Aalto University
School of Science

MS-E2135
Decision Analysis
Lecture 5

- *Elicitation of attribute-specific value functions*
- *Preferential and difference independence*
- *Aggregation of values with an additive value function*
- *Interpretation of attribute weights*

Salo, Liesiö, Punkka, Vilkkumaa

Last time

- Most decisions involve multiple several objectives

- Given the axioms of Lecture 4, the DM's preferences for a single attribute can be represented by a cardinal value function $v_i(x_i)$ such that

$$v_i(x_i) \geq v_i(y_i) \Leftrightarrow x_i \succcurlyeq y_i$$
$$v_i(x_i) - v_i(x_i') \geq v_i(y_i) - v_i(y_i') \Leftrightarrow (x_i \leftarrow x_i') \succcurlyeq_d (y_i \leftarrow y_i').$$

- **<u>Next</u>**: How should one elicit the cardinal value function $v_i(x_i)$?

- **<u>Later today</u>**: Can these attribute-specific value functions be combined into an overall value function? $V(x_1, x_2, \dots x_n) = f\big(v(x_1), \dots, v(x_n)\big) = \sum_{i=1}^{n} w_i v_i^N(x_i)$?

Elicitation of value functions

- **Phases:**
  1. Define a wide enough measurement scale $X_i = [a_i^0, a_i^*]$
     (or $\left[a_i^*, a_i^0\right]$ if smaller achievements levels are preferred to higher ones)
  2. Ask a series of elicitation questions
  3. Check that the value function gives realistic results

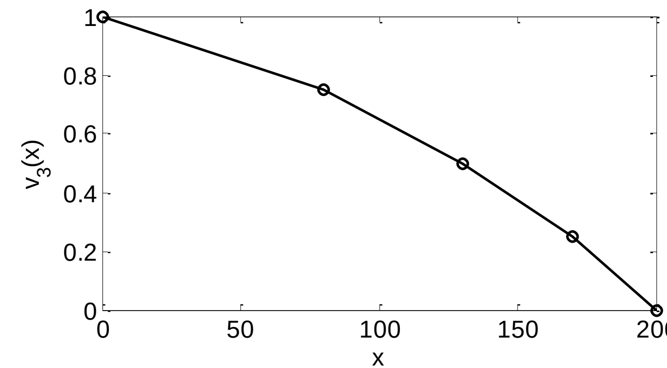
Elicitation of value functions: Indifference methods

❑ Bisection method:

- Ask the DM to assess level $x_{0.5} \in [a_i^0, a_i^*]$ such that she is indifferent between the changes $x_{0.5} \leftarrow a^0$ and $a^* \leftarrow x_{0.5}$.
- Then, ask her to assess levels $x_{0.25}$ and $x_{0.75}$ such that she is indifferent between
  - *changes* $x_{0.25} \leftarrow a^0$ *and* $x_{0.5} \leftarrow x_{0.25}$, *and*
  - *changes* $x_{0.75} \leftarrow x_{0.5}$ *and* $a^* \leftarrow x_{0.75}$.
- Continue until sufficiently many points have been obtained
  - *Use, e.g, linear interpolation between elicited points if needed*
- The value function can be obtained by fixing $v_i(a_i^0)$ and $v_i(a_i^*)$ at 0 and 1, respectively

Elicitation of value functions: Indifference methods

- Example of the bisection method
  - Attribute a_3 : Traveling days per year
  - Measurement scale $[a_3^*, a_3^0]$, where $a_3^* = 0$ and $a_3^0 = 200$; fix $v_3(a_3^0)$ =0 and $v_3(a_3^*)$ =1
    - *"What would be the number $x_{0.5}$ of traveling days such that you would be indifferent between a decrease from 200 to $x_{0.5}$ days a year and a decrease from $x_{0.5}$ to zero days a year?" (**Answer e.g., "130"**)*
    - *"What would be the number $x_{0.25}$ of traveling days such that you would be indifferent between a decrease from 200 to $x_{0.25}$ days a year and a decrease from $x_{0.25}$ to 130 days a year?" (**Answer e.g., "170"**)*
    - *"What would be the number $x_{0.75}$ of traveling days such that you would be indifferent between a decrease from 130 to $x_{0.75}$ days a year and a decrease from $x_{0.75}$ to zero days a year?" (**Answer e.g., "80"**)*



