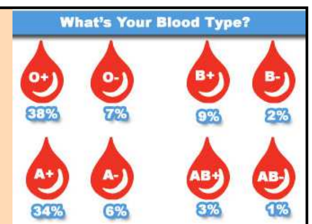
# Normal distributions

PSLS chapter 11



***REMINDER***

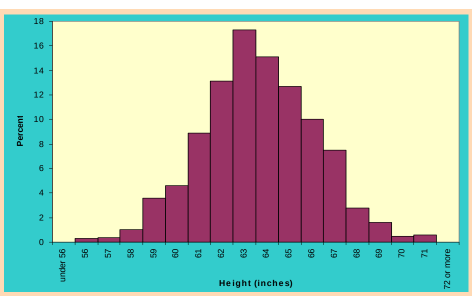
**Discrete sample space:**




For a random person:

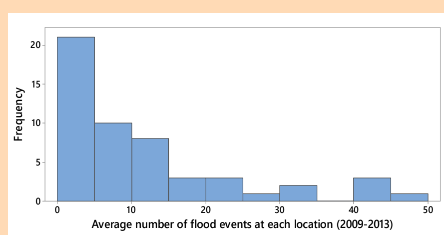
S = {O+, O-, A+, A-, B+, B-, AB+, AB-}

**Continuous sample space:**




For a random woman:

S = any value within a realistic range



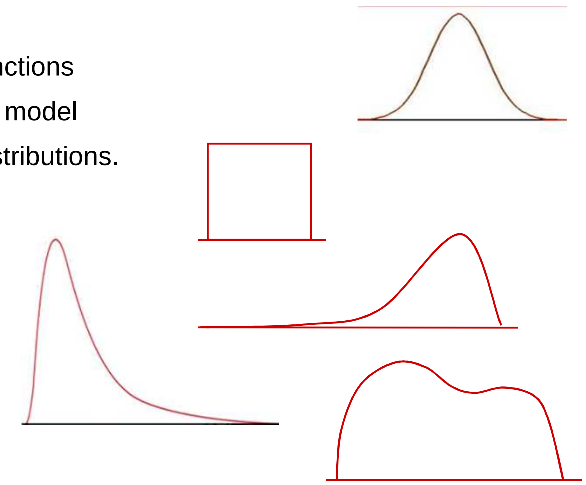

For a random coastal community:

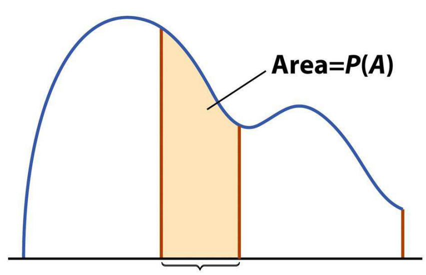
S = any positive value

# Continuous random variables

**Continuous sample spaces** contain an infinite number of outcomes over an interval of values.

We use mathematical functions called **density curves** to model continuous probability distributions.





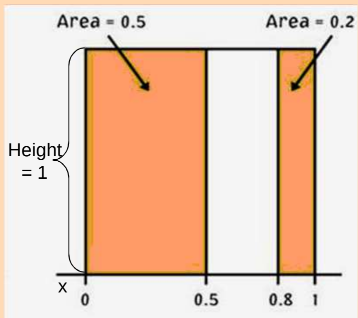

Events are defined over **intervals** of values.

Probability are computed as **areas** under the corresponding portion of the density curve.

The total area under a density curve represents the whole population (sample space) and equals 1 (100%).

***probability of randomly drawing one individual ⇔ population frequency***

Software generates at random with uniform probability a number between 0 and 1.
What is the probability $P(0 \le x \le 0.5)$ = ?




