Časovne vrste - seminar

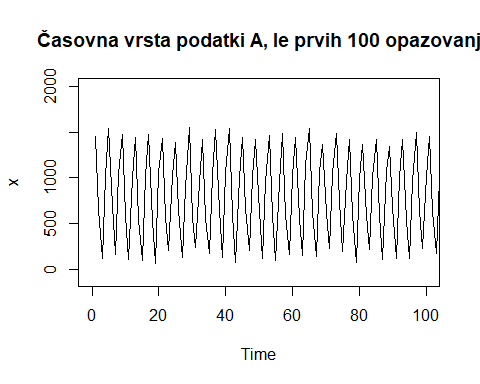
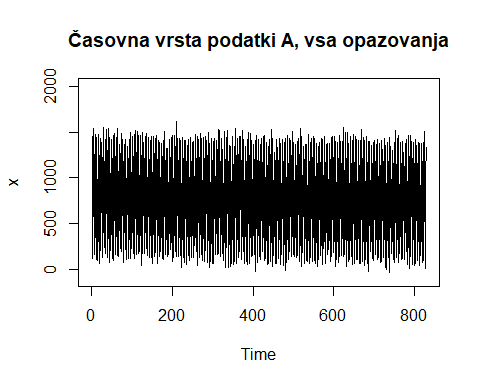
Poročilo o analizi časovnih vrst

Brina Pirc in Anja Trobec

Maj 2022

DATOTEKA A

1. Narišite graf in komentirajte, ali se iz njega vidi kakšen trend ali sezonskost.

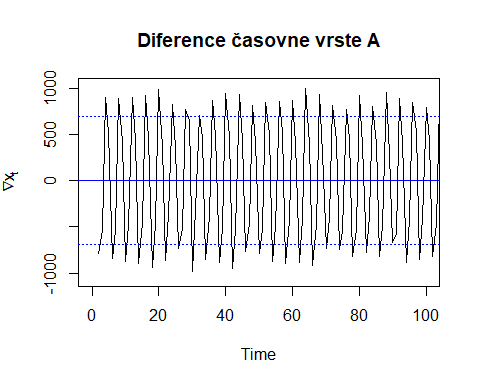
* 

## [1] "Ne opazimo trenda, opazimo pa sezonskost."

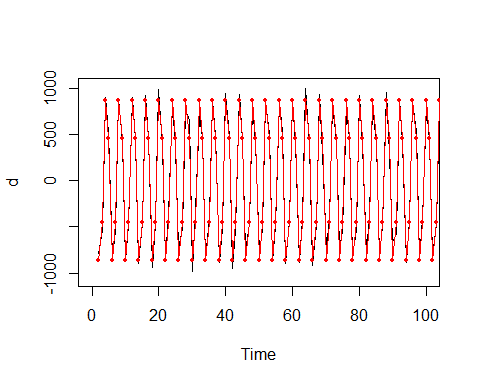
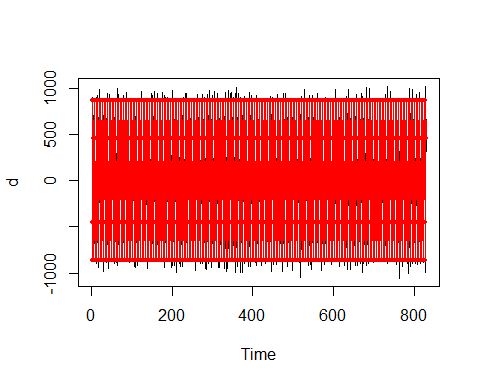
1. Odstranite morebiten trend in sezonskost z metodami, uporabljenimi pri tečaju: (zaporedno) diferenciranje, logaritmiranje, neposredna ocena sezonskih komponent, polinomski trend stopnje največ 3 ali prileganje periodične funkcije (ali kakšna kombinacija teh metod). Potem ko odstranite morebiten trend, narišite tudi surovi in zglajeni periodogram ter komentirajte, ali se vidi kakšna sezonskost in kakšna naj bi bila perioda.

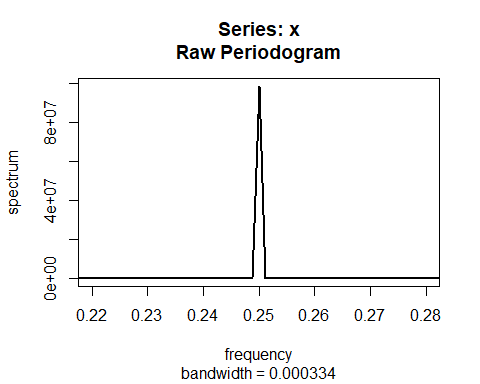
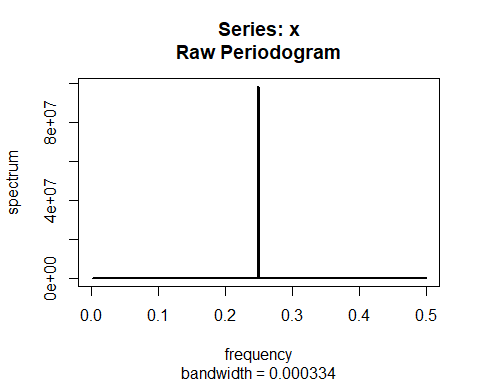
REŠEVANJE: logaritmiranja ne moremo uporabiti, ker imamo negativne podatke diferenciranje:

