

第3回 テスト理論

基本情報 / Basic Information	詳細情報 / Detailed Information	授業計画詳細情報 / Class Schedule Details
講義題目 / Subtitle	心理測定のための数学的道具 / Mathematical Tools for Psychometrics	
授業の目標、概要 / Course Objectives / Overview	心理統計学 (psychometrics) の様々な手法を理解・開発するために必要となる数学的道具を学ぶ。	
授業のキーワード / Keywords	日本語用 / Japanese	心理統計学、テスト理論、行列演算、統計計算
	英語用 / English	psychometrics, test theory, matrix operation, statistical computation
授業計画 / Schedule	2020 年度は主に行列演算は添え字演算の基礎と、それらを用いた回帰モデルに関して学んだが、今年度は、それらを発展させた内容を紹介する。具体的には心理統計学でよく用いられる確率分布とそれらを利用した母数の推定方法、ならびにテスト理論の基礎を中心的話題として紹介する。	
授業の方法 / Teaching Methods	講義形式と R を用いた演習並びに討論	
成績評価方法 / Grade Evaluation	数回の課題と学期末のレポート	
教科書 / Textbook	未定	
参考書 / Reference	以下の書籍を用いる可能性がある。 鈴木謙 (2020) 統計的機械学習の数理共立出版	
履修上の注意 / Notes on Taking the Course	以下の文書を一読しておいてほしい。(すべてを理解している必要はない。) http://mayekawa.in.coocan.jp/titech/docs/summation2.pdf http://mayekawa.in.coocan.jp/titech/docs/mat11.pdf また、昨年度の授業の概要は以下に掲示してある。(すべてを理解している必要はない。) 2020 年度の資料 http://mayekawa.in.coocan.jp/u-tokyo/index.html	
関連ホームページ / Course-Related Websites	http://mayekawa.in.coocan.jp/titech/docs/docs.html	
その他 / Others	上記のサイトに載っている文書を適宜利用する可能性がある。	
メールアドレス / e-mail address	mayekawa@nifty.com	

古典的テスト理論

$$X_i = T_i + E_i, \quad i = 1, 2, \dots, n$$

$$X_{ik} = T_i + E_{ik}, \quad i = 1, 2, \dots, n, \quad k = 1, 2, \dots, m$$

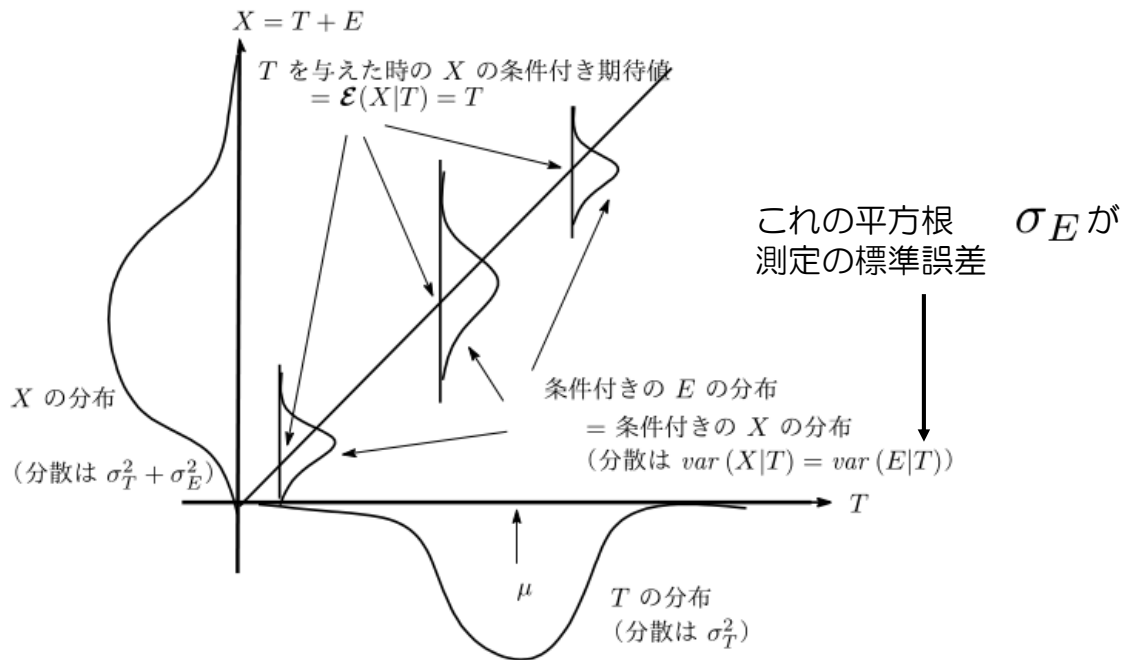


図1 古典的テスト理論

古典的テスト理論

基本的仮定

$$X = T + E$$

$$\mathcal{E}(E|T) = 0$$

どちらか一方

$$\mathcal{E}(X|T) = T$$

定義

$$\sigma_{E|T}^2 \equiv \text{var}(E|T)$$

測定の標準誤差 (SEM)

$$\mathcal{E}(T) = \mu \quad \text{and} \quad \text{var}(T) = \sigma_T^2$$

結果

$$\mathcal{E}(X) = \mu \quad \text{and} \quad \text{cov}(T, E) = 0$$

$$\text{var}(X) \equiv \sigma_X^2 = \sigma_T^2 + \sigma_E^2$$

$$\sigma_E^2 \equiv \mathcal{E}(\text{var}(E|T)) = \mathcal{E}(\text{var}(X|T))$$

テストの信頼性

2.1 信頼性係数

テスト X の信頼性係数 (reliability coefficient) $\rho(X)$ は

$$\rho(X) = \frac{\text{var}(T)}{\text{var}(X)} = \frac{\sigma_T^2}{\sigma_X^2} \quad (23)$$

と定義される。またこれは、(6) 式を用いると

$$\rho(X) = 1 - \frac{\text{var}(E)}{\text{var}(X)} = 1 - \frac{\sigma_E^2}{\sigma_X^2} \quad (24)$$

と表現される。いずれにしても信頼性係数は

$$0 \leq \rho(X) \leq 1 \quad (25)$$

を満たす。

$$\sigma_E = \sqrt{1 - \rho(X)} \times \sigma_X \quad (21)$$

5

$$\text{cov}(X, T) = \text{cov}(T + E, T) = \text{cov}(T, T) + \text{cov}(E, T) = \text{var}(T) + 0 = \sigma_T^2 \quad (26)$$

したがって、

$$\text{corr}(X, T) = \frac{\text{cov}(X, T)}{\sqrt{\text{var}(X)}\sqrt{\text{var}(T)}} = \frac{\sigma_T^2}{\sqrt{\sigma_X^2}\sqrt{\sigma_T^2}} = \frac{\sqrt{\sigma_T^2}}{\sqrt{\sigma_X^2}} = \sqrt{\frac{\sigma_T^2}{\sigma_X^2}} \quad (27)$$

であるため、

$$\rho(X) = \text{corr}(X, T)^2 \quad (28)$$

とも書ける。

X を T に線形に回帰したときの回帰式 付録 A.2

$$R^2 = \text{corr}(X, T)^2 = \rho(X) \quad \text{分散説明率} \quad (32)$$

6

複数のテスト

m 個のテスト

$$X_j = T_j + E_j \quad (33)$$

$$\mathcal{E}(E_j) = 0 \quad \text{and} \quad \text{var}(E_j) = \sigma_{E_j}^2 \quad (34)$$

$$\mathcal{E}(X_j) = \mathcal{E}(T_j) = \mu_j \quad (35)$$

$$\text{var}(X_j) = \text{var}(T_j) + \text{var}(E_j) \quad (36)$$

$$\rho(X_j) = \frac{\text{var}(T_j)}{\text{var}(X_j)} = 1 - \frac{\text{var}(E_j)}{\text{var}(X_j)} \quad (37)$$

$$\text{cov}(E_j, E_k) = 0, \quad j \neq k, \quad \text{and} \quad \text{cov}(T_j, E_k) = 0, \quad j, k = 1, 2, \dots, m$$

7

テスト間の共分散

$$\text{cov}(X_j, X_k) = \text{cov}(T_j + E_j, T_k + E_k) \quad (39)$$

$$= \text{cov}(T_j, T_k) + \text{cov}(T_j, E_k) + \text{cov}(E_j, T_k) + \text{cov}(E_j, E_k) \quad (40)$$

$$= \text{cov}(T_j, T_k) \quad (41)$$

観測される得点間の共分散は真値間の共分散に等しい。

$$\text{cov}(X_j, T_k) = \text{cov}(T_j, X_k) = \text{cov}(T_j, T_k) \quad (43)$$

