数理統計 補助資料
～統計学入門の復習～

2023年度2学期:　　月曜1限, 水曜3限

担当教員:　　　石垣　司

本日の授業の意図

大学水準の統計学の入門的な事項を復習

- スライドの内容は2022年度の経済学部の基礎専門科目「統計学入門」の抜粋 + α
- この内容を最低限の基礎として, 以降の授業を進めていく

具体的な数式よりも(もちろん, 数式の理解も重要であり, 疎かにしてはいけないが), **統計学の舞台設定を思い出してほしい**

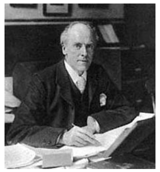
- 例えば, 記述統計と推測統計のそれぞれの意義, 確率とは, 確率を含む現象の記述方法, 期待値の計算方法, 標本の考え方とその仮定, 標本の分布の性質, 推定量とその性質, 仮説検定の設定

スライドの枚数が多量なため各内容は紹介程度に留める。もしも, 理解が曖昧な概念などがあれば各自で復習してほしい。

近代統計学(19世紀以降)

記述統計学(K. Pearsonが大成)

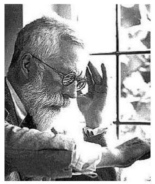
- データの平均・分散, 可視化などから分布を議論し, 対象の傾向や性質を把握
  - ヒストグラム, 標準偏差, 相関係数など



Karl Pearson 1857-1936

推測統計学(R. Fisherが体系化)

- 部分の標本から全体の構造を推定
  - 仮説検定, 最尤推定, 実験計画など



Ronald Fisher 1890-1962

#メモ　この2つの統計学の違い、特に、推測統計学のモチベーションを思い出してほしい。推測統計学では、興味のある対象の全てのデータを観測することはできないという前提に立つ。その上で、観測される一部のデータから興味のある対象全体(母集団)に対して分かることを確率の概念を用いて明らかにしたい。

様々な確率の定義 #1

確率とは？

- 偶然性を伴う事象が生じる可能性の尺度の表現

古典的確率(数学的確率)

$$P(A) = \frac{\text{事象}A\text{に含まれる根源事象の数}}{\text{標本空間に含まれる根源事象の数}}$$

Pierre-Simon Laplace 1746-1827

統計的確率

- 試行の回数 n で 事象 A の生じる回数 r のとき

$$P(A) = \lim_{n\to\infty} \frac{r}{n}$$

主観的確率

- 信念の度合いを $0 \sim 1$ の数値で表現
  - 例: 彼/彼女への告白が実る確率はせいぜい10 %程度だろう
  - 例: その新規事業は8割以上の確率で成功するだろう

様々な確率の定義 #2

確率の公理 （公理　理論を展開する前提）

1. 任意の事象 A に対し $0 \le P(A) \le 1$
2. 全事象 S に対し $P(S) = 1$
3. $\underbrace{A \cap B = \varphi}_{A\text{と}B\text{が排反}}$ ならば $P(A \cup B) = P(A) + P(B)$

Andrey Kolmogorov 1903-1987

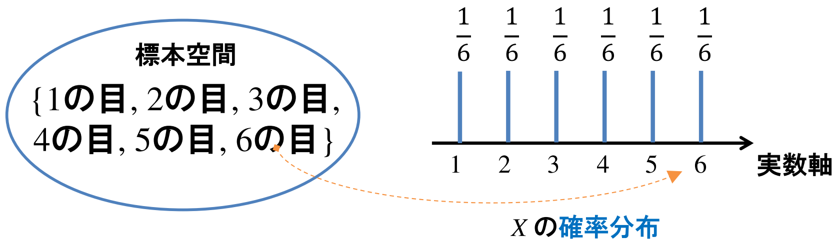
- 上の3つを満たす関数 P を確率として扱う
- 古典的・統計的・主観的な各確率は確率の公理を満たすように定義可能

  ※コルモゴロフの公理(1933)の正確な理解には確率測度が必要。その内容は本講義では扱わない

離散型確率変数

標本空間の根元事象の数が整数の場合の確率変数

- 例：サイコロの確率変数　X(1の目) = 1, …, X(6の目) = 6
- 事象の生じる確率：$P(X=1)=\frac{1}{6}, \ldots, P(X=6)=\frac{1}{6}$



- 矢印：確率変数 X
- 実数軸上の青い縦線：確率関数 $p_X(x_i)$

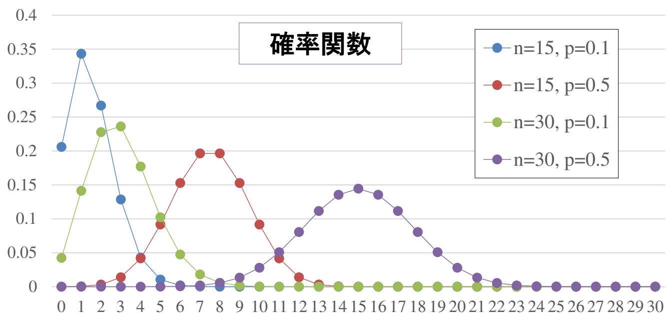
  縦線の上の数値：各事象の生じる確率

離散型確率分布の例～二項分布

n 回の独立なベルヌーイ試行の成功回数の分布

例：10回のコイン投げで表が出る回数

- 確率変数　$X = X_1 + , \cdots , + X_n$ （X_i は独立なベルヌーイ試行）
- 確率関数　$\mathrm{Binomial}(X = x) = {}_nC_x p^x (1-p)^{n-x}$
- パラメータ　n, p

