Theorem Let $p>1$, $q>1$, so that $1/p + 1/q = 1$. Then

$$E(|XY|) \le [E(|X|^p)]^{1/p} [E(|Y|^q)]^{1/q}.$$

Proof. ~~Fo~~ In the Holder's inequality,

put $x = \left|X [E(|X|^p)]^{-1/p}\right|$, $y = \left|Y [E(|Y|^q)]^{-1/q}\right|$

then we get

$$\left|X [E(|X|^p)]^{-1/p} \, Y [E(|Y|^q)]^{-1/q}\right|$$

$$= |XY| \, (E(|X|^p))^{-1/p} \, [E(|Y|^q)]^{-1/q}$$

$$\le \frac{1}{p} \left|X^p [E(|X|^p)]^{-1}\right| + \frac{1}{q} \left|Y^q [E(|Y|^q)]^{-1}\right|$$

with probability 1. ~~take expectation both sides~~ to get

$$\cancel{E[|XY|] \, [E(|X|^p)]^{-1/p} [E(|Y|^q)]^{-1/q}}$$

$$\cancel{\le \frac{1}{p} \, E(|X|^p)}$$

That is,

$$|XY| \le \frac{1}{p} \, \frac{|X|^p}{E(|X|^p)} [E(|X|^p)]^{1/p} [E(|Y|^q)]^{1/q}$$

$$+ \frac{1}{q} \, \frac{|Y|^q}{E(|Y|^q)} [E(|X|^p)]^{1/p} [E(|Y|^q)]^{1/q}$$

with probability 1. Now, we can take expectation on both the sides to get the desired inequality.

Functions of several random variables.

Let $X_1, X_2, \ldots, X_n$ be random variables defined on a probability space $(\Omega, \mathcal{F}, P)$. In practice, we deal with functions of $X_1, X_2, \ldots, X_n$ such as i) $X_1 + X_2$ ii) $X_1 - X_2$ iii) $X_1 X_2$ iv) $\min(X_1, \ldots X_n)$ v) $\max(X_1, X_2, \ldots, X_n)$ vi) $X_1^{\alpha} X_2^{\beta}$ and so on.

Qn Given the joint distribution of $(X_1, X_2, \ldots X_n)$ can we derive distribution functions of their functions?

Theorem

Let $g: \mathbb{R}^n \to \mathbb{R}^m$ be a Borel measurable function, that is, if $B \in \mathcal{B}(\mathbb{R}^m)$, then $g^{-1}(B) \in \mathcal{B}(\mathbb{R}^n)$. If $\underline{X} = (X_1, \ldots, X_n)$ is an n-dimensional random vector, then $g(\underline{X})$ is an m-dimensional random vector.

Idea: Measurable function of a random vector is a random vector.

Remark If $g: \mathbb{R}^n \to \mathbb{R}^m$ is a continuous function, then $g(X_1, X_2, \ldots, X_n)$ is a random vector when $(X_1, \dots X_n)$ is a random vector.

Now, we would like our question to be answered. ~~Unfortunate~~ Unfortunately, there is no simple and single answer to the question. There are several methods and tricks to arrive at our solution. You ~~can~~ are free to choose your favourite one.

The master equations for the functions of random vectors.

i) $\mathbb{P}\left(g(X_1, \dots X_n) \leq (y_1, \dots y_m) \right)$

$$= \mathbb{P}\left(g(X_1, \dots X_n) \in (y_1, \infty) \times (y_2, \infty) \times \dots \times (y_m, \infty) \right)$$

[$g : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a measurable function and $(Y_1, Y_2, \dots, Y_m) := g(X_1, X_2, \dots, X_n)$.]

$$= \mathbb{P}\left((Y_1, \dots Y_m) \in (y_1, \infty) \times (y_2, \infty) \times \dots \times (y_m, \infty) \right)$$

$$= \mathbb{P}\left(\bigcap_{i=1}^{m} \{ X_i \leq y_i \} \right) = \mathbb{P}\left(\bigcap_{i=1}^{m} \{ Y_i \in (y_i, \infty) \} \right).$$

$$= \begin{cases} \displaystyle\sum_{\{(x_1, \dots x_n) : g(x_1, x_2, \dots, x_n) \in (y_1, \infty) \times (y_2, \infty) \times \dots \times (y_m, \infty)\}} \mathbb{P}(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) \\ \displaystyle\int\int \dots \int_{\{(x_1, x_2, \dots, x_n) : g(x_1, \dots x_n) \in (y_1, \infty) \times (y_2, \infty) \times \dots \times (y_m, \infty)\}} f(x_1, x_2, \dots x_n) \, dx_1 \dots dx_n \end{cases}$$

