Indian Institute of Technology, Kanpur

Weibull Distribution

Parameter Estimation

Instructor: Prof. Debasis Kundu
By: Medha Agarwal and Kushagra Gupta
IIT Kanpur
May-June 2019

Table of Contents

Modified Maximum Likelihood Method I

N	Alpha=1 MSE	Lambda=1 MSE	Mu=0 MSE
N=10	1.289778 0.2222859	1.263701 0.9373971	-0.0220419 0.00569825
N=25	1.072025 0.02337532	1.031736 0.03233327	-0.004406659 0.001243337
N=50	1.051093 0.01086262	1.022025 0.01714685	-0.00167184 0.0003978323
N=150	1.011396 0.003092481	1.001243 0.008124393	-0.0001960437 4.345961e-05
N=1000	1.003818 0.0007279983	1.001352 0.001216654	-3.673703e-05 1.032726e-06

Modified Maximum Likelihood Method II

N	Alpha=1 MSE	Lambda=1 MSE	Mu=0 MSE
N=10	1.18358 0.2775532	1.10116 0.1686073	0.00118 0.0157545
N=25	1.028782 0.04037525	1.068441 0.06353261	0.01313 0.00279033
N=50	0.9982208 0.0173581	1.03563 0.02474953	0.01222 0.00075894
N=150	1.004141 0.005403722	1.021777 0.01146452	0.00275 4.963e-05
N=1000	1.002173 0.0004236811	1.002071 0.001342691	-0.00034 1.2e-06

Modified Moment Method

N	Alpha=1 MSE	Lambda=1 MSE	Mu=0 MSE
N=10	1.225984 0.1358397	1.090897 0.3944333	-0.002535711 0.009407445
N=25	1.098357 0.04583176	1.022553 0.05818679	-0.003856826 0.001319399
N=50	1.054168 0.02160289	1.013998 0.02569172	-5.392384e-05 0.0004016688
N=150	1.016897 0.00668958	1.00244 0.008271582	1.611801e-05 4.9162e-05
N=1000	1.003477 0.001004	0.999893 0.001184215	6.523252e-06 1.025181e-06

Least Square Estimation

N	alpha = 1 (MSE)	lambda = 1 (MSE)	mu = 0.001 (MSE)
25	0.95258391 (0.002248285)	0.92011014 (0.006382390)	0.03409797 (0.001095475)
150	0.966001730 (1.155882e-03)	0.969752616 (9.149042e-04)	0.006622271 (3.160993e-05)
1000	0.99785847 (4.586141e-06)	0.99588767 (1.691129e-05)	0.00193791 (8.796756e-07)

N	alpha = 0.5 (MSE)	lambda = 1 (MSE)	mu = 0.001 (MSE)
25	0.527187469 (0.0007391585)	0.860688990 (0.0194075576)	0.004863988 (0.0000149304)
150	0.502472491 (6.113209e-06)	0.980536407 (3.788315e-04)	0.001092873 (8.625389e-09)
1000	0.503732050 (1.392820e-05)	0.991230074 (7.691161e-05)	0.001001867 (3.486508e-12)

Weighted Least Square Estimation

N	alpha = 1 (MSE)	lambda = 1 (MSE)	mu = 0.001 (MSE)
25	0.88672189 (0.012831929)	0.93336852 (0.004439754)	0.04050816 (0.001560895)
150	0.998539133 (2.134132e-06)	0.985379374 (2.137627e-04)	0.008276092 (5.294151e-05)
1000	1.000356549 (1.271275e-07)	1.002145366 (4.602595e-06)	0.002160981 (1.347876e-06)

N	alpha = 0.5 (MSE)	lambda = 1 (MSE)	mu = 0.001 (MSE)
25	0.490678270 (8.689466e-05)	0.894527678 (1.112441e-02)	0.004849586 (1.481931e-05)
150	0.497919639 (4.327903e-06)	0.988412827 (1.342626e-04)	0.001064324 (4.137577e-09)
1000	0.500661087 ()	0.991719605 ()	0.001002039 ()

MAP estimate using BFGS model in PyMC3

N	alpha	lambda
N=25	1.75281443	1.23679377
N=100	1.03823311	0.94018592
N=500	0.95274871	0.95886669
N=1000	1.00202081	1.01139949

MAP estimate using Powell method in PyMC3

N	alpha	Lambda
N=25	0.85221853	0.67099244
N=100	1.03823311	0.94018592
N=500	0.9527487	0.95886671
N=1000	1.00181734	1.01049776

MCMC Algorithm: Gibbs Sampling

Priors:

$\alpha \sim Gamma(1, 1)$
$\lambda \sim Gamma(1, 1)$

Gibbs sampling iteration: 500000

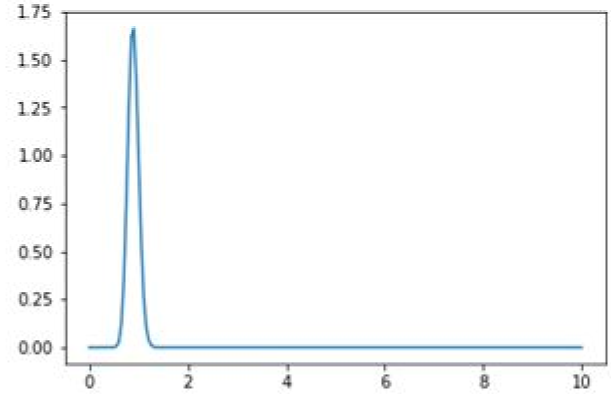


Fig: Plotting the conditional probability distribution of alpha, P(alpha|data) (non-normalized)

```
Estimated alpha is:   0.8945029575699471
Estimated lambda is: 1.1302730920315731
MSE for alpha:        0.022954689272950753
MSE for lambda:       0.04960441638245374
```

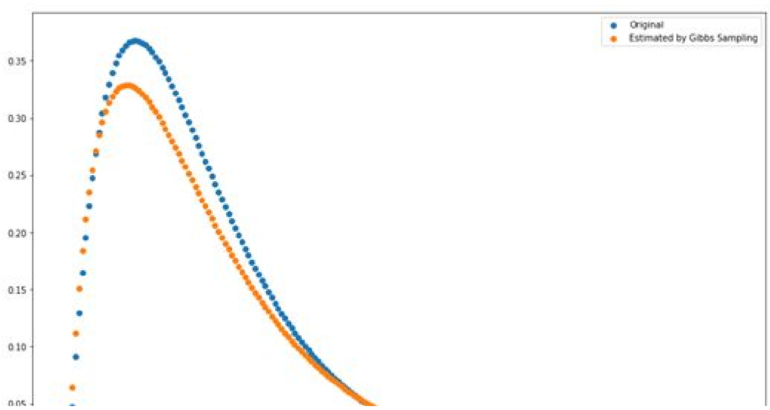


Fig: Comparison of Weibull distribution over 200 points with original and estimated alpha and lambda.

MCMC algorithm: Metropolis Hasting

The conditional probability for the unknown parameters is directly proportional to the product of prior and likelihood function.

$P(\alpha|X,\lambda) \sim P(X|\alpha,\lambda)\, P(\alpha)$

Similarly, $P(\lambda|X,\alpha) \sim P(X|\alpha,\lambda)\, P(\lambda)$

The proposal distribution for randomly walking in distribution space is taken as a normal distribution with mean as the previous sample value and standard deviation (proposal width) equal to 1.

$Q(\theta|\theta o) \sim Normal(\theta, 1)$

Initiate the value of α, λ by their MAP estimates.

Step 1: Generate α_{t+1} from $Normal(\alpha_t, 1)$

Similarly. Generate λ_{t+1} from $Normal(\lambda_t, 1)$

Step 2: Calculate the acceptance ratio for both the parameters:

$$R^1_{t+1} = \frac{P(\alpha_{t+1}|X,\lambda_{t-1})}{P(\alpha_t|X,\lambda_{t-1})} \quad \text{and} \quad R^2_{t+1} = \frac{P(\lambda_{t+1}|X,\alpha_t)}{P(\lambda_t|X,\alpha_{t-1})}$$

Step 3: Generate U^1_{t+1} from Uniform(0,1) and U^2_{t+1} from Uniform (0,1)

If $min(1,\ R^1_{t+1} = \frac{P(\alpha_{t+1}|X,\lambda_{t-1})}{P(\alpha_t|X,\lambda_{t-1})}) > U^1_{t+1}$, then accept α_{t+1} as the (t+1)th sample, else, α_t remains the (t+1)th sample.

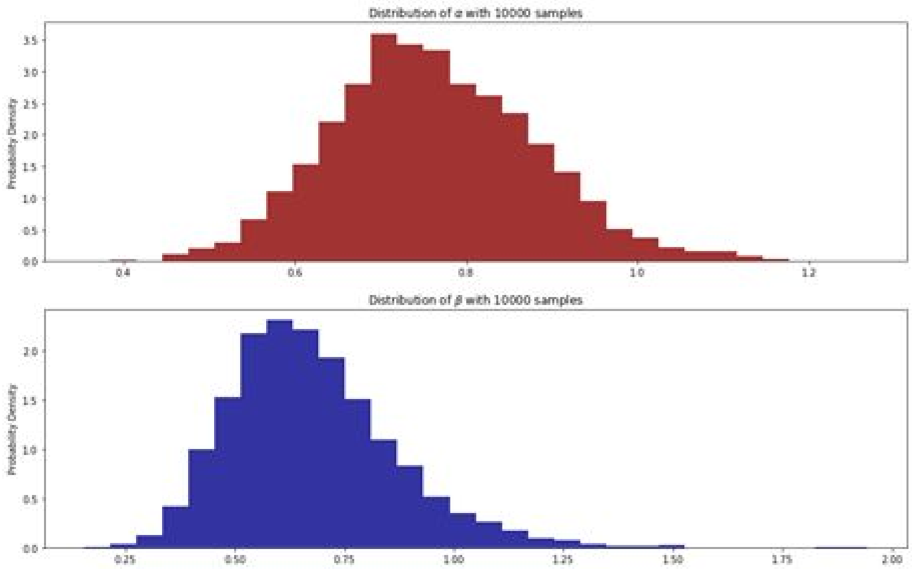
Repeat the above three steps for 10000 number of times.

Two chains are maintained simultaneously for each unknown parameter sampling.

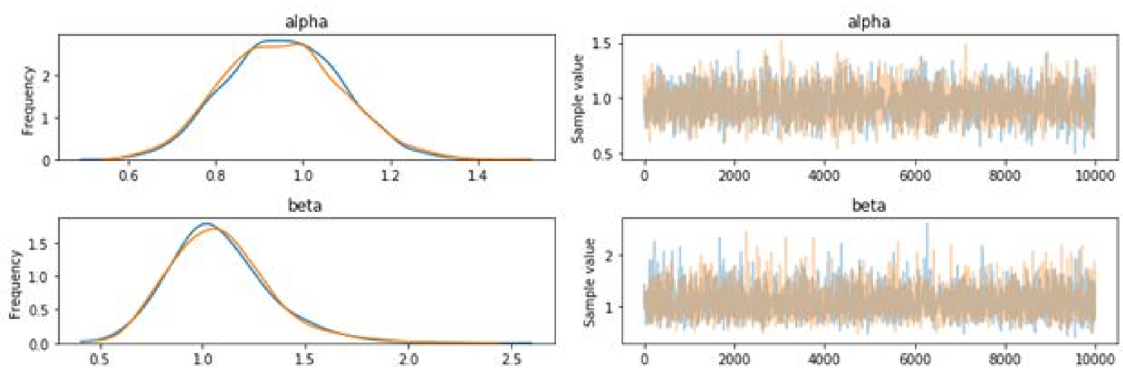
The histogram given below is plotted for the end half of the samples values.

Note - N stands for number of data points

N=25

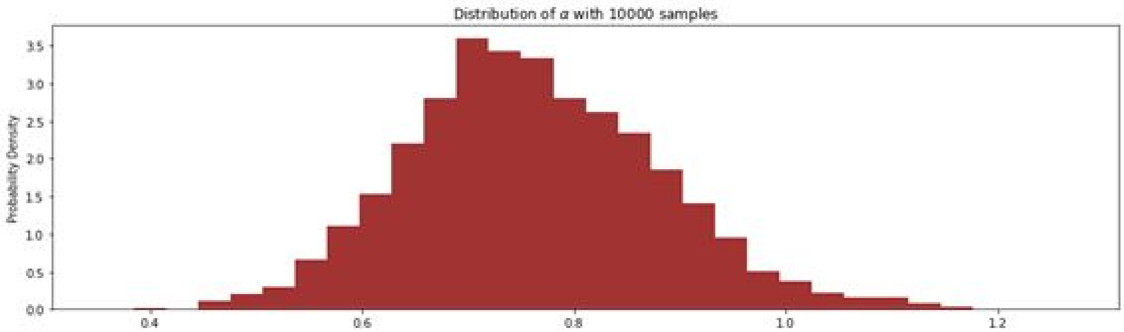


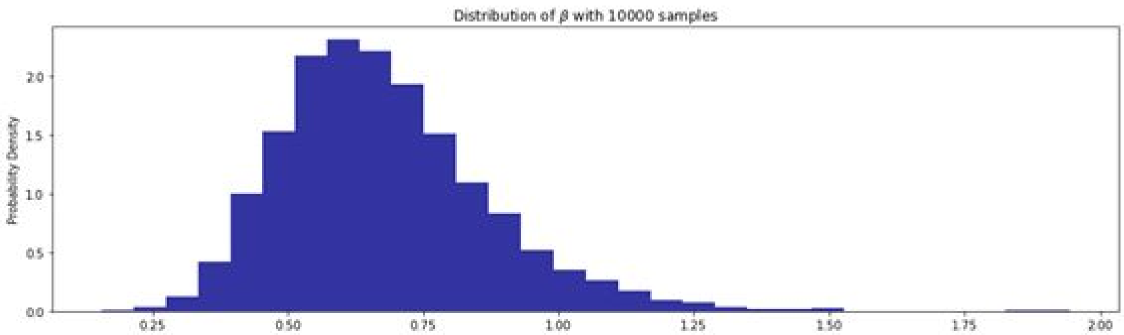
	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
alpha	0.763624	0.120185	0.002452	0.536529	1.002482	2678.670777	1.000106
beta	0.676136	0.190359	0.003020	0.331861	1.051015	3444.256254	1.000024



N=40

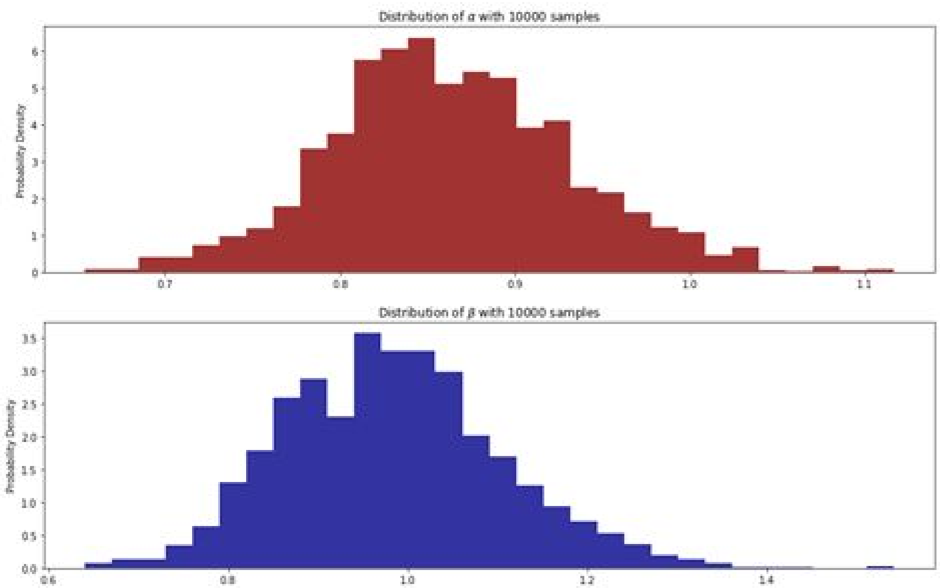
	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
alpha	0.763624	0.120185	0.002452	0.536529	1.002482	2678.670777	1.000106
beta	0.676136	0.190359	0.003020	0.331861	1.051015	3444.256254	1.000024

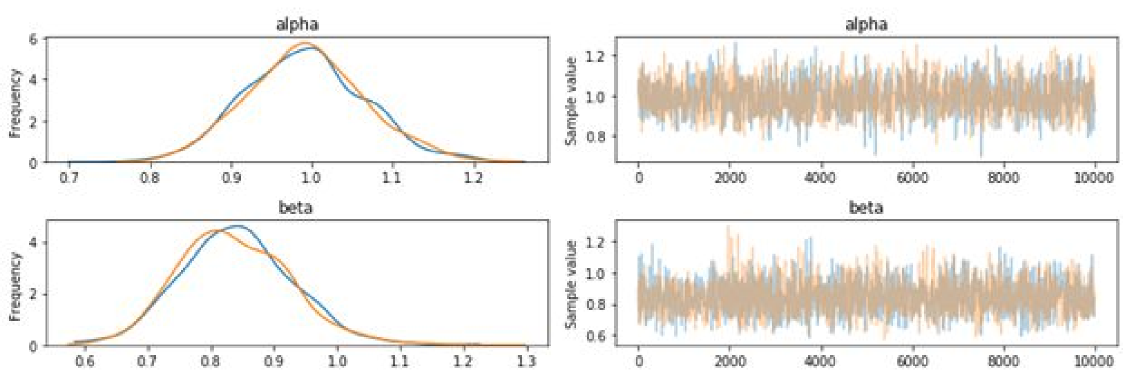




N=100

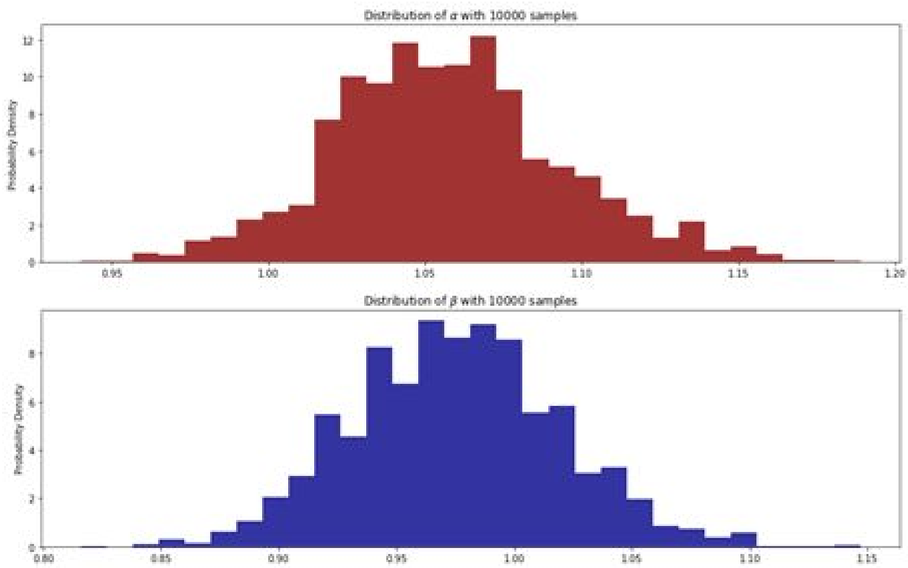
	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
alpha	0.865833	0.066550	0.001885	0.731164	0.994168	1498.776167	1.000905
beta	0.983499	0.120671	0.002491	0.758736	1.224548	2129.065360	0.999951

