

D. HEMON

**Estimation et dispositifs expérimentaux dans le
cas d'observations tronquées en nombre aléatoire.
Cas Poisson-normal**

Revue de statistique appliquée, tome 25, n° 3 (1977), p. 43-54

http://www.numdam.org/item?id=RSA_1977__25_3_43_0

© Société française de statistique, 1977, tous droits réservés.

L'accès aux archives de la revue « *Revue de statistique appliquée* » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

ESTIMATION ET DISPOSITIFS EXPÉRIMENTAUX DANS LE CAS D'OBSERVATIONS TRONQUÉES EN NOMBRE ALÉATOIRE. CAS POISSON-NORMAL

D. HEMON

RESUME

Soient (X_1, \dots, X_N) N réalisations indépendantes de la variable normale X . On suppose que le nombre d'observations N suit une loi de Poisson. Si X est tronquée à droite d'une valeur x_0 on ne peut observer ni la valeur de N ni celle des X qui sont supérieurs à x_0 . Ce travail décrit l'estimation du maximum de vraisemblance des paramètres de distribution de X et de N et propose un schéma expérimental pour optimiser cette estimation.

I – LE MODELE INITIATION-PROMOTION

De très nombreuses expériences de carcinogénèse chimique sont organisées autour de l'hypothèse "initiation-promotion" émise par Berenblum [1, 2]. Dans ces expériences les animaux subissent successivement :

- un traitement "initiateur" qui induit la transformation de certaines cellules de l'organe traité en cellules latentes cancéreuses : les cellules "initiées" ;
- un traitement "promoteur" qui permet aux cellules initiées de se développer en tumeurs observables.

Des observations recueillies sur chacun des animaux : nombre de tumeurs et délais d'apparition de celles-ci, l'expérimentateur cherche à tirer des informations concernant d'une part le "pouvoir initiateur" du premier traitement, d'autre part le "pouvoir promoteur" du second traitement.

On peut préciser le sens de ces termes en explicitant le modèle probabiliste qui résulte de l'hypothèse de Berenblum :

- le nombre de cellules initiées portées par animal est aléatoire, la moyenne λ de sa distribution caractérise le pouvoir initiateur du premier traitement ;
- les différentes cellules initiées évoluent indépendamment les unes des autres ;
- le temps de développement d'une cellule initiée (non observable) en tumeur (observable) est aléatoire, la moyenne μ de sa distribution caractérise le pouvoir promoteur du second traitement.

Ce modèle est couramment utilisé en carcinogénèse chimique expérimentale. Il est également utilisé en bactériologie où on est amené à s'intéresser à la fois au nombre de microorganismes clonogènes ensemencés dans un milieu de culture et

au délai de développement des ceux-ci en colonies observables. Plus généralement, ce modèle correspond à toute situation où le nombre d'observations concernant une variable X est aléatoire et où on s'intéresse aussi bien à la distribution de X qu'à celle du nombre d'observations. Pour plus de commodité nous garderons cependant dans la suite du texte le vocabulaire des expériences de carcinogénèse chimique : animaux, cellules initiées, tumeurs.

L'estimation des paramètres λ et μ du modèle ne poserait pas de problèmes particuliers si la durée de l'expérience était illimitée. On disposerait, en effet, de deux informations indépendantes :

- d'une part le nombre de tumeurs portées par un animal correspondant exactement au nombre de cellules initiées de cet animal et permettant d'estimer λ ;
- d'autre part les délais de promotion de toutes les cellules initiées permettant d'estimer μ .

En fait, la durée de l'expérience est généralement limitée et les observations sont tronquées à droite d'un point t_0 (au sens défini par exemple par Cohen dans [4]) :

– on ne connaît pas le nombre de cellules initiées portées par chaque animal mais seulement le nombre de celles qui se sont développées en tumeur avant la fin de l'expérience ;

– on n'observe que les délais de promotion inférieurs à la durée de l'expérience. C'est dans ces conditions particulières que nous nous proposons d'étudier l'estimation des paramètres λ et μ . Nous envisagerons ici le cas "Poisson-normal" :

– le nombre de cellules initiées est distribué selon une loi de Poisson de moyenne λ ;

– le délai de promotion d'une cellule initiée en tumeur observable est distribué selon une loi normale de moyenne μ et d'écart-type σ . Dans le cas de la carcinogénèse chimique expérimentale, ces hypothèses semblent vérifiées par les observations si les animaux utilisés sont suffisamment homogènes du point de vue génétique et si on utilise le logarithme des délais d'apparition des tumeurs.

