مجلّه‌ی پژوهش‌های آماری ایران

J. Statist. Res. Iran **16** (2019): 447–464

دوره‌ی ۱۶، شماره‌ی ۲، پاییز و زمستان ۱۳۹۸، صص ۴۶۴–۴۴۷

DOI: 10.52547/jsri.16.2.447

DOR: 20.1001.1.25385771.2020.16.2.1.1

A New Control Chart Based on a Bivariate Beta Distribution

Shohreh Enami[†,⋆], Hamzeh Torabi[†,*] and Seyyed Taghi Akhavan Niaki[‡]

[†] Yazd University
[‡] Sharif University of Technology
[⋆] Payame noor University

Received: 2020/15/10 Approved: 2021/28/07

Abstract. In many practical situations, the quality of a product can be measured based on some quality characteristics in terms of the sum of the percentage of these characteristics utilities. In these cases, control charts based on multivariate beta distribution can be used to monitor the process. This study aims to introduce a new control chart for monitoring the quality of products when two quality characteristics follow a bivariate Beta distribution. The efficacy of the proposed control chart is evaluated using the average run length criterion using a simulation study. In the case that the parameters of this distribution are unknown, the maximum likelihood method is applied. Then, using a simulation study, the performance of the proposed charts, in two cases known and unknown parameters are compared.

Keywords. Bivariate beta distribution; maximum likelihood method; average run length.

MSC 2010: 62D05.

1 Introduction

Control charts are one of the essential techniques for ensuring the quality of production. In many practical situations, researchers are interested in

[*] Corresponding author

investigating the number of defects, or the proportions of defective components in a product expressed in the interval $[0, 1]$, both of which often come from Bernoulli experiments. Examples are the number of defective parts in a production lot (percentage) or the proportion of medicine components of medicine in pharmaceutical industries (proportion). In this case, the np and p control charts are applied to monitoring and improving the production process. However, there are many situations in which the proportions are not the results of Bernoulli's experiments, although they fall within the $[0, 1]$-interval. The famous probability distribution that applies to this type of random variable is the beta distribution. There are some works performed in the literature in the field of control charts for beta distribution Bayer et al. (2018), Lee Ho and Bourguignon (2018) and Sant'Anna and Ten Caten (2012).

Sometimes researchers are interested in the investigation of more than one correlated quality characteristic simultaneously. redAdamski proposed a control chart for multivariate beta distribution (Adamski et al. (2012)). In this paper, we will propose a control chart that is different from the proposed control chart by Adamski. This paper focuses on the investigation of two correlated quality characteristics that follow the beta distribution. red Some applications of this distribution were introduced in Nadarajah (2007a), Sarabia et al. (2014) and Ghosh, I. (2019). Also, some bivariate beta distributions were introduced in Nadarajah and Kotz (2005), Nadarajah (2007a), Nadarajah (2007b), Nadarajah et al. (2017), Olkin and Liu (2003) and Orozco-Castaneda et al. (2012). Nevertheless, the one introduced by Libby and Novick Libby and Novick (1982) is used in the current work to design a control chart for monitoring the proportion of defective components in a product (the reason for this choice will be explained in Section 2).

The remainder of this paper is organized as follows:
In Section 2, the structure of a bivariate beta distribution (BB) introduced by Libby and Novick Libby and Novick (1982) is first reviewed. Then, by defining a new statistic and calculating its exact distribution, a control chart for an important parameter of this distribution is proposed. In Section 3, the efficacy of the proposed control chart will be investigated in terms of the average run length (ARL) criterion. The way to estimate the parameters of the bivariate beta distribution is explained in Section 4. In Section 5, the performance of the proposed chart when the parameters are estimated is compared with the one when the parameters are assumed known. Finally, some conclusions are presented in Section 6.

2 A Bivariate Beta Control Chart

Let Y_j, $j = 0, 1, 2$, be random variables following a gamma distribution with the shape parameter α_j and the scale parameter β. Assume that Y_0, Y_1 and Y_2 are mutually independent. Defining $X_1 = \dfrac{Y_1}{Y_1 + Y_0}$ and $X_2 = \dfrac{Y_2}{Y_2 + Y_0}$, the vector $\boldsymbol{X} = (X_1, X_2)$ follows a bivariate beta distribution, since here each X_j, $j = 1, 2$, follows a beta distribution marginally with the shape parameter α_j and the scale parameter α_0. As such, their means and variances are obtained by $\mu_j = E(X_j) = \dfrac{\alpha_j}{\alpha_j + \alpha_0}$ and $V(X_j) = \dfrac{\alpha_j \alpha_0}{(\alpha_j + \alpha_0 + 1)(\alpha_j + \alpha_0)^2}$, respectively.

