國立交通大學

統計學研究所

碩 士 論 文

類別資料在兩個經驗貝氏模型中的模型選取技術

A Model Selection Technique between Two Empirical Bayes Models for Categorical Data

研 究 生：劉振滎

指導教授：陳志榮　教授

中 華 民 國 九 十 四 年 六 月

類別資料在兩個經驗貝氏模型中的模型選取技術

A Model Selection Technique between Two Empirical Bayes Models for Categorical Data

研 究 生：劉振滎　　Student：Chen-Ying Liu

指導教授：陳志榮 博士　　Advisor：Dr. Chih-Rung Chen

國 立 交 通 大 學
統 計 學 研 究 所
碩 士 論 文

A Thesis

Submitted to Department of Computer and Information Science

College of Electrical Engineering and Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Statistics

June 2005

Hsinchu, Taiwan, Republic of China

中華民國九十四年六月

類 別 資 料 在 兩 個 經 驗 貝 氏 模 型 中 的 模 型 選 取 技 術

學生：劉振榮　　　　　　　　　　　　指導教授：陳志榮　教授

國立交通大學統計學研究所碩士班

摘　　　要

在本篇論文中，首先我們提出一個對於製程中的類別資料在兩個經驗貝氏模型中的模型選取技術。然後我們簡介可用於製程中類別資料的兩個有用的經驗貝氏模型。最後舉一個例子並透過模擬實驗來展示所提出的方法之表現。

關鍵字: 經驗貝氏; 製程監控; 類別資料; beta-二項式; Dirichlet-多項式; 變換-常態-二項式; 變換-常態-多項式; 管制圖; 品質管制.

A Model Selection Technique between Two Empirical Bayes Models for Categorical Data

student：Chen-Ying Liu Advisors：Dr. Chih-Rung Chen

Institute of Statistics
National Chiao Tung University

ABSTRACT

In the paper, first of all, a model selection technique between two empirical Bayes models for categorical data in manufacturing is proposed. Next, two useful empirical Bayes models for categorical data in manufacturing are introduced. Finally, the performance of the proposed method is illustrated by an example through simulations.

KEY WORDS: Empirical Bayes; Process monitoring; Categorical data; Beta-binomial; Dirichlet-multinomial; Transformed-normal-binomial; Transformed-normal-multinomial; Control chart; Quality control.

誌　　謝

光陰荏苒，轉眼已至鳳凰花開和驪歌清唱時分，興奮愉悅與悵然不捨的情緒參雜交錯，心中感慨萬千。回首統研所兩年歲月的風風雨雨，品嚐到生命的酸甜苦辣，仍記憶猶新，歷歷在目。研究室的熱烈討論聲跟球場上揮灑的汗水，隨著時間緣起緣滅，終究要劃下一個美麗的句點。

首先，所上每位老師孜孜不倦的教導，如沐春風，除了富含高深的專業學問和不懈的研究精神外，其待人接物跟處事哲學，皆惠我良多且開闊了寬廣的視野，誠可作為學習的榜樣。特別要由衷感激指導教授 陳志榮老師，不但教學認真與治學嚴謹，更熱心地為學生解惑。感謝這一年多來不遺餘力的悉心教誨和指點，鉅細靡遺地傾囊相授，不厭其煩地耐心幫忙審稿校正，實在是令人敬佩不已。師恩浩瀚，永銘於心。同時，承蒙 洪慧念老師、 黃榮臣老師以及 許文郁老師能撥冗抽空來蒞臨指教，並提供許多寶貴的建議，使論文更臻完美。

其次，感謝助理們於行政事務跟電腦方面的支援，使其能順利畢業。對於已畢業的小慧與姿吟學姊所提供的協助及經驗分享，銘感五內，而學長姐、同學、朋友跟學弟妹們的陪伴與關心，互相切磋和砥礪，培養出彌足珍貴的友誼，在艱辛的生活中備感溫暖且不顯孤單。另外，還需感謝家人們的照顧與栽培，尤其是含辛茹苦的母親，多年來無怨無悔的付出，在我學習低潮、碰到瓶頸或挫敗氣餒時，適時給予了精神上的關懷跟鼓勵，化作堅實的後盾，令我提起勇氣堅強面對諸多挑戰，得以心無旁騖地焚膏繼晷、披荊斬棘，潛心於研究，以致完成論文。也因為有您，才能造就今日的我。

最後，感謝交大優良的求學環境、師長父母的教養恩澤及同窗好友的勉勵扶持，不過心中的謝意難以用隻字片語表達，只好奉上最誠摯的祝福，願各位順心如意，謹將此拙著獻給大家，一同分享這份成就和喜悅。此刻，即將揮別研究所的生涯，懷抱著美好的回憶繼續邁入人生的另一段旅程，雖說天下無不散的筵席，但冀望未來的路途能有緣與你們相聚。

劉 振 熒　謹誌于

國立交通大學統計學研究所

中華民國九十四年六月

目　　　　　錄

1. INTRODUCTION

In a manufacturing process, suppose that there are k possible types of defects in a product for some known positive integer k. For each tested product item, the result could be classified as one and only one of the following $k+1$ disjoint categories: {the first defect type, ..., the kth defect type, pass}. Such data are called either binary for $k = 1$ or polytomous for $k \geq 2$. In the paper, categorical data denote either binary data for $k = 1$ or polytomous data for $k \geq 2$. See, e.g., McCullagh and Nelder (1989, Chapters 4 and 5) or Agresti (2002) for the categorical data analysis.

In the Bayesian framework, it is assumed that the unknown random parameters have a known prior distribution. In practice, choosing an appropriate subjective or objective prior distribution is usually a non-trivial task for practitioners. Instead of a Bayesian approach, an empirical Bayes approach is commonly used in the literature. For an empirical Bayes inference, the marginal distribution of the observed data is utilized to estimate the unknown hyperparameters and then a Bayesian inference is made for the random parameters as if the estimated prior distribution were the prior distribution.

There are some researches for the empirical Bayes process monitoring techniques for categorical data in manufacturing. For example, Yousry *et al.* (1991) used the beta-binomial empirical Bayes model for binary data utilizing the method of moments for estimation of the hyperparameters. Recently, Shiau *et al.* (2005) used the Dirichlet-multinomial empirical Bayes model for polytomous data utilizing both the method of moments and the pseudolikelihood method for estimation of the hyperparameters. Chen *et al.* (2004) used the beta-binomial or Dirichlet-multinomial empirical Bayes model for categorical data utilizing the maximum likelihood (ML) method for estimation of the hyperparameters and the likelihood ratio (LR) method for monitoring the manufacturing process. Similarly, Chen *et al.* (2005) used the transformed-normal-binomial or transformed-normal-multinomial empirical Bayes model for categorical data utilizing the same methods as Chen *et al.* (2004).

To proceed the discussion, we first briefly introduce a Bayesian inference as follows: In the Bayesian framework, it is assumed that the unknown random parameter vector θ has a known prior probability density function (p.d.f.) or probability mass function (p.m.f.) $\pi(\theta)$ and that the response vector $\mathbf{y}$ has a known conditional p.d.f. or p.m.f. $f(\mathbf{y}|\theta)$ given θ. Then a Bayesian inference is based on the posterior p.d.f. or p.m.f., $p(\theta|\mathbf{y})$, of θ given $\mathbf{y}$, where

$$p(\theta|\mathbf{y}) \propto f(\mathbf{y}|\theta)\,\pi(\theta).$$

It is common practice to estimate θ by the posterior mean, $E(\theta|\mathbf{y})$, or the posterior mode, $\mathrm{mode}(\theta|\mathbf{y})$, of θ given $\mathbf{y}$, where

$$E(\theta|\mathbf{y}) = \frac{\int_\Theta \theta\, f(\mathbf{y}|\theta)\,\pi(\theta)\,d\theta}{\int_\Theta f(\mathbf{y}|\theta)\,\pi(\theta)\,d\theta} \text{ or } \frac{\sum_{\theta\in\Theta} \theta\, f(\mathbf{y}|\theta)\,\pi(\theta)}{\sum_{\theta\in\Theta} f(\mathbf{y}|\theta)\,\pi(\theta)}$$

and

$$\mathrm{mode}(\theta|\mathbf{y}) = \arg\sup_{\theta\in\Theta} p(\theta|\mathbf{y}) = \arg\sup_{\theta\in\Theta} f(\mathbf{y}|\theta)\,\pi(\theta)$$

with $P(\{\theta \in \Theta\}) = 1$. See, e.g., Gelman *et al.* (2004) or O'Hagan and Forster (2004) for the Bayesian data analysis.

Next, we briefly introduce an empirical Bayes inference as follows: In the empirical Bayes framework, it is assumed that the unknown random parameter vector θ has a prior p.d.f. or p.m.f. $\pi(\theta; \lambda)$ and that the response vector $\mathbf{y}$ has a known conditional p.d.f. or p.m.f. $f(\mathbf{y}|\theta)$ given θ, where λ is an unknown hyperparameter vector and $\pi(\cdot;\cdot)$ is a known function. Then an empirical Bayes inference is based on the estimated posterior p.d.f. or p.m.f., $p(\theta|\mathbf{y}; \lambda)|_{\lambda=\hat{\lambda}(\mathbf{y})}$, of θ given $\mathbf{y}$, where

$$p(\theta|\mathbf{y}; \lambda) \propto f(\mathbf{y}|\theta)\, \pi(\theta; \lambda)$$

and $\hat{\lambda}(\mathbf{y})$ is an estimator of λ. In practice, $\hat{\lambda}(\mathbf{y})$ is frequently chosen as the maximum likelihood estimator (MLE) or a method-of-moments estimator (MME) of λ. Similarly, it is common practice to estimate θ by the estimated posterior mean, $E(\theta|\mathbf{y}; \lambda)|_{\lambda=\hat{\lambda}(\mathbf{y})}$, or the estimated posterior mode, $\text{mode}(\theta|\mathbf{y}; \lambda)|_{\lambda=\hat{\lambda}(\mathbf{y})}$, of θ given $\mathbf{y}$, where

$$E(\theta|\mathbf{y}; \lambda) = \frac{\int_\Theta \theta\, f(\mathbf{y}|\theta)\, \pi(\theta; \lambda)\, d\theta}{\int_\Theta f(\mathbf{y}|\theta)\, \pi(\theta; \lambda)\, d\theta} \text{ or } \frac{\sum_{\theta\in\Theta} \theta\, f(\mathbf{y}|\theta)\, \pi(\theta; \lambda)}{\sum_{\theta\in\Theta} f(\mathbf{y}|\theta)\, \pi(\theta; \lambda)}$$

and

$$\text{mode}(\theta|\mathbf{y}; \lambda) = \arg \sup_{\theta\in\Theta} p(\theta|\mathbf{y}; \lambda) = \arg \sup_{\theta\in\Theta} f(\mathbf{y}|\theta)\, \pi(\theta; \lambda)$$

with $P(\{\theta \in \Theta\}; \lambda) = 1$. See, e.g., Carlin and Louis (2000) for the empirical Bayes data analysis.

The remaining part of the paper is organized as follows. A model selection technique between two empirical Bayes models for categorical data in manufacturing is proposed in Section 2. In Section 3, two useful empirical Bayes models for categorical data are introduced. The performance of the proposed method is illustrated by an example through simulations in Section 4. Some concluding remarks and future work are given in Section 5.

2. A MODEL SELECTION TECHNIQUE

Assume that each tested product item is classified as one and only one of the following $k+1$ categories: {the first defect type, ..., the kth defect type, pass}, where k is a known positive integer. Let t be any positive integer. Suppose that there are n_t tested product items manufactured at time t, where n_t is a known positive integer. For $i \in \{1, \ldots, k\}$, let θ_{it} denote the probability that a product item manufactured at time t is of the ith defect type. Then $1 - \sum_{i=1}^{k} \theta_{it}$ $(\equiv \theta_{k+1,t})$ is the probability that a product item manufactured at time t passes the test. Assume that $\theta_{it} > 0$ for $i \in \{1, \ldots, k+1\}$. For $i \in \{1, \ldots, k\}$, let y_{it} denote the number of the tested product items which are of the ith defect type among the n_t tested product items manufactured at time t. Then $n_t - \sum_{i=1}^{k} y_{it}$ $(\equiv y_{k+1,t})$ is the number of the tested product items which pass the test among the n_t tested product items manufactured at time t. Set $\theta_t \equiv (\theta_{1t}, \ldots, \theta_{kt})^T$, $\mathbf{y}_t \equiv (y_{1t}, \ldots, y_{kt})^T$, $\Theta \equiv \{\theta_t : \theta_{1t}, \ldots, \theta_{kt} > 0 \text{ and } \sum_{i=1}^{k} \theta_{it} < 1\}$, and $\mathcal{Y}_{n_t} \equiv \{\mathbf{y}_t : y_{1t}, \ldots, y_{kt} \in \{0, 1, \ldots, n_t\} \text{ and } \sum_{i=1}^{k} y_{it} \leq n_t\}$.

Assume that $\mathbf{y}_t$ has the conditional binomial$(n_t; \theta_t)$ or multinomial$(n_t; \theta_t)$ distribution given θ_t. Let F_{θ_t} and $F_{\mathbf{y}_t|\theta_t}$ denote, respectively, the prior cumulative distribution function (c.d.f.) of θ_t and the conditional c.d.f. of $\mathbf{y}_t$ given θ_t. Then $\mathbf{y}_t$ has the conditional p.m.f.

$$f(\mathbf{y}_t|\theta_t) = 1_{\mathcal{Y}_{n_t}}(\mathbf{y}_t) \cdot \frac{n_t!}{\prod_{i=1}^{k+1} y_{it}!} \cdot \prod_{i=1}^{k+1} \theta_{it}^{y_{it}} \tag{1}$$

given θ_t, where $1_{\mathcal{Y}_{n_t}}(\mathbf{y}_t) = 1$ for $\mathbf{y}_t \in \mathcal{Y}_{n_t}$ and 0 otherwise. Thus, $\mathbf{y}_t$ has the marginal p.m.f.

$$f(\mathbf{y}_t; F_{\theta_t}) = 1_{\mathcal{Y}_{n_t}}(\mathbf{y}_t) \cdot \frac{n_t!}{\prod_{i=1}^{k+1} y_{it}!} \cdot \int_{\Theta} \prod_{i=1}^{k+1} \theta_{it}^{y_{it}} \, dF_{\theta_t}(\theta_t). \tag{2}$$

Throughout the paper, we say that the manufacturing process is in control at time t when $F_{\theta_t} = F$, where F is a c.d.f. on Θ with some unknown p.d.f. $\pi(\cdot)$. For any positive integer m, set $\mathcal{R}^m \equiv (-\infty, \infty)^m$, let $0_{m\times 1}$ denote the $m \times 1$ vector $(0, \ldots, 0)^T$, and let $1_{m\times 1}$ denote the $m \times 1$ vector $(1, \ldots, 1)^T$.