Le paragraphe II décrit l'estimation du maximum de vraisemblance (en abrégé M.V.) des paramètres (λ, μ, σ) . Le paragraphe III concerne un dispositif expérimental susceptible d'améliorer la précision des estimations lorsque la troncature des observations est importante.

II - ESTIMATION DU MAXIMUM DE VRAISEMBLANCE DES PARAMETRES (λ, μ, σ)

II.A – Vraisemblance des observations

Notons N le nombre d'animaux, n_i le nombre de tumeurs portées par le i ème animal en fin d'expérience et $(t_{ij}/j = 1, n_i)$ les délais d'apparition de chacune de ces tumeurs.

La vraisemblance des observations s'écrit :

$$L(\lambda, \mu, \sigma) = \prod_{i=1}^N \left\{ \exp \left[-\lambda F \left(\frac{t_0 - \mu}{\sigma} \right) \right] \frac{\left[\lambda F \left(\frac{t_0 - \mu}{\sigma} \right) \right]^{n_i}}{n_i!} \prod_{j=1}^{n_i} \frac{1}{\sigma} f \left(\frac{t_{ij} - \mu}{\sigma} \right) \right\}$$

où f et F sont respectivement la fonction de densité et la fonction cumulative de la distribution normale centrée réduite.

En effet :

- n_i est distribué suivant la loi de Poisson de moyenne $\lambda F \left(\frac{t_0 - \mu}{\sigma} \right)$;
- conditionnellement à n_i , $(t_{ij}/j = 1, n_i)$ sont n_i observations normales de paramètres μ et σ tronquées à droite du point t_0 .

II.B – Détermination des estimations du maximum de vraisemblance

L'estimation du maximum de vraisemblance de (λ, μ, σ) est le point $(\hat{\lambda}, \hat{\mu}, \hat{\sigma})$ qui maximise la quantité :

$$\begin{aligned} \text{Log } L = & -N\lambda F \left(\frac{t_0 - \mu}{\sigma} \right) + \left(\sum_{i=1}^N n_i \right) \text{Log}(\lambda) + \\ & + \sum_{i=1}^N \sum_{j=1}^{n_i} \text{Log} \left[\frac{1}{\sigma} f \left(\frac{t_{ij} - \mu}{\sigma} \right) \right] + \text{Cste} \quad (1) \end{aligned}$$

En annulant la dérivée de (1) par rapport à λ on obtient l'expression de $\hat{\lambda}$ en fonction de $\hat{\mu}$ et $\hat{\sigma}$:

$$\lambda = \frac{\bar{n}}{F \left(\frac{t_0 - \hat{\mu}}{\hat{\sigma}} \right)} \quad \text{où} \quad \bar{n} = \frac{\sum_{i=1}^N n_i}{N}.$$

En substituant cette expression dans (1), on obtient

$$\text{Log } L = \sum_{i=1}^N \sum_{j=1}^{n_i} \text{Log} \left\{ \frac{1}{\sigma} \frac{f \left(\frac{t_{ij} - \mu}{\sigma} \right)}{F \left(\frac{t_0 - \mu}{\sigma} \right)} \right\} + \text{Cste}$$

Cette dernière expression est le Log vraisemblance des valeurs $(t_{ij}/j = 1, n_i/i = 1, N)$ considérées comme un échantillon de la distribution normale de paramètres μ et σ tronquée à droite de t_0 .

On peut déterminer μ et σ en mettant en œuvre les méthodes itératives classiques décrites par exemple dans Box, Davies et Swann [3] en utilisant les expressions données en annexe concernant les dérivées d'ordre 1 et 2 du Log-vraisemblance. Alternativement on peut utiliser les tables et/ou les méthodes graphiques décrites par divers auteurs pour estimer les paramètres d'une distribution normale tronquée, par exemple par Cohen [4, 5].

