Gulf Journal of Mathematics ISSN: 2309-4966
Vol 20, Issue 1 (2025) 222-241 https://doi.org/10.56947/gjom.v20i.2820

ESTIMATES OF SOME ENTROPIES AND THEIR RELEVANCE IN FINANCIAL MARKET

DHIRAJ KUMAR SINGH[1*] and HIMANSHU BHATT[2]

Abstract. Using various types of bivariate beta priors such as bivariate beta type - I prior, Conner and Mosimann bivariate beta prior and bivariate beta type - III prior, the Bayesian estimate of the entropy of type - (α) and entropy of type - (α, β) are derived for multinomial likelihood. Bayesian estimates of Shannon entropy for different bivariate beta priors can be represented as a limiting function of Bayesian estimates of entropy of type - (α). In this paper, the parameters of the bivariate beta type-III prior under multinomial likelihood in Moody's corporate bond default rates data are selected using the Bayesian estimate of the entropy type -(α).

1. Introduction

For $m = 1, 2, \ldots$; let $\Gamma_m = \left\{ (\mathsf{p_1}, \ldots, \mathsf{p_m}) : \mathsf{p_i} \geq 0, i = 1, \ldots, m; \sum_{i=1}^{m} \mathsf{p_i} = 1 \right\}$ denote the set of all m-component complete probability distributions with nonnegative elements. For any probability distribution, the entropies

$$H_m(\mathsf{p_1}, \ldots, \mathsf{p_m}) = -\sum_{i=1}^{m} \mathsf{p_i} \log_2 \mathsf{p_i} \tag{1.1}$$

are known as Shannon entropies [21], where $H_m : \Gamma_m \to \mathbb{R}$, $m = 1, 2, \ldots$; $0 \log_2 0 := 0$. The entropies (1.1) are known as additive entropy. Havrda and Charvat [12] defined the entropies of degree α, $\alpha > 0$, $\alpha \neq 1$, $\alpha \in \mathbb{R}$ as

$$H_m^{\alpha}(\mathsf{p_1}, \ldots, \mathsf{p_m}) = (1 - 2^{1-\alpha})^{-1} \left(1 - \sum_{i=1}^{m} \mathsf{p_i}^{\alpha} \right) \tag{1.2}$$

where $H_m^{\alpha} : \Gamma_m \to \mathbb{R}$, $(\mathsf{p_1}, \ldots, \mathsf{p_m}) \in \Gamma_m$, $m = 1, 2, \ldots$. The entropies $H_m^{(\alpha,\beta)} : \Gamma_m \to \mathbb{R}$, $m = 1, 2, \ldots$, of type (α, β) by Behara and Nath [2], defined as

$$H_m^{(\alpha,\beta)}(\mathsf{p_1}, \ldots, \mathsf{p_m}) = \begin{cases} (2^{1-\alpha} - 2^{1-\beta})^{-1} \left(\sum_{i=1}^{m} \mathsf{p_i}^{\alpha} - \sum_{i=1}^{m} \mathsf{p_i}^{\beta} \right) & \text{if } \alpha \neq \beta \\ -2^{\beta-1} \sum_{i=1}^{m} \mathsf{p_i}^{\beta} \log_2 \mathsf{p_i} & \text{if } \alpha = \beta \end{cases} \tag{1.3}$$

where $0^{\alpha} := 0$, $0^{\beta} := 0$, $0^{\beta} \log 0 := 0$, $1^{\alpha} := 1$, $1^{\beta} := 1$ and $\alpha > 0, \beta > 0$.

Date: Received: Mar 31, 2025; Accepted: May 16, 2025.
* Corresponding author.
2020 *Mathematics Subject Classification.* Primary 94A17; Secondary 62F15, 60-11.
Key words and phrases. Entropy, Bayesian estimates of entropy, Multinomial likelihood.

The use of the multinomial distribution and Bayesian methods for entropy estimation has found extensive application across diverse fields such as ecology, genetics, linguistics, survey analysis, marketing, and credit risk analysis. A significant body of literature explores Bayesian approaches to estimate Shannon and Rényi entropies, utilizing various likelihood functions tailored to specific distributions. Notable contributions include works by Bodvin [3] on the Shannon entropy estimation using bayesian method for the bivariate beta distribution, Ren [20] discussing Bayesian entropy estimates for the Inverse Weibull Distribution, and Al-Omari [1] addressing a bayesian approach to determine the Shannon entropy of generalized inverse exponential distribution. Hassan [11] focused on the Bayesian estimate of Shannon entropy for the Lomax distribution based on upper record values, while Chacko and Asha [6] explored estimation of Shannon entropy using maximum likelihood and Bayesian estimation method for generalized exponential distribution. Other studies by Liu and Gui [17] and Shrahili [22] delved into entropy estimation for the two-parameter Lomax distribution and log-logistic distribution and Kumar [13] focused on estimating Tsalli's Entropy for exponential distribution and Tahlan [23] introduced the entropy for doubly truncated random variables known as generalized entropy of order (α, β) respectively. These works collectively demonstrate the adaptability of statistical techniques in capturing the intricacies of different data sets, reflecting a nuanced understanding of entropy in real-world applications.

Traditionally, entropy estimation methods, often based on frequency counts, can exhibit significant bias, particularly with limited data. This can lead to inaccurate assessments of the true entropy. Bayesian methods offer a valuable alternative by incorporating prior knowledge about the underlying probability distribution. This prior information can significantly improve the accuracy and robustness of entropy estimates, especially when dealing with noisy or limited data. Accurate parameter estimation is crucial for many distributions, and entropy estimates often play a key role in this process, motivating the pursuit of accurate Bayesian entropy estimates. As we know, the conjugate prior for multinomial likelihood is the Dirichlet distribution and for the bivariate case, Dirichlet distribution reduces to the bivariate beta type - I distribution.

In this paper, Section 2 is devoted to discussing the Bayesian estimate of entropy of type - (α) and type - (α, β) for different types of priors, such as bivariate beta type - I prior, Conner and Mosimann bivariate beta prior and bivariate beta type - III prior with multinomial likelihood function. Section 3 discusses parameter estimation for bivariate beta type - III distribution on corporate bond default rate data using the Bayesian estimate of the entropy of type - (α) for $\alpha = 2$ as the measure of certainty. In the last section, we discuss the findings of the paper.

2. Bayesian Estimator of Entropy of Type - (α) and type - (α, β)

Let $\mathbf{x} = (x_1, x_2, x_3)$ denote the frequency data generated from the multinomial distribution $MultN(\mathsf{p}_1, \mathsf{p}_2, \mathsf{p}_3)$. Then, the likelihood function is given by

$$\mathsf{f}(\mathbf{x}|\mathsf{p}_1, \mathsf{p}_2) = \frac{n!}{x_1!x_2!x_3!}\mathsf{p}_1{}^{x_1}\mathsf{p}_2{}^{x_2}(1 - \mathsf{p}_1 - \mathsf{p}_2)^{n - x_1 - x_2} \tag{2.1}$$

where $x_1 + x_2 + x_3 = n; 0 < \mathsf{p}_\mathsf{i} < 1$ for $i = 1, 2, 3$ and $0 < \mathsf{p}_1 + \mathsf{p}_2 < 1$, $\mathsf{p}_3 = 1 - \mathsf{p}_1 - \mathsf{p}_2$.

The Loss functions are derived from Bayesian decision theory and try to quantify the loss caused by estimating parameters. It is observed that less loss occurs with better estimators. Martz and Waller [19]; and Lee [16] established that if we assume the quadratic loss function, then for any specified prior distribution, the Bayesian estimate of entropy of type - (α) and entropy of type - (α, β) can be given by the expected value of entropy under the posterior distribution, respectively.

2.1. Bivariate Beta Type - I Prior.

Assume a bivariate beta type - I distribution [4] denoted by $BBeta^I(u_1, u_2, u_3)$ as a prior distribution for parameters in the multinomial distribution $MultN(\mathsf{p}_1, \mathsf{p}_2, \mathsf{p}_3)$ with the corresponding density function:

$$\mathsf{f}(\mathsf{p}_1, \mathsf{p}_2) = \frac{\Gamma(u_1 + u_2 + u_3)}{\Gamma(u_1)\Gamma(u_2)\Gamma(u_3)}\mathsf{p}_1{}^{u_1 - 1}\mathsf{p}_2{}^{u_2 - 1}(1 - \mathsf{p}_1 - \mathsf{p}_2)^{u_3 - 1} \tag{2.2}$$

where $0 < \mathsf{p}_\mathsf{i} < 1$ for $i = 1, 2, 3, 0 < \mathsf{p}_1 + \mathsf{p}_2 < 1, \mathsf{p}_3 = 1 - \mathsf{p}_1 - \mathsf{p}_2; u_1, u_2, u_3 > 0$ and Γ represents the Gamma function.

