

Lecture III: Simulation and Bayesian Estimation

Charles B. Moss

May 23, 2013

- 1 Drawing the Gamma
 - Empirical Moments
- 2 Simple Estimation by Simulation

Drawing the Gamma

- Recalling the Gamma distribution introduced in Lecture 2

$$f(V|\alpha, \beta) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^\alpha} V^{\alpha-1} \exp\left[-\frac{V}{\beta}\right] & \text{such that } 0 < v < \infty \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where $\alpha > 0$ and $\beta > 0$. The mean of this distribution is $E[V] = \alpha\beta$ and its variance is $V[V] = \alpha\beta^2$.

- First, we use R to draw 750 random draws from the Gamma distribution with $\alpha = 2.5$ and $\beta = 2$. The R code for this process is

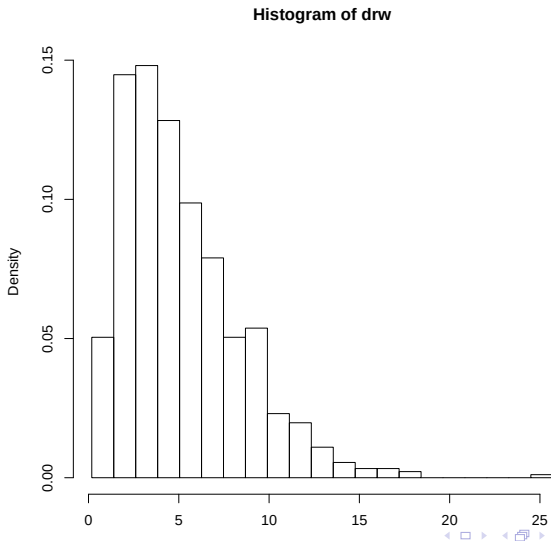
```
library(histogram)
drw      <- rgamma(750,2.5,scale=2)
gamhist <- histogram(drw)
```

- The first 120 of these draws are presented in Table 4 and histogram for all points is presented in Figure 5.

Random Gamma Draws

	0	20	40	60	80	100
1	2.61503	5.17145	5.42339	9.19676	3.68125	8.81904
2	5.13324	3.28600	5.35471	4.43099	5.51113	3.26221
3	4.69766	3.58568	5.21370	1.59593	1.30632	8.72694
4	3.60631	1.66896	2.99410	3.38592	6.43797	3.86667
5	3.94366	2.44866	2.27548	2.44204	2.74299	4.28329
6	4.89291	8.86468	3.45237	5.14626	5.75664	2.16090
7	3.45956	2.45509	7.22872	0.46441	5.15725	0.74350
8	1.71173	8.87103	25.71109	3.45818	11.10078	2.26246
9	1.38569	5.28757	9.04270	2.36622	7.58544	1.03688
10	6.58418	13.03351	4.11789	6.67662	4.05619	3.48332
11	2.79558	2.75118	1.94155	6.28265	3.73187	10.67169
12	0.80865	7.80284	2.59735	7.66718	4.89064	5.51004
13	8.07029	1.28752	6.94453	3.70194	9.11318	4.17981
14	4.63156	3.16253	2.58286	5.36301	6.34131	2.12394
15	5.12068	5.44437	0.52644	3.08291	2.59227	13.41238
16	8.73106	3.13356	1.68183	10.87351	4.67188	5.89577
17	9.68522	8.12020	2.24163	4.38590	6.36913	8.86390
18	1.96429	2.67012	1.50678	2.56803	2.11862	3.98088
19	5.15683	6.15837	8.71453	15.44049	1.73490	1.60140
20	4.30616	1.14511	9.49274	3.82978	3.77089	9.91295

Empirical Gamma Distribution



- As a starting point for our discussion, I would to consider the sample mean and variance for the simulated draws

$$\bar{V} = \frac{1}{750} \sum_{i=1}^{750} V_i = 5.1450 \quad (2)$$

$$V[V] = \frac{1}{750 - 1} \sum_{i=1}^{750} (V_i - 5.1450)^2 = 10.7940$$

Comparing these to the theoretical values $\mu_V = \alpha\beta = 5.0000$ and $\sigma_V^2 = \alpha\beta^2 = 10.0$, the simulated data are consistent with the original distribution.

