ProbStat Forum, Volume 11, April 2018, Pages 53–66
ISSN 0974-3235

ProbStat Forum is an e-journal.
For details please visit www.probstat.org.in

A family of bivariate discrete distributions on $\mathbb{Z}^2$ based on the Rademacher distribution

Christophe Chesneau[a], Maher Kachour[b], Hassan S. Bakouch[c]

[a]*Laboratoire de Mathématiques Nicolas Oresme (LMNO), Université de Caen Nomandie.*
[b]*Ecole supérieure de commerce IDRAC Lion.*
[c]*Department of Mathematics, Faculty of Science, Tanta University, Tanta, Egypt.*

Abstract. In this paper we introduce a new family of bivariate discrete distributions on $\mathbb{Z}^2$, called the $\text{Rad}(\alpha_1, \alpha_2, \theta)$ class. This new family can be considered as an extension (on $\mathbb{Z}^2$) of some standard bivariate discrete (non negative valued) distributions and it is used to generate bivariate random variables with possible negative or positive values for the covariance. Some properties of the family are investigated. Also, a statistical approach is described to estimate the unknown parameters of the family. Finally, we apply this general family of distributions to the bivariate Poisson and geometric distributions with simulation studies.

1. Introduction

Paired count observations appear in many situations and research fields. For example, in insurance, one can be interested in studying the evolution of the number of accidents in a site before and after some new regulations. Bivariate discrete distributions were introduced to fit such data. Over the last three decades, a large contribution on discrete bivariate distributions has been accumulated. For a literature review, we refer to (10), (8), (11), (14), (1), (12), (3), and (15).

In contrast to the well-known situation when the paired data are counts, i.e. observed on N^2, sometimes the data are observed on Z^2. For example, when analyzing intra-daily stock prices the changes take both positive and negative integer values, known also as ticks (the price can go up or down on certain predefined ranges of value). The price change is therefore characterized by discrete jumps. To the best of our knowledge there is a shortage of bivariate discrete distribution defined on Z^2. See for instance (4) who constructed bivariate Skellam distributions by a simple trivariate reduction scheme with additive functions. In a recent article, (5) introduced a family of distributions based on the generalized trivariate (multivariate) reduction technique and the Rademacher distribubtion. In this paper, we intend to contribute to this literature, by presenting a new family of bivariate discrete distributions, based also on the Rademacher distribubtion. This new family, denoted by $\text{Rad}(\alpha_1, \alpha_2, \theta)$, is defined on Z^2 and with possible negative or positive values for the covariance. Furthermore, it can be considered as an extension on Z^2 of some standard bivariate distributions such as the Poisson or the geometric.

The remainder of the paper proceeds as follows. Firstly, in Section 2, we present the considered $\text{Rad}(\alpha_1, \alpha_2, \theta)$ family using a couple of independent Rademacher random variables and a bivariate random

2010 *Mathematics Subject Classification.* 62G30, 62E10.
Keywords. Bivariate Geometric distribution ; Rademacher random variable ; Maximum likelihood estimator; Method of moments; Correlation.
Received: 09 October 2017; Revised: 05 April 2018, Re-revised: 05 April 2018, Accepted: 27 April 2018.
Email addresses: christophe.chesneau@unicaen.fr (Christophe Chesneau), maher.kachour@idraclyon.com (Maher Kachour), hassan.bakouch@science.tanta.edu.eg (Hassan S. Bakouch)

variable with support in $\mathbb{N}^2$, and investigate some of its properties. Section 3 describes statistical approach to estimate the unknown parameters efficiently. Finally, in Section 4, we apply our general distributions to the bivariate Poisson and Geometric distributions.

2. The $\mathbf{Rad}(\boldsymbol{\alpha_1, \alpha_2, \theta})$ class

First we present the (generalized) Rademacher distribution. We say a random variable R follows the (generalized) Rademacher distribution with parameter $\alpha \in (0,1)$ denoted by $\mathcal{R}(\alpha)$ if and only if

$$\mathbb{P}(R = -1) = 1 - \alpha, \quad \mathbb{P}(R = 1) = \alpha.$$

We adopt the notation $R \sim \mathcal{R}(\alpha)$. This distribution is symmetric (around 0) for the special case $\alpha = 1/2$. It has many well-known applications in bootstrapping, symmetrization of random variables and random walk theory. Let us now introduce a new class of distribution in $\mathbb{Z}^2$, denoted by $\mathrm{Rad}(\alpha_1, \alpha_2, \theta)$ and investigate some of its basic properties.

Definition 2.1. *Let $(\alpha_1, \alpha_2) \in (0,1)^2$, (X, Y) be a bivariate random variables with $(X, Y)(\Omega) \subseteq \mathbb{N}^2$, depending on a parameter or a vector of parameters denoted by θ, and, for any $j \in \{1, 2\}$, $R_j \sim \mathcal{R}(\alpha_j)$. Let us suppose that R_1 and R_2 are mutually independent and independent of X and Y. We say that a bivariate random variable (S, T) belongs to the Rad$(\alpha_1, \alpha_2, \theta)$ class, if and only if*

$$(S, T) = (R_1 X, R_2 Y).$$

Remark 2.2. *The Rad$(\alpha_1, \alpha_2, \theta)$ class can be interpreted as the extension to $\mathbb{Z}^2$ of the class of bivariate discrete distributions defined on $\mathbb{N}^2$.*

Let us now introduce some basic properties of (S, T):

- The support of (S, T) is given by $(S, T)(\Omega) = \{(s, t) \in \mathbb{Z}^2;\ (|s|, |t|) \in (X, Y)(\Omega)\}$.

