JAN ODELSTAD

Förväntad nytta i ljuset av alternativa numeriska representationer

1. Inledning

Principen om förväntad nytta spelar en viktig roll i beslutsteorin. Enligt denna princip är nyttan hos ett handlingsalternativ den förväntade nyttan hos handlingens utfall. Om $\{c_i \mid i \in I\}$ är de möjliga konsekvenserna av en handling a och för alla i, $i \in I$, c_i inträffar med sannolikheten p_i, så gäller enligt förväntade-nyttoprincipen att nyttan hos a är

$$\sum_{i \in I} p_i \cdot u(c_i)$$

där $u(c_i)$ är värdet eller nyttan hos utfallet c_i. Förväntade-nyttoprincipen kombineras ofta med tanken att man ska välja den handling som har den största [förväntade] nyttan.

Det tycks vara en vanlig uppfattning att principen om förväntad nytta kan ges ett formellt stöd. von Neumanns och Morgensterns klassiska teori om sannolikhetsblandningar från 1947 anses t ex ofta ge en axiomatisk grundval för denna princip. Per-Erik Malmnäs har dock nyligen (se Malmnäs (1991) och (1994)) argumenterat för att varken von Neumanns och Morgensterns eller andra existerande axiomatiska karakteriseringar av nyttobegreppet ger en formell grund för förväntade-nyttoprincipen. Jag ska i denna not diskutera några aspekter på frågan om den axiomatiska grundvalen för förväntade-nyttoprincipen. Som en bakgrund för denna diskussion ges i nästa sektion en kortfattad skiss av en variant av von Neumanns och Morgensterns teori om sannolikhetsblandningar.

2. Blandningsrum

$\langle A, \geq, \theta \rangle$ är ett *blandningsrum* (mixture space, se Roberts (1979) s 353) om $\geq$ är en binär relation på A och θ är en funktion sådan att

Filosofisk tidskrift 1995 nr 4, 160-168

$$\theta : A \times [0,1] \times A \rightarrow A$$

där [0,1] är det slutna intervallet mellan 0 och 1. Som brukligt är använder vi "apb" för att beteckna "θ(a,p,b)". Vi kallar apb för en [sannolikhets]blandning (med utfallen a och b). Låt

$$\theta_p(a,b) = \theta(a,p,b) = apb.$$

Observera att θ_p är en binär operation på A. Vi kan alltså se θ som en familj av binära operationer på A.

Den intenderade tolkningen av en [sannolikhets]blandning apb är ett lotteri (spel) som med sannolikheten p ger a som utfall och med sannolikheten 1-p ger b som utfall. Observera att utfallen a och b själva kan vara lotterier (blandningar). Enligt den intenderade tolkningen förstås $\geq$ som en preferensrelation, dvs. a $\geq$b innebär att a är åtminstone lika bra som b.

Låt oss säga att $\langle A, \geq, \theta\rangle$ är ett N-M-blandningsrum (uppkallat efter von Neumann och Morgenstern) om $\langle A, \geq, \theta\rangle$ är ett blandningsrum och följande axiom är uppfyllda för alla a,b,c i A (se Roberts (1979) s 354).

(A1) $\geq$ är en svag ordning på A.

$>$ är motsvarande strikta svaga ordning och $=$ är motsvarande indifferensrelation.

(A2) (apb) = [b(1-p)a] för alla p sådana att $0 \leq p \leq 1$.

(A3) [(apb)qb] = a(pq)b för alla p,q sådana att $0 \leq p,q \leq 1$.

(A4) Om a $>$ b, så (apc) $>$ (bpc) för alla p sådana att $0<p<1$.

(A5) Om a $>$ b $>$ c, så existerar p,q sådana att $0<p,q<1$ och (apc) $>$ b och b $>$ (aqc).

I von Neumann och Morgenstern (1947) visades att följande representations- och entydighetsteorem gäller för N-M-blandningsrum (se Roberts (1979) s 356).

Representationsteoremet: Om $\langle A, \geq, \theta\rangle$ är ett N-M-blandningsrum så finns en reellvärd funktion u definierad på A sådan att

$$a \geq b \text{ om och endast om } u(a) \geq u(b) \quad (1)$$

$$u(apb) = pu(a)+(1-p)u(b) \quad (2)$$

Entydighetsteoremet: Om u och v är funktioner som uppfyller (1) och (2) ovan så finns det reella tal r och s där r är större än 0, sådana att

$v(a) = ru(a)+s$ för alla a i A. (u och v är således intervallskalor.)

