

Soubor úloh – analýza závislostí¹

1. V následující tabulce jsou uvedené ceny nových bytů postavených několika stavebními firmami (v tis. Kč):

typ bytu	cena bytu
1+kk	221, 255, 264, 201
1+1	321, 329, 374, 335, 309
1+2	592, 543, 526, 505, 528, 564
1+3	640, 621, 687, 643, 672
1+4	934, 948, 962, 895

Sestavte korelační tabulku, vypočítejte podmíněné a celkové charakteristiky, určete poměr determinace, proveďte F-test a sestrojte graf podmíněných průměrů a rozptylů. Všechny výsledky interpretejte.

[0,988; 399,263]

2. U 40 pracovníků byla sledována závislost počtu chybných operací za směnu (Y) na délce zapracování v hodinách (X) s těmito výsledky:

x_i	y_i	x_i	y_i	x_i	y_i	x_i	y_i	x_i	y_i	x_i	y_i	x_i	y_i	x_i	y_i
3	2	2	4	1	5	2	6	4	3	3	6	2	5	3	5
4	3	2	6	3	3	4	4	1	6	5	2	4	1	4	2
1	4	4	1	2	7	4	3	2	4	3	3	1	6	1	7
4	5	2	5	3	4	3	5	2	5	2	3	5	4	4	4
4	2	5	3	5	1	3	4	5	1	3	4	3	4	5	2

a) Sestavte korelační tabulku, vypočítejte podmíněné a celkové charakteristiky, určete poměr determinace, proveďte F-test a sestrojte graf podmíněných průměrů a rozptylů. b) Určete regresní přímkou a její hodnotu v bodě $x = 3$, vypočítejte koeficient determinace a reziduální rozptyl.

[a) 0,518; 9,400 b) $\hat{y} = 6,718 - 0,933 \cdot x$; $\hat{y}(3) = 3,92$; $r^2 = 0,510$; $s_r^2 = 1,381$]

3. V rámci průzkumu kvality pohonných hmot bylo provedeno měření spotřeby benzínu při použití 4 různých druhů benzínu (např. od různých výrobců). Testování se provádělo na 16 vozidlech stejného modelu tak, že vždy 4 auta použila stejný benzin. Výsledky měření spotřeby v l/100 km při jednotlivých jízdách jsou uvedené v následující datové tabulce:

typ A	typ B	typ C	typ D
6,10	6,23	5,96	6,08
5,95	6,10	6,00	5,99
6,00	6,31	5,82	5,80
6,12	6,15	6,04	5,91

Sestavte korelační tabulku, vypočítejte podmíněné a celkové charakteristiky, určete poměr determinace, proveďte F-test a sestrojte graf podmíněných průměrů a rozptylů. Všechny výsledky interpretejte.

[0,587; 5,692]

4. V rámci průzkumu využívání volného času odpovědělo 50 oslovených osob na 2 otázky:

¹ není-li uvedeno jinak, u testů a odhadů volte hladinu významnosti $\alpha = 0,05$

kolik hodin v průměru denně strávíte u TV jednak ve všední dny (po–pá, prom. X), a jednak o víkendu (so–ne, prom. Y). Vstupní informace jsou uvedené v následující datové tabulce:

osoba	x_i	y_i	osoba	x_i	y_i	osoba	x_i	y_i	osoba	x_i	y_i	osoba	x_i	y_i
1	1	1	11	2	3	21	3	4	31	5	5	41	5	4
2	1	2	12	2	1	22	4	4	32	5	2	42	5	2
3	1	1	13	2	3	23	3	6	33	5	6	43	5	4
4	1	3	14	3	2	24	3	3	34	5	4	44	2	5
5	2	1	15	3	2	25	4	3	35	5	3	45	3	5
6	2	4	16	3	1	26	4	4	36	5	6	46	4	2
7	2	2	17	3	2	27	1	4	37	3	6	47	4	2
8	2	2	18	2	3	28	2	5	38	2	6	48	5	6
9	1	2	19	1	1	29	4	5	39	3	6	49	5	6
10	1	3	20	2	4	30	5	1	40	4	3	50	4	6

a) Sestavte korelační tabulku, vypočítejte podmíněné a celkové charakteristiky, určete poměr determinace, proveďte F-test a sestrojte graf podmíněných průměrů a rozptylů. b) Určete regresní ...

