確率・統計
第7回 分散分析

兵庫県立大学　社会情報科学部
川嶋宏彰
kawashima@sis.u-hyogo.ac.jp

本日の講義内容

- テキスト
  - 「統計学入門」には本日の内容がありません
- 分散分析
  - 3群以上の母平均の差の検定
- 連絡
  - 12/2 中間テスト（30-40分ぐらい？）＋解説＋講義

今日のポイント (Webサイトより)

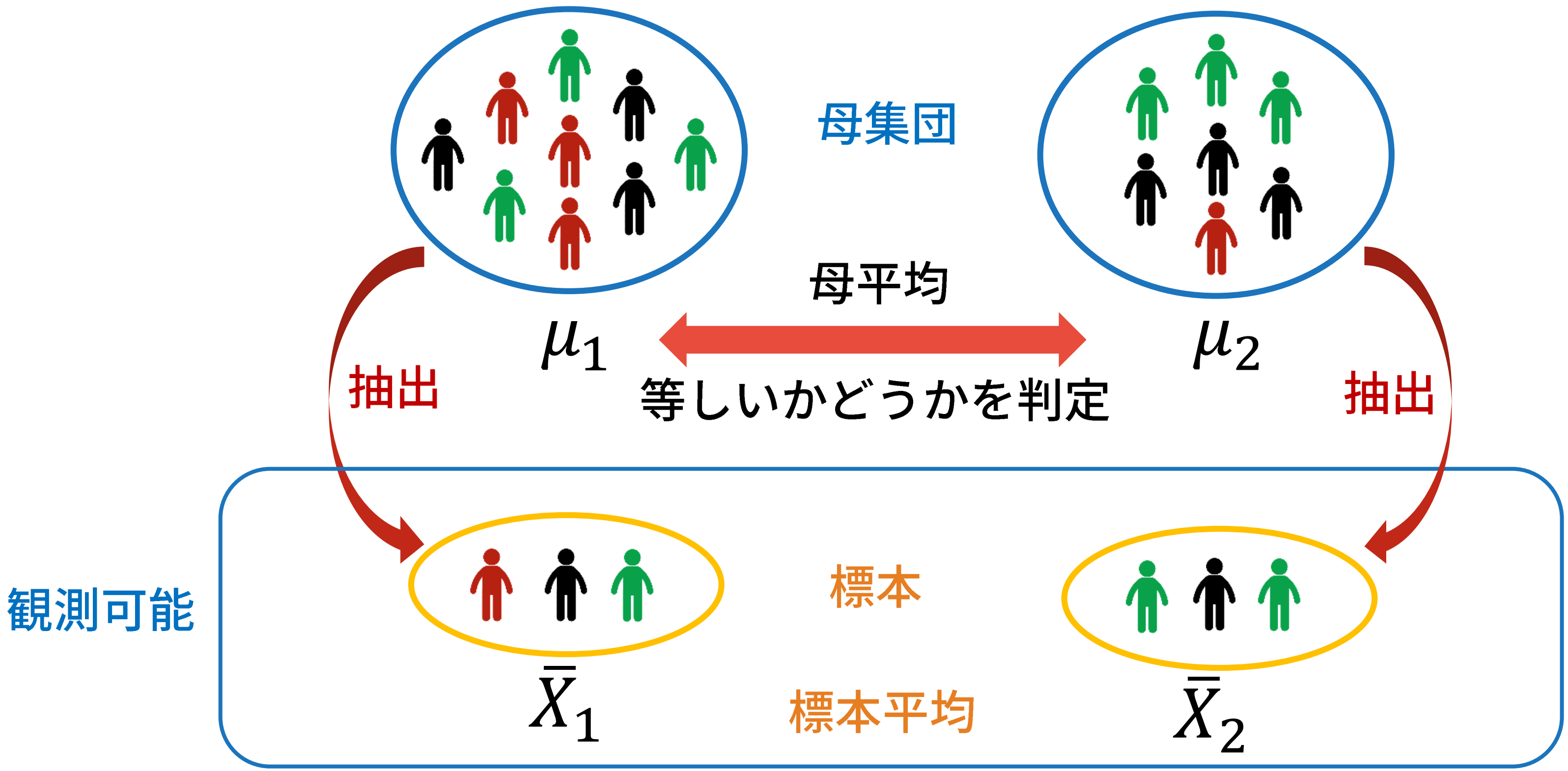
- F分布を用いたF検定の流れ
  - 何の検定に用いられるか?
    - 「等分散性の検定」： 2群の母分散が等しいか? （前回スライド）
    - 「分散分析」
  - F 分布の統計量は何か?：標本から得られた分散の比
- 多重検定の問題（なぜ分散分析を行うのか？）
  - (例) 3群について2群ずつ有意水準5%の検定を3回行った場合，<u>第一種の過誤が起こる確率</u>（つまり実質的な有意水準）は?
- 分散分析
  - 何を検定しているのか?（帰無仮説や対立仮説）
  - 大まかな考え方は?（群間の変動が群内の変動に対して…）
  - 分散分析表を埋めることができるか?（中間テストでは一部でよい）

統計量の差に関する検定

- 2群の標本に対する検定
  - 対応のある2標本群の差の検定 (t 検定)
  - 対応のない2標本群の差の検定
    - 母集団の平均に差があるか (t 検定)
    - 母集団の分散に差があるか (F 検定) → 第6回のスライドを使用
- 3群以上の標本に対する検定（分散分析）
  - 分散分析 (一元配置法)：複数の母集団の平均に差があるかを検定
  - 分散分析後の多重比較： どの母集団の平均に差があるか
  - 等分散性の検定
  - 対応のある場合の一元配置（繰り返しのない二元配置）
  - 二元配置法

(復習) 2群の標本に対する検定

- 二つの母集団からそれぞれ抽出された標本から母数が等しいかどうかを検定



(復習) 対応のない2群の標本に対する検定

- 母集団の平均に差があるかどうかを検定したい
  - 例： AクラスとBクラスの成績に差があるか？

Aクラス	Bクラス
69	49
52	40
68	52
46	37
72	55
40	38
45	45
62	
53	

	Aクラス	Bクラス
標本サイズ n	$n_1 = 9$	$n_2 = 7$
標本平均 $\bar{X}$	$\bar{X}_1 = 56.33$	$\bar{X}_2 = 45.14$
標準偏差 s	$s_1 = 11.76$	$s_2 = 7.105$
不偏分散 s^2	$s_1^2 = 138.3$	$s_2^2 = 50.48$
母平均	？	？
母分散	？	？

差があるか？

(前回) 対応のない2群の標本に対する平均の差の検定

- 仮説を設定

等分散性が仮定できる場合

  - 帰無仮説：母平均は等しい ($\mu_1 = \mu_2$)
  - 対立仮説：母平均は異なる ($\mu_1 \neq \mu_2$)　(片側検定なら $\mu_1 > \mu_2$ や $\mu_1 < \mu_2$)

- t 統計量を計算
  - 標本サイズ n_1, n_2 の2群の標本の母平均 μ_1, μ_2 が等しければ標本平均の差は自由度 $n_1 + n_2 - 2$ の t 分布に従う

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \sim t(n_1 + n_2 - 2)$$

2標本を合併 (pooling)