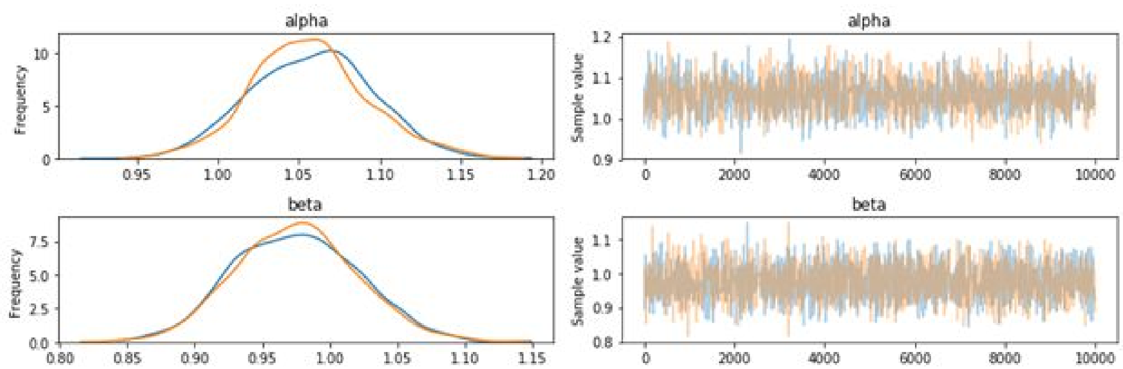




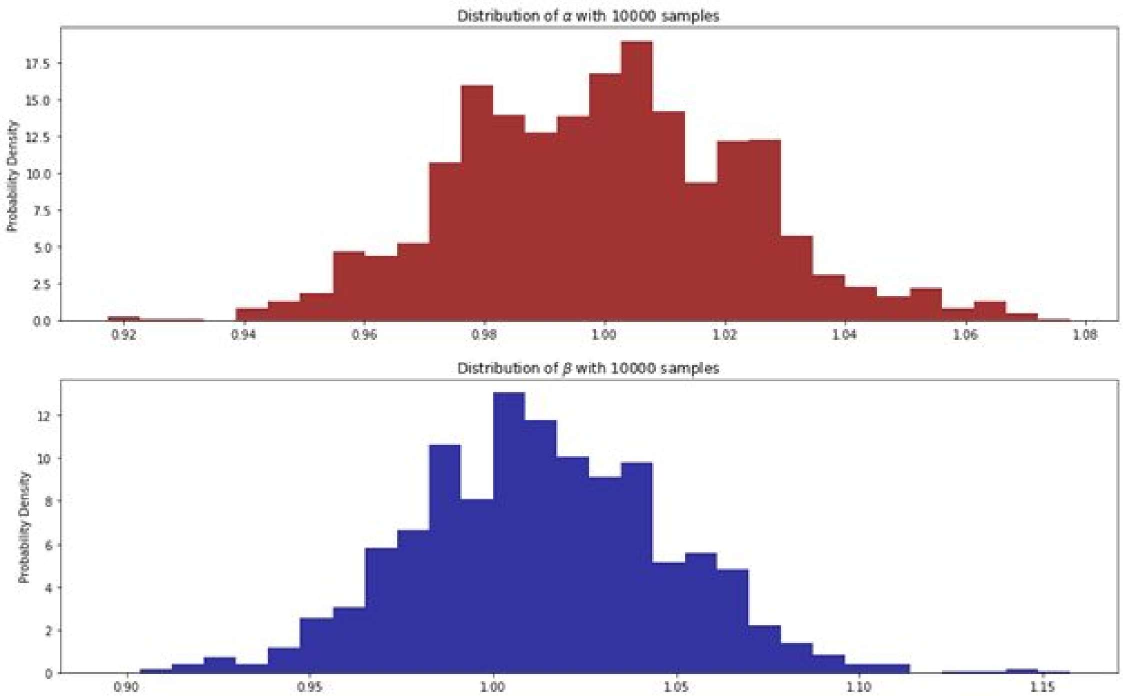
N=500

	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
alpha	1.057399	0.036855	0.001038	0.991939	1.137974	1254.158388	1.000198
beta	0.974905	0.044875	0.000948	0.889902	1.061426	1703.691877	1.000012

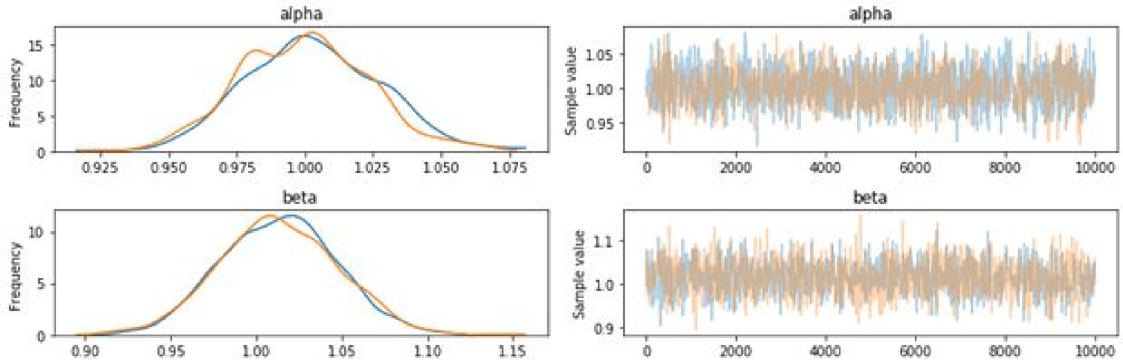




N=1000



	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff
alpha	1.001360	0.000602	0.000776	0.952100	1.047280	1032.047613
beta	1.013345	0.001148	0.000872	0.952771	1.081823	1389.367838



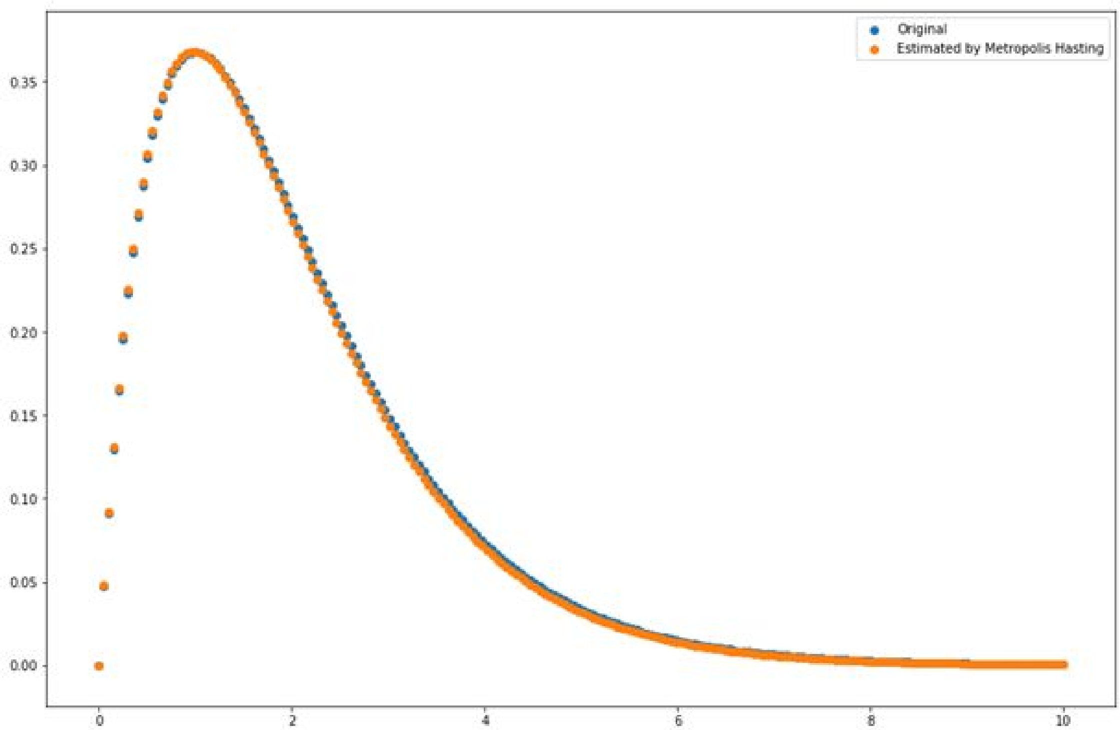


Fig: Comparison of Weibull distribution over 200 points with original and estimated alpha and lambda.

MCMC Algorithm using Metropolis Hasting for three variables: $\alpha,\ \lambda$ *and* μ

Priors: $\alpha \sim Gamma(1,1)$

$\lambda \sim Gamma(1,1)$

$\mu \sim Uniform(0,z1)$

$$P(\alpha|X,\lambda,\mu) \propto P(X|\alpha,\lambda,\mu)P(\alpha)$$
$$P(\lambda|\alpha,\mu,X) \propto P(X|\alpha,\lambda,\mu)P(\lambda)$$
$$P(\mu|\alpha,\lambda,X) \propto P(X|\alpha,\lambda,\mu)P(\mu)$$

The proposal distribution for randomly walking in distribution space is taken as a normal distribution with mean as the previous sample value and standard deviation (proposal width) equal to 1.

$$Q(\theta|\theta o) \sim Normal(\theta_o, 1)$$

Initiate the value of $\alpha,\ \lambda$ *and* μ by their MAP estimates.

Algorithm:

$\alpha_0,\ \lambda_0,\ \mu_0$ are the MAP estimates

Step 1: Generate α_{t+1} from $Normal(\alpha_t, 1)$

Generate λ_{t+1} from $Normal(\lambda_t, 1)$

Generate μ_{t+1} from $Normal(\mu_t, 1)$

Step 2: Calculate the acceptance ratio for all three parameters:

$$R^1_{t+1} = \frac{P(\alpha_{t+1}|X,\lambda_t,\mu_t))}{P(\alpha_t|X,\lambda_t,\mu_t)}\ ;\ R^2_{t+1} = \frac{P(\lambda_{t+1}|X,\alpha_t,\lambda_t)}{P(\lambda_t|X,\alpha_t,\lambda_t)}\ ;\ R^3_{t+1} = \frac{P(\mu_{t+1}|X,\alpha_t,\lambda_t)}{P(\mu_t|X,\alpha_t,\lambda_t)}$$

Step 3: Generate U_{t+1}^{1} from Uniform(0,1), U_{t+1}^{2} from Uniform (0,1) and U_{t+1}^{3} from Uniform(0,1)

If $min(1,\ R_{t+1}^{1}) > U_{t+1}^{1}$, then accept α_{t+1} as the (t+1)th sample, else, α_{t} remains the (t+1)th sample.

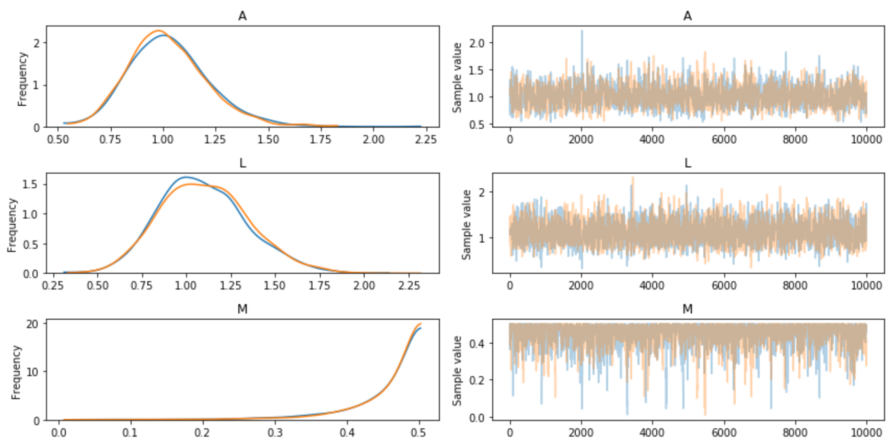
Similarly for R_{t+1}^{2} and R_{t+1}^{3}

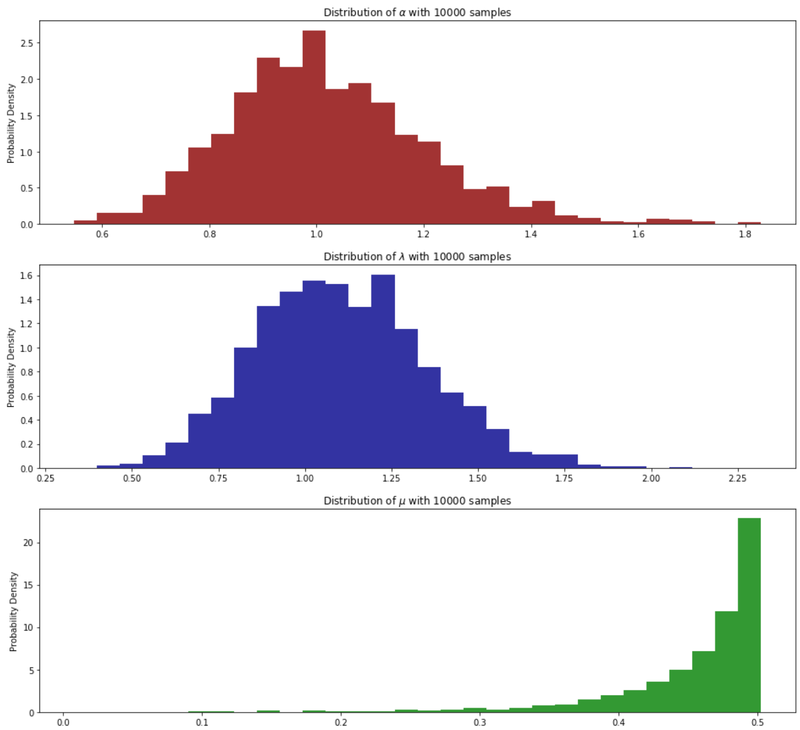
The above steps are repeated 10,000 times in two parallel chains.
The estimated values of $\alpha,\ \lambda\ and\ \mu$ are taken to be average of later half of samples.

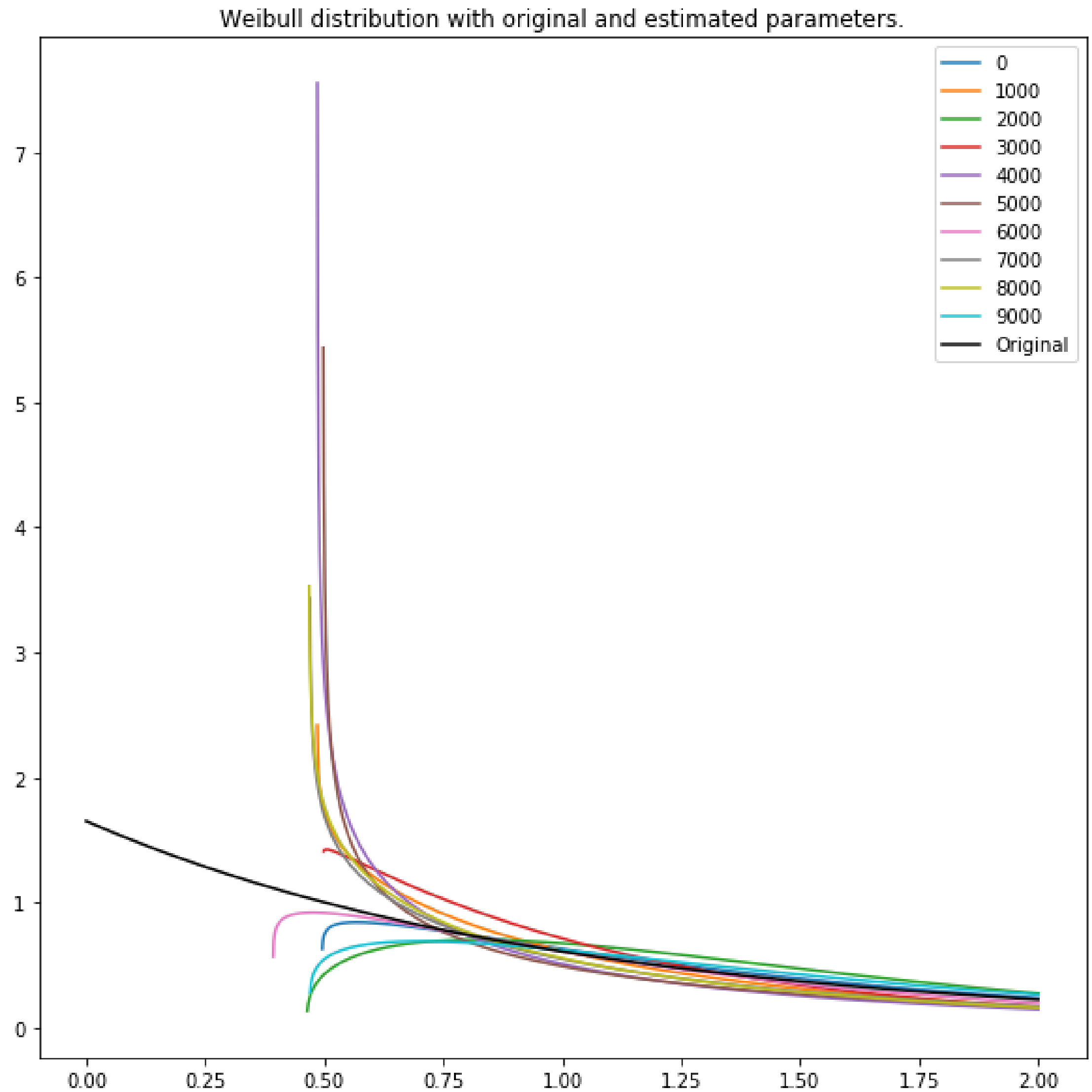
N=25, Alpha=1, Lambda=1, Mu=0.5

priors - λ ~Gamma(1, 1) α ~Gamma(1, 1) μ ~Uniform(0,z1)

	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
A	1.024349	0.187515	0.005313	0.669103	1.399845	1438.401336	1.000622
L	1.098582	0.243619	0.005859	0.648590	1.585815	2035.866805	1.000786
M	0.454896	0.058247	0.001573	0.339016	0.502424	1480.760463	1.000730




Distribution of α with 10000 samples
Probability Density
0.0
0.5
1.0
1.5
2.0
2.5
0.6
0.8
1.0
1.2
1.4
1.6
1.8
Distribution of λ with 10000 samples
Probability Density
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
Distribution of μ with 10000 samples
Probability Density
0
5
10
15
20
0.0
0.1
0.2
0.3
0.4
0.5

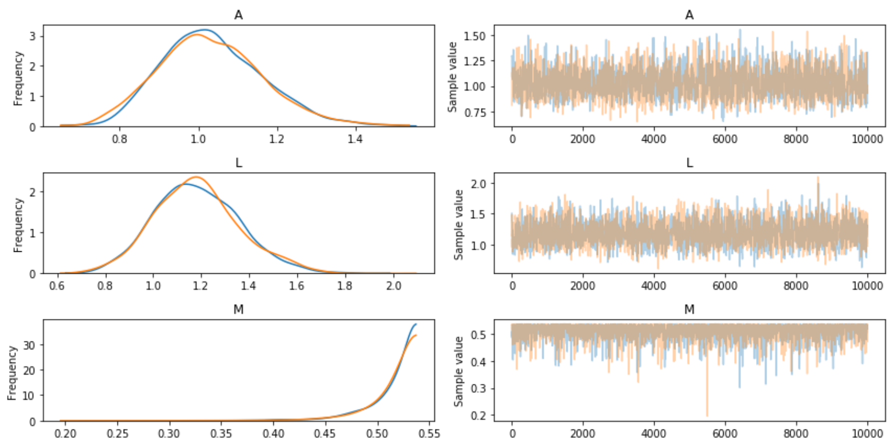


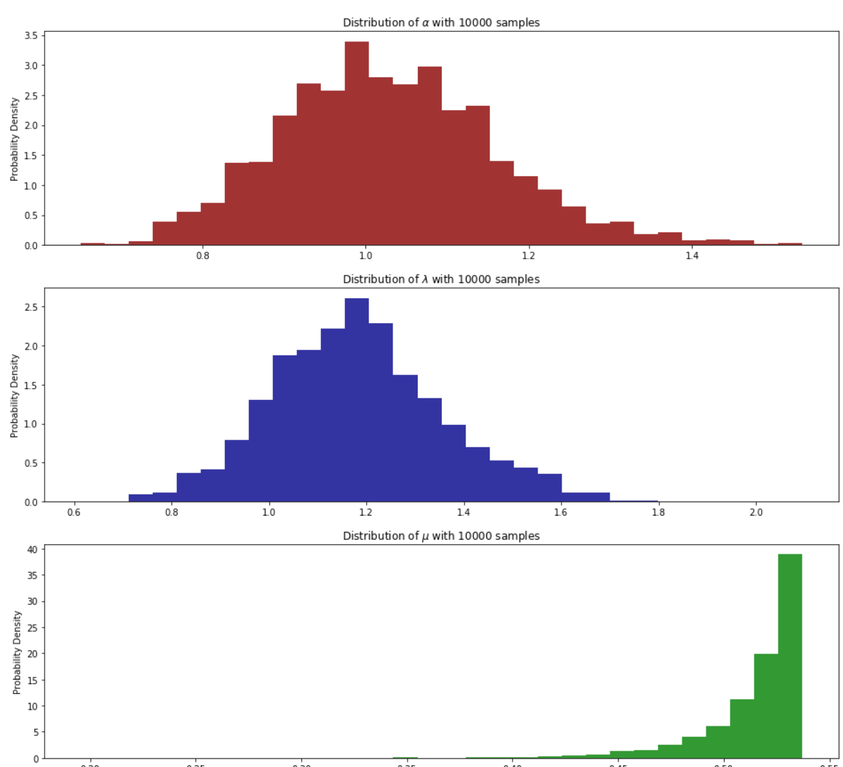
MSE for Alpha is: 0.0346368049430880 25
MSE for Lambda is: 0.07111768259417793
MSE for Mu is: 0.005068672879138695

