UTILIZACIÓN DE MÉTRICAS RIEMANNIANAS EN ANALISIS DE DATOS MULTIDIMENSIONALES Y SU APLICACIÓN A LA BIOLOGÍA

JOSE Mª OLLER SALA

BARCELONA, 25 de NOVIEMBRE de 1982.

debido a que la función de Bessel de indice ν:

$$J_{\nu}(x) = (\frac{x}{2})^{\nu} \sum_{n=o}^{\infty} \frac{(-1)^{n} (\frac{x}{2})^{2n}}{n!\,\Gamma(n+\nu+1)} \tag{110}$$

puede ser expresada en forma de integral como:

$$J_{\nu}(x) = \frac{(\frac{x}{2})^{\nu}}{\sqrt{\pi}\,\Gamma(\nu+\frac{1}{2})} \int_{o}^{\pi} \operatorname{sen}^{2\nu}\theta \; e^{-ix\cos\theta}\, d\theta \tag{111}$$

al integrar (109) entre 0 y π obtenemos el resultado:

<u>3.5.2.</u> La función de densidad, asintótica, de la distancia estimada entre dos poblaciones, conocida Δ, viene dada por:

$$f(D) = (\frac{-i}{\Delta})^{\frac{n-2}{2}} \frac{N_A N_B}{N_A+N_B} D^{n/2} e^{-\frac{1}{2}\frac{N_A N_B}{N_A+N_B}(D^2+\Delta^2)} J_{\frac{n-2}{2}}(i \frac{N_A N_B}{N_A+N_B} D\Delta) \tag{112}$$

$$D \geq 0$$

donde i es el número complejo $\sqrt{-1}$.

Nótese que a partir de (112) podremos resolver el contraste de hipótesis:

$$\begin{aligned} H_o&: \Delta = 0 \\ H_1&: \Delta > 0 \end{aligned} \tag{113}$$

Bajo la hipótesis nula, (112) se convierte en:

$$f(D|H_o) = \frac{1}{2^{\frac{n}{2}-1}} (\frac{N_A N_B}{N_A+N_B})^{n/2} D^{n-1} \frac{1}{\Gamma(\frac{n}{2})} e^{-\frac{1}{2}\frac{N_A N_B}{N_A+N_B} D^2} \qquad D \geq 0 \tag{114}$$

y si queremos hallar la función de densidad de la variable $\frac{N_A N_B}{N_A+N_B} D^2$, obtenemos:

$$f\left(\frac{N_A N_B}{N_A+N_B} D^2 \mid H_o\right) = \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} \left(\frac{N_A N_B}{N_A+N_B} D^2\right)^{\frac{n}{2}-1} e^{-\frac{1}{2} \frac{N_A N_B}{N_A+N_B} D^2} \qquad D \geq 0 \qquad (115)$$

en otras palabras, podemos enunciar el resultado:

3.5.3. Bajo la hipótesis nula, y asintóticamente, el estadístico:

$$U = \frac{N_A N_B}{N_A+N_B} D^2 \qquad (116)$$

converge a una distribución ji-cuadrado con n grados de libertad.

Por tanto, la región crítica asociada al contraste (113) será:

$$W_\varepsilon = \{x \in \mathbb{R}^{N_A+N_B} / U(x) > \chi^2_\varepsilon\}^2 \qquad (117)$$

Finalmente cabe destacar que (112) puede ser utilizado para el cálculo de la función de potencia (en función de Δ), siempre que consideremos tamaños muestrales elevados.

4. DISTANCIAS ENTRE DISTRIBUCIONES DISCRETAS.

Resumen:

En el presente capítulo se calcula la distancia, definida a través de la matriz de información de Fisher, a varias distribuciones de probabilidad discretas, entre otras, la distribución multinomial, las distribuciones de Poisson independientes y la distribución multinomial negativa.

4.1. DISTRIBUCION MULTINOMIAL

Sean $X^1,\dots,X^{n+1}$ variables aleatorias discretas, cuya función de densidad conjunta viene definida por:

$$f(x^1,\dots,x^{n+1}) = \frac{N!}{x^1!\cdot\ldots\cdot x^{n+1}!}\,(p^1)^{x^1}\cdot\ldots\cdot(p^{n+1})^{x^{n+1}} \qquad (1)$$

$$x^1,\dots,x^{n+1}\in \mathbb{N}\cup\{0\}$$

con las restricciones:

$$\sum_{i=1}^{n+1} x^i = N \qquad \text{y} \qquad \sum_{i=1}^{n+1} p^i = 1 \qquad (2)$$

El espacio paramétrico lo definiremos por:

$$E = \{(p^1,\dots,p^n)\in \mathbb{R}^n \,/\, p^i > 0 \quad i=1,\dots,n \quad \sum_{i=1}^{n} p^i < 1\} \qquad (3)$$

para evitar posteriores problemas en la definición del tensor métrico. Puede comprobarse de inmediato que E es una variedad n-dimensional arco-conexa.

El logaritmo de la función de densidad viene dado por:

$$\ln f = \sum_{h=1}^{n} x^h \ln p^h + (1-x^1-\dots-x^n)\ln(1-p^1-\dots-p^n) + cte \qquad (4)$$

derivando parcialmente respecto p^i y p^j obtenemos:

$$\frac{\partial^2 \ln f}{\partial p^i \partial p^j} = \frac{\partial}{\partial p^i}\left(\frac{x^j}{p^j} - \frac{(1-x^1-\dots-x^n)}{1-p^1-\dots-p^n}\right) = -\frac{\delta_{ij}x^j}{(p^j)^2} - \frac{(1-x^1-\dots-x^n)}{(1-p^1-\dots-p^n)^2} \qquad (5)$$

$$i,j = 1,\dots,n$$

siendo δ_{ij} las deltas de Kcroneker. Por tanto:

$$-E\left(\frac{\partial^2 \ln f}{\partial p^i \partial p^j}\right) = E\left(\frac{\delta_{ij} x^j}{(p^j)^2} + \frac{1-x^1-\ldots-x^n}{(1-p^1-\ldots-p^n)^2}\right) = N\left(\frac{\delta_{ij}}{p^j} + \frac{1}{1-p^1-\ldots-p^n}\right) \quad (6)$$

luego el tensor métrico viene dado por:

$$g_{\mu\nu} = N\left(\frac{\delta_{\mu\nu}}{p^{\mu}} + \frac{1}{1-p^1-\ldots-p^n}\right) \qquad \mu,\nu = 1,\ldots,n \quad (7)$$

y puede comprobarse con facilidad que es definido positivo en todo E y su determinante es igual a:

$$|g_{\mu\nu}| = N^n \frac{1}{p^1 \cdot p^2 \cdot \ldots \cdot p^n (1-p^1-\ldots-p^n)} \quad (8)$$

E es pues una variedad riemanniana.

El siguiente paso que podemos dar es calcular el recíproco del tensor métrico, $g_{\mu\nu}$, y obtenemos:

$$g^{\mu\nu} = \frac{1}{N}\left(\delta^{\mu\nu} p^{\mu} - p^{\mu} p^{\nu}\right) \quad (9)$$

En efecto,

$$\sum_{\nu=1}^{n} g_{\mu\nu} g^{\nu\lambda} = \sum_{\nu=1}^{n} N\left(\frac{\delta_{\mu\nu}}{p^{\mu}} + \frac{1}{1-p^1-\ldots-p^n}\right) \frac{1}{N}\left(\delta^{\nu\lambda} p^{\nu} - p^{\nu} p^{\lambda}\right) =$$

$$= \sum_{\nu=1}^{n} \delta_{\mu\nu}\delta^{\nu\lambda} \frac{p^{\nu}}{p^{\mu}} + \sum_{\nu=1}^{n} \frac{\delta^{\nu\lambda} p^{\nu}}{1-p^1-\ldots-p^n} - \sum_{\nu=1}^{n} \frac{\delta_{\mu\nu}}{p^{\mu}} p^{\nu} p^{\lambda} - \frac{\sum_{\nu=1}^{n} p^{\nu} p^{\lambda}}{1-p^1-\ldots-p^n} =$$

$$= \delta_{\mu}^{\lambda} \frac{p^{\lambda}}{p^{\mu}} + \frac{p^{\lambda}}{1-p^1-\ldots-p^n} - p^{\lambda} - \frac{p^{\lambda}(p^1+\ldots+p^n)}{1-p^1-\ldots-p^n} =$$

$$= \delta_{\mu}^{\lambda} + \frac{p^{\lambda}(p^1+\ldots+p^n)}{1-p^1-\ldots-p^n} - \frac{p^{\lambda}(p^1+\ldots+p^n)}{1-p^1-\ldots-p^n} = \delta_{\mu}^{\lambda} \quad (10)$$

Procedamos al cálculo de los símbolos de Christoffel de primera especie. Para ello:

$$\frac{\partial g_{ik}}{\partial p^j} = \frac{\partial}{\partial p^j} \left(N \left(\frac{\delta_{ik}}{p^i} + \frac{1}{1-p^1-\ldots-p^n}\right)\right) =$$

$$= N\left(-\delta_{ik}\delta_{ij} \frac{1}{(p^i)^2} + \frac{1}{(1-p^1-\ldots-p^n)^2}\right) \qquad (11)$$

