レポート提出票

科目名: 情報工学実験2

実験テーマ: 実験テーマ4 統計的推測と単回帰分析

実施日: 2020年 9月 21日

学籍番号: 4619055

氏名: 辰川力駆

共同実験者:

1　はじめに

推定と検定の考え方と手順を理解することを目標とする。

2　目的

1. 推定と検定の考え方と手順

   推定と検定の目的、考え方、手順を理解する。

2. t 分布

   t 分布の定義とその性質を理解する。

3. 母分散が未知の場合の母平均の推定と検定

   母分散が未知の場合の母平均の推定と検定の考え方と手順を理解する。さらに、推定と検定の関係を理解し、各方法の特徴を理解する。

3　実験方法

3.1　実験 1 推定と検定の考え方と手順

1. 平均 $\mu = 20.0$, 分散 $\sigma^2 = 0.4^2$ の正規分布 $N(20.0, 0.4^2)$ に従う乱数データを 10 個 × 1000 セット 発生させる。

2. 10 個の乱数データの平均値 $\bar{X}_s = \frac{\Sigma_{i=1}^{n} X_{is}}{n}$ をそれぞれ (s ごとに) 計算する。

3. 1000 個の $\bar{X}_s$ の平均値 $\bar{X} = \frac{\Sigma_{s=1}^{S} \bar{X}_s}{S}$ と次式の分散を求める。

$$\frac{\Sigma_{s=1}^{S}(\bar{X}_s - \bar{X})^2}{S-1}$$

   $\bar{X}_s$ が理論的に平均 20.0, 分散 $\frac{0.4^2}{n}$ の正規分布に従うことを踏まえて、結果を考察する。

4. 次式に基づいて、μ に対する 95 % 信頼区間をそれぞれ計算する。分散は σ^2 は既知 $(= 0.4^2)$ とする。

$$\text{下限} : \bar{X}_s - z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}}, \quad \text{上限} : \bar{X}_s + z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}}$$

   $z_{\alpha/2}$ は標準正規分布の上側 $\alpha/2$ % 点である。ここで $\alpha = 0.05$ とする。

5. 95 % 信頼区間に母平均 $\mu = 20.0$ が含まれる回数と割合を調べて、結果を考察する。

6. 両側検定における帰無仮説を $H_0 : \mu = \mu_0 = 20.0$, 対立仮説を $H_1 : \mu \neq \mu_0 = 20.0$ とする。分散は σ^2 は既知 $(= 0.4^2)$ とする。10 個の乱数データから、帰無仮説 H_0 に対する次式の検定統計量 Z_s をそれぞれ計算する。

$$Z_s = \frac{\bar{X}_s - \mu_0}{\sqrt{\sigma^2/n}}$$

7. 検定の有意水準 α は 5 % とする。$|Z_s| > z_{\alpha/2}$ のとき、帰無仮説 H_0 を棄却する。帰無仮説 H_0 が棄却された回数と割合を調べて、結果を考察する。

8. 手順 5 と手順 7 の結果を班のメンバーで共有 (合計) して、結果を考察する。

9. 平均 $\mu = 20.1 + a$, 分散 $\sigma^2 = 0.4^2$ の正規分布 $N(20.1 + a, 0.4^2)$ に従う乱数データを 10 個 × 1000 セット 発生させて、手順 6,7 を再度行なう。ただし、$a =$ 学籍番号の下一桁/10 とする。

10. 手順 9 の結果を班の他のメンバーの結果と比較して、その違いを考察する。

3.2 実験 2 t 分布

1. 平均 $\mu = 0$, 分散 $\sigma^2 = 1$ の正規分布 $N(0, 1)$ に従う乱数データを $c = b + 4$ 個 × 1000 セット発生させる。b は学籍番号の下一桁とする。

2. 最初の $\phi = c - 1 = b + 3$ 個の乱数データ $X_{ij}(j = 1..., \phi)$ から、次式により、自由度 ϕ のカイ二乗分布に従う乱数データを 1000 個作成する。

$$V_i(\phi) = X_{i1}^2 + \cdots + X_{i\phi}^2$$

3. 最後の c 番目の乱数データを U とする。式 (1) から、自由度 ϕ の t 分布に従う乱数データを 1000 個作成する。

$$T = \frac{U}{\sqrt{V/\phi}} \tag{1}$$

4. 得られた乱数データの上側 2.5 % 点を計算し、理論値と比較する。

3.3 実験 3 母分散が未知の場合の母平均の推定と検定

1. データシートから、学籍番号の下一桁に対応したデータ (7 個) を抽出する。

2. 7 個のデータを基本統計量 (データ数、平均、標準偏差、最小値、最大値) を計算する。

3. 次式に基づいて、母平均 μ に対する 95 % の信頼区間を求める。

$$\text{下限} : \bar{X} - t_{\alpha/2,v}\sqrt{\frac{s^2}{n}}, \quad \text{上限} : \bar{X} + t_{\alpha/2,v}\sqrt{\frac{s^2}{n}}$$

ここで、

$$\bar{X} = \frac{\Sigma_{i=1}^{n} X_i}{n}, \quad s^2 = \frac{\Sigma_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$$

である。$t_{\alpha/2,v}$ は自由度 $v = n - 1$ の t 分布の上側 $\alpha/2$ % 点である。ここで $\alpha = 0.05$ とする。

4. 両側検定における帰無仮説を $H_0 : \mu = \mu_0 = 12.5$, 対立仮説を $H_0 : \mu \neq \mu_0 = 12.5$ とする。7 個のデータから、帰無仮説 H_0 に対する次式の検定統計量 T を計算する。

5. 両側検定における P 値 を計算する。

6. 帰無仮説 H_0 が棄却されるかどうか判定して、検定の結果を解釈する。

7. 信頼区間と検定の結果にどのような関係があるかを数式に基づいて考察する。

4 結果・考察

4.1 実験 1 推定と検定の考え方と手順

1. Excel で乱数データを発生させた。

2.

