Quelques Exemples de calculs approchés de probabilités

1 Anniversaires

1. Quelle est la probabilité $p(r)$ que r personnes ont tous des dates d'anniversaires différentes? On considère qu'un année fait 365 jours.

2. Si r est petit, montrez que $p(r) \simeq 1 - \dfrac{1+2+\cdots+r-1}{365} = 1 - \dfrac{r(r-1)}{730}$. Vérifier cette approximation pour $r = 10$.

3. Pour r plus grand (mais pas trop) on obtient une meilleure approximation en passant au logarithme. Montrez que $\ln p(r) \simeq -\dfrac{1+2+\cdots+r-1}{365} = -\dfrac{r(r-1)}{730}$. Vérifier cette approximation pour $r = 30$.

4. A partir de combien de personnes est-il plus probables qu'au moins deux personnes aient la même date d'anniversaire?

2 Formule de Stirling: $n! \sim \left(\dfrac{n}{e}\right)^n \sqrt{2\pi n}$

On va utiliser des encadrements de $\int_1^n \ln x dx$ et les intégrales de Wallis pour démontrer la formule de Stirling. On représente dans le plan muni d'un repère orthonormé la courbe d'équation $y = \ln x$, les points A_k de coordonnées $(k, \ln k)$, $n > 1$, B_n de coordonnées $(n, 0)$, et $A_1 A_2 \cdots A_n B_n$ le polygone (convexe) P_n.

1. En comparant des aires par la méthode des rectangles, montrez que :
   $e(n/e)^n < n! < n^n$.

2. Montrez que $a_n := \ln(2) + \ln(3) + \cdots + \ln(n-1) + \frac{1}{2}\ln n$ représente l'aire de P_n et que $a_n < \int_1^n \ln x dx$.
   En déduire que $n! < e\sqrt{n}(n/e)^n$.

3. Montrez que $\int_{x-h/2}^{x+h/2} \ln x dx < h \ln x$ pour $0 < h < 2x$.
   En déduire que $\ln(n-1)! > \int_{3/2}^{n-1/2} \ln x dx$.

4. En déduire que $\int_{3/2}^n \ln x dx < a_n < \int_1^n \ln x dx$.

5. On pose $\delta_n = \ln n! - (n+1/2)\ln n + n$. Montrez que $0 < 1.5(1 - \ln 1.5) < \delta_n < 1$.

6. Expliquez pourquoi $1 - \delta_n$ représente l'aire du domaine compris entre P_n et la courbe $y = \ln x$. En déduire que que δ_n est strictement décroissante et converge vers un nombre strictement positif que l'on peut noter $\ln c$.

7. En déduire que $n! \sim c\sqrt{n}\frac{n^n}{e^n}$, $2.45 < c < 2.72$.

8. Estimation de c par les intégrales de Wallis. On pose $I_n = \int_0^{\pi/2}(\sin x)^n dx = \int_0^{\pi/2}(\cos x)^n dx$. Calculer I_0, I_1, trouver une relation de récurrence entre I_{n+2} et I_n. Calculer explicitement I_n et l'exprimer en fonction de $n!$ et $(2n)!$. Montrez que que I_n décroît et que $I_{n+1}/I_n \to 1$. Utiliser la formule de la question précédente pour évaluer I_{n+1}/I_n. Conclure.

9. Etudiez la validité de la formule de Stirling pour $n = 2, 3, 10, 100$.

3 Gaz dans une boîte

Une boîte de volume V comprend un zone Z de volume v. On note $p := \frac{v}{V} \in]0, 1[$. Dans la boîte se trouvent N molécules de gaz ($N \simeq 10^{23}$). On suppose que les molécules se distribuent au hasard dans la boîte. Au hasard signifiera ici que la probabilité de trouver une molécule donnée dans la zone Z est proportionelle au volume de Z. On cherche la proportion de gaz présent dans la zone Z.

1. Soit Ω l'ensemble des configurations possibles des N molécules de gaz. Chaque molécule est ou bien n'est pas dans Z. Montrez que le cardinal de Ω est 2^N.

2. On note Y_N la variable aléatoire représentant le nombre de molécules de gaz présentes dans la zone Z. Montrez que Y_N suit une loi binomiale.

3. On commence par le cas où Z représente la moitié de la boîte.

   (a) Montrez que si N est pair l'événemt $B_N = \{Y_N = N/2\}$ est le plus probable au sens où, si $k \neq N/2$ alors $P(B_N) > P(Y_N = k)$. Statistiquement, à quoi correspond l'événement B_N.

   (b) Montrez que si N est impair l'événement $B_N = \{Y_N = N/2 + 1/2\}$ (et l'événement symétrique $C_N = \{Y_N = N/2 - 1/2\}$) est le plus probable au sens où, si $k \neq 0$,$k \neq 1$ alors $P(B_N) = P(C_N) > P(Y_N = (N - 1/2 + k)$.

   (c) Cependant, à l'aide de la formule de Stirling, montrez que $P(B_N) \to 0$

   (d) Pensez vous qu'il y exactement autant de gaz dans Z que hors de Z si N est pair?

   (e) Cependant en utilisant le fait que $W_N := \frac{Y_N - Np}{\sqrt{Np(1-p)}}$ suit approximativement une loi normale pour N assez grand ($N > 30$), montrez que :
   $P\left(\left|\frac{Y_N}{N} - 1/2\right| < 10^{-10}\right) > 1 - 10^{-10}$.

4. Reprendre la question précédente dans le cas général: $0 < p < 1$, puis conclure.