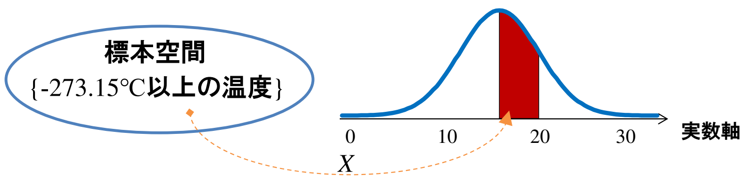


#メモ　グラフの見やすさのために点と点の間に補間直線を引いているが、離散確率分布のため点（整数）以外の実数部分の値は意味を持たないことに注意

連続型確率変数

標本空間が連続の確率変数

- 例：明日の最高気温　（確率変数 X）
- 事象がある範囲に入る確率：$P(15℃ \le X \le 20℃)$



- 矢印：確率変数 X
- 実数軸上の青い曲線：確率密度関数 $f(x)$
- 赤色部分の面積：確率 $P(15℃ \le X \le 20℃)$

確率密度関数

確率密度関数 $f(x)$

- 確率変数がとる全ての値の相対的な出現のしやすさを記述した関数
- 連続型確率分布の表現

連続型確率変数 X が範囲$[a, b]$に値をとる確率

$$P(a \le X \le b) = \int_a^b f(x)dx$$

※離散型確率変数と異なり $P(X = 15℃)$ は意味を持たない

$\because P(X = a) = 0$ （a は任意の実数）

確率密度関数の性質

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

正規分布（連続分布）

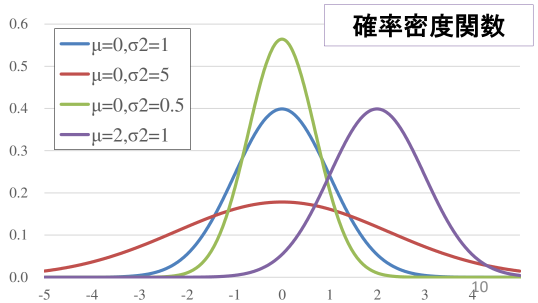
統計学で最も重要な連続分布

- 実現象や観測誤差よく表現
- 確率密度関数

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

Carolus Fridericus Gauss
(1777-1855)

- パラメータ　μ, σ^2
- $E[X] = \mu, V[X] = \sigma^2$



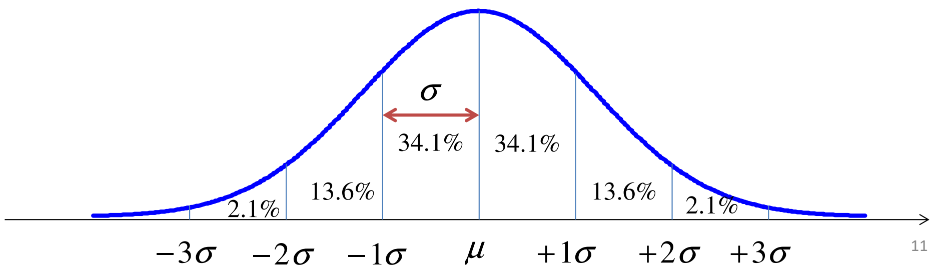
正規分布の確率密度関数の性質

左右対称の分布形

- 平均, 中央値, 最頻値は同じ
- 平均 μ と分散 σ^2 のみで分布が決まる（表記法：$N(\mu, \sigma^2)$）

標準偏差 σ

- $\pm\sigma$ 以内に事象が生じる確率 約68.2%
  ($\pm 2\sigma$は約95.5%, $\pm 3\sigma$ は約99.7%)



2変数間の離散型同時確率関数

同時確率関数　同時確率の確率関数

周辺確率関数　確率変数 X や Y の独自の確率関数

- 事象 A の確率変数　X(赤=奇数) = 0, X(赤=偶数) = 1
- 事象 B の確率変数　Y (青=2以下) = 0, Y (赤=3以上) = 1
- P_X　X の周辺確率関数
- $P_{X,Y}$　X と Y の同時確率関数

X ＼ Y	0	1	Xの周辺確率
0	$P_{X,Y}(0,0) = {}^1/_6$	$P_{X,Y}(0,1) = {}^1/_3$	$P_X(0) = {}^1/_2$
1	$P_{X,Y}(1,0) = {}^1/_6$	$P_{X,Y}(1,1) = {}^1/_3$	$P_X(1) = {}^1/_2$
Yの周辺確率	$P_Y(0) = {}^1/_3$	$P_Y(1) = {}^1/_3$	1

独立性

確率変数 X と Y はお互いに影響を与えない

- 例：青(X)と赤(Y)のサイコロの出る目は独立
  例：若年者層の割合(X)と婚姻率(Y)は非独立

独立と相関

- 独立であれば無相関
- 無相関であっても独立とは限らない

無相関で独立の例　　　無相関であるが独立ではない例

Wikipedia

独立のときの同時確率と条件付き確率

独立のときの同時確率

- 独立の定義　確率変数 X と Y の同時確率はそれぞれの確率の積に分解できるとき, X と Y は独立であるという

$$P(X,Y) = P(X)P(Y)$$

独立のときの条件付き確率

$$P(Y|X) = P(Y)$$

確率変数 X の取る値によって Y の確率は変化しない

- 2変数の同時確率関数の独立性の条件

$$p(0,0)p(1,1) = p(0,1)p(1,0)$$

X / Y	0	1
0	$p(0,0)$	$p(0,1)$
1	$p(1,0)$	$p(1,1)$

確率変数の共分散と相関係数

確率変数 X と Y の共分散

$$Cov[X,Y] = E[(X - E[X])(Y - E[Y])]$$

- 確率変数間の共変動の大きさ
- 確率変数 X と Y が独立のとき $Cov[X,Y] = 0$

確率変数 X と Y の相関係数

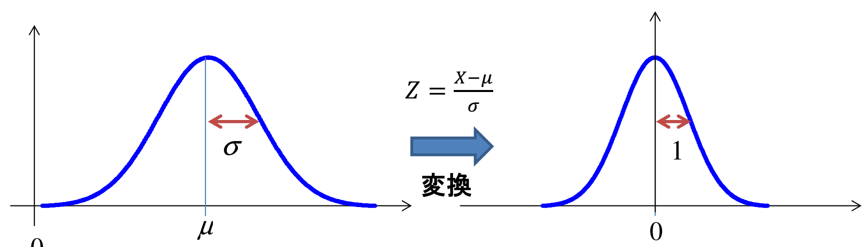
$$\rho = \frac{Cov[X,Y]}{\sqrt{V[X]V[Y]}}$$

- 確率変数 X と Y が独立のとき $\rho = 0$
- 値の範囲　$-1 \leq \rho \leq 1$

正規分布の標準化

確率変数 X を $X \sim N(\mu, \sigma^2)$ とすると,
確率変数 $Z = \frac{X-\mu}{\sigma}$ は, $Z \sim N(0,1)$ となる

- 正規分布は標準正規分布 $N(0,1)$ に変換可能
- "~"は, 確率変数 X は確率分布 $N(\mu, \sigma^2)$ から発生するという意味の記号



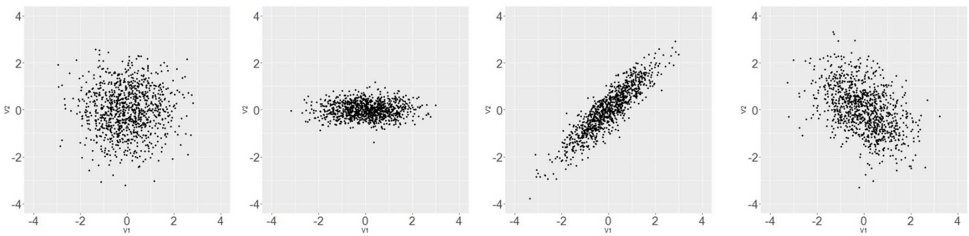
正規分布 $X \sim N(\mu, \sigma^2)$　　　標準正規分布　$Z \sim N(0,1)$