The probability density function (pdf) of the above bivariate beta distribution can be written as

$$f_{X_1,X2}(x_1, x_2) = \frac{x_1^{\alpha_1 - 1} x_2^{\alpha_2 - 1} (1 - x_1)^{(\alpha_0 + \alpha_2 - 1)} (1 - x_2)^{(\alpha_0 + \alpha_1 - 1)}}{\beta(\alpha_0, \alpha_1, \alpha_2)(1 - x_1 x_2)^{\sum_{i=0}^{2} \alpha_i}}, \qquad (1)$$

where $0 < x_j < 1$, $\alpha_j > 0$, $\beta(\alpha_0, \alpha_1, \alpha_2) = \dfrac{\Gamma(\alpha_0)\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_0 + \alpha_1 + \alpha_2)}$. Interested readers are referred to Nadarajah (2007a), Sarabia et al. (2014) and Wright (1937) for different applications of this distribution.

Let the random variable X_j, $j = 1, 2$, be the percent of defective parts in mass productions concerning a quality characteristic j. Also, assume that $\boldsymbol{X} = (X_1, X_2)$ follows a bivariate beta distribution with the above-introduced structure.. Define the statistic D as follows:

$$D = X_1 + X_2. \qquad (2)$$

The bivariate beta distribution structure and its application in the control chart can be described with the bellow example: Hepatitis C is an infectious disease caused by the blood-borne Hepatitis C virus. The disease is not seasonal and the most common routes of transmission include unsafe hypodermic needle injection practices, inadequate sterilization of medical equipment, and unscreened blood and blood products World Health Organization (2002). In monitoring Hepatitis C cases, the goal of health organizations is to quickly trigger alarms once an upward trend outbreak occurs so that public health events can be handled promptly. Assume that X_1 be the percent of disease in Victoria and X_2 be the percent of disease in South Australia. These two states share borders, which potentially gives rise to dependence on disease occurrences across those states. In this case, The statistic D, indicates the sum of the percentage of disease.

In this paper, we use the structure proposed by Libby and Novick Libby and Novick (1982), because, in this structure, each X_i can take value in the interval $[0,1]$, separately. So, the value of the statistic D is in the interval $[0,2]$ and there are not any limitations in the percent of defective parts.

In Theorem 1, the pdf of the statistic D is given.

***Theorem* 1.** *The pdf of D is given by*

$$f_D(d) = \begin{cases} \int_0^d f_{(D,T)}(d,t)dt, & 0 < d < 1, \\ \int_{d-1}^1 f_{(D,T)}(d,t)dt, & 1 < d < 2, \end{cases}$$

where

$$f_{D,T}(d,t) = \frac{t^{\alpha_1-1}(1-t)_2^{\alpha_0+\alpha_2-1}(d-t)^{(\alpha_2-1)}(1-d+t)^{(\alpha_0+\alpha_1-1}}{\beta(\alpha_0,\alpha_1,\alpha_2)(1-t(d-t))^{\sum_{i=0}^2 \alpha_i}},$$

and $\beta(\alpha_0,\alpha_1,\alpha_2)$ *defined in* (1).

Proof. Let $D = X_1 + X_2$ and $T = X_1$. Then $X_1 = T$ and $X_2 = D - T$. From 1 and using the change of variables technique, the joint probability density function of D and T is obtained as the following form

$$\begin{aligned} f_{D,T}(d,t) &= f_{X_1,X_2}(t, d-t)|J| \\ &= \frac{t^{\alpha_1-1}(1-t)^{\alpha_0+\alpha_2-1}(d-t)^{\alpha_2-1}(1-d+t)^{\alpha_0+\alpha_1-1}}{\beta(\alpha_0,\alpha_1,\alpha_2)(1-t(d-t))^{\sum_{i=0}^2 \alpha_i}}, \end{aligned}$$

As a result, the proof is completed when the marginal distribution of D is obtained by integrating the above joint probability density function with respect to t.

To design the control chart, Type-I error is first fixed at α to design the control chart. Then, the lower control limit (LCL) is determined by solving the following equation

$$P(D < \text{LCL}) = \int_0^{\text{LCL}} f_D(x)dx = \int_0^{\text{LCL}} \int_0^x f_{D,T}(x,t)dtdx = \frac{\alpha}{2}.$$

If the value of LCL cannot be obtained from the above equation (i.e., when LCL is between 1 and 2), then the lower control limit (LCL) is determined by solving the following equation

$$\begin{aligned} P(D < \mathrm{LCL}) &= \int_0^{\mathrm{LCL}} f_D(x)dx \\ &= \int_0^1 \int_0^x f_{D,T}(x,t)dtdx + \int_1^{\mathrm{LCL}} \int_{x-1}^1 f_{D,T}(x,t)dtdx = \frac{\alpha}{2}. \end{aligned}$$

Besides, one can obtain the upper control limit (UCL) by solving the following equation

$$P(D > \mathrm{UCL}) = \int_{\mathrm{UCL}}^2 f_D(x)dx = \int_{\mathrm{UCL}}^2 \int_{x-1}^1 f_{D,T}(x,t)dtdx = \frac{\alpha}{2}.$$

Again, when UCL cannot be obtained by the above equation (i.e., when UCL is between 0 and 1), then it is determined by solving the following equation

$$\begin{aligned} P(D > \mathrm{UCL}) &= \int_{\mathrm{UCL}}^2 f_D(x)dx \\ &= \int_{\mathrm{UCL}}^1 \int_0^x f_{D,T}(x,t)dtdx + \int_1^2 \int_{x-1}^1 f_{D,T}(x,t)dtdx = \frac{\alpha}{2}. \end{aligned}$$

□

3 Performance Evaluation of the Control Chart

In this section, Types I and II errors of the proposed chart are first discussed. Then, its performance is assessed using the average run length (ARL) criterion.

