SOBRE UNA GENERALIZACION DE LAS CURVAS DE PEARSON AL CASO BIDIMENSIONAL

SEBASTIAN NAVARRO SAGRISTA

INTRODUCCIÓN

El presente trabajo tiene por objeto la obtención de superficies de probabilidad en forma correlativa a las curvas que obtuvo Pearson mediante integración de su ecuación diferencial.

Así planteado el problema en toda su generalidad, se comprende que la solución es mucho más difícil y compleja, ya que, en lugar de una ecuación diferencial de primer orden, necesitamos resolver un sistema de ecuaciones en derivadas parciales, limitándonos, por tanto, a estudiar algunos casos particulares.

Así como la ecuación diferencial de Pearson, para la obtención de las curvas que llevan su nombre, se puede obtener mediante generalización de la ecuación diferencial de la curva normal, de la misma manera, mediante generalización del sistema de ecuaciones en derivadas parciales a que satisface la superficie normal, se pueden obtener otras tantas superficies de probabilidad.

Mediante una primera generalización obtenemos la ecuación

$$z = k\left(1 - \frac{x^2}{a^2} - \frac{y^2}{b^2} + \frac{2\,rxy}{ab}\right)^n \qquad n+1>0, \qquad r^2<1$$

De esta ecuación no se puede obtener, como caso particular, la superficie normal, pero las líneas de regresión son rectilíneas, el cieficiente r, es precisamente el coeficiente de correlación lineal, y la línea de regresión de la desviación típica condicional es una elipse, y, finalmente, la kurtosis, tanto de la x como de la y, sólo depende de n.

Mediante una segunda generalización, obtenemos la ecuación

$$z = ke^{-\frac{ax^2+2bxy+cy^2}{2}}(h^2+ax^2+2\,bxy+cy^2)^n$$

con $n+1>0$, si $h=0$.

Esta ecuación tiene, como caso particular, la superficie normal cuando $n=0$. Las líneas de regresión son rectilíneas, y tanto éstas como el coeficiente de correlación son independientes de h y n, y tienen la misma expresión, como en el caso de la superficie normal, y la kurtosis de la x o de la y sólo depende de n.

La última generalización la hemos hecho por un procedimiento distinto a los dos anteriores, y de esta forma obtenemos una superficie que es caso particular de la superficie normal, para valores convenientes de los parámetros.

Este método es general para la obtención de otras superficies de probabilidad, teniendo como meta el poder sistematizarlo en forma análoga a la obtenida para las curvas de Pearson.

La ecuación de la superficie normal es

$$z = ke^{-\frac{\alpha x^2 + 2\beta xy + \gamma y^2}{2}}$$

de la que se deduce

$$\frac{\partial z}{\partial x} = -z(\alpha x + \beta y)$$

$$\frac{\partial z}{\partial y} = -z(\beta x + \gamma y)$$

o sea

$$-\frac{1}{z}dz = (\alpha x + \beta y)\,dx + (\beta x + \gamma y)\,dy$$

y, recíprocamente, la superficie normal es la única solución de esta ecuación, pues resolviéndola se tiene que para y constante

$$-\frac{dz}{z} = (\alpha x + \beta y)\,dx, \qquad -lz = \frac{\alpha}{2}x^2 + \beta xy + \Phi(y)$$

y derivando respecto y

$$-\frac{1}{z}\frac{\partial z}{\partial y} = \beta x + \Phi'(y), \qquad \beta x + \gamma y = \beta x + \Phi'(y), \qquad \gamma y = \Phi'(y)$$

$$\Phi(y) = \frac{\gamma}{2}y^2 + k_1, \qquad -lz = \frac{\alpha}{2}x^2 + \beta xy + \frac{\gamma}{2}y^2 + k_1$$

es decir,

$$z = ke^{-\frac{\alpha x^2 + 2\beta xy + \gamma y^2}{2}}$$

Una primera generalización del sistema anterior puede ser

$$\frac{\partial z}{\partial x} = - z\, f(xy)$$

$$\frac{\partial z}{\partial y} = - z\, \varphi(xy)$$

es decir,

$$- \frac{1}{z}\, dz = f(xy)\, dx + \varphi(xy\, dy$$

cuyo segundo miembro debe cumplir la condición de integrabilidad que es $f'_y = \varphi'_x$.

Así, para $f(x, y)$ y $\varphi(x, y)$, polinomios homogéneos de segundo grado, y con la condición de integrabilidad

$$\frac{\partial z}{\partial x} = - z(\alpha x^2 + 2\beta xy + \gamma y^2)$$

$$\frac{\partial z}{\partial y} = - z(\beta x^2 + 2\gamma xy + \delta y^2)$$

o la ecuación equivalente

$$- \frac{dz}{z} = (\alpha x^2 + 2\beta xy + \gamma y^2)\, dx + (\beta x^2 + 2\gamma xy + \delta y^2)\, dy$$

y resolviéndola, como en el caso anterior, resulta

$$z = k e^{-\frac{\alpha x^3 + 3xy(\beta x + \gamma y) + \delta y^3}{-\ 3}}$$

y, en general, si $f(xy)$ y $\varphi(xy)$ son polinomios homogéneos que cumplen la condición de integrabilidad, la solución es una superficie de tipo exponencial de exponente un polinomio también homogéneo y de un grado superior.

PRIMERA GENERALIZACION

Consideremos como una primera generalización del sistema en derivadas parciales al que satisface la superficie normal, el siguiente:

$$\left.\begin{aligned} \frac{\partial z}{\partial x} &= - z^m f(xy) \\ \frac{\partial x}{\partial y} &= - z^m \varphi(xy) \end{aligned}\right\} m \neq 1, \qquad f'_y = \varphi'_x$$

Hagamos $m=1-\frac{1}{n}$; $f(xy)$ y $\varphi(xy)$ funciones lineales, o sea

$$\frac{\partial z}{\partial x}=-z^{1-\frac{1}{n}}(\alpha x+\beta y)$$

$$\frac{\partial z}{\partial y}=-z^{1-\frac{1}{n}}(\beta x+\gamma y)$$

que es equivalente a la ecuación

$$-z^{-1+\frac{1}{n}}dz=(\alpha x+\beta y)\,dx+(\beta x+\gamma y)\,dy$$

que integrada, da para y constante

$$-z^{\frac{1}{n}-1}dz=(\alpha x+\beta y)\,dx, \qquad -nz^{\frac{1}{n}}=\frac{\alpha}{2}x^2+\beta xy+\Phi(y)$$

y derivando respecto y

$$\beta x+\gamma y=\beta x+\Phi'(y), \qquad \Phi'(y)=\frac{\gamma}{2}y^2-k_1$$

$$-nz^{\frac{1}{n}}=\frac{\alpha}{2}x^2+\beta xy+\frac{\gamma}{2}y^2-k_1, \qquad z=\frac{1}{n^n}\left(-\frac{\alpha}{2}x^2-\beta xy-\frac{\gamma}{2}y^2+k_1\right)^n$$

que pondremos en la forma

$$z=f(xy)=k\left(1-\frac{x^2}{a^2}-\frac{y^2}{b^2}+\frac{2\,rxy}{ab}\right)^n$$

Esta ecuación es la que vamos a considerar para nuestro primer estudio.

Analicemos, con toda detalle, la función de densidad dada con la ecuación anterior, y considerar luego los casos particulares que de ella se derivan.

Veamos cuáles son las condiciones que debe cumplir $f(xy)$ para que pueda representar una función de densidad.

Supondremos que $\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{2\,rxy}{ab}$ es una forma cuadrática positiva, es decir $r^2<1$, y, por tanto, la cónica

$$\text{(C)} \qquad \varphi(xy)=\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{2\,rxy}{ab}-1=0$$

representa una elipse de centro el origen.

La región de los puntos (xy), tales que es $z \geqslant 0$, viene dada por

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{2\,rxy}{ab} < 1$$

los cuales son los puntos interiores a dicha elipse.

Por otra parte, para que pueda representar una función de densidad debe ser

$$\int\!\!\int_{(c)} f(xy)\,dx\,dy = 1$$

Sean las elipses

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{2\,rxy}{ab} = 1 - u$$

y considerando como áreas elementales las de dichas elipses, es

$$S = \frac{\pi\,ab\,(1-u)}{\sqrt{1-r^2}}$$

luego

$$1 = k\int_1^0 u^n\,ds = \frac{\pi\,k\,ab}{\sqrt{1-r^2}}\int_0^1 u^n\,du$$

la cual, para que sea convergente, debe ser $n+1>0$

$$1 = \frac{\pi\,kab}{(n+1)\sqrt{1-r^2}}, \qquad k = \frac{(n+1)\sqrt{1-r^2}}{\pi\,ab}$$

Por tanto, llegamos a la conclusión de que la función

$$z = k\left(1 - \frac{x^2}{a^2} - \frac{y^2}{b^2} + \frac{2\,rxy}{ab}\right)^n$$

es una función de densidad, siempre que

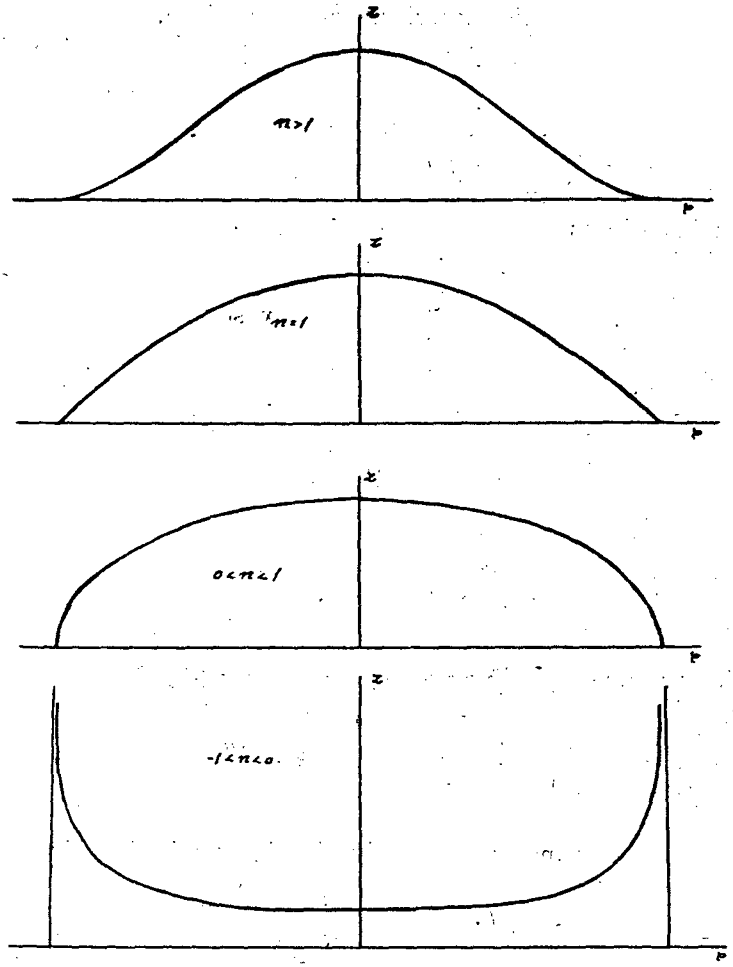
$$n+1>0 \qquad \text{con} \qquad r^2<1$$

y siendo

$$k = \frac{(n+1)\sqrt{1-r^2}}{\pi\,ab} \qquad \text{y} \qquad \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{2\,rxy}{ab} < 1$$

Forma de la superficie.—Cortando la superficie

$$z = k\left(1 - \frac{x^2}{a^2} - \frac{y^2}{b^2} + \frac{2\,rxy}{ab}\right)^n, \qquad r^2 < 1, \qquad n + 1 > 0$$



por un plano paralelo al xy, $z=\lambda$, $0<\lambda<k$ se tiene

$$\frac{x^2}{a^2\left(1-\dfrac{\lambda}{k}\right)^{\frac{1}{n}}} + \frac{y^2}{b^2\left(1-\dfrac{\lambda}{k}\right)^{\frac{1}{n}}} - \frac{2\,rxy}{ab\left(1-\dfrac{\lambda}{k}\right)^{\frac{1}{n}}} = 1$$

que son elipses homotéticas y de centro en el eje z.

Para $n>0$ si λ aumenta los ejes disminuyen.