- The distribution of (S, T) is given by

$$\begin{aligned} &\mathbb{P}(\{S = s\} \cap \{T = t\}) = \\ &\quad \alpha_1^{\mathbf{1}_{\{s>0\}}} (1 - \alpha_1)^{\mathbf{1}_{\{s<0\}}} \alpha_2^{\mathbf{1}_{\{t>0\}}} (1 - \alpha_2)^{\mathbf{1}_{\{t<0\}}} \mathbb{P}(\{X = |s|\} \cap \{Y = |t|\}), \\ &\qquad (s, t) \in \mathbb{Z}^2, \end{aligned}$$

  where, for any $A \subseteq \mathbb{R}$, $\mathbf{1}_A = \mathbf{1}_A(x)$ denotes the indicator function on A : $\mathbf{1}_A = 1$ if $x \in A$ and 0 elsewhere. Let us remark that, for $(s, t) = (0, 0)$, the powers satisfy : $\mathbf{1}_{\{s<0\}} = \mathbf{1}_{\{s>0\}} = \mathbf{1}_{\{t<0\}} = \mathbf{1}_{\{t>0\}} = 0$, so $\mathbb{P}(\{S = 0\} \cap \{T = 0\}) = \mathbb{P}(\{X = 0\} \cap \{Y = 0\})$. Let us now briefly show that the sum of the joint probability is 1. By using a decomposition according to the sign of s and t and noticing that $\alpha_1\alpha_2 + (1 - \alpha_1)\alpha_2 + \alpha_1(1 - \alpha_2) + (1 - \alpha_1)(1 - \alpha_2) = 1$, we have:

$$\begin{aligned}
\sum_{(s,t)\in\mathbb{Z}^2} \mathbb{P}(\{S=s\}\cap\{T=t\}) &= \mathbb{P}(\{X=0\}\cap\{Y=0\}) \\
&+ \alpha_2 \sum_{t>0} \mathbb{P}(\{X=0\}\cap\{Y=t\}) + (1-\alpha_2)\sum_{t<0} \mathbb{P}(\{X=0\}\cap\{Y=-t\}) \\
&+ \alpha_1 \sum_{s>0} \mathbb{P}(\{X=s\}\cap\{Y=0\}) + (1-\alpha_1)\sum_{s<0} \mathbb{P}(\{X=-s\}\cap\{Y=0\}) \\
&+ \alpha_1\alpha_2 \sum_{s>0}\sum_{t>0} \mathbb{P}(\{X=s\}\cap\{Y=t\}) \\
&+ (1-\alpha_1)\alpha_2 \sum_{s<0}\sum_{t>0} \mathbb{P}(\{X=-s\}\cap\{Y=t\}) \\
&+ \alpha_1(1-\alpha_2) \sum_{s>0}\sum_{t<0} \mathbb{P}(\{X=s\}\cap\{Y=-t\}) \\
&+ (1-\alpha_1)(1-\alpha_2) \sum_{s<0}\sum_{t<0} \mathbb{P}(\{X=-s\}\cap\{Y=-t\}) \\
&= \mathbb{P}(\{X=0\}\cap\{Y=0\}) + \sum_{t>0} \mathbb{P}(\{X=0\}\cap\{Y=t\}) \\
&+ \sum_{s>0} \mathbb{P}(\{X=s\}\cap\{Y=0\}) + \sum_{s>0}\sum_{t>0} \mathbb{P}(\{X=s\}\cap\{Y=t\}) = 1.
\end{aligned}$$

- The marginal distributions are given by

$$\mathbb{P}(S=s) = \alpha_1^{\mathbf{1}_{\{s>0\}}}(1-\alpha_1)^{\mathbf{1}_{\{s<0\}}}\mathbb{P}(X=|s|), \qquad s\in\mathbb{Z},$$

and

$$\mathbb{P}(T=t) = \alpha_2^{\mathbf{1}_{\{t>0\}}}(1-\alpha_2)^{\mathbf{1}_{\{t<0\}}}\mathbb{P}(Y=|t|), \qquad t\in\mathbb{Z}.$$

- The conditional distributions are given by

$$\mathbb{P}(\{S=s\} \mid \{T=t\}) = \alpha_1^{\mathbf{1}_{\{s>0\}}}(1-\alpha_1)^{\mathbf{1}_{\{s<0\}}}\mathbb{P}(\{X=|s|\} \mid \{Y=|t|\}),$$

$(s,t)\in\mathbb{Z}^2$, and

$$\mathbb{P}(\{T=t\} \mid \{S=s\}) = \alpha_2^{\mathbf{1}_{\{t>0\}}}(1-\alpha_2)^{\mathbf{1}_{\{t<0\}}}\mathbb{P}(\{Y=|t|\} \mid \{X=|s|\}),$$

$(s,t)\in\mathbb{Z}^2$.

- Moreover, since $R_1^n = 1$ if n is even and $R_1^n = R_1$ if n is odd, with the analog for R_2, using the independance between R_1, R_2 and (X,Y), we have

$$\begin{aligned}
\mathbb{E}(S^nT^m) &= \mathbb{E}(R_1^nX^nR_2^mY^m) = \mathbb{E}(R_1^n)\mathbb{E}(R_2^m)\mathbb{E}(X^nY^m) \\
&= \begin{cases} \mathbb{E}(X^nY^m) & \text{if } (n \text{ is even, } m \text{ even}), \\ \mathbb{E}(R_1)\mathbb{E}(X^nY^m) & \text{if } (n \text{ is odd, } m \text{ even}), \\ \mathbb{E}(R_2)\mathbb{E}(X^nY^m) & \text{if } (n \text{ is even, } m \text{ odd}), \\ \mathbb{E}(R_1)\mathbb{E}(R_2)\mathbb{E}(X^nY^m) & \text{if } (n \text{ is odd, } m \text{ odd}), \end{cases} \\
&= \begin{cases} \mathbb{E}(X^nY^m) & \text{if } (n \text{ is even, } m \text{ even}), \\ (2\alpha_1-1)\mathbb{E}(X^nY^m) & \text{if } (n \text{ is odd, } m \text{ even}), \\ (2\alpha_2-1)\mathbb{E}(X^nY^m) & \text{if } (n \text{ is even, } m \text{ odd}), \\ (2\alpha_1-1)(2\alpha_2-1)\mathbb{E}(X^nY^m) & \text{if } (n \text{ is odd, } m \text{ odd}). \end{cases}
\end{aligned}$$

Putting $n = 0$ or $m = 0$, we immediately obtain

$$\mathbb{E}(S^m) = \begin{cases} \mathbb{E}(X^m) & \text{if } m \text{ even}, \\ (2\alpha_1 - 1)\mathbb{E}(X^m) & \text{if } m \text{ odd}, \end{cases}$$

and

$$\mathbb{E}(T^m) = \begin{cases} \mathbb{E}(Y^m) & \text{if } m \text{ even}, \\ (2\alpha_2 - 1)\mathbb{E}(Y^m) & \text{if } m \text{ odd}. \end{cases}$$

- The covariance of S and T is given by

$$\mathbb{C}(S,T) = (2\alpha_1 - 1)(2\alpha_2 - 1)\mathbb{C}(X,Y),$$

where $\mathbb{C}(X,Y)$ denotes the covariance of X and Y.