観測される得点と真値の間の共分散は真値間の共分散に等しい。

8

希薄化の修正公式

$$\text{corr}(X_j, T_k) = \frac{\text{cov}(X_j, T_k)}{\sqrt{\text{var}(X_j)\text{var}(T_k)}} \quad (44)$$

$$= \frac{\text{cov}(X_j, X_k)}{\sqrt{\text{var}(X_j)\rho(X_k)\text{var}(X_k)}} \quad (43) \quad (45)$$

$$= \frac{\text{corr}(X_j, X_k)}{\sqrt{\rho(X_k)}} \geq \text{corr}(X_j, X_k) \quad (46) \quad (25)$$

$$\text{corr}(T_j, T_k) = \frac{\text{cov}(T_k, T_j)}{\sqrt{\text{var}(T_j)\text{var}(T_k)}} \quad (41) \quad (47)$$

$$= \frac{\text{std}(X_j)\text{std}(X_k)\text{corr}(X_j, X_k)}{\sqrt{\rho(X_j)\text{std}(X_j)\sqrt{\rho(X_k)\text{std}(X_k)}}} \quad (48)$$

$$= \frac{\text{corr}(X_j, X_k)}{\sqrt{\rho(X_j)\rho(X_k)}} \geq \text{corr}(X_j, X_k) \quad (49) \quad (25)$$

$$\text{corr}(X_j, X_k) \leq \frac{\text{corr}(X_j, T_k)}{\text{corr}(T_j, X_k)} \leq \text{corr}(T_j, T_k) \quad (50)$$

信頼性と妥当性

$$\text{corr}(T_j, T_k) = \frac{\text{corr}(X_j, X_k)}{\sqrt{\rho(X_j)\rho(X_k)}} \quad (49)$$

$$\text{corr}(X_j, X_k) = \text{corr}(T_j, T_k)\sqrt{\rho(X_j)\rho(X_k)} \quad (51)$$

相関係数は 1 を超えない & (25)

$$\text{corr}(X_j, X_k) \leq \sqrt{\rho(X_j)\rho(X_k)} = \sqrt{\rho(X_j)}\sqrt{\rho(X_k)} \leq \sqrt{\rho(X_j)} \quad (52)$$

他のテストとの相関(妥当性)は自分自身(j)の信頼性係数の平方根を超えない。

テストの信頼性は妥当性の必要条件である

合計点に関して

$$\begin{aligned}
 X &= \sum_{j=1}^m X_j \\
 &= \sum_{j=1}^m (T_j + E_j) \\
 &= \sum_{j=1}^m T_j + \sum_{j=1}^m E_j \\
 &= T + E
 \end{aligned}
 \qquad \text{var}(X) = \text{var}(T) + \text{var}(E)$$

$$\mathcal{E}(X) = \sum_{j=1}^m \mu_j \quad \text{and} \quad \mathcal{E}(E) = 0 \tag{239} \tag{57}$$

$$\text{var}(E) = \sum_{j=1}^m \text{var}(E_j) \tag{61}$$

$$= \sum_{j=1}^m \text{var}(X_j)(1 - \rho(X_j)) \tag{21} \tag{62}$$

11

$$\text{var}(X) = \sum_{j=1}^m \sum_{k=1}^m \text{cov}(X_j, X_k) \tag{240} \tag{63}$$

$$= \sum_{j=1}^m \text{var}(X_j) + \sum_{j \neq k} \text{cov}(X_j, X_k) \tag{64}$$

$$= \sum_{j=1}^m \text{var}(X_j) + \sum_{j \neq k} \text{cov}(T_j, T_k) \tag{65}$$

$$\text{var}(T) = \sum_{j=1}^m \sum_{k=1}^m \text{cov}(T_j, T_k) \tag{58}$$

$$= \sum_{j=1}^m \text{var}(T_j) + \sum_{j \neq k} \text{cov}(T_j, T_k) \tag{59}$$

$$= \sum_{j=1}^m \text{var}(X_j)\rho(X_j) + \sum_{j \neq k} \text{cov}(X_j, X_k) \tag{60}$$

合計点の信頼性

$$\rho(X) = \frac{\text{var}(T)}{\text{var}(X)} \tag{67}$$

$$= \frac{\sum_{j=1}^m \text{var}(T_j) + \sum_{j \neq k} \text{cov}(T_j, T_k)}{\text{var}(X)} \tag{68}$$

$$= \frac{\sum_{j=1}^m \text{var}(X_j)\rho(X_j) + \sum_{j \neq k} \text{cov}(X_j, X_k)}{\text{var}(X)} \tag{69}$$

$$\rho(X) = 1 - \frac{\text{var}(E)}{\text{var}(X)} \tag{62} \tag{70}$$

$$= 1 - \frac{\sum_{j=1}^m \text{var}(X_j)(1 - \rho(X_j))}{\text{var}(X)} \tag{71}$$

(71) より、分母が大きいと信頼性は大きくなる。
 そのためには (63) より、テスト間の共分散そして相関が大きければ良い。

内的整合性が高いテストの合計点（合計点）の信頼性係数は高くなる。

内部一貫性

各テストの信頼性が高いことも大事。

合計点と各テストの共分散は

$$\text{cov}(X, X_j) = \text{cov}\left(\sum_{k=1}^m X_k, X_j\right) \tag{229} \tag{72}$$

$$= \sum_{k=1}^m \text{cov}(X_k, X_j) = \sum_{k=1}^m \text{cov}(X_j, X_k) \tag{73}$$

$$\text{var}(X) = \sum_{j=1}^m \sum_{k=1}^m \text{cov}(X_j, X_k) = \sum_{j=1}^m \text{cov}(X, X_j)$$

$$= \sum_{j=1}^m \text{corr}(X, X_j) \sqrt{\text{var}(X)\text{var}(X_j)} \tag{74}$$

各テストと合計点の相関が高いほど合計点の信頼性は高くなる。

テストのモデル

2.4.0.1 平行テスト

いま、これらの m 個のテストの真値が全て等しく、誤差分散も全て等しい

$$T_j = T_* \quad \text{and} \quad \text{var}(E_j) = \sigma_E^2, \quad j = 1, 2, \dots, m \quad (76)$$

と仮定できる場合、それらの m 個のテストを**平行テスト** (**強平行テスト parallel test**) と

$$\mathcal{E}(X_j) = \mathcal{E}(T_* + E_j) = \mathcal{E}(T_*) = \mu \quad (78)$$

$$\text{var}(X_j) = \text{var}(T_* + E_j) = \text{var}(T_*) + \text{var}(E_j) = \sigma_{T_*}^2 + \sigma_E^2 \equiv \sigma_{X_*}^2 \quad (79)$$

$$\text{cov}(X_j, X_k) = \text{cov}(T_j, T_k) = \text{cov}(T_*, T_*) = \text{var}(T_*) = \sigma_{T_*}^2 \quad (80)$$

$$\text{corr}(X_j, X_k) = \frac{\text{cov}(X_j, X_k)}{\sqrt{\text{var}(X_j)\text{var}(X_k)}} = \frac{\sigma_{T_*}^2}{\sigma_{X_*}^2} = \rho^* \quad (81)$$

15

平行テストの合計点

$$\rho(X) = \frac{m\rho^*}{1 + (m-1)\rho^*} \quad \text{Spearman-Brown の公式} \quad (82) \quad (89)$$

$$X = \sum_{i=1}^m X_i = \sum_{i=1}^m (T_* + E_i) = mT_* + \sum_{i=1}^m E_i = T + E \quad (83)$$

$$\text{cov}(E_j, E_k) = 0, \quad j \neq k \quad (38)$$

$$\text{var}(T) = \text{var}(mT_*) = m^2\sigma_{T_*}^2 \quad (84)$$

$$\text{var}(X) = \text{var}(mT_* + \sum_{i=1}^m E_i) = \text{var}(mT_*) + \text{var}(\sum_{i=1}^m E_i) \quad (85)$$

$$= \text{var}(mT_*) + \sum_{i=1}^m \text{var}(E_i) = m^2\sigma_{T_*}^2 + m\sigma_E^2 \quad (86)$$

16

Spearman Brown の公式

$$\text{var}(T) = m^2 \sigma_{T^*}^2 \quad \text{var}(X) = m^2 \sigma_{T^*}^2 + m \sigma_E^2$$

