Detta är en hjälptext med ursprung ur FRAM-KLIV, ett projekt i forskningsprogrammet Kraft och liv i vatten. FRAM-KLIV utfördes av Anthesis Enveco AB tillsammans med forskare från Handelshögskolan i Stockholm, Luleå tekniska universitet och Sveriges lantbruksuniversitet i Umeå.

7-III. Analys av osäkerhet

Osäkerhet kan analyseras i Excelprogrammet på flera olika sätt. Användaren kan ange värden med hjälp av intervall och användaren kan göra en känslighetsanalys genom att studera hur resultaten ändras om indata ändras på något sätt. En mer avancerad möjlighet är att ange en statistisk sannolikhetsfördelning för de monetära värdena för en eller flera konsekvensposter. Om användaren väljer denna möjlighet kommer Excelprogrammet att använda sig av en simuleringsanalys. De här olika ansatserna till att analysera osäkerhet introduceras nedan.

Om projektet exempelvis gäller en investering är det naturligt att tänka sig att flera parametrar, inte minst framtida kostnader och elpriser, inte är kända med säkerhet vid beslutstillfället. Den enklaste varianten av en känslighetsanalys syftar till att försöka utröna om någon parameter/variabel är särskilt central för utfallet av kalkylen. För projekt med lång tidshorisont är exempelvis den valda diskonteringsräntan ofta viktig för utfallet.

Antag att en investering I_0 görs i tidpunkten 0 och att ett projekt genererar en enhet av en vara/tjänst i varje tidpunkt fram till projektets slut vid tidpunkten T. Intäkterna av projektet i varje tidpunkt t=1,2,3…T är p_t. Antag på samma sätt att kostnaderna för att producera en enhet är w_t i tidpunkterna t=1,2,3…T. Diskonteringsräntan antas konstant över projektets livslängd och betecknas r. Alla ekonomiska storheter är angiva i reala termer. Nettonuvärdet av investeringen är:

$$NNV = -I_0 + \sum(pt - wt) * (1 + r)^{-t}$$

En känslighetsanalys innebär i det enklaste fallet att vi beräknar projektets nettonuvärde för ett basfall $NNV_0 = NNV(I_0, pt, r, wt)$, där de ingående parametrarna ansätts baskalkylens värden, I_0, p_t, r, w_t (där det är underförstått att intäkter och kostnader kan variera över projektperioden, de skall alltså betraktas som vektorer). Detta kan enkelt göras i ett Excelark. För att studera hur känslig kalkylen är för valet av ränta görs samma beräkning för t.ex. ett ”lågränteval” och ett ”högränteval”; $NNV_{låg} = NNV(I_0, pt, r_{låg}, wt)$ samt $NNV_{hög} = NNV(I_0, pt, r_{hög}, wt)$. Denna analys görs ofta för en parameter i taget, så att p_t, varieras givet värdena på de övriga parametrarna osv.

Låt oss illustrera med några påhittade siffror.

Tabell 1. Exempel på en investeringskalkyl med indikerade värden för en känslighetsanalys

Investeringens ekonomiska parametrar	Baskalkyl	Känslighetsanalys
$-I_0$, grundinvestering	10	
T, projektlängd	10	
P_t intäkt i tidpunkt t=1,2,3…9,10	2	P(”låg”)=1, p(”hög”)=3
w_t kostnad i tidpunkt t=1,2,3…9,10	1	w(”låg”)=0,5, w(”hög”)=1,5
r konstant diskonteringsränta	3%	R(”låg”)=1 %, r(”hög”)=5 %

I detta ytterst förenklade fall, låter vi en parameter variera i taget. Följande resultattabell tolkas så att vi först beräknar projektets nettonuvärde med baskalkylens parametervärden, kallad ”bas” i tabellen. Vi låter sedan en parameter variera i taget och studerar hur projektutfallet ändras. Exempelvis betyder ”r=0,01” i tabellen att nuvärden beräknas med baskalkylens parametervärden, utom räntan som sänks till 0,01.

Tabell 2. Utfallet av en känslighetsanalys tillämpad på data i tabell 4.

	nettonuvärde
bas	-1,469
r=0,01	-0,528
r=0,05	-2,278
p=1	-10,0000000
p=3	7,060
w=0,5	2,795
w=1,5	-5,734

I baskalkylen ger investeringen ett negativt nettonuvärde (-1,469). Investeringen kan dock rekommenderas om intäkterna blir högre eller kostnaderna lägre (7,060 resp. 2,795), allt annat lika.

