We demonstrate that distributions of human response times have power-law tails and, among closed-form distributions, are best fit by the generalized inverse gamma distribution. We speculate that the task difficulty tracks the half-width of the distribution and show that it is related to the exponent of the power-law tail.

I. INTRODUCTION

Human response time (RT) is defined as the time delay between a signal and the starting point of human action. For example, one can measure time interval from a word appearing on a computer screen to when a participant pushes a keyboard button to indicate his or her response. Two well established empirical facts of RT are the power law tails of RT distributions \cite{1} and 1/f noise of RT time series \cite{2–5}, to which any theoretical description must conform.

The generalized inverse gamma (GIGa) function (Appendix A) belongs to a family of distributions (Appendix B), which includes inverse gamma (IGa), lognormal (LN), gamma (Ga) and generalized gamma (GGa). The remarkable property of GIGa is its power-law tail; for a general three-parameter case, the power-law exponent is given by the negative $1+\alpha\gamma$, so that GIGa($x; \alpha, \beta, \gamma$) $\propto x^{-1-\alpha\gamma}, x \to \infty$. GIGa emerges as a steady state distribution in a number of systems, from a network model of economy, \cite{6} to ontogenetic mass growth, \cite{7} to stock volatility \cite{8}. This common feature can be traced to a birth-death phenomenological model subject to stochastic perturbations (Appendix C).

Here we argue that among closed form distributions the GIGa best describes RT distribution. GIGa has a natural scale parameter, which determines the onset of the power law tail, and two shape parameters, which determine the exponent of the tail. As such, our argument is an extension of previous approaches, such as “cocktail” model, \cite{11} which effectively contains shape and scale parameters as well. Furthermore, we speculate that the difficulty of a cognitive task tracks the half-width of the RT distribution and discuss it within the GIGa framework.

Our numerical analysis is performed on the following data (explained in text): ELP (English Lexicon Project), HE (Hick’s Experiments) and LDT (Lexical Decision Time). Two key features distinguish our approach. First, in addition to usual individual participant fitting, we perform distribution fitting on combined participants’ data. While in line with individual fitting, this creates considerably less noisy sets of data. Second, we develop a procedure for fitting the tails of the distribution directly (Appendix D), which unequivocally proves the existence of power law tails.

This paper is organized as follows. In Section II, we provide description of the experimental setup and data acquisition. In Section III, we conduct log-log tail fitting and RT distribution fitting with GIGa. In Section IV, we conclude with the discussion of task difficulty.

II. DATA ACQUISITION

A. Data sources and description

ELP data is from the English Lexicon Project \cite{9, 10}. HE and LDT data was collected under the supervision of J. G. Holden.

1. ELP

ELP (English Lexicon Project) studies pronunciation latencies to visually presented words; participants sampled from six different Universities. \cite{9, 10}.

Data: Two sessions, 470 participants each: session 1 (ELP1), 1500 trials; session 2 (ELP2), 1030 trials.

2. HE

HE (Hick’s Choice RT Experiment) – given a stimulus selected from a finite set of stimuli, participants try to respond with an action from a set of actions corresponding to this set of stimuli. Original HE is described in \cite{11}.

Data: 11 participants completed 1440 trials of 2, 4, 6, 8, and 10 options, approximately 16 000 combined datapoints for each condition.
3. LDT

LDT (Lexical Decision Time).

Data: Three groups of 60 participants completed 100 word and 100 nonword trials of 1, 2, and 4 word LDT respectively, only the correct word trials are depicted, approximately 6000 datapoints for each group.

B. Data preprocessing

To enhance our efforts to understand the distribution’s tail behavior, we combined all participants’ data from each experiment into a single distribution.

III. DATA ANALYSIS

A. Tail fitting

Log-log plot of power law tail fitting is discussed in Appendix D. In Figs. 1 and 2 we show the results for RT experiments. With the exception of LDT, trials for most of the tasks timed out by 4 or 5 seconds. This requirement has the potential to distort RT distributions, especially their slow tails, as log-log plot bends downward when RT is close to 4 seconds. (In the future, the requirement of maximum time limit should be dropped or, at least, the limiting time cutoff must increase to reflect the natural RT distribution.) In contrast, the maximum RT for LTD is approximately 10 seconds and, as seen in Figs. 3 and 4 the log-log plots are closer to straight lines and GIGa fit is good.