Remark 2.3. *Note that $\mathbb{C}(S,T)$ can be positive or negative according to the values of (α_1, α_2).*

- The correlation of S and T is given by

$$\rho(S,T) = \frac{(2\alpha_1 - 1)(2\alpha_2 - 1)\mathbb{C}(X,Y)}{\sqrt{\mathbb{E}(X^2) - (2\alpha_1 - 1)^2(\mathbb{E}(X))^2}\sqrt{\mathbb{E}(Y^2) - (2\alpha_2 - 1)^2(\mathbb{E}(Y))^2}}.$$

- The characteristic function of (S,T) is given by

$$\begin{aligned} \phi_*(s,t) &= \mathbb{E}(e^{isS}e^{itT}) \\ &= \alpha_1\alpha_2\phi(s,t) + (1-\alpha_1)\alpha_2\phi(-s,t) + \alpha_1(1-\alpha_2)\phi(s,-t) \\ &+ (1-\alpha_1)(1-\alpha_2)\phi(-s,-t), \qquad (s,t) \in \mathbb{R}^2, \end{aligned}$$

where $\phi(x,y) = \mathbb{E}(e^{ixX}e^{iyY})$ denotes the characteristic function of (X,Y).

3. Parameter estimation

From the general definition of the Rad$(\alpha_1, \alpha_2, \theta)$ class, we are able to determine the Maximum Likelihood estimator (MLE) of the unknown parameters. The details are given below.

Let $\theta = (\theta_1, \ldots, \theta_m)$ be the vector of unknown parameters in (X,Y) and $\upsilon = (\alpha_1, \alpha_2, \theta)$. For any $(s,t) = ((s_1,t_1), \ldots, (s_n,t_n))$ with $(s_i,t_i) \in \mathbb{Z}^2$, the likelihood function associated to (S,T) is given by

$$\begin{aligned} &L_\upsilon(s,t) = \prod_{i=1}^{n} \mathbb{P}(\{S = s_i\} \cap \{T = t_i\}) = \\ &\quad \alpha_1^{\sum_{i=1}^n \mathbf{1}_{\{s_i>0\}}}(1-\alpha_1)^{\sum_{i=1}^n \mathbf{1}_{\{s_i<0\}}}\alpha_2^{\sum_{i=1}^n \mathbf{1}_{\{t_i>0\}}}(1-\alpha_2)^{\sum_{i=1}^n \mathbf{1}_{\{s_i<0\}}}\tilde{L}_\theta(|s|,|t|), \end{aligned}$$

where $\tilde{L}_\theta$ denotes the likelihood function associated to (X,Y). The log-likelihood function is given by

$$\begin{aligned} &\ell_\upsilon(s,t) = \ln L_\upsilon(s,t) = \\ &\qquad \ln(\alpha_1)\sum_{i=1}^{n} \mathbf{1}_{\{s_i>0\}} + \ln(1-\alpha_1)\sum_{i=1}^{n} \mathbf{1}_{\{s_i<0\}} + \ln(\alpha_2)\sum_{i=1}^{n} \mathbf{1}_{\{t_i>0\}} \\ &+ \quad \ln(1-\alpha_2)\sum_{i=1}^{n} \mathbf{1}_{\{s_i<0\}} + \tilde{\ell}_\theta(|s|,|t|), \end{aligned}$$

where $\tilde{\ell}_\theta$ is the log-likelihood function associated to (X,Y).

We have

$$\frac{\partial \ell_v}{\partial \alpha_1}(s,t) = \frac{1}{\alpha_1}\sum_{i=1}^{n} \mathbf{1}_{\{s_i>0\}} - \frac{1}{1-\alpha_1}\sum_{i=1}^{n} \mathbf{1}_{\{s_i<0\}}.$$

Hence

$$\frac{\partial \ell_v}{\partial \alpha_1}(s,t) = 0 \Leftrightarrow \alpha_1 = \frac{\sum_{i=1}^{n} \mathbf{1}_{\{s_i>0\}}}{n - \sum_{i=1}^{n} \mathbf{1}_{\{s_i=0\}}}$$

and this value is a maximum for $\alpha_1 \to \ell_\theta(s,t)$.

Then the MLE of α_1 and α_2 are given by

$$\hat{\alpha}_1 = \frac{\sum_{i=1}^{n} \mathbf{1}_{\{S_i>0\}}}{n - \sum_{i=1}^{n} \mathbf{1}_{\{S_i=0\}}} \tag{1}$$

and

$$\hat{\alpha}_2 = \frac{\sum_{i=1}^{n} \mathbf{1}_{\{T_i>0\}}}{n - \sum_{i=1}^{n} \mathbf{1}_{\{T_i=0\}}}. \tag{2}$$

On the other hand, since $|S| = X$ and $|T| = Y$ and

$$\frac{\partial \ell_v}{\partial \theta_i}(s,t) = \frac{\partial \tilde{\ell}_\theta}{\partial \theta_i}(|s|,|t|),$$

the MLE of θ have the same definitions of those associated to (X,Y) but defined with $(|S_1|,|T_1|),\ldots,(|S_n|,|T_n|)$ instead of $(X_1,Y_1),\ldots,(X_n,Y_n)$.

4. Applications to bivariate case of Poisson and geometric distribution

In this section, we apply our general family of distributions to the bivariate Poisson and geometric distributions. We then obtain two new bivariate distributions on $\mathbb{Z}^2$ having possible negative or positive values for the covariance function.