Si definimos:

$$\delta_{ijk} = \begin{cases} 1 & \Leftrightarrow \quad i=j=k \\ 0 & \Leftrightarrow \quad i\neq j \text{ ó } i\neq k \text{ ó } k\neq j \end{cases} \qquad (12)$$

podemos escribir:

$$\frac{\partial g_{ik}}{\partial p^j} = -\delta_{ijk} \frac{N}{(p^i)^2} + \frac{N}{(1-p^1-\ldots-p^n)^2} \qquad (13)$$

además:

$$\frac{\partial g_{ik}}{\partial p^j} = \frac{\partial g_{jk}}{\partial p^i} = \frac{\partial g_{ij}}{\partial p^k} \qquad (14)$$

por tanto, los símbolos de Christoffel de primera especie son:

$$[i\ j,k] = -\delta_{ijk} \frac{N}{2(p^i)^2} + \frac{N}{2(1-p^1-\ldots-p^n)^2} \qquad (15)$$

$$i,j,k = 1,\ldots,n$$

Obtengamos a continuación los símbolos de Christoffel de segunda especie:

$$\Gamma^k_{ij} = \sum_{\alpha=1}^{n} \frac{1}{N} (\delta^{k\alpha}p^k - p^k p^\alpha)\ N\left(-\frac{1}{2}\,\delta_{ij\alpha} \frac{1}{(p^i)^2} + \frac{1}{2(1-p^1-\ldots-p^n)}\right) =$$

$$= \sum_{\alpha=1}^{n} \left(- \frac{1}{2} \delta^{k\alpha} \delta_{ij\alpha} \frac{p^k}{(p^i)^2} + \frac{1}{2} p^k p^\alpha \delta_{ij\alpha} \frac{1}{(p^i)^2} + \frac{\delta^{k\alpha} p^k}{2(1-p^1-\ldots-p^n)^2} - \right.$$

$$\left. - \frac{p^k p^\alpha}{2(1-p^1-\ldots-p^n)^2} \right) =$$

$$= - \frac{1}{2} \delta_{ijk} \frac{1}{p^i} + \frac{1}{2} p^k \delta_{iji} \frac{1}{p^i} + \frac{p^k}{2(1-p^1-\ldots-p^n)^2} - \frac{p^k(1-p^1-\ldots-p^k)}{2(1-p^1-\ldots-p^k)^2} \qquad (16)$$

teniendo en cuenta que $\delta_{iji} = \delta_{ij}$ agrupando términos, y simplificando obtenemos:

$$\Gamma^k_{ij} = - \frac{1}{2} \delta_{ijk} \frac{1}{p^i} + \frac{1}{2} \delta_{ij} \frac{p^k}{p^i} + \frac{p^k}{2(1-p^1-\ldots-p^n)} \qquad (17)$$

A continuación vamos a calcular el tensor de Riemann-Christoffel de segunda especie. Calculemos en primer lugar:

$$\Gamma^\beta_{ik} \Gamma^a_{\beta j} = \sum_{\beta=1}^{n} \left(- \frac{1}{2} \delta_{ik\beta} \frac{1}{p^i} + \frac{1}{2} \delta_{ik} \frac{p^\beta}{p^i} + \frac{p^\beta}{2(1-p^1-\ldots-p^n)} \right) \cdot$$

$$\cdot \left(- \frac{1}{2} \delta_{\beta ja} \frac{1}{p^\beta} + \frac{1}{2} \delta_{\beta j} \frac{p^a}{p^\beta} + \frac{p^a}{2(1-p^1-\ldots-p^n)} \right) =$$

$$= \sum_{\beta=1}^{n} \left(\frac{1}{4} \delta_{ik\beta} \delta_{\beta ja} \frac{1}{p^i p^\beta} - \frac{1}{4} \delta_{ik\beta} \delta_{\beta j} \frac{p^a}{p^i p^\beta} - \frac{1}{4} \delta_{ik\beta} \frac{p^a}{p^i (1-p^1-\ldots-p^n)} - \right.$$

$$- \frac{1}{4} \delta_{ik} \delta_{\beta ja} \frac{1}{p^i} + \frac{1}{4} \delta_{ik} \delta_{\beta j} \frac{p^a}{p^i} + \frac{1}{4} \delta_{ik} \frac{p^\beta p^a}{p^i (1-p^1-\ldots-p^n)} -$$

$$\left. - \frac{1}{4} \delta_{\beta ja} \frac{1}{(1-p^1-\ldots-p^n)} + \frac{1}{4} \delta_{\beta j} \frac{p^a}{(1-p^1-\ldots-p^n)} + \frac{1}{4} \frac{p^a p^\beta}{(1-p^1-\ldots-p^n)^2} \right) =$$

$$= \frac{1}{4} \delta_{ikja} \frac{1}{(p^i)^2} - \frac{1}{4} \delta_{ikj} \frac{p^a}{(p^i)^2} - \frac{1}{4} \delta_{ik} \frac{p^a}{p^i (1-p^1-\ldots-p^n)} -$$

$$- \frac{1}{4} \delta_{ik} \delta_{ja} \frac{1}{p^i} + \frac{1}{4} \delta_{ik} \frac{p^a}{p^i} + \frac{1}{4} \delta_{ik} \frac{p^a \sum_{\beta=1}^{n} p^\beta}{p^i (1-p^1-\ldots-p^n)} -$$

$$- \frac{1}{4} \delta_{ja} \frac{1}{(1-p^1-\ldots-p^n)} + \frac{1}{4} \frac{p^a}{(1-p^1-\ldots-p^n)} + \frac{1}{4} \frac{p^a \sum_{\beta=1}^{n} p^\beta}{p^i (1-p^1-\ldots-p^n)^2} \qquad (18)$$

definiendo δ_{ikja} a partir de:

$$\delta_{ikja} = \begin{cases} 1 \Leftrightarrow i=k=j=a \\ 0 \Leftrightarrow \text{en caso contrario} \end{cases} \qquad (19)$$

agrupando términos podemos escribir:

$$\Gamma^\beta_{ik} \Gamma^a_{\beta j} = \frac{1}{4} \delta_{ikja} \frac{1}{(p^i)^2} - \frac{1}{4} \frac{\delta_{ikj}\, p^a}{(p^i)^2} - \frac{1}{4} \delta_{ik}\, \delta_{ja} \frac{1}{p^i} +$$

$$+ \frac{1}{4} \delta_{ik} \left(\frac{-p^a}{p^i (1-p^1-\ldots-p^n)} + \frac{p^a}{p^i} + \frac{p^a \sum_{\beta=1}^{n} p^\beta}{p^i (1-p^1-\ldots-p^n)} \right) -$$

$$- \frac{1}{4} \delta_{ja} \frac{1}{1-p^1-\ldots-p^n} + \frac{1}{4} \frac{p^a}{(1-p^1-\ldots-p^n)^2} = \frac{1}{4(p^i)^2} (\delta_{ikja}-\delta_{ikj} p^a) -$$

$$- \frac{1}{4p^i} (\delta_{ik} \delta_{ja}) + \frac{p^a - \delta_{ja} (1-p^1-\ldots-p^n)}{4(1-p^1-\ldots-p^n)^2} \qquad (20)$$

por tanto, cambiando la j por la k, podemos exribir:

$$\Gamma^\beta_{ij} \Gamma^a_{\beta k} = \frac{1}{4(p^i)^2} (\delta_{ijka}-\delta_{ijk} p^a) - \frac{1}{4p^i} (\delta_{ij} \delta_{ka}) + \frac{p^a - \delta_{ka} (1-p^1-\ldots-p^n)}{4(1-p^1-\ldots-p^n)^2} \qquad (21)$$

luego resulta:

$$\Gamma^\beta_{ik} \Gamma^a_{\beta j} - \Gamma^\beta_{ij} \Gamma^a_{\beta k} = \frac{1}{4p^i} (\delta_{ij} \delta_{ka}-\delta_{ik} \delta_{ja}) + \frac{1}{4} \frac{(\delta_{ka} - \delta_{ja})}{1-p^1-\ldots-p^n} \qquad (22)$$

