2024年3月5日

@高知大学 物部キャンパス

令和5年DS/DXセミナー

統計基礎

大阪大学 数理・データ科学教育研究センター

野島　陽水

点推定・区間推定・仮説検定

1. 点推定：１点のみでの推定する。情報量は少なく、計算コストも少ない。推定値はどれくらい信頼できる値なのかが不明。

   標本平均、標本不偏分散、標本標準偏差

2. 区間推定：信頼性を持たせた区間（信頼区間）でもって推定する。点推定よりも情報量はやや多いが、計算コストも多い。（標本平均はバラつく、点推定のみではバラツキの情報はわからない）

3. 仮説検定：標本の値（平均や分散など）を使って、母集団の値（平均や分散など）に差があるかどうかを確率的に検出することであり、標本で表された差が本質的な差なのか、偶然によるものなのかを確率的に（対応する分布を使って）判断すること。

- 加納克己/高橋秀人共著、基礎医学統計学改訂第７版（南江堂）
- 資料内で出てくる図等は左記の教科書から引用しています。

点推定・区間推定・仮説検定

1. 点推定：１点のみでの推定する。情報量は少なく、計算コストも少ない。推定値はどれくらい信頼できる値なのかが不明。

   標本平均、標本不偏分散、標本標準偏差

2. 区間推定：信頼性を持たせた区間（信頼区間）でもって推定する。点推定よりも情報量はやや多いが、計算コストも多い。（標本平均はバラつく、点推定のみではバラツキの情報はわからない）

3. 仮説検定：標本の値（平均や分散など）を使って、母集団の値（平均や分散など）に差があるかどうかを確率的に検出することであり、標本で表された差が本質的な差なのか、偶然によるものなのかを確率的に（対応する分布を使って）判断すること。

量的変数の記述と用いる指標

標本の代表値

標本平均

標本の総和を標本の大きさで割った値

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

平均はよく使われるが、他と極端に異なった値（外れ値）があると代表値として不適当となる。

母平均はμ（ミュー）で表す。

加重平均

特定の要素が異なる複数の標本間で平均が知りたい場合、加重平均（$\bar{X}w$）として算出する。

$$\bar{X}w = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}$$

$f_1 \cdots f_n$: 重み

量的変数の記述と用いる指標

標本の代表値

中央値

標本の各個体を小さい順に並べたとき中央に位置する値。
外れ値の影響をあまり受けない。

標本の大きさ（n）が偶数の場合は中央に最も近い２つの値の平均で、

奇数の場合は小さい順に$\dfrac{n+1}{2}$番目の値で定められる。

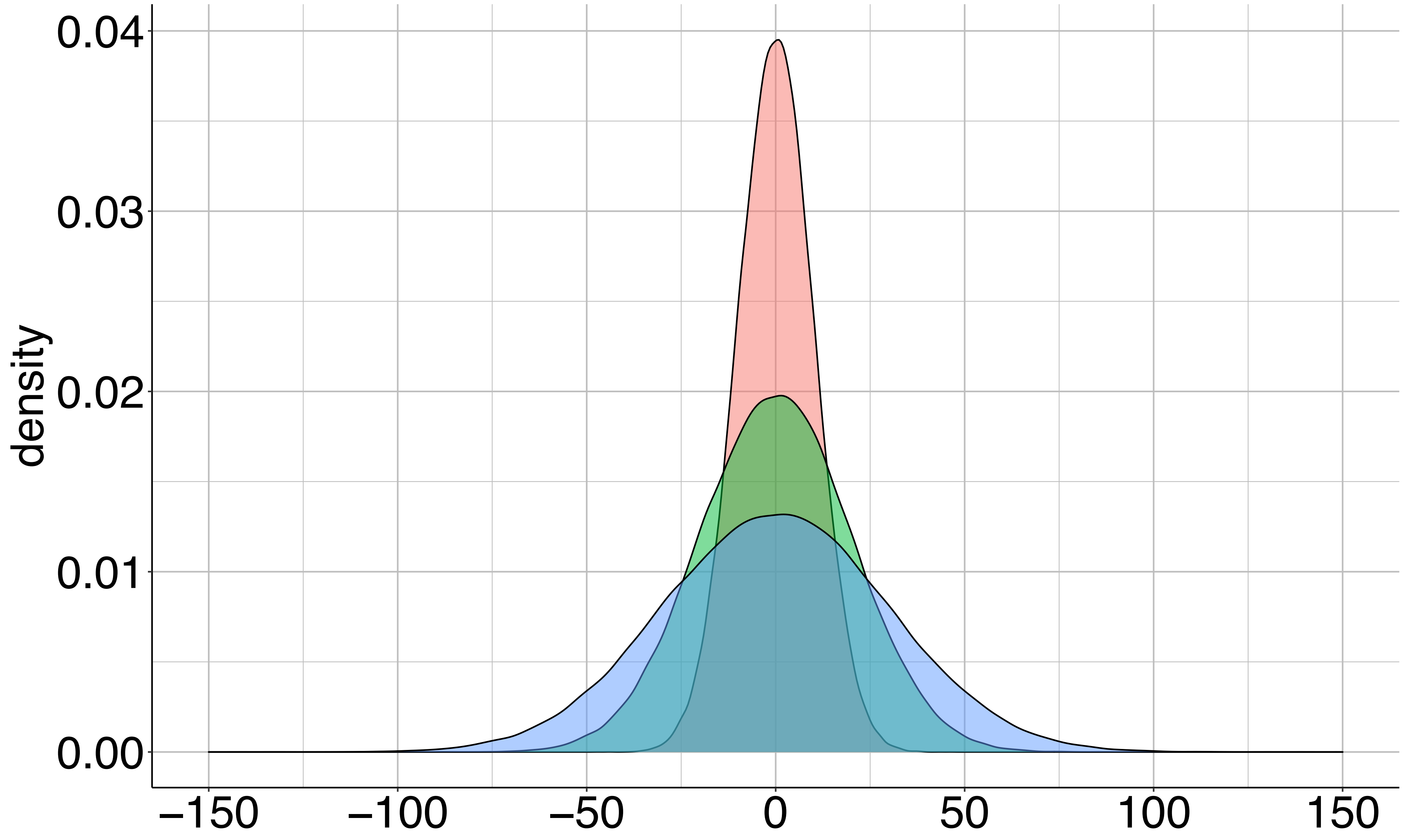
$$M_e = \begin{cases} x_{(\frac{n+1}{2})} & \cdots n\text{が奇数の場合} \\ \dfrac{1}{2}\left(x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}\right) & \cdots n\text{が偶数の場合} \end{cases}$$

標本の散布度

- 散らばりを示す指標（散布度）

下図に３つのデータの確率密度分布がある。

平均値（および中央値）はいずれも同じであるが、データの散布度が異なる。



代表値のみでは、データの特徴は理解できない。

代表値の周辺にどのように分布しているかを理解することも重要である。

標本の散布度

- 偏差：平均からどれだけ離れているかを指し示す値、$Xi - \bar{X}$で表す。

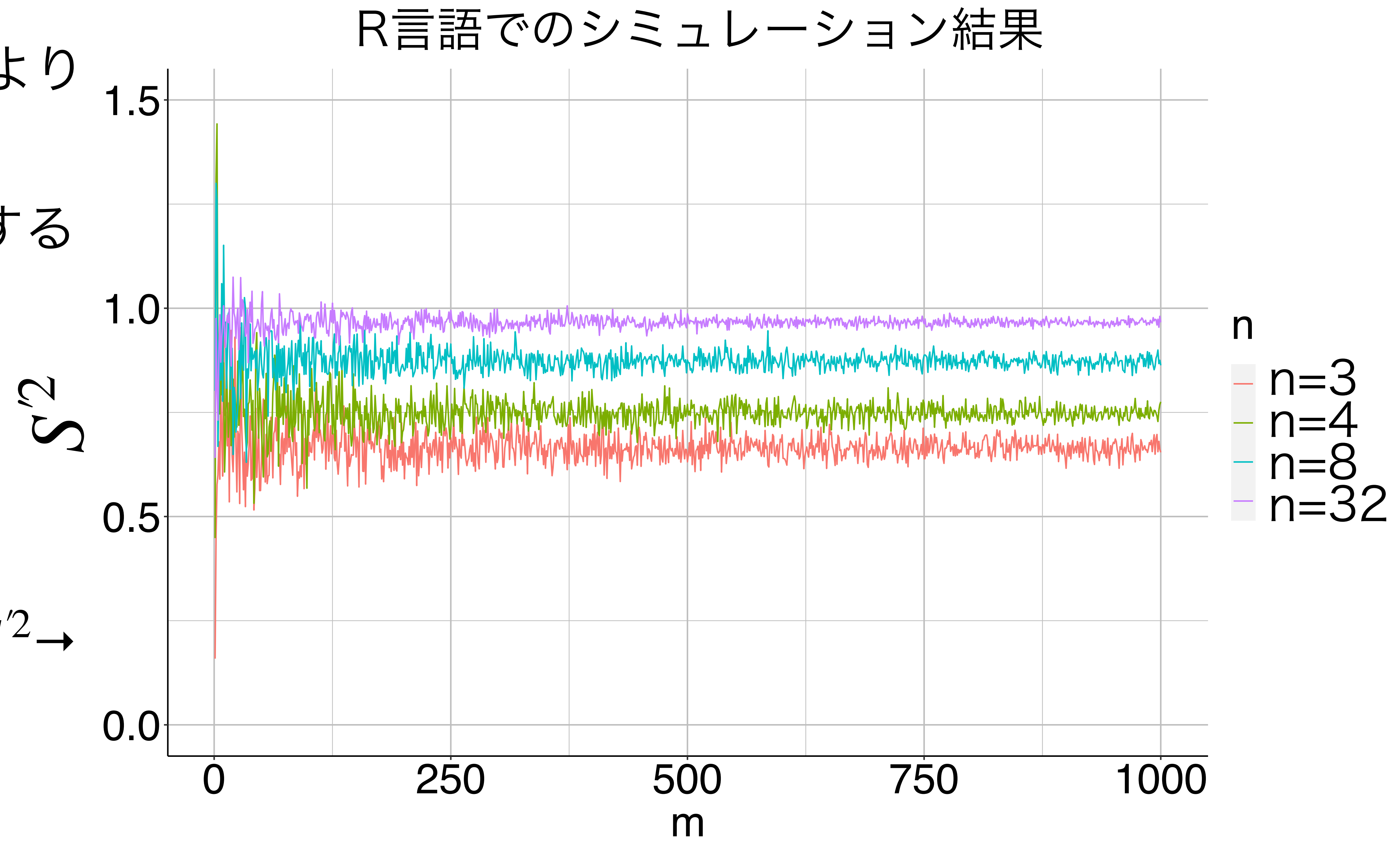
- 偏差の和：$\sum_{i=1}^{n}(X_i - \bar{X})$：常に０

- 偏差平方和：偏差の２乗の総和、$\sum_{i=1}^{n}(X_i - \bar{X})^2$　なぜ、2乗？：偏差の和は0。偏差を2乗→符号の影響をなくし偏差のばらつきが分かる。

標本の散布度

- 標本分散：「平均からの差の二乗」の平均。$S'^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$
- 標本分散の平方根を標本標準偏差という　$S' = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}$
- 標本分散を$S'^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$とすると、$S'^2$の分布は、母分散$\sigma^2$より小さい値を中心に分布する。したがって、$S'^2$では母分散を推定することはできない。

分散が１である母集団からn=3, 4, 8, 32の標本をm回抽出したときのS'^2→



標本の散布度

- そこで、S'^2に$\dfrac{n}{n-1}$を乗じてこれをS^2とすると、$S^2 = \dfrac{n}{n-1}S'^2 = \dfrac{n}{n-1} \times \dfrac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$