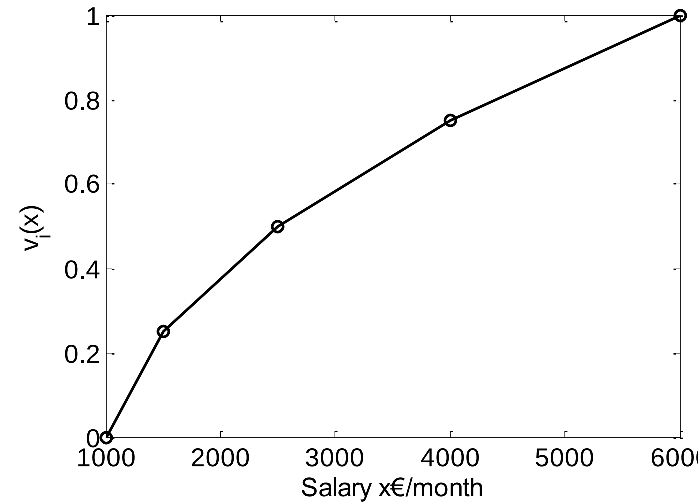
$$v_3(130) - v_3(200) = v_3(0) - v_3(130) \Rightarrow$$
$$v_3(130) = \frac{v_3(0) + v_3(200)}{2} = 0.5$$

$$v_3(170) - v_3(200) = v_3(130) - v_3(170) \Rightarrow$$
$$v_3(170) = \frac{v_3(130) + v_3(200)}{2} = 0.25$$

$$v_3(80) - v_3(130) = v_3(0) - v_3(80) \Rightarrow$$
$$v_3(80) = \frac{v_3(0) + v_3(130)}{2} = 0.75$$

Elicitation of value functions: Indifference methods

- **Sequence of equally preferred differences:**
  - Set $x_0 \in (a_i^0, a_i^*)$
  - Ask the DM to assess level $x_1 \in (x_0, a_i^*]$ such that he is indifferent between changes $x_0 \leftarrow a_i^0$ and $x_1 \leftarrow x_0$
    - $v_i(x_0) - v_i(a_i^0) = v_i(x_1) - v_i(x_0) \Rightarrow v_i(x_1) = 2v_i(x_0)$
  - Then, ask him to assess level $x_2 \in (x_1, a_i^*]$ such that he is indifferent between change $x_1 \leftarrow x_0$ and $x_2 \leftarrow x_1$
    - $v_i(x_1) - v_i(x_0) = v_i(x_2) - v_i(x_1) \Rightarrow v_i(x_2) = 3v_i(x_0)$
  - Continue until $x_N = a_i^*$ and solve the system of linear equations
    - $v_i(x_0) = \frac{v_i(x_N)}{N+1} = \frac{1}{N+1} \Rightarrow v_i(x_1) = \frac{2}{N+1}$ *etc.*
  - If $x_N > a_i^*$ (see exercises!)
    - *Change* a_i^* *to* x_N *and interpolate, or*
    - *Interpolate to get* $v_i(a_i^*) - v_i(a_i^0)$



Example:

$$[a_i^0, a_i^*] = [1000, 6000],\ x_0 = 1500$$
$$x_1 = 2500, x_2 = 4000, x_3 = 6000 = a_i^* \Rightarrow$$
$$v_i(1500) = \frac{1}{4},\ v_i(2500) = \frac{1}{2},\ v_i(4000) = \frac{3}{4}.$$

Elicitation of value functions: Indifference methods

- Indifference methods help specify a cardinal value function that captures the DM's preferences
- These methods are based on explicit statements about preferences ➔ They are theoretically sound

- Yet: indifference methods cannot be used if the measurement scale is discrete
  - E.g., Fit with interest: $X_4 = \{poor, fair, good, excellent\}$
  - Cf. Axiom A6

Elicitation of value functions: direct methods

- Direct rating
  - Ask the DM to directly attach a value to each attribute level
  - E.g. "Assume that the value of *poor fit* with interests is 0 and the value of *excellent fit* with interests is 1. What is the value of *fair fit* with interests? How about *good fit*?"
- Class rating
  - Divide the measurement scale into classes and ask the DM to attach a value to each class
- Ratio evaluation
  - Take one attribute level as a reference point and ask the DM to compare the other levels to this
  - E.g., "How many times more valuable is 1000€ than 900€?"

