Revista Colombiana de Estadística - Theoretical Statistics
Julio 2020, volumen 43, no. 2, pp. 173 a 182
http://dx.doi.org/10.15446/rce.v43n2.81979

On Predictive Distribution of K-Inflated Poisson Models with and Without Additional Information

Acerca de la distribución predicitiva de modelos Poisson K-inflados

Abdolnasser Sadeghkhani[1,a], S. Ejaz Ahmed[2,b]

[1]Departmento de Estadística, Instituto Tecnológico Autónomo de México, Ciudad de México, México

[2]Department of Mathematics and Statistics, Brock University, Ontario, Canada

Abstract

This paper addresses different approaches in finding the Bayesian predictive distribution of a random variable from a Poisson model that can handle count data with an inflated value of $K \in \mathbb{N}$, known as the KIP model. We explore how we can use other source of additional information to find such an estimator. More specifically, we find a Bayesian estimator of future density of random variable Y_1, based on observable X_1 from the $\mathrm{K_1IP}(p_1, \lambda_1)$ model, with and without assuming that there exists another random variable X_2, from the $\mathrm{K_2IP}(p_2, \lambda_2)$ model, independent of X_1, provided $\lambda_1 \geq \lambda_2$, and compare their performance using simulation method.

***Key words*:** KIP model; Bayesian statistics; Bayesian predictive distribution; Simulation.

Resumen

Este artículo presenta diferentes enfoques para buscar la distribución bayesiana predictiva de una variable aleatoria con un valor inflado $k \in \mathbb{N}$ conocido como el modelo KIP. Se explora como usar una fuente de información adicional para encontrar el estimador. Específicamente, se busca un estimador Bayesiano de la densidad futura de una variable aleatoria Y_1, basada en una variable observable X_1 a partir del modelo $\mathrm{K_1IP}(p_1, \lambda_1)$, con y sin el supuesto de que existe otra variable aleatoria X_2 del modelo $\mathrm{K_2IP}(p_2, \lambda_2)$, independiente de X_1, si $\lambda_1 \geq \lambda_2$, y se compara su desempeño usando un método de simulación.

***Palabras clave*:** Modelo KIP; Estadísticas bayesianas; Distribución predictiva bayesiana; Simulación.

[a]Ph.D. E-mail: nasser@itam.mx
[b]Ph.D. E-mail: sahmed5@brocku.ca

1. Introduction

The probability mass function (pmf) of a count variable X that follows a Poisson model, $\text{Po}(\lambda)$, is given by

$$P(X = x \mid \lambda) = \frac{e^{-\lambda}\lambda^x}{x!}\,, \;\; x = 0, 1, \dots\,, \lambda > 0. \tag{1}$$

Model (1) is used to represent count data, especially rare events such as highway accidents, earthquakes, incidents of terrorism, or airplane crashes. Often this type of dataset exhibits more zero or other observations than the Poisson model would anticipate. In these situations, a k-inflated Poisson (KIP) model is better applicable. Let us suppose that the probability of a random variable X is inflated at the value k. This model is a two–component mixture model combining a point mass at $x = k$ with a Poisson model and it has the pmf

$$P(X = x \mid p,\, \lambda) = \begin{cases} p + (1-p)\frac{e^{-\lambda}\lambda^k}{k!} & \text{if } x = k \\ (1-p)\frac{e^{-\lambda}\lambda^x}{x!} & \text{if } x \in \mathbb{N} \setminus \{k\}. \end{cases} \tag{2}$$

We denoted this model by $\text{KIP}(p, \lambda)$. The KIP model reduces to the ZIP model, the zero inflated Poisson, when $k = 0$, and to the Poisson model when $p = 0$ (see Mullahy, 1986 and Lambert, 1992). Unhapipat, Tiensuwan & Pal (1986) studied the Bayesian properties of the ZIP model with application in the public health. There has been some research which considers multiple inflations. Lin & Tsai (2013) have discussed a model that can be applied to both excessive zeros and ones known as the zero–and–one–inflated Poisson, or ZOIP model, and Melkersson & Rooth (2000) proposed a zero–and–two–inflated Poisson model, which accounts for a relative excess of both zero and two children in modeling complete female fertility. Sadeghkhani & Ahmed (1986) studied the Bayesian properties of the K_1-and-K_2–inflated Poisson, or $\text{K}_1\text{K}_2\text{P}$ model, for any $K_i \in \mathbb{N}$ for $i = 1, 2$ in the ice hockey. In this paper, we address the problem of estimating the future density of random variable Y_1, based on observable X_1 from the $\text{K}_1\text{IP}(p_1, \lambda_1)$ model, with (and without) assuming that there exists another random variable X_2, from the $\text{K}_2\text{IP}(p_2, \lambda_2)$ model, independent of X_1, provided $\lambda_1 \geq \lambda_2$. We also study different kinds of predictive distributions for Y_1.

