Two random variables (B)

Multiple random variables

Francesco Corona

Department of Computer Science
Federal University of Ceará, Fortaleza

Conditional distributions and expectations

Two random variables

Conditional distributions and expectations

We have discussed the joint probability distribution of a pair of RVs

$\rightsquigarrow$ How to get individual (marginal) distributions from the joint

We discuss conditional distributions

- Distribution of one of the two random variables
- ..., when the other one has taken a specific value

First the discrete case, as it follows from conditional probability

- Then, the continuous case

Conditional distributions and expectations (cont.)

Discrete case

Let X_1 and X_2 be two random variables of the discrete type

Let $p_{X_1,X_2}(x_1, x_2)$ indicate their joint PMF

- Positive over $\mathcal{S}$, zero elsewhere

Let $p_{X_1}(x_1)$ and $p_{X_2}(x_2)$ be the PMFs of X_1 and X_2

- Positive over $\mathcal{S}_{X_1}$ and $\mathcal{S}_{X_2}$, zero elsewhere

Conditional distributions and expectations (cont.)

Let x_1 be a point in the support $\mathcal{S}_{X_1}$ of X_1, $p_{X_1}(x_1) > 0$

By using the concept of conditional probability,

$$P(X_2 = x_2 | X_1 = x_1) = \frac{P(X_1 = x_1, X_2 = x_2)}{P(X_1 = x_1)} = \frac{p_{X_1,X_2}(x_1, x_2)}{p_{X_1}(x_1)}$$

This holds true for all x_2 in the support $\mathcal{S}_{X_2}$ of X_2

Conditional distributions and expectations (cont.)

Example

$$P(X_2 = x_2 | X_1 = x_1) = \frac{P(X_1 = x_1, X_2 = x_2)}{P(X_1 = x_1)} = \frac{p_{X_1,X_2}(x_1, x_2)}{p_{X_1}(x_1)}$$

Consider the discrete random vector (X_1, X_2) with the tabulated PMF

x_1/x_2	0	1	2	3	$p_{X_1}(x_1)$
0	1/8	1/8	0	0	2/8
1	0	2/8	2/8	0	4/8
2	0	0	1/8	1/8	2/8
$p_{X_2}(x_2)$	1/8	3/8	3/8	1/8	

Conditional distributions and expectations (cont.)

We can define the function

$$p_{X_2|X_1}(x_2|x_1) = \frac{p_{X_1,X_2}(x_1,x_2)}{p_{X_1}(x_1)}, \quad x_2 \in \mathcal{S}_{X_2} \tag{1}$$

This function is a valid PMF of the discrete RV

- It is non-negative
- It sums up to one

$$\sum_{x_2} p_{X_2|X_1}(x_2|x_1) = \sum_{x_2} \frac{p_{X_1,X_2}(x_1,x_2)}{p_{X_1}(x_1)} = \frac{1}{p_{X_1}(x_1)} \sum_{x_2} p_{X_1,X_2}(x_1,x_2)$$
$$= \frac{1}{p_{X_1}(x_1)} p_{X_1}(x_1) = 1$$

This is true for any fixed x_1 with $p_{X_1}(x_1) > 0$

$p_{X_2|X_1}(x_2|x_1)$ is the **conditional probability mass function** (**conditional PMF**) of discrete type RV X_2, given that discrete type RV $X_1 = x_1$

Conditional distributions and expectations (cont.)

Given $x_2 \in \mathcal{S}_{X_2}$, we can similarly define the symbol $p_{X_1|X_2}(x_1|x_2)$

$$p_{X_1|X_2}(x_1|x_2) = \frac{p_{X_1,X_2}(x_1,x_2)}{p_{X_2}(x_2)}, \quad x_1 \in \mathcal{S}_{X_1}$$

This is true for any fixed x_2 with $p_{X_2}(x_2) > 0$

$p_{X_1|X_2}(x_1|x_2)$ is the **conditional probability mass function** (**conditional PMF**) of discrete type RV X_1, given the discrete type RV $X_2 = x_2$

Conditional distributions and expectations (cont.)

Example

$$p_{X_2|X_1}(x_2|x_1) = p_{X_1,X_2}(x_1,x_2)/p(x_1)$$

x_1/x_2	0	1	2	3	
0	1/2	1/2	0	0	1
1	0	1/2	1/2	0	1
2	0	0	1/2	1/2	1

$$p_{X_1|X_2}(x_1|x_2) = p_{X_1,X_2}(x_1,x_2)/p(x_2)$$

x_1/x_2	0	1	2	3	
0	1	1/3	0	0	
1	0	2/3	1/2	2/3	
2	0	0	1/2	1/3	
	1	1	1	1	

Conditional distributions and expectations (cont.)

We may abbreviate the notation

- $p_{X_1|X_2}(x_1|x_2)$ by $p_{1|2}(x_1|x_2)$
- $p_{X_2|X_1}(x_2|x_1)$ by $p_{2|1}(x_2|x_1)$

Similarly, we may use $p_1(x_1)$ and $p_2(x_2)$ to indicate the marginal PMFs

Conditional distributions and expectations (cont.)

Continuous case

Let X_1 and X_2 indicate two random variables of the continuous type

- Let $f_{X_1}(x_1)$ and $f_{X_2}(x_2)$ be the marginal density functions
- Let $f_{X_1,X_2}(x_1, x_2)$ be the joint PDF

A definition of conditional probability density functions for continuous RVs

$\rightsquigarrow$ When $f_{X_1}(x_1) > 0$, we define the symbol $f_{X_2|X_1}(x_2|x_1)$

$$f_{X_2|X_1}(x_2|x_1) = \frac{f_{X_1,X_2}(x_1, x_2)}{f_{X_1}(x_1)} \tag{2}$$

Conditional distributions and expectations (cont.)

$$f_{X_2|X_1}(x_2|x_1) = \frac{f_{X_1,X_2}(x_1,x_2)}{f_{X_1}(x_1)}$$

x_1 is understood of as having taken a(ny) fixed value for which $f_{X_1}(x_1) > 0$

$$\begin{aligned}\int_{-\infty}^{\infty} f_{X_2|X_1}(x_2|x_1)\mathrm{d}x_2 &= \int_{-\infty}^{\infty} \frac{f_{X_1,X_2}(x_1,x_2)}{f_{X_1}(x_1)}\mathrm{d}x_2 \\ &= \frac{1}{f_{X_1}(x_1)}\int_{-\infty}^{\infty} f_{X_1,X_2}(x_1,x_2)\mathrm{d}x_2 \\ &= \frac{1}{f_{X_1}(x_1)} f_{X_1}(x_1) = 1\end{aligned}$$

Conditional distributions and expectations (cont.)

For $f_{X_2}(x_2) > 0$, the conditional PDF of the random variable of the continuous type X_1, given that the continuous type random variable $X_2 = x_2$

$$f_{X_1|X_2}(x_1|x_2) = \frac{f_{X_1,X_2}(x_1, x_2)}{f_{X_2}(x_2)}, \quad f_{X_2}(x_2) > 0$$

Conditional distributions and expectations (cont.)

Again, we may abbreviate the notation

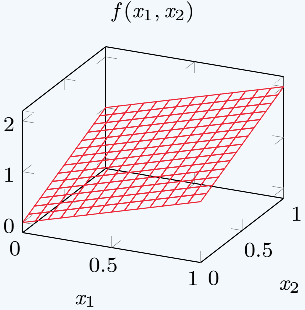
- $f_{X_1|X_2}(x_1|x_2)$ by $f_{1|2}(x_1|x_2)$
- $f_{X_2|X_1}(x_2|x_1)$ by $f_{2|1}(x_2|x_1)$

Similarly, we may use $f_1(x_1)$ and $f_2(x_2)$ to indicate the marginal PDFs

Conditional distributions and expectations (cont.)

