The Power of Two-Choices in reducing discrepancy

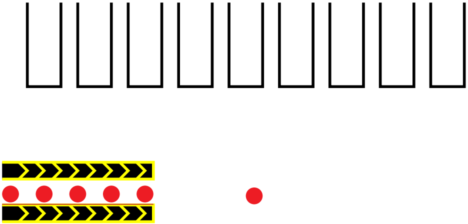
Ohad N. Feldheim

Based on joint works with:
Ori Gurel-Gurevich (HUJI)
Raaz Dwivdi & Aaditya Ramdas (Berkeley)

June 2017

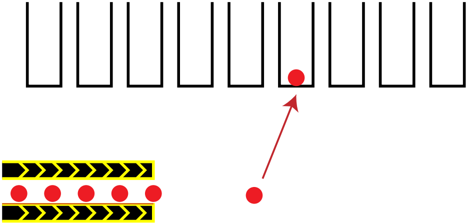
Balls and bins model

Consider an online process in which N balls are randomly assigned, one by one to M bins.



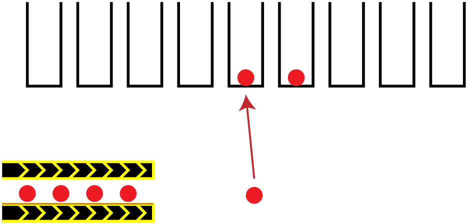
Balls and bins model

Consider an online process in which N balls are randomly assigned, one by one to M bins.



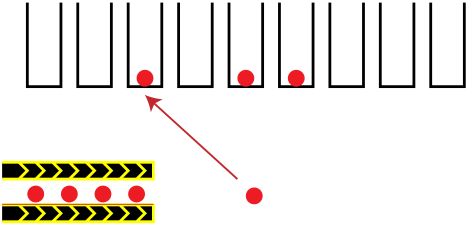
Balls and bins model

Consider an online process in which N balls are randomly assigned, one by one to M bins.



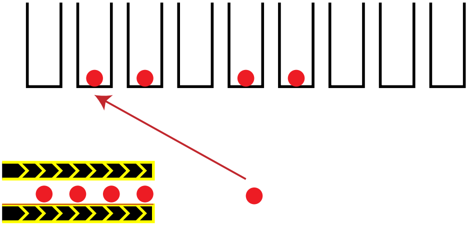
Balls and bins model

Consider an online process in which N balls are randomly assigned, one by one to M bins.



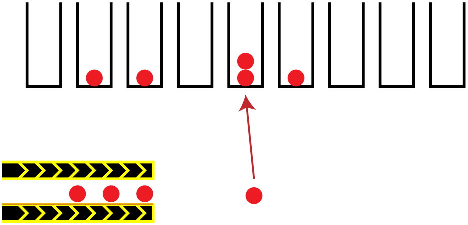
Balls and bins model

Consider an online process in which N balls are randomly assigned, one by one to M bins.



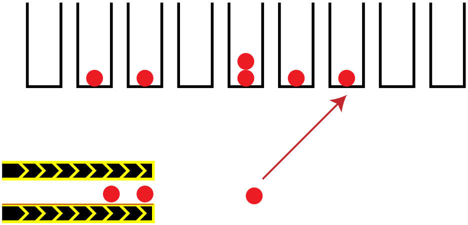
Balls and bins model

Consider an online process in which N balls are randomly assigned, one by one to M bins.



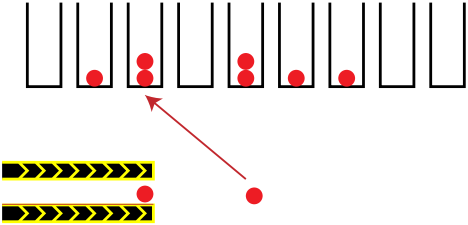
Balls and bins model

Consider an online process in which N balls are randomly assigned, one by one to M bins.



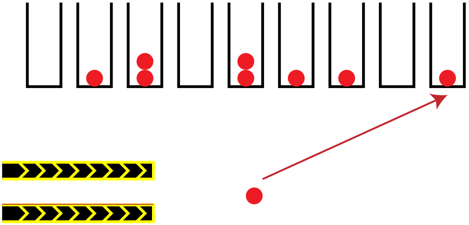
Balls and bins model

Consider an online process in which N balls are randomly assigned, one by one to M bins.



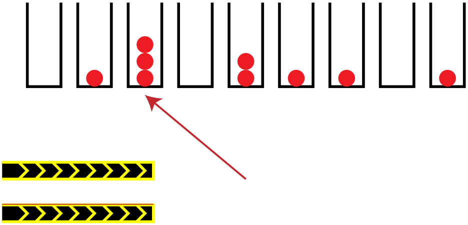
Balls and bins model

Consider an online process in which N balls are randomly assigned, one by one to M bins.



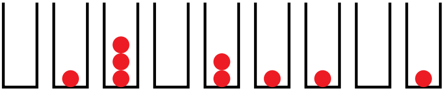
Balls and bins model

Consider an online process in which N balls are randomly assigned, one by one to M bins.



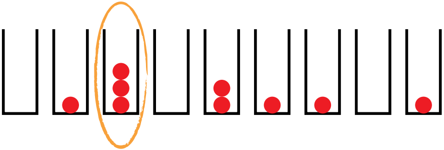
Balls and bins model

Consider an online process in which N balls are randomly assigned, one by one to M bins.



Balls and bins model

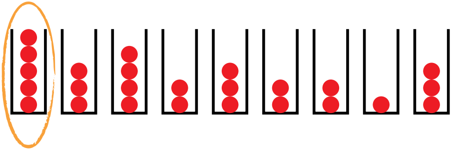
Consider an online process in which N balls are randomly assigned, one by one to M bins.



- After M balls, highest occupancy is a.a.s. $(1+o(1))\frac{\log M}{\log\log M}$

Balls and bins model

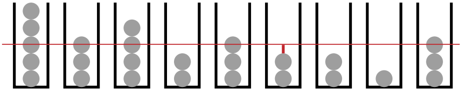
Consider an online process in which N balls are randomly assigned, one by one to M bins.



