区間推定入門

きむあき

推測統計の要

- 点推定と区間推定

- 点推定は、未知の値を一点で推測すること

＞例　標本平均から母平均を推定する（ただ1つの値で）。これは標本平均の値をそのまま当てはめるというべたな方法でOK

・区間推定は、未知の値を、一定の確率で取りうる幅をもって推測することが求められる。やや　手順が多い。でも、それだけ。

これらの作業には、前提としていくつかの仮定が用いられる。

それが、次に示すもの・・・

何のこと？

- 繰り返すとすべての自然に得られるデータは正規分布になる

（大数の法則、中心極限定理）

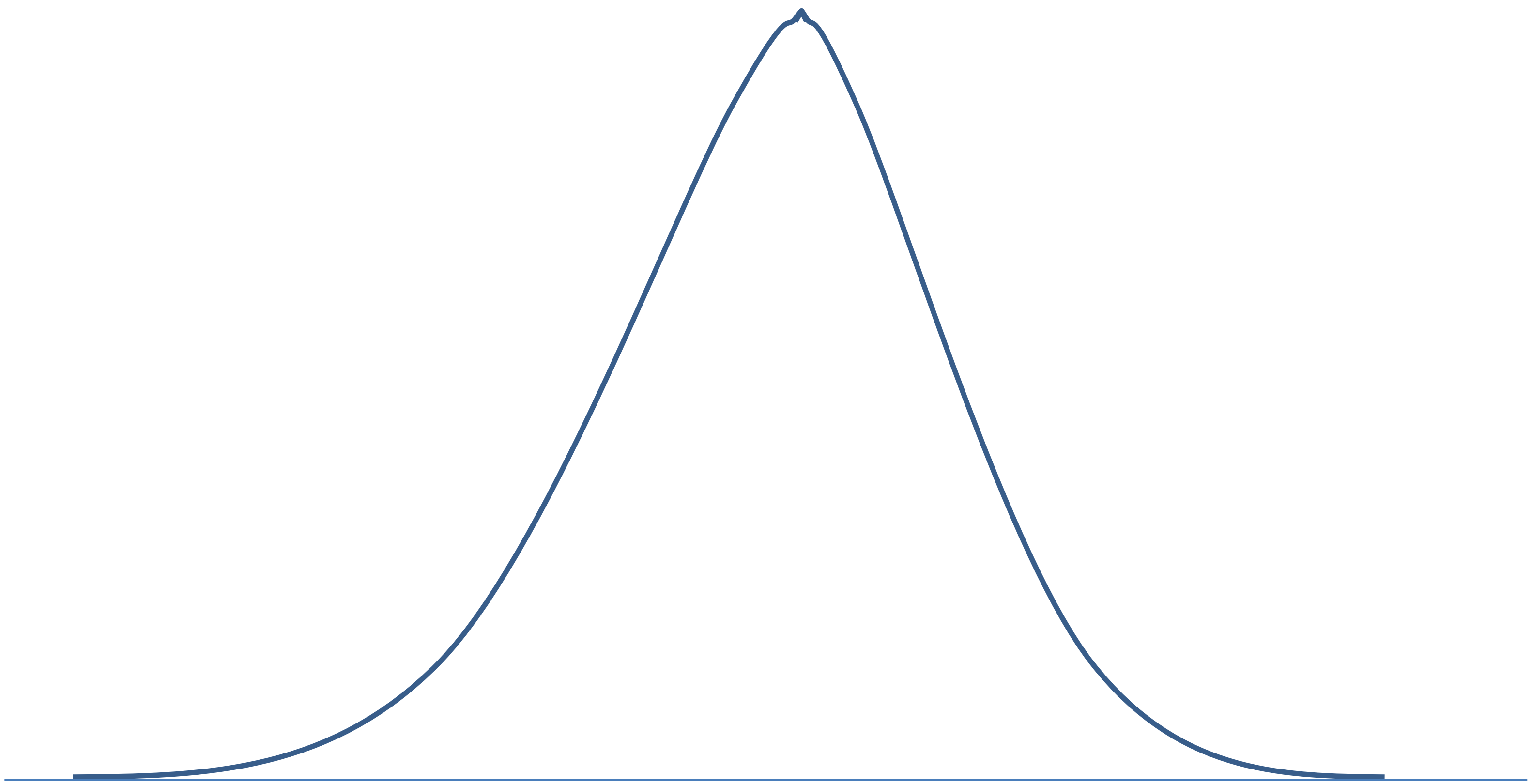
- この性質を利用して、サンプルからその母平均を仮定して、その取りうる範囲を調べること

（確率分布）

無作為（正しいデタラメ）こそ自然

自然、理想の数はその集合を集めてその単位の大きさを小さい方から大きい方に並べて、縦軸に頻度をとれば、正規分布という形になる。

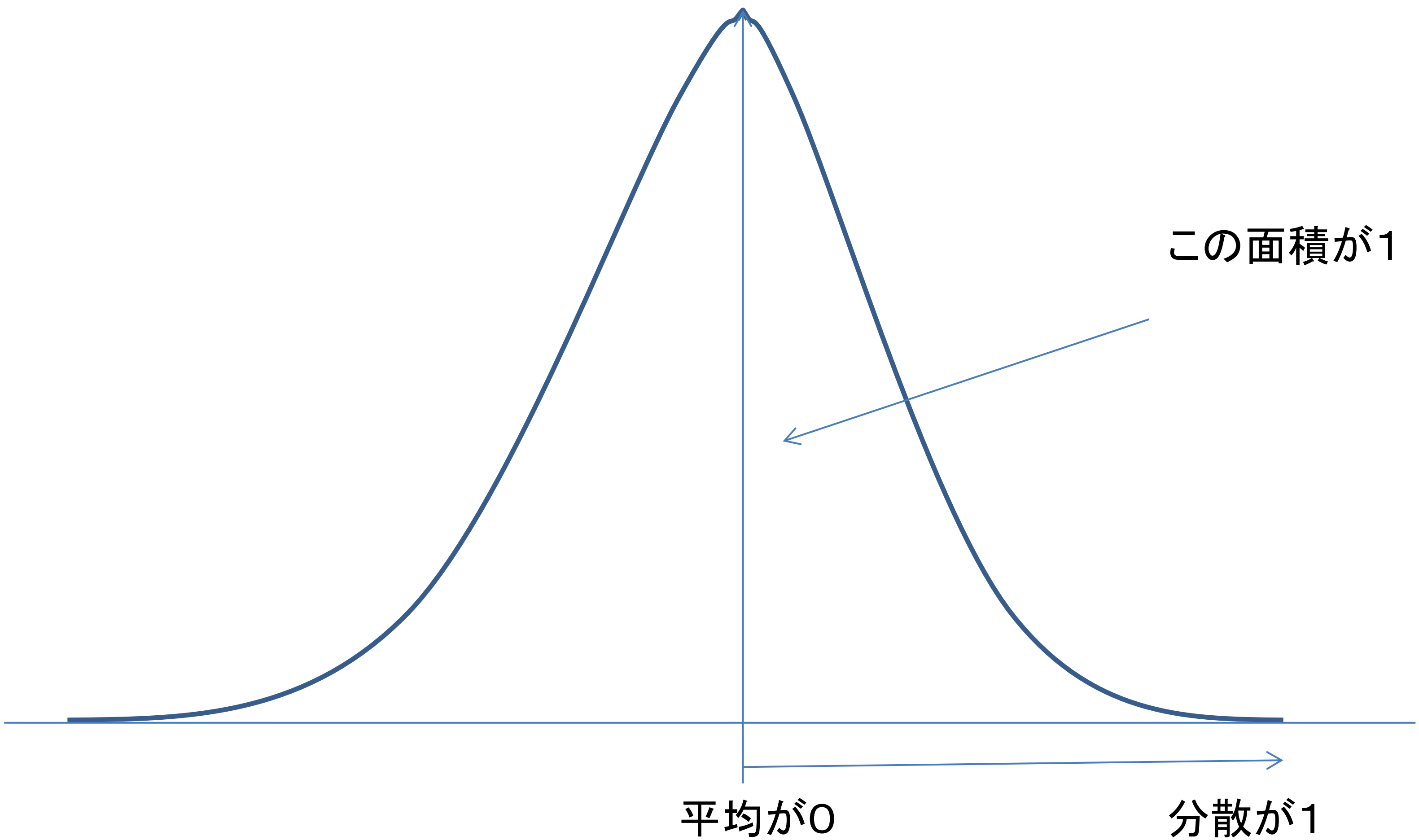
- ここで正規分布についておさらいしよう



頂上がただ一つあって、左右に均等な裾（すそ）がある山の形（データの分布で描かれる）のことだな

で、正規分布曲線で囲まれた部分は確率を表すんだね。これを1として考えると標準正規分布曲線になるんだね。この面積は0～1になるということだ。これが危険率5％＝0.05（この面積の5％）ということを表す基本的な考え方なんだな。

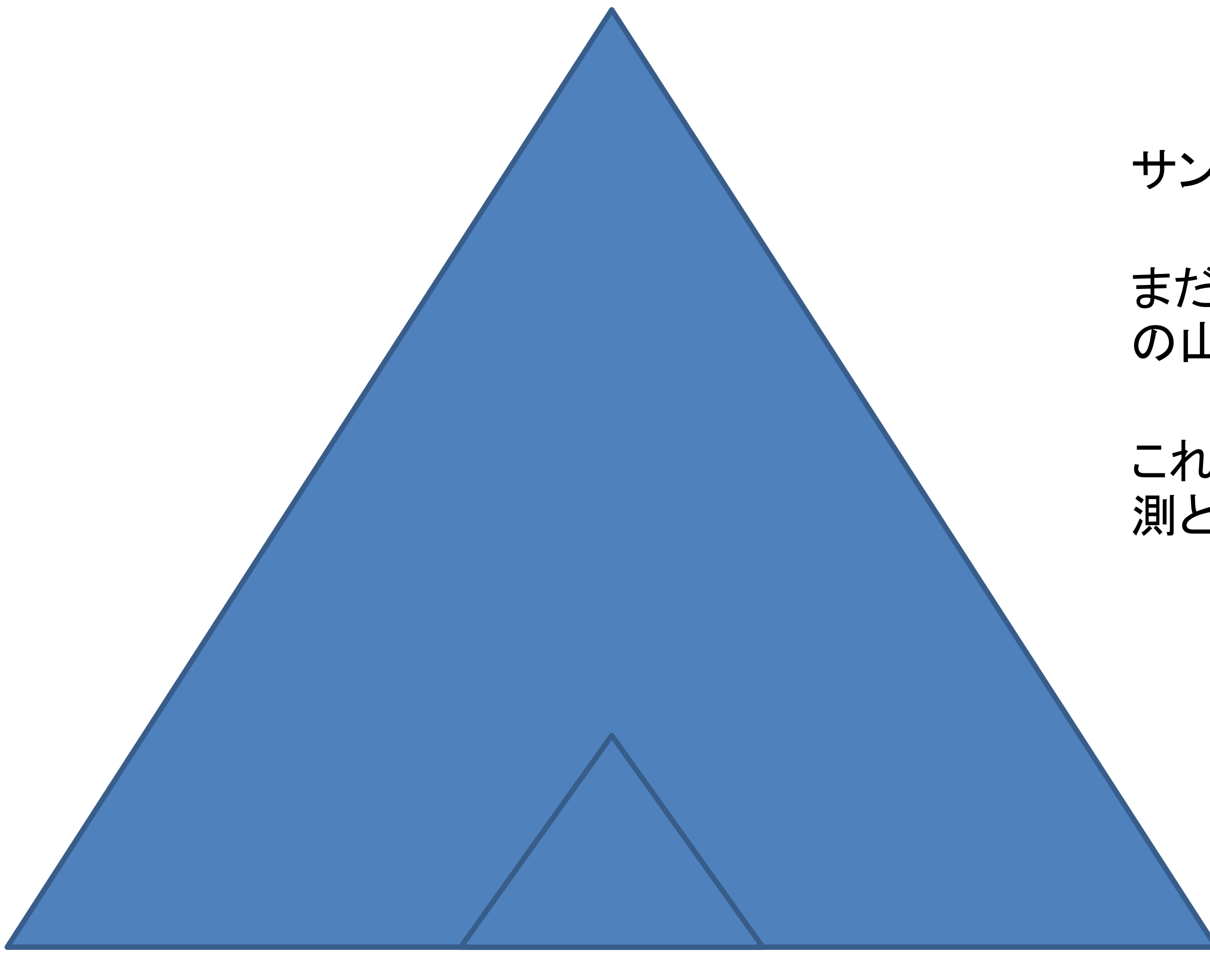
- さらに標準正規分布というものを知っておこう



- 以上　準備終了

- そして30例以下のデータなんかで、母集団の平均＝母平均が、誤差を考慮した場合＝誤差が正規分布すると仮定した場合、
- サンプル＝標本の平均から推定した母平均の取りうる値を示すことを区間推定と呼ぶのだ。

区間推定は、データの推定

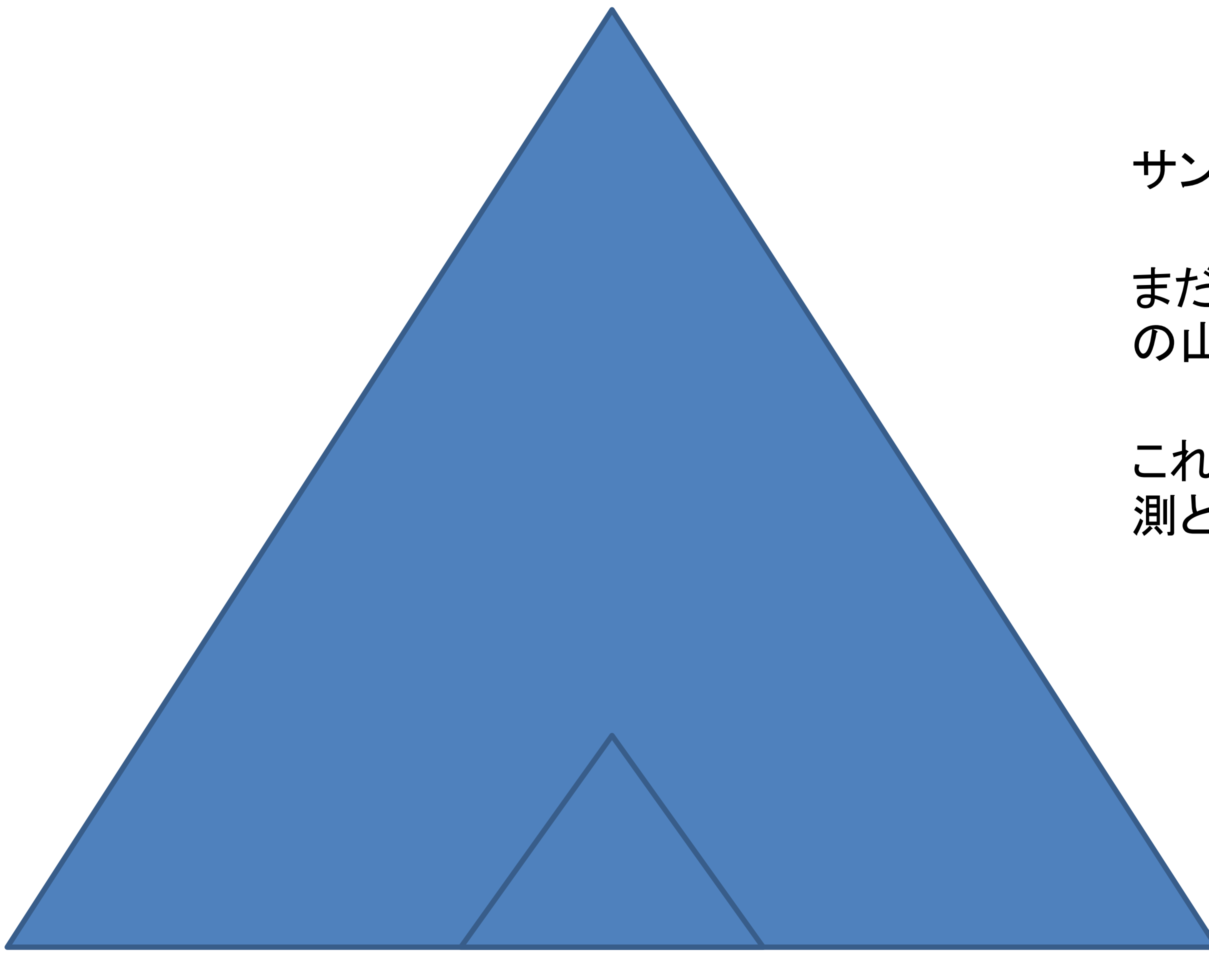


サンプルの山から

まだ見ぬ母集団の山を描くこと

これをデータの推測といった

サンプルの山の頂上と母集団の山の頂上が一致するならば、苦労なし



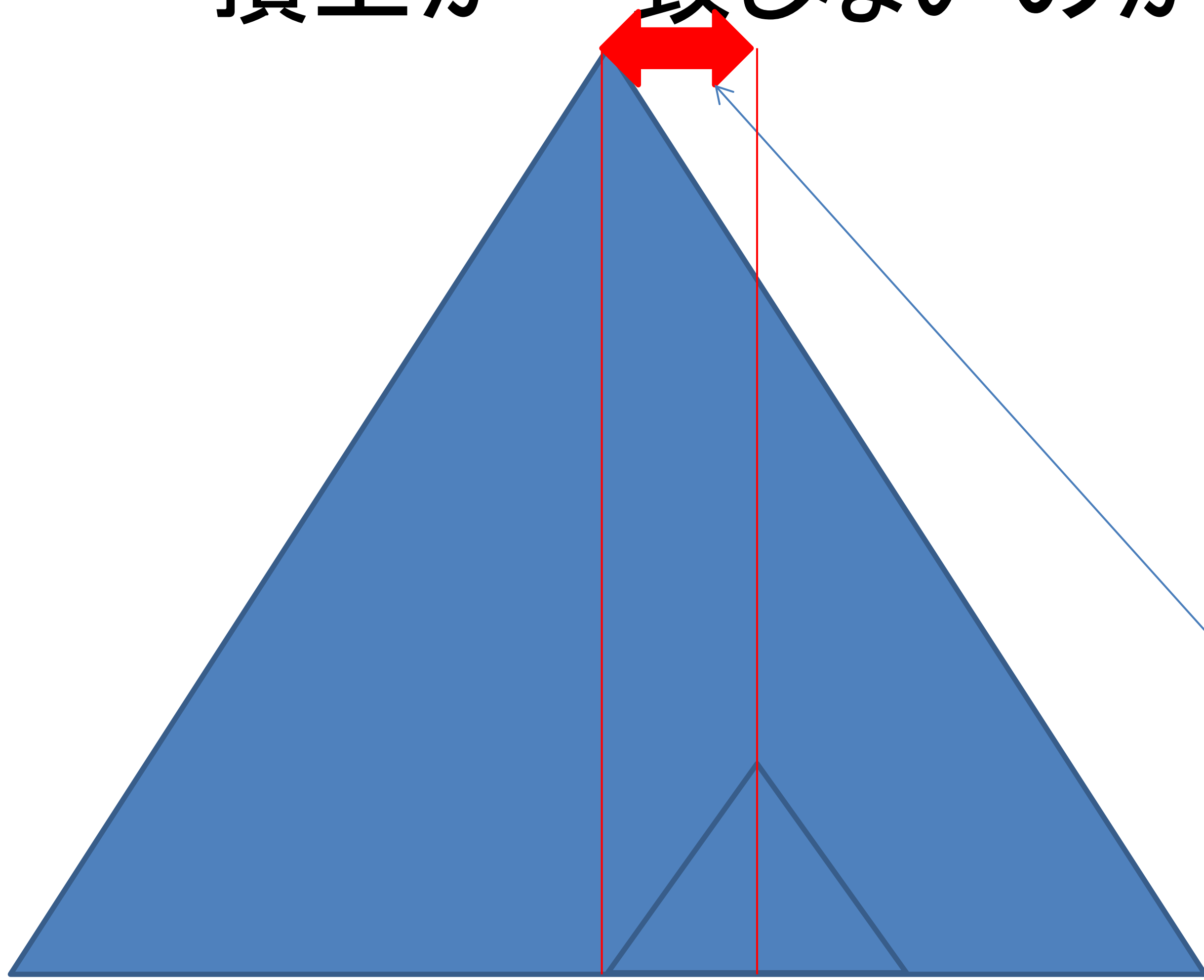
サンプルの山から

まだ見ぬ母集団の山を描くこと

これをデータの推測といった

これを理想の数の山（分布）と考える

サンプルの山の頂上と母集団の山の頂上が一致しないのが現実



サンプルの山から

まだ見ぬ母集団
の山を描くこと

これをデータの推
測といった

現実は、こんだ
けずれてる！

これを現実の数の山（分布）と考える

**誤差を含
むのさ**

そこで

知りたいのは未知の母集団の平均値（頂上）

（サンプルの頂上のx座標値）－
（母集団の頂上のx座標値）

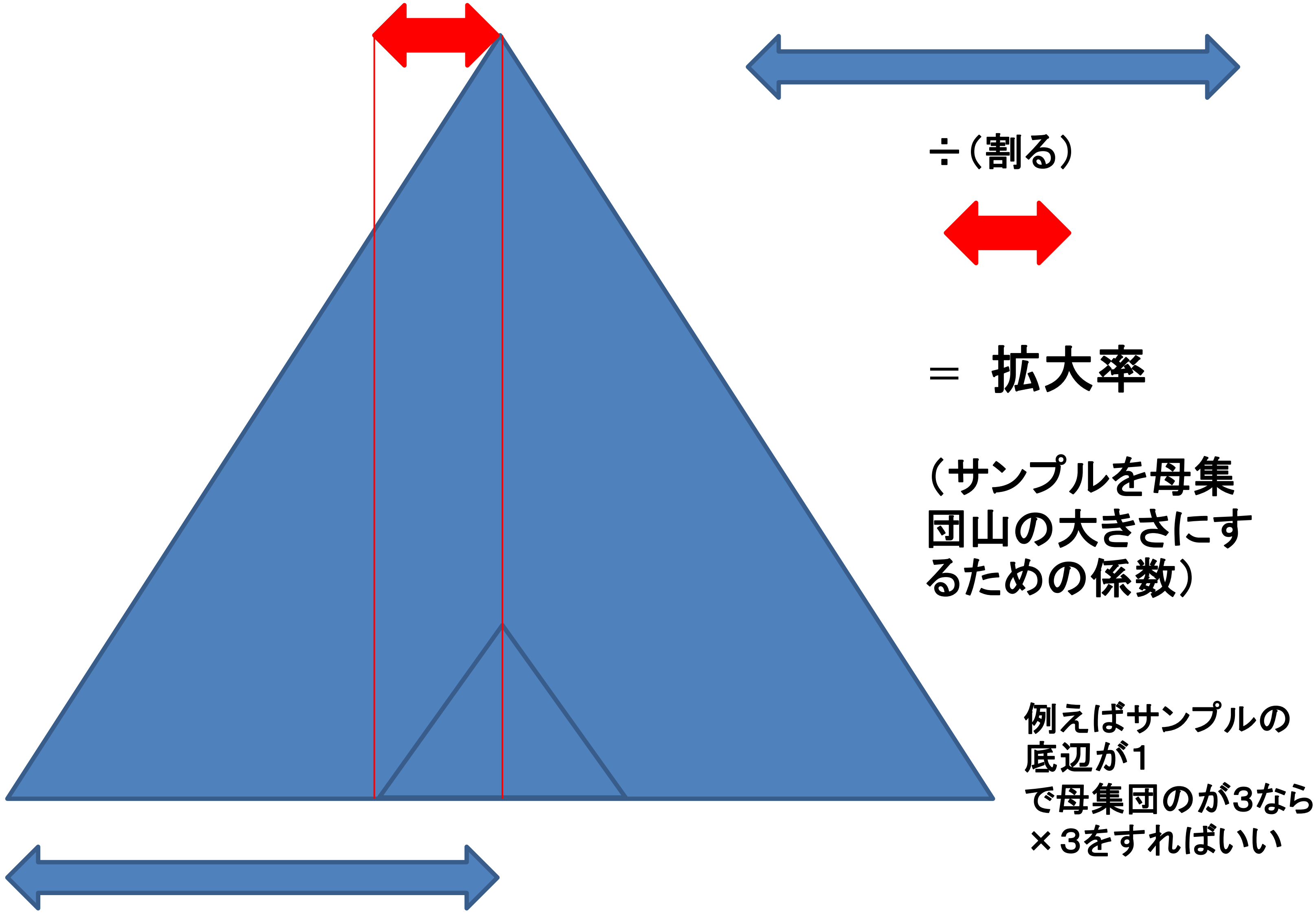
↓

頂上のずれ

誤差がもたらす、幅を知りたいわけだ

山の頂上と底辺が山の形を決める

サンプルの山の底辺と母集団の山の底辺の割合が拡大率



頂上のずれ
を
拡大率で割る

サンプルの山の底辺と理想の山の底辺のデータ1個当たりの割合で割った値

が確率△の面積の95%に当たる位置を求めるのが信頼区間の推定となる

- ここでいう拡大率を支えているのは確率分布の考え方であり、中心極限定理であり、大数の法則である。

スッキリ

９５%信頼区間はこうなる

４７．５%の面積を占める部分の一番右側の底辺のx座標の値

頂上のずれ
を
拡大率で割る
（ただしデータ1個当たりの値で計算してね）

４７．５%の面積を占める部分の一番右側の底辺のx座標の値

—1.96（　標準正規分布の場合）

—2.25（　t分布の場合）

1.96 （標準正規分布の場合）

2.25 （t分布の場合）

これが標準正規分布を仮定した場合の母平均の95%信頼区間

- －１．９６＜｛（xbar－μ）/（√（s^2/n）｝＜１．９６

注　t分布を仮定した場合には、±1.96部分が2.25になる

- －1．96＜{（xbar－μ）/（√（s^2/n）}＜1．96

は2つの不等式を合わせた表現であるから

－1．96＜{（xbar－μ）/（√（s^2/n）}

と

{（xbar－μ）/（√（s^2/n）}＜1．96

にわけて計算すればよい

Xbarは標本の平均

S^2は不偏分散（n-1に注意）

nは標本の個数

計算機を使うときは小数点第3位を切り捨てて、2位で計算を続けよう。（便宜的にorこの試験時は）

- 例題
- ｛4、1、2、6｝の95％信頼区間を求めよ。

（ただし、このデータの分布は正規分布に従うとする。答えは小数点第2位まで書いてください。）

μの不等式を解くことと同じ

手順

- ①　xbarを求める

{　　　}の中の数字を足して、個数で割る

- ②　偏差を求める

各データからxbarを引く

- ③　偏差平方和を求める

　②のデータをそれぞれ2乗して足し合わせる

- ④　不偏分散をもとめる

③のデータを個数nから1引いた値で割る　これをまたnで割ってルートを計算する。＞これをt統計量と呼ぶ。

- ⑤　これ（t統計量）を不等式に代入してμの範囲＜＞を求める

1.96をt統計量に掛けて、標本平均に足すのと、引く作業を行うと解が導かれる。