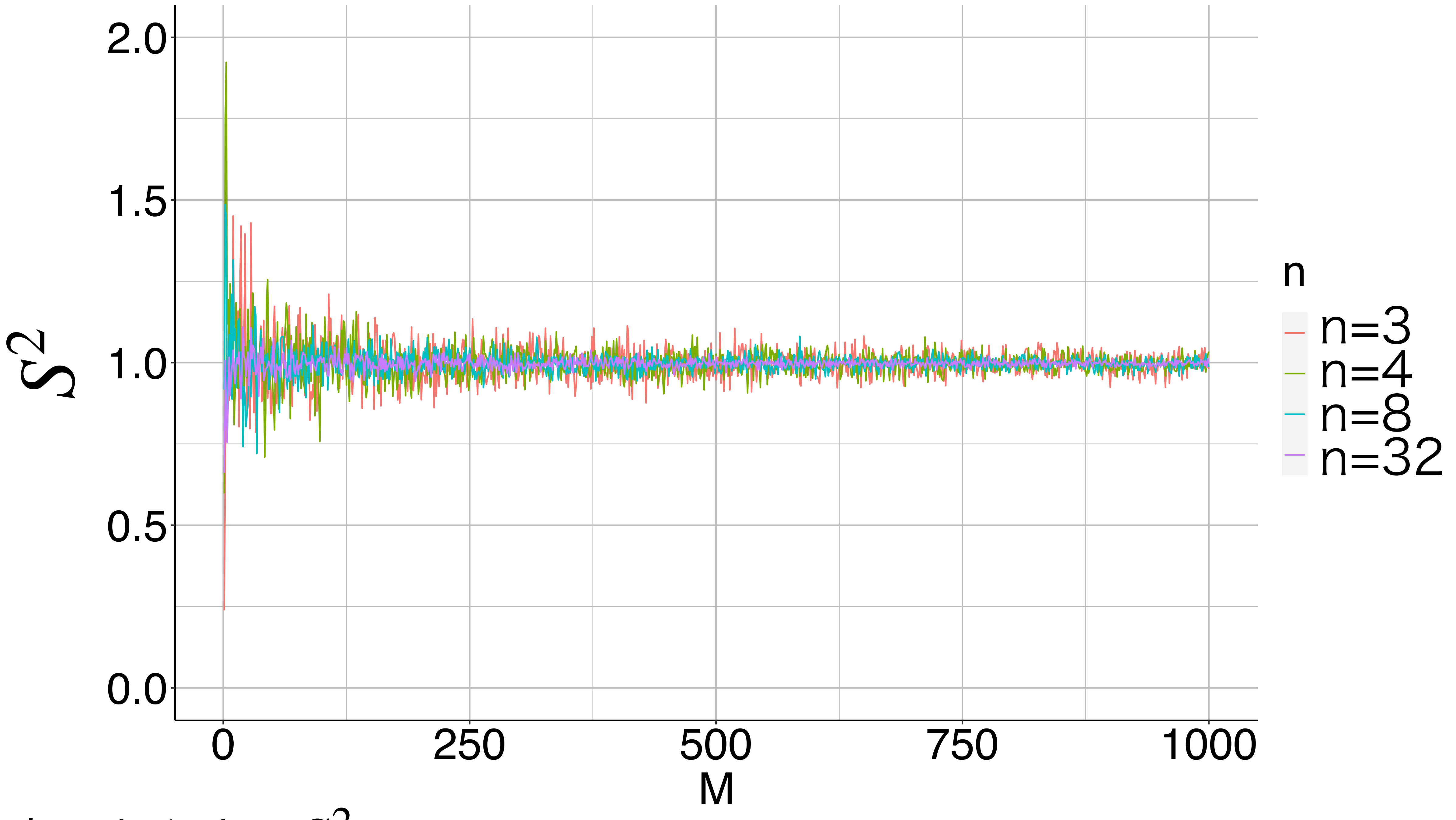
S^2を標本不偏分散という。また母分散はσ^2（シグマ二乗）で表す。

- なぜ$n-1$なのか？

1 自由度による説明

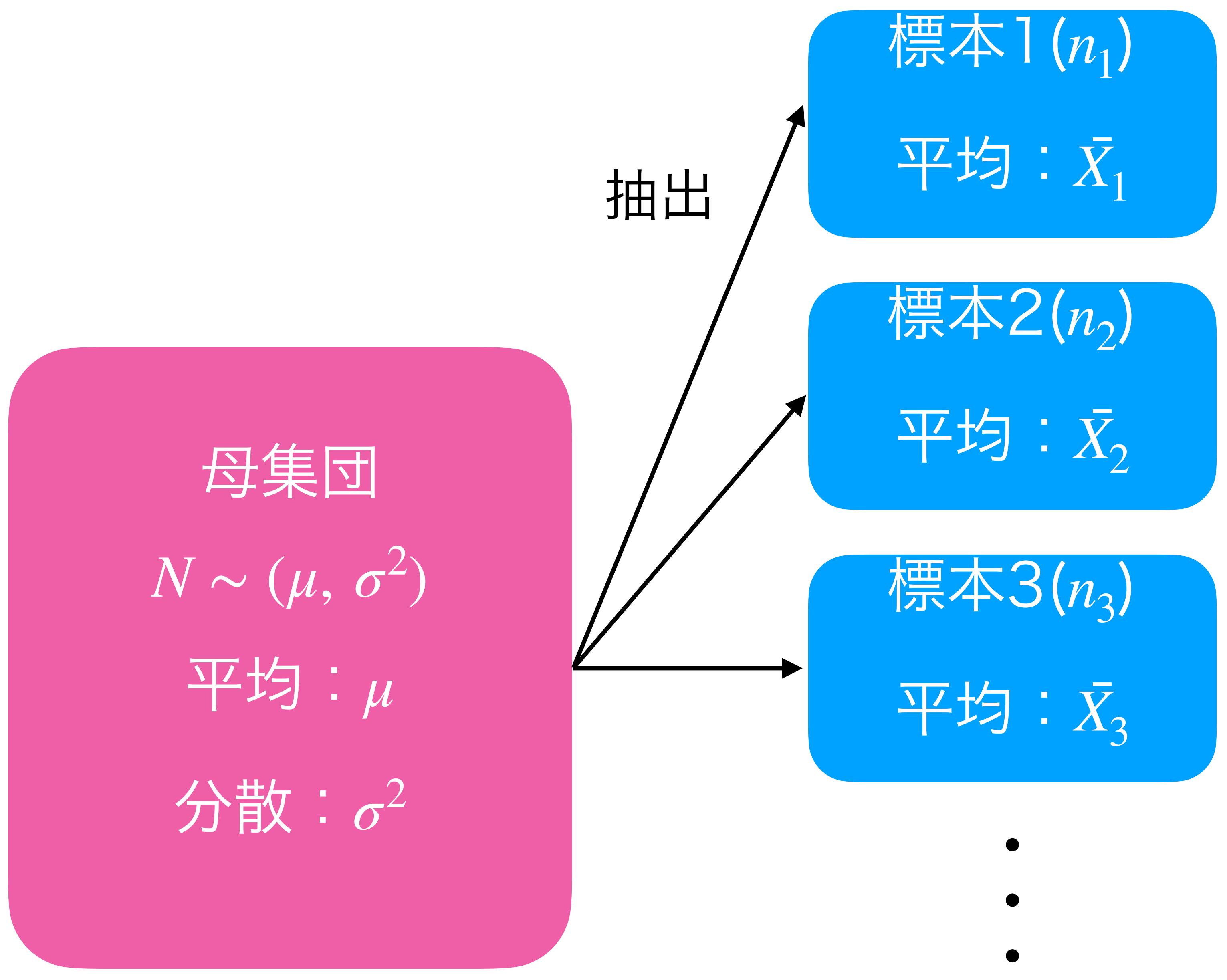
n 個の偏差$(X_i - \bar{X})$の和は 0 であるため，$n-1$ 個の偏差が与えられると残りの 1 つは自動的に決まる[*7]．つまり自由に決められる偏差の数は $n-1$ である．したがって偏差を用いた分散を定めるためには，偏差平方和 SS を n でなく $n-1$ で割るほうが合理的である．自由に決められる変数の数 $n-1$ を**自由度**という．

R言語でのシミュレーション結果



分散が 1 である母集団からn=3, 4, 8, 32の標本をm回抽出したときのS^2↑

平均値の散布度（バラツキ）について



母集団から複数の標本を抽出し、その標本ごとに標本平均を算出する

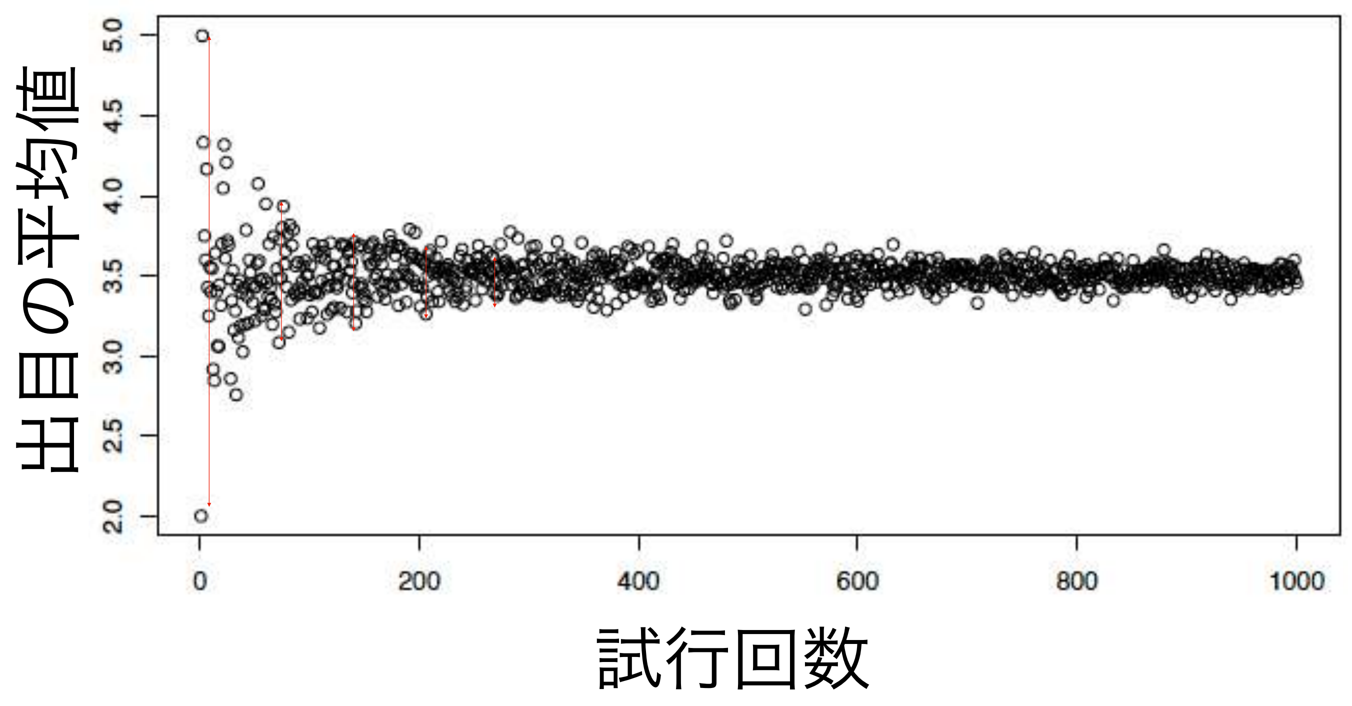
→適切な方法で抽出すれば、常に同じ集団が抽出されるわけではないため標本平均はバラつく

大数の法則：

標本の大きさが大きくなるにつれて、標本平均は母平均μに近づく

→標本の大きさが大きくなると標本平均の分散は小さくなる。

→したがって、標本平均の分散は$\frac{\sigma^2}{n}$、標準偏差は$\frac{\sigma}{\sqrt{n}}$



標本の散布度

- なぜ$n-1$なのか？

母分散(σ^2)は、標本分散(S'^2)に平均の分散($\frac{\sigma^2}{n}$)を加えることで推定できることから、

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 + \frac{\sigma^2}{n}$$

$$\sigma^2(1 - \frac{1}{n}) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

$$\sigma^2(\frac{n-1}{n}) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

$$\sigma^2 = (\frac{n}{n-1})\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

$$\hat{\sigma}^2 = S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

点推定・区間推定・仮説検定

1. 点推定：１点のみでの推定する。情報量は少なく、計算コストも少ない。推定値はどれくらい信頼できる値なのかが不明。

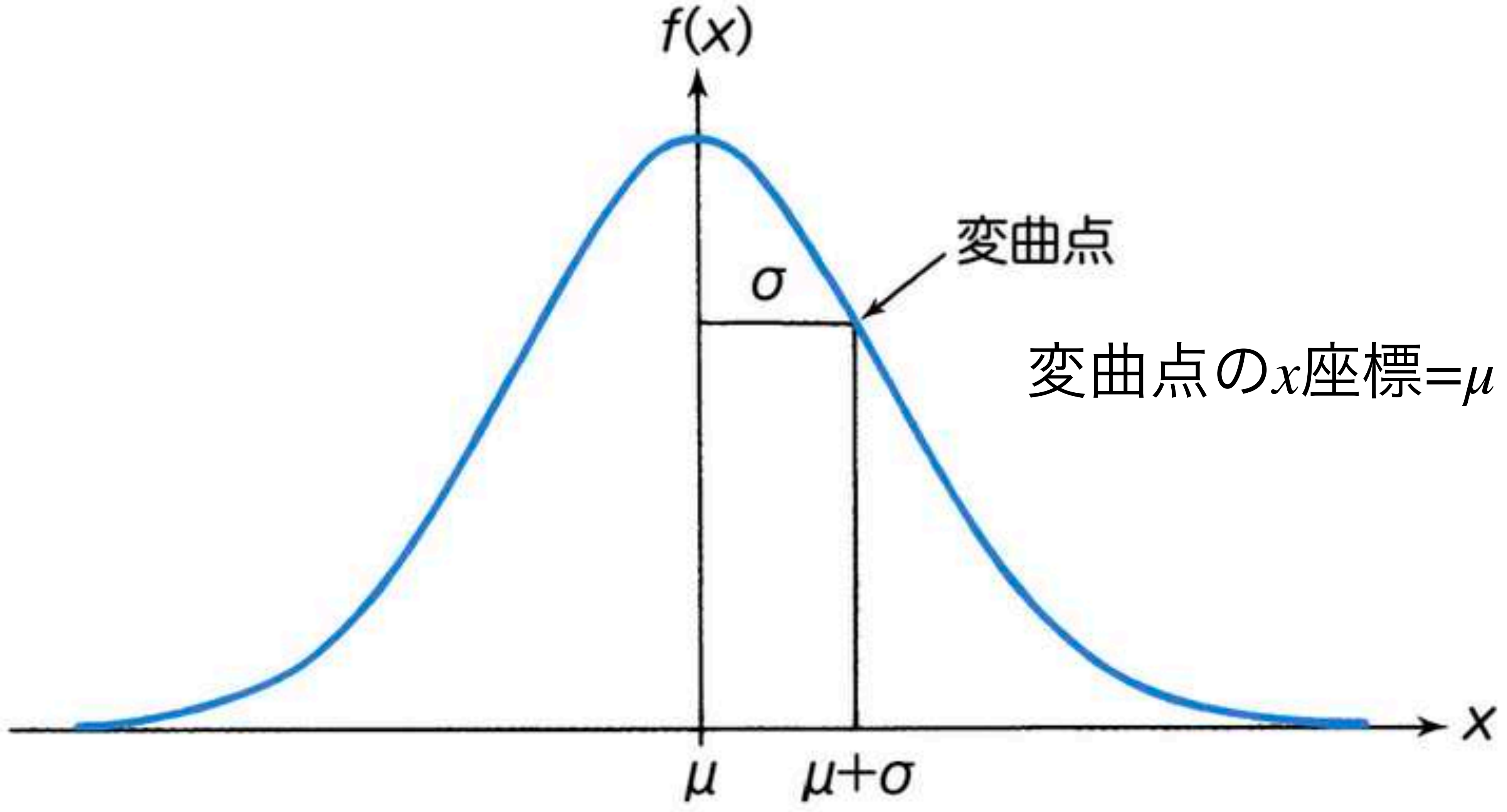
   標本平均、標本不偏分散、標本標準偏差

2. 区間推定：信頼性を持たせた区間（信頼区間）でもって推定する。点推定よりも情報量はやや多いが、計算コストも多い。（標本平均はバラつく、点推定のみではバラツキの情報はわからない）

3. 仮説検定：標本の値（平均や分散など）を使って、母集団の値（平均や分散など）に差があるかどうかを確率的に検出することであり、標本で表された差が本質的な差なのか、偶然によるものなのかを確率的に（対応する分布を使って）判断すること。

正規分布（ガウス分布）

- 自然界で観察される多くの現象は、特定の値の度数が最も多く、その度数から正負の方向に離れるに従って、度数は小さくなる。
- このような分布を正規分布と呼ぶ。
- 正規分布の確率密度関数は、次式で表される。
- $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \ (-\infty < x < \infty)$



変曲点のx座標=$\mu + \sigma$

図 5·6　正規分布

- 平均μ、分散σ^2の正規分布は$N(\mu, \sigma^2)$で表される。
- 実際の研究データでは、対数変換後に正規分布に従う場合が多い。このような分布を対数正規分布と呼ぶ。e.g.遺伝子発現解析など

