＜ 2 次元正規分布の周辺分布 ＞

$(X,\ Y)$ を一般の 2 次元正規分布に従う確率変数とする (p15 定理 2)。

確率密度関数 $p(x,\ y)$ は

$$p(x,\ y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-\frac{Q(x,\ y)}{2}}$$

$$Q(x,\ y) = \frac{1}{1-\rho^2}\left\{\left(\frac{x-m_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-m_1}{\sigma_1}\right)\left(\frac{y-m_2}{\sigma_2}\right) + \left(\frac{y-m_2}{\sigma_2}\right)^2\right\}$$

$$= \frac{1}{(1-\rho^2)\sigma_2^2}\left\{(y-m_2)^2 - 2\rho\frac{\sigma_2}{\sigma_1}(x-m_1)(y-m_2) + \left(\frac{\sigma_2}{\sigma_1}\right)^2(x-m_1)^2\right\}$$

$$= \frac{1}{(1-\rho^2)\sigma_2^2}\left[\{(y-m_2) - \rho\frac{\sigma_2}{\sigma_1}(x-m_1)\}^2 + (1-\rho^2)\left(\frac{\sigma_2}{\sigma_1}\right)^2(x-m_1)^2\right]$$

$$= \frac{1}{(1-\rho^2)\sigma_2^2}\left\{(y-m_2) - \rho\frac{\sigma_2}{\sigma_1}(x-m_1)\right\}^2 + \frac{(x-m_1)^2}{\sigma_1^2}$$

となる。よって $X = x$ が起こったとき $Y = y$ の起こる条件付確率密度は

$$p(Y = y \mid X = x) = \frac{p(x,\ y)}{p_X(x)} = \frac{\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-\frac{Q(x,\ y)}{2}}}{\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-m_1)^2}{2\sigma_1^2}}}$$

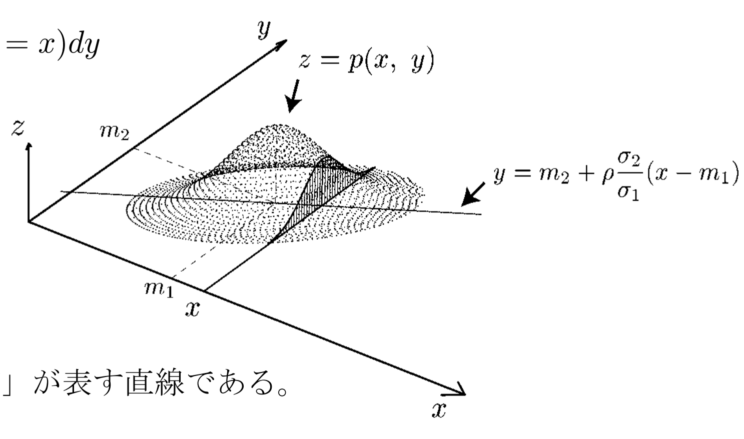
$$= \frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}}e^{-\frac{1}{2\sigma_2^2(1-\rho^2)}\{y-m_2-\rho\frac{\sigma_2}{\sigma_1}(x-m_1)\}^2}$$

である。これは平均 $m_2 + \rho\frac{\sigma_2}{\sigma_1}(x-m_1)$，分散 $\sigma_2^2(1-\rho^2)$ の

1 次元正規分布密度である。従って $X = x$ が起こったとき，Y の平均は

$$E[Y \mid X = x] = \int_{-\infty}^{\infty} yp(Y = y \mid X = x)dy$$

$$= m_2 + \rho\frac{\sigma_2}{\sigma_1}(x-m_1)$$



となる。この直線

$$y = m_2 + \rho\frac{\sigma_2}{\sigma_1}(x-m_1)$$

は「$X = x$ のときの Y の条件付平均値」が表す直線である。