##   
## Call:  
## lm(formula = x ~ I(sin(p \* t)) + I(cos(p \* t)) + I(sin(p2 \* t)) +   
## I(cos(p2 \* t)))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -583.92 -146.58 -7.66 145.46 545.64   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 770.8550 7.6106 101.287 <2e-16 \*\*\*  
## I(sin(p \* t)) 287.0049 10.7703 26.648 <2e-16 \*\*\*  
## I(cos(p \* t)) 550.2589 10.7703 51.090 <2e-16 \*\*\*  
## I(sin(p2 \* t)) 18.6738 10.7656 1.735 0.0832 .   
## I(cos(p2 \* t)) -0.4413 10.7751 -0.041 0.9673   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 219.1 on 824 degrees of freedom  
## Multiple R-squared: 0.8012, Adjusted R-squared: 0.8002   
## F-statistic: 830.1 on 4 and 824 DF, p-value: < 2.2e-16



##   
## Call:  
## lm(formula = d ~ I(sin(perioda \* t)) + I(cos(perioda \* t)))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -204.725 -50.142 -0.614 49.743 226.055   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.1355 2.5258 -0.054 0.957   
## I(sin(perioda \* t)) 450.9594 3.5720 126.247 <2e-16 \*\*\*  
## I(cos(perioda \* t)) 862.4393 3.5720 241.442 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 72.68 on 825 degrees of freedom  
## Multiple R-squared: 0.989, Adjusted R-squared: 0.989   
## F-statistic: 3.712e+04 on 2 and 825 DF, p-value: < 2.2e-16



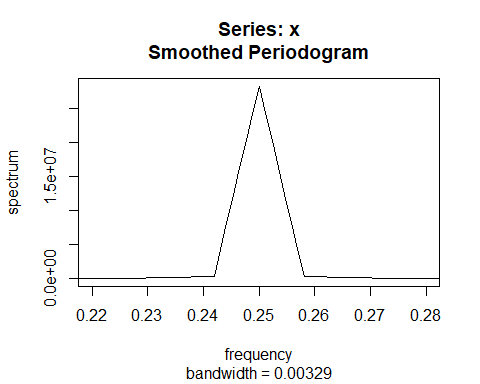
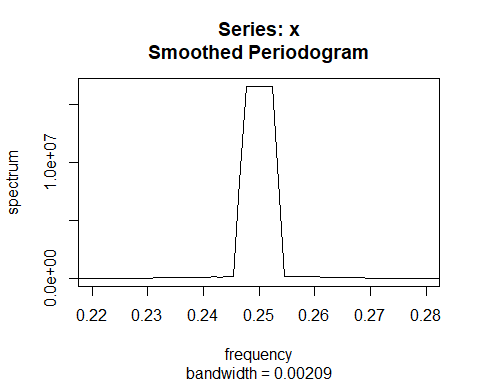
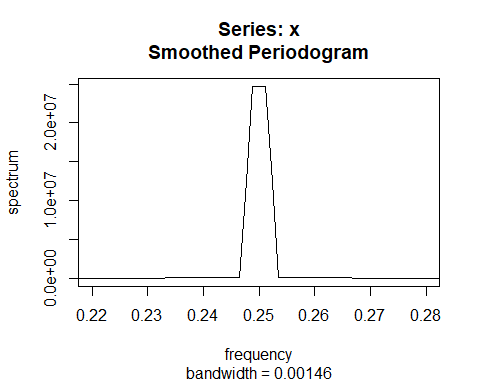


## Daniell(2)   
## coef[-2] = 0.2  
## coef[-1] = 0.2  
## coef[ 0] = 0.2  
## coef[ 1] = 0.2  
## coef[ 2] = 0.2

## mDaniell(1)   
## coef[-1] = 0.25  
## coef[ 0] = 0.50  
## coef[ 1] = 0.25

## Daniell(1,1)   
## coef[-2] = 0.1111  
## coef[-1] = 0.2222  
## coef[ 0] = 0.3333  
## coef[ 1] = 0.2222  
## coef[ 2] = 0.1111

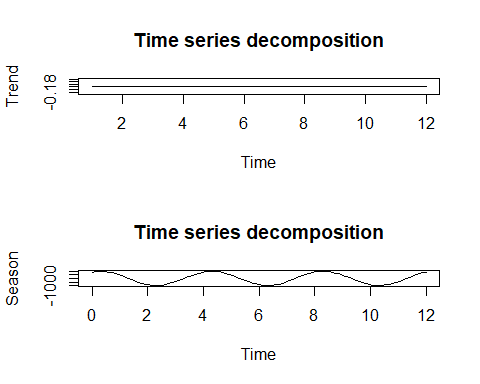
## mDaniell(1,1)   
## coef[-2] = 0.0625  
## coef[-1] = 0.2500  
## coef[ 0] = 0.3750  
## coef[ 1] = 0.2500  
## coef[ 2] = 0.0625



