# INTRODUCTION TO MEASURES OF INEQUALITY AND CONCENTRATION IN TOURISM

Antonio Fernández Morales

University of Málaga, 2016

## INTRODUCTION

There are many statistical measures available in the literature for the analysis of inequality and concentration.

The most common measure in this field is the Gini index (Roselló Nadal *et al.,* 2004), which is associated to the Lorenz curve. Both were developed at the beginning of the 20th century: the Lorenz curve in 1905 and the Gini index in 1914 (Giorgi, 1993).

A revision of recent applications of these measures in the field of tourism can be found in Cisneros-Martín and Fernández-Morales (2015) and Duro (2016).

## APPLICATIONS

Inequality measures have many different applications in the field of tourism, either specific or derived from applications in the field of Economics. Some examples:

**INCOME DISTRIBUTION INEQUALITY**

It was the first application of inequality measures, and still the most important.

**WAGES DISTRIBUTION INEQUALITY**

Closely related with the previous one.

**TOURIST DEMAND CONCENTRATION UN CERTAIN MONTHS**

Gini index can measure the degree of annual concentration of tourist demand.

**TERRITRIAL CONCENTRATION OF TOURISM DEMAND/SUPPLY**

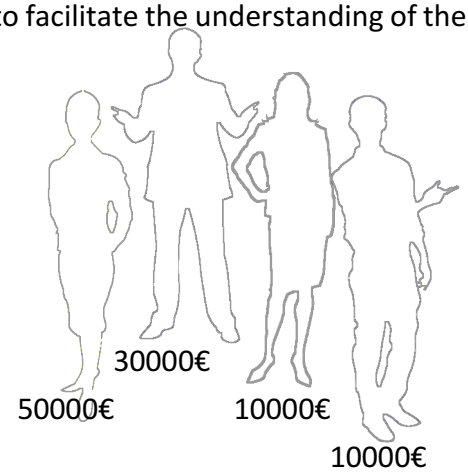
The level of concentration in territorial areas can be used to measure saturation levels.

Thus, the Gini index, and the associated Lorenz curve, is a useful measure with different applications in tourism research, to get a better understanding of the distribution of wages (Marcouiller and Xia, 2008), the seasonal concentration of the demand (Cisneros-Martínez and Fernández-Morales, 2015; Fernández-Morales and Cisneros-Martínez, 2015) or the supply (Lopez Bonilla and López Bonilla, 2006), of the seasonality of air passengers traffic (Cisneros-Martínez and Fernández-Morales, 2016), the seasonality of cruise tourism (Fernández-Morales and Martín-Carrasco, Y., 2014), etc.

## CASE STUDY 1

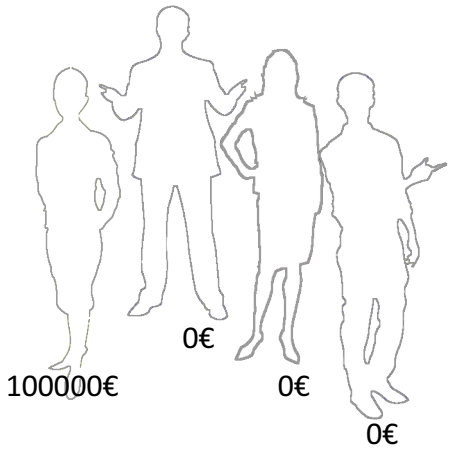
The first case study consists of a very simple situation, in order to facilitate the understanding of the Gini index and the Lorenz curve.

There are four workers in a small travel agency with annual salaries (see figure):
50000€, 30000€, 10000€ and 10000€

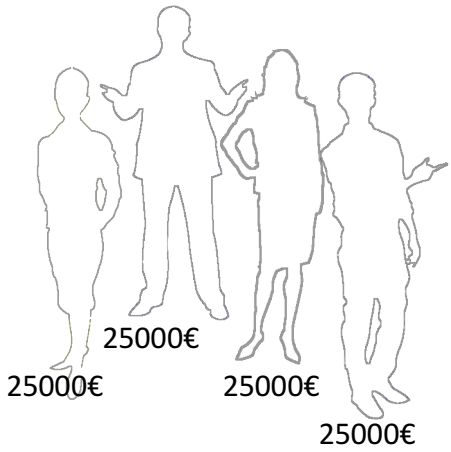



If the salaries in the travel agency change to the distributions shown in the figures below, is there a change in inequality?

