Université de PARIS SUD
NFI Optronique $4^{\text{ième}}$ année

Probabilités et statistique

Corrigé succinct de l'examen de janvier 2005

Exercice I

Parmi une population de n lapins, le nombre d'albinos suit une loi binomiale $\mathcal{B}(n,p)$ avec $p = 1/100$. Lorsque $n = 100$, cette loi est très proche de celle de Poisson de paramètre $\lambda = np = 1$. D'où une probabilité $1/e \simeq 0.368$ d'avoir exactement un albinos, et $1-1/e \simeq 0.632$ pour un ou plus.

Exercice II

Le théorème de Bayes permet d'écrire, en notant $A \equiv$ paquet accepté

$$\begin{aligned} P(A) &= P(A|4 \text{ défauts})\, 0.3 + P(A|1 \text{ défaut})\, 0.7 \\ &= \frac{\mathrm{C}_6^3}{\mathrm{C}_{10}^3}\, 0.3 + \frac{\mathrm{C}_9^3}{\mathrm{C}_{10}^3}\, 0.7 \\ &= 0.54 \end{aligned}$$

Exercice III

On note $M \equiv$ "malade" et $P \equiv$ "test positif". On a $Pr(P|M) = 0.95$ et $Pr\left(\overline{P}|\overline{M}\right) = 0.95$. Par ailleurs, $Pr(M) = 0.01$ et l'on cherche $Pr\,(M|P)$. D'après le théorème de Bayes,

$$Pr(M|P) = \frac{Pr(P|M)Pr(M)}{Pr(P|M)Pr(M) + Pr\left(P|\overline{M}\right)Pr\left(\overline{M}\right)}.$$

Enfin, $Pr(\overline{M}) = 0.99$ et $\left(P|\overline{M}\right) = 0.05$, d'où $Pr\,(M|P) \simeq 0.16$. En dépit du résultat positif, la probabilité d'être malade reste relativement faible, essentiellement parce que la maladie est peu répandue.

Exercice IV

Traduisons les données de l'énoncé : la densité conditionnelle de ℓ s'écrit

$$g(\ell|T=t) = \alpha \exp(-\ell/\bar{\ell}),$$

où la condition de normalisation impose $\alpha = 1/\bar{\ell}$. De plus, T suit une loi $\mathcal{G}(\varepsilon,\sigma)$. Les théorèmes de l'espérance et de la variance totale permettent d'écrire $E[E(L|T)] = E(L)$ et $E[V(L|T)] + V[E(L|T)] = V(L)$. Compte tenu de

$$E(L|T=t) = \bar{\ell} = \beta t \qquad \text{et} \qquad V(L|T=t) = \beta^2 t^2,$$

il vient $E[V(L|T)] = \beta^2 E(T^2) = \beta^2(\sigma^2+\varepsilon^2)$ et $V[E(L|T)] = \beta^2\sigma^2$, d'où

$$\boxed{E(L) = \beta\varepsilon \qquad \text{et} \qquad V(L) = \beta^2(\varepsilon^2 + 2\sigma^2)}.$$

Exercice V

Lorsque $X = n$, Y suit une loi $\mathcal{B}(n,p)$. On en déduit

$$P(X\!=\!n, Y\!=\!k) \;=\; P(Y\!=\!k|X\!=\!n)P(X = n) \;=\; C_n^k\, p^k\, q^{n-k}\, e^{-\lambda}\, \frac{\lambda^n}{n!} \qquad \text{où} \qquad p + q = 1.$$

Pour obtenir la loi marginale de Y, on remarque que l'expression précédente est valable pour $0 \leqslant k \leqslant n$ et $n \in \mathbb{N}$.

$$\begin{aligned} P(Y = k) &= \sum_{n=k}^{\infty} P(X\!=\!n, Y\!=\!k) \;=\; \frac{1}{k!}\, p^k\, e^{-\lambda} \sum_{n=k}^{\infty} \frac{n!}{(n-k)!}\, q^{n-k}\, \frac{\lambda^n}{n!} \;=\; \frac{p^k\, e^{-\lambda}}{k!}\, \lambda^k \sum_{l=0}^{\infty} \frac{(q\lambda)^l}{l!} \\ &= \frac{(\lambda p)^k\, e^{-\lambda}}{k!}\, \exp(q\lambda) \;=\; \frac{1}{k!}\, (\lambda p)^k\, e^{-\lambda p}. \end{aligned}$$

Y suit donc une loi de Poisson de paramètre λp. On retrouve par le calcul un résultat attendu : Y suit un processus de Poisson de valeur moyenne $p\lambda$.

La variable Z représente le nombre de particules non détectées. Il s'agit donc d'une variable aléatoire de même loi que Y, à condition d'interchanger p et $1-p$: loi de Poisson de paramètre $(1-p)\lambda$. Pour calculer la loi jointe de (Z, Y), on remarque que

$$Z = z \quad \text{et} \quad Y = k \qquad \Leftrightarrow \qquad X = z + k \quad \text{et} \quad Y = k,$$

d'où

$$P(Z\!=\!z, Y\!=\!k) \;=\; P(X\!=\!z+k, Y\!=\!k) \;=\; C_{z+k}^k\, p^k\, q^{z+k-k}\, e^{-\lambda}\, \frac{\lambda^{z+k}}{(z+k)!} \;=\; e^{-\lambda}\, \frac{(\lambda q)^z}{z!}\, \frac{(\lambda p)^k}{k!}.$$

Il apparaît que $P(Z = z, Y = k) = P(Z = z)P(Y = k)$: les variables Y et Z sont donc indépendantes. Pour terminer, si l'on triplait le temps d'observation ($T \to 3T$), le nombre de particules observées suivrait une loi de Poisson $\mathcal{P}(3\lambda)$.

Exercice VI : corrélation et causalité

1. La grande majorité des déplacements ayant lieu dans le voisinage (30 km) du domicile, le "raisonnement" tenu est une plaisanterie. On note $A \equiv$ "accident" et $V \equiv$ "être dans le voisinage du domicile". L'énoncé indique que $P(V|A) = 0.8$, mais le propos est de comparer $P(A|V)$ à $P(A|\overline{V})$, ou, ce qui revient au même (à justifier), $P(A|V)$ à $P(A)$. Le théorème de Bayes indique que

$$\frac{P(A|V)}{P(A)} = \frac{P(V|A)}{P(V)}.$$

Ainsi, la question est de savoir si la probabilité d'effectuer un trajet dans le voisinage de son domicile est supérieure ou non à 0.8. Si $P(V) = 0.75$, on peut effectivement conclure à un assoupissement de l'attention pour les trajets de proximité, mais si $P(V) = 0.85$, la conclusion est inversée...

2. Cet énoncé montre le résultat profond que les personnes de 60 ans en suffisamment bonne santé pour pouvoir faire du jogging ont une espérance de vie supérieure à celle de la population du même âge. Si le jogging est bénéfique pour la santé, ce n'est pas avec ce type de données qu'on pourra le montrer ; le vice de fond est de même nature que celui discuté dans la question 1.

Les deux exemples proposés ici sont typiques des raccourcis erronés qui caractérisent parfois le maniement des probabilités. Ils relèvent d'une "logique" que Coluche a poussé dans ses retranchements, en proposant de ne jamais aller à l'hôpital en cas de maladie, puisque les statistiques démontrent sans contestation possible qu'il est plus probable de décéder dans un lit d'hôpital que chez soi !