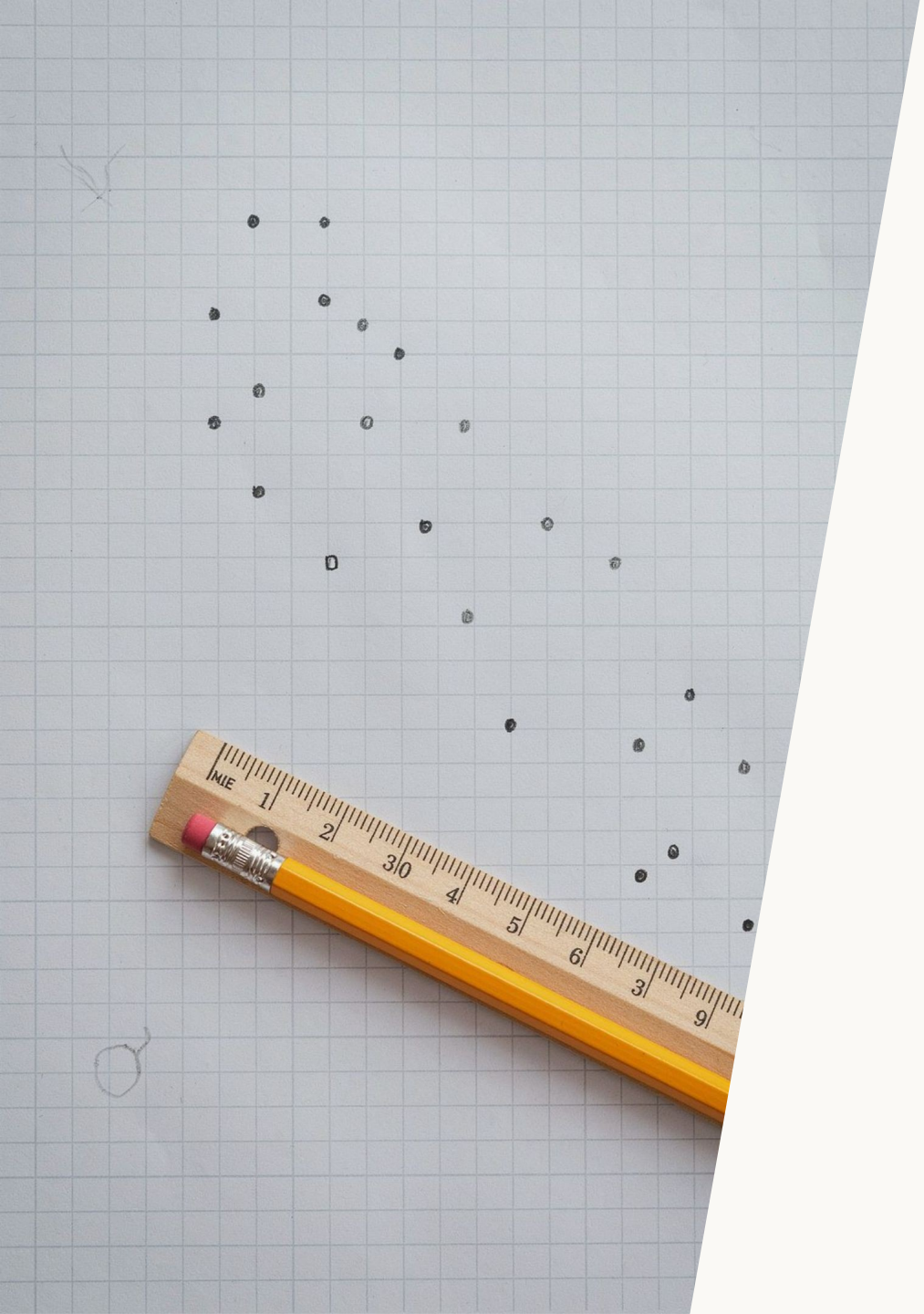
СӘТБАЕВ УНИВЕРСИТЕТІ

SATBAYEV UNIVERSITY

Статистика және деректерді талдау

Ержан Асел Ануаркызы,Ph.D., «Ақпараттық жүйелер» кафедрасының қауымдастырылған профессоры

Assel.yerzhan@satbayev.university



Шашырау көрсеткіштері

Деректер тек орташа санмен сипатталмайды — олардың қалай шашырайтынын білу маңызды шешімдердің негізі болып табылады.



Шашырау деген не?

Шашырау көрсеткіштері — өлшенетін мөлшерлердің ортадан қаншалықты алыс немесе жақын орналасқанын сандық түрде сипаттайтын статистикалық параметрлер.

Тек орташа мән берілген кезде деректер жайлы толық сурет алу мүмкін емес. Шашырауды білу арқылы деректің нақты құрылымы анықталады.

