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：母分散が利用できない

②　$\bar{x}$の標本分布の形状：正規分布ではない

「ユー」

①　→母分散 σ^2 の代わりに「不偏分散 u^2」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、標本分布にかたよりがないこと。

・平均の場合

標本平均$\bar{x}$は、母平均μの「不偏推定値」とみなされる。

・分散の場合　　標本分散 $s^2 = \boxed{\dfrac{1}{n}}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$：$\bar{x}$からの差を取り、符号の影響を除去するために2乗して総和したものの平均

母分散に対して偏りを持つ
（常にσ^2より小さくなる）

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

母分散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 $u = \sqrt{\dfrac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}$：

→厳密には母標準偏差σの不偏推定値ではない（かたよりを持つ）→ 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分布: $t = \dfrac{\bar{x}-\mu}{u/\sqrt{n}}$ 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の値により図5.1のように分布形状が変化する.
(n=∞で正規分布と一致)

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	$\boxed{1.321}$	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

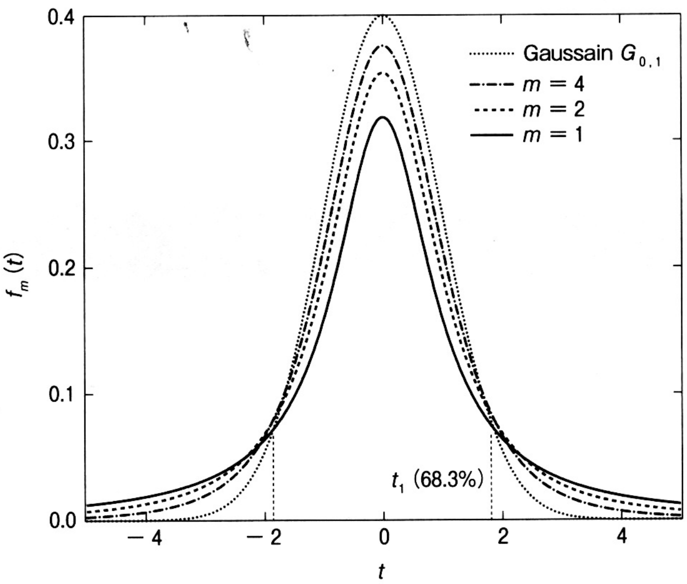


図 5.1　t分布

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

t分布表からnとZ%に基づき決定される.

・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

$\bar{x} = \frac{1}{n}\sum x_i = \frac{1}{3}(80.0+81.1+80.5) = 80.53\ldots = \underline{80.5}$

$$u^2 = \frac{1}{n-1}\Sigma(x_i - \bar{x})^2$$

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

$$= \frac{1}{2}\{(-0.5)^2 + 0.6^2\}$$

$$= 0.305$$

$t_{n-1}(2\%) = 1.321$, $t_{n-1}(2\%)\sqrt{\frac{u^2}{n}} = 0.421\cdots = 0.4$

$$\therefore 80.1 \leqq \mu \leqq 80.9$$

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$N(0, 1^2)$に従う

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

（nが大きければ、正規母集団でなくても成り立つ）

母平均μ の区間推定

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

正規分布表から求められる

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

$$u = \pm\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

母平均μ の区間推定

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

信頼度2%，標本数nに基づきt分布表から読み取る.

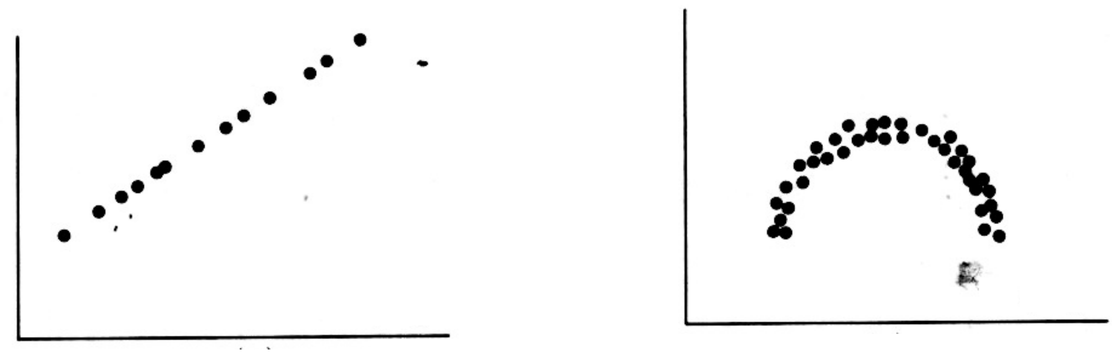
5.3　相関と回帰

●相関関係： 一方が増加すると、他方が増加もしくは減少する二つ（以上）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
- 数学の成績と物理の成績