3. 1000 個の $\bar{X}_s$ の平均値 $\bar{X} = \frac{\Sigma_{s=1}^{S}\bar{X}_s}{S}$ と次式の分散を求める。

$$\frac{\Sigma_{s=1}^{S}(\bar{X}_s - \bar{X})^2}{S-1}$$

$\bar{X}_s$ が理論的に平均 20.0, 分散 $\frac{0.4^2}{n}$ の正規分布に従うことを踏まえて、結果を考察する。

4. 次式に基づいて、μ に対する 95 % 信頼区間をそれぞれ計算する。分散は σ^2 は既知 $(= 0.4^2)$ とする。

$$\text{下限} : \bar{X}_s - z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}}, \quad \text{上限} : \bar{X}_s + z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}}$$

$z_{\alpha/2}$ は標準正規分布の上側 $\alpha/2$ % 点である。ここで $\alpha = 0.05$ とする。

5. 95 % 信頼区間に母平均 $\mu = 20.0$ が含まれる回数と割合を調べて、結果を考察する。

6. 両側検定における帰無仮説を $H_0 : \mu = \mu_0 = 20.0$, 対立仮説を $H_1 : \mu \neq \mu_0 = 20.0$ とする。分散は σ^2 は既知 $(= 0.4^2)$ とする。10 個の乱数データから、帰無仮説 H_0 に対する次式の検定統計量 Z_s をそれぞれ計算する。

$$Z_s = \frac{\bar{X}_s - \mu_0}{\sqrt{\sigma^2/n}}$$

7. 検定の有意水準 α は 5 % とする。$|Z_s| > z_{\alpha/2}$ のとき、帰無仮説 H_0 を棄却する。帰無仮説 H_0 が棄却された回数と割合を調べて、結果を考察する。

8. 手順 5 と手順 7 の結果を班のメンバーで共有 (合計) して、結果を考察する。

9. 平均 $\mu = 20.1 + a$, 分散 $\sigma^2 = 0.4^2$ の正規分布 $N(20.1 + a, 0.4^2)$ に従う乱数データを 10 個 $\times$ 1000 セット 発生させて、手順 6,7 を再度行なう。ただし、$a =$ 学籍番号の下一桁/10 とする。

10. 手順 9 の結果を班の他のメンバーの結果と比較して、その違いを考察する。

4.2 実験 2 t 分布

1. 平均 $\mu = 0$, 分散 $\sigma^2 = 1$ の正規分布 $N(0, 1)$ に従う乱数データを $c = b + 4$ 個 $\times$ 1000 セット発生させる。b は学籍番号の下一桁とする。

2. 最初の $\phi = c - 1 = b + 3$ 個の乱数データ $X_{ij}(j = 1..., \phi)$ から、次式により、自由度 ϕ のカイ二乗分布に従う乱数データを 1000 個作成する。

$$V_i(\phi) = X_{i1}^2 + \cdots + X_{i\phi}^2$$

3. 最後の c 番目の乱数データを U とする。式 (1) から、自由度 ϕ の t 分布に従う乱数データを 1000 個作成する。

$$T = \frac{U}{\sqrt{V/\phi}} \tag{2}$$

4. 得られた乱数データの上側 2.5 % 点を計算し、理論値と比較する。

4.3 実験 3 母分散が未知の場合の母平均の推定と検定

1. データシートから、学籍番号の下一桁に対応したデータ (7 個) を抽出する。

2. 7 個のデータを基本統計量 (データ数、平均、標準偏差、最小値、最大値) を計算する。

3. 次式に基づいて、母平均 μ に対する 95 % の信頼区間を求める。

$$\text{下限} : \bar{X} - t_{\alpha/2,v}\sqrt{\frac{s^2}{n}}, \quad \text{上限} : \bar{X} + t_{\alpha/2,v}\sqrt{\frac{s^2}{n}}$$

   ここで、

$$\bar{X} = \frac{\Sigma_{i=1}^{n} X_i}{n}, \quad s^2 = \frac{\Sigma_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$$

   である。$t_{\alpha/2,v}$ は自由度 $v = n - 1$ の t 分布の上側 $\alpha/2$ % 点である。ここで $\alpha = 0.05$ とする。

4. 両側検定における帰無仮説を $H_0 : \mu = \mu_0 = 12.5$, 対立仮説を $H_0 : \mu \neq \mu_0 = 12.5$ とする。7 個のデータから、帰無仮説 H_0 に対する次式の検定統計量 T を計算する。

5. 両側検定における P 値 を計算する。

6. 帰無仮説 H_0 が棄却されるかどうか判定して、検定の結果を解釈する。

7. 信頼区間と検定の結果にどのような関係があるかを数式に基づいて考察する。

5 まとめ

実験結果は，その性質をよく考えて，表 または 図 (グラフ) にする．グラフの場合は，縦軸や横軸が何を示すかを明記する．表の数値などは，有効数字に留意する．

6 感想

参考文献

[1] J. J. Collins et al., *PRE*, **52**(4):R3321, 1995.

[2] E. M. Izhikevich, *IEEE Trans. NN*, **14**(6):1569, 2003.