Hallemos ahora:

$$\frac{\partial \Gamma^a_{ik}}{\partial p^j} = \frac{\partial}{\partial p^j}\left[-\frac{1}{2}\delta_{ika}\frac{1}{p^i} + \frac{1}{2}\delta_{ik}\frac{p^a}{p^i} + \frac{p^a}{2(1-p^1-\ldots-p^n)} \right] =$$

$$= \frac{1}{2}\delta_{ika}\frac{\delta_{ij}}{(p^i)^2} + \frac{1}{2}\delta_{ik}\frac{p^i\delta_{aj}-p^a\delta_{ij}}{(p^i)^2} + \frac{1}{2}\frac{(1-p^1-\ldots-p^n)\delta_{aj}+p^a}{(1-p^1-\ldots-p^n)^2} =$$

$$= \frac{1}{2}\delta_{ikaj}\frac{1}{(p^i)^2} + \frac{1}{2}\delta_{ik}\delta_{aj}\frac{1}{p^i} - \frac{1}{2}\delta_{ijk}\frac{p^a}{(p^i)^2} + \frac{\delta_{aj}}{2(1-p^1-\ldots-p^n)} +$$

$$+ \frac{1}{2}\frac{p^a}{(1-p^1-\ldots-p^n)^2} \qquad (23)$$

por tanto, cambiando la k por la j, obtenemos:

$$\frac{\partial \Gamma^a_{ij}}{\partial p^k} = \frac{1}{2}\delta_{ijak}\frac{1}{(p^i)^2} + \frac{1}{2}\delta_{ij}\delta_{ak}\frac{1}{p^i} - \frac{1}{2}\delta_{ikj}\frac{p^a}{(p^i)^2} + \frac{\delta_{ak}}{2(1-p^1-\ldots-p^n)} +$$

$$+ \frac{1}{2}\frac{p^a}{(1-p^1-\ldots-p^n)^2} \qquad (24)$$

luego resulta:

$$\frac{\partial \Gamma^a_{ik}}{\partial p^j} - \frac{\partial \Gamma^a_{ij}}{\partial p^k} = \frac{1}{2p^i}(\delta_{ik}\delta_{aj}-\delta_{ij}\delta_{ak}) + \frac{\delta_{aj}-\delta_{ak}}{2(1-p^1-\ldots-p^n)} \qquad (25)$$

por tanto resulta:

$$R^a_{ijk} = \frac{1}{4p^i}(\delta_{ik}\delta_{aj} - \delta_{ij}\delta_{ak}) + \frac{\delta_{aj}-\delta_{ak}}{4(1-p^1-\ldots-p^n)} \qquad (26)$$

$$i,j,k,a = 1,\ldots,n$$

Al no anularse el tensor de Riemann-Christoffel de segunda especie, no existe una transformación admisible de coordenadas (parámetros) que reduzca al tensor métrico a un tensor constante.

A continuación procederemos a calcular la curvatura Riemanniana de la variedad E, para ello deberemos hallar el tensor de Riemmann-Christoffel de primera especie.

$$R_{hijk} = \sum_{a=1}^{n} N\left(\delta_{ha}\frac{1}{p^h} + \frac{1}{1-p^1-\ldots-p^n}\right)\left(\frac{1}{4p^i}(\delta_{ik}\delta_{aj}-\delta_{ij}\delta_{ak}) + \frac{\delta_{aj}-\delta_{ak}}{4(1-p^1-\ldots-p^n)}\right) =$$

$$= N\sum_{a=1}^{n}\left(\frac{\delta_{ha}(\delta_{aj}-\delta_{ak})}{4p^h(1-p^1-\ldots-p^n)} + \frac{\delta_{aj}-\delta_{ak}}{4(1-p^1-\ldots-p^n)^2} + \right.$$

$$\left. + \frac{\delta_{ha}}{4p^h p^i}(\delta_{ik}\delta_{aj}-\delta_{ij}\delta_{ak}) + \frac{\delta_{ik}\delta_{aj}-\delta_{ij}\delta_{ak}}{4p^i(1-p^1-\ldots-p^n)}\right) =$$

$$= N\left(\frac{\delta_{hj}-\delta_{hk}}{p^h(1-p^1-\ldots-p^n)} + \frac{\delta_{ik}\delta_{hj}-\delta_{ij}\delta_{hk}}{4p^h p^i} + \frac{\delta_{ik}-\delta_{ij}}{4p^i(1-p^1-\ldots-p^n)}\right) \qquad (27)$$

que podemos escribir como:

$$R_{hijk} = N\left(\frac{\delta_{ik}\delta_{hj}-\delta_{ij}\delta_{hk}}{4p^h p^i} + \frac{(\delta_{hj}-\delta_{hk})p^i+(\delta_{ik}-\delta_{ij})p^h}{4p^h p^i(1-p^1-\ldots-p^n)}\right) \qquad (28)$$

$$h,i,j,k = 1,\ldots,n$$

Por otra parte:

$$g_{hj}\,g_{ik} = N\left(\frac{\delta_{hj}}{p^h} + \frac{1}{1-p^1-\ldots-p^n}\right) N\left(\frac{\delta_{ik}}{p^i} + \frac{1}{1-p^1-\ldots-p^n}\right) =$$

$$= N^2\left(\frac{\delta_{hj}\delta_{ik}}{p^h p^i} + \frac{\delta_{hj}p^i+\delta_{ik}p^h}{p^h p^i(1-p^1-\ldots-p^n)} + \frac{1}{(1-p^1-\ldots-p^n)^2}\right) \qquad (29)$$

por tanto, cambiando la j por la k podemos escribir:

$$g_{hk}g_{ij} = N^2\left(\frac{\delta_{hk}\delta_{ij}}{p^h p^i} + \frac{\delta_{hk}p^i+\delta_{ij}p^h}{p^h p^i(1-p^1-\ldots-p^n)} + \frac{1}{(1-p^1-\ldots-p^n)^2} \right) \qquad (30)$$

luego resulta:

$$g_{hj}g_{ik}-g_{hk}g_{ij} = N^2\left(\frac{\delta_{hj}\delta_{ik}-\delta_{hk}\delta_{ij}}{p^h p^i} + \frac{(\delta_{hj}-\delta_{hk})p^i+(\delta_{ik}-\delta_{ij})p^h}{p^h p^i(1-p^1-\ldots-p^n)} \right) \qquad (31)$$

por tanto:

$$R_{hijk} = \frac{1}{4N}\left(g_{hj}g_{ik} - g_{hk}g_{ij}\right) \qquad (32)$$

$$h,i,j,k = 1,\ldots,n$$

por tanto la curvatura Riemanniana es constante en todo el espacio e independiente de la orientación del plano que consideremos (espacio isótropo):

$$k = \frac{1}{4N} \qquad (33)$$

La variedad riemanniana E será, localmente, isométrica a la superficie de una hiperesfera de radio $2\sqrt{N}$ en un espacio euclideo n+1 dimensional.

Por no ser el espacio euclideo, para hallar la distancia entre dos puntos de E, tendremos que calcular la geodésica que une a ambos. En nuestro caso las ecuaciones de las geodésicas son:

$$\frac{d^2p^\gamma}{ds^2} + \sum_{\alpha=1}^{n}\sum_{\beta=1}^{n}\left(-\frac{1}{2}\delta_{\alpha\beta\gamma}\frac{1}{p^\alpha} + \frac{1}{2}\delta_{\alpha\beta}\frac{p^\gamma}{p^\alpha} + \frac{p^\gamma}{2(1-p^1-\ldots-p^n)}\right)\frac{dp^\alpha}{ds}\frac{dp^\beta}{ds} = 0$$

$$\gamma=1,\ldots,n \qquad (34)$$

El sistema puede escribirse como:

$$\frac{d^2p^\gamma}{ds^2} - \frac{1}{2p^\gamma}\frac{dp^\gamma}{ds}\frac{dp^\gamma}{ds} + \frac{p^\gamma}{2}\sum_{\beta=1}^{n}\frac{1}{p^\beta}\frac{dp^\beta}{ds}\frac{dp^\beta}{ds} +$$

$$+ \frac{p^\gamma}{2(1-p^1-\ldots-p^n)}\sum_{\alpha=1}^{n}\sum_{\beta=1}^{n}\frac{dp^\alpha}{ds}\frac{dp^\beta}{ds} = 0 \qquad \gamma=1,\ldots,n \tag{35}$$

y hay que resolver con la condición:

$$\sum_{\mu=1}^{n}\sum_{\nu=1}^{n} N\left(\frac{\delta_{\mu\nu}}{p^\mu} + \frac{1}{1-p^1-\ldots-p^n}\right)\frac{dp^\mu}{ds}\frac{dp^\nu}{ds} = 1 \tag{36}$$

desarrollando (36) obtenemos:

$$N\sum_{\mu=1}^{n}\frac{1}{p^\mu}\frac{dp^\mu}{ds}\frac{dp^\mu}{ds} + N\frac{1}{1-p^1-\ldots-p^n}\sum_{\mu=1}^{n}\sum_{\nu=1}^{n}\frac{dp^\mu}{ds}\frac{dp^\nu}{ds} = 1 \tag{37}$$

por lo que el sistema (35) puede escribirse como:

$$\frac{d^2p^\gamma}{ds^2} - \frac{1}{2p^\gamma}\frac{dp^\gamma}{ds}\frac{dp^\gamma}{ds} + \frac{p^\gamma}{2N} = 0 \tag{38}$$

$$\gamma=1,\ldots,n$$

Teniendo en cuenta que $p^{n+1} = 1-p^1-\ldots-p^n$, se cumple que:

$$\sum_{\gamma=1}^{n}\frac{d^2p^\gamma}{ds^2} = -\frac{d^2p^{n+1}}{ds^2} \qquad \sum_{\gamma=1}^{n}\frac{dp^\gamma}{ds} = -\frac{dp^{n+1}}{ds} \tag{39}$$

y por tanto, si sumamos las n ecuaciones del sistema (38) resulta:

$$-\frac{d^2p^{n+1}}{ds^2}-\sum_{\gamma=1}^{n}\frac{1}{2p^\gamma}\frac{dp^\gamma}{ds}\frac{dp^\gamma}{ds}+\frac{1-p^{n+1}}{2N}=0 \tag{40}$$

y teniendo en cuenta que, debido a (37) podemos escribir:

$$\sum_{\gamma=1}^{n}\frac{1}{p^\gamma}\frac{dp^\gamma}{ds}\frac{dp^\gamma}{ds}=\frac{1}{N}-\frac{1}{p^{n+1}}\frac{dp^{n+1}}{ds}\frac{dp^{n+1}}{ds} \tag{41}$$

resulta finalmente, al sustituir en (40):

$$\frac{d^2p^{n+1}}{ds^2}-\frac{1}{2p^{n+1}}\frac{dp^{n+1}}{ds}\frac{dp^{n+1}}{ds}+\frac{p^{n+1}}{2N}=0 \tag{42}$$

por lo que el sistema (38) puede ser escrito de forma mas simétrica como:

$$\frac{d^2p^\gamma}{ds^2}-\frac{1}{2p^\gamma}\frac{dp^\gamma}{ds}\frac{dp^\gamma}{ds}+\frac{p^\gamma}{2N}=0 \qquad \gamma=1,\ldots,n+1 \tag{43}$$

