4.D.4. Jointly Distributed Stochastic Variables

Stochastic variables $\boldsymbol{X} = \{X_j;\ j = 1, \ldots, n\}$ are jointly distributed if they are defined on the same sample space S.

The definition (4.11) of §4.D.1 can be generalized to give the joint accumulative distribution function for $\boldsymbol{X} = \{X_j\}$ as

$$F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} d y_1 \ldots \int_{-\infty}^{x_n} d y_n\, P_{X_1,\ldots,X_n}(y_1, \ldots, y_n) \tag{4.25}$$

$$= \text{Prob}\left[\left(X_1 < x_1\right) \,\&\&\, \ldots \,\&\&\, \left(X_n < x_n\right)\right] \tag{4.23}$$

= probability of X_j having realization less than x_j, for all j, simultaneously.

where && is the logical operator "AND".

As with any multi-variate problems, formulae in this section can be written more succinctly in vector form. For example, (4.25) can be written as

$$F_{\boldsymbol{X}}(\boldsymbol{x}) = \int_{-\infty}^{x_1} d y_1 \ldots \int_{-\infty}^{x_n} d y_n\, P_{\boldsymbol{X}}(\boldsymbol{y}) \tag{4.25a}$$

which also shows the limits of the vector formulation, namely, we cannot write the integrals as $\int_{-\infty}^{\boldsymbol{x}} d\boldsymbol{y}$ since its meaning is not commonly recognized.

(4.11) implies

$$P_X(x) = \frac{d F_X(x)}{d x}$$

which can be generalized to give

$$P_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \frac{\partial^n F_{X_1,\ldots,X_n}(x_1, \ldots, x_n)}{\partial x_1 \ldots \partial x_n} \tag{4.24}$$

or

$$P_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{\partial^n F_{\boldsymbol{X}}(\boldsymbol{x})}{\partial x_1 \ldots \partial x_n} \tag{4.24a}$$

where

$$P_{X_1,\ldots,X_n}(x_1, \ldots, x_n)\, d x_1 \ldots d x_n = \text{Prob}\left[\left(x_1 \le X_1 < x_1 + d x_1\right) \,\&\& \ldots \&\&\, \left(x_n \le X_n < x_n + d x_n\right)\right] \tag{4.24b}$$

= probability of X_j having a realization in interval $(x_j,\ x_j + d x_j)$, for all j, simultaneously.

$= P_{\boldsymbol{X}}(\boldsymbol{x})\, d^n x$

For the sake of clarity, we switch to the special case of two stochastic variables $\{X, Y\}$.

(4.24-5) thus simplify to

$$P_{XY}(x, y) = \frac{\partial^2 F_{XY}(x, y)}{\partial x \partial y} \tag{4.24c}$$

$$F_{XY}(x, y) = \int_{-\infty}^{x} d x' \int_{-\infty}^{y} d y'\, P_{XY}(x' y') = \text{Prob}\left[\left(X < x\right) \&\& \left(Y < y\right)\right] \tag{4.25a}$$

(4.25a) implies

$$F_{XY}(-\infty, y) = 0$$

$$= F_{XY}(x, -\infty) = F_{XY}(-\infty, -\infty) \tag{4.25b}$$

Applying the rule of addition, we have

$$P_X(x) = \int_{-\infty}^{\infty} dy\, P_{XY}(x, y) \qquad \& \qquad P_Y(y) = \int_{-\infty}^{\infty} dx\, P_{XY}(x, y) \tag{4.29}$$

Putting these back into (4.25a) gives

$$F_{XY}(x, \infty) = \int_{-\infty}^{x} dx'\, P_X(x') = F_X(x) \,\& \qquad F_{XY}(\infty, y) = F_Y(y) \tag{4.28}$$

$$\rightarrow \quad F_{XY}(\infty, \infty) = \int_{-\infty}^{\infty} dx'\, P_X(x') = 1 \tag{4.28a}$$

The n^{th} moments are

$$\langle x^n \rangle = \int_{-\infty}^{\infty} dx\, P_X(x)\, x^n$$

$$= \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\, P_{XY}(x, y)\, x^n \qquad \text{[(4.29) used.]} \tag{4.30}$$

$$\langle y^n \rangle = \int_{-\infty}^{\infty} dy\, P_Y(y)\, y^n$$

$$= \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dx\, P_{XY}(x, y)\, y^n \qquad \text{[(4.29) used.]} \tag{4.30a}$$

which can be combined to give

$$\langle x^m y^n \rangle = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\, P_{XY}(x, y)\, x^m y^n \tag{4.31}$$

Similarly, the variances are

$$\text{Var}(X) = \int_{-\infty}^{\infty} dx\, P_X(x) \left(x - \langle x \rangle \right)^2$$

$$= \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\, P_{XY}(x, y) \left(x - \langle x \rangle \right)^2 = \left\langle \left(x - \langle x \rangle \right)^2 \right\rangle$$

$$\text{Var}(Y) = \int_{-\infty}^{\infty} dy\, P_Y(y) \left(y - \langle y \rangle \right)^2$$

$$= \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dx\, P_{XY}(x, y) \left(y - \langle y \rangle \right)^2 = \left\langle \left(y - \langle y \rangle \right)^2 \right\rangle$$

which suggests the covariance

$$\text{Cov}(X, Y) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\, P_{XY}(x, y) \left(x - \langle x \rangle \right) \left(y - \langle y \rangle \right)$$

$$= \left\langle \left(x - \langle x \rangle \right) \left(y - \langle y \rangle \right) \right\rangle \tag{4.32a}$$

$$= \left\langle x y - \langle x \rangle y - x \langle y \rangle + \langle x \rangle \langle y \rangle \right\rangle$$

$$= \langle x y \rangle - \langle x \rangle \langle y \rangle \tag{4.32}$$

and its normalized version, the correlation function

$$\text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} \tag{4.33}$$

$$= \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\, \text{Var}(Y)}}$$

$$= \frac{\left\langle \left(x - \langle x \rangle \right) \left(y - \langle y \rangle \right) \right\rangle}{\sqrt{\left\langle \left(x - \langle x \rangle \right)^2 \right\rangle \left\langle \left(y - \langle y \rangle \right)^2 \right\rangle}} \tag{4.33a}$$

Some properties of $\text{Cor}(X, Y)$ are

(i) $\text{Cor}(X, Y) = \text{Cor}(Y, X)$ [$X \leftrightarrow Y$ in (4.32) used.]

