Continuous Probability Distributions

Normal Probability Distribution

Dr.Mamatha.H.R

Professor

Department of Computer Science and Engineering

PES University

Bangalore

Course material created using various Internet resources and text book

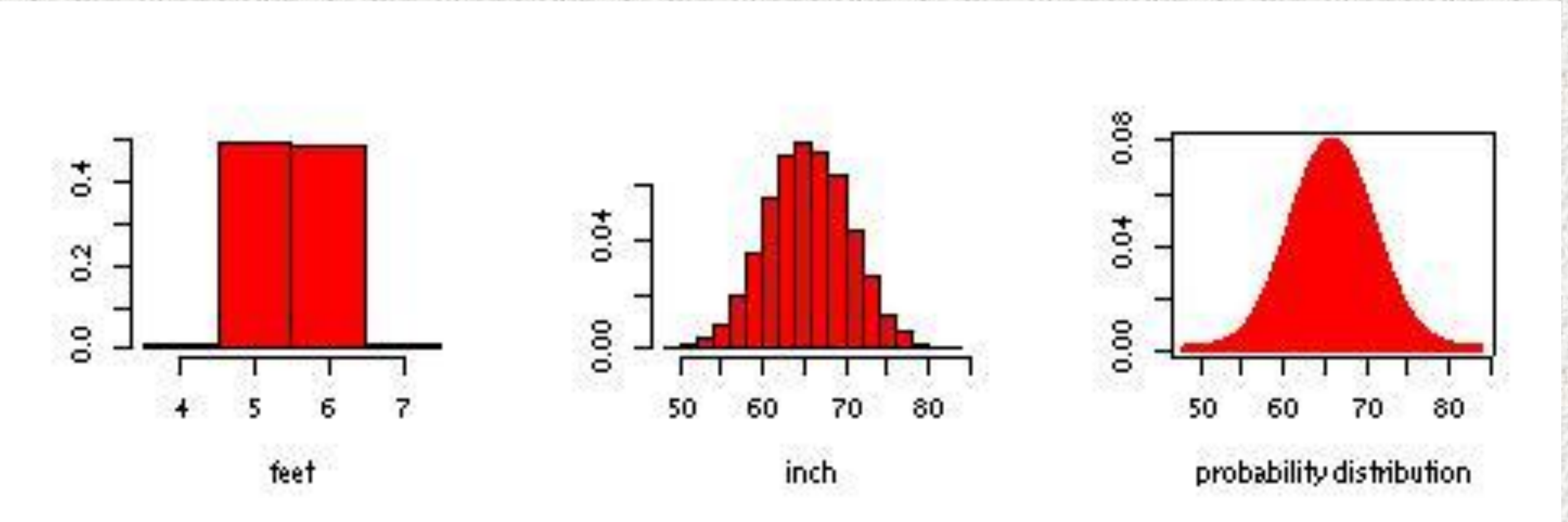
Continuous Random Variables

A random variable is continuous if it can assume the infinitely many values corresponding to points on a line interval.

- Examples:
  - **Heights, weights**
  - **length of life of a particular product**
  - **experimental laboratory error**

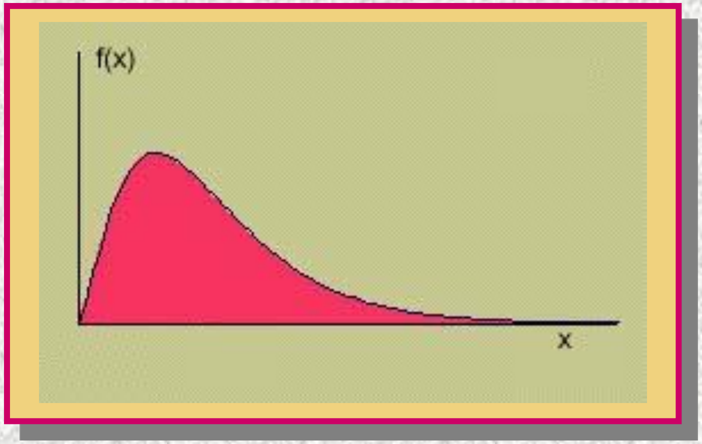
Continuous Probability Distribution

Suppose we measure height of students in this class. If we "discretize" by rounding to the nearest feet, the discrete probability histogram is shown on the left. Now if height is measured to the nearest inch, a possible probability histogram is shown in the middle. We get more bins and much smoother appearance. Imagine we continue in this way to measure height more and more finely, the resulting probability histograms approach a smooth curve shown on the right.



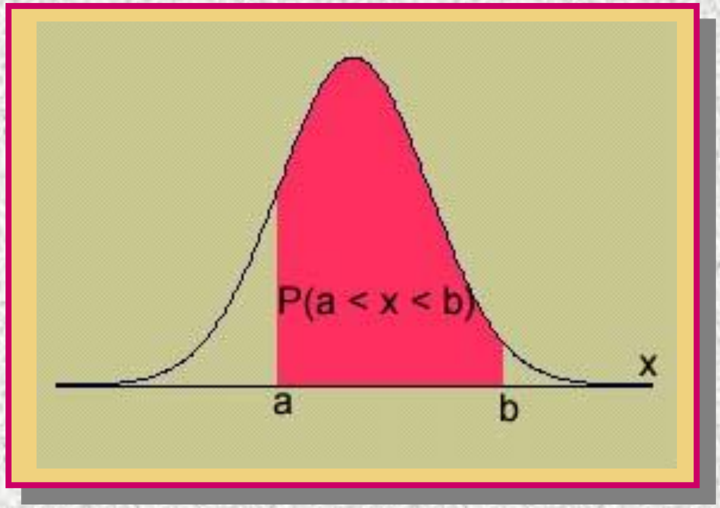
Probability Distribution for a Continuous Random Variable

Probability distribution **describes how the probabilities are distributed over all possible values. A** probability distribution **for a** continuous random variable x **is specified by a mathematical function denoted by f(x) which is called the** density function. **The graph of a density function is a smooth curve.**

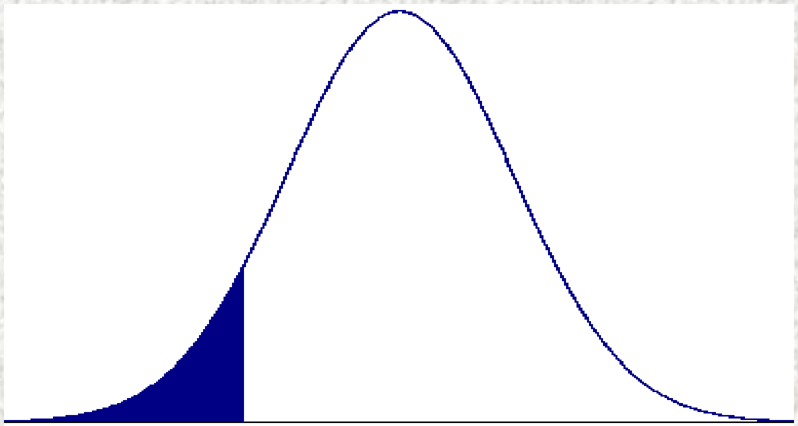


Properties of Continuous Probability Distributions

- **$f(x) \geq 0$**
- **The area under the curve is equal to 1.**
- **$P(a \leq x \leq b)$ = area under the curve between a and b.**

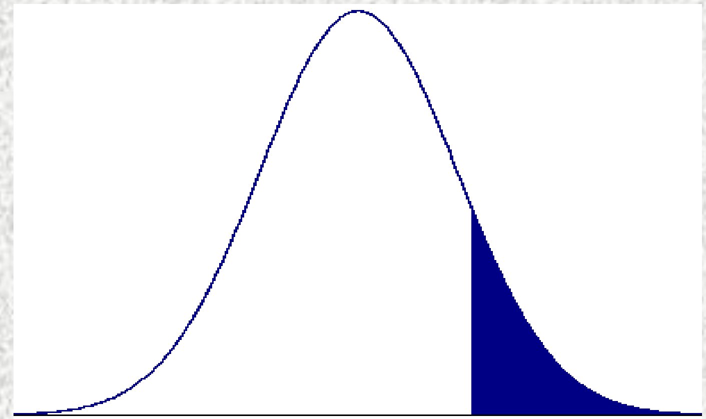


Some Illustrations



a

$P(x<a)$



b

$P(x>b)$

Notice that for a continuous random variable x,

$$P(x = a) = 0$$

for any specific value a because the "area above a point" under the curve is a line segment and hence has 0 area. Specifically this means

$$P(x<a) = P(x \le a)$$

$$P(a<x<b) = P(a\le x<b) = P(a<x\le b) = P(a \le x\le b)$$

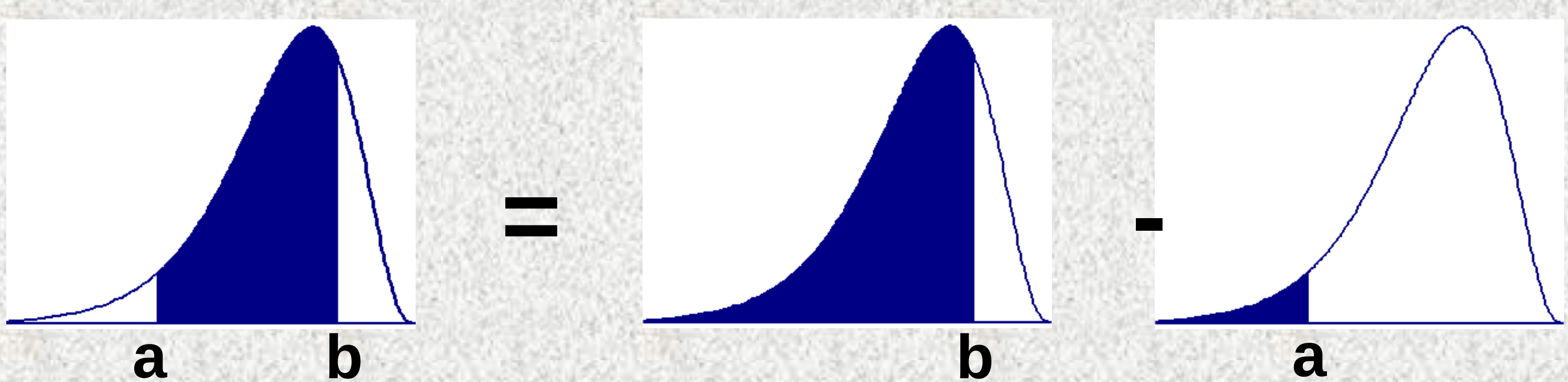
Method of Probability Calculation

The probability that a continuous random variable **x** lies between a lower limit **a** and an upper limit **b** is

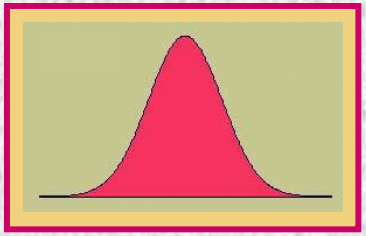
$P(a<x<b)$ = (cumulative area to the left of b) –

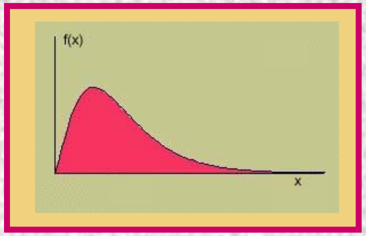
(cumulative area to the left of a)

$= P(x < b) - P(x < a)$



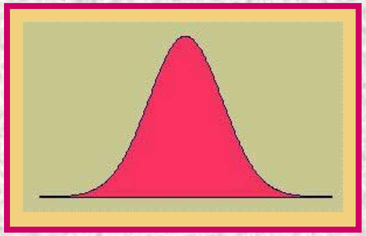
Continuous Probability Distributions





- There are many different types of continuous random variables
- We try to pick a model that
  - Fits the data well
  - Allows us to make the best possible inferences using the data.
- One important continuous random variable is the normal random variable.

