Multivariate Distribution: Joint Distribution

Part I: Distributional Level Concepts

Le Wang

Questions

We are often interseted in the relationships between two random variables, for example,

1. Whether changes in interest rates have any impacts on stock prices
2. Wages and Years of Schooling
3. Income and the Number of Children
4. The level of air pollution and child mortality
5. etc.

Again, these variables are random variables because before their values are realized, we do not know what will happen.

Multivariate Distribution:

Extension to Random vectors (multiple random variables)

1. univariate (one-dimensional random variable)
2. bivariate (two-dimensional random variable)
3. multivariate (random variables of arbitrary dimension)

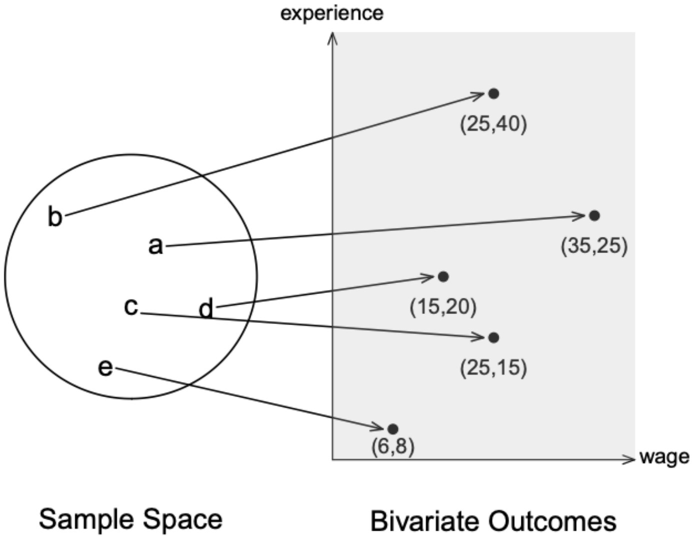
Bivariate Random Variables

Bivariate Randon Variable is a pair of numerical outcomes; a function from the sample space to $\mathbb{R}^2$.

Suppose that X and Y are two discrete random variables and that X can take on values $\{x_{1,}x_{2,}, \ldots, x_n\} and Y \{y_1, y_2, \ldots, y_m\}$. Then, we have $n \times m$ combinations of possible events.

The ordered pair (X, Y) take values in the product $\{(x_1, y_1), (x_1, y_2\}, \ldots, (x_n, y_1), \ldots, (x_n, y_m)\}$

Let's visualize it (wage and experience)



Definition (Joint) Probability Mass Function (p.m.f) of X and Y is defined as follows

$$\Pr[X = x_i, Y = y_j]$$

Intuition: You can think of a ordered pair (combination) as an unit, the joint distribution is nothing but the probability of a combination.

If you know the joint distribution of (X, Y), you know everything about X and Y, not only the individual one, but also their relationship.

$X \backslash Y$	y_1	y_2	$\dots$	y_j	$\dots$	y_m
x_1	$p(x_1, y_1)$	$p(x_1, y_2)$	$\dots$	$p(x_1, y_j)$	$\dots$	$p(x_1, y_m)$
x_2	$p(x_2, y_1)$	$p(x_2, y_2)$	$\dots$	$p(x_2, y_j)$	$\dots$	$p(x_2, y_m)$
$\vdots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
x_i	$p(x_i, y_1)$	$p(x_i, y_2)$	$\dots$	$p(x_i, y_j)$	$\dots$	$p(x_i, y_m)$
$\vdots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
x_n	$p(x_n, y_1)$	$p(x_n, y_2)$	$\dots$	$p(x_n, y_j)$	$\dots$	$p(x_n, y_m)$

A joint probability mass function must satisfy the following two properties:

1.$0 \leq p\left(x_i, y_j\right) \leq 1$ 2.The total probability is 1. In other words,

$$\sum_{i=1}^{n} \sum_{j=1}^{m} p\left(x_i, y_j\right) = 1$$

$$
\begin{aligned}
\sum_{i=1}^{n} \sum_{j=1}^{m} p\left(x_i, y_j\right) &= \sum_{i=1}^{n} \left[\sum_{j=1}^{m} p\left(x_i, y_j\right)\right] \\
&= \sum_{i=1}^{n} \left[p\left(x_i, y_1\right) + p\left(x_i, y_2\right) + p\left(x_i, y_3\right) + \cdots + p\left(x_i, y_m\right)\right] \\
(\text{ first row }) &= \left[p\left(x_1, y_1\right) + p\left(x_1, y_2\right) + p\left(x_1, y_3\right) + \cdots + p\left(x_1, y_m\right)\right] \\
(\text{ second row }) &+ \left[p\left(x_2, y_1\right) + p\left(x_2, y_2\right) + p\left(x_2, y_3\right) + \cdots + p\left(x_2, y_m\right)\right] \\
&+ \ldots \\
(\text{ last row }) &+ \left[p\left(x_n, y_1\right) + p\left(x_n, y_2\right) + p\left(x_n, y_3\right) + \cdots + p\left(x_n, y_m\right)\right]
\end{aligned}
$$

Example: Play Craps Let X be the value on the first dice and Y the value on the second dice. Both X and Y take on values from 1 to 6. What is the joint distribution for these two variables?

Example: Play Craps Let X be the value on the first dice and Y the value on the second dice. Both X and Y take on values from 1 to 6. What is the joint distribution for these two variables?

$X \backslash Y$	1	2	3	4	5	6
1	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
2	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
3	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
4	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$

How to evaluate this?

Two-step summation (fix one first and then compute along that dimension): Let X be the value on the first dice and Y the value on the second dice. Both X and Y take on values from 1 to 6 .

$$\sum_{i=1}^{n}\sum_{j=1}^{m} p\left(x_i, y_j\right)$$

$X \backslash Y$	1	2	3	4	5	6	Total Probability
1	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$=\frac{1}{6}$
2	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$=\frac{1}{6}$
3	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$=\frac{1}{6}$
4	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$=\frac{1}{6}$
5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$=\frac{1}{6}$
6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$=\frac{1}{6}$
							$=1$

Cumulative Distribution Function

Joint CDF: $F(x, y) = \mathbb{P}[X \leq x, Y \leq y] = \mathbb{P}[\{X \leq x\} \cap \{Y \leq y\}]$

Cumulative Distribution Function

Joint CDF: $F(x, y) = \mathbb{P}[X \leq x, Y \leq y] = \mathbb{P}[\{X \leq x\} \cap \{Y \leq y\}]$

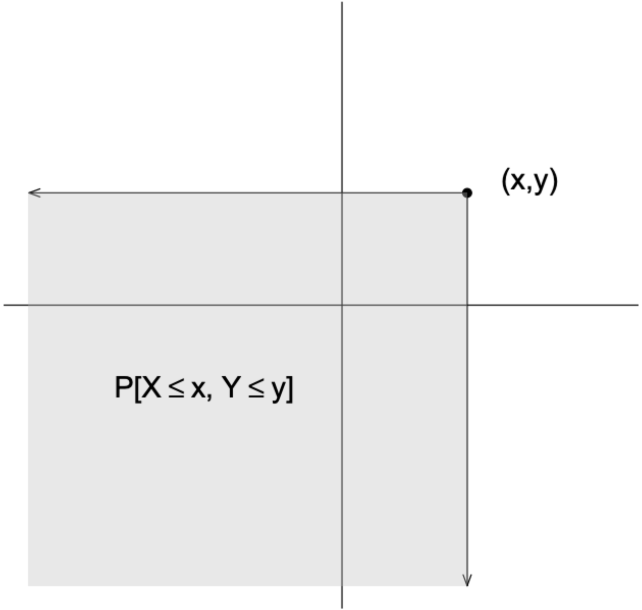
when we want to be clear that the distribution refers to the pair (X, Y), we add subscripts, e.g., $F_{X,Y}(x, y)$. When they are clear from the context, we omit the subscripts.

