra at spørsmål kan være stilt på en måte som styrer eksperten. Det legges i teksten stor vekt på å etablere en strategi som reduserer effekten mulige heuristikker og feilkilder.

Som tidligere beskrevet er det ikke en bestemt prosess som er best for alle typer analyseobjekter. Noen aktiviteter er imidlertid selvsagte, slik som utvelgelse av ekspert(er).

For å komme frem til en god prosess beskriver forfatterne at det en rekke vurderinger bør gjøres.

- I hvilken grad skal eksperter involveres og motta informasjon før selve ekspertvurderingen starter?
- Hva slags informasjon, og i hvilken mengde er eventuelt hensiktsmessig å videreformidle til eksperten før vurderingen?
- Hvor lang tid og hvor mye ressurser (budsjett, praktiske forhold) er tilgjengelig til forberedelser?
- Hva er hensiktsmessig møtested, skal eksperten arbeide på sitt vanlige sted (hjem/kontor) eller skal de innkalles til prosjektets kontor/ annet sted?
- Er det nødvendig med opplæring, i så fall hva slags, og hvordan skal den gjennomføres?
- Er det hensiktsmessig at navnene på ekspertene og deres vurdering og resultat blir offentliggjort?

Etter å ha vurdert de nevnte spørsmål designes ekspertvurderingen i en såkalt "protokoll". Encyclopedia viser til to eksempler for prosesser, en til bruk ved kun en ekspert, og en til bruk ved flere eksperter:

En ekspert (Morgan and Henrion, 1990):

- (Utvalgelse av ekspert)
- Motivering – bli kjent med ekspertene og forsøke å avdekke eventuelle motivasjoner hos ekspertene som kan påvirke resultatet
- Strukturering – strukturering av problemstillinger
- Kondisjonering – ekspertene gjennomgår helt grunnleggende sin evne til bedømmelse for å unngå kognitive feilkilder
- Kvantifisering - fastsettelse av sannsynligheter på det aktuelle tema
- Verifisering – kontroll av konsistens (ikke "riktighet", men matematisk, at sannsynligheter summerer til 1 osv)

I denne prosessen er analytikerens rolle å hjelpe ekspertene til å unngå psykologiske feilkilder

Flere eksperter (US Nuclear Regulatory Commission, NUREG 1150):

- (Utvalgelse av eksperter)
- Fellesmøte nr 1
  - Presentasjon av tema og bakgrunnsmateriale
  - Diskusjon mellom ekspertene vedrørende tema, og tilbakemelding på spørsmål
  - Opplæring, inkludert tilbakemelding på (ekspertens) vurderinger
- Individuelle studier (ca en måned)
- Fellesmøte nr 2
  - Diskusjon vedrørende metoder, modeller og datakilder som er benyttet
  - Individuell vurdering av hver ekspert (det er ikke tillatt med informasjonsutveksling mellom ekspertene i denne fasen)
- Etterarbeid/ dokumentasjon
  - Dokumentasjon av forklaringer
  - Feedback fra eksperter
  - Aggregering av resultater ved hjelp av matematiske teknikker (enkle gjennomsnittsberegninger).

#### 4.8 Terje Aven

Terje Aven er professor ved UiS, og en betydelig bidragsyter innenfor risikoanalyser i Norge generelt, og subjektive prosesser som benyttes i risikoanalysene spesielt. Han gjennomgår hvordan man kan fastsette sannsynligheter i kapittel 4 i sin bok "Foundations of risk analysis" (2003).

Han presiserer at det ikke finnes "sanne" sannsynligheter, og gir dermed klart uttrykk for en bayesiansk tilnærming, ikke en frekventistisk. Han påpeker at man ikke kan verifisere sannsynligheter i en risikoanalyse, da disse representerer risikoanalytikerens usikkerhet knyttet til hendelsen/ a priori sannsynlighet.

Han presiserer at ekspertvurderinger bør benyttes når det ikke er gode data tilgjengelig på en lettere måte. Han gjør oppmerksom på at slike prosesser kan være svært ressurskrevende.

Aven definerer ikke spesifikke roller. Av teksten fremgår det allikevel følgende roller i prosessen:

- Risikoanalytiker/ ekspert
- Evaluerer
- Analytiker

Aven har en svært interessant betraktning, at det til syvende og sist er analytikeren som er ansvarlig for de tall som benyttes i analysen. Av andre kan man få en forståelse (mulig ubevisst) at det er eksperten som er ansvarlig, men Aven påpeker at eksperten kun er et instrument hvorfra vi får kunnskap til å sette sannsynlighet. Han påpeker at eksperten ofte har avansert kunnskap på et svært snevert område, og at det sjelden er mulig for vedkommende å holde tritt med en god analytiker innenfor de aktuelle vurderings spørsmålene.

Aven har også nevnt Kahnemans og Tverskys tre heuristikker som nevnt i 2.5.4.:

- Availability
- Representativeness
- Adjustment and anchoring

Aven har ikke definert aktivitetene som foregår i en konverteringsprosess spesifikt steg for steg på samme måte som flere andre, men har mer fokus på hvordan man skal komme frem til gode og brukbare sannsynligheter. Av teksten fremgår det allikevel noen av aktivitetene i en konverteringsprosess, og slik det fremstår, som han mener prosessen bør inneholde. Disse er:

- Oppøving/ trening av eksperter
- Konvertering

Han er kritisk til kalibreringsøvelser dersom hensikten er at man skal få en ekspert til å kalibrere sin sannsynlighetsvurdering innenfor spesielle temaer, men han mener at disse øvelsene kan være gode for øve opp evnen til å gi input i form av sannsynligheter generelt, og at slik trening øker troverdigheten til eksperten. Han beskriver hva man bør legge vekt på ved trening av eksperter, slik at trening/ opplæring av eksperter regnes som en av aktivitetene som bør inngå i en prosess. Videre beskrives det at treningen foruten talltrening også bør inneholde informasjon om heuristikker og bias.

Han påpeker videre at eksperter enten kan oppgi sannsynligheter direkte, eller, dersom de ikke er vant til sannsynligheter, kan de prate i sitt eget språk, og la risikoanalytikeren "oversette" deres språk til sannsynligheter. Den sistnevnte fremgangsmåten har den svakheten at en risikoanalytiker neppe fullt og helt kjenner alle detaljer i eksperten fag og språkterminologi, og at der en fare for at han misforstår noe, noe som ikke nødvendigvis oppdages dersom ikke eksperten følger han tilsvarende innenfor sannsynlighetsregning. Aven mener at resultatet av en ekspertvurdering må tilfredsstillende noen gitte prinsipper (Cooke 1991) for at de skal ses på som akseptable og danne konsensus:

- Reproduerbarhet – alle beregninger må kunne gjennomføres på nytt, og man må få det samme resultatet.
- Forklarende (Accountability) – grunnlaget for sannsynlighetene som er brukt må kunne identifiseres og forklares
- Empirisk kontroll – sannsynlighetsvurderingene må kunne være gjenstand for/ mottakelig for kontroll
- Nøytralitet – metodene for å kombinere og evaluere ekspertmeningene må oppfordre ekspertene til å uttale seg med sin fulle sannhet/ overbevisning.
- Rettferdighet – alle eksperter skal behandles likt

Videre kommer Aven kort inn på hvordan motivasjonen kan variere hos de enkelte personer, og han nevner som eksempel at en selgers mening kan gis mindre tyngde enn en ingeniør, da man kan tro at selgeren har flere motiver som kan påvirke hans dømmekraft. Aven påpeker at ingeniører oftere ses på som mer objektive, men at ingeniører også kan ha insentiver som påvirker deres bedømming.



#### 4.9 Rausand og Utne

Rausand og Utne har i sin bok *"Risikoanalyse – teori og metoder"* en gjennomgang av stegene risikoanalysen. Kapittel 5 gir de en beskrivelse av de ulike trinnene i risikoanalysen. Boken har ikke et eget kapittel om bayesianske metoder, eller bruk av ekspertvurderinger, men har mer fokus på kilder hvor "objektive" data kan hentes fra, teknikker som kan brukes i en risikoanalyse (hendelsestrær, feiltrær etc.), og stegene i standarder for risikovurdering.

Rausand og Utne kommer ikke inn på roller i sin bok, men henviser til analysegruppen som må inneholde nødvendig kompetanse for å gjennomføre risikoanalysen på en riktig måte.

Rausand og Utne kommenterer ikke heuristikker og bias relatert til ekspertvurderinger. Kompetanse om disse er heller ikke nevnt som nødvendig i analysen. Under forklaring på hvordan man sikrer kvalitet på analysen nevnes det at analysen skal være objektiv, systematisk, strukturert og i størst mulig grad skal være basert på fakta. Vi aner en mer frekventistisk holdning på Rausands og Utnes tilnærming i denne boken.

Rausand og Utne kommer ikke inn på prosessen med hvordan kunnskap konverteres til sannsynligheter i detalj, men nevner følgende:

- Personell/ kompetanse er input til trinn 1 i risikoanalysen, som blant annet er å etablere analysegruppe og tilstrebe involvering.
- I hvilken grad en risikoanalyse oppfyller de målsettingene som ble satt for analysen, er i stor grad avhengig av den kompetansen analysegruppa har. Nødvendig kompetanseelementer kan omfatte flere nevnte fagområder.

#### 4.10 Oppsummering av funn

I dette kapittelet er åtte ulike publikasjoner (bøker og artikler) gjennomgått. Under presenteres de åtte i stikkord i egnet tabell, slik at leseren grovt kan sammenligne de ulike forfatters tilnærming.

**Tabell 2: Oversikt over funn i litteratur utover de som evalueres i kap 5.**

Navn	Aktiviteter	Roller	Heuristikker	Aggregering
O'Hagan et al	-Forberedelser -Identifisere og rekruttere eksperter -Motivere og lære opp eksperter -Strukturering og dekomponering -Konvertering med aggregering.	Beslutningstaker Ekspert Statistiker Fasilitator	Tversky & Kahneman	-Matematisk -Konsensus i gruppe. -Delphiteknikk
Meyer & Booker	-Definering av spørsmålsonråder og konkrete spørsmål -Strukturere spørsmål -Velge og motivere eksperter -Velge elementer for konverteringsprosess	Ekspert Analytiker Beslutningstaker (Prosjekt-personell som datainnsamler etc.)	Knyttet til de 4 aktiviteter en ekspert gjennomfører: -Forstå spørsmålet -Huske relevant informasjon -Gjøre seg opp en mening -Formulere og	-Matematisk -Konsensus i gruppe. -Delphiteknikk

Navn	Aktiviteter	Roller	Heuristikker	Aggregering
	<ul style="list-style-type: none"> <li>-Designe selve ekspertvurderingsprosessen</li> <li>-Øving prosjektpersonell</li> <li>-Gjennomføre ekspertvurderingsprosess</li> <li>-Data analyse</li> </ul>		<p>rapportere et svar</p> <p>Tversky &amp; Kahneman relevant i pkt 3</p>	
Merkhofer	<ul style="list-style-type: none"> <li>-Motivering</li> <li>-Strukturering</li> <li>-Foredling</li> <li>-Koding (utlokking)</li> <li>-Verifisering</li> <li>-Aggregering</li> <li>-Diskretisering</li> </ul>	<p>Ekspert</p> <p>Analytiker</p>	<p>Tversky &amp; Kahneman</p> <p>Motivasjonspåvirkning</p>	<ul style="list-style-type: none"> <li>-Konsensus i gruppe</li> <li>-Mekanisk (matematisk)</li> </ul>
Vose	<p>Forberedelser</p> <p>Dekomponering</p> <p>Fastsettelse av sannsynlighet</p> <p>Dokumentering</p>	<p>Analytiker</p> <p>Ekspert</p>	<p>Tversky &amp; Kahneman</p> <p>Feil ekspert</p> <p>Organisasjonskultur</p> <p>Skjult agenda</p> <p>Uvillighet til å si de rette tingene</p> <p>Ønsker å si de rette tingene</p> <p>Ekspert for travel</p> <p>Enhetsforvirring</p> <p>Viser ikke usikkerhet</p>	<ul style="list-style-type: none"> <li>-Matematisk aggregering</li> <li>-Delphiteknikk (men først i gruppe så intervju med hver enkelt, ikke likt som O'Hagan)</li> </ul>
Clemen & Reilley	<ul style="list-style-type: none"> <li>-Bakgrunn</li> <li>-Identifisere og rekruttere ekspert</li> <li>-Motivering</li> <li>-Strukturering o dekomponering</li> <li>-Trening på konvertering og konvertering</li> <li>-Aggregering</li> </ul>	<p>Beslutningstaker</p> <p>Fasilitator</p> <p>Ekspert</p>	<p>Tversky &amp; Kahneman</p>	<ul style="list-style-type: none"> <li>-Matematisk</li> <li>-Konsensus i gruppe</li> </ul>
Encyclopedia	<p>Ikke definert en enkelt prosess, flere mulige avhengig av analyseobjekt</p>	<p>Ekspert</p> <p>Administrator</p> <p>Intervjuer</p>	<p>Tversky &amp; Kahneman</p> <p>Motivasjonspåvirkning</p>	
Aven	<p>Ikke definert en enkelt prosess, resultatet må</p>	<p>Risikoanalytiker/ ekspert</p> <p>Evaluerer</p>	<p>Tversky &amp; Kahneman</p>	

Navn	Aktiviteter	Roller	Heuristikker	Aggregering
	tilfredsstill et sett kriterier, prosess designes etter det.	Analytiker		
Rausand og Utne	-	-	-	-

Som det fremkommer av gjennomgangen har Rausand og Utne ikke vektlagt konverteringsprosesser. Videre omtale av Rausand og Utne utelates derfor i denne oppgaven.

