



پیش بینی نرخ قتل در ایالات جنوبی آمریکا (۱۹۹۰-۱۹۶۰)

پروژه درس آنالیز پیشرفته داده ها

حسین شکوهی زاده و ساجده اسماعیل زاده

استاد : دکتر ریخته گران

ترم اول سال تحصیلی ۱۴۰۱_۱۴۰۰



تصویر سازی داده‌ها

Data Visualization

جرائم، اغلب مانعی بزرگ در رشد و توسعه اقتصادی به شمار می‌رود، زیرا تمایل به افزایش با عدم اطمینان اقتصادی، دلسردی برای سرمایه‌گذاری بلند مدت و فرصت‌های شغلی جدید و فرسایش حاکمیت قانون دارد. در عین حال، عدم رشد اقتصادی در ارتباط با نابرابری اقتصادی و اجتماعی بالا، باعث افزایش سطح جرم و خشونت می‌شود.

شرایط اقتصادی اجتماعی و نابرابری اجتماعی نقش مهمی در تأثیر یا عدم برخورد افراد خاص با رفتار مجرمانه دارد.

هدف از این تجزیه و تحلیل، استفاده از داده‌های هر شهر در ایالات جنوبی آمریکا برای تحمین نرخ قتل است. از این برآوردها می‌توان نرخ قتل در سال‌های آتی را پیش‌بینی کرد و عوامل تاثیرگذار بر نرخ قتل در ایالت جنوبی آمریکا را شناسایی نمود.

این مجموعه داده شامل اطلاعات ۱۴۱۲ شهر در ایالات جنوبی آمریکا می‌باشد که ۱۷ فاکتور مختلف (متغیرها) در این شهرها بررسی شده است.

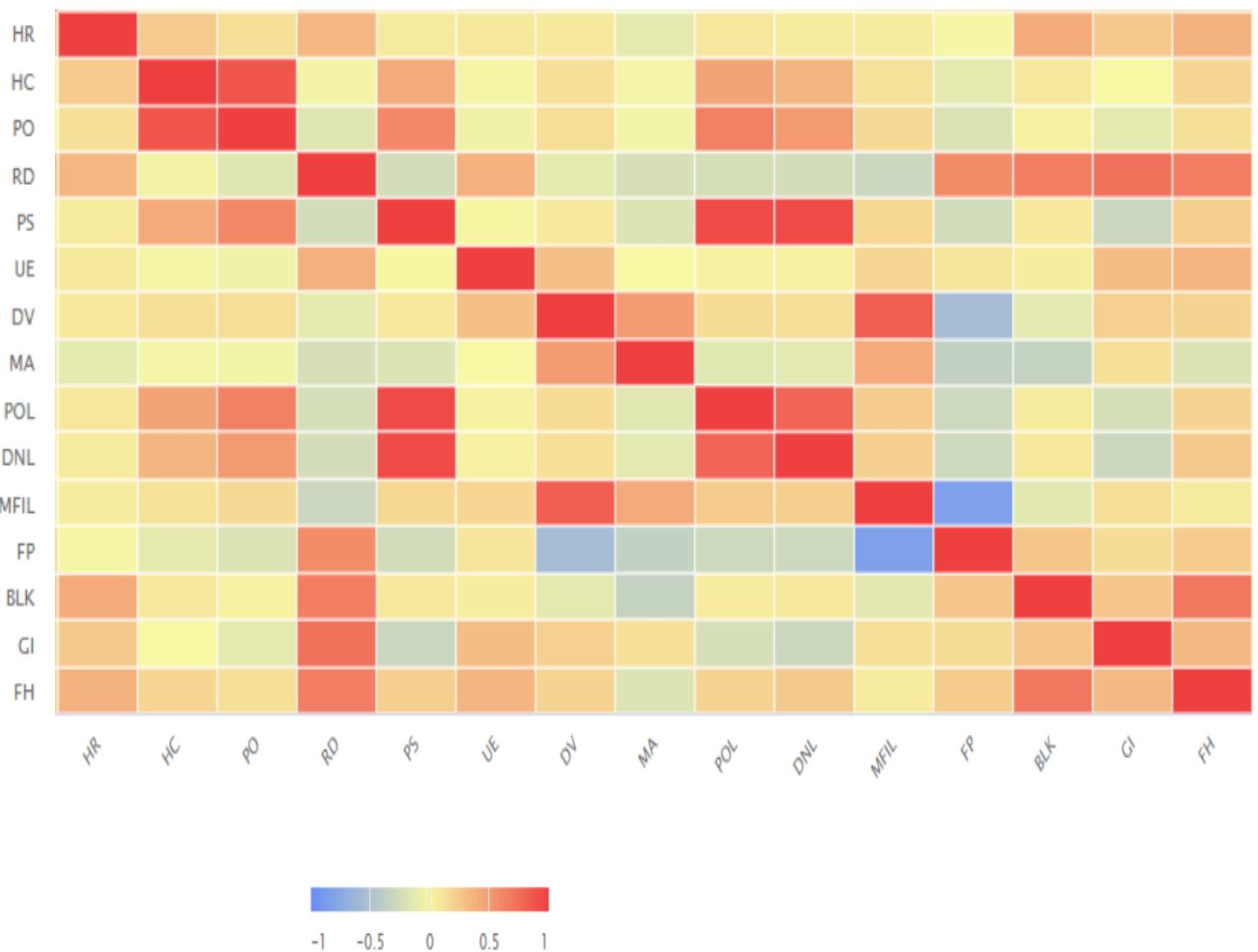
<https://geodacenter.github.io/data-and-lab/south>

معرفی متغیر ها

نام متغیر	نوع متغیر	توضیح متغیر
Name	رده ای	نام شهر ها
State_Name	رده ای	نام ایالت مربوط به شهرمورد نظر
HR**	عددی و اعشاری	نرخ قتل در هر ۱۰۰۰۰۰ نفر
HC**	عددی و اعشاری	میانگین سه ساله شمارش قتل
PO**	عددی و صحیح	جمعیت شهرستان
RD**	عددی و اعشاری	محرومیت از منابع
PS**	عددی و اعشاری	ساختمان جمعیت
UE**	عددی و اعشاری	نرخ بیکاری
DV**	عددی و اعشاری	نرخ طلاق
MA**	عددی و اعشاری	میانه سن
POL**	عددی و اعشاری	گزارش جمعیت
DNL**	عددی و اعشاری	گزارش تراکم جمعیت
MFIL**	عددی و اعشاری	گزارش میانه درآمد خانواده
FP**	عددی و اعشاری	درصد خانواده های زیر فقر
BLK**	عددی و اعشاری	درصد سیاه پوست ها
GI**	عددی و اعشاری	شاخص جینی نابرابری درآمد خانواده
FH**	عددی و اعشاری	درصد زنان سرپرست خانواده

NAME	HR	HC	PO	RD	PS	UE	DV	MA	POL	DNL	MFIL	FP	BLK	GI	FH
Hancock	1.682864235	0.666666667	39615	-1.394676863	1.218684208	3.1	2.277989394	28.9	10.586963	6.168123	8.841014	9.6	3.839454752	0.2236450	9.981297372
Brooke	4.607233356	1.333333333	28940	-1.207106945	0.955640335	4.8	1.504362652	28.5	10.272980	5.795643	8.696845	13.5	1.416724257	0.2204073	10.92933737
Ohio	0.974131927	0.666666667	68437	-0.46337022	1.557221283	7.3	1.993838781	34.1	11.133669	6.470230	8.599326	20.6	3.052442392	0.2723981	15.62164302
Marshall	0.876247558	0.333333333	38041	-0.901189691	0.739644675	11	3.479828088	31.2	10.546420	4.829392	8.548110	19.8	0.946347362	0.2276472	11.96283391
New Castle	4.228384822	13	307446	-0.985395233	2.229292133	4.7	1.751699519	29.5	12.636055	6.551555	8.828055	11.3	11.72205851	0.2561062	12.03571382
Washington	1.827104733	1.666666667	91219	-0.697130662	1.274076721	8.3	2.020578818	30.5	11.421018	5.291968	8.545781	21.3	2.833839441	0.2572662	12.27321238
Allegany	0.79205725	0.666666667	84169	-0.52400984	1.235971903	6.7	1.949102931	32.8	11.340582	5.281459	8.538367	23.2	1.382931958	0.2599558	14.50330233
Cecil	6.197322757	3	48408	-0.803565216	0.865846555	6.5	1.73869846	25.2	10.787420	4.895776	8.568266	19.9	5.941166749	0.2536662	10.90588454
Garrett	0	0	20420	0.308776649	-0.008407433	10.7	1.299802204	27.8	9.924270	3.433547	8.127995	43.3	0.190989226	0.3396934	12.14037951
Harford	1.303407106	1	76722	-0.957619702	1.144737913	4.5	1.576576577	24.8	11.247944	5.132052	8.676417	16.7	9.671280728	0.2563840	9.680950036

در جدول بالا میتوانید تعدادی از داده ها را مشاهده کنید.

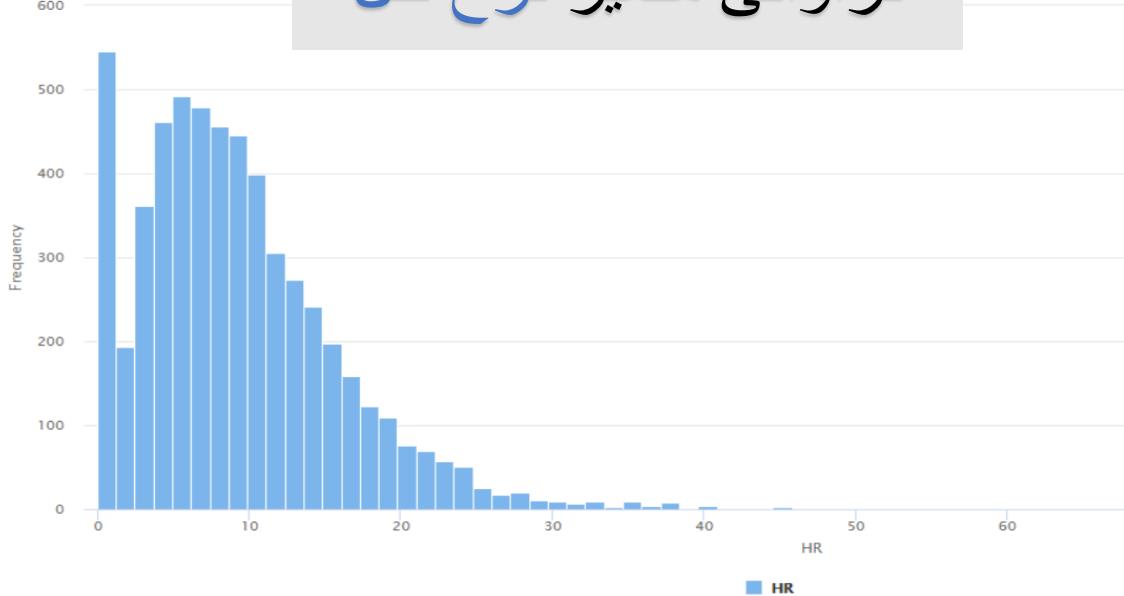


در این شکل همبستگی بین متغیر های را بررسی میکنیم . ضریب همبستگی عددی بین منفی یک و یک است و هر چه این عدد از صفر دورتر باشد ، میزان همبستگی دو متغیر بیشتر است . اعداد منفی به این معناست که دو متغیر با یک دیگر رابطه عکس دارند .

مطابق شکل ، نرخ قتل بیشترین همبستگی را با متغیر های **محرومیت از منابع** و **درصد سیاه پوست ها** و **درصد زنان سرپرست خانواده** دارد .

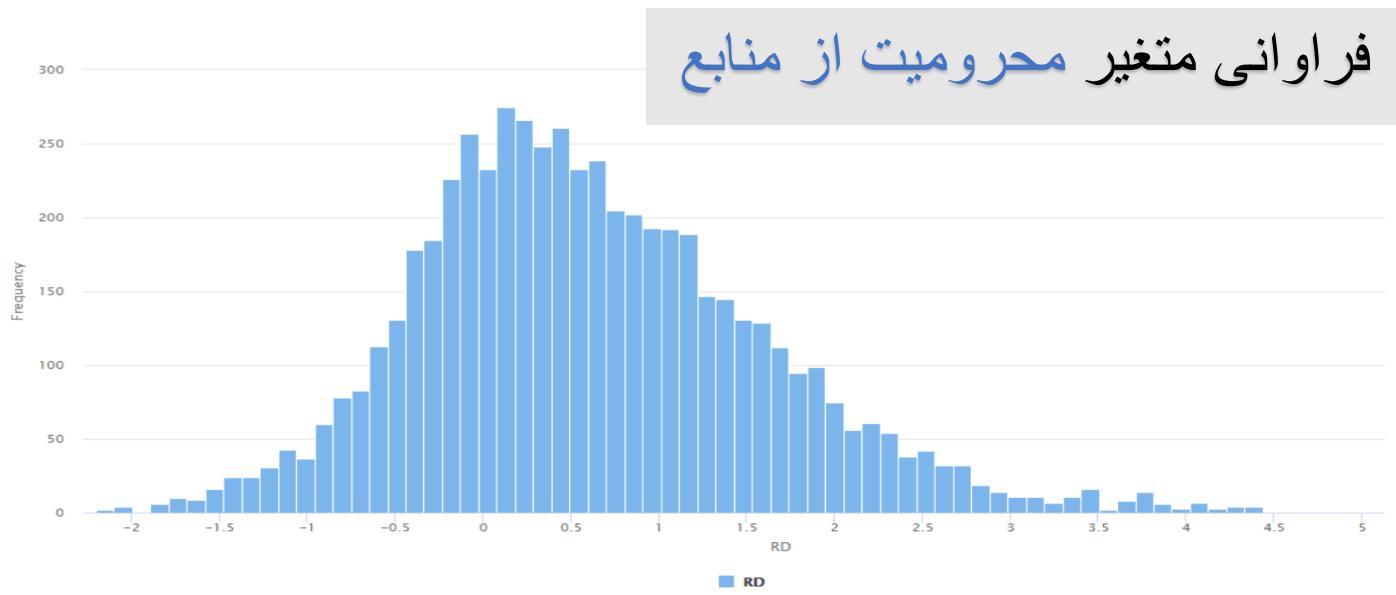
طبق بررسی داده ها ، در این مجموعه داده هیچ داده گم شده ای وجود ندارد!

فراوانی متغیر نرخ قتل



همان طور که در شکل بالا مشاهده می کنید ، بیشتر داده ها در محدوده ۰ تا ۱۱ بوده است .
همچنین میتوان دید توزیع آن به صورت نرمال چوله به راست می باشد .

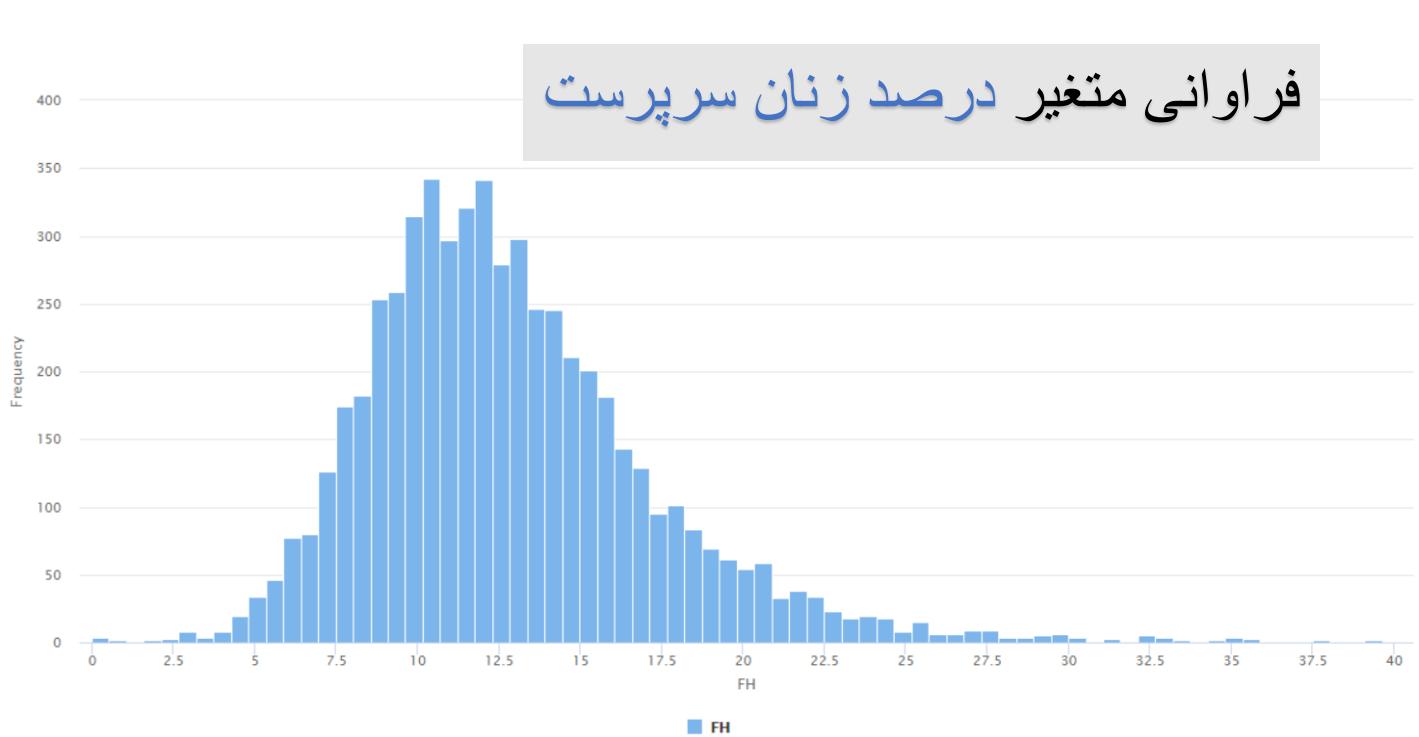
فراوانی متغیر محرومیت از منابع



فراوانی متغیر در صد سیاه پوستان



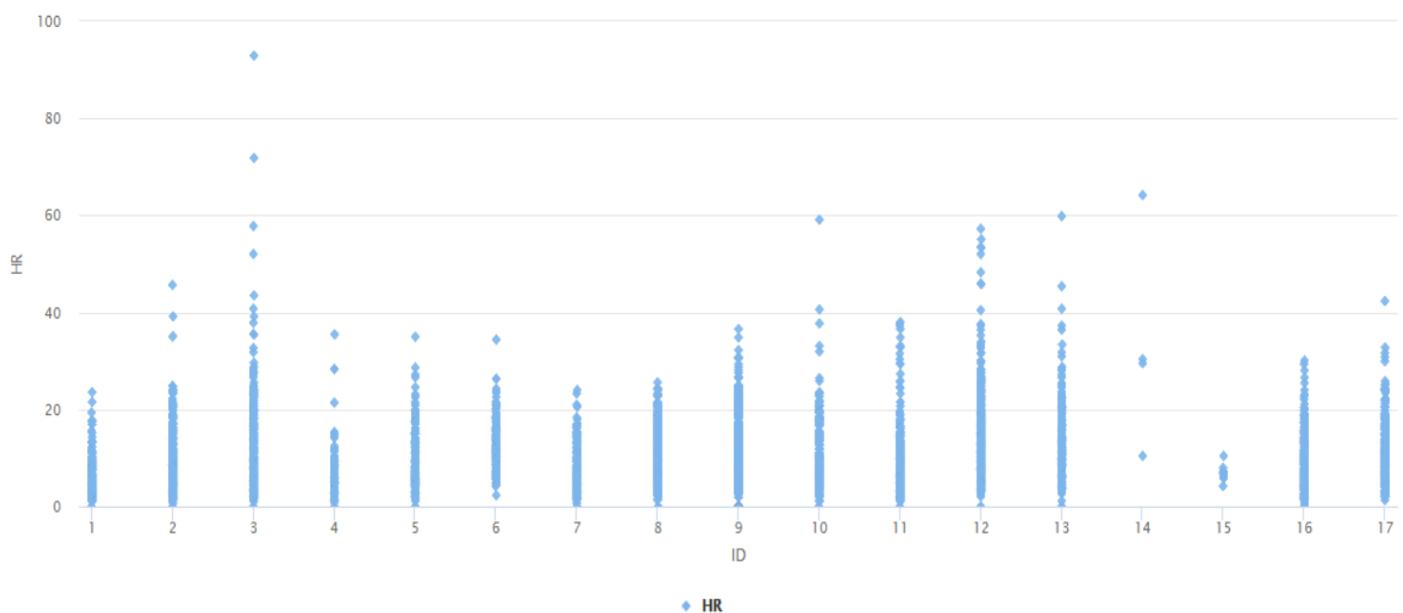
فراوانی متغیر در صد زنان سرپرست



فراوانی متغیر ایالات جنوبی



با توجه به این دو نمودار میتوان گفت که بعضی ایالات بیشتر در این داده ها مشاهده شدند و نرخ قتل در آنها بالاتر است و برخی کمتر مشاهده شده اند و نرخ قتل در آنها پایین تر است.



مدل رگرسیون خطی

Linear Regression
model

هدف از این گزارش ، پیش بینی نرخ قتل با سایر متغیرهای توضیحی می باشد . از آنجایی که متغیرها زیاد می باشند و برخی از آنها به صورتی دیگر در متغیرها هستند از آنها صرف نظر می کنیم .

متغیر هایی که از آنها صرف نظر می کنیم عبارتند از : میانگین سه ساله شمارش قتل(HC) ، گزارش جمعیت(POL) و گزارش تراکم جمعیت(DNL) و شاخص جینی نابرابری درآمد خانواده (GI).

ابتدا به صورت رندوم ۷۵ درصد داده ها را به عنوان داده های آموزشی و ۲۵ درصد داده ها را به عنوان داده آزمایشی در نظر می گیریم.

تمام مدل هایی که روی داده ها اجرا می کنیم ، تنها روی داده های آموزشی ما هستند و از داده های آزمایشی تنها برای بررسی دقیق مدل نهایی استفاده می کنیم و هیچ دخالتی در انتخاب مدل نهایی ما ندارند و در نهایت با استفاده از آماره های DIC و MSE آموزشی مدل ها را با هم مقایسه می کنیم .

تخمین تمامی پارامترهای مجھول با استفاده از رویکرد بیز در نرم افزار open bugs صورت می گیرد.

بخش اول: مدل های رگرسیون بدون اثر تصادفی

مدل اول:

اولین مدل را با استفاده از متغیر های مورد نظر و با در نظر گرفتن توزیع نرمال برای جمله خط اجرا می کنیم.

$$HR_i = X_i * \beta + \epsilon_i$$

$$X_i = [\text{MFIL}, \text{PO}, \text{RD}, \text{PS}, \text{UE}, \text{DV}, \text{MA}, \text{FP}, \text{BLK}, \text{FH}]$$

$$\beta = [\beta_0, \dots, \beta_{11}]$$

$$\epsilon_i \sim N(0, \sigma^2)$$

توزیع پیشین β را نرمال چند متغیره و σ^2 را معکوس گاما در نظر می گیریم.

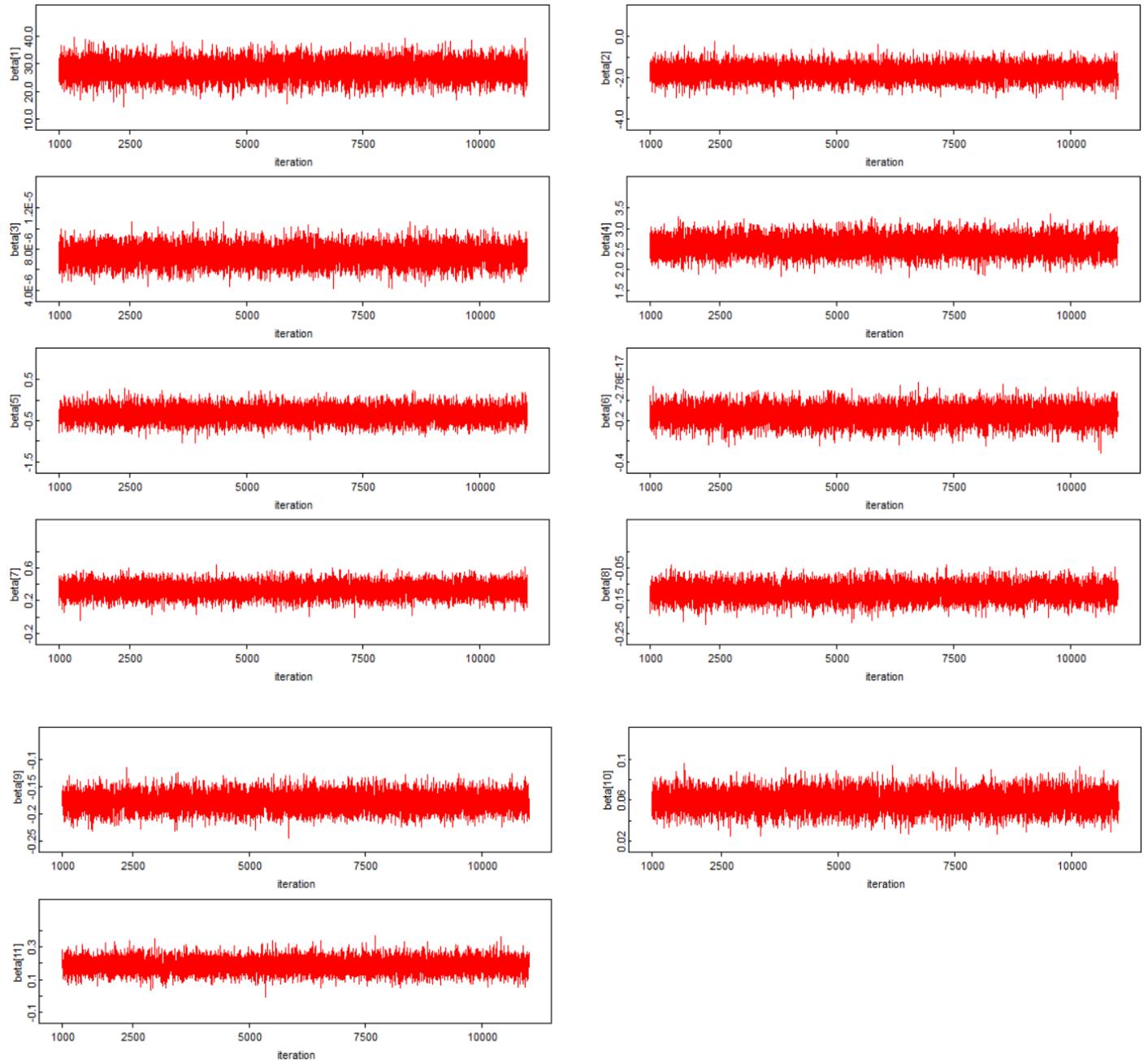
هر دو توزیع را آگاهی نابخش (واریانس زیاد و میانگین صفر) در نظر گرفتیم.

برآورده بیز β ها و واریانس آنها را می توان در جدول زیر مشاهده کرد.

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta[1]	27.8	3.36	0.03111	21.27	27.8	34.27	1001	10000
beta[2]	-1.746	0.3579	0.003289	-2.434	-1.743	-1.055	1001	10000
beta[3]	7.371E-6	8.911E-7	8.107E-9	5.65E-6	7.376E-6	9.119E-6	1001	10000
beta[4]	2.593	0.2119	0.002105	2.186	2.592	3.012	1001	10000
beta[5]	-0.3419	0.178	0.001627	-0.6901	-0.3429	0.01326	1001	10000
beta[6]	-0.1711	0.04255	4.487E-4	-0.2549	-0.1705	-0.0889	1001	10000
beta[7]	0.3307	0.08157	8.045E-4	0.1715	0.331	0.4907	1001	10000
beta[8]	-0.1241	0.02403	2.841E-4	-0.1714	-0.1238	-0.07697	1001	10000
beta[9]	-0.178	0.01575	1.536E-4	-0.2087	-0.1782	-0.1472	1001	10000
beta[10]	0.05894	0.009423	9.085E-5	0.04026	0.05891	0.07746	1001	10000
beta[11]	0.1859	0.0425	3.747E-4	0.1035	0.1858	0.268	1001	10000

همانطور که مشاهده می کنید واریانس تخمین عرض از مبدأ و ضریب گزارش میانه در آمد خانواده (MFIL) زیاد است پس کمتر می توان به آنها اعتماد کرد. همچنین ضریب محرومیت از منابع (RD) نیز خیلی به صفر نزدیک است ولی با توجه به آزمون فرض انجام شده، به وجود این متغیر در حضور سایر متغیرها نیاز است.

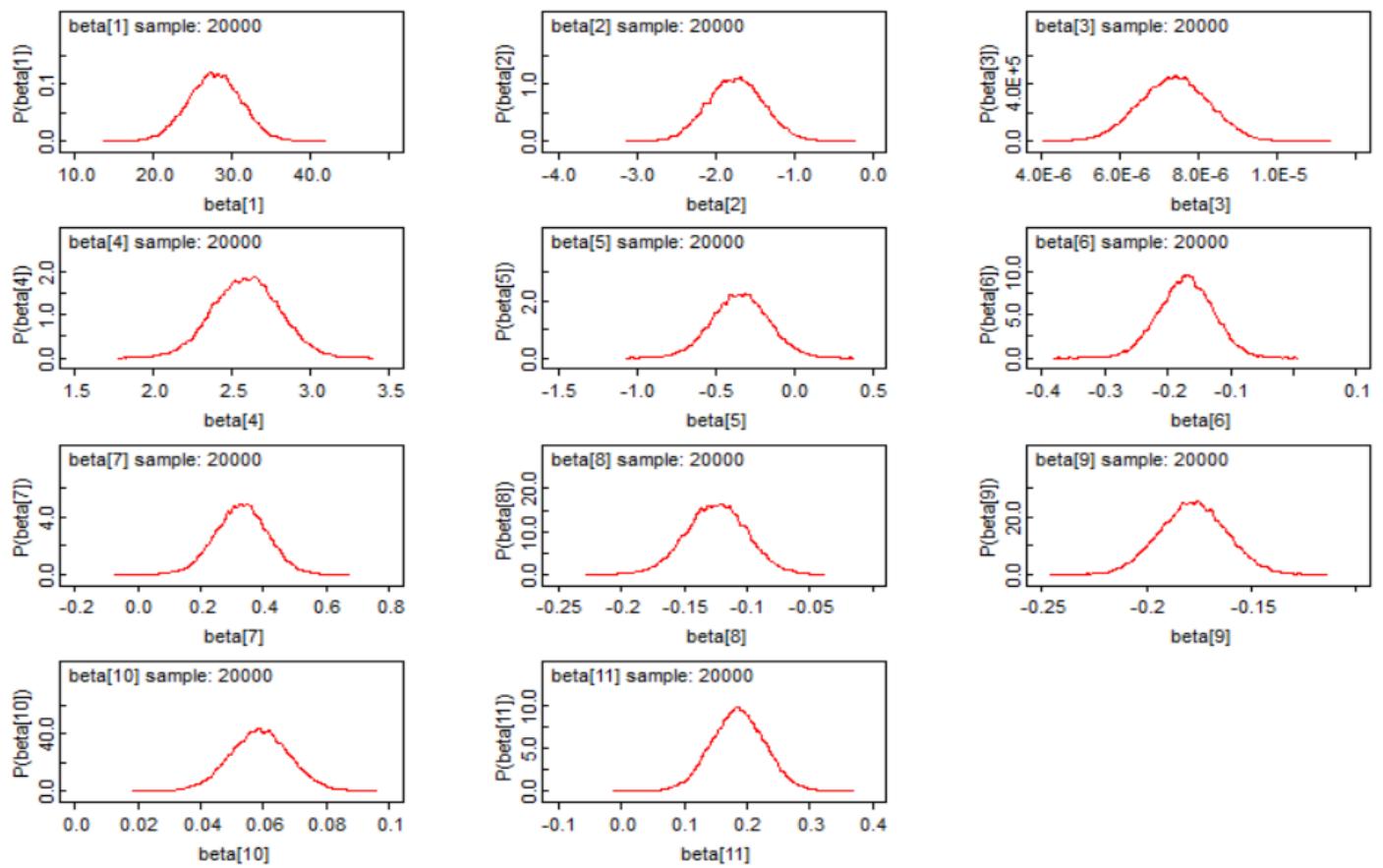
نمودار های تاریخچه β را در شکل زیر مشاهده می کنید:



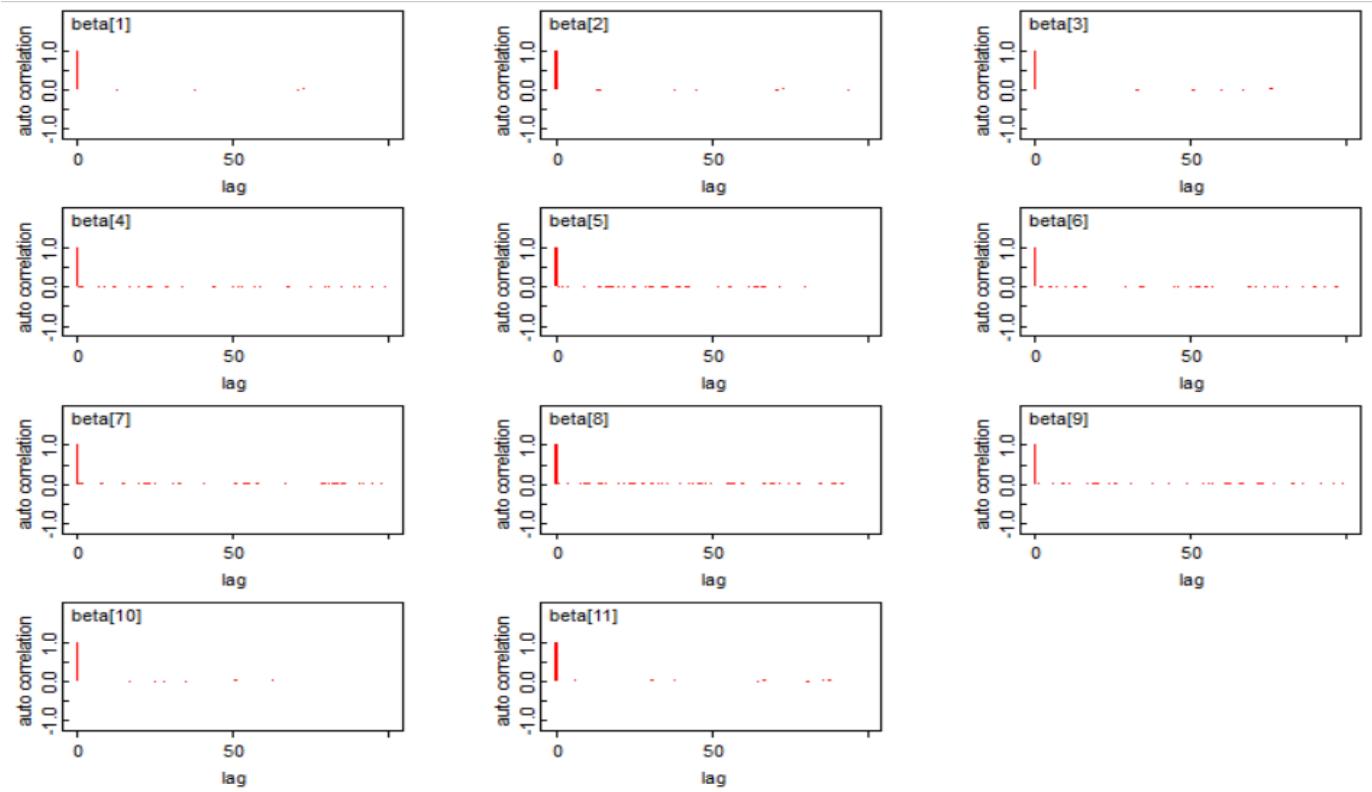
با توجه به نمودار های بالا می توان دید که β های تولید شده توسط این مدل همگرا شده اند و مدل با موفقیت پیاده سازی شده است.

در صفحه بعد چگالی های β را مشاهده می کنید:

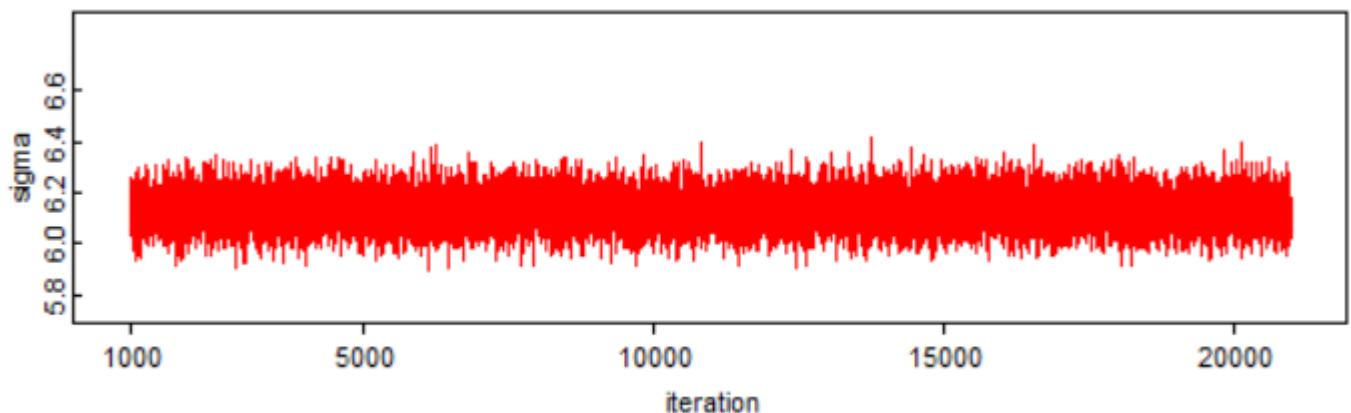
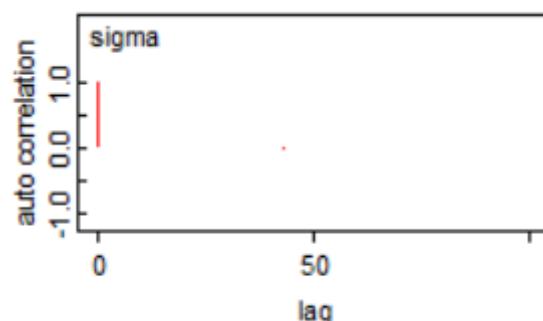
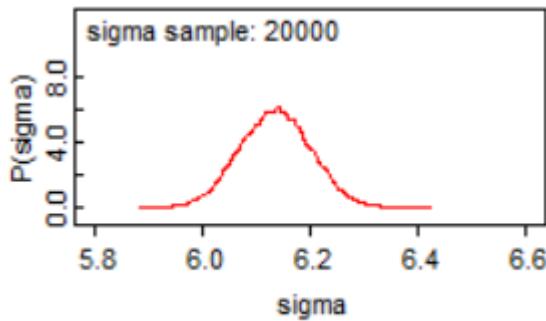
با توجه به نمودار های زیر می توان گفت که نمونه ها به صورت غیروابسته تولید شده اند.



با توجه به این نمودارها می‌توان گفت که توزیع‌های نرمال به خوبی شبیه‌سازی شده‌اند.



	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
sigma	6.135	0.06712	4.871E-4	6.004	6.135	6.267	1001	20000



با توجه به نمودار های بالا واریانس ۱۳۵/۶ تخمین زده شده و نمودار تاریخچه آن همگرا شده است.

توزیع آن که گاما وارون است به خوبی شبیه سازی شده است.

با توجه به جدول زیر DIC مدل ۲۷۴۰۰ می باشد.

	Dbar	Dhat	DIC	pD
HR	27390.0	27380.0	27400.0	11.97
total	27390.0	27380.0	27400.0	11.97

مدل دوم:

حال مدل دوم را با استفاده از همان متغیرها و با در نظر گرفتن توزیع SQ_N برای جمله خطأ اجرا می‌کنیم.

$$Y_i = X_i * \beta + \epsilon_i$$

$$X_i = [\text{MFIL}, \text{PO}, \text{RD}, \text{PS}, \text{UE}, \text{DV}, \text{MA}, \text{FP}, \text{BLK}, \text{FH}]$$

$$\beta = [\beta_0, \dots, \beta_{11}]$$

$$\epsilon_i \sim SN(0, \sigma^2, \delta)$$

توزیع پیشین β را نرمال چند متغیره، σ^2 را معکوس گاما و δ را نرمال در نظر می‌گیریم.

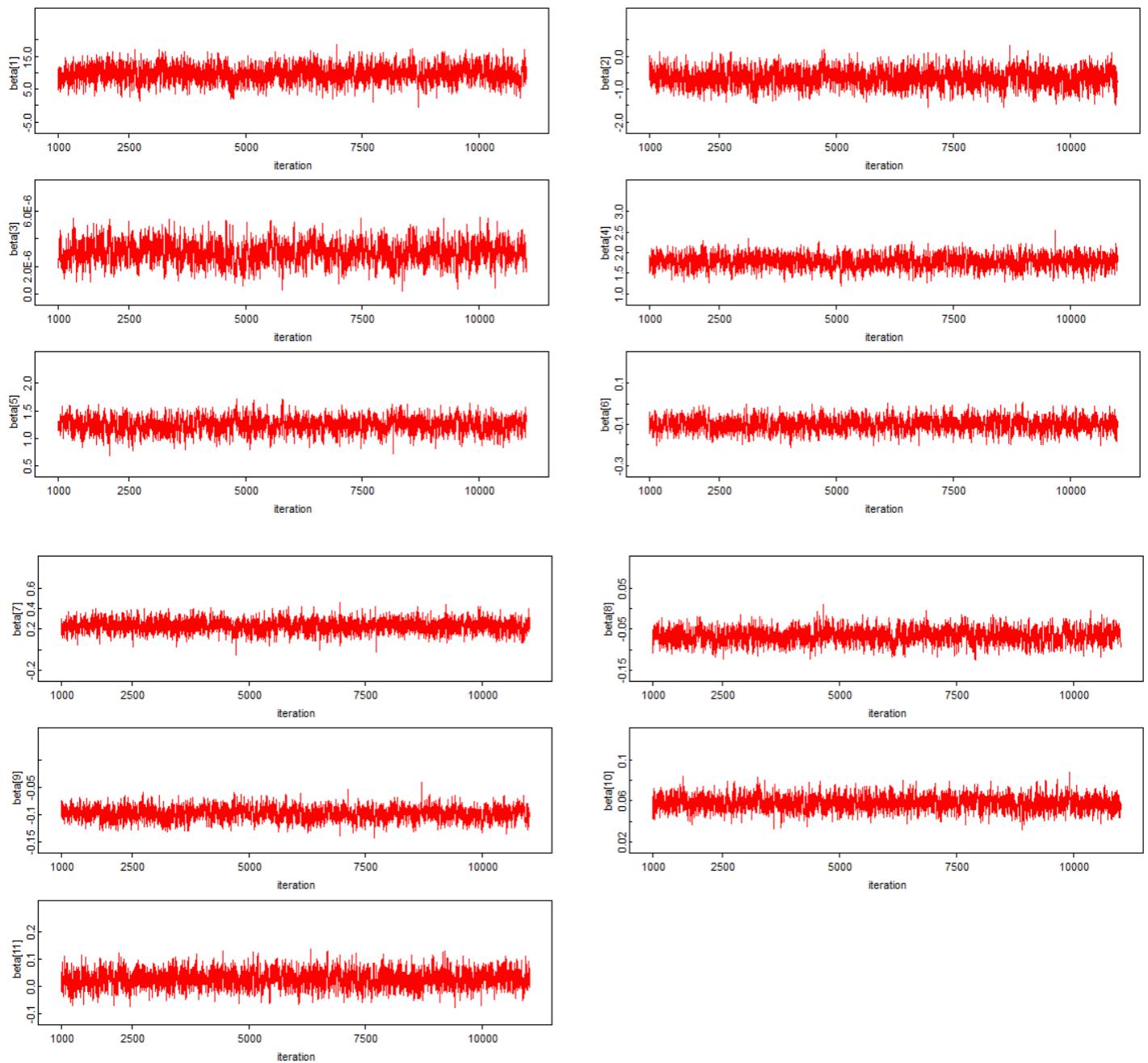
هر سه توزیع را آگاهی نابخش (واریانس زیاد و میانگین صفر) در نظر گرفتیم.

