Theory

- Introduction to simulation
- Flow of simulation – generating process realizations
- Random number generation from given distribution
- Collection and analysis of simulation data
- Variance reduction techniques

Variance reduction techniques

- Simulation is a technique which often requires a lot of time / computational effort
- If the sampling can somehow be changed such that the variance of a sample becomes smaller,
  - one gets a more accurate estimate with a given number of repetitions
  - the number of repetitions required for a given accuracy is reduced (inversely proportional to the reduction in variance or squared reduction of the standard deviation)
- Methods by which this can be achieved are called variance reduction techniques
  - and are valuable provided that the technique itself does not require much additional computation
- Variance reduction techniques use
  - intelligence
  - known results or knowledge about the system

  to make such observations which have a smaller variance

Variance reduction techniques

- Use of common random number sequences
- Antithetic variables
- Control variables
- Conditioning – exploiting known expectations
- Importance sampling (relates mostly to static simulation; will not be covered here)
- RESTART

WARNING

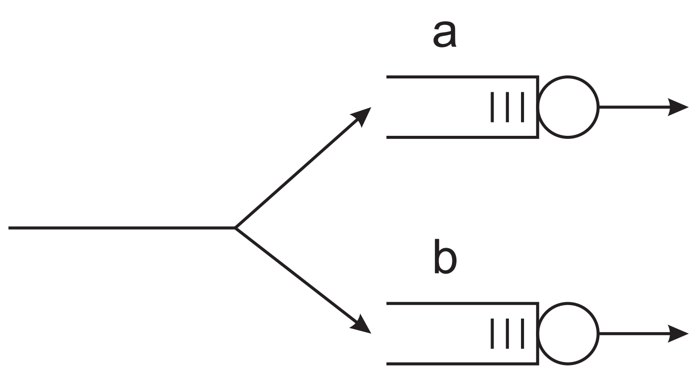
- Uneducated (erroneous) use of variance reduction techniques may have an effect which is opposite to the desired one, viz. may lead to an increase of variance

Use of common random number sequences

- Can be used when the purpose of simulation is to compare the performances of two similar systems
- In a typical case, one explores the behavior of the system using two different control or operation policies A and B, for example
  - the behavior of a queue with two different scheduling policies:
  - for instance, what is the difference in waiting time in the system of two queues depicted in the picture, when the arriving customers are directed to the queues a and b

    A) cyclically

    B) randomly, with a 50% probability to either queue a or queue b



Use of common random number sequences (continued)

- We are interested in the difference of the expectation of some performance parameter X in these two politics
$$\Delta\mu = \mu_A - \mu_B = \mathrm{E}[X^A] - \mathrm{E}[X^B]$$
where X^A and X^B denote the random variable X under the respective policies, and
$$\mu_A = \mathrm{E}[X^A] \qquad \mu_B = \mathrm{E}[X^B]$$

- In order to estimate the difference $\Delta\mu$ it is advantageous to simulate the system under policies A and B using exactly the same realizations for the arrival processes
  - if in the simulation of policy A the arrival pattern happens to contain a burst of arrivals, then the same burst occurs also when simulating policy B
  - the difference does not arise from random differences in the arrival patterns but only from the differences in the arrival processes

Use of common random number sequences (continued)

- In simulations which are repeated n times we get the samples

$$\begin{cases} X_1^A, X_2^A, X_3^A, \ldots, X_n^A \\ X_1^B, X_2^B, X_3^B, \ldots, X_n^B \end{cases}$$

- Under the policy A the sample average and sample variance are

$$\overline{X}_A = \frac{1}{n}\sum_{i=1}^{n} X_i^A \qquad S_A^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i^A - \overline{X}_A)^2$$

and similar equations hold for the policy B

- For the difference $\Delta\mu$ we get the estimator

$$\Delta\hat{\mu} = \overline{X}_A - \overline{X}_B$$

with the variance

$$\mathrm{V}[\Delta\hat{\mu}] = \mathrm{V}[\overline{X}_A] + \mathrm{V}[\overline{X}_B] - 2\mathrm{Cov}[\overline{X}_A, \overline{X}_B]$$

