

# 区間推定

樋口さぶろお

龍谷大学理工学部数理情報学科

使える統計! L13(2014-01-15 Wed)

今日の目標

- ① 標本から, 母平均値, 母比率を区間推定できる.



<http://hig3.net>

## L11-Q2

## Quiz 解答:正規分布の確率

標準正規分布表を見ると,  $Q(z) = 1/40 = 0.025$  となるのは,  $z = 1.96$  くらいであるとわかる ( $1 - 0.025 \times 2 = 0.95$  だから図からもわかる).

$\mu = 2000, \sigma = 200$  と, 標準得点  $z = \frac{x-\mu}{\sigma}$  の定義から,  $x = 2392$  円程度いれておけばよい.

## L12-S2

## Quiz 解答:推定

母平均値の推定値は, 標本平均値で与えられ,

$$\bar{x} = \frac{1}{5}[10 + 20 + 30 + 30 + 110] = 40(\text{分})$$

母分散の推定値は, 標本 (不偏) 分散で与えられ,

$$s^2 = \frac{1}{5-1}[(10-40)^2 + (20-40)^2 + (30-40)^2 + (30-40)^2 + (110-40)^2] = 1600(\text{分}^2)$$

母標準偏差の推定値は, 標本 (不偏) 分散の平方根で与えられ,

$$s = \sqrt{1600} = 40(\text{分})$$

## L12-S3

## Quiz 解答:推定

これはサイズ 10 の標本.

標本平均値は

$$\frac{1}{10}[0 + 0 + 0 + 0 + 0 + 0 + 10 + 10 + 30 + 100] = 15(\text{円})$$

. よって, 母平均値は 15 円と推定される.

標本 (不偏) 分散は

$$\frac{1}{10-1}[(0-15)^2 \times 6 + (10-15)^2 \times 2 + (30-15)^2 + (100-15)^2] = 930.6(\text{円}^2)$$

. よって, 母分散は 930.6 円と推定される.

母標準偏差は  $\sqrt{930.6} = 30.5$  円と推定される.

## 点推定と区間推定

### 点推定

真の母平均値はわからないが、標本平均値を使って、

母平均値は  $A$  円と推定される

それどのくらい正確なの？ 実は

### 区間推定

母平均値が、 $B$  円以上  $C$  円以下である '確率' は 0.95.

ここで '確率' というのは不誠実. '信頼区間' という概念.

母平均値の **信頼係数** 95% の **信頼区間** は  $B$  円以上  $C$  円以下

というのが正しい言葉遣い. 以下で意味と求め方.

## 標本平均値の分布

標本平均値は標本ごとに異なる…標本平均値自身が確率分布を持つ.

### 標本平均値の分布

- 標本平均値  $\bar{X}$  の平均値  $= \mu =$  母平均値.
- 標本平均値  $\bar{X}$  の分散  $= \frac{\sigma^2}{n} = \frac{\text{母分散}}{\text{標本の大きさ}}$

で, 中心極限定理より, 標本サイズ  $n \rightarrow +\infty$  で, 正規分布に従う.

なぜなら, 中心極限定理から,  $X_1 + \cdots + X_n$  は

$$\text{母平均値 } n\mu, \quad \text{母分散 } n\sigma^2$$

の正規分布に近づく.

よって標本平均値  $\bar{X} = \frac{1}{n}[X_1 + \cdots + X_n]$  は

$$\text{母平均値 } \frac{1}{n} \times n\mu = \mu, \quad \text{母分散 } \frac{1}{n^2} \times n\sigma^2 = \frac{1}{n}\sigma^2$$

の正規分布に近づく.

母平均値  $\mu$ , 母分散  $\sigma^2$  の母集団から, サイズ  $n$  の標本を何回も取り出して, 毎回, 標本平均値  $\bar{X}$  を計算すると, 確率 0.95 で,

$$\mu - 1.96 \times \sqrt{\frac{\sigma^2}{n}} < \bar{X} < \mu + 1.96 \times \sqrt{\frac{\sigma^2}{n}}$$

の範囲に収まる.

$\bar{X}, \mu$  を移項すると, 確率は 0.95 で

$$\bar{X} - 1.96 \times \sqrt{\frac{\sigma^2}{n}} < \mu < \bar{X} + 1.96 \times \sqrt{\frac{\sigma^2}{n}}.$$

## 区間推定

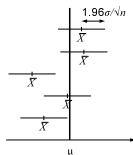
標本平均値  $\bar{X}$  として  $m$  が得られたとき、  
母平均値の、信頼係数 95% の信頼区間は

$$m - 1.96 \times \sqrt{\frac{\sigma^2}{n}} < \mu < m + 1.96 \times \sqrt{\frac{\sigma^2}{n}}$$

母平均値の、信頼係数 99% の信頼区間は

$$m - 2.58 \times \sqrt{\frac{\sigma^2}{n}} < \mu < m + 2.58 \times \sqrt{\frac{\sigma^2}{n}}$$

同じポリシーで何回も推定したときに、区間に含まれて‘当たる確率’



## 母分散なんてわかってないんですけど?

標本サイズ  $n$  が大きいとき,  $\sigma^2$  の代用として, 推定値である標本 (不偏) 分散  $s^2$  を使っちゃってもかまわない.

