Metody Bayesowskie - Projekt 2

Estera Saidło, Wiktoria Szczypka, Paweł Warchoł

25 maja 2020

Cel projektu i opis danych

Celem projektu jest zastosowanie wnioskowania bayesowskiego w regresji wielorakiej na podstawie danych “Boston Housing”, które są wbudowane w R i znajdują się w pakiecie mlbench. Dane dotyczą wartości domów na przedmieściach Bostonu.

Zostały podzielone na 3 części:

1. Dane, na podstawie których zostanie stworzony model MNK, który posłuży do wyznaczenia rozkładu a priori (100 obserwacji).
2. Dane użyte do części właściwej badania (400 obserwacji).
3. Dane użyte do prognozy (6 obserwacji).

Prognozy uzyskane za pomocą wnioskowania bayesowskiego zostaną porównane z prognozami Metody Najmniejszych Kwadratów.

Zmienna objaśniana:

MEDV - mediana wartości domów na przedmieściach Bostonu (w tysiącach dolarów).

Zmienne objaśniające:

RM - średnia liczba pokoi w domu,

CRIM - wskaźnik przestępczości na mieszkańca,

LSTAT - procent populacji o niższym statusie społecznym,

PTRATIO - stosunek liczby uczniów do nauczycieli w danym obszarze podmiejskim,

INDUS - odsetek niedetalicznych akrów biznesowych w danym obszarze podmiejskim.

```
##       medv             rm             crim              lstat
##  Min.   : 5.00   Min.   :3.561   Min.   : 0.00632   Min.   : 1.73
##  1st Qu.:17.02   1st Qu.:5.886   1st Qu.: 0.08204   1st Qu.: 6.95
##  Median :21.20   Median :6.208   Median : 0.25651   Median :11.36
##  Mean   :22.53   Mean   :6.285   Mean   : 3.61352   Mean   :12.65
##  3rd Qu.:25.00   3rd Qu.:6.623   3rd Qu.: 3.67708   3rd Qu.:16.95
##  Max.   :50.00   Max.   :8.780   Max.   :88.97620   Max.   :37.97
##     ptratio          indus
##  Min.   :12.60   Min.   : 0.46
##  1st Qu.:17.40   1st Qu.: 5.19
##  Median :19.05   Median : 9.69
##  Mean   :18.46   Mean   :11.14
##  3rd Qu.:20.20   3rd Qu.:18.10
##  Max.   :22.00   Max.   :27.74
```

Zmienna objaśniana znajduje się w przedziale od 5 do 50 tysięcy dolarów. Wartość maksymalna wydaje się być wartością skrajną, gdyż 3 kwantyl jest na poziomie 25 tysięcy dolarów. Oznacza, to, że 75% obserwacji przyjmuje wartość 25 tysięcy dolarów lub mniej. Średnia liczba pokoi w domu znajduje się w przedziale między 3.5 a 8.8. Średnio w domu znajduje się ok. 6 pokoi. Średni wskaźnik przestępczości na mieszkańca wynosi ponad 3, co oznacza, że średnio na 1 mieszkańca przedmieść Bostonu przypadają między 3 a 4 przestępstwa. Wartość maksymalna to prawie 89 przestępstw na jednego mieszkanca, jednak tutaj także można zauważyć, że wartość ta będzie outlierem. Średnio jest 12% ludzi o niższym statusie społecznym, maksymalnie jest to prawie 38%, a minimalnie mniej niż 2%. Stosunek liczby uczniów do liczby nauczycieli jest między 12.6 a 22. Średnio na 1 nauczyciela przypada ok. 18 uczniów. Odsetek niedetalicznych akrów biznesowych znajduje się w przedziale od 0.26% do 27.74%. Średnio jest to 11.14% akrów na miasto.

Rozkład a priori

Model stworzony za pomocą Metody Najmniejszych Kwadratów służący do oszacowania parametrów rozkładu a priori:

```
##
## Call:
## lm(formula = medv ~ ., data = data0)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.8136 -1.5335 -0.1079  1.2961  7.1484
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -13.93609    6.81186  -2.046   0.0436 *
## rm            7.02718    0.67941  10.343  < 2e-16 ***
## crim         -2.47209    1.10305  -2.241   0.0274 *
## lstat        -0.27686    0.06769  -4.090 9.09e-05 ***
## ptratio      -0.15066    0.21677  -0.695   0.4888
## indus        -0.15993    0.08693  -1.840   0.0689 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.458 on 94 degrees of freedom
## Multiple R-squared:  0.8368, Adjusted R-squared:  0.8281
## F-statistic: 96.41 on 5 and 94 DF,  p-value: < 2.2e-16
```

