Improved Simulation Methods for System Reliability Estimation

by

Arne B. Huseby, Univ. of Oslo

Crude Monte Carlo

Random vector: $\mathbf{X} = (X_1, \ldots, X_n)$

Goal: Estimate $E[\phi] = E[\phi(\mathbf{X})] = h$

Simulation data: $\mathbf{X}_1, \ldots, \mathbf{X}_N$
(sampled from the distribution of **X**)

Monte Carlo (unbiased) estimate:

$$\hat{h}_{MC} = \frac{1}{N}\sum_{r=1}^{N}\phi(\mathbf{X}_r) \qquad \mathrm{Var}(\hat{h}_{MC}) = \mathrm{Var}(\phi)/N$$

Conditional MC

$S = S(\mathbf{X})$, a discrete valued variable with known distribution, and values in the set $\{s_1, \dots, s_k\}$.

Introduce for $j = 1, \dots, k$:

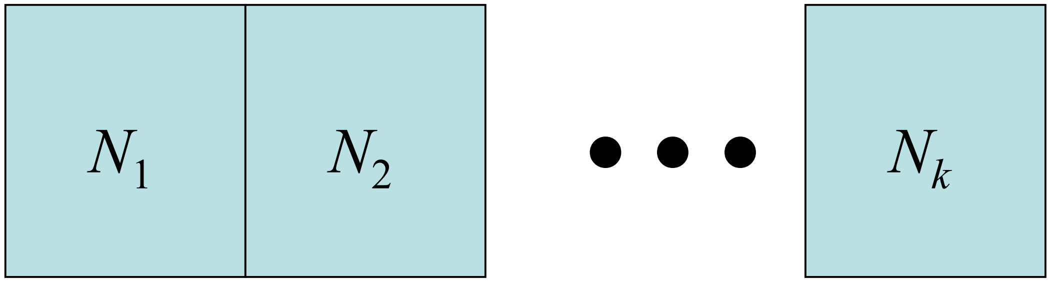
$$\theta_j = \mathrm{E}\left[\phi \middle| S = s_j\right]$$

We can then write:

$$h = \sum_{j=1}^{k} \theta_j \Pr\left(S = s_j\right)$$

Partitioning the MC Sample

We then partition the MC-sample set into k groups:



such that: $N = N_1 + \cdots + N_k$

The samples in the j-th group are sampled from the conditional distribution of $\mathbf{X}$ given that $S = s_j$, and are used to estimate θ_j, $j = 1, \ldots, k$.

The Combined Estimate

We denote the data in the j-th group by $\{\mathbf{X}_{r,j} : r = 1, \dots, N_j\}$. θ_j is then estimated by the unbiased Monte Carlo estimate:

$$\hat{\theta}_j = \frac{1}{N_j} \sum_{r=1}^{N_j} \phi\left(\mathbf{X}_{r,j}\right)$$

and by combining these estimates we get the conditional MC estimate:

$$\hat{h}_{CMC} = \sum_{j=1}^{k} \hat{\theta}_j \Pr\left(S = s_j\right)$$

with variance:

$$\mathrm{Var}\left(\hat{h}_{CMC}\right) = \sum_{j=1}^{k} \mathrm{Var}\left(\phi \middle| S = s_j\right)\left[\Pr\left(S = s_j\right)\right]^2 / N_j$$

Sample Group Sizes

Assume e.g., that the sample groups are sized so that:

$N_j \approx N \cdot \Pr(S = s_j),\ j = 1, \ldots, k.$

With this choice we get the improved variance:

$$\mathrm{Var}\left(\hat{h}_{CMC}\right) \approx \sum_{j=1}^{k} \mathrm{Var}\left(\phi \middle| S = s_j\right) \Pr\left(S = s_j\right) / N = \mathrm{E}\left[\mathrm{Var}\left(\phi \middle| S\right)\right] / N$$

$$= \left\{\mathrm{Var}\left(\phi\right) - \mathrm{Var}\left[\mathrm{E}\left(\phi \middle| S\right)\right]\right\} / N \leq \mathrm{Var}\left(\phi\right) / N = \mathrm{Var}\left(\hat{h}_{MC}\right)$$

How to choose S

We observe that the conditional estimate has smaller variance than the crude MC estimate provided that:

$$\mathrm{Var}[E(\phi \mid S)] > 0$$

This quantity can be interpreted as a measure of how much information S contains about ϕ.