Cumulative Distribution Function

Joint CDF: $F(x, y) = \mathbb{P}[X \leq x, Y \leq y] = \mathbb{P}[\{X \leq x\} \cap \{Y \leq y\}]$

when we want to be clear that the distribution refers to the pair (X, Y), we add subscripts, e.g., $F_{X,Y}(x, y)$. When they are clear from the context, we omit the subscripts.

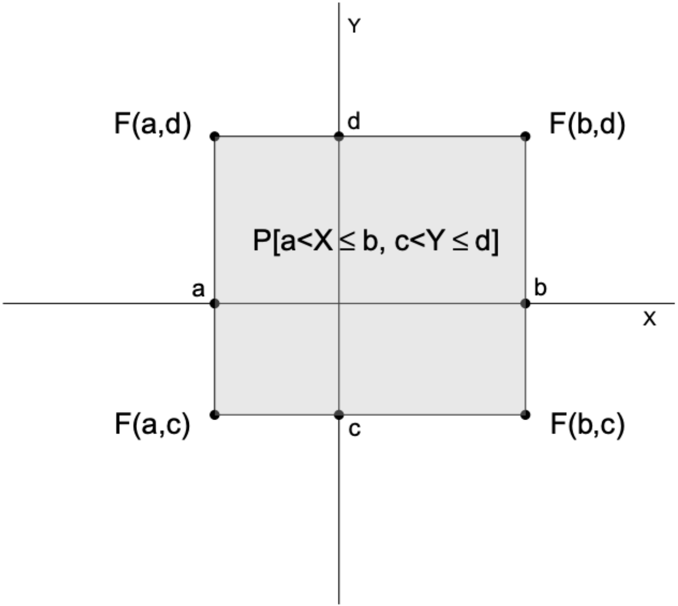
Example: $F(x, y) = (1 - e^{-x})(1 - e^{-y})$ for $x, y \geq 0$


(x,y)
P[X ≤ x, Y ≤ y]

$$
\begin{aligned}
F(x, y) &= \Pr[X \le x, Y \le y] \\
&= \sum_{x_i \le x, y_j \le y} p\left(x_i, y_j\right)
\end{aligned}
$$

Properties of CDF:

1. The distribution function is weakly increasing in each argument
2. $0 \leq F(x, y) \leq 1$
3. $\mathbb{P}[a < X \leq b, c < Y \leq d] =$
   $F(b, d) - F(b, c) - F(a, d) + F(a, c)$


Y
F(a,d)
d
F(b,d)
P[a<X ≤ b, c<Y ≤ d]
a
b
X
F(a,c)
c
F(b,c)

Example: (Play Craps)

$X \backslash Y$	1	2	3	4	5	6
1	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
2	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
3	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
4	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$

What is $F(3.5, 4) = Pr[X \leq 3.5, Y \leq 4]$?

Example: (Play Craps)

$X \backslash Y$	1	2	3	4	5	6
1	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
2	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
3	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
4	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$

What is $F(3.5, 4) = Pr[X \leq 3.5, Y \leq 4]? = \frac{12}{36}$

Implication 1: From Joint to Marginal Distribution

Although X and Y are jointly distributed random variables, we sometimes are only interested in each variable itself (either X or Y).

In this case, we need to find out the p.m.f. of X without Y, or the p.m.f. of Y without X. This is called marginal p.m.f..

It is nothing but

$\Pr[X = x_i]$

X	Probability
1	
2	
3	
4	
5	
6	

$X \backslash Y$	1	2	3	4	5	6	X	Probability
1	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	1	
2	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	2	
3	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	3	
4	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	4	
5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	5	
6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	6	

What is $\Pr[X = 1]$?