- Direct methods should be employed only if indifference methods are not viable
  - Direct methods do not explicitly involve preferences for changes between achievement levels
  - Direct methods do not necessarily lead to a cardinal value function

Aggregation of values

- **Problem:** How to measure the overall value of alternative $x = (x_1, x_2, \dots x_n)$?

$$V(x_1, x_2, \dots x_n) = ?$$

- **Question:** Can the overall value be obtained by aggregating attribute-specific values?

$$V(x_1, x_2, \dots x_n) = f\big(v(x_1), \dots, v(x_n)\big)?$$

- **Answer:** Yes, if the attributes are
  - Mutually preferentially independent and
  - Difference independent

Preferential independence

- **Definition:** The attribute X is preferentially independent of the other attributes $\boldsymbol{Y}$, if for any $x, x' \in X$ and $\boldsymbol{y}' \in \boldsymbol{Y}$

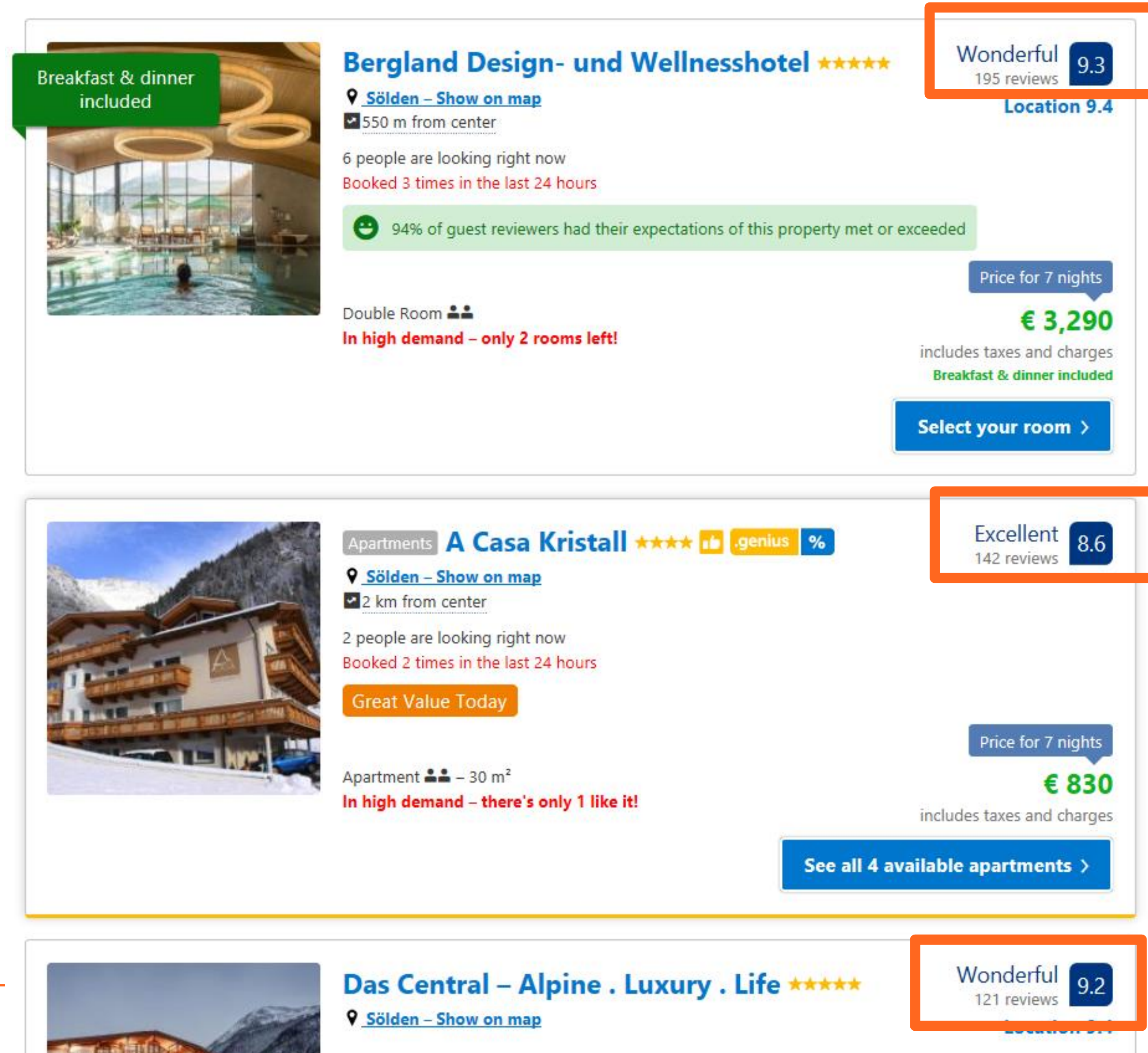
$$(x, \boldsymbol{y}') \succcurlyeq (x', \boldsymbol{y}') \Rightarrow (x, \boldsymbol{y}) \succcurlyeq (x', \boldsymbol{y}) \text{ for all } \boldsymbol{y} \in \boldsymbol{Y}$$

