
 **QHN**
QUALITY HEALTH NETWORK

Variation – It Exists Everywhere!

Fred Koeman
North York
General
Hospital


QHN – Variation, It Exists Everywhere. Page: 1 Prepared by Fred Koeman, NYGH

 **QHN**
QUALITY HEALTH NETWORK

Learning Objectives

- Understand the concept and types of variation
- Learn the differences between quantitative and qualitative data
- Learn to construct frequency distributions
- Understand the measures of central tendency and dispersion
- Learn about the attributes of a normal distributions
- Making predictions with normal distributions

QHN – Variation, It Exists Everywhere. Page: 2 Prepared by Fred Koeman, NYGH

 **QHN**
QUALITY HEALTH NETWORK

Important!

**Understand
concepts,
don't memorize
formulas**

QHN – Variation, It Exists Everywhere. Page: 3 Prepared by Fred Koeman, NYGH

QHN **Concept of Variation**


- The smaller the variation the more consistent and predictable your process will be.
- Although variation can never be eliminated it does impact quality, costs, and customer satisfaction.
- No two objects, product or actions are alike, therefore, variation exists everywhere!
 - Service quality
 - Product performance
 - Process outputs

To control and reduce variation you must first understand, quantify, AND interpret variation in a data set.


QHN – Variation, It Exists Everywhere. Page: 4 Prepared by Fred Koeman, NYGH

QHN **Variation - What is it?**


Neither accurate or precise



Accurate but not precise



Accurate and precise



Reduced Variation →

QHN – Variation, It Exists Everywhere. Page: 5 Prepared by Fred Koeman, NYGH

QHN **Types of Variation**

<p>Common-Cause Variation</p> <p>Is always present in a process.</p> <p>A process with only common-cause variation is:</p> <ul style="list-style-type: none"> – controlled and consistent – stable and predictable over time – capable of performing within the acceptable range <p>Process changes are required to reduce common-cause variation.</p>	<p>Special-Cause Variation</p> <p>Is caused by unusual or external events.</p> <p>A process with special-cause variation is:</p> <ul style="list-style-type: none"> – uncontrolled and inconsistent – unstable and unpredictable – not capable of performing to requirements <p>Special-cause variation requires immediate action.</p>
---	--

QHN – Variation, It Exists Everywhere. Page: 6 Prepared by Fred Koeman, NYGH

QHN **Populations versus Samples**

Population:

- contains every possible observation or census
- difficult to capture the entire population in data collection

Sample:

- information from samples "infer" or approximate population characteristics
- samples, or subsets of populations are much easier to collect

Population Parameters
 μ = Population mean
 σ = Population standard deviation

Sample Statistics
 \bar{x} = Sample mean
 s = Sample standard deviation

QHN - Variation, It Exists Everywhere. Page: 7 Prepared by Fred Koeman, NYGH

QHN **Discrete Data**

- Data that is counted or classified and not measured on a scale
- Only a finite number of values is possible
- Values cannot be subdivided meaningfully
- The counts can be made up of either:
 - data either meets or fails to meet some criteria
 - yes/no
 - pass/fail
 - agree/disagree
 - the sum of a particular event
 - number of patient falls
 - number of C-sections
 - number of incorrect prescriptions

Also known as qualitative or attribute

QHN - Variation, It Exists Everywhere. Page: 8 Prepared by Fred Koeman, NYGH

QHN **Continuous Data**

- Information that can take be measured on a continuous scale
- Can take on almost any numeric value and be broken down into finer increments:
 - waiting time in minutes
 - body weight in kilograms
 - length of stay in days
 - temperature
 - cost
- Meaningful information exists between variables

Also known as quantitative or variable data

QHN - Variation, It Exists Everywhere. Page: 9 Prepared by Fred Koeman, NYGH

Distributions

Parts or activities vary within every process.

Size → Size → Size → Size →

But they form a pattern that, if stable, can be described as a distribution.

