TWO–DIMENSIONAL RANDOM VARIABLE

A PAIR OF RVs: X, Y

CUMULATIVE DISTRIBUTION FUNCTION:

$$F(x,y) = \mathcal{P}(X \le x,\, Y \le y) = \begin{cases} \sum_{x_i \le x, y_k \le y} \mathcal{P}(X = x_i, Y = y_k) & \textit{(discrete RV)} \\ \int_{-\infty}^{x} \int_{-\infty}^{y} f(x,y)\, dx dy & \textit{(continuous RV)} \end{cases}$$

for a RV of discrete type we may define:

$$p_{ik} = \mathcal{P}(X = x_i,\, Y = y_k)$$

and for a RV of continuous type we have:

$$f(x,y) = \frac{\partial^2 F}{\partial x \partial y}; \qquad \mathcal{P}(X \in [x, x{+}dx] \bigcap Y \in [y, y{+}dy]) = f(x,y)\, dx dy$$

Marginal Distribution Functions:

$$\mathcal{P}(a \le x \le b;\ y\ \text{ any value }) = \mathcal{P}(a \le x \le b; -\infty \le y \le \infty)$$
$$= \int_a^b \left[\int_{-\infty}^{\infty} f(x,y)dy\right] dx \equiv \int_a^b g(x)dx$$

$$g(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$$

in an exactly analogous way:

$$h(y) = \int_{-\infty}^{\infty} f(x,y)\,dx$$

we call $g(x),\ h(y)$ marginal distribution functions of x and y , respectively. Their role is exactly the same as the role of the pdf of a single RV.

conditional distributions:

$$f(y|x_0) \overset{*}{=} \frac{f(x_0, y)}{\int_{-\infty}^{\infty} f(x_0, y)\, dy} = \frac{f(x_0, y)}{g(x_0)}$$

$$f(x|y_0) \overset{*}{=} \frac{f(x, y_0)}{\int_{-\infty}^{\infty} f(x, y_0)\, dx} = \frac{f(x, y_0)}{h(y_0)}$$

Note: the * equalities follow from the normalisation condition, i.e.:

$$\int_{-\infty}^{\infty} f(y|x_0)\, dy = 1 = \int_{-\infty}^{\infty} f(x|y_0)\, dx$$

Now, let's put simply $y_0 = y$ and $x_0 = x$. We have:

$$h(y) = \int_{-\infty}^{\infty} f(x, y)\, dx = \int_{-\infty}^{\infty} f(y|x) g(x) dx$$

$$g(x) = \int_{-\infty}^{\infty} f(x, y)\, dy = \int_{-\infty}^{\infty} f(x|y) h(y) dy$$

NOW IF RVs X AND Y ARE INDEPENDENT WE HAVE:

$$f(y|x) = h(y) \quad f(x|y) = g(x)$$

conditional distributions:

On the other hand we have:

$$f(y|x) = \frac{f(x,y)}{g(x)} = h(y)$$

so

$$f(x,y) = g(x) \cdot h(y)$$

for the **independent** *RVs X and Y the joint probability (density) function is a product of two corresponding marginal (density) distributions !*
Note: for a discrete RV we have also marginal distributions.
Let $p_{ik} = P(X = x_i; Y = y_k)$; $\sum_{i,k} p_{ik} = 1$. Then the marginal probability for X, $p_{i\cdot}$ and Y, $p_{\cdot k}$ will be defined, respectively, as:

$$\begin{aligned} p_{i\cdot} &= P(X = x_i; Y = \text{ any value}) \\ p_{\cdot k} &= P(Y = y_k; X = \text{ any value}) \end{aligned}$$

Of course, we have:

$$\sum_i p_{i\cdot} = \sum_k p_{\cdot k} = 1.$$

THE PARAMETERS OF A 2D RV (X,Y)

(case of a continuous variable):

$$E\{H(X,Y)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H(x,y)f(x,y)\,dxdy$$

The moments:

$$\lambda_{lm} = E\{X^l Y^m\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^l y^m\, f(x,y)\,dxdy$$

$$\alpha_{lm} = E\{(X-a)^l (Y-b)^m\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-a)^l (y-b)^m\, f(x,y)\,dxdy$$

THE PARAMETERS OF A 2D RV (X,Y), cntd.

the expected (mean) value of RV X:

$$E\{X\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f(x,y)\,dxdy = \lambda_{10} = \int_{-\infty}^{\infty} x\,g(x)dx$$

$$E\{Y\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f(x,y)\,dxdy = \lambda_{01} = \int_{-\infty}^{\infty} y\,h(y)dy$$

central moments:

$$a = \lambda_{10} \quad b = \lambda_{01}$$

$$\mu_{lm} = E\left\{(X-\lambda_{10})^l (Y-\lambda_{01})^m\right\}$$

$$\mu_{00} = 1; \quad \mu_{10} = \mu_{01} = 0$$

$$\mu_{11} = COV(X,Y)$$

$$\mu_{20} = VAR(X)$$

$$\mu_{02} = VAR(Y)$$

THE COVARIANCE AND CORRELATION OF A 2D RV:

