

UNIVERSITETET I OSLO

Det matematisk-naturvitenskapelige fakultet

Obligatorisk oppgave: STK 2400 - Elementær innføring i risiko- og pålitelighetsanalyse
Innleveringsfrist: Torsdag 8. november 2012, kl. 14.30
Innleveringssted: Matematisk institutts ekspedisjonskontor, 7. etasje, N. H. Abels hus.

Oppgave 1

En fordeling som benyttes mye i risikoanalyse er *lognormalfordelingen*. En stokastisk variabel, Y , sies å være lognormalfordelt med parametre μ og σ dersom $X = \ln(Y)$ er normalfordelt med forventning μ og standardavvik σ . Vi skriver i så fall $Y \sim \text{lognorm}(\mu, \sigma)$.

(a) Anta at $U \sim N(0, 1)$, dvs. U er standard normalfordelt, og la $Y = \exp(\sigma U + \mu)$. Forklar kort hvorfor det da følger at $Y \sim \text{lognorm}(\mu, \sigma)$.

(b) Anta at $Y \sim \text{lognorm}(\mu, \sigma)$, og la y_α være slik at:

$$(1.1) \quad \Pr(Y \leq y_\alpha) = \alpha \%, \quad \text{der } 0 \leq \alpha \leq 100$$

[Med denne notasjonen er altså f.eks. y_{10} lik 10%-fraktilen i $\text{lognorm}(\mu, \sigma)$ -fordelingen.]

Vis at:

$$(1.2) \quad y_\alpha = \exp(\sigma u_\alpha + \mu),$$

der u_α er α %-fraktilen i standard normalfordelingen.

(c) Vis at:

$$(1.3) \quad \frac{y_{90}}{y_{50}} = \frac{y_{50}}{y_{10}}$$

(d) Anta at $y_{10} = 5.0$ og at $y_{90} = 20.0$. Finn den tilhørende verdien av y_{50} . Bestem også verdiene på μ og σ for denne fordelingen. Du kan her benytte at $u_{90} = 1.28155$. [Svar: $\mu = 2.30259$, og $\sigma = 0.54087$.]

(e) Benytt Riscue til å lage en simuleringsmodell bestående av kun én node. Kall denne noden Y . (Se Figur 1 under.)



Figur 1. Simuleringsmodell med kun én node

Som funksjon i denne noden benyttes:

$$(1.4) \quad \text{EXP}(\text{SUM}(\text{PRODUCT}(\text{NORMAL}(); 0.54087); 2.30259))$$

Kjør en simulering på denne modellen. Benytt f.eks. 10000 iterasjoner. Lag et S-kurve-plot og et tetthetsplot for resultatnoden. Kontroller at 10%-fraktilen og 90%-fraktilen for

resultatnoden blir (omtrent) det ønskede (dvs. henholdsvis 5.0 og 20.0). Lag også et plot som viser hvordan disse fraktilverdiene konvergerer. Gi en kort kommentar til resultatene.

Vi skal nå se hvordan vi kan konstruere en *tidsrekkemodell* der alle variablene er lognormalt fordelte. For å konstruere denne starter vi først med en rekke $\{W_n\}$ av *uavhengige* standard normalfordelte variable. Basert på denne rekken danner vi så en ny rekke $\{U_n\}$ som følger:

$$(1.5) \quad U_1 = W_1, \quad U_n = aU_{n-1} + bW_n, \quad n = 2, 3, \dots$$

Vi bemerker at siden U_{n-1} inngår i uttrykket for U_n , $n = 2, 3, \dots$, så blir variablene i denne rekken stokastisk avhengige.

(f) Bestem konstantene a og b slik at $U_n \sim N(0, 1)$ og $\text{Corr}(U_n, U_{n-1}) = c$, $n = 2, 3, \dots$, der c er en passende konstant slik at $|c| \leq 1$. [Svar: $a = c$ og $b = \text{sqrt}(1-c^2)$.]

(g) Vi skal nå benytte **Riscue** til å lage en simuleringsmodell bygget opp som i punktet over. For at vi på en enkel måte skal kunne variere korrelasjonen c , skal vi la denne være representert i modellen som en parameter. En slik parameter opprettes ved først å åpne parametervinduet, ved å benytte meny-kommandoen: "Model>Parameter table...". Deretter oppretter vi en parameter ved å benytte meny-kommandoen: "Model>Add Parameter". Til slutt fyller vi inn informasjon i de ulike kolonner slik at vi ender opp som vist i Figur 2.