- Interpretation: Preference over the level of attribute X does not depend on the levels of the other attributes, as long as they stay the same
  - "All other things $\boldsymbol{Y}$ being equal (no matter what they are), an alternative with performance level x w.r.t. X is preferred to an alternative with level $x' \in X$"

Example

- Consider choosing accommodation for a (downhill) skiing vacation trip

- How do the accommodation alternatives differ from each other?
  - *What are the **attributes** that influence your decision?*



Preferential independence: example 1

- Attribute X is preferentially independent of the other attributes $\boldsymbol{Y}$, if for any $x,x' \in X$ and $\boldsymbol{y}' \in \boldsymbol{Y}$

$$(x,\boldsymbol{y}') \succcurlyeq (x',\boldsymbol{y}') \Rightarrow (x,\boldsymbol{y}) \succcurlyeq (x',\boldsymbol{y}) \text{ for all } \boldsymbol{y} \in \boldsymbol{Y}$$

- 2 Attributes
  - $X=\{1,\dots,500\}$ number of reviews
  - $Y=[1,10]$ average of reviews
- Is X preferentially independent of Y?
  - *No:* $(500,10) \succcurlyeq (5,10)$, but $(500,1) \prec (5,1)$
- Is Y preferentially independent of X?
  - *Yes (if higher average is preferred independently of #reviews,* ***as long there are equally many reviews****):* $(500,10) \succcurlyeq (500,9) \Rightarrow (x,10) \succcurlyeq (x,9)$ *for any* x

Preferential independence: example 2

- Consider choosing a meal using two attributes:
  1. Food ∈ {beef, fish}
  2. Wine ∈ {red, white}
- Preferences:
  1. Beef is preferred to fish (no matter what the wine is):
     - (beef, red) $\succcurlyeq$ (fish, red)
     - (beef, white) $\succcurlyeq$ (fish, white)
  2. White wine is preferred with fish and red wine with beef
     - (fish, white) $\succcurlyeq$ (fish, red)
     - (beef, red) $\succcurlyeq$ (beef, white)
- Food **is** preferentially independent of wine
  - *Beef* is preferred to *fish*, no matter what the *wine* is: $(x, \boldsymbol{y}') \succcurlyeq (x', \boldsymbol{y}') \Rightarrow (x, \boldsymbol{y}) \succcurlyeq (x', \boldsymbol{y})$ *for all* $y \in Y$
- Wine **is not** preferentially independent of food
  - Attribute-specific valuation of wine is not meaningful from the meal's perspective

Mutual preferential independence

- **Definition:** Attributes A are mutually preferentially independent, if any subset of attributes **X** ⊂ A is preferentially independent of the other attributes **Y** = A**X** so that for any **x, x'** ∈ ***X*, *y'*** ∈ ***Y***

  $$(\boldsymbol{x}, \boldsymbol{y}') \succcurlyeq (\boldsymbol{x}', \boldsymbol{y}') \Rightarrow (\boldsymbol{x}, \boldsymbol{y}) \succcurlyeq (\boldsymbol{x}', \boldsymbol{y}) \text{ for all } \boldsymbol{y} \in \boldsymbol{Y}$$

- Interpretation: Preference over the levels of attributes ***X*** does not depend on the levels of the other attributes **Y**, as long as they stay the same

Mutual preferential independence: example

- Consider choosing a meal using three attributes:
  1. Food ∈ {beef, fish}
  2. Side dish ∈ {potato, rice}
  3. Wine ∈ {red, white}
- Preferences:
  1. **All other things being equal**, red ≽ white, beef ≽ fish, potato ≽ rice
  2. Full meals:
     - (**beef**, rice, red)≽(**beef**, potato, white)
     - (**fish**, potato, white) ≽ (**fish**, rice, red)