Zmiennymi najbardziej wpływającymi na kształtowanie się wartości domów w Bostonie w modelu nr. 1 są: średnia liczba pokoi w domu oraz odsetek ludzi o niższym statusie. Średnia liczba pokoi jest jedyną zmienną, której wzrost powoduje wzrost zmiennej objaśnianej. Wzrost pozostałych zmiennych wpływa negatywnie na wzrost wartości domów. Jednak stosunek liczby do nauczycieli nie jest istotny w kontekście kształtowania się wartości domów w Bostonie. Współczynnik determinacji R^2 jest na poziomie prawie 84%. Można więc przyjąć, że taki zestaw zmiennych objaśniających całkiem dobrze opisuje zmienność zmiennej objaśnianej.

Parametry rozkładu a priori zostały dobrane na podstawie modelu MNK stworzonym z pierwszej cześci danych. Wektor β_0 to oszacowane współczynniki modelu. Macierz X_0 to zmienne objaśniające pierwszej części danych, które poprzedza kolumna 1, reprezentująca stałą. Macierz $SIGMA_0$ to odwrotność macierzy $X_0^T X_0$, α_0 to różnica liczby obserwacji i liczby kolumn pierwszej czesci danych, a δ_0 to suma kwadratów reszt z modelu.

Model MNK

Został stworzony model MNK na danych przeznaczonych do właściwej części badania. Na jego podstawie zostanie sprawdzone spełnienie założeń dotyczących reszt oraz dokonana będzie prognoza dla danych z części 3.

```
##
## Call:
## lm(formula = medv ~ ., data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -14.370  -3.592  -1.033   1.752  28.576
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 20.16234    4.60905   4.375 1.56e-05 ***
## rm           4.18465    0.49222   8.502 3.92e-16 ***
## crim        -0.07333    0.03449  -2.126   0.0341 *
## lstat       -0.59202    0.06030  -9.818  < 2e-16 ***
## ptratio     -0.86674    0.14633  -5.923 6.88e-09 ***
## indus        0.01442    0.05306   0.272   0.7860
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.642 on 394 degrees of freedom
## Multiple R-squared:  0.679,  Adjusted R-squared:  0.675
## F-statistic: 166.7 on 5 and 394 DF,  p-value: < 2.2e-16
```

Model MNK stworzony na drugiej części danych wskazuje na dużą istotność zmiennych: średnia liczba pokoi w domu, odsetek ludzi o niższym statusie oraz stosunek uczniów do nauczycieli. W modelu nr. 1 stosunek uczniów do nauczycieli okazał się być nieistotny. Taka rozbieżność wynika najprawdopodobniej z nierównomiernego rozkładu danych. Można więc wnioskować, że informacja dotycząca tej zmiennej w rozkładzie a priori nie będzie pomocna. Nieistotna w tym modelu okazała się zmienna dotycząca odsetka niedetalicznych akrów biznesowych na danym obszarze podmiejskim, natomiast w poprzednim modelu była istotna na poziomie 10%. Wskaźnik przestępczości jest istotny na poziomie 5% w obu modelach.

Interpretacja oszacowań parametrów:

- Przy wzroście średniej liczby pokoi w domu o 1 jednostkę, mediana wartości domów na przedmieściach Bostonu wzrasta o ok. 4.18 tysięcy dolarów przy założeniu, że pozostałe zmienne są na stałym poziomie.
- Przy wzroście wskaźnika przestępczości na mieszkańca o 1 jednostkę, mediana wartości domów na przedmieściach Bostonu maleje o ok. 0.07 tysięcy dolarów przy założeniu, że pozostałe zmienne są na stałym poziomie.
- Przy wzroście procentu populacji o niższym statusie społecznym o 1 jednostkę, mediana wartości domów na przedmieściach Bostonu maleje o ok. 0.59 tysięcy dolarów przy założeniu, że pozostałe zmienne są na stałym poziomie.
- Przy wzroście stosunku liczby uczniów do nauczycieli na danym obszarze podmiejskim o 1 jednostkę, mediana wartości domów na przedmieściach Bostonu maleje o ok. 0.87 tysięcy dolarów przy założeniu, że pozostałe zmienne są na stałym poziomie.
- Przy wzroście odsetka niedetalicznych akrów biznesowych na danym obszarze podmiejskim o 1 jednostkę, mediana wartości domów na przedmieściach Bostonu wzrasta o ok. 0.01 tysięcy dolarów przy założeniu, że pozostałe zmienne są na stałym poziomie.

