

TÖBBSZÖRÖS REGRESZIÓS ANALÍZIS I. Többszörös lineáris regresszió

Füst György

Többszörös regresszió I.

miért elengedhetetlen a többszörös regressziós számítás?

- a többszörös regressziós számítások fajtái
- a többszörös lineáris regresszió egyenlete
- többszörös lineáris regressziós számítás elvégzése számítógépen

KÉT VÁLTOZÓ KÖZÖTTI KAPCSOLAT MÉRÉSI MÓDJAI: A KORRELÁCIÓ ÉS A REGRESSZIÓ

- **Az alapvető kérdés: van-e kapcsolat két, ugyanabban az egyénben, állatban, kísérleti mintában, stb. mért különböző változó között?**
- **Ha csak arra vagyunk kíváncsiak, hogy ilyen kapcsolat fennáll-e, akkor korrelációt számítunk, ha arra is, hogy ha fennáll ilyen kapcsolat, akkor az egyik változó értékeiből hogyan lehet előre jelezni a másik változó értékeit, akkor regressziós, általában lineáris regressziós számítást végzünk. A korreláció és a regresszió között sok a hasonlóság, ha a korreláció mérőszáma az un. korrelációs koefficiens szignifikáns, akkor mindig szignifikáns lesz a lineáris regresszió is.**
- **A leggyakrabban használt és az orvosi irodalomban igen gyakran megtalálható eljárások.**

A korrelációs számítás legfontosabb szabálya: a szignifikáns korreláció sem jelent ok-okozati kapcsolatot

- Ha x és y között erős korreláció van, akkor az lehet azért, mert
- 1. az y változásai okozzák az x változásait
- 2. a x változásai okozzák az y változásait
- 3. egy harmadik faktor mind az x -et, mind az y -t egy irányba (vagy ellenkező irányba) befolyásolja. Ez a leggyakoribb!!!

EGYSZERŰBB MODEL (KATZ)

Kimenetel (függő változó)	Példa a kimenetelre	A használandó többszörös analitikai módszer
Folyamatos	Vérnyomás, testsúly, hőmérséklet	Többszörös lineáris regresszió
Dichotóm (igen-nem)	Halál, rák, felvétel intenzív osztályra	Többszörös logisztikus regresszió
Az igen eseményig eltelt idő	A halálig, a rák dg-ig eltelt idő	Cox regresszió (proportional hazard analízis)

Példa a többszörös lineáris regresszióra (Burián et al, Circulation 2001)

	Súlyos ISZB (n=248)	Kontroll be- tegek (n=53)	p
HDL chol, mmol/l	1.22 (0.67- 2.05)	1.29 (1.14- 1.37)	0.006
triglicerid, mmol/l	2.5 (0.3- 16.6)	1.96 (0.9- 6.3)	0.016
anti-hsp60, AU/ml	102 (0-2410)	57 (0-722)	0.0001
Chl.pneum poz., %	79.4	64.2	0.021

1. kérdés: van-e összefüggés az anti-hsp60 és a páros összehasonlításnál szignifikáns különséget adó másik 3 változó között? **nincs vagy gyenge**

	Spearman r	p
HDL chol.	- 0.082	0.271
triglicerid	0.137	0.022
Chl. pneu pozitivitás	0.006	0.917

STATISTICA OUTPUT I

N=289	BETA	ST.ERR OF BETA	B	ST.ERR OF B	t (284)	p level
inter- cept			1.65	0.097	17.16	0.00000
csoport	0,213	0.059	0.30	0.083	3.61	0.00036
HDL chol	-0.018	0.059	-0.018	0.050	0.31	0.75482
trigl.	0.052	0.060	0.015	0.017	0.87	0.38299
Chl. pneum.	0.030	0.058	0.037	0.072	0.51	0.60772

STATISTICA OUTPUT II

Regression Summary for Dependent Variable:

LOGHSP60

R= ,22932224

R²= ,05258869

Adjusted R²= ,03924487

F(4,284)=3,9411 p<,00395

Std.Error of estimate: ,51360

STATISTICA OUTPUT III

	Sum of squares	df	Mean squares	F	p-level
Regression	4.1583	4	1.03959	3.94105	0.00395
Residual	74.9149	284	0.26379		
Total	79.0732				

A regressziós egyenes egyenlete

- $Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \epsilon$

a használt egyenlet a minta alapján:

- $Y = a + b_1 X_1 + b_2 X_2 + b_3 X_3 + b_4 X_4 \dots$

TÖBB VÁLTOZÓ HATÁSÁT EGYETLEN ÉRTÉKBEN ÖSSZEGEZZÜK (súlyozott átlag)

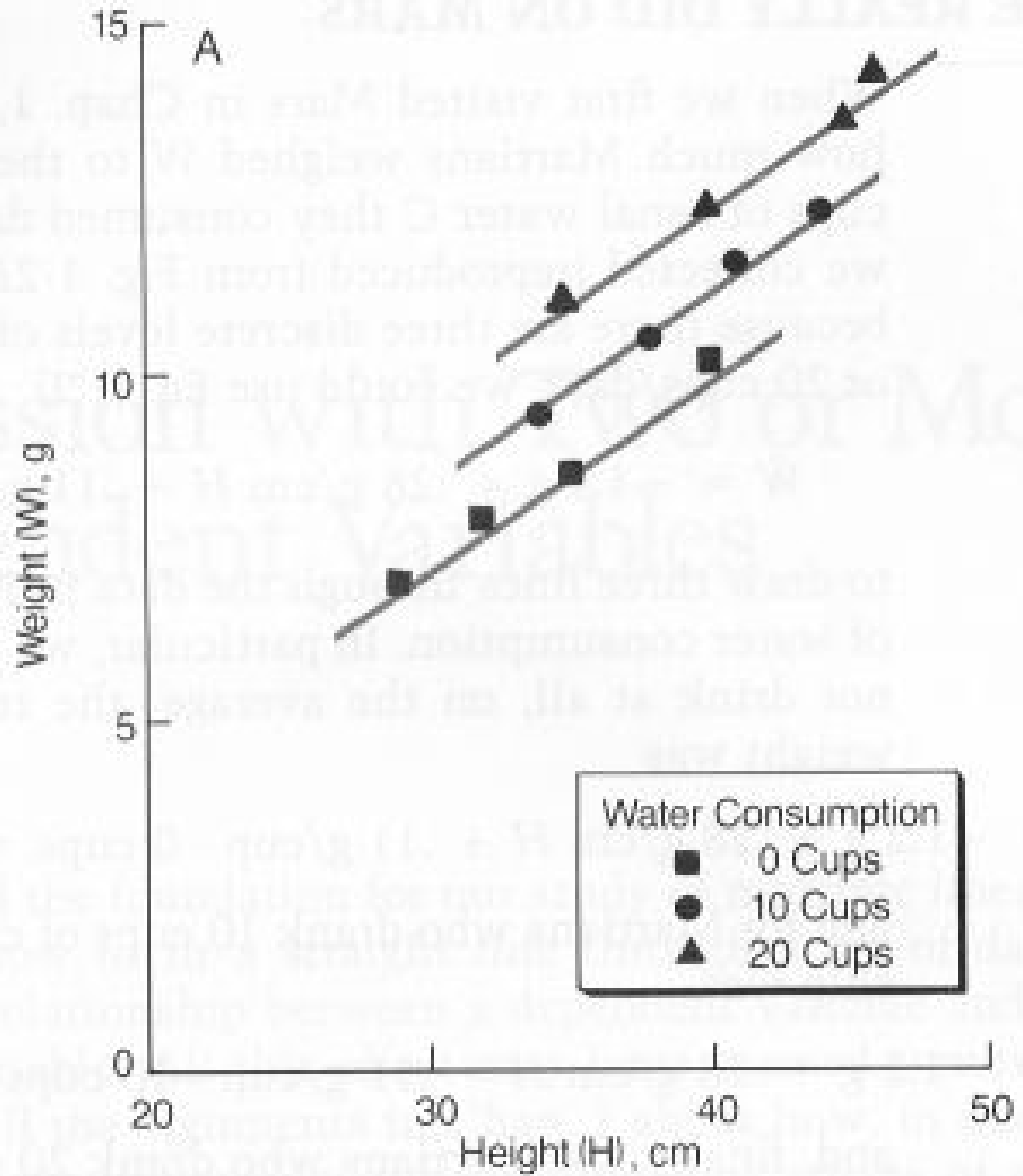
ahol az X_1 az első független változó és a b_1 a hozzá tartozó regressziós koefficiens, az X_2 a második független változó és a b_2 a hozzá tartozó regressziós koefficiens, stb.

