Bayesian Posterior Predictive Probability Happiness

Gabriela Rodríguez-Hernández, Galileo Domínguez-Zacarías*, Carlos Juárez Lugo

Universidad Autónoma del Estado de México, Ecatepec, Mexico
Email: ´grodriquezh@uaemex.mx, ´gadoza3@gmail.com

Received 15 March 2016; accepted 17 May 2016; published 20 May 2016

Abstract

We propose to determine the underlying causal structure of the elements of happiness from a set of empirically obtained data based on Bayesian. We consider the proposal to study happiness as a multidimensional construct which converges four dimensions with two different Bayesian techniques, in the first we use the Bonferroni correction to estimate the mean multiple comparisons, on this basis it is that we use the function t and a z-test, in both cases the results do not vary, so it is decided to present only those shown by the t test. In the Bayesian Multiple Linear Regression, we prove that happiness can be explained through three dimensions. The technical numerical used is MCMC, of four samples. The results show that the sample has not atypical behavior too and that suitable modifications can be described through a test. Another interesting result obtained is that the predictive probability for the case of sense positive of life and personal fulfillment dimensions exhibit a non-uniform variation.

Keywords

Bayesian Inference, Posterior Predictive Distribution, MCMC, Happiness

1. Introduction

Bayesian networks emerged about three decades ago as alternatives to conventional systems-oriented decision-making and forecasting under uncertainty in probabilistic terms [1]. A Bayesian network is a statistical tool that represents a set of associated uncertainties given conditional independence relationships established between them [2] [3].

The rule Bayes is a rigorous method for interpreting evidence in the context of previous experience or knowledge. The Bayes rule has recently emerged as a powerful tool with a wide range to applications which include:

*Corresponding author.

genetics, image processing, ecology, physics and engineering. The essential characteristics of Bayesian methods are their explicit use of probability for quantifying uncertainty in inferences base on statistical data analysis. A Bayesian probability interval for an unknown quantity of interest can be directly regarded as having a high probability of containing the unknown quantity, in contrast to a frequentist confidence interval which may strictly be interpreted only in relation to a sequence of similar inferences that might be made in repeated practice.

Using Bayesian networks as a tool for data analysis is not widespread in the context of Psychology. However, the benefits of using this tool in all areas of psychology are identified in several ways: on the economic front with Bayesian networks could develop systems to make judgments appropriate chance to improve diagnosis and psychological treatment. On the scientific level, Bayesian networks cannot be overlooked if psychology strives to clarify the mechanisms by which people evaluate, decide and make inferences; as they may serve as analytical and theoretical reference in the development of models of reasoning, learning and perception of uncertainty [3].

In psychology, the area’s most prolific work in the use of Bayesian networks has been the causal learning [4]-[7], where they test the hypothesis that people represent causal knowledge similarly to as it does a Bayesian network. Other theoretical developments have made possible to extend the use of Bayesian network outside the domain of causal knowledge, for instance, in inductive generalization of concepts and learning words [8]. Another facet of Bayesian networks which have proved useful tools is in market research [9]. Others found that a Bayesian network classifies customers of a company relative to the perspective of long-term purchases [10]. On the other hand, it analyzed the nature of human emotions and the negative behavior resulting from overcrowding during mass events. Utilizing the Bayesian network, his model shows the dependence structure between different emotions and negative behaviors of pilgrims in the crowd [11].

Finally, by its shaping power, the Bayesian networks could be used to generate more and better models of how organizations, groups or social aggregate which is the subject of study for psychology. Thus the aim of this study was to determine the underlying causal structure of the elements of happiness from a set of empirically obtained data. It is considered a Bayesian statistical parameter, which is inferred as an uncertain event, in this study the knowledge about happiness is not accurate and is subject to uncertainty, therefore happiness can be described by a probability distribution. Whereby the Metropolis algorithm used is a specific type of process Monte Carlo, which generates a random way so that every step along the way is completely independent of the previous steps of the current position and generates Markov chains. Process in each step has not memory of the previous states. This is known as Markov Chain Monte Carlo (MCMC)

Happiness has been defined as an entity that can be described by a specific set of measures [12], [13], a mental state that people can gain control over in a cognitive way to perceive and conceive both themselves and their world as an experience of joy, satisfaction or positive welfare [14]. Unfortunately, terms like happiness have been used frequently in daily discourse and may now have vague and somewhat different meanings.