Орта жеткіліксіз

Орташа мән бір ғана нүктені сипаттайды, бірақ деректердің "жайылымын" көрсетпейді.

Шашырау — толықтырушы

Шашырау параметрлері деректің қаншалықты бірыңғай немесе алшақ екенін ашады.

Негізгі параметр 1: Орташа арифметикалық мөлшер ($\bar{x}$)



Орташа арифметикалық мөлшер барлық бақылау нәтижелерінің қосындысын олардың санына бөлу арқылы табылады:

$$\bar{x} = \frac{\sum_{i=1}^{N} x_i}{N}$$

Бұл параметр деректер тобының "орта жағдайын" анықтайды — шашырау орталығы ретінде қабылданады және барлық кейінгі есептеулердің негізі болады.

Негізгі параметр 2: Шашырау ауқымы (R)

Шашырау ауқымы (Range) — деректер жиынындағы ең үлкен мен ең кіші мөлшердің айырмасы:

$$R = x_{\max} - x_{\min}$$

R мәні деректің жалпы "жайылым шегін" жылдам бағалауға мүмкіндік береді. Алайда ол тек екі шеткі нүктеге сүйенетіндіктен, аралықтағы деректер туралы мәлімет бермейді.

Артықшылығы

Есептеу өте қарапайым, нәтиже бірден түсінікті.

Шектеуі

Тек шеткі мәндерге сүйенеді, аномалияларға сезімтал.



Негізгі параметр 3: Орташа квадратты ауытқу (σ)

Орташа квадратты ауытқу шашырауды ең дәл сипаттайтын параметр. *N* бойынша есептеледі:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N}}$$

1 **σ кіші болса**

Деректер орталыққа жинақталған, шашырау аз.

2 **σ үлкен болса**

Деректер кең таралған, нәтижелер алшақ.

