Monotone couplings: Stein's method and beyond

Fraser Daly

Heriot–Watt University

Based on joint work with Oliver Johnson, Claude Lefèvre and Sergey Utev

May 2015

Outline

We will use stochastic orderings to extend well-known ‘monotone coupling’ results in Stein’s method for Poisson approximation. The same ideas may be employed in many other settings.

- The Poisson case
  - Size biasing and stochastic ordering
  - Negative dependence and monotone couplings
  - Poisson approximation
  - Bounding the Poincaré constant
  - Convex ordering, with applications
- The geometric case
  - Monotone failure rate
  - Applications to queues

Size biasing

Throughout this talk, let W be a non–negative, integer–valued random variable with $\mathbb{E}[W] = \lambda > 0$ and $\mathbb{P}(W = 0) = p$.

Size biasing

We let $W^\star$ have the W–size–biased distribution, so that for $j \in \mathbb{Z}^+$

$$\mathbb{P}(W^\star = j) = \frac{j\mathbb{P}(W = j)}{\lambda}\,.$$

Equivalently, we define $W^\star$ by

$$\mathbb{E}[Wg(W)] = \lambda\mathbb{E}[g(W^\star)]\,,$$

for all functions g for which the expectations exist.

Stochastic ordering

Stochastic ordering

We write $X \geq_{st} Y$ if $\mathbb{E}f(X) \geq \mathbb{E}f(Y)$ for all increasing functions f.

Throughout this talk, we consider random variables satisfying

$$W + 1 \geq_{st} W^\star \,. \tag{1}$$

This is our 'negative dependence' (or 'monotone coupling') condition.

Negative dependence

Our negative dependence condition holds if

1. $W \sim \text{Po}(\lambda)$. Equality holds in (1) if and only if $W \sim \text{Po}(\lambda)$. This is the usual characterization of the Poisson distribution used in Stein's method.

Negative dependence

Our negative dependence condition holds if

1. $W \sim \text{Po}(\lambda)$. Equality holds in (1) if and only if $W \sim \text{Po}(\lambda)$. This is the usual characterization of the Poisson distribution used in Stein's method.
2. W is *ultra log–concave* (of degree ∞). That is,

$$\frac{(j+1)\mathbb{P}(W=j+1)}{\mathbb{P}(W=j)} \quad \text{is increasing in } j\,.$$

   Or, equivalently, $W+1 \geq_{lr} W^\star$. Here '$\geq_{lr}$' is the likelihood ratio ordering. The ULC(∞) class was introduced by Liggett (1997) to capture negative dependence.

3. $W = X_1 + \cdots + X_n$ is a sum of (dependent) Bernoulli random variables with

$$\mathbb{E}[\phi(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)|X_i = 1] \\ \leq \mathbb{E}[\phi(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)]\,,$$

for each $i = 1, \ldots, n$ and increasing function $\phi : \{0,1\}^{n-1} \mapsto \{0,1\}$. That is, $X_1, \ldots, X_n$ are *negatively related*. See Barbour, Holst and Janson (1992).

3. $W = X_1 + \cdots + X_n$ is a sum of (dependent) Bernoulli random variables with

$$\mathbb{E}[\phi(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)|X_i = 1] \\ \leq \mathbb{E}[\phi(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)]\,,$$

for each $i = 1, \ldots, n$ and increasing function $\phi : \{0,1\}^{n-1} \mapsto \{0,1\}$. That is, $X_1, \ldots, X_n$ are *negatively related*. See Barbour, Holst and Janson (1992).

4. $W = X_1 + \cdots + X_n$ is a sum of (dependent) Bernoulli random variables with

$$\mathrm{Cov}(f(X_i), g(W - X_i)) \leq 0\,,$$

for each $i = 1, \ldots, n$ and increasing functions f and g. That is, $X_1, \ldots, X_n$ are *totally negatively dependent*. See Papadatos and Papathanasiou (2002).

Some examples

The hypergeometric distribution

Distribute m balls uniformly into N urns (each with capacity for up to one ball), and let $W = X_1 + \cdots + X_n$ count the number of the first n urns that are occupied. Then $X_1, \ldots, X_n$ are negatively related and so W satisfies our negative dependence condition (1).

Small spacings on a circle

Distribute n points uniformly on a circle and let $S_1, \ldots, S_n$ be the arc–length distances between successive points. Let $X_i = I(S_i < a)$ for some threshold $a > 0$. Then $W = X_1 + \cdots + X_n$ counts the number of small spacings. $X_1, \ldots, X_n$ are negatively related, and so W satisfies our negative dependence condition (1).