☐YES
☐NO




☐YES
☐NO




Now, we analyse some numeric features of the salaries distribution:

| $X_i$ | $n_i$ | $x_i \cdot n_i$ | $p_i$ | $q_i$ | $P_i$ | $Q_i$ |
|---|---|---|---|---|---|---|
| 10000 | 2 | 20000 | 50 | 20 | 50 | 20 |
| 30000 | 1 | 30000 | 25 | 30 | 75 | 50 |
| 50000 | 1 | 50000 | 25 | 50 | 100 | 100 |
| | 4 | 100000 | 100 | 100 | | |

A) What is the total amount of salaries received by the travel agency's workers? _____€
B) What percentage of workers are paid 10000 €? _____%
C) What percentage of the total amount of salaries is earned by the worker who is paid 50000 €? ____%
D) What percentage of the total amount of salaries is earned by the 50% of less-paid workers? _____%

## THE LORENZ CURVE

The Lorenz curve is the graphical representation of the points ($P_i$, $Q_i$), joined by straight segments. (Fernández Morales and Lacomba Arias, 2003)

Where:
$P_i$ : cumulative percentages of frequencies
$Q_i$ : cumulative percentages of the total amount of the variable (cumulative $q_i$)

$$q_i = 100 \frac{x_i n_i}{\sum_{i=1}^{k} x_i n_i}$$

PROPERTIES

- It starts at (0,0)
- It ends at (100,100)
- It is increasing
- The greater the distance between the Lorenz curve and the diagonal starting at (0,0), the greater concentration of the variable
- The closer to the diagonal, the less concentration in the distribution.

EXTREME CASES

- PERFECT EQUALITY (EQUIDISTRIBUTION): The Lorenz curve coincides with the diagonal. This means that there is no inequality in the distribution of the variable, since all the observed values are equal to the arithmetic mean (Gini index = 0)
- MAXIMUM INEQUALITY: The Lorenz curve connects points (0,0), (100 (N-1)/N,0) and (100,100). This means that all the observations but one equal 0, so that the concentration level is maximum. (Gini index approaches 1).

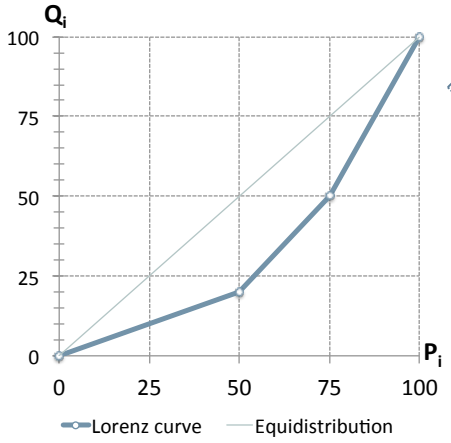
Gini index equals 1 only with infinite observations.

RELATION BETWEEN THE LORENZ CURVE AND THE GINI INDEX

The area between the Lorenz curve and the line of perfect equality divided by the area of the triangle determined by the points (0,0), (100,0), (100,100), corresponds exactly with the Gini index.

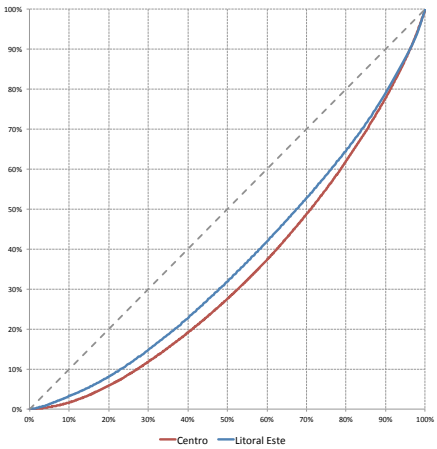
SMOOTHNESS OF THE CURVE

The Lorenz curve polygon becomes smoother as the number of data (or intervals) increases.