  Each attribute is preferentially independent of the other two, but the attributes are not mutually preferentially independent:

  $$(y', \boldsymbol{potato}, \boldsymbol{white}) \succcurlyeq (y', \boldsymbol{rice}, \boldsymbol{red}) \nRightarrow (y, \boldsymbol{potato}, \boldsymbol{white}) \succcurlyeq (y, \boldsymbol{rice}, \boldsymbol{red})$$

Mutual pref. independence: example 2

- Choosing a car w.r.t. attributes A={*top speed*, *price*, CO_2 *emissions*}
  - Attributes defined on continuous scales

- Are all *A*'s subsets (X) preferentially independent of the other attributes (Y=A\X)?

- Each single attribute is preferentially independent of the other attributes, because
  - *Lower price is preferred to higher price independent of other attributes (if other attributes are equal)*
  - *Higher top speed is preferred to lower*
  - *Smaller emissions are preferred to bigger ones*

Mutual pref. independence: example 2

- Is X = {*price*, *CO_2 emissions*} pref. independent of Y = {*top speed*}?
  - Consider two cars which **differ in price** (e.g., 30000 e, 25000 e) and **emissions** (150 g/km, 200 g/km) so that one of the alternatives is better in emissions and the other in price. Set the same top speed for the alternatives (e.g. 230 km/h). Which one is better?
    - DM says (230 km/h, 30000 e, 150 g/km) $\succ$ (230 km/h, 25000 e, 200 g/km)
    - = when top speed is 230 km/h, she is willing to pay extra 5000 € on top of 25000 € for this emission reduction
  - Change the top speed. Is the first car still preferred to the second? e.g. **does (150 km/h, 30000 e, 150 g/km) $\succ$ (150 km/h, 25000 e, 200 g/km) hold**?
    - "No matter what the top speed, (30000 e, 150 g/km) $\succ$ (25000 e, 200 g/km)"
  - Consider other prices and emissions; does your preference hold for all top speeds?
  - If varying the top speed does not influence preference between alternatives, then {*price*, *CO_2 emissions*} is preference independent of {*top speed*}

Difference independence

- **Definition**: Attribute *X* is difference independent of the other attributes ***Y*** if and only if $(x, \boldsymbol{y}') \leftarrow (x', \boldsymbol{y}') \sim_d (x, \boldsymbol{y}) \leftarrow (x', \boldsymbol{y})$ for any *x*,*x*' ∈ X, $\boldsymbol{y}, \boldsymbol{y}' \in \boldsymbol{Y}$

- Interpretation: The preference over a <u>change</u> in attribute X does not depend on the levels of the other attributes **Y**, as long as they stay the same

Difference independence: example

- Is {*top speed*} difference independent of the other attributes {*price, CO_2 emissions*}?
  - Construct ***y*** and ***y'*** from any two levels of price and CO_2 emissions; ***y***=(25000 e, 150 g/km) and ***y'***=(30000 e, 200 g/km)
  - Consider any two levels of top speed; ***x'***=160 km/h, ***x***=200 km/h
  - Does (200 km/h, 30000 e, 200 g/km) ← (160 km/h, 30000 e, 200 g/km) $\sim_d$ (200 km/h, 25000 e, 150 g/km) ← (160 km/h, 25000 e, 150 g/km) hold?
    - If yes (for all *x*,*x*',***y***,***y'***), then difference independence holds
    - That is, does the value of increased top speed depend on the levels of other attributes or not?
    - Is the "amount of" value added by a fixed change in top speed independent of the other attributes?

Difference independence: example of implications

- We are choosing downhill skiing accommodation with regard to 6 attributes, which include *cost per night* (in €) and *possibility to go to sauna* (binary)
  - We think that (170 e, sauna, x_3, x_4, ...)~(145 e, no sauna, x_3, x_4, ...) with some $x_3,\ldots,x_6$ = we would pay an additional 25 € on top of 145 € for the sauna, with some $x_3,\ldots,x_6$
  - Then, if difference independence holds (for each attribute):

    (145 e, no sauna, x_3, x_4, ...) ← (170e, no sauna, x_3, x_4, ...) $\sim_d$

    (170 e, sauna, x_3, x_4, ...) ← (170 e, no sauna, x_3, x_4, ...) **for any $\boldsymbol{x_3,\ldots,x_6}$**
  - For any $x_3,\ldots,x_6$ = "No matter how close to nearest ski lifts , no matter how fancy the breakfast, how bad the reviews, etc."