Size → Size → Size →

Distributions can differ in:

Location **Spread** **Shape**

Size → Size → Size →

QHN - Variation, It Exists Everywhere. Page: 16 Prepared by Fred Koeman, NYGH

Common Frequency Distribution Shapes

QHN - Variation, It Exists Everywhere. Page: 17 Prepared by Fred Koeman, NYGH


Histograms

Histograms are used to summarize data sets in terms of:

- **Shape**
 - the shape of the histogram can visually show key characteristics about the data
- **Central tendency**
 - the 3 measures of central tendency are the mean, median, and mode
- **Dispersion**
 - the 3 measures of dispersion are the range, variance, and standard deviation

Frequency Distributions are used to construct Histograms

QHN - Variation, It Exists Everywhere. Page: 18 Prepared by Fred Koeman, NYGH

QHN **Measures of Central Tendency** 


Measures of Central Tendency defines the location where the data is centered. The 3 common measures of central tendency are:

Mode: The data point that occurs with the most frequency

Median: The middle value in an ordered set of data. If the data set has an even number of observations, the median is the average of the middle two.

Mean: The average of all the data points which is equal to the sum of all the data points divided by the number of data points.

QHN - Variation, It Exists Everywhere. Page: 19 Prepared by Fred Koeman, NYGH

QHN **The 3 M's** 

Data set: 50, 51, 53, 40, 54, 51, 53, 42, 52, 53

The ordered data set is: 40, 42, 50, 51, 51, 52, 53, 53, 53, 54

Mode = 53

- Useful for qualitative type data and for identifying clusters; unimodal, bimodal, and multimodal. Limited use for any other data analysis.

Median = $51 + 52 / 2 = 51.5$


- Good estimator when extreme values are present.

Mean = $40 + 42 + 50 + 51 + 51 + 52 + 53 + 53 + 53 + 54 / 10 = 49.9$

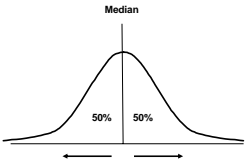
- Accounts for all values in a data set and is the basic building block for statistical analysis. Susceptible to extreme outliers in data sets.

In Excel use the formulas "MODE()", "Median()", "AVERAGE()"

QHN - Variation, It Exists Everywhere. Page: 20 Prepared by Fred Koeman, NYGH

QHN **The Median and Mean** 

- 50% of the area is to the left and right of the median
- For a symmetric histogram both the mean and the median are approximately the same
- The mean is more sensitive to outliers than the median
- Means from different data sets can be combined directly to get the overall mean
- All of the data sets are needed to get the median of the combined data



QHN - Variation, It Exists Everywhere. Page: 21 Prepared by Fred Koeman, NYGH

QHN **Impact of Outliers**

Data Set 4.2, 4.4, 5.1, 5.6, 6.0, 6.4, 6.8, 7.1, 7.4, 7.4, 7.9, 8.2, 8.2, 8.7, 9.1, 9.6, 9.6, 10.0, 10.5, 11.6

Mean = 7.69
Median = 7.65

By changing the last 2 values in the data set to 20.5 and 21.6 you get...

Median = 7.65
Mean = 8.89

QHN - Variation, It Exists Everywhere. Page: 22 Prepared by Fred Koeman, NYGH

QHN **Measures of Dispersion**

Dispersion is the degree to which the data tends to spread about the mean. The 3 most common measures of dispersion are:

Range: The difference between the smallest and largest observations in a data set.

Variance: The measure of how spread out a distribution is. It is calculated as the average squared deviation of each data point from the mean.

Standard Deviation: Calculated as the square root of the variance, it is the most commonly used measurement to quantify variation in a distribution. The standard deviation is in the same units as the data collected.

QHN - Variation, It Exists Everywhere. Page: 23 Prepared by Fred Koeman, NYGH

QHN **The Spread of a Distribution**

- The "spread" of a distribution is the degree to which the data tends to differ from each other compared to the mean.
- If all the data in a sample were about equal, the spread would be small. Conversely, data with large differences would result in a larger spread.
- The terms variability and dispersion are synonyms for spread.
- Standard deviation is the most common measurement of variation.**


Small Spread

Large Spread

Spread

QHN - Variation, It Exists Everywhere. Page: 24 Prepared by Fred Koeman, NYGH

QHN **Range**



Range is the difference between the largest and smallest values in a data set.