The main objective is how can we use another variable X_2 independent of X_1 (seemingly irrelevant at first glance, but have some linkage through the unknown parameters) to predict the future distribution of Y_1 based on X_1. In here we have assumed that $\lambda_1 \geq \lambda_2$. For instance, from previous information or historical data we know, say, the mean number of the accidents in region A is greater than region B. (Accidents in region A and B are assumed to be independent) How can we predict the future number of accidents in region A by knowing that additional information?

The remainder of this paper is organized as follows: Section 2 presents a review of the problem, the likelihood function, and Bayesian setups. In Section 3, we study different Bayesian predictive densities of random variables, while in Section 4, we compare the results using simulation studies. Finally, we make some concluding remarks in Section 5.

2. Problem Setup

Let $X_1 = (X_{11}, \dots, X_{1n_1})$ be a random sample of size n_1 from $\mathrm{K_1IP}(p_1, \lambda_1)$, independent of X_2. Likewise, let $X_2 = (X_{21}, \dots, X_{2n_2})$, a random sample of size n_2 from $\mathrm{K_2IP}(p_2, \lambda_2)$. Let $X = (X_1, X_2)$ with parameters p_i and λ_i, for $i = 1, 2$ are unknown. Based on previous study, we will assume $\lambda_1 \geq \lambda_2$.

If we, suppose n_{1k_1} and n_{2k_2}, are the number of $x_{1i} = k_1$ and $x_{2m} = k_2$ for $i = 1, \dots, n_1$ and $m = 1, \dots, n_2$, are quite large, then likelihood function can be given by

$$
\begin{aligned}
&L(p_1, \lambda_1, p_2, \lambda_2 \mid x) \\
&\propto \left(p_1 + (1-p_1)\, e^{-\lambda_1} \frac{\lambda_1^{k_1}}{k_1!}\right)^{n_{1k_1}} \left(p_2 + (1-p_2)\, e^{-\lambda_2} \frac{\lambda_2^{k_2}}{k_2!}\right)^{n_{2k_2}} (1-p_1)^{n_1 - n_{1k_1}} (1-p_2)^{n_2 - n_{2k_2}} \\
&\qquad \times e^{-(n_1 - n_{1k_1})\lambda_1} \lambda_1^{\sum_{\{i: x_{1i} \neq k_1\}} x_{1i}} e^{-(n_2 - n_{2k_2})\lambda_2} \lambda_2^{\sum_{\{m: x_{2m} \neq k_2\}} x_{2m}} \\
&\qquad \propto \sum_{j=0}^{n_{1k_1}} \binom{n_{1k_1}}{j} p_1^j (1-p_1)^{n_1 - j} e^{-\lambda_1 (n_1 - j)} \lambda_1^{\sum\limits_{\{i: x_{1i} \neq k_1\}} x_{1i} + k_1 (n_{1k_1} - j)} \\
&\qquad\qquad \times \sum_{l=0}^{n_{2k_2}} \binom{n_{2k_2}}{l} p_2^l (1-p_2)^{n_2 - l} e^{-\lambda (n_2 - l)} \lambda_2^{\sum\limits_{\{m: x_{2m} \neq k_2\}} x_{2m} + k_2 (n_{2k_2} - l)} .
\end{aligned} \tag{3}
$$

We also consider the following prior densities:

$$
p_1, p_2 \sim \mathit{independent}\ U(0,1)\,, \tag{4}
$$

$$
(\lambda_1, \lambda_2) \sim \left(\frac{1}{\lambda_1 \lambda_2}\right)^{\alpha - \beta + 1}, \ \ \beta > 0, \ 0 < \alpha < \beta, \quad \text{and} \quad \lambda_1 \geq \lambda_2 > 0. \tag{5}
$$

The prior in (5) was introduced by Komaki (2004) as a class of improper shrinkage prior for the mean of Poisson distribution that the Jeffreys prior corresponds to $\alpha = 0$ and $\beta = \frac{1}{2}$. Lemma 1 gives the joint posterior distribution under prior densities in (4) and (5).

