科目名	数理統計学	担当者名	渕野　昌	所要時間	90 分	2014 年 1 月 28 日 13:20–14:50 施行
持込	すべて可					
添付する解答用紙	1　枚配付（問題用紙の回収　要・否）		計算用紙　0　枚配付			

I. K 市の家庭で飼われている 5 匹の犬の背丈 (普通の姿勢での頭頂までの高さ) を計ったところ，520mm, 470mm, 170mm, 430mm, 300mm という値が得られた．このデータについて，a) 平均，　b) 分散，　c) 標準偏差，　d) 不偏分散を求めよ．

II. O 市のあるアイスクリーム店でのある 7 日について，その日の日中の平均気温とアイスクリームの売上は右のようなものであった．このデータでの，a) 温度の平均，　b) 売上の平均，　c) 共分散，　d) 相関係数を求めよ．

日中の温度 (℃)	売上 (円)
14.2	21500
16.2	32500
11.9	18500
15.2	33200
18.5	40600
22.1	55200
25.1	61400

III. N 区の 6 つの家庭のそれぞれの飼い猫 6 匹について体重を計ったところ，この 6 匹の平均は 3.2kg で，分散は 4.8 だった．

(a) 家猫の体重は正規分布に従うと仮定して，N 区の家猫の体重の平均の 95 %での信頼区間を求めよ．

(b) (おまけの問題) 上の (a) での「家猫の体重は正規分布に従う」という仮定の妥当性について論ぜよ.「飼い犬の背丈は正規分布に従う」という仮定より「家猫の体重は正規分布に従う」の方が妥当性が高いと言えるか.

IV. 上の **III.** でのデータに基いて,「N 区の家猫の体重の平均は 3.8kg である」とう主張の検定を行なえ.

V. ある製品の重さ (mg) は正規分布に従い，母分散が 25 になることが知られている．この製品の重さの平均を 95 %の信頼区間の幅が 4mg 以内になるように推定するためには，最低何個のサンプルを調べる必要があるか.

VI. ある全国テストでのある問題の正解率は，0.9 であった．ある試験会場では 160 人の試験参加者のうち，正解者は 131 人だった．この試験会場でのこの問題の成績は全国平均に比べて有意に悪いと言えるか.

---

* 試験の実施後，この試験問題と解答例を

http://kurt.scitec.kobe-u.ac.jp/~fuchino/kobe/suuritoukeigaku-ws13-14-kimatsu.pdf

として掲示します.

解 答 用 紙

数理統計学 (2014 年 1 月 28 日)

学生番号: 名前:

I. a)： $\frac{520+470+170+430+300}{5} = 378$，平均： 378mm b)： $\frac{520^2+470^2+170^2+430^2+300^2}{5} - (378)^2 = 16136$，分散： 16136
c:） $\sqrt{16136} \approx 127$ 標準偏差： 127mm
d)： $\frac{5}{5-1} \times 16136 = 20170$，不変分散： 20170

II. a): $\frac{14.2+16.2+11.9+15.2+18.5+22.1+25.1}{7} = 17.6$ 温度の平均: 17.6 ℃，b) $\frac{21500+32500+18500+33200+40600+55200+61400}{7} = 37557.142857142855\cdots$ 売上の平均: 37557 円
c) $\frac{1}{7}\cdot(14.2\cdot 21500+16.2\cdot 32500+11.9\cdot 18500+15.2\cdot 33200+18.5\cdot 40600+22.1\cdot 55200+25.1\cdot 61400)-17.6\cdot 37557.142857142855 = 63101.42857142852\cdots$ 共分散: $63101.42857142852\cdots$
d): 温度の分散: $\frac{14.2^2+16.2^2+11.9^2+15.2^2+18.5^2+22.1^2+25.1^2}{7} - 17.6^2 = 18.44$
売上の分散: $\frac{21500^2+32500^2+18500^2+33200^2+40600^2+55200^2+61400^2}{7} - 37557.142857142855\cdots^2 = 222082448.97959208\cdots$
$\frac{63101.42857142852\cdots}{\sqrt{18.44}\cdot\sqrt{222082448.97959208}} = 0.9860560774312182\cdots$ 相関係数: ≈ 0.986; データからも一見して明らかだが，高い正の相関があることがわかる．