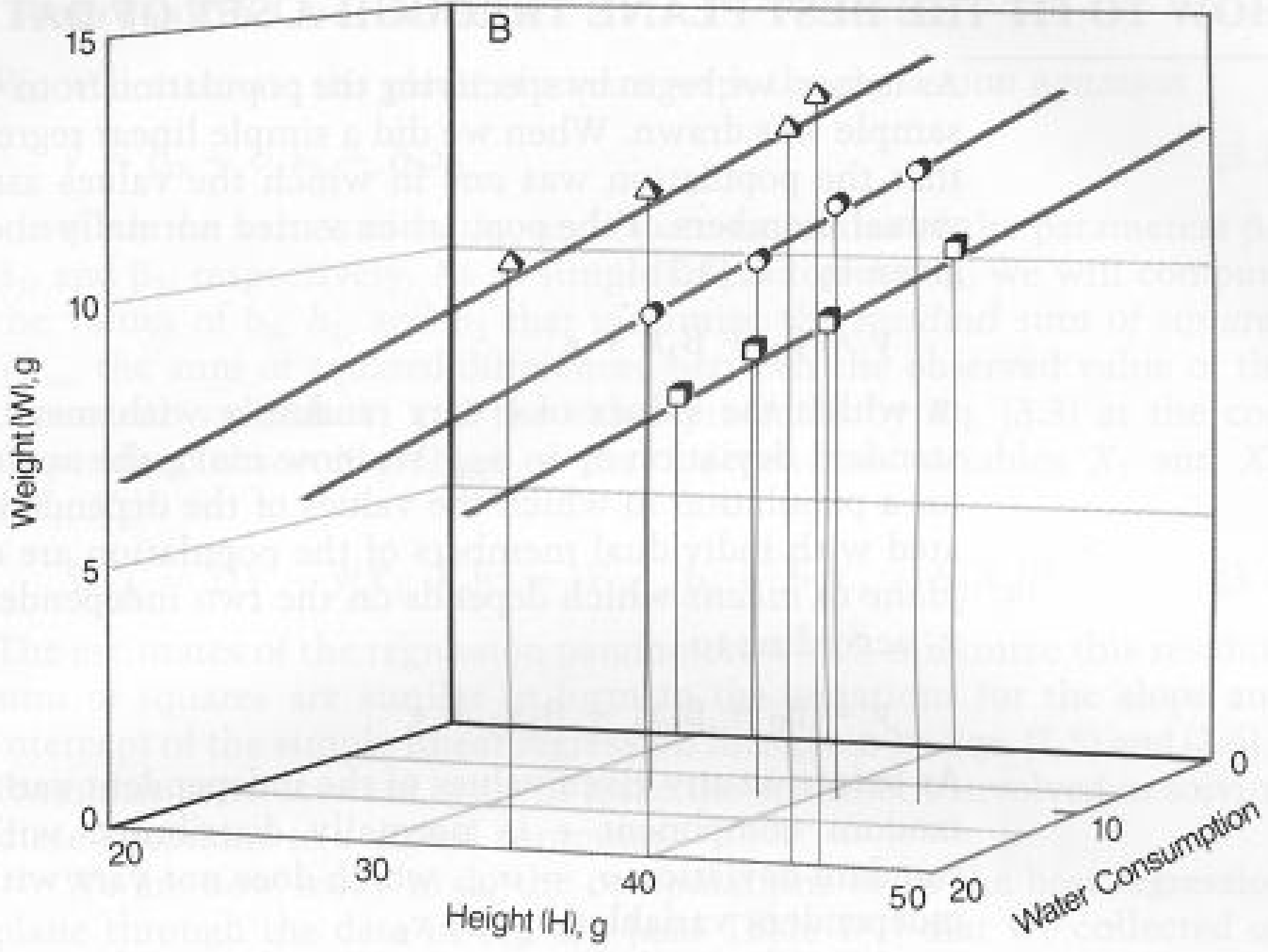
A regressziós egyenes egyenlete (folyt.)

- **A számítás hasonló az egyszerű lineáris regresszióhoz, a legkisebb átlagos négyzetes távolság kiszámításán alapul.**
- **Két független változó esetén egy síktól való távolságot minimalizálunk, több független változónál ez már nem szemléltethető**

Glanzt SA, Slinker BK: Primer of Applied Regression and Analysis of Variance, McGraw-Hill, 1990

- Látogatás a Marson. Összefüggés a marslakók magassága és testsúlya között. (egyszerű regresszió).
Befolyásolja-e ezt az összefüggést az, hogy a marslakók naponta hány csésze, a Mars csatornáiból származó vizet fogyasztanak (0, 10 vagy 20)?





