

i	Pohl	Vyska	Hmotnost	Vi - mV	Hi - mH	(Vi-mV)(Hi-H.stř.i	ei
1	0	172	86	-2,5	11,04545	-27,61364	81,2244 4,775604
2	1	169	58	-5,5	-16,95455	93,25	64,92978 -6,929785
3	0	183	80	8,5	5,045455	42,886364	85,97131 -5,971305
4	1	170	69	-4,5	-5,954545	26,795455	65,36132 3,638678
5	0	180	85	5,5	10,04545	55,25	84,67669 0,323306
6	1	173	76	-1,5	1,045455	-1,568182	66,65593 9,344066
7	0	190	89	15,5	14,04545	217,70455	88,99207 0,007934
8	1	174	62	-0,5	-12,95455	6,4772727	67,08747 -5,087471
9	1	160	55	-14,5	-19,95455	289,34091	61,04595 -6,04595
10	0	182	75	7,5	0,045455	0,3409091	85,53977 -10,53977
11	0	198	101	23,5	26,04545	612,06818	92,44436 8,555637
12	1	153	48	-21,5	-26,95455	579,52273	58,02519 -10,02519
13	1	174	65	-0,5	-9,954545	4,9772727	67,08747 -2,087471
14	1	162	76	-12,5	1,045455	-13,06818	61,90902 14,09098
15	0	171	69	-3,5	-5,954545	20,840909	80,79286 -11,79286
16	1	159	48	-15,5	-26,95455	417,79545	60,61441 -12,61441
17	0	192	78	17,5	3,045455	53,295455	89,85514 -11,85514
18	1	170	59	-4,5	-15,95455	71,795455	65,36132 -6,361322
19	1	181	76	6,5	1,045455	6,7954545	70,10823 5,891769
20	0	179	95	4,5	20,04545	90,204545	84,24516 10,75484
21	0	165	101	-9,5	26,04545	-247,4318	78,20364 22,79636
22	0	182	98	7,5	23,04545	172,84091	85,53977 12,46023

m 0,5 174,5 74,95455 vzorcem 0,6516525
sd - vzorec 11,18566 16,15248 funkcí 0,6516525
sd - funkcí 0,511766 11,18566 16,15248

rPV -0,619731
rPH -0,763283
rVH 0,651653

a 7
b1 0,4315372
b2 -15

R 0,7954403
R2 0,6327252

SS 2061,7969

rHe 0,7214391
rVe 0,0950207

korelace

$$r_{XY} = \frac{1}{(N-1)} \text{SUMA}(i=1 \dots N) \left\{ \frac{(x_i - m_x)(y_i - m_y)}{s_x s_y} \right\}$$

úprava vzorce korelace

$$= \frac{1}{(N-1)} \text{SUMA}(i=1 \dots N) [z_{xi} z_{yi}]$$

průměr

$$m_X = (1/N) \sum x_i$$

rozptyl (variance)

$$s_x^2 = [1/(N-1)] \sum (x_i - m_x)^2$$

směrodatí

s = odmoc

kovariance

$c_{XY} = s_X r_{XY} s_Y$ z toho

$$r_{XY} = c_{XY} / (s_X s_Y)$$

Vícenásobná regrese (v tomto případě tří proměnných)

Cíl: předpovědět hodnoty závislé proměnné (zde Y)

Regresní rovnice:

$$Y = a + b_1 X_1 + b_2 X_2 + \dots + b_n X_n$$

nebo $\hat{Y} = a + b_1 X_1 + b_2 X_2 + \dots + b_n X_n$

Y se stříškou: Hodnota proměnné
 a: Tzv. regresní konstanta, anglicky intercept
 b1 a b2: Tzv. regresní koeficient

R^2 : tzv. mnohonásobná korelace

Postup

V případě, že máme v regresi více než dvě proměnné
 Nastupuje řešení pomocí metod tzv. numerické metody
 Zkusíme odhadnout regresní model $H = a + b_1 V_1 + b_2 V_2 + \dots + b_n V_n$

1. Předstíráme, že známe hodnoty a, b1 a b2 - vle
2. Vypočítáme H se stříškou pro hodnoty pohlaví
3. Musíme si stanovit nějaká kritéria kvality toho modelu
 1. bude to
 2. další bu
4. Měníme, přepisujeme hodnoty a, b1 a b2 tak, a

) Suma(1...N) xi

ariance)

-1)] Suma(1...N) (xi - mx)²

ná odchylka

nina(s²)

iěnných)

ř) pomocí hodnot nezávislých proměnných (zde X1 a X2)

2+?

2 pro každou individuální hodnotu

ie Y předpovězená pomocí hodnot X1 a X2 a regresních koeficientů a, b1 a b2

cky intercept, v podstatě hodnota proměnné Y pro X = 0

y, které vyjadřují, o kolik vzroste hodnota závislé proměnné (zde Y), pokud se hodnota nezávislých

se rovná korelaci hodnot Y a hodnot Y předpovězených pomocí regresní rovnice (Y se stříškou)

nné, **NELZE** koeficienty najít analyticky na základě vzorců (tzn. neexistují žádné vzorce pro výpočet matematiky. Abychom si prakticky vyzkoušeli, jak tyto metody mohou postupovat, vyzkoušíme si to + b2 P (tzn. odhadneme hmotnost na základě výšky a pohlaví).

ivo v příslušných políčkách jsem napsal nějaké odhady hodnot, které jsem určil zcela "od oka"

a výšky každého respondenta a tyto "od oka" stanovené hodnoty a, b1 a b2

odelu

jednak již známý koeficient R (potažmo R²), tzn. korelace H a H se stříškou

čím vyšší, tím lepší model (odhad hmotnosti na základě výšky a pohlaví)
de tzv. suma čtverců reziduí, neboli chyb odhadu, neboli hodnot ei (značí se často jako SS - z angličtiny)

