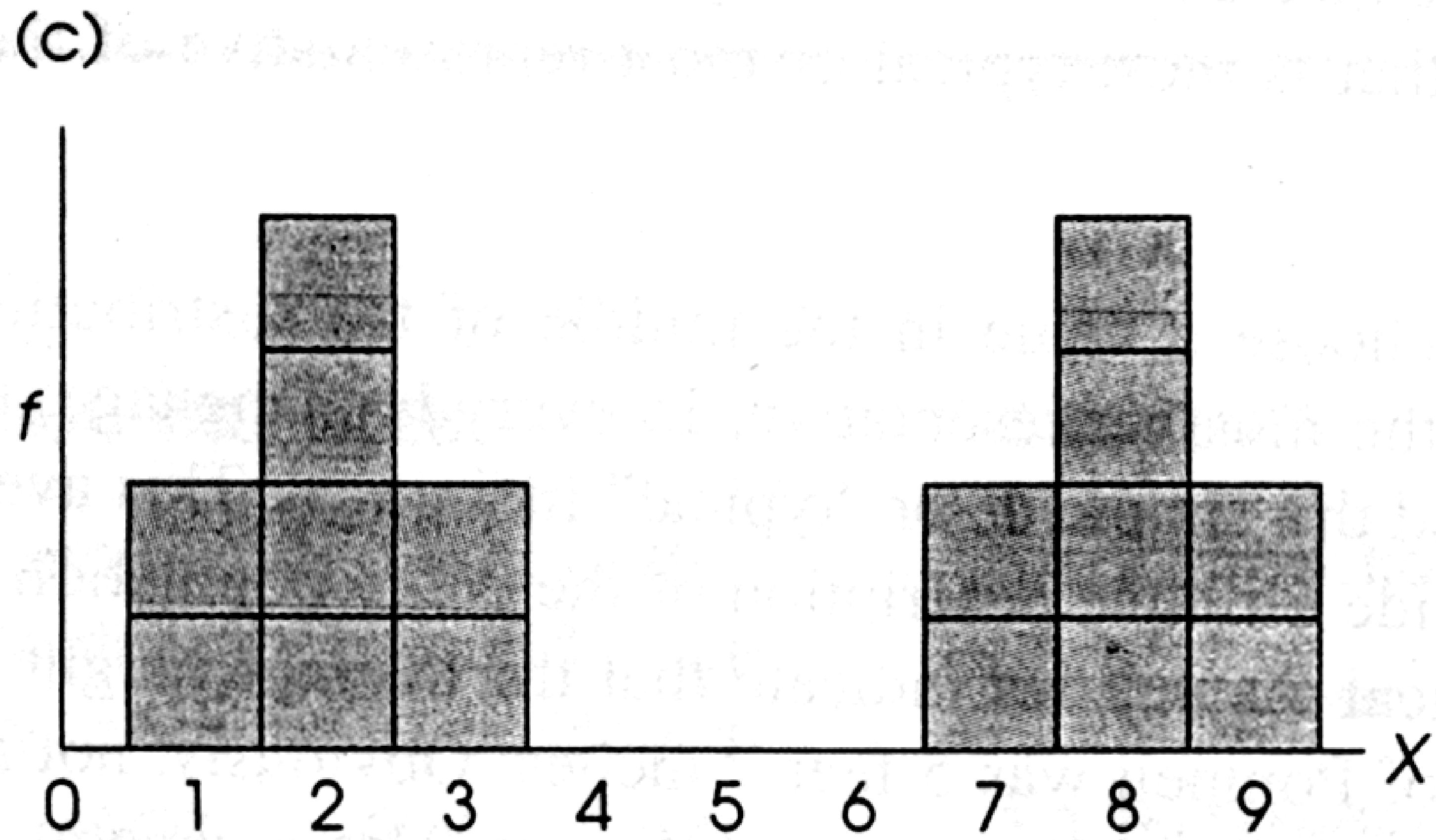
Measure of central tendency

- Central tendency
  - A statistical measure that identifies a single score as representative for an entire distribution. The goal of central tendency is to find the single score that is most typical or most representative of the entire group.

Measure of central tendency



Measure of central tendency

- The mean
  - Population mean vs. sample mean

$$\mu = \frac{\Sigma X}{N} \qquad \bar{x} = \frac{\Sigma X}{n}$$

  - N=4: 3,7,4,6

$$\bar{x} = \frac{\Sigma X}{n} = \frac{20}{4} = 5$$

Measure of central tendency

- The weighted mean
- Group A: $\overline{x} = 6$ n=12
- Group B: $\overline{x} = 7$ n=8
- Weighted mean = $\dfrac{\Sigma X_1 + \Sigma X_2}{n_1 + n_2}$ = $\dfrac{72 + 56}{12 + 8}$ = 6.4

- Seriously sensitive to extreme scores.

Measure of central tendency

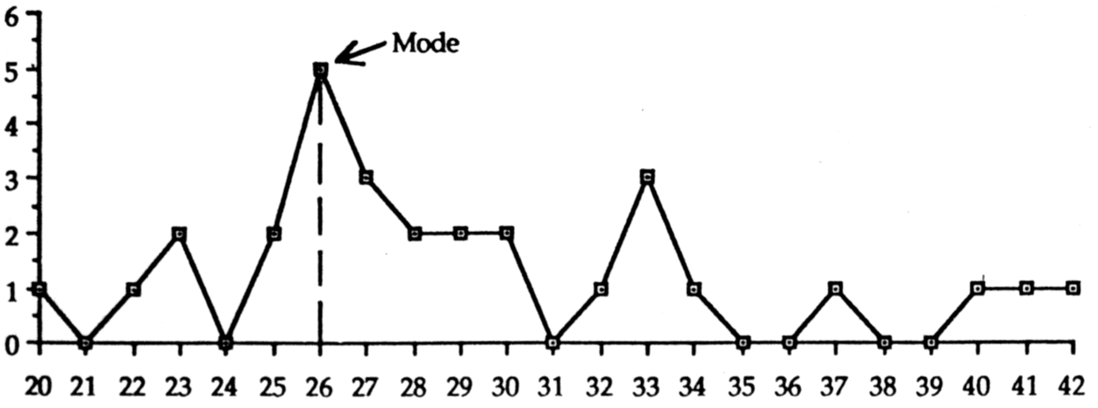
- Median
  - The score that divides a distribution exactly in half. Exactly 50 percent of the individuals in a distribution have scores at or below the median.
  - odd: 3, 5, 8, 10, 11 → median=8
  - even: 3, 3, 4, 5, 7, 8 → median=(4+5)/2=4.5

Measure of central tendency

- Median
  - The median is often used as a measure of central tendency when the number of scores is relatively small, when the data have been obtained by rank-order measurement, or when a mean score is not appropriate.

