

1. V tabulce jsou uvedeny údaje o hodnotě produkce v 100 000Kč (y) a o výši investic v 10 000Kč (x) v průběhu 12 po sobě jdoucích letech ve strojírenské firmě.
- a) Stanovte rovnici regresní přímky modelující závislost hodnoty produkce na výši investic.
 - b) Pomocí koeficientu determinace zhodnoťte výstižnost regresní funkce.
 - c) Vypočítejte rezidua modelu.
 - d) Na hladině významnosti 0,05 testujte, zda je model zatížen autokorelací. Využijte Durbin-Watsonův test.

Rok	x	y
1	16.3	44.4
2	16.8	48.4
3	18.5	54.2
4	16.42	50
5	17.9	54.9
6	17.4	53.9
7	15.7	47
8	16.2	52.4
9	17	53
10	16.7	52.9
11	17.5	53.1
12	19.1	62

Rok	x	y
1	16.3	44.4
2	16.8	48.4
3	18.5	54.2
4	16.42	50
5	17.9	54.9
6	17.4	53.9
7	15.7	47
8	16.2	52.4
9	17	53
10	16.7	52.9
11	17.5	53.1
12	19.1	62

Výzkumný zemědělský ústav zkoušel vliv nově vyvinutého typu hnojiva na výnosy keříčkových rajčat. Výsledky průzkumu jsou uvedeny v tabulce.

- Zvolte vhodný typ regresního modelu popisujícího závislost výnosů na množství použitého hnojiva - stačí pro popis lineární funkce, nebo je vhodnější zvolit parabolou?
- Určete rovnici regresní funkce
- Zhodnoťte výstižnost zvolené regresní funkce
- Určete optimální množství hnojiva, tj. takové množství, při kterém dosáhneme maximálních výnosů

Hnojivo (100 kg/h)	0.7	1.3	3.8	4.2	2.5	5	5.3	6	5.8	6.2
Výnosy (t/ha)	12	14.6	40.3	43.7	28	45	41.7	39	35.4	29.6

V sociologické studii okresu Karviná byla také zkoumána souvislost ročních úspor s ročními příjmy rodin s dvěma dětmi školou povinnými. Výsledky studie uvádí tabulka.

- Najděte lineární regresní model popisující závislost úspor na příjmech
- Odhadněte model polynomicou funkcí řádu 2 (kvadratickou funkcí). Napište rovnici této funkce a označte koeficienty
- Odhadněte model exponenciální funkcí. Napište rovnici této funkce a odpovídající hodnotu koeficientu
- Odhadněte model mocninnou funkcí. Napište rovnici této funkce a odpovídající hodnotu koeficientu
- Odhadněte model logaritmickou funkcí. Napište rovnici této funkce a odpovídající hodnotu koeficientu
- Který z modelů a)-e) vystihuje data nejlépe? Proč?

Příjem (tis. Kč)	Uspory(tis. Kč)
104	6
125	5.6
146	9.2
167	14
111	8
135	9.1
189	20.5
196	29
205	23.2
210	38.5
170	25
230	40

dpovídající hodnotu koeficientu determinace.
tu determinace.
determinace.
tu determinace.

Tabulka zachycuje čtvrtletnou míru nezaměstnanosti v německu v letech 1969-2014.

Zdroj: <https://research.stlouisfed.org>

a) Načrtněte graf této časové řady.

b) Tuto časovou řadu vyrovnejte pomocí klouzavých průměrů.

c) Vyrovnanou časovou řadu načrtněte, porovnejte graf s výsledkem úlohy a).

DATE	Pořadí	VALUE
1969-01-01	1	1.23
1969-04-01	2	0.50
1969-07-01	3	0.40
1969-10-01	4	0.53
1970-01-01	5	0.97
1970-04-01	6	0.43
1970-07-01	7	0.40
1970-10-01	8	0.53
1971-01-01	9	0.97
1971-04-01	10	0.53
1971-07-01	11	0.57
1971-10-01	12	0.80
1972-01-01	13	1.27
1972-04-01	14	0.80
1972-07-01	15	0.73
1972-10-01	16	0.93
1973-01-01	17	1.23
1973-04-01	18	0.83
1973-07-01	19	0.80
1973-10-01	20	1.33
1974-01-01	21	2.23
1974-04-01	22	1.77
1974-07-01	23	1.97
1974-10-01	24	3.00
1975-01-01	25	4.30
1975-04-01	26	3.87
1975-07-01	27	3.80
1975-10-01	28	4.27
1976-01-01	29	4.90
1976-04-01	30	3.73
1976-07-01	31	3.50
1976-10-01	32	3.80
1977-01-01	33	4.47
1977-04-01	34	3.70
1977-07-01	35	3.63
1977-10-01	36	3.87
1978-01-01	37	4.50
1978-04-01	38	3.57
1978-07-01	39	3.43
1978-10-01	40	3.57
1979-01-01	41	4.13
1979-04-01	42	3.03
1979-07-01	43	2.93
1979-10-01	44	3.07
1980-01-01	45	3.67
1980-04-01	46	2.97
1980-07-01	47	3.20
1980-10-01	48	3.70

