Two discrete random variables

Anis Rezgui
Mathematics & Computer Sciences Department
Carthage University - INSAT

October 18, 2019

anis.rezguii@gmail.com

Table of contents

anis.rezguii@gmail.com

Introduction

After careful study of this chapter you should be able to:

1. recognize mass distribution functions of the sum of usual random variables.
2. understand the concept of joint mass distribution function of two discrete random variables.
3. use the joint mass distribution function to compute probabilities.
4. extract marginal and conditional mass distribution functions from the joint mass distribution.
5. determine the mass distribution function of a function of two discrete random variables.
6. understand the total expectation theorem.
7. understand the total variance theorem.

anis.rezguii@gmail.com

Motivation

In many practical situations we need to observe **several** phenomenon or parameters **simultaneously** which leads to many measures at the same time.

So it is natural to consider random vector $(X_1, \cdots, X_n)$ or easier a pair of random variables (X_1, X_2), the latter pair can be either discrete, continuous or mixed. In this chapter we focus on a pair of discrete random variables.

Independence of random variables

Two discrete random variables may be independent or dependent.

Let X and Y be two discrete random variables. They are said to be independent if for any value x_i of X and y_j of Y:

$$\mathbb{P}\{X = x_i\,,\ Y = y_j\} = \mathbb{P}\{X = x_i\} \times \mathbb{P}\{Y = y_j\}.$$

anis.rezguii@gmail.com

Joint probability distribution

Definition

Let (X, Y) be a pair of discrete random variables with $\{x_i \, : \, i = 1, \cdots, p\}$ and $\{y_j \, : \, j = 1, \cdots, q\}$ as values sets.
The joint probability distribution of the random pair (X, Y) is given by the matrix

$$\mathbb{P}_{(X,Y)} = \begin{bmatrix} p_{11} & \cdots & p_{1q} \\ \vdots & \vdots & \vdots \\ p_{p1} & \cdots & p_{pq} \end{bmatrix}$$

where $p_{ij} = \mathbb{P}\{X = x_i \, , \, Y = y_j\}$.

anis.rezguii@gmail.com

Notes

1. The joint probability distribution matrix is also said the joint probability mass function or the contingency table of the two random variables X and Y.

2. A matrix (p_{ij}) is a the joint probability distribution of a given random pair X, Y if and only if it satisfies:
   - $p_{ij} \geq 0$
   - $\sum_{1 \leq i \leq p} \sum_{1 \leq j \leq q} p_{ij} = 1.$

3. If X and Y are independent, we have

$$p_{ij} = \mathbb{P}\{X = x_i\} \times \mathbb{P}\{Y = y_j\}.$$

anis.rezguii@gmail.com

Joint probability distribution

Key rules

Proposition

Let (X, Y) be a discrete random pair with values set $\{x_1, \cdots, x_p\} \times \{y_1, \cdots, y_q\}$.

1. If $C \subset \{x_1, \cdots, x_p\} \times \{y_1, \cdots, y_q\}$ the probability

$$\mathbb{P}\{(X, Y) \in C\} = \sum_{(x,y) \in C} \mathbb{P}\{X = x \text{ and } Y = y\}.$$

2. Let $Z = g(X, Y)$ where g is a given function defined on $\{x_1, \cdots, x_p\} \times \{y_1, \cdots, y_q\}$, then its probability mass function is defined as follows:

$$\mathbb{P}\{Z = z\} = \sum_{\{(x,y):g(x,y)=z\}} \mathbb{P}\{X = x \text{ and } Y = y\}.$$

Example 2

Consider the matrix, (P), below:

$$\begin{bmatrix} 2\% & 5\% & 10\% \\ 5\% & 10\% & 10\% \\ 30\% & 14\% & 14\% \end{bmatrix}$$

1. Check that (P) is a joint probability distribution matrix.
2. Suppose now that (P) equals $\mathbb{P}_{(X,Y)}$ where $X = \{-1, 0, 1\}$ and $Y = \{1, 2, 3\}$
    1. Compute the probability $\mathbb{P}\Big\{X \in \{-1, 0\} \text{ and } Y \in \{1, 2\}\Big\}$.
    2. Let $Z = X + Y$ determine its probability mass function $\mathbb{P}_Z$.
    3. Compute both expectations of X and Y.

anis.rezguii@gmail.com

Example 2

Consider two successive number draws, defined as follows: The first draw is done in an equiprobable manner from the set $\{1, 2, 3, 4\}$,
The second one is done in an equiprobable manner from the set $\{1, \cdots, k\}$, where k is the result of the first draw. This two draws define the random pair (X, Y).