[0,164; 2,211 b) $\hat{y} = 2,083 + 0,454 \cdot x$; 0,145; $\hat{y}(4) = 3,899$]

5. U prodaných ojetých automobilů Škoda Felicia byla zjištěná cena (v tis. Kč) a počet ujetých km (v tis. km). Údaje jsou sestavené do následující tabulky:

ujeté km	cena
0–20	55,0 54,6 50,6 51,1 50,0 43,0
20–40	47,0 43,6 41,3 34,0
40–60	39,9 34,0 29,0 31,6
60–80	31,0 28,0 24,6 27,0
80–100	25,6 17,6

a) Sestavte korelační tabulku, vypočítejte podmíněné a celkové charakteristiky, určete poměr determinace, proveďte F-test a sestrojte graf podmíněných průměrů a rozptylů. b) Určete regresní přímkou, vypočítejte koeficient determinace a reziduální rozptyl, odhadněte průměrnou cenu auta s 50 tis. km. Všechny výsledky interpretujte.

[a) 0,872; 25,540 b) $\hat{y} = 53,567 - 0,372 \cdot x$; -0,929; 34,967]

6. Ve 12 supermarketech byl testován vliv vystavení zboží na jeho odbyt. Zboží bylo umístěno 3 různými způsoby, vždy ve 4 supermarketech stejně, a sledoval se počet prodaných kusů za určitou dobu. Rozhodněte, zda umístění zboží ovlivňuje počet prodaných kusů.

způsob uložení	počet prodaných kusů
A	42, 46, 37, 48
B	53, 92, 79, 104
C	71, 89, 122, 115

Sestavte korelační tabulku, vypočítejte podmíněné a celkové charakteristiky, určete poměr determinace, proveďte F-test a sestrojte graf podmíněných průměrů a rozptylů. Všechny výsledky interpretujte.

[0,675; 9,331]

7. Lze na základě cen 16 náhodně vybraných dvoupokojových bytů ve městech A, B, C prohlásit, že v cenách bytů jsou mezi městy rozdíly?

město	cena bytu (v tis. Kč)
A	1040, 950, 1100, 1570, 870
B	900, 1700, 1290, 900, 1250, 1460
C	1550, 1300, 1220, 1660, 1340

Sestavte korelační tabulku, vypočítejte podmíněné a celkové charakteristiky, určete poměr determinace, proveďte F-test a sestrojte graf podmíněných průměrů a rozptylů.

8. Z velkého souboru zaměstnaneckých domácností bylo vybráno po deseti domácnostech jednočlenných, dvoučlenných, tříčlenných, čtyřčlenných a pětičlenných. Byly sledovány jejich měsíční výdaje v Kč za potraviny a nápoje připadající na jednoho člena domácnosti. Získané údaje jsou uvedeny v následující tabulce:

počet členů domácnosti	výdaje za potraviny a nápoje připadající na 1 člena domácnosti v Kč
1	2687, 3936, 3283, 3564, 3252, 3440, 4044, 4014, 3776, 2829
2	3031, 2350, 4143, 3236, 3800, 2901, 2656, 2878, 4352, 3624
3	3529, 3325, 3731, 3313, 2303, 2565, 2777, 3899, 2755, 3254
4	2137, 2201, 2786, 3132, 2223, 3433, 2760, 2763, 2257, 2661
5	2062, 2239, 2448, 2137, 2032, 3101, 3121, 2674, 2526, 2125

Pomocí analýzy rozptylu zjistěte, zda v uvedených domácnostech závisí vydání za potraviny a nápoje na počtu členů domácnosti.

