Efficiency of email communication

Prepared by James Hogg and Kerrie Mengersen for Good Directions

April 3, 2024

Researchers: James Hogg Kerrie Mengersen

EXECUTIVE SUMMARY

1. Instruction
2. Method
3. Outputs
4. What do we find

Introduction

Instructions

Method

Results

Description of Dashboard

Statistical methodology

Basic

Uncertainty

To relax the assumption of a fixed read time, we simulate read times from appropriate Gamma Distributions, where the mean is based on the input value. To allow Good Directions to provide their uncertainty we allow the user to input a 'best guess' (aka mean) and a range of possible values (aka 95% interval). Assuming the user is imaging read times as approximately normally distributed, the 95% interval provided is converted to a standard deviation via $(U - L)/3.92$.

By first observing the number of emails of each type, we simulate and then sum the read times. We conduct this process 1000 times and then summarise the distribution of the total read times using medians and quantiles.

Visualise

Case studies

References

Appendixes

Basic Calculator

The following notation is used:

- N_i: Number of emails of type i, where $i \in \{1, 2, 3\}$ corresponds to the three different types of emails.
- R_i: Fixed time to read each email of type i.

- A_i: Fixed time to take action on each email of type i.
- W_i: Fixed time to write or respond to each email of type i.
- P_i: Probability that an email of type i is essential.
- T_{total}: Total time spent on all emails.
- $T_{essential}$: Total time spent on essential emails.
- $T_{nonessential}$: Total time spent on nonessential emails.

The total time spent on emails of each type is calculated by summing the product of the number of emails and the respective times for reading, action, and writing. The formula for the total time spent on each type i is given by:

$$T_i = N_i \times (R_i + A_i + W_i)$$

This is calculated for each type i, and then summed to find the total time spent on all emails:

$$T_{total} = \sum_{i=1}^{3} T_i$$

To differentiate between essential and nonessential times, we use the probability P_i:

$$T_{essential,i} = P_i \times T_i$$

$$T_{nonessential,i} = (1 - P_i) \times T_i$$

Then, to find the aggregate time spent on essential and nonessential emails across all types, we sum across the respective times for each type:

$$T_{essential} = \sum_{i=1}^{3} T_{essential,i}$$

$$T_{nonessential} = \sum_{i=1}^{3} T_{nonessential,i}$$

The cost associated with the time spent on emails can be determined by multiplying the time by the pay rate. Let Pay denote the average pay per hour. The cost of time spent on all emails and the cost of time spent on essential and nonessential emails are calculated as follows:

$$\text{Cost}_{total} = T_{total} \times \text{Pay}$$

$$\text{Cost}_{essential} = T_{essential} \times \text{Pay}$$

$$\text{Cost}_{nonessential} = T_{nonessential} \times \text{Pay}$$

These formulas provide the basis for the deterministic calculations used to simulate the total time and cost associated with email handling within an organization.

Intermediate Calculator

The intermediate calculator introduces variability into the derivations, both in the count of received emails and the time taken for the email activities (e.g., read, action, respond). The time for the email activities are modeled using the Gamma distribution, parameterized by the means and standard deviations derived from user inputs. The notation is as follows:

- N_i: Random variable for the number of emails of type i, with $i \in \{1, 2, 3\}$ denoting different email types.
- R_i, A_i, W_i: Random variables for the time to read, action, and write emails of type i, respectively.
- P_i: Probability that an email of type i is essential.
- T_{total}: Aggregated time spent on all emails across simulations.
- $T_{essential}$: Aggregated time spent on essential emails across simulations.
- $T_{nonessential}$: Aggregated time spent on nonessential emails across simulations.
- $R_{lower}, R_{upper}, A_{lower}, A_{upper}, W_{lower}, W_{upper}$: User-specified lower and upper bounds for reading, action, and writing times, respectively.

1. **Sample the number of emails**:

- For each type i:
$$N_i \sim \text{Uniform}(N_{lower,i}, N_{upper,i})$$

2. **Parameterisation of email activities**:

- The parameters for the Gamma distribution are derived from the user-specified lower and upper bounds as follows:

$$\mu_i = \frac{R_{upper,i} + R_{lower,i}}{2}$$

$$\sigma_i^2 = \left(\frac{R_{upper,i} - R_{lower,i}}{3.92} \right)^2$$

- Similar calculations are made for action and writing times.

3. **Gamma distribution for time activities**:

- For each type i and activity (reading R_i, action A_i, writing W_i):

$$R_i, A_i, W_i \sim \text{Gamma}(k = \frac{\mu_i^2}{\sigma_i^2}, \theta = \frac{\sigma_i^2}{\mu_i})$$

- The shape k and scale θ parameters are calculated using the derived mean μ_i and variance σ_i^2.

4. **Calculate total time**:

- The total time for each type i and for each simulation iteration s is the sum of the sampled times:
$$T_i = \sum (R_i + A_i + W_i)$$

5. **Stratify by essential and nonessential emails**:

- The distinction between essential and nonessential times is made by applying the probability P_i via a Bernoulli trial for each email.

$$T_{essential,i} = T_i \times \text{Bernoulli}(P_i)$$

$$T_{nonessential,i} = T_i \times (1 - \text{Bernoulli}(P_i))$$

6. **Calculate yearly hours and cost**

- The calculators required to derive the yearly hours and cost values follow those given in the basic calculator.

By running the above steps 1000 times, the intermediate calculator can estimate the distribution of the output values, including total, essential, and nonessential times, and yearly hours and cost. To summarise the simulations a point estimate (e.g., median) and interval (2.5% and 97.5% quantiles) are reported in the dashboard.

Deliverables

Objectives

- Report that describes the email communication model, including methodologies used, exemplar outcomes, description of the dashboard and illustrations of outputs from the dashboard (approximately 10 pages)
- Interactive email efficiency dashboard (software tool)

Background

https://www.marketingsherpa.com/article/average-email-open-time-is

https://www.radicati.com/wp/wp-content/uploads/2015/02/Email-Statistics-Report-2015-2019-Executive-Summary.pdf states that "In 2015, the number of business emails sent and received per user per day totals 122 emails per day. This figure continues to show growth and is expected to average 126 messages sent and received per business user by the end of 2019." - Article states that on average a person sends 30 emails a day and recieves about 96

Fantastic article with lots of statistics https://hbr.org/2019/01/how-to-spend-way-less-time-on-email-every-day#:~:text=The%20average%20professional%20spends%2028,120%20messages%20received% states "The average professional spends 28% of the work day reading and answering email, according to a McKinsey analysis. For the average full-time worker in America, that amounts to a staggering 2.6 hours spent and 120 messages received per day."

Chronic emailers "These are the people that copy everyone on everything they do. It is usually done as a vain attempt to look important or busy. What these people are really doing is wasting a LOT of other people's time. Company emails should be limited and should conveyed on a "need to know" basis. That means ONLY the relevant information is disseminated, and ONLY to those that need to know it." "

Methodology

In this section, we describe the formula used to derive the key outputs of the dashboard:

- number of emails recieved
- total reading time
- staff cost