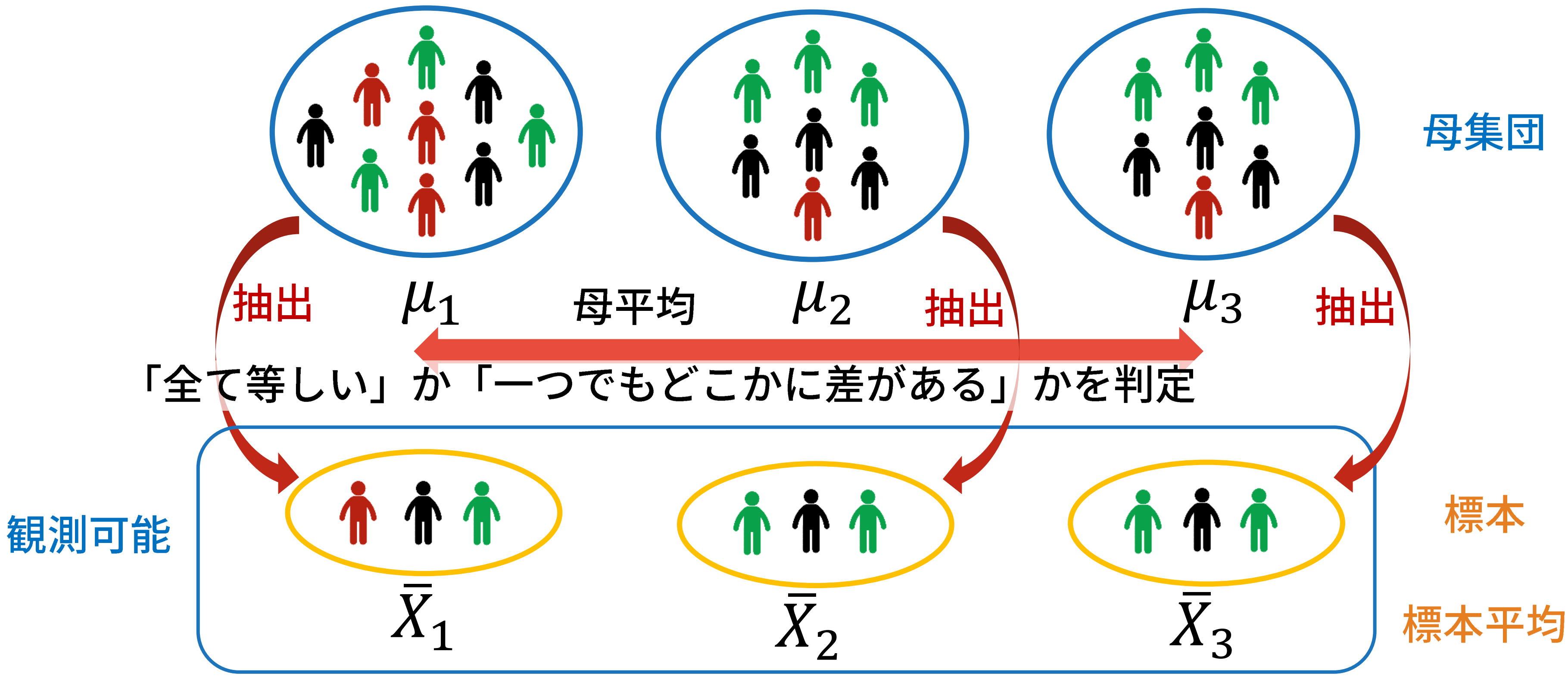
$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

- p 値を計算
  - 帰無仮説の下で，標本から得られた $|t^*|$ の上側確率を求め，両側検定なので2倍する

$$p = 2\,P(t \geq |t^*|; n_1 + n_2 - 2\,)$$

3群 (以上) の標本に対する検定

- 三つの母集団からそれぞれ抽出された標本から母数が等しいかどうかを検定



第一種の過誤

- <u>帰無仮説が正しいにも関わらず</u>帰無仮説を棄却してしまう<u>誤り</u>を第一種の過誤 (Type-I error) という
  - 要するに，本当は差がないのに差があると言ってしまう
  - 例：（本当は効果がないのに）薬を飲む場合と飲まない場合で「差がある」（つまり薬に効果がある）と判定してしまう
- 問題（中間テストで出題される可能性が高い）
  - 第一種の過誤とは？
  - 有意水準5%で検定を行ったときに，第一種の過誤が起きる確率は？
- 答え
  - 第一種の過誤が起きる確率は，実は有意水準そのもの（別名「危険率」）
  - つまり「有意水準 α」とは，第一種の過誤をどれくらい許すかの水準
    - 有意水準5% ($\alpha = 0.05$)で検定したならば，差がないときに，100回に5回は「差がある」と言ってしまう(第一種の過誤が起きる)のを許す

3群以上の標本に対する平均の差の検定

- 3群以上の母集団の平均に差があるか調べたい
  → 分散分析 (analysis of variance: ANOVA)

「群」を「水準」と呼ぶことも多い(例：1要因3水準)

水準　要因 (因子)

3種類の肥料それぞれで育てたスイカの重さの測定結果 (kg)

群1 (肥料1)	群2 (肥料2)	群3 (肥料3)
9.5	10.1	11.3
9.7	10.5	10.7
10.1	9.6	10.2
9.8	9.3	
9.3		
標本平均 9.7	9.9	10.7

3群以上まとめて分散分析
≠ 2群に対する t 検定の繰り返し

問題：3通りの2群(ペア)それぞれに対して有意水準5%のt検定を行うことを考える．このとき第一種の過誤が生じる確率は？（すべて等しいのに少なくとも1組に有意差を認めてしまう確率は？）

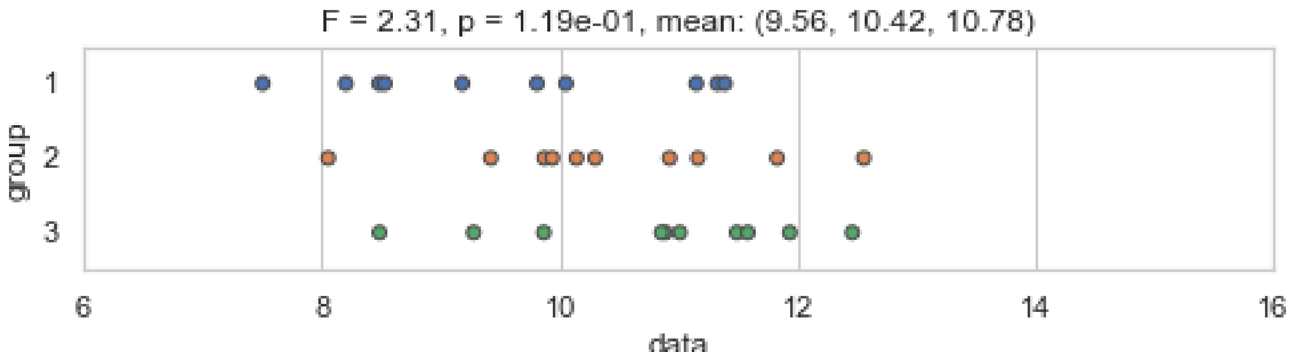
$$1 - 0.95 \times 0.95 \times 0.95 = 0.1426$$

