МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

ДЕРЖАВНИЙ ВИЩИЙ НАВЧАЛЬНИЙ ЗАКЛАД

ДОНЕЦЬКІЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ

Кафедра АТ

МЕТОДИЧНІ ВКАЗІВКИ

ДО ВИКОНАННЯ ІНДИВІДУАЛЬНИХ ДОМАШНІХ ЗАВДАНЬ

З КУРСУ

"ТЕОРІЯ ІМОВІРНОСТЕЙ І МАТЕМАТИЧНА СТАТИСТИКА"

для студентів спеціальностей
6.092401 "Телекомунікаційні мережі і системи"
і 6.091401 "Системи управління та автоматики"
денної форми навчання

Затверджено на засіданні кафедри
"Автоматика і телекомунікації"
протокол № 4 від 22.04.2009

Затверджено на засіданні
навчально-видавничої ради ДонНТУ
протокол № _ від _________

Донецьк - 2009

Методичні вказівки до виконання індивідуальних домашніх завдань з курсу “Теорія імовірностей і математична статистика” для спеціальностей 6.092401 “Телекомунікаційні мережі і системи” та 6.091401 “Системи управління та автоматики” денної форми навчання / Воропаєва В.Я., Долгіх І.П., Зайцева Е.Є., Коваленко Є.Г. – Донецьк, ДонНТУ, 2009. – 46 с.

Містять основні теоретичні відомості та рекомендації до самостійного виконання задач двох індивідуальних завдань за матеріалами першого та другого змістовних модулей з курсу “Теорія імовірностей і математична статистика” відповідно.

Укладачі: Воропаєва В. Я., к.т.н., доцент

Долгіх І.П., ст. викладач

Зайцева Е.Є., асистент

Коваленко Є.Г., асистент

Відповідальний за випуск: Бессараб В.І., зав. кафедрою АТ, к.т.н., доцент

ПЕРЕДМОВА
ДО ІНДИВІДУАЛЬНОГО ЗАВДАННЯ №1

Індивідуальне завдання №1 спрямоване на закріплення теоретичних знань та набуття практичних навичок за матеріалами першого змістовного модуля курсу «Теорія імовірностей і математична статистика» та містить задачі за темами:

1. Основні поняття теорії імовірностей, формули комбінаторики і геометрична імовірність.
2. Теореми додавання і множення імовірностей. Формула повної імовірності. Формули Бейеса. Повторення випробувань. Формула Бернуллі.
3. Дискретні випадкові величини.
4. Функції та щільність розподілу імовірностей дискретних і безперервних випадкових величин.

Індивідуальне завдання №1 складається з восьми задач, сім з яких містяться у збірнику [2]. Варіанти завдання наведені в таблиці в кінці методичних вказівок. При виконанні індивідуального завдання для отримання повної інформації варто користуватись літературою, наведеною у переліку посилань. Теоретичні відомості в методичних вказівках до виконання індивідуального завдання викладено в нататій послідовності:

- для розв'язання задач з першого завдання (№№12-42) слід використовувати теоретичні відомості 1-ого розділу – «Основні поняття теорії імовірностей». Для отримання додаткової інформації див. [1, ст.17-30], [2, ст. 8-18];
- для розв'язання задач з другого завдання (№№47-88) слід використовувати теоретичні відомості пунктів 2.1-2.2 2-ого розділу – «Основні теореми». Для отримання додаткової інформації див. [1, ст.31-48], [2, ст.18-30];

- для розв’язання задач з третього завдання (№№90-118) слід використовувати теоретичні відомості пунктів 2.3-2.4 2-ого розділу «Основні теореми». Для отримання додаткової інформації див. [1, ст. 50-55], [2, ст.31-37];
- для розв’язання задач з четвертого, п’ятого та шостого завдань (№№167-229) слід використовувати теоретичні відомості 3-ого розділу – «Дискретні випадкові величини». Для отримання додаткової інформації див. [1, ст. 64-100], [2, ст. 52-79];
- для розв’язання задач з сьомого завдання (№№263-289) слід використовувати теоретичні відомості пунктів 4.1-4.2 4-ого розділу – «Функції та щільність розподілу імовірностей випадкових величин». Для отримання додаткової інформації див. [1, ст. 111-1119], [2, ст. 87-97];
- для розв’язання задач з восьмого завдання (№№ 330-370) слід використовувати теоретичні відомості пунктів 4.3-4.4 4-ого розділу – «Функції та щільність розподілу імовірностей випадкових величин». Для отримання додаткової інформації див. [1, ст. 122-154], [2, ст. 106-119].

ЗМІСТ ІНДИВІДУАЛЬНОГО ЗАВДАННЯ №1

Завдання 1-5 і 7-8 наведені у збірнику В.Е. Гмурман. Руководство к решению задач по теории вероятностей и математической статистике. М.-2003. Завдання 6 має наступні варіанти:

A. З ящика, у якому 3 чорні та 2 білі кульки послідовно виймають кульки, доки не з'явиться біла кулька. Визначити ряд розподілу, функцію розподілу, МО, дисперсію, СКВ випадкової величини, що дорівнює кількості вийнятих чорних кульок. Побудувати багатокутник розподілу і графік функції розподілу.

B. У мішень роблять 3 постріли. Імовірність влучення при кожному - 0,8. У разі влучення стрілок отримує 10 балів. Для ВВ, яка дорівнює кількості отриманих балів, визначити ряд розподілу, функцію розподілу, МО, дисперсію, СКВ. Побудувати багатокутник розподілу і графік функції розподілу.

C. Визначити закон розподілу ВВ – кількості випадання 6 очок на гральній кісті, якщо її кидають 3 рази. Розрахувати функцію розподілу, МО, дисперсію, СКВ. Побудувати багатокутник розподілу і графік функції розподілу.

D. Автомобіль проїжджає 3 світлофори, кожний із імовірністю 0,5 може або забороняти, або дозволяти проїзд. Для ВВ, яка дорівнює кількості світлофорів, що пройдені до першої зупинки, знайти ряд розподілу, функцію розподілу, МО, дисперсію, СКВ. Побудувати багатокутник розподілу і графік функції розподілу.

E. Монету підкинули 4 рази. Імовірність випадання герба 0,5. Для ВВ, яка дорівнює кількості появ герба знайти ряд розподілу, функцію розподілу, МО, дисперсію, СКВ. Побудувати багатокутник розподілу і графік функції розподілу.

F. Імовірність влучення рухомої цілі одним пострілом дорівнює 0,3. Стрільба триває до першого влучення. Доки ціль знаходиться в зоні обстрілу, у неї встигають зробити 4 постріли. Які середні витрати снарядів при стрільбі у 100 цілей. Для ВВ, яка дорівнює кількості використаних снарядів, знайти ряд розподілу, функцію розподілу, МО, дисперсію, СКВ. Побудувати багатокутник розподілу і графік функції розподілу.

G. Імовірність влучення по цілі одним пострілом дорівнює 2/3. Для ВВ, яка дорівнює кількості влучень при 4 пострілах, знайти ряд розподілу, функцію розподілу, МО, дисперсію, СКВ. Побудувати багатокутник розподілу і графік функції розподілу.

H. Серед 20 приладів 6 неточних. Для ВВ, яка дорівнює кількості точних приладів серед 5 довільно обраних приладів, знайти ряд розподілу, функцію розподілу, МО, дисперсію, СКВ. Побудувати багатокутник розподілу і графік функції розподілу.

I. При підгонці деталей може виконуватись декілька спроб. Імовірність того, що потрібна 1 спроба - 0,3; 2 спроби - 0,5; 3 спроби – 0,1; 4 спроби – 0,1. Скільки спроб потрібно для оброблення 200 деталей? Для ВВ, яка дорівнює кількості спроб, знайти ряд розподілу, функцію розподілу, МО, дисперсію, СКВ. Побудувати багатокутник розподілу і графік функції розподілу.