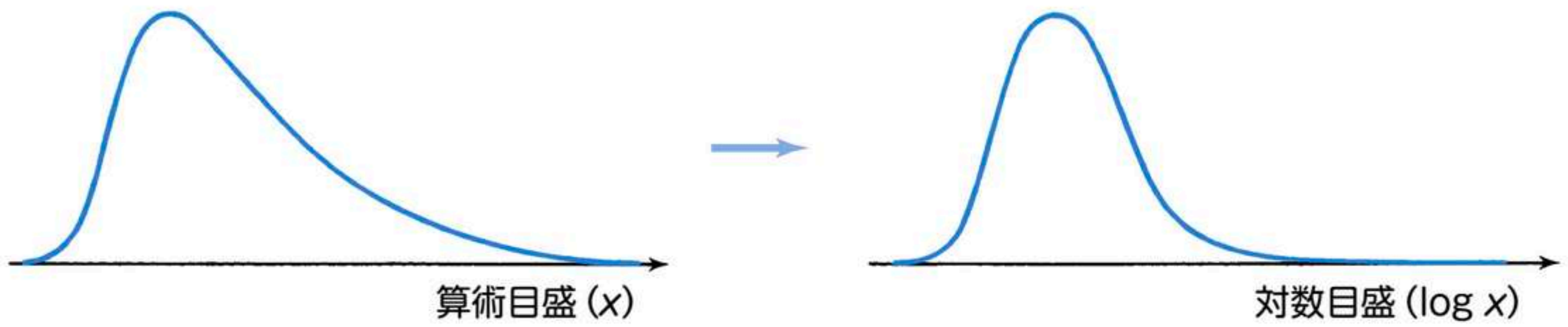


図 5·7　対数による正規分布化

正規分布（ガウス分布）の特徴

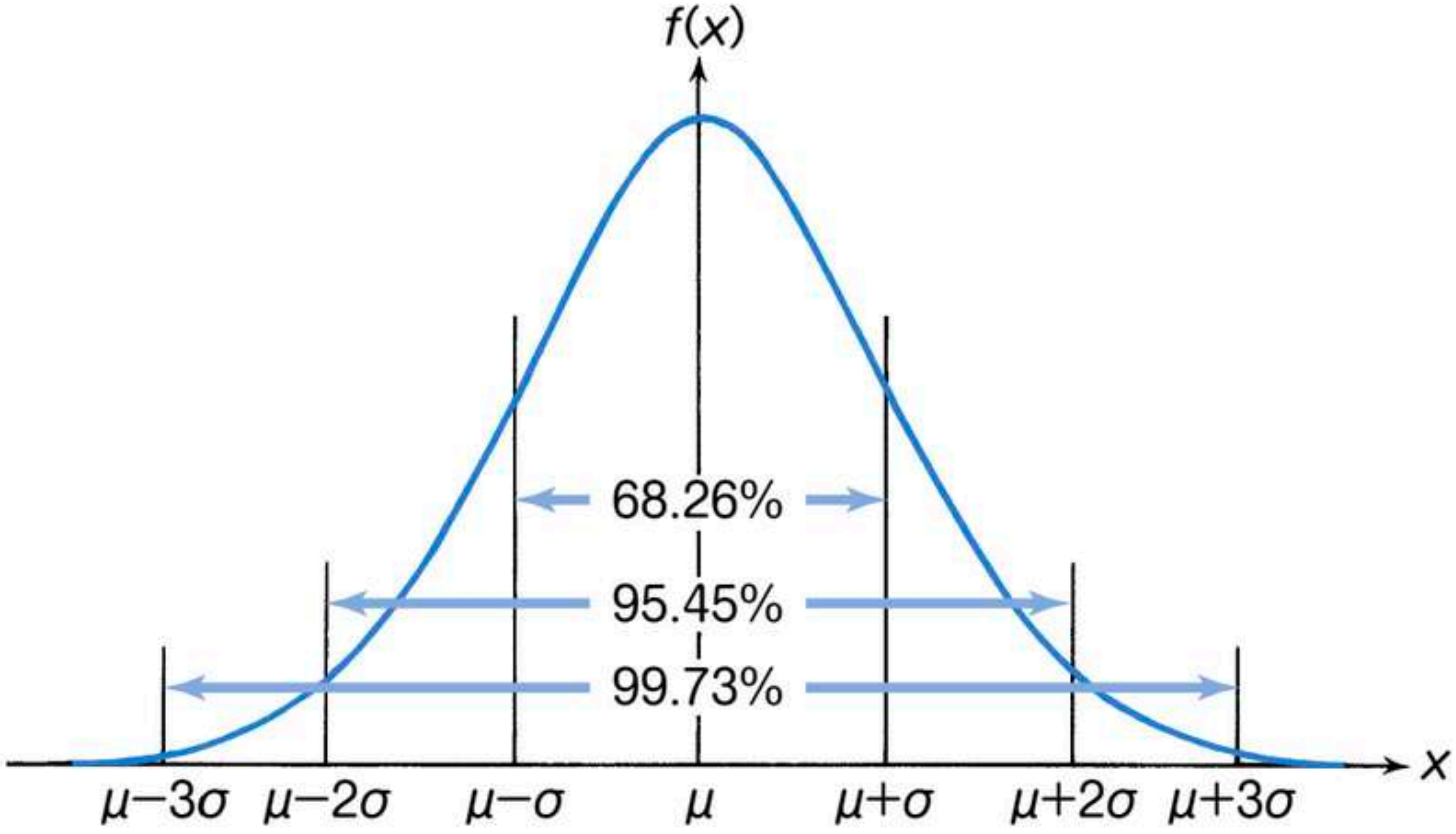
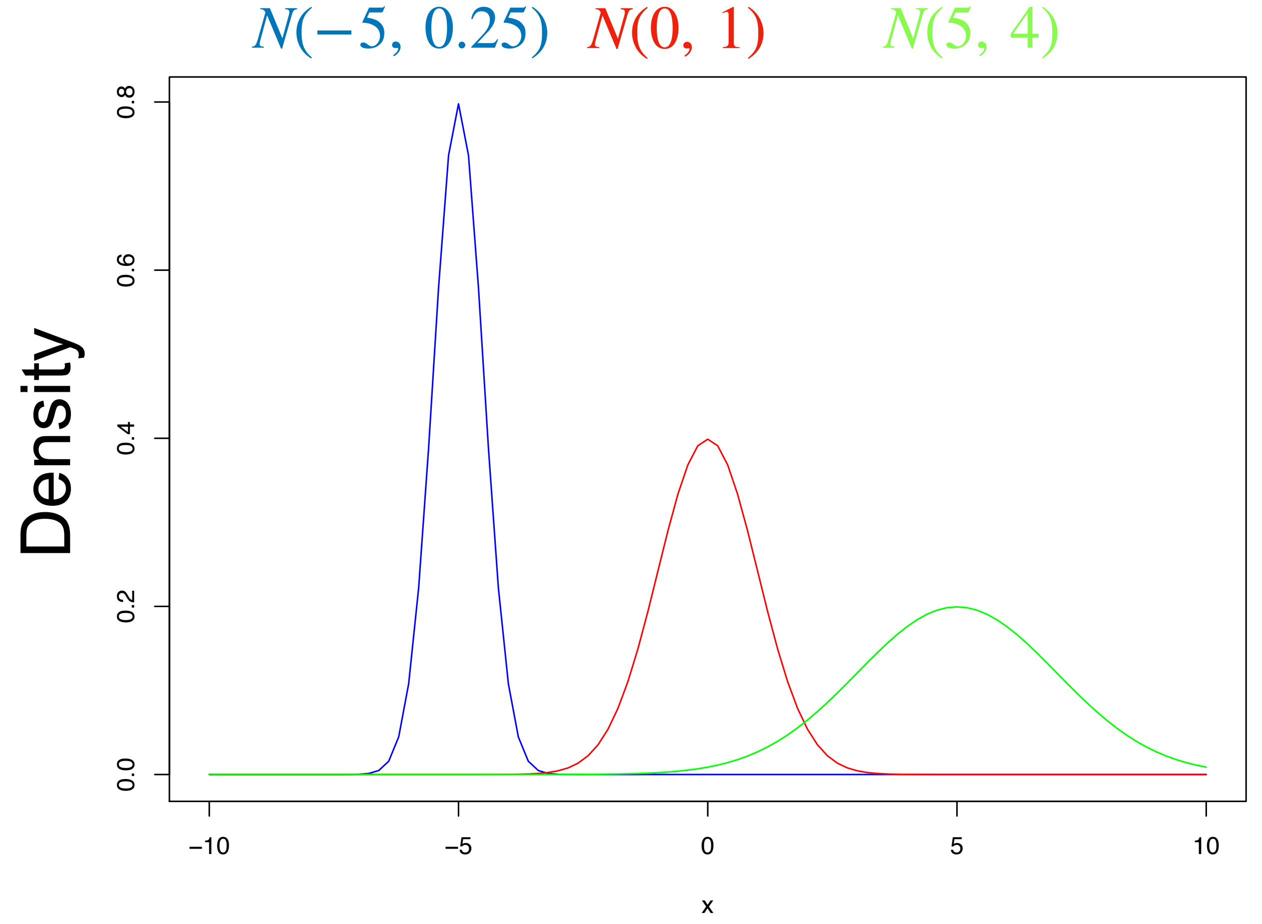


図 5·9　正規分布の特徴

標準化

- 単位や平均値が異なる標本間で比較したい

→任意の正規分布$N(\mu, \sigma^2)$について、平均を０、分散を１にすればよい（つまり、$N(0,1)$にすればよい）

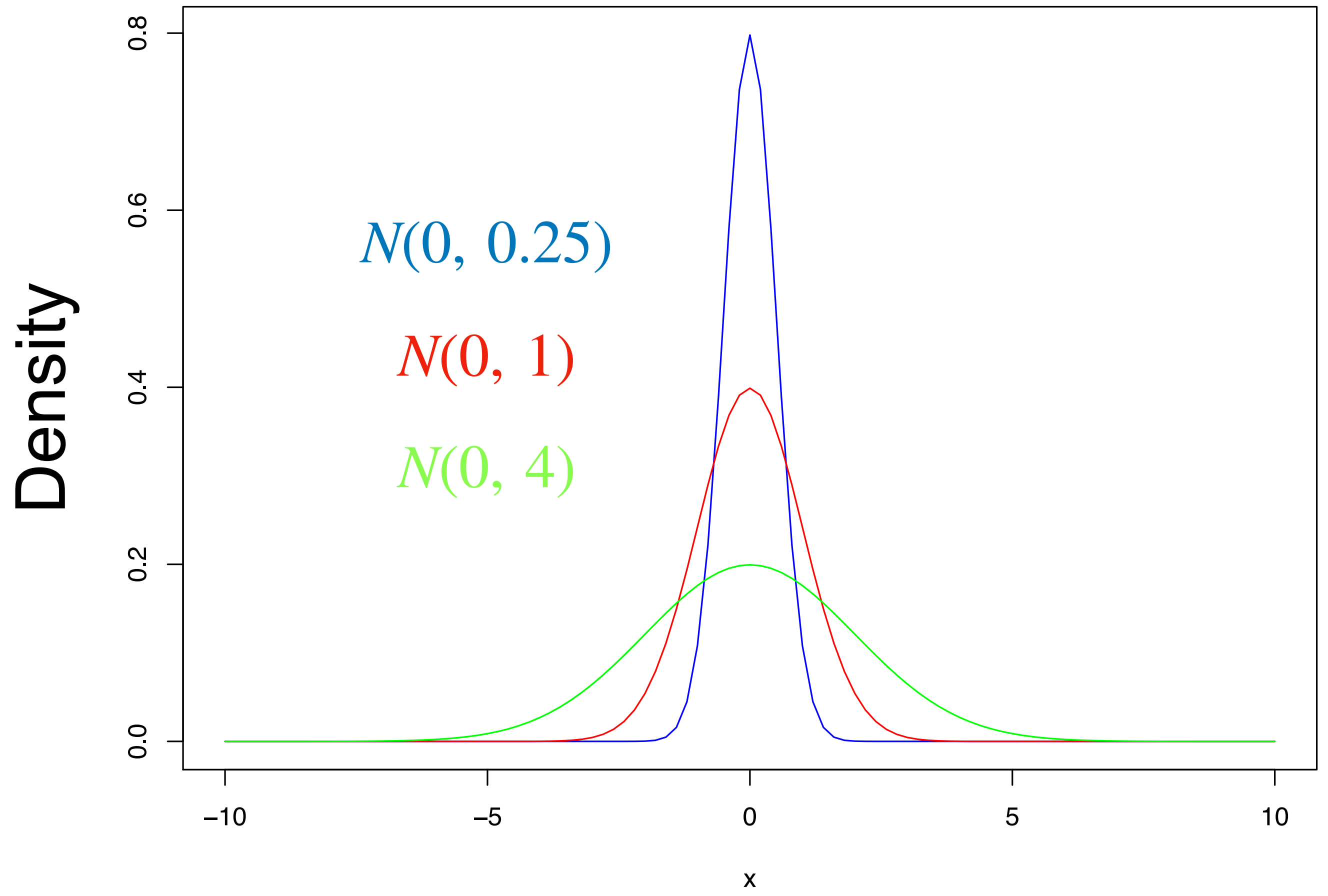


元の分布

- $z = \dfrac{X-\mu}{\sigma}$の式を用いて変換.

