The Beta Distribution

Gary Schurman, MBE, CFA

April, 2016

The Beta Distribution is a model for random variables which can be constrained to the interval $[0, 1]$ and is therefore useful in modeling continuous random variables that lie between 0 and 1, such as probabilities, proportions and percentages. The Beta Distribution is also used as the conjugate prior distribution for binomial probabilities in Bayesian statistics (Gelman and Carlin 2004). When used in this Bayesian setting, a - 1 may be thought of as the prior number of successes and b - 1 may be thought of as the prior number of failures for the phenomena of interest (a and b are parameters to the beta function $\beta(a, b)$).

Our Hypothetical Problem

We are given annualized net charge-off rate monthly data for Commercial Banks for the years 1985 to 2015 (Source: Federal Reserve Bank of St. Louis). Net charge-off data is presented in the graph below...

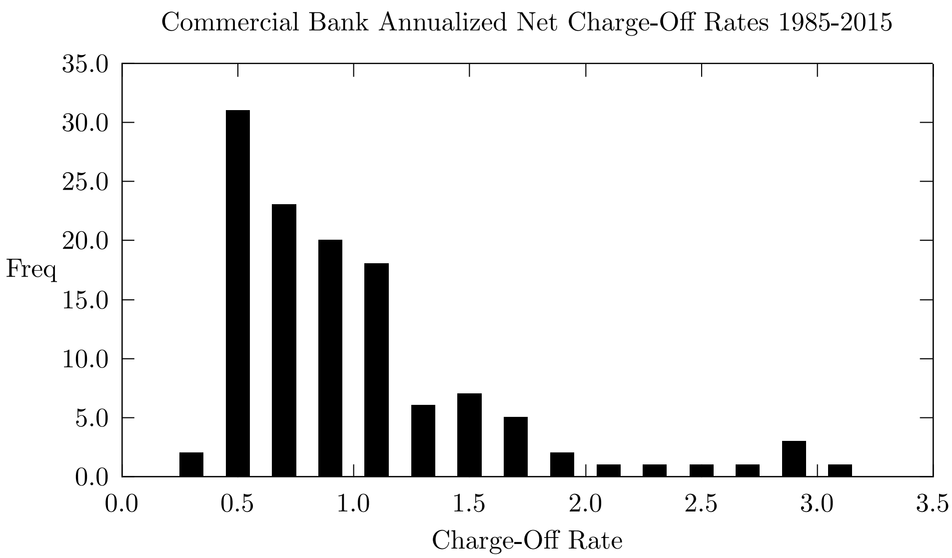

Table 1: Commercial Bank Net Charge-Off Rate

Net charge-off rate mean	0.9936%
Net charge-off rate standard deviation	0.5744%

Question: What is the probability that the annualized net charge-off rate for any given month will be greater than 2.00%?

The Probability Density Function

In the white paper **The Beta Function** we defined the beta function of the variable α and λ to be the following equation...

$$\beta(\alpha, \lambda) = \int\limits_{u=0}^{u=1} u^{\alpha-1} (1-u)^{\lambda-1} \delta u \;\; \text{...where...} \;\; \{\alpha \,|\, \alpha \in \mathbb{R}, \alpha > 0\} \;\; \text{...and...} \;\; \{\lambda \,|\, \lambda \in \mathbb{R}, \lambda > 0\} \tag{1}$$

The solution to Equation (1) above per that white paper is...

$$\beta(\alpha,\lambda)=\frac{\Gamma(\alpha)\,\Gamma(\lambda)}{\Gamma(\alpha+\lambda)} \tag{2}$$

We will define the function $f(\alpha,\lambda)$ to be the probability density function (PDF) for the Beta Distribution. Using Equations (1) and (2) above the equation for the PDF is...

$$f(\alpha,\lambda)=\frac{1}{\beta(\alpha,\lambda)}\,u^{\alpha-1}\,(1-u)^{\lambda-1}\ \text{...noting that...}\ \int\limits_{u=0}^{u=1} f(\alpha,\lambda)\,\delta u=\frac{\beta(\alpha,\lambda)}{\beta(\alpha,\lambda)}=1 \tag{3}$$

Using Equation (3) above the equation for the probability that the Beta-distributed random variable z is less than $x \in \{0,1\}$ given the Beta Distribution parameters α and λ is...

$$P\Bigg[z<x\Bigg]=\int\limits_{u=0}^{u=x} f(\alpha,\lambda)\,\delta u=\beta(\alpha,\lambda)^{-1}\int\limits_{u=0}^{u=x} u^{\alpha-1}\,(1-u)^{\lambda-1}\,\delta u \tag{4}$$

To solve Equation (4) above we can use the following Excel function...

$$P\Bigg[z<x\Bigg]=\text{BETA.DIST}(x,\alpha,\lambda,\text{TRUE}) \tag{5}$$

Note that the Beta Distribution will take the following shapes given that either or both of the parameters α and λ are less than one, equal to one, or greater than one...

