Logistic regression: log excess loss with GD, SGD, scikit
n=10000, d=80
$\gamma_{SGD0} = 0.1$, schedule: decreasing 2
$\gamma_{GD} = 0.5$
$\mu = 1$

- GD empirical risk (5.716s)
- SGD empirical risk (0.569s)
- SGA empirical risk (0.606s)

x-axis: log10 number of sample queries (computation time)
y-axis: log10 excess empirical risk