برآورد بیز β ها و واریانس آنها را می‌توان در جدول زیر مشاهده کرد:

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta[1]	9.653	2.411	0.08282	4.847	9.67	14.37	1001	10000
beta[2]	-0.6542	0.2549	0.008516	-1.15	-0.6533	-0.1513	1001	10000
beta[3]	3.014E-6	7.615E-7	2.837E-8	1.546E-6	3.004E-6	4.52E-6	1001	10000
beta[4]	1.784	0.1588	0.004873	1.461	1.788	2.084	1001	10000
beta[5]	1.239	0.1333	0.004602	0.9701	1.243	1.492	1001	10000
beta[6]	-0.09958	0.03185	9.826E-4	-0.161	-0.0994	-0.03836	1001	10000
beta[7]	0.2301	0.0585	0.001729	0.1157	0.2311	0.3446	1001	10000
beta[8]	-0.066667	0.0173	6.1E-4	-0.1003	-0.0663	-0.03279	1001	10000
beta[9]	-0.09834	0.01149	4.283E-4	-0.1207	-0.09824	-0.07622	1001	10000
beta[10]	0.05794	0.006778	1.98E-4	0.04467	0.05794	0.07125	1001	10000
beta[11]	0.02678	0.02948	8.815E-4	-0.03081	0.02653	0.08513	1001	10000

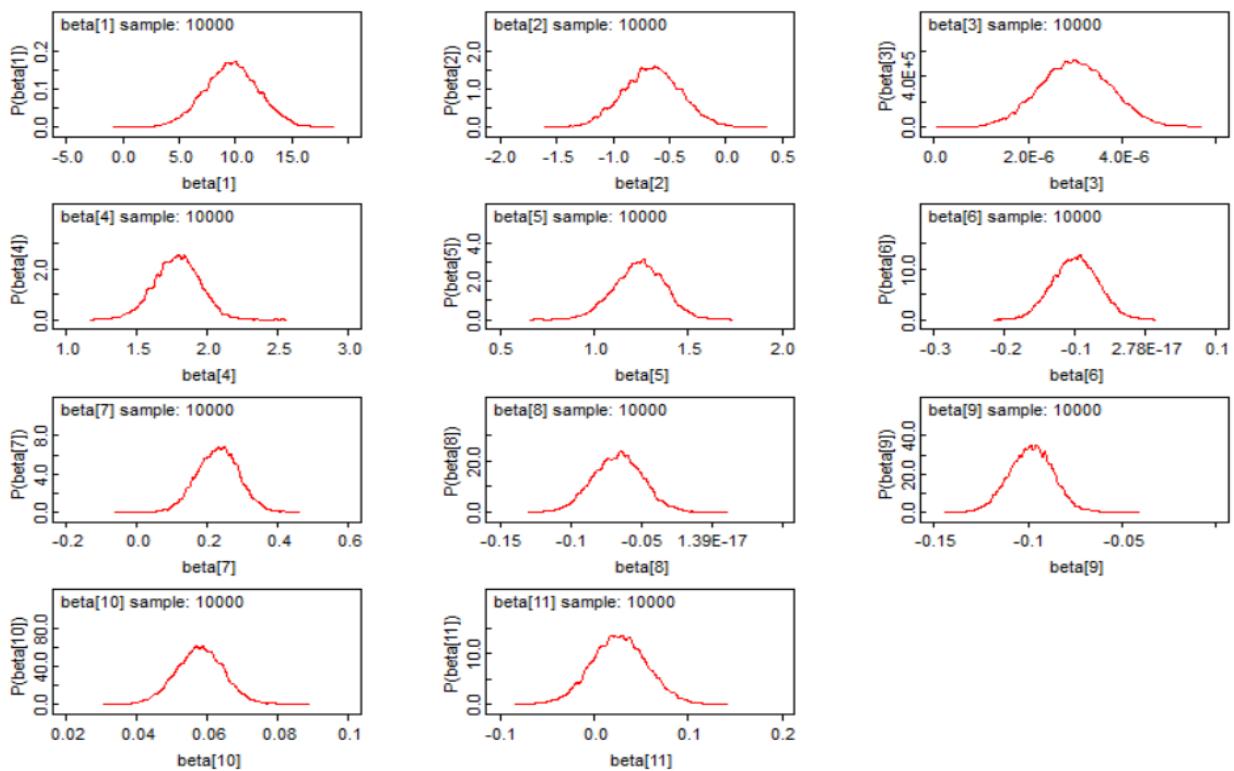
همانطور که مشاهده می‌کنید واریانس تخمین عرض از مبدا و ضریب گزارش میانه درآمد خانواده (MFIL) زیاد است پس کمتر می‌توان به آن‌ها اعتماد کرد.

نمودار های تاریخچه β را در شکل زیر مشاهده می‌کنید:

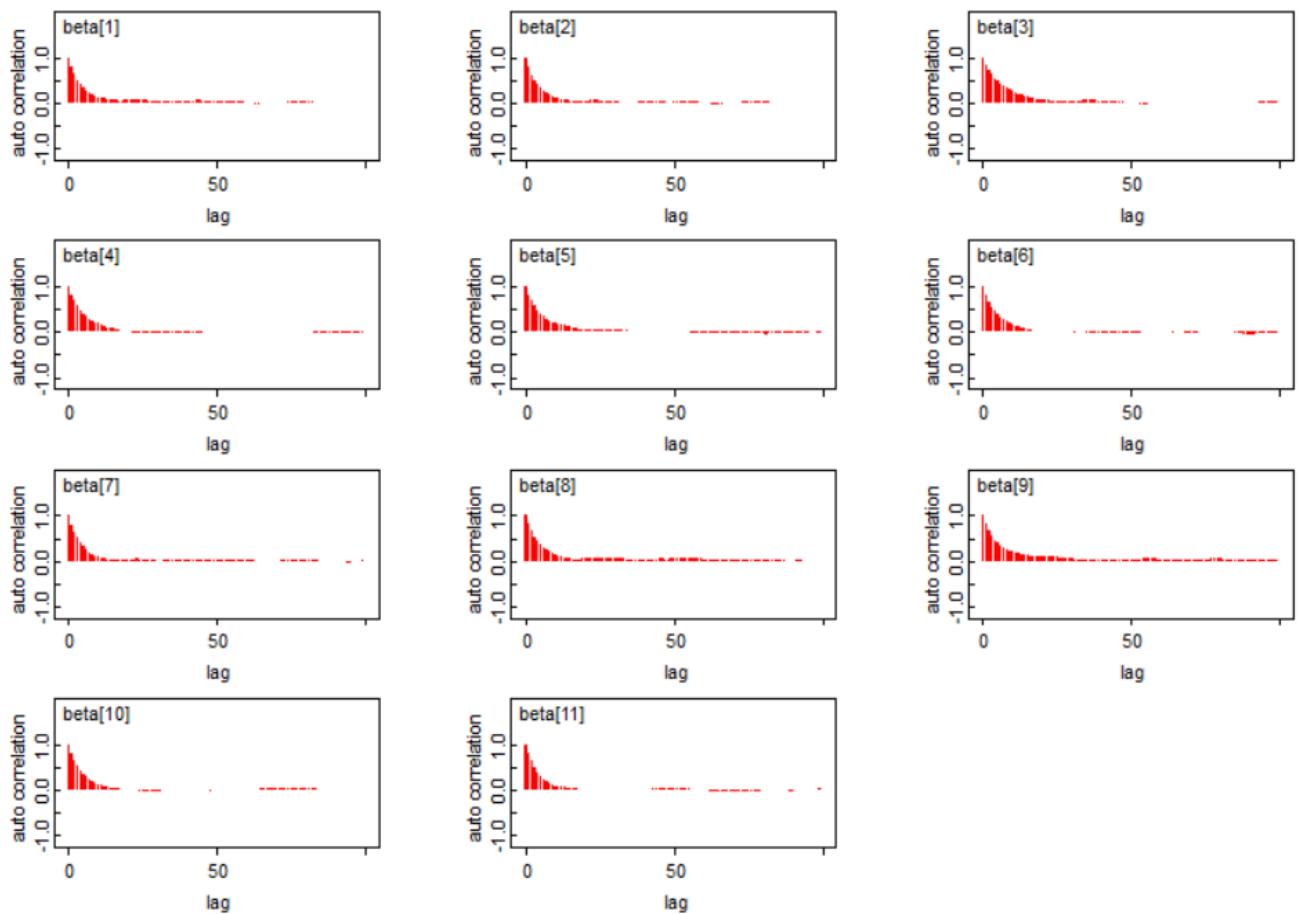


با توجه به نمودار های بالا میتوان دید که β های تولید شده توسط این مدل همگرا شده اند و مدل با موفقیت پیاده سازی شده است.

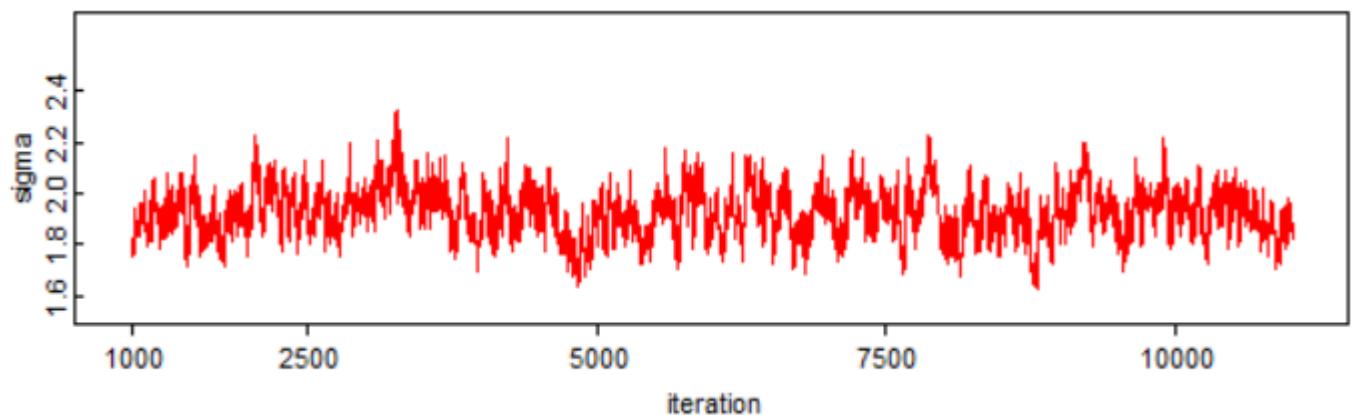
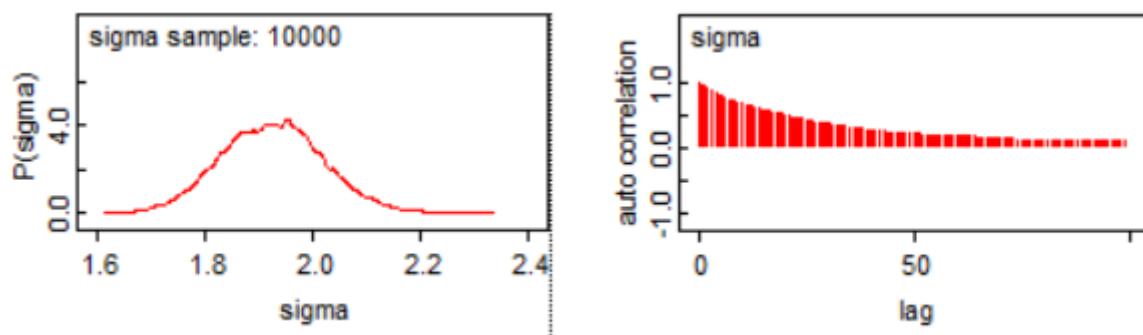
در ادامه نمودار های چگالی β و خودهمبستگی نمونه ها رسم شده است.



با توجه به این نمودارها می‌توان گفت که توزیع‌های نرمال به خوبی شبیه سازی شده‌اند.



	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
sigma	1.925	0.09543	0.006159	1.743	1.924	2.116	1001	10000



با توجه به نمودار های بالا میانگین واریانس $1/925$ تخمین زده شده و توزیع آن که نرمال است ولی همگرا نشده و به خوبی شبیه سازی نشده است. (با افزایش تولید نمونه نیز این مشکل برطرف نشد).

همگایی سایر متغیرهای موجود در مدل نیز بررسی شد و به خوبی همگرا بودند.
با توجه به جدول زیر DIC کل مدل ۲۶۹۴۰ می باشد.

	Dbar	Dhat	DIC	pD
HR	17560.0	14060.0	21060.0	3500.0
z	5872.0	5872.0	5872.0	1.559E-9
total	23440.0	19940.0	26940.0	3500.0

مدل سوم:

حال مدل سوم را با استفاده از همان متغیرها و با در نظر گرفتن توزیع t برای جمله خطأ اجرا می کنیم.

$$HR_i = X_i * \beta + \epsilon_i$$

$$X_i = [\text{MFIL}, \text{PO}, \text{RD}, \text{PS}, \text{UE}, \text{DV}, \text{MA}, \text{FP}, \text{BLK}, \text{FH}]$$

$$\beta = [\beta_0, \dots, \beta_{11}]$$

$$\epsilon_i \sim t(0, \sigma^2, \nu)$$

توزیع پیشین β را نرمال چند متغیره ، σ^2 را معکوس گاما و ν را نمایی در نظر می گیریم.

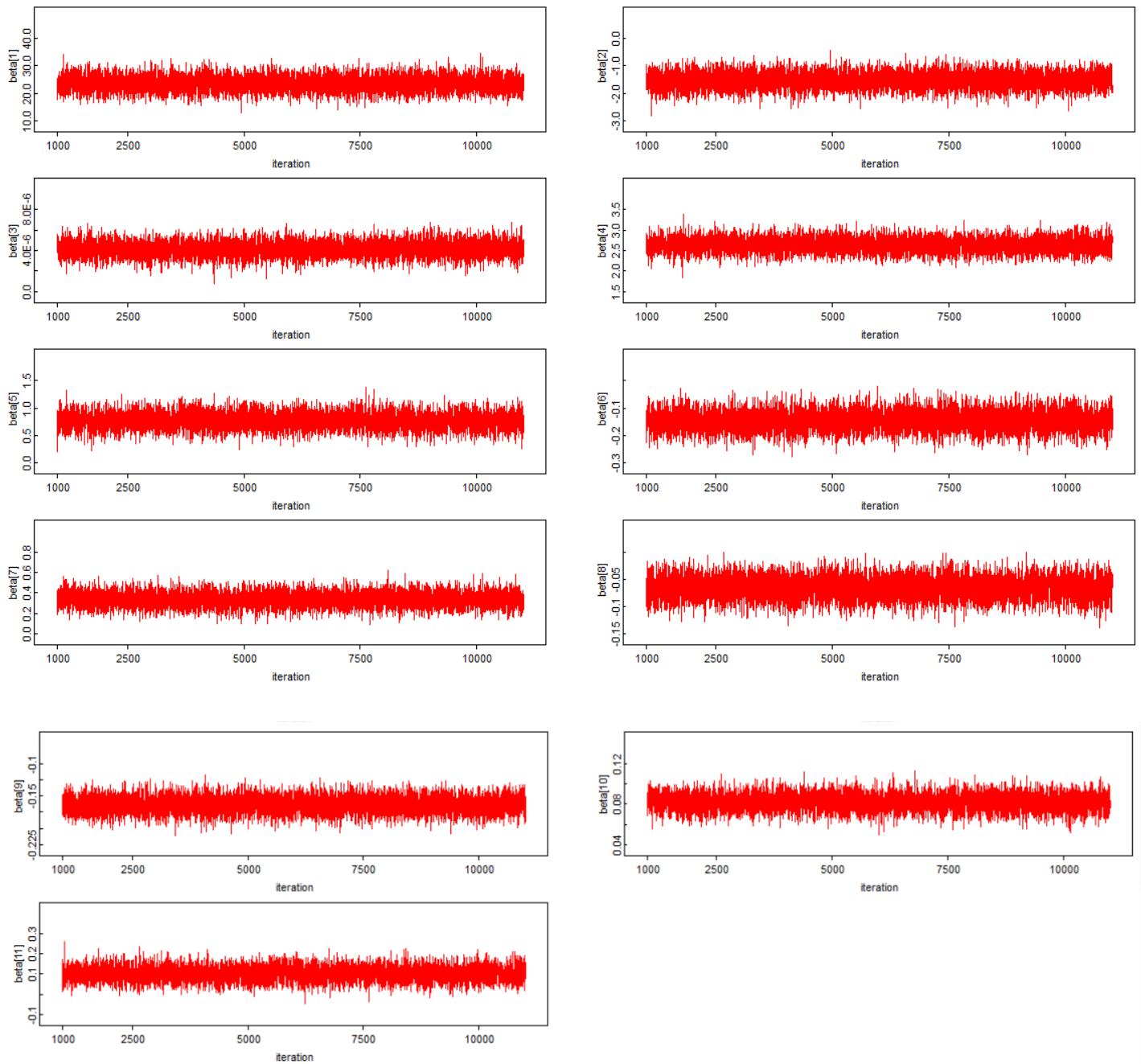
هر سه توزیع را آگاهی نابخش(واریانس زیاد و میانگین صفر) در نظر گرفتیم.

تخمین β ها و واریانس آنها را میتوان در جدول زیر مشاهده کرد .

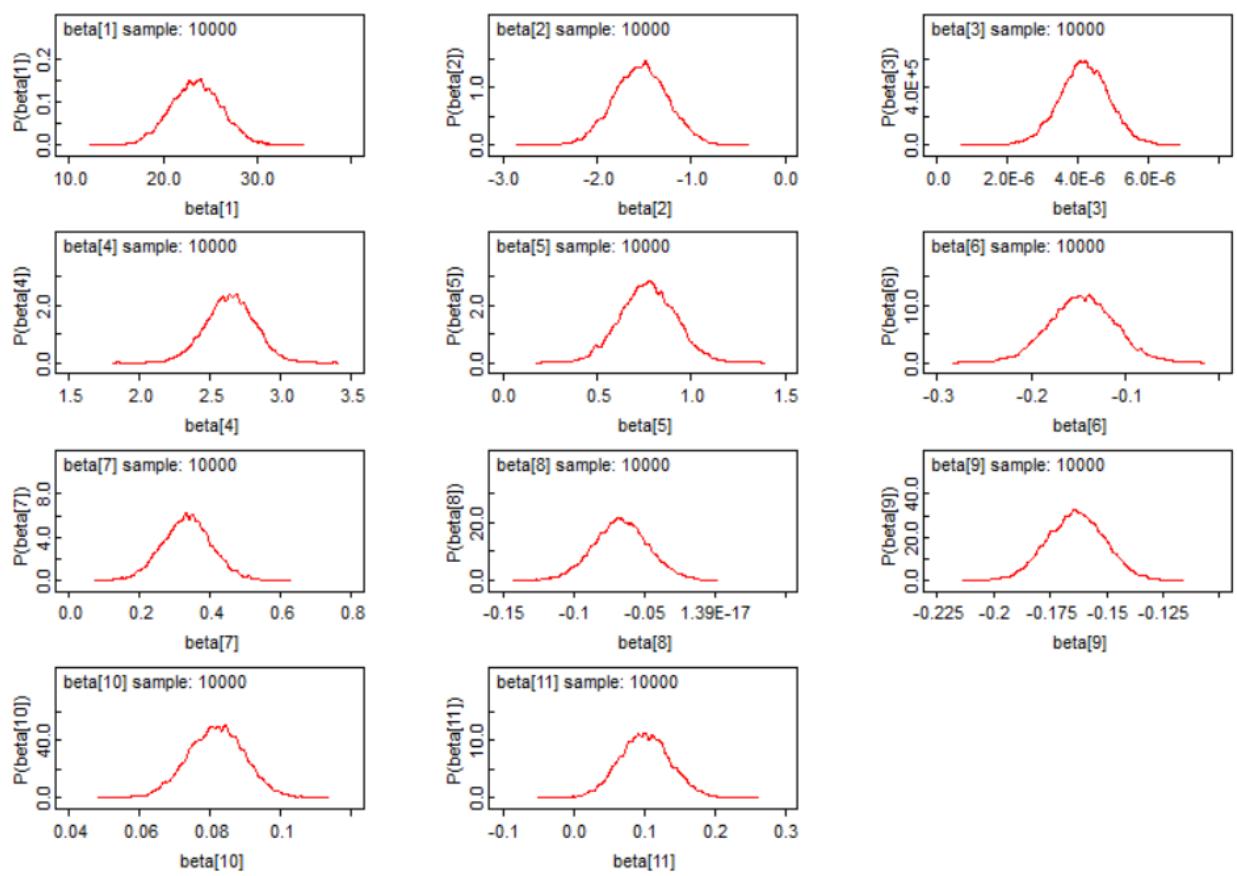
	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta[1]	23.52	2.697	0.033	18.28	23.53	28.8	1001	10000
beta[2]	-1.537	0.2877	0.003522	-2.099	-1.536	-0.9748	1001	10000
beta[3]	4.167E-6	7.105E-7	9.859E-9	2.741E-6	4.17E-6	5.546E-6	1001	10000
beta[4]	2.642	0.1702	0.002576	2.303	2.643	2.971	1001	10000
beta[5]	0.7692	0.1457	0.002287	0.4847	0.7692	1.056	1001	10000
beta[6]	-0.1459	0.03472	4.954E-4	-0.2134	-0.146	-0.07693	1001	10000
beta[7]	0.3348	0.06879	9.765E-4	0.1991	0.3352	0.4721	1001	10000
beta[8]	-0.0664	0.01919	2.372E-4	-0.1039	-0.06651	-0.0284	1001	10000
beta[9]	-0.1635	0.01256	1.698E-4	-0.188	-0.1635	-0.1391	1001	10000
beta[10]	0.08178	0.007994	1.36E-4	0.06612	0.08191	0.09717	1001	10000
beta[11]	0.1011	0.03555	5.16E-4	0.03319	0.1007	0.1713	1001	10000

همانطور که مشاهده می کنید واریانس تخمین عرض از مبدأ زیاد می باشد. همچنین ضریب محرومیت از منابع (RD) نیز خیلی به صفر نزدیک می باشد.

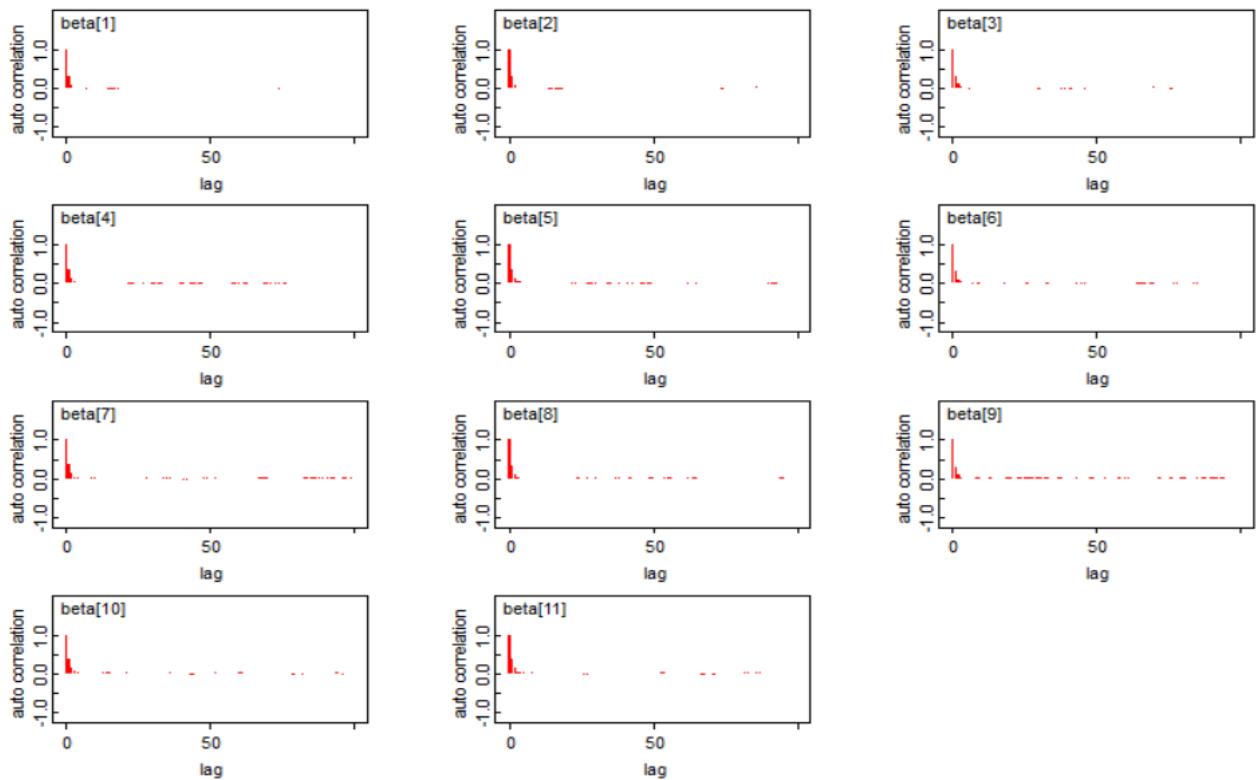
نمودار های تاریخچه β در صفحه بعد مشخص است:



با توجه به نمودار های بالا میتوان دید که β های تولید شده توسط این مدل همگرا شده اند و مدل با موفقیت پیاده سازی شده است .



با توجه به این نمودار ها میتوان گفت که توزیع های نرمال به خوبی شبیه سازی شده اند.



همگرایی سایر متغیرهای موجود در مدل نیز بررسی شد و به خوبی همگرا بودند.

با توجه به جدول زیر DIC کل مدل ۲۶۱۱۰ می باشد .

	Dbar	Dhat	DIC	pD
HR	25090.0	24070.0	26110.0	1019.0
total	25090.0	24070.0	26110.0	1019.0

مدل چهارم:

مدل چهارم را با استفاده از همان متغیر ها و با در نظر گرفتن توزیع SQ_t برای جمله خطاب اجرا می کنیم .

$$HR_i = X_i * \beta + \epsilon_i$$

$$X_i = [\text{MFIL}, \text{PO}, \text{RD}, \text{PS}, \text{UE}, \text{DV}, \text{MA}, \text{FP}, \text{BLK}, \text{FH}]$$

$$\beta = [\beta_0, \dots, \beta_{11}]$$

$$\epsilon_i \sim St(0, \sigma^2, \nu, \delta)$$

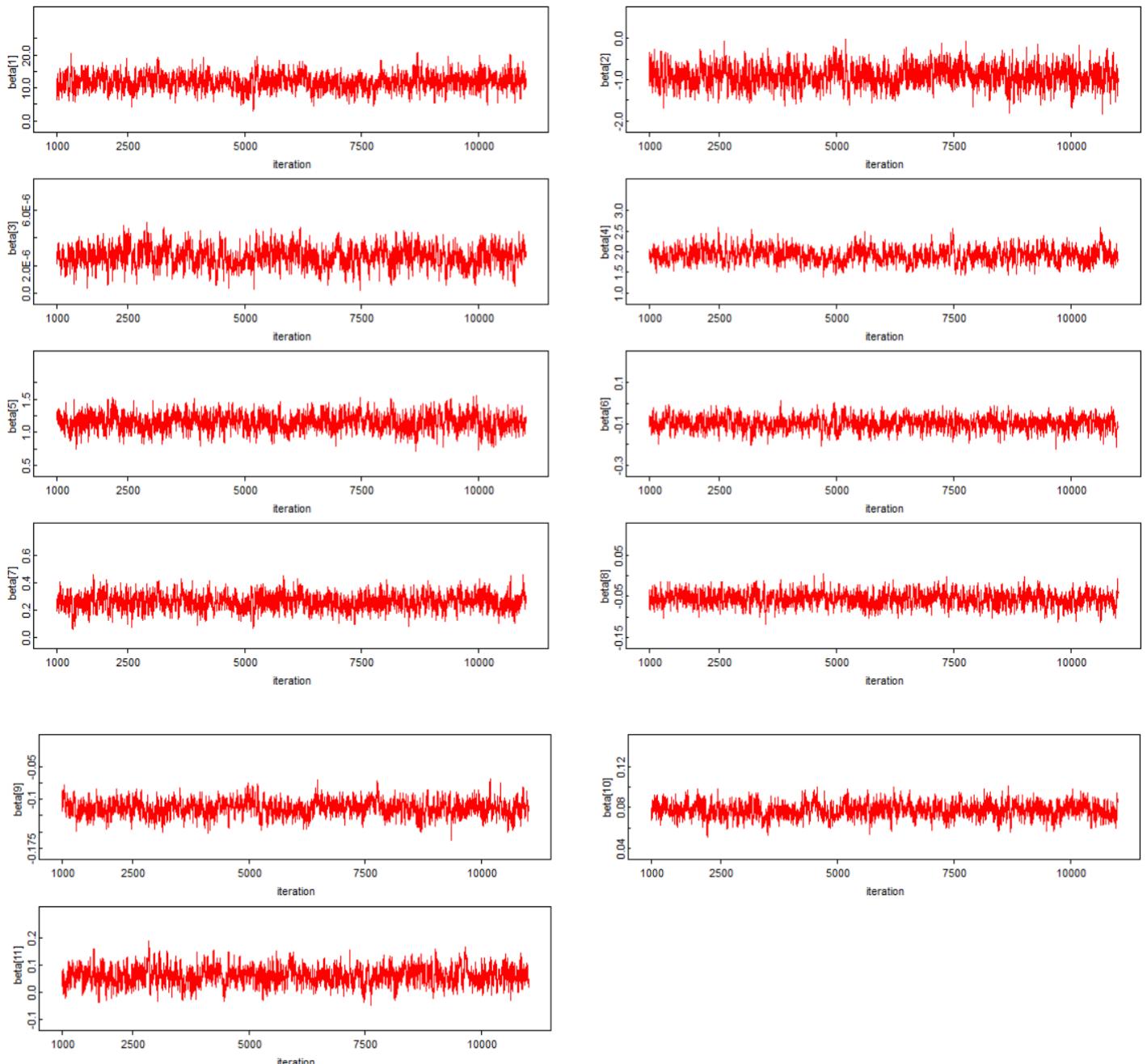
توزیع پیشین β را نرمال چند متغیره ، σ^2 را معکوس گاما ، ν را نمایی و δ را نرمال در نظر می گیریم.

هر چهار توزیع را آگاهی نابخش(واریانس زیاد و میانگین صفر) در نظر گرفتیم.
تخمین β ها و واریانس آنها را میتوان در جدول زیر مشاهده کرد .

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta[1]	11.89	2.269	0.1013	7.402	11.85	16.38	1001	10000
beta[2]	-0.9089	0.2359	0.01007	-1.379	-0.907	-0.433	1001	10000
beta[3]	2.667E-6	7.142E-7	3.343E-8	1.244E-6	2.677E-6	4.036E-6	1001	10000
beta[4]	1.922	0.1713	0.008832	1.605	1.918	2.277	1001	10000
beta[5]	1.153	0.1205	0.005189	0.9188	1.151	1.392	1001	10000
beta[6]	-0.09783	0.03106	0.001276	-0.1611	-0.09678	-0.03788	1001	10000
beta[7]	0.2589	0.05528	0.002398	0.1517	0.2587	0.3678	1001	10000
beta[8]	-0.05683	0.0158	5.579E-4	-0.08768	-0.05655	-0.02622	1001	10000
beta[9]	-0.1135	0.0114	5.463E-4	-0.1361	-0.1134	-0.09088	1001	10000
beta[10]	0.07732	0.006899	3.335E-4	0.06365	0.07734	0.09091	1001	10000
beta[11]	0.06037	0.03076	0.001377	8.603E-4	0.06026	0.1224	1001	10000

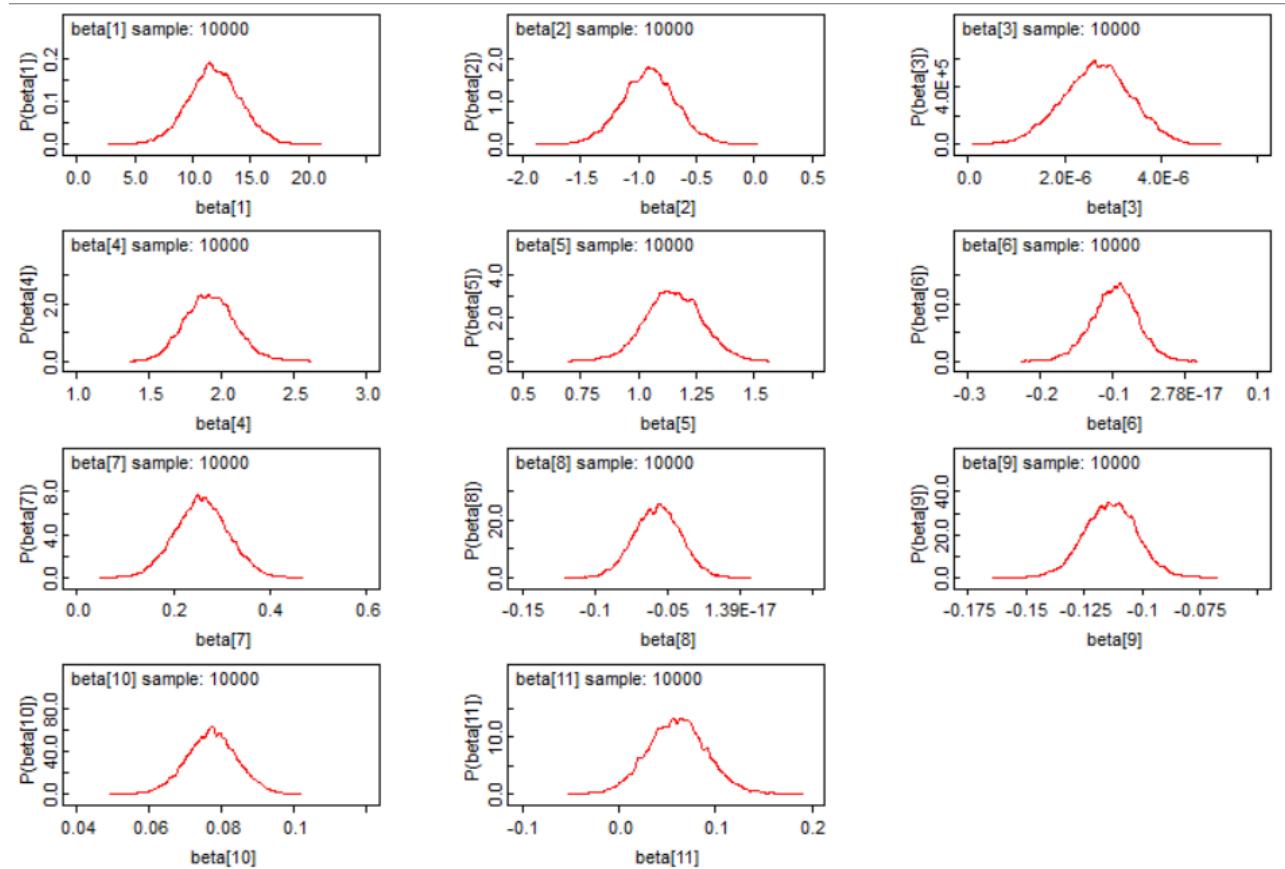
همانطور که مشاهده می کنید واریانس تخمین عرض از مبدأ زیاد است . همچنین ضریب محرومیت از منابع (RD) نیز خیلی به صفر نزدیک می باشد.

نمودار های تاریخچه :

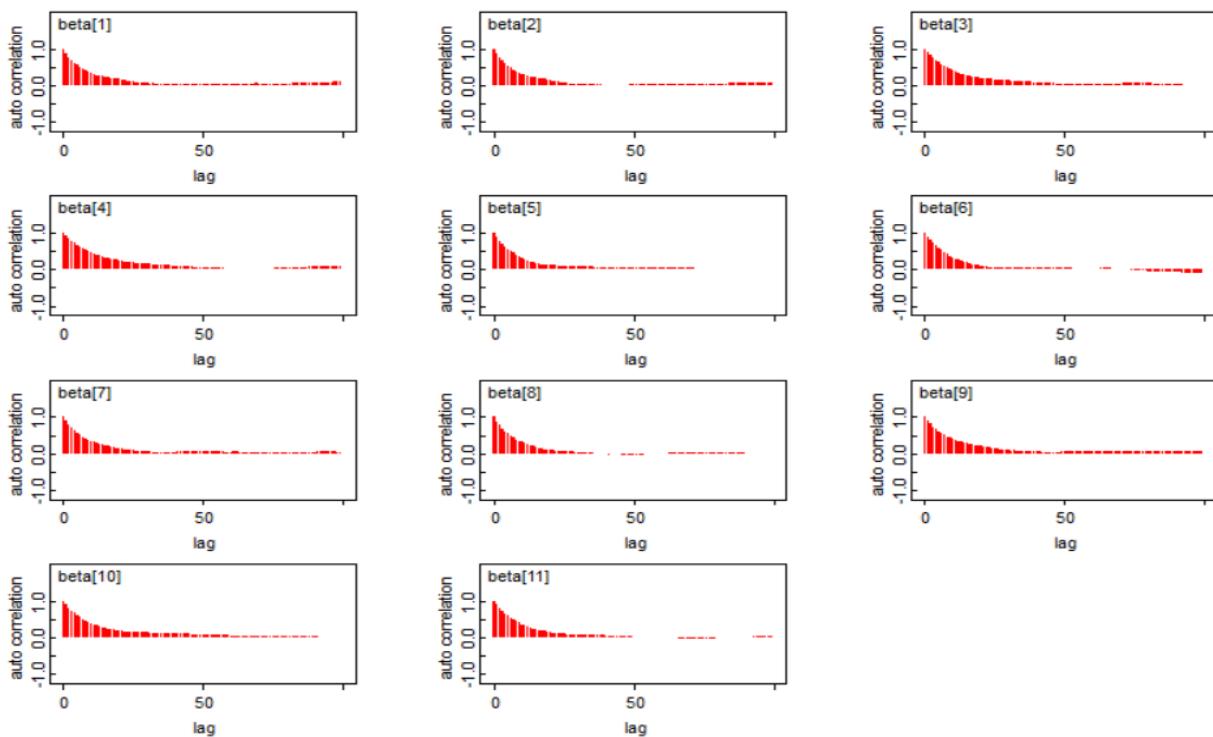


با توجه به نمودار های بالا میتوان دید که β های تولید شده توسط این مدل همگرا شده اند و مدل با موفقیت پیاده سازی شده است.

نمودار های چگالی در پایین مشخص است:



با توجه به این نمودار ها میتوان گفت که توزیع های نرمال به خوبی شبیه سازی شده اند.



با توجه به جدول زیر DIC کل مدل ۲۶۱۲۰ می‌باشد.

	Dbar	Dhat	DIC	pD
HR	17250.0	14250.0	20250.0	2998.0
z	5872.0	5872.0	5872.0	-6.659E-10
total	23120.0	20120.0	26120.0	2998.0

بخش دوم: مدل های رگرسیون با اثر تصادفی

مدل پنجم:

در این مدل با در نظر گرفتن هر ایالت به عنوان یک پانل و توزیع نرمال برای جمله خطای داریم:

$$HR_i = X_i * \beta + a[id[i]]$$

$$X_i = [\text{MFIL}, \text{PO}, \text{RD}, \text{PS}, \text{UE}, \text{DV}, \text{MA}, \text{FP}, \text{BLK}, \text{FH}]$$

$$\beta = [\beta_0, \dots, \beta_{11}]$$

$$\epsilon_i \sim N(0, \sigma^2)$$

توزیع پیشین β را نرمال چند متغیره، σ^2 را معکوس گاما در نظر می‌گیریم.

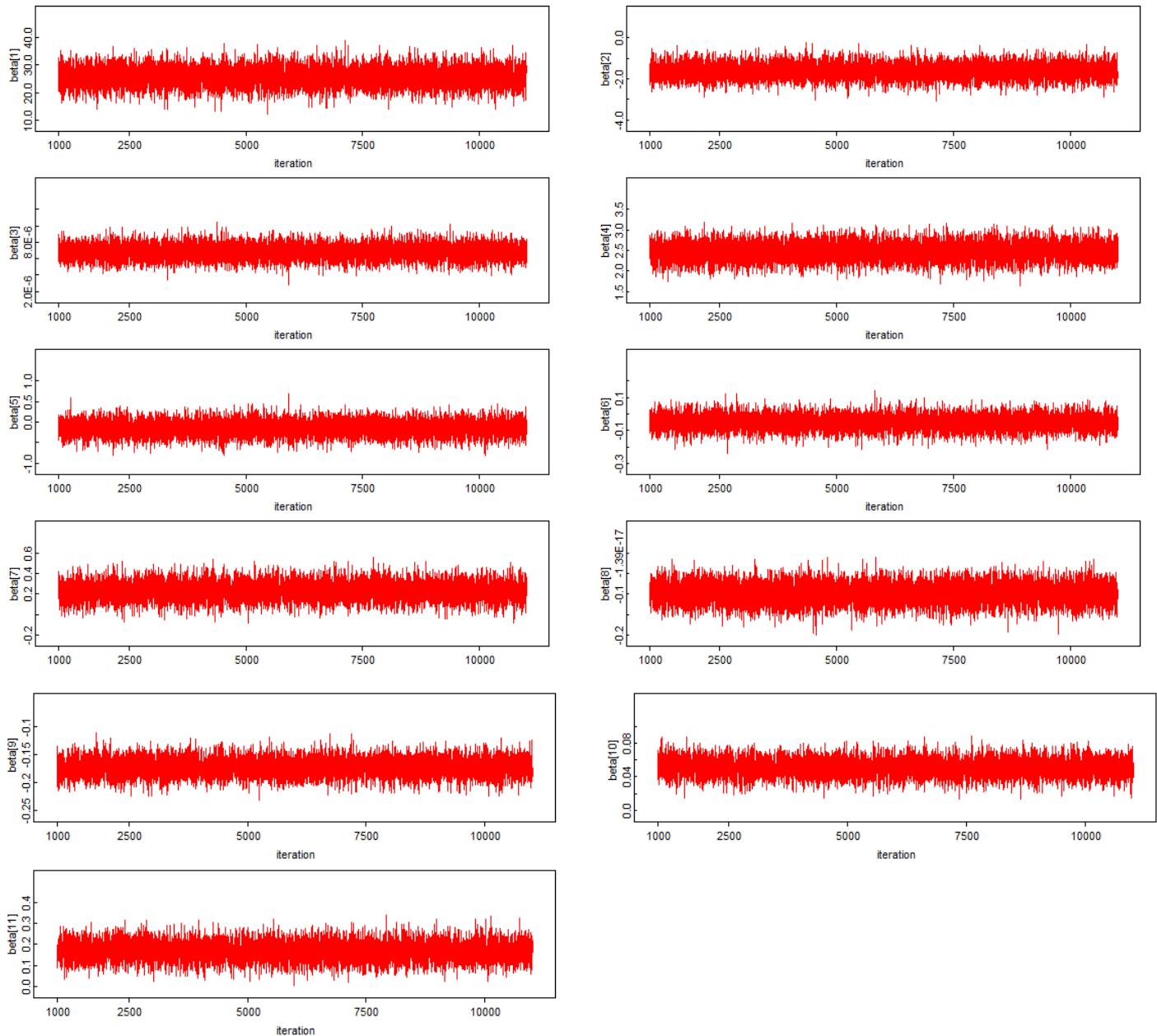
هردو توزیع را آگاهی نابخش (واریانس زیاد و میانگین صفر) در نظر گرفتیم.