For $u \in \{1, 2\}$, let model u denote the parametric family $\{F_{u,\lambda_u} : \lambda_u \in \Lambda_u\}$, where λ_u is a $q_u \times 1$ hyperparameter vector for some known positive integer q_u, each F_{u,λ_u} is a c.d.f. on Θ with known p.d.f. $\pi_u(\cdot; \lambda_u)$, and Λ_u is a known open subset of $\mathcal{R}^{q_u}$. Without loss of generality, assume that $q_1 \le q_2$. Assume that $\partial^2 \pi_u(\theta_t; \lambda_u)/\partial\lambda_u \partial\lambda_u^T$ exists for $\theta_t \in \Theta$, $\lambda_u \in \Lambda_u$, and $u \in \{1, 2\}$. For $\lambda_u \in \Lambda_u$ and $u \in \{1, 2\}$, let $F_{\mathbf{y}_t;u,\lambda_u}$ denote the marginal c.d.f. of $\mathbf{y}_t$ when $F_{\theta_t} = F_{u,\lambda_u}$. Then $\mathbf{y}_t$ has the marginal p.m.f.

$$f(\mathbf{y}_t; F_{u,\lambda_u}) = 1_{\mathcal{Y}_{n_t}}(\mathbf{y}_t) \cdot \frac{n_t!}{\prod_{i=1}^{k+1} y_{it}!} \cdot \int_{\Theta} \prod_{i=1}^{k+1} \theta_{it}^{y_{it}} \, dF_{u,\lambda_u}(\theta_t) \tag{3}$$

when $F_{\theta_t} = F_{u,\lambda_u}$ for some $\lambda_u \in \Lambda_u$ and $u \in \{1, 2\}$.

For $\lambda_u \in \Lambda_u$ and $u \in \{1, 2\}$, the Kullback-Leibler distance between F and F_{u,λ_u} is

$$d\left(F, F_{u,\lambda_u}\right) \equiv \int_{\Theta} \log\left[\frac{\pi(\theta_t)}{\pi_u(\theta_t; \lambda_u)}\right] dF(\theta_t) \quad (\equiv d_u(\lambda_u)). \tag{4}$$

By the Jensen inequality,

$$\begin{aligned} d_u(\lambda_u) &= \int_{\Theta} -\log\left[\frac{\pi_u(\theta_t; \lambda_u)}{\pi(\theta_t)}\right] dF(\theta_t) \;\ge\; -\log\left[\int_{\Theta} \frac{\pi_u(\theta_t; \lambda_u)}{\pi(\theta_t)} \cdot \pi(\theta_t)\, d\theta_t\right] \\ &= -\log\left[\int_{\{\theta_t : \pi(\theta_t) > 0\}} \pi_u(\theta_t; \lambda_u)\, d\theta_t\right] \;\ge\; -\log\left[\int_{\Theta} \pi_u(\theta_t; \lambda_u)\, d\theta_t\right] \;=\; 0, \end{aligned}$$

where $d_u(\lambda_u) = 0$ if and only if $F_{u,\lambda_u} = F$.

For $\lambda_u \in \Lambda_u$ and $u \in \{1,2\}$, assume that all of the following conditions hold: $d_u(\lambda_u) < \infty$, $\partial^2 d_u(\lambda_u)/\partial\lambda_u\partial\lambda_u^T$ exists,

$$\frac{\partial d_u(\lambda_u)}{\partial\lambda_u} = \int_\Theta \frac{\partial}{\partial\lambda_u}\left\{\log\left[\frac{\pi(\theta_t)}{\pi_u(\theta_t;\lambda_u)}\right]\right\} dF(\theta_t),$$

and

$$\frac{\partial^2 d_u(\lambda_u)}{\partial\lambda_u\partial\lambda_u^T} = \int_\Theta \frac{\partial^2}{\partial\lambda_u\partial\lambda_u^T}\left\{\log\left[\frac{\pi(\theta_t)}{\pi_u(\theta_t;\lambda_u)}\right]\right\} dF(\theta_t).$$

For $u \in \{1,2\}$, assume that there exists a unique $\lambda_u^0 \in \Lambda_u$ such that

$$\lambda_u^0 = \arg\inf_{\lambda_u\in\Lambda_u} d_u(\lambda_u). \tag{5}$$

Suppose that we are interested in choosing either model 1 or model 2 as an approximate model for monitoring the manufacturing process. For this purpose, we would like to consider the hypothesis testing problem with the null hypothesis $H_0 : d_1(\lambda_1^0) \le d_2(\lambda_2^0)$ versus the alternative $H_1 : d_1(\lambda_1^0) > d_2(\lambda_2^0)$. Then we choose model 2 if and only if we reject H_0 in favor of H_1.

Note that $\partial d_u(\lambda_u)/\partial\lambda_u|_{\lambda_u=\lambda_u^0} = 0_{q_u\times 1}$ for $u \in \{1,2\}$. For $\lambda_u \in \Lambda_u$ and $u \in \{1,2\}$, set $g_u(\lambda_u) \equiv -\partial d_u(\lambda_u)/\partial\lambda_u$ and $h_u(\lambda_u) \equiv -\partial g_u(\lambda_u)/\partial\lambda_u^T$. Then, for $\lambda_u \in \Lambda_u$ and $u \in \{1,2\}$,

$$\begin{aligned} g_u(\lambda_u) &= \int_\Theta \frac{\partial\pi_u(\theta_t;\lambda_u)/\partial\lambda_u}{\pi_u(\theta_t;\lambda_u)}\, dF(\theta_t) \\ &\equiv \int_\Theta S_u(\lambda_u;\theta_t)\, dF(\theta_t) \equiv E\left(S_u(\lambda_u;\theta_t);F\right) \end{aligned} \tag{6}$$

and

$$\begin{aligned} h_u(\lambda_u) &= \int_\Theta -\frac{\partial S_u(\lambda_u;\theta_t)}{\partial \lambda_u^T}\, dF(\theta_t) \\ &\equiv \int_\Theta J_u(\lambda_u;\theta_t)\, dF(\theta_t) \;\equiv\; E\left(J_u(\lambda_u;\theta_t);F\right). \end{aligned} \tag{7}$$

When both $g_u(\lambda_u)$ and $h_u(\lambda_u)$ have closed-form formulas for $\lambda_u \in \Lambda_u$ and $u \in \{1,2\}$ in a simulation study, we may utilize the following Newton-Raphson method to obtain λ_u^0: First choose a *good* initial value $\lambda_u^{0(0)}$ for λ_u^0 and then iterate the following equations

$$\lambda_u^{0(v+1)} = \lambda_u^{0(v)} + \left[h_u\left(\lambda_u^{0(v)}\right)\right]^{-1} g_u\left(\lambda_u^{0(v)}\right) \tag{8}$$

for $v = 0, 1, \ldots$ until $\lambda_u^{0(v)}$ converges to λ_u^0. When $g_u(\lambda_u)$ or $h_u(\lambda_u)$ does not have a closed-form formula for some $\lambda_u \in \Lambda_u$ and $u \in \{1,2\}$ in a simulation study, we may first simulate an i.i.d. sample $\{\theta_t^{(1)}, \ldots, \theta_t^{(R)}\}$ of size R, e.g., $R = 100\,000$, from the c.d.f. F and then numerically evaluate $g_u(\lambda_u)$ and $h_u(\lambda_u)$ by $R^{-1}\cdot\sum_{r=1}^R S_u(\lambda_u;\theta_t^{(r)})$ and $R^{-1}\cdot\sum_{r=1}^R J_u(\lambda_u;\theta_t^{(r)})$, respectively.

Suppose that there is an available *in-control* historical data set $\{\mathbf{y}_1, \ldots, \mathbf{y}_T\}$ in the manufacturing process for some known positive integer T, where $(\theta_1^T, \mathbf{y}_1^T)^T, \ldots, (\theta_T^T, \mathbf{y}_T^T)^T$ are independent $2k \times 1$ random vectors. Set $\theta \equiv (\theta_1^T, \ldots, \theta_T^T)^T$, $\mathbf{y} \equiv (\mathbf{y}_1^T, \ldots, \mathbf{y}_T^T)^T$, and $\mathcal{Y} \equiv \mathcal{Y}_{n_1} \times \cdots \times \mathcal{Y}_{n_T}$.

Given $\mathbf{y}$ and under model u for $u \in \{1,2\}$, the log-likelihood function for λ_u is

$$\ell_u(\lambda_u;\mathbf{y}) \;\equiv\; \log\left[\prod_{t=1}^T f(\mathbf{y}_t;F_{u,\lambda_u})\right] \;=\; \sum_{t=1}^T \log[f(\mathbf{y}_t;F_{u,\lambda_u})] \;\equiv\; \sum_{t=1}^T \ell_u(\lambda_u;\mathbf{y}_t), \tag{9}$$

the score function for λ_u is

$$\begin{aligned} S_u(\lambda_u;\mathbf{y}) &\equiv \frac{\partial \ell_u(\lambda_u;\mathbf{y})}{\partial \lambda_u} = \sum_{t=1}^{T} \frac{\partial \ell_u(\lambda_u;\mathbf{y}_t)}{\partial \lambda_u} \\ &= \sum_{t=1}^{T} \frac{\partial f(\mathbf{y}_t;F_{u,\lambda_u})/\partial \lambda_u}{f(\mathbf{y}_t;F_{u,\lambda_u})} \equiv \sum_{t=1}^{T} S_u(\lambda_u;\mathbf{y}_t), \end{aligned} \tag{10}$$

and the observed (Fisher) information for λ_u is

$$J_u(\lambda_u;\mathbf{y}) \equiv -\frac{\partial S_u(\lambda_u;\mathbf{y})}{\partial \lambda_u^T} = \sum_{t=1}^{T} -\frac{\partial S_u(\lambda_u;\mathbf{y}_t)}{\partial \lambda_u^T} \equiv \sum_{t=1}^{T} J_u(\lambda_u;\mathbf{y}_t). \tag{11}$$

Given $\mathbf{y}$ and under model u for $u \in \{1,2\}$, the MLE $\hat{\lambda}_u(\mathbf{y})$ $(\equiv \hat{\lambda}_u)$ of λ_u solves the score equation $S_u(\lambda_u) = 0_{q_u \times 1}$ for λ_u. That is, $S_u(\hat{\lambda}_u) = 0_{q_u\times 1}$ for $u \in \{1,2\}$. We may utilize the following Newton-Raphson method to obtain $\hat{\lambda}_u$ for $u \in \{1,2\}$: First choose a *good* initial value $\hat{\lambda}_u^{(0)}$ for $\hat{\lambda}_u$ and then iterate the following equations

$$\hat{\lambda}_u^{(v+1)} = \hat{\lambda}_u^{(v)} + \left[J_u\left(\hat{\lambda}_u^{(v)};\mathbf{y}\right)\right]^{-1} S_u\left(\hat{\lambda}_u^{(v)};\mathbf{y}\right) \tag{12}$$

for $v = 0, 1, \dots$ until $\hat{\lambda}_u^{(v)}$ converges to $\hat{\lambda}_u$.

Let $F_{\mathbf{y}}$ denote the c.d.f. of $\mathbf{y}$ with p.m.f. $f(\mathbf{y};F)$. For $\lambda_u \in \Lambda_u$ and $u \in \{1,2\}$, let $F_{\mathbf{y};u,\lambda_u}$ denote the c.d.f. of $\mathbf{y}$ with p.m.f. $f(\mathbf{y};F_{u,\lambda_u})$ when $F_{\theta_1} = \ldots = F_{\theta_T} = F_{u,\lambda_u}$. For $\lambda_u \in \Lambda_u$ and $u \in \{1,2\}$, the Kullback-Leibler distance between $F_{\mathbf{y}}$ and $F_{\mathbf{y};u,\lambda_u}$ is

$$\begin{aligned} d\left(F_{\mathbf{y}}, F_{\mathbf{y};u,\lambda_u}\right) &\equiv \sum_{\mathbf{y}\in\mathcal{Y}} \log\left[\frac{f(\mathbf{y};F)}{f(\mathbf{y};F_{u,\lambda_u})}\right] f(\mathbf{y};F) \\ &= \sum_{t=1}^{T}\left\{\sum_{\mathbf{y}_t\in\mathcal{Y}_{n_t}} \log\left[\frac{f(\mathbf{y}_t;F)}{f(\mathbf{y}_t;F_{u,\lambda_u})}\right] f(\mathbf{y}_t;F)\right\} \\ &\equiv \sum_{t=1}^{T} d\left(F_{\mathbf{y}_t}, F_{\mathbf{y}_t;u,\lambda_u}\right) \quad (\equiv d_u^{n_1,\ldots,n_T}(\lambda_u)). \end{aligned} \tag{13}$$

For $u \in \{1, 2\}$, assume that there exists a unique $\lambda_u^{n_1,\ldots,n_T} \in \Lambda_u$ such that

$$\lambda_u^{n_1,\ldots,n_T} = \arg \inf_{\lambda_u \in \Lambda_u} d_u^{n_1,\ldots,n_T}(\lambda_u). \tag{14}$$

When $n_1 = \ldots = n_T$, set $d_u^{n_1}(\lambda_u) \equiv d_u^{n_1,\ldots,n_T}(\lambda_u)$ and $\lambda_u^{n_1} \equiv \lambda_u^{n_1,\ldots,n_T}$ for $\lambda_u \in \Lambda_u$ and $u \in \{1, 2\}$. Then $d_u^{n_1}(\lambda_u) = T \cdot d\left(F_{\mathbf{y}_1}, F_{\mathbf{y}_1;u,\lambda_u}\right)$ for $\lambda_u \in \Lambda_u$ and $u \in \{1, 2\}$.

Note that $\partial d_u^{n_1,\ldots,n_T}(\lambda_u)/\partial\lambda_u|_{\lambda_u=\lambda_u^{n_1,\ldots,n_T}} = 0_{q_u\times 1}$ for $u \in \{1, 2\}$. For $\lambda_u \in \Lambda_u$ and $u \in \{1, 2\}$, set $g_u^{n_1,\ldots,n_T}(\lambda_u) \equiv -T^{-1}\cdot\partial d_u^{n_1,\ldots,n_T}(\lambda_u)/\partial\lambda_u$ and $h_u^{n_1,\ldots,n_T}(\lambda_u) \equiv -T^{-1}\cdot\partial g_u^{n_1,\ldots,n_T}(\lambda_u)$ $/\partial\lambda_u^T$. Then, for $\lambda_u \in \Lambda_u$ and $u \in \{1, 2\}$,

$$g_u^{n_1,\ldots,n_T}(\lambda_u) = \frac{1}{T}\cdot\sum_{t=1}^{T}\left[\sum_{\mathbf{y}_t\in\mathcal{Y}_{n_t}} S_u(\lambda_u; \mathbf{y}_t)\, f(\mathbf{y}_t; F)\right] = E\left(\frac{1}{T}\cdot\sum_{t=1}^{T} S_u(\lambda_u; \mathbf{y}_t); F\right) \tag{15}$$

and

$$h_u^{n_1,\ldots,n_T}(\lambda_u) = \frac{1}{T}\cdot\sum_{t=1}^{T}\left[\sum_{\mathbf{y}_t\in\mathcal{Y}_{n_t}} J_u(\lambda_u; \mathbf{y}_t)\, f(\mathbf{y}_t; F)\right] = E\left(\frac{1}{T}\cdot\sum_{t=1}^{T} J_u(\lambda_u; \mathbf{y}_t); F\right). \tag{16}$$

When $n_1 = \ldots = n_T$, set $g_u^{n_1}(\lambda_u) \equiv g_u^{n_1,\ldots,n_T}(\lambda_u)$ and $h_u^{n_1}(\lambda_u) \equiv h_u^{n_1,\ldots,n_T}(\lambda_u)$ for $\lambda_u \in \Lambda_u$ and $u \in \{1, 2\}$. Then

$$g_u^{n_1}(\lambda_u) = \sum_{\mathbf{y}_1\in\mathcal{Y}_{n_1}} S_u(\lambda_u; \mathbf{y}_1)\, f(\mathbf{y}_1; F) = E\left(S_u(\lambda_u; \mathbf{y}_1); F\right) \tag{17}$$

and

$$h_u^{n_1}(\lambda_u) = \sum_{\mathbf{y}_1\in\mathcal{Y}_{n_1}} J_u(\lambda_u; \mathbf{y}_1)\, f(\mathbf{y}_1; F) = E\left(J_u(\lambda_u; \mathbf{y}_1); F\right) \tag{18}$$

for $\lambda_u \in \Lambda_u$ and $u \in \{1, 2\}$. When $n_1 = \ldots = n_T$, it can be shown that $\hat{\lambda}_u = \lambda_u^{n_1} + O_p(1/\sqrt{T})$ as $T \to \infty$ for $\lambda_u \in \Lambda_u$ and $u \in \{1, 2\}$. Thus, it is very likely that $\lambda_u^{n_1,\ldots,n_T} \approx \lambda_u^0$ for large $\min\{n_1, \ldots, n_T\}$ and $\hat{\lambda}_u \approx \lambda_u^0$ for large T and $\min\{n_1, \ldots, n_T\}$.