B. GIGa distribution fitting

In Figs. 4 and 5 we show GIGa distribution (Appendix A) fitting of RT. In the figures, the distance from the origin to the blue dot is rightward shift of GIGa distribution. The RTs to the left of the red lines are cut in the fitting of GIGa distribution. \( \alpha, \beta, \gamma \), the cut and shift parameters are all found by minimizing the chi-squared test statistic as follows. We choose the cut and shift parameters, find \( \alpha, \beta, \gamma \) through maximum likelihood estimation and compute the chi-squared test statistic. We repeat this process for another group of cut and shift parameters. In the end, we obtain the parameters that minimize the chi-squared test statistic.

Visually, GIGa fitting is good, yet p-values are all zero with the exception of LDT. As discussed above, a possible explanation is that the participants are not given enough time to respond, which distorts RT distributions. Also, Ref. 12 argues that chi-squared statistic yields poor results for goodness-of-fit – we used chi-squared statistic because, due to the cut parameter, the total number of RTs is not fixed in our parameter fitting. Lastly, in Fig. 6 we show the relationship between the tail exponent parameter \( \alpha \gamma \) and log-log fitted exponent parameter – with the exception of 4 LDT (which is one of the hardest tasks – see below), the correspondence is quite good.

IV. TASK DIFFICULTY

In Fig. 5 we plot the power law exponent from the best fit GIGa above as a function of their half-width. With the exception of Hick 6, there is a clear tracking between the two (notice that by eye HE PDFs seemingly show decrease of modal PDF and increase of PDF half-width
with the increase of Hick’s number).

We speculate that the half width of the distribution would be a natural measure of a task difficulty. This is easily analyzed in terms of the GIGa distribution, which we believe is well suited to description of RT distributions. In Appendix A it is explained that due to GIGa’s scaling property, it is sufficient to consider the $\gamma = 1$ case, that is IGa. Furthermore, we can eliminate one more parameter by setting mean to unity. In some cognitive tasks, the mean may not be a good indicator of difficulty since an easy cognitive task may require a more idiosyncratic response and vice versa.

For such IGa, a single parameter $\alpha$ then defines both scale and shape, that is the half width is directly related to the exponent of the power law tail. As seen in Fig. 9 it has a maximum as a function of this parameter, which also marks a crossover between IGa limiting behaviors. This opens up an interesting possibility that depending on the magnitude of $\alpha$, increase in the task difficulty may either increase or decreases the magnitude of the power law exponent. This subject, including sufficient data to analyze the aforementioned scaling property, requires further investigation.

V. ACKNOWLEDGMENTS

J.G. Holden’s work was supported by the National Science Foundation Award BCS-0642718.

We repeat a number of Appendices from [8] verbatim, given that the expected audiences for these two papers are vastly different.
FIG. 5: GIGa fitting of one, two, four word LDT. One word LDT: GIGa(0.754, 357, 3.96) with $\alpha \gamma = 3.0$. Two word LDT: GIGa(1.96, 1424, 2.39) with $\alpha \gamma = 4.7$. Four word LDT: GIGa(25.1, 7.37 x 10^6, 0.376) with $\alpha \gamma = 9.4$. The p-values are 0.97, 0.82, and 0.87 respectively.

whose mode is $(\alpha - 1)/2$ and the magnitude of the maximum is $4 \exp[-(\alpha - 1)^2/2]/(\alpha - 1) \approx 4/(\alpha - 1)$, from the regime of large $\alpha$, $\alpha \to \infty$, where

$$P_{\text{Scaled}}(x) \to \delta(x - 1).$$  \hspace{1cm} (A5)

We now turn to GIGa distribution and the effect of parameter $\gamma$. In Fig. 11 we give the contour plots of modal PDF and total half-widths in the $(\eta, \gamma)$ plane, where $\eta = \alpha \gamma$ and $-1 - \eta$ is the exponent of the power law tail. We observe an interesting scaling property of GIGa: for $\gamma \approx 2.1/\eta$, the dependence of the PDF on $\eta$ is very weak, as demonstrated in Fig. 12 where it is plotted for integer $\eta$ from 2 to 7. An alternative way to illustrate this is to plot PDF for a fixed $\eta$ and variable $\gamma$, as shown

FIG. 6: GIGa fitting of Hick’s experiment. The parameters $\{\alpha, \beta, \gamma\}$ of GIGa are \{0.731, 115, 3.41\}, \{1.57, 275, 2.48\}, \{1.64, 430, 3.07\}, and \{7.80, 2922, 1.10\} respectively. $\alpha \gamma$ are 2.5, 3.9, 5.0, and 8.6 respectively. The p-values are all 0.
Appendix B: Parametrization of the GIGa family of distributions

