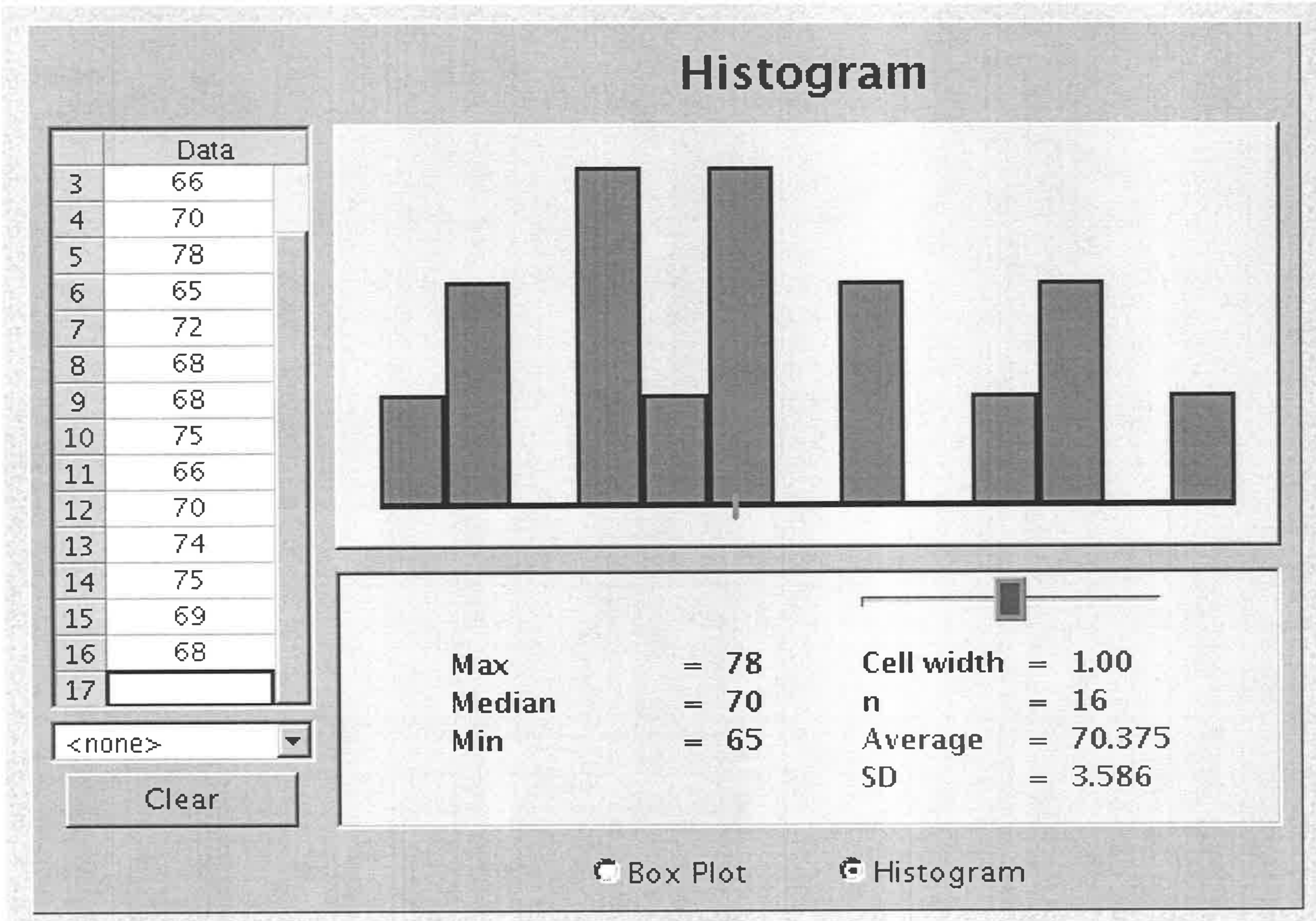
USU Women's Basketball Team, 2013-2014

Average Height = 71.625

Height	**Deviation from AV**	**Squared Deviation**
75	3.38	11.39
74	2.38	5.64
72	0.38	0.14
71	-0.63	0.39
76	4.38	19.14
76	4.38	19.14
69	-2.63	6.89
72	0.38	0.14
67	-4.63	21.39
67	-4.63	21.39
69	-2.63	6.89
67	-4.63	21.39
73	1.38	1.89
73	1.38	1.89
74	2.38	5.64
71	-0.63	0.39
Average Squared Deviation =		8.98
Standard Deviation =		3.00

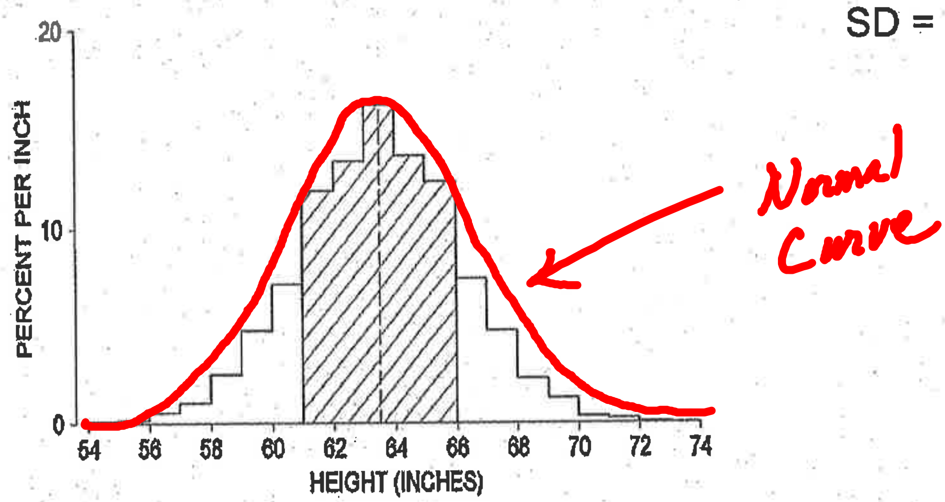
The SD is the RMS of the deviations from AV


Histogram
Data
3 66
4 70
5 78
6 65
7 72
8 68
9 68
10 75
11 66
12 70
13 74
14 75
15 69
16 68
17
<none>
Clear
Max = 78
Median = 70
Min = 65
Cell width = 1.00
n = 16
Average = 70.375
SD = 3.586
Box Plot
Histogram

827.

