3 Some Applications of the Delta Method

We consider some further aspects and extensions of the MC method in the notation of II.1.

3a Estimating a Function

In some cases, one is interested in estimating not $z = \mathbb{E}Z$, but some function $f(z)$.

Ex31.3a **Example 3.1** Let $Z_1, Z_2, \ldots$ be i.i.d. failure times of some system. Then $f(z) = 1/z = 1/\mathbb{E}Z$ is the long–run failure intensity. □

Consider thus an estimator $\widehat{z} = \widehat{z}_R$ obeying a CLT of the form (2.2) (not necessarily based upon the CMC method and with a possible non–zero bias δ_R). We can then write

$$f(\widehat{z}_R) - f(z) \;=\; f'(z)(\widehat{z}_R - z) + \frac{f''(z)}{2}(\widehat{z}_R - z)^2 + \mathrm{o}\big((\widehat{z}_R - z)^2\big). \qquad (3.1)$$ 12.2g

Multiplying by $\sqrt{R}$, the two last terms on the r.h.s. are negliglible from the point of view of distributions, and so we conclude that

$$\sqrt{R}(f(\widehat{z}_R) - f(z)) \;\xrightarrow{\mathscr{D}}\; \mathrm{N}\big(0, \omega^2\big)$$

where $\omega^2 = f'(z)^2\sigma^2$; this is the essence of the Delta method in its simplest form. We have the obvious estimator $\widehat{\omega}^2 = f'(\widehat{z})^2\widehat{\sigma}^2$ for ω^2, and hence $\widehat{z}_R \pm 1.96\,\hat{\omega}/\sqrt{R}$ is an asymptotic 95% confidence interval.

If the bias δ_R is of order δ/R, taking expectations in (3.1) yields the following asymptotics for the bias of $\widehat{z}_R$[3]

$$\mathbb{E}\big[f(\widehat{z}_R) - f(z)\big] \;\sim\; \frac{1}{R}\Big(\delta f'(z) + \frac{\sigma^2 f''(z)}{2}\Big). \qquad (3.2)$$ 12.2h

One implication is that even if $\widehat{z}_n$ is unbiased, $f(\widehat{z}_R)$ is not: taking a non–linear function typically introduces a bias of order $1/R$.

3b Multivariate Output

Let next $\boldsymbol{Z} = \big(Z^{(1)}, \ldots, Z^{(k)}\big)$ be a random vector with possibly dependent component so that $\sigma_{ij} = \mathbb{C}ov\big(Z^{(i)}, Z^{(j)}\big)$ may be non-zero for $i \neq j$, and assume that we want to estimate $f = f\big(z^{(1)}, \ldots, z^{(k)}\big)$ where $z^{(i)} = \mathbb{E}Z^{(i)}$ (we will meet the example $\mathbb{E}Z^{(2)}/\mathbb{E}Z^{(1)}$ in connection with regenerative

[3] Whereas the derivation of the CLT is rigorous, this argument is heuristic; one needs various types of uniform integrability assumptions. The point is just to give an idea of what type of result can typically be expected.

simulation). We then simulate R i.i.d. replications

$$\boldsymbol{Z}_1 = \bigl(Z_1^{(1)},\dots,Z_1^{(k)}\bigr),\ \dots,\ \boldsymbol{Z}_R = \bigl(Z_n^{(1)},\dots,Z_R^{(k)}\bigr)$$

of $\boldsymbol{Z}$ and let $\hat{f} = f\bigl(\hat{z}^{(1)},\dots,\hat{z}^{(k)}\bigr)$ where

$$\widehat{z}^{(i)} \;=\; \frac{Z_1^{(i)}+\cdots+Z_R^{(i)}}{R}, \quad \widehat{\sigma}_{ij} \;=\; \frac{1}{R-1}\sum_{r=1}^{R}\bigl(Z_r^{(i)}-\widehat{z}^{(i)}\bigr)\bigl(Z_r^{(j)}-\hat{z}^{(j)}\bigr).$$

Then up to $\mathrm{o}\bigl(R^{-1/2}\bigr)$ terms we have

$$\widehat{f}-f \;=\; \sum_{i=1}^{k}\frac{\partial f}{\partial z^{(i)}}\bigl(z^{(1)},\dots,z^{(k)}\bigr)\bigl(\widehat{z}^{(i)}-z^{(i)}\bigr)$$

from which it follows that $\sqrt{R}\bigl(\widehat{f}-f\bigr) \xrightarrow{\mathscr{D}} \mathrm{N}\bigl(0,\omega^2\bigr)$ where

$$\omega^2 \;=\; \sum_{i,j=1}^{k}\frac{\partial f}{\partial z^{(i)}}\bigl(z^{(1)},\dots,z^{(k)}\bigr)\frac{\partial f}{\partial z^{(j)}}\bigl(z^{(1)},\dots,z^{(k)}\bigr)\sigma_{ij}$$

and our confidence interval is $\hat{f}\pm 1.96\,\hat{\omega}/\sqrt{R}$ where

$$\widehat{\omega}^2 \;=\; \sum_{i,j=1}^{k}\frac{\partial f}{\partial z^{(i)}}\bigl(\hat{z}^{(1)},\dots,\hat{z}^{(k)}\bigr)\frac{\partial f}{\partial z^{(j)}}\bigl(\hat{z}^{(1)},\dots,\hat{z}^{(k)}\bigr)\hat{\sigma}_{ij}\,.$$

3c The Variance of the Variance Estimator $\hat{\sigma}^2$

Given the choice between two CMC schemes based upon (one-dimensional) r.v.'s Z, Z' with variances $\sigma_Z^2, \sigma_{Z'}^2$, the first concern would be to obtain minimal variance for a given CPU time (simulation budget) t as discussed in Section 1.