This Appendix is a self-contained re-derivation of a LN limit of GIGa. The three-parameter GIGa distribution is given by

$$\text{GIGa}(x; \alpha, \beta, \gamma) = \frac{\gamma}{\beta \Gamma(\alpha)} e^{-\frac{\beta}{x}} \left( \frac{\beta}{x} \right)^{1+\alpha \gamma}$$ \quad (B1)

for $x > 0$ and 0 otherwise. We require that $\alpha, \beta, \gamma > 0$. IGa is the the $\gamma = 1$ case of GIGa:

$$\text{IGa}(x; \alpha, \beta) = \frac{1}{\beta \Gamma(\alpha)} e^{-\frac{\beta}{x}} \left( \frac{\beta}{x} \right)^{1+\alpha}$$ \quad (B2)

Note that GIGa and IGa have power-law tails $x^{-1-\alpha \gamma}$ and $x^{-1-\alpha}$ respectively for $x \gg \beta$.

We proceed to rewrite GIGa in the following form:

$$\text{GIGa}(x; \alpha, \beta, \gamma) = \frac{\gamma}{x \Gamma(\alpha)} \exp \left[ \alpha \ln \left( \frac{x}{\beta} \right)^{-\gamma} - \left( \frac{x}{\beta} \right)^{-\gamma} \right].$$ \quad (B3)

A re-parameterization

$$\mu = \ln \beta - \frac{1}{\gamma} \ln \frac{1}{\lambda^2}$$ \quad (B4)

$$\sigma = \frac{1}{\gamma \sqrt{\alpha}}$$ \quad (B5)

$$\lambda = \frac{1}{\sqrt{\alpha}},$$ \quad (B6)

with $\sigma > 0$ and $\lambda > 0$, allows to express the old parameters in terms of the new:

$$\alpha = \frac{1}{\lambda^2}$$ \quad (B7)

$$\beta = e^\mu \lambda^{-2\sigma}$$ \quad (B8)

$$\gamma = \frac{\lambda}{\sigma}$$ \quad (B9)

leading, in turn, to

$$\left( \frac{x}{\beta} \right)^{-\gamma} = e^{-\frac{1}{\sigma}(\ln x - \mu)} \lambda^{-2}$$ \quad (B10)

$$\ln \left( \frac{x}{\beta} \right)^{-\gamma} = -\frac{\lambda}{\sigma}(\ln x - \mu) + \ln(\lambda^{-2})$$ \quad (B11)
FIG. 11: Top: contours of modal PDF of GIGa distributions with mean 1. Thin lines: contours of modal PDF. Thick line: \( \gamma = 2.1/\eta \). Bottom: contours of total half-widths of GIGa distributions with mean 1. Thick line: \( \gamma = 2.5/\eta \).

FIG. 12: Scaled PDF of GIGa distributions with mean 1. In the plots, \( \gamma = 2.1/\eta \). Six lines correspond to \( \eta = 2, 3, ..., 7 \), and

\[
\alpha \log \left( \frac{x}{\beta} \right)^{-\gamma} = \left( \frac{x}{\beta} \right)^{-\gamma} \approx \frac{\ln(\lambda^{-2}) - 1}{\lambda^2} - \frac{(\ln x - \mu)^2}{2\sigma^2},
\]

where we have used the Taylor expansion of the \( \exp \) term in Eq. (B10), which depends on \( \lambda/\sigma = \gamma \to 0^+ \). We can also prove that

\[
\frac{\gamma}{\Gamma(\alpha)} \exp \left[ \frac{\ln(\lambda^{-2}) - 1}{\lambda^2} \right] = \frac{1}{\sqrt{2\pi\sigma}},
\]

based on the Stirling’s approximation when we let \( \lambda^{-2} = \alpha \to +\infty \).

Upon substitution of Eqs. (B12) and (B13) into Eq. (B1), we obtain the LN distribution

\[
\text{LN}(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma x}} \exp \left[ -\frac{(\ln x - \mu)^2}{2\sigma^2} \right].
\]

In conclusion, GIGa has the limit of LN when \( \lambda \) tends to 0 in such a way that \( \alpha \) tends to \( +\infty \) quadratically and \( \gamma \) tends to 0 linearly.