（引き算の項：どの2群比較でも誤らない確率）

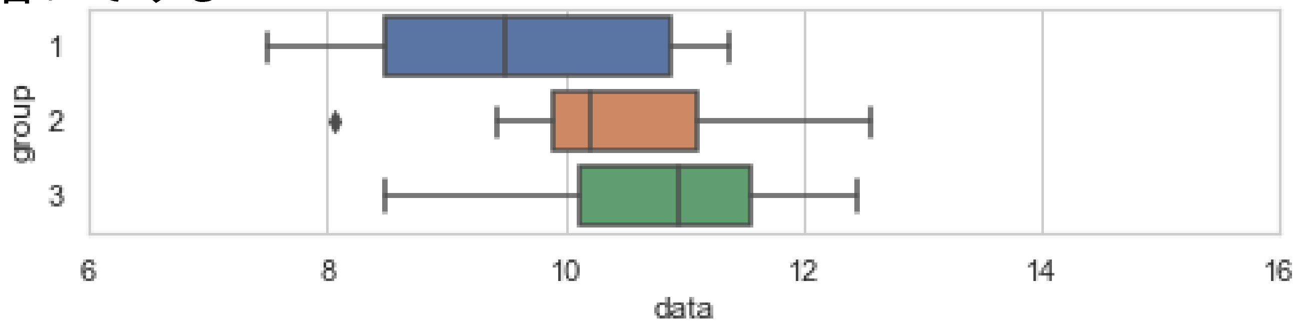
→ 実質，有意水準14%の検定になってしまう！

分散分析の考え方

- 群間で差はあるか？
  - 標本平均の差 (群2-群1): 0.86, (群3-群2): 0.36, (群3-群1): 1.22

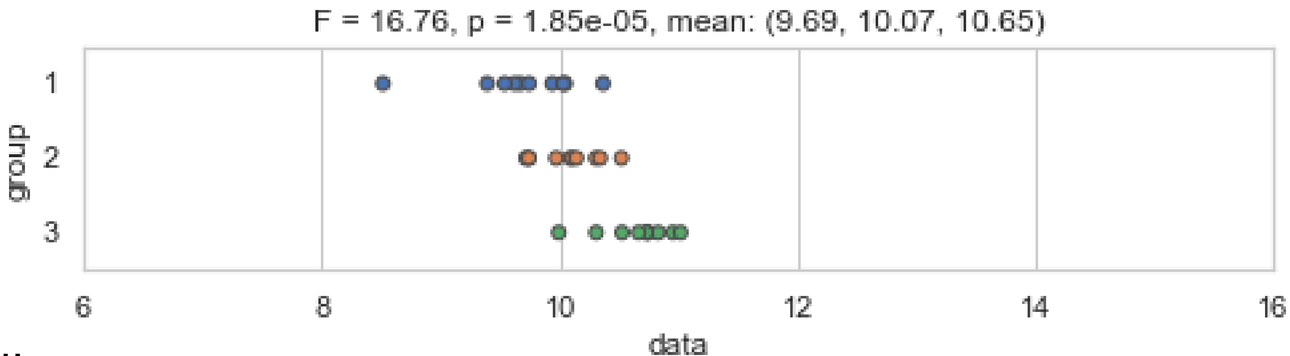


箱ひげ図も描いてみる

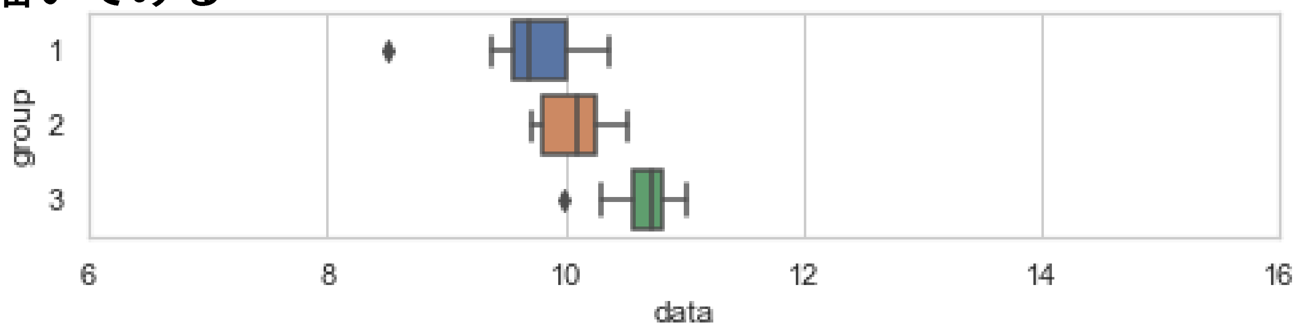


分散分析の考え方

- 群間で差はあるか？
  - 標本平均の差 (群2-群1): 0.38, (群3-群2): 0.58, (群3-群1): 0.96



箱ひげ図も描いてみる

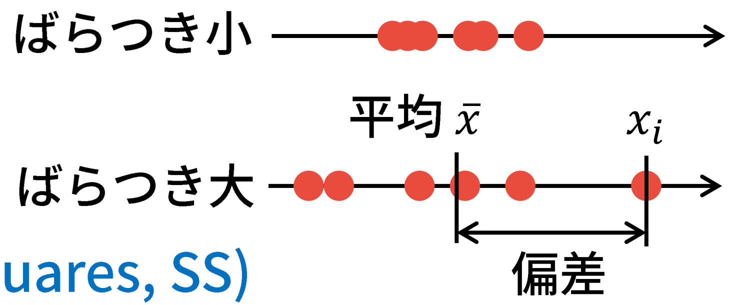


データのばらつきを表す指標

- データ (観測値) $x_1, \ldots, x_n$ のばらつきを定量化したい
  - 偏差 (deviation)
    - $x_i - \bar{x}$: 平均値からのずれ
    - ただし $\bar{x} = \frac{x_1 + \cdots + x_n}{n} = \frac{1}{n} \sum_{i=1}^{n} x_i$
  - 変動 (variation) or 平方和 (sum of squares, SS)
    - $Q = (x_1 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2$: 偏差の平方和

      $= \sum_{i=1}^{n} (x_i - \bar{x})^2$
  - 不偏分散 (unbiased variance)
    - $Q/(n-1)$: 不偏性のある標本分散
    - $n - 1$: 自由度 (平方和で独立に動かせる成分の数)

      - $(x_1 - \bar{x}) + \cdots + (x_n - \bar{x}) = 0 \quad \Leftrightarrow \bar{x} = \frac{x_1 + \cdots + x_n}{n}$



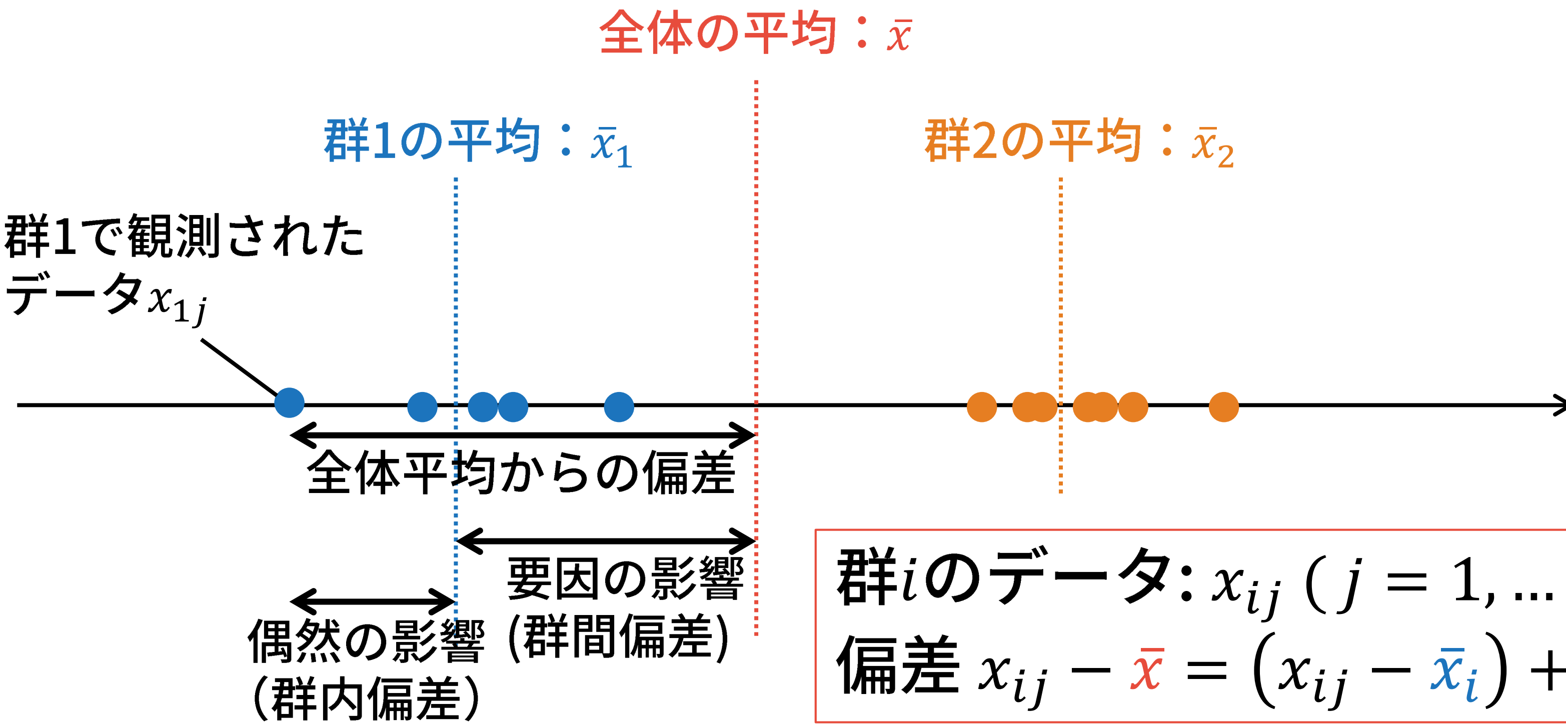
(練習: 変動の計算)