Efectuemos el cambio:

$$(u^\gamma)^2=p^\gamma \qquad \gamma=1,\ldots,n+1 \tag{44}$$

entonces:

$$\frac{dp^\gamma}{ds}=2u^\gamma\frac{du^\gamma}{ds} \qquad \frac{d^2p^\gamma}{ds^2}=2\frac{du^\gamma}{ds}\frac{du^\gamma}{ds}+2u^\gamma\frac{d^2u^\gamma}{ds^2} \tag{45}$$

por tanto resulta:

$$\frac{d^2u^\gamma}{ds^2}+\frac{1}{4N}u^\gamma=0 \qquad \gamma=1,\ldots,n+1 \tag{46}$$

cada una de estas ecuaciones es una ecuación diferencial lineal de segundo orden, homogenea, cuya solución viene dada por:

$$u^{\gamma} = A_{\gamma} \cos\left(\frac{1}{2\sqrt{N}} s\right) + B_{\gamma} \operatorname{sen}\left(\frac{1}{2\sqrt{N}} s\right) \qquad \gamma=1,\ldots n+1 \tag{47}$$

o equivalentemente, las ecuaciones de las geodésicas vendrán dadas por:

$$p^{\gamma} = \left(A_{\gamma} \cos\left(\frac{1}{2\sqrt{N}} s\right) + B_{\gamma} \operatorname{sen}\left(\frac{1}{2\sqrt{N}} s\right)\right)^2 \qquad \gamma=1,\ldots,n+1 \tag{48}$$

donde las A_{γ} y B_{γ} son constantes de integración.

Hemos de determinar $A_1,\ldots,A_{n+1}$, $B_1,\ldots,B_{n+1}$ y s de forma que las geodésicas pasen por $a^1,\ldots,a^{n+1}$ y $b^1,\ldots,b^{n+1}$. En primer lugar, hallaremos las condiciones que debemos imponer a las A_i y B_i para que la suma de las p^i sea igual a 1. Por tanto:

$$\sum_{\gamma=1}^{n+1} \left(A_{\gamma} \cos \frac{1}{2\sqrt{N}} s + B_{\gamma} \operatorname{sen} \frac{1}{2\sqrt{N}} s\right)^2 = \sum_{\gamma=1}^{n+1} \left(A_{\gamma}^2 \cos^2 \frac{s}{2\sqrt{N}} + B_{\gamma}^2 \operatorname{sen}^2 \frac{s}{2\sqrt{N}} + \right.$$

$$\left. + 2 A_{\gamma} B_{\gamma} \operatorname{sen} \frac{s}{2\sqrt{N}} \cos \frac{s}{2\sqrt{N}}\right) = \cos^2 \frac{s}{2\sqrt{N}} \left(\sum_{\gamma=1}^{n+1} A_{\gamma}^2\right) + \operatorname{sen}^2 \frac{s}{2\sqrt{N}} \left(\sum_{\gamma=1}^{n+1} B_{\gamma}^2\right) +$$

$$+ 2 \operatorname{sen} \frac{s}{2\sqrt{N}} \cos \frac{s}{2\sqrt{N}} \left(\sum_{\gamma=1}^{n+1} A_{\gamma} B_{\gamma}\right) = 1 \tag{49}$$

y esta igualdad debe cumplirse para cualquier s. Esta igualdad es equivalente a:

$$\operatorname{sen}^2 \frac{s}{2\sqrt{N}} \left(\sum_{\gamma=1}^{n+1} B_{\gamma}^2 - \sum_{\gamma=1}^{n+1} A_{\gamma}^2\right) + \operatorname{sen} \frac{s}{\sqrt{N}} \sum_{\gamma=1}^{n+1} A_{\gamma} B_{\gamma} = 0 \tag{50}$$

que por verificarse para cualquier s, equivale a:

$$\sum_{\gamma=1}^{n+1} B_\gamma^2 = \sum_{\gamma=1}^{n+1} A_\gamma^2 \qquad \sum_{\gamma=1}^{n+1} A_\gamma B_\gamma = 0 \tag{51}$$

pero además, sustituyendo s=0 en (49) obtenemos:

$$\sum_{\gamma=1}^{n+1} A_\gamma^2 = 1 \tag{52}$$

Además, debera verificarse que para s=0, la geodésica pase por el punto $a^1,\ldots,a^{n+1}$. De esta condición se sigue:

$$A_\gamma = \pm\sqrt{a^\gamma} \qquad \gamma=1,\ldots,n \tag{53}$$

Para un cierto s, la geodésica pasará por el punto $b^1,\ldots,b^{n+1}$, por tanto:

$$\pm\sqrt{b^\gamma} = A_\gamma \cos\frac{s}{2\sqrt{N}} + B_\gamma \operatorname{sen}\frac{s}{2\sqrt{N}} \qquad \gamma=1,\ldots,n+1 \tag{54}$$

por lo que resulta:

$$B_\gamma = \frac{\left(\pm\sqrt{b^\gamma} - A_\gamma \cos\frac{s}{2\sqrt{N}}\right)}{\operatorname{sen}\frac{s}{2\sqrt{N}}} \qquad \gamma=1,\ldots,n+1 \tag{55}$$

debido a (51) resulta:

$$\sum_{\gamma=1}^{n+1} A_\gamma\left(\pm\sqrt{b^\gamma} - A_\gamma \cos\frac{s}{2\sqrt{N}}\right) = 0 \tag{56}$$

equivalente a:

$$\sum_{\gamma=1}^{n+1} \pm \sqrt{a^\gamma}\sqrt{b^\gamma} - \sum_{\gamma=1}^{n+1} A_\gamma^2 \cos\frac{s}{2\sqrt{N}} = 0 \tag{57}$$

y por (52):

$$s = 2\sqrt{N} \text{ arc cos } \left(\sum_{\gamma=1}^{n+1} \sqrt{a^\gamma b^\gamma}\right) \tag{58}$$

eligiéndose los signos de $\sqrt{a^\gamma b^\gamma}$ positivos para obtener el mínimo valor posible para s.

Vamos a comprobar que si $(a^1,\ldots,a^n)\in E$ y $(b^1,\ldots,b^n)\in E$ entonces siempre se cumple:

$$0 \leq \sum_{\gamma=1}^{n+1} \sqrt{a^\gamma b^\gamma} \leq 1 \tag{59}$$

siendo $a^{n+1} = 1-a^1-\ldots-a^n$ y $b^{n+1} = 1-b^1-\ldots-b^n$. En efecto, la positividad es trivial, además:

$$0 \leq \sum_{\gamma=1}^{n+1} (\sqrt{a^\gamma} - \sqrt{b^\gamma})^2 = \sum_{\gamma=1}^{n+1} (a^\gamma + b^\gamma - 2\sqrt{a^\gamma b^\gamma}) =$$

$$= 2\left(1 - \sum_{\gamma=1}^{n+1} \sqrt{a^\gamma b^\gamma}\right) \tag{60}$$

ello significa que dados dos puntos de E siempre existe una geodésica que los une.

La distancia entre dos puntos de la variedad E siempre estará comprendida entre 0 y $\sqrt{N}\;\pi$, como pueda comprobarse sustituyendo en (58) $\sum_{\gamma=1}^{n+1} \sqrt{a^\gamma b^\gamma}$ por su valor mínimo, 0.

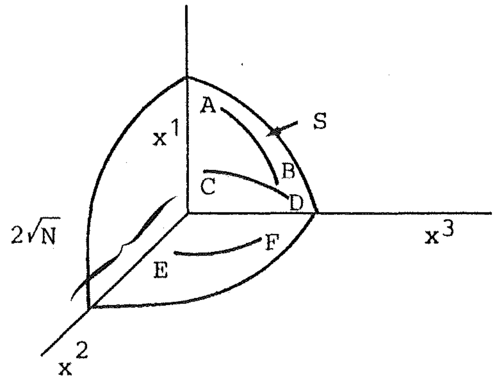
Es posible también distanciar puntos situados en la frontera de E, con un paso al límite, usando la misma fórmula (58) por razones de continuidad en la frontera.