(Joint) P.m.f

$X \backslash Y$	1	2	3	4	5	6	X	Probability
1	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	1	
2	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	2	
3	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	3	
4	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	4	
5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	5	
6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	6	

What is $\Pr[X = 1]$?

$$\Pr[X = 1] = \Pr[(1,1) \cup (1,2), \ldots, \cup(1,6)] = \frac{1}{6}$$

Finite addtivity result

$$p_X(x_i) = \sum_{j=1}^{m} p(x_i, y_j)$$
$$p_Y(y_j) = \sum_{i=1}^{n} p(x_i, y_j)$$

Important Note: When you sum over the entire support of a variable, then the function is no longer a function of that variable!!!!

Integrating out the other variables.

The marginal CDF of X is

$$F_X(x) = \mathbb{P}[X \leq x] = \mathbb{P}[X \leq x, Y \leq \infty] = \lim_{y \to \infty} F(x, y)$$

Continuous Variables

1. CDF: $F(x, y) = \mathbb{P}[X \leq x, Y \leq y] = \mathbb{P}[\{X \leq x\} \cap \{Y \leq y\}]$
2. PDF: When $F(x, y)$ is continuous and differentiable its joint density $f(x, y)$ equals

$$f(x, y) = \frac{\partial^2}{\partial x \partial y} F(x, y)$$

When we want to be clear that the density refers to the pair (X, Y) we add subscripts, $f_{X,Y}(x, y)$

3. Marginal distribution:

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

and

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

Properties of PDF

1. non-negative function
2. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = 1$
3. $\int_{c}^{d}\int_{a}^{b} f(x,y)dxdy = \int_{c}^{d}\int_{a}^{b} \frac{\partial^2}{\partial x \partial y} F(x,y)dxdy = \mathbb{P}[a \leq X \leq b, c \leq Y \leq d]$

$\mathbb{P}[(X,Y) \in A] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbb{I}\{(x,y) \in A\} f(x,y)dxdy$

Homework: Marginal Distribution How to derive the marginal PDF from the marginal CDF? **Hint: Leiniz Integral Rule:**

$$\begin{aligned} F(x) &= \Pr[X \leq x] \\ &= \Pr[X \leq x, -\infty < Y < \infty] \\ &= \int_{-\infty}^{x} \int_{-\infty}^{\infty} f(u,v) du dv \\ &= \int_{-\infty}^{x} \left[\int_{-\infty}^{\infty} \boldsymbol{f(u,v) dv} \right] du \\ &= \int_{-\infty}^{x} \boldsymbol{g(u)} du \end{aligned}$$

Mixed Discrete and Continuous Variables

In modern micro-econometrics, this frequently arises. For example, you are interested in estimating the impact of college education (binary; whether or not you have college education) on wages (continuous). The most convenient way to characterize the joint distribution of mixed discrete and continuous variables is to look at

$$F(x, y) = \Pr[X \leq x, Y \leq y]$$

Joint, Marginal Distribution and Independence

The need to detec and properly measure association and dependence is an essential task in economic model building and forecasting. This is the main activity in empirical economics whether one is searching for contemporaneous relations among economic variables, working with cross-section data or time series, or determining dynamic structure or any patterns.

Numerous diagnostic procedures, such as the DW test, LM tests, are used to examine model residuals for departure from independence, i.i.d., martingale difference property, etc.

We can at least define when two variables are NOT related, or independent of each other.

1. When two events are said to be independent of each other, what this means is that the probability that one event occurs in no way affects the probability of the other event occurring.
2. When two events are said to be dependent, the probability of one event occurring influences the likelihood of the other event (either smaller or larger than the likelihood of the other event alone!).