The difficulty of defining happiness has led pioneer psychologists in the study of happiness propose the term subjective well-being (SWB). SWB refers to people's evaluations of their own lives and encompasses both cognitive judgments of satisfaction and affective appraisals of moods and emotions. This conceptualization emphasizes the subjective nature of happiness and holds individual human beings to be the best judges of their own happiness [15]-[17].

There is empirical evidence indicating that well-being is a much broader construct than stability of emotions and subjective judgment about life satisfaction; e.g. situational models consider that the sum of happy moments in life results in the satisfaction of people [18], that is, a person exposed to a greater amount of happy events will be more satisfied with his or her life. People briefly react to good and bad events, but in a short time they return to neutrality. Thus, happiness and unhappiness are merely short-lived reactions to changes in people [19] and it depends strongly on intentional activity [20], [21]. But also it has been identified the temperament is suggested to influence happiness [22], [23] and other personality traits such as optimism and self-esteem [24]-[26], the self-determination [27].

Scholars have noticed that happiness is not a single thing, but it can be broken down into its constituent elements. Considering this information, Alarcón [28] proposes to study happiness as a multidimensional construct which converges satisfaction of what has been achieved, positive attitudes toward life, (experiences that reflect positive feelings concerning one’s self and life) personal fulfillment, and joy of living.

Happiness Scale of Lima (HSL) [28], consists of 27 items and reported item-scale correlations were highly
significant and high internal consistency ($\alpha = 0.92$). The factorial analysis of principal components and varimax rotation, revealed that happiness is a multidimensional behavior, consisting of four dimension: $X_1$ Sense positive of life, $X_2$ Satisfaction with life, $X_3$ Personal fulfillment and $X_4$ Joy of living.

2. The Model

Definition 1
The $\sigma$-algebra is used for measured definition. A probability measure is a mapping $P: \mathcal{F} \rightarrow [0,1]$ that assigns a probability $P(A) \in [0,1]$ to any event, with the properties $P(\Omega) = 1$ and $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ whenever $A_i \cap A_j = \emptyset$. A real value random variable is a mapping $X: \Omega \rightarrow \mathbb{R}$ and $P(X(\omega) \in A) = P(\omega \in X^{-1}(A))$.

Definition 2
A random variables es given by $X: \Omega \rightarrow \mathbb{R}$ where $A$ is measurable and subset of $\mathbb{R}$, we denote by $\{X \in A\}$ then the event is $\{X \in A\} = \{\omega \in \Omega: X(\omega) \in A\}$.

Definition 3
The random samples, say $X_1 = \{X_{i1}, \ldots, X_{il}\}$, $X_2 = \{X_{i2}, \ldots, X_{il}\}$, $X_3 = \{X_{i3}, \ldots, X_{il}\}$ and $X_4 = \{X_{i4}, \ldots, X_{il}\}$, have the same underlying distribution.

We assume that we have four data samples $\{X_{i1}, X_{i2}, X_{i3}, X_{i4}\}$. If the data have the same distributions, then $X_1 \equiv X_2 \equiv X_3 \equiv X_4$. Uncertainty about the true value of parameter it described by a measurement $\Pi$ conditional observed datasets. The posterior predictive distribution, is defined as

$$P_X(A \in \Omega) = \sum_{i=1}^{\infty} P(A \in \Omega) \Pi_X$$

Lema 1.
Let $P$ be a probability measure on $(\Omega, x_0, x_1, x_2, x_3, x_4, \mathcal{F}_1 \otimes \mathcal{F}_2 \otimes \mathcal{F}_3 \otimes \mathcal{F}_4)$ such that $P\left(\{\omega_1, \omega_2, \omega_3, \omega_4\}: I_i(\omega) = 1\right) = P_i(I_i = 1)$ for $i = 1, 2, 3, 4$.