補足:2変量正規分布

確率変数 X と Y に関する正規分布

- $E[X], E[Y], V[X], V[Y], Cov[X,Y]$ で分布が決まる
  - ※学部上級・大学院以上の計量経済学では多変量扱いが必須

$V[X]=1$ $V[Y]=1$ $Cov[X,Y]=0$	$V[X]=1$ $V[Y]=0.1$ $Cov[X,Y]=0$	$V[X]=1$ $V[Y]=1$ $Cov[X,Y]=0.9$	$V[X]=1$ $V[Y]=0.5$ $Cov[X,Y]=-0.5$



各代表値を持つ2変量正規分布から1000個の乱数を発生

期待値

確率分布の代表値(平均や分散)**の計算を一般化**

離散確率分布の期待値

$$E[g(X)] = \sum_{i=1}^{m} g(x_i)p(x_i)$$

連続確率分布の期待値

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)\,dx$$

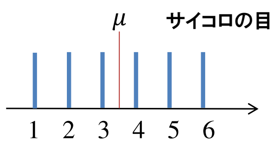
確率変数の期待値(確率分布の平均 μ)

確率分布の中心(正確には重心)**の位置を表す値**

離散確率分布の平均

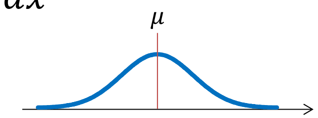
※高校数学で学習済み

$$\mu = E[X] = \sum_{i=1}^{m} x_i p(x_i)$$



連続確率分布の平均

$$\mu = E[X] = \int_{-\infty}^{\infty} xf(x)\,dx$$



確率分布の分散 σ^2

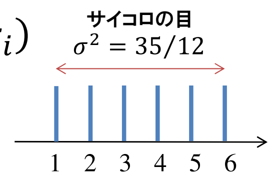
確率分布のばらつき(平均からのズレ)**を表す指標**

- $g(X) = (X-\mu)^2$ の期待値が分散

#メモ 直観的には $g(X) = |X-\mu|$ の方が理解しやすいが、絶対値よりも2乗の方が数学的な性質がよいため、このように定義される

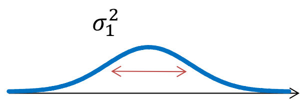
離散確率分布の分散 σ^2

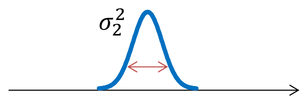
$$V[X] = E[(X-\mu)^2] = \sum_{i=1}^{m}(x_i-\mu)^2 p(x_i)$$



連続確率分布の分散 σ^2

$$V[X] = E[(X-\mu)^2] = \int_{-\infty}^{\infty}(x-\mu)^2 f(x)\,dx$$





$\sigma_1^2 > \sigma_2^2$

期待値計算の性質

定数の期待値(a, b は定数)

$$E[a] = a$$

線形性

$$E[ag_1(X) + bg_2(X)] = aE[g_1(X)] + bE[g_2(X)]$$

分散の分解 check!

$$V[X] = E[X^2] - E[X]^2$$

確率変数の1次式の分散 check!

$$V[aX + b] = a^2V[X]$$

分散共分散行列

#メモ 図のデータは分散共分散行列 V の多変量正規分布から1000個の乱数を発生

分散共分散行列 V

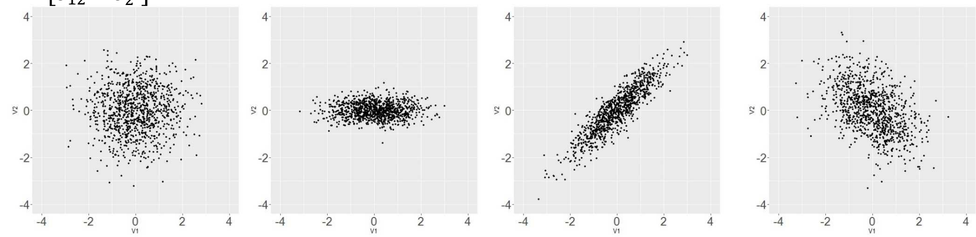
- $x_1, \dots, x_M$ を確率変数, 期待値を E で表すとき、

$$V = \begin{bmatrix} E[(x_1 - E[x_1])(x_1 - E[x_1])] & \cdots & E[(x_M - E[x_M])(x_1 - E[x_1])] \\ \vdots & \ddots & \vdots \\ E[(x_M - E[x_M])(x_1 - E[x_1])] & \cdots & E[(x_M - E[x_M])(x_M - E[x_M])] \end{bmatrix}$$

対角成分は各 x_i の分散(σ_i^2)

対角成分以外の要素は各 x_i と x_j の共分散(σ_{ij})(変数間の線形関係の強さを表す)

$V = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ $\quad V = \begin{bmatrix} 1 & 0 \\ 0 & 0.1 \end{bmatrix}$ $\quad V = \begin{bmatrix} 1 & 0.9 \\ 0.9 & 1 \end{bmatrix}$ $\quad V = \begin{bmatrix} 1 & -0.5 \\ -0.5 & 0.5 \end{bmatrix}$

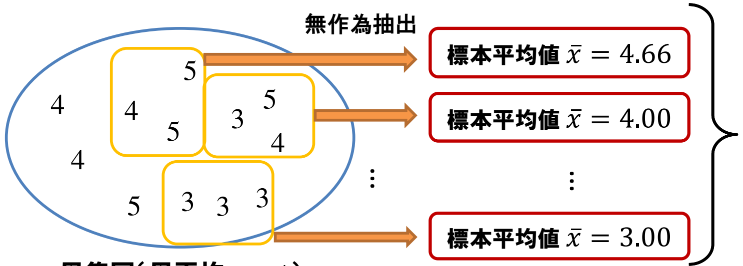


標本分布

標本から計算される代表値(など)が従う確率分布

- 例: $\mu = 4$ の母集団から $n = 3$ の標本抽出

抽出される標本は確率的に決まる



抽出された標本に依存して、実現値の値はばらつく

<u>標本平均値はある確率分布に従い発生する</u>

【これ以降の授業での設定】

- <u>標本 $\{X_1, \dots, X_n\}$ は確率変数</u>
- <u>標本は独立同一分布から無作為抽出</u>

$\{x_1, \dots, x_n\}$ は確率変数の実現値(観測値・データ)

用語の整理

独立同一分布(i.i.d.)

