統計学
(2005 年度前期　講義ノート)

平成 17 年 4 月 30 日 (土) 版

教科書『基本統計学 (第２版)』(豊田・大谷・小川・長谷川・谷﨑著，東洋経済新報社，2002 年)

谷﨑　久志
神戸大学・経済学部

- この講義ノートは，
  http://ht.econ.kobe-u.ac.jp/~tanizaki/class
  からダウンロード可。
- この講義ノートの文中のページは教科書『基本統計学 (第 2 版)』のページに対応。

序説 (P.1)

1. 統計的記述：

   資料の収集と整理 (平均値・分散・メディアン等の計算) $\Longrightarrow$ 第 1, 2 章

2. 統計的推測：

   標本から母集団の特徴をつかむこと

   (a) 標本： データを標本と考える

   (b) 母集団： 標本を含む全体

   (c) 母集団の特徴： 母集団の特性を表すパラメータ (母数という)

   (d) パラメータ (母数)： 平均，分散

   $\Longrightarrow$ 母数 (パラメータ) の推定と仮説検定が主な内容

1　度数分布 (P.3)

1.1　変数 (P.3)

変数の種類 (P.3)

1. 連続型変数： ある区間内の任意の実数値をとりうる変数 (身長，体重，温度，$\cdots$)

2. 離散型変数： 不連続な値しかとらない変数 (サイコロの出た目，家族数，$\cdots$)
   ただし，離散型変数を連続型変数とみなす場合も多い (例： 金額は離散型変数，1997 年の GNP は 514343.1 $\times$ 10 億円で，1 円に対して，GNP の値はあまりにも大きい)

データの種類 (P.8)

1. 時系列データ： 時間に依存するデータ (P.5 の表 1.1, P.8 の表 1.4)

2. クロスセクション・データ (横断面データ)： 家計，企業等の一時点でのデータの系列 (P.8 の表 1.5，P.9 の表 1.6)

1.2　度数分布 (P.4)

表 1.2 (P.5) のデータ (20 個の物体の重さ):

4.3	5.2	7.2	6.4	3.5	5.6	6.7	6.1	4.1	6.8
5.0	5.6	3.8	4.6	5.8	5.1	6.2	5.3	7.4	5.9

このデータを整理する。
$\Longrightarrow$ 表 1.3 (P.7)

階級値	階級境界値	度数
3.45	2.95〜3.95	2
4.45	3.95〜4.95	3
5.45	4.95〜5.95	8
6.45	5.95〜6.95	5
7.45	6.95〜7.95	2
	合計	20

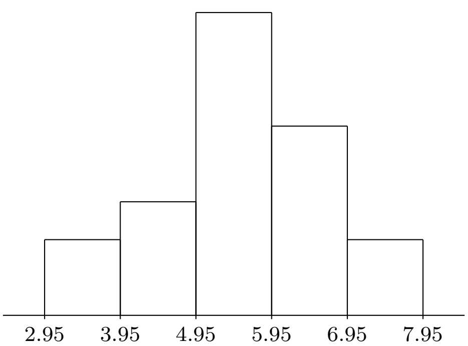
をもとにして,

表 1.3　20 個の物体の重さの度数分布表

階級値	階級境界値	度数	相対度数	累積度数	累積相対度数
3.45	2.95〜3.95	2	0.10	2	0.10
4.45	3.95〜4.95	3	0.15	5	0.25
5.45	4.95〜5.95	8	0.40	13	0.65
6.45	5.95〜6.95	5	0.25	18	0.90
7.45	6.95〜7.95	2	0.10	20	1.00
	合計	20	1.000		

を得る。小数第 2 位の 0.05 の単位で区間を分けている理由
$\longrightarrow$ 四捨五入の関係
小数第 1 位の 0.1 の単位で区間を分けた場合，境界値がどの階級に属するか区別できなくなる。(例えば，5.0 は 4.95 以上から 5.05 未満の間の数値)