Para $-1<n<0$ si λ aumenta los ejes aumentan.

Veamos la intersección con el plano zx.

Se obtiene la curva

$$z = k\left(1 - \frac{x^2}{a^2}\right)^n$$

$$z' = -\frac{2\,kn}{a^2}\left(1 - \frac{x^2}{a^2}\right)^{n-1} x, \qquad z'' = \frac{2\,kn}{a^2}\left(1 - \frac{x^2}{a^2}\right)^{n-2}\left(\frac{2\,n-1}{a^2}\,x^2 - 1\right)$$

Cortando por un plano que pase por el eje z y formando el ángulo α con el plano zx, y tomando como eje x' la traza de este plano con el xy, se tiene

$$x = x' \cos\alpha, \qquad y = x' \operatorname{sen}\alpha$$

luego

$$z = k\left(1 - \frac{b^2\cos^2\alpha + a^2\operatorname{sen}^2\alpha - 2\,rab\operatorname{sen}\alpha\cos\alpha}{a^2 b^2}\,x'^2\right)^n$$

cuyo numerador es positivo ya que se puede poner en la forma

$$(a\operatorname{sen}\alpha - br\cos\alpha)^2 + b^2(1 - r^2)\cos^2\alpha$$

luego esta curva tiene la misma forma que las vistas anteriormente.

Cálculo de los momentos.—Por la simetría de la elipse, respecto del origen, se deduce que éste es el centro de gravedad.

El momento central de orden $f+g$ es

$$\mu_{f,g} = \int\int_{(C)} f(xy)\, x^f y^g\, dx\, dy$$

y, por tanto, los momentos de orden impar son nulos.

Pongamos la ecuación de la función de densidad en la forma

$$f(xy) = k\left[\left(1 - \frac{1-r^2}{a^2}\,x^2\right) - \left(\frac{y}{b} - \frac{rx}{a}\right)^2\right]^n$$

Sean y_1 e y_2 los valores de y deducidos de la ecuación de la elipse (C) para x constante, es decir,

$$\frac{y_2}{b} - \frac{rx}{a} = +\sqrt{1 - \frac{1-r^2}{a^2}\,x^2}, \qquad \frac{y_1}{b} - \frac{rx}{a} = -\sqrt{1 - \frac{1-r^2}{a^2}\,x^2}$$

El campo de variación de la x tiene por extremos los puntos de tangente paralela al eje y, o sean los puntos tales que $\frac{\partial \varphi}{\partial y}=0$, lo que nos da $x=\pm \frac{a}{\sqrt{1-r^2}}$ y, por tanto,

$$\mu_{f,g}=k\int_{-\frac{a}{\sqrt{1-r^2}}}^{\frac{a}{\sqrt{1-r^2}}} x^f dx \int_{y_1}^{y_2} y^g \left[\left(1-\frac{1-r^2}{a^2}x^2\right)-\left(\frac{y}{b}-\frac{rx}{a}\right)^2\right]^n dy$$

Haciendo el cambio de variables $\frac{y}{b}-\frac{rx}{a}=\sqrt{1-\frac{1-r^2}{a^2}x^2}$

$$\mu_{f,g}=kb^{g+1}\int_{-\frac{a}{\sqrt{1-r^2}}}^{\frac{a}{\sqrt{1-r^2}}} x^f\left(1-\frac{1-r^2}{a^2}x^2\right)^{n+\frac{1}{2}} dx\int_{-1}^{1}\left[\frac{rx}{a}+\rho\sqrt{1-\frac{1-r^2}{a^2}x^2}\right]^g (1-\rho^2)^n d\rho$$

pero

$$\left[\frac{rx}{a}+\rho\sqrt{1-\frac{1-r^2}{a^2}x^2}\right]^g=\sum_{\nu=0}^{g}\binom{g}{\nu}\left(\frac{rx}{a}\right)^{g-\nu}\rho^\nu\left(1-\frac{1-r^2}{a^2}x^2\right.$$

Para ν impar la última integral es nula y, por tanto,

$$\mu_{f,g}=kb^{g+1}\int_{-\frac{a}{\sqrt{1-r}}}^{\frac{a}{\sqrt{1-r}}} x^f\left(1-\frac{1-r^2}{a^2}x^2\right)^{n+\frac{1}{2}} dx\sum_{0}^{g/2}\binom{g}{2\nu}\left(\frac{rx}{a}\right)^{g-2\nu}\left(1-\frac{1-r^2}{a^2}x^2\right)^\nu\int_{-1}^{1}\rho^{2\nu}\left(1-\rho^2\right)^n d\rho$$

La última integral, con el cambio $\rho^2=u$, es $B\left(\nu+\frac{1}{2},\ n+1\right)$ o sea la integral Euleriana de primera especie

$$\mu_{f,g}=\frac{2\,kb^{g+1}r^g}{a^g}\sum_{0}^{g/2}\binom{g}{2\nu}\left(\frac{r}{a}\right)^{-2\nu}B\left(\nu+\frac{1}{2},\,n+1\right)\int_{0}^{\frac{a}{\sqrt{1-r^2}}} x^{f+g-2\nu}\left(1-\frac{1-r^2}{a^2}x^2\right)^{n+\nu+\frac{1}{2}}dx$$

que con el cambio $\frac{1-r^2}{a^2}x^2=t$

$$\mu_{f,g}=\frac{ka^{f+1}b^{g+1}r^g}{(1-r^2)^{\frac{f+g+1}{2}}}\sum_{0}^{g/2}\binom{g}{2\nu}r^{-2\nu}(1-r^2)^\nu B\left(\nu+\frac{1}{2},\,n+1\right)\int_{0}^{1} t^{\frac{f+g-2\nu-1}{2}}(1-t)^{n+\nu+\frac{1}{2}}dt$$

$$=\frac{ka^{f+1}b^{g+1}r^g}{(1-r^2)^{\frac{f+g+1}{2}}}\sum_{0}^{g/2}\binom{g}{2\nu}r^{-2\nu}(1-r^2)^\nu B\left(\nu+\frac{1}{2},\,n+1\right)B\left(\frac{f+g-2\nu+1}{2},\,n+\nu+\frac{3}{2}\right)$$

y, finalmente, en función de la Euleriana de segunda especie

$$\mu_{f,g} = \frac{k a^{f+1} b^{g+1}}{(1-r^2)^{\frac{f+g+1}{2}}} \frac{\Gamma(n+1)}{\Gamma\left(\frac{f+g}{2}+n+2\right)} \sum_0^{g/2} \binom{g}{2\nu} r^{g-2\nu} (1-r^2)^\nu \, \Gamma\left(\nu+\frac{1}{2}\right) \Gamma\left(\frac{f+g-2\nu+1}{2}\right)$$

y por ser $n+1>0$ se deduce que existen todos los momentos de orden par. Sustituyendo el valor de k resulta

$$\mu_{f,g} = \frac{a^f b^g}{\pi (1-r^2)^{\frac{f+g}{2}}} \frac{\Gamma(n+2)}{\Gamma\left(\frac{f+g}{2}+n+2\right)} \sum_0^{g/2} \binom{g}{2\nu} r^{g-2\nu} (1-r^2)^\nu \, \Gamma\left(\nu+\frac{1}{2}\right) \Gamma\left(\frac{f+g-2\nu+1}{2}\right)$$

De esta fórmula se deduce

$$\mu_{2,0} = \quad = \frac{a^2}{2(1-r^2)(n+2)}, \qquad \mu_{0,2} = \sigma_2^2 = \frac{b^2}{2(1-r^2)(n+2)}$$

$$\mu_{1,1} = \frac{abr}{2(1-r^2)(n+2)}, \qquad \mu_{4,0} = \frac{3a^4}{4(1-r^2)^2(n+2)(n+3)}$$

Entre los momentos se verifica la relación $\dfrac{\mu_{f,1}}{\mu_{f+1,0}} = \dfrac{\mu_{1,1}}{\mu_{2,0}}$.
En efecto

$$\mu_{f,1} = \frac{a^f b}{\pi(1-r^2)^{\frac{f+1}{2}}} \frac{\Gamma(n+2)}{\Gamma\left(\frac{f+1}{2}+n+2\right)} r\,\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{f+2}{2}\right)$$

$$\mu_{f+1,0} = \frac{a^{f+1}}{\pi(1-r^2)^{\frac{f+1}{2}}} \frac{\Gamma(n+2)}{\Gamma\left(\frac{f+1}{2}+n+2\right)} \Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{f+2}{2}\right)$$

y por división $\dfrac{\mu_{f,1}}{\mu_{f+1,0}} = \dfrac{br}{a}$ y por ser independiente de f debe ser

$$\frac{\mu_{f,1}}{\mu_{f+1,0}} = \frac{\mu_{1,1}}{\mu_{2,0}}$$

Coeficiente de correlación.—De las fórmulas obtenidas anteriormente para los primeros momentos se deduce que

$$r = \frac{\mu_{1,1}}{\sqrt{\mu_{2,0}\,\mu_{0,2}}}$$

es decir, el coeficiente r de la ecuación es precisamente el coeficiente de correlación lineal.

Función marginal.—La función de frecuencia marginal de la x es

$$f_1(x) = \int_{y_1}^{y_2} f(xy)\, dy$$

que, con el cambio de variables visto anteriormente para la obtención de los momentos, resulta

$$f_1(x) = k \int_{y_1}^{y_2} \left[\left(1 - \frac{1-r^2}{a^2} x^2\right) - \left(\frac{y}{b} - \frac{rx}{a}\right)^2\right]^n dy =$$

$$= kb\left(1 - \frac{1-r^2}{a^2} x^2\right)^{n+\frac{1}{2}} \int_{-1}^{1} (1-\rho^2)^n\, d\rho$$

$$f_1(x) = kb\, \mathrm{B}\left(\frac{1}{2},\quad n+1\right) \cdot \left(1 - \frac{1-r^2}{a^2} x^2\right)^{n+\frac{1}{2}}$$

y finalmente

$$f_1(x) = kb\sqrt{\pi}\, \frac{\Gamma(n+1)}{\Gamma\left(n+\frac{3}{2}\right)} \left(1 - \frac{1-r^2}{a^2} x^2\right)^{n+\frac{1}{2}}$$

de rango $-\dfrac{a}{\sqrt{1-r^2}}$, $\dfrac{a}{\sqrt{1-r^2}}$.

Análogamente, la función marginal de la y sería

$$f_2(y) = ka\sqrt{\pi}\, \frac{\Gamma(n+1)}{\Gamma\left(n+\frac{3}{2}\right)} \left(1 - \frac{1-r^2}{b^2} y^2\right)^{n+\frac{1}{2}}$$

de rango $-\dfrac{b}{\sqrt{1-r^2}}$, $\dfrac{b}{\sqrt{1-r^2}}$.

El centro de gravedad es el origen, y los momentos impares nulos. El coeficiente de asimetría es nulo y el de exceso es

$$\gamma = \frac{\mu_{4,0}}{\mu_{2,0}^2} - 3 = 3\,\frac{n+2}{n+3} - 3 = \frac{-3}{n+3} < 0$$

que sólo depende de n y teniendo el mismo valor para el de la x y de la y.

De esta fórmula se deduce que $n+1=\frac{-2\gamma-3}{\gamma}$ y por ser $n+1>0$ se debe cumplir la condición

$$-1,5<\gamma<0$$

Función condicional.—La función de frecuencia condicional de y sobre x es

$$f(y/x)=\frac{f(xy)}{f_1(x)}$$

El centro de gravedad está dado por la fórmula

$$m_1(x)=\frac{\int_{y_1}^{y_2} y\,f(xy)\,dy}{f_1(x)}$$

cuyo numerador es

$$\int_{y_1}^{y_2} y\,f(xy)\,dy = k\int_{y_1}^{y_2} y\left[\left(1-\frac{1-r^2}{a^2}x^2\right)-\left(\frac{y}{b}-\frac{rx}{a}\right)^2\right]^n dy$$

y con el cambio de variables ya utilizado

$$\frac{y}{b}-\frac{rx}{a}=\rho\sqrt{1-\frac{1-r^2}{a^2}x^2}$$

se tiene

$$=kb^2\int_{-1}^{1}\left(\frac{rx}{a}+\rho\sqrt{1-\frac{1-r^2}{a^2}x^2}\right)\left(1-\frac{1-r^2}{a^2}x^2\right)^{n+\frac{1}{2}}(1-\rho^2)^n\,d\rho=$$

$$=\frac{kb^2\,rx}{a}\left(1-\frac{1-r^2}{a^2}x^2\right)^{n+\frac{1}{2}}\int_{-1}^{1}(1-\rho^2)^n\,d\rho$$

o sea

$$=\frac{brx}{a}f_1(x),\qquad \text{luego}\qquad m_1(x)=\frac{br}{a}x$$

Por tanto, la línea de regresión de y sobre x es

$$y=\frac{bx}{a}x$$

es decir, la regresión es rectilínea e independinete del valor de n.

