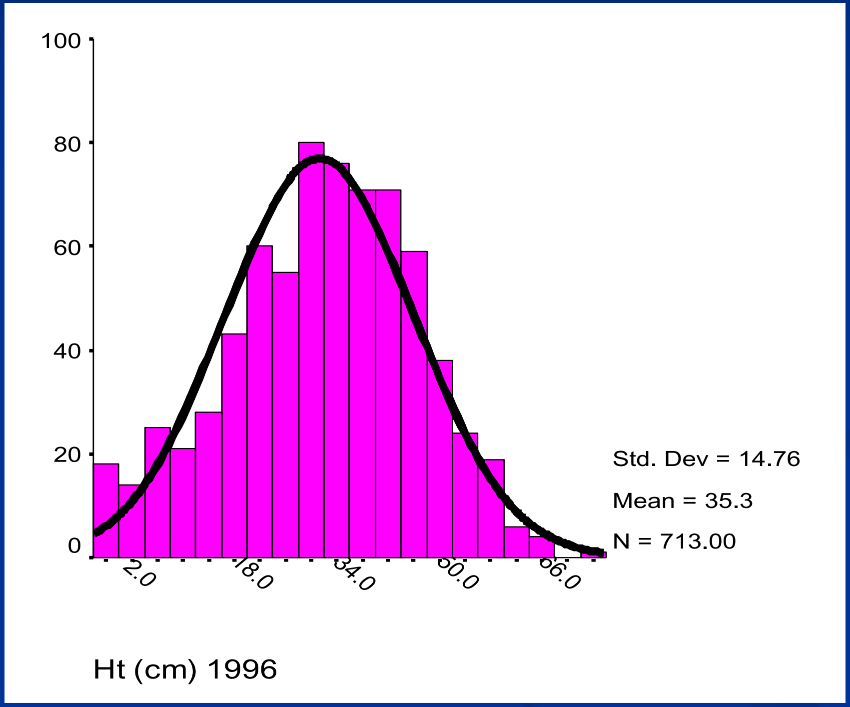
The Binomial, Poisson, and Normal Distributions

Probability distributions

- We use probability distributions because they work –they fit lots of data in real world



Height (cm) of *Hypericum cumulicola* at Archbold Biological Station

Random variable

- The mathematical rule (or function) that assigns a given numerical value to each possible outcome of an experiment in the sample space of interest.

Types of Random variables

- Discrete random variables
- Continuous random variables

The Binomial Distribution
Bernoulli Random Variables

- Imagine a simple trial with only two possible outcomes
  - Success (*S*)
  - Failure (*F*)

Jacob Bernoulli (1654-1705)

- Examples
  - Toss of a coin (heads or tails)
  - Sex of a newborn (male or female)
  - Survival of an organism in a region (live or die)

The Binomial Distribution
Overview

- Suppose that the probability of success is p
- What is the probability of failure?
  - $q = 1 - p$
- Examples
  - Toss of a coin (S = head): $p = 0.5 \Rightarrow q = 0.5$
  - Roll of a die ($S = 1$): $p = 0.1667 \Rightarrow q = 0.8333$
  - Fertility of a chicken egg (S = fertile): $p = 0.8 \Rightarrow q = 0.2$

The Binomial Distribution
Overview

- Imagine that a trial is repeated n times

- Examples
  - A coin is tossed 5 times
  - A die is rolled 25 times
  - 50 chicken eggs are examined

- Assume p remains constant from trial to trial and that the trials are statistically independent of each other

The Binomial Distribution
Overview

- What is the probability of obtaining x successes in n trials?

- Example
  - What is the probability of obtaining 2 heads from a coin that was tossed 5 times?

$$P(HHTTT) = (1/2)^5 = 1/32$$

The Binomial Distribution
Overview

- But there are more possibilities:

HHTTT	HTHTT	HTTHT	HTTTH
	THHTT	THTHT	THTTH
		TTHHT	TTHTH
			TTTHH

$P(2 \text{ heads}) = 10 \times 1/32 = 10/32$

The Binomial Distribution
Overview

- In general, if trials result in a series of success and failures,

$$FFSFFFFSFSFSSFFFFFSF\ldots$$

Then the probability of x successes in that order is

$$\begin{aligned} P(x) &= q \cdot q \cdot p \cdot q \cdot \ldots \\ &= p^x \cdot q^{n-x} \end{aligned}$$

The Binomial Distribution
Overview

- However, if order is not important, then

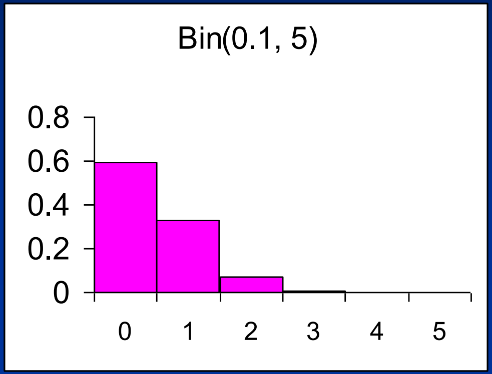
$$P(x) = \frac{n!}{x!(n-x)!} p^x \cdot q^{n-x}$$