図 1.1 20 個の物体の重さのグラフ



グラフの形

- 右の裾野が広い $\Longrightarrow$ 右に歪んでいる
- 左の裾野が広い $\Longrightarrow$ 左に歪んでいる

グラフの作り方

1. 階級境界値：階級の境界を定める値
2. 階級値：階級境界値の中点
3. 度数：ある階級に属するデータの数
4. 度数分布表：各階級とその度数を表に表したもの
5. ヒストグラム：度数分布をグラフに表す
6. 相対度数：各階級の度数をデータの総数で割ったもの，すなわち，各階級に属するデータの割合
7. 累積度数：ある階級以下の度数を合計したもの
8. 累積相対度数：ある階級以下の相対度数を合計したもの

2　代表値 (P.15)

度数分布表，ヒストグラム： 統計データを整理し，母集団に関する情報を得る一つの方法。
分布の状態を数値で表したい。
代表値： データを代表する値 $\Longrightarrow$ 平均値，分散，標準偏差，中央値 (メディアン)，最頻値 (モード)，$\cdots$

2.1　平均値 (P.15)

n 個のデータ： $x_1, x_2, \cdots, x_n$

算術平均 (P.15)：

$$\overline{x} = \frac{1}{n}(x_1 + x_2 + \cdots + x_n) = \frac{1}{n}\sum_{i=1}^{n} x_i$$

表 1.2 (P.5) のデータから

$$\overline{x} = \frac{1}{20}(4.3 + 5.2 + \cdots + 5.9) = 5.53$$

となる。

加重平均 (P.16)：

階級値	階級境界値 (以上)　(未満)	度数
m_1	$a_0 \sim a_1$	f_1
m_2	$a_1 \sim a_2$	f_2
$\vdots$	$\vdots$	$\vdots$
m_k	$a_{k-1} \sim a_k$	f_k
	合計	n

ただし，$m_1 = \dfrac{a_0 + a_1}{2}$, $m_2 = \dfrac{a_1 + a_2}{2}$, $\cdots$,
$m_k = \dfrac{a_{k-1} + a_k}{2}$ とする。

上のような度数分布表が利用可能なとき，

$$\overline{x} = \frac{1}{n}(f_1 m_1 + f_2 m_2 + \cdots + f_k m_k) = \frac{1}{n}\sum_{i=1}^{k} f_i m_i$$

として，平均値を計算することが出来る。$\Longrightarrow$ 加重平均 (各階級値を度数でウエイトづけして平均したもの)

$$\overline{x} = \sum_{i=1}^{k} \frac{f_i}{n} m_i$$

$\dfrac{f_i}{n}$ は相対度数である。
上の表のデータの平均を求めると，

$$\begin{aligned}\overline{x} &= \frac{1}{20}\Big(2 \times 3.45 + 3 \times 4.45 \\ &\quad +8 \times 5.45 + 5 \times 6.45 + 2 \times 7.45\Big) \\ &= 5.55\end{aligned}$$

階級の幅の選び方によって，多少，値は異なる。

2.2　分散，標準偏差 (P.18)

分散，標準偏差： データの散らばり具合を表す
分散，標準偏差が大きければ，データの存在する範囲が広い
標準偏差＝分散の平方根
分散 (s^2 で表す) の定義：

$$\begin{aligned}s^2 &= \frac{1}{n}\Big((x_1 - \overline{x})^2 + (x_2 - \overline{x})^2 + \cdots + (x_n - \overline{x})^2\Big) \\ &= \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2\end{aligned}$$

ただし，$\overline{x} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} x_i$ とする。

標準偏差： s
分散の実際の計算には，

$$s^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{x}^2$$

を用いる。
なぜなら，

$$\begin{aligned}s^2 &= \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 \\ &= \frac{1}{n}\sum_{i=1}^{n}(x_i^2 - 2\overline{x}x_i + \overline{x}^2) \\ &= \frac{1}{n}\Big(\sum_{i=1}^{n} x_i^2 - 2\overline{x}\sum_{i=1}^{n} x_i + \sum_{i=1}^{n}\overline{x}^2\Big) \\ &= \frac{1}{n}\Big(\sum_{i=1}^{n} x_i^2 - 2n\overline{x}^2 + n\overline{x}^2\Big) \\ &= \frac{1}{n}\Big(\sum_{i=1}^{n} x_i^2 - n\overline{x}^2\Big) \\ &= \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{x}^2\end{aligned}$$