III. a): t-分布表の $\nu = 6-1$, $\alpha = 0.05$ での値 2.571 を用いて，$\left(3.2 - 2.571\frac{\sqrt{4.8}}{\sqrt{6-1}}, 3.2 + 2.571\frac{\sqrt{4.8}}{\sqrt{6-1}}\right) \approx (0.68, 5.72)$ つまり 0.68kg ∼ 5.72kg が 95 %の信頼区間となる．

もし $\alpha = 0.01$ を用いると，$\left(3.2 - 4.032\frac{\sqrt{4.8}}{\sqrt{6-1}}, 3.2 + 4.032\frac{\sqrt{4.8}}{\sqrt{6-1}}\right) \approx (-0.75, 7.15)$ となるが，体重はマイナスの値をとることはないから，0kg ∼ 7.15kg が 99 %の信頼区間となる．

b): 中心極限定理 (やその一般化) を思い出してみると，正規分布を示すと考えてよい特性値はそれが一様な多数の要因の和として実現しているものである．一方，飼い猫の集団には，ペルシャ猫，シャム猫，和猫など，明確に区別できる部分集団があることから，一様性が阻害されている可能性がある．一方，例えば，犬の背丈での，チワワの背丈とドーバーマンの背丈のような部分集団ごとの明確な違いは猫の体重にはないように思えるので，犬の背丈に対しての正規分布の仮定よりは，猫の体重に対する正規分布の仮定の方により高い妥当性があるように思える．

ウラへ

学生番号:　　　　　　　　　　　　　　　　　名前:

IV. N 区の家猫の体重の分布が正規分布に従うと仮定し，その平均値が 3.8kg だったとするとサイズが 6 の標本の平均 $\bar{X}$ とその分散 S^2 に対し，$T = \dfrac{\bar{X} - 3.8}{\sqrt{\frac{S^2}{6-1}}}$ は t-分布 t_{6-1} に従うので，たとえば 95 %の棄却域は，教科書 p.174 の数表から，$(-\infty, -2.571] \cup [2.571, \infty)$ となる．T の式に標本の実測値を代入すると，$\frac{3.2-3.8}{\sqrt{\frac{4.8}{6-1}}} \approx -0.612$ となり，棄却域には含まれない．このことから，得られた標本の値からは「N 区の家猫の体重の平均は 3.8kg である」が正しくないとは結論できない．(”正しいと結論できる” と言っているわけではないことに注意!)

V. この製品の，サイズが n の標本の平均を $\bar{X}$ とすると，$Z = \dfrac{\bar{X} - \mu}{\sqrt{\frac{25}{n}}}$ は標準正規分布に従うので，95 %の信頼区間は，$P(|Z| < 1.96) \approx 0.95 \Leftrightarrow \mu - 1.96\frac{5}{\sqrt{n}} < X < \mu + 1.96\frac{5}{\sqrt{n}}$ となる．この区間の幅は $2 \times 1.96\frac{5}{\sqrt{n}}$ だから，$2 \times 1.96\frac{5}{\sqrt{n}} < 4 \Leftrightarrow 24.01 < n$ だから，データのサイズは 25 個以上でなければいけないことがわかる．

VI. この会場での試験参加者も全国平均の正解率を持っていたと仮定すると，会場での正解者の数を返す確率変数を X すると，X は二項分布 $B_N(160, 0.9)$ に従うから，$E(X) = 160 \times 0.9$, $S^2(X) = 160 \times 0.9 \times 0.1$ となる．したがって，$Z = \dfrac{X - 160 \times 0.9}{\sqrt{160 \times 0.9 \times 0.1}} = \dfrac{\frac{X}{n} - 0.9}{\sqrt{\frac{0.9 \times 0.1}{160}}}$ は近似的に標準正規分布に従う．Z の式の X に実測値 131 を代入すると，$\dfrac{131 - 160 \times 0.9}{\sqrt{160 \times 0.9 \times 0.1}} = -3.4258$ となり，この値は有意水準を 1 %としても，棄却域に入ることがわかる．このことから，このクラスの平均は全国平均より有意に低いと結論することができる．(クラスの平均が全国平均より低いことはいずれにしても事実であるが，このクラスの平均が，全国平均より (ばらつきの許容範囲を越えて) 本質的に低い (有意に低い) と言える，ということを主張している点がポイントである．)