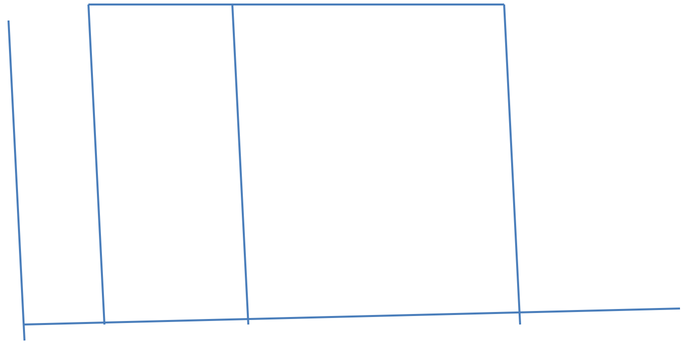
Continuous Probability Distributions

Uniform Distribution

- The density function of the continuous uniform random variable *X* on the interval [*A, B*] is

$$f(x; A, B) = \begin{cases} \frac{1}{B-A}, & A \leq x \leq B, \\ 0, & \text{elsewhere.} \end{cases}$$

- often called the **rectangular distribution**



- Suppose that a large conference room at a certain company can be reserved for no more than 4 hours. Both long and short conferences occur quite often. In fact, it can be assumed that the length X of a conference has a uniform distribution on the interval $[0, 4]$.
- (a) What is the probability density function?
- (b) What is the probability that any given conference lasts at least 3 hours?

- The appropriate density function for the uniformly distributed random variable *X* in this situation is

$$f(x) = \begin{cases} \frac{1}{4}, & 0 \le x \le 4, \\ 0, & \text{elsewhere.} \end{cases}$$

$$P[X \ge 3] = \int_3^4 \frac{1}{4}\, dx = \frac{1}{4}.$$

The mean and variance of the uniform distribution are

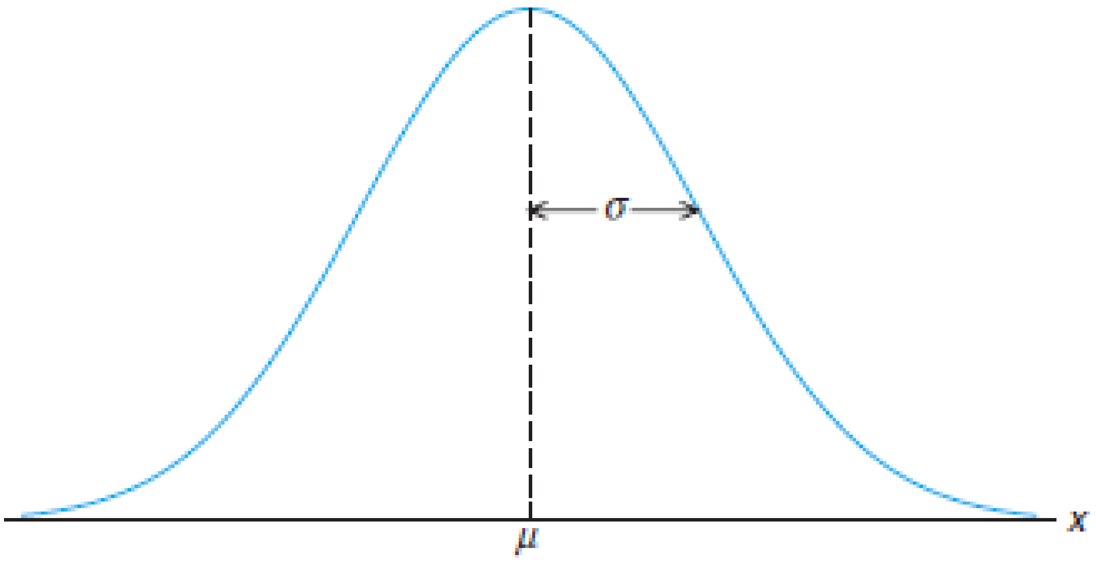
$$\mu = \frac{A+B}{2} \text{ and } \sigma^2 = \frac{(B-A)^2}{12}.$$

Normal Distribution

- The most important continuous probability distribution in the entire field of statistics

- Approximately describes many phenomena that occur in nature, industry, and research. For example, physical measurements in areas such as meteorological experiments, rainfall studies, and measurements of manufactured parts

- The normal distribution is often referred to as the **Gaussian distribution.**



Normal Curve

- A continuous random variable X having the bell-shaped distribution of like above Figure is called a **normal random variable**.