Uniform distribution over [0,1]

**The probability $P(0 < x < 0.5)$ is**

**A) smaller than the probability $P(0 \le x \le 0.5)$.**

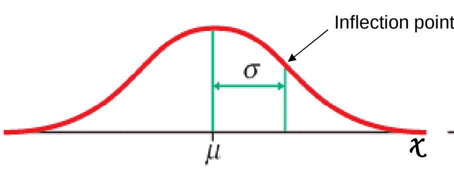
**B) equal to the probability $P(0 \le x \le 0.5)$.**

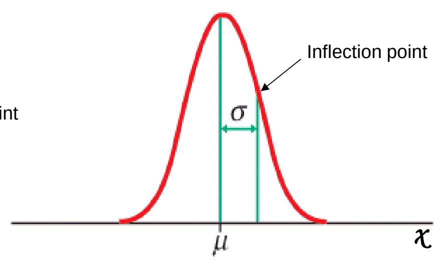
**C) greater than the probability $P(0 \le x \le 0.5)$.**

**The probability of a single numerical value is meaningless when the sample space is continuous.**

# Normal distributions

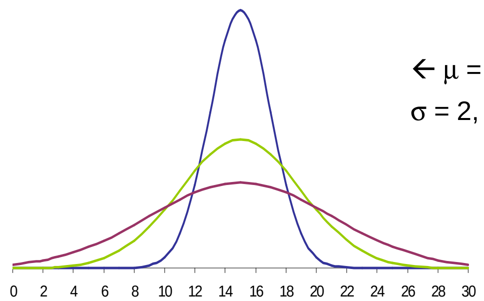
**Normal** – or Gaussian – **distributions** (defined from –∞ to +∞):




$$N(\mu,\sigma): f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

## Effect of changing $\mu$ or $\sigma$

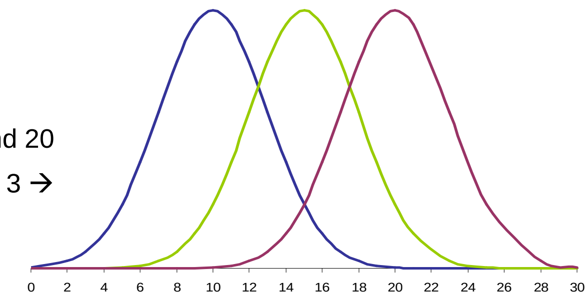



← $\mu = 15$

$\sigma = 2, 4, \text{ and } 6$

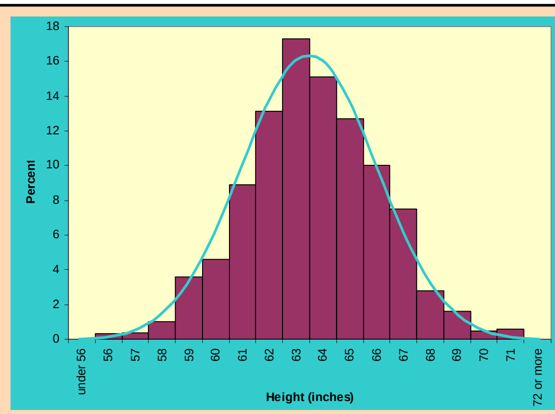
$\mu = 10, 15, \text{ and } 20$

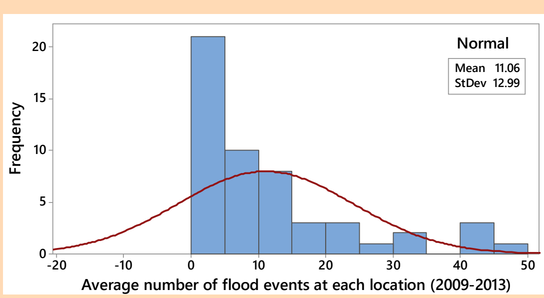
$\sigma = 3$ →




Women's heights in inches →

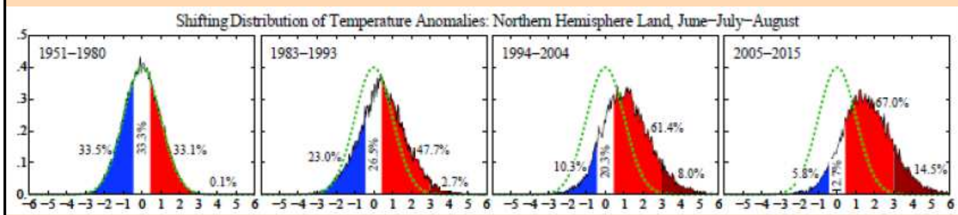
✓




~~← Average number of flood events in 50 ocean-side communities in the U.S (2009-13)~~

www.columbia.edu/~jeh1/mailings/2016/20160120_Temperature2015.pdf




*"The frequency distribution of summer mean temperature anomalies in Northern Hemisphere land areas has shifted by more than one standard deviation in the past decade relative to the bell curve of natural variability in the base period 1951-1980."*

