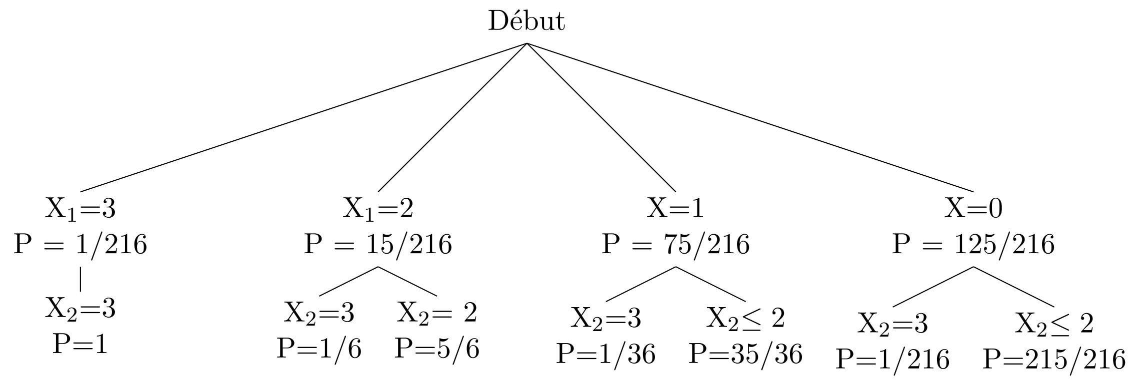
Correction de l'examen du mercredi 14 Mai 2014

Exercice 1 : Questions de cours.

1. Pour deux événements A et B d'un espace des issues Ω, la formule des probabilités totales dit que
$$P(A) = P(A|B)\,P(B) + P(A|B^c)\,P(B^c).$$
2. Une V.A. réelle Z qui suit la loi $\mathcal{N}(0,1)$ admet comme densité la fonction
$$\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}, \qquad \forall a < b, \qquad P\big(Z \in [a,b]\big) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{1}{2}x^2}\,dx.$$

Exercice 2 : Jeux de dés.

1. La meilleure façon de compter est de distinguer les trois dés. On peut alors introduire un ordre dans le résultat obtenu : (résultat du dé 1, résultat du dé 2, résultat du dé 3). Avec l'ordre, il y a $6 = 3!$ permutations possibles qui donnent le résultat 421 : $(1,2,4)$, $(1,4,2)$, $(2,1,4)$, $(2,4,1)$, $(4,1,2)$, $(4,2,1)$. Pour des lancers indépendants avec des dés non faussés, chacune d'entre elles a une probabilité $\left(\frac{1}{6}\right)^3$ de sortir. On a donc une probabilité $\frac{6}{216} = \frac{1}{36}$ d'obtenir le 421 du premier coup.
2. En comptant avec l'ordre, il n'y a que le résultat $(1,1,1)$ qui corresponde à trois 1. On a donc $\frac{1}{216}$ chances d'obtenir trois 1 du premier coup.
3. X suit la loi $\mathcal{B}(3, \frac{1}{6})$, loi binomiale de paramètre 3 et $\frac{1}{6}$. On a les formules :
   — $P(X=0) = \binom{3}{0} \times \left(\frac{1}{6}\right)^0 \times \left(\frac{1}{6}\right)^3 = \frac{125}{216}$,
   — $P(X=1) = \binom{3}{1} \times \left(\frac{1}{6}\right)^1 \times \left(\frac{1}{6}\right)^2 = \frac{75}{216}$,
   — $P(X=2) = \binom{3}{2} \times \left(\frac{1}{6}\right)^2 \times \left(\frac{1}{6}\right)^1 = \frac{15}{216}$,
   — $P(X=3) = \binom{3}{3} \times \left(\frac{1}{6}\right)^3 \times \left(\frac{1}{6}\right)^0 = \frac{1}{216}$
4. On note X_1 et X_2 le nombre de 1 obtenu après respectivement au premier et second coup. Cela donne le graphe suivant (dans lequel les probabilités indiquées sur la dernière ligne sont les probabilités conditionelles)