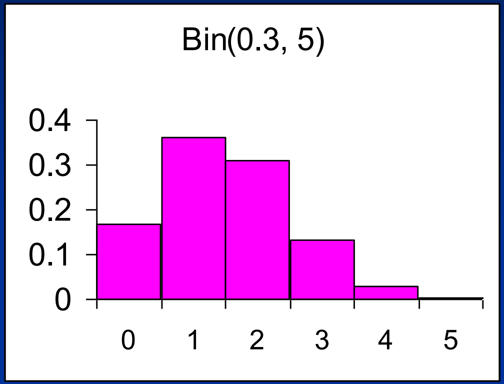
where $\frac{n!}{x!(n-x)!}$ is the number of ways to obtain x successes

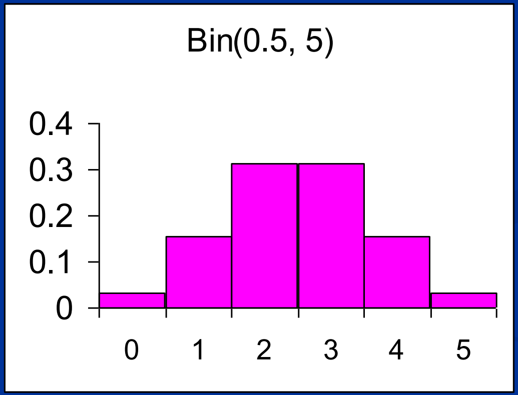
in n trials, and $i! = i \cdot (i-1) \cdot (i-2) \cdot \ldots \cdot 2 \cdot 1$

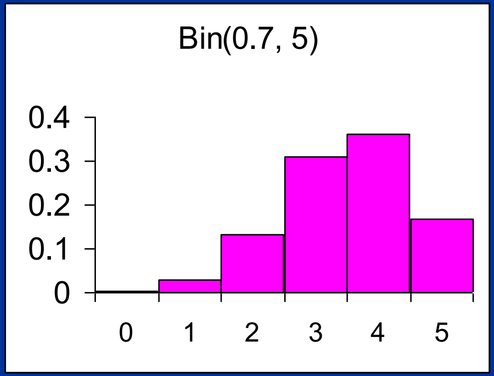
The Binomial Distribution

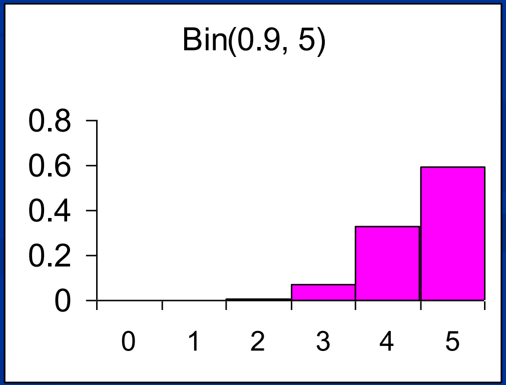
Overview





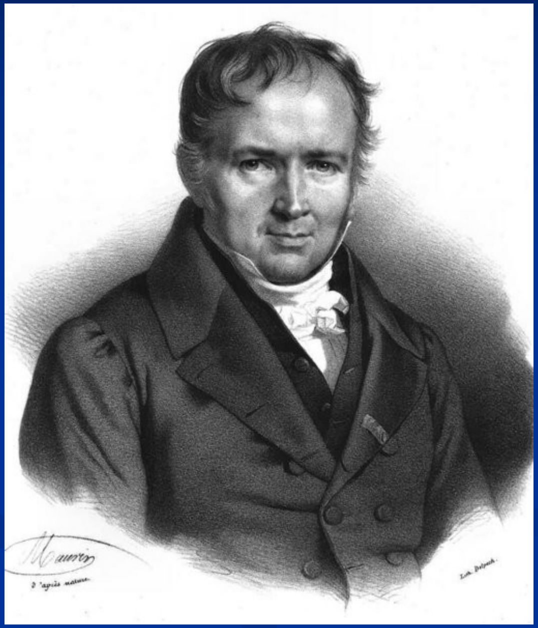






The Poisson Distribution
Overview

- When there is a large number of trials, but a small probability of success, binomial calculation becomes impractical
  - Example: Number of deaths from horse kicks in the Army in different years



Simeon D. Poisson (1781-1840)

- The mean number of successes from n trials is $\mu = np$
  - Example: 64 deaths in 20 years from thousands of soldiers

The Poisson Distribution
Overview

- If we substitute μ/n for p, and let n tend to infinity, the binomial distribution becomes the Poisson distribution:

$$P(x) = \frac{e^{-\mu}\mu^{x}}{x!}$$

The Poisson Distribution
Overview

- Poisson distribution is applied where random events in space or time are expected to occur
- Deviation from Poisson distribution may indicate some degree of non-randomness in the events under study
- Investigation of cause may be of interest

The Poisson Distribution
Emission of α-particles

- Rutherford, Geiger, and Bateman (1910) counted the number of α-particles emitted by a film of polonium in 2608 successive intervals of one-eighth of a minute
  - What is n?
  - What is p?

- Do their data follow a Poisson distribution?