(ii) $\text{Cor}(X, X) = 1$ & $\text{Cor}(X, -X) = -1$ [(4.33a) used.]

(iii) $-1 \le \text{Cor}(X, Y) \le 1$

Proof: (4.32a) gives

$$\left|\text{Cov}(X, Y)\right| \le \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\, P_{XY}(x, y) \left|x - \langle x\rangle\right| \left|y - \langle y\rangle\right| \quad \text{[Schwarz's inequality used.]} \qquad (4.33b)$$

where the equal sign applies if $\left|x - \langle x\rangle\right| = \left|y - \langle y\rangle\right|$. In which case,

$$|\text{Cov}(X, X)| = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\, P_{XY}(x, y) \left(x - \langle x\rangle\right)^2$$

$$= \text{Var}(X) = \text{Var}(Y) = \sqrt{\text{Var}(X)\,\text{Var}(Y)}$$

Putting this back into (4.33b) gives

$$|\text{Cor}(X, Y)| \le 1 \qquad \text{QED.}$$

(iv) $\text{Cor}(a X + b,\ c Y + d) = \text{Cor}(X, Y)$ $a, b, c, d = \text{consts}$ & $a, c \neq 0$

Proof: Let

$$Z = a X + b \qquad z = a x + b$$

then (4.12) gives

$$P_Z(z) = \int_{-\infty}^{\infty} dx\, P_X(x)\, \delta\left[z - (a x + b)\right]$$

$$= \int_{-\infty}^{\infty} dx\, P_X(x) \frac{1}{|a|} \delta\left[x - \frac{1}{a}(z - b)\right]$$

$$= \frac{1}{|a|} P_X\left[\frac{1}{a}(z - b)\right]$$

$$\therefore \quad \langle z^n \rangle = \int_{-\infty}^{\infty} dz\, P_Z(z)\, z^n$$

$$= \frac{1}{|a|} \int_{-\infty}^{\infty} dz\, P_X\left[\frac{1}{a}(z - b)\right] z^n$$

$$= \int_{-\infty}^{\infty} dx\, P_X(x) \left(a x + b\right)^n$$

$$= \left\langle \left(a x + b\right)^n \right\rangle$$

which could have been obtained directly by treating z^n as a function of x.

Thus,

$$\text{Var}(Z) = \left\langle \left(a x + b - \langle a x + b \rangle\right)^2 \right\rangle$$

$$= a^2 \left\langle \left(x - \langle x \rangle\right)^2 \right\rangle$$

$$= a^2\, \text{Var}(X)$$

Similarly,

$$\text{Var}(c y + d) = c^2\, \text{Var}(Y)$$

$$\text{Cov}(a X + b,\ c Y + d) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\, P_{XY}(x, y) \left(a x + b - \langle a x + b \rangle\right)\left(c y + d - \langle c y + d \rangle\right)$$

$$= \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\, P_{XY}(x, y)\; a c \left(x - \langle x \rangle\right)\left(y - \langle y \rangle\right)$$

$$= a c\, \text{Cov}(X, Y)$$

$$\to \quad \mathrm{Cor}(a X+b,\ c Y+d)=\frac{a\,c\,\mathrm{Cov}(X,Y)}{\sqrt{a^2\,\mathrm{Var}(X)\,c^2\,\mathrm{Var}(Y)}}=\mathrm{Cor}(X,Y) \qquad \text{QED}$$

In case X & Y are statistically independent, we have

(*i*′) $P_{XY}(x,y)=P_X(x)\,P_Y(y)$ [Definition (4.4) used.]

(ii′) $\langle x\,y\rangle=\langle x\rangle\langle y\rangle$

Proof: $\langle x\,y\rangle=\int_{-\infty}^{\infty}dx\int_{-\infty}^{\infty}dy\,P_{XY}(x,y)\,x\,y$

$=\int_{-\infty}^{\infty}dx\int_{-\infty}^{\infty}dy\,P_X(x)\,P_Y(y)\,x\,y$ [(i') used.]

$=\langle x\rangle\langle y\rangle$ QED

(iii′) $\langle (x+y)^2\rangle-\langle (x+y)\rangle^2=\langle x^2\rangle-\langle x\rangle^2+\langle y^2\rangle-\langle y\rangle^2$

Proof: $\langle (x+y)^2\rangle-\langle (x+y)\rangle^2=\langle (x^2+2\,x\,y+y^2)\rangle-\left(\langle x\rangle+\langle y\rangle\right)^2$

$=\langle x^2\rangle+2\langle x\rangle\langle y\rangle+\langle y^2\rangle-\left(\langle x\rangle^2+2\langle x\rangle\langle y\rangle+\langle y\rangle^2\right)$

$=\langle x^2\rangle+\langle y^2\rangle-\langle x\rangle^2-\langle y\rangle^2$ QED.

(iv′) $\mathrm{Cor}(X,Y)=0$

Proof: $\mathrm{Cov}(X,Y)=\int_{-\infty}^{\infty}dx\int_{-\infty}^{\infty}dy\,P_{XY}(x,y)\,(x-\langle x\rangle)(y-\langle y\rangle)$ [(4.32a) used.]