Lo cual es también consecuencia de la relación que existe entre los momentos

$$\frac{\mu_{f,1}}{\mu_{f+1,0}} = \frac{\mu_{1,1}}{\mu_{2,0}}$$

la cual es precisamente la condición para que la regresión sea rectilínea.

Análogamente la regresión de x sobre y es

$$x = \frac{ar}{b} y$$

El punto de intersección de la recta $y = \frac{br}{a} x$ con la elipse (C) puesta en la forma

$$\left(1 - \frac{1 - r^2}{a^2} x^2\right) - \left(\frac{y}{b} - \frac{rx}{a}\right)^2 = 0$$

es $1 - \frac{1-r^2}{a^2} x^2 = 0$, o sea $x = \pm \frac{a}{\sqrt{1-r^2}}$ que es la tangente paralela al eje y; por tanto, las rectas de regresión son los diámetros conjugados de los ejes coordenados respecto la elipse (C).

La varianza condicional de y sobre x es

$$D^2(y/x) = \frac{\int_{y_1}^{y_2} \left(y - \frac{br}{a} x\right)^2 f(xy)\, dy}{f_1(x)}$$

Calculemos el numerador mediante el cambio

$$\frac{y}{b} - \frac{rx}{a} = \rho \sqrt{1 - \frac{1 - r^2}{a^2} x^2}$$

y se tiene que es igual a

$$k \int_{y_1}^{y_2} \left(y - \frac{brx}{a}\right)^2 \left[\left(1 - \frac{1 - r^2}{a^2} x^2\right) - \left(\frac{y}{b} - \frac{rx}{a}\right)^2\right]^n dy =$$

$$= kb^3 \left(1 - \frac{1 - r^2}{a^2} x^2\right)^{n+\frac{3}{2}} \int_{-1}^{1} \rho^2 (1 - \rho^2)^n \, d\rho$$

es decir,

$$= kb^3\left(1 - \frac{1-r^2}{a^2}x^2\right)^{n+\frac{3}{2}} B\left(\frac{3}{2}, n+1\right) =$$

$$= \frac{kb^3\sqrt{\pi}\,\Gamma(n+1)}{2\,\Gamma\left(n+\frac{5}{2}\right)}\left(1 - \frac{1-r^2}{a^2}x^2\right)^{n+\frac{3}{2}}$$

luego

$$D^2(y/x) = \frac{b^2}{2}\,\frac{\Gamma\left(n+\frac{3}{2}\right)}{\Gamma\left(n+\frac{5}{2}\right)}\left(1 - \frac{1-r^2}{a^2}x^2\right)$$

y finalmente

$$D^2(y/x) = \frac{b^2}{2n+3}\left(1 - \frac{1-r^2}{a^2}x^2\right)$$

La línea de regresión de la desviación típica de y sobre x es, por tanto,

$$y^2 = \frac{b^2}{2n+3}\left(1 - \frac{1-r^2}{a^2}x^2\right)$$

es decir, es la elipse referida a sus ejes de ecuación

$$\frac{x^2}{\frac{a^2}{1-r^2}} + \frac{y^2}{\frac{b^2}{2n+3}} = 1$$

Traslademos la y (o sea σ) paralelamente al eje y a partir de la recta de regresión de y con x que tomaremos como nuevo eje x'.

Se tiene $x = x'\cos\alpha$ $\quad \operatorname{tg}\alpha = \dfrac{br}{a}$ luego $\cos\alpha = \dfrac{a}{\sqrt{a^2+b^2r^2}}$

y la ecuación de la elipse referida a los nuevos ejes x' y es

$$\frac{x'^2\cos^2\alpha}{\frac{a^2}{1-r^2}} + \frac{y^2}{\frac{b^2}{2n+3}} = 1 \qquad \frac{x'^2}{\frac{a^2+b^2r^2}{1-r^2}} + \frac{y^2}{\frac{b^2}{2n+3}} = 1$$

la cual nos da gráficamente el valor de la σ llevada a partir del centro de gravedad de la distribución de y sobre x.

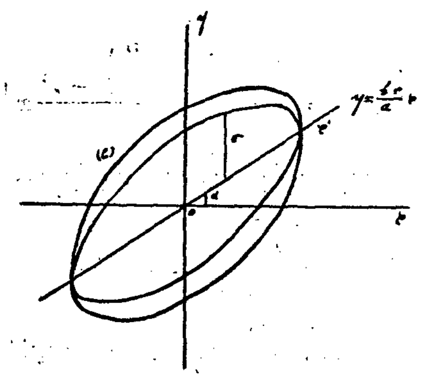
Ya vimos que en la elipse

$$\text{(C)} \qquad \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{2\,rxy}{ab} = 1$$

el eje y y la recta de regresión de y sobre x son diámetros conjugados, luego su ecuación referida a estos ejes es

$$\frac{x'^2}{\dfrac{a^2 + b^2 r^2}{1 - r^2}} + \frac{y^2}{b^2} = 1$$

luego aquella elipse y ésta son homológicas afines en dirección al eje y, y de eje la recta de regresión de y con x.



En general, el momento de orden h condicional de y con x es

$$\frac{\displaystyle\int_{y_1}^{y_2} \left(y - \frac{brx}{a}\right)^h f(xy)\,dy}{f_1(x)}$$

cuyo numerador es, efectuando el mismo cambio de variables,

$$= 2\,kb^{h+1}\left(1 - \frac{1 - r^2}{a^2}x^2\right)^{\frac{h+1}{2}+n} \int_0^1 \rho^h (1 - \rho^2)^n\, d\rho =$$

$$= kb^{h+1}\left(1 - \frac{1 - r^2}{a^2}x^2\right)^{\frac{h+1}{2}+n} \mathrm{B}\left(\frac{h+1}{2},\quad n + 1\right)$$

$$= kb^{h+1}\,\frac{\Gamma\left(\dfrac{h+1}{2}\right)\Gamma(n+1)}{\Gamma\left(n + \dfrac{h+3}{2}\right)}\left(1 - \frac{1 - r^2}{a^2}x^2\right)^{\frac{h+1}{2}+n}$$

luego el momento de orden h vale

$$b^h \frac{\Gamma\left(\frac{h+1}{2}\right)\Gamma\left(n+\frac{3}{2}\right)}{\sqrt{\pi}\,\Gamma\left(n+\frac{h+3}{2}\right)}\left(1-\frac{1-r^2}{a^2}\right)^{\frac{h}{2}}$$

Como caso particular se tiene:

momento de orden 3 $\quad b^3 \frac{\Gamma\left(n+\frac{3}{2}\right)}{\sqrt{\pi}\,\Gamma(n+3)}\left(1-\frac{1-r^2}{a^2}x^2\right)^{\frac{3}{2}}$

momento de orden 4 $\quad b^4 \frac{\Gamma\left(\frac{5}{2}\right)\Gamma\left(n+\frac{3}{2}\right)}{\sqrt{\pi}\,\Gamma\left(n+\frac{7}{2}\right)}\left(1-\frac{1-r^2}{a^2}x^2\right)^2$

Por tanto, el coeficiente de asimetría vale

$$\frac{b^3\,\Gamma\left(n+\frac{3}{2}\right)}{\sqrt{\pi}\,\Gamma(n+3)} : \frac{b^3}{(2n+3)^{\frac{3}{2}}} = \frac{(2n+3)^{\frac{3}{2}}\,\Gamma\left(n+\frac{3}{2}\right)}{\sqrt{\pi}\,\Gamma(n+3)} > 0$$

y el de exceso

$$b^4 \frac{\Gamma\left(\frac{5}{2}\right)\Gamma\left(n+\frac{3}{2}\right)}{\sqrt{\pi}\,\Gamma\left(n+\frac{7}{2}\right)} : \frac{b^4}{(2n+3)^2} - 3 = \frac{3(2n+3)^2}{4\left(n+\frac{5}{2}\right)\left(n+\frac{3}{2}\right)} - 3$$

es decir,

$$\frac{-6}{2n+5} < 0$$

y también sólo depende de n.

Cálculo de las constantes.—Puesto que en la ecuación de la superficie

$$z = k\left(1-\frac{x^2}{a^2}-\frac{y^2}{b^2}+\frac{2\,rxy}{ab}\right)^n$$

entran las cinco constantes k, a, b, r y n, necesitamos cinco ecuaciones para determinarlas. Vimos que los cinco primeros momentos no nulos tienen por valor

$$k \pi ab = (n+1)\sqrt{1-r^2} \qquad \mu_{2,0} = \sigma_1^2 = \frac{a^2}{2(1-r^2)(n+2)}$$

$$\mu_{0,2} = \sigma_2^2 = \frac{b^2}{2(1-r^2)(n+2)} \qquad \mu_{1,1} = \frac{abr}{2(1-r^2)(n+2)}$$

$$\mu_{4,0} = \frac{3a^4}{4(1-r^2)^2(n+2)(n+3)}$$

Ya vimos que

$$r = \frac{\mu_{1,1}}{\sqrt{\mu_{2,0}\,\mu_{0,2}}}$$

por otra parte $\dfrac{\mu_{4,0}}{\mu_{2,0}^2} = 3\dfrac{n+2}{n+3}$, luego $n = 3\dfrac{\mu_{4,0} - 2\mu_{2,0}^2}{3\mu_{2,0}^2 - \mu_{4,0}}$

Sustituyendo en la primera ecuación, estos valores y el de $a \cdot b$, deducido del valor de $\mu_{1,1}$, se tiene

$$k = \frac{2\mu_{4,0} - 3\mu_{2,0}^2}{2\pi\mu_{4,0}\sqrt{\mu_{2,0}\,\mu_{0,2} - \mu_{1,1}^2}}$$

Finalmente, de la segunda y tercera ecuación se obtienen los valores de a^2 y b^2

$$a^2 = 2\frac{\mu_{4,0}(\mu_{2,0}\,\mu_{0,2} - \mu_{1,1}^2)}{\mu_{0,2}(3\mu_{2,0}^2 - \mu_{4,0})} \qquad b^2 = 2\frac{\mu_{4,0}(\mu_{2,0}\,\mu_{0,2} - \mu_{1,1}^2)}{\mu_{2,0}(3\mu_{2,0}^2 - \mu_{4,0})}$$

Pongamos la ecuación de la superficie en función de n, r, σ_1 y σ_2. Se tiene

$$a^2 = 2\sigma_1^2(1-r^2)(n+2) \qquad b^2 = 2\sigma_2^2(1-r^2)(n+2)$$

y además

$$k = \frac{(n+1)\sqrt{1-r^2}}{\pi ab} = \frac{n+1}{2\pi\sigma_1\sigma_2(n+2)\sqrt{1-r^2}}$$

quedando finalmente

$$z = \frac{n+1}{2\pi\sigma_1\sigma_2(n+2)\sqrt{1-r^2}}\left[1 - \frac{1}{2(1-r^2)(n+2)}\left(\frac{x^2}{\sigma_1^2} - \frac{2rxy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}\right)\right]^n$$

ecuación usada por Pearson en «Biométrika» para el ajuste de algunas distribuciones.

Elipses equiprobables.—La probabilidad de que un punto esté en el interior de la elipse

$$\text{(E)} \qquad \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{2rxy}{ab} = 1 - u$$

es

$$P = \iint_{(E)} f(xy)\, dx\, dy$$

y tomando como áreas elementales las de dichas elipses es $s = \frac{\pi ab(1-u)}{\sqrt{1-r^2}}$ luego

$$P = k \int_1^u u^n\, d s = \frac{\pi kab}{\sqrt{1-r^2}} \int_u^1 u^n\, du = \frac{\pi kab}{(n+1)\sqrt{1-r^2}}\,(1 - u^{n+1})$$

y sustituyendo el valor de k queda finalmente

$$P = 1 - u^{n+1}$$

Para un dado valor de P es $n = (1-P)^{\frac{1}{n+1}}$, con lo que queda definida la elipse (E).