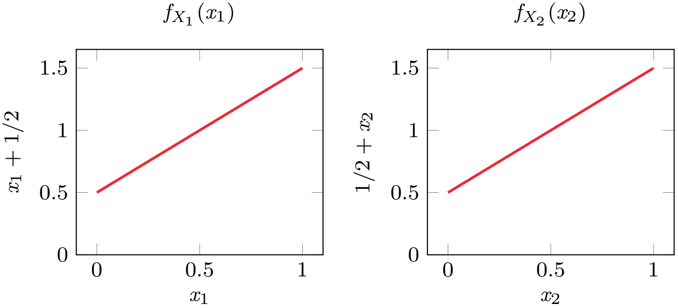
Example

Let X_1 and X_2 have joint PDF



$$f(x_1, x_2) = \begin{cases} x_1 + x_2, & 0 < x_1, x_2 < 1 \\ 0, & \text{elsewhere} \end{cases}$$

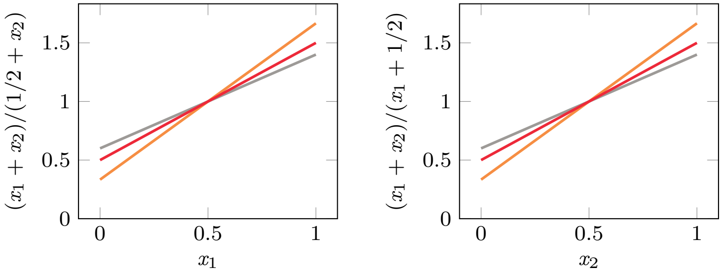
Conditional distributions and expectations (cont.)



Conditional distributions and expectations (cont.)

$$\rightsquigarrow f_{X_2|X_1}(x_2|x_1) = \frac{f_{X_1,X_2}(x_1,x_2)}{f_{X_1}(x_1)}$$

$f_{X_1|X_2}(x_1|X_2 = 1/4|1/2|3/4)$ $\quad$ $f_{X_2|X_1}(x_2|X_1 = 1/4|1/2|3/4)$



Conditional distributions and expectations (cont.)

Each of $f_{2|1}(x_2|x_1)$ and $f_{1|2}(x_1|x_2)$ is a PDF of a random variable

- $\rightsquigarrow$ Each has all the properties of a PDF

This means that we can calculate probabilities and expectations

Conditional distributions and expectations (cont.)

Suppose that the random variables X_1 and X_2 are of the continuous type

The conditional probability that $a < X_2 < b$, given that $X_1 = x_1$

$$p(a < X_2 < b | X_1 = x_1) = \int_a^b f_{2|1}(x_2|x_1) \mathrm{d}x_2$$

The conditional probability that $c < X_1 < d$, given that $X_2 = x_2$

$$P(c < X_1 < d | X_2 = x_2) = \int_c^d f_{1|2}(x_1|x_2) \mathrm{d}x_1$$

Conditional distributions and expectations (cont.)

Let $u(X_2)$ be a function of X_2

The expectation of $u(X_2)$ (if it exists)

$$E\big[u(X_2)\big] = \int_{-\infty}^{\infty} u(x_2) f_2(x_2) \mathrm{d}x_2$$

The conditional expectation of $u(X_2)$, given that $X_1 = x_1$ (if it exists)

$$E\big[u(X_2)|x_1\big] = \int_{-\infty}^{\infty} u(x_2) f_{2|1}(x_2|x_1) \mathrm{d}x_2$$

$E\big[u(X_2)|x_1\big]$ is a function of x_1

Conditional distributions and expectations (cont.)

Mean and variance of the conditional distribution of X_2, given $X_1 = x_1$

$$E(X_2|x_1) = \int_{-\infty}^{\infty} x_2 f_{2|1}(x_2|x_1) \mathrm{d}x_2$$

$$\mathrm{Var}(X_2|x_1) = E\big\{\big[X_2 - E(X_2|x_1)\big]^2 \big| x_1\big\} = E(X_2^2|x_1) - \big[E(X_2|x_1)\big]^2$$

(If they exist)

AKA the conditional mean and conditional variance of X_2, given $X_1 = x_1$

Conditional distributions and expectations (cont.)

In a similar fashion, the conditional expectation of $u(X_1)$, given $X_2 = x_2$

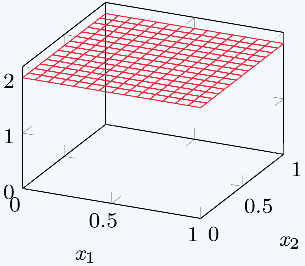
$$E\big[u(X_1)|x_2\big] = \int_{-\infty}^{\infty} u(x_1) f_{1|2}(x_1|x_2) \mathrm{d}x_1$$

Conditional distributions and expectations (cont.)

Example

Let X_1 and X_2 have the joint PDF

$f(x_1, x_2)$ with $(x_1, x_2) \in \mathcal{R}^2$



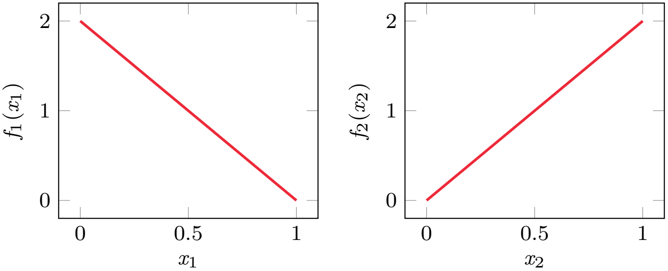
$$f(x_1, x_2) = \begin{cases} 2, & 0 < x_1 < x_2 < 1 \\ 0, & \text{elsewhere} \end{cases}$$

Conditional distributions and expectations (cont.)

The marginal probability density functions

$$f_1(x_1) = \begin{cases} \int_{x_1}^{1} f_{12}(x_1, x_2)\mathrm{d}x_2 = \int_{x_1}^{1} 2\mathrm{d}x_2 = 2(1-x_1), & 0 < x_1 < 1 \\ 0, & \text{elsewhere} \end{cases}$$

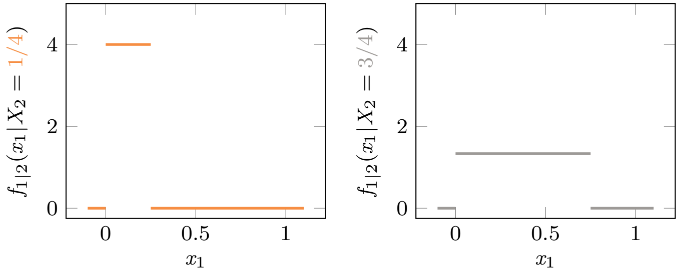
$$f_2(x_2) = \begin{cases} \int_{0}^{x_2} f_{12}(x_1, x_2)\mathrm{d}x_1 = \int_{0}^{x_2} 2\mathrm{d}x_1 = 2x_2, & 0 < x_2 < 1 \\ 0, & \text{elsewhere} \end{cases}$$



Conditional distributions and expectations (cont.)

The conditional PDF of X_1, given $X_2 = x_2$ with $0 < x_2 < 1$

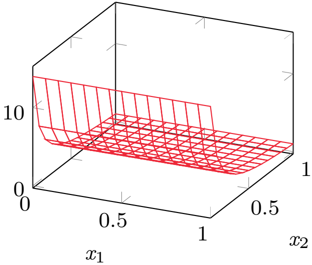
$$f_{1|2}(x_1|x_2) = \begin{cases} 2/2x_2 = 1/x_2, & 0 < x_1 < x_2 < 1 \\ 0, & \text{elsewhere} \end{cases}$$



Conditional distributions and expectations (cont.)

$$f_{1|2}(x_1|x_2) = \begin{cases} 2/2x_2 = 1/x_2, & 0 < x_1 < x_2 < 1 \\ 0, & \text{elsewhere} \end{cases}$$

$f_{1|2}(x_1|x_2)$ with $(x_1, x_2) \in \mathcal{R}^2$

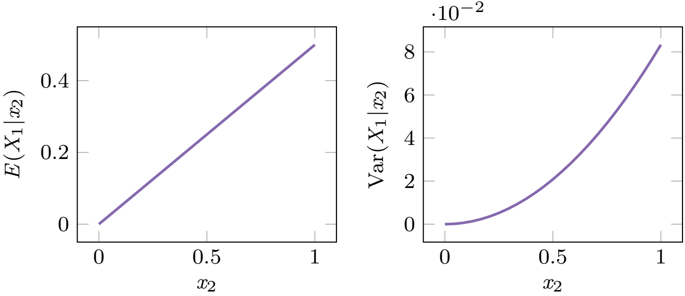


Conditional distributions and expectations (cont.)

