

Indexing Makes Your Book Perfect

**(and also, the common way of
arranging words in Japanese)**

Keiichiro Shikano

Book editor for Ohmsha, Ltd.

K16.shikano@gmail.com

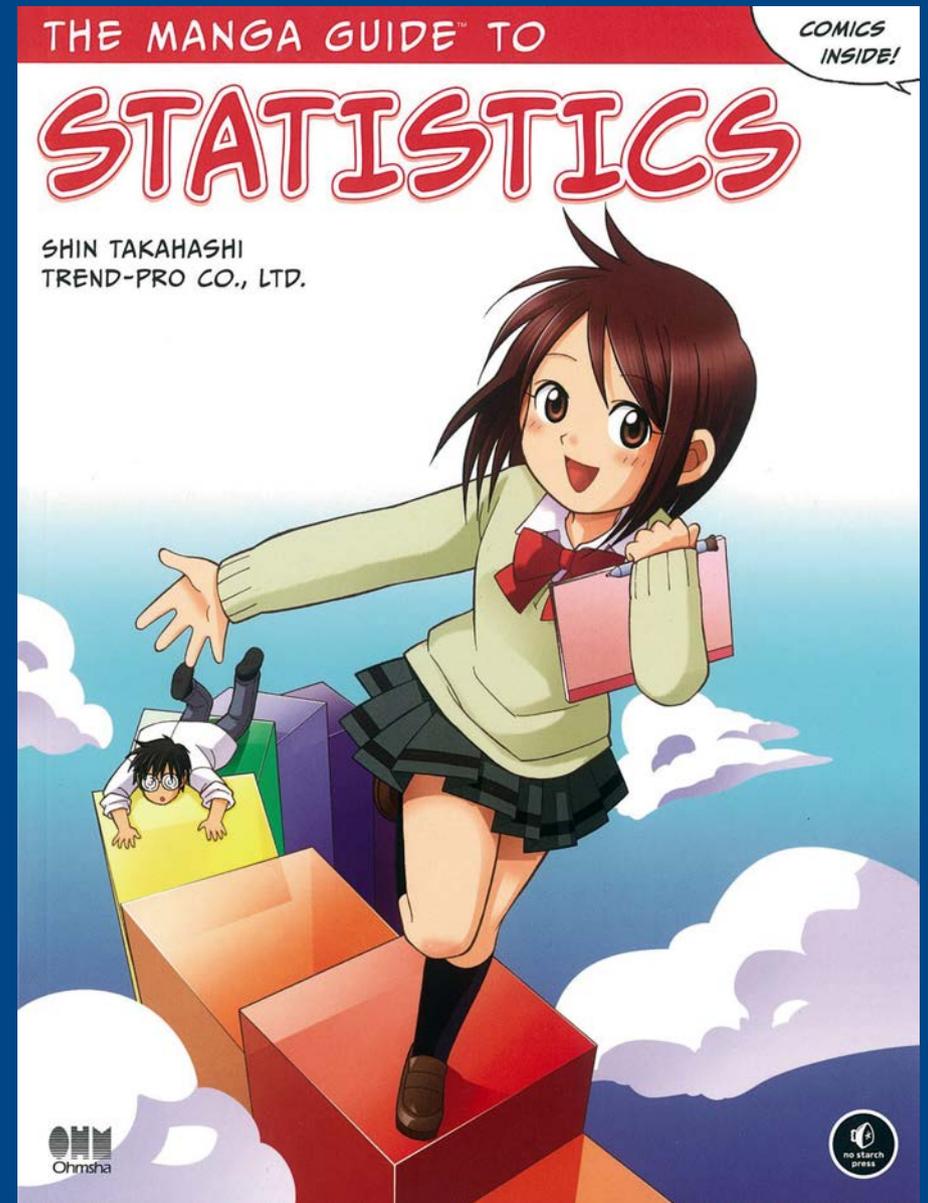
TUG 2013 October 23

**Indexes from
around
the world**

Example

The Manga Guide to Statistics

(in English)



Index in English

INDEX

A

actual measurement frequencies,
130, 131
alpha value (α), 159, 163
alternative hypothesis
accuracy of, 166
considerations, 174
Cramer's coefficient, 186
examples of, 161, 171-173
overview, 170-174
P-value and, 175-179
test of difference between popula-
tion ratios, 173
arithmetic mean, 43, 73, 74
average (mean). See mean
AVERAGE function, 196
average savings, 46-47

C

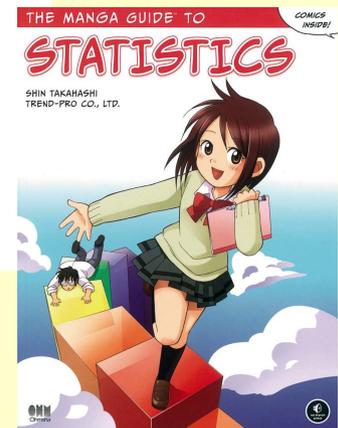
calculations. See Excel calculations
categorical data, 14-29
correlation ratio, 121

CHITEST function, 210-211
class midpoint, 36-39, 54, 56
classes
calculating with Sturges' Rule, 55,
56, 58
intraclass variance, 117, 123,
124, 126
range of, 39, 54-57, 84
coefficient
correlation, 116-120, 206-207
Cramer's. See Cramer's coefficient
CORREL function, 207
correlation, 115, 119
correlation coefficient, 116-120,
206-207
correlation ratio, 117, 121-127, 207
COUNTIF function, 197-198
Cramer's coefficient, 127-138
accuracy of, 147
alternative hypothesis, 186
calculating, 130-135, 141
examples of, 127-136
Excel and, 207

data types, 13-29, 117
degree of relation, 115, 116-120
degrees of freedom (df), 99-108
descriptive statistics, 57-58
deviation, standard, 48-53, 70-79
deviation scores, 74-80, 199-203
df (degrees of freedom), 99-108
distributions
chi-square. See chi-square
distribution
Excel and, 107-109
F, 106-107
normal, 86-91
standard normal, 89-98, 204-205
t, 106

E

estimation theory, 57-58
Euler's number, 86
Excel calculations, 191-211
chi-square distribution, 205-206
correlation coefficient, 206-207



How to use index

INDEX

A

actual measurement frequencies, 130, 131
alpha value (α), 159, 163
alternative hypothesis
accuracy of, 166
considerations, 174
Cramer's coefficient, 186
examples of, 161, 171-173
overview, 170-174
P-value and, 175-179
test of difference between population ratios, 173
arithmetic mean, 43, 73, 74
average (mean). See mean
AVERAGE function, 196
average savings, 46-47

C

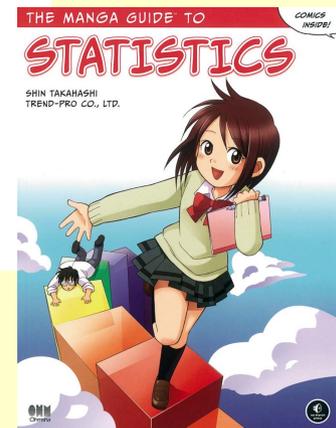
calculations. See Excel calculations
categorical data, 14-29
correlation ratio, 121

CHITEST function, 210-211
class midpoint, 36-39, 54, 56
classes
calculating with Sturges' Rule, 55, 56, 58
intraclass variance, 117, 123, 124, 126
range of, 39, 54-57, 84
coefficient
correlation, 116-120, 206-207
Cramer's. See Cramer's coefficient
CORREL function, 207
correlation, 115, 119
correlation coefficient, 116-120, 206-207
correlation ratio, 117, 121-127, 207
COUNTIF function, 197-198
Cramer's coefficient, 127-138
accuracy of, 147
alternative hypothesis, 186
calculating, 130-135, 141
examples of, 127-136
Excel and, 207