When both $g_u^{n_1,\ldots,n_T}(\lambda_u)$ and $h_u^{n_1,\ldots,n_T}(\lambda_u)$ have closed-form formulas for $\lambda_u \in \Lambda_u$ and $u \in \{1,2\}$ in a simulation study, we may utilize the following Newton-Raphson method to obtain $\lambda_u^{n_1,\ldots,n_T}$: First choose a *good* initial value $\lambda_u^{n_1,\ldots,n_T(0)}$ for $\lambda_u^{n_1,\ldots,n_T}$ and then iterate the following equations

$$\lambda_u^{n_1,\ldots,n_T(v+1)} = \lambda_u^{n_1,\ldots,n_T(v)} + \left[h_u^{n_1,\ldots,n_T}\left(\lambda_u^{n_1,\ldots,n_T(v)}\right)\right]^{-1} g_u^{n_1,\ldots,n_T}\left(\lambda_u^{n_1,\ldots,n_T(v)}\right) \tag{19}$$

for $v = 0, 1, \ldots$ until $\lambda_u^{n_1,\ldots,n_T(v)}$ converges to $\lambda_u^{n_1,\ldots,n_T}$. When $g_u^{n_1,\ldots,n_T}(\lambda_u)$ or $h_u^{n_1,\ldots,n_T}(\lambda_u)$ does not have a closed-form formula for some $\lambda_u \in \Lambda_u$ and $u \in \{1,2\}$ in a simulation study, we may first simulate an i.i.d. sample $\{\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(R)}\}$ of size R, e.g., $R = 100\,000$, from the c.d.f. $F_{\mathbf{y}}$ and then numerically evaluate $g_u^{n_1,\ldots,n_T}(\lambda_u)$ and $h_u^{n_1,\ldots,n_T}(\lambda_u)$ by $R^{-1} \cdot \sum_{r=1}^{R}[T^{-1} \cdot \sum_{t=1}^{T} S_u(\lambda_u; \mathbf{y}_t^{(r)})]$ and $R^{-1} \cdot \sum_{r=1}^{R}[T^{-1} \cdot \sum_{t=1}^{T} J_u(\lambda_u; \mathbf{y}_t^{(r)})]$, respectively.

When $n_1 = \ldots = n_T$ and both $g_u^{n_1}(\lambda_u)$ and $h_u^{n_1}(\lambda_u)$ have closed-form formulas for $\lambda_u \in \Lambda_u$ and $u \in \{1,2\}$ in a simulation study, we may utilize the following Newton-Raphson method to obtain $\lambda_u^{n_1}$: First choose a *good* initial value $\lambda_u^{n_1(0)}$ for $\lambda_u^{n_1}$ and then iterate the following equations

$$\lambda_u^{n_1(v+1)} = \lambda_u^{n_1(v)} + \left[h_u^{n_1}\left(\lambda_u^{n_1(v)}\right)\right]^{-1} g_u^{n_1}\left(\lambda_u^{n_1(v)}\right) \tag{20}$$

for $v = 0, 1, \ldots$ until $\lambda_u^{n_1(v)}$ converges to $\lambda_u^{n_1}$. When $g_u^{n_1}(\lambda_u)$ or $h_u^{n_1}(\lambda_u)$ does not have a closed-form formula for some $\lambda_u \in \Lambda_u$ and $u \in \{1,2\}$ in a simulation study, we may simply simulate an i.i.d. sample $\{\mathbf{y}_1^{(1)}, \ldots, \mathbf{y}_1^{(R)}\}$ of size R, e.g., $R = 100\,000$, from the c.d.f. $F_{\mathbf{y}_1}$ and then numerically evaluate $g_u^{n_1}(\lambda_u)$ and $h_u^{n_1}(\lambda_u)$ by $R^{-1} \cdot \sum_{r=1}^{R} S_u(\lambda_u; \mathbf{y}_1^{(r)})$ and $R^{-1} \cdot \sum_{r=1}^{R} J_u(\lambda_u; \mathbf{y}_1^{(r)})$, respectively.

Now, consider the simple case where F belongs to either model 1 or model 2. For $\lambda_1 \in \Lambda_1$, $\lambda_2 \in \Lambda_2$, and $\mathbf{y} \in \mathcal{Y}$, set

$$\phi^*_{\lambda_1,\lambda_2}(\mathbf{y}) \equiv \begin{cases} 1 & \text{for } f(\mathbf{y}; F_{1,\lambda_1}) < f(\mathbf{y}; F_{2,\lambda_2}), \\ 0 & \text{otherwise.} \end{cases} \tag{21}$$

Then $\phi^*_{\lambda_1,\lambda_2}|_{\lambda_1=\lambda_1^0,\lambda_2=\lambda_2^0}$ $(\equiv \phi^*_{\lambda_1^0,\lambda_2^0})$ is the likelihood ratio test (LRT) for testing the new hypothesis testing problem with the null hypothesis $H_0' : F = F_{1,\lambda_1^0}$ versus the alternative $H_1' : F = F_{2,\lambda_2^0}$. Let ϕ be any randomized test, i.e., $0 \le \phi(\mathbf{y}) \le 1$ for $\mathbf{y} \in \mathcal{Y}$. When $\mathbf{y}$ is observed and the randomized test ϕ is used for this new hypothesis testing problem, we reject H_0' in favor of H_1' with probability $\phi(\mathbf{y})$. For any randomized test ϕ, let α_ϕ and β_ϕ denote, respectively, the type I error and the type II error of ϕ for this new hypothesis testing problem. Then, for any randomized test ϕ,

$$\begin{aligned} \alpha_\phi + \beta_\phi &= \sum_{\mathbf{y}\in\mathcal{Y}} \phi(\mathbf{y}) f\left(\mathbf{y}; F_{1,\lambda_1^0}\right) + \sum_{\mathbf{y}\in\mathcal{Y}} [1-\phi(\mathbf{y})] f\left(\mathbf{y}; F_{2,\lambda_2^0}\right) \\ &= 1 + \sum_{\mathbf{y}\in\mathcal{Y}} \phi(\mathbf{y}) \left[f\left(\mathbf{y}; F_{1,\lambda_1^0}\right) - f\left(\mathbf{y}; F_{2,\lambda_2^0}\right)\right] \\ &\ge 1 + \sum_{\mathbf{y}\in\mathcal{Y}} \phi^*_{\lambda_1^0,\lambda_2^0}(\mathbf{y}) \left[f\left(\mathbf{y}; F_{1,\lambda_1^0}\right) - f\left(\mathbf{y}; F_{2,\lambda_2^0}\right)\right] = \alpha_{\phi^*_{\lambda_1^0,\lambda_2^0}} + \beta_{\phi^*_{\lambda_1^0,\lambda_2^0}}. \end{aligned} \tag{22}$$

Thus, $\phi^*_{\lambda_1^0,\lambda_2^0}$ is a test which minimizes $\alpha_\phi + \beta_\phi$ among all randomized tests for this new hypothesis testing problem.

Note that $d_u^{n_1,\ldots,n_T}(\lambda_u) \to 0$ as $d_u(\lambda_u) \to 0$ for $u \in \{1,2\}$ and that

$$d_1^{n_1,\ldots,n_T}(\lambda_1) - d_2^{n_1,\ldots,n_T}(\lambda_2) = E\left(\log\left[\frac{f(\mathbf{y}; F_{2,\lambda_2})}{f(\mathbf{y}; F_{1,\lambda_1})}\right]; F\right) \tag{23}$$

for $\lambda_1 \in \Lambda_1$ and $\lambda_2 \in \Lambda_2$. When $f(\mathbf{y}; F_{1,\lambda_1})|_{\lambda_1=\hat{\lambda}_1} < f(\mathbf{y}; F_{2,\lambda_2})|_{\lambda_2=\hat{\lambda}_2}$, it is very likely that $d_1^{n_1,\ldots,n_T}(\lambda_1^{n_1,\ldots,n_T}) > d_2^{n_1,\ldots,n_T}(\lambda_2^{n_1,\ldots,n_T})$ and $d_1(\lambda_1^0) > d_2(\lambda_2^0)$. Thus, in the paper, we suggest to use the test $\phi^*_{\lambda_1,\lambda_2}|_{\lambda_1=\hat{\lambda}_1,\lambda_2=\hat{\lambda}_2}(\equiv \phi^*_{\hat{\lambda}_1,\hat{\lambda}_2})$ for the original hypothesis testing problem with the null hypothesis $H_0 : d_1(\lambda_1^0) \le d_2(\lambda_2^0)$ versus the alternative H_1 : $d_1(\lambda_1^0) > d_2(\lambda_2^0)$. That is, we choose model 2 for $f(\mathbf{y}; F_{1,\lambda_1})|_{\lambda_1=\hat{\lambda}_1} < f(\mathbf{y}; F_{2,\lambda_2})|_{\lambda_2=\hat{\lambda}_2}$ and model 1 otherwise.

3. AN EXAMPLE

For $\lambda_1 \in \Lambda_1$, let F_{1,λ_1} denote the c.d.f. of the beta(λ_1) or Dirichlet(λ_1) distribution, a conjugate prior of the binomial($n_t; \theta_\mathbf{t}$) or multinomial($n_t; \theta_t$) distribution, where $\lambda_1 \equiv (\lambda_{11}, \ldots, \lambda_{1,k+1})^T$ and $\Lambda_1 = (0, \infty)^{k+1}$. In this case, $q_1 = k + 1$. For $\lambda_1 \in \Lambda_1$,

$$\pi_1(\theta_t; \lambda_1) = 1_\Theta(\theta_t) \cdot \frac{\Gamma(\lambda_{1s})}{\prod_{i=1}^{k+1} \Gamma(\lambda_{1i})} \cdot \prod_{i=1}^{k+1} \theta_{it}^{\lambda_{1i}-1},$$

where $1_\Theta(\theta_t) = 1$ for $\theta_t \in \Theta$ and 0 otherwise. Set $\lambda_{1s} \equiv \sum_{i=1}^{k+1} \lambda_{1i}$ and $\lambda_1' \equiv \lambda_1/\lambda_{1s}$.

Set $\eta_t \equiv (\log(\theta_{1t}/\theta_{k+1,t}), \ldots, \log(\theta_{kt}/\theta_{k+1,t}))^T (\equiv (\eta_{1t}, \ldots, \eta_{kt})^T)$. Then $\theta_{it} = \exp(\eta_{it})/ [1 + \sum_{i'=1}^k \exp(\eta_{i't})]$ for $i \in \{1, \ldots, k\}$. Let $N(\mu, \mathbf{\Sigma})$ denote the k-variate normal distribution with mean vector μ and $k \times k$ positive definite covariance matrix $\mathbf{\Sigma}$. When η_t has the $N(\mu, \mathbf{\Sigma})$ distribution for some μ $(\equiv (\mu_1, \ldots, \mu_k)^T) \in \mathcal{R}^k$ and positive definite covariance matrix $\mathbf{\Sigma}$ $(\equiv (\Sigma_{ii'})_{k\times k})$, we say that θ_t has the transformed-normal(λ_2) distribution, where $\lambda_2 \equiv (\mu^T, \Sigma^{11}, \ldots, \Sigma^{1k}, \Sigma^{22}, \ldots, \Sigma^{2k}, \ldots, \Sigma^{kk})^T (\equiv (\lambda_{21}, \ldots, \lambda_{2,k(k+3)/2})^T)$ with $(\Sigma^{ii'})_{k\times k} = \mathbf{\Sigma}^{-1}$. For $\lambda_2 \in \Lambda_2$, let F_{2,λ_2} denote the c.d.f. of the transformed-normal(λ_2) distribution, where $\Lambda_2 = \mathcal{R}^k \times \{(\Sigma^{11}, \ldots, \Sigma^{1k}, \Sigma^{22}, \ldots, \Sigma^{2k}, \ldots, \Sigma^{kk})^T : (\Sigma^{ii'})_{k\times k}$ is a $k \times k$ positive definite covariance matrix$\}$. Then Λ_2 is an open subset of $\mathcal{R}^{k(k+3)/2}$. In this case,

$q_2 = k(k+3)/2 = q_1 + (k-1)(k+2)/2 \geq q_1$, where $q_1 = q_2$ if and only if $k = 1$. For $\lambda_2 \in \Lambda_2$,

$$\begin{aligned}\pi_2(\theta_t;\lambda_2) &= \frac{1}{(2\pi)^{k/2}|\boldsymbol{\Sigma}|^{1/2}} \cdot \exp\left[-\frac{1}{2}(\eta_t-\mu)^T\boldsymbol{\Sigma}^{-1}(\eta_t-\mu)\right] \cdot \left|\det\left(\frac{\partial\eta_t}{\partial\theta_t^T}\right)\right| \\ &= \frac{1}{(2\pi)^{k/2}|\boldsymbol{\Sigma}|^{1/2}\prod_{i=1}^{k+1}\theta_{it}} \cdot \exp\left[-\frac{1}{2}(\eta_t-\mu)^T\boldsymbol{\Sigma}^{-1}(\eta_t-\mu)\right],\end{aligned}$$

where

$$\frac{\partial\eta_t}{\partial\theta_t^T} = \operatorname{diag}\left\{\frac{1}{\theta_{1t}},\ldots,\frac{1}{\theta_{kt}}\right\} + \frac{1}{\theta_{k+1,t}} \cdot 1_{k\times1}\,1_{k\times1}^T.$$

For $\lambda_1 \in \Lambda_1$, it follows from Johnson *et al.* (1997, pages 80 and 81) that

$$f(\mathbf{y}_t;F_{1,\lambda_1}) = 1_{\mathcal{Y}_{n_t}}(\mathbf{y}_t) \cdot \exp\left[\sum_{j=0}^{n_t-1}\log\left(\frac{j+1}{\lambda_{1s}+j}\right) - \sum_{i=1}^{k+1}\sum_{j=0}^{y_{it}-1}\log\left(\frac{j+1}{\lambda_{1i}+j}\right)\right].$$

For $\lambda_2 \in \Lambda_2$, let ϕ_{λ_2} and Φ_{λ_2} denote, respectively, the p.d.f. and the c.d.f. of the $N(\mu,\boldsymbol{\Sigma})$ distribution. For $\lambda_2 \in \Lambda_2$,

$$\begin{aligned}f(\mathbf{y}_t;F_{2,\lambda_2}) &= 1_{\mathcal{Y}_{n_t}}(\mathbf{y}_t) \cdot \frac{n_t!}{\prod_{i=1}^{k+1}y_{it}!} \cdot \int_{\mathcal{R}^k}\frac{\exp(\mathbf{y}_t^T\,\eta_t)}{[1+\sum_{i=1}^k\exp(\eta_{it})]^{n_t}}\,d\,\Phi_{\lambda_2}(\eta_t) \\ &\equiv 1_{\mathcal{Y}_{n_t}}(\mathbf{y}_t) \cdot \frac{n_t!}{\prod_{i=1}^{k+1}y_{it}!} \cdot a(\lambda_2;\mathbf{y}_t).\end{aligned}$$

For $\lambda_2 \in \Lambda_2$, set $b(\lambda_2;\mathbf{y}_t) \equiv \partial a(\lambda_2;\mathbf{y}_t)/\partial\lambda_2$ and $c(\lambda_2;\mathbf{y}_t) \equiv \partial b(\lambda_2;\mathbf{y}_t)/\partial\lambda_2^T$. Then

$$b(\lambda_2;\mathbf{y}_t) = \int_{\mathcal{R}^k}\frac{\partial\phi_{\lambda_2}(\eta_t)/\partial\lambda_2}{\phi_{\lambda_2}(\eta_t)} \cdot \frac{\exp(\mathbf{y}_t^T\,\eta_t)}{[1+\sum_{i=1}^k\exp(\eta_{it})]^{n_t}}\,d\,\Phi_{\lambda_2}(\eta_t)$$

and

$$c(\lambda_2;\mathbf{y}_t) = \int_{\mathcal{R}^k}\frac{\partial^2\phi_{\lambda_2}(\eta_t)/\partial\lambda_2\partial\lambda_2^T}{\phi_{\lambda_2}(\eta_t)} \cdot \frac{\exp(\mathbf{y}_t^T\,\eta_t)}{[1+\sum_{i=1}^k\exp(\eta_{it})]^{n_t}}\,d\,\Phi_{\lambda_2}(\eta_t)$$

for $\lambda_2 \in \Lambda_2$. A quick way to numerically evaluate $a(\lambda_2;\mathbf{y}_t)$, $b(\lambda_2;\mathbf{y}_t)$, and $c(\lambda_2;\mathbf{y}_t)$ for $t \in \{1,\ldots,T\}$ is to utilize the method of the multivariate Gauss-Hermite integration, e.g., see Fahrmeir and Tutz (2001, pages 447-449). All of nodes and weights of the Hermite polynomial of 32 degrees are shown in the appendix for the method of the multivariate Gauss-Hermite integration.