J. Монету підкидають до появи решки, але не більше 5 разів. Для ВВ, яка дорівнює кількості підкидань, знайти ряд розподілу, функцію розподілу, МО, дисперсію, СКВ. Побудувати багатокутник розподілу і графік функції розподілу.

K. З ящика, у якому 5 чорних і 3 білих кульки послідовно виймають кульки, доки не з'явиться біла кулька. Визначити ряд розподілу, функцію розподілу, МО, дисперсію, СКВ випадкової величини, що дорівнює кількості вийнятих чорних кульок. Побудувати багатокутник розподілу і графік функції розподілу.

L. У мішень роблять 3 постріли. Імовірність влучення при кожному - 0,6. У разі влучення стрілок отримує 5 балів. Для ВВ, яка дорівнює кількості отриманих балів, визначити ряд розподілу, функцію розподілу, МО, дисперсію, СКВ. Побудувати багатокутник розподілу і графік функції розподілу.

M. Визначити закон розподілу ВВ – кількості випадання 3 очок на гральній кісті, якщо її кидають 4 рази. Розрахувати функцію розподілу, МО, дисперсію, СКВ. Побудувати багатокутник розподілу і графік функції розподілу.

N. Автомобіль проїжджає 3 світлофори, кожний із імовірністю 0,5 може або забороняти, або дозволяти проїзд. Для ВВ, яка дорівнює кількості світлофорів, що пройдені до першої зупинки, знайти ряд розподілу, функцію розподілу, МО, дисперсію, СКВ. Побудувати багатокутник розподілу і графік функції розподілу.

O. Монету підкинули 3 рази. Імовірність випадання герба 0,5. Для ВВ, яка дорівнює кількості появ герба знайти ряд розподілу, функцію розподілу, МО, дисперсію, СКВ. Побудувати багатокутник розподілу і графік функції розподілу.

P. Імовірність влучення рухомої цілі одним пострілом дорівнює 0,4. Стрільба триває до першого влучення. Доки ціль знаходиться в зоні обстрілу, у неї встигають зробити 3 постріли. Які середні витрати снарядів при стрільбі у 200 цілей. Для ВВ, яка дорівнює кількості використаних снарядів, знайти ряд розподілу, функцію розподілу, МО, дисперсію, СКВ. Побудувати багатокутник розподілу і графік функції розподілу.

Q. Імовірність влучення по цілі одним пострілом дорівнює 0,4. Для ВВ, яка дорівнює кількості влучень при 5 пострілах, знайти ряд розподілу, функцію розподілу, МО, дисперсію, СКВ. Побудувати багатокутник розподілу і графік функції розподілу.

R. Серед 15 приладів 5 неточних. Для ВВ, яка дорівнює кількості точних приладів серед 6 довільно обраних приладів, знайти ряд розподілу, функцію розподілу, МО, дисперсію, СКВ. Побудувати багатокутник розподілу і графік функції розподілу.

S. При налагодженні приладу необхідно виконувати від 1 до 3 тестів. Імовірність того, що потрібен 1 тест - 0,3; 2 тести - 0,5. Скільки тестів потрібно для налагодження 200 приладів? Для ВВ, яка дорівнює кількості спроб, знайти ряд розподілу, функцію розподілу, МО, дисперсію, СКВ. Побудувати багатокутник розподілу і графік функції розподілу.

T. Монету підкидають до появи решки, але не більше 3 разів. Для ВВ, яка дорівнює кількості підкидань монети, знайти ряд розподілу, функцію розподілу, МО, дисперсію, СКВ. Побудувати багатокутник розподілу і графік функції розподілу.

U. З ящика, у якому 4 чорні і 3 білих кульки послідовно виймають кульки, доки не з'явиться біла кулька. Визначити ряд розподілу, функцію розподілу, МО, дисперсію, СКВ випадкової величини, що дорівнює кількості вийнятих чорних кульок. Побудувати багатокутник розподілу і графік функції розподілу.

V. У мішень роблять 3 постріли. Імовірність влучення при кожному - 0,6. У разі влучення стрілок отримує 10 балів. Для ВВ, яка дорівнює кількості отриманих балів, визначити ряд розподілу, функцію розподілу, МО, дисперсію, СКВ. Побудувати багатокутник розподілу і графік функції розподілу.

W. Визначити закон розподілу ВВ – кількості випадання 6 очок на гральній кісті, якщо її кидають 5 разів. Розрахувати функцію розподілу, МО, дисперсію, СКВ. Побудувати багатокутник розподілу і графік функції розподілу.

X. Імовірність влучення рухомої цілі одним пострілом дорівнює 0,3. Стрільба триває до першого влучення. Доки ціль знаходиться в зоні обстрілу, у неї встигають зробити 5 пострілів. Які середні витрати снарядів при стрільбі у 100 цілей. Для ВВ, яка дорівнює кількості використовуваних снарядів, знайти ряд розподілу, функцію розподілу, МО, дисперсію, СКВ. Побудувати багатокутник розподілу і графік функції розподілу.

Y. Імовірність влучення по цілі одним пострілом дорівнює 0,3. Для ВВ, яка дорівнює кількості влучень при 3 пострілах, знайти ряд розподілу, функцію розподілу, МО, дисперсію, СКВ. Побудувати багатокутник розподілу і графік функції розподілу.

Z. Серед 10 приладів 3 неточні. Для ВВ, яка дорівнює кількості точних приладів серед 4 навздогад обраних приладів, знайти ряд розподілу, функцію розподілу, МО, дисперсію, СКВ. Побудувати багатокутник розподілу і графік функції розподілу.

ТАБЛИЦЯ ВАРІАНТІВ ДО ІЗ №1

№	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
1	12	13	14*а*	14*б*	15	16	18	19	20	21*а*	21*б*	21*в*	22	32	33*а*	33б	34	31	28	29	38	37	36	30	42	26
2	52	53	54	55	56*б*	58	57*а*	61	67	68	70	81	82	83	57*б*	87	50	88	51	47	57*б*	56*в*	56*а*	51	50	47
3	90	91	92	93	94	95	96	98	99	100	101	102	111 *б*	113 *а*	112 *б*	113 *б*	115 *б*	115 *в*	115 *г*	118	96	94	93	90	92	91
4	167	168	169	175	178	180 *а*	180 *б*	186 *а*	186 *б*	171	173	175	168	178	169	180*г*	180 *в*	177	173	167	175	168	177	169	178	168
5	191	193	194	200	209	211 *а*	211 *б*	214	211 *б*	217	219	220	229	201 *б*	206	214	201 *а*	198	229	220	211 *а*	191	194	200	209	193
6	A	B	C	D	E	F	G	I	K	H	L	M	N	O	P	R	S	T	U	J	Q	V	W	X	Y	Z
7	263	268	269	276	265	263	266	268	269	270	272	273	274	276	263	279	283	284	288	263	272	289	270	263	269	270
8	330	332	333	335	338	341	342	343	330	338	351	352	332	338	368 *а*	368 *б*	370 *а*	370 *б*	370 *в*	368 *а*	330	351	335	343	338	352

ПЕРЕДМОВА
ДО ІНДИВІДУАЛЬНОГО ЗАВДАННЯ № 2

Однією з центральних задач математичної статистики є задача оцінки теоретичного розподілу випадкової величини на основі вибіркових даних. При цьому часто припускається, що закон розподілу генеральної сукупності відомий, але невідомі параметри цього закону (наприклад, математичне сподівання, дисперсія). Потрібно знайти наближене значення цих параметрів, тобто отримати їх статистичні оцінки.

Індивідуальне завдання №2 спрямоване на закріплення теоретичних знань та набуття практичних навичок за матеріалами другого змістовного модуля курсу «Теорія імовірностей і математична статистика» та містить задачі за темами:

1. Побудова емпіричної функції розподілу. Побудова гістограми та полігону частот (відносних частот).
2. Точкові оцінки параметрів розподілу. Інтервальні оцінки параметрів розподілу.
3. Перевірка статистичних гіпотез про параметри розподілу. Перевірка статистичних гіпотез про вид розподілу.