II.C – Variances et covariances asymptotiques des estimateurs du maximum de vraisemblance

La matrice des variances et covariances asymptotiques de l'estimateur du M.V. est donnée par :

$$V = \frac{M^{-1}}{N}$$

où M est la matrice des informations de Fisher apportées par une unité expérimentale sur (λ, μ, σ) .

Le calcul de M est détaillé en annexe. Il conduit à la structure suivante de M^{-1} :

$$M^{-1} = \begin{pmatrix} \lambda v_{\lambda\lambda} & \sigma v_{\lambda\mu} & \sigma v_{\lambda\sigma} \\ \sigma v_{\lambda\mu} & \frac{\sigma^2}{\lambda} v_{\mu\mu} & \frac{\sigma^2}{\lambda} v_{\mu\sigma} \\ \sigma v_{\lambda\sigma} & \frac{\sigma^2}{\lambda} v_{\mu\sigma} & \frac{\sigma^2}{\lambda} v_{\sigma\sigma} \end{pmatrix} \quad (2)$$

les coefficients v ne dépendant que de $z_0 = \frac{t_0 - \mu}{\sigma}$ ou, de façon équivalente, du taux de troncature des observations $\pi = 1 - F(z_0)$.

La table 1 donne les valeurs numériques des coefficients $v_{\lambda\lambda}$, $v_{\mu\mu}$, $v_{\sigma\sigma}$ et des corrélations des estimateurs du maximum de vraisemblance (λ, μ, σ) pour différentes valeurs de π . Ces valeurs ont été calculées à partir de l'expression analytique des coefficients de M donnés en annexe. On constate que la variance des estimateurs augmente très rapidement avec le taux de troncature des observations. Si celui-ci est important on est donc dans de mauvaises conditions d'estimation. On constate également que la corrélation des estimateurs augmente rapidement avec le taux de troncature des observations. Ce fait conduit à penser que l'amélioration de l'estimation de l'un des paramètres du modèle est susceptible d'améliorer également l'estimation des autres paramètres. A partir de cette idée, nous proposons au paragraphe suivant un dispositif expérimental permettant d'améliorer la qualité des estimations du M.V. lorsque le taux de troncature des observations est important.

TABLE 1

Coefficients de la matrice des variances et covariances asymptotique de l'estimateur du M.V. ($\hat{\lambda}$, $\hat{\mu}$, $\hat{\sigma}$) en fonction de la troncature des

$$\text{observations } \pi = 1 - F\left(\frac{t_0 - \mu}{\sigma}\right).$$

π	$v_{\lambda\lambda}$	$v_{\mu\mu}$	$v_{\sigma\sigma}$	$\rho_{\lambda\mu}^{(*)}$	$\rho_{\lambda\sigma}^{(*)}$	$\rho_{\mu\sigma}^{(*)}$
.00	1.00	1.00	.50	.000	.000	.000
.20	2.64	5.32	2.20	.705	.653	.773
.40	13.05	22.17	5.20	.933	.864	.907
.60	73.51	92.09	13.05	.982	.930	.956
.80	657.80	571.29	47.58	.994	.961	.980

(*) Coefficients de corrélation linéaire.

III – UTILISATION DE DEUX GROUPES EXPERIMENTAUX EN VUE D'AMELIORER L'ESTIMATION DE (λ , μ , σ)

En carcinogénèse chimique expérimentale on dispose de traitements initiateurs "puissants" c'est-à-dire induisant l'initiation d'un nombre important de cellules de l'organe traité, et de traitements promoteurs "puissants" c'est-à-dire induisant un développement rapide des cellules initiés en tumeurs observables. On peut mettre à profit l'existence de tels traitements pour améliorer l'estimation des paramètres du modèle.