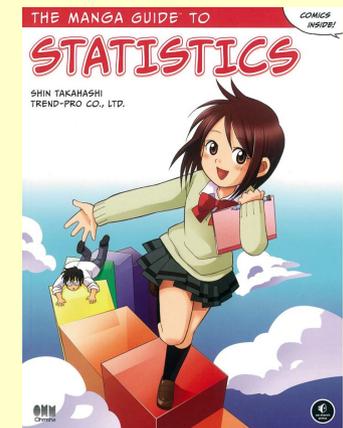
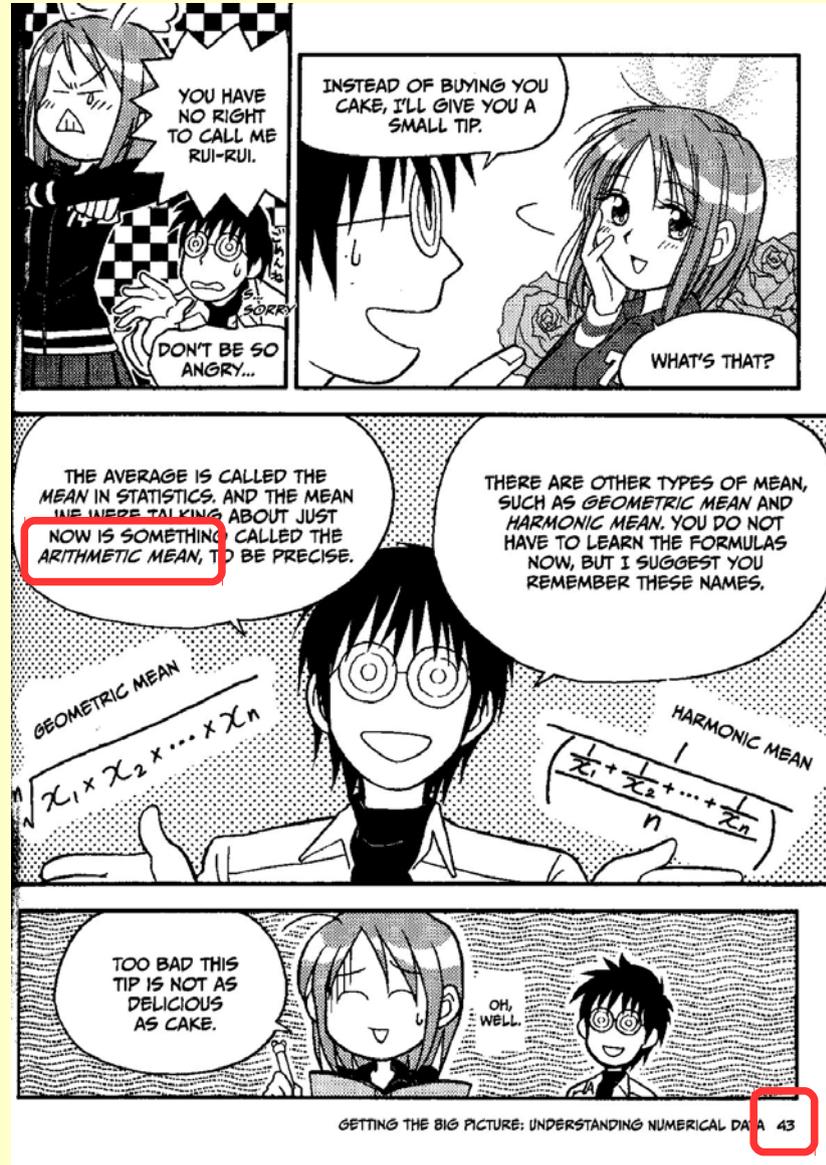
data types, 13-29, 117
degree of relation, 115, 116-120
degrees of freedom (df), 99-108
descriptive statistics, 57-58
deviation, standard, 48-53, 70-79
deviation scores, 74-80, 199-203
df (degrees of freedom), 99-108
distributions
chi-square. See chi-square distribution
Excel and, 107-109
F, 106-107
normal, 86-91
standard normal, 89-98, 204-205
t, 106

E

estimation theory, 57-58
Euler's number, 86
Excel calculations, 191-211
chi-square distribution, 205-206
correlation coefficient, 206-207

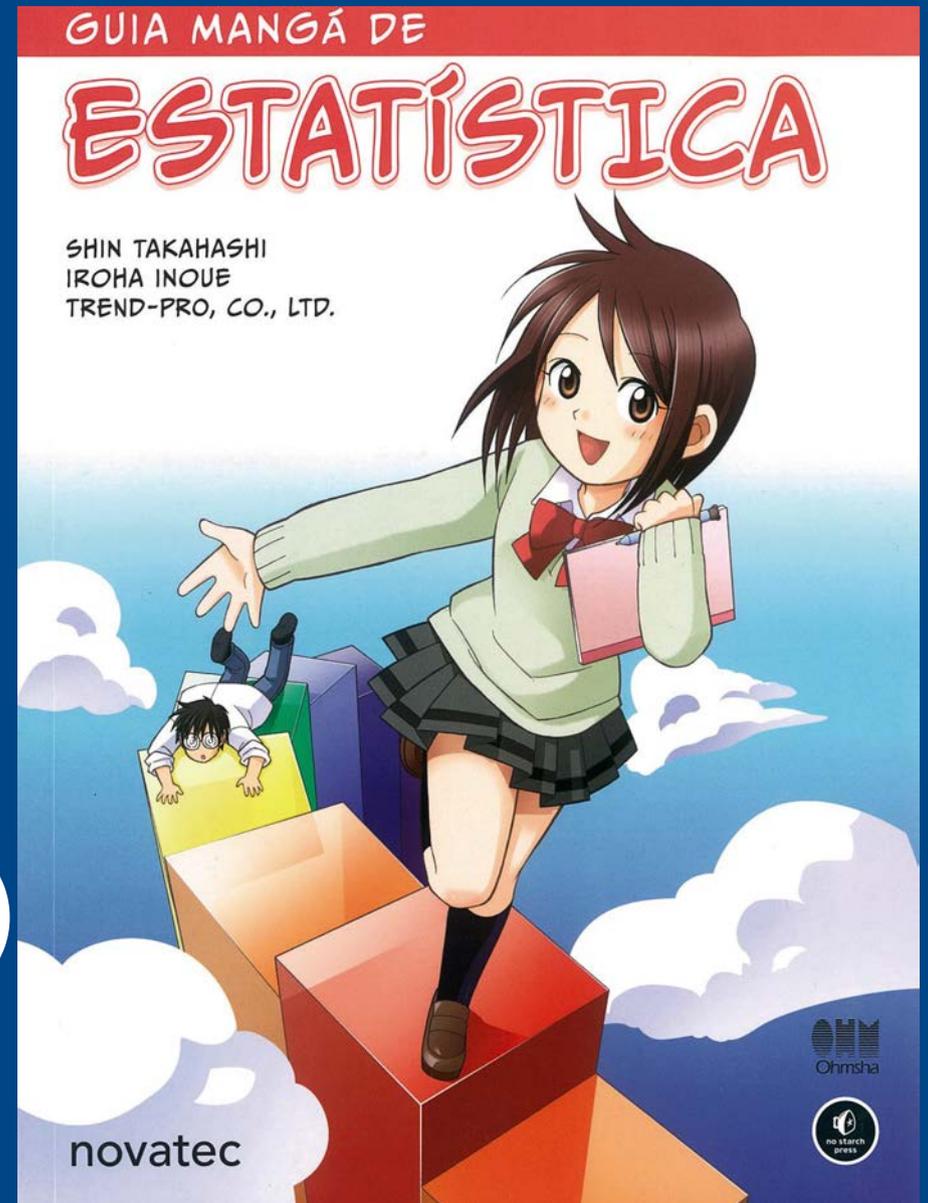


How to use index



Example

**The
Manga Guide
to Statistics
(in Portuguese)**



Index in Portuguese

ÍNDICE

A

alfa (α), valor de, 159, 163
amostras, 6, 7, 52, 57
arquivos do Excel, download, 192

