

Lecture I: Mixed Estimators

Charles B. Moss

May 19, 2013

1 Basic Distributions

2 Mixed Estimator

Basic Distributions

- In Lecture 1 I played a little fast and loose with the variance coefficients of the normal distribution. Let us return to the prior distribution for the mean

$$f_1(\mu | \tilde{\mu}, \sigma_\mu^2) = \frac{1}{\sqrt{2\pi\sigma_\mu^2}} \exp \left[-\frac{(\mu - \tilde{\mu})^2}{2\sigma_\mu^2} \right] \quad (1)$$

- Here we explicitly recognize the fact that our prior for μ (based on the normal distribution) has a variance parameter σ_μ^2 .

- For the next wrinkle, we want to recognize the variance of the observed variable (y)

$$f_2(y | \mu, \sigma_y^2) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp \left[-\frac{(y - \mu)^2}{2\sigma_y^2} \right] \quad (2)$$

where σ_y^2 is the variance of the normal distribution for our observed variable.

- Based on the formulation in Equation 2 we have acquired a new unknown parameter – σ_y^2 . Like the mean, we need to formulate a prior for this variable.
 - In formulating this variable we need to consider the characteristics of parameter – it needs to be always positive (say $V \equiv \sigma_y^2 \gg 0$).
 - One parameter that fits the mold for this distribution is the Gamma Distribution. Following Moss (2014) Section 3.5 the Gamma distribution can be written as

$$f(V|\alpha, \beta) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^\alpha} V^{\alpha-1} \exp\left[-\frac{V}{\beta}\right] & \text{such that } 0 < v < \infty \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

Letting $\alpha = 2.5$ and $\beta = 2$ the graph of the Gamma function is depicted in Figure 1

Gamma Distribution Function

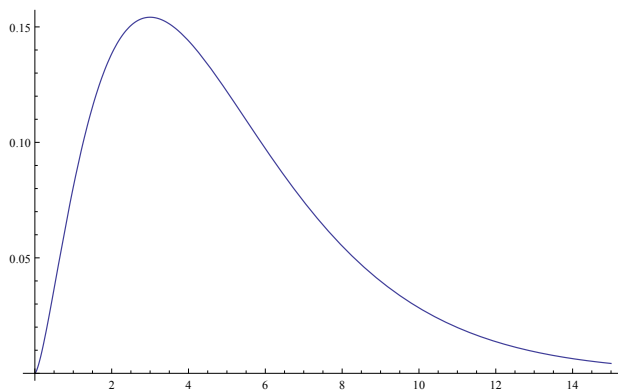


Figure : Gamma Distribution Function

Complete Prior

- Folding Equation 1 and Equation 3 together gives the complete prior (i.e., with the variance terms)

$$\begin{aligned}
 f_1(\mu, V | \tilde{\mu}, \sigma_\mu^2, \alpha, \beta) &= f_{11}(\mu | \tilde{\mu}, \sigma_\mu^2) \times f_{12}(V | \alpha, \beta) \\
 &= \frac{1}{\sqrt{2\pi\sigma_\mu^2}} \exp\left[-\frac{(\mu - \tilde{\mu})^2}{2\sigma_\mu^2}\right] \times \\
 &\quad \frac{1}{\Gamma(\alpha)\beta^\alpha} V^{\alpha-1} \exp\left[-\frac{V}{\beta}\right]
 \end{aligned} \tag{4}$$

The sample distribution then becomes

$$f_2(y | \mu, \sigma_y^2) = \frac{1}{V\sqrt{2\pi}} \exp\left[-\frac{(y - \mu)^2}{2V}\right] \tag{5}$$

- The posterior distribution then becomes

$$\begin{aligned}
 f(\tilde{\mu}, \sigma_{\mu}^2, \alpha, \beta | y) &= \frac{1}{\sqrt{2\pi\sigma_{\mu}^2} (\sqrt{2\pi V}) \Gamma(\alpha) \beta^{\alpha}} V^{\alpha-1} \\
 &\times \exp \left[-\frac{(\mu - \tilde{\mu})^2}{2\sigma_{\mu}^2} - \frac{(y - \mu)^2}{2V} - \frac{V}{\beta} \right] \\
 &= \frac{1}{2\pi\sigma_{\mu} V^{1/2} \Gamma(\alpha) \beta^{\alpha}} V^{\alpha-1} \\
 &\times \exp \left[-\frac{(\mu - \tilde{\mu})^2}{2\sigma_{\mu}^2} - \frac{(y - \mu)^2}{2V} - \frac{V}{\beta} \right]
 \end{aligned} \tag{6}$$

For simplification purposes, the Gamma distribution is typically replaced with the inverse Gamma.