Thus, when looking for good choices for S we should look for variables containing as *much information* about ϕ as possible.

Restrictions on S

- S must have a distribution that can be derived in polynomial time
- The number of possible values of S, i.e., k, must be bounded by a polynomial in n
- It must be possible to sample efficiently from the distribution of $\mathbf{X}$ given S

System Reliability

We now assume that **X** is a vector of independent bernoulli variables, the component state vector relative to a system of n components. The function, ϕ is also bernoulli and is referred to as the structure function of the system.

If a state variable, X_i, is 1, the corresponding component is functioning, otherwise it is failed. The size of a component state vector is defined as the number of '1's in the vector.

If the structure function ϕ is 1 the system is functioning, otherwise it is failed.

The expection, h, is called the system reliability and is equal to $\Pr(\phi = 1)$.

Some possible choices for S

- $S = (X_1, \ldots, X_m)$, where m is a suitable fixed number
- $S = (\phi_L, \phi_U)$, where ϕ_L and ϕ_U are two simpler structure functions

  (i.e., the expectations of these two functions can be evaluated in polynomial time)

  such that $\phi_L \leq \phi \leq \phi_U$
- $S = X_1 + \cdots + X_n$

In the remaining part of this talk, we assume that S is chosen as the sum of the component states.

Conditioning on the Sum

Introduce:

$p_i = \Pr(X_i = 1),\ i = 1, \ldots, n$

$$S_m = \sum_{i=m}^{n} X_i,\ m = 1, \ldots, n$$

(so $S = S_1$, and $S_n = X_n$)

The distributions of $S_1, \ldots, S_n$ can be calculated in $O(n^2)$-time using the following recursive relation:

$$\Pr(S_m = s) = p_m \Pr(S_{m+1} = s-1) + (1 - p_m) \Pr(S_{m+1} = s)$$

Sampling the X_i's given S

Step 1. Sample X_1 from the cond. distrib. of $X_1 \mid S = s$

Step 2. Sample X_2 from the cond. distrib. of $X_2 \mid S = s, X_1 = x_1$

.....

In the *m*-th step we use the following formula:

$$\Pr\left(X_m = x_m \middle| X_1 = x_1, \ldots, X_{m-1} = x_{m-1}, S = s\right) = \frac{\Pr\left(X_m = x_m, S = s \middle| X_1 = x_1, \ldots, X_{m-1} = x_{m-1}\right)}{\Pr\left(S = s \middle| X_1 = x_1, \ldots, X_{m-1} = x_{m-1}\right)}$$

$$= \frac{\Pr(X_m = x_m, S_{m+1} = s - \sum_{j=1}^{m} x_j)}{\Pr(S_m = s - \sum_{j=1}^{m-1} x_j)} = \frac{p_m^{x_m}(1 - p_m)^{1 - x_m} \Pr(S_{m+1} = s - \sum_{j=1}^{m} x_j)}{\Pr(S_m = s - \sum_{j=1}^{m-1} x_j)}$$

The Reliability Polynomial

We now consider the special case where all the component reliabilities are equal, i.e., $p_1 = p_2 = \cdots = p_n = p$, and introduce:

b_s = The number of path vectors of size s, $s = 0, \ldots, n$.

The system reliability, expressed as a function of p is then given by:

$$h(p) = \sum_{s=0}^{n} \theta_s \binom{n}{s} p^s (1-p)^{n-s}$$

where:

$$\theta_s = \mathrm{E}[\phi | S = s] = b_s \Big/ \binom{n}{s}, \quad s = 0, 1, \ldots, n.$$

Two Observations

- The unknown quantities, $\theta_0, \ldots, \theta_n$ do not depend on p. Thus, by estimating these quantities, we get an estimate of the entire $h(p)$-function.

- The quantity θ_s can be interpreted as the fraction of path vectors among all state vectors of size s, $s = 0, \ldots, n$. Thus, θ_s can be estimated by sampling random vectors of size s and calculating the frequency of path vectors among the sampled vectors.