Use of common random number sequences (continued)

- The samples $X_1^A, X_2^A, X_3^A, \ldots, X_n^A$ are independent, so are the samples $X_1^B, X_2^B, X_3^B, \ldots, X_n^B$
- By using the same realizations for systems A and B, however, the samples X_i^A and X_i^B, $i = 1 \ldots n$, are correlated
  - the correlation is positive, because the same bursts or quiet times occur in both simulations
  - correspondingly the sample averages $\overline{X}_A$ and $\overline{X}_B$ are positively correlated
  - the variance of the estimator is smaller than in the case of using independent arrival realizations for policies A and B
- The estimator of the variance (of the estimator) is

$$\frac{1}{n}(S_A^2 + S_B^2 - 2S_{AB}^2)$$

where S_{AB} is the sample covariance

$$S_{AB}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i^A - \overline{X}_A)(X_i^B - \overline{X}_B)$$

Use of common random number sequences (continued)

- In the trivial special case, where the policies A and B are the same, the use of common random number sequences leads to the correct estimator for the difference $\Delta\hat{\mu} = 0$
  - while, when using independent realizations, the difference estimator usually is $\neq 0$

In comparing the performance of two different systems, the simulations should be made in conditions which are identical as far as possible

- The use of common random number sequences has a slight impact on the implementation of the simulation program
  - for instance, in comparing the FIFO and LIFO policies, the service times have to be drawn at the time a customer arrives at the system, while in the case of independent realizations, one can defer the drawing of the service times to the moment when the service starts

Antithetic variables

- The task is to estimate the expected value $\mu = \mathrm{E}[X]$ of the random variable X
- The idea: make two different runs such that "small" values of the observed variable in one run are replaced by "large" values of this variable in the other run and vice versa, and use the average over the two runs as the estimator
- Let X be the observed value of the studied variable in one experiment
- Let Y be the result from another experiment having the same distribution as X
- Then $Z = \frac{1}{2}(X+Y)$ is a random variable, which has the same expectation as X

$$\mu = \mathrm{E}[Z]$$

- The variance of Z

$$\begin{aligned} \mathrm{V}[Z] &= \tfrac{1}{4}(\mathrm{V}[X] + \mathrm{V}[Y] + 2\mathrm{Cov}[X,Y]) \\ &= \tfrac{1}{2}(\mathrm{V}[X] + \mathrm{Cov}[X,Y]) \end{aligned}$$

- If X and Y are negatively correlated, $\mathrm{Cov}[X,Y] < 0$, then the variance of Z is smaller than in the case where X and Y are independent ($\frac{1}{2}\mathrm{V}[X]$)

Antithetic variables (continued)

- Let the cdf of X be $F(x)$ and let the samples of X be generated by the inverse transform method

$$X = F^{-1}(U)$$

  where $U \sim \mathrm{U}(0,1)$ (uniform distribution)

- Then Y, which is generated using the formula

$$Y = F^{-1}(1-U)$$

  obeys the same distribution as X

  - when $U \sim \mathrm{U}(0,1)$ then also $1-U \sim \mathrm{U}(0,1)$

- It is clear that X and Y are negatively correlated
  - if U happens to be small (≈ 0) then $1-U$ takes a large value (≈ 1)
  - X and Y take values at the opposite ends of the range
- X and Y constitute a *antithetic pair of random variables*

Antithetic variables (continued)

Example

- Let $X \sim \mathrm{Exp}(1)$, i.e.

$$X = -\log U \qquad Y = -\log(1-U)$$

- One can show that in this case $\mathrm{Cov}[X,Y] = 1 - \pi^2/6 = -0.645$
- Variance is reduced by 64 % in comparison with the use of independent random number sequences (and costs nothing!)
- If the expectation of X is to be estimated by simulations
  - which, of course, does not make sense in the problem of the example
  - it advantageous to use after each drawn random number U_i to use as the next random number the value $1 - U_i$