$n$  が小さいとき 正規分布でなく  $t$  分布というものを使って考えるのが本当は正しいけど, この授業ではやらない.



## クラスの身長之母平均値を区間推定してみよう!



or

<http://hig3.net> → (最初 or 左上) 生活の中の統計技術データ収集

## L13-Q1

## Quiz(区間推定)

ある賞金が確率分布に従うスピードくじを 10 回ひいたところ、賞金は、0 円, 0 円, 0 円, 0 円, 0 円, 0 円, 10 円, 10 円, 30 円, 100 円だった。賞金の母平均値を、信頼係数 95% の信頼区間で区間推定しよう。

## L13-Q2

## Quiz(区間推定)

瀬田学舎の龍大生の通学時間の分布を知るために、無作為に5人を選んで質問したところ次のようだった。通学時間の母平均値を、信頼係数99%の信頼区間で区間推定しよう。

10分, 20分, 30分, 30分, 110分。

## L13-Q3

## Quiz(区間推定)

某ハンバーガーチェーンのポテト M の重さを, ショップ A,B,C で調べたところ次のようになった (単位は g)

A 135,136,133,136

B 136,138,132,138

C 135,136,133,136,135,136,133,136

それぞれのショップのポテト M の重さの母平均値を, 信頼係数 95% で区間推定しよう.

## 母比率の区間推定

- クラスの中で、血液型 A 型の人々の比率は？  $n$  人に質問しただけで推定したい。
- 候補者 A の得票率は何%？  $n$  人に質問しただけで推定したい。
- 工場から出荷する製品のうち、何% が不良品？  $n$  個だけ抜き出して調査したい。
- このコインの表が出る確率は？  $n$  回投げるだけで推定したい。

母比率 = (1 人取り出したときに A 型である) 確率  $p$ 。

### 比率の点推定

$n$  人中  $k$  人が A 型だったとき

$$\frac{k}{n}$$

が母比率  $p$  のよい推定値になっている。

## 比率の区間推定

標本の中で該当するものの個数  $k$  の分散  $= np(1 - p)$ .

$\frac{k}{n}$  の分散  $= \frac{1}{n^2} np(1 - p) = \frac{1}{n} p(1 - p)$ .

$p$  を推定値  $\frac{k}{n}$  で代用する.

### 母比率の区間推定

母比率の信頼係数 95% の信頼区間は

$$\frac{k}{n} - 1.96 \times \sqrt{\frac{1}{n} \frac{k}{n} (1 - \frac{k}{n})} < p < \frac{k}{n} + 1.96 \times \sqrt{\frac{1}{n} \frac{k}{n} (1 - \frac{k}{n})}$$

母比率の信頼係数 99% の信頼区間は

$$\frac{k}{n} - 2.58 \times \sqrt{\frac{1}{n} \frac{k}{n} (1 - \frac{k}{n})} < p < \frac{k}{n} + 2.58 \times \sqrt{\frac{1}{n} \frac{k}{n} (1 - \frac{k}{n})}$$

## L13-Q4

## Quiz(比率の区間推定)

クラスから学生のサイズ( )の標本をとったところ、( )人中( )人が血液型 A であると答え、残りが A 型でないと答えた。

- ① クラスの、血液型 A である学生の母比率を (点) 推定しよう
- ② クラスの、血液型 A である学生の母比率の、信頼係数 95% の信頼区間を求めよう。
- ③ クラスの、血液型 A である学生の母比率の、信頼係数 99% の信頼区間を求めよう。



or

<http://hig3.net> → (最初 or 左上) 生活の中の統計技術データ収集



## L13-Q5

## Quiz(区間推定の性質)

標本からの母平均値の区間推定について、正しいのはどれ?