Using the Bayes theorem, the posterior distribution is given by

$$\mathsf{f}(\mathsf{p}_1, \mathsf{p}_2|\mathbf{x}) = \frac{\mathsf{f}(\mathbf{x}|\mathsf{p}_1, \mathsf{p}_2)\mathsf{f}(\mathsf{p}_1, \mathsf{p}_2)}{\int\int \mathsf{f}(\mathbf{x}|\mathsf{p}_1, \mathsf{p}_2)\mathsf{f}(\mathsf{p}_1, \mathsf{p}_2)d\mathsf{p}_1 d\mathsf{p}_2}. \tag{2.3}$$

From Equations (2.1) and (2.2), we obtain

$$\mathsf{f}(\mathbf{x}|\mathsf{p}_1, \mathsf{p}_2)\mathsf{f}(\mathsf{p}_1, \mathsf{p}_2) = \frac{\Gamma(u_1 + u_2 + u_3)}{\Gamma(u_1)\Gamma(u_2)\Gamma(u_3)}\frac{n!}{x_1!x_2!x_3!}\prod_{i=1}^{2}\mathsf{p}_\mathsf{i}{}^{u_i + x_i - 1}(1 - \mathsf{p}_1 - \mathsf{p}_2)^{u_3 + x_3 - 1}. \tag{2.4}$$

This implies

$$\begin{aligned}\int\int \mathsf{f}(\mathbf{x}|\mathsf{p}_1, \mathsf{p}_2)\mathsf{f}(\mathsf{p}_1, \mathsf{p}_2)d\mathsf{p}_1 d\mathsf{p}_2 = {} & \frac{\Gamma(u_1 + u_2 + u_3)}{\Gamma(u_1)\Gamma(u_2)\Gamma(u_3)}\frac{n!}{x_1!x_2!x_3!} \\ & \times \left[B(u_1 + x_1, u_2 + x_2, u_3 + x_3)\right].\end{aligned} \tag{2.5}$$

From Equations (2.3), (2.4) and (2.5)

$$\mathsf{f}(\mathsf{p}_1, \mathsf{p}_2|\mathbf{x}) = [B(u_1 + x_1, u_2 + x_2, u_3 + x_3)]^{-1}\prod_{i=1}^{2}\mathsf{p}_\mathsf{i}{}^{u_i + x_i - 1}(1 - \mathsf{p}_1 - \mathsf{p}_2)^{u_3 + x_3 - 1} \tag{2.6}$$

where $B(\alpha_1, \alpha_2, \alpha_3) = \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)}{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}$ denotes the beta function.

Theorem 2.1. *Under the quadratic loss function, the Bayesian estimate of entropy of type - (α) $(\alpha > 0, \alpha \neq 1)$ for the multinomial likelihood with respect to the bivariate beta type - I distribution as prior is given by*

$$\hat{H}^I_\alpha = \frac{1}{2^{1-\alpha}-1} \left[\frac{B(\xi_1+\alpha, \xi_2, \xi_3) + B(\xi_1, \xi_2+\alpha, \xi_3) + B(\xi_1, \xi_2, \xi_3+\alpha)}{B(\xi_1, \xi_2, \xi_3)} - 1\right] \tag{2.7}$$

where $\xi_i = u_i + x_i$ for $i = 1, 2, 3$.

Proof. Let the bivariate beta type - I distribution as prior for parameters in multinomial likelihood, then the Bayesian estimate of the entropy of type - (α) $(\alpha > 0, \alpha \neq 1)$ under the quadratic loss function denoted by $\hat{H}^I_\alpha$ is equal to the expected value of entropy of type - (α) with respect to the posterior distribution. Therefore

$$\begin{aligned}
\hat{H}^I_\alpha &= E_{\mathsf{f}(\mathsf{p}_1,\mathsf{p}_2|\mathbf{x})}[H^I_\alpha] \\
&= \int_0^1 \int_0^{1-\mathsf{p}_2} \frac{1}{2^{1-\alpha}-1} \left[\sum_{i=1}^{3} {\mathsf{p}_\mathsf{i}}^\alpha - 1\right] \mathsf{f}(\mathsf{p}_1, \mathsf{p}_2|\mathbf{x}) d\mathsf{p}_1 d\mathsf{p}_2 \\
&= \frac{1}{2^{1-\alpha}-1} \int_0^1 \int_0^{1-\mathsf{p}_2} \left[\sum_{i=1}^{3} {\mathsf{p}_\mathsf{i}}^\alpha - 1\right] [B(u_1+x_1, u_2+x_2, u_3+x_3)]^{-1} \\
&\quad \times {\mathsf{p}_1}^{u_1+x_1-1} {\mathsf{p}_2}^{u_2+x_2-1} (1-\mathsf{p}_1-\mathsf{p}_2)^{u_3+x_3-1} d\mathsf{p}_1 d\mathsf{p}_2 \,.
\end{aligned}$$

If $\xi_i = u_i + x_i$ $(i = 1, 2, 3)$, then

$$\begin{aligned}
\hat{H}^I_\alpha = \frac{[B(\xi_1, \xi_2, \xi_3)]^{-1}}{2^{1-\alpha}-1} &\left[\int_0^1 \int_0^{1-\mathsf{p}_2} {\mathsf{p}_1}^{\xi_1+\alpha-1} {\mathsf{p}_2}^{\xi_2-1} (1-\mathsf{p}_1-\mathsf{p}_2)^{\xi_3-1} d\mathsf{p}_1 d\mathsf{p}_2 \right. \\
&+ \int_0^1 \int_0^{1-\mathsf{p}_2} {\mathsf{p}_1}^{\xi_1-1} {\mathsf{p}_2}^{\xi_2+\alpha-1} (1-\mathsf{p}_1-\mathsf{p}_2)^{\xi_3-1} d\mathsf{p}_1 d\mathsf{p}_2 \\
&+ \int_0^1 \int_0^{1-\mathsf{p}_2} {\mathsf{p}_1}^{\xi_1-1} {\mathsf{p}_2}^{\xi_2-1} (1-\mathsf{p}_1-\mathsf{p}_2)^{\xi_3+\alpha-1} d\mathsf{p}_1 d\mathsf{p}_2 \\
&\left. - \int_0^1 \int_0^{1-\mathsf{p}_2} {\mathsf{p}_1}^{\xi_1-1} {\mathsf{p}_2}^{\xi_2-1} (1-\mathsf{p}_1-\mathsf{p}_2)^{\xi_3-1} d\mathsf{p}_1 d\mathsf{p}_2 \right] \\
= \frac{1}{2^{1-\alpha}-1} &\left[\frac{B(\xi_1+\alpha, \xi_2, \xi_3) + B(\xi_1, \xi_2+\alpha, \xi_3) + B(\xi_1, \xi_2, \xi_3+\alpha)}{B(\xi_1, \xi_2, \xi_3)} - 1\right].
\end{aligned}$$

This completes the proof. □

Corollary 2.2. *Under the quadratic loss function, the Bayesian estimate of the Shannon entropy for the multinomial likelihood with respect to the bivariate beta type - I prior can be given as limiting function of Bayesian estimate of entropy of type - (α) i.e.* $\lim_{\alpha\to 1} \hat{H}^I_\alpha = \hat{H}^I$.

Proof. Let $\hat{H}^I_\alpha$ denote the Bayesian estimate of entropy of type - (α) for multinomial likelihood using bivariate beta type - I prior. Bodvin [3] presented the Bayesian estimate of Shannon entropy $\hat{H}^I$ for multinomial likelihood using bivariate beta type - I prior as

$$\hat{H}^I = \frac{1}{-\log 2}\left[\sum_{i=1}^{3}\frac{\xi_i}{\sum_{j=1}^{3}\xi_j}\left\{\psi(\xi_i+1)-\psi(\sum_{j=1}^{3}\xi_j+1)\right\}\right]$$

where $\xi_i = u_i + x_i$ $(i = 1, 2, 3)$ and $\psi(x) = \frac{\Gamma'(x)}{\Gamma(x)}$.
Now, as α approaches to 1, equation (2.7) gives

$$\begin{aligned}
\lim_{\alpha\to 1}\hat{H}^I_\alpha &= \lim_{\alpha\to 1}\frac{1}{2^{1-\alpha}-1}\\
&\quad\times\left[\frac{B(\xi_1+\alpha,\xi_2,\xi_3)+B(\xi_1,\xi_2+\alpha,\xi_3)+B(\xi_1,\xi_2,\xi_3+\alpha)}{B(\xi_1,\xi_2,\xi_3)}-1\right]\\
&= \lim_{\alpha\to 1}\frac{d}{d\alpha}\left[\frac{B(\xi_1+\alpha,\xi_2,\xi_3)+B(\xi_1,\xi_2+\alpha,\xi_3)+B(\xi_1,\xi_2,\xi_3+\alpha)}{B(\xi_1,\xi_2,\xi_3)}-1\right]\\
&\qquad\times\frac{1}{\frac{d}{d\alpha}(2^{1-\alpha}-1)}\\
&= \lim_{\alpha\to 1}\frac{[B(\xi_1,\xi_2,\xi_3)]^{-1}}{-2^{1-\alpha}\log 2}\left[B(\xi_1+\alpha,\xi_2,\xi_3)\left\{\psi(\xi_1+\alpha)-\psi(\sum_{i=1}^{3}\xi_i+\alpha)\right\}\right.\\
&\quad+B(\xi_1,\xi_2+\alpha,\xi_3)\left\{\psi(\xi_2+\alpha)-\psi(\sum_{i=1}^{3}\xi_i+\alpha)\right\}\\
&\quad\left.+B(\xi_1,\xi_2,\xi_3+\alpha)\left\{\psi(\xi_3+\alpha)-\psi(\sum_{i=1}^{3}\xi_i+\alpha)\right\}\right]\\
&= \frac{[B(\xi_1,\xi_2,\xi_3)]^{-1}}{-\log 2}\frac{\Gamma(\xi_1)\Gamma(\xi_2)\Gamma(\xi_3)}{\Gamma(\xi_1+\xi_2+\xi_3+1)}\\
&\quad\times\left[\sum_{i=1}^{3}\xi_i\left\{\psi(\xi_i+1)-\psi(\sum_{j=1}^{3}\xi_j+1)\right\}\right]\\
&= \frac{1}{-\log 2}\left[\sum_{i=1}^{3}\frac{\xi_i}{\sum_{j=1}^{3}\xi_j}\left\{\psi(\xi_i+1)-\psi(\sum_{j=1}^{3}\xi_j+1)\right\}\right]\\
&= \hat{H}^I.
\end{aligned}$$

