Exercices sur les algorithmes probabilités
Nom du rédacteur : Johanne Cohen
https://www.lri.fr/~jcohen/cours-probabiliste.html

Algorithmes probabilistes : paradoxe des anniversaires

Supposons qu'il y a 30 personnes dans la salle. Est-il plus probable que deux personnes dans la salle partagent le même anniversaire ou que deux personnes dans la salle ne partagent pas le même anniversaire ?

Nous supposerons que l'anniversaire de chaque personne est un jour de calendrier sur une année de 365 jours est choisi indépendamment et uniformément au hasard pour chaque personne.

Focalisons nous sur la date d'anniversaire d'une personne p.

Soit $\mathcal{A}$ l'ensemble des jours de l'année. Notons $X(p)$ la variable aléatoire désignant le jour de naissance de que la personne p soit née le jour j.

Question 1. Quelle est la probabilité que une personne soit née le jour j : *i. e.* calculer $\Pr(X_p = j)$?

> Comme il est supposé que le jour de naissance de la personne p est choisi uniformément au hasard. Cela signifie que
>
> $$\forall j \in \mathcal{A}, i \in \mathcal{A}, \quad \Pr(X_p = j) = \Pr(X_p = i)$$
>
> L'univers est $\Omega = \mathcal{A}$ et donc $\Pr(\Omega) = 1$. De plus comme deux évènements $X_p = j$ et $X_p = i$ sont disjoints si $i \neq j$, nous avons $\Pr(\Omega) = \sum_{j \in \mathcal{A}} \Pr(X_p = j)$
>
> Comme nous avons $\sum_{j \in \mathcal{A}} \Pr(X_p = j) = 1$ (car une personne a forcement un anniversaire et donc elle est née un jour). Donc nous avons $\Pr(X_p = j) = \frac{1}{|\mathcal{A}|}$.

Question 2. Quelle est la probabilité que une personne ne soit pas née le jour j : *i. e.* calculer $\Pr(X_p \neq j)$?

> Comme $\Pr(X_p \neq j) + \Pr(X_p = j) = 1$, on a
>
> $$\Pr(X_p \neq j) = 1 - \frac{1}{|\mathcal{A}|}$$

Notons $X_p \notin \mathcal{J}$ l'évènement représentant le fait que l'anniversaire de la personne p n'est pas un jour de $\mathcal{J}$.

Question 3. Quelle est la probabilité $\Pr(X_p \notin \mathcal{J})$?

Nous allons introduire l'évènement $X_p = \mathcal{J}$ représentant le fait que l'anniversaire de la personne p est un jour de $\mathcal{J}$. Comme $\Pr(X_p \notin \mathcal{J}) + \Pr(X_p = \mathcal{J}) = 1$, Il suffit de calculer la probabilité $\Pr(X_p = \mathcal{J})$

$$\Pr(X_p = \mathcal{J}) = \sum_{j \in \mathcal{J}} \Pr(X_p = j) = \sum_{j \in \mathcal{J}} \frac{1}{|\mathcal{A}|} = \frac{|\mathcal{J}|}{|\mathcal{A}|}$$

Donc, on obtient $\Pr(X_p = \mathcal{J}) = \left(1 - \frac{|\mathcal{J}|}{|\mathcal{A}|}\right)$.

Focalisons nous sur les dates d'anniversaires de deux personnes p_1 et p_2.

Question 4. Supposons que la personne p_1 est née le jour j. Quelle est la probabilité que la personne p_2 soit née le jour j : *i. e.* calculer $\Pr(X_{p_2} \neq j | X_{p_1} = j)$?

$$\Pr(X_{p_2} \neq j | X_{p_1} = j) = \frac{\Pr(X_{p_2} \neq j \wedge X_{p_1} = j)}{\Pr(X_{p_1} = j)}$$

Comme $X_{p_1} = j$ et $X_{p_2} \neq j$ sont deux évènements indépendants, on a

$$\Pr(X_{p_2} \neq j \wedge X_{p_1} = j) = \Pr(X_{p_2} \neq j) \cdot \Pr(X_{p_1} = j).$$

$$\Pr(X_{p_2} \neq j | X_{p_1} = j) = \frac{\Pr(X_{p_2} \neq j) \cdot \Pr(X_{p_1} = j)}{\Pr(X_{p_1} = j)} = \Pr(X_{p_2} \neq j)$$

Question 5. Supposons que la personne p_1 est née le jour j. Quelle est la probabilité que la personne 2 ne soit pas née le jour j : *i. e.* calculer $\Pr(X_{p_2} = j | X_{p_1} = j)$?

