EDUR 8131
Chat 3
Notes 2 Normal Distribution and Standard Scores

Questions

Standard Scores: Z score

Z = (X – M) / SD

Z = deviation score divided by standard deviation

Z score indicates how far a raw score deviates from the sample mean in SD units.

Scores are 1, 2, 3, 4, 5, 6, 7, 8

M = 4.50
SD = 2.449

From SPSS

Descriptives

			Statistic	Std. Error
score	Mean		4.5000	.86603
	95% Confidence Interval for Mean	Lower Bound	2.4522	
		Upper Bound	6.5478	
	5% Trimmed Mean		4.5000	
	Median		4.5000	
	Variance		6.000	
	Std. Deviation		2.44949	
	Minimum		1.00	
	Maximum		8.00	
	Range		7.00	
	Interquartile Range		4.50	
	Skewness		.000	.752
	Kurtosis		-1.200	1.481

Find Z for a score of 7
Z = (X – M) / SD
M = 4.50
SD = 2.449

Z = (7 – 4.5) / 2.449 = 2.5 / 2.449

Z = 1.02

Interpretation:

The raw score of 7 is 1.02 standard deviations above the mean.

Find Z for a score of 4
Z = (X – M) / SD
M = 4.50
SD = 2.449

Z = (4 – 4.5) / 2.449 = -.5 / 2.449 = -.20
Z = -.20

Interpretation:

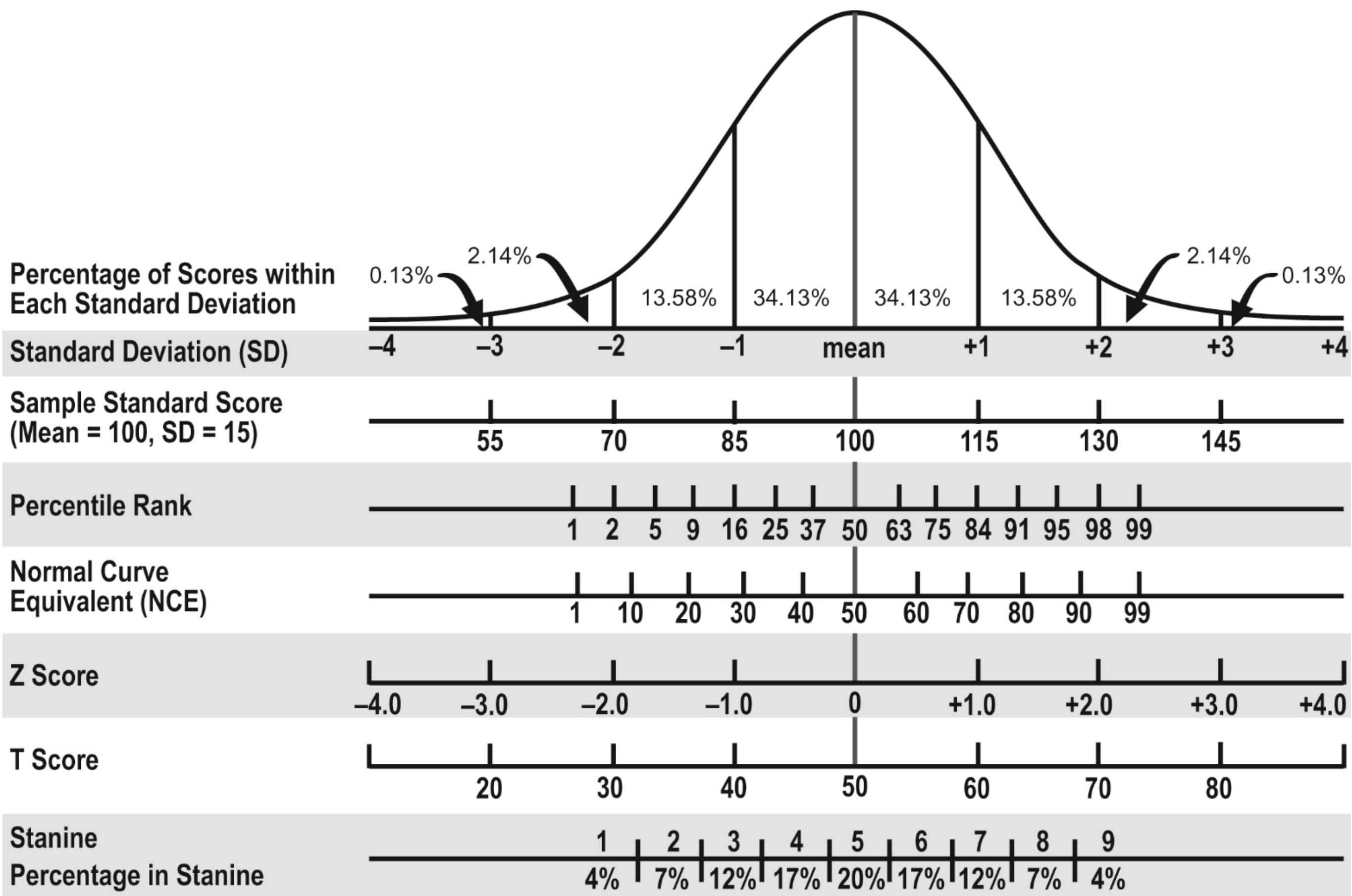
The raw score of 4 is -0.20 standard deviations below the mean.

