Contrôle des connaissances du cours
« Bruits et Signaux »
de l'Ecole Doctorale d'Astrophysique.
$\boxed{\textbf{Corrigé}}$

Le mercredi 16 décembre 1998 de 14h à 16h amphi de l'IAP.

Le contrôle est noté sur 20, le barème est indiqué à la fin de chaque groupe de questions. Les exercices sont indépendants.

Le χ^2 réduit.

On a réalisé l'ajustement de données expérimentales y_i, ($i = 1, \ldots, n$) à l'aide d'un modèle μ_i dépendant de k paramètres inconnus. Suite à cette estimation on a construit la variable aléatoire suivante :

$$X = \sum_{i=1}^{n} \frac{(y_i - \mu_i)^2}{\sigma_i^2} \,. \tag{1}$$

① Les données y_i sont entachées d'un bruit, mais on suppose que le choix des paramètres est tel que l'espérance mathématique des y_i est égale à μ_i ($\mathrm{E}\{y_i\} = \mu_i$). Les μ_i sont donc les moyennes des y_i, mais d'après vous que représentent les σ_i ?

② L'expérimentateur ne connait pas les σ_i mais il pose : $\sigma_i^2 = \mu_i$, pouvez-vous donner les considérations théoriques qui peuvent justifier son choix ?

③ L'expérimentateur a de bonnes raisons de penser que la variable X suit une loi du χ^2 à $f = n - k$ degrés de liberté. Comme c'est souvent l'usage, il considère plutôt le χ^2 *réduit* : X/f. Quels sont la moyenne et l'écart type de ce χ^2 réduit ? (On rapelle que la moyenne d'une variable aléatoire suivant la loi du χ^2 à f degrés de liberté est f et que sa variance est $2f$.)

④ Le nombre de points est de $n = 1024$ et le nombre de paramètres est de $k = 10$, le χ^2 réduit vaut 1.1, l'ajustement vous semble-t-il satisfaisant ?

($1 + 1 + 2 + 2$ points.)

Corrigé.

❶ Les paramètres σ_i sont les écarts types des données y_i, en gros ce sont les « barres d'erreurs ».

❷ σ_i^2 est la variance des données, on peut dire que la variance est égale à la moyenne si les y_i sont des variables aléatoires de Poisson. L'observateur a vraisemblablement compté des photons.

❸ On a pour la moyenne $\mathrm{E}\{\chi^2/f\} = \mathrm{E}\{\chi^2\}/f = f/f = 1$ et pour la variance $\mathrm{Var}(\chi^2/f) = \mathrm{Var}(\chi^2)/f^2 = 2f/f^2 = 2/f$ (voir poly § 5.6.3 par exemple).

❹ On a $f = 1024 - 10 = 1014$, on peut alors confondre le loi du χ^2 réduit ou non avec une loi normale. Pour le montrer on peut utiliser le théorème central limite en se souvenant que la loi du χ^2 à f degrés de liberté est la somme de f variables aléatoires normales réduites indépendantes. L'écart type de cette loi est $\sqrt{2/1014} \approx 0.0444$ sa moyenne est 1. On est à $|1.1 - 1|/0.0444 \approx 2.25$ « sigmas » de la moyenne : l'ajustement est acceptable suivant le critère classique des trois-sigmas.

Loi suivie par un quotient.

On sait que le quotient de deux variables aléatoires normales indépendantes de moyenne nulle et d'écart type un, suit une loi de Cauchy :

$$f(x) = \frac{1}{\pi}\frac{1}{1+x^2} \,.$$

Le but de cet exercice est de montrer que ce résultat n'est pas une propriété caractéristique de la loi normale. Considérons en effet les variables indépendantes X_1 et X_2 qui suivent la même loi dont la densité de probabilité $g(x)$ est donnée par l'expression suivante :

$$g(x) = \frac{\sqrt{2}}{\pi}\frac{1}{1+x^4}\,, \quad (-\infty < x < \infty)\,.$$

① Donner l'expression de la densité de probabilité du rapport $Y = X_1/X_2$ sous la forme d'une intégrale.

② Pour évaluer cette intégrale, le premier reflexe du chercheur est de se précipiter sur "Table of Integrals, Series and Products" de Gradshteyn et Ryzhik. On y trouve la formule 3.264.2 qui nous apprend que :

$$\int_0^\infty \frac{x^{\mu-1}\,dx}{(\beta+x^2)(\gamma+x^2)} = \frac{\pi}{2}\,\frac{\gamma^{\frac{\mu}{2}-1} - \beta^{\frac{\mu}{2}-1}}{\beta-\gamma}\,\frac{1}{\sin\frac{\pi}{2}\mu}\,,$$

pour $\beta, \gamma > 0$ et $0 < \mu < 4$. A l'aide de cette formule montrez que Y suit aussi une loi de Cauchy.