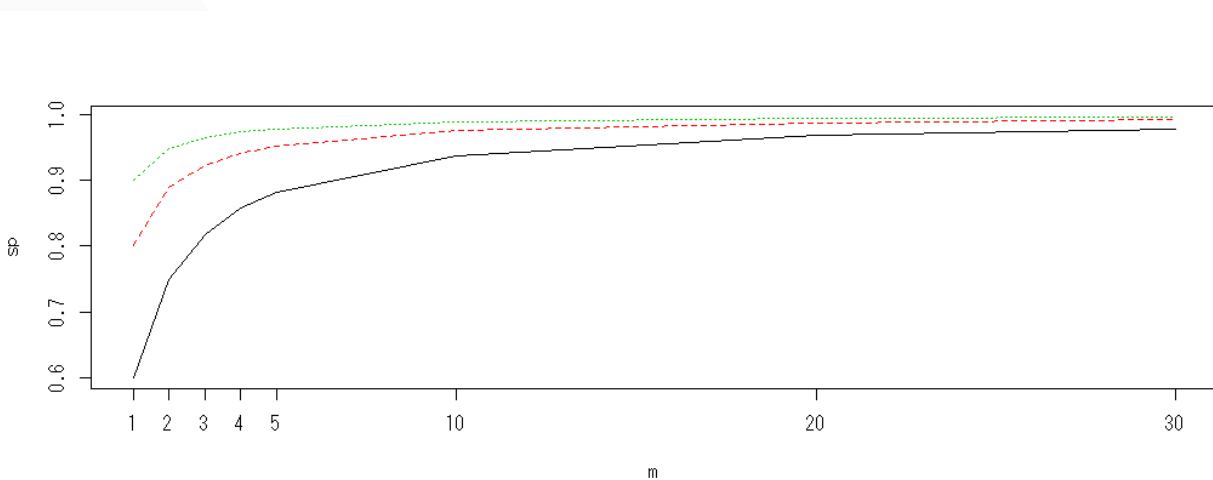
$$\rho(X) = \frac{\text{var}(T)}{\text{var}(X)} = \frac{m^2 \sigma_{T^*}^2}{m^2 \sigma_{T^*}^2 + m \sigma_E^2} \quad (87)$$

$$= \frac{m^2 \sigma_T^2 / \sigma_{X^*}^2}{m^2 \sigma_{T^*}^2 / \sigma_{X^*}^2 + m \sigma_E^2 / \sigma_{X^*}^2} = \frac{m^2 \rho^*}{m^2 \rho^* + m(1 - \rho^*)} \quad (88)$$

$$= \frac{m \rho^*}{m \rho^* + (1 - \rho^*)} = \frac{m \rho^*}{1 + (m - 1) \rho^*} \quad (89)$$

Spearman Brown の公式

$\rho = 0.6, 0.8, 0.9$



タウ τ 等価テスト

2.4.0.2 τ 等価テスト

次に、誤差分散が等しいという仮定を落とし、真値が等しいこと

$$T_j = T_*, j = 1, 2, \dots, m \quad (90)$$

のみを仮定したモデルを τ 等価テスト (tau-equivalent test)、また、真値が定数だけずれているということ

$$T_j = b_j + T_*, j = 1, 2, \dots, m \quad (91)$$

のみを仮定したモデルを本質的 τ 等価テスト (essentially tau-equivalent test) と呼ぶ。この両者は弱平行テストとも呼ばれることがある。本質的 τ 等価テストなテストの場合、

$$\mathcal{E}(X_j) = \mathcal{E}(b_j + T_* + E_j) = \mathcal{E}(b_j + T_*) = b_j + \mu \quad (92)$$

$$\text{var}(X_j) = \text{var}(b_j + T_* + E_j) = \text{var}(T_*) + \text{var}(E_j) = \sigma_{T_*}^2 + \text{var}(E_j) \quad (93)$$

$$\text{cov}(X_j, X_k) = \text{var}(T_*) = \sigma_{T_*}^2, j \neq k \quad (94)$$

$$\text{corr}(X_j, X_k) = \frac{\sigma_{T_*}^2}{\sqrt{\text{var}(X_j)}\sqrt{\text{var}(X_k)}} = \sqrt{\frac{\sigma_{T_*}^2}{\text{var}(X_j)}}\sqrt{\frac{\sigma_{T_*}^2}{\text{var}(X_k)}} \quad (95)$$

$$= \sqrt{\rho(X_j)}\sqrt{\rho(X_k)} = \text{corr}(X_j, T_j)\text{corr}(X_k, T_k) \quad (96)$$

となる。(28) 式を利用。)

$$\rho(X) = \text{corr}(X, T)^2$$

弱平行テストの合計点

$$X = \sum_{i=1}^m X_i = \sum_{i=1}^m (T_* + E_i) = mT_* + \sum_{i=1}^m E_i = T + E \quad (97)$$

$$\text{var}(T) = \text{var}(mT_*) = m^2\sigma_{T_*}^2 \quad (98)$$

$$\text{var}(X_i) = \text{var}(T_*) + \text{var}(E_i) = \sigma_{T_*}^2 + \text{var}(E_i) \quad (99)$$

$$\text{var}(X) = \text{var}(mT_* + \sum_{i=1}^m E_i) = \text{var}(mT_*) + \text{var}(\sum_{i=1}^m E_i) \quad (100)$$

$$= m^2\sigma_{T_*}^2 + \sum_{i=1}^m \text{var}(E_i) \quad (101)$$

(99) 式の両辺の和を取ることで

$$\sum_{i=1}^m \text{var}(X_i) = m\sigma_{T_*}^2 + \sum_{i=1}^m \text{var}(E_i) \quad (102)$$

したがって

$$\sum_{i=1}^m \text{var}(E_i) = \sum_{i=1}^m \text{var}(X_i) - m\sigma_{T_*}^2 \quad (103)$$

$$\text{var}(X) = m^2 \sigma_{T^*}^2 + \sum_{i=1}^m \text{var}(E_i) \quad (101)$$

$$\sum_{i=1}^m \text{var}(E_i) = \sum_{i=1}^m \text{var}(X_i) - m \sigma_{T^*}^2 \quad (103)$$

$$\begin{aligned} &= m^2 \sigma_{T^*}^2 + \sum_{i=1}^m \text{var}(X_i) - m \sigma_{T^*}^2 \\ &= m(m-1) \sigma_{T^*}^2 + \sum_{i=1}^m \text{var}(X_i) \end{aligned} \quad (105)$$

$$\sigma_{T^*}^2 = \frac{\text{var}(X) - \sum_{i=1}^m \text{var}(X_i)}{m(m-1)} \quad (106)$$

観測されない真値の分散が観測されるテスト得点で表現できる！！

21

アルファ係数

$$\sigma_{T^*}^2 = \frac{\text{var}(X) - \sum_{i=1}^m \text{var}(X_i)}{m(m-1)} \quad (106)$$

$$\rho(X) = \frac{\text{var}(T)}{\text{var}(X)} = \frac{m^2 \sigma_{T^*}^2}{\text{var}(X)} \quad (107)$$

$$= \frac{m}{m-1} \frac{\text{var}(X) - \sum_{i=1}^m \text{var}(X_i)}{\text{var}(X)} \quad (108)$$

$$= \frac{m}{m-1} \left(1 - \frac{\sum_{i=1}^m \text{var}(X_i)}{\text{var}(X)} \right) \equiv \alpha(X) \quad (109)$$

Cronbach の α 係数

$$\text{cov}(X_j, X_k) = \text{var}(T_*) = \sigma_{T^*}^2, \quad j \neq k \quad (94)$$

(94) 式が満たされない場合には α 係数は真の信頼性係数の下限

$$\rho(X) = \frac{\text{var}(T)}{\text{var}(X)} \geq \frac{m^2 \sigma_{T^*}^2}{\text{var}(X)} = \alpha(X) \quad (116)$$

同族テスト Congeneric Test == FA

各テストの真値間の相関が 1 の場合、すなわち、

$$T_j = b_j + \lambda_j T_*, j = 1, 2, \dots, m \quad (117)$$

$$X_j = b_j + \lambda_j T_* + E_j \quad (118)$$

$$\mathcal{E}(X_j) = b_j + \lambda_j \mu \quad \text{and} \quad \text{var}(X_j) = \lambda_j^2 \sigma_{T_*}^2 + \text{var}(E_j) \quad (119)$$

$$\text{cov}(X_j, X_k) = \lambda_j \lambda_k \sigma_{T_*}^2 \quad (120)$$

$$\mu = 0, \quad \text{and} \quad \sigma_{T_*}^2 = 1$$

$$\rho(X_j) = \frac{\text{var}(T_j)}{\text{var}(X_j)} = \frac{\lambda_j^2 \sigma_{T_*}^2}{\lambda_j^2 \sigma_{T_*}^2 + \text{var}(E_j)} \quad (123)$$