تخمین β ها و واریانس آنها را میتوان در جدول زیر مشاهده کرد:

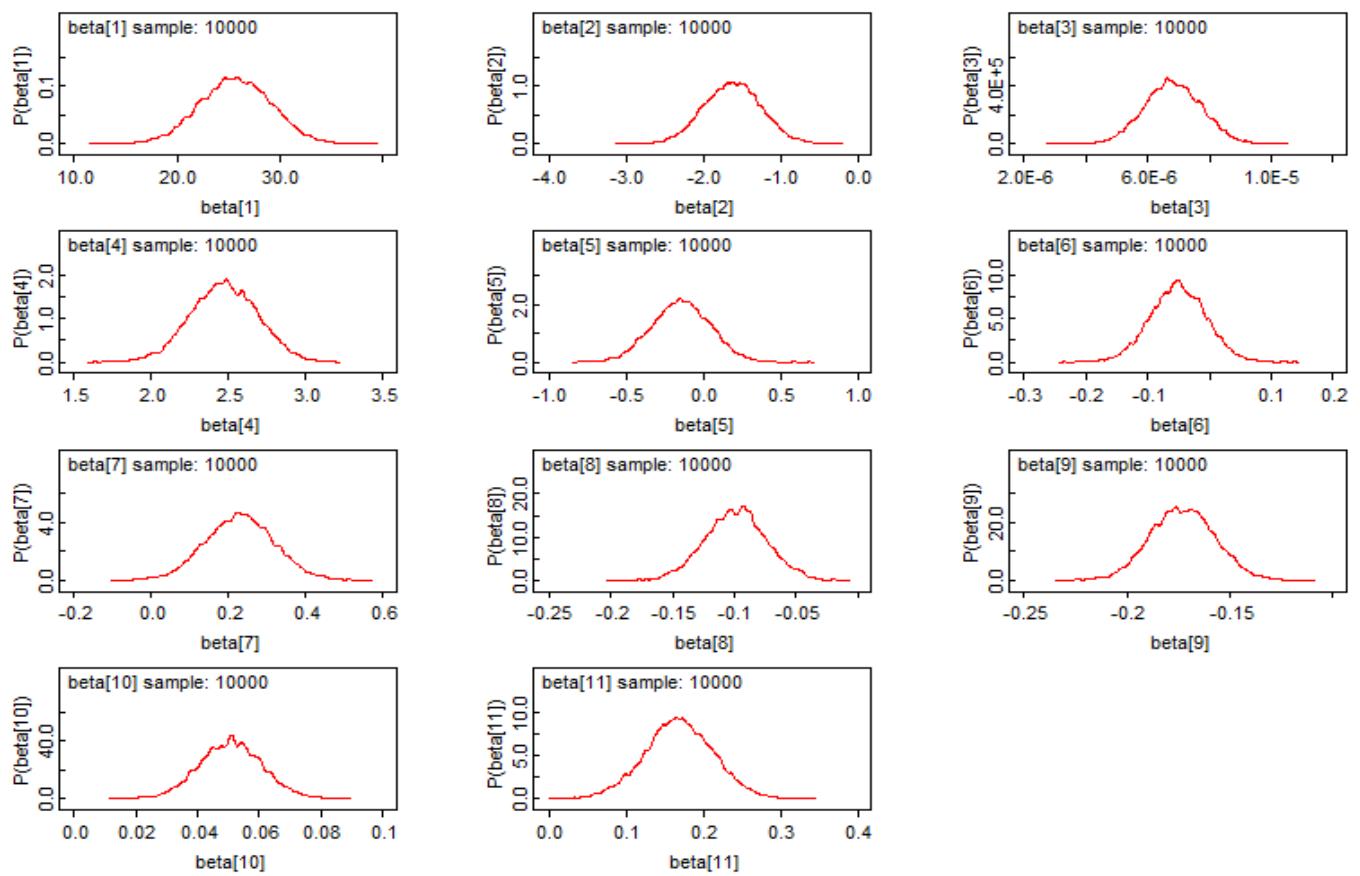
	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta[1]	25.82	3.429	0.0382	19.08	25.81	32.49	1001	10000
beta[2]	-1.631	0.366	0.003945	-2.351	-1.631	-0.9102	1001	10000
beta[3]	6.852E-6	9.242E-7	1.01E-8	5.065E-6	6.853E-6	8.701E-6	1001	10000
beta[4]	2.473	0.2159	0.002751	2.046	2.471	2.891	1001	10000
beta[5]	-0.1503	0.1853	0.00205	-0.5147	-0.1469	0.2124	1001	10000
beta[6]	-0.05015	0.04542	4.851E-4	-0.1405	-0.05008	0.03776	1001	10000
beta[7]	0.2303	0.08701	0.001031	0.05994	0.2314	0.4022	1001	10000
beta[8]	-0.09868	0.02416	2.595E-4	-0.1464	-0.0986	-0.05175	1001	10000
beta[9]	-0.1733	0.01557	1.61E-4	-0.2039	-0.1731	-0.1431	1001	10000
beta[10]	0.05061	0.01011	1.215E-4	0.03098	0.05068	0.07044	1001	10000
beta[11]	0.1694	0.04357	4.786E-4	0.08377	0.1695	0.2544	1001	10000

همانطور که مشاهده می‌کنید واریانس تخمین عرض از مبدأ و ضریب گزارش میانه در آمد خانواده (MFIL) زیاد است پس کمتر می‌توان به آن‌ها اعتماد کرد.

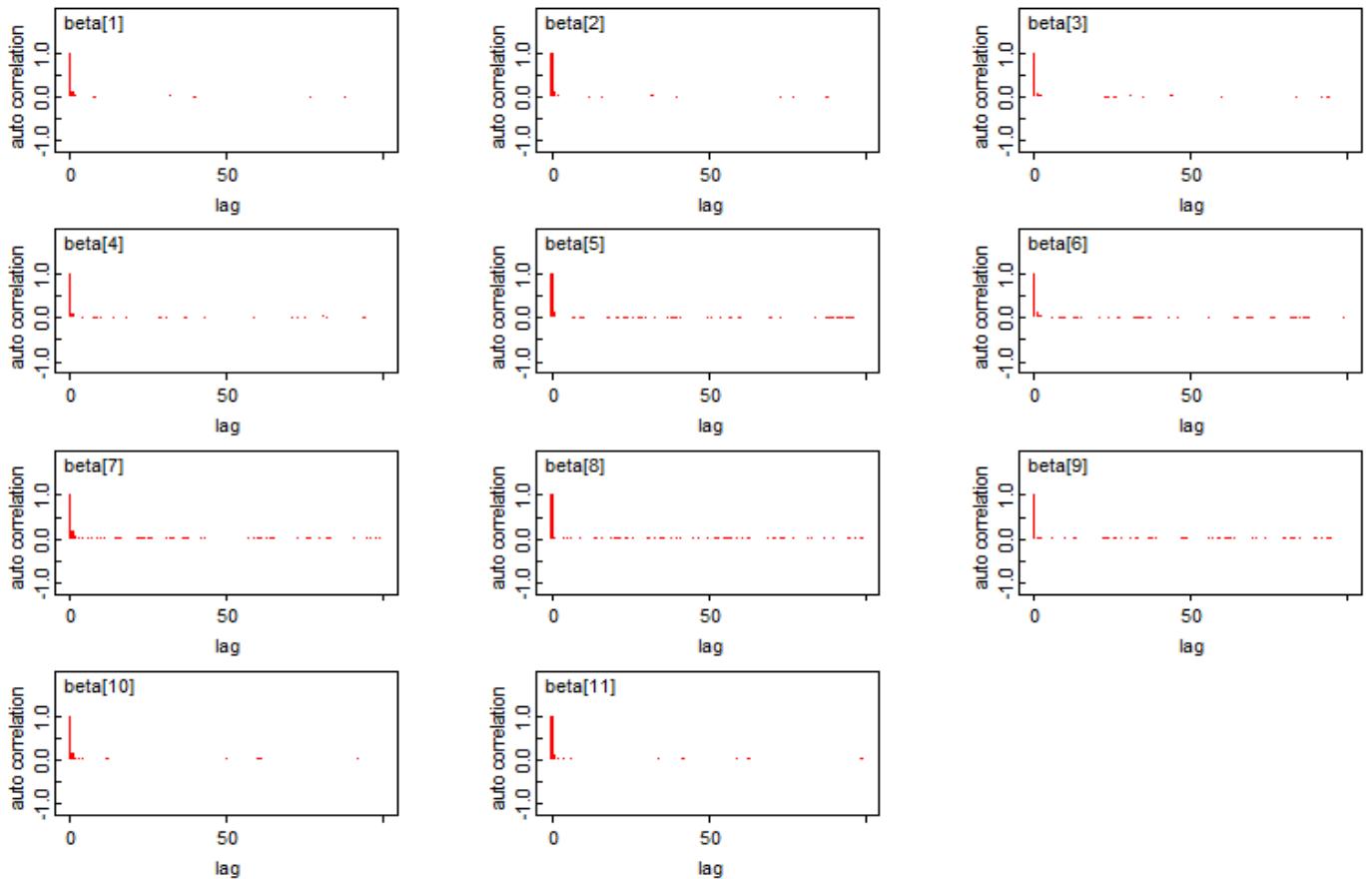
در شکل زیر نمودارهای تاریخچه را مشاهده می‌کنید:



با توجه به نمودارهای بالا می‌توان دید که β ‌های تولید شده توسط این مدل همگرا شده‌اند و مدل با موفقیت پیاده سازی شده است.

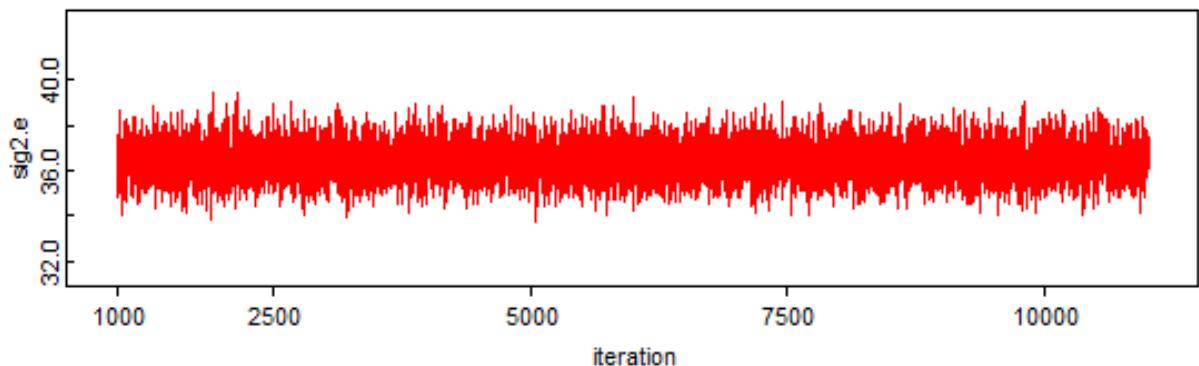


با توجه به این نمودارها میتوان گفت که توزیع های نرمال به خوبی شبیه سازی شده اند.



با توجه به نمودارهای بالا میتوان گفت که نمونه ها به صورت غیروابسته تولید شده اند.

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
sig2.e	36.46	0.799	0.008594	34.93	36.45	38.07	1001	10000



با توجه به نمودارهای بالا واریانس σ^2_e تخمین زده شده و نمودار تاریخچه آن همگرا شده است.

توزیع آن که گاما وارون است به خوبی شبیه سازی شده است.

در ادامه نمودارهای تاریخچه برای اثر تصادفی رسم شده اند.

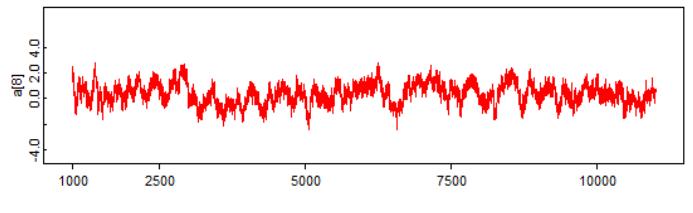
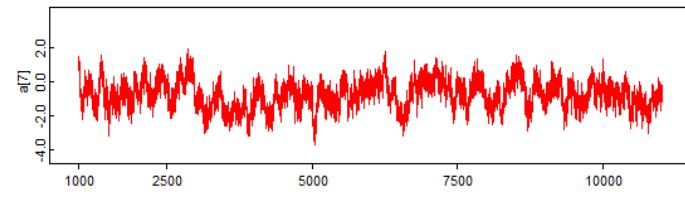
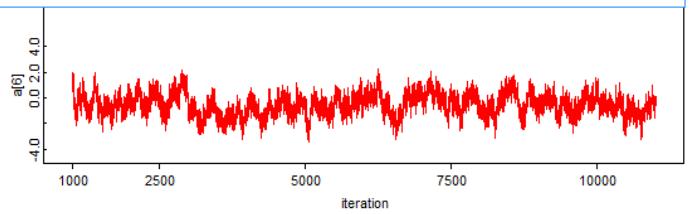
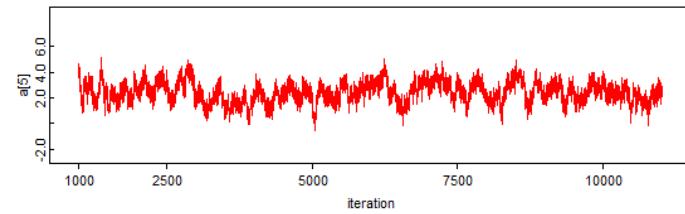
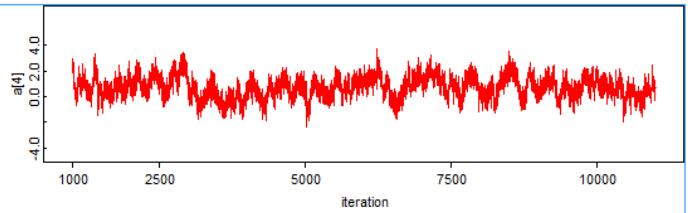
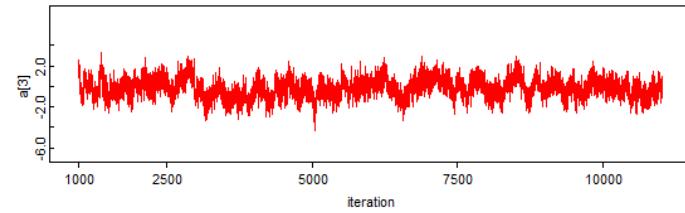
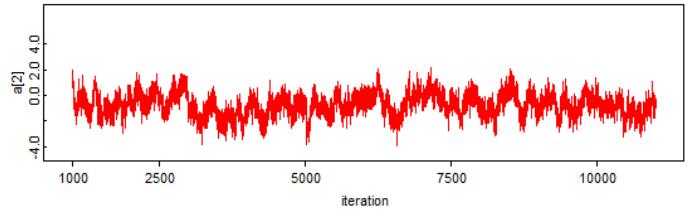
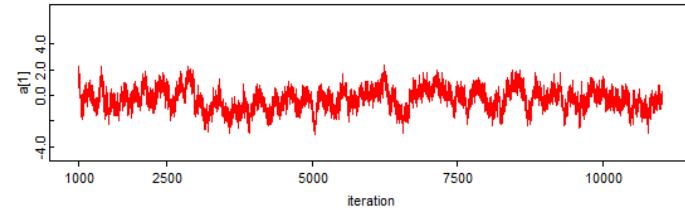
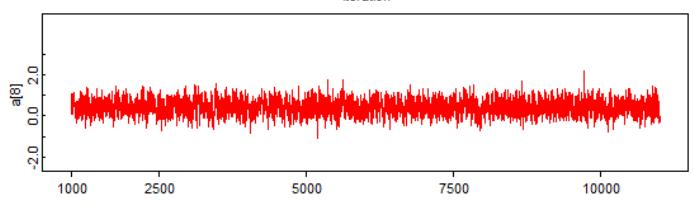
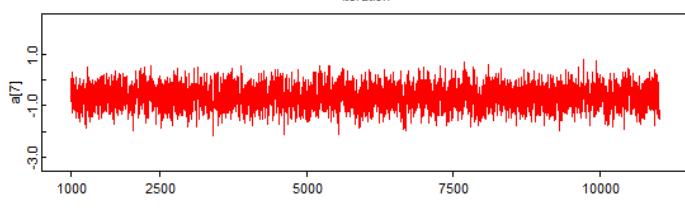
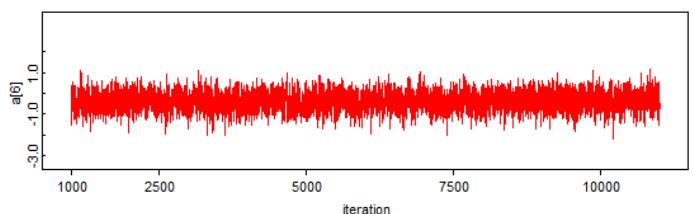
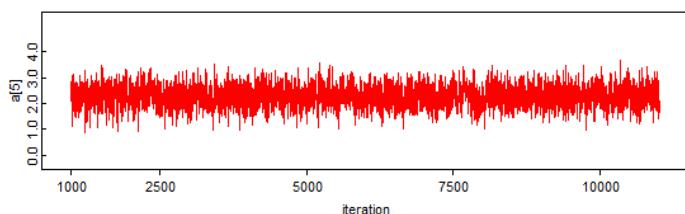
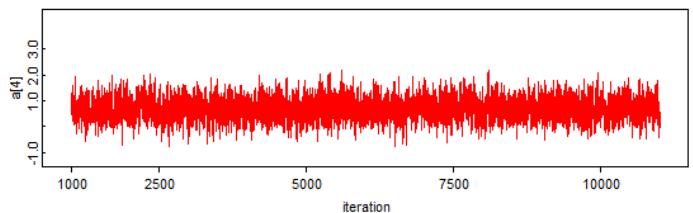
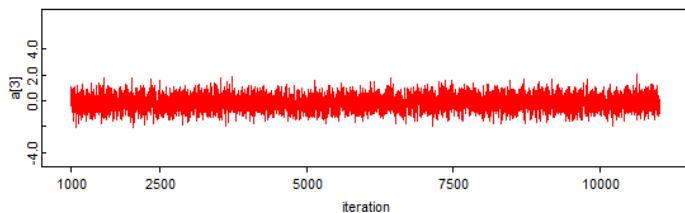
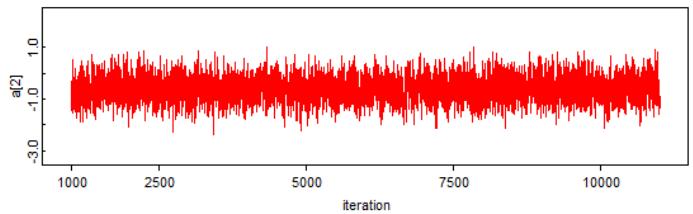
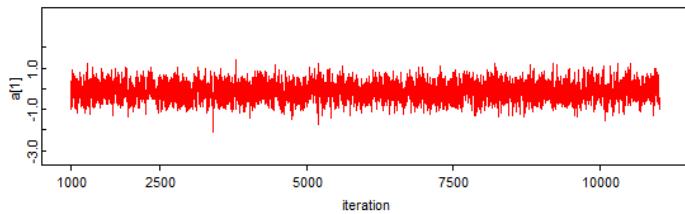
در اولین حالت توزیع پیشین اثر تصادفی هر پانل را مستقل از پانل دیگر و نرمال صفر و یک گرفتیم:

همانطور که مشاهده می کنید به خوبی همگرا شده.

با توجه به اینکه ۱۷ پانل داریم و اثرات تصادفی شبیه به هم است پس فقط نمودار ۶ پانل اول را رسم کردیم.

در حالت دوم توزیع پیشین اثر تصادفی هر پانل را مستقل از پانل دیگر و نرمال ۰ و ۰/۱ گرفتیم.(آگاهی نابخش)

مشاهده می کنید که همگرایی مدل بدتر شده است.



با توجه به جدول زیر DIC کل مدل ۲۷۲۸۰ می‌باشد.

	Dbar	Dhat	DIC	pD
HR	27250.0	27230.0	27280.0	23.92
total	27250.0	27230.0	27280.0	23.92

مدل ششم:

در این مدل با در نظر گرفتن هر ایالت به عنوان یک پانل و اثر تصادفی هر پانل و توزیع SQ_N برای جمله خطای داریم:

$$HR_i = X_i * \beta + a[id[i]]$$

$$X_i = [MFIL, PO, RD, PS, UE, DV, MA, FP, BLK, FH]$$

$$\beta = [\beta_0, \dots, \beta_{11}]$$

$$\epsilon_i \sim SN(0, \sigma^2, \delta)$$

توزیع پیشین β را نرمال چند متغیره، σ^2 را معکوس گاما و δ را نرمال در نظر می‌گیریم.

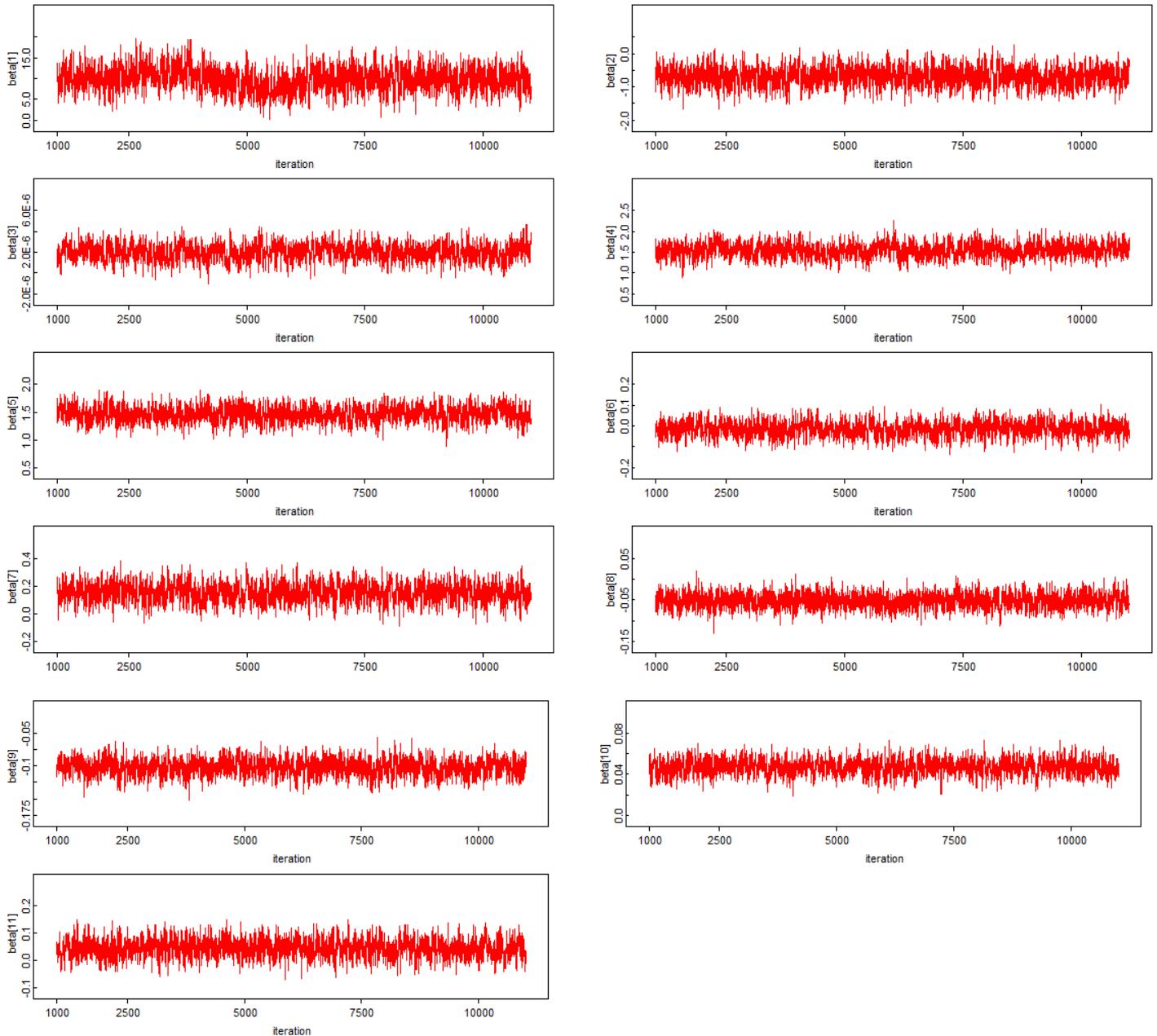
هر سه توزیع را آگاهی نابخش(واریانس زیاد و میانگین صفر) در نظر گرفتیم.

تخمین β ها و واریانس آنها را می‌توان در جدول زیر مشاهده کرد.

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta[1]	9.888	2.711	0.1182	4.573	9.869	15.22	1001	10000
beta[2]	-0.691	0.272	0.008379	-1.228	-0.6901	-0.1522	1001	10000
beta[3]	1.996E-6	7.85E-7	3.02E-8	4.578E-7	2.007E-6	3.489E-6	1001	10000
beta[4]	1.55	0.1639	0.006285	1.23	1.55	1.88	1001	10000
beta[5]	1.47	0.1353	0.004922	1.203	1.473	1.729	1001	10000
beta[6]	-0.01407	0.03442	0.001164	-0.08272	-0.01339	0.05125	1001	10000
beta[7]	0.1515	0.06649	0.002005	0.02046	0.1526	0.2787	1001	10000
beta[8]	-0.05215	0.0168	5.239E-4	-0.08458	-0.05218	-0.01932	1001	10000
beta[9]	-0.1023	0.01193	4.303E-4	-0.1256	-0.1022	-0.07888	1001	10000
beta[10]	0.04689	0.007183	2.43E-4	0.03264	0.04696	0.06086	1001	10000
beta[11]	0.04404	0.0317	9.838E-4	-0.01822	0.04375	0.107	1001	10000

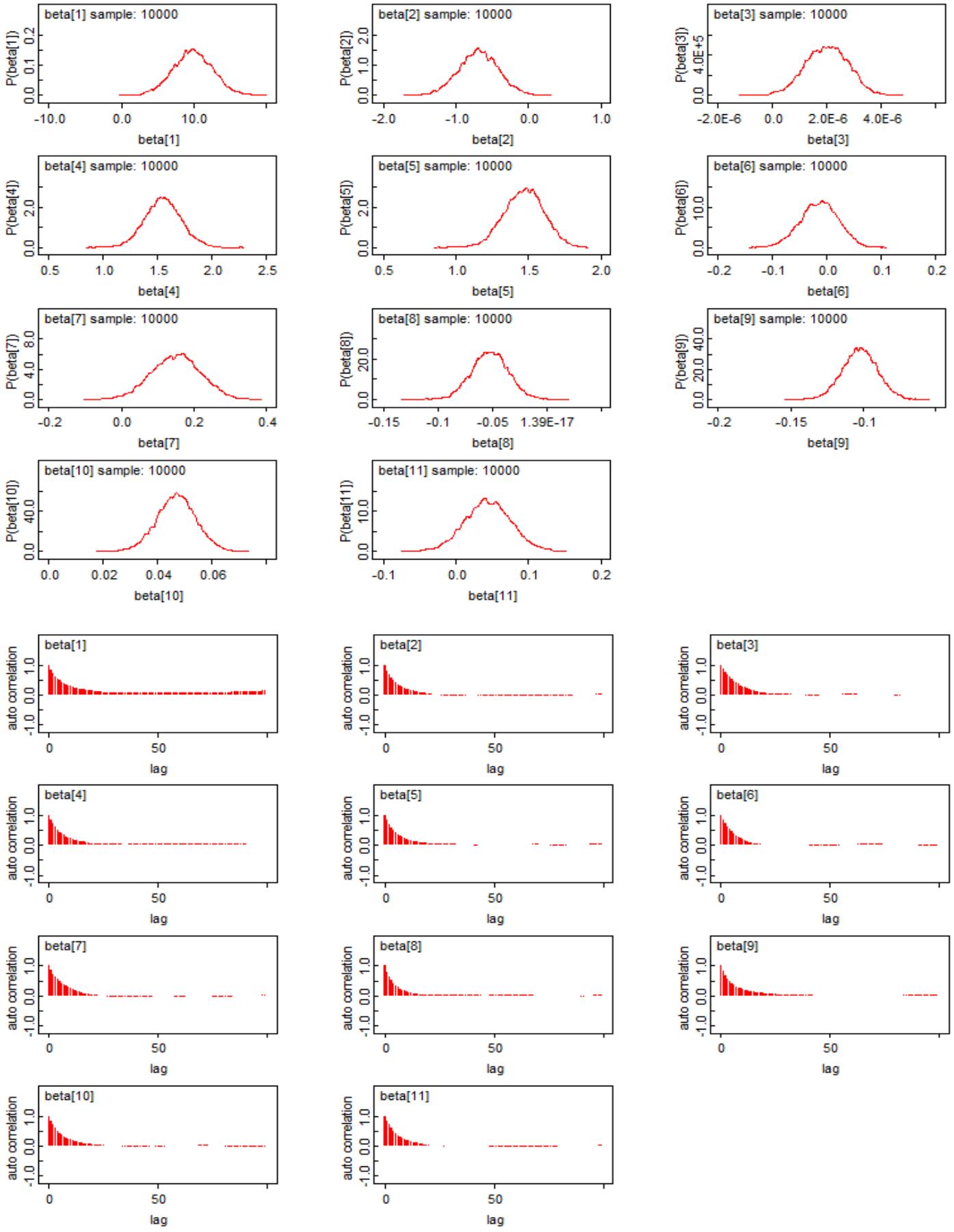
همانطور که مشاهده می‌کنید واریانس تخمین عرض از مبدأ زیاد می‌باشد بنابراین به آن کمتر می‌توان اعتماد کرد.

در ادامه نمودارهای تاریخچه بتا را رسم می‌کنیم:

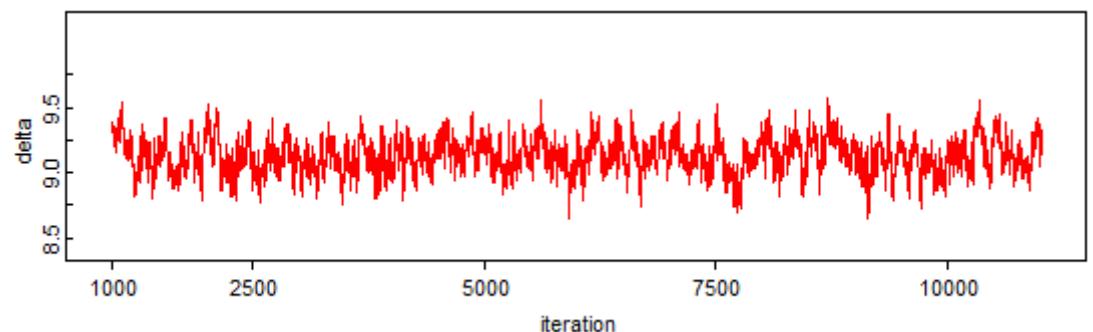


می‌توان گفت همگرایی نمودارهای بالا قابل قبول است.

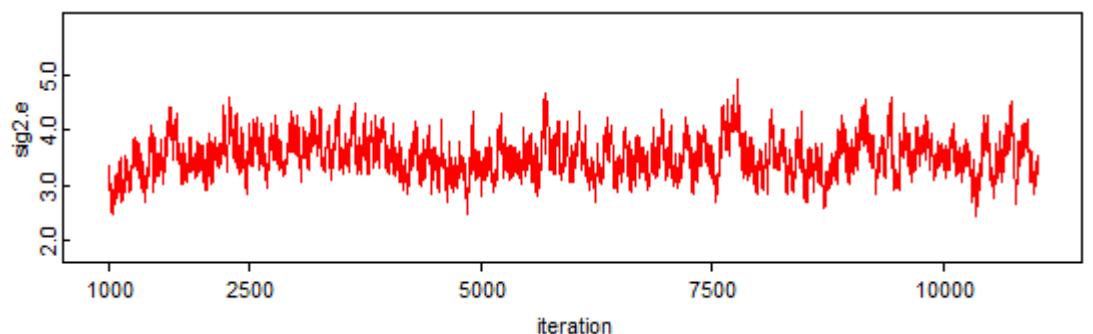
در صفحه بعد نمودارهای توزیع و همبستگی نمونه‌ها رسم شده است:



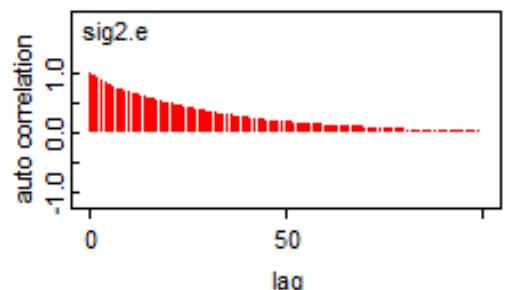
نمودار تاریخچه δ :



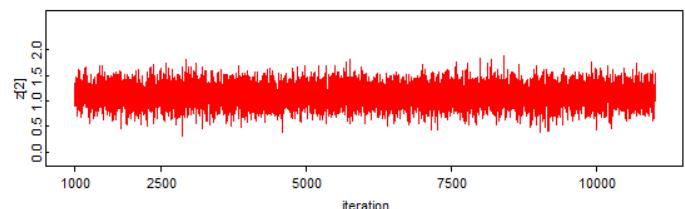
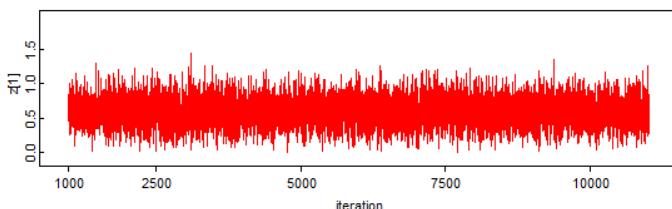
نمودار و برآورد بیز واریانس:



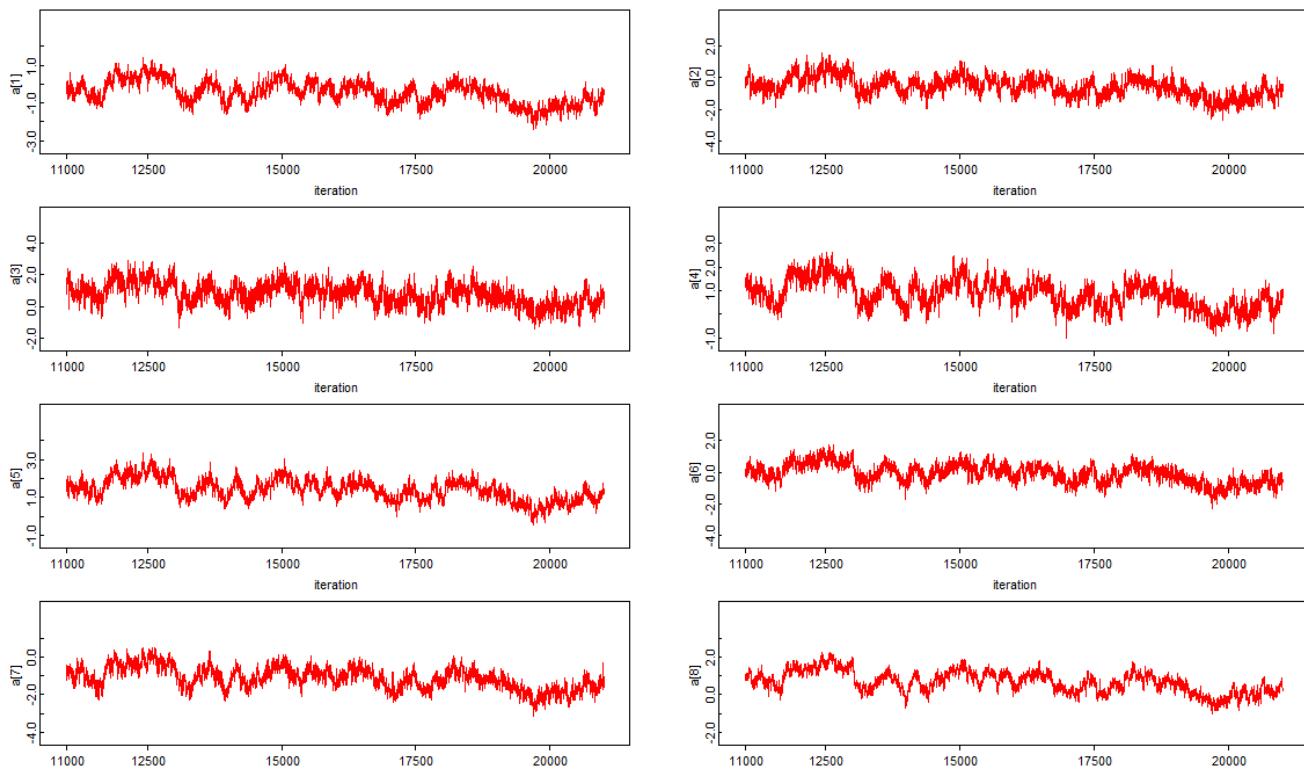
	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
sig2.e	3.517	0.3451	0.02208	2.898	3.498	4.233	1001	10000



نمودار متغیر پنهان Z برای دو رکورد اول:

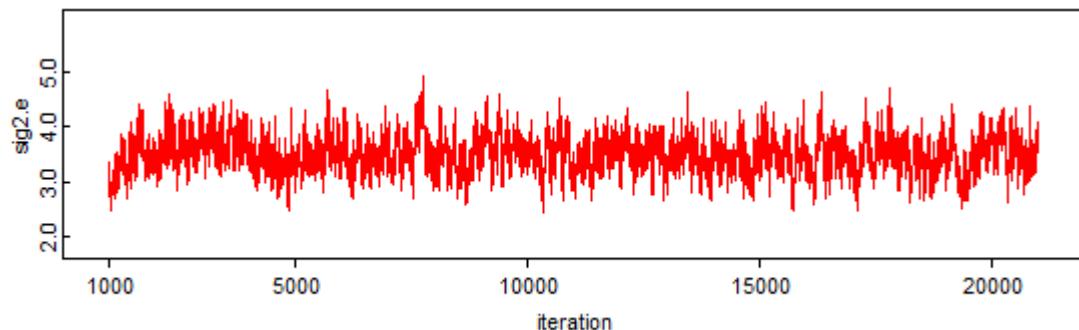


نمودار های تاریخچه اثرات تصادفی با پیشین $N(0,0.1)$



همانطور که مشاهده می‌کنید نمونه‌ها به خوبی شبیه سازی نشده‌اند.

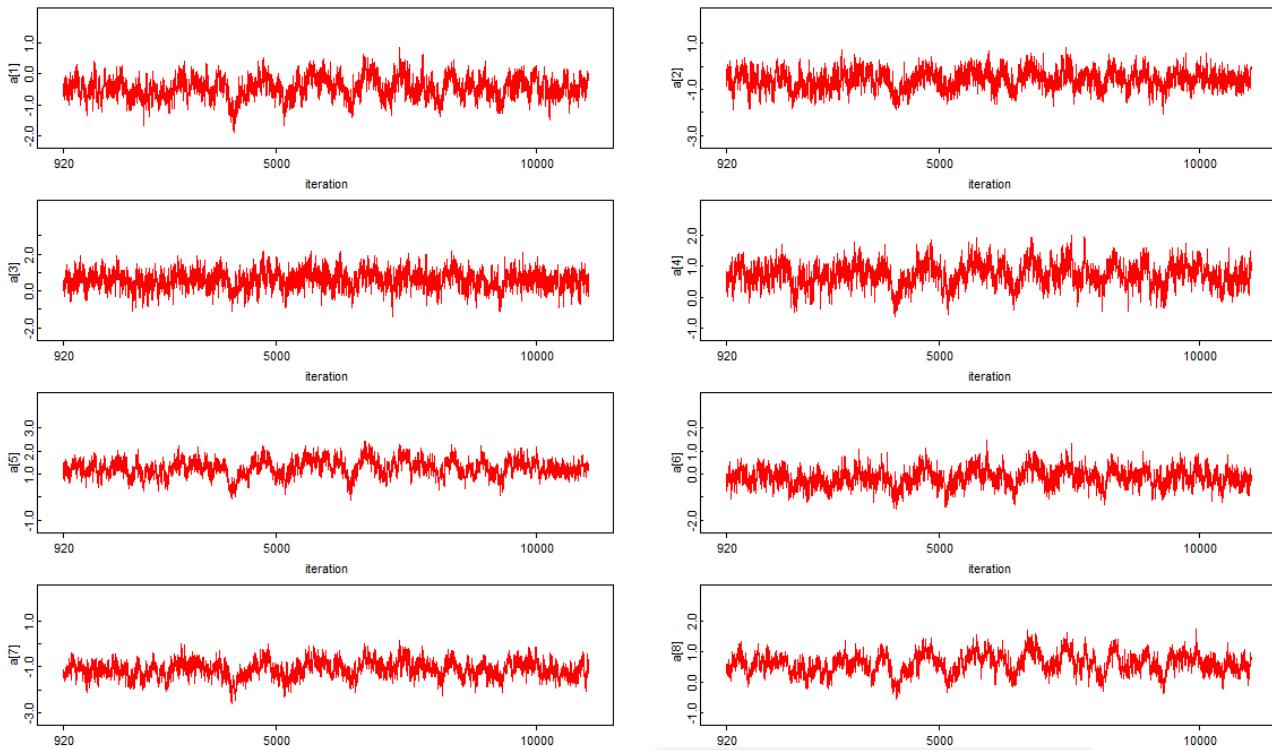
بعد از تولید ۱۰۰۰۰ نمونه دیگر همگرایی واریانس و متغیرهای پنهان بهتر می‌شود.



همانطور که در شکل زیر مشاهده می‌کنید DIC این مدل برابر ۲۶۷۱۰ است.

