Normalized cumulated regrets for different bandit algorithms, averaged 100 times
3 arms: Non-Stationary MAB, Bernoulli with $\Upsilon = 4$ break-points

Normalized non-stationary regret
$\frac{R_t}{\log(t)} = \frac{1}{\log(t)} \sum_{s=1}^{t} \max_k \mu_k(t) - \frac{1}{\log(t)} \sum_{s=1}^{t} \sum_{k=1}^{3} \mu_k(t) E_{100}[\mathbb{1}(I(t) = k)]$

Time steps $t = 1 \ldots T$, horizon $T = 10000$

UCB