

В. М. Руденко

МАТЕМАТИЧНА СТАТИСТИКА

НАВЧАЛЬНИЙ ПОСІБНИК

*Рекомендовано
Міністерством освіти і науки України
для студентів вищих навчальних закладів*

Київ
«Центр учбової літератури»
2012

УДК 519.22(075.8)
ББК 22.171я73
Р 83

*Гриф надано
Міністерством освіти і науки України
(Лист № 1/11-2918 від 08.04.2010 р.)*

Рецензенти:

Бомба А. Я. – доктор фізико-математичних наук, професор;
Ямницький В. М. – доктор психологічних наук, професор.

Руденко В. М.

Р 83 Математична статистика. Навч. посіб. – К.: Центр учбової літератури, 2012. – 304 с.

ISBN 978-611-01-0277-3

Розкриваються основи теорії ймовірностей та математичної статистики: предмет, методи, базові категорії, показники тенденцій і мінливості сукупностей, статистичне оцінювання, перевірка статистичних гіпотез з використанням параметричних і непараметричних критеріїв, кореляційний, регресійний, дисперсійний аналіз. Розглядаються технологічні прийоми і способи комп'ютерної реалізації статистичної обробки даних на базі табличного процесора MS Excel, організація діалогового інтерфейсу.

Пропонується матеріал для самостійної підготовки (лабораторно-практичні та тестові завдання). Приклади супроводжуються розрахунками, графічним ілюстративним матеріалом.

Посібник може бути корисним для аспірантів і магістрантів, учителів і психологів, студентів і викладачів вищих навчальних закладів.

УДК 519.22(075.8)
ББК 22.171я73

ISBN 978-611-01-0277-3

© Руденко В. М., 2012.
© Центр учбової літератури, 2012.

ЗМІСТ

ВСТУП	7
1. ПРЕДМЕТ МАТЕМАТИЧНОЇ СТАТИСТИКИ	9
Основні завдання та методи математичної статистики	9
2. СТАТИСТИЧНІ ПОКАЗНИКИ ВИБІРКИ	15
2.1. Емпіричні розподіли	15
Варіаційні ряди та статистичні розподіли	15
Незгруповані розподіли	18
Згруповані розподіли	26
Атрибутивні розподіли	31
Ранжировані розподіли	32
2.2. Показники вибірки	36
Міри центральної тенденції (МЦТ)	36
Міри мінливості (ММ)	39
Розрахунки та інтерпретація МЦТ і ММ	43
Початкові та центральні моменти	49
Кванти лі	51
Нормовані дані	53
2.3. Кореляційний аналіз	56
Сутність кореляції	56
Лінійна кореляція	58
Нелінійна кореляція	62
Коефіцієнти взаємної зв'язаності	64
2.4. Регресійний аналіз	68
Одномірна лінійна регресія	68
Множинна регресія	71

3. ОСНОВИ ТЕОРІЇ ЙМОВІРНОСТЕЙ	74
3.1. Випробування та події	75
Основні поняття і означення	75
Операції над подіями	77
Ймовірність подій	80
Умовна ймовірність	84
Формула повної ймовірності	85
Формула Байєса	89
Елементи комбінаторики	91
3.2. Випадкові величини	94
Розподіли випадкових величин	94
Характеристики випадкових величин	106
3.3. Закон великих чисел	118
Повторні випробування	118
Теорема Бернуллі	121
Теорема Чебишева	123
Центральна гранична теорема	125
3.4. Теоретичні розподіли ймовірностей	129
Біноміальний розподіл	129
Нормальний розподіл	135
Розподіли «хі-квадрат», Фішера і Стьюдента	142
4. СТАТИСТИЧНЕ ОЦІНЮВАННЯ	151
Поняття статистичного оцінювання параметрів	151
Точкове оцінювання, властивості статистичних оцінок	154
Методи статистичного оцінювання параметрів	156
Інтервальне оцінювання	160
5. ПЕРЕВІРКА СТАТИСТИЧНИХ ГІПОТЕЗ	165
5.1. Характеристика методів перевірки статистичних гіпотез	165
Поняття статистичної гіпотези	165
Статистичні критерії	167

Параметричні і непараметричні критерії	167
Рівень статистичної значущості	168
Правила прийняття статистичних рішень	169
Помилки прийняття статистичних рішень	170
Статистичні рішення на основі p -значень	171
Типи і загальна схема перевірки статистичних гіпотез	172
5.2. Гіпотези щодо нормального розподілу ознак	174
Критерії асиметрії та ексцесу t_A і t_E	175
Критерій згоди χ^2	178
Критерій Шапіро-Віллка W	180
5.3. Перевірка однорідності вибірок	184
Критерій Стьюдента t	185
Критерій Крамера-Велча T	189
Критерій Колмогорова-Смірнова λ	191
Критерій Вілкоксона-Манна-Вітні U	193
Критерій Лемана-Розенблатта	196
5.4. Перевірка гіпотез про чисельні значення параметрів	201
Значущість середнього (критерій z , дисперсія відома)	202
Значущість середнього (критерій t , дисперсія невідома)	204
Значущість дисперсії (критерій χ^2)	206
Відмінності у значеннях середніх	
(t -критерій для двох зв'язаних вибірок)	208
Відмінності у значеннях дисперсій	
(F -критерій Фішера для двох незв'язаних вибірок)	210
Відмінності у значеннях дисперсій	
(t -критерій Стьюдента для двох зв'язаних вибірок).	214
Відмінності у значеннях дисперсій 3-х і більш сукупностей	
(критерій Кохрана q для вибірок однакових обсягів)	216
Відмінності у значеннях дисперсій 3-х і більш сукупностей	
(критерій Бартлета M для вибірок різних обсягів)	217