Sprawdzenie założeń dotyczących reszt modelu:

Normalność reszt została zbadana za pomocą testu Shapiro-Wilka, którego H_0 stanowi o normalności reszt.

```
##
##  Shapiro-Wilk normality test
##
## data:  modelMNK$residuals
## W = 0.88456, p-value < 2.2e-16
```

P-value jest mniejsze od 5%, co oznacza, że H_0 należy odrzucić na rzecz H_1, a więc reszty nie mają rozkładu normalnego. Jednak ze względu na dużą liczbę obserwacji, można założyć asymptotyczną normalność reszt.

W celu zbadania autokorelacji zostanie przeprowadzony test Durbina-Watsona, którego H_0 stanowi o braku autokorelacji reszt.

```
## Non-constant Variance Score Test
## Variance formula: ~ fitted.values
## Chisquare = 0.5711502, Df = 1, p = 0.4498
```

Wartość p-value wskazuje na odrzucenie H_0 na rzecz H_1, co oznacza, że występuje autokorelacja reszt.

Heteroskedastycznośc została zbadana na podstawie testu Breucha-Pagana o następujących hipotezach:

H_0: Homoskedastyczność.

H_1: Heteroskedastyczność.

```
##
##  Durbin-Watson test
##
## data:  modelMNK
## DW = 0.88942, p-value < 2.2e-16
## alternative hypothesis: true autocorrelation is greater than 0
```

Wartość p-value wskazuje na to, że nie ma podstaw by odrzucić H_0, co oznacza homoskedastyczność modelu.

Badane dane są przekrojowe, więc nie ma sensu sprawdzać dla nich autokorelacji.

Rozkład a posteriori

Na podstawie rozkładu a priori wyznaczone zostaną parametry rozkładu a posteriori według poniższych wzorów.

$$\Sigma_1 = (X'X + \Sigma_0^{-1})^{-1}$$

$$\beta_1 = \Sigma_1(X'y + \Sigma_0^{-1}\beta_0) = \Sigma_1\left((X'X)\hat{\beta} + \Sigma_0^{-1}\beta_0\right)$$

$$\alpha_1 = \alpha_0 + n$$

$$\delta_1 = \delta_0 + y'y - \beta_1{}'\Sigma_1^{-1}\beta_1 + \beta_0'\Sigma_0^{-1}\beta_0$$

Funkcja `calc_pars` zwraca obliczone parametry rozkładu.

Wartość parametru α_1:

```
## [1] 494
```

Wartość parametru δ_1:

```
## [1] 13450.4
```

Macierz Σ_1:

```
##               [,1]          [,2]          [,3]          [,4]          [,5]
## [1,]  0.5881812108 -5.373649e-02  8.664709e-04 -2.981236e-03 -1.173988e-02
## [2,] -0.0537364911  6.746705e-03 -5.669291e-05  3.912600e-04  3.504934e-04
## [3,]  0.0008664709 -5.669291e-05  3.595983e-05 -1.619702e-05 -1.872534e-05
## [4,] -0.0029812362  3.912600e-04 -1.619702e-05  9.190145e-05 -1.216054e-05
## [5,] -0.0117398763  3.504934e-04 -1.872534e-05 -1.216054e-05  5.508685e-04
## [6,]  0.0002230869  1.139019e-05 -8.172026e-06 -3.240795e-05 -3.513546e-05
##               [,6]
## [1,]  2.230869e-04
## [2,]  1.139019e-05
## [3,] -8.172026e-06
## [4,] -3.240795e-05
## [5,] -3.513546e-05
## [6,]  7.135106e-05
```

Macierz β_1:

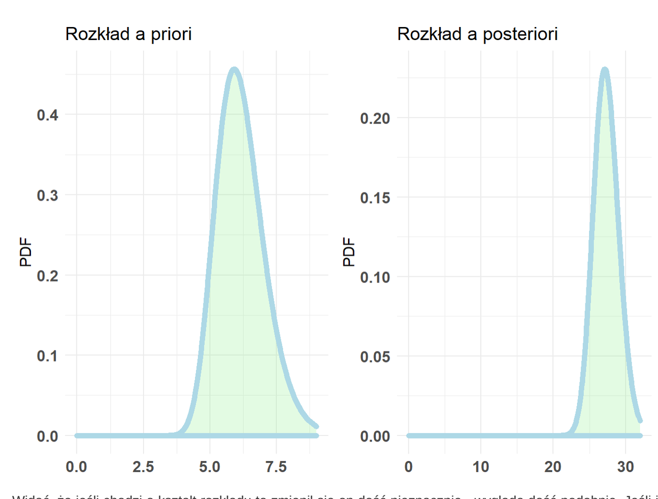
```
##             [,1]
## [1,] 16.56966954
## [2,]  4.61586479
## [3,] -0.07057019
## [4,] -0.55315826
## [5,] -0.86868535
## [6,]  0.02587626
```