La distancia obtenida en (58) es igual, salvo constantes, a la distancia propuesta por Bhattacharyya, (1946), heurísticamente, considerando las probabilidades como los cosenos directores al cuadrado de un vector en un espacio euclídeo y tomando como distancia, entre dos de estos vectores, el angulo entre ellos. Fue usada posteriormente por Cavalli-Sforza en genética de poblaciones.

La variedad E es globalmente isométrica a una parte de la superfiie de una hiperesfera de radio $2\sqrt{N}$ en un espacio euclídeo n+1 dimensional, siendo la métrica de dicha hipersuperficie la métrica euclídea inducida en ella. Concretamente sería parte de la superficie definida por:

$$S = \{(x^1,\ldots,x^{n+1}) \in \mathbb{R}^{n+1} \ / \ x^i = 2\sqrt{N}\ \sqrt{p^i} \quad i=1,\ldots,n+1\} \qquad (61)$$

En el caso de una distribución trinomial: n+1 = 3, podemos ilustrarlo en la figura que sigue:



4.2. ALGUNAS PROPIEDADES DE LA DISTANCIA ASOCIADA A LA DISTRIBUCION MULTINOMIAL.

Vamos a examinar a continuación el comportamiento de la distancia al refinar la partición asociada a la distribución multinomial. Sea $A_1,\ldots$ $\ldots,A_n$ una partición de un conjunto Ω de sucesos elementales, verificándose:

$$P(A_i) = a^i \quad \text{(en población A)}$$
$$P(A_i) = b^i \quad \text{(en población B)} \tag{62}$$
$$i = 1,\ldots,n$$

Subdividamos A_n en $B_1,\ldots,B_k$ conjuntos disjuntos, con:

$$P(B_i) = \alpha^i \quad \text{(en población A)}$$
$$P(B_i) = \beta^i \quad \text{(en población B)} \tag{63}$$
$$i = 1,\ldots,k$$

verificando:

$$a^n = \alpha^1 + \ldots + \alpha^k$$
$$b^n = \beta^1 + \ldots + \beta^k \tag{64}$$

Entonces la distancia entre las poblaciones A y B utilizando la partición $A_1,\ldots,A_n$ es menor que la obtenida utilizando la partición refinada:

$$s = 2\sqrt{N}\,\text{arccos}\Big(\sum_{\gamma=1}^{n} \sqrt{a^\gamma b^\gamma}\Big) \leq 2\sqrt{N}\,\text{arccos}\Big(\sum_{\gamma=1}^{n-1} \sqrt{a^\gamma b^\gamma} + \sum_{\delta=1}^{k} \sqrt{\alpha^\delta \beta^\delta}\Big) \tag{65}$$

En efecto, bastará probar que:

$$\sqrt{a^n b^n} \geq \sum_{\delta=1}^{k} \sqrt{\alpha^\delta \beta^\delta} \tag{66}$$

equivalente a:

$$a^n b^n \geq \left(\sum_{\delta=1}^{k} \sqrt{\alpha^\delta \beta^\delta}\right)^2 \tag{67}$$

desarrollando el segundo término:

$$\left(\sum_{\delta=1}^{k} \sqrt{\alpha^\delta \beta^\delta}\right)^2 = \sum_{\delta=1}^{k} \alpha^\delta \beta^\delta + \sum_{\mu\neq\delta} \sqrt{\alpha^\delta \beta^\delta \alpha^\mu \beta^\mu} \tag{68}$$

y si tenemos en cuenta que:

$$\sum_{\delta=1}^{k} \alpha^\delta \sum_{\mu=1}^{k} \beta^\mu = \sum_{\delta=1}^{k} \alpha^\delta \beta^\delta + \sum_{\mu\neq\delta} \alpha^\delta \beta^\mu \tag{69}$$

debido a que se cumple (64), será suficiente probar:

$$\sum_{\alpha<\mu} (\alpha^\delta \beta^\mu + \alpha^\mu \beta^\delta) \geq 2 \sum_{\delta<\mu} \sqrt{\alpha^\delta \beta^\delta \alpha^\mu \beta^\mu} \tag{70}$$

y ello es cierto debido a que:

$$(\alpha^\delta \beta^\mu + \alpha^\mu \beta^\delta)^2 = (\alpha^\delta \beta^\mu)^2 + (\alpha^\mu \beta^\delta)^2 + 2\ \alpha^\delta \alpha^\mu \beta^\delta \beta^\mu \geq 4\ \alpha^\delta \alpha^\mu \beta^\delta \beta^\mu \tag{71}$$

ya que:

$$(\alpha^\delta \beta^\mu - \alpha^\mu \beta^\delta)^2 \geq 0 \tag{72}$$

Con ello hemos probado que al subdividir un átomo de la partición aumenta (o se queda igual) la distancia entre dos poblaciones, hecho natural pues aumenta nuestra información sobre las analogías y diferencias de las dos poblaciones.

Esta distancia "valora" más las diferencias cuando las probabilidades son bajas que cuando son altas, este hecho queda patente contemplando la diferencial de s:

$$ds = \sum_{\gamma=1}^{n+1} \left(\frac{\partial s}{\partial a^\gamma} da^\gamma + \frac{\partial s}{\partial b^\gamma} db^\gamma \right) = -\sum_{\gamma=1}^{n+1} \frac{2\sqrt{N}}{\sqrt{1-\left(\sum_{\gamma=1}^{n+1}\sqrt{a^\gamma b^\gamma}\right)^2}} \frac{1}{2}(a^\gamma b^\gamma)^{-1/2} \cdot$$

$$\cdot\ (b^\gamma da^\gamma + a^\gamma db^\gamma) = -\frac{\sqrt{N}}{\sqrt{1-\left(\sum_{\gamma=1}^{n+1}\sqrt{a^\gamma b^\gamma}\right)^2}} \sum_{\gamma=1}^{n+1} \frac{1}{\sqrt{a^\gamma b^\gamma}} (b^\gamma da^\gamma + a^\gamma db^\gamma) \tag{73}$$

y debido a que $\sum_{\gamma=1}^{n+1} \sqrt{a^\gamma b^\gamma} = \cos s$ resulta:

$$ds = -\frac{\sqrt{N}}{\operatorname{sen} s} \sum_{\gamma=1}^{n} \frac{1}{\sqrt{a^\gamma b^\gamma}} (b^\gamma da^\gamma + a^\gamma db^\gamma) \tag{74}$$

Consideremos a continuación la distribución multinomial construida sobre el producto cartesiano de k particiones, finitas e independientes del conjunto de sucesos elementales asociado a una experiencia aleatoria, sean

$$\mathcal{P}_i = \{A_{i1}, \ldots, A_{in_i}\} \qquad i=1,\ldots,k \tag{75}$$

particiones finitas de Ω , independientes. Supongamos que:

$$p_i^j = \text{Prob}\ [A_{ij}] \tag{76}$$

El espacio paramétrico resultante será:

$$E = \{(p_1^1 \ldots p_1^{n_1-1}, \ldots, p_k^1, \ldots p_k^{n_k-1}) \in \mathbb{R}^{n_1+\ldots+n_k-k} \;/$$

$$p_i^j > 0 \quad \sum_{j=1}^{n_i-1} p_i^j < 1 \quad i=1,\ldots,k\} \tag{77}$$

Procedamos al cálculo del tensor métrico. Consideremos las variables aleatorias X_i, definidas a partir de:

$$X_i(\omega) = j \quad \leftrightarrow \quad \omega \in A_{ij} \qquad i=1,\ldots,k \tag{78}$$

Las variables X_i son estocásticamente independientes ya que:

$$\text{Prob}\left[X_1=j_1 \ldots X_k=j_k\right] = \text{Prob}\left[A_{1j_1} \times \ldots \times A_{kj_k}\right] = p_1^{j_1} \ldots p_k^{j_k} \tag{79}$$

por tanto, la función de densidad conjunta de las mismas será de la forma:

$$f(x^1,\ldots,x^k,p_1^1\ldots p_1^{n_1-1},\ldots,p_k^1\ldots p_k^{n_k-1}) = f_1(x^1,p_1^1\ldots p_1^{n_1-1})\cdot\ldots$$

$$\ldots\cdot f_k(x^k,p_k^1\ldots p_k^{n_k-1}) \tag{80}$$

y estaremos en condiciones de aplicar los resultados 2.3.5 y 2.3.6. La distancia en cada uno de los espacios paramétricos E_i:

$$E_i = \{(p_i^1,\ldots,p_i^{n_i-1}) \in \mathbb{R}^{n_i-1} \;/\; p_i^j > 0 \quad \sum_{j=1}^{n_i-1} p_i^j < 1\} \tag{81}$$

vendrán dadas por:

$$s_i = 2\sqrt{N} \arccos\left(\sum_{j=1}^{n_i} \sqrt{a_i^j b_i^j}\right) \qquad i=1\ldots k \tag{82}$$

Nótese que hemos considerado la experiencia aleatoria de extraer con remplazamiento N individuos de Ω y observar a que elemento de la partición P_i pertenece cada uno. Las a_i^j y las b_i^j son las probabilidades de los sucesos A_{ij} en dos poblaciones distintas A y B.