Observe that, for $\lambda_1 \in \Lambda_1$, $\lambda_2 \in \Lambda_2$, and $\mathbf{y}_t \in \mathcal{Y}_{n_t}$,

$$\begin{aligned}
\ell_1(\lambda_1;\mathbf{y}_t) &= \sum_{j=0}^{n_t-1} \log\left(\frac{j+1}{\lambda_{1s}+j}\right) - \sum_{i=1}^{k+1}\sum_{j=0}^{y_{it}-1} \log\left(\frac{j+1}{\lambda_{1i}+j}\right), \\
\ell_2(\lambda_2;\mathbf{y}_t) &= \log(n_t!) - \sum_{i=1}^{k+1} \log(y_{it}!) + \log\left[a(\lambda_2;\mathbf{y}_t)\right], \\
S_1(\lambda_1;\mathbf{y}_t) &= \left(\sum_{j=0}^{y_{1t}-1} \frac{1}{\lambda_{11}+j}, \ldots, \sum_{j=0}^{y_{k+1,t}-1} \frac{1}{\lambda_{1,k+1}+j}\right)^T - \left[\sum_{j=0}^{n_t-1} \frac{1}{\lambda_{1s}+j}\right] \cdot 1_{(k+1)\times 1}, \\
S_2(\lambda_2;\mathbf{y}_t) &= \frac{b(\lambda_2;\mathbf{y}_t)}{a(\lambda_2;\mathbf{y}_t)},
\end{aligned}$$

$$\begin{aligned}
J_1(\lambda_1;\mathbf{y}_t) &= \operatorname{diag}\left\{\sum_{j=0}^{y_{1t}-1} \frac{1}{(\lambda_{11}+j)^2}, \ldots, \sum_{j=0}^{y_{k+1,t}-1} \frac{1}{(\lambda_{1,k+1}+j)^2}\right\} \\
&\quad - \left[\sum_{j=0}^{n_t-1} \frac{1}{(\lambda_{1s}+j)^2}\right] \cdot 1_{(k+1)\times 1}\, 1^T_{(k+1)\times 1},
\end{aligned}$$

and

$$J_2(\lambda_2;\mathbf{y}_t) = \frac{b(\lambda_2;\mathbf{y}_t)\, b^T(\lambda_2;\mathbf{y}_t) - a(\lambda_2;\mathbf{y}_t)\cdot c(\lambda_2;\mathbf{y}_t)}{[a(\lambda_2;\mathbf{y}_t)]^2}.$$

For $\lambda_1 \in \Lambda_1$ and $\mathbf{y} \in \mathcal{Y}$, set $J_1(\lambda_1;\mathbf{y}) \equiv \operatorname{diag}\{b_1(\mathbf{y}), \ldots, b_{k+1}(\mathbf{y})\} - b_s \cdot 1_{(k+1)\times 1}\, 1^T_{(k+1)\times 1}$. Then $b_s = \sum_{t=1}^{T}\sum_{j=0}^{n_t-1} 1/(\lambda_{1s}+j)^2$ and $b_i(\mathbf{y}) = \sum_{t=1}^{T}\sum_{j=0}^{y_{it}-1} 1/(\lambda_{1i}+j)^2$ for $i \in \{1, \ldots, k+1\}$ and $\mathbf{y} \in \mathcal{Y}$. When $b_1(\mathbf{y}), \ldots, b_{k+1}(\mathbf{y}) > 0$ and $1/b_s \neq \sum_{i=1}^{k+1} 1/b_i(\mathbf{y})$ for $\mathbf{y} \in \mathcal{Y}$, we have

$$\begin{aligned}
[J_1(\lambda_1;\mathbf{y})]^{-1} &= \operatorname{diag}\left\{\frac{1}{b_1(\mathbf{y})}, \ldots, \frac{1}{b_{k+1}(\mathbf{y})}\right\} \\
&\quad + \frac{1}{1/b_s - \sum_{i=1}^{k+1} 1/b_i(\mathbf{y})} \left(\frac{1}{b_1(\mathbf{y})}, \ldots, \frac{1}{b_{k+1}(\mathbf{y})}\right)^T \left(\frac{1}{b_1(\mathbf{y})}, \ldots, \frac{1}{b_{k+1}(\mathbf{y})}\right).
\end{aligned}$$

4. A SIMULATION STUDY

In this section, consider the situation where $F = p^* \cdot F_{1,\lambda_1^*} + (1 - p^*) \cdot F_{2,\lambda_2^*}$, i.e., $\pi(\cdot) = p^* \cdot \pi_1(\cdot; \lambda_1^*) + (1 - p^*) \cdot \pi_2(\cdot; \lambda_2^*)$, for some $p^* \in [0, 1]$, $\lambda_1^* \in \Lambda_1$, and $\lambda_2^* \in \Lambda_2$.

For $\lambda_u \in \Lambda_u$ and $u \in \{1, 2\}$,

$$g_u(\lambda_u) = p^* \cdot E\left(S_u(\lambda_u; \theta_t); F_{1,\lambda_1^*}\right) + (1 - p^*) \cdot E\left(S_u(\lambda_u; \theta_t); F_{2,\lambda_2^*}\right)$$

and

$$h_u(\lambda_u) = p^* \cdot E\left(J_u(\lambda_u; \theta_t); F_{1,\lambda_1^*}\right) + (1 - p^*) \cdot E\left(J_u(\lambda_u; \theta_t); F_{2,\lambda_2^*}\right).$$

For the simulation study, we choose $T = 300$ and $n_1 = \ldots = n_T = 35$. Consider the following three possible cases.

Case 1 : $F = F_{1,\lambda_1^*}$, i.e., $p^* = 1$ and $\lambda_1^* = \lambda_1^0$.

Observe that, for $\lambda_2 \in \Lambda_2$,

$$g_2(\lambda_2) = \frac{1}{2} \cdot \left[\frac{\partial \log(|\boldsymbol{\Sigma}^{-1}|)}{\partial \lambda_2} - E\left(\frac{\partial \left(\eta_t - \mu\right)^T \boldsymbol{\Sigma}^{-1} \left(\eta_t - \mu\right)}{\partial \lambda_2}; F_{1,\lambda_1^0}\right)\right]$$

and

$$h_2(\lambda_2) = -\frac{1}{2} \cdot \left[\frac{\partial^2 \log(|\boldsymbol{\Sigma}^{-1}|)}{\partial \lambda_2 \partial \lambda_2^T} - E\left(\frac{\partial^2 \left(\eta_t - \mu\right)^T \boldsymbol{\Sigma}^{-1} \left(\eta_t - \mu\right)}{\partial \lambda_2 \partial \lambda_2^T}; F_{1,\lambda_1^0}\right)\right],$$

where

$$E\left(\eta_{it}; F_{1,\lambda_1^0}\right) = E\left(\log(\theta_{it}); F_{1,\lambda_1^0}\right) - E\left(\log(\theta_{k+1,t}); F_{1,\lambda_1^0}\right)$$

and

$$
\begin{aligned}
& E\left(\eta_{it}\,\eta_{i't}; F_{1,\lambda_1^0}\right) \\
= \; & E\left(\log(\theta_{it})\log(\theta_{i't}); F_{1,\lambda_1^0}\right) - E\left(\log(\theta_{k+1,t})\left[\log(\theta_{it}) + \log(\theta_{i't})\right]; F_{1,\lambda_1^0}\right) \\
& + E\left(\left[\log(\theta_{k+1,t})\right]^2; F_{1,\lambda_1^0}\right)
\end{aligned}
$$

for $i, i' \in \{1, \ldots, k\}$ with $k \geq 2$. When θ_t has the beta(λ_1^0) or Dirichlet(λ_1^0) distribution, θ_{it} has the beta($\lambda_{1i}^0, \lambda_{1s}^0 - \lambda_{1i}^0$) distribution and

$$
\int_0^1 \frac{\Gamma(\lambda_{1s}^0)}{\Gamma(\lambda_{1i}^0)\,\Gamma(\lambda_{1s}^0 - \lambda_{1i}^0)}\,\theta_{it}^{\lambda_{1i}^0 - 1}\,(1-\theta_{it})^{\lambda_{1s}^0 - \lambda_{1i}^0 - 1}\,d\,\theta_{it} = 1
$$

for $i \in \{1, \ldots, k+1\}$. Taking the derivative with respect to λ_{1i}^0 for $i \in \{1, \ldots, k+1\}$, we have

$$
E\left(\log(\theta_{it}); F_{1,\lambda_1^0}\right) = \psi\left(\lambda_{1i}^0\right) - \psi\left(\lambda_{1s}^0\right),
$$

where $\psi(x) \equiv d\log\left[\Gamma(x)\right]/dx$ for $x > 0$. For $x > 0$,

$$
\psi(x) = -c + (x-1)\cdot\sum_{i=1}^{\infty}\frac{1}{i(i+x-1)},
$$

where c (≈ 0.5772156649) is the Euler constant. See, e.g., Abramowitz and Stegun (1964, page 259). Taking the derivative with respect to λ_{1i}^0 twice for $i \in \{1, \ldots, k+1\}$, we have

$$
E\left(\left[\log(\theta_{it})\right]^2; F_{1,\lambda_1^0}\right) = \psi'\left(\lambda_{1i}^0\right) - \psi'\left(\lambda_{1s}^0\right) + \left[\psi\left(\lambda_{1i}^0\right) - \psi\left(\lambda_{1s}^0\right)\right]^2,
$$

where $\psi'(x) \equiv d\psi(x)/dx$ for $x > 0$, For $x > 0$.

$$
\psi'(x) = \sum_{i=0}^{\infty}\frac{1}{(x+i)^2}.
$$

See, e.g., Abramowitz and Stegun (1964, page 260). Since θ_t has the Dirichlet(λ_1^0) distribution for $k \geq 2$, $(\theta_{it}, \theta_{i't})^T$ has the Dirichlet$(\lambda_{1i}^0, \lambda_{1i'}^0, \lambda_{1s}^0 - \lambda_{1i}^0 - \lambda_{1i'}^0)$ distribution and

$$\int_0^1 \left[\int_0^{1-\theta_{i't}} \frac{\Gamma(\lambda_{1s}^0) \cdot \theta_{it}^{\lambda_{1i}^0 - 1} \theta_{i't}^{\lambda_{1i'}^0 - 1} (1 - \theta_{it} - \theta_{i't})^{\lambda_{1s}^0 - \lambda_{1i}^0 - \lambda_{1i'}^0 - 1}}{\Gamma(\lambda_{1i}^0)\, \Gamma(\lambda_{1i'}^0)\, \Gamma(\lambda_{1s}^0 - \lambda_{1i}^0 - \lambda_{1i'}^0)} \, d\theta_{it} \right] d\theta_{i't} = 1$$

for $i \neq i'$ and $i, i' \in \{1, \ldots, k+1\}$ with $k \geq 2$. Taking the derivative with respect to λ_{1i}^0 and then $\lambda_{1i'}^0$, we have

$$E\left(\log(\theta_{it}) \log(\theta_{i't}); F_{1,\lambda_1^0}\right) = -\psi'\left(\lambda_{1s}^0\right) + \left[\psi\left(\lambda_{1i}^0\right) - \psi\left(\lambda_{1s}^0\right)\right] \left[\psi\left(\lambda_{1i'}^0\right) - \psi\left(\lambda_{1s}^0\right)\right]$$

for $i \neq i'$ and $i, i' \in \{1, \ldots, k+1\}$ with $k \geq 2$. Finally, λ_2^0 can be numerically evaluated by utilizing the Newton-Raphson method.

We simulate an i.i.d. sample $\{\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(R)}\}$ of size R, e.g., $R = 100\,000$, from the c.d.f. $F_{\mathbf{y};1,\lambda_1^0}$. Since all of λ_1^0, λ_2^0, and $\lambda_2^{n_1}$ are known in a simulation study, we can numerically evaluate $P(\{\phi^*_{\lambda_1^0,\lambda_2^0}(\mathbf{y}) = 1\}; F_{1,\lambda_1^0})$ and $P(\{\phi^*_{\lambda_1^{n_1},\lambda_2^{n_1}}(\mathbf{y}) = 1\}; F_{1,\lambda_1^0})$ by $|\{r : \phi^*_{\lambda_1^0,\lambda_2^0}(\mathbf{y}^{(r)}) = 1\}|/R$ and $|\{r : \phi^*_{\lambda_1^{n_1},\lambda_2^{n_1}}(\mathbf{y}^{(r)}) = 1\}|/R$, respectively, where $|S|$ denotes the number of elements in S for any set S. Since both λ_1^0 and λ_2^0 are unknown in a real problem, we can numerically evaluate $P(\{\phi^*_{\hat{\lambda}_1,\hat{\lambda}_2}(\mathbf{y}) = 1\}; F_{1,\lambda_1^0})$ by $|\{r : \phi^*_{\hat{\lambda}_1(\mathbf{y}^{(r)}),\hat{\lambda}_2(\mathbf{y}^{(r)})}(\mathbf{y}^{(r)}) = 1\}|/R$.