Table 2: Shape of the Distribution

$\alpha>1$	$\lambda>1$	the distribution is bell shaped
$\alpha<1$	$\lambda<1$	the distribution is U shaped
$\alpha>1$	$\lambda<1$	the distribution is J shaped
$\alpha<1$	$\lambda>1$	the distribution is inverse J shaped
$\alpha=1$	$\lambda=1$	the distribution is an uniform distribution

Distribution Moments

We will define the variable M_1 to be the first moment of the Beta Distribution. Using Appendix Equation (16) below the equation for the first moment of the Beta Distribution is...

$$M_1=\mathbb{E}\Bigg[u\Bigg]=\frac{\alpha}{\alpha+\lambda} \tag{6}$$

We will define the variable M_2 to be the second moment of the Beta Distribution. Using Appendix Equation (18) below the equation for the second moment of the Beta Distribution is...

$$M_2=\mathbb{E}\Bigg[u^2\Bigg]=\frac{\alpha\,(\alpha+1)}{(\alpha+\lambda)\,(\alpha+\lambda+1)} \tag{7}$$

Using Equation (6) above the equation for the mean of the Beta Distribution is...

$$\text{Mean}=M_1=\frac{\alpha}{\alpha+\lambda} \tag{8}$$

Using Appendix Equation (19) below the equation for the variance of the Beta Distribution is...

$$\text{Variance}=M_2-M_1^2=\frac{\alpha\lambda}{(\alpha+\lambda)^2(\alpha+\lambda+1)} \tag{9}$$

Parameter Estimation

To use the Beta Distribution we need estimates of the parameters α and λ. If we have the first and second moments of the sample data and determine that the distribution of the sample data is consistent with a Beta Distribution then we can estimate the parameters α and λ via **The Method of Moments**.

Using Appendix Equation (23) below the equation for the Beta Distribution's α parameter estimate given the first two moments of the sample distribution is...

$$\alpha = (M_1 - M_2)/((1+\theta)\,M_2 - M_1) \text{ ...where... } \theta = \frac{1-M_1}{M_1} \tag{10}$$

Using Appendix Equation (20) below the equation for the Beta Distribution's λ parameter estimate given the first moment of the sample distribution and the estimate of the α parameter via Equation (10) above is...

$$\lambda = \frac{\alpha\,(1-M_1)}{M_1} \tag{11}$$

The Answer To Our Hypothetical Problem

Using the data in Table 1 above (net charge-off rates 1985 to 2015) the first two moments of the distribution of the sample data where the variable D represents the sample data points are...

$$M_1 = \mathbb{E}\Big[D\Big] = 0.009936 \text{ ...and... } M_2 = \mathbb{E}\Big[D^2\Big] = 0.005744^2 + 0.009936^2 = 0.000132 \tag{12}$$

Using the sample moments as determined by Equation (12) above and Equations (10) and (11) above the estimates of the Beta Distribution parameters α and λ are...

$$\alpha = 2.95 \text{ ...and... } \lambda = 294.21 \tag{13}$$

Using Equations (14) and (13) above the answer to our hypothetical problem is...

$$P\Big[z < 0.0200\Big] = \text{BETA.DIST}(0.0200, 2.95, 294.21, \text{TRUE}) = 6.05\% \tag{14}$$

Appendix

A. Using Equation (3) above the equation for the first moment of the Beta Distribution is...

$$\begin{aligned}
\mathbb{E}\Big[u\Big] &= \int_{u=0}^{u=1} \frac{1}{\beta(\alpha,\lambda)}\,u^{\alpha-1}\,(1-u)^{\lambda-1}\,u\,\delta u \\
&= \beta(\alpha,\lambda)^{-1} \int_{u=0}^{u=1} u^{\alpha}\,(1-u)^{\lambda-1}\,\delta u \\
&= \beta(\alpha,\lambda)^{-1} \times \beta(\alpha+1,\lambda)
\end{aligned} \tag{15}$$

Using Equation (2) above the solution to Equation (15) above is...

$$\begin{aligned}
\mathbb{E}\Big[u\Big] &= \frac{\Gamma(\alpha+\lambda)}{\Gamma(\alpha)\,\Gamma(\lambda)} \times \frac{\Gamma(\alpha+1)\,\Gamma(\lambda)}{\Gamma(\alpha+\lambda+1)} \\
&= \frac{\Gamma(\alpha+\lambda)}{\Gamma(\alpha)\,\Gamma(\lambda)} \times \frac{\alpha\,\Gamma(\alpha)\,\Gamma(\lambda)}{(\alpha+\lambda)\,\Gamma(\alpha+\lambda)} \\
&= \frac{\alpha}{\alpha+\lambda}
\end{aligned} \tag{16}$$