Det ble tidligere sagt at Rausand og Utne representerte NTNU-miljøet, hvilket er riktig, men for å påpeke at NTNU-miljøet også er involvert i arbeid knyttet til disse gjøres det oppmerksom på at Conceptrapport nr 17 fra NTNU omhandler "Ekspertvurderinger for å avdekke og strukturere kvalitativ informasjon" (Sunnevåg, 2007). Denne er ikke gjengitt detaljert her, men er benyttet som henvisning.

## 5. DISKUSJON

### 5.1 Sammenligning av metoder

For å få et godt grunnlag for å lage en ny "best practice" strukturmessig og innholdsmessig, gjennomgås de enkelte variablene i de ulike metodene i det følgende:

- Forutsetninger for prosess
- Innhold i prosessen
- Ekspertrollen
- Heuristikker og bias
- Aggregering

Risikoanalyser er svært variable i omfang. Når litteraturen i kapittel 4 gjennomgås, kan det virke som om man må engasjere flere titalls eksperter for at eksperter og konverteringsprosesser skal være relevant. Ekspertvurderinger benyttes imidlertid også i risikoanalyser som er små i omfang, og det ligger ikke som en forutsetning i det videre arbeidet at det nødvendigvis må være mange eksperter i et stort prosjekt.

#### 5.1.1 Forutsetninger for prosess

Utgangspunktet for oppgaven er risikoanalyser hvor det er behov for ekspertvurderinger for å få data på sannsynligheter. Som kapittel 4 viser, er det flere av de nevnte prosessene som omhandler ekspertvurderinger generelt (Meyer og Booker), og hvor også ekspertvurderinger av andre forhold enn sannsynligheter (eks konsekvenser) vil være dekket av innholdet. Aven og Vose omhandler risikoanalyser, hvor de også omtaler ekspertvurderinger som hjelpemiddel for å etablere sannsynligheter, mens O'Hagan og Merkhofer fokuserer kun på sannsynligheter.

Mens de fleste nevnte forfattere går i dybden på prosess og bias direkte, har Encyclopedia og Cooke en annen fremgangsmetode som skiller seg fra de andre. Før man starter en konverteringsprosess med eksperter må man vurdere hva slags objekt man ser på, og de skiller mellom hendelser, system/prosesser, konkurranse/ spill (med flere aktører med motsatt vilje) og det vi kalte samfunnsforhold.

Meyer & Bookers prosess tar utgangspunkt i at konverteringsprosessen skjer som del av et prosjekt, og de starter med å identifisere prosjektmål og hensikten med prosjektet. Det er et fellestrekk med Meyer og Booker og Encyclopedia ved at de starter med en overordnet tilnærming, og forsøker å forstå hvordan konverteringsprosessen best kan støtte et større mål, og de erkjenner at avhengig av prosjektmålet, kan det være ulike måter hvor ekspertvurderinger og konverteringsprosesser bør gjennomføres, og det kan være hensiktsmessig å tilpasse prosessen etter hva man ønsker å oppnå. Resterende litteratur fokuserer mer på detaljer i konverteringsprosessen.

Rent pedagogisk ser Meyers og Bookers litteratur bra ut, spesielt fra en prosjektleders ståsted. Mye av den resterende litteraturen er imidlertid skrevet for analytikere/ statistikere som kanskje ikke er i en prosjektlederposisjon, men som kommer inn for å gjøre en risikoanalyse som del av et stort prosjekt. Det er da viktig å legge merke til at flere av de trinn som Meyer og Booker beskriver i sin litteratur er noe mer overordnet, og kanskje må ses på som forutsetninger, da planer allerede er vedtatt, og flere andre aktiviteter er gjennomført. Desto viktigere å forstå hvordan risikoanalysen passer inn, og kan skape verdi for prosjektet. Meyer og Booker kommer derimot ikke inn på, som Encyclopedia, at ulike analyseobjekter krever ulike fremgangsmåter. Det blir uriktig å si at de ikke mener det, men årsaken til at det ikke sies så direkte er nok at

Meyers og Bookers bok omhandler ekspertvurderinger generelt, ikke ekspertvurderinger i risikoanalyser spesielt.

Encyclopedias gjennomgang av ulike typer analyseobjekt viser at det er hensiktsmessig å være bevisst hvilket analyseobjekt man vurderer, da dette kan ha påvirkning på hva som er essensen i svaret eksperten skal gi. Fra den reelle verden har vi tilsvarende eksempler som kan være nyttig:

Når vi gjennomfører pålitelighetsanalyser kan det være for å kontrollere at vi tilfredsstillir riktig SIL (Safety Integrity Level)-nivå i forhold til et regelverk/ en standard. Ofte har vi da feildata fra OREDA etc. som gir oss relevant input. Stadig økende bruk av software, som det ikke i samme omfang finnes feildata på, i kombinasjon med "kjente" tekniske komponenter gjør imidlertid at ekspertvurderinger kan være hensiktsmessig i slike analyser. Vi vet at standarder setter krav til SIL-nivå (1-4), og at et system typisk må tilfredsstillir nivå 3 for å være godkjent. Grenseverdier for nivå 3 er oppgitt, og målet med en ekspertvurdering vil være å konvertere all tilgjengelig kunnskap til et tall vi tror på, som kan brukes i pålitelighetsvurderingen videre, for å vise at systemet samlet oppnår nivå 3. Her er altså selve tallstørrelsen kritisk, og den som designer prosessen med eksperter må ha det i mente.

Direktoratet for samfunnssikkerhet og Beredskap (DSB) har gitt ut en veiledning for bruk av ROS-analyser i arealplanlegging. Med endring i siste utgave av plan- og bygningsloven er det gitt pålegg om at utbygger skal gjennomføre en ROS-analyse i en reguleringsplanfase: *"Analysen skal vise alle risiko- og sårbarhetsforhold som har betydning for om arealet er egnet til utbyggingsformål, og eventuelle endringer i slike forhold som følge av planlagt utbygging"*. I ROS-analyser er det aktuelt med flere typer eksperter, det kan være innenfor flom og vannføring, ras og skred, biologisk mangfold og naturpåvirkninger etc. Av teksten fra plan- og bygningsloven leser vi at man primært pålegges å vurdere *endringen* i risiko, det står ikke at risikoen skal kvantifiseres. En prosess i en slik sammenheng skal således ha fokus på endring av sannsynlighet (og konsekvens), ikke nødvendigvis å fastslå sannsynligheten (risikoen) nøyaktig (det finnes ikke akseptkriterier som kan brukes for å evaluere et risikonivå, slik som standarden setter krav i SIL-analyser).

Encyclopedia viste til flere typer objekter, blant annet konkurransesituasjoner med aktører med motsatt vilje, og viste at disse bør håndteres forskjellig fra hendelser. Det virker således fornuftig slik Encyclopedia skisserer, at man tar hensyn til analyseobjektet før man planlegger prosessen videre i detalj.

Av og til finnes det få eksperter som kan bidra med god og relevant kunnskap. Av og til har disse ekspertene en interesse i saken også, som eier, ansatt eller på annen måte en berørt part. Det er ønskelig å avdekke slike interesser raskt, da slike interesser nettopp kan svekke tilliten til vurderingene som gjøres. I valget mellom å bruke en ikke-ekspert, eller en ekspert som har interesse i saken, foretrekkes det siste, men det anbefales da å dokumentere prosessen for å imøtegå senere spekulasjoner som måtte komme i etterkant (Meyer & Booker, 1991).

### 5.1.2 Innhold i prosessen

I kapittel 4 så vi at de ulike forfatterne har ulikt antall trinn i sin konverteringsprosess, og at innholdet i noen grad varierer. Imidlertid er det noen fellestrekk ved inndelingen som jeg vil forsøke å illustrere i det følgende.

O'Hagan, Clemen & Reilly og Vose starter alle med en oppstarts-/ forberedelsesfase, og Meyer & Booker og Encyclopedia viser til aktiviteter som det er hensiktsmessig å kalle en oppstarts-/ forberedelsesfase. Her inngår blant annet det å forstå hensikten med prosjektet, og få oversikt over hvilke områder man trenger ekspertkunnskap på. Merkhofer starter med motivering, og han

nevner ikke en initieringsfase eksplisitt, men hans artikkel starter noe lenger inn i konverteringsprosessen enn de andre. Som det fremgår av pkt 5.1.1 er det hensiktsmessig å tilpasse prosessen til analyseobjektet. Vi konkluderer med at det bør være en oppstartsfase i begynnelsen av prosessen.

Som neste hovedbolk av aktiviteter varierer det litt hva de ulike forfatterne beskriver:

- O'Hagan identifiserer, rekrutterer og motiverer eksperter før han strukturerer
- Meyer og Booker strukturerer spørsmål, velger og motiverer eksperter før konverteringsprosessen designes og personell øves.
- Merkhofer motiverer og strukturerer
- Vose dekomponerer problemstillingen
- Clemen & Reilly velger og motiverer eksperter før strukturering og dekomponering og øving gjennomføres

Aven har ikke satt opp en egen prosess, men sier at den tilpasses behovet og designes ut fra behov. Han sier imidlertid også at det er analytikeren som er ansvarlig, ikke eksperten, eksperten skal kun gi tallgrunnlaget. O'Hagan skisserer at eksperter kan være med på å dekomponere problemstillinger. Hvem skal så dekomponere problemstillingen, og lage spørsmålene? Man kan dels argumentere for at dette til en viss grad er en iterativ prosess, og dels for at det må tilpasses hvert enkelt tilfelle.

I henhold til Aven er det analytikeren som produserer selve analysen, og som dekomponerer problemstillinger og strukturerer arbeidet. Noen må uansett først identifisere et behov for en ekspertvurdering, og det er analytikeren. Dette stemmer også med Meyer og Booker som først identifiserer spørsmålsmråder, det vil si å avklare innenfor hvilke områder man trenger eksperter i konverteringsprosesser.

Mens O'Hagan tilsynelatende først rekrutterer og motiverer eksperter, og så bruker eksperter til å strukturere og dekomponere, mener Meyer & Booker at det er hensiktsmessig først å strukturere spørsmål før man velger og motiverer eksperter. Dette kan se ut som en uenighet, men slik denne forfatteren forstår de to, kan begge ha rett hvis man ser det miljøet de to skisserer prosessen i.

Meyer og Booker skisserer et stort prosjekt med en stor stab analytikere og prosjektmedarbeidere, samt flere dusin eksperter. O'Hagan gir et bilde av et mindre prosjekt, hvor man er dypere i en enkelt problemstilling. I det prosjektet Meyer & Booker skisserer er det store analytiker-ressurser som det forutsettes selv også har en god kjennskap til fagområdet. Prosjektet er stort, og det er mange ekspertvurderinger. Analytikerne dekomponerer problemstillinger og strukturerer spørsmål. Når man kjenner disse spørsmålene kan man identifisere relevante eksperter som så kan kalles inn, motiveres, gis opplæring, og delta i prosjektet. O'Hagan har identifisert de aktuelle områdene hvor det er behov for eksperter allerede i oppstartsfasen, og kaller nå inn disse ekspertene. Han har ikke tilsynelatende så stor stab som Meyer og Booker har, men kanskje et mindre prosjekt. Han mener at ekspertenes kunnskap er nødvendig for å dekomponere problemstillinger, og finne årsakssammenhenger som man så setter sammen til sannsynligheter.

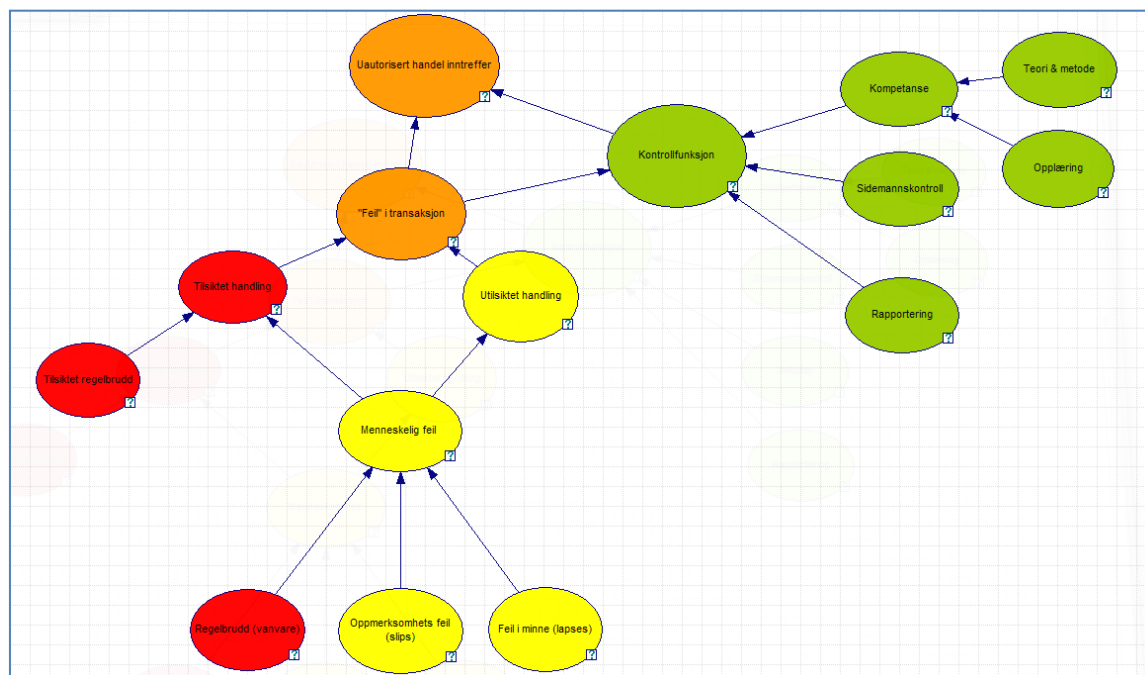
Det er altså ingen som har rett eller galt, men de beskriver noe ulike forutsetninger. Begge er enig om at man først må identifisere hvilke områder man trenger eksperter på, begge er enig i at dekomponering og strukturering er hensiktsmessig for å få håndterlige spørsmål, og begge er enige i at eksperter må identifiseres, rekrutteres og motiveres. Vose og Clemen & Reilly har de samme elementene, men ikke så fint oppdelt som Meyer og Booker.