Antithetic variables (continued)

Example (continued)

i	X_i	Y_i	Z_i
1	2.3340	0.1019	1.2180
2	0.0569	2.8945	1.4757
3	1.1121	0.3988	0.7554
4	0.1235	2.1527	1.1381
5	2.2724	0.1088	1.1906
6	2.0813	0.1333	1.1073
7	0.4510	1.0132	0.7321
8	2.2548	0.1108	1.1828
9	0.1761	1.8231	0.9996
10	0.8221	0.5789	0.7005

Sample averages

$\bar{X}$	$\bar{Y}$	$\bar{Z}$
1.1684	0.9316	1.0500

Sample variances

S_X^2	S_Y^2	S_Z^2
0.9489	1.0256	0.0634

Antithetic variables (continued)

- In a repeated experiment one can generate values $X_1, X_2, \ldots$ with independent uniformly distributed random numbers $U_1, U_2, \ldots$
  - $X_1, X_2, \ldots$ are independent
- Similarly, one can generate a sequence of values $Y_1, Y_2, \ldots$ with independent uniformly distributed random numbers $1 - U_1, 1 - U_2, \ldots$
  - $Y_1, Y_2, \ldots$ are independent
- Then also the values $Z_i = (X_i + Y_i)/2,\ i = 1, 2, \ldots$ are independent
- In each pair, X_i and Y_i are negatively correlated
- The variance of Z_i is smaller than it would be if X_i and Y_i were independent
- Correspondingly the variance of the estimator

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} Z_i$$

  is smaller than in the case of independent random variables

Antithetic variables (continued)

- In the simulation of a system, one has to us the sequences $U_1, U_2, \ldots$ and $1 - U_1, 1 - U_2, \ldots$ have to be used for "the same purpose"
  - for instance, for the generation of service times
  - one has to be careful in implementing the program to get proper synchronization
- The amount by which the variance reduction is reduced is difficult to estimate in advance
  - in practice, one has to study this by test runs
- The computational cost of the use of antithetic variables is minimal, so give the method a chance!

Control variables

- In the method of control variables the idea is in a way opposite to that of antithetic variables: in this method one tries to utilize (strong) positive correlation between the observed variable X and some auxiliary variable Y, so called control variable
- A prerequisite for the method is that the expectation $\nu = \mathrm{E}[Y]$ of the control variable Y is known exactly
- The task is again to estimate the expectation $\mu = \mathrm{E}[X]$
- In the simulation one collects data for both X and Y
- With the aid of these one forms the actual observed variable

$$Z = X - (Y - \nu)$$

  which has the expected value

$$\mathrm{E}[Z] = \mathrm{E}[X] - (\mathrm{E}[Y] - \nu) = \mu$$

- By observing the values of Z in repeated simulations, one gets by their average an unbiased estimate for μ

Control variables (continued)

- The variance of Z is

$$V[Z] = V[X - Y] = V[X] + V[Y] - 2\mathrm{Cov}[X, Y]$$

- The variance is less than that of X provided that the correlation is strong enough

$$\mathrm{Cov}[X, Y] > V[Y]/2$$

Control variables: example

- Consider an M/M/1 queue which is empty at time $t = 0$
- Let X be the average waiting time of first 100 customers
- We take as the control variable Y the average service time of 99 first customers
  - because the service time of the 100^{th} customer has no bearing to X there is no idea to include it in the control variable
  - The expectation of Y $\nu = \mathrm{E}[Y]$ of the average service time is known with a given service time distribution
- If the service times are (happen to be) long, then one can anticipate that also the waiting times are long
  - thus, it is reasonable to assume that X and Y are positively correlated
- Y is presumably a good candidate for a control variable
  - this has to be checked experimentally
  - if the control variable is poorly chosen, this may result in variance increase instead of reduction

