Indistinguishability / Birthday Bounds

Big Picture:

<u>Previously</u>: Impossible **in principle** to distinguish libraries
$\Rightarrow$ information perfectly hidden

"<u>Modern crypto</u>" Just **really hard** to distinguish
(comparable to probability of blindly guessing a long key)

$\mathcal{L}_{\text{samp-L}}$

```
SAMP():
  r ← {0,1}^λ
  return r
```

"with replacement"

$\approx$

$\mathcal{L}_{\text{samp-R}}$

```
R := ∅
SAMP():
  r ← {0,1}^λ \ R
  add r to R
  return r
```

all strings except ones in R

"without replacement"

Since 2 libs are $\approx$, I can design a crypto scheme that samples values uniformly from $\{0,1\}^\lambda$

Adv can't tell whether I'm choosing w/ or w/o replacement

$\Rightarrow$ can "just assume" I won't see a repeated value

"Obvious" Distinguisher

A:

```
call SAMP q times
If repeated value, return 1
else               return 0
```

$\Pr[A \diamond \mathcal{L}_{\text{samp-R}} \Rightarrow 1] = 0$

$\Pr[A \diamond \mathcal{L}_{\text{samp-L}} \Rightarrow 1] > 0$

Def: $\text{CollProb}(q,\lambda)$ = probability of seeing a repeated output of SAMP in $A \diamond \mathcal{L}_{\text{samp-L}}$ (q calls to SAMP)

= Advantage of "obvious distinguisher" A

Claim: No calling program that makes q calls to SAMP can distinguish with advantage better than $\text{CollProb}(q,\lambda)$ (book)

Q: What is $\text{CollProb}(q,\lambda)$?

---

say $s_1 \cdots s_q$ are indep. samples from $\{0,1\}^\lambda$

$$
\begin{aligned}
\Pr[s_1 \cdots s_q \text{ all distinct}] &= \Pr[s_2 \neq s_1] \\
&\times \Pr[s_3 \notin \{s_1, s_2\}] \\
&\times \Pr[s_4 \notin \{s_1 \ldots s_3\}] \\
&\times \cdots
\end{aligned}
$$

$$= \left(1 - \frac{1}{2^\lambda}\right)\left(1 - \frac{2}{2^\lambda}\right)\left(1 - \frac{3}{2^\lambda}\right) \cdots \left(1 - \frac{q-1}{2^\lambda}\right)$$

$$
\boxed{
\begin{aligned}
\text{CollProb}(q,\lambda) &= 1 - \Pr[\text{distinct}] \\
&= 1 - \prod_{i=1}^{q}\left(1 - \frac{i-1}{2^\lambda}\right)
\end{aligned}
}
$$

Approximations

$$0.632 \, \frac{q(q-1)}{2^{\lambda+1}} \leq \text{CollProb}(q, \lambda) \leq \frac{q(q-1)}{2^{\lambda+1}}$$

$$\left(\text{for} \quad q \leq \sqrt{2^{\lambda+1}} \right)$$

In summary:

take q samples from $\{0,1\}^\lambda$

$\Rightarrow$ probability $\Theta\left(\frac{q^2}{2^\lambda}\right)$ of repeat