- The mathematical equation for the probability distribution of the normal variable depends on the two parameters μ and σ, its mean and standard deviation, respectively. Hence, we denote the values of the density of X by $n(x; \mu, \sigma)$.

The Normal Distribution: as mathematical function (pdf)

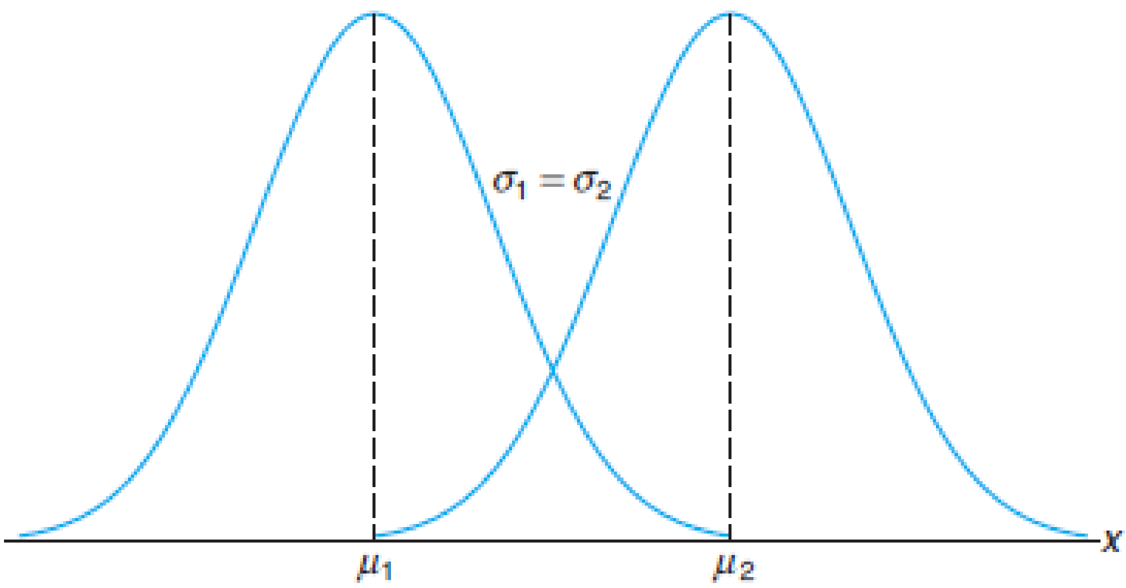
$$f(x) = \frac{1}{\boxed{\sigma}\sqrt{2\pi}} \cdot e^{-\frac{1}{2}(\frac{x-\boxed{\mu}}{\sigma})^2}$$