The conditional mean and the conditional variance of X_1, given $X_2 = x_2$

$$E(X_1|x_2) = \int_{-\infty}^{\infty} x_1 f_{1|2}(x_1|x_2)\mathrm{d}x_1 = \int_0^{x_2} x_1 \big(1/x_2\big)\mathrm{d}x_1$$
$$= x_2/2, \quad 0 < x_2 < 1$$

$$\mathrm{Var}(X_1|x_2) = E\big\{\big[X_1|x_2 - E(X_1|x_2)\big]^2\big\} = \int_0^{x_2} \big(x_1 - x_2/2\big)^2 \big(1/x_2\big)\mathrm{d}x_1$$
$$= x_2^2/12, \quad 0 < x_2 < 1$$



Conditional distributions and expectations (cont.)

We can also compare these values

- $P(0 < X_1 < 1/2|X_2 = 3/4)$
- $P(0 < X_1 < 1/2)$

We have,

$$P(0 < X_1 < 1/2|X_2 = 3/4) = \int_0^{1/2} f_{1|2}(x_1|3/4)\mathrm{d}x_1 = \int_0^{1/2} (4/3)\mathrm{d}x_1$$
$$= 2/3$$

But,

$$P(0 < X_1 < 1/2) = \int_0^{1/2} f_1(x_1)\mathrm{d}x_1 = \int_0^{1/2} 2(1 - x_1)\mathrm{d}x_1$$
$$= 3/4$$

Conditional distributions and expectations (cont.)

Since, $E(X_2|x_1)$ is a function of x_1, then $E(X_2|X_1)$ is a random variable

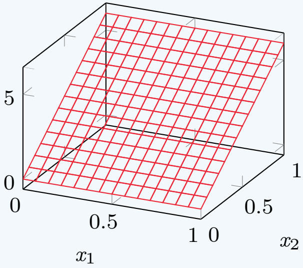
- It has its own distribution, mean and variance

Conditional distributions and expectations (cont.)

Example

Let X_1 and X_2 be two random variables with the joint PDF

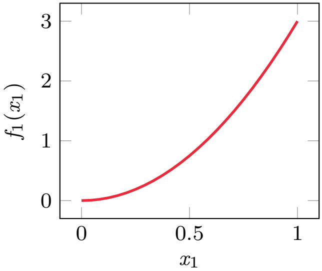
$f(x_1, x_2)$ with $(x_1, x_2) \in \mathcal{R}^2$



$$f(x_1, x_2) = \begin{cases} 6x_2, & 0 < x_2 < x_1 < 1 \\ 0, & \text{elsewhere} \end{cases}$$

Conditional distributions and expectations (cont.)

The marginal PDF of X_1



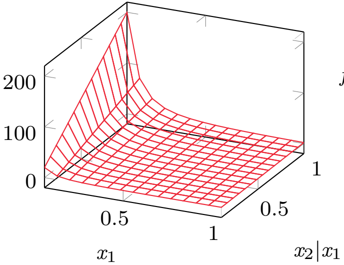
$$f_1(x_1) = \int_0^{x_1} 6x_2 \mathrm{d}x_2 = 3x_1^2$$

for $0 < x_1 < 1$

Conditional distributions and expectations (cont.)

The conditional PDF of X_2, given $X_1 = x_1$

$f_{2|1}(x_2|x_1)$ with $(x_1, x_2) \in \mathcal{R}^2$



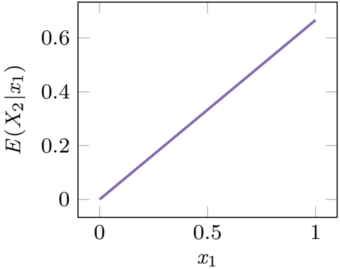
$$f_{2|1}(x_2|x_1) = 6x_2/3x_1^2 = 2x_2/x_1^2$$

for $0 < x_2 < x_1$, with $0 < x_1 < 1$

Conditional distributions and expectations (cont.)

The conditional mean of X_2, given $X_1 = x_1$

$$E(X_2|x_1) = \int_0^{x_1} x_2 \Big(\frac{2x_2}{x_1^2}\Big) \mathrm{d}x_2 = 2/3x_1, \quad 0 < x_1 < 1$$



Conditional distributions and expectations (cont.)

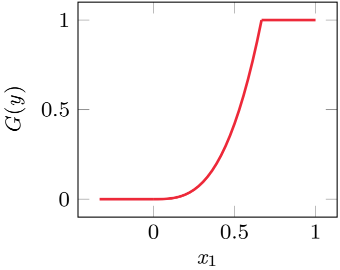
$E(X_2|X_1)$ is a random variable, let it be Y, the CDF of $Y = 2X_1/3$

$$G(y) = P(Y \leq y) = P\Big(X_1 \leq \frac{3y}{2}\Big), \quad 0 \leq y < 2/3$$

From the PDF $f_1(x_1)$, we have

$$G(y) = \int_0^{3y/2} 3x_1^2 \mathrm{d}x_1 = 27y^3/8, \quad 0 \leq y < 2/3$$

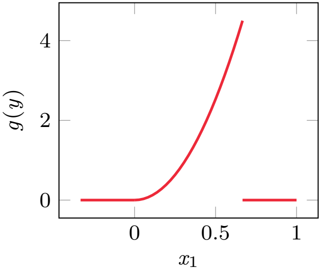
$G(y) = 0$ if $y < 0$ and $G(y) = 1$ if $2/3 < y$



Conditional distributions and expectations (cont.)

The PDF of $Y = 2X_1/3$

$$g(y) = 81y^2/8, \quad 0 \leq y < 2/3 \text{ (zero elsewhere)}$$



The mean and variance of $Y = 2X_1/3$

$$E(Y) = \int_0^{2/3} y\Big(81y^2/8\Big)\mathrm{d}y = 1/2$$

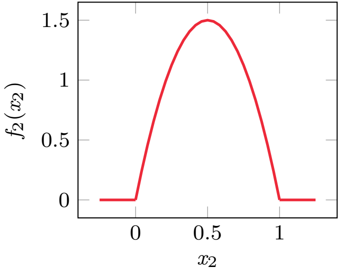
$$\mathrm{Var}(Y) = \int_0^{2/3} y^2\Big(81y^2/8\Big)\mathrm{d}y - (1/2)(1/2) = 1/60$$

Conditional distributions and expectations (cont.)

The marginal PDF of X_2

$$f_2(x_2) = \int_{x_2}^{1} 6x_2 \mathrm{d}x_1 = 6x_2(1 - x_2), \quad 0 < x_2 < 1$$

and zero elsewhere



Conditional distributions and expectations (cont.)

We can show that $E(X_2) = 1/2$ and $\text{Var}(X_2) = 1/20$

That is,

$$E(Y) = E\big[E(X_2|X_1)\big] = E(X_2)$$

and

$$\text{Var}(Y) = \text{Var}\big[E(X_2|X_1)\big] \leq \text{Var}(X_2)$$

■

Conditional distributions and expectations (cont.)

Theorem 1.1

Let (X_1, X_2) be a random vector such that the variance of X_2 is finite

Then,

(a) $E\big[E(X_2|X1)\big] = E(X_2)$

(b) $Var\big[E(X_2|X_1)\big] \leq Var(X_2)$

Proof

For the continuous case

We first prove (a)

$$E(X_2) = \int_{-\infty}^{\infty} \int_{\infty}^{\infty} x_2 f(x_1, x_2) \mathrm{d}x_2 \mathrm{d}x_1 = \int_{-\infty}^{\infty} \Big[\int_{-\infty}^{\infty} x_2 \frac{f(x_1, x_2)}{f_1(x_1)} \Big] f_1(x_1) \mathrm{d}x_1$$
$$= \int_{-\infty}^{\infty} E(X_2|x_1) f_1(x_1) \mathrm{d}x_1 = E\big[E(X_2|X_1)\big]$$

Conditional distributions and expectations (cont.)