$=\int_{-\infty}^{\infty}dx\int_{-\infty}^{\infty}dy\,P_X(x)\,P_Y(y)\,(x-\langle x\rangle)(y-\langle y\rangle)$ [(i') used.]

$=0$ QED.

Caution: $\mathrm{Cor}(X,Y)=0$ does not imply X & Y are independent.

The generalization of (4.12) to the case $Z=G(X,Y)$ is obviously

$$P_Z(z)=\int_{-\infty}^{\infty}dx\int_{-\infty}^{\infty}dy\,P_{XY}(x,y)\ \delta[z-G(x,y)] \tag{4.34}$$

so that the characteristic function (4.15) becomes

$$f_Z(k)=\int_{-\infty}^{\infty}dz\,e^{ikz}\int_{-\infty}^{\infty}dx\int_{-\infty}^{\infty}dy\,P_{XY}(x,y)\ \delta[z-G(x,y)]$$

$$=\int_{-\infty}^{\infty}dx\int_{-\infty}^{\infty}dy\,P_{XY}(x,y)\,e^{ikG(x,y)} \tag{4.35}$$

On the other hand, generalization of the characteristic function itself to the multi-variable case gives

$$f_{X_1\ldots X_n}(k_1,\ \ldots,k_n)=\int_{-\infty}^{\infty}dx_1\ \ldots\int_{-\infty}^{\infty}dx_n\,P_{X_1\ldots X_n}(x_1,\ \ldots x_n)\ e^{i(k_1x_1+\ldots+k_nx_n)} \tag{4.36}$$

or

$$f_{\boldsymbol{X}}(\boldsymbol{k})=\int d^n x\,P_{\boldsymbol{X}}(\boldsymbol{x})\,e^{i\boldsymbol{k}\cdot\boldsymbol{x}} \tag{4.36a}$$

so that the joint moments are given by

$$\langle x_1^{m_1}\ldots x_p^{m_p}\rangle=(-i)^{m_1+\ldots+m_p}\frac{\partial^{m_1}}{\partial k_1^{m_1}}\ldots\frac{\partial^{m_p}}{\partial k_p^{m_p}}f_{X_1\ldots X_n}(k_1,\ \ldots,k_n)\Big|_{k_1=\ldots=k_n=0} \qquad p\le n \tag{4.37}$$

$$=(-i)^{m_1+\ldots+m_p}\frac{\partial^{m_1}}{\partial k_1^{m_1}}\ldots\frac{\partial^{m_p}}{\partial k_p^{m_p}}f_{\boldsymbol{X}}(\boldsymbol{k})\Big|_{\boldsymbol{k}=0}$$

Ex.4.7.

Show that $\mathrm{Cor}(X, Y)$ is a measure of the degree to which X depends on Y.

Answer

From linear algebra, two functions $X(x)$ & $Y(x)$ are linearly independent if

$$a\,X + b\,Y = 0 \qquad \forall\, x \tag{1}$$

has only the trivial solution

$$a = b = 0 \tag{1a}$$

The analog for two stochastic variables X & Y is that they are linearly independent if

$$\left\langle \left[a\,(x - \langle x \rangle) + b\,(y - \langle y \rangle) \right]^2 \right\rangle = 0 \tag{2}$$

has only the trivial solution

$$a = b = 0 \tag{2a}$$

Central moments are used since our results must be coordinate (or realization) independent. The square is used because we want to turn the problem into a variational one.

Consider now the positive quantity

$$e = \left\langle \left[a\,(x - \langle x \rangle) + b\,(y - \langle y \rangle) \right]^2 \right\rangle \geq 0 \tag{3}$$

Finding the solution of (2) is then equivalent to finding the minimum of e.

Now,

$$e = a^2 \left\langle (x - \langle x \rangle)^2 \right\rangle + 2\,a\,b \left\langle (x - \langle x \rangle)(y - \langle y \rangle) \right\rangle + b^2 \left\langle (y - \langle y \rangle)^2 \right\rangle$$

$$= a^2\,\mathrm{Var}(X) + 2\,a\,b\,\mathrm{Cov}(X, Y) + b^2\,\mathrm{Var}(Y) \tag{4}$$

The extrema of e with respect to (a, b) are given by

$$\delta e = 0 \;= 2\left(a\,\mathrm{Var}(X) + b\,\mathrm{Cov}(X, Y) \right) \delta a + 2\left(a\,\mathrm{Cov}(X, Y) + b\,\mathrm{Var}(Y) \right) \delta b \tag{5}$$

If the extremum is a minimum,

$$\delta^2 e = 2\,\mathrm{Var}(X)\,(\delta a)^2 + 4\,\mathrm{Cov}(X, Y)\,\delta a\,\delta b + 2\,\mathrm{Var}(Y)\,(\delta b)^2 \;> 0 \tag{5a}$$

Note:

$\delta^n e$ = changes in e due to the n^{th}-power products of δa & δb.

Since we are looking for the variation of e due to changes δa & δb in a & b, respectively, terms $\delta^n a$ or $\delta^n b$ with $n \geq 2$ are meaningless in this context.

Since δa & δb are arbitrary, (5) can be satisfied for all δa & δb if and only if

$$a\,\mathrm{Var}(X) + b\,\mathrm{Cov}(X, Y) = 0 \qquad \& \qquad a\,\mathrm{Cov}(X, Y) + b\,\mathrm{Var}(Y) = 0$$

which can have a non-trivial solution if

$$\begin{vmatrix} \mathrm{Var}(X) & \mathrm{Cov}(X, Y) \\ \mathrm{Cov}(X, Y) & \mathrm{Var}(Y) \end{vmatrix} = 0 \;= \mathrm{Var}(X)\,\mathrm{Var}(Y) - \mathrm{Cov}(X, Y)^2$$

$$= \mathrm{Var}(X)\,\mathrm{Var}(Y) \left[1 - \mathrm{Cor}(X, Y)^2 \right] \qquad [\,(4.33)\text{ used.}\,] \tag{6}$$

Thus

$$\text{Cor}(X, Y) = \pm 1 \tag{6a}$$

are extrema of e.