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

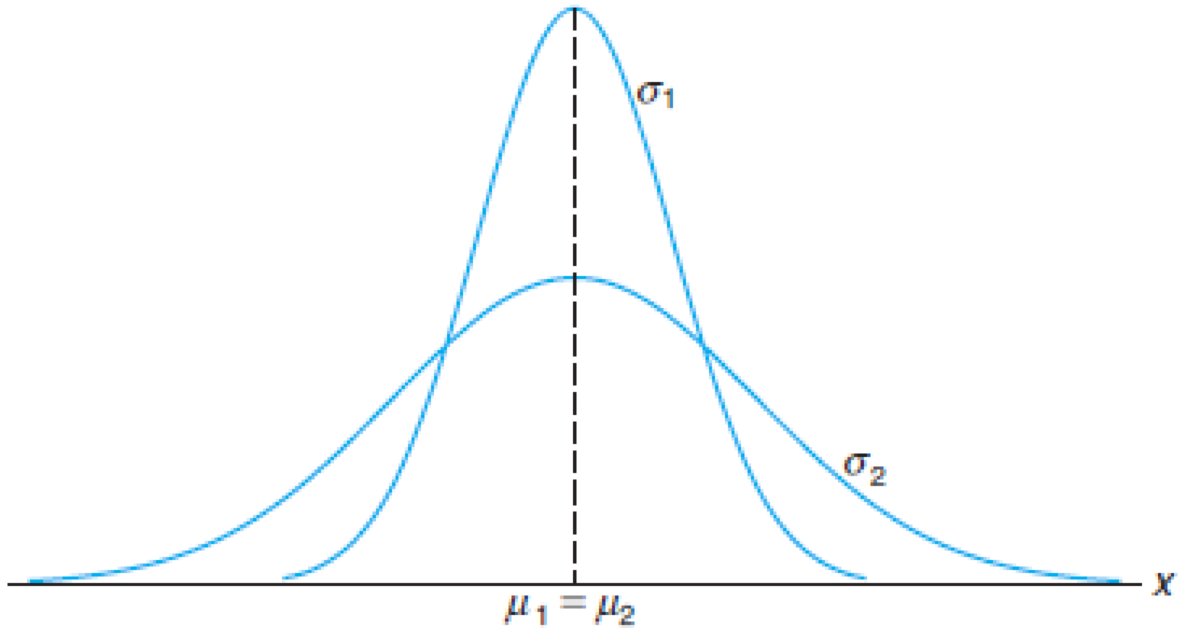
<u>Note constants:</u>
π=3.14159
e=2.71828

Impact of μ and σ *on shape and location*

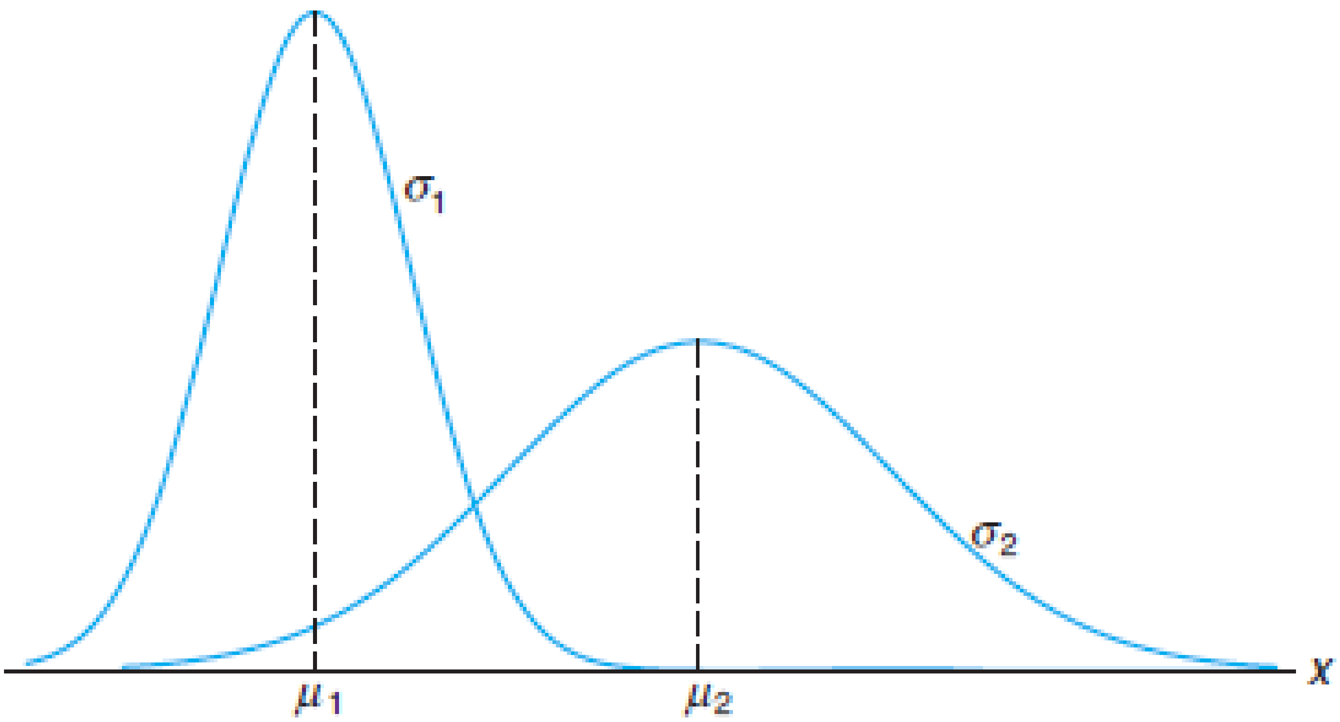
- Normal curves with $\mu 1 < \mu 2$ and $\sigma 1 = \sigma 2$.



