経済統計概論　5

標本分布と推定

1．ランダムサンプリング

- 無作為抽出法
  - 調査対象のすべてを調査せず，その一部を無作為に選び，それのみを調査する
  - 無作為に選ぶ場合，多くは乱数表を使用し，乱数表の示す番号に該当する対象を標本とする（単純無作為抽出）．
    - 学生生活実態調査の場合は，乱数表は抽出開始点を決めるためにのみ使用し，学籍番号順に今回の場合5人おき，または，3人おきに選んでもよい（系統抽出）．
    - この場合，学籍番号の振り方と，調査対象の調査属性が独立であると仮定している．

2

1．1　無作為抽出とは何か？（1）

- 無作為かどうかは，標本の選び方と，調査する属性が独立であるかどうかによって決まる．
  - 無作為ではない例
    - 週当たりの授業出席数という属性と，ある授業に出席しているということは独立ではない．
      - 標本に選ばれる確率は週当たりの授業出席数に比例する．
    - 通学時間という属性とある授業に出席しているということは独立ではない．
  - よって，授業でこのような質問を含むアンケートを採ることは厳密に言うと無作為抽出にはならない．
    ー＞セレクションバイアスがあるという．

3

1．1　無作為抽出とは何か？（2）

- 標本同士は独立でなければならない
  - 無作為ではない例
    - ある標本の値が特に高いので，値が低そうな標本を次は選んだ
    - 生活スタイルが似た人々のグループでまとめて標本を抽出
- 無作為抽出ではない場合を有意抽出と呼ぶ
- 学生の調査を行う場合
  - 学籍番号を無作為に抽出し，その学生を呼び出しアンケートに記入してもらう

4

1．2　有意抽出に意味があるか

- 前者のセレクションバイアスがある場合
  - ほとんど意味がない．
    - 補助調査を行いバイアス補正する
    - 重み付き分布などの理論モデルで補正
- 後者の独立性に問題がある
  - 調査結果は平均的にはバイアスがない
  - 誤差評価に問題がある

5

1．3　無作為抽出の確率モデル

- 各調査対象の確率モデル
  - 各調査対象の属性値の分布
    - （母）平均　μ，分散 σ^2　とする
    - この段階では調査対象同士が独立でなくてもよい
      - 集落，ゼミ，同一の授業に出席などの集団属性
- 各標本の確率モデル
  - 各標本の属性値の分布
    - 平均　μ，分散 σ^2
      - 平均，分散に関してバイアスが発生しないよう抽出
    - 各標本は独立（抽出によって独立化）

6

2．標本平均の分布
2．1　標本平均の期待値

- i番目の標本値を X_i とする．
- 標本平均　$\overline{X} = (X_1 + \cdots + X_n)/n$
- 標本平均の期待値（6．3（1）の法則を使う）

$$E[\overline{X}] = E\left[\frac{X_1 + \cdots + X_n}{n}\right] = \frac{1}{n}E[X_1 + \cdots + X_n]$$

$$= \frac{1}{n}\{E[X_1] + \cdots + E[X_n]\} = \frac{1}{n}\{\mu + \cdots + \mu\} = \mu$$

  – 標本平均の期待値は調査対象の分布の平均と一致

2．2　標本平均の（分布の）分散

- 標本平均の分散（6．3（2）の法則を使う）
  – Xiは各独立だから

$$V[\overline{X}] = V\left[\frac{X_1 + \cdots + X_n}{n}\right] = \frac{1}{n^2}V[X_1 + \cdots + X_n]$$

$$= \frac{1}{n^2}\{V[X_1] + \cdots + V[X_n]\} = \frac{1}{n^2}\{\sigma^2 + \cdots + \sigma^2\}$$

$$= \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

2．3　標本平均の分布の意味

- **標本平均の分布**
  – **期待値** μ　　（**調査対象の母平均**）
  – **分散** σ^2/n　（**調査対象の分散/標本数**）
- 標本平均で母平均を推定するのは平均的には正しい（バイアスがない，**不偏**である）
- 標本数が増えるとそれに反比例して標本平均のばらつき＝分散が小さくなる
  – 標本数が大きくなると標本平均は調査対象の平均にどんどん近づく（**一致性**がある）