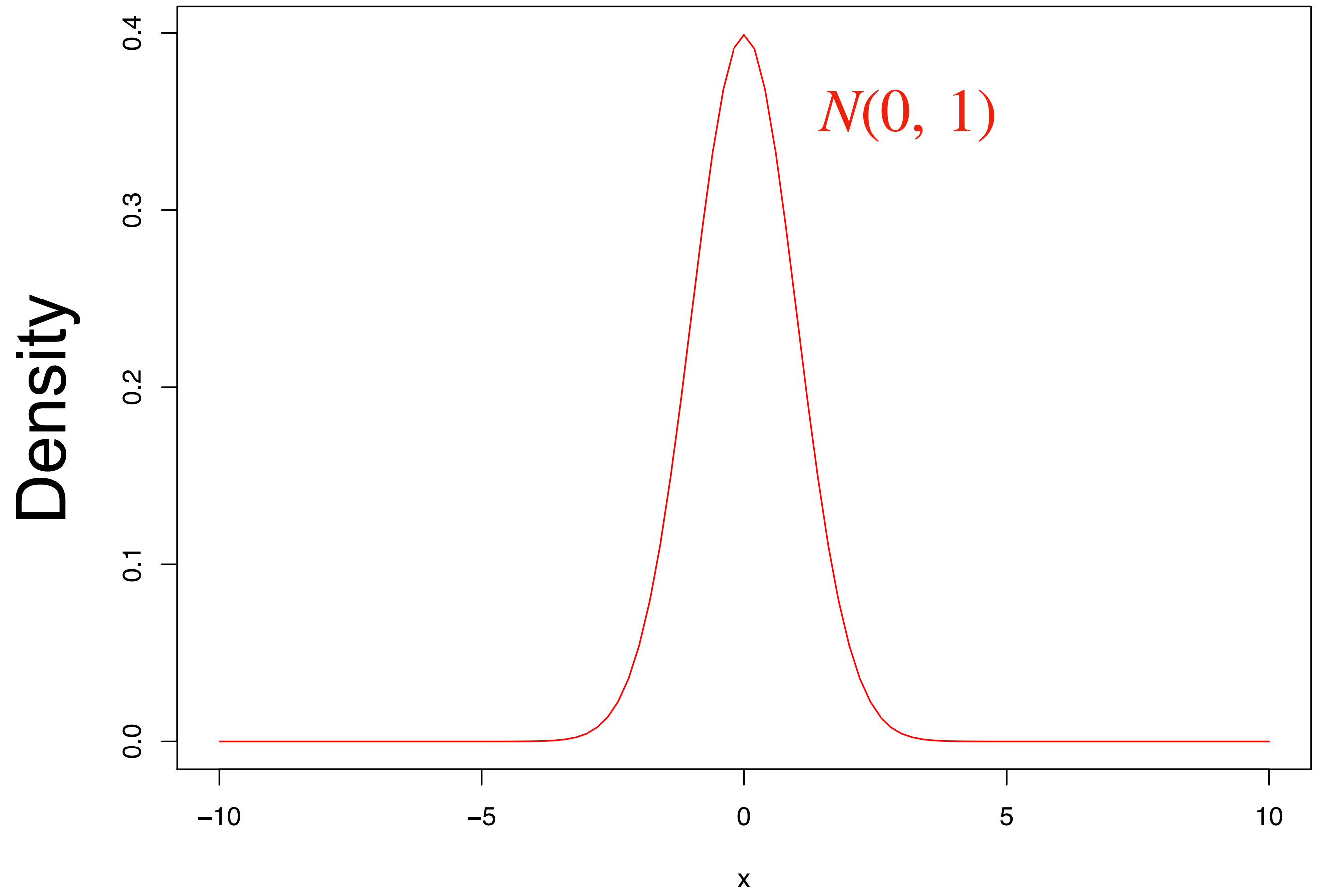
（標準得点, z得点, z-scoreなどと呼ぶ）

$X - \mu$



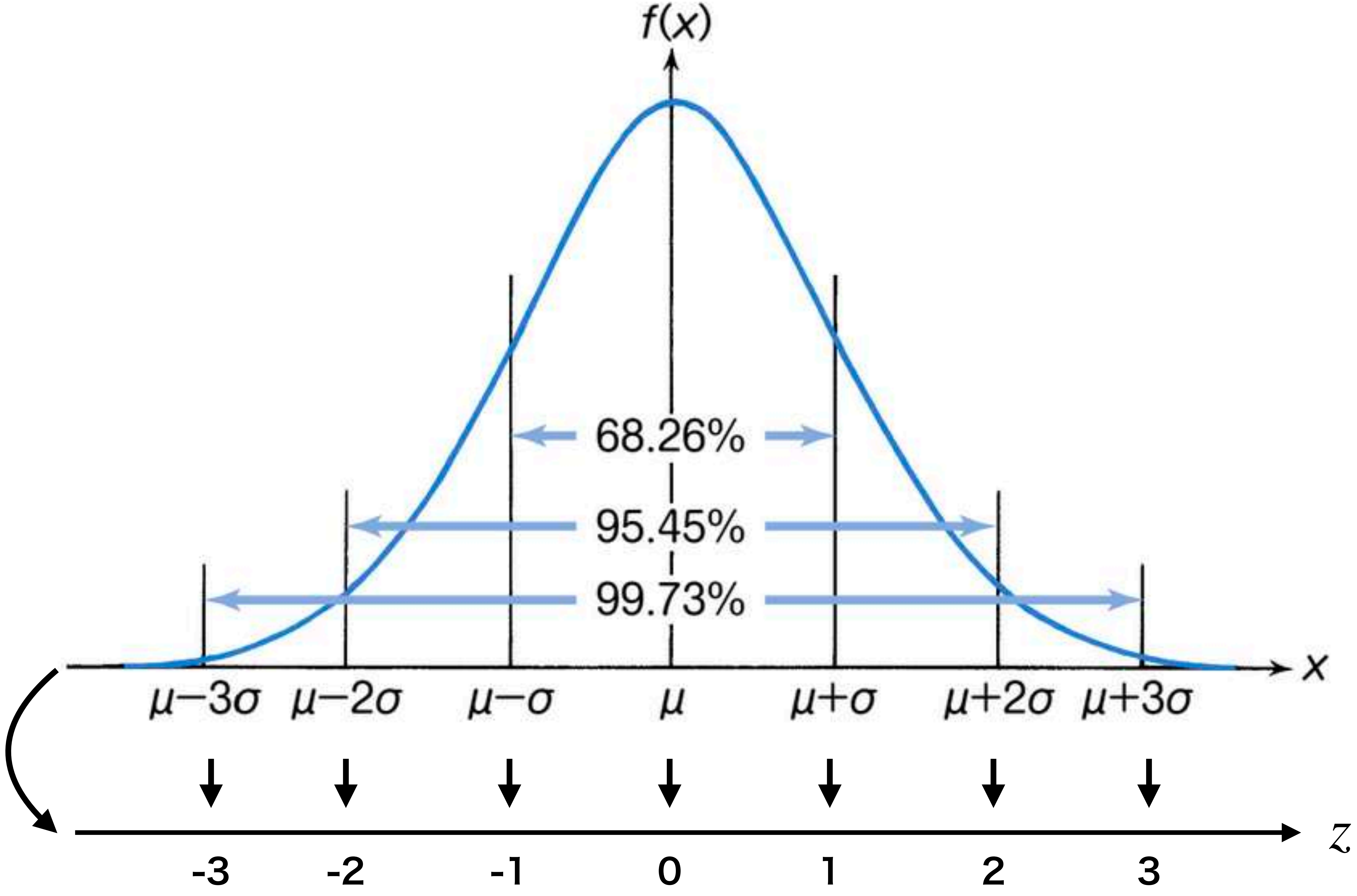
全ての値から母平均μを引いた分布

$\dfrac{X-\mu}{\sigma}$



全ての値から母平均μを引き標準偏差σで除した分布

標準正規分布の特徴

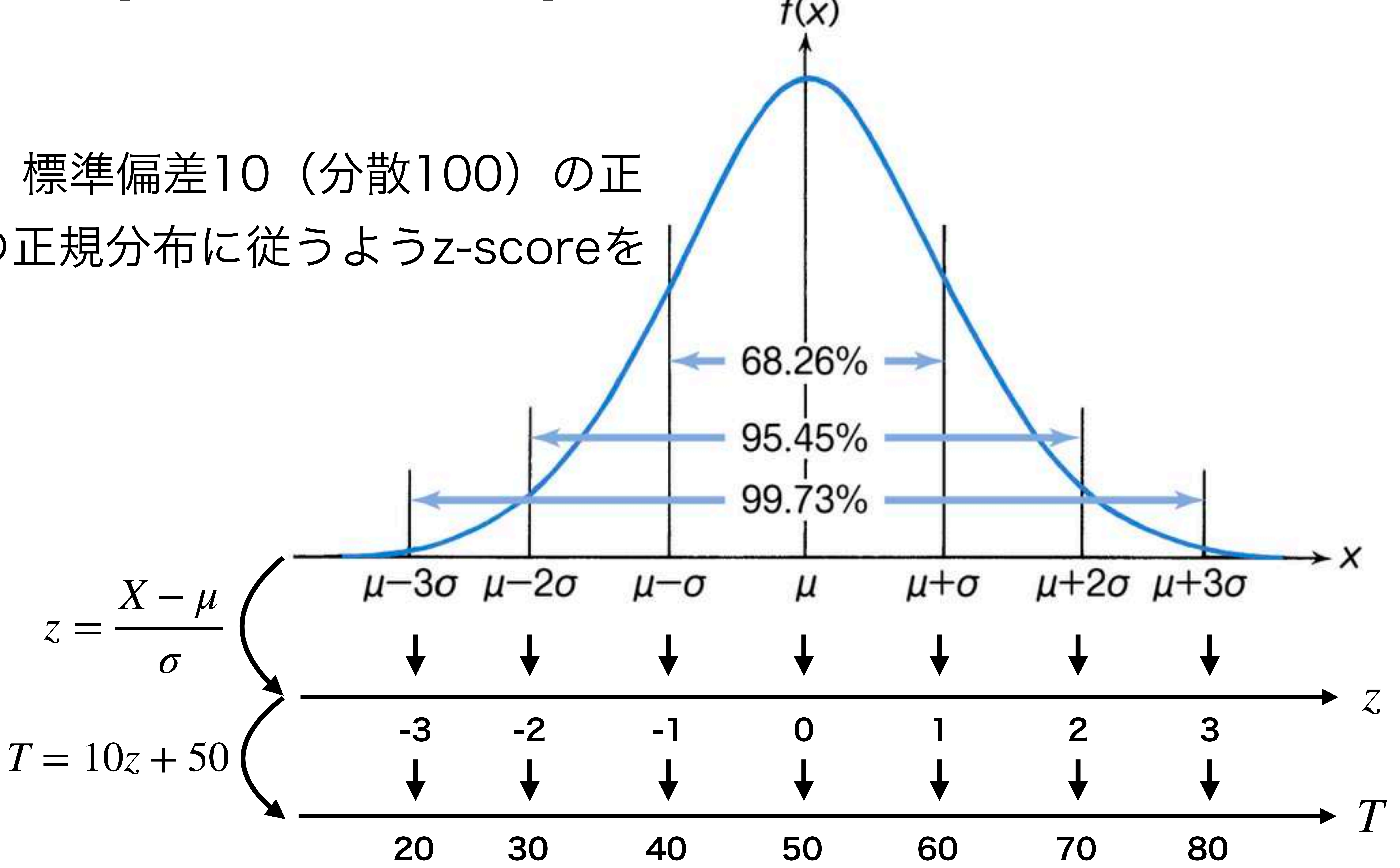


$$z = \frac{X - \mu}{\sigma}$$

偏差値（T-score）

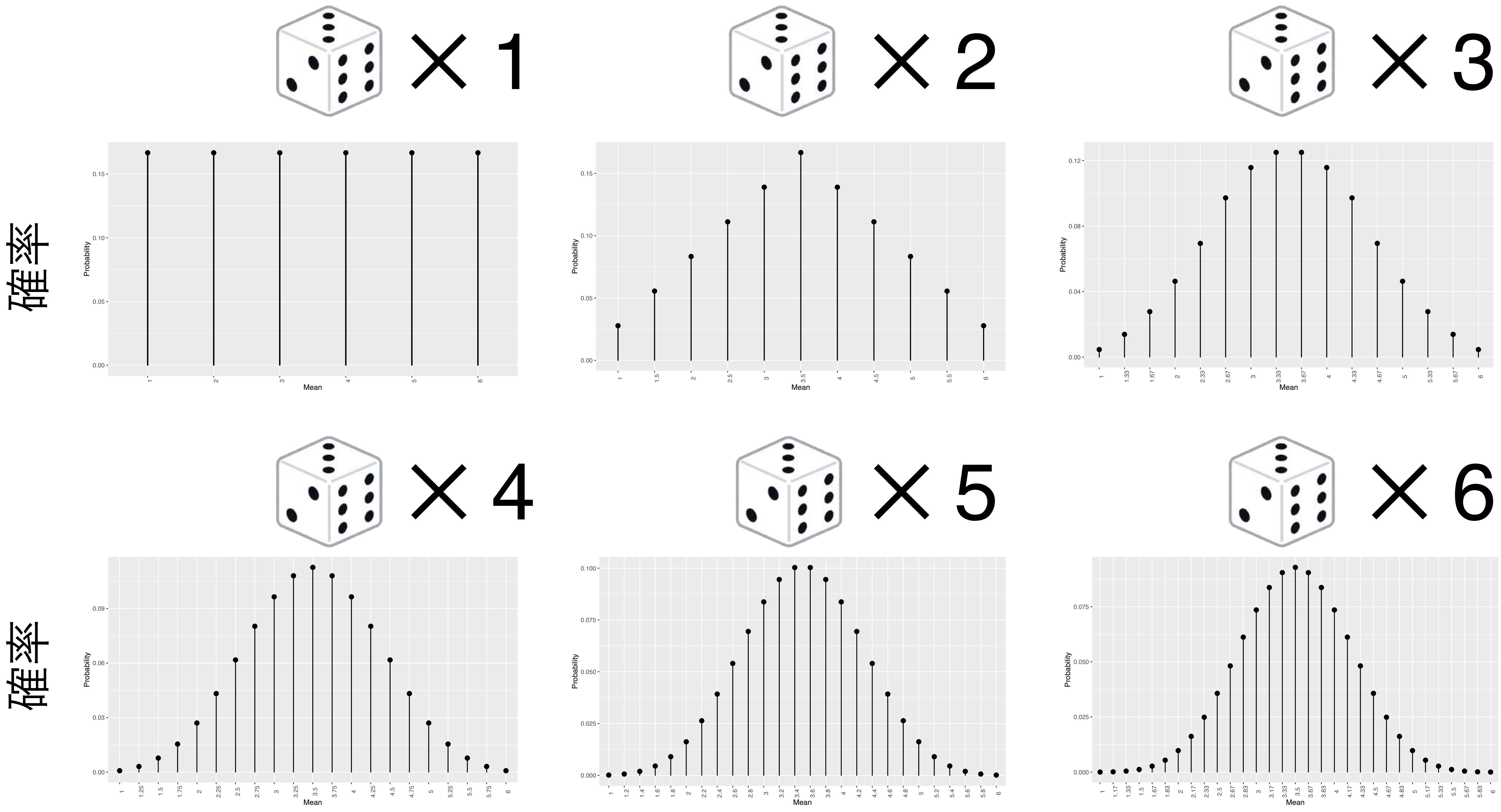
- 偏差値：平均50、標準偏差10（分散100）の正規分布$N(50,100)$の正規分布に従うようz-scoreを変換した値。

- $T = 10z + 50$



中心極限定理

中心極限定理：「標本を抽出する母集団が平均μ、分散σ^2の正規分布$N(\mu, \sigma^2)$に従う場合においても、従わない場合においても、抽出する標本の大きさnが大きくなるにつれて標本平均の分布は「平均μ、分散$\dfrac{\sigma^2}{n}$」の正規分布$N(\mu, \dfrac{\sigma^2}{n})$に近づく」



６面体サイコロ×１個〜６個の出目の平均値

母平均の区間推定（母分散既知）

・手順

1. 標本平均を求める

$$\bar{X} = \frac{X_1 + X_2 + \ldots X_n}{n}, \bar{X} \sim N(\mu, \frac{\sigma^2}{n})$$

2. 変数の標準化を行う。

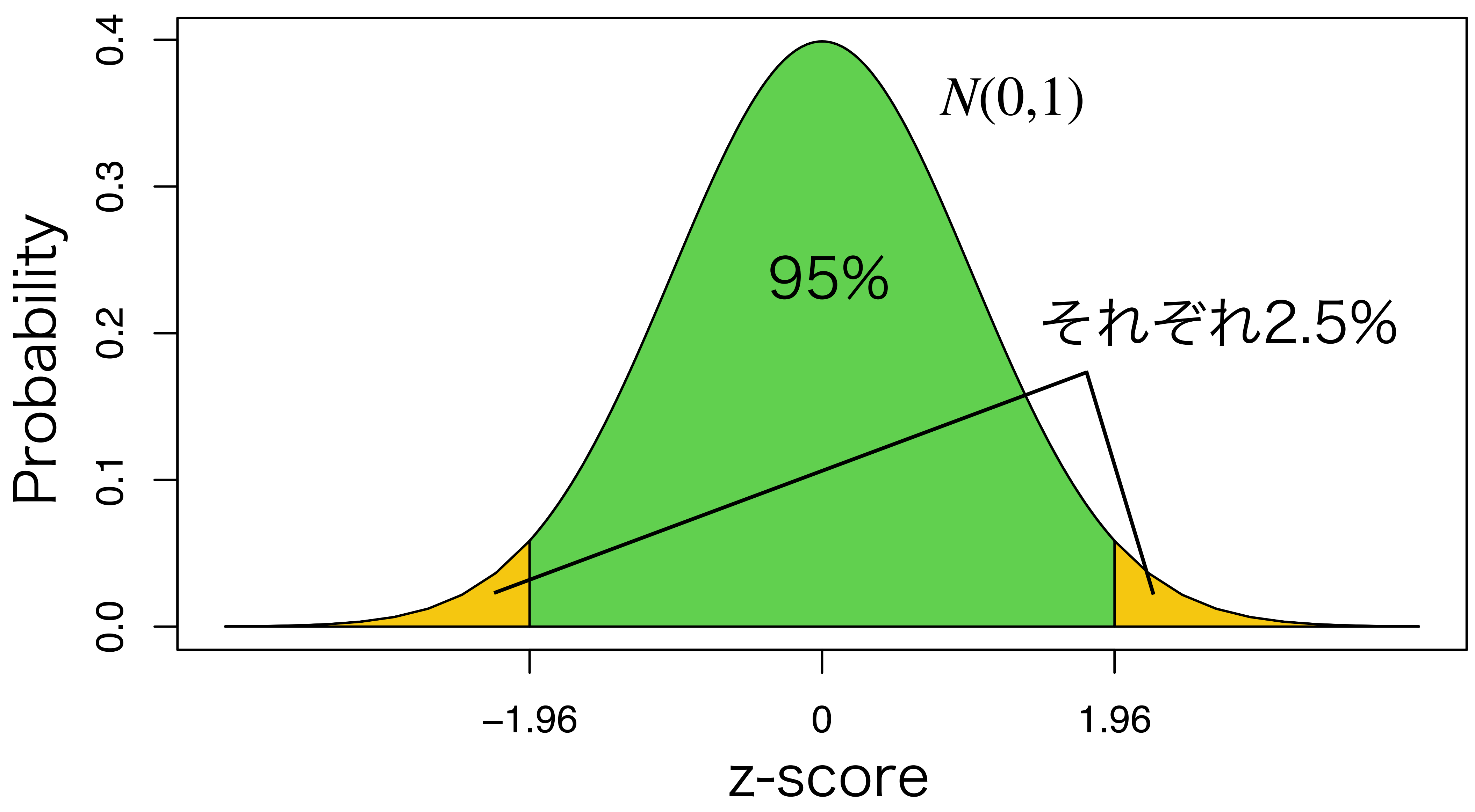
$$z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}, z \sim N(0,1)$$

