

## 序章 統計学を学ぶ意義と サンプリングの重要性

### 序.1. 統計学を学ぶ意義

～統計学により何を明らかにするのか

#### A. 褥瘡問題の基礎研究

～看護研究の一例として

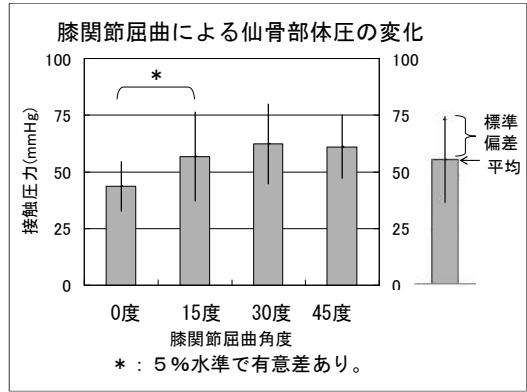
佐久総合病院の褥瘡対策研究会は、軽度の膝拘縮であっても、褥瘡の危険因子になりうるのではないかと考え、写真のようにシーネで固定することにより膝拘縮を模擬し、体圧を測定した。

仙骨部における体圧は、右図のようになり、膝角度0度と15度の間でかなりの差がみられた。

研究グループは、この差が医学的に意味のあるものかどうかを、『対応のあるt検定』という方法を用いて検討し、15度程度の膝拘縮でも仙骨圧が上昇するとの結論を出した。

以上のように、『対応のあるt検定』は、看護研究などに用いられることがしばしばである。

この解析方法は、第5章後半で紹介する。



#### B. 健康調査の結果を日本の平均と比較する

保健師さんがしばしば出会う問題である。

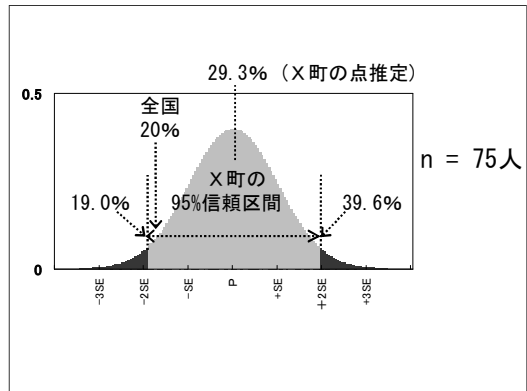
例えば、X町の保健師さんが、40才代の成人を適切に選んで、75人の血圧測定を行なってみたところ、22人が高血圧であったという結果が出た。

高血圧者割合は、 $22/75 = 0.293 \rightarrow 29.3\%$ であったことになる。一方、全国的調査では、20%であるという。

これを、単純に比較すれば、X町の高血圧者の割合は日本全国の平均より高いことになるが、75人程度の調査により断言してよいだろうか？

医学統計学には、『比率の検定』という方法があり、右上の図のように、X町の高血圧者の割合の『95%信頼区間』というものを算出し、全国平均がその範囲に入ると、『医学的に意味のある差ではなかった』などと考察する。

『比率の検定』については、第4章後半～第5章前半で紹介する。



C. ガンの原因、危険度を探る

近年の日本において、癌は、死亡原因の第1位であり、癌の原因を明らかにして、予防措置を講じることが必要であると、ほとんどの方が考えるところであろう。

このために、疫学的に調査したデータを解析して、要因と発癌の関係を評価するのが、統計学の役割の一つである。

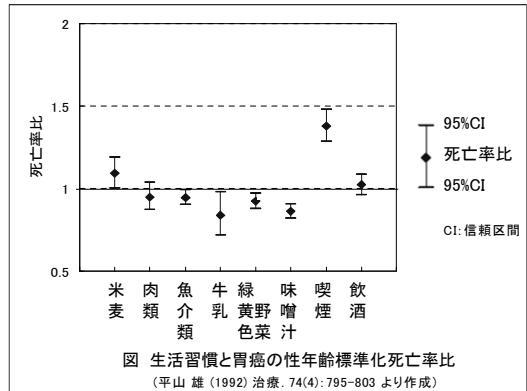
癌の原因、危険度を明らかにした調査結果の一例が、右の新聞記事である。



さらには右図のように、大規模な調査に基づき様々な要因を同時に勘案して、危険因子であるか抑制効果があるかを評価する解析方法も広く使用されている。

(平山雄(1992) 治療. 74(4):795-803)

この解析方法の一部を、第3章後半で紹介する。



D. 薬害を明らかにする (サリドマイドの場合)

1957年 に睡眠薬サリドマイドが販売されるようになると、世界各国で、短上肢を特徴とする奇形の子供が生まれるようになった。

西ドイツのウィドゥキント・レンツ(Widukind Lenz)博士は、奇形の子供を出産した母親と健常児を出産した母親がサリドマイドを服用したかどうかを調べ、右表のような結果を得た。このような結果を解析して、「奇形とサリドマイドの服用との間に関係がある」といえるかどうかを判定するのが、『 $\chi^2$ 検定』(カイ・ジ・ジョウ・ケンテイ)という統計手法である。

1961年11月に、奇形とサリドマイドの服用との間に関係があるとする、レンツ博士の警告により、世界のほとんどで、販売が中止された。

その後、サリドマイドによる出生障害が、大規模に追跡され、妊娠初期に妊婦が服用した場合、出生児に短上肢、聴力障害などの奇形・障害が発生することが明らかにされ、現在はサリドマイド胎芽症と呼ばれている。(片平洸彦「ノーモア薬害」(1997) 30p~)