Hence, as $\alpha\to 1$, $\hat{H}^I_\alpha$ approaches to $\hat{H}^I$. □

Theorem 2.3. *Under the quadratic loss function, the Bayesian estimate of the entropy of type - (α,β) $(\alpha,\beta > 0, \alpha\neq\beta)$ for the multinomial likelihood with respect to bivariate beta type - I distribution as prior is given by*

$$\begin{aligned}
\hat{H}^I_{(\alpha,\beta)} =& \frac{[B(\xi_1,\xi_2,\xi_3)]^{-1}}{2^{1-\alpha}-2^{1-\beta}}\left[B(\xi_1+\alpha,\xi_2,\xi_3)+B(\xi_1,\xi_2+\alpha,\xi_3)+B(\xi_1,\xi_2,\xi_3+\alpha)\right.\\
&\left.-B(\xi_1+\beta,\xi_2,\xi_3)-B(\xi_1,\xi_2+\beta,\xi_3)-B(\xi_1,\xi_2,\xi_3+\beta)\right]
\end{aligned}$$

where $\xi_i = u_i + x_i$ *for* $i = 1, 2, 3$ *and*

$$\hat{H}^I_\alpha = \hat{H}^I_{(\alpha,1)} \text{ and } \lim_{(\alpha,\beta)\to(1,1)} \hat{H}^I_{(\alpha,\beta)} = \hat{H}^I .$$

Proof. Let the bivariate beta type - I distribution as prior for parameters in multinomial likelihood, then the Bayesian estimate of the entropy of type - (α,β) $(\alpha,\beta > 0, \alpha \neq \beta)$ under the quadratic loss function denoted by $\hat{H}^I_{(\alpha,\beta)}$ is equal to the expected value of entropy of type - (α,β) with respect to the posterior distribution. Therefore using equation (2.6), we have

$$\begin{aligned}
\hat{H}^I_{(\alpha,\beta)} &= \int_0^1 \int_0^{1-\mathsf{p}_2} \frac{1}{2^{1-\alpha} - 2^{1-\beta}} \left[\sum_{i=1}^3 \mathsf{p_i}^\alpha - \sum_{i=1}^3 \mathsf{p_i}^\beta \right] \mathsf{f}(\mathsf{p}_1, \mathsf{p}_2 | \mathbf{x}) d\mathsf{p}_1 d\mathsf{p}_2 \\
&= \frac{[B(u_1 + x_1, u_2 + x_2, u_3 + x_3)]^{-1}}{2^{1-\alpha} - 2^{1-\beta}} \int_0^1 \int_0^{1-\mathsf{p}_2} \left[\sum_{i=1}^3 \mathsf{p_i}^\alpha - \sum_{i=1}^3 \mathsf{p_i}^\beta \right] \\
&\qquad \times \mathsf{p}_1^{u_1+x_1-1} \mathsf{p}_2^{u_2+x_2-1} (1 - \mathsf{p}_1 - \mathsf{p}_2)^{u_3+x_3-1} d\mathsf{p}_1 d\mathsf{p}_2 .
\end{aligned}$$

If $\xi_i = u_i + x_i$ $(i = 1, 2, 3)$, then

$$\begin{aligned}
\hat{H}^I_{(\alpha,\beta)} = \frac{[B(\xi_1, \xi_2, \xi_3)]^{-1}}{2^{1-\alpha} - 2^{1-\beta}} &[B(\xi_1 + \alpha, \xi_2, \xi_3) + B(\xi_1, \xi_2 + \alpha, \xi_3) + B(\xi_1, \xi_2, \xi_3 + \alpha) \\
&-B(\xi_1 + \beta, \xi_2, \xi_3) - B(\xi_1, \xi_2 + \beta, \xi_3) - B(\xi_1, \xi_2, \xi_3 + \beta)] .
\end{aligned}$$

Put $\beta = 1$ in the above equation, we obtain $\hat{H}^I_{(\alpha,1)} = \hat{H}^I_\alpha$.
Let

$$\begin{aligned}
\mathsf{f}(\alpha, \beta) =& B(\xi_1 + \alpha, \xi_2, \xi_3) + B(\xi_1, \xi_2 + \alpha, \xi_3) + B(\xi_1, \xi_2, \xi_3 + \alpha) \\
&- B(\xi_1 + \beta, \xi_2, \xi_3) - B(\xi_1, \xi_2 + \beta, \xi_3) - B(\xi_1, \xi_2, \xi_3 + \beta), \\
\mathsf{g}(\alpha, \beta) =& B(\xi_1, \xi_2, \xi_3)(2^{1-\alpha} - 2^{1-\beta}).
\end{aligned}$$

Thus, using the L'Hopital's rule for multivariable function [15], we have

$$\begin{aligned}
\lim_{(\alpha,\beta)\to(1,1)} \hat{H}^I_{(\alpha,\beta)} = \lim_{(\alpha,\beta)\to(1,1)} \frac{\frac{\partial}{\partial\beta}\mathsf{f}(\alpha,\beta)}{\frac{\partial}{\partial\beta}\mathsf{g}(\alpha,\beta)} &= \frac{\frac{\partial}{\partial\beta}\mathsf{f}(1,1)}{\frac{\partial}{\partial\beta}\mathsf{g}(1,1)} \\
&= \frac{1}{-\log 2} \left[\sum_{i=1}^3 \frac{\xi_i}{\sum_{j=1}^3 \xi_j} \left\{ \psi(\xi_i + 1) - \psi(\sum_{j=1}^3 \xi_j + 1) \right\} \right] .
\end{aligned}$$

Hence, as $(\alpha,\beta) \to (1,1)$, $\hat{H}^I_{(\alpha,\beta)}$ approaches to $\hat{H}^I$. □

2.2. **Conner and Mosimann Bivariate Beta Prior.** Assume Conner and Mosimann [5] bivariate beta distribution denoted by $BBeta^{CM}(u_1, u_2, u_3, d)$ as a prior distribution for parameters in multinomial distribution $MultN(\mathsf{p}_1, \mathsf{p}_2, \mathsf{p}_3)$ with the corresponding density function is

$$\begin{aligned}
\mathsf{f}(\mathsf{p}_1, \mathsf{p}_2) = \frac{\Gamma(u_1 + d)\Gamma(u_2 + u_3)}{\Gamma(u_1)\Gamma(u_2)\Gamma(u_3)\Gamma(d)} \sum_{l=0}^\infty \binom{d - u_2 - u_3}{l} (-1)^l \\
\times \mathsf{p}_1^{u_1+l-1} \mathsf{p}_2^{u_2-1} (1 - \mathsf{p}_1 - \mathsf{p}_2)^{u_3-1}
\end{aligned}$$

where $0 < \mathsf{p_i} < 1$ for $i = 1, 2, 3; 0 < \mathsf{p}_1 + \mathsf{p}_2 < 1; , \mathsf{p}_3 = 1 - \mathsf{p}_1 - \mathsf{p}_2; u_1, u_2, u_3, d > 0$. Using the above equation, Bodvin [3] established that the posterior distribution for the parameters in the multinomial likelihood under the Conner and Mosimann bivariate beta prior is given by

$$\mathsf{f}(\mathsf{p}_1, \mathsf{p}_2|\mathbf{x}) = K \sum_{l=0}^{\infty} \binom{d - u_2 - u_3}{l} (-1)^l \mathsf{p}_1{}^{\xi_1+l-1} \mathsf{p}_2{}^{\xi_2-1} (1 - \mathsf{p}_1 - \mathsf{p}_2)^{\xi_3-1} \tag{2.8}$$

where $K = \frac{\Gamma(\xi_2+\xi_3)\Gamma(u_1+x_1+x_2+x_3+d)}{\Gamma(\xi_1)\Gamma(\xi_2)\Gamma(\xi_3)\Gamma(x_2+x_3+d)}$; $0 < \mathsf{p_i} < 1$, $(i = 1, 2, 3); 0 < \mathsf{p}_1 + \mathsf{p}_2 < 1$; $u_1, u_2, u_3 > 0$ and $\xi_i = u_i + x_i$ $(i = 1, 2, 3)$.

Theorem 2.4. *Under the quadratic loss function, the Bayesian estimate of the entropy of type - (α) $(\alpha > 0, \alpha \neq 1)$ for the multinomial likelihood with respect to the Conner and Mosimann bivariate beta distribution as prior is given by*

$$\begin{aligned}\hat{H}_{\alpha}^{CM} = \frac{K}{2^{1-\alpha} - 1} \sum_{l=0}^{\infty} \binom{d - u_2 - u_3}{l} (-1)^l \, [B(\xi_1 + l + \alpha, \xi_2, \xi_3) \\ + B(\xi_1 + l, \xi_2 + \alpha, \xi_3) + B(\xi_1 + l, \xi_2, \xi_3 + \alpha) - B(\xi_1, \xi_2, \xi_3)]\end{aligned} \tag{2.9}$$

where $\xi_i = u_i + x_i$ $(i = 1, 2, 3)$.