4.1. Application to the bivariate Poisson distribution

The bivariate Poisson distribution is one of the most famous bivariate discrete distributions. It is define as follows. Let $(\lambda_0,\lambda_1,\lambda_2) \in (0,\infty)^3$ be three parameters and U_0, U_1 and U_2 be three independent random variables such that, for any $j \in \{0,1,2\}$, $U_j \sim \mathcal{P}(\lambda_j)$, i.e.

$$\mathbb{P}(U_j = u) = e^{-\lambda_j}\frac{\lambda_j^u}{u!}, \qquad u \in \mathbb{N}.$$

Let us now consider the two random variables:

$$X = U_1 + U_0, \qquad Y = U_2 + U_0.$$

We say thus that (X,Y) follows a bivariate Poisson distribution. Thus, the probability mass function of (X,Y) is given by

$$\mathbb{P}(\{X=x\}\cap\{Y=y\}) = e^{-(\lambda_0+\lambda_1+\lambda_2)}\frac{\lambda_1^x}{x!}\frac{\lambda_2^y}{y!}\sum_{i=0}^{\min(x,y)}\binom{x}{i}\binom{y}{i} i!\left(\frac{\lambda_0}{\lambda_1\lambda_2}\right)^i, \quad (x,y)\in\mathbb{N}^2.$$

Note that

$$X \sim \mathcal{P}(\lambda_1+\lambda_0), \qquad Y \sim \mathcal{P}(\lambda_2+\lambda_0), \qquad \mathbb{C}(X,Y) = \mathbb{V}(U_0) = \lambda_0.$$

For a comprehensive treatment of the bivariate Poisson distribution and its multivariate extensions see (10), and (8). Set now

$$(S, T) = (R_1 X, R_2 Y), \tag{3}$$

with $R_j \sim \mathcal{R}(\alpha_j)$, for any $j \in \{1, 2\}$. Moreover, we suppose that R_1 and R_2 are mutually independent and independent of X and Y.

Plot of a special case of the joint probability of (S, T) is given by Figure 1. Figures 2 and 3 give bivariate plots that illustrate positive and negative correlations for this distribution.

Let us now investigate the estimation of the unknown parameters $(\alpha_1, \alpha_2, \lambda_0, \lambda_1, \lambda_2)$ from a n-sample $(S_1, T_1), \ldots, (S_n, T_n)$ via two estimation methods: the method of moments and the MLE.

Method of moments. Let $\theta_1 = \lambda_1 + \lambda_0$. Observe that

$$\mathbb{E}(S^2) = \theta_1^2 + \theta_1 \quad \Leftrightarrow \quad \theta_1 = \frac{\sqrt{1 + 4\mathbb{E}(S^2)} - 1}{2}$$

which yields the following estimator:

$$\hat{\theta}_1 = \frac{\sqrt{1 + 4\overline{S^2}} - 1}{2}, \qquad \overline{S^2} = \frac{1}{n} \sum_{i=1}^{n} S_i^2.$$

Since

$$\mathbb{E}(S) = (2\alpha_1 - 1)\theta_1 \quad \Leftrightarrow \quad \alpha_1 = \frac{1}{2} \left(\frac{\mathbb{E}(S)}{\theta_1} + 1 \right)$$

an estimator for α_1 is

$$\hat{\alpha}_1 = \frac{1}{2} \left(\frac{\overline{S}}{\hat{\theta}_1} + 1 \right), \qquad \overline{S} = \frac{1}{n} \sum_{i=1}^{n} S_i.$$

Let $\theta_2 = \lambda_2 + \lambda_0$. Hence, using similar arguments, we have

$$\hat{\theta}_2 = \frac{\sqrt{1 + 4\overline{T^2}} - 1}{2}, \qquad \overline{T^2} = \frac{1}{n} \sum_{i=1}^{n} T_i^2$$

and

$$\hat{\alpha}_2 = \frac{1}{2} \left(\frac{\overline{T}}{\hat{\theta}_2} + 1 \right), \qquad \overline{T} = \frac{1}{n} \sum_{i=1}^{n} T_i.$$

Let us now investigate the estimators of λ_j with $j \in \{0, 1, 2\}$. Observe that

$$\mathbb{P}\left(\{S = 0\} \cap \{T = 0\}\right) = e^{-(\lambda_1 + \lambda_2 + \lambda_0)},$$

$$\mathbb{P}(S = 0) = e^{-(\lambda_1 + \lambda_0)}, \qquad \mathbb{P}(T = 0) = e^{-(\lambda_2 + \lambda_0)}.$$

Therefore natural estimators for λ_1, λ_2 and λ_0 are respectively given by

$$\hat{\lambda}_1 = -\log \left(\frac{\sum_{i=1}^{n} \mathbf{1}_{\{(S_i, T_i) = (0,0)\}}}{\sum_{i=1}^{n} \mathbf{1}_{\{T_i = 0\}}} \right),$$

$$\hat{\lambda}_2 = -\log \left(\frac{\sum_{i=1}^{n} \mathbf{1}_{\{(S_i, T_i) = (0,0)\}}}{\sum_{i=1}^{n} \mathbf{1}_{\{S_i = 0\}}} \right)$$

and

$$\hat{\lambda}_0 = -\log \left(\frac{\sum_{i=1}^{n} \mathbf{1}_{\{(S_i, T_i) = (0,0)\}}}{n e^{-(\hat{\lambda}_1 + \hat{\lambda}_2)}} \right).$$

MLE. The MLE of α_1 (resp. α_2) are given by (1) (resp. (2)). To calculate the MLE of $\lambda = (\lambda_0, \lambda_1, \lambda_2)$ we use the EM-algorithm for the bivariate Poisson models, originally proposed by (9).

Remark 4.1. *Note that, (7) showed the asymptotic property of MLE for the bivariate Poisson distribution.*

Parameter	actual value	Mean (MM)	Sd (MM)	Mean (ML)	Sd (ML)
α_1	0.4	0.3998928	0.01849841	0.3999116	0.01739304
α_2	0.7	0.6994876	0.01677994	0.7001840	0.0145807
λ_0	1	0.9731342	0.18982206	0.9989756	0.06292195
λ_1	1	1.0313487	0.19963600	1.0006694	0.06299054
λ_2	2	2.0355057	0.23531690	1.9981334	0.07110126

Table 1: Parametric estimation of the unknown parameters $(\alpha_1, \alpha_2, \lambda_0, \lambda_1, \lambda_2)$ from the distribution of (3) (using the Poisson distributions) via the MM and MLE.

A short simulation experiment

We simulate $n = 1000$ observations of (S, T) (where (X, Y) follows a bivariate Poisson distribution) and we compute both moment method (MM) and maximum likelihood (ML) estimators. After 1000 independent replications, we calculate the average and the standard deviation of the sequence of the obtained estimates. Results are summarized in Table 1.

One can see from Table 1 that the ML estimation gives a smaller standard deviation than the MM estimation for all parameters. Fitting to a Gaussian distribution is illustrated in Figure 7, for the ML estimator, showing its good asymptotic normality properties.