La distancia en E vendrá dada por:

$$s = 2\sqrt{N}\sqrt{\sum_{i=1}^{k}(\arccos(\sum_{j=1}^{n_i}\sqrt{a_i^j b_i^j}))^2} \tag{83}$$

4.3. CASO DE VARIABLES BINARIAS

Consideremos a continuación el caso (conocido como "caso binario") que las k particiones sean de la forma:

$$P_i = \{A_i, \bar{A}_i\} \qquad i=1,\ldots,k \tag{84}$$

entonces cada E_i es euclídeo y por tanto E es euclídeo. Habrá que hallar una transformación de coordenadas que reduzca al tensor métrico a un tensor constante, basándonos en el resultado 3.2.3, por tanto:

$$y^i = A_i \int \sqrt{g_{ii}}\, dp^i + B_i \qquad i=1,\ldots,k \tag{85}$$

define la transformación buscada. Nótese que hemos llamado p^i a la probabilidad de A_i. El tensor métrico es diagonal, y los elementos g_{ii} vienen dados por:

$$g_{ii} = \frac{N}{p^i(1-p^i)} \qquad i=1,\ldots,k \tag{86}$$

habrá que integrar $\sqrt{g_{ii}}$, obteniendo:

$$\int \sqrt{g_{ii}}\, dp^i = \int \frac{dp^i}{\sqrt{p^i(1-p^i)}} = \int \frac{2dp^i}{\sqrt{1-(1-2p^i)^2}} = \text{arcsen}(2p^i-1) + \text{cte.} \qquad (87)$$

por tanto la transformación:

$$y^i = A_i \text{ arcsen } (2p^i-1) + B_i \qquad (88)$$
$$i=1,\ldots,k$$

reduce al tensor métrico a un tensor constante. Podemos hallar las constantes A_i, B_i $i=1,\ldots,k$ para que el tensor métrico sea la identidad, δ_{ij}. Teniendo en cuenta que:

$$p^i = \frac{1}{2} \text{ sen } \left(\frac{y^i-B_i}{A_i}\right) + \frac{1}{2} \qquad (89)$$

y recordando la forma como se transforman los tensores covariantes al cambiar de coordenadas, podremos escribir:

$$1 = \delta_{\mu\mu} = \frac{\partial p^\mu}{\partial y^\mu} \frac{\partial p^\mu}{\partial y^\mu} g_{\mu\mu} = \frac{1}{4A_\mu^2} \cos^2\left(\frac{y^\mu-B_\mu}{A_\mu}\right) \frac{4N}{1-\text{sen}^2\left(\frac{y^\mu-B_\mu}{A_\mu}\right)} =$$

$$= \frac{1}{4\, A_\mu^2} \cos^2\left(\frac{y_\mu-B_\mu}{A_\mu}\right) \frac{4}{\cos^2\left(\frac{y_\mu-B_\mu}{A_\mu}\right)} = \frac{N}{A_\mu^2} \qquad (90)$$

por tanto $A_\mu = \sqrt{N}$ y B_μ podemos tomarlo igual a cero. Una transformación que reduce al tensor métrico a la identidad es:

$$y^i = \sqrt{N} \text{ arcsen } (2p^i-1) \qquad i=1,\ldots,k \qquad (91)$$

Por tanto, la distancia entre dos puntos de E, de coordenadas (antiguas), $p^1,\ldots,p^k$ y $q^1,\ldots,q^k$, vendrá dada por:

$$s = \sqrt{N \sum_{i=1}^{k} (\mathrm{arcsen}\,(2p^i-1) - \mathrm{arcsen}(2q^i-1))^2} \tag{92}$$

dada la isometría global de E con un subconjunto de un espacio euclídeo.

4.4. DISTRIBUCIONES DISCRETAS FINITAS

Todos estos resultados pueden relacionarse con las distribuciones de probabilidad discretas y finitas, que dependen de un número finito de parámetros:

$$p^i = p^i(\theta^1,\ldots,\theta^k) \qquad i=1,\ldots,n \quad k\leq n \tag{93}$$

con p^i diferenciables, $p^i > 0$ y $\sum_{i=1}^{n} p^i = 1$ y la aplicación $(\theta^1\ldots$ $\ldots\theta^k) \to (p^1\ldots p^n)$ inyectiva. Entonces la métrica definida en el espacio paramétrico

$$E = \{(\theta^1,\ldots,\theta^k)\in \mathbb{R}^k \;/\; p^i > 0 \;\sum_{i=1}^{n} p^i = 1\} \tag{94}$$

a través de la matriz de información de Fisher, es idéntica a la métrica inducida por la métrica euclídea en la variedad:

$$V = \{(x^1,\ldots,x^n)\in\mathbb{R}^n \;/\; x^i = 2\sqrt{p^i(\theta^1\ldots\theta^k)} \;\; i=1,\ldots,n \quad (\theta^1,\ldots,\theta^k)\in E\} \tag{95}$$

En efecto, en V la métrica inducida, usando el convenio de sumación de los índices repartidos, viene dada por:

$$ds^2 = dx^i \cdot dx^i = 4d\sqrt{p^i}\, d\sqrt{p^i} = \frac{1}{\sqrt{p^i}} \left(\frac{\partial p^i}{\partial \theta^\mu} d\theta^\mu \right) \cdot \frac{1}{\sqrt{p^i}} \left(\frac{\partial p^i}{\partial \theta^\nu} d\theta^\nu \right) =$$

$$= \frac{1}{p^i} \frac{\partial p^i}{\partial \theta^\mu} \frac{\partial p^i}{\partial \theta^\nu} d\theta^\mu\, d\theta^\nu \tag{96}$$

luego el tensor métrico en V es:

$$g_{\mu\nu} = \frac{1}{p^i} \frac{\partial p^i}{\partial \theta^\mu} \frac{\partial p^i}{\partial \theta^\nu} \qquad \mu,\nu = 1,\ldots,k \tag{97}$$

En E, el tensor métrico definido a partir de la matriz de información de Fisher es:

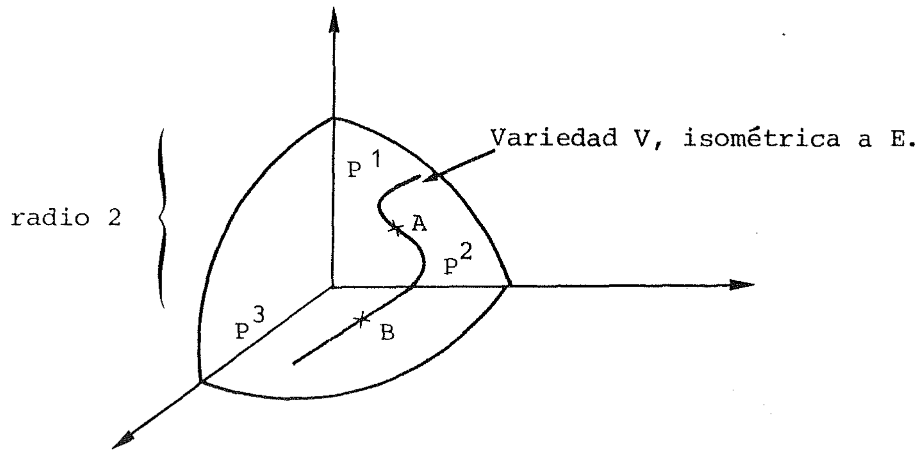
$$g_{\mu\nu} = E\left(\frac{1}{f^2} \frac{\partial f}{\partial \theta^\mu} \frac{\partial f}{\partial \theta^\nu} \right) \qquad \mu,\nu = 1,\ldots,k \tag{98}$$

siendo f la función de densidad, que por ser discreta, podemos escribir:

$$g_{\mu\nu} = \sum_{i=1}^{n} \frac{1}{(p^i)^2} \frac{\partial p^i}{\partial \theta^\mu} \frac{\partial p^i}{\partial \theta^\nu} p^i = \sum_{i=1}^{n} \frac{1}{p^i} \frac{\partial p^i}{\partial \theta^\mu} \frac{\partial p^i}{\partial \theta^\nu} \qquad \mu,\nu=1,\ldots,k \tag{99}$$

expresión idéntica a (97), teniendo en cuenta que en ésta usabamos el convenio de sumación de los índices repetidos. Debido a que $(\theta^1,\ldots,\theta^k) \to (p^1 \ldots p^n)$ es inyectiva, existe una aplicación biyectiva entre E y V tal que el tensor métrico en un punto de E, coincide con el tensor métrico, en V, de la imagen de dicho punto por la citada aplicación biyectiva.