Den intenderade tolkningen av funktionen u, vilken uppfyller villkoren (1) och (2) ovan, är att u är ett nyttomått för blandningsrummet $\langle A, \geq,\theta\rangle$.

Notera att (2) är en formulering av den förväntade nyttoprincipen; enligt (2) är nyttan hos blandningen apb den förväntade nyttan av utfallen a och b.

Vi ska nu se närmare på innebörden i representationsteoremet. Låt för den skull ω vara den funktion sådan att

$$\omega : Re \times [0,1] \times Re \rightarrow Re$$

(där [0,1] är det slutna intervallet mellan 0 och 1) *och* det gäller att $\omega(x,p,y) = px+(1-p)y$. Låt $\omega_p(x,y) = \omega(x,p,y) = px+(1-p)y$. ω_p är en binär operation på Re. Vi kan alltså se ω som en familj av binära operationer på Re. Det kan vara bekvämt att använda $+_p$ som beteckning för ω_p och vi skriver $\omega_p(x,y) = x+_py$. Således gäller att

$$x+_py = px+(1-p)y.$$

Representationsteoremet innebär alltså att ett N-M-blandningsrum $\langle A, \geq,\theta\rangle$ kan homomorft representeras med $\langle Re,\geq,\omega\rangle$ i följande mening: Det finns $u : Re \rightarrow Re$ sådan att (1) ovan gäller samt

$$u(\theta(a,p,b)) = \omega(u(a),p,u(b)) \quad (3)$$

Notera att (3) är ekvivalent med att

$$u(\theta_p(a,b)) = \omega_p(u(a),u(b)) \quad (4)$$

(3) är ju ett sätt att formulera den förväntade nyttoprincipen, vilken alltså utsäger att θ representeras med ω.

3. Alternativa numeriska representationer av blandningsrum

Det är lätt att inse att ett N-M-blandningsrum $\langle A, \geq,\theta\rangle$ kan homomorft representeras med andra numeriska strukturer än $\langle Re,\geq,\omega\rangle$. I så fall kommer θ att representeras med en annan familj av binära operationer än ω, och således kommer den förväntade nyttoprincipen att ersättas med en annan princip. Låt oss se mer i detalj på detta. Först en definition.

Ett N-M-blandningsrum $\langle A, \geq,\theta\rangle$ kallas *icke-negativt* om det finns en homomorfi u från $\langle A, \geq, \theta \rangle$ till $\langle Re, \geq, \omega \rangle$ sådan att u(a) är ett icke-negativt reellt tal för alla $a \in A$. $\langle A, \geq, \theta \rangle$ är alltså icke-negativt om det finns en homomorfi från $\langle A, \geq, \theta \rangle$ till $\langle Re^*, \geq, \omega \rangle$ där $Re^* = Re^+ \cup \{0\}$, dvs. Re* är mängden av icke-negativa reella tal.

Vi ska nu se på en annan numerisk representation av ett icke-negativt blandningsrum $\langle A, \geq, \theta \rangle$ än $\langle Re, \geq, \omega \rangle$. Låt μ vara den funktion sådan att

$$\mu : Re^* \times [0,1] \times Re^* \rightarrow Re^*$$

och $\mu(x,p,y) = \sqrt{(px^2+(1-p)y^2)}$. Vi kan alltså se μ som en familj av binära operationer på Re*. Låt $\mu_p(x,y) = \mu(x,p,y) = \sqrt{(px^2+(1-p)y^2)}$. Det kan vara bekvämt att använda $\otimes_p$ som beteckning för μ_p och vi skriver $\mu_p(x,y) = x \otimes_p y$. Således gäller att

$$x \otimes_p y = \sqrt{(px^2+(1-p)y^2)}.$$

Notera att om $f : Re^* \rightarrow Re^*$ sådan att $f(x)=\sqrt{x}$ så är f en isomorfi från $\langle Re^*, \geq, \omega \rangle$ till $\langle Re^*, \geq, \mu \rangle$. Ty f är omvändbar samt

$$x \geq y \text{ omm } f(x) \geq f(y)$$

och

$$\begin{aligned} f(\omega(xpy)) &= \sqrt{(px+(1-p)y)} = \sqrt{(p(\sqrt{x})^2+(1-p)(\sqrt{y})^2)} = \\ &= (\sqrt{x}) \otimes_p (\sqrt{y}) = \mu(f(x),p,f(y)) \end{aligned}$$

