計測工学　第 5 回

前回：
- 測定における平均
- 標本分布
- 母平均の区間推定

今回：
- 母平均の区間推定（続き）
- 相関と回帰

5.1 母平均μ の推定（母分散σ^2が未知の場合）

●母分散σ^2が既知の場合との違い

① $\bar{x}$の標本分布の分散: 母分散を利用できない.

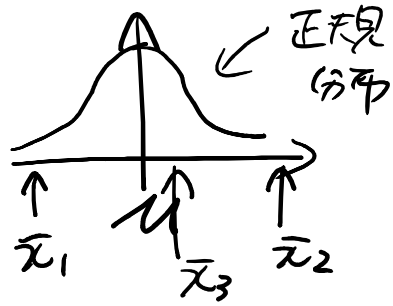
② $\bar{x}$の標本分布の形状: 正規分布ではない.

① →母分散σ^2の代わりに「不偏分散 u^2 (アルファベットの「ユー」)」を用いる

注:「不偏」とは？・・・ 標本により母集団の推定を行う際に、標本分布に偏りが生じないこと.

・平均の場合

標本平均$\bar{x}$は
母平均μの「不偏推定値」である.



・分散の場合　　　　標本分散$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$:

偏りを示す.
(常に母分散σ^2より
小さくなる)

↓

不偏分散$u^2 = \frac{n}{n-1}s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$:

$\frac{1}{n}$ではなく
$\frac{1}{n-1}$を取る.

母分散σ^2の
不偏推定値とみなせる.

・標準偏差の場合　　不偏分散からの標準偏差$u = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$:

↓

厳密には母標準偏差σの不偏推定値
ではない → nが十分大きい(n>30)なら
実用上問題ない.

② →$\bar{x}$の標本分布は「t分布　　」を示す

● $t=\dfrac{\bar{x}-\mu}{u/\sqrt{n}}$ の標準化で規定される分布.
(σではなくuを用いる)　またnの値によって分布形状が変化する.
(n=∞で正規分布)

表 5.1　**t** 分布表

自由度 ($N-1$)	標本数 (N)	$t_{N-1}(z\,\%)$, $z(\%)$：信頼度 68.3%	90%	95%	99%
		P(有意水準), 両側検定 31.7%	10%	5%	1%
1	2	1.837	6.314	12.706	63.657
2	3	1.321	2.920	4.303	9.925
3	4	1.197	2.353	3.182	5.841
4	5	1.141	2.132	2.776	4.604
5	6	1.110	2.015	2.571	4.032
6	7	1.090	1.943	2.447	3.707
7	8	1.077	1.895	2.365	3.500
8	9	1.066	1.860	2.306	3.355
9	10	1.059	1.833	2.262	3.250
10	11	1.052	1.812	2.228	3.169
15	16	1.034	1.753	2.131	2.947
20	21	1.026	1.725	2.086	2.845
25	26	1.020	1.708	2.060	2.787
30	31	1.017	1.697	2.042	2.750
40	41	1.013	1.684	2.021	2.704
60	61	1.008	1.671	2.000	2.660
120	121	1.004	1.658	1.980	2.617
∞	∞	1.000	1.645	1.960	2.576

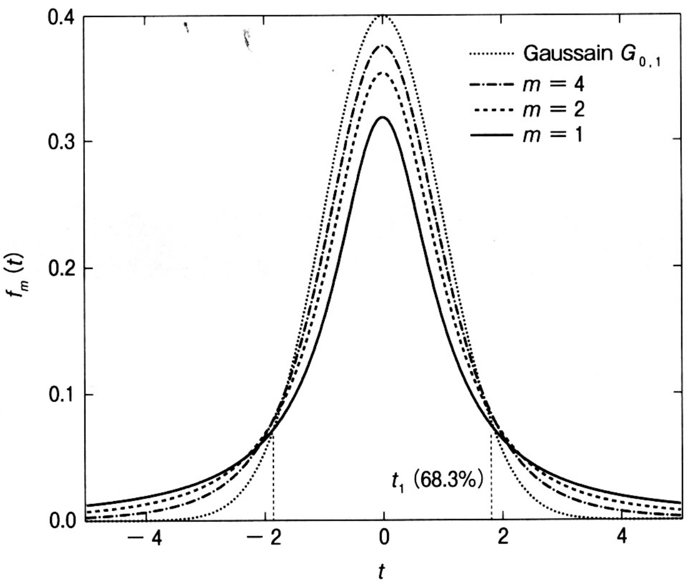


図 5.1　**t** 分布

●母平均 μ の区間推定結果を $\bar{x}-t_{n-1}(Z\%)\dfrac{u}{\sqrt{n}} \le \mu \le \bar{x}+t_{n-1}(Z\%)\dfrac{u}{\sqrt{n}}$ **と表す.**

→ t分布表から n および Z% に基づき
「$t_{n-1}(Z\%)$」の値を読み取る.