Improved Sampling Method

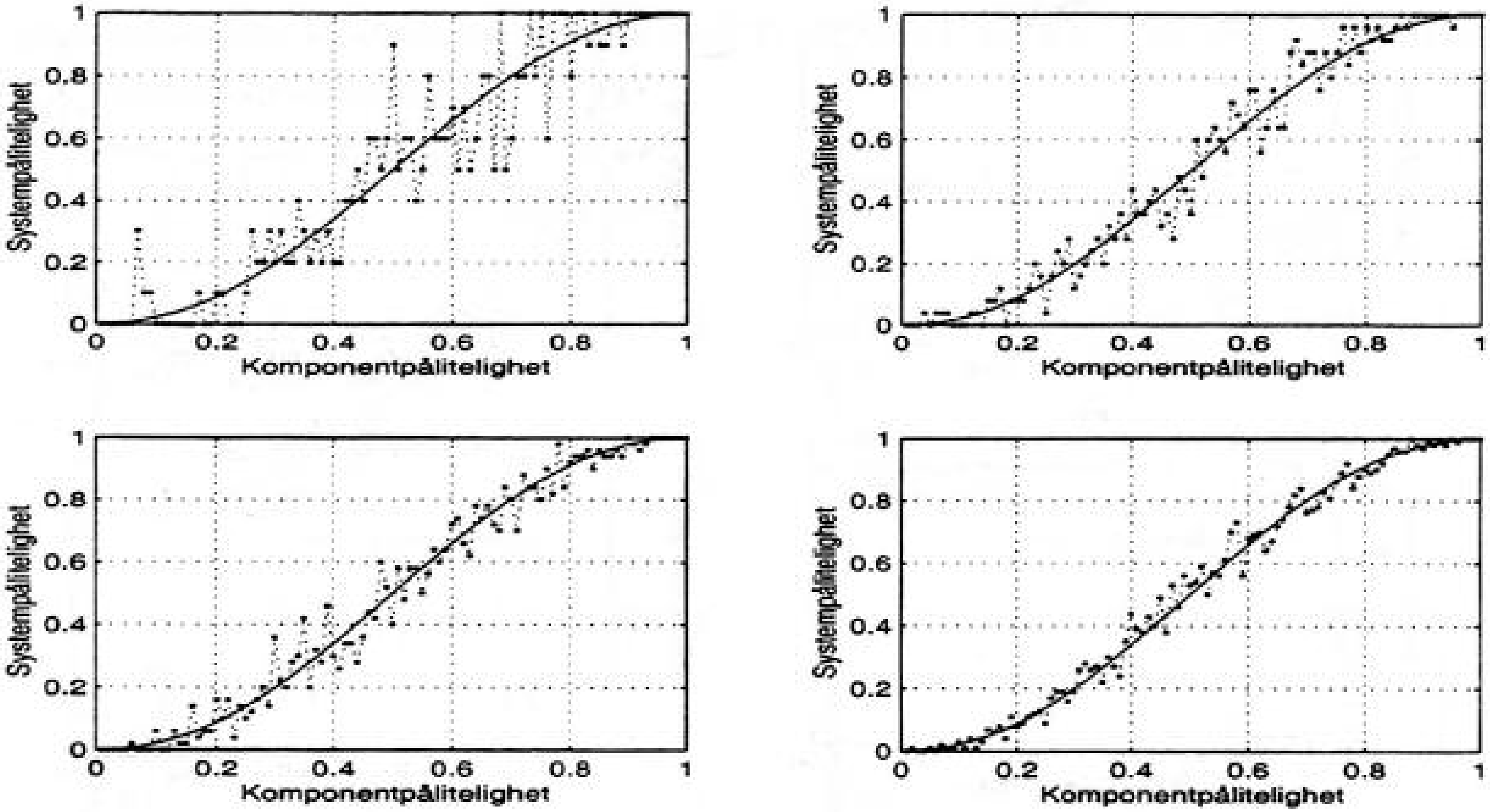
Step 1. Generate a random permutation $(i_1, \ldots , i_n)$ of the component set $\{1, \ldots , n\}$

Step 2. Define an increasing sequence of component state vectors, $\mathbf{X}_0, \mathbf{X}_1, \ldots , \mathbf{X}_n$, where: $\mathbf{X}_0 = (0, \ldots , 0)$, and where $\mathbf{X}_j$, is obtained from $\mathbf{X}_{j-1}$, by changing entry i_j from 0 to 1, $j = 1, \ldots , n$.