Rozkłady brzegowe parametrów

Na podstawie powyżej obliczonych parametrów wielowymiarowego rozkładu a posteriori obliczone zostały parametry rozkładów brzegowych σ^2 oraz poszczególnych β.

Poniżej przedstawiono rozkład a priori i a posteriori σ^2. W tym celu skorzystano z poniższego wzoru.

$$f(\sigma^2|y) \sim IG\left(\frac{\alpha_1}{2}, \frac{\delta_1}{2}\right)$$

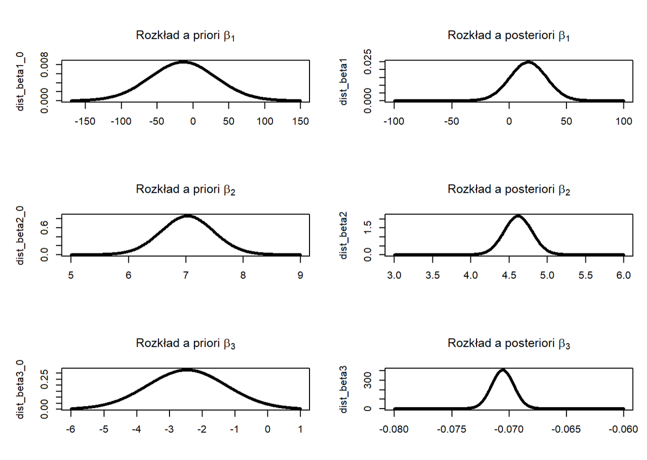


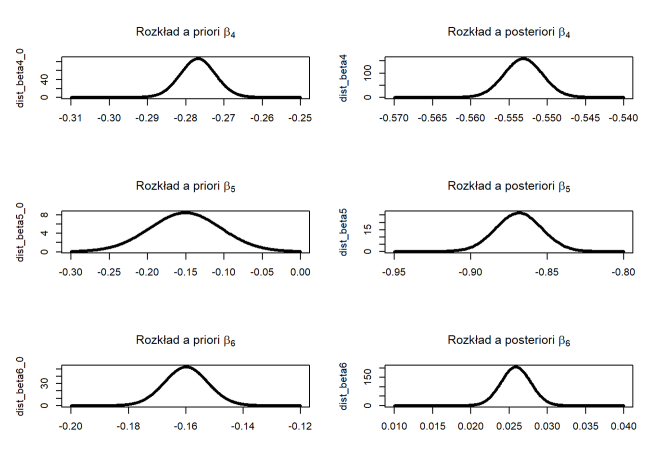
Widać, że jeśli chodzi o kształt rozkładu to zmienił się on dość nieznacznie - wygląda dość podobnie. Jeśli jednak spojrzymy na oś X, to widać, ze uległy one znacznej zmianie. Na podstawie rozkładu a priori spodziewalibyśmy się, ze przyjmuje on wartość między 3.5, a 8.5, natomiast rozkład a posteriori sugeruje, ze jest to przedział 20-35.

Funkcja `rozkl_brzeg_beta` oblicza paramtery rozkładów brzegowych dla poszczególnych β. Funkcja oblicza rozkłady na podstawie poniższego wzoru.

$$f(\beta_i|y) \sim t\left(\beta_{1i}, \frac{\delta_1}{\alpha_1}\Sigma_{1,ii}, \alpha_1\right)$$

Funkcja ta zostanie zastosowana zarówno do obliczenia brzegowych rozkładów (dla poszczególnych β) a priori jak i a posteriori. Uzyskane zostaną 3 parametry rozkładu t: wartość oczekiwana, skala oraz stopnie swobody. W celu ich wizualizacji posłużono się wbudowaną funkcją `dt.scaled`.





Z powyższych wykresów można stwierdzić, iż postacie rozkładu uległy zmianie. Gęstości rozkładu a posteriori są 'węższe' niż w przypadku gęstości a priori, tzn. rozstępy przedziałów, w których na X% spodziewamy się wartości danego parametru są mniejsze.