The probability for $X_1 \equiv X_2 \equiv X_3 \equiv X_4$ denoted $P\left(X_1 \equiv X_2 \equiv X_3 \equiv X_4\right)$ is equal to the joint probability $P\left(I_1 = 1, I_2 = 1, I_3 = 1, I_4 = 1\right)$ for any fixed $P_{x_1} \wedge P_{x_2} \wedge P_{x_3} \wedge P_{x_4}$, the probability is bounded by

$$\max\left\{0, 4 \parallel P_{x_1} \wedge P_{x_2} \wedge P_{x_3} \wedge P_{x_4} \parallel - 1\right\} \leq P\left(X_1 \equiv X_2 \equiv X_3 \equiv X_4\right) \leq \parallel P_{x_1} \wedge P_{x_2} \wedge P_{x_3} \wedge P_{x_4} \parallel$$

Consider the problem of selecting independent samples from several populations for the purpose of between-group comparisons, either through hypothesis testing or estimation of mean differences. A companion problem is the estimation of within-group mean levels. Together, these problems form the foundation for the very common analysis of variance framework, but also describe essential aspects of stratified sampling, cluster analysis, empirical Bayes, and other settings. Procedures for making between-group comparisons are known as multiple comparisons methods. The goal of determining which groups have equal means requires testing a collection of related hypotheses [29].

Consider independent samples from $l$ normally distributed populations with equal variances $(3a$ y $3b)$

$$X_{i1}, X_{i2}, \ldots, X_{il} \sim N\left(\mu, \sigma^2\right)$$

Then $H^{(a,b)}(\mu_a = \mu_b)$ for each $(a, b)$ then there is $l(l-1)/2$ tests. For example, for two sample there is 1 test, for three sample there is 3 test, and 4 there is 6 test. When two or more means are taken as equal, we merely combine all relevant samples into one.

Let $Y_n$ denote the observed data. Assume that $Y_n$ is to be described using a model $M_k$ selected from a set of candidate models $\{M_1, \ldots, M_j\}$. Assume that each $M_k$ is uniquely parameterized by $\theta_k$, an element of the
parameter space $\Theta(k)$. In the multiple comparisons problem, the class of candidate models consists of all possible mean level clustering. Each candidate model is parameterized by the mean vector $\mu = (\mu_1, \ldots, \mu_l)$ and the common variance $\sigma^2$, with the individual means restricted by the model defined clustering of equalities. That is, each model determines a corresponding parameter space where particular means are taken as equal.

Let $L_k(\theta_i | Y_i), \ldots, L_k(\theta_k | Y_k)$, where $\theta$ take values in a $d$-dimensional parameter space $\Theta$, be likelihood functions associated with the samples and $\pi(k)$ denote a prior density on $\Theta$ over the model $\{M_1, \ldots, M_l\}$. Then $g(\theta_k | k)$ denote a prior on $\theta_k$ given the model $M_k$. The posterior probability for $M_k$ and $\theta_k$ can be written as

$$P(M_k | Y_n) = \frac{\pi(k) g(\theta_k | k) L_k(\theta_k | Y_n)}{\sum_{k} P(M_k) p_{M_k}(Y_k)}$$  (4)

$$B_k = -\ln L(\hat{\theta} | Y_n) + \text{dim}(\theta) \ln(n)$$  (5)

You can write the posterior probability (4) y (5) as,

$$\pi_k(k | Y_n) = h(Y_n)^{-1} \pi(k) \int_{\Theta_k} g(\theta | k) L(\theta | Y_n) d\theta_k$$  (6)

We use a uniform prior for $\pi_k = \frac{1}{L} \forall k$ (6) it can be written as

$$\pi_k(k | Y_n) = h(Y_n)^{-1} \int_{\Theta_k} L(\theta | Y_n) d\theta_k$$

we using equations (5,6) and the previous development, we write the following relationship

$$-2 \ln \pi_k(k | Y_n) = -2 \ln \int L(\theta | Y_n) d\theta + c_n$$  (7)

**Definition 4**

The Taylor series of a real or complex-valued function $f(x)$ that is infinitely differentiable at a real or complex number $x$ is the power series

$$f(x) = \sum_{n=0}^{\infty} \frac{d^n f(x)}{n!} (x-a)^n$$

Then, applying the definition 4 to (7), we obtain

$$-\int \left[L(\hat{\theta}_k | Y_n) + \frac{\partial \ln L(\hat{\theta}_k | Y_n)}{\partial \theta_k} + \frac{\partial^2 \ln L(\hat{\theta}_k | Y_n)}{\partial \theta_k \partial \theta_k} + \cdots \right] d\theta_k + c_n$$  (8)

The Fisher information [30] is a way of measuring the amount of information that an observable random variable $X$ carries out unknown parameter $\theta$.

$$E \left[ \frac{\partial}{\partial \theta} \log f(X; \theta) | \theta \right] = 0$$  (9)