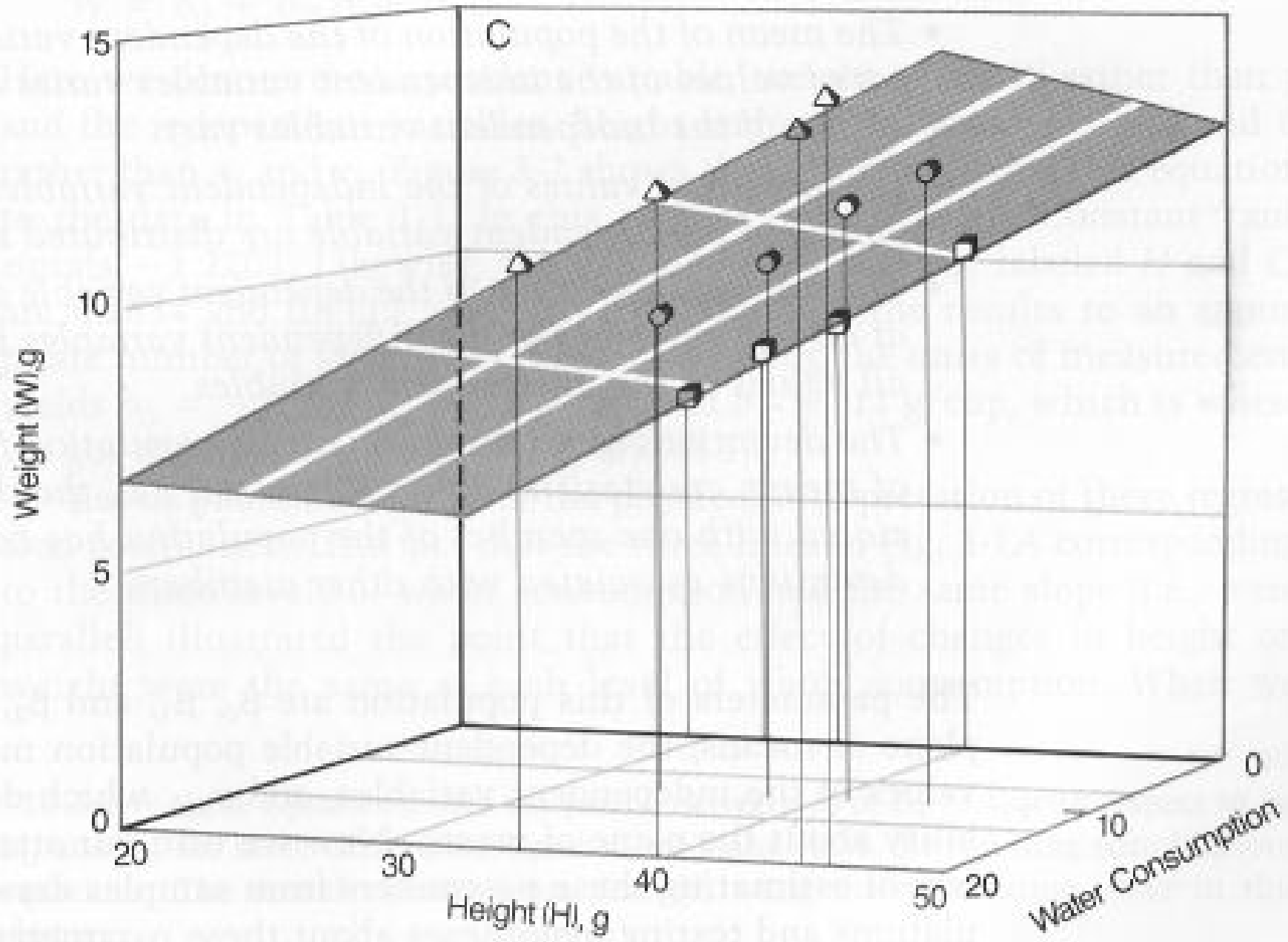


FIGURE 3-1 (continued)

A regressziós egyenes egyenlete (folyt.)

- A függő változó mindig folyamatos, a független változó lehet folyamatos és nominális**

a kétféle értékű nominális változók kódolása: 0 vagy 1 (DUMMY VARIABLE)

**pl. kontroll: 0, beteg: 1,
Chl. pn. neg: 0, poz: 1**

A többszörös regresszió eredményeinek interpretálása

- A beta regressziós koefficiens: többszörös regresszió esetében ez az jelenti, hogy **ha a többi független változó értéke állandó**, akkor a vizsgált független változó egy egységnyi változásának a függő változó milyen mértékű változása felel meg.
- $\text{PI log(anti-hsp65 AU/ml)} = 0.213 \text{ csoport} - 0.018 \text{ mmol/l HDL-chol} + 0.052 \text{ mmol/l trigl} + 0.03 \text{ Chl. pneumoniae} + 1.65$
- Tehát a 0-ról 1 egységre való növelés (kontrollról betegre) a log-anti-hsp60 szintet 0.213-al növeli. A 0.213 antilogja: 1.63, tehát a betegek anti-hsp60 szintje átlagosan 1.63 –szor magasabb lenne akkor, ha nem lenne a kontrollok és a betegek között különbség a HDL koleszterin, a triglicerid szintben, ill. a Chl. pneumoniae pozitívitás %-ában (**észlelt érték: 102 vs. 57, $102/57=1,79$**).

A regressziós koefficiens szignifikanciája

a koefficiens szignifikanciája kiszámítható

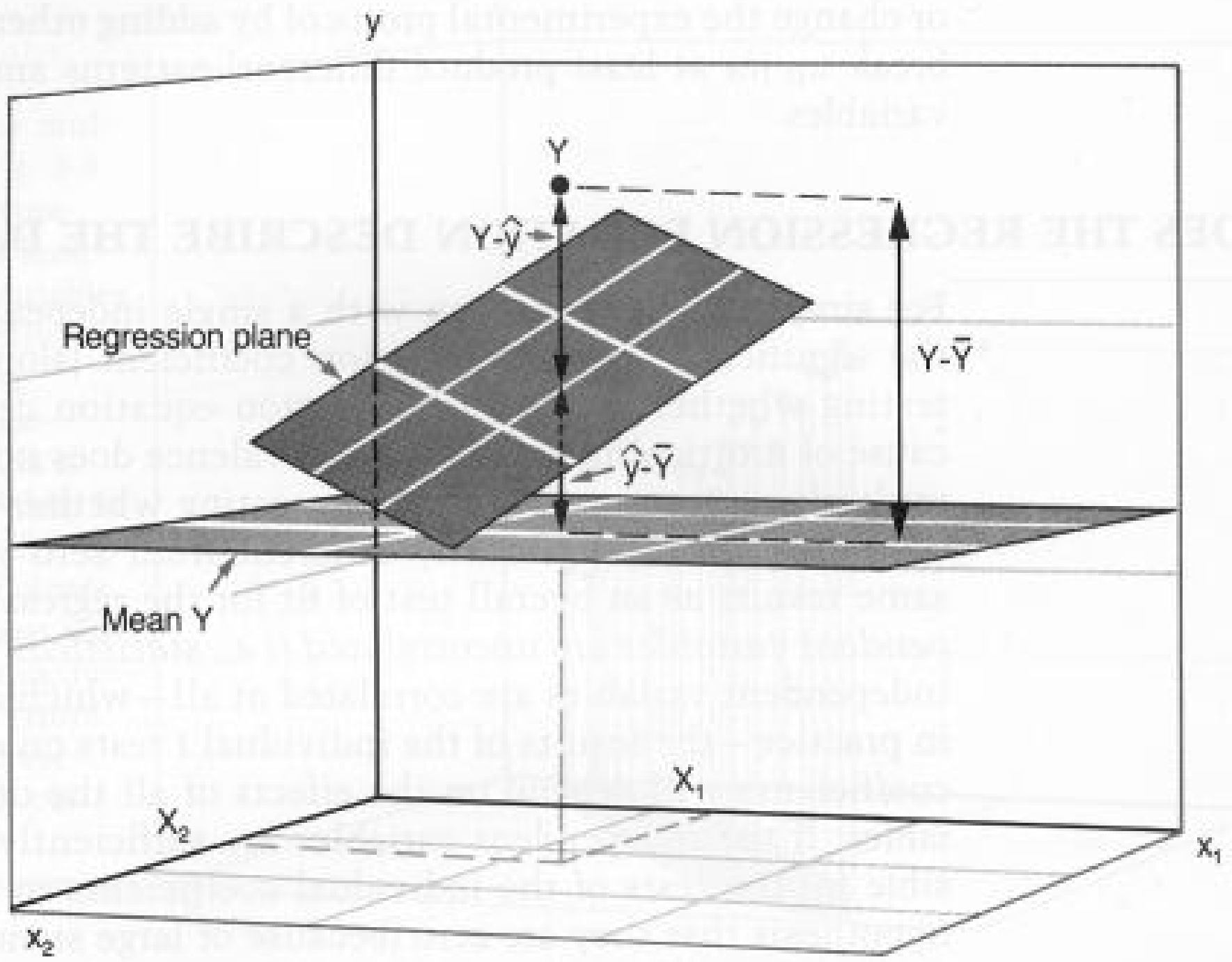
- t teszttel
t teszt: a **b** regr. koeff. értéke osztva ennek S.E.-jével, a megfelelő df-nél t táblázatban keresem (keresi a gép) az értéket.
- Standardizált regressziós koefficiens: **beta**: a változó minden értékéből levonjuk az X átlagértékét és elosztjuk a SD-val, így az átlag: 0, a SD: 1 lesz. Ekkor a regressziós koefficiensek összehasonlíthatók, az van nagyobb hatással a függő változóra, amelyik nagyobb.

