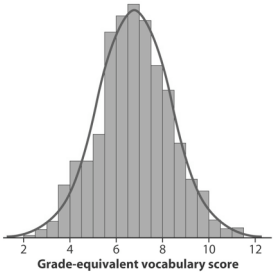
Chapter 3

The Normal Distributions

Density Curves

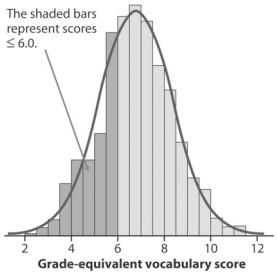
Example: here is a histogram of vocabulary scores of 947 seventh graders.
The smooth curve drawn over the histogram is a *mathematical "idialization"* for the distribution.

It is what the histogram "looks" like when we have LOTS of data.



Density Curves

Example: the areas of the shaded bars in this histogram represent the proportion of scores in the observed data that are less than or equal to 6.0. This proportion is equal to 0.303.



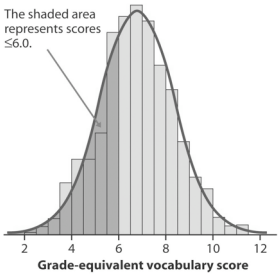
Density Curves

Example: now the area under the smooth curve to the left of 6.0 is shaded. Its proportion to the total area is now equal to 0.293 (not 0.303).

This is what the proportion on the previous slide would equal to if we had LOTS of data.

Like tossing a fair coin.

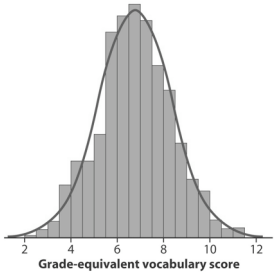
In reality, we get fractions near 50%.



Density Curves

If the scale is adjusted so the total area under the curve is exactly 1, then this curve is called a density curve.

This means heights of bars in histogram are proportions instead of frequencies.



Density Curves

- Always on or above the horizontal axis
- Have area exactly 1 underneath curve
- Area under the curve and above any range of values is the "theoretical" proportion of all observations that fall in that range

Density Curves

- The ***median*** of a density curve is the equal-areas point, the point that divides the area under the curve in half
- The ***mean*** of a density curve is the balance point, at which the curve would balance if made of solid material

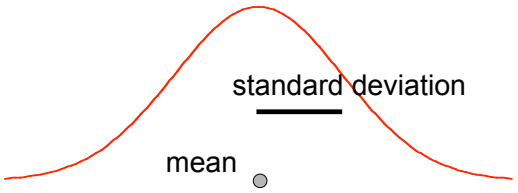
Density Curves

- The mean and standard deviation computed from actual observations (data) are denoted by $\bar{x}$ and *s*, respectively.
- The mean and standard deviation of the "theoretical" distribution represented by the density curve are denoted by μ ("mu") and σ ("sigma"), respectively.

Question

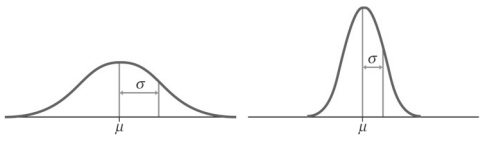
Data sets consisting of physical measurements (heights, weights, lengths of bones, and so on) for adults of the same species and sex tend to follow a similar pattern. The pattern is that most individuals are clumped around the average, with numbers decreasing the farther values are from the average in either direction. Describe what shape a histogram (or density curve) of such measurements would have.

Bell-Shaped Curve: The Normal Distribution



The Normal Distribution

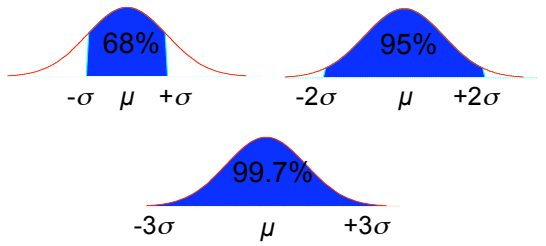
Knowing the mean (μ) and standard deviation (σ) allows us to make various conclusions about Normal distributions. Notation: N(μ,σ).