Proof. Let the Conner and Mosimann bivariate beta distribution as prior for parameters in the multinomial likelihood, then the Bayesian estimate of the entropy of type - (α) $(\alpha > 0, \alpha \neq 1)$ under the quadratic loss function is denoted by $\hat{H}_{\alpha}^{CM}$ is equal to the expected value of entropy of type - (α) with respect to the posterior distribution i.e.

$$\begin{aligned}\hat{H}_{\alpha}^{CM} &= E_{\mathsf{f}(\mathsf{p}_1,\mathsf{p}_2|\mathbf{x})}[H_{\alpha}^{CM}] \\ &= \int_0^1 \int_0^{1-\mathsf{p}_2} \frac{1}{2^{1-\alpha} - 1} \left[\sum_{i=1}^{3} \mathsf{p_i}^{\alpha} - 1 \right] \mathsf{f}(\mathsf{p}_1, \mathsf{p}_2|\mathbf{x}) d\mathsf{p}_1 d\mathsf{p}_2 \,.\end{aligned}$$

Now, using equation (2.8), the above equation can be written as

$$\begin{aligned}\hat{H}_{\alpha}^{CM} &= \frac{K}{2^{1-\alpha} - 1} \sum_{l=0}^{\infty} \binom{d - u_2 - u_3}{l} (-1)^l \left[\int_0^1 \int_0^{1-\mathsf{p}_2} \left(\sum_{i=1}^{3} \mathsf{p_i}^{\alpha} - 1 \right) \right. \\ &\qquad \left. \times \, \mathsf{p}_1{}^{\xi_1+l-1} \mathsf{p}_2{}^{\xi_2-1} (1 - \mathsf{p}_1 - \mathsf{p}_2)^{\xi_3-1} \right] \\ &= \frac{K}{2^{1-\alpha} - 1} \sum_{l=0}^{\infty} \binom{d - u_2 - u_3}{l} (-1)^l \, [B(\xi_1 + l + \alpha, \xi_2, \xi_3) \\ &\qquad + B(\xi_1 + l, \xi_2 + \alpha, \xi_3) + B(\xi_1 + l, \xi_2, \xi_3 + \alpha) - B(\xi_1, \xi_2, \xi_3) \,] \,.\end{aligned}$$

This completes the proof. □

Corollary 2.5. *Under the quadratic loss function, the Bayesian estimate of the Shannon entropy for the multinomial likelihood with respect to the Conner and Mosimann bivariate beta prior can be written as limiting function of Bayesian estimate of entropy of type - (α) i.e. $\lim_{\alpha \to 1} \hat{H}_{\alpha}^{CM} = \hat{H}^{CM}$.*

Proof. Let $\hat{H}_{\alpha}^{CM}$ denote the Bayesian estimate of entropy of type-(α) for multinomial likelihood using Conner and Mosimann bivariate beta prior. As Bodvin [3] discussed, the Bayesian estimate of Shannon entropy $\hat{H}^{CM}$ for multinomial likelihood using the Conner and Mosimann bivariate beta prior is equal to

$$\hat{H}^{CM} = -\frac{K}{\log 2}\sum_{l=0}^{\infty}\binom{d-u_2-u_3}{l}(-1)^l\frac{\Gamma(\tau_1)\Gamma(\tau_2)\Gamma(\tau_3)}{\Gamma(\tau_1+\tau_2+\tau_3+1)} \times \sum_{i=1}^{3}\tau_i[\psi(\tau_i+1)-\psi(\sum_{j=1}^{3}\tau_j+1)]$$

where $\tau_1 = u_1 + x_1 + l, \tau_2 = u_2 + x_2, \tau_3 = u_3 + x_3$ and $\psi(x) = \frac{\Gamma'(x)}{\Gamma(x)}$.
From Equation (2.9), we obtain

$$\begin{aligned}
\lim_{\alpha\to 1}\hat{H}_{\alpha}^{CM} &= \lim_{\alpha\to 1}\frac{K}{2^{1-\alpha}-1}\sum_{l=0}^{\infty}\binom{d-u_2-u_3}{l}(-1)^l\,[B(\xi_1+l+\alpha,\xi_2,\xi_3) \\
&\quad + B(\xi_1+l,\xi_2+\alpha,\xi_3) + B(\xi_1+l,\xi_2,\xi_3+\alpha) - B(\xi_1,\xi_2,\xi_3)\,] \\
&= \lim_{\alpha\to 1}\frac{K}{-2^{1-\alpha}\log 2}\sum_{l=0}^{\infty}\binom{d-u_2-u_3}{l}(-1)^l\left[\frac{d}{d\alpha}B(\xi_1+l+\alpha,\xi_2,\xi_3)\right. \\
&\quad \left. + \frac{d}{d\alpha}B(\xi_1+l,\xi_2+\alpha,\xi_3) + \frac{d}{d\alpha}B(\xi_1+l,\xi_2,\xi_3+\alpha) - \frac{d}{d\alpha}B(\xi_1,\xi_2,\xi_3)\right] \\
&= \lim_{\alpha\to 1}\frac{K}{-2^{1-\alpha}\log 2}\sum_{l=0}^{\infty}\binom{d-u_2-u_3}{l}(-1)^l \\
&\quad \times\left[B(\xi_1+l+\alpha,\xi_2,\xi_3)\left\{\psi(\xi_1+l+\alpha)-\psi(\sum_{i=1}^{3}\xi_i+l+\alpha)\right\}\right. \\
&\quad + B(\xi_1+l,\xi_2+\alpha,\xi_3)\left\{\psi(\xi_2+\alpha)-\psi(\sum_{i=1}^{3}\xi_i+l+\alpha)\right\} \\
&\quad \left. + B(\xi_1+l,\xi_2,\xi_3+\alpha)\left\{\psi(\xi_3+\alpha)-\psi(\sum_{i=1}^{3}\xi_i+l+\alpha)\right\}\right] \\
&= \frac{K}{-\log 2}\sum_{l=0}^{\infty}\binom{d-u_2-u_3}{l}(-1)^l\frac{\Gamma(\xi_1+l)\Gamma(\xi_2)\Gamma(\xi_3)}{\Gamma(\xi_1+\xi_2+\xi_3+l+1)} \\
&\quad \times\left[(\xi_1+l)\left\{\psi(\xi_1+l+1)-\psi(\sum_{i=1}^{3}\xi_i+l+\alpha)\right\}\right. \\
&\quad + \xi_2\left\{\psi(\xi_2+1)-\psi(\sum_{i=1}^{3}\xi_i+l+1)\right\}
\end{aligned}$$

$$
\begin{aligned}
&+ \xi_3 \left\{ \psi(\xi_3+1) - \psi(\sum_{i=1}^{3} \xi_i + l + 1) \right\} \Bigg] \\
= \ & - \frac{K}{\log 2} \sum_{l=0}^{\infty} \binom{d-u_2-u_3}{l} (-1)^l \frac{\Gamma(\tau_1)\Gamma(\tau_2)\Gamma(\tau_3)}{\Gamma(\tau_1+\tau_2+\tau_3+1)} \\
&\times \sum_{i=1}^{3} \tau_i [\psi(\tau_i+1) - \psi(\sum_{j=1}^{3} \tau_j + 1)]
\end{aligned}
$$

where $\tau_1 = u_1 + x_1 + l, \tau_2 = u_2 + x_2, \tau_3 = u_3 + x_3$. □

Theorem 2.6. *Under the quadratic loss function, the Bayesian estimate of the entropy of type - (α, β) $(\alpha, \beta > 0, \alpha \neq \beta)$ for multinomial likelihood with respect to Conner and Mosimann bivariate beta distribution as prior is given by*

$$
\begin{aligned}
\hat{H}^{CM}_{(\alpha,\beta)} = & \frac{K}{2^{1-\alpha} - 2^{1-\beta}} \sum_{l=0}^{\infty} \binom{d-u_2-u_3}{l} (-1)^l \\
& \times [B(\xi_1+l+\alpha, \xi_2, \xi_3) + B(\xi_1+l, \xi_2+\alpha, \xi_3) + B(\xi_1+l, \xi_2, \xi_3+\alpha) \\
& \quad - B(\xi_1+l+\beta, \xi_2, \xi_3) - B(\xi_1+l, \xi_2+\beta, \xi_3) - B(\xi_1+l, \xi_2, \xi_3+\beta)]
\end{aligned} \tag{2.10}
$$

where $\xi_i = u_i + x_i$ for $i = 1, 2, 3$ and

$$\hat{H}^{CM}_{\alpha} = \hat{H}^{CM}_{(\alpha,1)} \ \textit{and} \lim_{(\alpha,\beta)\to(1,1)} \hat{H}^{CM}_{(\alpha,\beta)} = \hat{H}^{CM}.$$