- Normal curves with $\mu 1 = \mu 2$ and $\sigma 1 < \sigma 2$.



- Normal curves with $\mu 1 < \mu 2$ and $\sigma 1 < \sigma 2$



Properties of normal curve

1. The mode, which is the point on the horizontal axis where the curve is a maximum, occurs at $x = \mu$.
2. The curve is symmetric about a vertical axis through the mean μ.
3. The curve has its points of inflection at $x = \mu \pm \sigma$; it is concave downward if $\mu - \sigma < X < \mu + \sigma$ and is concave upward otherwise.
4. The normal curve approaches the horizontal axis asymptotically as we proceed in either direction away from the mean.
5. The total area under the curve and above the horizontal axis is equal to 1.

E(X) = μ

- To evaluate the mean, we first calculate

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

- Setting $z = (x - \mu)/\sigma$ and $dx = \sigma\, dz$, we obtain

$$E(X-\mu) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z e^{-\frac{1}{2}z^2} dz = 0,$$

since the integrand above is an odd function of z.

we conclude that $E(X) = \mu$.

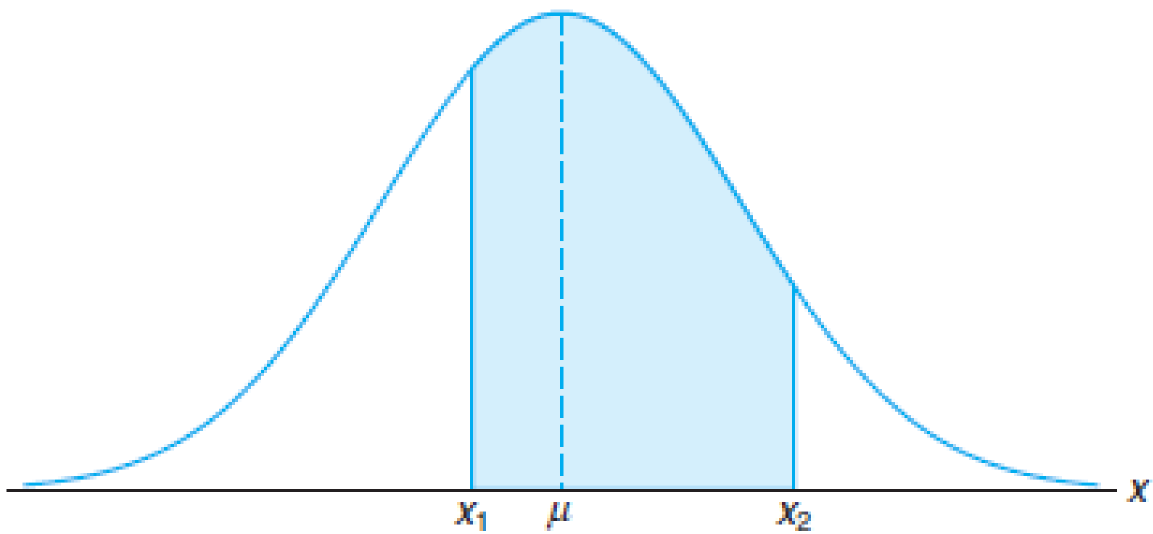
Home Task

- Verify following
  - It's a probability function, so no matter what the values of μ and σ, must integrate to 1!
  - $Var(X)=\sigma^2$

Areas under the Normal Curve

$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} n(x;\mu,\sigma)\,dx = \frac{1}{\sqrt{2\pi}\sigma}\int_{x_1}^{x_2} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}dx$$