Lemma 1. *Suppose that $X_1 = (X_{11}, \dots, X_{1n_1})$ is a random sample of size n_1 from $\mathrm{K_1IP}(p_1, \lambda_1)$ and that $X_2 = (X_{21}, \dots, X_{2n_2})$ is a random sample of size n_2 from $\mathrm{K_2IP}(p_2, \lambda_2)$ with prior density in (4) and (5). Then by assuming*

$$
w_j(x_1) = \sum_{\{i: x_{1i} \neq k_1\}} x_{i1} + k_1 (n_{1k_1} - j) + \beta - \alpha\,, \ \ \textit{and} \tag{6}
$$

$$
w_l(x_2) = \sum_{\{m: x_{2m} \neq k_2\}} x_{m2} + k_2 (n_{2k_2}) + \beta - \alpha\,, \tag{7}
$$

1. *the joint posterior $\pi(p_1, p_2, \lambda_1, \lambda_2 \mid x)$ density is given by*

$$
\frac{\begin{array}{r} \sum_{j=0}^{n_{1k_1}} \sum_{l=0}^{n_{2k_2}} \binom{n_{1k_1}}{j} \binom{n_{2k_2}}{l} p_1^j\, (1-p_1)^{n_1 - j} p_2^l\, (1-p_2)^{n_2 - l} \\ e^{-\lambda_1 (n_1 - j)} \lambda_1^{w_j(x_1) - 1} e^{-\lambda_2 (n_2 - l)} \lambda_2^{w_l(x_2) - 1} \end{array}}{\begin{array}{l} \sum_{j=0}^{n_{1k_1}} \sum_{l=0}^{n_{2k_2}} \binom{n_{1k_1}}{j} \binom{n_{2k_2}}{l} \mathrm{Bet}(j+1, n_1 - j + 1) \\ \mathrm{Bet}(l+1, n_2 - l + 1)\, \mathrm{Q}(n_1 - j, w_j(x_1), w_l(x_2)) (n_2 - l)^{-w_l(x_2)} \end{array}}, \tag{8}
$$

where

$$\mathrm{Q}(a,b,c)=\int_0^{\infty} e^{-at}\, t^{b-1}\gamma(c;t)\,dt\,, \tag{9}$$

and $\gamma(c;d)$ *is the upper incomplete gamma function and is given as* $\int_0^d t^{b-1}e^{-t}\,dt$.

2. *The joint posterior* $\pi(p_1,\lambda_1 \mid x)$ *density is given by*

$$\frac{\sum_{j=0}^{n_{1k_1}}\sum_{l=0}^{n_{2k_2}}\binom{n_{1k_1}}{j}\binom{n_{2k_2}}{l}p_1^j\,(1-p_1)^{n_1-j}e^{-\lambda_1(n_1-j)}\lambda_1^{w_j(x_1)-1}\ \mathrm{Bet}(l+1,n_2-l+1)\gamma(w_l(x_2);\lambda_1)(n_2-l)^{-w_l(x_2)}}{\sum_{j=0}^{n_{1k_1}}\sum_{l=0}^{n_{2k_2}}\binom{n_{1k_1}}{j}\binom{n_{2k_2}}{l}\,\mathrm{Bet}(j+1,n_1-j+1)\ \mathrm{Bet}(l+1,n_2-l+1)\,\mathrm{Q}(n_1-j,w_j(x_1),w_l(x_2))(n_2-l)^{-w_l(x_2)}}. \tag{10}$$

This implies

a) the marginal posterior $\pi(\lambda_1 \mid x)$, *given by*

$$\frac{\sum_{j=0}^{n_{1k_1}}\sum_{l=0}^{n_{2k_2}}\binom{n_{1k_1}}{j}\binom{n_{2k_2}}{l}e^{-\lambda_1(n_1-j)}\lambda_1^{w_j(x_1)-1}\,\mathrm{Bet}(j+1,n_1-j+1)\ \mathrm{Bet}(l+1,n_2-l+1)\gamma(w_l(x_2);\lambda_1)(n_2-l)^{-w_l(x_2)}}{\sum_{j=0}^{n_{1k_1}}\sum_{l=0}^{n_{2k_2}}\binom{n_{1k_1}}{j}\binom{n_{2k_2}}{l}\,\mathrm{Bet}(j+1,n_1-j+1)\ \mathrm{Bet}(l+1,n_2-l+1)\,\mathrm{Q}(n_1-j,w_j(x_1),w_l(x_2))(n_2-l)^{-w_l(x_2)}}. \tag{11}$$

b) the marginal posterior $\pi(p_1 \mid x)$, *given by*

$$\frac{\sum_{j=0}^{n_{1k_1}}\sum_{l=0}^{n_{2k_2}}\binom{n_{1k_1}}{j}\binom{n_{2k_2}}{l}p_1^j(1-p_1)^{n_1-j}\ \mathrm{Bet}(l+1,n_2-l+1)\,\mathrm{Q}(n_1-j,w_j(x_1),w_l(x_2))(n_2-l)^{-w_l(x_2)}}{\sum_{j=0}^{n_{1k_1}}\sum_{l=0}^{n_{2k_2}}\binom{n_{1k_1}}{j}\binom{n_{2k_2}}{l}\,\mathrm{Bet}(j+1,n_1-j+1)\ \mathrm{Bet}(l+1,n_2-l+1)\,\mathrm{Q}(n_1-j,w_j(x_1),w_l(x_2))(n_2-l)^{-w_l(x_2)}}. \tag{12}$$

