

Kontingenční tabulky a korelační koeficienty

1. V následující tabulce je výsledek žákovského průzkumu o oblíbě přírodovědných předmětů na gymnáziu ve třech čtvrtých ročnících. Zajímá nás, zda obliba předmětů závisí na pohlaví dotazovaných studentů. Ověřte χ^2 -testem, zda jsou veličiny: obliba přírodovědných předmětů a pohlaví nezávislé. Formulujte nulovou hypotézu a výsledek testu komentujte slovně. Zkontrolujte, zda četnosti splňují podmínku pro platnost χ^2 -test a zdůvodněte slovně.

	Matematika	Fyzika	Chemie	Biologie	Celkem
Dívky	26	18	12	24	80
Chlapci	8	16	12	4	40
Celkem	34	34	24	28	120

2. Ve studii nozokomiálních infekcí se sledovaly veličiny, které obecně platí za rizikový faktor. V následující tabulce je výsledek šetření ve 100 nemocnicích v ČR a za rizikový faktor je považováno pohlaví pacienta. Zajímá nás, zda se potvrdí závislost výskytu NI na pohlaví. Ověřte testem χ^2 -testem, zda jsou tyto veličiny nezávislé. Formulujte nulovou hypotézu a výsledek testu komentujte slovně.

Pohlaví	Pacienti s NI	Pacienti bez NI	Celkem
Muž	238	715	953
Žena	131	531	662
Celkem	369	1246	1615

3. Z dlouhodobého sledování dětí u dětského lékaře vyšly tyto údaje:

Pohlaví	Vrozené vady kyčlí		Celkem
	Ano	Ne	
Chlapci	325	1589	1914
Dívky	389	1525	1914
Celkem	714	3114	3828

Ověřte χ^2 -testem, zda jsou vrozená vada kyčlí a pohlaví dítěte nezávislé veličiny. Formulujte nulovou hypotézu a výsledek testu komentujte slovně. Určete Pearsonův a Cramerův koeficient kontingence.

4. V následující tabulce je výsledek žákovského průzkumu o oblíbě předmětů na jazykové škole ve druhých ročnících. Zajímá nás, zda obliba předmětů závisí na pohlaví dotazovaných studentů. Ověřte χ^2 -testem, zda jsou veličiny: obliba humanitních předmětů a pohlaví nezávislé. Formulujte nulovou hypotézu a výsledek testu komentujte slovně. Zkontrolujte, zda četnosti splňují podmínku pro platnost testu χ^2 -testem a zdůvodněte slovně. Určete Pearsonův, Cramerův a Čuprovův koeficient kontingence.

	Čeština	Francouzština	Angličtina	Latina	Celkem
Dívky	24	18	21	4	67
Chlapci	15	14	22	2	53
Celkem	39	32	43	6	120

5. Byla zjištěna výška otců a výška jejich nejstarších synů [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

Určete Pearsonův, Spearmanův a Kendallův korelační koeficient. Provedte test významnosti Pearsonova korelačního koeficientu. [Datový soubor: vyska.otec_syn.txt]

6. 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

Určete Pearsonův, Spearmanův a Kendallův korelační koeficient. Provedte test významnosti Pearsonova korelačního koeficientu. [Datový soubor: stari_stroje_naklady.txt]

7. U 30 žáků byly zjištěny hodnoty znaků X – známka z fyziky, Y – známka z chemie a Z – pohlaví (0... dívka, 1... chlapec).

Fyzika	Chemie	Pohlaví	Fyzika	Chemie	Pohlaví	fyzika	Chemie	Pohlaví
1	2	1	2	3	1	3	4	0
2	3	1	2	3	0	5	4	0
2	3	0	2	2	0	2	1	0
4	5	1	4	5	0	2	2	0
2	1	1	3	3	1	3	1	1
4	3	1	3	4	1	3	4	0
2	2	1	2	3	1	2	1	1
4	2	0	2	2	0	1	1	0
3	3	0	5	3	1	1	1	1
4	5	0	3	2	1	1	2	0

Určete Pearsonův, Spearmanův a Kendallův korelační koeficient pro známky z fyziky a chemie. Výpočty proveďte zvlášť pro dívky, pro hochy a dohromady bez závislosti na pohlaví.

[Datový soubor: znamky.txt]

8. U 40 pracovníků byla sledována závislost počtu chybných operací za směnu (Y) na délce zpracování 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

Určete Pearsonův, Spearmanův a Kendallův korelační koeficient. Provedte test významnosti Pearsonova korelačního koeficientu. [Datový soubor: chyby_zpracovani.txt]

9. Data popisují výsledky vstupních zdravotních testů uchazečů o službu u policie.

Tlak	66	87	85	59	76	77	70	66	75	66
Hmotnost	87,36	117,6	82,85	62,32	82	102	70,12	88,07	77,96	74,33
Tuk	16,98	27,6	6,61	3,26	19	27	6,88	18,8	18,87	8,15
Tlak	74	68	72	76	94	63	80	67	77	78
Hmotnost	56,2	81,75	80,24	74,81	61,98	95,23	72,48	92,45	104,56	66,2
Tuk	3,44	20,31	12,96	12,42	3,58	12,91	11,34	17,5	18,93	10,94
Tlak	77	67	78	78	80	95	76	78	73	80
Hmotnost	87,16	82,42	64,11	81,57	99,85	78,49	87,13	65,64	51,76	67,14
Tuk	17,72	9,55	9,54	13,1	17,75	9,57	18,52	6,4	2,86	4,31
Tlak	81	61	65	69	66	75	72	66	93	77
Hmotnost	78,74	86,83	70,48	72,67	85,86	84,86	66,97	68,33	63,34	85,72
Tuk	16,26	9,72	6,29	4,37	14,43	17	5,8	8,14	3,63	23,61
Tlak	68	71	84	81	74	79	89	79	80	67
Hmotnost	89	95,17	84,19	63,12	70,01	82,11	71	94,56	70,91	79,19
Tuk	18,83	19,16	15,83	8,77	6,61	22,22	8,29	26,82	9,32	19,9

Určete Pearsonův, Spearmanův a Kendallův korelační koeficient pro jednotlivé dvojice proměnných.
 Spočtete koeficienty mnohonásobné korelace. [Datový soubor: vstupni_testy.txt]