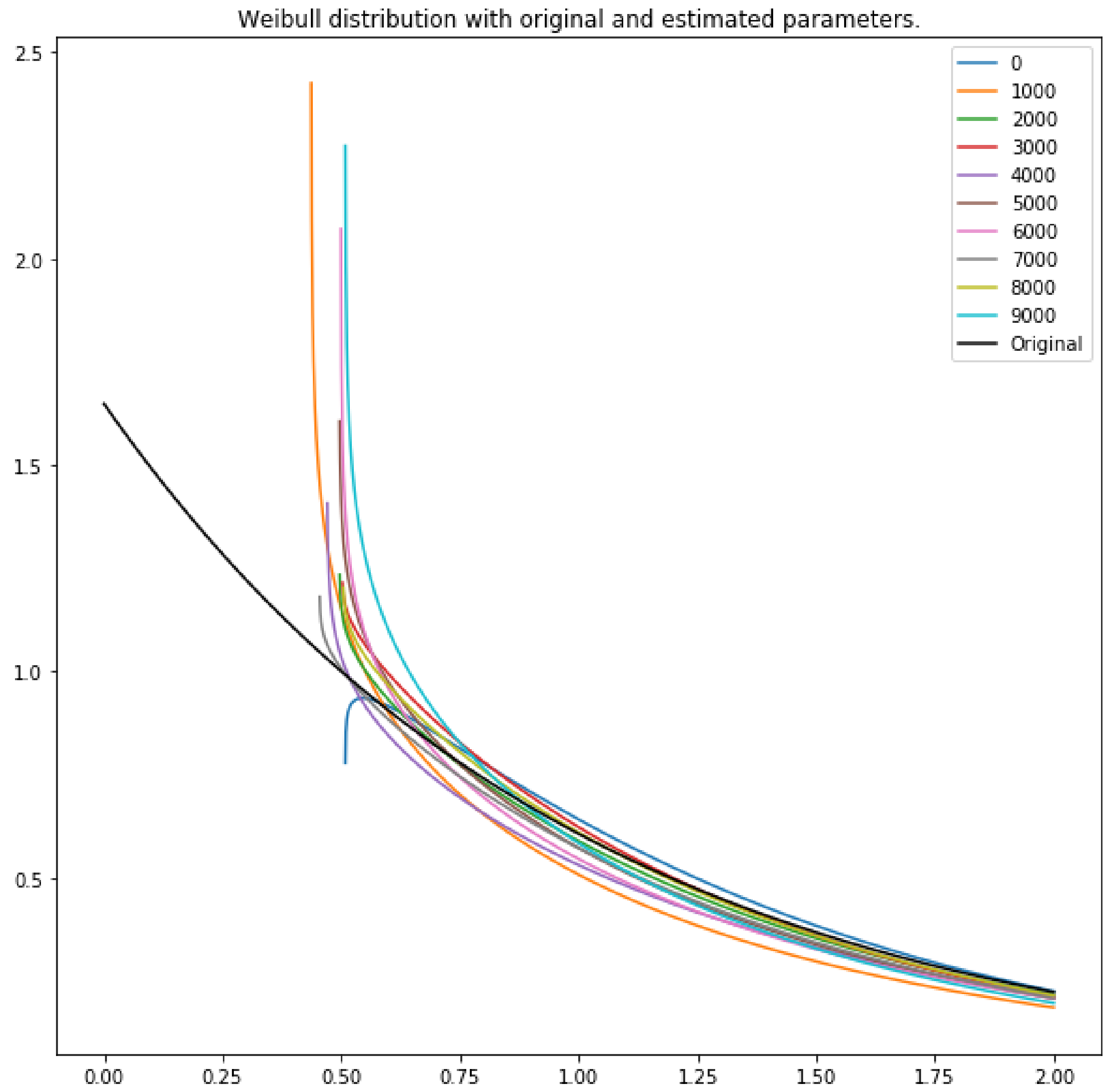
N=50, Alpha=1, Lambda=1, Mu=0.5

priors - λ ~Gamma(1, 1) α ~Gamma(1, 1) μ ~Uniform(0,z1)

	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
A	1.032968	0.129581	0.003020	0.783033	1.282453	1718.248156	1.000743
L	1.183090	0.174727	0.003361	0.831437	1.524481	2238.143134	0.999951
M	0.514889	0.025433	0.000499	0.464587	0.537297	2207.451154	1.000138





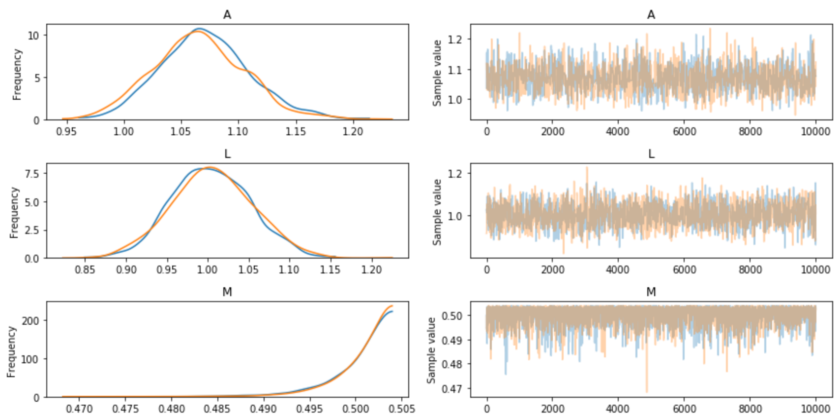


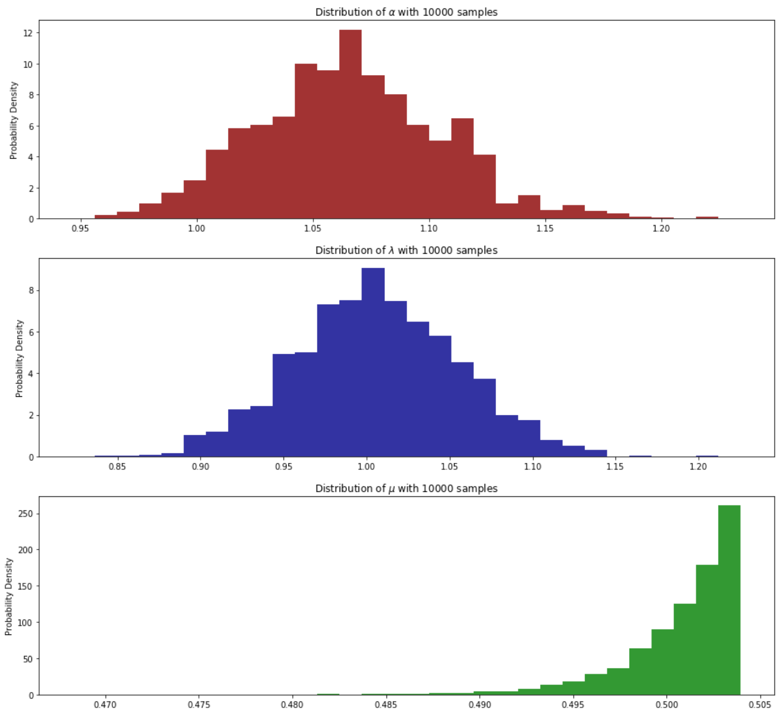
MSE for Alpha is: 0.00286614962153955
MSE for Lambda is: 0.0024795360312085592
MSE for Mu is: 5.698112478449439e-06

N=500, Alpha=1, Lambda=1, Mu=0.5

priors - λ ~Gamma(1, 1) α ~Gamma(1, 1) μ ~Uniform(0,z1)

	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
A	1.068223	0.039533	0.001027	0.993006	1.147914	1361.821208	1.004472
L	1.005229	0.048241	0.001286	0.915218	1.101866	1574.695514	1.000380
M	0.500669	0.003219	0.000063	0.494299	0.503967	2961.822293	1.002158



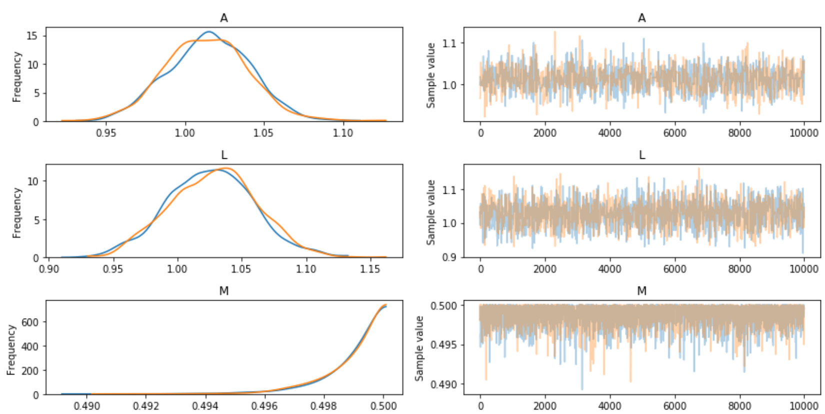


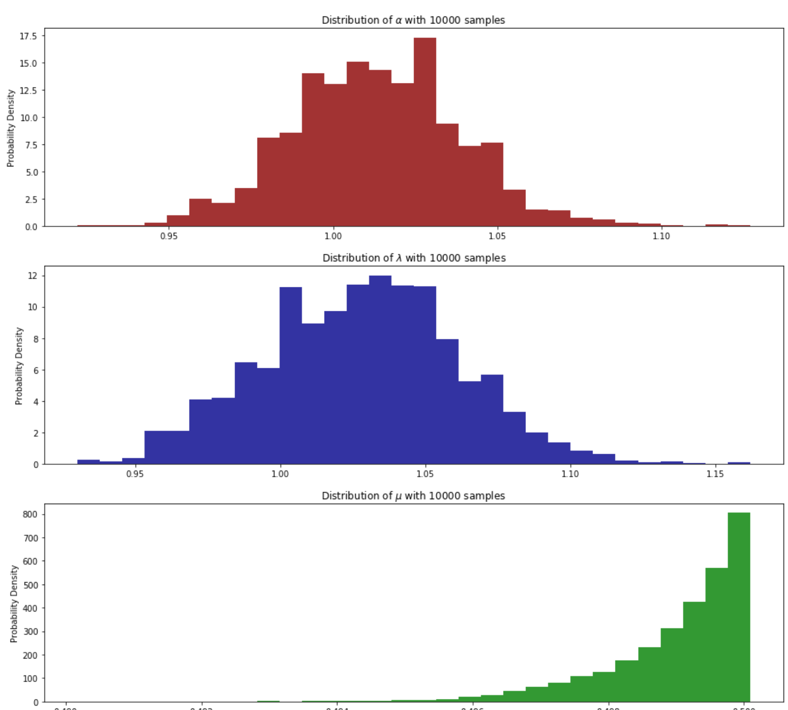
MSE for Alpha is: 0.0013092400121756857
MSE for Lambda is: 0.005922173557944803
MSE for Mu is: 1.4544933343337147e-05

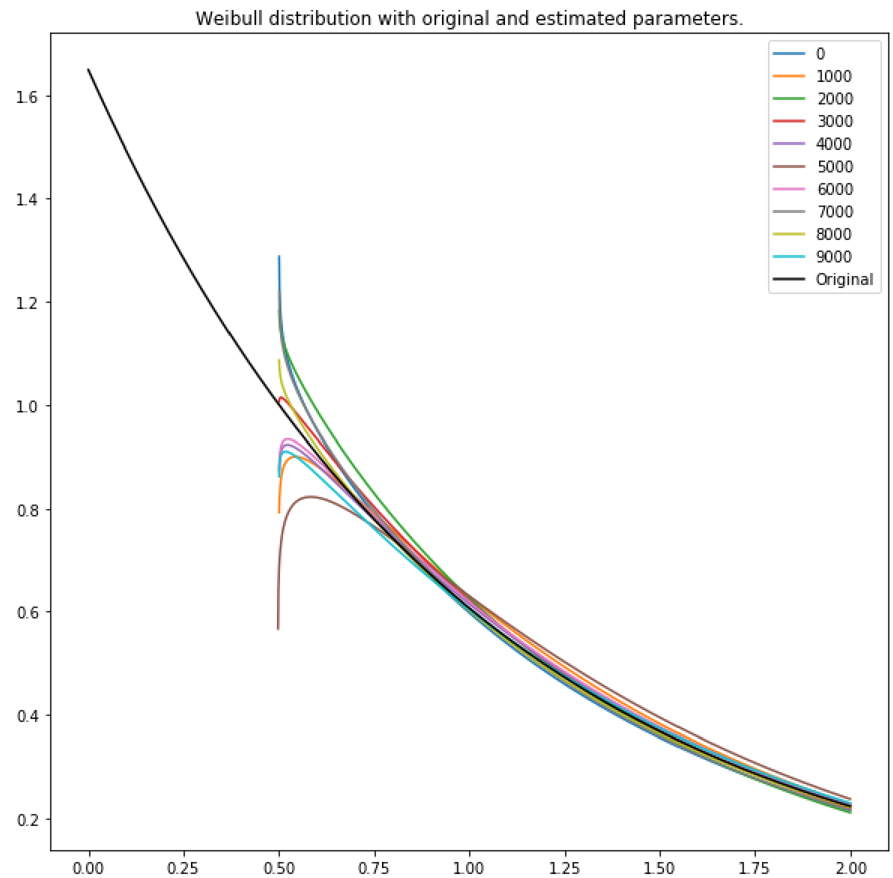
N=1000, Alpha=1, Lambda=1, Mu=0.5

priors - λ ~Gamma(1, 1) α ~Gamma(1, 1)
μ ~Uniform(0,z1)

	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
A	1.014535	0.025851	0.000804	0.958565	1.058926	1009.240908	1.001272
L	1.026761	0.033529	0.000928	0.957254	1.086602	1273.326722	1.001227
M	0.499013	0.001075	0.000019	0.496872	0.500093	3648.075043	0.999961





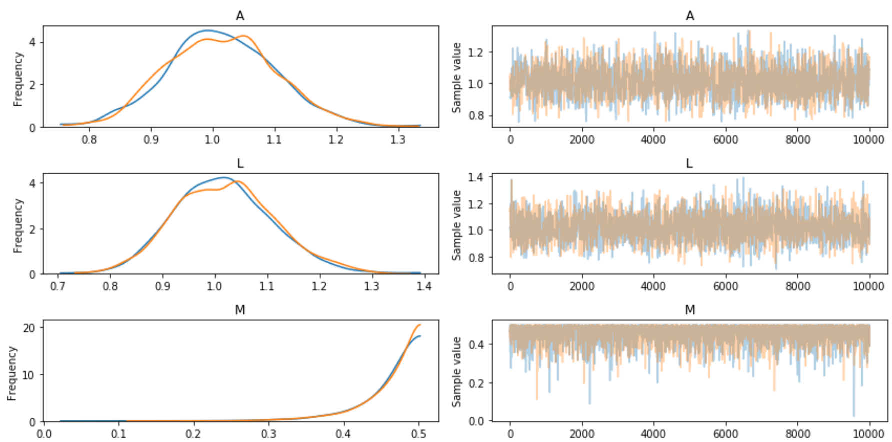


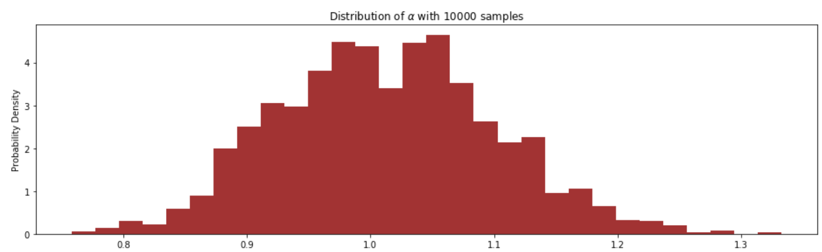
MSE for Alpha is: 0.00286614962153955
MSE for Lambda is: 0.0024795360312085592
MSE for Mu is: 5.698112478449439e-06

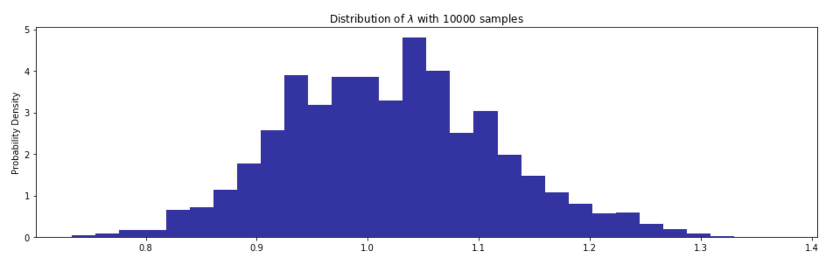
N=25, Alpha=1, Lambda=1, Mu=0.5

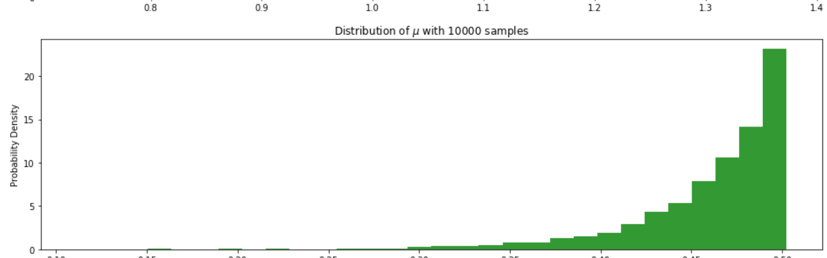
priors - λ ~Gamma(100, 100) α ~Gamma(100, 100) μ ~Uniform(0,z1)