Proof. The numerator of (8) in (1) is the product of likelihood function (3) and prior densities (4) and (5) and denominator is obtained as follows:

$$\begin{aligned}&\int_0^1\!\!\int_0^{\infty}\sum_{j=0}^{n_{1k_1}}\binom{n_{1k_1}}{j}p_1^j(1-p_1)^{n_1-j}e^{-\lambda_1(n_1-j)}\lambda_1^{w_j(x_1)-1}\,d\lambda_1\,dp_1\\&\qquad\times\int_0^1\!\!\int_0^{\lambda_1}\sum_{l=0}^{n_{2k_2}}\binom{n_{2k_2}}{l}p_2^l(1-p_2)^{n_2-l}e^{-\lambda_2(n_2-l)}\lambda_2^{w_l(x_2)-1}\,d\lambda_2\,dp_2.\end{aligned} \tag{13}$$

The second line in (13) is

$$\sum_{l=0}^{n_{2k_2}}\binom{n_{2k_2}}{l}\mathrm{Bet}(l+1,n_2-l+1)\gamma(w_l(x_2);\lambda_1)(n_2-l)^{-w_l(x_2)}. \tag{14}$$

By using (9), (13) can be written as

$$\sum_{j=0}^{n_{1k_1}} \sum_{l=0}^{n_{2k_2}} \binom{n_{1k_1}}{j}\binom{n_{2k_2}}{l} \mathrm{Bet}(j+1, n_1-j+1)$$
$$\mathrm{Bet}(l+1, n_2-l+1)\, \mathrm{Q}(n_1-j, w_j(x_1), w_l(x_2))(n_2-l)^{-w_l(x_2)},$$

which completes the proof. The proof of (2) can be done similarly. □

Note that by using (11) and (10), one can obtain the Bayes parameters of p_1 and λ_1, given observed $x = (x_1, x_2)$, restricted to $\lambda_1 \geq \lambda_2$, as is demonstrated in Lemma 2.

Lemma 2. *The Bayes estimator of parameters p_1 and λ_1, under the squared error loss function, are given by*

$$\hat{p}_{1,\pi} = \frac{\sum_{j=0}^{n_{1k_1}} \sum_{l=0}^{n_{2k_2}} \binom{n_{1k_1}}{j}\binom{n_{2k_2}}{l} \mathrm{Bet}(j+2, n_1-j+1)\; \mathrm{Bet}(l+1, n_2-l+1)\, \mathrm{Q}(n_1-j, w_j(x_1), w_l(x_2))(n_2-l)^{-w_l(x_2)}}{\sum_{j=0}^{n_{1k_1}} \sum_{l=0}^{n_{2k_2}} \binom{n_{1k_1}}{j}\binom{n_{2k_2}}{l} \mathrm{Bet}(j+1, n_1-j+1)\; \mathrm{Bet}(l+1, n_2-l+1)\, \mathrm{Q}(n_1-j, w_j(x_1), w_l(x_2))(n_2-l)^{-w_l(x_2)}} \tag{15}$$

$$\hat{\lambda}_{1,\pi} = \frac{\sum_{j=0}^{n_{1k_1}} \sum_{l=0}^{n_{2k_2}} \binom{n_{1k_1}}{j}\binom{n_{2k_2}}{l} \mathrm{Bet}(j+1, n_1-j+1)\; \mathrm{Bet}(l+1, n_2-l+1)\, \mathrm{Q}(n_1-j, w_j(x_1)+1, w_l(x_2))(n_2-l)^{-w_l(x_2)}}{\sum_{j=0}^{n_{1k_1}} \sum_{l=0}^{n_{2k_2}} \binom{n_{1k_1}}{j}\binom{n_{2k_2}}{l} \mathrm{Bet}(j+1, n_1-j+1)\; \mathrm{Bet}(l+1, n_2-l+1)\, \mathrm{Q}(n_1-j, w_j(x_1), w_l(x_2))(n_2-l)^{-w_l(x_2)}}. \tag{16}$$