Measure of central tendency

- Mode
  - Most frequently obtained score in the data
  - Problems:
    - No mode

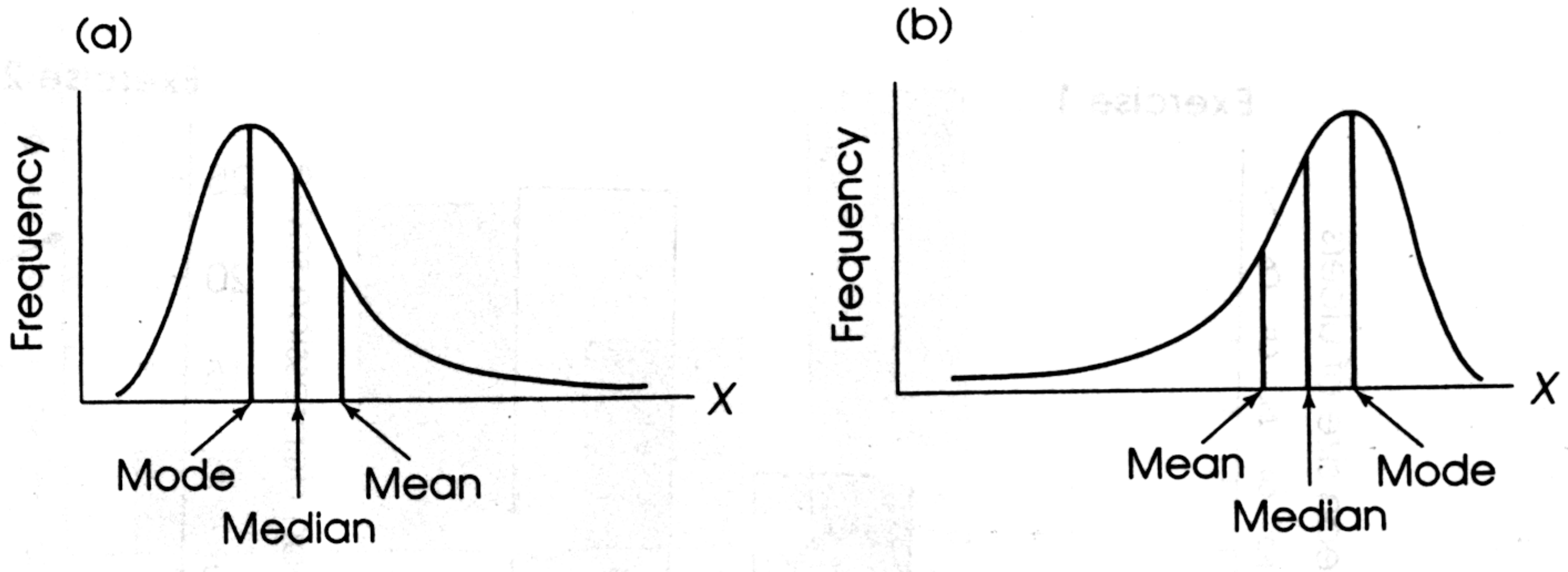


Measure of central tendency

- Choosing a measure of central tendency
  - the level of measurement of the variable concerned (nominal, ordinal, interval or ratio);
  - the shape of the frequency distribution;
  - what is to be done with the figure obtained.
  - The **mean** is really suitable only for **ratio and interval data**. For **ordinal variables**, where the data can be ranked but one cannot validly talk of `equal differences' between values, the **median**, which is based on ranking, may be used. Where it is not even possible to rank the data, as in the case of a **nominal variable**, the **mode** may be the only measure available.

Measure of central tendency

- Central tendency and the shape of the distribution



Summary

1. The purpose of central tendency is to determine the single value that best represents the entire distribution of scores. The three standard measures of central tendency are the mode, the median, and the mean.
2. The mean is the arithmetic average. It is computed by summing all the scores and then dividing by the number of scores. Conceptually, the mean is obtained by dividing the total (IX) equally among the number of individuals (N or n). Although the calculation is the same for a popula-tion or a sample mean, a population mean is identified by the symbol and a sample mean is identified by X.
3. Changing any score in the distribution will cause the mean to be changed. When a constant value is added to (or sub-tracted from) every score in a distribution, the same con-stant value is added to (or subtracted from) the mean. If every score is multiplied by a constant, the mean will be multiplied by the same constant. In nearly all circum-stances, the mean is the best representative value and is the preferred measure of central tendency.

Summary

1. The median is the value that divides a distribution exactly in half. The median is the preferred measure of central tendency when a distribution has a few extreme scores that displace the value of the mean. The median also is used when there are undetermined (infinite) scores that make it impossible to compute a mean.
2. The mode is the most frequently occurring score in a dis-tribution. It is easily located by finding the peak in a frequency distribution graph. For data measured on a nominal scale, the mode is the appropriate measure of central ten-dency. It is possible for a distribution to have more than one mode.
3. For symmetrical distributions, the mean will equal the me-dian. If there is only one mode, then it will have the same value, too.
4. For skewed distributions, the mode will be located toward the side where the scores pile up, and the mean will be pulled toward the extreme scores in the tail. The median will be located between these two values.

Homework

Imagine that you received the following data on the vocabulary test mentioned earlier:

20	22	23	23	23
23	23	23	24	25
28	29	30	30	30
30	30	30	31	32
32	33	33	34	35
35	36	36	37	37

1. Chart the data and draw the frequency polygon.
2. Compute the mean, mode, and median of the data and decide which of the three you believe to be best for the central tendency of the data.