W przypadku β_6 można zauważyć zmianę znaku parametru. Pozostałe rozkład parametrów ulegały mniejszym lub większym zmianom jeśli chodzi o ich wartości, ale spodziewany znak parametru pozostawał ten sam.

Bayesowskie estymatory

Bayesowskie estymatory, przy założeniu kwadratowej funkcji straty, są równe wartości oczekiwanej rozkładów brzegowych poszczególnych parametrów. Są już więc one wyznaczone.

```
## (Intercept)          rm        crim       lstat     ptratio       indus
## 16.56966954  4.61586479 -0.07057019 -0.55315826 -0.86868535  0.02587626
```

Interpretacja oszacowań parametrów uzyskanych za pomocą metod Bayesowskich:

- Przy wzroście średniej liczby pokoi w domu o 1 jednostkę, mediana wartości domów na przedmieściach Bostonu wzrasta o ok. 4.61 tysięcy dolarów przy założeniu, że pozostałe zmienne są na stałym poziomie.
- Przy wzroście wskaźnika przestępczości na mieszkańca o 1 jednostkę, mediana wartości domów na przedmieściach Bostonu maleje o ok. 0.07 tysięcy dolarów przy założeniu, że pozostałe zmienne są na stałym poziomie.
- Przy wzroście procentu populacji o niższym statusie społecznym o 1 jednostkę, mediana wartości domów na przedmieściach Bostonu maleje o ok. 0.55 tysięcy dolarów przy założeniu, że pozostałe zmienne są na stałym poziomie.
- Przy wzroście stosunku liczby uczniów do nauczycieli na danym obszarze podmiejskim o 1 jednostkę, mediana wartości domów na przedmieściach Bostonu maleje o ok. 0.87 tysięcy dolarów przy założeniu, że pozostałe zmienne są na stałym poziomie.
- Przy wzroście odsetka niedetalicznych akrów biznesowych na danym obszarze podmiejskim o 1 jednostkę, mediana wartości domów na przedmieściach Bostonu wzrasta o ok. 0.03 tysięcy dolarów przy założeniu, że pozostałe zmienne są na stałym poziomie.

W celu zbadania istotności obliczonych parametrów wyznaczono ich HPDI (*highest posterior density interval*), użyto w tym celu wbudowanej funkcji `hdi`.

```
## $`hpdi_(Intercept)`
##     lower     upper
## -14.84282  47.76004
##
## $hpdi_rm
##    lower    upper
## 4.254188 4.969500
##
## $hpdi_crim
##       lower       upper
## -0.07241834 -0.06856020
##
## $hpdi_lstat
##      lower      upper
## -0.5582556 -0.5485173
##
## $hpdi_ptratio
##      lower      upper
## -0.8988634 -0.8413924
##
## $hpdi_indus
##      lower      upper
## 0.02200444 0.02955220
```

Powyżej przedstawiono Bayesowski odpowiednik 95% przedziału ufności. Można zauważyć iż w przypadku β_1 do przedziału należy 0 - oznacza to, iż wyraz wolny jest niestotny. Pozostałe parametry są istotne.

Poniżej przedstawiono parametry uzyskane za pomocą metody MNK oraz metod Bayesowskich.

MNK:

```
## (Intercept)          rm        crim       lstat     ptratio       indus
## 20.16233510  4.18465106 -0.07333107 -0.59202480 -0.86673573  0.01441730
```

Bayes:

```
## (Intercept)          rm        crim       lstat     ptratio       indus
## 16.56966954  4.61586479 -0.07057019 -0.55315826 -0.86868535  0.02587626
```

Można zauważyć, że uzyskane współczynniki parametrów są do siebie podobne. Większą różnicę można zauważyć jedynie w przypadku wyrazu wolnego.

Prognozowanie

W celu wyznaczenia Bayesowskich prognoz należy wyznaczyć rozkłady predykcyjne. Celem będzie stworzenie prognoz dla odłożonych 6 obserwacji.