Proof. Let the Conner and Mosimann bivariate beta distribution as prior for parameters in the multinomial likelihood, then the Bayesian estimate of the entropy of type - (α, β) $(\alpha, \beta > 0, \alpha \neq \beta)$ under the quadratic loss function is denoted by $\hat{H}^{CM}_{(\alpha,\beta)}$ is equal to the expected value of entropy of type - (α, β) with respect to the posterior distribution. Therefore using equation (2.8), we have

$$
\begin{aligned}
\hat{H}^{CM}_{(\alpha,\beta)} = & \frac{K}{2^{1-\alpha} - 2^{1-\beta}} \left[\int_0^1 \int_0^{1-\mathsf{p}_2} \left(\sum_{i=1}^{3} \mathsf{p_i}^{\alpha} - \sum_{i=1}^{3} \mathsf{p_i}^{\beta} \right) \right. \\
& \left. \times \sum_{l=0}^{\infty} \binom{d-u_2-u_3}{l} (-1)^l \mathsf{p_1}^{\xi_1+l-1} \mathsf{p_2}^{\xi_2-1} (1-\mathsf{p}_1-\mathsf{p}_2)^{\xi_3-1} d\mathsf{p}_1 d\mathsf{p}_2 \right].
\end{aligned}
$$

where $0 < \mathsf{p}_1, \mathsf{p}_2 < 1, 0 < \mathsf{p}_1 + \mathsf{p}_2 < 1, \mathsf{p}_3 = 1 - \mathsf{p}_1 - \mathsf{p}_2$ and the infinite series involved in the above converges uniformly for all $(\mathsf{p}_1, \mathsf{p}_2, \mathsf{p}_3) \in \Gamma_3$. Therefore, we can interchange the summation and integral sign. Thus, we obtain the equation (2.10). Put $\beta = 1$ in the equation (2.10), we have $\hat{H}^{CM}_{(\alpha,1)} = \hat{H}^{CM}_{\alpha}$. Let

$$
\begin{aligned}
\mathsf{f}(\alpha, \beta) = K \sum_{l=0}^{\infty} & \binom{d-u_2-u_3}{l} (-1)^l [B(\xi_1+l+\alpha, \xi_2, \xi_3) + B(\xi_1+l, \xi_2+\alpha, \xi_3) \\
& + B(\xi_1+l, \xi_2, \xi_3+\alpha) - B(\xi_1+l+\beta, \xi_2, \xi_3) \\
& - B(\xi_1+l, \xi_2+\beta, \xi_3) - B(\xi_1+l, \xi_2, \xi_3+\beta)],
\end{aligned}
$$

$\mathsf{g}(\alpha, \beta) = 2^{1-\alpha} - 2^{1-\beta}$.

Using the L'Hopital's rule for multivariable function [15], we have

$$\lim_{(\alpha,\beta)\to(1,1)} \hat{H}^{CM}_{(\alpha,\beta)} = \lim_{(\alpha,\beta)\to(1,1)} \frac{\mathsf{f}(\alpha,\beta)}{\mathsf{g}(\alpha,\beta)} = \lim_{(\alpha,\beta)\to(1,1)} \frac{\frac{\partial \mathsf{f}(\alpha,\beta)}{\partial\beta}}{\frac{\partial \mathsf{g}(\alpha,\beta)}{\partial\beta}} = \frac{\frac{\partial \mathsf{f}}{\partial\beta}(1,1)}{\frac{\partial \mathsf{g}}{\partial\beta}(1,1)}$$
$$= -\frac{K}{\log 2}\sum_{l=0}^{\infty}\binom{d-u_2-u_3}{l}(-1)^l\frac{\Gamma(\tau_1)\Gamma(\tau_2)\Gamma(\tau_3)}{\Gamma(\tau_1+\tau_2+\tau_3+1)}$$
$$\times\sum_{i=1}^{3}\tau_i[\psi(\tau_i+1)-\psi(\sum_{j=1}^{3}\tau_j+1)]$$

where $\tau_1 = u_1 + x_1 + l, \tau_2 = u_2 + x_2, \tau_3 = u_3 + x_3$. Hence, as $(\alpha, \beta) \to (1, 1)$, $\hat{H}^{CM}_{(\alpha,\beta)}$ approaches to $\hat{H}^{CM}$. □

2.3. **Bivariate Beta Type - III Prior.** Assume that bivariate beta type - III distribution given by Ehlers [8] as a prior distribution for parameters in multinomial likelihood $MultN(\mathsf{p}_1, \mathsf{p}_2, \mathsf{p}_3)$ denoted by $BBeta^{III}(u_1, u_2, u_3, c)$ with the corresponding density function is

$$\mathsf{f}(\mathsf{p}_1,\mathsf{p}_2) = \frac{\Gamma(u_1+u_2+u_3)}{\Gamma(u_1)\Gamma(u_2)\Gamma(u_3)}c^{u_1+u_2}{\mathsf{p}_1}^{u_1-1}{\mathsf{p}_2}^{u_2-1}(1-\mathsf{p}_1-\mathsf{p}_2)^{u_3-1}$$
$$\times[1-(1-c)\mathsf{p}_1-(1-c)\mathsf{p}_2]^{-(u_1+u_2+u_3)}$$

where $0 < \mathsf{p}_i < 1$ for $i = 1, 2; 0 < \mathsf{p}_1 + \mathsf{p}_2 < 1, u_1, u_2, u_3, c > 0$. From the above equation, Bodvin [3] established that the posterior distribution for the multinomial likelihood under bivariate beta type - III prior is given by

$$\mathsf{f}(\mathsf{p}_1,\mathsf{p}_2|\mathbf{x}) = K\sum_{l=0}^{\infty}\sum_{m=0}^{l}\binom{-(u_1+u_2+u_3)}{l}\binom{l}{m}$$
$$\times(c-1)^l{\mathsf{p}_1}^{\xi_1+l-m-1}{\mathsf{p}_2}^{\xi_2+m-1}(1-\mathsf{p}_1-\mathsf{p}_2)^{\xi_3-1}$$

where

$$K = \frac{\Gamma(\xi_1+\xi_2+\xi_3)}{\Gamma(\xi_1)\Gamma(\xi_2)\Gamma(\xi_3)}c^{(u_1+u_2+u_3)}\left[{}_2F_1(u_1+u_2+u_3,\xi_3;\xi_1+\xi_2+\xi_3;\frac{c-1}{c})\right]^{-1}$$

and ${}_2F_1(.)$ represents the Gauss hypergeometric function [10, 18] and $\xi_i = u_i + x_i (i = 1, 2, 3)$ and $|c - 1| < \frac{1}{\mathsf{p}_1+\mathsf{p}_2}$.

Theorem 2.7. *Under the quadratic loss function, the Bayesian estimate of the entropy of type - (α) $(\alpha > 0, \alpha \neq 1)$ for the multinomial likelihood using bivariate beta type - III distribution as prior is given by*

$$\begin{aligned}\hat{H}_{\alpha}^{III} =& \frac{K}{2^{1-\alpha}-1}\sum_{l=0}^{\infty}\sum_{m=0}^{l}\binom{-(u_1+u_2+u_3)}{l}\binom{l}{m}(c-1)^l \\ & \times\Big[B(\xi_1+l+\alpha-m,\xi_2+m,\xi_3)+B(\xi_1+l-m,\xi_2+\alpha+m,\xi_3) \\ & \quad + B(\xi_1+l-m,\xi_2+m,\xi_3+\alpha)-B(\xi_1+l-m,\xi_2+m,\xi_3)\Big]\end{aligned} \tag{2.11}$$

where $\xi_i = u_i + x_i\,(i = 1, 2, 3)$.

Proof. Let the bivariate beta type - III distribution as prior for parameters in multinomial likelihood, then the Bayesian estimate of the entropy of type - (α) ($\alpha > 0, \alpha \neq 1$) under the quadratic loss function denoted by $\hat{H}_{\alpha}^{III}$ is equal to the expected value of entropy of type - (α) with respect to the posterior distribution i.e.

$$\begin{aligned}\hat{H}_{\alpha}^{III} =& E_{\mathsf{f}(\mathsf{p}_1,\mathsf{p}_2|\mathbf{x})}[H_{\alpha}^{III}] \\ =& \int_0^1\int_0^{1-\mathsf{p}_2}\frac{1}{2^{1-\alpha}-1}\left[\sum_{i=1}^{3}\mathsf{p_i}^{\alpha}-1\right]\mathsf{f}(\mathsf{p}_1,\mathsf{p}_2|\mathbf{x})d\mathsf{p}_1 d\mathsf{p}_2 \\ =& \frac{K}{2^{1-\alpha}-1}\sum_{l=0}^{\infty}\sum_{m=0}^{l}\binom{-(u_1+u_2+u_3)}{l}\binom{l}{m}(c-1)^l \\ & \times\Big[\int_0^1\int_0^{1-\mathsf{p}_2}\mathsf{p}_1^{\xi_1+l+\alpha-m-1}\mathsf{p}_2^{\xi_2+m-1}(1-\mathsf{p}_1-\mathsf{p}_2)^{\xi_3-1}d\mathsf{p}_1 d\mathsf{p}_2 \\ & \quad +\int_0^1\int_0^{1-\mathsf{p}_2}\mathsf{p}_1^{\xi_1+l-m-1}\mathsf{p}_2^{\xi_2+\alpha+m-1}(1-\mathsf{p}_1-\mathsf{p}_2)^{\xi_3-1}d\mathsf{p}_1 d\mathsf{p}_2 \\ & \quad +\int_0^1\int_0^{1-\mathsf{p}_2}\mathsf{p}_1^{\xi_1+l-m-1}\mathsf{p}_2^{\xi_2+m-1}(1-\mathsf{p}_1-\mathsf{p}_2)^{\xi_3+\alpha-1}d\mathsf{p}_1 d\mathsf{p}_2 \\ & \quad -\int_0^1\int_0^{1-\mathsf{p}_2}\mathsf{p}_1^{\xi_1+l-m-1}\mathsf{p}_2^{\xi_2+m-1}(1-\mathsf{p}_1-\mathsf{p}_2)^{\xi_3-1}d\mathsf{p}_1 d\mathsf{p}_2\Big] \\ =& \frac{K}{2^{1-\alpha}-1}\sum_{l=0}^{\infty}\sum_{m=0}^{l}\binom{-(u_1+u_2+u_3)}{m}\binom{l}{m}(c-1)^l \\ & \times\Big[B(\xi_1+l+\alpha-m,\xi_2+m,\xi_3)+B(\xi_1+l-m,\xi_2+\alpha+m,\xi_3) \\ & \quad + B(\xi_1+l-m,\xi_2+m,\xi_3+\alpha)-B(\xi_1+l-m,\xi_2+m,\xi_3)\Big]\end{aligned}$$

where $\xi_i = u_i + x_i$ for $i = 1, 2, 3$. Hence, equation (2.11) established. □

Corollary 2.8. *Under the quadratic loss function, the Bayesian estimate of the Shannon entropy for the multinomial likelihood with respect to the bivariate beta* *type - III prior can be written as the limiting function of Bayesian estimate of entropy of type - (α) i.e.* $\lim_{\alpha\to 1}\hat{H}_{\alpha}^{III} = \hat{H}^{III}$.