En utilisant la formule des probabilités totales de Bayes, on obtient :

$$\begin{aligned} P(X_2 = 3) &= P(X_2 = 3|X_1 = 3)\, P(X_1 = 3) + P(X_2 = 3|X_1 = 2)\, P(X_1 = 2) \\ &\qquad + P(X_2 = 3|X_1 = 1)\, P(X_1 = 1) + P(X_2 = 3|X_1 = 0)\, P(X_1 = 0) \\ P(X_2 = 3) &= 1 \times \frac{1}{216} + \frac{1}{6} \times \frac{15}{216} + \frac{1}{36} \times \frac{75}{216} + \frac{1}{216} \times \frac{125}{216} \\ P(X_2 = 3) &= \frac{216 + 36 \times 15 + 6 \times 75 + 125}{216^2} = \frac{1331}{46656} \\ P(X_2 = 3) &\approx 2,9\% \end{aligned}$$

5. Pour $i = 1, \dots 6$, il y a 1 chance sur 216 d'obtenir trois i du premier coup (voir question 2). Donc en additionnant les 6 événements, cela donne une probabilité $\frac{1}{36}$ d'obtenir trois fois la même valeur.
6. N suit la loi géométrique de paramètre $p = \frac{1}{36}$: $\mathcal{P}(\frac{1}{36})$.
7. Pour avoir $N > 10$, il faut rater les dix premières tentatives, donc

$$P(N > 10) = \Big(\frac{35}{36}\Big)^{10} \approx 0,755.$$

8. On calcule l'espérance du gain moins la mise, noté G ici. C'est une variable aléatoire.

$$\begin{aligned} \mathbb{E}(G) &= 3 \times P(N \geq 10) + (-1) \times P(N > 10) = 3\Big[1 - \Big(\frac{35}{36}\Big)^{10}\Big] - \Big(\frac{35}{36}\Big)^{10} \\ &= 3 - 4\Big(\frac{35}{36}\Big)^{10} \approx -0,018 \end{aligned}$$

Le jeu est donc très légérement défavorable. Si on joue 100 fois de suite, on perdra en moyenne $1,8$ euros. Pour une fois que le jeu proposé n'est pas clairement défavorable, on peut être tenté d'y jouer. Mais il a l'air aussi très monotone.

Exercice 3 : Publicité pour un pizzaiolo.

1. X suit la loi binomiale de paramètres $N = 5000$ et $p = 0,001$: $\mathcal{B}(5000; 0,001)$. Son éspérance et sa variance sont donnés par :

$$\mathbb{E}(X) = N\,p = 5000 \times 0,001 = 5, \qquad \mathbb{V}(X) = N\,p(1-p) = 5000 \times 0,001 \times 0,999 = 4,995.$$

2. On est dans le cas d'une loi binomiale avec N grand et p très petit (très infiérieur à 15%). On peut donc l'approcher par une loi de Poisson de paramètre $\lambda = N\,p = 5$: $\mathcal{P}(5)$.
L'espérance et la variance d'une loi de Poisson de paramètre λ sont égales à λ. Ici, on a donc

$$\mathbb{E}(X) = \mathbb{V}(X) = 5,$$

ce qui est très proche des vrai valeur (et exact pour l'espérance).

3. Sous l'approximation par la loi de Poisson $\mathcal{P}(5)$, on a les formules : $P(X = k) = e^{-5}\frac{5^k}{k!}$ pour tout k. Cela permet de calculer

$$\begin{aligned} P(X \leq 3) &= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) \\ &= e^{-5}\Big(1 + 5 + \frac{25}{2} + \frac{125}{6}\Big) = e^{-5}\frac{236}{6} = e^{-5}\frac{118}{3} \\ &\approx 0,27. \end{aligned}$$

Il a donc 27% de chances d'avoir moins de trois client en une soirée.

4. Y suit la loi binomiale de paramètres $N = 2000$ et $p = 0,2$: $\mathcal{B}(2000; 0,2)$. Son éspérance et sa variance sont donnés par :

$$\mathbb{E}(Y) = N\,p = 2000 \times 0,2 = 400, \qquad \mathbb{V}(Y) = N\,p(1-p) = 2000 \times 0,2 \times 0,8 = 320.$$

5. On est dans le cas d'une loi binomiale avec N grand et p supérieur à 15%. On peut donc l'approcher par une loi normale de paramètre $\mu = N\,p = 400$ et $\sigma^2 = 320$: $\mathcal{N}(400, 320)$. On a alors $\mathbb{E}(Y) = 400$ et $\mathbb{V}(Y) = 320$, comme pour la loi exacte.
6. On calcule une valeur approchée de l'écart-type : $\sigma = \sqrt{320} \approx 17,9$. On introduit la variable pivôt $Z = \frac{Y-400}{17,9}$ qui suit elle approxiamtivement une loi $\mathcal{N}(0,1)$.
7. On remarque que $Y \geq 440$ correspond à $Z \geq \frac{440-400}{17,9} = 2,24$. D'après la table de la loi $\mathcal{N}(0,1)$, on a