We then prove (b)

$$\begin{aligned}\text{Var}(X_2) =& E\big[(X_2-\mu_2)^2\big] = E\big\{\big[X_2 - E(X_2|X_1) + E(X_2|X_1) - \mu_2\big]^2\big\}\\ =& E\big\{\big[X_2 - E(X_2|X_1)\big]^2\big\} + \big\{\big[E(X_2|X_1)-\mu_2\big]^2\big\} + \\ & 2E\big\{\big[X_2 - E(X_2|X_1)\big]\big[E(X_2|X_1)-\mu_2\big]\big\}\end{aligned}$$

Conditional distributions and expectations (cont.)

The last term on the LHS is zero

$$2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Big[x_2 - E(X_2|x_1)\Big]\Big[E(X_2|x_1) - \mu_2\Big]f(x_1, x_2)\mathrm{d}x_2\mathrm{d}x_1$$
$$= 2\int_{-\infty}^{\infty}\Big[E(X_2|x_1) - \mu_2\Big]\Big\{\Big[x_2 - E(X_2|x_1)\Big]\frac{f(x_1, x_2)}{f_1(x_1)}\mathrm{d}x_2\Big\}f_1(x_1)\mathrm{d}x_1$$

$E(X_2|x_1)$ is the conditional mean of X_2, given $X_1 = x_1$

Consider the expression within inner brackets

$$E(X_2|x_1) - E(X_2|x_1) = 0$$

Thus, the double integral is zero

Conditional distributions and expectations (cont.)

Accordingly, we have

$$\text{Var}(X_2) = E\Big\{\big[X_2 - E(X_2|X_1)\big]^2\Big\} + E\Big\{\big[E(X_2|X_1) - \mu_2\big]^2\Big\}$$

The first term on the RHS is non-negative

- It is the expected value of a non-negative function

$\rightsquigarrow$ $\big[X_2 - E(X_1|X_1)\big]^2$

Since $E\big[E(X_2|X_1)\big] = \mu_2$, it follows that the second term is $\text{Var}\big[E(X_2|X_1)\big]$

Hence,

$$\text{Var}(X_2) \geq \text{Var}[E(X_2|X_1)]$$

■

Conditional distributions and expectations (cont.)

This result has an intuitive interpretation

Both the random variables X_2 and $E(X_2|X_1)$ have the same mean μ_2

Suppose μ_2 is unknown

⇝ We could use either of the two RVs to guess at μ_2

We proved that $\text{Var}(X_2) \geq \text{Var}\big[E(X_2|X_1)\big]$

⇝ We would put more belief in $E(X_2|X_1)$

Remark

If we observe (X_1, X_2) to be (x_1, x_2), we prefer to use $E(X_2|x_1)$ to x_2

- A less wobbly a guess at μ_2

The correlation coefficient

Two random variables

The correlation coefficient

The next results are commonly shown in terms of X and Y

- We use those letters also here (rather than X_1 and X_2)

We discuss the concepts for continuous and discrete cases jointly

- We only use notation for the continuous case
- Properties hold for the discrete case

The correlation coefficient (cont.)

Let X and Y have the joint PDF $f(x, y)$

Let $u(x, y)$ be a function of x and y

- We defined the mathematical expectation $E\big[u(X, Y)\big]$
- The definition of E is subjected to its existence[1]

The means of X and Y (μ_1 and μ_2)

$\rightsquigarrow$ Set function $u(x, y)$ equal to x and y, respectively

The variances of X and Y (σ_1^2 and σ_2^2)

$\rightsquigarrow$ Set function $u(x, y)$ equal to $(x - \mu_1)^2$ and $(y - \mu_2)^2$, respectively

[1] From now on, existence of all mathematical expectations is assumed.

The correlation coefficient (cont.)

Consider the mathematical expectation

$$
\begin{aligned}
E\big[(X-\mu_1)(Y-\mu_2)\big] &= E(XY - \mu_2 X - \mu_1 Y + \mu_1\mu_2) \\
&= E(XY) - \mu_2 E(X) - \mu_1 E(Y) + \mu_1\mu_2 \\
&= E(XY) - \mu_1\mu_2
\end{aligned}
$$

We call this number the **covariance** of X and Y

- ⇝ It is indicated by $\text{cov}(X, Y)$

The correlation coefficient (cont.)

Suppose each of σ_1 and σ_2 is positive

Then, we can define

$$\rho = \frac{E\big[(X-\mu_1)(Y-\mu_2)\big]}{\sigma_1\sigma_2} = \frac{\text{cov}(X,Y)}{\sigma_1\sigma_2}$$

We call this number the **correlation coefficient** of X and Y

The correlation coefficient (cont.)

$$\rightsquigarrow \quad E\big[(X-\mu_1)(Y-\mu_2)\big] = E(XY) - \mu_1\mu_2$$

$$\rightsquigarrow \quad \rho = \frac{E\big[(X-\mu_1)(Y-\mu_2)\big]}{\sigma_1\sigma_2} = \frac{\text{cov}(X,Y)}{\sigma_1\sigma_2}$$

Remark

The expected value of the product of two random variables

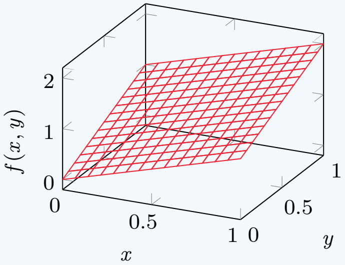
$$\rightsquigarrow \quad E(XY) = \mu_1\mu_2 + \rho\sigma_1\sigma_2 = \mu_1\mu_2 + \text{cov}(X,Y)$$

It is equal to the product of their expectations, plus their covariance

The correlation coefficient (cont.)

Example

Let the random variables X and Y have the joint PDF



$$f(x, y) = \begin{cases} x + y, & 0 < x, y < 1 \\ 0, & \text{elsewhere} \end{cases}$$

We want to compute the correlation coefficient ρ of X and Y

The correlation coefficient (cont.)

$$\rightsquigarrow \begin{cases} \mu_1 = E(X) = \int_0^1 \int_0^1 x(x+y)\mathrm{d}x\mathrm{d}y = 7/12 \\ \sigma_1^2 = E(X^2) - \mu_1^2 = \int_0^1 \int_0^1 x^2(x+y)\mathrm{d}x\mathrm{d}y - (7/12)^2 = 11/144 \end{cases}$$

$$\rightsquigarrow \begin{cases} \mu_2 = E(Y) = \int_0^1 \int_0^1 y(x+y)\mathrm{d}y\mathrm{d}x = 7/12 \\ \sigma_1^2 = E(Y^2) - \mu_2^2 = \int_0^1 \int_0^1 y^2(x+y)\mathrm{d}y\mathrm{d}x - (7/12)^2 = 11/144 \end{cases}$$

The covariance of X and Y

$$E(XY) - \mu_1\mu_2 = \int_0^1 \int_0^1 xy(x+y)\mathrm{d}x\mathrm{d}y - \big(7/12\big)\big(7/12\big) = -1/144$$

The correlation coefficient of X and Y follows

$$\rho = \frac{E\big[(X-\mu_1)(Y-\mu_2)\big]}{\sigma_1\sigma_2} = \frac{-1/144}{\sqrt{\big(11/144\big)}\sqrt{\big(11/144\big)}} = -1/11$$

■

The correlation coefficient (cont.)

Remark

For certain kinds of distributions of two random variables X and Y

- ρ proves to be a very useful characteristic of the distribution
- The formal definition of ρ does not reveal this explicitly

We make some superficial observations about ρ

- They will be explored more deeply later

The correlation coefficient (cont.)

Consider a joint distribution of two random variables X and Y

Suppose that a correlation coefficient of the two RVs exists

- That is, both of the variances are positive

Then, it can be shown that ρ satisfies the relation

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

Consider the case in which $\rho = 1$

This can be understood as there is a line with equation $y = a + bx, b > 0$ whose graph contains all of the probability of the distribution of X and Y

- In this extreme case, we have $P(Y = a + bX) = 1$

In the case that $\rho = -1$, we have the same state of things, except that $b < 0$

The correlation coefficient (cont.)

Consider the case that ρ does not have one of its extreme values

Can this be understood as there a line in the $x - y$ plane such that the probability for X and Y tends to be concentrated in a band about this line?

- This is the case, under certain restrictive conditions

Under those same conditions, we understand ρ as a measure of the strength of the concentration of the probability for X and Y about such a line

The correlation coefficient (cont.)