Applying (9) to (8), is obtained

$$\int L(\theta_k | Y_n) d\theta_k \approx L(\theta_k | Y_n) \exp \left\{ -\frac{n}{2} \left( \theta_k - \hat{\theta}_k \right) \left[ \frac{\partial^2 \ln L(\hat{\theta}_k | Y_n)}{\partial \theta_k \partial \theta_k} \right] \left( \theta_k - \hat{\theta}_k \right) \right\} d\theta_k$$
taken

$$-2 \ln \pi (k \mid Y_n) = -\ln L(\hat{\theta}_k \mid Y_n) + \dim(\theta_k) \ln(n)$$

Then

$$\pi(k \mid Y_n) \approx \exp\left(\frac{-B_k}{2}\right)$$

With respect to the candidate model class \(\{M_1, M_2, \ldots, M_L\}\), we obtain, the posterior model probabilities

$$\pi(k \mid Y_n) \equiv \frac{\exp\left(\frac{-B_k}{2}\right)}{\sum_{l=1}^{L} \exp\left(-\frac{B_l}{2}\right)}$$

Many clever methods have been devised for constructing and sampling from arbitrary posterior distributions. Markov chain simulation (also called Markov chain Monte Carlo, or MCMC) is a general method based on drawing values of \(\theta\) from approximate distributions and then correcting those draws to better approximate the target posterior distribution, \(p(\theta \mid y)\) \cite{34} \cite{35}. The sampling is done sequentially, with the distribution of the sampled draws depending on the last value drawn; hence, the draws form a Markov chain. (As defined in probability theory, a Markov chain is a sequence of random variables \(\theta_1, \theta_2, \ldots\), for which, for any \(t\), the distribution of \(\theta_t\) given all previous \(\theta\)'s depends only on the most recent value) The key to the method’s success, however, is not the Markov property but rather that the approximate distributions are improved at each step in the simulation, in the sense of converging to the target distribution \cite{35}.

A z-statistic should be calculated when the standard deviation of the population(s) is known. If the standard deviation is not known, then the standard error must be estimated using the standard deviation of the sample(s). Due to this estimation, we must use the t-distribution which is thicker in the tails to account for estimating the standard error with the sample standard deviation \cite{34}.

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

$$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

Until now, we have built the theory to apply to the case of four sample data, now what we will do, will be a particular case study, and see that Lema and definitions 1, 2 and 3 are applied naturally like the Equation (14)

Example particular case

Suppose \(H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4\). Select a random sample of size \(n_i\). From \(i\)th group \((i = 1, 2, 3, 4)\) so sample sizes are \((n_1, n_2, n_3, n_4)\)

$$\bar{x}_i = \frac{\sum_{j=1}^{n_i} x_{ij}}{n_i}$$

\(\bar{x}\) is the sample mean for the observations in all group combined

$$\bar{x} = \frac{\sum_{i=1}^{4} \sum_{j=1}^{n_i} x_{ij}}{n_1 + n_2 + n_3 + n_4}$$
Variability in the data, the deviation of an individual observation

\[ x_{ij} - \bar{x} = (x_{ij} - \bar{x}_j) + (\bar{x}_j - \bar{x}) \]

\[ \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} (x_{ij} - \bar{x})^2 = \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} (x_{ij} - \bar{x}_j)^2 + \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} (\bar{x}_j - \bar{x})^2. \]

To test the null hypothesis that the population means are all the same, use the test statistic

\[ t = \frac{[n_i (\bar{x}_i - \bar{x})^2 + n_z (\bar{x}_z - \bar{x})^2 + n_s (\bar{x}_s - \bar{x})^2 + n_4 (\bar{x}_4 - \bar{x})^2][n_1 + n_z + n_s + n_4 - 4]}{3[(n_1 - 1)s_1^2 + (n_z - 1)s_z^2 + (n_s - 1)s_s^2 + (n_4 - 1)s_4^2]} \]

Under \( H_0 \), this statistic has \( t \) distribution with \( k-1 \) and \( n-k \) degree freedom.