Az R^2 érték többszörös regressziónál

- **Akár az egyedi, az egyenletbe bevett változóra, akár ennek egy részére vagy az összesre vonatkozóan az R^2 érték azt mutatja, hogy az adott független változó(k) hány százalékban határozzák meg a független változót. Ha az R^2 érték:1,00, akkor teljes mértékben, ha 0.00, akkor egyáltalán nem, ha 0.50. akkor erősen.**
- **Példánkban a 4 változó (csoport, HDL-chol, trig, Chl.pneum) együttesen 0.0526 (Statistica), R^2 értéket ad, tehát a négy tényező igen gyengén határozza meg a természetes anti-hsp60 antitestek titerét. Szakmailag O.K.**

Kapcsolat a többszörös regresszió és a variancia analízis között

- $R = \sqrt{1 - (SS_{reg}/SS_{tot})}$ és
- $SS_{tot} = SS_{reg} + SS_{res}$, ezért
- $R^2 = 1 - (SS_{res}/SS_{tot}) = 1 - (SS_{tot} - SS_{reg})/SS_{tot} = 1 - 1 + SS_{reg}/SS_{tot} = SS_{reg}/SS_{tot}$
- ennek szignifikanciáját az F eloszlás szerint határozzuk meg (variancia analízis).
- Az adjusztált R^2 figyelembe veszi a több változó egyenletbe vitelekor bekövetkező szabadságfok csökkenést. Példánkban (SPSS), az R^2 : 0.073, az adjusztált R^2 ehhez igen hasonló: 0.052



A lépcsőzetes többszörös regresszió (stepwise multiple regression)

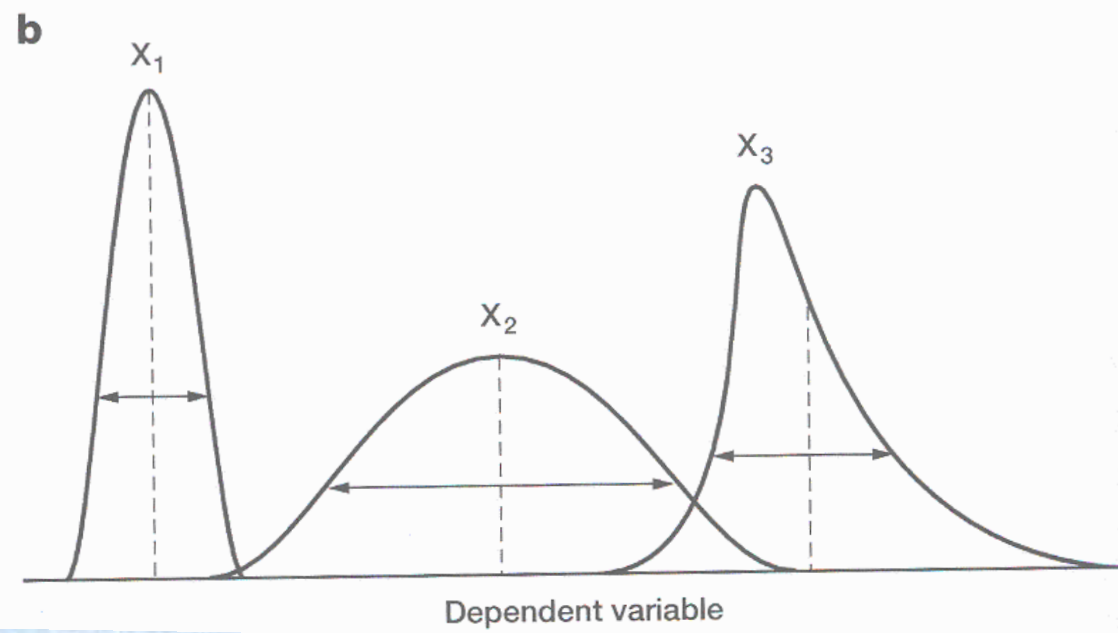
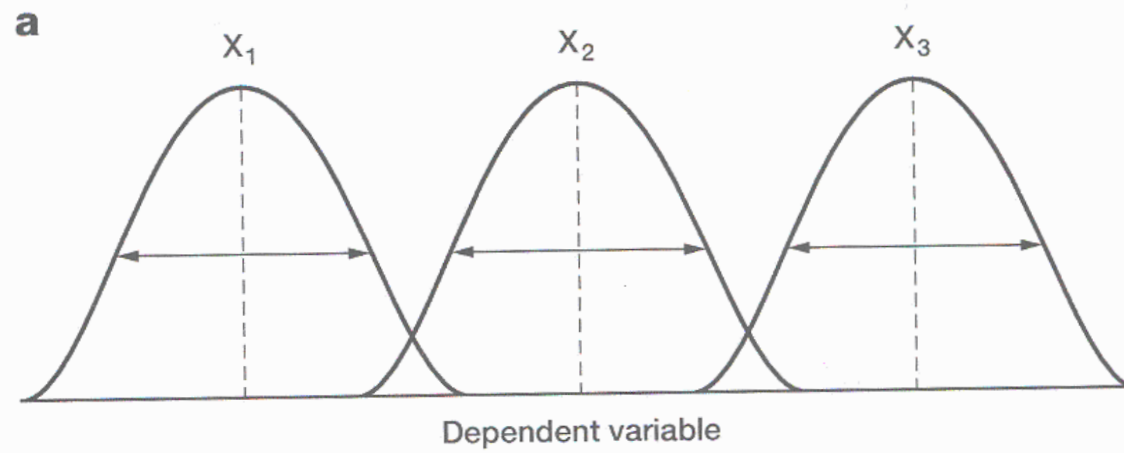
- A cél: minél jobb, a függő változót minél jobban előrejelző modelt építsünk fel: legegyszerűbb mód: minden szakmailag értelmes változót figyelembe veszünk, kiszámítjuk a b értékeket, majd azokat, melyek nem szignifikánsak, kihagyjuk és újra számolunk. Ha jól dolgoztunk, akkor az egyes változókhoz tartozó R^2 értékeknek nőnie kell.**
- A módszert automatikusan is el lehet végezni, ennek három módja a forward selection, a backward elimination és a stepwise regression**

A többszörös lineáris regressziós analízis esetében leggyakrabban felmerülő problémák és ezek megoldása

- A függő változó normális eloszlású kell legyen és az átlag körül azonos varianciának kell fennállnia
- Milyen független változókat enged meg a számítás?
- A független változók közötti túl szoros kapcsolat, a multicollinearitás problémája
- Hiányzó értékek
- Kiugró (outlier) értékek
- Mintaszám követelmények

A függő változó normális eloszlású kell legyen és az átlag körül azonos varianciának kell fennállnia

- Tehát a független változó bármelyik értékét is vesszük (p. életkor: 20-29, 30-39, 40-49 év), az ehhez tartozó függő változó értékek Gauss görbét kell, hogy adjanak
- Azonos variancia: a független változó minden értékénél mért függő változó értékeknek az átlagtól való távolsága azonos kell legyen (homoscedasticitás=homogén eloszlás)
- A kettő eltérése gyakran, de nem mindig együtt jár
- Mindig rajzoljunk!!!