4.2. Application to the bivariate geometric distribution

Let $(p_0, p_1, p_2) \in (0, 1)^3$ be three parameters and U_0, U_1 and U_2 be three independent random variables such that, for any $j \in \{0, 1, 2\}$, $U_j \sim \mathcal{G}_0(p_j)$, i.e.

$$\mathbb{P}(U_j = u) = p_j q_j^u, \qquad u \in \mathbb{N},$$

where $q_j = 1 - p_j$. Let now

$$X = U_1 + U_0, \qquad Y = U_2 + U_0.$$

Thus, one can say that (X, Y) follows a bivariate Geometric distribution (modified in 0). Then, the probability mass function of (X, Y) is given by

$$\mathbb{P}(\{X = x\} \cap \{Y = y\}) = (p_0 p_1 p_2) q_1^x q_2^y \sum_{k=0}^{\min\{x,y\}} \left(\frac{q_0}{q_1 q_2}\right)^k, \qquad (x, y) \in \mathbb{N}^2.$$

Remark that

- for any $x \in \mathbb{N}$ and $y \in \mathbb{N}$,

$$\mathbb{P}(X = x) = \frac{p_1 p_0}{q_0 - q_1}(q_0^{x+1} - q_1^{x+1}), \qquad \mathbb{P}(Y = y) = \frac{p_2 p_0}{q_0 - q_2}(q_0^{y+1} - q_2^{y+1}).$$

- we have

$$\mathbb{C}(X, Y) = \mathbb{V}(U_0) = \frac{q_0}{p_0^2}.$$

For a comprehensive treatment of the bivariate geometric distribution, we refer to (13), (2), and (6). Set now

$$(S, T) = (R_1 X, R_2 Y), \tag{4}$$

with $R_j \sim \mathcal{R}(\alpha_j)$, for any $j \in \{1, 2\}$. Moreover, we suppose that R_1 and R_2 are mutually independent and independent of X and Y.

Plot of a special case of the joint probability of (S, T) is given by Figure 4. Figures 5 and 6 give bivariate plots that illustrate positive and negative correlations for this distribution.

Let us now investigate the estimation of the unknown parameters $(\alpha_1, \alpha_2, p_0, p_1, p_2)$ from a n-sample $(S_1, T_1), \ldots, (S_n, T_n)$ via two estimation methods : the method of moments and the MLE.

Parameter	actual value	Mean (MM)	Sd (MM)	Mean (ML)	Sd (ML)
α_1	0.1	0.09994431	0.027364810	0.10058339	0.010938576
α_2	0.6	0.59918124	0.021549369	0.59987719	0.016319864
p_0	0.3	0.29942310	0.01728777	0.29933596	0.008043778
p_1	0.8	0.80401244	0.037735241	0.80076329	0.015544756
p_2	0.4	0.39954472	0.032864295	0.39998300	0.010265953

Table 2: Parametric estimation of the unknown parameters $(\alpha_1, \alpha_2, p_0, p_1, p_2)$ from the distribution of (4) (using the geometric distributions) via the MM and MLE.

Method of moments. One can observe that

$$\mathbb{P}(\{S=0\} \cap \{T=0\}) = p_0 p_1 p_2, \quad \mathbb{P}(S=0) = p_0 p_1, \quad \mathbb{P}(T=0) = p_0 p_2.$$

Natural estimators for p_1, p_2 and p_0 are respectively given by

$$\hat{p}_1 = \frac{\sum_{i=1}^n \mathbf{1}_{\{(S_i,T_i)=(0,0)\}}}{\sum_{i=1}^n \mathbf{1}_{\{T_i=0\}}}, \qquad \hat{p}_2 = \frac{\sum_{i=1}^n \mathbf{1}_{\{(S_i,T_i)=(0,0)\}}}{\sum_{i=1}^n \mathbf{1}_{\{S_i=0\}}}$$

and

$$\hat{p}_0 = \frac{\sum_{i=1}^n \mathbf{1}_{\{(S_i,T_i)=(0,0)\}}}{n\hat{p}_1\hat{p}_2}.$$

Now remark that

$$\mathbb{E}(S) = (2\alpha_1 - 1)\left(\frac{q_1}{p_1} + \frac{q_0}{p_0}\right), \qquad \mathbb{E}(T) = (2\alpha_2 - 1)\left(\frac{q_2}{p_2} + \frac{q_0}{p_0}\right).$$

Therefore natural estimators for α_1 and α_2 are respectively given by

$$\hat{\alpha}_1 = \frac{1}{2}\left(\frac{\hat{p}_1 + \hat{p}_0 - (2 - \overline{S})\hat{p}_1\hat{p}_0}{\hat{p}_1 + \hat{p}_0 - 2\hat{p}_1\hat{p}_0}\right)$$

and

$$\hat{\alpha}_2 = \frac{1}{2}\left(\frac{\hat{p}_2 + \hat{p}_0 - (2 - \overline{T})\hat{p}_2\hat{p}_0}{\hat{p}_2 + \hat{p}_0 - 2\hat{p}_2\hat{p}_0}\right),$$

where

$$\overline{S} = \frac{1}{n}\sum_{i=1}^n S_i, \qquad \overline{T} = \frac{1}{n}\sum_{i=1}^n T_i.$$

MLE. The MLE of α_1 and α_2 are given by (1) and (2). To calculate the MLE of $\lambda = (\lambda_0, \lambda_1, \lambda_2)$, it suffices to program a function that calculates the MLE of a bivariate Geometric distribution, which is straightforward with standard statistical packages.

A short simulation experiment

We simulate $n = 1000$ observations of (S, T) (where (X, Y) follows a bivariate Geometric distribution) and we compute both moment method (MM) and maximum likelihood (ML) estimators. After 1000 independent replications, we calculate the average and the standard deviation of the sequence of the estimates obtained. Outputs of this simulation are reported in Table 2.

Remark that from Table 2 the ML estimation gives a smaller standard deviation than the MM estimation for all parameters. Fitting to a Gaussian distribution is illustrated in Figure 8, for the ML estimator. The good fits illustrate nice asymptotic normality properties of the considered estimator.

Acknowledgments.

The authors are thankful to the editor and the referee for their fruitful suggestions and comments.