Now four three

Suppose \( H_0 : \mu_1 = \mu_2 = \mu_3 \). Select a random sample of size \( n_i \). From \( i \)th group \( (i = 1, 2, 3) \) so sample sizes are \((n_1, n_2, n_3)\)

\[ \bar{x}_i = \frac{\sum_{j=1}^{n_i} x_{ij}}{n_i} \]

\[ \bar{x} = \frac{\sum_{i=1}^{k} \sum_{j=1}^{n_i} x_{ij}}{n_1 + n_2 + n_3} \]

\[ \bar{x} = \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2 + n_3 \bar{x}_3}{n_1 + n_2 + n_3} \]

\[ \sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \bar{x})^2 = \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} (x_{ij} - \bar{x}_j)^2 + \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} (\bar{x}_j - \bar{x})^2. \]

\[ t = \frac{[n_1 (\bar{x}_1 - \bar{x})^2 + n_z (\bar{x}_z - \bar{x})^2 + n_s (\bar{x}_s - \bar{x})^2][n_1 + n_z + n_s - 3]}{2[(n_1 - 1)s_1^2 + (n_z - 1)s_z^2 + (n_s - 1)s_s^2]} \]

Then \( H_0 : \mu_1 = \mu_2 \)

\[ t = \frac{[n_1 (\bar{x}_1 - \bar{x})^2 + n_z (\bar{x}_z - \bar{x})^2][n_1 + n_z - 2]}{[(n_1 - 1)s_1^2 + (n_z - 1)s_z^2]} \]

and

\[ H_0 : \mu_3 \]

\[ t = \frac{[n_1 (\bar{x}_1 - \bar{x})^2][n_1 - 1]}{[(n_1 - 1)s_1^2]} \]

The principle of Bayes model is to compute posteriors based on specified priors and the likelihood function of data, the four groups of size 1110.

We began with a descriptive model of data from four groups, wherein the parameters were meaningful measures of central tendency, variance, and normality. Bayesian inference reallocates credibility to parameter values that are consistent with the observed data. The posterior distribution across the parameter values gives complete information about which combinations of parameter values are credible. In particular, from the posterior distri-
bution we can assess the credibility of specific values of interest, such as zero difference between means, or zero difference between standard deviations. We can also decide whether credible values of the difference of means are practically equivalent to zero, so that we accept the null value for practical purposes.

The Bayesian posterior distribution can also be used as a complete hypothesis for assessing power, that is, the probabilities of achieving research goals such as rejecting a null value, accepting a null value, or reaching a desired precision of estimation. The power estimation incorporates all the information in the posterior distribution by integrating across the credible parameter values, using each parameter-value combination to the extent it is credible. Figure 1 shows histograms of data that are labeled with $X_i$ on their abscissas, and these data are fixed at their empirically observed values.

3. Bayesian Multiple Linear Regression

$(x_1, \cdots, x_n)$ now, we need to find an adjustment function for the above data. A linear regression model where more than one variable involved is called multiple regression model [36]-[38]

\[ Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon \]  

(15)

where, we have $k$ regressors, parameters $\beta$ they are regression coefficients, then the Approach Bayesian Multiple Linear Regression is [36] [37]

\[
(Y \mid X, \beta, \sigma^2) \sim N(X, \beta, \sigma^2, \delta) 
\]  

(16)

The prior distribution of $(\beta, \sigma^2)$ is NIG (Normal-Inverse-Gamma) and it is given by $(b_0, B_0, n_0, S_0)$ thus $b_0, B_0, n_0, S_0$ they are hyperparameters.

\[
(\beta \mid \sigma^2, Y, X) \sim N(b_1, \sigma^2, B_1) 
\]

(17)

\[
(\beta \mid Y, X) \sim t_{n_0}(b_1, S, B). 
\]

\[
(\sigma^2 \mid \beta, Y, X) \sim IG\left(\frac{n_0}{2}, \frac{n_0 S_1}{2}\right). 
\]

Figure 1. Posterior predictive probability dimensions.
where IG (Inverse Gaussian)

If we denote by \( \omega = (Y, X) \) \( \theta = (\mu, \delta, \beta, \gamma) \)

We can express the density function as

\[
    f(w) = g(\theta) h(w) \exp\left( \sum_{j=0}^{4} \varphi_j(\theta) t_j \right).
\]

The conjugate prior distribution will be given by

\[
    p(\theta) \approx g(\theta)^{\omega} \exp\left( \sum_{j=0}^{4} \varphi_j(\theta) a_j \right).
\]