dvs för alla p, $1 \leq p \leq 1$,

$$f(x +_p y) = f(x) \otimes_p f(y).$$

Av detta följer att om $\langle A, \geq, \theta \rangle$ kan homomorft representeras med $\langle Re^*, \geq, \omega \rangle$ så kan $\langle A, \geq, \theta \rangle$ också homomorft representeras med $\langle Re^*, \geq, \mu \rangle$. Ty om u är en homomorfi från $\langle A, \geq, \theta \rangle$ till $\langle Re^*, \geq, \omega \rangle$, så är $f^\circ u$ (kompositionen av u och f) en homomorfi från $\langle A, \geq, \theta \rangle$ till $\langle Re^*, \geq, \omega \rangle$. Detta kan inses på följande sätt:

$$a \geq b \text{ omm } u(a) \geq u(b) \text{ omm } \sqrt{u(a)} \geq \sqrt{u(b)} \text{ omm } (f^\circ u)(a) \geq (f^\circ u)(b).$$

$$\begin{aligned} (f^\circ u)(apb) &= f(u(a) +_p u(b)) = (f(u(a)) \otimes_p (f(u(b)) \\ &= \mu((f^\circ u)(a),p,(f^\circ u)(b)). \end{aligned}$$

θ representeras alltså då med μ, vilket innebär att den princip som nu svarar mot principen om förväntad nytta är följande:

$$h(\theta(a,p,b)) = \mu(h(a),p,h(b)) = \sqrt{(ph(a)^2+(1-p)h(b)^2)}$$

där h är en homomorfi från $\langle A, \geq,\theta\rangle$ till $\langle Re^*,\geq,\mu\rangle$. Principen säger alltså att nyttan hos en blandning (ett lotteri) är kvadratroten ur det förväntade värdet av nyttan i kvadrat hos blandningens (lotteriets) utfall.

Av det som sagts följer att om $\langle A, \geq,\theta\rangle$ är ett icke-negativt N-M-blandningsrum, så kan $\langle A, \geq,\theta\rangle$ homomorft representeras med $\langle Re^*, \geq,\omega\rangle$. Vidare gäller att om h och h′ är homomorfier från $\langle A, \geq,\theta\rangle$ till $\langle Re^*,\geq,\omega\rangle$, så finns reella tal r och s sådana att

$$h' = rh\otimes s = \sqrt{(rh)^2+s^2}).$$

En annan numerisk struktur som kan användas för att representera N-M-blandningsrum och som vi ska återkomma till är $\langle Re^*,\geq,\nu\rangle$, där

$$\nu : Re^* \times [0,1] \times Re^* \rightarrow Re^*$$

och

$$\nu(x,p,y) = p^2x+(1-p)^2y+2p(1-p)\sqrt{(xy)}.$$

Vi kan, i analogi med vad vi gjort tidigare, se ν som en familj av binära operationer på Re*. Det är lätt att förvissa sig om att g: $Re^* \rightarrow Re^*$ sådan att $g(x)=x^2$ är en isomorfi från $\langle Re^*,\geq,\omega\rangle$ till $\langle Re^*,\geq,\nu\rangle$. Det innebär att om u är en homomorfi från $\langle A,\geq,\theta\rangle$ till $\langle Re^*,\geq,\omega\rangle$ så är u^2 en homomorfi från $\langle A,\geq,\theta\rangle$ till $\langle Re^*,\geq,\nu\rangle$. Notera att

$$\nu(x,p,y) = (p\sqrt{x}+(1-p)\sqrt{y})^2.$$

Vi kommer att använda oss av detta längre fram.

4. Förväntad nytta som ett specialfall av nyttoreduktion

Med ett *nyttomått* för ett blandningsrum $\langle A, \geq,\theta\rangle$ avser vi här en homomorfi h från $\langle A, \geq,\theta\rangle$ till en struktur $\langle B,\geq,\chi\rangle$, där $B\subseteq Re$ och χ en familj av binära operationer på B. Således gäller för h att

$$h(\theta(a,p,b)) = \chi(h(a),p,h(b)).$$

Ett annat sätt att uttrycka detta är att säga att nyttan hos en blandning är en funktion av utfallens nytta och sannolikhet. Detta är alltså en generell princip gällande nyttomått för blandningsrum. Man skulle kanske kunna kalla denna för en generell *nyttoreduktionsprincip;* nyttan hos blandningen reduceras till en funktion av nyttan och sannolikheten hos utfallen. I det specialfall då nyttomåttet representerar blandningsrummet med $\langle R,\geq,\omega\rangle$ där $R\subseteq Re$, så får nyttoreduktionsprincipen den mer specifika formen av principen om förväntad nytta. Således gäller att principen om förväntad nytta är den form nyttoreduktionsprincipen antar då blandningsrummet representeras på det vanliga sättet.

