Probability in the Engineering and Informational Sciences, **14**, 2000, 1–7. Printed in the U.S.A.

ON SIMULATION OF STOCHASTICALLY ORDERED LIFE-LENGTH VARIABLES

Torgny Lindvall
Department of Mathematical Statistics
University of Göteborg
41296 Göteborg, Sweden
E-mail: lindvall@math.chalmers.se

Let F and G be life-length distributions such that $F \overset{\mathcal{D}}{\leq} G$. We solve the following problem: How should (X,Y) be generated in order to maximize $\mathbb{P}(X = Y)$, under the conditions $X \overset{\mathcal{D}}{=} F$, $Y \overset{\mathcal{D}}{=} G$, and $X \leq Y$? We also find a necessary and sufficient condition for the existence of such a maximal coupling with the property that X and Y are independent, conditioned that $X < Y$. It is pointed out that using familiar Poisson process thinning methods does not produce (X,Y) which maximizes $\mathbb{P}(X = Y)$.

1. INTRODUCTION

Let F and G be the distribution functions of the nonnegative random variables X and Y, assumed to denote life lengths of some objects. In order to reveal properties of F and G and relations between them by a simulation, we consider the outcomes of i.i.d. pairs

$$(X_i, Y_i) \overset{\mathcal{D}}{=} (X,Y), \qquad i = 1,2,\ldots,n.$$

We shall pay no attention to how a simulation improves as n increases. Our concern is rather to study how (X,Y) should be generated when we know that F is stochastically dominated by G. What use could be made of that qualitative information?

We denote stochastic domination by $F \overset{\mathcal{D}}{\leq} G$; recall that this means

$$F(x) \geq G(x) \quad \text{for all } x \geq 0, \tag{1}$$

 0269-9648/00 $12.50

https://doi.org/10.1017/S0269964800141014 Published online by Cambridge University Press

which is equivalent to

$$\int f\,dF \leq \int f\,dG \tag{2}$$

for all increasing and bounded functions f on $[0,\infty)$.

Throughout, assume $F \overset{\mathcal{D}}{\leq} G$. It is common knowledge how to achieve $X \leq Y$: For a distribution function H on $[0,\infty)$, let H^* be the generalized inverse function defined by

$$H^*(u) = \inf\{s \geq 0; H(s) \geq u\}$$

for $0 < u < 1$. (Below, we make use once of the obvious extension of this definition to subprobability distribution functions.) Let

$$X = F^*(U) \quad \text{and} \quad Y = G^*(U), \tag{3}$$

where U is uniformly distributed on (0,1) ($U \overset{\mathcal{D}}{=} \mathrm{Uni}(0,1)$). We get $X \overset{\mathcal{D}}{=} F$ and $Y \overset{\mathcal{D}}{=} G$, and we certainly have $X \leq Y$ due to (1).

What further use could be made of the inequality $F \overset{\mathcal{D}}{\leq} G$? Let us set up a list of requests on (X,Y):

i. We have $X \leq Y$.
ii. In addition to (i), $\mathbb{P}(X = Y)$ is maximized.
iii. In addition to (ii), the pair (X,Y), conditioned that $X < Y$, has some desirable property.

We have managed request (i).

Section 2 handles request (ii). It is found, perhaps surprisingly, that the maximal possible value of $\mathbb{P}(X = Y)$ under the condition $X \leq Y$ is the same as the maximum without that restriction.

In Section 3, we present a case where request (iii) can be met: If F and G have densities f and g such that $f - g$ changes sign at most once, then (X,Y) may be produced so that X and Y become independent, conditioned that $X < Y$.

In Section 4, we point out that if F and G have failure rate functions r_1 and r_2, respectively, satisfying $r_1 \geq r_2$, then the standard method to produce $X \overset{\mathcal{D}}{=} F$ and $Y \overset{\mathcal{D}}{=} G$ by a thinning of a Poisson process yields $X \leq Y$ if a common Poisson process is used, but it does not satisfy (ii).

There are several fine accounts on simulation, among them Devroye [1] and Ripley [4], but surprisingly little attention has been paid to our type of topics. One exception is Devroye [2]; in Section 2 of that paper, the problem of how to produce (X,Y) with $\mathbb{P}(X = Y)$ maximized is treated, but with no attention paid to extra conditions, such as $F \overset{\mathcal{D}}{\leq} G$.