Measure of variability

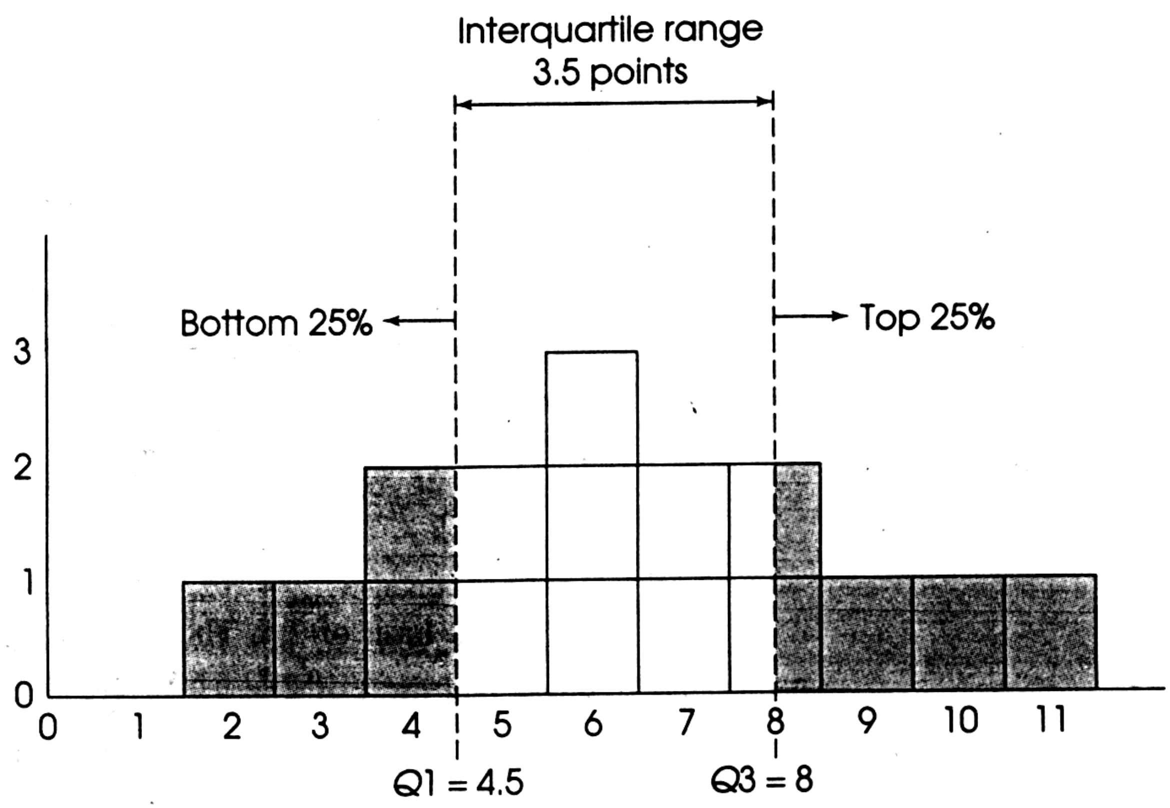
- Variability provides a quantitative measure of the degree to which scores in a distribution are spread out or clustered together.

Measure of variability

- Range
  - $range = X_{highest} - X_{lowest}$
  - Quartile:
    - A statistical term describing a division of observations into four defined intervals based upon the values of the data and how they compare to the entire set of observations.
      Each quartile contains 25% of the total observations. Generally, the data is ordered from smallest to largest with those observations falling below 25% of all the data analyzed allocated within the 1st quartile, observations falling between 25.1% and 50% and allocated in the 2nd quartile, then the observations falling between 51% and 75% allocated in the 3rd quartile, and finally the remaining observations allocated in the 4th quartile.
  - Interquartile: The interquartile range is a measure of spread or dispersion. It is the difference between the 75th percentile (often called Q3) and the 25th percentile (Q1). The formula for interquartile range is therefore: Q3-Q1.
  - Semi-interquartile: The semi-interquartile range is a measure of spread or dispersion. It is computed as one half the difference between the 75th percentile [often called (Q3)] and the 25th percentile (Q1). The formula for semi-interquartile range is therefore: (Q3-Q1)/2.
  - TOEFL: (560-470)/2=45

Measure of variability

Frequency distribution for a population of $N = 16$ scores. The first quartile is $Q1 = 4.5$. The third quartile is $Q3 = 8.0$. The interquartile range is 3.5 points. Note that the third quartile ($Q3$) divides the two boxes at $X = 8$ exactly in half so that a total of 4 boxes is above $Q3$ and 12 boxes are below it.



Measure of variability

- Variance
  - Deviation: deviation of one score from the mean
  - Variance: taking the distribution of all scores into account.

$$\sigma^2 = \frac{\Sigma(X - \mu)^2}{N} \longrightarrow s^2 = \frac{\Sigma(X - M)^2}{N-1}$$

	SCORES	SORTED SCORES	SCORES-MEAN	SQUARED SCORES
	15	23	6.33	40.11
	11	22	5.33	28.44
	19	20	3.33	11.11
	13	19	2.33	5.44
	16	19	2.33	5.44
	16	18	1.33	1.78
	17	18	1.33	1.78
	14	18	1.33	1.78
	22	17	0.33	0.11
	15	17	0.33	0.11
	16	17	0.33	0.11
	19	17	0.33	0.11
	18	16	-0.67	0.44
	12	16	-0.67	0.44
	16	16	-0.67	0.44
	15	16	-0.67	0.44
	18	16	-0.67	0.44
	20	15	-1.67	2.78
	23	15	-1.67	2.78
	18	15	[illegible]	2.78
	17	14	[illegible]	7.11
	17	13	[illegible]	13.44
	17	12	[illegible]	21.78
	16	11	-5.67	32.11
SUM=	400	400	0.00	181.33

Sum of square (SS)

n=24

MEAN= 16.67
VARIANCE= 7.56
SD= 2.75

Note: Mean + SD = 19.42
Mean − SD = 13.92
Mean + 2SD = 22.16
Mean − 2SD = 11.17

Measure of variability

- Standard deviation

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

score	mean	deviation*	squared deviation
8	9.67	- 1.67	2.79
25	9.67	+15.33	235.01
7	9.67	- 2.67	7.13
5	9.67	- 4.67	21.81
8	9.67	- 1.67	2.79
3	9.67	- 6.67	44.49
10	9.67	+ .33	.11
12	9.67	+ 2.33	5.43
9	9.67	- .67	.45
	sum of	squared dev=	320.01

Standard Deviation = Square root (sum of squared deviations / (N-1)