- After M balls, highest occupancy is a.a.s. $(1 + o(1))\frac{\log M}{\log \log M}$
- After $N >> M$ balls, highest occupancy is a.a.s. $\Theta\left(\sqrt{\frac{N \log M}{M}}\right)$

Balls and bins model

Consider an online process in which N balls are randomly assigned, one by one to M bins.

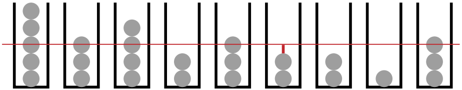


- After M balls, highest occupancy is a.a.s. $(1+o(1))\frac{\log M}{\log\log M}$
- After $N >> M$ balls, highest occupancy is a.a.s. $\Theta\left(\sqrt{\frac{N\log M}{M}}\right)$

  typical deviation from expectation is $\Theta\left(\sqrt{\frac{N}{M}}\right)$

Balls and bins model

Consider an online process in which N balls are randomly assigned, one by one to M bins.

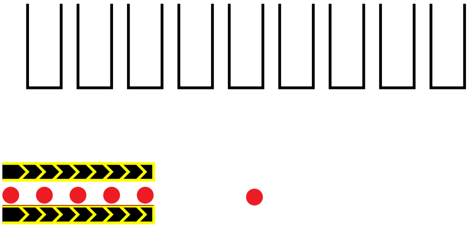


- After M balls, highest occupancy is a.a.s. $(1+o(1))\frac{\log M}{\log\log M}$
- After $N >> M$ balls, highest occupancy is a.a.s. $\Theta\left(\sqrt{\frac{N\log M}{M}}\right)$

  typical deviation from expectation is $\Theta\left(\sqrt{\frac{N}{M}}\right)$
- *Load balancing* is an effort to reduce these quantities. (possible with control over the distribution of the balls.)

Power of two choices

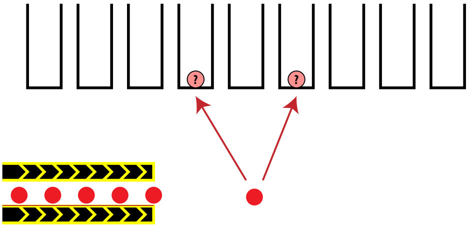
Azar, Broder, Karlin, and Upfal (1994): Very little choice is sufficient for rather balanced allocation - choice between two random cells per ball.



Greedy strategy: choose the least occupied cell.

Power of two choices

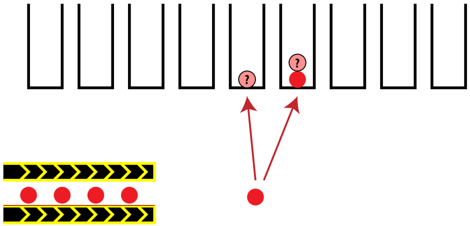
Azar, Broder, Karlin, and Upfal (1994): Very little choice is sufficient for rather balanced allocation - choice between two random cells per ball.



Greedy strategy: choose the least occupied cell.

Power of two choices

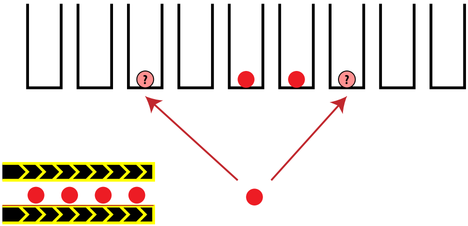
Azar, Broder, Karlin, and Upfal (1994): Very little choice is sufficient for rather balanced allocation - choice between two random cells per ball.



Greedy strategy: choose the least occupied cell.

Power of two choices

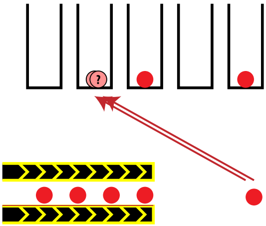
Azar, Broder, Karlin, and Upfal (1994): Very little choice is sufficient for rather balanced allocation - choice between two random cells per ball.



Greedy strategy: choose the least occupied cell.

Power of two choices

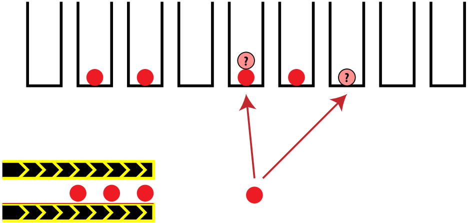
Azar, Broder, Karlin, and Upfal (1994): Very little choice is sufficient for rather balanced allocation - choice between two random cells per ball.



Greedy strategy: choose the least occupied cell.

Power of two choices

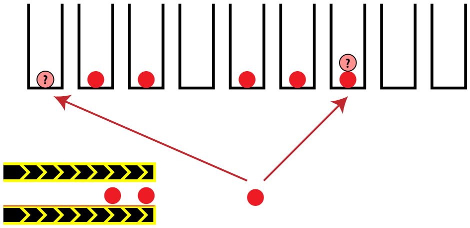
Azar, Broder, Karlin, and Upfal (1994): Very little choice is sufficient for rather balanced allocation - choice between two random cells per ball.



Greedy strategy: choose the least occupied cell.

Power of two choices

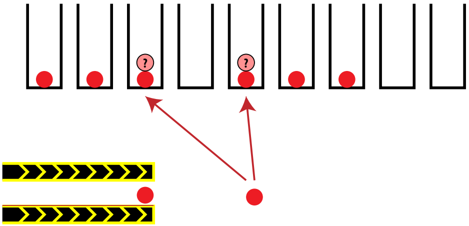
Azar, Broder, Karlin, and Upfal (1994): Very little choice is sufficient for rather balanced allocation - choice between two random cells per ball.



Greedy strategy: choose the least occupied cell.

Power of two choices

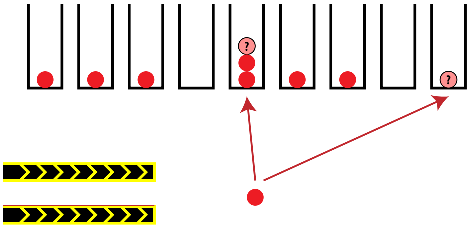
Azar, Broder, Karlin, and Upfal (1994): Very little choice is sufficient for rather balanced allocation - choice between two random cells per ball.