The Normal Distribution



The formula that generates the normal probability distribution is:

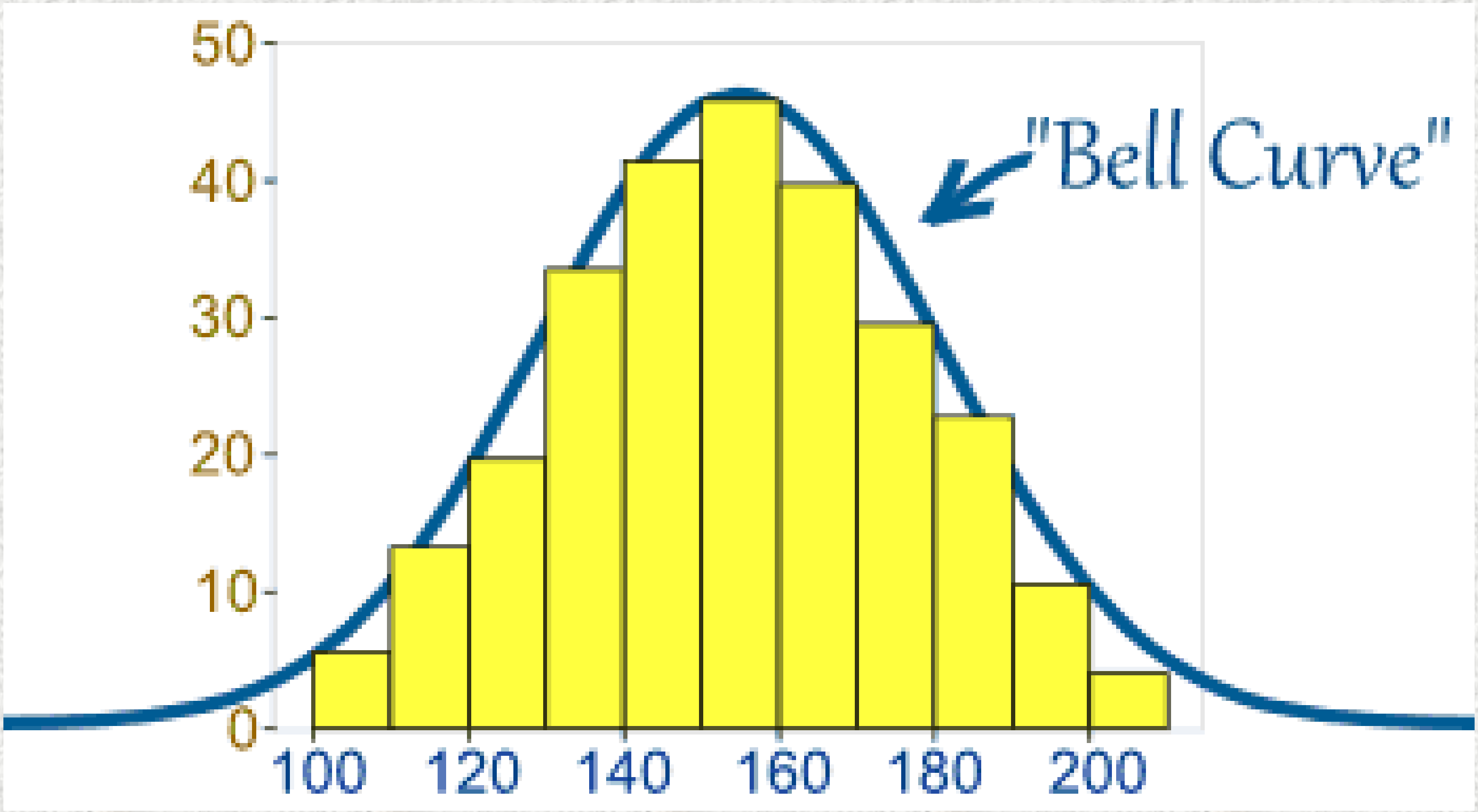
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad \text{for } -\infty < x < \infty$$

$e = 2.7183 \qquad \pi = 3.1416$

μ and σ are the population mean and standard deviation.

Two parameters, mean and standard deviation, completely determine the Normal distribution. The shape and location of the normal curve changes as the mean and standard deviation change.

A Bell Curve

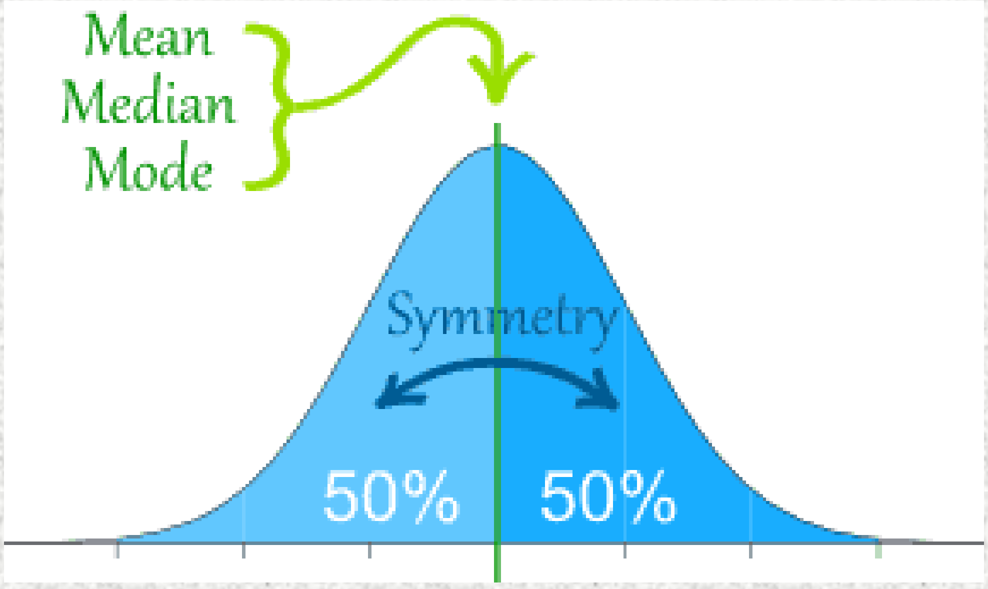


What are some examples of things that follow a Normal Distribution?

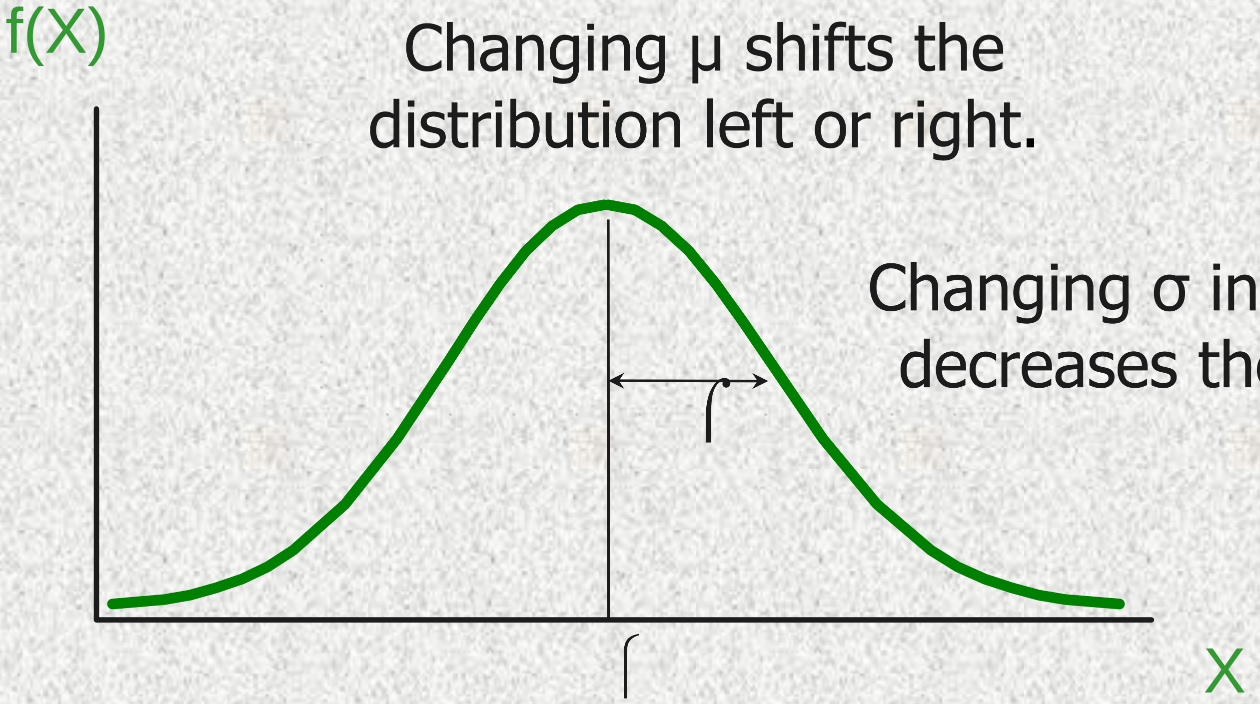
- Heights of people
- Size of things produced by machines
- Errors in measurements
- Blood Pressure
- Test Scores

Normal Distribution Curve

- mean=median=mode
- Symmetry about the center
- 50% of the values less than the mean and 50% greater than the mean



The Normal Distribution



Changing μ shifts the distribution left or right.

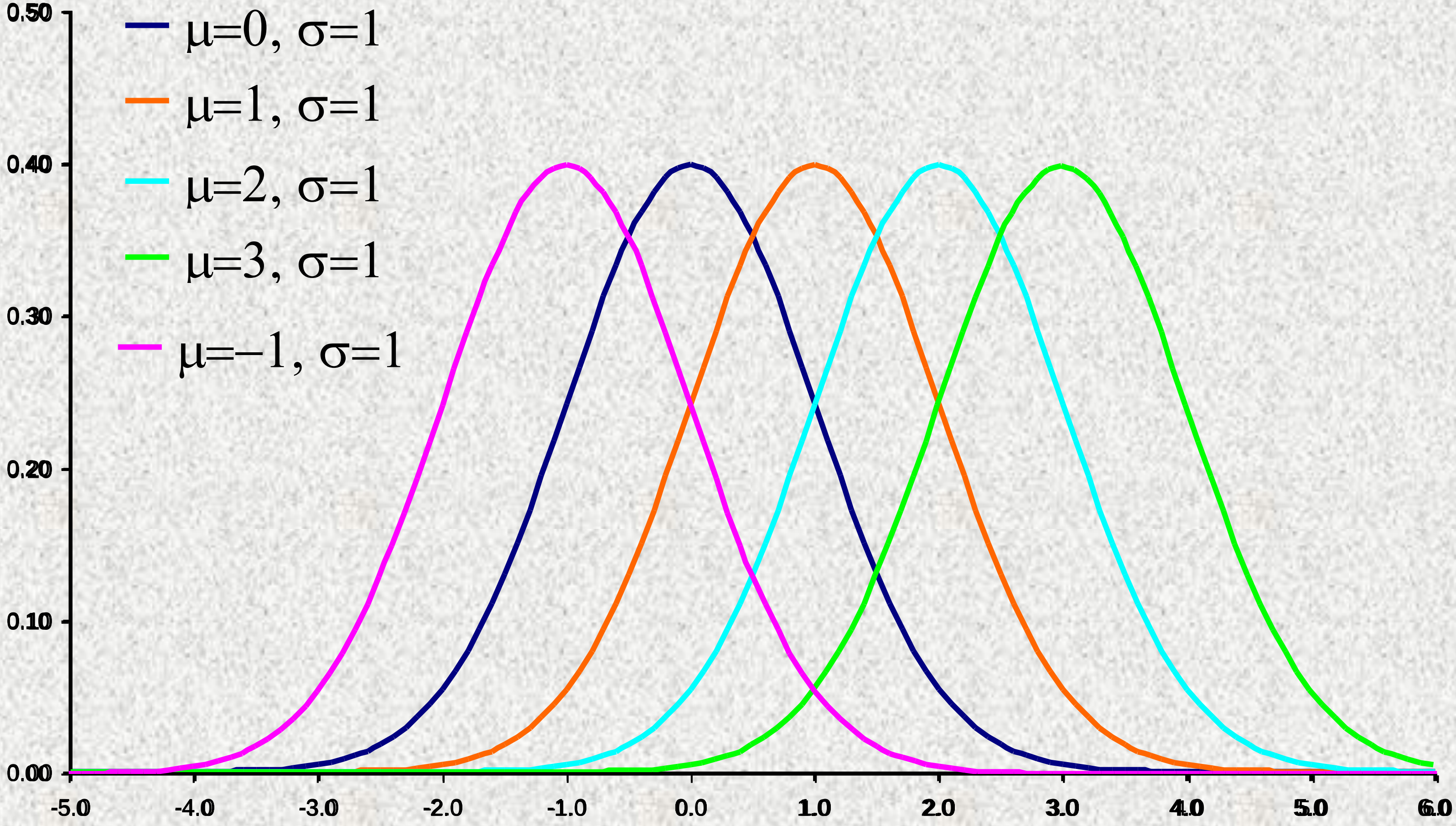
Changing σ increases or decreases the spread.

