

U samom startu važno je definisati razlike u značenju dva pojma: korelacija i regresija. Pod **korelacijom** se podrazumeva prisustvo uzročno-posledične veze između dve veličine X i Y. Promena jedne veličine može dovesti do porasta odnosno opadanja druge. Ova veza je kvantifikovana tzv. Korelacionim koeficijentom r .

Zbog svega toga kažemo da su ove dve veličine u međusobnoj zavisnosti. Pri tome se obično podrazumeva linearna zavisnost, a postupak utvrđivanja date linearne krive (prave) naziva se **regresijom**.

$$y = a + bx$$

- Koeficijent korelacije:

$$r = \frac{\sum_i \{(x_i - \bar{x})(y_i - \bar{y})\}}{\left\{ \left[\sum_i (x_i - \bar{x})^2 \right] \left[\sum_i (y_i - \bar{y})^2 \right] \right\}^{1/2}} \quad -1 \leq r \leq +1$$

- Test za proveru značajnosti korelacije:

- ✓ Dvosmerni t-test
- ✓ **Nulta hipoteza** – ne postoji korelacija između x i y
- ✓ Broj stepeni slobode – $(n-2)$

$$t = \frac{|r|\sqrt{n-2}}{\sqrt{1-r^2}}$$

- Određivanje nagiba i odsečka prave:

$$b = \frac{\sum_i \{(x_i - \bar{x})(y_i - \bar{y})\}}{\sum_i (x_i - \bar{x})^2}$$

$$a = \bar{y} - b\bar{x}$$

- Greške nagiba i odsečka:

$s_{y/x} = \sqrt{\frac{\sum_i (y_i - \hat{y}_i)^2}{n-2}}$	$s_a = s_{y/x} \sqrt{\frac{\sum_i x_i^2}{n \sum_i (x_i - \bar{x})^2}}$	$s_b = \frac{s_{y/x}}{\sqrt{\sum_i (x_i - \bar{x})^2}}$
	$a \pm t_{(n-2)} s_a$	$b \pm t_{(n-2)} s_b$

- Izračunavanje nepoznate koncentracije i njene greške:

x_0 - vrednost dobijena zamenom izmerene y -vrednosti u jednačinu prave

$$s_{x_0} = \frac{s_{y/x}}{b} \sqrt{1 + \frac{1}{n} + \frac{(y_0 - \bar{y})^2}{b^2 \sum_i (x_i - \bar{x})^2}}$$

$$s_{x_0} = \frac{s_{y/x}}{b} \sqrt{\frac{1}{m} + \frac{1}{n} + \frac{(y_0 - \bar{y})^2}{b^2 \sum_i (x_i - \bar{x})^2}}$$

$$x_0 \pm t_{(n-2)} s_{x_0}$$

- Granica detekcije: vrednost koncentracije (x), za koju važi uslov:

$$\text{granica detekcije} = y_B + 3s_B$$

$$y_B (= a), \quad s_B (= s_{y/x})$$

Šta ćete raditi danas?

- Koeficijent korelacije.
- Test za proveru značajnosti korelacije.
- Određivanje nagiba i odsečka prave.
- Greške nagiba i odsečka.
- Izračunavanje nepoznate koncentracije i njene greške.
- Granica detekcije.
- Upoređivanje dve analitičke metode regresionom analizom.

Primer 1. Regresija i korelacija

Standardni rastvor fluoresceina je ispitivan fluorescentnom spektrofotometrijom i dobijeni su sledeći rezultati:

Intenzitet fluorescencije: 2,1 5,0 9,0 12,6 17,3 21,0 24,7
 Koncentracija, pg/cm³: 0 2 4 6 8 10 12

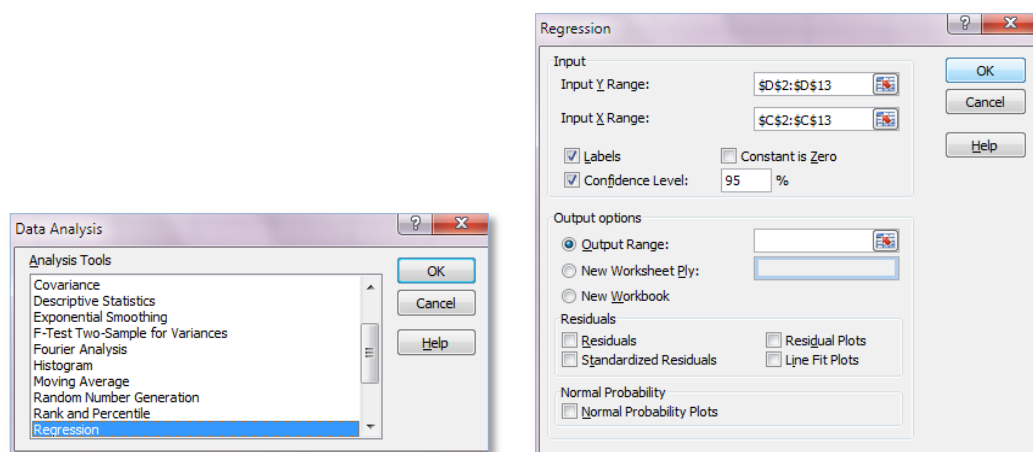
Određiti:

- Koeficijent korelacije.
- Proveriti da li je korelacija značajna.
- Nagib i odsečak.
- Standardnu devijaciju i interval pouzdanosti nagiba i odsečka.
- Nepoznatu koncentraciju i njenu grešku za intenzitet fluorescencije od 13,5.
- Granicu detekcije određivanja fluoresceina.

U okviru Data Analysis ToolPack-a postoji alatka **Regression**. Odaberite opciju sa padajućeg menija Tools/Data Analysis; starujte komandu Regression.

U polje **Input Range** unesite opseg ćelija između kojih su smešteni vaši podaci a koji se odnose na odgovarajuću osu.

U polje **Output Range** unesite ćeliju ispod koje i desno do koje nema nikakvih podataka na radnom listu, u suprotnom excel će vam saopštiti da će rezultate prepisati preko već postojećih podataka.



Ukoliko ste sve ispravno uradili trebalo bi da konačan rezultat izgleda ovako:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1													
2		0	2.1		SUMMARY OUTPUT								
3		2	5.0										
4		4	9.0		Regression Statistics								
5		6	12.6		Multiple R	0.99887957							
6		8	17.3		R Square	0.99776039							
7		10	21.0		Adjusted R Square	0.99731246							
8		12	24.7		Standard Error	0.43284771							
9					Observations	7							
10					ANOVA								
11						df	SS	MS	F	Significance F			
12					Regression	1	417.3432143	417.3432	2227.528	8.066E-08			
13					Residual	5	0.936785714	0.187357					
14					Total	6	418.28						
15													
16													
17						Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
18					Intercept	1.51785714	0.294936001	5.146395	0.003626	0.75970002	2.276014	0.7597	2.276014
19					X Variable 1	1.93035714	0.040900264	47.19669	8.07E-08	1.82521967	2.035495	1.82522	2.035495
20													

- a) Koeficijent korelacije – $r = 0,9989$
b) $t = 47,19669$
 $t_k = 2,57$
 $t > t_k - H_0$ se odbacuje, tj. postoji korelacija između koncentracije i intenziteta fluorescencije.
c) Nagib i odsečak – $b = 1,93$; $a = 1,52$
d) Standardna devijacija i interval pouzdanosti nagiba i odsečka –
 $b = 1,93 \pm 0,11$; $a = 1,52 \pm 0,76$
e) Nepoznata koncentracija i njena grešku za intenzitet fluorescencije od 13,5 –
 $x_0 = 6,21 \pm 0,62$
f) Granic detekcije određivanja fluoresceina –
 $0,67 \text{ pg/cm}^3$.

- Upoređivanje dve analitičke metode regresionom analizom.

Primer 2. Poređenje dve analitičke metode regresijom

Nivo kiseline u 20 uzoraka urina određivan je novom fluorimetrijskom metodom i rezultati upoređivani rezultatima dobijenim standardnom fotometrijskom tehnikom. Dobijeni su sledeći podaci:

Uzorak	Fluorimetrija	Fotometrija
1	1,87	1,98
2	2,20	2,31
3	3,15	3,29
4	3,42	3,56
5	1,10	1,23
6	1,41	1,57
7	1,84	2,05
8	0,68	0,66
9	0,27	0,31
10	2,80	2,82
11	0,14	0,13
12	3,20	3,15
13	2,70	2,72
14	2,43	2,31
15	1,78	1,92
16	1,53	1,56
17	0,84	0,94
18	2,21	2,27
19	3,10	3,17
20	2,34	2,36

Da li se metode razlikuju po preciznosti?