	Dbar	Dhat	DIC	pD
HR	17330.0	13820.0	20840.0	3512.0
z	5872.0	5872.0	5872.0	-6.659E-10
total	23200.0	19690.0	26710.0	3512.0

اگر توزیع پیشین اثرات تصادفی را $N(0,1)$ بگیریم نمودارهای تاریخچه بهتر می‌شود.



مدل هفتم:

در این مدل با در نظر گرفتن اثر تصادفی هر پانل و توزیع t برای جمله خطای داریم:

$$HR_i = X_i * \beta + a[id[i]]$$

$$X_i = [\text{MFIL}, \text{PO}, \text{RD}, \text{PS}, \text{UE}, \text{DV}, \text{MA}, \text{FP}, \text{BLK}, \text{FH}]$$

$$\beta = [\beta_0, \dots, \beta_{11}]$$

$$\epsilon_i \sim t(0, \sigma^2, \nu)$$

توزیع پیشین β را نرمال چند متغیره، σ^2 را معکوس گاما و ν را نمایی در نظر می‌گیریم.

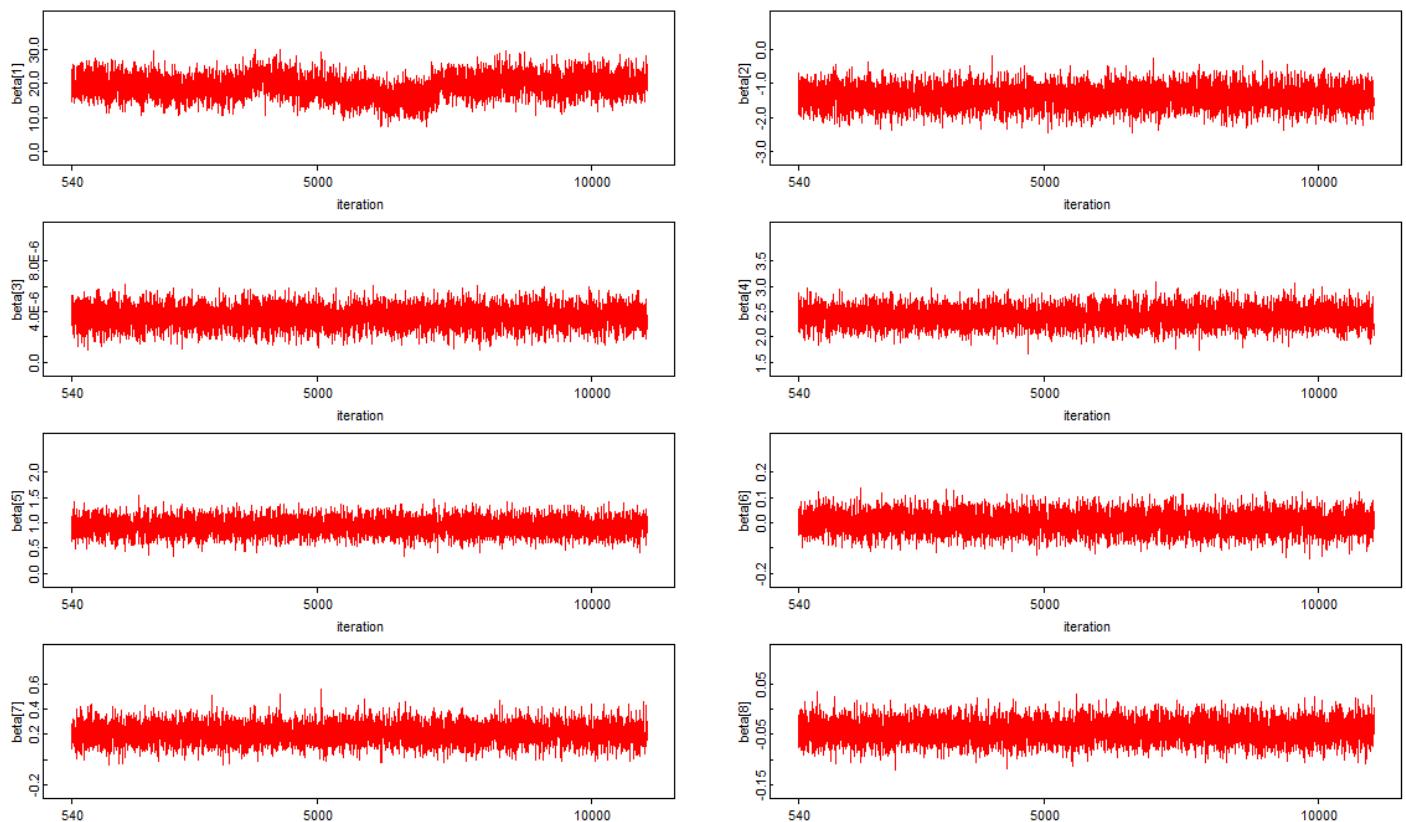
هر سه توزیع را آگاهی نابخش (واریانس زیاد و میانگین صفر) در نظر گرفتیم.

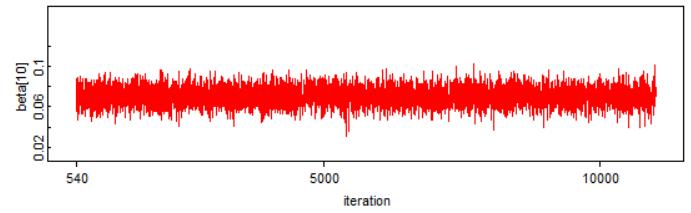
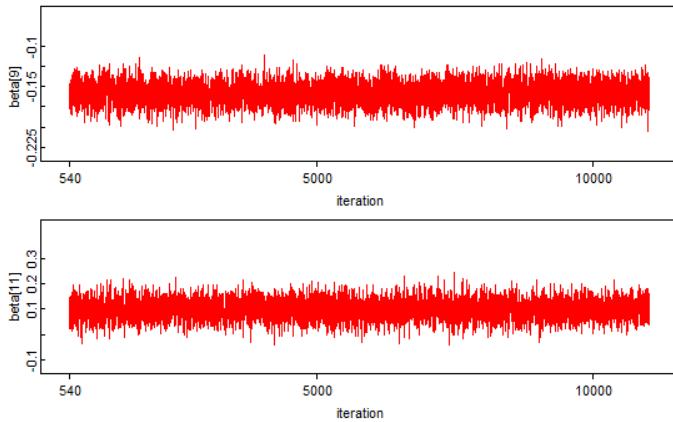
تخمین β ها و واریانس آنها را می‌توان در جدول زیر مشاهده کرد:

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta[1]	19.19	3.169	0.1592	12.87	19.27	25.18	541	10460
beta[2]	-1.4	0.2975	0.004884	-1.976	-1.4	-0.8101	541	10460
beta[3]	3.619E-6	7.484E-7	1.158E-8	2.092E-6	3.634E-6	5.048E-6	541	10460
beta[4]	2.403	0.1732	0.002595	2.07	2.402	2.746	541	10460
beta[5]	0.9328	0.1515	0.002525	0.6397	0.9333	1.234	541	10460
beta[6]	0.001382	0.03749	5.11E-4	-0.07165	0.001309	0.0748	541	10460
beta[7]	0.2087	0.07224	0.001186	0.06803	0.2079	0.3492	541	10460
beta[8]	-0.0436	0.0197	2.814E-4	-0.08158	-0.04392	-0.004568	541	10460
beta[9]	-0.1589	0.01248	2.111E-4	-0.183	-0.159	-0.1341	541	10460
beta[10]	0.07041	0.0086	1.374E-4	0.05364	0.07039	0.087	541	10460
beta[11]	0.09842	0.03699	5.809E-4	0.02616	0.09829	0.1716	541	10460

همانطور که مشاهده می‌کنید واریانس تخمین عرض از مبدأ زیاد می‌باشد بنابراین به آن کمتر می‌توان اعتماد کرد.

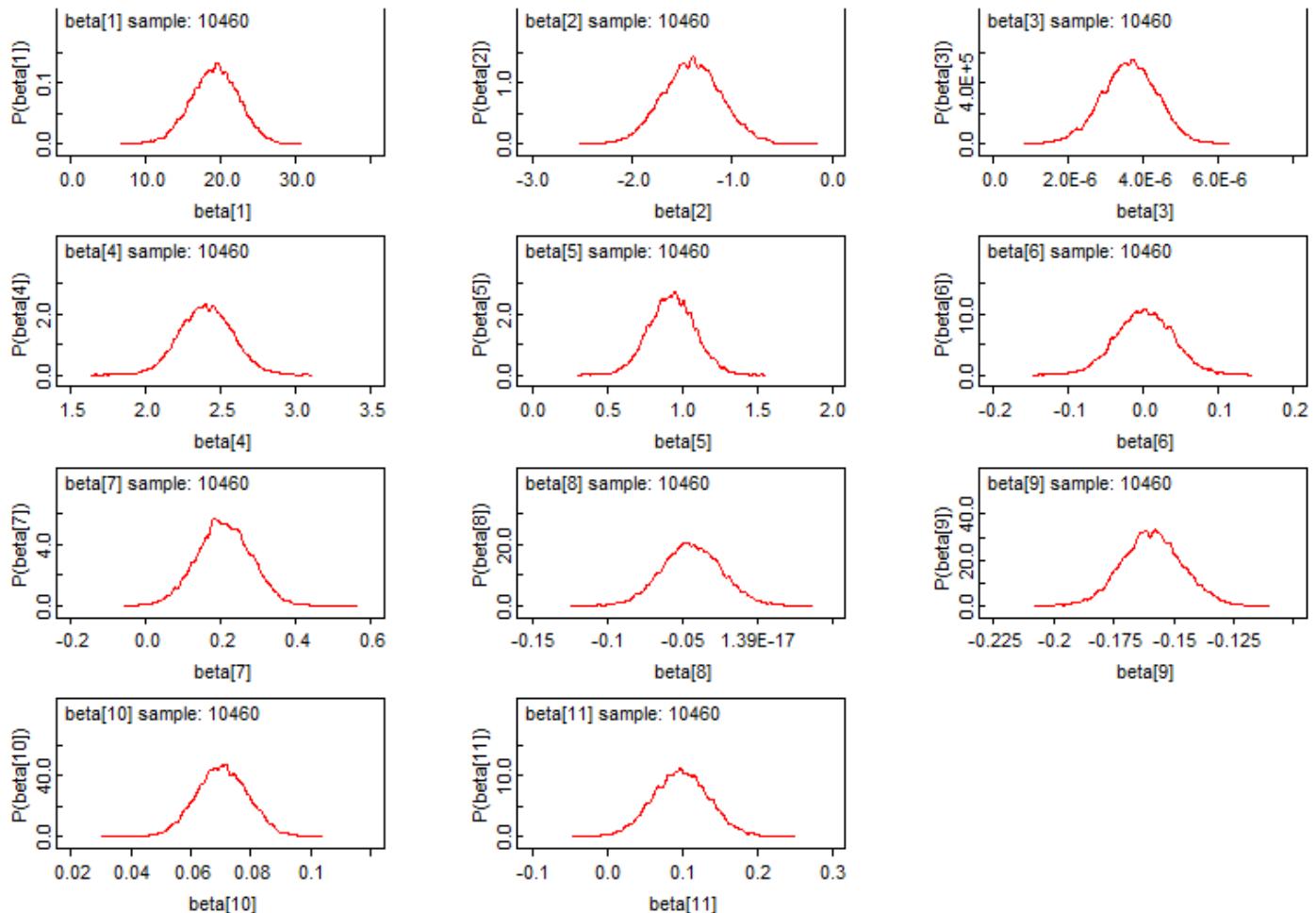
در ادامه نمودارهای تاریخچه بتا را رسم می‌کنیم.

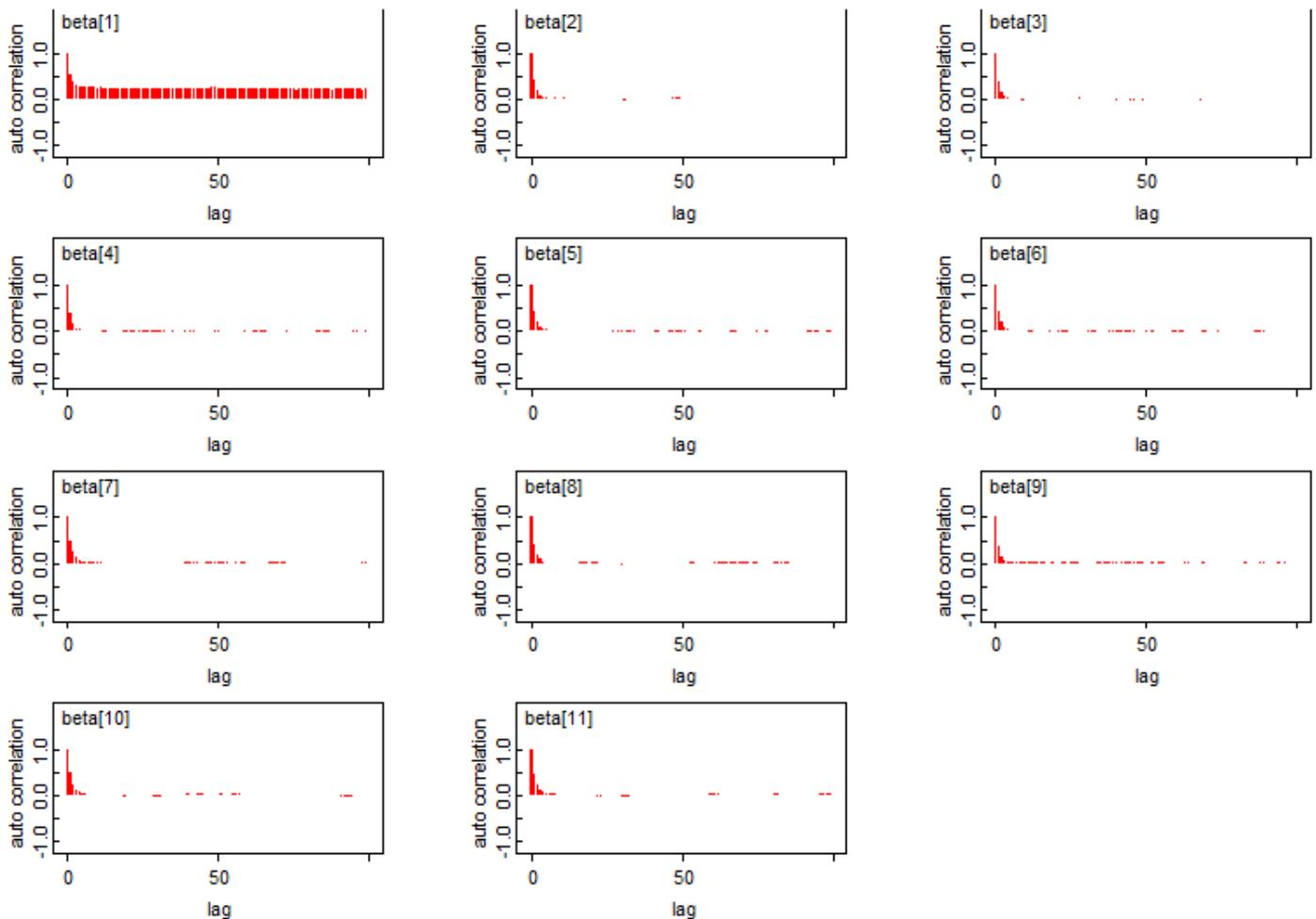




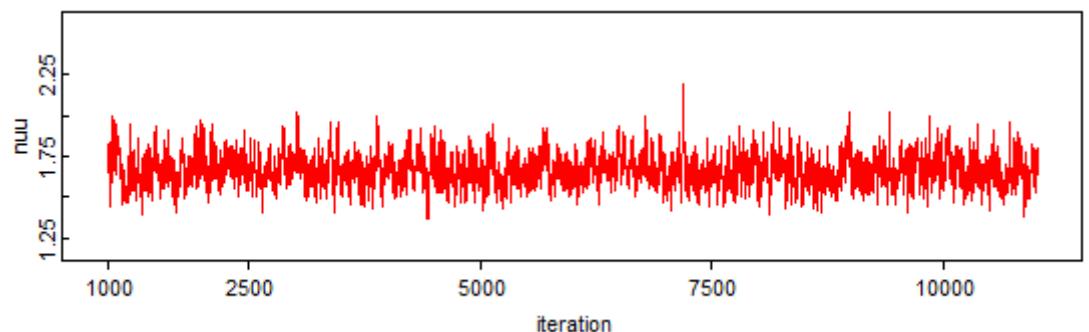
همگرایی ها بسیار خوب است.

در ادامه نمودارهای توزیع و همبستگی نمونه ها رسم شده است.

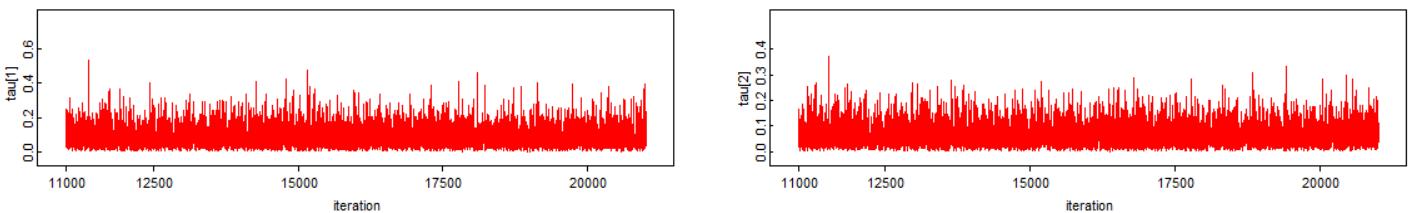




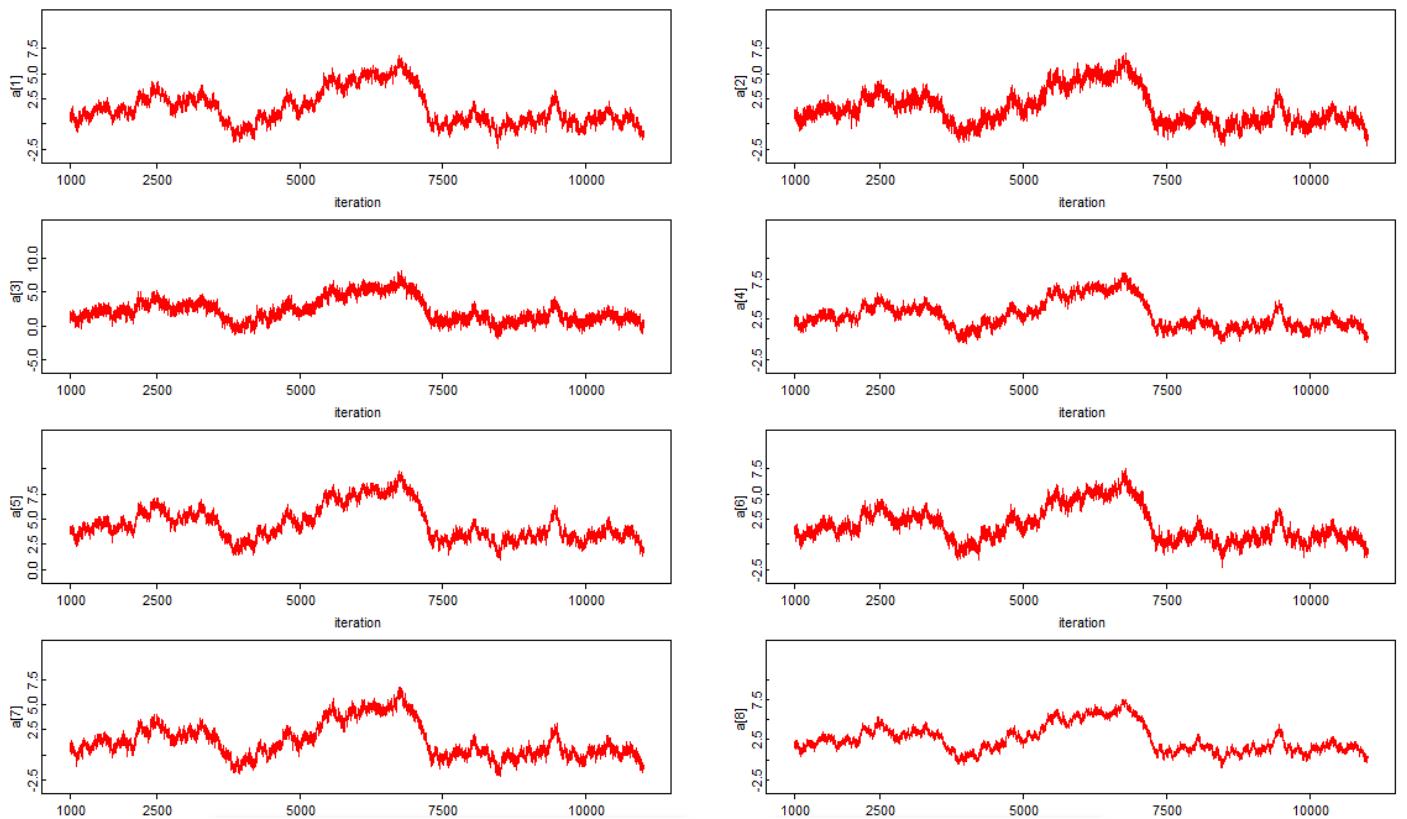
نمودار متغیر پنجم: nuu



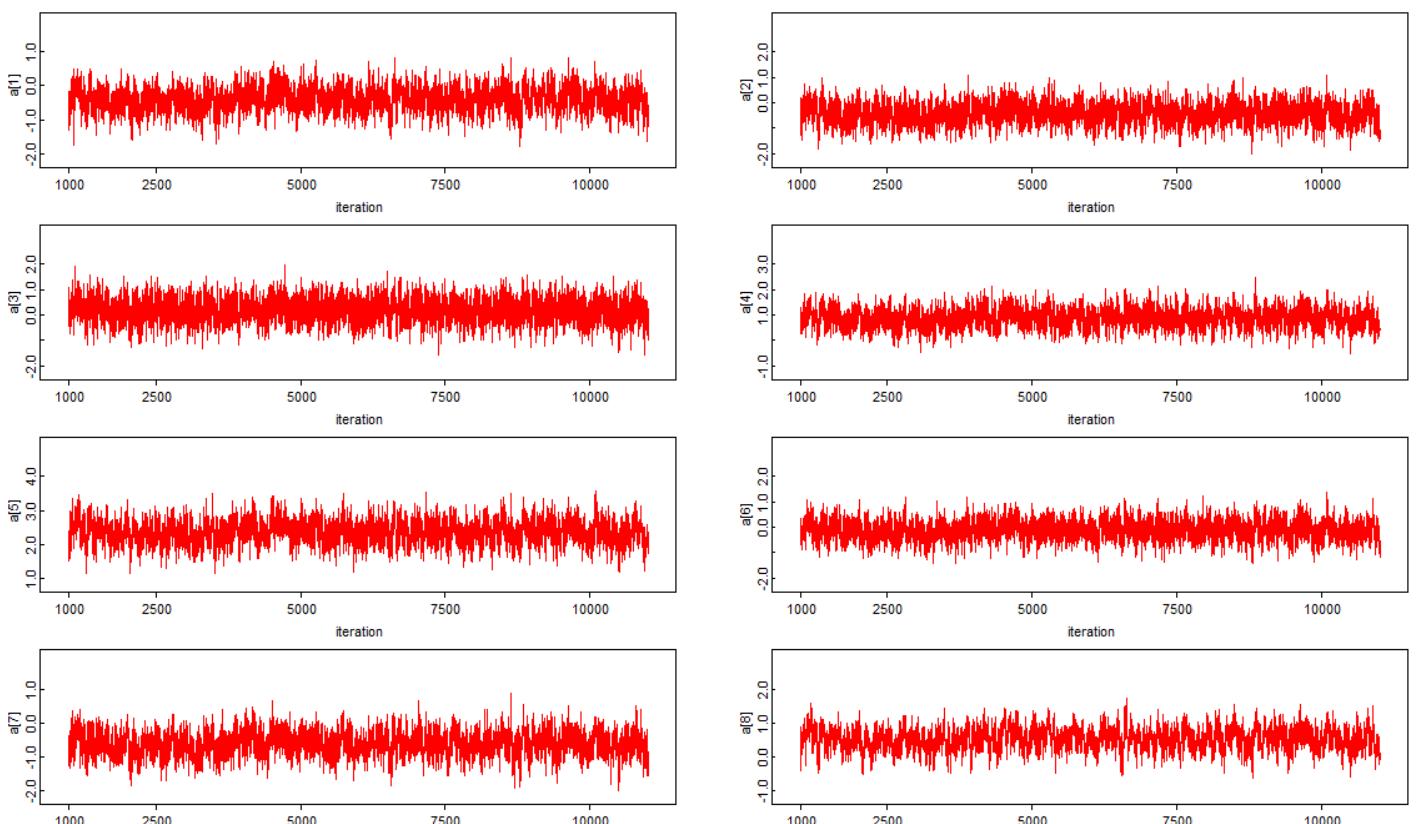
نمودار تاریخچه واریانس برای دو رکورد اول:



با در نظر گرفتن پیشین $N(0, 0.01)$ برای اثرات تصادفی داریم:



که اصلا همگرا نیست ولی اگر $N(0,1)$ بگیریم به فرم زیر است:



در پایان مشاهده می‌کنیم که DIC مدل برابر ۲۵۹۲۰ است:

	Dbar	Dhat	DIC	pD
HR	24840.0	23770.0	25920.0	1071.0
total	24840.0	23770.0	25920.0	1071.0

مدل هشتم:

در این مدل با در نظر گرفتن هر ایالت به عنوان یک پانل و توزیع SQ_t برای جمله خطای داریم:

$$HR_i = X_i * \beta + a[id[i]]$$

$$X_i = [\text{MFIL}, \text{PO}, \text{RD}, \text{PS}, \text{UE}, \text{DV}, \text{MA}, \text{FP}, \text{BLK}, \text{FH}]$$

$$\beta = [\beta_0, \dots, \beta_{11}]$$

$$\epsilon_i \sim St(0, \sigma^2, \nu, \delta)$$

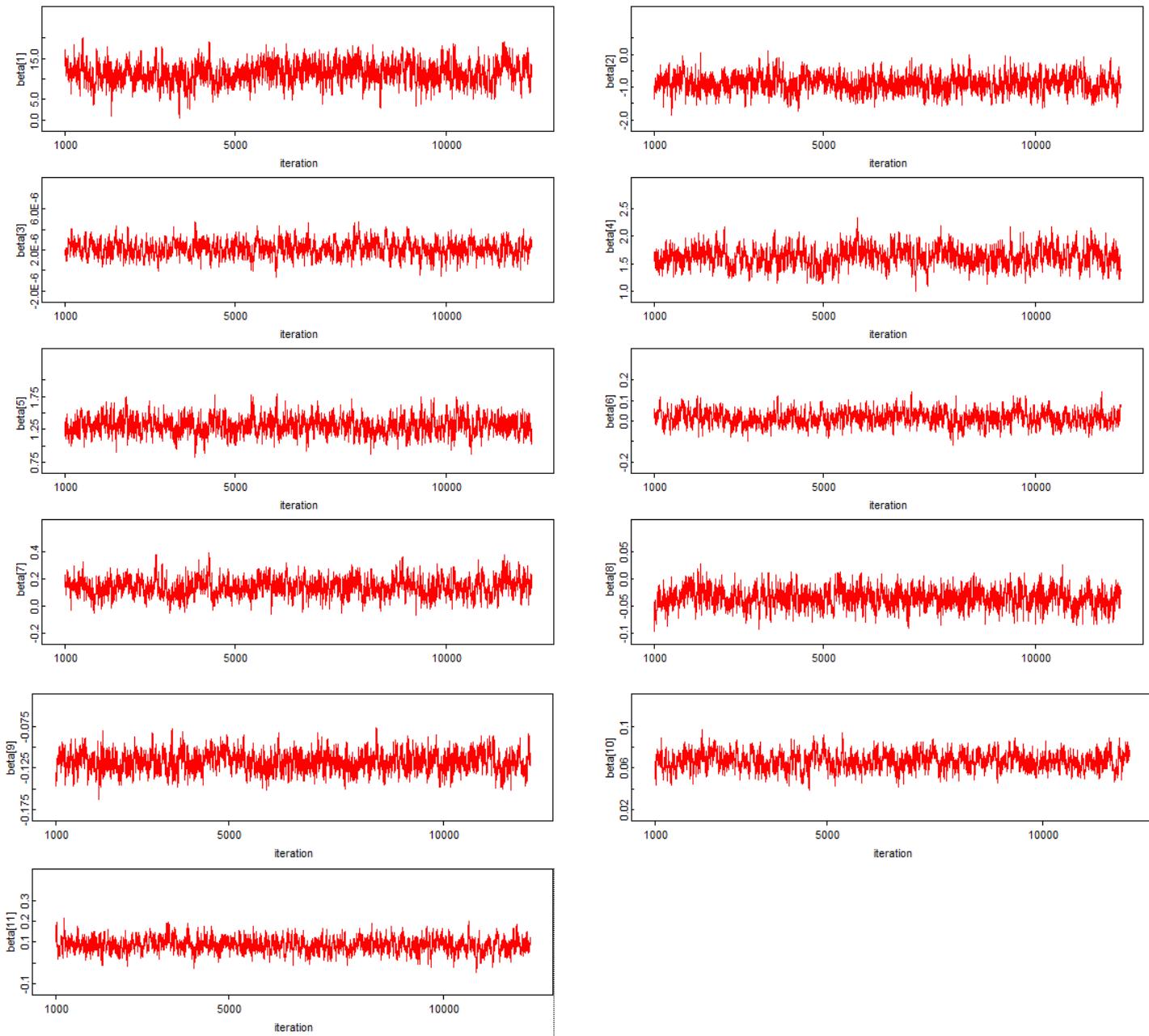
توزیع پیشین β را نرمال چند متغیره ، σ^2 را معکوس گاما ، ν را نمایی و δ را نرمال در نظر می‌گیریم.

هر چهار توزیع را آگاهی نابخش(واریانس زیاد و میانگین صفر) در نظر گرفتیم.
تخمین β ها و واریانس آنها را می‌توان در جدول زیر مشاهده کرد :

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta[1]	11.44	2.465	0.1235	6.578	11.44	16.27	1001	11000
beta[2]	-0.8878	0.2535	0.01159	-1.38	-0.8923	-0.3796	1001	11000
beta[3]	2.137E-6	7.335E-7	3.399E-8	6.885E-7	2.137E-6	3.582E-6	1001	11000
beta[4]	1.63	0.1707	0.008978	1.304	1.63	1.971	1001	11000
beta[5]	1.306	0.1254	0.005775	1.066	1.303	1.563	1001	11000
beta[6]	0.0163	0.03399	0.00175	-0.04816	0.01584	0.08369	1001	11000
beta[7]	0.139	0.06328	0.003101	0.01373	0.1389	0.2665	1001	11000
beta[8]	-0.03538	0.01645	7.653E-4	-0.06817	-0.03543	-0.003146	1001	11000
beta[9]	-0.1175	0.01119	5.311E-4	-0.1393	-0.1176	-0.09519	1001	11000
beta[10]	0.06672	0.007794	3.788E-4	0.05122	0.06673	0.08173	1001	11000
beta[11]	0.08657	0.03238	0.001452	0.02341	0.08602	0.1504	1001	11000

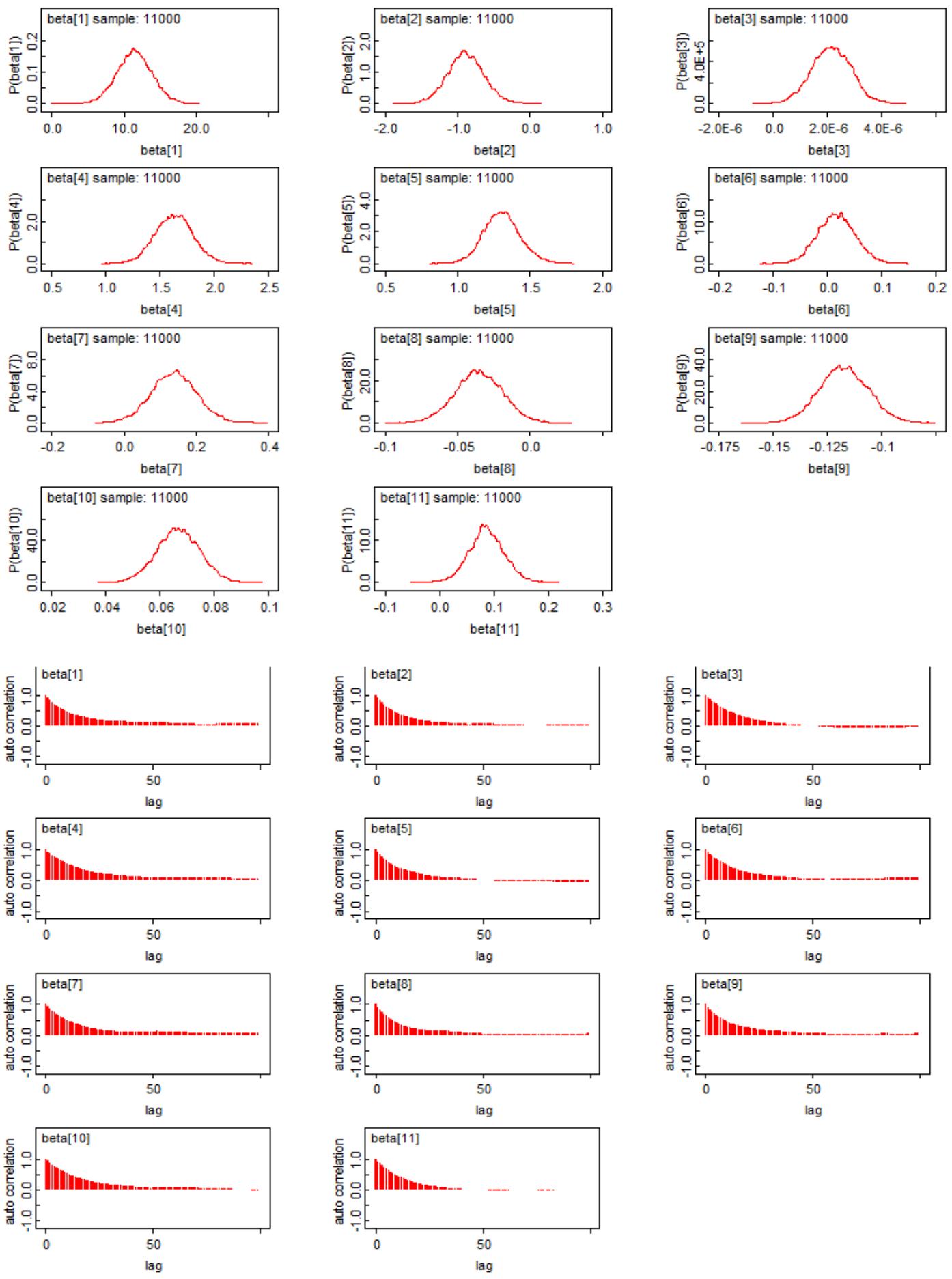
همانطور که مشاهده می‌کنید واریانس تخمین [عرض از مبدأ](#) زیاد می‌باشد بنابراین به آن کمتر می‌توان اعتماد کرد.

در ادامه نمودارهای تاریخچه بتا را رسم می‌کنیم:

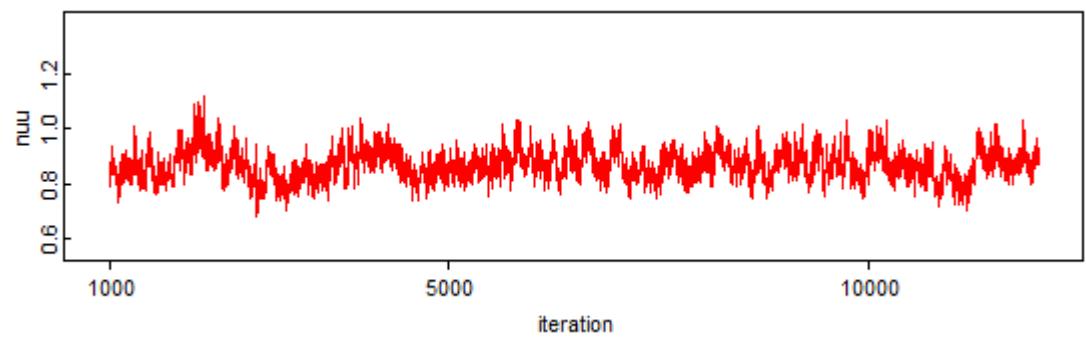


همگرایی ها نسبتا خوب است.

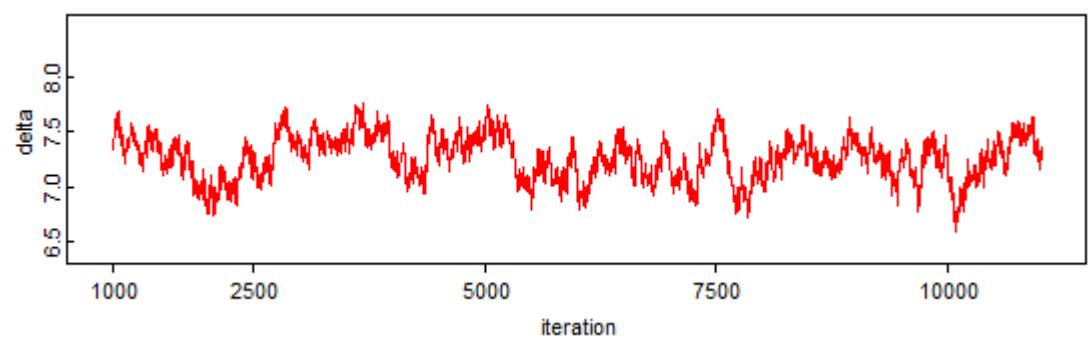
در ادامه نمودارهای توزیع و همبستگی نمونه ها رسم شده است:



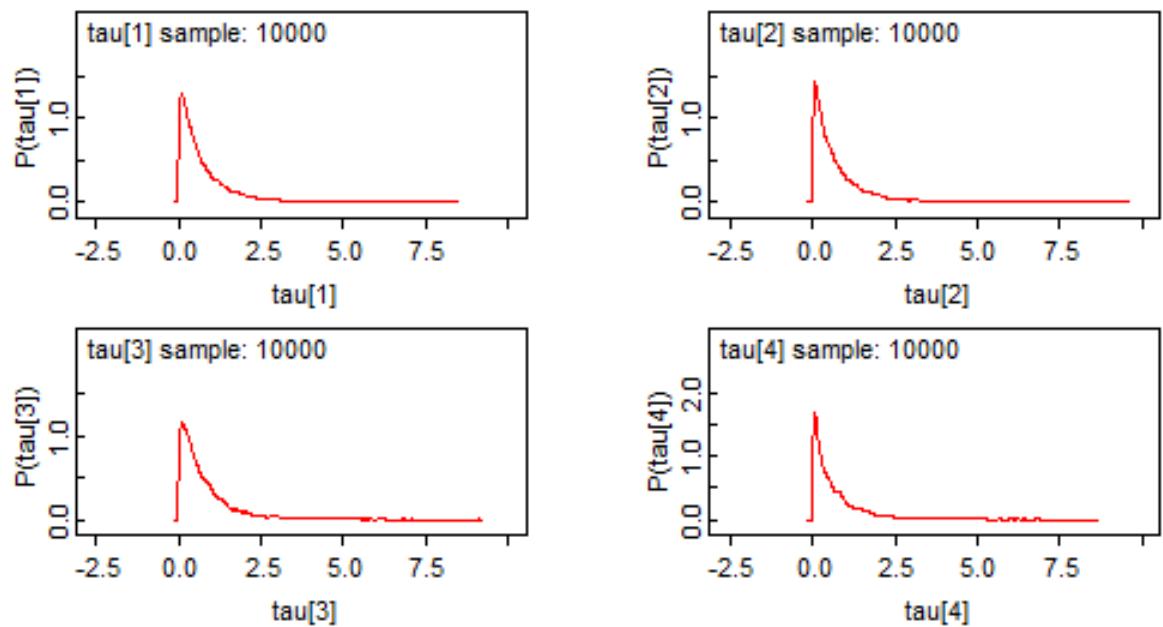
نمودار تاریخچه ν :



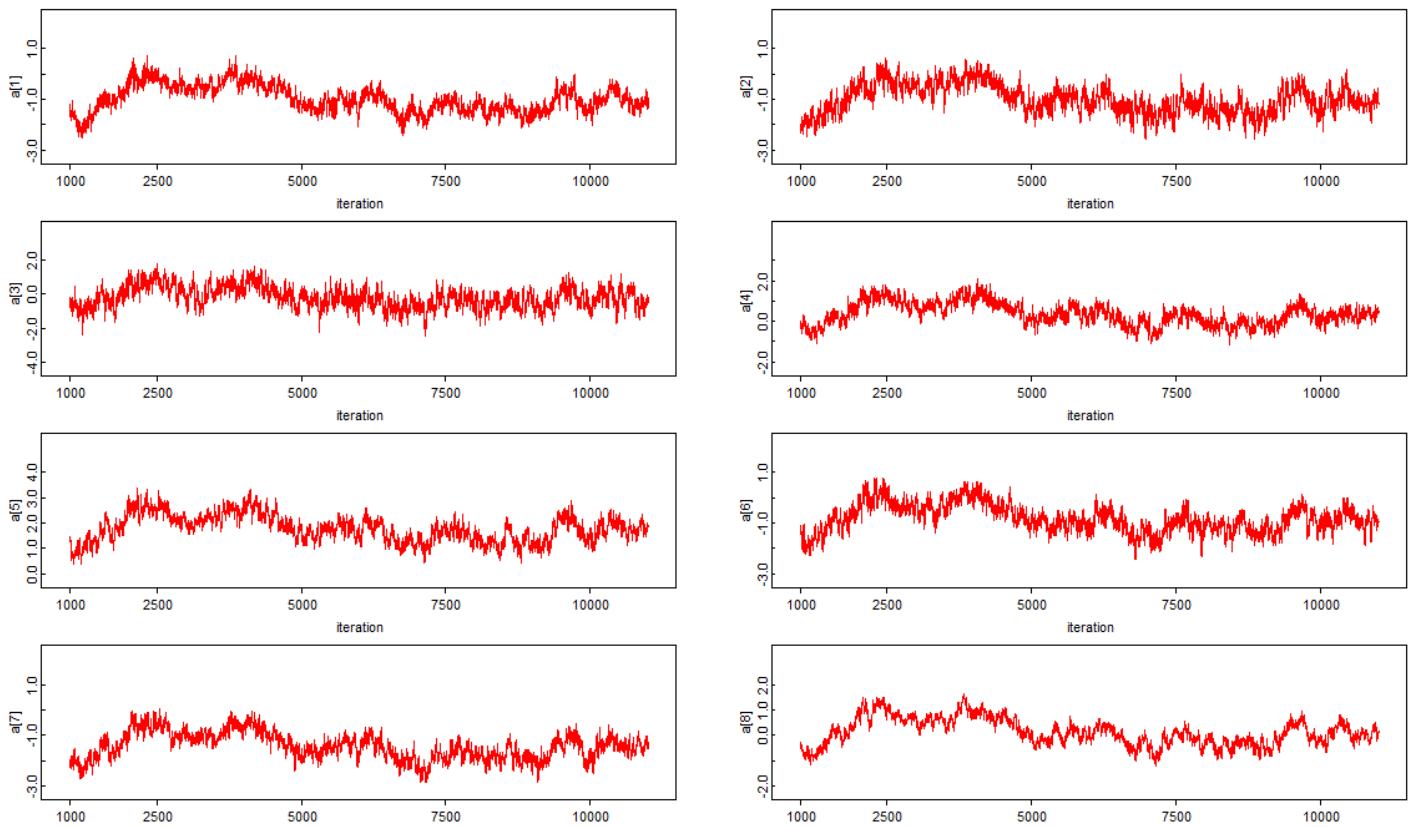
نمودار تاریخچه δ :



تابع چگالی بردار واریانس:



تاریخچه اثرات تصادفی با توزیع پیشین $N(0,0.1)$

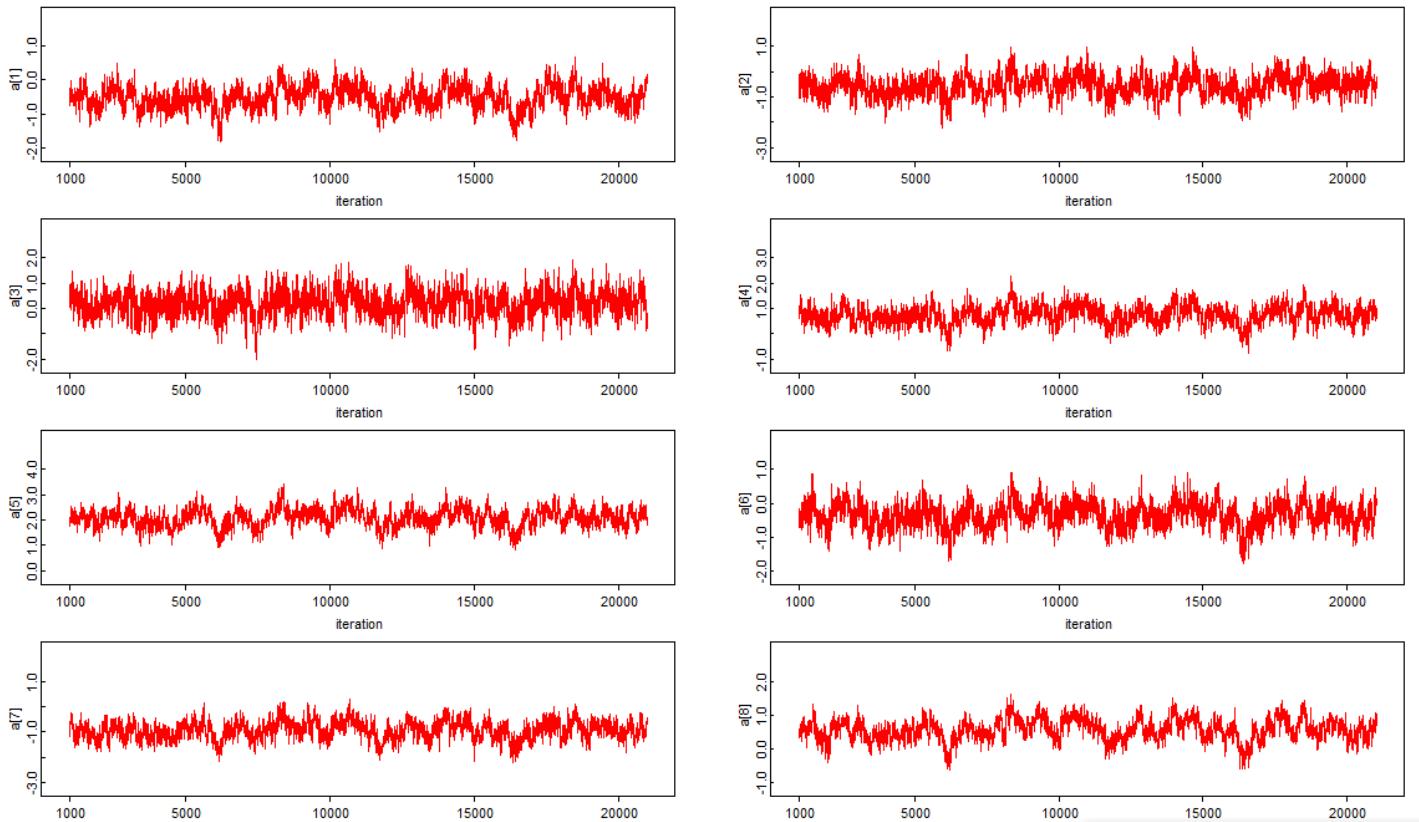


نمونه های اثرات تصادفی همگرا نیستند.