Control variables: in improved method

- The reduction can be improved by including an undetermined coefficient a in the definition of Z

$$Z = X - a(Y - \nu)$$

- Again we have $\mathrm{E}[Z] = \mu$
- The parameter a is determined by the condition that the variance of Z is minimized
- This leads to the optimal value of a, denoted by a^*,

$$a^* = \frac{\mathrm{Cov}[X, Y]}{\mathrm{V}[Y]}$$

- The minimized variance is

$$\mathrm{V}[Z] = \mathrm{V}[X] - \frac{\mathrm{Cov}[X, Y]^2}{\mathrm{V}[Y]} = (1 - \rho_{X,Y}^2)\mathrm{V}[X]$$

  where

$$\rho_{X,Y} = \frac{\mathrm{Cov}[X, Y]}{(\mathrm{V}[X]\mathrm{V}[Y])^{1/2}}$$

  is the correlation coefficient of X and Y

Control variables: in improved method (continued)

- In the improved method, the variance of the observed variable Z is always smaller than (or at most as great as) the variance of X
- Also, now it does not matter whether the correlation is negative or positive
  - in the case of negative correlation the sign of correcting term is changed, i.e. $a < 0$
- To determine the optimal value of a, one needs both $\mathrm{V}[Y]$ and $\mathrm{Cov}[X, Y]$
  - the values of these are usually not known in advance
  - the values have to be estimated by the sample variance and sample covariance in preliminary simulation runs

$$S_Y^2 = \frac{\sum_{i=1}^{n} (Y_i - \bar{Y})^2}{n-1} \qquad S_{X,Y}^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})}{n-1}$$

$$a^* \approx \frac{S_{X,Y}^2}{S_Y^2}$$

Control variable (continuation of the example)

i	X_i	Y_i
1	3.84	0.92
2	3.18	0.95
3	2.26	0.88
4	2.76	0.89
5	4.33	0.93
6	1.35	0.81
7	1.82	0.84
8	3.01	0.92
9	1.68	0.85
10	3.60	0.88

- X = the average waiting time of first 100 customers
- Y = the average service time of first 99 customers

$$\nu = \mathrm{E}[Y] = 0.9$$

- The average interarrival time 1 (load 0.9)
- Ten repetitions: $\bar{X} = 3.78, \quad \bar{Y} = 0.89$
- The sample variances and covariances:

$$S_X^2 = 13.33$$
$$S_Y^2 = 0.002$$
$$S_{X,Y}^2 = 0.07$$

$\Rightarrow \rho_{X,Y} = 0.43$

$\Rightarrow a^* = 30$

$\Rightarrow \bar{Z} = \bar{X} - a^*(\bar{Y} - \nu) = 4.13$

- This is precisely the correct value (while $\bar{X} = 3.78$)

Control variables (continued)

- The use of control variables may require some additional work (in particular, the use of the improved method) that the previously discussed reduction methods
  - the method, however, may be quite effective
- The control variable can also be taken from another, auxiliary system
- The problem may, for instance, be to study the average sojourn time in an M/M/1 queue under some advanced scheduling algorithm
  - then one can as the auxiliary system the ordinary M/M/1-FIFO queue
  - the control variable is now the sojourn time in this auxiliary system
  - its expectation is known, $1/(\mu - \lambda)$
  - the studied system and the auxiliary system are run with the same input streams (the same interarrival times and service times)

Conditioning

- The idea: replace a random variable by its conditional expectation
- The data analysis of a simulation usually means calculating the average value of some random variable X
  - sojourn time in a queue; the cumulative time the queue is empty etc; the cumulative time all servers are busy . . .
- In some cases, the conditional expectation of the observed random variable (conditioned on the value of some other random variable Y observed in the system) is known exactly
- In such cases it is advantageous to collect statistics, not from the variable X itself but from the conditional expectation $\mathrm{E}[X \mid Y]$
  - the observed variable is now $Z = \mathrm{E}[X \mid Y]$
  - this is a random variable, the values of which vary as Y varies
  - with a fixed value Y, it is a known number (assumed to be known)