Find Z for a score of 1
Z = (X – M) / SD
M = 4.50
SD = 2.449

Z = (1 – 4.5) / 2.449 = -3.5 / 2.449 = -1.429
Z = -1.429 or -1.43

Interpretation:
The raw score of 1 is -1.43 standard deviations below the mean.

Area under Z score (assuming normal distribution and using Z scores to find area)



Scores are Verbal SAT, minimum and maximum score of 200 to 800 with M = 500 and SD = 100

Step 1: Convert raw score X to Z score
Step 2: Find area below the Z score using table provided (or use your own table/calculator)

X = 600 (M = 500, SD = 100)

Z = (600-500) / 100 = 1

What proportion of scores are below Z = 1.00?

P(Z ≤ 1.00) = .8413

Also, what proportion of scores will be above Z = 1.00?

1 - .8413 = .1587

Suppose the Z = 1.05, what area is below this Z score?

P(Z ≤ 1.05) = .8531

Find in table like this:

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.00	.5000	.5040	.5080	.5120	.5160	[illegible]	.5239	.5279	.5319	.5359
0.10	.5398	.5438	.5478	.5517	.5557	[illegible]	.5636	.5675	.5714	.5753
0.20	.5793	.5832	.5871	.5910	.5948	[illegible]	.6026	.6064	.6103	.6141
0.30	.6179	.6217	.6255	.6293	.6331	[illegible]	.6406	.6443	.6480	.6517
0.40	.6554	.6591	.6628	.6664	.6700	[illegible]	.6772	.6808	.6844	.6879
0.50	.6915	.6950	.6985	.7019	.7054	[illegible]	.7123	.7157	.7190	.7224
0.60	.7257	.7291	.7324	.7357	.7389	[illegible]	.7454	.7486	.7517	.7549
0.70	.7580	.7611	.7642	.7673	.7704	[illegible]	.7764	.7794	.7823	.7852
0.80	.7881	.7910	.7939	.7967	.7995	[illegible]	.8051	.8078	.8106	.8133
[illegible]	.8159	.8186	.8212	.8238	[illegible]	[illegible]	.8315	.8340	.8365	.8389
1.00	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	.8531	.8554	.8577	.8599	.8621
[illegible]	.8643	.8665	.8686	.8708	[illegible]	[illegible]	.8770	.8790	.8810	.8830
1.20	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.30	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.40	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.50	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.60	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.70	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.80	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706

1.05

What proportion of scores below Z = -2.53?

P(Z ≤ -2.53) = .0057

What proportion of scores below Z = -0.32?

P(Z ≤ -0.32) = .3745

What proportion of scores above Z = -0.32?

