Matematikai geodéziai számítások 6.

Lineáris regresszió számítás elektronikus távmérőkre

Dr. Bácsatyai, László

Matematikai geodéziai számítások 6.: Lineáris regresszió számítás elektronikus távmérőkre

Dr. Bácsatyai, László
Lektor: Dr. Benedek, Judit

Ez a modul a TÁMOP - 4.1.2-08/1/A-2009-0027 „Tananyagfejlesztéssel a GEO-ért” projekt keretében készült. A projektet az Európai Unió és a Magyar Állam 44 706 488 Ft összegben támogatta.

v 1.0

Publication date 2010

Kivonat

Ez a modul bemutatja, hogyan lehet megállapítani, van-e és ha van, milyen a kapcsolat az elektronikus távmérővel mért távolságok nagysága, s a mérési hibák között.

Tartalom

6. fejezet - Lineáris regresszió számítás elektronikus távmérőkre

1. 6.1 A feladat megfogalmazása

Elektronikus távmérővel különböző nagyságú, ismert valódi értékű távolságokat mértünk. Adottak a mért d_i távolságok 0,1 km élességgel és a mért távolságok Δ_i valódi hibáinak $|\Delta_i|$ abszolút értékei.

Határozzuk meg a korrelációs együtthatót és írjuk fel a regressziós egyenes egyenletét! Ítéljük meg, hogy a kapott összefüggés alapján következtethetünk-e összefüggésre a mért távolságok nagysága és a valódi hibák $|\Delta_i|$ abszolút értékei között.

Leadandók különálló borítólapba foglalva:

- A feladatkiírás és a kiinduló adatok (feladatlapba foglalva),
- Számítások listája a részeredményekkel együtt,
- A regressziós egyenes grafikonja

A feladatot az EXCEL használata nélkül, manuálisan, zsebkalkulátorral kell megoldani, s a felhasznált képletekkel és tájékoztató szöveges információkkal együtt – különálló borítólapba foglalva - kézzel írott, vagy Microsoft Word formátumban kell leadni.

2. 6.2 A feladatban szereplő fogalmak

Két valószínűségi változó kapcsolatának jellemzésére a kovarianciát és a korrelációs együtthatót használjuk.

$$\mathrm{cov}_{uz} = \frac{\sum_{i=1}^{n}(u_i - U)(z_i - Z)}{n} = \frac{\sum_{i=1}^{n}\Delta_{u_i}\cdot\Delta_{z_i}}{n}.$$

A fenti képletben cov_{uz} - az előzetes középhibához hasonlóan – az *előzetes kovariancia*, amely számítható akkor, ha ismerjük az u_i és z_i valószínűségi változók várható értékeit. A továbbiakban valószínűségi változó helyett a *mérési eredmény*, várható érték helyett a *valódi érték* kifejezéseket fogjuk használni. Ekkor a Δ_{ui} és Δ_{zi} mennyiségek a mérési eredmények (valódi) véletlen hibái. Belátható, hogy egyetlen mennyiségnek saját magával alkotott kovarianciája a *variancia* (az *előzetes középhiba* négyzete):

$$\mu_u^2 = \frac{\sum_{i=1}^{n}(u_i - U)(u_i - U)}{n} = \frac{\sum_{i=1}^{n}\Delta_{u_i}^2}{n}.$$

A korrelációs együttható a

$$r_{uz} = \frac{\sum_{i=1}^{n}(u_i - U)(z_i - Z)}{n\cdot\mu_u\cdot\mu_z} = \frac{\sum_{i=1}^{n}\Delta_{u_i}\cdot\Delta_{z_i}}{\sqrt{\sum_{i=1}^{n}\Delta_{u_i}^2\cdot\sum_{i=1}^{n}\Delta_{z_i}^2}}$$

képlettel fejezhető ki, ahol μ_u és μ_z az u és z mérendő mennyiségek *előzetes középhibái*.

