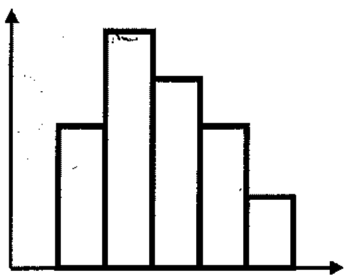
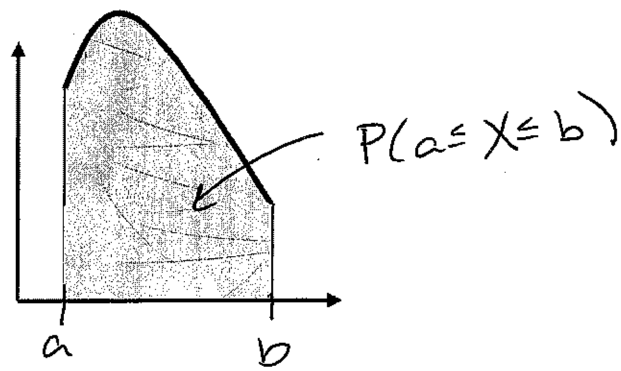
Ch 5 Notes Continuous Random Variables

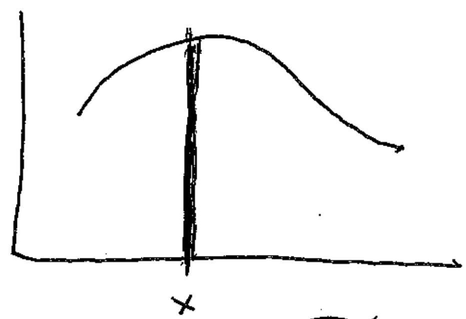
Characteristics of Continuous Random Variables

- **Outcomes are** measured   * Recall: discrete data v. cont. data
- **Geometrically, probability is an** area under a curve





- **P(X=x)=** 0



$P(X=x)$
$= P(x \le X \le x)$

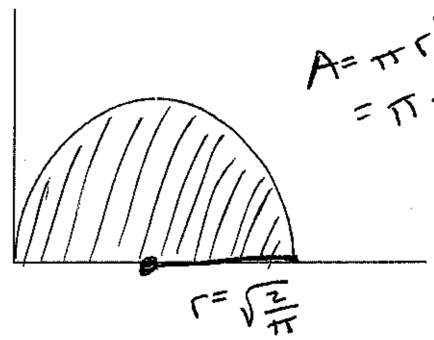
Continuous Probability Density Function/ Continuous PDF

A continuous PDF is a function $y = f(x)$

$* \ f(x) \geq 0$

- **f(X) is defined so that**

  the total area between the curve and the x-axis is one



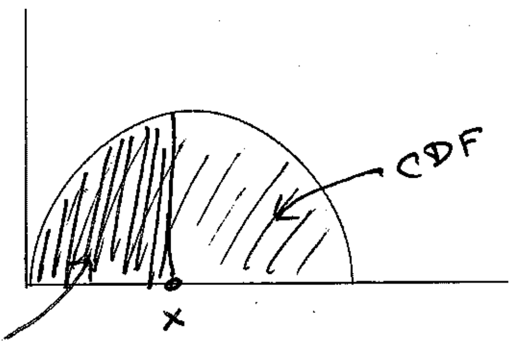
$A = \pi r^2 \cdot \frac{1}{2}$
$= \pi \cdot \frac{2}{\pi} \cdot \frac{1}{2} = 1$

- **The PDF is the** curve

Continuous Cumulative Density Function/ Continuous CDF

The CDF is the <u>area</u> under the curve

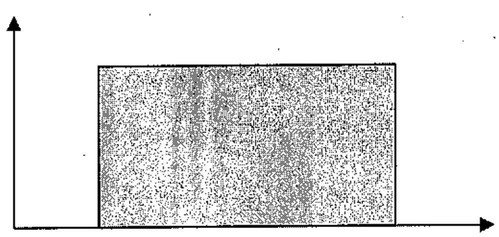
- $P(X \leq x)$



Area = $P(X \leq x)$

Some Types of Continuous PDF's

- **Uniform Distribution (Ch 5)**



*all X-values are equally likely

ex) random # generator

- **Exponential Distribution (Ch 5)**

* ex) The amount of time a USPS clerk spends with a customer

PDF is an exp. fn.