مقدار DIC کل این مدل برابر ۲۵۲۵۰ است.

	Dbar	Dhat	DIC	pD
HR	16610.0	13850.0	19380.0	2763.0
z	5872.0	5872.0	5872.0	-6.659E-10
total	22490.0	19720.0	25250.0	2763.0

تاریخچه اثرات تصادفی با توزیع پیشین $N(0,1)$ و بعد از گرفتن ۱۰۰۰۰ نمونه دیگر به فرم زیر است:



رگرسیون خطی با
تأخیر مرتبه اول

Linear Regression
with First time
delay

مدل نهم:

در این مدل هر ایالت به عنوان یک پانل در ۴ سال با متغیر هدف HR و وابسته به سال قبل از آن است.

توزیع جمله خطأ نرمال است.

$$HR_{it} = X_i * \beta + \gamma HR_{it-1} + \epsilon_i$$

$$X_i = [\text{MFIL}, \text{PO}, \text{RD}, \text{PS}, \text{UE}, \text{DV}, \text{MA}, \text{FP}, \text{BLK}, \text{FH}]$$

$$\beta = [\beta_0, \dots, \beta_{11}]$$

$$\gamma \sim N(0, 0.01)$$

$$\epsilon_i \sim N(0, \sigma^2)$$

توزیع پیشین β را نرمال چند متغیره ، σ^2 را معکوس گاما در نظر می‌گیریم.

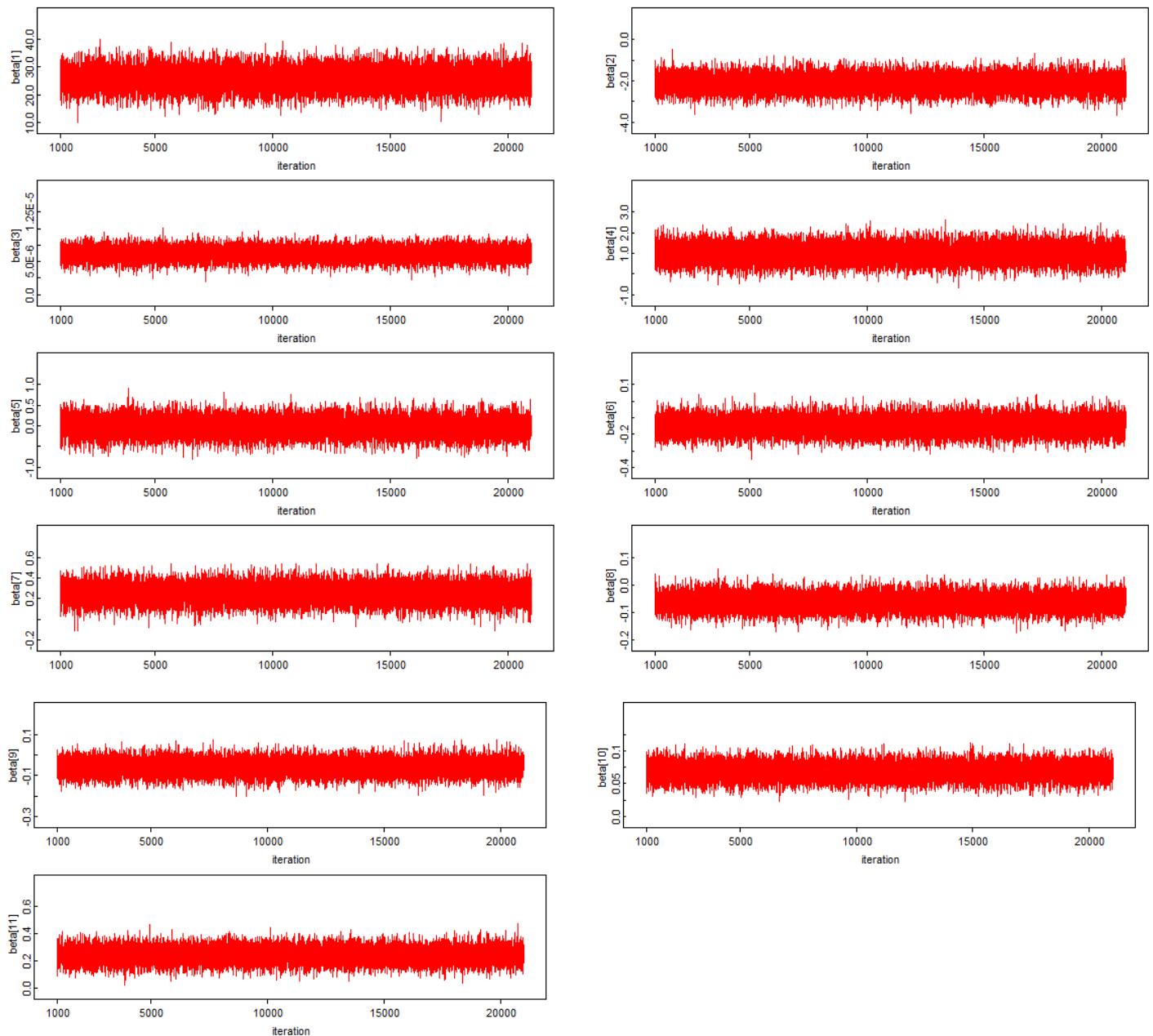
هر دو توزیع را آگاهی نابخش(واریانس زیاد و میانگین صفر) در نظر گرفتیم

تخمین β ها و واریانس آنها را می‌توان در جدول زیر مشاهده کرد :

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta[1]	25.76	3.683	0.02828	18.59	25.76	32.99	1001	20000
beta[2]	-2.121	0.3817	0.002963	-2.868	-2.119	-1.368	1001	20000
beta[3]	6.163E-6	9.382E-7	6.713E-9	4.331E-6	6.163E-6	8.023E-6	1001	20000
beta[4]	0.9976	0.4114	0.003104	0.1886	0.9989	1.806	1001	20000
beta[5]	-0.02078	0.2081	0.001469	-0.4275	-0.02082	0.384	1001	20000
beta[6]	-0.1465	0.04936	3.682E-4	-0.2434	-0.1463	-0.0489	1001	20000
beta[7]	0.2497	0.08549	6.219E-4	0.08168	0.251	0.4163	1001	20000
beta[8]	-0.06327	0.0283	1.81E-4	-0.1187	-0.06325	-0.007589	1001	20000
beta[9]	-0.05863	0.03567	2.623E-4	-0.1293	-0.05859	0.01087	1001	20000
beta[10]	0.06865	0.01191	7.83E-5	0.04494	0.06872	0.0919	1001	20000
beta[11]	0.2408	0.05368	3.338E-4	0.1356	0.2407	0.3453	1001	20000

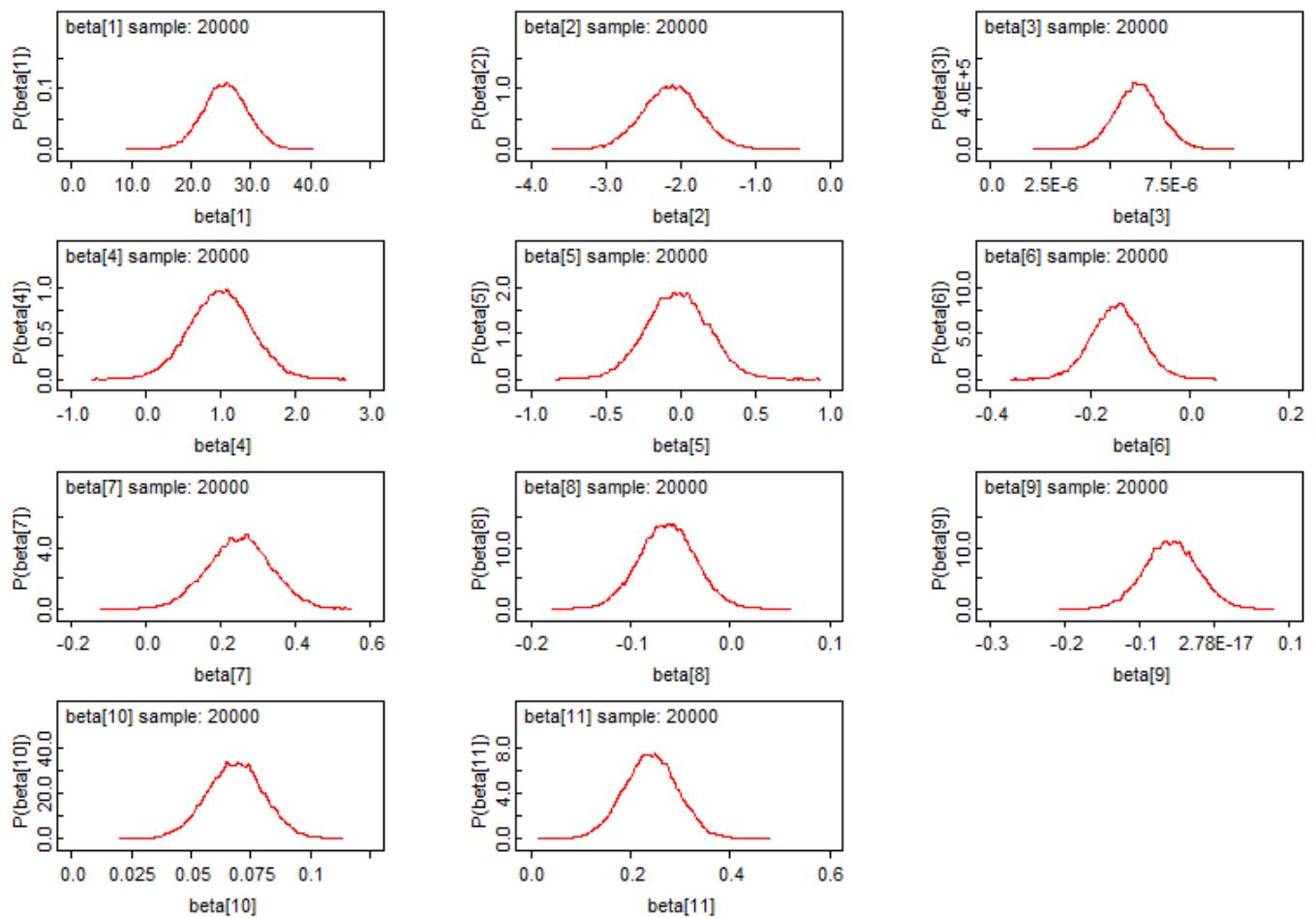
همانطور که مشاهده می‌کنید واریانس تخمین عرض از مبدأ زیاد می‌باشد بنابراین به آن کمتر می‌توان اعتماد کرد.

نمودارهای تاریخچه β در صفحه بعد مشخص است:



با توجه به نمودار های بالا میتوان دید که β های تولید شده توسط این مدل همگرا شده اند و مدل با موفقیت پیاده سازی شده است.

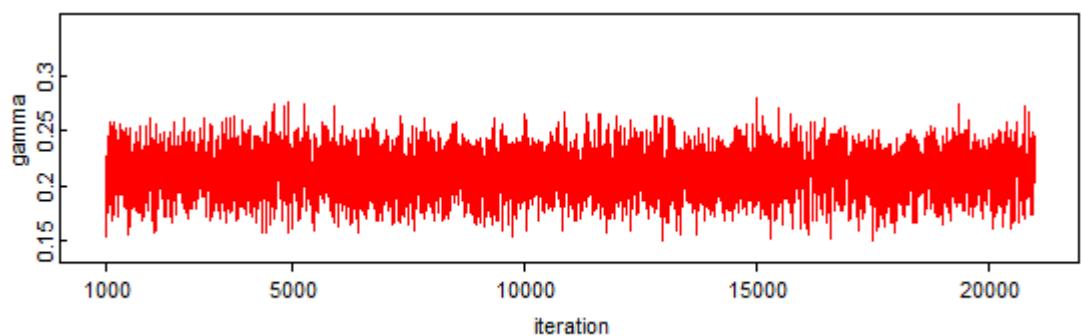
نمودار های چگالی β :

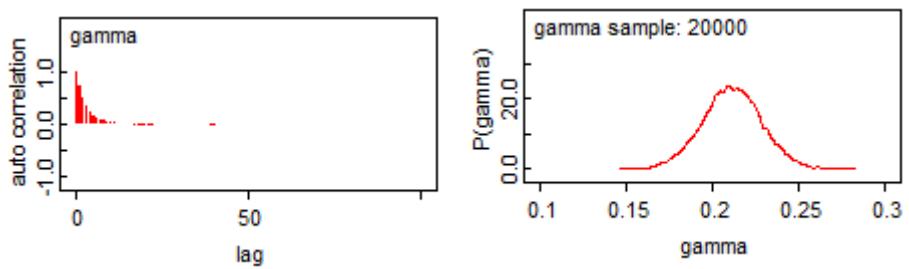


با توجه به این نمودارها می‌توان گفت که توزیع‌های نرمال به خوبی شبیه سازی شده‌اند.

در ادامه برآورد بیز و نمودار تاریخچه گاما را نشان می‌دهیم.

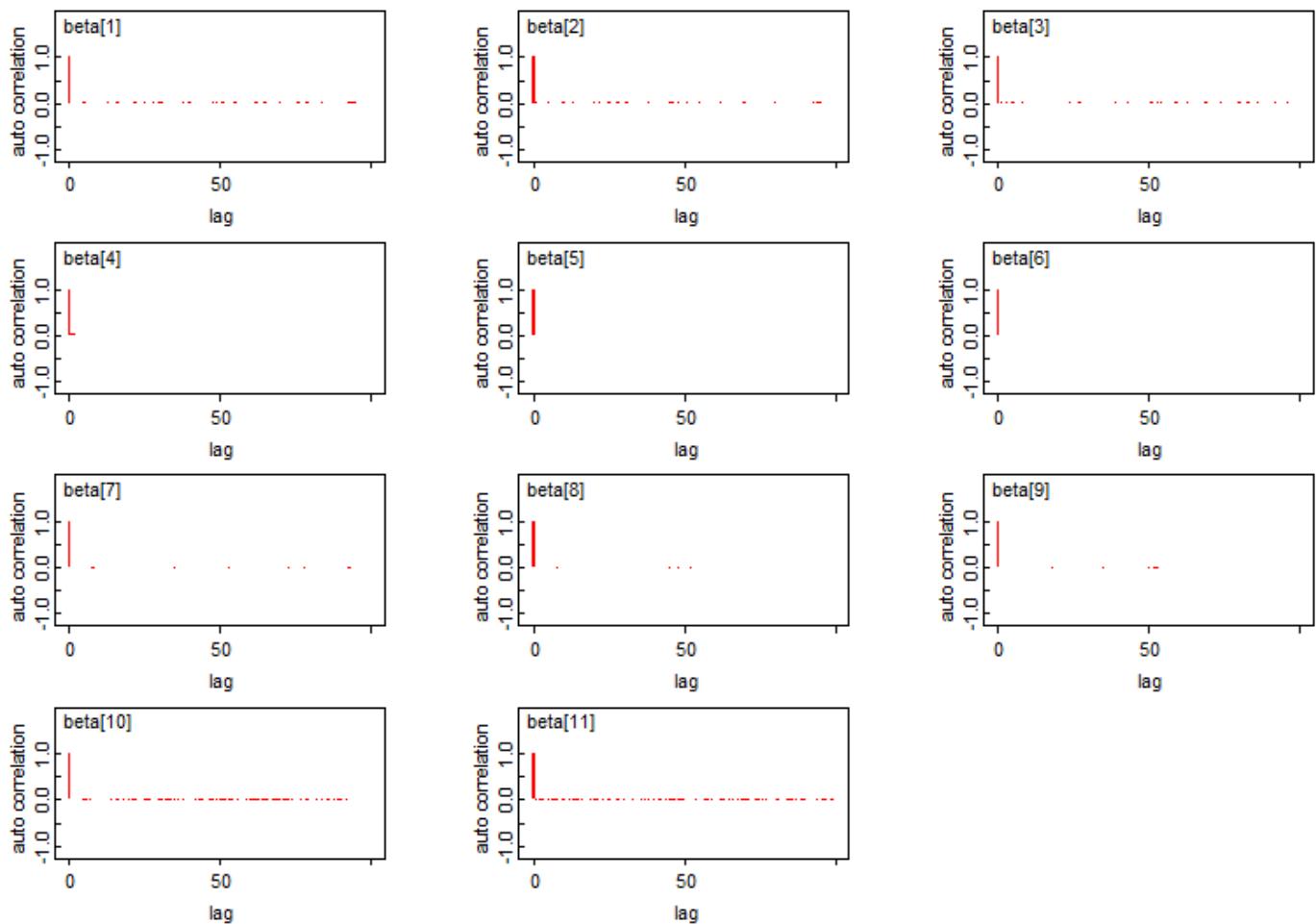
	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
gamma	0.2111	0.01718	2.899E-4	0.177	0.2112	0.245	1001	20000





بنابراین ضریب گاما به خوبی شبیه سازی شده است.

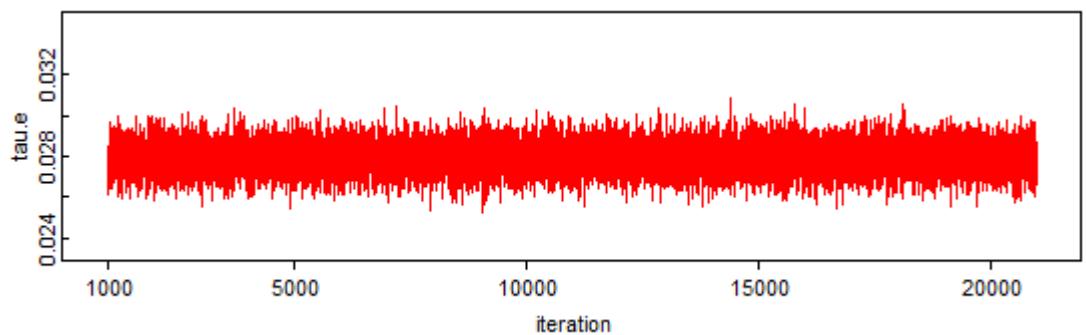
نمودار های خود همبستگی β



با توجه به نمودار های بالا میتوان گفت که نمونه ها به صورت غیروابسته تولید شده اند.

نمودار تاریخچه واریانس را در شکل زیر مشاهده می کنید:

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
tau.e	0.0279	7.03E-4	4.608E-6	0.02654	0.02788	0.02929	1001	20000



برای زمان اول که همان سال ۱۹۶۰ است مدل رگرسیون ساده می‌زنیم که نتایج در زیر مشخص است.

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta1[1]	-4.371	9.635	0.0669	-23.46	-4.383	14.42	1001	20000
beta1[2]	0.776	1.234	0.009599	-1.645	0.7704	3.24	1001	20000
beta1[3]	5.061E-6	2.668E-6	1.842E-8	-2.153E-7	5.077E-6	1.03E-5	1001	20000
beta1[4]	-2.267	2.339	0.01681	-6.842	-2.275	2.377	1001	20000
beta1[5]	-0.8656	0.3411	0.002188	-1.532	-0.8662	-0.1911	1001	20000
beta1[6]	-0.108	0.09183	6.88E-4	-0.2868	-0.1074	0.0717	1001	20000
beta1[7]	0.8098	0.2753	0.001999	0.2741	0.8102	1.351	1001	20000
beta1[8]	-0.1885	0.04619	3.304E-4	-0.2801	-0.1886	-0.09781	1001	20000
beta1[9]	0.114	0.09457	6.35E-4	-0.07131	0.1141	0.2998	1001	20000
beta1[10]	0.1093	0.03321	2.384E-4	0.04382	0.1094	0.1744	1001	20000
beta1[11]	0.288	0.1817	0.001399	-0.06886	0.2866	0.6439	1001	20000
	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
tau.e1	0.02774	0.001209	9.21E-6	0.02541	0.02773	0.03011	1001	20000

با توجه به شکل زیر DIC مدل برابر ۲۷۲۰ است.

	Dbar	Dhat	DIC	pD
HR	20390.0	20370.0	20400.0	12.91
HR1	6802.0	6791.0	6813.0	11.04
gamma	6.446	6.446	6.446	7.523E-13
total	27200.0	27170.0	27220.0	23.95

مدل دهم:

حال مدل دهم را با استفاده از همان متغیرها و با در نظر گرفتن توزیع SQ_N برای جمله خطأ اجرا می کنیم.

$$HR_{it} = X_i * \beta + \gamma HR_{it-1} + \epsilon_i$$

$$X_i = [\text{MFIL}, \text{PO}, \text{RD}, \text{PS}, \text{UE}, \text{DV}, \text{MA}, \text{FP}, \text{BLK}, \text{FH}]$$

$$\beta = [\beta_0, \dots, \beta_{11}]$$

$$\gamma \sim N(0, 0.01)$$

$$\epsilon_i \sim SN(0, \sigma^2, \delta)$$

توزیع پیشین β را نرمال چند متغیره ، σ^2 را معکوس گاما و δ را نرمال در نظر می گیریم.

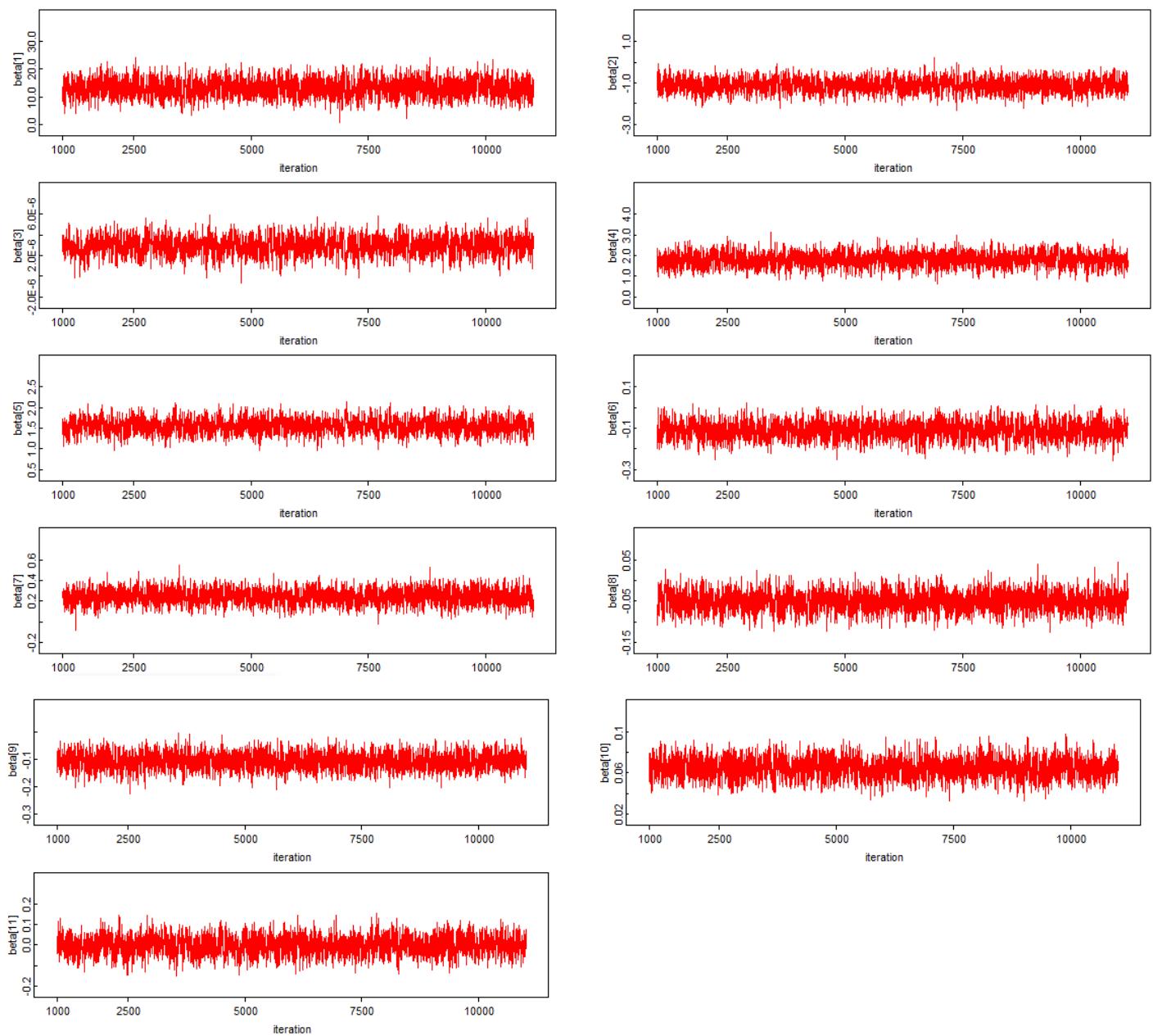
هر سه توزیع را آگاهی نابخش (واریانس زیاد و میانگین صفر) در نظر گرفتیم.

تخمین β ها و واریانس آنها را میتوان در جدول زیر مشاهده کرد :

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta[1]	13.23	2.961	0.07451	7.421	13.26	19.05	1001	10000
beta[2]	-1.125	0.3116	0.008232	-1.736	-1.127	-0.5211	1001	10000
beta[3]	2.947E-6	8.544E-7	2.813E-8	1.168E-6	2.982E-6	4.508E-6	1001	10000
beta[4]	1.799	0.3161	0.009287	1.176	1.798	2.424	1001	10000
beta[5]	1.553	0.1676	0.0053	1.232	1.549	1.884	1001	10000
beta[6]	-0.1098	0.03862	0.001106	-0.1856	-0.1094	-0.0343	1001	10000
beta[7]	0.2429	0.06663	0.0019	0.1102	0.2426	0.3724	1001	10000
beta[8]	-0.05043	0.02188	6.354E-4	-0.09376	-0.05062	-0.007735	1001	10000
beta[9]	-0.1081	0.02799	6.814E-4	-0.162	-0.1081	-0.05295	1001	10000
beta[10]	0.06474	0.009126	2.683E-4	0.04684	0.06482	0.08284	1001	10000
beta[11]	-0.002169	0.04238	0.001315	-0.08448	-0.002361	0.08074	1001	10000

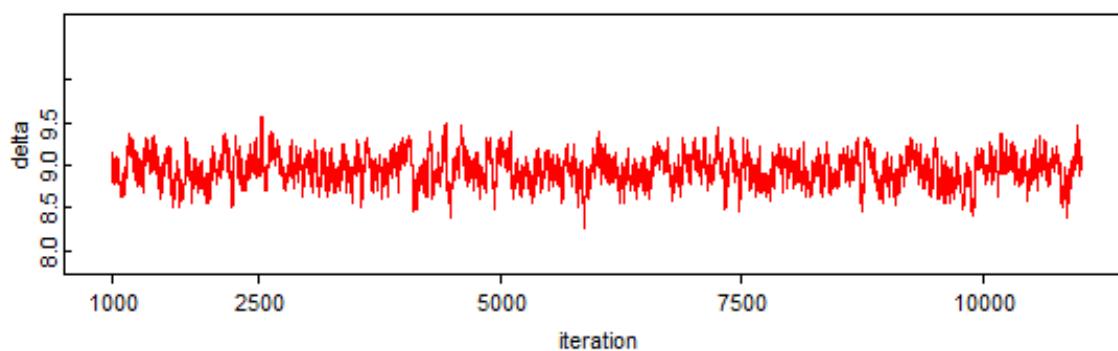
همانطور که مشاهده می کنید واریانس تخمین عرض از مبدأ زیاد می باشد بنابراین به آن کمتر می توان اعتماد کرد.

نمودار های تاریخچه β



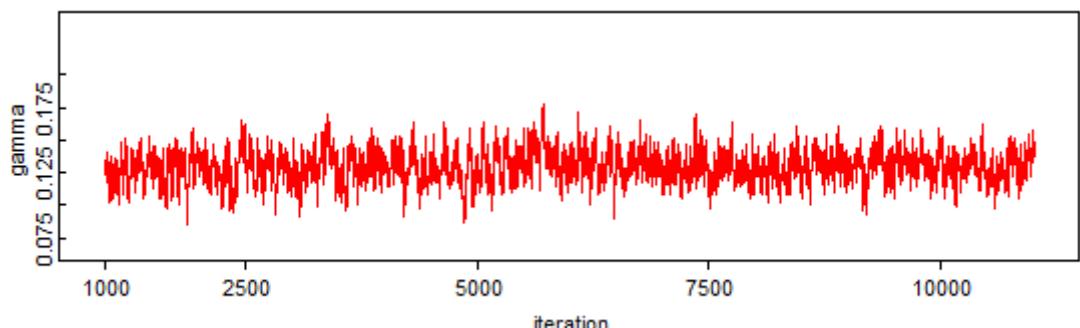
همگرایی نمودار های بالا کمی کمتر از مدل ششم است.

نمودار تاریخچه ۵ در شکل زیر مشخص است که همگرایی نسبتا خوبی دارد.

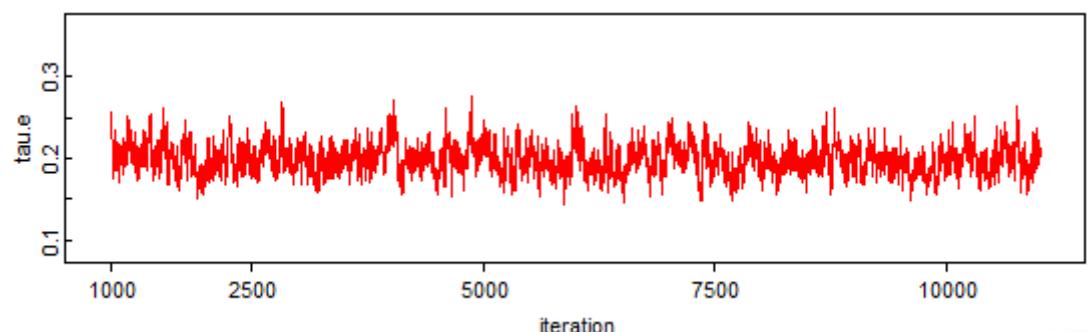


نمودار تاریخچه گاما:

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
gamma	0.1296	0.01219	5.707E-4	0.1055	0.13	0.1531	1001	10000



نمودار واریانس:



نمودار های سال ۱۹۶۰ نیز دقیقا مشابه است.

با توجه به شکل زیر DIC کل مدل برابر ۲۶۵۳۰ است.

	Dbar	Dhat	DIC	pD
HR	14170.0	11680.0	16650.0	2482.0
HR1	3006.0	2013.0	3999.0	993.4
gamma	6.446	6.446	6.446	6.954E-13
z	4404.0	4404.0	4404.0	-4.994E-10
z1	1468.0	1468.0	1468.0	-1.665E-10
total	23050.0	19570.0	26530.0	3476.0

مدل یازدهم:

حال مدل یازدهم را با استفاده از همان متغیرها و با در نظر گرفتن توزیع t برای جمله خطای اجرا می کنیم:

$$HR_{it} = X_i * \beta + \gamma HR_{it-1} + \epsilon_i$$

$$X_i = [\text{MFIL}, \text{PO}, \text{RD}, \text{PS}, \text{UE}, \text{DV}, \text{MA}, \text{FP}, \text{BLK}, \text{FH}]$$

$$\beta = [\beta_0, \dots, \beta_{11}]$$

$$\gamma \sim N(0, 0.01)$$

$$\epsilon_i \sim t(0, \sigma^2, \nu)$$

توزیع پیشین β را نرمال چند متغیره ، σ^2 را معکوس گاما و ν را نمایی در نظر می گیریم.

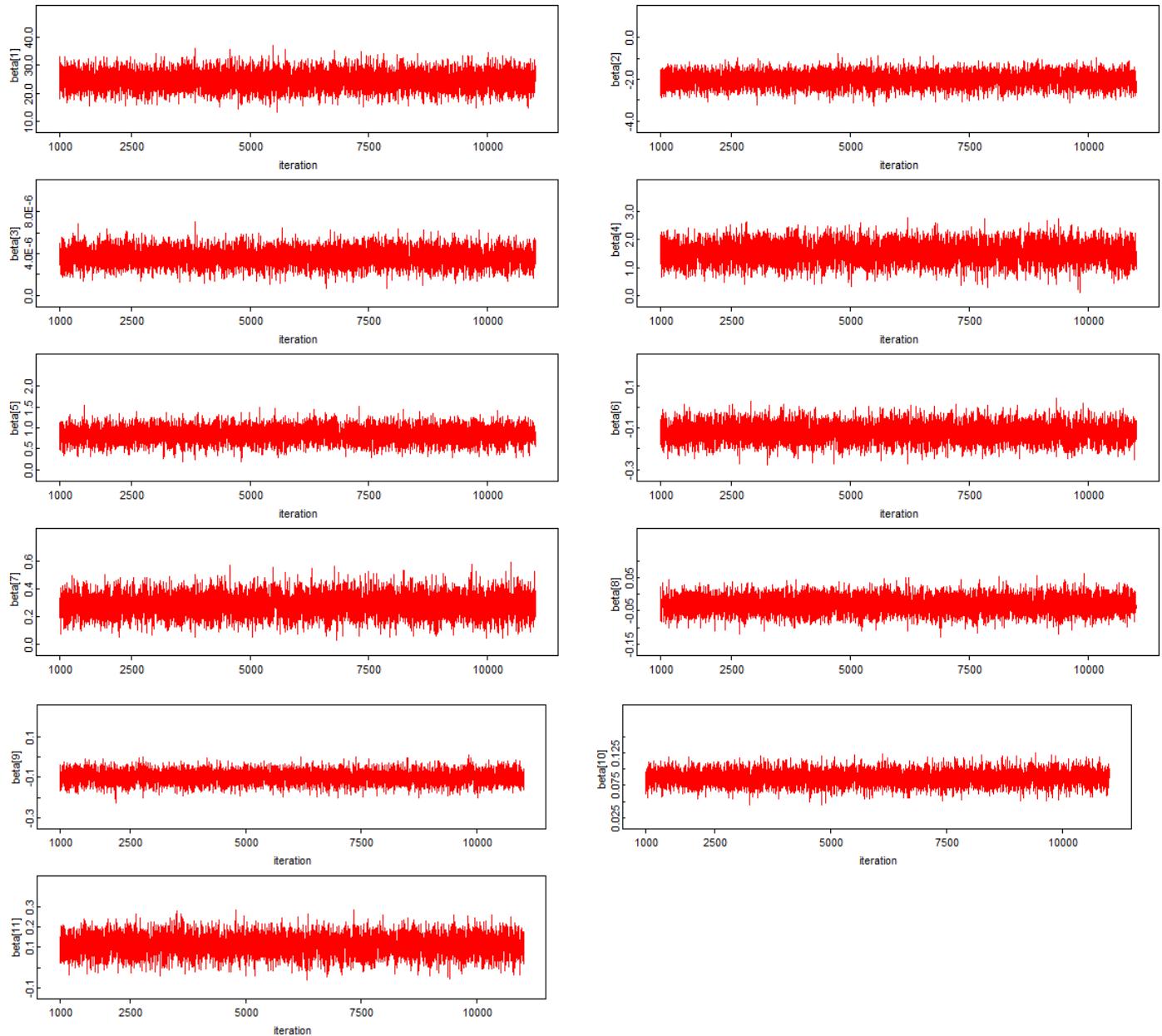
هر سه توزیع را آگاهی نابخش(واریانس زیاد و میانگین صفر) در نظر گرفتیم.

تخمین β ها و واریانس آنها را می توان در جدول زیر مشاهده کرد .

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta[1]	24.59	3.065	0.04038	18.72	24.56	30.63	1001	10000
beta[2]	-2.026	0.3239	0.004308	-2.664	-2.026	-1.405	1001	10000
beta[3]	3.672E-6	7.924E-7	9.759E-9	2.1E-6	3.678E-6	5.202E-6	1001	10000
beta[4]	1.525	0.3453	0.005303	0.8457	1.527	2.2	1001	10000
beta[5]	0.8444	0.1768	0.002366	0.4919	0.8459	1.188	1001	10000
beta[6]	-0.1223	0.04153	6.639E-4	-0.2025	-0.1228	-0.03983	1001	10000
beta[7]	0.283	0.07257	0.001109	0.1425	0.2817	0.4277	1001	10000
beta[8]	-0.03135	0.02368	3.201E-4	-0.07752	-0.03104	0.01437	1001	10000
beta[9]	-0.1005	0.02928	4.689E-4	-0.1585	-0.1003	-0.04306	1001	10000
beta[10]	0.08665	0.01044	1.8E-4	0.06617	0.08655	0.1074	1001	10000
beta[11]	0.1098	0.04729	6.679E-4	0.0175	0.1098	0.2014	1001	10000

همانطور که مشاهده می‌کنید واریانس تخمین [عرض از مبدأ](#) زیاد می‌باشد بنابراین به آن کمتر می‌توان اعتماد کرد.

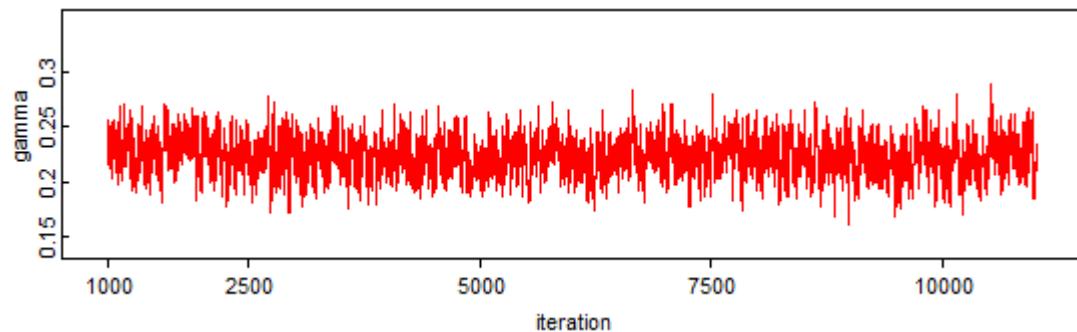
نمودار های تاریخچه β



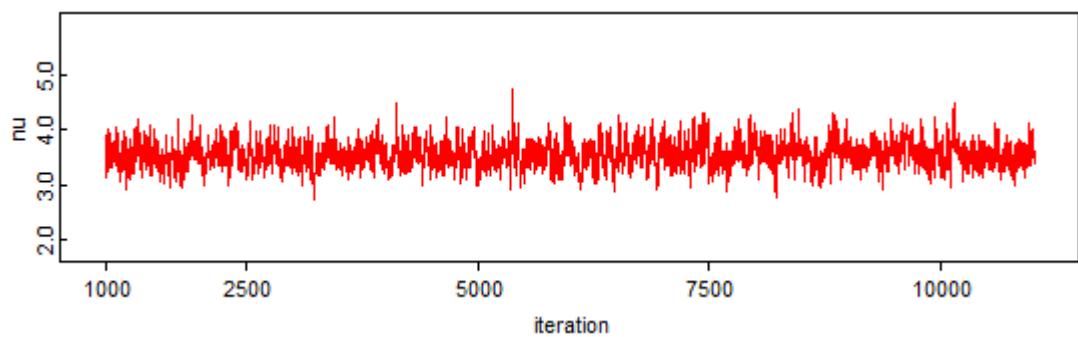
همانطور که مشخص است مدل به خوبی همگرا شده است.

