

13-1 条件が整えば可能です。

11-1 で二項分布は「玉数が多ければ正規分布に近似」することを、また 12-3 で標本から母集団における平均と標準偏差の推定が可能であることを指摘しました。これらと 12-3 で言及した「正規分布では標本平均が母平均の最良の推定値」であることを考え合わせると、『標本の平均値が正規分布すること』さえ仮定できれば、全体 (=母集団) の平均 (=母平均) が存在する範囲について部分 (標本) から一定の判断をすることが可能です。

13-2 まず、正規分布では

平均値±1.96 標準偏差の範囲内に全体の 95%が存在します。従って、ある方法で得られた標本平均値の集団が正規分布するならば、その集団の平均値±1.96 標準誤差 (=標本平均値の分布の標準偏差) の範囲内には「その方法で得られる全ての標本平均値」の 95%があるはず。他方、母平均の最良の推定値は個々の標本平均値ですから、上、(=ある標本平均値±1.96 標準誤差) の範囲内には真の「母平均」が 95%の確率で存在することが期待できます。この範囲が「母平均の 95%信頼区間」です。

※標準誤差=不偏標準偏差/√ 標本の大きさ

そこで今回は、表計算ソフトの列 data から、標本平均値の正規分布を前提として母平均の 95%信頼区間を求める script を作成します。

```
<title>母平均の信頼区間@大標本法</title>
<!-- saved from url=(0008)about:internet -->
<form>
  <textarea rows="9" cols="28"></textarea><br>1. 一列@表計算ソフトを paste※
  →2.<input type="button" value="OK" onClick="OK()">を click<br>
  ※最終 data 以降の改行は欠測とみなされるため要削除。
</form>
<script>
  素="得点%1%2%3%4%5%6%7%8%9%10";document.forms[0].elements[0].value=素
  function OK() {
    素=document.forms[0].elements[0].value.replace(/%r%n/g, "").split("%n")
    名=素.shift();有=0;欠=0;総和=0;二次=0;桁=1000
    for (i in 素){if(素[i].replace(/%s/g, "")=="") {欠++;continue}
      有++;値=(素[i])-0;総和=総和+値
    }
    平均=総和/有
    for (i in 素){if(素[i].replace(/%s/g, "")=="")continue
      差=素[i]-平均;二次=二次+差*差
    }
    SD=Math.sqrt(二次/有);SE=Math.sqrt(二次/(有-1))
    SE=SE/Math.sqrt(有);p=new Array(95,99,90);z=new Array(1.96,2.576,1.645)
    r0="%n 不偏標準偏差="+Math.round(桁*素)/桁+"%n 標準誤差="+Math.round(桁*SE)/桁
    r0=r0+"%n ☆正規分布に基づく信頼区間:"
    for (i=0; i<3; i++){
      幅=z[i]*SE;r0=r0+"%n +p[i]+%":Math.round(桁*(平均-幅))/桁+" ~"
      r0=r0+Math.round(桁*(平均+幅))/桁+"%n · 標本平均 ± "+Math.round(桁*幅)/桁
    }
    平均=Math.round(桁*平均)/桁;SD=Math.round(桁*SD)/桁
    r1="変数: "+名+"%n 有効 "+有+"%n 欠測 "+欠+"%n = "+(有+欠)+"%n 平均="+平均
    document.forms[0].elements[0].value=r1+"%n 標準偏差="+SD+r0
  }
</script>
```

「丸め」の桁指定を一貫して簡潔に行うための変数。

空白%r%nを""に揃えてskin。

正規分布で確率が95,99,90%になる平均±xσの値が1.96,2.576,1.645。

13-3 列 data から SD を求める

ところまでは 12 と同様です。標本平均値の分布の標準偏差である標準誤差は (2 次の積率 / (件数-1)) / √ 件数 で求め、最も一般的な 95%に加え、90%と 99%についても ±1.645・2.576 の各値を用いて信頼区間を求めています。

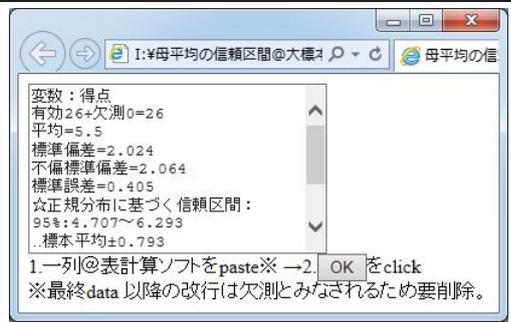
13-4 不偏標準偏差と標準誤差について少し説明すると

不偏標準偏差 (u) は標本の標準偏差 (SD) から推測した母集団の標準偏差で、SD が平均の周りの 2 次の積率を「件数」で割った値の平方根であるのに対し、「件数-1」で割って求めます。1 を引く理由については「件数∞の母集団に対し件数有限の標本には極端な値が含まれにくい傾向の補正」という説明 (?) が納得しやすいのでは。

また、標準誤差 (SE : Standard Error) は一般的には「推定量の分布の標準偏差」です (標本平均値は母平均値の推定量)。母標準偏差 (σ) が分かっているならば σ / √ 件数 で求めますが、一般に σ は未知なのでその推定値の u を使って u / √ 件数 で求めます。式が示すとおり、件数が多いほど標準誤差は小さくなり、信頼区間は狭く (=推測が特定の) になります。

【発展課題】

「標本の平均値が正規分布する」(13-1) 条件について考えて & 調べてみましょう。



資料 33 資料 32 の内容の実行結果の一例