23

真値のベイズ推定

2.7.0.2 真値の信用区間

真値と誤差の分布に独立な正規分布を仮定した場合、すなわち、

$$X|T \sim N(T, \sigma_E^2) \quad \text{and} \quad T \sim N(\mu, \sigma_T^2) \quad (183)$$

とした場合、 X を観測した時の T の条件付き分布 (事後分布) は、ベイズ統計学における正規分布の母平均に関する議論 (分散が既知の場合で自然共役事前分布を用いた場合) から、

$$T|X \sim N(\mu_{T|X}, \sigma_{T|X}^2) \quad (184)$$

ただし、

$$\sigma_{T|X}^2 = \frac{1}{\frac{1}{\sigma_E^2} + \frac{1}{\sigma_T^2}} = \frac{\sigma_E^2 \sigma_T^2}{\sigma_T^2 + \sigma_E^2} \quad (185)$$

$$\mu_{T|X} = \sigma_{T|X}^2 \left(\frac{X}{\sigma_E^2} + \frac{\mu}{\sigma_T^2} \right) = \sigma_{T|X}^2 \frac{\sigma_T^2 X + \sigma_E^2 \mu}{\sigma_T^2 \sigma_E^2} = \frac{\sigma_T^2 X + \sigma_E^2 \mu}{\sigma_T^2 + \sigma_E^2} \quad (186)$$

となる。したがって、 $X = x$ を観測した時の T の信用区間は

$$x \pm z_{(1-\alpha/2)} \sigma_{T|X} \quad (187)$$

24

真値の回帰推定

$$\mu_{T|X} = \rho(X) X + (1 - \rho(X))\mu \quad (191)$$

$$\sigma_{T|X}^2 = \sigma_X^2 (1 - \rho(X)) \rho(X) = \sigma_E^2 \rho(X) \quad (192)$$

付録 A.2 に示した回帰の公式

$$T \rightarrow V, \quad X \rightarrow U$$

ベイズ推定値

$$\sigma_{T|X}^2 = \frac{1}{\frac{1}{\sigma_E^2} + \frac{1}{\sigma_T^2}} = \frac{\sigma_E^2 \sigma_T^2}{\sigma_T^2 + \sigma_E^2} \quad (185)$$

$$\mu_{T|X} = \sigma_{T|X}^2 \left(\frac{X}{\sigma_E^2} + \frac{\mu}{\sigma_T^2} \right) = \sigma_{T|X}^2 \frac{\sigma_T^2 X + \sigma_E^2 \mu}{\sigma_T^2 \sigma_E^2} = \frac{\sigma_T^2 X + \sigma_E^2 \mu}{\sigma_T^2 + \sigma_E^2} \quad (186)$$

25

IRTにおける項目特性曲線の導出

26

項目反応の数理モデル

- どうして項目特性曲線はロジスティック曲線なのか

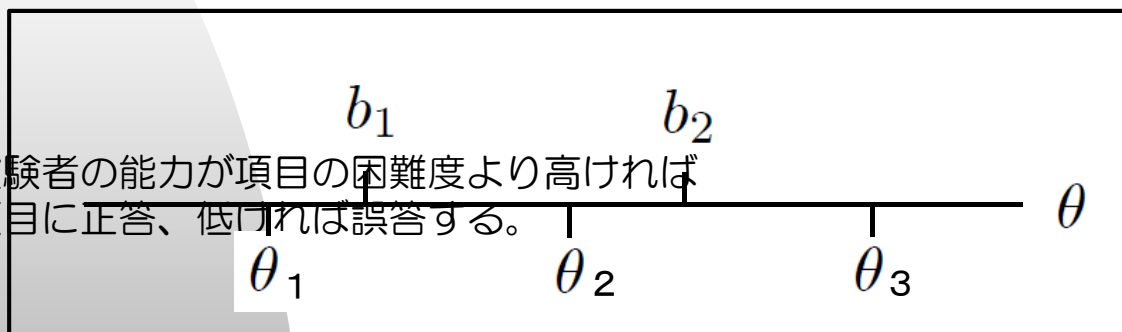
$$P_j(\theta) = \frac{1}{1 + \exp(-1.7a_j(\theta - b_j))}$$

27

Guttman モデル

- 被験者の能力 θ と、項目の困難度 b を同一尺度を用いて表す。 θ -scale

- 被験者の能力が項目の困難度より高ければ項目に正答、低ければ誤答する。



28

Guttman Scale (理想型)

	45	46	47	48	49	50	51	52	53	54	55	56	...
a...	1	1	0	0	0	0	0	0	0	0	0	0	...
b...	1	1	1	0	0	0	0	0	0	0	0	0	...
c...	1	1	1	0	0	0	0	0	0	0	0	0	...
d...	1	1	1	1	0	0	0	0	0	0	0	0	...
e...	1	1	1	1	0	0	0	0	0	0	0	0	...
f...	1	1	1	1	1	1	0	0	0	0	0	0	...
g...	1	1	1	1	1	1	0	0	0	0	0	0	...
h...	1	1	1	1	1	1	1	1	0	0	0	0	...
i...	1	1	1	1	1	1	1	1	1	0	0	0	...
j...	1	1	1	1	1	1	1	1	1	0	0	0	...
k...	1	1	1	1	1	1	1	1	1	0	0	0	...
l...	1	1	1	1	1	1	1	1	1	1	0	0	...
m...	1	1	1	1	1	1	1	1	1	1	1	1	...

実際に得られる項目反応データ

	45	46	47	48	49	50	51	52	53	54	55	56	...
a...	0	1	0	1	1	0	0	0	0	0	0	0	...
b...	1	1	1	0	0	1	0	0	0	0	0	0	...
c...	1	1	1	0	0	1	1	0	0	0	0	0	...
d...	1	1	1	1	1	1	0	1	0	0	0	0	...
e...	1	1	1	1	0	1	0	1	0	0	0	0	...
f...	1	1	1	1	0	1	1	0	0	0	0	0	...
g...	1	1	1	1	1	1	1	0	0	0	0	0	...
h...	1	1	1	1	0	1	1	1	0	0	0	0	...
i...	1	1	1	1	0	1	0	1	0	1	0	0	...
j...	1	1	1	1	1	1	1	0	1	0	0	0	...
k...	1	1	1	1	0	1	0	1	0	0	0	0	...
l...	1	1	1	1	0	1	1	1	1	1	0	0	...
m...	1	0	1	1	1	1	1	0	0	1	1	1	...

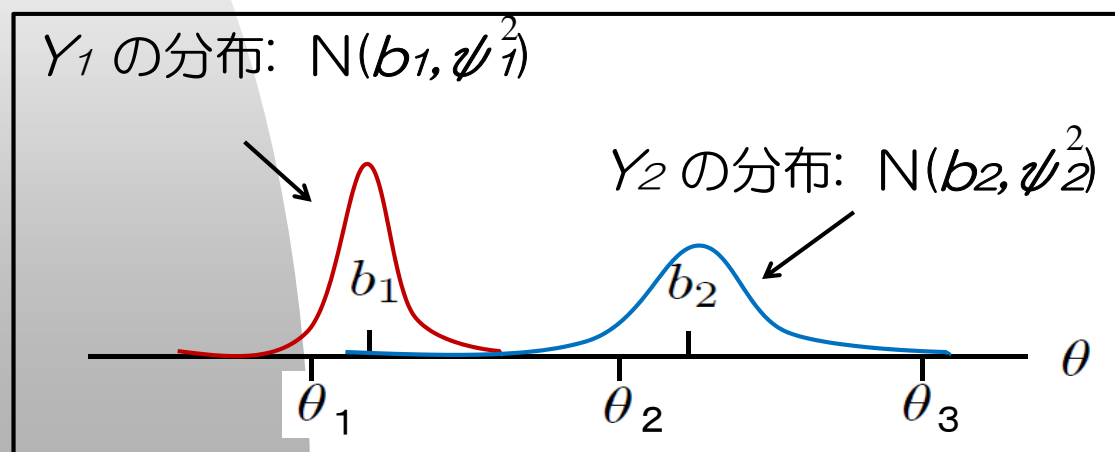
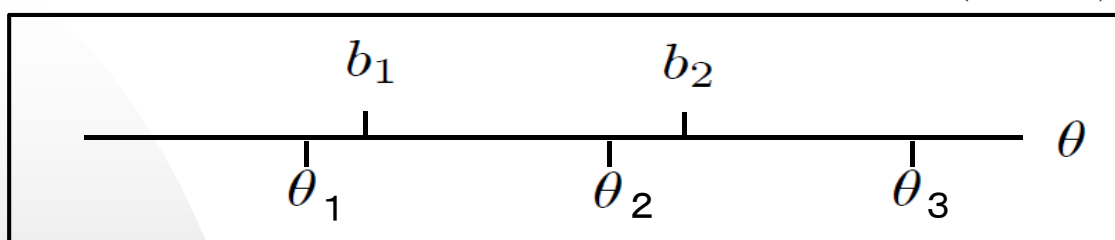
どう頑張っても綺麗に並べられない。