## [1] 216 217 215 218 214 219 213 220 212 221 211 222 210 223 209 224 208 225  
## [19] 207 206 226 205 227 204 228 203 229 202 230 201 231 200 232 199 233 198  
## [37] 234 235 197 236 196 248 250 249 252 251 247 237 253 246 195 254 184 185  
## [55] 183 182 255 181 245 186 256 180 238 179 244 187 194 257 178 177 258 188  
## [73] 398 305 399 397 304 176 385 400 306 307 243 384 386 308 303 401 396 259  
## [91] 387 239 302 321 175 260 193 388 299 322 277 301 298 189 420 300 320 279  
## [109] 278 276 174 309 297 395 419 421 383 261 242 402 156 275 389 280 155 319  
## [127] 173 323 157 418 262 154 394 296 363 422 240 417 390 310 153 190 274 158  
## [145] 393 382 416 324 364 263 281 172 241 403 423 392 295 318 362 152 159 192  
## [163] 391 264 273 191 294 171 381 415 361 365 424 160 311 282 151 374 272 380  
## [181] 325 293 375 265 317 376 360 404 170 266 271 414 366 373 132 169 425 150  
## [199] 379 283 133 292 161 377 267 270 378 333 168 316 372 359 312 131 134 268  
## [217] 269 334 413 284 313 367 291 354 351 326 405 353 162 332 355 352 135 149  
## [235] 412 285 289 290 371 130 426 358 167 288 315 335 314 368 350 286 163 287  
## [253] 165 336 166 406 331 327 357 369 356 411 330 136 129 164 370 337 328 329  
## [271] 148 427 349 407 338 137 410 342 345 348 339 346 347 408 343 409 428 344  
## [289] 341 340 147 128 138 429 431 430 432 139 146 127 140 108 109 107 145 110  
## [307] 106 141 111 105 144 112 142 113 143 104 114 126 103 85 115 84 82 83  
## [325] 86 81 87 88 102 89 80 90 116 125 91 101 92 79 117 64 124 65  
## [343] 93 63 100 62 66 94 61 118 78 95 67 99 123 96 60 97 98 119  
## [361] 68 12 36 13 35 11 77 59 37 34 10 14 33 38 58 69 15 9  
## [379] 120 39 122 32 76 57 16 70 8 121 56 40 31 71 17 7 75 41  
## [397] 55 72 30 18 54 42 6 74 73 29 53 19 43 5 52 28 44 20  
## [415] 4 51 45 27 21 50 46 3 49 22 47 48 26 25 23 24 2 1

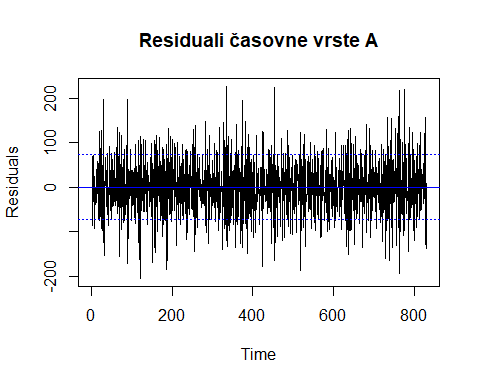
## [1] 0.2500000 0.2511574 0.2488426 0.2523148

## [1] 4.000000 3.981567 4.018605 3.963303



1. Narišite graf rezidualov in komentirajte, ali so videti stacionarni. Stacionarnost tudi preizkusite z uporabo ustreznih statističnih metod.

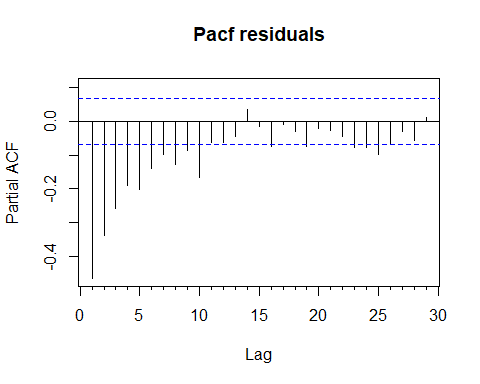
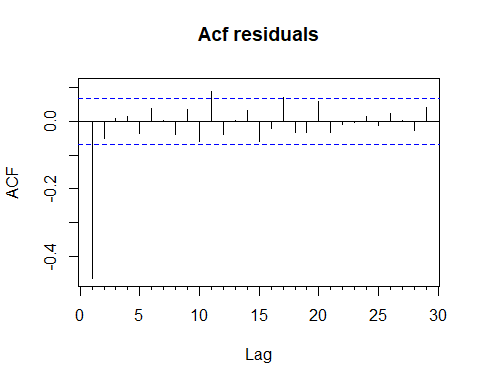
## [1] 0.9889818



##   
## Augmented Dickey-Fuller Test  
##   
## data: d.res  
## Dickey-Fuller = -16.866, Lag order = 9, p-value = 0.01  
## alternative hypothesis: stationary

## [1] "Augmented Dickey-Fuller Test ne zavrne stacionarnosti."

1. Na rezidualih naredite grafikona ACF in PACF in na njuni podlagi predlagajte vsaj en model vrste AR(p) ali MA(q).



## [1] "Predlagava izbiro modela MA(1) in AR(10), torej ARMA(1,10)."

1. Na podlagi Yule–Walkerjevih cenilk in kriterija AIC izberite najboljši model AR(p). Primerjajte ga z najboljšim modelom ARMA(p, q) za p + q ≤ 3 po kriteriju AIC (pozor: kriterij AIC je lahko definiran drugače od postopka do postopka). Če je videti smiselno, pa namesto tega uporabite model GARCH.

