Exam Final

IFT6561

Gabriel C-Parent

DIRO
UdeM

Part 1

a

Experimental Results

Beta vector with CV
(0.00628, 0.000347, 0.00138, 0.00501, 0.00458, 0.00271, 0.00276, 0.00703)
Average without CV : 0.11448
Average with CV : 0.11291
Variance without CV : 0.025091
Variance with CV : 0.010342
Variance reduction factor : 2.4261
Confidence interval student 95 without CV : 0.11448 ± 0.0031046
Confidence interval student 95 with CV : 0.11291 ± 0.0019932

As shown in the data, the confidence interval with control variates (CV) is about 1.5 times tighter than without CV. The average of the estimator without CV is also contained within the 95% student confidence interval of the estimator with CV. Overall, the strategy using CV is pretty good, decreasing the variance by a factor of 2.4. This was expected, as stated in example 6.26.

b

The whole concept of stochastic derivatives is explained at p.101 of the book, I'm not rewriting it here, we'll just state the core of the matter.
The first necessary observation is that both θ_2 and θ_4 are parameters for the mean of exponential distributions. Using the same proof as given in example 1.49 of the book, we can explain why our estimator is unbiased.
Suppose we want to estimate $\frac{\delta \mathbb{E}[T]}{\delta \theta_i}$, we can write T as a function of θ_i.
Say f is the function. $f'(\theta_i, U) = V_i'(\theta_i)$ if the link i is in the longest path and 0 otherwise. Since, in this case, both θ_i are the means of exponential random variables, we have $Y_i = Y_i(\theta_i) = -\theta_i ln(1 - U_i)$ and its derivative is $-ln(1 - U_i)$. Using the corollary 6.6 to prove that we can interchange the expectation and derivative operators to obtain that our estimator of $f'(\theta_i)$ is unbiased.

Basically, this is the rest of the example 1.49.

c

As stated in example 1.50 of the book, if the function $f_j(\theta_j, U) = \mathbb{1}[T > x]$, the estimator can only take two values, 0 and 1. Its derivative is either undefined, when the function jumps exactly at θ_j or is 0. It cannot be an unbiased estimator of the derivative because the original function is discontinuous at some point as a function of θ_j.

d

First thing, we need to condition on all activities except the ones depending on θ_2 and θ_4.
The value of the equation is $\mathbb{P}[T > x] = 1 - \prod l \in LF_l[x - \alpha_l - \beta_l]$

e

f

g

2

a

The point sets (PS) are in corresponding order: Sob + S, Sob + LMS + S, Kor + S, Kor + S + B
PS1 BGSS : 25
PS1 BGBS : 57
PS1 DGBS : 297
PS2 BGSS : 12
PS2 BGBS : 77
PS2 DGBS : 561
PS3 BGSS : 10
PS3 BGBS : 18
PS3 DGBS : 116
PS4 BGSS : 12

PS4 BGBS : 127

PS4 DGBS : 118

The results for Sob+S are lower than expected, especially for the bridge sampling, but the results for Sob + LMS + S match quite well with table 6.12.

For the Korobov lattice, Kor+S matches very well the expected results while Kor+S+B yields much greater variance reduction factor.

I think the cause of this is perhaps an error in the calculation of the variance reduction factor, or otherwise the variation is too important.

The important point is that we reach the same observation that DGBS usually yields superior reduction of variance.

b

$$\pi(x) = \frac{\lambda^\alpha x^{\alpha-1} \exp(-\lambda x)}{\Gamma(\alpha)}$$

$$\pi_\theta(x) = \frac{\exp(\theta x)\pi(x)}{\mathbb{M}(\theta)}$$

where M is the mgf of the Gamma distribution. The exponentially twisted pdf is then

$$\pi_\theta(x) = \frac{\exp(\theta x)}{(1 - \frac{\theta}{\lambda})^{-\alpha}} \frac{\lambda^\alpha x^{\alpha-1} \exp(-\lambda x)}{\Gamma(\alpha)}$$

this yields the following new densities

$$\pi_\theta(x) = \frac{\exp(-(\lambda - \theta))x^{\alpha-1}}{\Gamma(\alpha)(\lambda - \theta)^{-\alpha}}$$

We can generate pseudo-random variable by sampling from the new distributions $Gamma(\alpha, \lambda - \theta)$, since only the scale parameter was changed.

The likelihood ratio to multiply by will be the product of the individual likelihoods:

$$L(x) = \exp(-\theta x + \alpha(\ln(\frac{\lambda}{\lambda - \theta})))$$

$$L(\omega) = L(x_1) * L(x_2)$$

We must simply apply these to the Gamma processes (substituting for the real parameters).

c

d

e