References

[1] Balakrishnan, N. Discrete multivariate distributions. In: Encyclopedia of Statistical Sciences, 2nd ed. (Eds. N. Balakrishnan, C. Read and B. Vidakovic). John Wiley and Sons, New York, 2005.
[2] Basu, A. P. and Dhar, S. Bivariate Geometric Distribution. Journal Applied Statistical Science, 2(1):33-44, 1995.
[3] Bishop, Y. M., Fienberg, S. E., and Holland, P. W. Discrete multivariate analysis: theory and practice. Springer Science and Business Media, 2007.
[4] Bulla, J. Chesneau, C. and Kachour, M. On the bivariate Skellam distribution. Communications in Statistics - Theory and Method. DOI : 10.1080/03610926.2013.837925, 2015.
[5] Chesneau, C. Kachour, M. and Karlis, D. On some distributions arising from a generalized trivariate reduction scheme. Statistical Methodology, Vol. 25, pp. 36-50, July 2015.
[6] Dhar, S. and Balaji, S. On the characterization of a bivariate geometric distribution. Communications in Statistics, Theory and Methods, 35(5):759-766, 2006.
[7] Holgate, P. Estimation for the bivariate Poisson distribution. Biometrika, 51(1-2), 241-287, 1964.
[8] Johnson, N., Kotz, S. and Balakrishnan, N. Multivariate Discrete Distributions. Wiley, New York, 1997.
[9] Karlis, D. and Ntzoufras, I. Bivariate Poisson and Diagonal Inflated Bivariate Poisson Regression Models in R. Journal of Statistical Software. Volume 14, Issue 10, pp. 1-36, 2005.
[10] Kocherlakota, S. and Kocherlakota, K. Bivariate Discrete Distributions, Statistics: textbooks and monographs, volume 132. Markel Dekker, New York, 1992.
[11] Kocherlakota, S. and Kocherlakota, K. Bivariate discrete distributions, in S. Kotz, C. B. Read and D. L. Banks, eds, Encyclopedia of Statistical Sciences- Update, Vol. 2, Wiley, New York, pp. 6883, 1997.
[12] Lai, C.D. Construction of discrete bivariate distributions. In Advances in distribution theory, order statistics, and inference, Birkhauser, Boston, pp. 29-58, 2006.
[13] Nair, K. R. M. and Nair, U. On characterizing the bivariate exponential and geometric distributions. Annals of the Institute of Statistical Mathematics, 40(2): 267-271, 1988.
[14] Papageorgiou, H. Multivariate discrete distributions, in C. B. Kotz, S. Read and D. L. Banks, eds, Encyclopedia of Statistical Sciences-Updated, Vol. 1, Wiley, New York, pp. 408419, 1997.
[15] Sarabia J. and Gómez-Déniz E. Construction of multivariate distributions: a review of some recent results. Stat Operat Res Trans 32:435, 2008.

Figures

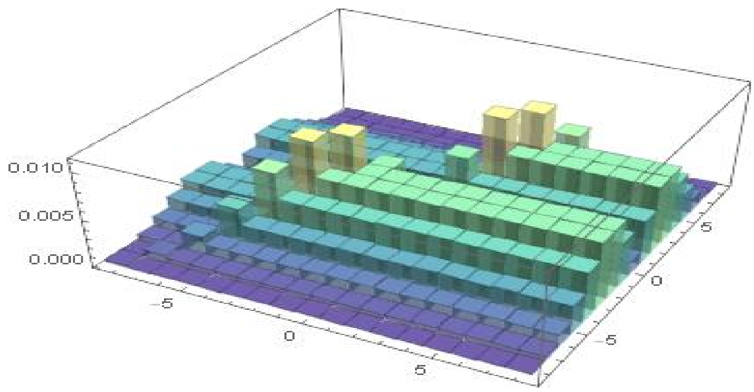

Figure 1: Plot of the joint probability of (S, T) from the $\text{Rad}(\alpha_1, \alpha_2, \theta)$ class when (X, Y) follows the bivariate Poisson distribution, with $\alpha_1 = 0.5$, $\alpha_2 = 0.33$, $\lambda_0 = 1$, $\lambda_1 = 2$ and $\lambda_2 = 3$.

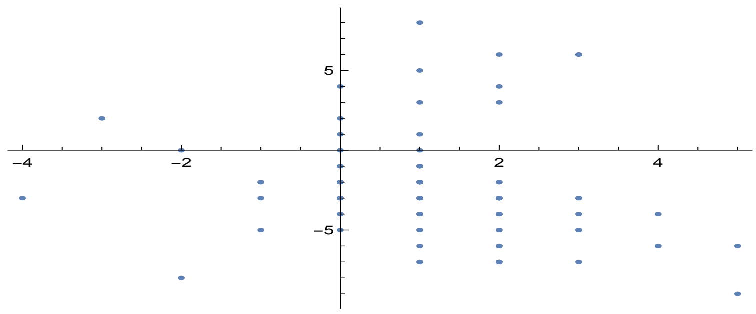

Figure 2: Bivariate plot of a sample (s, t) of length 100 of (S, T) from the $\text{Rad}(\alpha_1, \alpha_2, \theta)$ class when (X, Y) follows the bivariate Poisson distribution, with $\alpha_1 = 0.9$, $\alpha_2 = 0.1$, $\lambda_0 = 1$, $\lambda_1 = 0.6$ and $\lambda_2 = 3$; negative correlation of -0.204911.

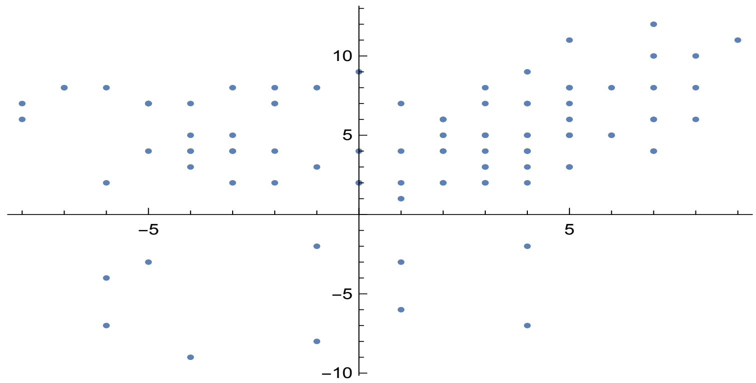

Figure 3: Bivariate plot of a sample (s, t) of length 100 of (S, T) from the Rad$(\alpha_1, \alpha_2, \theta)$ class when (X, Y) follows the bivariate Poisson distribution, with $\alpha_1 = 0.7$, $\alpha_2 = 0.9$, $\lambda_0 = 2$, $\lambda_1 = 1.6$ and $\lambda_2 = 3$; positive correlation of 0.230737.