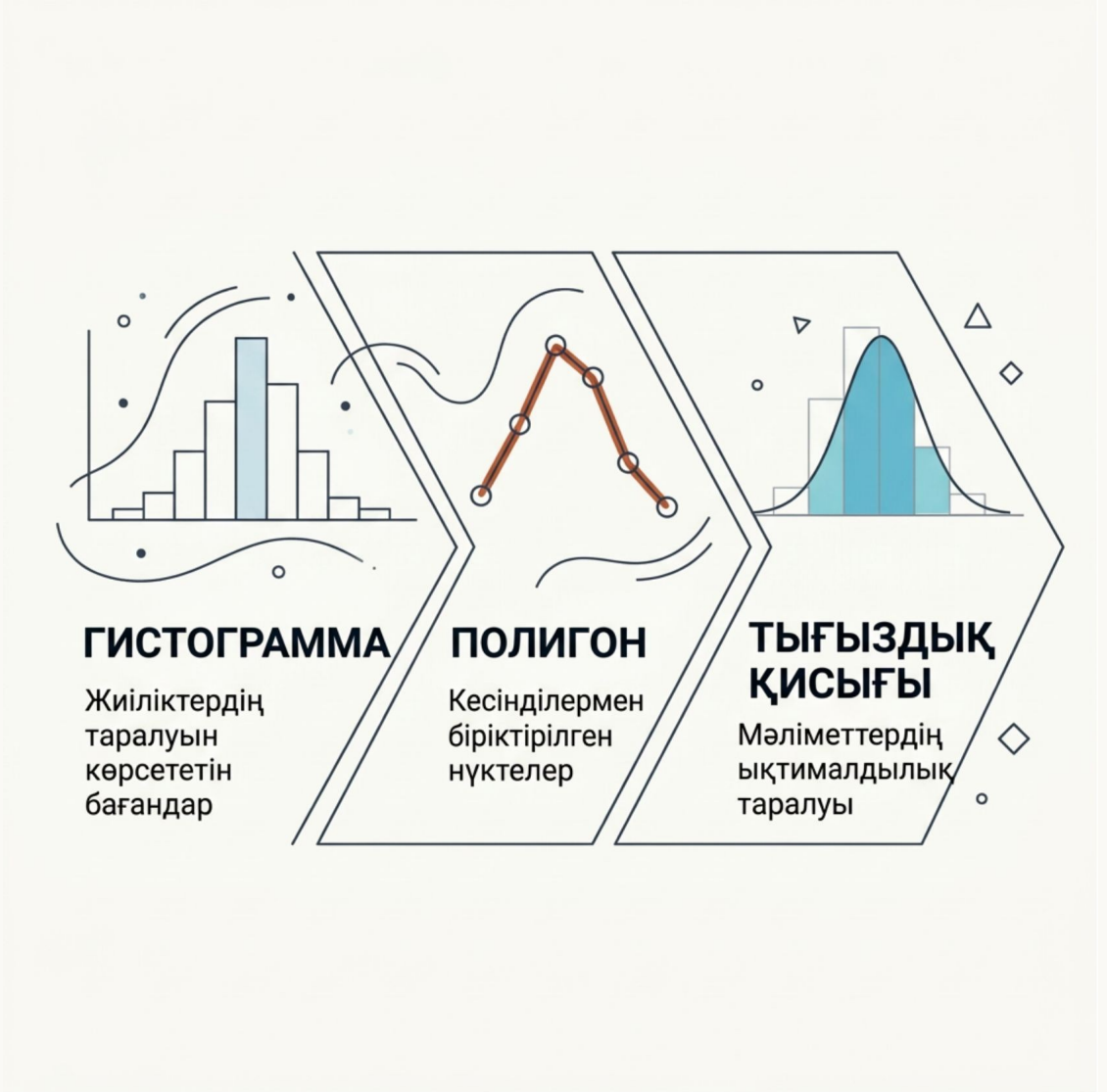
3 **N > 25 кезінде**

σ мәні тұрақтанып, сенімді нәтиже береді.

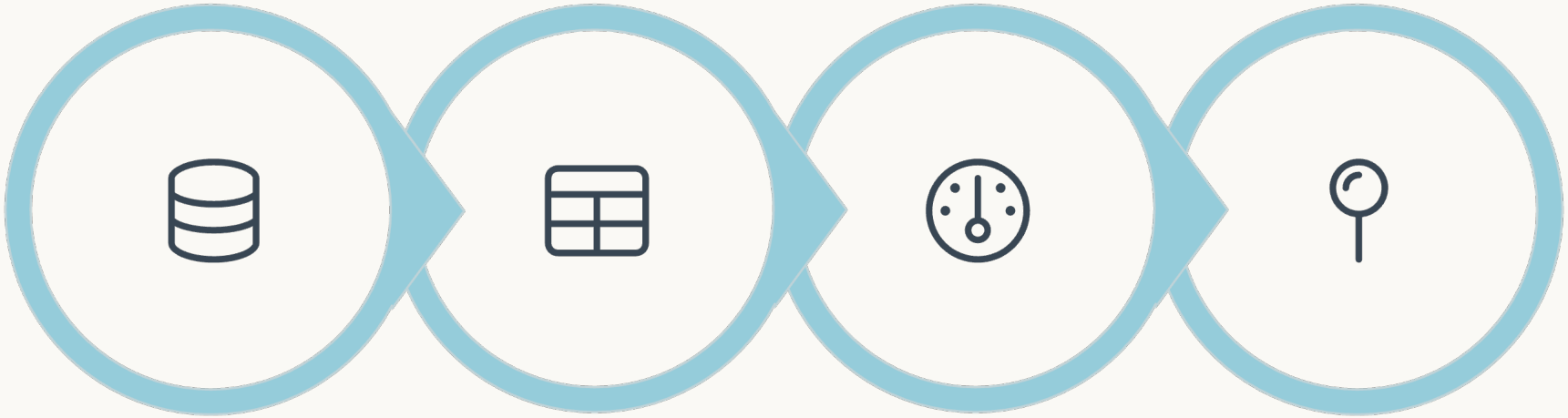
Қисық таралу: шашырауды графикпен көру

Шашырауды тек сандармен емес, графикалық түрде де бейнелеуге болады. Гистограмма, полигон және тығыздық қисығы — деректердің қалай таралғанын көзбен шолу құралдары.

Ординат өсіне теориялық ықтималдық тығыздығы орналасады. Графиктің астындағы аудан жиілік аралықтарына ықтималдыққа сәйкес келеді.



Шашырауды практикалық есептеу реті



Алғашқы жинау

Кесте жасау

Шашырау С анықтау

Координатты белгілеу

Есептеуді дұрыс бастау үшін деректерді жүйелі түрде кестеге орналастыру және нақты нүктені — шашырау басын — анықтау маңызды.

Есептеу блогы: орта, дисперсия және ауытқу

1

Орташа мән $\bar{t}$

Барлық бақылаулардың орталық нүктесі.

2

Орташа арифм. ауытқу ρ

Мәндердің ортадан абсолют алшақтығы.

3

Дисперсия D

Ауытқулардың квадраттық орташасы.