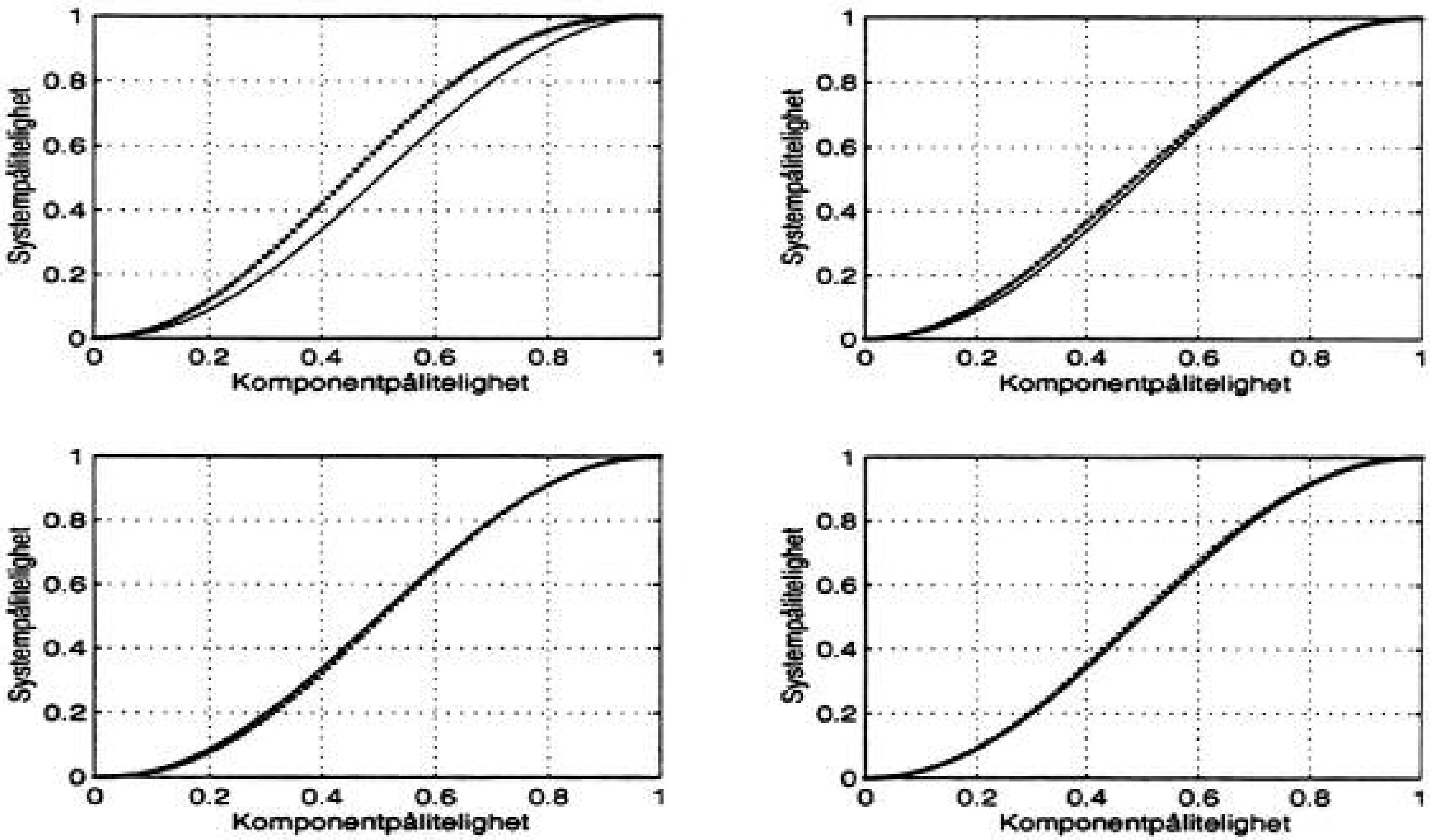
It is easy to verify that the vector $\mathbf{X}_s$ generated this way is uniformly distributed on the set of all state vectors of size s.