1981-01-01	49	4.80
1981-04-01	50	4.23
1981-07-01	51	4.73
1981-10-01	52	5.70
1982-01-01	53	7.00
1982-04-01	54	6.13
1982-07-01	55	6.50
1982-10-01	56	7.50
1983-01-01	57	8.97
1983-04-01	58	7.90
1983-07-01	59	7.80
1983-10-01	60	7.97
1984-01-01	61	8.93
1984-04-01	62	7.77
1984-07-01	63	7.83
1984-10-01	64	7.93
1985-01-01	65	9.13
1985-04-01	66	7.90
1985-07-01	67	7.80
1985-10-01	68	7.90
1986-01-01	69	9.03
1986-04-01	70	7.60
1986-07-01	71	7.47
1986-10-01	72	7.47
1987-01-01	73	8.73
1987-04-01	74	7.53
1987-07-01	75	7.63
1987-10-01	76	7.70
1988-01-01	77	8.70
1988-04-01	78	7.60
1988-07-01	79	7.53
1988-10-01	80	7.33
1989-01-01	81	7.87
1989-04-01	82	6.87
1989-07-01	83	6.77
1989-10-01	84	6.87
1990-01-01	85	7.27
1990-04-01	86	6.33
1990-07-01	87	6.17
1990-10-01	88	5.90
1991-01-01	89	6.13
1991-04-01	90	5.40
1991-07-01	91	5.57
1991-10-01	92	6.37
1992-01-01	93	8.07
1992-04-01	94	7.43
1992-07-01	95	7.67
1992-10-01	96	7.80
1993-01-01	97	8.90
1993-04-01	98	8.50
1993-07-01	99	9.03
1993-10-01	100	9.30
1994-01-01	101	10.37
1994-04-01	102	9.57
1994-07-01	103	9.37
1994-10-01	104	9.00

1995-01-01	105	9.87
1995-04-01	106	9.13
1995-07-01	107	9.30
1995-10-01	108	9.47
1996-01-01	109	10.90
1996-04-01	110	10.10
1996-07-01	111	10.17
1996-10-01	112	10.40
1997-01-01	113	12.03
1997-04-01	114	11.13
1997-07-01	115	11.33
1997-10-01	116	11.43
1998-01-01	117	12.43
1998-04-01	118	10.93
1998-07-01	119	10.53
1998-10-01	120	10.40
1999-01-01	121	11.40
1999-04-01	122	10.33
1999-07-01	123	10.23
1999-10-01	124	10.07
2000-01-01	125	10.83
2000-04-01	126	9.40
2000-07-01	127	9.20
2000-10-01	128	9.03
2001-01-01	129	9.97
2001-04-01	130	9.13
2001-07-01	131	9.13
2001-10-01	132	9.27
2002-01-01	133	10.27
2002-04-01	134	9.57
2002-07-01	135	9.60
2002-10-01	136	9.73
2003-01-01	137	11.17
2003-04-01	138	10.47
2003-07-01	139	10.30
2003-10-01	140	10.13
2004-01-01	141	11.00
2004-04-01	142	10.40
2004-07-01	143	10.43
2004-10-01	144	10.40
2005-01-01	145	12.57
2005-04-01	146	11.83
2005-07-01	147	11.50
2005-10-01	148	11.00
2006-01-01	149	12.10
2006-04-01	150	10.97
2006-07-01	151	10.37
2006-10-01	152	9.67
2007-01-01	153	10.03
2007-04-01	154	9.13
2007-07-01	155	8.67
2007-10-01	156	8.10
2008-01-01	157	8.53
2008-04-01	158	7.80
2008-07-01	159	7.50
2008-10-01	160	7.20

2009-01-01	161	8.43
2009-04-01	162	8.27
2009-07-01	163	8.10
2009-10-01	164	7.70
2010-01-01	165	8.57
2010-04-01	166	7.77
2010-07-01	167	7.47
2010-10-01	168	7.00
2011-01-01	169	7.80
2011-04-01	170	7.07
2011-07-01	171	6.87
2011-10-01	172	6.50
2012-01-01	173	7.30
2012-04-01	174	6.77
2012-07-01	175	6.70
2012-10-01	176	6.57
2013-01-01	177	7.37
2013-04-01	178	6.83
2013-07-01	179	6.73
2013-10-01	180	6.57
2014-01-01	181	7.23
2014-04-01	182	6.63

Predikce

1. V tabulce jsou uvedeny údaje o hodnotě produkce v 100 000Kč (y) a o výši investic v 10 000Kč (x) v průběhu 12 po sobě jdoucích letech ve strojířenské firmě.
- a) Stanovte rovnici regresní přímky modelující závislost hodnoty produkce na výši investic.
 - b) Pomocí koeficientu determinace zhodnoťte výstižnost regresní funkce.
 - c) Na hladině významnosti 0,05 ověřte nenulovost regresního parametru β_1 .
 - d) V jakém rozmezí se bude s pravděpodobností 0,05 pohybovat regresní koeficient b_1 ?
 - e) Jakou produkci je možné očekávat při investici 2 050 000 Kč?
V jakém intervalu se bude tento odhad pohybovat? Volte $\alpha = 0,05$.