Para $P = 1/2$ se tiene $n = 2^{-\frac{1}{n+1}}$, el cual define la elipse probable, es decir, la elipse tal que existe la misma probabilidad de que un punto sea interior o exterior a ella.

Para proceder a las comprobaciones experimentales, basta dividir la región del plano interior a la elipse (C) en diez regiones de igual probabilidad, para lo cual basta hacer en la relación

$$P = 1 - u^{n+1}, \qquad P = \frac{\nu}{10} \qquad \nu = 1, 2, 3 \ldots 10$$

a las cuales corresponden las elipses de parámetro

$$u = \left(1 - \frac{\nu}{10}\right)^{\frac{1}{n+1}}$$

debiéndose verificar que en el interior de la primera elipse y en cada una de las nueve coronas restantes debe haber la décima parte del total.

Casos particulares.—Para $n=1$ la función de densidad tiene por ecuación

$$z = k\left(1 - \frac{x^2}{a^2} - \frac{y^2}{b^2} + \frac{2\,rxy}{ab}\right)$$

y el coeficiente de exceso por ser $\gamma = \dfrac{-3}{n+3}$ es en este caso $\gamma = -\dfrac{3}{4}$ condición que debe cumplirse para que la distribución pueda ser de este tipo.

Para $n=1/2$ la superficie es el elipsoide

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{k^2} - \frac{2\,rxy}{ab} = 1$$

y las funciones de frecuencia marginal son parábolas

$$f_1(x) = \frac{k\pi b}{2}\left(1 - \frac{1-r^2}{a^2}x^2\right), \qquad f_2(y) = \frac{k\pi a}{2}\left(1 - \frac{1-r^2}{b^2}y^2\right)$$

y el coeficiente de exceso es $\gamma = -\dfrac{6}{7}$.

Para $r=0$ es

$$z = k\left(1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}\right)^n$$

En este caso no hay correlación entre las variables, pero no son independientes.

Para $n=-1/2$ las funciones de frecuencia marginal son

$$f_1(x) = kb\pi, \qquad f_2(y) = ka\pi$$

es decir, son uniformes.

SEGUNDA GENERALIZACION

Siguiendo nuestro estudio, consideremos el sistema

$$\frac{\partial z}{\partial x} = z\left(-ax - by + n\,\frac{2\,ax + 2\,by}{h^2 + ax^2 + 2\,bxy + cy^2}\right)$$

$$\frac{\partial z}{\partial y} = z\left(-bx - cy + n\,\frac{2\,bx + 2\,cy}{h^2 + ax^2 + 2\,bxy + cy^2}\right)$$

el cual constituye una generalización de las ecuaciones a que satisface la superficie normal y comprende a ésta como caso particular para $n=0$.

Este sistema es equivalente a la ecuación

$$dz = z\left(-ax - by + n\frac{2ax + 2by}{h^2 + ax^2 + 2bxy + cy^2}\right)dx +$$

$$+ z\left(-bx - cy + n\frac{2bx + 2cy}{h^2 + ax^2 + 2bxy + cy^2}\right)dy$$

la cual cumple la condición de integrabilidad.

Integrando para y constante

$$\frac{dz}{z} = \left(-ax - by + n\frac{2ax + 2by}{h^2 + ax^2 + 2bxy + cy^2}\right)dx$$

$$lz = -\frac{ax^2}{2} - bxy + nl(h^2 + ax^2 + 2bxy + cy^2) + \Phi(y)$$

derivando respecto y

$$-bx - cy + n\frac{2bx + 2cy}{h^2 + ax^2 + 2bxy + cy^2} =$$

$$= -bx + n\frac{2bx + 2cy}{h^2 + ax^2 + 2bxy + cy^2} + \Phi'(y)$$

$$\Phi'(y) = -cy, \qquad \Phi(y) = k_1 - \frac{cy^2}{2}$$

luego

$$lz = k_1 - \frac{ax^2 + 2bxy + cy^2}{2} + l(h^2 + ax^2 + 2bxy + cy^2)^n$$

$$z = f(xy) = ke^{-\frac{ax^2 + 2bxy + cy^2}{2}}(h^2 + ax^2 + 2bxy + cy^2)^n$$

Esta ecuación es la que vamos a estudiar como función de densidad, y según ya dijimos constituye una generalización de la superficie normal y comprende a ésta para $n=0$, luego en todos los cálculos posteriores que obtengamos, si hacemos $n=0$, debe eliminarse h y debe obtenerse el valor correspondiente de la superficie normal.

Supondremos que $ax^2+2bxy+cy^2$ es una función cuadrática positiva, es decir,

$$ac - b^2 > 0 \qquad a > 0 \qquad c > 0$$

Determinemos k para que aquella función constituya una función de densidad, para lo cual debe ser

$$k \iint_{-\infty}^{+\infty} e^{-\frac{ax^2+2bxy+cy^2}{2}} \cdot (h^2 + ax^2 + 2bxy + cy^2)^n \, dx \, dy = 1$$

Considerando como áreas elementales, las elipses

$$h^2 + ax^2 + 2bxy + cy^2 = u \qquad S = \frac{\pi(u - h^2)}{\sqrt{ac - b^2}}$$

se obtiene

$$\frac{\pi k e^{\frac{h^2}{2}}}{\sqrt{ac-b^2}} \int_{h^2}^{\infty} e^{-\frac{u}{2}} \cdot u^n \, du = 1; \qquad k = \frac{\sqrt{ae - b^2}}{\pi e^{\frac{h^2}{2}}} \cdot \frac{1}{\int_{h^2}^{\infty} e^{-\frac{u}{2}} \cdot u^n \, du}$$

y designando por

$$I_n = \int_{h^2}^{\infty} e^{-\frac{u}{2}} \cdot u^n \, du \qquad k = \frac{\sqrt{ac - b^2}}{\pi e^{\frac{h}{2}} I_n}$$

para $n=0$ es $k = \dfrac{\sqrt{ac - b^2}}{2\pi}$ que coincide con el de la superficie normal.

La integral

$$I_n = \int_{h^2}^{\infty} e^{-\frac{u}{2}} \cdot u^n \, du$$

es convergente para cualquier valor de n siempre que $h \neq 0$.

En efecto

$$I_n = \int_{h^2}^{\infty} \frac{e^{-\frac{u}{2}} \cdot u^{n+2}}{u^2} \, du$$

Para $n+2 \leq 0$ la función $e^{-\frac{u}{2}} \cdot u^{n+2}$ es constantemente decreciente, pues su derivada es

$$e^{-\frac{u}{2}} \left[-\frac{u^{n+2}}{2} + (n+2)\, u^{n+1} \right] < 0$$

y, por tanto, la función $e^{-\frac{u}{2}} \cdot u^{n+2}$ está acotada en el intervalo $h^2 \infty$.

Para $n+2>0$ dicha función tiene un máximo para $u=2(n+2)>0$ y, por tanto, está también acotada.

Para $h=0$ es

$$I_n = \int_0^\infty e^{-\frac{u}{2}} \cdot u^n \, du = 2^{n+1} \, \Gamma(n+1)$$

y, por tanto, es convergente para $n+1>0$.

Respecto al cálculo de I_n se tiene que para $h=0$ su cálculo depende de la función Euleriana de primera especie, la cual está ya tabulada.

Para $h \neq 0$ y $n=\frac{r}{2}-1$, siendo r entero y positivo, vamos a ver que el cálculo de I_n se puede hacer mediante la tabla de la χ^2 de Pearson.

En efecto, se tiene

$$P=P(\chi^2>h^2)=\frac{1}{2^{\frac{r}{2}}\,\Gamma\left(\frac{r}{2}\right)}\int_{h^2}^\infty e^{-\frac{x}{2}} x^{\frac{r}{2}-1} dx = \frac{1}{2^{n+1}\,\Gamma(n+1)}\int_{h^2}^\infty e^{-\frac{u}{2}} u^n du$$

luego $I_n=2^{n+1}\cdot\Gamma(n+1)\cdot P$, y siendo $P=\text{Prob}(\chi^2>h^2)$ y calculada con $2(n+1)$ grados de libertad, y entonces k toma la forma

$$k=\frac{\sqrt{ac-b^2}}{\pi e^{\frac{h^2}{2}}\, 2^{n+1}\,\Gamma(n+1)\,P}$$

En general, para $h \neq 0$ la integral I_n viene determinada por la función Γ incompleta.

Resumiendo, la función

$$z = k e^{-\frac{ax^2+2bxy+cy^2}{2}} (h^2+ax^2+2bxy+cy^2)^n$$

es una función de densidad, si $h \neq 0$ para cualquier valor de n, siendo

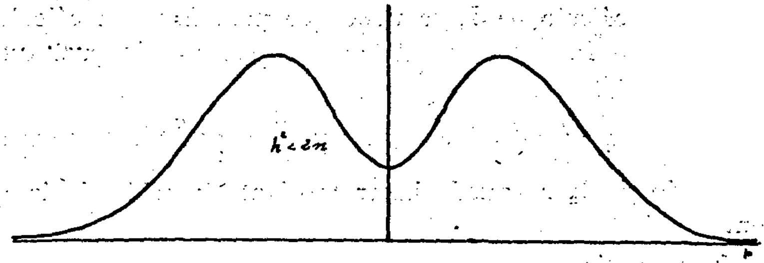
$$k=\frac{\sqrt{ac-b^2}}{\pi e^{-\frac{h^2}{2}}\int_{h^2}^\infty e^{-\frac{u}{2}}\cdot u^n du}$$

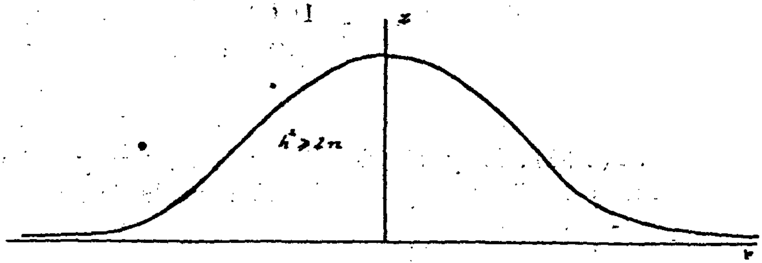
y para $h=0$ si $n+1>0$ y entonces

$$k=\frac{\sqrt{ac-b^2}}{2^{n+1}\pi\,\Gamma(n+1)}$$

Forma de la superficie.—Cortando la superficie por el plano $y=0$ queda

$$z = ke^{-\frac{ax^2}{2}}(h^2+ax^2)^n \quad a>0; \qquad z'=kae^{-\frac{ax^2}{2}}(h^2+ax^2)^{n-1}x(2n-h^2-ax^2)$$





Momentos.—Por la simetría, el centro de gravedad es el origen. Designemos por μ' los momentos correspondientes cuando es $b=0$. Tenemos, pues,

$$\mu'_{2,0} = k \int\int_{-\infty}^{+\infty} e^{-\frac{a'x^2+c'y^2}{2}} \cdot (h^2 + a'x^2 + c'y^2)^n x^2\, dx\, dy$$

Haciendo el cambio

$$x = \frac{1}{\sqrt{a'}} \rho \cos w, \qquad y = \frac{1}{\sqrt{c'}} \rho \operatorname{sen} w$$

$$\mu'_{2,0} = \frac{k}{a'\sqrt{a'c'}} \int\int e^{-\frac{\rho^2}{2}} (h^2 + \rho^2)^n \rho^3 \cos^2 \omega\, d\rho\, d\omega =$$

$$= \frac{k}{a'\sqrt{a'c'}} \int_0^{2\pi} \cos^2 \omega\, d\omega \int_0^\infty e^{-\frac{\rho^2}{2}} (h^2 + \rho^2)^n \rho^3\, d\rho$$

$$\mu'_{2,0} = \frac{k\pi}{a'\sqrt{a'c'}} \int_0^\infty e^{-\frac{\rho^2}{2}} (h^2 + \rho^2)^n \rho^3\, d\rho$$

y finalmente con el cambio $h^2 + \rho^2 = u$

$$\mu'_{2,0} = \frac{k \pi e^{\frac{h^2}{2}}}{2 a' \sqrt{a' c'}} \int_{h^2}^{\infty} e^{-\frac{u}{2}} \cdot u^n (u - h^2)\, du =$$

$$= \frac{k \pi e^{\frac{h^2}{2}}}{2 a' \sqrt{a' c'}} I_{n+1} - \frac{k \pi e^{\frac{h^2}{2}} h^2}{2 a' \sqrt{a' c'}} I_n$$