Putting (6a) into (5a) gives

$$\begin{aligned}\delta^2 e &= 2\,\text{Var}(X)\,(\delta a)^2 \pm 4\sqrt{\text{Var}(X)\,\text{Var}(Y)}\;\delta a\,\delta b + 2\,\text{Var}(Y)\,(\delta b)^2 \\ &= 2\left(\sigma_X\,\delta a \pm \sigma_Y\,\delta b\right)^2 \\ &> 0\end{aligned} \tag{6b}$$

Thus, (6a) are minima of e, and hence non-trivial solutions of (2). In other word,

$$|\,\text{Cor}(X, Y)\,| = 1 \quad \rightarrow \quad X \,\&\, Y \text{ are linearly dependent.} \tag{7}$$

For $\text{Cor}(X, Y) \neq \pm 1$, (2) has only the trivial solution $a = b = 0$. Therefore,

$$|\,\text{Cor}(X, Y)\,| \neq 1 \quad \rightarrow \quad X \,\&\, Y \text{ are linearly independent.} \tag{7a}$$

According to property (iii),

$$|\,\text{Cor}(X, Y)\,| \leq 1$$

Therefore, (7a) reduces to

$$|\,\text{Cor}(X, Y)\,| < 1 \quad \rightarrow \quad X \,\&\, Y \text{ are linearly independent.} \tag{7b}$$

Now, (7) can be interpreted as

$$|\,\text{Cor}(X, Y)\,| = 1 \quad \rightarrow \quad X \,\&\, Y \text{ are completely correlated.} \tag{8a}$$

in which case, we can interpret the opposite end as

$$|\,\text{Cor}(X, Y)\,| = 0 \quad \rightarrow \quad X \,\&\, Y \text{ are completely uncorrelated.} \tag{8b}$$

while

$$|\,\text{Cor}(X, Y)\,| < 1 \quad \rightarrow \quad X \,\&\, Y \text{ are partially correlated.} \tag{8c}$$

Caution: The linear independence discussed here is NOT the same as the (statistical) independence defined in (4.4) and property (i').

Ex.4.8.

Let $X \,\&\, Y$ be two (statistically) independent stochastic variables that are Gaussian distributed with

$$\langle x \rangle = \langle y \rangle = 0 \qquad \& \qquad \sigma_X = \sigma_Y = 1 \tag{0}$$

Find the joint distribution function for the stochastic variables

$$V = X + Y \qquad \& \qquad W = X - Y$$

Are $V \,\&\, W$ (statistically) independent?

Answer

According to the specification (0), we have [c.f. Ex.4.4]

$$P_X(x) = \frac{1}{\sqrt{2\pi}}\,e^{-x^2/2} \qquad P_Y(y) = \frac{1}{\sqrt{2\pi}}\,e^{-y^2/2} \tag{1a}$$

Since $X \,\&\, Y$ are independent,

$$P_{XY}(x, y) = P_X(x)\,P_Y(y) = \frac{1}{2\pi}\,e^{-\left(x^2 + y^2\right)/2} \tag{1b}$$

(4.34) then gives [see §Code]

$$P_V(v) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \frac{1}{2\pi} e^{-(x^2+y^2)/2} \delta[v-(x+y)]$$
$$= \int_{-\infty}^{\infty} dx \frac{1}{2\pi} e^{-[x^2+(v-x)^2)]/2} = \frac{1}{2\sqrt{\pi}} e^{-v^2/4}$$
$$P_W(w) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \frac{1}{2\pi} e^{-(x^2+y^2)/2} \delta[w-(x-y)]$$
$$= \int_{-\infty}^{\infty} dx \frac{1}{2\pi} e^{-[x^2+(x-w)^2)]/2} = \frac{1}{2\sqrt{\pi}} e^{-w^2/4}$$

Generalizing (4.34) to

$$P_{VW}(v,w) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\, P_{XY}(x,y)\, \delta[v-G(x,y)] \times \delta[w-H(x,y)] \quad (1c)$$

we get

$$P_{VW}(v,w) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \frac{1}{2\pi} e^{-(x^2+y^2)/2} \delta[v-(x+y)] \times \delta[w-(x-y)] \quad (1d)$$
$$= \int_{-\infty}^{\infty} dx \frac{1}{2\pi} e^{-[x^2+(v-x)^2)]/2} \delta[w-(2x-v)]$$
$$= \frac{1}{4\pi} e^{-(v^2+w^2)/4} \quad (3)$$

Since

$$P_{VW}(v,w) = P_V(v)\, P_W(w)$$

V & W are statistically independent.

