

統計学 課題その 7

担当: 社会情報科学部 山本 岳洋 (t.yamamoto@sis.u-hyogo.ac.jp)

締切: 6 月 25 日 (金) 21:00

解答は以下の URL にある Google フォームから提出すること.

<https://forms.gle/HKr2bhMAd2ajV1xY9>

講義資料「母集団と標本」や教科書・参考書を学習し、本課題に取り組むこと.

問1.

ある調査によると、神戸市の高校生の平日睡眠時間は 400 分、標準偏差 80 分である. いま、神戸市の高校生を母集団とし、その睡眠時間が正規分布 $N(400, 80^2)$ に従うと仮定しよう. このとき、以下の問い a., b. に答えよ. 単位は分とし、小数点第 1 位まで求めればよい. (なお、このデータは架空のデータです. 全国の高校生の平日睡眠時間は 420 分だそうです. 参考:[NHK, 2015 年国民生活時間調査](#))

必要があれば以下の確率を用いよ. Excel や標準正規分布表は用いないこと.

確率変数 Z が標準正規分布 $N(0,1)$ に従うとき、以下が知られている

$$P(-1.96 \leq Z \leq 1.96) = 0.95$$

$$P(-2.58 \leq Z \leq 2.58) = 0.99$$

- a. 神戸市の高校生の 95% はどのくらい睡眠時間をとっているかを求めよ. すなわち、母集団におけるある学生の睡眠時間を Y とするとき、 $P(a \leq Y \leq b) = 0.95$ となる a, b をそれぞれ求めよ.
- b. この母集団からサンプルサイズ 100 の標本を無作為復元抽出し、標本平均 \bar{X} を求める. このとき、次の 1. - 4. に答えよ.
 1. 標本平均 \bar{X} の期待値 $E(\bar{X})$ を答えよ.
 2. 標本平均 \bar{X} の標準偏差 $\sqrt{V(\bar{X})}$ を答えよ.
 3. $P(c \leq \bar{X} \leq d) = 0.95$ となる c, d をそれぞれ求めよ.
 4. サンプルサイズを 100 ではなく 10,000 とし標本を抽出し標本平均 \bar{X} を求めたとき、 $P(c \leq \bar{X} \leq d) = 0.95$ となる c, d をそれぞれ求めよ.

問2.

確率 $\frac{1}{4}$ で当選し $\frac{3}{4}$ ではずれるくじを無作為復元抽出で n 回引くことを考える。これは、母集団が二項分布 $B(1, 0.25)$ に従うとき、この母集団からサンプルサイズ n の標本 X_1, \dots, X_n を無作為復元抽出していると考えることができる。このとき、 X_i ($1 \leq i \leq n$) は引いたくじが当たりであれば 1, そうでなければ 0 となるような確率変数である。また、標本中の当選回数をサンプルサイズ n で割った変数、すなわち $\bar{X} = \frac{1}{n}(X_1 + \dots + X_n)$ を定義するとこれは標本平均の定義そのものである。これに関して、以下の問い a., b., c. に答えよ。なお必要であれば二項分布 $B(n, p)$ の期待値 $= np$, 分散 $= np(1 - p)$ であることを用いて良い。 答えは分数でよく、必要があれば n を用いてよい。

a. 母平均 μ を答えよ。

b. X_i の期待値 $E(X_i)$ と分散 $V(X_i)$ について答えよ。

ヒント: 先週の課題の問 4 と考え方は同じ。結局、 X_i が従う分布は...

c. $X_1 + \dots + X_n$ の期待値 $E(X_1 + \dots + X_n)$ と分散 $V(X_1 + \dots + X_n)$ を答えよ。

ヒント: 複数の考え方が存在します。1つの例: 結局、 $X_1 + \dots + X_n$ は $\frac{1}{4}$ で当たるクジを n 回引いたときの当選回数なので、 $X_1 + \dots + X_n$ が従う分布は...

問 3(自由記述):

今回の扱った内容について、特に説明してほしいところがあれば記述してください。また、講義への要望や感想、山本へのコメントなどもあればここに自由に記述してください。 **中間レポートの感想やコメントなどもあればここにどうぞ。** 書きにくい内容はメールや Teams のチャットでかまいません。

最後に、本課題を解く上で参考にした書籍（教科書・参考書以外）・ウェブサイト等があればここに記載してください。

以上.