RT分布のex-Gaussianによる分析
―データ解析演習―

教育学研究科教育科学専攻
教育認知心理学講座
廣井隆志
hiroi.takashi.52u@st.kyoto-u.ac.jp

2013/6/12 発表

X. 実習の準備

- Rを用いて行なうので、Rを起動する
- ex-Gaussian分布に関する関数を使用するため、”exGauss.R”を読み込む
  - ”exGauss.R”は「反応時間解析の理論と応用」から利用できる（http://www7b.biglobe.ne.jp/~homunculus/neuro/RTanalysis/RTanalysis.html）

“exGauss.R”の読み込み方

1．ファイル”exGauss.R”があるディレクトリに変更

2．>source(“exGauss.R”)　　#打ち込む

または、「ファイル」->「Rコードのソースを読み込み」からファイル”exGauss.R”を選択する。

0.発表概要

- 平均値による分析は、操作（要因）の効果を誤って解釈する可能性がある

- ex-Gaussian分布はこの問題を解決する

- ex-Gaussian分布を理解して、そのパラメータを求める（推定する）練習をする

1. We all love the mean!

- Balota & Yap (2011)
  - 3つの一流雑誌から2010年に285の論文が出版された。
  - そのうち、49%が反応時間(RT)を尺度として用いた。
  - 重要なことだが、RTを従属変数として用いた論文のうち、95%が主に平均RTに頼っていた。