標本平均の分布実験（1）

- 平均1，分散1の正規分布から標本をとる

	標本1	標本2	標本3	標本4	標本5	標本6	標本7	標本8	標本9	標本10
X1	0.448	0.904	1.209	1.659	-0.358	2.662	1.130	1.831	0.582	1.898
X2	1.445	1.896	0.072	1.440	0.903	-2.061	0.214	0.041	3.515	0.858
X3	-1.539	0.699	1.230	-0.477	2.603	-0.589	0.724	1.195	1.578	0.842
X4	1.594	1.853	1.214	2.657	0.383	1.003	-1.003	3.021	0.265	1.575
X5	-0.667	1.890	0.495	-0.790	1.902	1.105	-1.172	1.525	1.298	1.653
X6	0.685	-0.118	0.895	1.792	-0.362	0.117	1.391	0.739	1.047	1.265
X7	0.269	-0.440	-0.009	2.698	0.629	1.653	0.645	0.159	0.554	1.459
X8	0.384	1.216	0.457	1.337	-0.217	-1.329	0.653	1.023	3.062	0.108
X9	2.503	0.120	0.589	2.047	0.448	1.782	1.075	0.316	1.462	1.097
X10	0.564	0.009	0.648	0.383	2.577	2.640	1.538	1.073	1.527	1.854
標本平均	0.568	0.803	0.680	1.275	0.851	0.698	0.519	1.092	1.489	1.261

標本平均の標本平均	0.9237	標本平均の分散	0.1004

標本平均の分布実験（2）

- 標本数が増えると推定の正確さがます
  – 標本数１００の場合

X94	1.781	1.243	-0.295	1.719	0.561	-0.069	-0.029	1.049	-0.697	2.217
X95	1.999	0.257	3.495	1.756	0.871	-1.350	1.858	0.605	1.293	-0.882
X96	1.425	1.247	0.691	-0.822	1.356	0.832	2.463	2.325	0.779	0.377
X97	1.484	1.457	0.194	0.484	1.466	2.409	-0.033	2.658	1.599	0.823
X98	0.000	1.350	1.186	-0.105	1.085	1.701	2.774	2.594	0.829	0.167
X99	0.304	0.216	-0.824	0.180	1.273	0.170	2.433	-1.676	-0.352	0.301
X100	0.741	1.371	3.149	-1.660	2.477	0.975	0.864	1.816	2.318	-0.238
標本平均	1.084	0.999	1.032	1.052	1.032	0.967	1.174	1.021	1.001	0.738

標本平均の標本平均	1.010	標本平均の分散	0.0111

ひさびさの国試問題

- 3個のランダム標本 X_1, X_2, X_3 に基づき，母集団平均を推定したい．このとき，次の記述のうち，妥当なのはどれか．
  – 1．加重平均 $(3X_1 + 2X_2 + X_3)/6$ を推定量として用いる方が，単純平均 $(X_1 + X_2 + X_3)/3$ を推定量として用いるよりもよい．その理由は，この二つの推定量の分散は等しいが，単純平均は不偏推定量ではないからである．
  – 2．単純平均を推定量として用いる方が，加重平均を推定量として用いるよりもよい．その理由は，この二つの分散は等しいが，加重平均は不偏推定量ではないからである．

国試問題（続き）

– 3．加重平均を推定量として用いる方が，単純平均を推定量として用いるよりもよい．その理由は，加重平均も単純平均も共に不偏推定量であるが，加重平均の方が推定量の分散が小さいからである．
– 4．単純平均を推定量として用いる方が，加重平均を推定量として用いるよりもよい．その理由は，加重平均も単純平均も共に不偏推定量ではあるが，単純平均の方が推定量の分散が小さいからである．
– 5．加重平均を推定量として用いても，単純平均を推定量として用いても，結果に違いはない．その理由は，加重平均も単純平均も共に不偏推定量であり，推定量の分散も等しいからである．（H.14国家公務員試験1種経済職）

13

問題のポイント

- 加重平均 $(3X_1+2X_2+X_3)/6$ と単純平均 $(X_1+X_2+X_3)/3$ の比較
  - 期待値
    - 不偏性　期待値が調査対象の分布平均と一致するか
  - 分散
    - どちらの分散が小さいか？