Alternatively,

$$\int dv \int dw\, \delta(v)\, \delta(w) = \int dx \int dy\, \delta(x)\, \delta(y) = 1$$
$$= \int dv \int dw\, |J|\, \delta(x)\, \delta(y)$$

→ $\delta(v)\,\delta(w) = |J|\ \delta(x)\,\delta(y)$

or more generally,

$$\rightarrow \quad \delta[v-g(x,y)] \times \delta[w-h(x,y)] = |J|\ \delta[x-G(v,w)] \times \delta[y-H(v,w)] \quad (2)$$

where the Jacobian of the transformation $(v,w) \to (x,y)$ is defined as

$$J = \frac{\partial(x,y)}{\partial(v,w)} = \begin{vmatrix} \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \\ \frac{\partial x}{\partial w} & \frac{\partial y}{\partial w} \end{vmatrix} \quad (2a)$$

Solving

$$v = x + y = g(x,y) \qquad w = x - y = h(x,y)$$

gives

$$x = \frac{1}{2}(v+w) = G(v,w) \qquad y = \frac{1}{2}(v-w) = H(v,w)$$

so that (2a) becomes

$$J = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}$$

and (2) evaluates to

$$\delta[v-(x+y)] \times \delta[w-(x-y)] = \frac{1}{2}\ \delta\left[x - \frac{1}{2}(v+w)\right] \times \delta\left[y - \frac{1}{2}(v-w)\right] \tag{2b}$$

(1d) then becomes

$$P_{VW}(v, w) = \frac{1}{4\pi} e^{-\left[(v+w)^2+(v-w)^2\right]/8}$$
$$= \frac{1}{4\pi} e^{-\left(v^2+w^2\right)/4}$$

in agreement with (3).

Code

In[∘]:= $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)/2}$ **DiracDelta[v - (x + y)] dx dy**

Out[∘]= $e^{-\frac{v^2}{4}}\sqrt{\pi}$

In[∘]:= $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)/2}$ **DiracDelta[w - (x - y)] dx dy**

Out[∘]= $e^{-\frac{w^2}{4}}\sqrt{\pi}$

(* Mathematica fails *)
$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)/2}$ **DiracDelta[v - (x + y)] × DiracDelta[w - (x - y)] dx dy**

Out[∘]= $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{\frac{1}{2}(-x^2-y^2)}$ DiracDelta[v - x - y] × DiracDelta[w - x + y] dx dy

In[∘]:= $\int_{-\infty}^{\infty} e^{-(x^2+(v-x)^2)/2}$ **DiracDelta[w - (2 x - v)] dx**

Out[∘]= ConditionalExpression[$\frac{1}{2} e^{-\frac{v^2}{4}-\frac{w^2}{4}}$, Im[v] + Im[w] == 0]

In[∘]:= $\int_{-\infty}^{\infty} e^{-(x^2+(v-x)^2)/2}$ **dx**

Out[∘]= $e^{-\frac{v^2}{4}}\sqrt{\pi}$

In[∘]:= $\int_{-\infty}^{\infty} e^{-(x^2+(x-w)^2)/2}$ **dx**

Out[∘]= $e^{-\frac{w^2}{4}}\sqrt{\pi}$

Ex. 4.9.

The multivariate Gaussian distribution with zero mean is given by

$$P_{X_1 \ldots X_n}(x_1, \ldots, x_n) = P_{\boldsymbol{X}}(\boldsymbol{x}) = \sqrt{\frac{\det \mathbb{g}}{(2\pi)^n}}\ e^{-\boldsymbol{x}^T \cdot \mathbb{g} \cdot \boldsymbol{x}/2} \tag{1}$$

where $\mathbb{g}$ is a $n \times n$ real, symmetric, positive definite matrix so that

$$\boldsymbol{x}^T \cdot \mathbb{g} \cdot \boldsymbol{x} = \sum_{i,j=1}^{n} x_i g_{ij} x_j > 0 \qquad \forall \boldsymbol{x} \tag{2}$$

(a) Show that $P_{\boldsymbol{X}}(\boldsymbol{x})$ is normalized to 1.

(b) Compute the characteristic function $f_{\boldsymbol{X}}(\boldsymbol{k})$.

(c) Show that $\langle x_i \rangle = 0$ as claimed, while all linear joint moments of the form $\langle x_1 \ldots x_{2n} \rangle$ can be expressed as

a linear combination of products of $\langle x_i x_j \rangle$.

Answer (a)

Since $\mathbb{g}$ is real symmetric, it can be diagonalized by a orthogonal transform so that

$$\mathbb{O}^T \cdot \mathbb{g} \cdot \mathbb{O} = \mathbb{d} = \text{diag}(d_1, \ldots, d_n) \tag{3}$$

where $\{d_j > 0\}$ are the eigenvalues of $\mathbb{g}$ and $\mathbb{O}$ is a $n \times n$ orthogonal matrix so that

$$\mathbb{O}^{-1} = \mathbb{O}^T \qquad \& \qquad \det \mathbb{O} = \det \mathbb{O}^T = 1 \tag{3a}$$

Let

$$\boldsymbol{\alpha} = \mathbb{O}^T \cdot \boldsymbol{x} \qquad \rightarrow \qquad \boldsymbol{x} = \mathbb{O} \cdot \boldsymbol{\alpha} \tag{3b}$$

then

$$J = \frac{\partial(\alpha_1, \ldots, \alpha_n)}{\partial(x_1, \ldots, x_n)} = \det\left(\frac{\partial \alpha_i}{\partial x_j}\right) = \det \mathbb{O}^T = 1$$

$$\rightarrow \quad d^n \alpha = d\alpha_1 \ldots d\alpha_n = |J| \, d^n x = d^n x = dx_1 \ldots dx_n \tag{4}$$

Also,

$$\begin{aligned} \boldsymbol{x}^T \cdot \mathbb{g} \cdot \boldsymbol{x} &= (\mathbb{O} \cdot \boldsymbol{\alpha})^T \cdot \mathbb{g} \cdot (\mathbb{O} \cdot \boldsymbol{\alpha}) \\ &= \boldsymbol{\alpha}^T \cdot \mathbb{O}^T \cdot \mathbb{g} \cdot \mathbb{O} \cdot \boldsymbol{\alpha} \\ &= \boldsymbol{\alpha}^T \cdot \mathbb{d} \cdot \boldsymbol{\alpha} \qquad (4a) \\ &= \sum_{i,j=1}^{n} \alpha_i d_i \delta_{ij} \alpha_j \\ &= \sum_{i=1}^{n} \alpha_i^2 d_i \qquad (5) \end{aligned}$$

and

$$\det \mathbb{g} = \det(\mathbb{O} \cdot \mathbb{d} \cdot \mathbb{O}^T) = \det(\mathbb{O}^T \cdot \mathbb{O} \cdot \mathbb{d}) = \det(\mathbb{d}) = d_1 \cdots d_n \tag{6}$$