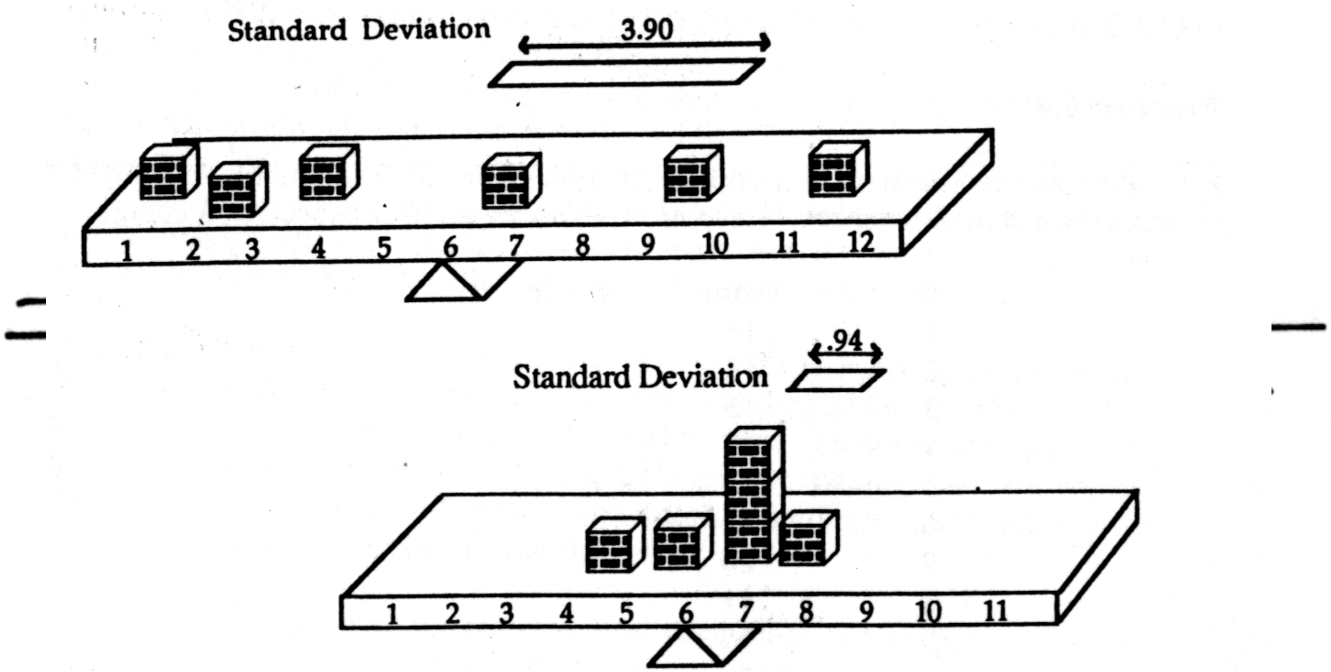
= Square root (320.01/(9-1))

= Square root (40)

= 6.32

Measure of variability

- The larger the standard deviation figure, the wider the range of distribution away from the measure of central tendency



Measure of variability

- Adding a constant to each score does not change the standard deviation.
- Multiplying each score by a constant causes the standard deviation to be multiplied by the same constant.

Measure of variability

Group A	Group B
11	20
8	10
10	1
9	8
8	0
12	30
10	13
11	6

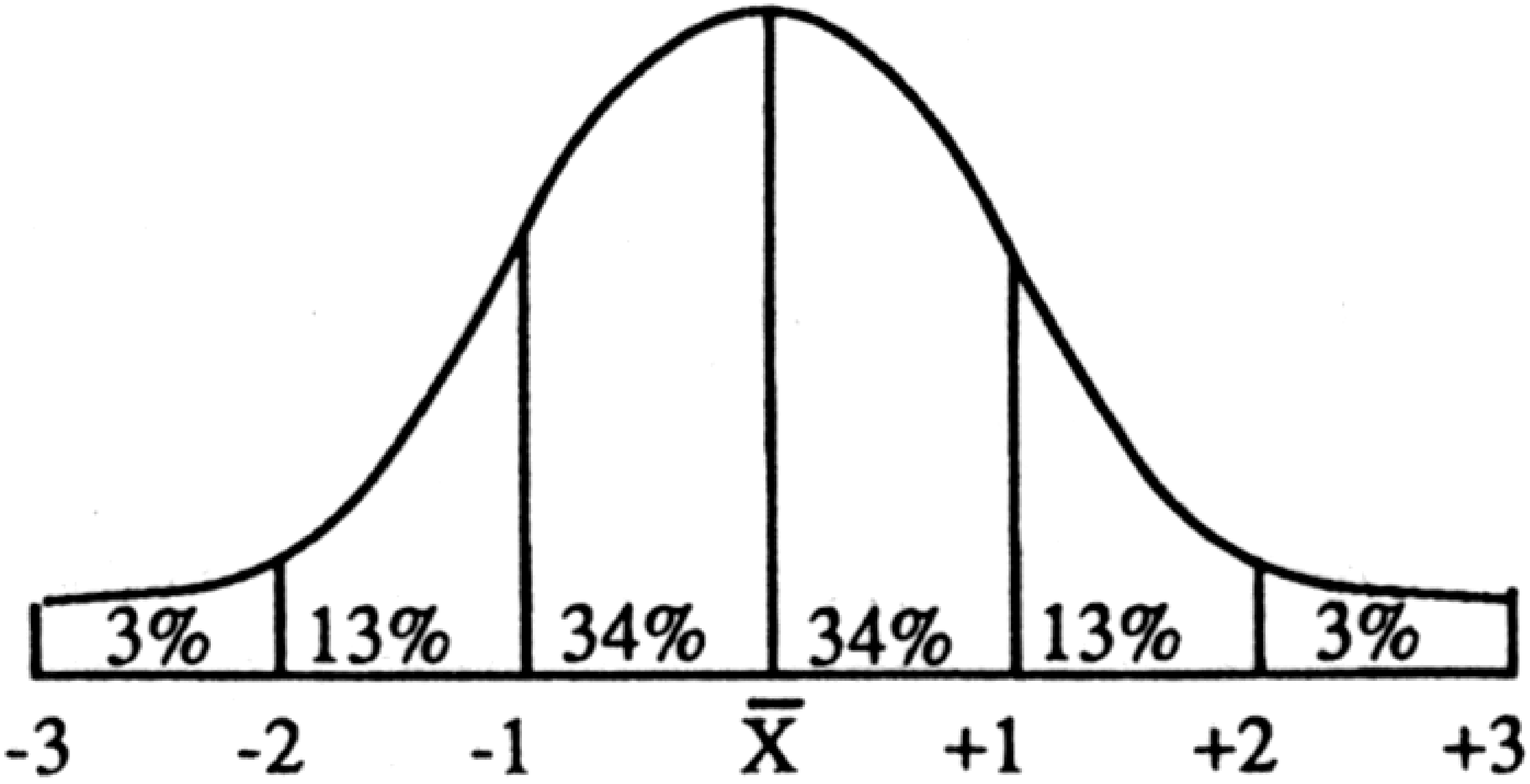
Measure of variability

Reporting the standard deviation (APA):