Індивідуальне завдання №2 складається з шести задач, які містяться у збірнику [2]. Варіанти завдання наведені в таблиці в кінці методичних вказівок. При виконанні індивідуального завдання для отримання повної інформації, варто користуватись літературою, наведеною у переліку посилань. Теоретичні відомості в методичних вказівках до виконання індивідуального завдання викладені в наступній послідовності:

- для розв'язання задач з першого та другого завдання (№№ 440-494, 635-658) слід використовувати теоретичні відомості 5-ого розділу – «Вибірковий метод». Для отримання додаткової інформації див. [1, С.192-196], [2, С.151-156];
- для розв'язання задач з третього та четвертого завдання (№№ 451-528) слід використовувати теоретичні відомості 6-ого розділу – «Статистичні

оцінки параметрів розподілу». Для отримання додаткової інформації див. [1, С.197-206, 213-222], [2, С.157-163, 174-180];

для розв'язання задач з п'ятого та шостого завдання (№№ 555-667) слід використовувати теоретичні відомості 7-ого розділу «Статистична перевірка статистичних гіпотез». Для отримання додаткової інформації див. [1, С.281-310, 329-332], [2, С.206-230, 251-282].

ЗМІСТ ІНДИВІДУАЛЬНОГО ЗАВДАННЯ №2

Завдання 1-6 наведені у збірнику В.Е. Гмурман. Руководство к решению задач по теории вероятностей и математической статистике. М. − 2003.

Завдання № 1 передбачає виконання одного з 3 типів задач:

1. Побудова емпіричної функції розподілу.
2. Побудова полігону частот.
3. Побудова полігону відносних частот.

Тому в таблиці варіантів вказано № задачі в збірнику, з якої необхідно взяти вихідні дані − ряд розподілу вибірки, і в дужках цифра 1, 2 або 3 − що необхідно зробити для цих вихідних даних. Наприклад, 442а (3) − для вихідних даних задачі 442а необхідно побудувати полігон відносних частот.

Завдання № 2 передбачає виконання одного з двох типів задач:

1. Побудова гістограми частот.
2. Побудова гістограми відносних частот.

Завдання виконується аналогічно попередньому.

Решта завдань передбачає розв’язок вказаних у збірнику задач.

ТАБЛИЦЯ ВАРІАНТІВ ДО ІЗ №2

№	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
1	442 а (1)	442 б (1)	442 а (2)	442 б (2)	442 а (3)	442 б (3)	444 а (1)	444 б (1)	444 а (2)	444 б (2)	444 а (3)	444 б (3)	445 а (1)	445 б (1)	445 в (1)	445 б (3)	445 в (3)	441 (2)	441 (3)	440 (2)	440 (3)	469 (1)	469 (2)	469 (3)	470 (2)	470 (3)	470 (1)
2	446 (2)	447 а (1)	447 б (1)	447 а (2)	447 б (2)	448 (1)	449 а (1)	449 б (1)	449 а (2)	449 б (2)	459 (1)	459 (2)	494 (1)	494 (2)	477 (1)	477 (2)	635 (1)	635 (2)	636 (1)	636 (2)	649 (1)	649 (2)	650 (1)	650 (2)	651 (1)	651 (2)	658 (2)
3	451	454	456	458	459	461	462	464	465	467	469	470	472	473	475	477	479	482	484	486	488	524 а	524 б	527 а	527 б	528 а	528 б
4	502 а	502 б	503	504	505	507	509	511	513 а	513 б	515	517	519	520	521	522 а	522 б	503	511	507	520	515	509	504	521	511	517
5	555	557	559	561	562	564	566	568	569	571	573	575 а	575 б	575 в	576 в	577 в	579 б	582	583	584	585	589	590	555	568	573	585
6	649	650	651	653	654	655	659	660	661	663	664	665	666	667	649	650	651	653	654	655	659	660	661	663	664	665	667

1 ОСНОВНІ ПОНЯТТЯ ТЕОРІЇ ІМОВІРНОСТЕЙ

1.1 Визначення імовірності

За класичним визначенням *імовірність події* визначається як

$$P(A) = m/n,$$

де m - кількість елементарних результатів випробувань, що сприяють появі події A; n - загальна кількість можливих елементарних результатів випробувань. Передбачається, що елементарні випробування складають повну групу подій та є рівноймовірними.

Із визначення імовірності випливають такі її властивості:

Властивість 1. *Імовірність достовірної події дорівнює одиниці*.

$$P(A) = m/n = n/n = 1.$$

Властивість 2. *Імовірність неможливої події дорівнює нулю.*

$$P(A) = m/n = 0/n = 0.$$

Властивість 3. *Імовірність випадкової події – додатне число, що знаходиться в інтервалі від нуля до одиниці.*

$$0 < P(A) < 1.$$

Властивість 4. *Імовірність будь-якої події – додатне число, що знаходиться в інтервалі від нуля до одиниці, включно.*

$$0 \le P(A) \le 1.$$

Відносна частота події A визначається, як

$$W(A) = \frac{w(A)}{N},$$

де $w(A)$ - кількість випробувань, у яких відбулась подія A; N - загальна кількість випробувань, що проводилися. При збільшенні кількості випробувань N відносна частота $W(A)$ наближається до значення імовірності $P(A)$.

1.2 Основні формули комбінаторики

Комбінаторика вивчає комбінації, що відповідають певним умовам, які можна скласти з елементів будь-якої природи заданої кінцевої множини. При безпосередньому визначенні імовірностей часто використовують формули комбінаторики.

Перестановки – комбінації, що складаються з одних і тих самих *n* будь-яких елементів і відрізняються лише їх порядком. Кількість можливих перестановок:

$$P_n = n!,$$

де $n! = 1 \cdot 2 \cdot 3 ... n$.

Розміщення – комбінації, що утворені з *n* різних елементів по *m* елементів, які відрізняються складом елементів, або їх порядком. Кількість можливих розміщень:

$$A_n^m = n(n-1)(n-2)...(n-m+1).$$

Сполучення – комбінації, що утворені з *n* різних елементів по *m* елементів, які відрізняються хоча б одним елементом.

$$C_n^m = \frac{n!}{m!(n-m)!}.$$

При розв'язанні задач комбінаторики використовують наступні правила:

Правило суми. Якщо об'єкт *A* може бути обраний із сукупності об'єктів *m* способами, а інший об'єкт *B* може бути обраний *n* способами, то обрати або *A*, або *B* можливо *m+n* способами.

Правило добутку. Якщо об'єкт *A* може бути обраний із сукупності об'єктів *m* способами і при кожному такому виборі об'єкт *B* може бути обраний *n* способами, то пара *(A,B)* може бути обрана $m \cdot n$ способами.

1.3 Геометричні імовірності

Припустимо, що відрізок l є частиною відрізка L. На відрізок L довільно поставили точку. Якщо припустити, що імовірність попадання точки на відрізок l пропорційна довжині цього відрізка и не залежить від його розташування відносно відрізку L, то імовірність попадання точки на відрізок l визначається, як

$$P = \frac{\text{Довжина l}}{\text{Довжина L}}.$$

Припустимо, що плоска фігура g є частиною плоскої фігури G. На фігуру G навмання поставили точку. Якщо припустити, що імовірність попадання точки на фігуру g пропорційна площині цієї фігури й не залежить ні від її розташування відносно G, ні від форми g, то імовірність попадання точки у фігуру g визначається, як

$$P = \frac{\text{Площина g}}{\text{Площина G}}.$$

Аналогічно визначається імовірність попадання точки у просторову фігуру v, яка є частиною фігури V:

$$P = \frac{\text{Об`єм v}}{\text{Об`єм V}}.$$

2 ОСНОВНІ ТЕОРЕМИ

2.1 Теорема додавання імовірностей

Сумою A+B *двох подій* A і B є подія, яка полягає у появі або події A, або події B, або обох цих подій. Якщо A і B несумісні, то A+B – подія, яка складається з появи будь-якої однієї з цих подій.