Hence,

$$\begin{aligned} \int d^n x \, e^{-\boldsymbol{x}^T \cdot \mathbb{g} \cdot \boldsymbol{x}/2} &= \int_{-\infty}^{\infty} dx_1 \ldots \int_{-\infty}^{\infty} dx_n \, e^{-\boldsymbol{x}^T \cdot \mathbb{g} \cdot \boldsymbol{x}/2} \\ &= \prod_{i=1}^{n} \int_{-\infty}^{\infty} d\alpha_i \, e^{-\alpha_i^2 d_i/2} \qquad [\text{ (5) used. }] \\ &= \prod_{i=1}^{n} \sqrt{\frac{2\pi}{d_i}} \qquad [\textit{ Mathematica} \text{ used. }] \end{aligned}$$

$$= \sqrt{\frac{(2\pi)^n}{\det \mathbb{g}}} \qquad \text{[(6) used.]} \qquad (7)$$

$$\rightarrow \quad \int d^n x\, P_{\boldsymbol{X}}(\boldsymbol{x}) = 1 \qquad \text{[(1) used.]}$$

In[ ]:= `Assuming[d > 0, ∫_{-∞}^{∞} e^{- a^2 d / 2} ⅆa]`

Out[ ]= $\frac{\sqrt{2\pi}}{\sqrt{d}}$

Answer (b)

Using (4.36), we have

$$f_{\boldsymbol{X}}(\boldsymbol{k}) = \sqrt{\frac{\det \mathbb{g}}{(2\pi)^n}} \int d^n x\, e^{i \boldsymbol{k}^T \cdot \boldsymbol{x}}\, e^{-\boldsymbol{x}^T \cdot \mathbb{g} \cdot \boldsymbol{x}/2}$$

$$= \sqrt{\frac{\det \mathbb{g}}{(2\pi)^n}} \int d^n \alpha\, e^{i \boldsymbol{k}^T \cdot \mathbb{O} \cdot \boldsymbol{\alpha}}\, e^{-\boldsymbol{\alpha}^T \cdot \mathbb{d} \cdot \boldsymbol{\alpha}/2} \qquad \text{[(3b) \& (4a) used.]}$$

$$= \sqrt{\frac{\det \mathbb{g}}{(2\pi)^n}} \prod_{i=1}^{n} \int_{-\infty}^{\infty} d\,\alpha_i\, e^{i \tilde{k}_i \alpha_i}\, e^{-\alpha_i^2 d_i/2} \qquad [\, \tilde{\boldsymbol{k}} = \mathbb{O}^T \cdot \boldsymbol{k}.\ \ \text{(5) used.]} \qquad (8)$$

$$= \sqrt{\frac{\det \mathbb{g}}{(2\pi)^n}} \prod_{i=1}^{n} \sqrt{\frac{2\pi}{d_i}}\, e^{-\tilde{k}_i^2 / 2 d_i} \qquad \text{[\textit{Mathematica} used.]}$$

$$= \exp\left(-\frac{1}{2} \sum_{i=1}^{n} \frac{\tilde{k}_i^2}{d_i} \right) \qquad \text{[(6) used.]} \qquad (8a)$$

$$= \exp\left(-\frac{1}{2} \tilde{\boldsymbol{k}}^T \cdot \mathbb{d}^{-1} \cdot \tilde{\boldsymbol{k}} \right)$$

$$= \exp\left(-\frac{1}{2} \boldsymbol{k}^T \cdot \mathbb{O} \cdot \mathbb{d}^{-1} \cdot \mathbb{O}^T \cdot \boldsymbol{k} \right) \qquad \text{[(8) used.]}$$

$$= \exp\left(-\frac{1}{2} \boldsymbol{k}^T \cdot \mathbb{g}^{-1} \cdot \boldsymbol{k} \right) \qquad \text{[(1) used.]} \qquad (8b)$$

In[ ]:= `Assuming[d > 0 && k > 0, ∫_{-∞}^{∞} e^{i k a - a^2 d / 2} ⅆa]`

Out[ ]= $\frac{e^{-\frac{k^2}{2 d}} \sqrt{2\pi}}{\sqrt{d}}$

Answer (c)

(8b) can be written as

$$f_{\boldsymbol{X}}(\boldsymbol{k}) = \exp\left[-\frac{1}{2} \sum_{i,j=1}^{n} k_i \left(\mathbb{g}^{-1}\right)_{ij} k_j \right] = e^{-\Phi} \qquad (9)$$

where

$$\Phi = \frac{1}{2}\sum_{i,j=1}^{n} k_i \left(\mathbb{g}^{-1}\right)_{ij} k_j \tag{9a}$$