1. Determine the joint probability distribution of (X, Y).
2. Compute the probability: $\mathbb{P}\Big\{X \in \{1, 2, 3\} \text{ and } Y \in \{1, 2\}\Big\}$.
3. Let $Z = X + Y$ determine its probability mass function $\mathbb{P}_Z$.

anis.rezguii@gmail.com

Marginal probability distribution

Definition

The marginal probability distribution of X (resp. of Y) is simply the probability distribution of $\mathbb{P}_X$ (resp. of $\mathbb{P}_Y$), when determined from the joint probability distribution of the random pair (X, Y). In this case $\mathbb{P}_X$ expresses as follows:

$$\mathbb{P}\{X = x_i\} = \sum_{j=1}^{q} p_{ij}. \quad (\text{resp. } \mathbb{P}\{Y = y_j\} = \sum_{i=1}^{p} p_{ij}.)$$

anis.rezguii@gmail.com

Marginal probability distribution

Note

1. Note that the given of the joint probability distribution contains more information than the only knowledge of the two marginal probability distributions. In fact the joint distribution gives information about the simultaneous random behavior of both X and Y, however the only knowledge of the probability distributions of X and Y gives information about the random behavior of each independently as the other one doesn't exist.
2. The only knowledge of the marginal probability distributions does never determine the joint probability distribution.

anis.rezguii@gmail.com

Marginal probability distribution

Example 3

Reconsider the last two examples and deduce the two marginal probability distributions for each.

Conditional probability distribution

Definition

Let (X, Y) be a random pair. The conditional probability distribution of Y such that $X = x_i$ describes the random behavior of Y knowing that $X = x_i$, is denoted by $\mathbb{P}_{Y|X=x_i}$, and is defined as follows:

$$\begin{aligned}\mathbb{P}_{Y|X=x_i}(y_j) &= \mathbb{P}\{Y = y_j|X = x_i\} \\ &= \frac{\mathbb{P}\{Y = y_j, X = x_i\}}{\mathbb{P}\{X = x_i\}} \\ &= \frac{p_{ij}}{\sum_{l=1}^{q} p_{il}}\end{aligned}$$

anis.rezguii@gmail.com

Conditional probability distribution

Example 4

Reconsider the example 1.

1. Determine the the conditional probability distribution of Y such that $X = 0$, $\mathbb{P}_{Y|X=0}$.
2. Determine the the conditional probability distribution of X such that $Y = 2$, $\mathbb{P}_{X|Y=2}$.

anis.rezguii@gmail.com

Conditional probability distribution

Example 5

Reconsider the example 2 and Find out the conditional probability distribution of $Y|X = k$ for $k = 1, \cdots, 4$.

anis.rezguii@gmail.com

Note

It very important to note that the knowledge of the marginal probability distributions as well as the knowledge of the conditional probability distributions determine the joint probability distribution. In fact

$$\begin{aligned}\mathbb{P}\{X = x_i\} \times \mathbb{P}\{Y = y_j | X = x_i\} &= \mathbb{P}\{X = x_i\,,\ Y = y_j\} \\ &= p_{ij}.\end{aligned}$$

anis.rezguii@gmail.com

The conditional expectation $\mathbb{E}(Y|X=x_i)$

Definition

The conditional expectation of Y such that $X = x_i$ is the expectation of Y with respect to the conditional probability distribution $\mathbb{P}_{Y|X=x_i}$, given by the following:

$$\mathbb{E}(Y|X=x_i) = \sum_{j=1}^{p} y_j \, \mathbb{P}\{Y = y_i|X = x_i\} = \sum_{j=1}^{q} y_j \, \frac{p_{ij}}{\sum_{j=1}^{q} \, p_{ij}}$$

The conditional expectation $\mathbb{E}(Y|X)$

Definition

For $x \in \{x_1, \cdots, x_p\}$ we define the function $\varphi(x) = \mathbb{E}(Y|X = x)$. And so, we define the random variable conditional expectation by

$$\mathbb{E}(Y|X) = \varphi(X).$$

anis.rezguii@gmail.com

Notes

1. The function $\varphi(x) = \mathbb{E}(Y|X = x)$ is always implicit but it visualizes the part of Y that depends on X as a function of this latter.
2. It is easy to show that $\mathbb{E}(Y_1 + Y_2|X) = \mathbb{E}(Y_1|X) + \mathbb{E}(Y_2|X)$.