Det finns möjlighet att bedöma hur sannolika de olika utfallen är. Om utfallen som ger positivt nettonuvärde är osannolika talar det inte för att projektet är samhällsekonomiskt lönsamt. En ytterligare begränsning är att denna känslighetsanalys inte ger en fullständig bild av projektets potential om fler än en parameter är osäker, vilket för övrigt torde vara normalfallet för de flesta investeringsprojekt. En s.k. simuleringsanalys, eller mer allmänt en Monte-Carlo analys, är ett populärt sätt att ge en mer fullständig beskrivning av projektets möjligheter när flera parametrar är osäkra. Det mest naturliga sättet att införa osäkerhet i detta sammanhang är att, i termer av ovan exempel, ersätta de enskilda parametervärdena med sannolikhetsfördelningar. Detta kan göras på många olika sätt och med mer kraftfulla metoder än vad som presenteras här. Den grundläggande idén är att vi låter parametrarna variera simultant och få en bild av projektet när alla parameter tänks variera inom givna intervall. Inledningsvis förklaras några statistiska begrepp via ett stiliserat exempel.

Inledning till simuleringsanalysen

För att förklara några viktiga statistiska begrepp som används i simuleringarna, utgår vi ifrån ett exempel. Betrakta ett projekt där såväl intäkter som kostnader är osäkra. Vi bortser inledningsvis från diskontering. Låt oss anta att intäkterna från projektet är 1,2 eller 3 (motsvarande säg, ett ”dåligt”/”medel”/”bra” år), och att sannolikheten för varje utfall är lika stort, dvs 1/3. Låt oss vidare anta att kostnaderna har samma struktur, dvs de antas vara 1,2 eller 3 och att vi inte heller där vet mer än att de har samma sannolikhet 1/3. Kostnaderna och intäkterna är två **stokastiska variabler**. Vi antar vidare att intäkter och kostnader är **oberoende**, så att kunskap om vad intäkten verkligen blir, inte ger anledning till att revidera sannolikheten för att kostnaderna antar ett visst värde. Det är analogt med att kasta två tärningar i följd. Om den första tärningen som kastas ger utfallet ”6”, påverkar det utfallet inte sannolikheterna för att få ett visst utfall på den andra tärningen. Det är precis samma chans att kastet med tärning nummer 2 ger värdet ”6”, som att värdet råkar bli ”1”.

Vad vi precis har beskrivit är två **sannolikhetsfördelningar,** som beskriver sannolikheten för att de olika värdena skall realiseras. Vi kan kombinera intäkter och kostnader till vinst = intäkt – kostnad. Det ger en annan **stokastisk variabel**, med ett annat **utfallsrum**. Om intäkterna fixeras till 1 kan vinsten bli 0, -1, eller -2, om intäkterna sätts till 2 blir vinsten 1,0,-1 och om intäkterna alltid skulle vara 3 blir vinsten 2,1,0. Ett motsvarande resonemang kring en given kostnad och en osäker intäkt, ger oss att den stokastiska variabeln vinst kan anta värdena -2,-1,0,1,2, vilket är dess utfallsrum. Notera att dessa värden på vinst inte har identiska sannolikheter, därför att vinsten bara kan bli -2 på ett sätt, medan vinsten kan bli -1/+1 på fler än ett sätt. Det är endast ett sätt som vinsten kan bli 2, nämligen att intäkten blev 3 och kostnaden 1. Om man går igenom de olika utfallsmöjligheterna finner man att sannolikheterna för de olika utfallen, från lägsta till högsta, är (1/9, 2/9, 3/9, 2/9, 1/9) eftersom det finns

9 olika utfall (3*3) för de oberoende variablerna intäkt och kostnad (för varje värde på intäkten kan kostnaden anta 3 värden osv).