In the scheme MCMC

\[
    Y | \theta_i \sim N_n\left( A_i \theta_i, C_i \right). 
\]

where \( A_i \) is a matrix \((n \times p)\) and \( C_i \) is the covariance matrix and \( \theta_i \) is a vector dimension \( p \) of regression coefficients to do Bayesian estimation. Assume the prior distribution of \( \theta_i \) is

\[
    \theta_i \sim N_p\left( A_i \theta_2, C_i \right). 
\]

where \( A_i \) and \( C_i \) are matrices representing our beliefs about average and covariance of prior distribution. Take \( A_2 = I \) and using Bayesian approach, we obtain

\[
    \theta_i | Y \sim N_p\left( D d, D \right). 
\]

where

\[
    d = A_i C_i^{-1} Y + C_i A_i \theta_2. \\
    D^{-1} = A_i C_i^{-1} A_i + C_i^{-1}. 
\]

We can see that both matrices have two terms, one that only it depends on the prior and other that only it depends of data. This is very useful because in each iteration. We have to update only the last term. The question to be dealt is the choice of hyperparameters \( C_2 \) \( \theta_2 \)

\[
    y_i | X_i, z_i \sim N\left( X_i^{\prime} b + \beta_{k+1, i}, z_i \right). \\
    z_i \sim IG(\gamma, \delta)
\]

Then, we can write the equation as follows [38]

\[
    v(i) = \beta_0 + \sum_{j=1}^{p} \beta_j X_j(i) + \beta_{cov} X_{cov}(i)
\]

Figures 2-5 show the range HDI, which means Higher Density Interval. Values inside HDI have a greater probability density (credibility) that values outside this. Therefore, the 95% HDI includes the most incredible parameters values. There is a way that the posterior 95% HDI could exclude zero even when the data have a frequency of zero. It can happen if the prior already excludes zero. This interval is useful as a summary of the distribution and decision tool. The decision rule is simple. Any value outside of the 95% HDI is rejected [38].

In all four cases the focus is on assessing if the predictors were differentially predictive, for which we examine the posterior distribution of the differences standardized regression coefficients, given that the comparison is based in the normalization using the single sample.

As shown in Figure 2, the data to \( X_1 \sim f\left( X_2, X_3, X_4 \right) \), none of the coefficients are within the supposed.

We observed than for \( X_1 \) (sense positive of life) for differences in \( \beta \) are out of interval to be credible, indicating than \( X_1 \sim f\left( X_2, X_3, X_4 \right) \), cannot be a linear combination of the other dimensions.

In Figure 4, the outcomes for dataset \( X_2 \sim f\left( X_1, X_3, X_4 \right) \) where \( X_2 \) (satisfaction with life) \( \beta_2 - \beta_3 = 0.301, \)
Figure 2. Posterior distribution for $X_1$.

Figure 3. Posterior distribution for $X_2$.

Figure 4. Posterior distribution for $X_3$. 
HDI goes from 0.209 to 0.391. Therefore $X_1$ and $X_3$ are equally costly and if we want to avoid double cost of measuring both, then, is probable to be more effective assess $X_2$ than $X_3$.

In Figure 5 the outcomes for dataset $X_3 \sim f(X_1, X_2, X_3)$, where $X_3$ (personal fulfillment) $\beta_3 - \beta_1 = 0.238$, HDI goes from 0.149 to 0.325. Therefore $X_1$ and $X_2$ are equally costly and if we want to avoid double cost of measuring both, then, is probable to be more effective assess $X_1$ than $X_4$.

Figure 5 shows for dataset $X_4 \sim f(X_1, X_2, X_3)$, where $X_4$ (joy of living) $\beta_4 - \beta_1 = 0.106$, HDI goes from 0.0205 to 0.191. Therefore $X_1$ and $X_2$ are equally costly, and if we want to avoid double cost of measuring both, then, is probable to be more effective assess $X_1$ than $X_3$.

The hypotheses of work, was that if through sample analysis could infer that happiness, only one is affected by three dimensions $X_2$ (sense positive of life), $X_3$ (personal fulfillment) and $X_4$ (satisfaction with life). In this context, we note that the hypothesis was tested, the marked tendency on distributions in recent factors was sufficient to support this theory, on the other hand, through the Multilinear Regression Bayesian, also tested this hypothesis.

Due to recent revolutionary advances in Bayesian posterior computation via computer-intensive MCMC simulation techniques, difficulties with posterior computations can be overcome. A Bayesian state-space model is readily implemented using standard Bayesian software such as JAGS, BUGS, NIMBLE and STAN. One can therefore avoid writing one-off programs in a low-level language. Any modifications, such as different prior
distributions, applications to different data sets, or the use of different sampling distributions, require the change of just a single line in the code.

References