5 5 4 4 5 4 : $Q =$

1 7 6 3 2 8 : $Q =$

分散分析の考え方

- データのずれを<u>要因の影響</u>と<u>偶然の影響</u>に分けて説明する



群iのデータ: $x_{ij}\ (j = 1, \ldots, n_i)$

偏差 $x_{ij} - \bar{x} = (x_{ij} - \bar{x}_i) + (\bar{x}_i - \bar{x})$

群内偏差　　群間偏差

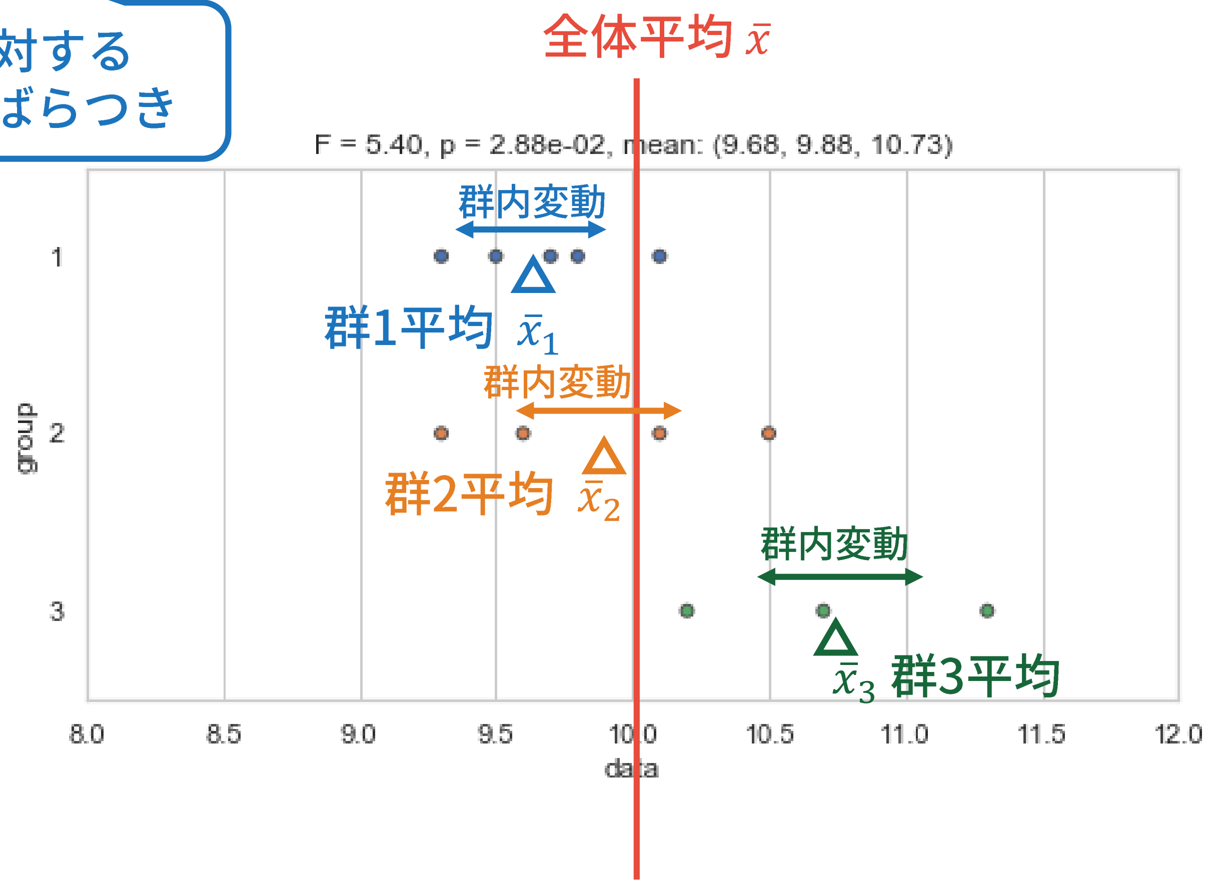
分散分析の考え方

- 全体の変動 ＝群間の変動（要因効果）＋群内の変動（誤差）

全体平均 $\bar{x}$ に対する
各群の平均$\bar{x}_i$のばらつき

群1 (肥料1)	群2 (肥料2)	群3 (肥料3)
9.5	10.1	11.3
9.7	10.5	10.7
10.1	9.6	10.2
9.8	9.3	
9.3		

平均 9.7 9.9 10.7



各データの群間偏差・群内偏差を計算

- 全体平均からの変動を二つの成分に分ける
  - 群間の変動：全体平均と各群平均の差（群間偏差）の平方和
  - 群内の変動：各群内の平均からの差（群内偏差）の平方和

ここに興味がある

これが大きいと
群間の変動が埋もれる

まず各データの偏差を求める

$$x_{ij} = \bar{x} + (\bar{x}_i - \bar{x}) + (x_{ij} - \bar{x}_i)$$

元のデータ（i：列方向、j：行方向）

	群1	群2	群3
	9.50	10.10	11.30
	9.70	10.50	10.70
	10.10	9.60	10.20
	9.80	9.30	
	9.30		
平均	9.68	9.88	10.73

=

全体の平均 $\bar{x}$

群1	群2	群3
10.01	10.01	10.01
10.01	10.01	10.01
10.01	10.01	10.01
10.01	10.01	
10.01		

+

群間偏差

群1	群2	群3
-0.33	-0.13	0.73
-0.33	-0.13	0.73
-0.33	-0.13	0.73
-0.33	-0.13	
-0.33		

+

群内偏差

群1	群2	群3
-0.18	0.23	0.57
0.02	0.63	-0.03
0.42	-0.28	-0.53
0.12	-0.58	
-0.38		

群１の平均9.68の全体平均10.01からの差

群１の平均9.67からの差

変動（偏差平方和）を計算

元のデータ

群1	群2	群3
9.50	10.10	11.30
9.70	10.50	10.70
10.10	9.60	10.20
9.80	9.30	
9.30		
平均 9.68	9.88	10.73

=

全体の平均 $\bar{x}$

群1	群2	群3
10.01	10.01	10.01
10.01	10.01	10.01
10.01	10.01	10.01
10.01	10.01	
10.01		

＋

群間偏差

群1	群2	群3
-0.33	-0.13	0.73
-0.33	-0.13	0.73
-0.33	-0.13	0.73
-0.33	-0.13	
-0.33		

＋

群内偏差

群1	群2	群3
-0.18	0.23	0.57
0.02	0.63	-0.03
0.42	-0.28	-0.53
0.12	-0.58	
-0.38		

各セルが偏差の平方（二乗）
（例）$(-0.33)^2 = 0.11$

群間偏差の平方

群1	群2	群3
0.11	0.02	0.53
0.11	0.02	0.53
0.11	0.02	0.53
0.11	0.02	
0.11		

＋

群内偏差の平方

群1	群2	群3
0.03	0.05	0.32
0.00	0.39	0.00
0.18	0.08	0.28
0.01	0.33	
0.14		

総和を求める

群間の変動：2.19
(群間偏差の平方和)

群内の変動：1.82
(郡内偏差の平方和)

分散分析の考え方

- 群間の差の大きさを測りたい
  - 適切な統計量を定義して，それが（帰無仮説の下で）何かの分布に従うなら検定可能 → どのような統計量を定義すればよいだろうか？
  - F分布のことを思い出すと・・・不偏分散の比が従う分布
    - 不偏分散は，変動 (平均からの差の平方和) を自由度で割ったもの
    - F分布に従う「群間の差の大きさの指標」が作れるかもしれない

帰無仮説の下で，この比は
自由度 (ϕ_A, ϕ_e) の F 分布に従う

<u>分散分析が用いる統計量</u>

$$F = \frac{\text{群間の変動}/\text{自由度}\phi_A\text{（水準間の差異，要因効果）}}{\text{群内の変動}/\text{自由度}\phi_e\text{（偶然性，誤差）}}$$

自由度とF値の計算

- **自由度**（実質的に自由に動かせる変数の数）

標本サイズ
5+4+3

  - 群間自由度は 2 (= 3群−1平均)，群内自由度は 9 (= 12要素−3平均)

- **統計量Fの値を計算**

(変動/自由度)を平均平方と呼ぶ

  - 変動 (偏差平方和) を自由度で割るとそれぞれ $\frac{2.19}{2} = 1.09$ と $\frac{1.82}{9} = 0.21$
  - この比が $F = 1.09/0.21 = 5.40$ → 自由度 (2,9) の F 分布で検定

自由度2：
$(\bar{x}_1 - \bar{x})$と$(\bar{x}_2 - \bar{x})$
が決まると
$(\bar{x}_3 - \bar{x})$は自動的
に決まる
$(\sum n_i(\bar{x}_i - \bar{x}) = 0)$

群間の変動
(群間偏差の平方和)：2.19

群1	群2	群3
0.11	0.02	0.53
0.11	0.02	0.53
0.11	0.02	0.53
0.11	0.02	
0.11		

+

群内の変動
(郡内偏差の平方和)：1.82

群1	群2	群3
0.03	0.05	0.32
0.00	0.39	0.00
0.18	0.08	0.28
0.01	0.33	
0.14		

自由度9：
各群の自由度の和
群1は $5-1=4$
群2は $4-1=3$
群3は $3-1=2$
（スライド13）

一元配置分散分析

- 仮説を設定
  - 帰無仮説：すべての母平均は等しい $(\mu_1 = \mu_2 = \cdots = \mu_I)$
  - 対立仮説：いずれかの母平均は異なる $(\mu_i \neq \mu_{i'}$ for some $i, i')$