Example: The histogram of heights of 6,588 women aged 18-74 in NHANES II.

Average = 63.5"
SD = 2.5"

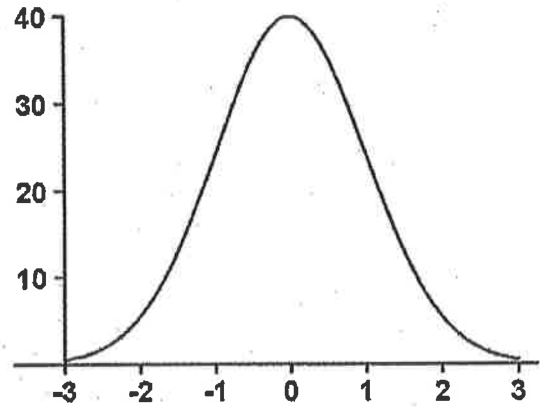


Average − 1 SD = 61″
Average + 1 SD = 66″
} 67% of the women were between 61″ and 66″ tall (shaded)

Average − 2 SDs = 58.5″
Average + 2 SDs = 68.5″
} 94% of the women were between 58.5″ and 68.5″ tall

Chapter 5: The Normal Approximation for Data

The standard normal curve

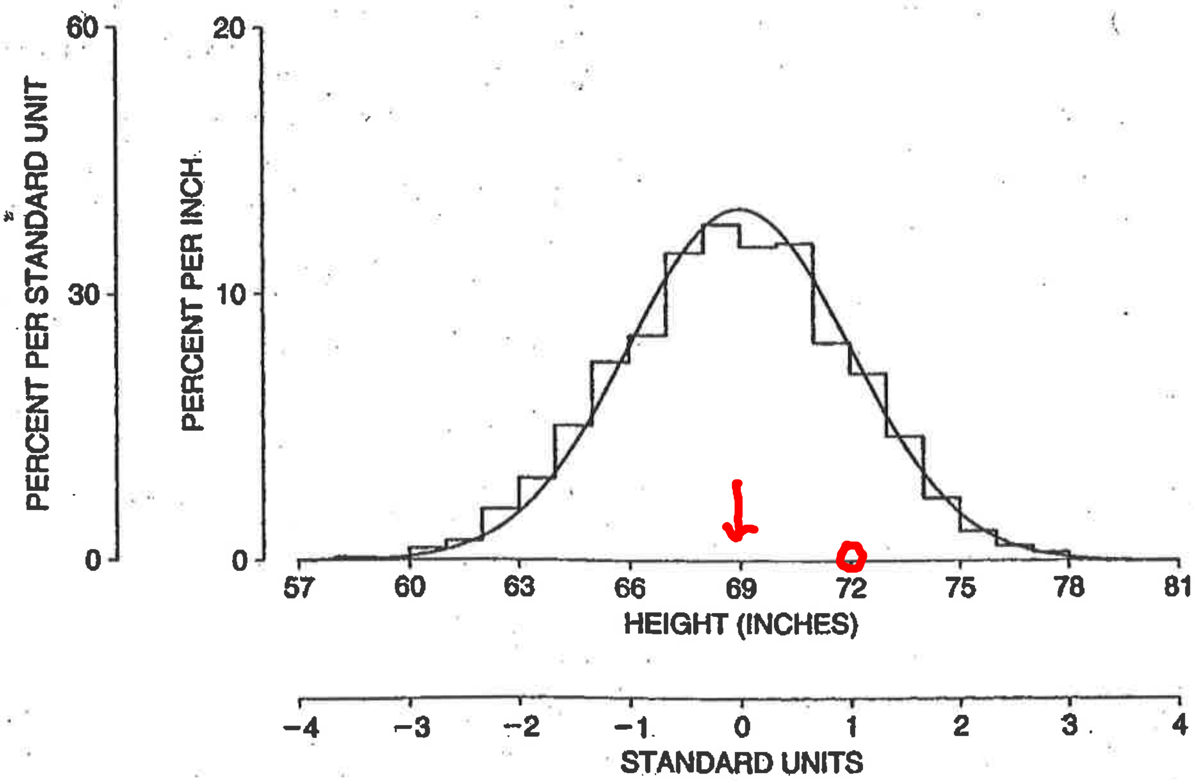


$$y = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

Gauss

2Ⓑ a₁

HEIGHTS OF MEN



AV = 69

SD = 3

10

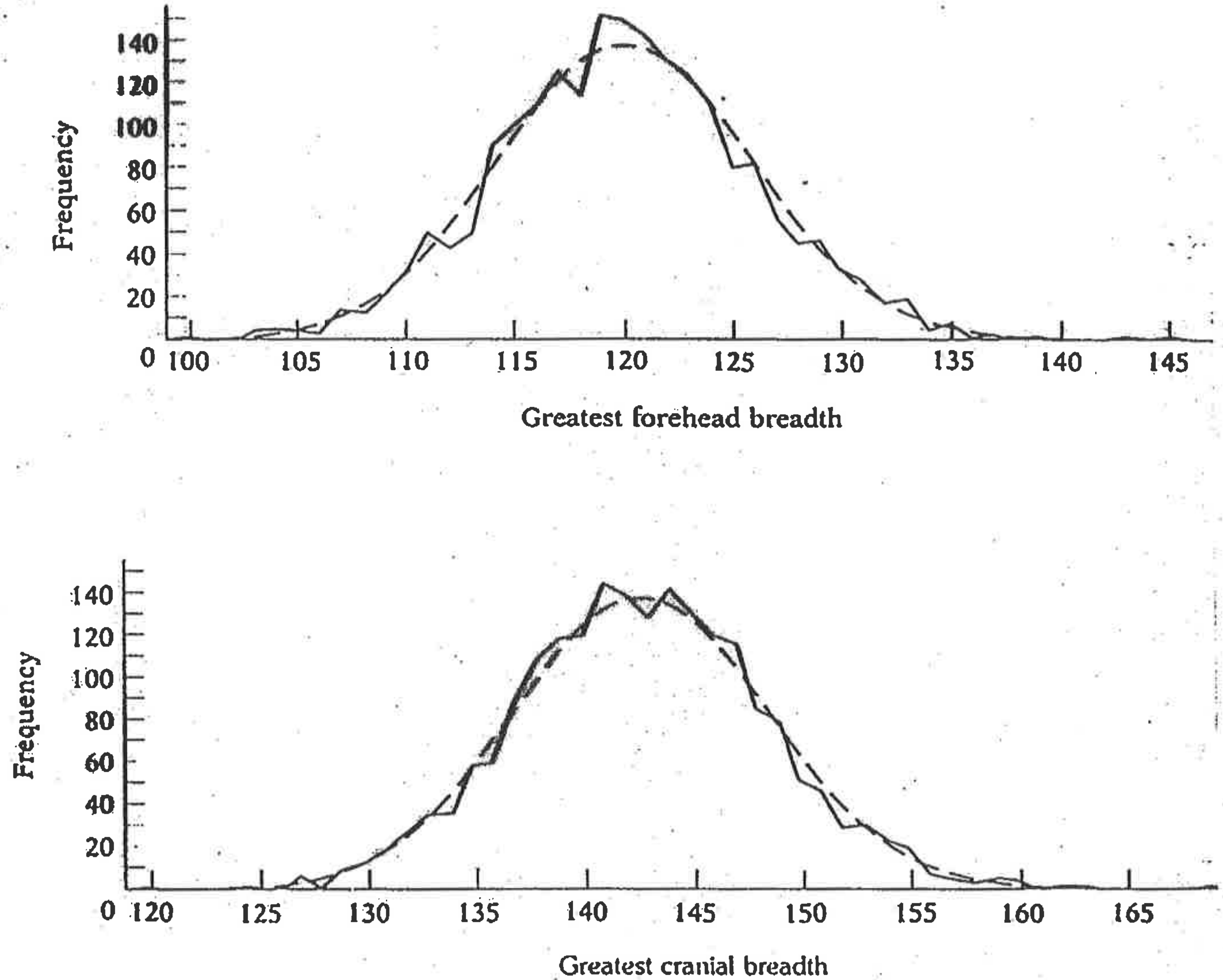
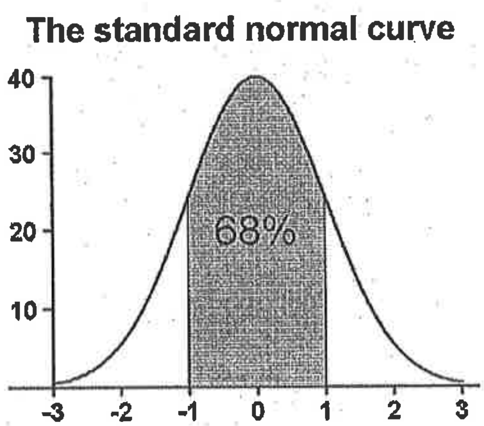