	Type of instrument			
	Listening		Watching	
	Mean	SD	Mean	SD
Males	15.72	4.43	6.94	2.26
Females	3.47	1.12	2.61	0.98

Measure of variability

- Standard deviation and normal distribution



Homework

1. Calculate the mean, median, mode, range and standard deviation for the following sample:

Midterm Exam

X	X
100	85
88	82
83	96
105	107
78	102
98	113
126	94
85	119
67	91
88	100
88	72
77	88
114	85

Homework

2. Suppose that the following scores were obtained on administering a language proficiency test to ten aphasics who had undergone a course of treatment, and ten otherwise similar aphasics who had not undergone the treatment:

Experimental group	*Control group*
15	31
28	34
62	47
17	41
31	28
58	54
45	36
11	38
76	45
43	32

Calculate the mean score and standard deviation for each group, and comment on the results.

Locating scores and finding scales in a distribution

TOEFL SCORE COMPARISON TABLE
(based on the score of 759,768 examinees
who took the test from July 1985 through June 1987)

TOTAL		SECTION SCORES					
		Sec 1		Sec 2		Sec 3	
Your Score	%ile lower	Your Score	%ile lower	Your Score	%ile lower	Your Score	%ile lower
660	99	66	98	66	97	66	98
640	97	64	95	64	94	64	96
620	93	62	90	62	90	62	92
600	89	60	85	60	85	60	87
580	83	58	78	58	77	58	80
560	74	56	71	56	69	56	71
540	64	54	62	54	59	54	61
520	52	52	52	52	50	52	50
500	41	50	41	50	40	50	39
480	29	48	30	48	31	48	29
460	20	46	20	46	22	46	22
440	13	44	12	44	15	44	15
420	8	42	7	42	10	42	10
400	4	40	4	40	7	40	6
380	2	38	2	38	4	38	4
360	1	36	1	36	2	36	2
340		34		34	1	34	1
320		32		32	1	32	1
300		30		30		30	

Percentiles, quartiles, deciles

	Score	Frequency (f)	Relative freq.	Cumulative freq. (F)	%ile
	50	6	.08	75	96
→	40	18	.24	69	80
	30	27	.36	51	50
	20	18	.24	24	20
	10	6	.08	6	4

$$Percentile = (100)\frac{\text{no. below} + 1/2 \text{ same}}{N}$$

$$= (100)\frac{51 + 1/2(18)}{75}$$

$$= 80$$

Mind work

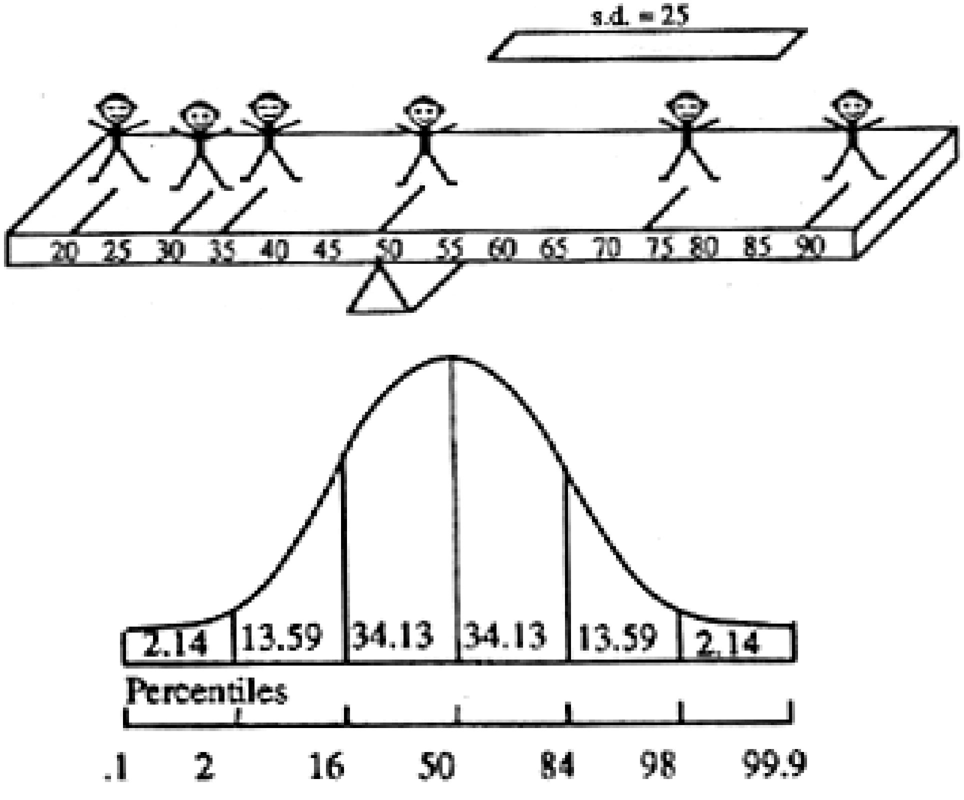
Imagine that you conducted an in-service course for ESL teachers. To receive university credit for the course, the teachers must take examinations--in this case, a midterm and a final. The midterm was a multiple-choice test of 50 items and the final exam presented teachers with 10 problem situations to solve. Sue, like most teachers, was a whiz at taking multiple-choice exams, but bombed out on the problem-solving final exam. She received a 48 on the midterm and a 1 on the final. Becky didn't do so well on the midterm. She kept thinking of exceptions to answers on the multiple-choice exam. Her score was 39. However, she really did shine on the final, scoring a 10. Since you expect students to do well on both exams, you reason that Becky has done a creditable job on each and Sue has not. Becky gets the higher grade. Yet, if you add the points together, Sue has 49 and Becky has 49. The question is whether the points are really equal.