- 標本 $\{X_1, \dots, X_n\}$ の各 X_i は独立である
- 標本 $\{X_1, \dots, X_n\}$ は同一の母集団から抽出

母平均, 母分散

- 母集団の平均 μ と分散 σ^2

統計量(statistic)

- 標本から目的の値を計算する関数 $T(X_1, \dots, X_n)$

統計量自体も確率変数

- 例: 標本平均の統計量 $T(X_1, \dots, X_n) = \frac{1}{n}\sum_{i=1}^{n} X_i$

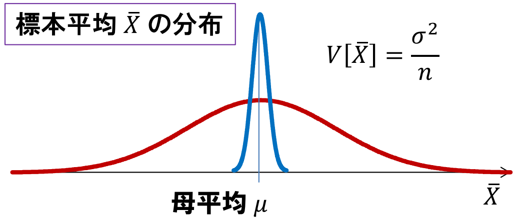
標本平均の分布

標本平均 $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

$\bar{X}$ は確率変数（サンプルサイズ n の標本抽出を1回の試行とみなすと，試行ごとに標本平均の実現値は確率的に決まる）

標本平均の期待値 $E[\bar{X}] = \mu$ check!

標本平均の分散 $V[\bar{X}] = \frac{\sigma^2}{n}$ check!



n が大きいと分散が小さい
$|\mu - \bar{x}|$ が小さくなる確率が高い

n が小さいと分散が大きい
$|\mu - \bar{x}|$ が大きくなる確率が高い

大数の法則

統計的推測の根幹となる定理


Jakob Bernoulli
(1654-1705)

$$\lim_{n\to\infty} P(|\bar{X} - \mu| \geq \varepsilon) = 0$$ check!

標本平均に対するチェビシェフの不等式

- サンプルサイズ n が無限大に近づくと，$\bar{X}$ は μ に限りなく近づくことを数学的に保証

  （より正確な意味）母平均と母分散が有限であれば，n を無限大に近づけていくことで，標本平均 $\bar{X}$ と母平均 μ のズレの大きさが ε よりも大きくなる確率は 0 である

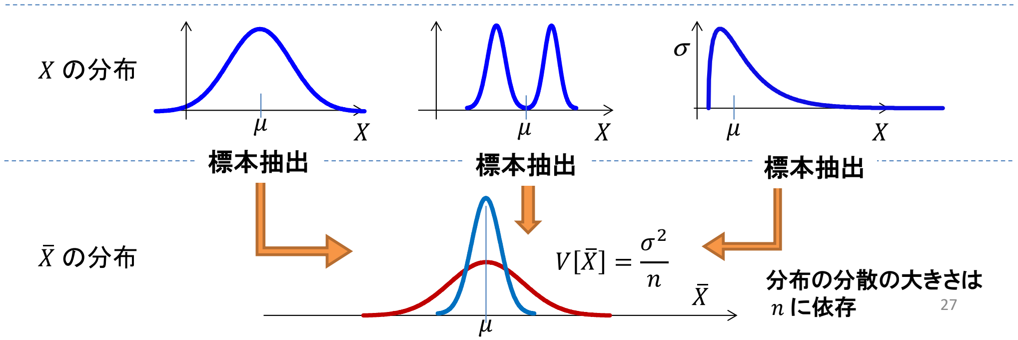
- 統計的確率と数学的確率の一致

  例：サイコロを無限回投げ続ける(統計的確率)と 1 の目の出る割合は，理論値の $1/6$（数学的確率）に一致する

中心極限定理

<u>任意の確率分布</u>(平均 μ, 分散 σ^2 は有限)の標本平均の分布は，サイズ n が十分大きい時 $N\left(\mu, \frac{\sigma^2}{n}\right)$ に近似できる

統計学の定理の中で最も重要かつ非自明な定理の一つ
積率母関数(モーメント母関数)による証明 check!



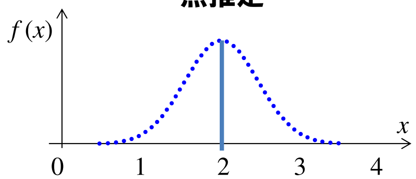
推定

標本から母集団の分布のパラメータ(母数)を求める

- **点推定**：未知のパラメータ θ を1点で推定
- **区間推定**：未知のパラメータ θ が存在する区間を推定
- データ(標本の実現値)を利用して未知 θ の値を推定

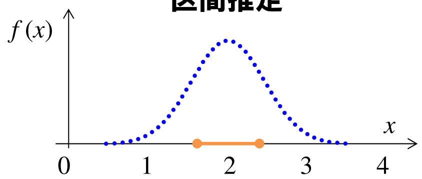
  例　データ$\{2, 3, 1, 2, 3, 4, \ldots, 2\}$から母集団の母平均を推定

点推定



$\hat{\theta} = 2$

区間推定



$P(1.7 \leq \hat{\theta} \leq 2.3)$

推定量 #1

推定量(estimator)

- **標本からパラメータを推定する関数**
  - **統計量と同じ。推定が目的の場合の統計量。推定量も確率変数**
  - **推定値：観測値から推定量を計算した実際の値**
- **無作為標本**　$\boldsymbol{X} = \{X_1, \ldots, X_n\}$
- **パラメータ**　θ
- **推定量**　$T(\boldsymbol{X})$ or $\hat{\theta}$

- **例：標本平均の推定量**

$$\hat{\theta} = T(\boldsymbol{X}) = \bar{X} = \underbrace{\frac{1}{n}\sum_{i=1}^{n} X_i}_{\text{推定量}}$$

推定量 #2

推定量は自由に設定可能

- **例：母平均 μ の推定量の候補**

$$T_1(\boldsymbol{X}) = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$T_2(\boldsymbol{X}) = \frac{1}{n+1}\left(X_1 + \sum_{i=1}^{n} X_i\right)$$

$$T_3(\boldsymbol{X}) = \frac{1}{2}(X_1 + X_n)$$

$T_1(\boldsymbol{X})$ が標本平均の推定量として用いられるのは，
不偏性，一致性，よい効率性をもつため

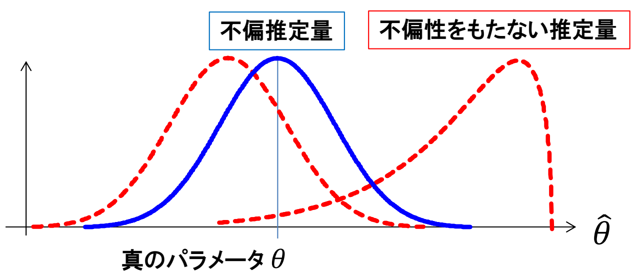
不偏性 (Unbiasedness)