Situasjonen må således avgjøre rekkefølgen her, dersom analytikerne behersker fagområdene så godt at de selv kan dekomponere og strukturere, så gjøres det før ekspertene involveres, men dersom det er behov for ekspertenes bistand til dekomponering må det ses i sammenheng med konverteringsprosessen. I praksis, for den store mengden av noe mindre analyser, er det sannsynlig at ekspertene involveres i mer enn kun å fastsette sannsynligheter på et papir, men de deltar i idedugnader hvor hendelser identifiseres, årsaker kartlegges, og sannsynligheter fastsettes. Hvorvidt man kaller en årsaksanalyse for dekomponering kan kanskje diskuteres, men dersom vi definerer dekomponering som et høyere nivå, er det naturlig å se på ekspertenes bistand i årsaksanalyser kun som en del av selve konverteringsprosessen.

Av dette følger at en aktivitet med dekomponering og strukturering gjennomført av analytikere gjennomføres før eksperter identifiseres og rekrutteres, og deretter bringes inn for konverteringsprosessen.

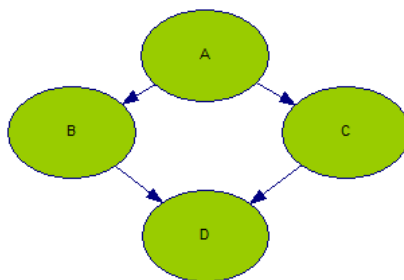
Selve konverteringsprosessen er selvfølgelig med hos alle forfattere. Den benevnes noe ulikt av de ulike forfattere, men inneholder i hovedtrekk de samme elementer – en prosess der eksperter er til stede for å bidra med sin kunnskap og konvertere denne til sannsynligheter.

O'Hagan og Merkhofer beskriver at konverteringen kan være med hjelp av indirekte eller direkte metoder. Encyclopedia beskriver kort at Bayesianske nettverk kan benyttes i denne prosessen. Aven beskriver hvordan dette kan gjøres i praksis. Sannsynligheter kan finnes direkte dersom man har relevante data. Imidlertid er årsaksbildet og den nødvendige bakgrunnskunnskapen ofte så komplekst at det er vanskelig uten å tegne dette i en sannsynlighetsmodell.



Figur 9: Eksempel på enkelt Bayesiansk nettverk som viser årsaker til en uautorisert handel i en bank

Ved et enkelt eksempel kan utregningene vises. La A være forelder til B og C som igjen er foreldre til A, se Figur 10: Eksempel



Figur 10: Eksempel

Sannsynligheten for D er da:

$$P(D)=P(D,C,B)*P(B,C)+P(A|\bar{C},B)*P(\bar{C},B)+ P(D|C,\bar{B})*P(C,\bar{B})+P(D|\bar{C},\bar{B})*P(\bar{C},\bar{B})$$

Det er sammenfallende synspunkt blant forfatterne at en vesentlig del av selve konverteringsprosessen er opplæring i forkant av konverteringen. Innholdet i opplæringen varierer noe, men alle har fokus på at ekspertene må forstå hvilke heuristikker og andre påvikninger de kan være eksponert for under prosessen, slik at dette kan redusere sannsynligheten for skjevfordelinger.

Det kan synes som forfatterne ikke har et likt syn på om det er riktig eller ikke å trene på konvertering, for deretter å kontrollere om ekspertene har svart "riktig". Her kommer vi inn på kjernen i forskjellen på frekventister og subjektivister. Aven argumenterer med at sannsynlighet er en måleenhet på usikkerhet, det eksisterer ingen sann verdi, kun vår grad av tro på hvorvidt en hendelse inntreffer eller ikke. Aven mener imidlertid at trening på konvertering har en viss verdi, men da for å gjøre eksperten mer sikker på å håndtere sannsynligheter, ikke i noen form for å kalibrere eksperten til en såkalt "riktig" sannsynlighet. Se også evaluering

Videre beskriver flere at prosessen kan gjennomføres med eksperter en og en eller med ekspertene samlet. Dette ses i sammenheng med aggregering av sannsynligheter og diskuteres ytterligere i 5.1.55.1.5.

### 5.1.3 Roller

Den litteraturen som er gjennomgått viser at det er forholdsvis stor likhet i bruk av roller. Følgende roller er identifisert:

Rolle	Litteratur
Ekspert(er)	O'Hagan et al, Meyer & Booker, Merkhofer, Vose, Clemen & Reilly, Encyclopedia, Aven
Analytiker (e)	O'Hagan et al, Meyer & Booker, Merkhofer, Vose, Aven
Statistiker (e)	O'Hagan et al
Administrator/fasilitator / intervjuer/ evaluerer/ prosjektpersonell	O'Hagan et al, Meyer & Booker, Clemen & Reilly, Encyclopedia, Aven
Beslutningstager	Meyer & Booker, Clemen & Reilly

Som vi ser er det konsensus om rollen ekspert. Selv om kun to beskriver beslutningstager som en rolle, er det opplagt at det er en beslutningstager eller oppdragsgiver, som skal ha resultatet av analysen, selv om han ikke deltar i selve prosessen. Nesten alle har en rolle som er ansvarlig



for gjennomføringen av konverteringsprosessen. Denne rollen kalles ved flere navn, slik som fasilitator, intervjuer etc. Flere har også rollen analytiker.

Som flere har påpekt kan noen av rollene dekkes av en og samme person. Analytikeren kan også være statistiker, intervjuer og fasilitator, samt planlegge og administrere prosessen.

For en "best practice" prosess forutsettes det derfor tre roller:

- Beslutningstager – den som skal bruke produktet, oppdragsgiver.
- Ekspert – den som har kunnskap som skal konverteres til sannsynlighet. Det kan være flere eksperter ved behov.
- Analytiker – den som planlegger og gjennomfører konverteringsprosessen. Kan være flere personer ved behov.

#### 5.1.4 Heuristikker og Bias

Den gjennomgåtte litteraturen er enstemmig i sin vurdering av at eksperter som gjennomfører konverteringsprosesser utsettes for heuristikker. Den fremtredende litteraturen innen området er laget av Tversky, Slovic og Kahneman, men både Vose og Encyclopedia beskriver praktiske forhold utover dette.

Encyclopedia, Meyer & Booker og delvis Aven beskriver at man i konverteringsprosesser bør etablere en strategi mot heuristikker. En strategi mot huristikker kan inneholde følgende:

- Forsøk å forutse hvilke bias som er aktuelle, og som det er sannsynlig at kan inntreffe i den kommende konverteringsprosessen
- Ved behov – gjør om den planlagte prosessen slik at den blir mindre eksponert for bias
- Informer ekspertene om heuristikker og bias før konverteringen starter, hva årsakene er og hvordan de blir definert
- Overvåk konverteringsprosessen med hensyn til om bias inntreffer
- Juster prosessen der og da for å imøtegå bias, eksempelvis, ta pauser dersom eksperter er utmattede.
- Analyser data for å finne eventuelle tilfeller av bias (dette trinnet kan erstatte 4 og 5 dersom en enklere analyse gjennomføres, men trinn 6 bør aldri utelates dersom man legger kognitive definisjon av bias til grunn, mener forfatterne).

Ut fra den gjennomgåtte litteraturen forstår vi at håndtering av heuristikker og bias er en svært viktig aktivitet i en konverteringsprosess, og at en "best practice" prosess må inkludere flere underaktiviteter fra start til slutt i konverteringsprosessen:

- Design for å unngå heuristikker
- Opplæring/ informasjon
- Overvåking
- Kontroll av resultater

#### 5.1.5 Aggregering

Når flere eksperter benyttes i samme konverteringsprosess, er det ønskelig å aggregere svare til en verdi som går inn som variabel i risikoanalysen. O'Hagan, Meyer & Booker, Merkhofer, Clemen & Reilly og Vose beskriver at dette kan skje i gruppe, hvor man skal nå konsensus, eller mekanisk (matematisk).

##### Gruppeteknikk

Delphi-teknikken ble etablert tidlig på 50-tallet som en spin-off fra at Air Force-prosjekt i USA, og er en av de kjente metoder for å få tak i eksperters meninger, og aggregere (syntetisere) denne.

Metoden går ut på at flere eksperter, som ikke kjenner hverandres identitet, fyller ut et spørreskjema. I enkelte tilfeller bes ekspertene om å rangere hvor god de er innenfor det aktuelle fagfeltet spørsmålet dreier seg om. Dette samles inn, analyseres, og sendes så ut til deltagerne sammen med medianen og 25 og 75 % persentilen. Ekspertene blir så spurt om de vil revurdere svarene sine, og dersom de beholder et svar utenfor +/- 25 % utenfor medianen, må de begrunne hvorfor. Dette gjennomføres som en iterativ prosess, typisk med 3-4 iterasjoner.

Flere er kritiske til Delphiteknikken av følgende årsaker:

- De som treffer innenfor +/-25 % blir "belønnet" med mindre arbeid, mens de som er utenfor må argumentere for sitt syn.
- Stor mengde svarark blir aldri levert, uten at svarprosent rapporteres i svar
- Det er ikke gjort forskning på hvorfor folk dropper ut, og ikke leverer svarark.
- Flere iterasjoner, gjør at man spør seg om svarene konvergerer mer av kjedsomhet (for å bli ferdig) enn på grunn av konsensus.

Delphi teknikken som beskrevet over anbefales av Meyer og Booker. O'Hagan og Encyclopedia beskriver imidlertid en litt annerledes gjennomføring av teknikken, der ekspertene i tillegg til svaret også legger ved en begrunnelse for svaret de gir, og denne informasjonen anonymiseres og gis de andre ekspertene.

Det synes å være en stor forskjell på disse to fremgangsmåtene. I den førstnevnte måten å gjennomføre Delphi-teknikken på, hva er årsaken til at ekspertene skal revurdere og gi sitt svar? Den eneste nye opplysningen er hva andre tror, og hvordan egen tro er i forhold til det.

På måten O'Hagan beskriver Delphiteknikken gis det imidlertid en begrunnelse for svaret. Det kan jo hende at det er forhold en ekspert ikke er kjent med, eller ikke har tenkt på, men som nevnes som forklaring til en annens ekspertvurdering. Det utvider kunnskapsgrunnlaget til eksperten, som så ut fra den nyervervede kunnskapen gir en korrigerende på første estimat.

Alle forfattere har beskrevet Tversky og Kahnemans arbeid vedrørende Heuristikker, og viser til den som gjeldende i sin metode. "Anchoring and adjusting" er en av de nevnte heuristikkerne som man ønsker unngått. Slik Meyer og Booker beskriver sin teknikk, blir det etter denne forfatterens syn en organisert heuristisk øvelse hvor den eneste kilden til påvirkning/ endring av eget tall er vissheten om hva de andre har svart. For at Delphi-teknikken skal ha verdi, i stedet for å virke som en heuristikk, er det derfor viktig at ekspertene begrunner sine valg godt, slik at andre eksperter med bakgrunn i eventuell ny innsikt kan vurdere å endre sin sannsynlighet.

Aggregering i gruppe kan være utsatt for gruppedynamikker som påvirker resultatet for mye. Merkhofer har en alternativ metode hvor man i ettertid diskuterer det gruppen kom frem til i intervjuer med en og en ekspert. På den måten får man frem hva den enkelte ekspert egentlig mener, noe som gjør at man får et bilde på om gruppedynamikken har påvirket resultatet vesentlig.

Av og til blir ikke gruppen enig, og analytikeren får et uløst problem i fanget. Da kan en mekanisk vektning bli løsningen. Hvis det er noen forskjell på ekspertene som tilsier at man skal høre mer på den ene enn den andre, er det mulig å lage en vektingsmodell, der man vekter de ekspertene man har mest tro på mer enn de resterende, og så bruke en mekanisk vektning.

Av og til er eksperter også helt uenig. Dersom man ikke kjenner fagområdet, og ikke kan skille ekspertene, bør man forhøre seg med flere eksperter for å få noe mer sikkerhet for tallene.

Mekanisk aggregering

Matematisk aggregering kan i enkleste form skje ved å regne ut det aritmetiske gjennomsnitt  $\bar{x}$ .

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n X_i$$

Meyer & Booker beskriver at også median<sup>3</sup> og geometrisk<sup>4</sup> gjennomsnitt er hyppig brukt som estimator for sentraltendensen. Det geometriske gjennomsnittet kan skrives:

$$\left( \prod_{i=1}^n a_i \right)^{1/n} = \sqrt[n]{a_1 a_2 \dots a_n}$$

Det er alltid likt eller mindre enn det aritmetiske gjennomsnittet. La oss se på et tenkt eksempel for å vise nytten. Fem eksperter har gitt hver sin verdi på en sannsynlighet. Som vi ser varierer svarene ekstremt mye. Kun ett av tallene er høyere enn det aritmetiske gjennomsnittet, og fordi dette ene tallet er så høyt, dras gjennomsnittet også opp. Det aritmetiske gjennomsnittet gir her ikke et godt bilde av meningene til de 5 ekspertene. I slike tilfeller kan medianen være en bedre estimator enn det aritmetiske gjennomsnittet. Vi ser også at det geometriske gjennomsnitt ligger nær medianen.

