Homework 1

Christian Tillich

September 24, 2017

Setup

Problem 1

	Mean	SD	Naive SE	Time-series SE
mu	9.310872	0.5433104	0.0024298	0.0024298
sigma	1.680099	0.3813382	0.0017054	0.0018608
sigma2	2.968150	1.4886810	0.0066576	0.0072940

	2.5%	25%	50%	75%	97.5%
mu	8.236610	8.968636	9.313526	9.655361	10.400754
sigma	1.127616	1.412654	1.615312	1.875431	2.601137
sigma2	1.271518	1.995592	2.609231	3.517240	6.765914

The lower point of the 95% confidence interval is 1.125, so below 2.

Problem 2

	Mean	SD	Naive SE	Time-series SE
mu	9.300918	0.5406104	0.0024177	0.0023879
pt.sigma.gt	0.034500	0.1825114	0.0008162	0.0008506
sigma	1.677932	0.3810938	0.0017043	0.0018347

	2.5%	25%	50%	75%	97.5%
mu	8.220961	8.960129	9.302855	9.64440	10.369537
pt.sigma.gt	0.000000	0.000000	0.000000	0.00000	1.000000
sigma	1.127660	1.410542	1.613000	1.87264	2.598689

The posterior probability of sigma > 2.5 is 0.035, so above 0.02.

Problem 3

	Mean	SD	Naive SE	Time-series SE
cv	0.1813649	0.0430078	0.0001923	0.0002083

	Mean	SD	Naive SE	Time-series SE
mu	9.3073900	0.5429842	0.0024283	0.0024415
sigma	1.6815746	0.3783754	0.0016921	0.0018327

	2.5%	25%	50%	75%	97.5%
cv	0.120138	0.1512155	0.174016	0.2026466	0.2861655
mu	8.226063	8.9601631	9.307054	9.6518162	10.3882916
sigma	1.129945	1.4137514	1.620360	1.8791738	2.5930134

The 95% confidence interval for the coefficient of variation is [0.120, 0.285], so the upper point is above 0.25.

Problem 4

	Mean	SD	Naive SE	Time-series SE
cv	0.1813331	0.0432396	0.0001934	0.0002133
mu	9.3072984	0.5445513	0.0024353	0.0024435
p.cv.gt	0.2688000	0.4433402	0.0019827	0.0021064
sigma	1.6813372	0.3802945	0.0017007	0.0018708

	2.5%	25%	50%	75%	97.5%
cv	0.120145	0.151152	0.1738692	0.2028085	0.285788
mu	8.216208	8.961710	9.3057129	9.6548548	10.393368
p.cv.gt	0.000000	0.000000	0.0000000	1.0000000	1.000000
sigma	1.127368	1.414114	1.6176863	1.8787393	2.591994

The probability of the cv > 0.20 is 27%. So that probability is below 30%.

Problem 5

	Mean	SD	Naive SE	Time-series SE
mu	8.417775	0.9434623	0.0042193	0.0042527
p.rep.gt[1]	0.988900	0.1047712	0.0004686	0.0004699
p.rep.gt[2]	0.420720	0.4936796	0.0022078	0.0022078
p.rep.gt[3]	0.423400	0.4941025	0.0022097	0.0022169
p.rep.gt[4]	0.559020	0.4965094	0.0022205	0.0021795
p.rep.gt[5]	0.353440	0.4780426	0.0021379	0.0021379
p.rep.gt[6]	0.681460	0.4659148	0.0020836	0.0020836
p.rep.gt[7]	0.118920	0.3236976	0.0014476	0.0014561
p.rep.gt[8]	0.195220	0.3963739	0.0017726	0.0017727
p.rep.gt[9]	0.557460	0.4966923	0.0022213	0.0022213
p.rep.gt[10]	0.240760	0.4275492	0.0019121	0.0019121
sigma	2.924985	0.6572176	0.0029392	0.0031836

	2.5%	25%	50%	75%	97.5%
mu	6.534727	7.818066	8.422957	9.016504	10.312873
p.rep.gt[1]	1.000000	1.000000	1.000000	1.000000	1.000000
p.rep.gt[2]	0.000000	0.000000	0.000000	1.000000	1.000000
p.rep.gt[3]	0.000000	0.000000	0.000000	1.000000	1.000000
p.rep.gt[4]	0.000000	0.000000	1.000000	1.000000	1.000000
p.rep.gt[5]	0.000000	0.000000	0.000000	1.000000	1.000000
p.rep.gt[6]	0.000000	0.000000	1.000000	1.000000	1.000000
p.rep.gt[7]	0.000000	0.000000	0.000000	0.000000	1.000000
p.rep.gt[8]	0.000000	0.000000	0.000000	0.000000	1.000000
p.rep.gt[9]	0.000000	0.000000	1.000000	1.000000	1.000000
p.rep.gt[10]	0.000000	0.000000	0.000000	0.000000	1.000000
sigma	1.970375	2.462789	2.817195	3.265763	4.511555

1) Observation 1 has the highest probability of y_rep > y, at p = 98.7%.
2) That probability is over 95%.
3) It's a problematic model. Per the model, the probability of recreating the first observation is < 2%.

Problem 6

	Mean	SD	Naive SE	Time-series SE
mu	8.898756	0.7439233	0.0033269	0.0043005
p.rep.gt[1]	0.998700	0.0360324	0.0001611	0.0001647
p.rep.gt[2]	0.479340	0.4995780	0.0022342	0.0022651
p.rep.gt[3]	0.482880	0.4997118	0.0022348	0.0022348
p.rep.gt[4]	0.676460	0.4678314	0.0020922	0.0021525
p.rep.gt[5]	0.386640	0.4869849	0.0021779	0.0022104
p.rep.gt[6]	0.826280	0.3788723	0.0016944	0.0017677
p.rep.gt[7]	0.071580	0.2577938	0.0011529	0.0011849
p.rep.gt[8]	0.155480	0.3623653	0.0016205	0.0016795
p.rep.gt[9]	0.677560	0.4674150	0.0020903	0.0021452
p.rep.gt[10]	0.215760	0.4113527	0.0018396	0.0018566
sigma	1.988394	0.5644666	0.0025244	0.0036237

	2.5%	25%	50%	75%	97.5%
mu	7.391607	8.434722	8.910118	9.374002	10.338518
p.rep.gt[1]	1.000000	1.000000	1.000000	1.000000	1.000000
p.rep.gt[2]	0.000000	0.000000	0.000000	1.000000	1.000000
p.rep.gt[3]	0.000000	0.000000	0.000000	1.000000	1.000000
p.rep.gt[4]	0.000000	0.000000	1.000000	1.000000	1.000000
p.rep.gt[5]	0.000000	0.000000	0.000000	1.000000	1.000000
p.rep.gt[6]	0.000000	1.000000	1.000000	1.000000	1.000000
p.rep.gt[7]	0.000000	0.000000	0.000000	0.000000	1.000000
p.rep.gt[8]	0.000000	0.000000	0.000000	0.000000	1.000000
p.rep.gt[9]	0.000000	0.000000	1.000000	1.000000	1.000000
p.rep.gt[10]	0.000000	0.000000	0.000000	0.000000	1.000000
sigma	1.185520	1.588405	1.887503	2.279136	3.368682

1) Observeration 1 still has the highest probability of y_rep > y, now at p = 99.8%
2) It's still over 95%
3) The mean went from 8.41 in Problem 5 to 8.91 in Problem 6. But that would help explain why the it became less likely to replicate observation 1 despite having the fatter tails (not gonna lie, this still seems counter-intuitive)