III.A – Estimation du paramètre d'initiation λ

Pour améliorer l'estimation du paramètre d'initiation λ , on a intérêt à utiliser un traitement promoteur (μ_0 , σ_0) le plus puissant possible pour réduire le taux de troncature des observations. Il se peut cependant que la durée de l'expérience soit telle que les observations restent tronquées. Dans ce cas les estimateurs λ et ($\hat{\mu}_0$, $\hat{\sigma}_0$) étant corrélés, on peut penser qu'un supplément d'information concernant la valeur de (μ_0 , σ_0) serait susceptible d'améliorer l'estimation de λ . Les informations supplémentaires concernant (μ_0 , σ_0) peuvent être obtenues en utilisant un deuxième groupe expérimental recevant un traitement initiateur puissant λ_0 et le traitement promoteur précédant (μ_0 , σ_0). L'utilisation d'un initiateur puissant devant multiplier le nombre d'observations donnant des informations sur (μ_0 , σ_0).

Nous nous intéresserons donc au dispositif expérimental suivant qui comporte deux groupes expérimentaux :

– *1er groupe* dit "principal" : N . x animaux recevant le traitement initiateur étudié (λ) et un traitement promoteur (μ_0 , σ_0) ;

– 2ème groupe dit “secondaire” : $N \cdot (1 - x)$ animaux recevant un second traitement initiateur λ_0 et le même traitement promoteur (μ_0, σ_0) . N est le nombre total d’animaux et x est la proportion d’animaux consacrée au groupe principal.

L’optimisation de la valeur x dans ce type de schéma expérimental a été étudiée par Hemon [6]. En utilisant les notations de [6], la structure du dispositif expérimental est la suivante :

– un premier groupe expérimental fournit $N \cdot x$ observations indépendantes Y dont la distribution dépend du paramètre α qui intéresse l’expérimentateur et d’un paramètre nuisible β ,

– un second groupe expérimental fournit $N \cdot (1 - x)$ observations indépendantes Z dont la distribution ne dépend pas de α mais seulement de β et, éventuellement, d’un second paramètre nuisible γ . On montre dans (6) que la matrice des variances et covariances asymptotiques de l’estimateur du M.V. de α dans ce schéma à deux groupes est donnée par :

$$D(x) = \frac{1}{N} \left\{ \begin{array}{c} \frac{D_{\alpha\alpha}}{x} - \frac{D_{\alpha\beta}}{x} \left[\frac{D_{\beta\beta}}{x} + \frac{\bar{D}_{\beta\beta}}{1-x} \right]^{-1} \frac{D_{\beta\alpha}}{x} \\ \frac{D_{\beta\alpha}}{x} \end{array} \right\} \quad (3)$$

où

$$\cdot \left\{ \begin{array}{cc} D_{\alpha\alpha} & D_{\alpha\beta} \\ D_{\beta\alpha} & D_{\beta\beta} \end{array} \right\}$$

est la matrice des variances et covariances asymptotiques de l’estimateur du M.V. de (α, β) fondé sur une observation du terme principal (Y)

$$\text{et } \cdot \bar{D}_{\beta\beta}$$

est la matrice des variances et covariances asymptotiques de l’estimateur du M.V. de β fondé sur une observation du terme secondaire (Z) :

Dans le cas présent :

$$\left. \begin{array}{l} D(x) = \text{var}(\hat{\lambda}) \\ D_{\alpha\alpha} = v_{\lambda\lambda}(\pi_0) \quad \text{avec} \quad \pi_0 = 1 - F\left(\frac{t_0 - \mu_0}{\sigma_0}\right) \\ D_{\alpha\beta} [\sigma_0 v_{\lambda\mu}(\pi_0) \quad \sigma_0 v_{\lambda\sigma}(\pi_0)] \\ D_{\beta\beta} = \frac{\sigma_0^2}{\lambda} \left\{ \begin{array}{cc} v_{\mu\mu}(\pi_0) & v_{\mu\sigma}(\pi_0) \\ v_{\mu\sigma}(\pi_0) & v_{\sigma\sigma}(\pi_0) \end{array} \right\} \\ D_{\beta\beta} = \frac{\sigma_0^2}{\lambda_0} \left\{ \begin{array}{cc} v_{\mu\mu}(\pi_0) & v_{\mu\sigma}(\pi_0) \\ v_{\mu\sigma}(\pi_0) & v_{\sigma\sigma}(\pi_0) \end{array} \right\} \end{array} \right\} \quad (4)$$

