Forecasting payment terminals failure

Patrick

1/19/2020

Project summary

Say the company we work for manufactures and sells electronic devices in a B2B context while offering after sales services. In some sectors, shutdown time might be extremely costly for clients and handling device reparations requires maintening a sufficient stock of spare parts at hand. Yet, producing and storing those parts comes with a cost which is why inventory levels should be about just high enough to cover the demand. This project explores the idea of quantifying and forecasting the occurence of device failures and therefore the future reparation needs using as input historical data on product sales and reparations (as a proxy for failure occurences).

Literature reviews on the topic (Auweraer, Boute, and Syntetos 2019; Dekker et al. 2013) stress the poor performence of traditional time series forecasting methods for the task of predicting spare parts demand. A first challenge is that the occurence of failures is erratic and intermitent. Another is that our expectations on the number of failures should obviously take into account the number of devices currently on the market, which we will refer to as the *installed base*, as well as its expected evolution, rather than naïvely extrapolating patterns observed in historical failure data.

The method explored here consists in three steps: First, a time-to-failure probability distribution will be estimated for each device part and failure type based on historical failure data. In a second phase, a forecast of future failures is performed based on the current state of the installed base by exploiting the fact that the expected number of part failures in time `t` is the sum of the individual probabilities to fail in time `t` for all parts in service. Finally, the exercise will be repeated in a setting where the future size and composition of the installed base is estimated.

Data

The data used as source for the analysis is generated for the purpose of this project (see `Dataset generation: commented code`). The minimum requirement for this methodology to work is that one can gather the following information for (most) devices that have been sold in the past:

- its model;
- the date at which it has been sold;
- the reparation date, if any;
- the description of the failure (i.e. part that failed and reason), if any.

The dataset used here contains 5000 records of that type (see first 5 records on Table 1). In total, 3000 devices of model `A` were sold between January 01 2017 and December 31 2019 and 2000 devices of model `B` were sold between January 01 2018 and December 31 2019. In total, 4 different types of failures were recorded with the following frequencies listed in Table 2.

Table 1: Dataset

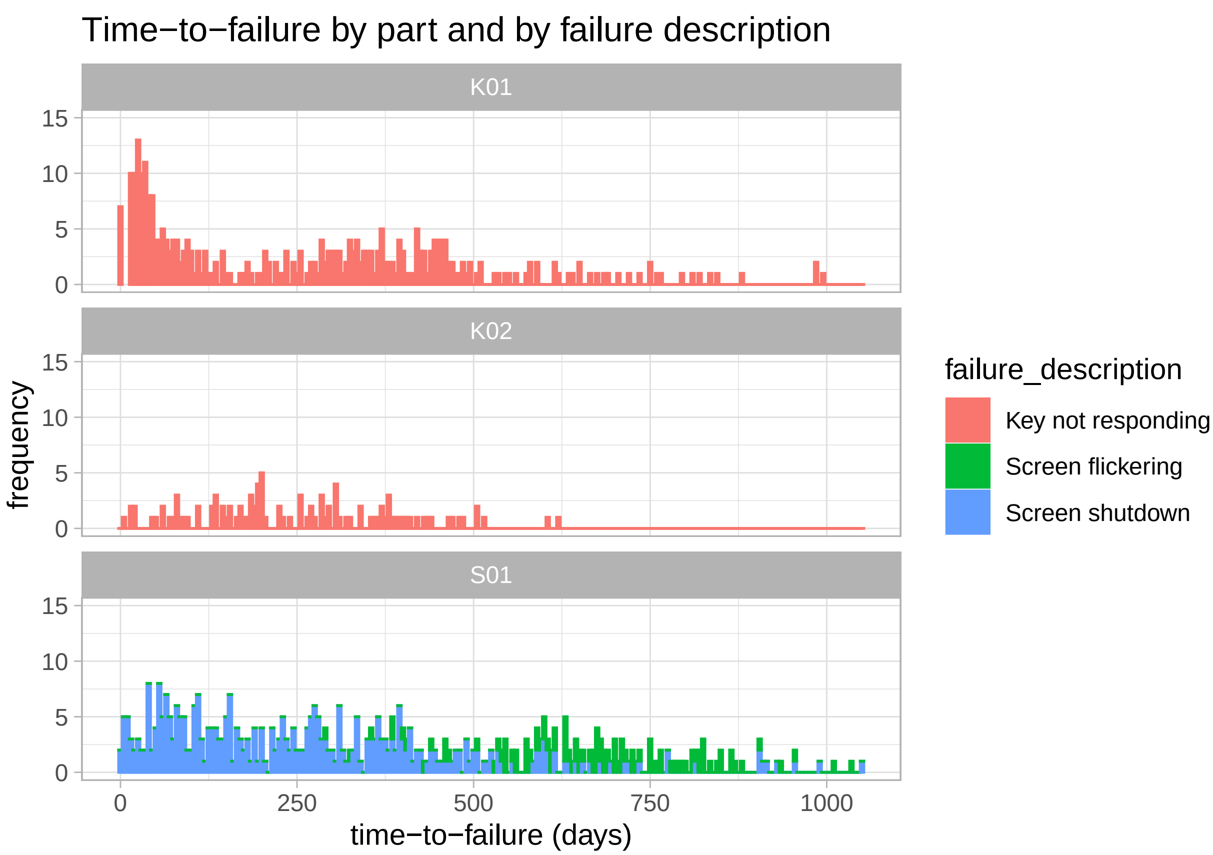
X	terminal_model	date_sold	date_repaired	part_code	failure_description
1	A	2018-10-08	2018-11-09	K01	Key not responding
2	A	2019-10-29	NA	S01	NA
3	A	2019-02-01	NA	S01	NA
4	A	2018-12-31	NA	K01	NA
5	A	2017-11-16	NA	K01	NA