群i の平均
$$\bar{x}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}$$

全体平均
$$\bar{x} = \frac{1}{n}\sum_{i=1}^{I}\sum_{j=1}^{n_i} x_{ij}$$

- 分散分析表 を計算 $(x_{ij}$：群iのデータ, $j = 1, \ldots, n_i)$

変動要因 SV	平方和(変動) SS	自由度 df	平均平方MS	分散比 (F値)
グループ間変動	$SS_A = \sum_{i=1}^{I} n_i(\bar{x}_i - \bar{x})^2$	$\phi_A = I - 1$	$MS_A = \frac{SS_A}{\phi_A}$	$F_{Ae} = \frac{MS_A}{MS_e}$
グループ内変動	$SS_e = \sum_{i=1}^{I}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)^2$	$\phi_e = \sum_{i=1}^{I}(n_i - 1)$	$MS_e = \frac{SS_e}{\phi_e}$	
合計	$SS_T = \sum_{i=1}^{I}\sum_{j=1}^{n_i}(x_{ij} - \bar{x})^2$	$\phi_T = \sum_{i=1}^{I} n_i - 1$		

$SS_A + SS_e = SS_T$　　$\phi_A + \phi_e = \phi_T$

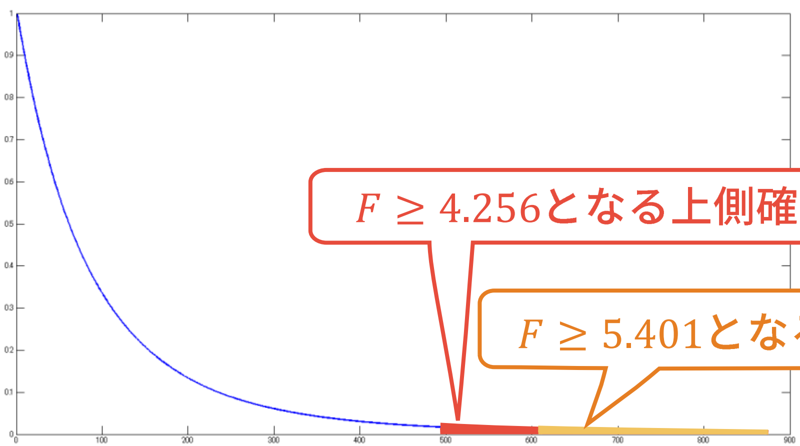
自由度(ϕ_A, ϕ_e)のF分布に従う

一元配置分散分析の実行例

- 自由度(2,9)のF分布を用いて有意水準5%で検定

$\bar{x}_1 = 9.680$ $\bar{x}_2 = 9.875$ $\bar{x}_3 = 10.73$ $\bar{x} = 10.01$

変動要因SV	平方和SS	自由度df	平均平方MS	分散比 (F値)
グループ間変動	2.187	2	1.094	5.401
グループ内変動	1.822	9	0.202	
合計	4.009	11		

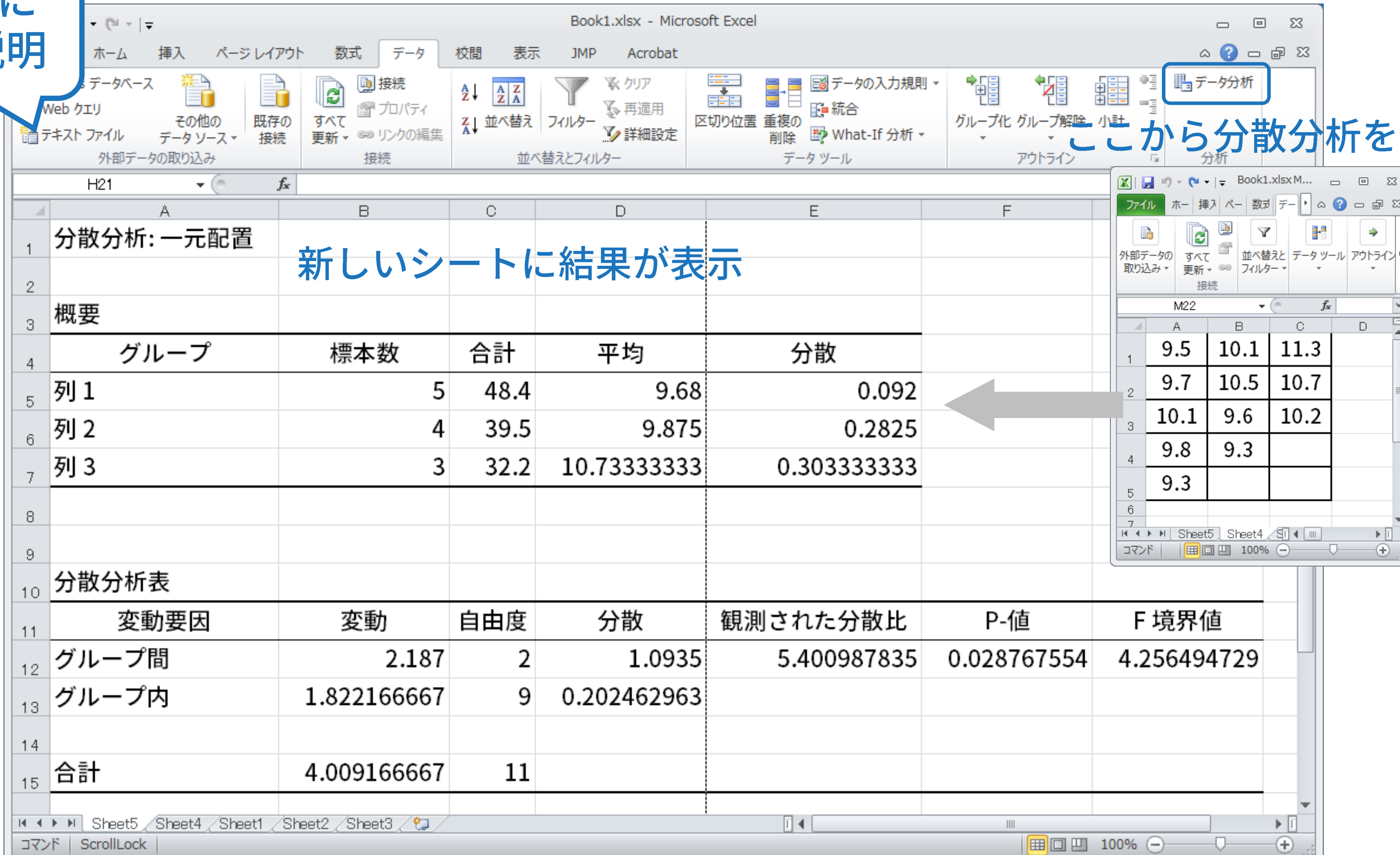


「分散分析の結果、有意差が認められた
($p = 0.029 < 0.05$)」

つねに片側検定：
グループ間変動のほうが
小さい場合は興味なし

一元配置分散分析の実行例 (Excel)

第9回に
詳細説明



分散分析: 一元配置

概要

グループ	標本数	合計	平均	分散
列 1	5	48.4	9.68	0.092
列 2	4	39.5	9.875	0.2825
列 3	3	32.2	10.73333333	0.303333333

分散分析表

変動要因	変動	自由度	分散	観測された分散比	P-値	F 境界値
グループ間	2.187	2	1.0935	5.400987835	0.028767554	4.256494729
グループ内	1.822166667	9	0.202462963			
合計	4.009166667	11				

A	B	C
9.5	10.1	11.3
9.7	10.5	10.7
10.1	9.6	10.2
9.8	9.3	
9.3		

分散分析のまとめと補足

- 分散分析とは3群以上の母平均の差の検定
  - 分散分析表を用いて，Fが大きければ (すべての群の母平均が等しいという帰無仮説を棄却し) 要因に効果があると判定

$$F = \frac{\text{群間の変動/自由度}\phi_A\text{（水準間の差異，要因効果）}}{\text{群内の変動/自由度}\phi_e\text{（偶然性，誤差）}}$$

  - 補足
    - 母集団は正規分布（正規母集団）を仮定（t 検定と同様）
    - いずれの群の母分散も等しいことを仮定
    - 実は，2群に対する分散分析 ＝ 2群に対する両側 t 検定
    - 群内，群間の「群（グループ）」は「水準」「級」とも呼ばれる
- 第9回に t 検定や分散分析の演習(PC)を予定