Сумою декількох подій називають подію, яка полягає у появі хоча б однієї з цих подій.

Теорема. Імовірність появи однієї з двох несумісних подій, будь-якої, дорівнює сумі імовірностей цих подій:

$$P(A+B)=P(A)+P(B).$$

Теорема. Сума імовірностей подій $A_1, A_2, ..., A_n$, що складають повну групу, дорівнює одиниці:

$$P(A_1)+P(A_2)+...+P(A_n)=1.$$

Протилежними є дві єдино можливі події, які утворюють повну групу.

Теорема. Сума імовірностей протилежних подій дорівнює одиниці:

$$P(A)+P(\overline{A})=1.$$

2.2 Теорема множення імовірностей

Добутком двох подій A і B є подія AB, яка полягає у сумісній появі цих подій. *Добутком декількох подій* є подія, що полягає у сумісній появі всіх цих подій.

Умовною імовірністю $P_A(B)$ називають імовірність події B, розраховану при умові, що подія A вже відбулась. Умовну імовірність можна розрахувати наступним чином:

$$P_A(B)=P(AB)/P(A) \quad (P(A)>0).$$

Теорема. Імовірність сумісної появи двох подій є добутком імовірності однієї з них на умовну імовірність іншої, розраховану за умови, що перша подія вже відбулась:

$$P(AB) = P(A)P_A(B).$$

Висновок. Імовірність сумісної появи декількох подій *n* дорівнює добутку імовірності однієї з них на умовні імовірності усіх інших, при чому імовірність кожної наступної події розраховується за умови, що всі попередні події вже відбулись:

$$P(A_1A_2A_3...A_n) = P(A_1)P_{A_1}(A_2)P_{A_1A_2}(A_3)...P_{A_1A_2...A_{n-1}}(A_n),$$

де $P_{A_1A_2...A_{n-1}}(A_n)$- імовірність події A_n, розрахованої за умови, що події $A_1, A_2, A_3,..., A_{n-1}$ відбулись.

Подія B є *незалежною* від події A, якщо поява події A не змінює імовірності події B, тобто якщо умовна імовірність події B дорівнює її безумовній імовірності:

$$P_A(B) = P(B).$$

Отже, якщо подія B не залежить від події A, то й подія A не залежить від події B; це означає, що властивість незалежності подій є взаємною.

Дві події називають *незалежними*, якщо імовірність їх сполучення дорівнює добутку імовірностей цих подій; інакше події називають *залежними*. Декілька подій називають *попарно незалежними*, якщо кожні дві з них незалежні. Декілька подій називають *незалежними у сукупності* (або *незалежними*), якщо незалежні кожні дві з них і незалежні кожна подія і всі можливі добутки інших.

Висновок. Імовірність сумісної появи подій, незалежних у сукупності, дорівнює добутку цих подій:

$$P(A_1A_2...A_n) = P(A_1)P(A_2)...P(A_n).$$

Примітка. Якщо події $A_1, A_2,..., A_n$ незалежні у сукупності, то й протилежні їм події $\overline{A}_1, \overline{A}_2,..., \overline{A}_n$ також незалежні у сукупності.

2.3 Висновки з теорем множення і добутку

Дві події називають *сумісними*, якщо поява однієї з них не виключає появи іншої в одному випробуванні.

Теорема. Імовірність появи хоча б однієї з двох сумісних подій дорівнює сумі імовірностей цих подій без імовірності їх сумісної появи:

$$P(A+B)=P(A)+P(B)-P(AB).$$

Теорема. Імовірність події А, яка може відбутись лише за умови появи однієї з несумісних подій $B_1, B_2, ..., B_n$, які створюють повну групу, дорівнює сумі добутків імовірностей кожної з цих подій на відповідну умовну імовірність події А:

$$P(A)=P(B_1)P_{B_1}(A)+P(B_2)P_{B_2}(A)+...+P(B_n)P_{B_n}(A).$$

Це рівняння називають *формулою повної імовірності*.

Формула Бейєса. Якщо подія А може відбутись лише за умови появи однієї з несумісних подій (гіпотез) $B_1, B_2, ..., B_n$, які складають повну групу подій, то після появи події А, то імовірності гіпотез можна визначити за формулами Бейєса:

$$P_A(B_i)=\frac{P(B_i)\cdot P_{B_i}(A)}{P(A)}\ (i=1,2,...,n),$$

де $P(A)=P(B_1)\cdot P_{B_1}(A)+P(B_2)\cdot P_{B_2}(A)+...+P(B_n)\cdot P_{B_n}(A).$

2.4 Повторення випробувань. Формула Бернуллі.

Якщо відбуваються випробування, при яких імовірність появи події А у кожному випробуванні не залежить від результатів інших випробувань, то такі випробування називають *незалежними* відносно події А.

Теорема. Імовірність появи хоча б однієї з подій $A_1, A_2, ..., A_n$ незалежних в сукупності, дорівнює різниці поміж одиницею і добутком імовірностей протилежних подій $\overline{A}_1, \overline{A}_2, ..., \overline{A}_n$

$$P(A) = 1 - P\left(\overline{A_1}\,\overline{A_2}\ldots\overline{A_n}\right).$$

Якщо події $A_1, A_2, \ldots, A_n$ мають однакову імовірність, що дорівнює p, то імовірність появи хоча б однієї з цих подій обчислюється за формулою:

$$P(A) = 1 - (1 - p)^n$$

Формула Бернуллі. Імовірність того, що у n незалежних випробуваннях, у кожному з яких імовірність появи події дорівнює p(0<p<1),подія відбудеться k разів (у будь-якій послідовності), дорівнює

$$P_n(k) = C_n^k p^k q^{n-k},$$

або

$$P_n(k) = \frac{n!}{k!(n-k)!} p^k q^{n-k},$$

де q=1-p.

3 ДИСКРЕТНІ ВИПАДКОВІ ВЕЛИЧИНИ

Дискретною називають випадкову величину, можливі значення якої – окремі ізольовані числа, які ця величина приймає з визначеними ймовірностями. *Законом розподілу* дискретної випадкової величини називають перелік її можливих значень і відповідних ним ймовірностей. Закон розподілу можна задавати у вигляді таблиці, аналітичним та графічним способами. Для вираження графічним способом у прямокутній системі координат строять крапки $M_1(x_1;p_1), M_2(x_2;p_2),...,M_n(x_n;p_n)$ (x_i - можливі значення X, p_i – відповідні імовірності), які поєднують відрізками прямих. Отриману фігуру називають *багатокутником розподілу*.

Біноміальним називають закон розподілу дискретної випадкової величини X- кількості появ події у n незалежних випробуваннях, у кожному з яких імовірність появи події – p. Імовірність можливого значення X=k (кількості k появ подій) розраховують за формулою Бернуллі. Якщо велика кількість випробувань, а імовірність з появи події у кожному випробуванні дуже мала, то використовують приблизну формулу

$$P_n(k) = \lambda^k e^{-\lambda} / k!,$$

де k – кількість появ події у n випробуваннях, $\lambda = np$ (середня кількість появ подій у n випробуваннях), і кажуть, що випадкова величина розподілена за законом *Пуассона*.

3.1 Найпростіший потік подій

Потоком подій називають послідовність подій, які відбуваються у випадкові моменти часу.

Найпростіший (пуассонівський) потік має такі властивості: *стаціонарність, відсутність післядії та ординарність.*

Стаціонарність означає, що з плином часу імовірнісні характеристики потоку не змінюються. Стаціонарність потоку рівносильна постійній щільності імовірності появи подій в будь-який момент часу, інакше кажучи, для стаціонарного потоку імовірність появи *k* подій за проміжок довжиною Δt залежить тільки від величини проміжку і не залежить від його розташування на вісі часу.

Відсутність післядії означає незалежність імовірносних характеристик потоку від попередніх подій. Тобто, імовірність появи *k* викликів за проміжок $[t_1,t_2]$ не залежить від появи подій до моменту t_1.