If the variances $\sigma_Z^2 T$, $\sigma_{Z'}^2 T'$ per unit CPU time are roughly the same (in particular, if σ_Z^2 and $\sigma_{Z'}^2$ are roughly the same and T, T' are roughly the same), the next concern might be to choose the method with the most reliable variance estimate. We go next through the computations relevant for this comparison. Write $m_k = \mathbb{E}Z^k$ (then $z = m_1$, $\sigma^2 = m_2 - m_1^2$).

`Prop17.1a` **Proposition 3.2** $\sqrt{R}\begin{pmatrix}\hat{z}-z\\ \hat{\sigma}^2-\sigma^2\end{pmatrix} \xrightarrow{\mathscr{D}}$

$$N\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}\sigma^2 = m_2-m_1^2 & 2m_1^3+m_3-3m_1m_2\\ 2m_1^3+m_3-3m_1m_2 & -4m_1^4+8m_1^2m_2+m_4-m_2^2-4m_1m_3\end{pmatrix}\right)$$

Proof. Let $\overline{Z} = \sum_1^R Z_r/R = \hat{z}$, $\overline{Z^2} = \sum_1^R Z_r^2/R$. Obviously,

$$\sqrt{R}\begin{pmatrix}\overline{Z}-m_1\\ \overline{Z^2}-m_2\end{pmatrix} \xrightarrow{\mathscr{D}} N\left(\begin{pmatrix}0\\0\end{pmatrix},\boldsymbol{\Sigma}\right)$$

Gompertz-Makeham distribution
insurance
life insurance

where

$$\boldsymbol{\Sigma} = \begin{pmatrix} \mathbb{V}ar(Z) & \mathbb{C}ov(Z,Z^2) \\ \mathbb{C}ov(Z,Z^2) & \mathbb{V}ar(Z^2) \end{pmatrix} = \begin{pmatrix} \sigma^2 = m_2 - m_1^2 & m_3 - m_1 m_2 \\ m_3 - m_1 m_2 & m_4 - m_2^2 \end{pmatrix}$$

Letting $\boldsymbol{f}(x,y) = \big(x, y - x^2\big)$, $\big(\widehat{z}, \widehat{\sigma}^2\big)$ has the same asymptotics as $\boldsymbol{f}\big(\overline{Z}, \overline{Z^2}\big)$ so the result follows by the Delta method, with the asymptotic covariance matrix

$$\boldsymbol{Df}(m_1,m_2)\,\boldsymbol{\Sigma}\,\boldsymbol{Df}(m_1,m_2)' = \begin{pmatrix} 1 & 0 \\ -2m_1 & 1 \end{pmatrix} \boldsymbol{\Sigma} \begin{pmatrix} 1 & -2m_1 \\ 0 & 1 \end{pmatrix}$$

which is the same as asserted. □

Exercises

oAss2

3.1 (TP) A system develops in i.i.d. cycles. According to whether a certain catastrophic event occurs or not within a cycle, the cycle is classified as failed or non–failed. Denote by p the probability that a cycle is failed, by ℓ_1 the expected length of a cycle given it does not fail, and by ℓ_2 the expected time until failure in a cycle given it is failed. We are interested in ℓ, the expected time until a failure occurs.

1. Express ℓ in terms of p, ℓ_1, ℓ_2.
2. You are presented with statistics of 1000 cycles, of which 87 failed. The non–failed cycles had an empirical mean of 20.2 and an empirical variance of 18.6, and the average time until failure in the failed cycles was 5.4 with an empirical variance of 3.1. Give a confidence interval for ℓ.

OutPr8.9a

3.2 (A) A man of age 45 signs a life annuity contract, according to which he has to pay premium at (continuous) rate π until age 67 or the time T of death, and if still alive at age 67, he will receive a payment at rate 1 until death. Assuming a constant interest rate r, the equation determining π is therefore

$$\mathbb{E}\int_{55}^{67\wedge T} \pi \mathrm{e}^{-rt}\mathrm{d}t = \mathbb{E}\int_{T\vee 67}^{T} \mathrm{e}^{-rt}\mathrm{d}t$$

(the two sides are the present values of payments, resp. benefits). Your assignment is to give a simulation estimate of π and an associated confidence interval, assuming $r = 4\%$ and that T follows the *Gompertz-Makeham distribution* (conditioned on $T > 55$). This is given by its hazard rate $\beta(t) = a + b\mathrm{e}^{ct}$ so that the the tail and the density are

$$\overline{F}(t) = \exp\{-at - \mathrm{e}^{ct}b/c\}, \quad f(t) = \big(a + b\mathrm{e}^{ct}\big)\overline{F}(t)\,,$$

and is widely used as lifetime distribution in life insurance. Use the parameters $a = 5{\cdot}10^{-4}$, $b = 7.5858{\cdot}10^{-5}$, $c = \log(1.09144)$ from the so-called G82 base widely employed by Danish insurance companies (the unit is years). The r.v. T may be generated by acceptance-rejection with a uniform(55,110) proposal (neglecting the possibility $T > 110$), using the fact that $f(t) \le 0.035$ for all t.