9. Byla zjištěna výška otců a výška jejich nejstarších synů, údaje jsou v tabulce [v cm]:

otec	165	178	158	170	180	160	170	167	185	165	173	175
syn	162	184	163	170	189	165	177	170	187	176	171	183

a) Sestrojte bodový graf. b) Určete regresní přímku a vypočítejte koeficient determinace resp. korelační koeficient. c) Odhadněte průměrnou výšku syna při výšce otce 178 cm. Všechny výsledky interpretujte.

[b) $\hat{y} = 0,123 + 1,024 \cdot x$; 0,783; 0,885 c) 182,432]

10. V rámci srovnání několika autobazarů byly v jednom z nich zjištěné údaje o stáří a ceně (prom. X a Y) 10 ojetých aut značky Felicia Combi. Údaje jsou uvedené v následující tabulce:

stáří (roky)	3	4	5	6	7	7	8	8	9	9
cena (tis.Kč)	167	165	139	149	119	129	89	115	76	89

a) Sestrojte bodový graf. b) Určete regresní přímku a vypočítejte koeficient determinace resp. korelační koeficient. c) Odhadněte ... Všechny výsledky interpretujte.

11. U náhodně vybraných řidičů, kteří mají evidované dopravní přestupky, byl zjišťován vztah mezi výší pokut (v tis. Kč) a počtem přestupků. Výsledky jsou uvedené v tabulce:

pokuty	1	1	2	2	2	3	3	3	3	4
přestupky	5	5	3	4	5	1	2	2	3	2

- a) Sestrojte bodový graf. b) Určete regresní přímku a vypočítejte koeficient determinace. c) Určete regresní hyperbolu a vypočítejte index determinace. d) Pro oba regresní modely odhadněte průměrný počet přestupků při úhrnné pokutě 2,5 tis. Kč a pomocí reziduálních rozptylů stanovte vhodnější model. Všechny výsledky interpretujte.

[b) $\hat{y} = 6,286 - 1,286 \cdot x$; 0,708 c) $\hat{y} = 1,009 + 4,310/x$; 0,638 d) $\hat{y}(2,5) = 3,071$; 2,733; $s_r^2 = 0,714$; 0,888]

12. O 7 vybraných strojích v určitém podniku máme informace o jejich stáří (v letech) a týdenních nákladech na jejich údržbu (v Kč):

stáří stroje	1	1	3	3	5	6	7
náklady	35	52	81	105	100	125	120

- a) Sestrojte bodový graf. b) Určete regresní přímku a vypočítejte koeficient determinace. c) Určete regresní logaritmickou křivku a vypočítejte index determinace. d) Pro oba regresní modely odhadněte průměrné náklady na údržbu strojů 4 roky starých a pomocí reziduálních rozptylů stanovte vhodnější model pro popis závislosti nákladů na stáří strojů. Všechny výsledky interpretujte.

[b) $\hat{y} = 39,444 + 13,150 \cdot x$; $r^2 = 0,828$ c) $\hat{y} = 44,65 + 40,49 \cdot \ln x$; $i^2 = 0,909$ d) $\hat{y}(4) = 92,043$; $\hat{y}(4) = 100,78$; $s_r^2 = 239,85$; $s_r^2 = 126,8$]

13. Při sociologickém průzkumu názorů mladých lidí na manželství byla sledována mimo jiné závislost mezi hodnocením manželství rodičů (X) a rodičovským vzorem (Y). Zpracované odpovědi z 200 dotazníků jsou uvedené v následující tabulce:

Hodnocení manželství rodičů	Hodnocení rodičovského vzoru			$n_{i\cdot}$
	rozhodně ano	výhrady	rozhodně ne	
spokojené	52	17	0	69
problémy	34	54	9	97
nespokojené	2	12	5	19
neuvedeno	6	7	2	15
$n_{\cdot j}$	94	90	16	200

14. xxx