Thurstone の確率モデル

- 被験者の能力 θ と、項目の困難度を同一尺度を用いて表す。 θ -scale
- 項目の困難度 Y は、平均が b 、標準偏差が ψ の正規分布をする。
- 被験者の能力が項目の困難度 Y より高ければ項目に正答、低ければ誤答する。 ψ

Thurstone モデル

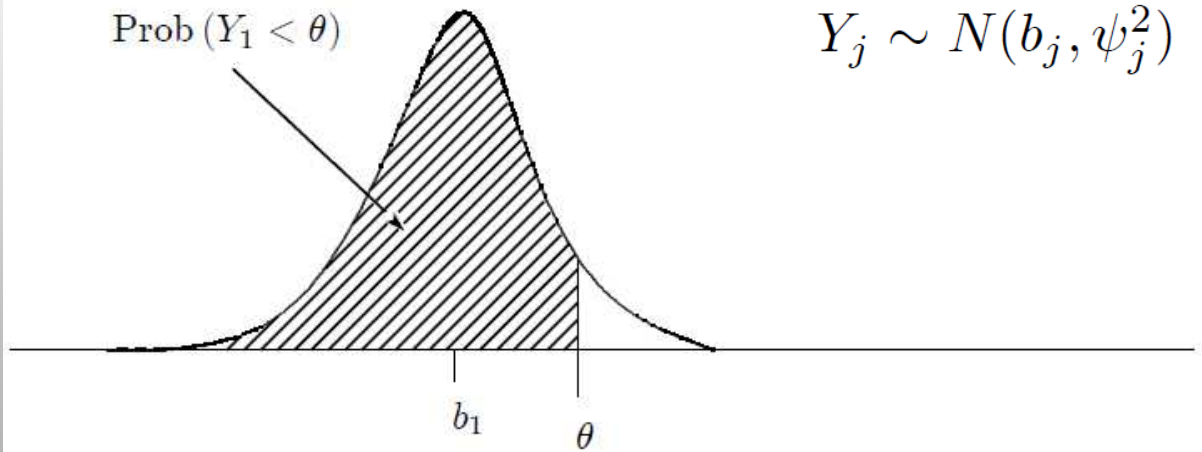
(Guttman)



易しめの項目 $b_1 < b_2$

$\theta \geq Y_j$ ならば $U_j = 1$ (項目 j に正答)

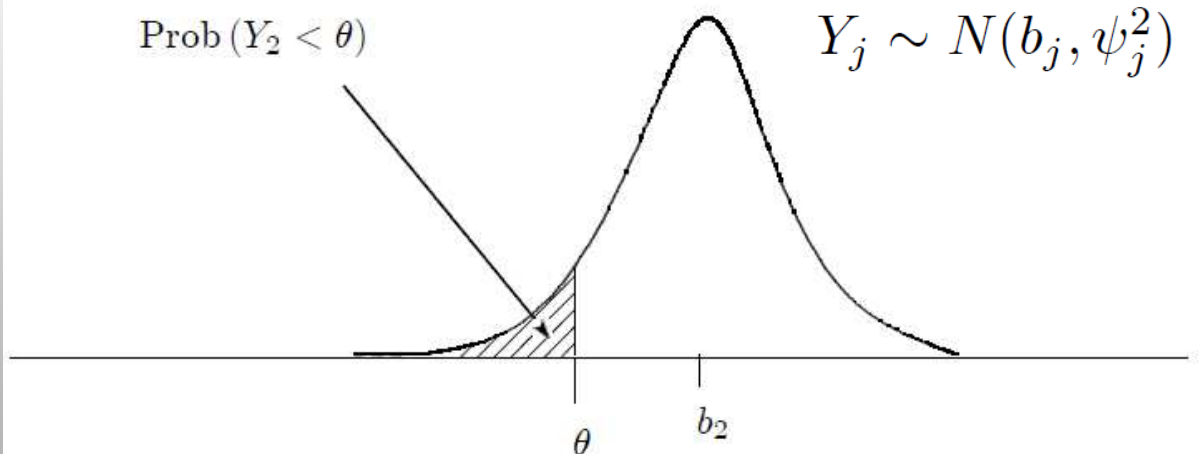
$\theta < Y_j$ ならば $U_j = 0$ (項目 j に誤答)



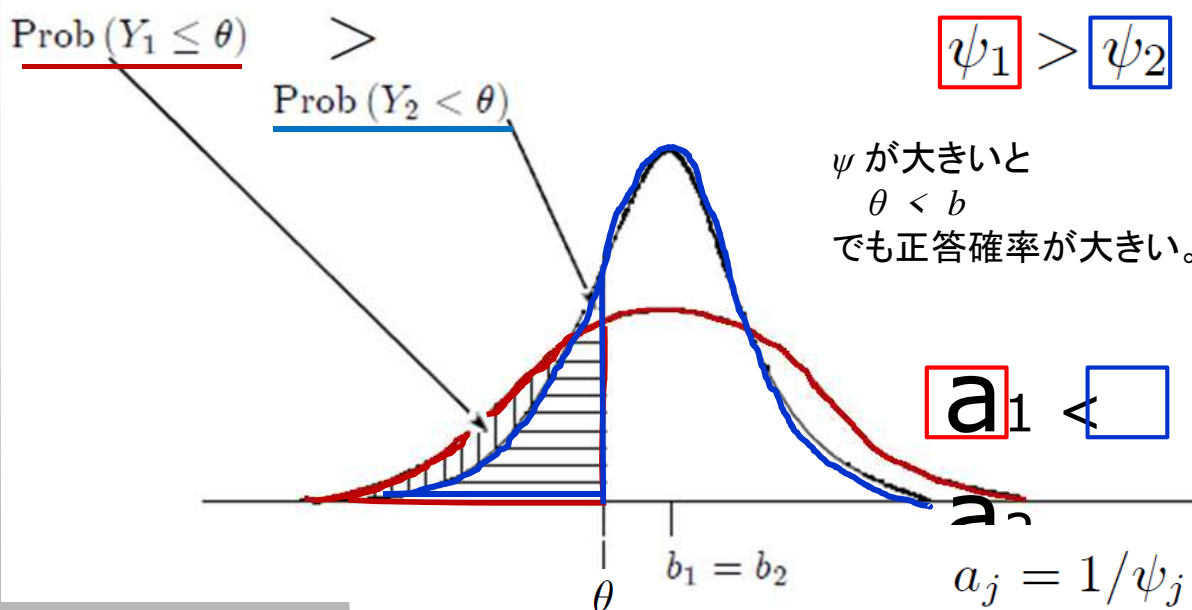
難しめの項目 $b_1 < b_2$

$\theta \geq Y_j$ ならば $U_j = 1$ (項目 j に正答)

$\theta < Y_j$ ならば $U_j = 0$ (項目 j に誤答)