Moreover, the marginal posterior of parameter of λ_1, given x_1, can be obtained as follows:

$$\pi(p_1 \mid x_1) = \frac{\sum_{j=0}^{n_{1k_1}} \binom{n_{1k_1}}{j} p_1^j (1-p_1)^{n_1-j}\, \Gamma(w_j(x_1))(n_1-j)^{-w_j(x_1)}}{\sum_{j=0}^{n_{1k_1}} \binom{n_{1k_1}}{j} \mathrm{Bet}(j+1, n-j+2)\, \Gamma(w_j(x_1))(n_1-j)^{-w_j(x_1)}}, \tag{17}$$

$$\pi(\lambda_1 \mid x_1) = \frac{\sum_{j=0}^{n_{1k_1}} \binom{n_{1k_1}}{j} \mathrm{Bet}(j+1, n-j+2)\, e^{-\lambda_1(n_1-j)} \lambda_1^{w_j(x_1)-1}}{\sum_{j=0}^{n_{1k_1}} \binom{n_{1k_1}}{j} \mathrm{Bet}(j+1, n-j+2)\, \Gamma(w_j(x_1))(n_1-j)^{-w_j(x_1)}}. \tag{18}$$

therefore, one can use (17) to obtain (unrestricted) Bayes estimator of parameter λ_1,

$$\hat{p}_1 = \frac{\sum_{j=0}^{n_{1k_1}} \binom{n_{1k_1}}{j} \mathrm{Bet}(j+2, n-j+2)\, \Gamma(w_j(x_1))(n_1-j)^{-w_j(x_1)}}{\sum_{j=0}^{n_{1k_1}} \binom{n_{1k_1}}{j} \mathrm{Bet}(j+1, n-j+2)\, \Gamma(w_j(x_1))(n_1-j)^{-w_j(x_1)}}, \tag{19}$$

$$\hat{\lambda}_1 = \frac{\sum_{j=0}^{n_{1k_1}} \binom{n_{1k_1}}{j} \mathrm{Bet}(j+1, n-j+2)\, \Gamma(w_j(x_1)+1)(n_1-j)^{-w_j(x_1)-1}}{\sum_{j=0}^{n_{1k_1}} \binom{n_{1k_1}}{j} \mathrm{Bet}(j+1, n-j+2)\, \Gamma(w_j(x_1))(n_1-j)^{-w_j(x_1)}}. \tag{20}$$

3. Bayesian Predictive Densities

We consider the problem of constructing Bayes predictive density for the future observation y, based on observable x. We use the Kullback Leibler (KL) loss function (divergence)

$$\mathrm{KL}(q(y \mid \theta)), \hat{q}_\pi(y;x)) = \sum_y q(y \mid \theta) \log \frac{q(y \mid \theta)}{\hat{q}_\pi(y;x)}\,, \tag{21}$$

where $\hat{q}_\pi(y;x)$ is the Bayes predictive distribution in estimating the density $Y \sim q(\cdot \mid \theta)$, and θ is $a(n)$ (vector of) unknown parameter(s). The corresponding risk function given as follows:

$$\begin{aligned} r_{\hat{q}}(\theta) &= \mathbb{E}^{X|\theta}\,\mathrm{KL}(q,\hat{q}) \\ &= \sum_x q(x \mid \theta) \sum_y q(y \mid \theta) \log \frac{q(y \mid \theta)}{\hat{q}_\pi(y;x)}. \end{aligned} \tag{22}$$

Previous studies (see, Corcuera & Giummolé, 1999), indicates that under the KL, the Bayes predictive distribution for Y based on observed value x, and posterior density $\pi(\cdot \mid x)$, is given as

$$\hat{q}_\pi(y;x) = \int_\Theta q(y \mid \theta)\,\pi(\theta \mid x)\,d\theta. \tag{23}$$

The following theorem provides the Bayes predictive distribution under the KL loss function.

Theorem 1. *Let* $X_1 = (X_{11}, \ldots, X_{1n_1})$ *be a random sample from from* $\mathrm{K_1IP}(p_1, \lambda_1)$*, independent of* X_2*. Let* $X_2 = (X_{21}, \ldots, X_{2n_2})$ *be random sample from* $\mathrm{K_2IP}(p_2, \lambda_2)$*. Then the Bayes predictive distribution of future random variable* $Y_1 \sim \mathrm{K_1IP}(p_1, \lambda_1)$*, by considering the additional information* $\lambda_1 \geq \lambda_2$*,* $\hat{q}(Y_1 = y_1;\, x)$ *for* $y_1 \in \mathbb{N} \setminus \{k_1\}$ *is given by*

