Nonnormality Effects on the Economic-Statistical Design of $\bar{X}$ Charts with Weibull In-Control Time

Huifen Chen* and Yuyen Cheng

Department of Industrial Engineering, Chung-Yuan Christian University,

22 Pu Jen, Pu-Chung Li, Chung Li 320, TAIWAN

December 30, 2005

*Corresponding author. *Fax: +886-3-2654499; E-mail address: huifen@cycu.edu.tw*

Nonnormality Effects on the Economic-Statistical Design of $\bar{X}$ Charts with Weibull In-Control Time

December 30, 2005

ABSTRACT:

We consider the economic-statistical design of $\bar{X}$-control charts for nonnormal quality measurements. Specifically, we assume that the sample average $\bar{X}$ has a Johnson distribution. The Johnson distribution is general in that it can be made to fit all possible values of skewness and kurtosis. The cost model, proposed by McWilliams, is used to determine the optimal design parameters—the sample size, time between successive samples, and number of standard deviations away from the center line. This work is a generalization of Rahim's models; for example, it combines mainly three of Rahim's models: (i) economic design of $\bar{X}$ chart under non-normality (Rahim 1985), (ii) economic design of $\bar{X}$ chart under Weibull shock models (Banerjee and Rahim 1988), and (iii) economic-statistical design of $\bar{X}$ charts with non-Markovian in-control times (Al-Oraini and Rahim 2003).

Our sensitivity analysis shows that nonnormality has a significant effect on the design parameters and hence should not be ignored. Sensitivity to the Weibull shape and the process-mean shift are also considered. We also compare the economic-statistical and fully economic designs for nonnormal data.

Keywords: Economic-Statistical Design; Johnson Distribution; Nonnormality; $\bar{X}$ Control Chart.

1 Introduction

We consider the economic-statistical design of $\bar{X}$ control charts, assuming that the control chart point $\bar{X}$ (i.e., the sample average of n quality measurements) has a Johnson distribution (Johnson 1949) and that the time until the process is out of control has a Weibull distribution. In designing a control chart, three parameters—the sample size n, time h between successive samples, and the number k of standard deviations away from the center line—must be determined. In economic-statistical design, these parameters are chosen so that the expected cost per hour is minimized under constraints, e.g., minimum allowable values of the Type I error probability (probability that $\bar{X}$ falls outside control limits while the process is in control) and the Type II error probability (probability that $\bar{X}$ falls within control limits while the process out of control).

The $\bar{X}$ control chart is used to control the process mean at the desired level μ_0. The purpose is to maintain the required production quality. During the production process, the product quality characteristic X may vary because of assignable causes, resulting in a shift of the process mean to an inadequate level. Rahim (1985) lists some typical assignable causes, including defective raw materials, faulty setup, untrained operators, and the cumulative effects of heat, vibration, shock, etc. In this research, we assume that the process of interest has a single assignable cause. Studies of economic models assuming multiple assignable causes can be found in Duncan (1971) and Tagaras and Lee (1988). For surveys of the literature of $\bar{X}$ control charts, see Ho and Case (1994), Montgomery (1980), and Vance (1983).

The three design parameters n, h, and k are chosen so that the expected hourly cost (Equation 5) is minimized under constraints on the Type I and II error probabilities, or equivalently, on the average run length (ARL). The expected hourly cost equals the ratio of the expected cycle cost to the expected cycle time, where the production cycle is illustrated in Figure 1. The cost model used here is based on the model proposed by McWilliams (1989), which is an extension of the work of Lorenzen and Vance (1986). Woodall (1985, 1986) points out that a fully economic design ignores the statistical performance of control charts. On one hand, this may result in too many defectives. On the other hand, it may cause too many false alarms. If there are too many such false alarms, the control chart may be ignored in practice. Saniga (1989) first proposes the economic-statistical design. Al-Oraini and Rahim (2003) show that the economic-statistical design significantly improves the statistical performance at only a slight cost increase. Also, there is the statistical design, which chooses only the values of n and k under the desired Type I and II error probability constraints. However, n and k may not be chosen optimally to reduce the expected hourly cost.

Most literature assumes that the sample average $\bar{X}$ and the process in-control time follow normal and exponential distributions, respectively. (See, for example, Duncan 1956, and Lorenzen and Vance 1986.) A large sample size would result in an approximately normal distribution for $\bar{X}$. In practice, however, the sample size for the $\bar{X}$ control chart is usually small (e.g., Duncan, 1956, suggests $2 \leq n \leq 10$ for $\Delta \geq 2$) and therefore $\bar{X}$ is not necessarily normally distributed.

Some nonnormal literature exists. Lashkari and Rahim (1982) consider the CUSUM chart, and Nagendra and Rai (1971) and Rahim (1985) consider the economic design of $\bar{X}$ charts. All of this work models the probability density function of the quality measurement X by the first four terms of the Edgeworth series, where the four terms are functions of the first four moments. The normal distribution is a special case of Edgeworth series. Burr (1967), Yourstone and Zimmer (1992) and Chou et al. (2000) model the distribution of $\bar{X}$ as a Burr distribution (Burr 1942) using the moment method (see also Section 2). The advantage of the Burr distribution is that it has a closed-form cdf (cumulative distribution function), which simplifies computations of the Type I and II error probabilities (Equations 2 and 3). The disadvantage, however, is that because the Burr distribution is right skewed, unlike the Edgeworth series, it strictly limits $\bar{X}$ to a nonnormal distribution.

This research considers the economic-statistical design to minimize the cost of control charts while keeping reasonable Type I and II error probabilities. We assume that the sample average $\bar{X}$ has a Johnson distribution (see the Appendix). We choose the Johnson distribution because as shown in Figure 2, the Johnson family (including normal, lognormal, bounded, and unbounded types) covers the entire feasible part of the (β_1, β_2) plane, where β_1 stands for the squared skewness and β_2 the kurtosis. All lognormal (β_1, β_2) fall on the lognormal curve in the figure. The region above the lognormal curve consists of bounded Johnson distributions, denoted by S_B. The region below, which consists of unbounded Johnson distributions, we denote S_U. For each point (β_1, β_2), there is one corresponding Johnson distribution (Johnson et al. 1994, p. 36). All (β_1, β_2) for the Edgeworth series fall on a curve below the lognormal curve; all (β_1, β_2) for the Burr distribution form a region that is above the Weibull curve in the Pearson $(\sqrt{\beta}_1, \beta_2)$ plane (Johnson et al., 1994, pages 29 and 687). We choose the Johnson family for a wide range of the skewness and kurtosis.