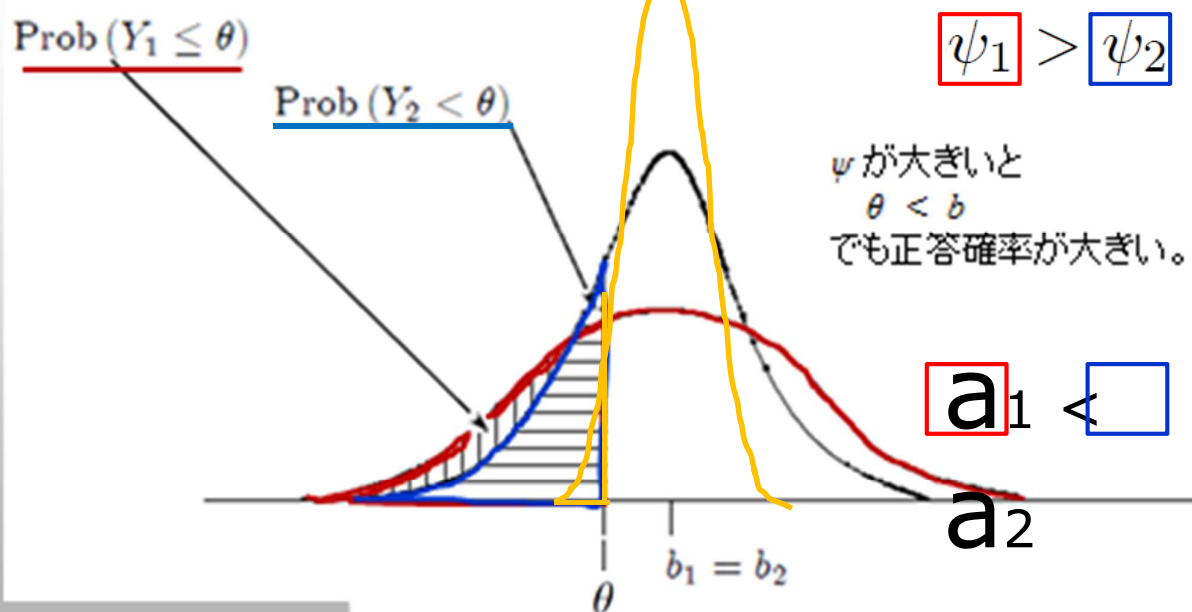


項目の精度



項目の精度

精度がとても良い項目



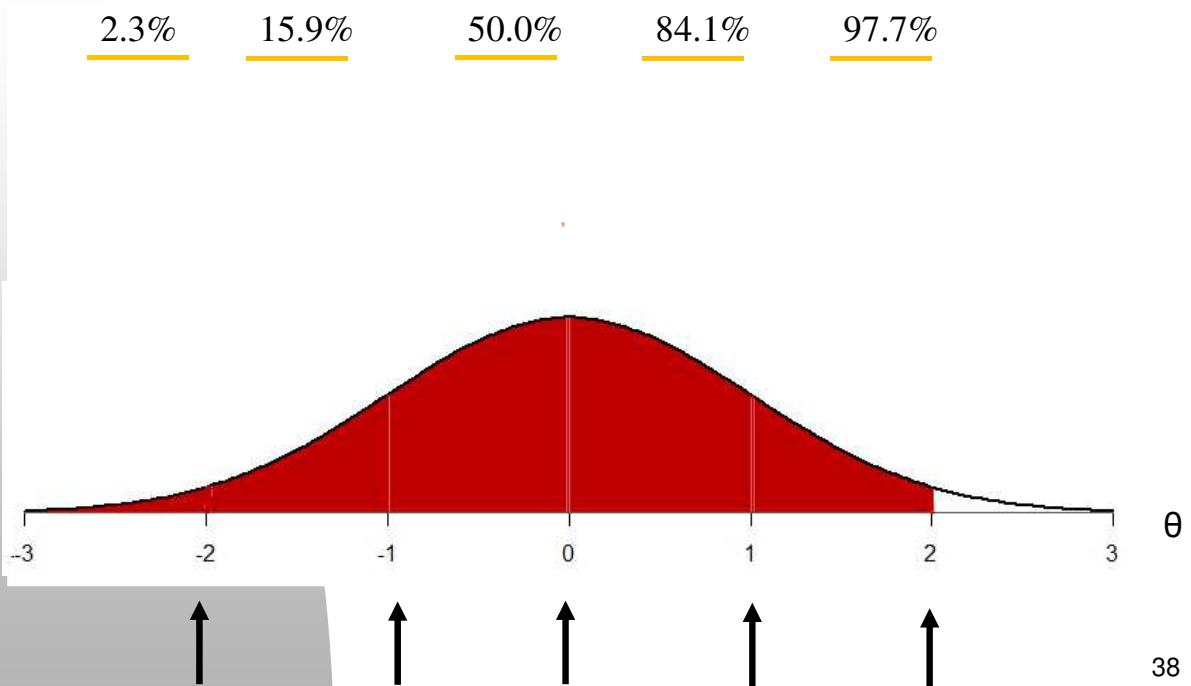
項目特性曲線

$$\text{Prob}(U_j = 1|\theta) = \text{Prob}(Y_j < \theta)$$

能力が θ の人が
項目 j に正答する確率を
 θ の関数として表したものの。

37

$\text{Pr}(Y_j < \theta)$ $Y_j \sim N(0, \psi^2)$
 b, ψ^2

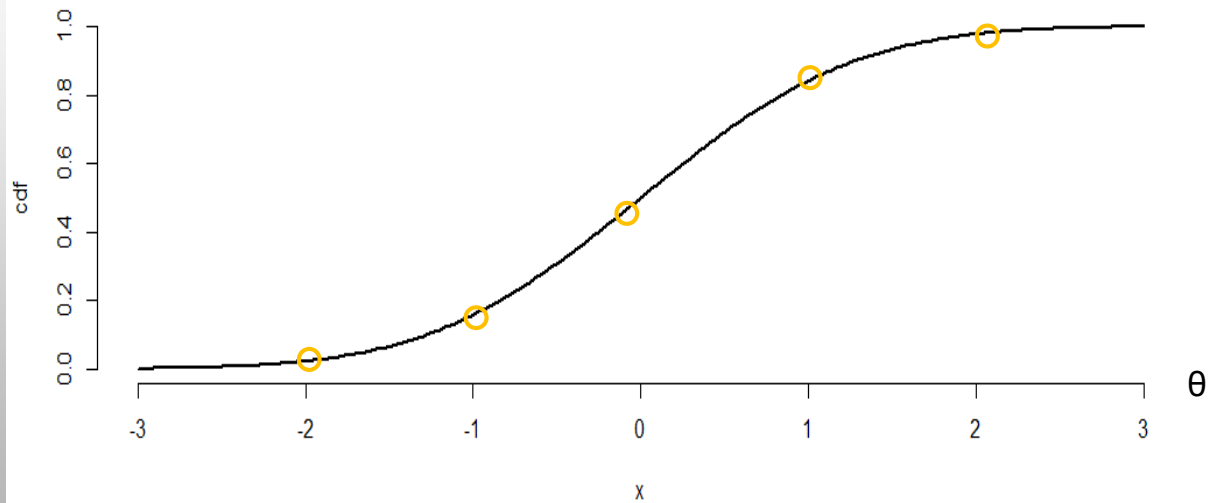


38

$$\Pr(Y_j < \theta)$$

$$Y_j \sim N(0, 1.0)$$

b, ψ^2



39

$$\Pr(Y_j < \theta)$$

$$Y_j \sim N(0, 1.5)$$

2

b, ψ^2

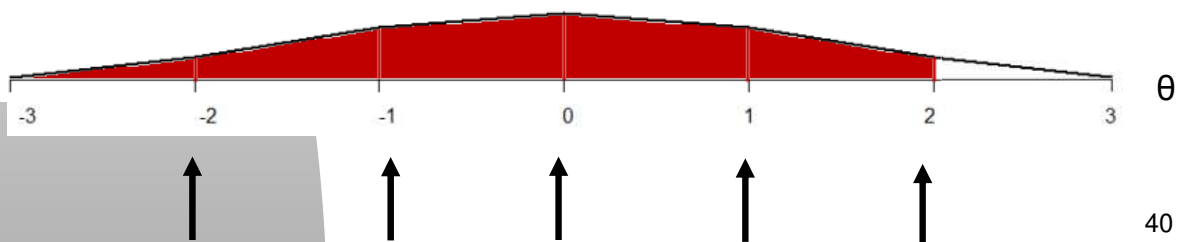
9.1%

25.2%

50.0%

74.8%

90.9%

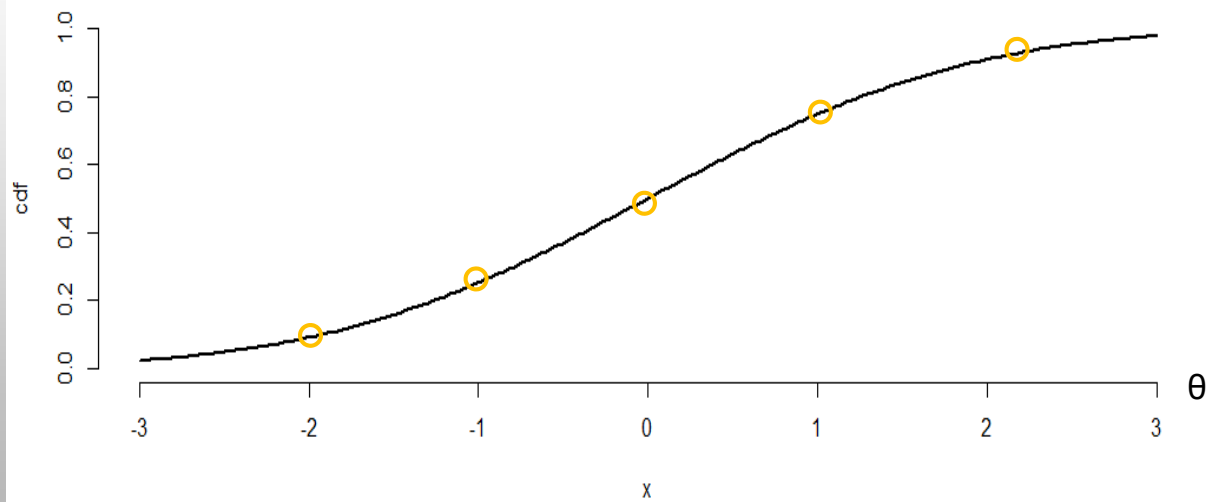


40

$$\Pr(Y_j < \theta)$$

$$Y_j \sim N(0, 1.5^2)$$

b, ψ^2

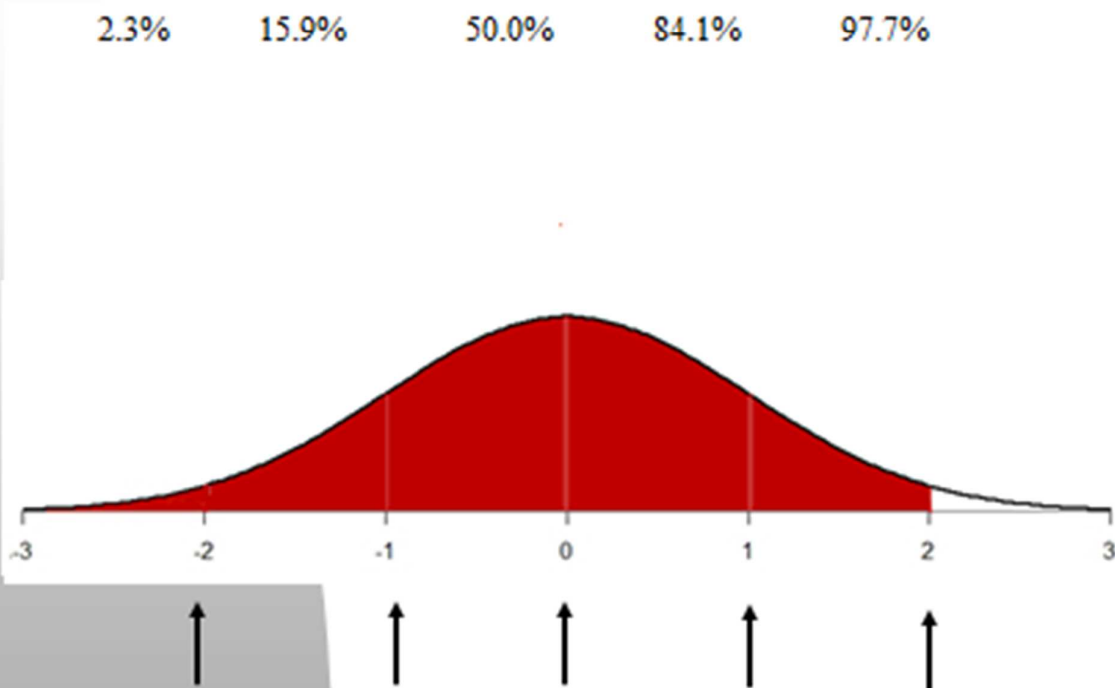


41

$$\Pr(Y_j < \theta)$$

$$Y_j \sim N(0, 1.0)$$

b, ψ^2



$$\Pr(Y_j < \theta)$$

$$Y_j \sim N(0, 0.6^2)$$

2

b, ψ^2

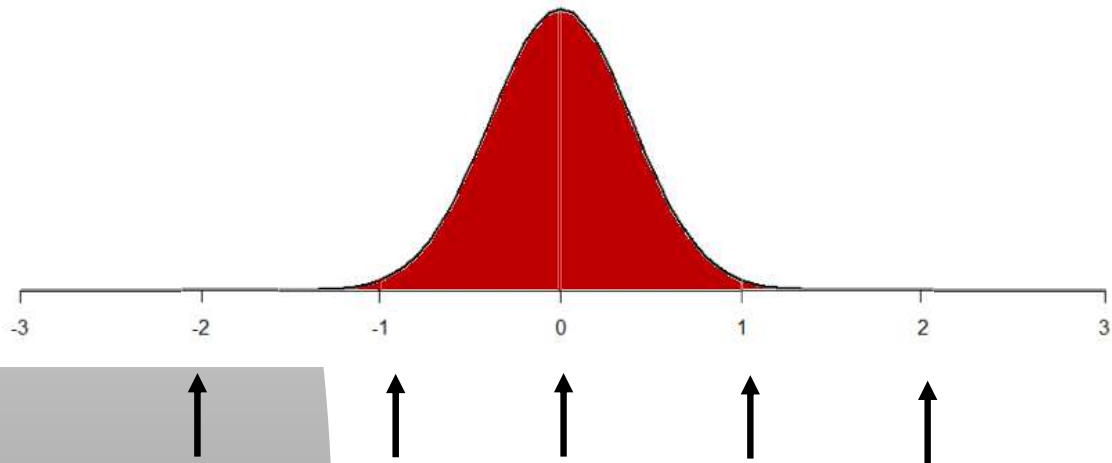
0.00%

2.23%

50.0%

97.73%

99.99%