**Tabell 3: eksempel på bruk av ulike mål for sentraltendensen**

<u>Ekspert</u>	<u>Estimat</u>
1	0,00001
2	0,0001
3	0,0001
4	0,1
5	0,00001
<hr/>	
Aritmetisk gjennomsnitt	0,02004
Median	0,0001
Geometrisk gjennomsnitt	1,6E-04

Kahneman og Tversky (1982) har funnet en annen begrunnelse for å bruke medianen i stedet for aritmetisk gjennomsnitt i ekspertvurderinger. De fant at eksperter, når de konverterte kunnskap til tall, benyttet en verdi som egentlig var medianen, ikke det aritmetiske gjennomsnittet. Hvis alle verdiene fra en ekspertvurdering til sammen utgjør en symmetrisk form, så er medianen og det aritmetiske gjennomsnittet det samme. Tversky og Kahneman fant imidlertid at formen alltid var skjev, med en topp på den ene siden, og en lang hale på den andre, slik at medianen og det aritmetiske gjennomsnittet ikke var samme verdi. Av den bakgrunn anbefales det å se på ekspertens svar som medianverdi fremfor gjennomsnittsverdi (Meyer og Booker).

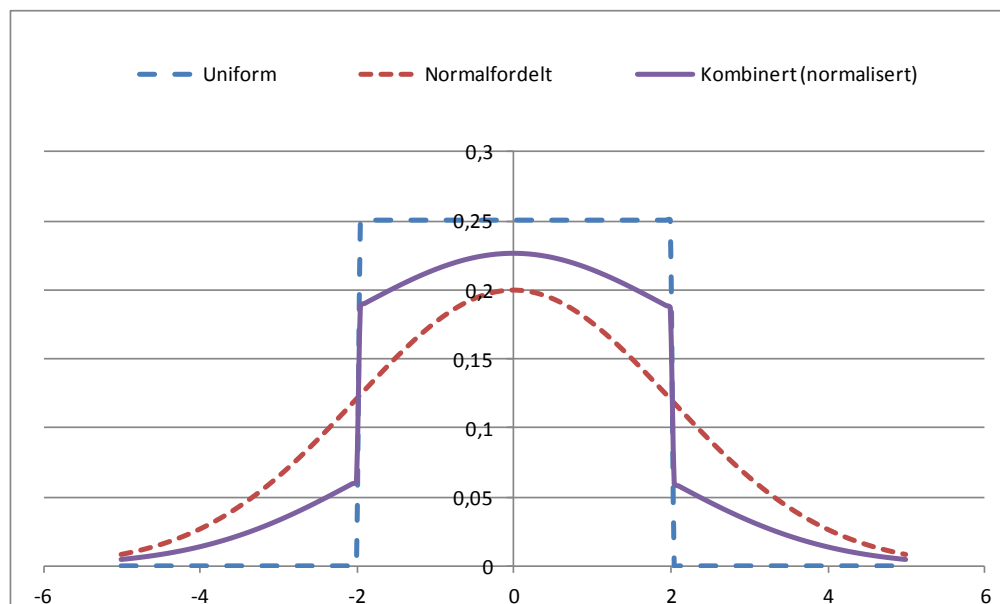
For fordelinger er det også svært enkelt å aggregere disse og få en ny aggregert fordeling. La oss tenke oss at to eksperter fant hver sin fordeling, ekspert A foreslår en uniform fordeling, og ekspert B foreslår en normalfordeling. Vi legger disse sammen punktvis, og normaliserer ved å dele på det samlede arealet under fordelingene. Resultatet blir en normalisert kombinert fordeling.

<sup>3</sup> Medianen er det midterste tallet i utvalget. Dersom utvalget er et partall, er medianen snittet av de to midterste.

<sup>4</sup> Det geometriske gjennomsnittet er n'te roten til produktet man finner ved å gange alle tallene i tallrekken.

Areal	1	0,988014312	1,988014	1
	Uniform	Normalfordelt	Kombinert	Kombinert (normalisert)
x	U(x)	N(x)	N(x)+U(x)	
-5	0	0,00876415	0,008764	0,00440849
-4,95	0	0,009326474	0,009326	0,00469135
-4,9	0	0,009918677	0,009919	0,00498924
-4,85	0	0,010541892	0,010542	0,00530272
-4,8	0	0,011197265	0,011197	0,00563239
-4,75	0	0,01188595	0,011886	0,00597881
-4,7	0	0,01260911	0,012609	0,00634256
-4,65	0	0,01336791	0,013368	0,00672425
-4,6	0	0,014163519	0,014164	0,00712446

Figur 11: utsnitt fra excel – grunnlag for figur 10.



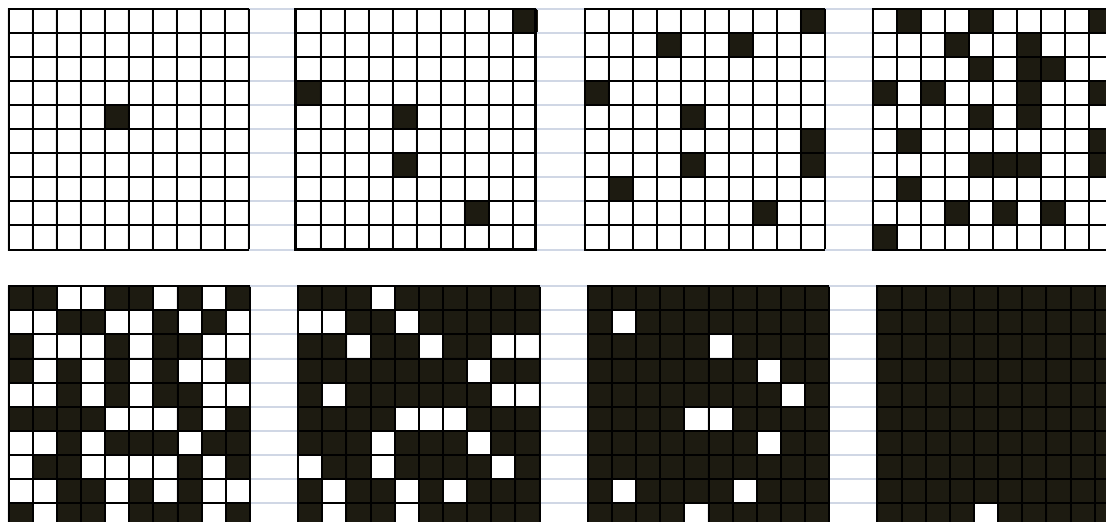
Figur 12: eksempel på aggregering av fordelinger

### 5.1.6 Om svært små sannsynligheter

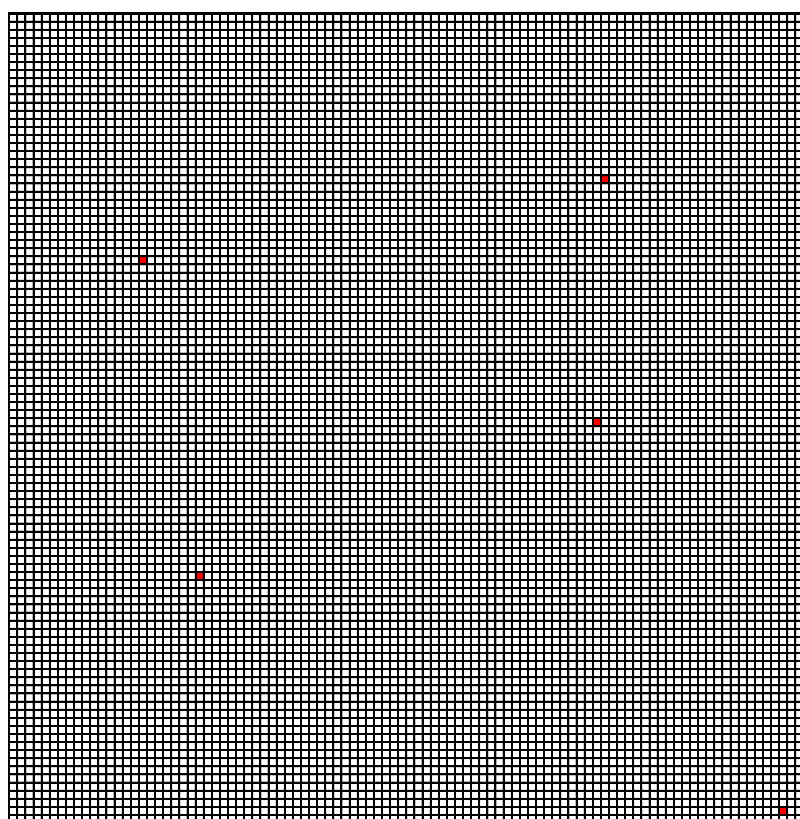
Å gjennomføre konverteringsprosesser med direkte metoder for svært lave sannsynligheter er en krevende oppgave. I risikoanalyser er det vanlig at vi opererer med hendelser i spekteret  $10^{-4}$ ,  $10^{-5}$  og  $10^{-6}$ . I stedet for direkte metoder kan sannsynlighetsmodeller benyttes, der man dekomponerer problemstillinger til ekspertene føler seg komfortabel med nivået (Clemen & Reilly, 2001). Dekomponering anbefales som en av flere hovedaktiviteter av Clemen & Reilly, O'Hagan og Vose, men også Meyer & Booker og Merkhofer benytter dekomponering i sine prosesser. Dekomponering bør være en del av analytikerens oppgaver i en designfase eller som del av selve ekspertvurderingen hvor ekspertene dekomponerer en hendelse i en årsaksanalyse.

Flere kollegaer som arbeider med risikoanalyser innrømmer at det er vanskelig å forholde seg til små tall, man mister følelsen med tallene i størrelsesorden under  $10^{-2}$  til  $10^{-3}$ . Frekvenser i størrelsesorden hvert 100. år, og sannsynligheter med hundredeler kan vi assosiere med årene i et liv eller i et århundre. Vi har kontroll på 1900-tallet i grovt med tanke på historien og hendelser, hvilket medfører at tall i størrelsesorden  $x \cdot 10^{-2}$  til har mange et forhold til.

Vose har i sin bok en god indirekte metode hvor han viser figurer som visuelt hjelpemiddel for å anslå sannsynligheter. For sannsynligheter mellom 1 % og 100 % kan den være et godt hjelpemiddel, men for små verdier blir den noe upraktisk i bruk.



Figur 13: Eksempler på figurer for å illustrere verdiene (oppe fra venstre) 1 %, 5 %, 10 %, 25 %, 50 %, 75 %, 90 %, 99 %. (Vose 1996)



Figur 14: 5 av 10000 ruter - 0,05 % ( $5 \cdot 10^{-4}$ )

Det å kunne relatere små sannsynligheter til "kjente" tall, slik som 7 rette i Lotto, kan være hensiktsmessig. Imidlertid kan det være en fare ved at man ikke helt erkjenner hvor liten sannsynligheten er i Lotto og lignede spill, siden det omtrent hver lørdag er 3-4 personer som blir Lottomillionærer.

Mange av oss har kjørt E6 nordover, og har gjennom det et inntrykk av hvor langt det er fra Oslo til Mo i Rana (via Dovre). I følge gulesiders kart er distansen 993 km, nesten 1000 km. 1000 km er en million meter. Tenk deg at vi har sett oss ut en bestemt av disse en millioner meterne. La oss så kjøre fra Oslo til Mo, og du kan sette en krittstrek ett sted på den strekningen. Sannsynligheten for at du treffer den ene meteren vi så oss ut på forhånd kan uttrykkes ved en sannsynlighet lik  $1 \cdot 10^{-6}$ .

7 rette i Lotto har en sannsynlighet på  $1,859 \cdot 10^{-7}$ . Det er således 5 ganger mer sannsynlig at man skal lykkes med en strek på strekningen Oslo-Mo, enn i Lotto. Det spilles imidlertid i størrelsesorden 17-18 millioner rekker med Lotto hver uke (Norsk Tipping, 2011). Selv om det er helt usannsynlig at en enkelt rekke skal være riktig, vil det, når det tippes så mange rekker, som regel være noen som har 7 rette. Det er ganske vanlig å få 4+1 i Lotto, men husk at mange normalt spiller hver uke, med 10 rekker, og 520 rekker i året vil statistisk medføre 3 premier for 4+1 i året.

**Tabell 4: Små sannsynligheter og eksempler på størrelser**

Hva	Sannsynlighet	Strekning	Relativt
Sette et merke innenfor riktig meter i løpet av en meter.	$P=1$	1 meter ( $1^1$ )	Helt sikkert
Sette et merke innenfor riktig meter i løpet av 10 meter (husvegg).	$P=0,1$	10 meter ( $10^1$ )	
Sette et merke innenfor riktig meter i løpet av 100 meter (langside på idrettsbane).	$P=0,01$	100 meter ( $10^2$ )	4+ 1 i Lotto er 0,00586 pr rekke
Sette et merke innenfor riktig meter i løpet av 1 km (2,5 runder på en idrettsbane).	$P=0,001$	1 000 meter ( $10^3$ )	
Sette et merke innenfor riktig meter i løpet av 10 km (Skøyen-Sandvika langs E18).	$P=0,0001$	10 000 meter ( $10^4$ )	
Sette et merke innenfor riktig meter i løpet av 100 km (Oslo-Tønsberg langs E18).	$P=0,00001$	100 000 meter ( $10^5$ )	12 rette i tipping på en rekke $=1,88 \cdot 10^{-6}$ 7 rette i Lotto på en rekke $=1,859 \cdot 10^{-7}$
Sette et merke innenfor riktig meter i løpet av 1000 km (Oslo- Mo i Rana langs E6).	$P=0,000001$	1 000 000 meter ( $10^6$ )	

Aven anbefaler jevnlig trening med små tall og lite sannsynlige hendelser for å bli vant til så lave sannsynligheter, både for analytikere og eksperter. Med jevnlig trening vil man gradvis bli fortrolig med tallene, men dette alene vil ikke løse utfordringen mener Aven.