Regresionu pravu možemo da upotrebimo i za upoređivanje dve metode upotrebljene za određivanje različitih koncentracija analita. Na x -osu se nanose rezultati preciznije metode. Regresionom analizom izračunava se nagib (b), odsečak (a) i korelacioni koeficijent (r) regresione prave. Ukoliko 0 ulazi u interval pouzdanosti odsečka, 1 u interval pouzdanosti nagiba i korelacioni koeficijent je blizak 1 , može se zaključiti da se metode statistički značajno ne razlikuju.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1														
2		1.87	1.98		SUMMARY OUTPUT									
3		2.20	2.31											
4		3.15	3.29		Regression Statistics									
5		3.42	3.56		Multiple R	0.996647								
6		1.10	1.23		R Square	0.993305								
7		1.41	1.57		Adjusted R Squ	0.992933								
8		1.84	2.05		Standard Error	0.082875								
9		0.68	0.66		Observations	20								
10		0.27	0.31											
11		2.80	2.82		ANOVA									
12		0.14	0.13			df	SS	MS	F	gnificance F				
13		3.20	3.15		Regression	1	18.34127	18.34127	2670.439	5.03E-21				
14		2.70	2.72		Residual	18	0.123629	0.006868						
15		2.43	2.31		Total	19	18.4649							
16		1.78	1.92											
17		1.53	1.56											
18		0.84	0.94											
19		2.21	2.27											
20		3.10	3.17											
21		2.34	2.36											
22														
23	st.dev.	0.985818	0.990021											
24														

r = 0,9966

Odsečak iznosi -0,0497, sa donjom i gornjom granicom intervala pouzdanosti od -0,1399 i 0,0404 ⇒ interval sadrži 0.

Nagib iznosi 0,9924, donjom i gornjom granicom intervala pouzdanosti od 0,9521 i 1,0328 ⇒ interval sadrži 1.

Zaključak: metode se statistički značajno ne razlikuju.

Zadaci

- U cilju određivanja sadržaja izo-oktana u određenoj smeši ugljovodonika vršena je kalibracija standardnim rastvorima izo-oktana i dobijene sledeće vrednosti:

Molski procenat izo-oktana	Oblast pika
0,352	1,09
0,803	1,78
1,080	2,60
1,380	3,03
1,750	4,01

- Utvrđiti jednačinu kalibracione prave. Odrediti standardnu devijaciju i interval pouzdanosti nagiba i odsečka ($P = 0,05$).
- Nakon hromatografskog određivanja izo-oktana u ispitivanoj smeši ugljovodonika dobijen je pik u oblasti od 2,65. Odrediti sadržaj izo-oktana u smeši, kao i standardnu devijaciju, ako je:
 - dobijeni pik rezultat jednog merenja;
 - dobijeni pik srednja vrednost četiri merenja.
- Odrediti granicu detekcije ispitivane metode.

- U seriji standardnih rastvora puferovanih na pH = 4,6 određivan je Cd^{2+} i dobijeni su sledeći rezultati:

$[\text{Cd}^{2+}](\text{nM})$	15,4	30,4	44,9	59,0	72,7	86,0
$S_{\text{izm}}(\text{nA})$	4,8	11,4	18,2	26,6	32,3	37,7

- Da li je S_{izm} linearna funkcija koncentracije?
- Odrediti jednačinu odgovarajuće regresione prave.
- Izračunati interval pouzdanosti nagiba i odsečka dobijene prave.

Za rastvore puferovane na pH = 3,7 dobijeni su sledeći rezultati:

$[\text{Cd}^{2+}](\text{nM})$	15,4	30,4	44,9	59,0	72,7	86,0
$S_{\text{izm}}(\text{nA})$	15,0	42,7	58,8	77,0	101,0	118,0

- Da li je veća osetljivost metode na nižoj pH-vrednosti?

- e) Za nepoznati uzorak puferovan na pH = 3,7 dobijen je intenzitet signala od 66,3 nA. Odrediti koncentraciju Cd u uzorku ($P = 0,05$).

3. Pri kolorimetrijskom određivanju sadržaja glukoze u seriji standardnih rastvora dobijeni su sledeći rezultati:

Koncentracija, mM:	0,0	2,0	4,0	6,0	8,0	10,0
Apsorbancija, A:	0,002	0,150	0,294	0,434	0,570	0,704

- Proveriti da li je apsorbancija linearna funkcija koncentracije.
 - Odrediti nagib i odsečak dobijene regresione prave.
 - Odrediti standardnu devijaciju nagiba i odsečka ove prave.
 - Na osnovu ovih rezultata odrediti granicu detekcije glukoze
 - Dobijena kalibraciona prava upotrebljena je za određivanje sadržaja glukoze u krvnom serumu. Za četiri paralelna određivanja glukoze u nepoznatom uzorku dobijene su sledeće vrednosti A: 0,325; 0,327; 0,339; 0,326. Koliki je sadržaj glukoze u nepoznatom uzorku?
 - Primenom Theil-ove „nepotpune“ metode odrediti nagib i odsečak regresione prave.
4. Dve instrumentalne metode su poređene grafički i dobijena je linearna zavisnost koja je data sledećom jednačinom: $y = 3,87 \pm 15,34 + 0,86 x \pm 0,08$ ($r = 0,9945$). Rezultati koji su dobijeni ovim metodama statistički se razlikuju na nivou značajnosti od $P = 0,05$.
- da
 - ne