Súlyozá,

Tóth Gergely
ELTE-TÁTK, Doktori iskola
Statisztika tanszék

Mi az a súlyozás?

Egyszerű válasz

- Egy eljárás, amely segítségével a mintavételt korrigáljuk, az eredmények reprezentativitását garantáljuk.
- Analóg példa: kreditpontokkal súlyozott átlag kiszámítása ... Mi a tétje?

Összetettebb válasz

történő becslés javítása érdekében.

Egy olyan matematikai transzformáció, amely segítségével

 az alapsokaság és a minta azonos paramétereinek
 ismeretében – korrekciót lehet végezni a mintavételből

Egy példa a súlyozásra – I.

Alapsokaság eloszlása									
		Ne	em		Nem				
	<u>Férfi</u>		Nő	TOTAL		Férfi	Nő	TOTAL	
	fiatal	16 667	16 667	33 334	fiatal	17%	17%	33%	
kor	középkorú	16 667	16 667	33 334	<u>ই</u> középkorú	17%	17%	33%	
	idős	16 667	16 667	33 334	idős	17%	17%	33%	
	TOTAL	50 001	50 001	100 002	TOTAL	50%	50%	100%	

Minta eloszlása									
		Ne	em		Nem				
	Férfi Nő TOTAL					TOTAL			
	fiatal	180	100	280	fiatal	18%	10%	28%	
kor	középkorú	212	80	292	<u>ই</u> középkorú	21%	8%	29%	
	idős	228	200	428	idős	23%	20%	43%	
3	TOTAL	620	380	1 000	TOTAL	62%	38%	100%	

Egy példa a súlyozásra – II.

		Eltéré	S		Eltérés százalékosan				
		Ne	em		Nem				
	<u>Férfi</u>		Nő	TOTAL		Férfi	Nő	TOTAL	
	fiatal	-1%	7%	5%	fiatal	93%	167%	119%	
kor	középkorú	-5%	9%	4%	ই középkorú	79%	208%	114%	
	idős	-6%	-3%	-9%	idős	73%	83%	78%	
	TOTAL	-12%	12%	0%	TOTAL	81%	132%	100%	

		Súlyc	k		Súlyozott értékek				
		N	em		Nem				
	Férfi Nő T			TOTAL		TOTAL			
	fiatal	0.93	1.67	1.19	fiatal	166.7	166.7	333.3	
kor	középkorú	0.79	2.08	1.14	<u>ই</u> középkorú	166.7	166.7	333.3	
	idős	0.73	0.83	0.78	idős	166.7	166.7	333.3	
4	TOTAL	0.81	1.32	1.00	TOTAL	500.0	500.0	1 000.0	

Egy példa a súlyozásra – III.

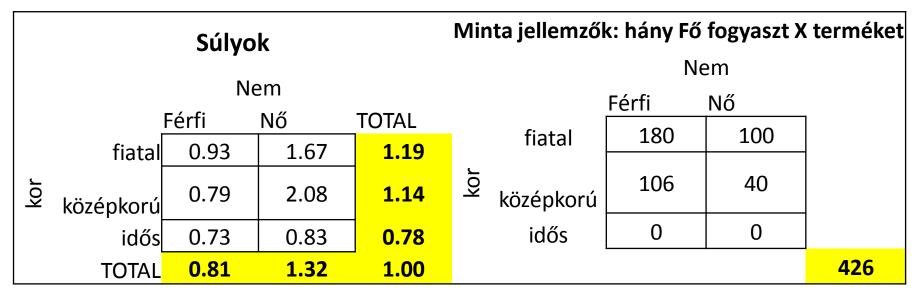
Va	alódi jellemző	k: hány % fogya	Való	Valódi jellemzők: hány Fő fogyaszt X terméke				
	Nem				Nem			
		Férfi	Nő	_	_	Férfi	Nő	
	fiatal	100%	100%		fiatal	16667	16667	
kor	középkorú	50%	50%	kor	középkor ú	8334	8334	
	idős	0%	0%		idős	0	0	
				50%				50001

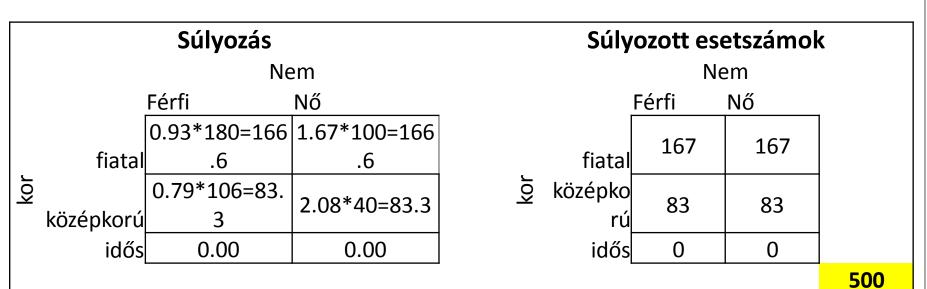
Termékhasználók aránya= 50 001/100 002 = 50%

M	linta jellemző	k: hány % fogya	szt X terméket	Minta jellemzők: hány Fő fogyaszt X terméket					
		N	em		Nem				
	_	Férfi	Nő	_	Férfi	Nő			
	fiatal	100%	100%	fiatal	180	100			
kor	középkorú	50%	50%	<u>১</u> középk orú	106	40			
	idős	0%	0%	idős	0	0			
E							426		

5

Egy példa a súlyozásra - IV.





Becslés= 500/1000 = 50%

Milyen mértékben lehet súlyozni?

A súlyozás nem "csodaszer"

Ha az eltérés oka:

- hibás az alkalmazott a mintavétel
- nagy a válaszmegtagadás
- egyéb adathiány

akkor valójában a hibák elfedését szolgálja!

- Súlyok értékei
 - =1 tökéletes megfelelés
 - <1 adott csoport felülreprezentált</p>
 - >1 adott csoport alulreprezentált

Milyen hibákat lehet korrigálni?

Valószínűségi mintákat:

1. Teljes véletlen minta



2. Rétegzett minta 🧭



- I. Eltérés mértéke nem túl nagy (súlyok eloszlása)
- II. Véletlen mintavételnek betudható hibák korrigálására

De nem alkalmas/Nem szabad használni:

A minta valószínűtlenül nagy hibáinak "eltűntetésére"



 Ennek okai sokfélék lehetnek: rossz mintavétel, véletlen, válaszmegtagadás, stb.

További fontos megjegyzés:

A súlyozás NEM garantálja a becslés jóságát



KÖSZÖNÖM A FIGYELMET!

Tóth Gergely toth.gergo@gmail.com