となる。
表 1.2 (P.5) のデータの分散を求めると，

$$\begin{aligned}s^2 &= \frac{1}{20}\Big((4.3 - 5.53)^2 + (5.2 - 5.53)^2 + \cdots \\ &\quad +(5.9 - 5.53)^2\Big) \\ &= 1.1591\end{aligned}$$

または，

$$\begin{aligned}s^2 &= \frac{1}{20}(4.3^2 + 5.2^2 + \cdots + 5.9^2) - 5.53^2 \\ &= 1.1591\end{aligned}$$

$s = 1.0766$ ===＞ 標準偏差
表 2.1 (P.16) の度数分布表からの計算では，

$$s^2 = \frac{1}{n}\sum_{i=1}^{k} f_i (m_i - \overline{x})^2$$

となる。ただし，$\overline{x} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{k} f_i m_i$ とする。
実際の計算には，

$$s^2 = \frac{1}{n}\sum_{i=1}^{k} f_i m_i^2 - \overline{x}^2$$

を使う。
なぜなら,

$$
\begin{aligned}
s^2 &= \frac{1}{n}\sum_{i=1}^{k} f_i(m_i-\overline{x})^2 \\
&= \frac{1}{n}\sum_{i=1}^{k} f_i(m_i^2-2\overline{x}m_i+\overline{x}^2) \\
&= \frac{1}{n}\Big(\sum_{i=1}^{k} f_i m_i^2 - 2\overline{x}\sum_{i=1}^{k} f_i m_i + \overline{x}^2\sum_{i=1}^{k} f_i\Big) \\
&= \frac{1}{n}\Big(\sum_{i=1}^{k} f_i m_i^2 - 2n\overline{x}^2 + n\overline{x}^2\Big) \\
&= \frac{1}{n}\Big(\sum_{i=1}^{k} f_i m_i^2 - n\overline{x}^2\Big) \\
&= \frac{1}{n}\sum_{i=1}^{k} f_i m_i^2 - \overline{x}^2
\end{aligned}
$$

となる。
上の表のデータの分散を求めると,

$$
\begin{aligned}
s^2 = \frac{1}{20}\Big(&2(3.45-5.55)^2+3(4.45-5.55)^2 \\
&+8(5.45-5.55)^2+5(6.45-5.55)^2 \\
&+2(7.45-5.55)^2\Big) \\
= 1.19&
\end{aligned}
$$

または,

$$
\begin{aligned}
s^2 = \frac{1}{20}(&2\times 3.45^2+3\times 4.45^2 \\
&+8\times 5.45^2+5\times 6.45^2+2\times 7.45^2)-5.55^2 \\
= 1.19&
\end{aligned}
$$

すなわち，$s=1.0909$,

2.3　メディアン，モード (P.17)

- 範囲： 最大値ー最小値
- 四分位点：
  25 %点 (第 1 四分位点)，50 %点 (第 2 四分位点)，75 %点 (第 3 四分位点) のこと
- 四分位範囲： 第 3 四分位点ー第 1 四分位点
- メディアン（中央値）：
  大きい順に並べて，真ん中の値 (第 2 四分位点)
- モード（最頻値）：
  最も多い度数の階級値 (表 1.3 のデータでは 5.45，階級の幅によって変わる)

3　正規分布と正規分布表 (P.65)

確率変数

- 離散型確率変数 $\Longrightarrow$ 2 項分布，$\cdots$
- 連続型確率変数 $\Longrightarrow$ 正規分布，カイ 2 乗 (χ^2) 分布，t 分布，$\cdots$

3.1　正規分布の特性 (P.65)