En trening med små tall med ekspertene før selve konverteringsprosessen vil derfor ha bedre nytte, for å sikre at ekspertene forstår verdiene, enn det som foreslås i mye av litteraturen, kalibreringsøvelser for å oppnå "riktige" verdier (i et frekventistisk syn).

## 5.2 Evaluering av ressursbruk og tillit

I kapittel 3 ble det definert to kriterier for å evaluere konverteringsprosessene, ressursbruk og tillit til resultatet. Som det fremgikk er tillit et absolutt krav, mens ressursbruk blir en relativ verdi.

De gjennomgåtte prosesser er i litteraturen ulikt beskrevet, har ikke helt like forutsetninger, og det er vanskelig å sammenligne prosessen med hverandre direkte. Det har vist seg at flere tar forbehold om at deler av innholdet må tilpasses situasjonen og analyseobjektet, hvilket gjør at prosessene ikke er så statiske som i hvert fall et par av dem kan fremstå før man setter seg grundig inn i dem. En evaluering har derfor begrenset nytte i en form for konkurranse mellom de ulike prosessene dersom man er ute etter en "vinner". En evaluering har bedre nytte dersom man kan gå gjennom sterke og svake sider for hver prosess knyttet til kriteriene, som et hjelpemiddel for å etablere en "best practice".

For å evaluere gjennomgås derfor prosessene i forhold til i hvilken grad deres prosess påvirker ressursbruk positivt (mindre ressursbruk enn snittet), nøytral (som snittet) eller negativt (mer ressursbruk enn snittet), og om prosessen er egnet for å skape tillit til resultatet. Ved evaluering av tillit benyttes de samme uttrykkene, negativ, nøytral, positiv.

Dette gjennomføres som en subjektiv vurdering av prosessene av forfatteren i denne oppgaven, ut fra erfaring med ekspertvurderinger i risikoanalyser, og med den kunnskap som er opparbeidet under utarbeidelsen.

**Tabell 5: Evaluering av prosesser**

Prosess	Ressursbruk	Tillit (subjektiv, reell ekspert, håndtere heuristikker og bias ++)
O'Hagan et al	O'Hagan har en 5-steps prosess som omhandler selve konverteringen. Det administrative er nedtonet, og fokus er på det tekniske. Det er allikevel ingen forhold som skiller prosessen positivt eller negativt fra de andre prosessene. Konklusjon: Nøytral	O'Hagan har en strukturert prosess som har fokus på å håndtere heuristikker og bias, men som har relativt lite fokus på å tilpasse seg analyseobjektet. Det legges vekt på at eksperten skal være reell ekspert. De har en klart frekventistisk tilnærming, og oppsummerer konverteringsprosessen i boka med å si blant annet: <i>"subjective probabilities can be well calibrated, but often they are not",</i> og <i>"it is important that feedback is given on training exercises to help the expert learn how to make better probabilities".</i> Det fremkommer tydelig at forfatterne har et syn som tilsier at noen verdier er riktige og andre feil, noe som ikke er forenlig med en subjektivistisk tilnærming. Tillit nevnes ikke som mål eller som viktig å oppnå ved planlegging av prosess. Konklusjon: Negativ
Meyer & Booker	Meyer og Booker har en svært omfattende prosess. Den starter med å forstå hensikten i prosjektet der risikoanalysen inngår. Det er	Meyer og Booker har en omfattende prosess, men denne omhandler ekspertvurderinger generelt, og ikke sannsynlighetsvurderinger spesielt.

Prosess	Ressursbruk	Tillit (subjektiv, reell ekspert, håndtere heuristikker og bias ++)
	<p>mye fokus også på det administrative, og det er den eneste av prosessene som faktisk omhandler informasjon vedrørende kostnader. De anbefaler at eksperter skal stille uten honorar, men vi ser bort fra den muligheten i risikoanalyser i dagens marked.</p> <p>Meyer &amp; Bookers prosess er langt mer omfattende enn de resterende. Den kan virke noe for omfattende for mindre analyser, og er nok i større grad tiltenkt store prosjekter med omfattende ekspertvurderinger, som utvikling av ny type atomkraftverk, månelandinger etc.</p> <p>Brukt på riktig måte i forhold til et lite prosjekts behov er den antagelig like god som de andre, men i den grad noen av prosessene er mer ressurskrevende, er det denne.</p> <p>Konklusjon: Negativ</p>	<p>De legger betydelig vekt på at man i en oppstartsfasen skal forstå hensikten med det man gjør, slik at man gjør de riktige aktivitetene i prosessen. Dette vil være med på å skape tillit til prosessen.</p> <p>Det legges vekt på at ekspertene skal være reell ekspert.</p> <p>De har fokus på å håndtere heuristikker og bias.</p> <p>Meyer og Booker har også en frekventistisk tilnærming til sannsynligheter, der de "riktige" verdiene skal tilstrebes, gjerne med trening og kalibrering av ekspertene. Dette er en vurdering ikke alle deler, og styrker ikke tilliten til resultatet av prosessen.</p> <p>Konklusjon: Nøytral/ Negativ</p>
Merkhofer	<p>Merkhofers prosess er som O'Hagans kun omfattet av selve konverteringen, når ekspertene er til stede. Det administrative er nedtonet, og fokus er på det tekniske. Det er ingen forhold som skiller prosessen positivt eller negativt fra de andre prosessene.</p> <p>Konklusjon: Nøytral</p>	<p>Merkhofers prosess er avgrenset til det tekniske. Han problematiserer ikke at prosess må tilpasses analyseobjektet, men fremstiller prosess som generell uten at dette fremstår unyansert med hans vinkling, siden den fokuserer på det tekniske.</p> <p>Han har fokus på å håndtere heuristikker og bias.</p> <p>Ikke opplysninger vedrørende sannsynlighetsoppfatning.</p> <p>Konklusjon: Nøytral</p>
Clemen & Reilly	<p>Clemen og Reillys prosess har 7 hovedaktiviteter, og starter før selve konverteringen. Det administrative er nedtonet, og fokus er på det tekniske. Det er ingen forhold som skiller prosessen positivt eller negativt fra de andre prosessene</p> <p>Konklusjon: Nøytral</p>	<p>Clemen og Reilly har en oversiktlig prosess hvor de har et subjektivistisk syn på sannsynligheter hvor sannsynligheten representerer ekspertens grad av tro.</p> <p>Det redegjøres for heuristikker og bias, også under aggregering.</p> <p>Legger vekt på å rekruttere riktige eksperter.</p> <p>De problematiserer ikke at prosessen må tilpasses analyseobjektet direkte,</p>



Prosess	Ressursbruk	Tillit (subjektiv, reell ekspert, håndtere heuristikker og bias ++)
		men legger stor vekt på å forstå behovet for og verdien i ekspertvurderingene slik at prosessen blir hensiktsmessig. Konklusjon: Positiv
Vose	Vose har ikke en definert prosess steg for steg, med fokus er på det tekniske. Det er ingen forhold som skiller prosessen positivt eller negativt fra de andre prosessene. Konklusjon: Nøytral	Vose erkjenner behov for subjektive beregninger, men har fokus på en frekventistisk tilnærming. Sier eksempelvis <i>"dersom man ikke har nok data til å spesifisere usikkerhetene til en variabel fullstendig..."</i> . Usikkerhet defineres som forskjellen mellom en sannsynlighet og hendelsenes <i>"sanne"</i> verdi, det vil si en typisk frekventistisk tankemåte som mange benytter, men som altså bygger på en forutsetning ikke alle deler. Vose beskriver på en god måte mange av de forhold som kan inntreffe underveis i prosessen, og som kan svekke resultatet (og dermed tilliten), og hvordan dette kan forbygges, men han nevner ikke tillit som et mål eller noe som må sikres oppnådd gjennom de tekniske metodene han skisserer. Konklusjon: Negativ
Encyclopedia	Encyclopedia starter noe tidligere i prosessen enn selve konverteringen, og starter med avklaringer av analyseobjekt etc. Det administrative er nedtonet, og fokus er på det tekniske. Konklusjon: Nøytral	Avdekker ikke syn på sannsynlighet. Fokus på at eksperter skal være riktige eksperter. Encyclopedia peker på at en strategi for prosessen må utarbeides basert på situasjonen, noe som øker tilliten. Det sies også at etikken i gjennomføringen må sikres, blant annet ved å evaluere om ekspertenes velferd er sikret under og etter prosessen. Avhengig av hva analyseobjektet er vil målet og resultatet av ekspertvurderingen være forskjellig. Det foreslås en strategi for å motvirke bias, noe som øker tilliten. Konklusjon: Nøytral
Aven	Aven kommenterer at ekspertvurdering kan være kostnadskreven av natur, men det er ikke opplysninger i hans	Aven vektlegger i stor grad tillit til resultatet. Han har imidlertid ikke laget en kokebok for prosessen som skal sørge for dette, men sier at en

Prosess	Ressursbruk	Tillit (subjektiv, reell ekspert, håndtere heuristikker og bias ++)
	litteratur som tilsier at ressursbruken skal være større eller mindre enn snittet av prosessene. Konklusjon: Nøytral	prosess bør lage grunnlag for konsensus, og bør tilpasses deretter. Aven mener det er umulig å skulle skaffe seg komplett kunnskap av "verden" for å kunne gjennomføre risikoanalyser, man vil aldri ha et slikt komplett bilde. Ekspertvurderinger er en god måte å få vite noe om det man ikke vet fra før. Selv om ekspertvurderingen er subjektiv er den ikke gal, tvert i mot. Man må forholde seg til usikkerhet, og den bruker han sannsynlighetsbegrepet til å si noe om. Aven vekker tillit når han legger vekt på å unngå heuristikker og bias, og sier at subjektive vurderinger må bygge på all tilgjengelig kunnskap. Konklusjon: Positiv

De forholdene som vekker tillit i konverteringsprosessene, og som det er ønskelig å implementere i en "best practice" hvor hensikten er å oppnå at oppdragsgiver har tillit til resultatet er:

- Erkjenne at det ikke finnes tilstrekkelig informasjon omkring hendelsene til å ha en frekvensbasert tilnærming, men heller bruke en bayesiansk tilnærming hvor vi betrakter sannsynlighetene som vår grad av tro på hendelsen som bygge spå all tilgjengelig kunnskap (Aven 2003, Clemen & Reilly 2001)
- Forstå oppdragsgivers nytte av vurderingen, og tilpasse konverteringsprosessen slik at resultatet kan brukes av oppdragsgiver.
- Vurdere analyseobjektet, og situasjonen for øvrig, med ressurser og tid, og sikre at den prosessen som designes tilpasses dette.
- Sikre at ekspertenes velferd er sikret, ved at de ikke eksponeres for risiko som følge av deltagelse i oppdraget (mer ved politiske, samfunnsmessige risikoanalyser).
- Håndtere heuristikker og bias proaktivt, lage strategi for håndtering og følge denne.

### 5.3 Oppsummering diskusjon

Forskning innen psykologi har vist at presise subjektive sannsynlighetsvurderinger ikke kan gjøres så enkelt som å spørre en ekspert om en sannsynlighet, helt uten videre. Vi har sett at i en utspørringssituasjon utsettes eksperten for flere uønskede påvirkninger, fra intervjueren, andre eksperter, skjemaer etc., noe som kan føre til skjevfordelinger i ekspertsvarene. For å redusere effekten av dette, er det nødvendig med en strukturert metode som ivaretar de kjente heuristikker og andre uønskede effekter av påvirkning (Merkhofer 1987).

Fra kapittel 2 vet vi at aktuelle standarder innen risikoanalyser krever at det i en risikoanalyse skal være kompetanse innen risikoanalytiske metoder (analytikerens) og kompetanse innen analyseobjektet (eksperten) og andre relevante fag (flere eksperter). Det kreves også at sannsynligheter beskrives i form av sannsynligheter eller frekvenser, hvilket nødvendiggjør en kvantifisering av kunnskapen (konvertering).

Sannsynlighetene fra en ekspertvurdering er subjektive, og vi legger en tilnærming til grunn der sannsynligheten reflekterer ekspertens grad av tro, ikke at sannsynligheten skal være så nær en "sann" sannsynlighet, da det ikke eksiterer sanne sannsynligheter, bare grad av tro (Clemen & Reilly 2001, Aven 2003).