14

期待値の比較

- 期待値
  - 単純平均の期待値は母平均と一致する
  - 加重平均の期待値

$$
\begin{aligned}
E[(3X_1+2X_2+X_3)/6] &= (1/6)\times E[3X_1+2X_2+X_3] \\
&= (1/6)\times\{E[3X_1]+E[2X_2]+E[X_3]\} \\
&= (1/6)\times\{3E[X_1]+2E[X_2]+E[X_3]\} \\
&= (1/6)\times(3\mu+2\mu+\mu)=\mu
\end{aligned}
$$

  - 加重平均も期待値は母平均と一致（不偏性）

15

分散の比較

- 分散
  - 単純平均の分散は $\sigma^2/n=\sigma^2/3$
  - 加重平均の分散

$$
\begin{aligned}
V[(3X_1+2X_2+X_3)/6] &= (1/36)\times V[3X_1+2X_2+X_3] \\
&= (1/36)\times\{V[3X_1]+V[2X_2]+V[X_3]\} \\
&= (1/36)\times\{9V[X_1]+4V[X_2]+V[X_3]\} \\
&= (1/36)\times(9\sigma^2+4\sigma^2+\sigma^2)=14\sigma^2/36=7\sigma^2/18
\end{aligned}
$$

  - 加重平均の分散は単純平均の分散以上になる
    - 加重平均の重みが同じなり単純平均と同じになる場合は等しい

16

2．4　標本数が十分大きいときの標本平均の分布（1）

- 中心極限定理
  - 独立な確率変数の和の分布を考える
  - $Y_1,Y_2,\cdots,Y_n$ を独立で期待値0，分散1の確率変数の列とする
  - $S_n=Y_1+Y_2+\cdots+Y_n$
  - $S_n/\sqrt{n}$ はnが大きくなるとその分布は正規分布に近づく
- これを利用して標本平均の分布を近似する

17

2．4　標本数が十分大きいときの標本平均の分布（2）

- Xiを変換して中心極限定理を適用する

$$
Y_i=\frac{X_i-\mu}{\sigma}
$$

$$
E[Y_i]=E\left[\frac{X_i-\mu}{\sigma}\right]=\frac{E[X_i]-\mu}{\sigma}=\frac{\mu-\mu}{\sigma}=0
$$

$$
V[Y_i]=\frac{1}{\sigma^2}V[X_i-\mu]=\frac{V[X]+(-1)^2V[\mu]}{\sigma^2}=\frac{\sigma^2+0}{\sigma^2}=1
$$

18

2．4　標本数が十分大きいときの標本平均の分布（3）

- Yiに対して中心極限定理を適用する

$$\frac{Y_1+\cdots+Y_n}{\sqrt{n}}=\frac{(X_1-\mu)/\sigma+\cdots+(X_1-\mu)/\sigma}{\sqrt{n}}$$

$$=\frac{X_1+\cdots+X_n-n\mu}{\sigma\sqrt{n}}=\frac{\sqrt{n}}{\sigma}\left(\frac{X_1+\cdots+X_n}{n}-\mu\right)$$

これがN（0，1）に近づく．

$$\frac{\sqrt{n}}{\sigma}\left(\frac{X_1+\cdots+X_n}{n}-\mu\right)\longrightarrow Z$$

19

2．4　標本数nが十分大きいときの標本平均の分布（4）

- 変換

$$\overline{X}=\frac{X_1+\cdots+X_n}{n}\longrightarrow\frac{\sigma}{\sqrt{n}}Z+\mu\sim N\left(\mu,\frac{\sigma^2}{n}\right)$$

- 標本平均の分布は近似的には平均μ，分散 σ^2/n の正規分布で近似できる．
- 上記の式の考え方
  - $\overline{X}-\mu$ の分布が $(\sigma/\sqrt{n})Z\sim N(0,\sigma^2/n)$
  - あるいは，$(\sqrt{n}/\sigma)(\overline{X}-\mu)$ の分布が $Z\sim N(0,1)$

20

3．母平均の推定（nが大きい場合）3．1　「古典」統計学における推定

- 「古典」統計学の考え
  - μは確率変数ではなく定数（$\overline{X}$ が解っていても $\overline{X}$ はあくまでも確率変数．原因μ，結果 $\overline{X}$）
  - $P(\overline{X}-c\le\mu\le\overline{X}+c)=0.95$ となるcを求める．
    - あたる確率が95%になる区間を求める．
    - $P(\mu-c\le\overline{X}\le\mu+c)=0.95$ となるcを求める．
- これと対立する考え方にベイズ統計学がある
  - 一旦 $\overline{X}$ が解ってしまえば，その情報から再出発する．従って，μを確率変数と考えることになる．