Låt mig försöka uttrycka denna tanke något mer precist. För den skulle behövs några definitioner.

Antag att $\langle A,\geq,\theta\rangle$ är ett blandningsrum och att $f:A \rightarrow Re$ är en homomorfi från $\langle A, \geq\rangle$ till $\langle Re,\geq\rangle$. Vi säger då att:

(1) f uppfyller *förväntade-nyttovillkoret* om för alla $a,b \in A$ och $p \in [0,1]$

$$f(\theta(a,p,b)) = \omega(f(a),p,f(b)).$$

(2) f uppfyller *nyttoreduktionsvillkoret* om det finns en funktion

$$\chi : f[A] \times [0,1] \times f[A] \rightarrow Re$$

sådan att för alla $a,b \in A$ och $p \in [0,1]$

$$f(\theta(a,p,b)) = \chi(f(a),p,f(b)).$$

Vi kan nu uttrycka tankegången ovan på följande sätt: Varje nyttomått för ett N-M-blandningsrum uppfyller nyttoreduktionsvillkoret men inte förväntade-nyttovillkoret. Däremot uppfyller varje nyttomått som representerar ett N-M-blandningsrum med $\langle R,\geq,\omega\rangle$ där $R \subseteq Re$ förväntade-nyttovillkoret.

von Neumanns och Morgensterns axiomatisering av nyttoteorin görs i s k kvalitativa termer, dvs i termer av relationer och operationer på blandningarna. Axiomatiseringen görs således inte i termer av nyttomåttet, vilket i stället först kommer in i representationsteoremet för teorin. Det ligger i karaktären hos denna typ av axiomatisering att olika numeriska strukturen kan användas för representationen av blandningarna. Eftersom förväntad-nyttooperationen sammanhänger med en viss typ av representationen kommer alla nyttomått inte att uppfylla förväntade-nyttovillkoret. Detta är ingen defekt hos von Neumanns och Morgensterns teori utan en följd av den kvalitativa axiomatiseringens natur.

5. Axiomatisk grundval för förväntade-nyttoprincipen

Tanken bakom strävan att ge en axiomatisk grundval för förväntade-nyttoprincipen tycks vara följande. Ett axiomsystem för nyttoteorin utgör ett formellt stöd för förväntade-nyttoprincipen om axiomen garanterar nyttomått som uppfyller förväntade-nyttovillkoret. Det är viktigt här att skilja på en stark och en svag variant av denna tanke. Låt oss säga att ett axiomsystem ger ett *starkt formellt stöd* för nyttoprincipen om varje nyttomått för ett godtyckligt blandningsrum som satisfierar axiomen uppfyller förväntade-nyttovillkoret. Däremot ger ett axiomsystem ett *svagt formellt stöd* för nyttoprincipen om det för varje blandningsrum som satisfierar axiomen finns ett nyttomått som uppfyller förväntade-nyttovillkoret.

Det som tycks intressera Malmnäs är om axiomatisering kan ge ett starkt formellt stöd för förväntade-nyttoprincipen. Enligt Malmnäs gäller nämligen att en agent uppfyller förväntade-nyttoprincipen om hans eller hennes nyttomått uppfyller förväntade-nyttovillkoret. (Malmnäs (1991) sid 21.) Vidare tycks han tänka sig att en axiomatisering av nyttoteorin ger ett formellt stöd för förväntade-nyttoprincipen om följande gäller: Om en agents preferenser uppfyller axiomen så uppfyller agenten förväntade-nyttoprincipen.

En av slutsatserna i Malmnäs (1991) och (1994) är att von Neumanns och Morgensterns axiomatisering lika lite som någon annan axiomatisering i kvalitativa termer kan ge ett formellt stöd för förväntade-nyttoprincipen. Av vad som tidigare har sagts i denna not framgår att jag helt instämmer i denna slutsats. Vi kan alltid välja att representera ett N-M-blandningsrum med en numerisk struktur sådan att homomorfierna från blandningsrummet till denna struktur inte uppfyller förväntade-nyttovillkoret.

Men av att von Neumanns och Morgensterns axiomatisering av nyttoteorin inte ger ett starkt stöd för förväntade-nyttoprincipen följer inte att den inte ger något stöd alls. Tvärt om ger von Neumanns och Morgensterns teori ett svagt formellt stöd i ovan angivna mening. Om blandningsrummet uppfyller villkoren för N-M-blandningsrum, kan vi representera det med en struktur, så att alla homomorfier från blandningsrummet till denna struktur uppfyller förväntade-nyttovillkoret.