Mediante una integración por partes

$$I_{n+1} = \int_{h^2}^{\infty} e^{-\frac{u}{2}} u^{n+1}\, du = \int_{h^2}^{\infty} u^{n+1}\, d - 2 e^{-\frac{u}{2}} =$$

$$= \left[-2\, u^{n+1} e^{-\frac{u}{2}} \right]_{h^2}^{\infty} + 2 (n+1) \int_{h^2}^{\infty} e^{-\frac{u}{2}} u^n\, du$$

$$I_{n+1} = 2\, h^{2(n+1)} e^{-\frac{h^2}{2}} + 2 (n+1)\, I_n$$

$$\mu'_{2,0} = \frac{k \pi h^{2(n+1)}}{a' \sqrt{a' c'}} + \frac{k \pi e^{\frac{h^2}{2}} I_n}{\sqrt{a' c'}} \cdot \frac{2 (n+1) - h^2}{2 a'}$$

y teniendo presente el valor de k

$$\mu'_{0,2} = \left[\frac{k \pi h^{2(n+1)}}{\sqrt{a' c'}} + \frac{2 (n+1) - h^2}{2} \right] \frac{1}{a'}$$

y análogamente

$$\mu'_{0.2} = \left[\frac{k \pi h^{2(n+1)}}{\sqrt{a' c'}} + \frac{2 (n+1) - h^2}{2} \right] \frac{1}{c'}$$

y $\mu'_{1,1} = 0$.

Con el mismo cambio de variables hecho anteriormente, se tiene

$$\mu'_{4,0} = k \iint_{-\infty}^{+\infty} e^{-\frac{a' x^2 + c' y^2}{2}} (h^2 + a' x^2 + c' y^2)^n x^4\, dx\, dy$$

$$\mu'_{4,0} = \frac{k}{a'^2 \sqrt{a' c'}} \int_0^{2\pi} \cos^4 \omega\, d\omega \int_0^{\infty} e^{-\frac{\rho^2}{2}} (h^2 + \rho^2)^n \rho^5\, d\rho$$

$$\mu'_{4,0} = \frac{3 k \pi}{4 a'^2 \sqrt{a' c'}} \int_0^{\infty} e^{-\frac{\rho^2}{2}} (h^2 + \rho^2)^n \rho^5\, d\rho$$

y haciendo $h^2 + \rho^2 = u$

$$\mu'_{4,0} = \frac{3\,k\,\pi\,e^{\frac{h^2}{2}}}{8\,a'^2\sqrt{a'\,c'}} \int_{h^2}^{\infty} e^{-\frac{u}{2}}\, u^n\,(u - h^2)^2\,du =$$

$$= \frac{3\,k\,\pi\,e^{\frac{h^2}{2}}}{8\,a'^2\sqrt{a'\,c'}} \left(I_{n+2} - 2\,h^2\,I_{n+1} + h^4\,I_n\right)$$

mediante la integración por partes vista anteriormente, se tiene

$$I_{n+2} = 2\,h^{2(n+2)}\,e^{-\frac{h^2}{2}} + 2\,(n+2)\,I_{n+1}$$

$$I_{n+1} = 2\,h^{2(n+1)}\,e^{-\frac{h^2}{2}} + 2\,(n+1)\,I_n$$

luego

$$I_{n+2} = 2\,h^{2(n+2)}\,e^{-\frac{h^2}{2}} + 4\,(n+2)\,h^{2(n+1)}\,e^{-\frac{h^2}{2}} + 4\,(n+1)\,(n+2)\,I_n$$

y, por tanto,

$$I_{n+2} - 2\,h^2\,I_{n+1} + h^4\,I_n = 2\,h^{2(n+1)}\,e^{-\frac{h^2}{2}}\,(2\,n + 4 - h^2) +$$
$$+ \left[4\,(n+1)\,(n+2-h^2) + h^4\right] I_n$$

y finalmente, teniendo presente el valor de k,

$$\mu'_{4,0} = \left[\frac{k\,\pi\,h^{2(n+1)}}{4\sqrt{a'\,c'}}\,(2n + 4 - h^2) + \frac{(n+1)\,(n+2-h^2)}{2} + \frac{h^4}{8}\right]\frac{3}{a'^2}$$

y análogamente

$$\mu'_{0,4} = \left[\frac{k\,\pi\,h^{2(n+1)}}{4\sqrt{a'c'}}\,(2\,n + 4 - h^2) + \frac{(n+1)\,(n+2-h^2)}{2} + \frac{h^4}{8}\right]\frac{3}{c'^2}$$

Veamos, en general, cuál es la relación que existe entre estos momentos del mismo orden.

Mediante el cambio $x = \frac{1}{\sqrt{a'}}\,\rho\cos\omega$, $y = \frac{1}{\sqrt{c'}}\,\rho\,\mathrm{sen}\,\omega$, se tiene, en general,

$$\mu'_{2p,2q} = k \iint_{-\infty}^{+\infty} e^{-\frac{a'\,x^2 + c'\,y^2}{2}}\,(h^2 + a'\,x^2 + c'\,y^2)^n\,x^{2p}\,y^{2q}\,dx\,dy$$

$$\mu'_{2p,2q} = \frac{k}{a'^p\,c'^q\sqrt{a'c'}} \int_0^{2\pi} \cos^{2p}\omega\,\mathrm{sen}^{2q}\,\omega\,d\,\omega \int_0^{\infty} e^{-\frac{\rho^2}{2}}\,(h^2+\rho^2)^n\,\rho^{2p+2q+1}\,d\rho$$

y análogamente

$$\mu'_{2p+2g,\,2q-2g} = \frac{k}{a'^{\,p+g}\,c'^{\,q-g}\sqrt{a'c'}} \int_0^{2\pi} \cos^{2p+2g}\omega\,\mathrm{sen}^{2q-2g}\omega\,d\omega \int_0^{\infty} e^{-\frac{\rho^2}{2}}(k^2+\rho^2)^n \rho^{2p+2q+}$$

y por división

$$\frac{\mu_{2p\;2q}}{\mu'_{2p+2g,\,2q-2g}} = \left(\frac{a'}{c'}\right)^g \frac{\int_0^{2\pi}\cos^{2p}\omega\,\mathrm{sen}^{2q}\omega\,d\omega}{\int_0^{2\pi}\cos^{2p+2g}\omega\,\mathrm{sen}^{2q-2g}\,d\omega} = \left(\frac{a'}{c'}\right)^g \frac{L_{2p,\,2q}}{L_{2p+2g,\,2q-2g}}$$

Siendo

$$L_{2p,\,2q} = \int_0^{2\pi}\cos^{2p}\omega\,\mathrm{sen}^{2q}\omega\,d\omega = \int_0^{2\pi}\mathrm{sen}^{2q-1}\omega\cdot d\,\frac{\cos^{2p+1}\omega}{2p+1} =$$

$$= \frac{2q-1}{2p+1}\int_0^{2\pi}\cos^{2p+2}\omega\,\mathrm{sen}^{2q-2}\omega\,d\omega$$

y de esta fórmula se deducen las siguientes

$$L_{2p,\,2q} = \frac{2q-1}{2p+1} L_{2p+2,\,2q-2}$$

$$L_{2p+2,\,2q-2} = \frac{2q-3}{2p+3} L_{2p+4,\,2q-4}$$

. . . .

$$L_{2p+2g-2,\,2q-2g+2} = \frac{2q-2g+1}{2p+2g-1} L_{2p+2q,\,2q-2g}$$

y, por tanto,

$$L_{2p,\,2q} = \frac{(2q-1)(2q-3)\ldots(2q-2g+1)}{(2p+1)(2p+3)\ldots(2p+2g-1)} L_{2p+2g,\,2q-2g}$$

y sustituyendo este valor

$$\frac{\mu'_{2p,\,2q}}{\mu'_{2p+2g,\,2q-2g}} = \frac{(2q-1)(2q-3)\ldots(2q-2g+1)}{(2p+1)(2p+3)\ldots(2p+2g-1)}\left(\frac{a'}{c'}\right)^g$$

y análogamente

$$\frac{\mu'_{2p,\,2q}}{\mu'_{2p-2g,\,2q+2g}} = \frac{(2p-1)(2p-3)\ldots(2p-2g+1)}{(2q+1)(2q+3)\ldots(2q+2g-1)}\left(\frac{c'}{a'}\right)^g$$

Veamos ahora la relación entre los momentos de distinto orden.

Según hemos visto ya anteriormente

$$\mu'_{2p,\,2q,\,n} = \frac{k_n}{a'^{p+\frac{1}{2}}\,c'^{q+\frac{1}{2}}} \int_0^{2\pi} \cos^{2p}\omega \,\mathrm{sen}^{2q}\omega\, d\omega \int_0^{\infty} e^{-\frac{\rho^2}{2}} (h^2+\rho^2)^n \rho^{2p+2q+1}\, d\rho$$

pero se tiene que

$$\int_0^{2\pi} \cos^{2p}\omega\,\mathrm{sen}^{2q}\omega\, d\omega = \frac{2q-1}{2(p+q)} \int_0^{2\pi} \cos^{2p}\omega\,\mathrm{sen}^{2q-2}\omega\, d\omega$$

y por ser

$$\int_0^{\infty} e^{-\frac{\rho^2}{2}} (h^2+\rho^2)^n \rho^{2p+2q+1}\, d\rho = \int_0^{\infty} e^{-\frac{\rho^2}{2}} (h^2+\rho^2)^n (h^2+\rho^2-h^2)\, \rho^{2p+2q-1}\, d\rho$$

$$= \int_0^{\infty} e^{-\frac{\rho^2}{2}} (h^2+\rho^2)^{n+1} \rho^{2p+2q-1}\, d\rho - h^2 \int_0^{\infty} e^{-\frac{\rho^2}{2}} (h^2+\rho^2)^n \rho^{2p+2q-1}\, d\rho$$

y sustituyendo

$$\mu'_{2p,\,2q,\,n} = \frac{1}{c'}\,\frac{2q-1}{2(p+q)}\,\frac{k_n}{a'^{p+\frac{1}{2}}\,c'^{q+\frac{1}{2}}} \int_0^{2\pi} \cos^{2p}\omega\,\mathrm{sen}^{2q-2}\omega\, d\omega \left[\int_0^{\infty} e^{-\frac{\rho^2}{2}} (h^2+\rho^2)^{n+1} \rho^{2p+2q-1} d\rho \right.$$

$$\left. - h^2 \int_0^{\infty} e^{-\frac{\rho^2}{2}} (h^2+\rho^2)^n \rho^{2p+2q-1}\, d\rho \right]$$

es decir,

$$\mu'_{2p,\,2q,\,n} = \frac{1}{c'}\,\frac{2q-1}{2(p+q)} \left(\frac{k_n}{k_{n+1}}\, \mu'_{2p,\,2q-2,\,n+1} - h^2 \mu'_{2p,\,2q} \right)$$

Según vimos ya

$$k_n = \frac{\sqrt{a'c'}}{\pi e^{\frac{h^2}{2}} I_n} \qquad I_{n+1} = 2 e^{-\frac{h^2}{2}} h^{2(n+1)} + 2(n+1) I_n$$

y sustituyendo en la segunda el valor de I en función de k

$$\frac{\sqrt{a'c'}}{\pi e^{\frac{h^2}{2}} k_{n+1}} = 2 e^{-\frac{h^2}{2}} h^{2(n+1)} + 2(n+1) \frac{\sqrt{a'c'}}{\pi e^{\frac{h^2}{2}} k_n}$$

de donde

$$\frac{k_n}{k_{n+1}} = 2(n+1) + \frac{2\pi h^{2(n+1)}}{\sqrt{a'c'}}\, k_n$$

Así para $p=0$, $q=2$

$$\mu'_{0,4} = \frac{3}{4\,c'}\left(\frac{h_n}{k_{n+1}}\,\mu'_{0,2,n+1} - h^2\,\mu'_{0,2,n}\right)$$

$$\mu'_{0,4} = \frac{3}{4\,c'^2}\,\frac{k_n}{k_{n+1}}\left(\frac{k_{n+1}\,\pi\,h^{2(n+2)}}{\sqrt{a'c'}} + \frac{2(n+2)-h^2}{2}\right) -$$

$$- \frac{3\,h^2}{4\,c'}\left(\frac{k_n\,\pi\,h^{2(n+1)}}{\sqrt{a'c'}} + \frac{2(n+1)-h^2}{2}\right)$$

$$= \frac{3}{4\,c'^2}\,\frac{k_n}{k_{n+1}}\,\frac{2(n+2)-h^2}{2} - \frac{3\,h^2(2\,n+2-h^2)}{8\,c'^2}$$

o sea

$$\mu'_{0,4} = \left[\frac{k_n\,\pi\,h^{2(n+1)}}{4\sqrt{a'c'}}(2\,n+4-h^2) + \frac{(n+1)(n+2-h^2)}{2} + \frac{h^4}{8}\right]\frac{3}{c'^2}$$

conforme obtuvimos anteriormente.