See Barbour, Holst and Janson (1992) for these, and many more, examples.

Poisson approximation

We begin with Poisson approximation for W.

We use the total variation distance

$$d_{TV}(\mathcal{L}(W), \mathcal{L}(Y)) = \sup_{A \subseteq \mathbb{Z}^+} |P(W \in A) - P(Y \in A)| \,.$$

Theorem

Let $W + 1 \geq_{st} W^\star$. Then

$$d_{TV}(\mathcal{L}(W), Po(\lambda)) \leq \frac{1 - e^{-\lambda}}{\lambda} \{\lambda - Var(W)\} \,.$$

See Daly, Lefèvre and Utev (2012).

Sketch proof

Using Stein's method for Poisson approximation, we may write

$$d_{TV}(\mathcal{L}(W), \mathrm{Po}(\lambda)) = \lambda \sup_{A \subseteq \mathbb{Z}^+} \left| \mathbb{E} f_A(W+1) - \mathbb{E} f_A(W^\star) \right|,$$

where, for each $A \subseteq \mathbb{Z}^+$, $\sup_x |f_A(x+1) - f_A(x)| \leq \frac{1-e^{-\lambda}}{\lambda}$.

Sketch proof

Using Stein's method for Poisson approximation, we may write

$$d_{TV}(\mathcal{L}(W), \mathrm{Po}(\lambda)) = \lambda \sup_{A \subseteq \mathbb{Z}^+} \left| \mathbb{E} f_A(W+1) - \mathbb{E} f_A(W^\star) \right| ,$$

where, for each $A \subseteq \mathbb{Z}^+$, $\sup_x |f_A(x+1) - f_A(x)| \leq \frac{1-e^{-\lambda}}{\lambda}$. Now,

$$\begin{aligned} \mathbb{E} f_A(W+1) - \mathbb{E} f_A(W^\star) &= \sum_{j=0}^{\infty} f_A(j) \left[\mathbb{P}(W+1=j) - \mathbb{P}(W^\star = j) \right] \\ &= \sum_{j=0}^{\infty} \left\{ f_A(j+1) - f_A(j) \right\} \left[\mathbb{P}(W+1 > j) - \mathbb{P}(W^\star > j) \right] . \end{aligned}$$

The result follows.

Theorem

Let $W = X_1 + \cdots + X_n$, $p_i = \mathbb{E}X_i$ and $\lambda = \mathbb{E}W$. Suppose that for each $1 \leq i \leq n$ there are random variables Y_{ji} $(1 \leq j \leq n)$ defined on the same probability space as the X_i with

$$\mathcal{L}(Y_{ji} : 1 \leq j \leq n) = \mathcal{L}(X_j : 1 \leq j \leq n | X_i = 1)\,.$$

Suppose also that for each i the set $\{1, \ldots, n\} \setminus \{i\}$ is partitioned into Γ_i^-, Γ_i^+ and Γ_i^0 such that $Y_{ji} \leq X_j$ if $j \in \Gamma_i^-$ and $Y_{ji} \geq X_j$ if $j \in \Gamma_i^+$. Then

$$d_{TV}(\mathcal{L}(W), Po(\lambda)) \leq \frac{1 - e^{-\lambda}}{\lambda} \left\{ \sum_{i=1}^{n} p_i^2 + \sum_{i=1}^{n} \sum_{j \in \Gamma_i^-} |Cov(X_i, X_j)| \right.$$
$$\left. + \sum_{i=1}^{n} \sum_{j \in \Gamma_i^+} Cov(X_i, X_j) + \sum_{i=1}^{n} \sum_{j \in \Gamma_i^0} (\mathbb{E}X_i X_j + p_i p_j) \right\}.$$

See Barbour, Holst and Janson (1992).

Generalizations and extensions

- Positive dependence: suppose $W = X_1 + \cdots + X_n$ is a sum of indicators. There are many interesting applications where we have

$$W + 1 - X_J \leq_{st} W^\star ,$$

where J is a random index with

$$\mathbb{P}(J = i) = \frac{\mathbb{E}X_i}{\mathbb{E}W} , \qquad 1 \leq i \leq n .$$

This happens, for example, if $X_1, \dots, X_n$ are positively related. See Daly, Lefèvre and Utev (2012).