($1 + 2$ points.)

Corrigé.

❶ On utilise la formule du § 4.3.3 du cours qui donne la densité de probabilité d'un quotient de variables aléatoires. Soit h cette densité, il vient :

$$\begin{aligned} h(x) &= \frac{2}{\pi^2}\int_{-\infty}^{+\infty} \frac{|y|\,dy}{(1+y^4)(1+x^4y^4)}\,, \\ &= \frac{2}{\pi^2}\int_0^\infty \frac{2y\,dy}{(1+y^4)(1+x^4y^4)}\,, \\ &= \frac{2}{\pi^2}\frac{1}{x^4}\int_0^\infty \frac{du}{(1+u^2)(x^{-4}+u^2)}\,. \end{aligned}$$

❷ Il suffit d'utiliser la formule avec $\mu = 1$, $\beta = 1$ et $\gamma = x^{-4}$ pour obtenir le résultat.

« Sphère céleste. »

On désire trouver un moyen de répartir des objets sur la sphère céleste en vue de comparer une distribution simulée par rapport à une distribution observée. A titre d'exemple on peut songer à la répartition des "bursters γ" dans l'univers.

On veut que la répartition simulée soit *uniforme* sur la sphère céleste. Par uniforme sur la sphère céleste, on entend que la probabilité de trouver un objet dans l'angle solide Ω soit égale à $\Omega/4\pi$.

Il existe de nombreuses façons d'atteindre ce but, on trouvera ci-dessous un moyen qui présente de très agréables propriétés de symétries. C'est la méthode que je vous recommande si d'aventure vous vous trouviez face à ce problème.

① On considère un point M de $\mathbb{R}^3$, on le repère grâce à ses coordonnées cartésiennes (X, Y, Z) ou ses coordonnées sphériques (R, Θ, Φ).

Soit f_C la densité de probabilité du vecteur OM de coordonnées (X, Y, Z), quelle est la densité de probabilité f_S de ce même vecteur mais exprimée en coordonnées sphériques ?

② On « tire » au hasard les 3 nombres X, Y et Z suivant la loi normale réduite, ces 3 variables aléatoires sont indépendantes.

Quelle est la densité de probabilité f_C du triplet (X, Y, Z) ?

Quelle est la densité de probabilité f_S du triplet (R, Θ, Φ) ?

③ Montrer que la densité de probabilité de la loi marginale $f(\theta, \phi)$ est bien uniforme au sens défini plus haut.

($1 + 2 + 2$ points.)

Corrigé.

❶ On peut utiliser la formule (4.20) p. 52, mais on peut aussi raisonner comme suit. Le passage de (X, Y, Z) à (R, Θ, Φ) est une bijection, la conservation de la probabilité s'exprime alors ainsi $f_C(x, y, z)dxdydz = f_S(r, \theta, \phi)drd\theta d\phi$. (Notez bien qu'en procédant ainsi il peut y avoir un problème de signe qui est levé en remarquant qu'une probabilité est nécessairement positive.) Il vient $f_S(r, \theta, \phi) = f_C(x, y, z)|J|^{-1}$, où J est le jacobien du changement de variable. Mais $|J|^{-1} = |J^{-1}|$ qui est l'élément de volume en sphérique : $r^2 dr \sin\theta d\theta d\phi = r^2 dr d\Omega$. Finalement :

$$f_S(r, \theta, \phi) = f_C(r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta)r^2\sin\theta\,.$$

❷ Le résultat est immédiat :

$$f_C(x, y, z) = \frac{1}{(2\pi)^{3/2}} e^{-\frac{1}{2}(x^2+y^2+z^2)}\,,$$
$$f_S(r, \theta, \phi) = \frac{1}{(2\pi)^{3/2}} e^{-\frac{1}{2}r^2} r^2 \sin\theta\,.$$

❸ La densité marginale de (Θ, Φ) pour R quelconque est par définition :

$$\begin{aligned} f(\theta, \phi) &= \int_0^\infty f_S(r, \theta, \phi)\, dr\,, \\ &= \frac{\sin\theta}{(2\pi)^{3/2}} \int_0^\infty r^2 e^{-\frac{1}{2}r^2}\, dr\,. \end{aligned}$$