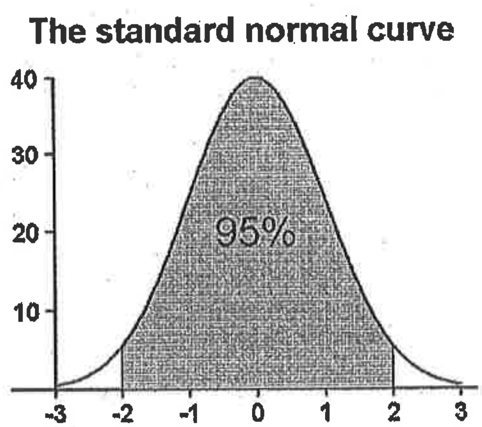
Figure 4-24. 2000 Hungarian skulls. The distribution of each of two measurements on a large sample of male skulls is approximately normal. The solid line is the actual frequency curve, and the broken line is a normal curve that approximates the distribution. [From Karl Pearson, "Craniological notes," *Biometrika*, June 1903, p. 344. Reproduced by permission of the Biometrika Trustees.]

11

Facts about the standard normal curve:

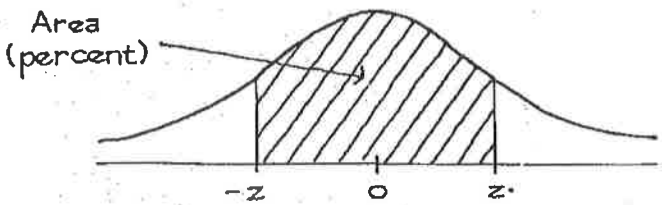
- Total area under the curve is 100%
- The curve is symmetric about 0





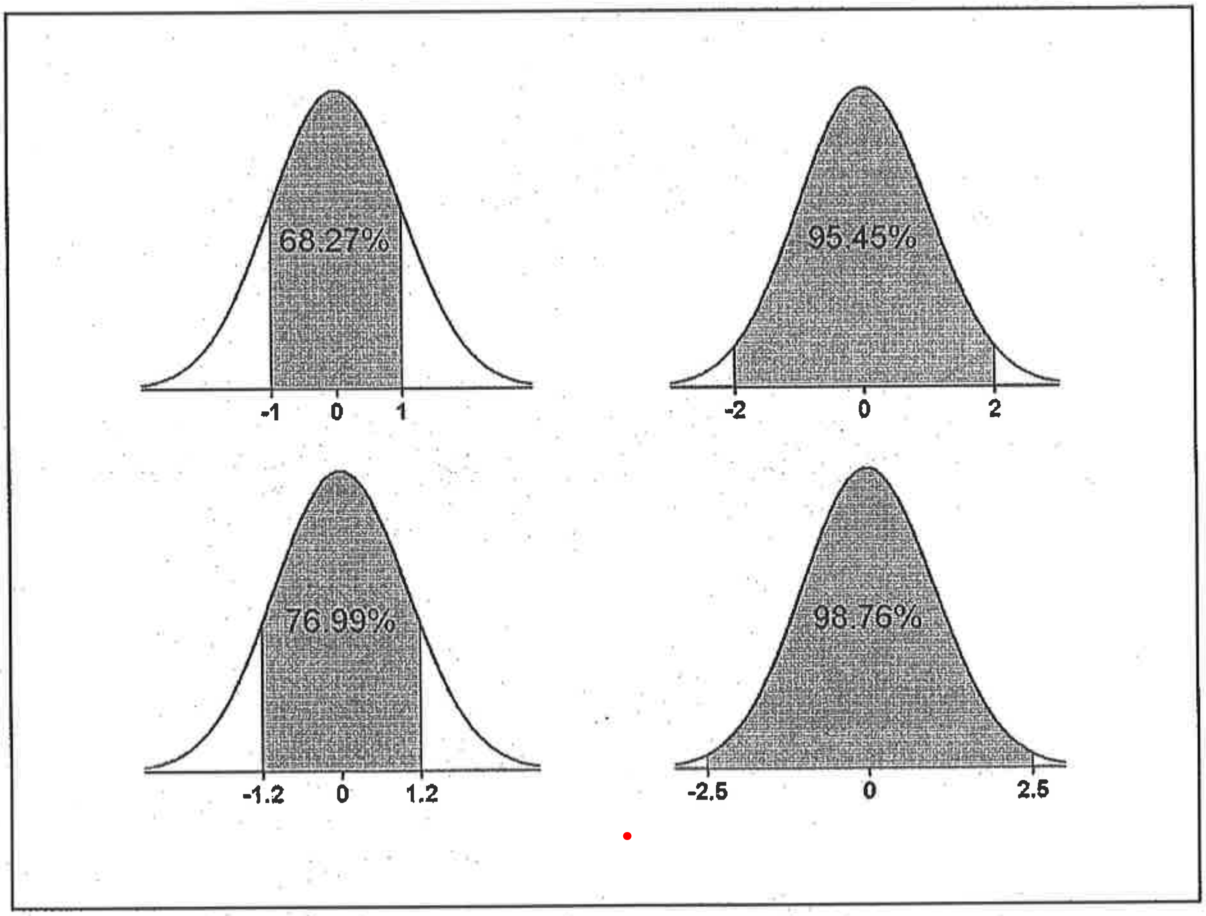
12.