I kapittel 3 så vi at et absolutt krav til prosessen er at beslutningstager har tillit til resultatet. Sannsynligvis er det viktigste for å oppnå tillit hos en beslutningstager at ekspertene er faglig dyktig, og faktisk vet hva han prater om. Av og til finnes det få eksperter som kan bidra med god og relevant kunnskap, og de kan også ha en interesse i saken, slik vi prøver å unngå. I valget med å bruke en ikke-ekspert, eller en ekspert som har interesse i saken, foretrekkes det siste, men det anbefales å dokumentere prosessen for å imøtegå senere spekulasjoner som måtte komme i etterkant (Meyer & Booker, 1991). En forutsetning for tillit er en subjektiv tolkning av sannsynlighetsbegrepet (Mengel, 2011)

Videre har vi i kapittel 4 og 5 sett at det er hensiktsmessig å tilpasse metoden etter situasjonen og analyseobjektet. Dette vil sammen med å forebygge skjevfordelinger være med på å skape tillit til resultatet fra ekspertvurderingene. Vi har sett at en mekanisk aggregering av både sannsynligheter og sannsynlighetsfordeling er enkelt, og kan benyttes dersom ikke aggregering skjer i en gruppe.

Det vekker tillit når prosessene håndterer heuristikker og bias proaktivt, og dette kan gjøres ved å lage en strategi for håndtering av heuristikker og bias og følge denne.

Det fremgår også at flere av prosessene bruker en terminologi som viser at forfatterne ser på sannsynlighet som riktige eller feil. Dette er ikke god bruk av ord, da man får inntrykk av at det er mulig å lage objektive sannsynligheter. Vi kan si at sannsynligheter er gode når de er fri for heuristikker og bias, men det blir feil når man sier at en sannsynlighet er "riktig".

## 6. REVIDERT PROSESS TIL BRUK I RISIKOANALYSER

Med bakgrunn i funn i kapittel 2,4 og 5 har denne forfatteren utarbeidet et forslag til en prosess som har til hensikt å konvertere kunnskap til sannsynligheter, på en måte som sikrer at beslutningstager har tillit til det resultat som konverteringsprosessen gir. Prosessen er gjengitt i prosesskart med forklaringer. Det er benyttet et såkalt ISOR- format (ISOR – Input, Styring, Output og Ressurser).

### 6.1 Overordnet modell

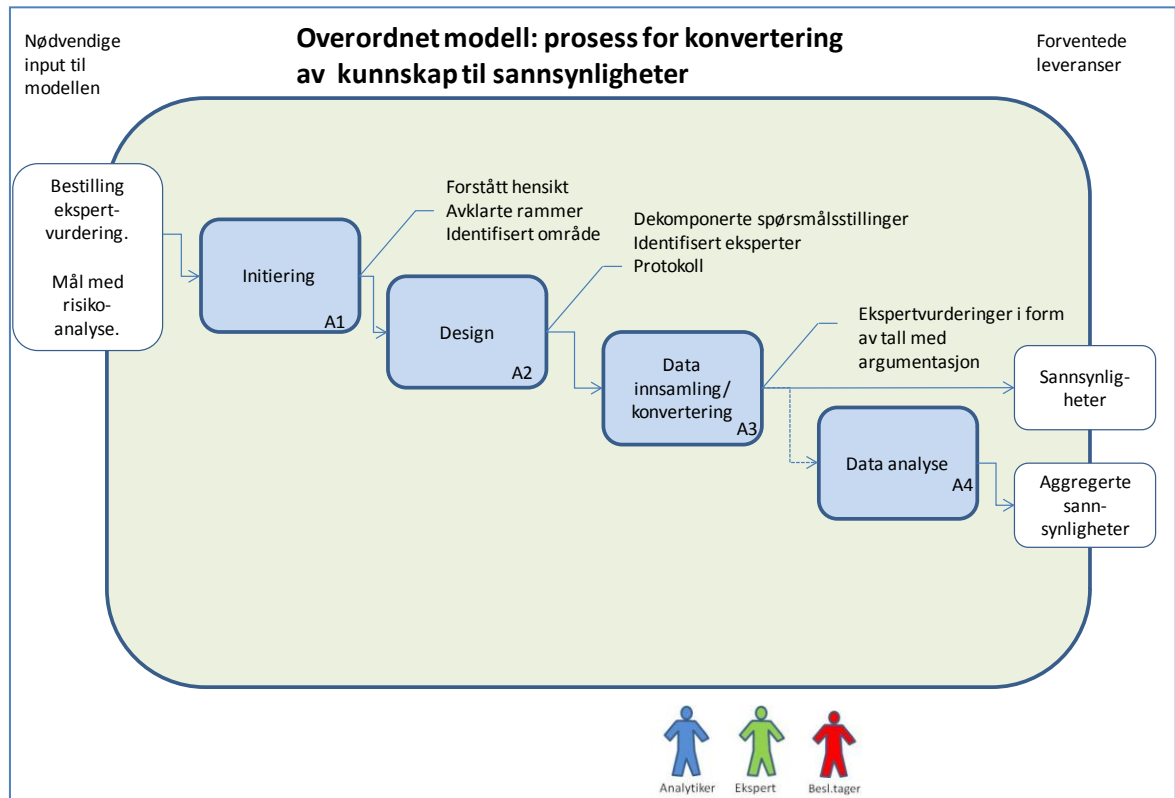
I prosessen er input i boksen til venstre, deretter følger aktiviteter (delaktiviteter) i blå bokser i midten, og resultat (output) i boks til høyre. Ressurser er tegnet i underkant.

I henhold til avklaring i dette kapittelet er det benyttet tre roller, ekspert, analytiker og beslutningstager. Analytiker brukes som fellesbetegnelse på intervjuer, statistiker, organisator og risikoanalytiker da dette ofte er samme person.

Den overordnede prosessen starter med et behov (bestilling) for en ekspertvurdering i en risikoanalyse. Av denne bør det fremgå en målsetting med analysen, slik at analytikeren kan designe en protokoll (uttrykk for den skrevne planen for konverteringsprosessen). Prosessen starter når behov for ekspertvurdering mottas, og avsluttes når en sannsynlighet eller aggregerte sannsynligheter leveres til oppdragsgiver.

Følgende aktiviteter inngår i prosessen:

- Aktivitet 1 (A1): Initierting
- Aktivitet 2 (A2): Design
- Aktivitet 3 (A3): Data innsamling/ konvertering
- Aktivitet 4 (A4): Data analyse



**Figur 15: Prosess for konvertering av kunnskap til sannsynligheter**

Initieringsaktiviteten har til hensikt å avklare hva hensikten og målet med ekspertvurderingen er, avklare hvilke ressurser som er tilgjengelig slik at protokollen designes hensiktsmessig i forhold til de rammene.

Designaktiviteten har til hensikt å sikre at prosessen gjennomføres på en hensiktsmessig måte som forebygger heuristikker og bias, og som er tilpasset analyseobjektet og de praktiske rammene.

Data innsamlings-/ konverteringsaktiviteten har til hensikt å fremskaffe ekspertkunnskap og konvertere denne til sannsynlighet(er). I en bayesiansk tilnærming kan denne sees på som data som for å oppdatere en a priori modell. Dersom man kun skal ha data for en ekspert avsluttes prosessen etter denne aktiviteten.

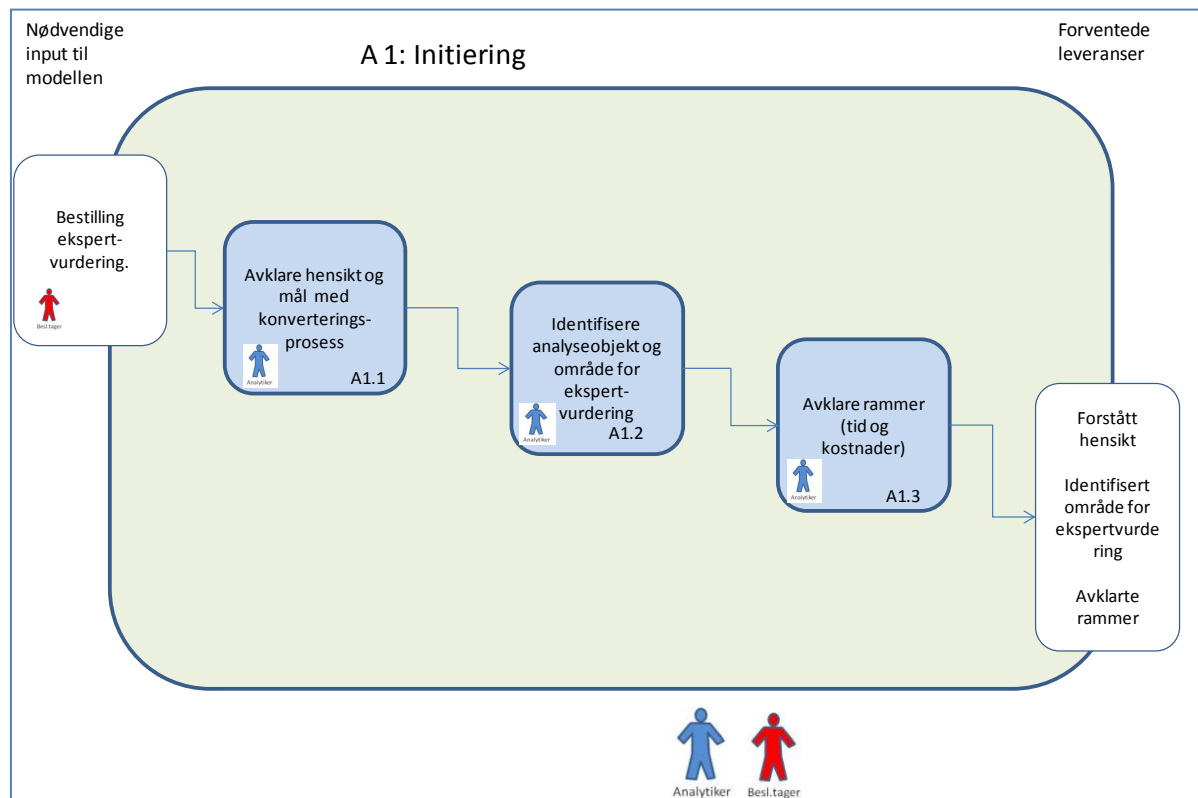
Data analyseaktiviteten benyttes når det er flere eksperter som har ulike sannsynligheter (data), og hvor dataene skal kontrolleres for systematiske bias, og deretter aggregeres mekanisk.

I det følgende gjennomgås de enkelte aktivitetene.

## 6.2 Initiering

Initieringsaktiviteten består av tre delaktiviteter:

- Avklare hensikt og mål med konverteringsprosess
- Identifisere områder for ekspertvurdering
- Avklare rammer (tid og kostnader)



Figur 16: Aktivitet 1 – Initiering – med delaktiviteter

### Avklare hensikt og mål med konverteringsprosess

I denne delaktiviteten mottar analytikeren en bestilling eller et behov fra en oppdragsgiver (beslutningstager eller person på vegne av beslutningstager). Det er vesentlig at analytiker forsikrer seg om at hensikten med bestillingen er forstått, at han forstår hva resultatet skal brukes til, slik at han sikrer et resultat beslutningstageren kan bruke. Han må forstå hva slags type analyseobjekt det er snakk om, og vurdere om det er behov for tilpasninger av prosessen.

### Identifisere områder for ekspertvurdering

Ut fra vurdering av analyseobjektet gjøres den første avklaringen av hvilket fagområde man skal søke eksperter i, om det er et enkelt, eller flere. En første dekomponering av problemstillingen kan skje herfor å finne hvilke ulike fagområder man trenger eksperter på.

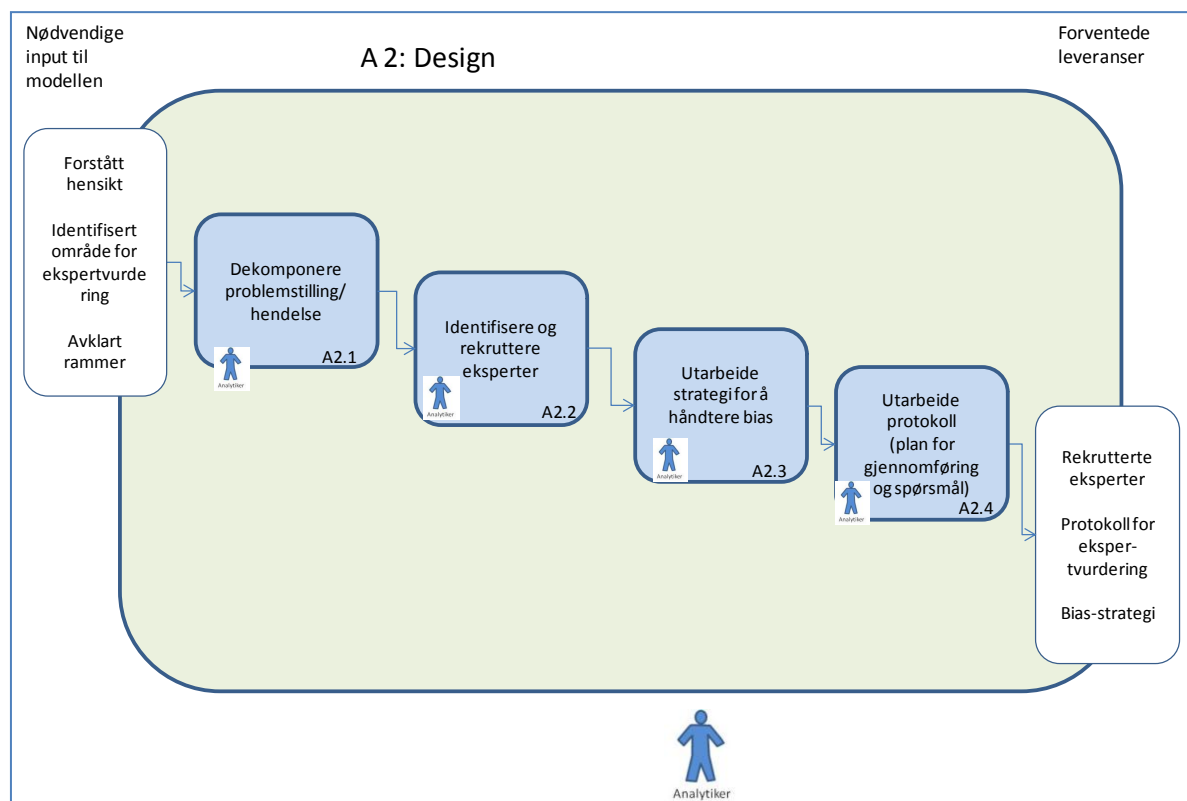
### Avklare rammer (tid og kostnader)

Analytiker må her avklare hvilke rammer som gjelder for oppdraget, for eksempel når resultatet skal være klart, og hvor store ressurser som kan benyttes på honorar til innleide eksperter, møtefasiliteter etc. I dialogen med beslutningstager (oppdragsgiver) har analytiker oversikt over områder man trenger eksperter på, og kan melde behov for å få en fornuftig prosess.

