Rare event estimation: Quiz 2

Email answers & code to Patrick Laub before class on March 22

1. Imagine performing the cross-entropy method where the family of distributions to search inside are the $\mathsf{Pareto}(v)$ distributions, with p.d.f. $f(x; v) = vx^{-(v+1)}$ for $x \geq 1$. Solve algebraically

$$\sum_{r=1}^{R} 1\{X_r > \gamma_i\} \frac{f(X_r)}{f(X_r; v_{i-1})} \frac{\mathrm{d}}{\mathrm{d}v} \Big\{ \log\big[f(X_r; v) \big] \Big\} = 0.$$

2. Estimate $\ell = \mathbb{P}(S > 10^5)$ where $S = X_1 + X_2 + X_3$ where $X_i \sim \mathsf{Pareto}(i + 10)$ and the X_i's are independent of each other. Use the cross-entropy method to find the best IS proposal distribution. Look in the family of distributions

$$f(\boldsymbol{x}; \boldsymbol{v}) = f_{\mathsf{Pareto}(v_1)}(x_1) f_{\mathsf{Pareto}(v_2)}(x_2) f_{\mathsf{Pareto}(v_3)}(x_3)$$

where $\boldsymbol{x}$ denotes (x_1, x_2, x_3) and $f_{\mathsf{Pareto}(v)}(x) = vx^{-(v+1)}$ is the p.d.f. of a $\mathsf{Pareto}(v)$ variable.

Hint: In R, the `actuar` package has some simple methods for Pareto variables (and more complex ones for generalised Pareto). However, note that their definition of Pareto variables is shifted so that they have support $[0, \infty)$ whereas the ones above are defined on $[1, \infty)$.

3. Use cross-entropy to find the minimum of Rastrigin's function

$$S(\boldsymbol{x}) = 20 + \big(x_1^2 - 10\cos(2\pi x_1)\big) + \big(x_2^2 - 10\cos(2\pi x_2)\big).$$

Use the family of bivariate normal distributions, i.e.

$$f(x; v_i) = f(x; (\mu_i, \Sigma_i)) = (2\pi)^{-1} |\Sigma_i|^{-1/2} \exp\left\{ -\frac{1}{2}(x - \mu_i)^\top \Sigma_i^{-1} (x - \mu_i) \right\}.$$

Start with $\mu_0 = (150, 150)$, and choose your own Σ_0 starting matrix. Recreate the some of the plots from the lecture slides:

- scatterplots of the samples for a few different iterations, marking the elite samples differently to the others,
- a plot of the best objective value found across each of the iterations (on log scale for y axis).