Math 140
Introductory Statistics

Professor Silvia Fernández

Lectures 7
Based on the book *Statistics in Action*
by A. Watkins, R. Scheaffer, and G. Cobb.

The Two Main Problems in the Standard Normal Distribution.

Unknown Percentage. (Given z, find P)

- With Table A (end of the textbook)
  - Use the units and the first decimal to locate the row and the closest hundredths digits to locate the column. The number found is the percentage of the number of values below z.
- With Calculator
  - Enter `normalcdf`(-99999, z) to get the percentage of the number of values below z.

Example (given z find P)

Find the percentage, P, of values below $z = 1.23$.

Calculator

P = normalcdf(-99999,1.23)

= .8906513833

~ 89.07%

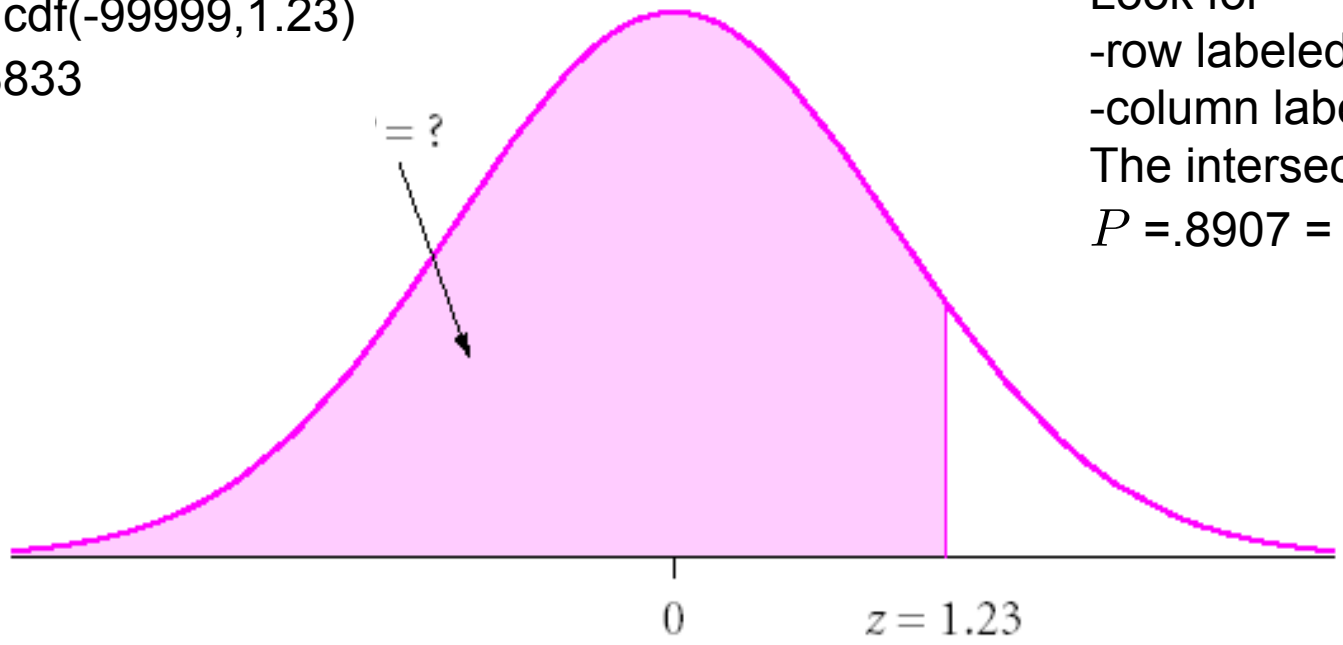


Table A

Look for

-row labeled 1.2

-column labeled .03

The intersection shows

P =.8907 = **89.07%**

Display 2.77 The percentage of values below $z = 1.23$.