Range (R) = Xmax - Xmin


Attributes and Uses:

- Simplest measure of dispersion or variation.
- Does not account for each numerical value.
- Most affected by extreme scores.
- Affected by sample size, validity strongest for small data sets (fewer than ten.)
- Simple to calculate (no need for calculator or Excel.)
- Provides "rough" estimate of variation.

In Excel use the formula "MAX()" – "MIN()"

QHN – Variation, It Exists Everywhere. Page: 25 Prepared by Fred Koeman, NYGH

QHN **The Standard Deviation**



Standard Deviation, known as the standard measure of dispersion, is approximately the average distance a data set is from the mean.

Sample Standard Deviation (s)


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

Where:
n = A subset of the population
x_i = ith data value

In Excel use the formula "STDEV()"

QHN – Variation, It Exists Everywhere. Page: 26 Prepared by Fred Koeman, NYGH

QHN **Calculating Standard Deviation**



n	X	$\bar{X} - \bar{X}$	$(X - \bar{X})^2$
1	65	2	4
2	57	-6	36
3	69	6	36
4	58	-5	25
5	55	-8	64
6	67	4	16
7	63	0	0
8	58	-5	25
9	68	5	25
10	71	8	64
11	62	-1	1
Σ	693		296
\bar{X}	63		
S^2			29.60
S			5.44

Range (R) = 16

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)}$$

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

$$S = \sqrt{\frac{(63 - 60)^2 + (63 - 57)^2 + \dots + (71 - 63)^2 + (62 - 63)^2}{11 - 1}}$$

QHN – Variation, It Exists Everywhere. Page: 27 Prepared by Fred Koeman, NYGH

Predictions Using Z Distributions

From the ED Waiting Time Data calculate the % of patients that waited between 61 & 69 minutes:

$\bar{X} = 63, s = 5.44$

$\bar{X} = 63$

35.57% 13.57%

61 69

$$Z_L = \frac{\bar{X} - X}{s} = \frac{63 - 61}{5.44} = 0.3676 \qquad Z_U = \frac{X - \bar{X}}{s} = \frac{69 - 63}{5.44} = 1.102$$

From the Z tables: $P(Z_L < 0.3676) = 0.3557$ From the Z tables: $P(Z_U > 1.102) = 0.1357$

Therefore the % of patients that waited between 61 & 69 minutes is:

$$1 - (0.3557 + 0.1357) = 0.5086 = 50.86\%$$

QHN - Variation, It Exists Everywhere. Page: 31 Prepared by Fred Koeman, NYGH

Z-Score to Six Sigma

A Six Sigma process has six standard deviations between the mean and the closest specification limit for a short-term capability study.

$\bar{X} = 63$

LL UL

$63 - 6(5.44) = 30.36$ $63 + 6(5.44) = 95.64$

1 σ_{st}

6 σ_{st}

Minutes

Z Scale

QHN - Variation, It Exists Everywhere. Page: 32 Prepared by Fred Koeman, NYGH

The Six Sigma Measurement

Three Sigma Process

Mean = 0
St. Dev. = 2
Target = 12

Shift the Mean 1.5σ

By shifting a 3 sigma process 1.5σ results in 66,807 defects per million opportunities.

Six Sigma Process

Mean = 0
St. Dev. = 1
Target = 12

Shift the Mean 1.5σ

By shifting a 6 sigma process 1.5σ results in 3.4 defects per million opportunities.

QHN - Variation, It Exists Everywhere. Page: 33 Prepared by Fred Koeman, NYGH



My Contact Info



Fred Koeman
Corporate Performance Improvement Specialist
North York General Hospital
(416) 756-6898
fkoeman@nygh.on.ca