En remplaçant dans (3) les expressions données en (4) pour les matrices D, il vient :

$$\text{var}(\lambda) = \frac{\lambda v_{\lambda\lambda}(\pi_0)}{N} \left\{ \frac{1}{x} - \frac{R^2}{x^2 \left[\frac{1}{x} + \frac{\lambda}{\lambda_0} \cdot \frac{1}{1-x} \right]} \right\} \quad (5)$$

où $R^2 = D_{\alpha\beta} D_{\beta\beta}^{-1} D_{\beta\alpha} / D_{\alpha\alpha}$ est le carré du coefficient de corrélation multiple entre les estimateurs $\hat{\lambda}$ et $(\hat{\mu}, \hat{\sigma})$ fondés sur une observation du groupe principal.

En annulant la dérivée de (5) par rapport à x, on détermine facilement la valeur optimale x^* comme racine d'un trinôme du second degré. A partir de cette valeur et de (5) on peut calculer la quantité :

$$\varphi = \text{var}(\hat{\lambda})_{x=x^*} / \text{var}(\hat{\lambda})_{x=1}$$

cette quantité mesure la réduction de variance de l'estimateur $\hat{\lambda}$ due à l'adoption du schéma optimal à la place du schéma simple $x = 1$.

La table 2 donne les valeurs numériques de x^* et φ pour quelques valeurs de π_0 et de λ/λ_0 . D'après les valeurs prises par φ on peut améliorer l'estimation de λ lorsque les observations sont tronquées et si l'initiateur λ_0 est nettement plus puissant que l'initiateur λ .

TABLE 2

Dispositif $N[x(\lambda, \mu_0, \sigma_0) + (1-x)(\lambda_0, \mu_0, \sigma_0)]$. Dans chaque case, la valeur supérieure est la valeur x^* de x qui minimise la variance de l'estimateur du M.V. $\hat{\lambda}$ et la valeur inférieure est celle de $\varphi = \text{var}(\lambda)_{x=x^*} / \text{var}(\lambda)_{x=1}$. Dans les cases blanches $x^* = 1, \varphi = 1$.

λ/λ_0 $\pi_0(*)$	1.00	.80	.60	.40	.20
.00					
.20				.90 (.98)	.82 (.88)
.40		.80 (.99)	.60 (.90)	.53 (.75)	.54 (.54)
.60		.43 (.92)	.33 (.77)	.31 (.58)	.34 (.37)
.80		.21 (.86)	.17 (.68)	.16 (.49)	.19 (.27)

$$(*) \pi_0 = 1 - F\left(\frac{t_0 - \mu}{\sigma}\right)$$

Par exemple si 40% des observations sont tronquées et si l'initiateur λ_0 est cinq fois plus puissant que l'initiateur λ ($\lambda/\lambda_0 = .20$), la valeur optimale de x est 0,54 et la réduction de variance de l'estimateur du M.V. $\hat{\lambda}$ du à l'adoption du schéma optimal est de 54%.

III.B - Estimation du paramètre de promotion μ

Pour améliorer l'estimation du paramètre de promotion μ on a intérêt à utiliser un traitement initiateur λ_0 puissant pour multiplier les observations fournissant des informations sur μ . Si le promoteur étudié est peu puissant et/ou si la durée de l'expérience est courte, les observations sont tronquées. Dans ce cas, les estimateurs $\hat{\lambda}_0$ et $\hat{\mu}$ sont corrélés, on peut penser qu'un supplément d'informations concernant la valeur de λ_0 serait susceptible d'améliorer l'estimation de μ . Les informations supplémentaires concernant λ_0 peuvent être obtenues en utilisant un deuxième groupe expérimental recevant l'initiateur précédant λ_0 et un traitement promoteur puissant (μ_0, σ_0). L'utilisation d'un promoteur puissant fournit des observations moins tronquées et donc plus d'informations sur la valeur de λ_0 .