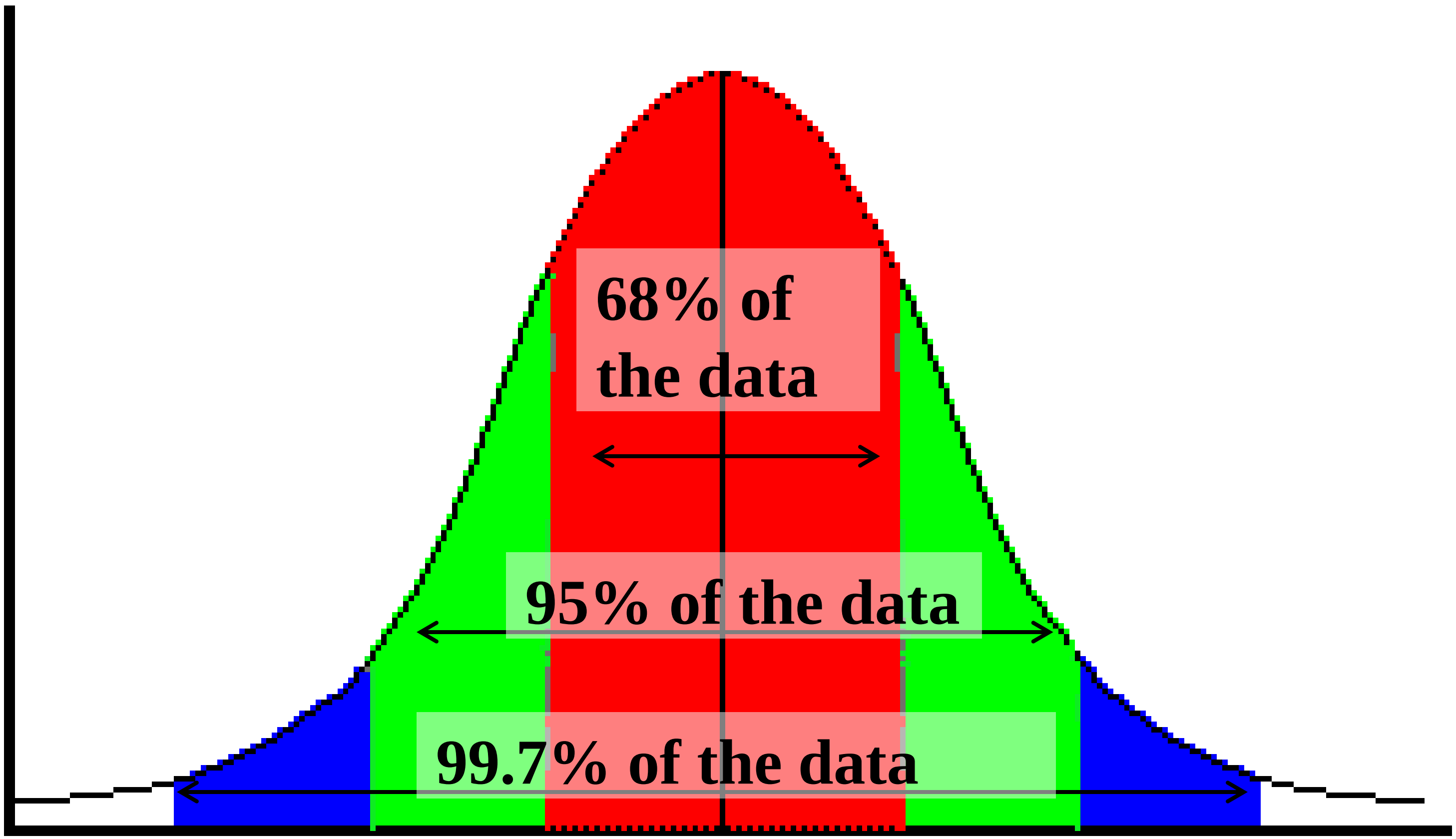
is represented by the area of the shaded region.