Should Sue also do this bit of arithmetic, she might come to your office to complain of the injustice of it all. How will you show her that the value of each point on the two tests is different?

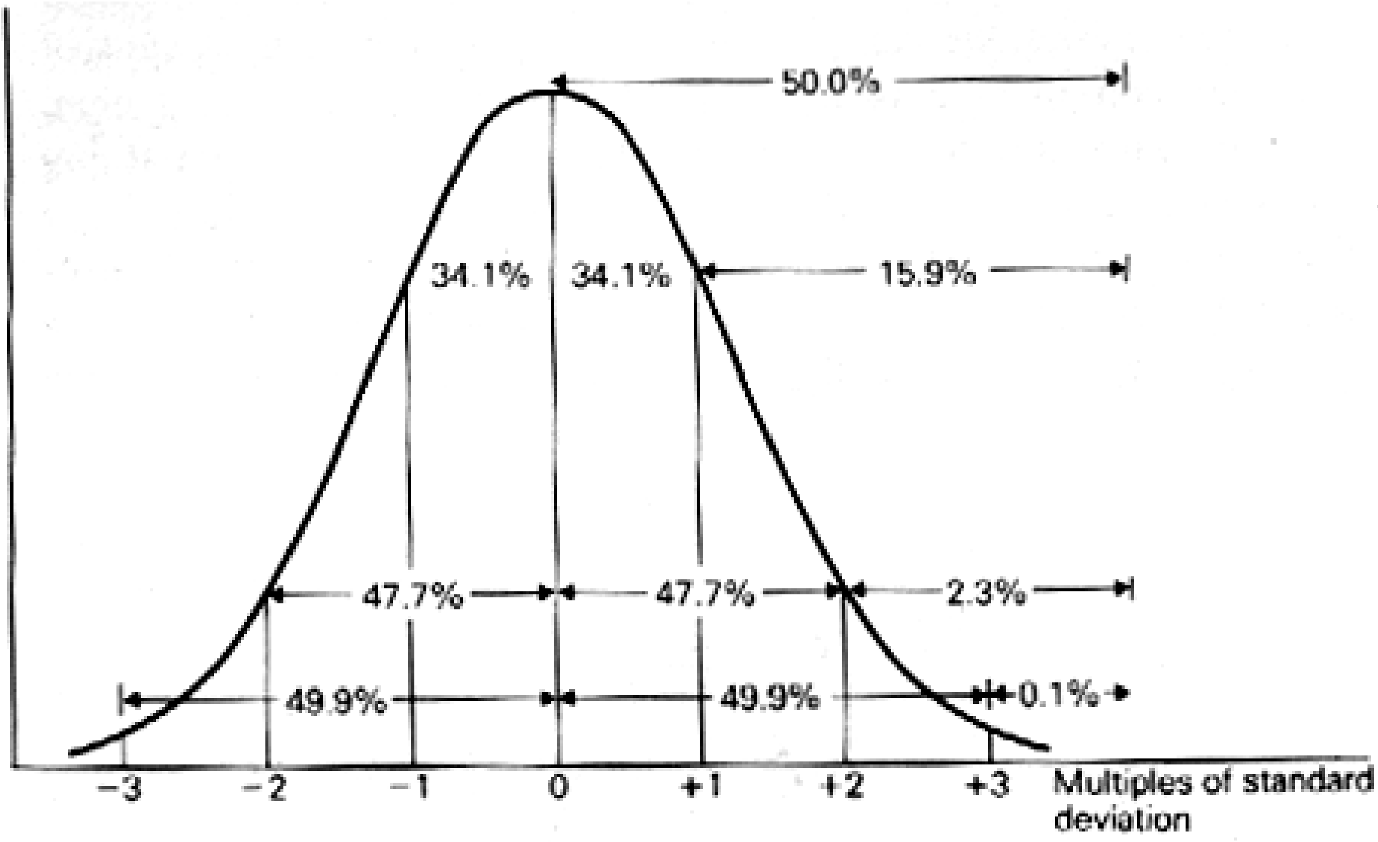
Locating scores and finding scales in a distribution

- Standard score (z-scores)

$$z = \frac{X - \bar{x}}{s}$$



Locating scores and finding scales in a distribution



Areas under the normal distribution curve

Mind work

Suppose that we have measured the times taken by a very large number of people to utter a particular sentence, and have shown these times to be normally distributed with a mean of 3.45 sec and a standard deviation of 0.84 sec. Armed with this information, we can answer various questions.

1. What proportion of the (potentially infinite) population of utterance times would be expected to fall below 3 sec?
2. What proportion would lie between 3 and 4 sec?
3. What is the time below which only 1 per cent of the times would be expected to fall?

Mind work

- 1. *z*-score for 3 sec. $z = \frac{3 - 3.45}{0.84} = -0.54$
- 2. check the normal distribution table
- 3. *z*-score for 4 sec. $z = \frac{4 - 3.45}{0.84} = -0.66$
- 4. 100-29.46-25.46=45.1 per cent
- 5. *z*-score for 1 per cent: 2.33
- 6. $-2.33 = \frac{x - 3.45}{0.84}$ *x*=(-2.33x0.84)+3.45=1.49 sec

Normal Distribution table

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
3.5	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002

Locating scores and finding scales in a distribution

- *T*-score
  - *T* score = 10(*z*) + 50
  - *Z*=(*T*-score-500)/100

$$X = z \times s + \overline{x}$$

Mind work

某外语学院在其研究生教学中规定，只要有一门课程的考试成绩低于75分，即取消其撰写论文的资格。显然，这是不科学的。因为这实质上也是把不同质的考试硬拉在一起进行比较。同是75分，在不同考试中的意义是不一样的。在一个非常容易的考试中，它可能是比较低的分数，而在一个难度较大的考试中，它却可能是比较高的考分。如果凡是低于该分数的都不让写论文，这是不科学的，也是不公平的。科学的做法是把各科的考试分数换算成标准分，然后规定多少标准分以下的没有资格写论文。同上例一样，有了标准分之后，也可以把各科的成绩合成一个总分，或求平均分，排出名次，再制定一个标准，以确定总分或平均分为多少的人才有资格撰写论文

Locating scores and finding scales in a distribution

- Distributions with nominal data
  - Implicational scaling (Guttman scaling)
  - Coefficient of scalability