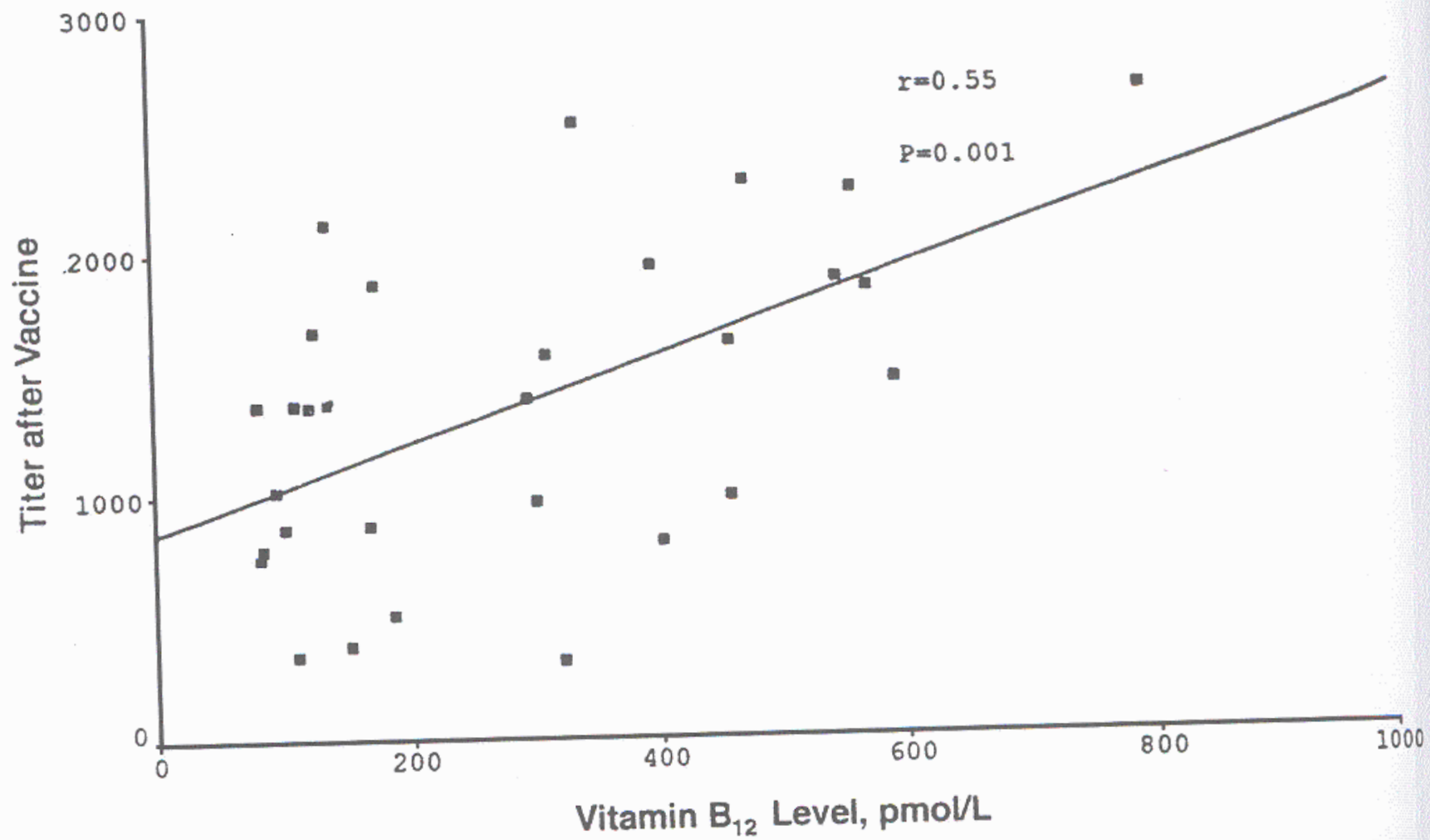


Hogyan vizsgáljuk meg a normál eloszlást?

- Rajz, statisztikai programok: összehasonlítás a normál eloszlással
- Statisztikai tesztek (pl. Kolmogorov-Smirnov)
- Nagy elemszámnál (>100) feltételezhető a normális eloszlás, ha nagyon eltérő kilógó értékek nincsenek közöttük

Milyen független változókat enged meg a számítás?

- Fő követelmény: a független változókkal lineárisan változik a függő változó átlaga is.
- Mit jelent ez a többszörös lineáris regresszió esetében? : A függő változó átlaga lineárisan változik a független változók súlyozott átlagával: súly: milyen erős az adott független változó kapcsolata a függő változóval?



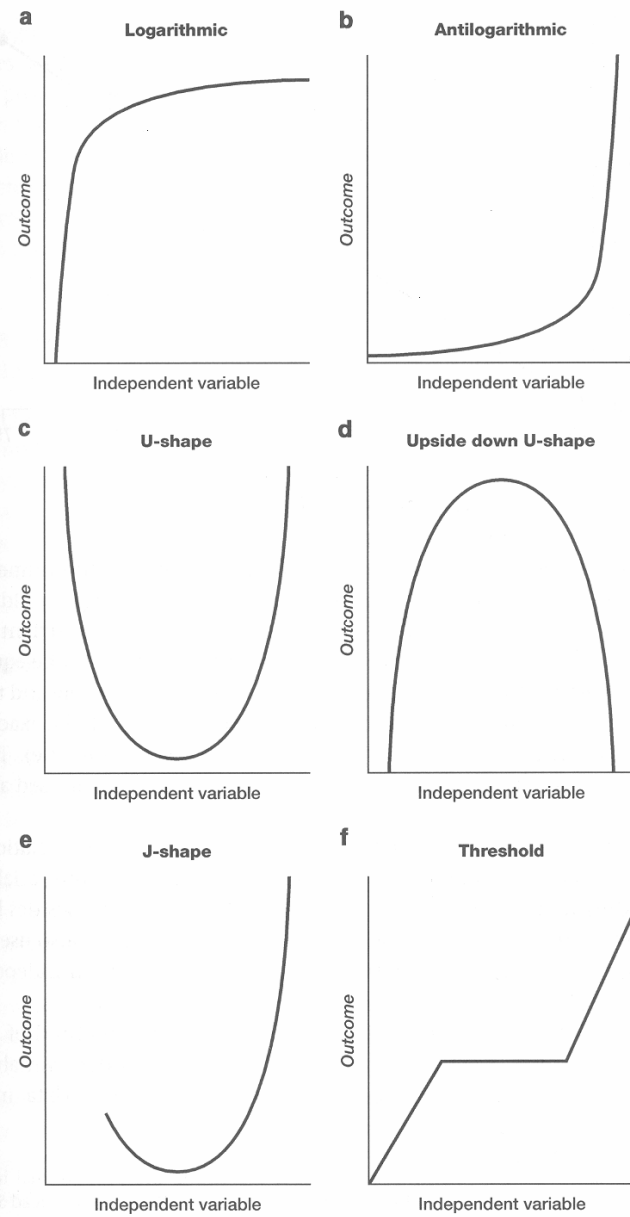
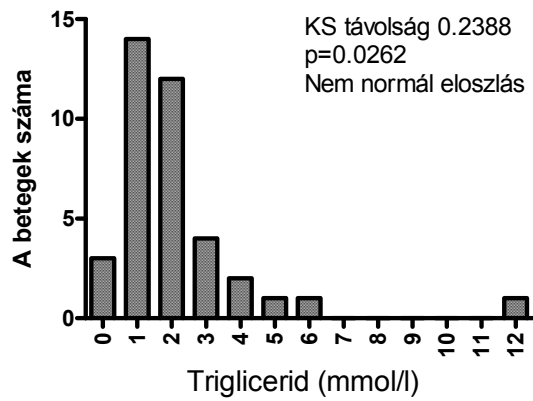
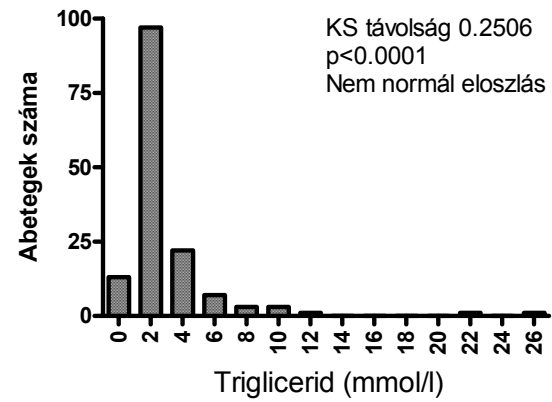


