Online Supplement to "A Study on the Cross-Entropy Method for Rare-Event Probability Estimation"

INFORMS Journal on Computing

Tito Homem-de-Mello
Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, Illinois 60208, USA, tito@northwestern.edu

---

Proof of Proposition 3. Let L_p be an arbitrary path. As mentioned in the paper, L_p consists of S vertical arcs and $J-1$ horizontal ones. Notice that the dth arc down in L_p corresponds to some random variable Y_{dj}. By the assumption that the random variables $\{Y_{sj}\}_{j=1,\ldots,J}$ are identically distributed, we have that

$$L_p \stackrel{d}{=} \sum_{s=1}^{S} Y_{s1} + \sum_{j=1}^{J-1} Y_{s_j j},$$

where $s_1, \ldots, s_{J-1}$ is some sequence of indices in $\{1, \ldots, S\}$ and $\stackrel{d}{=}$ denotes equality in distribution. Now, the assumption of stochastic dominance implies that

$$L_p \leq_{st} \sum_{s=1}^{S} Y_{s1} + \sum_{j=1}^{J-1} Y_{s_{max} j}. \tag{A–1}$$

Finally, let $L_{p_{max}}$ denote the path constructed in the following way: go down $s_{max}-1$ arcs, then go right $J-1$ arcs, and then go down $S+1-s_{max}$ arcs. Then, $L_{p_{max}}$ is equal in distribution to the right-hand side of (A–1).

We shall prove next that (22) in the paper holds. Under the assumptions of the proposition, we have that $L_p \leq_{st} L_{p_{max}}$ for all $p = 1, \ldots, T$, i.e., $P(L_p \geq x) \leq P(L_{p_{max}} \geq x)$ for all $p = 1, \ldots, T$ and all x. Since $C_{SJ} = \max_{p=1,\ldots,T} L_p$ we have that $P(C_{SJ} \geq x) \geq P(L_{p_{max}} \geq x)$. Moreover,

$$P(C_{SJ} \geq x) \;=\; 1 - P(C_{SJ} < x) \;=\; 1 - P(L_p < x,\; p = 1, \ldots, T). \tag{A–2}$$

Using Bonferroni's inequality, we have

$$P(L_p < x,\; p = 1, \ldots, T) \;\geq\; 1 - \sum_{p=1}^{T} P(L_p \geq x).$$

By substituting the above expression into (A–2) and using the fact that $P(L_p \geq x) \leq P(L_{p_{max}} \geq x)$ we obtain the bounds

$$P(L_{p_{max}} \geq x) \;\leq\; P(C_{SJ} \geq x) \;\leq\; T\,P(L_{p_{max}} \geq x).$$

A combinatorial argument shows that the total number of feasible paths is given by $T = \binom{S+J-2}{J-1}$.

The final assertion of the proposition follows immediately from the construction of $L_{p_{max}}$.

■

Table A–1: Progression of the Algorithm for the Discrete-Distribution Case, $J = 4$, $S = 3$, Data Set as in Table 5 in the Paper

k	$\widehat{\gamma}^k$	(s,j)	$\widehat{p}^k_{sj,1}$	$\widehat{p}^k_{sj,2}$	$\widehat{p}^k_{sj,3}$	$\widehat{p}^k_{sj,4}$
		(1,1)	0.000	0.000	0.218	0.782
		(2,1)	0.019	0.275	0.127	0.579
		(3,1)	0.131	0.338	0.042	0.489
		(1,2)	0.210	0.062	0.220	0.508
		(2,2)	0.016	0.272	0.107	0.605
		(3,2)	0.110	0.283	0.035	0.572
1	216	(1,3)	0.262	0.077	0.241	0.420
		(2,3)	0.016	0.294	0.139	0.551
		(3,3)	0.079	0.202	0.032	0.687
		(1,4)	0.300	0.088	0.262	0.350
		(2,4)	0.024	0.318	0.132	0.527
		(3,4)	0.000	0.032	0.030	0.939
2	236			...		
		(1,1)	0.000	0.000	0.000	1.000
		(2,1)	0.000	0.000	0.000	1.000
		(3,1)	0.131	0.338	0.042	0.489
		(1,2)	0.132	0.039	0.139	0.690
		(2,2)	0.000	0.000	0.000	1.000
		(3,2)	0.110	0.283	0.035	0.572
3	237	(1,3)	0.262	0.077	0.241	0.420
		(2,3)	0.000	0.000	0.000	1.000
		(3,3)	0.079	0.202	0.032	0.687
		(1,4)	0.300	0.088	0.262	0.350
		(2,4)	0.000	0.000	0.000	1.000
		(3,4)	0.000	0.000	0.000	1.000