- Compound Poisson approximation. See Daly (2013).
- Approximation by the equilibrium of a birth–death process. See Daly, Lefèvre and Utev (2012).

Bounds on the Poincaré (inverse spectral gap) constant

Let

$$R_W = \sup_{g \in \mathcal{G}(W)} \left\{ \frac{\mathbb{E}[g(W)^2]}{\mathbb{E}[\{g(W+1) - g(W)\}^2]} \right\},$$

where

$$\mathcal{G}(W) = \{g : \mathbb{Z}^+ \mapsto \mathbb{R} \text{ with } \mathbb{E}[g(W)^2] < \infty \text{ and } \mathbb{E}[g(W)] = 0\}.$$

The lower bound $R_W \geq \mathrm{Var}(W)$, with equality iff W is (shifted) Poisson, is well-known. Under our negative dependence assumption, we also get a straightforward upper bound.

Theorem

If $W + 1 \geq_{st} W^\star$,

$$R_W \leq \lambda .$$

See Daly and Johnson (2013).

Sketch proof

Following Klaassen (1985), define

$$\chi(i,j) = I(\lfloor\lambda\rfloor \leq j < i) - I(i \leq j < \lfloor\lambda\rfloor) - (\lambda - \lfloor\lambda\rfloor)I(j = \lfloor\lambda\rfloor)\,.$$

Sketch proof

Following Klaassen (1985), define

$$\chi(i,j) = I(\lfloor\lambda\rfloor \le j < i) - I(i \le j < \lfloor\lambda\rfloor) - (\lambda - \lfloor\lambda\rfloor)I(j = \lfloor\lambda\rfloor)\,.$$

Using the Cauchy-Schwarz inequality, we can show that for $g \in \mathcal{G}(W)$,

$$\mathbb{E}[g(W)^2] \le \lambda \sum_{j=0}^{\infty} [g(j+1) - g(j)]^2 \left\{\mathbb{E}\chi(W^\star, j) - \mathbb{E}\chi(W, j)\right\}\,.$$

Now use the fact that $\chi(i,j)$ is increasing in i for fixed j and apply our stochastic ordering assumption.

Sketch proof

Following Klaassen (1985), define

$$\chi(i,j) = I(\lfloor\lambda\rfloor \leq j < i) - I(i \leq j < \lfloor\lambda\rfloor) - (\lambda - \lfloor\lambda\rfloor)I(j = \lfloor\lambda\rfloor)\,.$$

Using the Cauchy-Schwarz inequality, we can show that for $g \in \mathcal{G}(W)$,

$$\mathbb{E}[g(W)^2] \leq \lambda \sum_{j=0}^{\infty}[g(j+1) - g(j)]^2 \left\{\mathbb{E}\chi(W^\star, j) - \mathbb{E}\chi(W, j)\right\}\,.$$

Now use the fact that $\chi(i,j)$ is increasing in i for fixed j and apply our stochastic ordering assumption.
Finally, we have that

$$\chi(w+1,j) - \chi(w,j) = I(w = j)\,.$$

Negative dependence and convex ordering

For random variables X and Y, we write that $X \geq_{cx} Y$ if $\mathbb{E}f(X) \geq \mathbb{E}f(Y)$ for all convex functions f.

Theorem

If $W + 1 \geq_{st} W^\star$ and $Z \sim Po(\lambda)$ then

$$W \leq_{cx} Z\,.$$

Technical note: we prove that W is smaller than Z in the *increasing* convex order. Convex ordering then follows since $\mathbb{E}[W] = \mathbb{E}[Z]$.

This result can be generalized to the s–convex orderings for any integer $s \geq 1$ ($s = 1$: usual stochastic order; $s = 2$: increasing convex order). See Lefèvre and Utev (1996) for a definition.

Sketch proof

Define the thinning operator: $T_\alpha W = \sum_{i=1}^{W} \eta_i$, where $\eta_1, \eta_2 \ldots$ are iid Be(α). Let $W_\alpha = T_\alpha W + T_{1-\alpha} Z$. We can show that if $W + 1 \geq_{st} W^\star$ then $W_\alpha + 1 \geq_{st} W_\alpha^\star$ for all $\alpha \in [0, 1]$.

