Ignacio Cascos, Probability
Master in Statistics for Data Science
Universidad Carlos III de Madrid

5. Random vectors (2/2)

Problem 1. Assume heights for women follow a normal distribution with mean of 64 inches and standard deviation 3 inches, while men's heights follow a normal distribution with mean of 70 inches and standard deviation 3 inches. Assume further that 51% of the individuals in a population are women and the joint distribution of the heights of a couple of siblings (one sister and one brother) is normal with correlation 0.6.

1. What is the probability that a woman selected at random is taller than 68 inches?
2. What height is exceeded by 65% of women?
3. Compute the mean and variance of the height of an individual selected at random.
4. What is the probability that an individual selected at random is taller than 68 inches?
5. What height is exceeded by 65% of individuals?
6. Consider now a couple of siblings (one sister and one brother). What is the distribution of their average height?
7. What is the probability that the midheight from part 7. is greater than 68 inches?
8. What midheight is exceeded by 65% of couples of siblings (one sister and one brother)?

Problem 2. A firm has workers of 5 types with five different salaries

number of workers	10	5	2	3	1
salary	1000	3000	5000	10000	50000

1. Represent the Lorenz curve of the salaries. What is the Gini index?
2. The firm will spend 10000 extra monetary units in salaries, how should they spend that amount of money in order that no worker will earn less salary than now and the distribution of salaries is as egalitarian as possible. What will the new Gini index be?
3. Answer the previous question if instead of 10000, the firm can dispose of 35000 extra monetary units.

Problem 3. When events occur in a time period $(0, a)$ with regard to a Poisson process, it is well known that, conditioned on the total number of events k, the joint distribution of the times at with the events occur follows a uniform distribution in $(0, a)^k$. That is, if X_i represents the arrival time of one of them, then $X_i \sim \mathrm{U}(0, a)$ and it is independent of the other X's.

1. If ten patients visit the Emergency Room (ER) of a hospital between 9 and 10, what is the probability that fifth of them reaches the ER before 9:30?
2. If ten patients visit the ER between 9 and 10, what is the probability that the first 3 of them arrive before $9:20$, 4 of them between $9:20$ and $9:40$, and the last 3 of them after $9:40$?
3. If two patients visit the ER of a hospital between 9 and 10, what is the conditional distribution of the time at which the second patient arrived at the ER given that the first patient arrived at time t_1?