C

Cálculos do Excel, 191-211
coeficiente de correlação,
206-207
desvio-padrão, 195-196
distribuição normal padrão,
204-205
distribuição qui-quadrado,
205-206
distribuições e, 107-109
mediana, 195-196
média, 195-196
padrões de desvio, 74-80,
199-203
tabelas de frequência, 192, 195

correlação negativa, 119
correlação positiva, 119
correlação zero, 119

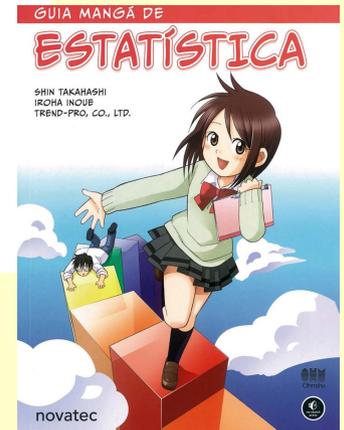
D

dados
conjunto de, 186
inadequados para o coeficiente de
correlação, 120
incomensuráveis. Ver dados
categóricos
qualitativos. Ver dados categóricos
quantitativos. Ver dados numéri-
cos
que não podem ser medidos. Ver
dados categóricos
“dispersão de”, 49, 58, 69, 70, 80
dados categóricos, 14-29
como resultado de pesquisa,
60-64

dispersão de dados, 49, 58, 69, 70,
80
distribuição
Excel, 107-109
F, 106-107
normal, 86-91
padrão normal 89-98, 204-205
t, 106
distribuição qui-quadrado, 99-105
cálculo, 130-133
descrição, 99
exemplos de, 99-105, 152
graus de liberdade, 99-108
pontos no eixo horizontal,
205-206

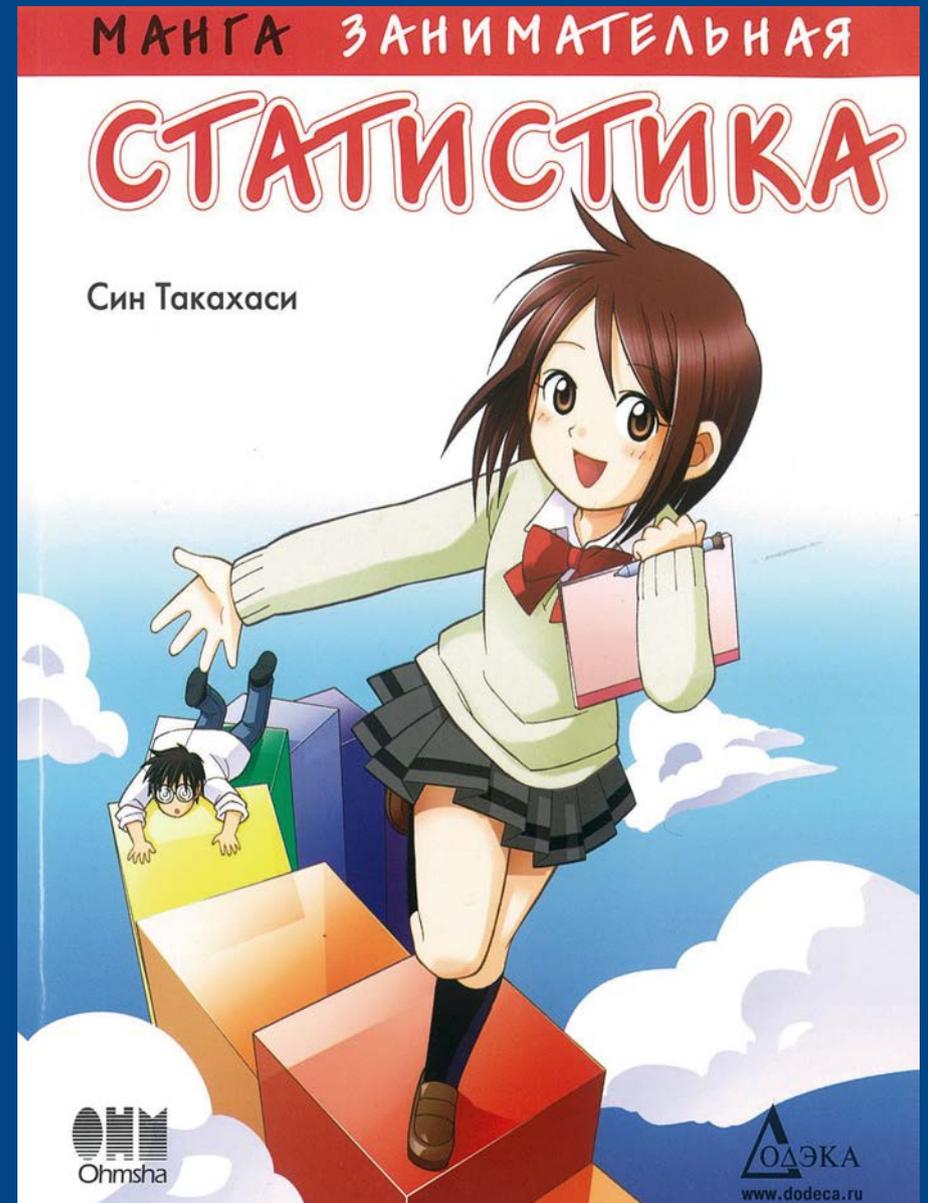
E

eixo horizontal, 39, 102, 107, 109,
125
cálculo de pontos em, 107



Example

**The
Manga Guide
to Statistics
(in Russian)**



Index in Russian

Latin
alphabet

Предметный указатель

Russian
alphabet

α

α , 163, 166

A

AVERAGE, функция, 196

C

CHIDIST, функция, 107

CHIINV, функция, 107, 205-206

CHITEST, функция, 210-211

CORREL, функция, 207

COUNTIF, функция, 197-198

E

Eiken, экзамен, 23-25

F

FDIST, функция, 107

FINV, функция, 107

FREQUENCY, функция, 193-194

FRACП, функция, 107

FRACПОБР, функция, 107

F-распределение, 106-107

M

Microsoft Excel, см. *вычисления в Excel, функции Excel*

N

NORMDIST, функция, 107

A

альтернативная гипотеза

P-значение, 175-179

коэффициент корреляции Крамера, 186

обзор, 170-174

определение, 174

примеры, 161, 171-173

проверка гипотезы о равенстве долей в совокупностях, 173

анкетирование, 4-6

качественные данные, 60-64

ограничения, 4-7

проверка независимости, 137, 208-211

таблицы распределения, 62-64

B

вероятность, 81-109

F-распределение, 106-107

нормальное распределение, 86-89

определение, 82

распределение и Excel, 107-109

распределение Стьюдента, 106

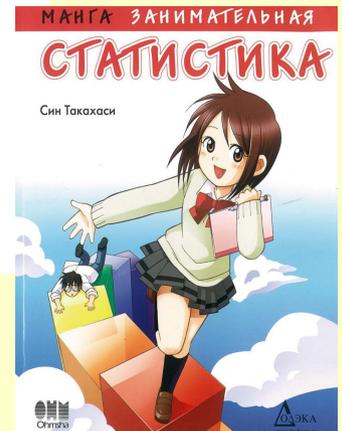
распределение хи-квадрат, 99-105, 205-206