Greedy strategy: choose the least occupied cell.

Power of two choices

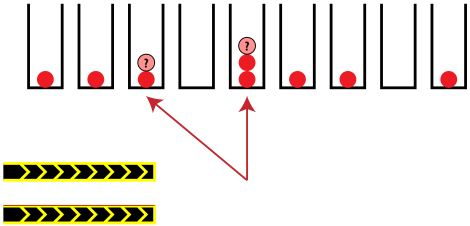
Azar, Broder, Karlin, and Upfal (1994): Very little choice is sufficient for rather balanced allocation - choice between two random cells per ball.



Greedy strategy: choose the least occupied cell.

Power of two choices

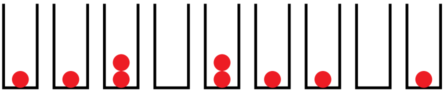
Azar, Broder, Karlin, and Upfal (1994): Very little choice is sufficient for rather balanced allocation - choice between two random cells per ball.



Greedy strategy: choose the least occupied cell.

Power of two choices

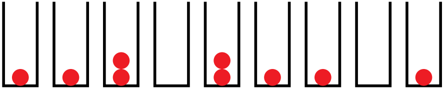
Azar, Broder, Karlin, and Upfal (1994): Very little choice is sufficient for rather balanced allocation - choice between two random cells per ball.



Greedy strategy: choose the least occupied cell.

Power of two choices

Azar, Broder, Karlin, and Upfal (1994): Very little choice is sufficient for rather balanced allocation - choice between two random cells per ball.



Greedy strategy: choose the least occupied cell.

Balls	no-choice max	2-choices max	no-choice typ. dev.	2-choices typ. dev.
M	$\frac{\log M}{\log\log M}$	$\frac{\log\log M}{2}$	O(1)	O(1)
$N >> M$	$\Theta\left(\sqrt{\frac{N \log M}{M}}\right)$	$\log M$	$\Theta\left(\sqrt{\frac{N}{M}}\right)$	O(1)

Power of two choices - remarks

This observation had many applications

- Server load-balancing
- Distributed shared memory
- Efficient on-line hashing
- Low-congestion circuit routing

Power of two choices - remarks

This observation had many applications

- Server load-balancing
- Distributed shared memory
- Efficient on-line hashing
- Low-congestion circuit routing

It is interesting to note that

Power of two choices - remarks

This observation had many applications

- Server load-balancing
- Distributed shared memory
- Efficient on-line hashing
- Low-congestion circuit routing

It is interesting to note that

- More choice does not yield significant improvement.

Power of two choices - remarks

This observation had many applications

- Server load-balancing
- Distributed shared memory
- Efficient on-line hashing
- Low-congestion circuit routing

It is interesting to note that

- More choice does not yield significant improvement.
- If balls keep appearing and dying at rate 1 the phenomenon persists (Luczak & McDiarmid '05)

Power of two choices - remarks

This observation had many applications

- Server load-balancing
- Distributed shared memory
- Efficient on-line hashing
- Low-congestion circuit routing

It is interesting to note that

- More choice does not yield significant improvement.
- If balls keep appearing and dying at rate 1 the phenomenon persists (Luczak & McDiarmid '05)
- However if one can't keep track of the number of balls per bin (due to having $M^{1-\epsilon}$ bits of memory), then no asymptotic improvement over no-choice is possible (Alon, Gurel-Gurevich, Lubetzky '09)

Interval partition

Consider an online process in which an interval is partitioned into smaller and smaller subintervals by selecting new partition points uniformly at random.

Interval partition

Consider an online process in which an interval is partitioned into smaller and smaller subintervals by selecting new partition points uniformly at random.

Interval partition

Consider an online process in which an interval is partitioned into smaller and smaller subintervals by selecting new partition points uniformly at random.

Interval partition

Consider an online process in which an interval is partitioned into smaller and smaller subintervals by selecting new partition points uniformly at random.

Interval partition

Consider an online process in which an interval is partitioned into smaller and smaller subintervals by selecting new partition points uniformly at random.

Interval partition

Consider an online process in which an interval is partitioned into smaller and smaller subintervals by selecting new partition points uniformly at random.

Interval partition

Consider an online process in which an interval is partitioned into smaller and smaller subintervals by selecting new partition points uniformly at random.

Interval partition

Consider an online process in which an interval is partitioned into smaller and smaller subintervals by selecting new partition points uniformly at random.

We view the points at time n as each having $1/n$ mass, call this μ^n

Interval partition

Consider an online process in which an interval is partitioned into smaller and smaller subintervals by selecting new partition points uniformly at random.

We view the points at time n as each having $1/n$ mass, call this μ^n

Convergence to uniform measure:

$\lim_{n\to\infty} \mu^n \stackrel{\mathrm{T}.V.}{=} \mathcal{U}[0,1]$

Interval partition

Consider an online process in which an interval is partitioned into smaller and smaller subintervals by selecting new partition points uniformly at random.

We view the points at time n as each having $1/n$ mass, call this μ^n

Convergence to uniform measure:

$\lim_{n\to\infty} \mu^n \stackrel{T.V.}{=} \mathcal{U}[0,1]$

Three natural ways to measure discrepancy/rate of convergence:

Geometric:

Interval partition

Consider an online process in which an interval is partitioned into smaller and smaller subintervals by selecting new partition points uniformly at random.

We view the points at time n as each having $1/n$ mass, call this μ^n

Convergence to uniform measure:

$\lim_{n\to\infty} \mu^n \overset{\mathsf{T}.V.}{=} \mathcal{U}[0,1]$

Three natural ways to measure discrepancy/rate of convergence:

Geometric:

- Interval variation - $\max_{a,b} |\mu^n((a,b)) - \mu((a,b))|$

Interval partition

Consider an online process in which an interval is partitioned into smaller and smaller subintervals by selecting new partition points uniformly at random.