Proof. Let $\hat{H}_{\alpha}^{III}$ denote the Bayesian estimate of entropy of type - (α) for multinomial likelihood using bivariate beta type - III prior. As Bodvin [3] discussed, the Bayesian estimate of Shannon entropy $\hat{H}^{III}$ for multinomial likelihood using bivariate beta type - III prior is given by

$$\hat{H}^{III} = \frac{-K}{\log 2}\sum_{l=0}^{\infty}\sum_{m=0}^{l}\binom{-(u_1+u_2+u_3)}{l}\binom{l}{m}(c-1)^l \times \frac{\Gamma(\eta_1)\Gamma(\eta_2)\Gamma(\eta_3)}{\Gamma(\eta_1+\eta_2+\eta_3+1)}\left[\sum_{i=1}^{3}\eta_i\left\{\psi(\eta_i+1)-\psi(\sum_{j=1}^{3}\eta_j+1)\right\}\right]$$

where $\eta_1 = u_1+x_1+l-m, \eta_2 = u_2+x_2+m, \eta_3 = u_3+x_3$ and $\psi(x) = \frac{\Gamma'(x)}{\Gamma(x)}$.
From Equation (2.11), we have

$$\begin{aligned}
\lim_{\alpha\to 1}\hat{H}_{\alpha}^{III} &= \lim_{\alpha\to 1}\frac{K}{2^{1-\alpha}-1}\sum_{l=0}^{\infty}\sum_{m=0}^{l}\binom{-(u_1+u_2+u_3)}{l}\binom{l}{m}(c-1)^l \\
&\quad\times [B(\eta_1+\alpha,\eta_2,\eta_3)+B(\eta_1,\eta_2+\alpha,\eta_3)+B(\eta_1,\eta_2,\eta_3+\alpha) \\
&\qquad -B(\eta_1,\eta_2,\eta_3)] \\
&= \lim_{\alpha\to 1}\frac{K}{-2^{1-\alpha}\log 2}\sum_{l=0}^{\infty}\sum_{m=0}^{l}\binom{-(u_1+u_2+u_3)}{l}\binom{l}{m}(c-1)^l \\
&\quad\times\Bigg[B(\eta_1+\alpha,\eta_2,\eta_3)\left\{\psi(\eta_1+\alpha)-\psi(\sum_{i=1}^{3}\eta_i+\alpha)\right\} \\
&\qquad + B(\eta_1,\eta_2+\alpha,\eta_3)\left\{\psi(\eta_2+\alpha)-\psi(\sum_{i=1}^{3}\eta_i+\alpha)\right\} \\
&\qquad + B(\eta_1,\eta_2,\eta_3+\alpha)\left\{\psi(\eta_3+\alpha)-\psi(\sum_{i=1}^{3}\eta_i+\alpha)\right\}-0\Bigg] \\
&= \frac{-K}{\log 2}\sum_{l=0}^{\infty}\sum_{m=0}^{l}\binom{-(u_1+u_2+u_3)}{l}\binom{l}{m}(c-1)^l\frac{\Gamma(\eta_1)\Gamma(\eta_2)\Gamma(\eta_3)}{\Gamma(\eta_1+\eta_2+\eta_3+1)} \\
&\quad\times\Bigg[\eta_1\left\{\psi(\eta_1+1)-\psi(\sum_{i=1}^{3}\eta_i+1)\right\}+\eta_2\left\{\psi(\eta_2+1)-\psi(\sum_{i=1}^{3}\eta_i+1)\right\} \\
&\qquad + \eta_3\left\{\psi(\eta_3+1)-\psi(\sum_{i=1}^{3}\eta_i+1)\right\}\Bigg] \\
&= \frac{-K}{\log 2}\sum_{l=0}^{\infty}\sum_{m=0}^{l}\binom{-(u_1+u_2+u_3)}{l}\binom{l}{m}(c-1)^l
\end{aligned}$$

$$\times \frac{\Gamma(\eta_1)\Gamma(\eta_2)\Gamma(\eta_3)}{\Gamma(\eta_1+\eta_2+\eta_3+1)}\left[\sum_{i=1}^{3}\eta_i\left\{\psi(\eta_i+1)-\psi(\sum_{j=1}^{3}\eta_j+1)\right\}\right]$$

where $\eta_1 = u_1 + x_1 + l - m, \eta_2 = u_2 + x_2 + m, \eta_3 = u_3 + x_3$. Thus, we have concluded that as $\alpha \to 1$, $\hat{H}^{III}_{\alpha}$ approaches to $\hat{H}^{III}$. □

Theorem 2.9. *Under the quadratic loss function, the Bayesian estimate of entropy of type - (α,β) $(\alpha,\beta > 0,\ \alpha \neq \beta)$ for the multinomial likelihood using bivariate beta type - III distribution as prior is given by*

$$\begin{aligned}\hat{H}^{III}_{(\alpha,\beta)} = \frac{K}{2^{1-\alpha}-2^{1-\beta}}\sum_{l=0}^{\infty}\sum_{m=0}^{l}\binom{-(u_1+u_2+u_3)}{l}\binom{l}{m}(c-1)^l\\ \times [\ B(\xi_1+l+\alpha-m,\xi_2+m,\xi_3)+B(\xi_1+l-m,\xi_2+m+\alpha,\xi_3)\\ +B(\xi_1+l-m,\xi_2+m,\xi_3+\alpha)-B(\xi_1+l+\beta-m,\xi_2+m,\xi_3)\\ -B(\xi_1+l-m,\xi_2+m+\beta,\xi_3)-B(\xi_1+l-m,\xi_2+m,\xi_3+\beta)\ ]\end{aligned} \tag{2.12}$$

where $\xi_i = u_i + x_i$ for $i = 1, 2, 3$ are the parameter of the posterior distribution and

$$\hat{H}^{III}_{\alpha} = \hat{H}^{III}_{(\alpha,1)} \text{ and } \lim_{(\alpha,\beta)\to(1,1)}\hat{H}^{III}_{(\alpha,\beta)} = \hat{H}^{III}.$$

Proof. Let the bivariate beta type - III distribution as prior for parameters in multinomial likelihood, then the Bayesian estimate of the entropy of type - (α,β) $(\alpha,\beta > 0,\alpha \neq \beta)$ under the quadratic loss function denoted by $\hat{H}^{III}_{(\alpha,\beta)}$ is equal to the expected value of entropy of type - (α,β) with respect to the posterior distribution i.e.

$$\begin{aligned}\hat{H}^{III}_{(\alpha,\beta)} = \int_0^1\int_0^{1-\mathsf{p}_2}\frac{1}{2^{1-\alpha}-2^{1-\beta}}\left[\sum_{i=1}^{3}\mathsf{p_i}^{\alpha}-\sum_{i=1}^{3}\mathsf{p_i}^{\beta}\right]\\ \times K\sum_{l=0}^{\infty}\sum_{m=0}^{l}\binom{-(u_1+u_2+u_3)}{l}\binom{l}{m}(c-1)^l\mathsf{p_1}^{\xi_1+l-m-1}\\ \times \mathsf{p_2}^{\xi_2+m-1}(1-\mathsf{p}_1-\mathsf{p}_2)^{\xi_3-1}d\mathsf{p}_1 d\mathsf{p}_2.\end{aligned}$$