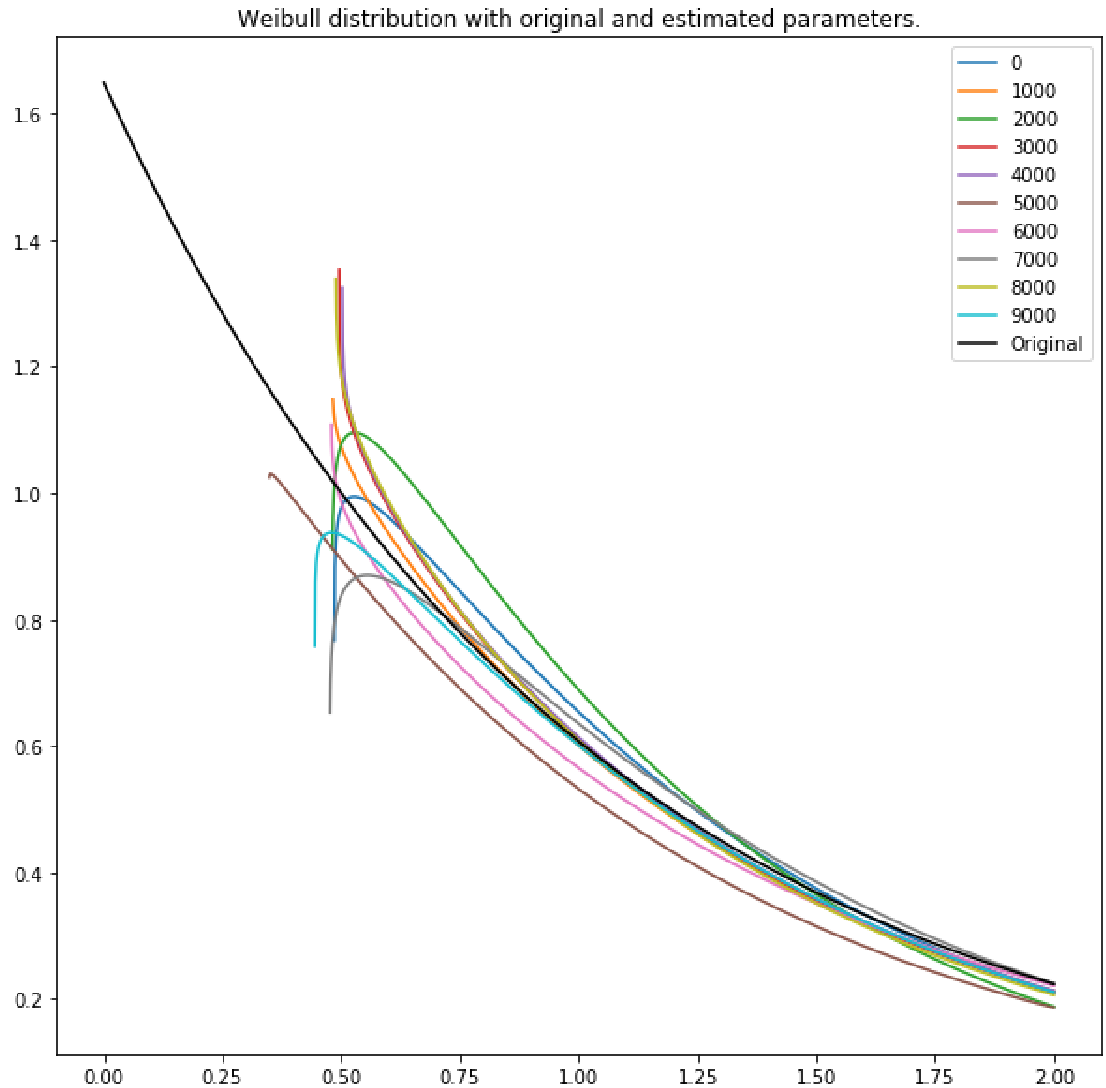
	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
A	1.013801	0.087354	0.001897	0.845811	1.186640	2158.045175	0.999983
L	1.018393	0.094562	0.002003	0.840676	1.210296	2370.352901	1.000495
M	0.461479	0.042637	0.000829	0.374385	0.502420	2646.209373	1.000055









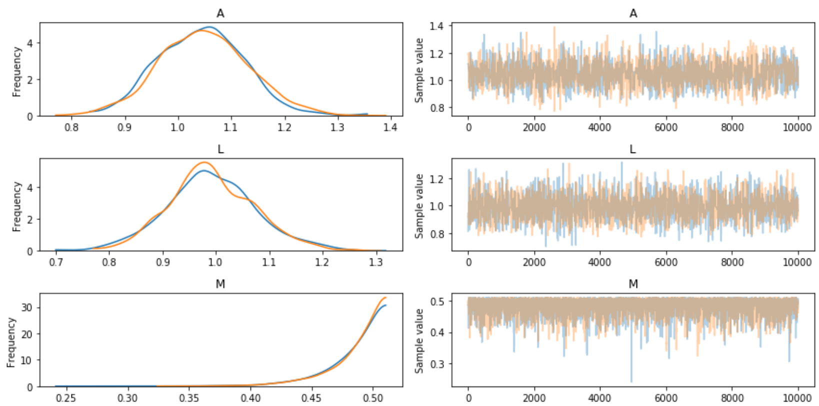


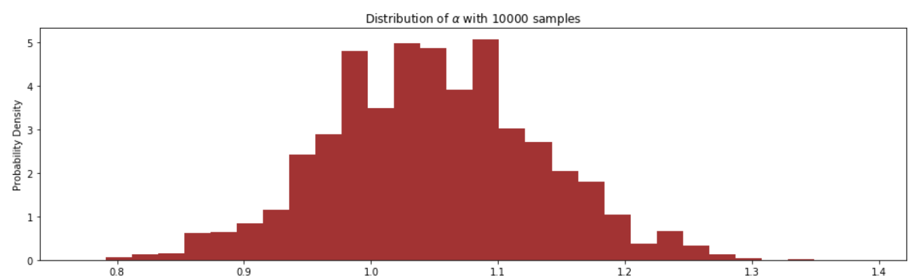
MSE for Alpha is: 0.007892422711518262
MSE for Lambda is: 0.009490949744963468
MSE for Mu is: 0.003275644212439761

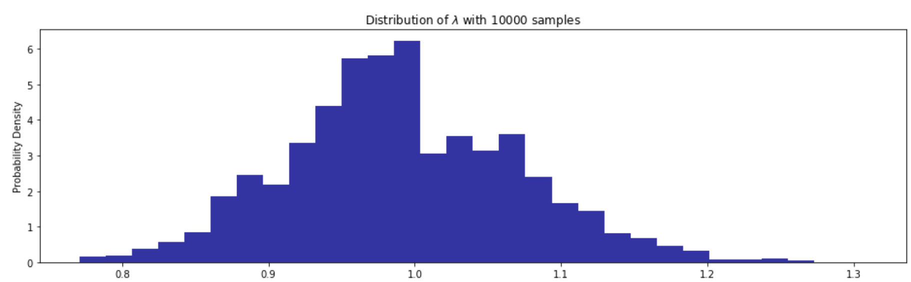
N=50, Alpha=1, Lambda=1, Mu=0.5

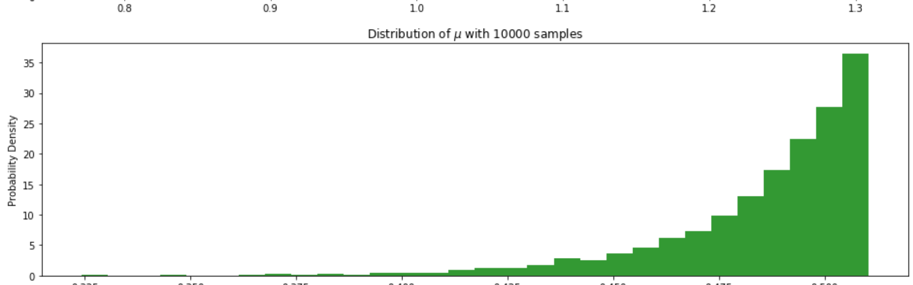
priors - λ ~Gamma(100, 100) α ~Gamma(100, 100) μ ~Uniform(0,z1)

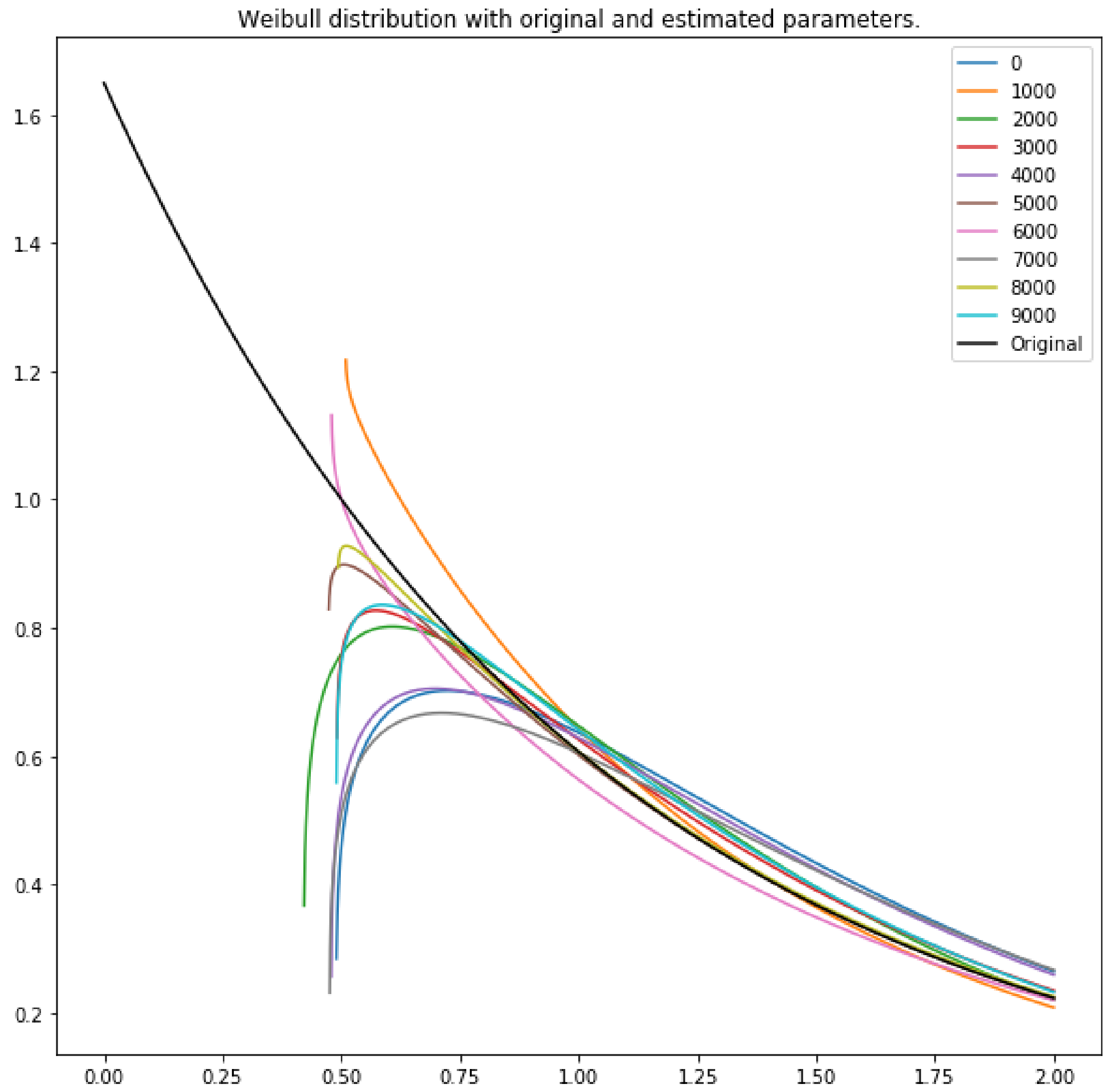
	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
A	1.046149	0.081533	0.001801	0.886892	1.203415	1671.849419	1.001482
L	0.991525	0.081886	0.001533	0.828728	1.159005	2077.692400	0.999950
M	0.485830	0.024501	0.000394	0.436986	0.510463	3253.322812	1.000170









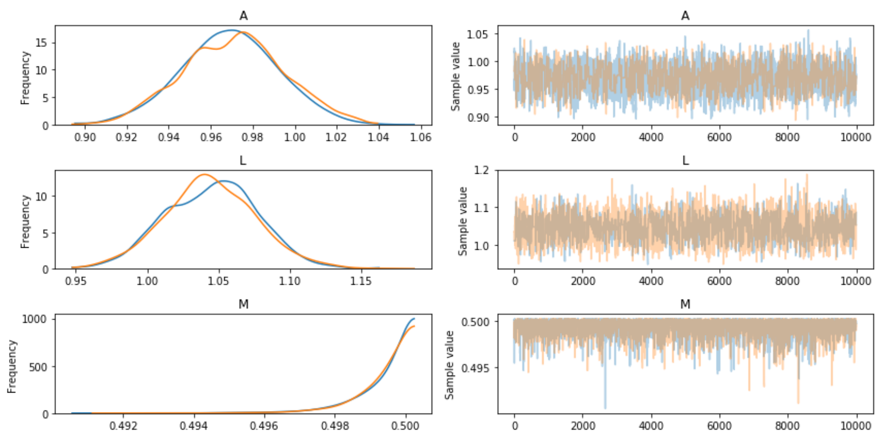


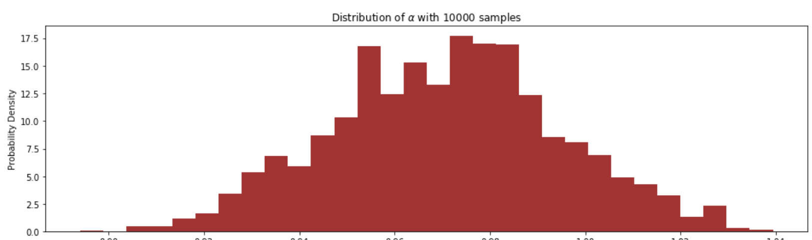
MSE for Alpha is: 0.010586863025173988
MSE for Lambda is: 0.009207450566624927
MSE for Mu is: 0.00036445467929207415

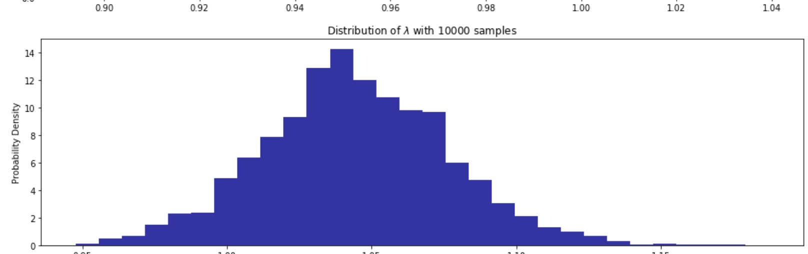
N=1000, Alpha=1, Lambda=1, Mu=0.5

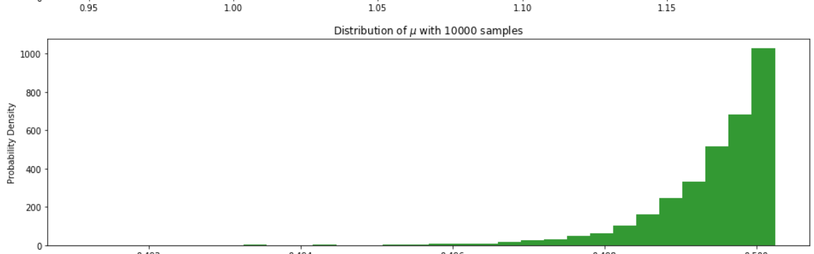
priors - λ ~Gamma(100, 100) α ~Gamma(100, 100) μ ~Uniform(0,z1)

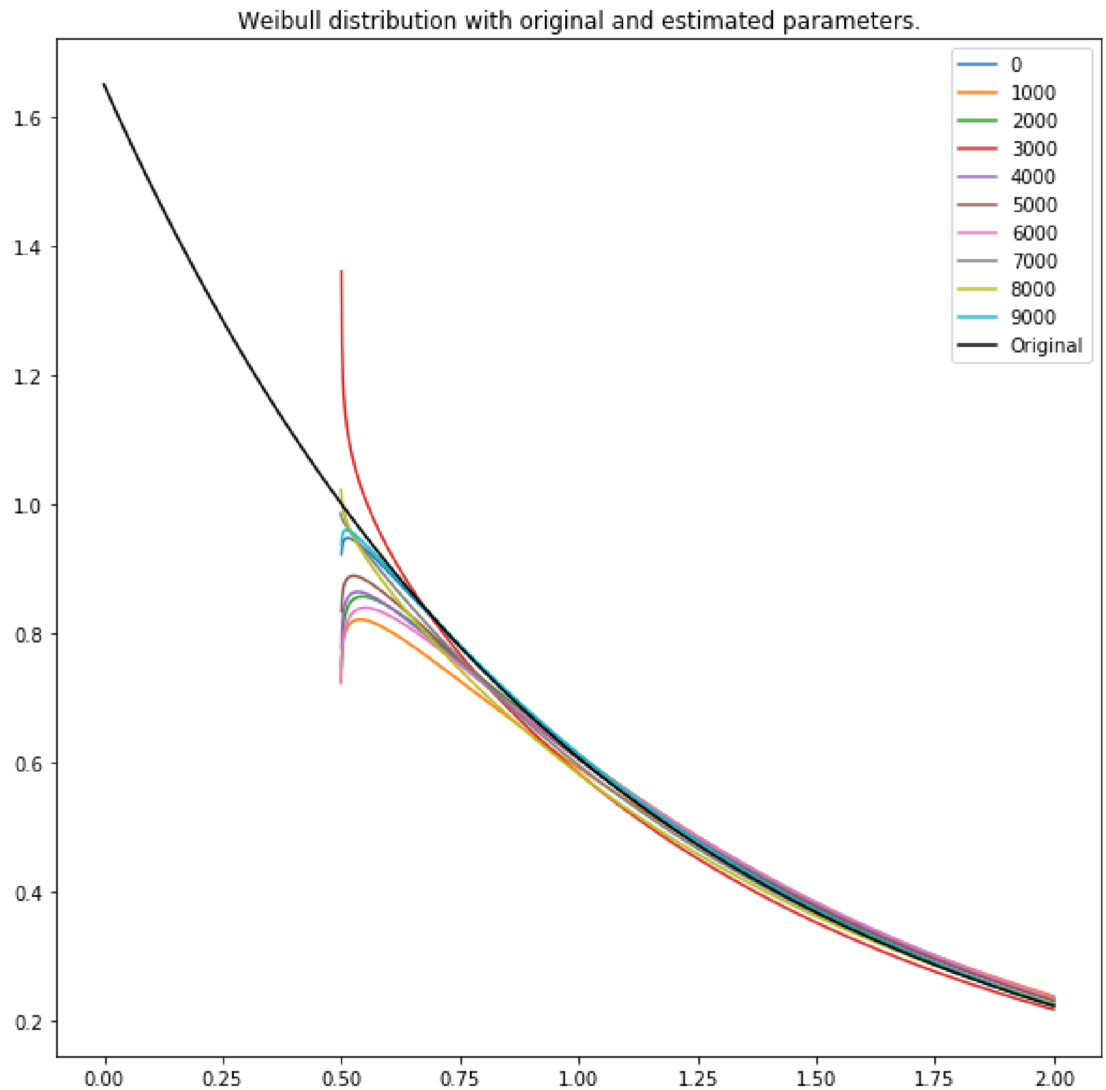
	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
A	0.970064	0.023412	0.000549	0.923616	1.015203	1720.722924	1.001504
L	1.045115	0.032267	0.000760	0.981730	1.106723	1816.337761	1.000118
M	0.499440	0.000844	0.000014	0.497803	0.500248	3641.704624	1.000084









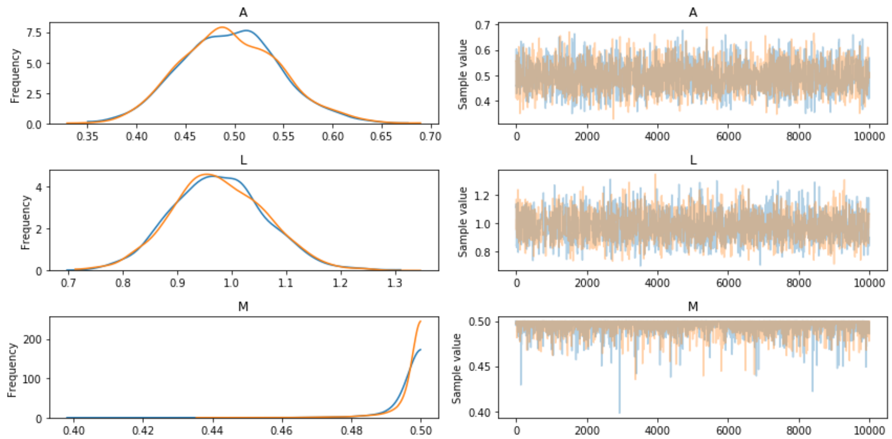


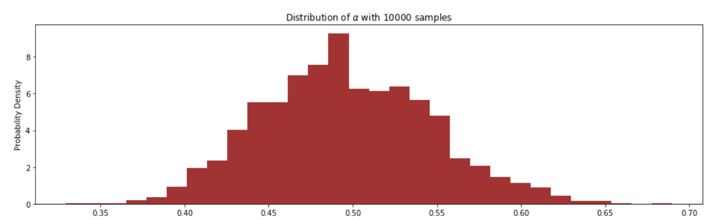
MSE for Alpha is: 0.0007065086825127311
MSE for Lambda is: 0.0024411372007294
MSE for Mu is: 2.4865561346760305e-06