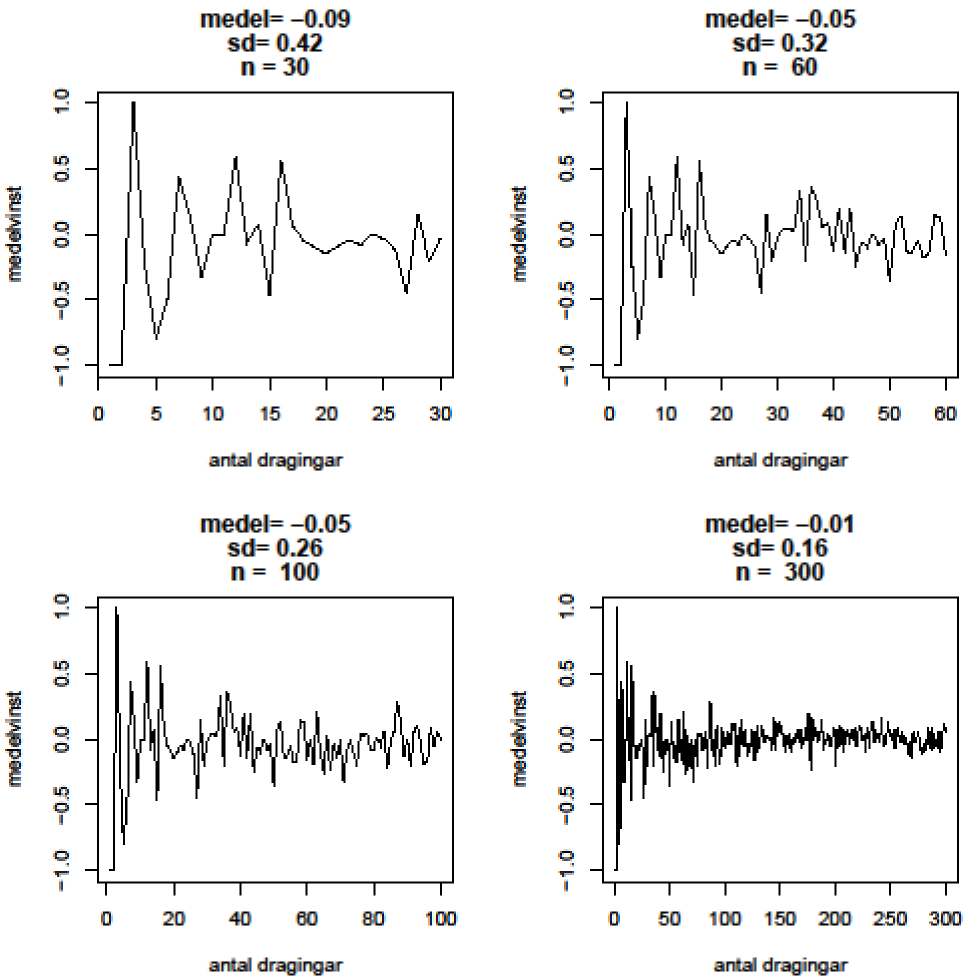
Vi kan nu räkna ut **medelvärdena** för de tre stokastiska variablerna. För intäkterna har vi 1/3*(1+2+3) = 2 och samma medelvärde får vi för kostnaderna. För vinsten har vi att summera (1/9, 2/9, 3/9, 2/9, 1/9)*(-2,-1,0,1,2) vilket blir = 0, vilket vi redan visste då intäkterna är i genomsnitt 2 och kostnaderna i genomsnitt 2.

En **Monte Carlo simulering** innebär i vårt fall att vi drar slumptal ur en given fördelning och räknar ut vinsten för varje dragning. Exempelvis kan vi dra 10 olika utfall för intäkterna ur fördelningen ({1,2,3},{1/3,1/3,1/3}). Vi måste då ”lägga tillbaka” efter varje dragning, eftersom vi bara har tre olika värden. Om vi drar 10 tal ur denna fördelning, kan de representera en simulerad utveckling för, säg, intäkterna 10 år framåt. Det är möjligt, men osannolikt, att vi drar intäkten = 3 för de 10 dragningarna. Den exakta sannolikheten för det utfallet kan enkelt beräknas (för det första året är sannolikheten 1/3, att vi skall få två stycken 3 i rad har sannolikheten 1/9, osv, så att sannolikheten blir $(1/3)^{10}$ eller 1 på 59 049).