Since convergence in the above infinite series is uniform for all for all $(\mathsf{p}_1,\mathsf{p}_2,\mathsf{p}_3) \in \Gamma_3$. Therefore, we can interchange the integral and summation sign. Thus, we obtain the equation (2.12). Put $\beta = 1$ in the equation (2.12) , we have $\hat{H}^{III}_{(\alpha,1)} = \hat{H}^{III}_{\alpha}$. Let

$$\begin{aligned}\mathsf{f}(\alpha,\beta) = K\sum_{l=0}^{\infty}\sum_{m=0}^{l}\binom{-(u_1+u_2+u_3)}{l}\binom{l}{m}(c-1)^l\\ \times [\ B(\xi_1+l+\alpha-m,\xi_2+m,\xi_3)+B(\xi_1+l-m,\xi_2+m+\alpha,\xi_3)\\ +B(\xi_1+l-m,\xi_2+m,\xi_3+\alpha)-B(\xi_1+l+\beta-m,\xi_2+m,\xi_3)\\ -B(\xi_1+l-m,\xi_2+m+\beta,\xi_3)-B(\xi_1+l-m,\xi_2+m,\xi_3+\beta)\ ],\end{aligned}$$

$$\mathsf{g}(\alpha,\beta) = 2^{1-\alpha} - 2^{1-\beta}.$$

Using the L'Hopital's rule for multivariable function [15], we have

$$\begin{aligned}\lim_{(\alpha,\beta)\to(1,1)} \hat{H}^{III}_{(\alpha,\beta)} &= \lim_{(\alpha,\beta)\to(1,1)} \frac{\mathsf{f}(\alpha,\beta)}{\mathsf{g}(\alpha,\beta)} = \lim_{(\alpha,\beta)\to(1,1)} \frac{\frac{\partial}{\partial\beta}\mathsf{f}(\alpha,\beta)}{\frac{\partial}{\partial\beta}\mathsf{g}(\alpha,\beta)} = \frac{\frac{\partial \mathsf{f}}{\partial\beta}(1,1)}{\frac{\partial \mathsf{g}}{\partial\beta}(1,1)} \\ &= \frac{-K}{\log 2} \sum_{l=0}^{\infty}\sum_{m=0}^{l} \binom{-(u_1+u_2+u_3)}{l}\binom{l}{m}(c-1)^l \\ &\quad \times \frac{\Gamma(\eta_1)\Gamma(\eta_2)\Gamma(\eta_3)}{\Gamma(\eta_1+\eta_2+\eta_3+1)}\left[\sum_{i=1}^{3}\eta_i\left\{\psi(\eta_i+1)-\psi(\sum_{j=1}^{3}\eta_j+1)\right\}\right]\end{aligned}$$

where $\eta_1 = \xi_1 + l - m, \eta_2 = \xi_2 + m, \eta_3 = \xi_3$. Hence, as $(\alpha,\beta) \to (1,1)$, $\hat{H}^{III}_{(\alpha,\beta)}$ approaches to $\hat{H}^{III}$. $\square$

3. Corporate Bonds Default Risk

Corporate bonds are debt securities issued by private and public corporations. There are two kinds of corporate bonds: Investment and Speculative. Investment Grade bonds are thought to have a lesser risk of default and obtain higher ratings from credit rating agencies. Speculative Grade bonds often provide higher interest rates to reward investors for taking on more risk. The likelihood that a firm may fail to make timely interest or principal installments and consequently default on its bonds is known as credit or default risk. The corporate bond default rate is the actual number of corporate bonds that have defaulted out of a total number of corporate bonds from a certain population of companies.

This section will analyze corporate bond default risks using Bayesian criteria. We are using the Entropy of type - (α) to choose the prior distribution in the Bayesian credit risk model calibration technique. In the context of corporate bonds, it is assumed that there are three events:

- Event 1: Default occurs in investment grade.
- Event 2: Default occurs in speculative grade.
- Event 3: Default does not occur.

We assume that the above three events follow the multinomial distribution as in equation (2.1) and bivariate beta distribution is followed by the parameters in the multinomial distribution. As far as the occurrence of default in a particular grade is concerned, selecting the best accurate bivariate beta distribution for the parameters in multinomial distribution is critical because it directly impacts a bond default risk.

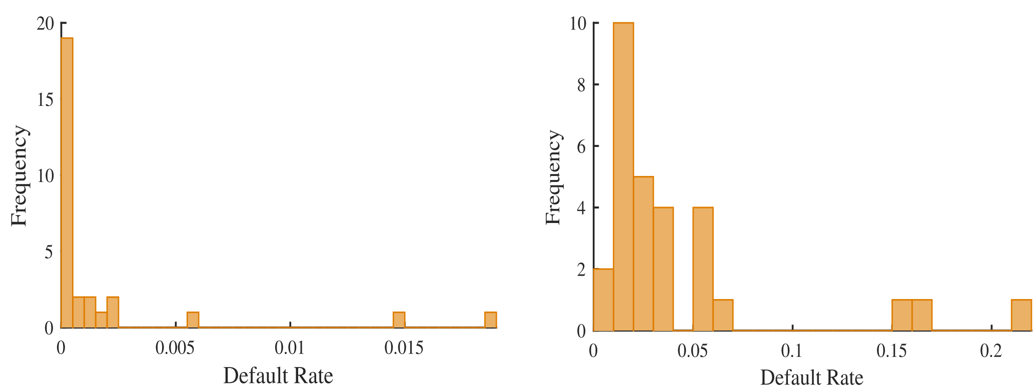
3.1. **Data.** The data for annual volume-weighted corporate bond default rates in Table: 1 is used for analysis obtained from the Moody's "Default Trends – Global Annual Default Study: Corporate default rate will rise in 2023 and peak in early 2024" [7] and corporate bonds default rate for each year is used from 1994 to 2022.

For the analysis, default rates of investment grade and speculative grade are used. It is to be noted that there are a few years in which no default occurred in investment grade, whereas the assumption in bivariate beta distribution is $p_1, p_2 > 0$. However, it does not violate the assumption because, in practice, it is believed that a bond can not have a default rate of 0.

TABLE 1. Moody's Corporate Bond Default Rates (1994-2022)

Year	Investment Grade	Speculative Grade	Year	Investment Grade	Speculative Grade
1994	0	0.0213	2009	0.0022	0.1615
1995	0	0.0308	2010	0.0008	0.017
1996	0	0.0234	2011	0.0015	0.0146
1997	0	0.0191	2012	0.0001	0.0205
1998	0	0.0287	2013	0.0004	0.0109
1999	0.0003	0.0588	2014	0.0001	0.0174
2000	0.0014	0.0563	2015	0	0.0348
2001	0.0055	0.1524	2016	0	0.0327
2002	0.0186	0.2156	2017	0	0.0182
2003	0	0.0595	2018	0	0.0192
2004	0	0.0188	2019	0.0022	0.0282
2005	0.0007	0.0379	2020	0.0001	0.0626
2006	0	0.0105	2021	0	0.0031
2007	0	0.0082	2022	0.0014	0.0155
2008	0.0147	0.058			

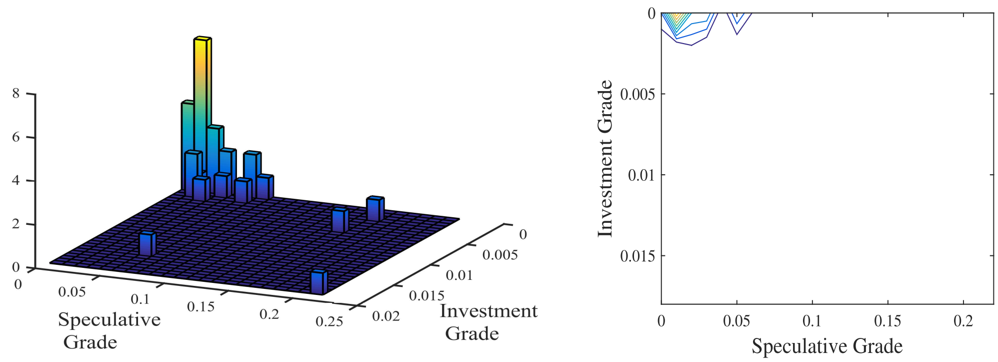
3.2. **Analysis.** For visualization of default rates, Figure: 1 depicts the univariate distribution of corporate bond default rates, and the figure conveys a very high density of the distribution towards the low default rates in the investment grade. Likewise, speculative-grade data is also highly distributed to low default rates. From the joint distribution of the corporate bond default



(A) Investment Grade (B) Speculative Grade

FIGURE 1. Univariate Default Rate Distribution

rates, the linear correlation between the default rates of investment and speculative grade is positive and equal to 0.6885. Bodvin [3] discussed correlations for different types of bivariate beta priors. Here, the correlation between investment and speculative grade is positive; therefore, bivariate beta type - III and bivariate beta type - V distributions [9] are suitable priors for the model. We do not consider bivariate beta type - I and Conner and Mosimann bivariate beta prior because they induce negative correlation between p_1 and p_2.



(A) Joint Data Distribution (B) Contour Plot

FIGURE 2. Joint Distribution of Investment and Speculative Grade Default Rates

3.3. **Bayesian Estimation of Entropy of Type - (α).** The proposal is to use the Entropy of type - (α) for $\alpha = 2$, available data, and expert judgment to identify the best values of the parameters for the bivariate beta type - III prior with the multinomial model having dimension 3. Bayesian priors are typically generated in the financial industries by utilizing opinions from professionals. However, for illustration in the study, the prior will be created using default rate data. Because of the capacity to adjust for positive correlation, bivariate beta type - III is used as viable prior in this application. The parameters are determined using the steps outlined below:

- Identify the order of magnitude of the parameters based on the results of the shape studies for the bivariate beta type - III distribution. As for the corporate bonds default rate data, the distribution is concentrated toward small values of investment-grade default rates (p_1) and small values of speculative-grade default rates (p_2). This depicts that for the bivariate beta type - III distribution, the parameters u_1, u_2 and u_3 be smaller than c.
- Establish parameter bands by applying quantitative analyst expert's suggestion and trial and error techniques. For the bivariate beta type - III distribution, for example, c must be substantially greater than u_1, u_2 and u_3 to get the required concentration for the default rate data.
- Compute the Entropy type - (α) with the help of Bayesian estimate from equation (2.11) using grid search method by providing the values of the parameters of the bivariate beta type - III distributions in the grid and

multinomial distribution observations and α. As our aim is to select the prior distribution, therefore, $x_1 = 2, x_2 = 6, x_3 = 16$ is used for observations for multinomial distribution and we calculate entropy type - (α) for $\alpha = 2$.