So

$$0 \leq \int_0^1 \frac{\lambda}{\alpha} \left[\mathbb{P}(W_\alpha + 1 \geq j) - \mathbb{P}(W_\alpha^\star \geq j)\right] d\alpha$$

Sketch proof

Define the thinning operator: $T_\alpha W = \sum_{i=1}^{W} \eta_i$, where $\eta_1, \eta_2 \dots$ are iid $\mathrm{Be}(\alpha)$. Let $W_\alpha = T_\alpha W + T_{1-\alpha} Z$. We can show that if $W + 1 \geq_{st} W^\star$ then $W_\alpha + 1 \geq_{st} W_\alpha^\star$ for all $\alpha \in [0,1]$.

Let $h_j(X) = \mathbb{E}(X - j + 1)_+$. By a lemma of Johnson (2007)

$$-\frac{\partial}{\partial\alpha} h_j(W_\alpha) = \frac{\lambda}{\alpha}\left[\mathbb{P}(W_\alpha + 1 \geq j) - \mathbb{P}(W_\alpha^\star \geq j)\right].$$

So

$$0 \leq \int_0^1 \frac{\lambda}{\alpha}\left[\mathbb{P}(W_\alpha + 1 \geq j) - \mathbb{P}(W_\alpha^\star \geq j)\right] d\alpha$$

Sketch proof

Define the thinning operator: $T_\alpha W = \sum_{i=1}^{W} \eta_i$, where $\eta_1, \eta_2 \dots$ are iid $\mathrm{Be}(\alpha)$. Let $W_\alpha = T_\alpha W + T_{1-\alpha} Z$. We can show that if $W + 1 \geq_{st} W^\star$ then $W_\alpha + 1 \geq_{st} W_\alpha^\star$ for all $\alpha \in [0,1]$.

Let $h_j(X) = \mathbb{E}(X - j + 1)_+$. By a lemma of Johnson (2007)

$$-\frac{\partial}{\partial \alpha} h_j(W_\alpha) = \frac{\lambda}{\alpha} \left[\mathbb{P}(W_\alpha + 1 \geq j) - \mathbb{P}(W_\alpha^\star \geq j) \right].$$

So

$$\begin{aligned} 0 &\leq \int_0^1 \frac{\lambda}{\alpha} \left[\mathbb{P}(W_\alpha + 1 \geq j) - \mathbb{P}(W_\alpha^\star \geq j) \right] d\alpha \\ &= -\int_0^1 \frac{\partial}{\partial \alpha} h_j(W_\alpha)\, d\alpha = h_j(Z) - h_j(W), \end{aligned}$$

which proves the (increasing) convex ordering.

Application: upper bounds on entropy

Let $q_j = \mathbb{P}(W = j)$, and

$$H(W) = -\sum_{j=0}^{\infty} q_j \log(q_j)\,,$$

the entropy of W.

Corollary

If $W + 1 \geq_{st} W^\star$,

$$H(W) \leq H(Z)\,.$$

This follows from our theorem and a lemma of Yu (2009), since Z is log–concave. See also Johnson (2007).

Application: concentration inequalities

Corollary

Let $W+1 \geq_{st} W^\star$ and $s>0$. Then

$$\begin{aligned} \mathbb{P}(W \geq \lambda + s) &\leq e^s \left(1+\frac{s}{\lambda}\right)^{-(s+\lambda)}, \\ \mathbb{P}(W \leq \lambda - s) &\leq e^{-s} \left(1-\frac{s}{\lambda}\right)^{s-\lambda}, \end{aligned}$$

where the latter bound applies if $s<\lambda$.

See also Arratia and Baxendale (2015).

Application: Poisson approximation

If X and Y have distribution functions F and G, respectively, define the distance

$$d_n(\mathcal{L}(X), \mathcal{L}(Y)) = \|\Delta^n F - \Delta^n G\|_\infty ,$$

where $\Delta f(x) = f(x+1) - f(x)$ is the forward difference operator.

Corollary

If $W + 1 \geq_{st} W^\star$ then

$$d_n(\mathcal{L}(W), \mathcal{L}(Z)) \leq 2^{(1+n)_+ - 1} \left(\lambda - Var(W)\right) ,$$

for $n \in \{-2, -1, 0, 1\}$.

This includes bounds on the Kolmogorov and stop–loss distances, and a distance useful in local limit results.

Compound geometric approximation

Define the failure rate of W:

$$r_W(j) = \frac{\mathbb{P}(W = j)}{\mathbb{P}(W > j)},$$

for $j \in \mathbb{Z}^+ = \{0, 1, 2, \ldots\}$.

We let W be such that $\mathbb{P}(W = 0) = p$ and $r_W(j) \geq \delta$ for all j.