Ординарність означає практичну неможливість появи двох або більше подій за будь-який нескінченно малий проміжок часу Δt.

Формула Пуассона – імовірність появи k подій найпростішого потоку за час t:

$$P_t(k) = \frac{(\lambda t)^k \cdot e^{-\lambda t}}{k!},$$

де λ – інтенсивність потоку - середня кількість подій, які з’являються за одиницю часу.

3.2 Основні характеристики дискретних випадкових величин

Математичним очікуванням випадкової величини називають суму добутків усіх її можливих значень на їх імовірності:

$$M(X) = x_1 p_1 + x_2 p_2 + ... + x_n p_n = \sum_{i=1}^{n} x_i p_i.$$

Математичне очікування біноміального розподілу дорівнює добутку кількості випробувань на імовірність появи події в одному випробуванні:

$$M(X) = np.$$

Дисперсією випадкової величини X називають математичне очікування квадрат відхилення випадкової величини від її математичного очікування:

$$D(X) = M[X - M(X)]^2 = M(X)^2 - [M(X)]^2.$$

Дисперсія біноміального розподілу дорівнює добутку кількості випробувань на імовірність появи і не появи події в одному випробуванні:

$$D(X) = npq.$$

Властивості математичного очікування та дисперсії наведені у таблиці 1.1.

Таблиця 1.1 - Властивості математичного очікування та дисперсії

Математичне очікування	Дисперсія
1. Математичне очікування константи дорівнює самій константі: $M(C) = C$	1. Дисперсія константи дорівнює нулю: $D(C) = 0$
2. Постійний множник можна виносити за знак математичного очікування: $M(CX) = CM(X)$	2. Постійний множник можна виносити за знак дисперсії, попередньо возвести його в квадрат: $D(CX) = C^2 D(X)$
3. Математичне очікування добутку взаємно незалежних випадкових подій дорівнює добутку математичних очікувань множників: $M(X_1 X_2 ... X_n) = M(X_1) \cdot M(X_2) ... M(X_n)$	3. Дисперсія суми незалежних випадкових величин дорівнює сумі дисперсій доданків: $D(X_1 + X_2 + ... + X_n) =$ $= D(X_1) + D(X_2) + ... + D(X_n)$
4. Математичне очікування суми незалежних випадкових величин дорівнює сумі математичних очікувань доданків: $M(X_1 + X_2 + ... + X_n) =$ $= M(X_1) + M(X_2) + ... + M(X_n)$	

Середньо-квадратичне відхилення випадкової величини є корінь квадратний з дисперсії:

$$\sigma(X) = \sqrt{D(X)}.$$

3.3 Теоретичні моменти

Початковим моментом порядку k випадкової величини X називають математичне очікування величини X^k:

$$v_k = M(X^k).$$

Початковий момент першого порядку дорівнює математичному очікуванню:

$$v_1 = M(X).$$

Центральним моментом порядку k випадкової величини X називають математичне очікування величини $[X - M(X)]^k$:

$$\mu_k = M[X - M(X)]^k.$$

Центральний момент першого порядку дорівнює нулю:

$$\mu_1 = M[X - M(X)] = 0.$$

Центральний момент другого порядку дорівнює дисперсії:

$$\mu_2 = M[X - M(X)]^2 = D(X).$$

Центральні моменти зручно розраховувати використовуючи формули, які виражають центральні моменти крізь початкові:

$$\mu_2 = v_2 - v_1^2,$$
$$\mu_3 = v_3 - 3v_1 v_2 + 2v_1^3,$$
$$\mu_4 = v_4 - 4v_1 v_3 + 6v_1^2 v_2 - 3v_1^4.$$

4 ФУНКЦІЇ ТА ЩІЛЬНІСТЬ РОЗПОДІЛУ ІМОВІРНОСТЕЙ ВИПАДКОВИХ ВЕЛИЧИН

4.1 Функції та щільність імовірностей випадкових величин

Функцією розподілу називають функцію F(x), яка визначає для кожного значення *x* імовірність того, що випадкова величина X приймає значення, менше ніж *x,* тобто

$$F(x) = P(X < x).$$

Щільність розподілу імовірностей неперервної випадкової величини називають першу похідну від функції розподілу: $f(x) = F'(x)$.

Властивості функції розподілу і щільності розподілу наведено в таблиці 4.1.

Таблиця 4.1 - Властивості функції розподілу і щільності розподілу

Функція розподілу	Щільність розподілу
1. Значення функції розподілу належить відрізку [0;1]: $0 \le F(x) \le 1$.	1. Щільність розподілу додатна (позитивна): $f(x) \ge 0$.
2. Функція розподілу – не убиваюча функція: $F(x_2) \ge F(x_1)$, якщо $x_2 > x_1$.	2. Невласний інтеграл від щільності розподілу в межі від $-\infty$ до ∞ дорівнює одиниці: $\int_{-\infty}^{\infty} f(x)dx = 1$.
3. Якщо всі можливі значення випадкової величини X належать проміжку (a,b), то $F(x) = 0$ при $x \le a$; $F(x) = 1$ при $x \ge b$.	
4. Функція розподілу неперервна зліва: $\lim\limits_{x \uparrow x_0} F(x) = F(x_0)$.	

4.2 Характеристики неперервних випадкових величин

Математичне очікування неперервної випадкової величини X, можливі значення якої належать всій осі *Ox*, визначається рівнянням:

$$M(X) = \int_{-\infty}^{\infty} x f(x) dx,$$

де f(x) – щільність розподілу випадкової величини X.

Якщо всі можливі значення належать інтервалу (a,b), то

$$M(X) = \int_{a}^{b} x f(x) dx.$$

Модою $M_0(X)$ неперервної випадкової величини *X* називають те її можливе значення, якому відповідає локальний максимум щільності імовірності. Якщо розподілення має два однакових максимуми, то його називають *бімодальним*.

Медіаною $M_e(X)$ неперервної випадкової величини X називають те її можливе значення, яке визначається рівнянням:

$$P[X < M_e(X)] = P[X > M_e(X)].$$

Геометрично медіану можна пояснити як точку в якій ордината f(x) ділить площу навпіл, яка обмежена кривою розподілу.

Дисперсія неперервної випадкової величини X, можливі значення якої належать всій осі *Ox*, визначається рівнянням:

$$D(X) = \int_{-\infty}^{\infty} [x - M(X)]^2 f(x) dx.$$

Якщо всі можливі значення належать інтервалу (a,b), то

$$D(X) = \int_{a}^{b} [x - M(X)]^2 f(x) dx.$$

4.3 Види розподілу імовірностей неперервних випадкових величин

- *Рівномірним* називають розподіл імовірностей неперервної випадкової величини X, якщо на інтервалі (a,b), якому належать всі можливі значення X, щільність зберігає постійне значення, а саме $f(x) = \frac{1}{b - a}$; поза інтервалом f(x)=0.

- *Нормальним* називають розподіл імовірностей неперервної випадкової величини X, щільність якого має вигляд:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}},$$

де a – математичне очікування, σ – середньо-квадратичне відхилення X.

Імовірність того, що X приймає значення, які належать інтервалу (α, β),

$$P(\alpha < X < \beta) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right),$$

де $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{x^2}{2}} dx$ - функція Лапласа.

Імовірність того, що абсолютна величина відхилу менша за число δ:

$$P(|X - a| < \delta) = 2\Phi(\delta/\sigma).$$

При a=0:

$$P(|X| < \delta) = 2\Phi(\delta/\sigma).$$

Асиметрія, ексцес, мода і медіана нормального розподілу відповідно дорівнюють:

$$A_s = 0, E_k = 0, M_0 = a, M_e = a, \text{де } a = M(X).$$

- *Експонентним* законом називають розподіл імовірностей неперервної випадкової величини X, щільність якого має вид:

$$f(x) = \begin{cases} 0 \text{ при } x < 0 \\ \lambda e^{-\lambda x} \text{при } x \geq 0, \end{cases}$$

де λ – постійна позитивна величина.