正規分布の確率密度関数 $f(x)$

$$
f(x)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)
$$

ただし，$\exp(x)=e^x$ とする。$\pi=3.141592$ (円周率)，$e=2.718282$ (自然対数の底) に注意。

$$
\mathrm{E}(X)=\mu,\quad \mathrm{V}(X)=\sigma^2
$$

$\Longrightarrow$ 平均 μ，分散 σ^2 の正規分布
$\Longrightarrow N(\mu,\sigma^2)$
確率変数 X が正規分布 $N(\mu,\sigma^2)$ に従う
$\Longrightarrow X\sim N(\mu,\sigma^2)$
正規分布の確率密度関数 $\Longrightarrow$ 図 5.1
性質：

1. $x=\mu$ に関して左右対称
2. 正規分布の平均，メディアン (中央値)，モード (最頻値) はすべて等しく μ
3. 下側の面積の合計は 1 $\Longrightarrow$ 連続型確率密度関数の性質

$$
\int_{-\infty}^{\infty} f(x)\mathrm{d}x=1
$$

$N(0,1)\Longrightarrow$ 標準正規分布
重要：
$X\sim N(\mu,\sigma^2)$ とする。このとき，基準化 (標準化) すると，$Z=\dfrac{X-\mu}{\sigma}\sim N(0,1)$ となる。(P.52 の定理 4.4 を参考に)
重要：

$X_1, X_2, \cdots, X_n$ は互いに独立で，X_i は平均 μ，分散 σ^2 の正規分布に従うとする。
このとき，$\overline{X} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i \sim N(\mu, \dfrac{\sigma^2}{n})$ となる。(P.59，定理 4.9 を参考に)
さらに，基準化 (標準化) すると，$\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ となる。

3.2 正規分布表の使い方 (P.67)

分布関数 $F(x)$

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$$

$f(t)$ が正規分布の確率密度関数のとき，積分の計算は手計算は不可能
$\Longrightarrow$ 正規分布表 (P.68, P.245) の利用
正規分布表 (P.68, P.245) $\Longrightarrow$ 標準正規分布 $N(0,1)$ の確率を求める
$Z \sim N(0,1)$ について，

$$P(Z > 1.96) = ?$$

正規分布表では，標準正規分布 $N(0,1)$ の上側確率が計算されている。
上側確率： Z がある値 z より大きくなる確率 $P(Z > z)$
$P(Z > z) = \alpha$ となるとき，z のことを 100α % 点という。
$P(|Z| > z) = \alpha$ を両側確率と呼ぶ。
$P(|Z| > z) = \alpha$ となるとき，z のことを $100\alpha/2$ % 点という。
$P(Z > 1.96) = 0.0250$

例題 5.1 (P.68)： $P(Z \ge 1.64) = P(Z > 1.64) = 0.0505$

例題 5.2 (P.69)： $P(Z < 1.96) = 1 - P(Z \ge 1.96) = 1 - 0.0250 = 0.9750$

例題 5.3 (P.69)： $P(Z < -1.96) = P(Z > 1.96) = 0.0250$

例題 5.4 (P.69)： $P(-1.96 < Z < 1.64)$
$= 1 - P(Z > 1.64) - P(Z > 1.96)$
$= 1 - 0.0505 - 0.0250 = 0.9245$

例題 5.5 (P.70)： $P(0.25 < Z < 1.96)$
$= P(Z > 0.25) - P(Z > 1.96)$
$= 0.4013 - 0.0250 = 0.3763$

例題 5.6 (P.71)： $X \sim N(5, 2^2)$ のとき，$P(6 < X < 8) = ?$

解答： $Z = \dfrac{X-\mu}{\sigma} \sim N(0,1)$ を利用する。
$Z = \dfrac{X-5}{2} \sim N(0,1)$ なので，
$P(6 < X < 8)$
$= P(\dfrac{6-5}{2} < \dfrac{X-5}{2} < \dfrac{8-5}{2})$
$= P(0.5 < Z < 1.5)$
$= P(Z > 0.5) - P(Z > 1.5)$
$= 0.3085 - 0.0668 = 0.2417$