For the simulation study, consider the case where $F = F_{1,\lambda_1^0}$ with $\lambda_1^0 = (7, 2, 1)^T$. We simulate an i.i.d. sample $\{\theta_t^{(1)}, \ldots, \theta_t^{(R)}\}$ of size R, e.g., $R = 100\,000$, from the c.d.f. F_{1,λ_1^0}. Set $\mu^{0(0)} \equiv (\mu_1^{0(0)}, \ldots, \mu_k^{0(0)})^T$ and $\mathbf{\Sigma}^{0(0)} \equiv (\Sigma_{ii'}^{0(0)})_{k\times k}$, where

$$\mu_i^{0(0)} \equiv \frac{1}{R} \cdot \sum_{r=1}^{R} \log\left(\frac{\theta_{it}^{(r)}}{\theta_{k+1,t}^{(r)}}\right)$$

and

$$\Sigma_{ii'}^{0(0)} \equiv \frac{1}{R-1} \cdot \sum_{r=1}^{R} \left[\log \left(\frac{\theta_{it}^{(r)}}{\theta_{k+1,t}^{(r)}} \right) - \mu_i^{0(0)} \right] \left[\log \left(\frac{\theta_{i't}^{(r)}}{\theta_{k+1,t}^{(r)}} \right) - \mu_{i'}^{0(0)} \right]$$

for $i, i' \in \{1, \ldots, k\}$. Iterating equation (8), we obtain $\lambda_2^0 = (-1.450, -2.450, 1.273, -0.109, 0.565)^T$. Iterating equation (20), $\lambda_2^{n_1}$ is obtained and shown in Table 1 for $n_1 \in \{35, 70, 140\}$. It is easily seen from Table 1 that $||\lambda_2^{n_1} - \lambda_2^0||_2$ decreases as n_1 increases, where $||\lambda_2^{n_1} - \lambda_2^0||_2 \equiv [(\lambda_2^{n_1} - \lambda_2^0)^T(\lambda_2^{n_1} - \lambda_2^0)]^{1/2}$. Finally, we obtain $P(\{\phi^*_{\lambda_1^0,\lambda_2^0}(\mathbf{y}) = 1\}; F_{1,\lambda_1^0}) \approx 0.007$, $P(\{\phi^*_{\lambda_1^{35},\lambda_2^{35}}(\mathbf{y}) = 1\}; F_{1,\lambda_1^0}) \approx 0.064$, and $P(\{\phi^*_{\hat{\lambda}_1,\hat{\lambda}_2}(\mathbf{y}) = 1\}; F_{1,\lambda_1^0}) \approx 0.127$ which are less than 0.5 and shown in Table 3.

Case 2 : $F = F_{2,\lambda_2^*}$, i.e., $p^* = 0$ and $\lambda_2^* = \lambda_2^0$, where $F_{2,\lambda_2^*} \notin \{F_{1,\lambda_1} : \lambda_1 \in \Lambda_1\}$.

Observe that, for $\lambda_1 \in \Lambda_1$,

$$\begin{aligned} g_1(\lambda_1) &= \psi(\lambda_{1s}) \cdot 1_{(k+1)\times 1} - (\psi(\lambda_{11}), \ldots, \psi(\lambda_{1,k+1}))^T \\ &\quad + \left(E\left(\log(\theta_{1t}); F_{2,\lambda_2^0}\right), \ldots, E\left(\log(\theta_{k+1,t}); F_{2,\lambda_2^0}\right)\right)^T \end{aligned}$$

and

$$h_1(\lambda_1) = \mathrm{diag}\left\{\psi'(\lambda_{11}), \ldots, \psi'(\lambda_{1,k+1})\right\} - \psi'(\lambda_{1s}) \cdot 1_{(k+1)\times 1} 1_{(k+1)\times 1}^T,$$

where

$$E\left(\log(\theta_{it}); F_{2,\lambda_2^0}\right) = \mu_i^0 + E\left(\log(\theta_{k+1,t}); F_{2,\lambda_2^0}\right)$$

for $i \in \{1, \ldots, k\}$ and

$$E\left(\log(\theta_{k+1,t}); F_{2,\lambda_2^0}\right) = -E\left(\log\left[1 + \sum_{i=1}^{k} \exp(\eta_{it})\right]; F_{2,\lambda_2^0}\right).$$

Here $E(\log[1+\sum_{i=1}^{k}\exp(\eta_{it})]; F_{2,\lambda_2^0})$ can be numerically evaluated by the method of the multivariate Gauss-Hermite integration. Finally, λ_1^0 can be numerically evaluated by utilizing the Newton-Raphson method.

We simulate an i.i.d. sample $\{\mathbf{y}^{(1)},\ldots,\mathbf{y}^{(R)}\}$ of size R, e.g., $R = 100\,000$, from the c.d.f. $F_{\mathbf{y};2,\lambda_2^0}$. Since all of λ_1^0, λ_2^0, and $\lambda_1^{n_1}$ are known in a simulation study, we can numerically evaluate $P(\{\phi^*_{\lambda_1^0,\lambda_2^0}(\mathbf{y}) = 0\}; F_{2,\lambda_2^0})$ and $P(\{\phi^*_{\lambda_1^{n_1},\lambda_2^{n_1}}(\mathbf{y}) = 0\}; F_{2,\lambda_2^0})$ by $|\{r : \phi^*_{\lambda_1^0,\lambda_2^0}(\mathbf{y}^{(r)}) = 0\}|/R$ and $|\{r : \phi^*_{\lambda_1^{n_1},\lambda_2^{n_1}}(\mathbf{y}^{(r)}) = 0\}|/R$, respectively. Since both λ_1^0 and λ_2^0 are unknown in a real problem, we can numerically evaluate $P(\{\phi^*_{\hat\lambda_1,\hat\lambda_2}(\mathbf{y}) = 0\}; F_{2,\lambda_2^0})$ by $|\{r : \phi^*_{\hat\lambda_1(\mathbf{y}^{(r)}),\hat\lambda_2(\mathbf{y}^{(r)})}(\mathbf{y}^{(r)}) = 0\}|/R$.

For the simulation study, consider the case where $F = F_{2,\lambda_2^0}$ with $\lambda_2^0 = (-1.450, -2.450, 1.273, -0.109, 0.565)^T$. We simulate an i.i.d. sample $\{\theta_t^{(1)},\ldots,\theta_t^{(R)}\}$ of size R, e.g., $R = 100\,000$, from the c.d.f. F_{2,λ_2^0}. Set $\lambda_1^{0(0)} \equiv \lambda_{1s}^{0(0)} \cdot \lambda_1'^{0(0)}$, where $\lambda_{1s}^{0(0)}$ is the MME of λ_{1s} proposed in Shiau *et al.* (2005) and

$$\lambda_1'^{0(0)} \equiv \frac{1}{R} \cdot \sum_{r=1}^{R} \theta_t^{(r)}.$$

Iterating equation (8), we obtain $\lambda_1^0 = (5.771, 1.707, 0.884)^T$. Iterating equation (20), $\lambda_1^{n_1}$ is obtained and shown in Table 2 for $n_1 \in \{35, 70, 140\}$. Similarly, it is easily seen that $||\lambda_1^{n_1} - \lambda_1^0||_2$ decreases as n_1 increases, where $||\lambda_1^{n_1} - \lambda_1^0||_2 \equiv [(\lambda_1^{n_1} - \lambda_1^0)^T(\lambda_1^{n_1} - \lambda_1^0)]^{1/2}$. Finally, we obtain $P(\{\phi^*_{\lambda_1^0,\lambda_2^0}(\mathbf{y}) = 0\}; F_{2,\lambda_2^0}) \approx 0.021$, $P(\{\phi^*_{\lambda_1^{n_1},\lambda_2^{n_1}}(\mathbf{y}) = 0\}; F_{2,\lambda_2^0}) \approx 0.008$, and $P(\{\phi^*_{\hat\lambda_1,\hat\lambda_2}(\mathbf{y}) = 0\}; F_{2,\lambda_2^0}) \approx 0.014$ which are all less than 0.5 and shown in Table 3.

Case 3 : $F = p^* \cdot F_{1,\lambda_1^*} + (1 - p^*) \cdot F_{2,\lambda_2^*}$ for some $0 < p^* < 1$, $\lambda_1^* \in \Lambda_1$, and $\lambda_2^* \in \Lambda_2$, where $p^* \cdot F_{1,\lambda_1^*} + (1-p^*) \cdot F_{2,\lambda_2^*} \notin \{F_{u,\lambda_u} : \lambda_u \in \Lambda_u \text{ and } u \in \{1,2\}\}$.

We simulate an i.i.d. sample $\{\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(R)}\}$ of size R, e.g., $R = 100\,000$, from the c.d.f. $F_{\mathbf{y}}$. The initial value of λ_u^0 for $u \in \{1,2\}$ can be obtained by the same methods in *Case 1* and *2*. Iterating equation (8), λ_u^0 can be numerically evaluated. When $d_1(\lambda_1^0) \leq d_2(\lambda_2^0)$, we can numerically evaluate $P(\{\phi^*_{\lambda_1^0,\lambda_2^0}(\mathbf{y}) = 1\}; F)$, $P(\{\phi^*_{\lambda_1^{35},\lambda_2^{35}}(\mathbf{y}) = 1\}; F)$, and $P(\{\phi^*_{\hat{\lambda}_1,\hat{\lambda}_2}(\mathbf{y}) = 1\}; F)$ by $|\{r : \phi^*_{\lambda_1^0,\lambda_2^0}(\mathbf{y}^{(r)}) = 1\}|/R$, $|\{r : \phi^*_{\lambda_1^{35},\lambda_2^{35}}(\mathbf{y}^{(r)}) = 1\}|/R$, and $|\{r : \phi^*_{\hat{\lambda}_1(\mathbf{y}^{(r)}),\hat{\lambda}_2(\mathbf{y}^{(r)})}(\mathbf{y}^{(r)}) = 1\}|/R$, respectively. When $d_1(\lambda_1^0) > d_2(\lambda_2^0)$, we can numerically evaluate $P(\{\phi^*_{\lambda_1^0,\lambda_2^0}(\mathbf{y}) = 0\}; F)$, $P(\{\phi^*_{\lambda_1^{35},\lambda_2^{35}}(\mathbf{y}) = 0\}; F)$, and $P(\{\phi^*_{\hat{\lambda}_1,\hat{\lambda}_2}(\mathbf{y}) = 0\}; F)$ by $|\{r : \phi^*_{\lambda_1^0,\lambda_2^0}(\mathbf{y}^{(r)}) = 0\}|/R$, $|\{r : \phi^*_{\lambda_1^{35},\lambda_2^{35}}(\mathbf{y}^{(r)}) = 0\}|/R$, and $|\{r : \phi^*_{\hat{\lambda}_1(\mathbf{y}^{(r)}),\hat{\lambda}_2(\mathbf{y}^{(r)})}(\mathbf{y}^{(r)}) = 0\}|/R$, respectively.

First, consider the case where $F = p^* \cdot F_{1,\lambda_1^*} + (1-p^*) \cdot F_{2,\lambda_2^*}$ for $p^* \in \{1/6, 1/3, 1/2, 2/3, 5/6\}$ with $\lambda_1^* = (7,2,1)^T$ and $\lambda_2^* = (-1.450, -2.450, 1.273, -0.109, 0.565)^T$. Iterating equation (8), we can numerically evaluate λ_u^0 for $u \in \{1,2\}$. $d_1(\lambda_1^0)$, $d_2(\lambda_2^0)$, and the probability of choosing the wrong hypothesis for each of $\phi^*_{\lambda_1^0,\lambda_2^0}$, $\phi^*_{\lambda_1^{35},\lambda_2^{35}}$, and $\phi^*_{\hat{\lambda}_1,\hat{\lambda}_2}$ are shown in Table 3. It is easily seen from Table 3 that $d_1(\lambda_1^0)$ decreases and $d_2(\lambda_2^0)$ increases as p^* increases. The probability of choosing the wrong hypothesis is less than 0.5 except for $\phi^*_{\hat{\lambda}_1,\hat{\lambda}_2}$ with $p^* = 2/3$.

Next, consider the case where $F = p^* \cdot F_{1,\lambda_1^*} + (1-p^*) \cdot F_{2,\lambda_2^*}$ for $p^* \in \{1/6, 1/3, 1/2, 2/3, 5/6\}$ with $\lambda_1^* = (5.771, 1.707, 0.884)^T$ and $\lambda_2^* = (-1.450, -2.450, 1.273, -0.109, 0.565)^T$. Iterate equation (8), we can numerically evaluate λ_u^0 for $u \in \{1,2\}$. $d_1(\lambda_1^0)$, $d_2(\lambda_2^0)$, and the probability of choosing the wrong hypothesis for each of $\phi^*_{\lambda_1^0,\lambda_2^0}$, $\phi^*_{\lambda_1^{35},\lambda_2^{35}}$, and $\phi^*_{\hat{\lambda}_1,\hat{\lambda}_2}$ are shown in Table 4. Similarly, it is easily seen from Table 4 that $d_1(\lambda_1^0)$ decreases and $d_2(\lambda_2^0)$ increases as p^* increases. The probability of choosing the wrong hypothesis is less than 0.5 except for $\phi^*_{\lambda_1^0,\lambda_2^0}$ with $p^* = 1/2$. The main reason is that $d_1(\lambda_1^0)$ (≈ 2.980) and $d_2(\lambda_2^0)$ (≈ 2.944) are nearly the same.

Finally, we would like to investigate the results of the empirical Bayes process monitoring scheme proposed in Chen *et al.* (2004) and Chen *et al.* (2005) by ignoring the fact that the model used is only an approximate model. Let γ denote the false alarm rate, i.e., the probability that an out-of-control signal occurs when the manufacturing process is in control. Conventionally, γ is taken to $2\Phi(-3)$ (≈ 0.0026998).

For $u \in \{1,2\}$, we order the 100 000 $\hat{\lambda}_u$'s in decreasing order of $d(F_{u,\lambda_u^0}, F_{u,\lambda_u})|_{\lambda_u=\hat{\lambda}_u}$ ($\equiv d(F_{u,\lambda_u^0}, F_{u,\hat{\lambda}_u})$), a measure of how close $F_{u,\hat{\lambda}_u}$ is to F_{u,λ_u^0} in our study, where

$$\begin{aligned} & d\left(F_{1,\lambda_1^0}, F_{1,\lambda_1}\right) \\ = \;& \log\left[\frac{\Gamma(\lambda_{1s}^0)}{\prod_{i=1}^{k+1}\Gamma(\lambda_{1i}^0)}\right] - \log\left[\frac{\Gamma(\lambda_{1s})}{\prod_{i=1}^{k+1}\Gamma(\lambda_{1i})}\right] + \sum_{i=1}^{k+1}\left(\lambda_{1i}^0 - \lambda_{1i}\right)\left[\psi\left(\lambda_{1i}^0\right) - \psi\left(\lambda_{1s}^0\right)\right] \end{aligned}$$

for $\lambda_1 \in \Lambda_1$ and

$$\begin{aligned} & d\left(F_{2,\lambda_2^0}, F_{2,\lambda_2}\right) \\ = \;& \frac{1}{2}\left\{\log\left[\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{\Sigma}^0|}\right] + E\left((\eta_t-\mu)^T\boldsymbol{\Sigma}^{-1}(\eta_t-\mu) - \left(\eta_t-\mu^0\right)^T\left(\boldsymbol{\Sigma}^0\right)^{-1}\left(\eta_t-\mu^0\right); F_{2,\lambda_2^0}\right)\right\} \end{aligned}$$

for $\lambda_2 \in \Lambda_2$. Set $\xi_t \equiv \boldsymbol{\Sigma}^{1/2}(\eta_t - \mu^0)$, $\mu' \equiv \boldsymbol{\Sigma}^{1/2}(\mu - \mu^0)$, and $V \equiv (\boldsymbol{\Sigma}^0)^{1/2}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\Sigma}^0)^{1/2}$. Then

$$\begin{aligned} & E\left(\left(\xi_t-\mu'\right)^T V\left(\xi_t-\mu'\right); F_{2,\lambda_2^0}\right) \\ = \;& tr\left(V\cdot E\left(\left(\xi_t-\mu'\right)\left(\xi_t-\mu'\right)^T; F_{2,\lambda_2^0}\right)\right) \;=\; tr\left(V\left(I+\mu'\mu'^T\right)\right) \end{aligned}$$

and

$$
\begin{aligned}
& d\left(F_{2,\lambda_2^0}, F_{2,\lambda_2}\right) \\
= \; & \frac{1}{2}\left\{\log\left(|\mathbf{\Sigma}|\right) - \log\left(|\mathbf{\Sigma}^0|\right) + E\left(\left(\xi_t - \mu'\right)^T V\left(\xi_t - \mu'\right); F_{2,\lambda_2^0}\right) - E\left(\xi_t^T \xi_t; F_{2,\lambda_2^0}\right)\right\} \\
= \; & \frac{1}{2}\left\{\log\left(|\mathbf{\Sigma}|\right) - \log\left(|\mathbf{\Sigma}^0|\right) + tr\left(\mathbf{\Sigma}^0 \mathbf{\Sigma}^{-1}\right) + \left(\mu - \mu^0\right)^T \mathbf{\Sigma}^{-1}\left(\mu - \mu^0\right) - k\right\}.
\end{aligned}
$$

Thus, we pick the MLE's corresponding to the best 10th, 30th, 50th, 70th, and 90th percentiles of these 100 000 MLE's based on this measure. For the true λ_u and each MLE picked, compute the *in-control* probability and the average run length ARL_0 when the process is in control. When the process is out of control, compute the *out-of-control* probability and the average run length ARL_1.