The Normal Distribution: as mathematical function (pdf)

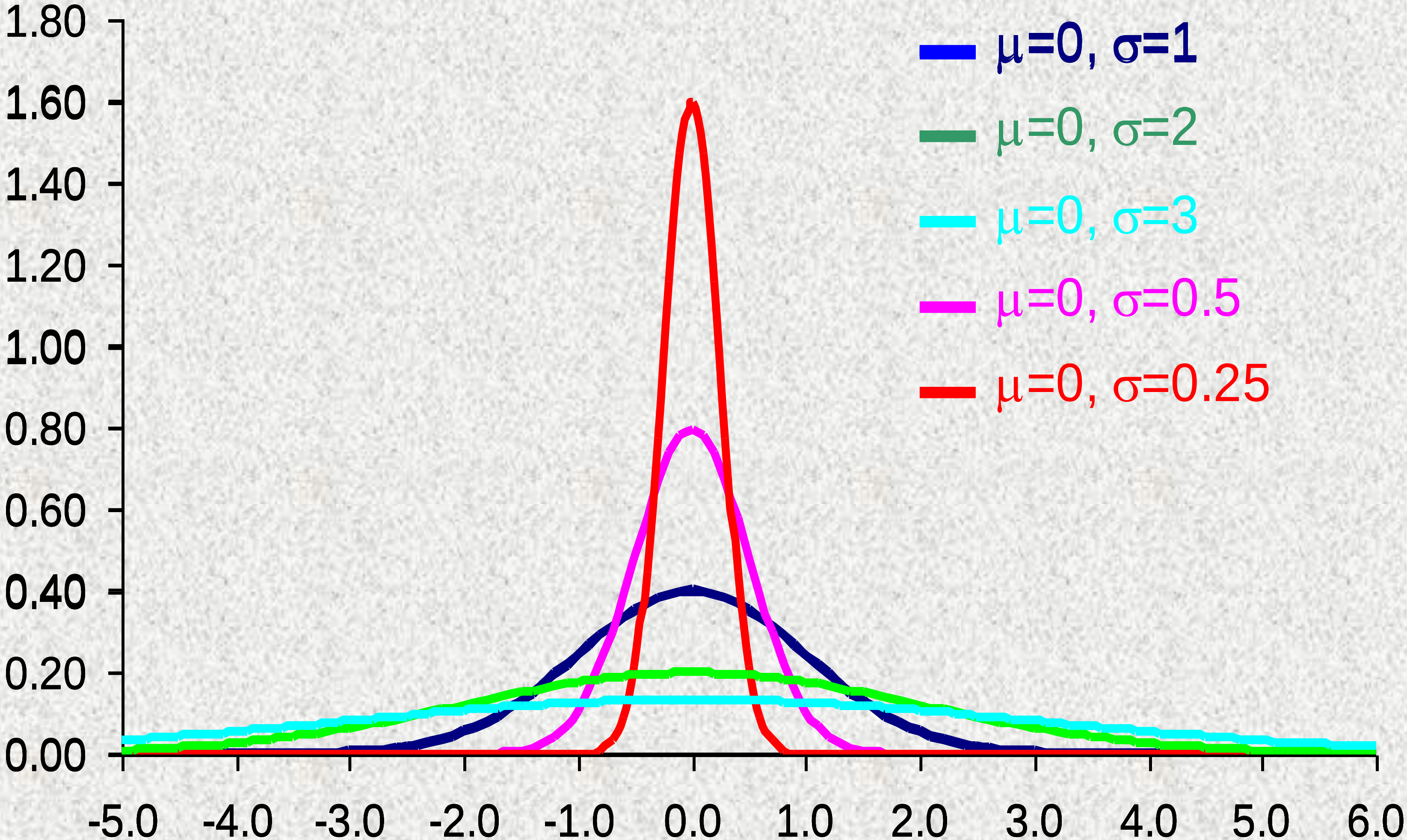
$$f(x) = \frac{1}{\boxed{\sigma}\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-\boxed{\mu}}{\sigma}\right)^2}$$

This is a bell shaped curve with different centers and spreads depending on μ and σ

Normal Distributions: σ=1



Normal Distributions: $\mu=0$



The Normal PDF

It's a probability function, so no matter what the values of μ and σ, must integrate to 1!

$$\int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = 1$$

Normal distribution is defined by its mean and standard dev.

$E(X)=\mu = \int_{-\infty}^{+\infty} x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$

$Var(X)=\sigma^2 = \left(\int_{-\infty}^{+\infty} x^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx\right) - \mu^2$

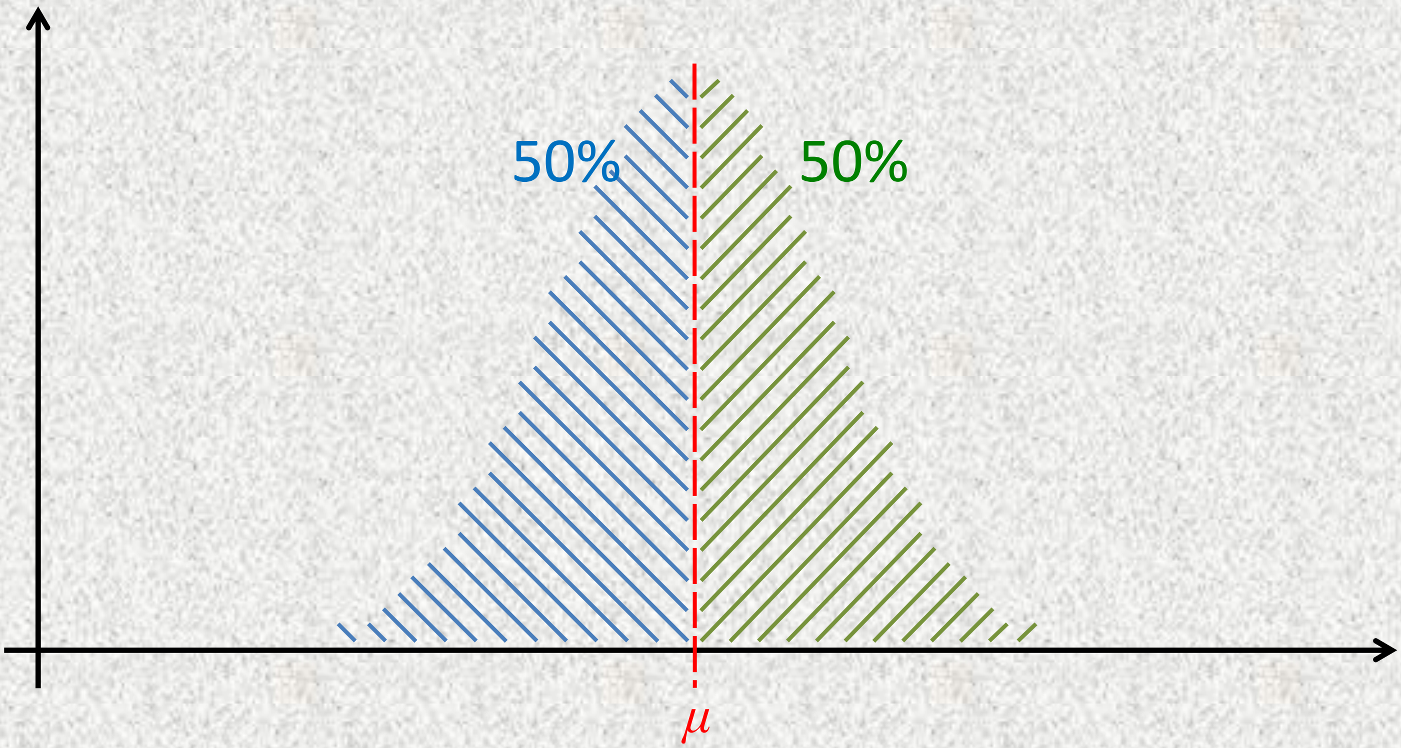
Standard Deviation(X)=σ

Why do we need to know Standard Deviation?

- Any value is
  - likely to be within 1 standard deviation of the mean
  - very likely to be within 2 standard deviations
  - almost certainly within 3 standard deviations

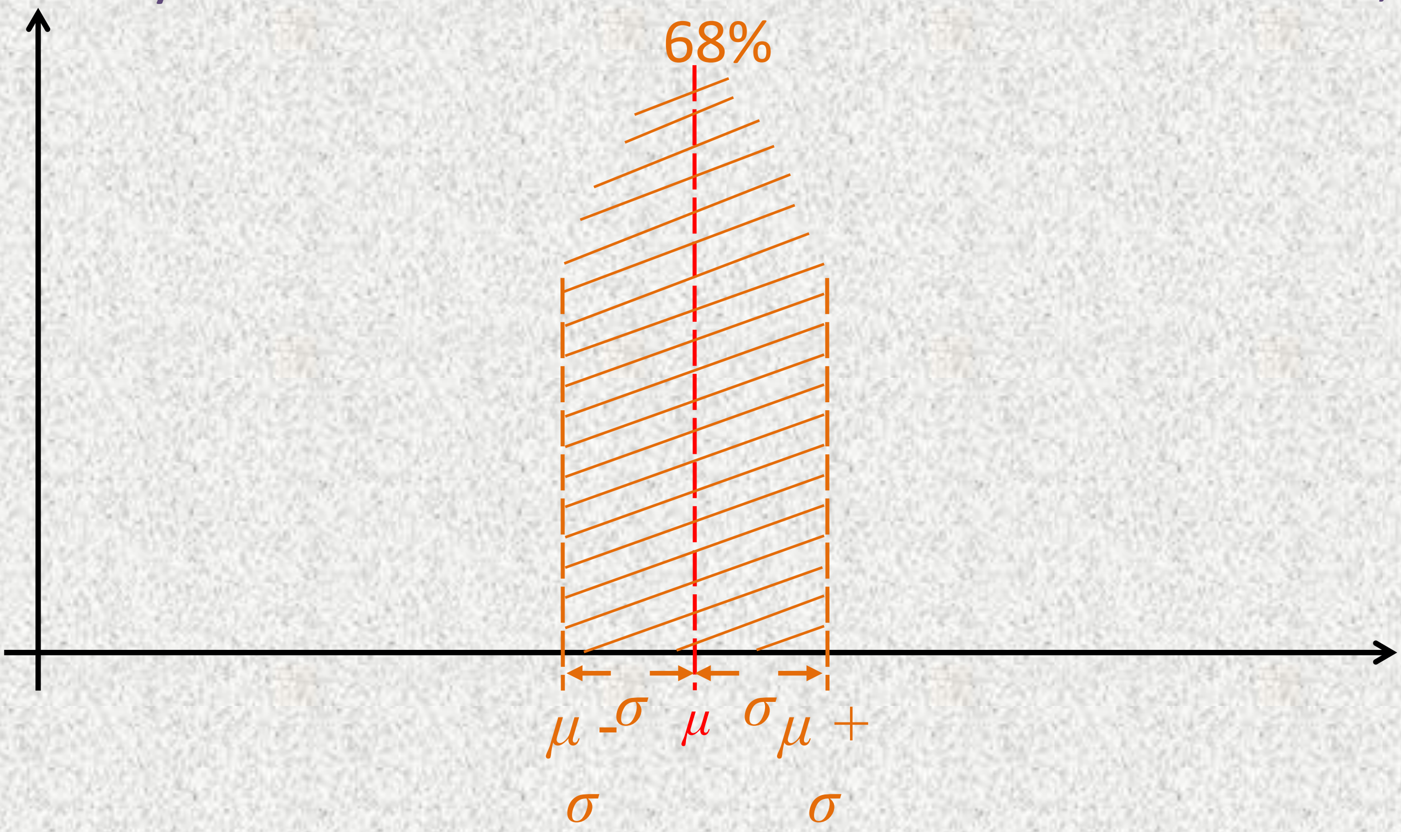
The properties of a normal distribution:

- It is a bell-shaped curve.
- It is symmetrical about the mean, μ.
- (The mean, the mode and the median all have the same value).
- The total area under the curve is 1 (or 100%).
- 50% of the area is to the left of the mean, and 50% to the right.