復習の参考

- 「確率・統計」のウェブサイトで重要ポイントを列挙
  - 中間テストは主にこのポイント＋宿題やレポートから出題
- 参考図書
  - 教科書「統計学入門」（授業スライドで範囲を指定）
  - 第3～5回の内容：前期「統計学」や高校数学Bの教科書も参考になる
  - 分散分析：涌井良幸，涌井貞美「統計解析がわかる」技術評論社など
- Rコマンダーの復習方法
  - 第1回からタイトルに[Rcmdr]があるスライドを順に行う
  - 第1回のスライドで紹介したブルーバックスを参考にする

練習問題

表1は，それぞれ異なる薬a, b, c を投与した3群において，薬に対する反応を測定した数値とする（以下、単に反応と呼ぶ）．これより3種類の薬に対する反応の差を検定したい．ただし，それぞれの薬に対する反応は，母分散の等しい正規分布を仮定できるものとする．

1. 各群の標本サイズ，平均，変動，不偏分散を求めよ．
2. 表1のデータで3群の一元配置分散分析を行うための分散分析表を作成せよ．
3. 分散分析に用いる分布とその自由度を述べよ．
4. 棄却域か p値を求めよ．両側と片側検定のどちらか？検定結果も述べよ．

表1. 薬に対する反応

薬a	薬b	薬c
5	10	5
4	8	10
4	9	6
7	9	

各群の統計量，分散分析表

1. 各群の標本サイズ，平均，変動，不偏分散を求めよ．

	薬a	薬b	薬c
標本サイズ	n_a = 4	n_b = 4	n_c = 3
標本平均			
変動（偏差平方和）			
不偏分散			

2. 表1のデータで3群の一元配置分散分析を行うための分散分析表を作成せよ．

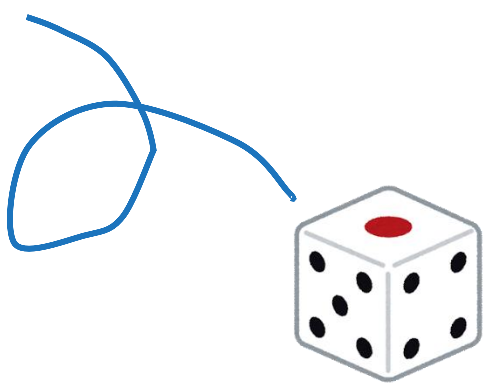
	変動（偏差平方和）	自由度	平均平方	分散比
群間変動				
群内変動				
全体	54	10		

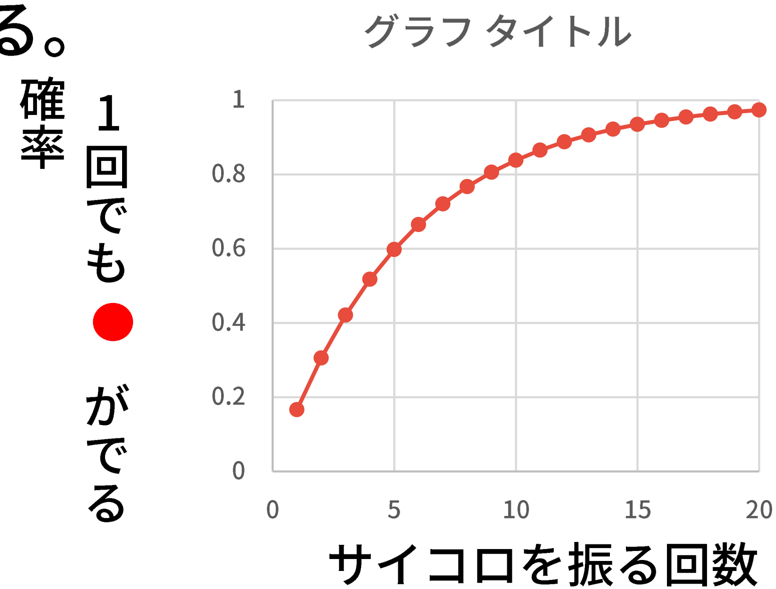
多重検定で起こりうること

- サイコロを振って ● がでたら、第1種過誤（α=0.167）と判断することにする。

有意水準

- サイコロを振る回数が増えれば、1回でも ● がでる（第1種過誤が起こる）可能性は増える。





多重検定の問題

- ３群（A・B・C）での平均の差を比較するとき
  - A群 ⇔ B群、 B群 ⇔ C群、 C群 ⇔ A群の３回検定を行うと $1-0.95^3=0.14$ と第１種過誤が大きくなる

例） 新薬Ａ、新薬Ｂと同効既存薬Ｃの比較など

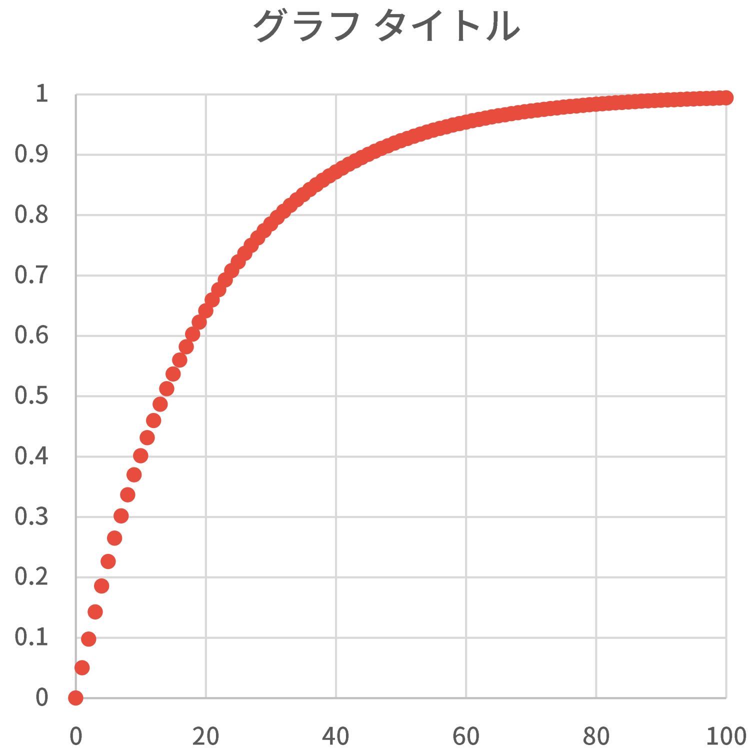
- 4群（A・B・C・D）での平均の差を比較するとき
  - A群 ⇔ B群、 A群 ⇔ C群、 A群 ⇔ D群
    B群 ⇔ C群、 B群 ⇔ D群、 C群 ⇔ D群の6回検定を行うと $1-0.95^6=0.26$ と第１種過誤が大きくなる