Conditioning (continued)

- Identically holds

$$\mathrm{E}[Z] = \mathrm{E}[\mathrm{E}\,[X \,|\, Y]] = \mathrm{E}[X]$$

- On the other hand, there is a conditioning formula for variance

$$\mathrm{V}[X] = \mathrm{E}[\mathrm{V}[X \,|\, Y]] + \mathrm{V}[\mathrm{E}\,[X \,|\, Y]]$$

  from which

$$\mathrm{V}[Z] = \mathrm{V}[\mathrm{E}\,[X \,|\, Y]] = \mathrm{V}[X] - \mathrm{E}[\mathrm{V}[X \,|\, Y]]$$

- Variance is reduced
  - the conditioning eliminates the variability of X within all the cases where the conditioning variable Y has a fixed value

Conditioning: example 1

- A loss system receives call arrivals with a Poissonian intensity λ
- The problem is to estimate how many call per time unit are blocked on the average
- In a straight forward simulation one observes the development of the system over a period T and records the number of blocked calls L
  - the estimator for the desired quantity is L/T
- Because calls are blocked only when the system is in a blocking state, it holds

$$L = \sum_i L_i$$

where L_i is the number of blocked calls in the i^{th} blocking period

- Let the length of the i^{th} blocking period in the simulation be B_i
- Since we have assumed Poissonian arrivals, we have

$$\mathrm{E}\left[L_i \,|\, B_i\right] = \lambda B_i$$

Conditioning: example 1 (continued)

- Instead of recording the number of blocked calls, it is more advantageous to record the total time spent in the blocking state $\sum_i B_i$
- The conditional expectation of blocked calls is now

$$\bar{L} = \lambda \sum_i B_i$$

  ands the new estimator for the desired quantity is $\bar{L}/T$

- The flow of the simulation is not altered at all
  - now one only records instead of L_i the values of $\mathrm{E}[L_i \mid B_i] = \lambda B_i$
  - the variance caused by random generation of the arrival stream during the blocking period is now eliminated

Conditioning: example 2

- The problem is to estimate the average sojourn time in an */M/1-FIFO queue
- In a straight forward simulation one observes the development of the system until n customers have been served
  - and records the sojourn time $T_i,\ i = 1, 2, \ldots, n$ of each customer
  - and calculates the estimator $\frac{1}{n}\sum_i T_i$
- If upon the arrival of the i^{th} customer there are m_i customers in the queue, then conditioned on this number one has

$$\mathrm{E}\left[T_i \,|\, m_i\right] = (m_i + 1)/\mu$$

where $1/\mu$ is the average service time of one customer (note, memorylessness)

Conditioning: example 2 (continued)

- Instead of T_i is advantageous to take $\mathrm{E}[T_i \,|\, m_i]$ or essentially m_i as the observed variable
- One gets the estimator

$$\frac{1}{n}\sum_{i=1}^{n}(m_i+1)/\mu$$

- Again, the flow of simulation remains the same
  - each customer arriving in the queue sees the same queue length as before
  - now just the recording of the sojourn times does not contain the variability caused by the random drawing of the service times of the customers in the queue
  - all the service times have been replaced by their expectation $1/\mu$ (in the data collection, not in the actual simulation)

Conditioning (continued)

In data collection it is advantageous to replace, whenever possible, the observed variable by its conditional expectation (conditioned on the dynamic state of the system)

- A small reservation to the above: the variance reduction is not always guaranteed
  - samples of a variable observed in a simulation are not independent but usually they are positively correlated, which leads to a larger variance of their average than in the case of independent variables
  - the conditioning may make the positive correlation even stronger
  - in principle, it is possible that this effect overrides the variance reduction of a single sample (in practice, this seldom happens)

