

Az Ökonometria és a Statisztikai Modellelés Felelőssége

Hajdu Ottó

egyetemi tanár

MTA doktora

Eötvös Loránd Tudományegyetem

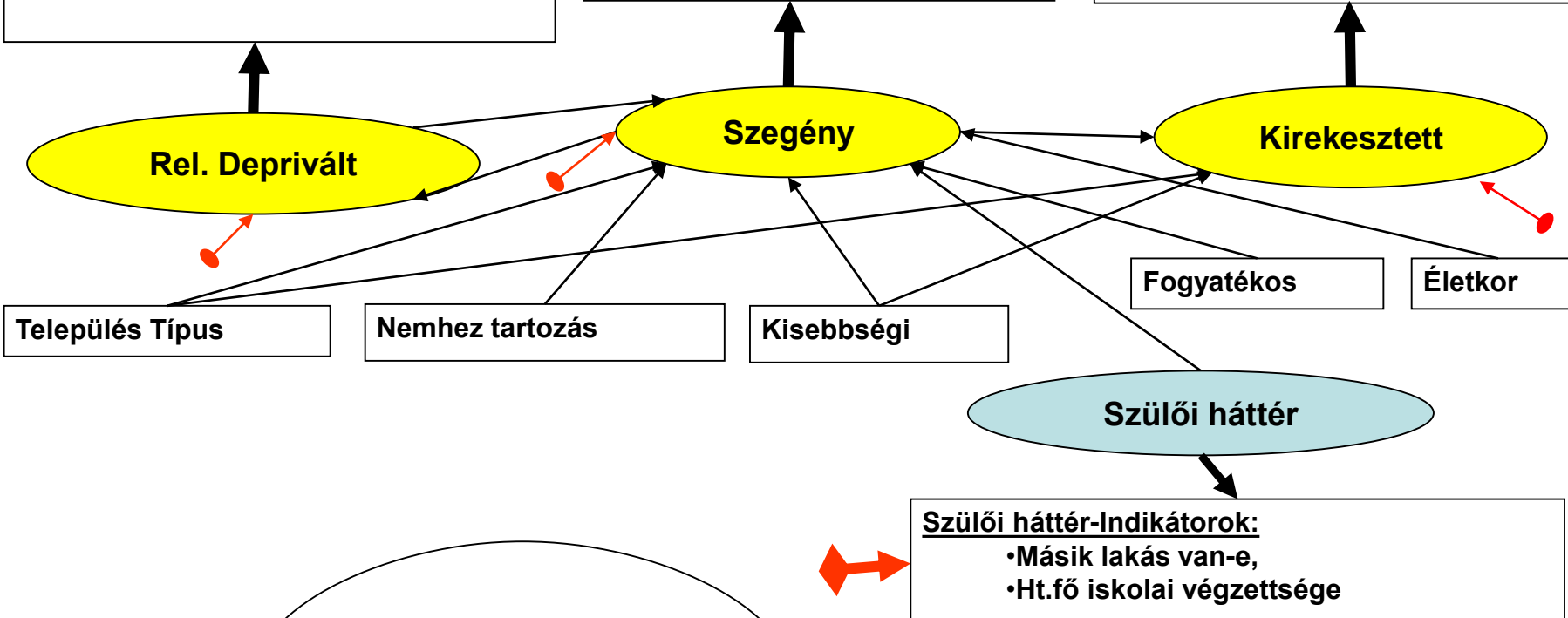
Gazdálkodástudományi Intézet

Egy Háztartásstatisztikai, fogyasztási H_0 modell körvonalai

- Depriváltság-Indikátorok:**
- Ht. Tartós fogy. cikkek száma?
 - Ht. Saját gépkocsi értéke (MFt)?
 - Ht. Lakás értéke (MFt)?
 - Ht. Másik lakás van-e?
 - Ht. Épület környezete?
 - Ht. Munkanélküliek száma?
 - Ht. Fő iskolai végzettsége?