$\perp\!\!\!\perp$

Equivalent Definitions $X \perp Y$

1

$$F(x, y) = F_X(x) F_Y(y)$$

2

$$\text{Discrete Variable } p(x_i, y_j) = p_X(x_i)\, p_Y(y_j), \forall x, y$$

$$\text{Continuous Variable } f(x, y) = f(x) f(y) \quad \forall x, y$$

3

$$\begin{aligned} \Pr[X \in A, Y \in B] &= \Pr[X \in A \cap Y \in B] \\ &= \Pr[X \in A] \cdot \Pr[Y \in B] \end{aligned}$$

✓

Example:

$$\Pr[X \in (x, x+\epsilon), Y = 1] = \Pr[X \in (x, x+\epsilon)] \cdot \Pr[Y = 1]$$

$$p(x,y) = p(x)\,p(y)$$ ✓

$X \backslash Y$	1	2	3	4	5	6	$p(x_i)$
1	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
2	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
3	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
4	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
$p(y_j)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	

(x, y)

$$p(x=1, y=1) = \frac{1}{36}$$

$$p(x=1)\cdot p(y=1) = \frac{1}{6}\,\frac{1}{6} = \frac{1}{36}$$

Homework Show that if $X \perp Y$, then $X \perp -Y$. You will use this result for next homework question.

Pr[ ].

4

$X \perp Y$ [Factor Theorem.]

Casella and Berger, p.153

Lemma 4.2 .7 Let (X, Y) be a bivariate random vector with joint pdf or pmf $f(x, y)$ Then X and Y are independent random variables if and only if there exist functions $g(x)$ and $h(y)$ such that, for every $x \in \Re$ and $y \in \Re$

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

$$f(x,y) = \underbrace{\boxed{\exp(-x)}}_{g(x)}\ \underbrace{\boxed{\exp(-y)}}_{h(y)}$$

$X \perp Y$

Casella and Berger, p.155

Let $g(x)$ be a function only of x and $h(y)$ be a function only of y, then

$$\mathbf{E}(g(X)h(Y)) = (\mathrm{E}g(X))(Eh(Y))$$

$$\iint g(x)h(y)f(x,y)\,dx\,dy \qquad E[g(x)]\,E[h(y)] = \int h(y)f(y)\,dy$$

$$\iint g(x)h(y)f(x)\cdot f(y)\,dx\,dy$$

$$\int \boxed{\left[\int g(x)f(x)\,dx\right]} \cdot h(y)f(y)\,dy$$

Proof:

$$\begin{aligned}
\mathrm{E}(g(X)h(Y)) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f(x,y)dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f_X(x)f_Y(y)dxdy \\
&= \int_{-\infty}^{\infty} h(y)f_Y(y)\int_{-\infty}^{\infty} g(x)f_X(x)dxdy \\
&= \left(\int_{-\infty}^{\infty} g(x)f_X(x)dx\right)\left(\int_{-\infty}^{\infty} h(y)f_Y(y)dy\right) \\
&= (\mathrm{E}g(X))(\mathrm{E}h(Y))
\end{aligned}$$

Convolution

Homework: Moment generating functions for $Z = \boxed{X + Y}$ (convolution), when X, Y are independent.

$$\mathbb{E}[e^{tZ}] = \mathbb{E}[e^{tX}]\mathbb{E}[e^{tY}]$$

✓

$X \perp Y \Rightarrow X \perp -Y.$

Homework: Show that if X and Y are independent and follow the same distribution, then $Z = X - Y$ is symmetric around zero. **Hint: Just show that** Z and $-Z$ have the same distribution, which then implies symmetry around zero. Further hint: show that they have the same moment generating functions.

Application: This has been applied a lot in the recent development, e.g., in semiparametric models of censored outcomes or rank estimation.

$Z \quad -Z \Rightarrow E[\varepsilon_i - \varepsilon_j | x] = 0$

$\varepsilon_i - \varepsilon_j | x$ is symmetric around zero.

$$E\left[e^{tX}\right] = \boxed{e^{t\mu + \frac{t^2\sigma^2}{2}}}$$

Homework: Show that the sum of two independent, normal variables are still normally distributed.

Distribution

Note that independence, be it unconditional or conditional, is the most important concept in our course and later microeconometrics as well!

$$Y = H(\varepsilon, X)$$

Wage

$$\varepsilon \perp Z$$

$$\varepsilon \perp X$$