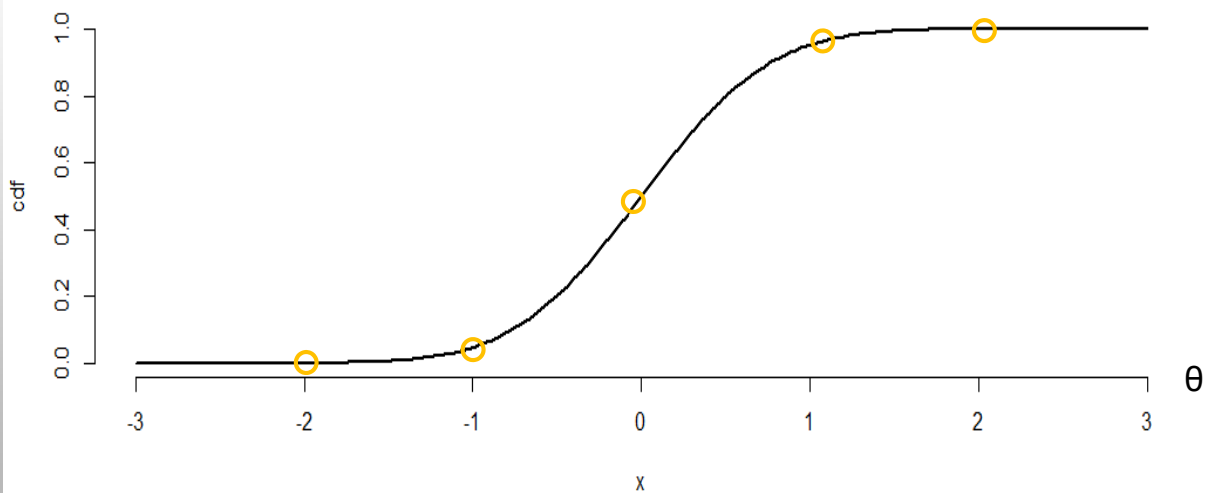


43

$$\Pr(Y_j < \theta)$$

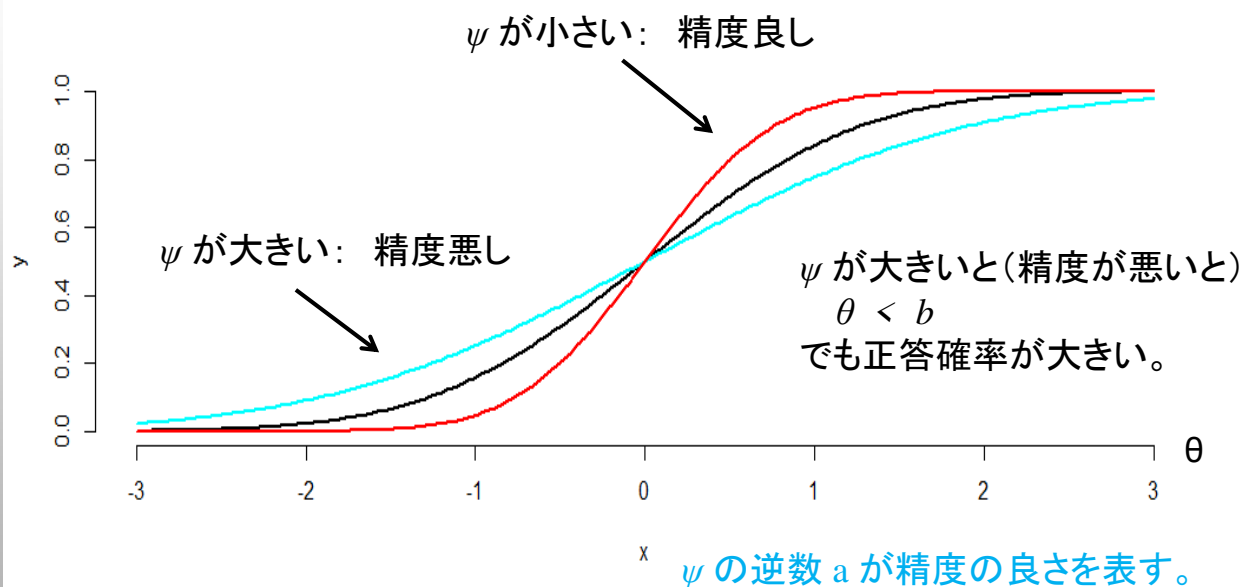
$$Y_j \sim N(0, 0.6^2)$$

b, ψ^2



44

精度の比較 ($b = 0$ の場合)



45

項目特性曲線の導出

$$\text{Prob}(Y_j < \theta) = \text{Prob}\left(\frac{Y_j - b_j}{\psi_j} < \frac{\theta - b_j}{\psi_j}\right)$$

$$Y_j \sim N(b_j, \psi_j^2)$$

$$Z_j = \frac{Y_j - b_j}{\psi_j} \sim N(0, 1)$$

$$= \text{Prob}\left(Z_j < \frac{\theta - b_j}{\psi_j}\right)$$

$$= \Phi\left(\frac{\theta - b_j}{\psi_j}\right)$$

標準正規分布の分布関数 (統計の教科書の付表)

46

項目特性曲線の導出

$$\text{Prob}(Y_j < \theta) = \Phi\left(\frac{\theta - b_j}{\psi_j}\right)$$

項目識別力 $a_j = 1/\psi_j$

$$= \Phi(a_j(\theta - b_j))$$

より簡単なロジスティック関数で
 Φ を近似する。

$$\approx \frac{1}{1 + \exp(-1.7a_j(\theta - b_j))}$$

47

項目特性曲線の導出

$$\text{Prob}(Y_j < \theta)$$

$$\approx P_j(\theta) = \frac{1}{1 + \exp(-1.7a_j(\theta - b_j))}$$

a_j は項目識別力 (Y の標準偏差 ψ の逆数)