- ① 母分散が大きいほど、信頼区間は大きくなる
- ② 標本サイズが大きいほど、信頼区間は大きくなる
- ③ 母平均値が大きいほど、信頼区間は小さくなる
- ④ 信頼係数が大きいほど、信頼区間は小さくなる

## L13-Q6

## Quiz(母比率の区間推定の性質)

母平均値の区間推定で、標本サイズを2倍にしたとき、信頼区間の長さはいくらになるか？

- ① 4倍になる
- ② 2倍になる
- ③ 1.4倍になる
- ④ 変わらない
- ⑤ 0.7倍になる
- ⑥ 1/2倍になる
- ⑦ 1/4倍になる

## L13-Q7

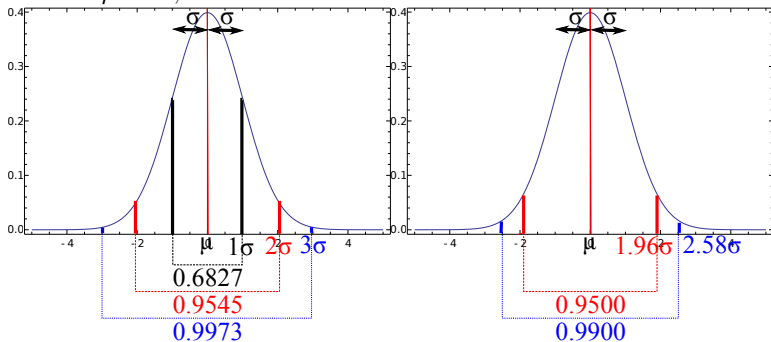
## Quiz(母比率の区間推定)

選挙で出口調査をしたところ、50人中35人がA候補に投票したと答えた。母比率、すなわち有権者全体でのA候補の得票率を考える。

- ① A候補の得票率を、(点)推定しよう
- ② A候補の得票率を、信頼係数95%で区間推定しよう。
- ③ A候補の得票率を、信頼係数99%で区間推定しよう。

## 標準正規分布 (ガウス分布) のグラフに関係した面積をおぼえよう!

いまは  $\mu = 0, \sigma = 1$  と思ってね.

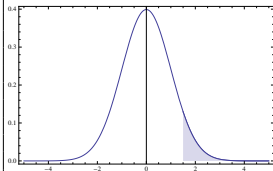


次のページの表さえあれば, 導ける.

標準正規確率表 (上側確率  $Q(z)$ )

$z \geq z_0$  となる確率  $= Q(z_0) = \frac{1}{2}\text{erfc}(z_0/\sqrt{2})$ . よく統計の教科書の付録に表が載ってる.

$z_0$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010



## 連絡

- きょうは紙1枚+アンケート2枚提出+携帯データ送信
- これまでの紙の配布資料や未返却の Quiz は, 1号館5階1-503向かいの掲示板前に置いています. 配布資料のファイルは <http://hig3.net> からダウンロードできます.
- 台風の分の補講やります! eラーニングシステム <https://moodle.media.ryukoku.ac.jp> 上で. 2014-01-16 木以降にメールで開始をお知らせします. 受講締切は 2014-01-28 火 23:55. ここで学習した分も出題計画に含まれます.

# ファイナルトライアル計画 I

日時 2014-01-29 Wed 水 3 (60 分)

場所 大注意. 教室変更あり. 8-102

形式 ペーパーテスト. 計算問題が多いけど選択肢問題群もあるかも. 関数電卓持込可. 正規分布の確率の表と, グラフ-面積の図は問題とともに印刷して配ります.

参照 A4 × 1 枚両面持込可.

- 専用用紙を使ってください. コピー可.
- 縮小コピー, 貼り付けは不可.
- 回収します.

持込 ファイナルトライアルでは携帯電話 (のアプリの電卓) は使えません.

配点 100 点 50 ピーナッツ

公欠 通常の追試規定に従います.

# ファイナルトライアル出題計画

2014-01-15 水 の授業中に最終的に下のように確定しました。

- (離散的) 確率分布が与えられたとき, 母平均値, 母分散, 母標準偏差をもとめる (L08)
- (連続的) 確率分布 (のグラフ) が与えられたとき, 確率, 母平均値, 母標準偏差を求める (L09)
- 標準とは限らない正規分布について, どの範囲ならこんな確率, こんな確率ならこの範囲, を求める (L11)
- 標本が与えられたとき, 標本平均値, 標本分散を計算することにより, 母平均値, 母分散, 母標準偏差, 母比率を点推定する (L12)
- 標本が与えられたとき, 母集団の母平均値, 母比率を区間推定する (L13)
- 母集団, 標本, 標本抽出, 推定などの意味 (L12) たぶん選択肢. 定義を暗記しておけばいいというものでなく, 心から納得しておいてね.
- (プチテスト範囲の再出題) データが与えられたときに, 母平均値, 母分散, 母標準偏差を求めよう (L02,L03). 標本ナントカ, との違いにご注意.
- ピアソンの  $\chi^2$ , クラメールの連関係数 (補講=L14 の内容) プチテストでやった相関係数とは違います.