4 DATA POINTS
(Case study 1)

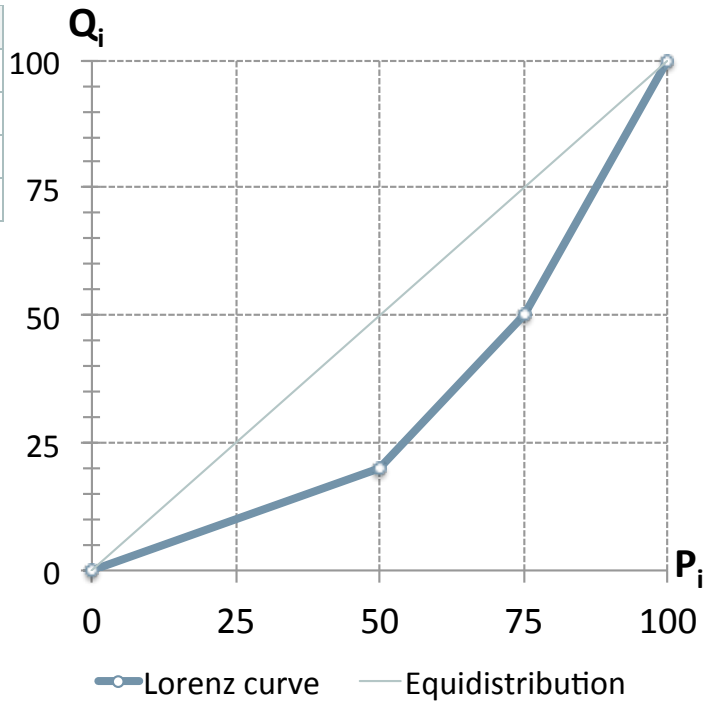
MORE THAN 500 DATA POINTS
Income distribution in two areas of Málaga
Source: Martín Reyes *et al*. (2012)
Gini index: Centro 0.3214,
Litoral Este 0.266



## CASE STUDY 1

LORENZ CURVE

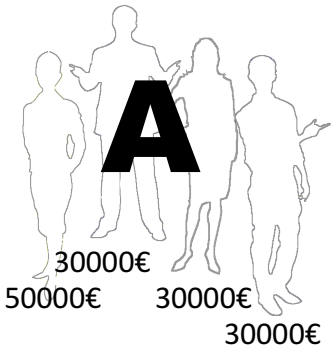
| $X_i$ | $n_i$ | $x_i \cdot n_i$ | $p_i$ | $q_i$ | $P_i$ | $Q_i$ |
|---|---|---|---|---|---|---|
| | | | | | 0 | 0 |
| 10000 | 2 | 20000 | 50 | 20 | 50 | 20 |
| 30000 | 1 | 30000 | 25 | 30 | 75 | 50 |
| 50000 | 1 | 50000 | 25 | 50 | 100 | 100 |
| | 4 | 100000 | 100 | 100 | | |

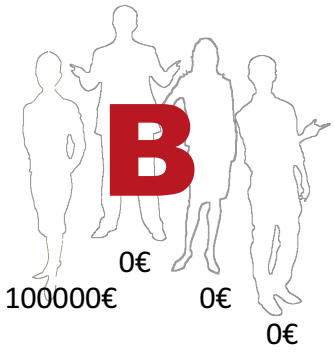


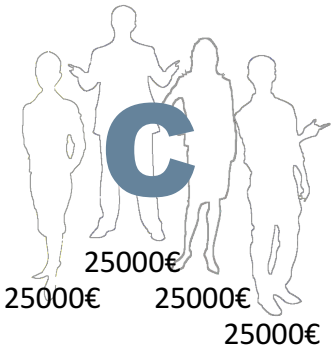

E) What percentage of the total amount of salaries is earned by the 50% of less-paid workers?_____%

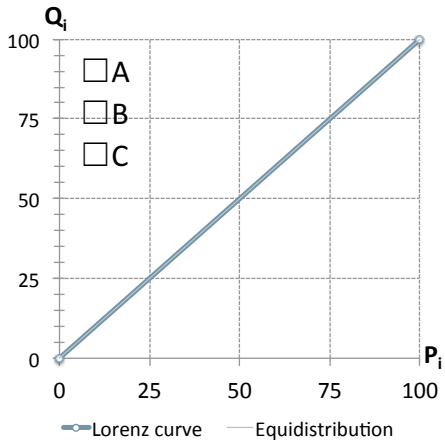
F) What percentage of the total amount of salaries is received by the 25% of highest-paid workers?_____%