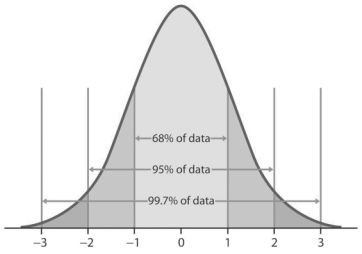
68-95-99.7 Rule for Any Normal Curve

- 68% of the observations fall within one standard deviation of the mean
- 95% of the observations fall within two standard deviations of the mean
- 99.7% of the observations fall within three standard deviations of the mean

68-95-99.7 Rule for Any Normal Curve



68-95-99.7 Rule for Any Normal Curve



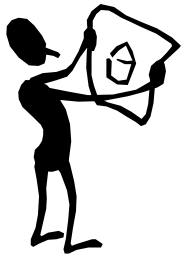
Health and Nutrition Examination Study of 1976-1980

- Heights of adult men, aged 18-24
  - mean: 70.0 inches
  - standard deviation: 2.8 inches
  - heights follow a normal distribution, so we have that heights of men are N(70, 2.8).

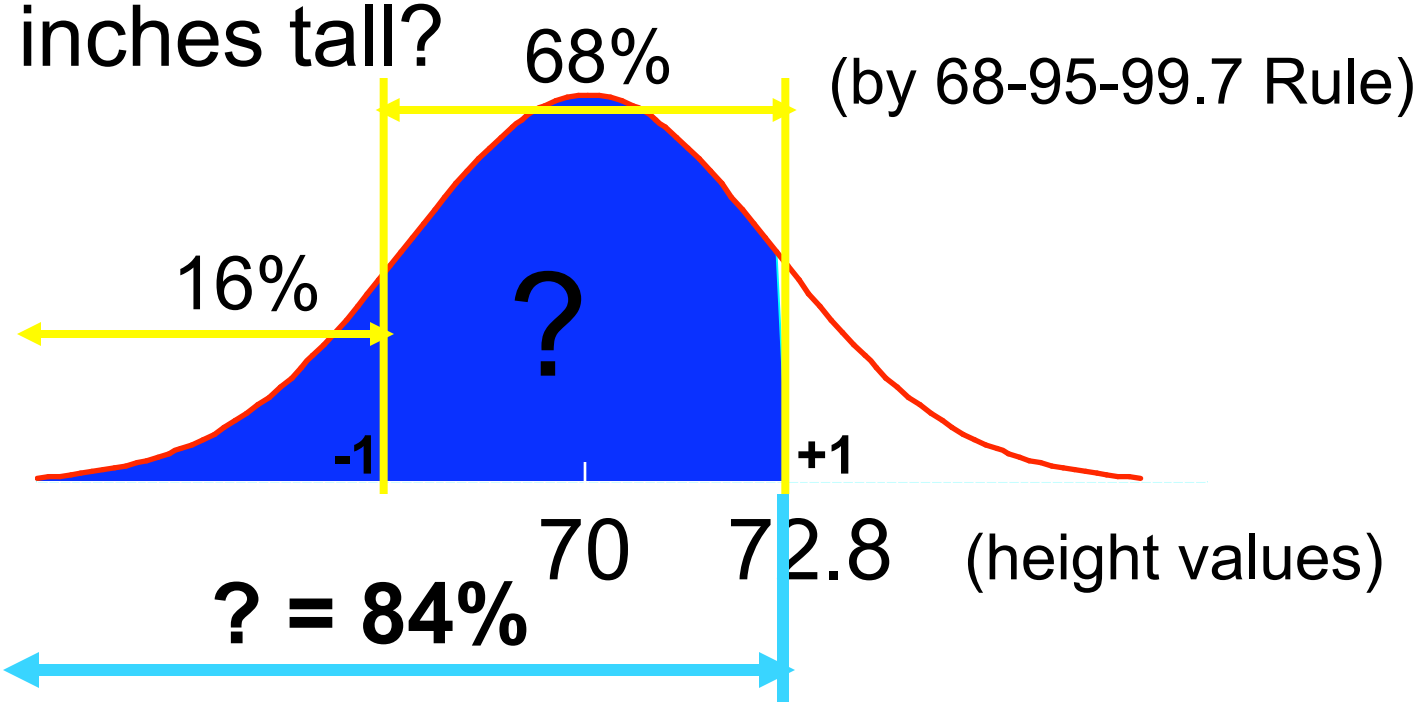
Health and Nutrition Examination Study of 1976-1980

- 68-95-99.7 Rule for men's heights
  - 68% are between 67.2 and 72.8 inches
    [$\mu \pm \sigma$ = 70.0 ± 2.8]
  - 95% are between 64.4 and 75.6 inches
    [$\mu \pm 2\sigma$ = 70.0 ± 2(2.8) = 70.0 ± 5.6]
  - 99.7% are between 61.6 and 78.4 inches
    [$\mu \pm 3\sigma$ = 70.0 ± 3(2.8) = 70.0 ± 8.4]

Health and Nutrition Examination Study of 1976-1980

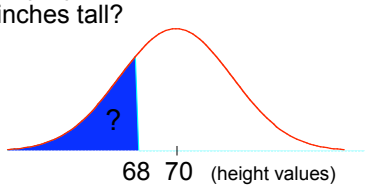


- What proportion of men are less than 72.8 inches tall?