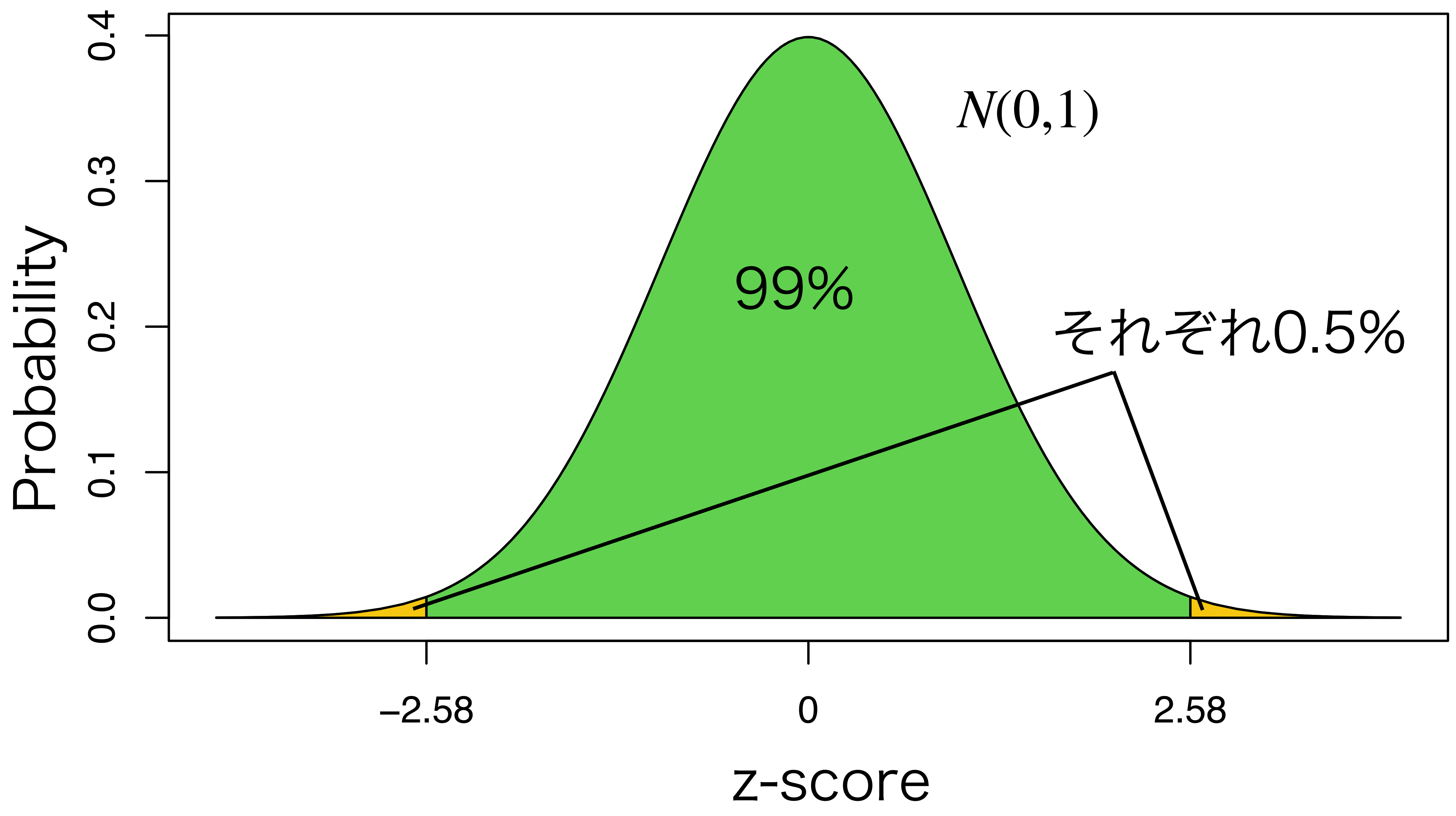
3. 標準正規分布の95%点（信頼係数/信頼度）からzの信頼区間を算出

$$-1.96 < \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} < 1.96$$

4. 3．をμについての区間に変換

$$\bar{X} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}$$





t分布

- 連続型確率変数の分布の一つ
- t分布：「正規分布に従う母集団の分散が未知」で「nが小さい場合」に平均の区間を推定する問題に利用される。
- 母分散が未知である場合には統計量tに関する確率密度分布(t分布)を使って区間推定を行う。統計量tは、次式で表される。

$$t = \frac{\bar{X} - \mu}{\sqrt{\frac{S^2}{n}}}$$

- このとき、tは自由度$n-1$のt分布に従う、という。
- t分布の確率密度関数は、次式で表される。

$$f(t) = \frac{1}{\sqrt{df}\, B\left(\frac{1}{2}, \frac{df}{2}\right)\left(1 + \frac{t^2}{df}\right)^{(df+1)/2}} \qquad (-\infty < t < \infty)$$

t分布

- tは正規分布に従わないのか？

$$t = \frac{\bar{X} - \mu}{\sqrt{\frac{S^2}{n}}}$$

►

標本を抽出し標本平均$\bar{X}$を計算することを何度も繰り返した場合、大数の法則より標本の大きさnが小さいと標本平均$\bar{X}$の標本間でのバラツキは大きくなる
→この場合、tは正規分布よりも裾野が広い（つまりバラツキの大きい）分布に従い、正規分布には従わない
→この正規分布に従わない分布をt分布と呼ぶ。
→t分布はnの大きさによって分布の形状が変化し、nが十分大きいときはt分布は正規分布に近づく。

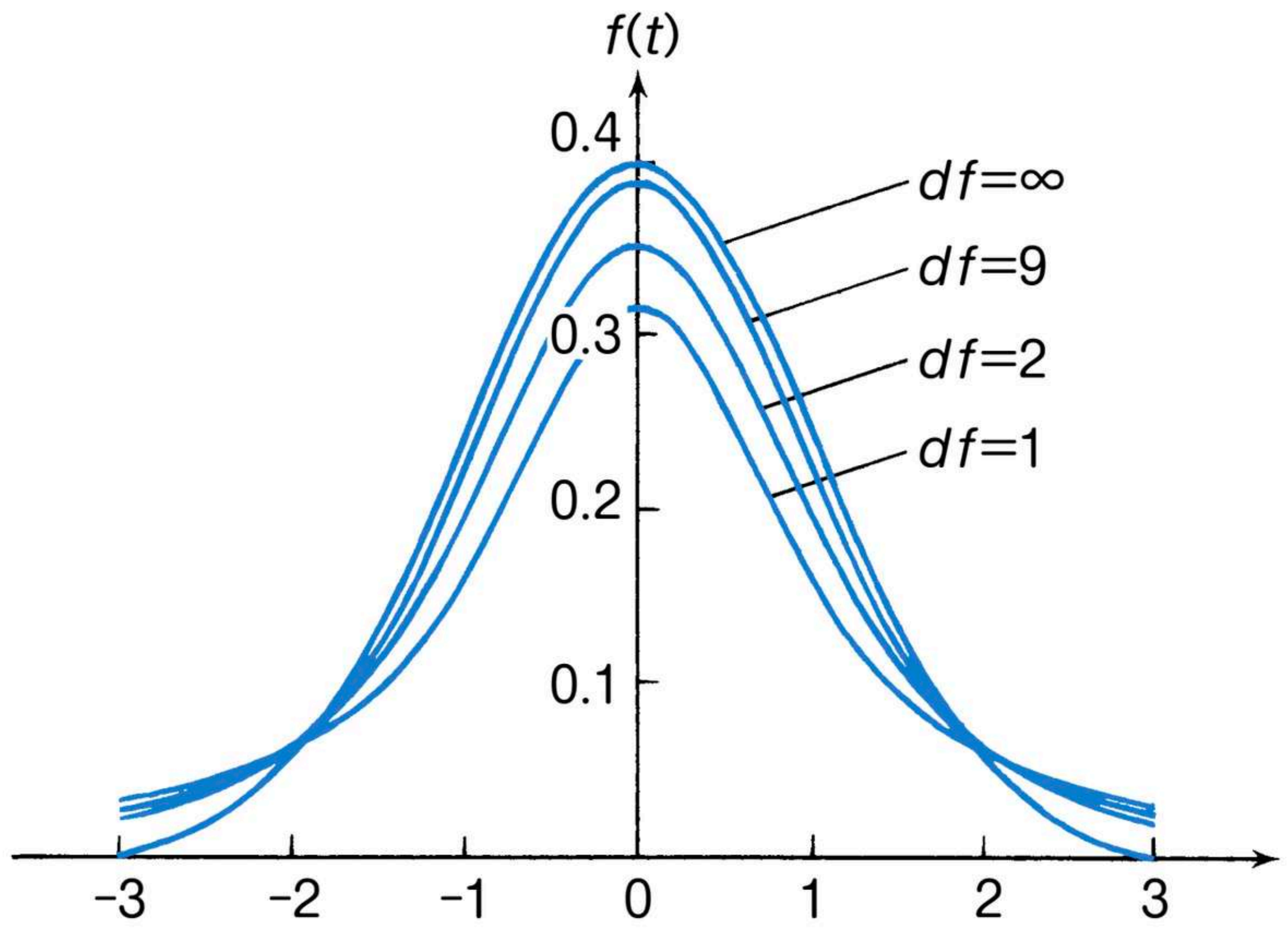


図 5･13　種々の自由度の t 分布

母平均の区間推定（母分散未知）

・手順

1. 標本平均を求める

$$\bar{X} = \frac{X_1 + X_2 + \ldots X_n}{n}$$

2. 標本不偏分散S^2を算出。

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

3. 信頼係数$(1-\alpha)100\%$のとき、自由度$n-1$のt分布の両側α%点を$t_{n-1}(\frac{\alpha}{2})$

　とすると、t分布の信頼区間は以下で表される。αを有意水準という。

$$-t_{n-1}(\frac{\alpha}{2}) < \frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}} < t_{n-1}(\frac{\alpha}{2})$$

4. μについて解くと、

$$\bar{X} - t_{n-1}(\frac{\alpha}{2})\frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{n-1}(\frac{\alpha}{2})\frac{S}{\sqrt{n}}$$