GIGa (IGa) are also transparently related to GGa (Ga) distribution: \( \text{GGa}(x; \alpha, \beta, \gamma) \leftrightarrow \text{GIGa}(x; \alpha, \beta, \gamma) \) and \( \text{GGa}(x; \alpha, \beta, \gamma) \leftrightarrow \text{GGa}(1/x; \alpha, 1/\beta, \gamma) \). Note, finally, that Lawless [14] derived the LN limit of GGa in a manner similar to ours, which solidifies the concept of the “family” that unites these distributions.

Appendix C: Stochastic “birth-death” model

Many natural and social phenomena fall into a stochastic “birth-death” model, described by the equation

\[
dx = c_1 x^{1-\gamma} dt - c_2 x dt + \sigma x dW
\]
where $x$ can alternatively stand for such additive quantities as wealth, mass of a species, volatility variance, etc., and cognitive response times here.

The second term in the rhs describes an exponentially fast decay, such as the loss of wealth and mass due to the use of one’s own resources, or the reduction of volatility in the absence of competing inputs and of response times due to learning. The first rhs term may alternatively describe metabolic consumption, acquisition of wealth in economic exchange, plethora of market signals, and variability of cognitive inputs.

The third, stochastic term is the one that changes the otherwise deterministic approach, characterized by the saturation to a final value of the quantity, with the probabilistic distribution of the values - as it were, GIGa in the steady-state limit. Furthermore, just as the wealth model has microscopic underpinnings in a network model of economic exchange, it is likely that stochastic ontogenetic mass growth could be described by analogous network model based on capillary exchange. A network analogy may be possible for cognitive response times and volatility as well.

**Appendix D: Log-log plot of distribution tails**

The exponent of a power law tail can be easily calculated once we notice that

$$1 - \text{CDF}(x) = \int_{x}^{\infty} \text{PDF}(x) dx.$$ \hspace{1cm} (D1)

If $\text{PDF}(x) \propto C x^{-\rho}$ with $x \gg 1$, then

$$\log(1 - \text{CDF}(x)) \propto \text{const} - \rho \log x.$$ \hspace{1cm} (D2)

In Figs. 14 and 15 we show the log-log plot of the tail of LN and IGa distributions respectively. Clearly, a straight line fit is considerably better for the latter, even though the fitted slope does not coincide with the theoretical value. Towards this end, in Fig. 16 we show log-log plots of simulated data sampled from the IGa distributions. Below $-1$ of the y-axis, the left blue, middle green, and right red curves correspond to $\sigma = 0.2, 0.5,$ and 1 respectively. The dashed lines are fitting of $\log_{10}(1 - \text{CDF}(x))$ vs. $\log_{10} x$ in a range of CDF from 0.9 to 0.99.

To understand this $\gamma$-dependence the difference between the theoretical and fitted slope, we consider the local slope of the log-log plot,

$$\frac{d \log(1 - \text{CDF}(x))}{d \log x}.$$ \hspace{1cm} (D3)

For GIGa (and IGa, $\gamma = 1$), the local slope is given by

$$\gamma e^{-\left(\frac{x}{\beta}\right)^\gamma} \left(\frac{\beta}{\tau}\right)^{\alpha \gamma} \frac{\Gamma(\alpha)}{\Gamma(\gamma)} \left(Q(\alpha, \left(\frac{x}{\beta}\right)^\gamma) - 1\right),$$ \hspace{1cm} (D4)

with the regularized gamma function $Q(s, x) = \Gamma(s, x)/\Gamma(s)$, where $\Gamma(s, x) = \int_x^\infty t^{s-1} e^{-t} dt$ is the incomplete gamma function. The local slopes are shown, as

![Log-log plots of distribution tails](image)

**FIG. 14:** Top: plots of PDF of LN($x; \mu, \sigma$) with mean 1. The left red, middle green, and right blue curves correspond to parameter $\sigma = 1, 0.5,$ and 0.2 respectively. Bottom: log-log plots of simulated data sampled from the LN distributions. Below $-1$ of the y-axis, the left blue, middle green, and right red curves correspond to $\sigma = 0.2, 0.5,$ and 1 respectively. The dashed lines are fitting of $\log_{10}(1 - \text{CDF}(x))$ vs. $\log_{10} x$ in a range of CDF from 0.9 to 0.99.