Health and Nutrition Examination Study of 1976-1980

- What proportion of men are less than 68 inches tall?



How many standard deviations is 68 from 70?

Standardized Scores

- How many standard deviations is 68 from 70?
- ***standardized score*** =
  (observed value minus mean) / (std dev)
  [= (68 − 70) / 2.8 = −0.71]
- The value 68 is 0.71 standard deviations *below* the mean 70.

Standardized Scores

Jane is taking 1070-1. John is taking 1070-2.
Jane got 81 points. John got 76 points.
Question: Did Jane do slightly better?

Acount for difficulty: subtract class average.
Jane: 81-71=10; John: 76-56=20
Question: Did John do way better?

Acount for variability: divide by standard deviation.
Jane: (81-71)/2=5; John: (76-56)/10=2
Answer: Jane did way better!

Standard Normal Distribution

- The standard Normal distribution is the Normal distribution with mean 0 and standard deviation 1: N(0,1).
- Useful Fact: If data has Normal distribution with mean μ and standard deviation σ, then the following *standardized data* has the standard Normal distribution:

$$z_i = \frac{x_i - \mu}{\sigma}$$

Health and Nutrition Examination Study of 1976-1980

- What proportion of men are less than 68 inches tall?

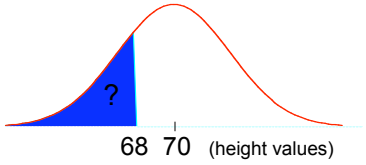


Table A: Standard Normal Probabilities

- See pages 690-691 in text for Table A.
  (the *"Standard Normal Table"*)
- Look up the closest standardized score (*z*) in the table.
- Find the probability (area) to the left of the standardized score.

Table A: Standard Normal Probabilities

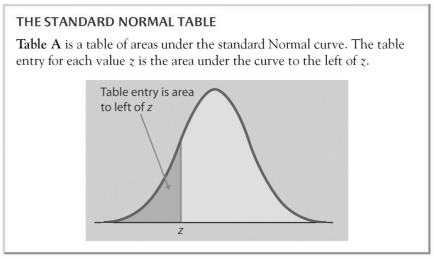

THE STANDARD NORMAL TABLE

Table A is a table of areas under the standard Normal curve. The table entry for each value z is the area under the curve to the left of z.

Table entry is area to left of z

z

Table A: Standard Normal Probabilities

	.01
-0.7	.2389

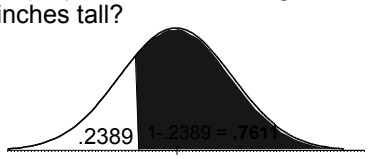
Health and Nutrition Examination Study of 1976-1980

- What proportion of men are less than 68 inches tall?

68 70 (height values)
-0.71 0 (standardized values)

Health and Nutrition Examination Study of 1976-1980

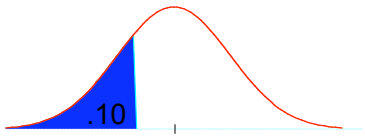
- What proportion of men are greater than 68 inches tall?



68 70 (height values)
-0.71 0 (standardized values)

Health and Nutrition Examination Study of 1976-1980

- How tall must a man be to place in the lower 10% for men aged 18 to 24?



? 70 (height values)

Table A: Standard Normal Probabilities

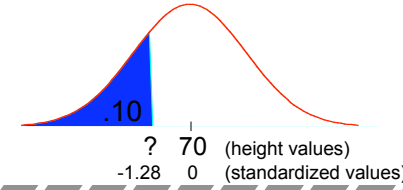
- See pages 690-691 in text for Table A.
- Look up the closest probability (to .10 here) inside the table.
- Find the corresponding standardized score.
- The value you seek is that many standard deviations from the mean.

Table A: Standard Normal Probabilities

	.08
-1.2	.1003

Health and Nutrition Examination Study of 1976-1980

- How tall must a man be to place in the lower 10% for men aged 18 to 24?



Observed Value for a Standardized Score

- Need to "unstandardize" the z-score to find the observed value (x) :

$$z = \frac{x - \mu}{\sigma} \Rightarrow x = \mu + z\sigma$$

- ***observed value*** =
  mean plus [(standardized score) × (std dev)]

Observed Value for a Standardized Score

- *observed value* =
  mean plus [(standardized score) × (std dev)]
  = 70 + [(−1.28) × (2.8)]
  = 70 + (-3.58) = 66.42
- A man would have to be approximately 66.42 inches tall or less to place in the lower 10% of all men in the population.