результаты теста, 83-84

соответствующая, 104

стандартное нормальное распределение, 89-98, 204-205

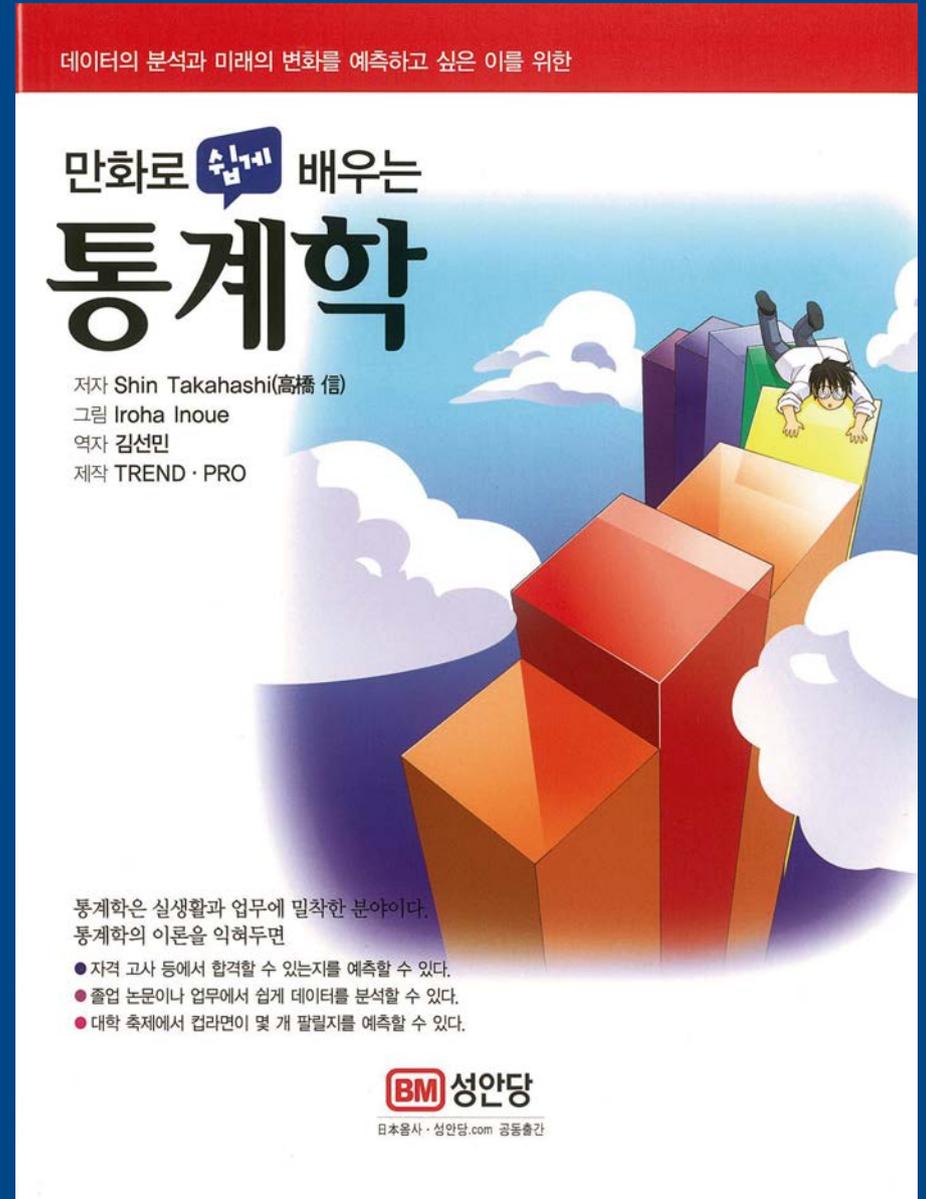
степень свободы, 99-108



Example

The Manga Guide to Statistics

(in Korean)



Index in Korean

Index **찾·아·보·기**

Latin
alphabet

AVERAGE

196

TDIST

107

CHIDIST

107

TINV

107

CHIINV

206

검정

149

CHITEST

210

검정통계량

128

CORREL

207

계급

35

F분포

106

계급값

36

FDIST

107

귀무가설

150, 161, 170

FINV

107

급내변동

123

FREQUENCY

193

기각역

150, 159

기대도수

130

Hangul
alphabet



Index in Korean

동일성의 검정	184
모비율 차의 검정	149
모집단	4, 6
모평균 차의 검정	149
무상관	119
무상관의 검정	149
산술평균	43
상가분포	43
상가평균	43
상관	119
상관비	117, 121

산술 평균
sansul pyeong-gyun

arithmetic
mean

Word group beginning
with consonant letter “人”



Example

The Manga Guide to Statistics

(in Thai)



Index in Thai

Latin
alphabet



Thai
script

A	
AVERAGE	202
C	
CHIDIST	113
CHIINV	212
CHITEST	216
CORREL	213
Cramer's V	135
F	
FDIST	113

ก	
กลุ่มประชากร (population)	10, 12
การแจกแจงแบบไคกำลังสอง (chi-square distribution)	105
การแจกแจงที (T distribution)	112
การแจกแจงแบบปกติมาตรฐาน (standard normal distribution)	95
การแจกแจงปกติ (normal distribution)	92, 94
การแจกแจงเอฟ (F distribution)	112
การทดสอบ (test)	155
การทดสอบความต่างของค่าเฉลี่ยของ ประชากร	155



Index in Thai

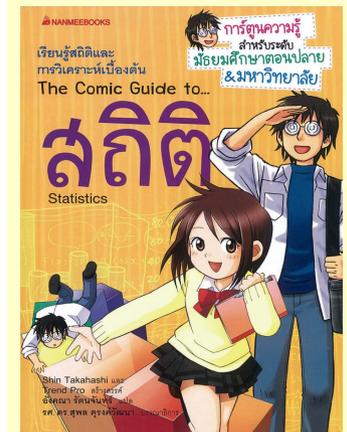
- The index is arranged as the same manner as the alphabetical languages, although Thai is not technically an alphabet.

Word group beginning with letter “ค”

ค่าเฉลี่ยเลขคณิต
Khā chelīy lekḥ

ข	
ขอบเขตวิกฤต (critical region)	156, 165
ข้อมูลเชิงคุณภาพ (qualitative data)	25
ข้อมูลเชิงปริมาณ (quantitative data)	25
ค	
ความถี่ (frequency)	42, 143 157
ความถี่คาดหวัง (expected frequency)	136
ค่า P-value	181
ค่าเฉลี่ย (mean)	47
ค่าเฉลี่ยเลขคณิต (arithmetic mean)	49
ค่าเฉลี่ยฮาร์มอนิก (harmonic mean)	49
ค่าเบี่ยงเบน (deviation Score)	72, 80
ค่าเบี่ยงเบนมาตรฐาน (standard deviation)	55
ค่ามาตรฐาน, คะแนนมาตรฐาน (standard score)	78
ค่าสถิติไคกำลังสองของเพียร์สัน (Pearson's statistical chi-square)	138
ด	
ตัวอย่าง (sample)	10, 12
ตารางแจกแจงความถี่ (frequency distribution table)	42