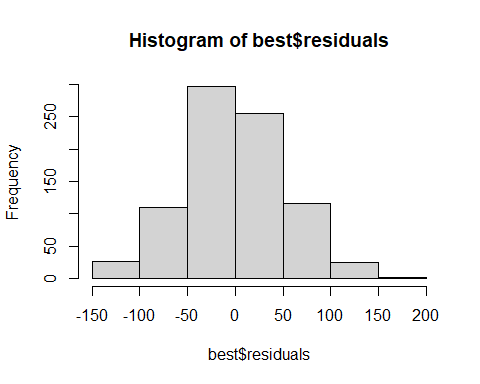
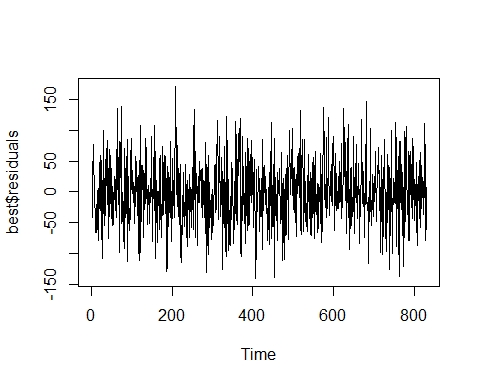
##   
## Call:  
## ar(x = d.res, aic = TRUE, arg = "yule–walker")  
##   
## Coefficients:  
## 1 2 3 4 5 6 7 8   
## -0.9220 -0.9094 -0.8658 -0.8001 -0.7618 -0.6768 -0.6231 -0.5965   
## 9 10 11 12 13 14 15 16   
## -0.5208 -0.4867 -0.3520 -0.3175 -0.2951 -0.2842 -0.3573 -0.3849   
## 17 18 19 20 21 22 23 24   
## -0.3418 -0.3625 -0.3654 -0.3199 -0.3217 -0.3150 -0.2899 -0.2264   
## 25 26   
## -0.1609 -0.0685   
##   
## Order selected 26 sigma^2 estimated as 2852

##   
## Call:  
## arima(x = d.res, order = c(0, 0, 1))  
##   
## Coefficients:  
## ma1 intercept  
## -0.9897 0.0537  
## s.e. 0.0060 0.0214  
##   
## sigma^2 estimated as 2748: log likelihood = -4455.07, aic = 8916.14

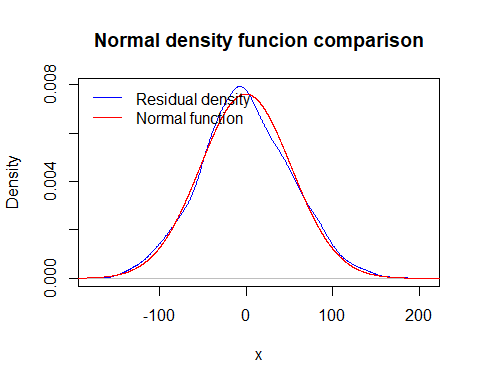
## [1] "Izbrali sva model MA(1)."

1. Izberite »optimalni« model in ocenite vse njegove parametre.

## [1] "OPTIMALNI MODEL: izberemo tisti model, ki ima najnižji aic. V najinem primeru je to MA(1)."

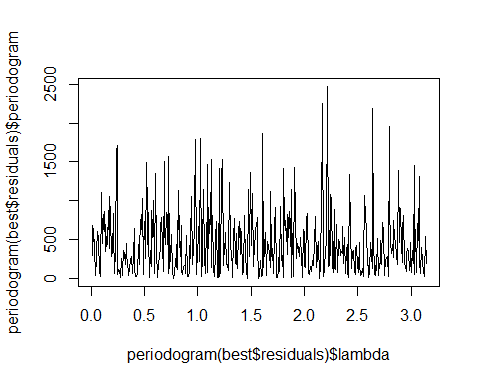


##   
## Shapiro-Wilk normality test  
##   
## data: best$residuals  
## W = 0.99795, p-value = 0.4167



## [1] "Shapirov test ne zavrne hipoteze, torej gre za normalno porazdelitev kar je očitno tudi iz grafa."

1. Oglejte si ostanke po vašem modelu in komentirajte, ali so videti kot beli šum. Primerjajte njihovo porazdelitev z normalno.

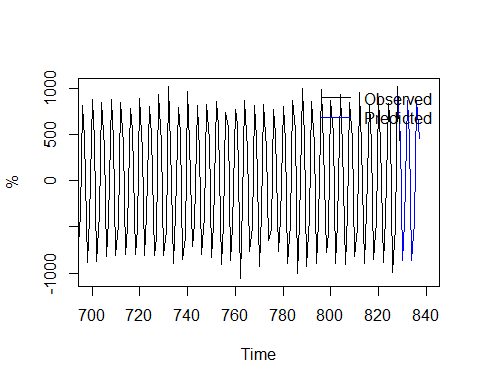
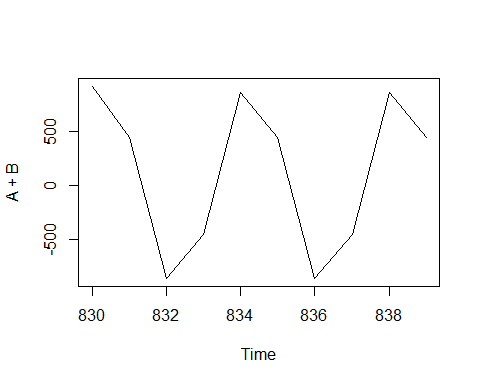
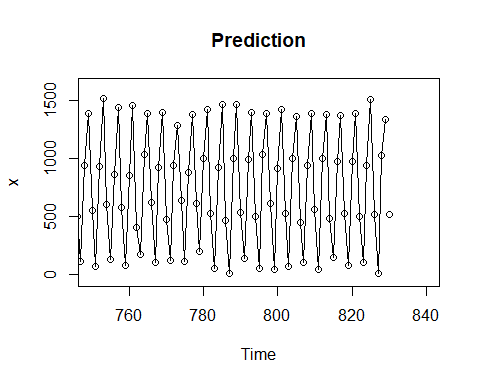
* 