- **「平均」を用いて何か問題があるのか?**

1.1.「平均」の問題

その１　ある操作（要因）の効果が生じた場合、どのような分布の変化によって生じたものかがわからない

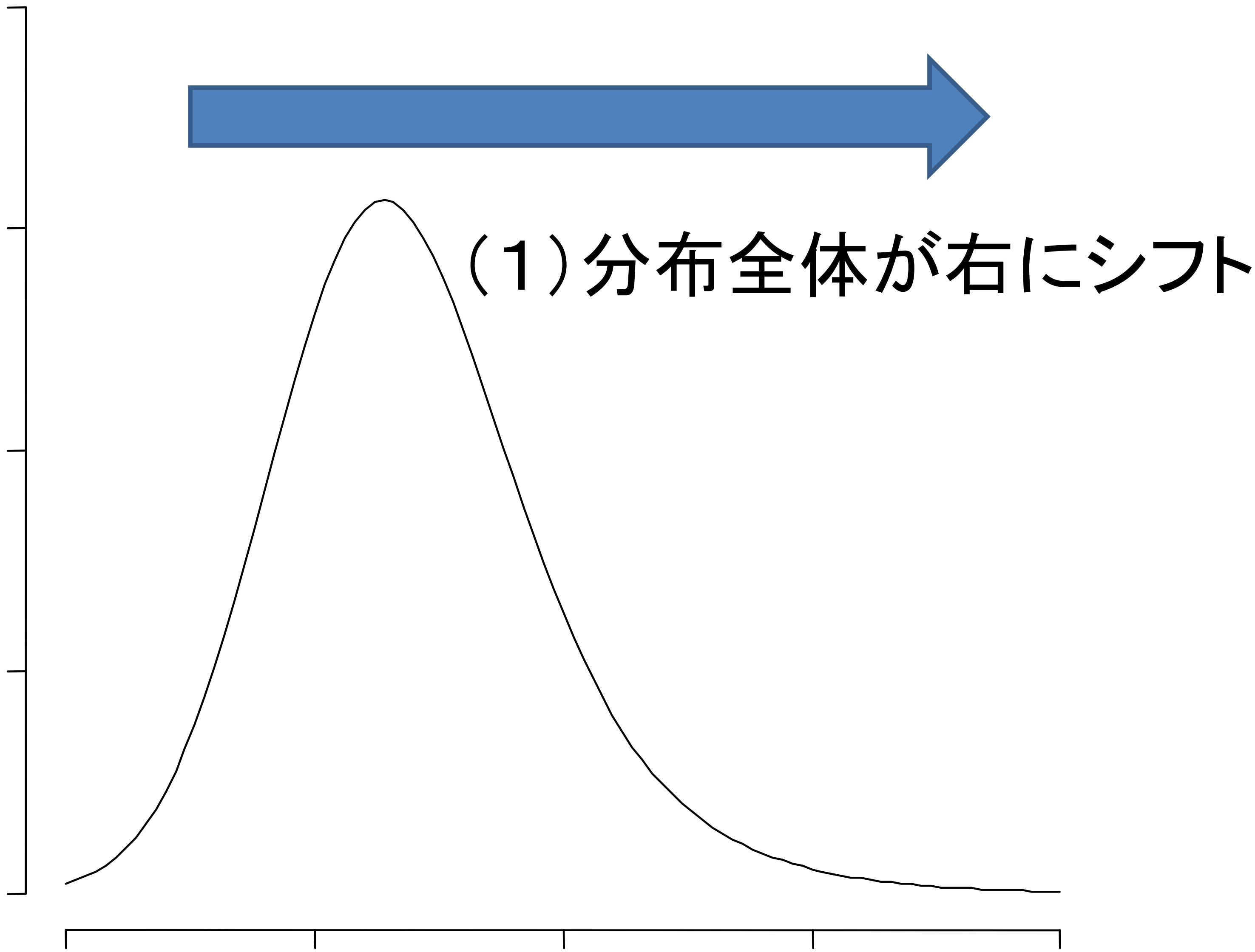
例　初婚年齢が10年ごとに平均で〇〇歳遅くなっている。

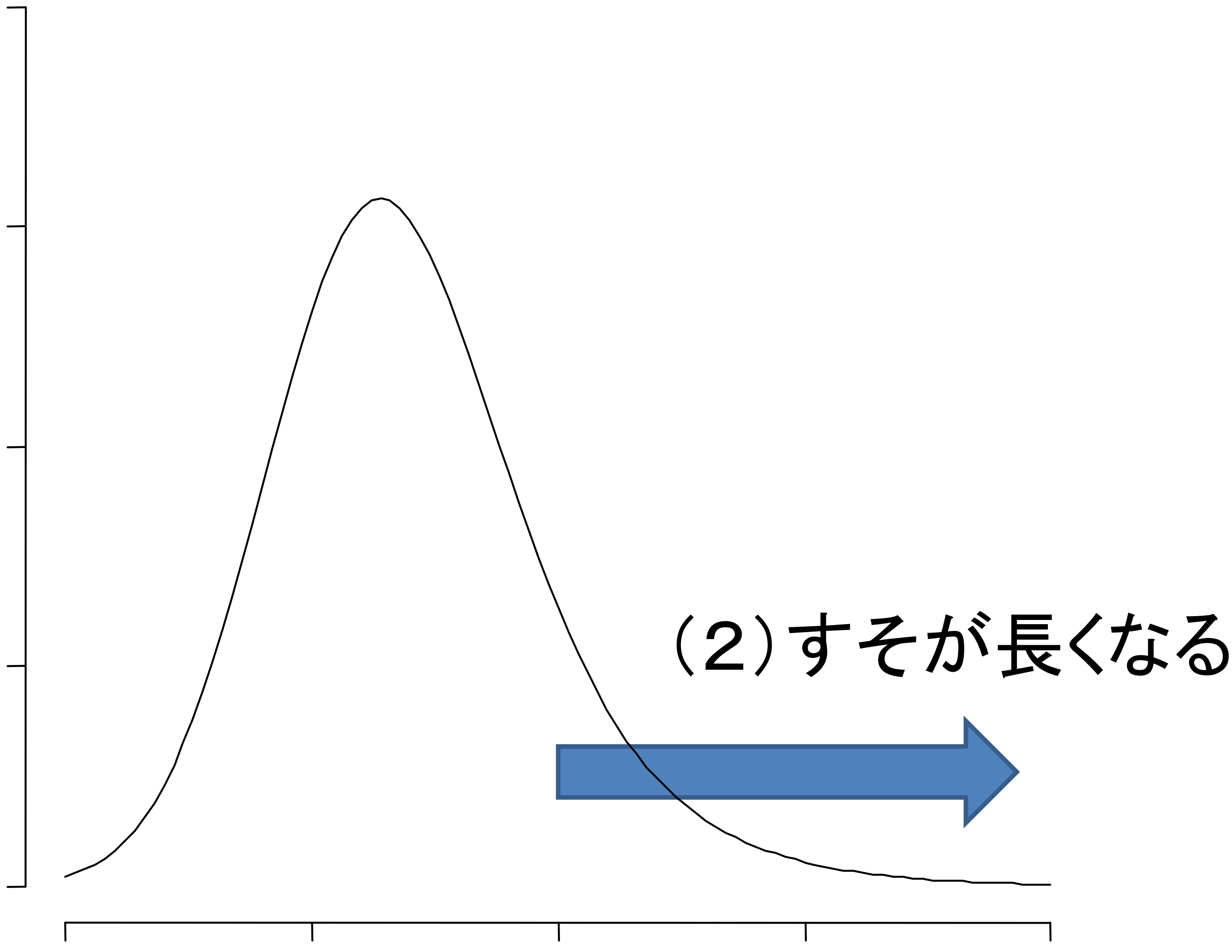
– （１）分布全体が右にシフトした
– （２）右の尾（すそ）が長くなった
– （３）その両方が生じている