21

3．2　推定のための戦略

- 既知のもの
  - 標本値が解って後は，
    $\overline{X}$（標本平均），n（標本数），
    Zの分布（標準正規分布）
- 推測したいこと
  - 誤差を考えたμの推定値
- 分かっていないこと　σまたは σ^2

22

3．3　σ（σ^2）の点推定

- 分散を推定する
  - （標本）分散（全数調査）　$\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$
  - 標本分散（標本調査）　$\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$
- 今の場合nが大きいので，この二つの値はほとんど変わらない．
- $\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$ を採用しよう．

23

3．4　推定法（1）

- 結局σを $\hat{\sigma}$ で推定し（置き換えて），
  $(\sqrt{n}/\hat{\sigma})(\overline{X}-\mu)\cong Z$
- よって，z_α を標準正規分布の100α%点（すなわち，$P(Z\le z_\alpha)=\alpha$ を満たす z_α）とすると，

$$P\left[z_{0.025}\le\frac{\sqrt{n}(\overline{X}-\mu)}{\hat{\sigma}}\le z_{0.975}\right]=P[z_{0.025}\le Z\le z_{0.975}]$$

$$=P(Z\le z_{0.975})-P(Z\le z_{0.025})=0.975-0.025=0.95$$

- ちなみに，$z_{0.975}=1.96$

24

3．4 推定法(3)

- $z_{0.025} \le \dfrac{\sqrt{n}\left(\bar{X}-\mu\right)}{\hat{\sigma}} \le z_{0.975} \Leftrightarrow z_{0.025}\dfrac{\hat{\sigma}}{\sqrt{n}} \le \bar{X}-\mu \le z_{0.975}\dfrac{\hat{\sigma}}{\sqrt{n}}$

  $\Leftrightarrow \bar{X}-z_{0.975}\dfrac{\hat{\sigma}}{\sqrt{n}} \le \mu \le \bar{X}-z_{0.025}\dfrac{\hat{\sigma}}{\sqrt{n}}$

- $z_{0.025}=-z_{0.975}=-1.96$ **を考慮すると，**

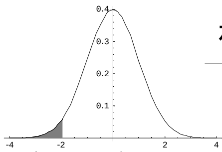
  $P\left[\bar{X}-z_{0.975}\dfrac{\hat{\sigma}}{\sqrt{n}} \le \mu \le \bar{X}+z_{0.975}\dfrac{\hat{\sigma}}{\sqrt{n}}\right]=0.95$

- **95%信頼区間は，**

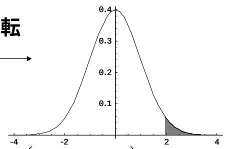
  $\boxed{\bar{X}-1.96\dfrac{\hat{\sigma}}{\sqrt{n}} \le \mu \le \bar{X}+1.96\dfrac{\hat{\sigma}}{\sqrt{n}}}$

25

$z_{0.025}=-z_{0.975}=-1.96$ の根拠



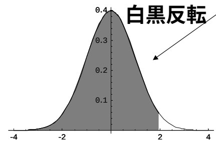
左右反転



$P(Z \le z_{0.025})=0.025$

$P(Z \ge -z_{0.025})=0.025$

白黒反転



$P(Z \le -z_{0.025})=1-0.025$

$=0.975$

$-z_{0.025}=z_{0.975}$

26

3．5 nが十分大きいとは？

- **親指ルール**
  - n＝30以上
- **厳密に考える**
  - Berry-Esseenの限界
  - $(\bar{X}-\mu)/(\sigma\sqrt{n})$の分布関数と標準正規分布の差は $3E\left[|X_i-\mu|^3\right]/(\sigma^3\sqrt{n})$ 以下である．
  - $E\left[|X_i-\mu|^3\right]$ は $\sum_{i=1}^{n}|X_i-\bar{X}|^3/n$，$\sigma^3$ は $\hat{\sigma}^3$で近似．
  - 例えば誤差1%ポイント以下かは下の式で判定