Under these conditions, we consider the approximation of W by the geometric sum $U = \sum_{i=1}^{N} Y_i$, where $N \sim \mathrm{Ge}(p)$ has a geometric distribution (supported on $\mathbb{Z}^+$) and $Y, Y_1, Y_2, \ldots$ are i.i.d. positive, integer–valued random variables with

$$p \leq (1 - p)\delta\mathbb{E}Y.$$

Compound geometric approximation

Theorem

With the above definitions and conditions,

$$d_{TV}(\mathcal{L}(W), \mathcal{L}(U)) \leq \frac{(1-p)\mathbb{E}Y}{p} - \mathbb{E}W\,.$$

Under the conditions of the theorem, this upper bound is always nonnegative.

We will consider some applications of this result to the M/G/1 queue.

It will be easiest to evaluate a suitable lower bound δ if W has either increasing failure rate (IFR) or decreasing failure rate (DFR).

Sketch proof

Let V be such that $V + Y \stackrel{st}{=} W|W > 0$. Using Stein's method,

$$d_{TV}(\mathcal{L}(W), \mathcal{L}(U)) = (1-p) \sup_{A \subseteq \mathbb{Z}^+} \left| \mathbb{E} f_A(W+Y) - \mathbb{E} f_A(V+Y) \right| ,$$

where, for each $A \subseteq \mathbb{Z}^+$, $\sup_{j,k} |f_A(j) - f_A(k)| \leq p^{-1}$.

Sketch proof

Let V be such that $V + Y \stackrel{st}{=} W | W > 0$. Using Stein's method,

$$d_{TV}(\mathcal{L}(W), \mathcal{L}(U)) = (1-p) \sup_{A \subseteq \mathbb{Z}^+} |\mathbb{E} f_A(W+Y) - \mathbb{E} f_A(V+Y)| \,,$$

where, for each $A \subseteq \mathbb{Z}^+$, $\sup_{j,k} |f_A(j) - f_A(k)| \leq p^{-1}$.

Analogously to the Poisson case, we find conditions under which $W + Y \geq_{st} V + Y$. By the definition of V, this holds if

$$(1-p)\mathbb{P}(W + Y > j) \geq \mathbb{P}(W > j) \,, \qquad \forall j \in \mathbb{Z}^+ \,.$$

Sketch proof

Let V be such that $V + Y \stackrel{st}{=} W | W > 0$. Using Stein's method,

$$d_{TV}(\mathcal{L}(W), \mathcal{L}(U)) = (1-p) \sup_{A \subseteq \mathbb{Z}^+} |\mathbb{E} f_A(W+Y) - \mathbb{E} f_A(V+Y)| \,,$$

where, for each $A \subseteq \mathbb{Z}^+$, $\sup_{j,k} |f_A(j) - f_A(k)| \leq p^{-1}$.

Analogously to the Poisson case, we find conditions under which $W + Y \geq_{st} V + Y$. By the definition of V, this holds if

$$(1-p)\mathbb{P}(W+Y > j) \geq \mathbb{P}(W > j) \,, \qquad \forall j \in \mathbb{Z}^+ \,.$$

Rearranging this, we require that

$$\frac{1}{\mathbb{P}(W > j)} \mathbb{E}\left[\sum_{k=j+1-Y}^{j} \mathbb{P}(W = k)\right] \geq \frac{p}{1-p} \,,$$

which holds under our lower bound on the failure rate. The bound follows after some straightforward algebra.

Geometric approximation for IFR distributions

If the failure rate of W is increasing, a lower bound is given by $r_W(0)$. In this case, we may take $Y = 1$ a.s., and we have geometric approximation:

Corollary

Let W be IFR with $\mathbb{P}(W = 0) = p$. Then

$$d_{TV}(\mathcal{L}(W), Ge(p)) \leq \frac{1-p}{p} - \mathbb{E}W$$

Applications to the M/G/1 queue

Consider a single server queue with

- customers arriving at rate η, and
- i.i.d. customer service times with the same distribution as S.

Let $\rho = \eta \mathbb{E} S$ and assume $\rho < 1$.

We will present approximation results for

1. the number of customers in the system in equilibrium, and
2. the number of customers served during a busy period.