Let $f(x, y)$ denote the joint PDF of two random variables X and Y

Let $f_1(x)$ denote the marginal PDF of X

The conditional PDF of Y, given $X = x$

$$f_{2|1}(y|x) = \frac{f(x, y)}{f_1(x)}, \quad \text{at points in which } f_1(x) > 0$$

The conditional mean of Y, given $X = x$

$$E(Y|x) = \int_{-\infty}^{\infty} y f_{2|1}(y|x) \mathrm{d}y = \frac{\int_{-\infty}^{\infty} y f(x, y) \mathrm{d}y}{f_1(x)}$$

The correlation coefficient (cont.)

$$E(Y|x) = \int_{-\infty}^{\infty} y f_{2|1}(y|x) \mathrm{d}y = \frac{\int_{-\infty}^{\infty} y f(x,y) \mathrm{d}y}{f_1(x)}$$

This conditional mean of Y, given $X = x$, is a function of x, $u(x)$

Let $u(x)$ be a linear function, $u(x) = a + bx$

- ⇝ The conditional mean of Y is linear in x

When $u(x) = a + bx$, the constants a and b have simple values

The correlation coefficient (cont.)

Theorem

Suppose (X, Y) have a joint distribution

Let the variances of X and Y be finite and positive

Indicate the means and variances of X and Y by the usual symbols

- μ_1 *and* μ_2*, for the means*
- σ_1^2 *and* σ_2^2*, for the variances*

Let ρ be the correlation coefficient between X and Y

If $E\big(Y|X\big)$ is linear in X, then

$$E(Y|X) = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(X - \mu_1) \tag{3}$$

and

$$E\big[Var\big(Y|X\big)\big] = \sigma_2^2(1 - \rho^2) \tag{4}$$

The correlation coefficient (cont.)

Proof

The proof is given for the continuous case

Let $E(Y|x) = a + bx$

From

$$E(Y|x) = \frac{\int_{-\infty}^{\infty} y f(x,y) \mathrm{d}y}{f_1(x)} = a + bx,$$

we have

$$\int_{-\infty}^{\infty} y f(x,y) \mathrm{d}y = (a + bx) f_1(x) \qquad (5)$$

The correlation coefficient (cont.)

$$\int_{-\infty}^{\infty} yf(x,y)\mathrm{d}y = (a+bx)f_1(x)$$

If both members are integrated on x, we have

$$E(Y) = a + bE(X)$$

or

$$\mu_2 = a + b\mu_1 \tag{6}$$

We know that $\mu_1 = E(X)$ and $\mu_2 = E(Y)$

The correlation coefficient (cont.)

$$\int_{-\infty}^{\infty} y f(x, y) \mathrm{d}y = (a + bx) f_1(x)$$

If both members are first multiplied by x and then integrated on x, we have

$$E(XY) = aE(X) + bE(X^2)$$

or

$$\rho\sigma_1\sigma_2 + \mu_1\mu_2 = a\mu_1 + b(\sigma_1^2 + \mu_1^2) \quad (7)$$

We know that $\rho\sigma_1\sigma_2$ is the covariance of X and Y

The correlation coefficient (cont.)

The simultaneous solution of Equation (6) and (7) yields

$$\begin{cases} a = \mu_2 - \rho \dfrac{\sigma_2}{\sigma_1} \mu_1 \\ b = \rho \dfrac{\sigma_2}{\sigma_1} \end{cases}$$

These values give the first result, Equation (3)

The correlation coefficient (cont.)

The conditional variance of Y

$$\begin{aligned} \text{Var}(Y|x) &= \int_{-\infty}^{\infty} \Big[y - \mu_2 - \rho \frac{\sigma_2}{\sigma_1}(x - \mu_1) \Big]^2 f_{2|1}(y|x) \text{d}y \\ &= \frac{\int_{-\infty}^{\infty} \Big[(y - \mu_2) - \rho \frac{\sigma_2}{\sigma_1}(x - \mu_1) \Big]^2 f(x, y) \text{d}y}{f_1(x)} \end{aligned} \quad (8)$$

This variance is non-negative and it is (at most) a function of x alone

Independence (cont.)

If multiplied by $f_1(x)$ and integrated on x

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Big[(y-\mu_2)-\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\Big]^2 f(x,y)\mathrm{d}y\mathrm{d}x =$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Big[(y-\mu_2)^2-2\rho\frac{\sigma_2}{\sigma_1}(y-\mu_2)(x-\mu_1)+\rho^2\frac{\sigma_2^2}{\sigma_1^2}(x-\mu_1)^2\Big]f(x,y)\mathrm{d}y\mathrm{d}x$$

$$= E\big[(Y-\mu_2)^2\big]-2\rho\frac{\sigma_2}{\sigma_1}E\big[(X-\mu_1)(Y-\mu_2)\big]+\rho^2\frac{\sigma_2^2}{\sigma_1^2}E\big[(X-\mu_1)^2\big]$$

$$= \sigma_2^2-2\rho\frac{\sigma_2}{\sigma_1}\rho\sigma_1\sigma_2+\rho^2\frac{\sigma_2^2}{\sigma_1^2}\sigma_1^2$$

$$= \sigma_2^2-2\rho^2\sigma_2^2+\rho^2\sigma_2^2=\sigma_2^2(1-\rho^2)$$

The result is non-negative

■

The correlation coefficient (cont.)

Let the variance $\text{Var}(Y|x)$ be denoted by $k(x)$

Then,

$$E\big[k(x)\big] = \sigma_2^2(1-\rho^2) \geq 0$$

Accordingly, $\rho^2 \leq 1$ or $-1 \leq \rho \leq +1$

Remark

$-1 \leq \rho \leq +1$, whether the conditional mean is linear or is not ($\star$)

The correlation coefficient (cont.)

Suppose that the variance $\text{Var}(Y|x)$ is positive, but not a function of x

$\rightsquigarrow$ The variance is a constant $k > 0$

Let k be multiplied by $f_1(x)$ and integrated on x

- The result is k, so that $k = \sigma_2^2(1 - \rho^2)$

In this case, the variance of each conditional distribution of Y, given $X = x$

$$\rightsquigarrow \quad \sigma_2^2(1 - \rho^2)$$

If $\rho = 0$, the variance of each conditional distribution of Y, given $X = x$

$$\rightsquigarrow \quad \sigma_2^2$$

This is also the variance of the marginal distribution of Y

The correlation coefficient (cont.)

Suppose $\rho^2 \simeq 1$

The variance of each conditional distribution of Y, given $X = x$, is small

- The concentration of probability for this conditional distribution near the mean $E(Y|x) = \mu_2 + \rho(\sigma_2/\sigma_1)(x - \mu_1)$

The correlation coefficient (cont.)

Similar comments can be made about $E(X|y)$, if it is linear

Specifically,

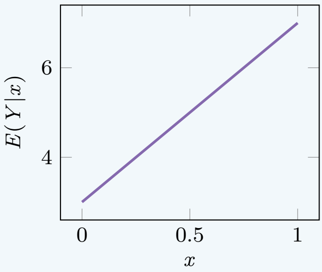
- $E(X|y) = \mu_1 + \rho(\sigma_1/\sigma_2)(y - \mu_2)$
- $E\big[\mathrm{Var}(X|Y)\big] = \sigma_1^2(1 - \rho^2)$

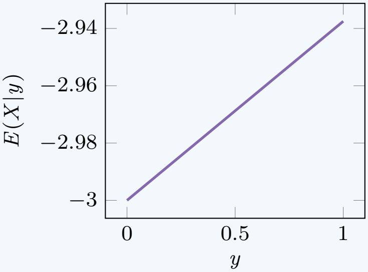
The correlation coefficient (cont.)

Example

Let the random variables X and Y have the linear conditional means

- $E(Y|x) = 4x + 3$
- $E(X|y) = 1/16y - 3$





The correlation coefficient (cont.)

From the general expression for linear conditional means

- ⇝ $E(Y|x) = \mu_2$, if $x = \mu_1$
- ⇝ $E(X|y) = \mu_1$, if $y = \mu_2$

Thus,

- $\mu_2 = 4\mu_1 + 3 = -15/4$
- $\mu_1 = 1/16\mu_2 - 3 = -12$

The correlation coefficient (cont.)