In this work, we use the Johnson family to model the distribution of $\bar{X}$, rather than X, for better computational efficiency. If we were to model the distribution of X as Johnson, the distribution of $\bar{X}$ could possibly not be Johnson. One disadvantage of this approach is that it complicates the evaluation of the Type I and Type II error probabilities, which are n dimensional integrals and need to be computed numerically or estimated by Monte Carlo simulation. As a result, finding the optimal values of n, h, and k becomes more complex. Although not reported here, our simulation results show that the optimal values of n, h, and k do not vary much in the derivation of the Johnson model for X.

We also assume that the in-control time (i.e., the elapsed time before the assignable cause occurs) follows a Weibull distribution. Duncan (1956) and Lorenzen and Vance (1986) assume that the in-control time distribution is exponential. However, the memoryless property of the exponential distribution may not be applicable in practice. The Weibull distribution has the advantages that it offers many distribution shapes (including exponential), a nonconstant hazard rate function, and a closed-form cdf. Literature on the control chart applications of the Weibull distribution includes Banerjee and Rahim (1988), McWilliams (1989), and Zhang and Berardi (1997).

The rest of this paper is organized as follows. In Section 2, we discuss the cost function and constraints under the Johnson and Weibull distribution assumptions. In Section 3, we perform sensitivity analysis to study the effects of nonnormality, Weibull shape, and shift on the optimal parameters. Comparisons of the fully economic and economic-statistical designs are also discussed. In Section 4, we give our conclusions.

2 The Cost Model

The cost model used here is identical to the McWilliams (1989) model except that the distribution of the sample average $\bar{X}$ is assumed to have a Johnson distribution (see the Appendix). The production process is assumed to start in an in-control state, where the quality measurement X has mean $\mu = \mu_0$, standard deviation σ, skewness α_3 and kurtosis α_4. (Notice that $\alpha_3 = \mathrm{E}[(X-\mu)/\sigma]^3$ and $\alpha_4 = \mathrm{E}[(X-\mu)/\sigma]^4$). When the assignable cause occurs, the process mean μ shifts from μ_0 to $\mu_0 + \Delta\sigma$, $\Delta \in R$, but the standard deviation, skewness, and kurtosis remain unchanged. The value of Δ is assumed known so that the optimal values of n, h, and k can be computed. The in-control time V is assumed to have a Weibull distribution with shape parameter θ and scale parameter λ. The Weibull cdf is $F(v) = 1 - \exp\{-(\lambda v)^\theta\}$ and its mean is $\lambda^{-1}\Gamma(1+\theta^{-1})$. When $\theta = 1$, the Weibull distribution is an exponential distribution. In order to detect a shift in the process mean, a sample of n independent quality characteristic measurements $X_1, ..., X_n$ is taken at intervals of h hours. If the sample average $\bar{X}$ falls outside the control limits $\mu_0 \pm k\sigma/\sqrt{n}$ for a positive constant k, then an out-of-control signal is recorded in the control chart. In this case, the quality-control engineers try to determine whether there is an assignable cause. If such

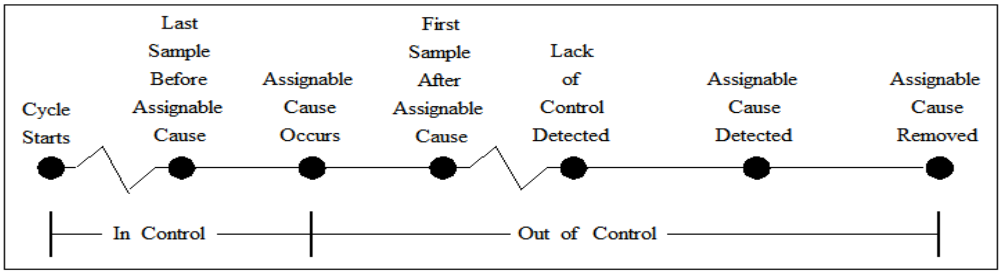

Figure 1: Plot of a production cycle (from Lorenzen and Vance, 1986)

an assignable cause is identified, appropriate action is taken to the production process and restore the in-control state.

For a given value of the sample size n, the distribution of $\bar{X}$ can be modeled as a Johnson distribution using the first four moments of X. Since $\bar{X}$ is the sample average of n independent observations of X, the first four moments of $\bar{X}$ are

$$\mathrm{E}(\bar{X}) = \mu, \qquad \mathrm{V}(\bar{X}) = \sigma^2/n, \qquad \alpha_{3,\bar{X}} = \alpha_3/\sqrt{n}, \qquad \alpha_{4,\bar{X}} = \frac{\alpha_4 - 3}{n} + 3 \ ,$$

where $\alpha_{3,\bar{X}}$ and $\alpha_{4,\bar{X}}$ denote the skewness and kurtosis of $\bar{X}$. Given n, we can fit a Johnson distribution to the four moments of $\bar{X}$ (see the Appendix). Notice that σ, α_3, and α_4 are assumed known to the quality engineer. In practice these values can be estimated, e.g., via the moment method, using many realizations of X. Though the mean μ is unknown, it is assumed that μ equals a known target mean μ_0 when the process is in control and equals $\mu_0 + \Delta\sigma$ otherwise. To compute the Type I and II error probabilities α and β in Equations (2) and (3), we can fit a Johnson distribution to $\bar{X}$ based on its conditional mean value (i.e., μ_0 for α and $\mu_0 + \Delta\sigma$ for β) and then compute the Johnson cumulative probability.

The $\bar{X}$ control chart is designed to detect whether the process is out of control. The design parameters n, h, and k are chosen to minimize the expected hourly cost, i.e., $\mathrm{E}(C)/\mathrm{E}(T)$, where C is the cycle cost, T is the cycle time, and $\mathrm{E}(\cdot)$ is the mean function. A quality cycle is defined as the time until the next in-control period. The sequence of quality cycles follows a renewal reward process because the in-control times in each cycle are identically and independently distributed. Therefore, the expected hourly cost $\mathrm{E}(C/T)$ equals the ratio of the expected cycle cost to the expected cycle time (Ross 1970, page 53), i.e., $\mathrm{E}(C)/\mathrm{E}(T)$. Computations of the expected cycle time and cycle cost are described as follows.