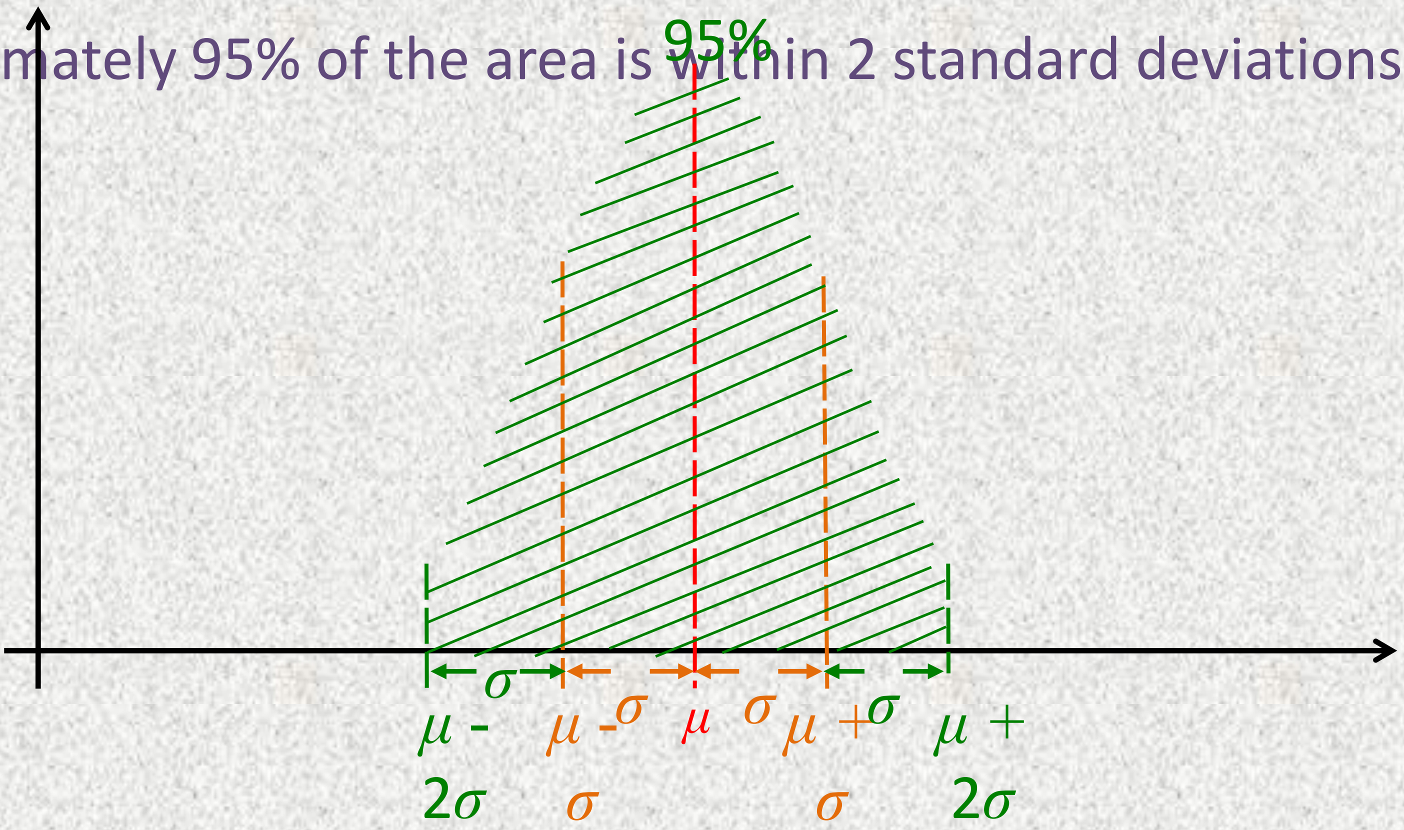
The properties of a normal distribution:

- It is a bell-shaped curve.
- It is symmetrical about the mean, μ.
- (The mean, the mode and the median all have the same value).
- The total area under the curve is 1 (or 100%).
- 50% of the area is to the left of the mean, and 50% to the right.
- Approximately 68% of the area is within 1 standard deviation, σ, of the mean.



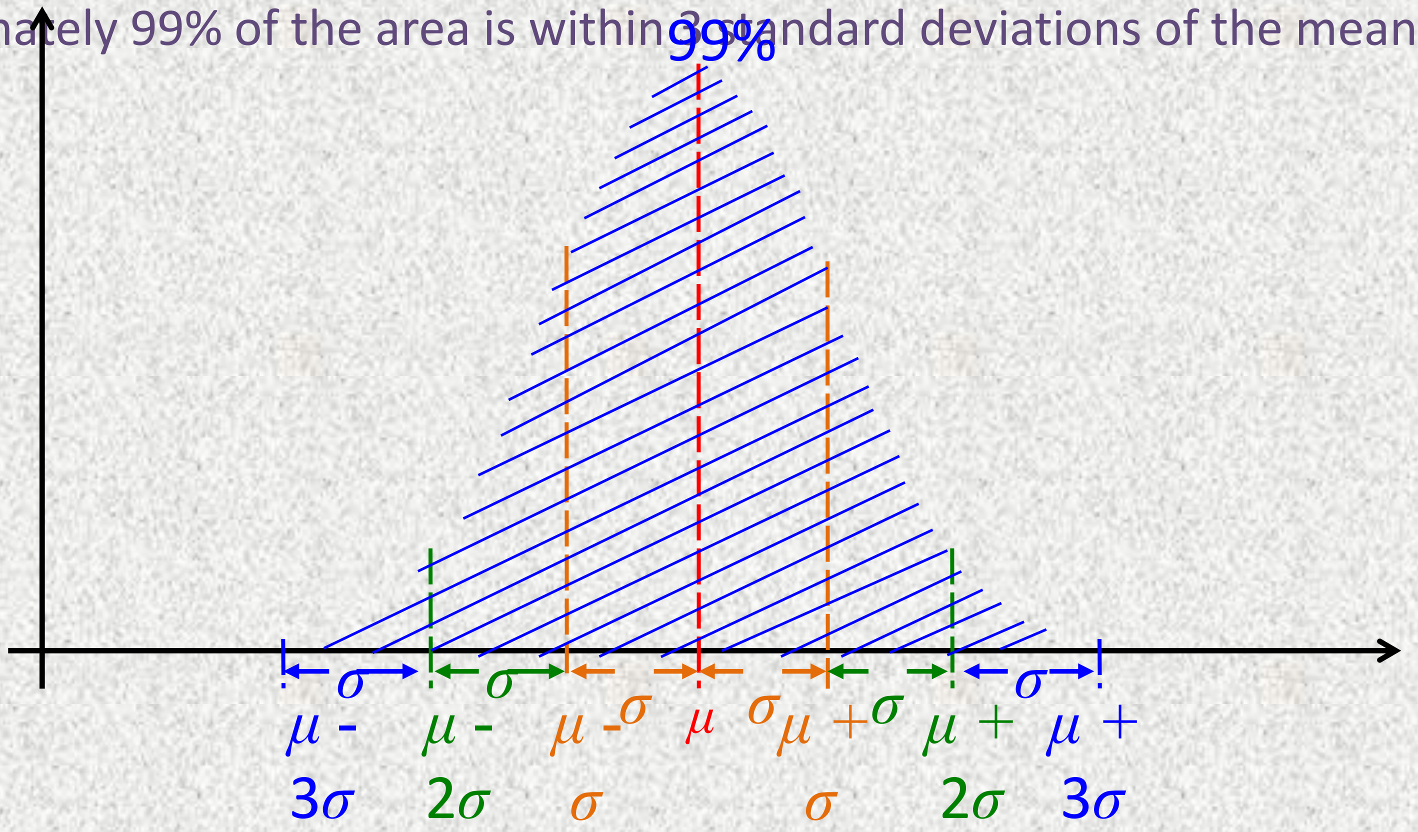
The properties of a normal distribution:

- It is a bell-shaped curve.
- It is symmetrical about the mean, μ.
- (The mean, the mode and the median all have the same value).
- The total area under the curve is 1 (or 100%).
- 50% of the area is to the left of the mean, and 50% to the right.
- Approximately 68% of the area is within 1 standard deviation, σ, of the mean.
- Approximately 95% of the area is within 2 standard deviations of the mean.



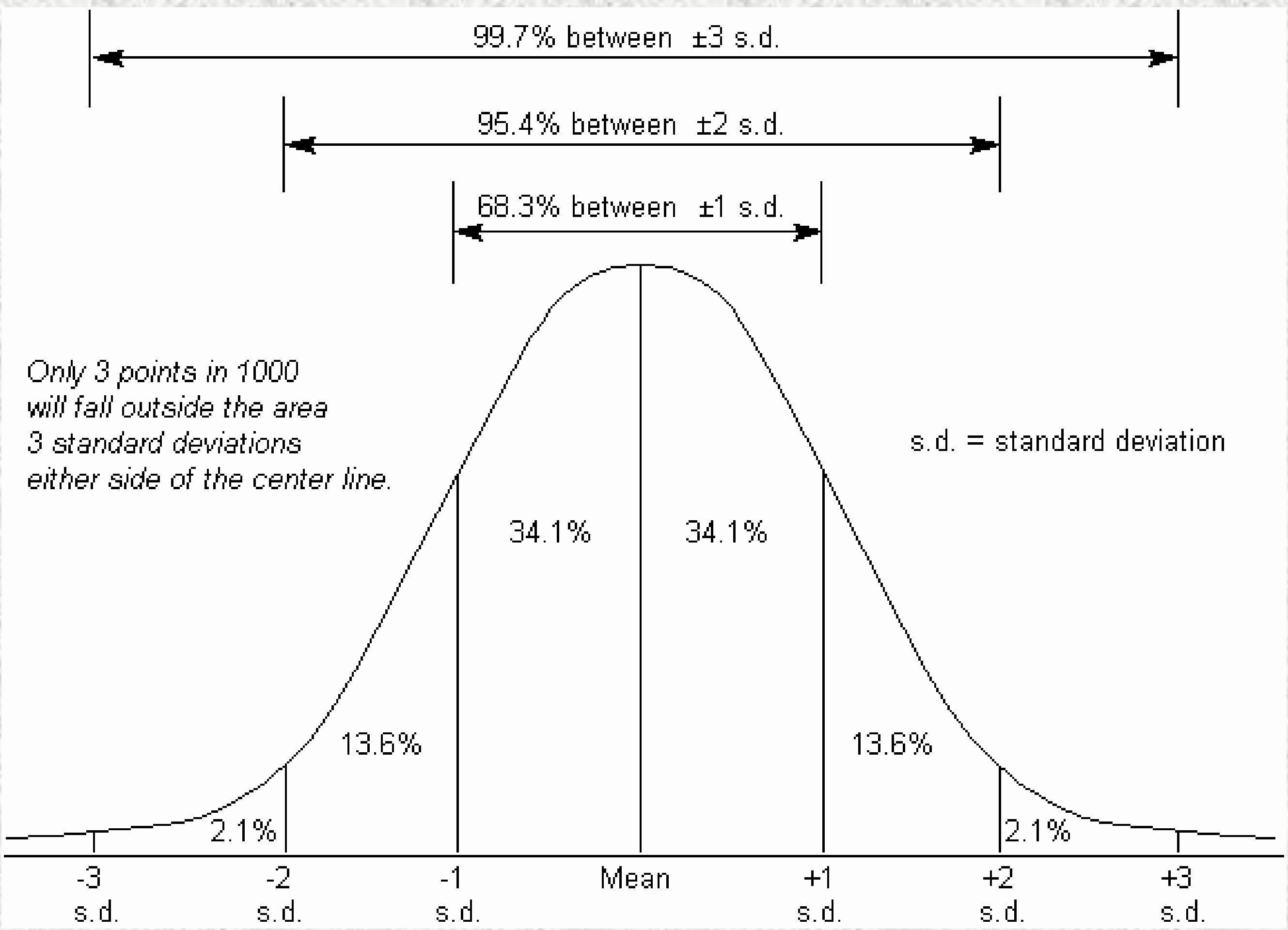
The properties of a normal distribution:

- It is a bell-shaped curve.
- It is symmetrical about the mean, μ.
- (The mean, the mode and the median all have the same value).
- The total area under the curve is 1 (or 100%).
- 50% of the area is to the left of the mean, and 50% to the right.
- Approximately 68% of the area is within 1 standard deviation, σ, of the mean.
- Approximately 95% of the area is within 2 standard deviations of the mean.
- Approximately 99% of the area is within 3 standard deviations of the mean.