RESTART (Repetitive Simulation Trials After Reaching Thresholds)

- RESTART is an acceleration method intended for the simulation of rare events
- The basic idea of the method is simple and is easy to apply to any system (it does not require a complex analysis)
- This is also a kind of conditioning method
  - but now the conditional expectation is determined by simulations
- Assume that we are interested in a rare event A
  - foe instance, A may mean that a variable X_t exceeds some (high) level L

$$A = \{X_t > L\}$$

- Let $C \supset A$ be another, less rare event, for which it holds

$$1 \gg \mathrm{P}\{C\} \gg \mathrm{P}\{A\}$$

  - for instance, C may mean that the variable X_t exceeds some lower threshold T

$$C = \{X_t > T\}$$

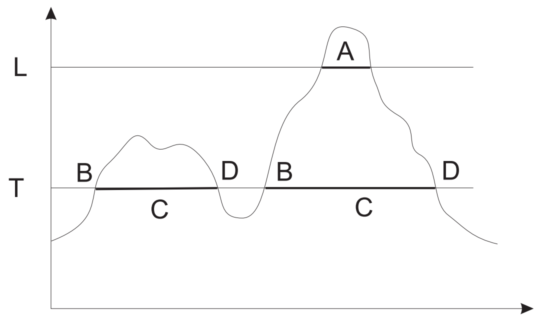
RESTART (continued)

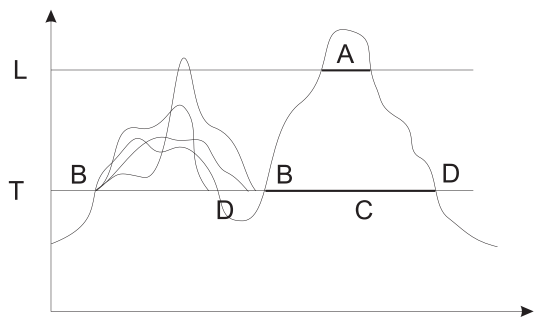
- By the definition of conditional probability

$$\mathrm{P}\{A\} = \mathrm{P}\{C\}\mathrm{P}\{A \,|\, C\}$$

- By an ordinary simulation it is easy to obtain an accurate estimate for the probability $\mathrm{P}\{C\}$
- But it is difficult to estimate the probability $\mathrm{P}\{A \,|\, C\}$, because C occurs relatively infrequently
- In the RESTART method the estimate for $\mathrm{P}\{A \,|\, C\}$ is improved by making several repetitions of those part of the process where C occurs
  - because C is relatively rare, the repetitions of the events C do not make the simulation too heavy

RESTART (continued)





Ordinary simulation (upper figure) and the RESTART method (lower figure)

– upon occurrence of B the whole state of the system is saved

– when the point D is reached, i.e. the event C is over, the state of the previous event B is restored and the period B–D is simulated anew

– this procedure is repeated n times

– after the n^{th} repetition the simulation is continued as usually from the point D on, until one reaches a new point B, from which one again makes n repetition to the next point D

– to observe the rare event A, those periods (period C) where it at all can occur are repeated n times

RESTART (continued)

- The RESTART method can lead to quite dramatic reductions in the simulation time
- If the system is complex, however, the storing and restoring of the system state at point B may represent a remarkable computational cost
  - the whole state of the system (including all the variables) has to be stored
- In order to further improve the method, one has developed variants, where several threshold levels and/or hysteresis are used

Importance sampling (IS) (1)

- Importance sampling is one of the most effective ways to reduce variance in the system
  - but it also requires the most careful analysis of the system
- Consider a discrete r.v. X with state space $\mathcal{S}$ and let $X \in \mathcal{S} \sim p(x)$
- Assume we are interested in

$$\alpha = \mathrm{E}[1(X \in \mathcal{A})] = \mathrm{P}\{X \in \mathcal{A}\}$$

- Our estimator is then

$$\hat{\alpha} = \frac{1}{N} \sum_{n=1}^{N} 1(X_n \in \mathcal{A}),$$

  where X_n are i.i.d.