5.5. Виявлення відмінностей і зсуву у рівні ознаки	221
Критерій Крускала-Волліса H	222
Критерій Фрідмана χ^2_r	224
Критерій тенденцій Пейджа L	227
5.6. Перевірка значущості коефіцієнтів кореляції	231
Коефіцієнт лінійної кореляції Пірсона r_{xy}	231
Коефіцієнт рангової кореляції Спірмена r_s	233
Дихотомічний коефіцієнт кореляції Пірсона ϕ	234
Точково-бісеріальний коефіцієнт кореляції r_{pb}	236
6. ДИСПЕРСІЙНИЙ АНАЛІЗ	239
Дисперсійний однофакторний аналіз	239
Дисперсійний двофакторний аналіз	244
7. ЗАВДАННЯ ДЛЯ САМОСТІЙНОЇ РОБОТИ	252
7.1. Лабораторні роботи	252
7.2. Тестові завдання	262
7.3. Практичні контрольні завдання	280
ДОДАТКИ	285
ЛІТЕРАТУРА	298

ВСТУП

Психолог у своїй діяльності нерідко має справу з масивами емпіричної інформації і змушений будувати свої висновки в умовах невизначеності. Така ситуація зумовлена особливостями психологічних об'єктів. Як правило, вони є стохастичними за своєю природою, їхній поведінці притаманна деяка невизначеність, а станам – випадковість. Тому психологічні дослідження з залучення математики здебільшого виконуються з використанням методів *математичної статистики*. Ось чому «провести компетентне дослідження або прочитати і зрозуміти звіти про дослідження, не опанувавши імовірнісним і статистичним мисленням, неможливо» – пише професор Фред Керлінгер у своїй книзі «Основи дослідження поведінки»¹.

Математична статистика вивчає і одночасно відображає імовірнісну (випадкову) природу процесів і подій, що значною мірою є характерною рисою соціальної, політичної, педагогічної та інших сфер життя та діяльності.

Математична статистика і теорія ймовірностей зіграли значну роль у впровадженні математичних методів в психологію. Саме вони з моменту виникнення постійно прагнули поширити сферу свого застосування на проблеми імовірнісного аналізу особистісних рис, здібностей та поведінки. Прикладами є «Досвід моральної арифметики» Ж. Бюффона (кінець XVIII ст.), «Людина і розвиток її здібностей або досвід суспільної фізики» А. Кетле (середина XIX ст.) і чимало інших робіт.

Проте залучення математичної статистики стимулювалося не внутрішньою логікою розвитку наукових психологічних ідей, а суб'єктивним бажанням математиків застосувати імовірнісно-статистичні методи до широкого кола явищ гуманітарного характеру. І тільки з XX ст. починається цілеспрямована розробка й ефективне використання статистичних методів для аналізу емпіричних даних (при вимірюванні, оцінюванні, перевірці причинно-

¹ *Kerlinger F.N. Foundations of Behavioral Research. Holt, Rinehart and Winston, New York, 1964, p. IX.*

наслідкових гіпотез і т.д.). З'являються такі потужні багатомірні статистичні методи, як факторний аналіз, суттєвий внесок у розробку якого на початковій стадії зробив Чарльз Спірмен (1863-1945), дисперсійний аналіз, пов'язаний з ім'ям Рональда Фішера (1890-1962), та інші методи.

Вивчення математичної статистики у вищих навчальних закладах спрямоване на формування теоретичної й методологічної бази студентів для поглибленого опанування знаннями та вміннями щодо організації, проведення та інтерпретації результатів психологічних досліджень. Засвоєння курсу є важливим як в аспекті розуміння важливості основ математичної статистики в гуманітарних дослідженнях, так і в оволодінні методикою і технікою статистичних обчислень з використанням сучасних комп'ютерних технологій.

При формуванні змісту посібника і виборі наочних засобів доведення його до свідомості студентів було також враховано доволі скромний початковий рівень математичної підготовки гуманітаріїв, не завжди високу умотивованість до вивчення точних наук. Для досягнення прийнятних результатів у навчанні залучено доступний для розуміння, не перевантажений складними математичними викладками, посильний і цікавий студентам у фаховому плані матеріал без зайвих спрощень і вульгаризацій математичних положень.

Користуючись нагодою, автор висловлює глибоку вдячність ректорові Рівненського державного гуманітарного університету професору Руслану Михайловичу Постолювському за підтримку і сприяння у підготовці посібника до видання. Щира подяка професорам Андрію Ярославовичу Бомбі та Вадиму Марковичу Ямницькому за допомогу при апробації рукопису.

Особиста подяка завідувачеві кафедри теоретичної та прикладної статистики Львівського національного університету імені Івана Франка професору Ярославу Івановичу Єлейко за поради і зауваження, що були корисними під час доопрацювання рукопису.