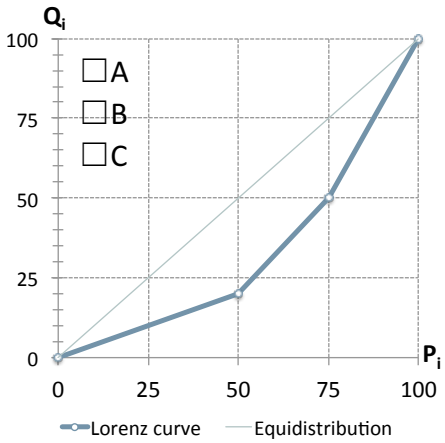
G) Match each distribution of salaries with its Lorenz curve

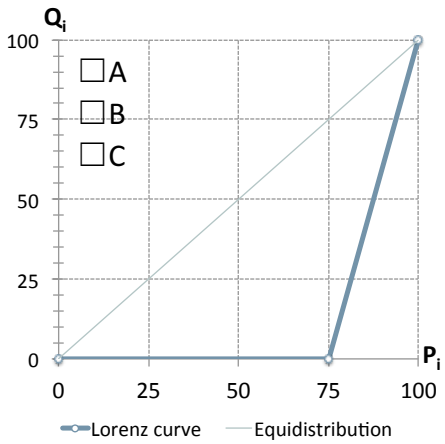


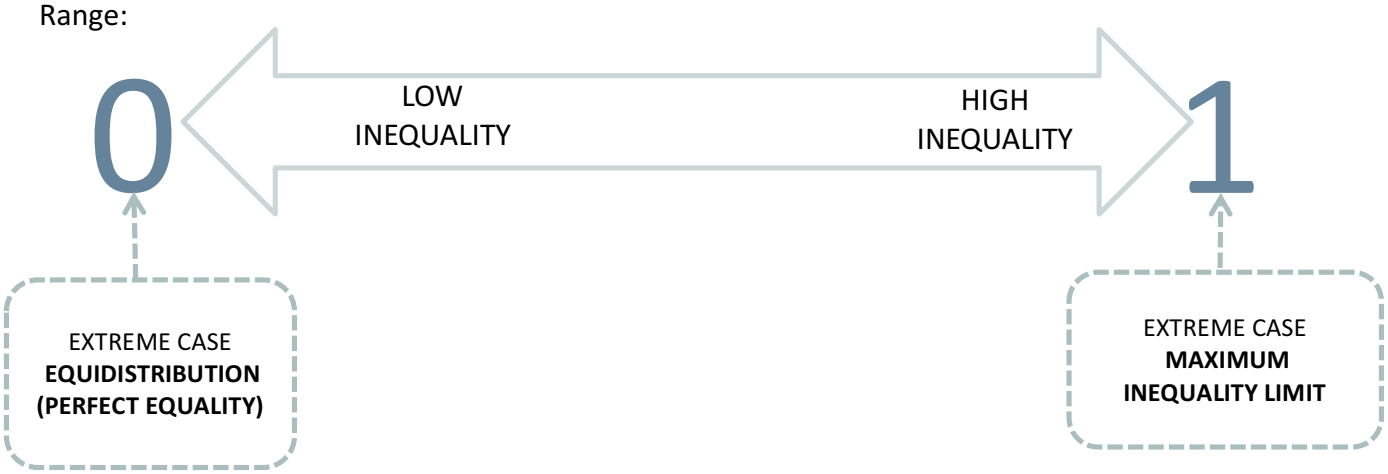



## GINI INDEX

The Gini index is the most commonly used statistical indicator for measuring inequality in the frequency distribution of a variable.

Range:




CALCULATION

There are many formulas available. We will use:

$$G = \frac{\sum_{i=1}^{k-1} P_i Q_{i+1} - \sum_{i=1}^{k-1} P_{i+1} Q_i}{10000}$$

Where:

$P_i$ : cumulative percentages of frequencies,

$Q_i$ : cumulative percentages of the total amount of the variable (cumulative $q_i$)

$$q_i = 100 \frac{x_i n_i}{\sum_{i=1}^{k} x_i n_i}$$

PROPERTIES

- It is not affected by scales changes (scale independence)
- Very sensitive to outliers