Implication: "the improvement needed in an attribute to compensate a loss in another attribute does not depend on the levels of other attributes"

Additive value function

Theorem: If all attributes are mutually preferentially independent and each attribute is difference independent of the others, then there exists an additive value function

$$V(x) = V(x_1, \dots, x_n) = \sum_{i=1}^{n} v_i(x_i)$$

which represents preference relations $\succcurlyeq$, $\succcurlyeq_d$ in the sense that

$$V(x) \geq V(y) \Leftrightarrow x \succcurlyeq y$$
$$V(x) - V(x') \geq V(y) - V(y') \Leftrightarrow (x \leftarrow x') \succcurlyeq_d (y \leftarrow y')$$

Note: The additive value function is unique up to positive affine transformations, i.e., V(x) and V'(x)=αV(x)+β, α>0 represent the same preferences

… but where are the attribute weights w_i?

Theorem: If all attributes are (…) , then there exists an additive value function

$$V(x) = V(x_1, \ldots, x_n) = \sum_{i=1}^{n} v_i(x_i)$$

- Can these attribute-specific value functions be combined into an overall value function?

$$V(x_1, x_2, \ldots x_n) = f\big(v(x_1), \ldots, v(x_n)\big) = \sum_{i=1}^{n} w_i v_i^N(x_i)\,?$$

Normalized form of the additive value function $V(x) = V(x_1, \dots, x_n) = \sum_{i=1}^{n} v_i(x_i)$

- **Denote**
  - x_i^0 = Least preferred level w.r.t to attribute i
  - x_i^* = Most preferred level w.r.t to attribute i
- **Then,**

$$V(x) = V(x) - V(x^0) + V(x^0)$$

$$= \sum_{i=1}^{n} v_i(x_i) - \sum_{i=1}^{n} v_i\left(x_i^0\right) + V(x^0) = \sum_{i=1}^{n}\left[v_i(x_i) - v_i\left(x_i^0\right)\right] + V(x^0)$$

$$= \sum_{i=1}^{n} \underbrace{\left[v_i(x_i^*) - v_i\left(x_i^0\right)\right]}_{W_i > 0} \frac{v_i(x_i) - v_i(x_i^0)}{v_i(x_i^*) - v_i(x_i^0)} + V(x^0)$$

$$= \sum_{i=1}^{n} W_i \left[\underbrace{\frac{1}{v_i(x_i^*) - v_i(x_i^0)}}_{\alpha_i > 0} v_i(x_i) + \underbrace{\frac{-v_i(x_i^0)}{v_i(x_i^*) - v_i(x_i^0)}}_{\beta_i} \right] + V(x^0) \dots$$

Normalized form of the additive value function (cont'd)

$$\ldots = \sum_{i=1}^{n} W_i \underbrace{[\alpha_i v_i(x_i) + \beta_i]}_{v_i^N \in [0,1]} + V(x^0)$$

Normalized attribute-specific value function

$$v_i^N(x_i) = \frac{v_i(x_i) - v_i(x_i^0)}{v_i(x_i^*) - v_i(x_i^0)} \in [0,1]$$

$$= \sum_{i=1}^{n} \left[\left(\sum_{i=1}^{n} W_i\right) \cdot \underbrace{\frac{W_i}{\sum_{i=1}^{n} W_i}}_{= w_i > 0, \sum_{i=1}^{n} w_i = 1} \cdot v_i^N(x_i) \right] + V(x^0)$$

$$= \underbrace{\left(\sum_{i=1}^{n} W_i\right)}_{\chi > 0} \underbrace{\sum_{i=1}^{n} w_i v_i^N(x_i)}_{V^N(x)} + \underbrace{V(x^0)}_{\delta}$$

$$= \chi V^N(x) + \delta$$

Normalized additive value function
$V^N(x) = \sum_{i=1}^{n} w_i v_i^N(x_i) \in [0,1]$

$V(x) = \chi V^N(x) + \delta$ is a positive affine transformation of $V^N(x)$; they represent the same preferences!