2. THE MAXIMUM OF $\mathbb{P}(X = Y)$

Let F and G have densities f and g, respectively, with respect to the measure λ. We may always take $\lambda = F + G$; the reader who wishes to restrict attention to F and G

https://doi.org/10.1017/S0269964800141014 Published online by Cambridge University Press

with densities with respect to the Lebesque measure does not need to pay attention to the choice of λ.

The total variation distance $\|F-G\|$ between F and G is the L^1 distance between f and g:

$$\|F - G\| = \int |f-g|\,d\lambda. \tag{4}$$

We have

$$\begin{aligned}\int &|f-g|\,d\lambda \\ &= \int (f - f\wedge g)\,d\lambda + \int (g - f\wedge g)\,d\lambda \\ &= 2\left(1 - \int f\wedge g\,d\lambda\right).\end{aligned} \tag{5}$$

Now the coupling inequality (cf. Lindvall [3, Sect. I.2]) tells us that

$$\|F-G\| \le 2\mathbb{P}(X \ne Y) \tag{6}$$

for any (X,Y) such that $X \overset{\mathcal{D}}{=} F$ and $Y \overset{\mathcal{D}}{=} G$. Hence,

$$\mathbb{P}(X=Y) = 1 - \mathbb{P}(X\ne Y) \le \int f\wedge g\,d\lambda. \tag{7}$$

Denote $\int f\wedge g\,d\lambda$ by γ. Throughout, we assume that $\gamma < 1$; since $\gamma = 1$ means $F = G$, we may exclude that case. Any coupling (X,Y) satisfying

$$\mathbb{P}(X=Y) = \gamma \tag{8}$$

is called a γ-coupling (cf. Lindvall [3, Sect. I.5], an account in which the definition of γ-coupling is more restrictive than the one here). We understand from (7) that γ is the best possible coupling probability.

Let H denote the subprobability measure with density $f\wedge g$ with respect to λ:

$$H(A) = \int_A f\wedge g\,d\lambda \quad \text{for Borel sets } A \subset [0,\infty).$$

The total mass of H is, of course, equal to γ. Without much hesitation, we write $H(x) = H[0,x]$ for $x \ge 0$, and we find that in order to generate (X,Y) such that $\mathbb{P}(X=Y)=\gamma$ with the use of $U \overset{\mathcal{D}}{=} \text{Uni}(0,1)$, it is crucial that the following hold:

$$\text{if } U < \gamma, \text{ then } X = Y = H^*(U) \tag{9}$$

and

$$\text{if } U > \gamma, \text{ see to it that } X \overset{\mathcal{D}}{=} F \text{ and } Y \overset{\mathcal{D}}{=} G. \tag{10}$$

https://doi.org/10.1017/S0269964800141014 Published online by Cambridge University Press

The last requirement is met by, inter alia, the following rule:

$$\text{if } U > \gamma, \text{ then } X = (F-H)^*(U-\gamma)$$

and

$$Y = (G-H)^*(U-\gamma). \tag{11}$$

To verify that, let U be defined as the identity mapping on (0,1) and let l be the Lebesque measure restricted to that interval. For example, for X given by (10) and (11), we get for any $x \geq 0$ that

$$\begin{aligned}\mathbb{P}(X \leq x) &= \mathbb{P}(H^*(U) \leq x, U < \gamma) + \mathbb{P}((F-H)^*(U-\gamma) \leq x, U > \gamma)\\ &= l\{u; u < \gamma, H^*(u) \leq x\} + l\{u; u > \gamma, (F-H)^*(u-\gamma)\}\\ &= l\{u; u < \gamma, H^*(u) \leq x\} + l\{v; 0 < v < 1-\gamma, (F-H)^*(v) \leq x\}\\ &= H(x) + (F-H)(x) = F(x).\end{aligned}$$

Any simulation satisfying (9) and (10) yields a pair (X,Y) such that $\mathbb{P}(X = Y)$ is maximized. Now recall request (ii). The assumption $F \overset{\mathcal{D}}{\leq} G$ implies that $F(x) - H(x) \geq G(x) - H(x)$ for $x \geq 0$, and we get $(F-H)^* \leq (G-H)^*$. Hence, rule (11) renders $X \leq Y$, and we have managed request (ii).