Este resultado se interpreta de la siguientes forma: bajo la hipóte sis anterior, es posible considerar que E es una subvariedad incluida en la superficie de una hiperesfera de radio 2 en un espacio euclídeo. La métrica en E es la métrica euclídea inducida en esta hipersuperficie. Para el caso de que tengamos una distribución trinomial uniparamétrica: n=3 , k=1, resulta:



Un corolario evidente es que, en las condiciones dadas por (93), si A y B son dos poblaciones estadísticas de parámetros $\theta_A^1,\ldots,\theta_A^k$ y $\theta_B^1,\ldots$ $\ldots,\theta_B^k$ entonces la distancia de A a B en E obtenida a través de la matriz de información de Fisher, d_{AB}, verifica:

$$d_{AB} \geq 2 \arccos \left(\sum_{i=1}^{n} \sqrt{p^i(\theta_A^1\ldots\theta_A^k)\cdot p^i(\theta_B^1\ldots\theta_B^k)} \right) \qquad (100)$$

4.5. DISTRIBUCIONES DISCRETAS NUMERABLES

Sea X una variable aleatoria discreta que puede tomar un número infinito numerable de valores con probabilidad estrictamente mayor que cero. El espacio paramétrico vendrá definido por:

$$E = \{(p^1,\dots,p^n,\dots)\in \mathbb{R}^\infty \ / \ p^i > 0 \quad i\in \mathbb{N} \quad \sum_{i=1}^{\infty} p^i = 1\} \tag{101}$$

En estas condiciones al tensor métrico en E, obtenido a partir de la matriz de información de Fisher, viene definido por:

$$g_{\mu\nu} = \delta_{\mu\nu}\,\frac{1}{p^\mu} \qquad \mu,\nu \in \mathbb{N} \tag{102}$$

La longitud de una curva c, de parámetro t, en E vendrá dada por:

$$\int_c \sqrt{\sum_{i=1}^{\infty} \frac{1}{p^i}\,(\dot{p}^i)^2}\; dt \tag{103}$$

donde las $\dot{p}^i$ indican las derivadas de p^i respecto al parámetro t. Para hallar una geodésica que una dos puntos habrá que resolver el problema variacional con ligaduras holónomas:

$$\begin{gathered}\delta \int_a^b \sqrt{\sum_{i=1}^{\infty} \frac{1}{p^i}\,(\dot{p}^i)^2}\; dt = 0 \\ \sum_{i=1}^{\infty} p^i - 1 = 0\end{gathered} \tag{104}$$

Consideremos:

$$G = \Big(\sum_{i=1}^{\infty} \frac{1}{p^i}\,(\dot{p}^1)^2\Big)^{1/2} + \lambda(t)\Big(\sum_{i=1}^{\infty} p^i - 1\Big) \tag{105}$$

y hallemos una extremal de:

$$\int_a^b G\; dt \tag{106}$$

A partir de las ecuaciones de Euler y de la restricción, habrá que determinar las $p^i(t)$ y $\lambda(t)$.

$$\frac{\partial G}{\partial p^k} = -\frac{1}{2}\left(\sum_{i=1}^{\infty} \frac{1}{p^i} (\dot{p}^i)^2\right)^{-1/2} \frac{(\dot{p}^k)^2}{(p^k)^2} + \lambda(t)$$

$$\frac{\partial G}{\partial \dot{p}^k} = \left(\sum_{i=1}^{\infty} \frac{1}{p^i} (\dot{p}^i)^2\right)^{-1/2} \frac{\dot{p}^k}{p^k} \tag{107}$$

Cuando $t=s$, s longitud de la curva, entonces:

$$\sum_{i=1}^{\infty} \frac{1}{p^i} (\dot{p}^i)^2 = 1 \tag{108}$$

condición de que el vector tangente a la curva sea unitario. Por tanto:

$$\frac{\partial G}{\partial p^k} = -\frac{1}{2}\left(\frac{\dot{p}^k}{p^k}\right)^2 + \lambda(s) \qquad \frac{\partial G}{\partial \dot{p}^k} = \frac{\dot{p}^k}{p^k} \tag{109}$$

luego resulta:

$$\frac{d}{ds}\left(\frac{\partial G}{\partial \dot{p}^k}\right) = -\frac{1}{(p^k)^2}(\dot{p}^k)^2 + \frac{\ddot{p}^k}{p^k} \tag{110}$$

por tanto las ecuaciones de Euler son finalmente:

$$-\frac{1}{2}\left(\frac{\dot{p}^k}{p^k}\right)^2 + \lambda(s) + \left(\frac{\dot{p}^k}{p^k}\right)^2 - \frac{\ddot{p}^k}{p^k} = 0 \qquad k \in \mathbb{N} \tag{111}$$

que hay que resolver con la restricción:

$$\sum_{i=1}^{\infty} p^i - 1 = 0 \tag{112}$$

Efectuando el cambio de variable $(u^k)^2 = p^k$, resulta:

$$\dot{p}^k = 2u^k\dot{u}^k \qquad \ddot{p}^k = 2\dot{u}^k\dot{u}^k + 2u^k\ddot{u}^k \tag{113}$$

sustituyendo:

$$2\left(\frac{\dot{u}^k}{u^k}\right)^2 - 2\left(\frac{\dot{u}^k}{u^k}\right)^2 - 2\,\frac{\ddot{u}^k}{u^k} + \lambda(s) = -\,\frac{2\ddot{u}^k}{u^k} + \lambda(s) = 0 \tag{114}$$

por tanto resulta el sistema:

$$\begin{aligned} &\ddot{u}^k - \frac{1}{2}\lambda(s)\,u^k = 0 \qquad k \in N \\ &\sum_{i\in \mathbb{N}} (u^i)^2 = 1 \end{aligned} \tag{115}$$

Observemos que de la última ecuación de (115) obtenemos:

$$\sum_{i\in \mathbb{N}} u^i \cdot \dot{u}^i = 0 \tag{116}$$

Además, la condición (108) se convierte en:

$$\sum_{i\in \mathbb{N}} \frac{1}{(u^i)^2}\,(2u^i \cdot \dot{u}^i)^2 = 4 \sum_{i\in \mathbb{N}} (\dot{u}^i)^2 = 1 \tag{117}$$

y si derivamos (116) resulta:

$$\sum_{i\in \mathbb{N}} \dot{u}^i\dot{u}^i + \sum_{i\in \mathbb{N}} u^i\ddot{u}^i = 0 \tag{118}$$

por lo que:

$$\sum_{i\in\mathbb{N}} u^i \ddot{u}^i = -\frac{1}{4} \qquad (119)$$

sumando las ecuaciones del sistema (115) (excepto la restricción), multiplicadas cada una por u^k, obtendremos:

$$\sum_{k\in\mathbb{N}} u^k \ddot{u}^k - \frac{1}{2}\lambda(s) \sum_{k\in\mathbb{N}} (u^k)^2 = 0 \qquad (120)$$

igual a:

$$-\frac{1}{4} - \frac{1}{2}\lambda(s) = 0 \qquad (121)$$

por lo que:

$$\lambda(s) = -\frac{1}{2} \qquad (122)$$

por tanto resulta el sistema:

$$\ddot{u}^k + \frac{1}{4} u^k = 0 \qquad k\in\mathbb{N} \qquad (123)$$

sistema lineal de segundo orden, con raices complejas $\frac{1}{2}i$, $-\frac{1}{2}i$, cuya solución general es:

$$u^k = A_k \cos\frac{1}{2}s + B_k \operatorname{sen}\frac{1}{2}s \qquad k\in\mathbb{N} \qquad (124)$$

equivalente a:

$$p^k = (A_k \cos\frac{1}{2}s + B_k \operatorname{sen}\frac{1}{2}s)^2 \qquad k\in\mathbb{N} \qquad (125)$$

habrá que determinar A_k y B_k para que para $s=0$, $p^k=a^k$ $k\in\mathbb{N}$ y para un cierto s, $p^k=b^k$ $k\in\mathbb{N}$.

Puede demostrarse, análogamente el caso de una multinomial con N=1, que:

$$A_k = \sqrt{a^k} \qquad B_k = (\sqrt{b^k} - \sqrt{a^k} \cos \tfrac{1}{2} s)/\mathrm{sen}\ \tfrac{1}{2} s \qquad k \in \mathbb{N} \tag{126}$$

y finalmente:

$$s = 2 \arccos \left(\sum_{k \in \mathbb{N}} \sqrt{a^k b^k}\right) \tag{127}$$

Equivale a representar a las poblaciones estadísticas como puntos en una hiperesfera de radio 2 en $\mathbb{R}^{\infty}$. En efecto, si $a=(a^1,\ldots,a^n,\ldots)\in$ $\in E$ y $b=(b^1,\ldots,b^n,\ldots)\in E$ se representarían en $\mathbb{R}^{\infty}$ por puntos de coordenadas: $(2\sqrt{a^1},\ldots,2\sqrt{a^n},\ldots) \in \mathbb{R}^{\infty}$ y $(2\sqrt{b^1},\ldots,2\sqrt{b^n},\ldots)\in \mathbb{R}^{\infty}$, ambos puntos en S_2^{∞}, definido por:

$$S_2^{\infty} = \{(x^1,\ldots,x^n,\ldots)\in \mathbb{R}^{\infty} / \sum_{i=1}^{\infty} (x^i)^2 = 4\} \tag{128}$$

con la distancia en S_2^{∞} inducida por la "distancia" euclídea en $\mathbb{R}^{\infty}$:

$$d(x,y) = \sqrt{\sum_{k \in \mathbb{N}} (x^k - y^k)^2} \tag{129}$$

supuesta existente.