The expected cycle time consists of four parts as shown in Figure 1: (i) the expected time elapsed before the assignable cause occurs, (ii) the expected time between the occurrence of the assignable cause and the next out-of-control signal, (iii) the expected time T_1 required to identify the assignable cause, and (iv) the expected time T_2 to repair the process. As described in McWilliams (1989), the expected cycle time is

$$\begin{aligned} \mathrm{E}(T) &= \left[\lambda^{-1}\Gamma(1+\theta^{-1}) + (1-\delta_1)T_0 s/\mathrm{ARL}_0\right] \\ &\quad + \left[(s+\mathrm{ARL}_1)h - \lambda^{-1}\Gamma(1+\theta^{-1}) + nE\right] + T_1 + T_2 \\ &= (s+\mathrm{ARL}_1)h + (1-\delta_1)T_0 s/\mathrm{ARL}_0 + nE + T_1 + T_2, \end{aligned} \tag{1}$$

where δ_1 equals 1 if production continues during the assignable-cause search and 0 otherwise, T_0 is the expected assignable-cause search time for a false alarm, ARL_0 and ARL_1 are the average run lengths when the process is in control and out of control, E is the expected sampling time for one observation, and $s = \sum_{i=0}^{\infty} i\ \mathrm{P}\{ih \le V < (i+1)h\} = \sum_{i=1}^{\infty} e^{-(\lambda i h)^{\theta}}$ is the expected number of samples taken during the in-control state.

For independent observations, both ARL_0 and ARL_1 are related to Type I and II error probabilities as follows. The average run length ARL_0 is equal to $1/\alpha$, where α is the Type I error probability defined as

$$\begin{aligned} \alpha &= \mathrm{P}\{\ \bar{X} < \mu_0 - k\sigma/\sqrt{n} \quad \text{or} \quad \bar{X} > \mu_0 + k\sigma/\sqrt{n} \mid \mu = \mu_0\} \\ &= 1 + F_{\mu_0}(\mu_0 - k\sigma/\sqrt{n}) - F_{\mu_0}(\mu_0 + k\sigma/\sqrt{n}), \end{aligned} \tag{2}$$

and $F_\mu(\cdot)$ is the Johnson cdf with mean μ, standard deviation $\sigma/\sqrt{n}$, skewness $\alpha_{3,\bar{X}}$, and kurtosis $\alpha_{4,\bar{X}}$. (Computation of the Johnson cdf is described in the Appendix.) Similarly, $\mathrm{ARL}_1 = 1/(1-\beta)$, where

$$\begin{aligned} \beta &= \mathrm{P}\{\mu_0 - k\sigma/\sqrt{n} \le \bar{X} \le \mu_0 + k\sigma/\sqrt{n} \mid \mu = \mu_0 + \Delta\sigma\} \\ &= F_{\mu_0+\Delta\sigma}(\mu_0 + k\sigma/\sqrt{n}) - F_{\mu_0+\Delta\sigma}(\mu_0 - k\sigma/\sqrt{n}) \end{aligned} \tag{3}$$

is the Type II error probability. If the sample average $\bar{X}$ has a normal distribution (i.e., $\alpha_{3,\bar{X}} = 0$ and $\alpha_{4,\bar{X}} = 3$), then $\alpha = 2\Phi(-k)$ and $\beta = \Phi(k - \Delta\sqrt{n}) - \Phi(-k - \Delta\sqrt{n})$, where

Φ is the standard normal cdf.

There are costs associated with each of the four parts of the cycle time. The cost of the entire cycle includes the cost of nonconformities (while in control and out of control), false alarms, and repair. Let C_0 and C_1 ($> C_0$) denote the hourly costs due to nonconformities produced while the process is in and out of control, respectively. Furthermore, let δ_2 be 1 if production continues during the repair process, and 0, otherwise. The expected cost per cycle arising from nonconformities is then $C_0\lambda^{-1}\Gamma(1+\theta^{-1}) + C_1[(s+\mathrm{ARL}_1)h - \lambda^{-1}\Gamma(1+\theta^{-1}) + nE + \delta_1 T_1 + \delta_2 T_2]$. If we denote the cost per false alarm by c_f, then the total expected cost for false alarms is $c_f s/\mathrm{ARL}_0$. Let a be the fixed cost per sample and b be the variable cost per unit sampled. The cost per sample is then $a+bn$. Therefore, the expected cost for sampling and charting the result is $(a+bn)h^{-1}[(s+\mathrm{ARL}_1)h + nE + \delta_1 T_1 + \delta_2 T_2]$. Let W be the cost for locating and repairing the assignable cause. The expected cost per cycle is therefore

$$\begin{aligned}
\mathrm{E}(C) &= C_0\lambda^{-1}\Gamma(1+\theta^{-1}) + C_1\Big[(s+\mathrm{ARL}_1)h - \lambda^{-1}\Gamma(1+\theta^{-1}) + nE + \delta_1 T_1 + \delta_2 T_2\Big] \\
&\quad + c_f s/\mathrm{ARL}_0 + (a+bn)h^{-1}\Big[(s+\mathrm{ARL}_1)h + nE + \delta_1 T_1 + \delta_2 T_2\Big] + W. \qquad (4)
\end{aligned}$$

In economic-statistical design, the $\bar{X}$-chart design parameters n, h, and k are chosen to minimize the expected cost per hour for a quality cycle under constraints on the Type I and II error probabilities α and β. Let p_1 and p_2 denote the upper bounds for α and β, respectively. The design parameters n, h, and k are determined by solving the optimization problem:

$$\begin{aligned}
\min \quad & \mathrm{E}(C)/\mathrm{E}(T) \qquad (5) \\
\text{s.t.} \quad & \alpha < p_1, \\
& \beta < p_2, \\
n \in \{2,3,...\}, \quad & h > 0, \quad k > 0.
\end{aligned}$$

The objective function $\mathrm{E}(C)/\mathrm{E}(T)$, however, is not monotonic with respect to the design parameters n, h, and k. For the sensitivity analysis in Section 3, we use the grid search method to compute the optimal values n^*, h^*, and k^* of n, h, and k, respectively. The procedure for computing n^*, h^*, and k^* can be found in Nagendra and Rai (1971) and von

Collani (1986, 1988).

3 Sensitivity Analysis

Here we perform sensitivity analysis to investigate how nonnormality, Weibull shape, and shift affect the optimal values of the design parameters—that is, the sample size n^*, sampling interval h^*, and factor k^*. We also compare the economic-statistical design to the fully economic design.