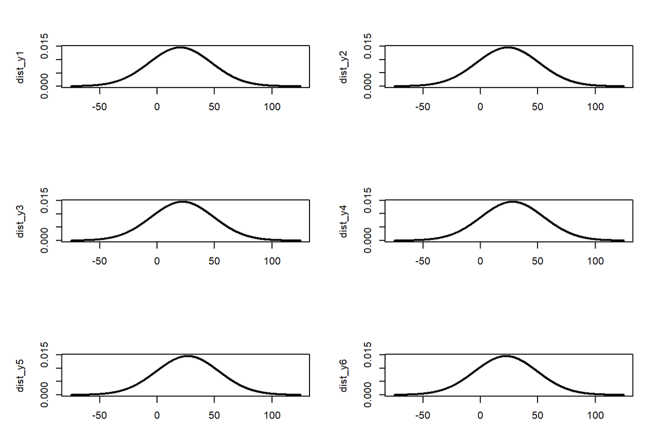
Poniżej wzór na m-wymiarowy rozkład predykcyjny. Analogicznie jak w przypadku wcześniejszych kroków, należy wyznaczyć rozkłady brzegowe poszczególnych y.

$$f(y_\tau|\mathbf{y}) \sim t_m\left(x_\tau\beta_1, \frac{\delta_1}{\alpha_1}[I_m + x_\tau\Sigma_1 x_\tau'], \alpha_1\right)$$

Tak przedstawiają się obserwacje na podstawie których wykonane zostanie prognozowanie:

```
##        rm    crim lstat ptratio indus
## 501 6.027 0.22438 14.33    19.2  9.69
## 502 6.593 0.06263  9.67    21.0 11.93
## 503 6.120 0.04527  9.08    21.0 11.93
## 504 6.976 0.06076  5.64    21.0 11.93
## 505 6.794 0.10959  6.48    21.0 11.93
## 506 6.030 0.04741  7.88    21.0 11.93
```

Poniżej przedstawiono wykresy rozkładów predykcyjnych dla $y_i, i = 1...6$.



Poniżej prognoza poszczególnych obserwacji, która jest równa wartości oczekiwanej (przy założeniu kwadratowej funkcji straty). Przedstawiono również prognozę uzyskaną metodą MNK oraz rzeczywistą wartość.

	MNK	Bayes	rzeczywiste
501	20.38144	20.01888	16.8
502	23.99282	23.71492	22.4
503	22.36404	21.85920	20.6
504	27.98153	27.71215	23.9
505	26.71905	26.40397	22.0
506	22.69770	22.10741	11.9

Można zobserwować, iż prognozy uzyskane za pomocą MNK oraz metod Bayesowskich nie różnią się zbytnio - maksymalna różnica między nimi to ok. 0.6. Jednakże porównując uzyskane prognozy z rzeczywistymi wartościami jednoznacznie widać duże różnice. Prognozy w każdym przypadku są zawyżone.

Dla uzyskanych prognoz obliczony został HPDI. Dla uzyskanego modelu MNK obliczono przedziały ufności na poziomie 95%. Poniżej przedstawiono uzyskane wyniki.

	MNK.lwr	MNK.upr	Bayes.lwr	Bayes.upr
501	19.58899	21.17388	-32.49670	75.53295
502	22.88282	25.10281	-31.12983	79.52882
503	21.22571	23.50237	-29.86800	79.97894
504	26.63723	29.32583	-22.01343	85.56901
505	25.45558	27.98251	-28.05711	81.54663
506	21.46207	23.93333	-26.36014	77.60854

Widzimy, że przedziały modelu MNK są zdecydowanie mniejsze niż HPDI wyznaczone dla podejścia Bayesowskiego. HPDI ma bardzo duży rozstęp - wskazuje, iż średnia cena mieszkania może być ujemna!

	RMSE	MAE	MAPE
MNK	5.385074	4.422762	0.2770948
BAYES	5.029819	4.036088	0.2548130

Pomimo otrzymania ogromnych przedziałów HPDI dla modelu bayesowskiego, błędy prognoz w przypadku obu podejść są podobne - co nie dziwi biorąc pod uwagę, że prognozy były na bardzo zbliżonym poziomie w przypadku obydwu modeli. Nienzacznie mniejsze w każdym przypadku są one jednak dla modelu otrzymaego na podstawie podejścia Bayesowskiego.

Istotność zmiennych crim oraz indus

W kolejnym etapie zbadana zostanie łączna istotność zmiennych crim i indus poprzez zastosowanie tradycyjnego testu F dla modelu MNK oraz poprzez porównanie modeli dla sposobu bayesowskiego.