The Poisson Distribution

Emission of α-particles

- Calculation of μ:

$$\mu = \text{No. of particles per interval}$$
$$= 10097/2608$$
$$= 3.87$$

- Expected values:

$$2680 \times P(x) = 2608 \times \frac{e^{-3.87}(3.87)^x}{x!}$$

No. α-particles	Observed
0	57
1	203
2	383
3	525
4	532
5	408
6	273
7	139
8	45
9	27
10	10
11	4
12	0
13	1
14	1
Over 14	0
Total	2608

The Poisson Distribution

Emission of α-particles

No. α-particles	Observed	Expected
0	57	54
1	203	210
2	383	407
3	525	525
4	532	508
5	408	394
6	273	254
7	139	140
8	45	68
9	27	29
10	10	11
11	4	4
12	0	1
13	1	1
14	1	1
Over 14	0	0
Total	2608	2680

The Poisson Distribution

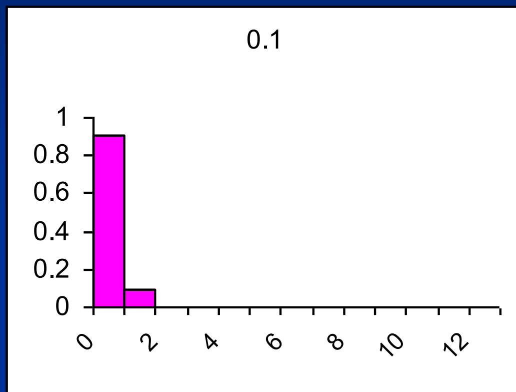
Emission of α-particles

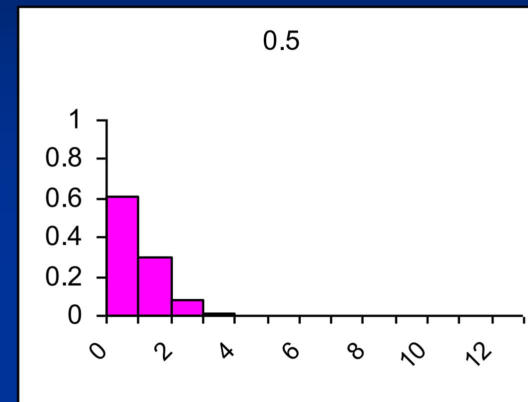
Random events

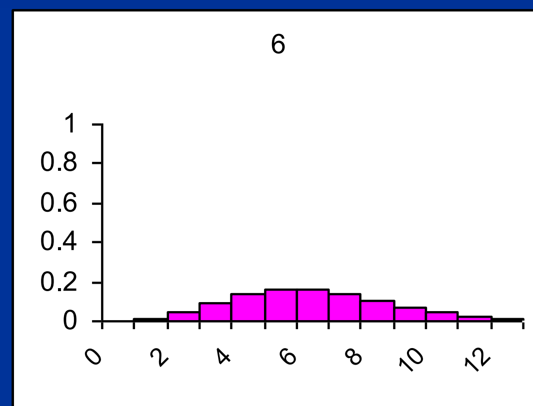
Regular events

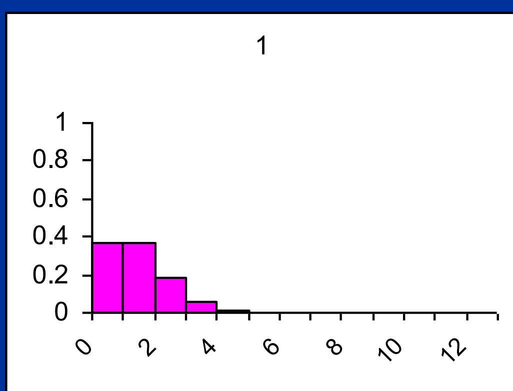
Clumped events

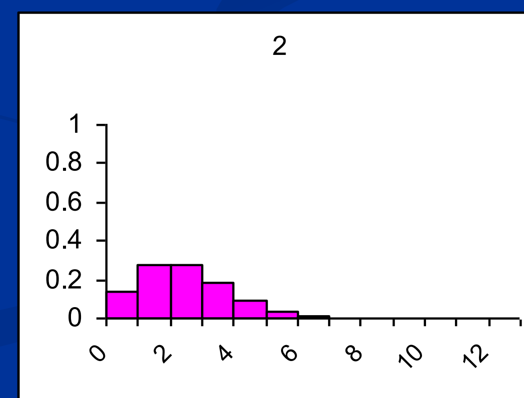
The Poisson Distribution











The Expected Value of a Discrete Random Variable

$$E(X) = \sum_{i=1}^{n} a_i p_i = a_1 p_1 + a_2 p_2 + \ldots + a_n p_n$$

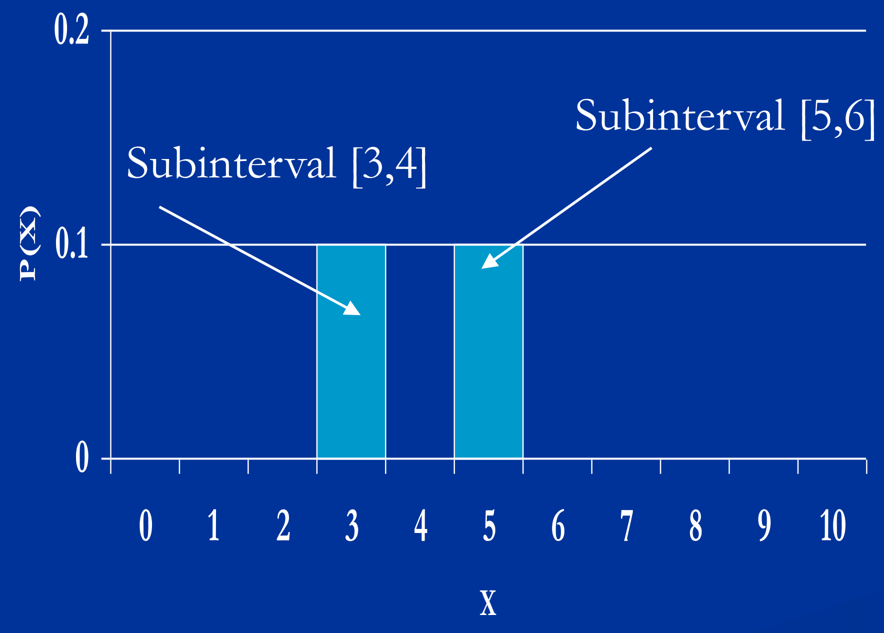
The Variance of a Discrete Random Variable

$$\sigma^2(X) = E\left[X - E(X)\right]^2$$