$$\frac{\begin{array}{r}\sum_{j=0}^{n_{1k_1}}\sum_{l=0}^{n_{2k_2}} \binom{n_{1k_1}}{j}\binom{n_{2k_2}}{l} \mathrm{Bet}(j+1, n_1-j+2)\,\mathrm{Bet}(l+1, n_2-l+1) \\ \mathrm{Q}(n_1-j+1, w_j(x_1)+y_1, w_l(x_2))(n_2-l)^{-w_l(x_2)}\end{array}}{\begin{array}{r}y_1!\sum_{j=0}^{n_{1k_1}}\sum_{l=0}^{n_{2k_2}} \binom{n_{1k_1}}{j}\binom{n_{2k_2}}{l} \mathrm{Bet}(j+1, n_1-j+1)\,\mathrm{Bet}(l+1, n_2-l+1) \\ \mathrm{Q}(n_1-j, w_j(x_1), w_l(x_2))(n_2-l)^{-w_l(x_2)}\end{array}}, \tag{24}$$

and for $y_1 = k_1$*,* $\hat{q}(Y_1 = k_1;\, x)$ *is given by*

$$\frac{\begin{array}{r}\sum_{j=0}^{n_{1k_1}}\sum_{l=0}^{n_{2k_2}} \binom{n_{1k_1}}{j}\binom{n_{2k_2}}{l} \mathrm{Bet}(j+1, n_1-j+2)\,\mathrm{Bet}(l+1, n_2-l+1) \\ \mathrm{Q}(n_1-j+1, w_j(x_1)+k_1, w_l(x_2))(n_2-l)^{-w_l(x_2)}\end{array}}{\begin{array}{r}k_1!\sum_{j=0}^{n_{1k_1}}\sum_{l=0}^{n_{2k_2}} \binom{n_{1k_1}}{j}\binom{n_{2k_2}}{l} \mathrm{Bet}(j+1, n_1-j+1)\,\mathrm{Bet}(l+1, n_2-l+1) \\ \mathrm{Q}(n_1-j, w_j(x_1), w_l(x_2))(n_2-l)^{-w_l(x_2)}\end{array}}, \tag{25}$$

$$\frac{\begin{array}{r}\sum_{j=0}^{n_{1k_1}}\sum_{l=0}^{n_{2k_2}} \binom{n_{1k_1}}{j}\binom{n_{2k_2}}{l} \mathrm{Bet}(j+2, n_1-j+1)\,\mathrm{Bet}(l+1, n_2-l+1) \\ \mathrm{Q}(n_1-j+1, w_j(x_1), w_l(x_2))(n_2-l)^{-w_l(x_2)}\end{array}}{\begin{array}{r}\sum_{j=0}^{n_{1k_1}}\sum_{l=0}^{n_{2k_2}} \binom{n_{1k_1}}{j}\binom{n_{2k_2}}{l} \mathrm{Bet}(j+1, n_1-j+1)\,\mathrm{Bet}(l+1, n_2-l+1) \\ \mathrm{Q}(n_1-j, w_j(x_1), w_l(x_2))(n_2-l)^{-w_l(x_2)}\end{array}}. \tag{26}$$

Proof. According to (23), the proof of (24) is given by

$$\int_0^1\int_0^\infty\int_0^1\int_0^{\lambda_1} \frac{1-p_1}{y_1!} e^{-\lambda_1}\lambda_1^{y_1}\, \pi(p_1,p_2,\lambda_1,\lambda_2 \mid x)\, d\lambda_2\, d\lambda_1\, dp_2\, dp_1\,, \tag{27}$$

replacing joint posterior density function (8) in (27). Doing calculations similar to those in the proof of Lemma (1), gives the proof. Equation (25) can be proven likewise using the following formula:

$$\int_0^1\int_0^\infty\int_0^1\int_0^{\lambda_1} \left(p_1 + \frac{1-p_1}{y_1!} e^{-\lambda_1}\lambda_1^{y_1}\right) \pi(p_1,p_2,\lambda_1,\lambda_2 \mid x) d\lambda_2\, d\lambda_1\, dp_2\, dp_1.$$

□

If we ignore the additional information $\lambda_1 \geq \lambda_2$, we can find the $\hat{q}_{mre}$, the minimum risk equivariant (mre) distribution, which is the Bayes estimator under KL loss function and non–informative prior without contemplating the restriction of the parameters. The following lemma, gives the mre distribution.

