

Understanding Variability in Statistics What is lasiability

differ from that conter.

the values in a classest are: It provides ensight into how much the number didfer from each other in simplex terms, variability measures the extent to which individual desta points in a distribution deviate from the lenter (such as the mean as median)

Key Torm

Variability, spread and dispossion all describe to some content Law spokad the data is Just like contral dependency discribes the "enter" of the doto vasigbility describes how much Individual data points

Range: A simple Measure of veriability The range is one of the most basic and simplest nearures of vasiability. It gives us an idea of ten spead out the values in a dataset by calculating the difference between the highest and lovest score

Limitation. The range only uses the extens values (highest and lonest) which reams it is very sensitive to outliers Is the obtaset has extreme values, the range can give a misleading inpression as how nuch the values are spead out.

Interquatile Perge (IBR)

The interquastile range is a measure of variability
that downer on the michle 50% of the other. The
IBR is particularly useful because it is not adjected
by the extreme values or outliers, making it a now
robust measure of variability compared to the
sange

How to calculate the Interquartile Range IBR = 75th poseentile - 25th percentile

Real world Example

Suppose you have the jost scores and students in the different classes, class A and class B. you calculate IBR for both classes

Class A IBR 4

Class B J 6R 8.

The class B IBP is higher, Indicating that the rejudile tois.

of the test scores in class B are nove spread out

compared to class A class A has less variability in its cert

scores, while class B shows greater variability

Summasy

I BP docuses on the speed of the middle (50%) of the difference it now so bust to putliers than the range.

The I BP is useful for understanding the configuration or dispossion of class without being influenced by extreme values.

1	
	Variance
	Vasiance is a measure of the spead of a distribution
	of data. It quantities how much the individual data points
	devide from the mean cos certes) of the distribution.
	A nigh resignce nears the date points one more speed
	out from the nean:
	Formula for variance
(D)	Population Vaciance $\sigma^2 = \{(\chi -)u)^2$
	. N
	or is the population variance
	& sepsesent injuisidual data point
	I is the population mean
	N is the total number of data point in the population
	, .
	How to compute versionce (Example)
	· FOS QUIZI data: 9,9,9,8,8,8,8,7,7,7,7,6,6,6,6,6,55
	Step1 Mean (21): 7.0
	step? compute the deviation from the mean afor ouch offa
	Point
	(9-7) (6-7)
	(9-7) $(6-7)$
	$(9-7) \qquad (5-7)$
	$(8-7) \qquad (5-7)$
_	(8-7)
	(8-4)
	(7-7)
	(6-7)
	$\begin{pmatrix} 6-7 \\ 6-7 \end{pmatrix}$
	(0-7)

the it

Step? Square Each deviation Soundes $2^2 = 4$ $3^2 = 1$ $0^2 = 0$ $(-1)^2 = 1$ $(-2)^2 = 4$

Step y find the avosage of squared deviation, by adding all squared deviation

Step5 Divide the total number of scores (N-20)

£1x->6) 2 - 30 - 15

·· Vosiance dos Quiz I is 15

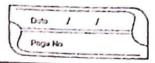
For a sample; lets assume we have a sample scare of 1,2,45

Sample Mean (M) = (1+2+4+5)/4 = 3

Step 1 : compute the deviation from the mean (1-3) = -7 (2-3) = -1 (4-3) = 1

(5-3) = 7

Step Squared each deviation

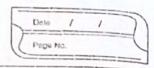


Step3 Lind the any of squared deviation.

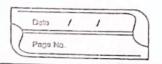
Add square deviation - Io Step4 Divido by N-I = 3 (sna + is a sample) $\frac{10}{3} \sim 3.33$ Tresedose to sample varsiance is appeximately 3.B key Concepts · Variance reasures how spread out the values in distribution · Re formula dos variance for a population uses (Nos the denominator, while for a sample, the formal was population vasiance Sample vasiance is slightly larger than population with the account for the fact that a sample might not separated the sull cline scity of a population Standard Doviation The standard Deviation is a newsure of the spreador dispossion of a set of dita points around the mean.

It provides a may of quentifying low much the individual points devide of som the mean of the distribution of the variance of its square root of the variance of its has the same unit of reasurement as the data points, which makes it casies to interpret Dompare

to the variance



	Formula dos standard deviation
	Formula dos standard deviation Population Standard deviation = \\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
	where or is the population standard deviation
	X sepresent sindividual data points
	It is the population Mean
	N is the total number od data points in the page
2	Sample standard deviation = - (= (X-M)>
- 1,	Sample standard deviation = = \{(X-M)>
	whos s is the sample shandred deviation
	X sepresents individual data points
1	Mis the somption
	N is the number of duta points in the cample
	How to compute standard doviction (Example).
	for Quil I Duta: 9,9,9,8,8,8,8,7,7,7,7,6,6,6,6,6,6,5,5,5,6,6,6,6,6,6,6
	step1 (alculate the vasiunce which is 1.5
	Step 2 Jake the square rock of variance to get the standard devigtion - VI5 - 7.257
	Stancard dovigtion - VID ~ 7.25+
	50 the Handred deviation for quiz I is approximately



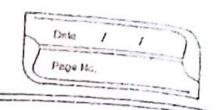
Exercises deviation is a undul neasure of unsightly because it is expressed in the same units as the desta it solid; matring it easies to interpret then unance of the larges the (sta) , the now spread out the data points are around the near

they Take anay: (Mean, Median and Mode)
The mean, median and mode are all neasures of
certical location, each answering the question "were in
the center of the data sel?

Mean: Provides the Asithmetic average. It is the most comen used measure but can be influenced by authors Or skored distribution

Median: The middle value when the dat it ordered. It is loss at fected by outliers and is often the bost chains of skew ad distributions

Mode: The most disequent value in the data. It is especially used for cote gorical data and situations when you've interested in identifying the most common value.



key Takenway (The range, Standord obviation, variance)

The range standord deviation and variance all

provide neasures of the variability or spoeded of a

data set. They help ansnex the question, "How

variable are the data:

- Range gies a simple neasure of how tax a post the maximum and minimum values are

and minimum values are

variance gies a neasure of how much each data

point differs from the near, but in squared units.

Standard deviation is the square root of the laxime and provides a neasure of spread in the Griginal units of the data