Consider the case where $F = F_{1,\lambda_1^0}$ with $\lambda_1^0 = (7, 2, 1)^T$. Utilizing model 2 with $\lambda_2^0 = (-1.450, -2.450, 1.273, -0.109, 0.565)^T$ for monitoring, the in control probability and ARL_0 for λ_2^0 and the best 10%, 30%, 50%, 70%, 90% $\hat{\lambda}_2$'s are shown in Table 5. It is easily seen from Table 5 that the $\gamma_{\lambda_2^0}$ and all the $\gamma_{\hat{\lambda}_2}$ are less than γ. When θ_t has an out of control c.d.f. $F_{1,\tilde{\lambda}_1}$ with $\tilde{\lambda}_1 = (5, 3, 2)^T$ different from the in-control c.d.f. F_{1,λ_1^0}. The out of control probability and ARL_1 for λ_2^0 and the best 10%, 30%, 50%, 70%, 90% $\hat{\lambda}_2$'s are shown in Table 6.

Consider the case where $F = F_{2,\lambda_2^0}$ with $\lambda_2^0 = (-1.450, -2.450, 1.273, -0.109, 0.565)^T$. Utilizing model 1 with $\lambda_1^0 = (5.771, 1.707, 0.884)^T$ for monitoring, the in control probability and ARL_0 for λ_2^0 and the best 10%, 30%, 50%, 70%, 90% $\hat{\lambda}_2$'s are shown in Table 7. It is easily seen from Table 7 that the $\gamma_{\lambda_1^0}$ and all the $\gamma_{\hat{\lambda}_1}$ are larger than γ.

First, consider the case where $F = p^* \cdot F_{1,\lambda_1^*} + (1-p^*) \cdot F_{2,\lambda_2^*}$ for $p^* \in \{1/6, 1/3, 1/2, 2/3, 5/6\}$ with $\lambda_1^* = (7,2,1)^T$ and $\lambda_2^* = (-1.450, -2.450, 1.273, -0.109, 0.565)^T$. Utilizing model 2 for monitoring, the in control probability and ARL_0 for λ_2^0 and the best 10%, 30%, 50%, 70%, 90% $\hat{\lambda}_2$'s are shown in Table 8. It is easily seen from Table 8 that the $\gamma_{\lambda_2^0}$ and all the $\gamma_{\hat{\lambda}_2}$ are less than γ. Utilizing model 1 for monitoring, the in control probability and ARL_0 for λ_1^0 and the best 10%, 30%, 50%, 70%, 90% $\hat{\lambda}_1$'s are shown in Table 9. It is easily seen from Table 9 that the $\gamma_{\lambda_1^0}$ and all the $\gamma_{\hat{\lambda}_1}$ are large than γ. When θ_t has an out of control c.d.f. $F = \tilde{p} \cdot F_{1,\tilde{\lambda}_1} + (1-\tilde{p}) \cdot F_{2,\tilde{\lambda}_2}$ for $\tilde{p} \in \{1/6, 1/3, 1/2, 2/3, 5/6\}$ with $\tilde{\lambda}_1 = (5,3,2)^T$ and $\tilde{\lambda}_2 = (-0.583, -1.083, 1.787, -0.456, 1.271)^T$ different from the in control c.d.f. $F = p^* \cdot F_{1,\lambda_1^*} + (1-p^*) \cdot F_{2,\lambda_2^*}$. Utilizing model 2 for monitoring, the out of control probability and ARL_1 for λ_2^0 and the best 10%, 30%, 50%, 70%, 90% $\hat{\lambda}_2$'s are shown in Table 10 and 11, respectively.

Next, consider the case where $F = p^* \cdot F_{1,\lambda_1^*} + (1-p^*) \cdot F_{2,\lambda_2^*}$ for $p^* \in \{1/6, 1/3, 1/2, 2/3, 5/6\}$ with $\lambda_1^* = (5.771, 1.707, 0.884)^T$ and $\lambda_2^* = (-1.450, -2.450, 1.273, -0.109, 0.565)^T$. Utilizing model 2 for monitoring, the in control probability and ARL_0 for λ_2^0 and the best 10%, 30%, 50%, 70%, 90% $\hat{\lambda}_2$'s are shown in Table 12. It is easily seen from Table 12 that the $\gamma_{\lambda_2^0}$ and all the $\gamma_{\hat{\lambda}_2}$ are less than γ. Utilizing model 1 for monitoring, the in control probability and ARL_0 for λ_1^0 and the best 10%, 30%, 50%, 70%, 90% $\hat{\lambda}_1$'s are shown in Table 13. It is easily seen from Table 13 that the $\gamma_{\lambda_1^0}$ and all the $\gamma_{\hat{\lambda}_1}$ are large than γ. When θ_t has an out of control c.d.f. $F = \tilde{p} \cdot F_{1,\tilde{\lambda}_1} + (1-\tilde{p}) \cdot F_{2,\tilde{\lambda}_2}$ for $\tilde{p} \in \{1/6, 1/3, 1/2, 2/3, 5/6\}$ with $\tilde{\lambda}_1 = (5,3,2)^T$ and $\tilde{\lambda}_2 = (-0.583, -1.083, 1.787, -0.456, 1.271)^T$ different from the in-control c.d.f. $F = p^* \cdot F_{1,\lambda_1^*} + (1 - p^*) \cdot F_{2,\lambda_2^*}$. Utilizing model 2 for monitoring, the out of control probability and ARL_1 for the λ_2^0 and the best 10%, 30%, 50%, 70%, 90% $\hat{\lambda}_2$'s are shown in Table 14 and 15, respectively.

5. CONCLUSIONS AND FUTURE WORK

In the paper, we develop a model selection technique for categorical data in manufacturing process. Then an example of choosing two empirical Bayes models for categorical data is discussed. If $F = p^* \cdot F_{1,\lambda_1^*} + (1-p^*) \cdot F_{2,\lambda_2^*}$, for $p^* \in [0,1]$, $\lambda_1^* \in \Lambda_1$, and $\lambda_2^* \in \Lambda_2$, then the probability of choosing the wrong hypothesis are almost less than 0.5. For the process monitoring, since these two parametric models under consideration are only approximate models, the critical point is incorrect such that the probability of signaling out-of-control is different from the γ. What we want to do next is to utilize the resampling method to find an approximate critical point for monitoring.

APPENDIX

All of nodes and weights of the Hermite polynomial of 32 degrees are shown in the following table. This table is obtained from the following website:

http://www.efunda.com/math/num_integration/findgausshermite.cfm

No.i	abscissas x_i	weights w_i
1	-7.12581390983	$7.31067642754 \times 10^{-23}$
2	-6.40949814928	$9.23173653482 \times 10^{-19}$
3	-5.81222594946	$1.19734401957 \times 10^{-15}$
4	-5.27555098664	$4.21501019491 \times 10^{-13}$
5	-4.77716450334	$5.93329148347 \times 10^{-11}$
6	-4.30554795347	$4.09883215841 \times 10^{-9}$
7	-3.85375548542	$1.57416779440 \times 10^{-7}$
8	-3.41716749282	$3.65058512533 \times 10^{-6}$
9	-2.99249082501	$5.41658405999 \times 10^{-5}$
10	-2.57724953773	$5.36268365495 \times 10^{-4}$
11	-2.16949918361	$3.65489032677 \times 10^{-3}$
12	-1.76765410946	$1.75534288315 \times 10^{-2}$
13	-1.37037641095	$6.04581309559 \times 10^{-2}$
14	-0.97650046359	$1.51269734077 \times 10^{-1}$
15	-0.58497876544	$2.77458142303 \times 10^{-1}$
16	-0.19484074157	$3.75238352593 \times 10^{-1}$
17	0.19484074157	$3.75238352593 \times 10^{-1}$
18	0.58497876544	$2.77458142303 \times 10^{-1}$
19	0.97650046359	$1.51269734077 \times 10^{-1}$
20	1.37037641095	$6.04581309559 \times 10^{-2}$
21	1.76765410946	$1.75534288315 \times 10^{-2}$
22	2.16949918361	$3.65489032677 \times 10^{-3}$
23	2.57724953773	$5.36268365495 \times 10^{-4}$
24	2.99249082501	$5.41658405999 \times 10^{-5}$
25	3.41716749282	$3.65058512533 \times 10^{-6}$
26	3.85375548542	$1.57416779440 \times 10^{-7}$
27	4.30554795347	$4.09883215841 \times 10^{-9}$
28	4.77716450334	$5.93329148347 \times 10^{-11}$
29	5.27555098664	$4.21501019491 \times 10^{-13}$
30	5.81222594946	$1.19734401957 \times 10^{-15}$
31	6.40949814928	$9.23173653482 \times 10^{-19}$
32	7.12581390983	$7.31067642754 \times 10^{-23}$

REFERENCES

1. Abramowitz, M. and Stegun, I. A. (1964). *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables.* Dover Publications, New York.

2. Agresti, A. (2002). *Categorical Data Analysis*, 2nd ed. John Wiley & Sons, New York.

3. Carlin, B. P. and Louis, T. A. (2000). *Bayes and Empirical Bayes Methods for Data Analysis*, 2nd ed. Chapman & Hall/CRC, Boca Raton.

4. Chen, C.-R., Shiau, J.-J. H., Liao, H.-H., and Feltz, C. J. (2004). A process monitoring technique for categorical data under the beta-binomial or Dirichlet-multinomial empirical Bayes model. Technical Report, Institute of Statistics, National Chiao Tung University, Hsinchu, Taiwan.

5. Chen, C.-R., Shiau, J.-J. H., Lin, T.-Y., and Feltz, C. J. (2005). A process monitoring technique for categorical data under the transformed-normal-binomial or transformed-normal-multinomial empirical Bayes model. Technical Report, Institute of Statistics, National Chiao Tung University, Hsinchu, Taiwan.

6. Fahrmeir, A. and Tutz, G. (2001). *Multivariate Statistical Modelling Based on Generalized Linear Models*, 2nd ed. Springer-Verlag, New York.

7. Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2004). *Bayesian Data Analysis*, 2nd ed. Chapman & Hall/CRC, Boca Raton.

8. Johnson, N. L., Kotz, S., and Balakrishnan, N. (1997). *Discrete Multivariate Distributions*. John Wiley & Sons, New York.

9. McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models*, 2nd ed. Chapman and Hall, London.

10. O'Hagan, A. and Forster, J. (2004). *Kendall's Advanced Theory of Statistics, Volume 2B: Bayesian Inference*, 2nd ed. Arnold, London.

11. Shiau, J.-J. H., Chen, C.-R., and Feltz, C. J. (2005). An empirical Bayes process monitoring technique for polytomous data. *Quality and Reliability Engineering International*, **21**, 13-28.

12. Yousry, M. A., Sturm, G. W., Feltz, C. J., and Noorossana, R. (1991). Process monitoring in real time: empirical Bayes approach - discrete case. *Quality and Reliability Engineering International*, **7**, 123-132.

Table 1: $\lambda_2^{n_1}$ and $||\lambda_2^{n_1} - \lambda_2^0||_2$ for $n_1 \in \{35, 70, 140\}$ with $\lambda_2^0 = (-1.450, -2.450, 1.273, -0.109, 0.565)^T$.

n_1	$\lambda_2^{n_1}$	$\|\|\lambda_2^{n_1} - \lambda_2^0\|\|_2$
35	$(-1.430, -2.344, 1.475, -0.194, 0.862)^T$	0.385
70	$(-1.435, -2.366, 1.423, -0.172, 0.798)^T$	0.296
140	$(-1.440, -2.387, 1.376, -0.154, 0.741)^T$	0.218

Table 2: $\lambda_1^{n_1}$ and $||\lambda_1^{n_1} - \lambda_1^0||_2$ for $n_1 \in \{35, 70, 140\}$ with $\lambda_1^0 = (5.771, 1.707, 0.884)^T$.

n_1	$\lambda_1^{n_1}$	$\|\|\lambda_1^{n_1} - \lambda_1^0\|\|_2$
35	$(6.935, 1.991, 1.004)^T$	0.229
70	$(6.779, 1.953, 0.990)^T$	0.172
140	$(6.597, 1.909, 0.973)^T$	0.121

Table 3: $d_1(\lambda_1^0)$, $d_2(\lambda_2^0)$, and the probability of choosing the wrong hypothesis for each of $\phi^*_{\lambda_1^0,\lambda_2^0}$, $\phi^*_{\lambda_1^{n_1},\lambda_2^{n_1}}$, and $\phi^*_{\hat{\lambda}_1,\hat{\lambda}_2}$ for $p^* \in \{0, 1/6, 1/3, 1/2, 2/3, 5/6, 1\}$ with $\lambda_1^* = (7, 2, 1)^T$, $\lambda_2^* = (-1.450, -2.450, 1.273, -0.109, 0.565)^T$, $T = 300$, and $n_1 = \ldots = n_T = 35$.

p^*	$d_1(\lambda_1^0)$	$d_2(\lambda_2^0)$	the probability of choosing the wrong hypothesis		
			$\phi^*_{\lambda_1^0,\lambda_2^0}$	$\phi^*_{\lambda_1^{n_1},\lambda_2^{n_1}}$	$\phi^*_{\hat{\lambda}_1,\hat{\lambda}_2}$
0	10.049	0.000	0.021	0.008	0.014
1/6	7.665	0.231	0.071	0.064	0.041
1/3	5.499	0.941	0.189	0.147	0.100
1/2	3.575	2.167	0.409	0.296	0.219
2/3	1.933	3.966	0.317	0.488	0.589
5/6	0.659	6.441	0.097	0.247	0.351
1	0.000	9.859	0.007	0.064	0.127

Table 4: $d_1(\lambda_1^0)$, $d_2(\lambda_2^0)$, and the probability of choosing the wrong hypothesis for each of $\phi^*_{\lambda_1^0,\lambda_2^0}$, $\phi^*_{\lambda_1^{n_1},\lambda_2^{n_1}}$, and $\phi^*_{\hat{\lambda}_1,\hat{\lambda}_2}$ for $p^* \in \{0, 1/6, 1/3, 1/2, 2/3, 5/6, 1\}$ with $\lambda_1^* = (5.771, 1.707, 0.884)^T$, $\lambda_2^* = (-1.450, -2.450, 1.273, -0.109, 0.565)^T$, $T = 300$, and $n_1 = \ldots = n_T = 35$.

p^*	$d_1(\lambda_1^0)$	$d_2(\lambda_2^0)$	the probability of choosing the wrong hypothesis		
			$\phi^*_{\lambda_1^0,\lambda_2^0}$	$\phi^*_{\lambda_1^{n_1},\lambda_2^{n_1}}$	$\phi^*_{\hat{\lambda}_1,\hat{\lambda}_2}$
0	10.049	0.000	0.021	0.008	0.014
1/6	7.290	0.333	0.085	0.072	0.049
1/3	4.930	1.319	0.249	0.172	0.132
1/2	2.980	2.944	0.515	0.340	0.285
2/3	1.466	5.209	0.211	0.429	0.497
5/6	0.433	8.141	0.049	0.190	0.264
1	0.000	11.841	0.004	0.049	0.116

Table 5: $\gamma_{\lambda_2^0}$, $\gamma_{\hat{\lambda}_2}$, $\mathrm{ARL}_{0,\lambda_2^0}$, and $\mathrm{ARL}_{0,\hat{\lambda}_2}$ for the best $10\%, 30\%, 50\%, 70\%, 90\%$ of these 100 000 i.i.d. experiments in a decreasing order by $d(F_{2,\lambda_2^0}, F_{2,\hat{\lambda}_2})$ with $\lambda_2^0 = (-1.450, -2.450, 1.273, -0.109, 0.565)^T$, $T = 300$, and $n_1 = \ldots = n_T = 35$ when $F = F_{1,\lambda_1^0}$ with $\lambda_1^0 = (7, 2, 1)^T$.