برآورد بیز و نمودار تاریخچه گاما در شکل زیر مشخص است.

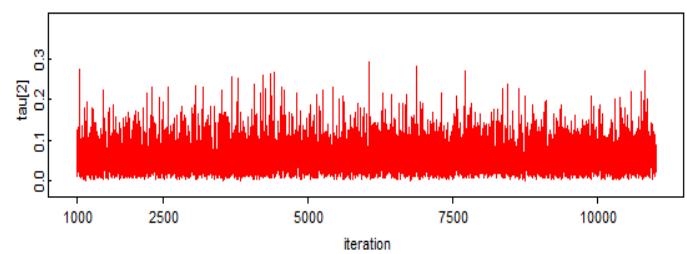
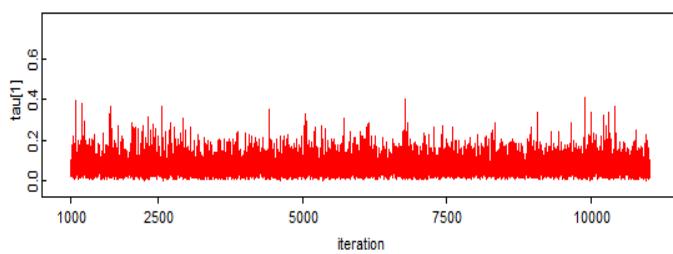
	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
gamma	0.2237	0.01728	6.019E-4	0.1905	0.2237	0.2579	1001	10000



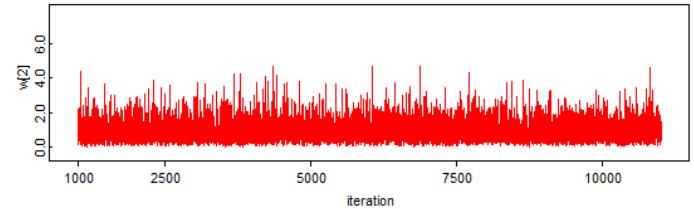
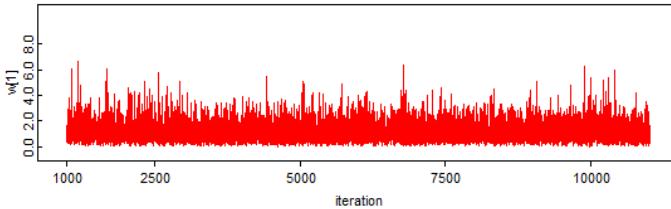
نمودار تاریخچه γ :



نمودار تاریخچه واریانس ها که برای نمونه فقط دو رکورد اول آورده شده است :



نمودار تاریخچه متغیر پنهان w که برای نمونه فقط دو رکورد اول آورده شده است :



همه نمودار های مدل زمان اول مشابه نمودار های مدل انتقال است.

در پایان DIC کل مدل را بدست می آوریم:

	Dbar	Dhat	DIC	pD
HR	18860.0	18100.0	19610.0	757.2
HR1	5935.0	5645.0	6225.0	290.0
gamma	6.446	6.446	6.446	6.954E-13
total	24800.0	23750.0	25850.0	1047.0

همانطور که مشاهده می کنید DIC به ۲۵۸۵۰ رسیده است.

مدل دوازدهم:

مدل چهارم را با استفاده از همان متغیرها و با در نظر گرفتن توزیع SQ_t برای جمله خطأ اجرا می کنیم .

$$HR_{it} = X_i * \beta + \gamma HR_{it-1} + \epsilon_i$$

$$X_i = [MFIL, PO, RD, PS, UE, DV, MA, FP, BLK, FH]$$

$$\beta = [\beta_0, \dots, \beta_{11}]$$

$$\gamma \sim N(0, 0.01)$$

$$\epsilon_i \sim St(0, \sigma^2, \nu, \delta)$$

توزیع پیشین β را نرمال چند متغیره ، σ^2 را معکوس گاما ، ν را نمایی و δ را نرمال در نظر می گیریم.

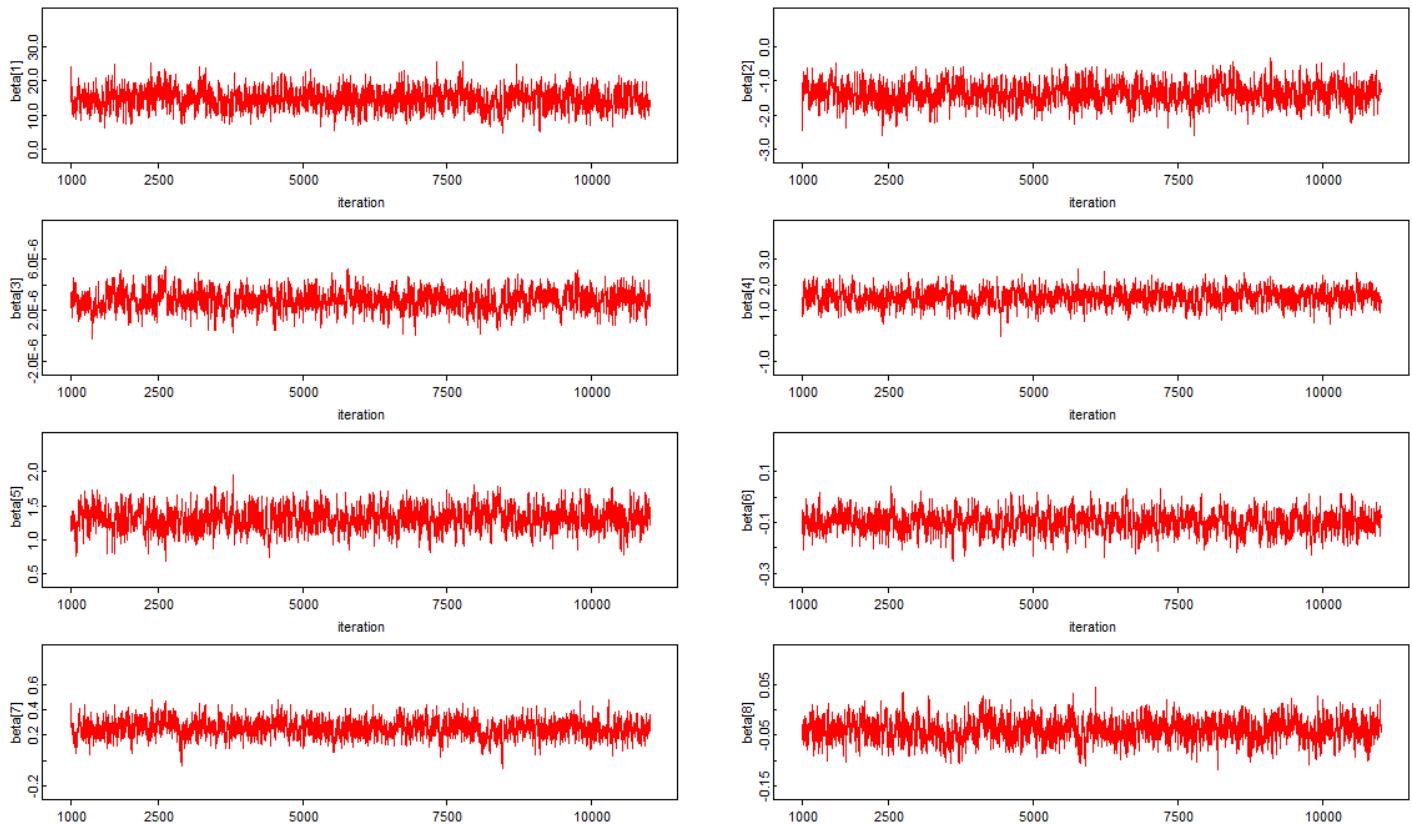
هر چهار توزیع را آگاهی نابخش(واریانس زیاد و میانگین صفر) در نظر گرفتیم.

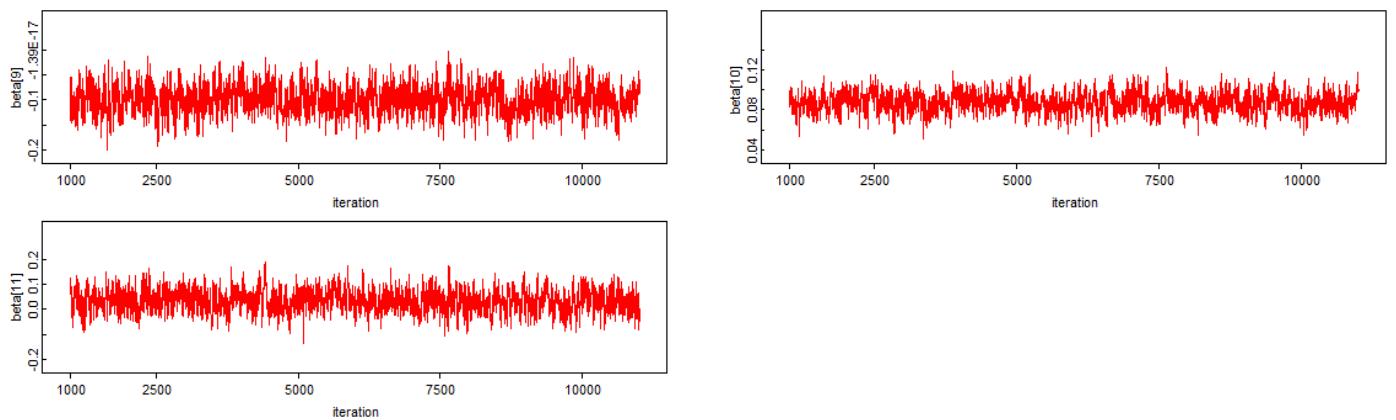
تخمین β ها و واریانس آنها را می‌توان در جدول زیر مشاهده کرد.

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta[1]	14.73	2.846	0.1221	9.126	14.71	20.42	1001	10000
beta[2]	-1.375	0.3006	0.01355	-1.971	-1.378	-0.7865	1001	10000
beta[3]	2.787E-6	7.557E-7	3.317E-8	1.225E-6	2.8E-6	4.246E-6	1001	10000
beta[4]	1.509	0.3057	0.01205	0.9029	1.514	2.097	1001	10000
beta[5]	1.314	0.1621	0.006247	0.9887	1.314	1.633	1001	10000
beta[6]	-0.09891	0.0391	0.001533	-0.1781	-0.09802	-0.02269	1001	10000
beta[7]	0.2538	0.06581	0.002706	0.1211	0.2548	0.3782	1001	10000
beta[8]	-0.04026	0.02109	8.969E-4	-0.08222	-0.04021	6.642E-5	1001	10000
beta[9]	-0.0997	0.02716	0.001077	-0.1519	-0.1001	-0.04652	1001	10000
beta[10]	0.08733	0.009438	4.495E-4	0.06871	0.08712	0.106	1001	10000
beta[11]	0.03338	0.04207	0.001634	-0.04912	0.03344	0.1152	1001	10000

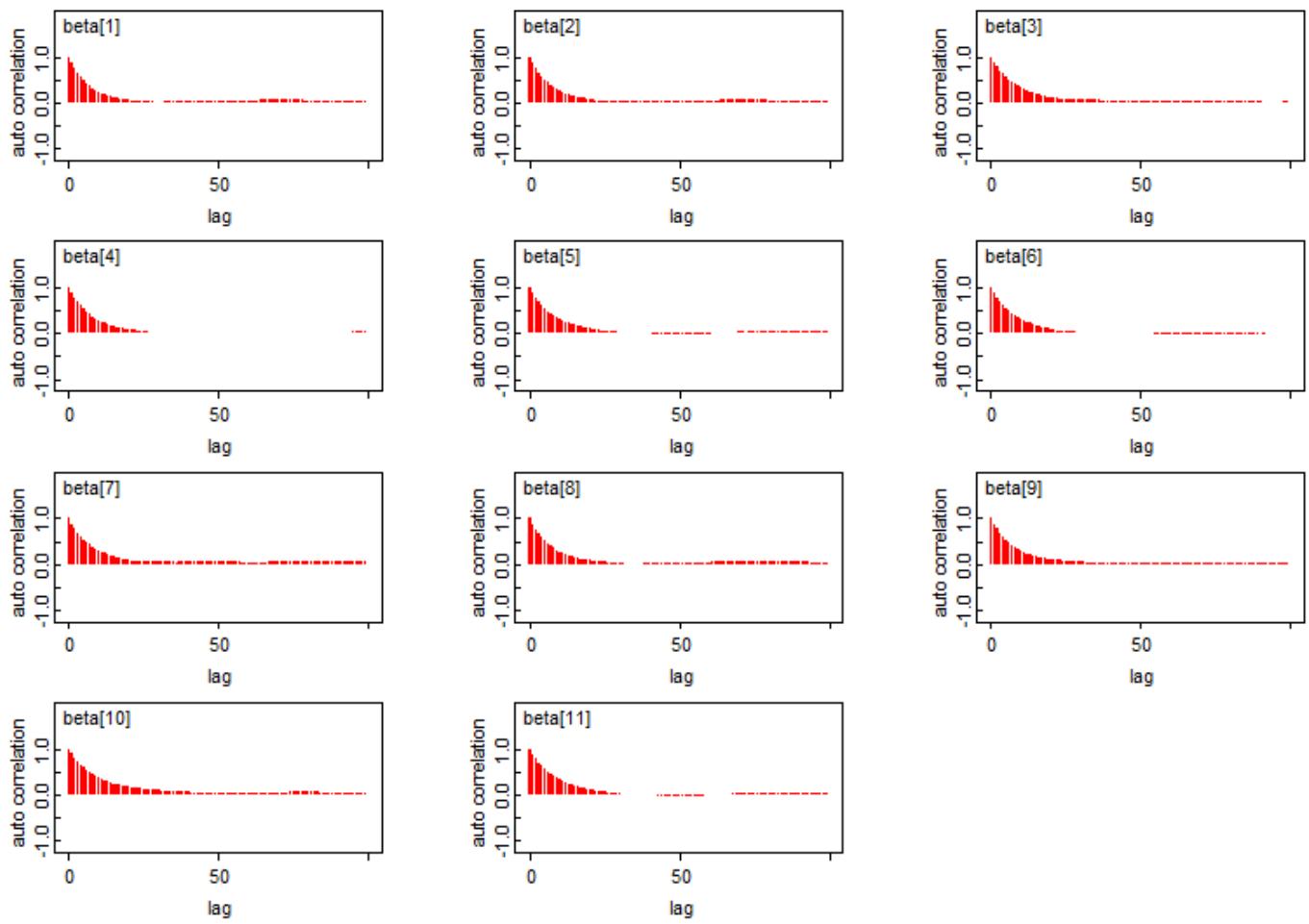
همانطور که مشاهده می‌کنید واریانس تخمین عرض از مبدأ زیاد می‌باشد بنابراین به آن کمتر می‌توان اعتماد کرد.

نمودار های تاریخچه β



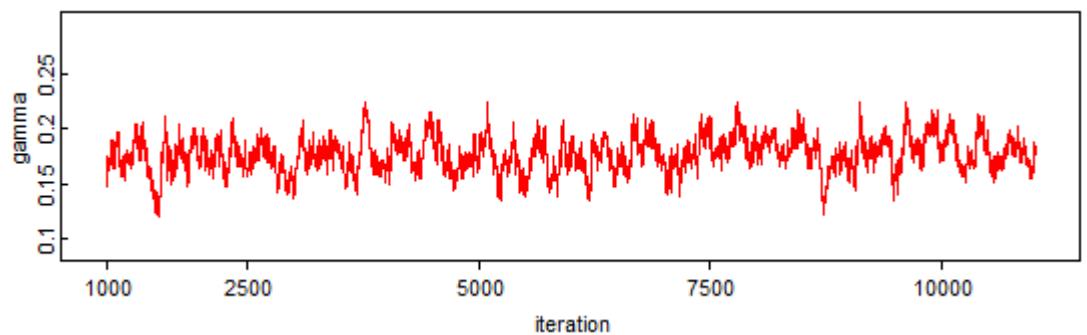


همانطور که در شکل بالا مشاهده می‌کنید همگرایی‌ها نسبت به مدل قبلی بدتر شده ولی هنوز در حد قابل قبولی می‌تواند پیش‌بینی مناسبی برای β ارائه دهد.
باتوجه به شکل زیر نمونه‌ها غیر وابسته تولید شده‌اند.



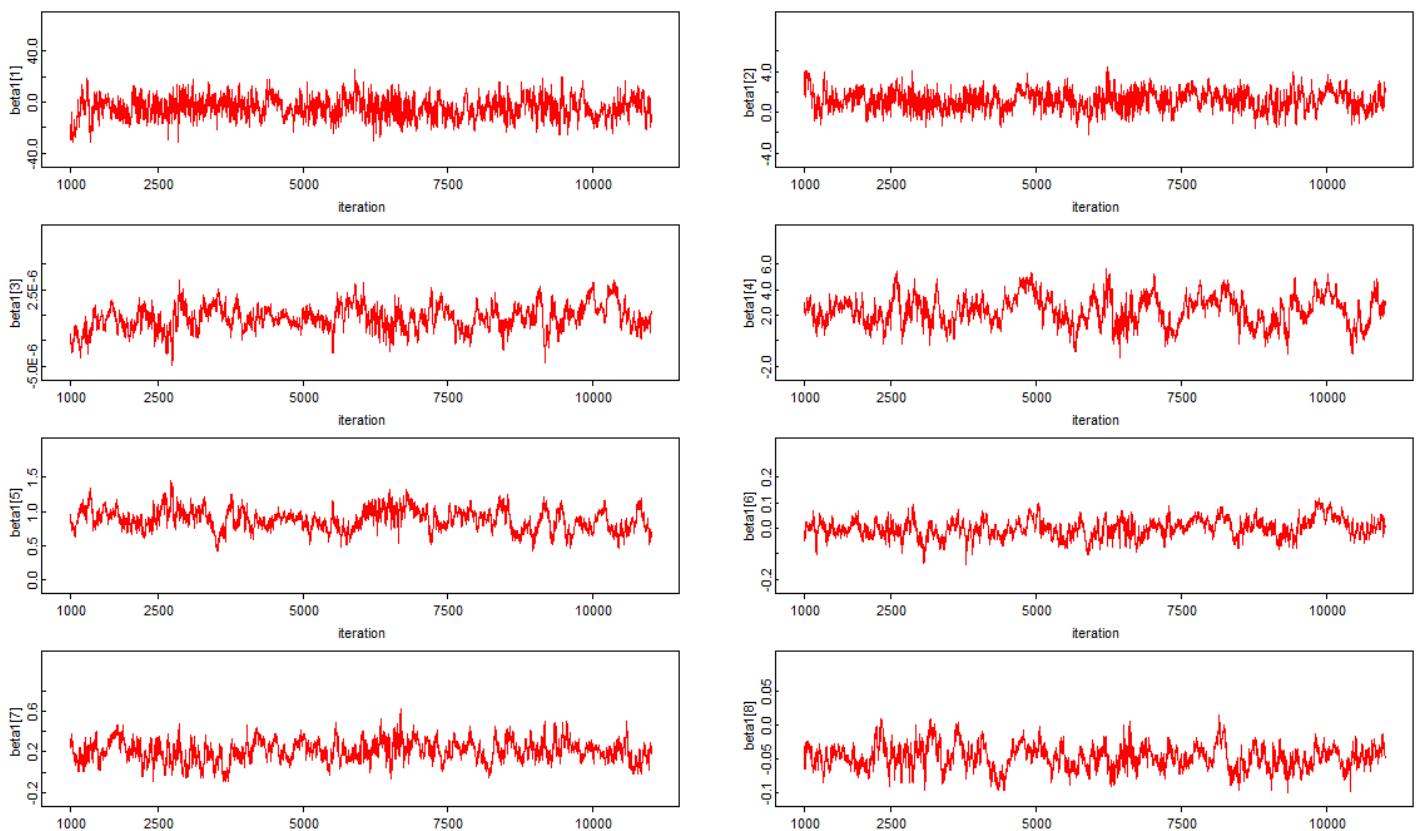
برآورد بیز و نمودار تاریخچه گاما در شکل زیر مشخص است.

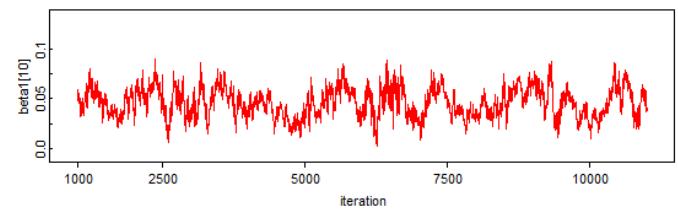
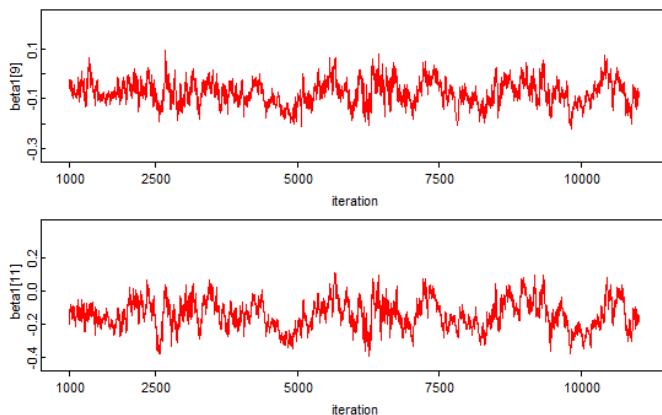
	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
gamma	0.1769	0.01601	0.001164	0.1439	0.1765	0.2079	1001	10000



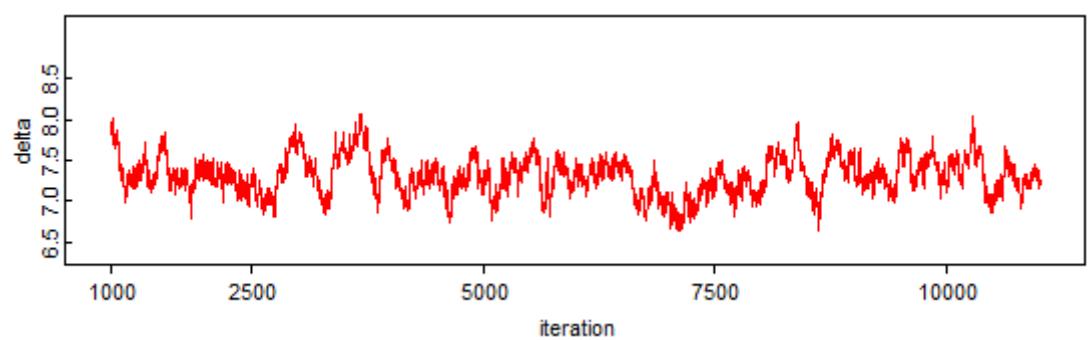
بنابراین گاما به خوبی شبیه سازی نشده است.

نمودار های مدل سال ۱۹۶۰ به صورت زیر است.

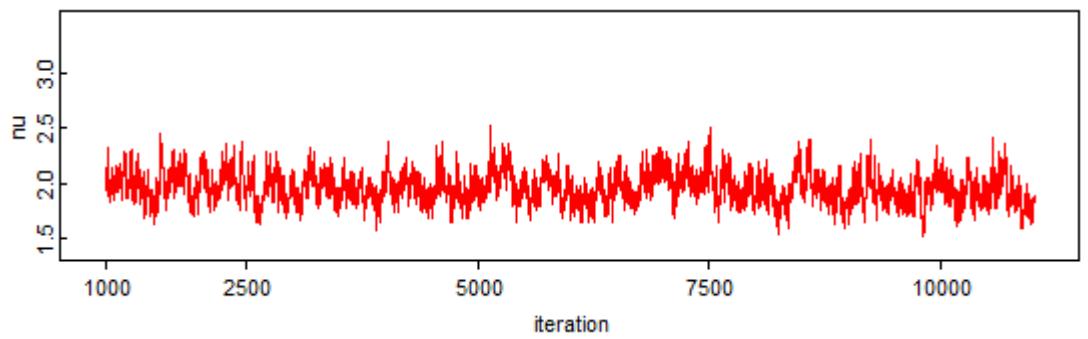




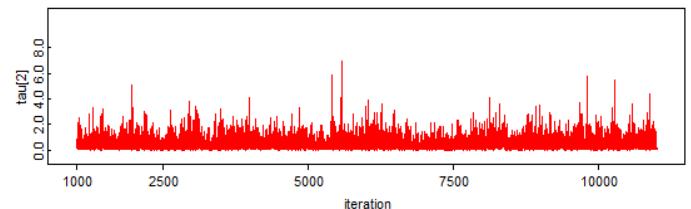
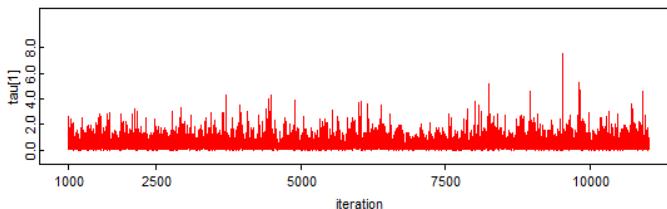
نمودار تاریخچه متغیر δ



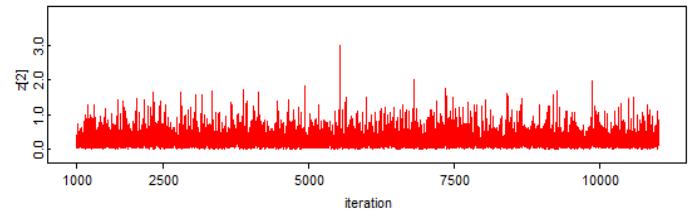
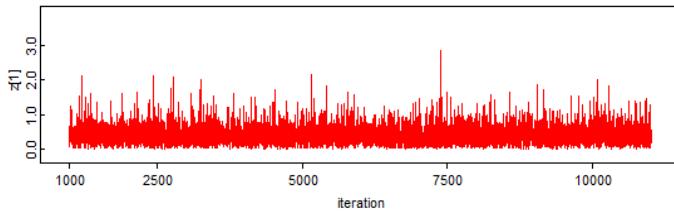
نمودار تاریخچه متغیر پنهان ν



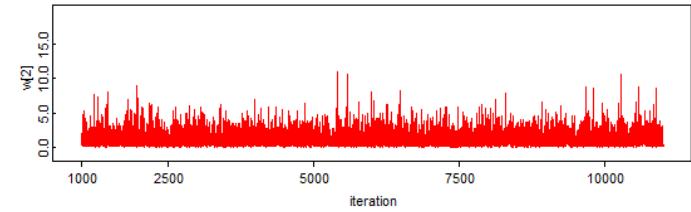
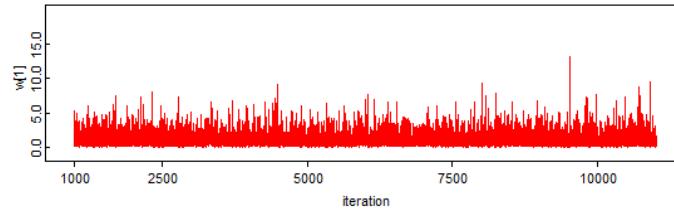
نمودار های واریانس برای دو رکورد اول:



نمودار های متغیر پنهان z



نمودار های متغیر پنهان w



مقدار DIC کل این مدل برابر 180.90 است.

:	Dbar	Dhat	DIC	pD
HR	13610.0	11420.0	15810.0	2193.0
HR1	1692.0	6978.0	-3593.0	-5285.0
gamma	6.446	6.446	6.446	6.954E-13
z	4404.0	4404.0	4404.0	-4.994E-10
z_1	1468.0	1468.0	1468.0	-1.665E-10
total	21180.0	24280.0	18090.0	-3092.0

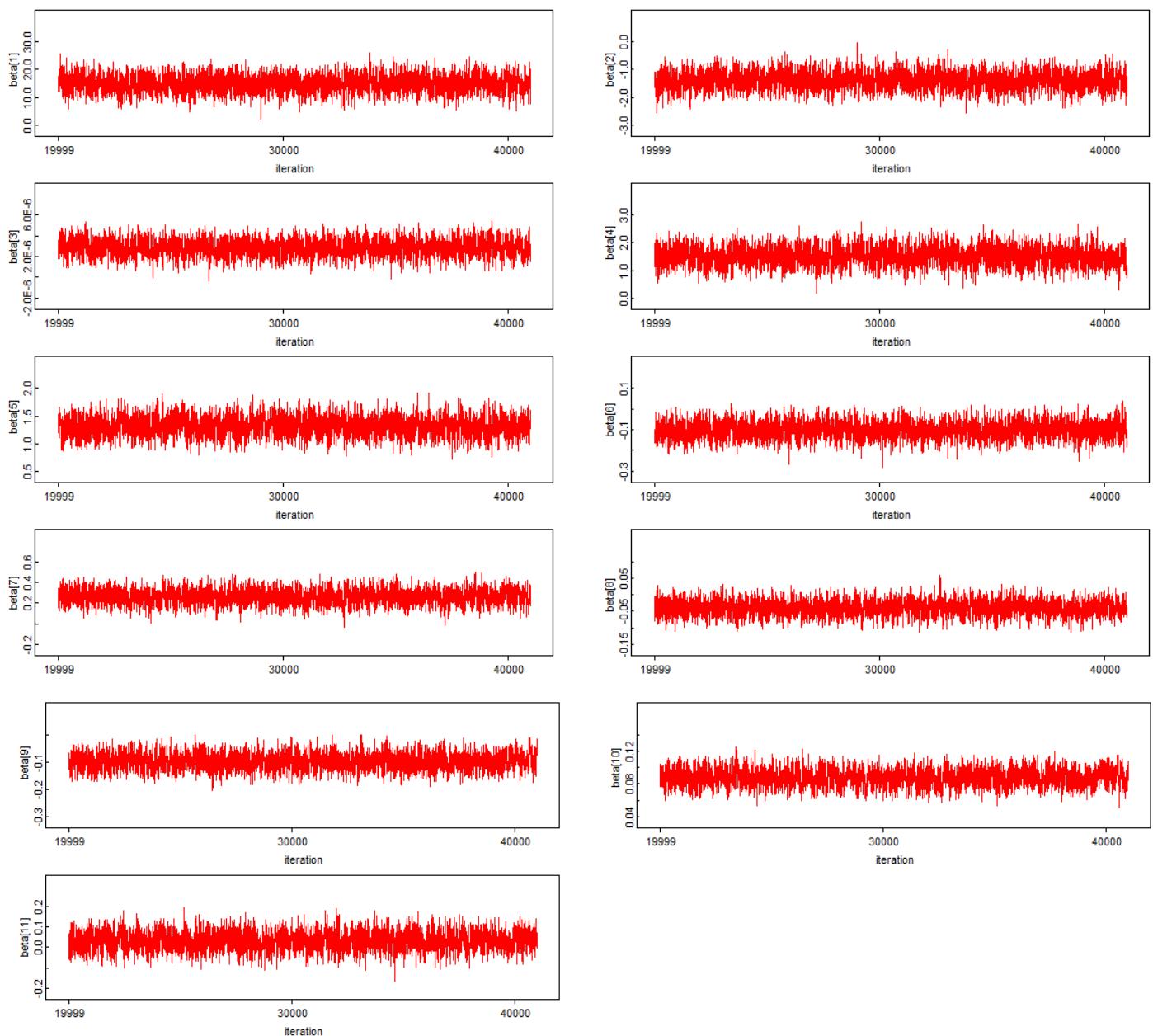
با وجود اینکه DIC کم است ولی نمودارهای تاریخچه همگرایی خوبی نداشته است.

برای بهبود همگرایی ها تولید نمونه را افزایش داده و دورریز انجام می دهیم:

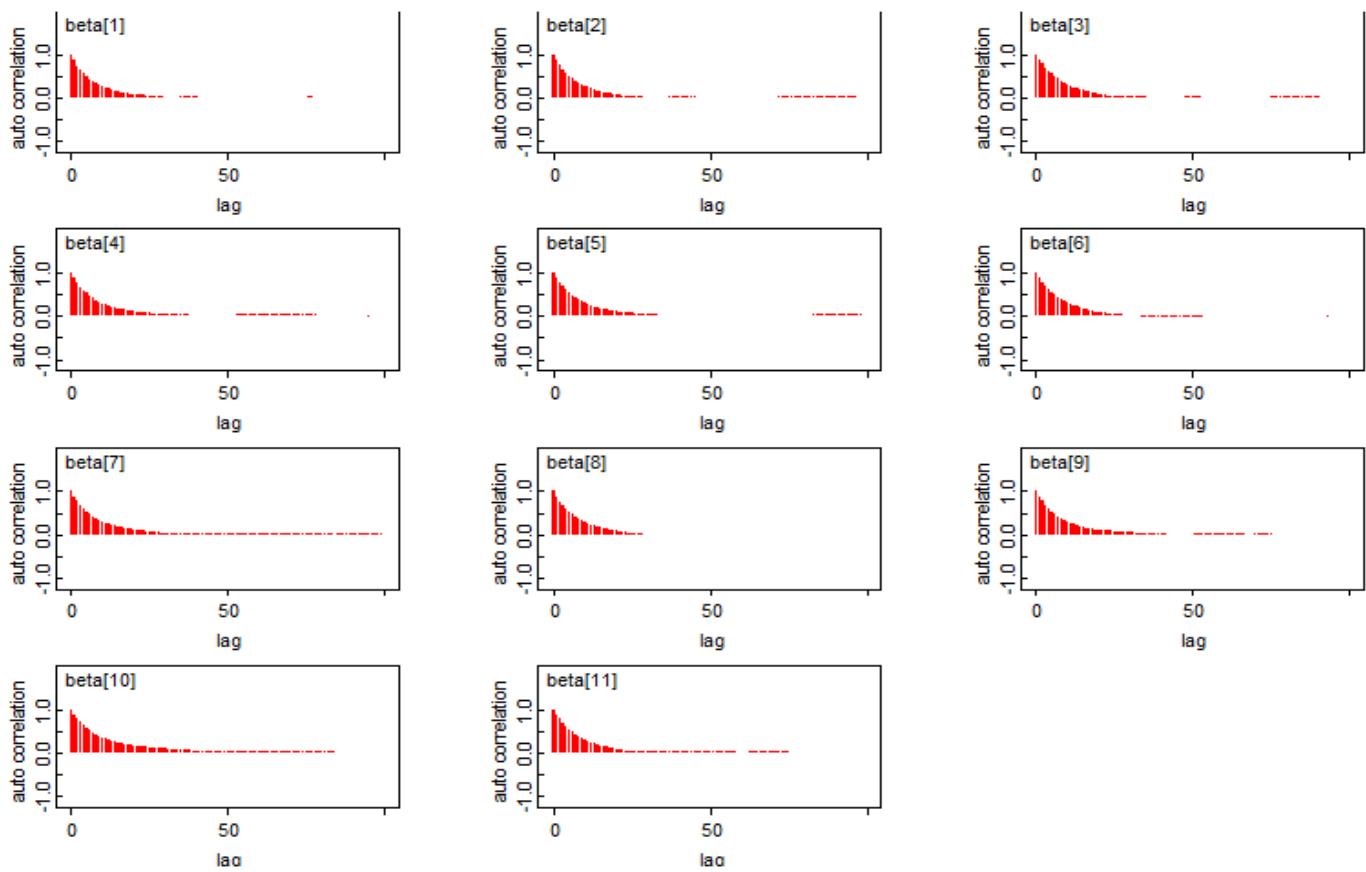
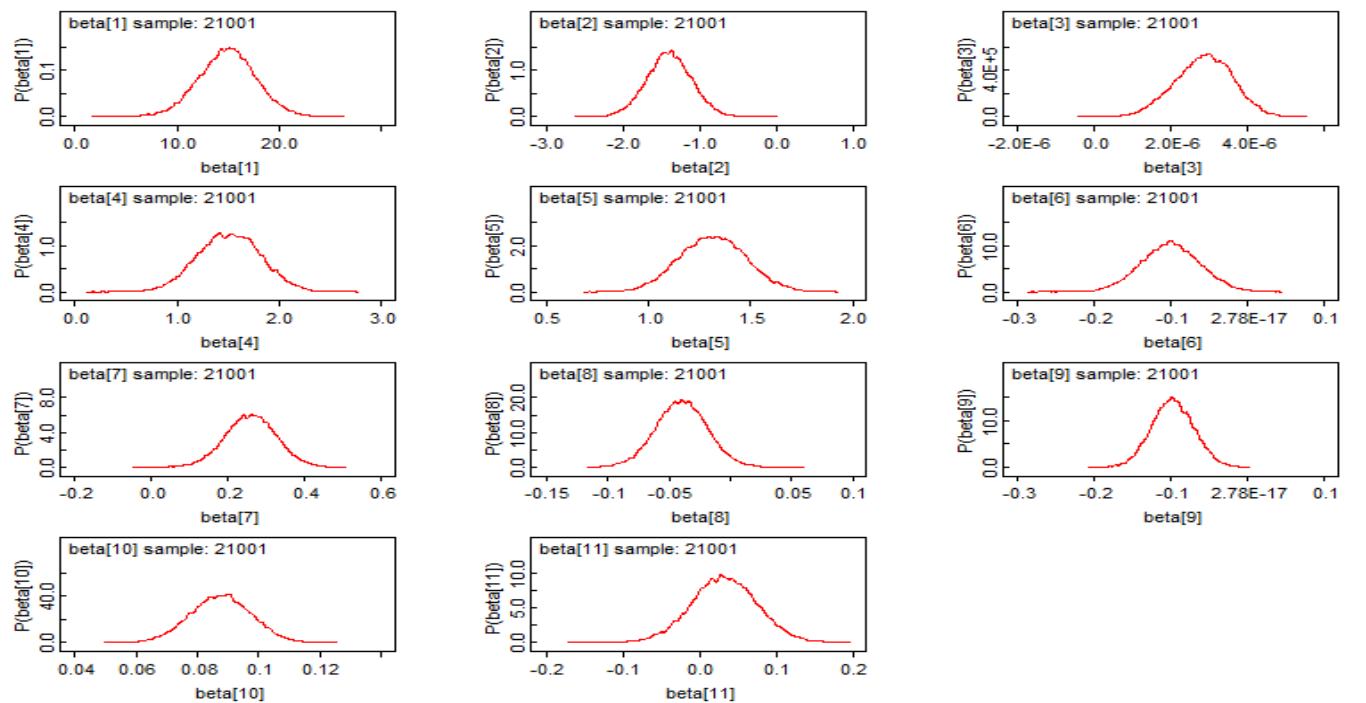
مقدار β بعد از تولید ۴۰۰۰۰ نمونه و دورریز ۲۰۰۰۰ تا از آن ها:

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta[1]	14.94	2.785	0.07977	9.395	14.98	20.4	20000	21001
beta[2]	-1.398	0.2911	0.008206	-1.972	-1.399	-0.8266	20000	21001
beta[3]	2.848E-6	7.462E-7	2.084E-8	1.355E-6	2.874E-6	4.251E-6	20000	21001
beta[4]	1.513	0.3074	0.008927	0.9185	1.514	2.111	20000	21001
beta[5]	1.316	0.1622	0.004782	1.004	1.315	1.639	20000	21001
beta[6]	-0.1011	0.03856	0.001052	-0.1767	-0.1009	-0.02613	20000	21001
beta[7]	0.2596	0.06575	0.001997	0.1296	0.2599	0.3891	20000	21001
beta[8]	-0.0397	0.02089	5.398E-4	-0.08104	-0.0397	0.001429	20000	21001
beta[9]	-0.09797	0.02735	7.616E-4	-0.151	-0.09839	-0.04407	20000	21001
beta[10]	0.08768	0.009644	3.062E-4	0.0688	0.0877	0.1064	20000	21001
beta[11]	0.03218	0.04225	0.001181	-0.05145	0.03206	0.1146	20000	21001

نمودارهای تاریخچه:

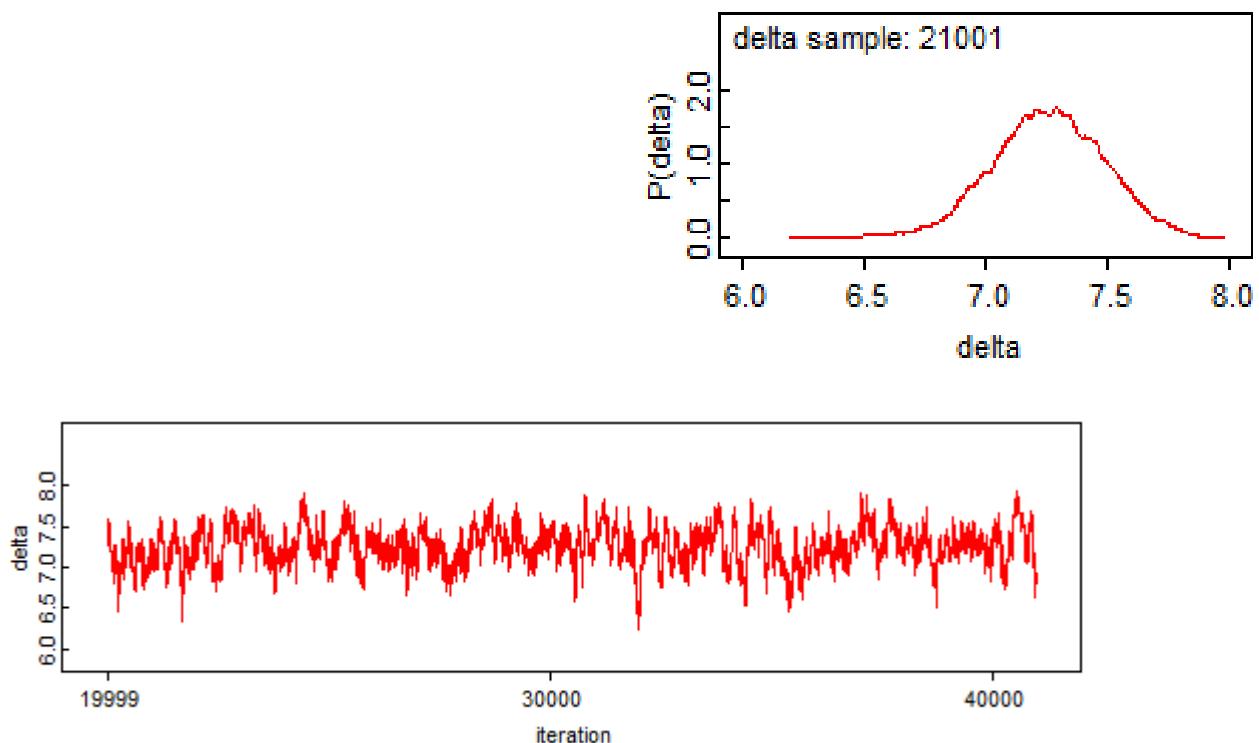


نمودارهای چگالی و خودهمبستگی:

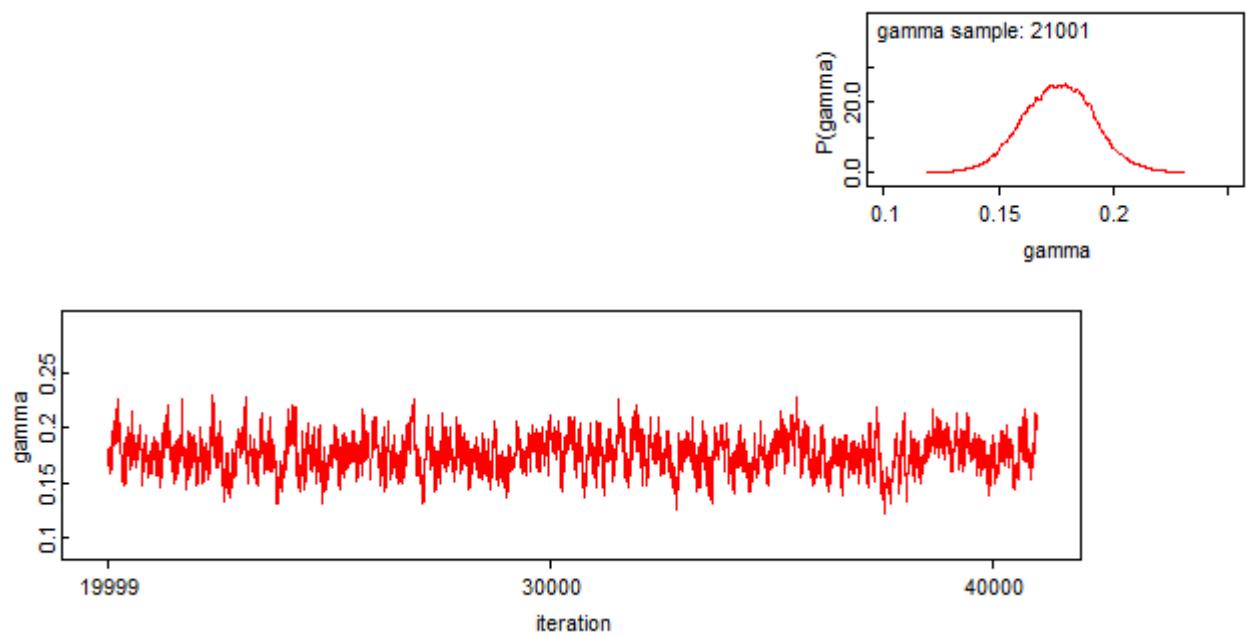


بنابراین نمونه‌ها ناهمبسته تولید شده اند.