例） 経過観察群Ａ、手術実施群Ｂ、抗がん剤使用群Ｃ、手術・抗がん剤併用群Dの比較など

手術で体力が落ちて抗がん剤の副作用に
耐えられない人が多くなるかも？

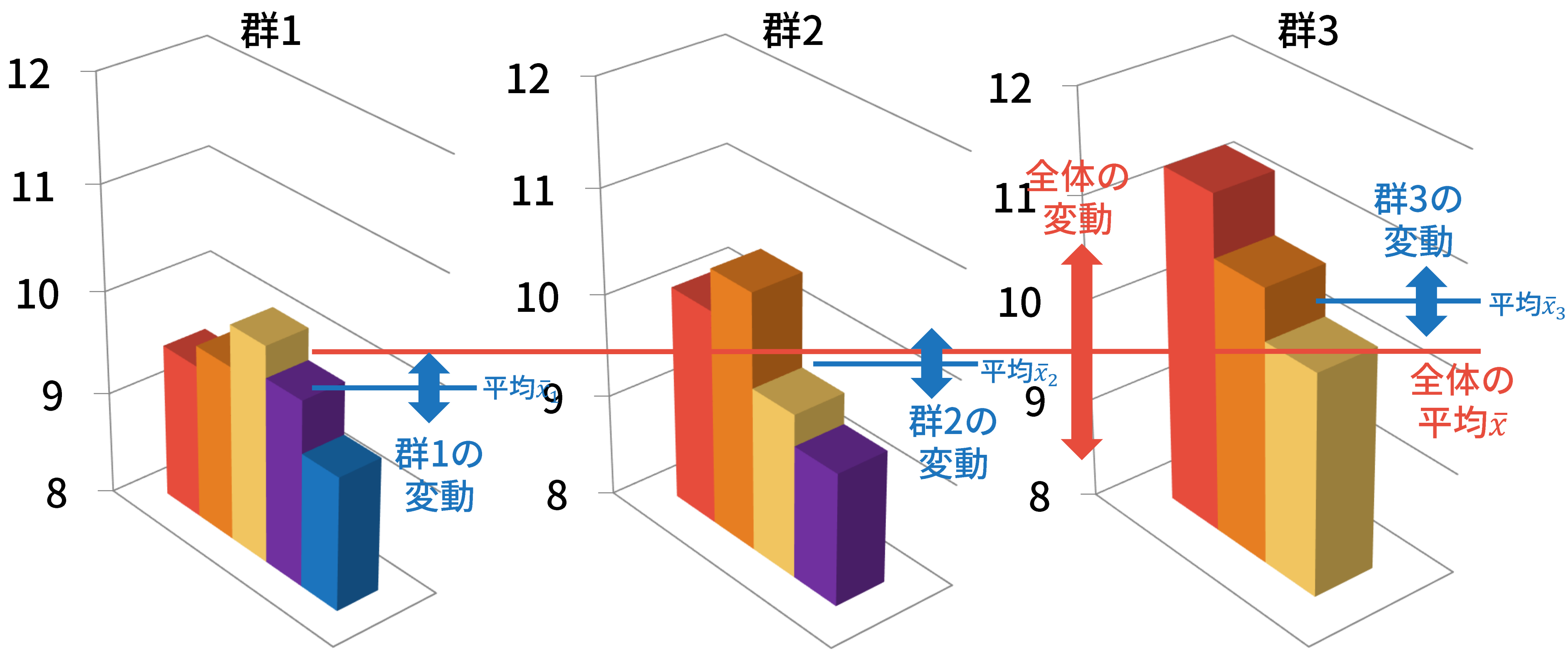
多重検定の結果

- 有意水準0.05の検定を繰り返した時に、1回でも第1種過誤が起こる確率



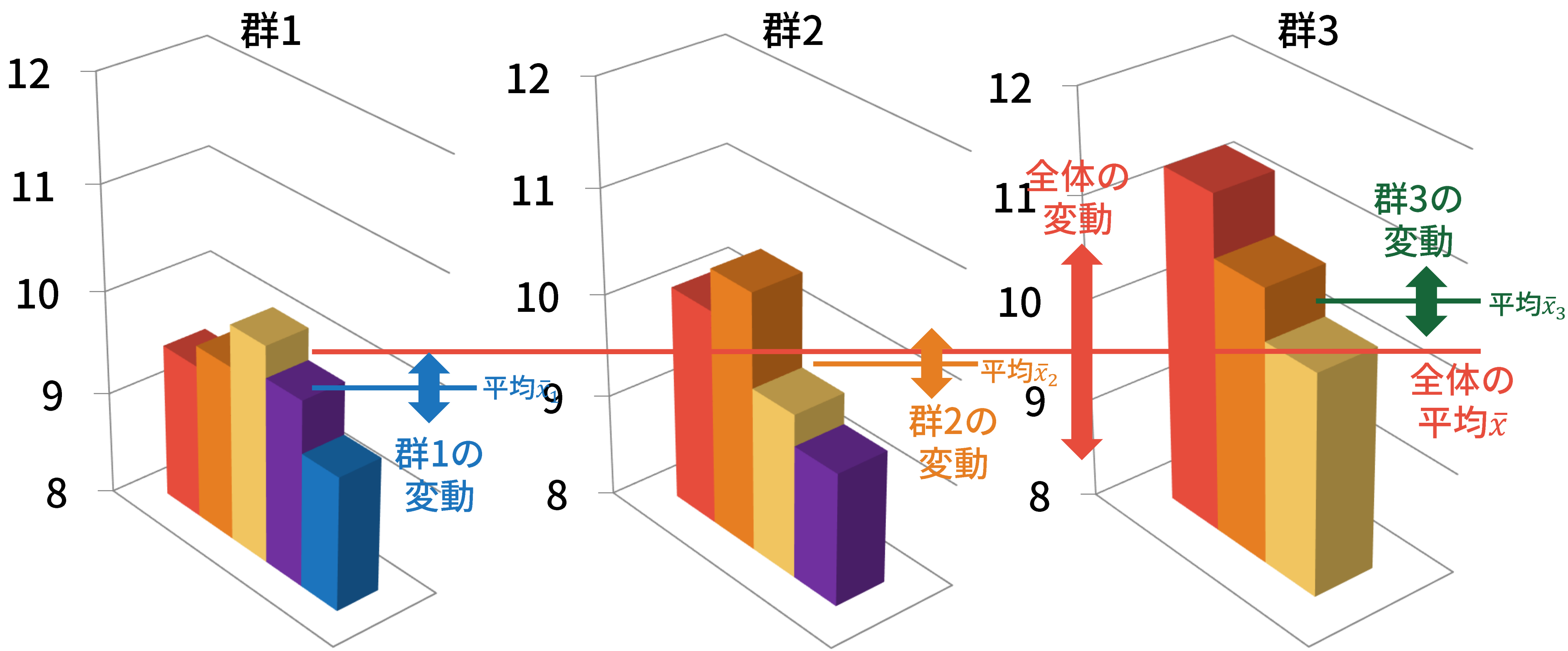
分散分析の考え方

- 全体の変動 ＝群間の変動（要因効果）＋群内の変動（誤差）



分散分析の考え方

- 全体の変動 ＝群間の変動（要因効果）＋群内の変動（誤差）



3群以上の標本に対する検定

- 標本群の性質によって使い分けが必要
  1. 母集団の分布が正規分布と仮定できるか
     - 仮定できる → パラメトリック検定
     - 仮定できない → ノンパラメトリック検定
  2. 群間に対応があるか

Bartlett検定の結果
標本群の分散が異なる場合にも

パラメトリック検定		ノンパラメトリック検定	
対応なし	対応あり	対応なし	対応あり
一元配置 分散分析	二元配置 分散分析	Kruskal-Wallis 検定	Friedman検定

講義では扱わないが重要

事後検定(post-hoc test)

- 分散分析（ANOVA）によって、帰無仮説「すべての母平均は等しい」が棄却され、「いずれか群の母平均が異なる」と分かったら？
  - 「どの群が異質なのか知りたい！」
  - 「でも、0.05のt検定を繰り返すのは危険そうだ・・・」

多重比較

- 多重検定の問題を考慮した方法
  - 複数の2群間の差の検定を同時に行っても、一つ以上の2群間の差が有意となる確率をあらかじめ定めた有意水準以内にする検定方法
    - Bonferroni法
      - 全体として有意水準が満たされるよう有意水準を下げてすべての群間でそれぞれ個別に検定

改良手法：Holm法・Shaffer法

    - Tukey法
      - 母平均について群間ですべての対比較を同時に検定
    - Dunnett法
      - １つの対照群と２つ以上の処理群があって、母平均について対照群と処理群の対比較のみを同時に検定
      - 各処理群の母平均が対照群の母平均と比べ 「異なるかどうか」だけでなく「小さいといえるか」または「大きい といえるか」を判定