By generating N random permutations, we get N sequences of vectors, and by counting the number of path vectors among these, we get unbiased (but correlated) estimates for $\theta_0, \ldots , \theta_{n.}$

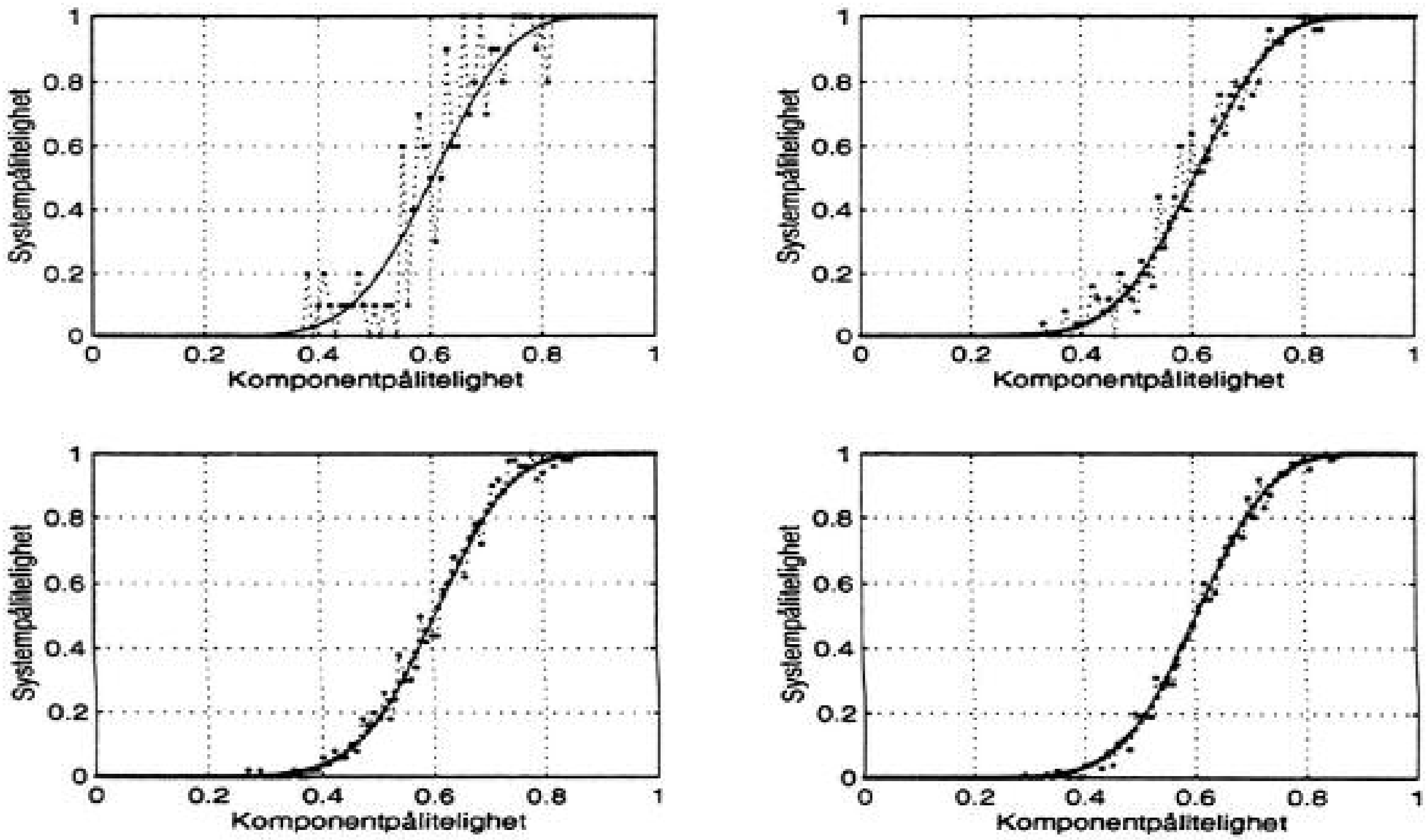
Crude MC estimates 1



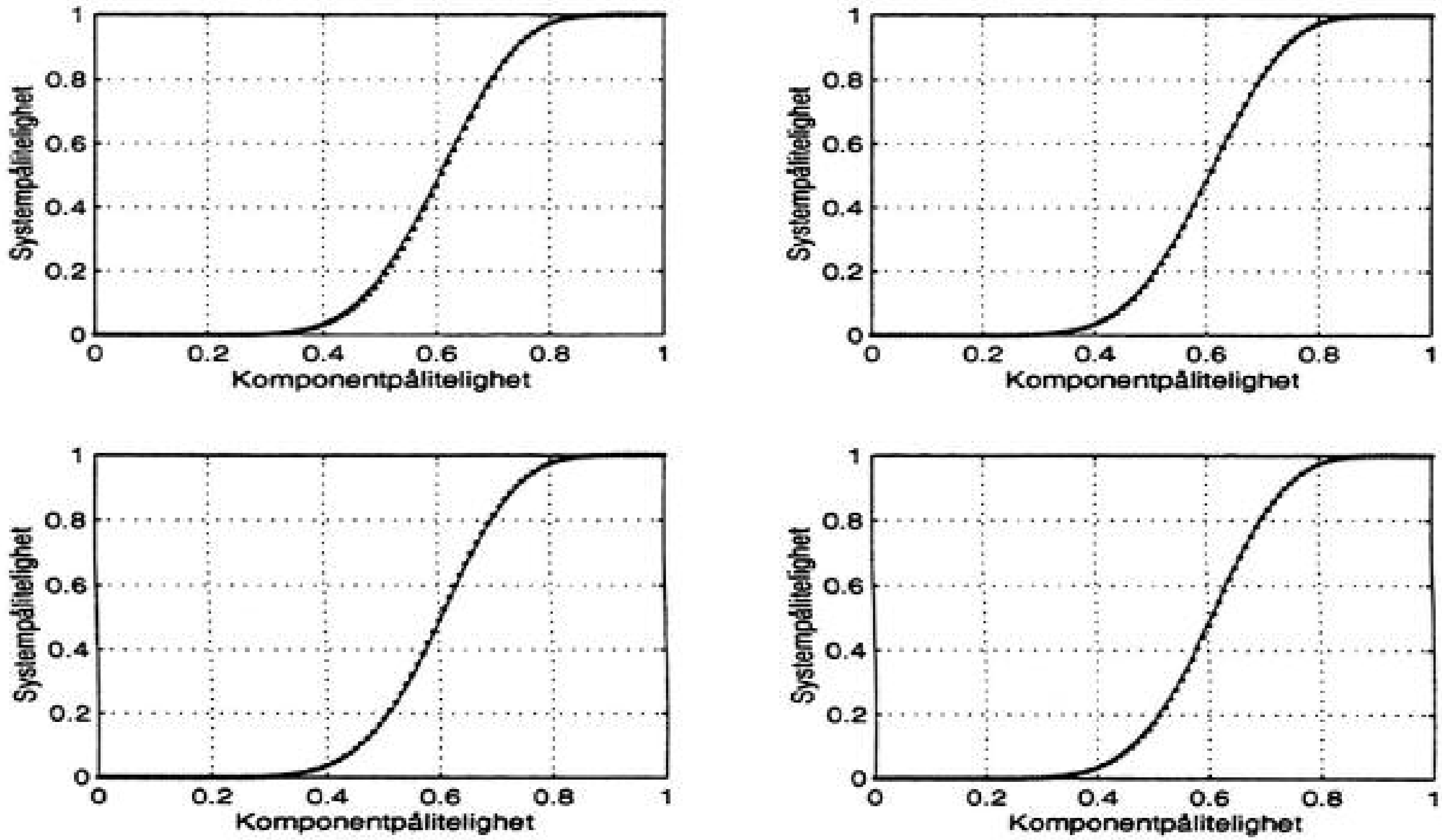
Improved estimates 1



Crude MC estimates 2



Improved estimates 2



Two Methods:

- **Direct Sampling Method:** Partition the N samples into (n-1) groups. In the j-th group generate state vectors $\mathbf{X}_r$ such that $\mathbf{X}_r \sim (\mathbf{X} \mid S = j)$, $j = 0, 1, \ldots, n$.