Table 1 shows the effects of nonnormality on the optimal design values n^*, h^*, and k^*. Specifically, we study how n^*, h^*, and k^* change as the skewness α_3 and kurtosis α_4 of the quality measurement X vary. There are 55 design points, corresponding to two values of the bound parameters ($p_1 = 1$, .01 and $p_1 = p_2$), as shown in column 1 of the table, three values of the shift parameter ($\Delta = .5$, 1, 2), as shown in column 2, and 11 values of the population skewness, kurtosis combinations as shown in column 3. We test at three skewness values ($\alpha_3 = 0$, 2, 5) and three kurtosis values ($\alpha_4 = 6$, 36, 100). Excluding the invalid point (5, 6), these values yield 8 points. In addition, we include the point (0, 3), which corresponds to the normal distribution, and the points (2, 10.9) and (5, 68.3), which represent lognormal distribution shapes. Figure 2 illustrates these 11 combinations of skewness and kurtosis in the (β_1, β_2) plane. Furthermore, only positive skewness is considered because the control limits are symmetric. Notice that when $p_1 = p_2 = .01$, some distribution shapes may have no feasible solutions for $\Delta = .5$, and hence, only $\Delta = 1$ and 2 are considered.

The other parameters are set as follows. The Weibull shape parameter θ is set to .5 and the scale parameter λ is chosen so that the mean in-control time $E(V) = 100$.

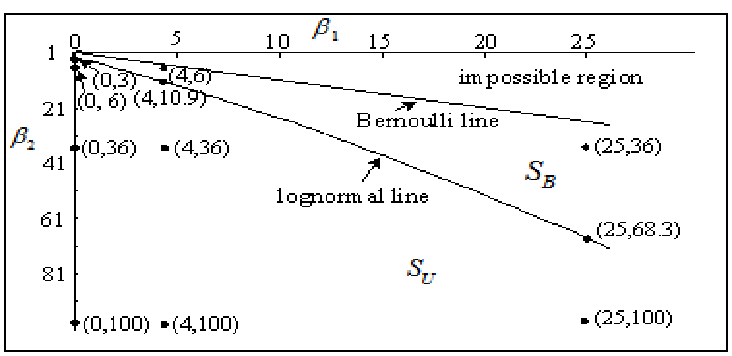

Figure 2: The (β_1, β_2) plane

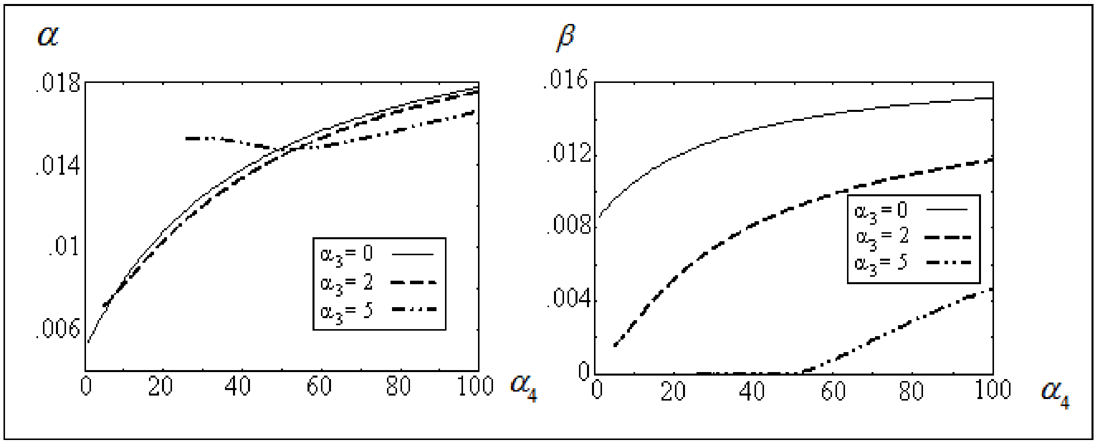

Figure 3: Plots of α and β as functions of the kurtosis α_4 for $n = 26$, $k = 2.74$, $\Delta = 1$, and $\alpha_3 = 0, 2, 5$ (the other cost parameters are the same as in Table 1)

The nonconformity costs C_0 and C_1 are made proportional to the defect rate, defined as $\mathrm{P}\{|X - \mu_0| > 3.5\sigma\}$. Here, we set $C_0 = 1000\ \mathrm{P}\{|X - \mu_0| > 3.5\sigma|\mu = \mu_0\}$ and hence, $C_1 = 1000\ \mathrm{P}\{|X - \mu_0| > 3.5\sigma|\mu = \mu_0 + \Delta\sigma\}$. The other cost parameters are given by $\delta_1 = \delta_2 = 1$, $E = 0$, $T_0 = 0$, $T_1 = 2$, $T_2 = 0$, $c_f = 50$, $a = .5$, $b = .1$, and $W = 25$. Arbitrarily we set $\mu_0 = 0$ and $\sigma = 1$ for simplicity because μ_0 and σ affect neither the mean hourly cost $\mathrm{E}(C)/\mathrm{E}(T)$ nor the Type I and II error probabilities α and β.

Table 1 lists n^*, h^*, k^*, the Johnson distribution type of $\bar{X}$ (given in the "type" column), C_0, C_1, α, β, and the expected hourly cost. We see that nonnormality affects the values of n^*, h^*, and k^*. The more (α_3, α_4) deviates from (0, 3), the more n^*, h^*, and k^* deviate from the normal values, which are, for examples, (25, 10.63, 2.67) for $p_1 = p_2 = .01$ and $\Delta = 1$, and (8, 1.95, 3.33) for $\Delta = 2$. For the sensitivity to the skewness, usually k^* increases and n^* and h^* decrease as the skewness α_3 goes up. Furthermore, the optimal parameters are sensitive to kurtosis α_4. When the skewness α_3 is small and p_1=p_2 = .01, n^* and k^* increase as α_4 increases. This is because in an economic-statistical design with low p_1 and p_2 values, the Type I and II error probabilities α and β strongly affect the optimal design parameters n^* and k^*. Figure 3 shows that for a small skewness, α and β increase as the kurtosis increases, and hence, n^* and k^* vary monotonically with respect to α_4. However, when the skewness α_3 is large (e.g., 5), α, and hence n^* and k^*, may not vary monotonically with respect to the kurtosis α_4. Furthermore, the expected hourly cost increases with α_4 when α_3 is small and decreases otherwise.