Mediante estas fórmulas, se conocen todos los momentos cuando es $b=0$.

Hallemos los momentos en el caso $b \neq 0$.

Girando los ejes coordenados hasta hacerlos coincidir con los de la elipse

$$ax^2 + 2\,bxy + cy^2 = \gamma$$

las fórmulas de transformación son

$$x = \mathrm{X}\cos\varphi - \mathrm{Y}\,\mathrm{sen}\,\varphi$$

$$y = \mathrm{X}\,\mathrm{sen}\,\varphi + \mathrm{Y}\cos\varphi$$

con lo que aquélla se transforma en

$$a'\,\mathrm{X}^2 + c'\,\mathrm{Y}^2 = \gamma$$

siendo

$$a' = a\cos^2\varphi + b\,\mathrm{sen}\,2\,\varphi + c\,\mathrm{sen}^2\varphi$$

$$c' = a\,\mathrm{sen}^2\varphi - b\,\mathrm{sen}\,2\,\varphi + c\cos^2\varphi$$

de las cuales se deduce

$$a'\,c' = ac - b^2, \qquad \mathrm{tg}\,2\,\varphi = \frac{2\,b}{a-c}$$

Efectuando, pues, aquel cambio de variables se tiene

$$\mu'_{2,0} = k \int\int_{-\infty}^{+\infty} e^{-\frac{ax^2+2bxy+cy^2}{2}} (h^2 + ax^2 + 2bxy + cy^2)^n x^2 \, dx \, dy =$$

$$= k \int\int_{-\infty}^{+\infty} e^{-\frac{a'X^2+c'Y^2}{2}} (h^2+a'X^2+c'Y^2)^n (X^2 \cos^2\varphi - XY \operatorname{sen} 2\varphi + Y^2 \cos^2\varphi) \, dX \, dY$$

$$\mu_{2,0} = \cos^2\varphi \, \mu'_{2,0} + \operatorname{sen}^2\varphi \, \mu'_{0,2}$$

y sustituyendo los valores ya calculados de $\mu'_{2,0}$ y $\mu'_{0,2}$

$$\mu_{2,0} = \left[\frac{k \pi h^{2(n+1)}}{\sqrt{a'c'}} + \frac{2(n+1) - h^2}{2} \right] \frac{c' \cos^2\varphi + a' \operatorname{sen}^2\varphi}{a'c'}$$

Si en las fórmulas de transfromación hacemos $X = \operatorname{sen}\varphi$, $Y = \cos\varphi$ se obtiene $x=0$, $y=1$, y como

$$a'X^2 + c'Y^2 = ax^2 + 2bxy + cy^2$$

resulta

$$a' \operatorname{sen}^2\varphi + c' \cos^2\varphi = c$$

quedando finalmente

$$\mu_{2,0} = \left[\frac{k \pi h^{2(n+1)}}{\sqrt{ac - b^2}} + \frac{2(n+1) - h^2}{2} \right] \frac{c}{ac - b^2}$$

Para $n=0$ resulta $\mu_{2,0} = \dfrac{c}{ac - b^2}$ que coincide con el valor correspondiente de la superficie normal.

Análogamente resulta

$$\mu_{0,2} = \left[\frac{k \pi h^{2(n+1)}}{\sqrt{ac - b^2}} + \frac{2(n+1) - h^2}{2} \right] \frac{a}{ac - b^2}$$

Efectuando el mismo cambio de variables, se obtiene para la covarianza

$$\mu_{1,1} = k \int\int_{-\infty}^{+\infty} e^{-\frac{ax^2+2bxy+cy^2}{2}} (h^2 + ax^2 + 2bxy + cy^2)^n xy \, dx \, dy =$$

$$= k \int\int_{-\infty}^{+\infty} e^{-\frac{a'X^2+c'Y^2}{2}} (h^2 + a'X^2 + c'Y^2)^n (X^2 \operatorname{sen}\varphi \cos\varphi + XY \cos 2\varphi -$$

$$- Y^2 \operatorname{sen}\varphi \cos\varphi) \, dX \, dY$$

$$\mu_{1,1} = (\mu'_{2,0} - \mu'_{0,2}) \operatorname{sen} \varphi \cos \varphi = \left[\frac{k \pi h^{2(n+1)}}{\sqrt{a' c'}} + \frac{2(n+1) - h^2}{2} \right] \frac{(c'-a') \operatorname{sen} \varphi \cos \varphi}{a' c'}$$

pero

$$(c' - a') \operatorname{sen} \varphi \cos \varphi = \frac{1}{2} \operatorname{sen} 2\varphi (- a \cos 2\varphi - 2b \operatorname{sen} 2\varphi + c \cos 2\varphi) =$$

$$= - \frac{1}{2} [(a-c) \operatorname{sen} 2\varphi \cos 2\varphi + 2b \operatorname{sen}^2 2\varphi] = - \frac{1}{2} (2b \cos^2 2\varphi + 2b \operatorname{sen}^2 2\varphi) = -b$$

con lo que

$$\mu_{1,1} = \left[\frac{h \pi h^{2(n+1)}}{\sqrt{ac - b^2}} + \frac{2(n+1) - h^2}{2} \right] \frac{-b}{ac - b^2}$$

Para $n=0$ es $\mu_{1,1} = \dfrac{ac - b^2}{-b}$ que coincide con el correspondiente de la superficie normal.

Veamos una relación que existe entre los elementos μ' y los μ. Se tiene

$$\mu_{2p,0} = k \int\int_{-\infty}^{+\infty} e^{-\frac{ax^2 + 2bxy + cy^2}{2}} (h^2 + ax^2 + 2bxy + cy^2)^n x^{2p} \, dx \, dy$$

o sea

$$\mu_{2p,0} = k \int\int_{-\infty}^{+\infty} e^{-\frac{a' X^2 + c' Y^2}{2}} (h^2 + a' X^2 + c' Y^2)^n (X \cos \varphi - Y \operatorname{sen} \varphi)^{2p} \, dX \, dY$$

pero

$$(X \cos \varphi \; Y \operatorname{sen} \varphi)^{2p} = \sum_{\nu=0}^{2p} (-1)^\nu \binom{2p}{\nu} X^{2p-\nu} Y^\nu \cos^{2p-\nu} \varphi \operatorname{sen}^\nu \varphi$$

y teniendo presente que los momentos impares son nulos

$$\mu_{2p,0} = k \sum_{\nu=0}^{p} \binom{2p}{2\nu} \cos^{2p-2\nu} \varphi \operatorname{sen}^{2\nu} \varphi \int\int_{-\infty}^{+\infty} e^{-\frac{a' X^2 + c' Y^2}{2}} (h^2 + a' X^2 +$$

$$+ c' Y^2)^n X^{2p-2\nu} Y^{2\nu} \, dX \, dY = \sum_{\nu=0}^{p} \binom{2p}{2\nu} \cos^{2p-2\nu} \varphi \operatorname{sen}^{2\nu} \varphi \cdot \mu'_{2p-2\nu, 2\nu}$$

Por la relación que vimos anteriormente entre los momentos del mismo orden, se verifica

$$\mu'_{2p-2\nu, 2\nu} = \frac{1 . 3 \ldots . (2\nu - 1)}{(2p - 1)(2p - 3) \ldots (2p - 2\nu + 1)} \left(\frac{a'}{c'} \right)^\nu \mu'_{2p,0}$$

y por ser

$$\binom{2p}{2\nu} \frac{1,3, \ldots (2\nu - 1)}{(2p-1)(2p-3)\ldots(2p-2\nu+1)} = \binom{p}{\nu}$$

se obtiene

$$\mu_{2p,0} = \mu'_{2p,0} \sum_{\nu=0}^{p} \binom{p}{\nu} \cos^{2p-2\nu} \varphi \operatorname{sen}^{2\nu} \varphi \left(\frac{a'}{c'}\right)^{\nu} =$$

$$= \mu'_{2p,0} \left(\cos^2 \varphi + \frac{a'}{c'} \operatorname{sen}^2 \varphi\right)^{p} = \frac{\mu'_{2p,0}}{c'^{p}} (c' \cos^2 \varphi + a' \operatorname{sen}^2 \varphi)^{p}$$

y como ya dijimos que $c' \cos^2 \varphi + a' \operatorname{sen}^2 \varphi = c$ queda finalmente

$$\mu_{2p,0} = \left(\frac{c}{c'}\right)^{p} \mu'_{2p,0}$$

En particular

$$\mu_{4,0} = \left(\frac{c}{c'}\right)^{2} \mu'_{4,0}$$

y sustituyendo el valor de $\mu'_{4\,0}$ ya calculado

$$\mu_{4,0} = \frac{3\,c^2\,k\,\pi\,h^{2(n+1)}}{4\,(a'\,c')^{2+\frac{1}{2}}} (2n+4-h^2) + \frac{3\,c^2}{(a'\,c')^2} \left[\frac{(n+1)(n+2-h^2)}{2} + \frac{h^4}{8}\right]$$

o sea

$$\mu_{4,0} = \left[\frac{k\,\pi\,h^{2(n+1)}}{4\sqrt{ac-b^2}} (2n+4-h^2) + \frac{(n+1)(n+2-h^2)}{2} + \frac{h^4}{8}\right] \frac{3\,c^2}{(ac-b^2)^2}$$

Para $n=0$ resulta $\mu_{4,0} = \dfrac{3\,c^2}{(ac-b^2)^2}$ como en el caso de la superficie normal.

Coeficiente de correlación.—El valor de dicho coeficiente es

$$r = \frac{\mu_{1,1}}{\sqrt{\mu_{2,0}\,\mu_{0,2}}}$$

y en nuestro caso

$$r = \frac{-b}{\sqrt{ac}}$$

es decir, independiente de h y n, teniendo además el mismo valor que en el caso de la superficie normal.

Funciones marginales.—La expresión de la función marginal de la x es

$$f_1(x) = \int_{-\infty}^{\infty} f(xy)\,dy$$

o sea

$$f_1(x) = k \int_{-\infty}^{\infty} e^{-\frac{ax^2 + 2bxy + cy^2}{2}} (h^2 + ax^2 + 2bxy + cy^2)^n\,dy$$

Haciendo

$$\frac{ac - b^2}{c} = d, \qquad ax^2 + 2bxy + cy^2 = \left(y\sqrt{c} + \frac{b}{\sqrt{c}}\,x\right)^2 + dx^2$$

y efectuando el cambio de variables

$$y\sqrt{c} + \frac{b}{\sqrt{c}}\,x = \sqrt{h^2 + dx^2} \cdot t$$

$$f_1(x) = \frac{2k}{\sqrt{c}}\,e^{-\frac{dx^2}{2}} (h^2 + dx^2)^{n+\frac{1}{2}} \int_0^{\infty} e^{-\frac{h^2+dx^2}{2}t^2} \cdot (1 + t^2)^n\,dt$$

que es la expresión de la función marginal de la X. Si hacemos

$$\varphi(x, n) = \int_0^{\infty} e^{-\frac{h^2+dx^2}{2}t^2} (1 + t^2)^n\,dt$$

que es la integral de Laplace, resulta

$$f_1(x) = \frac{2k}{\sqrt{c}}\,e^{-\frac{dx^2}{2}} (h^2 + dx^2)^{n+\frac{1}{2}}\,\varphi(x, n)$$

Expresión análoga para la función marginal de la y. Para $n = 0$ resulta $f_1(x) = \frac{k\sqrt{2\pi}}{\sqrt{c}}\,e^{-\frac{dx^2}{2}}$ que coincide con la correspondiente de la superficie normal.