**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.

68-95-99.7 Rule



68-95-99.7 Rule in Math terms…

$$\int_{\mu-\sigma}^{\mu+\sigma} \frac{1}{\sigma\sqrt{2\pi}} \bullet e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx = 0.68$$

$$\int_{\mu-2\sigma}^{\mu+2\sigma} \frac{1}{\sigma\sqrt{2\pi}} \bullet e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx = 0.95$$

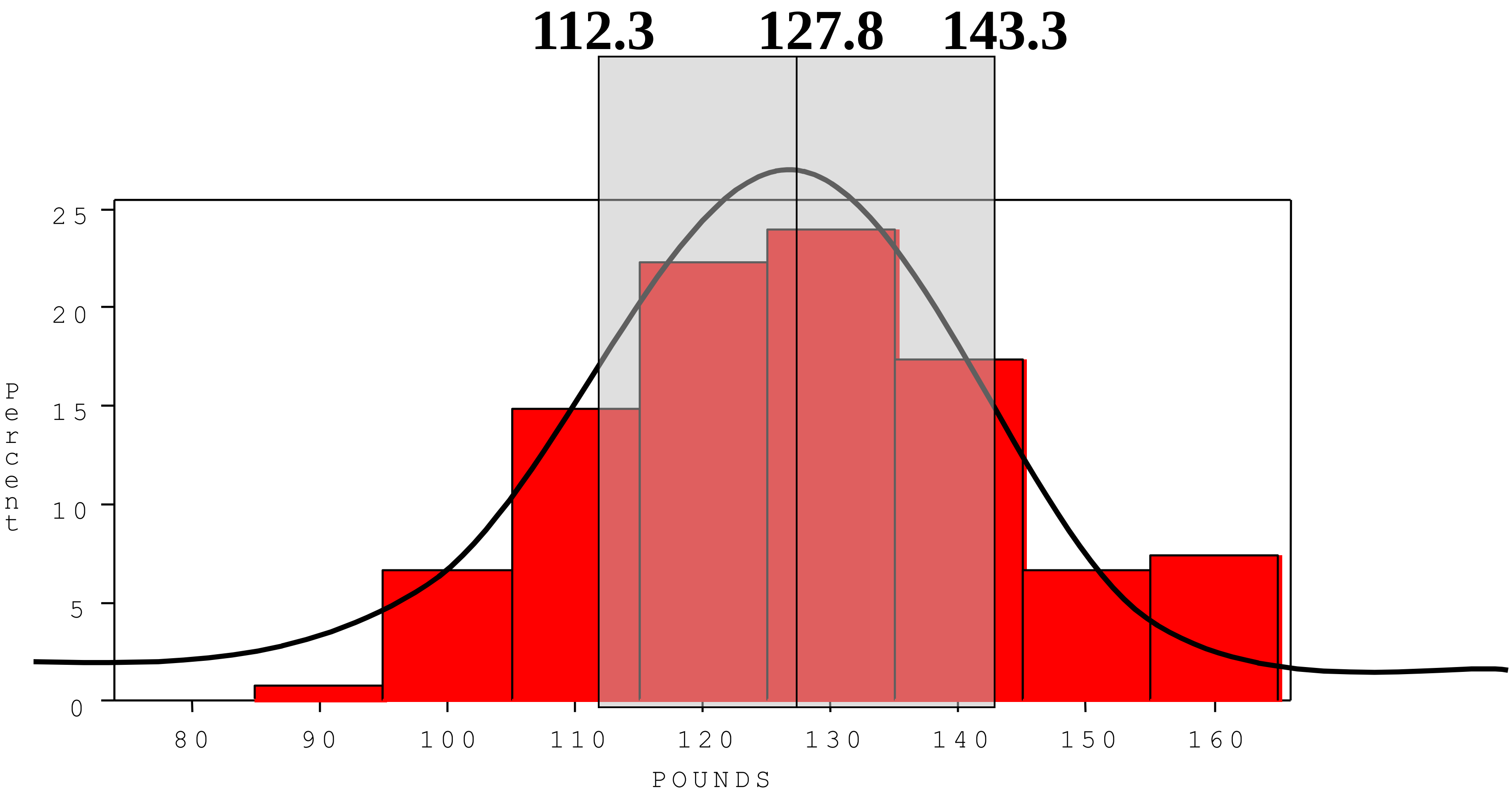
$$\int_{\mu-3\sigma}^{\mu+3\sigma} \frac{1}{\sigma\sqrt{2\pi}} \bullet e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx = 0.997$$

How good is rule for real data?