    $3\frac{1}{n}\sum_{i=1}^{n}|X_i-\mu|^3/\left(\hat{\sigma}^3\sqrt{n}\right) \le 0.01$

27

3．6 例(視聴率調査)＜1＞

- **全世帯中何%がある番組を見ているか調査**
- **各調査対象の確率モデル**
  - 各調査対象の視聴しているかどうかの分布
  - 見ていたら1，見ていなかったら0の値を各家庭に割り当てる＝確率変数化
    - (母)平均 μ，分散 σ^2 とする
    - (母)平均の意味
      - 全体の内視聴している家庭の割合をpとすると母平均は $p\times 1+(1-p)\times 0=p$ つまり母視聴率
    - 分散 $p\times(1-p)^2+(1-p)\times(0-p)^2$

      $=(1-p)p\{(1-p)+p\}=(1-p)p$

28

3．6 例(視聴率調査)＜2＞

- **視聴率調査は限られた数の世帯をランダムサンプリングで調査し母視聴率を推定する．**
- **無作為抽出しているので，標本＝調査対象家庭の視聴状況は各々独立**

  $X_i=1(\text{見ている時}),0(\text{見ていないとき})$
- **(母)視聴率＝母平均を標本平均で推定する．**

  $\bar{X}=\{X_1+\cdots+X_n\}/n$

  $=\{X_i=1,\text{つまり},\text{視聴世帯の数}\}/\text{標本数}$

29

3．6 例(視聴率調査)＜3＞

- **95%信頼区間は，** $\bar{X}-1.96\dfrac{\hat{\sigma}}{\sqrt{n}} \le \mu \le \bar{X}+1.96\dfrac{\hat{\sigma}}{\sqrt{n}}$
- $\hat{\sigma}$ **の計算**（視聴世帯数$=m$ **とすると**）

  $\hat{\sigma}^2=\dfrac{1}{n}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2=\dfrac{1}{n}\left\{\sum_{X_i=1\text{の場合}}\left(1-\bar{X}\right)^2+\sum_{X_i=0\text{の場合}}\left(0-\bar{X}\right)^2\right\}$

  $=\dfrac{1}{n}\left\{m\times(1-m/n)^2+(n-m)\times(0-m/n)^2\right\}$

  $=m/n\times(1-m/n)^2+(1-m/n)\times(m/n)^2$

  $=(m/n)(1-m/n)$

30

3.6 例（視聴率調査）＜4＞

- ９５%信頼区間は，

$$\frac{m}{n}-1.96\sqrt{\frac{m}{n}\left(1-\frac{m}{n}\right)}\Big/\sqrt{n} \le \mu \le \frac{m}{n}+1.96\sqrt{\frac{m}{n}\left(1-\frac{m}{n}\right)}\Big/\sqrt{n}$$

- ９５%信頼区間の誤差範囲表

視聴率	標本数					
	40000	1500	1000	600	250	200
5%，95%	0.2%	1.1%	1.4%	1.7%	2.7%	3.0%
10%，90%	0.3%	1.5%	1.9%	2.4%	3.7%	4.2%
20%，80%	0.4%	2.0%	2.5%	3.2%	5.0%	5.5%
30%，70%	0.4%	2.3%	2.8%	3.7%	5.7%	6.4%
40%，60%	0.5%	2.5%	3.0%	3.9%	6.1%	6.8%
50%，50%	0.5%	2.5%	3.1%	4.0%	6.2%	6.9%

31

参考Web

- ヴィデオリサーチ
  - http://www.videor.co.jp/rating/wh/contents.htm
- 日本テレビ視聴率不正操作事件について
  - http://www.ntv.co.jp/info/news/20031118.html
  - http://www.videor.co.jp/press/
  - https://www.nhk.or.jp/kdns/_wakaran/03/1122.html
  - http://www.mainichi.co.jp/eye/interview/200311/14-1.html
  - １７世帯に不正働きかけー＞２．８%の操作
  - ６００世帯での調査で視聴率１０%程度の場合，上下２．４%の誤差がある．不正働きかけは誤差範囲を上回る
  - 「．．．本件視聴率工作を思い立った。これにより担当番組の視聴率が０．１%しか上がらなかったとしても、１４．９%と１５．０%では大違いであり．．．」
    - をいをい，０．１%の差なんて誤差範囲内だろう．統計学の勉強をしろよ．＞日本テレビと当該元プロデューサ，広告屋など業界関係者