B. Using Equation (3) above the equation for the second moment of the Beta Distribution is...

$$\mathbb{E}\Bigg[u^2\Bigg] = \int\limits_{u=0}^{u=1} \frac{1}{\beta(\alpha,\lambda)}\, u^{\alpha-1}\,(1-u)^{\lambda-1}\, u^2\,\delta u$$
$$= \beta(\alpha,\lambda)^{-1} \int\limits_{u=0}^{u=1} u^{\alpha+1}\,(1-u)^{\lambda-1}\,\delta u$$
$$= \beta(\alpha,\lambda)^{-1} \times \beta(\alpha+2,\lambda) \tag{17}$$

Using Equation (2) above the solution to Equation (17) above is...

$$\mathbb{E}\Bigg[u^2\Bigg] = \frac{\Gamma(\alpha+\lambda)}{\Gamma(\alpha)\,\Gamma(\lambda)} \times \frac{\Gamma(\alpha+2)\,\Gamma(\lambda)}{\Gamma(\alpha+\lambda+2)}$$
$$= \frac{\Gamma(\alpha+\lambda)}{\Gamma(\alpha)\,\Gamma(\lambda)} \times \frac{\alpha\,(\alpha+1)\,\Gamma(\alpha)\,\Gamma(\lambda)}{(\alpha+\lambda)\,(\alpha+\lambda+1)\,\Gamma(\alpha+\lambda)}$$
$$= \frac{\alpha\,(\alpha+1)}{(\alpha+\lambda)\,(\alpha+\lambda+1)} \tag{18}$$

C. Using Equations (6) and (7) above the equation for the variance of the Beta Distribution is...

$$\text{Variance} = \frac{\alpha\,(\alpha+1)}{(\alpha+\lambda)\,(\alpha+\lambda+1)} - \left(\frac{\alpha}{\alpha+\lambda}\right)^2$$
$$= \frac{\alpha\,(\alpha+1)\,(\alpha+\lambda)\,(\alpha+\lambda) - a^2(\alpha+\lambda)\,(\alpha+\lambda+1)}{(\alpha+\lambda)\,(\alpha+\lambda)\,(\alpha+\lambda)\,(\alpha+\lambda+1)}$$
$$= \frac{\alpha\,(\alpha+1)\,(\alpha+\lambda) - a^2(\alpha+\lambda+1)}{(\alpha+\lambda)\,(\alpha+\lambda)\,(\alpha+\lambda+1)}$$
$$= \frac{\alpha^3+\alpha^2\lambda+\alpha^2+\alpha\lambda-\alpha^3-\alpha^2\lambda-\alpha^2}{(\alpha+\lambda)\,(\alpha+\lambda)\,(\alpha+\lambda+1)}$$
$$= \frac{\alpha\lambda}{(\alpha+\lambda)^2(\alpha+\lambda+1)} \tag{19}$$

D. Using Equation (6) above the equation for the parameter λ in terms of the parameter α where the variable M_1 is the known first moment of the Beta Distribution is...

$$M_1 = \frac{\alpha}{\alpha+\lambda} \text{ ...such that... } \lambda = \frac{\alpha\,(1-M_1)}{M_1} \tag{20}$$

E. Using Equations (6) and (7) above the equation for the parameter

$$M_2 = \frac{\alpha\,(\alpha+1)}{(\alpha+\lambda)\,(\alpha+\lambda+1)}$$
$$M_2 = M_1\,\frac{(\alpha+1)}{(\alpha+\lambda+1)}$$
$$\alpha\,M_2 + \lambda\,M_2 + M_2 = \alpha\,M_1 + M_1$$
$$\alpha\,(M_2-M_1) + \lambda\,M_2 = M_1 - M_2 \tag{21}$$

Using Equation (20) above we can make the following definition...

$$\text{if... } \theta = \frac{1-M_1}{M_1} \text{ ...then... } \lambda = \alpha\,\theta \tag{22}$$

Using Equation (22) above we can rewrite Equation (21) above as...

$$\alpha\,(M_2-M_1) + \alpha\,\theta\,M_2 = M_1 - M_2$$
$$\alpha\,(M_2-M_1+\theta\,M_2) = M_1 - M_2$$
$$\alpha = (M_1-M_2)/((1+\theta)\,M_2 - M_1) \tag{23}$$