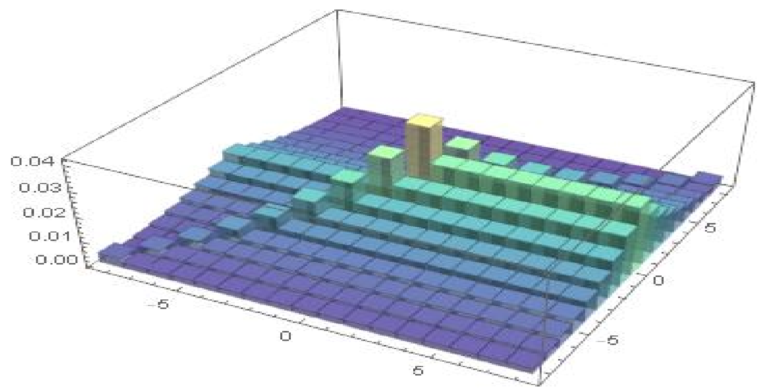

Figure 4: Plot of the joint probability of (S, T) from Rad$(\alpha_1, \alpha_2, \theta)$ class when (X, Y) follows the bivariate Geometric distribution with $\alpha_1 = 0.5$, $\alpha_2 = 0.33$, $p_0 = 0.5$, $p_1 = 0.33$ and $p_2 = 0.25$.

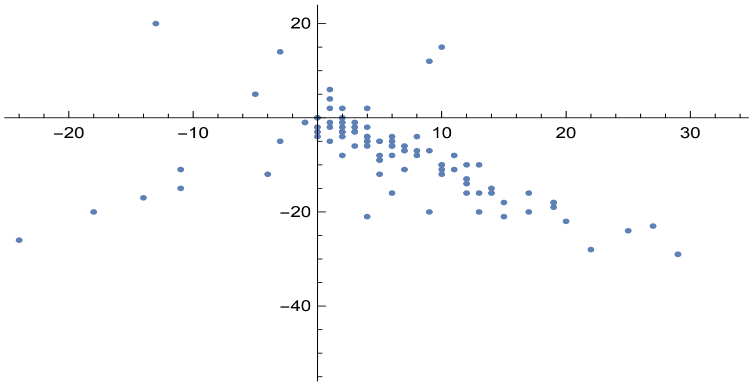

Figure 5: Bivariate plot of a sample (s,t) of length 100 of (S,T) from the $\text{Rad}(\alpha_1,\alpha_2,\theta)$ class when (X,Y) follows the bivariate Geometric distribution with $\alpha_1 = 0.9$, $\alpha_2 = 0.1$, $p_0 = 0.1$, $p_1 = 0.6$ and $p_2 = 0.3$; negative correlation of -0.66291.

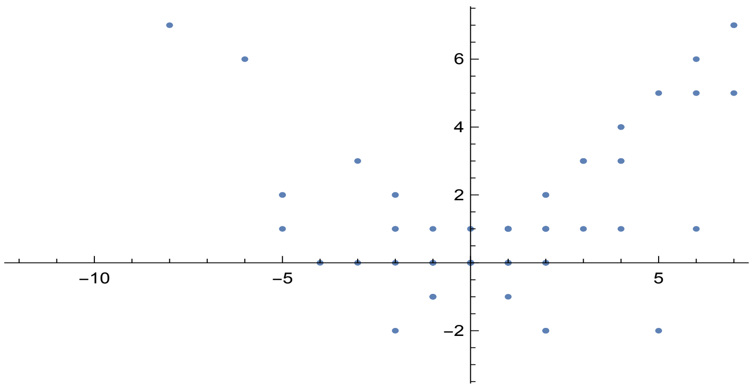

Figure 6: Bivariate plot of a sample (s,t) of length 100 of (S,T) from the $\text{Rad}(\alpha_1,\alpha_2,\theta)$ class when (X,Y) follows the bivariate Geometric distribution with $\alpha_1 = 0.7$, $\alpha_2 = 0.8$, $p_0 = 0.4$, $p_1 = 0.6$ and $p_2 = 0.9$; positive correlation of 0.470974.