推定量の期待値がパラメータ θ となる性質

$$E[T(\boldsymbol{X})] = \theta$$

不偏推定量　不偏性をもつ推定量

- **「推定量 $T(\boldsymbol{X})$ は不偏である」**
- **平均的に過大評価も過小評価もしない推定量**



一致性 (Consistency)

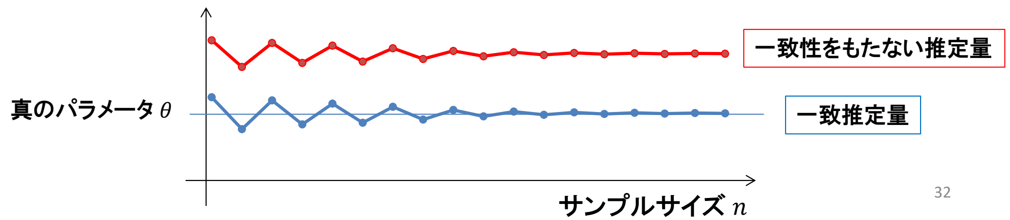
サンプルサイズ n が大きくなると推定量がパラメータ θ に近づく性質

$$\lim_{n\to\infty} P(|T(\boldsymbol{X}) - \theta| > \varepsilon) = 0 \quad (\varepsilon \text{ は正の定数})$$

推定量 $T(X)$ と母数 θ のズレが ε よりも大きい確率はゼロに近づく

一致推定量：一致性をもつ推定量

- **推定量が一致性を持たない場合，サンプルサイズ n をどれだけ大きくしても，母集団のパラメータ(一点)に一致しない**

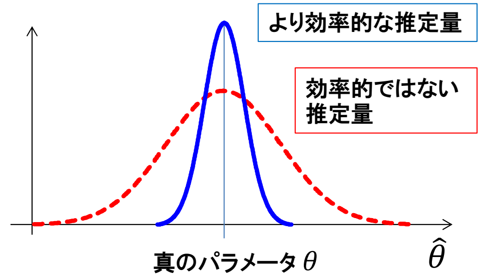


効率性(Efficiency)

効率的な推定量: 分散のより小さな推定量

最小分散推定量: 最も分散の小さな推定量

不偏推定量かつ一致推定量が複数ある場合, さらに良い推定量を選ぶ基準として用いられる



問題

$$T_1(\boldsymbol{X}) = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$T_2(\boldsymbol{X}) = \frac{1}{n+1}\left(X_1 + \sum_{i=1}^{n} X_i\right)$$

– $T_1(\boldsymbol{X})$と $T_2(\boldsymbol{X})$ は不偏推定量かつ一致推定量である。$V[T_1(\boldsymbol{X})]$ と $V[T_2(\boldsymbol{X})]$ からどちらがより効率的な推定量となるか確かめよ

その他の性質

最小分散不偏推定量

– あらゆる不偏推定量の中で, 最も効率的な推定量

正規母集団 $N(\mu, \sigma^2)$ からの標本に関して,

「標本平均 $\bar{X}$ は母平均 μ の最小分散不偏推定量」

「標本分散 S^2 は母分散 σ^2 の最小分散不偏推定量」

"平均"や"分散"を当然のように推定量として使用しているが, その理論的正当性の保証

その他の性質

– 漸近正規性, 漸近有効性, 十分統計量, 頑健性など

標本の誤差の見積もり～信頼区間 #1

母集団が正規分布 $N(\mu, \sigma^2)$ のときの標本平均

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

例: 正規母集団の母平均 μ の95%信頼区間

$$\left[\bar{X} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right]$$ check!

– **標準誤差(SE)** 標本平均の分布の標準偏差 $\sqrt{V[\bar{X}]} = \frac{\sigma}{\sqrt{n}}$

– "1.96×標準誤差"で誤差の大きさを見積もり

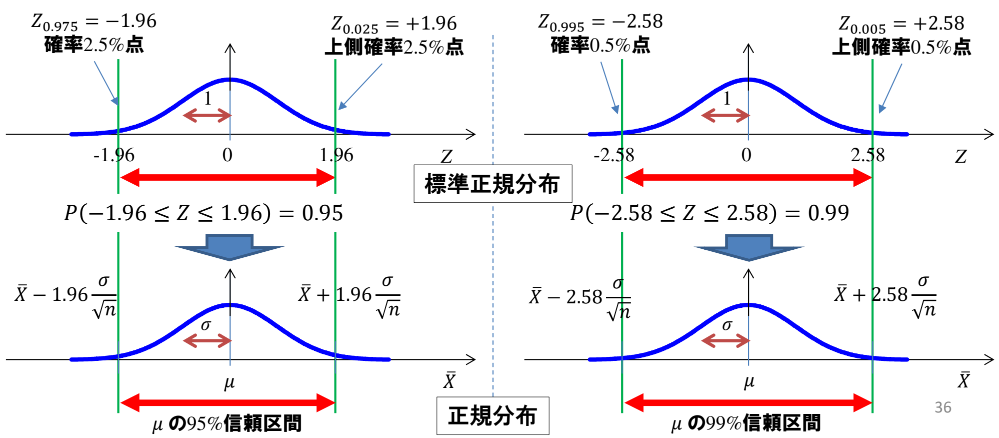
$$P(-1.96 \le Z \le 1.96) = 0.95$$

数値1.96は標準正規分布表より求まる

標本の誤差の見積もり～信頼区間 #2

95%信頼区間　　Z_α: 標準正規分布の上側確率 $100 \times \alpha$%点

– サンプルサイズ n の標本抽出を100回繰り返したら場合, 約95回はその区間$\left[\bar{X} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right]$の中に μ が含まれる



標本の誤差の見積もり～信頼区間 #3

95%信頼区間(μに関する標本平均)の意味・解釈

- 【正】 サイズ n の標本抽出を100回繰り返した場合, 約95回はその区間の中に μ が含まれる
- 1つのサイズ n の標本から計算された95%信頼区間を考える
- 【誤】 その区間の中に μ が95%の確率で出現する
  母平均 μ は定数である(確率変数ではない)
- 【正】 確率95%で, その区間は μ を含む
  より正確には「その区間は μ を含む or 含まないのどちらかであり, “μ を含む”という事象が生じる確率は95%である」
- 【誤】 その区間の中心ほど, μ に近い可能性が高い

小標本の区間推定に必要な確率分布

William Sealy Gosset
1876-1937
ギネスビール社員
“Student” はペンネーム

Student の t 分布 (1908)