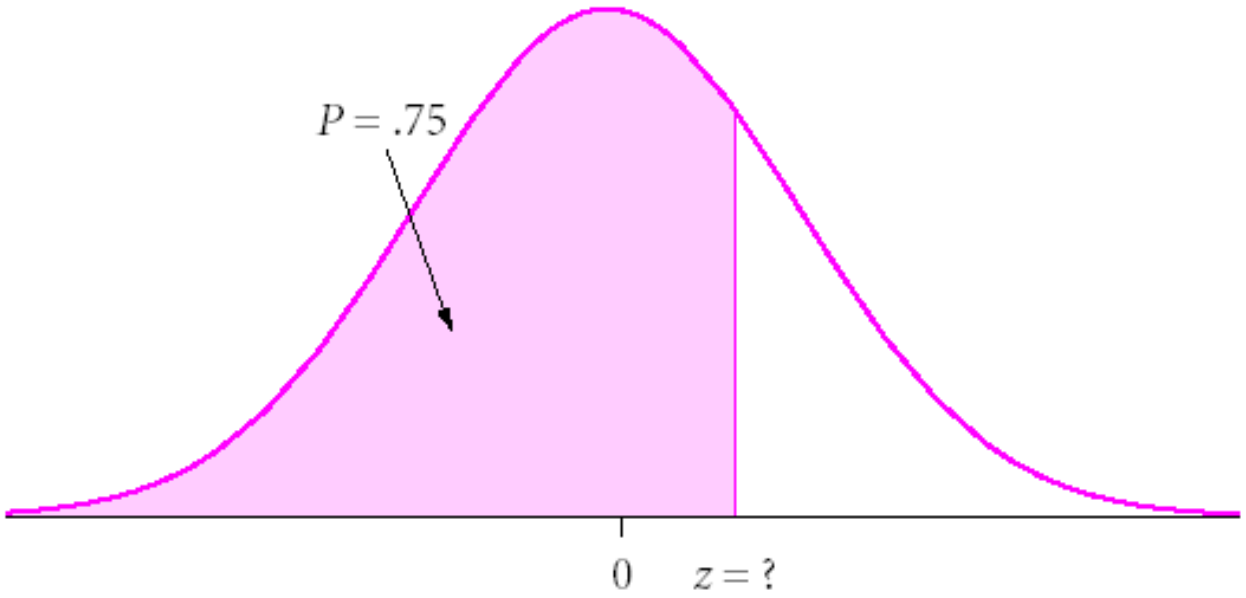
The Two Main Problems in the Standard Normal Distribution.

Unknown Value Problem. (Given P, find z)

- With Table A
  - Look for P in the body of the table. (or the number closest to it). Read back the row and column for that number. Use the row as the units and tenths of z, and the column as the hundredths digits of z. Note that P must be a percentage (written as a proportion, that is, a number between 0 and 1) of the number of values below a certain value z.
- With Calculator
  - Enter `invNorm`(P) to get the value z such that P equals the percentage of the number of values below z.

Example (given P find z-score)

Find the z-score that falls at the 75th percentile of the standard normal distribution; that is, the z-score that divides the bottom 75% of the values from the rest.



Calculator

z=invNorm(.75)

= .6744897495

~ .67

Table A

The value closest to .75 in the body of table A is .7486, which is in row .6 and column .07.

Then the z-score is **.67**

Value $\leftrightarrow$ z-score $(x \leftrightarrow z)$

- Standardizing (from x to z)

$$z = \frac{x - \overline{x}}{SD}$$

- Unstandardizing (from z to x)

$$x = \overline{x} + z(SD)$$

The two main problems (summary)

Unknown percentage

given x, find P

x to z to P

$$z = \frac{x - \overline{x}}{SD}$$

`normalcdf`(-99999, z)

Table: row and column

Unknown value

given P, find x

P to z to x

`invNorm`(P)

$$x = \overline{x} + z(SD)$$

Table: body

Example

Example

For groups of similar individuals, heights are often approximately normal in their distribution. For example, the heights of 18- to 24-year-old males in the United States are approximately normal, with mean 70.1 inches and standard deviation 2.7 inches. What percentage of these males are more than 74 inches tall?