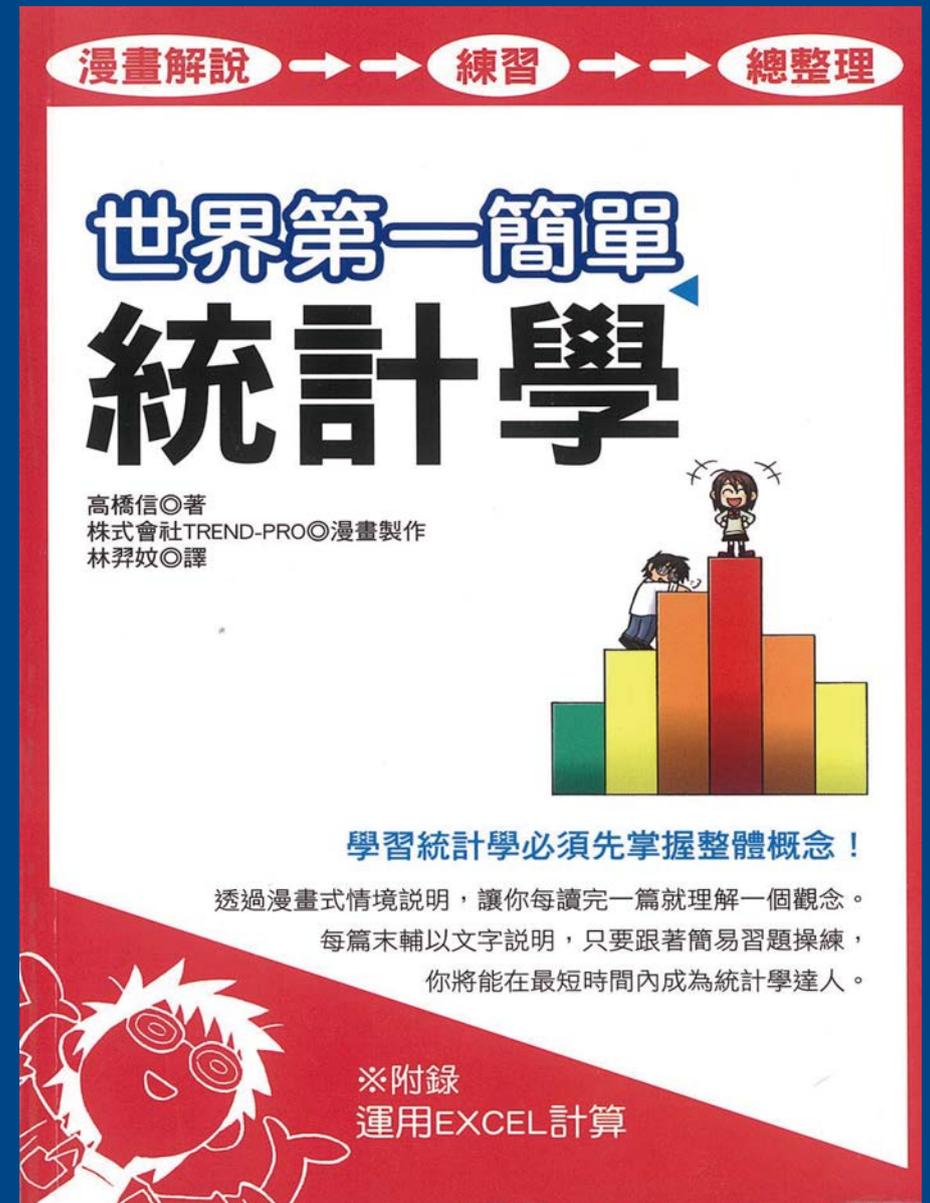
arithmetic mean



Example

The Manga Guide to Statistics

(in traditional Chinese)



Index in Chinese (Traditional)



A

AVERAGE 196

C

CHIDIST 107
CHIINV 206
CORREL 207
CHITEST 210

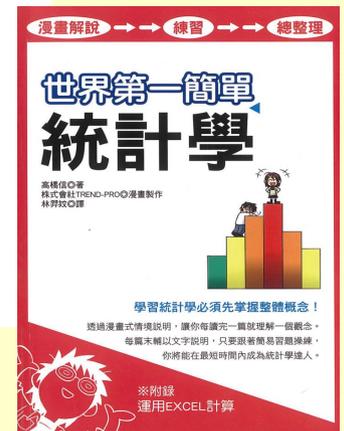
NORMSINV 107

P

P 值 175

S

STDEVP 197

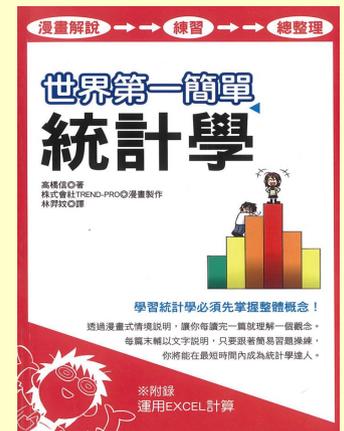


Index in Chinese (Traditional)

Stroke counts of
算 is fourteen.



四劃	← 4 strokes
中位數 ······ 44	
五劃	← 5 strokes
母體 ······ 4, 6	
平均數 ······ 41	
史特吉斯公式 ······ 55	
卡方分配 ······ 99	
卡方分配表 ······ 103	
虛無假說 ······ 150, 161, 170	
十四劃	← 14 strokes
算術平均數 ······ 43	arithmetic mean
對立假說 ······ 150, 161, 170	
齊一性檢定 ······ 184	
十五劃	← 15 strokes
樣本 ······ 4, 6	
數值資料 ······ 10	



Pinyin based order

算術平均數 (for “arithmetic mean”)



Romanize

Suànshù píngjūn shù

You can find the term in
the “S” section.

Example

The Manga Guide to Statistics

(in simplified
Chinese)



Example

**The
Manga Guide
to Statistics
(in Japanese)**



Index in Japanese

索引

A
AVERAGE 196

C
CHIDIST 107
CHIINV 206
CHITEST 210
CORREL 207

F
F分布 106

S
STDEVP 197

T
t分布 106
TDIST 107
TINV 107

カ行
階級 35
階級値 36



Index in Japanese

- The index is arranged in gojuon order, which is roughly based on the pronunciation.

TDIST.....	107
TINV.....	107
カ行	
階級.....	35
階級値.....	36
カイ二乗分布.....	99
カイ二乗検定.....	151
カイ二乗分布表.....	103
確率密度関数.....	85
カテゴリカルデータ.....	19
カテゴリーデータ.....	19
検定.....	149
検定統計量.....	150, 164
サ行	
算術平均.....	43
自然対数の底.....	86
実測度数.....	130
自由度.....	99, 101

← **か** to **こ** of gojuon

← **さ** to **ぞ** of gojuon

arithmetic mean



Kana based order

算術平均 (for “arithmetic mean”)



Turn into kana (syllables)

さんじゅつへいきん

(“sanjutsuheikin” in romanization)

You can find the term in the “さ” section of gojuon, instead of the “S” section.

Gojuon order

1 あ a	2 い i	3 う u	4 え e	5 お o										
6 か ka	7 が ga	8 き ki	9 ぎ gi	10 く ku	11 ぐ gu	12 け ke	13 げ ge	14 こ ko	15 ご go					
16 さ sa	17 ざ za	18 し si	19 じ ji	20 す su	21 ず zu	22 せ se	23 ぜ ze	24 そ so	25 ぞ zo					
26 た ta	27 だ da	28 ち ti	29 ぢ di	30 つ tu	31 づ du	32 て te	33 で de	34 と to	35 ど do					
36 な na	37 に ni	38 ぬ nu	39 ね ne	40 の no										
41 は ha	42 ば ba	43 ぱ pa	44 ひ hi	45 び bi	46 ぴ pi	47 ふ fu	48 ぶ bu	49 ぷ pu	50 へ he	51 べ be	52 ぺ pe	53 ほ ho	54 ぼ bo	55 ぽ po
56 ま ma	57 み mi	58 む mu	59 め me	60 も mo										
61 や ya	62 ゆ yu	63 よ yo		64 ら ra	65 り ri	66 る ru	67 れ re	68 ろ ro		69 わ wa	70 を wo	71 ん nn		

- Red ones can be used as the header for the group containing followings. (Light-weight ones often omitted.)
- Numbered only as a guide. There is more syllables.