・例題：正規分布する母集団（母平均：μ，母分散：未知）から以下の標本を取出したとき，母平均 μ を信頼係数 68.3%で区間推定せよ.　→測定結果　80.0, 81.1, 80.5

$$\bar{x} = 80.53\cdots \fallingdotseq 80.5$$

$$u^2 = \frac{1}{2}\{(80.0-80.5)^2 + (81.1-80.5)^2 + (80.5-80.5)^2\}$$

$$= 0.305$$

t分布表より $t_{n-1}(2\%) = 1.321 \rightarrow t_{n-1}(2\%)\dfrac{u}{\sqrt{n}}$

$$= t_{n-1}(2\%)\sqrt{\frac{u^2}{n}}$$

$$= 0.427\cdots$$

$$= 0.4$$

$$80.1 \le \mu \le 80.9$$

5.2 母平均μ の区間推定（まとめ）

●母平均μ の推定（正規母集団$N(\mu, \sigma^2)$の場合）

σ^2 既知

$\bar{x}$の標本分布は標準化正規分布$N(0,1)$に従う.

$$z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$$

母平均μ の区間推定

$$\bar{x} - z\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + z\frac{\sigma}{\sqrt{n}}$$

正規分布表から算出可能,

信頼度 68.3% → $z = 1.000$

95% → $z = 1.960$

σ^2 未知

$\bar{x}$の標本分布は自由度$n-1$のt分布に従う.

$$t = \frac{\bar{x}-\mu}{u/\sqrt{n}}$$

u^2：不偏分散

u：不偏分散からの標準偏差

母平均μ の区間推定

$$\bar{x} - t_{n-1}(2\%)\frac{u}{\sqrt{n}} \le \mu \le \bar{x} + t_{n-1}(2\%)\frac{u}{\sqrt{n}}$$

信頼度2%, 標本数nより t分布表から読み取る.

5.3　相関と回帰

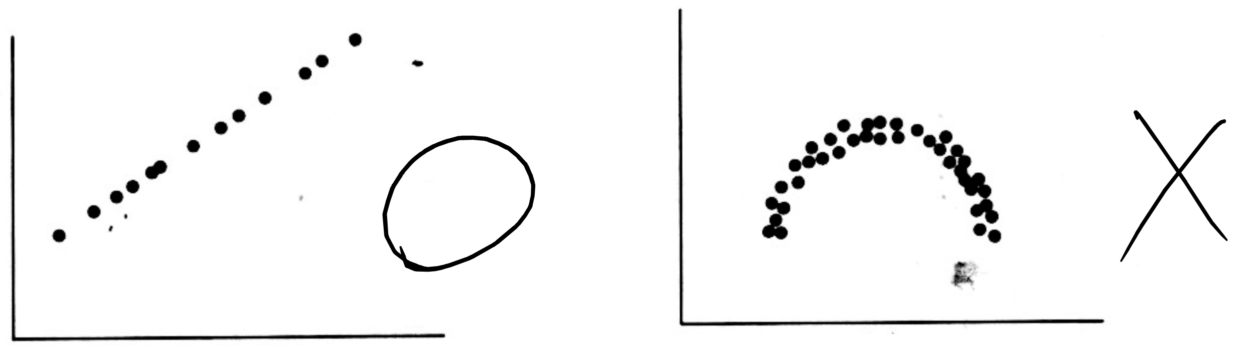
●相関関係： 一方が増加すると、他方も増加または減少する一組の変数の関係．

・関心のある 1 組の変数の間に，何か関連があるかどうかを調べたい．

例：・喫煙と心臓病の患者数
・数学と物理の成績　etc...

線形関係を指すのが一般的

・問い：次の 2 つのグラフで示されるような関係は相関関係があると見なせるのか？



・相関係数：r，相関関係の程度を定量的に示す指標．

・一般的には変数間の線形関係を示す．
(非線形の相関係数もある)

・因果関係を示すものではない．

●回帰分析： 変数間の関係を回帰式として定式化し，それを用いて変数の値を予測すること．

回帰式を求めるための一般的手法：
最小二乗法

5.4 第 5 回講義に関する意見･感想･質問のまとめ

●意見・感想

・小テストで式の形を忘れてしまった，小テストでヘマをしないように頑張る，小テストがわからなかった，n を 80 人にしてしまった，小テストを通して有効数字や分散について理解が深まった，小テストでデータ数の取り方を間違えた，小テストはできた感じがした，分散と標準偏差を間違えた：15←今回の小テストは，平均 7.1 点，満点 29 名でした．引き続き良い結果だったと思います．