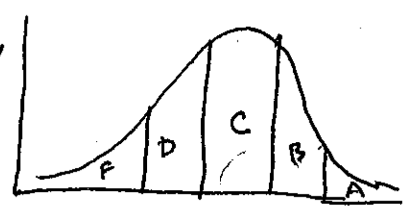
- Normal distr.

Ch. 6 Notes
Normal Distribution

The Normal Distribution is also called a bell curve

Examples: IQ scores
heights
grades (sometimes)

Graph:



The graph of the Normal Distribution is bell-shaped

Notation: $X \sim N(\mu, \sigma)$

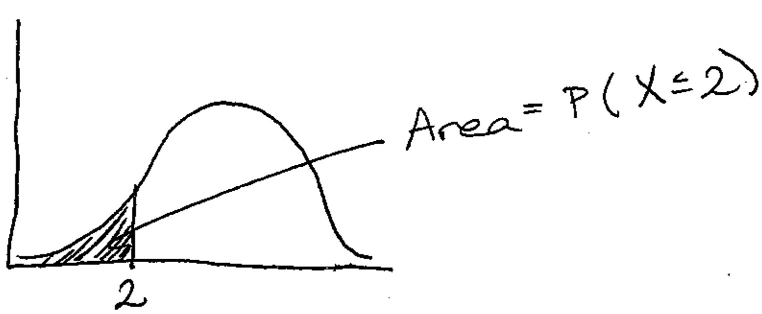
μ = mean
σ = std. dev.

μ tells us
where the "bell" is centered

σ tells us
how tall + wide the bell is

PDF = equation of the curve

CDF: $P(X \leq x)$ = area to the left of x



Example: The heights of De Anza women students are approximately normally distributed with a mean of 65 inches and a standard deviation of 3 inches.

1. Define the random variable in words.
   X = ht. of a female DA student

2. State the approximate distribution of X $\mu = 65$
   $X \sim N(65, 3)$ $\sigma = 3$

3. If one De Anza woman student is chosen at random, what is the probability that she is between 64 and 67 inches tall?

* Calc. instructions p. 95

$$P(\underbrace{64}_{a} < X < \underbrace{67}_{b}) = \overset{\text{(DISTR)}}{\text{normalcdf}}(\underset{a}{64}, \underset{b}{67}, \underset{\mu}{65}, \underset{\sigma}{3})$$

$$= 0.3781$$

So there's a 37.81% chance that a randomly chosen female DA student is between 5'4" and 5'7"