Example against gojuon

according to the gojuon order ...

but often ...

久保田五郎 (くぼたごろう)
ku bo ta go ro u

窪田三郎 (くぼたさぶろう)
ku bo ta sa bu ro u

久保田四郎 (くぼたしろう)
ku bo ta shi ro u

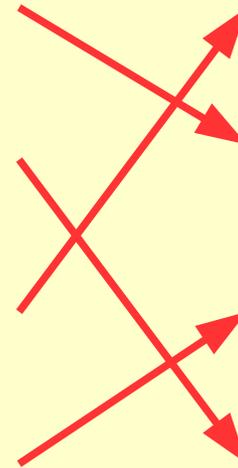
窪田二郎 (くぼたじろう)
ku bo ta ji ro u

久保田四郎

久保田五郎

窪田二郎

窪田三郎



If you'd like to know more about arranging Japanese index...

- **SIST 13 – Indexes and Indexing**
 - One of the Standards for Information of Science and Technology by the Ministry of Education, Culture, Sports, Science and Technology
 - http://sti.jst.go.jp/sist/handbook/sist13/sist13_m.htm

**Making
(Japanese)
index with
LaTeX**

Embed index entry with kana

算術平均 (for “arithmetic mean”)

↓ Turn into kana (syllables)

さんじゅつへいきん

↓ Put it in `\index` command

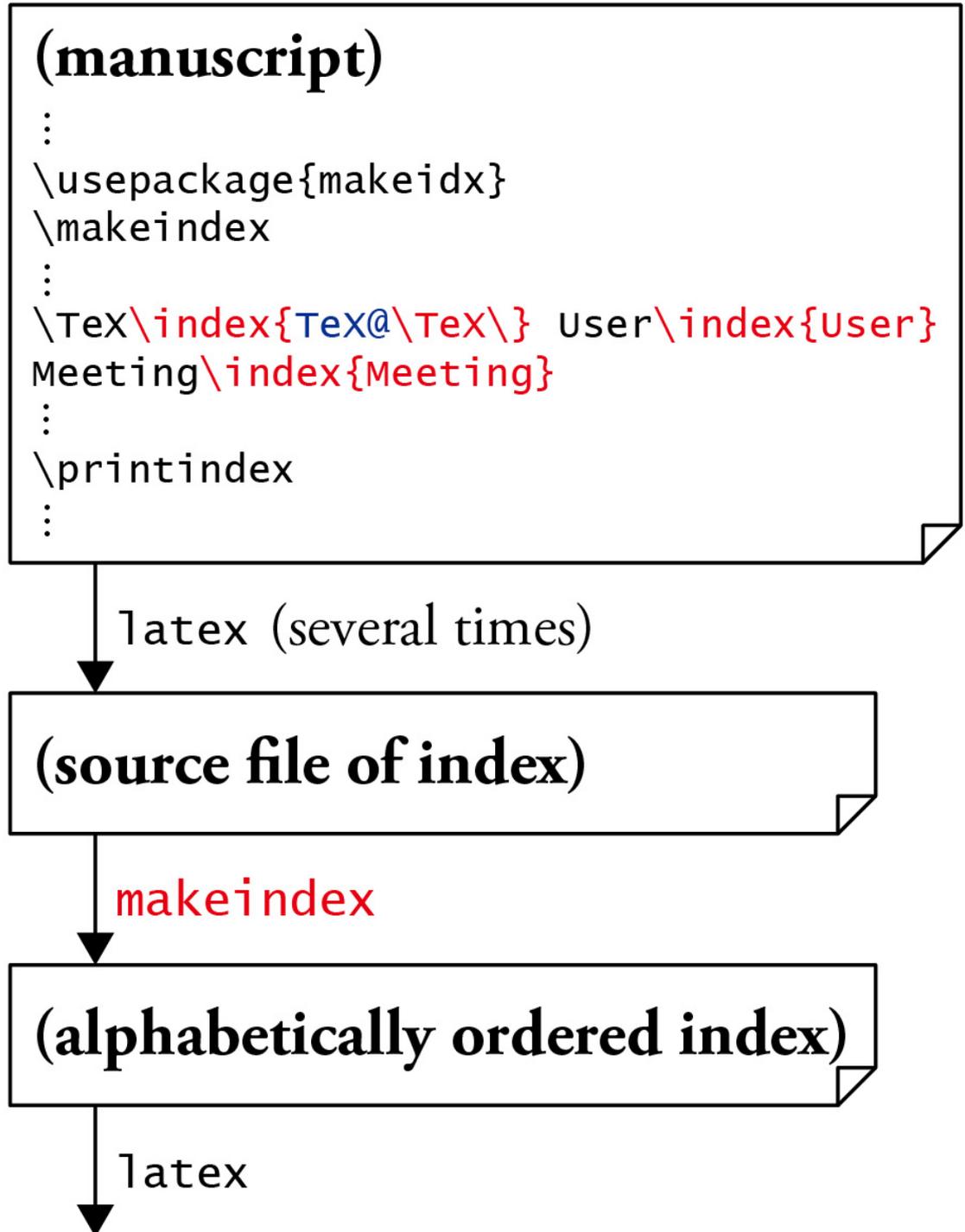
`\index{さんじゅつへいきん@算術平均}`

kana(syllables)