3.1 Types I and II Errors

When the process is in control (i.e., $\mu_j = \mu_{j0}$), but the control chart plots a sample out-of-control, then Type I error occurs. The probability of Type I error can be obtained as follows

$$\alpha = P(D > \mathrm{UCL}|\mu_1 = \mu_{10}, \mu_2 = \mu_{20}) + P(D < \mathrm{LCL}|\mu_1 = \mu_{10}, \mu_2 = \mu_{20}),$$

where μ_{10} and μ_{20} are average proportions when the process is in control. On the other hand, when the parameters shift from the in-control values to any other undesirable values such as μ_{j1}, but the control chart plots a sample inside the control limits, then Type II error occurs. The probability of the Type II error can be obtained as follows

$$\beta = P(\text{UCL} < D < \text{UCL}|\mu_1 = \mu_{11}, \mu_2 = \mu_{21}),$$

where μ_{11} and μ_{21} are the average proportions when the process is out-of-control.

3.2 Average Run Length

The performance of the control charts is usually evaluated by the ARL criterion. There are two types of ARLs, i.e., in-control (ARL_0) and out-of-control (ARL_1) [6]. When the process is in control, large values of $\text{ARL}_0 = \dfrac{1}{\alpha}$ is desirable whereas in a case of out-of-control, small values of $\text{ARL}_1 = \dfrac{1}{1-\beta}$ is interested. In what follows, the performance of the control chart is assessed using the above ARLs.

Consider two customary targets ARL_0s 200 and 370 when $\alpha_{00} = 6$ and $\alpha_{00} = 9$ and assume $\mu_{11} = \mu_{10} + c$ and $\mu_{21} = \mu_{20} + c$, where c is a shift value. Tables 1-4 present the performance of the control chart in terms of ARL_1 when the average proportion shifts (Tables 1 and 2, for $\text{ARL}_0 = 200$ and Tables 3 and 4, for $\text{ARL}_0 = 370$). In these tables, some different shift values (c) are shown in the first column. We consider the correlation coefficients $0.1, 0.35, 0.4, 0.5, 0.7$ and 0.8 to consider different dependence (small, moderate, and strong dependence). So, for some values of α_{10}, α_{20}, proportional to these correlation coefficients, calculated ARL_1. For example, for $\alpha_{00} = 6$ and $\rho = 0.35$, α_{10} and α_{20} obtain 5 and 2.2, respectively. The ARL_1s are obtained using the $MATEMATICA$ software.

Note that $\mu_{10}, \mu_{11}, \mu_{20}$ and μ_{21} must be between 0 and 1, based on some values of c, μ_{11} and μ_{21} can be obtained negative or more than 1. For example, in the last column of Table 4, for $c = 0.5$, μ_{11} is equal to 1.31, or in the first column of Table 4, for $c = -0.1$, μ_{21} is equal to -0.037. The conclusions made based on the results in Tables 1-4 are summarized as follows:

(i) The ARL_1 decreases as c increases for all cases.

(ii) With increasing ρ, the values of ARL_1 increases, when $0 < \rho < 0.45$.

Table 1. The ARL_1 of the proposed control chart, when $\alpha_{00} = 6$ and $ARL_0 = 200$.

	ρ					
	0.1	0.35	0.4	0.5	0.71	0.81
	$\binom{\alpha_{10}}{\alpha_{20}} = \binom{1}{0.4}$	$\binom{\alpha_{10}}{\alpha_{20}} = \binom{5}{2.2}$	$\binom{\alpha_{10}}{\alpha_{20}} = \binom{5}{3.2}$	$\binom{\alpha_{10}}{\alpha_{20}} = \binom{5}{8}$	$\binom{\alpha_{10}}{\alpha_{20}} = \binom{15}{16.5}$	$\binom{\alpha_{10}}{\alpha_{20}} = \binom{25.5}{30}$
c	$\binom{\mu_{10}}{\mu_{20}} = \binom{0.14}{0.16}$	$\binom{\mu_{10}}{\mu_{20}} = \binom{0.45}{0.27}$	$\binom{\mu_{10}}{\mu_{20}} = \binom{0.45}{0.35}$	$\binom{\mu_{10}}{\mu_{20}} = \binom{0.45}{0.57}$	$\binom{\mu_{10}}{\mu_{20}} = \binom{0.71}{0.73}$	$\binom{\mu_{10}}{\mu_{20}} = \binom{0.81}{0.83}$
−0.1	...	20.49	22.25	22.45	14.04	7.52
−0.09	...	25.89	27.91	27.91	17.55	9.39
−0.08	...	33.02	35.31	35.01	22.32	12.03
−0.06	...	54.87	57.69	56.67	38.1	21.56
−0.04	...	91.99	94.89	92.96	68.64	44.2
0.00	**200.3**	**200.1**	**200.1**	**200.11**	**200.1**	**200.03**
0.04	123.05	183.53	184.6	182.45	147.98	99.53
0.06	78.45	140.95	142.64	137.87	92.01	46.28
.08	51.83	104.08	105.66	99.54	54.84	20.13
0.09	42.72	89.1	90.48	84.07	41.96	12.97
0.1	35.52	76.25	77.43	70.91	31.94	8.24
0.15	15.69	35.59	35.84	30.11	7.64	...
0.18	10.31	22.93	22.85	18.04	3.17	...
0.2	7.99	17.23	17.03	12.84	1.84	...
0.25	4.56	8.67	8.35	5.57	...	...
0.3	2.86	4.58	4.27	2.56	...	...
0.35	1.98	2.58	2.34	1.36	...	...
0.4	1.49	1.6	1.44	1.06	...	...
0.48	1.12	1.04	1.01	...	...	...
0.5	1.08	1	1	...	...	...