Figure 5.3. Variety of nonlinear relationships between an independent variable and an outcome.

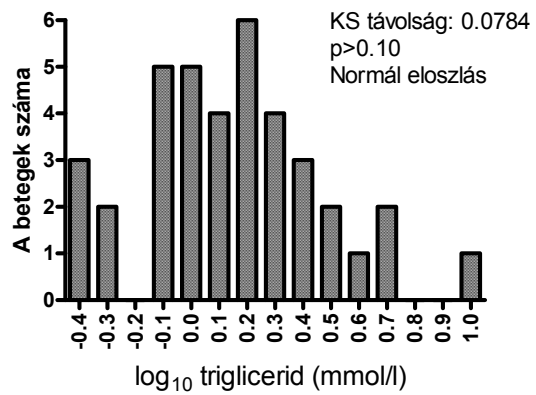
I-es típusú cukorbeteg, n=38



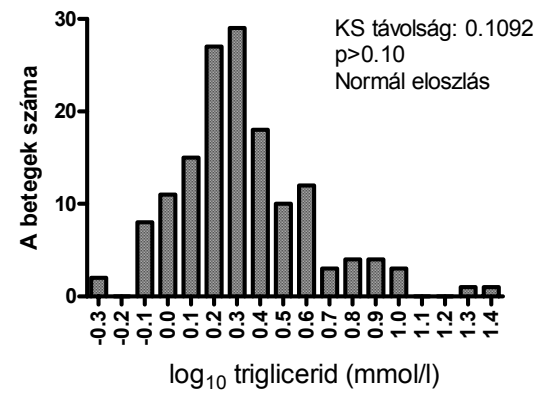
II-es típusú cukorbeteg, n=148



I-es típusú cukorbeteg, n=38



II-es típusú cukorbeteg, n=148



Milyen típusú független változók szerepelhetnek és melyek nem szerepelhetnek?

TABLE 4.1

Independent variables and multivariable analysis.

Type of independent variable	Example of independent variable	Multiple linear regression	Multiple logistic regression	Proportional hazards analysis
Interval	Age, blood pressure	Yes	Yes	Yes
Dichotomous	Gender	Yes	Yes	Yes
Ordinal	Cancer stage	No	No	No
Nominal	Ethnicity	No	No	No

Megoldás: többszöri dichotóm (igen/nem) változók képzése

TABLE 4.2

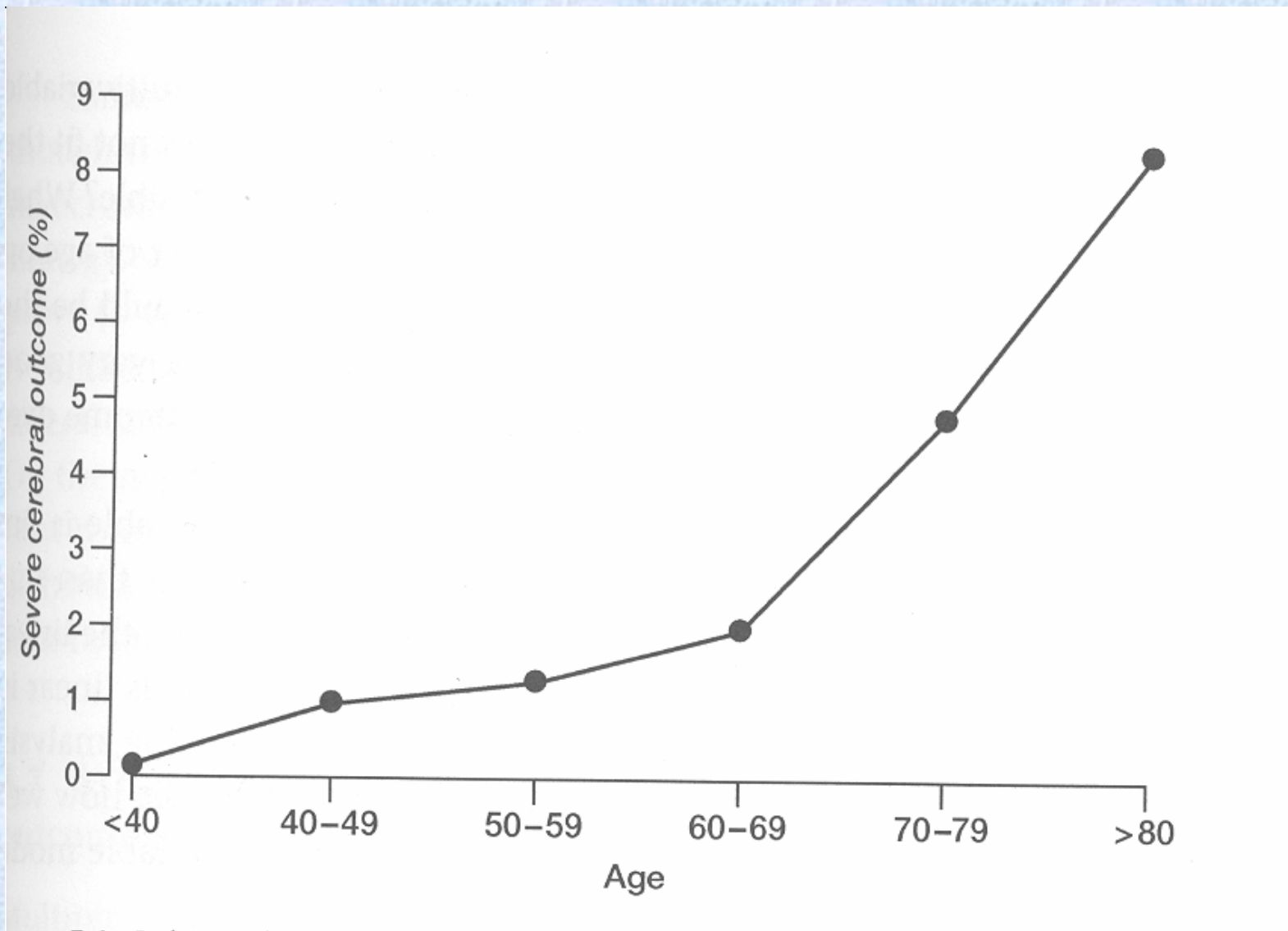
Creation of multiple dichotomous variables to represent a nominal independent variable