A NORMAL TABLE

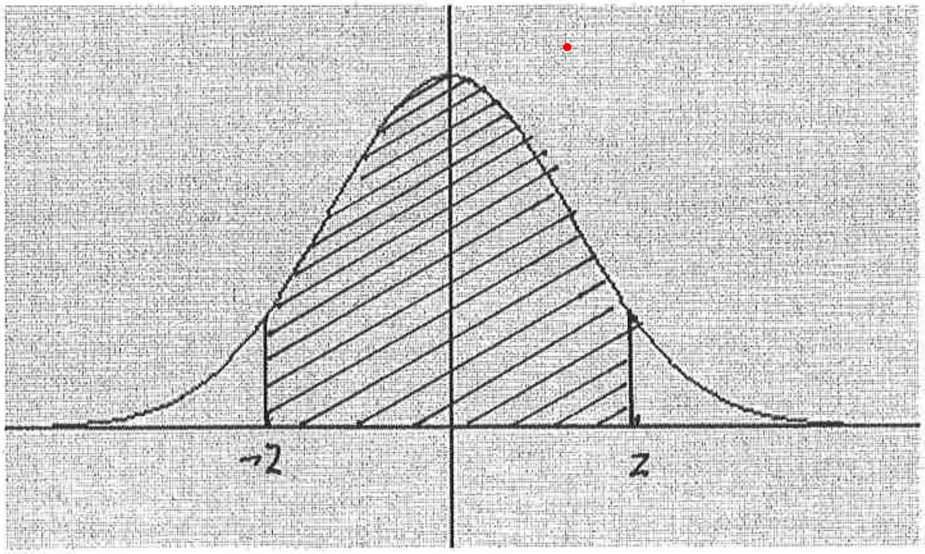


z	Area	z	Area	z	Area
0.00	0	1.50	86.64	3.00	99.730
0.05	3.99	1.55	87.89	3.05	99.771
0.10	7.97	1.60	89.04	3.10	99.806
0.15	11.92	1.65	90.11	3.15	99.837
0.20	15.85	1.70	91.09	3.20	99.863
0.25	19.74	1.75	91.99	3.25	99.885
0.30	23.58	1.80	92.81	3.30	99.903
0.35	27.37	1.85	93.57	3.35	99.919
0.40	31.08	1.90	94.26	3.40	99.933
0.45	34.73	1.95	94.88	3.45	99.944
0.50	38.29	2.00	95.45	3.50	99.953
0.55	41.77	2.05	95.96	3.55	99.961
0.60	45.15	2.10	96.43	3.60	99.968
0.65	48.43	2.15	96.84	3.65	99.974
0.70	51.61	2.20	97.22	3.70	99.978
0.75	54.67	2.25	97.56	3.75	99.982
0.80	57.63	2.30	97.86	3.80	99.986
0.85	60.47	2.35	98.12	3.85	99.988
0.90	63.19	2.40	98.36	3.90	99.990
0.95	65.79	2.45	98.57	3.95	99.992
1.00	68.27	2.50	98.76	4.00	99.9937
1.05	70.63	2.55	98.92	4.05	99.9949
1.10	72.87	2.60	99.07	4.10	99.9959
1.15	74.99	2.65	99.20	4.15	99.9967
1.20	76.99	2.70	99.31	4.20	99.9973
1.25	78.87	2.75	99.40	4.25	99.9979
1.30	80.64	2.80	99.49	4.30	99.9983
1.35	82.30	2.85	99.56	4.35	99.9986
1.40	83.85	2.90	99.63	4.40	99.9989
1.45	85.29	2.95	99.68	4.45	99.9991

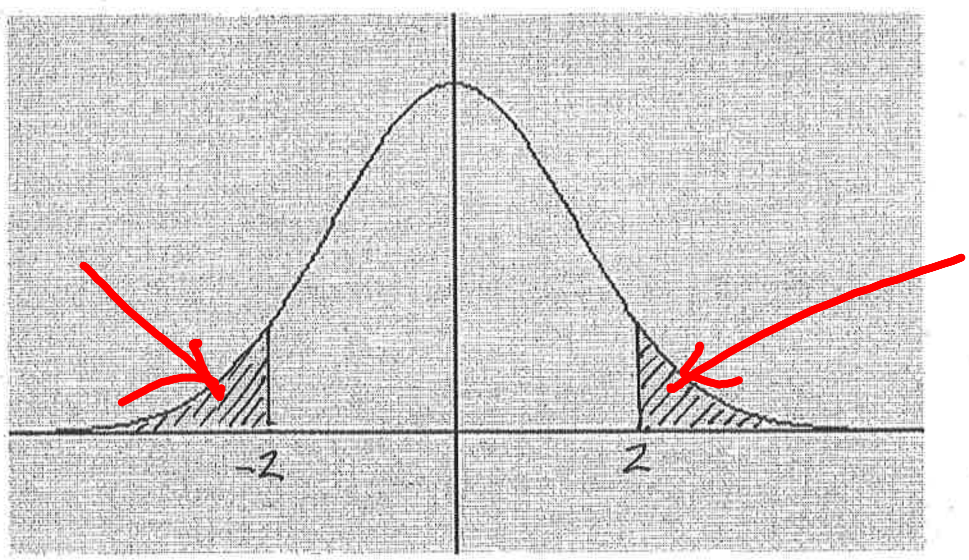
13.