4. Find the probability that a randomly chosen woman student is less than 62 inches tall.

$P(X<62) = \text{normalcdf}(\underbrace{-1E99}_{=-1\times10^{99}}, 62, 65, 3)$

$= 0.1587$

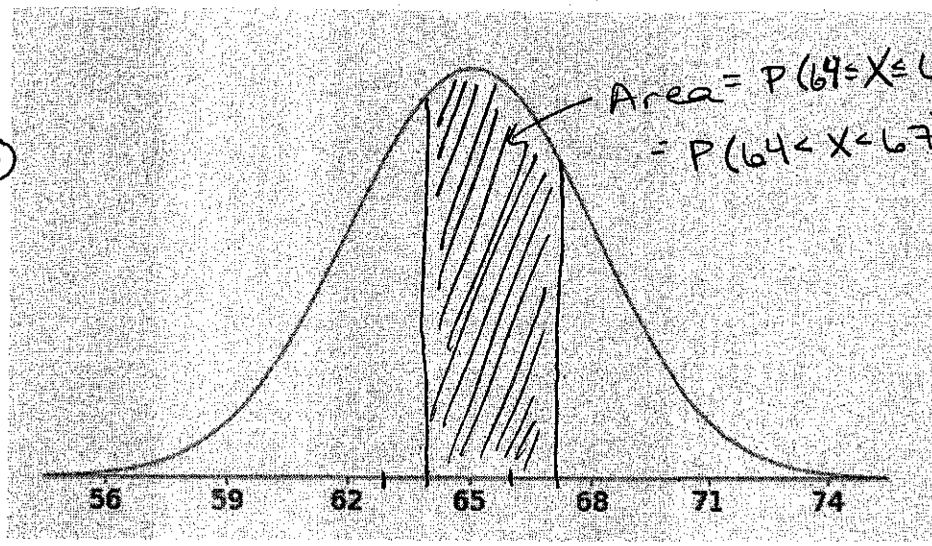
5. Find the probability that a randomly chosen woman student is more than $\underbrace{\text{72 inches}}_{\text{i.e. 6 ft.}}$ tall.