 b_j は項目困難度 (Y の平均)

48

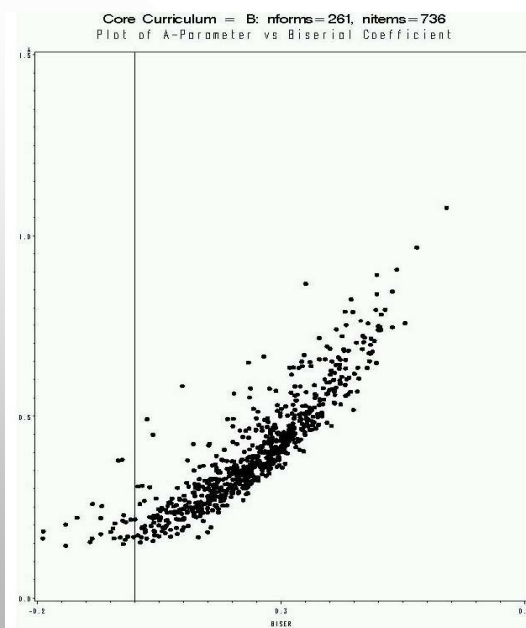
古典的統計量との関係

- 項目通過率（正答率）と項目困難度 b パラメタ
- 項目識別力（識別指数）と a パラメタ
- 合計点と能力値 θ

49

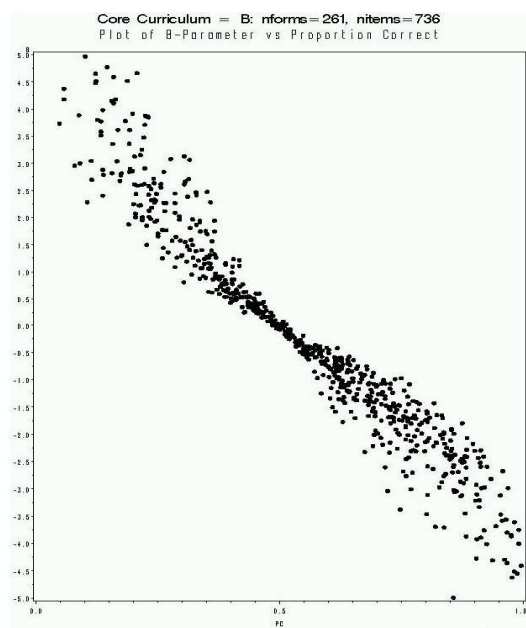
古典的項目統計量との関係

a



項目合計点相関

b

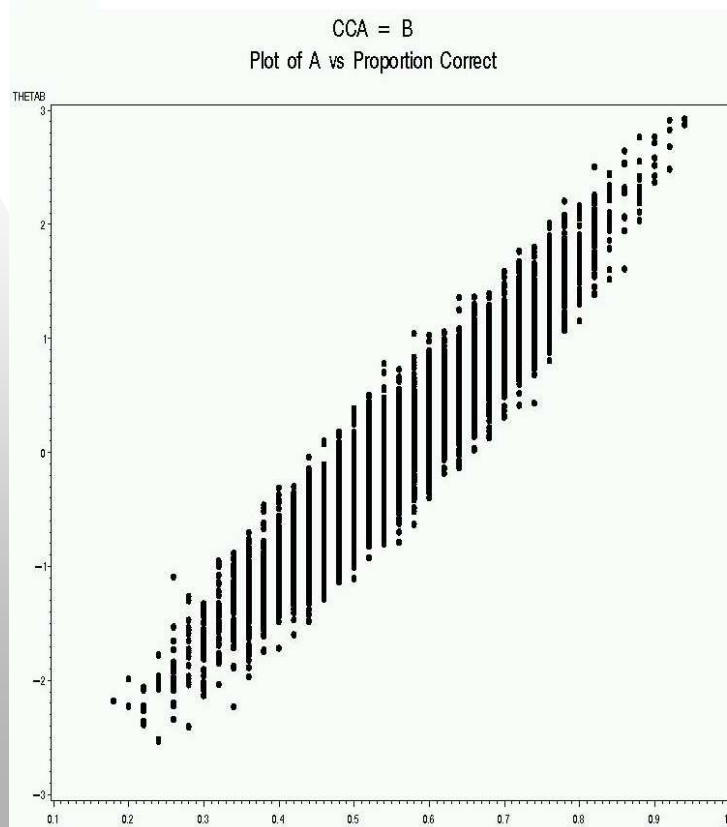


項目通過率

50

正答数得点と θ の関係の例

θ



51

参考文献

- 南風原 (2012) 尺度の作成・使用と妥当性の検討
教育心理学年報 vol. 51 213-217

52

Quiz

Quiz 1 ふたつの確率変数の和の分散の公式

$$\text{var}(X_1 + X_2) = \text{var}(X_1) + \text{var}(X_2) + \text{cov}(X_1, X_2)$$

を証明し、それを3変数に拡張せよ。

ヒントは (251) や (252) 式を使うこと。(A.1.0.4 を使うのであればそれを証明しなければならない。)

また、(251) や (252) を多変量に拡張した公式

$$\mathbf{D}(\mathbf{x}) = \mathbf{E}(\mathbf{x}\mathbf{x}') - \mathbf{E}(\mathbf{x})\mathbf{E}(\mathbf{x})'$$

を証明せよ。

ただし、 $\mathbf{D}(\mathbf{x})$ の対角要素には $\text{var}(X_j)$ が入り、その ij 要素には $\text{cov}(X_i, X_j)$ が入っている。また、 $\mathbf{E}(\mathbf{x})' = \mathbf{E}(\mathbf{x}')$ である。