Empirical Moments

- However, we can gain a little more intuition from these results. For example, consider the mean and variance based on theoretical expectations

$$\begin{aligned}\mu_V &= \int_0^\infty V \frac{1}{\Gamma(\alpha) \beta^\alpha} V^{\alpha-1} \exp\left[\frac{-V}{\beta}\right] dV \\ \sigma_V^2 &= \int_0^\infty (V - \mu_V)^2 \frac{1}{\Gamma(\alpha) \beta^\alpha} V^{\alpha-1} \exp\left[\frac{-V}{\beta}\right] dV\end{aligned}\tag{3}$$

- Next, consider an approximation for Equation 3 based on a discrete sample

$$\mu_V \approx \sum_{i=1}^N \tilde{V}_i \times P[V_i] \quad (4)$$

$$\sigma_V^2 \approx \sum_{i=1}^N \left(\tilde{V}_i - \mu_V \right)^2 \times P[V_i]$$

- In the case of sample statistics we have $N \rightarrow 750$ and $P[V_i] = 1/750 \forall i$. Consider the parameters for the histogram presented in Table 9. The probability for each band is defined as

Histogram Probability Density Functions

i	Mid-Point V_i	Breaks a_i	Density h_i	Probability $P[V_i]$	Mean	Variance
0		0.1813				
1	0.7891	1.3970	0.0505	0.0613	0.0484	1.1748
2	2.0048	2.6127	0.1448	0.1760	0.3528	1.7584
3	3.2205	3.8284	0.1481	0.1800	0.5797	0.6810
4	4.4362	5.0441	0.1283	0.1560	0.6921	0.0830
5	5.6519	6.2598	0.0987	0.1200	0.6782	0.0284
6	6.8676	7.4755	0.0790	0.0960	0.6593	0.2781
7	8.0834	8.6912	0.0505	0.0613	0.4958	0.5221
8	9.2991	9.9069	0.0537	0.0653	0.6075	1.1162
9	10.5148	11.1226	0.0230	0.0280	0.2944	0.8012
10	11.7305	12.3383	0.0197	0.0240	0.2815	1.0343
11	12.9462	13.5540	0.0110	0.0133	0.1726	0.8072
12	14.1619	14.7697	0.0055	0.0067	0.0944	0.5395
13	15.3776	15.9854	0.0033	0.0040	0.0615	0.4171
14	16.5933	17.2011	0.0033	0.0040	0.0664	0.5224
15	17.8090	18.4169	0.0022	0.0027	0.0475	0.4263
16	19.0247	19.6326	0.0000	0.0000	0.0000	0.0000
17	20.2404	20.8483	0.0000	0.0000	0.0000	0.0000
18	21.4561	22.0640	0.0000	0.0000	0.0000	0.0000
19	22.6718	23.2797	0.0000	0.0000	0.0000	0.0000
20	23.8875	24.4954	0.0000	0.0000	0.0000	0.0000
21	25.1032	25.7111	0.0011	0.0013	0.0335	0.5300
Sum				1.000	5.1657	10.7200

- Empirical Variance

$$P[V_i] = h_i \times (a_i - a_{i-1}) \quad (5)$$

Thus for the first band

$$P[0.7891] = 0.0505 \times (1.3970 - 0.1813) = 0.0613 \quad (6)$$

- The sixth column in Table 9 is then

$$V_i \times P[V_i] = V_i \times h_i \times (a_i - a_{i-1}) \quad (7)$$

and the sum of the sixth column is the sample mean. Similarly, the seventh column is the probability weighted variance. Notice that mean based on this probability weighting is 5.1657 which is extremely close to both the unweighted sample mean and the true mean. Similarly, the probability weighted variance is 10.7200 which is also close to both the unweighted sample variance and the true variance. Hence, there is a correspondence between the true integrals in Equation 3 and the numerical approximations in Equation 4. This correspondence will be extensively exploited in applied Bayesian analysis.