Source: *Statistical Abstract of the U.S. 1991.*

Standardize (get z)

$x = 74$

$$z = \frac{x - \bar{x}}{SD} = \frac{74 - 70.1}{2.7} = 1.4444$$

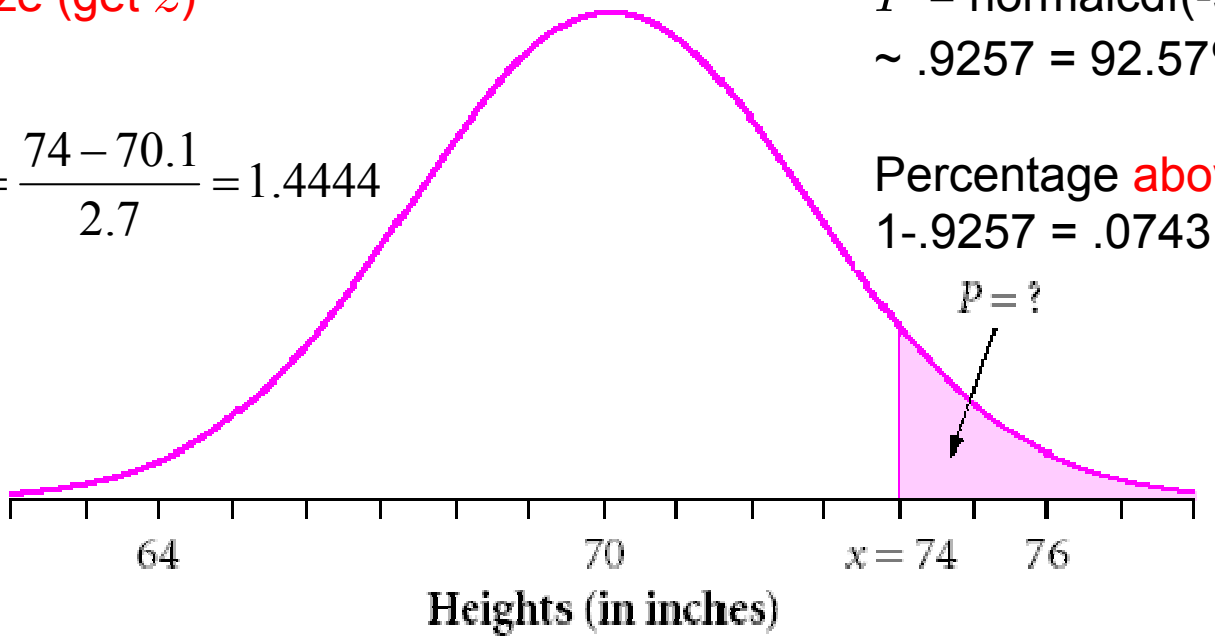
Percentage below 74 in

P = normalcdf(-99999,1.4444)