We view the points at time n as each having $1/n$ mass, call this μ^n

Convergence to uniform measure:

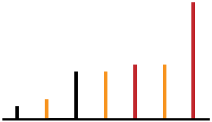
$\lim_{n\to\infty} \mu^n \overset{T.V.}{=} \mathcal{U}[0,1]$

Three natural ways to measure discrepancy/rate of convergence:
Geometric:

- Interval variation - $\max_{a,b} |\mu^n((a,b)) - \mu((a,b))|$
- Largest/smallest interval - $\max_{a,b} |(\mu^n((a,b)) = 0) - 1/n|$

Interval partition

Consider an online process in which an interval is partitioned into smaller and smaller subintervals by selecting new partition points uniformly at random.



We view the points at time n as each having $1/n$ mass, call this μ^n

Convergence to uniform measure:

$\lim_{n\to\infty} \mu^n \overset{\text{T.}V.}{=} \mathcal{U}[0,1]$

Three natural ways to measure discrepancy/rate of convergence:
Geometric:

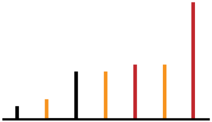
- Interval variation - $\max_{a,b} |\mu^n((a,b)) - \mu((a,b))|$
- Largest/smallest interval - $\max_{a,b} |(\mu^n((a,b)) = 0) - 1/n|$

Non-geometric:

- Empirical normalized interval distribution.

Interval partition

Consider an online process in which an interval is partitioned into smaller and smaller subintervals by selecting new partition points uniformly at random.



We view the points at time n as each having $1/n$ mass, call this μ^n

Convergence to uniform measure:

$\lim_{n\to\infty} \mu^n \overset{\mathrm{T.V.}}{=} \mathcal{U}[0,1]$

Three natural ways to measure discrepancy/rate of convergence:

Geometric:

- Interval variation - $\max_{a,b} |\mu^n((a,b)) - \mu((a,b))|$ -
- Largest/smallest interval - $\max_{a,b} |(\mu^n((a,b)) = 0) - 1/n|$

Non-geometric:

- Empirical normalized interval distribution.

Two-choices and interval partitions

Benjamini: Can two choices regulate interval partitions?

Two-choices and interval partitions

Benjamini: Can two choices regulate interval partitions?

In particular what if we partition the largest interval?

What if we partition the point furthest from neighbour?

Two-choices and interval partitions

Benjamini: Can two choices regulate interval partitions?

In particular what if we partition the largest interval?

What if we partition the point furthest from neighbour?

Two-choices and interval partitions

Benjamini: Can two choices regulate interval partitions?

In particular what if we partition the largest interval?

What if we partition the point furthest from neighbour?

Two-choices and interval partitions

Benjamini: Can two choices regulate interval partitions?

In particular what if we partition the largest interval?

What if we partition the point furthest from neighbour?

Cf. *Kakutani process* - uniform partition of the largest interval.

Two-choices and interval partitions

Benjamini: Can two choices regulate interval partitions?

In particular what if we partition the largest interval?

What if we partition the point furthest from neighbour?

Cf. *Kakutani process* - uniform partition of the largest interval.

Two-choices and interval partitions

Benjamini: Can two choices regulate interval partitions?

In particular what if we partition the largest interval?

What if we partition the point furthest from neighbour?

Cf. *Kakutani process* - uniform partition of the largest interval.

Two-choices and interval partitions

Benjamini: Can two choices regulate interval partitions?

In particular what if we partition the largest interval?

What if we partition the point furthest from neighbour?

Cf. *Kakutani process* - uniform partition of the largest interval.

Two-choices and interval partitions

Benjamini: Can two choices regulate interval partitions?

In particular what if we partition the largest interval?

What if we partition the point furthest from neighbour?

Cf. *Kakutani process* - uniform partition of the largest interval.

The empirical interval distribution:

Uniform interval partition $\rightarrow Exp(1)$
Kakutani interval partition $\rightarrow \mathcal{U}(0,2)$
Max-2 interval partition $\rightarrow ???$

Two-choices and interval partitions

Benjamini: Can two choices regulate interval partitions?

In particular what if we partition the largest interval?

What if we partition the point furthest from neighbour?

Cf. *Kakutani process* - uniform partition of the largest interval.

The empirical interval distribution:

Uniform interval partition $\to Exp(1)$
Kakutani interval partition $\to \mathcal{U}(0,2)$
Max-2 interval partition $\to ???$

However, even Kakutani offers no significant improvement in terms of interval variation.

Convergence of 2-Max interval partition process

Studying 2-Max is a rather difficult task:

Maillard & Paquette '14: 2-Max converges to some limit distribution.

Convergence of 2-Max interval partition process

Studying 2-Max is a rather difficult task:

Maillard & Paquette '14: 2-Max converges to some limit distribution.
Junge '15: 2-Max converges to $U[0,1]$.

Convergence of 2-Max interval partition process

Studying 2-Max is a rather difficult task:

Maillard & Paquette '14: 2-Max converges to some limit distribution.
Junge '15: 2-Max converges to $U[0, 1]$.

Both experimental and heuristic arguments suggest that 2-Max offers no improvement in interval variation when compared with uniform.

Convergence of 2-Max interval partition process

Studying 2-Max is a rather difficult task:

Maillard & Paquette '14: 2-Max converges to some limit distribution.
Junge '15: 2-Max converges to $U[0,1]$.

Both experimental and heuristic arguments suggest that 2-Max offers no improvement in interval variation when compared with uniform.

In a sense, corresponds to Alon, Gurel-Gurevich, Lubetzky.

Global strategy that regulates interval variation

We show that a global strategy can reduce discrepancy significantly.

Global strategy that regulates interval variation

We show that a global strategy can reduce discrepancy significantly.

Power of two-choices in regulating interval discrepancy (F. & Gurel-Gurevich 2016+)

In a power of two-choices process on $\mathcal{U}[0,1]$, the chooser can obtain

$$\lim_{n\to\infty} \Pr\left(\max_{a,b} |\mu^n((a,b)) - \mu((a,b))| < \frac{C \log^3 N}{N} \right) = 1.$$

Global strategy that regulates interval variation

We show that a global strategy can reduce discrepancy significantly.