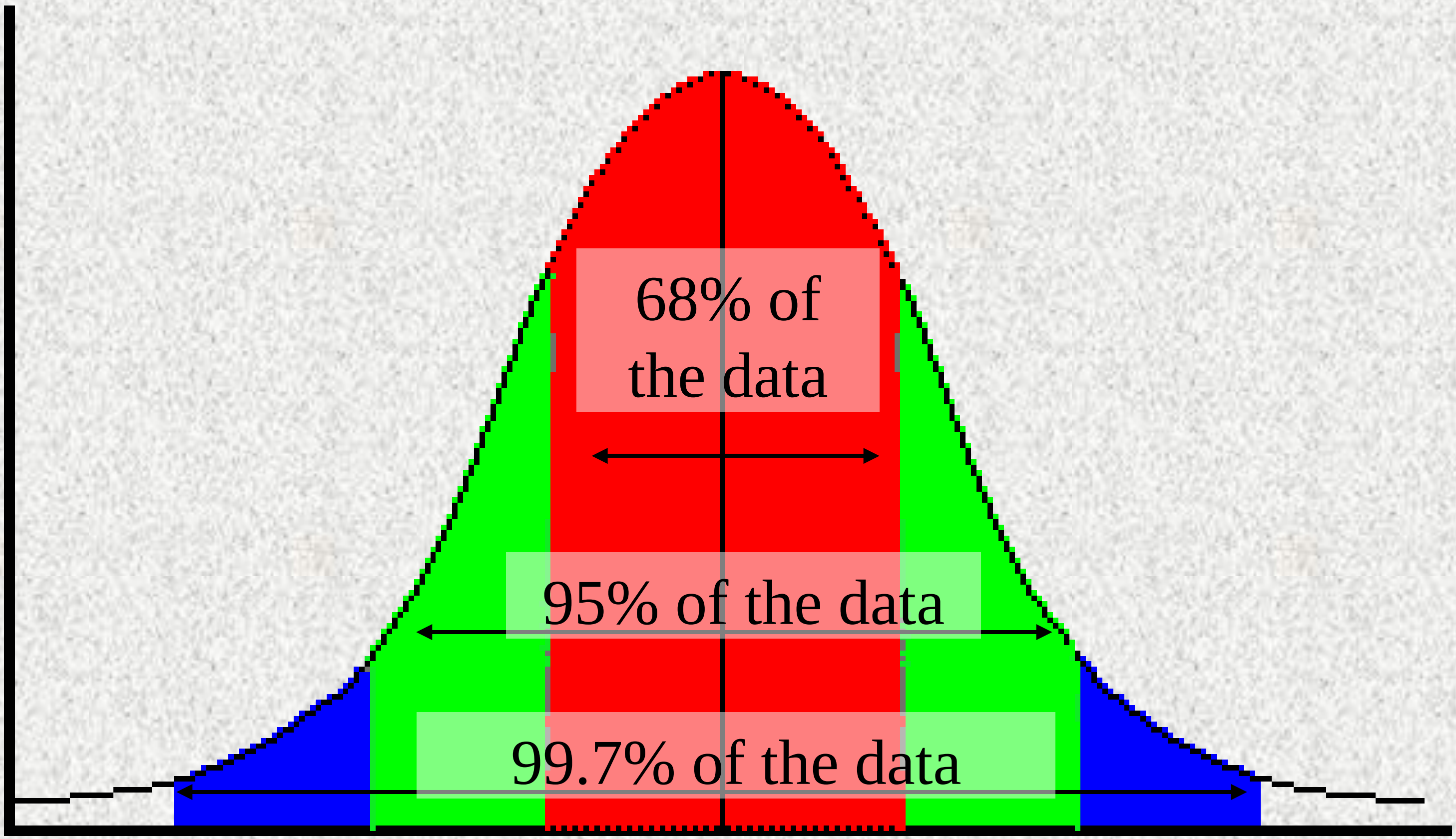


**The beauty of the normal curve:

No matter what μ and σ are, the area between $\mu-\sigma$ and $\mu+\sigma$ is about 68%; the area between $\mu-2\sigma$ and $\mu+2\sigma$ is about 95%; and the area between $\mu-3\sigma$ and $\mu+3\sigma$ is about 99.7%. Almost all values fall within 3 standard deviations.


99.7% between ±3 s.d.
95.4% between ±2 s.d.
68.3% between ±1 s.d.
Only 3 points in 1000
will fall outside the area
3 standard deviations
either side of the center line.
s.d. = standard deviation
34.1%
34.1%
13.6%
13.6%
2.1%
2.1%
-3
s.d.
-2
s.d.
-1
s.d.
Mean
+1
s.d.
+2
s.d.
+3
s.d.

68-95-99.7 Rule



68-95-99.7 Rule in Math terms…

$$\int_{\mu-\sigma}^{\mu+\sigma} \frac{1}{\sigma\sqrt{2\pi}} \bullet e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}\, dx = .68$$

$$\int_{\mu-2\sigma}^{\mu+2\sigma} \frac{1}{\sigma\sqrt{2\pi}} \bullet e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}\, dx = .95$$

$$\int_{\mu-3\sigma}^{\mu+3\sigma} \frac{1}{\sigma\sqrt{2\pi}} \bullet e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}\, dx = .997$$

How good is rule for real data?

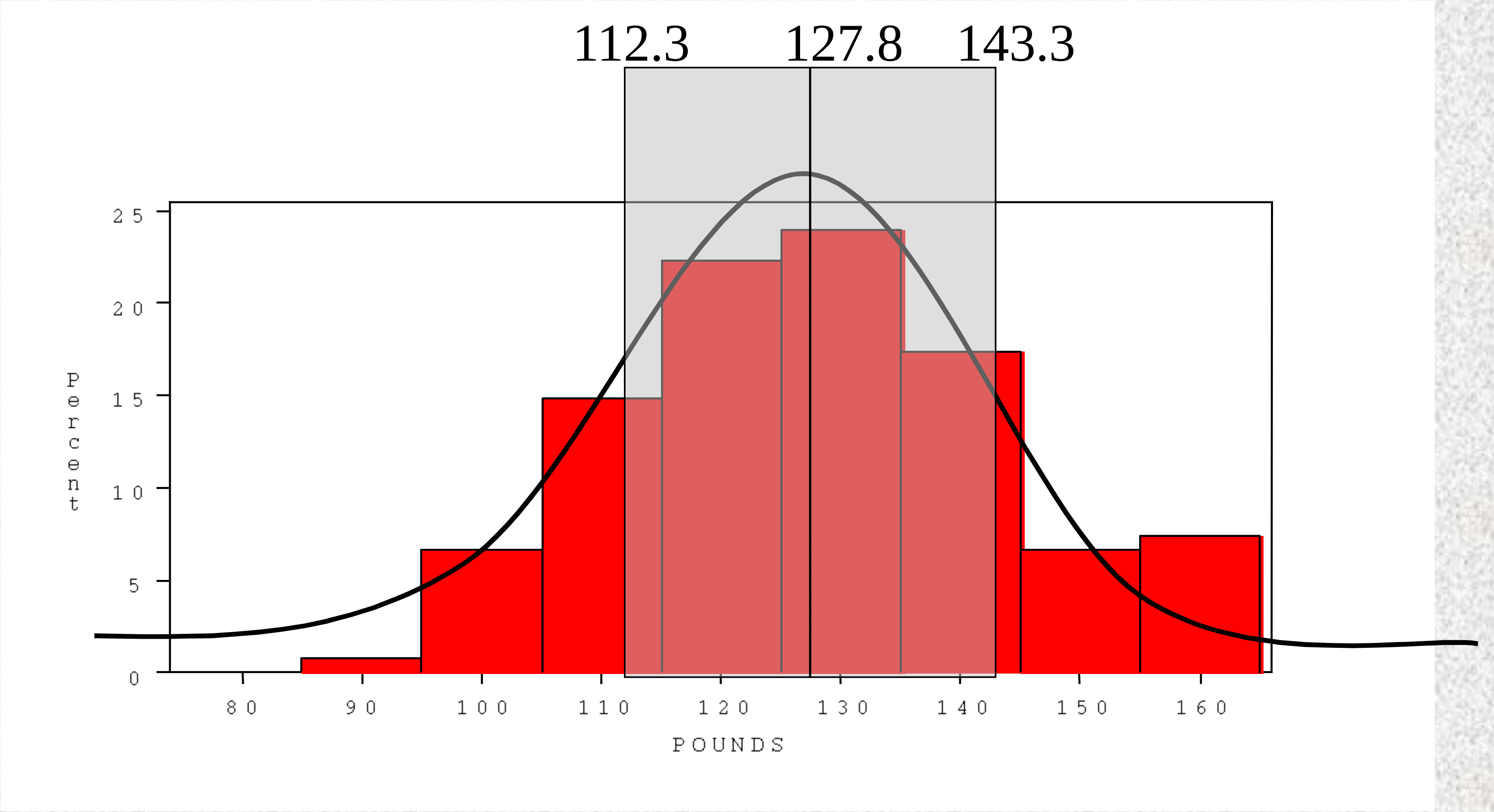
Check some example data:

The mean of the weight of the runners = 127.8

The standard deviation (SD) = 15.5

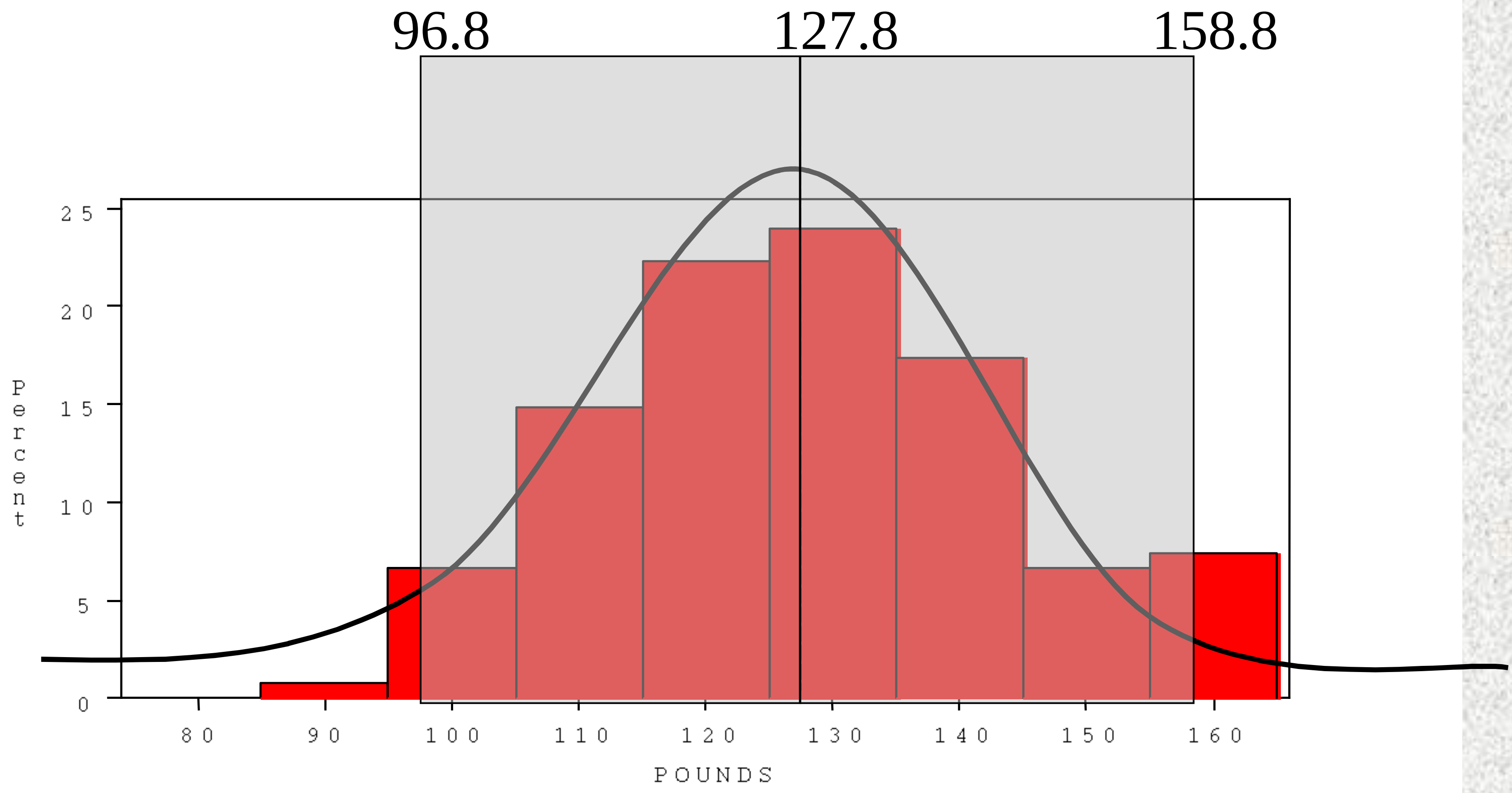
68% of 120 = .68x120 = ~ 82 runners

In fact, 79 runners fall within 1-SD (15.5 lbs) of the mean.