$$= \sum_{i=1}^{n} p_i \left(a_i - \sum_{i=1}^{n} a_i p_i \right)^2$$

Uniform random variables

- The closed unit interval, which contains all numbers between 0 and 1, including the two end points 0 and 1



$$f(x) = \begin{cases} 1/10, 0 \le x \le 10 \\ 0, otherwise \end{cases}$$

The probability density function

The Expected Value of a continuous Random Variable

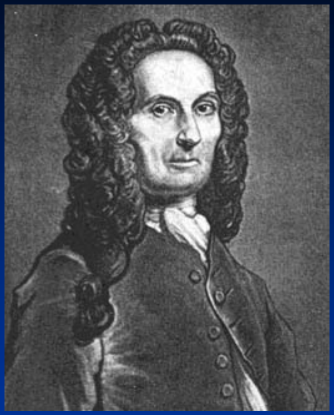
$$E(X) = \int xf(x)dx$$

For an uniform random variable *x*, where *f(x)* is defined on the interval [a,b], and where a<b,

$$E(X) = (b+a)/2 \quad \text{and} \quad \sigma^2(X) = \frac{(b-a)^2}{12}$$

The Normal Distribution

Overview



Abraham de Moivre
(1667-1754)

- Discovered in 1733 by de Moivre as an approximation to the binomial distribution when the number of trails is large
- Derived in 1809 by Gauss
- Importance lies in the Central Limit Theorem, which states that the sum of a large number of independent random variables (binomial, Poisson, etc.) will approximate a normal distribution
  - Example: Human height is determined by a large number of factors, both genetic and environmental, which are additive in their effects. Thus, it follows a normal distribution.

Karl F. Gauss
(1777-1855)

