Birthday bounds

Classical birthday bound.
Let $y_1, \ldots, y_q$ be uniformly and independently drawn from a size-N set. Then

$$1 - e^{q(q-1)/2N} \leq \Pr\left[\exists i \neq j, y_i = y_j\right] \leq \frac{q(q-1)}{2N}.$$

If furthermore $q \leq \sqrt{2N}$, the lower bound is $\geq \frac{q(q-1)}{4N}$.

Proof of the upper bound. The probability that two independent elements drawn from a size-N set are equal is at most $1/N$. Therefore, using the union bound, $\Pr\left[\exists i \neq j, y_i = y_j\right] \leq \sum_{i \neq j} 1/N = q(q-1)/2N$. □

Proof of the lower bound. Let N_i be the event "there is no collision amongst $y_1, \ldots, y_i$." We want to lower bound $\Pr\left[\neg N_q\right]$, that is to upper bound $\Pr\left[N_q\right]$. Using the law of total probability, $\Pr\left[N_q\right] = \Pr\left[N_q \middle| N_{q-1}\right] \Pr\left[N_{q-1}\right] + \Pr\left[N_q \middle| \neg N_{q-1}\right] \Pr\left[N_{q-1}\right]$. But $\Pr\left[N_q \middle| N_{q-1}\right] = 0$. Therefore, $\Pr\left[N_q\right] = \Pr\left[N_q \middle| N_{q-1}\right] \Pr\left[N_{q-1}\right]$. Hence, by induction

$$\Pr\left[N_q\right] = \Pr[N_1] \cdot \Pr[N_2 | N_1] \cdots \Pr\left[N_q \middle| N_{q-1}\right].$$

Now, $\Pr[N_1] = 1$ (since there cannot be a collision between one element). And for $i > 1$, $\Pr[N_i | N_{i-1}]$ is the probability that y_i collides with a y_j, $j < i$, assuming they are all different. Since y_i is drawn independently of the previous y_j's, $\Pr[N_i | N_{i-1}] = 1 - (i-1)/N$.
Together, we obtain that $\Pr\left[N_q\right] = \prod_{i=1}^{q-1} 1 - i/N$. Since $1 + x \leq e^x$ for all x,

$$\Pr\left[N_q\right] \leq \prod_{i=1}^{q-1} e^{-i/N} = e^{-\frac{1}{N}\sum_{i=1}^{q-1} i} = e^{-q(q-1)/2N}.$$

Hence $\Pr\left[\exists i \neq j, y_i = y_j\right] \geq 1 - e^{-q(q-1)/2N}$. Finally, if $q \leq \sqrt{2N}$, $q(q-1)/2N \leq 1$ and since $e^{-x} \leq 1 - x/2$ for $0 \leq x \leq 1$, the probability is at least $q(q-1)/4N$. □

Variant of the birthday bound.
Let $y_1, \ldots, y_q, z_1, \ldots, z_q$ be uniformly and independently drawn from a size-N set. Then

$$1 - e^{-q^2/N} \leq \Pr\left[\exists i, j, y_i = z_j\right] \leq \frac{q^2}{N}.$$

If furthermore $q \leq \sqrt{N}$, the lower bound is $\geq \frac{q^2}{2N}$.

Proof of the upper bound. The probability that two independent elements drawn from a size-N set are equal is at most $1/N$. Therefore, using the union bound, $\Pr\left[\exists i, j, y_i = z_j\right] \leq \sum_{i,j} 1/N = q^2/N$. □

Proof of the lower bound. Let N_i be the event "for all j, $y_i \neq z_j$. We want to lower bound $\Pr\left[\bigvee_i \neg N_i\right]$, that is to upper bound $\Pr\left[\bigwedge_i N_i\right]$. Assume that we first draw all the values z_j, $1 \leq j \leq q$. Then N_i is the event"$y_i \notin Z$" where $Z = \{z_1, \ldots, z_q\}$. Since the y_i's are drawn independently, the N_i's are independent events. Therefore, $\Pr\left[\bigwedge_i N_i\right] = \prod_i \Pr[N_i]$. Now, $N_i = \bigvee_j [y_i = z_j]$. Assume now on the contrary that we first draw y_i, then the z_j's independently: We see that the events $[y_i = z_j]$ are independent. That is $\Pr[N_i] = \prod_j \Pr\left[y_i = z_j\right] = (1 - 1/N)^q$. Now using $1 + x \leq e^x$, $\Pr[N_i] \leq e^{-q/N}$ and $\Pr\left[\bigwedge_i N_i\right] \leq e^{-q^2/N}$. Therefore, $\Pr\left[\exists i, j, y_i = z_j\right] \geq 1 - e^{-q^2/N}$. If $q \leq \sqrt{N}$, $q^2/N \leq 1$ and since $e^{-x} \leq 1 - x/2$ for $0 \leq x \leq 1$, the probability is at least $q^2/2N$. □