##   
## Box-Pierce test  
##   
## data: d.res  
## X-squared = 179.27, df = 1, p-value < 2.2e-16

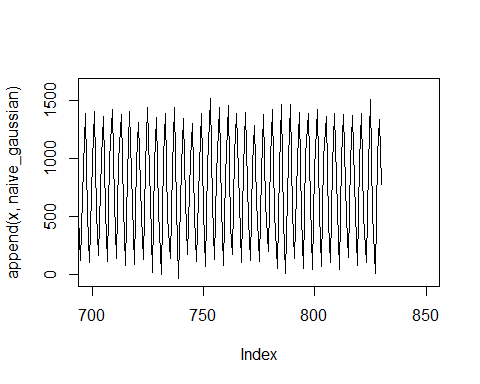
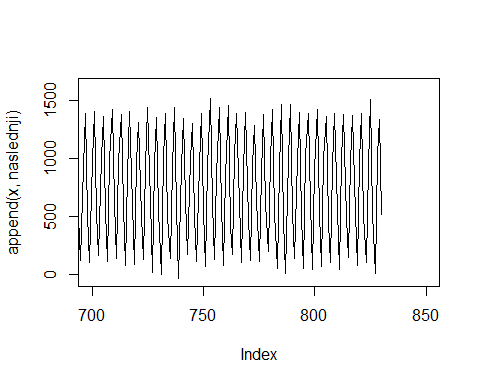
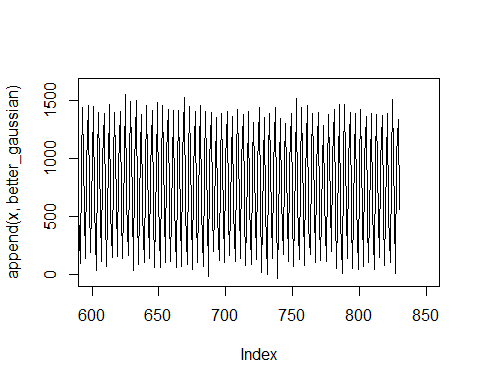
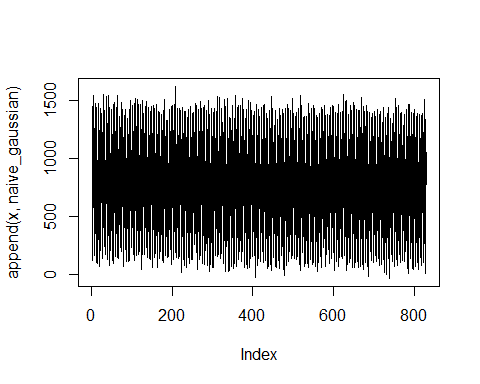
##   
## Box-Ljung test  
##   
## data: d.res  
## X-squared = 179.92, df = 1, p-value < 2.2e-16

## [1] "Ne gre za white noise."

1. Z uporabo izbranega modela in pod predpostavko normalnosti z R-ovo funkcijo predict konstruirajte 90% napovedni interval za naslednjo vrednost. Ne pozabite vračunati tudi odstranjenega trenda in sezonskosti.

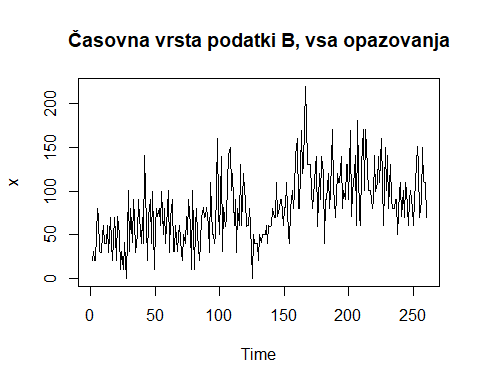


1. Dobljeni napovedni interval primerjajte z napovednim intervalom, ki bi ga dobili, če bi naivno privzeli, da so podatki kar Gaussov beli šum – pred in po odstranitvi trenda in sezonskosti.



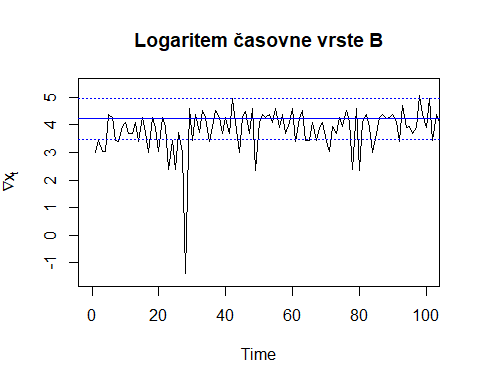
DATOTEKA B, če logaritmiramo

1. *Narišite graf in komentirajte, ali se iz njega vidi kakšen trend ali sezonskost.*

* 

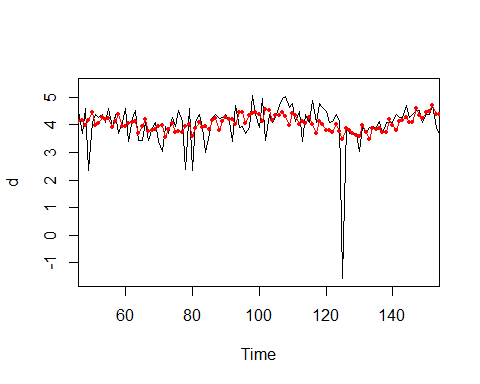
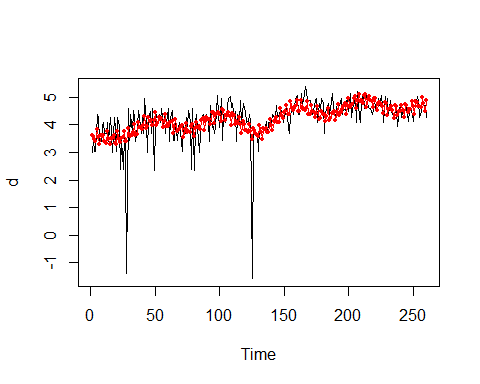
## [1] "Opazimo trend, na prvi pogled ne opazimo sezonskosti."