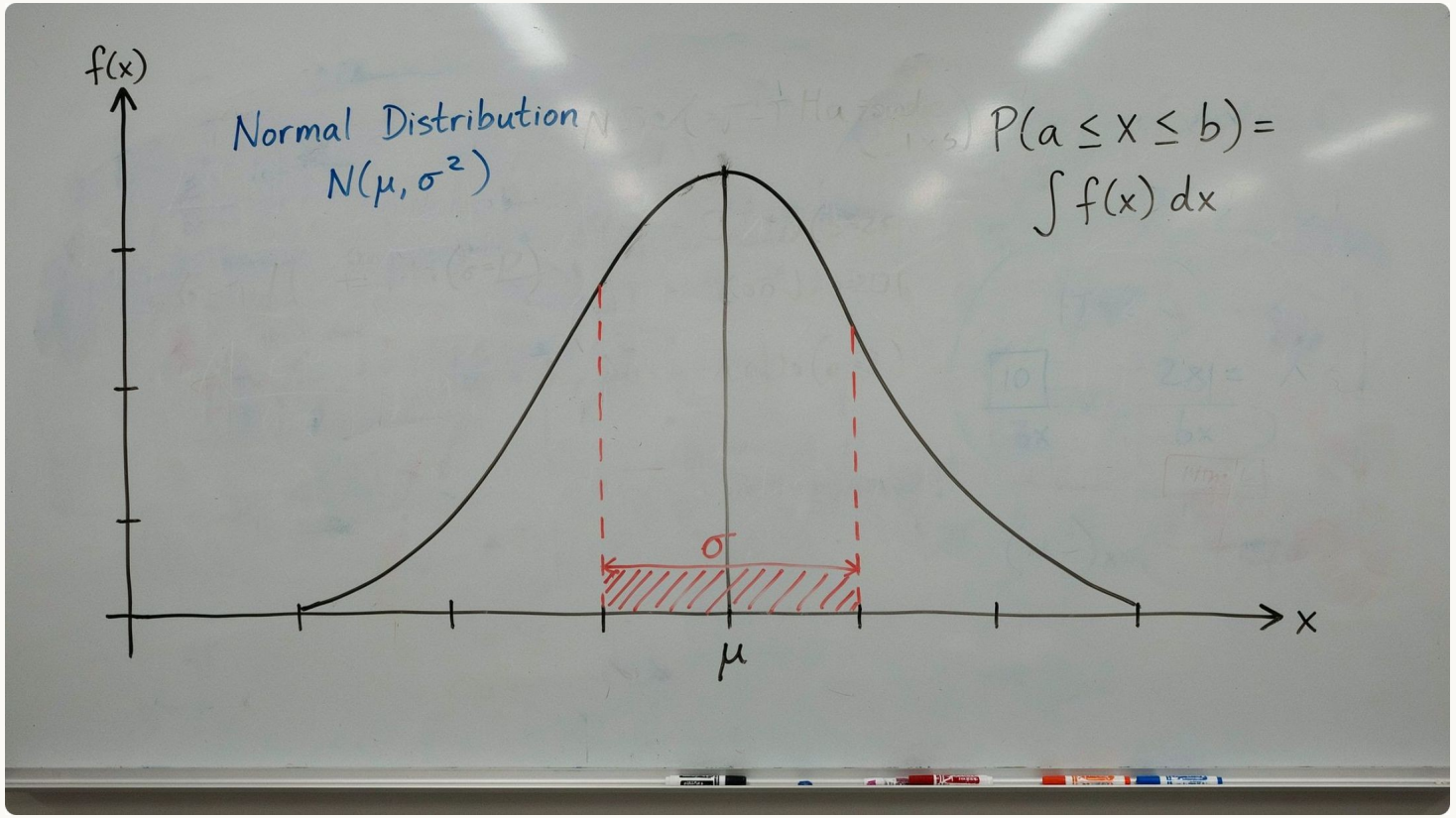
4

Орташа квадраттық ауытқу σ

Дисперсияның квадрат түбірі — негізгі параметр.



Тексеру және сәйкестендіру: график + теория



Есептеу аяқталғаннан кейін нәтижелерді графикалық және теориялық тұрғыдан **тексеру міндетті** қадам болып табылады.

Графиктер тұрғызу

Гистограмма, полигон, тәжірибелік ықтималдық қисығы салынады.

Вариация коэффициенті *v*

σ-ны орташаға бөліп, салыстырмалы шашырауды анықтайды.

Теориялық заң таңдау

f(t) және *F(t)* графиктері арқылы үлестіру заңы расталады.

Қорытынды: шашырау көрсеткіштері қалай шешім береді?

$\bar{x}$ — не күту керек?

Ортадан не күтілетінін, типтік нәтижені анықтайды.

R — айырма шегі

Деректердің жалпы жайылым шекарасын жылдам бағалайды.

σ — шашырау күші

Деректердің орташадан нақты алшақтығын дәл өлшейді.

Теория + тәжірибе

Екеуінің үйлесуі деректің сенімділік шекарасын дәлелдейді.