čím nižší, tím lepší model

bychom dosáhli co nejvyšší hodnoty R (nebo R²) a nejnižší hodnoty SS

h proměnných (zde X1 a X2) změni o 1

čet a, b1 a b2)
formou hry.

ví
jl. Sum of Squares)

i	Pohlaví (P)	Výška (V)	Hmotnost (H)	H.stř.i.zV	ei.HzV	P.stř.i.zV	ei.PzV
1	0	172	86	72,60202	13,39798	0,5708849	-0,570885
2	1	169	58	69,779	-11,779	0,6559467	0,344053
3	0	183	80	82,95312	-2,953118	0,2589914	-0,258991
4	1	170	69	70,72001	-1,720007	0,6275928	0,372407
5	0	180	85	80,13009	4,869907	0,3440533	-0,344053
6	1	173	76	73,54303	2,456967	0,5425309	0,457469
7	0	190	89	89,54018	-0,540178	0,0605138	-0,060514
8	1	174	62	74,48404	-12,48404	0,514177	0,485823
9	1	160	55	61,30992	-6,309921	0,9111323	0,088868
10	0	182	75	82,01211	-7,01211	0,2873454	-0,287345
11	0	198	101	97,06825	3,931753	-0,166318	0,166318
12	1	153	48	54,72286	-6,722861	1,1096099	-0,10961
13	1	174	65	74,48404	-9,484041	0,514177	0,485823
14	1	162	76	63,19194	12,80806	0,8544244	0,145576
15	0	171	69	71,66102	-2,661015	0,5992388	-0,599239
16	1	159	48	60,36891	-12,36891	0,9394862	0,060514
17	0	192	78	91,4222	-13,4222	0,0038059	-0,003806
18	1	170	59	70,72001	-11,72001	0,6275928	0,372407
19	1	181	76	81,0711	-5,071101	0,3156993	0,684301
20	0	179	95	79,18908	15,81092	0,3724072	-0,372407
21	0	165	101	66,01496	34,98504	0,7693625	-0,769363
22	0	182	98	82,01211	15,98789	0,2873454	-0,287345

m	0,5	174,5	74,954545455	H predikovaná z V	P predikované z V
sd	0,5117663	11,18566	16,152479605	r	0,651653 r
					-0,619731
rPV	-0,61973			a	-89,25145 a
rPH	-0,76328			b	0,941009 b
rVH	0,651653				5,447764
					-0,028354

Parciální korelace PH.V

vzorcem -0,60379
korelace e -0,60379

H.stř.i.zP	ei.HzP	V.stř.i.zP	ei.VzP	V.stř.i.zH	ei.VzH	P.stř.i.zH	ei.PzH
87	-1	181,2727	-9,272727	179,4845	-7,484507	0,232883	-0,232883
62,90909	-4,909091	167,7273	1,272727	166,8489	2,151115	0,910019	0,089981
87	-7	181,2727	1,727273	176,7769	6,223127	0,377984	-0,377984
62,90909	6,090909	167,7273	2,272727	171,8129	-1,812879	0,644001	0,355999
87	-2	181,2727	-1,272727	179,0332	0,966765	0,257066	-0,257066
62,90909	13,09091	167,7273	5,272727	174,9718	-1,971785	0,474717	0,525283
87	2	181,2727	8,727273	180,8383	9,161676	0,160333	-0,160333
62,90909	-0,909091	167,7273	6,272727	168,654	5,346027	0,813286	0,186714
62,90909	-7,909091	167,7273	-7,727273	165,4951	-5,495068	0,98257	0,01743
87	-12	181,2727	0,727273	174,5205	7,479488	0,498901	-0,498901
87	14	181,2727	16,72727	186,2536	11,74641	-0,129869	0,129869
62,90909	-14,90909	167,7273	-14,72727	162,3362	-9,336162	1,151854	-0,151854
62,90909	2,090909	167,7273	6,272727	170,0078	3,99221	0,740735	0,259265
62,90909	13,09091	167,7273	-5,727273	174,9718	-12,97178	0,474717	0,525283
87	-18	181,2727	-10,27273	171,8129	-0,812879	0,644001	-0,644001
62,90909	-14,90909	167,7273	-8,727273	162,3362	-3,336162	1,151854	-0,151854
87	-9	181,2727	10,72727	175,8743	16,12567	0,42635	-0,42635
62,90909	-3,909091	167,7273	2,272727	167,3002	2,699843	0,885836	0,114164
62,90909	13,09091	167,7273	13,27273	174,9718	6,028215	0,474717	0,525283
87	8	181,2727	-2,272727	183,546	-4,545957	0,015232	-0,015232
87	14	181,2727	-16,27273	186,2536	-21,25359	-0,129869	0,129869
87	11	181,2727	0,727273	184,8998	-2,899774	-0,057318	0,057318

?=?+?_1

^_?=?+?_1

H predikovaná z P V predikovaná z P V predikovaná z H P predikované z H

r	-0,763283 r	-0,619731 r	0,651653 r	-0,763283
a	87 a	181,2727 a	140,6751 a	2,312659
b	-24,09091 b	-13,54545 b	0,451272 b	-0,024183

Parciální korelace VH.P

Parciální korelace PV.H

vzorcem 0,352284
korelace e 0,352284

vzorcem -0,249639
korelace e -0,249639

?_1+?_2 ?_2+?

?_?1+?_2 ?_?2