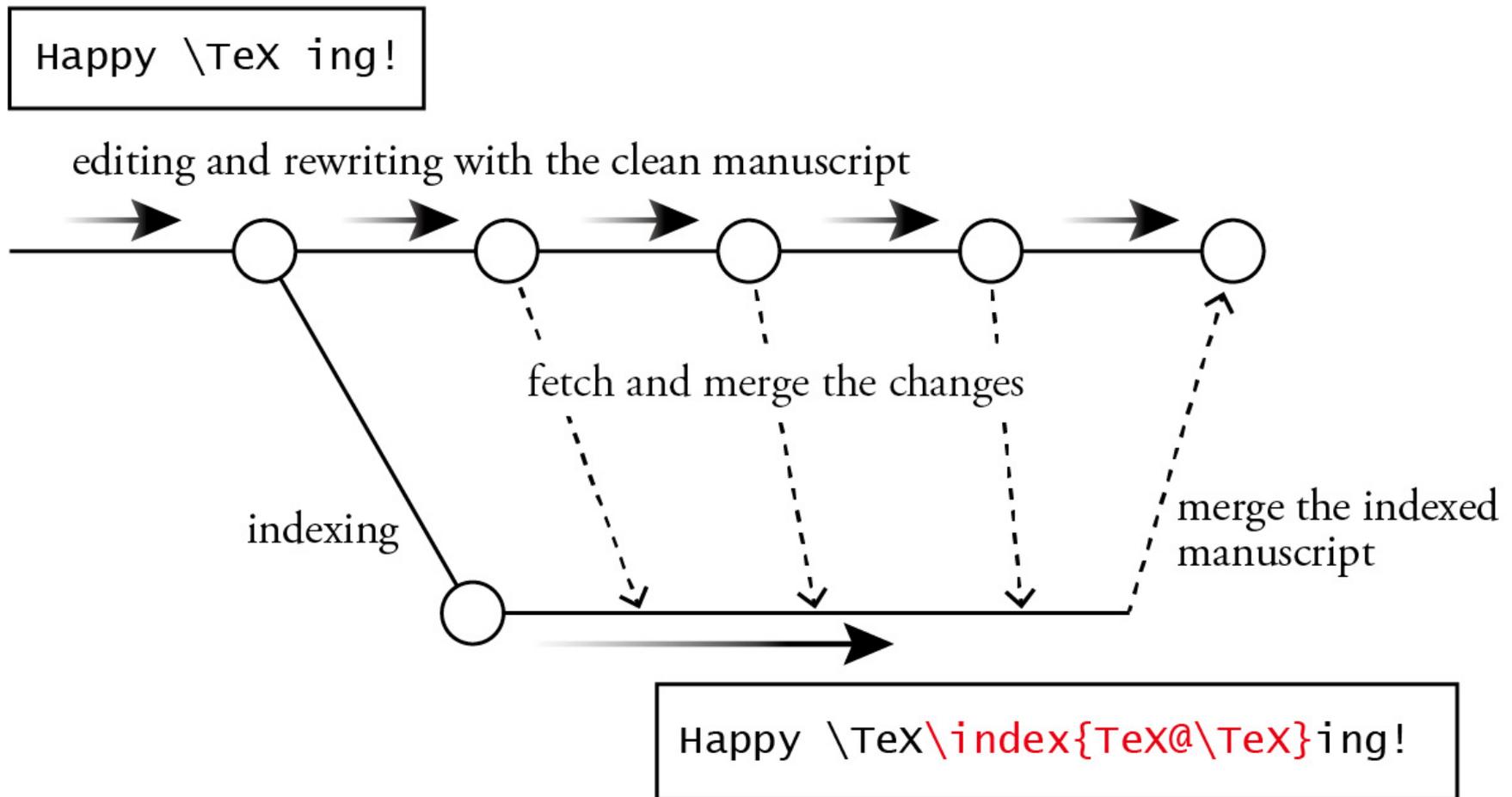
word appearing
in the index

Indexing in LaTeX

- Embed the index entries within the manuscripts using `\index` command.
- `makeindex` arrange the entries so as to be sorted and laid out.
- The manuscript tends to be messy.



Tips: How to avoid messy manuscripts



Indexing Japanese in LaTeX

- Include kana in each `\index` command.
- `mindex` arrange the entries so as to be sorted and laid out.
- The manuscript tends to be messier.

(manuscript)

```
⋮  
\usepackage{makeidx}  
\makeindex  
⋮  
\TeX\index{TeX@\TeX\} ユーザー \index{ゆーざー @ ユーザー }  
の集まり \index{あつまり @ 集まり }  
⋮  
\printindex  
⋮
```

↓ latex (several times)

(source file of index)

↓ mindex

(alphabetically ordered index)

↓ latex

Kana depends on the context (not just on kanji)

相性

Turn into kana

for “compatibility”

あいしょう
a i syo u

多相性

Turn into kana

for “polymorphism”

たそうせい
ta so u se i

多相性

Turn into kana

for “a lot of compatibilities”

たあいしょう?
ta a i syo u

Morphological analyzer and its corner case

多相的な $\xrightarrow{\text{morphological analyzer}}$ たそう / てき / な

success

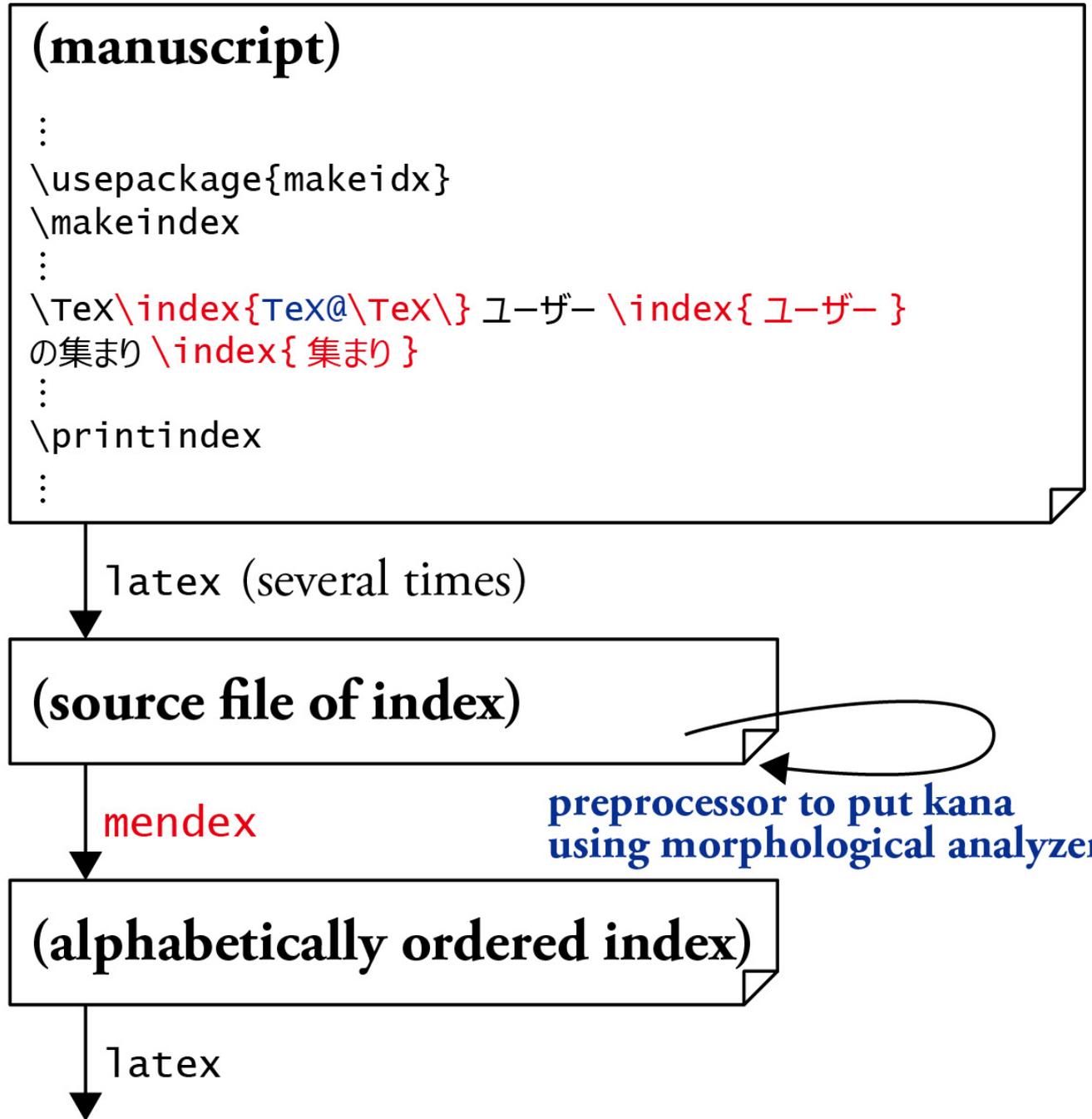
多相性 $\xrightarrow{\text{morphological analyzer}}$ た / あいしょう

たそう / せい

fail

Using morphological analyzer to put kana

- The manuscript is slightly tidier than before!