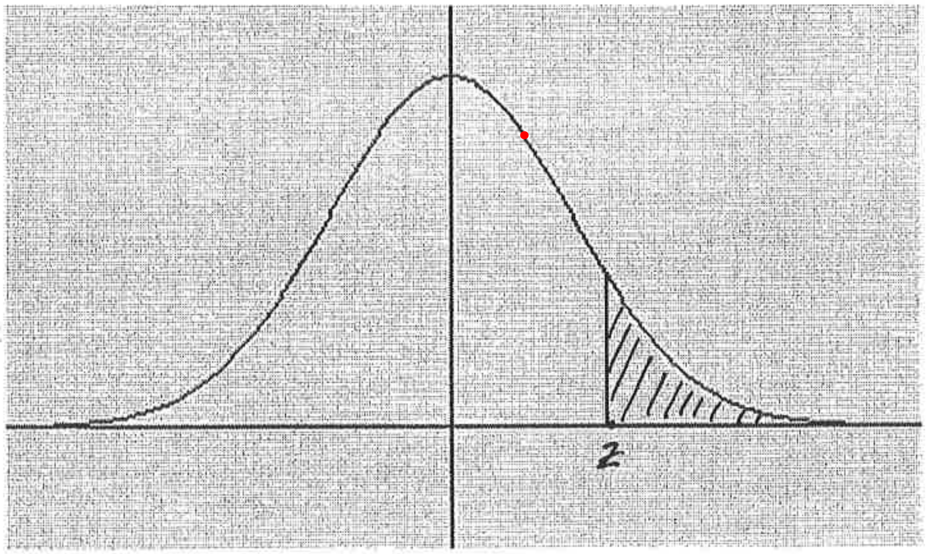
14.



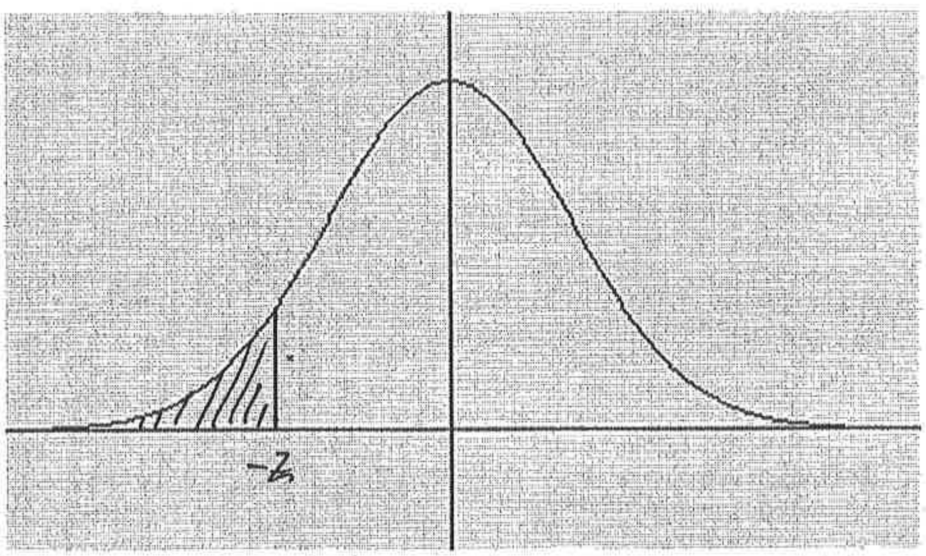
A(z)



100 - A(z)