W tym celu stworzono odpowiedni model bayesowski analogicznie jak we wcześniejszej fazie, z pominięciem 2 wymienionych wcześniej zmiennych. Przyjęto a priori, że nie mamy żadnego przekonania cd. wyższości któregoś z modeli nad drugim - prawdopodobieństwa a priori w takim wypadku są równe P(M1) = P(M2) = 1/2, gdzie jako M1 przyjęto model bez restrykcji, a jako model M2 - z restrykcjami. O wyniku tego porownania świadczył będzie w takim wypadku tzw. czynnik bayesowski. Jednak ze względu na jego charakterystykę (występowanie funkcji gamma) oraz spore wartości liczbowe dla naszych modeli (np wartości wsp. delta większe od 10 000 czy współczynniki alfa na poziomie niemal 500) niezbędne było odpowiednie uproszczenie tego wzoru poprzez skrócenie odpowiednich wartości. Otrzymano wyniki jak poniżej.

```
## [1] "Wartość logarytmu z ilorazu szans a posteriori"
```

```
## [1] 40.92157
```

Widać wyraźnie, że wartość ta jest bardzo duża, co oznacza, że możemy wnioskować o wyższości modelu 1 (bez restrykcji) nad modelem z restrykcjami. Dla modelu MNK przeprowadzono test F. Otrzymano następujące wyniki:

```
## Linear hypothesis test
##
## Hypothesis:
## indus = 0
## crim = 0
##
## Model 1: restricted model
## Model 2: medv ~ rm + crim + lstat + ptratio + indus
##
##   Res.Df   RSS Df Sum of Sq      F Pr(>F)
## 1    396 12686
## 2    394 12542  2    144.17 2.2645 0.1052
```

Powyższy wynik daje nie daje podstaw do odrzucenia H0 mówiącej o braku łączej istotności zmiennych indus i crim - jest to więc wynik odmienny do tego otrzymanego poprzez porównanie modeli bayesowskich. Jednak dla stworzonego modelu MNK zmienne indus i crim nie były istotne na poziomie isotności 1% (zmienna crim była istotna na poziomie 5%, natomiast indus nadal nie). W przypadku podejścia bayesowskiego przedziały HPDI nie wskazywały na brak istotności któregokolwiek z parametrów.

Algorytm Gibbsa

W celu wyznaczenia rozkadów brzegowych a posteriori można również użyć podejścia symulacyjnego i zastosować algorytm Gibbsa, przy założeniu odpowiednich rozkładów warunkowych parametrów $/Beta$ i $sigma^2$. Przeprowadzono 11 111 losowań, w których zgodnie z algorytmem Gibbsa losowano wartości pochodzące z odpowiednich rozkładoW warunkowych naszych nieznanych parametrów, po czym odrzucono 1111 obserwacji początkowych. Na podstawie pozostałych 10 tysięcy wylosowanych określono rozkłady brzegowe a posteriori.

W algorytmie wykorzystano poniższe wzory na rozkłady warunkowe parametrów:

$\beta^{(i)}$ z rozkładu $N_k(\beta_1^{(i)}, \Sigma_1^{(i)})$,

$\sigma^{2(i)}$ z rozkładu $IG\left(\frac{\alpha_1}{2}, \frac{\delta_1^{(i)}}{2}\right)$

Na podstawie otrzymanych rozkładów wyznaczono punktowe oceny parametrów (jako średnie) oraz ich błędy (odchylenia standardowe).

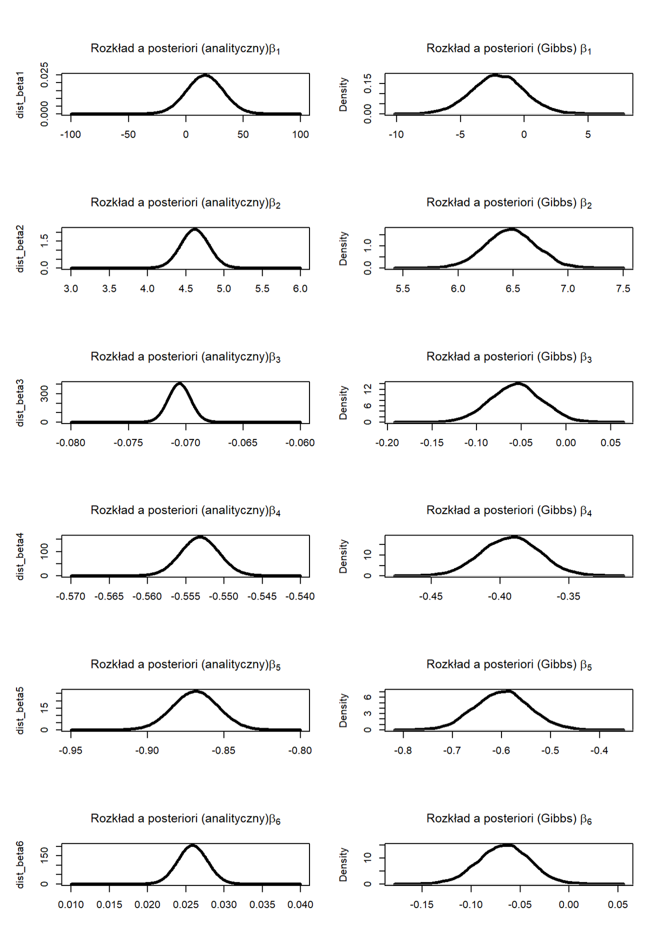
Wyznaczone prognozy:

```
## (Intercept)          rm        crim       lstat     ptratio       indus
##      -2.058       6.478      -0.056      -0.392      -0.598      -0.066
##       sigma
##      29.645
```