	10%	30%	50%	70%	90%
$\gamma_{\lambda_2^0}$	$\gamma_{\hat{\lambda}_2}$	$\gamma_{\hat{\lambda}_2}$	$\gamma_{\hat{\lambda}_2}$	$\gamma_{\hat{\lambda}_2}$	$\gamma_{\hat{\lambda}_2}$
0.0005	0.0003	0.001	0.001	0.0003	0.001

	10%	30%	50%	70%	90%
$\mathrm{ARL}_{0,\lambda_2^0}$	$\mathrm{ARL}_{0,\hat{\lambda}_2}$	$\mathrm{ARL}_{0,\hat{\lambda}_2}$	$\mathrm{ARL}_{0,\hat{\lambda}_2}$	$\mathrm{ARL}_{0,\hat{\lambda}_2}$	$\mathrm{ARL}_{0,\hat{\lambda}_2}$
2073.867	2921.422	1280.703	1168.032	3826.315	1000.012

Table 6: $P_{out,\lambda_2^0,\tilde{\lambda}_1}$, $P_{out,\hat{\lambda}_2,\tilde{\lambda}_1}$, $\mathrm{ARL}_{1,\lambda_2^0,\tilde{\lambda}_1}$, and $\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{\lambda}_1}$ for the best $10\%, 30\%, 50\%, 70\%, 90\%$ of these 100 000 i.i.d. experiments in a decreasing order by the $d(F_{2,\lambda_2^0}, F_{2,\hat{\lambda}_2})$ with $\lambda_2^0 = (-1.450, -2.450, 1.273, -0.109, 0.565)^T$, $T = 300$, and $n_1 = \ldots = n_T = 35$ when $F = F_{1,\tilde{\lambda}_1}$ with $\tilde{\lambda}_1 = (5, 3, 2)^T$.

	10%	30%	50%	70%	90%
$P_{out,\lambda_2^0,\tilde{\lambda}_1}$	$P_{out,\hat{\lambda}_2,\tilde{\lambda}_1}$	$P_{out,\hat{\lambda}_2,\tilde{\lambda}_1}$	$P_{out,\hat{\lambda}_2,\tilde{\lambda}_1}$	$P_{out,\hat{\lambda}_2,\tilde{\lambda}_1}$	$P_{out,\hat{\lambda}_2,\tilde{\lambda}_1}$
0.017	0.012	0.024	0.024	0.009	0.027

	10%	30%	50%	70%	90%
$\mathrm{ARL}_{1,\lambda_2^0,\tilde{\lambda}_1}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{\lambda}_1}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{\lambda}_1}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{\lambda}_1}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{\lambda}_1}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{\lambda}_1}$
60.267	80.561	42.014	41.456	115.074	36.492

Table 7: $\gamma_{\lambda_1^0}$, $\gamma_{\hat{\lambda}_1}$, $\mathrm{ARL}_{0,\lambda_1^0}$, and $\mathrm{ARL}_{0,\hat{\lambda}_1}$ for the best $10\%, 30\%, 50\%, 70\%, 90\%$ of these 100 000 i.i.d. experiments in a decreasing order by the $d(F_{1,\lambda_1^0}, F_{1,\hat{\lambda}_1})$ with $\lambda_1^0 = (5.771, 1.707, 0.884)^T$, $T = 300$, and $n_1 = \ldots = n_T = 35$ when $F = F_{1,\lambda_2^0}$ with $\lambda_2^0 = (-1.450, -2.450, 1.273, -0.109, 0.565)^T$.

	10%	30%	50%	70%	90%
$\gamma_{\lambda_1^0}$	$\gamma_{\hat{\lambda}_1}$	$\gamma_{\hat{\lambda}_1}$	$\gamma_{\hat{\lambda}_1}$	$\gamma_{\hat{\lambda}_1}$	$\gamma_{\hat{\lambda}_1}$
0.010	0.009	0.016	0.015	0.012	0.012

	10%	30%	50%	70%	90%
$\mathrm{ARL}_{0,\lambda_1^0}$	$\mathrm{ARL}_{0,\hat{\lambda}_1}$	$\mathrm{ARL}_{0,\hat{\lambda}_1}$	$\mathrm{ARL}_{0,\hat{\lambda}_1}$	$\mathrm{ARL}_{0,\hat{\lambda}_1}$	$\mathrm{ARL}_{0,\hat{\lambda}_1}$
97.912	110.413	62.723	68.348	86.875	83.515

Table 8: $\gamma_{\lambda_2^0}$, $\gamma_{\hat{\lambda}_2}$, $\mathrm{ARL}_{0,\lambda_2^0}$, and $\mathrm{ARL}_{0,\hat{\lambda}_2}$ for the best $10\%, 30\%, 50\%, 70\%, 90\%$ of these 100 000 i.i.d. experiments in a decreasing order by the $d(F_{2,\lambda_2^0}, F_{2,\hat{\lambda}_2})$ with $T = 300$ and $n_1 = \ldots = n_T = 35$ when $F = p^* \cdot F_{1,\lambda_1^*} + (1 - p^*) \cdot F_{2,\lambda_2^*}$ for $p^* \in \{1/6, 1/3, 1/2, 2/3, 5/6\}$ with $\lambda_1^* = (7, 2, 1)$ and $\lambda_2^* = (-1.450, -2.450, 1.273, -0.109, 0.565)^T$.

		10%	30%	50%	70%	90%
p^*	$\gamma_{\lambda_2^0}$	$\gamma_{\hat{\lambda}_2}$	$\gamma_{\hat{\lambda}_2}$	$\gamma_{\hat{\lambda}_2}$	$\gamma_{\hat{\lambda}_2}$	$\gamma_{\hat{\lambda}_2}$
1/6	0.002	0.002	0.002	0.002	0.002	0.002
1/3	0.002	0.002	0.003	0.002	0.002	0.003
1/2	0.002	0.001	0.001	0.001	0.003	0.001
2/3	0.001	0.001	0.001	0.001	0.002	0.002
5/6	0.001	0.001	0.001	0.001	0.001	0.001

		10%	30%	50%	70%	90%
p^*	$\mathrm{ARL}_{0,\lambda_2^0}$	$\mathrm{ARL}_{0,\hat{\lambda}_2}$	$\mathrm{ARL}_{0,\hat{\lambda}_2}$	$\mathrm{ARL}_{0,\hat{\lambda}_2}$	$\mathrm{ARL}_{0,\hat{\lambda}_2}$	$\mathrm{ARL}_{0,\hat{\lambda}_2}$
1/6	429.149	458.667	452.298	617.770	428.381	668.522
1/3	510.049	572.669	309.477	407.377	405.927	369.132
1/2	628.538	714.295	1096.739	828.135	316.158	681.330
2/3	818.737	1333.622	848.331	734.249	554.763	617.768
5/6	1173.996	1013.322	1345.436	1156.539	1151.260	689.321

Table 9: $\gamma_{\lambda_1^0}$, $\gamma_{\hat{\lambda}_1}$, $\mathrm{ARL}_{0,\lambda_1^0}$, and $\mathrm{ARL}_{0,\hat{\lambda}_1}$ for the best $10\%, 30\%, 50\%, 70\%, 90\%$ of these 100 000 i.i.d. experiments in a decreasing order by the $d(F_{1,\lambda_1^0}, F_{1,\hat{\lambda}_1})$ with $T = 300$ and $n_1 = \ldots = n_T = 35$ when $F = p^* \cdot F_{1,\lambda_1^*} + (1 - p^*) \cdot F_{2,\lambda_2^*}$ for $p^* \in \{1/6, 1/3, 1/2, 2/3, 5/6\}$ with $\lambda_1^* = (7, 2, 1)$ and $\lambda_2^* = (-1.450, -2.450, 1.273, -0.109, 0.565)^T$.

		10%	30%	50%	70%	90%
p^*	$\gamma_{\lambda_1^0}$	$\gamma_{\hat{\lambda}_1}$	$\gamma_{\hat{\lambda}_1}$	$\gamma_{\hat{\lambda}_1}$	$\gamma_{\hat{\lambda}_1}$	$\gamma_{\hat{\lambda}_1}$
1/6	0.010	0.013	0.011	0.007	0.008	0.008
1/3	0.008	0.010	0.009	0.009	0.008	0.005
1/2	0.007	0.013	0.008	0.006	0.006	0.006
2/3	0.006	0.006	0.006	0.006	0.005	0.009
5/6	0.004	0.006	0.003	0.009	0.004	0.004

		10%	30%	50%	70%	90%
p^*	$\mathrm{ARL}_{0,\lambda_1^0}$	$\mathrm{ARL}_{0,\hat{\lambda}_1}$	$\mathrm{ARL}_{0,\hat{\lambda}_1}$	$\mathrm{ARL}_{0,\hat{\lambda}_1}$	$\mathrm{ARL}_{0,\hat{\lambda}_1}$	$\mathrm{ARL}_{0,\hat{\lambda}_1}$
1/6	104.101	78.989	93.531	136.475	120.389	119.532
1/3	118.335	104.301	107.397	113.588	133.027	187.852
1/2	133.595	78.626	132.239	170.435	179.943	181.045
2/3	169.300	163.799	166.018	178.248	195.913	117.159
5/6	226.574	179.740	330.658	113.433	268.169	222.592

Table 10: $P_{out,\lambda_2^0,\tilde{p}}$ and $P_{out,\hat{\lambda}_2,\tilde{p}}$ for the best $10\%, 30\%, 50\%, 70\%, 90\%$ of these 100 000 i.i.d. experiments in a decreasing order by the $d(F_{2,\lambda_2^0}, F_{2,\hat{\lambda}_2})$ with $p^* \in \{1/6, 1/3, 1/2, 2/3, 5/6\}$, $T = 300$, and $n_1 = \ldots = n_T = 35$ when $F = \tilde{p} \cdot F_{1,\tilde{\lambda}_1} + (1-\tilde{p}) \cdot F_{2,\tilde{\lambda}_2}$ for $\tilde{p} \in \{1/6, 1/3, 1/2, 2/3, 5/6\}$ with $\tilde{\lambda}_1 = (5, 3, 2)^T$ and $\tilde{\lambda}_2 = (-0.583, -1.083, 1.787, -0.456, 1.271)^T$.

		10%	30%	50%	70%	90%
p^*	$P_{out,\lambda_2^0,\tilde{p}=1/6}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/6}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/6}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/6}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/6}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/6}$
1/6	0.018	0.019	0.016	0.010	0.016	0.008
1/3	0.018	0.017	0.032	0.017	0.024	0.027
1/2	0.018	0.015	0.010	0.014	0.038	0.010
2/3	0.018	0.010	0.017	0.019	0.024	0.014
5/6	0.018	0.022	0.015	0.019	0.014	0.029

		10%	30%	50%	70%	90%
p^*	$P_{out,\lambda_2^0,\tilde{p}=1/3}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/3}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/3}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/3}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/3}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/3}$
1/6	0.018	0.019	0.016	0.009	0.016	0.008
1/3	0.018	0.017	0.032	0.017	0.024	0.026
1/2	0.018	0.015	0.010	0.014	0.037	0.010
2/3	0.018	0.010	0.017	0.018	0.024	0.014
5/6	0.018	0.022	0.015	0.018	0.014	0.029

		10%	30%	50%	70%	90%
p^*	$P_{out,\lambda_2^0,\tilde{p}=1/2}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/2}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/2}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/2}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/2}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/2}$
1/6	0.017	0.018	0.016	0.009	0.016	0.008
1/3	0.017	0.017	0.032	0.017	0.024	0.026
1/2	0.017	0.015	0.010	0.014	0.037	0.010
2/3	0.017	0.010	0.017	0.018	0.024	0.014
5/6	0.017	0.022	0.014	0.018	0.013	0.028

		10%	30%	50%	70%	90%
p^*	$P_{out,\lambda_2^0,\tilde{p}=2/3}$	$P_{out,\hat{\lambda}_2,\tilde{p}=2/3}$	$P_{out,\hat{\lambda}_2,\tilde{p}=2/3}$	$P_{out,\hat{\lambda}_2,\tilde{p}=2/3}$	$P_{out,\hat{\lambda}_2,\tilde{p}=2/3}$	$P_{out,\hat{\lambda}_2,\tilde{p}=2/3}$
1/6	0.017	0.018	0.015	0.009	0.016	0.007
1/3	0.017	0.016	0.031	0.017	0.024	0.026
1/2	0.017	0.015	0.009	0.014	0.037	0.009
2/3	0.017	0.010	0.016	0.018	0.024	0.014
5/6	0.017	0.022	0.014	0.018	0.013	0.028

		10%	30%	50%	70%	90%
p^*	$P_{out,\lambda_2^0,\tilde{p}=5/6}$	$P_{out,\hat{\lambda}_2,\tilde{p}=5/6}$	$P_{out,\hat{\lambda}_2,\tilde{p}=5/6}$	$P_{out,\hat{\lambda}_2,\tilde{p}=5/6}$	$P_{out,\hat{\lambda}_2,\tilde{p}=5/6}$	$P_{out,\hat{\lambda}_2,\tilde{p}=5/6}$
1/6	0.017	0.018	0.015	0.009	0.015	0.007
1/3	0.017	0.016	0.031	0.016	0.023	0.026
1/2	0.017	0.015	0.009	0.014	0.036	0.009
2/3	0.017	0.010	0.016	0.017	0.023	0.013
5/6	0.017	0.022	0.014	0.018	0.013	0.028

Table 11: $\mathrm{ARL}_{1,\lambda_2^0,\tilde{p}}$ and $\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}}$ for the best $10\%, 30\%, 50\%, 70\%, 90\%$ of these 100 000 i.i.d. experiments in a decreasing order by the $d(F_{2,\lambda_2^0}, F_{2,\hat{\lambda}_2})$ with $p^* \in \{1/6, 1/3, 1/2, 2/3, 5/6\}$, $T = 300$, and $n_1 = \ldots = n_T = 35$ when $F = \tilde{p} \cdot F_{1,\tilde{\lambda}_1} + (1-\tilde{p}) \cdot F_{2,\tilde{\lambda}_2}$ for $\tilde{p} \in \{1/6, 1/3, 1/2, 2/3, 5/6\}$ with $\tilde{\lambda}_1 = (5,3,2)^T$ and $\tilde{\lambda}_2 = (-0.583, -1.083, 1.787, -0.456, 1.271)^T$.