Power of two-choices in regulating interval discrepancy (F. & Gurel-Gurevich 2016+)

In a power of two-choices process on $\mathcal{U}[0,1]$, the chooser can obtain

$$\lim_{n\to\infty} \Pr\left(\max_{a,b} |\mu^n((a,b)) - \mu((a,b))| < \frac{C\log^3 N}{N}\right) = 1.$$

- In fact power of one-retry is enough (to be explained shortly).

Global strategy that regulates interval variation

We show that a global strategy can reduce discrepancy significantly.

Power of two-choices in regulating interval discrepancy (F. & Gurel-Gurevich 2016+)

In a power of two-choices process on $\mathcal{U}[0,1]$, the chooser can obtain

$$\lim_{n\to\infty} \Pr\left(\max_{a,b} |\mu^n((a,b)) - \mu((a,b))| < \frac{C\log^3 N}{N} \right) = 1.$$

- In fact power of one-retry is enough (to be explained shortly).
- There exists a single universal strategy which obtains this for all N.

Global strategy that regulates interval variation

We show that a global strategy can reduce discrepancy significantly.

Power of two-choices in regulating interval discrepancy (F. & Gurel-Gurevich 2016+)

In a power of two-choices process on $\mathcal{U}[0,1]$, the chooser can obtain

$$\lim_{n\to\infty} \Pr\left(\max_{a,b} |\mu^n((a,b)) - \mu((a,b))| < \frac{C\log^3 N}{N}\right) = 1.$$

- In fact power of one-retry is enough (to be explained shortly).
- There exists a single universal strategy which obtains this for all N.
- Cf. lower bound of $\frac{C\log N}{N}$, no choice estimate $\sqrt{\frac{C\log N}{N}}$.

Global strategy that regulates interval variation

Discrete counterpart

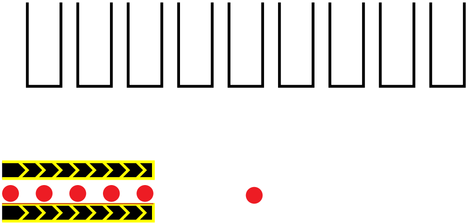
For N balls on $\mathcal{U}([M])$, a probabilistic retry strategy obtains

$$\Pr\left(\max_{a<b\in[M]} |\mu^n([a,b]) - \mu([a,b])| > \Delta \log^3 M\right) \leq Ce^{-c\Delta}.$$

Stochastic point of view on the power of two choices

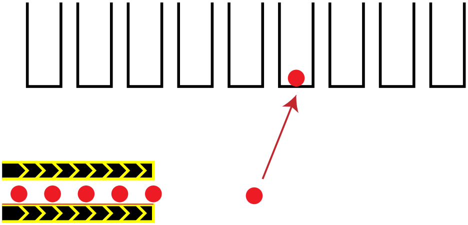
Two-choices and One-retry

One-retry: A related intermediate setup. The chooser is only offered a chance to re-roll the target bin once per ball.



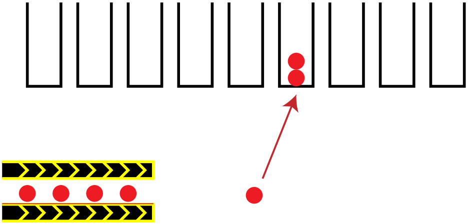
Two-choices and One-retry

One-retry: A related intermediate setup. The chooser is only offered a chance to re-roll the target bin once per ball.



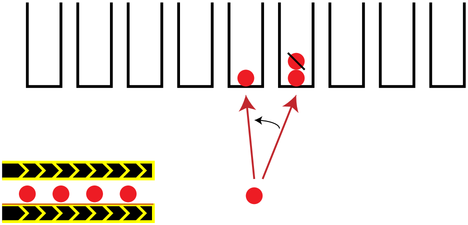
Two-choices and One-retry

One-retry: A related intermediate setup. The chooser is only offered a chance to re-roll the target bin once per ball.



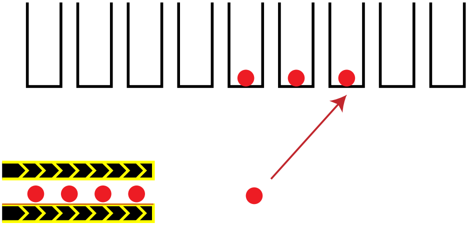
Two-choices and One-retry

One-retry: A related intermediate setup. The chooser is only offered a chance to re-roll the target bin once per ball.



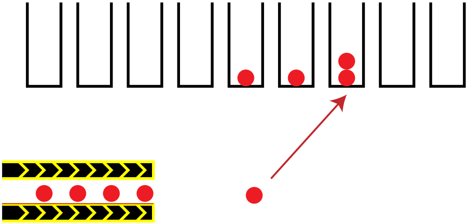
Two-choices and One-retry

One-retry: A related intermediate setup. The chooser is only offered a chance to re-roll the target bin once per ball.



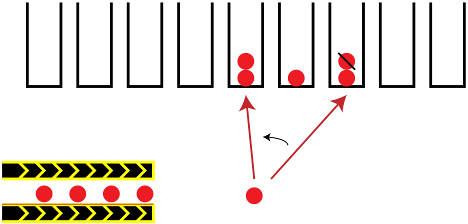
Two-choices and One-retry

One-retry: A related intermediate setup. The chooser is only offered a chance to re-roll the target bin once per ball.



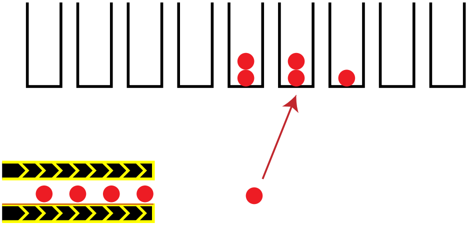
Two-choices and One-retry

One-retry: A related intermediate setup. The chooser is only offered a chance to re-roll the target bin once per ball.



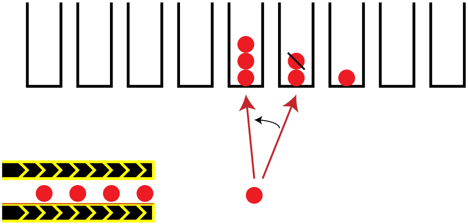
Two-choices and One-retry

One-retry: A related intermediate setup. The chooser is only offered a chance to re-roll the target bin once per ball.