Still from the general expression for linear conditional means

- ⇝ The product of the coefficients of x and y equals ρ^2
- ⇝ The ratio of the coefficients of x and y equals σ_2^2/σ_1^2

We have,

- ⇝ $\rho^2 = 4(1/16) = 1/4$, with $\rho = 1/2$ (not $-1/2$)
- ⇝ $\sigma_1^2 \sigma_2^2 = 64$

From the two linear conditional means, we find values for μ_1, μ_2, ρ and σ_2/σ_1

- Though not the values of σ_1 and σ_1 alone

■

The correlation coefficient (cont.)

Consider the MGF for the random vector (X, Y)

- The joint MGF gives explicit formulas for some moments

For random variables of the continuous type

$$\rightsquigarrow \quad \frac{\partial^{k+m} M(t_1, t_2)}{\partial t_1^k \partial t_2^m} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^k y^m e^{(t_1 x + t_2 y)} f(x, y) \mathrm{d}x \mathrm{d}y$$

So,

$$\frac{\partial^{k+m} M(t_1, t_2)}{\partial t_1^k \partial t_2^m}\Big|_{t_1 = t_2 = 0} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^k y^m f(x, y) \mathrm{d}x \mathrm{d}y = E(X^k, Y^m)$$

The correlation coefficient (cont.)

In a simplified notation, we have

$$\begin{aligned}
\mu_1 &= E(X) = \frac{\partial M(0,0)}{\partial t_1} \\
\mu_2 &= E(Y) = \frac{\partial M(0,0)}{\partial t_2} \\
\sigma_1^2 &= E(X^2) - \mu_1^2 = \frac{\partial^2 M(0,0)}{\partial t_1^2} - \mu_1^2 \\
\sigma_2^2 &= E(Y^2) - \mu_2^2 = \frac{\partial^2 M(0,0)}{\partial t_2^2} - \mu_2^2 \\
E\big[(X-\mu_1)(Y-\mu_2)\big] &= \frac{\partial^2 M(0,0)}{\partial t_1 \partial t_2} = \mu_1 \mu_2
\end{aligned} \tag{9}$$

From these we can compute the correlation coefficient ρ

The correlation coefficient (cont.)

The results hold if X and Y are RV of the discrete type

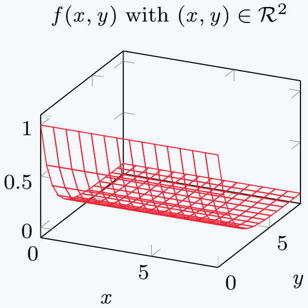
Correlation coefficients may be computed from the MGF of the joint

$\leadsto$ If that function is available

The correlation coefficient (cont.)

Example

Let the two random variables X and Y have the joint density



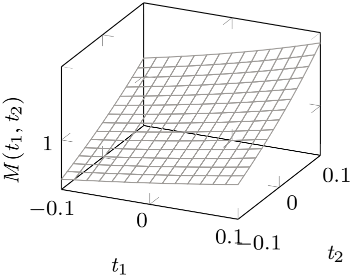
$$f(x,y) = \begin{cases} e^{-y}, & 0 < x < y < \infty \\ 0, & \text{elsewhere} \end{cases}$$

The correlation coefficient (cont.)

We have already determined the MGF

$$M(t_1, t_2) = \frac{1}{(1 - t_1 - t_2)(1 - t_2)}$$

for $t_1 + t_2 < 1$ and $t_2 < 1$



The correlation coefficient (cont.)

For this distribution, Equation (9) becomes (⋆)

$$\begin{aligned} \mu_1 &= 1 \\ \mu_2 &= 2 \\ \sigma_1^2 &= 1 \\ \sigma_2^2 &= 2 \\ E\big[(X-\mu)(Y-\mu_2)\big] &= 1 \end{aligned} \tag{10}$$

The correlation coefficient of X and Y is $\rho = 1/\sqrt{2}$

■

Independence

Two random variables

Independence

Let X_1 and X_2 be two random variables of the continuous type

- ⇝ The joint probability density function is indicated by $f(x_1, x_2)$
- ⇝ The marginal probability density functions are $f_1(x_1)$ and $f_2(x_2)$

Let the definition of conditional PDF be indicated by $f_{2|1}(x_2|x_1)$

The joint PDF $f(x_1, x_2)$

$$f(x_1, x_2) = f_{2|1}(x_2|x_1) f_1(x_1)$$

Independence (cont.)

Suppose that we have an instance where $f_{2|1}(x_2|x_1)$ does not depend on x_1

The marginal PDF of x_2 for random variables of the continuous type

$$f_2(x_2) = \int_{-\infty}^{\infty} f_{2|1}(x_2|x_1) f_1(x_1) \mathrm{d}x_1 = f_{2|1}(x_2|x_1) \int_{-\infty}^{\infty} f_1(x_1) \mathrm{d}x_1$$
$$= f_{2|1}(x_2|x_1)$$

Accordingly, when $f_{2|1}(x_2|x_1)$ does not depend upon x_1

$$f_2(x_2) = f_{2|1}(x_2|x_1)$$

and

$$f(x_1, x_2) = f_1(x_1) f_2(x_2)$$

Independence (cont.)

$$f(x_1, x_2) = f_{2|1}(x_2|x_1)f_1(x_1)$$

Let the conditional distribution of X_2, given $X_1 = x_1$, be independent of x_1

$\rightsquigarrow$ $f(x_1, x_2) = f_{2|1}(x_2|\cancel{x_1})f_1(x_1) = f_2(x_2)f_1(x_1)$

Independence (cont.)

Definition

Independence

Let the RVs X_1 and X_2 have joint PDF $f(x_1, x_2)$ [joint PMF $p(x_1, x_2)$]

Let the marginal PDFs [PMFs] be $f_1(x_1)$ [$p_1(x_1)$] and $f_2(x_2)$ [$p_2(x_2)$]

*The random variables X_1 and X_2 are said to be **independent** iff*

$$f(x_1, x_2) \equiv f_1(x_1) f_2(x_2) \qquad [p(x_1, x_2) \equiv p_1(x_1) p_2(x_2)]$$

*RVs that are not independent are said to be **dependent***

Independence (cont.)

Remark

On the product of two positive functions

The product $f_1(x_1)f_2(x_2)$ of two positive functions $f_1(x_1)$ and $f_2(x_2)$

- It refers to a function that is positive on the product space

Let $f_1(x_1)$ and $f_2(x_2)$ be two functions

Suppose that they are positive on, and only on, their spaces $\mathcal{S}_1$ and $\mathcal{S}_2$

Product of $f_1(x_1)$ and $f_2(x_2)$ is positive on and only on the product space

$$\rightsquigarrow \quad \mathcal{S} = \big\{(x_1, x_2) : x_1 \in S_1, x_2 \in S_2\big\}$$

Independence (cont.)

Example

Let $S_1 = \{x_1 : 0 < x_1 < 1\}$

Let $S_2 = \{x_2 : 0 < x_2 < 3\}$

Then,

$$\rightsquigarrow \quad \mathcal{S} = \{(x_1, x_2) : 0 < x_1 < 1, 0 < x_2 < 3\}$$

Independence (cont.)

Remark

On the identity

There may be certain points $(x_1, x_2) \in \mathcal{S}$ at which $f(x_1, x_2) \neq f_1(x_1) f_2(x_2)$

Let A be the set of points (x_1, x_2) at which the equality does not hold

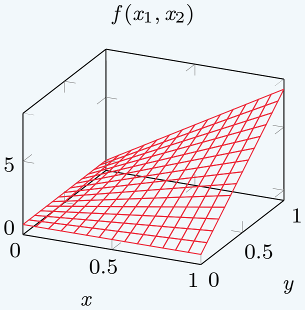
- $\rightsquigarrow$ In A, we have (interpret) that $P(A) = 0$

Products of non-negative functions and identities will be interpreted alike

Independence (cont.)

Example

Let X_1 and X_2 be random variables with the PDF



$$f(x_1, x_2) = \begin{cases} 8x_1x_2, & 0 < x_1 < x_2 < 1 \\ 0, & \text{elsewhere} \end{cases}$$

Expression $(8x_1x_2)$ may suggest that X_1 and X_2 are independent

Yet, $\mathcal{S} = \{(x_1, x_2) : 0 < x_1 < x_2 < 1\}$ is not a product space

Independence (cont.)