Over time, the distribution of Northern **summer** land temperatures got:

**A)** a greater mean and a greater standard deviation

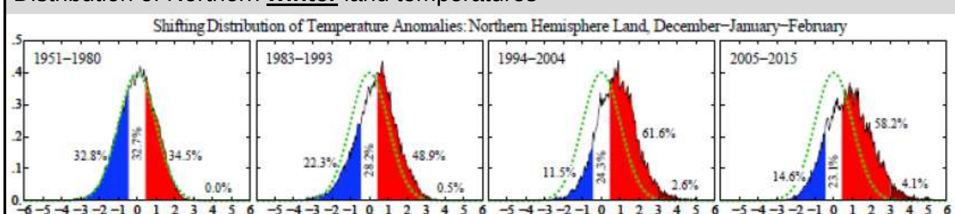
**B)** a greater mean and the same standard deviation

**C)** a greater mean and a smaller standard deviation

**D)** the same mean and a greater standard deviation

**E)** the same mean and a smaller standard deviation

Distribution of Northern **winter** land temperatures




**Climate Change Threatens an Iconic Desert Tree**

It's not just the polar bear. Animals and plants in Earth's other extreme environment—the desert—are endangered by rising temperatures.

At Joshua Tree National Park in California's Mojave Desert, these tough, gnarled plants are threatened by climate change. A survey of the park found few or no young trees in roughly 30 percent of their range.

PHOTOGRAPH BY KEVIN SCHAFER, MINDEN PICTURES/CORBIS

By **Osha Gray Davidson**, National Geographic

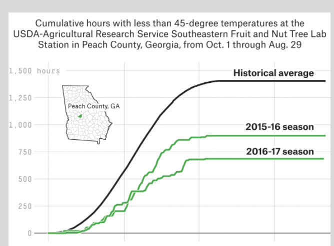
PUBLISHED OCTOBER 28, 2015

THE SALT

**If These Trees Don't Get Time To Chill, Farmers Will Be Out On A Limb**

by EZRA DAVID ROMERO

January 25, 2017 • KVPR Tree crops like pistachios, peaches and almonds need a certain amount of cold weather every year. But scientists say that California's climate may become too warm for them to grow there.

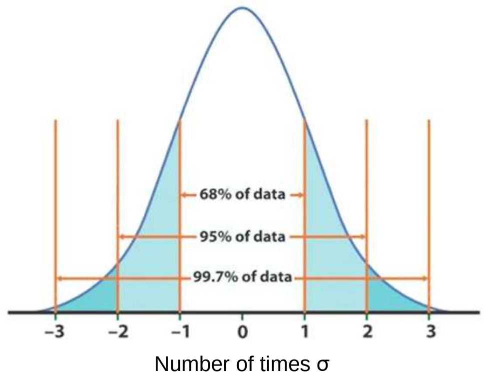



SEP. 14, 2017 FiveThirtyEight SOURCE: AGROCLIMATE

**How A Warm Winter Destroyed 85 Percent Of Georgia's Peaches**

## All normal curves N(μ,σ) share the same properties

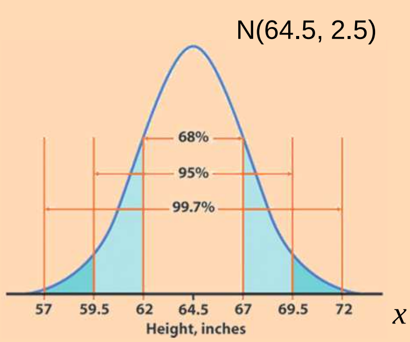
- About 68% of all observations under N(μ,σ) are within $\mu \pm \sigma$.
- About 95% of all observations under N(μ,σ) are within $\mu \pm 2\sigma$.
- Almost all (99.7%) observations under N(μ,σ) are within $\mu \pm 3\sigma$.