~ .9257 = 92.57%

Percentage above 74 in

1-.9257 = .0743 = **7.43%**



Example

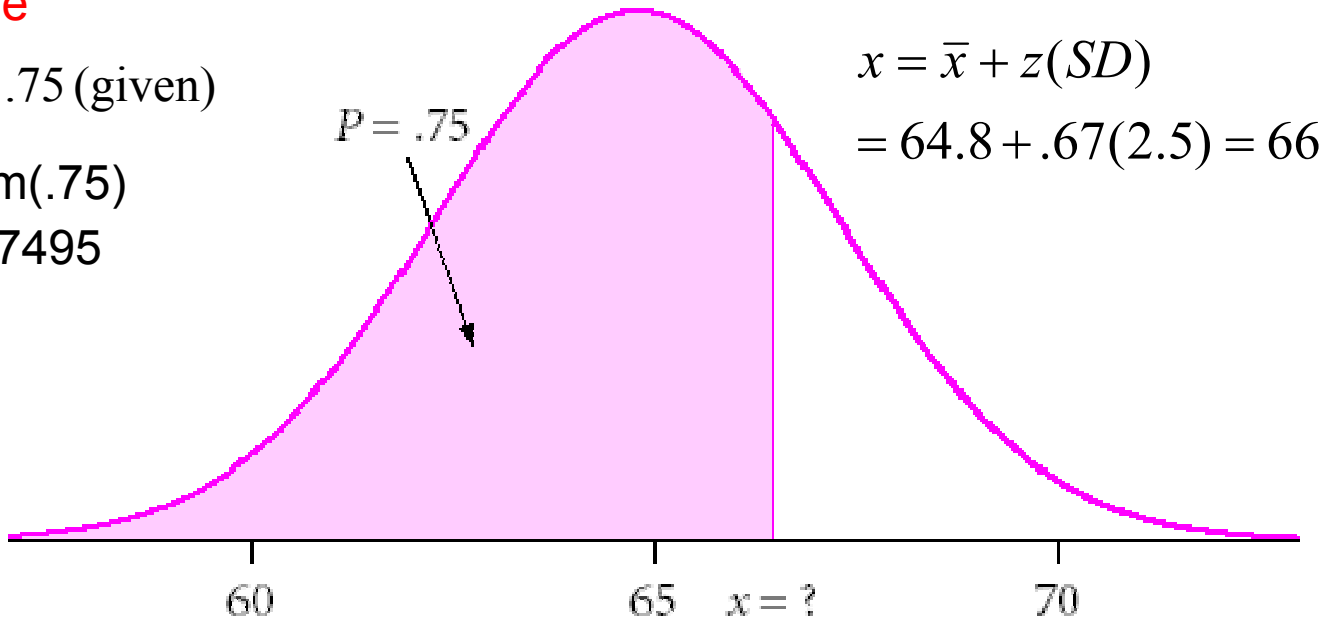
Example

The heights of females in the United States who are between the ages of 18 and 24 are approximately normally distributed, with mean 64.8 inches and standard deviation 2.5 inches. What height separates the shortest 75% from the tallest 25%?

Get z-score

$P = 75\% = .75$ (given)

z =invNorm(.75)
= .6744897495
~ .67



Unstandardize (get x)

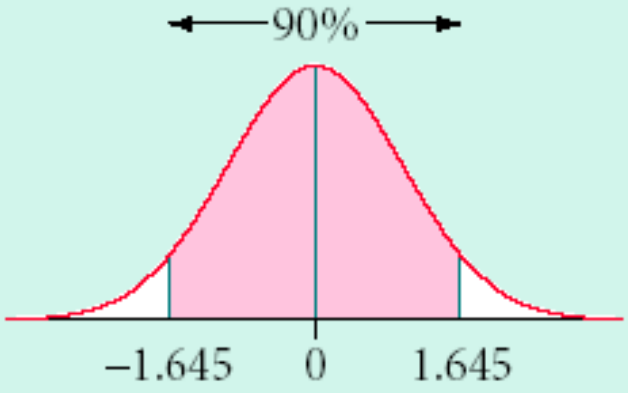
$x = \bar{x} + z(SD)$

$= 64.8 + .67(2.5) = 66.475$ in

Example

- According to the table on page 87, the distribution of death rates from cancer per 100,000 residents by state is approximately normal*, with mean 196 and *SD* 31. The middle 90% of death rates are between what two numbers?

90% of the values lie within 1.645 standard deviations of the mean.



*Provided that Alaska and Utah, which are outliers because of their unusually young populations, are left out.

Example (p. 91 – given P find x) cont.

- According to the table on page 87, the distribution of death rates from cancer per 100,000 residents by state is approximately normal*, with mean 196 and *SD* 31. The middle 90% of death rates are between what two numbers?
- Get z-scores (middle 90% is betwee 5% and 95%)

  5% =.05 corresponds to $z = \text{-}1.64485$

  95%= .95 corresponds to $z = 1.64485$
- Unstandardize

$$x = \bar{x} + z(SD) = mean + z(SD) = 196 + (-1.64485)(31) = 145.00965$$

$$x = x + z(SD) = mean + z(SD) = 196 + (1.64485)(31) = 246.99035$$

- So the middle 90% of states have death rates between 145.009 and 246.99 deaths per 1000,000 residents.

*Provided that Alaska and Utah, which are outliers because of their unusually young populations, are left out.

