

# 1 The probability of unique keys

To sample  $N$  numbers from a size  $M$  pool with replacement, there are in total  $M^N$  cases. If the sampling is without replacement, there are  $P(M, N)$  cases, where  $P(M, N)$ , also known as  $nPr$ , is defined as

$$P(M, N) = \frac{M!}{(M-N)!}.$$

So the probability of getting non-duplicating keys (the correct case) is

$$\Pr[\text{unique keys}] = \frac{P(M, N)}{M^N}$$

According to a narrowed version of Stirling's formula by Robbins<sup>1</sup>

$$\sqrt{2\pi n} n^{n+\frac{1}{2}} e^{-n} e^{\frac{1}{12n+1}} \leq n! \leq \sqrt{2\pi n} n^{n+\frac{1}{2}} e^{-n} e^{\frac{1}{12n}}$$

we then have

$$\frac{M^{M+\frac{1}{2}} e^{-M} e^{\frac{1}{12M+1}}}{(M-N)^{(M-N)+\frac{1}{2}} e^{-(M-N)} e^{\frac{1}{12(M-N)}}} \leq P(M, N) \leq \frac{M^{M+\frac{1}{2}} e^{-M} e^{\frac{1}{12M}}}{(M-N)^{(M-N)+\frac{1}{2}} e^{-(M-N)} e^{\frac{1}{12(M-N)+1}}}$$

which simplify to

$$\frac{M^{M+\frac{1}{2}} \cdot e^{-N}}{(M-N)^{(M-N)+\frac{1}{2}}} \cdot e^{\frac{1}{12M+1} - \frac{1}{12(M-N)}} \leq P(M, N) \leq \frac{M^{M+\frac{1}{2}} \cdot e^{-N}}{(M-N)^{(M-N)+\frac{1}{2}}} \cdot e^{\frac{1}{12M} - \frac{1}{12(M-N)+1}}$$

so

$$\begin{aligned} \frac{M^{(M-N)+\frac{1}{2}} \cdot e^{-N}}{(M-N)^{(M-N)+\frac{1}{2}}} \cdot e^{\frac{1}{12M+1} - \frac{1}{12(M-N)}} &\leq \\ \Pr[\text{unique keys}] &\leq \frac{M^{(M-N)+\frac{1}{2}}}{(M-N)^{(M-N)+\frac{1}{2}} e^N} \cdot e^{\frac{1}{12M} - \frac{1}{12(M-N)+1}} \end{aligned}$$

which simplify to

$$\begin{aligned} \left( \frac{M}{M-N} \right)^{(M-N)+\frac{1}{2}} \cdot e^{-N} \cdot e^{\frac{1}{12M+1} - \frac{1}{12(M-N)}} &\leq \\ \Pr[\text{unique keys}] &\leq \left( \frac{M}{M-N} \right)^{(M-N)+\frac{1}{2}} \cdot e^{-N} \cdot e^{\frac{1}{12M} - \frac{1}{12(M-N)+1}} \end{aligned}$$

we will later write the above inequality as

$$f(M, N) \leq \Pr[\text{unique keys}] \leq g(M, N)$$

---

<sup>1</sup><https://doi.org/10.2307/2308012>

## 2 Bound

We want to guarantee that randperm algorithm's succeed probability above a certain threshold  $\Pr[\text{unique keys}] \geq q$ . We can achieve this by requiring  $f(M, N) \geq q$ . Note that

$$\begin{aligned} f(M, N) &= \left( \frac{M}{M-N} \right)^{(M-N)+\frac{1}{2}} \cdot e^{-N} \cdot e^{\frac{1}{12M+1} - \frac{1}{12(M-N)}} \geq q \\ &= \left( 1 + \frac{N}{M-N} \right)^{(M-N)+\frac{1}{2}} \cdot e^{-N} \cdot e^{\frac{1}{12M+1} - \frac{1}{12(M-N)}} \geq q \end{aligned}$$

then

$$\log f(M, N) = \left( M - N + \frac{1}{2} \right) \log \left( 1 + \frac{N}{M-N} \right) - N + \frac{1}{12M+1} - \frac{1}{12(M-N)}$$

since

$$\begin{aligned} \log \left( 1 + \frac{N}{M-N} \right) &\geq \frac{N}{M-N} - \frac{1}{2} \left( \frac{N}{M-N} \right)^2 \\ M - N + \frac{1}{2} &\geq M - N \\ \frac{1}{12M+1} &\geq 0 \end{aligned}$$

we have

$$\log f(M, N) \geq (M-N) \left[ \frac{N}{M-N} - \frac{1}{2} \left( \frac{N}{M-N} \right)^2 \right] - N - \frac{1}{12(M-N)} = -\frac{1}{12} \cdot \frac{6N^2+1}{M-N}$$

So, as long as

$$-\frac{1}{12} \cdot \frac{6N^2+1}{M-N} \geq \log q$$

that is

$$M \geq N - \frac{6N^2+1}{12 \cdot \log q}$$

there is a guarantee that the probability of getting unique keys is above  $q$ .

### 3 Summary

So, in order to have  $N$  different random numbers at probability  $\geq q$ , you will need a

$$\left\lceil \log_2 \left( N - \frac{6N^2 + 1}{12 \cdot \log q} \right) \right\rceil$$

bit random number generator.

Plotting the above equation for  $q = 0.9$ . The plot of the above function is shown below, the  $x$ -axis is the number  $n$  has  $N = 2^n$

