of the parameters \( \rho_i \) and \( \beta_i \) is well defined, then we can model the collective call behavior of the users and see its evolution over time. From now on, we will call the meta-distribution of the parameters \( \rho_i \) and \( \beta_i \) the MetaDist distribution.

In Figure 5-a, we show the scatter plot of the parameters \( \rho_i \) and \( \beta_i \) of the CDD of each user \( i \) for the first month of our dataset. We can not observe any latent pattern due to the overplotting but, however, we can spot outliers. Moreover, by plotting the \( \rho_i \) and \( \beta_i \) parameters using isocontours, as shown in Figure 5-b, we automatically smooth the visualization by desconsidering low populated regions. While darker colors mean a higher concentration of pairs \( \rho_i \) and \( \beta_i \), white color mean that there are no users with CDDs with these values of \( \rho_i \) and \( \beta_i \).

![Fig. 5. Scatter plot of the parameters \( \rho_i \) and \( \beta_i \) of the CDD of each user \( i \) for the first month of our dataset. In (a) we can not see any particular pattern, but we can spot outliers. By plotting the isocontours (b), we can observe how well a bivariate Gaussian (c) fits the real distribution of the \( \rho_i \) and \( \beta_i \) of the CDDs ('meta-fitting').](image)

Surprisingly, we observe that the isocontours of Figure 5-b are very similar to the ones of a bivariate Gaussian. In order to verify this, we extracted from the MetaDist distribution the means \( P \) and \( B \) of the parameters \( \rho_i \) and \( \beta_i \), respectively, and also the covariance matrix \( \Sigma \). We use these values to generate the isocontours of a bivariate Gaussian distribution and we plotted it in Figure 5-c. We observe that the isocontours of the generated bivariate Gaussian distribution are similar to the ones from the MetaDist distribution, which indicates that both distributions are also similar. Thus, we conjecture that a bivariate Gaussian distribution fits the real distribution of \( \rho \) and \( \beta \)s, making the MetaDist a good model to represent the population of users with a determined calls duration behavior.

Given that the MetaDist is a good model for the group behavior of the customers in our dataset, we can now visualize and measure how them evolve over time. In Figure 6 we show the evolution of the MetaDist over the four months of our dataset. The first observation we can make is that the bivariate Gaussian shape stands well during the whole analyzed period, what validates the robustness of the MetaDist. Moreover, a primarily view indicates that the meta-parameters also have not change significantly over the months. This can be confirmed by the first 5 rows of Table 1, which describes the value of the meta-parameters \( P, B \) and \( \Sigma(\sigma^2_{\rho}, \sigma^2_{\beta}, \text{cov}(\rho_i, \beta_i)) \) for the four analyzed months. This indicates that the phone company already reached a stable state before its customers concerning its prices, plans and services. In fact, the only noticeable
Surprising Patterns for the Call Duration Distribution of Mobile Phone Users

difference occurs between the first month and the others. We observe that the meta-parameters of the first month have a slightly higher variance than the others, what indicates that this is probably an atypical month for the residents of the country in which our phone records were collected. But in spite of that, in general, the meta-parameters do not change through time. Then, we can state the following observation:

**Observation 1.** **TYPICAL BEHAVIOR.** The typical human behavior is to have a efficiency coefficient $\rho \approx 1.59$ and a weakness coefficient $\beta \approx -6.25$. Thus, the median duration for a typical mobile phone user is 51 seconds and the mode is 20 seconds.

![Fig. 6. Evolution of the MetaDist over the four months of our dataset. Note that the collective behavior of the customers is practically stable over time.](image)

**4.2 Focal Point**

An interesting observation we can derive from the MetaDist showed in Figure 5 is that there exists a significant negative correlation between the parameters $\rho_i$ and $\beta_i$. This negative correlation, more precisely of $-0.86$, lead us to the fact that the OR lines, i.e., the TLAC odds ratio plots of the customers of our dataset, when plotted together, should cross over a determined region. In order to verify this, we plotted in Figure 7-a the OR lines for some customers of our dataset. As we can observe, it appears that these lines are all crossing in the same region, when the duration is approximately 20 seconds and the odds ratio approximately 0.1. Then, in Figure 7-b, we plotted together the OR lines of 20,000 randomly picked customers and derived from them the isocontours to show the most populated areas. As we can observe, there is a highly populated point when the duration is 17 seconds and the OR is 0.15. By analyzing the whole month dataset, we verified that more than 50% of the users have OR lines that cross this point. From now on, we call this point the **Focal Point**.

Formally, the **Focal Point** is a point on the OR plot with two coordinates: a coordinate $FP_{duration}$ in the duration axis and a coordinate $FP_{OR}$ in the OR axis. When a set of customers have their OR plots crossing at a **Focal Point** with coordinates $(FP_{duration}, FP_{OR})$, it means that for all these customers the $\frac{FP_{OR}}{1+FP_{OR}}$th percentile of their CDD is on $FP_{duration}$ seconds. Thus, in the 2 bottom lines of Table 1, we describe the **Focal Point** coordinates for the four months of our analysis and, surprisingly, the **Focal Point** is stationary. Thus, we can make the following observation:

**Observation 2.** **UNIVERSAL PERCENTILE.** The vast majority of mobile phone users has the same 10th percentile, that is on 17 seconds.