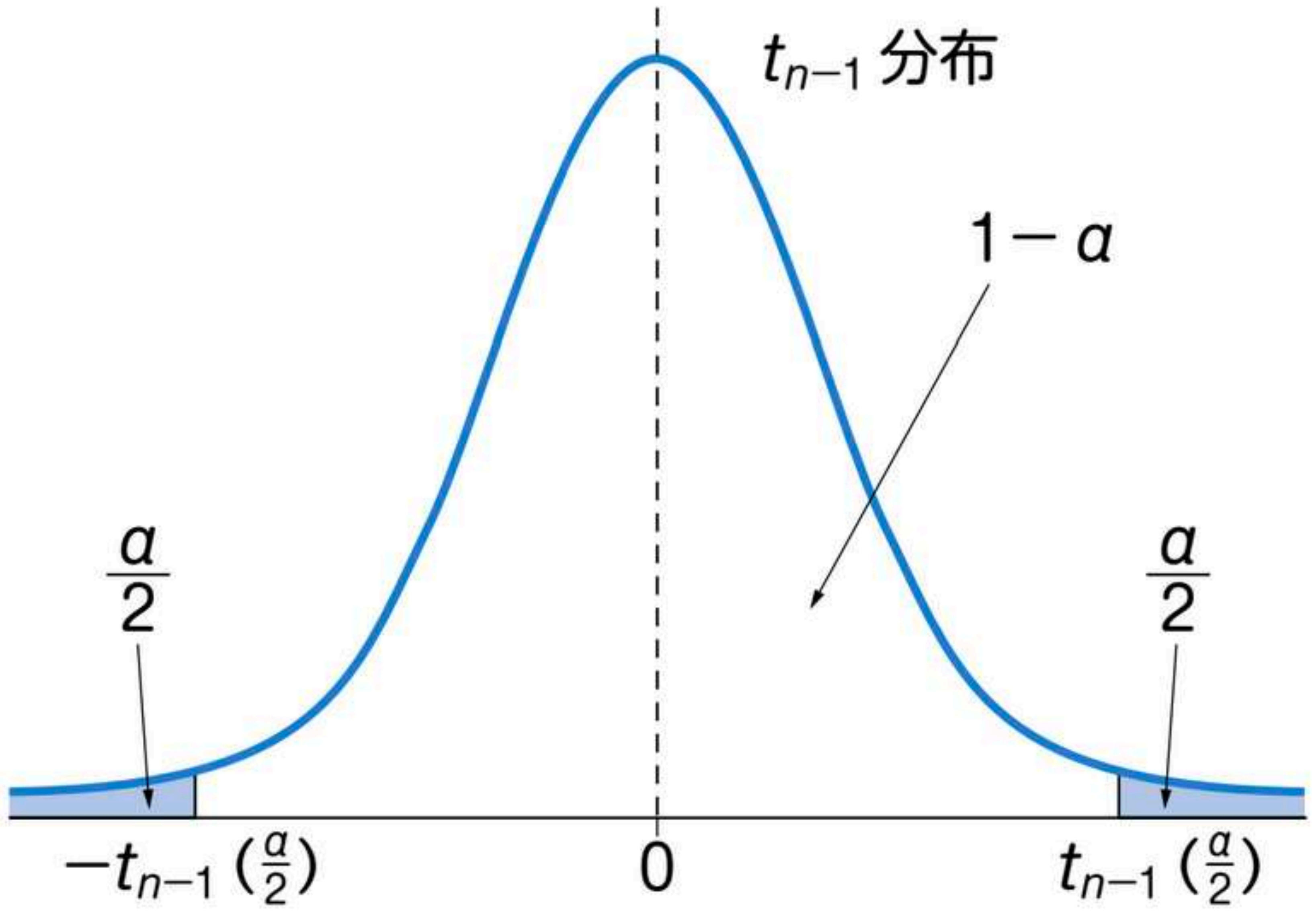
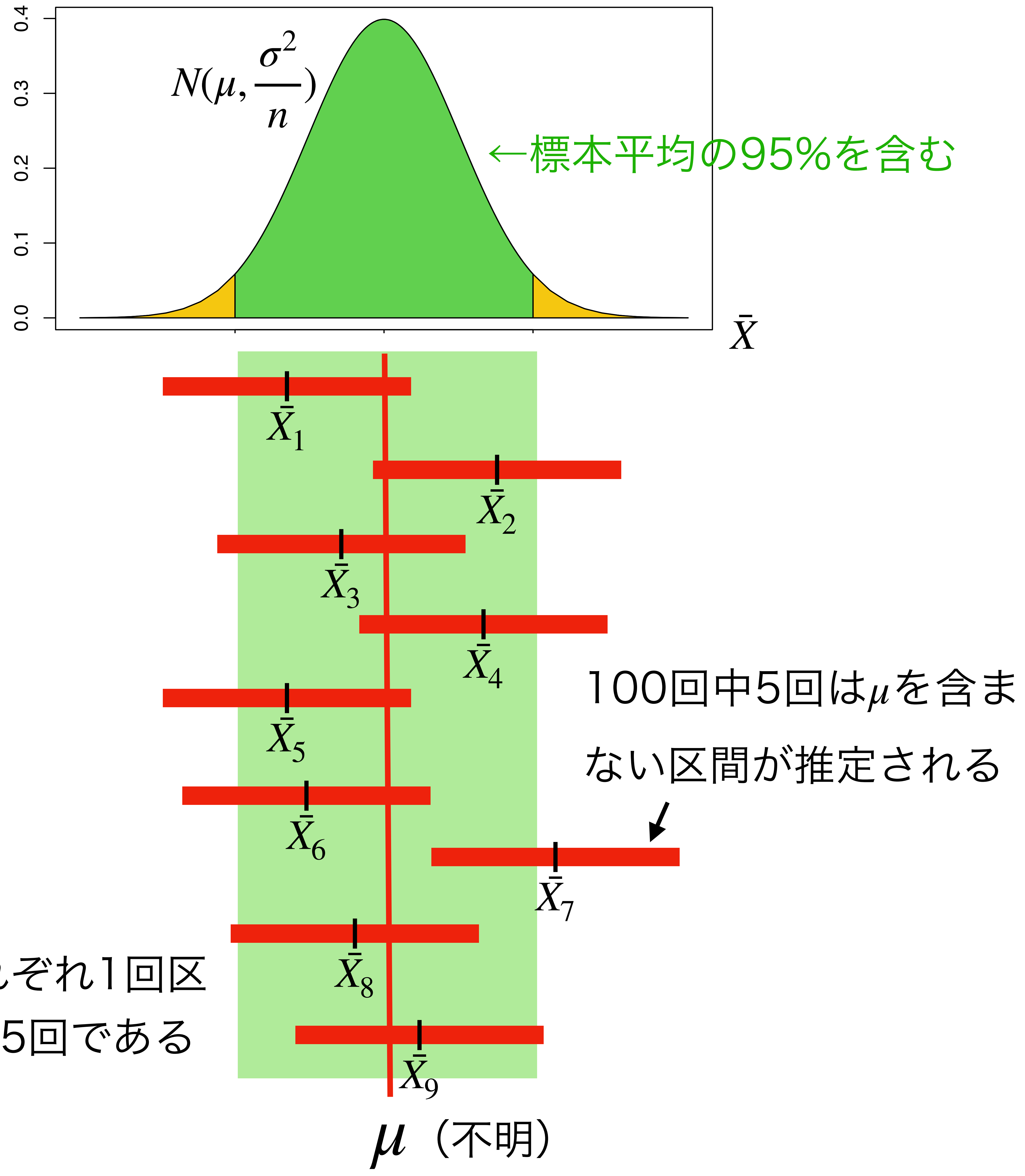


図 9·1　t_{n-1} 分布における上側 $\left(100\times\frac{\alpha}{2}\right)$ %点と下側 $\left(100\times\frac{\alpha}{2}\right)$ %点

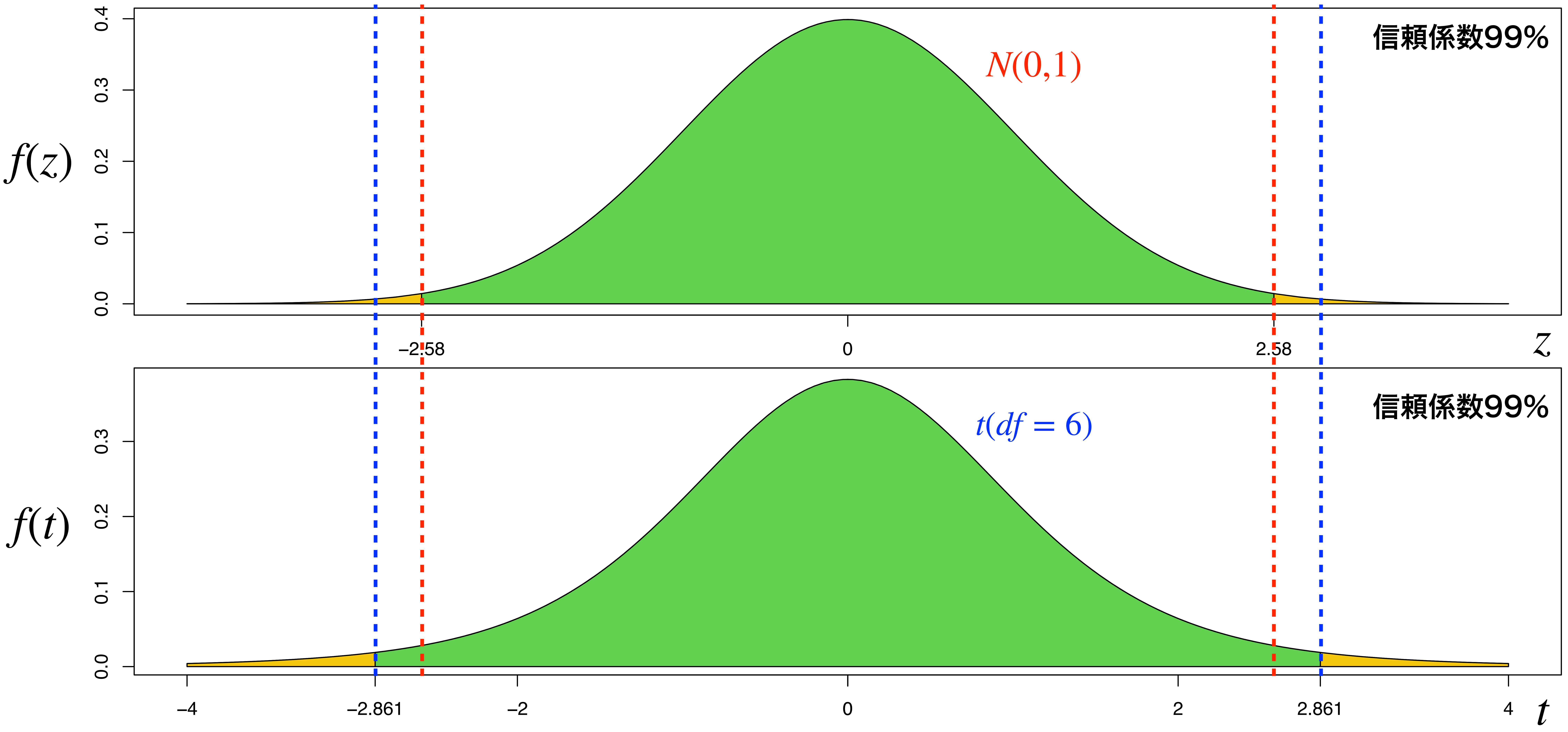
母平均の区間推定

- 推定した区間の解釈：95%の確率でa~bの区間は母平均を含む。
- 注意⚠：母平均が様々な値を取りうる（ばらつく）という前提で、母平均が推定された区間a,bに含まれる確率が95%であるという意味ではない。→母平均は定数であり誤差は生じない。誤差が生じるのは標本から求められるa,bである。



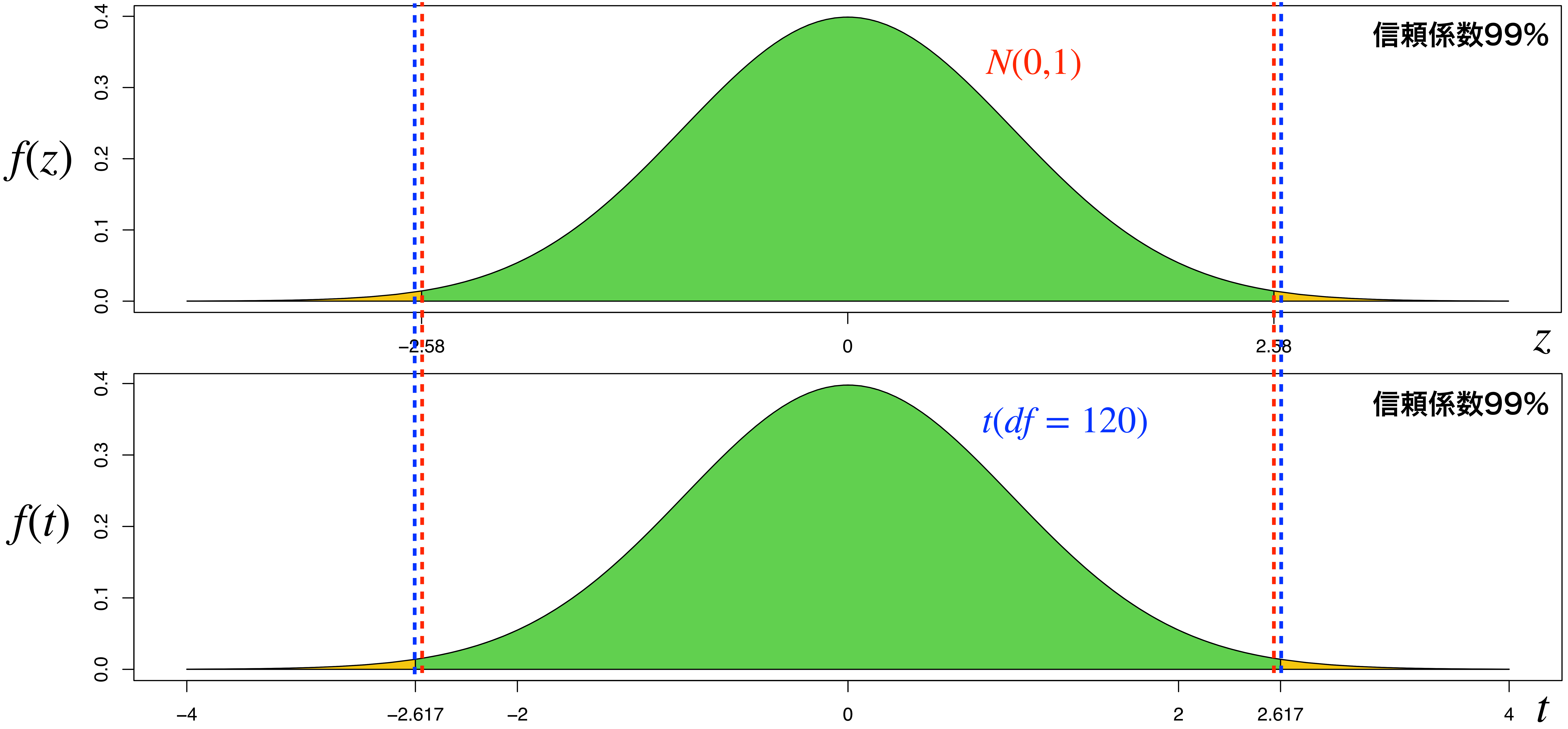
信頼係数が95%の場合、100個の標本についてそれぞれ1回区間推定すると（計100回）、μを含む区間の回数は95回である

母分散既知・未知での分布使い分けの理由



- 標準正規分布を用いた方が推定される区間幅が小さいため、推定精度が高い

母分散既知・未知での分布使い分けの理由



- ただし、nを大きくしていくとt分布を用いて推定した区間は標準正規分布を用いて推定した区間に近づく

点推定・区間推定・仮説検定

1. 点推定：１点のみでの推定する。情報量は少なく、計算コストも少ない。推定値はどれくらい信頼できる値なのかが不明。

   標本平均、標本不偏分散、標本標準偏差

2. 区間推定：信頼性を持たせた区間（信頼区間）でもって推定する。点推定よりも情報量はやや多いが、計算コストも多い。（標本平均はバラつく、点推定のみではバラツキの情報はわからない）

3. 仮説検定：標本の値（平均や分散など）を使って、母集団の値（平均や分散など）に差があるかどうかを確率的に検出することであり、標本で表された差が本質的な差なのか、偶然によるものなのかを確率的に（対応する分布を使って）判断すること。

仮説検定

- 標本の値（平均や分散など）を使って、母集団の値（平均や分散など）に差があるかどうかを確率的に検出することであり、標本で表された差が本質的な差なのか、偶然によるものなのかを確率的に（対応する分布を使って）判断すること。

表 6·1　血中コレステロール値(mg/dL)

	1	2	3	4	5	6	7	8	9	10	平均
A 群	212	147	156	176	191	252	187	357	163	200	204.1
B 群	200	153	147	175	182	246	194	355	161	193	200.6

- 標本平均を単純に比較すると差はあるようだが、本質的に差があるかどうか（母平均に差があるかどうか）はわからない。（外れ値による影響を受けている可能性など）

仮説検定

仮説検定の手順：①帰無仮説、対立仮説を立てる。

- 帰無仮説（Null hypothesis）
  - 棄却されることを前提として立てた仮説

    例）母集団Aと母集団Bの平均に差がない

- 対立仮説（Alternative hypothesis）
  - 帰無仮説が棄却された場合に採択する仮説

    例）母集団Aと母集団Bの平均に差がある

仮説検定

仮説検定の手順：①帰無仮説、対立仮説を立てる。

- 帰無仮説はゼロ仮説とも呼ばれることからH_0、対立仮説はH_1と表す。

e.g. ある２つの母集団の平均について仮説検定する場合、

- 帰無仮説は$H_0 : \mu_1 = \mu_2$

- 対立仮説

１．両側検定の場合：$H_1 : \mu_1 \neq \mu_2$

２．片側検定の場合：$H_1 : \mu_1 > \mu_2$、または$H_1 : \mu_1 < \mu_2$

仮説検定

仮説検定の手順：②有意水準αを定める。

- 一般的には、最小でも$\alpha = 0.05$で設定する。その他、$\alpha = 0.01$, $\alpha = 0.001$とする場合もある。

仮説検定の手順：③検定統計量に標本のデータを代入する。

- 検定統計量（下図ではT）の式に標本の値（平均、標準偏差など）を代入し、分布上の検定統計量の位置からP値(帰無仮説の下で実際に標本から計算された検定統計量以上（または以下）の検定統計量が観測される確率)を算出する。