Функція розподілу експонентного закону:

$$F(x) = \begin{cases} 0 & \text{при } x < 0 \\ 1 - e^{-\lambda x} & \text{при } x \geq 0. \end{cases}$$

Імовірність попадання в інтервал (a,b) неперервної випадкової величини X, яка розподілена за експонентним законом:

$$P(a < X < b) = e^{-\lambda a} - e^{-\lambda b}.$$

Математичне очікування, дисперсія та середньо-квадратичне відхилення експонентного розподілу відповідно дорівнюють:

$$M(X) = 1/\lambda,\ D(X) = 1/\lambda^2,\ \sigma(X) = 1/\lambda.$$

4.4 Функція надійності

Тривалість часу безвідмовної роботи елемента має експонентний розподіл, функція розподілу якого:

$$F(t) = P(T < t) = 1 - e^{-\lambda t}\ (\lambda > 0),$$

визначає імовірність відмови елемента за час t (λ – інтенсивність відмов, T- тривалість безвідмовної роботи елемента).

Функцією надійності R(t) називають функцію, яка визначає імовірність безвідмовної роботи елемента за час t:

$$R(t) = e^{-\lambda t}.$$

5 ВИБІРКОВИЙ МЕТОД

5.1 Статистичний закон розподілу вибірки

Нехай із генеральної сукупності взято вибірку. При цьому x_1 спостерігали n_1 раз, $x_2 - n_2$ раз, …, $x_k - n_k$ раз.

Варіантом називається значення ознаки x_i, що спостерігається.

Варіаційним рядом називається послідовність варіант, записана в порядку зростання або спадання.

Варіаційний ряд може бути дискретним або неперервним.

Частотою називається кількість раз, яку спостерігають ознаку x_i. Частота ознаки x_i позначається через n_i.

Відносною частотою називається відношення частоти до об'єма вибірки:

$$w_i = \frac{n_i}{n}.$$

Зауважимо, що об'єм вибірки дорівнює сумі частот:

$$n = \sum_{i=1}^{k} n_i .$$

Статистичним законом розподілу вибірки називається відповідність між значеннями досліджуваної ознаки та частотами або відносними частотами, з якими ця ознака з'являється.

Статистичний закон розподілу вибірки можна задавати

– таблицею,

– аналітично,

– графічно.

1.4 Табличне завдання статистичного закону розподілу вибірки

Табличне завдання статистичного закону розподілу вибірки здійснюється за допомогою таблиці, перший ряд якої містить перелік ознак, що приймає досліджувана ознака. Другий – відповідні ознакам частоти або відносні частоти.

x_i	x_1	x_2	$\dots$	x_k
n_i	n_1	n_2	$\dots$	n_k
w_i	w_1	w_2	$\dots$	w_k

Зауважимо, що у випадку, коли досліджувана ознака є неперервною величиною, статистичний закон розподілу вибірки записується у формі розподілу частинних інтервалів.

$x_i - x_{i+1}$	$x_1 - x_2$	$x_2 - x_3$	$\dots$	$x_k - x_{k+1}$
n_i	n_1	n_2	$\dots$	n_k

1.5 Емпірична функція розподілу

Нехай відомий статистичний розподіл частот кількісної ознаки X.

Позначимо через n_x число варіант, які менші за значення ознаки x, а через n – загальне число спостережень.

Відносна частота події $X < x$ дорівнює $\frac{n_x}{n}$. Якщо x змінюється, то змінюється і відносна частота. Отже, відносна частота є функцією від x. Ця функція знаходиться за допомогою випробування або емпіричним (розрахунковим) шляхом, тому називається *емпіричною функцією*.

Емпіричною функцією розподілу називається функція $F^*(x)$, яка для кожного значення x визначає відносну частоту події $X < x$:

$$F^*(x) = \frac{n_x}{n},$$

де: n_x – сума частот варіант, що менші за x;

n – об’єм вибірки.

Властивості емпіричної функції розподілу.

Властивість 1. Значення емпіричної функції належить відрізку $[0;1]$.

Властивість 2. $F^*(x)$ − не убутна функція.

Властивість 3. Якщо $x_{\min}$ − найменша варіанта, то $F^*(x)=0$, якщо $x \le x_{\min}$.

Якщо $x_{\max}$ − найбільша варіанта, то $F^*(x)=1$, якщо $x > x_{\max}$.

1.6 Графічне завдання статистичного закону розподілу вибірки

З метою наочності статистичний закон розподілу задають графічно за допомогою полігона та гістограми.

Полігоном частот називається ламана, відрізки якої поєднують точки з координатами (x_1, n_1), (x_2, n_2), …, (x_k, n_k), де x_i − варіанти вибірки, а n_i − відповідні частоти.

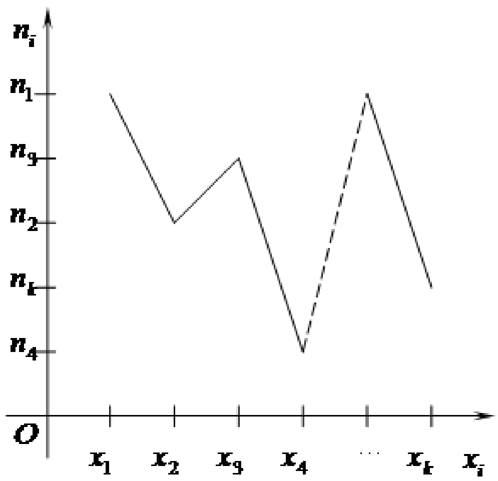

Рисунок 5.1 − Полігон частот

Полігоном відносних частот називається ламана, відрізки якої поєднують точки з координатами (x_1, w_1), (x_2, w_2), …, (x_k, w_k), де x_i − варіанти вибірки, а w_i − відповідні відносні частоти.

Якщо ознака розподілена неперервно, весь інтервал, в якому містяться всі спостережені значення ознаки, розбивають на ряд частинних інтервалів довжини $h = x_{i+1} - x_i$ і знаходять n_i − суму частот варіант, які потрапили до i-го інтервалу.

Гістограмою частот називається ступінчаста фігура, яка складається з прямокутників, основами яких є частинні інтервали довжиною h, а висоти дорівнюють відношенню $\frac{n_i}{h}$.

Відношення $\frac{n_i}{h}$ називається *щільністю частоти*.

Площа частинного i-го прямокутника дорівнює $h \cdot \frac{n_i}{h} = n_i$ − сумі частот варіант, що потрапили до i-го інтервалу. Площа гістограми частот дорівнює сумі всіх частот, тобто об’єму вибірки n.

Приклад. Побудувати гістограму частот вибірки.

$x_i - x_{i+1}$	1−5	5−9	9−13	13−17	17−21
n_i	10	20	50	12	8

Для заданої вибірки довжина частотних інтервалів $h = 4$. Отримуємо гістограму.

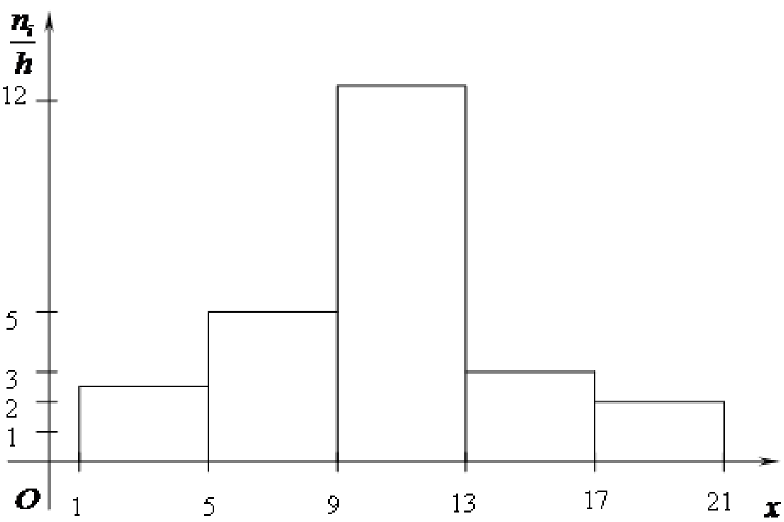

Рисунок 5.2 − Гістограма частот

Гістограмою відносних частот називається ступінчаста фігура, яка складається з прямокутників, основами яких є частинні інтервали довжиною h, а висоти дорівнюють відношенню $\frac{w_i}{h}$.