Simple Estimation by Simulation

- Consider the small sample on the share of capital cost for production agriculture from the KLEM data introduced in Lecture 2. Suppose we assume a prior normal distribution with a mean of $\tilde{\mu} = 0.25$ and a Coefficient of Variation of 1.25. The variance would then be $(0.25 \times 1.25)^2 = 0.097656$. The posterior distribution would then be

$$f[\alpha | y^0] = \frac{1}{\sqrt{2\pi \times 0.096757}} \exp \left[-\frac{(\alpha - 0.25)^2}{2 \times 0.096757} \right]$$

$$\left(\frac{1}{\sqrt{2\pi\sigma_y^2}} \right)^{10} \exp \left[-\frac{(0.1532 - \alpha)^2 + (0.1620 - \alpha)^2 + \cdots (0.1818 - \alpha)^2}{2\sigma_y^2} \right] \quad (8)$$

Capital Share in KLEM Data

Year	Share
1960	0.1532
1961	0.1620
1962	0.1548
1963	0.1551
1964	0.1343
1965	0.1534
1966	0.1587
1967	0.1501
1968	0.1459
1969	0.1818

- The concept is then to estimate α by integrating the probability density function by simulation. Specifically

$$\hat{\alpha}_{Bayes} = \int_{-\infty}^{\infty} \alpha f[\alpha | y^0] d\alpha. \quad (9)$$

As a starting point, assume that we draw a value of α based on the prior distribution. For example, in the first row of Table 15 we draw $\alpha = 0.8881$. Given this draw, we compute the likelihood function of the sample

Simulation Share Estimator

Draw	α	$L(\alpha y^0)$	$\alpha \times L(\alpha y^0)$
1	0.8881	8.5594E-13	7.6018E-13
2	0.1610	9.9130E-01	1.5960E-01
3	0.1159	9.1806E-01	1.0642E-01
4	0.3177	2.5264E-01	8.0261E-02
5	0.0058	3.1454E-01	1.8129E-03
6	-0.0174	2.1401E-01	-3.7269E-03
7	0.2817	4.3307E-01	1.2198E-01
8	0.9582	3.2764E-15	3.1395E-15
9	0.3789	7.4286E-02	2.8149E-02
10	-0.1751	3.5680E-03	-6.2479E-04
11	-0.0477	1.1906E-01	-5.6766E-03
12	0.2405	6.8021E-01	1.6360E-01
13	0.2461	6.4644E-01	1.5908E-01
14	0.2991	3.3906E-01	1.0143E-01
15	0.7665	3.9991E-09	3.0655E-09
16	0.0952	8.2591E-01	7.8617E-02
17	-0.0060	2.6062E-01	-1.5565E-03
18	0.2745	4.7428E-01	1.3020E-01
19	0.6151	1.7537E-05	1.0787E-05
20	0.3079	2.9626E-01	9.1228E-02

$$L(\alpha | y^0) \propto \exp \left[- \frac{(0.1532 - 0.8881)^2 + (0.1620 - 0.8881)^2 + \cdots (0.1818 - 0.8881)^2}{2 \times 0.096757} \right] \quad (10)$$

$$= 8.5594E - 13$$

- Summing over the 20 observations

$$\hat{\alpha}_{Bayes} = \frac{\sum_{i=1}^{20} \alpha_i L[\alpha_i | y^0]}{\sum_{i=1}^{20} L[\alpha_i | y^0]} = 0.1769 \quad (11)$$

Expanding the sample to 200 draws $\hat{\alpha}_{Bayes} = 0.1601$.

Sample R Code

```
dta <- rbind(0.1532,0.1620,0.1548,0.1551,0.1343,0.1534,0.1587,
             0.1501,0.1459,0.1818)

a <- rnorm(200,mean=0.25,sd=sqrt(0.096757))

for (i in 1:200) {
  for (t in 1:10) {
    if (t == 1) l <- 0
    l <- l -((dta[t]-a[i])^2)/(2*0.096757) }
    if (i == 1) res <- cbind(a[i],exp(l),a[i]*exp(l)) else \
      res <- rbind(res,cbind(a[i],exp(l),a[i]*exp(l))) }

ahat <- sum(res[,3])/sum(res[,2])

print(res)
print(ahat)
```