32

4. 母平均の推定（正規分布の標本の場合）

- もし母集団が正規分布だった場合はnが大きくなくても $\bar{X}$ を適当に標準化した物はある特定の分布になる
  - ある指標が様々な要因によって決まる．
  - 様々な要因はプラスマイナスの両方の値を取る
  - 平均＋様々な要因の和で決まり，中心極限定理から，指標自身が正規分布を示す
  - $X_i \sim N(\mu, \sigma^2)$（と仮定する）

33

4.1 標本平均の分布

- $\bar{X}$ の分布

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} = \frac{(X_1-\mu)+\cdots+(X_n-\mu)}{n\sigma/\sqrt{n}}$$

$$= \frac{(X_1-\mu)/\sigma+\cdots+(X_n-\mu)/\sigma}{\sqrt{n}}$$

  - $(X_i-\mu)/\sigma \sim N(0,1)$
  - 独立な正規分布の和も正規分布

$$(X_1-\mu)/\sigma+\cdots+(X_n-\mu)/\sigma \sim N(0,n)$$

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} = \frac{(X_1-\mu)/\sigma+\cdots+(X_n-\mu)/\sigma}{\sqrt{n}} \sim N(0,1)$$

34

4.2 標準化統計量とt分布

- σという解らない値で割らねばならない
- そこで，$S^2 = \sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2 \Big/ (n-1)$とすると，

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\Big/\frac{S}{\sigma} = \frac{\bar{X}-\mu}{S/\sqrt{n}} = T_{n-1}$$

$$= \frac{\bar{X}-\mu}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}/\sqrt{n}} \sim t(n-1)$$

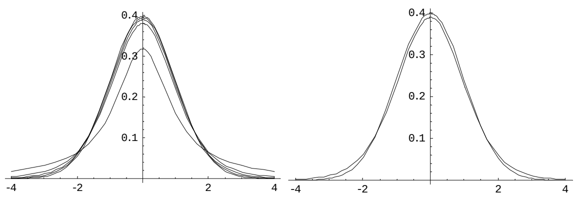
- 自由度ｎ－１のt分布になる

35

4.3 t分布の密度関数

- t分布と自由度
  - 密度関数＜左右対称＞

（左図：n=1,2,5,10,100,右図:n=10と正規分布）



36

４．４　信頼区間

- 自由度ｎ－１のt分布の100α％点 $t_{n-1,\alpha}$
  - $P(T_{n-1} \le t_{n-1,\alpha}) = \alpha$を満たす $t_{n-1,\alpha}$
  - 教科書ｐ．２８１のt分布表で調べる
- ９５％信頼区間

$$P\left(-t_{n-1,0.975} \le \frac{\overline{X}-\mu}{S/\sqrt{n}} \le t_{n-1,0.975}\right) = 0.95$$

$$\overline{X} - t_{n-1,0.975} S/\sqrt{n} \le \mu \le \overline{X} + t_{n-1,0.975} S/\sqrt{n}$$

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}$$

５．母分散の推定 (正規分布の標本の場合)

- カイ二乗分布

$$\sum_{i=1}^{n}\left(\frac{X_i-\overline{X}}{\sigma}\right)^2 = \frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2 = \mathrm{X}^2 \sim \chi^2(n-1)$$

  - 自由度ｎ－１のカイ二乗分布
  - カイ二乗分布の100α％点 $\chi^2_{n-1,\alpha}$
    $P\left[\mathrm{X}^2 \le \chi^2_{n-1,\alpha}\right] = \alpha$ を満たす $\chi^2_{n-1,\alpha}$
  - 教科書ｐ．２８０のカイ二乗分布表参照
    - ただし，αが右から取ってある点に注意

５．母分散の推定 (正規分布の標本の場合)

- ９５％信頼区間

$$P\left[\chi^2_{n-1,0.025} \le \frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2 \le \chi^2_{n-1,0.975}\right] = 0.95$$

$$P\left[\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2 \Big/ \chi^2_{n-1,0.975} \le \sigma^2 \le \sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2 \Big/ \chi^2_{n-1,0.025}\right]$$
$$= 0.95$$

$$\frac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{\chi^2_{n-1,0.975}} \le \sigma^2 \le \frac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{\chi^2_{n-1,0.025}}$$