Check some example data:
The mean of the weight of the women = 127.8
The standard deviation (SD) = 15.5
Total participant = 120

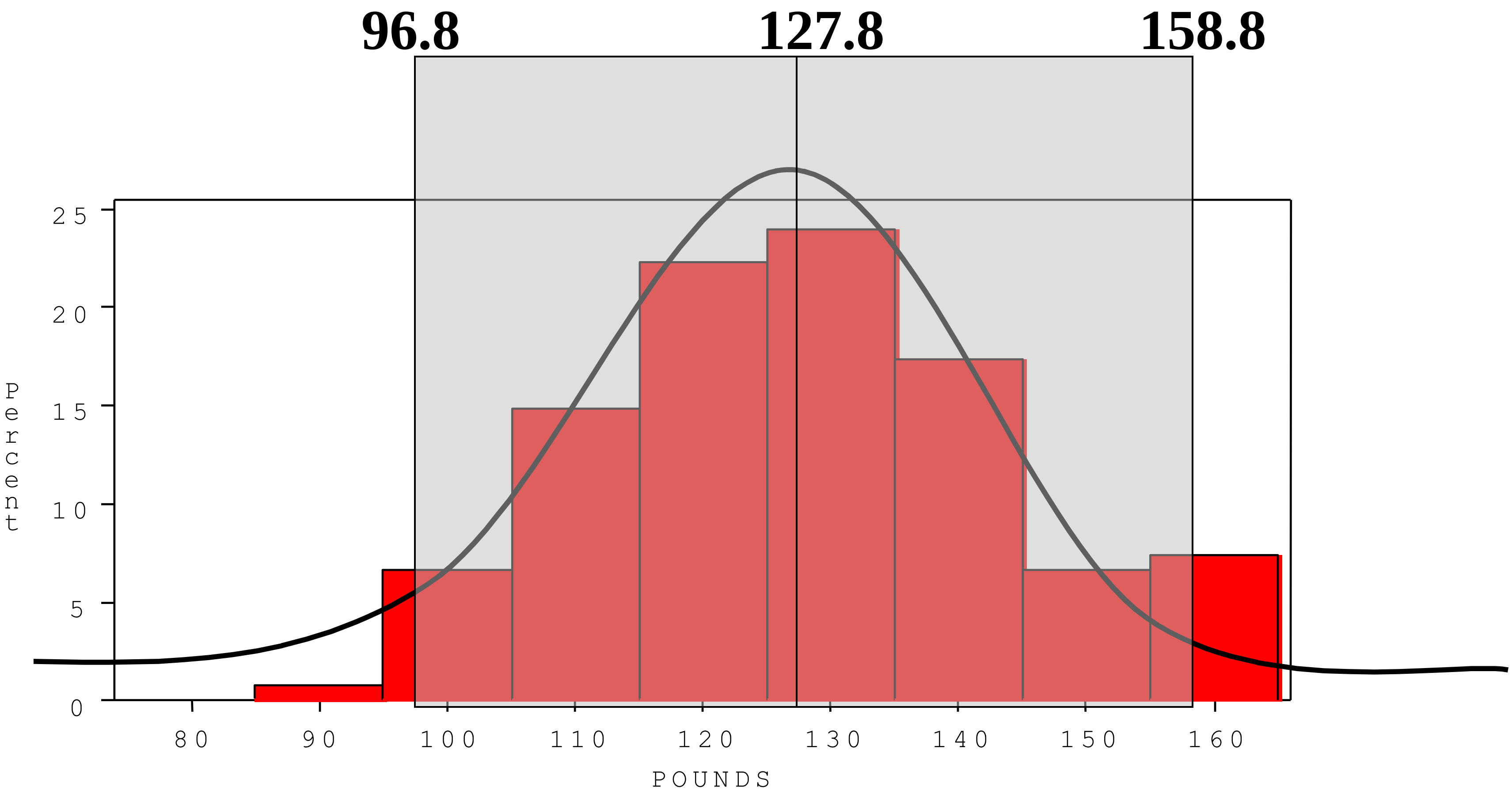
68% of 120 = 0.68x120 = ~ 82 runners

In fact, 79 runners fall within 1-SD of the mean.