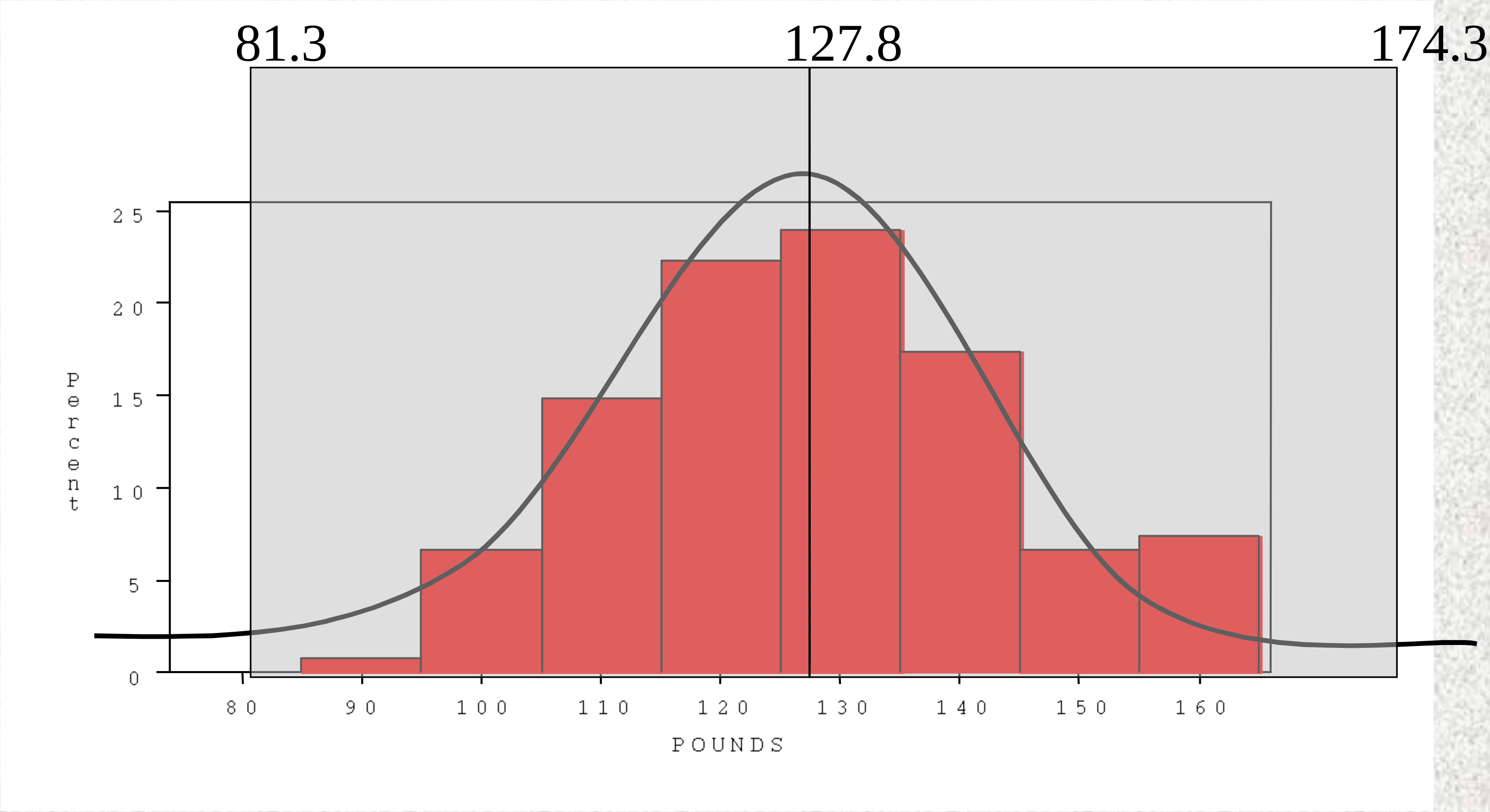
95% of 120 = .95 x 120 = ~ 114 runners

In fact, 115 runners fall within 2-SD's of the mean.



99.7% of 120 = .997 x 120 = 119.6 runners

In fact, all 120 runners fall within 3-SD's of the mean.



Example

- Suppose SAT scores roughly follows a normal distribution in the U.S. population of college-bound students (with range restricted to 200-800), and the average math SAT is 500 with a standard deviation of 50, then:
  - 68% of students will have scores between 450 and 550
  - 95% will be between 400 and 600
  - 99.7% will be between 350 and 650

Example

- **BUT…**
- **What if you wanted to know the math SAT score corresponding to the 90th percentile (=90% of students are lower)?**

P(X≤Q) = .90 →

$$\int_{200}^{Q} \frac{1}{(50)\sqrt{2\pi}} \bullet e^{-\frac{1}{2}\left(\frac{x-500}{50}\right)^2} dx = .90$$

Solve for Q?

The Standard Normal (Z): "Universal Currency"

The formula for the standardized normal probability density function is

$$p(Z) = \frac{1}{(1)\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{Z-0}{1}\right)^2} = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}(Z)^2}$$

The Standard Normal Distribution (Z)

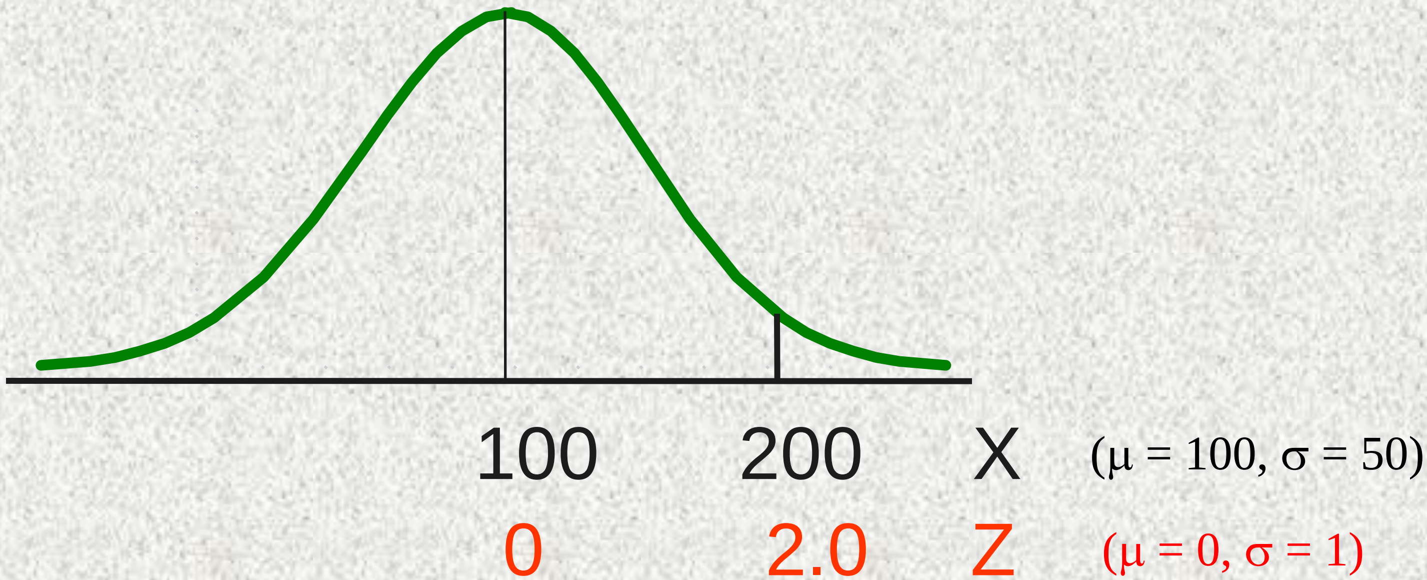
All normal distributions can be converted into the standard normal curve by subtracting the mean and dividing by the standard deviation:

$$\boxed{Z = \frac{X - \mu}{\sigma}}$$

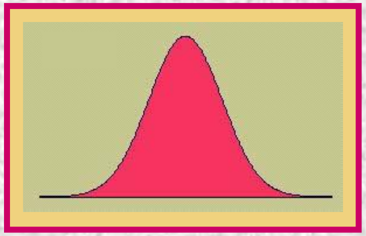
Somebody calculated all the integrals for the standard normal and put them in a table! So we never have to integrate!

Even better, computers now do all the integration.

Comparing X and Z units

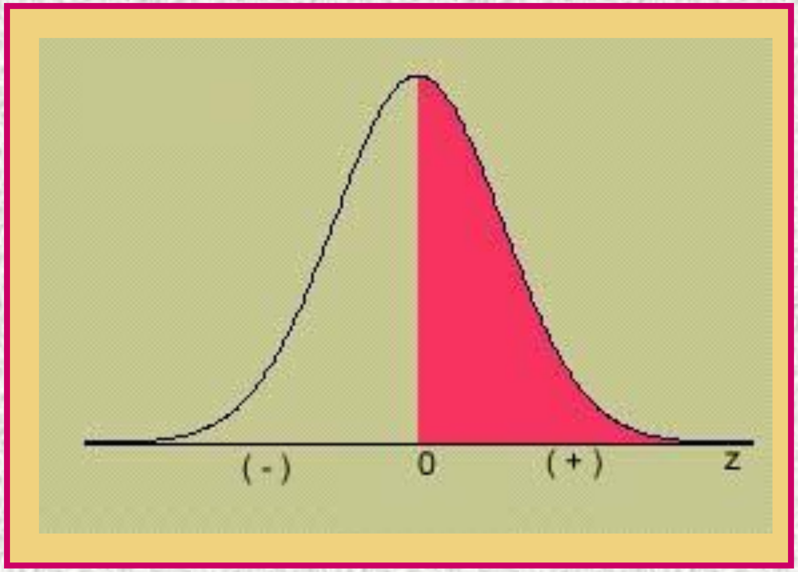


The Standard Normal Distribution



- **To find P(a < *x* < b), we need to find the area under the appropriate normal curve.**
- **To simplify the tabulation of these areas, we** standardize **each value of *x* by expressing it as a *z*-score, the number of standard deviations σ it lies from the mean μ.**

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



The Standard Normal (z) Distribution

- **Mean = 0; Standard deviation = 1**
- **When $x = \mu$, $z = 0$**
- **Symmetric about $z = 0$**
- **Values of z to the left of center are negative**
- **Values of z to the right of center are positive**
- **Total area under the curve is 1.**
- **Areas on both sides of center equal .5**

Example

- **For example: What's the probability of getting a math SAT score of 575 or less, μ=500 and σ=50?**

$$Z = \frac{575-500}{50} = 1.5$$

- i.e., A score of 575 is 1.5 standard deviations above the mean

$$\therefore P(X \leq 575) = \int_{200}^{575} \frac{1}{(50)\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-500}{50}\right)^2} dx \longrightarrow \int_{-\infty}^{1.5} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}Z^2} dz$$

But to look up Z= 1.5 in standard normal chart = .9332

Using Table

The four digit probability in a particular row and column of Table gives the area under the standard normal curve less than a value z. This is enough because the standard normal curve is symmetric.