- Szegénység-Indikátorok:**
- Folyóvíz van-e?
 - Rezsit nem fizet (Db?)
 - Lakásgondok?
 - Munkanélküliek száma?
 - Nyugdíjasok?
 - Jövedelem/fő ?

- Kirekesztettség-Indikátorok:**
- Gépkocsi futott km ?
 - Közgyógyellátás ?
 - Munkanélküliek száma ,
 - Ht.fő iskolai végzettsége?



Boksz: Manifeszt,

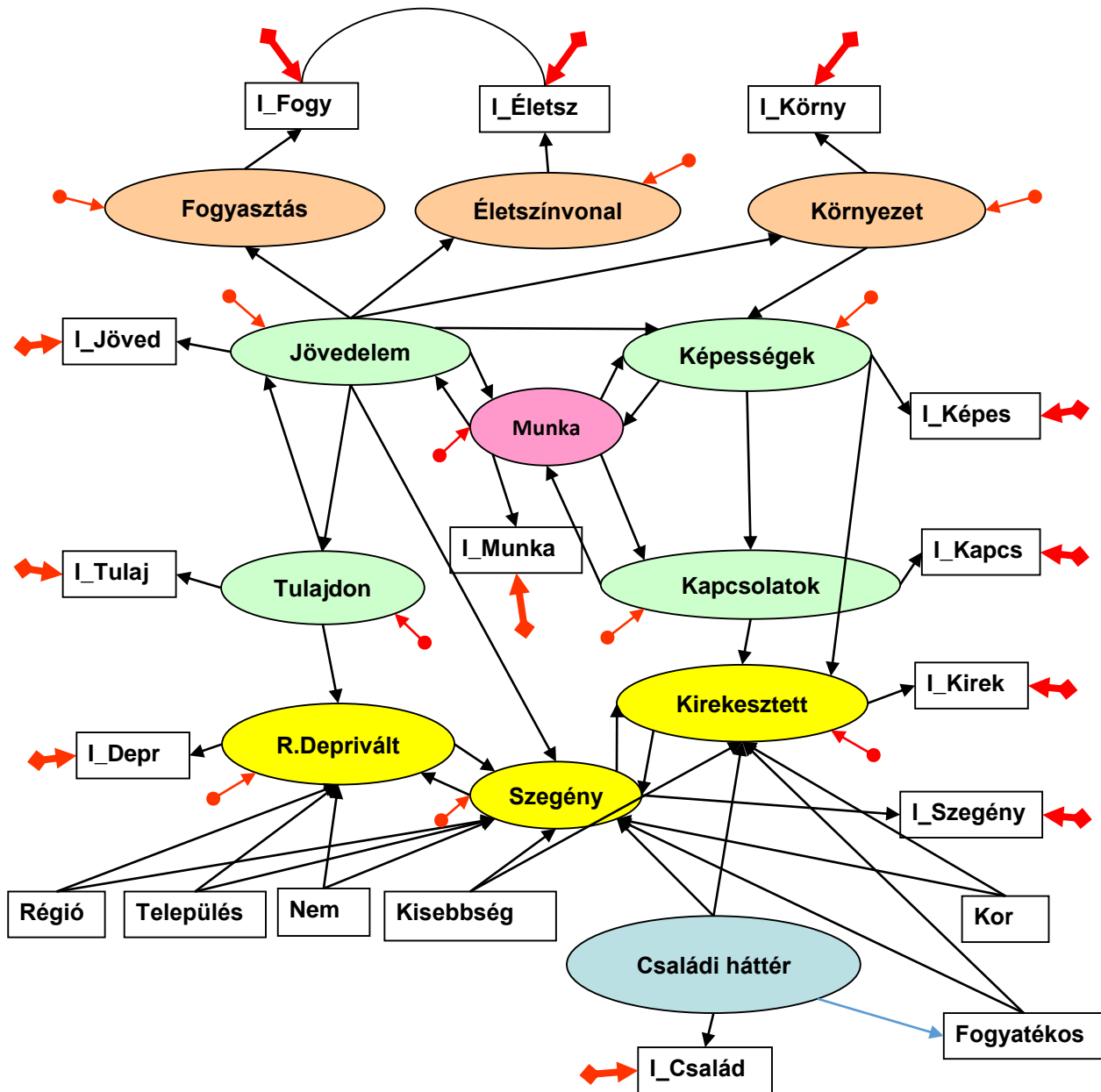
Ovális: Látens,

Írányított vastag fekete nyíl:
Látans → Manifeszt

Írányított vékony fekete nyíl:
Manifeszt → Látens

Vastag Piros nyíl:
Random Reziduum,

Vékony Piros nyíl:
Random, Unexplained Error.



Egy szofisztikáltabb H₁ modell keretei

Finomítási javaslatok:

Ív: korrelációt jelez,

Több paraméterrel jár, tehát:

1. Jobb illeszkedés,
2. Nehezebb paraméterértelmezés,
3. Szubjektív döntés a hipotézisek között
4. Kérdőjeles P-value alapján
5. Kérdés, milyen statisztikai teszt alapján alkalmazzuk a P-veget,
6. Vagy hagyjuk a P-value-t másra !

Az adatstruktúra kialakítása

1. A **Missing** value pótlása / nempótlása:
 1. Imputáció esetén: nem létező információt viszünk az adatállományba → torzítás,
 2. Üres cella megtartása: a Modell elveszti a cellát tartalmazó indikátor változót.
2. Az **Outlier** szűrés meg / meg_nem valószínűsítése:
 1. A nemszűrés **torzítja, elvezeti** a valódi tendenciát,
 2. Az outlierok *elhagyása információvesztéssel* jár,
 3. Az outlier-okozott *strukturális törés* kezelése **többlet_paramétereket** igényel.