The Johnson distribution type of $\bar{X}$ depends on the population skewness and kurtosis.

Table 1: The sensitivity analysis of non-normality on the optimal values n^*, h^*, and k^* with $\Delta = 1$, 2, $\mathrm{E}(V) = 100$, $\theta = .5$, $p_1 = 1, .01$, and $p_1 = p_2$

p_1	Δ	(α_3, α_4)	n^*	h^*	k^*	type	C_0	C_1	α	β	E(C)/E(T)
		(0, 3)	35	39.49	2.16	S_N	.465	1.382	.0308	.2124	1.026
		(0, 6)	38	27.85	2.22	S_U	5.951	7.432	.0273	.1926	6.646
		(0, 36)	36	34.28	2.15	S_U	11.304	12.475	.0375	.182	11.932
		(0, 100)	32	39.36	2.04	S_U	11.995	12.995	.0485	.1787	12.578
		(2, 6)	45	5.53	2.26	S_B	0	30.247	.0231	.1346	3.04
	.5	(2, 10.9)	43	12.52	2.24	S_U	10.182	16.365	.0255	.1466	11.518
		(2, 36)	41	18.81	2.21	S_U	11.089	14.164	.032	.1487	12.058
		(2, 100)	36	24.55	2.12	S_U	11.922	13.899	.0422	.1588	12.714
		(5, 36)	42	15.04	2.09	S_B	18.392	23.368	.0356	.1047	19.59
		(5, 68.3)	41	14.43	2.12	S_U	13.288	18.511	.0364	.1179	14.519
		(5, 100)	40	16.21	2.13	S_U	11.88	16.093	.0385	.1238	12.995
		(0, 3)	16	8.25	2.77	S_N	.465	6.213	.0056	.1093	1.397
		(0, 6)	16	7.52	2.8	S_U	5.951	12.904	.0066	.1124	6.974
		(0, 36)	17	9.39	2.95	S_U	11.304	16.518	.0108	.1025	12.23
		(0, 100)	16	10.47	2.94	S_U	11.995	16.406	.0152	.1061	12.866
		(2, 6)	17	2.82	2.74	S_B	0	55.288	.0083	.0668	3.157
1	1	(2, 10.9)	17	5.16	2.77	S_U	10.182	26.781	.0088	.0727	11.762
		(2, 36)	17	7.06	2.9	S_U	11.089	20.279	.0114	.0887	12.288
		(2, 100)	17	8.7	2.98	S_U	11.922	18.387	.0142	.0902	12.953
		(5, 36)	19	6.87	2.98	S_B	18.392	30.111	.0125	.0213	19.731
		(5, 68.3)	17	6.2	2.84	S_U	13.288	26.493	.0149	.0569	14.725
		(5, 100)	17	6.75	2.9	S_U	11.88	22.87	.0149	.0682	13.197
		(0, 3)	6	1.75	3.23	S_N	.465	66.807	.0012	.0476	3.32
		(0, 6)	6	2	3.39	S_U	5.951	57.673	.0023	.0632	8.408
		(0, 36)	7	2.7	3.87	S_U	11.304	44.874	.0048	.062	13.258
		(0, 100)	7	3.06	3.99	S_U	11.995	39.535	.0065	.0646	13.76
		(2, 6)	6	1.46	3.43	S_B	0	107.933	.0029	.0233	4.047
	2	(2, 10.9)	7	1.93	3.67	S_U	10.182	75.767	.0029	.025	13.085
		(2, 36)	7	2.43	3.85	S_U	11.089	53.386	.0048	.0503	13.33
		(2, 100)	7	2.84	3.98	S_U	11.922	44.15	.0065	.0583	13.853
		(5, 36)	8	2.8	4.49	S_B	18.392	53.75	.0033	.0017	20.362
		(5, 68.3)	7	2.42	3.92	S_U	13.288	60.332	.0065	.028	15.696
		(5, 100)	7	2.54	3.95	S_U	11.88	53.662	.0067	.0403	14.128
		(0, 3)	25	10.63	2.67	S_N			.0076	.0099	1.453
		(0, 6)	26	9.73	2.74	S_U			.0072	.0098	7.036
		(0, 36)	29	12.04	2.88	S_U			.0099	.0099	12.293
		(0, 100)	33	13.86	3.12	S_U			.0098	.0099	12.943
		(2, 6)	22	3.23	2.71	S_B			.0082	.0096	3.226
	1	(2, 10.9)	23	5.99	2.75	S_U			.0085	.0098	11.811
		(2, 36)	27	8.63	2.94	S_U			.0089	.0098	12.348
		(2, 100)	31	11.01	3.12	S_U			.0099	.0099	13.03
		(5, 36)	22	7.07	3.16	S_B			.0088	.0092	19.742
		(5, 68.3)	25	6.98	3.13	S_U	same as for p_1=1		.0092	.0098	14.769
.01		(5, 100)	27	7.92	3.14	S_U			.0099	.0098	13.259
		(0, 3)	8	1.95	3.33	S_N			.0009	.0099	3.373
		(0, 6)	8	2.28	3.25	S_U			.0027	.0099	8.452
		(0, 36)	10	3.11	3.68	S_U			.005	.0099	13.314
		(0, 100)	11	3.59	3.87	S_U			.0063	.0099	13.828
		(2, 6)	7	1.49	3.64	S_B			.0015	.0093	4.049
	2	(2, 10.9)	7	2	3.44	S_U			.0042	.0097	13.096
		(2, 36)	9	2.69	3.71	S_U			.005	.0099	13.368
		(2, 100)	10	3.26	3.79	S_U			.0069	.0099	13.911
		(5, 36)	8	2.8	4.49	S_B			.0033	.0017	20.362
		(5, 68.3)	8	2.52	3.96	S_U			.006	.0098	15.706
		(5, 100)	9	2.75	4	S_U			.0059	.0098	14.158

It is well known that the sample average of the normal population is also distributed as normal. Intuitively if the population (α_3, α_4) falls in the bounded (or unbounded) area in the (β_1, β_2) plane, the Johnson type of $\bar{X}$ is S_B (or S_U). However, if the population (α_3, α_4) falls on the lognormal line, $\bar{X}$ is distributed as S_U, rather than lognormal.