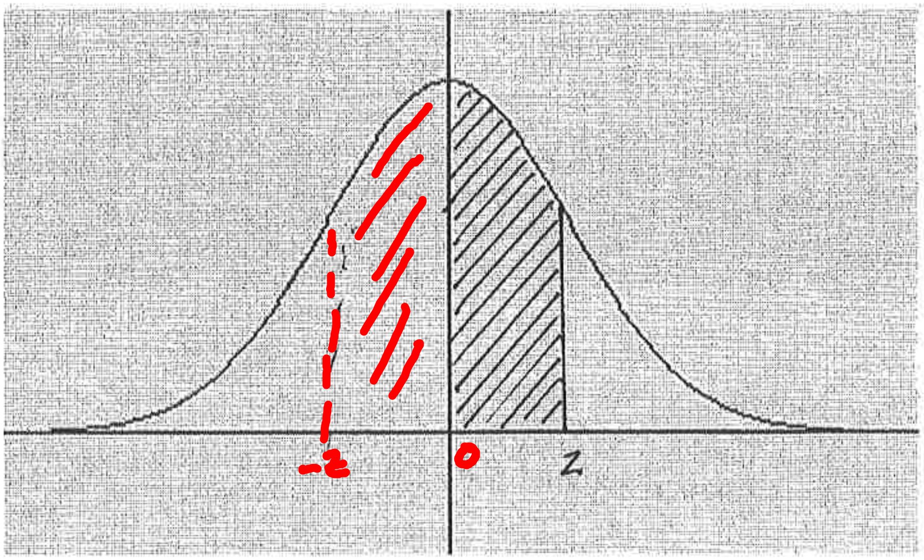


$$\frac{100 - A(z)}{2}$$

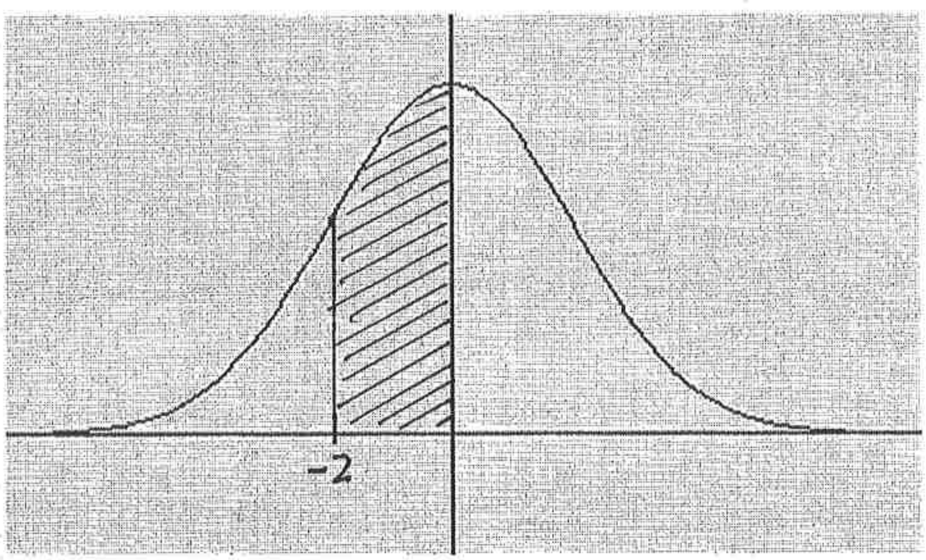


$$\frac{100 - A(z)}{2}$$

14

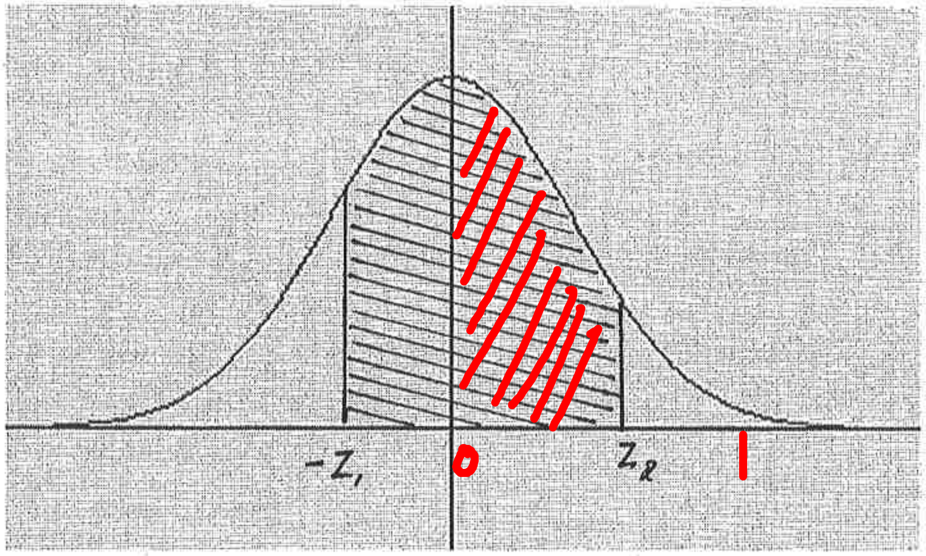


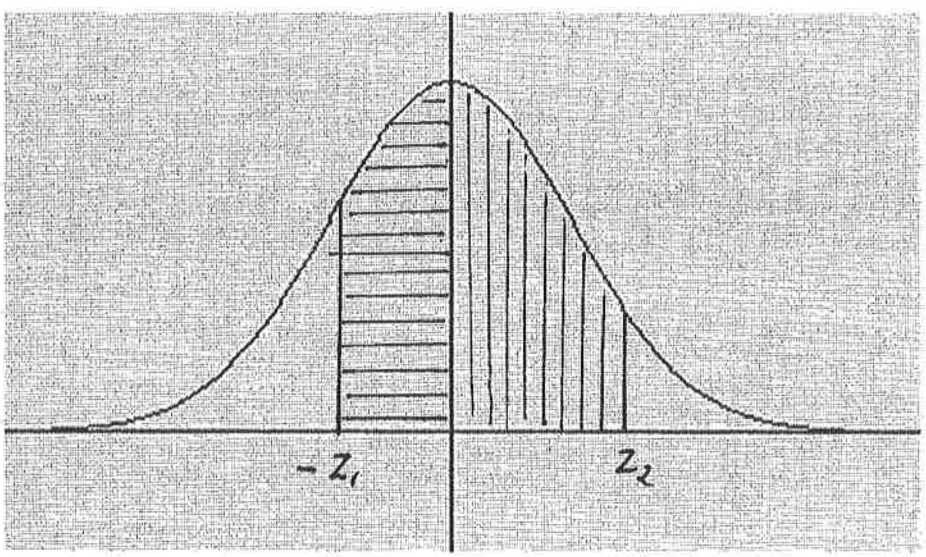
$\frac{A(z)}{2}$



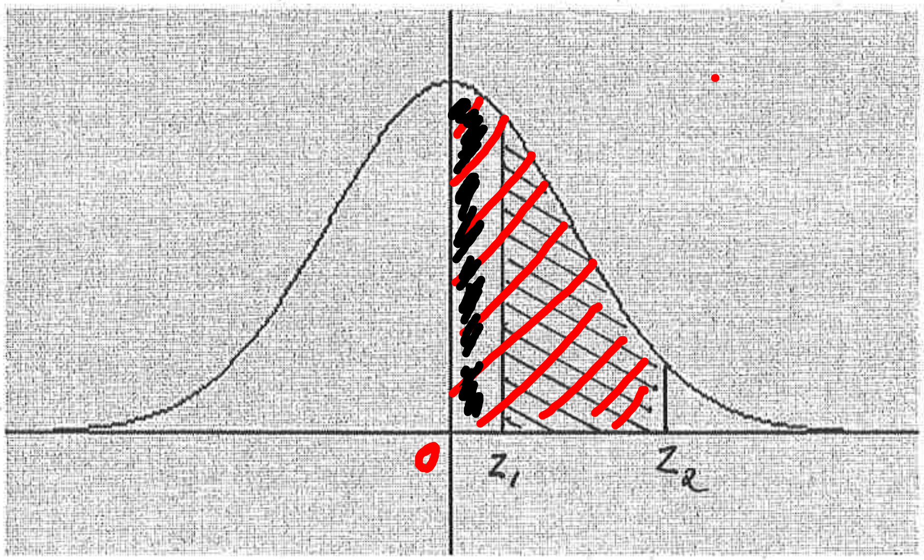
$\frac{A(z)}{2}$

17.