The Normal Distribution

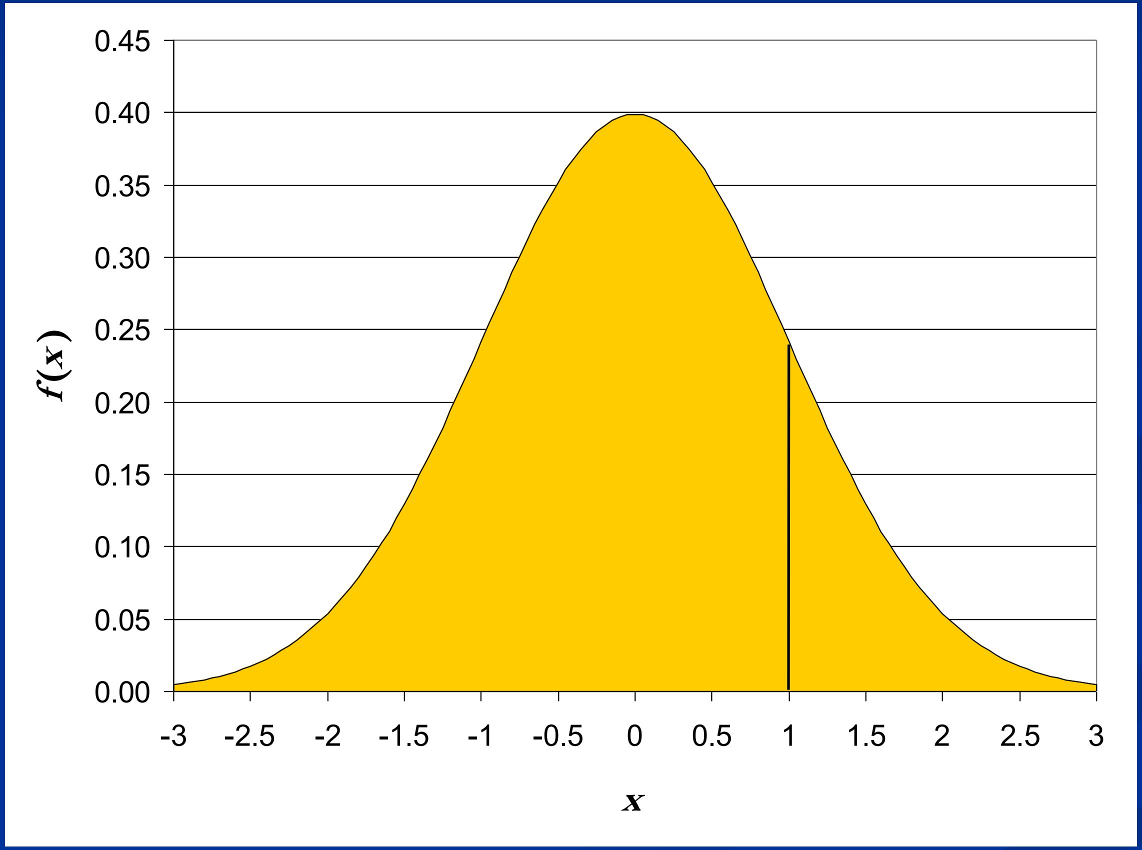
Overview

- A continuous random variable is said to be normally distributed with mean μ and variance σ^2 if its probability density function is

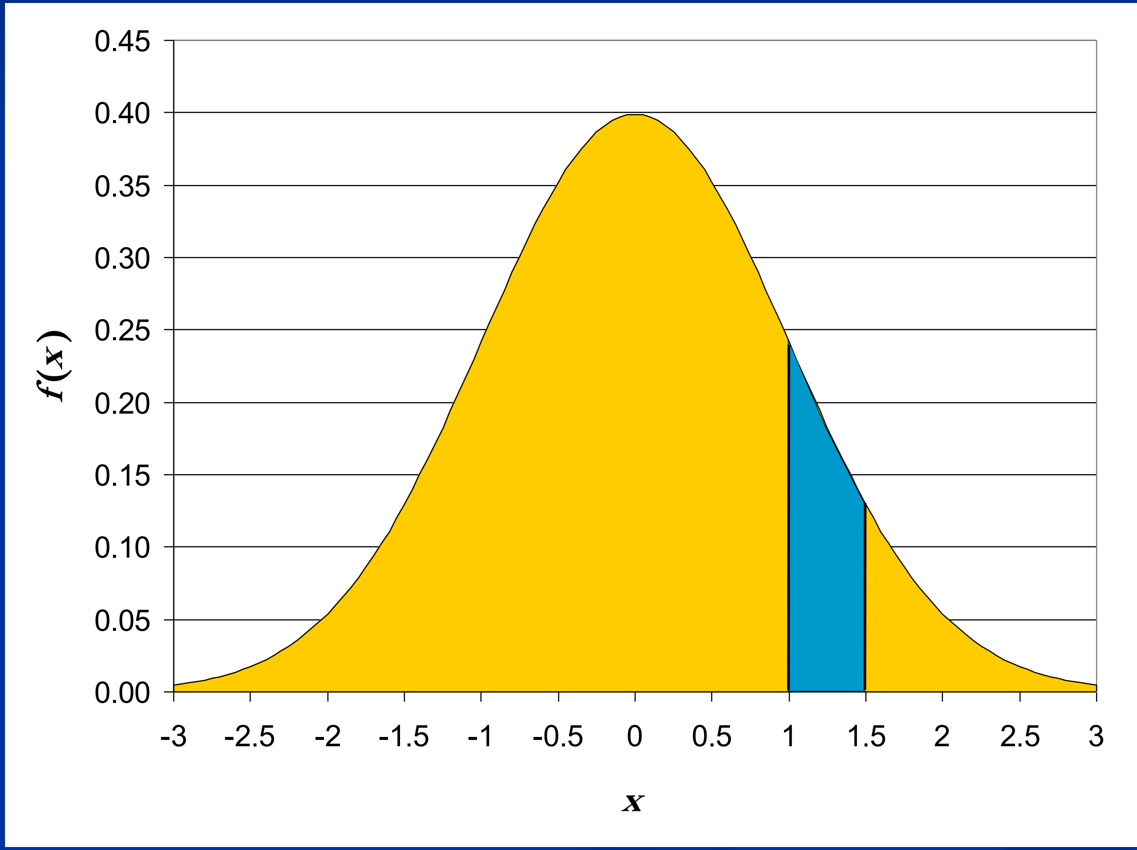
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

- $f(x)$ is not the same as $P(x)$
  - $P(x)$ would be 0 for every x because the normal distribution is continuous
  - However, $P(x_1 < X \leq x_2) = \int_{x_1}^{x_2} f(x)dx$