『 $\chi^2$ 検定』については、第3章前半で紹介する。

**母親のサリドマイド剤服用と奇形児の出産との関係**

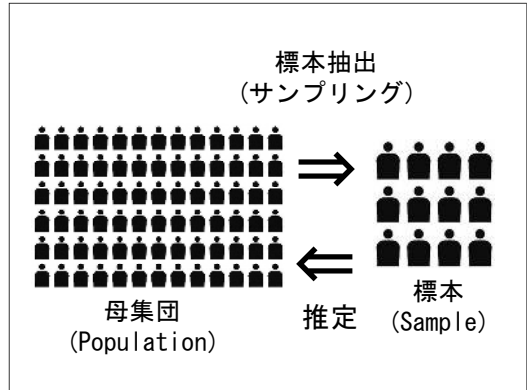
	奇形	非奇形	計
服用	90	2	92
非服用	22	186	208
計	112	188	300

レンツ博士の研究(1961)から

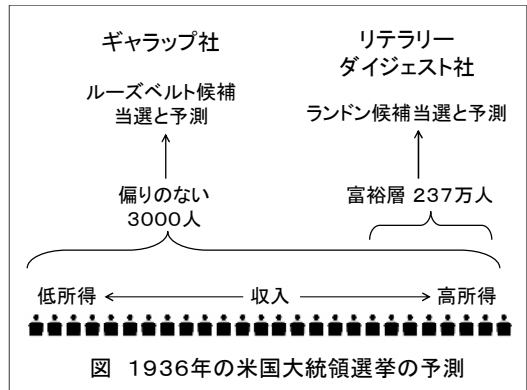
### 序.2.1. サンプルング(標本抽出)の重要性

A市の住民10万人について、糖尿病予備群の割合を調査するとして。10万人全てを調査すれば、正確な割合が出てくるが、膨大な費用と手間がかかり、現実的ではない。そこで、例えば100人を選んで調査することが、しばしば行なわれる。

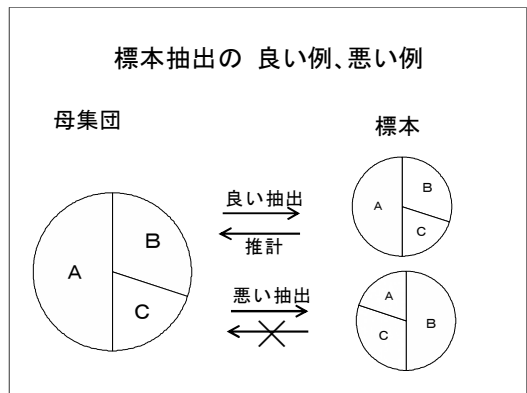
統計学では調査対象者を選ぶことを、サンプルング(標本抽出)といい、選ばれた調査対象者群を、標本(サンプル)と呼ぶ。



サンプルング(標本抽出)の良否が、調査結果を大きく左右した実例の一つ紹介させていただく。1936年のアメリカ大統領選挙は、ルーズベルト候補とランドン候補の対決であった。右図のように、ギャラップ社は偏りの少ないサンプルング方法で3000人を調査した。一方、リテラリー・ダイジェスト社(LD社)は237万人も調査したのだが、対象者が富裕層に偏っていた。ギャラップ社はルーズベルトの当選を予測し、LD社は予想を間違えてしまい、翌年に倒産したという。偏りのないサンプルングが重要であることが、お解かりいただけるだろう。

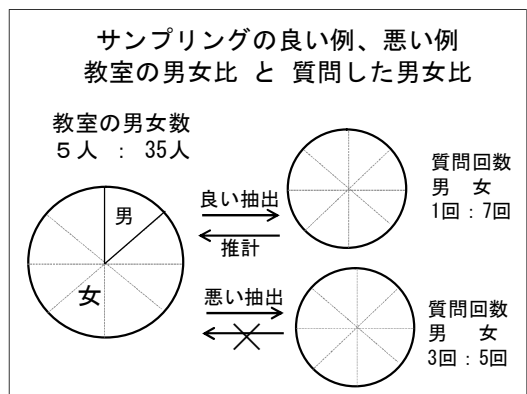


ある世論調査を考えてみよう。右のように、全体(母集団)としては、意見A : 50%、意見B : 30%、意見C : 20%であるが、世論調査のサンプルングが偏っていたため、意見A : 20%、意見B : 30%、意見C : 50%と、母集団を反映していない結果が出てしまったのでは、調査研究の意義を問われてしまう。



このようなことは、医学的調査でも起きうる事で、妥当なサンプルング(標本抽出)を行なうために、いろいろな方法が考案、使用されている。

これから、いろいろな統計解析の手法を学習してゆくが、もともとのサンプルング(標本抽出)の良否が、調査の成否を分けることを強調しておきたい。



<問題>  
右の円グラフに男女の領域を入れよ。

序.2.2. 外れ値をどう考えるか？

～新潟水俣病の場合

右図は、高齢者18人の収縮時血圧を測定した結果である。

この分布の場合、200mmHg以上の人が1人いるのだが、どのように考えたらよいだろう。

200mmHg以上の1人について、生活習慣などを調べ、血圧を低下させるために指導に力を入れることが重要である。

このような、集団から大きく離れた値を、医学統計学では「外れ値」(ハズレ)と呼ぶ。

そして、正常値の分布を把握するために、「異常値の除外」という考え方があり、200mmHg以上の1人を除外して解析が行なわれる場合もある。

この方法が、不適切に使用された例を右の図に示した。

四大公害病の一つ、新潟水俣病では、原因企業側の研究者が、右図のような分布があったときに、75ppm、187ppmは異常値として除外すべきであると無視してしまった。

このことについて、片平洸彦先生は、「自己に不都合なデータを学問的な装いをして抹殺しようとした」と指摘し、75ppm、187ppmを示した方々について、原因を究明し、対策を実行することが必要であると主張した。

(片平洸彦「やさしい統計学」(2000) 20p～)