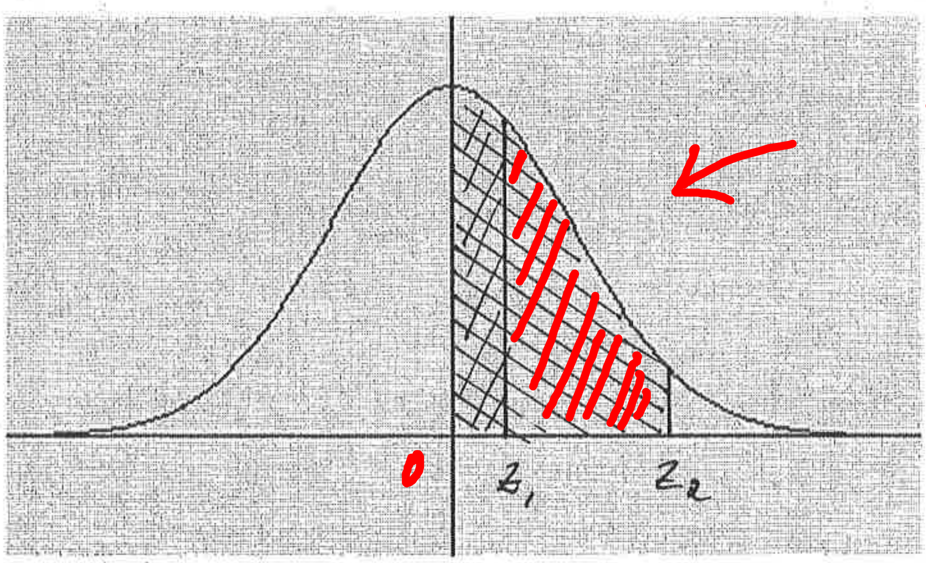




$$\frac{A(Z_1)}{2} + \frac{A(Z_2)}{2}$$



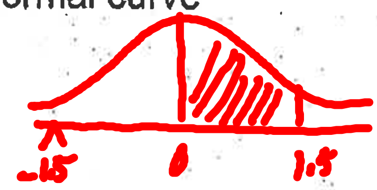
$$\frac{A(z_2)}{2} - \frac{A(z_1)}{2}$$



$$\underbrace{\frac{A(z_2)}{2}} - \frac{A(z_1)}{2}$$

19

1. Find the area between 0 and 1.5 under the standard normal curve



$\frac{A(1.5)}{2} = \frac{86.64}{2} = 43.32\% \approx 43\%$

2. Find the area between –3 and 0 under the standard normal curve

$\frac{A(3)}{2} = \frac{99.73}{2} = 49.865 \approx 49.9\%$

3. Find the area between –1 and 1.3 under the standard normal curve

$\frac{A(1)}{2} + \frac{A(1.3)}{2}$

$= 34\% + \frac{80.64}{2} = 34\% + 40.3\% \approx 74.3\%$

20

Standard Units

To convert to standard units, we do:

$$z = \frac{x - \text{average}}{SD}$$

Example 2. Test scores have an average of 70 and an SD of 10.

a) If a student scored 80, what is their score in standard units?

$$\frac{80-70}{10} = 1$$

b) If a student scored 60, what is their score in standard units?

$$\frac{60-70}{10} = -1$$

21

Example 2 (continued): average = 70, SD = 10

c) If a student scored 75, what is their score in standard units?

$$\frac{75-70}{10} = \frac{1}{2}$$

d) If a student scored 85, what is their score in standard units?

$$\frac{85-70}{10} = \frac{15}{10} = 1.5$$

e) If a student scored 55, what is their score in standard units?

$$\frac{55-70}{10} = -\frac{15}{10} = -1.5$$

22.

Standard Units

Standard units say how many SDs we are above (+) or below (-) the average.

Example 1. Heights of a group of women follow the normal curve with an average of 63.5" and an SD of 2.5". Convert the following to standard units:

a) 66 inches $\frac{66-63.5}{2.5} = \frac{2.5}{2.5} = 1$

b) 61 inches

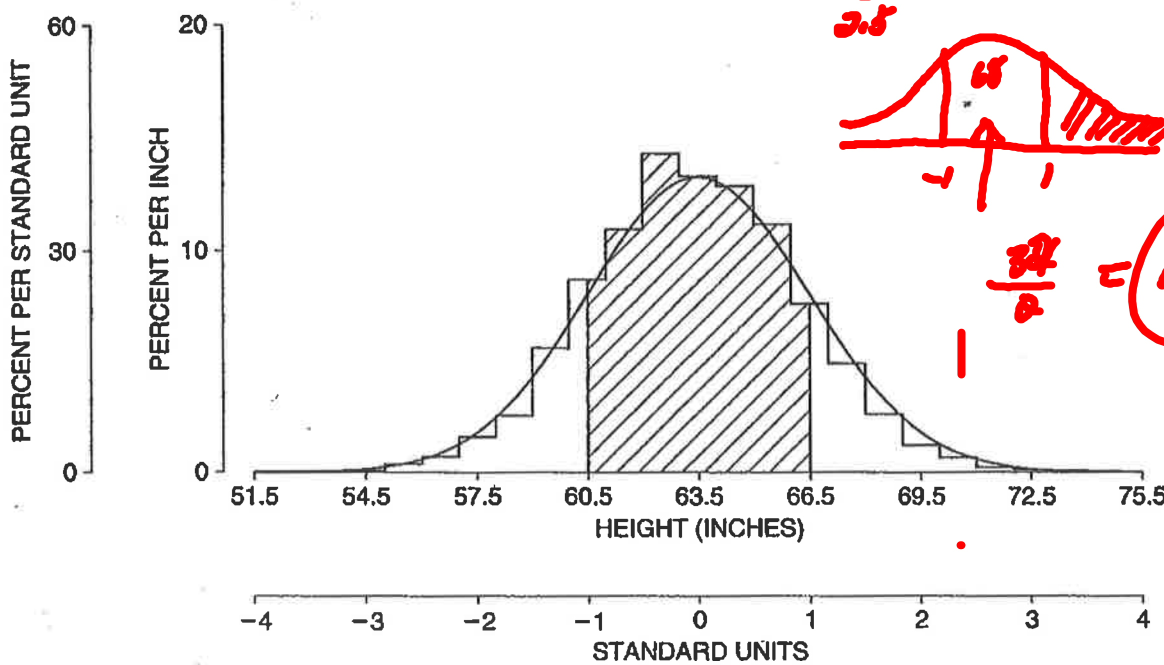
c) 63.5 inches $\frac{63.5-63.5}{2.5} = 0$

d) 62 inches

Figure 2. A histogram for heights of women compared to the normal curve. The area under the histogram between 60.5 inches and 66.5 inches (the percentage of women within one SD of average with respect to height) is about equal to the area between −1 and +1 under the curve—68%.