3. **Zajos** adatállomány kisimítása:

Feltárhat tendenciákat, és Elkendőzhet tendenciákat.
4. Az **Adatsűrítés / Adatbővítés** problémája:
 1. A *Maximum Likelihood* becslés Large-Sample követelményű,
 2. A *Big-data adatbőség* fölösleges, zajos, ismétlődő információkat tartalmaz,
 3. Az **Egzakt** statisztikai **következtetés** *kismintát* igényel.
5. **A mérési hiba: helytelen Proxy !**
 1. „Jövedelmezőség” → Eredmény/Sajátjövedelem, vagy Eredmény/Nettó árbevétel,
 2. „Budapesti lakásár Övezete” a Kerület → Kérdés: mi legyen a Proxy-kör?

Látens Dimenziók, Manifeszt Indikátorok, és a Prediktorok

1. Legyenek **látens** dimenziók: pl. a Jövedelmezőség, a Likviditás, az Adósság”
 1. Ahol a megfigyelhető proxy *Indikátor (okozott)* manifeszt változók lehetnek pl.:
 1. Eszközarányos adózott eredmény, stb.
 2. Forgóeszköz / rövid lejáratú kötelezettségek, stb.
 3. Bonitás = H.L.Kötelezettségek / Saját tőke, stb.
 2. A *manifeszt (ok)* prediktorok lehetnek pl.:
 1. Zrt., Kft., Bt., Menedzsment, Kapcsolati tőke, Stratégia, stb.
2. **A kutatási cél:** Látens dimenziók oksági kapcsolatait mérni, ahol
 1. Látens, és Manifeszt is lehet Függő és Független változó is egy többegyenletes modellben.
3. **A mérési hiba: helytelen Proxy !**
 1. „Jövedelmezőség” → Eredmény/Sajáttőke, vagy Eredmény/Nettó árbevétel,
 2. „Budapesti lakásár Övezete” a Kerület → Kérdés: mi legyen a Proxy-kör?