<u>Often</u>

$$\int\int \dots \int f(x_1, x_2, \dots x_n) \, dx_1 \dots dx_n = \int dx_1 \int dx_2 \dots \int dx_n \, f(x_1, \dots, x_n)$$

where we use the summation in case of discrete random vector and integration in case of continuous random vectors.

Of course, in case of contionous random vector $\underline{Y}=(Y_1, Y_2, \ldots Y_m) = g(X_1, X_2, \ldots, X_n)$, we can differentiate the ~~random vect~~ distribution function

$$\mathbb{F}_{\underline{Y}}(y_1, y_2, \ldots, y_m) = \mathbb{P}\left(\bigcap_{i=1}^{m} \{Y_i \le y_i\}\right)$$

with respect to $y_1, y_2, \ldots y_m$ to derive the joint density of the random vector $\underline{Y} = (Y_1, Y_2, \ldots Y_m)$.

$$f_{Y_1 \ldots Y_m}(y_1, y_2, \ldots y_m) = \frac{\partial^m}{\partial y_1 \partial y_2 \ldots \partial y_m} \mathbb{F}_{Y_1 \ldots Y_m}(y_1, y_2, \ldots y_m)$$

Let us consider a simpler situation with $m=n$. So $g: \mathbb{R}^n \to \mathbb{R}^n$ and let $(Y_1, \ldots Y_n) = g(X_1, X_2, \ldots, X_n)$ and $(y_1, y_2, \ldots y_n) = g(x_1, x_2, \ldots, x_n)$. Here g is a multi-valued function and so we use $\underline{g}$ instead of g and $\underline{g} = (g_1, g_2, \ldots g_n)$. So we can write

$$y_i = g_i(x_1, x_2, \ldots x_n)$$

where $g_i : \mathbb{R}^n \to \mathbb{R}$ for $i=1, 2, 3, \ldots n$. We shall use $\underline{x} = (x_1, x_2, \ldots, x_n)$. Let $f(x_1, x_2, \ldots x_n)$ be the ~~density of~~ joint density of $\underline{X}$.

Then for any $B \in \mathcal{B}(\mathbb{R}^n)$, we have

$$\mathbb{P}\left((Y_1, Y_2, \ldots, Y_n) \in B \right) = \mathbb{P}\left[g(X_1, X_2, \ldots X_n) \in B \right].$$

$$= \mathbb{P}\left[(X_1, X_2, \ldots, X_n) \in g^{-1}(B) \right].$$

$$= \int_{g^{-1}(B)} \left(\prod_{i=1}^{n} dx_i \right) f(x_1, \ldots, x_n)$$

$$= \int_{\{(x_1 \ldots x_n) \in \mathbb{R}^n : g(x_1 \ldots x_n) \in B\}} \left(\prod_{i=1}^{n} dx_i \right) f(x_1, \ldots, x_n)$$

[Recall that $\mathcal{B}(\mathbb{R}^n)$ is actually generated by the the semi-~~open~~ closed intervals

$$\left\{ (-\infty, u_1] \times (-\infty, u_2] \times \cdots \times (-\infty, u_n] : (u_1, u_2, \ldots, u_n) \in \mathbb{R}^n \right\}$$

and so ~~we ca~~ it is enough to focus on

$$B = (-\infty, y_1] \times (-\infty, y_2] \times \cdots \times (-\infty, y_n].$$

Note that

$$\left\{ (x_1, x_2, \ldots x_n) \in \mathbb{R}^n : g(x_1, \ldots x_n) \in B \right\}$$

$$= \left\{ (x_1, x_2, \ldots x_n) \in \mathbb{R}^n : g_1(x_1 \ldots x_n) \le y_1, \; g_2(x_1 \ldots x_n) \le y_2, \ldots, g_n(x_1 \ldots x_n) \le y_n \right\}$$

