

Guttmans lambda (λ)

Fyrirlestur í Aðferðafræði II

© 1998, 2000 Guðmundur Arnkelsson

All rights reserved. Copying or distribution prohibited without explicit permission. Students in Methodology II at the University of Iceland may print a copy for their own private use.

Íþróttaiðkun 15–75 ára

Íþrótt	Fjöldi	%
Ganga	165	21,1
Sund	140	17,9
Skokk/polfimi	157	20,1
Boltaíþróttir	136	17,4
Annað	185	23,3
Samtals	783	100,0

Íþróttaiðkun fólks á aldrinum 15–75 ára sem segist stunda íþróttir

- Hver er besta spá um íþróttaiðkun óþekktis einstaklings?

Íþróttaiðkun eftir kyni

Íþrótt	Kyn		Samtals
	Karlar	Konur	
Ganga	61	104	165
Sund	48	92	140
Skokk/polfimi	42	115	157
Boltaíþróttir	115	21	136
Annað	121	64	185
Samtals	387	396	783

266 mistök

281 mistök

598 mistök

- Frekari upplýsingar bæta forspána
 - „Annað“ er líklegasta íþróttin fyrir karla og skokk/polfimi fyrir konur
 - Forspáin batnar við að vita kyn

λ_{yx} sem mælikvarði á bætta spá

$$\lambda_{yx} = \frac{\left(\begin{array}{l} \text{Mistök þegar frum-} \\ \text{breytan er ekki þekkt} \end{array} \right) - \left(\begin{array}{l} \text{Mistök þegar frum-} \\ \text{breytan er þekkt} \end{array} \right)}{\left(\begin{array}{l} \text{Mistök þegar frum-} \\ \text{breytan er ekki þekkt} \end{array} \right)}$$

Í okkar tilfelli verður þá niðurstaðan eftirfarandi:

$$\lambda_{yx} = \frac{598 - 547}{598} = \frac{51}{598} = 0,09$$

Tvær leiðir að lambda

$$\lambda_{YX} = \frac{\text{Mistök ef } X \text{ óþekkt} - \text{Mistök ef } X \text{ þekkt}}{\text{Mistök ef } X \text{ óþekkt}} = \frac{\text{Rétt ef } X \text{ þekkt} - \text{Rétt ef } X \text{ óþekkt}}{\text{Heildarfjöldi} - \text{Rétt ef } X \text{ óþekkt}}$$

- Lambda má ýmist reikna út frá fjölda mistaka eða aukningu í réttri spá
 - Minnkun mistaka samsvarar bættri spá
 - Algengast er að setja formúluna fram á formi bættrar spár

Formúla fyrir Guttmans λ_{yx}

$$\lambda_{YX} = \frac{\sum L_{YX} - L_Y}{N - L_Y}$$

N : Heildarfjöldi staka í töflunni

L_Y : Fjöldinn í fjölmennasta hólfinu þegar litið er framhjá frumbreytunni

L_{YX} : Fjöldi í fjölmennasta hólfinu fyrir hvert gildi frumbreytunnar fyrir sig

- Formúlan miðast við að spáð sé um fylgibreytu á grundvelli frumbreytu
 - Þetta er ósamhverf (*assymmetric*) mælitala sbr. λ_{yx}
 - Einnig er til samhverft (*symmetric*) lambda

Dæmi um útreikning

$$\lambda_{YX} = \frac{\sum L_{YX} - L_Y}{N - L_Y} = \frac{(121+115) - 185}{783 - 185} =$$
$$\frac{236 - 185}{598} = \frac{51}{598} = 0,085 \approx 0,09$$

Íþrótt sem frumbreyta

- Ósamhverfi lambda leiðir til þess að $\lambda_{YX} \neq \lambda_{XY}$
 - Því getur skipt miklu máli hvor breytan er fylgibreyta
- Í okkar dæmi fáum við mun betri spá með því að spá kyni út frá íþrótt frekar en íþrótt með kyni
 - Í dæminu mætti tákna lambda með λ_{XY} eða jafnvel $\lambda_{\text{Kyn} \text{ Íþrótt}}$

$$\lambda_{YX} = \frac{\sum L_{YX} - L_Y}{N - L_Y} = \frac{(104 + 92 + 115 + 115 + 121) - 396}{783 - 396} =$$
$$\frac{547 - 396}{387} = \frac{151}{387} = 0,39$$

Samhverft lambda (λ)

- Einnig má reikna samhverft (*symmetric*) lambda
 - Samhverft lambda hentar ef ekki er hægt að gera upp á milli breytanna, þ.e. ekki vitað hvor sé frumbreyta
- Í reynd er λ vegið meðaltal af λ_{xy} og λ_{yx}
 - Því liggur λ alltaf einhvers staðar á milli λ_{xy} og λ_{yx}

$$\lambda = \frac{\sum L_{YX} + \sum L_{XY} - L_Y - L_X}{2N - L_Y - L_X}$$

Útreikningur á samhverfu lambda

$$\lambda = \frac{\sum L_{YX} + \sum L_{XY} - L_Y - L_X}{2N - L_Y - L_X} = \frac{236 + 547 - 185 - 396}{2 \cdot 783 - 185 - 396} = \frac{783 - 581}{1566 - 581} = \frac{202}{985} = 0,205 \approx 0,21$$

Hér miðast λ við sömu gögn og áður

λ liggur greinilega á milli λ_{yx} og λ_{xy}

λ_{yx}	0,09
λ_{xy}	0,39
λ	0,21

Veikleikar lambda sem mælitölu

Lambda metur aukningu í forspá við það að þekkja frumbreytuna—ekki tengsl eða orsakaáhrif frumbreytu.

Ef frumbreyta hefur engin áhrif á fylgibreytu, eru allar skilyrtu dreifingarinnar eins og því $\lambda = 0,0$.

En lambda getur verið núll, þótt frumbreytan hafi raunveruleg áhrif á fylgibreytuna. Taflan sýnir slíkt tilvik: Hlutfallslega færri eru í flokki *a* og fleiri í flokki *d* eftir því sem frumbreytan breytist frá *A* til *D*. En flokkur *c* er alltaf fjölmennastur og því verður lambda núll þrátt fyrir þessi tengsl.

Fylgibreyta	Frumbreyta				Heild
	A	B	C	D	
a	27	12	5	1	45
b	26	14	7	3	50
c	30	15	10	5	60
d	10	7	8	4	29
Heild	93	48	30	13	184

Flokkur *c* er algengastur í bæði jaðar- og öllum skilyrtu dreifingunum. Því er alltaf best að spá flokki *c* hvort sem frumbreyta er þekkt eða ekki og lambda því núll.

Skilyrtu dreifingarinnar eru samt ekki eins og því eru fyrir hendi tengsl frum- og fylgibreytu.