The central limit theorem has little effect on the optimal values of n^*, h^*, and k^*. Even when n^* is large, the values of h^* and k^* for nonnormal distributions deviate from the normal values. This is because the nonnormality affects C_0 and C_1. The central limit theorem brings the nonnormal α and β close to the normal values. However, the nonnormal C_0 and C_1 differ significantly from the normal C_0 and C_1. Hence, the nonnormal n^*, h^*, and k^* differ from the normal values even when n^* is large.

Next we study the sensitivity of the Weibull shape parameter θ. Tables 2 and 3 list the optimal values n^*, h^*, and k^* for $\theta = .5$, .75, 1, 2, E(V) = 10, 100, $\Delta = 1$, 2, $p_1 = .05$, 1, and $p_2 = p_1$. Four population shapes are shown in column 1 of both tables: (0, 36), (5, 36), (2, 10.9) and (0, 3). Column 2 shows two values of p_1, $p_1 = .05$ and $p_1 = 1$. However, for $(\alpha_3, \alpha_4) = (5,\ 36)$, the results for $p_1 = .05$ are not displayed because they are identical to the results for $p_1 = 1$ (see Table 4). The values of the remaining cost parameters are as given in Table 1. Both tables show that as θ increases (while E(V) remains constant, but λ decreases), h^* increases slightly because Var(V) gets smaller. (However, the values of n^* and k^* are insensitive to changes in θ.) The sensitivity is more obvious for $\theta < 1$ than $\theta \geq 1$. A large decrease in E(V) results in a modest decrease in h^*. We expect a decrease in h^* because when the assignable cause occurs more frequently, the sampling frequency should increase as well. However, the sampling frequency may not increase with the same speed because otherwise, E(T) would be much shorter and would result in a larger hourly cost.

Table 4 compares the fully economic and economic-statistical designs. There are in total 60 design points, corresponding to the parameter values $(\alpha_3, \alpha_4) \in \{(0,3), (0,36),\ (5,\ 36),\ (2,10.9)\}$, $\Delta \in \{.5,1,2\}$, $p_1 \in \{.05,.1,.15,.3,1\}$ and $p_2 = p_1$. The other cost parameters are as in Table 1. When $p_1 = p_2 = 1$ (the top line in each row of the table), the economic-statistical design is called the fully economic design because there are no limits on the Type I and Type II error probabilities. Table 4 shows that the economic-statistical design is preferable because the Type I and II error probabilities can be kept as low as .05 at only a slight increase in the mean hourly cost. By contrast, the fully economic design has the lowest cost but the Type II error probability varies as high as .19. If we require that α and β

Table 2: The effects of the Weibull shape θ on n^*, h^*, and k^* with θ = .5, .75, 1, 2, E(V) = 10, 100, Δ = 1, 2, p_1 = .05, 1, $p_2 = p_1$, and $(\alpha_3, \alpha_4) \in \{(0, 36), (5, 36)\}$

				$\Delta = 1$				$\Delta = 2$			
(α_3, α_4)	p_1	E(V)	θ	n^*	h^*	k^*	E(C)/E(T)	n^*	h^*	k^*	E(C)/E(T)
(0, 36)	1	100	.5	17	9.39	2.95	12.23	7	2.7	3.87	13.258
			.75	17	10.01	2.91	12.207	7	2.83	3.84	13.227
			1	18	10.47	2.94	12.199	7	2.86	3.84	13.22
			2	18	10.66	2.93	12.195	7	2.88	3.83	13.217
		10	.5	14	5.32	2.68	15.295	7	1.03	3.83	21.696
			.75	14	5.63	2.62	15.244	7	1.1	3.79	21.573
			1	14	5.89	2.58	15.214	7	1.13	3.77	21.531
			2	14	6.57	2.49	15.178	7	1.15	3.76	21.51
	.05	100	.5	20	10.38	2.86	12.24	8	2.78	4.05	13.26
			.75	20	11	2.86	12.213	8	2.91	4.02	13.228
			1	20	11.28	2.86	12.204	8	2.95	4.01	13.22
			2	20	11.48	2.86	12.199	8	2.96	4.01	13.218
		10	.5	17	6.09	2.51	15.321	7	1.04	3.73	21.7
			.75	17	6.31	2.51	15.263	7	1.11	3.73	21.574
			1	17	6.53	2.51	15.229	7	1.14	3.73	21.531
			2	17	7.13	2.51	15.186	7	1.15	3.73	21.51
(5, 36)	1	100	.5	19	6.87	2.98	19.731	8	2.8	4.49	20.362
			.75	19	7.3	2.96	19.694	8	2.93	4.48	20.327
			1	19	7.49	2.95	19.682	8	2.97	4.48	20.319
			2	19	7.58	2.95	19.677	8	2.99	4.48	20.316
		10	.5	16	2.9	2.7	24.431	7	1.05	4.19	29.05
			.75	16	3.11	2.67	24.315	7	1.12	4.19	28.915
			1	16	3.26	2.65	24.26	7	1.15	4.18	28.869
			2	16	3.45	2.63	24.214	7	1.17	4.18	28.845

are both under .05, the mean hourly cost increases by less than 2%. However, on the whole, the economic-statistical design is superior. When the shift Δ is large, the fully economic design has a low α and β, and in several cases, the fully economic and economic-statistical designs are identical. The sample size n^* decreases as the bounds p_1 and p_2 become larger, but decreases slowly when p_1 and p_2 are greater than .15.

Table 4 also shows the effect of the shift Δ on the optimal values n^*, h^*, and k^*. When Δ increases, the sample size n^* decreases, the time h^* between samples decreases, and the factor k^* increases. If Δ is large, it is easier to detect that the process is out of control and hence, the sample size n need not be large. Usually when n decreases, h decreases as well. For a smaller sample, the sampling cost is less, and hence more samples are allowed. Furthermore, since the Type II error probability β is reduced by a large Δ, the factor k increases, reducing α while maintaining an allowable value of β.

