T 9.1

a)

> Urn of N balls of N distinct colors.
>
> Pick a ball, record the color, put the ball back in the urn.
> Repeat $k \geq 1$ times.
>
> Compute $\mathbb{P}(\text{picked } k \text{ distinct balls})$

First note that since each ball is colored in a unique way, we identify a ball by its color.

Let $E = \{\text{picked } k \text{ distinct balls after } k \text{ iterations}\}$

If $k > N$, then $\mathbb{P}(E) = 0$.

Assume $k \in \{1, \dots, N\}$.

Although the drawn ball is re-inserted into the urn every time, we have that every single one of the k draws is dependent of the previous ones.

Let $X_i \in \{1, \dots, N\}$ be the r.v. representing the color of the k-th ball drawn.

We have that $\mathbb{P}(X_i = c) = \frac{1}{N} \quad \forall i \in \{1, \dots, k\}, \forall c \in \{1, \dots, N\}$

Moreover, we clearly also have that:

$$\mathbb{P}(E) = \mathbb{P}\left(X_i \neq X_j \quad \forall i, j \in \{1, \dots, k\} \text{ and } i \neq j\right)$$

$$= \mathbb{P}\left(X_2 \notin \{X_1\}, X_3 \notin \{X_2, X_1\}, X_4 \notin \{X_3, X_2, X_1\}, \dots, X_k \notin \{X_{k-1}, \dots, X_1\}\right)$$

$$= \mathbb{P}(X_2 \neq X_1) \cdot \mathbb{P}(X_3 \neq X_2 \mid X_2 \neq X_1) \cdot \dots \cdot \mathbb{P}(X_k \neq X_{k-1} \mid X_{k-1} \neq X_{k-2})$$

$$= \left(1 - \frac{1}{N}\right) \cdot \left(1 - \frac{2}{N}\right) \cdot \dots \cdot \left(1 - \frac{k-1}{N}\right)$$

$$= \frac{(N-1) \cdot (N-2) \cdot \dots \cdot (N-(k-1))}{N^{k-1}} = N \cdot \frac{(N-1) \cdot (N-2) \cdot \dots \cdot (N-k+1)}{N^k}$$

$$= \frac{N!}{N^k (N-k)!}$$

b) What is the minimal number of people in a room so that the probability of having two of them born on the same day exceeds 50% ?

Assuming all people are born in a non-leap year, the answer is 23 persons. Let's see why.

$E' = \{$ having 2 people born on the same day $\}$

It logically follows that

$$\mathbb{P}(E') = 1 - \mathbb{P}(E) = 1 - \frac{N!}{N^k (N-k)!}$$

where $N = 365$ is fixed and we will show that $k = 23$ so that $\mathbb{P}(E') > \frac{1}{2}$.

We can view $\mathbb{P}(E')$ as a function of k. Let $f(k) = 1 - \frac{N!}{N^k (N-k)!}$

We note that $f(k)$ is a increasing function w.r.t. $k > 0$.

Indeed, $f(k+1) > f(k) \iff 1 - \frac{N!}{N^{k+1}(N-k-1)!} > 1 - \frac{N!}{N^k (N-k)!}$

$$\iff \frac{1}{N(N-k-1)!} < \frac{1}{(N-k)!} \overset{(N-k)! \neq 0}{\iff} \frac{(N-k)!}{N(N-k-1)!} < 1$$

$$\iff N-k < N \iff k > 0 \text{ , that is true } \checkmark$$

Since $\mathbb{P}(E') = f(k)$ is an increasing function w.r.t. $k > 0$, we can simply evaluate it for increasing values of k until we find $\bar{k}$ s.t. $f(\bar{k}) > \frac{1}{2}$ (with $N = 365$ fixed).

- $k = 1$: $\mathbb{P}(E') = 0$
- $k = 2$: $\mathbb{P}(E') = 0.00274 = \frac{1}{365}$

  ⋮

- $k = 22$: $\mathbb{P}(E') = 0.4757$
- $k = 23$: $\mathbb{P}(E') = 0.5073$

So $\bar{k} = 23$ is the minimal number of people in a room so that the probability of having two of them born on the same day exceeds 50%.

c) For the setting in (a), show that the expected number of draws needed to record some color twice is $1 + Q(N)$, where

$$Q(N) = \sum_{k=1}^{N} \frac{N!}{(N-k)!\, N^k}$$

Let R be the r.v. representing the number of draws necessary to record some color twice.

We have to compute

from some widely known result in probabilities since $R \in \{0,1,2,\dots\}$

$$\mathbb{E}(R) := \sum_{k=0}^{\infty} k \cdot \mathbb{P}(R=k) = \sum_{k=0}^{\infty} \mathbb{P}(R>k)$$

Analogously to what was already mentioned in part (a), after $N+1$ draws we will necessarily have already picked the same ball twice, since there are only N distinct balls in the urn.

Therefore, we have that

$$\mathbb{E}(R) := \sum_{k=0}^{\infty} k \cdot \mathbb{P}(R=k) = \sum_{k=0}^{\infty} \mathbb{P}(R>k) = \sum_{k=0}^{N} \mathbb{P}(R>k)$$

since $\mathbb{P}(R>\tilde{k}) = 0 \quad \forall \tilde{k} > N$

Now we note that

$$\mathbb{P}(R>k) = \mathbb{P}(E)$$

since they both refer to the event where we have not picked the same ball twice for the first k draws.

Therefore

$$\mathbb{E}(R) = \sum_{k=0}^{N} \mathbb{P}(R>k) = \sum_{k=0}^{N} \frac{N!}{N^k (N-k)!} = 1 + \sum_{k=1}^{N} \frac{N!}{N^k (N-k)!} = 1 + Q(N) \quad \square$$