Soit $\mathcal{A}$ l'ensemble des jours de l'année.

$$\Pr(X_{p_2} \neq j | X_{p_1} = j) = \frac{\Pr(X_{p_2} \neq j \wedge X_{p_1} = j)}{\Pr(X_{p_1} = j)}$$

Comme $X_{p_1} = j$ et $X_{p_2} \neq j$ sont deux évènements indépendants,

$$\Pr(X_{p_2} \neq j | X_{p_1} = j) = \frac{\Pr(X_{p_2} \neq j) \cdot \Pr(X_{p_1} = j)}{\Pr(X_{p_1} = j)}$$

$$\Pr(X_{p_2} \neq j | X_{p_1} = j) = \Pr(X_{p_2} \neq j) = (1 - \frac{1}{|\mathcal{A}|})$$

Notons $\mathcal{B}_k$ l'événement que k personnes ne sont pas nées le même jour.

Question 6. Quelle est la probabilité que deux personnes ne sont pas nées le même jour.

$$\Pr(\mathcal{B}_2) = \sum_{j \in \mathcal{A}} (\Pr(X_{p_2} = j | X_{p_1} \neq j) \cdot \Pr(X_{p_1} = j)$$

$$\Pr(\mathcal{B}_2) = (1 - \frac{1}{|\mathcal{A}|})$$

La probabilité que la deuxième personne ait un anniversaire différent est de $(1 - 1/365)$.

Focalisons nous sur un ensemble de k personnes.

Nous allons introduire l'évènement suivant $\mathcal{E}(\mathcal{H}, J)$ qui désigne que

1. les anniversaires des personnes de $\mathcal{H}$ sont dans l'ensemble $\mathcal{J}$;
2. les personnes de $\mathcal{H}$ ne sont pas nées le même jour ;

Question 7. Définir l'évènement $\mathcal{B}_k$ en fonction des évènements $\mathcal{E}(\mathcal{H}, J)$. Donner sa probabilité en fonction de celles des évènements $\mathcal{E}(\mathcal{H}, J)$.

Notons $\mathcal{P}(\mathcal{A}, k)$ l'ensemble des parties de k éléments de l'ensemble $\mathcal{A}$. Observons que l'évènement se réalise $\mathcal{E}(\mathcal{H}, J)$ quand $X_{\mathcal{H}} = \mathcal{J}$ et $|\mathcal{H}| = |\mathcal{J}|$.
L'évènement $\mathcal{B}_k$ se réalise si l'un de ces évènements $\mathcal{E}(\mathcal{H}, J)$, pour tout élément $J \in \mathcal{P}(\mathcal{A}, k)$ se réalise :

$$\mathcal{B}_k = \bigcup_{J \in \mathcal{P}(\mathcal{A}, k)} \mathcal{E}(\mathcal{H}, J)$$

Soit J et J' deux éléments distincts de $\mathcal{P}(\mathcal{A}, k)$. Comme les évènements de $\mathcal{E}(\mathcal{H}, J)$ et $\mathcal{E}(\mathcal{H}, J')$ sont disjoints,

$$\Pr(\mathcal{B}_k) = \sum_{J \in \mathcal{P}(\mathcal{A}, k)} \Pr(\mathcal{E}(\mathcal{H}, J))$$

Question 8. Quelle est la probabilité que k personnes ne sont pas nées le même jour.