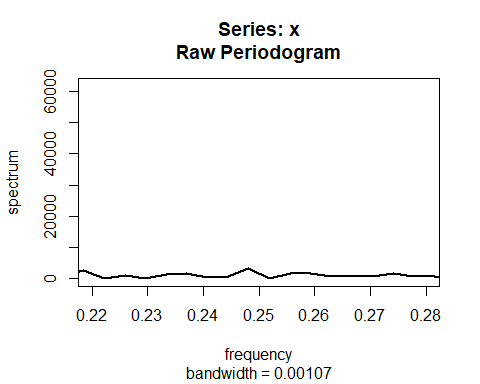
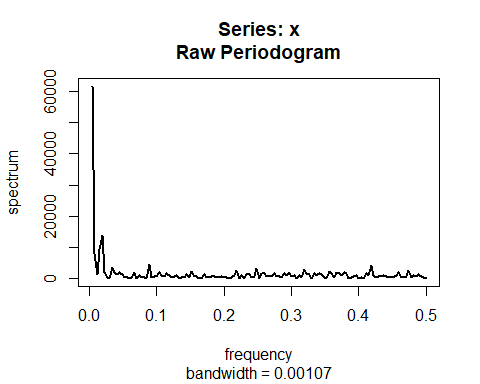
1. *Odstranite morebiten trend in sezonskost z metodami, uporabljenimi pri tečaju: (zaporedno) diferenciranje, logaritmiranje, neposredna ocena sezonskih komponent, polinomski trend stopnje največ 3 ali prileganje periodične funkcije (ali kakšna kombinacija teh metod). Potem ko odstranite morebiten trend, narišite tudi surovi in zglajeni periodogram ter komentirajte, ali se vidi kakšna sezonskost in kakšna naj bi bila perioda.*



##   
## Call:  
## lm(formula = d ~ t)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -5.7489 -0.2438 0.0435 0.3718 1.1512   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.598206 0.083546 43.069 < 2e-16 \*\*\*  
## t 0.004720 0.000555 8.506 1.48e-15 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.6716 on 258 degrees of freedom  
## Multiple R-squared: 0.219, Adjusted R-squared: 0.216   
## F-statistic: 72.35 on 1 and 258 DF, p-value: 1.481e-15

##   
## Call:  
## lm(formula = d ~ t + I(cos(perioda2 \* t)) + I(cos(perioda3 \*   
## t)) + I(cos(perioda4 \* t)) + I(sin(perioda5 \* t)) + I(cos(perioda6 \*   
## t)))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -5.0183 -0.1792 0.0701 0.3551 0.9637   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.5969896 0.0775436 46.387 < 2e-16 \*\*\*  
## t 0.0047298 0.0005151 9.182 < 2e-16 \*\*\*  
## I(cos(perioda2 \* t)) 0.2267701 0.0546668 4.148 4.58e-05 \*\*\*  
## I(cos(perioda3 \* t)) -0.1603372 0.0546668 -2.933 0.00366 \*\*   
## I(cos(perioda4 \* t)) -0.1244814 0.0546668 -2.277 0.02362 \*   
## I(sin(perioda5 \* t)) 0.1724847 0.0546669 3.155 0.00180 \*\*   
## I(cos(perioda6 \* t)) 0.1297305 0.0546668 2.373 0.01839 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.6233 on 253 degrees of freedom  
## Multiple R-squared: 0.3405, Adjusted R-squared: 0.3248   
## F-statistic: 21.77 on 6 and 253 DF, p-value: < 2.2e-16



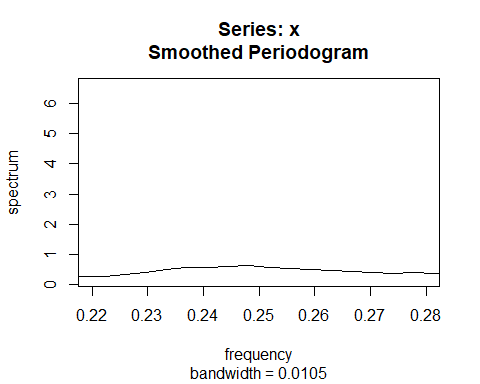
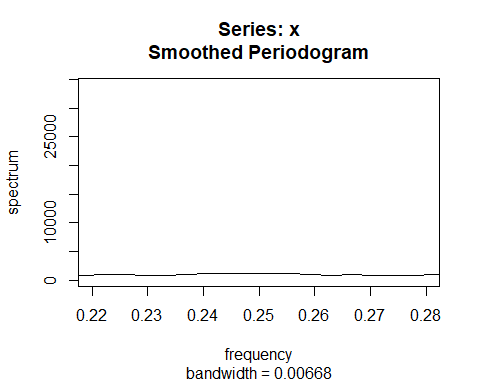
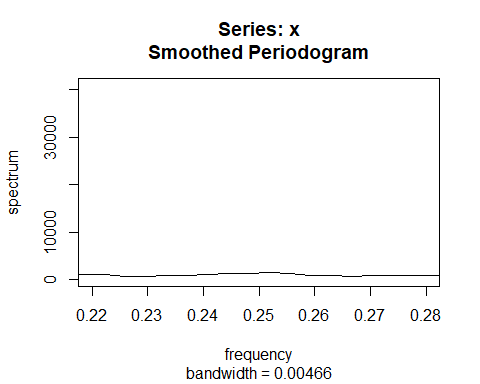