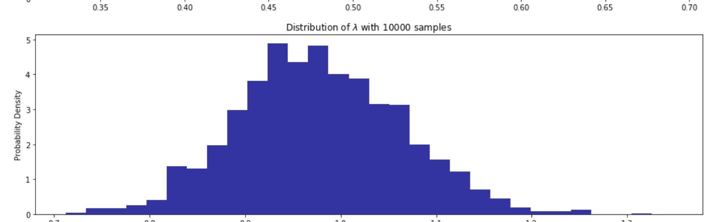
N=25, Alpha=0.5, Lambda=1, Mu=0.5

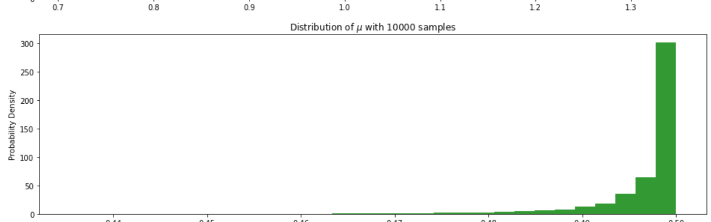
priors - λ ~Gamma(100, 100) α ~Gamma(50, 100) μ ~Uniform(0,z1)

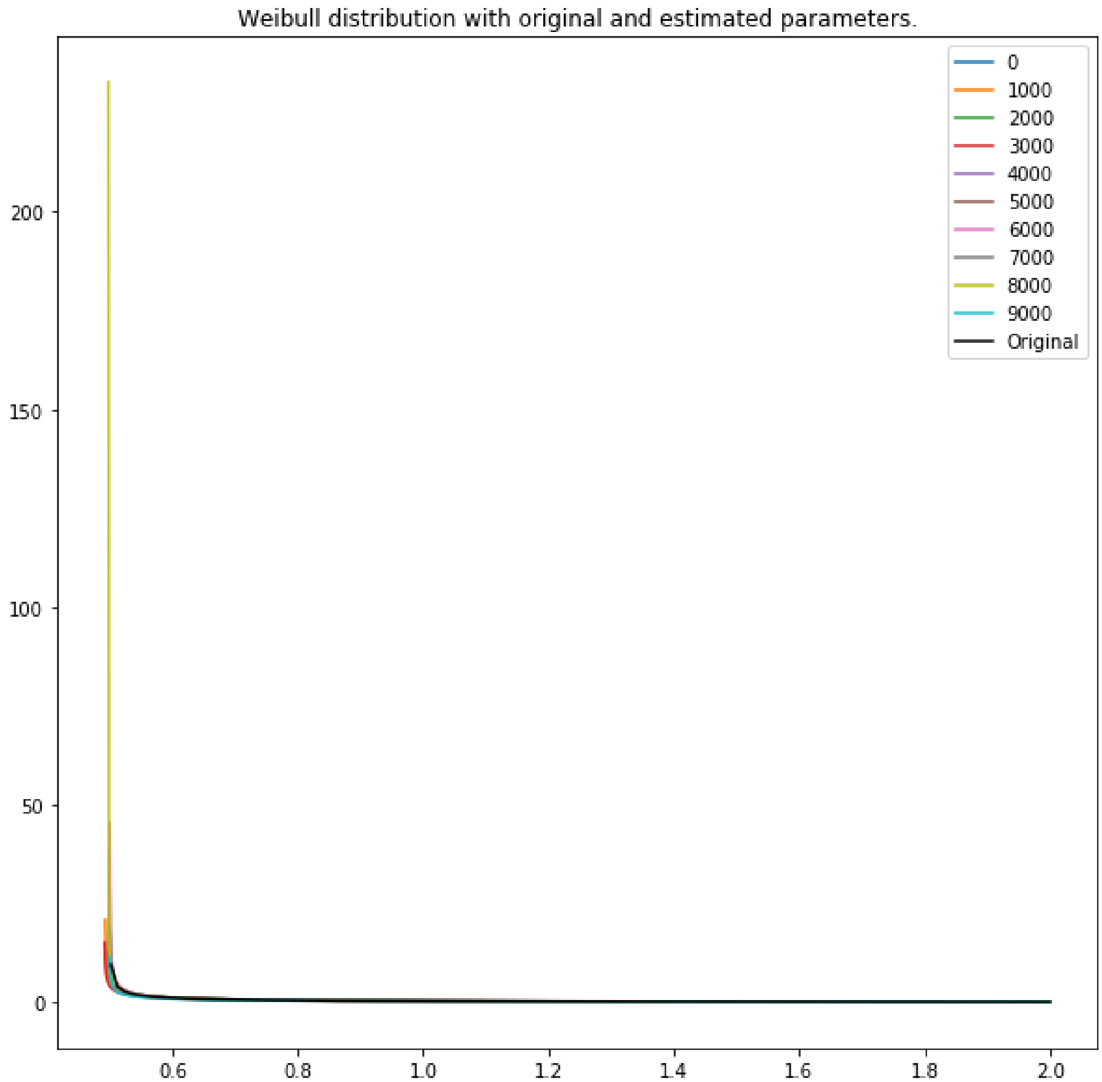
	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
A	0.495806	0.050328	0.001079	0.401477	0.597502	2260.816092	1.000044
L	0.977214	0.085774	0.001747	0.818970	1.148230	2348.187202	0.999983
M	0.496786	0.005598	0.000073	0.486276	0.500000	4602.798042	1.000088









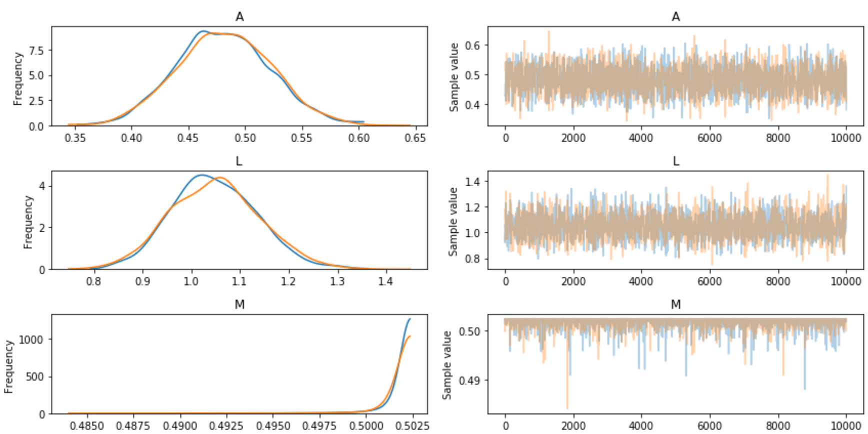


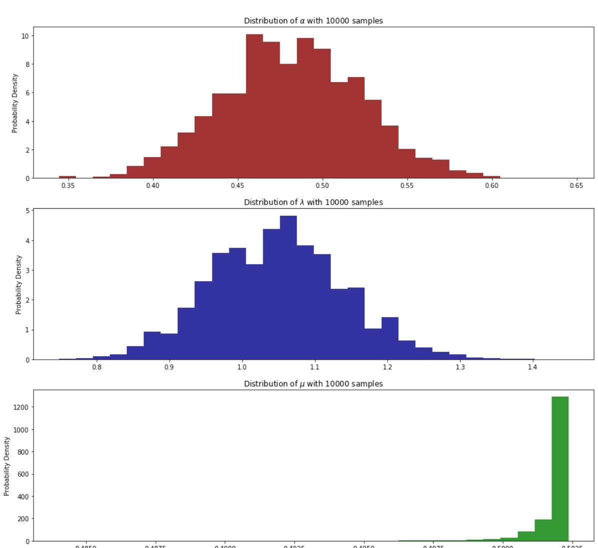
MSE for Alpha is: 0.25629753718499404
MSE for Lambda is: 0.007870020773745781
MSE for Mu is: 3.963593111729778e-05

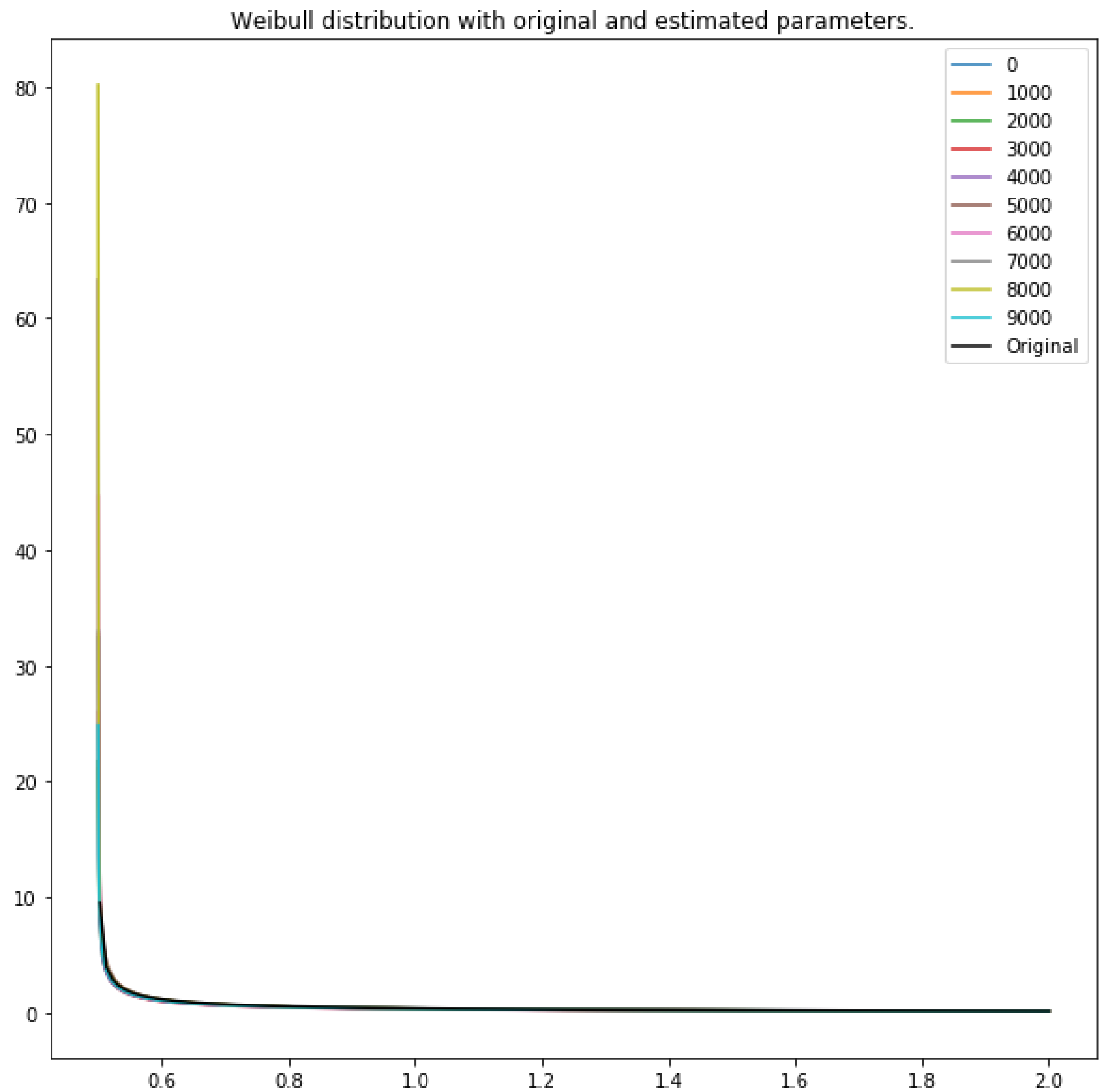
N=50, Alpha=0.5, Lambda=1, Mu=0.5

priors - λ ~Gamma(100, 100) α ~Gamma(50, 100) μ ~Uniform(0,z1)

	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
A	0.481224	0.040818	0.000836	0.402080	0.559746	2324.240240	1.000317
L	1.048928	0.089005	0.001835	0.879380	1.223715	2109.218711	1.000010
M	0.501924	0.000850	0.000013	0.500566	0.502366	4485.437337	0.999950





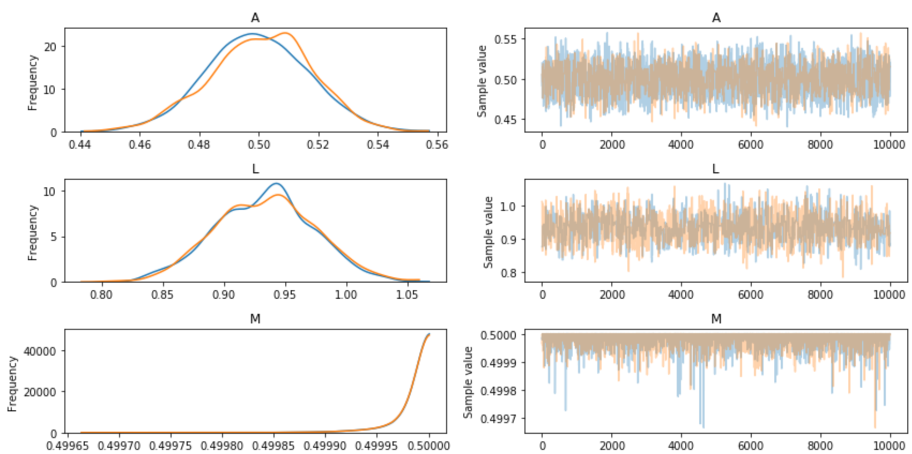


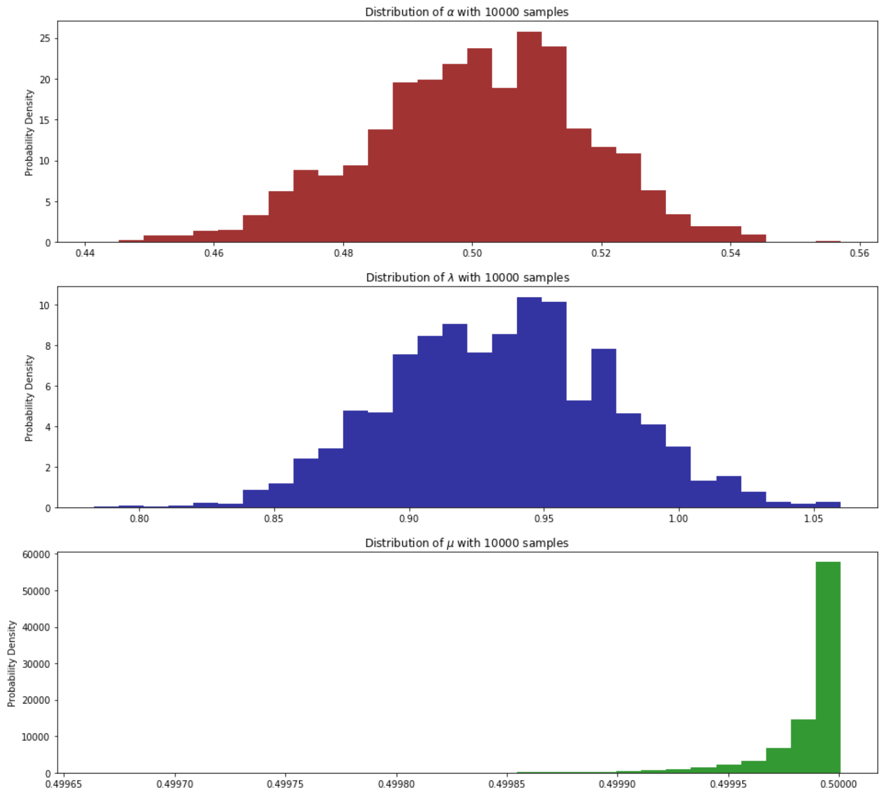
MSE for Alpha is: 0.21289034170535656
MSE for Lambda is: 0.008294030003785563
MSE for Mu is: 3.4263390781396535e-06

N=500, Alpha=0.5, Lambda=1, Mu=0.5

priors - λ ~Gamma(100, 100) α ~Gamma(50, 100) μ ~Uniform(0,z1)

	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
A	0.500156	0.016892	3.916714e-04	0.466120	0.531470	1863.905892	1.001149
L	0.933766	0.040247	1.161486e-03	0.854452	1.010333	1053.365078	1.000164
M	0.499987	0.000022	2.984752e-07	0.499944	0.500001	5100.058296	0.999950



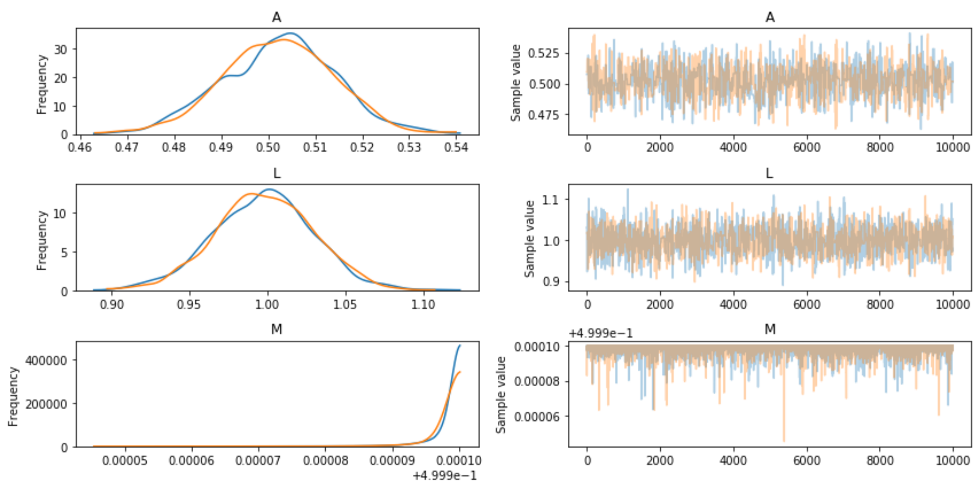


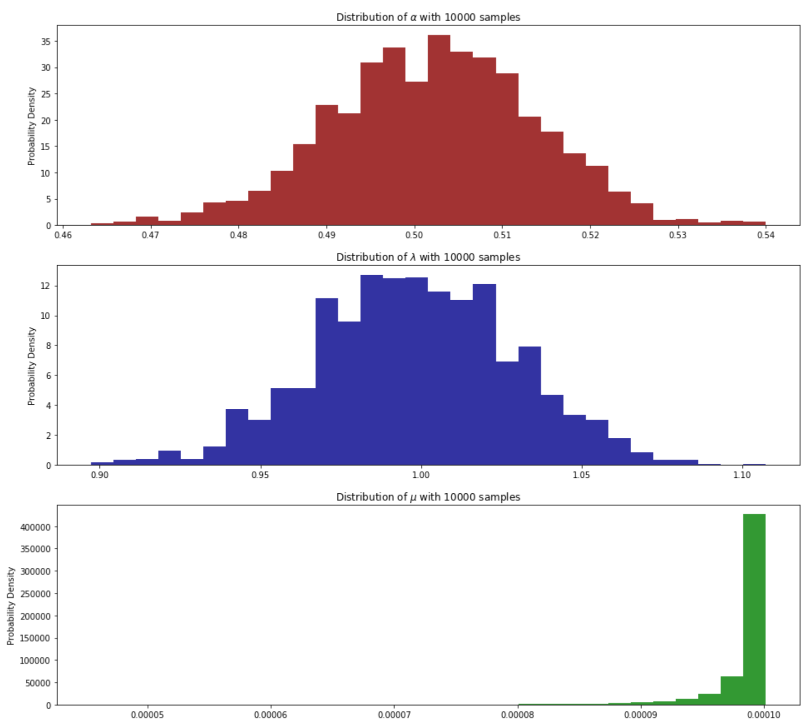
MSE for Alpha is: 0.28983542125003703
MSE for Lambda is: 0.003742219155835908
MSE for Mu is: 5.689966689048724e-13