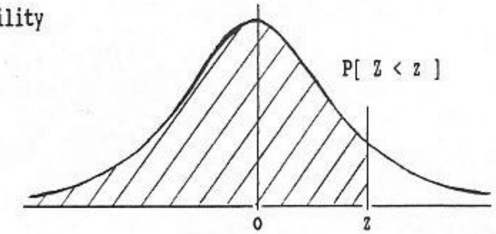
Looking up probabilities in the standard normal table

STANDARD STATISTICAL TABLES

1. Areas under the Normal Distribution

The table gives the cumulative probability up to the standardised normal value z
i.e.

$$P[\,Z < z\,] = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp(-\tfrac{1}{2}Z^2)\,dZ$$



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5159	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7854
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8804	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9773	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9865	0.9868	0.9871	0.9874	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9924	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964

Z=1.51

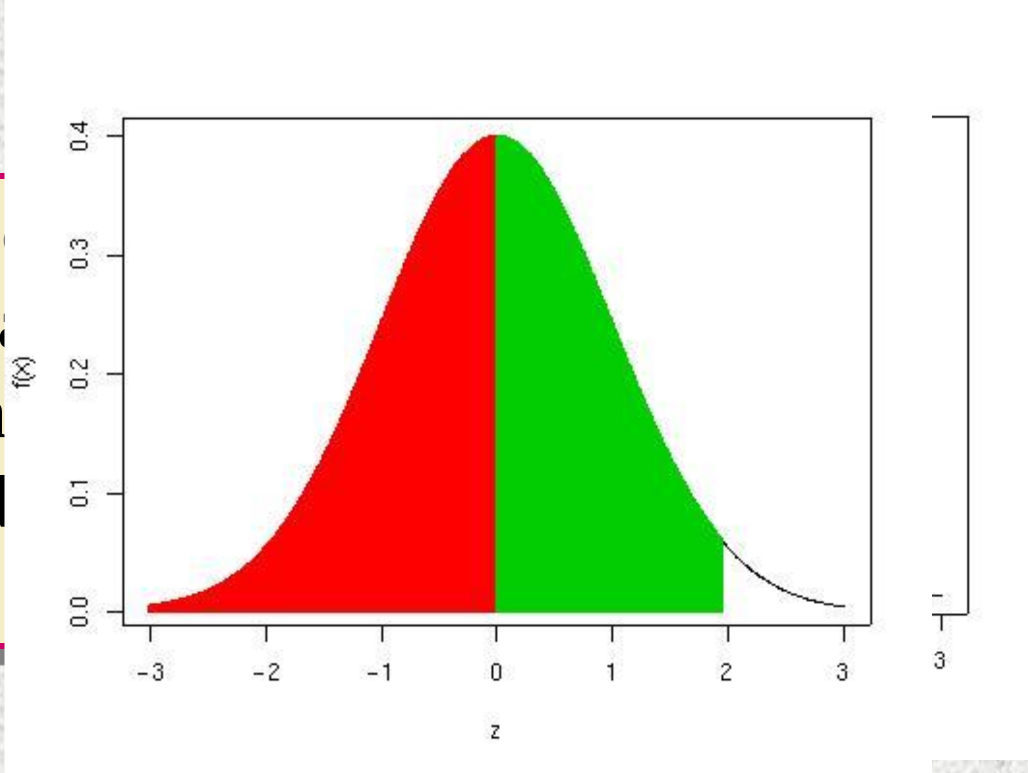
=1.51

What is the area to the left of Z=1.51 in a standard normal curve?

Area is 93.45%

Using Table

- ✓To find an area betwe
read directly from the ta
- ✓Use properties of stan
probability rules to find



$P(0<z<1.96) = P(z<1.96)-P(z<0)=.9750-0.5=.4750$

$P(-1.96<z<0)= P(0<z<1.96)=.4750$

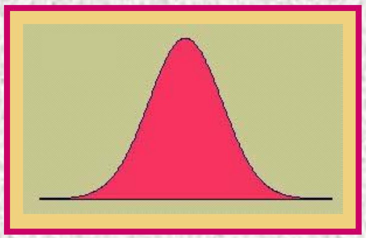
$P(z<1.96)=P(z<0)+ P(0<z<1.96)=.5+.4750=.9750$

$P(z<-1.96)=P(z>1.96) =.0250$

$P(-1.96<z<1.96)=P(z<1.96)-P(z<-1.96)$

$=.9750-.0250=.9500$

Working Backwards



Often we know the area and want to find the z-value that gives the area.

Example: Find the value of a positive *z* that has area .4750 between 0 and *z*.

1. **Add 0.5 to .4750.Look for the four digit area closest to .9750 in Table .**
2. **What row and column does this value correspond to?**
3. ***z* = 1.96**

Example

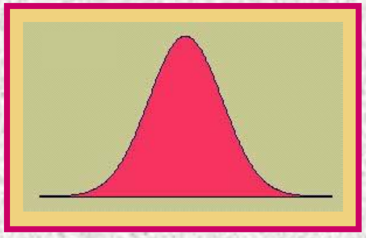
$P(z<?) = .75$

$z = .67$

What percentile does this value represent?

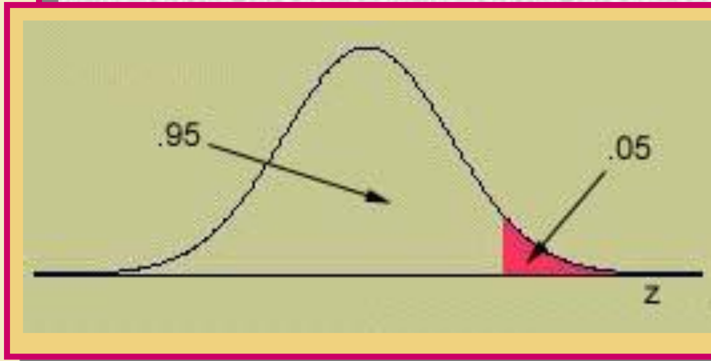
75th percentile, or the third quartile.

Working Backwards



Find the value of *z* that has area .05 to its right.

1. The area to its left will be 1 - .05 = .95
2. The area to its left will be .95
3. Look for the four digit area closest to .9500 in Table 3.
4. Since the value .9500 is halfway between .9495 and .9505, we choose *z* halfway between 1.64 and 1.65. $z=1.645$



Finding Probabilities for the General Normal Random Variable

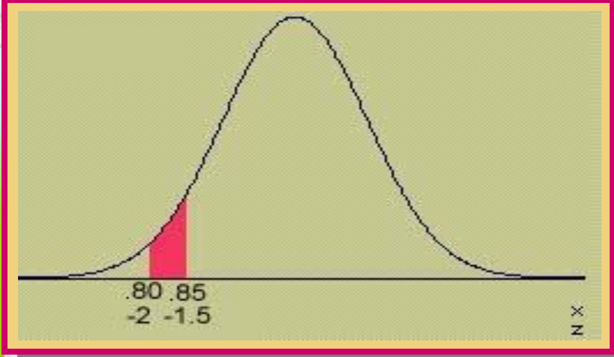
✓**To find an area for a normal random variable *x* with mean μ and standard deviation σ, *standardize or rescale* the interval in terms of *z*.**

✓**Find the appropriate area using Table .**

Example: ***x* has a normal distribution with mean = 5 and sd = 2. Find P(*x* > 7).**

$$P(x > 7) = P(z > \frac{7-5}{2}) = P(z > 1) = 1 - P(z < 1)$$

$$= 1 - P(z < 0) - P(0 < z < 1) = 1 - .5 - .3413 = .1587$$



Example

The weights of packages of ground nuts are normally distributed with mean 1 pound and standard deviation .10. What is the probability that a randomly selected package weighs between 0.80 and 0.85 pounds?

$$P(.80 < x < .85) = P\left(\frac{.80-1}{.1} < z < \frac{.85-1}{.1}\right)$$

$$= P(-2 < z < -1.5) = P(1.5 < z < 2)$$

$$= P(0 < z < 2) - P(0 < z < 1.5)$$

$$= .4772 - .4332 = .0440$$

Example

What is the weight of a package such that only 5% of all packages exceed this weight?

$$P(x > ?) = .05$$

$$P(z > \frac{? - 1}{.1}) = .05$$

$$P(0 < z < \frac{? - 1}{.1}) = .95 - .50 = .45$$

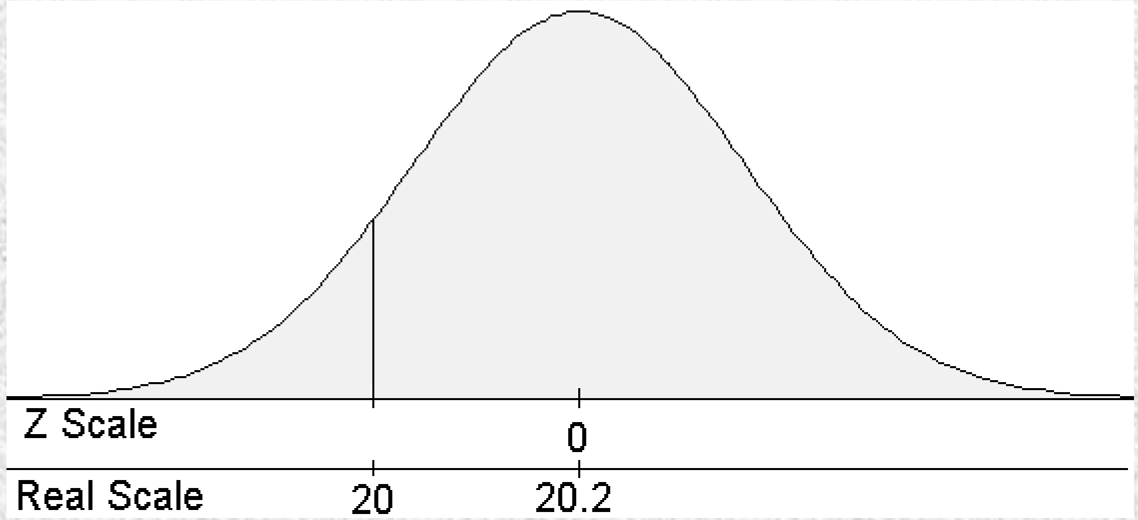
$$\text{From Table 3, } \frac{? - 1}{.1} = 1.645$$

$$? = 1.645(.1) + 1 = 1.16$$

Example

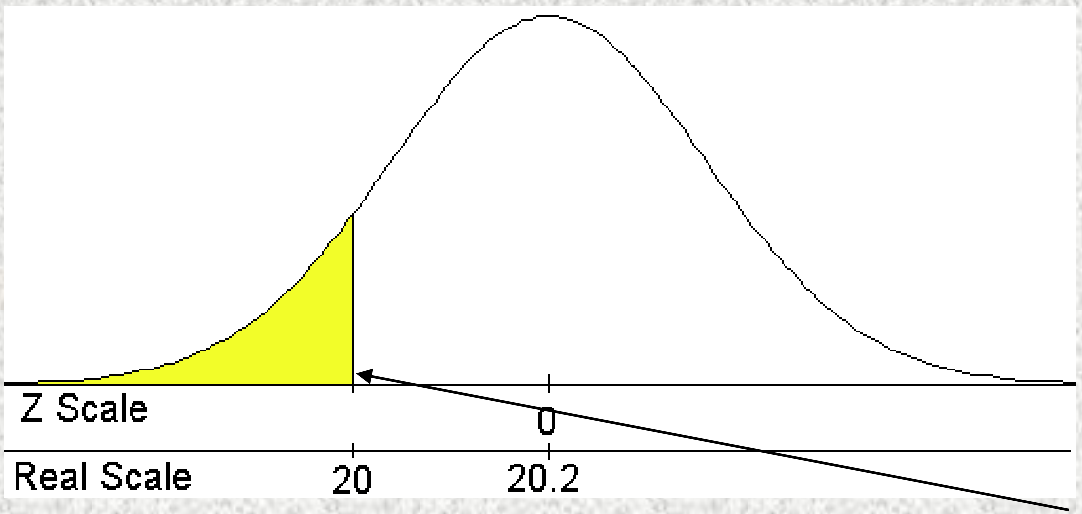
A Company produces “20 ounce” jars of a picante sauce. The true amounts of sauce in the jars of this brand sauce follow a normal distribution.

Suppose the companies “20 ounce” jars follow a normally distribution with a mean μ=20.2 ounces with a standard deviation σ=0.125 ounces.



Example

What proportion of the jars are under-filled (i.e., have less than 20 ounces of sauce)?



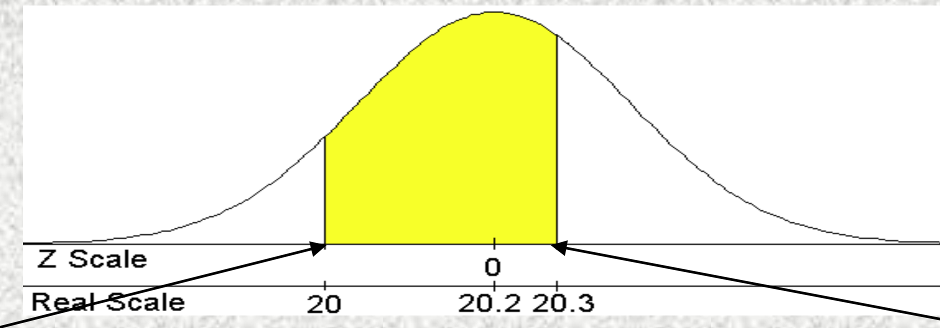
$$z = \frac{x - \mu}{\sigma}$$

$$= \frac{20 - 20.2}{0.125} = -1.60$$

P(z<-1.60) = P(z>1.60) = P(z>0)-P(0<z<1.60) = .5-.4452 = .0548. The proportion of the sauce jars that are under-filled is .0548

Example

What proportion of the sauce jars contain between 20 and 20.3 ounces of sauce.