Pの面積　　　　　　　　　　　　αの面積

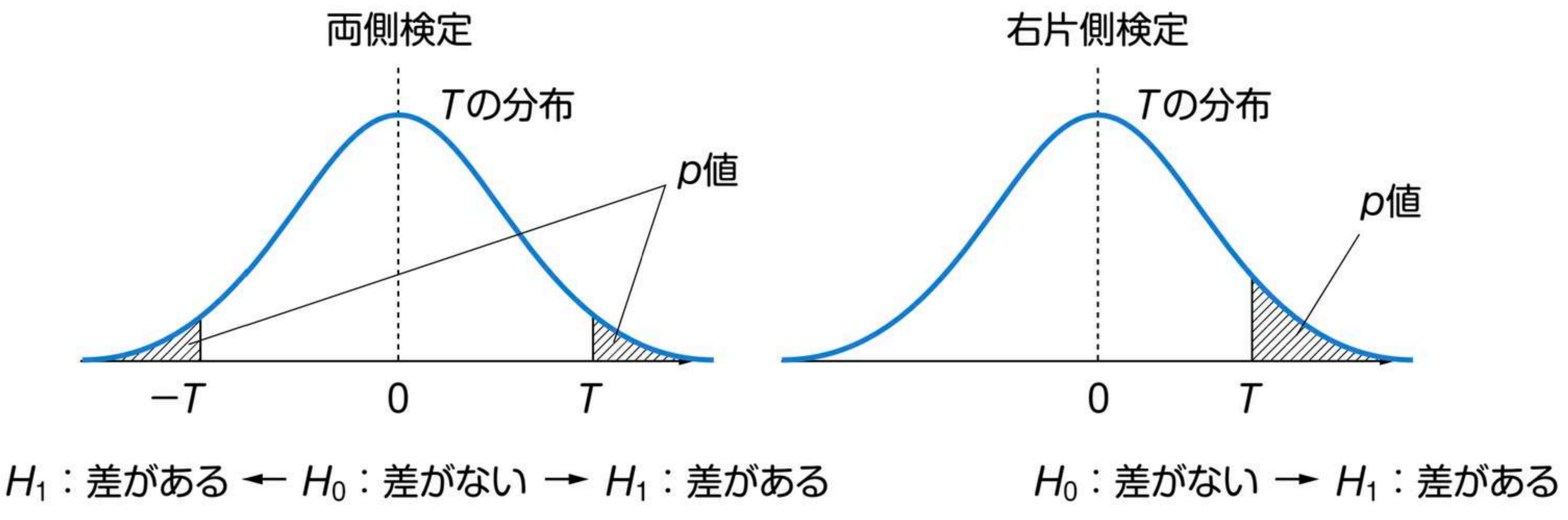


図 6·3　T と p 値の関係

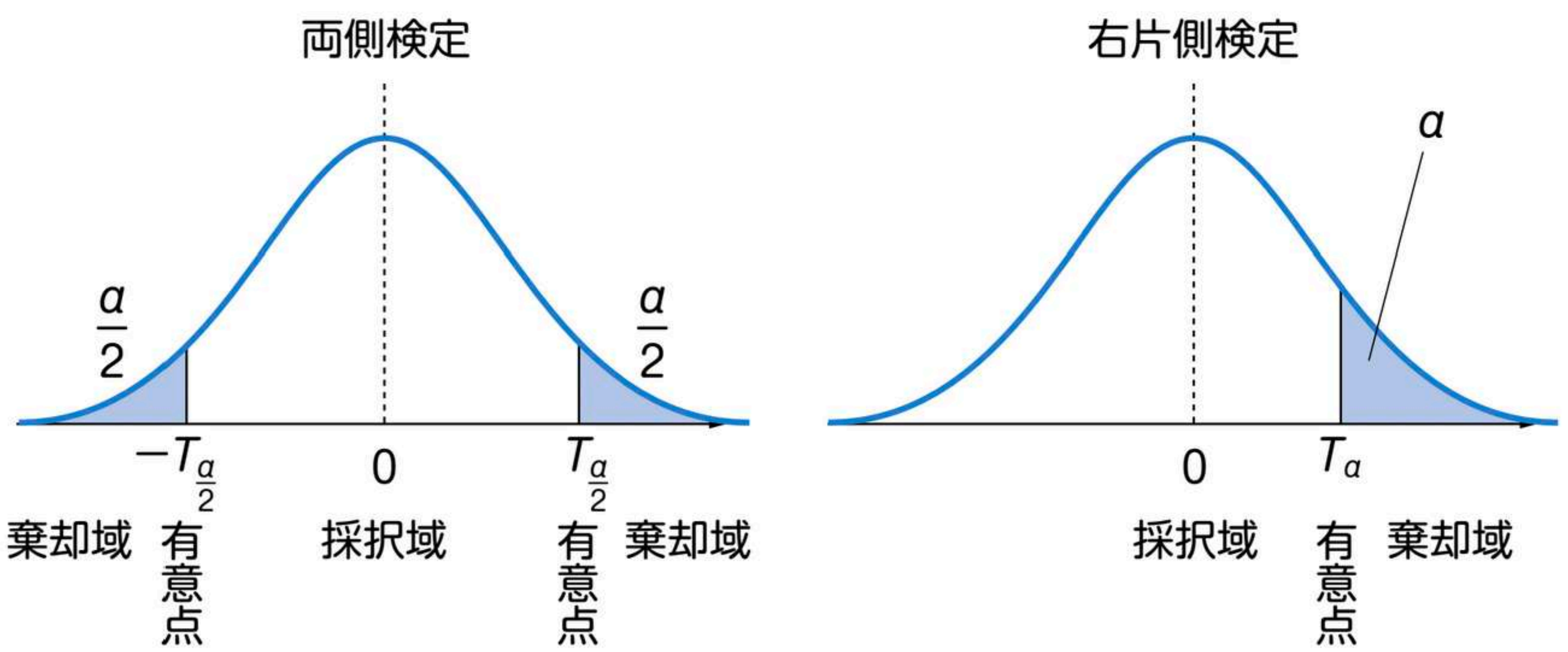


図 6·4　棄却域と採択域

仮説検定

仮説検定の手順：④Pとαを比較して、帰無仮説が棄却されるのか採択されるのかを判断する

- 計算したP値と設定したα見て棄却域・採択域を判断する。

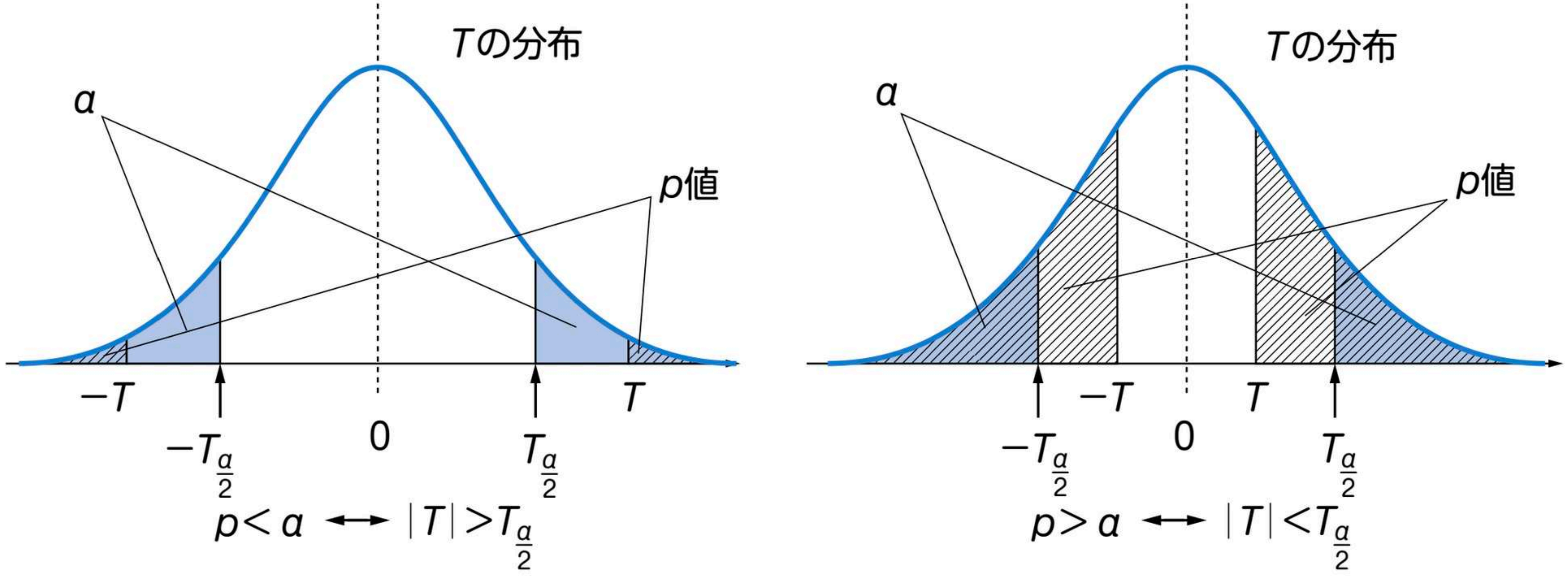


図 6·5　α, $T_{\frac{\alpha}{2}}$, T, p 値の関係(両側検定)

Pの面積とαの面積を比較してどちらが大きいかで、帰無仮説が棄却されるかどうかが決まる。

仮説検定の手順：⑤結果の解釈：

帰無仮説が棄却されれば→「A群とB群に差はある」

帰無仮説が採択されれば→「A群とB群に差があるとは言えない」

背理法と仮説検定

表 6·2　背理法と仮説検定

	背理法	仮説検定
示したいこと	「A である」	一般に対立仮説であることが多い
仮説	「A でない」 ↓ 推論	帰無仮説 ↓ 標本を用いた推論
結果	**矛盾**	**帰無仮説のもとでは標本から得られた結果が起こる確率が p である（0 に近い）**
客観的判断	「A でない」は誤りである	はじめに設定した確率 α（有意水準）と比較する[*1] $p<\alpha \Rightarrow p$ は非常に小さいと考える （帰無仮説は正しくない） $p \geqq \alpha \Rightarrow p$ は非常に小さいとはいえないと考える （帰無仮説は正しくないとはいえない）
結論	「A である」が正しい	（$p<\alpha$ のとき）対立仮説が正しい （$p \geqq \alpha$ のとき）帰無仮説が正しい

２群の比較

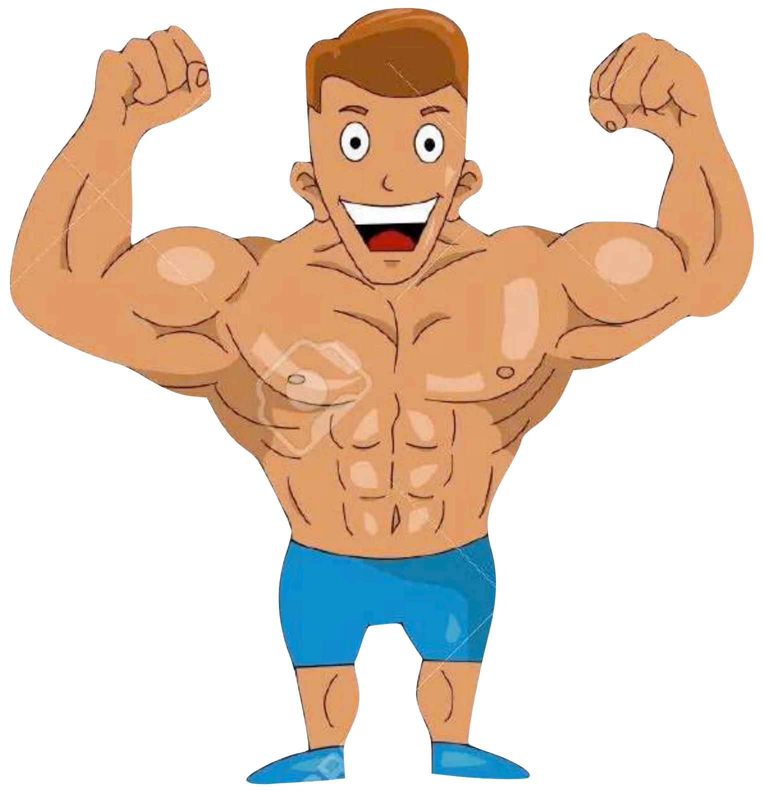
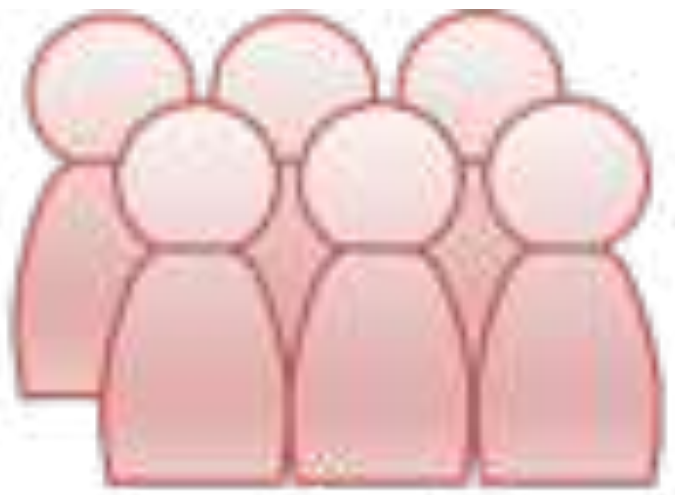
本日の内容

- **対応のない２群の差の検定**
  - 母集団が正規分布に従う → **パラメトリック検定**
    - 母分散が等しい → Student’s t 検定
    - 等しくない → Welch’s t 検定
  - 母集団が正規分布に従わない → **ノンパラメトリック検定**
    - Wilcoxon順位和検定
    - Mann-WhitneyのU検定

- **対応のある２群の差の検定**
  - 母集団が正規分布に従う → **パラメトリック検定** t 検定
  - 母集団が正規分布に従わない → **ノンパラメトリック検定** Wilcoxon符号付き順位検定

対応がある／ないとは

例えば、

対応がない		対応がある	
トレーニング法１	トレーニング法２	筋トレ前	筋トレ後
			
健常者	患者	薬物投与前	薬物投与後
			

等分散の確認（F検定）

２群のt検定では、等分散か等分散でないかで選択すべき検定方法が異なる。

そのため、事前に等分散かどうか調べる必要があり、これをF検定で行う。

手順

① 帰無仮説H_0、対立仮説H_1の設定。$H_0 : \sigma_1^2 = \sigma_2^2,\ H_1 : \sigma_1^2 \neq \sigma_2^2$