Lemma 3. *The minimum risk equivariant (mre) distribution for* $Y_1 \sim \mathrm{K_1IP}(p_1,\lambda_1)$*, under KL loss function and given* $X_1 = (X_{11},\ldots,X_{1n_1})$*, and prior densities* $(1/\lambda_1)^{\alpha-\beta+1}$*, for* $\beta > 0$, $0 < \alpha < \beta$*, independent of* $p_1 \sim U(0,1)$ *is given by*

$$\hat{q}(Y_1 = k_1;\, x_1) = \frac{\sum_{j=0}^{n_{1k_1}+1} \binom{n_{1k_1}+1}{j} \mathrm{Bet}(n_1-j+1, j+1)\, \Gamma(w_j(x_1)+k_1+1)(n-j+1)^{-w_j(x1)-k_1}}{k_1! \sum_{j=0}^{n_{1k_1}} \binom{n_{k_1}}{j} \mathrm{Bet}(n_1-j+2, j+1)\, \Gamma(w_j(x_1)+1)(n_1-j)^{-w_j(x_1)}}\,, \tag{28}$$

$$\hat{q}(Y_1 = y_1;\, x_1) = \frac{\sum_{j=0}^{n_{1k_1}} \binom{n_{1k_1}}{j} \mathrm{Bet}(j+1, n_1-j+2)\, \Gamma(w_j(x_1)+y)(n_1-j+1)^{-w_j(x_1)-y_1}}{y_1! \sum_{j=0}^{n_{1k_1}} \binom{n_{1k_1}}{j} \mathrm{Bet}(j+1, n_1-j+1)\, \Gamma(w_j(x_1))(n_1-j)^{-w_j(x_1)}}\,,\ y_1 \in \mathbb{N} \setminus \{k_1\}. \tag{29}$$

4. Comparison of Bayes and Plug-In Distributions

In this section, we simulated a random random variable of 1000 from two independent variables $X_1 \sim \mathrm{K_1IP}(0.2, 8)$ with $k_1 = 2$, and $X_2 \sim \mathrm{K_2IP}(0.3, 5)$ with $k_2 = 0$, respectively. Table 1 and Figure 1 are based on these two random samples.

Table 1: Simulation of 1000 observations form two independent variables $X_1 \sim \mathrm{K_1IP}(0.2, 8)$ with $k_1 = 2$, and $X_2 \sim \mathrm{K_2IP}(0.3, 5)$ with $k_2 = 0$.

Simulation from X_1	Simulation from X_2
$k_1 = 2$	$k_2 = 0$
$n_{1k_1} = 219$	$n_{2k_2} = 301$
$\bar{x}_1 = 6.66$	$\bar{x}_2 = 3.50$

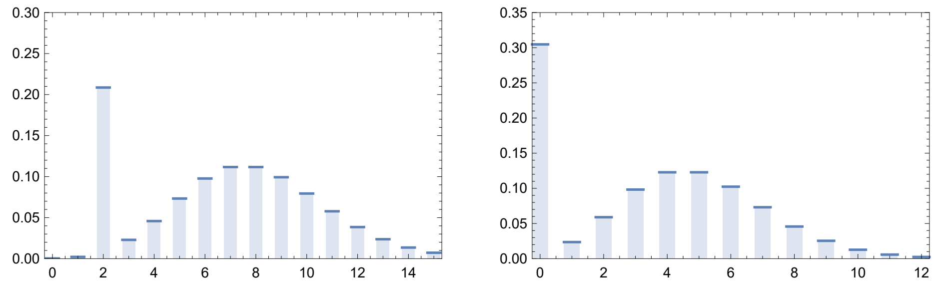

FIGURE 1: Pmf's of the K_1IP and K_2IP corresponding to Table 1.

The goal of this simulation is to find the density of future random variable $X_1 \sim \mathrm{KIP}(0.2, 8)$, by considering $\lambda_1 = 8 \geq \lambda_2 = 5$. Table 2 gives the Bayes predictive distribution (Theorem 1), the mre distribution (Lemma 3), the plug-in predictive distribution based on mle and the plug-in predictive distribution based on posterior expectations of the unknown parameters (Lemma 2).

TABLE 2: The Bayes, mre and plug-in predictive distributions (based on mle (equation 3) and posterior expectation) for future observation $y_1 \sim \mathrm{K_1IP}(0.2, 8)$, with $k_1 = 2$.