Suppose that g is a "nice" function (non-singular linear transform) and hence, invertible. This means that there exists a function $g^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ such that if $g(x_1, x_2, \ldots x_n) = (y_1, y_2, \ldots y_n)$ then $g^{-1}(y_1, y_2, \ldots, y_n) = (x_1, x_2, \ldots x_n)$. Then we can use the "Jacobian of tranformation" to write down the ~~der~~ partial derivatives etc. Furthermore for the notational simplicity we can use $g^{-1} = (g_1^{-1}, g_2^{-1}, \ldots, g_n^{-1})$. Then we have following simplified form.

$$= \int_{g_1^{-1}(-\infty, y_1)} dx_1 \int_{g_2^{-1}(-\infty, y_2)} dx_2 \ldots \int_{g_n^{-1}(-\infty, y_n)} dx_n \; f(x_1 \ldots x_n)$$

Once we reach here, we can use the change of variable formula for the integral of several variables and see

$$\int_{-\infty}^{y_1} du_1 \int_{-\infty}^{y_2} du_2 \ldots \int_{-\infty}^{y_n} du_n \; f\left(g_1^{-1}(u_1, u_2 \ldots u_n), g_2^{-1}(u_1, u_2, \ldots u_n), \ldots g_n^{-1}(u_1 \ldots u_n)\right) \det(J)$$

where

$$J = \frac{\partial(x_1 \ldots x_n)}{\partial(u_1 \ldots u_n)} = \begin{vmatrix} \frac{\partial x_1}{\partial u_1} & \frac{\partial x_1}{\partial u_2} & \ldots & \frac{\partial x_1}{\partial u_n} \\ \frac{\partial x_2}{\partial u_1} & \frac{\partial x_2}{\partial u_2} & \ldots & \frac{\partial x_2}{\partial u_n} \\ \frac{\partial x_n}{\partial u_n} & \frac{\partial x_n}{\partial u_n} & \ldots & \frac{\partial x_n}{\partial u_n} \end{vmatrix}$$

which can be computed using the explicit form of g^{-1}.

Theorem Let $(X_1, X_2, \ldots, X_n)$ be an n-dimensional random vector of continuous type with probability density function $f(x_1, x_2, \ldots, x_n)$.

a) Let

$$y_1 = g_1(x_1, x_2, \ldots, x_n)$$
$$y_2 = g_2(x_1, x_2, \ldots, x_n)$$
$$\vdots$$
$$y_n = g_n(x_1, x_2, \ldots, x_n)$$

be a one-to-one mapping of $\mathbb{R}^n$ into itself; that is, there exists the inverse transformation

$$x_1 = h_1(y_1, y_2, \ldots, y_n), \quad x_2 = h_2(y_1, y_2, \ldots, y_n)$$
$$\ldots, \quad x_n = h_n(y_1, y_2, \ldots, y_n)$$

defined over the range of transformation.

b) Assume that both the mapping $\underline{g} = (g_1, g_2, \ldots, g_n)$ and $\underline{h} = (h_1, h_2, \ldots, h_n)$ are continuous.

c) Assume that the partial derivatives

$$\frac{\partial x_j}{\partial y_i} = \frac{\partial h_j(y_1, \ldots, y_n)}{\partial y_i}, \quad 1 \le i \le n, \quad 1 \le j \le n$$

exist and continuous.

d) Assume that the Jacobian J of inverse transformation

$$J = \frac{\partial(x_1, x_2, \ldots, x_n)}{\partial(y_1, y_2, \ldots, y_n)} = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} & \cdots & \frac{\partial x_2}{\partial y_n} \\ \vdots & & & \\ \frac{\partial x_n}{\partial y_1} & \frac{\partial x_n}{\partial y_2} & \cdots & \frac{\partial x_n}{\partial y_n} \end{vmatrix}$$

$$= \begin{vmatrix} \frac{\partial h_1}{\partial y_1} & \frac{\partial h_1}{\partial y_2} & \cdots & \frac{\partial h_1}{\partial y_n} \\ \frac{\partial h_2}{\partial y_1} & \frac{\partial h_2}{\partial y_2} & \cdots & \frac{\partial h_2}{\partial y_n} \\ \vdots & & & \\ \frac{\partial h_n}{\partial y_1} & \frac{\partial h_n}{\partial y_2} & \cdots & \frac{\partial h_n}{\partial y_n} \end{vmatrix}$$

is different from zero for $(y_1, y_2, \ldots, y_n)$ in the range of transformation.