In general

X_1 and X_2 are dependent if the space of positive probability density of X_1 and X_2 is bounded by a curve that is neither a horizontal nor a vertical line

■

Independence (cont.)

Independence can be presented also in terms of cumulative distributions CDFs

- No need for reasoning with PDFs/PMFs

Theorem 3.1

Let (X_1, X_2) have the joint CDF $F(x_1, x_2)$

Let X_1 and X_2 have the marginal CDFs $F_1(x_1)$ and $F_2(x_2)$

Then, X_1 and X_2 are independent if and only if

$$F(x_1, x_2) = F_1(x_1) F_2(x_2), \quad \text{for all } (x_1, x_2) \in \mathcal{R}^2 \tag{11}$$

Independence (cont.)

Proof

For the continuous case

Suppose $F(x_1, x_2) = F_1(x_1)F_2(x_2)$ for all $(x_1, x_2) \in \mathcal{R}^2$ holds true

The mixed second partial derivative of the joint CDF

$$\frac{\partial^2 F(x_1, x_2)}{\partial x_1 \partial x_2} = f_1(x_1)f_2(x_2)$$

Hence, X_1 and X_2 are independent

Independence (cont.)

Suppose X_1 and X_2 are independent, $f(x_1, x_2) = f_1(x_1) f_2(x_2), \forall (x_1, x_2) \in \mathcal{R}^2$

By definition of joint CDF

$$\begin{aligned} F(x_1, x_2) &= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f_1(w_1) f_2(w_2) \mathrm{d}w_2 \mathrm{d}w_1 \\ &= \int_{-\infty}^{x_1} f_1(w_1) \mathrm{d}w_1 \cdot \int_{-\infty}^{x_2} f_2(w_2) \mathrm{d}w_2 \\ &= F_1(x_1) F_2(x_2) \end{aligned}$$

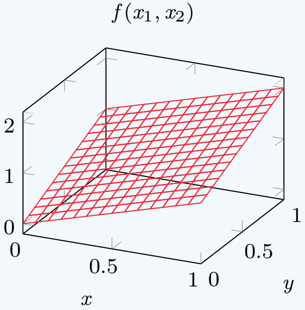
Hence, condition $F(x_1, x_2) = F_1(x_1) F_2(x_2)$ for all $(x_1, x_2) \in \mathcal{R}^2$ holds true

■

Independence (cont.)

Example

Let X_1 and X_2 be two random variables with the PDF



$$f(x_1, x_2) = \begin{cases} x_1 + x_2, & 0 < x_1, x_2 < 1 \\ 0, & \text{elsewhere} \end{cases}$$

We show that X_1 and X_2 are dependent

Independence (cont.)

The marginal densities

$$f_1(x_1) = \begin{cases} \int_{-\infty}^{\infty} f(x_1, x_2) \mathrm{d}x_2 = \int_0^1 (x_1 + x_2) \mathrm{d}x_2 = x_1 + 1/2, & 0 < x_1 < 1 \\ 0, & \text{elsewhere} \end{cases}$$

$$f_2(x_2) = \begin{cases} \int_{-\infty}^{\infty} f(x_1, x_2) \mathrm{d}x_2 = \int_0^1 (x_1 + x_2) \mathrm{d}x_2 = 1/2 + x_2, & 0 < x_2 < 1 \\ 0, & \text{elsewhere} \end{cases}$$

Since $f(x_1, x_2) \not\equiv f_1(x_1) f_2(x_2)$, random variables X_1 and X_2 are dependent

■

Independence (cont.)

We can assert that X_1 and X_2 are dependent, without the marginal PDFs

Theorem

Let the random variables X_1 and X_2 have supports $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively

Let the joint PDF be indicated by $f(x_1, x_2)$

Then, X_1 and X_2 are independent if and only if $f(x_1, x_2)$ can be written as product of a non-negative function of x_1 and a non-negative function of x_2

$$f(x_1, x_2) \equiv g(x_1)h(x_2)$$

- $g(x_1) > 0$ *for* $x_1 \in \mathcal{S}_1$ *and zero elsewhere*
- $h(x_2) > 0$ *for* $x_2 \in \mathcal{S}_2$ *and zero elsewhere*

Independence (cont.)

Proof

If X_1 and X_2 are independent, then $f(x_1, x_2) \equiv f_1(x_1) f_2(x_2)$

- $f_1(x_1)$ and $f_2(x_2)$, the marginal PDFs of X_1 and X_2

Thus, condition $f(x_1, x_2) \equiv g(x_1) h(x_2)$ is satisfied

Conversely, let $f(x_1, x_2) \equiv g(x_1) h(x_2)$

Then, for RVs of the continuous type

$$f_1(x_1) = \int_{-\infty}^{\infty} \underbrace{g(x_1) h(x_2)}_{f(x_1, x_2)} \mathrm{d}x_2 = g(x_1) \int_{-\infty}^{\infty} h(x_2) \mathrm{d}x_2 = c_1 g(x_1)$$

$$f_2(x_2) = \int_{-\infty}^{\infty} \underbrace{g(x_1) h(x_2)}_{f(x_1, x_2)} \mathrm{d}x_1 = h(x_2) \int_{-\infty}^{\infty} g(x_1) \mathrm{d}x_1 = c_2 g(x_2)$$

c_1 and c_2 are constants (not functions of x_1 and x_2)

Independence (cont.)

Moreover, I also know that $c_1 c_2 = 1$

$$1 = \int_{-\infty}^{\infty} \int_{\infty}^{\infty} \underbrace{g(x_1)h(x_2)}_{f(x_1,x_2)} \mathrm{d}x_1 \mathrm{d}x_2$$

$$= \Big[\int_{-\infty}^{\infty} g(x_1)\mathrm{d}x_1 \Big] \Big[\int_{-\infty}^{\infty} h(x_2)\mathrm{d}x_2 \Big]$$

$$= c_2 c_1$$

These results imply

$$f(x_1, x_2) \equiv g(x_1)h(x_2) \equiv c_1 g(x_1) c_2 h(x_2) \equiv f_1(x_1) f_2(x_2)$$

Accordingly, X_1 and X_2 are independent

■

Independence (cont.)

The theorem is true for the discrete case, too

- ⇝ Replace the joint PDF by the joint PMF

Independence (cont.)

Theorem

Let X_1 and X_2 be two random variables

X_1 and X_2 are independent RVs if and only if the following condition holds

$$P(a < X_1 \leq b, c < X_2 \leq d) = P(a < X_1 \leq b)P(c < X_2 \leq d) \qquad (12)$$

for every $a < b$ and $c < d$, where a, b, c and d are constants

Proof

If X_1 and X_2 are independent, then by the previous theorem

$$\begin{aligned} P(a < X_1 \leq b, c < X_2 \leq d) &= F(b,d) - F(a,d) - F(b,c) + F(a,c) \\ &= F_1(b)F_2(d) - F_1(a)F_2(d) - F_1(b)F_2(c) + F_1(a)F_2(c) \\ &= \big[F_1(b) - F_1(a)\big]\big[F_2(d) - F_2(c)\big] = P(a < X_1 \leq b)P(c < X_2 \leq d) \end{aligned}$$

The joint CDF of (X_1, X_2) factors into a product of marginal CDFs

⇝ This again implies that X_1 and X_2 are independent

■

Independence (cont.)

Example

Let the RVs X_1 and X_2 have supports $\mathcal{S}_1$ and $\mathcal{S}_2$ and joint PDF $f(x_1, x_2)$

We know that X_1 and X_2 are independent if and only if

$$f(x_1, x_2) \equiv f(x_1)h(x_2)$$

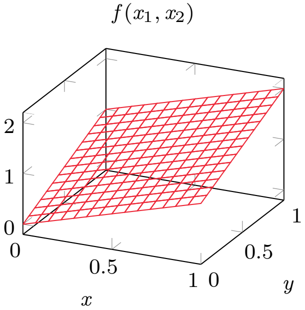
- $g(x_1) > 0$ for $x_1 \in \mathcal{S}_1$ and zero elsewhere
- $h(x_2) > 0$ for $x_2 \in \mathcal{S}_2$ and zero elsewhere

Independence is necessary for condition

$$p\big(a < X_1 \leq b, c < X_2 \leq d\big) = p\big(a < X_1 \leq b\big)p\big(c < X_2 \leq d\big)$$

Independence (cont.)