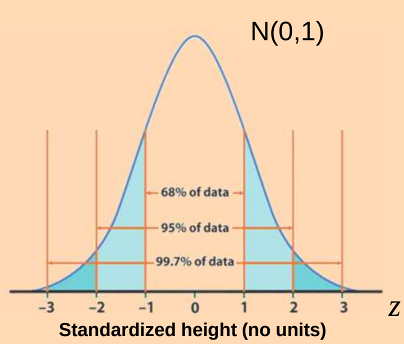

## The standard Normal distribution

We can ***standardize*** data by computing $\boxed{z = \frac{(x-\mu)}{\sigma}}$

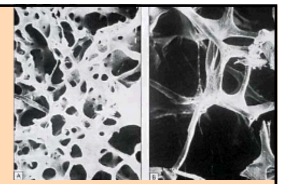
If $x$ has the N(μ,σ) distribution, then $z$ has the N(0,1) distribution, also called "the standard Normal distribution.".



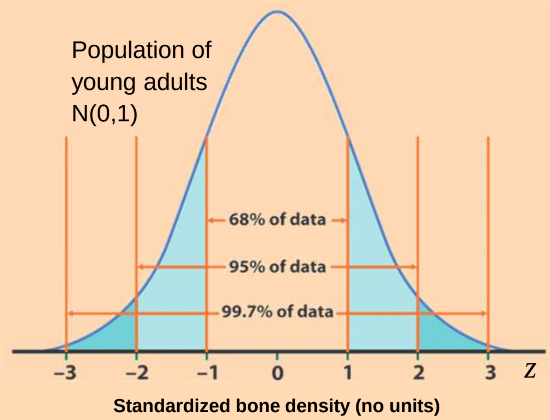

=>



| World Health Organization definitions of osteoporosis based on standardized bone density levels | |
|---|---|
| ***Normal*** | Bone density is within 1 standard deviation (> –1) of the young adult mean or above. |
| ***Low bone mass*** | Bone density is 1 to 2.5 standard deviation below the young adult mean (between –2.5 and –1). |
| ***Osteoporosis*** | Bone density is 2.5 standard deviation or more below the young adult mean (≤ –2.5). |



*Healthy bone* *Osteoporosis*




What percent of young adults have a **healthy** bone density (value > –1)?

**A)** 100%

**B)** 97.5%

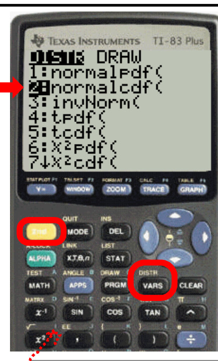
**C)** 95%

**D)** 84%

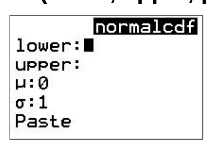
**E)** 68%

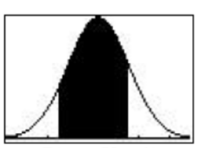
## TI 83/84

2nd DISTR



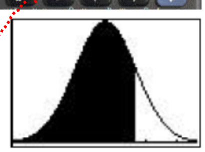

**normalcdf(lower, upper, μ, σ)**



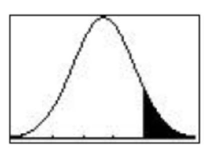
$P(a \le X \le b)$

normalcdf($a$, $b$, μ, σ).



$P(X < k) = P(X \le k)$

normalcdf(–1E99, $k$, μ, σ).



$P(X > k) = P(X \ge k)$

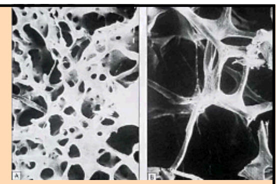
normalcdf($k$, 1E99, μ, σ).

To get "E" type "2nd" then "EE"

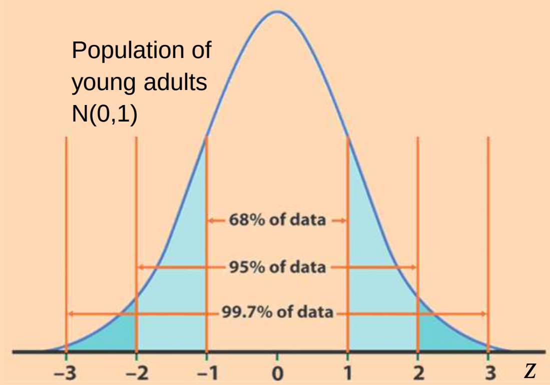
–1E99 and 1E99 are the equivalent of –∞ and +∞ for the calculator