・σ2 未知の場合の区間推定を理解できた，前回との結びつきがあって理解しやすかった，t 分布について理解できた，t 分布の区間推定について計算が難しかったが理解できた，まとめがあったのでわかりやすかった，相関関係のみなし方について理解できた，パターン別に説明されてわかりやすかった：13

・復習を頑張る，復習がコツコツできている，相関関係や回帰式を復習しておく，区間推定の式や用語を間違わないようにしっかり理解する，用語の意味を理解しながら復習する，それぞれの分散の違いを復習する実際に手を動かして計算して復習する，中間テストも近づいているので復習をしっかりする：11

・電卓を練習してきて満点を取る，電卓の機能を確認して計算できるようにする，手計算で分散を求めるとミスが起こるので電卓で行う，電卓で計算するとき母分散と標本分散を見間違えないようにする，自分の関数電卓は標本分散が表示されているので気をつける，応用数理では母分散と標本分散のどちらを使うかで迷ったことがある，今日の例題を電卓の統計モードでやろうと思ったがどうしても答えが合わなかった：8←

・改めて分布表をしっかり考えると問題が解きやすい，授業内容に問題なくついていけている，授業がスムーズに終わってよかった：3

・何気なく最小二乗法を使っていたが回帰分析という名前があることを初めて知った，相関と回帰については曖昧だったので改めて知ることができた：2

・授業早く終わって嬉しい，早く終わっても 4 限があるので同じだった：2←ギリギリまでしたくはないのですが，流石に今日は早すぎましたね･･･

・朝から気温が高くいっそう夏を感じた←明日は雨模様のようです．

・明日は大事なスウェーデン戦なので早く寝て明日に備える←8:00 からですよね，1 限がない人は精一杯応援してください！

・冷房で肌寒かったので上着を持ってくる←今日は 25℃設定でしたが，風の当たり方だけでもだいぶ違うと思うので席を移動するのも一つの手かもしれません．

・グラフを発明した人はすごい←今日，そんな話しましたっけ？

●質問

・例題の答えが $80.0 \leqq \mu \leqq 81.0$ になった（丸めの影響？）が正解になるか？←確かに丸め方次第で最小桁が 1 くらい変わりますので，その程度のずれは OK とします．

・計算途中で丸めは行わない方がいいか？←原則はしない方がいいですが，手計算ではかなり手間になりますよね．

・小テストで n が全員なのか 16 人なのか曖昧？←曖昧なのは理解できていないからです．「選んで」いるわけですから 16 人の方に決まっています．

・例題の記述で「･･･＝2.254＝2.3」と「･･･＝2.254≈2.3」はどちらも正解か？←この授業ではどちらでも可としますが，むしろ前者の方が適切です（有効数字として適切な桁に処理することは近似ではないため）．

・因果関係には相関係数のような指標はないのか？←現時点で私も詳しくは知りませんが，「因果関係が成立する」というための前提条件を満たすことの方が重要です．

・変数変換の式の形が「分母σ」と「分母σ/√n」の 2 つあって理解できていない←母集団の正規分布なのか，標本分布の正規分布なのかの違いです．

・相関関係と因果関係がどちらも認められる場合はあるのか？←基本的に因果関係が求められる時は相関関係も成り立っていると思います．

・母平均もわからない場合の区間推定はどうなるのか？←そもそも母平均がわからない前提なので推定しています．

・不偏分散はなぜ 1/(n-1)を取るのか？←むしろ応用数理 E で聞くべき質問です（私は統計の専門家ではないので）．一応ネットで調べたらきちんと数学的に説明している動画がありましたので，興味があったら見てみてください． https://www.youtube.com/watch?v=uzhopjGAiQg

5.5 第 4 回小テスト解答

Q.1　学生 80 名のクラスからランダムに選ばれた 16 人の平均身長は 157.7cm であった．クラス全員の身長の標準偏差が 4.6 であることが分かっている場合，クラス全員の平均身長を信頼係数 95%(Z=1.960)で区間推定せよ．

A.1　信頼区間の片側の幅：$Z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$ であり，ここで $n=16$，$Z=1.960$，$\sigma=4.6$

よって

$$Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}=1.960\cdot\frac{4.6}{\sqrt{16}}=2.254=2.3$$ （平均身長に合わせて小数第一位までとする）

$$\bar{x}-Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\le\mu\le\bar{x}+Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$157.7-2.3\le\mu\le157.7+2.3$$

従って　155.4≦μ≦160.0 cm