Regresión.—La regresión de y sobre x es

$$m_1(x) = \frac{\int_{-\infty}^{\infty} y\,f(xy)\,dy}{f_1(e)}$$

siendo el numerador

$$\int_{-\infty}^{\infty} y\, f(xy)\, dy = k \int_{-\infty}^{\infty} y e^{-\frac{ax^2+2bxy+cy^2}{2}} (h^2 + ax^2 + 2bxy + cy^2)^n\, dy$$

que con el cambio anterior se transforma en

$$\int_{-\infty}^{\infty} y\, f(xy)\, dy = \frac{k}{\sqrt{c}} e^{-\frac{dx^2}{2}} (h^2 + dx^2)^{n+\frac{1}{2}} \int_{-\infty}^{\infty} \left[t \frac{\sqrt{h^2 + dx^2}}{\sqrt{c}} - \right.$$

$$\left. - \frac{b}{c} x \right] e^{-\frac{h^2+dx^2}{2} t^2} (1 + t^2)^n\, dt = \left(- \frac{b}{c} x \right) \frac{2\, k e^{-\frac{dx^2}{2}}}{\sqrt{c}} (h^2 +$$

$$+ dx^2)^{n+\frac{1}{2}} \int_{0}^{\infty} e^{-\frac{h^2+dx^2}{2} t^2} (1 + t^2)^n dt = \left(- \frac{b}{c} x \right) f_1(x)$$

luego $m_1(x) = - \frac{b}{c} x$ y análogamente $m_2(y) = - \frac{b}{a} y$; por tanto, las líneas de regresión son independientes de h y n y son rectilíneas, teniendo además el mismo valor que en el caso de la superficie normal.

Varianza condicional.—Su expresión es, en el caso de y sobre x,

$$D^2(y/x) = \frac{\int_{-\infty}^{\infty} \left(y + \frac{b}{c} x \right)^2 f(xy)\, dy}{f_1(x)}$$

Efectuando el cambio de variables visto anteriormente, se tiene

$$f_1(x) \cdot D^2(y/x) = k \int_{-\infty}^{\infty} \left(y + \frac{b}{c} \right)^2 e^{-\frac{1}{2}\left(y\sqrt{c} + \frac{b}{\sqrt{c}} x \right)^2} \left[h^2 + \left(y \sqrt{c} + \frac{bx}{\sqrt{c}} \right)^2 + \right.$$

$$\left. + dx^2 \right]^n dy = \frac{2\, k}{c \sqrt{c}} e^{-\frac{dx^2}{2}} (h^2 + dx^2)^{n+\frac{3}{2}} \int_{0}^{\infty} e^{-\frac{h^2+dx^2}{2} t^2} (1 + t^2)^n\, t^2\, dt$$

luego

$$D^2(y/x) = \frac{h^2 + dx^2}{c} \frac{\int_{0}^{\infty} e^{-\frac{h^2+dx^2}{2} t^2} \cdot (1 + t^2)^n\, t^2\, dt}{\int_{0}^{\infty} e^{-\frac{h^2+dx^2}{2} t^2} (1 + t^2)^n\, dt}$$

que puede ponerse en la forma, en función de la integral de Laplace

$$D^2(y/x) = \frac{h^2 + dx^2}{c}\left[\frac{\varphi(x,\ n+1)}{\varphi(x,\ n)} - 1\right]$$

Para $n=0$ resulta $D^2(y/x) = \frac{1}{c}$ que coincide con el valor correspondiente de la superficie normal.

Elipses equiprobables.—Consideremos las elipses

$$\text{(E)} \qquad h^2 + ax^2 + 2\,bxy + cy^2 = u$$

La probabilidad de que un punto esté en el interior de dicha elipse es

$$P = k\int\!\!\int_{(E)} e^{-\frac{ax^2+2\,bxy+cy^2}{2}}\,(h^2 + ax^2 + 2\,bxy + cy^2)^n\,dx\,dy$$

y considerando como áreas elementales dichas elipses como hicimos al calcular el valor de k, resulta

$$P = \frac{\pi\,k\,e^{\frac{h^2}{2}}}{\sqrt{ac - b^2}}\int_{h^2}^{u} e^{-\frac{u}{2}}\,u^n\,du$$

y sustituyendo el valor de k

$$P = \frac{\int_{h^2}^{u} e^{-\frac{u}{2}}\,u^n\,du}{\int_{h^2}^{\infty} e^{-\frac{u}{2}}\,u^n\,du} = \frac{\int_{h^2}^{\infty} e^{-\frac{u}{2}}\,u^n\,du - \int_{u}^{\infty} e^{-\frac{u}{2}}\,u^n\,du}{\int_{h^2}^{\infty} e^{-\frac{u}{2}}\,u^n\,du}$$

quedando finalmente

$$P = 1 - \frac{\int_{u}^{\infty} e^{-\frac{u}{2}}\,u^n\,du}{\int_{h^2}^{\infty} e^{-\frac{u}{2}}\,u^n\,du}$$

Para un dado valor de P, podemos hacar el correspondiente valor de u, el cual, sustituído en (E), nos define la elipse tal que la probabilidad de que un punto esté en el interior de ella es precisamente P. En particular, para P=1/2 se tiene la elipse probable. Haciendo

$$P = \frac{\nu}{10} \qquad \nu = 1, 2, 3 \ldots 10$$

obtenemos diez elipses tales que la probabilidad de que un punto esté en el interior de cada corona es precisamente $\frac{1}{10}$, lo que constituye un método para comparar con la práctica.

En el caso de que n sea de la forma $n = \frac{r}{2} - 1$, siendo r entero y positivo ya vimos que

$$\int_u^\infty e^{-\frac{u}{2}} u^n \, du = 2^{n+1} \Gamma(n+1) P_1$$

siendo $P_1 = \text{Prob}\,(\chi^2 > u)$ con $2(n+1)$ grados de libertad.

Por tanto, si es $P_2 = \text{Prob}\,(\chi^2 > h^2)$ resulta

$$P = 1 - \frac{P_1}{P_2}$$

y la distribución de la χ^2 de Pearson nos proporciona en este caso el valor de P, conocido u y recíprocamente.

Para $h=0$ resulta $P_2 = 1$ y entonces

$$P = 1 - \text{Prob}\,(\chi^2 > u) = \text{Prob}\,(\chi^2 < u)$$

con $2(n+1)$ grados de libertad. En general, el cálculo de P depende de la Γ incompleta.

Caso particular.—El caso particular $h=0$ es interesante ya que las fórmulas obtenidas se simplifican extraordinariamente. Nuestra función de densidad es entonces

$$z = k e^{-\frac{ax^2 + 2bxy + cy^2}{2}} (ax^2 + 2bxy + cy^2)^n$$

con

$$k = \frac{\sqrt{ac - b^2}}{2^{n+1} \pi \Gamma(n+1)} \qquad n + 1 > 0$$

Haciendo $h=0$ en las fórmulas obtenidas para los momentos resulta

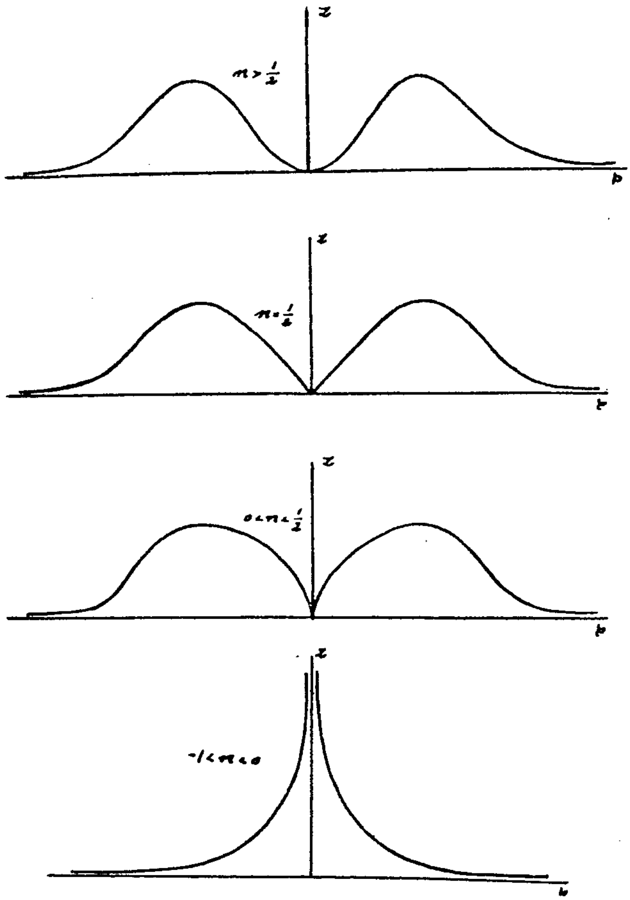
$$\sigma_1^2 = \frac{c(n+1)}{ac - b^2} \qquad \sigma_2^2 = \frac{a(n+1)}{ac - b^2} \qquad \mu_{4,0} = \frac{3(n+1)(n+2)c^2}{2(ac - b^2)^2}$$

y designando por γ_1 el coeficiente de exceso de la variable x

$$\gamma_1 = \frac{\mu_{4,0}}{\sigma_1^4} - 3 \qquad\qquad \gamma_1 = \frac{-3\,n}{2\,(n+1)}$$

dependiente solamente de n.

Pongamos la función de densidad en función de σ_1, σ_2, r y n. Del



sistema de ecuaciones

$$\sigma_1^2 = \frac{c\,(n+1)}{ac - b^2}, \qquad \sigma_2^2 = \frac{a\,(n+1)}{ac - b^2}, \qquad r = \frac{-b}{\sqrt{ac}}$$

se deduce

$$a = \frac{n+1}{\sigma_1^2\,(1-r^2)}, \qquad b = -\frac{r\,(n+1)}{\sigma_1\,\sigma_2\,(1-r^2)}, \qquad c = \frac{n+1}{\sigma_2^2\,(1-r^2)}$$

y sustituyendo

$$z = \frac{(n+1)^{n+1}}{\sigma_1 \sigma_2 (1-r^2)^{n+\frac{1}{2}} \cdot 2^{n+1} \cdot \pi \cdot \Gamma(n+1)} e^{-\frac{n+1}{2(1-r^2)}\left(\frac{x^2}{\sigma_1^2} - \frac{2rxy}{\sigma_1 \sigma_2} + \frac{y^2}{\sigma_1^2}\right)} \cdot \left(\frac{x^2}{\sigma_1^2} - \frac{2rxy}{\sigma_1 \sigma_2} + \frac{y^2}{\sigma_2^2}\right)^n$$

El valor de n se puede calcular de la fórmula que hemos visto para el exceso

$$\gamma_1 = \frac{-3n}{2(n+1)}, \qquad n = \frac{-2\gamma_1}{2\gamma_1 + 3}$$

Por la condición $n+1 = \dfrac{3}{2\gamma_1 + 3} > 0$ se deduce que $\gamma_1 > -1{,}5$.

Forma de la superficie.—Cortando por el plano $y=0$ resulta

$$z = ka^n e^{-\frac{ax^2}{2}} x^{2n}, \qquad z' = ka^n e^{-\frac{ax^2}{2}} x^{2n-1}(2n - ax^2)$$

Estas curvas, para los distintos valores de n, están representadas en la página anterior.

Cálculo de las constantes.—En la ecuación de nuestra función de densidad

$$z = ke^{-\frac{ax^2+2bxy+cy^2}{2}} (h^2 + ax^2 + 2bxy + cy^2)^n$$

vamos a determinar las cinco constantes k, a, b, c y h, suponiendo el valor de n ya fijado de antemano. Para ello necesitamos las cinco ecuaciones siguientes

$$k = \frac{\sqrt{ac-b^2}}{\pi e^{\frac{h^2}{2}} \int_{h^2}^{\infty} e^{-\frac{u}{2}} u^n \, du}, \qquad \sigma_1^2, \qquad \sigma_2^2, \qquad \mu_{4,0} \qquad \text{y} \qquad r = \frac{-b}{\sqrt{ac}}$$

y suponemos que se conocen los valores de σ_1^2, σ_2^2, $\mu_{4,0}$ y r deducidos de las observaciones.