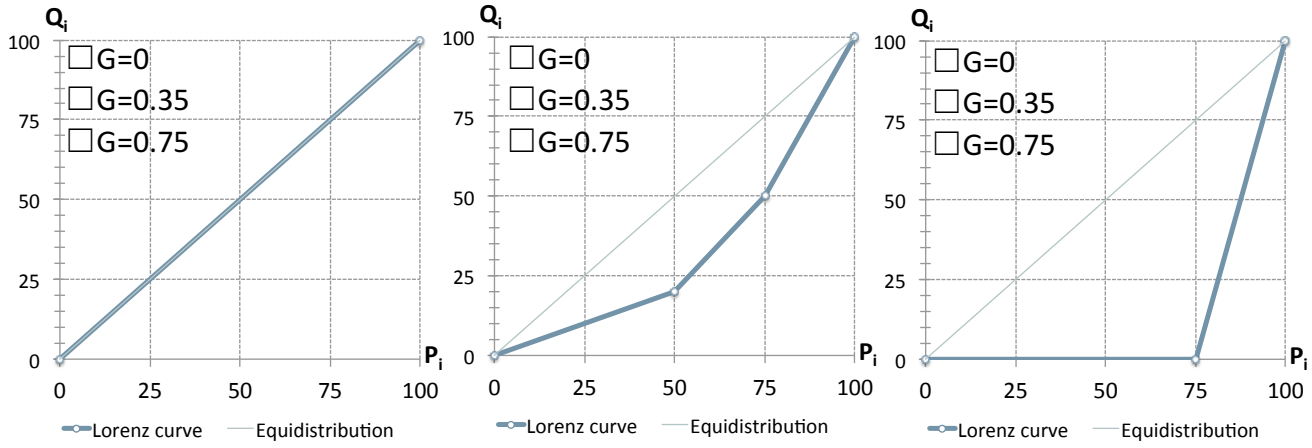
## CASE STUDY 1

Gini index

| $X_i$ | $n_i$ | $x_i \cdot n_i$ | $p_i$ | $q_i$ | $P_i$ | $Q_i$ | $P_i \cdot Q_{i+1}$ | $P_{i+1} \cdot Q_i$ |
|---|---|---|---|---|---|---|---|---|
| 10000 | 2 | 20000 | 50 | 20 | 50 | 20 | 2500 | 1500 |
| 30000 | 1 | 30000 | 25 | 30 | 75 | 50 | 7500 | 5000 |
| 50000 | 1 | 50000 | 25 | 50 | 100 | 100 | | |
| | 4 | 100000 | 100 | 100 | | | 10000 | 6500 |

$$G = \frac{10000 - 6500}{10000} = 0.35$$

H) Match each Lorenz curve with its Gini index




## CASE STUDY 2

In a recent study about salaries inequality in tourism related sectors in Wisconsin, U.S.A. (Marcouiller & Xia, 2008), the following Gini indexes have been estimated:

| Subsector | Gini index |
|---|---|
| A: *Performing Arts and Spectator Sports* | 0.2334 |
| B: *Museums, Parks and Historical Sites* | 0.2182 |
| C: *Food Services and Drinking Places* | 0.0418 |

Which subsector(s) fit(s) these descriptions?

A B C

☐☐☐ 'presence of a small number of highly paid employees and a relatively large number of lower paid employees'

☐☐☐ 'generally common and persistent problem of low wages experienced across this occupational category'

## CASE STUDY 3

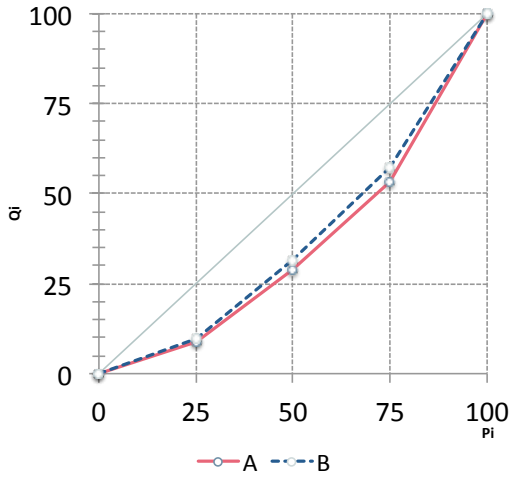
The following table shows the distribution of women's monthly salaries in the hospitality sector in Spain in 2010, according to the survey *Encuesta de estructura salarial 2010* by the National Statistics Institute (the last interval has been adjusted).