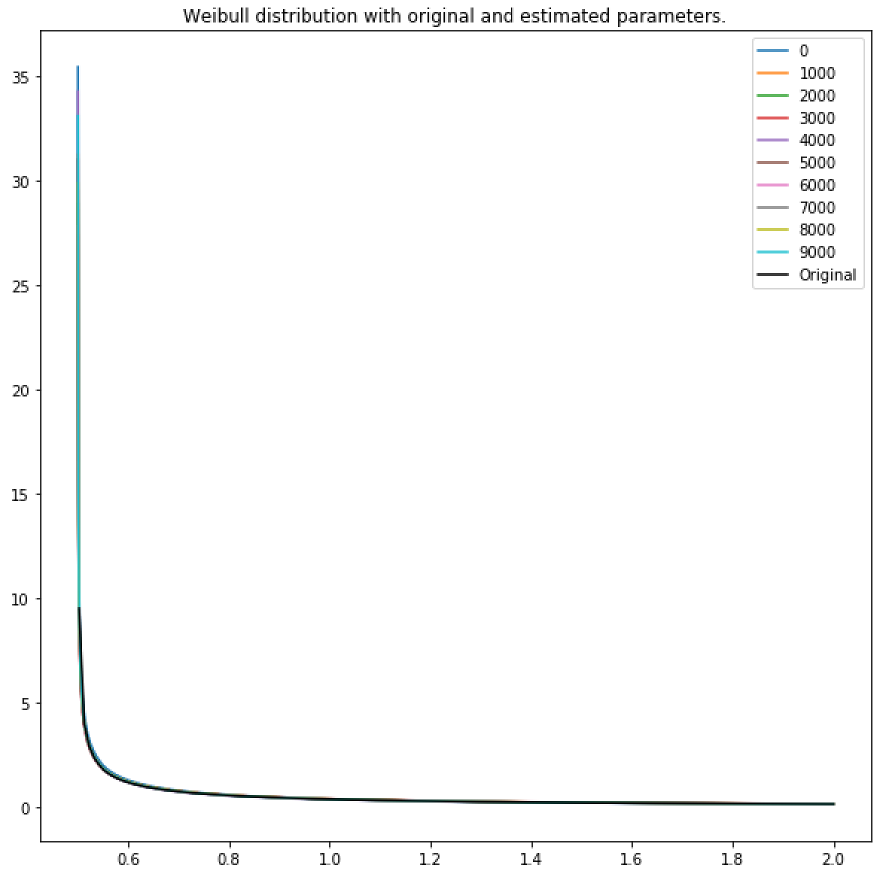
N=1000, Alpha=0.5, Lambda=1, Mu=0.5

priors - λ ~Gamma(100, 100) α ~Gamma(50, 100) μ ~Uniform(0,z1)

	mean	sd	mc_error	hpd_2.5	hpd_97.5	n_eff	Rhat
A	0.502256	0.011774	3.745509e-04	0.478174	0.524026	938.243835	1.000030
L	0.997896	0.031057	8.598886e-04	0.935464	1.055710	1399.044679	1.000117
M	0.499999	0.000003	3.894835e-08	0.499994	0.500000	4287.339101	1.000056







MSE for Alpha is: 0.2541403216700826
MSE for Lambda is: 0.0024025923581422123
MSE for Mu is: 2.1455812668016285e-11

Log concave posterior of alpha with gamma prior

The posterior distribution of parameters is proportional to the likelihood function times the priors of alpha and gamma.

$P(\alpha, \lambda | X) \propto P(X | \alpha, \lambda) P(\lambda) P(\alpha)$

$\lambda \sim Gamma(a,b)$ $\alpha \sim Gamma(c,d)$

$$P(\alpha, \lambda | X) \propto \frac{\lambda^{a+n-1} \, exp(-\lambda(b+\Sigma Xi^{\alpha})) \, \alpha^{c+n-1} \Pi Xi^{\alpha-1} \, exp(-d\alpha) \, b^{a} d^{c}}{\Gamma a \, \Gamma c}$$

$$\propto \frac{Gamma(a+n, \, b+\Sigma Xi^{\alpha}) \, \Pi(Xi^{\alpha-1}) \, \alpha^{c+n-1} \, exp(-d\alpha)}{(b+\Sigma Xi^{\alpha})^{a+n}}$$

$$P(\lambda | \alpha, X) = Gamma(a+n, \; b+\Sigma Xi^{\alpha})$$

$$P(\alpha | X) = K \frac{\Pi(Xi)^{\alpha-1} e^{-d\alpha} \alpha^{c+n-1}}{(b+\Sigma Xi^{\alpha})^{a+n}}$$ log-concave pdf of α

The shape of the posterior density function closely resembles that of the gamma density function, so we approximate the posterior density function of α with the gamma density function by equating the first two moments.

$\alpha | X \sim Gamma(p, q)$ $\frac{p}{q} = \int \alpha P(\alpha | X) \, d\alpha$ $\frac{p^2+p}{q} = \int \alpha^2 P(\alpha | X) \, d\alpha$

We have normalized the integration by multiplying with appropriate constants to prevent the problem of exploding values while estimation of parameters.

Algorithm for calculating Bayes estimate

Step 1: Sample a random sample (X) of a weibull distribution.

Step 2: Generate α1 from the log concave density P(α|X) ≃ Gamma(p,q)

Step 3: Generate λ1 from P(λ| α,X) using α1 generated in step 2.

Step 4: Repeat steps 2 and 3 M times and sort the values so that

$\alpha_{(1)} < \alpha_{(2)} \ldots. < \alpha_{(M)}$ and $\lambda_{(1)} < \lambda_{(1)} \ldots. < \lambda_{(M)}$.

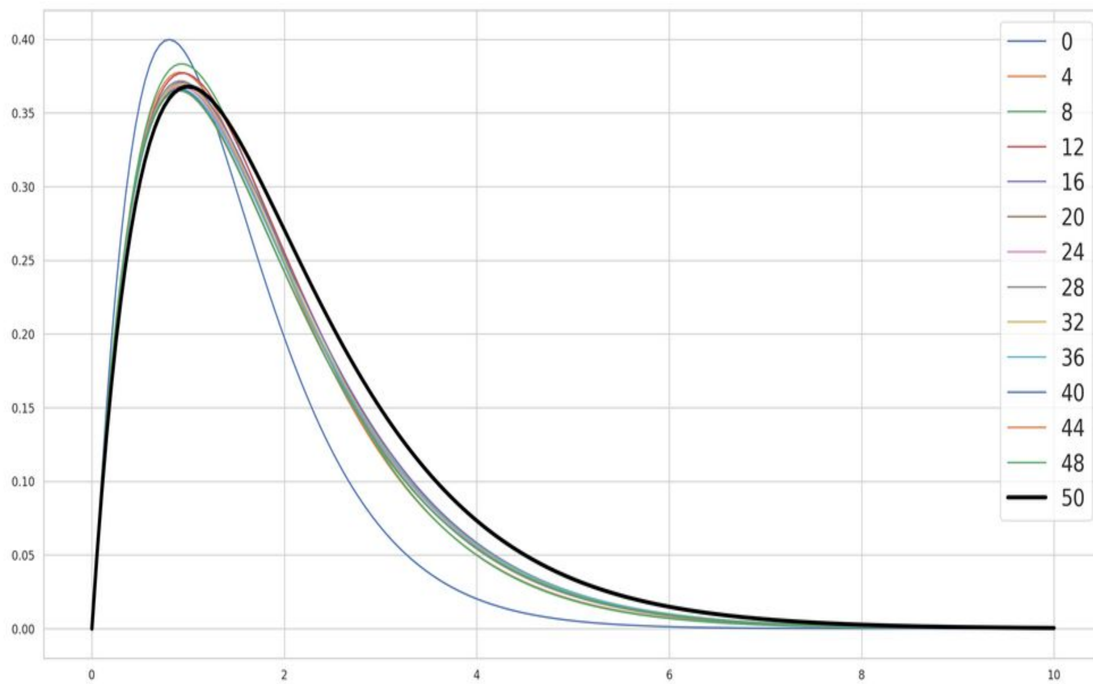
Step 4: The 100(1-2β)% symmetric confidence interval can be calculated as

$\alpha_{[M(1-\beta)]}$) and ($\lambda_{[M\beta]}, \lambda_{[M(1-\beta)]}$)

N=25 priors - λ ~Gamma(50, 50) α ~Gamma(50, 50)

Parameter : alpha
Mean : 1.0040254418789316
MSE : 0.012980101594989066
Credible interval : (0.6699993079787578, 1.0034025851377872)

Parameter : lambda
Mean : 1.1069636285211218
MSE : 0.038653317638978286
Credible interval : (0.8994571299774963, 1.4158608079355923)



N=40 priors - λ ~Gamma(1, 1) α ~Gamma(1, 1)

Parameter : alpha
Mean : 1.0296149445866158
MSE : 0.01681947697451171
Credible interval : (0.7606575710524227, 1.2542518312983209)

Parameter : lambda
Mean : 0.9789656731294296
MSE : 0.026789125678845463
Credible interval : (0.7138883029373431, 1.3605675717636805)

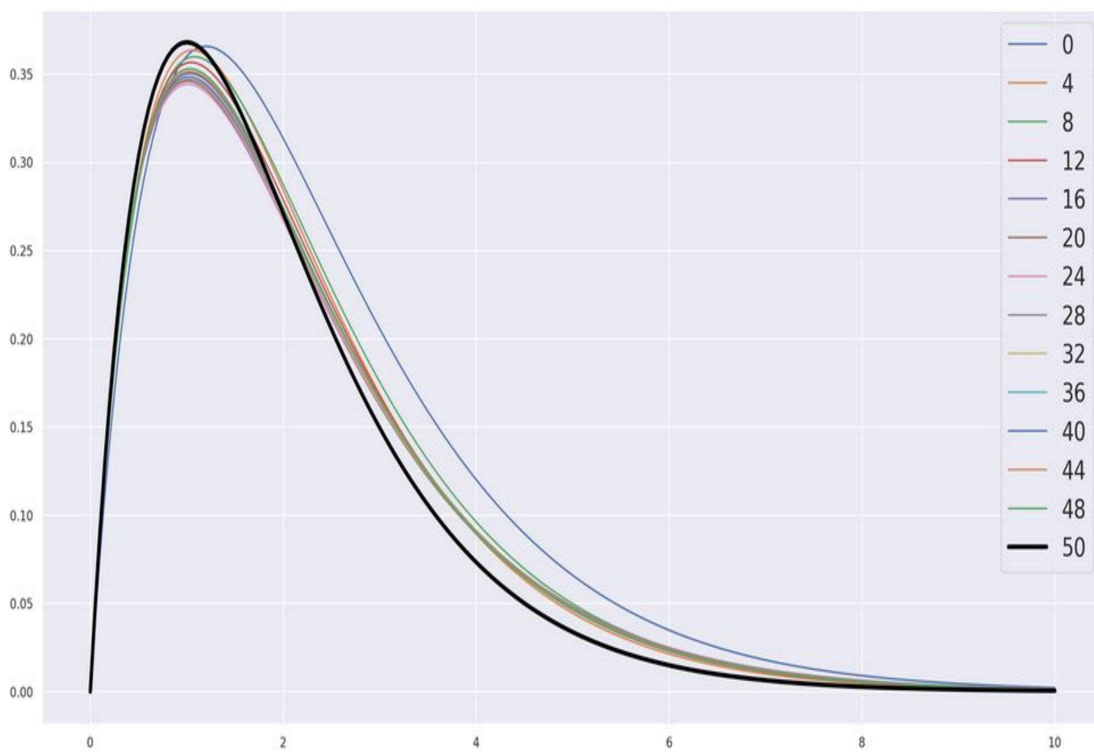


Fig: convergence

N=40 priors - λ ~Gamma(50, 50) α ~Gamma(50, 50)

Parameter : alpha
Mean : 0.9586332951581324
MSE : 0.007334818220872868 5
Credible interval : (0.8201215077291094, 1.176416299892396)

Parameter : lambda
Mean : 0.9696183174311593
MSE : 0.010779892707561039
Credible interval : (0.7785652878328362, 1.224936752811868)

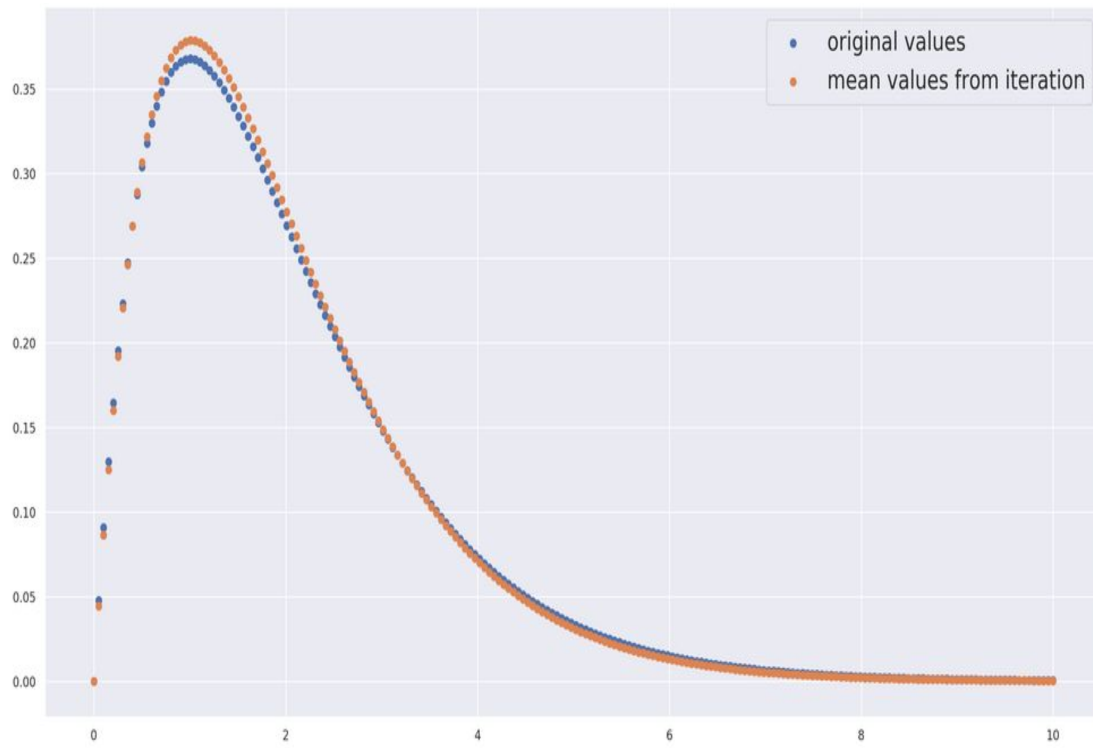
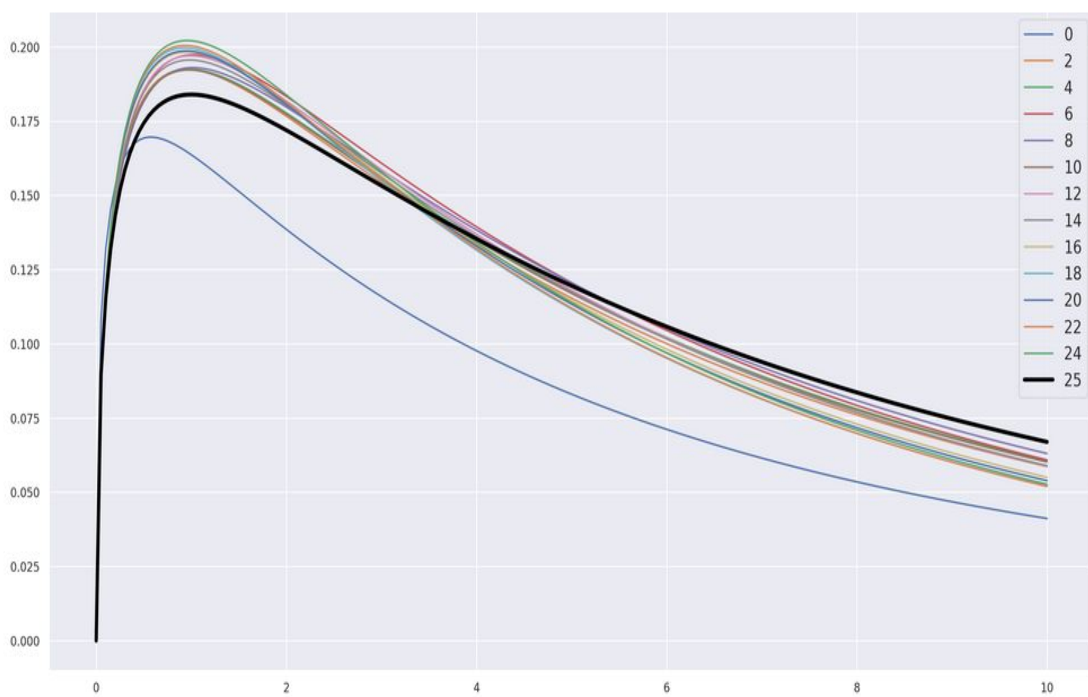


Fig: comparison of weibull distribution using 40 data points of the original values vs the final estimate for alpha = 1 lambda = 1

N=25 priors - λ ~Gamma(1, 1) α ~Gamma(1, 2)

Parameter : alpha
Mean : 0.5494484222921936
MSE : 0.008680547320258556
Credible interval : (0.48440098317202557, 0.9949250071180015)

Parameter : lambda
Mean : 1.0259753094401263
MSE : 0.03219408610445811
Credible interval : (0.7322559006364312, 1.6927029895389192)

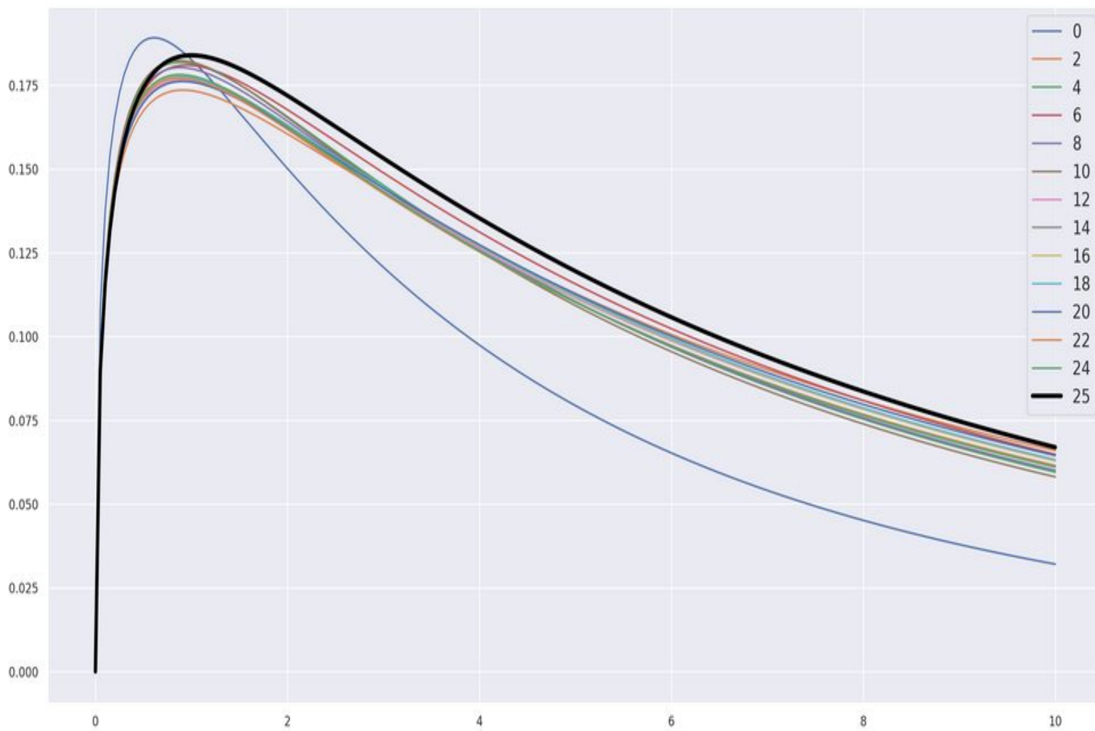


Observation - Variance of X ~ Gamma(p,q) ∝ $1/q^2$, therefore the variance can be reduced drastically while preserving the mean to improve the estimates.