	$y_1 \in \mathbb{N} \setminus \{2\}$	$y_1 = 2$
$\hat{\boldsymbol{q}}(\cdot)$	$\dfrac{\sum_{j=0}^{219}\binom{219}{j}\,\mathrm{Bet}(j+1, 1002-j)\;\Gamma(y-2(j-3330.25))(1001-j)^{2(j-3330.25)-y_1}}{y_1!\sum_{j=0}^{219}\binom{219}{j}\,\mathrm{Bet}(j+1, 1001-j)\;\Gamma(2(3330.25-j))(1000-j)^{2(j-3330.25)}}$	0.21
$\hat{\boldsymbol{q}}_{\boldsymbol{\pi}}(\cdot)$	$\dfrac{\sum_{j=0}^{219}\sum_{l=0}^{301}\binom{2019}{j}\binom{301}{l}\,\mathrm{Bet}(j+1, 1002-j)\,\mathrm{Bet}(l+1, 1001-l)\;\mathrm{Q}(1001-j, y_1-2(-3330.25+j), 3500.5)(1000-l)^{-3500.5}}{1.12635850689 \times 10^{16647}}$	0.23
$\hat{\boldsymbol{q}}_{\boldsymbol{plug,mle}}(\cdot)$	$\frac{1915.07}{y_1!}7.9^{y_1}$	0.21
$\hat{\boldsymbol{q}}_{\boldsymbol{plug,b}}(\cdot)$	$\frac{0.0003}{y_1!}7.9^{y_1}$	0.22

The p-values for Pearson's chi-squared tests for the Bayes, the mre, and two plug-in distributions in Table 2, are 0.86, 0.98 and 0.99 respectively, and therefore one can conclude that all the proposed distributions, estimate the future density well.

Moreover, We can compare the performance of the predictive distributions in terms of the risk function based on the KL loss function given in equation (22). Figure 2 illustrates the risk functions of the Bayes predictive distribution (using the additional information) and the mre predictive distribution.

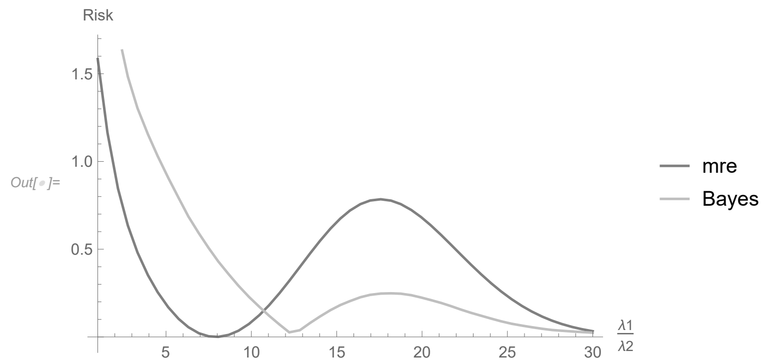

FIGURE 2: The risk functions of the Bayes and the mre predictive distributions (the predictive distributions with and without considering the additional information respectively) under the KL loss function.

5. Conclusions

In this paper, we introduced different predictive distributions for a future random variable from the KIP model, with and without considering additional information, and compared their performance using a simulation.

Acknowledgements

Abdolnasser Sadeghkhani acknowledges the ITAM and Asociación Mexicana de Cultura, A.C. for supporting this paper. S. Ejaz Ahmed acknowledges the Natural Sciences and the Engineering Research Council of Canada, and the Ontario Centre of Excellence for supporting this research.

The authors are grateful to the editor and the anonymous reviewers for their valuable comments and helpful suggestions.

[Recibido: septiembre de 2019 — Aceptado: abril de 2020]

References

Corcuera, J. M. & Giummolé, F. (1999), 'A generalized bayes rule for prediction', *Scandinavian Journal of Statistics* **26**(2), 265–279.

Komaki, F. (2004), 'Simultaneous prediction of independent poisson observables', *The Annals of Statistics* **32**(4), 1744–1769.

Lambert, D. (1992), 'Zero-inflated Poisson regression, with an application to defects in manufacturing', *Technometrics* **34**(1), 1–14.

Lin, T. H. & Tsai, M. H. (2013), 'Modeling health survey data with excessive zero and K responses', *Statistics in Medicine* **32**(9), 1572–1583.

Melkersson, M. & Rooth, D. O. (2000), 'Modeling female fertility using inflated count data models', *Journal of Population Economics* **13**(2), 189–203.

Mullahy, J. (1986), 'Specification and testing of some modified count data models', *Journal of Econometrics* **33**(3), 341–365.

Sadeghkhani, A. & Ahmed, S. E. (1986), 'The application of predictive distribution estimation in multiple–inflated poisson models to ice hockey data', *To be appeared in journal of Model assisted statistics & application* **33**(3), 341–365.

Unhapipat, S., Tiensuwan, M. & Pal, N. (1986), 'Bayesian predictive inference for zero–inflated poisson (zip) distribution with applications', *American Journal of Mathematical and Management Sciences* **37**(1), 66–79.