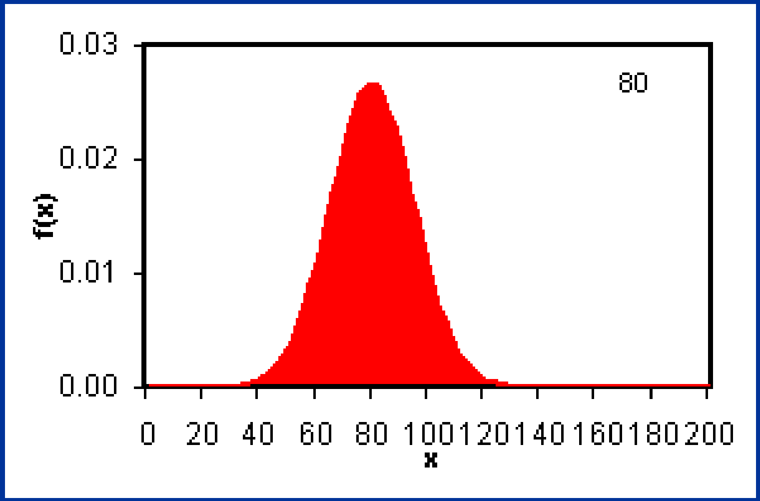
The Normal Distribution
Overview



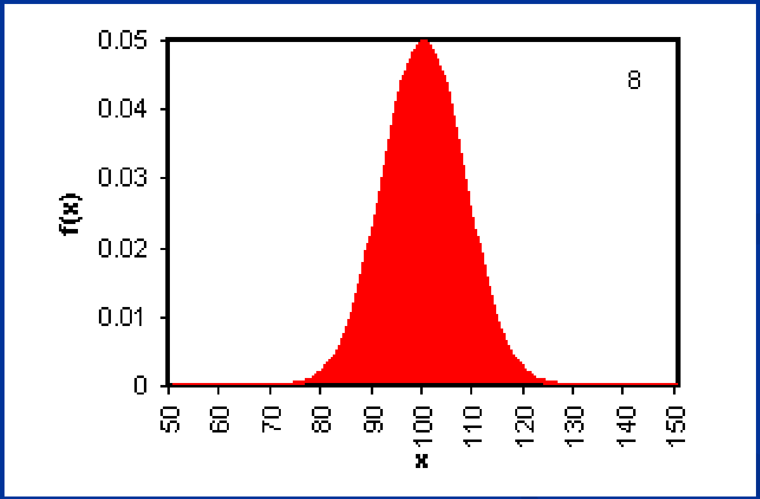
The Normal Distribution
Overview



The Normal Distribution
Overview



Mean changes



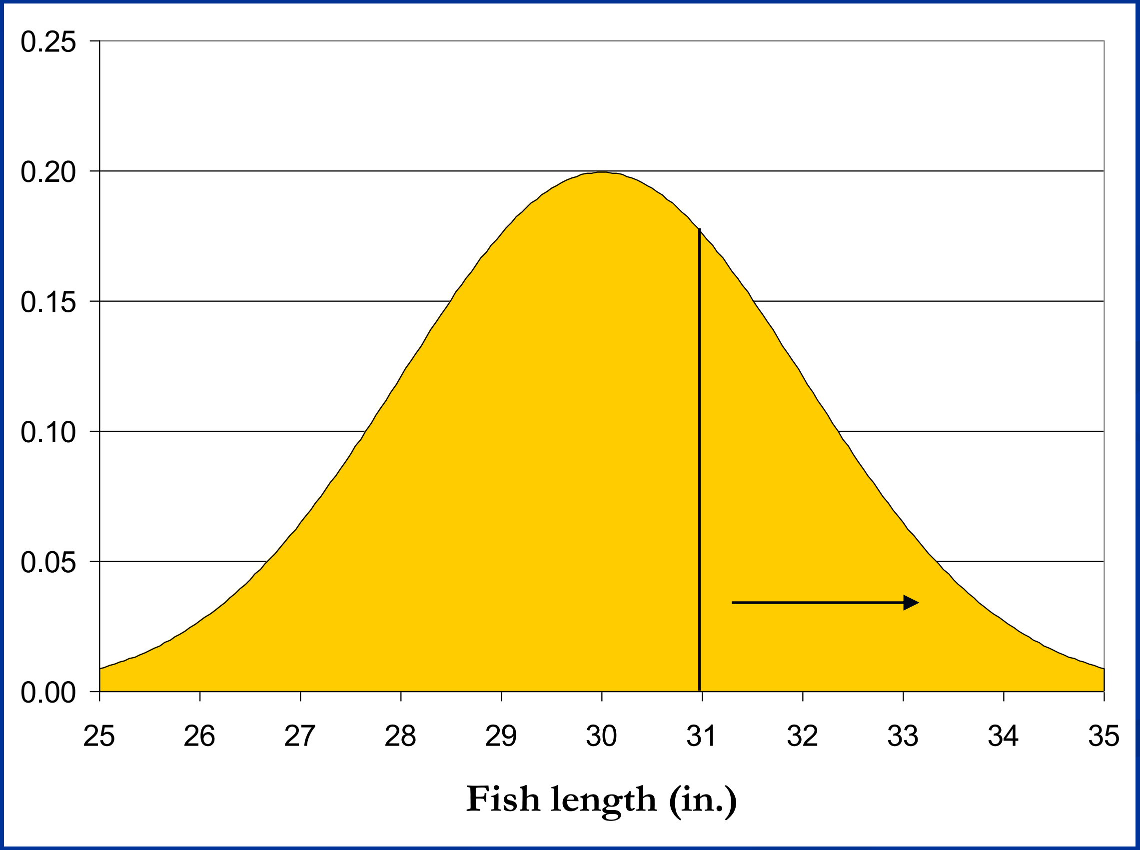
Variance changes

The Normal Distribution

- A sample of rock cod in Monterey Bay suggests that the mean length of these fish is $\mu = 30$ in. and $\sigma^2 = 4$ in.
- Assume that the length of rock cod is a normal random variable
- If we catch one of these fish in Monterey Bay,
  - What is the probability that it will be at least 31 in. long?
  - That it will be no more than 32 in. long?
  - That its length will be between 26 and 29 inches?