Number of customers in the system

Let W be the equilibrium number of customers in the system. We have

$$\mathbb{P}(W=0) = 1-\rho\,, \qquad \mathbb{E}W = \rho + \frac{\rho^2\mathbb{E}[S^2]}{2(1-\rho)(\mathbb{E}S)^2}\,,$$

$$r_W(j) \geq \frac{1-\rho}{\eta\sup_{t\geq 0}\mathbb{E}[S-t|S\geq t]}\,, \qquad \forall j\in\mathbb{Z}^+\,.$$

See, for example, Asmussen (2003) and Ross (2006).

S is said to be New Better than Used in Expectation (NBUE) if $\mathbb{E}[S-t|S\geq t]\leq \mathbb{E}S$ for all $t\geq 0$. In that case our theorem gives

$$d_{TV}(\mathcal{L}(W),\mathrm{Ge}(1-\rho)) \leq \frac{\rho^2}{1-\rho}\left(1-\frac{\mathbb{E}[S^2]}{2(\mathbb{E}S)^2}\right)\,.$$

As expected, this upper bound is zero if S has an exponential distribution.

Customers served during a busy period

Let $W+1$ be the number of customers served during a busy period of our system. Then $\mathbb{P}(W=0)=\mathbb{E}e^{-\eta S}=p$ and $\mathbb{E}W=\rho(1-\rho)^{-1}$.

If S is IFR, Shanthikumar (1988) shows that W is DFR, and we may find a lower bound on the failure rate using results due to Kyprianou (1972).

Let ψ be the Laplace transform of the density of S and let ξ be the real root of $1+\eta\psi'(s)$ nearest the origin. Let

$$\theta=\frac{\xi-\eta+\eta\psi(\xi)}{(\xi-\eta)\psi(\eta)}\,.$$

Then if $U=\sum_{i=1}^{N}Y_i$, where $N\sim\mathrm{Ge}(p)$ and $(1-p)\theta\mathbb{E}Y\geq 1-p\theta$,

$$d_{TV}(\mathcal{L}(W),\mathcal{L}(U))\leq\frac{(1-p)\mathbb{E}Y}{p}-\frac{\rho}{1-\rho}\,.$$

References

- R. Arratia and P. Baxendale, (2015). Bounded size bias couplings: a Gamma function bound, and universal Dickman-function behavior. *Probab. Theory Relat. Fields*, to appear.
- S. Asmussen, (2003). *Applied Probability and Queues* 2nd ed. Springer, New York.
- A. D. Barbour, L. Holst and S. Janson, (1992). *Poisson Approximation*. Oxford University Press, Oxford.
- F. Daly, (2013). Compound Poisson approximation with association or negative association via Stein's method. *Electron. Comm. Prob.* 18 (30): 1 –12.
- F. Daly, (2015a). Negative dependence and stochastic orderings. Preprint. `http://arxiv.org/abs/1504.06493`
- F. Daly, (2015b). Compound geometric approximation under a failure rate constraint. Preprint. `http://arxiv.org/abs/1504.06498`
- F. Daly and O. Johnson, (2013). Bounds on the Poincaré constant under negative dependence. *Stat. Prob. Lett.* 83:511 –518.
- F. Daly, C. Lefèvre and S. Utev, (2012). Stein's method and stochastic orderings. *Adv. Appl. Prob.* 44: 343–372.
- O. Johnson, (2007). Log–concavity and the maximum entropy property of the Poisson distribution. *Stochastic Processes Appl.* 117: 791 – 802.

References

- C. Lefèvre and S. Utev, (1996). Comparing sums of exchangeable Bernoulli random variables. *J. Appl. Prob.* 33: 285 – 310.
- C. Klaassen, (1985). On an inequality of Chernoff. *Ann. Probab* 13: 966–974.
- E. K. Kyprianou, (1972). The quasi–stationary distributions of queues in heavy traffic. *J. Appl. Probab.* 35: 821–831.
- T. M. Liggett, (1997). Ultra logconcave sequences and negative dependence. *J. Combin. Theory Ser. A* 79: 315–325.
- N. Papadatos and V. Papathanasiou, (2002). Approximation for a sum of dependent indicators: an alternative approach. *Adv. Appl. Prob.* 34: 609 – 625.
- S. Ross, (2006). Bounding the stationary distribution of the M/G/1 queue size. *Probab. Engrg. Inform. Sci* 20: 571–574.
- J. G. Shanthikumar, (1988). DFR property of first–passage times and its preservation under geometric compounding. *Ann. Probab.* 16: 397–406.
- Y. Yu, (2009). On the entropy of compound distributions on nonnegative integers. *IEEE Trans. Inf. Theory* 55: 3645 – 3650.