- t 統計量の従う確率分布
- ビールの品質管理から誕生
  醸造器内の酵母数を標本抽出し顕微鏡で数えて厳密管理
  当初は $\sigma^2 = S^2$ としてビール品質の監視
  しかし, 小標本のとき, S^2 の値がおかしいと気付く
  $\bar{X}$ の信頼性評価に正規分布(with $\sigma = S/\sqrt{n}$)は不適の疑念をもつ
- ⇒小標本のときの S^2 の正しい見積もりが必要

カイ2乗分布 $\chi^2_{(n)}$

- 標本分散 S^2 に関連する確率分布

東京大学出版会「統計学入門」p.203コラム、D.サルツブルク「統計学を拓いた異才たち」第3章

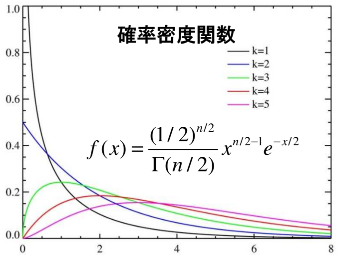
χ^2 分布

自由度 n のカイ2乗分布 $\chi^2_{(n)}$

- 確率変数 $Z_i \sim i.i.d.N(0,1)$ の2乗和 χ^2 の従う分布

$$\chi^2 = Z_1^2 +, \dots, + Z_n^2$$

- パラメータ　自由度 n (サンプルサイズ n に対応)
- 平均と分散　$E[\chi^2] = n, V[\chi^2] = 2n$
- 再生性をもつ
  $X_i \sim \chi^2_{(n)}$ (自由度 n の χ^2 分布)
  $Y_i \sim \chi^2_{(m)}$ (自由度 m の χ^2 分布)
  $X_i + Y_i \sim \chi^2_{(n+m)}$



Wikipedia (public domain)

t 分布と χ^2 分布の関係

t 統計量は, $\bar{Z}$ と χ^2 分布に従う確率変数の比

- 標本 $X_i \sim i.i.d.N(\mu, \sigma^2)$ のとき U は自由度 $n-1$ の χ^2 分布に従う

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

- t 統計量

check!

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} = \frac{\bar{Z}}{\sqrt{U/(n-1)}}$$

← 標準正規分布
← 自由度 $n-1$ の χ^2 分布

標準正規分布に従う確率変数と χ^2 分布に従う確率変数の比
両者の密度関数の比から t 統計量の密度関数を生成

<u>t 統計量が従う分布を自由度 $n-1$ の t 分布とよぶ</u>

$U \sim \chi^2_{(n)}$ の証明は煩雑なため省略。興味のある人は、「高杉&馬場, 演習統計学キャンパス・ゼミ, マセマ出版社」の p.168 などを参照

t 分布

自由度 n の t 分布

- $Z \sim N(0,1)$, $W \sim \chi^2_{(n)}$, Z と W は独立のとき, t が従う分布

$$t = \frac{Z}{\sqrt{W/n}}$$

- パラメータ　自由度 n
- 平均と分散

  $E[t] = 0,\ V[t] = \frac{n}{n-2}$ $(n > 2)$

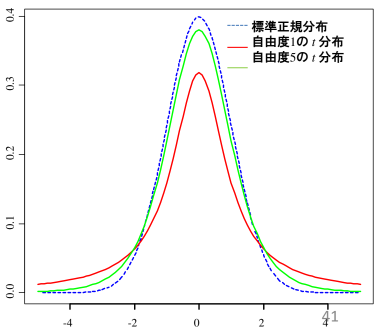
  $n = 1$ のときコーシー分布
- 原点に対象
- 自由度 n が大きい時($n > 30$ くらい)

  正規分布にほとんど一致

  $n \to \infty$ の時, 正規分布

確率密度関数

$$f(x) = \frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\,\Gamma(n/2)}(1+x^2/n)^{-(n+1)/2}$$



統計的仮説検定って？

母集団の分布の母数に関する仮説の妥当性を標本から検証する方法

- 例1：サイコロを600回投げたとき1の目が120回も出た。この回数は偶然の範囲内か？
- 例2：2020年8月に実施されたTOEFL IPTテストを自主的に受検した東北大学生のスコアの平均値は500点であった。ここで, 未受検の学生30名を無作為抽出して同じテストを受検させたときのスコアの平均値が480点であったとする。このスコアの違いは偶然の範囲内といえるか？
- 例3：2021年度の東北大学の入学者の割合（入学者数 / 志願者数）は宮城県出身者は29.4%, 北海道出身者は41.0%であった。この入学者の割合の違いは偶然の範囲内か？

  数学的内容は既学習+α。問題設定を理解することが重要

仮説検定のロジック（帰無仮説 vs 対立仮説）

帰無仮説：“仮説は正しくない”を表現する仮説

対立仮説：“仮説は正しい”を表現する仮説

- 例2：自ら受検した学生と未受検の学生の母集団のTOEFL IPTの平均スコアに差はあるか？

  帰無仮説：“平均スコアに差はない”

  対立仮説：“平均スコアに差はある”

仮説検定のロジック

- 仮説に対する標本が得られていて, 帰無仮説が正しいという設定の下でその標本が生じる確率を計算すると, その確率は有意水準よりも小さい

  ⇒ ゆえに, 対立仮説を支持する

検定統計量

仮説検定で使用する統計量

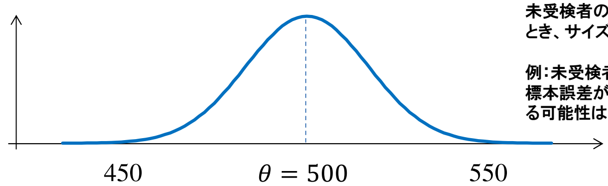
- 帰無仮説を表現する検定統計量の分布と実際の観測値とのズレの大きさを測るために利用
- 例2：TOEFL IPTの平均スコアは異なるか？

  ⇒ 検定統計量に標本平均 $\bar{X}$ を用いる

  帰無仮説：“平均スコアに差はない”

  対立仮説：“平均スコアに差はある”

帰無仮説が成立すると仮定したときの検定統計量 $\bar{X}$ の分布の例



未受検者の母集団の母平均も 500 であると仮定したとき、サイズ $n = 30$ の標本が従う標本平均 $\bar{X}$ の分布

例：未受検者の母集団の母平均も 500 であるならば、標本誤差があったとしても平均スコアが400や600となる可能性は小さい

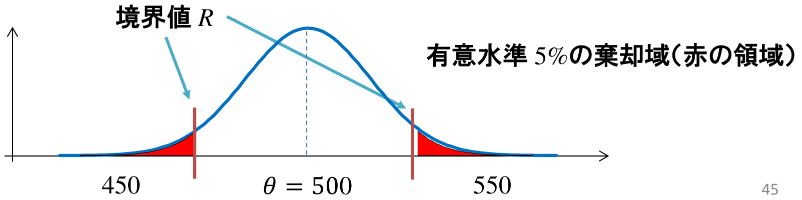
棄却域と有意水準

棄却域: 帰無仮説から離れたある領域(範囲)