Table 2: Failures frequency

failure_description	n
Key not responding	454
Screen flickering	106
Screen shutdown	345
NA	4095

Exploration and density estimation

Now let's start exploring the dataset as if we didn't know its generating process. First we'd be interested in understanding the lifetime at which failures generally occur. A first look reveals that the typical time-to-failure is quite different from part to part, hence from device to device.



We already see some interesting patterns here, for example a lot of failures are recorded in the first 50 days of life for the part *K01*, while failure seems more spead for other parts, but it's quite hard to tell more at this point. Besides, such a histogram suffers from a bias due to the right-censored nature of time-to-failure data. Suppose the vast majority of the devices containing the part *K01* have been sold in the last 50 days, then the relative peak we observe on the figure might simply reflect the fact that most of the devices about which we know the reparation date are relatively young.

A way to make better use of these data is to estimate probability density functions of time-to-failure for each device part. The obtained density functions will have three advantages:

- get a smooth distribution of lifetimes at which failures are likely to occur by getting rid of uninformative spikes;
- get a vector of (relative) failure probabilities instead of frequencies. This is what we'll need for the forecasting part.
- provided that a proper technique is used, get rid of the censoring bias.

We'll use a version of the Kernel density estimator introduced by Földes, Rejtő, and Winter (1981) in which kernels are weighted so as to make it robust to right-censoring. The R function `presmooth` implements that method.

A function will be applied that returns for each pair (*part*; *failure type*) a probability distribution in the part's time-to-failure.

References

Auweraer, Sarah Van der, Robert N Boute, and Aris A Syntetos. 2019. "Forecasting Spare Part Demand with Installed Base Information: A Review." *International Journal of Forecasting* 35 (1). Elsevier: 181–96.

Dekker, Rommert, Çerağ Pinçe, Rob Zuidwijk, and Muhammad Naiman Jalil. 2013. "On the Use of Installed Base Information for Spare Parts Logistics: A Review of Ideas and Industry Practice." *International Journal of Production Economics* 143 (2). Elsevier: 536–45.

Földes, A, L Rejtő, and B Winter. 1981. "Strong Consistency Properties of Nonparametric Estimators for Randomly Censored Data, Ii: Estimation of Density and Failure Rate." *Periodica Mathematica Hungarica* 12 (1). Akadémiai Kiadó, co-published with Springer Science+ Business Media BV ...: 15–29.