$$P(Y > 440) = P(Z > 2,24) = 1 - P(Z \leq 2,24) = 1 - 0,9875 = 1,25\%.$$

   Donc le pizzaiolo ne va pas pouvoir assurer toutes les commandes dans au moins 99% des semaines. Il faudrait qu'il commande plus de patôns.
8. On cherche le fractile à 99% de la loi $\mathcal{N}(0,1)$. D'après la table fournie, on trouve une valeur de $2,33$, et donc $P(Z \leq 2,33) = 99\%$. Retranscrit en terme de Y, on obtient avec $400 + 2,33 \times 17,9 \approx 441,7$

$$P(Y \leq 441,7) = P(Z \leq 2,33) = 1 - P(Z \leq 2,24) = 99\%$$

   Donc il faudrait commander 442 pâtons chaque semaine pour pouvoir assurer toutes les commandes sur au moins 99% des semaines.

Exercice 4 : Tests sur une source d'eau.
Suite à quelques intoxications, le propriétaire d'une source d'eau décide de faire des analyses à différents moment de l'année choisis au hasard. Sur 49 mesures éffectuées au total, 13 indiquent une eau non potable (c'est-à-dire avec les différentes pollutions possibles en dessous des seuils admissibles), le reste indiquant une eau potable. N'ayant jamais suivi de cours de statistique, il en conclut que l'eau est potable $26,5\%$ du temps.

1. Il ne devra pas donner de valeur exacte car il ne mesure pas la potabilité de sa source à chaque instant. Il n'a qu'un échantillon de mesures, et l'on sait que les fréquences empiriques obtenues sur des échantillons, fluctuent autour de la "vraie" moyenne. I ne peut donner qu'un intervalle de confiance, d'autant plus petit que l'échantillon est grand.
2. Si on note f_N la fréquence empirique obtenue, l'intervalle à 95% autour de celle-ci est donnée par

$$\left[f_N - 1,96\frac{\sqrt{f_N(1-f_N)}}{\sqrt{n}}; f_N + 1,96\frac{\sqrt{f_N(1-f_N)}}{\sqrt{n}}\right]$$

   Ici, on calcule $1,96\frac{\sqrt{0,0475}}{\sqrt{49}} = 1,96\frac{0,22}{7} = 0,061 = 6,1\%$. Et on peut donc dire qu'au risque de 5%, la proportion des moments pendant laquelle la source n'est pas potable est comprise dans l'intervalle

$$[20,4\%, 32,1\%].$$

3. On commence par ajouter au tableau une ligne et colonne totale et on calcule les efectifs attendus si la présence des cigales et la contamination de la source était indépendantes.

Eau \ Cigales	Présentes	Absentes	Total
Non potable	18 / 14	12 / 16	30
Potable	5 / 9	14 / 10	19
Total	23	26	49

On calcule ensuite la valeur de la variable pivôt associée :

$$
\begin{aligned}
z &= \frac{(18-14)^2}{14} + \frac{(12-16)^2}{16} + \frac{(5-9)^2}{9} + \frac{(14-10)^2}{10} \\
&= \frac{16}{13} + \frac{16}{15} + \frac{16}{10} + \frac{16}{11} = 1,23 + 1,06 + 1,60 + 1,45 \\
&= 5,35.
\end{aligned}
$$

Si on répétait l'expérience, la variable Z devrait suivre la loi du χ^2 à $(2-1)(2-1) = 1$ degré de liberté. D'après la table du χ^2 fourni, le fractile à 95% d'une χ^2 à 1 d.d.l. vaut $3,84$. Comme $5,35 > 3,84$, s'il y a indépendance, la valeur trouvée fait partie des 95% de valeurs extrèmes. Comme c'est peut probable, on peut rejeter l'hypothèse d'indépendance. La présence des cigales et la non-potabilité de la source sont corrélées.

4. Non, l'absence d'indépdance entre présence des cigales et potabilité de la source indique une corrélation entre celles-ci, mais pas forcémment une relation de cause à effet. En fait, il semble peu probable que les cigales, qui vivent principalement dans les arbres, influent sur l'eau de la source, qui vient du sous-sol. Par contre, on peut remarquer que les cigales sont présentes en été seulement, et que la corrélation observée veut plutôt dire que la source est plus souvent pollué en été (par exemple par la présence de troupeaux au dessus de celles-ci,...).