Nous pouvons prouver par récurrence que $\Pr(\mathcal{B}_k) = \Pi_{i=1}^{k-1}(1 - \frac{i}{|\mathcal{A}|})$.
Pour $k = 2$, l'hypothèse de récurrence est vrai. Supposons que l'hypothèse de récurrence est vrai pour k. Maintenant, nous allons vérifier que $\Pr(\mathcal{B}_{k+1}) = \prod_{i=1}^{k} \left(1 - \frac{i}{|\mathcal{A}|}\right)$.
Soit $\mathcal{H}$ un ensemble de $k + 1$ personnes.

$$\Pr(B_{k+1}) = \sum_{J \in \mathcal{P}(\mathcal{A}, k+1)} \Pr(\mathcal{E}(\mathcal{H}, J))$$

Concentrons nous sur l'évènement $\mathcal{E}(\mathcal{H}, J)$.

Soit p une personne de $\mathcal{H}$. Notons que $\mathcal{H}' = \mathcal{H}\backslash\{p\}$. Pour que $\mathcal{E}(\mathcal{H}, J)$ soit réalisable, il faut que p naisse un des jours de J et que les autres personnes naissent les autres jours de J

$$\mathcal{B}_{k+1} = \bigcup_{J \in \mathcal{P}(\mathcal{A},k)} \mathcal{E}(\mathcal{H}\backslash\{p\}, J) \wedge (X_p \notin J)$$

$$\begin{aligned}
\Pr(\mathcal{B}_{k+1}) &= \sum_{J \in \mathcal{P}(\mathcal{A},k)} \Pr(\mathcal{E}(\mathcal{H}\backslash\{p\}, J) \wedge (X_p \notin J)) \\
&= \sum_{J \in \mathcal{P}(\mathcal{A},k)} \Pr(\mathcal{E}(\mathcal{H}\backslash\{p\}, J)) \cdot \Pr(X_p \notin J)) \\
&= \sum_{J' \in \mathcal{P}(\mathcal{A},k)} \left(1 - \frac{k}{|\mathcal{A}|}\right) \cdot \Pr(\mathcal{E}(\mathcal{H}\backslash\{p\}, J)) \\
&= \left(1 - \frac{k}{|\mathcal{A}|}\right) \sum_{J' \in \mathcal{P}(\mathcal{A},k)} \Pr(\mathcal{E}(\mathcal{H}\backslash\{p\}, J)) \\
&= \left(1 - \frac{k}{|\mathcal{A}|}\right) \Pr(\mathcal{B}_k) \\
&= \left(1 - \frac{k}{|\mathcal{A}|}\right) \prod_{i=1}^{k-1} \left(1 - \frac{i}{|\mathcal{A}|}\right)
\end{aligned}$$

Question 9. En admettant que $e^x \geq 1 + x$ pour tout x, donner une borne supérieure de $\Pr(\mathcal{B}_k)$.

$$\begin{aligned}
\Pr(\mathcal{B}_k) &= \Pi_{i=1}^{k-1}(1 - \frac{i}{365}) \\
&\leq \Pi_{i=1}^{k-1} e^{-\frac{i}{365}}) \\
&\leq e^{-\frac{k(k-1)}{365}})
\end{aligned}$$

Question 10. Donner le plus grand nombre de personnes tels que la probabilité que k personnes ne soient pas nées le même jour soit inférieur à 1/2.

$$\begin{aligned}
\frac{1}{2} \leq & \qquad e^{-\frac{k(k-1)}{365})} \\
\ln(\frac{1}{2}) \leq & \qquad -\frac{k(k-1)}{2 \cdot 365}) \\
-\ln(2) \leq & \qquad -\frac{k(k-1)}{2 \cdot 365}) \\
\ln(2) \geq & \qquad \frac{k(k-1)}{2 \cdot 365}) \\
365 \ln(2) \geq & \qquad k(k-1) \leq k^2 \\
\sqrt{2 \cdot 365 \ln(2)} \geq & \qquad k \\
22.4943... \geq & \qquad k
\end{aligned}$$

Donc $k = 22$.