

DPCデータ分析からみた北里大学4病院の位置づけ

(図11)。これは特定機能病院として三次救急に特化しているためと考えられる。一方でこの地域では南町田病院の総入院患者数に占める救急搬送患者割合が突出して高い。このようにDPCデータから自院のみならず他院の経営方針まで伺い知ることができ、経営戦略を策定するうえでの重要な資料といえる。

4. DPCデータの限界

以上、公開されているDPCデータから地域医療の市場分析を試みたが、DPC分析には注意すべき限界もある。第一に、DPC病院のデータは地域医療の全てを反映しているわけではない。北里医療圏の場合、厚生省の患者調査とつぎ合わせるとDPC病院の患者数はこの医療圏全体の約1/2~1/3をカバーしているに過ぎない。第二に、DPC Codeとして医療資源を最も投入した1病名を選択することから、合併症の多い患者の状況が把握できない。さらにMDC分類と診療科の構成が必ずしも一致していない。例えばMDC06は消化器・肝臓胆嚢膵臓疾患であるが、内科と外科の区別はしていない。第三に、小児はそれぞれのMDCに含まれる症例が多く、MDC15小児疾患系として独立して扱われる症例数は少ないため経営データとして評価しにくい。さらに、効率性指標の算出の際には、対象病院に患者が存在しないDPC分類の平均在院日数については仮定が必要で、全国平均の平均在院日数を仮にあてはめて算出する方法が一般に用いられているが、DPC分類のカバー率が低い場合(北里大学病院で48.80%、北里大学東病院で14.81%)には信頼性が十分ではない。多くのDPC対象病院に当てはまるDPC分類を選択したコアのDPC分類群を元に効率性指標を算出する試みもあり、未だ、合理的な効率性指標の概念が確立している訳で

はない。以上のような限界を把握したうえでなお、DPCはこれまでにない全国基準のデータとして自院の位置づけや医療圏の環境を客観的に示すものである。このようにDPCは単なる診療報酬請求上の包括医療制度ではなく、医療情報の標準化のツールであり、自院のデータや公表されているデータを分析することにより、医療の標準化と質の向上、ベンチマーク、地域の患者マーケット調査など病院経営に応用することができる。ここに示した結果はDPC分析で得られるものの一部であり、今後は収支を含めたさらに精緻化したDPC分析により、エビデンスに基づく医療経営に役立つものと考えている。

謝辞

この研究は厚生労働科学研究費補助金「難治性疾患の医療費構造に関する研究」と、文部科学省北里大学・青山学院大学戦略的連携支援事業「ヘルスケア・ソリューション研究」の支援を得て行われた。

文 献

- 1) 松田晋也: DPCとこれからの医療政策—DPCと病院マネジメント—. 松田晋也 編, じほう, 東京, 2005, p.2-19
- 2) 松田晋也: 新しい病院マネジメントの考え方—DPCと病院マネジメント—. 松田晋也 編, じほう, 東京, 2005, p.20-31.
- 3) 橋本英樹: 医療経営概論, 医療経営の基本と実務 上巻 [戦略編]. KPMGヘルスケアジャパン編, 日経メディカル開発, 東京, 2006, p.1-16.
- 4) 阿部賢則: 経営戦略, 医療経営の基本と実務 上巻 [戦略編]. KPMGヘルスケアジャパン編, 日経メディカル開発, 東京, 2006, p.17-90.
- 5) 伏見清秀: DPCデータ活用ブック 第2版. じほう, 東京, 2008, p.31-76.

渋谷 明隆, 他

Title positioning of Kitasato University hospitals based on DPC data analyses

Akitaka Shibuya,^{1,3} Hirotsuke Kobayashi^{2,3}¹Department of Risk Management and Health Care Administration, Kitasato University School of Medicine²Department of Medical Engineering and Technology, Kitasato University School of Allied Health Sciences³Project Team A, Operations Management, Kitasato University Hospital

Background: The Diagnosis Procedure Combination, DPC, is the inclusive payment system in Japan, providing data useful to healthcare management bodies.

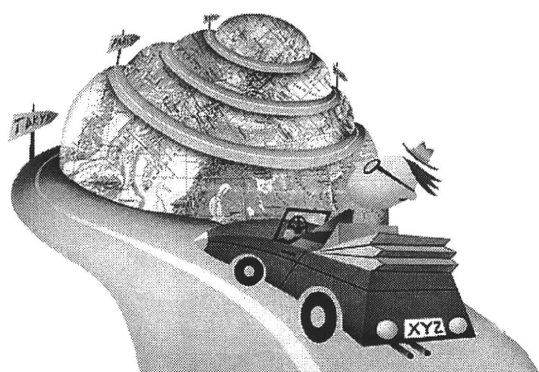
Method: The overall management of the four Kitasato University hospitals was evaluated on the basis of DPC data so far published by the Japanese Ministry of Health, Labour and Welfare in 2009. We also performed a market analysis and SWOT (Strength, Weakness, Opportunity, Threat) analysis using DPC data.

Results: Kitasato University Hospital treated serious illnesses with more efficiency than the average Japanese hospital, but Kitasato University East Hospital was inefficient at treating such illnesses. Kitasato Institute Hospital treated mild diseases efficiently but Kitasato Medical Center was not efficient in the treatment of mild cases. Within the medical market, Kitasato University Hospital played an important role in the treatment of seriously ill patients, especially in the field of cancer.

Conclusions: DPC is not only an inclusive payment system, but it can also serve as a tool for treatment standardization. DPC provides data usable for improving the quality of medical care, for benchmark studies and for market analyses, all of which contribute to hospital management.

Key words: DPC, health care management, medical quality, medical market analysis, SWOT analysis

シリーズ 第12回



本連載の第1回「偶然によるばらつきとは：論文のデータをシミュレートする」で、平均値と標準偏差が分かっている場合、その集団からのランダムサンプルの平均値がどのような分布をするかを、Microsoft Excelを用いたシミュレーションで知ることができることを紹介した。Personal Computer (PC)が広く使われるIT機器となっている現在、PCによるシミュレーションは統計解析において、有力なツールとなる。前回第11回「Rを用いた項目応答理論による解析の実際」で紹介したRは統計解析用のフリーソフトとして、さまざまな解析に用いることができるが、強力なシミュレーションの関数を備えており、Excelよりも容易にシミュレーションを行うことができる。ここでいうシミュレーションとはモンテカルロ・シミュレーションのことである。

PCを用いたマルコフ連鎖モンテカルロシミュレーションMarkov Chain Monte Carlo (MCMC) simulationはベイズ統計解析で使用され、データを得た際の事後確率分布を求めるのに用いられている。事後確率分布を得るためのアルゴリズムの一つであるGibbs samplerを用いるWin BUGSはフリーソフトであり、広く用いられているが、Rも同様の解析に用いることができる。他にも、市販のソフトも含め、PCによるシミュレーションは統計解析の分野で広く用いられつつある。

今回は、Rを用いていくつかのシミュレーションを行い、サンプルサイズの算出や信頼区間の算出に応用してみたい。RはThe Comprehensive R Archive Network (CRAN)のホームページ (<http://cran.r-project.org/>)からダウンロードできるが、インストールについては前回は参照されたい。

正規分布に従う母集団からランダムサンプルを得る

サンプルが正規分布に従う場合、その母集団からラ

Rを用いた シミュレーションによる サンプルサイズの算出

森實 敏夫

Morizane Toshio

神奈川歯科大学 内科

ンダムに抽出したサンプルを得ることを行ってみよう。rnorm (サンプル数, 平均値, 標準偏差) がそのための関数である。たとえば、母集団の平均値が120、標準偏差が10の正規分布に従う場合、ランダムサンプルを100得て、その分布をみる場合、Rでは次のように記述する：

```
> rnd_sample1 = rnorm(100, 120, 10)
```

rnd_sample1は任意に定めた変数である。このように記述し、Enterキーを押すと、この変数に100個の値がベクトルとして格納される。ベクトルは、コンピュータ・プログラミングでは、配列に相当し、100個の値が番号が付けられて順番に並んでいるものである。

たとえば、次のように記述して、Enterキーを押すと、このベクトルに格納されている値が100個表示される：

```
< rnd_sample1
[1] 107.99175 115.94839 119.75685 131.01354
    130.74986 136.51626 128.48748
[8] 113.11323 125.73524 111.96516 98.11553
    141.38978 125.29227 119.92777
途中省略
[92] 135.50827 123.75773 121.4938
    134.53572 119.71783 109.34538
    128.19721
```

```
[99] 116.20649 104.15110
```

もし、10番目の値を知りたいければ、次のように記述する：

```
< rnd_sample1[10]
[1] 111.9652
```

確かに10番目の値が表示される。

さらに、93番目から97番目までの5個の値を知りたいければ、次のように記述する：

```
> rnd_sample1[93:97]
[1] 123.7577 121.4938 134.5357 119.7178
    109.3454
```

確かに、93番目から97番目までの5個の値が表示されている。

rnd_sample1という変数に平均値120、標準偏差10の正規分布に従う母集団からのランダムサンプルが100個格納されているという意味が理解されたと思う。

この100個のランダムサンプルの分布をヒストグラムで見ると、hist()という関数によって、簡単にできる。

記述の仕方は、hist(変数名, xlab=" x軸ラベル", ylab=" y軸ラベル", main=" タイトル")で、さらに、add=TRUEを加えると、前の図に重ねて描画する。また、freq=FALSEを加えると、縦軸が率で表わされる。単に分布を確認するためにヒストグラムを表示するのであれば、hist(変数名)だけで十分である。

それでは、次のように記述してrnd_sample1の分布を見てみよう。より正確に表現するのであれば、rnd_sample1という変数に格納されている100個の値の分布である：

```
< hist(rnd_sample1)
別ウインドウに図1のようなグラフが表示される。
```

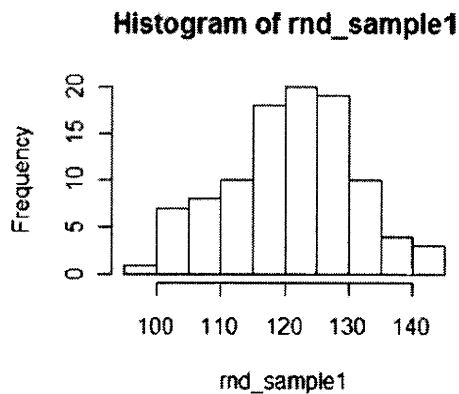


図1. 平均値120、標準偏差10の母集団からのランダムサンプル100個の分布を示すヒストグラム

それでは、これら100個のランダムサンプルの平均値やばらつきについて見てみよう。

まず、summary()関数である：

```
> summary(rnd_sample1)
Min. 1st Qu. Median Mean 3rd Qu. Max.
 98.12 114.70 121.30 120.90 127.40 144.60
最小値、第1四分位、中央値、平均値、第3四分位、最大値が表示される。
```

次に、mean()関数で平均値、sd()関数で標準偏差を求める：

```
> mean(rnd_sample1)
[1] 120.8771
> sd(rnd_sample1)
[1] 9.7015
```

なお、var()関数で分散を求めることができる。sqrt(var(rnd_sample1))で標準偏差を求めることもできる。

95%信頼区間は、quantile(変数名, c(0.025, 97.5))で求めることができる。

```
> quantile(rnd_sample1, c(0.025, 0.975))
 2.5% 97.5%
101.8213 139.5924
```

quantile(変数名, c(0.05, 0.95))とすれば、90%信頼区間となる。

以上の検討から、これら100個のランダムサンプルの平均値は母集団の平均値に近似し、標準偏差も母集団の標準偏差に近似していることが分かる

また、ヒストグラムの形からも、いわゆる"釣鐘型"をしており、正規分布に近似しているように見える。

それでは、正規分布に従った分布かどうかをShapiro-Wilk検定で見てみることにする。Rでは関数shapiro.test(変数名)が用意されている。次のように記述する：

```
> shapiro.test(rnd_sample1)
Shapiro-Wilk normality test
data: rnd_sample1
W = 0.9904, p-value = 0.6943
```

P値が0.6943と0.05よりは大きいので、正規分布に従うということがいえる。

もう一つの同様の関数としてKolmogorov-Smirnov検定用の関数ks.test(変数名, "pnorm", mean=平均値, sd=標準偏差)が用意されており、次のように記述する：

```
> ks.test(rnd_sample1, "pnorm", mean =
mean(rnd_sample1), sd = sqrt(var(rnd_sample1)))
One-sample Kolmogorov-Smirnov test
```

```
data: rnd_sample1
D = 0.0636, p-value = 0.8128
alternative hypothesis: two-sided
```

この検定でも、P値は0.8128なので有意とは言えず、正規分布に従うと判定することができる。これら2つの検定は、帰無仮説が「正規分布に従う」なので、それを棄却できなければ、正規分布に従うと言える。

以上の試行から、ランダムサンプルであれば、①その平均値と標準偏差は母集団の平均値と標準偏差に近似するということが、②母集団が正規分布に従う場合、ランダムサンプルも正規分布に従うということが示された。

サンプルの平均値の分布

それでは、それぞれのサンプルの値ではなく、サンプルの平均値の分布がどうなるかをシミュレーションしてみよう。上記のごとく、たとえば、100個のランダムサンプルを1回抽出して、その平均値を算出することを多数繰り返し、その多数の平均値の分布を見ればよいということになる。

Rには繰り返しの関数として、replicate(回数,

expression, simplify=TRUE)という関数がある。expressionを回数分実行し、ベクトルに収めるということを実行してくれる。たとえば、平均値120、標準偏差10の正規分布に従う母集団から100個のランダムサンプルを得て、その平均値を算出し、それを1万回繰り返して、変数m_distに格納することを行うには次のように記述する。ここでは、平均値、標準偏差、サンプルサイズ、回数をそれぞれm, s, n, kという変数にいったん代入してから計算を実行している：

```
> m = 120
> s = 10
> n = 100
> k = 10000
> m_dist = replicate(k, mean(rnorm(n, m, s)))
```

これで、ランダムサンプル100個の平均値10000個が変数m_distにベクトルとして格納された。

summary()関数でその10000個の値の平均値などを見てみよう：

```
> summary(m_dist)
Min. 1st Qu. Median Mean 3rd Qu. Max.
116.2 119.3 120.0 120.0 120.7 123.9
```

さらに、10000個の値の標準偏差は：

```
> sd(m_dist)
[1] 0.9996989
```

すなわち、標準偏差は小数点以下4桁目で四捨五入すると、1.000となる。

さて、ここで1つの平均値は100個のサンプルから算出した。すなわち、n=100のランダムサンプルの平均値である。母集団が正規分布に従い、平均値 μ 、標準偏差 σ の場合、n=100のサンプルの平均値は、平均値 μ 、標準偏差 σ/\sqrt{n} の正規分布に従うという中心極限定理が正しいことがここで示された。すなわち、 $s/\sqrt{n} = 10/\sqrt{100} = 10/10 = 1$ となり、上記の標準偏差値と一致する。平均値も120であるから、母集団の平均値 = mの値と一致している。

それでは、100個のサンプルの平均値の95%信頼区間を算出してみよう。上記のごとく関数quantile()を用いて：

```
> quantile(m_dist,c(0.025,0.975))
2.5% 97.5%
118.0255 121.9067
```

したがって、平均値120、標準偏差10の母集団から100個のランダムサンプルを得て、平均値を求めた場合、95%の場合は、118~122の範囲の値が得られるということが分かる。上記のごとく、100個のランダムサンプルの平均値は平均値120、標準偏差1の正規分布に従うので、正規分布の曲線下の面積が0.95となる範囲は $1.96 \times \text{標準偏差} = 1.96 \times 1 = 1.96$ なので、 $120 \pm 1.96 \approx 120 \pm 2.0$ となる。すなわち、118~122となり、上記の値と一致する。

実際には、1つの研究では100個のサンプルを1回だけ調べることしかできないので、そのサンプルの平均

値と標準偏差の値を母集団の平均値と標準偏差の近似値として用いて、サンプルの標準偏差s、サンプルサイズnとすると、 $1.96 \times s/\sqrt{n}$ の計算を行って、95%信頼区間を求めることが行われる。 s/\sqrt{n} はいわゆる標準誤差(Standard error, SE)と呼ばれている。すでに述べたように、サンプルの平均値と標準偏差はサンプルサイズが大きければ、母集団の平均値と標準偏差に近似するので、これが可能となる。

それでは、これら1万個の平均値の分布をヒストグラムで見てみよう。

```
> hist(m_dist)
```

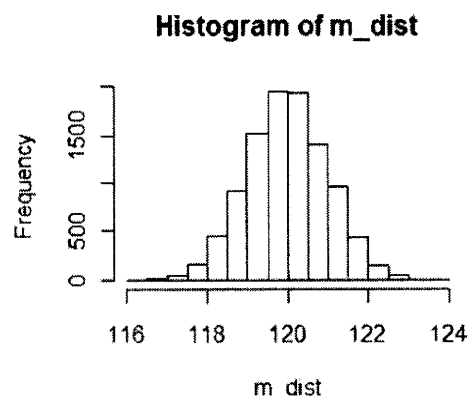


図2. 平均値120、標準偏差10の母集団からのランダムサンプル100個の平均値の分布を示すヒストグラム

図1と比べると、左右対称の釣鐘形で正規分布に従うことは見てとれるが、横軸はより狭い範囲に分布していることが分かる。なお、平均値の値が1万個あるので、Rではshapiro.test()を実行することはできない(最大5000個まで)。

単一群のサンプルサイズの算出

ある分布が想定される1つの母集団から1つの群をランダムサンプルとして抽出する場合に、サンプルサイズをどれくらいにするかによって、データの精度が決まってくる。上記の例のように、平均値 $\mu=120$ 、標準偏差 $\sigma=10$ の母集団の場合、100症例のサンプルサイズであれば、平均値は 120 ± 2 の範囲に95%の確率で入ることになるので、かなり精度の高いデータが得られる。この場合、本当の平均値、すなわち母集団の平均値からずれても、そのずれが最大2までは許容してもよいということであれば、100症例のサンプルサイズで調査をすればよいことになる。この、2という値は最大過誤(E)とよばれ、サンプルの平均値が $(1-\alpha)100\%$ の範囲に入るようにするためのサンプルサイズは次の式で算出される：

$$n = \frac{Z_{\alpha/2}^2 \sigma^2}{E^2} \quad \text{式 (1)}$$

この式は次の式から展開したものである：

$$(1-\alpha)100\% = \bar{X} \pm Z_{\alpha/2} \sigma / \sqrt{n}$$

ここで、たとえば、95%信頼区間であれば、 α は0.05となり、 $Z_{\alpha/2}$ は1.96となる。すなわち、標準正規分布に変換して、 $0 \pm 2SD$ の範囲にほぼ相当する。 \bar{X} はサンプルの平均値のことである。したがって、最大過誤は $\mu - \bar{X}$ である。

このように、データが連続変数の場合、想定される母集団の標準偏差、最大過誤、そして任意に設定される α 値によって、サンプルサイズを算出することができる。これは、正確度分析 Precision Analysis と呼ばれる方法である。実際には、母集団の分布の標準偏差を知ることはできないので、パイロット研究からのサンプルのデータに基づいて想定するしかない。ある程度のサンプルのデータがあれば、その標準偏差の値を母集団の標準偏差として用いればよい。

それでは、式(1)を用いて、 $E=1.96 \div 2$ の場合で、 $\alpha=0.05$ 、すなわち95%の確率でサンプルの平均値が 120 ± 2 の範囲に入るサンプルサイズを計算してみよう。

$$1.96^2 \times 10^2 / 1.96^2 = 100$$

つまり、100症例ということになる。この程度の計算は、普通の計算機でも、Excelでも容易に行うことができる。

そして、これは、上記のシミュレーションで得た結果と同じである。

それでは、Rを用いたシミュレーションで、任意の最大過誤Eを設定し、サンプルサイズを算出するにはどうしたらいいか考えてみよう。平均値mと標準偏差sは固定して、nをある値から1つずつ増加させ、quantile(変数、0.025)の値が $m-E$ を超えたときのnの値にすればよいと考えられる。これを手作業で行うことも可能であるが、手間がかかる。それではどうしたらよいか？

Rでは自分で関数をテキストファイルとして書いて、それを呼び出して実行させることが可能である。まず、適当なフォルダ名で新しいフォルダを作成する。RのFileメニューから、Change dir...を選び、そのフォルダを作業フォルダに指定する。以下に示す内容をメモ帳などのテキストエディタで書きこんで、テキストファイルとして、そのフォルダに保存する。ファイル名はsamp_size_m_sd_s.Rのようにする。すなわち、拡張子としてRを付ける。ファイル名は自分で分かりやすい他のものでも構わない。

```
sample.size.m.sd=function(m,s,E)
{
  lowlimit = m - E
  maxn = 1000
  k = 10000
  for (n in 1:maxn)
  {
    m.dist = replicate(k, mean(rnorm(n, m, s)))
    if (quantile(m.dist,0.025)>= lowlimit)
      break
  }
}
```

```
}
return(n)
}
```

このプログラムをRから呼び出して、実行させることができる。このプログラムは、m,s,Eの3つの引数を受け取り、95%信頼区間の下限値として、lowlimitという変数に $m-E$ の値を代入する。mは平均値、sは標準偏差である。サンプルサイズの上限值として、1000をmaxnという変数に代入する。rnorm()関数を用いてn個のランダムサンプルを得る作業を10000回実行させるが、この回数の値はkという変数に代入する。nの値を1から開始し、1つずつ増加させながら、平均値m、標準偏差sの正規分布に従う母集団からのランダムサンプルをn個抽出して、その平均値を求める作業を10000回行い、95%信頼区間の下限値がlowlimitの値を超えた時点で、この繰り返し作業を中断し、その時のnの値を返り値として戻す。以上である。

Rでは次のように記述する：

```
> source("samp_size_m_sd_s.R")
```

これにより、上記のプログラムを使用できる状態にする。これにより、自分で作成した関数sample.size.m.sd(平均値、標準偏差、最大過誤値)を用いることができるようになる。この関数には、カッコ内の3つの値をコンマで区切って入力する。

```
> sample.size.m.sd(120,10,1.96)
```

```
[1] 100
```

この例では、平均値が100、標準偏差が10の母集団からのランダムサンプルの平均値の分布で、95%信頼区間の下限値が $120-1.96$ 以上となった際の値が得られる。たしかに、100という値が得られた。式(1)で算出した値と同じであり、最初のシミュレーションの結果とも一致している。

層別化集団からのサンプルの場合

それでは、母集団が1つだけではなく、互いに異なる、平均値、標準偏差を持つ、複数の母集団からのランダムサンプルを扱う場合はどのようにしてサンプルサイズを求めたらいいだろうか。

たとえば、それぞれの医療機関ごとに、ある疾患の医療費が異なるような場合、それぞれの医療機関内でのデータ分布はある平均値とある標準偏差の正規分布に従うが、医療機関ごとにそれらの値が異なるような場合である。その場合には、いくつかの層に分かれていると考えることができ、各医療機関が層を構成している。そして、全体としての平均値をある一定の精度で知りたい場合に、どの程度のサンプル数を集めればよいかを知りたいということが目的である。すなわち、全医療機関を受診している全患者の医療費を調査するのではなく、一部の医療機関の一部の患者をサンプルとして抽出して、その平均値をある一定の精度で求めるという手順である。精度をどの程度にするかは、任

意に設定する要素である。

この場合、各医療機関を受診しているその疾患の患者数は異なるので、それぞれの医療機関の平均値を単純に平均することはできない。各医療機関のその疾患の患者数の比のデータが必要になる。できるだけ小さな値で、整数で比を表す必要がある。

ここで問題は、平均値と標準偏差を正確に求めることができないということである。これらのパラメータは母集団の値であるから、データから想定するしかない。たとえば、少数例のサンプルから求めた平均値と標準偏差の値を用いるのが一つの方法である。また、標準偏差については、その平均値から一番値の小さい例の値の差を求め、それを1.96で割り算した値を標準偏差として用いることも可能である。後者は、95%信頼区間の下限値が一番値の小さい例の値であるとみなして、算出する方法である。

それでは、表1に示す、3つの医療機関を調査しましょう。

それぞれの医療機関からの症例数として、2 : 3 : 5 になるように、10、15、25例を調査するとしよう。

Rで以下のように記述する：

```
> k = 10000
> n1 = 10
> m1 = 7000
> s1 = 650
> n2 = 15
> m2 = 8000
> s2 = 1200
> n3 = 25
> m3 = 9500
> s3 = 1000
> h_average = replicate(k, mean(c(rnorm(n1, m1,
s1), rnorm(n2, m2, s2), rnorm(n3, m3, s3))))
> quantile(h_average, c(0.025, 0.975))
> quantile(h_average, c(0.025, 0.975))
  2.5%  97.5%
8269.641 8822.601
> summary(h_average)
Min. 1st Qu. Median Mean 3rd Qu. Max.
8052 8455 8551 8551 8649 9094
> hist(h_average)
```

Histogram of h_average

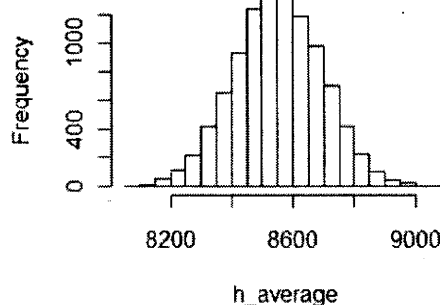


図3. それぞれ異なる分布の3つの層からのランダムサンプルの平均値の分布

平均値が8551円、95%信頼区間が8270～8823円という結果が得られた。すなわち、それぞれの医療機関から、10、15、25例ずつのランダムサンプルを抽出して、平均値を求めることを行った場合、95%の確率で、この範囲の値が得られることが分かる。すなわち、これらのサンプルサイズの場合、実際のデータとして、3医療機関全体の平均値は8551となる確率が最も高いが、もし偶然ずれたとしても、小さい場合に8270、大きい場合に8823となり、これらの範囲から外れることはほとんどないと考えてよいということになる。もし、この程度の精度で十分であるといえるなら、それぞれの医療機関から、10、15、25例ずつの調査をすれば十分であると言える。

それでは、3医療機関の平均値があるい任意に設定した最大過誤に収まるために必要なサンプルサイズの計算をするにはどうしたらいいか。上記のプログラムを拡張すればいいのであるが、それは次回に紹介したい。

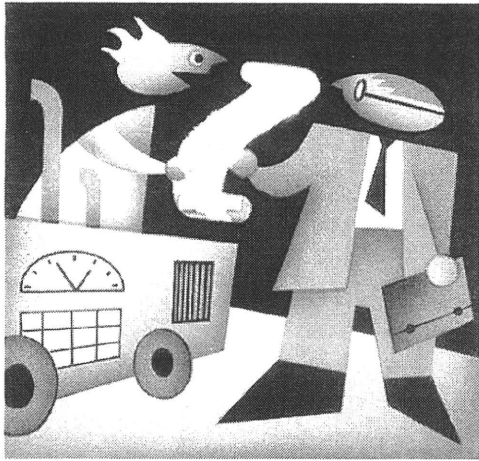
謝辞：

この論文は平成21年度「厚生労働省科学研究費補助金」[難治性疾患の医療費構造に関する研究]による支援を受けた。

表1. 3医療機関のデータ

	患者数 (N _i)	整数比 (n _i)	医療費平均値 (m _i) (円)	医療費標準偏差 (s _i)
医療機関 1	40	2	7000	650
医療機関 2	60	3	8000	1200
医療機関 3	100	5	9500	1000

シリーズ 第13回



Rでランダムサンプルを得る関数

前回、3つの医療機関のデータについて、それぞれの医療機関のその疾患の患者総数、医療費の平均値と標準偏差が分かっている場合、ある任意のサンプル数を調査した場合の、平均値の精度、すなわち95%信頼区間を、Rを用いたモンテカルロ・シミュレーションにより算出する方法について述べた。より一般的な言い方をすると、層別化された連続変数のデータがあり、それぞれの層における平均値と標準偏差が分かっている場合に、層全体のサンプルの平均値と信頼区間を算出することである。

この計算では、Rの関数の`rnorm(num, m, s)`を用いる。この関数は、平均値 m 、標準偏差 s の正規分布に従う母集団から、 num 個のサンプルをランダムに得る関数である。この関数を実行するたびに、 num 個のランダムサンプルが得られるので、その平均値を求めることを繰り返して行くと、この母集団からのサンプルサイズ num のサンプルの平均値を多数得ることができ、その分布を見たり、その平均値と標準偏差を求めることができる。この標準偏差はサンプルの平均値の標準偏差であり、通常標準誤差と呼ばれているものに相当する。

`rnorm()`関数で3つの医療機関1, 2, 3からのランダムサンプルを`rnorm(num1, m1, s1)`, `rnorm(num2, m2, s2)`, `rnorm(num3, m3, s3)`として得たうえで、全体の平均値を算出し、それを繰り返すことによって、3つ

Rを用いた シミュレーションによる サンプルサイズの算出—2

森實 敏夫

Morizane Toshio

国際医療福祉大学教授 塩谷病院内科

の層からのそれぞれのサンプル数のサンプルの平均値の分布、平均値、標準偏差を知ることが可能になる。

この計算の際に、サンプル数を増加させていくと、信頼区間の幅は徐々に狭くなっていく。そこで、各層のサンプル数を1ずつ増加させ、信頼区間が希望する任意の値以下になったら、その際の`num1`, `num2`, `num3`の値を知ることによって、一定の精度でデータを得るために必要なサンプル数を知ることができる。今回は、この作業を自動的に行い、`num1`, `num2`, `num3`の値を返す関数を作成してみる。

なお、母集団の平均値と標準偏差を正確に知ることとはできないので、少数例のサンプルから求めた平均値と標準偏差の値を用いることとする。なぜならば、ランダムサンプルであれば、その平均値も標準偏差も母集団の平均値と標準偏差に近似するからである。また、標準偏差については、その平均値から一番値の小さい例の値の差を求め、それを1.96で割り算した値を標準偏差として用いることも可能である。これは、95%のデータが含まれる範囲の下限値が一番値の小さい例の値であるとみなして、算出する方法である。

前回と同じく、表1に示す、3つの医療機関を調査するとして。患者数(N_i)はそれぞれの医療機関における調査対象となる疾患の全患者数である。その整数比(n_i)はこの場合、2 : 3 : 5となる。この整数比は、得られたサンプルの医療費の平均値から全体の医療費を推定したい様な場合には、必要となる。

表1. 3医療機関のデータ

	患者数 (N_i)	整数比 (n_i)	医療費平均値 (m_i) (円)	医療費標準偏差 (s_i)
医療機関 1	40	2	7000	650
医療機関 2	60	3	8000	1200
医療機関 3	100	5	9500	1000

3つの医療機関のデータに対するサンプルサイズ算出の関数

3つの医療機関の患者数の整数比、それぞれの医療費の平均値と標準偏差の値から、モンテカルロ・シミュレーションに基づき、必要なサンプルサイズを算出する関数を作成する。以下の内容をテキストファイルとして、ファイル名はsamp_size_m_sd_m.Rとして保存する。関数名はsample.size.m.sd.m()である。

```
sample.size.m.sd.m=function(m1,s1,n1,m2,s2,n2,m3,
  s3,n3,E)
{
  gm=(m1*n1+m2*n2+m3*n3)/(n1+n2+n3)
  #全体の平均値
  lowlimit=gm-E #95%信頼区間下限値
  maxn=1000 #繰り返しの最大数
  k=10000 #平均値の計算の回数
  for(n in 1:maxn) #n=1からmaxnまで1ずつ増加
  {
    num1=n1*n #各医療機関のサンプル数
    num2=n2*n
    num3=n3*n
    h_average=replicate(k,mean(c(rnorm(num1,m1,s1),
      rnorm(num2,m2,s2),rnorm(num3,m3,s3))))
    #3つの医療機関からのランダムサンプルの平均値の
    計算k回

    if(quantile(h_average,0.025)>=lowlimit)
      #95%信頼区間下限値を超えるかの判定
    break
  }
  return(c(n, gm)) #nと全体の平均値を返す
}
```

この関数は、医療機関1の平均値m1、標準偏差s1、整数比の値n1、医療機関2の平均値m2、標準偏差s2、整数比の値n2、医療機関3の平均値m3、標準偏差s3、整数比の値n3とし、全体の平均値の最大過誤をEとして代入する。返り値は、nと全体の平均値である。返り値nをそれぞれn1、n2、n3に掛け算した値が、それぞれの医療機関でのその精度に必要なサンプルサイズとなる。ランダムサンプルの平均値の計算を行う回数は、kで設定するが、10000回としている。計算時間を長くしてよいのであれば、精度を高めるために、k=100000と大きな値としてもよい。

上記の例のデータを用いた場合、Rでは次のように記述する。まず、ファイルsamp_size_m_sd_m.Rを保存したフォルダーをFileメニューのChange dir.から指定する。次に、>source()には、上記のファイル名を入力し、関数を読み込ませる。その後、各変数の値を代入

し、最後に、ファイル名ではなく、ファイルの中書き込んだ関数名とカッコ内にそれぞれの変数名を引数として入力する。引数とは、関数に値を渡すために用いる変数である。ここでは、Eを282として試算してみる：

```
> source("samp_size_m_sd_m.R")
> m1 = 7000
> s1 = 650
> n1 = 2
> m2 = 8000
> s2 = 1200
> n2 = 3
> m3 = 9500
> s3 = 1000
> n3 = 5
> E = 282
> sample.size.m.sd.m(m1,s1,n1,m2,s2,n2,m3,s3,n3,E)
[1] 5 8550
```

実際にn=5という値が得られたので、それぞれの医療機関でのサンプルサイズは、2、3、5に5を掛け算した値、すなわち、10、15、25例ずつということになる。このサンプルサイズで調査すれば、95%の確率で、全体の平均値は8550±282に含まれるということになる。

さて、さらに層の数が多い場合には、上記の関数を手直しして、それぞれの層の個数に対応した関数を作成すれば対応が可能である。ポイントは、for(n in 1:maxn){}で、nを1ずつ増加させながら、平均値の分布をみて、95%信頼区間の下限値が一定の値以上になったら、計算を中止して、その際のnの値を返すという点である。

2群の平均値の差のt検定に必要なサンプルサイズを求める

連続変数を独立した2群をt検定で比較する場合のサンプルサイズをモンテカルロ・シミュレーションで求める関数を作成してみよう。

それぞれの群gr1とgr2の平均値と標準偏差をm1, s1, m2, s2とする。それぞれの群からn1, n2個のランダムサンプルを得て、その値をt検定で比較するため、t統計値を算出し、自由度df=n1+n2-2の有意水準より大きいかどうかを調べる。すなわち、P<0.05となるかどうかを調べる。これを10000回繰り返し、その中で、有意水準より大きくなった回数の割合(prob)を算出する。すなわち、実際の有意水準をシミュレートして求める。

まず、引数として $m_1, s_1, n_1, m_2, s_2, n_2$ および α 水準 (cutoff) を与えると、ランダムサンプルを生成したうえで、10000回 t 検定を繰り返して、その中で、 t 統計値が有意水準を超える割合を返す関数を作成する。ファイル名をtest_t.Rとして保存する。関数名は、test.t()である。

```
test.t=function(m1,s1,n1,m2,s2,n2,cutoff)
{
  cnt = 0 #カウンターの設定
  k = 10000 #繰り返す回数の設定
  df = n1+n2-2 #自由度
  for (n in 1:k) #nを1ずつ増加させながら繰り返す
  {
    gr1 = rnorm(n1, m1, s1) #ランダムサンプルの生成
    gr2 = rnorm(n2, m2, s2) #ランダムサンプルの生成
    sp = sqrt(((n1-1)*sd(gr1)^2+(n2-1)*sd(gr2)^2)/df)
    #分散のプール
    t.stat =(mean(gr1)-mean(gr2))/(sp*sqrt(1/n1+1/n2))
    #t統計値の計算
    if (abs(t.stat)>qt(1-cutoff/2, df))
      #両側検定によるt統計値の判定
    cnt = cnt + 1 #有意な場合にカウントする
  }
  prob = cnt/k
  #1万回シミュレート後有意となった割合を算出
  return(prob) #割合の値を返す
}
```

この関数のプログラムでは、 t 統計値の計算は特別な関数を用いず、上記の式で行っている。また、qt()関数はRの関数で、確率と自由度を渡すと、対応する t 統計値を返す関数である。この関数では両側検定を行っているが、片側検定であれば、if (abs(t.stat)>=qt(1-cutoff/2, df))の部分を実はif (abs(t.stat)>=qt(1-cutoff, df))とする。あるいは、引数として渡すcutoffの値を2倍にする。後者の方が便利であろう。

まず、この関数test.t()を動かしてみよう。この関数には、それぞれの群の平均値、標準偏差、サンプル数、 α 水準を入力する。この α 水準は t 検定で有意と判定する値であり、通常は0.05とする。

たとえば、 $m_1=110, s_1=10, n_1=10, m_2=110, s_2=10, n_2=10, cutoff=0.05$ で計算してみよう。これは、2群の平均値が同じで、標準偏差が同じ場合に、10例ずつのサンプルを調べると偶然の偏りで、 $P < 0.05$ で有意と判定する率を求めることになる。

まず、関数を読み込ませる。なお、これらの関数のファイルを保存したフォルダをあらかじめ、FileメニューのChange dir.としておく。

```
> source("test_t.R")
```

次いで、test.t()関数を実行させる。

```
> test.t(110,10,10,110,10,10,0.05)
[1] 0.0486
```

この意味するところは、いずれも平均値100、標準偏差10の母集団、すなわち2群に差が無い母集団から、10個のランダムサンプルをそれぞれ得て、これらの値を α 水準0.05で t 検定で比較することを10000万回繰り返すと、486回は $P < 0.05$ で有意という判定をすることになるということを示している。

$m_1=110, s_1=10, n_1=10, m_2=120, s_2=10, n_2=10$ ではどうなるか見てみよう。

```
> test.t(110,10,10,120,10,10,0.05)
[1] 0.5726
```

標準偏差は10と同じであるが、平均値は110と120で10差がある2群を α 水準0.05で t 検定で比較することを10000回繰り返すと、5726回は $P < 0.05$ で有意差があるという結果を得ることを示している。これは検出力に相当する。

このように、シミュレーションで得たランダムサンプルで実際に t 検定を行った場合の結果について知ることができる。

それでは、次に、サンプル数 n を2から1つずつ増加させながら、test.t()関数を繰り返し呼び出して、帰無仮説： $m_1 = m_2$ の下で、 t 検定の結果が有意となる割合(proba)と、対立仮説： $m_1 \neq m_2$ の下で、 t 検定の結果が有意となる割合(prob)を順次計算し、probが検出力(power)以上、probaが α 水準(alpha)以下になった時点の n を返す関数を作成する。引数としては、 $m_1, s_1, m_2, s_2, alpha, power$ を与える関数である。ファイル名をsamp_size_t_test.Rとして保存する。関数名は、samp.size.t.test()である。

```
samp.size.t.test=function(m1,s1,m2,s2,alpha,power)
{
  maxn = 1000 #サンプル数の最大値の設定
  for (n in 2:maxn) #nを1ずつ増加させながら繰り返す
  {
    probp = test.t(m1,s1,n,m2,s2,n,0.05)
    #test.t()関数を呼び出し、返り値を得る
    # (対立仮説における有意となる割合)
    proba = test.t(m1,s1,n,m1,s2,n,0.05)
    #平均値が同じ場合 (帰無仮説における
    #有意となる割合)
```

```
if (probp >= power)
  # 検出力を以上およびα水準以下になったら中止する
if (proba <= alpha)
break
}
return(n) #nの値を返す
```

それでは、平均値がそれぞれ110, 120で標準偏差がいずれも10の2群をt検定で比較する場合に必要なサンプルサイズをこれら2つの関数を用いて算出してみよう。

まず関数をそれぞれ読み込ませる。

```
> source("test_t.R")
> source("samp_size_t_test.R")
```

次いで、必要な引数の値を次のように入力して計算を行う。α水準を0.05、検出力は0.9で計算を行ってみた。得られる値は、各群のサンプル数である。この計算には、PCの性能にもよるが、かなり時間がかかる。

```
> samp.size.t.test(110,10,120,10,0.05,0.9)
[1] 22
```

22例ずつ調べればよいという結果が得られた。

ちなみに、正規分布に従う独立した2群のt検定による比較の際に必要なサンプルサイズは次の式で計算される¹⁾：

$$\Delta = |\mu_1 - \mu_2|/\sigma$$

$$n = 2(Z_\alpha + Z_\beta)^2/\Delta^2 + Z_\alpha^2/4$$

σは母集団の標準偏差、μはそれぞれの平均値である。なお、α水準0.05の場合、Z_α = 1.96、検出力0.9の場合、Z_β = 1.28である。nは1群のサンプル数である。

2群をt検定で比較するような場合のサンプルサイズの算出は、モンテカルロ・シミュレーションによらなくても、計算式にしたがって、Excelなどで計算する方が簡単かもしれないが、標準偏差が2群で異なる場合、3群以上の比較や、3群以上の比較で、1群が他の2群より有意に高いあるいは低いということを示すためのサンプルサイズを算出するような場合など、複雑な系の場合には、モンテカルロ・シミュレーションによる方法は一つの有力な方法になるのではないかと考えられる。

また、モンテカルロ・シミュレーションを用いることによって、統計学の理解を深めることもできる。た

例えば、今回の例では、帰無仮説は2群の平均値が等しいということであるが、対立仮説は平均値がそれぞれ110と120で少なくとも10の差があるということであることが分かる。今回の関数を用いると、どのような平均値でどのような標準偏差の2群であっても、ある一定の値を超えるt統計値が得られる率を求めることができる。すなわち、さまざまな仮説の下でどのような結果が得られるかを知ることができる。新しい関数の作成も含めて、さまざまな応用が可能と考えられる。また、連続変数だけでなく、率やリスク比、オッズ比などにも応用が可能である。

文献・資料

1. Machin D, Cambel M, Fayers P, Pinol A: Comparing independent groups for binary, ordered categorical and continuous data. In: Sample size tables for clinical studies. 2nd edition, pp.18-78, 1997, Blackwell Science, Oxford, UK.

Ⅶ. 「難治性疾患の医療費構造に関する研究」

班会議

第1回 平成20年度 厚生労働科学研究費補助金難治性疾患克服研究事業
『難治性疾患の医療費構造に関する研究』班 プログラム

日時：平成20年4月30日（金）18：00～20：00

場所：スペースTOKU 会議室B

東京都中央区八重洲2-6-1 三徳八重洲ビル B1

TEL 03-3271-9403 FAX 03-3271-8587

出席者：荒井耕，伊藤道哉，内田智久，萩野美恵子，川合眞一，川島孝一郎，
高橋一司（鈴木則宏代理），武田輝行，中島孝，伏見清秀，美原盤，頼高朝子（服部信孝代
理），渡辺守

厚生労働省：井原正裕，林修一郎

研究班事務局：桑原淳子

（敬称略。50音順）

議題1 あいさつ（資料1 班員名簿、資料2 出席者名簿）

自己紹介

厚生労働省 林課長補佐 ごあいさつ

中島 プレゼンテーション

議題2 目的の確認および3年間の計画について（資料3 研究費交付申請書）

議題3 本年度の具体的計画について（資料4 本年度の課題）

3-1 課題の把握

3-2 症例抽出方法について（入院・外来で対応必要）

どのコーディングを基準とするか（資料5 コーディング対応表）

どの程度の確からしさを求めるか

次年度にむけて連結可能にするか

次年度にむけて重症度・病期を把握するか

3-3 調査対象の医療機関について

3-4 医療費構造データの区分分けについて

医療費構造の中身 診療報酬の分類？

3-5 データの処理方法について

議題4 既存の研究資料について

患者数の把握（資料6 患者数推計値）

重症度の選択

議題5 業務委託について

議題6 今後のスケジュールについて（スケジュール表）

議題7 事務・経理について

議題8 その他

第2回 平成20年度 厚生労働科学研究費補助金難治性疾患克服研究事業
『難治性疾患の医療費構造に関する研究』班会議 プログラム

日時：平成20年11月22日（土）13：00～16：00

場所：東京 八重洲ホール

東京都中央区日本橋3-4-13 新第一ビル 8階 811号室

TEL 03-3201-3631 FAX 03-3274-5111

出席者：

主任研究者：荻野美恵子

分担研究者：荒井耕、伊藤道哉、川合眞一、川島孝一郎、中島孝、西澤正豊、服部信孝、伏見清秀、
美原盤

研究協力者：相澤勝健、内田智久、亀井俊治、鈴木和久、高橋一司（鈴木則宏代理）、武田輝行、
長堀正和（渡辺守代理）、村岡 成、頼高朝子

厚生労働省：海老名英治

その他出席者：椿井富美恵

研究班事務局：桑原淳子

（敬称略。50音順）

13：00～13：05 開会のあいさつ 難治性疾患の医療費構造に関する研究班長 荻野 美恵子

13：05～13：10 ご挨拶 厚生労働省健康局疾病対策課 課長補佐 海老名 英治

第1部 各分担研究者からの報告

13：10～13：20 1. 神経難病医療費構造に関する研究～施設療養と在宅療養の比較

美原 盤、○内田 智久、相澤 勝健、栗原 真弓、高尾 昌樹
（財）脳血管研究所附属美原記念病院

13：20～13：30 2. 特定疾患患者が障害者自立支援法の療養介護への移行時の医療費構造

—個別事例の検討

中島 孝、○亀井 俊治、武田 輝行
独立行政法人国立病院機構新潟病院

13：30～13：40 3. 特定疾患における酵素補充療法の医療費構造

中島 孝、○武田 輝行
独立行政法人国立病院機構新潟病院

13：40～13：50 4. 在宅医療における各種疾患の医療費・介護費の推計

○川島孝一郎
仙台往診クリニック

- 13：50～14：00 5. 一大学病院（特定機能病院）における神経難病
 （脊髄小脳変性症、多系統萎縮症）
 ○西澤正豊
 新潟病院脳研究所神経内科
- 14：00～14：10 6. 全身性エリテマトーデス(SLE)のレセプト調査結果と今後のアンケート調査計画
 ○川合眞一
 東邦大学医学部内科学講座（大森）膠原病科
- 14：10～14：20 7. 難治性疾患患者のための医療費構造分析の可能性—がんとの比較から
 ○伊藤道哉
 東北大学大学院医学系研究科医療管理学
- 14：20～14：30 8. 地域患者データベースを用いた難病の地域疾病構造と
 医療費構造の推計手法について
 ○伏見清秀
 東京医科歯科大学大学院医療政策学講座医療情報・システム学
- 14：30～14：45 休憩（コーヒーブレイク）
- 第2部 緊急依頼調査について
- 14：45～15：00 9. 支払基金データによる医療費ヒストグラム
 ○荻野 美恵子
 北里大学医学部神経内科学
- 第3部 今後の調査について
- 15：00～15：30 平成20年度全体調査について
- 15：30～16：00 医療費構造分析の精緻化について
- (16：00) 閉会の挨拶

以上

第1回 平成21年度 厚生労働科学研究費補助金 難治性疾患克服研究事業
『難治性疾患の医療費構造に関する研究』班会議 プログラム

日時：平成21年7月19日（日）10：00 - 16：00

場所：東京 八重洲ホール

東京都中央区日本橋3-4-13 新第一ビル 2階 201号室

TEL 03-3201-3631 FAX 03-3274-5111

出席者：

研究代表者：荻野 美恵子

研究分担者：荒井 耕, 伊藤 道哉, 川合 眞一, 渋谷 明隆, 中島 孝, 西澤 正豊,
伏見 清秀, 美原 盤, 森實 敏夫, 山下 和彦

研究協力者：相澤 勝健, 荒井 康夫, 内田 智久, 小野沢 滋, 亀井 俊治, 川下 政幸,
楠 芳恵, 鈴木 和久, 高橋 一司, 武田 輝行, 長堀 正和, 松裏 裕行,
頼高 朝子

厚生労働省：

その他出席者：山口 治紀

研究班事務局：桑原 淳子

(敬称略. 50音順)

10：00～10：05 開会の挨拶 難治性疾患の医療費構造に関する研究班代表 荻野 美恵子

10：05～10：10 ご挨拶 厚生労働省健康局疾病対策課

10：10～10：30 新研究分担者紹介

第一部 研究分担者から研究進捗状況などのご報告

10：30～10：40 東北大学大学院医学系研究科 伊藤 道哉

10：40～10：50 東邦大学医療センター大森病院 川合 眞一

10：50～11：00 慶應義塾大学病院医療事務室 鈴木 和久

11：00～11：10 順天堂大学脳神経内科 頼高 朝子

11：10～11：20 脳血管研究所附属美原記念病院 美原 盤

11：20～11：30 神奈川歯科大学内科学講座 森實 敏夫

11：30～12：10 上記以外の先生方の進捗状況についてのご報告

(荒井 耕, 渋谷 明隆, 中島 孝, 西澤 正豊, 伏見 清秀, 長堀 正和,
小野沢 滋)

12：10～13：00 昼食休憩

第二部 全体調査の報告

13:00～14:45 東京医療保健大学 山下 和彦
 (株)健康保険医療情報総合研究所 山口 治紀
 北里大学医学部神経内科学 荻野 美恵子

1) 社会保険診療報酬支払基金とのやりとりについて

①平成19年11月データ

②平成20年11, 12月平成21年1月データ

2) 国民健康保険中央会とのやりとりについて

平成21年3月データ

3) 平成20年11月緊急調査データについて

北里大学, 慶應義塾大学, 東京医科歯科大学, 美原記念病院, 順天堂大学

4) 各研究班班長からの医療費推計値について

14:45～15:00 休憩 (コーヒープレイク)

第三部 今後の研究の進め方について

15:00～16:00 1) 各医療機関の倫理委員会 申請状況についての報告
 2) 平成21年度全体調査 (対象期間:平成20年7月～12月)
 3) 患者個人調査計画 (調査票作成について)

16:00 閉会の挨拶

以上

第2回 平成21年度 厚生労働科学研究費補助金 難治性疾患克服研究事業
『難治性疾患の医療費構造に関する研究』班会議 プログラム

日時：平成22年1月10日（日）10：00 - 16：00

場所：東京 八重洲ホール

東京都中央区日本橋3-4-13 新第一ビル 2階 201号室

TEL 03-3201-3631 FAX 03-3274-5111

出席者：

研究代表者：荻野 美恵子

研究分担者：荒井 耕, 伊藤 道哉, 川合 眞一, 渋谷 明隆, 中島 孝, 西澤 正豊,
伏見 清秀, 美原 盤, 森實 敏夫, 山下 和彦, 渡辺 守

研究協力者：内田 智久, 川下 政幸, 菊池 豊, 楠 芳恵, 鈴木 和久, 高橋 一司,
武田 輝行, 長堀 正和, 村岡 成, 吉田 誠, 頼高 朝子

厚生労働省：

その他出席者：山口 治紀, 濱谷 次郎

研究班事務局：桑原 淳子

(敬称略, 50音順)

10：00～10：05 開会の挨拶 難治性疾患の医療費構造に関する研究班代表 荻野 美恵子
10：05～10：10 ご挨拶 厚生労働省健康局疾病対策課

第一部 研究分担者からご報告

10：10～10：20 ライソゾーム病医療費構造について
国立病院機構新潟病院 中島 孝

10：20～10：30 パーキンソン病の医療費分布
順天堂大学 頼高 朝子

10：30～10：40 炎症性腸疾患の生涯医療費の算出
東京医科歯科大学 長堀 正和

10：40～10：50 病院における難病医療費—入院と外来の解析—
新潟大学脳研究所 西澤 正豊

10：50～11：00 当科外来における医療費調査の中間報告
東邦大学医療センター大森病院 川合 眞一

11：00～11：10 療養形態別の医療費および患者自己負担額の調査方法に関する提案
脳血管研究所附属美原記念病院 美原 盤

- 11：10～11：20 有害事象にともなう追加的医療費についての研究
北里大学医学部医療安全・管理学 渋谷 明隆
- 11：20～11：30 生涯医療費の推計に関する試み
東北大学大学院医学系研究科 伊藤 道哉
- 11：30～11：45 Rを用いたシミュレーションによるサンプルサイズの算出
神奈川歯科大学内科学講座 森實 敏夫
- 11：45～11：55 難治性疾患の医療費分析の方法と解析に関する検討
東京医療保健大学 山下 和彦
- 12：00～13：00 昼食休憩

第二部 全体調査の報告

- 13：00～13：15 平成21年度難治性疾患の医療費構造に関する研究
北里大学医学部神経内科学 荻野 美恵子
- 1) 社会保険診療報酬支払基金とのやりとりについて
平成21年11、12月平成21年1月データ
- 2) 国民健康保険中央会とのやりとりについて
平成21年8月～10月分データ
- 13：15～13：30 平成21年度データによる粗集計（国保と基金のデータを中心として）
（株）健康保険医療情報総合研究所 山口 治紀

第三部 来年度の研究の進め方について

- 13：30～16：00 途中、休憩（コーヒーブレイク）有り
- 1) 今後のデータ収集方法について
- 2) 各疾患の医療費の精緻化について
- ①患者ベースの医療費調査
調査フォームの作成
- ②生涯医療費の算出
生涯医療費を考える際の疾患モデルケースについて
パーキンソン病 順天堂大学 服部信孝（頼高朝子）
多発性硬化症 美原記念病院 美原盤
重力筋無力症 北里大学 荻野美恵子
多系統萎縮症 慶應大学 鈴木則宏（高橋一司）

ALS 仙台往診クリニック 川島孝一郎
遺伝性脊髄小脳変性症 新潟大学 西澤正豊
ライソゾーム病 新潟病院 中島孝
潰瘍性大腸炎 東京医科歯科大学 渡辺守 (長堀正和)
クローン病 東京医科歯科大学 渡辺守 (長堀正和)
肝臓疾患 北里大学 渋谷明隆
SLE, RA 東邦大学 川合眞一
生涯医療費を算出する場合の方法論
一橋大学 荒井耕
東北大学 伊藤道哉
東京医科歯科大学 伏見清秀
産業医科大学 松田晋哉

16:00 閉会の挨拶

以上