Nous nous intéresserons donc au dispositif expérimental suivant, qui comporte deux groupes expérimentaux :

- 1^{er} groupe dit "principal" : $N \cdot x$ animaux recevant un traitement initiateur λ_0 et le traitement promoteur étudié (μ, σ),

- 2^e groupe dit "secondaire" : $N \cdot (1 - x)$ animaux recevant le même traitement initiateur λ_0 et un second traitement promoteur (μ_0, σ_0)

où N est le nombre total d'animaux et x la proportion d'animaux consacrée au groupe principal.

Nous avons affaire à un dispositif analogue à celui décrit précédemment. En reprenant les notations définies plus haut, on a cette fois :

$$D(x) = \text{var}(\hat{\mu})$$

$$D_{\alpha\alpha} = \frac{\sigma^2}{\lambda_0} v_{\mu\mu}(\pi) \quad D_{\alpha\beta} = \sigma v_{\lambda\mu}(\pi) \quad D_{\beta\beta} = \lambda_0 v_{\lambda\lambda}(\pi) \quad \bar{D}_{\beta\beta} = \lambda_0 v_{\lambda\lambda}(\pi_0)$$

avec

$$\pi = 1 - F\left(\frac{t_0 - \mu}{\sigma}\right) \quad \text{et} \quad \pi_0 = 1 - F\left(\frac{t_0 - \mu_0}{\sigma_0}\right)$$

soit :

$$\text{var}(\hat{\mu}) = \frac{\sigma^2 v_{\mu\mu}(\pi)}{N \lambda_0} \left\{ \frac{1}{x} - \frac{\rho_{\lambda\mu}^2(\pi)}{x^2 \left[\frac{1}{x} + \frac{v_{\lambda\lambda}(\pi_0)}{v_{\lambda\lambda}(\pi)} \cdot \frac{1}{1-x} \right]} \right\} \quad (6)$$

En annulant la dérivée de (6) par rapport à x on détermine la valeur optimale x^* comme racine d'un trinôme du second degré, puis la valeur de $\varphi = \text{var}(\hat{\mu})$ $x = x^*/\text{var}(\hat{\mu})$ $x = 1$ en utilisant (6).

La table 3 donne les valeurs numériques de x^* et φ pour quelques valeurs de π et π_0 . D'après les valeurs prises par φ , l'utilisation d'un terme secondaire peut améliorer sensiblement la qualité de l'estimation de λ si les observations sont moins tronquées dans le groupe secondaire que dans le terme principal. Ainsi, si le taux de troncature des observations est de 40 % dans le groupe principal et de 20% seulement dans le groupe secondaire, la valeur optimale de x est 0,54 et la réduction de variance de l'estimateur du M.V. due à l'adoption du schéma optimal est de 55 %.

TABLE 3

Dispositif $N[x \cdot (\lambda_0, \mu, \sigma) + (1 - x) (\lambda_0, \mu_0, \sigma_0)]$. Dans chaque case, la valeur supérieure est la valeur x^* de x qui minimise la variance de l'estimateur du M.V. $\hat{\mu}$ et la valeur inférieure est celle de $\varphi = \text{var}(\mu) x = x^* / \text{var}(\mu) x = 1$. Dans les cases blanches $x^* = 1, \varphi = 1$.

π_0 (*) \ / \ π (*)	.00	.20	.40	.60	.80
.00		.91 (.99)	.62 (.37)	.63 (.09)	.74 (.02)
20			.54 (.55)	.51 (.14)	.63 (.03)
.40				.35 (.34)	.44 (.06)
.60					.26 (.19)
.80					

$$(*) \pi_0 = 1 - F\left(\frac{t_0 - \mu_0}{\sigma_0}\right) \text{ et } \pi = 1 - F\left(\frac{t_0 - \mu}{\sigma}\right)$$

REFERENCES

- [1] BERENBLUM I. — The cocarcinogenic action of croton oil. *Cancer Research*, 1 : 44-48, 1941.
- [2] BERENBLUM I. — Carcinogenesis and tumour pathogenesis — *Advances. Cancer Research*, 2 : 129-175, 1954.
- [3] BOX M.J., DAVIES D., SWANN W.H. — Non linear optimization techniques. *ICI Monograph N° 5*, Oliver and Boyd 1969.
- [4] COHEN A.C. Jr. — Simplified estimators for the normal distribution when samples are singly censored or truncated. *Technometrics*, Vol. 1, 1959, 217-237.