Table 3: The effects of the Weibull shape θ on n^*, h^*, and k^* with θ = .5, .75, 1, 2, E(V) = 10, 100, $\Delta = 1, 2$, p_1 = .05, 1, $p_2 = p_1$, and $(\alpha_3, \alpha_4) \in \{(2, 10.9), (0, 3)\}$

				$\Delta = 1$				$\Delta = 2$			
(α_3, α_4)	p_1	E(V)	θ	n^*	h^*	k^*	E(C)/E(T)	n^*	h^*	k^*	E(C)/E(T)
(2, 10.9)	1	100	.5	17	5.16	2.77	11.762	7	1.93	3.67	13.085
			.75	18	5.56	2.8	11.725	7	2.02	3.65	13.046
			1	18	5.67	2.79	11.715	7	2.04	3.64	13.038
			2	18	5.74	2.78	11.711	7	2.05	3.64	13.036
		10	.5	16	2.08	2.68	17.494	7	.72	3.64	26.842
			.75	16	2.23	2.64	17.369	7	.76	3.61	26.682
			1	16	2.32	2.62	17.316	7	.78	3.6	26.633
			2	16	2.42	2.59	17.279	7	.79	3.6	26.612
	.05	100	.5	18	5.39	2.73	11.765	same as for p_1=1			
			.75	18	5.67	2.73	11.726				
			1	18	5.77	2.73	11.716				
			2	18	5.82	2.73	11.711				
		10	.5	17	2.18	2.62	17.506				
			.75	17	2.32	2.62	17.375				
			1	17	2.4	2.62	17.319				
			2	17	2.48	2.62	17.28				
(0,3)	1	100	.5	16	8.25	2.77	1.397	6	1.75	3.23	3.32
			.75	16	8.74	2.75	1.375	6	1.83	3.21	3.286
			1	16	8.98	2.73	1.368	6	1.84	3.21	3.28
			2	17	9.37	2.76	1.364	6	1.85	3.21	3.278
		10	.5	14	4.35	2.62	4.614	5	.62	3.07	17.097
			.75	14	4.6	2.58	4.561	6	.69	3.18	16.957
			1	14	4.82	2.55	4.533	6	.71	3.17	16.914
			2	15	5.39	2.53	4.502	6	.72	3.16	16.897
	.05	100	.5	19	9.27	2.71	1.408	same as for p_1=1			
			.75	19	9.81	2.71	1.382				
			1	19	10.05	2.71	1.374				
			2	19	10.2	2.71	1.37				
		10	.5	17	5.09	2.47	4.645	6	.66	3.2	17.099
			.75	17	5.31	2.47	4.583	6	.69	3.18	16.957
			1	17	5.51	2.47	4.549	6	.71	3.17	16.914
			2	17	5.94	2.47	4.51	6	.72	3.16	16.897

4 Conclusions

We consider the economic-statistical design of $\bar{X}$ control charts under the assumption that the sample average $\bar{X}$ has a Johnson distribution and the time V until the assignable cause occurs has a Weibull distribution. We choose the Johnson probability model because we can fit it to any desired first four moments. The control-chart design parameters n, h, and k are determined so that the expected hourly cost is minimized under constraints on the Type I and II error probabilities. We discuss the cost model, computations of the expected hourly cost, and Johnson probability modeling. Sensitivity analysis is performed to investigate the effects of nonnormality, Weibull shape, and shift on the optimal design parameters n^*, h^*, and k^*. Four results follow: (i) The values of n^*, h^*, and k^* for nonnormal distributions deviate quite a bit from the values for the normal distribution. Therefore, in cost control, nonnormality can not be ignored. (ii) The Weibull shape parameter θ has a small effect on n^*, h^*, and k^*. When θ decreases while $\mathrm{E}(V)$ is held constant, $\mathrm{Var}(V)$ increases, and hence h^* decreases (i.e., more sampling). (iii) When the shift Δ increases, it is more easily detected, and hence n^* and h^* decrease. (iv) The economic-statistical design is superior to the fully economic design because the Type I and II error probabilities can be reduced to acceptable levels at only a slight increase in cost.

Acknowledgments

This research is supported by the National Science Council in Taiwan under grant NSC 91-2213-E-033-044. We thank Carol Troy for proofreading this paper.

Appendix: The Johnson Probability Model

The Johnson distribution family, proposed by Johnson (1949), includes three transformations of the standard normal distribution. Let Y and Z denote the Johnson and standard normal random variables, respectively. The three transformations are:

$$\begin{array}{lll} S_L: & Z = \gamma + \delta \ln(\dfrac{Y-\xi}{\eta}), & \eta(\ Y-\xi) \geq 0, \\ S_B: & Z = \gamma + \delta \ln(\dfrac{Y-\xi}{\xi+\eta-Y}), & 0 \leq Y-\xi \leq \eta, \\ S_U: & Z = \gamma + \delta \sinh^{-1}(\dfrac{Y-\xi}{\eta}), & -\infty < Y < \infty. \end{array} \tag{6}$$

The constants ξ and η are location and scale parameters, respectively; γ and δ are the shape parameters. The second transformation, S_B, provides a bounded random variable Y; the third transformation, S_U, results in an unbounded Y. For lognormal distributions, S_L, the range is bounded below if $\eta > 0$ and bounded above if $\eta < 0$. Furthermore, the normal distribution, denoted as S_N, is one of the types of the Johnson distribution besides S_U, S_B, and S_L. We can use the numerical routines of Hill et al. (1976) to find the Johnson distribution having the four desired moments mean, standard deviation, skewness, and kurtosis. To compute the Johnson cumulative probability $F(y) = \mathrm{P}\{Y \leq y\}$, we can transform y to z using Equation (6) and then let $F(y) = \Phi(z)$, where Φ is the standard normal cdf. For example, if Y is a lognormal distribution, $F(y) = \Phi[\gamma + \delta \ln((y-\xi)/\eta)]$.

References

Al-Oraini, H.A., Rahim, M.A., 2003. Economic Statistical Design of $\bar{X}$ Control Charts for Systems with gamma(λ, 2) In-control Times. Journal of Applied Statistics 30, 397–409.

Banerjee, P.K., Rahim, M.A., 1988. Economic Design of $\bar{X}$-Control Charts Under Weibull Shock Models. Technometrics 30, 407–414.

Burr, I.W., 1942. Cumulative Frequency Functions. Annals of Mathematical Statistics 13, 215–232.

Burr, I.W., 1967. The Effects of Non-normality on Constants for $\bar{X}$ and R Charts. Industrial Quality Control 23, 563–568.

Chou, C., Chen, C., Liu, H., 2000. Economic-Statistical Design of $\bar{X}$ Charts for Non-normal Data by Considering Quality Loss. Journal of Applied Statistics 27, 939–951.

v. Collani, E., 1986. A Simple Procedure to Determine the Economic Design of an $\bar{X}$ Control Chart. Journal of Quality Technology 18, 145–151.

v. Collani, E., 1988. A Unified Approach to Optimal Process Control. Metrika 35, 145–159.