### 6.3 Design

Designaktiviteten består av 4 delaktiviteter:

- Dekomponere problemstilling/ hendelse
- Identifisere og rekruttere eksperter
- Utarbeide strategi for å håndtere bias
- Utarbeide protokoll



Figur 17: Delaktivitet 2 – Design – med delaktiviteter

#### Dekomponere problemstilling/ hendelse

Problemstillingen / hendelsen dekomponeres ned til et nivå som er hensiktsmessig, og så langt analytikeren har kunnskap til. I denne fasen kan det være aktuelt å leie inn ekstern ekspert også til dette, men da en annen enn den som skal gjennomføre ekspertvurderingen. Dekomponeringen avsluttes når man har de spørsmål som skal stilles, eller problemstillinger som skal forelegges ekspertene for videre dekomponering/ årsaksbeskrivelse.

Dekomponeringen er nødvendig både for å utarbeide en protokoll – en datainnsamlingsplan for vi dokumenterer hvilke spørsmål som stilles hvilken ekspert osv., men også for å identifisere fagområde i detalj slik at riktige eksperter kan identifiseres.

#### Identifisere og rekruttere eksperter

I denne delaktiviteten gjennomføres det søk og opprettes kontakt med fagpersoner som kan stille som ekspert i en konverteringsprosess. Det er av vesentlig betydning at eksperten faktisk besitter god kunnskap om de aktuelle problemstillingene, og det må dokumenteres at eksperten virkelig er ekspert. Dette kan gjøres ved å kartlegge utdannelse, arbeidserfaring, høre med referanser og andre eksperter på samme område etc. At eksperten faktisk er ekspert er helt vesentlig for at beslutningstager skal ha tillit til resultatet, og dersom en person ikke har relevant kunnskap, skal han ikke velges som ekspert.

Eventuelle interesser som ekspertene har i saken må kartlegges og dokumenteres. Det inngås avtale om møte for selve konverteringen, og nødvendig informasjon oversendes ekspertene i forkant.

#### Utarbeide strategi for å håndtere bias

Det utarbeides en strategi for å håndtere bias proaktivt under prosessen. Denne må som minimum inneholde en oversikt over hvilke heuristikker og bias som det er sannsynlig kan inntreffe i prosessen, og en plan for å forhindre dette, slik som opplæring av eksperter om heuristikker og bias, og en plan for overvåking av prosessen og kontroll av resultater.

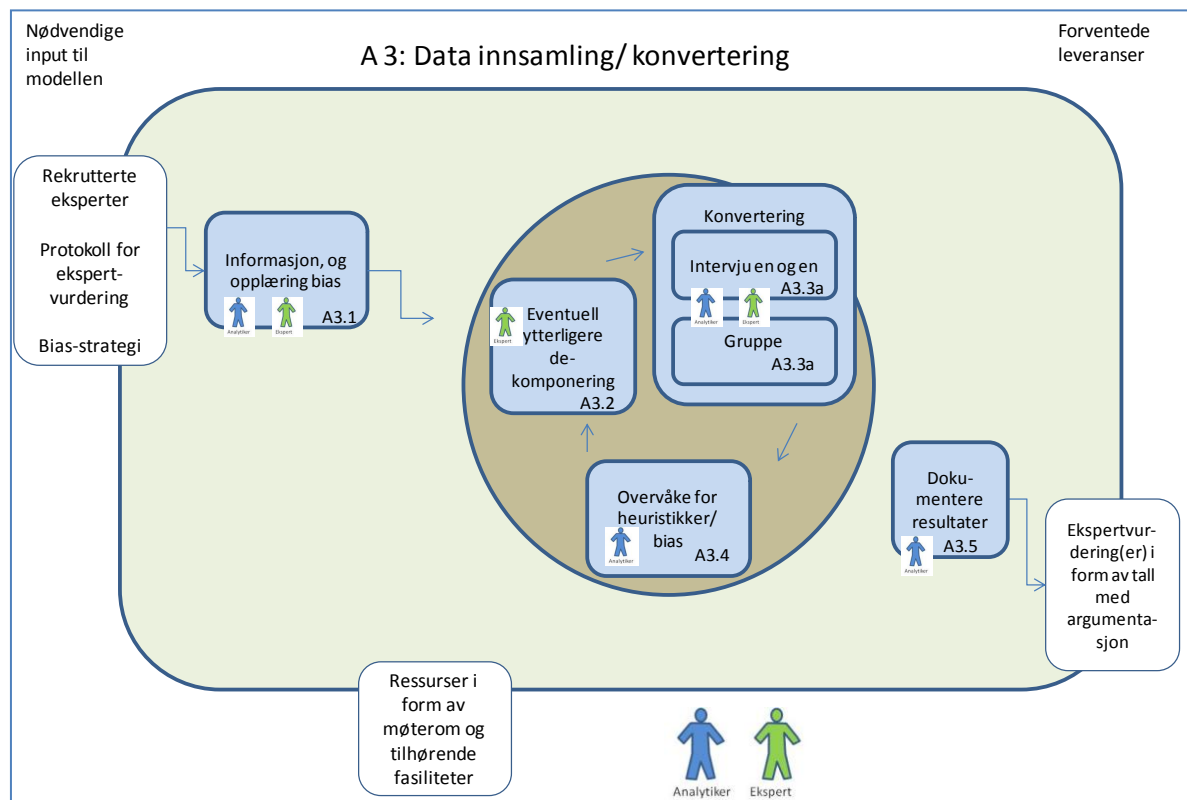
#### Utarbeide protokoll

Protokollen skal vise hvilke spørsmål som skal stilles til hvilke eksperter, hvilke analytikere som skal være hvor og gjøre hva, hvordan svarene skal håndteres og eventuelt aggregeres, og en tidsplan for dette. I store prosjekter kan dette være et omfattende dokument (data collection plan) som er et godt hjelpemiddel i prosessen, men dersom man skal ha en ekspert til å svare på et spørsmål, er dette mer en formalitet for å dokumentere prosessen.

## 6.4 Data innsamling/konvertering

Data innsamlingsaktiviteten består av fem delaktiviteter

- Informasjon og opplæring
- Eventuell ytterligere dekomponering
- Intervju/ gruppeprosess
- Overvåke for heuristikker
- Dokumentere resultater



Figur 18: Delaktivitet 3 – Data innsamling/ konvertering – med delaktiviteter



### Informasjon og opplæring

Når ekspertene er ankommet lokasjonen for konverteringen mottar de informasjon om prosessen. Dette kan være alt fra et eget møte første dag i en lang prosess, eller først punkt på en agenda i en times langt møte hvor konverteringen skjer – helt avhengig av situasjonen.

Det er helt vesentlig, for tilliten til prosessen, at det kan dokumenteres at opplæring av eksperter innen heuristikker og bias skjer før selve konverteringen. Det er også vesentlig at man får tak i ekspertens interesser i saken (motivasjonsheuristikker) her, dersom dette ikke allerede er kartlagt.

Videre skal opplæring av ekspertene i bruk av små sannsynligheter gjennomføres. Dette er ikke en kalibreringsøvelse, men en tilvenning i bruk av små tall, for å sikre at ekspertene er familiær med måten å uttrykke små tall på.

### Eventuell ytterligere dekomponering

Når ekspertene får problemstillingen eller spørsmålet forelagt, kan det hende at han/ de ønsker å gå videre i dekomponeringen, utover det som er gjort. Dette kan være for å kartlegge årsaker til en hendelse i detalj, eller for å komme på et nivå som er håndterbart for ekspertene for å sette sannsynligheter (ikke i nivå, men som ikke er for komplekst).

### Intervju/ gruppeprosess

Konverteringen kan enten gjennomføres med en og en i en intervjusituasjon eller ved hjelp av spørreskjema, eller som en gruppeprosess hvor flere eksperter er samlet. Målet er at ekspertene skal bruke all sin tilgjengelige kunnskap, og uttrykke sin grad av tro med hjelp av sannsynligheter. Det er vesentlig at det legges en subjektivistisk forståelse til grunn, og ikke en frekventistisk. Sannsynligheter skal ikke ses på som riktige eller gale, men et uttrykk for ekspertens grad av tro på om hendelsen inntreffer eller ikke. Det er allikevel mulig å si at det skal være god kvalitet på prosessen og sannsynlighetene, men da mener vi at de skal være fri fra heuristikker og bias. Det er viktig at analytikeren dokumenterer argumentasjonen/ begrunnelsen for ekspertens synspunkt, foruten svarene.

Med flere eksperter kan man ha en åpen diskusjon der ekspertene skal komme frem til en felles sannsynlighet, men slike prosesser utsettes for gruppedynamikk. Dersom man tror dette kan inntreffe bør man benytte Delphi-teknikker, der de ulike svarene er anonyme, men hvor ekspertene får tilgang til de andre ekspertenes svar og begrunnelse, og på bakgrunn av dette får mulighet til å justere sitt svar med bakgrunn i ny tilgjengelig kunnskap.

### Overvåke for heuristikker

Analytikeren må underveis i prosessen overvåke det som skjer med tanke på å oppdage heuristikker. Dette kan være om det fremkommer feil i spørsmål, om ekspert har motivasjoner som påvirker resultater, eller om ekspertene er så sliten, at man må gripe inn og justere tidsplanen for å gi ekspertene pause etc.

Ved behov for ytterligere dekomponering eller nye problemstillinger, gjennomføres de tre sistnevnte delaktiviteter igjen, til man er ferdig.

### Dokumentere resultater

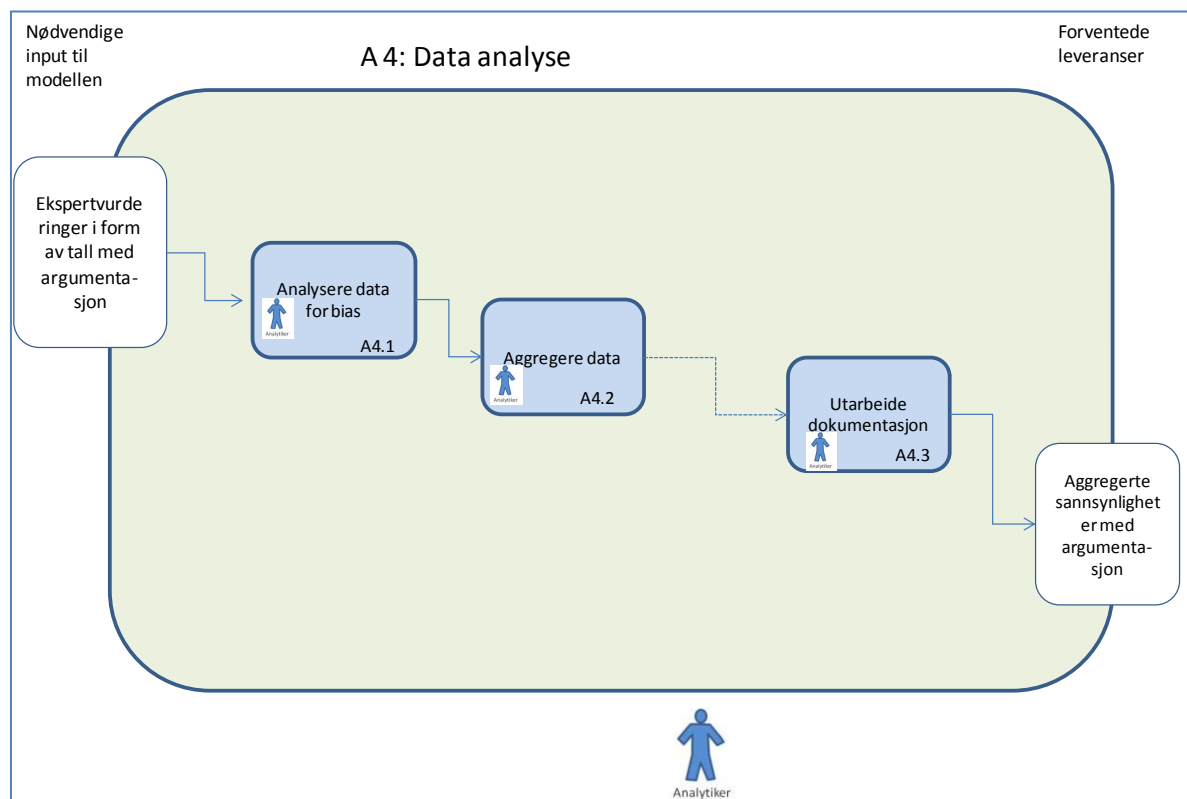
Resultatene fra konverteringen dokumenteres i et egnet format, eksempelvis et notat hvor det fremkommer hva problemstillingen er, hvem som har vært med på vurderingen, bakgrunnen til ekspertene, hvilken informasjon som er tilgjengelig samt svaret med en grundig begrunnelse for svaret. Ved utarbeidelse av en enkelt sannsynlighet rapporteres dette til beslutningstager – som ekspertens subjektive tro på at hendelsen inntreffer.