	African- American	Latino/ Hispanic	Asian/Pacific Islander	Native American	Other nonwhite
African-American	1	0	0	0	0
Latino/Hispanic	0	1	0	0	0
Asian/Pacific Islander	0	0	1	0	0
Native American	0	0	0	1	0
Other nonwhite	0	0	0	0	1
White/Caucasian	0	0	0	0	0

A dichotomizálás fő szabályai

- Annak elkerülésére, hogy az egész populációra nézve túl ne értékeljük az összefüggést
vagy természetes határokat veszünk (öreg/nem öreg (60 év))
vagy a független változó percentiliseit (tercilis, kvartilis, stb). Ebben az esetben a csoportok elemszáma azonos lesz

Szakmai probléma: a csoportban elég nagy elemszám legyen, ne legyen túl sok kiugró érték



Relatív kockázat a <40 évhez viszonyítva: 1.75/évtized

A többszörös lineáris regressziót legjobban torzító hiba: a multicollinearitás

Ha az egyes független változók erős ($R > 0.90$) korrelációt mutatnak egymással, akkor a modell erősen torzulhat (redundáns információk). Pl. vérnyomás előrejelzése az életkor, a testsúly és a testmagasság alapján. De a testsúly és a testmagasság erősen korrelál egymással. Nem biztos, hogy az automata szelekciónál nem marad-e bent mind a kettő. Előtte meg kell nézni, egyiket nem bevenni a modellbe!

További példák a multicollinearitásra, mikor okoz problémát, mit lehet tenni?

- Láz Celsius és Fahrenheit fokokban: hibaüzenet u.a változó
- Születési súly és a terhesség hossza
- A határok, ahol problémát okoz: $R < 0.8$: nincs probléma
- $R 0.8-0.9$ problémát okozhat
- $R > 0.9$ biztosan problémát okoz.
- Előzetes nem-paraméteres korreláció vizsgálat (Spearman)
- Számítógépes program: többszöri regressziós számításnál: korrelációs matrix (adjusztált matrix, pontosabb, 0.8, 0.9 itt is)

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)
szolhsphl	-,635	,252	6,337	1	,012	,530
Nem	,373	,248	2,255	1	,133	1,452
Életkorfelfedezés	,022	,014	2,502	1	,114	1,022
TNMT	,455	,216	4,451	1	,035	1,576

Correlation Matrix of Regression Coefficients

	szolhsphl	Nem	Életkorfelfedezés
Nem	-,061		
Életkorfelfedezés	,092	-,081	
TNMT	-,050	-,024	-,085

Bonyolultabb eset: több változó egyértelműen meghatároz egy további

TABLE 4.2

Creation of multiple dichotomous variables to represent a nominal independent variable

	African-American	Latino/Hispanic	Asian/Pacific Islander	Native American	Other nonwhite
African-American	1	0	0	0	0
Latino/Hispanic	0	1	0	0	0
Asian/Pacific Islander	0	0	1	0	0
Native American	0	0	0	1	0
Other nonwhite	0	0	0	0	1
White/Caucasian	0	0	0	0	0

A multicollinearitás mérőszáma a VIF

- A multicollinearitásból eredő redundáns információ csökkenti a független és függő változók között megállapított összefüggés pontosságát.
- Mérőszáma a variancia inflációs faktor (VIF) = $1/(1-R^2_i)$, hogyan torzít egy x_1 -el való összefüggését az y -al, ha a többi x az x_1 -el redundáns információkat tartalmaz
- Ha nincs ilyen akkor az $R^2 = 0$ és a $VIF = 1$. Ha R^2 nem = 0, akkor a $VIF > 1$, ha ez nagyon nagy akkor a regressziós koefficiens extrém módon eltérhet a valódi összefüggéstől. $VIF > 4$, vizsgálni kell, $VIF > 10$ biztosan értékelhetetlen eredmények

Coefficients^a

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
	B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1 (Constant)	55,311	8,117		6,814	,000					
szolhsp70	-,298	,854	-,030	-,349	,728	-,098	-,032	-,029	,961	1,041
Nem	,253	2,128	,010	,119	,906	-,061	,011	,010	,964	1,037
Életkor felfedezéskor	2,540	1,049	2,124	2,421	,017	-,126	,214	,202	,009	110,095
Életkor 0.mintavételkor	-2,670	1,044	-2,246	-2,556	,012	-,146	-,225	-,214	,009	110,390
TNM-T	-5,483	1,604	-,292	-3,418	,001	-,293	-,296	-,286	,958	1,044

a. Dependent Variable: survmonth

Mintaszám követelmények

- **Ma már erre számos komputeres program alkalmas, de van megközelítő szabály:
legalább 10-szer annyi megfigyelés (személy, állat, stb) legyen, mint ahány változó. Másrészt egy változónál minimálisan 5, de inkább 10 megfigyelés történjen.**

Specifikus szabályok a többszörös lineáris regresszióra

- Először el kell végezni az összefüggés számítást egyenként (kétváltozós analízis) pl. 3 év alatt a vizsgálatba bevonandó x emberből várhatóan hány hal meg, és annak a faktornak az értéke között, amelynek a befolyását vizsgálni szeretnénk Ha a számítás szerint az elemszám nem elég, akkor a többszörös elemzésnél se várható eredmény
- Egyszerűsített szabály: minden független változóra legalább 20 egyén (állat, minta) kell. Nagy S.E: hibalehetőség. <20 egyén, stb: óvatos értékelés.

Mit csináljunk, ha egy-egy független változóra nem elég az esetszám? I. Egy vagy több változó kihagyása

- Elméleti megfontolások
- Mérési megfontolások
 - Két változó erősen korrelál (még nincs collinearitás). Hagyjuk ki az egyiket, azt
 - Amelyikben több a hiányzó adat
 - Nagyobb a mérési hiba
 - Elméletleg kevésbé fontos
- Empirikus megfigyelések
 - A változó nem mutat összefüggést a kimenetellel a kétváltozós elemzés során
 - A változó nem mutat összefüggést a kimenetellel a többváltozós elemzés során
 - A változó nem függ össze a legfontosabb független változóval
 - A változó csak minimális mértékben befolyásolja a hatást a többváltozós elemzés során

Mit csináljunk, ha egy-egy független változóra nem elég az esetszám? II- Kombináljuk a változókat egyetlen változóba vagy skálába

- És/vagy konstrukciók
 - Pl. 3 változó, ha mindhárom vagy 2 vagy csak 1 is fennáll, akkor igen, ha egyik sem, akkor nem
- Összeadási skálák
 - Minden hasonló választ mutató változót u.a. skálán bejelölünk (pl. 1-10) majd az értékeket összeadjuk majd elosztjuk az összeadott változók számával: átlag válasz. vagy átlag score Ez csak erősen ($R > 0.65$) korreláló változók esetében hajtható végre
- Faktor analízis
 - Több, egymással szoros viszonyt mutató független változót összevon (pl. 15-t 2-3-ba. Klinikai vizsgálatnál szakmailag problematikus, az eredeti változók elvesznek.