3. THE THIRD REQUEST

In order to explore the possibilities beyond request (ii), we reconsider Eqs. (9) and (10). Put

$$\begin{aligned}M &= \{(x,y); 0 \leq x \leq y\},\\ M_0 &= \{(x,y); 0 \leq x < y\},\end{aligned}$$

and let P be the distribution of (X,Y) satisfying Eqs. (9) and (10). That P is a probability measure on M, and its restriction, Q say, to $\Delta = \{(x,x); x \geq 0\}$ is determined by (9). We have $Q(\Delta) = \gamma$.

Let $P_0 = P - Q$. We find that (10) means

$$\begin{aligned}&\text{the subprobability } P_0 \text{ on } M_0 \text{ has to have marginals}\\ &\text{with densities } f - f \wedge g \text{ and } g - f \wedge g \text{ w.r.t. } \lambda.\end{aligned} \tag{12}$$

For a hint to understand that, let A be a Borel set $\subset [0,\infty)$. We get

$$\begin{aligned}\int_A f(x)\, d\lambda(x) &= \mathbb{P}(X \in A)\\ &= \mathbb{P}(X \in A, X = Y) + \mathbb{P}(X \in A, X \neq Y)\\ &= H(A) + \int_{x \in A} dP_0(x,y)\\ &= \int_A (f \wedge g)(x)\, d\lambda(x) + \int_{x \in A} dP_0(x,y);\end{aligned}$$

hence,

$$\int_{x\in A} dP_0(x,y) = \int_A (f - f \wedge g)(x)\, d\lambda(x).$$

Denote the marginals of P_0 by ν and μ. When is it possible to let

$$P_0 = (\nu \times \mu)/(1-\gamma),$$

that is, when can we allow X and Y to be independent, conditioned that $X < Y$? Since P_0 is concentrated to M_0, we have that possibility if and only if $\nu(x,\infty)\cdot\mu[0,x] = 0$ for all $x \geq 0$, and this holds

$$\text{if and only if there exists an } a \geq 0 \text{ such that} \quad f(x) \geq g(x) \text{ for } x \leq a \text{ and } f(x) \leq g(x) \text{ for } x > a \tag{13}$$

due to (12).

We comply with (10) using the rule

$$\text{if } U > \gamma, \text{ then } X = K_1^*(U-\gamma) \text{ and } Y = K_2^*((1-\gamma)\cdot U'), \tag{14}$$

where $K_1(x) = \int_0^x (f-g)^+(s)\, d\lambda(s)$, $K_2(x) = \int_0^x (g-f)^+(s)\, d\lambda(s)$, and $U' \stackrel{\mathcal{D}}{=} \mathrm{Uni}(0,1)$ is independent of U.

In Figure 1a and 1b, we show the outcomes of simulations with $X \stackrel{\mathcal{D}}{=} \mathrm{Uni}(0,1)$ and $Y \stackrel{\mathcal{D}}{=} \mathrm{Uni}(0,1.25)$; $n = 20$. In Figure 1a, X_i and Y_i are independent; for Figure 1b we used Eq. (3), which renders $Y_i = 1.25\cdot X_i$. The simulations shown in Figure 1c and 1d demonstrate the γ-coupling. For (c), we used Eq. (11). Condition (13) is satisfied; hence, we may let X_i and Y_i be independent, conditioned that $X_i < Y_i$. Diagram (d) shows the outcome of such a simulation.

4. ON THE METHOD OF POISSON PROCESS THINNING

We are able to be brief in this section concerning the background theory since that is rather well established; for accounts, see Ripley [4, §4.3] and, in particular, Devroye [1, Chap. VI].