Ved bruk av flere eksperter fortsetter man til aktivitet 4 – data analyse

## 6.5 Data analyse

Data analyseaktiviteten består av tre delaktiviteter

- Analysere data for bias
- Aggregere data
- Utarbeide dokumentasjon



**Figur 19: Delaktivitet 4 – Data analyse – med delaktiviteter**

### Analysere data for bias

Når flere eksperter gir sine vurderinger, kan det være nyttig å analysere disse for å kontrollere om det er korrelasjon i resultatene når vi ser på den enkeltes person og verdier. Dersom noen er veldig mer positiv enn de resterende kan det være som en følge av motivasjonsheuristikk, misforståelser i benevnelser eller måleenheter etc. Ved funn bør disse sjekkes ut med den enkelte ekspert.

### Aggregere data

En mekanisk aggregering av data kan skje ved å finne aritmetisk gjennomsnitt, median og geometrisk gjennomsnitt. Som vi har sett er det i mange tilfeller riktig å bruke geometrisk gjennomsnitt i stedet fremfor det aritmetiske gjennomsnittet. De aggregerte data kan brukes som input i en risikoanalyse. Både enkeltsannsynligheter og fordelinger kan aggregeres.

### Utarbeide dokumentasjon

Resultatene fra konverteringen dokumenteres i et egnet format, eksempelvis et notat hvor det fremkommer hva problemstillingen(e) er, hvem som har vært med på vurderingen(e),

bakgrunnen til ekspertene, samt de enkelte svarene med begrunnelse for svarene, samt aggregerte data.

Dataene som oversendes beslutningstager (oppdragsgiver) kan nå sees på som nye data, og brukes til å oppdatere beslutningstager a priori sannsynlighet dersom denne foreligger, slik at han får en a posteriori sannsynlighet.

## 7. KONKLUSJON

Den overordnede problemstillingen for denne oppgaven er, med grunnlag i en risikoanalyse hvor det er behov for sannsynligheter, "hvordan bør en prosess for å konvertere kunnskap til sannsynlighet være for best mulig å skape tillit hos en beslutningstager"?

For å finne en argumentasjon og en løsning på dette ble først kravene i relevante standarder og veiledninger identifisert, hvor det fremkom at det stilles relativt få krav, bortsett fra at man skal ha relevant kompetanse involvert i prosessen, og at sannsynligheten skal kvantifiseres.

For å evaluere et sett med metoder som er representative for forskningsfronten ble det utarbeidet to evalueringskriterier; tillit og ressursbruk. Det ble vanskelig å benytte sistnevnte kriterium, da de fleste kriteriene tok forbehold om at prosessen måtte tilpasses situasjonen. For å ha tillit fant vi at det må legges en subjektiv fortolkning av sannsynlighetsbegrepet til grunn, hvor sannsynlighet er ekspertens grad av tro på at hendelsen inntreffer eller ikke, og at bruk av en frekventistisk tilnærming gir en falsk illusjon om objektiv sannsynlighet som gjør at prosessen ikke inngir tillit. Videre må prosessen forhindre heuristikker og bias, og vi må være sikre på at eksperten faktisk er ekspert innen det område vi trenger ekspertkompetanse på.

Vi kartla forskningsfronten, og fant ut at det er utgitt en mengde bøker og artikler om konvertering av kunnskap til sannsynligheter, men at dette neppe er det mest utforskede område i statistikken. Det er allikevel skrevet så mye at det ikke er mulig å komme gjennom alt i denne omgang, og det ble gjort et utvalg basert på de mest henviste, av oppslagsverk og lærebøker, de fremste norske innen risikoanalyse, samt flere for å vise noe av bredden.

Det ble avdekket at de fleste har fokus på å håndtere heuristikker og bias, og på å skaffe reelle eksperter i vurderingene. Flere av forfatterne har imidlertid en beskrivelse av konverteringsprosessen hvor man legger en tilnærming til sannsynlighet som er frekventistisk, og hvor det skapes en illusjon av objektivitet rundt konverteringsprosessene.

Videre fant vi at det er hensiktsmessig å tilpasse metoden etter situasjonen og analyseobjektet. Dette vil sammen med å forebygge skjevfordelinger være med på å skape tillit til resultatet fra ekspertvurderingene. Vi har sett at en mekanisk aggregering av både sannsynligheter og sannsynlighetsfordeling er enkelt, og kan benyttes dersom ikke aggregering skjer i en gruppe.

Denne oppgaven skulle bidra til mer viten omkring hvordan en prosess for å konvertere kunnskap til sannsynlighet for best mulig å skape tillit hos en beslutningstager skulle være. I stedet for å rangere metodene og kåre en "vinner", ble funnene benyttet til å lage et forslag til en ny "best practice". Denne er forholdsvis lik flere av de viste prosesser, men hensikten er at den på en tydeligere måte skal sikre at ekspertene er eksperter, at det tas hensyn analyseobjektet, og at heuristikker og bias håndteres proaktivt. En subjektivistisk forståelse av sannsynlighetsbegrepet, og hvor resultatet representerer ekspertens grad av tro, basert på den tilgjengelige informasjonen, er en forutsetning som ligger til grunn for prosessen.

Proessen består av fire aktiviteter (som igjen er delt i underaktiviteter);

- **Initiering:** har til hensikt å avklare hva hensikten og målet med ekspertvurderingen er, avklare hvilke ressurser som er tilgjengelig slik at protokollen designes hensiktsmessig i forhold til de rammene.

- Design: har til hensikt å sikre at prosessen gjennomføres på en hensiktsmessig måte som forebygger heuristikker og bias, og som er tilpasset analyseobjektet og de praktiske rammer.
- Data innsamlings-/ konvertering: har til hensikt å fremskaffe ekspertkunnskap og konvertere denne til sannsynlighet(er). I en bayesiansk tilnærming kan denne sees på som data som for å oppdatere en a priori modell. Dersom man kun skal ha data for en ekspert avsluttes prosessen etter denne aktiviteten.
- Data analyse: benyttes når det er flere eksperter som har ulike sannsynligheter (data), og hvor dataene skal kontrolleres for systematiske bias, og deretter aggregeres mekanisk.

## 8. REFERANSER

Aschehoug og Gyldendals. Store Norske leksikon. Aschehoug og Gyldendal.

Aven, Terje. Foundations of Risk Analysis. Wiley, 2003.  
— . Misconceptions of Risk. Wiley, 2010.

Bates, A.W., og Albert Sangra. Managing technology in higher education. Wiley/ Jossey-Bass, 2011.

Bø, Inge, og Lars Helle. Pedagogisk ordbok. Universitetsforlaget, 2006.

Clemen, Robert T., og Terence Reilly. Making hard decisions. Duxbury Thomson learning, 2001.

Cooke, Roger M. Experts in uncertainty. Oxford University press, 1991.

COSO - The committee of sponsoring Organizations of the Treadway Commission. Helhetlig risikostyring - et integrert rammeverk. Norges Interne Revisorers Forening, 2005.

Direktoratet for samfunnssikkerhet og beredskap (DSB). Samfunnssikkerhet i arealplanlegging - kartlegging av risiko og sårbarhet. Veileder, DSB, 2010.

Dr. Randrup-Thomsen, Søren, intervjuet av Tore Askeland. Avdelingsleder Risiko og sikkerhet i Rambøll Danmark (Mai 2012).

Dr. Thomsen, Jan, intervjuet av Tore Askeland. Chief Risk Officer, Norges Bank Investment Management (April 2012).

G.Løvås, Gunnar. Statistikk for universiteter og høyskoler. Universitetsforlaget, 2004.

Hacking, Ian. The emergence of probability. Cambridge, 2006.

ISO - the International Standardization for Organization. Risk Management - principles and guidelines (ISO 31000). ISO, 2009.

Kahneman, Daniel, Paul Slovic, og Amos Tversky. Judgment under uncertainty: Heuristics and biases. Cambridge, 1982.

Melnick, Edward L, og Brian S. Ever. Encyclopedia of quantitative risk analysis and assessment. Wiley, 2008.

Mengel, Marius. «Styring av operasjonell risiko - Tillit og beslutningsstøtte.» Masteroppgave ved UiS, 2011.

Merkhofer, Miley W. «Quantifying Judgemental Uncertainty: Methodology.» IEEE transactions on systems, man and cybernetics, 1987: 741-752.

Meyer, Mary A., og Jane M. Booker. Eliciting and analyzing expert judgement - a practical guide. SIAM/ASA.

Norsk Tipping. Omsetning 2010 og 2011. 2011. [https://www.norsk-tipping.no/selskapet/om\\_norsk\\_tipping/page?id=1070](https://www.norsk-tipping.no/selskapet/om_norsk_tipping/page?id=1070).

Oakley, Jeremy E., og Anthony O'Hagan. Probability is Perfect, but We Can't Elicit it perfectly. University of Sheffield, UK, 2002.

O'Hagan, Anthony, og Paul H. Gartwaite. Uncertain Judgements - eliciting expert probabilities. Wiley, 2006.

Rausand, Marvin, og Ingrid Bouwer Utne. Risikoanalyse - teori og metoder. Tapir akademisk forlag, 2009.

Røed, Willy, Terje Aven, og Hermann S. Wiencke. Risikoanalyse. Universitetsforlaget, 2008.

Schein, Edgar. Organizational Culture and Leadership. Bassey Bass, 2010.

Shafer, Glenn. «Lindley's paradox.» Journal of the American Statistical Association 77 (378):, 1982: 325–334.

Standard Norge. NS 5814, Krav til Risikovurderinger. Standard Norge, 2008.

Standard Norge. NS 5815, Risikovurdering av anleggsarbeid. Standard Norge, 2006.

Vose, David. Quantitative Risk Analysis. Wiley, 1996.

Walpole, Ronald E., Raymond H. Myers, Sharon L. Myers, og Keying Ye. Probability & statistics for engineers & scientists. Pearson, 2007.

## **VEDLEGG**



## **VEDLEGG 1 – LINDLEYS PARADOKS**

**"Lindley's paradox". Shafer, Glenn (1982). Journal of the American Statistical Association 77 (378): 325–334):**

Tenk deg et land med 49 581 guttefødsler og 48 870 jentefødsler i en gitt tidsperiode. Den observerte andelen guttefødsler  $X$  blir da  $49\,581/(49\,581 + 48\,870) = 0,50361$ . Vi forutsetter at guttefødslerne er binomial fordelte med parameter  $\theta$ . Vi ønsker å teste hvorvidt  $\theta$  er 0,5 eller om det er en annen verdi. Dette gir  $H_0: \theta=0,5$ , og alternativet  $H_1 \neq 0,5$ .

Frekventisters tilnærming for å hypoteseteste  $H_0$  er å beregne en p-verdi, det vil si sannsynligheten for å observere en andel av gutter som minst er like stor som  $X$ , forutsatt at  $H_0$  er sann. Fordi antall fødsler er svært høyt, vil man bruke forutsette en normalfordeling for en andel av guttefødsler:

$X \sim N(\theta, \sigma)$ , med  $\sigma^2 = \theta(1 - \theta)$ , for å beregne

$$P(X \geq x | \theta = 0,5) = \int_{x \approx 0,5036}^1 \frac{1}{\sqrt{2\pi\sigma}} e^{-\left(\frac{u-0,5}{\sigma}\right)^2} / 2 \, du \approx 0,0117$$

Vi ville blitt like overrasket om vi hadde sett 49,581 jentefødsler som det observerte antall guttefødsler, altså for eksempel  $X \approx 0,4964$ , så en frekventist ville vanligvis gjennomføre en tosidig test, hvor p-verdien ville være  $P \approx 2 \times 0,0117 = 0,0234$ . I begge tilfeller er p-verdien lavere enn signifikansnivået på 5 %, hvilket medfører at frekventister forkaster  $H_0$  da den ikke sammenfaller med de observerte data.

Forutsatt at man ikke har grunn til å favorisere den ene hypotesen fra den andre, ville man med en bayesiansk tilnærming angi a priori sannsynligheter:

$$\pi(H_0) = \pi(H_1) = 0,5$$

og deretter beregne a posteriori sannsynlighet  $H_0$  ved bruk av Bayes teorem:

$$P(H_0|k) = \frac{P(k|H_0) \pi(H_0)}{P(k|H_0) \pi(H_0) + P(k|H_1) \pi(H_1)}$$

Etter å ha observert  $k=49\,581$  gutter av  $n=98\,451$  fødsler, kan vi beregne a posteriori sannsynlighet av hver hypotese ved å bruke sannsynlighetsfordelingen for en binomial variabel.

$$P(k|H_0) = \binom{n}{k} 0,5^k (1 - 0,5)^{n-k} \approx 0,000195$$

$$P(k|H_1) = \int_0^1 \binom{n}{k} u^k (1 - u)^{n-k} \, du \approx \binom{n}{k} B(k + 1, n - k + 1) \approx 0,0000101$$

hvor  $B(a,b)$  er Betafunksjonen. Fra disse verdiene finner vi a posteriori sannsynlighet av  $P(H_0|k) \approx 0,95$ , noe som jo sterkt favoriserer  $H_0$  over  $H_1$ . Løsningene til de to fremgangsmåtene ser ut til å være i konflikt med hverandre, og det er dette som er "paradokset".