Name	ID	Def. value	Values	Multivalued	Info
Tidskorrelasjon	0	0.5	0.1; 0.2; ...; 1.0	<input type="checkbox"/>	Korrelasjon mellom påhverandre følgende ledd

Figur 2. Parametertabell

Tips: Vær spesielt nøye med hva du taster inn i kolonnen "Values". Som skilletegn mellom tallene skal det benyttes *semikolon*. Mellom 0.2 og 1.0 skal det stå: "semikolon, blank, punktum, punktum, punktum, blank, semikolon, blank".

Bemerk at verdien i kolonnen "ID", dvs. 0, blir fylt inn automatisk. Denne verdien benyttes når vi skal benytte parameteren i formler. Vi kan da bruke formelen $\text{PAR}(0)$, der argumentet er nettopp denne ID-verdien. Avkrysningsfeltet i kolonnen "Multivalued" skal foreløpig *ikke* stå avkrysset, hvilket betyr at modellen skal benytte parameterverdien som står i kolonnen "Def. value", dvs. 0.5.

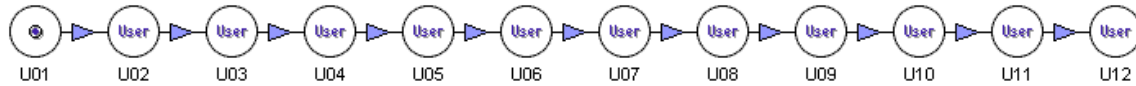
For å gjøre modellen vår litt lettere å lese, skal vi også lage noen brukerdefinerte funksjoner. Vi åpner derfor vinduet med slike funksjoner ved å benytte meny-kommandoen: "User>User Functions". Deretter oppretter vi to slike funksjoner ved å benytte meny-kommandoen: "User>Add User Function", og fyller inn detaljene slik at vi til slutt ender opp som i Figur 3.

Function name	Function
a	PAR(0)
b	SQRT(DIFFERENCE(1; SQR(PAR(0))))

Figur 3. Brukerdefinerte funksjoner

Vi legger merke til at siden $PAR(0)$ refererer til korrelasjonen, dvs. c , så blir de to brukerdefinerte funksjonene, kalt a og b henholdsvis, nettopp det vi fant som uttrykk for konstantene a og b i punktet over.

Vi kan nå lage vår tidsrekkemodell som en kjede av $n = 12$ noder som illustrert i Figur 4.



Figur 4. En tidsrekkemodell

For node **U01** benyttes følgende formel:

$$(1.6) \quad \text{NORMAL}()$$

For alle de øvrige nodene benyttes følgende formel:

$$(1.7) \quad \text{SUM}(\text{PRODUCT}(a()); \text{INPUT}()); \text{PRODUCT}(b()); \text{NORMAL}())$$

Kjør en simulering på denne modellen. Lag en korrelasjonstabell for resultatnodene og kontroller at korrelasjonen mellom to på hverandre følgende ledd er omtrent 0.5. Kommenter resultatene.

(h) Vi skal nå transformere tidsrekken fra punktet over til en tidsrekke, $\{Y_n\}$ der alle variablene er lognormalt fordelte (men ikke nødvendigvis identisk fordelte). Siden vi vet at $U_n \sim N(0, 1)$, $n = 1, 2, \dots$, kan dette oppnås ved å la:

$$(1.8) \quad Y_n = \exp(\sigma_n U_n + \mu_n), \quad n = 1, 2, \dots$$

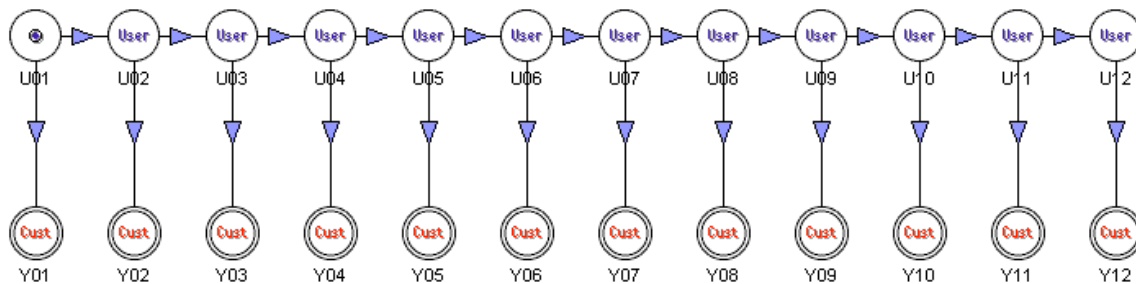
der σ_n og μ_n er beregnet basert på informasjon om 10%- og 90%-fraktilene til Y_n , $n = 1, 2, \dots$. I Tabell 1 har vi angitt disse fraktilene:

	10%	90%
1	5.0	8.2
2	8.0	11.0
3	10.0	12.5
4	10.0	12.5
5	10.0	12.5
6	8.0	11.0
7	6.5	10.0
8	5.0	8.2
9	3.3	7.1
10	2.5	5.5
11	2.1	4.6
12	1.7	3.1

Tabell 1. Fraktiler for tidsrekken

Beregn σ_n og μ_n , $n = 1, 2, \dots, 12$, basert på informasjonen i Tabell 1.

Ta så utgangspunkt i modellen i Figur 4, og lag en ny rad med noder under den første. De nye nodene gir du navn $Y01, \dots, Y12$ henholdsvis. Trekk så piler fra $U01$ til $Y01$, fra $U02$ til $Y02$, osv. til du får et diagram som vist i Figur 5.



Figur 5. To linkede tidsrekke modeller

For alle de nye nodene benyttes en formel av form:

$$(1.9) \quad \text{EXP}(\text{SUM}(\text{PRODUCT}(\langle \sigma \rangle; \text{INPUT}()); \langle \mu \rangle))$$

der $\langle \sigma \rangle$ og $\langle \mu \rangle$ substitueres de respektive verdiene av σ_n og μ_n , som du alt har beregnet ut fra Tabell 1.

Kjør en simulering på denne modellen. Lag en korrelasjonstabell for resultatnodene for de nye nodene. Synes du det ser ut til at korrelasjonen mellom to på hverandre følgende ledd fortsatt er omtrent 0.5? Kommenter resultatene.

Tips: For å unngå å få så mange irrelevante resultatnoder, kan du med fordel la være å lagre resultater for de gamle nodene, dvs. U01, ..., U12. Dette gjøres ved å selektere alle disse nodene, høyre-klikke på én av dem, og velge "Don't store results" fra menyen som dukker opp.

(i) Tidsrekken $\{Y_n\}$ representerer årlig produksjon (angitt i millioner fat) fra et oljefelt med "økonomisk levetid" på 12 år. For å finne inntekten (angitt i millioner norske kroner) fra feltet i denne perioden, må vi multiplisere med oljeprisen (angitt i dollar pr. fat) og dollarkurs (angitt i kroner pr. dollar). Både oljepris og dollarkurs er naturligvis usikre størrelser. Det er imidlertid vanlig å legge disse størrelsene inn i modellene som parametere slik at man kan foreta sensitivitetsanalyser.

I tillegg ønsker man gjerne å "neddiskontere" inntekter som kommer i fremtiden med en passende fast rentesats. Resultatet av dette omtales gjerne som "nåverdien" av inntektene.