N=25 priors - λ ~Gamma(50, 50) α ~Gamma(50, 100)

Parameter : alpha
Mean : 0.4907039782680261
MSE : 0.002454340241969291
Credible interval : (0.5023109251361051, 0.7768910639178358)

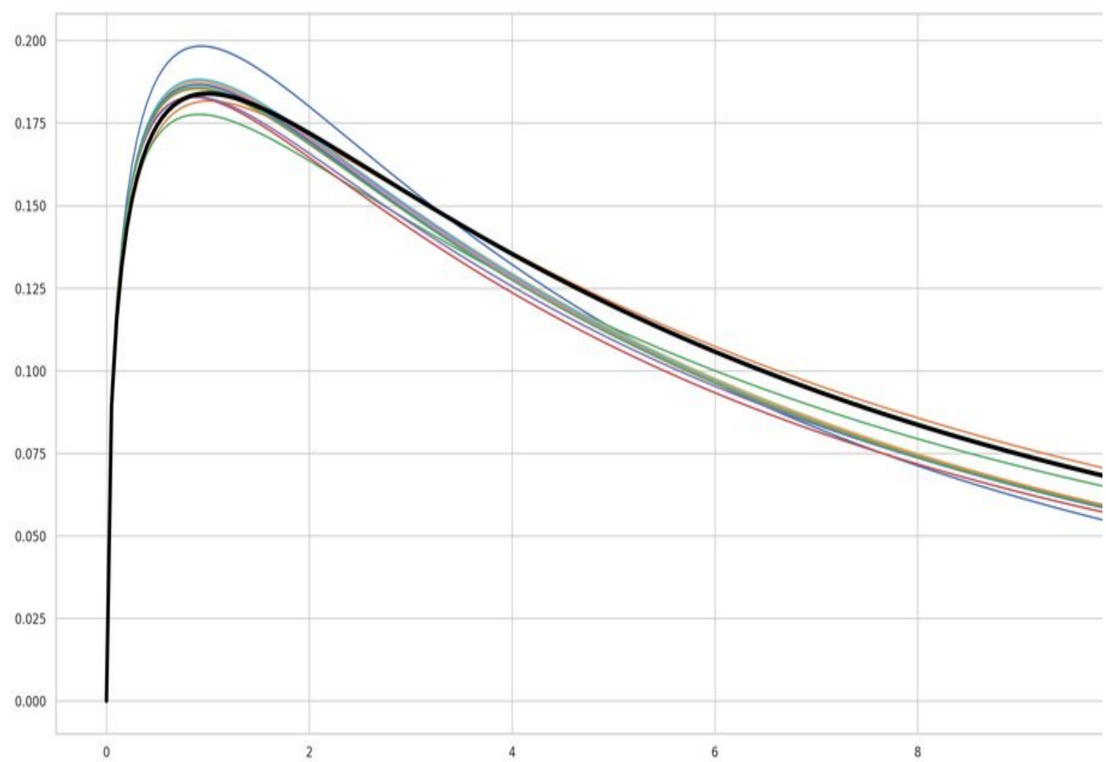
Parameter : lambda
Mean : 0.8801591678257367
MSE : 0.025339099790439742
Credible interval : (0.8291381009934877, 1.2866497083559905)



N=40 priors - λ ~Gamma(50, 50) α ~Gamma(50, 100)

Parameter : alpha
Mean : 0.4977446883462632
MSE : 0.0017444729847580377
Credible interval : (0.5696220777422055, 0.849876261246199)

Parameter : lambda
Mean : 1.015265285309922
MSE : 0.006843594079324681
Credible interval : (0.7828409349129712, 1.1956024856318899)



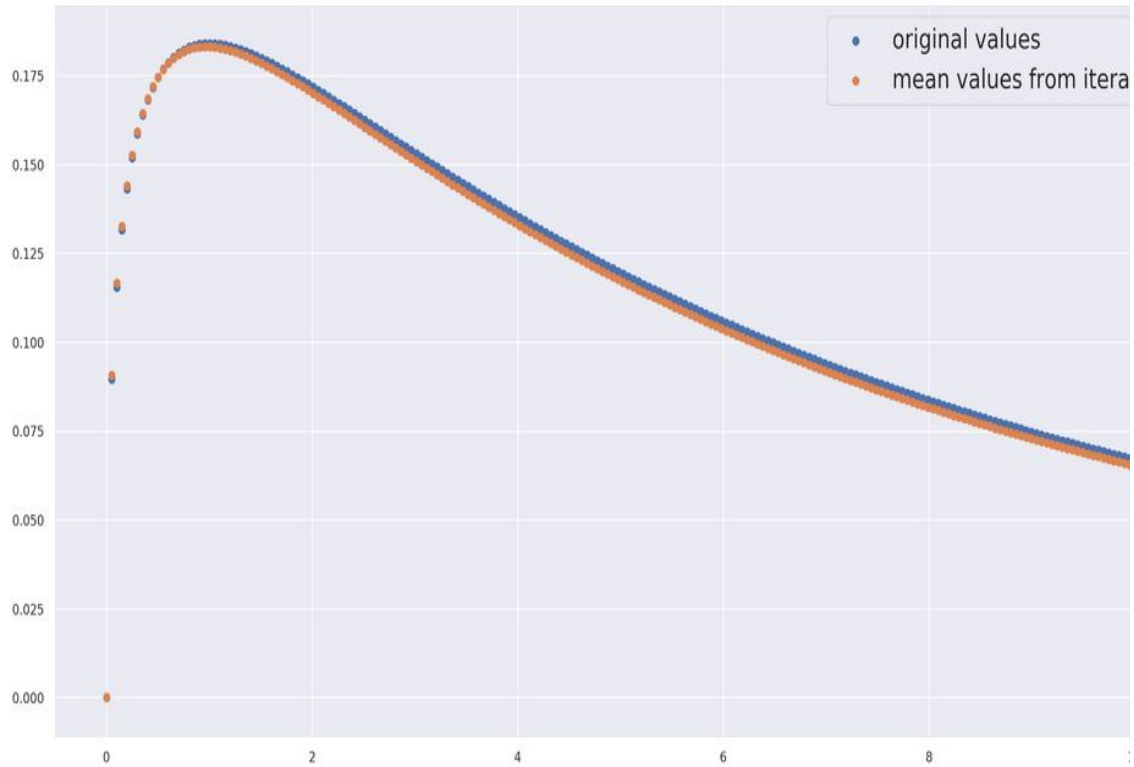


Fig: comparison of weibull distribution using 40 data points of the original values vs the final estimate for alpha = 0.5 lambda = 1

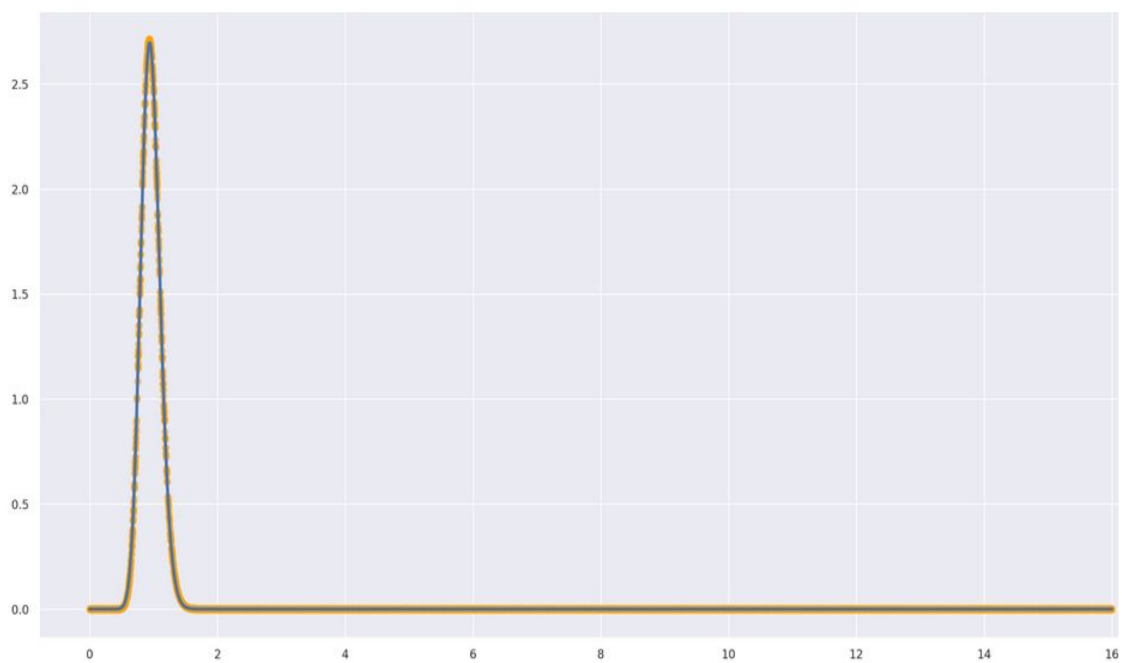


Fig: Gamma(p,q) closely resembles P(α|X)

Log concave posterior of alpha with Uniform prior

$P(\alpha, \lambda \mid X) \propto P(X \mid \alpha, \lambda) P(\lambda) P(\alpha)$

$\lambda \sim U(0,u1)$ $\alpha \sim U(0,u2)$

$$P(\alpha, \lambda|X) \propto \frac{\lambda^n \, exp(-\lambda(\Sigma Xi^\alpha)) \, \alpha^n \, \Pi Xi^{\alpha-1}}{u1u2}$$

$$\propto \frac{Gamma(a+n,\, b+\Sigma Xi^\alpha) \, \Pi(Xi^{\alpha-1}) \, exp(-d\alpha)}{(b+\Sigma Xi^\alpha)^{a+n}}$$

$$P(\lambda|\alpha, X) = Gamma(1+n,\ \Sigma Xi^\alpha)$$

$$P(\alpha|X) = K \frac{\Pi(Xi)^{\alpha-1} \, \alpha^n}{(\Sigma Xi^\alpha)^n}$$ log-concave pdf of α

The shape of the posterior density function closely resembles that of the gamma density function, so we approximate the posterior density function of α with the gamma density function by equating the first two moments.

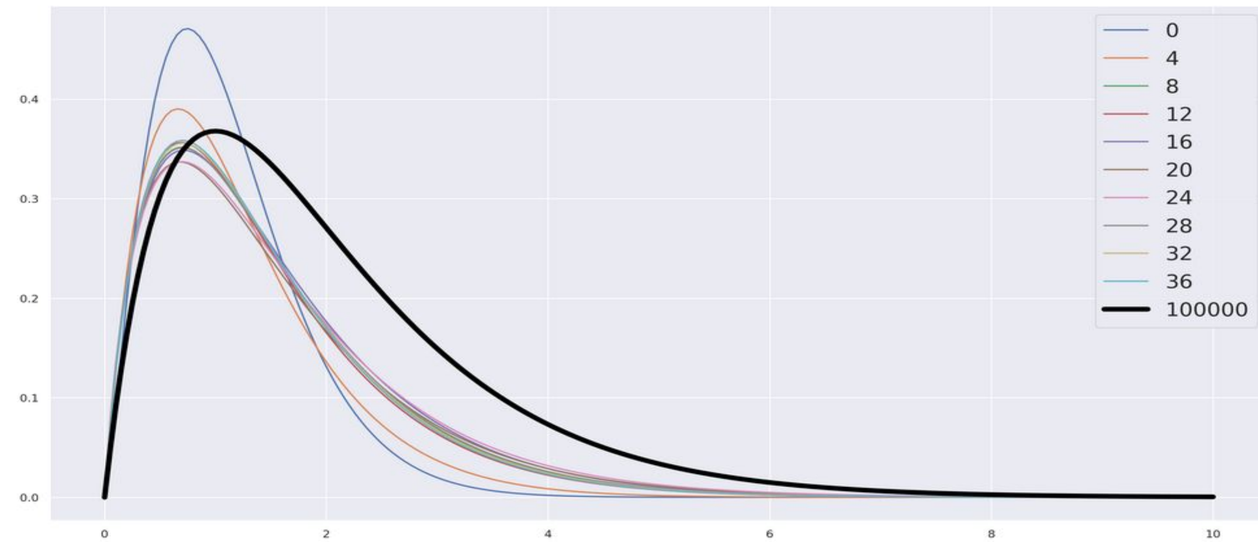
$$\alpha|X \sim Gamma(p, q)$$

$$\frac{p}{q} = \int \alpha P(\alpha|X)\, d\alpha \qquad \frac{p^2+p}{q} = \int \alpha^2 P(\alpha|X)\, d\alpha$$

N=25 priors - λ ~U(0,2) α ~U(0,2)

Parameter : alpha
Mean : 0.9563930370624473
MSE : 0.024012157019455724
Credible interval : (1.0611994192235505, 1.9286083419780182)

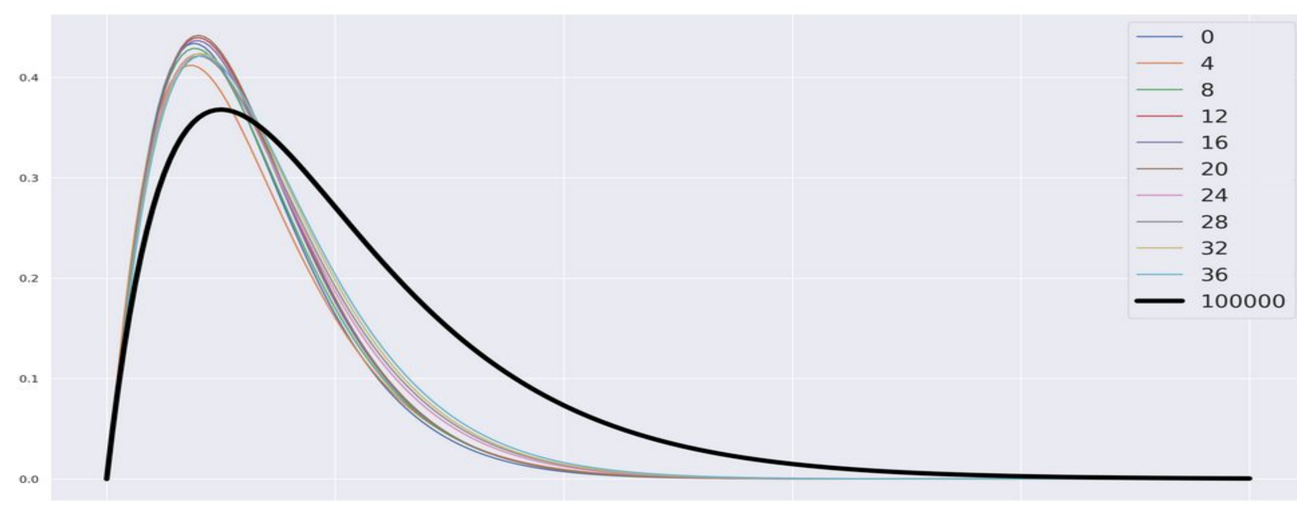
Parameter : lambda
Mean : 1.4885098791633855
MSE : 0.3245235533000572
Credible interval : (0.705231080312938, 1.6062268206543004)



N=40 priors - λ ~U(0,2) α ~U(0,2)

Parameter : alpha
Mean : 1.0968783547560028
MSE : 0.027780547462796866
Credible interval : (0.7795628342344206, 1.2867176305747083)

Parameter : lambda
Mean : 1.2074745641569902
MSE : 0.08031846032466992
Credible interval : (0.8457696884909994, 1.5793941476253668)



Log concave posterior of alpha for 3 parameters

The posterior distribution of parameters is proportional to the likelihood function times the priors of alpha and gamma.

$$P(\mu,\ \alpha, \lambda\,|X) \propto P(X|\ \alpha, \lambda, \mu) P(\lambda) P(\alpha)\, P(\mu)$$

$\lambda \sim Gamma(a,b)$ $\qquad$ $\alpha \sim Gamma(c,d)$ $\qquad$ $\mu \sim U(0, X_1)$

$$P(\mu, \alpha, \lambda|X) \propto \frac{\lambda^{\,a+n-1}\, exp(-\lambda(b+\Sigma(Xi-\mu)^{\alpha}))\ \alpha^{c+n-1}\Pi(Xi-\mu)^{\alpha-1}\ exp(-d\alpha)\ b^a d^c}{\Gamma a\ \Gamma c\ X1}$$

$$\propto \frac{Gamma(a+n,\ b+\Sigma(Xi-\mu)^{\alpha})\ \Pi(Xi-\mu)^{\alpha-1}\ \alpha^{c+n-1}\ exp(-d\alpha)}{(b+\Sigma(Xi-\mu)^{\alpha})^{a+n}\ X1}$$

$$P(\lambda|\alpha, \mu, X) = Gamma(a+n,\ b+\Sigma(Xi-\mu)^{\alpha})$$

$$P(\alpha|\mu, X) = K\frac{\Pi(Xi-\mu)^{\alpha-1}e^{-d\alpha}\alpha^{c+n-1}}{(b+\Sigma(Xi-\mu)^{\alpha})^{a+n}} \qquad \text{log-concave pdf of } \alpha$$

$$P(\mu|X) = 1/X1$$

The shape of the posterior density function closely resembles that of the gamma density function, so we approximate the posterior density function of α with the gamma density function by equating the first two moments.

$$\alpha|X \sim Gamma(p, q)$$

We have normalized the integration by multiplying with appropriate constants to prevent the problem of exploding values while estimation of parameters.

Algorithm for calculating Bayes estimate

Step 1: Sample a random sample (X) of a weibull distribution.

Step 2: Generate μ1 from U(0,X1).

Step 2: Generate α1 from the log concave density P(α|μ,X) ≈ Gamma(p,q) using μ1.

Step 3: Generate λ1 from P(λ| α,μ,X) using α1 generated in step 2.