Відношення $\frac{w_i}{h}$ називається *щільністю відносної частоти*.

Площа частинного i-го прямокутника дорівнює $h \cdot \frac{w_i}{h} = w_i$ − відносній частоті варіант, що потрапили до i-го інтервалу. Площа гістограми відносних частот дорівнює сумі всіх відносних частот, тобто одиниці.

Зауважимо, що виразність гістограми суттєво залежить від обрання довжини h частинних інтервалів.

6 СТАТИСТИЧНІ ОЦІНКИ ПАРАМЕТРІВ РОЗПОДІЛУ

6.1 Точкові та інтервальні оцінки

Статистичною оцінкою невідомого параметра теоретичного розподілу називається функція від вибіркових значень (варіант), яка в певному статистичному сенсі є близькою до справжнього значення цього параметра.

Отже, знайти статистичну оцінку невідомого параметра теоретичного розподілу означає знайти функцію від вибіркових значень, яка дає наближене значення невідомого параметра.

Усі оцінки діляться на точкові і інтервальні.

Точковою називається оцінка, яка визначається одним числом.

Інтервальною називається оцінка, яка визначається двома числами − початком і кінцем інтервалу.

Зауважимо, якщо вибірка має малий об’єм, то точкова оцінка значно відрізняється від параметра, що оцінюється, тобто веде до значних похибок. Точкові оцінки застосовуються тоді, коли за їх допомогою можна провести ще і інші розрахунки. Точкові оцінки не несуть інформації про точність конкретної оцінки. Точність і надійність оцінки дозволяють визначити інтервальні оцінки.

6.2 Вимоги до статистичних оцінок

Позначимо θ^* оцінку деякого теоретичного параметра θ. Якщо розглянути вибіркові значення $x_1, x_2, \ldots, x_n$ як реалізації випадкових величин $X_1, X_2, \ldots, X_n$, що отримали конкретні значення в результаті випробувань, то можна оцінку θ^* представити функцією від випадкових величин $X_1, X_2, \ldots, X_n$:

$$\theta^* = \varphi(X_1, X_2, \ldots, X_n).$$

Для того, щоб статистична оцінка давала добре наближення невідомого параметра, вона має бути

− незміщеною,

– ефективною,

– ґрунтовною.

Незміщеною статистичною оцінкою θ^* невідомого параметра θ називається статистична оцінка, математичне сподівання якої дорівнює параметру, що оцінюється:

$$M(\theta^*) = \theta .$$

В противному разі, тобто коли

$$M(\theta^*) \neq \theta ,$$

точкова статистична оцінка θ^* називається *зміщеною* відносно параметра генеральної сукупності θ. Різниця $|\theta^* - \theta|$ називається *зміщенням* статистичної оцінки θ^*.

Ефективною статистичною оцінкою невідомого параметра називається статистична оцінка, яка при заданому об'ємі вибірки має найменшу можливу дисперсію.

Ґрунтовною називається статистична оцінка, яка в разі необмеженого збільшення обсягу вибірки $n \to \infty$ наближається до оцінювального параметра θ, а саме:

$$\lim_{n\to\infty} P(|\theta^* - \theta| < \varepsilon) = 1 .$$

Зауважимо, що

1) оцінка генеральної середньої вибірковою середньою

$$\overline{x_B} = \frac{n_1 x_1 + n_2 x_2 + \ldots + n_k x_k}{n} = \frac{\sum_{i=1}^{k} n_i x_i}{n}$$

є незміщеною і ґрунтовною, якщо вибірка повторна, і незміщеною для безповторної вибірки;

2) в якості оцінки генеральної дисперсії приймають виправлену вибіркову дисперсію:

$$S^2 = \frac{n}{n-1} D_B ,$$

$$D_B = \frac{\sum_{i=1}^{k} n_i (x_i - \overline{x_B})^2}{n} = \overline{x_B^2} - [\overline{x_B}]^2 .$$

Ця оцінка задовольняє умові незміщенності;

3) для оцінки середньоквадратичного відхилення генеральної сукупності використовують виправлене вибіркове середньоквадратичне відхилення, яке дорівнює кореню квадратному з виправленої дисперсії.

6.3 Точність оцінки та довірчий інтервал

Нехай отримана за даними вибірки статистична характеристика θ^* є оцінкою невідомого параметра θ. Нехай θ − постійне число (воно може бути і випадковою величиною). θ^* тим краще визначає параметр θ, чим менша абсолютна величина відхилення $|\theta - \theta^*|$. Тобто, якщо $\delta > 0$ і $|\theta - \theta^*| < \delta$, то чим менше δ, тим оцінка точніша. Отже, додатне число δ характеризує *точність оцінки*.

Надійністю (*довірчою імовірністю*) *оцінки* θ за θ^* називається імовірність γ, з якою виконується нерівність

$$|\theta - \theta^*| < \delta .$$

Звичайно, надійність оцінки задається наперед. В якості надійності беруть число, близьке до одиниці.

Нехай відома надійність γ, така, що

$$P(|\theta - \theta^*| < \delta) = \gamma .$$

Тоді

$$P(\theta^* - \delta < \theta < \theta^* + \delta) = \gamma .$$

Тобто імовірність того, що інтервал $(\theta^* - \delta; \theta^* + \delta)$ містить невідомий параметр θ, дорівнює γ.

Довірчим інтервалом називається інтервал $(\theta^* - \delta; \theta^* + \delta)$, який покриває невідомий параметр із заданою надійністю γ.

7 СТАТИСТИЧНА ПЕРЕВІРКА СТАТИСТИЧНИХ ГІПОТЕЗ

7.1 Статистичні гіпотези

Статистичною гіпотезою називається гіпотеза про вид невідомого розподілу або про параметри невідомих розподілів генеральної сукупності, яка перевіряється через вибірку.

Простою називається статистична гіпотеза, яка містить тільки одне припущення.

Складною називається гіпотеза, яка складається з скінченого або нескінченого числа простих гіпотез.

Нульовою (*основною*) називають висунуту гіпотезу H_0.

Конкуруючою (*альтернативною*) називається гіпотеза H_1, яка суперечить нульовій.

7.2 Помилки I та II роду

Висунута гіпотеза може бути як вірною, так і хибною. Для перевірки її правильності використовують статистичні дані і статистичні методи, тому перевірку називають *статистичною*.

Оскільки висновок про правильність гіпотези робиться за результатами скінченої вибірки, то завжди існує ризик прийняти хибне рішення. При цьому в результаті перевірки можуть бути допущені помилки двох родів.

Помилка першого роду полягає в тому, що буде відхилена правильна гіпотеза. Імовірність помилки першого роду називають *рівнем значущості* і позначають α. Зауважимо, що найчастіше рівень значущості приймає значення 0.01 або 0.05.

Помилка другого роду полягає в тому, що буде прийнята неправильна гіпотеза. Імовірність помилки другого роду позначають β.

Вірне рішення може бути прийняте також в двох випадках:

– гіпотеза приймається і вона дійсно правильна,

– гіпотеза відхиляється і вона дійсно невірна.

7.3 Статистичний критерій перевірки гіпотез

Для перевірки нульової гіпотези використовують спеціально підібрану випадкову величину, точний або наближений розподіл якої відомий.

Статистичним критерієм (*статистикою*) називається випадкова величина K, яка використовується для перевірки нульової гіпотези і закон розподілу якої відомий.

Для кожного конкретного випадку величина K спеціально підбирається і може позначатися різними літерами: U або Z для нормального розподілу, F для закону Фішера-Снедекора, K для закону Колмогорова і т.д.