Table 2. The ARL_1 of the proposed control chart, when $\alpha_{00} = 9$ and $\text{ARL}_0 = 200$.

	ρ					
	0.1	0.35	0.4	0.5	0.71	0.81
	$\binom{\alpha_{10}}{\alpha_{20}} = \binom{1}{1}$	$\binom{\alpha_{10}}{\alpha_{20}} = \binom{6}{4}$	$\binom{\alpha_{10}}{\alpha_{20}} = \binom{6}{6}$	$\binom{\alpha_{10}}{\alpha_{20}} = \binom{10}{9}$	$\binom{\alpha_{10}}{\alpha_{20}} = \binom{22}{24}$	$\binom{\alpha_{10}}{\alpha_{20}} = \binom{40}{40}$
c	$\binom{\mu_{10}}{\mu_{20}} = \binom{0.14}{0.16}$	$\binom{\mu_{10}}{\mu_{20}} = \binom{0.45}{0.27}$	$\binom{\mu_{10}}{\mu_{20}} = \binom{0.45}{0.35}$	$\binom{\mu_{10}}{\mu_{20}} = \binom{0.45}{0.57}$	$\binom{\mu_{10}}{\mu_{20}} = \binom{0.71}{0.73}$	$\binom{\mu_{10}}{\mu_{20}} = \binom{0.81}{0.83}$
−0.1	...	15.18	16.55	16.82	9.94	5.26
−0.09	...	19.55	21.14	21.29	12.6	6.61
−0.08	...	25.5	27.35	27.33	16.3	8.54
−0.06	7.13	44.9	47.31	46.71	29.3	15.85
−0.04	23.76	81.18	83.9	82.6	57.84	34.38
0.00	**200.2**	**200.01**	**200.2**	**200.3**	**200.06**	**200.3**
0.04	109.21	158.16	158.68	153.5	114.77	70.11
0.06	64.48	110.12	110.63	103.68	62.13	27.66
.08	29.83	74.79	75	67.96	32.9	10.66
0.09	31.3	61.67	61.72	54.95	23.9	6.63
0.1	25.72	50.96	50.88	44.48	209.96	4.29
0.15	10.15	20.52	20.12	15.88	3.7	...
0.18	6.39	12.37	11.94	8.84	1.71	...
0.2	4.8	8.98	8.57	6.08	1.23	...
0.25	2.74	4.33	4.01	2.63	...	...
0.3	1.79	2.35	2.14	1.48	...	...
0.35	1.34	1.48	1.35	1.04	...	...
0.4	1.12	1.12	1.06	1	...	...
0.48	1.01	1	1	...	...	...
0.5	1	1	1	...	...	...

Table 3. The ARL_1 of the proposed control chart, when $\alpha_{00} = 6$ and $\text{ARL}_0 = 370$.