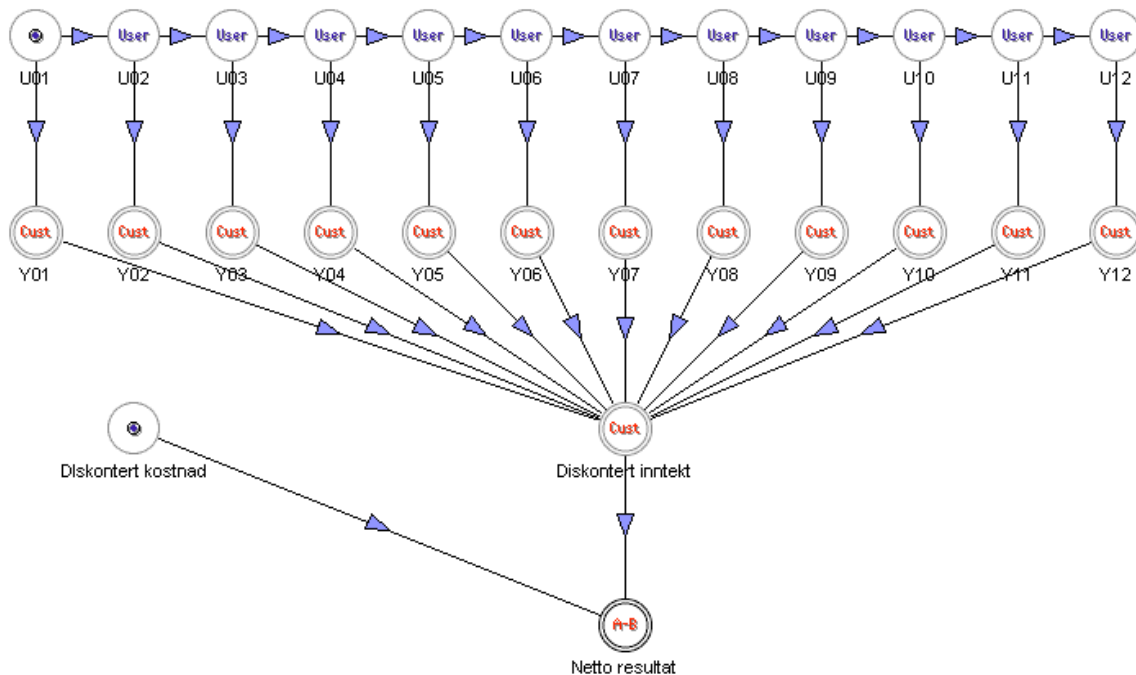
Opprett parametere for oljepris, dollarkurs og diskonteringsrente som vist i Figur 6.

Name	ID	Def. value	Values	Multivalued	Info
Tidskorrelasjon	0	0.5	0.1; 0.2; ... ; 1.0	<input type="checkbox"/>	Korrelasjon mellom påhverandre følgende ledd
Oljepris	1	18	8; 10; ... ; 20	<input type="checkbox"/>	Dollar pr. fat
Dollarkurs	2	7.7		<input type="checkbox"/>	Kroner pr. dollar
Diskonteringsrente	3	0.07		<input type="checkbox"/>	7 %

Figur 6. Parametertabell

Vi føyer så til tre nye noder i modellen: Diskontert inntekt, Diskontert kostnad og Netto resultat. Fra hver av nodene Y01, Y02, ..., Y12 trekker vi piler ned til Diskontert inntekt. Endelig trekker vi piler fra Diskontert inntekt og Diskontert kostnad til Netto resultat. Pass på at du oppretter pilen fra Diskontert inntekt til Netto resultat *før* du oppretter pilen

fra Diskontert kostnad til Netto resultat. Resultatet skal være et diagram som det som er illustrert i Figur 7.



Figur 7. Totalmodell

I Diskontert inntekt benyttes følgende formel:

$$(1.10) \quad \text{PRODUCT}(\text{INPUTNPV}(\text{PAR}(3)); \text{PAR}(1); \text{PAR}(2)),$$

mens i Diskontert kostnad skal du sette inn følgende formel:

$$(1.11) \quad \text{NORMAL}(6000; 1000).$$

Endelig skal du sette inn følgende formel i Netto resultat:

$$(1.12) \quad \text{INPUTDIFF}().$$

Sørg også for at den eneste noden som det lagres resultater for er **Netto resultat**. Gi en kort beskrivelse av hvordan formel (1.10) fungerer.

Kjør en simulering på denne modellen. Lag et S-kurve-plot og et tetthetsplot for resultatnoden (kun for **Netto resultat**). Prøv å anslå ut fra S-kurven hvor stor sannsynligheten er for at prosjektet gir et netto resultat som er lavere enn en milliard.

(j) Vi skal i dette punktet undersøke hvor følsomme resultatene er for variasjoner i tids-korrelasjonen, dvs. parameteren med ID = 0. Vi går derfor inn i parametervinduet igjen og krysser av for "Multivalued" for denne parameteren.