- [5] COHEN A.C. Jr. — Tables for maximum likelihood estimates : singly truncated and singly censored samples. *Technometrics*, Vol. 3, 535-541, 1961.
- [6] HEMON D. — Optimal use of extraneous information concerning a nuisance parameter. *Biometrika*, 1977, 64, 1, 51-57.

ANNEXE

MATRICE DES VARIANCES ET COVARIANCES ASYMPTOTIQUES DE L'ESTIMATEUR DU MAXIMUM DE VRAISEMBLANCE $(\hat{\lambda}, \hat{\mu}, \hat{\sigma})$.

Considérons $N = 1$ animal observé pendant une période de temps t_0 . Soit n le nombre de cellules initiées promues en tumeur avant la fin de l'expérience et $(t_j/j = 1, n)$ les délais d'apparition de ces tumeurs.

En posant $z_0 = \frac{t_0 - \mu}{\sigma}$ et $z_j = \frac{t_j - \mu}{\sigma}$ il vient, d'après (1) § II.B :

$$\text{Log } L = -\lambda F(z_0) + n \text{Log}(\lambda) - n \text{Log}(\sigma) + \sum_{j=1}^n \text{Log}[f(z_j)] + \text{Cste}$$

où f et F sont respectivement la fonction de densité et la fonction cumulative de la distribution normale centrée réduite.

Soit :

$$\text{Log } L = -\lambda F(z_0) + n \text{Log}(\lambda) - n \text{Log}(\sigma) - \frac{1}{2} \sum_{j=1}^n z_j^2 + \text{Cste}$$

Pour calculer les dérivés d'ordre 1 et 2 de $\text{Log } L$, on peut préalablement remarquer que :

$$z = \frac{t - \mu}{\sigma} \Rightarrow \frac{\delta z}{\delta \mu} = -\frac{1}{\sigma} \frac{\delta z}{\delta \sigma} = -\frac{z}{\sigma}$$

et $F'(z) = f(z) \quad f'(z) = -zf(z)$

Dès lors :

$$\frac{\delta \text{Log } L}{\delta \lambda} = -F(z_0) + \frac{n}{\lambda}$$

$$\frac{\delta \text{Log } L}{\delta \mu} = \frac{\lambda}{\sigma} f(z_0) + \frac{1}{\sigma} \sum_{j=1}^n z_j$$

$$\frac{\delta \text{Log } L}{\delta \sigma} = \frac{\lambda}{\sigma} z_0 f(z_0) - \frac{n}{\sigma} + \frac{1}{\sigma} \sum_{j=1}^n z_j^2$$

et :

$$\begin{aligned}
 \frac{\delta^2 \text{Log L}}{\delta \lambda^2} &= -\frac{n}{\lambda^2} \\
 \frac{\delta^2 \text{Log L}}{\delta \lambda \delta \mu} &= \frac{1}{\sigma} f(z_0) \\
 \frac{\delta^2 \text{Log L}}{\delta \lambda \delta \sigma} &= \frac{1}{\sigma} z_0 f(z_0) \\
 \frac{\delta^2 \text{Log L}}{\delta \mu^2} &= \frac{\lambda}{\sigma^2} z_0 f(z_0) - \frac{n}{\sigma^2} \\
 \frac{\delta^2 \text{Log L}}{\delta \mu \delta \sigma} &= -\frac{\lambda}{\sigma^2} f(z_0) [1 - z_0^2] - \frac{2}{\sigma^2} \sum_{j=1}^n z_j \\
 \frac{\delta^2 \text{Log L}}{\delta \sigma^2} &= -\frac{\lambda}{\sigma^2} z_0 f(z_0) [2 - z_0^2] + \frac{n}{\sigma^2} - \frac{n}{\sigma^2} \sum_{j=1}^n z_j^2
 \end{aligned} \tag{A}$$