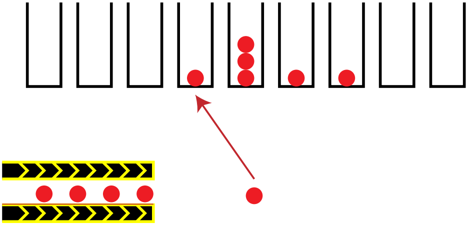
Two-choices and One-retry

One-retry: A related intermediate setup. The chooser is only offered a chance to re-roll the target bin once per ball.



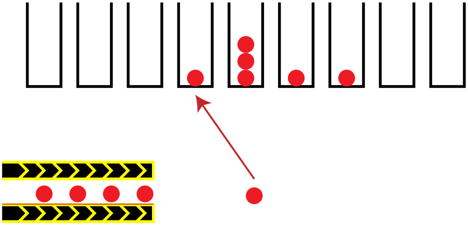
Two-choices and One-retry

One-retry: A related intermediate setup. The chooser is only offered a chance to re-roll the target bin once per ball.



Two-choices and One-retry

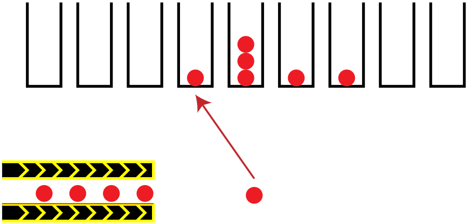
One-retry: A related intermediate setup. The chooser is only offered a chance to re-roll the target bin once per ball.



Equivalent to being oblivious to one of the two bins in the two choices setup.

Two-choices and One-retry

One-retry: A related intermediate setup. The chooser is only offered a chance to re-roll the target bin once per ball.



Equivalent to being oblivious to one of the two bins in the two choices setup. Asymptotic discrepancy like two-choices when $\#$balls $\gg$ $\#$ bins.

Stochastic Approach to Two-choices: Regulating Balls and Bins

Large family of one-retry distributions

What kind of distributions could be realized using one retry?

Large family of one-retry distributions

What kind of distributions could be realized using one retry?

Observation I

Every distribution whose density $\frac{1}{2M} < g(x) < \frac{3}{2M}$ on $[M]$ could be realized by a (probabilistic) one-retry strategy.

Large family of one-retry distributions

What kind of distributions could be realized using one retry?

Observation I

Every distribution whose density $\frac{1}{2M} < g(x) < \frac{3}{2M}$ on $[M]$ could be realized by a (probabilistic) one-retry strategy.

In general any distribution with Radon-Nikodim Derivative w.r.t the base distribution between 0.5 and 1.5 could be realized

Large family of one-retry distributions

What kind of distributions could be realized using one retry?

Observation I

Every distribution whose density $\frac{1}{2M} < g(x) < \frac{3}{2M}$ on $[M]$ could be realized by a (probabilistic) one-retry strategy.

In general any distribution with Radon-Nikodim Derivative w.r.t the base distribution between 0.5 and 1.5 could be realized

Proof. Write $f(x) = \frac{3}{2} - M \cdot g(x)$, and re-roll x with probability $f(x)$.

Large family of one-retry distributions

What kind of distributions could be realized using one retry?

Observation I

Every distribution whose density $\frac{1}{2M} < g(x) < \frac{3}{2M}$ on $[M]$ could be realized by a (probabilistic) one-retry strategy.

In general any distribution with Radon-Nikodim Derivative w.r.t the base distribution between 0.5 and 1.5 could be realized

Proof. Write $f(x) = \frac{3}{2} - M \cdot g(x)$, and re-roll x with probability $f(x)$.
The probability of a random bin to be re-rolled is $\sum_{i=1}^{M} \frac{f(x)}{M} = \frac{1}{2}$.

Large family of one-retry distributions

What kind of distributions could be realized using one retry?

Observation I

Every distribution whose density $\frac{1}{2M} < g(x) < \frac{3}{2M}$ on $[M]$ could be realized by a (probabilistic) one-retry strategy.

In general any distribution with Radon-Nikodim Derivative w.r.t the base distribution between 0.5 and 1.5 could be realized

Proof. Write $f(x) = \frac{3}{2} - M \cdot g(x)$, and re-roll x with probability $f(x)$.
The probability of a random bin to be re-rolled is $\sum_{i=1}^{M} \frac{f(x)}{M} = \frac{1}{2}$.
Hence the probability that x is chosen is now

$$\frac{1 - f(x)}{M} + \frac{1}{2M} = g(x)$$

.

Self regulating poisson point process

What kind of distributions do we wish to realize?

Self regulating poisson point process

What kind of distributions do we wish to realize?

- First - how to recover original balls and bins result with $N \gg M$.

Self regulating poisson point process

What kind of distributions do we wish to realize?
- First - how to recover original balls and bins result with $N \gg M$.
Consider a Poisson point process X_t with changing causal intensity $\lambda(t)$, defined by

$$\begin{aligned} \lambda_t^i &= 1+\theta & X_t &\leq t \\ \lambda_t^i &= 1-\theta & X_t &> t \end{aligned}$$

.

Self regulating poisson point process

What kind of distributions do we wish to realize?
- First - how to recover original balls and bins result with $N \gg M$.
Consider a Poisson point process X_t with changing causal intensity $\lambda(t)$, defined by

$$\lambda_t^i = 1 + \theta \qquad X_t \leq t$$
$$\lambda_t^i = 1 - \theta \qquad X_t > t$$

Proposition

For such a process $\Pr\left(|X_t - t| > \frac{\Delta}{\theta}\right) < C\theta^{-2}e^{-\Delta/3}$

.

Self regulating poisson point process

What kind of distributions do we wish to realize?
- First - how to recover original balls and bins result with $N \gg M$.
Consider a Poisson point process X_t with changing causal intensity $\lambda(t)$, defined by

$$\lambda_t^i = 1 + \theta \qquad X_t \leq t$$
$$\lambda_t^i = 1 - \theta \qquad X_t > t$$

Proposition

For such a process $\Pr\left(|X_t - t| > \frac{\Delta}{\theta}\right) < C\theta^{-2}e^{-\Delta/3}$

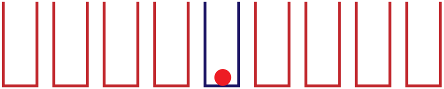
i.e. such a process has typical fluctuation $O(\theta)$.