Denna process kan tänkas representera utvecklingen av intäkterna 10 år framåt i tiden, där vi varje år slumpmässigt får en intäkt på 1, 2 eller 3 med samma sannolikhet. Om vi drar 100 slumptal kan det representera en utveckling 100 år framöver, eller i detta enkla fall, en utveckling över 10 år som upprepas 10 gånger. Vi kan göra denna tolkning eftersom slumptalen är oberoende av varandra. Ett ”bra” år följs inte nödvändigtvis av ett ”bra” år, tex. Givetvis kan vi modifiera vår modell så att vi bygger in beroenden och så vidare, men vi gör inte detta explicit här. Vi kan upprepa dragningen för kostnaderna och få, säg, 10 olika utfall. Skillnaden mellan intäkt och kostnad är vinsten och vi får då 10 olika utfall för vinsten. Alternativt kan vi direkt dra slumptal ur den konstruerade sannolikhetsfördelningen för vinst.

Om vi upprepar dragningarna tillräckligt många gånger kommer medelvinsten att konvergera mot det teoretiska medelvärdet 0. Det är en av huvudsatserna i statistiken och den används ju rutinmässigt av t.ex. försäkringsbolag.

I figur 1 illustreras utfallen för n=30, 60, 100 och 300 dragningar för slumpvariabeln vinst. För 30 dragningar blir medelvärdet av vinst, sett över just dessa dragningar, -0,09, med en **standardavvikelse** (sd) på 0,42; vi återkommer strax till begreppet standardavvikelse. För 60 dragningar blir det i just denna simulering ett medelvärde på -0,5 med en lägre standardavvikelse på 0,32. Vid 300 dragningar har medelvärdet nästan konvergerat mot det teoretiska medelvärdet på noll.

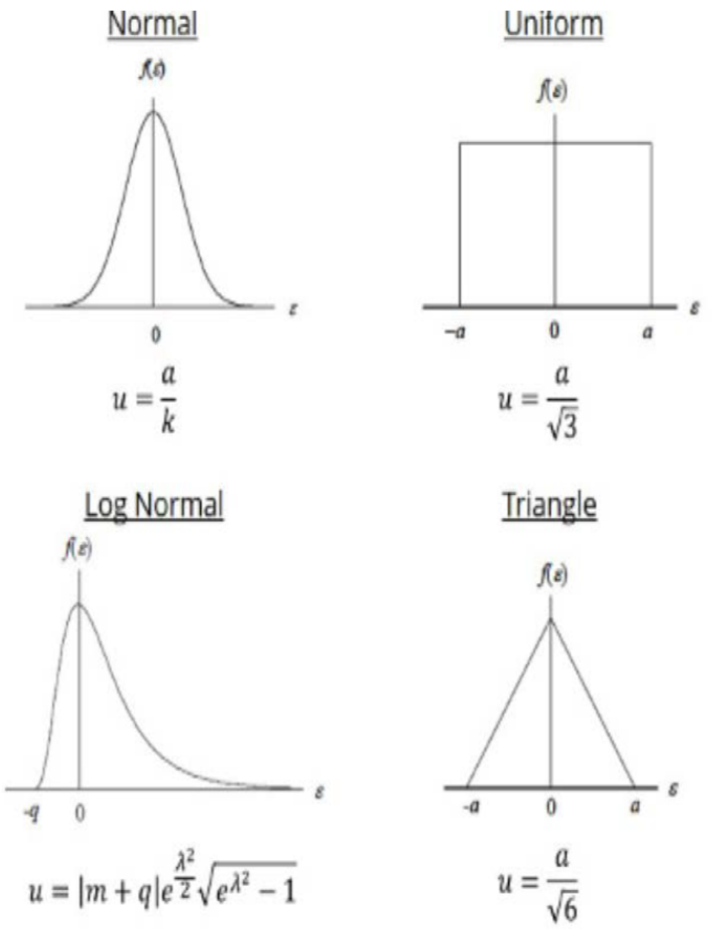


Figur 1. Illustration av hur upprepning av antalet dragningar gör att slumpvariabeln vinst konvergerar mot sitt medelvärde 0.