Hence,

$$\frac{\partial \Phi}{\partial k_m} = \frac{1}{2}\sum_{i,j=1}^{n}\left[\ \delta_{im}\left(\mathbb{g}^{-1}\right)_{ij} k_j + k_i\left(\mathbb{g}^{-1}\right)_{ij}\delta_{jm}\right]$$

$$= \frac{1}{2}\left[\sum_{j=1}^{n}\left(\mathbb{g}^{-1}\right)_{mj} k_j + \sum_{i=1}^{n} k_i\left(\mathbb{g}^{-1}\right)_{im}\ \right]$$

$$= \sum_{j=1}^{n}\left(\mathbb{g}^{-1}\right)_{mj} k_j \qquad [\ \mathbb{g}^T = \mathbb{g}\ \text{ used.}\ ]$$

$$= \left(\mathbb{g}^{-1}\cdot\boldsymbol{k}\right)_m \tag{10}$$

$$\rightarrow \quad \frac{\partial^2 \Phi}{\partial k_i \partial k_m} = \sum_{j=1}^{n}\left(\mathbb{g}^{-1}\right)_{mj}\delta_{ij} = \left(\mathbb{g}^{-1}\right)_{mi} = \left(\mathbb{g}^{-1}\right)_{im} \tag{10a}$$

$$\frac{\partial^n \Phi}{\partial k_{i_1} \ldots \partial k_{i_n}} = 0 \qquad \forall\ n \geq 3 \tag{10b}$$

Therefore,

$$\frac{\partial f_{\boldsymbol{X}}(\boldsymbol{k})}{\partial k_m} = -\frac{\partial \Phi}{\partial k_m} e^{-\Phi} = -\left(\mathbb{g}^{-1}\cdot\boldsymbol{k}\right)_m f_{\boldsymbol{X}}(\boldsymbol{k}) \tag{11}$$

$$\frac{\partial^2 f_{\boldsymbol{X}}(\boldsymbol{k})}{\partial k_i \partial k_m} = -\left(\frac{\partial^2 \Phi}{\partial k_i \partial k_m} - \frac{\partial \Phi}{\partial k_m}\frac{\partial \Phi}{\partial k_i}\right)e^{-\Phi} = -\left[\ \left(\mathbb{g}^{-1}\right)_{im} - \left(\mathbb{g}^{-1}\cdot\boldsymbol{k}\right)_i\left(\mathbb{g}^{-1}\cdot\boldsymbol{k}\right)_m\right] f_{\boldsymbol{X}}(\boldsymbol{k}) \tag{11a}$$

$$\frac{\partial^2 f_{\boldsymbol{X}}(\boldsymbol{k})}{\partial k_i^2} = -\left[\frac{\partial^2 \Phi}{\partial k_i^2} - \left(\frac{\partial \Phi}{\partial k_i}\right)^2\right]e^{-\Phi} = -\left[\ \left(\mathbb{g}^{-1}\right)_{ii} - \left(\mathbb{g}^{-1}\cdot\boldsymbol{k}\right)_i^2\ \right] f_{\boldsymbol{X}}(\boldsymbol{k}) \tag{11b}$$

$$\frac{\partial^3 f_{\boldsymbol{X}}(\boldsymbol{k})}{\partial k_i^3} = -\left\{\frac{\partial^3 \Phi}{\partial k_i^3} - 2\left(\frac{\partial \Phi}{\partial k_i}\right)\frac{\partial^2 \Phi}{\partial k_i^2} - \left[\frac{\partial^2 \Phi}{\partial k_i^2} - \left(\frac{\partial \Phi}{\partial k_i}\right)^2\right]\frac{\partial \Phi}{\partial k_i}\right\}e^{-\Phi}$$

$$= -\left(\frac{\partial \Phi}{\partial k_i}\right)\left[-3\frac{\partial^2 \Phi}{\partial k_i^2} + \left(\frac{\partial \Phi}{\partial k_i}\right)^2\ \right]e^{-\Phi} \qquad [\ \text{(10b) used.}\ ]$$

$$= -\left(\mathbb{g}^{-1}\cdot\boldsymbol{k}\right)_i\left[\ -3\left(\mathbb{g}^{-1}\right)_{ii} - \left(\mathbb{g}^{-1}\cdot\boldsymbol{k}\right)_i^2\ \right] f_{\boldsymbol{X}}(\boldsymbol{k}) \tag{11c}$$

from which we deduce

$$\frac{\partial^p f_{\boldsymbol{X}}(\boldsymbol{k})}{\partial k_i^p} \propto \left(\mathbb{g}^{-1}\cdot\boldsymbol{k}\right)_i \qquad \text{if } p \text{ is odd} \tag{11d}$$

Putting (11,a-c) into [see (4.37)]

$$\left\langle x_1^{m_1} \ldots x_p^{m_p}\right\rangle = (-i)^{m_1+\ldots+m_p}\frac{\partial^{m_1}}{\partial k_1^{m_1}}\ldots\frac{\partial^{m_p}}{\partial k_p^{m_p}} f_{\boldsymbol{X}}(\boldsymbol{k})\bigg|_{\boldsymbol{k}=0}$$

we can draw the following conclusions.

a. $\left\langle x_1^{m_1} \ldots x_p^{m_p}\right\rangle = 0$ if any m_j is odd. [(11d) used.]

b. $\left\langle x_1^{m_1} \ldots x_p^{m_p}\right\rangle = 0$ if $\sum_{j=1}^{p} m_p$ is odd. [Since one of the m_j's must be odd.]

c. $\langle x_i x_j\rangle = \left(\mathbb{g}^{-1}\right)_{ij}$

d. $\left\langle x_j^m\right\rangle \neq 0$ $\forall\ m$ even

Linear joint moments are of the form $\langle x_1 \ldots x_p\rangle$, i.e., $m_j = 1\ \ \forall\ j$.

According to (b),

$$\langle x_1 \ldots x_p \rangle = 0 \qquad \text{if} \quad p \text{ is odd.}$$

Therefore, we need consider only $\langle x_1 \ldots x_{2n} \rangle$.