例題 5.7 (P.71)： ある会社の従業員の通勤時間は平均 60 分，標準偏差 15 分の正規分布にしたがっている。この会社の 2.5 % の従業員が通勤時間の長さに不満を持っている。不満を持っている従業員の通勤時間は何分以上か？

解答： 従業員の通勤時間を X とする。
$X \sim N(60, 15^2)$
$Z = \dfrac{X-60}{15}$ とすると，
$Z \sim N(0,1)$
$P(Z > z) = 0.0250$ を満たす z は 1.96 なので，
$P(Z > 1.96) = 0.0250$
$\Longrightarrow P(\dfrac{X-60}{15} > 1.96) = 0.0250$
$\Longrightarrow P(X > 89.4) = 0.0250$
したがって，89.4 分以上の通勤時間の従業員が不満を持っていることになる。

問題 5.1 (P.72)： $Z \sim N(0,1)$

1. $P(Z \ge 1.57) = 0.0582$
2. $P(Z < 1.34)$

   $= 1 - P(Z > 1.34)$

   $= 1 - 0.0901 = 0.9099$

3. $P(-0.37 < Z \leq 1.6)$

$= 1 - P(Z > 0.37) - P(Z > 1.6)$

$= 1 - 0.3557 - 0.0548 = 0.5895$

4. $P(0.55 < Z < 1.67)$

$= P(Z > 0.55) - P(Z > 1.67)$

$= 0.2912 - 0.0475 = 0.2437$

5. $P(-2.08 < Z < -0.21)$

$= P(0.21 < Z < 2.08)$

$= P(Z > 0.21) - P(Z > 2.08)$

$= 0.4168 - 0.0188 = 0.3980$

問題 5.2 (P.72)： $X \sim N(2, 9)$,
i.e., $X \sim N(2, 3^2)$,
i.e., $Z = \dfrac{X-2}{3} \sim N(0,1)$

1. $P(X \geq 5.6)$

$= P(\dfrac{X-2}{3} \geq \dfrac{5.6-2}{3})$

$= P(Z < 1.2)$

2. $P(X < 10)$

$= P(\dfrac{X-2}{3} < \dfrac{10-2}{3})$

$= P(Z < 2.67)$

3. $P(1 < X \leq 4.7)$

$= P(\dfrac{1-2}{3} < \dfrac{X-2}{3} \leq \dfrac{4.7-2}{3})$

$= P(-0.33 < Z < 0.9)$

4. $P(3.2 < X < 7.7)$

$= P(\dfrac{3.2-2}{3} < \dfrac{X-2}{3} < \dfrac{7.7-2}{3})$

$= P(0.4 < Z < 1.9)$

5. $P(-1.3 < X < 1.19)$

$= P(\dfrac{-1.3-2}{3} < \dfrac{X-2}{3} < \dfrac{1.19-2}{3})$

$= P(-1.1 < Z < -0.27)$

表 2: t 分布表 $t(m)$： P.247

$$\alpha = P(T > t_\alpha) = \int_{t_\alpha}^{\infty} \frac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{m}{2})} \frac{1}{\sqrt{m\pi}} \frac{1}{(1+\frac{x^2}{m})^{\frac{m+1}{2}}} \mathrm{d}x$$