	ρ					
	0.1	0.35	0.4	0.5	0.71	0.81
	$\binom{\alpha_{10}}{\alpha_{20}} = \binom{1}{0.4}$	$\binom{\alpha_{10}}{\alpha_{20}} = \binom{5}{2.2}$	$\binom{\alpha_{10}}{\alpha_{20}} = \binom{5}{3.2}$	$\binom{\alpha_{10}}{\alpha_{20}} = \binom{5}{8}$	$\binom{\alpha_{10}}{\alpha_{20}} = \binom{15}{16.5}$	$\binom{\alpha_{10}}{\alpha_{20}} = \binom{25.5}{30}$
c	$\binom{\mu_{10}}{\mu_{20}} = \binom{0.14}{0.16}$	$\binom{\mu_{10}}{\mu_{20}} = \binom{0.45}{0.27}$	$\binom{\mu_{10}}{\mu_{20}} = \binom{0.45}{0.35}$	$\binom{\mu_{10}}{\mu_{20}} = \binom{0.45}{0.57}$	$\binom{\mu_{10}}{\mu_{20}} = \binom{0.71}{0.73}$	$\binom{\mu_{10}}{\mu_{20}} = \binom{0.81}{0.83}$
−0.1	...	29.55	32.18	32.61	19.4	9.79
−0.09	...	38.2	41.26	41.44	24.79	12.51
−0.08	...	49.88	53.41	53.17	32.26	16.42
−0.06	...	87.26	91.72	90.27	58	31.14
−0.04	...	154.68	159.38	156.29	112.48	68.5
0.00	**370.12**	**370.02**	**370.3**	**370.2**	**370.13**	**370.1**
0.04	226.64	335.56	340.33	335.08	269.5	177.19
0.06	141.54	253.13	258.57	248.70	163.92	79.71
.08	91.76	184.15	188.71	176.8	95.5	33.21
0.09	74.93	156.58	160.51	148.21	72.18	20.79
0.1	61.69	133.17	136.44	124.14	54.21	12.77
0.15	25.94	60.09	61	50.68	11.78	...
0.18	16.53	37.83	37.98	29.49	4.44	...
0.2	12.53	27.96	27.8	20.53	2.35	...
0.25	6.75	13.42	12.95	8.32	...	...
0.3	3.99	6.68	6.22	3.47	...	...
0.35	2.58	3.49	3.1	1.62	...	...
0.4	1.81	1.98	1.75	1.09	...	...
0.48	1.24	1.1	1.03	...	...	...
0.5	1.16	1	1	...	...	...

Table 4. The ARL_1 of the proposed control chart, when $\alpha_{00} = 6$ and $\text{ARL}_0 = 370$.

	ρ					
	0.1	0.35	0.4	0.5	0.71	0.81
	$\binom{\alpha_{10}}{\alpha_{20}} = \binom{1}{1}$	$\binom{\alpha_{10}}{\alpha_{20}} = \binom{6}{4}$	$\binom{\alpha_{10}}{\alpha_{20}} = \binom{6}{6}$	$\binom{\alpha_{10}}{\alpha_{20}} = \binom{10}{9}$	$\binom{\alpha_{10}}{\alpha_{20}} = \binom{22}{24}$	$\binom{\alpha_{10}}{\alpha_{20}} = \binom{40}{40}$
c	$\binom{\mu_{10}}{\mu_{20}} = \binom{0.14}{0.16}$,	$\binom{\mu_{10}}{\mu_{20}} = \binom{0.45}{0.27}$	$\binom{\mu_{10}}{\mu_{20}} = \binom{0.45}{0.35}$	$\binom{\mu_{10}}{\mu_{20}} = \binom{0.45}{0.57}$	$\binom{\mu_{10}}{\mu_{20}} = \binom{0.71}{0.73}$	$\binom{\mu_{10}}{\mu_{20}} = \binom{0.81}{0.83}$
−0.1	...	21.41	23.4	23.74	13.45	6.71
−0.09	...	28.24	30.6	30.73	17.45	8.64
−0.08	...	37.78	40.55	40.35	23.16	11.48
−0.06	8.9	70.27	73.92	72.59	44.01	22.63
−0.04	33.49	134.96	139.04	135.95	92.79	53.8
0.00	**370.16**	**370.1**	**370.1**	**370.2**	**370.04**	**370.3**
0.04	196.42	285.91	288.65	280.3	203.44	119.2
0.06	112.91	194.54	196.77	158.09	106.72	44.84
.08	67.98	129.53	130.75	118.75	54.68	16.23
0.09	53.65	105.78	106.56	94.99	38.98	9.71
0.1	42.79	86.58	86.98	76.03	27.74	6.05
0.15	15.79	33.13	32.59	25.5	5.15	...
0.18	9.54	19.28	18.64	13.55	2.1	...
0.2	7.05	13.66	13.04	9.02	1.37	...
0.25	3.69	6.13	5.66	3.52	...	...
0.3	2.24	3.06	2.75	1.65	...	...
0.35	1.55	1.76	1.57	1.08	...	...
0.4	1.22	1.21	1.12	1	...	...
0.48	1.03	1	1	...	...	...
0.5	1	1	1	...	...	...

(iii) With increasing ρ, the values of ARL_1 decreases, when $0.45 < \rho < 1$.

(iv) iv. The results in Tables 1 and 2 (or Tables 3 and 4) show that with increasing α_{00}, the values of ARL_1 decrease, for a constant correlation. For example, if $\alpha_{00} = 6, \text{ARL}_0 = 370, \rho = 0.35$ and $c = 0.1$, then $\text{ARL}_1 = 133.17$; and if $\alpha_{00} = 9, \text{ARL}_0 = 370, \rho = 0.35$ and $c = 0.1$, then $\text{ARL}_1 = 86.58$.