Let $\hat{N}$ be a Poisson process on $[0,\infty)$ with intensity $\lambda > 0$; $\hat{N}(A)$ is the number of points in the Borel set $A \subset [0,\infty)$. With $\hat{T}_n$ denoting the time of the nth point, we have $\hat{T}_n = \sum_1^n \hat{X}_i$, where $\hat{X}_1, \hat{X}_2, \ldots$, are i.i.d. and $\mathrm{Exp}(\lambda)$ distributed. To produce another Poisson process $\hat{N}'$ with intensity $\mu < \lambda$, we may use a thinning method. Each point of $\hat{N}$ is saved for $\hat{N}'$ with probability μ/λ, and the trials are independent for the different points. As is standard, let us abbreviate $\hat{N}[0,t]$ and $\hat{N}'[0,t]$ to $\hat{N}_t$ and $\hat{N}_t'$, respectively. This thinning method yields

i. $\hat{N}'(A) \leq \hat{N}(A)$ for all $A \subset [0,\infty)$.
ii. For a simulation, the program is very simple. **(15)**

Let $\hat{T}_n' = \sum_1^n \hat{Y}_i'$ be the time of the nth point of $\hat{N}'$. Writing $\hat{X} = \hat{X}_1$ and $\hat{Y} = \hat{Y}_1$ for convenience, we have $\mathbb{P}(\hat{X} = \hat{Y}) = \mu/\lambda$. However, if we had used a γ-coupling to

https://doi.org/10.1017/S0269964800141014 Published online by Cambridge University Press

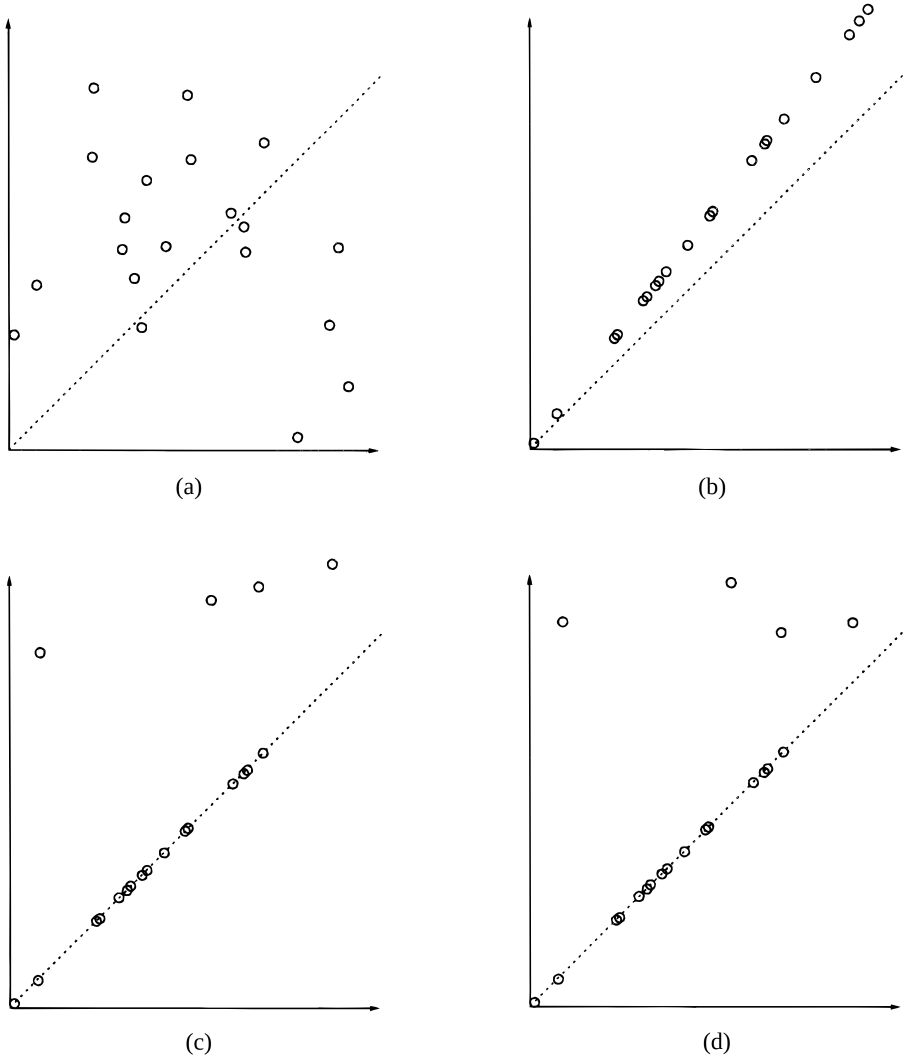

FIGURE 1. The simulations (a)–(d).