| **Monthly salary ($10^3$€)** | **$X_i$** | **ni** | **$x_i \cdot n_i$** | **$p_i$** | **$q_i$** | **$P_i$** | **$Q_i$** | **$P_i \cdot Q_{i+1}$** | **$P_{i+1} \cdot Q_i$** |
|---|---|---|---|---|---|---|---|---|---|
| 0.0 - 13,5 | 6.75 | 43650 | 294637.5 | 25 | 9.64 | 25.00 | 9.64 | | |
| 13.5 - 16,8 | 15.15 | 43650 | 661297.5 | 25 | 21.64 | 50.00 | 31.29 | | |
| 16.8 - 19,7 | 18.25 | 43650 | 796612.5 | 25 | 26.07 | 75.00 | 57.36 | | |
| 19.7 and more | 29.85 | 43650 | 1302952.5 | 25 | 42.64 | 100.00 | 100.00 | | |
| | | 174600 | 3055500.0 | 100 | 100 | | | | |

A) Complete the table, and calculate the average salary and the Gini index for women.
Average salary: _____ Gini index: _____

B) The men's average salary (from the same survey) is 22.6 ($10^3$€) and the estimated Gini index is 0.30.

| **Monthly salary ($10^3$€) MEN** | **$n_i$** |
|---|---|
| 0.0 – 16.3 | 31950 |
| 16.3 – 19.5 | 31950 |
| 19.5 – 24.7 | 31950 |
| 24.7 and more | 31950 |
| | 127800 |




Identify the Lorenz curves in the figure
Women ☐ A ☐ B
Men ☐ A ☐ B

C) What is the highest salary earned by the 25% less-paid workers?
Women: _____ Men: _____

D) What is the lowest salary earned by the highest-paid workers?
Women: _____ Men: _____

E) Do the 25% highest-paid women earn more than 40% of the total amount of women's salaries?
☐ Yes ☐ No

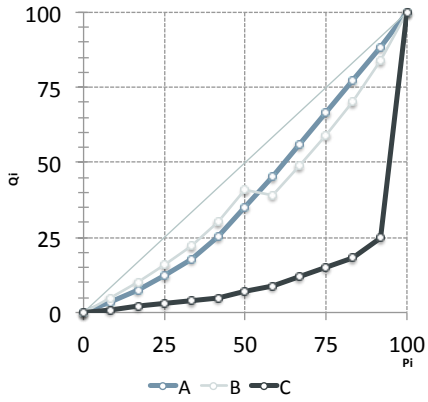
F) Do the 25% less-paid women earn less than 10% of the total amount of women's salaries?
☐ Yes ☐ No

## CASE STUDY 4

The table below shows the monthly distribution of visitors (thousands) to the Palazzo Ducale, Venice, in 2012 (source: Panciera, 2013). The distribution is mainly concentrated between April and October, with a Gini index = 0.1932.

A) Complete the table.

| Month | $X_i$ | $x_i \cdot n_i$ | $p_i$ | $q_i$ | $P_i$ | $Q_i$ | $P_i \cdot Q_{i+1}$ | $P_{i+1} \cdot Q_i$ |
|---|---|---|---|---|---|---|---|---|
| ENE | 48 | 48 | 8.3333 | 3.670 | 8.333 | 3.670 | 61.80 | 61.17 |
| DIC | 49 | 49 | 8.3333 | 3.746 | 16.667 | 7.416 | 205.15 | 185.40 |
| NOV | 64 | 64 | 8.3333 | 4.893 | 25.000 | 12.309 | 441.52 | 410.30 |
| FEB | 70 | 70 | 8.3333 | 5.352 | 33.333 | 17.661 | 835.90 | 735.87 |
| MAR | 97 | 97 | 8.3333 | 7.416 | 41.667 | 25.077 | 1462.16 | 1253.84 |
| OCT | 131 | 131 | 8.3333 | 10.015 | 50.000 | 35.092 | 2266.84 | 2047.03 |
| JUN | 134 | 134 | 8.3333 | 10.245 | 58.333 | 45.337 | 3255.63 | 3022.45 |
| SEP | 137 | 137 | 8.3333 | 10.474 | 66.666 | 55.811 | 4429.18 | 4185.81 |
| AGO | 139 | 139 | 8.3333 | 10.627 | 75.000 | 66.438 | 5785.55 | 5536.48 |
| JUL | 140 | 140 | 8.3333 | 10.703 | 83.333 | 77.141 | 7345.80 | 7071.23 |
| MAY | 144 | 144 | 8.3333 | | | | | |
| ABR | 155 | 155 | 8.3333 | 11.850 | 100.000 | 100.000 | | |
| | 1308 | 1308 | 100.000 | 100.000 | | | | |