Let X_1 and X_2 be two random variables with the PDF

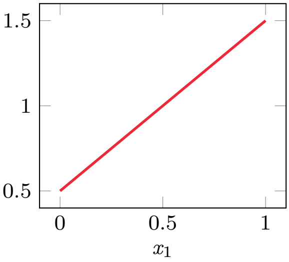


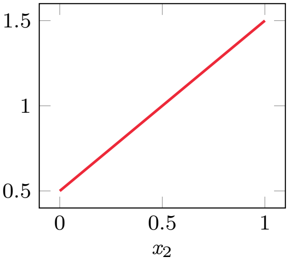
$$f(x_1, x_2) = \begin{cases} x_1 + x_2, & 0 < x_1, x_2 < 1 \\ 0, & \text{elsewhere} \end{cases}$$

$$P(0 < X_1 < 1/2, 0 < X_2 < 1/2) = \int_0^{1/2} \int_0^{1/2} (x_1 + x_2) \mathrm{d}x_1 \mathrm{d}x_2 = 1/8$$

Independence (cont.)

$f_{X_1}(x_1) = x_1 + 1/2, x_1 \in (0,1)$ $\qquad$ $f_{X_2}(x_2) = 1/2 + x_2, x_2 \in (0,1)$





$$P(0 < X_1 < 1/2) = \int_0^{1/2} (x_1 + 1/2)\mathrm{d}x_1 = 3/8$$

$$P(0 < X_2 < 1/2) = \int_0^{1/2} (1/2 + x_2)\mathrm{d}x_2 = 3/8$$

Thus,

$$\underbrace{p\big(a < X_1 \leq b, c < X_2 \leq d\big)}_{1/8} \neq \underbrace{p\big(a < X_1 \leq b\big)}_{3/8}\underbrace{p\big(c < X_2 \leq d\big)}_{3/8}$$

Independence (cont.)

Some probabilities are simpler when we have independent random variables

True also for some expectations (and some moment generating functions)

Independence (cont.)

Theorem

Let X_1 and X_2 be two independent random variables

Suppose that $E\big[u(X_1)\big]$ and $E\big[v(X_2)\big]$ exist

Then,

$$E\big[u(X_1)v(X_2)\big] = E\big[u(X_1)\big]E\big[v(X_2)\big]$$

Independence (cont.)

Proof

Independence of X_1 and X_2 implies factorisation of the joint PDF of (X_1, X_2)

$$f(x_1, x_2) = f_1(x_1) f_2(x_2)$$

Thus, by definition of expectation

$$\begin{aligned} E\big[u(X_1)v(X_2)\big] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \big[u(x_1)v(x_2)\big] \underbrace{f_1(x_1)f_2(x_2)}_{f(x_1,x_2)} \mathrm{d}x_1 \mathrm{d}x_2 \\ &= \Big[\int_{-\infty}^{\infty} u(x_1)f_1(x_1)\mathrm{d}x_1\Big]\Big[\int_{-\infty}^{\infty} v(x_2)f_2(x_2)\Big] \\ &= E\big[u(X_1)\big]E\big[v(X_2)\big] \end{aligned}$$

■

Independence (cont.)

Upon taking function $u(\cdot)$ and $v(\cdot)$ to be the identify functions

$$\rightsquigarrow \quad E(X_1 X_2) = E(X_1)E(X_2) \tag{13}$$

The expectation of the product of two independent variables X_1 and X_2

Independence (cont.)

Example

Let X and Y be two independent random variables

- Let μ_1 and μ_2 indicate the respective means
- Let σ_1^2 and σ_2^2 indicate the respective (positive) variances

We show that independence of X and Y implies correlation coefficient zero

The covariance of X and Y

$$E\big[(X-\mu_1)(Y-\mu_2)\big] = E(X-\mu_1)E(Y-\mu_2) = E(XY)-\mu_1\mu_2 = 0$$

■

Independence (cont.)

Theorem

Suppose the joint MGF $M(t_1, t_2)$ exists for random variables X_1 and X_2

Then, X_1 and X_2 are independent if and only if

$$M(t_1, t_2) = M(t_1, 0)M(0, t_2)$$

The joint MGF is identically equal to the product of the marginals MGFs

Independence (cont.)

Proof

Suppose that X_1 and X_2 are independent

Then,

$$\begin{aligned} M(t_1, t_2) &= E(e^{t_1 X_1 + t_2 X_2}) = E(e^{t_1 X_1} e^{t_2 X_2}) = E(e^{t_1 X_1}) E(e^{t_2 X_2}) \\ &= M(t_1, 0) M(t_2, 0) \end{aligned}$$

The independence of X_1 and X_2 implies that the MGF of the joint distribution factors into the product of the MGFs of the two marginal distributions

Independence (cont.)

Consider the MGF of the joint distribution of X_1 and X_2

$$M(t_1, t_2) = M(t_1, 0)M(0, t_2)$$

X_1 has a unique MGF

$$\rightsquigarrow M(t_1, 0) = \int_{-\infty}^{\infty} e^{t_1 x_1} f_1(x_1) \mathrm{d}x_1$$

X_2 has a unique MGF

$$\rightsquigarrow M(0, t_2) = \int_{-\infty}^{\infty} e^{t_2 x_2} f_2(x_2) \mathrm{d}x_2$$

Independence (cont.)

Thus,

$$\begin{aligned} M(t_1, 0)M(0, t_2) &= \Big[\int_{-\infty}^{\infty} e^{t_1 x_1} f_1(x_1) \mathrm{d}x_1 \Big] \Big[\int_{-\infty}^{\infty} e^{t_2 x_2} f_2(x_2) \mathrm{d}x_2 \Big] \\ &= \int_{-\infty}^{\infty} \int_{\infty}^{\infty} e^{t_1 x_1 + t_2 x_2} f_1(x_1) f_2(x_2) \mathrm{d}x_1 \mathrm{d}x_2 \end{aligned}$$

Independence (cont.)

We know that $M(t_1, t_2) = M(t_1, 0)M(0, t_2)$

$$\rightsquigarrow M(t_1, t_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{t_1 x_1 + t_2 x_2} f_1(x_1) f_2(x_2) \mathrm{d}x_1 \mathrm{d}x_2$$

We know that $M(t_1, t_2)$ is the MGF of X_1 and X_2

$$\rightsquigarrow M(t_1, t_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{t_1 x_1 + t_2 x_2} f(x_1, x_2) \mathrm{d}x_1 \mathrm{d}x_2$$

Independence (cont.)

The two distributions described by $f_1(x_1)f_2(x_2)$ and $f(x_1, x_2)$ are the same

- By the uniqueness of the MGF

$$\rightsquigarrow \quad f(x_1, x_2) \equiv f_1(x_1)f_2(x_2)$$

That is, if $M(t_1, t_2) = M(t_1, 0)M(0, t_2)$, then X_1 and X_2 are independent

Independence (cont.)

The proof relies on the assertion that a MGF (when it exists) is unique

- The MGF uniquely defines the distribution of probability

The proof for discrete-type RV uses summation instead of integration

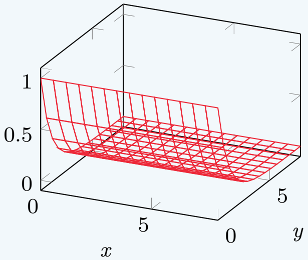
■

Independence (cont.)

Example

Let (X, Y) be a pair of random variables with the joint PDF

$f(x, y)$ with $(x, y) \in \mathcal{R}^2$



$$f(x, y) = \begin{cases} e^{-y}, & 0 < x < y < \infty \\ 0, & \text{elsewhere} \end{cases}$$

Independence (cont.)

We determined the that the MGF of (X, Y)

$$M(t_1, t_2) = \int_0^\infty \int_x^\infty \exp(t_1 x + t_2 x - y) \mathrm{d}y \mathrm{d}x$$
$$= \frac{1}{(1 - t_1 - t_2)(1 - t_2)}$$

(provided that $t_1 + t_2 < 1$ and $t_2 < 1$)

Because $M(t_1, t_2) \neq M(t_1, 0) M(0, t_2)$, the random variables are dependent

■