$$COV(X,Y) = E\{(X-\mu_X)((Y-\mu_Y)\} = \ldots = E\{XY\} - \mu_X\mu_Y$$

note (and remember): $COV(X,X) = VAR(X)$
CORRELATION

$$\rho(X,Y) \equiv CORR(X,Y) = \frac{COV(X,Y)}{\sigma(X)\sigma(Y)}$$

it's very easy to show that

$$-1 \leq \rho \leq +1$$

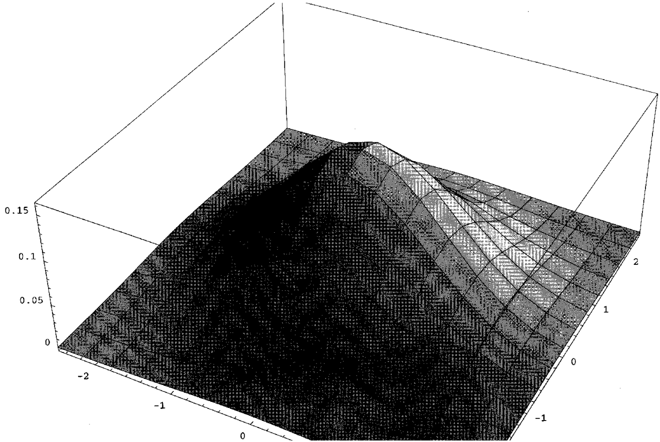
FOR 2 INDEPENDENT RVs:

$$f(x,y) = g(x)h(y)$$

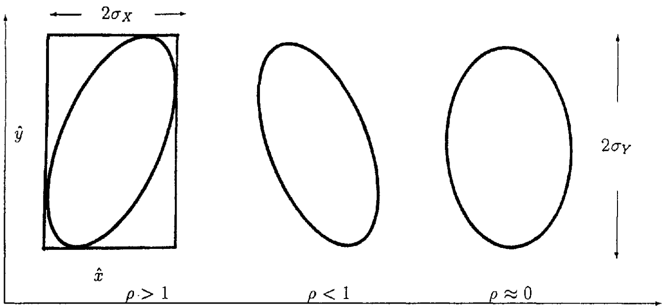
$$COV(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\hat{x})(y-\hat{y})g(x)h(y)\,dxdy = \ldots = 0!$$

independent variables cannot be correlated, but the reciprocal conjecture is false!!

bivariate normal distribution



bivariate normal distribution



bivariate discrete distribution

suppose we have a pair of RVs: X and Y. X – takes on the values: $0, 1, \ldots, 9$ and Y – $1, 2, 3, 4$ and 5. The data look like this:

Y X	0	1	2	3	4	5	6	7	8	9
1	0.0129	0.0149	0.0165	0.0175	0.0178	0.0175	0.0165	0.0149	0.0129	0.0108
2	0.0188	0.0217	0.024	0.0254	0.026	0.0254	0.024	0.0217	0.0188	0.0157
3	0.0214	0.0246	0.0272	0.0288	0.0294	0.0288	0.0272	0.0246	0.0214	0.0178
4	0.0188	0.0217	0.024	0.0254	0.026	0.0254	0.024	0.0217	0.0188	0.0157
5	0.0129	0.0149	0.0165	0.0175	0.0178	0.0175	0.0165	0.0149	0.0129	0.0108

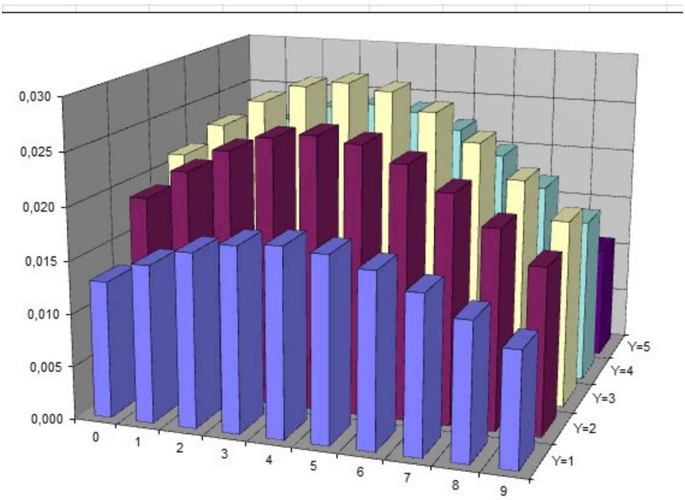
the marginal distribution $g(x)$:

g(x)	0.0848	0.0978	0.1082	0.1146	0.1170	0.1146	0.1082	0.0978	0.0848	0.0708

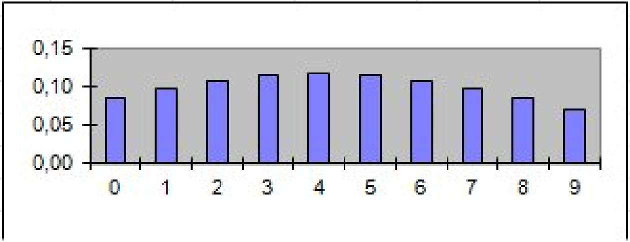
and the marginal distribution $h(y)$:

h(y)	0.1522	0.2215	0.2512	0.2215	0.1522

bivariate discrete distribution cntd.



marginal distribution of X



marginal distribution of Y