Ett annat sätt att se på hur upprepningar påverkar processen är att mäta standardavvikelsen, ett mått på spridningen runt medelvärdet. Ju fler observationer vi får, desto säkrare blir skattningen av medelvinsten; standardavvikelsen minskar. Vi ser detta tydligt i figuren ovan då standardavvikelsen sjunker från 0,42 till 0,16. Det är den beräknade standardavvikelsen, vars värde beror på de data som vi har genererat. I detta fall kommer standardavvikelsen att gå emot noll, dvs. medelvinsten kommer att konvergera mot noll om vi gör tillräckligt många upprepningar. Resonemanget är analogt med varför ett försäkringsbolag kan vara relativt säkra på utfallet för hela kollektivet försäkringstagare, däremot kan bolaget inte veta med säkerhet vilken individ som drabbas av en viss försäkringsbar händelse.

Val av fördelning

Låt oss nu tillämpa en simulering för vårt enkla exempel ovan. Vi skall då försöka ansätta olika sannolikhetsfördelningar för de osäkra parametrarna. Det finns en stor mängd fördelningar att välja mellan, t.ex. normalfördelning eller triangulär, se figur 2. På svenska brukar den triangulära fördelningen ibland kallas ”tältfördelningen” på grund av sitt utseende. Varje fördelning beskrivs av ett antal parametrar. För normalfördelningen (och den nära besläktade log-normalfördelningen) är det medelvärde och varians, för den uniforma (likformiga) behövs endast ett intervall [a,b]. Slutligen behöver man för den triangulära fördelningen ange ett intervall [a,b] samt ett värde där täthetsfunktionen har sitt maximum (”toppen på tältet”), eller det s.k. typvärdet.



Figur 2. Olika sannolikhetsfördelningar.

Vilken fördelning som skall väljas i varje enskilt fall går inte att svara på rent allmänt. En möjlighet är att inleda analysen med den enklaste fördelningen av alla, uniform fördelning, där varje utfall har samma sannolikhet inom ett givet intervall. Analysen förutsätter då att ett intervall kan ansättas för varje enskild parameter, samtidigt är det inte möjligt att göra någon bättre skattning än att varje utfall är lika sannolikt. I praktiken är det ofta möjligt att ha en uppfattning om hur sannolikt det är att en parameter, låt oss säga elpriset, ligger inom ett visst intervall samt sannolikheten för olika utfall. För elprisets del är det möjligt att använda observerade elpriser för att skatta fördelningen. Tältfördelningen är mycket användbar i detta sammanhang, inte minst därför att den liknar normalfördelningen (som utan tvekan är den mest använda för att beskriva data).

Känslighetsanalys och simulering: ett exempel

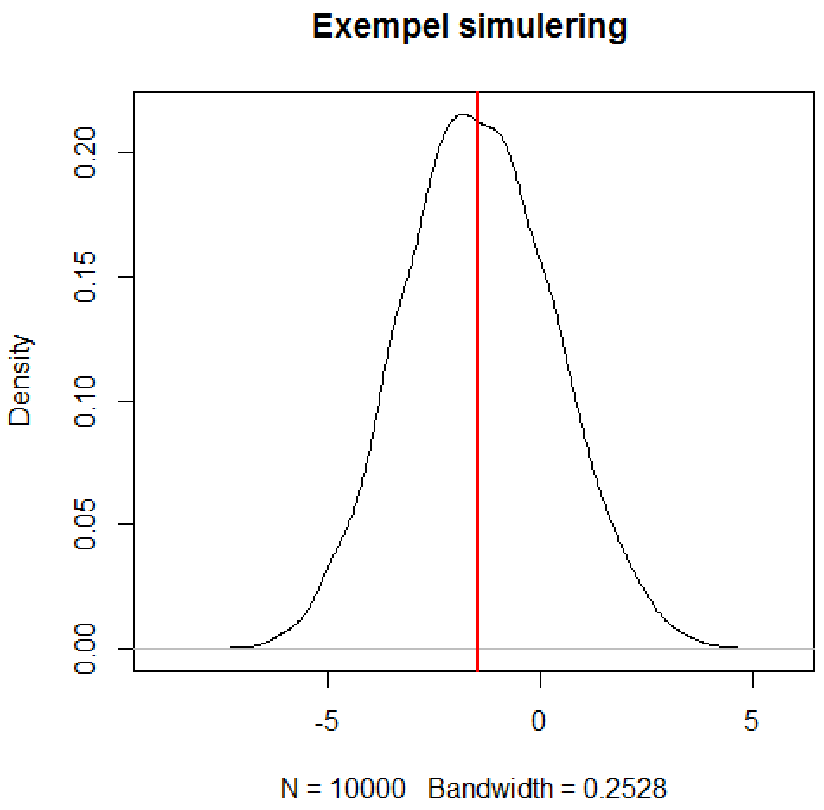
Låt oss återvända till vårt ursprungliga exempel, men nu utnyttja simulering för att beskriva projektet på ett mer intressant sätt. Vi skall använda den uniforma (likformiga) fördelningen, vilken kan sägas beskriva ett fall med mycket sparsam information; vi vet endast att variabeln ligger inom ett givet intervall, men har ingen annan uppfattning än att varje utfall inom det givna intervallet är lika sannolikt. Betrakta följande tabell:

Tabell 3. Data för simuleringarna.

Parameter	Basvärde	Antaget intervall	Antagen fördelning
$-I_0$, grundinvestering	10	(10,10) (ingen osäkerhet)	-
T, projektlängd	10	(10,10) (ingen osäkerhet)	--
P_t intäkt i tidpunkt t=1,2,3…9,10	2	[1,3]	Uniform (likformig)
w_t kostnad i tidpunkt t=1,2,3…9,10	1	[0.5, 1.5]	Uniform (likformig)
r konstant diskonteringsränta	3%	[1%, 5%]	Uniform (likformig)

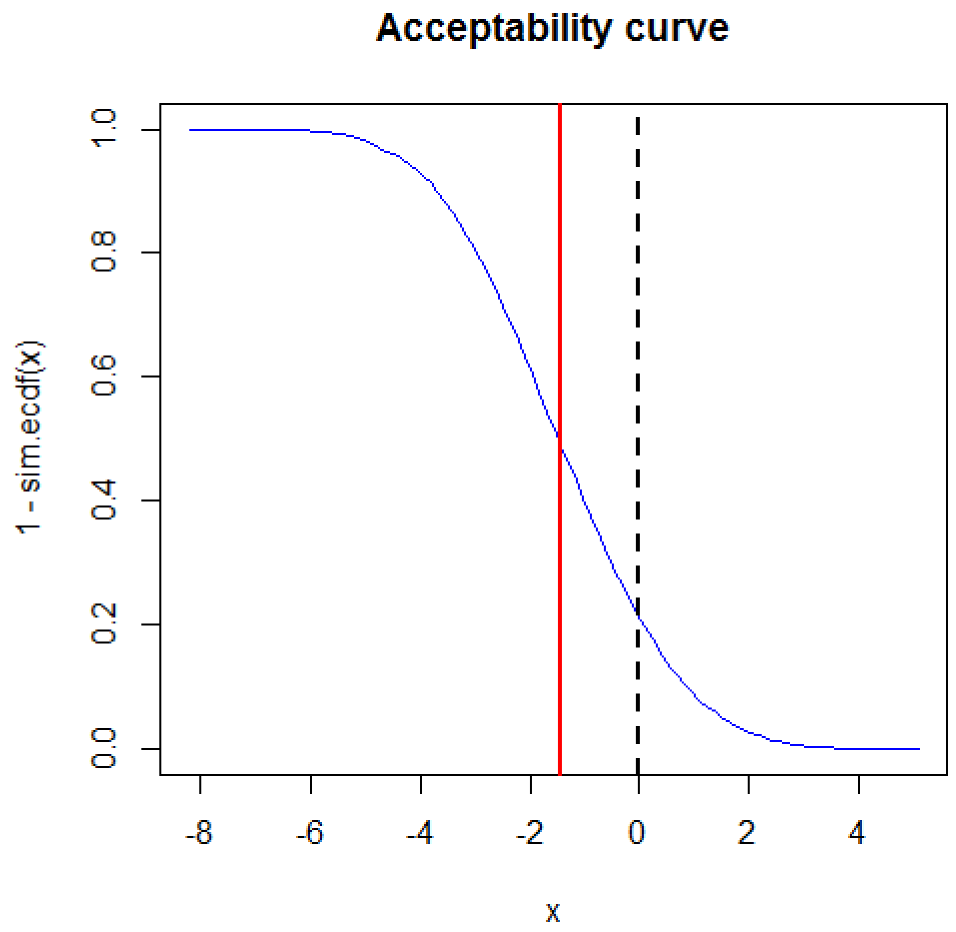
Notera att parametrarna i genomsnitt kommer att anta värdena från baskalkylen; om fördelningen är uniform och vi antar intervallet [1,3], kommer den variabeln att i genomsnitt vara 2, eftersom varje utfall i intervallet är lika sannolikt. Som vi redan sett, kan vi förvänta oss att vi med många upprepningar kommer att få ett utfall som liknar baskalkylen. Notera dock skillnaden att vi i simuleringen får olika värde varje period, emedan vi i känslighetsanalysen ansatte ett värde som fick vara konstant över hela perioden. En alternativ simulering vore att dra ett värde som antas vara konstant över hela perioden. Det kan vara ett rimligt tillvägagångssätt för t.ex. räntan som måhända inte varierar särskilt mycket på årsbasis i dessa sammanhang.