Step 4: Repeat steps 2 and 3 M times and sort the values so that

$\alpha_{(1)}<\alpha_{(2)}....<\alpha_{(M)}$ $\lambda_{(1)}<\lambda_{(1)}....<\lambda_{(M)}$ and $\mu_{(1)}<\mu_{(2)}....<\mu_{(M)}$

Step 4: The 100(1-2β)% symmetric confidence interval can be calculated as

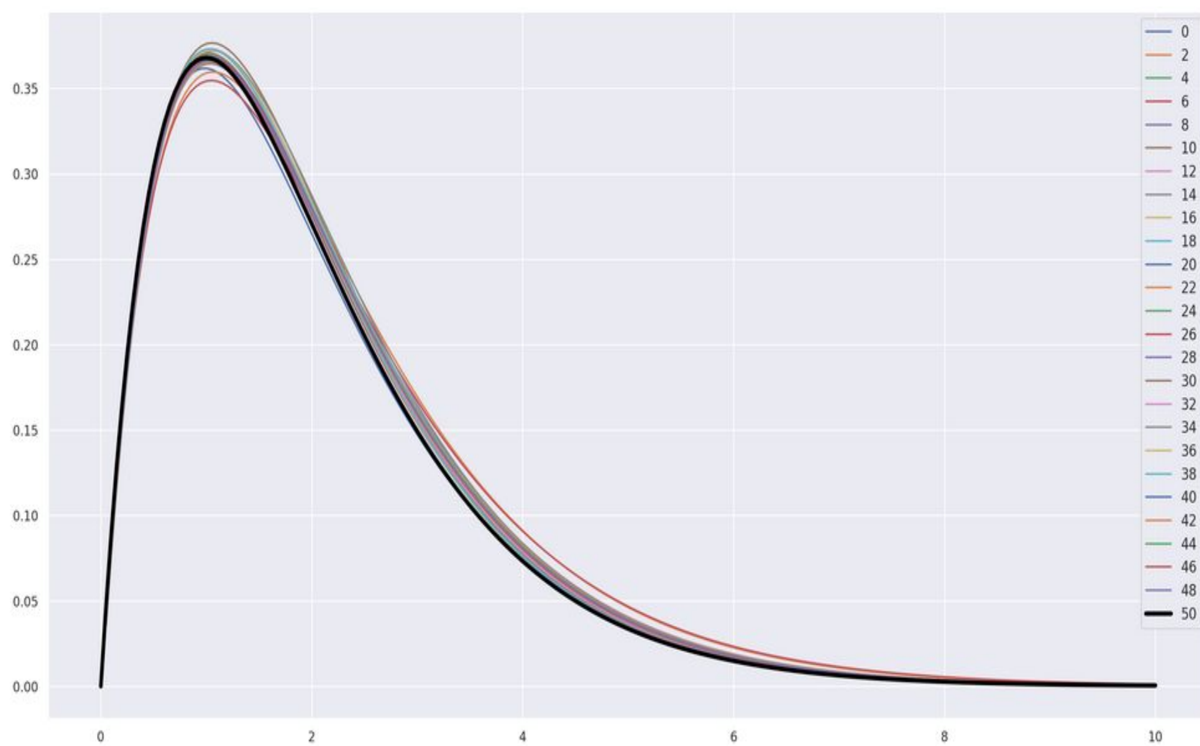
$(\alpha_{[M\beta]}, \alpha_{[M(1-\beta)]})$,$(\mu_{[M\beta]}, \mu_{[M(1-\beta)]})$ and $(\lambda_{[M\beta]}, \lambda_{[M(1-\beta)]})$

N=25 priors - λ ~Gamma(50, 50) α ~Gamma(50, 50) μ~U(0,X1)

Parameter : alpha
Mean : 0.9951597170939587
MSE : 0.015450828875074348
Credible interval : (0.7740062627404377, 1.2561235384678628)

Parameter : lambda
Mean : 0.9691014147742966
MSE : 0.016550533865813597
Credible interval : (0.7025364798000547, 1.2502946875545993)

Parameter : mu
Mean : 0.007879628118498135
MSE : 7.83404238892769e-05
Credible interval : (0.00089049393896781, 0.013892594677468485)

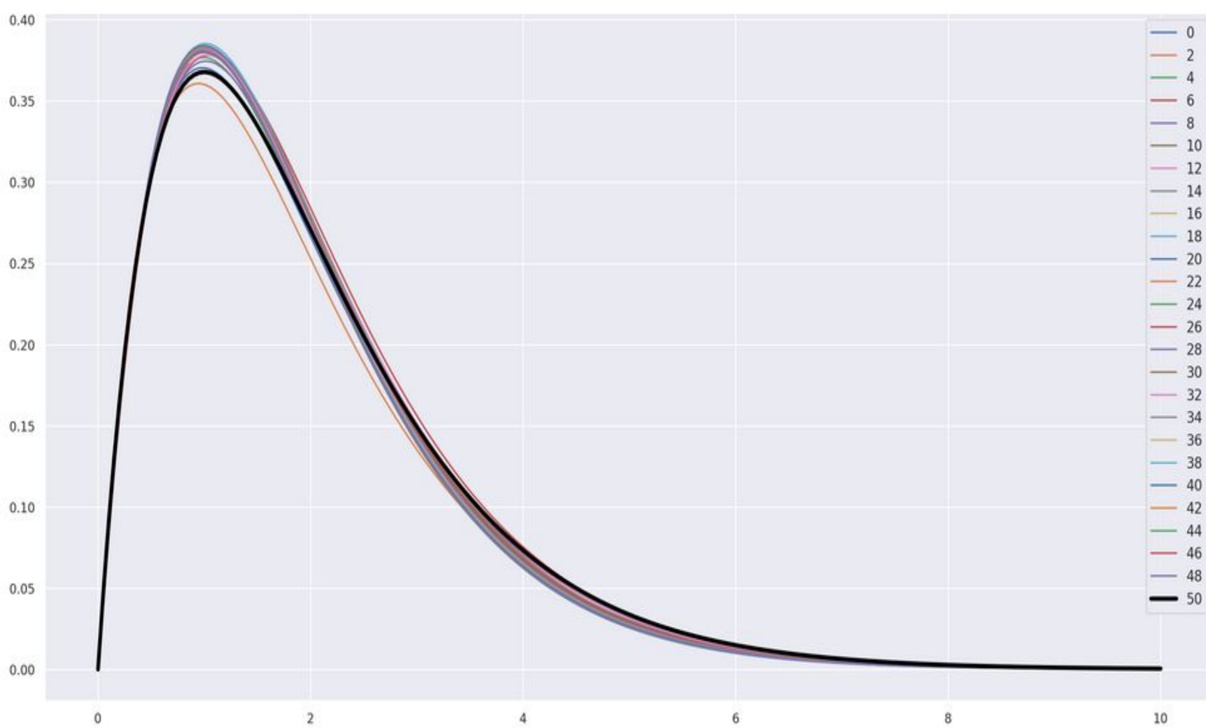


N=40 priors - λ ~Gamma(50, 50) α ~Gamma(50, 50) μ~U(0,X1)

Parameter : alpha
Mean : 1.019053392881811
MSE : 0.010129694370910807
Credible interval : (0.8775080919084305, 1.189752856665867)

Parameter : lambda
Mean : 1.0181873977433886
MSE : 0.013323046237448044
Credible interval : (0.776824787039948, 1.2625367840374693)

Parameter : mu
Mean : 0.005139478330592762
MSE : 3.610618685515901e-05
Credible interval : (0.00024146617260410318, 0.010527728152671136)

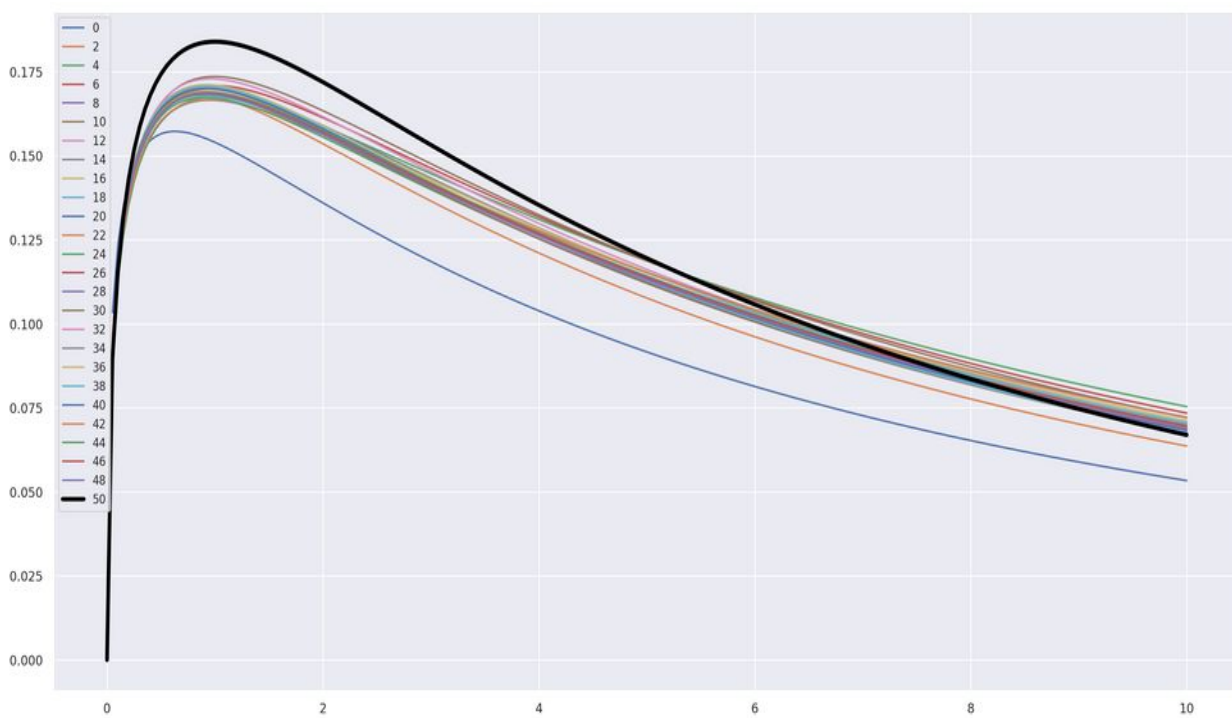


N=25 priors - λ ~Gamma(50, 50) α ~Gamma(50, 100) μ~U(0,X1)

Parameter : alpha
Mean : 0.4568094582909816
MSE : 0.003410863737340648З
Credible interval : (0.3871765791325081, 0.5343472976147146)

Parameter : lambda
Mean : 1.0532047655088443
MSE : 0.020541685553223013
Credible interval : (7973508093082481, 1.3213810508596895)

Parameter : mu
Mean : 5.96162420913552 16e-05
MSE : 4.78181852537756e-09
Credible interval : (2.5618840783464395e-06, 0.00011591185167233641)



N=40 priors - λ ~Gamma(50, 50) α ~Gamma(50, 100) μ~U(0,X1)

Parameter : mu = 0.2
Mean : 0.10995598635959622
MSE : 0.01361701495411246
Credible interval : (0.003873749963054199, 0.2562122452727793)

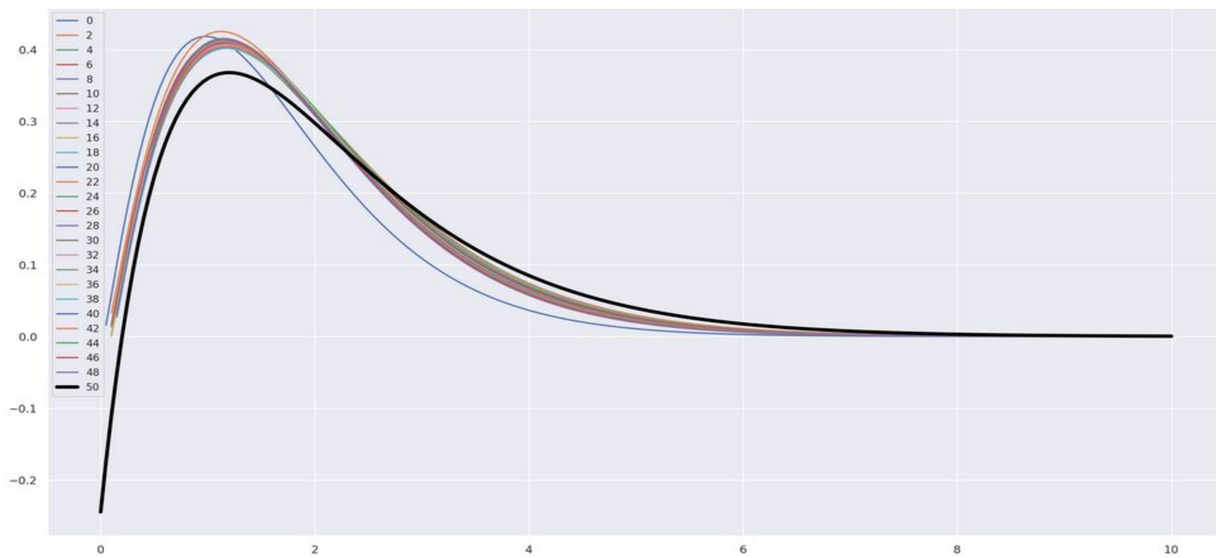


Fig: convergence of estimates of parameters for 40 data points for non zero μ

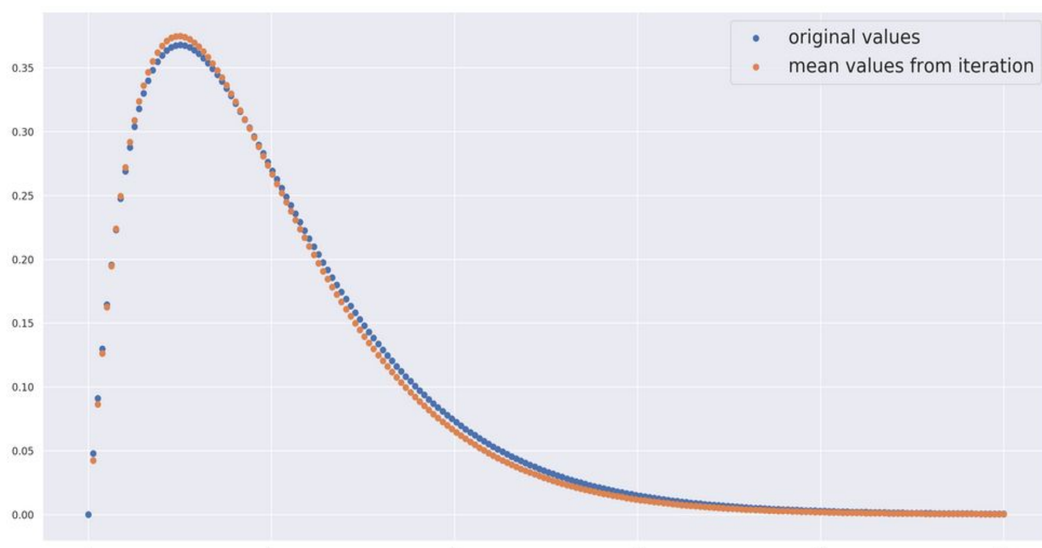


Fig: comparison of weibull distribution using 40 data points of the original values vs the final estimate for alpha = 1, lambda = 1 and mu=0

Bayesian estimation by splitting the posterior

Priors taken -
$\lambda \sim gamma(a,b)$ $\alpha \sim gamma(c,d)$ $\mu \sim uniform(0,Z1)$

Posterior distribution :

$$\pi(\alpha,\ \lambda,\ \mu|\ X)\ =\ (1/p)(\alpha)^{c+n-1}\ (\lambda)^{a+n-1} exp(-\lambda \sum(Xi-\mu)^{\alpha})\ exp(-b\lambda - d\alpha)\ \prod(Xi-\mu)^{\alpha-1}$$

By splitting the posterior, we obtain

$$\pi(\alpha,\ \lambda,\ \mu|\ X)\ =\ f(\lambda|\alpha,\ \mu,\ X)\, f(\alpha|\ \mu,\ X)\, f(\mu|X)\ h(\alpha,\ \lambda,\ \mu)$$

$$\lambda|\ \alpha,\ \mu,\ X\ \sim gamma(a+n,\ b+\sum(Xi-\mu)^{\alpha}$$

$$\alpha|\ \mu,\ X\ \sim gamma(c+n,\ d-\sum ln(Xi-\mu)$$

$$h(\alpha,\ \lambda,\ \mu)\ =\ \frac{\Gamma(a+n)\Gamma(c+n)Z1}{\Pi(Xi-\mu)\ p\ (b+\Sigma(Xi-\mu)^{\alpha})^{a+n}(d-\Sigma ln(Xi-\mu))^{c+n}}$$

$$\widehat{\alpha}\ = E_{\ \theta|X}(\alpha)$$

$$\widehat{\alpha} = \iiint h(\alpha,\ \lambda,\ \mu) g(\alpha,\ \lambda,\ \mu|X)\ d\alpha\ d\lambda\ d\mu\ =\ E_{\ g}(h(\alpha,\ \lambda,\ \mu))$$

Where $g(\alpha,\ \lambda,\ \mu|X)\ =\ f(\lambda|\alpha,\ \mu,\ X)\, f(\alpha|\ \mu,\ X)\, f(\mu|\ X)$

By weak law of large numbers,

$$\widehat{\alpha} = 1/n\ \Sigma h(\alpha i,\ \lambda i,\ \mu i)$$

Algorithm for calculating Bayes estimate

Step 1: Generate μ from f(μ).

Step 2: Using this μ, generate α from f(α|μ).

Step 3: Using generated α and μ, generate λ from f(λ| α, μ).

Step 4: Using these generated λ, α and μ get the value of $h(\alpha,\ \lambda,\ \mu)$.

Step 5: Repeat steps 1 to 4 100 times.

Step 5: Get estimates of λ, α and μ from the above generated values.

Note - algorithm took too long to run in R so the results couldn’t be compiled.

Conclusion

Comparison of MCMC Algorithms used:

- Metropolis Hasting
- Gibbs Sampling with log concave posterior for Alpha

	MCMC Algorithm: Metropolis Hasting	Gibbs Sampling with log concave posterior for Alpha
N=25 Alpha ~ Gamma(50,50) Lambda ~ Gamma(50,50) Mu ~ Uniform(0,z1)	ALPHA=1 Mean: 1.013801 MSE: 0.007892	ALPHA=1 Mean: 0.9951597 MSE: 0.0154508
	LAMBDA=1 Mean: 1.018393 MSE: 0.009491	LAMBDA=1 Mean: 0.9691014 MSE: 0.0165505
	MU=0.5 Mean: 0.461479 MSE: 0.003276	MU=0 Mean: 0.0078796 MSE: 7.8340424e-05
N=25 Alpha ~ Gamma(50,100) Lambda ~ Gamma(100,100) Mu ~ Uniform(0,z1)	ALPHA=0.5 Mean: 0.495806 MSE: 0.256297	ALPHA=0.5 Mean: 0.4568094 MSE: 0.0034109
	LAMBDA=1 Mean: 0.977214 MSE: 0.007870	LAMBDA=1 Mean: 1.0532048 MSE: 0.0205417
	MU=0.5 Mean: 0.496786 MSE: 3.963593e-05	MU=0 Mean: 5.9616242e-05 MSE: 4.7818185e-09

We can see that among the two Bayesian methods employing MCMC Algorithms, samples from Metropolis Hasting give a more accurate mean value and lesser variation from the actual value for the parameters **Alpha** and **Lambda**. However, Gibbs sampling approach with log concave posterior for alpha gives a better result for the parameter **Mu.**

Comparison for the two parameter estimations

	MCMC Algorithm: Metropolis Hasting	Gibbs Sampling with log concave posterior for Alpha
N=40 Alpha ~ Gamma(1,1) Lambda ~ Gamma(1,1)	ALPHA Mean: 1.076125 MSE: 0.0220139	ALPHA Mean: 0.8945029 MSE: 0.02295469
	LAMBDA Mean: 1.165857 MSE: 0.0581435	LAMBDA Mean: 1.1302731 MSE: 0.0496044

In the two parameter estimation, it is again not possible to rank one MCMC sampling algorithm over the other. The results get more accurate and precise with increasing sample size (N).

Bias in Bayesian estimation using log-concave method

With N=25 and N=40 data points, the bias associated with the estimation of alpha, lambda and mu seems to be increasing with an increase in the number of iterations. Therefore, better results are obtained with 100-200 iterations which is lower than what is generally used. The **high bias** can be attributed to the **low availability of data-points** (25 or 40 data points). An increase in the number of data points(500 or 1000 data points) results in **exploding values** in the algorithm. Possible future experiments can include alterations to the algorithm to take care of exploding values.

Comparison for 50 and 5000 iterations

Parameter	Mean(50)	Mean(5000)	MSE(50)	MSE(5000)
Alpha (1)	1.0190533928	0.9540325158	0.0101296943	0.0103225196
Lambda (1)	1.018187397	1.0401582260	0.0133230462	0.0142235072
Mu (0.01)	0.0051394783	0.0045431089	3.6106186e-05	2.7358532044

Comparison of all the Parameter Estimation Methods used

N=50 Alpha=1 Lambda=1 Mu=0

PRIOR DISTRIBUTIONS FOR BAYESIAN CASE (Alpha=1 Lambda=1 Mu=0.5)

Alpha ~ Gamma(100,100) Lambda ~ Gamma(100,100) Mu ~ Uniform(0,z1)

	Alpha (MSE)	Lambda (MSE)	Mu (MSE)
Maximum Likelihood - I	1.051093 0.01086262	1.022025 0.01714685	-0.00167184 0.0003978323
Maximum Likelihood - II	0.9982208 0.0173581	1.03563 0.02474953	0.01222 0.00075894
Modified Moment Method	1.054168 0.02160289	1.013998 0.02569172	-5.392384e-05 0.0004016688
Least Square Method	0.95258391 0.002248285	0.92011014 0.00638239	0.03409797 0.001095475
Weighted Least Square Method	0.88672189 (0.012831929)	0.93336852 (0.004439754)	0.04050816 (0.001560895)
MAP-BFGS model	1.75281443	1.23679377	
MAP-Powell model	0.85221853	0.67099244	
Bayesian Method	1.046149 0.0105868	0.991525 0.0092074	0.485830 0.0003644