**FIG. 15:** Top: plots of PDF of IGa($x; \alpha, \beta$) with mean 1. The left red, middle green, and right blue curves correspond to $\alpha = 3, 4,$ and 5 respectively. Bottom: log-log plots of simulated data sampled from the IGa distributions. Below $-1$ of the y-axis, the left blue, middle green, and right red curves correspond to $\alpha = 5, 4,$ and 5 respectively. The dashed lines with slopes $-3.5, -3.0,$ and $-2.5$ respectively are fitting of $\log_{10}(1 - \text{CDF}(x))$ vs. $\log_{10} x$ in a range of CDF from 0.9 to 0.99.
function of $x$ in Figs. [17] and [18] respectively. It is clear that the local slope can differ substantially from its limiting (saturation) value. As $\gamma$ becomes larger, the local slope tends closer to its limiting value.

For the LN distribution, the local slope is given by

$$\sqrt{\frac{2}{\pi}} e^{-\frac{(\log x - \mu)^2}{2\sigma^2}}$$

$$\sigma \left( 1 + \text{erf} \left( -\frac{\log x - \mu}{\sqrt{2}\sigma} \right) \right).$$

which slowly decreases with $x$. But as is clear from (D5) and Fig. [19] the local slope does not saturate when $x \to \infty$.

[1] J. G. Holden and S. Rajaraman, Frontiers in Psychology, 3, 209 (2012).
[2] G. C. Van Orden, J. G. Holden, and M. T. Turvey, Journal of Experimental Psychology: General 132, 331 (2003).
[3] G. C. Van Orden, J. G. Holden, and M. T. Turvey, Journal of Experimental Psychology: General 134, 117 (2005).
[4] C. T. Kello, G. G. Anderson, J. G. Holden, and G. C. Van Orden, Cognitive Science 32, 1217 (2010).
[5] E.-J. Wagenmakers, S. Farrell, and R. Ratcliff, Psychonomic Bulletin and Review 11, 579 (2004).
[6] T. Ma, J. G. Holden, and R. A. Serota, Physica A: Statistical Mechanics and its Applications 392, 2434 (2013).
[7] D. West and B. West, Int. J. Mod. Phys. B 26, 1230010 (2012).
[8] T. Ma and R. Serota, arXiv:1305.4173 (2013).
[9] D. A. Balota, M. J. Yap, K. A. Hutchison, M. J. Cortese, B. Kessler, B. Loftis, J. H. Neely, D. L. Nelson, G. B. Simpson, and R. Treiman, Behavior Research Methods 39, 445 (2007).
[10] http://elexicon.wustl.edu/ (2009).
[11] W. E. Hick, Quarterly Journal of Experimental Psychology, 4:1, 11 (1952).
[12] T. Van Zandt, Psychological Bulletin & Review, 7(3), 424 (2000).
[13] http://www.weibull.com/hotwire/issue15/hottopics15.htm (2002).
[14] J. F. Lawless, Statistical Models and Methods for Lifetime Data, (John Wiley & Sons, New York, 1982).
FIG. 17: Local slope of log-log plot of IGa distribution IGa(x; \( \alpha, \beta \)) with mean 1 (\( \beta = \Gamma(\alpha)/\Gamma(\alpha - 1) \)). The left column is the log-log plot and the right one is the local slope of the log-log plot from Eq. (D4). \( \alpha \) is 2, 3, 5, and 7 for the first, second, third, and fourth rows respectively. The red lines are \(-\alpha\): the limit of the local slope when \( x \to \infty \).
FIG. 18: Local slope of log-log plot of GIGa($x; \alpha, \beta, \gamma$) with mean 1 ($\beta = \Gamma(\alpha)/\Gamma(\alpha - 1/\gamma)$). The left column is the log-log plot and the right one is the local slope of the log-log plot from Eq. (D4). $\{\alpha, \gamma\}$ is $\{6, 0.5\}$, $\{10, 0.5\}$, $\{1.5, 2\}$, and $\{2.5, 2\}$ for the first, second, third, and fourth rows respectively. The red lines are $-\alpha\gamma$: the limit of the local slope when $x \to \infty$. 
FIG. 19: Local slope of log-log plot of lognormal distribution. The mean of the distribution is set as 1 through $\mu = -\sigma^2/2$. The left column is the log-log plot and the right one is the local slope of the log-log plot in Eq. (D5). $\sigma$ is 0.2, 0.5, 1, and 2 for the first, second, third, and fourth row respectively. The jagged part of the top right plot is due to computational precision.