・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：相関関係の程度を定量的に表わすもの．

・最も基本的な相関係数は、変数間の線形関係の程度を示す．

●回帰分析： 複数の変数間の関係を回帰式として記述し、それを用いて予測を行うこと．

最小二乗法を用いるのが一般的．

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

・応用数理 E でもここは苦手だったのでしっかり復習する，標本分布を使った区間推定を今までの区間推定と区別して覚える，前回と比べて難しくなってきたのでしっかり復習する，標本分散と不偏分散の違いを理解する，難しかったので復習する，記号（σと u，μ）を混同しないように確認する，授業に追いつききれなかったので復習する，前回の内容を十分に理解しておらず難しかったので復習する，似たような式が多くてこんがらがりそうなので復習する，不偏分散の出し方を身につけておく:11

・小テストの直前で誤りに気づいたが時間が足りなかった，またミスしてしまった，自信を持って答えられた，今までの小テストに自信がない，小テストで復習の成果が出た，小テストができた，小テストで記号が問題文になくて少し焦ったが問題なく解けた，小テストの出来が悪くなってきたので復習時間を延ばす:10←今回の小テストは平均 7.0 点，満点 21 名でした．前回よりは落ちましたが，以前として高いレベルの結果だと思います．

・母分散が既知と未知で整理して理解することができた，不偏分散を理解できた，t 分布に関して理解できた，母平均の区間推定を t 分布から示すことができた，非線形の相関関係について知った:9

・テストまで気づいたら 1 週間なので準備を進める，中間テストに向けて復習する，あっという間に中間試験 1 週間前になった:8←来週は月（通常授業）→水（補講）→木（中間試験）となります，タイトなスケジュールですが頑張ってください！

・例題の計算が複雑だったが電卓の機能で出せることを知ったので練習してみる，電卓を使いこなせるようになる，この機会に関数電卓の機能を確認する:4←電卓の機能を押さえておくのは大事かと思います．

・相関関係についての復習ができた，t 分布について復習できてよかった，新しいことはなかったが復習にはなった:3

・関数電卓で計算する度に Excel のすごさを思い知らされる，最近は Excel が自動で最小二乗法で計算してくれるため自分で求めることがなくなっていた:2←わかります，結局ちゃんとした計算をするには PC で Excel 使った方が楽ですよね．

・例題の不偏分散の計算ミスをなくしたい，例題の答えがあっていたのでよかった:2

・材料評価の 2 の舞になりたくないので気を引き締める←気持ちはわかります，頑張ってください！

・相関関係の右の図が相関なしにあのグラフの形状になることが考えにくい←仰る通りで，ただ一般的な線形関係のみを対象とした相関係数では検出できない，ということです．

・物理というより数学をやっている気がして楽しい←数学が好きなんですね，私からしたらマニアック！

・今バイトしている学習塾で教えている高校生が確率統計の勉強をしているので教えられるように自分も習得する←人に教えることで自分の理解も深まりますよね．

●質問

・最小二乗法の計算は実験でもよく使われるが求め方を完璧に覚えたほうがいいか？←次回やりますが，あれは覚えるものじゃないです．この科目の例題や試験においてもあれを暗記して計算させるような問題は出しません．

・相関関係の右の図に関して非線形の相関係数の式を使えば相関関係があると言える可能性があるのか？←そういうことです．

・相関係数と決定係数に違いはあるのか？←決定係数についてあまりよく理解していなかったので調べました．回帰式を求める際に表示される"R^2"のことなんですね．さらに，「最小二乗法による直線フィッティング

の場合，相関係数の二乗と決定係数は一致する。」という記述も見つけましたので，関連はある，と言えると思います．

・信頼区間と Z の関わりがいまいちわからない←授業でも説明しましたが，Z は信頼区間に含まれる確率を示す値です．

・標準偏差の有効数字がややこしい，区間推定でも有効数字の取り扱い方が大事なのでは？←例えば 2-3 ページの例題についても，仮に不偏分散を厳密に求めても結局信頼区間の幅を求める計算で大幅な丸め誤差は発生してしまう訳です．そういうことから考えると，あまり厳密な取り扱いをするのは意味がないように思います．

・以前不偏分散を用いるときには母集団の標本数が 10以上でないといけないと習ったが今回は 30 以上だった，この定義は明確にされているのだろうか？←

・不偏分散はなぜ n ではなく n-1 で割るのか？←（上の質問も含めて）うーん，この辺になると統計の専門的な知識が必要ですので，自身で調べるか授業で習った先生に教えてもらってください．

5.5 第 4 回小テスト解答

Q.1　学生 57 名のクラスの定期試験結果を，ランダムに抽出した 25 名のデータから分析したところ，標本平均は 76.2 点であった．クラス全員の成績における標準偏差が 5.9 点であることが分かっている場合，クラス全員の平均点を信頼係数 68.3%(Z=1.000)で区間推定せよ．

A.1　信頼区間の片側の幅：$Z\dfrac{\sigma}{\sqrt{n}}$ であり，ここで $n=25$， $\sigma=5.9$，また信頼係数 68.3%であるから

$Z=1.000$，よって

$$Z\frac{\sigma}{\sqrt{n}}=1.000\cdot\frac{5.9}{\sqrt{25}}=1.18=1.2$$ （標本平均に合わせて小数第一位までとする）

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

$$76.2-1.2 \le \mu \le 76.2+1.2$$

従って　75.0≦μ≦77.4 点