Questions

	Difficult					Easy
	Q6	Q5	Q4	Q3	Q2	Q1
S6	1	1	0	1	1	1
S5	0	1	1	1	1	1
S4	0	0	1	1	1	1
S3	0	0	0	1	1	1
S2	0	0	0	0	1	1
S1	0	0	0	0	0	1

"The Friend of the Earth"

Subject	Note Taking	Deduction	Transfer	Resourcing	Hypothesis Testing	Grouping	Uptaking	Clarifying	Repetition	Problem Identification	Reprise	Planning	Self-monitoring	Translation	Imagery	Feedback	Self-evaluation	Inferencing	Elaboration	Summarization	Total	Non-fit
S4	0	0	0	0	0	(1)	0	0	(1)	0	1	1	1	1	(0)	1	(0)	1	1	1	10	4
S1	0	0	0	0	0	0	0	0	0	0	(1)	(0)	1	1	1	1	1	1	1	1	9	2
S10	0	0	0	0	0	0	(1)	(1)	0	0	0	0	1	1	(0)	1	1	(0)	1	1	8	4
S11	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	7	0
S3	0	0	0	0	0	0	0	0	0	(1)	0	(1)	0	(0)	1	1	1	1	1	(0)	7	4
S8	0	0	0	0	0	0	0	0	0	0	0	0	(1)	0	1	1	(0)	1	1	1	6	2
S13	0	0	0	0	0	0	0	0	0	0	0	0	0	(1)	0	(0)	1	1	1	1	5	2
S7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	(1)	1	(0)	1	1	1	5	2
S6	0	0	0	0	0	0	0	0	0	(1)	0	(1)	0	(1)	0	(0)	1	(0)	(0)	1	5	6
S5	0	0	0	0	0	0	0	0	(1)	0	0	0	0	0	0	0	1	1	(0)	1	4	2
S19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	(1)	(1)	0	(0)	1	(0)	3	4
S12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	(1)	(0)	1	2	2
S16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	(1)	(0)	1	2	2
S9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	(1)	0	0	0	1	(0)	2	2
S20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	(1)	0	0	0	1	(0)	2	2
S17	0	0	0	0	0	0	0	0	0	0	0	0	(1)	0	0	0	(1)	0	(0)	(0)	2	4
S18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	(1)	0	0	(0)	1	2
S15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total	0	0	0	0	0	1	1	1	2	2	2	3	5	6	8	8	9	10	11	11	80	
Non-fit	0	0	0	0	0	1	1	1	2	2	1	3	2	3	6	3	5	5	5	6		46

Homework

I. The following scores are obtained by 50 subjects on a language aptitude test:

42	62	44	32	47	42	52	76	36	43
55	27	46	55	47	28	53	44	15	61
18	59	58	57	49	55	88	49	50	62
61	82	66	80	64	50	40	53	28	63
63	25	58	71	82	52	73	67	58	77

1. Draw a histogram to show the distribution of the scores.
2. Calculate the mean and standard deviation of the scores.
3. Suppose Lihua scored 55 in this test, what's her position in the whole class?

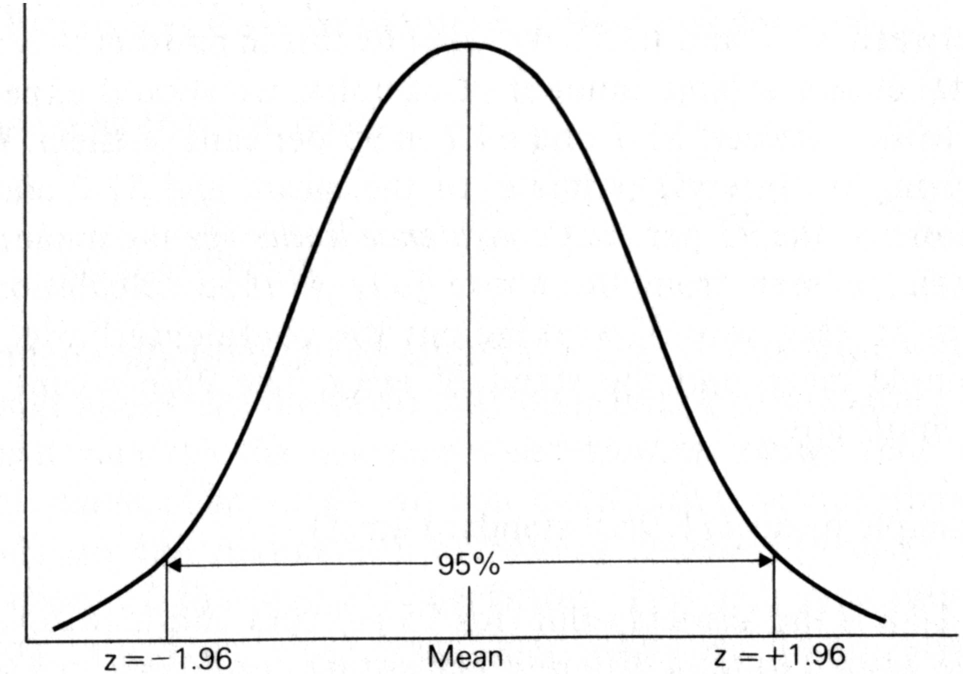
II. Suppose there will be 418,900 test takers for the NMET in 2006 in Guangdong, the key universities in China plan to enroll altogether 32,000 students in Guangdong. What score is the lowest threshold for a student to be enrolled by the key universities? (Remember the mean is 500, standard deviation is 100).

Sample statistics and population parameter: estimation

- Standard error
  - Sampling distribution of the mean
  - Standard error of mean
  - Standard error = $\frac{s}{\sqrt{N}}$
  - In order to halve the standard error, we should have to take a sample which was four times as big.
- Central limit theorem:
  - For any population with mean of μand standard deviation of σ, the distribution of sample means for sample size *n* will approach a normal distribution with a mean of μand a standard deviation of $\sigma / \sqrt{n}$ as *n* approaches infinity.
  - samples above 30

Sample statistics and population parameter: estimation

- Interpreting standard error: confidence limits



z-scores enclosing 95 per cent of the area under the normal curve

THE DIFFERENCE BETWEEN STANDARD DEVIATION AND STANDARD ERROR

A CONSTANT source of confusion for many students is the difference between standard deviation and standard error. You should remember that standard deviation measures the standard distance between a *score* and the population mean, $X - \mu$. Whenever you are working with a distribution of scores, the standard deviation is the appropriate measure of variability. Standard error, on the other hand, measures the standard distance between a *sample mean* and the population mean, $\overline{X} - \mu$. Whenever you have a question concerning a sample, the standard error is the appropriate measure of variability.