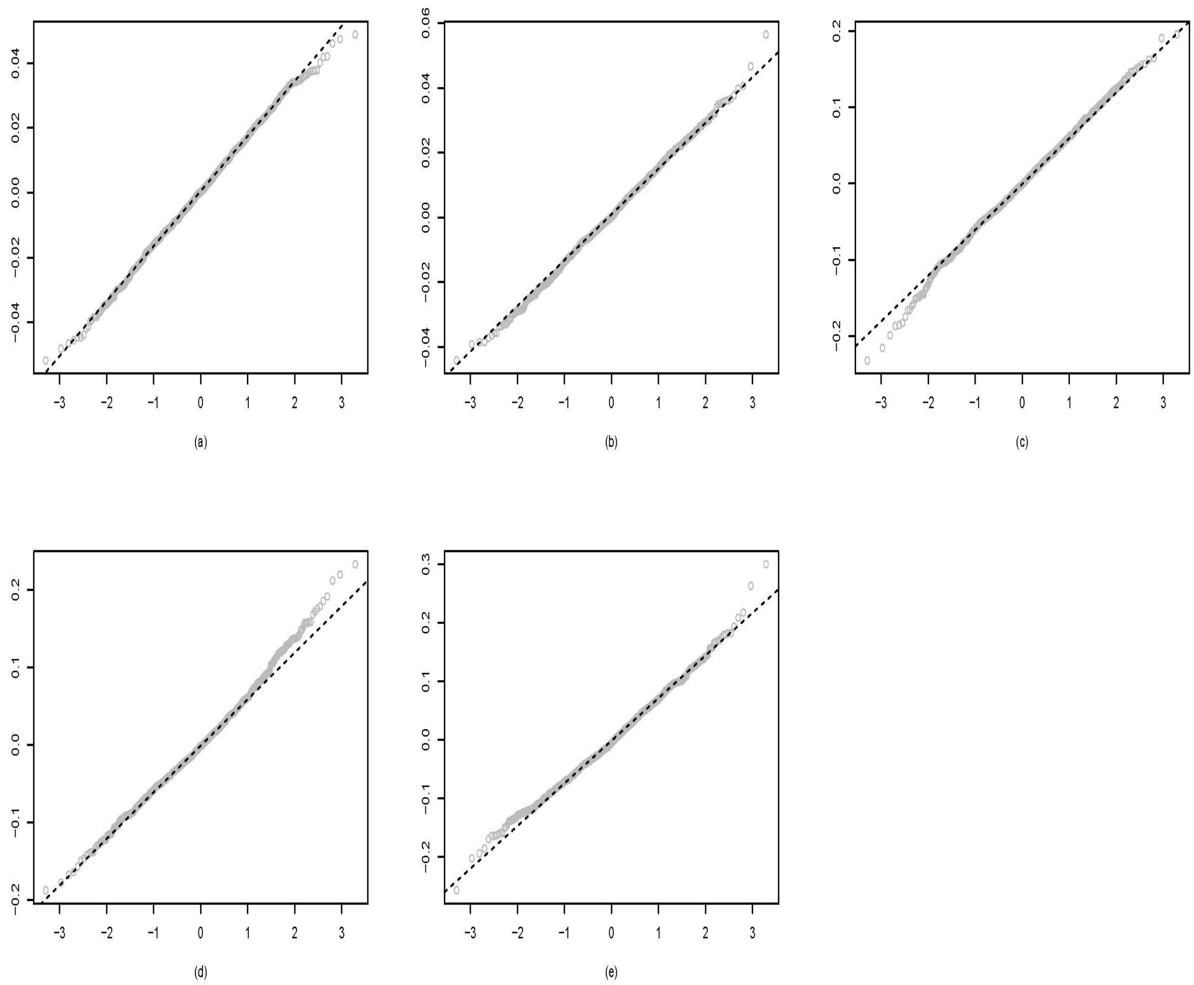

Figure 7: Normal Q-Q plots for the errors (ML estimator), when (X, Y) follows bivariate Poisson distribution and $n = 1000$. (a) Normal Q-Q plots of $\hat{\alpha}_1 - \alpha_1$. (b) Normal Q-Q plots of $\hat{\alpha}_2 - \alpha_2$. (c) Normal Q-Q plots of $\hat{\lambda}_0 - \lambda_0$. (d) Normal Q-Q plots of $\hat{\lambda}_1 - \lambda_1$. (e) Normal Q-Q plots of $\hat{\lambda}_2 - \lambda_2$.

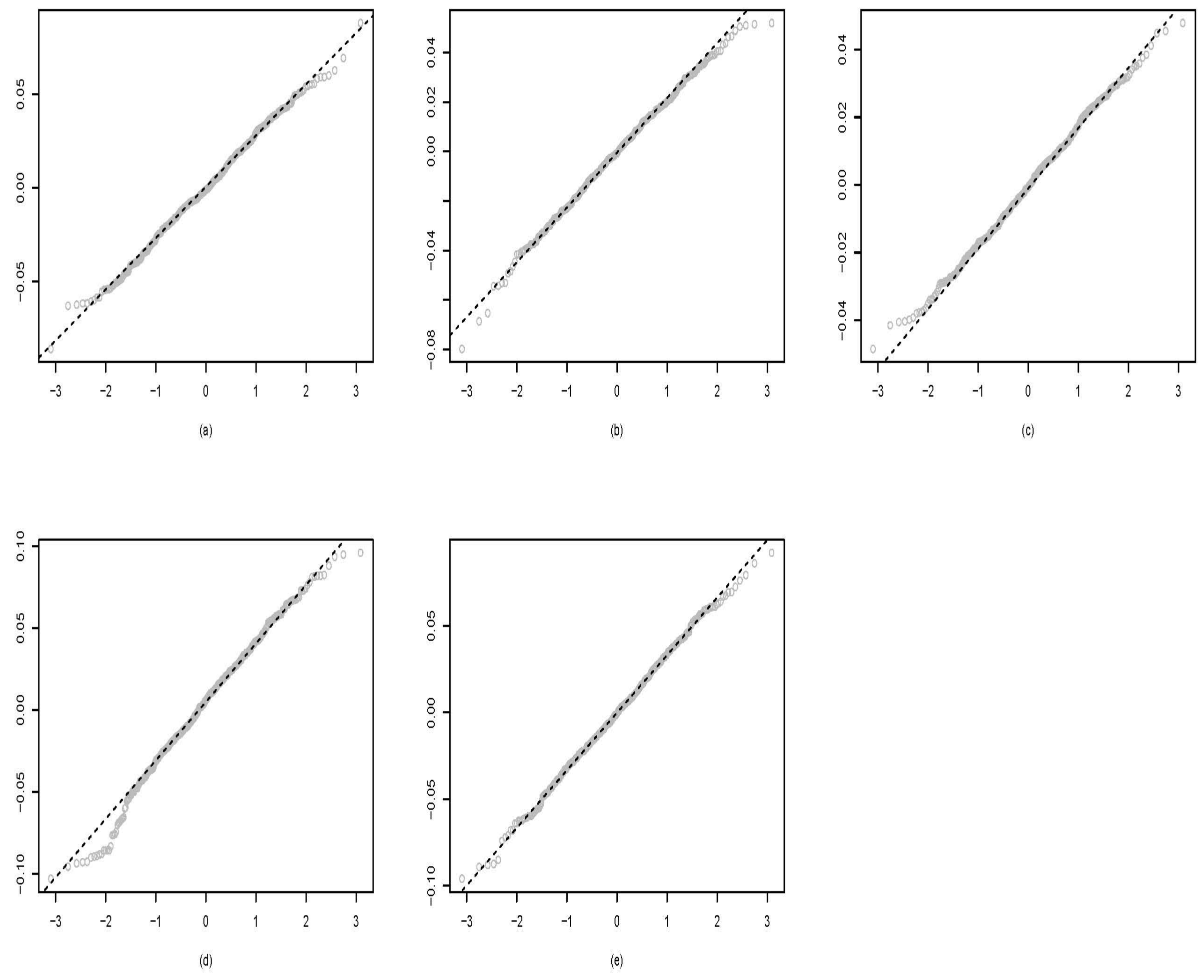

Figure 8: Normal Q-Q plots for the errors (ML estimator), when (X, Y) follows bivariate Geometric distribution and $n = 1000$. (a) Normal Q-Q plots of $\hat{\alpha}_1 - \alpha_1$. (b) Normal Q-Q plots of $\hat{\alpha}_2 - \alpha_2$. (c) Normal Q-Q plots of $\hat{p}_0 - p_0$. (d) Normal Q-Q plots of $\hat{p}_1 - p_1$. (e) Normal Q-Q plots of $\hat{p}_2 - p_2$.