Statistics Functions

THE NEW SCHOOL
UNIVERSITY LEARNING CENTER

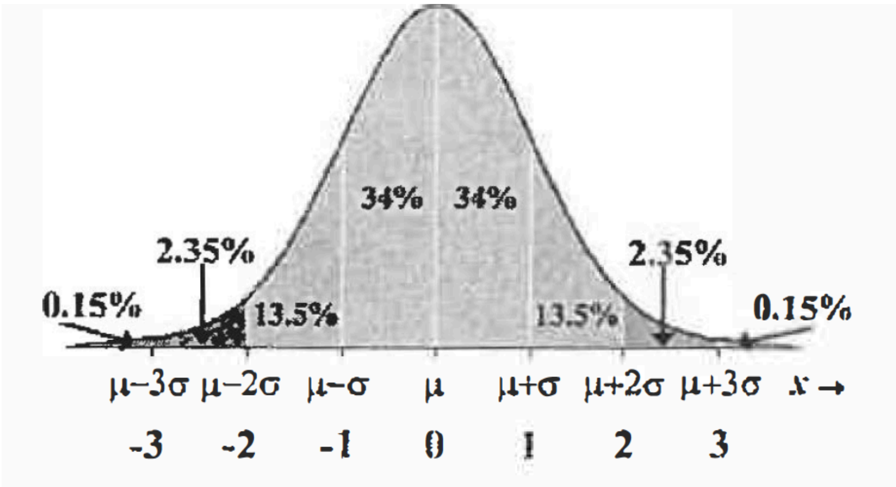
Statistics Formulas

Basic Terms, Notations & Formulas		
	Parameters (Population)	**Statististics (Sample**
Random Variable	X	x
Size	N	n
Maximum	highest value X	Highest value x
Minimum	lowest value X	Lowest value x
Range	maximum - minimum	
Mean	$\mu = \Sigma X/N$	$\bar{x} = \Sigma X/n$
Standard Deviation	$\sigma = \sqrt{\Sigma(X - \mu)^2/N}$	$s = \sqrt{\Sigma((x - \bar{x})^2/(n-1)}$
Standard Deviation Estimate	range/4	
Variance	$\sigma^2 = \Sigma(X - \mu)^2 /N$	$s^2 = \sqrt{\Sigma((x - \bar{x})^2/(n-1)}$
Z-Score (how many s.di's a value is away from the mean)	$z = (X - \mu)/\sigma$	$z = (x - \bar{x})/s$
*An easier way of thinking about **z-score***	→ Find the absolute difference between the value and the mean → See how many standard deviations are in that difference by dividing the number by the value of standard deviation → That's your z-score *When working with a mean for some sample, z=(x -u)/(o/vn)	
Proportion	P = (X/N)	$\hat{p} = (x/n)$

General & Empirical Rules - Normal Distribution

THE NEW SCHOOL

UNIVERSITY LEARNING CENTER



→ This *bell-shaped curv*e is a representation of a normally distributed data

Usual/unusual: A value that is within the 2-standard deviation boundaries is considered "usual"

→ As shown in the graph above, "usual" values take up 95% of a distribution

→ Unusual probability: $P(x) < 0.05$

Quartiles

5-Number Summary is composed of the 5 values: Minimum, Qi, Qz (Median), (4, and Maximum

Probability

P(A): Probability of event A occurring

P(A): Complement (the probability of event A NOT occurring} Addition rule: P(A or B) = P(A) + P(B) - P(A and B) Conditional probability: P(B I A) = P(A and B) / P(A) Permutation rule:

Random Variable

Size Maximum

highest value X

Highest value x **Minimum nnn**
lowest value X
Lowest value x Range
maximum - minimum **Mean**
H = *{X/*N
X = x/n **Standard Deviation**
o = VE(X - u)?/N
S = VE(x - x)*?*/(n-1) Standard Deviation Estimate
range / *4* **Variance**
o* = E(X - u) /N
s = VE(X-X)/n-1) Z-Score
z = (X - M*)/o*
z = (x-X*)/s An easier way of thinking abou**t z-score***
Proportion
P = (X/N)
P = (x/n*)*

General & Empirical Rules - Normal Distribution

→ This *bell-shaped curv*e is a representation of a normally distributed data
34%
34%
2.35%
2.35%
Usual/unusual: A value that is within the 2-standard deviation boundaries is considered "usual"
→ As shown in the graph above, "usual" values take up 95% of a distribution
→ Unusual probability: P(x) < 0.05
0.15%
13.5%
13.5%
0.15%

1-30 -20 -0
-3 -2 -1
1 0
leto 1+20 4+30 +

Statistics Functions

THE NEW SCHOOL
UNIVERSITY LEARNING CENTER

1 2 3

Quartiles **5-Number Summary** is composed of the 5 values**: Minimum, Q1, Q2 (Median), Q3, and Maximum**

Probability P(A): Probability of event A occurring
Range: 0 S **P(A) S 1** PĀ): Complement (the probability of event A NOT occurring) P(A) + P(Ā) = 1 Addition rule: P(A or B) = P(A) + P(B) - P(A and B)
Multiplication rule: P(A and B) = P(A) * P(B/*A*) Conditional probability: P(BIA) = P(A and B) / P(A)
At least-one probability: 1 - P(none) Permutation rule: Pr = n! / (n-r)!
Combination rule: Ce = n!/(n-r)!r! *Discrete Probability*
4 = {[x.P(x)]
o' = {[x?.P(x)] - H?
O = VE[x2.P(x)]-u? Expected value: E = {[x•P(x)] *Binomial Probability*
H = np
o = npq
0 **= Vnpa** Binomial probability formulas: P(x) = [n! / In-x)!x!] •p* .q*
P(x) = nC, p*.91**

Estimating Population Parameters

	Proportion p	**Mean (σ *known)***	***Mean (σ unknown)***
Sample	$\hat{p}$	$\bar{x}$	$\bar{x}$
Margin of Error	$E = z\alpha/2\sqrt{\hat{p}(1-\hat{p})/n}$	$E = z\alpha/2\sigma/\sqrt{n}$	$E = t\alpha/2\ (s/\sqrt{n}$
Confidence Interval	$\hat{p} - E < p < \hat{p} + E$	$\bar{x}-E<\mu<\bar{x}+E$	$\bar{x}-E<\mu<\bar{x}+E$

Statistics Functions

THE NEW SCHOOL
UNIVERSITY LEARNING CENTER

Sample Size	$n=[Z\alpha/2\ ]^2\ \hat{p}(1-\hat{p})/E^2$	$n=[Z_{a/2\sigma/E]}]^2$	
Test Statistic	$z=(\hat{p}-p)/(\sqrt{\ }pq/n)$	$z=(\bar{x}-\mu)/(\sigma/\sqrt{\ }n)$	
Critical Score	$z_{a/2}$ (2-tailed), z_a (1-tiled)	$t_{a/2\text{ or }}z_a$	$t_{a/2\text{ or }}t_a$ (Table A-3)

Hypothesis Testing Null hypothesis (Ho) Ho: p = 0.5 Ho: H = 98.6
Alternative hypothesis (H1) Hy: P = 0.5 Hiiu > 98.6 **Traditional Method (t**est statistic scores)
P-value method (area of the critical region) → Reject He if the test statistic falls within the critical region
→ Reject He if the p-value *sa* → Fail to reject H, if the test statistic does not fall within the critical region → Fail to reject Ho if the p-value > a