DATA 1010
In-class exercises
Samuel S. Watson
26 October 2018

Problem 1

The central limit theorem says that if S_n is a sum of i.i.d. finite-variance random variables is approximately normally distributed with mean $\mathbb{E}[S_n]$ and variance $\operatorname{Var}(S_n)$. Also, about 95% of the probability mass of a normal distribution is within two standard deviations of the mean.

If a million independent $\operatorname{Unif}([a,b])$'s are added, what is the shortest interval containing 95% of the probability mass of the distribution of the resulting sum?

Problem 2

The multivariate central limit theorem says that if $\mathbf{X}_1, \mathbf{X}_2, \ldots$ is an independent sequence of random vectors with a common distribution on $\mathbb{R}^n$, then the standardized mean

$$\mathbf{S}_n^* = \frac{X_n - n\mu}{\sqrt{n}}$$

converges in distribution to $\mathcal{N}(\mathbf{0}, \Sigma)$, where Σ is the covariance matrix of $\mathbf{X}_1$.

Investigate the multivariate central limit theorem using by making 2D histograms for i.i.d. sums of (i) uniform samples from the square, and (ii) samples from $((U+V)/2, V)$, where (U, V) is uniformly sampled from the square.

Problem 3

Find the mean and covariance of the random vector $[X, Y]$ defined by $X = \frac{1}{2}(U+V), Y = V$, where U and V are independent uniform random variables on $[0,1]$.

Use the result to find the density of the limiting distribution you plotted in the previous problem.

Problem 4

Find the conditional expectation of Y given X if the joint distribution has density

$$f(x,y) = \frac{3}{4000(3/2)\sqrt{2\pi}} x(20-x) e^{-\frac{1}{2(3/2)^2}\left(y-2-\frac{1}{50}x(30-x)\right)^2}.$$

on the strip $[0,20] \times \mathbb{R}$.