Az analízis megkezdésével kapcsolatos problémák I. Hogyan jelöljük az igen nem változót?

- 0-1, 1-2, 3-2-3 OK, -1-0 is OK de -1-1 vagy 1-3 nem OK
- Konvenció 0: nem történt meg az esemény
- 1: megtörtént az esemény
 - Előnye: nem keverem össze a dolgokat
 - Az átlag a prevalenciát mutatja (100 ember 10 esemény: az átlag 0.10)

Az analízis megkezdésével kapcsolatos problémák II. Melyik legyen a referencia csoport?

TABLE 8.1

Implications of changing the reference group for dichotomous variables.

	Odds ratio	Odds ratio
White/Caucasian	1.0 (reference)	4.0
African-American	0.25	1.0 (reference)
Latino	0.50	2.0
Asian/Pacific Islander	0.50	2.0
Native American	0.25	1.0
Other nonwhite ethnicity	0.50	2.0

Az analízis megkezdésével kapcsolatos problémák III A hiányzó adatok

- Hagyjuk ki azt a változót, amelyikben hiányzó adatok vannak
- Igen/nem változót készítünk a hiányzó adatok pótlására
- Próbáljunk meg újabb adatokat szerezni
- Csökkentsük a független változók számát az analízisben
- Becsüljük meg a hiányzó adatokat

A hiányzó adatok pótlásának módjai becsléssel

- A minta átlagát írjuk be
- Egy alcsoport átlagát írjuk be
- A hiányzó adatot egy másik kovariáns alapján modellezzük (egyszerű imputáció)
- A hiányzó adatot egy másik kovariáns alapján modellezzük, de egy random komponenst is beleveszünk az analízisbe (többszörös imputáció)

Az automatikus regressziós model építés három fő módszere

- **forward selection:** először egyetlen változót visz a program be az egyenletbe, azt, amelyiknek a legnagyobb a st. regr koefficiense, a következőnél megvizsgálja a program: szignifikánsan (F-teszt) növeli-e az R^2 értéket. Akkor van vége, ha nincs több ilyen változó.
- **backward elimination:** először minden változó bekerül a modelbe, majd lépésről lépésre eleminálja a program azokat a változókat, amelynél ez az elinináció az R^2 értéket nem csökkenti szignifikánsan.
- **stepwise regression (selection):** úgy kezdődik, mint a forward selection, de minden új változó beépítése után megvizsgálja a program, hogy a már beépített változók közül melyik eliminálható úgy, hogy az R^2 érték ne csökkenjen

A model megfelelőségének (fit) mérése a többszörös lineáris regressziós vizsgálatkor

Variables Entered/Removed^a

Model	Variables Entered	Variables Removed	Method
1	TMN-M, Életkor felfedezés _a kor, Nem	.	Enter

a. All requested variables entered.

b. Dependent Variable: survmonth

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,347 ^a	,120	,103	15,724

a. Predictors: (Constant), TMN-M, Életkor felfedezés_akor, Nem

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5246,296	3	1748,765	7,073	,000 ^a
	Residual	38322,773	155	247,244		
	Total	43569,069	158			

a. Predictors: (Constant), TMN-M, Életkor felfedezés_akor, Nem

b. Dependent Variable: survmonth

Hogyan interpretáljuk a többszörös lineáris regressziós elemzés eredményeit?

- Ha a regressziós koefficiens pozitív, akkor a független változó változásával párhuzamosan nő a függő változó
- Ha a regressziós koefficiens negatív, akkor a független változó változásával párhuzamosan csökken a függő változó

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	53,428	8,734		6,117	,000	36,175	70,680
	Nem	-1,544	2,559	-,046	-,603	,547	-6,599	3,511
	Életkor felfedezéskor	-,238	,123	-,147	-1,941	,054	-,481	,004
	TMN-M	-11,957	3,166	-,290	-3,777	,000	-18,211	-5,704

a. Dependent Variable: survmonth

Y (túlélés, hónap) = - 0.238 x életkor (év), 10 év különbség $-0.238 \times 10 = 2.38$ hónappal rövidebb túlélés

TNM-M 0 (nincs metastasis) vagy 1 (van metastasis)

Ha TNM-M 0-ról 1-re nő akkor $-11.957 \times 1 = -11.957$, tehát kb 12 hónappal rövidebb lesz a túlélés

1. feladat: az ólomkoncentráció és a kreatinin klírens (Stassen et al, NEJM, 327151, 1992)

- **Y: kreatinin klírens**
- **X1: log vér ólom koncentráció,**
- **X2: életkor,**
- **X3: BMI**
- **X4: log SGOT**
- **X5: használt-e diureticumot: 0: nem, 1: igen**

a regressziós koefficiens (b) a log ólom koncentrációra -9.5 ml/perc volt (CI: -18.1 - -0.9 ml/perc)

Kérdések: 1. szignifikáns volt-e a b érték?

2. hogyan függött össze az ólomkoncentráció a kreatinin klirensszel, ha az összes többi változó nem befolyásolhatta ezt?

Válaszok az 1. feladatra

- 1. igen: CI: -18.1 - -0.9 ml/perc, nincs közöttte 0**
- 2. ha a szérum ólomtartalma 1 egységgel nő (log érték: tehát tízszeresére), akkor a kreatinin klírens sz 9.5 ml/perccel csökken**

2. feladat: Feher et al. Beta blockers, lipoproteins and a non-insulin dependent diabetes (Postgrad. Med. 64, 927, 1988)

- Y (H): HDL2 szubfrakció
- X1 (B): beta blokkolót szedett 1: igen, 2: nem
- X2 (D).drink 1: alkoholt fogyasztott, 2: nem
- X3 (S) smoking 1: dohányzik, 0: nem
- X4 (A): életkor, év
- X5 (W) testsúly
- X6 (T) trigliceridek
- X7 (C) C-peptide
- X8 (G) vércukor

$H = 0.711 - 0.0824 B - 0.0173 D - 0.0399 S - 0.00455 A - 0.00214 W - 0.0444 T + 0.00463 C - 0.00391 G$. $R^2: 59.5\%$, $adj.R^2: 54.3\%$

Kérdés: mit jelentenek a piros számok?

Válaszok az 2. feladatra

- **-0.0824 B:** ha béta-blokkolót szed valaki, akkor a HDL2 (védő) frakció szérumszint **0.0824 mmol/l-el** kisebb lesz
- **-0.00455 A:** az öregedéssel párhuzamosan évente **0.00455 mmol/l-el** csökken a HDL2-frakció szérumszintje
- **-0.0444 T:** 1 mmol/l triglicerid szint növekedés **0.044 mmol/l HDL-csökkenéssel** jár együtt
- **+ 0.00463 C:** azoknak, akiknek magasabb a C peptid szintjük, as HDL2 koleszterin szintjük is magasabb, minden egységgel (ng/ml) **0.00463 mmol/l-el**