If you still find the distinction confusing, there is a simple solution. Namely, if you always use standard error, you always will be right. Consider the formula for standard error:

$$\text{standard error} = \sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$$

If you are working with a single score, then $n = 1$, and the standard error becomes

$$\text{standard error} = \sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{\sigma}{\sqrt{1}}$$

$$= \sigma = \text{standard deviation}$$

Thus, standard error always measures the standard distance from the population mean, whether you have a sample of $n = 1$ or $n = 100$.

Sample statistics and population parameter: estimation

- Normal distribution: sample is large
- *t*-distribution: sample is small
  - Degree of freedom: N-1
  - When sample is large, $t = z$

Sample statistics and population parameter: estimation

- Interpreting standard error: confidence limits
- Mean=58.2
- s=23.6
- N=50
- Standard error= $\frac{s}{\sqrt{N}} = \frac{23.6}{\sqrt{50}} = 3.3$
- $z = \frac{58.2 - \bar{x}}{3.3} \rightarrow -1.96 \le \frac{58.2 - \bar{x}}{3.3} \le 1.96$ $\quad 51.7 \le \bar{x} \le 64.7$

Sample statistics and population parameter: estimation

- Confidence limits for proportions
- Standard error = $\sqrt{\dfrac{p(1-p)}{N}}$
- Confidence limits=proportion in sample ±(critical value x standard error)

Sample statistics and population parameter: estimation

Suppose that we have taken a random sample of 500 finite verbs from a text, and found that 150 of them have present tense form. How can we set confidence limits for the proportion of present tense finite verbs in the whole text, the population from which the sample is taken?

95% confidence limits = proportion in sample

$\pm$ (1.96 X standard error) =0.30$\pm$(1.96x0.02)

= 0.30 $\pm$ 0.04 = 0.26 to 0.34.

We can thus be 95 per cent confident that the proportion of present tense finite verbs in the population lies between 26 and 34 per cent.

Sample statistics and population parameter: estimation

- Estimating required sample sizes
  - Standard error = $\sqrt{\frac{p(1-p)}{N}}$

In a paragraph there are 46 word tokens, of which 11 are two-letter words. The proportion of such words is thus 11/46 or 0.24. How big a sample of words should we need in order to be 95 per cent confident that we had measured the proportion to within an accuracy of 1 per cent?

0.01=1.96 x standard error

Standard error = 0.01 x 1.96

$$N = \frac{0.24 \times 0.76}{(0.01/1.96)^2} = 7007$$

Homework

I. The following are the times (in seconds) taken for a group of 30 subjects to carry out the detransformation of a sentence into its simplest form:

0.55	0.56	0.52	0.59	0.51	0.50
0.42	0.41	0.37	0.22	0.24	0.41
0.49	0.59	0.75	0.65	0.63	0.61
0.72	0.77	0.76	0.39	0.26	0.68
0.30	0.32	0.44	0.61	0.54	0.47

Calculate (i) the mean, (ii) the standard deviation, (iii) the standard error of the mean, (iv) the 99 per cent confidence limits for the mean.

II. A random sample of 300 finite verbs is taken from a text, and it is found that 63 of these are auxiliaries. Calculate the 95 per cent confidence limits for the proportion of finite verbs which are auxiliaries in the text as a whole.

III. Using the data in question II, calculate the size of the sample of finite verbs which would. be required in order to estimate the proportion of auxiliaries to within an accuracy of 1 per cent, with 95 per cent confidence.

Probability and Hypothesis Testing

- Null hypothesis (H_0)
  - The null hypothesis states that in the general population there is no change, no difference, or no relationship. In the context of an experiment, H_0 predicts that the independent variable (treatment) will have no effect on the dependent variable for the population. H_0: $\mu_A - \mu_B = 0$ or $\mu_A = \mu_B$
- Alternative hypothesis (H_1)
  - The alternative hypothesis (H_1) states that there is a change, a difference, or a relationship for the general population. H_1: $\mu_A \neq \mu_B$

Probability and Hypothesis Testing

- Null hypothesis (H_0)
  - When we reject the null hypothesis, we want the probability to be very low that we are wrong. If, on the other hand, we must accept the null hypothesis, we still want the probability to be very low that we are wrong in doing so.
- Type I error and Type II error
  - A type I error is made when the researcher rejected the null hypothesis when it should not have been rejected.
  - A type II error is made when the null hypothesis is accepted when it should have been rejected.
- In research, we test our hypothesis by finding the probability of our results. Probability is the proportion of times that any particular outcome would happen if the research were repeated an infinite number of times.

Probability and Hypothesis Testing

- Two-tailed and one-tailed hypothesis
  - When we specify no direction for the null hypothesis (i.e., whether our score will be higher or lower than more typical scores), we must consider both tails of the distribution. This is called two-tailed hypothesis.
  - If we have good reason to believe that we will find a difference (e.g., previous studies or research findings suggest this is so), then we will use a one-tailed hypothesis. One-tailed tests specify the direction of the predicted difference. We use previous findings to tell us which direction to select.

	.05	.01
1-tailed	1.64	2.33
2-tailed	1.96	2.57

Probability and Hypothesis Testing

- Steps in hypothesis testing

1. State the null hypothesis.
2. Decide whether to test it as a one- or two-tailed hypothesis. If there is no research evidence on the issue, select a two-tailed hypothesis. This will allow you to reject the null hypothesis in favor of an alternative hypothesis. If there is research evidence on the issue, select a one-tailed hypothesis. This will allow you to reject the null hypothesis in favor of a directional hypothesis.
3. Set the probability level (α level). Justify your choice.
4. Select the appropriate statistical test(s) for the data.
5. Collect the data and apply the statistical test(s).
6. Report the test results and interpret them correctly.

Probability and Hypothesis Testing

- Parametric vs. nonparametric
  - Parametric procedures
    - Make strong assumptions about the distribution of the data
    - Assume the data are NOT frequencies or ordinal scales but interval data
    - Data are normally distributed
  - Nonparametric procedures
    - Do not make strong assumptions about the shape of the distribution of the data
    - Work with frequencies and rank-ordered scales
    - Used when the sample size is small

Homework