4 Maximum Likelihood Estimation

The application of the proposed control chart involves two phases. In Phase I, the parameters of the bivariate Beta distribution are first estimated for an in-control process since they are unknown in many real-world applications. Then, the control limits are obtained based on the estimates. These control limits are used in Phase II to monitor the process. In what follows, the maximum likelihood estimation (MLE) method is employed to estimate the parameters of the BB distribution. The log-likelihood function is derived as follows, having n observations from the bivariate density in Equation (1)

$$\begin{aligned} l(\alpha_0, \alpha_1, \alpha_2) =& nln\Gamma(\alpha_0 + \alpha_1 + \alpha_2) - nln\Gamma(\alpha_0) - nln\Gamma(\alpha_2) \\ &+ (\alpha_1 - 1)\sum_{i=1}^{n} x_{i1}) + (\alpha_2 - 1)\sum_{i=1}^{n}\log(x_{i2}) \\ &+ (\alpha_0 + \alpha_2 - 1)\sum_{i=1}^{n}\log(1 - x_{i1}) \\ &+ (\alpha_0 + \alpha_1 - 1)\sum_{i=1}^{n}\log(1 - x_{i2}) \\ &- (\alpha_0 + \alpha_1 + \alpha_2)\sum_{i=1}^{n}\log(1 - x_{i1}x_{i2}). \end{aligned} \tag{3}$$

Differentiating (3) with respect to α_0, α_1 and α_2, we get

$$\begin{aligned}\frac{\partial \ell}{\partial \alpha_0} =& n\Psi(\alpha_0+\alpha_1+\alpha_2) - n\Psi(\alpha_0) + \sum_{i=1}^{n} \log(1-x_{i1}) \\ &+ \sum_{i=1}^{n} \log(1-x_{i2}) - \sum_{i=1}^{n} \log(1-x_{i1}x_{i2}), \\ \frac{\partial \ell}{\partial \alpha_1} =& n\Psi(\alpha_0+\alpha_1+\alpha_2) - n\Psi(\alpha_1) + \sum_{i=1}^{n} \log(x_{i1}) \\ &+ \sum_{i=1}^{n} \log(1-x_{i2}) - \sum_{i=1}^{n} \log(1-x_{i1}x_{i2}), \\ \frac{\partial \ell}{\partial \alpha_2} =& n\Psi(\alpha_0+\alpha_1+\alpha_2) - n\Psi(\alpha_2) + \sum_{i=1}^{n} \log(x_{i2}) \\ &+ \sum_{i=1}^{n} \log(1-x_{i1}) - \sum_{i=1}^{n} \log(1-x_{i1}x_{i2}), \end{aligned} \tag{4}$$

where $\Psi(\alpha) = \dfrac{\Gamma'(\alpha)}{\Gamma(\alpha)}$ is digamma function [3]. Setting Equations (4) to zero and solving them simultaneously results in the MLE estimates of α_0, α_1 and α_2.

Monte Carlo simulation experiments are performed in the next section to access the performance of the proposed control chart in various out-of-control conditions when the parameters of the BB distribution are unknown and have to be estimated based on some in-control samples of the process.

5 Simulation Experiments

In this section, the performance of the proposed control chart is evaluated when the parameters are unknown. When the parameters are unknown, in-control parameters have to be estimated, and then the control limits are obtained. In this case, ARL_0 is a random variable with mean AARL and standard deviation SDARL (Salehet al. (2015)). Table 5 displays the in-control AARL values for various values of n (number of samples) and five different processes with five different correlations. These values were obtained in $100,000$ replications of the process using the R software. Here, it

Table 5. In-control AARL for different values of n, when desired $\text{ARL}_0 = 370$.

n	ρ				
	0.1	0.35	0.4	0.5	0.7
60	371.7	375.81	376.4	380.18	381.17
65	**370.4**	375.21	376.35	376.17	380.78
100	367.44	**370.3**	370.28	372.39	374.00
110	366.87	369.6	**370.12**	370.7	373.06
140	366.42	368.8	369.00	**370.3**	371.29
180	365.46	367.7	366.95	368.17	**370.20**
200	365.11	367.61	366.93	367.51	369.81

is assumed that ARL_0 is 370 when the in-control parameters are known (i.e the desired ARL_0 is equal to 370). Table 5 shows that with increasing the correlation, the number of samples required to estimate the parameters until achieving the desired ARL_0, is increased. For example, if $\rho = 0.1$, then the number of samples required to estimate of parameters until achieving the desired ARL_0 is 65, whereas, for $\rho = 0.4$, we have n as 110.

Similarly, out-of-control AARL, as well as out-of-control ARL, were obtained for various shifts and different correlations (Table 6). Once again, 100,000 simulation experiments were performed for each out-of-control process. Table 6 shows, in most cases the performance of the proposed chart is better when the parameters are known rather than estimated (comparing AARL and ARL_1 in each cell). Except for $c > 0.2$, that the values of AARL and ARL are close to each other.

Table 7 gives the values of out-of-control SDARL and SDRL for different shifts and different correlations. The obtained results show that the performance of the proposed chart in terms of the standard deviation criterion is better when the parameters are estimated rather than they are known. In other words, SDARL values are smaller than SDRL values.