A Statisztikai Következtetés

Sokaság → Minta → P-érték

1. A mintavételi következtetés kényszere: nincs meg a teljeskörű, sokasági adat:
 1. Hipotézisek statisztikai tesztelése,
 2. Konfidencia tartományok szerkesztése.
2. Ha közel van a teljeskörűség esete, **okafogyott** a mintavételi következtetés → értelmét veszti:
 1. a döntést segítő P-value számítás,
 2. és a CI konfidencia intervallum számítása is.
3. A döntéshozó **felelőssége**:

Tudni kell *nem alkalmazni mechanikusan* a „misztifikált” Prob-értéket!
4. A modellt szakmai, közgazdasági alapokon specifikáljuk !

Modellszelekció

1. Mértékegységválasztás: **skálázási** kérdés, pl. pillanatnyi növekedési ütem.
2. A modelltípus kutatási kérdésnek megfelelő specifikálása pl.:

1. Lineáris modell,
2. Nemlineáris modellek:
 1. „Cobb-Douglas” típusú „termelési” log-log modell,
 2. „Növekedési” típusú exponenciális log-lin modell,
 3. „Keresleti-kínálati” típusú lin-log modell,
 4. „Telítődési” típusú, reciprok modell,

Ahol a *Marginális* hatás, a *Rugalmasság*, a *Semi-rugalmasság* **modell-specifikus!**

3. A modellező **felelőssége:**

Ha ezek elméleti lefutása ellentmond az empirikusnak, azt a típust nem illesztjük, és...

Újabb paraméter bevonása büntetendő, ha már **nem javítja** érdemben a modellt:

A modellszelekciós kritériumok szigorúsági rangsora:

Akaike → Hannan-Quinn → Bayes

Modellilleszkedés vs. Változófüggetlenség

1. A statisztikai függetlenségvizsgálat hipotéziseinek állítása:
Az aktuálisan **becsült** Modell **távolsága** a **Függetlenség** esetétől:
 1. *Null hipotézis*: Elhanyagolható
 2. *Alternatív hipotézis*: Jelentős → ez az eset preferált.
2. A statisztikai illeszkedésvizsgálat hipotéziseinek állítása:
Az aktuálisan **becsült** Modell **távolsága** a **Szaturált** Modelltől
 1. *Null hipotézis*: Elhanyagolható → ez az eset preferált,
 2. *Alternatív hipotézis*: Jelentős.
3. A döntéshozó **felelőssége** ezen problémáknál rendre:
 1. A modellspecifikus **Metrika** kiválasztása: **F**, vagy **Khi²** távolság,
 2. A mintaméretnek megfelelő szignifikancia szint korrekt megválasztása,
 3. A *nemfüggetlenség hipotézis* elfogadása **Kicsiny** Probability-értéket,
 4. Az *illeszkedés* hipotézis elfogadása **Magas** Probability-értéket preferál!

A becslési módszer következményei

??? Minél kisebbet hibázni \rightarrow OLS,
Minél hihetőbbet állítani \rightarrow ML,
Minél hübb lenni a minta momentumhoz \rightarrow MM ???

1. A Maximum Likelihood célfüggvény: eloszlásfüggő,
 1. A -2logLikelihood Chi²-modellszelektálást tesz lehetővé,
 2. A Pseudo R² jelentése, összehasonlíthatósága: ???
 3. Alapértelmezések alkalmazása, bár az X-eloszlás nemnormális,
 4. Van-e ADF lehetőség? (Asymptotically Distribution Free), pl. SEM-modell
2. A konvergenciaszint kalibrálása,
„Failed to converge!” \rightarrow Akkor is van közbülső eredmény !
3. OLS, GLS, GMM?
 1. Az OLS standard hiba is eloszlás (normalitás) érzékeny,
 2. A GLS módszer hozadéka nem praktikus,
 3. A momentum módszer túlzottan robusztus.

Köszönöm a figyelmet