Att en kvalitativ axiomatisering av nyttoteorin enbart kan ge ett svagt formellt stöd för förväntade-nyttoprincipen uppmärksammades tidigt av vissa beslutsteoretiker. Redan i Arrow (1951) påpekas att istället för att använda den nyttoskala som von Neumann och Morgenstern hade funnit, dvs ett nyttomått från ett N-M-blandningsrum till $\langle Re^{*}, \geq, \omega \rangle$, så kan man använda kvadraten på detta nyttomått. För, som Arrow framhåller, en individs beteende beskrivs då genom att säga att individen strävar att maximera det förväntade värdet av kvadratroten ur hans nytta. Arrow tänker sig alltså att representera blandningsrummet med $\langle Re^*,\geq,\nu\rangle$. Om u är en homomorfi från $\langle A,\geq,\theta\rangle$ till $\langle Re^*,\geq,\omega\rangle$ så är ju u^2 en homomorfi från $\langle A,\geq,\theta\rangle$ till $\langle Re^*,\geq,\nu\rangle$ (se sektion 3). Låt h vara en homomorfi från $\langle A,\geq,\theta\rangle$ till $\langle Re^*,\geq,\nu\rangle$. I sektion 3 konstaterades att

$$\nu(x,p,y) = (p\sqrt{x}+(1-p)\sqrt{y})^2.$$

vilket innebär att

$$h(a,p,b) = (p\sqrt{h(a)}+(1-p)\sqrt{h(b)})^2.$$

Att söka de a och b för vilka h(a,p,b) antar maximalt värde är alltså detsamma som att söka de a och b för vilka $p\sqrt{h(a)}+(1-p)\sqrt{h(b)}$ antar maximalt värde, och detta är precis vad Arrow noterar. Arrow understryker att detta påpekande inte ifrågasätter användbarheten av von Neumanns och Morgensterns resultat. Vad det säger är i stället att bland alla olika nyttomått kan man plocka ut en delmängd bestående av nyttomått vilka är linjärtransformationer av varandra och "which has the property of stating the laws of rational behavior in a particularly convenient way" (Arrow (1963) s 10). (I detta sammanhang är också s 98 i Savage (1972) av intresse. Se också Weymark (1991) s 264.)

Det svaga formella stöd som t ex von Neumanns och Morgensterns axiomatisering av nyttoteorin ger förväntade-nyttoprincipen är således av stor betydelse; det gör det möjligt att för N-M-blandningsrum hitta ett nyttomått som uppfyller förväntade-nyttovillkoret. Ett dylikt nyttomått är i många sammanhang bekvämt att använda.

Vid sidan av en axiomatisk grundval för förväntade-nyttoprincipen åberopas ofta stora talens lag som stöd för principen. Skulle kanske stora talens lag kunna ge det starka formella stöd för principen som axiomsystem inte kan ge? Jag förmodar att svaret är nej, ty stora talens lag tycks övergår i andra principer då den numeriska representationen ändras. Men det är en annan historia.

Not

Jag vill tacka Sven Danielsson, Lars Lindahl, Sten Lindström, Per-Erik Malmnäs och Wlodek Rabinowicz för värdefulla synpunkter och

intressanta diskussioner. Som den uppmärksamme läsaren redan har förstått används symbolerna ≥, > och = i denna uppsats med två olika innebörder, beroende på om de gäller sannolikhetsblandningar (spel) eller reella tal.

Litteratur

ARROW, K *Social choice and individual values.* 2:a uppl, New York: Wiley, 1963. (1:a uppl 1951.)

MALMNÄS, P-E "Axiomatiska argument för utilitetsprincipen. En formell undersökning". I Rabinowicz, W (red): *Valets vedermödor. Sex beslutsteoretiska studier.* Thales, 1991.

MALMNÄS, P-E "Axiomatic justifications of the utility principle: a formal investigation". *Synthese* 99: 233–249, 1994.

VON NEUMANN, J & MORGENSTERN, O *Theory of games and economic behavior.* 2:a uppl, Princeton: Princeton University Press, 1947.

ROBERTS, F *Measurement theory with applications to decisionmaking, utility and the social sciences.* Reading: Addison-Wesley, 1979.

SAVAGE, L *The foundations of statistics.* 2:a uppl, New York: Dover, 1972. (1:a uppl 1954).

WEYMARK, J "A reconsideration of the Harsanyi-Sen debate on utilitarianism". I Elster, J & Roemer, J (red): *Interpersonal comparisons of well-being.* Cambridge: Cambridge University Press, 1991.