Since the performance of the control chart was not the same using the AARL and SDARL criteria, we applied the coefficient of the variations of average run length (CVARL). CVARL is defined as

$$\text{CVARL} = \frac{\text{SDARL}}{\text{AARL}}.$$

Finally, Table 8 shows the values of out-of-control CVARL and CVRL for

Table 6. Comparing AARL and ARL_1 for various shifts, when $\alpha_{00} = 6$ and $ARL_0 = 370$.

c	(ρ, n)									
	$(0.1, 65)$		$(0.35, 100)$		$(0.4, 110)$		$(0.5, 140)$		$(0.7, 180)$	
	ARL_1	AARL	ARL_1	AARL	ARL_1	AARL	ARL_1	AARL	ARL_1	AARL
0.04	226.64	272.43	335.56	356.31	340.33	356.27	335.08	374.84	269.5	277.5
0.06	141.54	171.75	253.13	276.84	258.57	277.9	248.7	262.61	163.92	169.69
0.08	91.76	109.79	184.15	201.94	188.71	204.4	176.81	187.71	95.5	98.34
0.10	61.69	73.34	133.17	146.3	136.44	146.59	124.144	130.98	54.217	55.56
0.15	25.95	29.99	60.09	64.69	61	65.45	50.68	53.12	11.78	11.89
0.18	16.52	18.72	37.83	40.76	37.98	40.33	29.49	30.65	4.47	4.46
0.20	12.53	14.07	27.96	29.88	27.81	29.39	20.54	21.29	2.36	2.36
0.30	3.99	4.28	6.68	6.93	6.22	6.39	3.47	3.51	...	...
0.40	1.81	1.88	1.98	2.02	1.75	1.76	1.09	1.05	...	...
0.50	1.16	1.19	1	1.04	1	1.01	...	...	...	...

different shifts and different correlations. The obtained results show that except for small values of correlation and $c < 0.1$, in all cases, the performance of the proposed chart is better when the parameters are estimated. For example, if $\rho = 0.7$ and $c = 0.04$, then CVARL $= 0.516$ and CVRL $= 0.998$. When parameters are unknown, for estimating them, a large number of samples are needed to achieve the desired ARL0 (Table 5). On the other hand, taking more Phase I samples for the parameter estimation leads to better detection performance.

6 Conclusions and Future Researches

In the present article, the structure of a bivariate beta distribution introduced by Libby and Novick Libby and Novick (1982) was first reviewed. Then, by defining a statistic and obtaining its exact distribution, a control chart was proposed to monitor processes modelled by this distribution. The efficiency of this chart was evaluated in terms of out-of-control ARL (when ARL_0 remains constant), for different dependencies (small, moderate, and strong dependence). The simulation studies showed that in terms of the AARL criterion, in most cases, the performance of the proposed chart is better when the parameters are known than when the parameters are estimated. However, the opposite conclusion is made in terms of the SDARL criterion.

Table 7. Comparing **SDARL** and SDRL for various shifts, when $\alpha_{00} = 6$ and $\text{ARL}_0 = 370$.

c	(ρ, n)									
	$(0.1, 65)$		$(0.35, 100)$		$(0.4, 110)$		$(0.5, 140)$		$(0.7, 180)$	
	SDRL	SDARL	SDRL	SDARL	SDRL	SDARL	SDRL	SDARL	SDRL	SDARL
0.04	226.14	281.06	335.06	239.99	339.83	225.88	334.58	196.71	269	143.2
0.06	141.04	176.16	252.63	197.35	258.07	184.55	248.21	154.26	163.42	86.25
0.08	91.26	108.76	183.65	144	182.21	137.93	176.3	109.53	95	47.63
0.10	61.195	69.032	132.67	101.72	135.95	95.98	123.64	73.77	53.71	25.78
0.15	25.44	25.66	59.59	41.87	60.5	40.34	50.18	28.01	11.27	4.52
0.18	16.02	14.95	37.33	25.33	37.47	23.43	28.99	15.06	3.91	1.34
0.20	12.02	10.56	27.46	17.77	27.3	16.45	20.03	10.01	1.79	0.52
0.30	3.46	2.37	6.16	3.17	5.69	2.69	2.93	1.07	...	...
0.40	1.22	0.64	1.39	0.538	1.14	0.39	0.362	0.04	...	...
0.50	0.44	0.17	0.19	0.04	0.081	0.01	...	...	...	...