有意水準: ある事象が起こる確率が偶然とは考えにくい(有意である)と判断する基準となる確率(デジタル大辞泉)

– 検定統計量の分布(確率密度関数)の棄却域での面積
– 通常, 0.05 (5%) や 0.01 (1%)を設定
– **境界値**: 棄却域の境界線の値

帰無仮説が成立すると仮定したときの検定統計量 $\bar{X}$ の分布



帰無仮説の棄却と解釈

帰無仮説の「棄却」

– "帰無仮説は統計的に認められない"という結果
対立仮説が採択される, ともいう
– 例:「5%有意水準で棄却」,「5%水準で有意差あり」

帰無仮説が棄却された場合
⇒ 対立仮説が統計的に認められる

帰無仮説が棄却できない場合
⇒ **対立仮説に関して何もわからない**
※「対立仮説が棄却される」,「帰無仮説が採択される」ではないことに注意

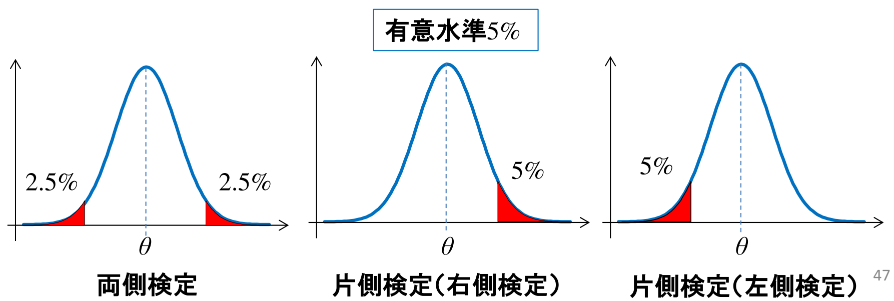
両側検定と片側検定 #1

両側検定: 棄却域を分布の両端に定める

– 例: 平均スコアは異なるか?

片側検定: 棄却域を分布の片端に定める

– 例: 未受検者の平均スコアは高いか?(右側検定)
– 例: 未受検者の平均スコアは低いか?(左側検定)



両側検定と片側検定 #2

帰無仮説 H_0 と対立仮説 H_1 の立て方

θ は母集団の母数の値(例:未受検者の本当の平均スコア)
θ_0 は帰無仮説での母数の値(例: $\theta_0 = 500$)

– 両側検定 $H_0: \theta = \theta_0, \quad H_1: \theta \neq \theta_0$
– 右側検定 $H_0: \theta = \theta_0 \text{ or } \theta \leq \theta_0, \quad H_1: \theta > \theta_0$
– 左側検定 $H_0: \theta = \theta_0 \text{ or } \theta \geq \theta_0, \quad H_1: \theta < \theta_0$

– 帰無仮説は適切な検定統計量を定義できるように定める
例 "$H_0: \theta \neq \theta_0$" では, 平均の検定統計量はうまく定義できない

– 帰無仮説に不等号を用いる場合でも, 実際には等号が成り立つ下での検定統計量を用いる

検定の例 #1

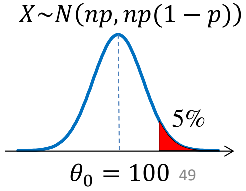
サイコロAを600回投げたとき1の目が120回も出た。これは1の目が多く出やすいイカサマサイコロか?

1. **仮説を立てる**
   - **帰無仮説:「イカサマサイコロではない」** $H_0 : \theta = \theta_0$
   - **対立仮説:「イカサマサイコロである」** $H_1 : \theta > \theta_0$

     1の目が出る回数が多すぎるという検定

     θ_0: 普通のサイコロで1の目が出る回数の理論値

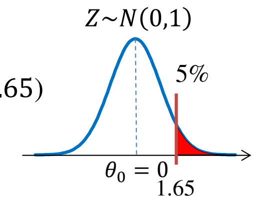
     θ: サイコロAの真の母数から計算される1の目が出る回数
2. **検定統計量 X を定める**
   - **1の目の出る回数の確率変数** $X \sim \mathrm{Binominal}(n, 1/6)$
   - **n が大きいので中心極限定理より正規分布で近似**
3. **有意水準を定める**
   - **5%有意水準($\alpha = 0.05$)の片側検定(右側検定)**



検定の例 #2

サイコロAを600回投げたとき1の目が120回も出た。これは1の目が多く出やすいイカサマサイコロか?

4. **標本から検定統計量を計算**
   - $E[X] = np = 100, V[X] = np(1-p) ≒ 83.3$
   - **帰無仮説の下では 1の目が出る回数は** $X \sim N(100, 83.3)$
   - $X = 120$ **のときの標準化統計量** $Z = (X - E[X])/\sqrt{V[X]} ≒ 2.19$
5. **検定統計量の分布から棄却域を求める**
   - $Z_{0.05} = 1.65$ **より棄却域は** $Z > 1.65$ (**境界値** $R = 1.65$)
6. **帰無仮説を棄却できるかの検定**
   - $Z > R$ **となり** Z **は棄却域の中にあるため帰無仮説は棄却される**
   - **結論　このサイコロAは有意水準5%でイカサマサイコロである**

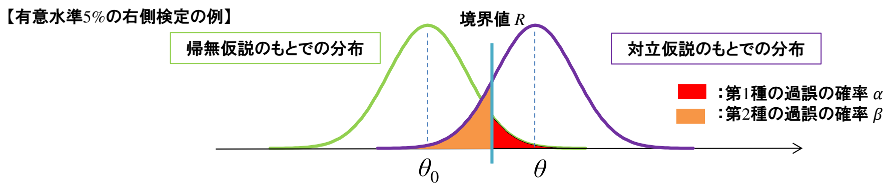


検定の誤り

第1種の過誤 (Type I error, False positive)
- **本当は帰無仮説を棄却できないのに, 検定では棄却してしまう間違え**
- P(**帰無仮説を棄却する | 帰無仮説が正しい**) $= \alpha$　(α **は有意水準**)

第2種の過誤(Type II error, False negative)
- **本当は帰無仮説を棄却できるのに, 検定では棄却できない間違え**
- P(**帰無仮説を棄却できない | 対立仮説が正しい**) $= \beta$

		検定結果	
		帰無仮説を棄却	棄却できない
真の構造	対立仮説が正しい	正しい検定	第2種の過誤
	帰無仮説が正しい	第1種の過誤	正しい検定



P 値

観測された有意水準

- **P 値が事前に設定された有意水準 α よりも小さければ, 帰無仮説は棄却できる**

- **右側検定での例**

  サイコロを600回投げたとき1の目が120回も出た。これは1の目が多く出やすいイカサマサイコロか?