		10%	30%	50%	70%	90%
p^*	$\mathrm{ARL}_{1,\lambda_2^0,\tilde{p}=1/6}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/6}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/6}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/6}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/6}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/6}$
1/6	56.099	53.492	63.070	104.086	61.277	129.189
1/3	56.099	58.840	30.879	57.219	41.165	37.370
1/2	56.099	64.748	102.242	71.053	26.646	99.071
2/3	56.099	98.407	58.573	53.786	41.032	69.115
5/6	56.099	44.761	67.726	53.632	72.064	34.400

		10%	30%	50%	70%	90%
p^*	$\mathrm{ARL}_{1,\lambda_2^0,\tilde{p}=1/3}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/3}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/3}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/3}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/3}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/3}$
1/6	56.886	54.063	63.773	105.968	62.250	131.093
1/3	56.886	59.566	31.224	58.061	41.538	37.813
1/2	56.886	65.411	103.591	71.634	26.864	101.090
2/3	56.886	99.995	59.329	54.676	41.509	70.506
5/6	56.886	45.038	68.392	54.315	73.222	34.747

		10%	30%	50%	70%	90%
p^*	$\mathrm{ARL}_{1,\lambda_2^0,\tilde{p}=1/2}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/2}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/2}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/2}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/2}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/2}$
1/6	57.695	54.646	64.492	107.918	63.255	133.053
1/3	57.695	60.310	31.577	58.927	41.919	38.266
1/2	57.695	66.087	104.976	72.224	27.086	103.192
2/3	57.695	101.634	60.105	55.596	41.997	71.954
5/6	57.695	45.318	69.072	55.017	74.418	35.100

		10%	30%	50%	70%	90%
p^*	$\mathrm{ARL}_{1,\lambda_2^0,\tilde{p}=2/3}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=2/3}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=2/3}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=2/3}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=2/3}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=2/3}$
1/6	58.528	55.242	65.227	109.942	64.292	135.073
1/3	58.528	61.072	31.938	59.820	42.307	38.730
1/2	58.528	66.778	106.399	72.825	27.313	105.383
2/3	58.528	103.328	60.901	56.547	42.497	73.463
5/6	58.528	45.602	69.765	55.736	75.654	35.461

		10%	30%	50%	70%	90%
p	$\mathrm{ARL}_{1,\lambda_2^0,\tilde{p}=5/6}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=5/6}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=5/6}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=5/6}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=5/6}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=5/6}$
1/6	59.385	55.852	65.979	112.044	65.364	137.155
1/3	59.385	61.855	32.307	60.740	42.702	39.206
1/2	59.385	67.484	107.860	73.435	27.542	107.670
2/3	59.385	105.080	61.718	57.531	43.009	75.036
5/6	59.385	45.889	70.471	56.475	76.931	35.829

Table 12: $\gamma_{\lambda_2^0}$, $\gamma_{\hat{\lambda}_2}$, $\text{ARL}_{0,\lambda_2^0}$, and $\text{ARL}_{0,\hat{\lambda}_2}$ for the best $10\%, 30\%, 50\%, 70\%, 90\%$ of these 100 000 i.i.d. experiments in a decreasing order by the $d(F_{2,\lambda_2^0}, F_{2,\hat{\lambda}_2})$ with $T = 300$ and $n_1 = \ldots = n_T = 35$ when $F = p^* \cdot F_{1,\lambda_1^*} + (1 - p^*) \cdot F_{2,\lambda_2^*}$ for $p^* \in \{1/6, 1/3, 1/2, 2/3, 5/6\}$ with $\lambda_1^* = (5.771, 1.707, 0.884)$ and $\lambda_2^* = (-1.450, -2.450, 1.273, -0.109, 0.565)^T$.

		10%	30%	50%	70%	90%
p^*	$\gamma_{\lambda_2^0}$	$\gamma_{\hat{\lambda}_2}$	$\gamma_{\hat{\lambda}_2}$	$\gamma_{\hat{\lambda}_2}$	$\gamma_{\hat{\lambda}_2}$	$\gamma_{\hat{\lambda}_2}$
1/6	0.002	0.002	0.002	0.004	0.004	0.001
1/3	0.002	0.002	0.002	0.001	0.001	0.004
1/2	0.001	0.001	0.002	0.001	0.002	0.003
2/3	0.001	0.001	0.001	0.0004	0.001	0.002
5/6	0.001	0.001	0.001	0.001	0.002	0.001

		10%	30%	50%	70%	90%
p^*	$\text{ARL}_{0,\lambda_2^0}$	$\text{ARL}_{0,\hat{\lambda}_2}$	$\text{ARL}_{0,\hat{\lambda}_2}$	$\text{ARL}_{0,\hat{\lambda}_2}$	$\text{ARL}_{0,\hat{\lambda}_2}$	$\text{ARL}_{0,\hat{\lambda}_2}$
1/6	470.175	595.580	443.966	276.241	239.030	769.612
1/3	592.735	416.511	536.236	846.050	970.737	224.911
1/2	756.675	813.934	649.492	856.887	508.406	305.835
2/3	955.090	959.679	1040.847	2301.318	1769.913	426.967
5/6	1447.122	1089.188	694.365	976.435	504.247	831.356

Table 13: $\gamma_{\lambda_1^0}$, $\gamma_{\hat{\lambda}_1}$, $\text{ARL}_{0,\lambda_1^0}$, and $\text{ARL}_{0,\hat{\lambda}_1}$ for the best $10\%, 30\%, 50\%, 70\%, 90\%$ of these 100 000 i.i.d. experiments in a decreasing order by the $d(F_{1,\lambda_1^0}, F_{1,\hat{\lambda}_1})$ with $T = 300$ and $n_1 = \ldots = n_T = 35$ when $F = p^* \cdot F_{1,\lambda_1^*} + (1 - p^*) \cdot F_{2,\lambda_2^*}$ for $p^* \in \{1/6, 1/3, 1/2, 2/3, 5/6\}$ with $\lambda_1^* = (5.771, 1.707, 0.884)$ and $\lambda_2^* = (-1.450, -2.450, 1.273, -0.109, 0.565)^T$.

		10%	30%	50%	70%	90%
p^*	$\gamma_{\lambda_1^0}$	$\gamma_{\hat{\lambda}_1}$	$\gamma_{\hat{\lambda}_1}$	$\gamma_{\hat{\lambda}_1}$	$\gamma_{\hat{\lambda}_1}$	$\gamma_{\hat{\lambda}_1}$
1/6	0.009	0.012	0.008	0.008	0.011	0.01
1/3	0.008	0.014	0.009	0.008	0.006	0.005
1/2	0.006	0.011	0.008	0.008	0.006	0.006
2/3	0.005	0.005	0.010	0.007	0.004	0.004
5/6	0.004	0.006	0.005	0.004	0.004	0.003

		10%	30%	50%	70%	90%
p^*	$\text{ARL}_{0,\lambda_1^0}$	$\text{ARL}_{0,\hat{\lambda}_1}$	$\text{ARL}_{0,\hat{\lambda}_1}$	$\text{ARL}_{0,\hat{\lambda}_1}$	$\text{ARL}_{0,\hat{\lambda}_1}$	$\text{ARL}_{0,\hat{\lambda}_1}$
1/6	111.595	86.314	131.813	117.975	92.973	86.373
1/3	129.723	69.299	106.714	127.448	155.697	182.420
1/2	154.883	87.955	133.345	128.392	168.520	162.005
2/3	192.150	208.536	102.868	138.559	232.638	244.044
5/6	253.034	164.246	188.869	259.491	274.559	313.108

Table 14: $P_{out,\lambda_2^0,\tilde{p}}$ and $P_{out,\hat{\lambda}_2,\tilde{p}}$ for the best $10\%, 30\%, 50\%, 70\%, 90\%$ of these 100 000 i.i.d. experiments in a decreasing order by the $d(F_{2,\lambda_2^0}, F_{2,\hat{\lambda}_2})$ with $p^* \in \{1/6, 1/3, 1/2, 2/3, 5/6\}$, $T = 300$, and $n_1 = \ldots = n_T = 35$ when $F = \tilde{p} \cdot F_{1,\tilde{\lambda}_1} + (1-\tilde{p}) \cdot F_{2,\tilde{\lambda}_2}$ for $\tilde{p} \in \{1/6, 1/3, 1/2, 2/3, 5/6\}$ with $\tilde{\lambda}_1 = (5, 3, 2)^T$ and $\tilde{\lambda}_2 = (-0.583, -1.083, 1.787, -0.456, 1.271)^T$.

		10%	30%	50%	70%	90%
p^*	$P_{out,\lambda_2^0,\tilde{p}=1/6}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/6}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/6}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/6}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/6}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/6}$
1/6	0.017	0.013	0.016	0.029	0.020	0.014
1/3	0.016	0.018	0.015	0.013	0.013	0.033
1/2	0.013	0.014	0.010	0.006	0.007	0.027
2/3	0.014	0.014	0.017	0.008	0.023	0.022
5/6	0.010	0.013	0.019	0.014	0.025	0.018

		10%	30%	50%	70%	90%
p^*	$P_{out,\lambda_2^0,\tilde{p}=1/3}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/3}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/3}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/3}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/3}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/3}$
1/6	0.017	0.013	0.016	0.029	0.019	0.014
1/3	0.016	0.018	0.015	0.013	0.013	0.033
1/2	0.013	0.014	0.010	0.006	0.007	0.027
2/3	0.014	0.013	0.017	0.008	0.022	0.022
5/6	0.010	0.013	0.019	0.014	0.025	0.018

		10%	30%	50%	70%	90%
p^*	$P_{out,\lambda_2^0,\tilde{p}=1/2}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/2}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/2}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/2}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/2}$	$P_{out,\hat{\lambda}_2,\tilde{p}=1/2}$
1/6	0.016	0.013	0.016	0.028	0.019	0.013
1/3	0.016	0.018	0.015	0.013	0.013	0.033
1/2	0.013	0.014	0.010	0.006	0.007	0.027
1/3	0.014	0.013	0.016	0.008	0.022	0.021
5/6	0.010	0.013	0.019	0.014	0.025	0.018

		10%	30%	50%	70%	90%
p^*	$P_{out,\lambda_2^0,\tilde{p}=2/3}$	$P_{out,\hat{\lambda}_2,\tilde{p}=2/3}$	$P_{out,\hat{\lambda}_2,\tilde{p}=2/3}$	$P_{out,\hat{\lambda}_2,\tilde{p}=2/3}$	$P_{out,\hat{\lambda}_2,\tilde{p}=2/3}$	$P_{out,\hat{\lambda}_2,\tilde{p}=2/3}$
1/6	0.016	0.013	0.016	0.028	0.019	0.013
1/3	0.015	0.017	0.015	0.013	0.013	0.032
1/2	0.013	0.014	0.010	0.006	0.007	0.027
2/3	0.014	0.013	0.016	0.008	0.022	0.021
5/6	0.009	0.013	0.019	0.014	0.025	0.018

		10%	30%	50%	70%	90%
p^*	$P_{out,\lambda_2^0,\tilde{p}=5/6}$	$P_{out,\hat{\lambda}_2,\tilde{p}=5/6}$	$P_{out,\hat{\lambda}_2,\tilde{p}=5/6}$	$P_{out,\hat{\lambda}_2,\tilde{p}=5/6}$	$P_{out,\hat{\lambda}_2,\tilde{p}=5/6}$	$P_{out,\hat{\lambda}_2,\tilde{p}=5/6}$
1/6	0.016	0.013	0.016	0.028	0.018	0.013
1/3	0.015	0.017	0.015	0.013	0.013	0.032
1/2	0.013	0.014	0.010	0.006	0.007	0.026
2/3	0.014	0.013	0.016	0.008	0.022	0.021
5/6	0.009	0.012	0.019	0.014	0.024	0.018

Table 15: $\mathrm{ARL}_{1,\lambda_2^0,\tilde{p}}$ and $\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}}$ for the best $10\%, 30\%, 50\%, 70\%, 90\%$ of these 100 000 i.i.d. experiments in a decreasing order by the $d(F_{2,\lambda_2^0}, F_{2,\hat{\lambda}_2})$ with $p^* \in \{1/6, 1/3, 1/2, 2/3, 5/6\}$, $T = 300$, and $n_1 = \ldots = n_T = 35$ when $F = \tilde{p} \cdot F_{1,\tilde{\lambda}_1} + (1 - \tilde{p}) \cdot F_{2,\tilde{\lambda}_2}$ for $\tilde{p} \in \{1/6, 1/3, 1/2, 2/3, 5/6\}$ with $\tilde{\lambda}_1 = (5, 3, 2)^T$ and $\tilde{\lambda}_2 = (-0.583, -1.083, 1.787, -0.456, 1.271)^T$.

		10%	30%	50%	70%	90%
p^*	$\mathrm{ARL}_{1,\lambda_2^0,\tilde{p}=1/6}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/6}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/6}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/6}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/6}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/6}$
1/6	59.411	76.442	60.680	34.507	50.469	73.589
1/3	62.966	54.730	65.855	75.563	76.964	29.862
1/2	75.744	72.086	98.724	63.165	140.586	36.486
2/3	69.842	74.069	59.545	126.069	44.370	45.000
5/6	102.325	78.660	52.349	71.022	39.992	55.137

		10%	30%	50%	70%	90%
p^*	$\mathrm{ARL}_{1,\lambda_2^0,\tilde{p}=1/3}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/3}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/3}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/3}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/3}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/3}$
1/6	60.055	77.196	61.391	34.947	51.613	73.996
1/3	63.508	55.608	66.462	75.772	77.100	30.261
1/2	76.058	72.273	100.223	163.352	140.832	36.807
2/3	70.247	74.443	60.108	127.651	44.570	45.864
5/6	103.520	79.045	52.758	71.581	40.242	55.116

		10%	30%	50%	70%	90%
p^*	$\mathrm{ARL}_{1,\lambda_2^0,\tilde{p}=1/2}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/2}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/2}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/2}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/2}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=1/2}$
1/6	60.714	77.964	62.118	35.397	52.812	74.408
1/3	64.059	56.515	67.079	75.983	77.236	30.670
1/2	76.375	72.460	101.768	163.539	141.079	37.134
2/3	70.657	74.821	60.681	129.273	44.772	46.762
5/6	104.744	79.433	53.173	72.149	40.496	55.096

		10%	30%	50%	70%	90%
p^*	$\mathrm{ARL}_{1,\lambda_2^0,\tilde{p}=2/3}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=2/3}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=2/3}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=2/3}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=2/3}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=2/3}$
1/6	61.387	78.748	62.863	35.859	54.067	74.824
1/3	64.620	57.452	67.708	76.194	77.373	31.091
1/2	76.694	72.648	103.361	163.726	141.326	37.467
2/3	71.072	75.202	61.265	130.937	44.976	47.696
5/6	105.996	79.826	53.595	72.726	40.753	55.076

		10%	30%	50%	70%	90%
p^*	$\mathrm{ARL}_{1,\lambda_2^0,\tilde{p}=5/6}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=5/6}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=5/6}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=5/6}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=5/6}$	$\mathrm{ARL}_{1,\hat{\lambda}_2,\tilde{p}=5/6}$
1/6	62.075	79.548	63.626	36.334	55.383	75.245
1/3	65.191	58.420	68.349	76.407	77.511	31.524
1/2	77.016	72.838	105.006	163.913	141.575	37.805
2/3	71.492	75.588	61.861	132.645	45.182	48.668
5/6	107.279	80.222	54.023	73.312	41.014	55.056