To begin, consider the case $n = 2$.

$$\langle x_1 x_2 x_3 x_4 \rangle = (-i)^4 \frac{\partial^4}{\partial k_1 \partial k_2 \partial k_3 \partial k_4} f_{\boldsymbol{x}}(\boldsymbol{k}) \bigg|_{\boldsymbol{k}=0} \tag{12}$$

Using

$$\frac{\partial}{\partial k_i} (\mathbb{g}^{-1} \cdot \boldsymbol{k})_j = \frac{\partial}{\partial k_i} [(\mathbb{g}^{-1})_{jm} k_m] = (\mathbb{g}^{-1})_{jm} \delta_{im} = (\mathbb{g}^{-1})_{ji}$$

and (11a), we have

$$\begin{aligned}
\frac{\partial^4 f_{\boldsymbol{x}}(\boldsymbol{k})}{\partial k_1 \partial k_2 \partial k_3 \partial k_4} &= -\frac{\partial^2}{\partial k_1 \partial k_2} \{ [(\mathbb{g}^{-1})_{34} - (\mathbb{g}^{-1} \cdot \boldsymbol{k})_3 (\mathbb{g}^{-1} \cdot \boldsymbol{k})_4] f_{\boldsymbol{x}}(\boldsymbol{k}) \} \\
&= \frac{\partial}{\partial k_1} \{ [(\mathbb{g}^{-1})_{32} (\mathbb{g}^{-1} \cdot \boldsymbol{k})_4 + (\mathbb{g}^{-1} \cdot \boldsymbol{k})_3 (\mathbb{g}^{-1})_{42} \\
&\qquad + ((\mathbb{g}^{-1})_{34} - (\mathbb{g}^{-1} \cdot \boldsymbol{k})_3 (\mathbb{g}^{-1} \cdot \boldsymbol{k})_4) (\mathbb{g}^{-1} \cdot \boldsymbol{k})_2] f_{\boldsymbol{x}}(\boldsymbol{k}) \} \\
&= \{ (\mathbb{g}^{-1})_{32} (\mathbb{g}^{-1})_{41} + (\mathbb{g}^{-1})_{31} (\mathbb{g}^{-1})_{42} - [(\mathbb{g}^{-1})_{31} (\mathbb{g}^{-1} \cdot \boldsymbol{k})_4 + (\mathbb{g}^{-1} \cdot \boldsymbol{k})_3 (\mathbb{g}^{-1})_{41}] (\mathbb{g}^{-1} \cdot \boldsymbol{k})_2 \\
&\qquad + [(\mathbb{g}^{-1})_{34} - (\mathbb{g}^{-1} \cdot \boldsymbol{k})_3 (\mathbb{g}^{-1} \cdot \boldsymbol{k})_4] (\mathbb{g}^{-1})_{21} \\
&\qquad + [(\mathbb{g}^{-1})_{32} (\mathbb{g}^{-1} \cdot \boldsymbol{k})_4 + (\mathbb{g}^{-1} \cdot \boldsymbol{k})_3 (\mathbb{g}^{-1})_{42} \\
&\qquad\quad + ((\mathbb{g}^{-1})_{34} - (\mathbb{g}^{-1} \cdot \boldsymbol{k})_3 (\mathbb{g}^{-1} \cdot \boldsymbol{k})_4) (\mathbb{g}^{-1} \cdot \boldsymbol{k})_2] (\mathbb{g}^{-1} \cdot \boldsymbol{k})_1 \} f_{\boldsymbol{x}}(\boldsymbol{k})
\end{aligned}$$

(12) thus becomes

$$\begin{aligned}
\langle x_1 x_2 x_3 x_4 \rangle &= (\mathbb{g}^{-1})_{32} (\mathbb{g}^{-1})_{41} + (\mathbb{g}^{-1})_{31} (\mathbb{g}^{-1})_{42} + (\mathbb{g}^{-1})_{34} (\mathbb{g}^{-1})_{21} \\
&= \langle x_3 x_2 \rangle \langle x_4 x_1 \rangle + \langle x_3 x_1 \rangle \langle x_4 x_2 \rangle + \langle x_3 x_4 \rangle \langle x_2 x_1 \rangle \qquad [\langle x_i x_j \rangle = \langle x_j x_i \rangle] \\
&= \text{sum of all possible averaged pairings of } x_1 x_2 x_3 x_4 .
\end{aligned} \tag{13}$$

Now, since [see (11a)]

$$\frac{\partial^2 f_{\boldsymbol{x}}(\boldsymbol{k})}{\partial k_i \partial k_j} \bigg|_{\boldsymbol{k}=0} = -(\mathbb{g}^{-1})_{ij}$$

the only nonvanishing items in a general linear product $\langle x_1 \ldots x_{2n} \rangle$ are the $(\mathbb{g}^{-1})_{ij}$'s. The only way to generalize (13) is therefore

$$\langle x_1 \ldots x_{2n} \rangle = \text{sum of all possible averaged pairings of } x_1 \ldots x_{2n} . \tag{14}$$

Note that although (14) resembles Wick's theorem, it is not appropriate to compare the two because, while the x_j's are all realizations, Wick's theorem is the result of ordering two kinds of objects, annihilators & creators, for which the average of a "normal" ordered product vanishes.

The number of terms in (14) is just the number of ways to pick, n pairs from $2n$ objects, without regard of the orders of the picking as well as that within each pair. For the 1st pair, there are $\frac{(2n)(2n-1)}{2}$ ways, where the factor $\frac{1}{2}$ comes from disregarding the order within the pair. For the 2nd pair, $\frac{(2n-2)(2n-3)}{2}$ ways. For the jth pair, $\frac{[2n-2(j-1)][2n-2(j-1)-1]}{2}$ ways, and so on. The number of ways for the *n*th pair is

$$\frac{[2n-2(n-1)][2n-2(n-1)-1]}{2}=1$$

The total number of ways is therefore

$$\frac{(2n)(2n-1)\cdots 2\cdot 1}{2^n\, n!}=\frac{(2n)!}{n!\,2^n}$$

where the factor $\frac{1}{n!}$ comes from disregarding the order of picking the pairs.