Для перевірки гіпотези за даними вибірки обчислюють спостережуване значення критерію.

Спостережуваним (*емпіричним*) *значенням критерію* $K_{спост}$ називається значення критерію, обчислене за даними вибірки.

Критичною областю називається сукупність значень критерію, за яких нульову гіпотезу відхиляють.

Областю прийняття гіпотези (*областю допустимих значень*) називається сукупність значень критерію, за яких гіпотезу приймають.

Основний принцип перевірки статистичних гіпотез: якщо спостережуване значення критерію належить критичній області, то нульову гіпотезу відхиляють; якщо спостережуване значення критерію належить області прийняття гіпотези, то гіпотезу приймають.

Критичними точками $k_{кр}$ називають точки, які відокремлюють критичну область від області прийняття гіпотези.

7.4 Критерій Фішера-Снедекора

За двома незалежними вибірками, об'єми яких n_1 і n_2, вибраним з нормальних генеральних сукупностей, були знайдені виправлені вибіркові дисперсії s_x^2 і s_y^2. Необхідно порівняти ці дисперсії.

Правило 1. Для того, щоб при заданому рівні значущості α перевірити нульову гіпотезу H_0: $D(X)=D(Y)$ про рівність генеральних дисперсій нормальних сукупностей при конкуруючій гіпотезі H_1: $D(X)>D(Y)$, необхідно обчислити спостережуване значення критерію (відношення більшої виправленої дисперсії до меншої)

$$F_{спост}=\frac{s_б^2}{s_м^2}$$

і в таблиці критичних точок розподілу Фішера-Снедекора за заданим рівнем значущості α і числам степенів свободи $k_1=n_1-1$, $k_2=n_2-1$ (k_1 − число степенів свободи більшої виправленої дисперсії), знайти критичну точку $F_{кр}(\alpha,k_1,k_2)$.

Якщо $F_{спост}<F_{кр}$ нема підстав відхиляти нульову гіпотезу.

Якщо $F_{спост}>F_{кр}$ нульову гіпотезу відхиляють.

Правило 2. При конкуруючій гіпотезі H_1: $D(X)\neq D(Y)$ критичну точку $F_{кр}(\alpha/2,k_1,k_2)$ шукають за рівнем значущості $\alpha/2$ і числам степенів свободи k_1, k_2 (k_1 − число степенів свободи більшої виправленої дисперсії).

Якщо $F_{спост}<F_{кр}$ нема підстав відхиляти нульову гіпотезу.

Якщо $F_{спост}>F_{кр}$ нульову гіпотезу відхиляють.

Приклад. За двома незалежними вибірками, об'єми яких $n_1=11$ і $n_2=14$, вилучених з нормальних генеральних сукупностей X і Y, знайдені виправлені вибіркові дисперсії $s_X^2=0{,}76$ і $s_Y^2=0{,}38$. При рівні значущості $\alpha=0{,}05$, перевірити нульову гіпотезу H_0: $D(X)=D(Y)$ про рівність генеральних дисперсій, при конкуруючий гіпотезі H_1: $D(X)>D(Y)$.

Розв’язок. Знайдемо відношення більшої виправленої дисперсії до меншої:

$$F_{спост} = \frac{0,76}{0,38} = 2 .$$

В таблиці критичних точок розподілу Фішера-Снедекора, за рівнем значущості $\alpha = 0,05$ і числам степенів свободи $k_1 = 11 - 1 = 10$, $k_2 = 14 - 1 = 13$ знаходимо критичну точку

$$F_{кр}(0,05;10;13) = 2,67 .$$

Оскільки $F_{спост} < F_{кр}$ – вибіркові виправлені дисперсії відрізняються незначно.

7.5 Критерій згоди Пірсона

Критерій згоди Пірсона не доводить справедливість гіпотези, а тільки на прийнятому рівні значущості, встановлює її згідність або не згідність з даними спостережень.

За критерієм Пірсона перевіряють гіпотезу про нормальний розподіл генеральної сукупності. Для цього використовують *емпіричні* (спостережувані) і *теоретичні* (обчислені у зв’язку з гіпотезою про нормальний розподіл) *частоти*.

Нехай з генеральної сукупності взято вибірку і складено її статистичний закон розподілу

x_i	x_1	x_2	$\dots$	x_s
n_i	n_1	n_2	$\dots$	n_s

Необхідно, використовуючи критерій Пірсона перевірити гіпотезу про те, що генеральна сукупність X розподілена нормально.

Для того, щоб при заданому рівні значущості α, перевірити гіпотезу про нормальний розподіл генеральної сукупності, необхідно:

1. Обчислити безпосередньо (при малому числі спостережень) або спрощеним методом (при великій кількості спостережень), наприклад, методом

добутків або сум, вибіркову середню $\overline{x_B}$ і вибіркове середньоквадратичне відхилення σ_B.

2. Обчислити теоретичні частоти

$$n_i' = \frac{nh}{\sigma_B} \cdot \varphi(u_i),$$

де n − об'єм вибірки (сума всіх частот), h − крок (різниця між двома сусідніми варіантами), $u_i = \frac{x_i - \overline{x_B}}{\sigma_B}$, $\varphi(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}$.

3. Порівняти емпіричні і теоретичні частоти за допомогою критерію Пірсона. Для цього:

а) обчислюють спостережуване значення критерію

$$\chi^2_{спост} = \sum \frac{(n_i - n_i')^2}{n_i'};$$

б) за таблицею критичних точок розподілу χ^2 за заданим рівнем значущості α та числом степенів свободи $k = s - 3$ (s − число груп у вибірці) знаходять критичну точку $\chi^2_{кр}(\alpha, k)$.

Якщо $\chi^2_{спост} < \chi^2_{кр}$ − немає підстав відхилити гіпотезу про нормальний розподіл генеральної сукупності, тому її приймають.

Якщо $\chi^2_{спост} > \chi^2_{кр}$ − гіпотезу відхиляють.

Приклад. Використовуючи критерій Пірсона, при рівні значущості 0,05, перевірити, чи узгоджується гіпотеза про нормальний розподіл генеральної сукупності X з емпіричним розподілом вибірки об'єму $n = 200$.

x_i	5	7	9	11	13	15	17	19	21
n_i	15	26	25	30	26	21	24	20	13

Розв'язок. Обчислимо спостережуване значення критерію $\chi^2_{спост}$. Для цього складемо розрахункову таблицю:

№	n_i	n'_i	$n_i - n'_i$	$(n_i - n'_i)^2$	$\frac{(n_i - n'_i)^2}{n'_i}$
1	15	9,1	−5,9	34,81	3,8
2	26	16,5	9,5	90,25	3,6
3	25	25,3	−0,3	0,09	0,0
4	30	32,0	−2,0	4,00	0,1
5	26	33,9	−7,9	62,41	1,9
6	21	29,8	−8,8	77,44	2,3
7	24	22,0	2,0	4,00	0,2
8	20	13,5	6,5	42,25	3,0
9	13	7,0	6,0	36,00	5,1
$\sum$	200				$\chi^2_{спост} = 20,0$

Отримали $\chi^2_{спост} = 20,0$.

Обчислюємо число степенів свободи $k = s - 3 = 9 - 3 = 6$.

За таблицею розподілу χ^2 шукаємо $\chi^2_{кр}(\alpha, k) = \chi^2_{кр}(0,05;6) = 12,6$.

Оскільки $\chi^2_{спост} > \chi^2_{кр}$, гіпотезу про нормальний розподіл генеральної сукупності відхиляємо.

ПЕРЕЛІК ПОСИЛАНЬ

1. В.Е. Гмурман. Теория вероятностей и математическая статистика: Учебное пособие для вузов.- М.: Высш шк.,-2002 – 479 с.: ил.
2. В.Е. Гмурман. Руководство к решению задач по теории вероятностей и математичесокй статистике. М.; Высш шк.,--2003.-404 с.: ил.