Haciendo

$$[1] \qquad \int_{h^2}^{\infty} e^{\frac{h^2-u}{2}} u^n \, du = \frac{1}{I} \qquad \text{resulta} \qquad \frac{k\pi}{\sqrt{ac-b^2}} = I$$

El coeficiente de kurtosis $\beta = \dfrac{\mu_{4.0}}{\sigma_1^4}$ se puede poner solamente en función de I y h^2, resultando una ecuación de segundo grado en I y, por tanto, se puede obtener éste en función de h^2, el cual, sustituído en [1], se obtiene una ecuación en h^2, y por tanteo se puede obtener el valor de éste y por tanto el de I. Estos valores, sustituídos en σ_1^2 y σ_2^2 junto con $r = \frac{-b}{\sqrt{ac}}$, nos proporciona los valores de a, b y c, así como el de k, por ser $k = \frac{\mathrm{I}\sqrt{ac - b^2}}{\pi}$.

En el caso particular de $n = -1$ resulta

$$\sigma_1^2 = \left(\mathrm{I} - \frac{h^2}{2}\right)\frac{c}{ac-b^2}\,,\quad \sigma_2^2 = \left(\mathrm{I} - \frac{h^2}{2}\right)\frac{a}{ac-b^2}\,,\quad \mu_{4,0} =$$

$$= \left[\mathrm{I}\,(2 - h^2) + \frac{h^4}{2}\right]\frac{3\,c^2}{4\,(ac - b^2)^2}$$

luego el coeficiente de kurtosis vale $\beta = \dfrac{3\left(2\,\mathrm{I} - h^2\,\mathrm{I} + \dfrac{h^4}{2}\right)}{4\left(\mathrm{I} - \dfrac{h^2}{2}\right)^2}$

$$4\,\beta\,\mathrm{I}^2 - (4\,\beta\,h^2 - 3\,h^2 + 6)\,\mathrm{I} + h^4\,\beta - \frac{3\,h^4}{2} = 0$$

esta ecuación tiene sus dos raíces reales ya que su discriminante es $(3h^2-6)^2 + 48\beta h^2 > 0$.

Por ser $\sigma_1^2 > 0$ debe ser $\mathrm{I} > \frac{h^2}{2}$ y como para $\mathrm{I} = \frac{h^2}{2}$ el primer miembro de la ecuación en I toma el valor $-3h^2 < 0$, luego de las dos raíces hay que tomar solamente la mayor.

Sustituído este valor de I en [1]

$$\int_{h^2}^{\infty} \frac{e^{-\frac{u}{2}}}{u}\,du = \frac{e^{-\frac{h^2}{2}}}{\mathrm{I}}$$

y disponiendo de las tablas de la función Γ incompleta, se puede obtener por tanteo el valor de h^2.

Conocidos los valores de I y h^2, si llamamos a $\mathrm{I} - \frac{h^2}{2} = s^2$ se tiene

$$\sigma_1^2 = \frac{cs^2}{ac - b^2}\,,\qquad \sigma_2^2 = \frac{as^2}{ac - b^2}\,,\qquad r = \frac{-b}{\sqrt{ac}}$$

de donde

$$a = \frac{s^2}{\sigma_1^2\,(1 - r^2)}\,,\qquad c = \frac{s^2}{\sigma_2^2\,(1 - r^2)}\,;\qquad b = \frac{-rs^2}{\sigma_1\,\sigma_2\,(1 - r^2)}\,,\qquad k = \frac{\mathrm{I}\,s^2}{\pi\,\sigma_1\,\sigma_2\sqrt{1 - r^2}}$$

y la ecuación toma la forma

$$z = \frac{I\,s^2}{\pi\,\sigma_1\,\sigma_2\sqrt{1-r^2}}\,e^{-\frac{s^2}{2(1-r^2)}\left(\frac{x^2}{\sigma_1^2}-\frac{2rxy}{\sigma_1\sigma_2}+\frac{y^2}{\sigma_2^2}\right)}\left[2\,(I-s^2)+\frac{s^2}{1-r^2}\left(\frac{x^2}{\sigma_1^2}-\frac{2\,rxy}{\sigma_1\,\sigma_2}+\frac{y^2}{\sigma_2^2}\right)\right]^{-1}$$

No queríamos terminar este trabajo sin hacer una aplicación práctica del mismo.

En el tratado de estadística matemática de Risser et Traynard se estudia el ajuste de la superficie normal a la tabla de frecuencia entre la longitud y anchura de las hojas de Hedera Helix, llegando a la conclusión de que no se ajusta lo conveniente. Teníamos la intención de ajustar una superficie del tipo que acabamos de esutdiar $n=-1$, ya que este caso es una generalización de la superficie normal, y aunque ya teníamos todos los datos preparados con este objeto, por no disponer de tablas de la función Γ, sin las cuales es imposible hacer dicho ajuste, según acabamos de ver, hemos tenido que desistir y esperar a una mejor ocasión para poder publicar el resultado que obtengamos mediante dicho estudio.

TERCERA GENERALIZACION

Si ponemos la ecuación de la superficie normal en unidades standard

$$z = \frac{1}{2\pi\sqrt{1-r^2}}\,e^{-\frac{1}{2(1-r^2)}(x^2-2\,rxy+y^2)}$$

y siguiendo el mismo proceso de derivación, se tiene

$$\frac{\partial z}{\partial x} = p = -z\,\frac{x-ry}{1-r^2}$$

$$\frac{\partial z}{\partial y} = q = -z\,\frac{y-rx}{1-r^2}$$

entre las cuales eliminando r se llega a la ecuación en derivadas parciales

$$p^2 - q^2 + z\,(px - qy) = 0$$

Recíprocamente veamos cuáles son las superficies de probabilidad que satisfacen a esta ecuación.

Haciendo $z = e^{u}$ se obtiene

$$p = e^{u} \frac{\partial u}{\partial x} = e^{u} \cdot p_1 \qquad q = e^{u} \frac{\partial u}{\partial y} = e^{u} q_1$$

y sustituyendo

$$p_1^2 - q_1^2 + p_1 x - q_1 y = 0$$

Hallemos una integral completa de esta ecuación por el método de Lagrange-Charpit, y se llega al sistema de ecuaciones

$$\frac{dx}{2 p_1 + x} = \frac{dy}{-2 q_1 - y} = \frac{dz}{q_1 y - p_1 x} = \frac{-dp_1}{p_1} = \frac{-dq_1}{-q_1}$$

Esta serie de razones iguales es igual a

$$\frac{dx + dp_1}{p_1 + x} = \frac{d(x + p_1)}{x + p_1}$$

la cual igualada con la última da

$$\frac{d(x + p_1)}{x + p_1} = \frac{dq_1}{q_1}; \qquad x + p_1 = -r q_1$$

la cual, junto con la ecuación anterior, nos proporciona el sistema en p y q.

$$\left.\begin{aligned} x + p_1 &= -r q_1 \\ p_1^2 - q_1^2 + p_1 x - q_1 y &= 0 \end{aligned}\right\}$$

y resolviéndolo se tiene

$$p_1 = \frac{x - ry}{r^2 - 1}, \qquad q_1 = \frac{y - rx}{r^2 - 1}$$

y, por tanto,

$$du = \frac{x - ry}{r^2 - 1} dx + \frac{y - rx}{r^2 - 1} dy$$

de donde

$$u = \frac{x^2 - 2 rxy + y^2}{2(r^2 - 1)} + b; \qquad z = K e^{-\frac{1}{2(1-r^2)}(x^2 - 2 rxy + y^2)}$$

que es la integral completa con las dos constantes k y r.

Para que represente una superficie de probabilidad debe ser

$$\int\int_{-\infty}^{+\infty} z\,dx\,dy = 1$$

y, por tanto,

$$1 = \mathrm{K} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy \int_{-\infty}^{\infty} e^{-\frac{(x-ry)^2}{2(1-r^2)}} dx$$

de donde

$$1 = \mathrm{K}\sqrt{2\pi(1-r^2)}\sqrt{2\pi}; \qquad \mathrm{K} = \frac{1}{2\pi\sqrt{1-r^2}}$$

y, por tanto, la superficie es

$$z = \frac{1}{2\pi\sqrt{1-r^2}} e^{-\frac{1}{2(1-r^2)}(x^2 - 2rxy + y^2)}$$

luego la superficie normal es la única superficie de probabilidad que satisface a aquella ecuación en derivadas parciales.

Tratemos de generalizar la ecuación en derivadas parciales obtenida anteriormente, al caso de ser

$$\mathrm{K}(p^2 - q^2) + z(px - qy) = 0$$

Haciendo como en el caso anterior $z = e^u$

$$\mathrm{K}(q_1^2 - q_1^2) + (p_1 x - q_1 y) = 0$$

y la integral completa por el método de Lagrange-Charpit nos la proporciona el sistema

$$\frac{dx}{2\mathrm{K}p_1 + x} = \frac{dy}{-2q_1 - y} = \frac{-dp_1}{p_1} = \frac{-dq_1}{-q_1}$$

que por combinación de la primera y tercera, esta serie de razones es igual a

$$\frac{dx + \mathrm{K}\,dp_1}{\mathrm{K}p_1 + x} = \frac{d(x + \mathrm{K}p_1)}{x + \mathrm{K}p_1}; \qquad \frac{d(x + \mathrm{K}p_1)}{x + \mathrm{K}p_1} = \frac{dq_1}{q_1}$$

de donde $e + \mathrm{K}p_1 = -rq_1$, obteniéndose el sistema

$$\left.\begin{aligned} \mathrm{K}(p_1^2 - q_1^2) + p_1 x - q_1 y &= 0 \\ x + \mathrm{K}p_1 &= -rq_1 \end{aligned}\right\}$$

que resuelto nos da

$$p_1 = \frac{Kx - ry}{r^2 - K^2} \qquad q_1 = \frac{Ky - rx}{r^2 - K^2}$$

y por tanto

$$du = \frac{Kx - ry}{r^2 - K^2} dx + \frac{Ky - rx}{r^2 - K^2} dy; \qquad u = \frac{Kx^2 - 2rxy + Ky^2}{2(r^2 - K^2)} + b$$

y finalmente

$$z = me^{-\frac{1}{2(K^2-r^2)}(Kx^2-2rxy+Ky^2)}$$

que con la condición para que represente una superficie de probabilidad

$$1 = m\int_{-\infty}^{\infty} e^{-\frac{y^2}{2K}} dy \int_{-\infty}^{\infty} e^{-\frac{K\left(x-\frac{ry}{K}\right)^2}{2(K^2-r^2)}} dx, \qquad m = \frac{1}{2\pi\sqrt{K^2 - r^2}}$$

y queda

$$z = \frac{1}{2\pi\sqrt{K^2 - r^2}} e^{-\frac{1}{2(K^2-r^2)}(Kx^2 - 2rxy + Ky^2)}$$

la cual poniéndola en la forma

$$z = \frac{1}{2\pi\sqrt{K}\sqrt{K}\sqrt{1 - \frac{r^2}{K^2}}} e^{-\frac{1}{2\left(1-\frac{r^2}{K^2}\right)}\left(\frac{x^2}{K} - \frac{2rxy}{K^2} + \frac{y^2}{K}\right)}$$

nos dice que es tambiéén una superficie normal con parámetros $\sigma_1 = \sigma_2 = \sqrt{K}$ y $\rho = \frac{r}{K}$.

BIBLIOGRAFIA

CRAMER.—Mathematical Methods of Statistics.
KENDALL.—The advanced theory of Statistics.
WILKS.—Mathematical Statistics.
JORDAN.—Statistique mathematique.
RISSER ET TRAYNARD.—Les principes de la statistique mathematique.
LA VALLEE POUSSIN.—Cours d'Analyse.
HUMBERT.—Cours d'Analyse.
GOURSAT.—Cours d'Analyse mathematique.
PEARSON.—The fifteen constant bivariate frecuency surface. Biometrika, 17.
ROMANOVSKY.—Generalisation of some types of the frecuency curves of Professor Pearson. Biometrika, 16.