- $X = 120$ **のときの標準化統計量** $Z = (X - E[X])/\sqrt{V[X]} ≒ 2.19$

$$p \text{ value} = P(Z > 2.19 | \theta = \theta_0)$$

- **標準正規分布表から P 値** $= 0.014$
- **5%水準では有意差あり, 1%水準では有意差なし**

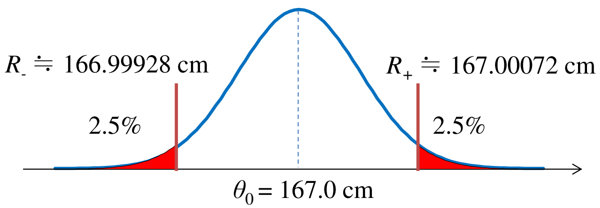
仮説検定の実用上の注意 #1

n を大きくすれば, ほとんどの帰無仮説が棄却される

- 検定の結果として有意差があった場合でも, 現実の問題としての意味を考察する必要あり
- 例　日本人の成人男性の平均身長が167.0cmと異なるかを $n = 100{,}000$ 人の標本から検定

  $X \sim N(167, 6^2)$ と仮定すると $V[\bar{X}] = 0.00036$ cm

  棄却域(±0.00072cm)に現実的な意味があるだろうか？



5%水準での両側検定

仮説検定の実用上の注意 #2

有意水準で測るのは仮説の統計的な確度のみ
⇒　仮説の効果の大きさは測っていない

誤用例

- 既存薬と比較し新薬Aは5%有意水準で効果の差あり
- 既存薬と比較し新薬Bは1%有意水準で効果の差あり
- よって, 新薬Bの方が新薬Aよりも効果が高い

誤用の可能性がある例

- $n = 3{,}000$ 人の調査結果から, 既存薬と比較して新薬Cは1%有意水準で効果の差あり
- よって, 新薬Cは既存薬よりも効果が高い

仮説検定の実用上の注意 #3

“なぜ有意差が生じたのか？” の情報はない

- 例1：サイコロを600回投げたとき1の目が120回も出た。この回数は偶然の範囲内か？

⇒ サイコロに何らの偏りがあるかどうかを調べることで比較的容易に原因を特定できる可能性がある

- 例3：2021年度の東北大学の入学者の割合（入学者数 / 志願者数）は宮城県出身者は29.4%, 北海道出身者は41.0%であった。この入学者の割合の違いは偶然の範囲内か？

⇒ この事象が生じる要因はいくつも考え得るし, 原因は複合的である可能性もある。加えて, ある仮説を検証・実証できるかは, また別の問題

「北海道から仙台へ越境するには何か強い動機があるはずだ。だから入学率が高いのでは？」という仮説を立てたときに, それを実証するのは容易ではない

＃　例3では、2021年度の大学入試の志願者と入学者は全数調査のため標本調査ではない。しかし, 各受験者の入学・非入学はベルヌーイ分布から発生する確率変数と見なすことができる。その上で、宮城県出身者の合否のベルヌーイ分布の平均母数 p_M と北海道出身者の合否のベルヌーイ分布の平均母数 p_H が異なるかどうかを検定（帰無仮説 $p_M = p_H$）するための検定統計量を適切に定義できる

① 1つの正規母集団で母分散が既知の平均の検定

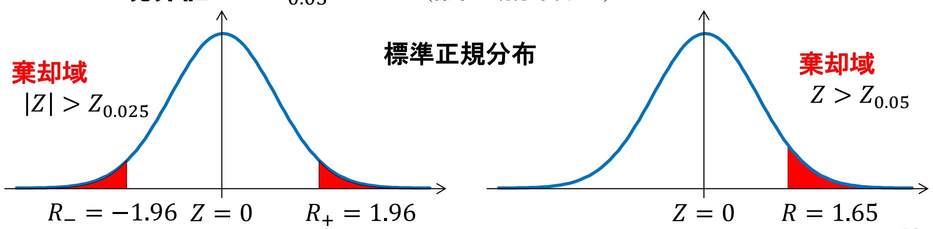
検定統計量　$Z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$　(標準正規分布)

- 帰無仮説　$\mu = \mu_0$　(μは母平均。μ_0：帰無仮説の下での平均)
- 例　有意水準5%の両側検定 (対立仮説 $\mu \neq \mu_0$)

  境界値 $R_+ = Z_{0.025} = 1.96$ (標準正規分布表より)

  境界値 $R_- = -Z_{0.025} = -1.96$
- 例　有意水準5%の右側検定 (対立仮説 $\mu > \mu_0$)

  境界値 $R = Z_{0.05} = 1.65$ (標準正規分布表より)



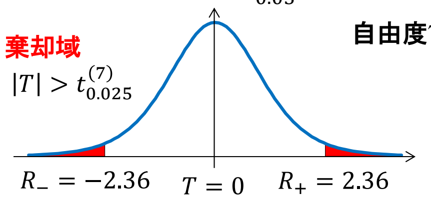
（両側検定）　（右側検定）

② 1つの正規母集団で母分散が未知の平均の検定

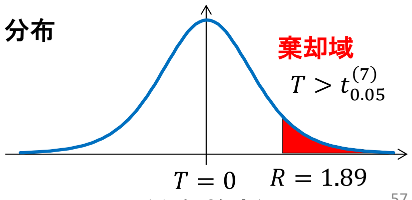
検定統計量 $T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$ （自由度 $n-1$ の t 分布）

- **帰無仮説** $\mu = \mu_0$
- **例** $n = 8$ **での有意水準**5%**の両側検定** (**対立仮説** $\mu \neq \mu_0$)
  - **境界値** $R_+ = t_{0.025}^{(7)} = 2.36$ （t 分布表より）
  - **境界値** $R_- = -t_{0.025}^{(7)} = -2.36$
- **例** $n = 8$ **での有意水準**5%**の右側検定** (**対立仮説** $\mu > \mu_0$)
  - **境界値** $R = t_{0.05}^{(7)} = 1.89$ （t 分布表より）

自由度7の t 分布



（両側検定）



（右側検定）