P(Z ≥ -0.32) = 1 - .3745 = .6255

Percentile Ranks

Definition = percentage of scores at or below a given score

PR = 50 = 50% of scores are equal to or below this score

Two ways to calculate PR

(a) **Frequency Distribution** - If given a set of scores, use cumulative relative frequency to find PR

Example: 1 2 3 4 5 6 7 8

Percentile Ranks

Scores

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	1	12.5	12.5	12.5
	2.00	1	12.5	12.5	25.0
	3.00	1	12.5	12.5	37.5
	4.00	1	12.5	12.5	50.0
	5.00	1	12.5	12.5	62.5
	6.00	1	12.5	12.5	75.0
	7.00	1	12.5	12.5	87.5
	8.00	1	12.5	12.5	100.0
	Total	8	100.0	100.0	

X = 7, what is the corresponding PR?
PR = 87.5

X = 3, what is the PR?
PR = 37.5

(b) **Normal Distribution** assumed, find Z, then find area under Z, then multiply by 100 to obtain PR

Three Steps to Finding PR if Data are Normally Distributed:
1. Convert raw score to Z
2. Find area under Z (to left of Z)
3. Multiply area (proportion) by 100 to obtain PR

Example 1 with Verbal SAT
M = 500 and SD = 100

X = 600

Step 1: convert to Z score
Z = (600-500)/ 100 = 100/100 = 1.00

Step 2: find area below Z score of 1
P = .8413

Step 3: Multiply proportion by 100 to obtain PR

PR = 100 * .8413 = 84.13

Thus, 84.13% of verbal SAT test takers will score 600 or less.

Example 2 with Verbal SAT
M = 500 and SD = 100

X = 350

Step 1: convert to Z score
Z = (350-500)/ 100 = -150/100 = -1.50

Step 2: find area below Z score of -1.5
P = .0668

Step 3: Multiply proportion by 100 to obtain PR
PR = 100 * .0668 = 6.68

Thus, 6.68% of verbal SAT test takers will score 350 or less.

Convert from Z to X

X = M + (Z*SD)

Example 1
Verbal SAT M = 500, SD = 100

Z = -2.13

What is the corresponding SAT score?

X = M + (Z * SD)
X = 500 + (-2.13 * 100) =

X = 500 + (-2.13 * 100) = 287

Example 2
Test 1 in EDUR 8131: M = 86.59, SD = 9.78

Your Z score is 1.63
What is your test score?

X = M + (Z*SD)

X = 86.59 + 1.63*9.78
= 86.59 + 15.9414

= 102.53

Area between two Z scores
Step 1 = Convert both scores to Z scores
Step 2 = Find areas below both Z scores
Step 3 = Find difference between the two proportions subtracting smaller area from the larger area

What proportion of students obtain Verbal SAT scores between 450 and 550?

Recall that for verbal SAT the M and SD are:
M = 500
SD = 100

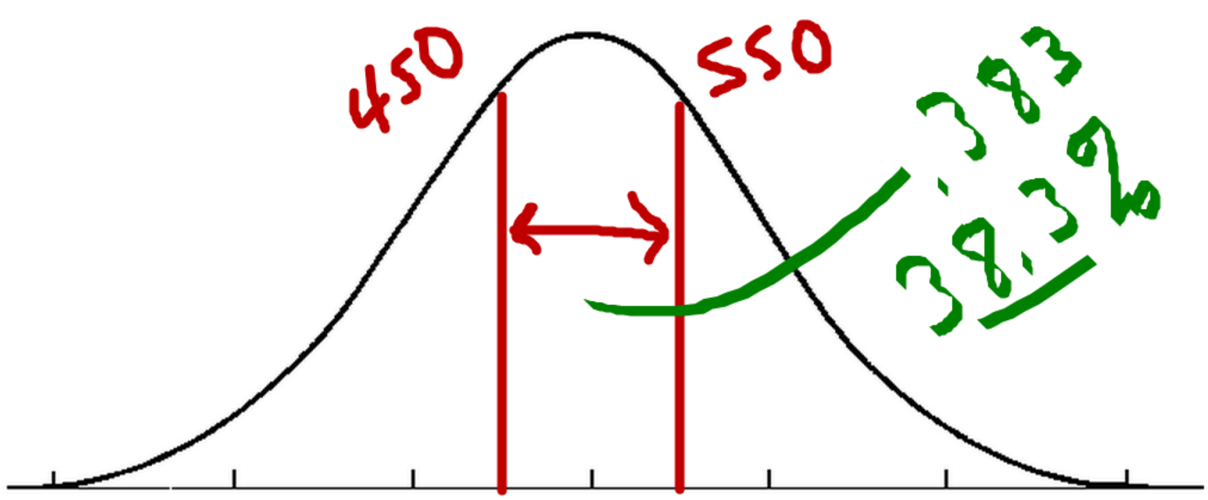
Step 1 = convert both scores to Z scores
450:
Z = (450 – 500) / 100 = -50 / 100 = -0.50
550:
Z = (550 – 500) / 100 = 50 / 100 = 0.50