نمودار تاریخچه δ

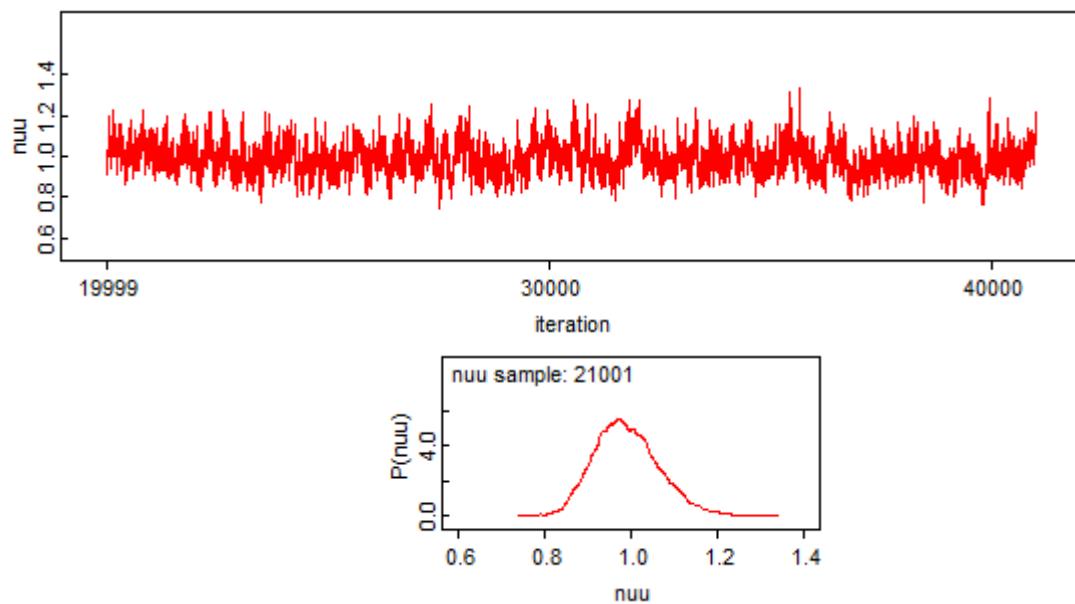


نمودار تاریخچه γ



	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
gamma	0.1763	0.01563	8.399E-4	0.1457	0.1764	0.2074	20000	21001

نمودار تاریخچه ν



بازوچه به شکل زیر DIC کل مدل برابر ۲۰۲۰۰ است.

	Dbar	Dhat	DIC	pD
HR	13620.0	11450.0	15790.0	2169.0
gamma	6.446	6.446	6.446	-6.374E-12
z	4404.0	4404.0	4404.0	-2.178E-9
total	18030.0	15870.0	20200.0	2169.0

مدل های فضایی

Spatial models

مدل سیزدهم :

اولین مدل فضایی را با ۱۱ متغیری که در مدل های قبلی در نظر گرفتیم و توزیع نرمال برای جملات خطابرازش می دهیم.

$$Y_i = X_i * \beta + u [ID(i)]$$

$$X_i = [\text{MFIL}, \text{PO}, \text{RD}, \text{PS}, \text{UE}, \text{DV}, \text{MA}, \text{FP}, \text{BLK}, \text{FH}]$$

$$\beta = [\beta_0, \dots, \beta_{11}]$$

$$\epsilon_i \sim N(0, \sigma^2)$$

توزیع پیشین β را نرمال چند متغیره ، σ^2 را معکوس گاما در نظر می گیریم.

هردو توزیع را آگاهی نابخش (واریانس زیاد و میانگین صفر) در نظر گرفتیم.

برآورد بیز β ها و واریانس آنها را می توان در جدول زیر مشاهده کرد:

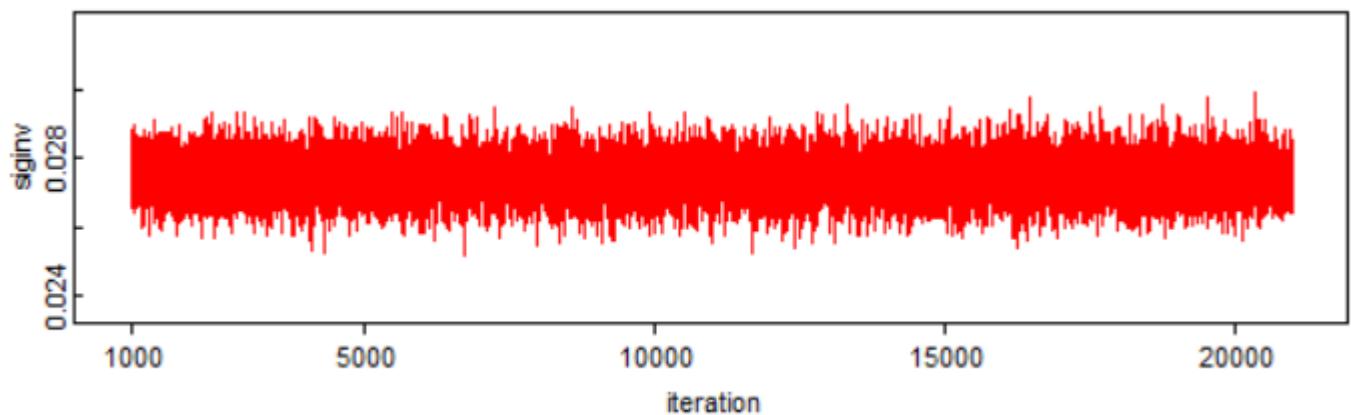
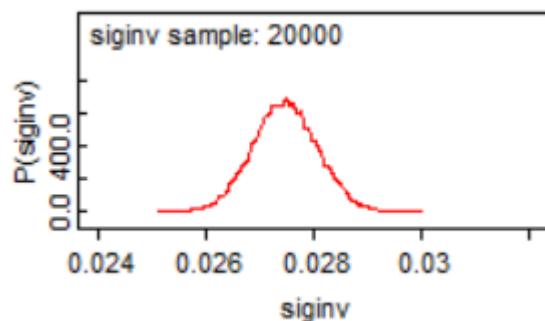
	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta[1]	25.46	3.452	0.02579	18.7	25.47	32.28	1001	20000
beta[2]	-1.599	0.3682	0.002794	-2.319	-1.597	-0.8776	1001	20000
beta[3]	6.873E-6	9.2E-7	7.734E-9	5.089E-6	6.873E-6	8.682E-6	1001	20000
beta[4]	2.463	0.2165	0.001717	2.037	2.462	2.888	1001	20000
beta[5]	-0.1648	0.1852	0.001448	-0.5276	-0.1656	0.1953	1001	20000
beta[6]	-0.04297	0.04593	3.758E-4	-0.132	-0.04328	0.04714	1001	20000
beta[7]	0.2181	0.08852	7.932E-4	0.04374	0.2187	0.3904	1001	20000
beta[8]	-0.09847	0.02423	1.78E-4	-0.1458	-0.09852	-0.05115	1001	20000
beta[9]	-0.1725	0.01566	1.112E-4	-0.203	-0.1726	-0.1415	1001	20000
beta[10]	0.04865	0.01013	8.465E-5	0.02886	0.0487	0.06865	1001	20000
beta[11]	0.1707	0.04357	3.762E-4	0.08585	0.1705	0.2575	1001	20000

همانطور که مشاهده می کنید واریانس تخمین عرض از مبدأ و ضریب گزارش میانه درآمد خانواده (MFIL) و ساختار جمعیت (PS) و نرخ بیکاری (UE) زیاد می باشد ، بنابراین به آن کمتر می توان اعتماد کرد.

با توجه به بررسی های انجام شده بر روی نمودار ها، β های تولید شده توسط این مدل همگرا شده بودند، پس مدل با موفقیت پیاده سازی شده است .

همچنین با توجه به نمودار های چگالی آن ها که بررسی شد توزیع های نرمال به خوبی تولید شده اند.

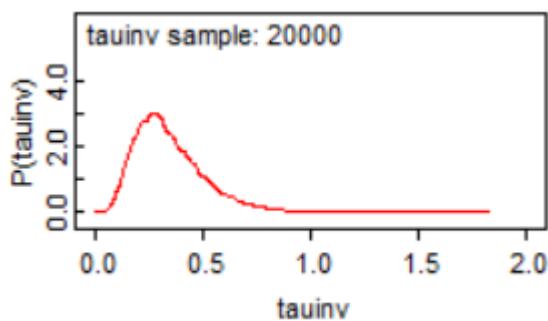
	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
signv	0.02746	6.001E-4	3.958E-6	0.02629	0.02746	0.02862	1001	20000

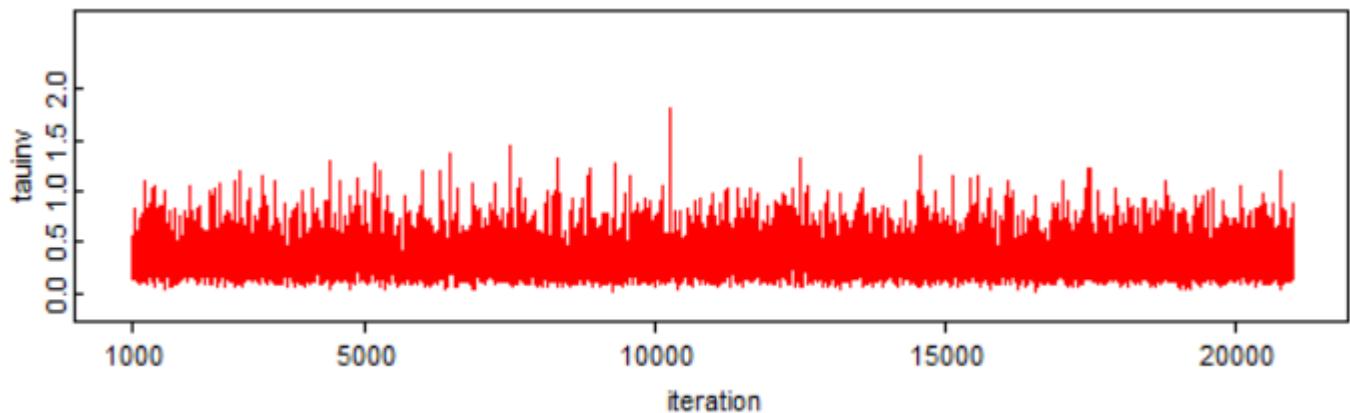


با توجه به تصاویر بالا میانگین واریانس 0.02746 ± 0.00018 شده است و این متغیر به خوبی همگرا شده و توزیع آن نرمال است.

به توجه به نمودار های زیر واریانس توزیع τ ها 0.3408 ± 0.1618 است و توزیع آن چوله به راست می باشد و به خوبی همگرا شده است.

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
tauinv	0.3408	0.1618	0.001801	0.1123	0.3105	0.729	1001	20000





همچنین خود U ها نیز با توجه به بررسی انجام شده به خوبی همگرا شده بودند و نتایج آن ها به صورت زیر می باشد.

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
u[1]	0.8578	0.3939	0.003409	0.07703	0.8652	1.619	1001	20000
u[2]	-0.3608	0.3827	0.003391	-1.124	-0.3597	0.3868	1001	20000
u[3]	0.9189	1.4	0.01435	-1.452	0.7975	4.048	1001	20000
u[4]	-1.241	1.545	0.01258	-4.316	-1.233	1.836	1001	20000
u[5]	2.278	0.4343	0.003494	1.431	2.277	3.125	1001	20000
u[6]	2.343	0.3186	0.002998	1.717	2.345	2.982	1001	20000
u[7]	-0.1871	0.3546	0.003414	-0.879	-0.1869	0.5132	1001	20000
u[8]	-0.5475	0.4125	0.003651	-1.365	-0.5442	0.2581	1001	20000
u[9]	-0.094	0.4051	0.003361	-0.8918	-0.0889	0.7046	1001	20000
u[10]	0.8101	0.3569	0.00252	0.1075	0.8094	1.506	1001	20000
u[11]	-1.116	0.4207	0.00392	-1.961	-1.113	-0.3029	1001	20000
u[12]	-0.02625	0.351	0.002785	-0.7235	-0.02091	0.6491	1001	20000
u[13]	0.415	0.2954	0.00273	-0.17	0.4169	0.9921	1001	20000
u[14]	-0.6093	0.317	0.002283	-1.24	-0.6074	0.01223	1001	20000
u[15]	-2.116	0.4746	0.004424	-3.075	-2.107	-1.215	1001	20000
u[16]	0.1973	0.5447	0.003938	-0.8871	0.2042	1.261	1001	20000
u[17]	-1.523	0.5838	0.005031	-2.704	-1.512	-0.3926	1001	20000

با توجه به جدول زیر DIC مدل ۲۷۲۵۰ می باشد .

	Dbar	Dhat	DIC	pD
HR	27250.0	27260.0	27250.0	-2.852
total	27250.0	27260.0	27250.0	-2.852

مدل چهاردهم :

دومین مدل فضایی را با ۱۱ متغیری که در مدل های قبلی در نظر گرفتیم و توزیع SQ نرمال برای جملات خطاب رازش می دهیم.

$$Y_i = X_i * \beta + u[ID(i)] + \epsilon_i$$

$$X_i = [\text{MFIL}, \text{PO}, \text{RD}, \text{PS}, \text{UE}, \text{DV}, \text{MA}, \text{FP}, \text{BLK}, \text{FH}]$$

$$\beta = [\beta_0, \dots, \beta_{11}]$$

$$\epsilon_i \sim SN(0, \sigma^2, \delta)$$

توزیع پیشین β را نرمال چند متغیره ، σ^2 را معکوس گاما و δ را نرمال در نظر می گیریم.

هر سه توزیع را آگاهی نابخش(واریانس زیاد و میانگین صفر) در نظر گرفتیم.

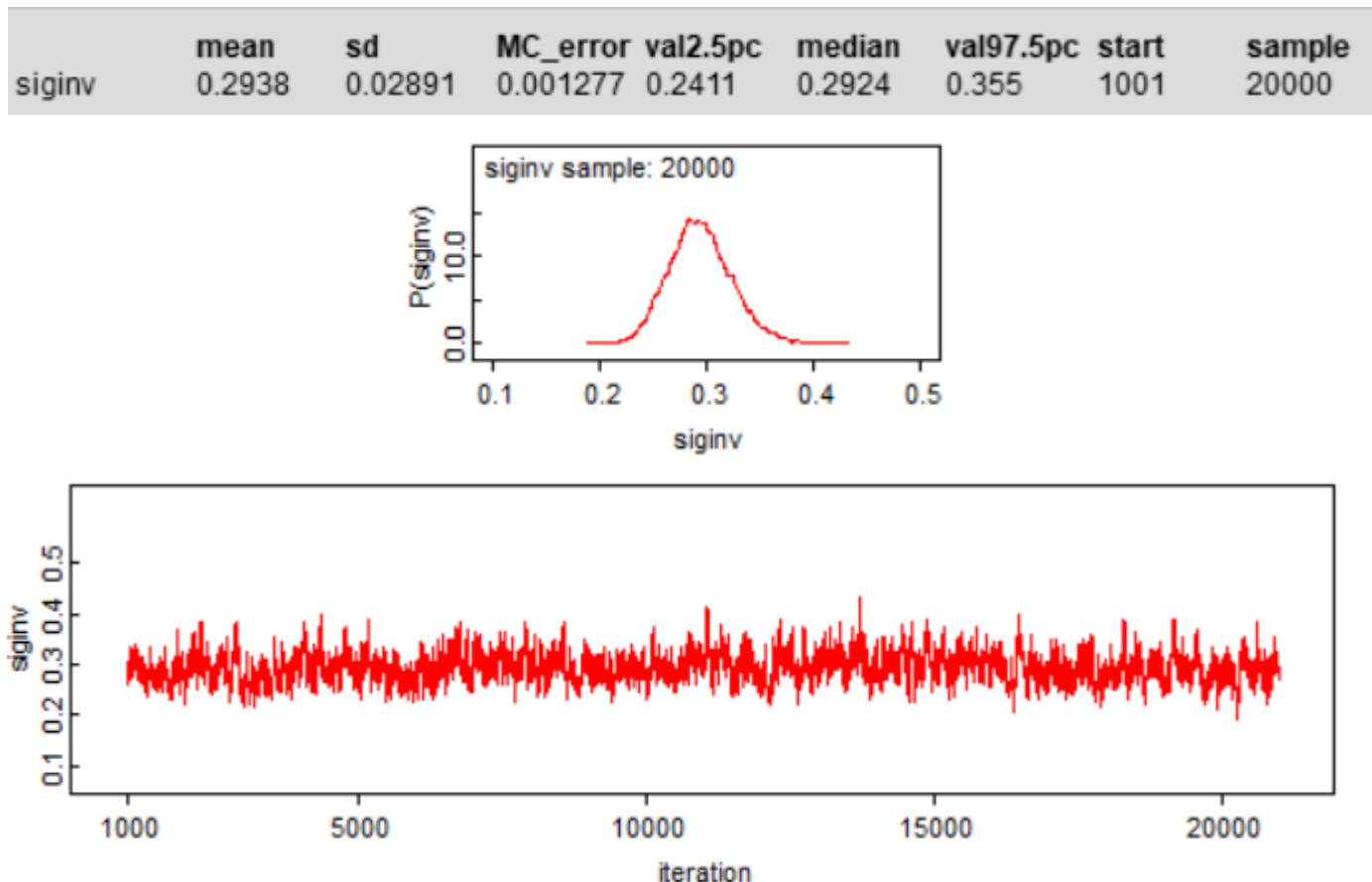
تخمین β ها و واریانس آنها را میتوان در جدول زیر مشاهده کرد :

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta[1]	9.464	2.467	0.05806	4.665	9.465	14.26	1001	20000
beta[2]	-0.6931	0.2603	0.006041	-1.206	-0.6919	-0.1873	1001	20000
beta[3]	2.057E-6	7.926E-7	2.328E-8	4.721E-7	2.079E-6	3.554E-6	1001	20000
beta[4]	1.542	0.1646	0.004486	1.227	1.541	1.864	1001	20000
beta[5]	1.454	0.14	0.003924	1.182	1.455	1.73	1001	20000
beta[6]	-0.01869	0.033	8.09E-4	-0.08351	-0.01846	0.04569	1001	20000
beta[7]	0.1579	0.06289	0.001558	0.03374	0.159	0.2799	1001	20000
beta[8]	-0.05415	0.01642	3.3E-4	-0.08704	-0.05428	-0.02196	1001	20000
beta[9]	-0.1018	0.01162	2.888E-4	-0.1248	-0.1017	-0.07924	1001	20000
beta[10]	0.04811	0.007106	1.632E-4	0.03431	0.048	0.06205	1001	20000
beta[11]	0.04506	0.03036	6.722E-4	-0.01529	0.04518	0.1048	1001	20000

همانطور که مشاهده می کنید واریانس تخمین عرض از مبدأ و ضریب گزارش میانه درآمد خانواده (MFIL) زیاد می باشد، بنابراین به آن کمتر می توان اعتماد کرد.

با توجه به بررسی های انجام شده بر روی نمودار ها، β های تولید شده توسط این مدل همگرا شده بودند، پس مدل با موفقیت پیاده سازی شده است .

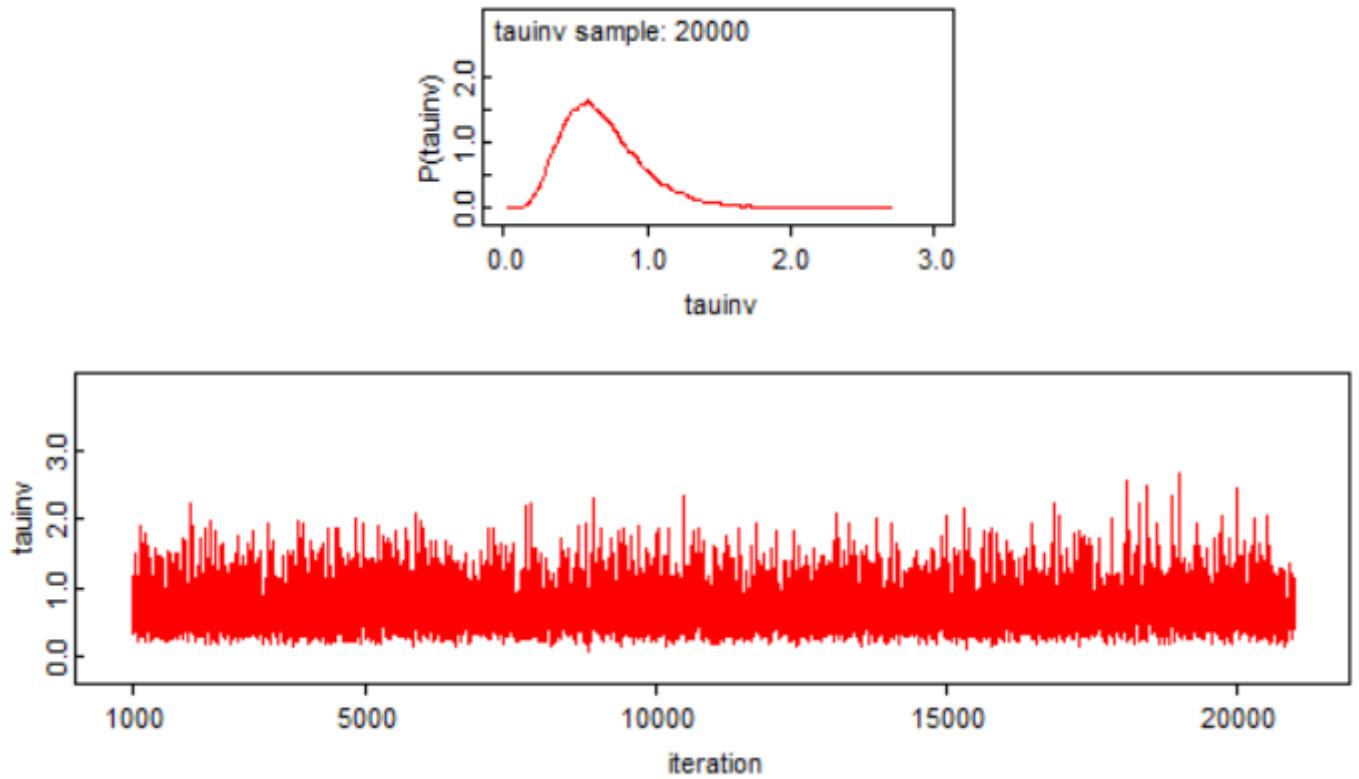
همچنین با توجه به نمودار های چگالی آن ها که بررسی شد، توزیع های نرمال به خوبی تولید شده اند.



با توجه به تصاویر بالا میانگین واریانس ۰/۲۹۳۸ شده است و توزیع آن نرمال است ولی همگرایی آن کمی بد شده است.

به توجه به نمودار های زیر واریانس توزیع τ_{α} ها ۰/۶۹۶۲ است و توزیع آن چوله به راست می باشد و به خوبی همگرا شده است.

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
tauinvs	0.6962	0.2875	0.004201	0.2743	0.6502	1.388	1001	20000



همچنین خود u ها نیز با توجه به بررسی انجام شده به خوبی همگرا شده بودند و نتایج آن ها به صورت زیر می باشد.

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
u[1]	1.15	0.3022	0.006819	0.5533	1.151	1.735	1001	20000
u[2]	0.05185	0.2745	0.00618	-0.4886	0.05253	0.5919	1001	20000
u[3]	-0.9926	0.8243	0.011	-2.583	-1.004	0.6578	1001	20000
u[4]	-1.448	1.206	0.02285	-3.854	-1.449	0.9766	1001	20000
u[5]	1.756	0.3198	0.007118	1.126	1.757	2.379	1001	20000
u[6]	1.429	0.2483	0.00636	0.9432	1.429	1.914	1001	20000
u[7]	-0.3409	0.2506	0.005411	-0.8249	-0.3416	0.1596	1001	20000
u[8]	-0.3026	0.304	0.007262	-0.9074	-0.2992	0.28	1001	20000
u[9]	0.761	0.3013	0.006892	0.1712	0.7573	1.357	1001	20000
u[10]	0.8499	0.2634	0.005771	0.3306	0.8489	1.367	1001	20000
u[11]	-0.3653	0.3072	0.008175	-0.9843	-0.3583	0.2212	1001	20000
u[12]	0.07708	0.2524	0.005622	-0.4203	0.07949	0.568	1001	20000
u[13]	0.7141	0.2198	0.005986	0.2827	0.7135	1.147	1001	20000
u[14]	-0.9578	0.2417	0.005447	-1.439	-0.9546	-0.4943	1001	20000
u[15]	-1.418	0.3183	0.007071	-2.058	-1.41	-0.8088	1001	20000
u[16]	0.9331	0.3972	0.009602	0.1418	0.9348	1.702	1001	20000
u[17]	-1.897	0.4085	0.008337	-2.724	-1.884	-1.129	1001	20000

طبق بررسی های انجام شده بقیه متغیرهای موجود در مدل نیز به خوبی همگرا شده اند.

با توجه به جدول زیر DIC کل مدل ۲۶۶۲۰ می باشد .

	Dbar	Dhat	DIC	pD
HR	17300.0	13860.0	20750.0	3445.0
z	5872.0	5872.0	5872.0	1.505E-11
total	23180.0	19730.0	26620.0	3445.0

مدل پانزدهم :

مدل فضایی بعدی را با ۱۱ متغیری که در مدل های قبلی در نظر گرفتیم و توزیع t برای جملات خطاب برآش می دهیم.

$$Y_i = X_i * \beta + u[ID(i)] + \epsilon_i$$

$$X_i = [\text{MFIL}, \text{PO}, \text{RD}, \text{PS}, \text{UE}, \text{DV}, \text{MA}, \text{FP}, \text{BLK}, \text{FH}]$$

$$\beta = [\beta_0, \dots, \beta_{11}]$$

$$\epsilon_i \sim t(0, \sigma^2, \nu)$$

توزیع پیشین β را نرمال چند متغیره ، σ^2 را معکوس گاما و ν را نمایی در نظر می گیریم.

هر سه توزیع را آگاهی نابخش(واریانس زیاد و میانگین صفر) در نظر گرفتیم.

تخمین β ها و واریانس آنها را میتوان در جدول زیر مشاهده کرد .

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta[1]	21.03	2.766	0.04	15.54	21.08	26.5	1001	10000
beta[2]	-1.388	0.297	0.004174	-1.974	-1.387	-0.8003	1001	10000
beta[3]	3.648E-6	7.545E-7	1.203E-8	2.143E-6	3.655E-6	5.117E-6	1001	10000
beta[4]	2.43	0.172	0.002724	2.095	2.429	2.769	1001	10000
beta[5]	0.9165	0.1507	0.002568	0.6231	0.9166	1.209	1001	10000
beta[6]	-0.01294	0.03759	6.42E-4	-0.08744	-0.01302	0.06088	1001	10000
beta[7]	0.2137	0.07214	0.001184	0.07305	0.2137	0.3549	1001	10000
beta[8]	-0.04485	0.0196	3.05E-4	-0.08376	-0.04481	-0.006825	1001	10000
beta[9]	-0.1588	0.0125	1.659E-4	-0.1835	-0.1588	-0.1342	1001	10000
beta[10]	0.07121	0.008657	1.318E-4	0.05442	0.07133	0.08804	1001	10000
beta[11]	0.101	0.03673	6.4E-4	0.02955	0.1008	0.1738	1001	10000

با توجه به بررسی های انجام شده تمامی متغیر های استفاده شده به خوبی همگرا شده بودند.

با توجه به جدول زیر DIC مدل 25900 می باشد .

	Dbar	Dhat	DIC	pD
HR	24850.0	23800.0	25900.0	1053.0
total	24850.0	23800.0	25900.0	1053.0

مدل شانزدهم :

آخرین مدل فضایی را با ۱۱ متغیری که در مدل های قبلی در نظر گرفتیم و توزیع SQ_t برای جملات خطاب را برآورد می دهیم.

$$Y_i = X_i * \beta + u[ID(i)] + \epsilon_i$$

$$X_i = [\text{MFIL}, \text{PO}, \text{RD}, \text{PS}, \text{UE}, \text{DV}, \text{MA}, \text{FP}, \text{BLK}, \text{FH}]$$

$$\beta = [\beta_0, \dots, \beta_{11}]$$

$$\epsilon_i \sim St(0, \sigma^2, \nu, \delta)$$

توزیع پیشین β را نرمال چند متغیره ، σ^2 را معکوس گاما ، ν را نمایی و δ را نرمال در نظر می گیریم.

هر چهار توزیع را آگاهی نابخش(واریانس زیاد و میانگین صفر) در نظر گرفتیم.

تخمین β ها و واریانس آنها را میتوان در جدول زیر مشاهده کرد.

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta[1]	10.39	2.452	0.1407	5.686	10.41	15.18	1001	10000
beta[2]	-0.8357	0.2558	0.01442	-1.338	-0.8327	-0.344	1001	10000
beta[3]	2.012E-6	7.388E-7	4.125E-8	4.95E-7	2.015E-6	3.455E-6	1001	10000
beta[4]	1.623	0.1703	0.009748	1.281	1.622	1.948	1001	10000
beta[5]	1.3	0.1269	0.006153	1.052	1.3	1.546	1001	10000
beta[6]	0.009723	0.03546	0.002145	-0.05786	0.009786	0.08023	1001	10000
beta[7]	0.13	0.06293	0.003371	0.006254	0.13	0.252	1001	10000
beta[8]	-0.0368	0.01627	6.36E-4	-0.06994	-0.03677	-0.005581	1001	10000
beta[9]	-0.1165	0.01113	6.201E-4	-0.1383	-0.1165	-0.0948	1001	10000
beta[10]	0.06666	0.007607	4.108E-4	0.05175	0.06674	0.08155	1001	10000
beta[11]	0.08894	0.03228	0.001743	0.02651	0.08909	0.15	1001	10000

با توجه به بررسی های انجام شده برخی متغیر های به کار رفته در مدل همگرایی خوب و قابل قبولی داشته و برخی همگرایی خوبی نداشته اند.

با توجه به جدول زیر DIC مدل ۲۴۸۰۰ می باشد.

I	Dbar	Dhat	DIC	pD
HR	16540.0	14140.0	18930.0	2391.0
z	5872.0	5872.0	5872.0	-5.907E-10
total	22410.0	20020.0	24800.0	2391.0

مدل های ترکیبی

combinatorial
models

مدل هدفم : ترکیب مدل زمانی و مکانی

اولین مدل با ۱۱ متغیری که در مدل های قبلی در نظر گرفتیم و توزیع نرمال برای جملات خطابرازش می دهیم.

$$HR_{it} = X_i * \beta + u[ID(i)] + HR_{it-1} + \epsilon_i$$

$$X_i = [\text{MFIL}, \text{PO}, \text{RD}, \text{PS}, \text{UE}, \text{DV}, \text{MA}, \text{FP}, \text{BLK}, \text{FH}]$$

$$\beta = [\beta_0, \dots, \beta_{11}]$$

$$\epsilon_i \sim N(0, \sigma^2)$$

توزیع پیشین β را نرمال چند متغیره ، σ^2 را معکوس گاما در نظر می گیریم.

هردو توزیع را آگاهی نابخش (واریانس زیاد و میانگین صفر) در نظر گرفتیم.

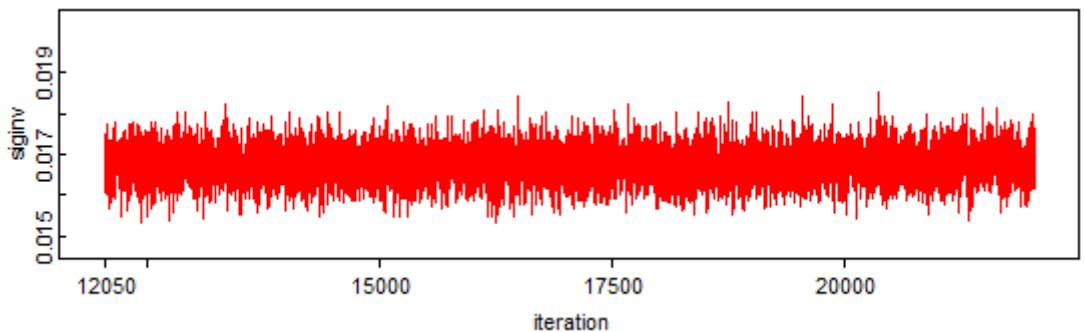
تخمین β ها و واریانس آنها را میتوان در جدول زیر مشاهده کرد :

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta[1]	18.77	4.659	0.03052	9.661	18.81	27.92	1051	21000
beta[2]	-2.487	0.4855	0.003376	-3.432	-2.49	-1.533	1051	21000
beta[3]	1.751E-6	1.239E-6	9.6E-9	-6.726E-7	1.747E-6	4.183E-6	1051	21000
beta[4]	-1.836	0.5377	0.004283	-2.895	-1.834	-0.7859	1051	21000
beta[5]	0.2554	0.274	0.001942	-0.2829	0.2564	0.7898	1051	21000
beta[6]	-0.05156	0.06549	4.837E-4	-0.1783	-0.05202	0.07668	1051	21000
beta[7]	-0.1266	0.114	8.683E-4	-0.3521	-0.1257	0.09644	1051	21000
beta[8]	0.05055	0.03696	2.511E-4	-0.02236	0.05061	0.1232	1051	21000
beta[9]	0.1143	0.04569	3.186E-4	0.02496	0.1139	0.2047	1051	21000
beta[10]	0.02754	0.01572	1.196E-4	-0.003208	0.02754	0.05849	1051	21000
beta[11]	0.2631	0.07092	5.608E-4	0.124	0.263	0.4019	1051	21000

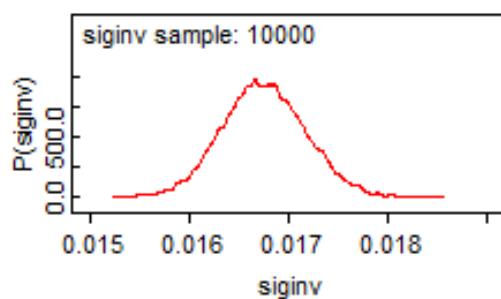
همانطور که مشاهده می کنید واریانس تخمین عرض از مبدا ، ضریب گزارش میانه درآمد خانواده (MFIL) ، ضریب ساختار جمعیت (PS) و ضریب محرومیت از منابع (RD) زیاد می باشد، بنابراین به آن کمتر می توان اعتماد کرد.

با توجه به بررسی های انجام شده بر روی نمودار ها، β های تولید شده توسط این مدل همگرا شده بودند، پس مدل با موفقیت پیاده سازی شده است .

نمودار تاریخچه واریانس:

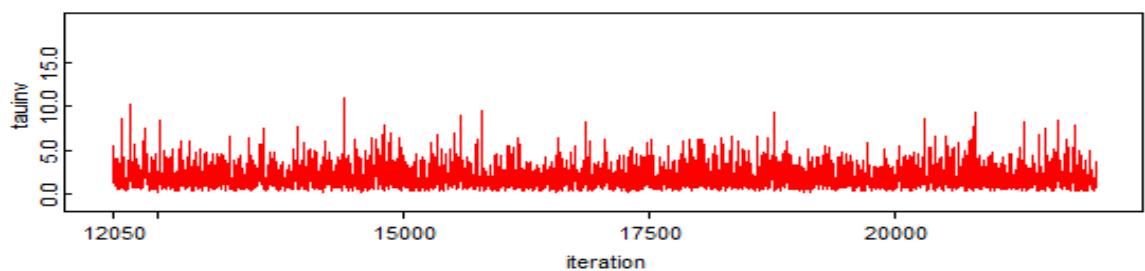


	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
siginv	0.01675	4.237E-4	4.349E-6	0.01593	0.01674	0.01759	12051	10000

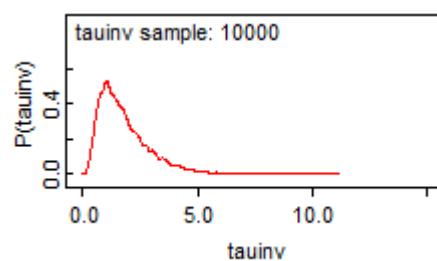


با توجه به تصاویر بالا میانگین واریانس 1657 ± 0.01 شده است و نمودار تاریخچه هم همگرا است.

به توجه به نمودار های زیر واریانس توزیع τ ها 1795 ± 1 است و توزیع آن چوله به راست می باشد و به خوبی همگرا شده است.



	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
tauinvs	1.795	1.104	0.02241	0.447	1.534	4.512	12051	10000



همچنین خود u ها نیز با توجه به بررسی انجام شده به خوبی همگرا شده بودند و نتایج آن ها به صورت زیر می باشد.

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
u[1]	-0.2296	0.3367	0.003079	-0.9305	-0.2197	0.4054	1051	21000
u[2]	0.3818	0.3495	0.003385	-0.2749	0.3685	1.112	1051	21000
u[3]	-0.2642	0.4988	0.00455	-1.285	-0.2467	0.6848	1051	21000
u[4]	-0.3108	0.4551	0.004043	-1.216	-0.305	0.5709	1051	21000
u[5]	-0.4365	0.3629	0.003023	-1.169	-0.4283	0.2529	1051	21000
u[6]	0.2618	0.3501	0.004065	-0.3742	0.2387	1.012	1051	21000
u[7]	-0.3584	0.3585	0.003589	-1.105	-0.3461	0.3154	1051	21000
u[8]	0.6279	0.3281	0.003536	0.005443	0.6187	1.292	1051	21000
u[9]	-0.1159	0.3679	0.003626	-0.8906	-0.09694	0.5627	1051	21000
u[10]	0.2517	0.3682	0.003547	-0.4318	0.2366	1.025	1051	21000
u[11]	0.04444	0.4113	0.003512	-0.7864	0.05166	0.8387	1051	21000
u[12]	0.08129	0.2965	0.002393	-0.4848	0.07492	0.6962	1051	21000
u[13]	0.1012	0.4183	0.003598	-0.7556	0.1125	0.9051	1051	21000
u[14]	-0.06662	0.3047	0.002358	-0.6655	-0.0703	0.5445	1051	21000
u[15]	-0.2426	0.4788	0.004907	-1.213	-0.2315	0.6877	1051	21000
u[16]	0.4989	0.6919	0.007521	-0.7154	0.439	2.051	1051	21000
u[17]	-0.2245	0.4765	0.004661	-1.221	-0.2115	0.7107	1051	21000

باتوجه به شکل زیر DIC کل مدل برابر ۲۲۰۲۰ است.

	Dbar	Dhat	DIC	pD
HR	22010.0	2.2E+4	22020.0	11.76
total	22010.0	2.2E+4	22020.0	11.76

اگر همین مدل را با یک ضریب برای زمان قبل انجام دهیم همه همگرایی‌ها خوب باقی می‌ماند ولی DIC به ۲۰۳۴۰ می‌رسد.