produce (X,Y), where $X \stackrel{\mathcal{D}}{=} \mathrm{Exp}(\lambda)$ and $Y \stackrel{\mathcal{D}}{=} \mathrm{Exp}(\mu)$, then we would get a larger probability. Indeed,

$$\mathbb{P}(X = Y) = \gamma = \int_0^\infty \min(\lambda e^{-\lambda x}, \mu e^{-\lambda x})\, dx > \mu/\lambda,$$

because $\min(\lambda e^{-\lambda x}, \mu e^{-\lambda x}) < (\mu/\lambda)\lambda e^{-\lambda x}$ for all $x > 0$, which is easily seen.

https://doi.org/10.1017/S0269964800141014 Published online by Cambridge University Press

Let $(X_i,Y_i), i = 1,2,\ldots,$ be i.i.d. γ-coupled pairs with $X_i \stackrel{\mathcal{D}}{=} \exp(\lambda)$ and $Y_i \stackrel{\mathcal{D}}{=} \exp(\mu)$; we notice that (13) holds, so we may let X_i and Y_i be independent, conditioned that $X_i < Y_i$.

We have found that natural requests on a simulation of a pair of Poisson processes are incompatible: If we prefer to use N and N', we lose 15(i) (but we still have that $N'_t \le N_t$ for all $t \ge 0$). Also, effort is needed for the programming.

Now, let F and G be life-length distributions with failure rate functions r_1 and r_2, respectively, satisfying $r_1(x) \ge r_2(x)$ for $x \ge 0$. Write $R_i(x) = \int_0^x r_i(s)\,ds$ for $i = 1$ and 2.

Using a well-known relation, we find that $F \stackrel{\mathcal{D}}{\le} G$:

$$\begin{aligned} F[x,\infty) &= \exp(-R_1(x)) \\ &\le \exp(-R_2(x)) = G[x,\infty) \quad \text{for } x \ge 0. \end{aligned}$$

For simplicity, let r_1 be bounded, and put $A = \sup_{x \ge 0} r_1(x)$. A well-established method to produce an $X \stackrel{\mathcal{D}}{=} F$ is as follows: Let $(T_i)_{i=1}^{\infty}$ be the points of a Poisson process with intensity A, and $(Z_i)_{i=1}^{\infty}$ a sequence of i.i.d. Uni$(0,A)$ variables, also independent of $(T_i)_{i=1}^{\infty}$. If we let

$$\hat{X} = \min\{T_i; Z_i \le r_1(T_i)\}, \qquad \textbf{(16)}$$

then indeed $\hat{X} \stackrel{\mathcal{D}}{=} F$. Using (16), but with r_1 replaced by r_2, we get a $\hat{Y} \stackrel{\mathcal{D}}{=} G$ and, certainly, $\hat{X} \le \hat{Y}$. By using a Poisson point process in $[0,\infty)^2$, a similar technique works for unbounded failure rate functions (cf. Lindvall [3, Sect. V.17]).

But in general this method does not maximize $\mathbb{P}(X = Y)$ among couplings such that $X \le Y$. We have understood that for exponentially distributed variables. Write $R_i(x) = \int_0^x r_i(s)\,ds$ for $i = 1$ and 2. For $r_1 \ge r_2$ and $r_1 \not\equiv r_2$ we get

$$\mathbb{P}(\hat{X} = \hat{Y}) = \int_0^\infty r_2(x)e^{-R_1(x)}\,dx$$

and this probability is strictly smaller than that of a γ-coupling (X,Y). Indeed,

$$\begin{aligned} \mathbb{P}(X = Y) &= \int_0^\infty \min(r_1(x)e^{-R_1(x)}, r_2(x)e^{-R_2(x)})\,dx \\ &\ge \int_0^\infty \min(r_1(x)e^{-R_1(x)}, r_2(x)e^{-R_1(x)})\,dx \\ &> \int_0^\infty r_2(x)e^{-R_1(x)}\,dx. \end{aligned}$$

References

1. Devroye, L. (1986). *Non-uniform random variate generation*. New York: Springer-Verlag.
2. Devroye, L. (1990). Coupled samples in simulation. *Operations Research* 38(1): 115–126.
3. Lindvall, T. (1992). *Lectures on the coupling method*. New York: Wiley.
4. Ripley, B. (1987). *Stochastic simulation*. New York: Wiley.