$P(X>72) = \text{normalcdf}(72, 1E99, 65, 3)$

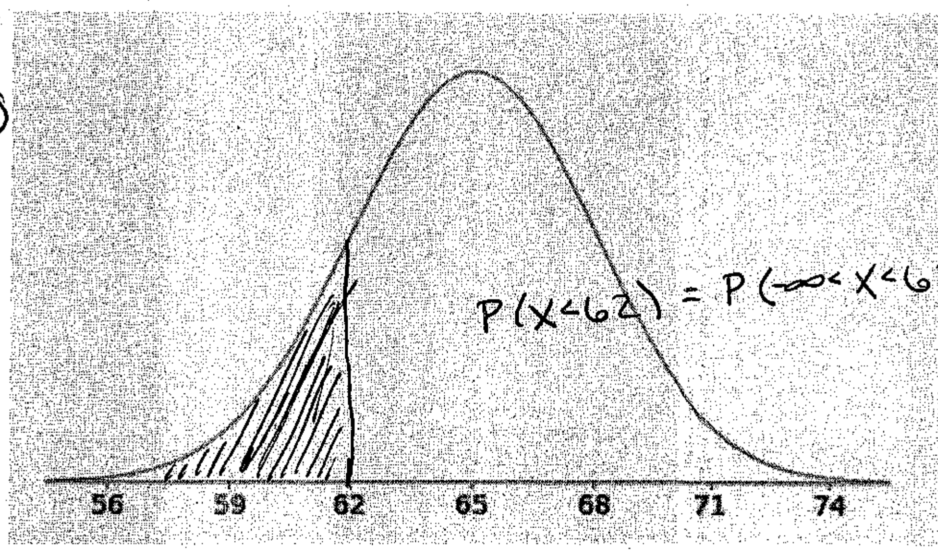
$= 0.0098$

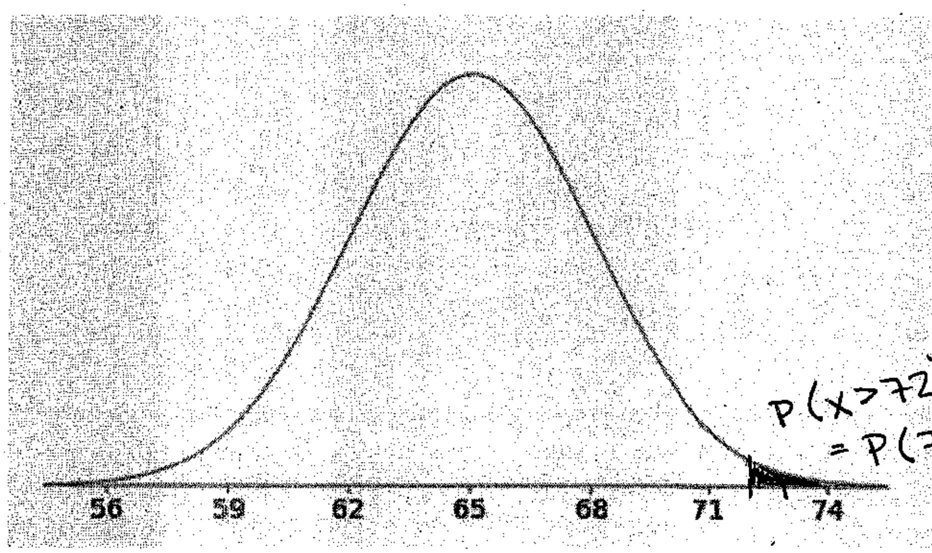
$X \sim N(65, 3)$

3


Area = P(64≤X≤67)
= P(64<X<67)
56
59
62
65
68
71
74

4


P(X<62) = P(-∞<X<62)
56
59
62
65
68
71
74


P(X>72)
= P(72<X<∞)
56
59
62
65
68
71
74

Percentiles ~~(continued)~~

The heights of De Anza women students are approximately normally distributed with a mean of 65 inches and a st dev of 3 inches.

X ~ N(65, 3)

*Calculator instr.

1. Find the 75th percentile of DA women's heights.

DISTR → invNorm (percentile, μ, σ)

invNorm (0.75, 65, 3) = 67.0235

75th percentile is about 67 inches
= 5'7"

2. The median of De Anza women's heights.

50th percentile

invNorm (0.5, 65, 3) = 65

*For bell curves
mean = median = mode

3. Find the middle 50% of De Anza women's heights.

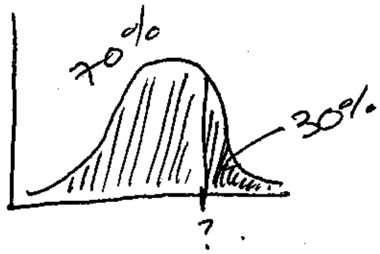
Q1 = 25th percentile = 62.9765 ≈ 63

Q3 = 75th percentile ≈ 67

The middle 50% is between 5'3" and 5'7"

4. 30% of all De Anza women are taller than how many inches?

70th percentile = 66.5732



Z-scores

A z-score tells us how many standard deviations a data value x is from the mean.

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

Example: Given X~N(27,5)

data value x	z-score z
17	-2 *2 std. devs. L of μ
27	0
32	1
20	-1.4
38.5	z=2.3

$\frac{17-27}{5}$ $\frac{27-27}{5}$ $\frac{32-27}{5}$ $\frac{20-27}{5}$

$$2.3 = \frac{x-27}{5}$$

$$11.5 = x - 27$$

$$38.5 = x$$

The closer a z-score is to 0, the closer x is to the mean μ.

Outliers

If a data value has a z-score of more than 3 or less than –3, it is an outlier.

Chapter 6: Calculating Z-scores

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

1. Given $X \sim N(8,2)$. Find z-scores for the following values of x. Then write a sentence stating what this means in terms of the mean and standard deviation.

a. $\mu =$ 8 $\sigma =$ 2

a. x = 10
z=1 x=10 is 1 std. dev. to the right of the mean

b. x = 5
z=-1.5 x=5 is 1.5 std. devs. to the left of the mean

$\frac{11.5-8}{2}$ c. x = 11.5
z=1.75 . . .

d. What value of x has a z-score of $z = -2.4$?
X=3.2 $-2.4 = \frac{x-8}{2}$ x=3.2 is 2.4 std. devs. to the left of the mean.
$-4.8 = x-8$

2. Given $X \sim N(100,10)$. Find z-scores for the following values of x. Then write a sentence stating what this means in terms of the mean and standard deviation.

a. $\mu =$______ $\sigma =$______

b. x = 105

c. x = 85

d. x = 142

e. What value of x has a z-score of $z = -2.4$?