Step 2 = Find areas below both Z scores
P(Z< .50) =
.6915

P(Z<-.50) =
.3085

Step 3 = Find difference between the two proportions subtracting small from larger
P(Z< .50) = .6915
P(Z<-.50) = .3085
Difference = .383

So about .383 or 38.3% of students will obtain verbal SAT scores between 450 and 550.



What proportion of students will obtain Verbal SAT scores between 500 and 400?
M = 500, SD = 100
(Compare with graphical chart of area)

Step 1 = convert both scores to Z scores
400:
Z = (400 – 500) / 100 = -100 / 100 = -1.00
500:
Z = (500 – 500) / 100 = 100 / 100 = 0.00

Step 2 = Find areas below both Z scores
P(Z< -1.00) =
.1587
P(Z< 0.00) =
.5000

Step 3 = Find difference between the two proportions subtracting small from larger
Difference = .5000 - .1587 = .3413

Population Distribution, Sample Distribution, and Sampling Distribution

Population – raw scores in population (census)
Sample – raw scores in sample taken from population
Sampling – distribution of a statistic taken from multiple samples

Population: Age of all 20 students in this class.

22	25	31	47	35	36	27	24	48	55
38	36	37	26	41	37	43	39	40	28

Mean = μ = 35.75

Sample: Age of students sampled from this class.

	1st Student Selected	2nd Student Selected	3rd Student Selected	Mean Age of Sampled Students, Mean	SD Age of Sampled Students, Standard Deviation
Sample 1:	22	25	31	**26**	**4.58**
Sample 2:	35	27	28	**30**	**4.36**
Sample 3:	48	55	38	**47**	**8.54**
Sample 4:	36	37	43	**38.67**	**3.79**
Sample 5:	47	24	36	**35.67**	**11.50**

Sampling Distribution: Distribution of statistic - Mean, SD, etc.

Sampling Distribution of the Sample Mean (M or $\overline{X}$)

Variance Error of Sample Mean: variance of sample means = $\sigma_{\bar{X}}^2 = \sigma^2/n$

Standard Error of Sample Mean: standard deviation of sample means = $\sigma_{\bar{X}} = \sqrt{\sigma^2/n} = \frac{\sigma}{\sqrt{n}}$

Why are estimates noted above referred to as error?

Statistics are estimates of parameters and have error (due to sampling error or bias). The variance error and standard error are estimates of how much error exists in the estimate.

Central Limit Theorem

If one is sampling independently from a population that has mean of μ and variance of σ^2, then as the sample size n approaches infinity, the sampling distribution of the sample mean $\overline{X}$ approaches normality without regard to the shape of the sampled population.

In short, as the sample size increases one may expect the sampling distribution of the mean to become normal in shape.

*Review illustration of central limit theorem from course web site

N = 2

M, M = 25% -----------------------> 475
M, F------------Mixed = 50%-----> 500
F, M
F, F = 25% --------------------------> 525

Males = 475
Females = 525