B) What curve corresponds to the table data? ☐ A ☐ B ☐ C

C) Simulate the distribution resulting from a 10% increment in each month. Compare the new Lorenz curve and Gini index with the original ones.

| ENE | FEB | MAR | ABR | MAY | JUN | JUL | AGO | SEP | OCT | NOV | DIC |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 53 | 77 | 107 | 171 | 158 | 147 | 154 | 153 | 151 | 144 | 70 | 54 |

Does the 10% monthly increment cause a change in the Gini index?

☐ Yes ☐ No

Why?

Continue this activity in the Seasonal Concentration Lab
http://goo.gl/xEcl97
(Fernández-Morales, 2015)

## SOLUTIONS

CASE STUDY 1

A) 100000 €

B) 50%

C) 50%

D) 20%

E) 20%

F) 50%

G) C, A, B

H) 0, 0.35; 0.75

CASE STUDY 2

A, B

C

CASE STUDY 3

A) Average: 17.5; G=0.26

B) B: Women A: Men
C) Women: 13500€ , Men: 16300€
D) Women: 19700€, Men: 24700€
E) Yes
F) Yes

CASE STUDY 4
A) 11.009, 91.666, 88.150, 9166.63, 8814.96
B) A
C) No

## REFERENCES


[1] Cisneros-Martínez, D. and Fernández-Morales, A. (2016). Understanding the seasonal concentration of tourist arrivals: the case of the south of Spain. In A. Artal-Tur and M. Kozak, M. (eds.) *Destination competitiveness, the environment and sustainability: challenges and cases* (131-143). Wallingford, U.K.: CABI.

[2] Cisneros-Martínez, D. and Fernández-Morales, A. (2015). Cultural tourism as tourist segment for reducing seasonality in a coastal area: the case study. *Curent Issues in Tourism* 18(8), 765-784.

[3] Duro, J. A. (2016) Seasonality of hotel demand in the main Spanish provinces: measurements and decomposition exercises. *Tourism Management* 52, 52-63.

[4] Fernández-Morales, A. (2015). Simulating seasonal concentration in tourism series. *Journal of Hospitality, Leisure, Sport & Tourism Education* 15, 116-123.

[5] Fernández-Morales, A. and Cisneros-Martínez, J.D. (2015). Seasonal inequalities in visitor distribution in Argentina's tourism regions. *E-Review of Tourism Research* 12(1), 115-126.

[6] Fernández-Morales, A. and Martín-Carrasco, Y. (2014). Concentración e impacto estacional del turismo de cruceros en Málaga. *Revista de Estudios Regionales* 101. 43-70.

[7] Fernández-Morales, A. and Lacomba-Arias, B. (2003). *Técnicas Estadísticas para el Turismo. Nociones teóricas y problemas resueltos*. Málaga, Spain: Editorial Ágora.

[8] Giorgi, G.M. (1993). A fresh look at the topical interest of the Gini concentration ratio, *Metron*, vol. LI (1-2), 93-98.

[9] López Bonilla, J.M. and López Bonilla, L.M. (2006). La concentración estacional en las regiones españolas desde una perspectiva de la oferta turística. *Revista de Estudios Regionales* 77, 77-104.

[10] Marcouiller, D.W. and Xia, X. (2008). Distribution of income from tourism-sensitive employment, *Tourism Economics*, 14 (3), 545–565.

[11] Martín Reyes, G., García Lizana, A., Fernández-Morales, A. (2012). *La distribución de la renta en la ciudad de Málaga. Índice de Gini, curva de Lorenz y pobreza en 2011*, Servicio de Programas (Observatorio de Medio Ambiente Urbano), Ayuntamiento de Málaga.

[12] Panciera, R. (2013). *Annuario Turismo 2012.* Comune di Venezia, Assessorato al Turismo, Venezia.

[13] Roselló Nadal, J., Riera Font, A. and Sansó Roselló, A. (2004). The economic determinants of seasonal patterns. *Annals of Tourism Research* 31(3), 697-711.