

標本の大きさの決め方

区間推定の考え方をえば、95%または99%信頼区間においてある決まった標準誤差（標準偏差）になるような、標準サイズを計算して求めることができます。

母平均の区間推定から標本の大きさを決める

母平均の区間推定の場合で、標本の大きさの決め方を考えます。

標本の大きさの求め方の考え

母分散が既知の場合は、正規分布を利用して、母平均 μ を推定できます。信頼度 $100(1-\alpha)\%$ の信頼区間は次のとおりです。

$$\bar{x} - z_{(\alpha/2)} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{(\alpha/2)} \frac{\sigma}{\sqrt{n}}$$

ここで、母平均と標本平均との誤差の限度を E とします。

$$E = z_{(\alpha/2)} \frac{\sigma}{\sqrt{n}}$$

この式を、標本サイズ n について解くと、次のようになります。

$$n = \left(\frac{z_{(\alpha/2)} \sigma}{E} \right)^2$$

これで、標本平均が $\bar{x} \pm E$ になる、標本サイズを n を求めることができます。

例題

ある県の大学生全体の1か月の平均の生活費について調査して推定するとき、全体の標準偏差が9,000円であれば、誤差を1,000円以内に95%信頼区間で推定するためには、何人に調査すればよいか。

1. 標本の大きさを n とする。
2. 誤差の限度 $E = 1000$ とすると、標本の大きさ n は、次のようになると考えられる。

$$\begin{aligned} n &= \left(\frac{z_{(\alpha/2)} \sigma}{E} \right)^2 \\ &= \left(\frac{1.96 \times 9000}{1000} \right)^2 \\ &= (17.64)^2 \\ &= 311.1696 \end{aligned}$$

つまり、312人の大学生に調査をすればよいことになります。

母比率の区間推定から標本の大きさを決める

母比率の区間推定の場合で、標本の大きさの決め方を考えます。

標本の大きさの求め方の考え

標本の大きさを n が十分大きい場合、正規分布による近似で母比率 P を推定できます。信頼度 $100(1-\alpha)\%$ の信頼区間は次のとおりです。

$$p \leq \bar{p} \pm z_{(\alpha/2)} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

ここで、母比率と標本比率との誤差の限度を D とします。

$$D = z_{(\alpha/2)} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

この式を、標本サイズ n について解くと、次のようになります。

$$n = \bar{p}(1-\bar{p}) \left(\frac{z_{(\alpha/2)}}{D} \right)^2$$

これで、標本比率が $\bar{p} \pm D$ になる、標本サイズを n を求めることができます。

例題

2010年6月29日に生中継された、サッカーのワールドカップの南アフリカ大会決勝トーナメント1回戦、日本対パラグアイ戦の平均視聴率は関東地区57.3%だった。この視聴率が、95%信頼区間において「57.3 ± 2%」となるような、標本サイズを求めよ。

1. 標本の大きさを n として、ある程度大きいと仮定する。
2. 標本数が大きい場合、正規分布による近似で母比率を推定できるを利用すると、標本サイズ n は、次のようになると考えられる。

$$\begin{aligned} n &= \bar{p}(1-\bar{p}) \left(\frac{z_{(\alpha/2)}}{D} \right)^2 \\ &= 57.3(1-57.3) \left(\frac{1.96}{0.02} \right)^2 \\ &= 2349.82 \end{aligned}$$

つまり、「57.3 ± 2%」で視聴率を調査するには、標本の大きさが**2350**以上である必要があります。

ちなみに、某視聴率調査会社の場合、関東地区で視聴率を調査する機械を取り付けてあるテレビの台数は、600台といわれています。