Mixed Estimator

- Focusing on the exponential term in Equation 6, we can derive a form of a mixed estimator.
- The first two terms in the exponent of Equation 6 can be combined

$$-\frac{(\mu - \tilde{\mu})^2}{2\sigma_\mu^2} - \frac{(y - \mu)^2}{2V} = -\frac{(\mu - \tilde{\mu})^2 V + (y - \mu)^2 \sigma_\mu^2}{2\sigma_\mu^2 V} \quad (7)$$

- With a little bit of flare, we are going to divide the numerator by σ_μ^2 and let $\phi = V/\sigma_\mu^2$ yielding

$$-\frac{(\mu - \tilde{\mu})^2}{2\sigma_\mu^2} - \frac{(y - \mu)^2}{2V} = -\frac{\phi (\mu - \tilde{\mu})^2 + (y - \mu)^2}{2V} \quad (8)$$

- Maximizing δ with respect to μ gives an estimator of μ conditional on $\tilde{\mu}$ and ϕ

$$\hat{\mu} = \frac{\phi \tilde{\mu} + y}{1 + \phi} \quad (9)$$

- Next, let us extend the sample distribution to include two observations

$$\begin{aligned}\tilde{f}_2(y_1, y_2 | \mu, \sigma_y^2) &= \prod_{i=1}^2 f_2(y_i | \mu, \sigma_y^2) \\ &= \left(\frac{1}{\sqrt{2\pi V}} \right)^2 \exp \left[-\frac{(y_1 - \mu)^2 + (y_2 - \mu)^2}{2V} \right]\end{aligned}\tag{10}$$

- Folding the results for Equation 10 into Equation 6 implies a two observation form of Equation 8

$$\begin{aligned}-\frac{(\mu - \tilde{\mu})^2}{2\sigma_\mu^2} - \frac{(y_1 - \mu)^2 + (y_2 - \mu)^2}{2V} = \\ -\frac{\phi(\mu - \tilde{\mu})^2 + (y_1 - \mu)^2 + (y_2 - \mu)^2}{2V}\end{aligned}\tag{11}$$

- With a little bit of effort, the value of μ that maximizes the exponent can be derived as

$$\hat{\mu} = \frac{\frac{\phi}{2}\tilde{\mu} + \frac{y_1 + y_2}{2}}{1 + \frac{\phi}{2}} \quad (12)$$

- Defining $\bar{y} = (y_1 + y_2) / 2$, this estimator becomes

$$\hat{\mu} = \frac{\frac{\phi}{2}\tilde{\mu} + \bar{y}}{1 + \frac{\phi}{2}} \quad (13)$$

This formulation clearly combines (or mixes) sample information with prior information.

- Next, consider extending the sample to n observations. The sample distribution becomes

$$\begin{aligned}\tilde{f}_2(y_1, y_2, \dots, y_n | \mu, \sigma_y^2) &= \prod_{i=1}^n f_2(y_i | \mu, \sigma_y^2) \\ &= \left(\frac{1}{\sqrt{2\pi V}} \right)^n \exp \left[-\frac{\sum_{i=1}^n (y_i - \mu)^2}{2V} \right]\end{aligned}\tag{14}$$

- The combining the exponent term for posterior distribution yields

$$-\frac{(\mu - \tilde{\mu})^2}{2\sigma_\mu^2} - \frac{\sum_{i=1}^n (y_i - \mu)^2}{2V} = -\frac{\phi(\mu - \tilde{\mu})^2 + \sum_{i=1}^n (y_i - \mu)^2}{2V} \quad (15)$$

Following our standard approach, the value of μ that maximizes the exponent term is

$$\hat{\mu} = \frac{\frac{\phi}{n}\tilde{\mu} + \bar{y}}{1 + \frac{\phi}{n}} \ni \bar{y} = \frac{\sum_{i=1}^n y_i}{n} \quad (16)$$

KLEM Data

- Consider the data on the share of cost spent on capital inputs for agriculture from Jorgenson's KLEM data presented in Table 1.

Table : 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

- As in Lecture 1, let us assume a prior for this share of 0.25 (i.e., 1/4 of overall cost).
- First, let us use only the first observation and assume a $\phi = 2.0$. The mixed estimator of the average share would be

$$\hat{\mu} = \frac{2.0 \times 0.25 + 0.1532}{1 + 2.0} = 0.2177. \quad (17)$$

- Next, consider an increase in ϕ to 3.0

$$\hat{\mu} = \frac{3.0 \times 0.25 + 0.1532}{1 + 3.0} = 0.2258. \quad (18)$$

- Finally, consider reducing the ϕ to 1.0

$$\hat{\mu} = \frac{1.0 \times 0.25 + 0.1532}{1 + 1.0} = 0.2016. \quad (19)$$

- As ϕ increases more weight is put on the prior. The smaller ϕ , the closer the estimate is to the sample value.

- Next, consider using the first two data points so that $\bar{y} = 0.1576$.
Assuming $\phi = 2.0$

$$\hat{\mu} = \frac{\frac{2.0}{2} \times 0.25 + 0.1576}{1 + \frac{2.0}{2}} = 0.2038 \quad (20)$$

Notice that this estimate is closer to the sample average than for the same ϕ with one observation.

- Expanding to the full sample $\phi = 2.0$ and $\bar{y} = 0.1549$ yields

$$\hat{\mu} = \frac{\frac{2.0}{10} \times 0.25 + 0.1549}{1 + \frac{2.0}{10}} = 0.1708. \quad (21)$$

- Finally, notice that as $n \rightarrow \infty$ the weight on the prior vanishes so that the estimate converges to the sample average.