En tillämpning av ovanstående resonemang på data i tabellen gav medelvärdet -1,449, sålunda ganska nära baskalkylens resultat, precis som förväntat. Simuleringarna ger data som kan användas för att beskriva projektet med fler karakteristika än genomsnittligt utfall. Vi kan exempelvis uppskatta täthetsfördelningen:



Figur 3. Uppskattad frekvensfördelning från simulerade data, N=10 000. Medelvärdet indikerat med röd linje.

Som förväntat är fördelningen symmetrisk runt medelvärdet, det behöver den inte vara i mer komplicerade och realistiska fall. Ett alternativt sätt att beskriva samma data är att räkna ut den empiriska fördelningsfunktionen (hur stor andel av data har ett mindre värde än x, där x varieras över ett intervall). Med hjälp av den kan man direkt avläsa hur stor sannolikheten är att projektet går med vinst. Det är en smaksak om man vill beskriva data med den empiriska fördelningsfunktionen eller den s.k. överlevnadsfunktionen, som helt enkelt är 1 minus fördelningsfunktionen. Johansson och Kriström (2012, 2016) använder den senare varianten.



Figur 4. Empirisk överlevnadsfunktion för simulerade data. Medelvärde och nollvinst indikerad med röd resp. streckad linje. Sannolikheten att projektet skall gå med vinst är ungefär 20%.

Vi ser ur figuren att sannolikheten att projektet skall vara lönsamt är ungefär 20 %. Denna typ av information kan vara värdefull för att förstå projektets potential. En lämplig alternativfördelning kan vara log-normal, den är endast definierad över de icke-negativa talen. Den triangulära fördelningen ger många möjligheter, då användaren själv anger maximum och minimum, samt skevhet.

I Excelprogrammet har en användare som vill använda sig av simuleringar möjlighet att välja mellan följande fyra olika sannolikhetsfördelningar: Uniform (likformig) fördelning, symmetrisk triangulär fördelning, asymmetrisk triangulär fördelning och lognormal fördelning.

Referenser

Johansson, P.-O., Kriström, B., 2012. The Economics of Evaluating Water Projects – Hydroelectricity Versus Other Uses. Springer, Heidelberg, Tyskland.

Johansson, P.-O., Kriström, B., 2016. Cost–Benefit Analysis for Project Appraisal. Cambridge University Press, Cambridge, UK.

Appendix. Programkod för simuleringarna

Vi använder det fritt tillgängliga programmet R (www.r-project.org) för simuleringsexemplet.

```
T=10  # projektlängd

p=vector()

w=vector()

f=function(T=10, I0=10,r=0.03,p=rep(2,10),w=rep(1,10)) {

r=runif(10,0.01,0.05) #uniform

p=runif(10,1,3)      #uniform

w=runif(10,0.5,1.5)   #uniform

beta=(1+r)^(1:T)

print(-I0+ sum((p-w)/beta))

}

# simulering

seed=1  # frö för sannolikhetsfunktion

sim=vector()

for (i in 1:10000) sim[i]=f()

summary(sim)


# figur 3

plot(density(sim),main="Exempel simulering ")

abline(v=-1.45,col="red",lwd=2)

# figur 4

sim.ecdf=ecdf(sim)

r <- range(min(sim),max(sim))

curve(1-sim.ecdf(x), from=r[1], to=r[2], col="blue", xlim=r,lwd=1.5,main="Acceptability curve")

abline(v=-1.45,col="red",lwd=2)

abline(v=0,col="black",lwd=2,lty=2)
```