- Determine the correlation for each parameter combination in the grid.
- During the selection of the parameters for the prior distribution, we choose them so that the Entropy of type - (α) and correlation are in the pre-defined range. A trial-and-error method is used to determine the entropy range. In the analysis for corporate bond default rate data range for the entropy of type - (α) for $\alpha = 2$ is set to be in between 0.2 to 0.3 and as the observed correlation of default rates between investment and speculative-grade is 0.6885, therefore, correlation range is set to be in between 0.6 to 0.75.

TABLE 2. Parameter Selection: Bivariate Beta Type - III Distribution

	Minimum	**Step Size**	**Maximum**	**Estimated Parameter Value**
u_1	1	1	10	4
u_2	20	1	30	30
u_3	1	1	10	6
c	300	10	400	390
Entropy of type - (α) (for $\alpha = 2$)				$= 0.25678$
Correlation				$= 0.60348$

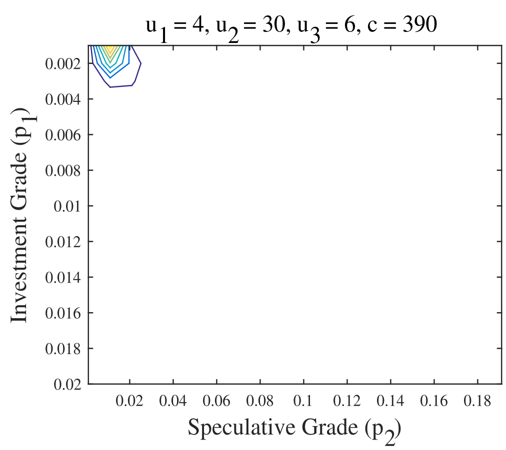


FIGURE 3. Bivariate Beta Type-III Fitted Distribution $BBeta^{III}(4, 30, 6, 390)$

The results obtained from the grid search approach for the bivariate beta type - III distribution are summarized in Table: 2 and Figure: 3. The first three columns in Table: 2 include information about the parameter boundaries and the parameter value for the bivariate beta type - III distribution is given in the last columns. The parameters appear to represent the features of the observed distribution very well.

4. Conclusions

This paper investigated methods for estimating type -(α) and entropy type -(α, β) for multinomial data. We employed various Bayesian approaches, utilizing bivariate beta Type-I, Connor and Mosimann bivariate beta, and bivariate beta Type-III distributions as prior distributions. The selection of the appropriate prior distribution was guided by the correlation observed between the data points. It was observed that the estimates of entropy type -(α) and entropy type -(α, β) can be used to derive the Bayesian estimates of Shannon entropy by approaching $\alpha \to 1$ and $(\alpha, \beta) \to (1, 1)$, respectively. By analyzing entropy type -(α) and entropy type -(α, β), we determined the optimal hyperparameter values for the bivariate beta Type-III prior. This methodology was applied to real-world data on investment-grade and speculative-grade bond default rates, demonstrating its effectiveness in selecting suitable hyperparameter values for the chosen prior distribution. Bayesian methods incorporate prior knowledge or beliefs about the underlying distribution, leading to more robust estimates. Here estimates were calculated under the quadratic loss function but for other loss functions such as linear exponential, and precautionary loss functions, estimates of entropies can be calculated. The quadratic loss function was employed for these estimations; however, entropy estimates can also be derived using different loss functions, such as the linear exponential and precautionary loss functions.

Acknowledgement. The authors are grateful to the referee for his valuable suggestions. The second author is supported with a senior research fellowship (File No.: 09/0045(11788)2021-EMR-I) from the Council of Scientific and Industrial Research, Ministry of Science & Technology, Government of India.

References

1. Al-Omari, A. I., Hassan, A. S., Nagy, H. F., Al-Anzi, A. R., & Alzoubi, L. (2021). Entropy Bayesian Analysis for the Generalized Inverse Exponential Distribution Based on URRSS. Computers, Materials and Continua, 69(3), 3795–3811. https://doi.org/10.32604/cmc.2021.019061
2. Behara, M., & Nath, P. (1974). Information and Entropy of Countable Measurable Partitions I. Kybernetika, 10, 491–503. Retrived from https://www.kybernetika.cz/content/1974/6/491
3. Bodvin, L. J. S., Bekker, A., & Roux, J. J. J. (2011). Shannon Entropy as a Measure of Certainty in a Bayesian Calibration Framework with Bivariate Beta Priors. South African Statistical Journal, 45, 171–204. Retrived from https://hdl.handle.net/10520/EJC99139
4. Cardeno, L., Nagar, D. K., & Sánchez, L. E. (2005). Beta Type 3 Distribution and Its Multivariate Generalization. Tamsui Oxford Journal of Mathematical Sciences, 21(2), 225–241. Retrived from https://www.researchgate.net/publication/268624540_Beta_type_3_distribution_and_its_multivariate_generalization

5. Connor, R. J., & Mosimann, J. E. (1969). Concepts of Independence for Proportions with a Generalization of the Dirichlet Distribution. Journal of the American Statistical Association, 64(325), 194–206. https://doi.org/10.1080/01621459.1969.10500963
6. Chacko, M., & Asha, P. S. (2018). Estimation of Entropy for Generalized Exponential Distribution Based on Record Values. Journal of the Indian Society for Probability and Statistics, 19, 79–96. https://doi.org/10.1007/s41096-018-0033-4
7. Default Trends – Global Annual Default Study: Corporate Default Rate will rise in 2023 and peak in early 2024. Retrived from https://www.moodys.com
8. Ehlers, R., Bekker, A., & Roux, J. J. J. (2009). The Central and Non-central Matrix Variate Dirichlet Type III Distribution. South African Statistical Journal, 43, 97–116. https://hdl.handle.net/10520/EJC99115
9. Ehlers, R., Bekker, A., & Roux, J. J. J. (2012). Triply Noncentral Bivariate Beta Type V Distribution. South African Statistical Journal, 46, 221–246. https://hdl.handle.net/10520/EJC-72f8e620b
10. Gradstein, I. S., & Ryzhik, I. M. (2014). Table of Integrals, Series, and Products. Eds. Jeffrey A., Zwillinger D. https://doi.org/10.1016/C2010-0-64839-5
11. Hassan, A. S., & Zaky, A. N. (2021). Entropy Bayesian Estimation for Lomax Distribution Based on Record. Thailand Statistician, 19(1), 95–114. Retrived from https://ph02.tci-thaijo.org/index.php/thaistat/article/view/242817
12. Havrda, J., & Charvat, F. (1967). Quantification Method of Classification Process. Concept of Structural α-entropy. Kybernetika (Prague), 3, 30–35. Retrived from https://www.kybernetika.cz/content/1967/1/30
13. Kumar, N., Dixit, A., & Vijay, V. (2024). Estimation of Tsallis Entropy for Exponentially Distributed Several Populations. ArXiv preprint arXiv:2401.09009. https://doi.org/10.48550/arXiv.2401.09009
14. Lai, C. D., & Balakrishnan, N. (2009). Continuous Bivariate Distributions. Springer-Verlag New York. https://doi.org/10.1007/b101765
15. Lawlor, G. R. (2020). L'Hôpital's Rule for Multivariable Functions. The American Mathematical Monthly, 127(8), 717–725. https://doi.org/10.1080/00029890.2020.1793635
16. Lee, P. M. (2004). Bayesian Statistics: An Introduction, 3rd ed., London: Wiley.
17. Liu, S., & Gui, W. (2019). Estimating the Entropy for Lomax Distribution Based on Generalized Progressively Hybrid Censoring. Symmetry, 11(10), 1219. https://doi.org/10.3390/sym11101219
18. Malik, S. H., & Qureshi, M. I. (2022). Certain Cubic Reduction Formulas Involving Hypergeometric Functions. Gulf Journal of Mathematics, 13(2), 94-105. https://doi.org/10.56947/gjom.v13i2.788
19. Martz, H. F., & Waller, R. A. (1982). Bayesian Reliability Analysis. John Wiley & Sons, New York. https://doi.org/10.1002/9780470061572.eqr081
20. Ren, H., & Hu, X. (2023). Bayesian Estimations of Shannon Entropy and Rényi Entropy of Inverse Weibull Distribution. Mathematics, 11(11), 2483. https://doi.org/10.3390/math11112483
21. Shannon, C. E. (1948). A Mathematical Theory of Communication. The Bell System Technical Journal, 27(3), 379–423, 623–656. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x
22. Shrahili, M., El-Saeed, A. R., Hassan, A. S., Elbatal, I., & Elgarhy, M. (2022). Estimation of Entropy for Log-logistic Distribution under Progressive Type II censoring. Journal of Nanomaterials, 2022, 1–10. https://doi.org/10.1155/2022/2739606
23. Tehlan, S., & Kumar, V. (2024). Quantile-based Generalized Entropy of Order (α, β) for Doubly Truncated Random Variable. Gulf Journal of Mathematics, 16(2), 135-150. https://doi.org/10.56947/gjom.v16i2.1875

[1] Department of Mathematics, Zakir Husain Delhi College (University of Delhi), Jawaharlal Nehru Marg, Delhi - 110002, India
Email address: dksingh@zh.du.ac.in

[2] Department of Mathematics, University of Delhi, Delhi - 110007, India
Email address: du.himanshu@gmail.com