Table 8. Comparing **CVARL** and CVRL for various shifts, when $\alpha_{00} = 6$ and $\text{ARL}_0 = 370$.

c	(ρ, n)									
	$(0.1, 65)$		$(0.35, 100)$		$(0.4, 110)$		$(0.5, 140)$		$(0.7, 180)$	
	CVRL	CVARL	CVRL	CVARL	CVRL	CVARL	CVRL	CVARL	CVRL	CVARL
0.04	0.997	1.03	0.998	0.673	0.998	0.634	0.998	0.565	0.998	0.516
0.06	0.996	1.02	0.998	0.713	0.998	0.664	0.998	0.587	0.997	0.508
0.08	0.994	0.99	0.997	0.713	0.997	0.647	0.997	0.583	0.994	0.48
0.10	0.991	0.94	0.996	0.695	0.996	0.654	0.996	0.563	0.991	0.464
0.15	0.98	0.85	0.991	0.65	0.991	0.616	0.99	0.527	0.956	0.38
0.18	0.969	0.798	0.986	0.621	0.986	0.58	0.983	0.491	0.88	0.3
0.20	0.995	0.75	0.981	0.594	0.982	0.56	0.975	0.469	0.759	0.22
0.30	0.86	0.55	0.922	0.459	0.916	0.42	0.843	0.306	...	...
0.40	0.67	0.34	0.704	0.267	0.654	0.22	0.297	0.04	...	...
0.50	0.377	0.142	0.187	0.04	0.081	0.01	...	...	...	...

So, we applied CVARL criterion. The obtained results show that except for small values of correlation and $c < 0.1$, in all cases the performance of the proposed chart is better when the parameters are estimated. The proposed chart can be used in the medical in monitoring the proportions of diseases and also in the industry for monitoring the proportions of defective components in a product. The proposed control chart can be extended for a multivariate beta distribution for more than two quality characteristics using other sampling schemes as future researches.

Acknowledgments

The authors sincerely thank the associate editor and anonymous reviewers for their constructive suggestions and comments, which helped us improve the article's presentation..

References

Adamski, K., Human, S., and Bekker, A. (2012). A Generalized Multivariate Beta Distribution: Control Charting When the Measurements are from an Exponential Distribution. *Statistical Papers*, **53**, 1045-1064.

Alt, F.B. (1985). *Multivariate Quality Control*. John Wiley and Sons, 110-122.

Bayer, F., Tondolo, C., and Müller, F. (2018). Beta Regression Control Chart for Monitoring Fractions and Proportions. *Computers and Industrial Engineering*, **119**, 416-426.

Ghosh, I. (2019). On the Reliability for Some Bivariate Dependent Beta and Kumaraswamy Distributions: A Brief Survey. *Stochastics and Quality Control*, **34**, 115-121.

Jolevska-Tuneska, B., and Jolevski, I. (2013). Some Results on the Digamma Function. *Applied Mathematics and Information Sciences*, **7**, 167-170.

Lee Ho, L., and Bourguignon, M. (2018). Control Charts to Monitor Rates and Proportions. *Quality and Reliability Engineering International*, **35**, 1-10.

Libby, D., and Novick, M. (1982). Multivariate Generalized Beta Distributions with Applications to Utility Assessment. *Educational Statistics*, **7**, 271-294.

Montgomery, D.C. (2013). *Design and Analysis of Experiments*, John Wiley and Sons, 110-122.

Nadarajah, S. (2007). A New Bivariate Beta Distribution with Application to Drought Data. *Metron*. **LXV (2)**, 153-174.

Nadarajah, S. (2007). The Bivariate F2 Beta Distribution. *American Journal of Mathematical and Management Sciences*. **27**, 351-368.

Nadarajah, S., Hsing Shih, S., and Nagar, D. (2017). A New Bivariate Beta Distribution. *Statistics*, **51**, 455-474.

Nadarajah, S., and Kotz, S. (2005). Some Bivariate Beta Distributions. *Statistics*, **39**, 457-466.

Olkin, I., and Liu, R. (2003). A Bivariate Beta Distribution. *Statistics and Probability Letters*, **62**, 407-412.

Orozco-Castaneda, J., Nagar, D., and Gupta, A. (2012). Generalized Bivariate Beta Distributions Involving Appell's Hypergeometric Function of the Second Kind. *Computers and Mathematics with Applications*, **64**, 2507-2519.

Sant'Anna, A., and Ten Caten, C. (2012). Beta Control Charts for Monitoring Fraction Data. *Expert Systems with Applications*, **39**, 10236-10243.

Saleh, N.A., Mahmoud, M.A., Keefe, M.J., and Woodall, W.H.(2015). The Difficulty in Designing Shewhart X and $\bar{X}$ Control Charts with Estimated Parameters. *Journal of Quality Technology*, **47**, 127-138.

Sarabia, M., Prieto, F., and Jordá, V. (2014). Bivariate Beta-Generated Distributions with Applications to Well-being Data. *Statistical Distributions and Applications*, **1**, 1-26.

World Health Organization (2002). Hepatitis C. Retrieved from: `http://www.who.int/csr/disease/hepatitis/Hepc.pdf`.

Wright, S. (1937). The Distribution of Gene Frequencies in Populations. *Proceedings of the National Academy of Sciences*, **23**, 307-320.

Shohreh Enami
Department of Statistics,
Yazd University,
Yazd, Iran.
email: *enami1387@yahoo.com*

Hamzeh Torabi
Department of Statistics,
Yazd University,
Yazd, Iran.
email: *hamzeh.torabi@gmail.com*

Seyyed Taghi Akhavan Niaki
Department of Industrial Engineering,
Sharif University of Technology,
Tehran, Iran.
email: *niaki@sharif.edu*