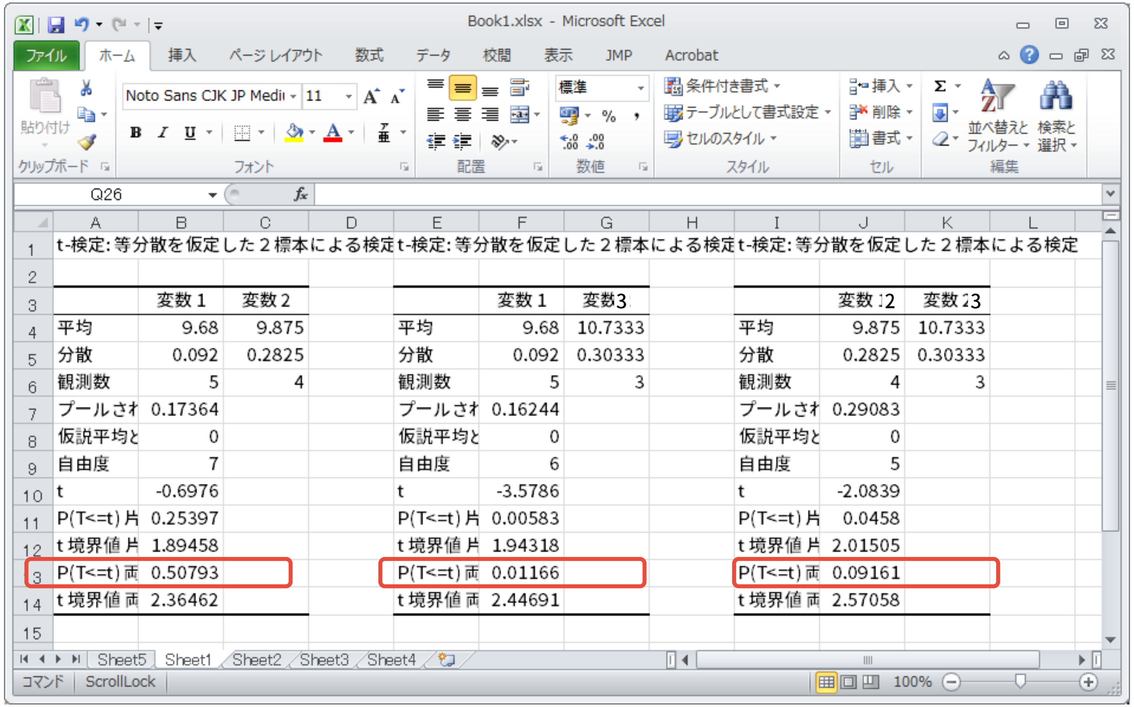
Bonferroni修正法

- 3群に対する多重検定 → 各群間に対してそれぞれ検定を行う
  - 有意水準 $\frac{\alpha}{3}$ として個別に検定 ($\alpha = 0.05 \rightarrow p < 0.0167$ならば有意差あり)
    - $1 - (1 - 0.167)^3 = 0.049$ とほぼ5%に

ある物質の血中濃度の測定結果

群1	群2	群3
9.5	10.1	11.3
9.7	10.5	10.7
10.1	9.6	10.2
9.8	9.3	
9.3		

有意差あり


Book1.xlsx - Microsoft Excel

	A	B	C	D	E	F	G	H	I	J	K
1	t-検定: 等分散を仮定した２標本による検定				t-検定: 等分散を仮定した２標本による検定				t-検定: 等分散を仮定した２標本による検定		
2											
3		変数 1	変数 2			変数 1	変数 3			変数 2	変数 3
4	平均	9.68	9.875		平均	9.68	10.7333		平均	9.875	10.7333
5	分散	0.092	0.2825		分散	0.092	0.30333		分散	0.2825	0.30333
6	観測数	5	4		観測数	5	3		観測数	4	3
7	プールされ	0.17364			プールされ	0.16244			プールされ	0.29083	
8	仮説平均と	0			仮説平均と	0			仮説平均と	0	
9	自由度	7			自由度	6			自由度	5	
10	t	-0.6976			t	-3.5786			t	-2.0839	
11	P(T<=t) 片	0.25397			P(T<=t) 片	0.00583			P(T<=t) 片	0.0458	
12	t 境界値 片	1.89458			t 境界値 片	1.94318			t 境界値 片	2.01505	
13	P(T<=t) 両	0.50793			P(T<=t) 両	0.01166			P(T<=t) 両	0.09161	
14	t 境界値 両	2.36462			t 境界値 両	2.44691			t 境界値 両	2.57058	
15											

例）ゲノム研究では何万という塩基配列の差をしらべる・・・

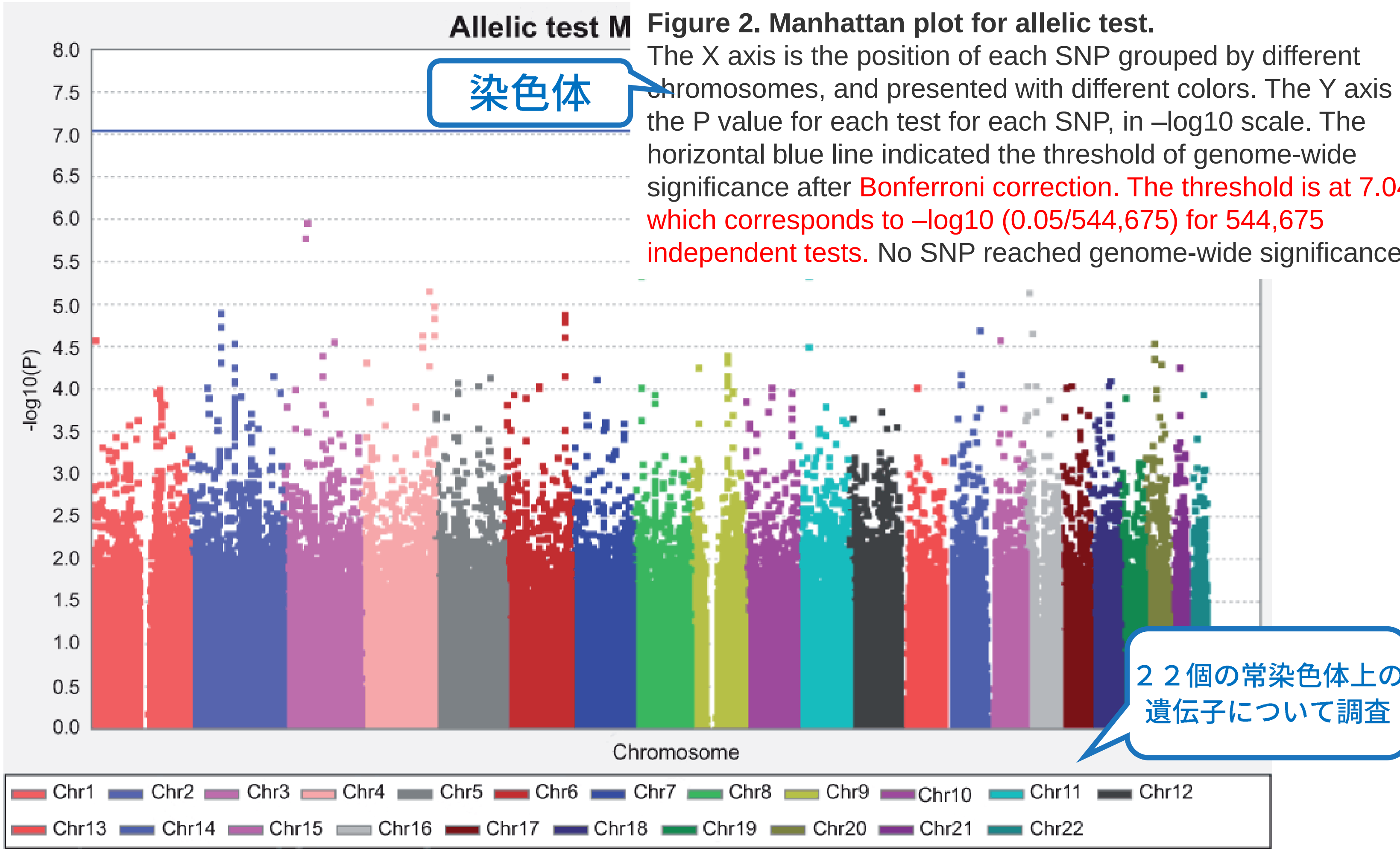


Figure 2. Manhattan plot for allelic test. The X axis is the position of each SNP grouped by different chromosomes, and presented with different colors. The Y axis is the P value for each test for each SNP, in –log10 scale. The horizontal blue line indicated the threshold of genome-wide significance after Bonferroni correction. The threshold is at 7.04, which corresponds to –log10 (0.05/544,675) for 544,675 independent tests. No SNP reached genome-wide significance.

3群以上の標本に対する検定

1. まずは分散分析で全群の母平均が等しいのかどうかをしらべる
2. 帰無仮説「すべての母平均は等しい」が棄却され、「いずれか群の母平均が異なる」と分かったら
3. 補正を加味した事後検定 (post-hoc test)で「どの群間に差があるのか」を調べる

（分散分析を省ける場合もあるが、分散分析を経た方が安全）

3群以上の標本に対する検定

- 標本群の性質によって使い分けが必要
  1. 母集団の分布が正規分布と仮定できるか
     - 仮定できる → パラメトリック検定
     - 仮定できない → ノンパラメトリック検定
  2. 群間に対応があるか

Bartlett検定の結果
標本群の分散が異なる場合にも

パラメトリック検定		ノンパラメトリック検定	
対応なし	対応あり	対応なし	対応あり
一元配置 分散分析	二元配置 分散分析	Kruskal-Wallis 検定	Friedman検定

$\underbrace{\text{Kruskal-Wallis検定 / Friedman検定}}_{\text{講義では扱わないが重要}}$