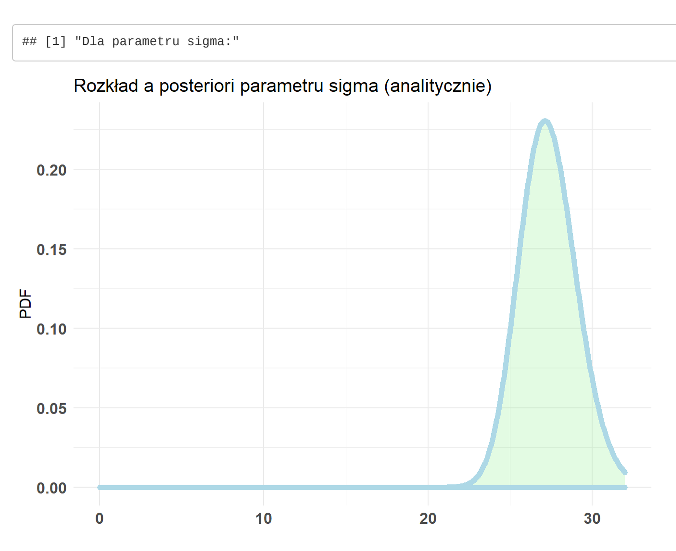
Wyznaczone odchylenia:

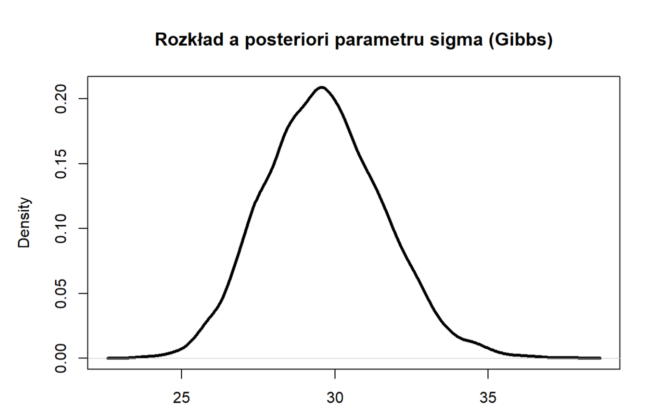
```
## (Intercept)          rm        crim       lstat     ptratio       indus
##      2.0515      0.2311      0.0287      0.0209      0.0559      0.0261
##       sigma
##      1.9397
```

Poniżej przedstawiono również wykresy rozkładów a posteriori uzyskane z podejścia analitycznego (z lewej) i te uzyskane poprzez zastosowanie algorytmu Gibbsa.




```
## [1] "Dla parametru sigma:"
```





Widać, że dla parametrów beta wyznaczone rozkłady są węższe po wyznaczeniu ich analitycznie (oprócz wyrazu wolnego, gdzie jest odwrotnie oraz $Beta_2$, którego postać rozkładu jest podobna). Widać jednak zmiany jeśli chodzi wartości dla poszczególnych parametrów.

```
##              (Intercept)       rm        crim      lstat    ptratio       indus
## Gibbs          -2.057704 6.477775 -0.05590243 -0.3921695 -0.5983451 -0.06628826
## Analitycznie   16.569670 4.615865 -0.07057019 -0.5531583 -0.8686854  0.02587626
##                 sigma
## Gibbs        29.64462
## Analitycznie 16.56967
```

Wartości oszacowań parametrów otzymane tymi 2 metodami rożnią się - widać, to szczególnie na przykładzie wyrazu wolnego, który jest zdecydowanie mniejszy w przypadku wyznaczania go algorytmem Gibbsa. Pozostałe parametry są jednak większe (oprócz indus) niż w przypadku wyznaczania ich analitycznie.