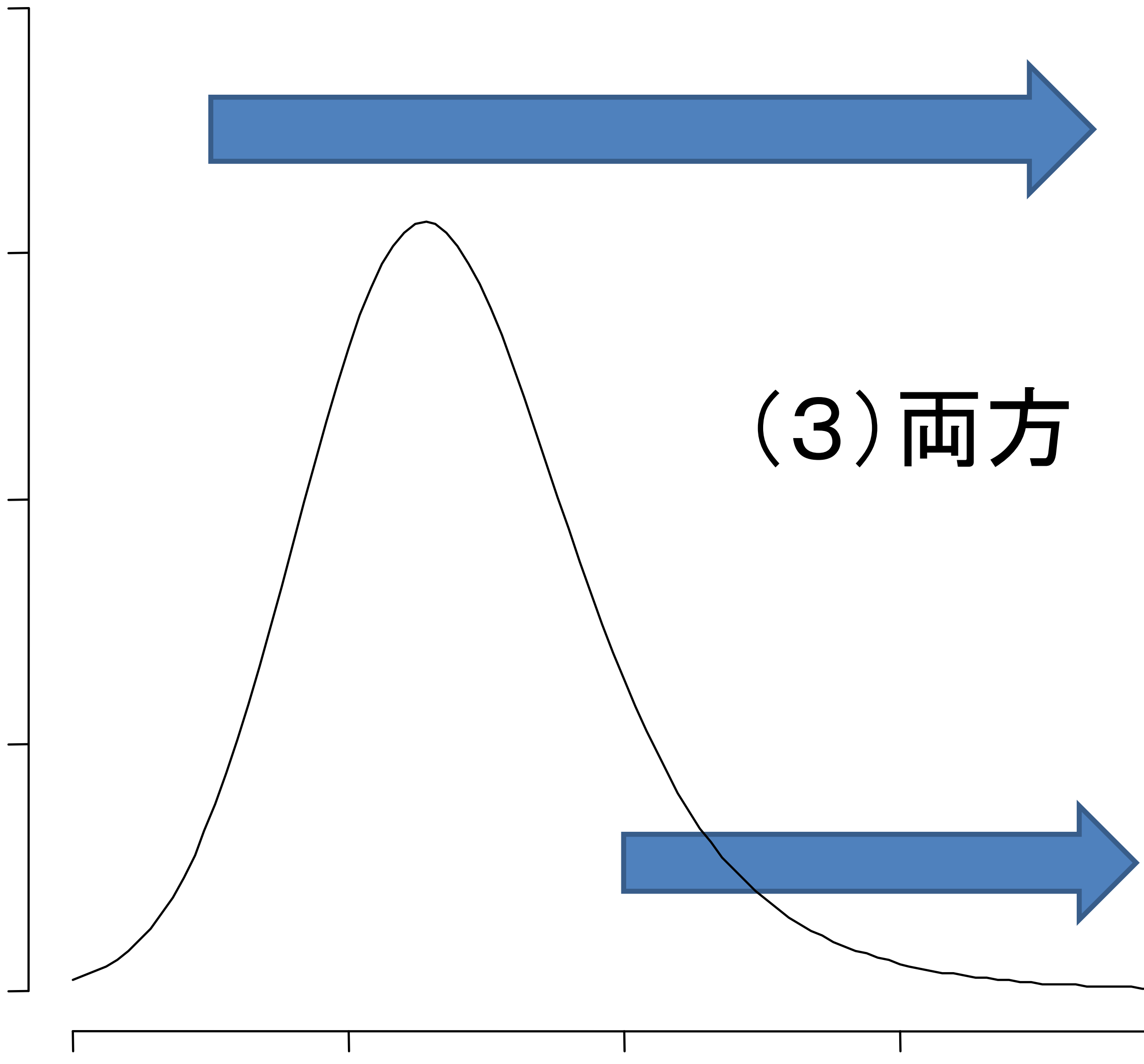
平均の変化では、この3つのケースのどれが生じているかはわからない

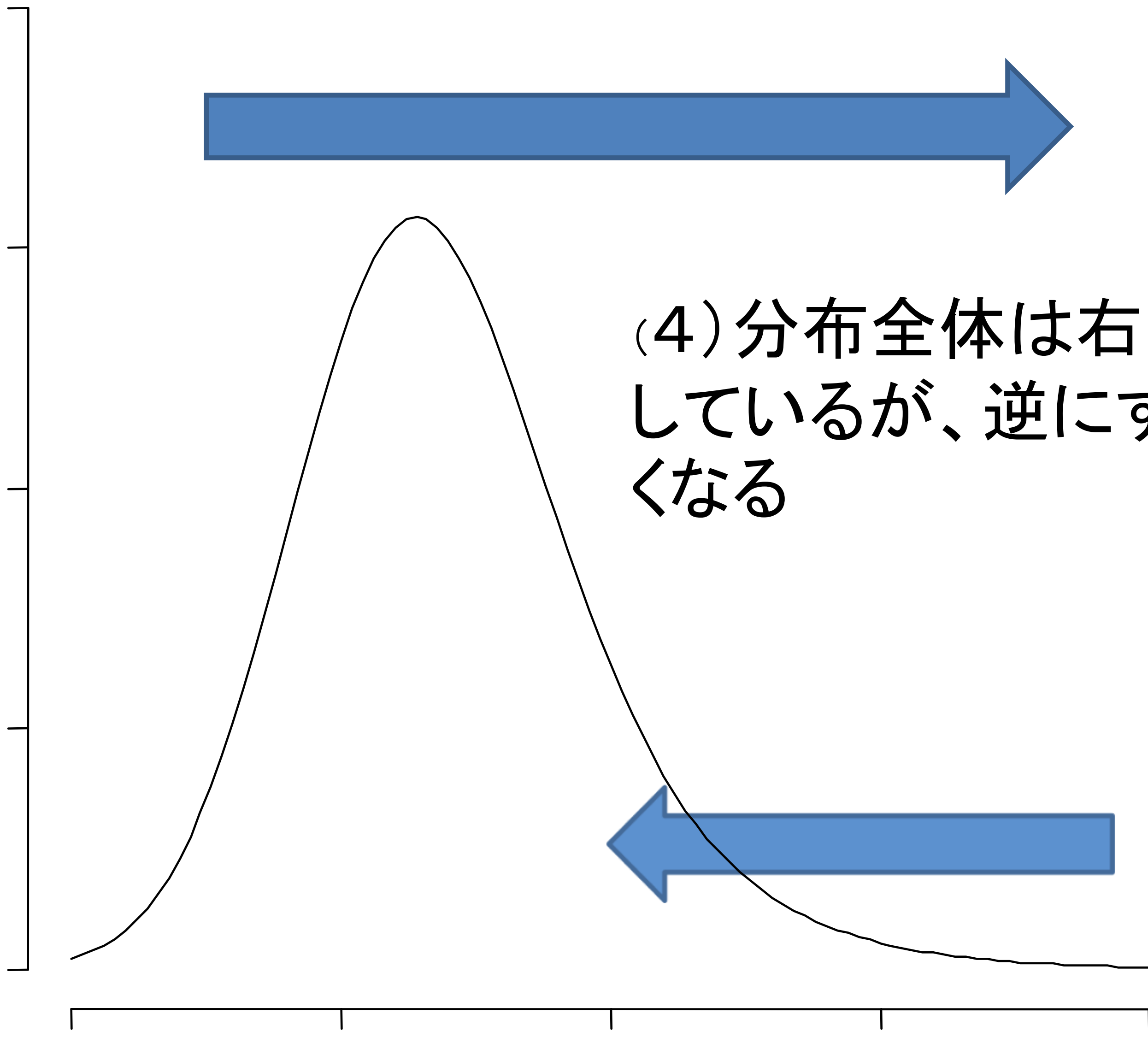
その２　ある操作（要因）の効果が平均値の差では生じていない場合でも、分布の形状には影響しているかもしれない

– **（４）分布全体は右にシフトしているが、右の尾（すそ）が短くなり、平均ではトレードオフされた**


（1）分布全体が右にシフト


（2）すそが長くなる


（3）両方



(4）分布全体は右にシフトしているが、逆にすそは短くなる

1.1.「平均」の問題

- ある操作の効果（平均の差分）が生じた場合、どのような分布の変換によって生じたものなのかがわからない

例　初婚年齢が10年ごとに平均で〇〇歳遅くなっている。
  – （1）分布全体が右にシフトした
  – （2）右の尾（すそ）が長くなった
  – （3）その両方が生じている

- ある操作の効果（平均の差分）が生じていない場合でも、分布には影響しているかもしれない
  – 例（4）分布全体は右にシフトしているが、右の尾（すそ）が短くなり、平均ではトレードオフされた

→**これらの問題を解決するには、ある操作（要因）の効果が分布（の形状）にどのような影響を与えるのかを分析することが必要である。**

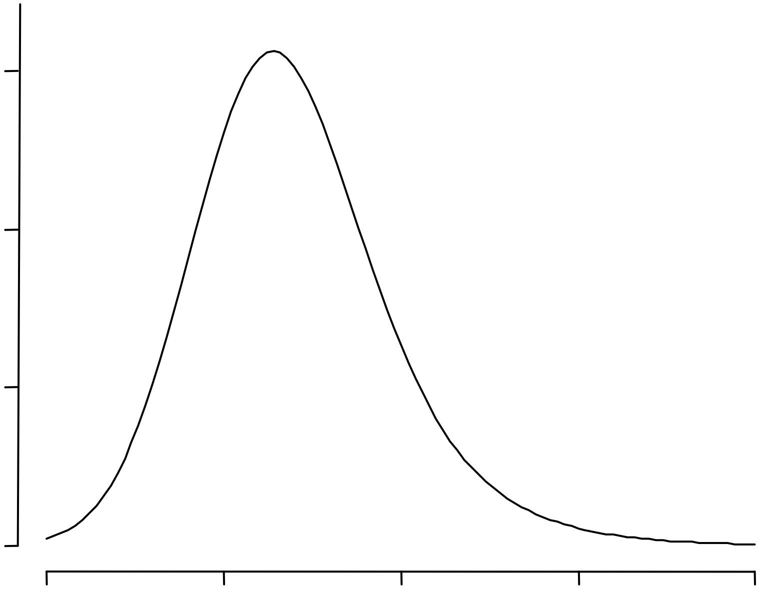
ex-Gaussian分布を用いた分析はその一つの方法である

2. ex-Gaussian分布とは

2. ex-Gaussian分布とは

- 正規分布、二項分布、ポワソン分布などと同じ確率分布

- 正規分布と指数分布から「作られる」分布（→2.2　実習）



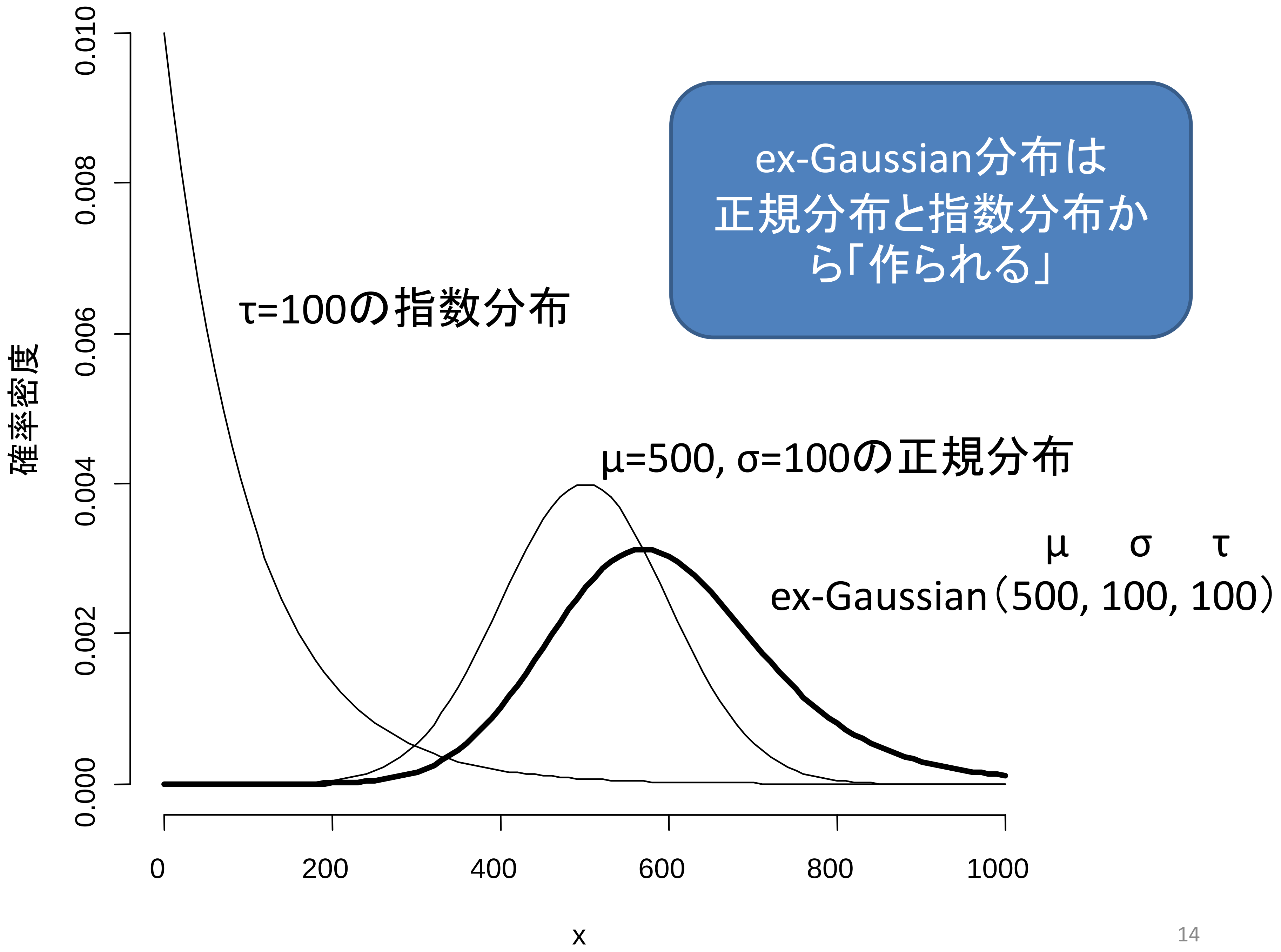
2. ex-Gaussian分布とは

正規分布は平均μと標準偏差σというパラメータで特徴づけられる

指数分布は平均τ、標準偏差τというパラメータで特徴づけられる

→ex-Gaussian分布は上記2つの分布から「作られる」ため、
ex-Gaussian分布は3つのパラメータ（μ , σ , τ）で特徴づけられる

ちなみに、 この分布の平均は μ＋τ となる


ex-Gaussian分布は
正規分布と指数分布から「作られる」
τ=100の指数分布
μ=500, σ=100の正規分布
μ σ τ
ex-Gaussian（500, 100, 100）
確率密度
0.000
0.002
0.004
0.006
0.008
0.010
0
200
400
600
800
1000
x

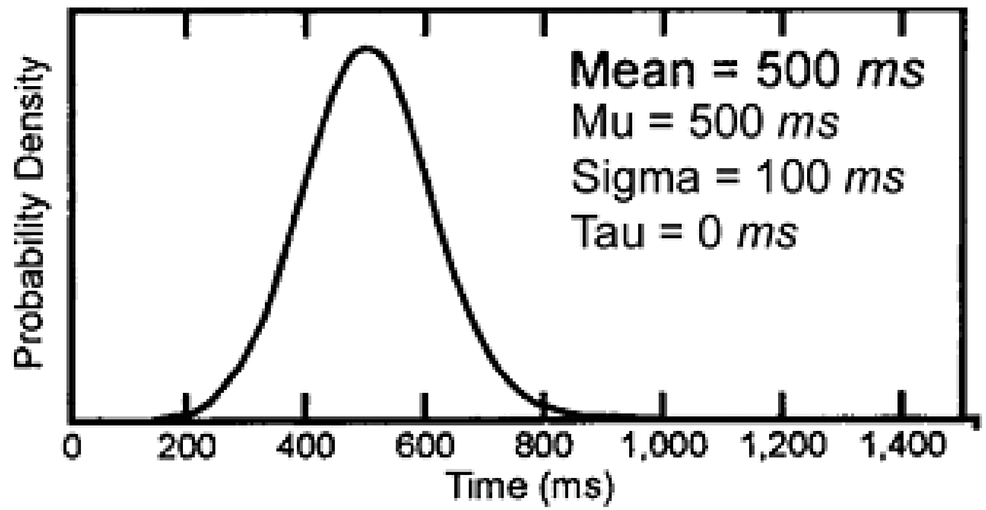
2. ex-Gaussian分布とは

- 「理由は分からないが、ex-Gaussianは反応時間によくフィットする」（Luce, 1986,p.439）
  - **手元のRTデータ分布はex-Gaussian分布であると考えれば、手元のRTデータ分布は3つのパラメータで表現できる**

- それぞれのパラメータは分布の特徴と大まかに対応する（2.1. を参照）
  - **操作（要因）のRT分布への効果はパラメータの変化によってとらえることができそうだ**

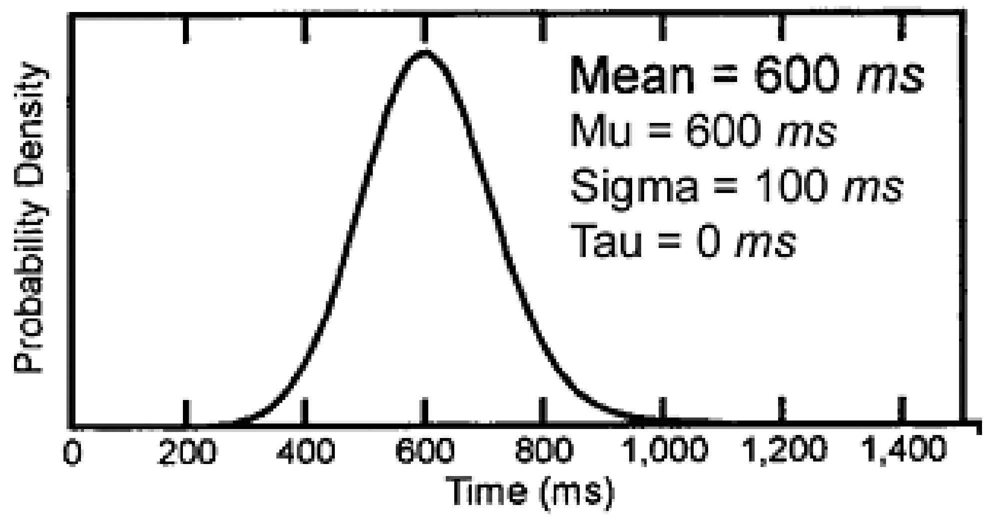
2.1. パラメータの変化と分布の変化の関係

A



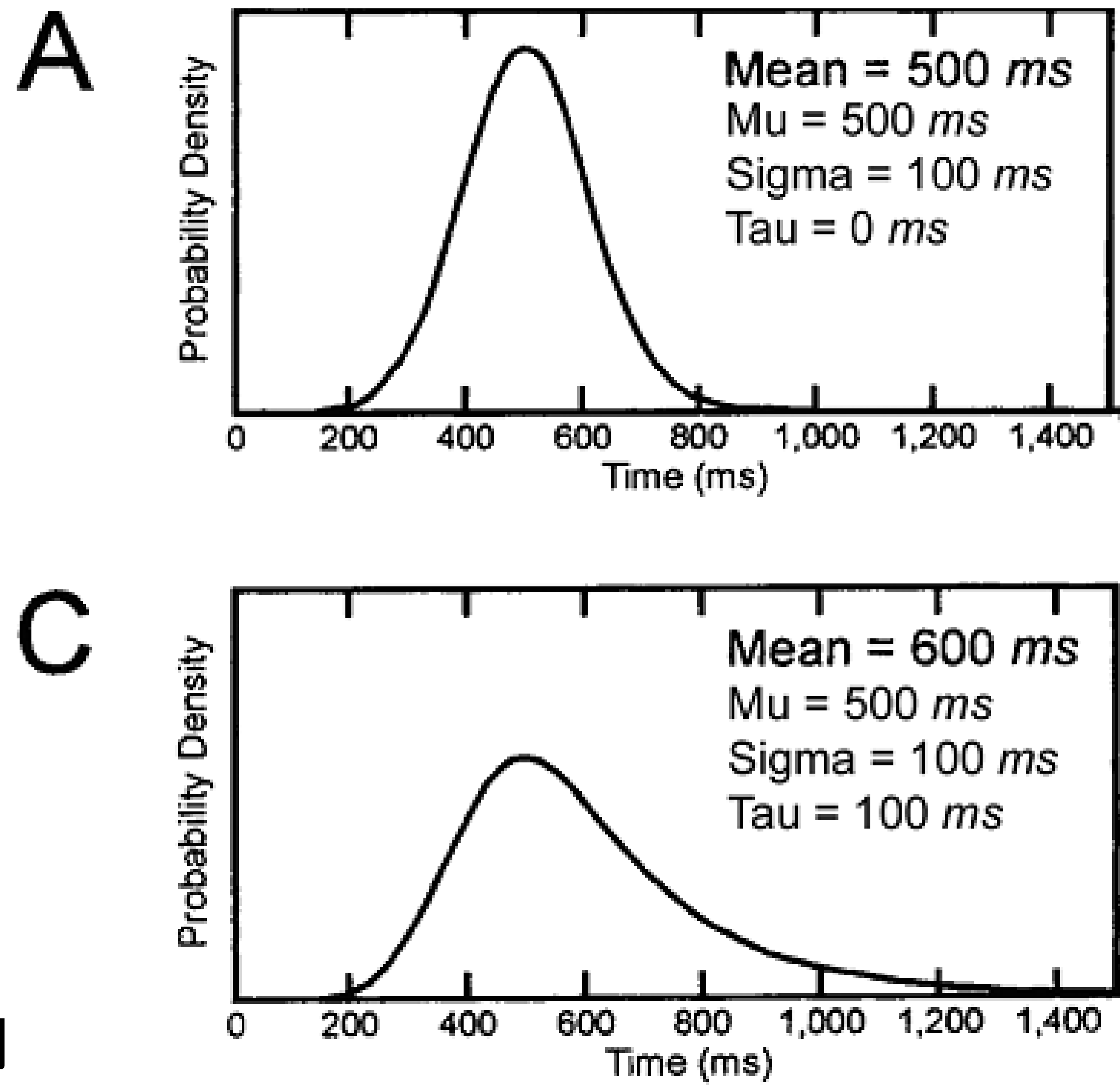
- μが変化
  ⇔分布全体が
  右にシフトする

B



図はBalota & Yap
(2011, p.161)から引用

- τが変化
  ⇔分布のすそ
  　が長くなる

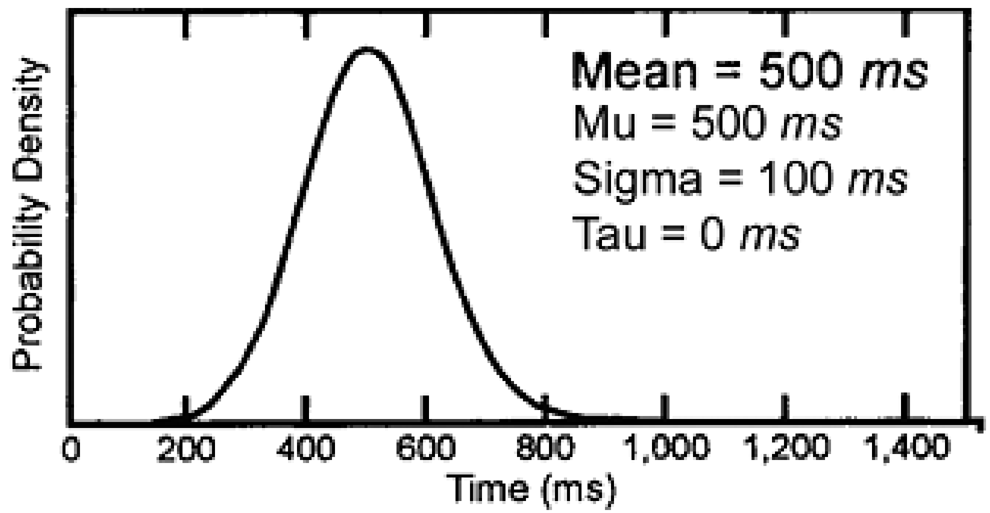


図はBalota & Yap
(2011, p.161)から引用

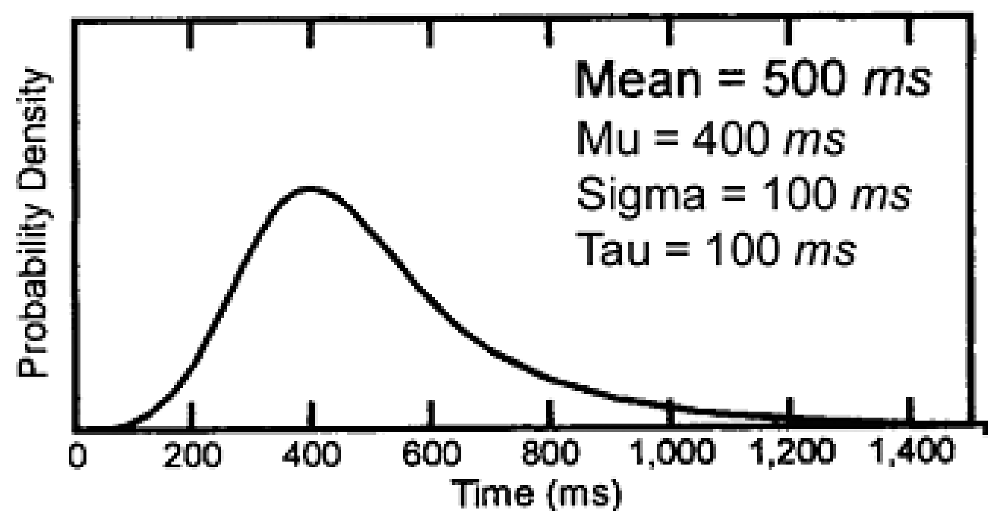
参考

- μとτの変化により平均の変化が相殺されている

A



D



図はBalota & Yap (2011, p.161)から引用

2.2. ex-Gaussian分布とは：実習1

ex-Gaussian分布は正規分布と指数分布から「作られる」とは、もう少し具体的にいえば、次の通り。

↓

平均μ、標準偏差σである正規分布に従う確率変数Xがあり、平均τ、標準偏差τである指数分布に従う確率変数Yがあるとすると、その和（X＋Y）が従う分布がex-Gaussian分布であり、そのパラメータはμ、σ、τであるということ。

（注意）ただし、XとYは独立。つまり、Xの値が大きくても小さくても、Yの値がそれに応じて小さい値が出やすくなったり、大きな値が出やすくなったりはしないということ。

2.2. ex-Gaussian分布とは：実習1

前述の内容を確かめてみる。

手順

（1）平均500、標準偏差100である正規分布から3000個の乱数を作る

（2）平均100、標準偏差100である指数分布から3000個の乱数を作る

（3）その和がつくる分布は、ex-Gaussian分布になっており、そのパラメータはμ=500、σ=100、τ=100であるはずだ！

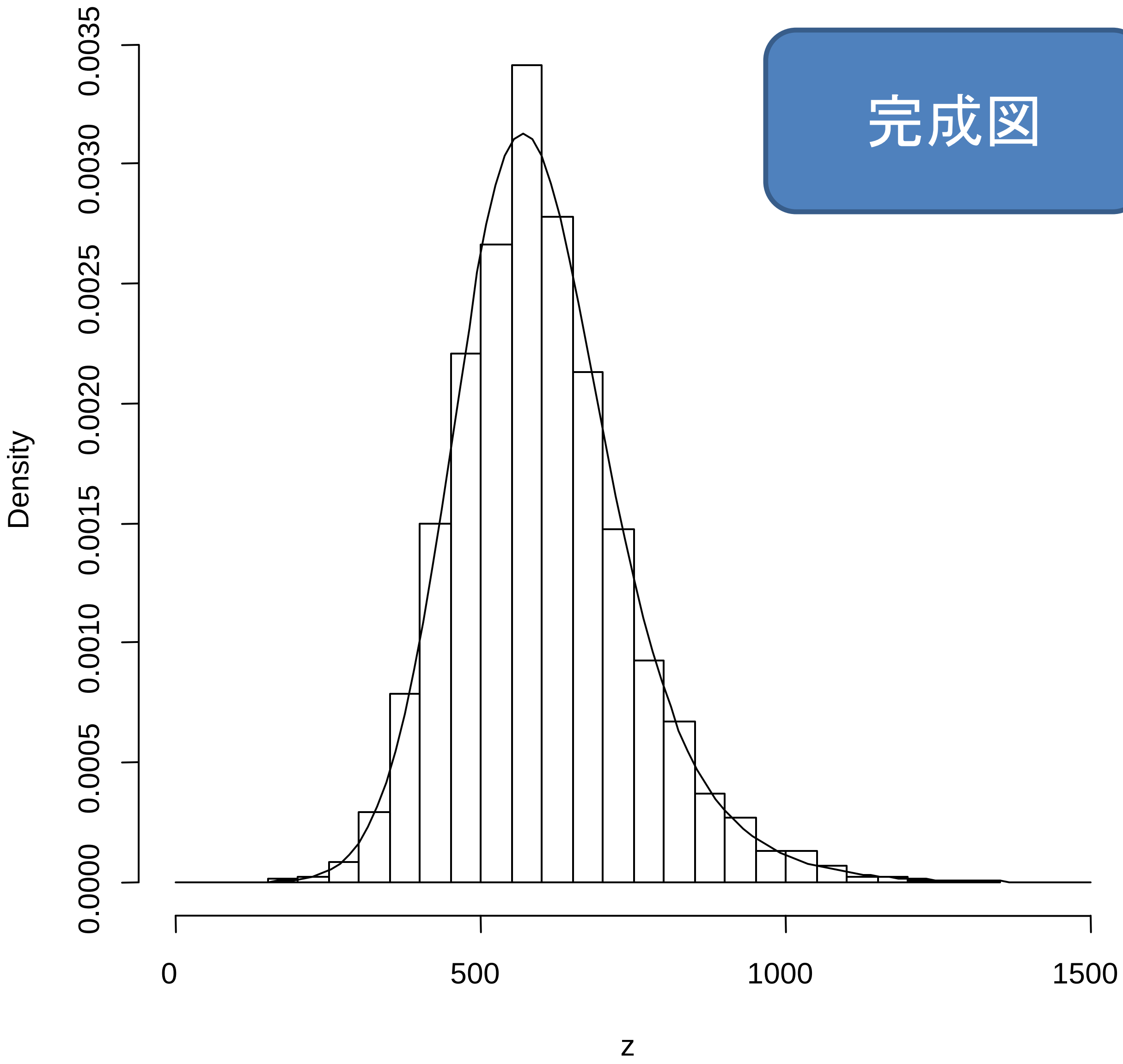
2.2. ex-Gaussian分布とは：実習1

```
x<- rnorm(3000, 500, 100)      #正規分布の乱数を作る
y<-rexp(3000, 1/100)           #指数分布の乱数を作る
                                （注意：Rでは逆数を入れる）

z<- x+y                        #xとyの和をzとする

hist(z, breaks=20, xlim=c(0,1500), prob=T)    #ヒストグラムを作る
curve(dexGauss(x, 500, 100, 100), add=T)      #ex-G分布の曲線
```

＊関数や引数については船尾（2009）などを参考にしてください。


Density
0.0000
0.0005
0.0010
0.0015
0.0020
0.0025
0.0030
0.0035
0
500
1000
1500
z

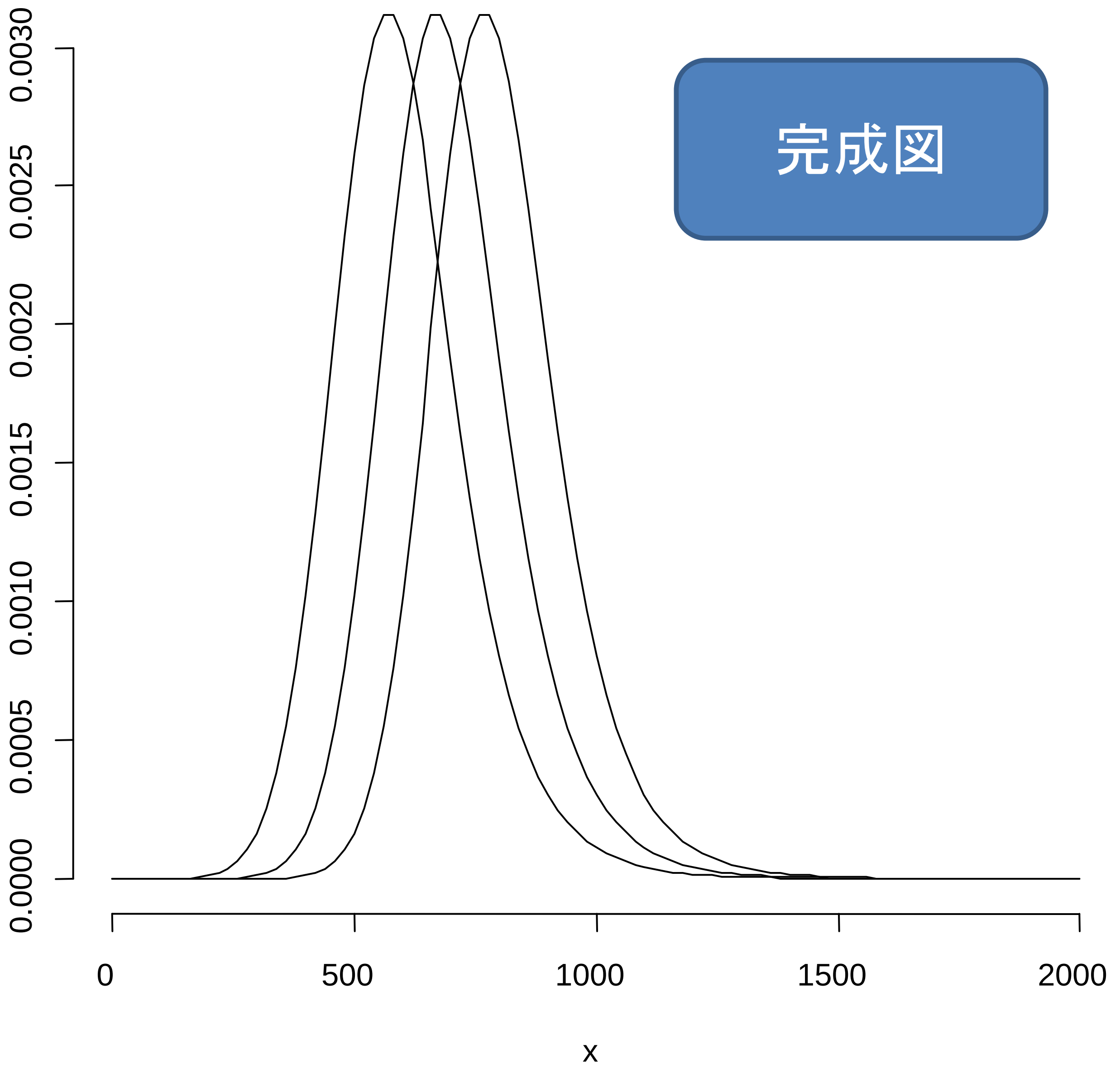
完成図

2.2. ex-Gaussian分布とは：実習2

- μを変化させると分布がどう変化するか、τを変化させると分布がどう変化するかを「図を重ねて」確かめる

- （1）μのみを変化させる（500→600→700）

```
curve(dexGauss(x,500, 100, 100), xlim=c(0,2000))
curve(dexGauss(x,600, 100, 100), xlim=c(0,2000), add=T)
curve(dexGauss(x,700, 100, 100), xlim=c(0,2000), add=T)
```

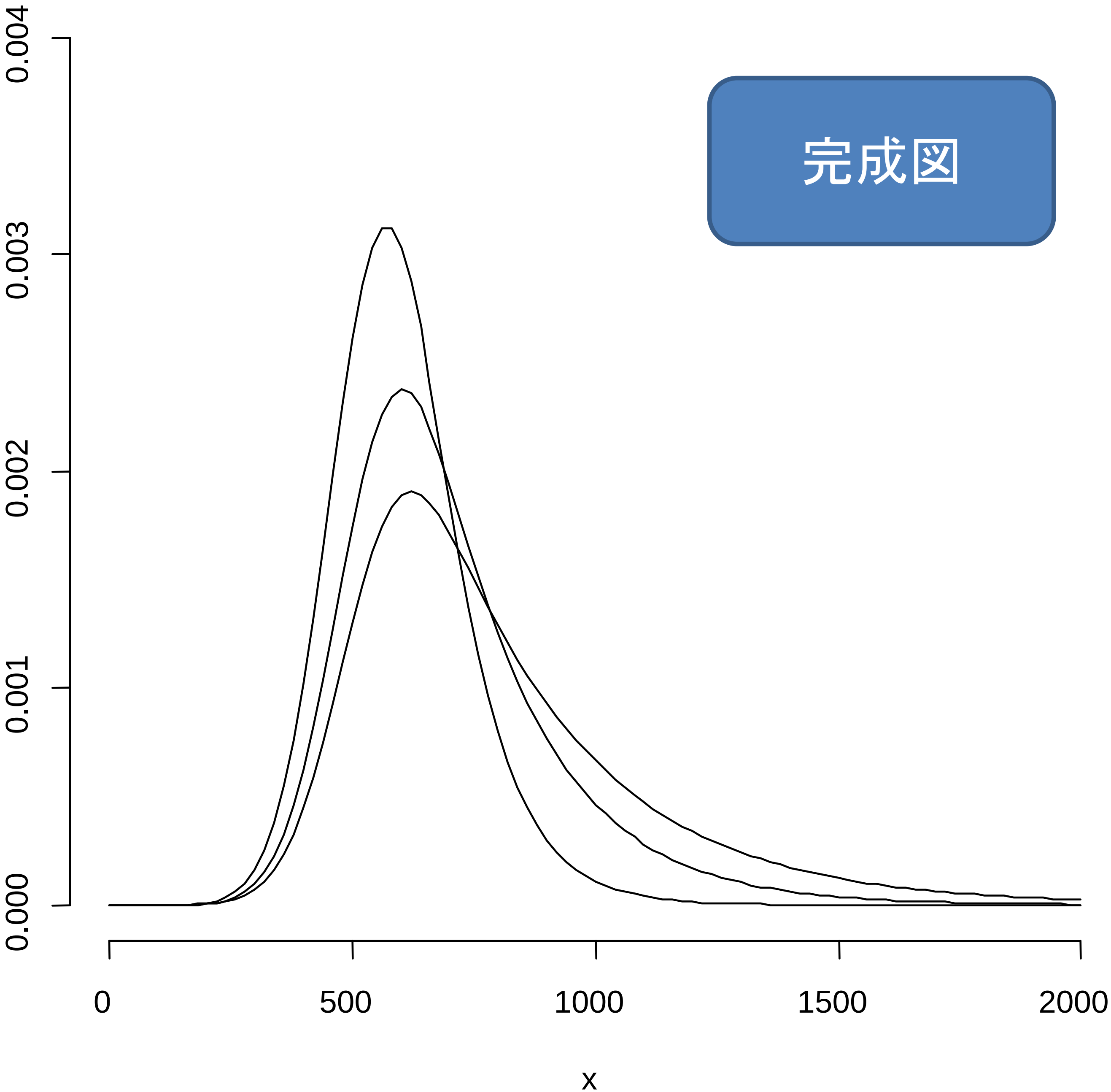


完成図
0.0000
0.0005
0.0010
0.0015
0.0020
0.0025
0.0030
0
500
1000
1500
2000
x

2.2. ex-Gaussian分布とは：実習2

- （２）τのみを変化させる（100→200→300）

```
curve(dexGauss(x,500, 100, 100), xlim=c(0,2000), ylim=c(0,0.004))
curve(dexGauss(x,500, 100, 200), add=T)
curve(dexGauss(x,500, 100, 300), add=T)
```


完成図
0.000
0.001
0.002
0.003
0.004
0
500
1000
1500
2000
x

3. ex-Gaussianのパラメータの推定

- 手元のRTデータがex-Gaussian分布に従っていると考える。次に行わなければならないことは、その従っているex-Gaussian分布のパラメータ（μ, σ, τ）を推定することである。

- ex-Gaussianのパラメータを推定する方法の一つは、最尤法（Most likelihood: ML）である。

- 最尤法の場合、パラメータの安定した推定値には、各条件で100ほどの観測しか必要としない（Heathcote, et al. 1991）
  - （注）Quantile maximum likelihood (QML) はサンプルサイズが40でも、パラメータの頑健な推定値を与えてくれる（Brown & Heathcote, 2003）。上記は標準的な最尤法を用いて推定する場合のことである。

3.1. 最尤法とは

- 母集団に特定の分布(e.g. 正規分布)を仮定したとき、観察された値(標本)が**最も確率的に起こり得る**ようにパラメータを推定する推定法のひとつ
  - cf. 回帰分析で使用される最小2乗法も母偏回帰係数の推定値を得るための推定法のひとつである

＊詳しくは、参考書を読んでください。

3. ex-Gaussianのパラメータの推定

- 最尤法によってパラメータを推定することはわかった
- 具体的には手元のRTデータからどうやってパラメータの推定値を出すのか？

➡ exGauss.Rの中にある関数 fitexGauss(x) のxにRTデータを入れれば、後はコンピュータがやってくれます

3.1. ex-Gaussianのパラメータの推定 :実習3

- ex-Gaussian分布に従う乱数を標本に見立て、その標本から最尤法でパラメータを推定する。

3.2. ex-Gaussianのパラメータの推定 :実習3

- μ=700, σ=100, τ=200であるex-Gaussian分布に従う乱数を100個作成し、それを変数xに入れる

```
x<-rexGauss(100, 700, 100, 200)
```

- fitexGauss(x)関数を用いて推定し、その結果を変数resultに入れる

```
result<-fitexGauss(x)
```

- いろいろ結果が出てくるので、パラメータの推定値のみを取り出す

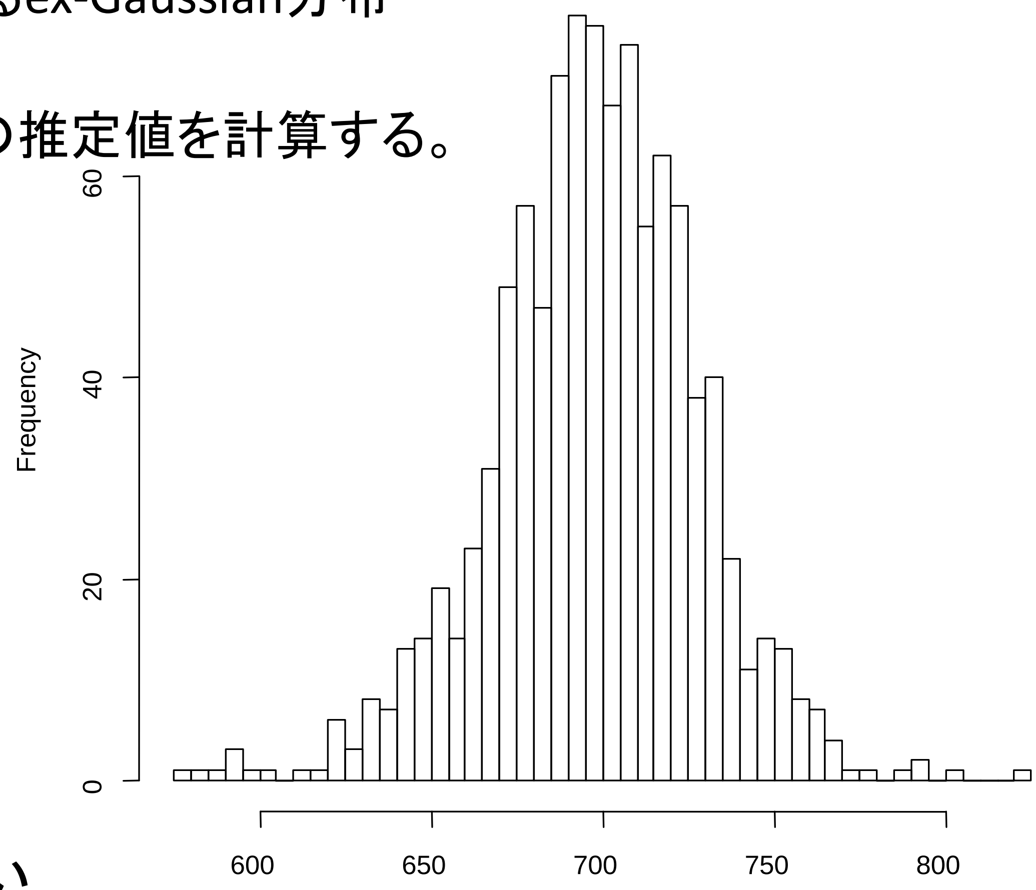
```
result$loglike
```

- 結果は母集団のパラメータ(μ=700, σ=100, τ=200)と近い値になったか？

（参考）μのsampling distribution

μのsampling distributionの作成手順

- （１）μ=700, σ=100, τ=200であるex-Gaussian分布から10000個の乱数を作る。
- （２）そこから100個抽出し、μの推定値を計算する。
- （３）（２）の手続きを1000個繰り返す。



分布の平均

697.95

分布の標準偏差

30.28

（注意）分布、値は近似にすぎない

0. 発表概要

- 平均値による分析は、操作(要因)の効果を誤って解釈する可能性がある
  - とりわけ操作(要因)の効果が平均では見られなくても、分布の形状は変化している(操作(要因)の効果がある)場合がある
- ex-Gaussian分布はこの問題を解決する
  - 3つのパラメータの変化によって、分布の変化がわかる
- ex-Gaussian分布を理解して、そのパラメータを求める(推定する)練習をする
  - すでに関数はある(自分で書くこともできる)ので、すぐにでも実戦投入可能

文献

Balota & Yap (2011) Moving beyond the mean in studies of mental chronometry: the power of response time distributional analyses. Current Directions in Psychological Science, 20, 160-166.

Brown & Heathcote (2003) QMLE: Fast, robust and efficient estimation of distribution functions based on quantiles. Behavior Resarch Methods, Instruments & Computers, 35, 485-492.

Heathcote, Poiel, & Mewhort (1991) Analysis of response time distributions: an example using the stroop task. Psychological Bulletin, 109, 340-347.

文献（続き）

Luce (1986) Response times : Their role in inferring elementary mental organization. New York: Oxford University Press.

反応時間解析の理論と応用（http://www7b.biglobe.ne.jp/~homunculus/neuro/RTanalysis/RTanalysis.html）

今回の実習で使用した“exGauss.R”はこのホームページにあり、今回利用させていただいた。感謝を申し上げます。

舟尾暢男 (2009) The R Tips—データ解析環境Rの基本技・グラフィックス活用集　第2版 オーム社