Duncan, A.J., 1956. The Economic Design of $\bar{X}$ Charts Used to Maintain Current Control of a Process. Journal of the American Statistical Association 51, 228–242.

Duncan, A.J., 1971. The Economic Design of $\bar{X}$ Charts Where There Is a Multiplicity of Assignable Causes. Journal of the American Statistical Association 66, 107–121.

Hill, I.D., Hill, R., Holder, R.L., 1976. Algorithm AS 99. Fitting Johnson Curves by Moments. Applied Statistics 25, 180–189.

Ho, C., Case, K.E., 1994. Economic Design of Control Charts: A Literature Review for 1981–1991. Journal of Quality Technology 26, 39–53.

Johnson, N.L., 1949. Systems of Frequency Curves Generated by Methods of Translation. Biometrika 36, 149–176.

Johnson, N.L., Kotz, S., Balakrishnan, N., 1994. Continuous Univariate Distributions—Volume 1, 2nd ed. NY: John Wiley & Sons.

Lashkari, R.S., Rahim, M.A., 1982. An economic design of cumulative sum charts to control non-normal process means. Computer and Industrial Engineering 6, 1–18.

Lorenzen, T.J., Vance, L.C., 1986. The Economic Design of Control Charts: A Unified Approach. Technometrics 28, 3–10.

McWilliams, T.P., 1989. Economic Control Chart Designs and the In-control Time Distribution: A Sensitivity Study. Journal of Quality Technology 21, 103–110.

Montgomery, D.C., 1980. The Economic Design of Control Charts: A Review and Literature Survey. Journal of Quality Technology 12, 75–87.

Nagendra, Y., Rai, G., 1971. Optimum sample size and sampling interval for controlling non-normal variables. Journal of the American Statistical Association 66, 637–646.

Rahim, M.A., 1985. Economic Model of $\bar{X}$-chart under Non-normality and Measurement Errors. Computer and Operations Research 12, 291–299.

Ross, S.M., 1970. Applied Probability Models with Optimization Applications. San Francisco: Holden-Day.

Saniga, E.M., 1989. Economic Statistical of Control-Chart Designs with an Application to $\bar{X}$ and R Charts. Technometrics 31, 313–320.

Tagaras, G., Lee., H.L., 1988. Economic Design of Control Charts with Different Control Limits for Different Assignable Causes. Management Science 34, 1347–1366.

Vance, L.C., 1983. A Bibliography of Statistical Quality Control Chart Techniques, 1970–1980. Journal of Quality Technology 15, 59–62.

Woodall, W.H., 1985. The Statistical Design of Quality Control Charts. The Statistician 34, 155–160.

Woodall, W.H., 1986. Weakness of the Economic Design of Control Charts. Technometrics 28, 408–409.

Yourstone, S.A., Zimmer, W.J., 1992. Non-normality and the Design of Control Charts for Averages. Decision Sciences 23, 1099–1113.

Zhang, G., Berardi, V., 1997. Economic Statistical Design of $\bar{X}$ Control Charts for Systems with Weibull In-control Times. Computers and Industrial Engineering 32, 575–586.

Table 4: Comparison of fully economic and economic-statistical designs for four population distribution shapes, $\Delta = .5,\ 1,\ 2$, $\mathrm{E}(V) = 100$, $\theta = .5$, $p_1 = .05, .1, .15, .3, 1$, and $p_2 = p_1$

Δ	(α_3, α_4)	p_1	n^*	h^*	k^*	α	β	$\mathrm{E}(C)/\mathrm{E}(T)$
		1	35	39.49	2.16	.0308	.2124	1.026
		⋮						
	(0, 3)	.3	35	39.49	2.16	.0308	.2124	1.026
		.15	40	43.75	2.12	.034	.1486	1.028
		.1	45	47.93	2.07	.0385	.0996	1.032
		.05	55	53.67	2.06	.0394	.0497	1.043
		1	36	34.28	2.15	.0375	.182	11.932
		⋮						
	(0, 36)	.3	36	34.28	2.15	.0375	.182	11.932
		.15	39	36.03	2.14	.0379	.1496	11.933
		.1	44	39.44	2.08	.0422	.0999	11.938
.5		.05	54	44.35	2.04	.0452	.0497	11.952
		1	42	15.04	2.09	.0356	.1047	19.59
		⋮						
	(5, 36)	.15	42	15.04	2.09	.0356	.1047	19.59
		.1	42	15.17	2.07	.0369	.0995	19.59
		.05	49	16.56	2.07	.0369	.0496	19.6
		1	43	12.52	2.24	.0255	.1466	11.518
		⋮						
	(2, 10.9)	.15	43	12.52	2.24	.0255	.1466	11.518
		.1	47	13.60	2.18	.0294	.0992	11.524
		.05	55	15.1	2.14	.0323	.0496	11.55
		1	16	8.25	2.77	.0056	.1093	1.397
		⋮						
	(0, 3)	.15	16	8.25	2.77	.0056	.1094	1.397
		.1	16	8.42	2.71	.0067	.0985	1.398
		.05	19	9.27	2.71	.0067	.0496	1.408
		1	17	9.39	2.95	.0108	.1025	12.23
		⋮						
	(0, 36)	.15	17	9.39	2.95	.0108	.1025	12.23
		.1	17	9.44	2.93	.0112	.0992	12.23
1		.05	20	10.38	2.86	.0118	.0498	12.24
		1	19	6.87	2.98	.0125	.0213	19.731
	(5, 36)	⋮						
		.05	19	6.87	2.98	.0125	.0213	19.731
		1	17	5.16	2.77	.0088	.0727	11.762
	(2, 10.9)	⋮						
		.1	17	5.16	2.77	.0088	.0727	11.762
		.05	18	5.39	2.73	.0094	.0488	11.765
		1	6	1.75	3.23	.0012	.0476	3.32
	(0, 3)	⋮						
		.05	6	1.75	3.23	.0012	.0476	3.32
		1	7	2.7	3.87	.0048	.062	13.258
	(0, 36)	⋮						
		.1	7	2.7	3.87	.0048	.062	13.258
2		.05	8	2.78	4.05	.0036	.047	13.26
		1	8	2.8	4.49	.0033	.0017	20.362
	(5, 36)	⋮						
		.05	8	2.8	4.49	.0033	.0017	20.362
		1	7	1.93	3.67	.0029	.025	13.085
	(2, 10.9)	⋮						
		.05	7	1.93	3.67	.0029	.025	13.085