Il est inutile de calculer l'intégrale (qui d'ailleurs s'intègre facilement par parties) car on sait que f est normalisée à un. On a nécessairement $\int f d\theta d\phi = 1$, d'où :

$$f(\theta, \phi) = \frac{\sin\theta}{4\pi}\,.$$

La probabilité P que M « tombe » dans un angle solide Ω quelconque est :

$$\begin{aligned} P &= \int_\Omega f(\theta,\phi)\,d\theta d\phi\,,\\ &= \int_\Omega \frac{\sin\theta}{4\pi}\,d\theta d\phi\,,\\ &= \frac{1}{4\pi}\int_\Omega d\Omega\,,\\ &= \frac{\Omega}{4\pi}\,. \end{aligned}$$

Détection d'ondes gravitationnelles.

Deux expériences indépendantes (Virgo en Italie et Ligo aux USA) tentent de détecter des ondes gravitationnelles. Pour cela elles ont mis en place des « triggers ». Un trigger est un logiciel qui examine les données et qui déclenche une procédure d'alarme lorsque quelque chose lui semble intéressant ou anormal. Le taux d'alarmes (vraies ou fausses) est de λ_1 par unité de temps pour Virgo et de λ_2 pour Ligo. On suppose que les flux d'alarmes sont, pour les deux expériences, des flux de Poisson. Le but de cet exercice est de voir comment on peut tirer parti des deux expériences afin de réduire le taux de fausses alarmes.

① Rappeler les conditions pour qu'un flux d'événements soit un flux de Poisson.

② Parmi les alarmes, certaines correspondent à de vraies ondes gravitationnelles et d'autres non. La probabilité pour qu'une alarme soit réelle est de p_1 pour Virgo et de p_2 pour Ligo. On imagine qu'à chaque alarme le « choix » entre vraie et fausse alarme a lieu de indépendemment des alarmes précédentes.

Quels sont les taux d'alarmes réelles pour les deux expériences ? Les flux correspondants sont-ils de Poisson ? (Surtout ne faire aucun calcul mais justifiez votre réponse.)

③ On décide de concentrer l'attention sur les alarmes qui semblent coïncider. On dit que deux alarmes coïncident si l'une issue de Virgo et l'autre issue de Ligo sont séparées de moins d'un temps Δt.

Pouvez-vous donner une idée de ce temps Δt ?

④ Plaçons-nous du point de vue de l'expérience Virgo par exemple. Une alarme est détectée, on regarde alors les données de Ligo (le travail est fait en temps différé et non en temps réel). Sous hypothèse nulle (pas de signal) quand peut-on dire qu'une alarme de Virgo coïncide avec une alarme de Ligo ? On demande une phrase pas une formule.

⑤ A cause de l'effet rétroactif, il n'est pas facile de donner la probabilité de l'événement précédent c'est pourquoi nous allons encore simplifier le problème en supposant que l'on cherche à detecter des ondes allant de Virgo à Ligo. Toujours sous hypothèse nulle, quelle est alors la probabilité pour qu'une alarme de Virgo coïncide avec une alarme de Ligo ?

($1+1+1+1+2$ points.)

Corrigé.

❶ Ordinaire, stationnaire et sans mémoire (voir § 9.1).

❷ Il s'agit d'un flux tamisé aléatoirement (voir poly § 9.4.2) le taux d'alarmes réelles est de $\lambda_1 p_1$ pour Virgo et de $\lambda_2 p_2$ pour Ligo.

On peut dire les choses ainsi, le tamisage laisse le flux ordinaire, stationnaire et sans-mémoire. Quand Δt est petit on a 0 ou 1 événement, en moyenne on en a λp_1 pour Virgo par exemple. Ce sont donc des flux de Poisson de paramètres λp_1 et λp_2.

❸ Le temps Δt doit être proche du temps mis par une onde gravitationnelle pour aller d'Italie aux USA. Pour fixer les idées on peut prendre le diamètre de la terre divisé par la vitesse de la lumière soit $\approx 0.04\,\mathrm{sec}$.

❹ Il faut qu'il y ait *au moins* une alarme Ligo dans l'intervalle de temps $[t-\Delta t, t+\Delta]$ autour du temps t où Virgo a detecté une alarme.

❺ Il y a coïncidence si au moins une alarme de Ligo est détectée dans un temps Δt *après* une alarme de Virgo. Sachant que les alarmes de Ligo arrivent au taux λ_2, la probabilité de l'événement précedent est $1-\mathrm{Pr}\{$qu'il n'y ait pas de d'alarmes$\} = 1 - e^{-\lambda_2 \Delta t}$.