Interpretation of attribute weights

- By construction, we have $w_i = \frac{v_i(x_i^*) - v_i(x_i^0)}{\sum_{i=1}^{n}(v_i(x_i^*) - v_i(x_i^0))} \propto v_i(x_i^*) - v_i(x_i^0)$
- Attribute weight w_i reflects the increase in overall value when the performance level on attribute a_i is changed from the worst level to the best – <u>relative to similar changes in other attributes</u>
- Weights thus reflect *trade-offs* between attributes; not their absolute "importance"
- The elicitation of attribute weights without this interpretation is not meaningful
  - Do not ask: "What is more important: environment or economy?"
  - Do ask: "How much is society willing to pay to save an insect species?"

Interpretation of attribute weights

- The correct interpretation and application of weights is crucial but far from trivial
- Let the least preferred and the most preferred levels in
  - *Cost savings be 0 € and 1 B€ ("money")*
  - *The number of insect species that are saved from extinction in Finland be 0 and 1 ("environmental aspects")*
  - In this setup, the environmental aspects would likely receive a very small weight (given that the weighting (0.5, 0.5) would mean that we equally prefer saving 1 B€ and saving 1 species)
- Cf. …. let the least preferred and the most preferred levels in
  - *Cost savings be 0 € and 1 B€*
  - *The number of insect species saved from extinction in Finland be 0 and 100*
- When the range between worst and best levels is wide, the attribute should have a correspondingly higher weight

Conditions

- What if the conditions (mutual preferential independence and difference independence) do not hold?
  - Reconsider the attribute ranges $[a_i^0, a_i^*]$: The conditions are more likely fulfilled when the ranges are small
  - Reconsider the attributes: Are you using the right measures? Use constructed attributes, reconsider your proxy attributes

- Even if the conditions do not hold, additive value function is in many cases employed to obtain approximate results

Example (Ewing et al. 2006*): military value of an installation

- “How to realign US Army units and which bases to close in order to operate more cost-efficiently?”

- Many attributes, including “*total heavy maneuver area*” (x_1) and “*largest contiguous area*” (x_2; a measure of heavy maneuver area quality)
  - *“Total heavy maneuver area”* x_1 *is not difference independent of the other attributes* $x_2 \cup \boldsymbol{Y}$ *because (1000 ha, 100 ha,* $\boldsymbol{y}$*”)* $\leftarrow$ *(100 ha, 100 ha,* $\boldsymbol{y}$*”)* $>$ *(1000 ha, 10 ha,* $\boldsymbol{y}$*”)* $\leftarrow$ *(100 ha, 10 ha,* $\boldsymbol{y}$*”) as the increase from 100 to 1000 ha in total area will be quite useless, if total area consists of over 100 small isolated pieces of land*

* Ewing, Tarantino, Parnell (2006): *Use of Decision Analysis in the Army Base Realignment and Closure (BRAC) 2005 Military Value Analysis*. Decision Analysis 3, 33-49

Example (Ewing et al. 2006*): military value of an installation

- Solution: unite the two attributes x_1 and x_2 into one attribute "*heavy maneuver area*"
  - Then *((1000 ha, 100 ha) ,* $\boldsymbol{y}$*') ← ((100 ha, 100 ha) ,* $\boldsymbol{y}$*')* $\succ_d$ *((1000 ha, 10 ha) ,* $\boldsymbol{y}$*) ← ((100 ha, 10 ha) ,* $\boldsymbol{y}$*)* does not violate required difference independence condition $(x, \boldsymbol{y}') \leftarrow (x', \boldsymbol{y}') \sim_d (x, \boldsymbol{y}) \leftarrow (x', \boldsymbol{y})$ for all $\boldsymbol{y} \in \boldsymbol{Y}$
  - BUT we need to elicit preferences between different 'pairs' (x_1, x_2)

Largest contiguous area (1,000s acres)	Total heavy maneuver area (1,000s acres)			
	≤10	>10 and ≤50	>50 and ≤100	>100
≤10	0.1	0.2	1.4	2.0
>10 and ≤50		3.2	4.3	5.2
>50 and ≤100			6.1	7.6
>100				10.0

Summary

- Elicitation of the attribute-specific value functions
  - *Use indifference methods if possible*
- The only meaningful interpretation for attribute weight w_i:

The improvement in overall value when attribute a_i is changed from its worst level to its best **relative to** similar changes in other attributes

- An additive value function captures the DM's preferences if and only if the attributes are mutually preferentially independent and each attribute is difference independent of the others