Recovering the $N \gg M$ balls and bins result

Now let us consider M such self regulating processes.

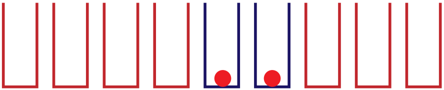
Recovering the $N \gg M$ balls and bins result

Now let us consider M such self regulating processes.



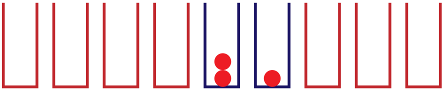
Recovering the $N \gg M$ balls and bins result

Now let us consider M such self regulating processes.



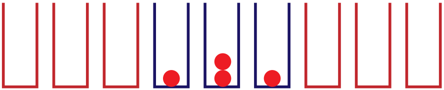
Recovering the $N \gg M$ balls and bins result

Now let us consider M such self regulating processes.



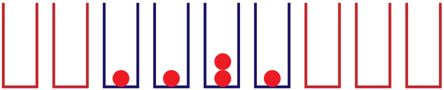
Recovering the $N \gg M$ balls and bins result

Now let us consider M such self regulating processes.



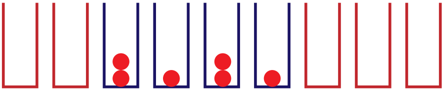
Recovering the $N \gg M$ balls and bins result

Now let us consider M such self regulating processes.



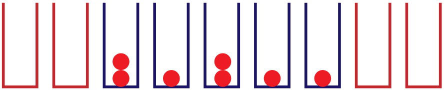
Recovering the $N \gg M$ balls and bins result

Now let us consider M such self regulating processes.



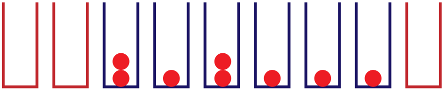
Recovering the $N \gg M$ balls and bins result

Now let us consider M such self regulating processes.



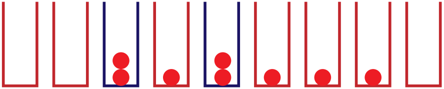
Recovering the $N \gg M$ balls and bins result

Now let us consider M such self regulating processes.



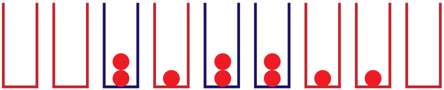
Recovering the $N \gg M$ balls and bins result

Now let us consider M such self regulating processes.



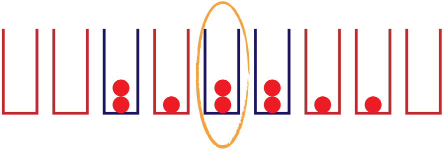
Recovering the $N \gg M$ balls and bins result

Now let us consider M such self regulating processes.



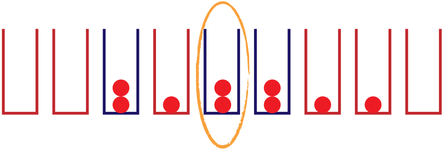
Recovering the $N \gg M$ balls and bins result

Now let us consider M such self regulating processes.



Recovering the $N \gg M$ balls and bins result

Now let us consider M such self regulating processes.



Among M such processes, the extremal fluctuation is $O(\theta \log M)$.

Recovering the $N \gg M$ balls and bins result

The plan: realize a self regulating process with $\theta = 1/5$ at every bin.

Recovering the $N \gg M$ balls and bins result

The plan: realize a self regulating process with $\theta = 1/5$ at every bin.

Can we do it?

Recovering the $N \gg M$ balls and bins result

The plan: realize a self regulating process with $\theta = 1/5$ at every bin.

Can we do it?
Yes - p_x, the probability of the next ball falling into a bin x is bounded by

$$\frac{1}{2M} \leq \frac{1-\theta}{M(1+\theta)} \leq p_x \leq \frac{1+\theta}{M(1-\theta)} \leq \frac{3}{2M}.$$

Recovering the $N \gg M$ balls and bins result

The plan: realize a self regulating process with $\theta = 1/5$ at every bin.

Can we do it?
Yes - p_x, the probability of the next ball falling into a bin x is bounded by

$$\frac{1}{2M} \leq \frac{1-\theta}{M(1+\theta)} \leq p_x \leq \frac{1+\theta}{M(1-\theta)} \leq \frac{3}{2M}.$$

Suppose we use this strategy until time N/M.
We expect N balls, Maximal load of $N/M + \Theta(\log M)$.

Recovering the $N \gg M$ balls and bins result

The plan: realize a self regulating process with $\theta = 1/5$ at every bin.

Can we do it?
Yes - p_x, the probability of the next ball falling into a bin x is bounded by

$$\frac{1}{2M} \le \frac{1-\theta}{M(1+\theta)} \le p_x \le \frac{1+\theta}{M(1-\theta)} \le \frac{3}{2M}.$$

Suppose we use this strategy until time N/M.
We expect N balls, Maximal load of $N/M + \Theta(\log M)$.
All that is left is to show that the same holds after N balls were distributed. - i.e. show concentration of the stopping time.