Rok	x	y
1	16.3	44.4
2	16.8	48.4
3	18.5	54.2
4	16.42	50
5	17.9	54.9
6	17.4	53.9
7	15.7	47
8	16.2	52.4
9	17	53
10	16.7	52.9
11	17.5	53.1
12	19.1	62

Durbin-Watsonův test na nepřítomnost autokorelace

V uvedeném modelu je ρ neznámý parametr – korelační koeficient nulové hypotézy, že model není zatížen autokorelací, proti alternativní hypotéze autokorelace ve tvaru AR(1). Test se provádí v několika krocích: 1. odhad parametrů původního regresního modelu časové řady metodou nejmenších čtverců, 2. získání vyrovnaných hodnot se vypočtením reziduální odchylky e_t , 3. se pak počítá testové kritérium

$$5-7 \quad T = \frac{\sum_{t=2}^T (e_t - e_{t-1})^2}{\sum_{t=1}^T e_t^2},$$

kde T udává délku časové řady hodnot, které jsou k dispozici. Pro testování se používají speciální statistické tabulky uvedené na konci tohoto učebního materiálu. Pro daný počet pozorování T , hladinu významnosti α a počet absolutního členu najde dolní hodnota d_L a horní hodnota d_H . Tyto hodnoty jsou párové korelace mezi rezidui regresního modelu. Tento odhad r n

$$5-8 \quad r = \frac{\sum_{t=2}^T e_t e_{t-1}}{\sum_{t=1}^T e_t^2}.$$

V případě, že výběrový korelační koeficient r je kladný, vyhodnotíme, že je-li testové kritérium 5-7 větší než d_H , nulová hypotéza o absenci autokorelace se zamítá, zatímco je-li kritérium menší než d_L , hypotéza se zamítá. Pokud je $d_L < T < d_H$, náhradní statistika $T^* = 4 - T$ a výše uvedené vyhodnocení se používá, že se aplikuje na testové kritérium T^* . Pokud se některá z testových hodnot d_L a d_H nelze na základě testu rozhodnout o platnosti hypotézy. Většinou se ale v takovém případě doporučuje vycházet z toho, že v modelu autokorelace je, protože v případě modelu časové řady je

cient. Test zkoumá platnost
ativní hypotéze, že v modelu
h. Nejprve se najdou odhady
ou nejmenších čtverců a ze
 e_t . Na základě těchto reziduí

ro toto kritérium jsou určeny
textu. V těchto tabulkách se
et parametrů modelu k bez
Dále je třeba vypočítat odhad
ná tvar

nocuje se statistický test tak,
senci autokorelace se přijímá,
ud je r záporný, vypočte se
rovádí stejně s tím rozdílem,
ovacích kritérií dostane mezi
nosti či neplatnosti nulové
ázet pro opatrnost z toho, že
je to dosti pravděpodobné.

4.7.2 Intervalové předpovědi

Bodová předpověď umožňuje pomocí jednoho čísla stanovit hodnotu předvídat. Intervaly spolehlivosti (konfidenční intervaly) nám dovolují stanovit příslušnou předpověď. Použijeme analogický postup, jaký jsme použili v kapitole 2, vzt (2.27). *Intervalová předpověď* vytvořená v čase n na časový horizont i je def oboustranný interval spolehlivosti:

$$[y(n+i) - t_{1-\alpha/2}(n-2) s_R \sqrt{Q_n(i)}, y(n+i) + t_{1-\alpha/2}(n-2) s_R \sqrt{Q_n(i)}],$$

kde s_R^2 je reziduální rozptyl definovaný vztahem, viz též (4.40):

$$s_R^2 = \frac{1}{n-p} \sum_{i=1}^n (y_i - \hat{Y}_i)^2,$$

p je počet odhadnutých parametrů modelu a

$$Q_n(i) = \sqrt{(1-R^2) \frac{n(n^2-1) + 12i^2}{(n^2-1)(n-2)}},$$

přičemž R^2 je koeficient determinace (3.20).

lané veličiny.
é intervalové
tahy (2.26) a
řinována jako

(4.68)

(4.69)

(4.70)