People with >50% homework (45/130 ~ 35%)

Aguilar, Evelyn
Assad, Michael
Basraon, Gurpreet
Blake, Eduardo
Bragado, Michelle
Butner, Christopher
Castillo, Clemence
Coney, Jocelyn
Curiel, Yessenia
Cuthbert, Michael
Davila, Andrea
Dong, Huiwen
Douglas, Arianna
Earley, Michele
Evans, Dena
Franco, Evelyn

Gibo, Katsumi
Gino, Rojeh
Guan, Bowen
Guevara, Kate
Hanner, Erik
Heredia, Elizabeth
Hernandez,Eric
Jackson, Patricia
Kaden, Sadie
Kim, Dong
Lang, Justin
Lin, Paulo
Martinez, Jeanett
Massey, Henry
Mou, Beixi
Muradyan, Marianna

Nelson, Emily
Nieto, Gregorio
Nunally, Richard
Orsini, Marianna
Poepping, Samantha
Ramirez, Andersson
Rasheed, Maisha
Rochana, Erica
Sanders, Mikita
Sparks, Adam
Tamayo, Raquel
Taylor-Dancy, William
Umali, Laura
Vernola, Kayla
Wang, Mengqi

65/130=50% of registered students do not have any worked saved.

Problem 19 – Homework 1

Introduced in 2000, the Honda Insight was the first hybrid car sold in the U.S. The mean gas mileage for the model year 2006 Insight with an automatic transmission is 53 miles per gallon on the highway. Suppose the gasoline mileage of this automobile is approximately normally distributed with a standard deviation of 2.2 miles per gallon.

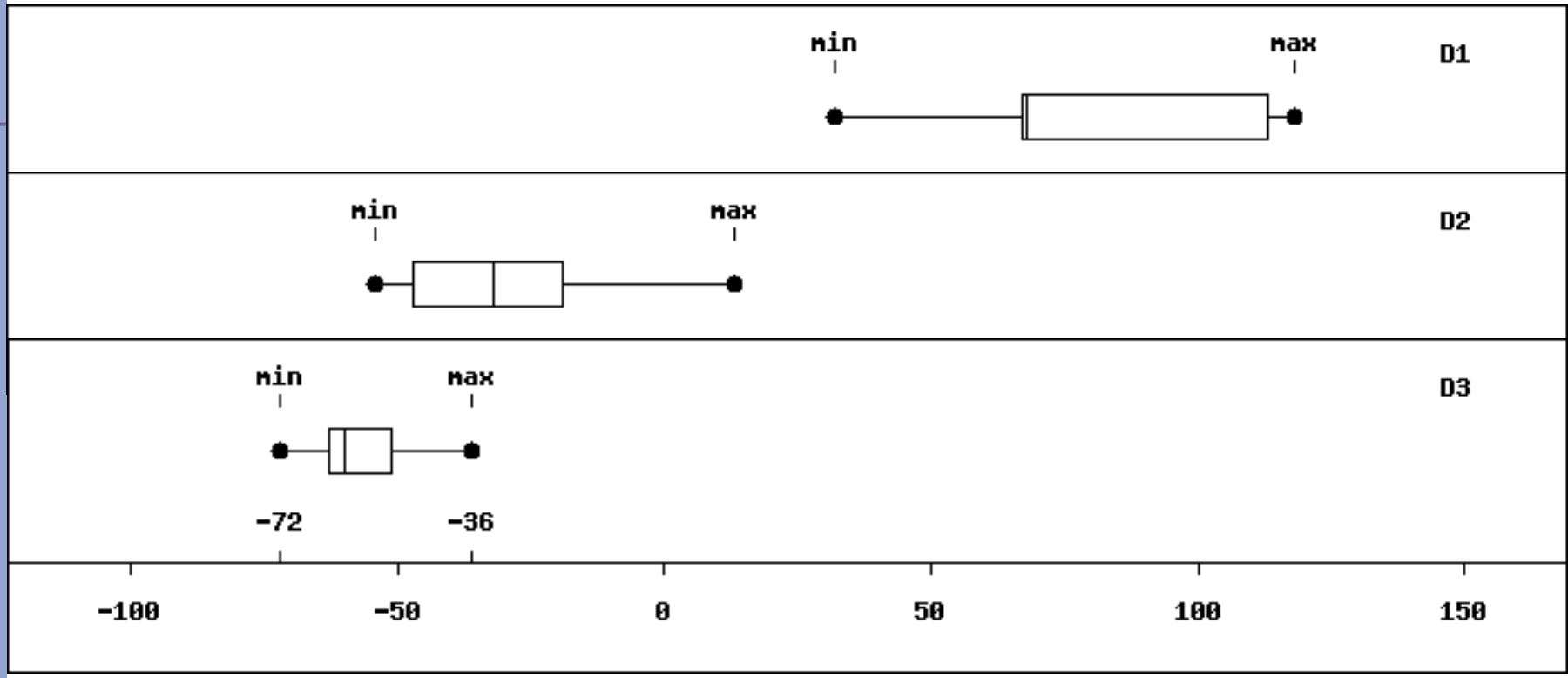
- (a) What proportion of 2006 Honda Insights with automatic transmission gets 60 miles per gallon or less on the highway?
- (b) What proportion of 2006 Honda Insights with automatic transmission gets between 58 and 62 miles per gallon on the highway?