Then $\underline{Y} = (Y_1, Y_2, \ldots, Y_n)$ has a joint absolutely continuous distribution function with probability density function given by

$$w(y_1, y_2, \ldots, y_n) = \det(J)\, f(h_1(y_1, \ldots, y_n), h_2(y_1, \ldots, y_n), \ldots, h_n(y_1, \ldots, y_n)).$$

Proof. For $(y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$, let

$$B = (-\infty, y_1] \times (-\infty, y_2] \times \cdots \times (-\infty, y_n].$$

Then $g^{-1}(B) = \{ \underline{x} = (x_1, \ldots x_n) : g_i(\underline{x}) \le y_i \text{ for } i = 1, 2, \ldots n \}$

$$\mathbb{P}(\underline{Y} \in B) = \mathbb{P}(\underline{X} \in g^{-1}(B))$$

$$= \int_{g^{-1}(B)} dx_1 \int dx_2 \cdots \int dx_n \; f(x_1, x_2, \ldots, x_n)$$

$$= \int_{-\infty}^{y_1} du_1 \int_{-\infty}^{y_2} du_2 \cdots \int_{-\infty}^{y_n} du_n \; f(h_1(u_1, \ldots u_n), h_2(u_1, u_2, \ldots, u_n), \ldots, h_n(u_1, u_2, \ldots u_n))$$

$$\left| \frac{\partial(x_1, x_2, \ldots, x_n)}{\partial(y_1, y_2, \ldots, y_n)} \right|_{(y_1, y_2, \ldots, y_n) = (u_1, u_2, \ldots u_n)}$$

(using change of variable formula for multiple integrals).

Remark. In actual applications, we may not have a mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$. Instead, we have mapping $\mathbb{R}^n \to \mathbb{R}^m$ where $m < n$. In that case, we add our favourite $(n-m)$ random variables ~~in the range~~ or function of random variables to get an one-to-one mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$. Then we integrate the extra variables to get the desired joint density.

$\underset{\sim}{g} : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ that is,

$$\underset{\sim}{g}(x_1, x_2, \ldots x_n) = \left(g_1(x_1, x_2, \ldots x_n), g_2(x_1, x_2, \ldots, x_n), \ldots g_m(x_1, x_2, \ldots, x_n)\right)$$

* Consider a few new functions

$$\overset{N}{g}_{m+1}(x_1 \ldots x_n), \overset{N}{g}_{m+2}(x_1, \ldots x_n), \ldots \overset{N}{g}_n(x_1 \ldots x_n).$$

* Define a new function $\overset{N}{\underset{\sim}{g}} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ such that

$$\overset{N}{\underset{\sim}{g}} = \left(g_1, g_2, \ldots, g_m, \overset{N}{g}_{m+1}, \overset{N}{g}_{m+2}, \ldots, \overset{N}{g}_n\right)$$

that is,

$$\overset{N}{\underset{\sim}{g}}(x_1 \ldots x_n) = \left(g_1(x_1, x_2, \ldots x_n), g_2(x_1, x_2, \ldots x_n), \ldots, g_m(x_1, x_2, \ldots x_n), \overset{N}{g}_{m+1}(x_1, x_2, \ldots, x_n), \ldots, \overset{N}{g}_n(x_1, x_2, \ldots x_n)\right)$$

* We make sure $\overset{N}{\underset{\sim}{g}} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is one-to-one and hence $\overset{N}{\underset{\sim}{g}}^{-1}$ exists. We use the theorem to derive the joint density function of

$$\left(g_1(X_1, \ldots X_n), g_2(X_1, X_2, \ldots, X_n), \ldots, g_m(X_1, X_2, \ldots, X_n), \overset{N}{g}_{m+1}(X_1, X_2, \ldots, X_n), \overset{N}{g}_{m+2}(X_1, X_2, \ldots, X_n), \ldots, \overset{N}{g}_n(X_1, X_2, \ldots, X_n)\right)$$

$$=: (Y_1, Y_2, \ldots, Y_m, Y_{m+1}, Y_{m+2}, \ldots, Y_n)$$

* From the density function of $(Y_1, Y_2, \ldots Y_m, Y_{m+1}, Y_{m+2}, \ldots Y_n)$ we integrate $(Y_{m+1}, Y_{m+2}, \ldots, Y_n)$.