$P$(-2 < z < 2) = normalcdf(-2, 2, 0, 1) = 0.9545 ≈ 95%

| World Health Organization definitions of osteoporosis based on standardized bone density levels | |
|---|---|
| ***Normal*** | Bone density is within 1 standard deviation (> –1) of the young adult mean or above. |
| ***Low bone mass*** | Bone density is 1 to 2.5 standard deviation below the young adult mean (between –2.5 and –1). |
| ***Osteoporosis*** | Bone density is 2.5 standard deviation or more below the young adult mean (≤ –2.5). |

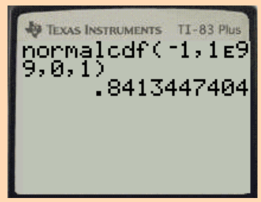


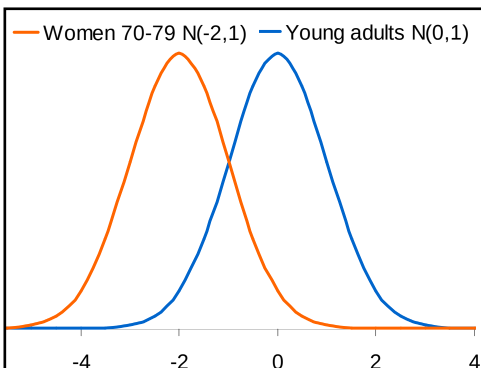
*Healthy bone* *Osteoporosis*



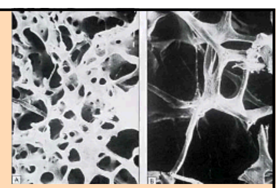

**Standardized bone density (no units)**

What percent of young adults have a **healthy** bone density (> –1)?




**Bone density (no units)**



*Healthy bone* *Osteoporosis*

Women aged 70 to 79 are NOT young adults. The mean bone density in this age is about −2 on the standard scale for young adults.

*What percent of women in their 70s have osteoporosis (BD ≤ –2.5)?
*What is the probability that a woman in her 70s has osteopenia (–2.5 < BD ≤ –1)?

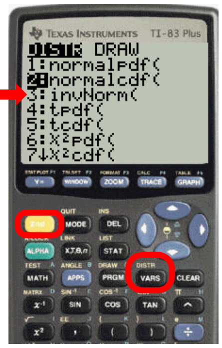
The answer is a value between

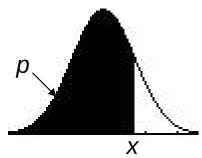
**A)** 0.8 and 1.0
**B)** 0.6 and 0.8
**C)** 0.4 and 0.6
**D)** 0.2 and 0.4
**E)** 0.0 and 0.2

# Inverse Normal calculations

We may also seek the range of values that correspond to a given proportion/ area under the curve.

**TI-83 computations**





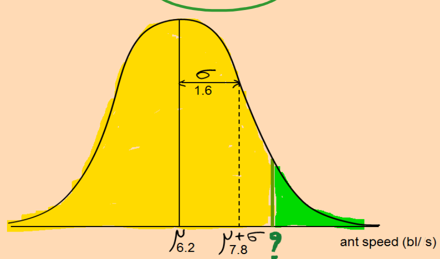
**$x$ = invNorm($p$, $\mu$, $\sigma$)**

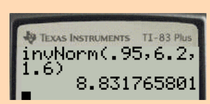
To find the value of $x$ such that there is area $p$ to the **left** of $x$.

➔ ***x* is the *p*[th] percentile**

A research paper examines the speed of ants on trails in their natural environment. The paper reports that, when traffic is light and not congested, ant speeds (X) vary roughly Normally with mean 6.2 body lengths per second (bl/s) and standard deviation 1.6 bl/s.

**How fast do the 5% fastest ants go in light traffic?**




**The 5% fastest ants have speeds of 8.83 bl/s and higher.**

**The 10% slowest ants in light traffic have speeds of**

$Q_1$ = ?

A) up to 4.15 bl/s.
B) at least 4.15 bl/s.
C) 8.25 bl/s.
D) up to 8.25 bl/s.
E) at least 8.25 bl/s.