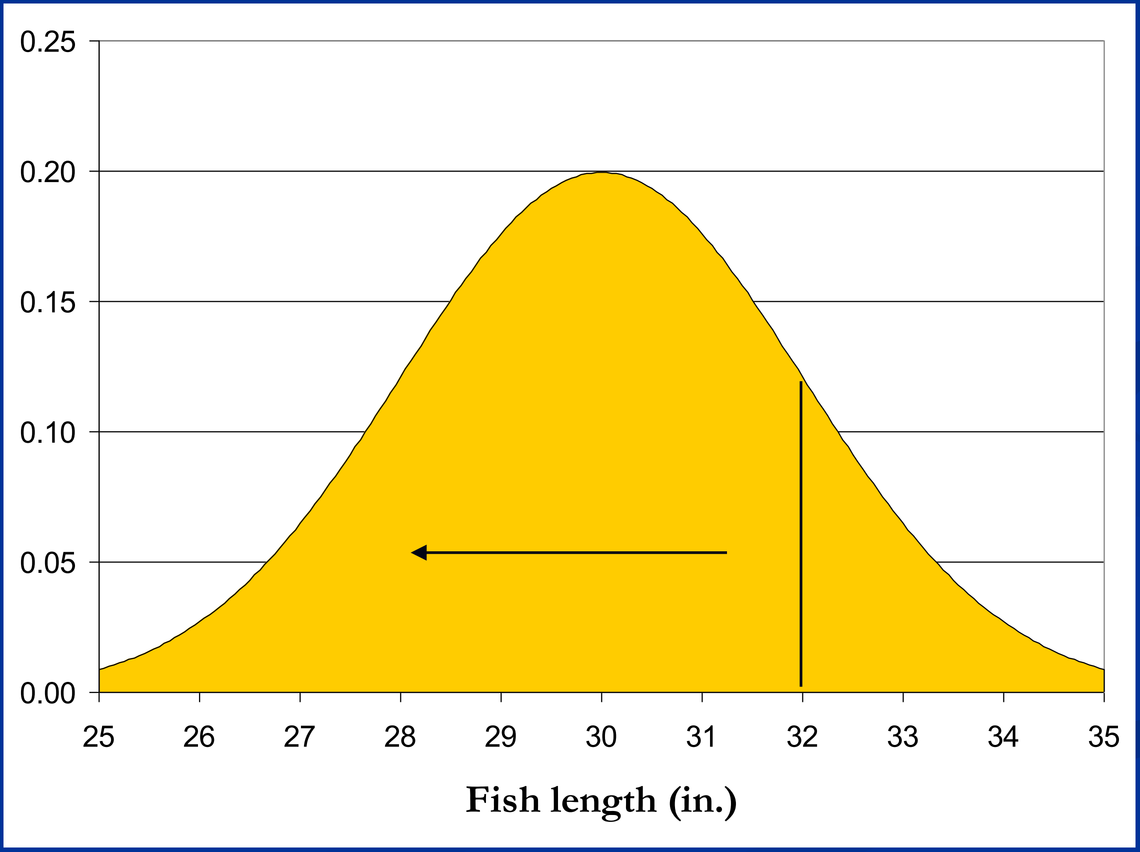
The Normal Distribution

- What is the probability that it will be at least 31 in. long?



