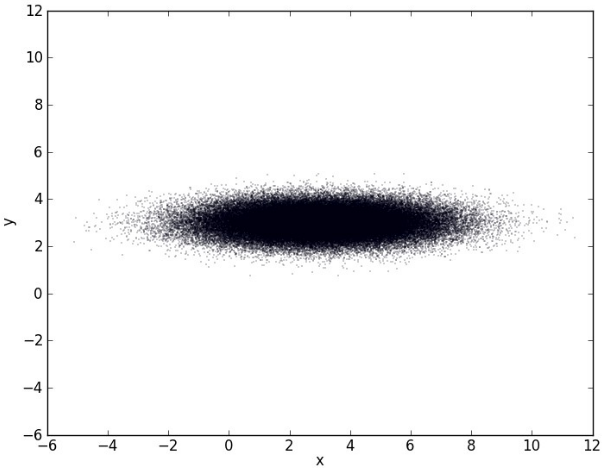
Tema 4.2.Distribución binormal y correlación

Generar 100000 sucesos de valores *(x,y)* distribuidos según una distribución binormal no correlacionada con:

$$\Upsilon_{x1} \Xi 3, \;\; ,_{x1} \Xi 2, \; \Upsilon_{x2} \Xi 3, \;\; ,_{x2}$$

a)Representar los puntos en un histograma de dos dimensiones y calcular la matriz de covarianza y el coeficiente de correlación.



$$La\ matiz\ de\ covarianza\ está: V_{ij} = E[(x_i + \mu_i)(x_j + \mu_j)]$$

Con los datos obtenemos:

$$V_{ij} = \begin{pmatrix} 3.97425102 & 0.00493506 \\ 0.00493506 & 0.2505779 \end{pmatrix}$$

con un coeficiente de correlación: $\rho(x_i, x_j) = \dfrac{V_{ij}}{(V_{ii} V_{jj})^{\frac{1}{2}}} = 0.00494530475502 \approx 0$

b) Contar los puntos *(x,y)* que pertenecen a sucesivas elipses cuyos semiejes vienen dados por un número entero de desviaciones típicas $,_{x1} \;\; ,_{x2}$ y comparar con los porcentajes para una distribución normal unidimensional.

Los puntos dentro de la elipse son:

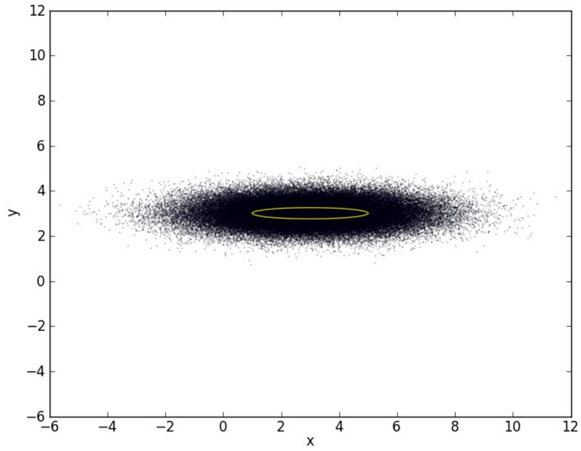
$$1\sigma \rightarrow 39403 \rightarrow 0.39403\%\ de\ los\ datos$$
$$2\sigma \rightarrow 86538 \rightarrow 0.86538\%\ de\ los\ datos$$
$$3\sigma \rightarrow 98886 \rightarrow 0.98886\%\ de\ los\ datos$$
$$4\sigma \rightarrow 99967 \rightarrow 0.99967\%\ de\ los\ datos$$

En comparacíon con los porcentajes para una distribución normal unidimensional:

$$P = \iint_{\sigma_1}^{\sigma_2} \frac{1}{2\pi} e^{\left[\left(\frac{x_1 - 3}{2}\right)^2 + \left(\frac{x_2 - 3}{0.5}\right)^2\right]} dx\, dy$$

$$\text{Obtenemos:}\quad \begin{array}{l} 1\,\sigma \rightarrow 0.46607\,\% \\ 2\,\sigma \rightarrow 0.91107\,\% \\ 3\,\sigma \rightarrow 0.99461\,\% \\ 4\,\sigma \rightarrow 0.99987\,\% \end{array}$$

Reconocemos que la aproximación con elipses no está suficiente para obtener los valores corectos.



c) Considerar la función *f(x,y) = 3x +5y*. Calcular la varianza de *f(x,y)* directamente usando propagación de errores y también a partir de los datos.

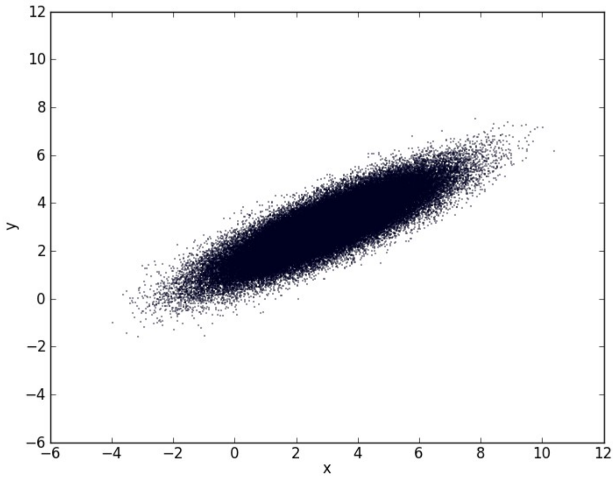
La varianza de f(x,y) de los datos es: 42.13

Con la formular de propagacion de errores podemos calcular la varianza tambien:

$$Var = (3\,\sigma_1)^2 + (5\,\sigma_2)^2$$

La varianza de f(x,y) de la propagacion de errores es: 42.14

d) Rotar ahora los sucesos respecto al centro de la distribución, no del origen, un ángulo θ = *30º* y repetir los pasos anteriores.



$$V_{ij}=\begin{pmatrix} 3.06710155 & 1.62735699 \\ 1.62735699 & 1.1887576 \end{pmatrix}$$

$$\rho=0.8522$$

$1\sigma\rightarrow 49693\rightarrow 0.49693\%\, de\, los\, datos$		$1\sigma\rightarrow 0.57734\%$
$2\sigma\rightarrow 85077\rightarrow 0.85077\%\, de\, los\, datos$		$2\sigma\rightarrow 0.93188\%$
$3\sigma\rightarrow 97105\rightarrow 0.97105\%\, de\, los\, datos$	con la formula	$3\sigma\rightarrow 0.99555\%$
$4\sigma\rightarrow 99635\rightarrow 0.99635\%\, de\, los\, datos$		$4\sigma\rightarrow 0.99989\%$

Para la varianza con la función obtenemos: 105.276
Con la formula que cambia en $V[f'(x',y')]=[(3\cos\theta+5\sin\theta)]^2+[(3\cos\theta-5\sin\theta)]^2$

obetenemos una varianza de: 105.96

e) A partir de los datos correlacionados estimar el ángulo θ y los valores de las $,_{u}$ $,_{v}$.

Los valores son: $\sigma_u=3.070$
$\sigma_v=1.186$

Con esos podemos calcular el angulo: $\sin\theta=\frac{(\rho\,\sigma_u\sigma_v)}{\sigma_{x_1}^2-\sigma_{x_2}^2}\rightarrow\theta=29{,}95$