Az r_{uz} korrelációs együttható értéke 0 és 1 közé esik. Ha r_{uz} = 0, azt mondjuk, hogy a két mérendő mennyiség korrelálatlan. A *korrelálatlanság* csak akkor jelent függetlenséget is, ha a mérési eredmények eloszlása *normális*. r = 1 esetén a korreláció maximális. Mivel a geodéziai mérési eredmények matematikai feldolgozásának egyszerűsítése céljából a méréseknek függetlennek kell lenniük, fontos, hogy milyen megbízhatósággal számítható ki a korrelációs együttható értéke, ill. milyen abszolút értéke mellett tekinthető szignifikánsnak a két mérendő mennyiség függőségére vonatkozóan. A feladatban a szigorú szignifikancia vizsgálattól eltekintünk, saját megítélésünk alapján kell megítélnünk a kapcsolat hiányát, ill. esetleges létezését.

Becsült értékük alapján két mérési sorozat (minta) közötti lineáris kapcsolat szorosságára következtethetünk [1]. Kapcsolat esetén az egyik mennyiség (pl. u) mért értékéből e függőségi kapcsolat - a regresszió - alapján *becsülhető* a másik mennyiség (pl. z) értéke. Lineáris regresszió esetén a z = f(u) függvény geometriai képe egyenes, amelyet ezért regressziós egyenesnek is nevezünk.

Legyen két mérési sorozatunk az alábbi mérési eredményekkel:

u_i	$u_1\, u_2 \ldots\ldots\ldots\, u_n$
z_i	$z_1\;\; z_2 \ldots\ldots\ldots\, z_n$

Az utólagos középhiba (négyzetének) analógiájára nevezzük a

$$c_{uz} = \frac{\sum_{i=1}^{n}(u_i - \bar{u})(z_i - \bar{z})}{n-1} = \frac{\sum_{i=1}^{n} v_{u_i} v_{z_i}}{n-1}$$

kifejezést *utólagos kovarianciának*.

Ekkor a korrelációs együttható becsült, tapasztalati értéke megadható az

$$r_{uz} = \frac{c_{uz}}{\mu_u \cdot \mu_z} = \frac{\sum\limits_{i=1}^{n} v_{u_i} \cdot v_{z_i}}{\sqrt{\sum\limits_{i=1}^{n} v_{u_i}^2 \cdot \sum\limits_{i=1}^{n} v_{z_i}^2}}$$

kifejezéssel.

A képletek jelölései:

$$v_{ui} = u_i - \bar{u},\; v_{zi} = z_i - \bar{z},$$

$$\bar{u} = \frac{\sum u_i}{n}; \quad \bar{z} = \frac{\sum z_i}{n}; \quad \mu_u = \sqrt{\frac{\sum (u_i - \bar{u})^2}{n-1}}; \quad \mu_z = \sqrt{\frac{\sum (z_i - \bar{z})^2}{n-1}}.$$

Az - empirikus - *regressziós egyenes* egyenletei kifejezhetők a

[1] A korrelációs együttható csak a lineáris kapcsolat jellemzésére megfelelő mérőszám, két valószínűségi változó közötti más függvénykapcsolatról nem kapunk információt. Egy meghatározott függvénykapcsolat szorosságát a korrelációs index méri.

$$z = \bar{z} + r_{uz} \cdot \frac{\mu_z}{\mu_u} \cdot (u - \bar{u}),$$

$$u = \bar{u} + r_{zu} \cdot \frac{\mu_u}{\mu_z} \cdot (z - \bar{z})$$

alakban, attól függően, hogy *z* függőségét vizsgáljuk *u* - tól, vagy *u* függőségét *z* - től. A fenti egyenletekben az eddigi jelölések mellett az egyenesek

$$b_z = r_{uz} \cdot \frac{\mu_z}{\mu_u}$$

$$b_u = r_{zu} \cdot \frac{\mu_u}{\mu_z}$$

iránytangenseit regressziós együtthatóknak nevezzük. A

$$z = a_z + b_z \cdot u,$$
$$u = a_u + b_u \cdot z$$

regressziós egyenesek megkaphatók a *legkisebb négyzetek elve* alapján is a

$$F(a_z, b_z) = \sum_{i=1}^{n} (z_i - a_z - b_z \cdot u_i)^2 = \min.,$$

$$G(a_u, b_u) = \sum_{i=1}^{n} (u_i - a_u - b_u \cdot z_i)^2 = \min.$$

feltételekből kiindulva. Az összefüggésekben adottak az u_i és z_i eredmény-párok, keressük a regressziós egyenesek a_z, b_z, ill. a_u, b_u együtthatóit.