Example 6

Reconsider the last example 1 and compute the random variable conditional expectation $\mathbb{E}(Y|X)$.

anis.rezguii@gmail.com

Example 7

Reconsider the last example 2 and compute the random variable conditional expectation $\mathbb{E}(Y|X)$.

anis.rezguii@gmail.com

The Total Expectation Theorem

Theorem

For any random pair (X, Y) we have

$$\mathbb{E}\Big(\mathbb{E}(Y|X)\Big) = \mathbb{E}(Y).$$

anis.rezguii@gmail.com

The conditional expectation

Proof

$$\begin{aligned}
\mathbb{E}\Big(\mathbb{E}(Y|X)\Big) &= \sum_i \mathbb{E}(Y|X = x_i) \times \mathbb{P}\{X = x_i\} \\
&= \sum_i \left\{ \sum_j y_j \mathbb{P}\{Y = y_j|X = x_i\} \right\} \times \mathbb{P}\{X = x_i\} \\
&= \sum_j y_j \left\{ \sum_i \mathbb{P}\{Y = y_j|X = x_i\} \times \mathbb{P}\{X = x_i\} \right\} \\
&= \sum_j y_j \mathbb{P}\{Y = y_j\} \ = \ \mathbb{E}(Y)
\end{aligned}$$

anis.rezguii@gmail.com

Note

For any function h defined on the values set of X we have:

$$\mathbb{E}(h(X)Y|X) = h(X)\mathbb{E}(Y|X),$$

in fact the random variable $h(X)$ is considered as a constant when we are conditioning on X.

anis.rezguii@gmail.com

The conditional variance

As in the case of the conditional expectation we start by defining, for any $x \in \{x_i,\ i = 1, \cdots, p\}$, the conditional variance

$$\mathbb{V}(Y|X = x_i) = \mathbb{E}\Big((Y - \mathbb{E}(Y|X = x_i))^2|X = x_i\Big).$$

So that we realize the function

$$\{x_1, \cdots, x_p\} \ni x \longmapsto \psi(x) = \mathbb{V}(Y|X = x)$$

and we define the random variable conditional variance $\mathbb{V}(Y|X)$ by

$$\mathbb{V}(Y|X) = \psi(X).$$

anis.rezguii@gmail.com

The conditional variance

Note

It is easy to check that

$$\begin{aligned}
\mathbb{V}(Y|X = x_i) &= \mathbb{E}\Big(\big(Y - \mathbb{E}(Y|X = x_i)\big)^2 |X = x_i\Big) \\
&= \mathbb{E}\Big(Y^2 - 2Y\mathbb{E}(Y|X = x_i) + \mathbb{E}(Y|X = x_i)^2 |X = x_i\Big) \\
&= \mathbb{E}(Y^2|X = x_i) - \mathbb{E}(Y|X = x_i)^2
\end{aligned}$$

and so

$$\begin{aligned}
\mathbb{V}(Y|X) &= \mathbb{E}\Big(\big(Y - \mathbb{E}(Y|X)\big)^2\Big) \\
&= \mathbb{E}(Y^2|X) - \mathbb{E}(Y|X)^2.
\end{aligned}$$

anis.rezguii@gmail.com

The Total Variance

Theorem

Let (X, Y) be a random pair then

$$\mathbb{V}(Y) = \mathbb{V}\Big(\mathbb{E}(Y|X)\Big) + \mathbb{E}\Big(\mathbb{V}(Y|X)\Big).$$

anis.rezguii@gmail.com

The conditional variance

Proof

$$\begin{aligned}
\mathbb{V}(Y) &= \mathbb{E}\Big((Y - \mathbb{E}(Y))^2\Big) \\
&= \mathbb{E}\Big((Y - \mathbb{E}(Y|X) + \mathbb{E}(Y|X) - \mathbb{E}(Y))^2\Big) \\
&= \mathbb{E}\Big((Y - \mathbb{E}(Y|X))^2\Big) + \mathbb{E}\Big((\mathbb{E}(Y|X) - \mathbb{E}(Y))^2\Big) + \\
&\quad 2\mathbb{E}\Big((Y - \mathbb{E}(Y|X))(\mathbb{E}(Y|X) - \mathbb{E}(Y))\Big) \\
&= \mathbb{E}\Big(\mathbb{V}(Y|X)\Big) + \mathbb{V}\Big(\mathbb{E}(Y|X)\Big).
\end{aligned}$$

anis.rezguii@gmail.com

Note

The total variance theorem enables to decompose the variance of Y into two parts:

$$\mathbb{V}(Y) = \mathbb{V}\Big(\mathbb{E}(Y|X)\Big) + \mathbb{E}\Big(\mathbb{V}(Y|X)\Big)$$

Total Variance = Explained Variance + Residual Variance.

anis.rezguii@gmail.com