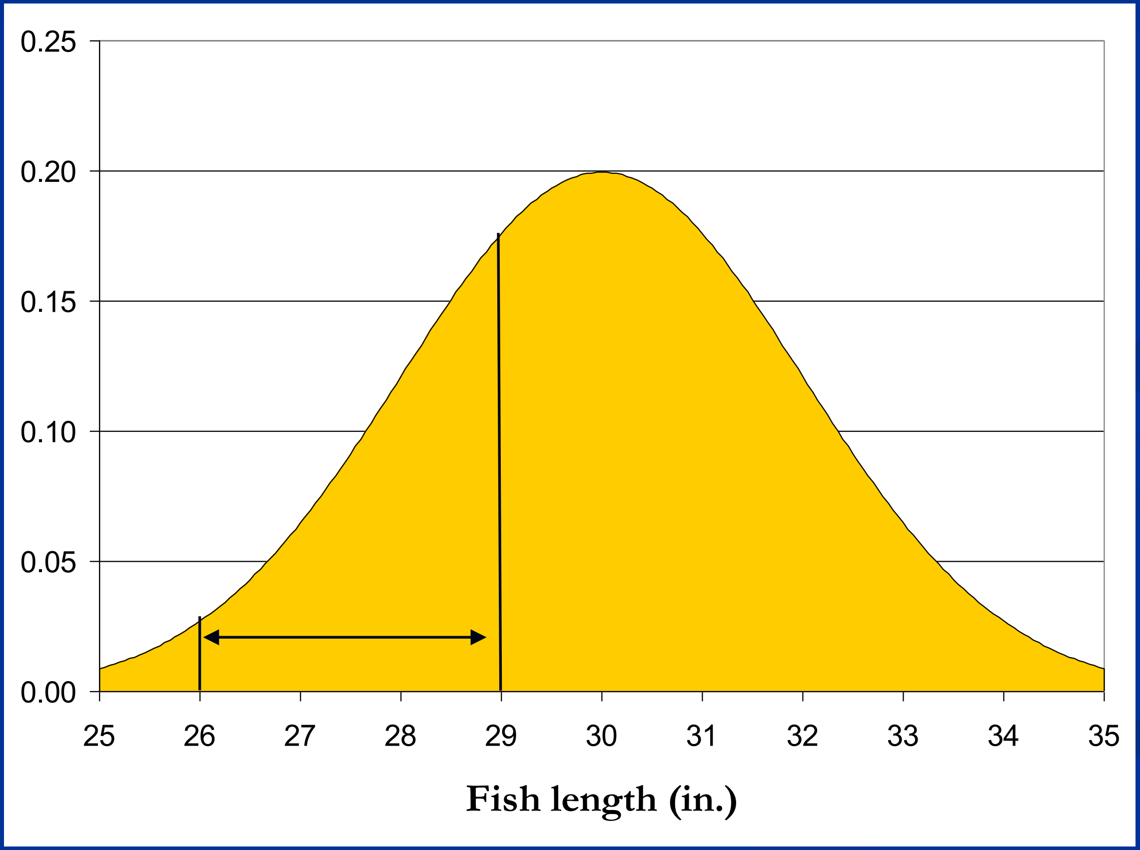
The Normal Distribution

- That it will be no more than 32 in. long?



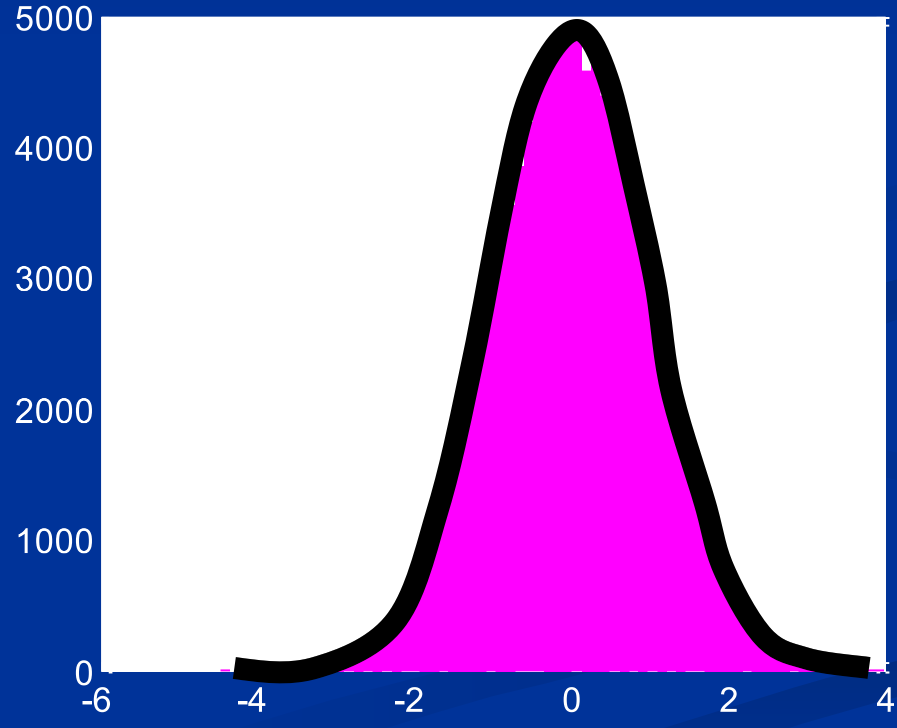
The Normal Distribution

- That its length will be between 26 and 29 inches?



Standard Normal Distribution

- $\mu=0$ and $\sigma^2=1$



Useful properties of the normal distribution

1. The normal distribution has useful properties:

- Can be added $E(X+Y) = E(X)+E(Y)$ and $\sigma 2(X+Y) = \sigma 2(X) + \sigma 2(Y)$
- Can be transformed with ***shift*** and ***change of scale*** operations

Consider two random variables X and Y

Let X~N(μ,σ) and let Y=aX+b where a and b area constants

Change of scale is the operation of multiplying *X* by a constant "***a***" because one unit of X becomes "a" units of Y.

Shift is the operation of adding a constant "*b*" to X because we simply move our random variable X "b" units along the x-axis.

If X is a normal random variable, then the new random variable Y created by this operations on X is also a random normal variable

For $X \sim N(\mu,\sigma)$ and $Y=aX+b$

- $E(Y) = a\mu+b$
- $\sigma^2(Y)=a^2 \sigma^2$
- A special case of a change of scale and shift operation in which $a = 1/\sigma$ and $b =-1(\mu/\sigma)$
- $Y=(1/\sigma)X-\mu/\sigma$
- $Y=(X-\mu)/\sigma$ gives
- $E(Y)=0$ and $\sigma^2(Y) =1$

The Central Limit Theorem

- That Standardizing any random variable that itself is a sum or average of a set of independent random variables results in a new random variable that is nearly the same as a standard normal one.
- The only caveats are that the sample size must be large enough and that the observations themselves must be independent and all drawn from a distribution with common expectation and variance.