$$\hat{\theta}_j = \sum_{r=1}^{N_j} \phi(\mathbf{X}_r) \Big/ N_j$$

- **Sequential Sampling Method:** In *each* sample generate an increasing sequence of state vectors $\mathbf{X}_{0,r}$, $\mathbf{X}_{1,r}$, ... , $\mathbf{X}_{n,r}$, such that $\mathbf{X}_{j,r} \sim (\mathbf{X} \mid S = j)$, $j = 0, 1, \ldots, n$.

$$\hat{\theta}_j = \sum_{r=1}^{N} \phi(\mathbf{X}_{j,r}) \Big/ N$$

Advantages /disadvantages

- **Advantages with sequential sampling:**
  - More information is gathered from each sample
  - Since the sequences of state vectors are increasing, in each sample we only need to evaluate the system state until we have generated a state vector, $\mathbf{X}_{j,r}$, which is a path vector (since each of the following state vectors must be path vectors as well)
- **Disadvantage with sequential sampling:**
  - The estimates of the θ_j's are correlated. These correlations increase the variance of the reliability estimate

Advantages /disadvantages (cont.)

- Simulation studies indicate that the advantages with the sequential sampling methods are more important than the disadvantage
- In the case where all the component reliabilities are equal, an exact sequential sampling method is obtained by generating random permutations of the index set. Such permutations can be generated *very fast*, i.e., in $O(n)$ time.
- What if the component reliabilities are unequal??

Sampling the X_i's given S (revisited)

In each sample:

1. Start out by generating n uniform variables, $U_1, \ldots, U_n$.
2. The state variables $X_1, \ldots, X_n$ are then calculated recursively by using the following formula:

$$X_m = \mathrm{I}\left(U_m \leq p_m \frac{\Pr(S_{m+1} = s - \sum_{j=1}^{m-1} X_j - 1)}{\Pr(S_m = s - \sum_{j=1}^{m-1} X_j)} \right)$$

$$= \mathrm{I}\left(U_m \leq \mu_m\left(s - \sum_{j=1}^{m-1} X_j\right)\right), \quad m = 1, \ldots, n$$

Observation:

- When $s = 1$, the generated state vector will have size 1, i.e., exactly *one* of the state variables is 1 while the others are 0.
- When $s = 2$, the generated state vector will have size 2, i.e., exactly *two* of the state variables are 1 while the others are 0.
- .....

IDEA: For each set of uniform variables, $U_1, \ldots, U_n$, we may generate a sequence of state vectors, $\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_n$, by using increasing values of s.

Increasing sequence?

- In order to optimize the performance of the simulation algorithm, we want the generated sequences of state vectors to be increasing (to minimize the number of system state evaluations)

To show that this holds for the generated sequence, we must show that:

$$\mu_m(r) = p_m \frac{\Pr(S_{m+1} = r-1)}{\Pr(S_m = r)}$$

is nondecreasing in r for all m.

Increasing sequence? (cont.)

Equivalently we must show that:

$$\frac{\Pr(S_m = r)}{\Pr\left(S_{m+1} = r-1\right)} = \frac{p_m \Pr\left(S_{m+1} = r-1\right) + (1-p_m)\Pr\left(S_{m+1} = r\right)}{\Pr\left(S_{m+1} = r-1\right)}$$

is nonincreasing in r for all m.

Equivalently:

$$\frac{\Pr\left(S_{m+1} = r\right)}{\Pr\left(S_{m+1} = r-1\right)} \geq \frac{\Pr\left(S_{m+1} = r+1\right)}{\Pr\left(S_{m+1} = r\right)}$$

for all r and m where the denominators are nonzero.

Increasing sequence? (cont.)

A sufficient condition for this to hold is that:

$$\Pr\left(S_{m+1} = r-1\right)\Pr\left(S_{m+1} = r+1\right) \le [\Pr(S_{m+1} = r)]^2$$

for all *r* and *m*.

If an integer valued random variable satisfies this condition, its distribution is said to be *log-concave*.

FORTUNATELY: It can be shown that a sum of independent binary variables always has a log-concave distribution.

Final comments

- Using the log-concavity property, we are able to construct an exact sequential sampling method for the case where not alle the component reliabilities are equal.
- Still *finding* the sequence of state vectors, given the U_j's, is somewhat less efficient than in the case with equal component reliabilities
  - Best cases: $O(n \log n)$ versus $O(n)$
  - Worst cases: $O(n^2)$ versus $O(n)$