## Daniell(2)   
## coef[-2] = 0.2  
## coef[-1] = 0.2  
## coef[ 0] = 0.2  
## coef[ 1] = 0.2  
## coef[ 2] = 0.2

## mDaniell(1)   
## coef[-1] = 0.25  
## coef[ 0] = 0.50  
## coef[ 1] = 0.25

## Daniell(1,1)   
## coef[-2] = 0.1111  
## coef[-1] = 0.2222  
## coef[ 0] = 0.3333  
## coef[ 1] = 0.2222  
## coef[ 2] = 0.1111

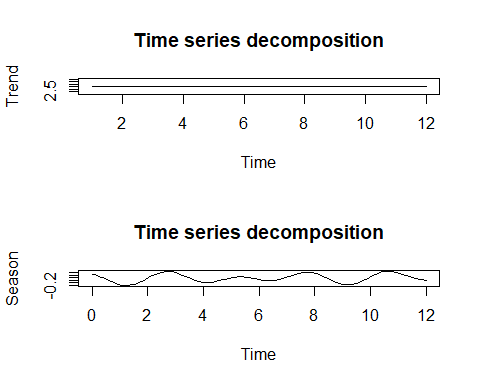
## mDaniell(1,1)   
## coef[-2] = 0.0625  
## coef[-1] = 0.2500  
## coef[ 0] = 0.3750  
## coef[ 1] = 0.2500  
## coef[ 2] = 0.0625



## [1] 1 2 3 4 5 6 7 8 9 10 11 102 101 100 103 114 99 67  
## [19] 113 98 66 12 115 68 104 112 65 13 97 64 69 116 111 105 14 24  
## [37] 70 63 117 96 23 106 110 25 22 71 109 72 108 15 107 118 21 95  
## [55] 26 20 62 73 94 16 119 19 93 75 74 76 92 27 130 131 90 50  
## [73] 18 89 17 129 77 128 91 51 132 52 88 53 133 120 48 49 61 134  
## [91] 127 87 78 47 135 126 125 28 86 124 80 79 85 123 122 54 46 81  
## [109] 121 45 84 42 44 41 39 82 40 43 29 83 55 38 60 59 37 58  
## [127] 56 36 30 57 35 34 31 33 32

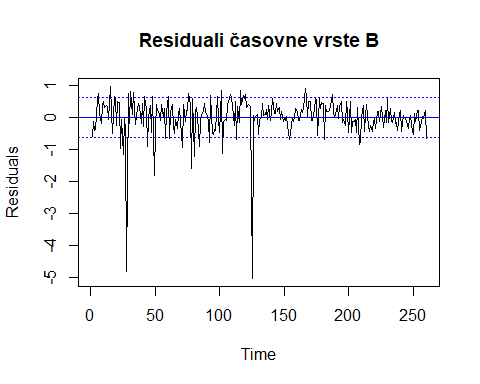
## [1] NA NA NA NA

## [1] NA NA NA NA



1. *Narišite graf rezidualov in komentirajte, ali so videti stacionarni. Stacionarnost tudi preizkusite z uporabo ustreznih statističnih metod.*

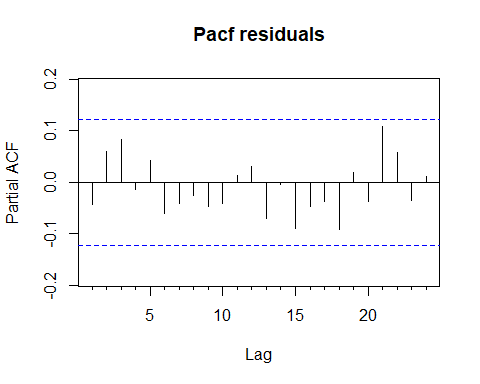
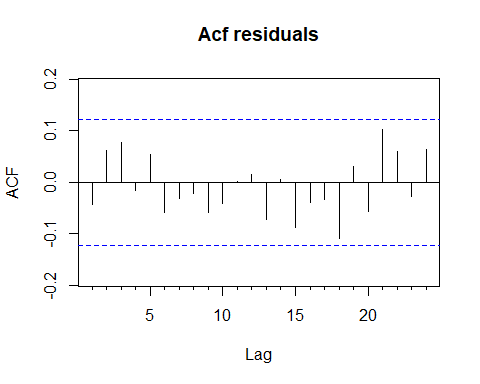
## [1] 0.324818



##   
## Augmented Dickey-Fuller Test  
##   
## data: d.res  
## Dickey-Fuller = -5.9954, Lag order = 6, p-value = 0.01  
## alternative hypothesis: stationary

## [1] "Augmented Dickey-Fuller Test ne zavrne stacionarnosti."

1. *Na rezidualih naredite grafikona ACF in PACF in na njuni podlagi predlagajte vsaj en model vrste AR(p) ali MA(q).*



## [1] "Videti je, da nimamo avtokorelacije."

1. *Na podlagi Yule–Walkerjevih cenilk in kriterija AIC izberite najboljši model AR(p). Primerjajte ga z najboljšim modelom ARMA(p, q) za p + q ≤ 3 po kriteriju AIC (pozor: kriterij AIC je lahko definiran drugače od postopka do postopka). Če je videti smiselno, pa namesto tega uporabite model GARCH.*