	Dbar	Dhat	DIC	pD
HR	20320.0	20310.0	20340.0	17.45
total	20320.0	20310.0	20340.0	17.45

تخمین ضریب و u و β

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
gamma	0.1874	0.01739	3.765E-4	0.1534	0.1875	0.2212	1001	10000

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
u[1]	-0.2105	0.3712	0.005636	-0.9659	-0.2008	0.4961	1001	10000
u[2]	-0.02271	0.3927	0.005155	-0.7965	-0.01105	0.7307	1001	10000
u[3]	-0.2702	0.5308	0.005666	-1.311	-0.2706	0.7767	1001	10000
u[4]	0.4031	0.4007	0.004575	-0.3806	0.4037	1.193	1001	10000
u[5]	1.624	0.3479	0.004901	0.961	1.619	2.311	1001	10000
u[6]	-0.00319	0.3725	0.004086	-0.7701	0.001745	0.715	1001	10000
u[7]	-0.984	0.3476	0.003705	-1.673	-0.9781	-0.3062	1001	10000
u[8]	1.003	0.2996	0.004199	0.4061	1.005	1.586	1001	10000
u[9]	0.1559	0.3897	0.003973	-0.6218	0.1556	0.916	1001	10000
u[10]	-0.119	0.3966	0.004455	-0.8799	-0.1186	0.6575	1001	10000
u[11]	-0.1047	0.3961	0.00542	-0.8915	-0.1013	0.6586	1001	10000
u[12]	-1.032	0.4661	0.008689	-2.026	-1.007	-0.2011	1001	10000
u[13]	1.182	0.4489	0.005501	0.318	1.174	2.094	1001	10000
u[14]	-0.9394	0.3678	0.004856	-1.689	-0.9339	-0.2479	1001	10000
u[15]	-1.403	0.6366	0.00904	-2.762	-1.377	-0.2139	1001	10000
u[16]	1.297	1.243	0.02412	-0.6853	1.131	4.207	1001	10000
u[17]	-0.5763	0.752	0.009151	-2.072	-0.5592	0.8683	1001	10000

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta[1]	24.69	3.797	0.04315	17.21	24.72	32.23	1001	10000
beta[2]	-2.144	0.397	0.004565	-2.933	-2.148	-1.371	1001	10000
beta[3]	5.441E-6	9.512E-7	9.819E-9	3.597E-6	5.446E-6	7.327E-6	1001	10000
beta[4]	0.6265	0.4217	0.004223	-0.2005	0.6268	1.447	1001	10000
beta[5]	0.2084	0.2151	0.002553	-0.2068	0.2101	0.6302	1001	10000
beta[6]	-0.04643	0.0538	7.541E-4	-0.151	-0.0462	0.05879	1001	10000
beta[7]	0.1906	0.09268	0.00118	0.008203	0.1905	0.3707	1001	10000
beta[8]	-0.04425	0.02895	3.364E-4	-0.1009	-0.04439	0.0132	1001	10000
beta[9]	-0.0421	0.03543	3.99E-4	-0.1109	-0.04247	0.0278	1001	10000
beta[10]	0.07152	0.01249	1.375E-4	0.04721	0.0714	0.09623	1001	10000
beta[11]	0.263	0.05542	5.865E-4	0.1542	0.2632	0.371	1001	10000

مدل هجدهم : ترکیب مدل زمانی و پانلی

در این مدل اثرات تصادفی هر پانل و لگ متغیر هدف را باهم در نظر می‌گیریم.

$$HR_{it} = X_i * \beta + a[ID(i)] + \gamma * HR_{it-1} + \epsilon_i$$

$$X_i = [\text{MFIL}, \text{PO}, \text{RD}, \text{PS}, \text{UE}, \text{DV}, \text{MA}, \text{FP}, \text{BLK}, \text{FH}]$$

$$\beta = [\beta_0, \dots, \beta_{11}]$$

$$\epsilon_i \sim N(0, \sigma^2)$$

توزیع پیشین β را نرمال چند متغیره ، σ^2 را معکوس گاما در نظر می‌گیریم.

هردو توزیع را آگاهی نابخش(واریانس زیاد و میانگین صفر) در نظر گرفتیم.

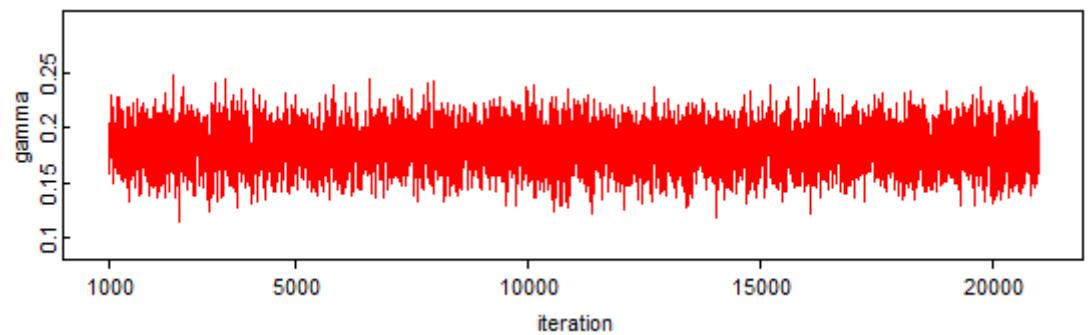
تخمین β ها و واریانس آنها را میتوان در جدول زیر مشاهده کرد :

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta[1]	24.64	3.93	0.05761	16.92	24.64	32.34	1001	20000
beta[2]	-2.158	0.4038	0.003356	-2.952	-2.156	-1.364	1001	20000
beta[3]	5.281E-6	9.665E-7	7.511E-9	3.383E-6	5.279E-6	7.164E-6	1001	20000
beta[4]	0.5357	0.422	0.003399	-0.2931	0.536	1.362	1001	20000
beta[5]	0.2307	0.2167	0.001868	-0.1902	0.2301	0.659	1001	20000
beta[6]	-0.00884	0.05329	4.581E-4	-0.1119	-0.008936	0.09548	1001	20000
beta[7]	0.1888	0.09422	9.324E-4	0.003315	0.1885	0.3724	1001	20000
beta[8]	-0.04414	0.02877	2.238E-4	-0.1008	-0.04415	0.01208	1001	20000
beta[9]	-0.03278	0.03553	2.905E-4	-0.1019	-0.03265	0.03744	1001	20000
beta[10]	0.06698	0.01289	1.204E-4	0.0416	0.06691	0.0923	1001	20000
beta[11]	0.2619	0.05632	4.657E-4	0.1511	0.2622	0.372	1001	20000

همانطور که مشاهده می کنید واریانس تخمین عرض از مبدا ، ضریب گزارش میانه درآمد خانواده (MFIL) ، ضریب ساختار جمعیت (PS) و ضریب محرومیت از منابع (RD) زیاد می باشد، بنابراین به آن کمتر می توان اعتماد کرد.

با توجه به بررسی های انجام شده بر روی نمودار ها، β های تولید شده توسط این مدل همگرا شده بودند، پس مدل با موفقیت پیاده سازی شده است .

نمودار تاریخچه و برآورد گاما در شکل صفحه بعد مشخص است:



	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
gamma	0.1823	0.0173	2.845E-4	0.1482	0.1823	0.2163	1001	20000

مقادیر اثرات تصادفی با پیشین $N(0, .01)$ به صورت زیر است:

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
a[1]	-0.7144	0.8502	0.04447	-2.376	-0.7054	0.9381	1001	20000
a[2]	-0.3776	0.8917	0.0438	-2.134	-0.3752	1.365	1001	20000
a[3]	-0.1344	0.9795	0.04265	-2.075	-0.1299	1.766	1001	20000
a[4]	0.3354	0.8587	0.04364	-1.335	0.3361	2.008	1001	20000
a[5]	1.703	0.8263	0.04354	0.07218	1.705	3.316	1001	20000
a[6]	-0.4197	0.8675	0.04367	-2.121	-0.4178	1.26	1001	20000
a[7]	-1.114	0.8338	0.04364	-2.766	-1.111	0.5037	1001	20000
a[8]	0.8353	0.8129	0.04439	-0.7603	0.8495	2.395	1001	20000
a[9]	0.07673	0.885	0.0436	-1.673	0.08527	1.789	1001	20000
a[10]	-0.2244	0.8864	0.04353	-1.974	-0.2131	1.477	1001	20000
a[11]	-0.5266	0.8561	0.04342	-2.196	-0.5237	1.158	1001	20000
a[12]	-2.095	0.919	0.04374	-3.897	-2.094	-0.2972	1001	20000
a[13]	1.308	0.9048	0.04355	-0.4536	1.305	3.099	1001	20000
a[14]	-1.335	0.8825	0.044	-3.088	-1.328	0.3884	1001	20000
a[15]	-2.083	1.106	0.04163	-4.271	-2.08	0.1122	1001	20000
a[16]	7.119	2.383	0.02673	2.424	7.144	11.76	1001	20000
a[17]	-0.5707	2.369	0.02623	-5.141	-0.5556	4.068	1001	20000

باتوجه به شکل زیر DIC مدل برابر ۲۰۳۵۰ است:

	Dbar	Dhat	DIC	pD
HR	20310.0	20280.0	20340.0	29.26
gamma	6.446	6.446	6.446	2.931E-13
total	20320.0	20290.0	20350.0	29.26

مدل نوزدهم : ترکیب مدل مکانی و پانلی

در این مدل اثرات تصادفی پانلی و مکانی را باهم درنظر می‌گیریم.

$$HR_{it} = X_i * \beta + a[ID(i)] + u[ID[i]] + \epsilon_i$$

$$X_i = [\text{MFIL}, \text{PO}, \text{RD}, \text{PS}, \text{UE}, \text{DV}, \text{MA}, \text{FP}, \text{BLK}, \text{FH}]$$

$$\beta = [\beta_0, \dots, \beta_{11}]$$

$$\epsilon_i \sim N(0, \sigma^2)$$

توزیع پیشین β را نرمال چند متغیره ، σ^2 را معکوس گاما در نظر می‌گیریم.

هردو توزیع را آگاهی نابخش(واریانس زیاد و میانگین صفر) در نظر گرفتیم.

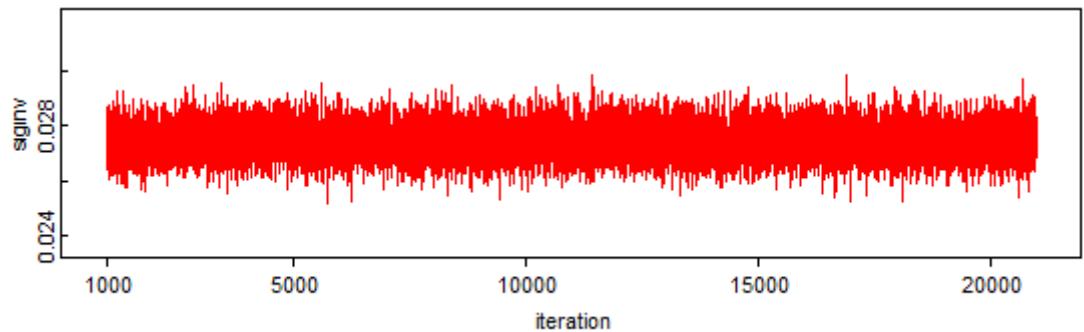
تخمین β ها و واریانس آنها را میتوان در جدول زیر مشاهده کرد :

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta[1]	25.38	3.452	0.02633	18.58	25.37	32.21	1001	20000
beta[2]	-1.587	0.3689	0.002681	-2.318	-1.587	-0.868	1001	20000
beta[3]	6.797E-6	9.174E-7	6.541E-9	4.979E-6	6.794E-6	8.598E-6	1001	20000
beta[4]	2.445	0.2149	0.001691	2.026	2.444	2.871	1001	20000
beta[5]	-0.1414	0.1856	0.001462	-0.5023	-0.1413	0.2291	1001	20000
beta[6]	-0.04475	0.04574	3.976E-4	-0.134	-0.04458	0.04473	1001	20000
beta[7]	0.2138	0.08833	6.953E-4	0.04077	0.2146	0.3854	1001	20000
beta[8]	-0.09722	0.02442	1.894E-4	-0.1444	-0.09724	-0.04872	1001	20000
beta[9]	-0.1724	0.01562	1.1E-4	-0.2031	-0.1724	-0.1417	1001	20000
beta[10]	0.05063	0.01007	9.066E-5	0.03096	0.05061	0.07053	1001	20000
beta[11]	0.1715	0.04353	3.204E-4	0.08664	0.1712	0.2575	1001	20000

همانطور که مشاهده می کنید واریانس تخمین عرض از مبدا ، ضریب گزارش میانه درآمد خانواده (MFIL) ، ضریب ساختار جمعیت (PS) و ضریب محرومیت از منابع (RD) زیاد می باشد، بنابراین به آن کمتر می توان اعتماد کرد.

با توجه به بررسی های انجام شده بر روی نمودار ها، β های تولید شده توسط این مدل همگرا شده بودند، پس مدل با موفقیت پیاده سازی شده است .

نمودار تاریخچه واریانس در شکل صفحه بعد مشخص است:



مقدادیر اثرات تصادفی پانلی:

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
a[1]	-0.2065	0.6441	0.01503	-1.473	-0.2109	1.067	1001	20000
a[2]	-0.5889	0.6445	0.01184	-1.823	-0.5955	0.7062	1001	20000
a[3]	0.1559	0.7558	0.01221	-1.328	0.1603	1.643	1001	20000
a[4]	0.4721	0.8476	0.02043	-1.239	0.4897	2.111	1001	20000
a[5]	1.295	0.8185	0.02347	-0.4462	1.347	2.774	1001	20000
a[6]	-0.4654	0.6376	0.01185	-1.721	-0.4644	0.7856	1001	20000
a[7]	-0.0298	0.6799	0.01798	-1.297	-0.05923	1.363	1001	20000
a[8]	0.1001	0.6694	0.02083	-1.24	0.1068	1.393	1001	20000
a[9]	0.4806	0.6489	0.01076	-0.7909	0.4803	1.753	1001	20000
a[10]	-0.08549	0.6597	0.01206	-1.395	-0.08186	1.218	1001	20000
a[11]	-0.2996	0.741	0.01551	-1.775	-0.2981	1.165	1001	20000
a[12]	-1.559	0.6493	0.01254	-2.777	-1.58	-0.227	1001	20000
a[13]	1.248	0.7525	0.01669	-0.3115	1.28	2.631	1001	20000
a[14]	-0.5244	0.6428	0.01395	-1.725	-0.5467	0.7789	1001	20000
a[15]	-0.7359	0.7987	0.01267	-2.281	-0.7495	0.8872	1001	20000
a[16]	1.003	0.9608	0.006562	-0.8856	1.006	2.876	1001	20000
a[17]	-0.02162	0.9491	0.006886	-1.886	-0.02717	1.842	1001	20000

مقدادیر اثرات تصادفی مکانی:

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
u[1]	0.08437	0.5696	0.01213	-1.11	0.09134	1.211	1001	20000
u[2]	-0.0211	0.5477	0.01046	-1.204	0.001371	1.011	1001	20000
u[3]	-0.3997	0.7017	0.01111	-1.824	-0.3828	0.9648	1001	20000
u[4]	0.2447	0.8262	0.01953	-1.343	0.2145	1.935	1001	20000
u[5]	1.066	0.7848	0.02239	-0.2948	0.9969	2.772	1001	20000
u[6]	0.1012	0.5306	0.009056	-0.9823	0.09981	1.149	1001	20000
u[7]	-0.7109	0.6226	0.01622	-2.036	-0.6637	0.4208	1001	20000
u[8]	0.3242	0.6294	0.01859	-0.9124	0.3146	1.583	1001	20000
u[9]	0.164	0.5583	0.008863	-0.9463	0.1586	1.288	1001	20000
u[10]	-0.1027	0.5612	0.009415	-1.243	-0.1046	1.022	1001	20000
u[11]	0.147	0.6992	0.01421	-1.284	0.1596	1.511	1001	20000
u[12]	-0.5407	0.5537	0.013	-1.869	-0.4553	0.3178	1001	20000
u[13]	0.8226	0.7089	0.01648	-0.4045	0.7632	2.375	1001	20000
u[14]	-0.6595	0.5403	0.01249	-1.877	-0.6053	0.2626	1001	20000
u[15]	-0.9818	0.7606	0.01551	-2.685	-0.9062	0.2984	1001	20000
u[16]	0.9133	1.172	0.02227	-0.8693	0.7249	3.744	1001	20000
u[17]	-0.4516	0.7277	0.01065	-2.03	-0.4097	0.9075	1001	20000

باتوجهه به شکل زیر DIC مدل

برابر ۲۷۲۴۰ است:

	Dbar	Dhat	DIC	pD
HR	27250.0	27260.0	27240.0	-7.066
total	27250.0	27260.0	27240.0	-7.066

مدل بیستم : ترکیب مدل مکانی و پانلی و زمانی

در این مدل اثرات تصادفی پانلی و مکانی را به همراه لگ زمان قبلی در نظر می‌گیریم.

$$HR_{it} = X_i * \beta + a[ID(i)] + u[ID[i]] + \gamma * HR_{it-1} + \epsilon_i$$

$$X_i = [\text{MFIL}, \text{PO}, \text{RD}, \text{PS}, \text{UE}, \text{DV}, \text{MA}, \text{FP}, \text{BLK}, \text{FH}]$$

$$\beta = [\beta_0, \dots, \beta_{11}]$$

$$\epsilon_i \sim N(0, \sigma^2)$$

توزیع پیشین β را نرمال چند متغیره ، σ^2 را معکوس گاما در نظر می‌گیریم.

هردو توزیع را آگاهی نابخش(واریانس زیاد و میانگین صفر) در نظر گرفتیم.

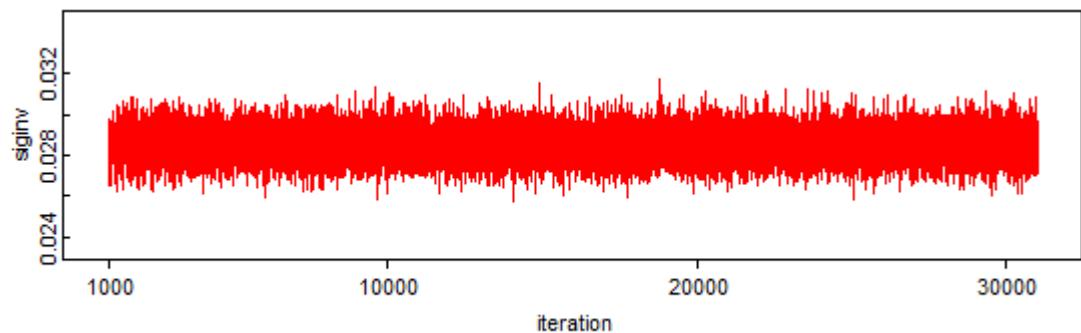
تخمین β ها و واریانس آنها را میتوان در جدول زیر مشاهده کرد :

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta[1]	24.84	3.875	0.02427	17.26	24.81	32.5	1001	30000
beta[2]	-2.171	0.4048	0.002438	-2.967	-2.167	-1.382	1001	30000
beta[3]	5.393E-6	9.636E-7	6.018E-9	3.506E-6	5.386E-6	7.292E-6	1001	30000
beta[4]	0.5959	0.4215	0.002607	-0.2275	0.5976	1.42	1001	30000
beta[5]	0.2216	0.215	0.001328	-0.1976	0.2222	0.642	1001	30000
beta[6]	-0.02762	0.05321	3.633E-4	-0.1318	-0.0276	0.07707	1001	30000
beta[7]	0.1921	0.09382	6.819E-4	0.008816	0.192	0.3771	1001	30000
beta[8]	-0.04352	0.02866	1.79E-4	-0.09924	-0.04336	0.01292	1001	30000
beta[9]	-0.03857	0.03573	2.07E-4	-0.1091	-0.03857	0.03153	1001	30000
beta[10]	0.06877	0.01268	9.192E-5	0.04392	0.06878	0.09337	1001	30000
beta[11]	0.2643	0.05569	3.365E-4	0.1557	0.2644	0.3749	1001	30000

همانطور که مشاهده می‌کنید واریانس تخمین عرض از مبدأ ، ضریب گزارش میانه درآمد خانواده (MFIL) ، ضریب ساختار جمعیت (PS) و ضریب محرومیت از منابع (RD) زیاد می باشد، بنابراین به آن کمتر می‌توان اعتماد کرد.

با توجه به بررسی های انجام شده بر روی نمودار ها، β های تولید شده توسط این مدل همگرا شده بودند، پس مدل با موفقیت پیاده سازی شده است .

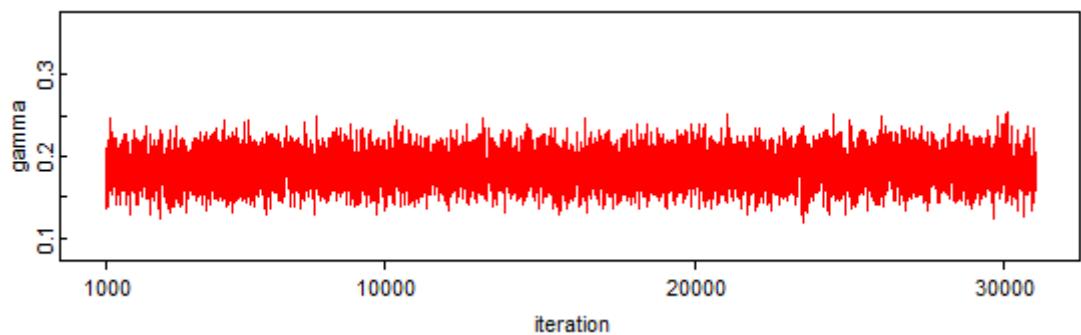
نمودار تاریخچه واریانس در شکل صفحه بعد مشخص است:



	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
siginv	0.02853	7.207E-4	4.25E-6	0.02714	0.02852	0.02995	1001	30000

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
tauinv	1.315	1.031	0.02166	0.1832	1.032	4.088	1001	30000

نمودار تاریخچه ضریب گاما:



	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
gamma	0.1851	0.01735	2.354E-4	0.1513	0.1849	0.2194	1001	30000

برآورد بیز عرض از مبدأ تصادفی :

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
a[1]	-0.4063	0.6208	0.01038	-1.603	-0.4123	0.8332	1001	30000
a[2]	-0.2573	0.6281	0.008829	-1.48	-0.2606	0.9948	1001	30000
a[3]	0.2013	0.7634	0.008629	-1.293	0.2001	1.706	1001	30000
a[4]	0.3912	0.7997	0.01333	-1.208	0.4006	1.948	1001	30000
a[5]	1.051	0.7479	0.01698	-0.4982	1.083	2.441	1001	30000
a[6]	-0.306	0.6225	0.009007	-1.542	-0.3082	0.9106	1001	30000
a[7]	-0.3034	0.6779	0.01148	-1.572	-0.3279	1.097	1001	30000
a[8]	0.5544	0.6638	0.01488	-0.8272	0.5815	1.818	1001	30000
a[9]	0.1016	0.6325	0.008632	-1.148	0.1016	1.341	1001	30000
a[10]	-0.08475	0.6511	0.008292	-1.372	-0.08191	1.204	1001	30000
a[11]	-0.3452	0.7125	0.01154	-1.733	-0.345	1.078	1001	30000
a[12]	-1.199	0.6366	0.00828	-2.406	-1.21	0.1086	1001	30000
a[13]	0.7871	0.7018	0.01028	-0.6399	0.804	2.118	1001	30000
a[14]	-0.5308	0.639	0.009472	-1.729	-0.5505	0.7666	1001	30000
a[15]	-0.7225	0.7883	0.007854	-2.243	-0.7284	0.8448	1001	30000
a[16]	1.133	0.9631	0.005905	-0.7676	1.128	3.008	1001	30000
a[17]	-0.05611	0.9692	0.005529	-1.96	-0.04986	1.83	1001	30000

برآورد بیز اثرات تصادفی مکانی:

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
u[1]	0.01209	0.529	0.008553	-1.123	0.03249	1.017	1001	30000
u[2]	0.1005	0.4858	0.006308	-0.908	0.1072	1.055	1001	30000
u[3]	-0.3202	0.6796	0.008757	-1.727	-0.3014	1.012	1001	30000
u[4]	0.03619	0.7666	0.01296	-1.462	0.02257	1.598	1001	30000
u[5]	0.6711	0.6945	0.01625	-0.5503	0.6157	2.173	1001	30000
u[6]	0.1536	0.4826	0.006937	-0.7791	0.1385	1.154	1001	30000
u[7]	-0.6252	0.6018	0.01116	-1.974	-0.5687	0.4177	1001	30000
u[8]	0.4457	0.5972	0.01254	-0.6632	0.4174	1.721	1001	30000
u[9]	0.1023	0.511	0.006931	-0.9321	0.09785	1.134	1001	30000
u[10]	0.008268	0.5127	0.006622	-1.026	3.577E-5	1.064	1001	30000
u[11]	0.09136	0.6584	0.01077	-1.248	0.09649	1.384	1001	30000
u[12]	-0.3297	0.4711	0.007792	-1.465	-0.2627	0.4268	1001	30000
u[13]	0.5047	0.6191	0.009882	-0.6135	0.4638	1.836	1001	30000
u[14]	-0.5049	0.487	0.008535	-1.592	-0.454	0.3269	1001	30000
u[15]	-0.7935	0.699	0.01205	-2.382	-0.7103	0.3662	1001	30000
u[16]	0.8553	1.104	0.01918	-0.7185	0.6595	3.634	1001	30000
u[17]	-0.4077	0.6582	0.008352	-1.848	-0.3699	0.807	1001	30000

همه نمودارهای تاریخچه a و u همگرا هستند.

باتوجه به شکل زیر DIC مدل ۲۰۳۰۰ است.

:	Dbar	Dhat	DIC	pD
HR	20320.0	20340.0	20300.0	-16.26
total	20320.0	20340.0	20300.0	-16.26

در صورتی که برای زمان اول نیز مدل مکانی و پانلی بزنیم DIC به صورت زیر است:

:	Dbar	Dhat	DIC	pD
HR	20320.0	20340.0	20300.0	-18.34
HR1	6744.0	6740.0	6748.0	4.134
total	27060.0	27080.0	27050.0	-14.2

ابتدا ۲۰ مدل را به طور خلاصه بیان کرده و سپس از روی جدول مرتبط با آن ها، بهترین مدل را انتخاب می کنیم.

مدل اول تا چهارم : مدل رگرسیون خطی چندگانه است با توزیع های نرمال، نرمال غیر مرکزی(SQ_N)، توزیع t و توزیع t غیر مرکزی.

مدل پنجم تا هشتم : مدل رگرسیون خطی چندگانه با عرض از مبدا تصادفی است با توزیع های نرمال، نرمال غیر مرکزی(SQ_N)، توزیع t و توزیع t غیر مرکزی.

مدل نهم تا دوازدهم : مدل رگرسیون خطی با تأخیر مرتبه اول با توزیع های نرمال، نرمال غیر مرکزی(SQ_N)، توزیع t و توزیع t غیر مرکزی.

مدل سیزدهم تا شانزدهم : مدل های فضایی با نرمال، نرمال غیر مرکزی(SQ_N)، توزیع t و توزیع t غیر مرکزی.

مدل هفدهم تا بیستم : مدل های ترکیبی با نرمال، نرمال غیر مرکزی(SQ_N)، توزیع t و توزیع t غیر مرکزی.

MSE_Train	آگاهی بخش بودن	همگرایی تاریخچه ها	DIC	مدل
۳۷/۵۲	x	به خوبی همگرا بودند.	۲۷۴۰۰	مدل اول
۹۳/۵۷	x	تنها همگرایی واریانس خوب نبود.	۲۶۹۴۰	مدل دوم
۳۹/۲۳	x	به خوبی همگرا بودند.	۲۶۱۱۰	مدل سوم
۷۸/۱۹	x	به خوبی همگرا بودند.	۲۶۱۲۰	مدل چهارم
۳۶/۲۴	✓	به خوبی همگرا بودند.	۲۷۲۸۰	مدل پنجم
۸۳/۰۶	✓	برخی همگرایی خوبی نداشتند.	۲۶۷۱۰	مدل ششم
۴۴/۸۷	✓	به خوبی همگرا بودند.	۲۵۹۲۰	مدل هفتم
۶۹/۲۱	✓	اثرات تصادفی همگرا نشدند.	۲۵۲۵۰	مدل هشتم
۳۷/۵۵	x	به خوبی همگرا بودند.	۲۷۷۲۰	مدل نهم
۱۰۲/۰۱	x	نسبتا همگرا بودند.	۲۶۵۳۰	مدل دهم
۳۹/۱۷	x	به خوبی همگرا بودند.	۲۵۸۵۰	مدل یازدهم
۸۹/۶۱	x	به خوبی همگرا بودند.	۲۷۷۲۰۰	مدل دوازدهم
۳۷/۹۹	x	به خوبی همگرا بودند.	۲۷۷۲۵۰	مدل سیزدهم
۸۹/۹۸	x	به خوبی همگرا بودند.	۲۶۶۲۰	مدل چهاردهم

مدل پانزدهم	۲۵۹۰۰	به خوبی همگرا بودند.	<input checked="" type="checkbox"/>	۳۹/۳۱
مدل شانزدهم	۲۴۸۰۰	برخی همگرایی خوبی نداشتند.	<input checked="" type="checkbox"/>	۷۸/۱۷
مدل هفدهم	۲۰۳۴۰	به خوبی همگرا بودند.	<input checked="" type="checkbox"/>	۲۷/۷۹
	۲۷۱۰۴	به خوبی همگرا بودند.	<input checked="" type="checkbox"/>	۳۵/۹۸
مدل هجدهم	۲۰۳۵۰	به خوبی همگرا بودند.	<input checked="" type="checkbox"/>	۲۷/۶۰
	۲۷۱۱۴	به خوبی همگرا بودند.	<input checked="" type="checkbox"/>	۳۷/۶۱
مدل نوزدهم	۲۷۲۴۰	به خوبی همگرا بودند.	<input checked="" type="checkbox"/>	۳۶/۲۱
مدل بیستم	۲۰۳۰۰	به خوبی همگرا بودند.	<input checked="" type="checkbox"/>	۲۷/۷
	۲۷۰۵۰	به خوبی همگرا بودند.	<input checked="" type="checkbox"/>	۳۶/۰۸

با توجه به جدول بالا، مدل هجدهم از آنجایی که **کمترین خطای DIC** را دارد و همچنین مدل آن به صورت **آگاهی نابخش** است و اطلاعات اضافه وارد مدل نشده است بهترین مدل در میان مدل ها می باشد. پس مدل نهایی به صورت زیر می باشد:

$$HR_{it} = 24.64 - 2.158 \times MFIL + 5.281 \times 10^{-6} \times PO + 0.5357 \times RD \\ + 0.2307 \times PS - 0.00884 \times UE + 0.1888 \times DV \\ - 0.04414 \times MA - 0.03278 \times FP + 0.06698 \times BLK \\ + 0.2619 \times FH + a[ID(i)] + 0.1823 \times HR_{it-1}$$

$$a=[-0.7144, -0.3776, -0.1344, 0.3354, 1.703, -0.4197, -1.114, 0.8353, \\ 0.07673, -0.2244, -0.5266, -2.095, 1.308, -1.335, -2.083, 7.119, -0.5707]$$

زمانی که همه ی متغیر ها صفر باشند، عرض از مبدا مقدار **نرخ قتل** را پیش بینی می کند. اما این ضریب با شهود ما مطابقت ندارد. زیرا در جامعه ای که جمعیت صفر باشد، **نرخ قتل** نیز معنایی ندارد. پس عرض از مبدا در داده های ما تفسیر درستی ندارد.

ضرایب بقیه ی متغیرها ، کاهش یا افزایش **نرخ قتل** با تغییر هر متغیر را نشان می دهد. برای مثال با افزایش یک واحد به **میانه درآمد خانواده**، در صورت ثابت بودن بقیه ی متغیر ها، به طور میانگین **نرخ قتل**، ۲/۱۵۸ کاهش خواهد داشت. همچنین اگر همه ی متغیر ها ثابت بماند، با افزایش یک واحد **نرخ طلاق**، به طور میانگین **نرخ قتل**، ۰/۱۸۸۸ افزایش خواهد داشت. به همین صورت میتوان بقیه ی ضرایب را نیز تفسیر کرد.

یکی از نتایج قابل توجه که کمی با تفکر اولیه ما در تناقض بود این است که نرخ بیکاری با نرخ قتل در این مدل رابطه‌ی عکس دارد!!!!

باتوجه به ضرایب متغیرهای موجود در مدل، میانه درآمد خانواده‌ها و محرومیت از منابع بیشترین تاثیر را در میانگین نرخ قتل دارد. پس میتوان با کنترل این دو مورد بیشترین کنترل را بر روی نرخ قتل داشت.

تفسیر بردار a و lag:

در صورت ثابت ماندن همه متغیرهای توضیحی، با افزایش یک واحد نرخ قتل در زمان قبل(lag)، به طور میانگین نرخ قتل، ۰.۱۸۲۳ افزایش خواهد داشت.

در پانل اول به طور میانگین عرض از مبدا ۰.۷۱۴۴ کاهش می‌یابد.

می‌توان برای هر پانل مقدار MSE مدل را محاسبه کرد. این مقادیر در جدول زیر قابل مشاهده است.

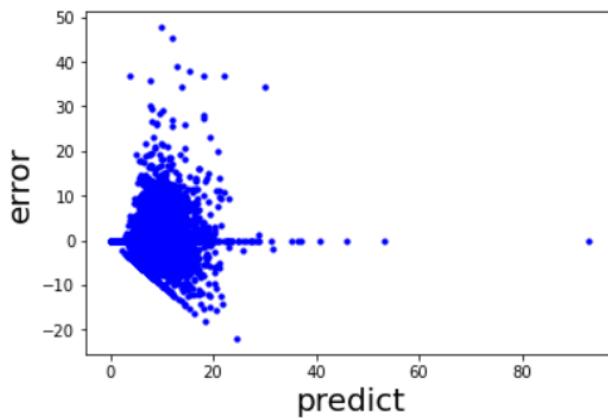
ایالت	Mississippi	Louisiana	Florida	Texas	West Virginia	Maryland	Kentucky	
	30,18	30,17	37,30	46,46	20,34	18,59	38,28	MSE

Delaware	Arkansas	Virginia	Tennessee	Oklahoma	South Carolina	Alabama	Georgia	
1,66	27,98	29,07	30,42	20,06	17,20	33,68	65,58	

DC	North Carolina
392,97	25,05

بنابراین دو ایالت واشنگتن DC و Delaware با توجه به اینکه دیتا کمتری داشتند نسبت به ایالت‌های دیگر خطای پرتی دارند.

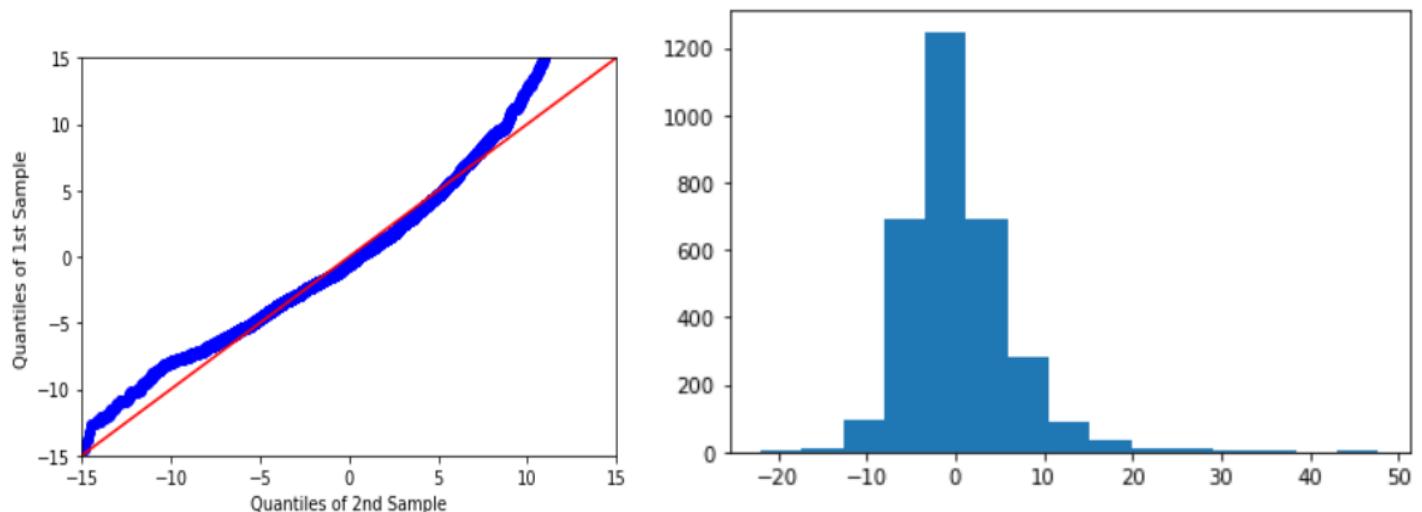
اکنون قصد داریم میزان دقیقیت مدل نهایی را بررسی کنیم.



نمودار بالا، تفاوت تخمین نرخ [قتل](#) را با خطأ نشان می‌دهد.

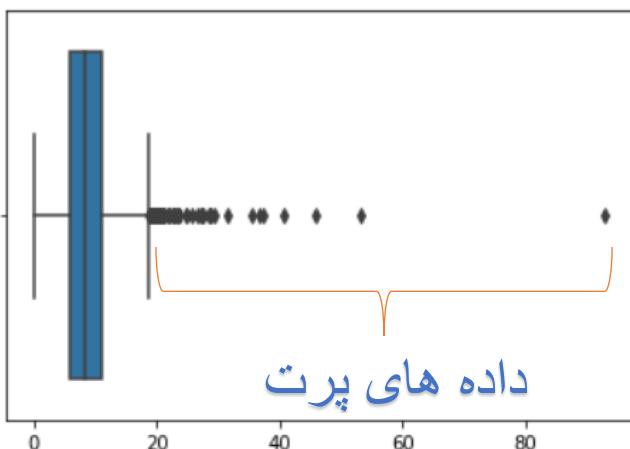
نرمال بودن مانده‌ها:

میانگین مانده‌ها عدد کوچکی (0.183) است و میتوان نتیجه گرفت داده‌ها به خوبی تخمین زده شده‌اند. همچنین با استفاده از نمودار $Q-Q$ normal می‌توان دید که داده‌ها تقریباً از فرض نرمال بودن مانده‌ها حمایت می‌کنند.

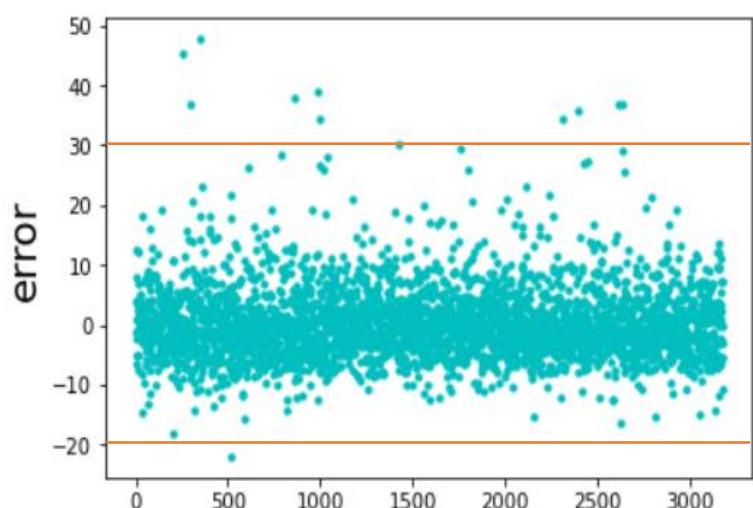


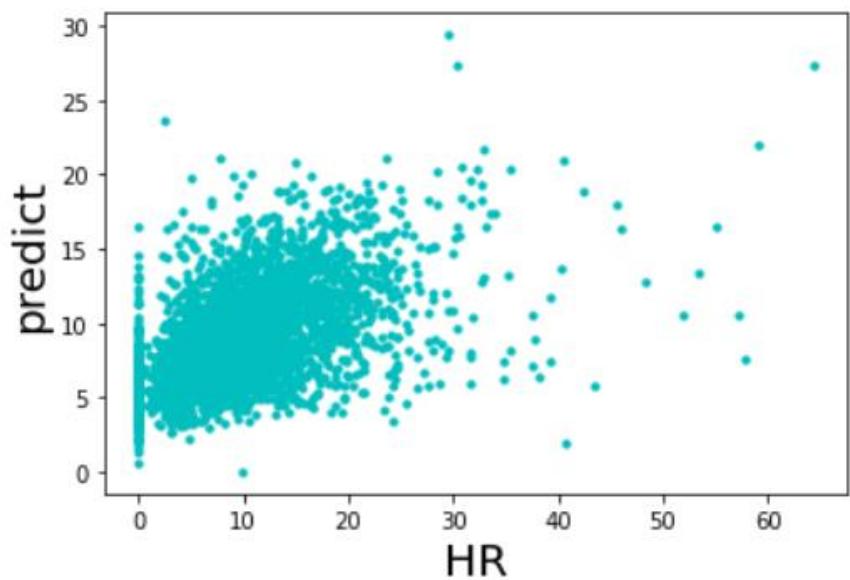
داده های پرت و تاثیرگذار:

با توجه به نمودارهای زیر میتوان دید که داده های پرت قابل توجهی در بین این داده ها وجود دارند. می توان با بررسی آمارهای دیگر مثل آماره کوک تاثیرگذار بودن آن ها را نیز بررسی کرد و سپس تصمیم به حذف کردن آن ها گرفت از آنجایی که هدف از انجام این پروژه کار با مدل های بیز است فرصت انجام این موارد را نداشته و فقط در حد اشاره کردن به آن اکتفا می کنیم.



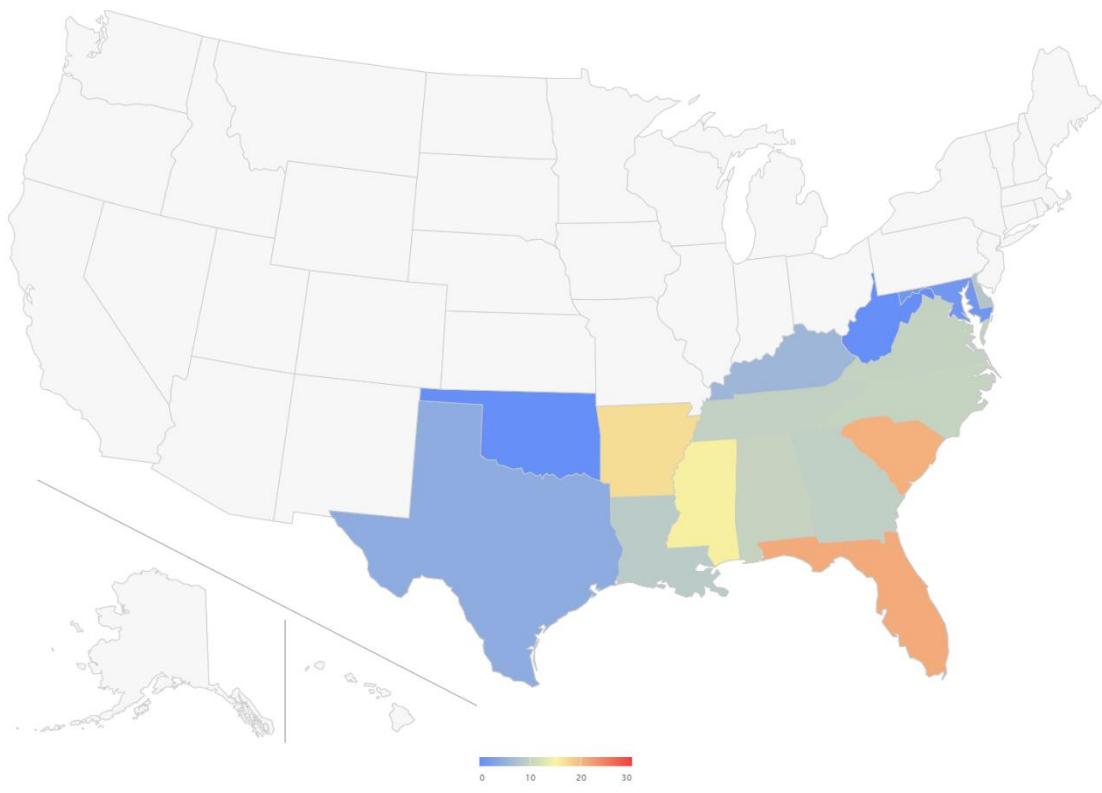
نمودار جعبه ای پیش بینی
های نرخ قتل





در آخر مدل را بر روی داده‌ی آزمایشی اجرا کرده و MSE آن $45/81$ شد.

HR



HR_hat