Importance sampling (2)

- Let $X^* \sim p^*(x)$ denote another r.v. such that $p^*(x) > 0, \forall x \in \mathcal{A}$
- Then

$$
\begin{aligned}
\alpha &= \mathrm{E}[1(X \in \mathcal{A}] \\
&= \sum_{x \in \mathcal{S}} p(x)\, 1(x \in \mathcal{A}) \\
&= \sum_{x \in \mathcal{S}} p^*(x)\, \frac{p(x)}{p^*(x)} 1(x \in \mathcal{A}) \\
&= \mathrm{E}_{p^*}\left[w(X^*)1(X^* \in \mathcal{A}\right],
\end{aligned}
$$

where $w(x) = p(x)/p^*(x)$ is the so called *likelihood ratio*

- Hence, we have a new estimator

$$\hat{\alpha} = \sum_{n=1}^{N} w(X_n^*)\, 1(X_n^* \in \mathcal{A})$$

- IS allows one to generate samples from distribution $p^*(x)$ and the bias is corrected by weighting the samples appropriately.

Optimal IS (1)

- Assume that

$$p^*(x) = \mathrm{P}\{X = x \,|\, X \in \mathcal{A}\} = \frac{p(x)}{\mathrm{P}\{X \in \mathcal{A}\}}$$

- Then $w(x) = p(x)/p^*(x) = \mathrm{P}\{X \in \mathcal{A}\}$ and

$$\begin{aligned}\hat{\alpha} &= \frac{1}{N}\sum_{n=1}^{N} w(X_n^*)1(X_n^* \in \mathcal{A}) \\ &= \frac{1}{N}(N \cdot \mathrm{P}\{X \in \mathcal{A}\} = \mathrm{P}\{X \in \mathcal{A}\} \qquad \text{(exactly!)}\end{aligned}$$

- Problem: to compute $w(x)$ one must know $\mathrm{P}\{X \in \mathcal{A}\}$ which is exactly what we need to estimate.
- Idea: a good IS distribution $p^*(x)$ tries to approximate the optimal solution as much as possible, but at the same time keeping the likelihood ratio $w(x)$ computable.

Optimal IS (2)

- Now consider $\mathrm{V}[w(X^*)\,1(X^* \in \mathcal{A})]$. It can be expressed as

$$\begin{aligned}
\mathrm{V}[w(X^*)\,1(X^* \in \mathcal{A})] &= \frac{\alpha^2}{\alpha^*} - \alpha^2 + \alpha^*(\sigma^*)^2, \quad \text{where} \\
\alpha &= \mathrm{E}[1(X \in \mathcal{A})], \\
\alpha^* &= \mathrm{E}_{p^*}\left[1(X^* \in \mathcal{A})\right], \\
(\sigma^*)^2 &= \mathrm{V}_{p^*}\left[w(X^*) \,|\, X^* \in \mathcal{A}\right].
\end{aligned}$$

- Observations
  - $p^*(x) = p(x) \Rightarrow \mathrm{V}[*] = \alpha - \alpha^2$
  - Increase $\alpha^* \Rightarrow \alpha^2/\alpha^*$ becomes smaller
  - Ideally, if $\alpha^* = 1$ and $\sigma^* = 0 \Rightarrow \mathrm{V}[*] = 0$
  - In practice, one tries to increase α^* and make sure that σ^* does not grow too large

Applications of IS

- Static simulation
  - The previous discussion dealt with the i.i.d. case. In simulation this corresponds to the static Monte Carlo situation.
  - Application examples: evaluation of multidimensional sums (integrals), for example blocking probabilities in loss systems
- Dynamic simulation
  - For some simple stochastic problems one can derive so called asymptotically optimal IS distributions (simple example: the M/M/1 queue)
  - More complex systems can not be analyzed and hence adaptive schemes have been proposed