2.1. 6.2.1 Számpélda

Kiinduló adatok:

A mérés sorszáma	A mérési eredmények	
	u_i (d_i) (km)	z_i ($\lvert \Delta_i \rvert$) (mm)
1	1,7	9,1
2	0,7	6,5
3	1,2	9,1
4	0,7	2,0
5	1,0	16,1
6	1,2	6,3

7	0,5	3,4
8	1,0	6,1
9	0,6	7,2
10	0,1	0,1
11	1,1	13,5
12	0,4	3,2
13	0,2	2,6
14	1,5	7,3
15	0,8	21,2
16	0,4	7,1
17	0,2	9,3
18	1,3	8,4
19	1,4	10,4
20	0,1	5,0

A számítások a következő táblázatban láthatók:

A mérés sorszáma	A mérési eredmények		Számítások				
	u_i (d_i) (km)	z_i ($\lvert\Delta_i\rvert$) (mm)	$v_u = u_i - \bar{u}$	$v_z = z_i - \bar{z}$	v_u^2	v_z^2	$v_u \cdot v_z$
1	1,7	9,1	0,89	1,4	0,792	1,96	1,246
2	0,7	6,5	-0,11	-1,2	0,012	1,44	0,132
3	1,2	9,1	0,39	1,4	0,152	1,96	0,546
4	0,7	2,0	-0,11	-5,7	0,012	32,49	0,627
5	1,0	16,1	0,19	8,4	0,036	70,56	1,596
6	1,2	6,3	0,39	-1,4	0,152	1,96	-0,546
7	0,5	3,4	-0,31	-4,3	0,096	18,49	1,333
8	1,0	6,1	0,19	-1,6	0,036	2,56	-0,304
9	0,6	7,2	-0,21	-0,5	0,044	0,250	0,105
10	0,1	0,1	-0,71	-7,6	0,504	57,76	5,396
11	1,1	13,5	0,29	5,8	0,084	33,64	1,682
12	0,4	3,2	-0,41	-4,5	0,168	20,25	1,845

13	0,2	2,6	-0,61	-5,1	0,372	26,01	3,111
14	1,5	7,3	0,69	-0,4	0,476	0,160	-0,276
15	0,8	21,2	-0,01	13,5	0,000	182,25	-0,135
16	0,4	7,1	-0,41	-0,6	0,168	0,36	0,246
17	0,2	9,3	-0,61	1,6	0,372	2,56	-0,976
18	1,3	8,4	0,49	0,7	0,240	0,490	0,343
19	1,4	10,4	0,59	2,7	0,348	7,290	1,593
20	0,1	5,0	-0,71	-2,7	0,504	7,290	1,917
Átlag: $\bar{u} = \frac{\sum u_i}{n} = 0,81$		$\bar{z} = \frac{\sum z_i}{n}$ = 7,7 mm	$\sum_{i=1}^{20} v_{u_i} =$ -0,1 km	$\sum_{i=1}^{20} v_{z_i} =$ -0,1 mm	Σ = 4,568	Σ = 469,730	Σ = 19,454

A táblázatban foglalt számítások alapján számíthatók a következő mennyiségek:

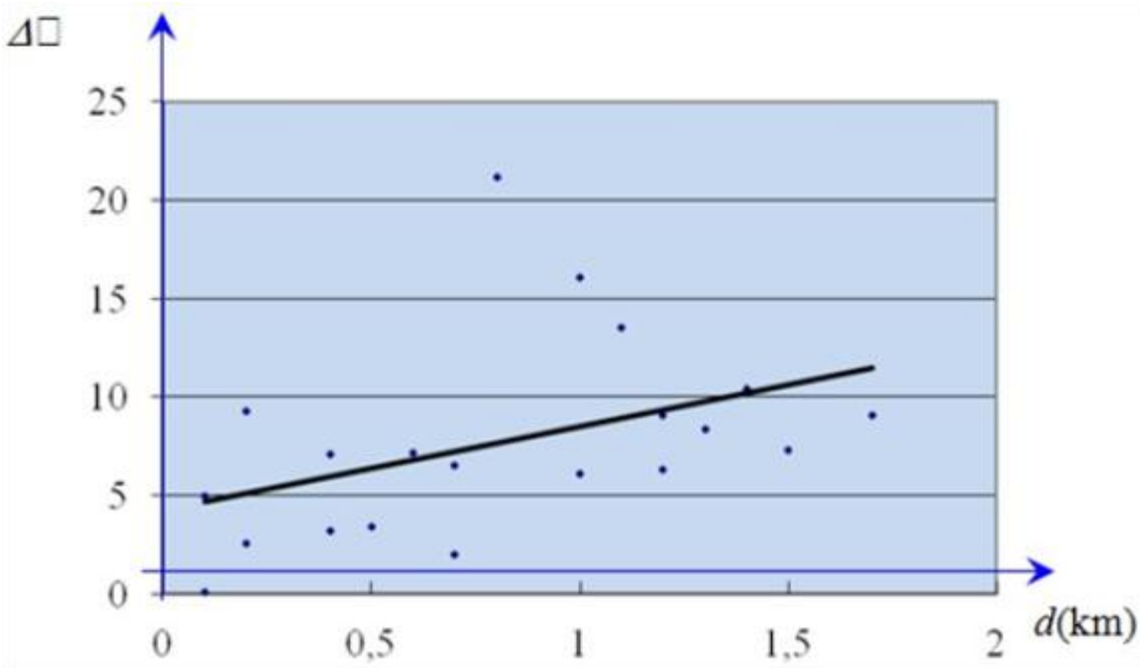
$$\mu_z = \mu_\Delta = \sqrt{\frac{\sum_{i=1}^{n}(z_i - \bar{z})^2}{n-1}} = \sqrt{\frac{469,73}{19}} = \pm 4,97\ \text{mm} \approx \pm 5\ \text{mm};$$

$$\mu_u = \mu_d = \sqrt{\frac{\sum_{i=1}^{n}(u_i - \bar{u})^2}{n-1}} = \sqrt{\frac{4,568}{19}} = \pm 0,49\ \text{km} \approx \pm 0,5\ \text{km};$$

$$r_{uz} = \frac{\sum_{i=1}^{n} v_{u_i} \cdot v_{z_i}}{\sqrt{\sum_{i=1}^{n} v_{u_i}^2 \cdot \sum_{i=1}^{n} v_{z_i}^2}} = \frac{19,454}{\sqrt{4,568 \cdot 469,73}} = \frac{19,454}{46,321} = +0,42$$

.

Az *ui (di), zi* ($|\Delta_i|$) pontok és a regressziós egyenes grafikus ábrázolása (az abszcissza tengelyen a távolságok km-ben, az ordináta tengelyen a mérési hibák mm dimenzióban szerepelnek).



A $|\Delta|$ valódi mérési hibáknak a távolságoktól való függőségét kifejező

$$z = \bar{z} + r_{uz} \cdot \frac{\mu_z}{\mu_u}\left(u - \bar{u}\right),$$

vagy, más jelölésekkel

$$|\Delta| = |\bar{\Delta}| + r_{uz} \cdot \frac{\mu_{|\Delta|}}{\mu_d}\left(d - \bar{d}\right)$$

regressziós egyenes egyenlete:

$$|\Delta| = 7{,}7\,\mathrm{mm} + 0{,}42 \cdot \frac{5}{0{,}5} \cdot (d - 0{,}81) = 7{,}7\,\mathrm{mm} + 4{,}2 \cdot (d - 0{,}81) = (4{,}3 + 4{,}2 \cdot d)\,\mathrm{mm}$$

ahol, hogy az eredményt mm dimenzióban kapjuk meg, a *d* értékét km-ben kell behelyettesíteni (a 4,2 regressziós együttható dimenziója ui. mm/km).

A grafikon alapján a két mennyiség között - nem túl erős - *korrelációt* lehet gyanítani. Megalapozottabb, meggyőzőbb válasz azonban csak nagyobb minta alapján lenne adható.

Fentiekből következik, s ezt igazolja a gyakorlat is, hogy az elektronikus távolságmérés eredményét egy távolságtól független és egy távolságfüggő hibatag befolyásolja. Megemlítjük, hogy e példában a fordított esetnek - a távolságnak a valódi hibáktól való függőségének - nincs értelme.

Irodalomjegyzék

Bácsatyai László: *Kiegyenlítő számítások, elektronikus jegyzet pdf formátumban,* NYME Geoinformatikai Kar, Székesfehérvár,