$$\frac{66-63.5}{2.5} = 1$$

$$\frac{32}{2} = 16\%$$



24

The Normal Approximation

If the histogram for the data values looks like the normal curve, we can estimate the percentage of people in an interval by:

1. Converting to standard units

2. Using the tables

This replaces the TRUE histogram with the appropriate NORMAL CURVE to give an APPROXIMATE area.

25

Normal Approximation

1. What percentage of the population has an IQ between 100 and 130?

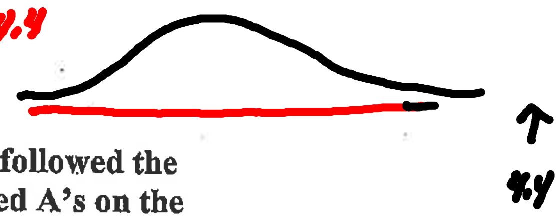
AV = 100
SD = 15

$\frac{100-100}{15} = 0$, $\frac{130-100}{15} = 2$

2. What percentage of USU students scored above 30 on the Math ACT?

AV = 19
SD = 2.5

$\frac{30-19}{2.5} = 4.4$



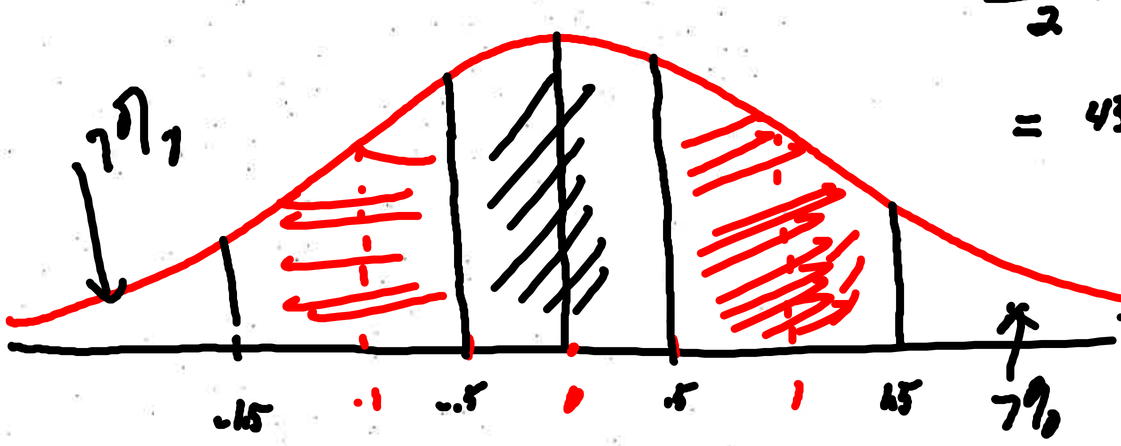
3. The scores on a recent Stat 1040 departmental final followed the normal curve with AV=72, SD=12. The top 15% earned A's on the final. Find the lowest score that still earned a grade of A.

4. What does it mean to be "graded on the normal curve"?

$\frac{A(1.5)}{2} - \frac{A(.5)}{2}$

$= 43\% - 19\%$

$\approx 24\%$



A(.5)

=

24%

C
38%

B
24%

48
38
4C

Caution! Not all histograms are normal!

We can ***not*** use the normal curve to figure areas or percentiles if the histogram does not follow the normal curve. Sometimes we can see that the histogram is not normal because the normal curve would give crazy answers!

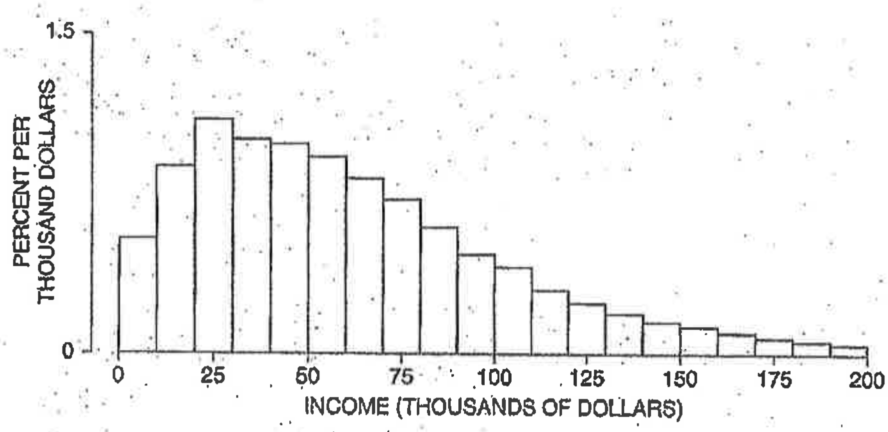
e.g. If we used the normal curve for incomes, it would tell us we had quite a large percentage of people below $0, which we know can't be true!

27

Percentiles

Percentiles are used to summarize histograms that are not symmetric, such as income histograms:

Figure 5. Distribution of families by income: the U.S. in 2004.



Source: March 2005 Current Population Survey: CD-ROM supplied by the Bureau of the Census. Primary families.

25

Percentiles

The 10^{th} percentile has 10% of the area to the left of it.

The 25th percentile has 25% of the area to the left of it.

The 75th percentile has 75% of the area to the left of it.

The 50^{th} percentile is the ________________

Table 1. Selected percentiles for family income in the U.S. in 2004.

1	$0
10	$15,000
25	$29,000
50	$54,000
75	$90,000
90	$135,000
99	$430,000

Source: March 2005 Current Population Survey; CD-ROM supplied by the Bureau of the Census. Primary families.

The pth percentile has p% of the area to the left of it.