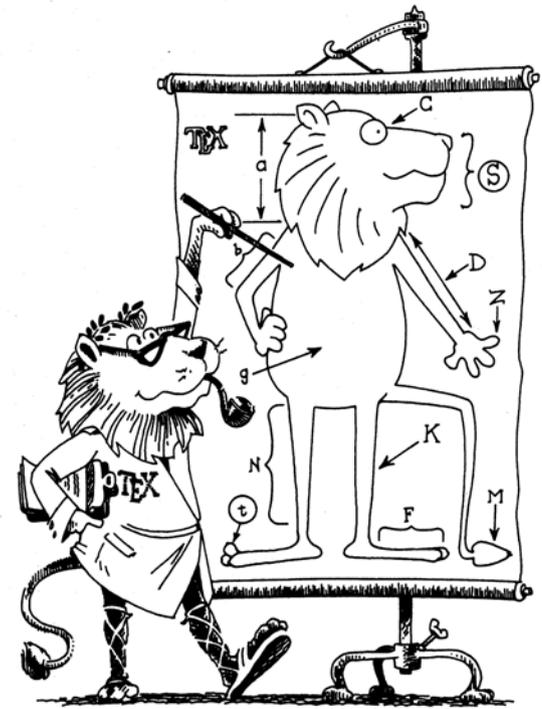


**Why bother
with
indexing?**

Example

The TeXbook

I
Index



Components of index

Components

- Headings** → exercises, vii, 1–303.
*`\exhyphenpenalty`, 96, 262, 272, 348.
`\exists` (\exists), 435.
`\exp` (exp), 162, 361.
- References** → *`\expandafter`, 40, 213, 215, 260, 308, 390, 348, 374–380.
expansion of expandable tokens, 212–216, 238, 267, 373–374.
- Subheadings** → avoiding, 216, 262–263, 377.
explicit hyphens, 4, 93, 454.
explicit kerns, 40, 280, 306, 454–455.
- Cross-references** → exponents, see superscripts.
extensible characters, 442.
extension fonts, 157, 351, 432–433, 441, 447.
extensions to TeX, 226, 228–229.
- Modifications** → eyestrain, reducing, 59.
faces, 13, 17, 390.
`\factor`, 270.
factorial, 169, 181.
*`\fam`, 154–159, 273, 289–290, 346–347, 351, 358, 414–415.
families, 153–159, 289–290, 346, 431, 442.
`\familyassignment`, 277.
`\familymember`, 271, 277.
family tree, 248.
- Qualifications** → `\family0` (math roman fonts), 153–157, 351.
`\family1` (math italic fonts), 155–157, 351.
`\family2` (math symbol fonts), 157, 351, 431, 441, 447.

Mandatory

Headings

References

- The reader will easily find what information the book has or not, just by revising the headings of the index.
- References are what make the index the index.

exercises, vii, 1–303.

*\exhyphenpenalty, 96, 262, 272, 348.

\exists (\exists), 435.

\exp (exp), 162, 361.

*\expandafter, 40, 213, 215, 260, 308,
390, 348, 374–380.

expansion of expandable tokens, 212–216,
238, 267, 373–374.

 avoiding, 216, 262–263, 377.

explicit hyphens, 4, 93, 454.

explicit kerns, 40, 280, 306, 454–455.

exponents, see superscripts.

extensible characters, 442.

extension fonts, 157, 351, 432–433, 441, 447.

extensions to T_EX, 226, 228–229.

eyestrain, reducing, 59.

faces, 13, 17, 390.

\langle factor \rangle, 270.

factorial, 169, 181.

*\fam, 154–159, 273, 289–290, 346–347,
351, 358, 414–415.

families, 153–159, 289–290, 346, 431, 442.

\langle family assignment \rangle, 277.

\langle family member \rangle, 271, 277.

family tree, 248.

family 0 (math roman fonts), 153–157, 351.

family 1 (math italic fonts), 155–157, 351.

family 2 (math symbol fonts), 157, 351,
431, 441, 447.

Subsidiary

- They help the reader learn the context of the headings.

Subheadings

Cross-references

Modifications

Qualifications

exercises, vii, 1–303.
*\exhyphenpenalty, 96, 262, 272, 348.
\exists (\exists), 435.
\exp (exp), 162, 361.
*\expandafter, 40, 213, 215, 260, 308,
390, 348, 374–380.
expansion of expandable tokens, 212–216,
238, 267, 373–374.
avoiding, 216, 262–263, 377.
explicit hyphens, 4, 93, 454.
explicit kerns, 40, 280, 306, 454–455.
exponents, see superscripts.
extensible characters, 442.
extension fonts, 157, 351, 432–433, 441, 447.
extensions to TeX, 226, 228–229.
eyestrain, reducing, 59.
faces, 13, 17, 390.
<factor>, 270.
factorial, 169, 181.
*\fam, 154–159, 273, 289–290, 346–347,
351, 358, 414–415.
families, 153–159, 289–290, 346, 431, 442.
<family assignment>, 277.
<family member>, 271, 277.
family tree, 248.
family 0 (math roman fonts), 153–157, 351.
family 1 (math italic fonts), 155–157, 351.
family 2 (math symbol fonts), 157, 351,
431, 441, 447.

Allowed

- Any remarks that help reader get the proper reference out of several alternatives.
- For example, each asterisk(*) indicates that the following heading is a primitive of TeX

exercises, vii, 1–303.

*`\exhyphenpenalty`, 96, 262, 272, 348.

`\exists` (\exists), 435.

`\exp` (\exp), 162, 361.

*`\expandafter`, 40, 213, 215, 260, 308, 330, 348, 374–380.

expansion of expandable tokens, 212–216, 238, 267, 373–374.

avoiding, 216, 262–263, 377.

explicit hyphens, 4, 93, 454.

explicit kerns, 40, 280, 306, 454–455.

exponents, see superscripts.

extensible characters, 442.

extension fonts, 157, 351, 432–433, 441, 447.

extensions to TeX, 226, 228–229.

eyestrain, reducing, 59.

faces, 13, 17, 390.

`\factor`, 270.

factorial, 169, 181.

*`\fam`, 154–159, 273, 289–290, 346–347, 351, 358, 414–415.

families, 153–159, 289–290, 346, 431, 442.

`\family assignment`, 277.

`\family member`, 271, 277.

family tree, 248.

family 0 (math roman fonts), 153–157, 351.

family 1 (math italic fonts), 155–157, 351.

family 2 (math symbol fonts), 157, 351, 431, 441, 447.

Unrecommended

- **Auxiliary information beyond the main purpose of the index.**
- **Exclude everything which gets the readers nowhere!**

Pointless Components

Deinocheirus	Osmólska & Roniewicz, 1970	158
Deinonychus	Ostrom, 1969	170
Deinosuchus	Holland, 1909	14
Dilong	Xu et al., 2004	149
Dilophosaurus	Welles, 1970	41, 124, 125
Dimetrodon	Cope, 1878	9
Dimorphodon	Buckland, 1829	18
Diplodocus	Marsh, 1878	8,21,41,102

- Excerpt from the index of “The Visual Book of Dinosaurs for Kids”
- Who cares the name of discoverers in the index?

**Indexing for
who?**

Three types of index users

- 1. Looking-up**
- 2. Reluctant**
- 3. Potential**

Looking-up readers

- **Already read the book.**
- **At least, know the subject.**
(Books on technologies or any sort of practical knowledge often have these kind of readers.)
- **They'd like to go directly to some specific part of the book.**

Reluctant index-users

- Readers would get lost during reading the book when ...
 - they forget the key words.
 - the book is poorly constructed.
- The latter could be avoided by indexing.

Potential readers

- You can use the index to grasp the contents of the book even before actually read it through.
- The index exhibits the contents of the book in another way than the book itself.

Index and ebook

**Is it necessary
anyway?**

Is it necessary anyway?

- **Definitely!**
- **They would supplement a full text literature search through ebooks.**
- **It may take a different shape from that of the printed books.**

Summing up

With bad index, readers ...

- would be just confused.
- have to read through the whole book anyway.
- would come to read an ill written or edited book.

With good index, readers ...

- could directly reach to the proper pages.
- could grasp the whole book.
- could get another aspect of the book.

What we have seen

- Indexes from the world
- How to arrange Japanese words in LaTeX realm
- Good index, bad index