Problem 16 – Homework 1

The scores of students on an exam are normally distributed with a mean of 581 and a standard deviation of 46.

- (a) What is the lower quartile score for this exam?
- (b) What is the upper quartile score for this exam?

Problem 4 – Homework 1



Which of the following are true?

A. At least three quarters of the data values represented in D1 are greater than the median value of D3 .
B. The data represented in D2 is symmetric.
C. The data for D1 has a greater median value than the data for D3 .
D. The data represented in boxplot D3 is skewed to the right.
E. All the data values for boxplot D1 are greater than the median value for D2 .
F. At least one quarter of the data values for D3 are less than the median value for D2

Problem 17 – Homework 1

IQ scores have a mean of 100 and a standard deviation of 15. Greg has an IQ of 127.

- What is the difference between Greg's IQ and the mean?
- Convert Greg's IQ score to a z score:

Problem 18 – Homework 1

Mike took 4 courses last semester: History, Spanish, Calculus, and Biology. The means and standard deviations for the final exams, and Mike's scores are given in the table below. Convert Mike's score into z scores.

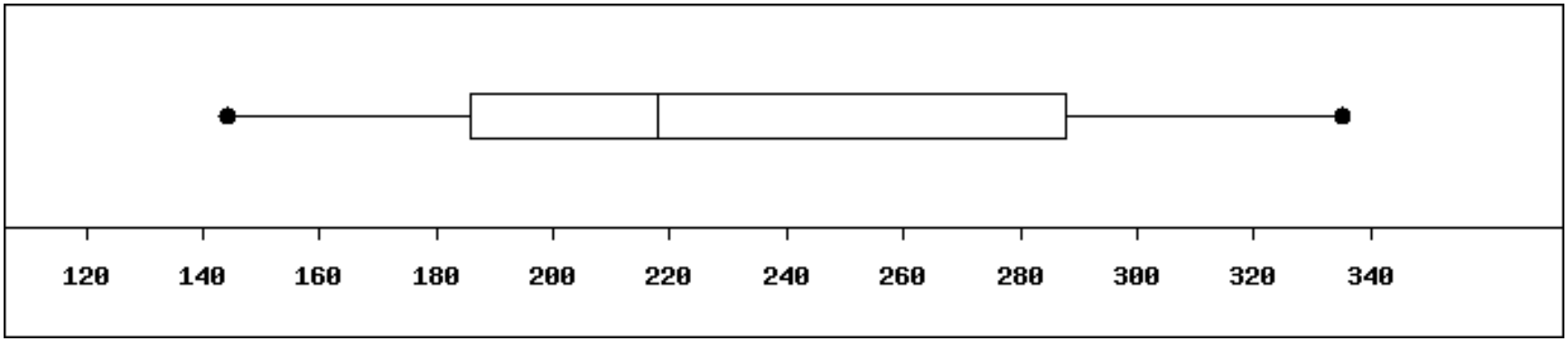
Subject	Mean	Standard deviation	Mike's score	Mike's z-score
History	53	16	49	
Spanish	44	12	38	
Calculus	70	12	88	
Biology	77	10	94.5	

- On what exam did Mike have the highest relative score?

Problem 5 – Homework 1

The boxplot below represents annual salaries of attorneys in thousands of dollars in Los Angeles. About what percentage of the attorneys have salaries between $186,000 and $288,000?

A. 20%
B. 50%
C. 25%
D. None of the Above



Problem 8 – Homework 1

Consider the following data set. Give the five number summary listing values in numerical order:

Data set: 27, 67, 26, 47, 78, 81, 73, 95, 88, 42, 96, 34, 82, 87, 37, 64, 56, 42, 100