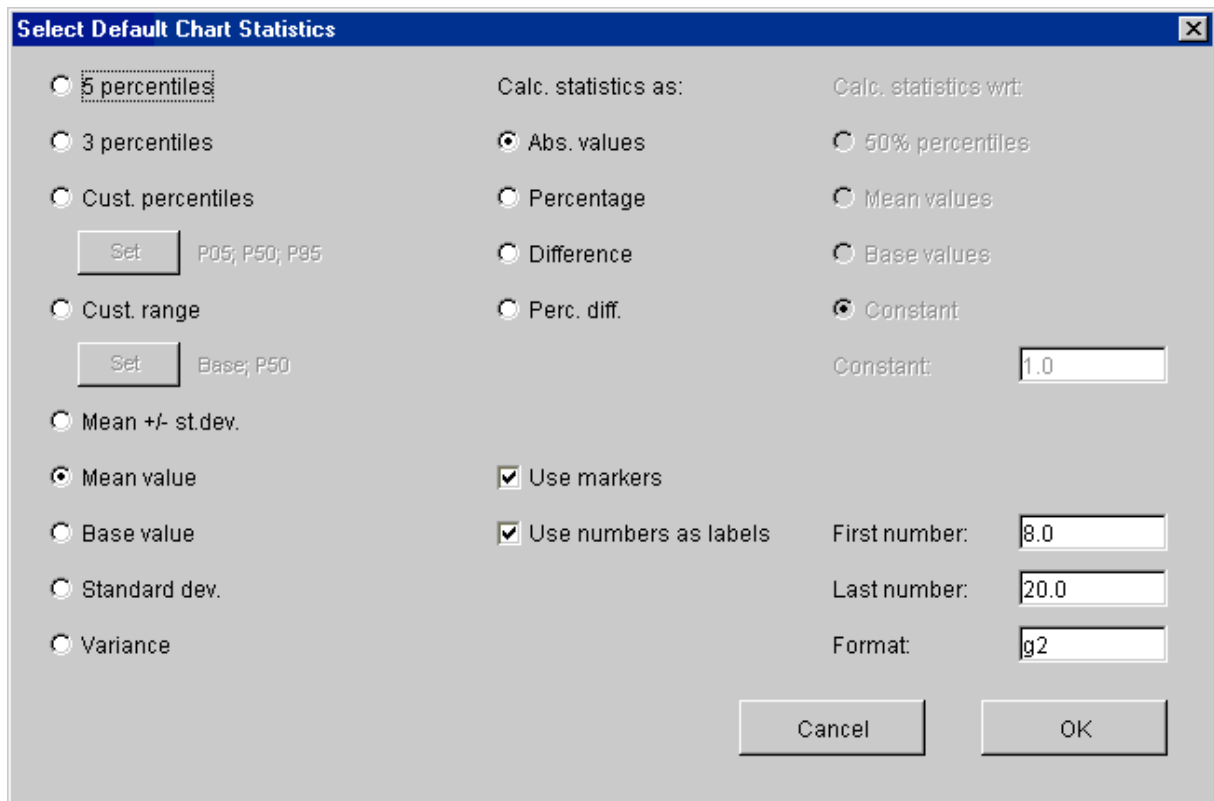
Kjør en simulering med dette oppsettet. Du skal da få opp et vindu med 10 resultatnoder, én resultatnode for hver korrelasjonsverdi, fra 0.1 opp til 1.0. Velg alle resultatnodene, og lag et felles S-kurveplot. På hvilken måte påvirker korrelasjonen fordelingen for **Netto resultat**? Kommentér dette resultatet.

(k) Vi skal til slutt prøve å finne ut hvor høy oljeprisen må være for at prosjektet skal ha positivt forventet netto resultat. For å undersøke dette krysser vi av for "Multivalued" for

oljeprisparameteren, dvs. parameteren med ID = 1. [Pass på at ingen andre parametere er krysset av som "Multivalued".]

Kjør en simulering med dette oppsettet. Du skal da få opp et vindu med 7 resultatnoder, én resultatnode for hver oljeprisverdi, dvs. for verdiene 8, 10, 12, ... , 20 (dollar pr. fat).

Benytt så meny-kommandoen "Statistics>Set Chart Statistics...". Du skal da få opp en dialogboks der du kan gjøre en del valg. Sørg for at oppsettet blir nøyaktig som vist i Figur 8.



Figur 8. Dialogboks for "Chart Statistics"

Velg så alle resultat nodene, og lag et linjeplot. Hvis du har gjort alt riktig, skal dette plottet vise forventet netto resultat som funksjon av oljeprisen. Hvor stor oljeprisen må være for at prosjektet skal ha positivt forventet netto resultat, kan nå leses av dette plottet som skjæringspunktet mellom kurven og null-linjen (grå stiptet vannrett linje). Kommenter resultatet.

SLUTT