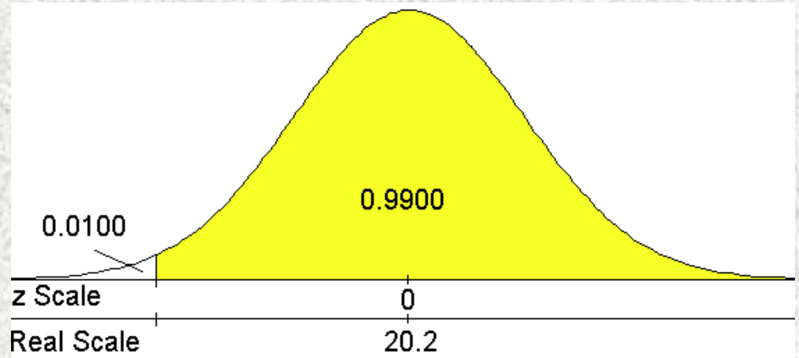
$$Z = \frac{20 - 20.2}{0.125} = -1.60 \qquad Z = \frac{20.3 - 20.2}{0.125} = 0.80$$

P(-1.60<z<.80) = P(-1.60<z<0)+P(0<z<.80) = P(0<z<1.60)+P(0<z<.80)=.4452+.2881=.7333

P(-1.60<z<.80) = P(z<.80)-P(z<-1.60)=.5+P(0<z<.80)-[.5-P(0<z<1.60)]=P(0<z<1.60)+P(0<z<.80)=.7333

Example

99% of the jars of this brand of picante sauce will contain more than what amount of sauce?



$$.99 = P(x > ?) = P(z > \frac{? - 20.2}{.125})$$

$$.01 = P(z < \frac{? - 20.2}{.125}) = P(z > \frac{20.2 - ?}{.125}) = .5 - P(0 < z < \frac{20.2 - ?}{.125})$$

$$P(0 < z < \frac{20.2 - ?}{.125}) = .49$$

$$\text{From Table 3, } \frac{20.2 - ?}{.125} = 2.33$$

$$? = 20.2 - 2.33(.125) = 19.91$$

Estimating the Parameters of a Normal Distribution

- **The parameters *μ and σ of a normal distribution represent its mean and variance, respectively.***
- **Therefore, if *$X1, \ldots, Xn$ are a random sample from a $N(\mu, \sigma^2)$ distribution,***
- ***μ is estimated with the sample mean X and σ^2* is estimated with the *sample variance s^2.***
- **As with any sample mean, the uncertainty in *X is $\sigma/\sqrt{n}$, which we replace with $s/\sqrt{n}$* if *σ is unknown.***
- ***In addition $\mu_X = \mu$, so X is unbiased for μ.***

Linear combinations of Normal random variables

Suppose two rats A and B have been trained to navigate a large maze.

X = Time of run for rat A $\quad X \sim \mathrm{N}(80, 10^2)$

Y = Time of run for rat B $\quad Y \sim \mathrm{N}(78, 13^2)$

On any given day what is the probability that rat A runs the maze faster than rat B?

Let $D = X - Y$ be the difference in times of rats A and B

If rat A is faster than rat B then $D < 0$ so we want $P(D < 0)$?

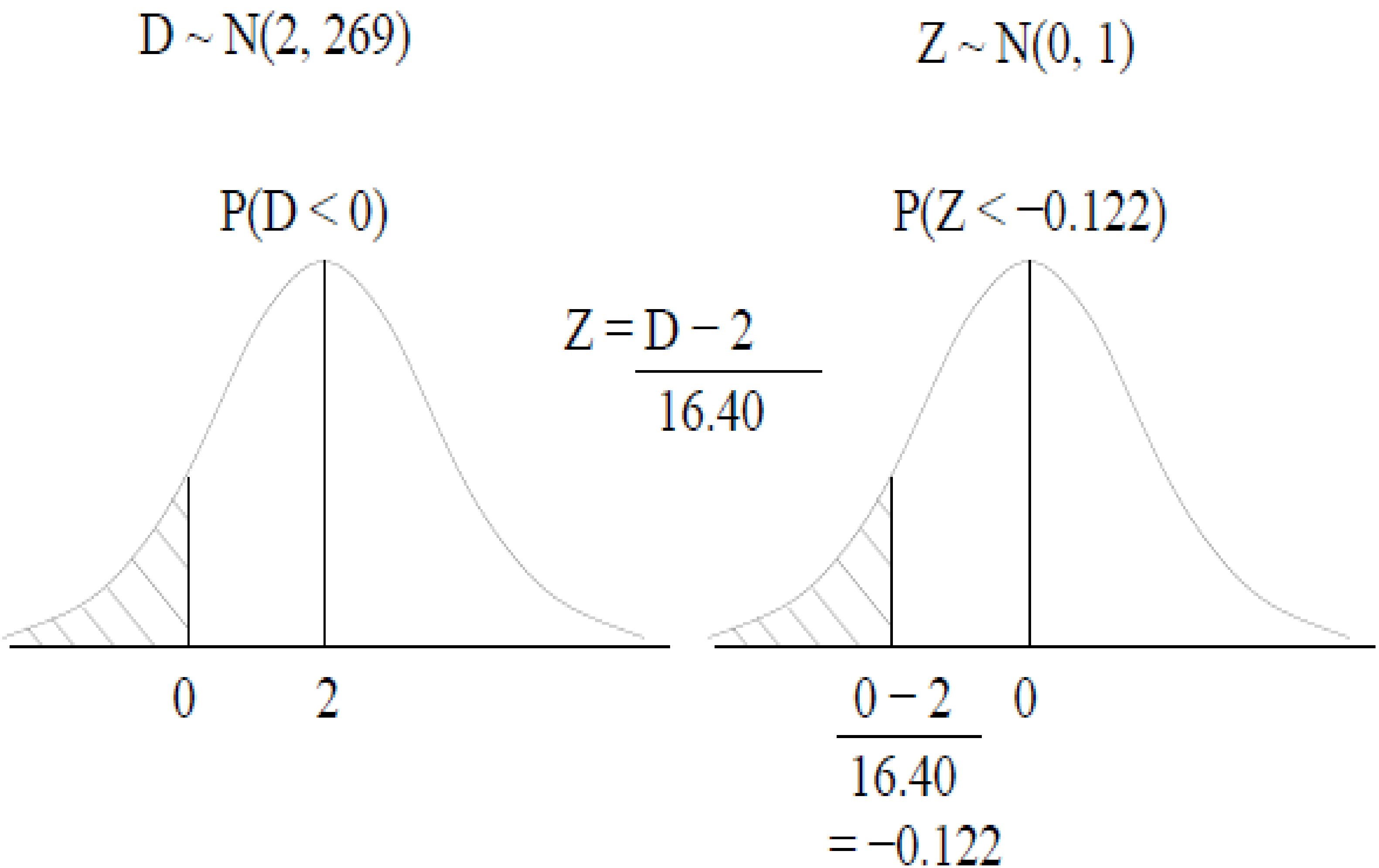
To calculate this probability we need to know the distribution of D. To do this we use the following rule

If X and Y are two independent normal variable such that $X \sim \mathrm{N}(\mu_1, \sigma_1^2)$ and $Y \sim \mathrm{N}(\mu_2, \sigma_2^2)$ then $\quad X - Y \sim \mathrm{N}(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$

In this example,

$$D = X - Y \sim \mathrm{N}(80 - 78, 10^2 + 13^2) = N(2, 269)$$

We can now calculate this probability through standardization



$$
\begin{aligned}
P(D < 0) &= P\left(\frac{D-2}{\sqrt{269}} < \frac{0-2}{\sqrt{269}}\right) \\
&= P(Z < -0.122) \qquad\qquad Z \sim N(0,1) \\
&= 1 - (0.2 \times 0.5478 + 0.8 \times 0.5517) \\
&= 0.45142
\end{aligned}
$$

Other rules that are often used are

If X and Y are two independent normal variables such that

$$X \sim \mathrm{N}(\mu_1, \sigma_1^2) \text{ and } Y \sim \mathrm{N}(\mu_2, \sigma_2^2)$$

then

$$X + Y \sim \mathrm{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$
$$aX \sim \mathrm{N}(a\mu_1, a^2\sigma_1^2)$$
$$aX + bY \sim \mathrm{N}(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2)$$

Let $X \sim N(\mu, \sigma^2)$, and let $a \neq 0$ and b be constants. Then

$$aX + b \sim N(a\mu + b, a^2\sigma^2).$$

A chemist measures the temperature of a solution in °C. The measurement is denoted C, and is normally distributed with mean 40°C and standard deviation 1°C. The measurement is converted to °F by the equation $F = 1.8C + 32$. What is the distribution of F?

Solution

Since C is normally distributed, so is F. Now $\mu_C = 40$, so $\mu_F = 1.8(40) + 32 = 104$, and $\sigma_C^2 = 1$, so $\sigma_F^2 = 1.8^2(1) = 3.24$. Therefore $F \sim N(104, 3.24)$.

Let $X_1, X_2, \ldots, X_n$ be independent and normally distributed with means $\mu_1, \mu_2, \ldots, \mu_n$ and variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$. Let $c_1, c_2, \ldots, c_n$ be constants, and $c_1X_1 + c_2X_2 + \cdots + c_nX_n$ be a linear combination. Then

$$c_1X_1+c_2X_2+\cdots+c_nX_n \sim N(c_1\mu_1+c_2\mu_2+\cdots+c_n\mu_n,\ c_1^2\sigma_1^2+c_2^2\sigma_2^2+\cdots+c_n^2\sigma_n^2)$$

Let $X_1, \ldots, X_n$ be independent and normally distributed with mean μ and variance σ^2. Then

$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \tag{4.27}$$

the strength, ductility, and hardness of welds made from nearly pure titanium. The equation is $E = 2C + 3.5N + O$, where E is the oxygen equivalence, and C, N, and

O are the proportions by weight, in parts per million, of carbon, nitrogen, and oxygen, respectively (a constant term involving iron content has been omitted). Assume that for a particular grade of commercially pure titanium, the quantities C, N, and O are approximately independent and normally distributed with means $\mu_C = 150$, $\mu_N = 200$, $\mu_O = 1500$, and standard deviations $\sigma_C = 30$, $\sigma_N = 60$, $\sigma_O = 100$. Find the distribution of E. Find $P(E > 3000)$.

Solution

Since E is a linear combination of independent normal random variables, its distribution is normal. We must now find the mean and variance of E. Using Equation (4.26), we compute

$$\begin{aligned}\mu_E &= 2\mu_C + 3.5\mu_N + 1\mu_O \\ &= 2(150) + 3.5(200) + 1(1500) \\ &= 2500 \\ \sigma_E^2 &= 2^2\sigma_C^2 + 3.5^2\sigma_N^2 + 1^2\sigma_O^2 \\ &= 2^2(30^2) + 3.5^2(60^2) + 1^2(100^2) \\ &= 57{,}700\end{aligned}$$

We conclude that $E \sim N(2500,\ 57{,}700)$.

To find $P(E > 3000)$, we compute the z-score: $z = (3000 - 2500)/\sqrt{57{,}700} = 2.08$. The area to the right of $z = 2.08$ under the standard normal curve is 0.0188. So $P(E > 3000) = 0.0188$.

Key Concepts

I. Continuous Probability Distributions

1. Continuous random variables

2. Probability distributions or probability density functions

a. Curves are smooth.

b. The area under the curve between a and b represents the probability that x falls between a and b.

c. $P(x = a) = 0$ for continuous random variables.

II. The Normal Probability Distribution

1. Symmetric about its mean μ.

2. Shape determined by its standard deviation σ.

Key Concepts

III. The Standard Normal Distribution

1. **The normal random variable *z* has mean 0 and standard deviation 1.**
2. **Any normal random variable *x* can be transformed to a standard normal random variable using**

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

3. **Convert necessary values of *x* to *z*.**
4. **Use Table 3 in Appendix I to compute standard normal probabilities.**
5. **Several important *z*-values have tail areas as follows:**

Tail Area:	.005	.01	.025	.05	.10
z-Value:	2.58	2.33	1.96	1.645	1.28