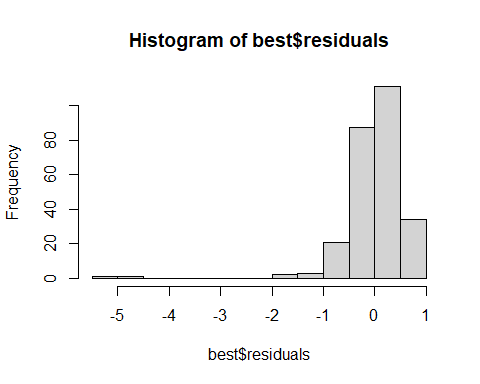
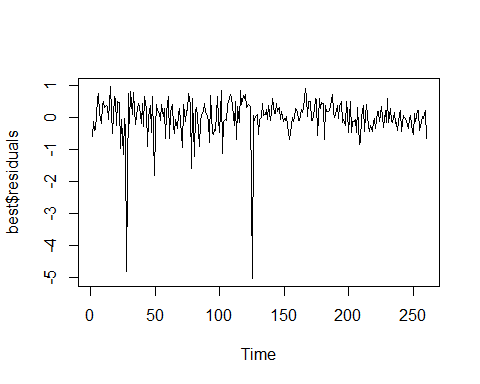
## [1] "Profesor Toman je bil bolj eleganten in se je iskanja rešitve lotil s preprosto zanko."

##   
## Call:  
## arima(x = d.res, order = c(0, 0, 0))  
##   
## Coefficients:  
## intercept  
## 0.0000  
## s.e. 0.0381  
##   
## sigma^2 estimated as 0.378: log likelihood = -242.45, aic = 488.91

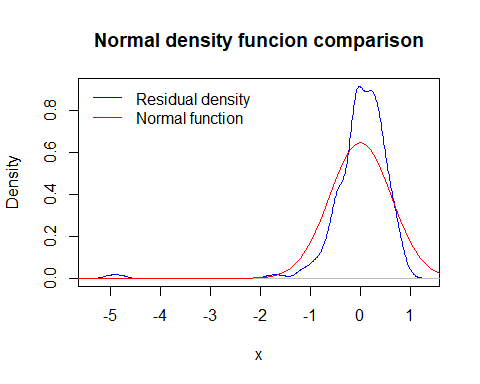
## [1] "Algoritem vrne predlog za model ARMA(0,0)."

1. *Izberite »optimalni« model in ocenite vse njegove parametre.*

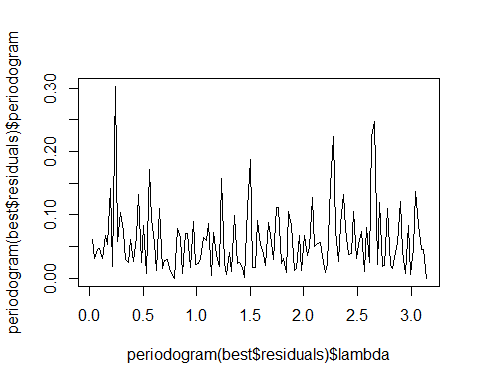
## [1] "OPTIMALNI MODEL: izberemo tisti model, ki ima najnižji aic. V najinem primeru je to ARMA(0,0)."



##   
## Shapiro-Wilk normality test  
##   
## data: best$residuals  
## W = 0.71579, p-value < 2.2e-16



## [1] "Shapirov test zavrne hipotezo, torej ne gre za normalno porazdelitev, kar je očitno tudi z grafa."

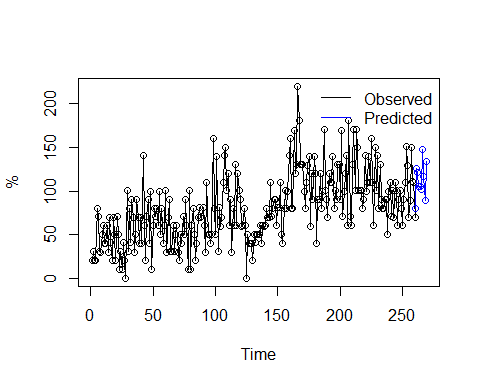
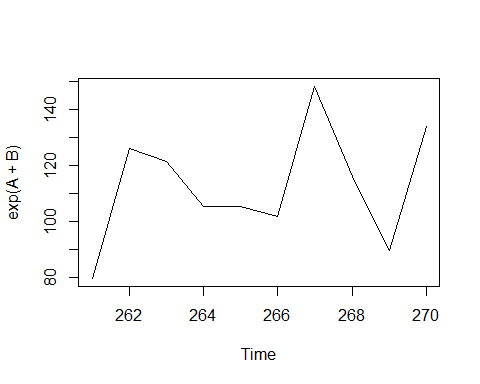
1. *Oglejte si ostanke po vašem modelu in komentirajte, ali so videti kot beli šum. Primerjajte njihovo porazdelitev z normalno.* 

##   
## Box-Pierce test  
##   
## data: d.res  
## X-squared = 0.46321, df = 1, p-value = 0.4961

##   
## Box-Ljung test  
##   
## data: d.res  
## X-squared = 0.46858, df = 1, p-value = 0.4936

## [1] "Imamo white noise!"

1. *Z uporabo izbranega modela in pod predpostavko normalnosti z R-ovo funkcijo predict konstruirajte 90% napovedni interval za naslednjo vrednost. Ne pozabite vračunati tudi odstranjenega trenda in sezonskosti.*



1. *Dobljeni napovedni interval primerjajte z napovednim intervalom, ki bi ga dobili, če bi naivno privzeli, da so podatki kar Gaussov beli šum – pred in po odstranitvi trenda in sezonskosti.*