② 有意水準αを定める

③ 帰無仮説$H_0 : \sigma_1^2 = \sigma_2^2$のもとで以下の検定統計量$T$を計算。２つの集団の分散が等しいかどうかを検定したいため、分散の比（分散比）を考える。

$$F = \frac{S_1^2}{S_2^2}$$

検定統計量Fは、自由度対$(n_1 - 1, n_2 - 1)$のF分布に従う。

④ P値とαの比較（検定統計量が棄却域なのか採択域なのかを確認）

⑤ 結果の解釈

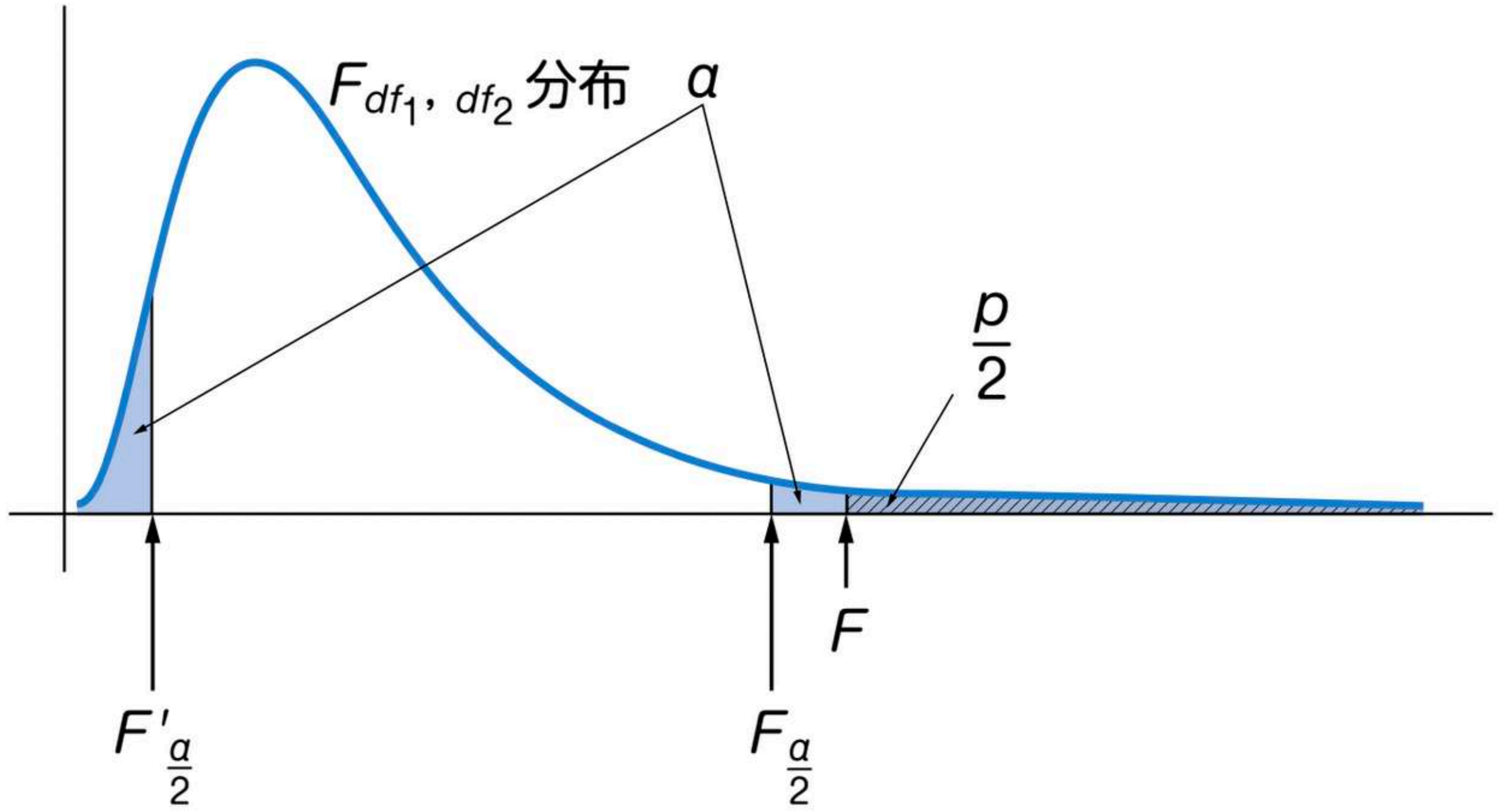


図 7·2　検定統計量 F の分布

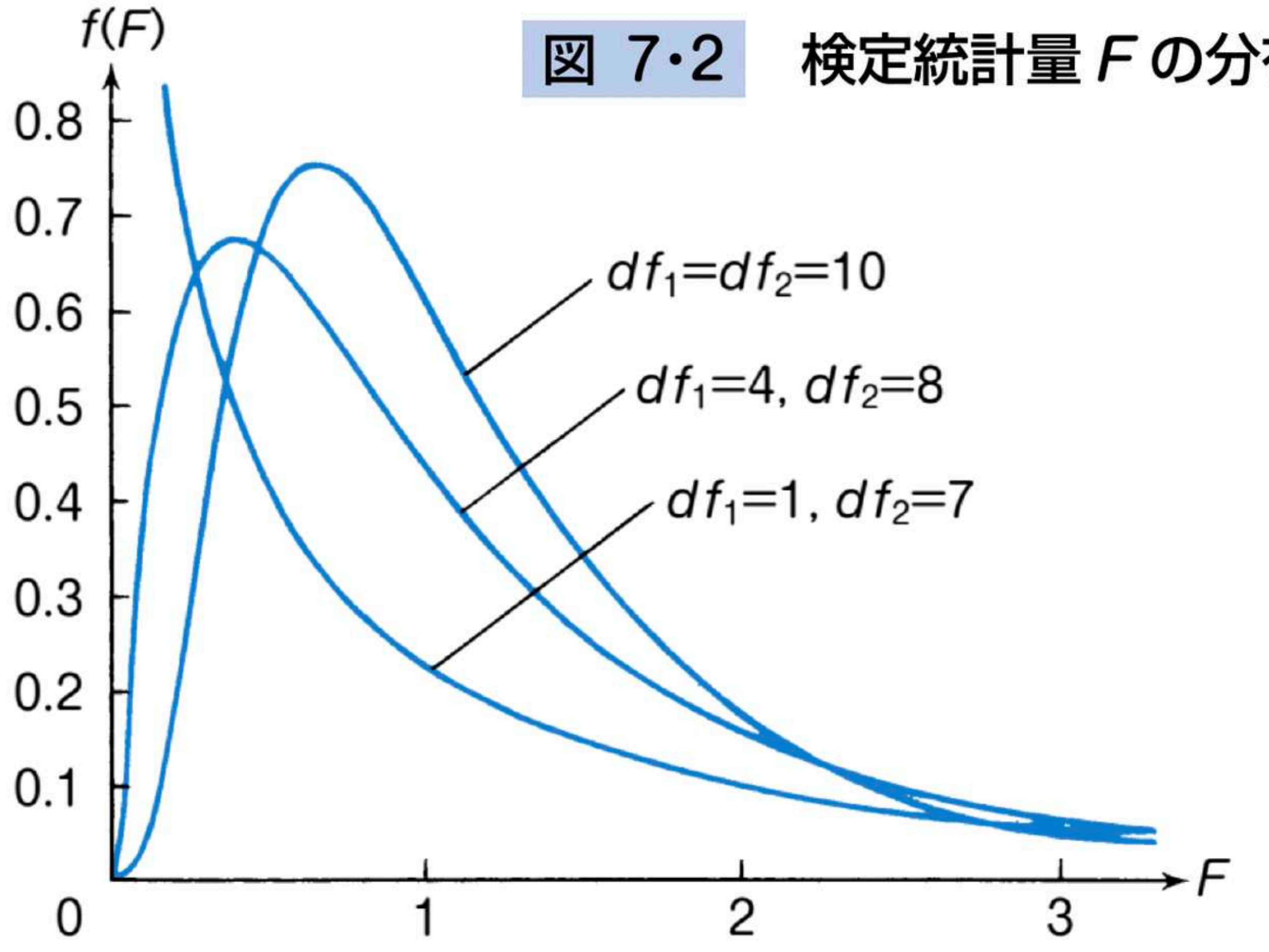


図 5·14　種々の自由度対の F 分布

2群のt検定

- 標本１：$N(\mu_1, \sigma_1^2)$の母集団から抽出した大きさn_1の標本X_1は正規分布$N(\mu_1, \sigma_1^2)$に従い、標本平均$\bar{X}_1$は$N(\mu_1, \frac{\sigma_1^2}{n_1})$の正規分布に従う。

- 標本２：$N(\mu_2, \sigma_2^2)$の母集団から抽出した大きさn_2の標本X_2は正規分布$N(\mu_2, \sigma_2^2)$に従い、標本平均$\bar{X}_2$は$N(\mu_2, \frac{\sigma_2^2}{n_2})$の正規分布に従う。

- これら２つの標本の標本平均の差$\bar{X}_1 - \bar{X}_2$は、正規分布の再生性より$N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$に従う。

- 従って、$\bar{X}_1 - \bar{X}_2$を標準化すると、

$$z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

正規分布の再生性

確率変数X_1, X_2が独立で$X_1 \sim N(\mu_1, \sigma_1^2),\ X_2 \sim N(\mu_2, \sigma_2^2)$であれば、

$$X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

$$X_1 - X_2 \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$$

正規分布に従う２つの独立な確率変数の和・差は、再び正規分布に従う

2群のt検定

- つまり、標本平均の差$\bar{X}_1 - \bar{X}_2$が正規分布に従っていると仮定できれば、t検定は使える

- したがって、母集団が正規分布に従わない場合でも、nが十分に大きければ、中心極限定理により以下の主張が可能となりt検定が使える

- 標本１：正規分布に従わない母集団から抽出した大きさn_1の標本X_1の標本平均$\bar{X}_1$は、n_1が十分大きければ中心極限定理により$N(\mu_1, \dfrac{\sigma_1^2}{n_1})$の正規分布に従う。

- 標本２：正規分布に従わない母集団から抽出した大きさn_2の標本X_2の標本平均$\bar{X}_2$は、n_2が十分大きければ中心極限定理により$N(\mu_2, \dfrac{\sigma_2^2}{n_2})$の正規分布に従う。

- したがって、 仮に２つの母集団が正規分布に従っていなくても、n_1, n_2が十分大きければ中心極限定理により、２つの標本の標本平均は正規分布に従うため、２つの標本の標本平均の差$\bar{X}_1 - \bar{X}_2$も正規分布の再生性より$N(\mu_1 - \mu_2, \dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2})$に従い、$t$検定が使える

Student’s t検定

対応のない２群のデータの分散が等しいときに、２つ平均値が等しいかどうかを判断する際に使う検定方法。

手順 ① 帰無仮説H_0、対立仮説H_1の設定。$H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 \neq \mu_2$
$H_1 : \mu_1 > \mu_2$
$H_1 : \mu_1 < \mu_2$

② 有意水準αを定める

③ 帰無仮説$H_0 : \mu_1 = \mu_2$のもとで検定統計量Tを計算

④ P値とαの比較（検定統計量が棄却域なのか採択域なのかを確認）

⑤ 結果の解釈

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

また母分散が等しいため、２つの標本を合体したときの標本分散S_p^2は次のように書ける。

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

したがって、検定統計量Tは以下と書ける。

$$T = \frac{\bar{X}_1 - \bar{X}_2}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

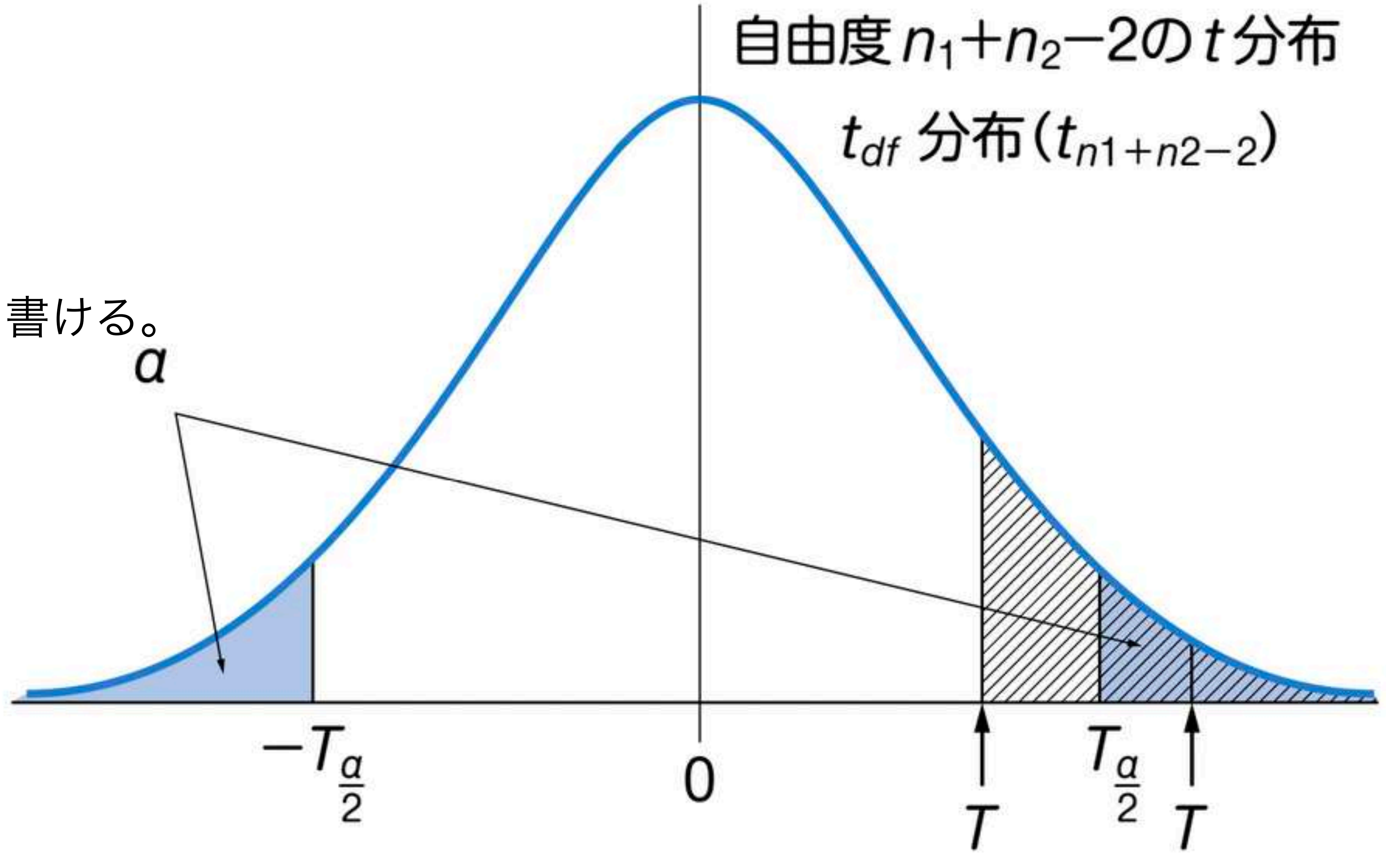


図 7·1 検定統計量 T の分布，$T_{\frac{\alpha}{2}}$ と α の関係（両側検定）

Welch's　t検定

対応のない２群のデータの分散が等しくない場合または等しいと断定できない場合に、２つ平均値が等しいかどうかを判断する際に使う検定方法。

① 帰無仮説H_0、対立仮説H_1の設定。$H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 \neq \mu_2$

$H_1 : \mu_1 > \mu_2$

$H_1 : \mu_1 < \mu_2$

② 有意水準αを定める

③ 帰無仮説$H_0 : \mu_1 = \mu_2$のもとで検定統計量Tを計算

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

このとき自由度dfは、以下の近似式から算出される（dfが整数にならない場合は、dfを超えない最大の整数をとる。）

$$df \approx \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{S_1^4}{n_1^2(n_1 - 1)} + \frac{S_2^4}{n_2^2(n_2 - 1)}}$$

④P値とαの比較（検定統計量が棄却域なのか採択域なのかを確認）

⑤結果の解釈