La matrice des informations de Fisher sur (λ, μ, σ) est donnée par :

$$M(\lambda, \mu, \sigma) = -E \left\{ \frac{D^2 \text{Log L}}{D(\lambda, \mu, \sigma)^2} \right\}$$

Pour calculer l'espérance de chacune des dérivées du deuxième ordre de Log L, on peut d'abord remarquer que si Z est une variable aléatoire normale centrée réduite et si z_0 est une valeur fixe, on a :

$$E(Z/Z \leq z_0) = \int_{-\infty}^{z_0} \frac{zf(z)}{F(z_0)} dz = -\frac{f(z_0)}{F(z_0)}$$

et

$$E(Z^2/Z \leq z_0) = \int_{-\infty}^{z_0} z^2 \frac{f(z)}{F(z_0)} dz = \left[1 - z_0 \frac{f(z_0)}{F(z_0)} \right]$$

en effet, on a remarqué plus haut que $f'(z) = -zf(z)$.

Dès lors :

$$E\left(\sum_{j=1}^n z_j\right) = E_n E\left(\sum_{j=1}^n z_j | n\right) = \lambda F(z_0) \left[-\frac{f(z_0)}{F(z_0)} \right] = -\lambda f(z_0)$$

et

$$E\left(\sum_{j=1}^n z_j^2\right) = E_n E\left(\sum_{j=1}^n z_j^2 | n\right) = \lambda F(z_0) \left[1 - z_0 \frac{f(z_0)}{F(z_0)} \right] \tag{B}$$

On déduit immédiatement de (A) et (B) l'expression analytique des éléments de la matrice M :

$$m_{\lambda\lambda} = E \left\{ - \frac{\delta^2 \text{Log L}}{\delta \lambda^2} \right\} = \frac{F(z_0)}{\lambda}$$

$$m_{\lambda\mu} = E \left\{ - \frac{\delta^2 \text{Log L}}{\delta \lambda} \right\} = - \frac{1}{\sigma} f(z_0)$$

$$m_{\lambda\sigma} = E \left\{ - \frac{\delta^2 \text{Log L}}{\delta \lambda \delta \sigma} \right\} = - \frac{1}{\sigma} z_0 \cdot f(z_0)$$

$$m_{\mu\mu} = E \left\{ - \frac{\delta^2 \text{Log L}}{\delta \mu^2} \right\} = \frac{\lambda}{\sigma^2} [F(z_0) - z_0 f(z_0)]$$

$$m_{\mu\sigma} = E \left\{ - \frac{\delta^2 \text{Log L}}{\delta \mu \delta \sigma} \right\} = \frac{\lambda}{\sigma^2} [1 + z_0^2] f(z_0)$$

La matrice M a donc la structure suivante :

$$M(\lambda, \mu, \sigma) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{\lambda}{\sigma} & 0 \\ 0 & 0 & \frac{\lambda}{\sigma} \end{pmatrix} M(1, t_0 - z_0, 1) \begin{pmatrix} \frac{1}{\lambda} & 0 & 0 \\ 0 & \frac{1}{\sigma} & 0 \\ 0 & 0 & \frac{1}{\sigma} \end{pmatrix}$$

et la matrice des variances et covariances asymptotiques de l'estimateur du M.V. s'écrit :

$$V(\lambda, \mu, \sigma) = M^{-1}(\lambda, \mu, \sigma) = \begin{pmatrix} \lambda v_{\lambda\lambda} & \sigma v_{\lambda\mu} & \sigma v_{\lambda\sigma} \\ \sigma v_{\lambda\mu} & \frac{\sigma^2}{\lambda} v_{\mu\mu} & \frac{\sigma^2}{\lambda} v_{\mu\sigma} \\ \sigma v_{\lambda\sigma} & \frac{\sigma^2}{\lambda} v_{\mu\sigma} & \frac{\sigma^2}{\lambda} v_{\sigma\sigma} \end{pmatrix}$$

où la matrice V est définie par :

$$V = M^{-1}(1, t_0 - z_0, 1)$$

et ne dépend que de z_0 ou, de façon équivalente, de la troncature des observations $\pi = 1 - F(z_0)$.