Stochastic Approach to Two-choices:

Regulating Interval Partition

Balancing poisson point process

Consider two point processes X_t^0, X_t^1 with changing causal intensities $\lambda^0(t), \lambda^1(t)$ (which may depend on other variables), which satisfy

$$
\begin{aligned}
\lambda_t^1, \lambda_t^0 &< 2 && \forall t \\
\lambda_t^0 - \lambda_t^1 &\geq \theta && X_t^0 \leq X_t^1 \\
\lambda_t^0 - \lambda_t^1 &\leq -\theta && X_t^0 > X_t^1
\end{aligned}
$$

Balancing poisson point process

Consider two point processes X_t^0, X_t^1 with changing causal intensities $\lambda^0(t), \lambda^1(t)$ (which may depend on other variables), which satisfy

$$\begin{aligned} \lambda_t^1, \lambda_t^0 &< 2 && \forall t \\ \lambda_t^0 - \lambda_t^1 &\geq \theta && X_t^0 \leq X_t^1 \\ \lambda_t^0 - \lambda_t^1 &\leq -\theta && X_t^0 > X_t^1 \end{aligned}$$

Proposition

For such a process $\Pr\left(|X_t^1 - X_t^0| > \frac{\Delta}{\theta}\right) < C\theta^{-2}e^{-\Delta/2}$

Balancing poisson point process

Consider two point processes X_t^0, X_t^1 with changing causal intensities $\lambda^0(t), \lambda^1(t)$ (which may depend on other variables), which satisfy

$$\begin{aligned} \lambda_t^1, \lambda_t^0 &< 2 && \forall t \\ \lambda_t^0 - \lambda_t^1 &\geq \theta && X_t^0 \leq X_t^1 \\ \lambda_t^0 - \lambda_t^1 &\leq -\theta && X_t^0 > X_t^1 \end{aligned}$$

Proposition

For such a process $\Pr\left(|X_t^1 - X_t^0| > \frac{\Delta}{\theta}\right) < C\theta^{-2}e^{-\Delta/2}$

We now define an intensity μ_t on $[0,1]$ which could be realized by a one-retry strategy. We will show that this intensity regulates all diadic intervals.

Balancing poisson point process

Consider two point processes X_t^0, X_t^1 with changing causal intensities $\lambda^0(t), \lambda^1(t)$ (which may depend on other variables), which satisfy

$$
\begin{aligned}
\lambda_t^1, \lambda_t^0 &< 2 && \forall t \\
\lambda_t^0 - \lambda_t^1 &\geq \theta && X_t^0 \leq X_t^1 \\
\lambda_t^0 - \lambda_t^1 &\leq -\theta && X_t^0 > X_t^1
\end{aligned}
$$

Proposition

For such a process $\Pr\left(|X_t^1 - X_t^0| > \frac{\Delta}{\theta}\right) < C\theta^{-2}e^{-\Delta/2}$

We now define an intensity μ_t on $[0,1]$ which could be realized by a one-retry strategy. We will show that this intensity regulates all diadic intervals.
Each $(a,b) \subset [0,1]$ has $(a,b) = \sum_{i=0}^{\lfloor \log_2 N \rfloor} I_i + R$ where I_i are diadic and $|R| < 1/N$.

Hierarchy of drifts

We build a hierarchy of $\log_2 N$ drifts of strength $C/\log N$, to control diadic discrepancies.

Hierarchy of drifts

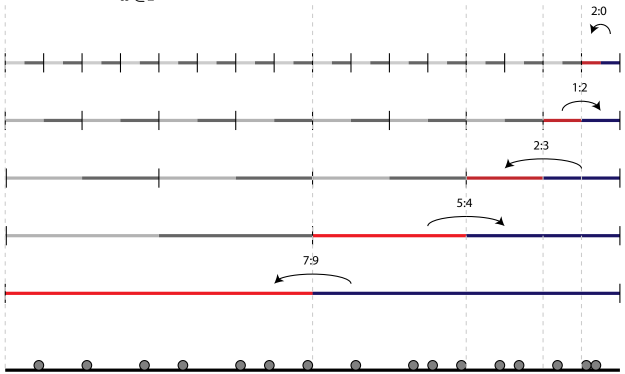
We build a hierarchy of $\log_2 N$ drifts of strength $C/\log N$, to control diadic discrepancies. The intensity of every point x is

$$1+\frac{C}{\log N}\underbrace{\sum_{x\in I}(-1)^{\mathbf{1}\{x\in I_{\text{left}}\}}(-1)^{\mathbf{1}\{\mu_t(I_{\text{left}}\leq I_{\text{right}})\}}}_{\text{drift hierarchy}}+\underbrace{c(-1)^{\mathbf{1}\{\mu_t([0,1])>t\}}}_{\text{rate regulating term}}$$

Hierarchy of drifts

We build a hierarchy of $\log_2 N$ drifts of strength $C/\log N$, to control diadic discrepancies. The intensity of every point x is

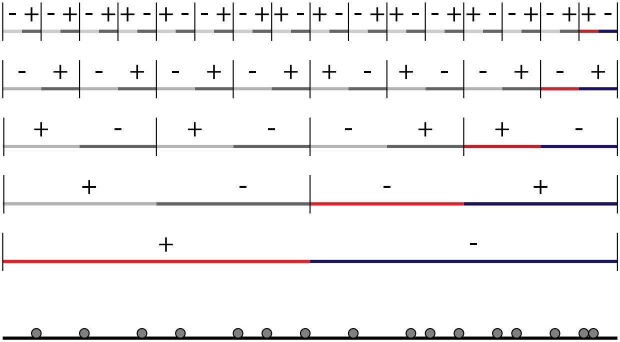
$$1+\frac{C}{\log N}\underbrace{\sum_{x\in I}(-1)^{\mathbf{1}\{x\in I_{\text{left}}\}}(-1)^{\mathbf{1}\{\mu_t(I_{\text{left}}\leq I_{\text{right}})\}}}_{\textit{drift hierarchy}}+\underbrace{c(-1)^{\mathbf{1}\{\mu_t([0,1])>t\}}}_{\textit{rate regulating term}}$$



Hierarchy of drifts

We build a hierarchy of $\log_2 N$ drifts of strength $C/\log N$, to control diadic discrepancies. The intensity of every point x is

$$1+\frac{C}{\log N}\overset{\textit{drift hierarchy}}{\sum_{x\in I}(-1)^{\mathbf{1}\{x\in I_{\text{left}}\}}(-1)^{\mathbf{1}\{\mu_t(I_{\text{left}}\le I_{\text{right}})\}}}+\overset{\textit{rate regulating term}}{c(-1)^{\mathbf{1}\{\mu_t([0,1])>t\}}}$$



Future directions

Future directions

- Other spaces

Future directions

- Other spaces
- Other measures of discrepancy.

Future directions

- Other spaces
- Other measures of discrepancy.
- Reducing the power of the log.

THANK YOU.