4.6. DISTRIBUCIONES DE POISSON INDEPENDIENTES

Sean $X_1,\ldots,X_n$ variables aleatorias discretas cuya función de densidad conjunta viene definida por:

$$f(x^1,\ldots,x^n,\lambda^1,\ldots,\lambda^n) = e^{-\lambda^1}\frac{(\lambda^1)^{x^1}}{x^1!}\cdots e^{-\lambda^n}\frac{(\lambda^n)^{x^n}}{x^n!} \qquad (130)$$

$$\lambda^1,\ldots,\lambda^n \in \mathbb{R}^+ \qquad x^1,\ldots,x^n \in \mathbb{N}\cup\{0\}$$

Nótese que cada X^i se distribuye según una distribución de Poisson de parámetro λ^i y que las X^i son variables aleatorias estocástica-- mente independientes. El espacio paramétrico vendrá dado por:

$$E = \{(\lambda^1,\ldots,\lambda^n)\in \mathbb{R}^n / \lambda^i > 0\} \qquad i=1,\ldots,n \qquad (131)$$

El tensor métrico viene definido por:

$$g_{\mu\nu} = \frac{\delta_{\mu\nu}}{\lambda^\mu} \qquad \mu,\nu = 1,\ldots,n \qquad (132)$$

como es fácil comprobar. Además, a partir del resultado 2.3.3 podemos afirmar que la variedad Riemanniana E es euclídea, y que hallaremos una transformación normalizante resolviendo:

$$y^i = A_i \int \sqrt{g_{ii}}\, d\lambda^i + B_i \qquad i=1,\ldots,n \qquad (133)$$

integrando obtenemos:

$$y^i = A_i \frac{1}{2} \sqrt{\lambda^i} + B_i \qquad i=1,\ldots,n \qquad (134)$$

donde vamos a determinar las A_i y B_i para que el tensor métrico no só- lo resulte constante sino que sea igual a la identidad. De (134) se si- gue:

$$\lambda^i = \frac{4(y^i - B_i)^2}{A_i^2} \tag{135}$$

por tanto:

$$1 = \delta_{\mu\mu} = \left(\frac{\partial\lambda^\mu}{\partial y^\mu}\right)^2 \cdot g_{\mu\mu} = 64\,\frac{(y^i - B_i)^2}{A_i^4} \cdot \frac{1}{\frac{4(y^i - B_i)^2}{A_i^2}} = \frac{16}{A_i^2} \tag{136}$$

por tanto habrá que elegir $A_i = 4$, y B_i arbitrario, por ejemplo $B_i = 0$. Por tanto la transformación buscada vendrá dada por:

$$y_i = 2\sqrt{\lambda^i} \qquad i=1,\ldots,n \tag{137}$$

la variedad E es globalmente isométrica a un semiespacio euclídeo y por tanto, la distancia entre dos poblaciones estadísticas de parámetros $\lambda^1,\ldots,\lambda^n$ y $\mu^1,\ldots,\mu^n$ será

$$d = 2\sqrt{\sum_{i=1}^{n} (\sqrt{\lambda^i} - \sqrt{\mu^i})^2} \tag{138}$$

4.7. DISTRIBUCION MULTINOMIAL COMPUESTA CON UNA POISSON

Consideremos a continuación el caso que N sea una variable aleatoria que sigue una distribución de Poisson de parámetro λ y sean $(Z_1,\ldots,Z_k)$ una variable aleatoria multivariante que sigue una multinomial (un ensayo) de probabilidades $p^1,\ldots,p^k$. Sea:

$$(X_1,\dots,X_k) = \sum_{j=1}^{N} (Z_1,\dots,Z_k)_j \tag{139}$$

donde la $(X_1,\dots,X_k)$ es la suma de variables aleatorias $(Z_1,\dots,Z_k)_j$ independientes, en número aleatorio N. La función de densidad conjunta de $X_1,\dots,X_k$ viene dada por:

$$P[X_1=n_1,\dots,X_k=n_k] = P[N=n_1+\dots+n_k]\cdot P[X_1=n_1,\dots,X_k=n_k/N=n_1+\dots+n_k] =$$

$$= e^{-\lambda}\,\frac{\lambda^{n_1+\dots+n_k}}{(n_1+\dots+n_k)!}\cdot\frac{(n_1+\dots+n_k)!}{n_1!\cdot\dots\cdot n_k!}\quad (p^1)^{n_1}\cdot\dots\cdot(p^k)^{n_k} =$$

$$= \prod_{j=1}^{k} e^{-\lambda p^j}\,\frac{(\lambda p^j)^{n_j}}{n_j!} \tag{140}$$

que es el producto de k variables aleatorias con distribución de Poisson, estocásticamente independientes, de parámetros $\lambda p^1,\dots,\lambda p^k$. Reparametrizando, llamando $\theta^i = \lambda p^i$, tendremos que la distancia entre dos poblaciones estadísticas de parámetros $(\theta_A^1,\dots,\theta_A^k)$ y $(\theta_B^1,\dots,\theta_B^k)$ viene dada por:

$$d(A,B) = 2\sqrt{\sum_{i=1}^{k} (\sqrt{\theta_A^i} - \sqrt{\theta_B^i})^2} \tag{141}$$

4.8. DISTRIBUCION MULTINOMIAL NEGATIVA

Sean $X_1,\dots,X_k$ variables aleatorias discretas cuya función de densidad conjunta viene definida por:

$$P[X_1=n_1,\ldots,X_k=n_k] = \frac{(n_1+\ldots+n_k+r-1)!}{n_1!\cdot\ldots\cdot n_k!(r-1)!}(p^1)^{n_1}\cdot\ldots\cdot(p^k)^{n_k}(1-p^1-\ldots-p^k)^r$$

$$n_1,\ldots,n_k,r \in \mathbb{N} \qquad p^i > 0 \qquad \sum_{i=1}^{k} p^i < 1 \tag{142}$$

Una situación que puede ser descrita por la multinomial negativa es la repetición de una experiencia aleatoria que admite k+1 resultados posibles, hasta la aparición del resultado k+1 ésimo r veces.

El espacio paramétrico asociado a dicha distribución es:

$$E = \{(p^1,\ldots p^k)\in \mathbb{R}^k \,/\, p^i > 0 \quad i=1,\ldots,k \quad \sum_{i=1}^{k} p^i < 1\} \tag{143}$$

El siguiente paso es hallar el tensor métrico:

$$\ln f = \ln \frac{(n_1+\ldots+n_k+r-1)!}{n_1!\cdot\ldots\cdot n_k!(r-1)!} + \sum_{i=1}^{k} n_i \ln p^i + r\ln(1-p^1-\ldots-p^k) \tag{144}$$

derivando respecto p^μ:

$$\frac{\partial \ln f}{\partial p^\mu} = \frac{n_\mu}{p^\mu} - \frac{r}{1-p^1-\ldots-p^k} \qquad \mu=1,\ldots k \tag{145}$$

derivando ahora respecto p^ν, obtenemos:

$$\frac{\partial^2 \ln f}{\partial p^\nu\,\partial p^\mu} = \frac{\partial}{\partial p^\nu}\left(\frac{n_\mu}{p^\mu} - \frac{r}{1-p^1-\ldots-p^k}\right) = -\frac{n_\mu}{(p^\mu)^2}\delta_{\mu\nu} - \frac{r}{(1-p^1-\ldots-p^k)^2}$$

$$\mu,\nu=1,\ldots,k \tag{146}$$

hallando la esperanza y cambiando de signo, como:

$$E(n_\mu) = \frac{rp^\mu}{1-p^1-\ldots-p^k} \tag{147}$$

resulta:

$$g_{\mu\nu} = \frac{r}{1-p^1-\ldots-p^k}\left[\frac{1}{p^\mu}\delta_{\mu\nu} + \frac{1}{1-p^1-\ldots-p^k}\right] \qquad (148)$$

$$\mu,\nu=1,\ldots,k$$

El recíproco del tensor métrico $g^{\mu\nu}$ será pues:

$$g^{\mu\nu} = \frac{1-p^1-\ldots-p^k}{r}\left[\delta^{\mu\nu} p^\mu - p^\mu p^\nu\right] \qquad (149)$$

$$\mu,\nu=1,\ldots,k$$

Procedamos a calcular los símbolos de Christoffel. Previamente:

$$\frac{\partial g_{in}}{\partial p^j} = \frac{r}{1-p^1-\ldots-p^k}\left(-\frac{1}{(p^i)^2}\delta_{ijn} + \frac{1}{(1-p^1-\ldots-p^k)^2}\right) +$$

$$+ \frac{r}{(1-p^1-\ldots-p^k)^2}\left(\frac{1}{p^i}\delta_{in} + \frac{1}{1-p^1-\ldots-p^k}\right) \qquad i,j,n=1,\ldots,k \qquad (150)$$

que puede escribirse como

$$\frac{\partial g_{in}}{\partial p^j} = \frac{r}{(1-p^1-\ldots-p^k)}\left(-\frac{1}{(p^i)^2}\delta_{ijn} + \frac{2}{(1-p^1-\ldots-p^k)^2}\right) +$$

$$+ \frac{r}{(1-p^1-\ldots-p^k)^2}\frac{1}{p^i}\delta_{in} \qquad i,j,n=1,\ldots,k \qquad (155)$$

por tanto, los símbolos de Christoffel de primera especie seran: