

百均の電卓で解ける 統計入門

数学的な厳密さよりも直感的に理解することに
重点を置いていきます・・・

全体を推理しよう!

ちょっと、復習

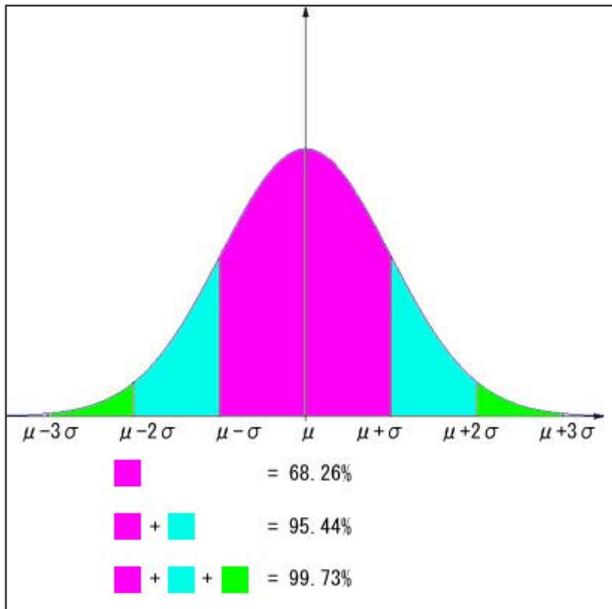
前号まで、あるデータ群が正規分布である時、 と  を合わせた範囲に入れば、その予言は 95% 当たるであろうことを説明してきました。

平均値が μ で S.D. が σ の一般正規分布の「95%的中」する範囲は、

$$(\mu - 1.96\sigma) \text{ 以上 } (\mu + 1.96\sigma) \text{ 以下}$$

となることを覚えていませんか？

平均が 60 点の全国テストで、S.D. が 15 だった場合、30.6 点から 89.4 点の間に 95% の人が入るということでしたね。上位 5% の人を優秀と評価するとした場合、90 点取れば、優秀と言っても良いですね。



逆に考えることも・・・

では、上位 5% の人を優秀と評価すると、90 点では優秀と評価できない場合を考えてみましょう。

まず、思いつくのが、平均点が 62 点以上の場合ですね。95% の上限が 91.4 点になってしまいますから、90 点の人は 95% に含まれてしまいます。

また、S.D. が 16 ではどうでしょう。同じように 95% の上限が 91.36 点になってしまいますから、90 点の人は 95% に含まれてしまいます。

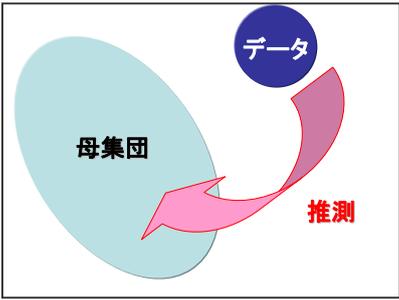
このように、同じデータでもそれが 95% に入るか、入らないかで全体像を推測することができるのです。

全体を考える

例えば、コインを投げる実験をして、表が 10 枚出ました。この時、投げたコインの枚数は何枚かを考えてみましょう。

いままでと同じように、95% 予言が当たる区間を考えるのです。当然、10 枚投げて 10 枚表が出るのは、考え難いですよね。

コイン投げについては、10 月号でコインを N 枚投げて、表の出た数を記録していきます。これを膨大な回数繰り返し、表の得た数で相対度数のヒストグラムを作ると、



$$\text{平均値} = \frac{N}{2}$$

$$\text{S.D.} = \frac{\sqrt{N}}{2}$$

の正規分布となることを書きました。

表の数が 10 枚の時、 $(\mu - 1.96\sigma)$ 以上 $(\mu + 1.96\sigma)$ 以下を満たすのは・・・

地道に当てはめていくと、母集団は 13 枚以上 30 枚以下となります。

この 13 枚以上 30 枚以下のような、「95% ありうる母数の入る区間」を「95% 信頼区間」と呼び、母数をこのような区間で推定することを「区間推定」と呼びます。

標準偏差のわかっている正規母集団の、平均に対する区間推定

前例では、コイン投げという平均値、S.D. の計算式が決まっているケースで説明しました。

しかし、よく遭遇するケースは標準偏差のわかっている正規母集団の、平均に対する区間推定です。例をあげてみましょう。

【例】自分の血圧を測った時のデータを x とします。このデータは、実際の血圧を平均 μ として、標準偏差 6 の正規分布をします。今、血圧の測定値が 130 とした場合のあなたの実際の血圧の範囲をどのように推定すればいいのでしょうか。95% 信頼区間を求めてみると・・・ $(\mu - 1.96\sigma)$ 以上 $(\mu + 1.96\sigma)$ 以下は・・・

$$\begin{aligned} -1.96 &\leq (x - \mu) / \delta \leq +1.96 \\ -1.96 &\leq (130 - \mu) / 6 \leq +1.96 \\ -11.76 &\leq 130 - \mu \leq +11.76 \\ 118.24 &\leq \mu \leq 141.76 \end{aligned}$$

となります。

【町田幸雄】

この項続く