α m	.10	.05	.025	.010	.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
31	1.309	1.696	2.040	2.453	2.744
32	1.309	1.694	2.037	2.449	2.738
33	1.308	1.692	2.035	2.445	2.733
34	1.307	1.691	2.032	2.441	2.728
35	1.306	1.690	2.030	2.438	2.724
36	1.306	1.688	2.028	2.434	2.719
37	1.305	1.687	2.026	2.431	2.715
38	1.304	1.686	2.024	2.429	2.712
39	1.304	1.685	2.023	2.426	2.708
40	1.303	1.684	2.021	2.423	2.704
41	1.303	1.683	2.020	2.421	2.701
42	1.302	1.682	2.018	2.418	2.698
43	1.302	1.681	2.017	2.416	2.695
44	1.301	1.680	2.015	2.414	2.692
45	1.301	1.679	2.014	2.412	2.690
46	1.300	1.679	2.013	2.410	2.687
47	1.300	1.678	2.012	2.408	2.685
48	1.299	1.677	2.011	2.407	2.682
49	1.299	1.677	2.010	2.405	2.680
50	1.299	1.676	2.009	2.403	2.678
60	1.296	1.671	2.000	2.390	2.660
80	1.292	1.664	1.990	2.374	2.639
120	1.289	1.658	1.980	2.358	2.617
240	1.285	1.651	1.970	2.342	2.596
∞	1.282	1.645	1.960	2.326	2.576

表 1: 正規分布表 $N(0,1)$： P.68, 245

$$\alpha = \mathrm{Prob}(Z > z_\alpha) = \int_{z_\alpha}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}x^2)\mathrm{d}x$$

z_α	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.4960	.4920	.4880	.4841	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4091	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.3946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2644	.2611	.2579	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2297	.2266	.2236	.2207	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1563	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1094	.1075	.1057	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0722	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0390	.0384	.0375	.0367
1.8	.0359	.0352	.0344	.0333	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0076	.0073	.0071	.0070	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0042	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002

α	.10	.05	.025	.010	.005
z_α	1.2816	1.6449	1.9600	2.3263	2.5758

3.3 t 分布

1. **定理 6.4 (P.86)：**

   $Z \sim N(0,1)$,

   $U \sim \chi^2(k)$,

   Z と U が独立

   このとき，

   $$T = \frac{Z}{\sqrt{U/k}} \sim t(k)$$

   自由度 k の t 分布： $t(k)$

t 分布 $\Longrightarrow$ 形状は自由度に依存する (図 6.3, P.87)
そのため，上側確率 0.10, 0.05, 0.025, 0.010, 0.005 の値のみが付表 3 になっている。
自由度もいくつか限定されている。
正規分布より裾野の広い分布 (図 6.3)
k が大きくなると，$t(k)$ は $N(0,1)$ に近づく。
$\Longrightarrow$ 付表 3 (P.247) の $m = \infty$ の数値を付表 1 の下の表 (P.245) と比較
例： $T \sim t(10)$ のとき，
$P(|T| > 3.169) = 0.01 \Longrightarrow$ 付表 3 (P.247)，図 6.4 (P.87)

すべての $i = 1, 2, \cdots, n$ について，$X_i \sim N(\mu, \sigma^2)$ で，
しかも，X_1, X_2, $\cdots$, X_n は互いに独立である
と仮定する。

1. **定理 6.5：標本平均の標本分布 (P.86)：**

   $$T = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$$

   自由度 $n-1$ の t 分布

   ただし，

   $$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

   とする。

   証明：

   すべての $i = 1, 2, \cdots, n$ について，$X_i \sim N(\mu, \sigma^2)$ で，

   しかも，X_1, X_2, $\cdots$, X_n は互いに独立である

   と仮定すると，

   $$\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$$

   となる。標準化によって，

   $$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

   を得る。

   一方，定理 6.3 (P.83) から，

   $$U = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

   となる。

   Z と U は独立となる。(証明略)

   したがって，

   $$\begin{aligned} T &= \frac{Z}{\sqrt{U/(n-1)}} \sim t(n-1) \\ &= \frac{\dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}}}{\sqrt{\dfrac{(n-1)S^2}{\sigma^2}\Big/ n-1}} \\ &= \frac{\dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}}}{\dfrac{S}{\sigma}} \\ &= \frac{\overline{X} - \mu}{S/\sqrt{n}} \end{aligned}$$

   なので，

   $$T = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$$

   を得る。

まとめ，

$$\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1) \qquad \Longrightarrow \text{(5.4) あたり (P.66)}$$

$$\frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t(n-1) \qquad \Longrightarrow \text{定理 6.5 (P.86)}$$