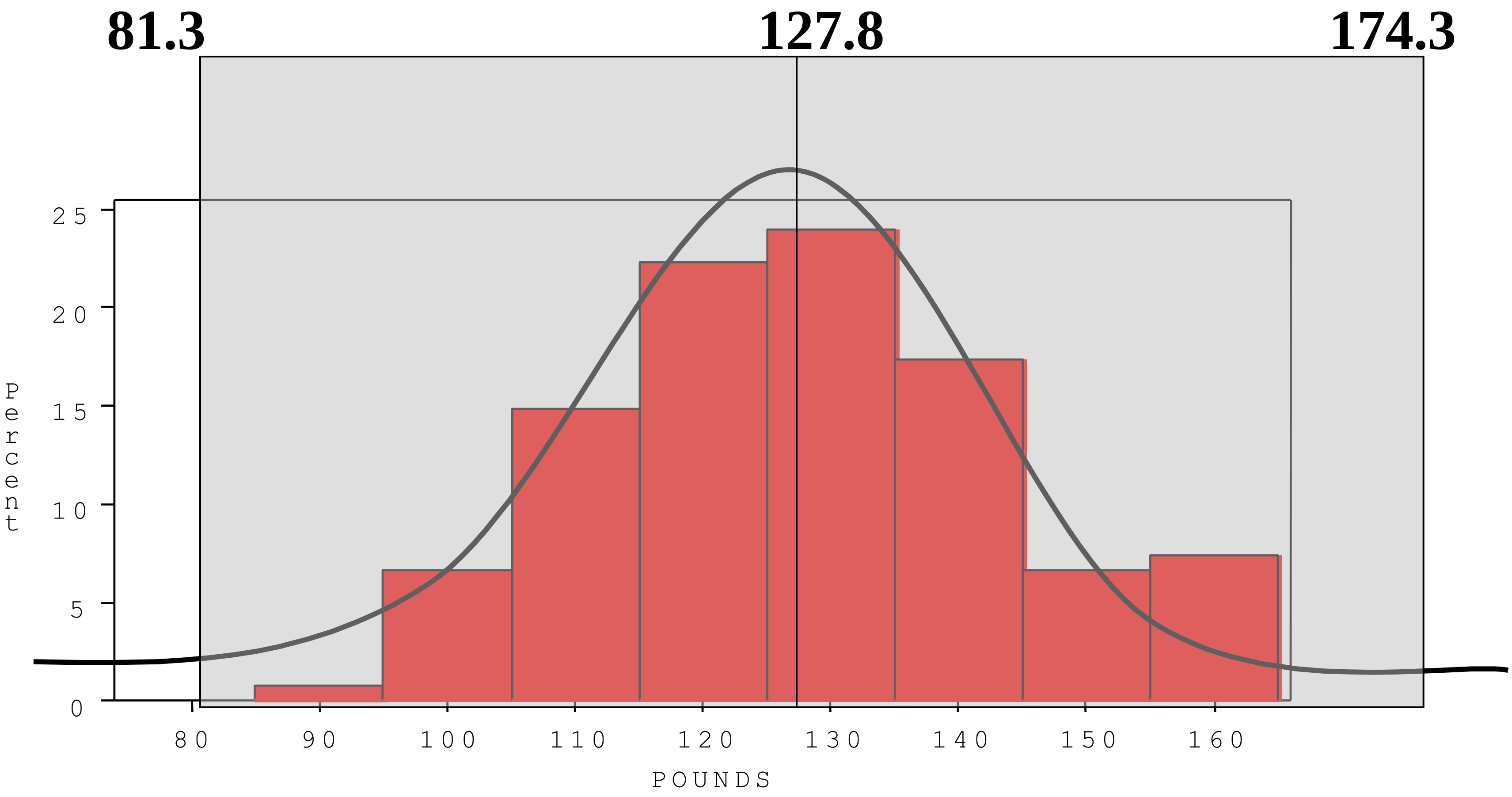
95% of 120 = 0.95 x 120 = ~ 114 runners

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



99.7% of 120 = 0.997 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 number of students who scores less than equal to 575

P(X≤575)

$$\int_{200}^{575} \frac{1}{(50)\sqrt{2\pi}} \bullet e^{-\frac{1}{2}(\frac{x-500}{50})^2} dx$$

The Standard Normal (Z):

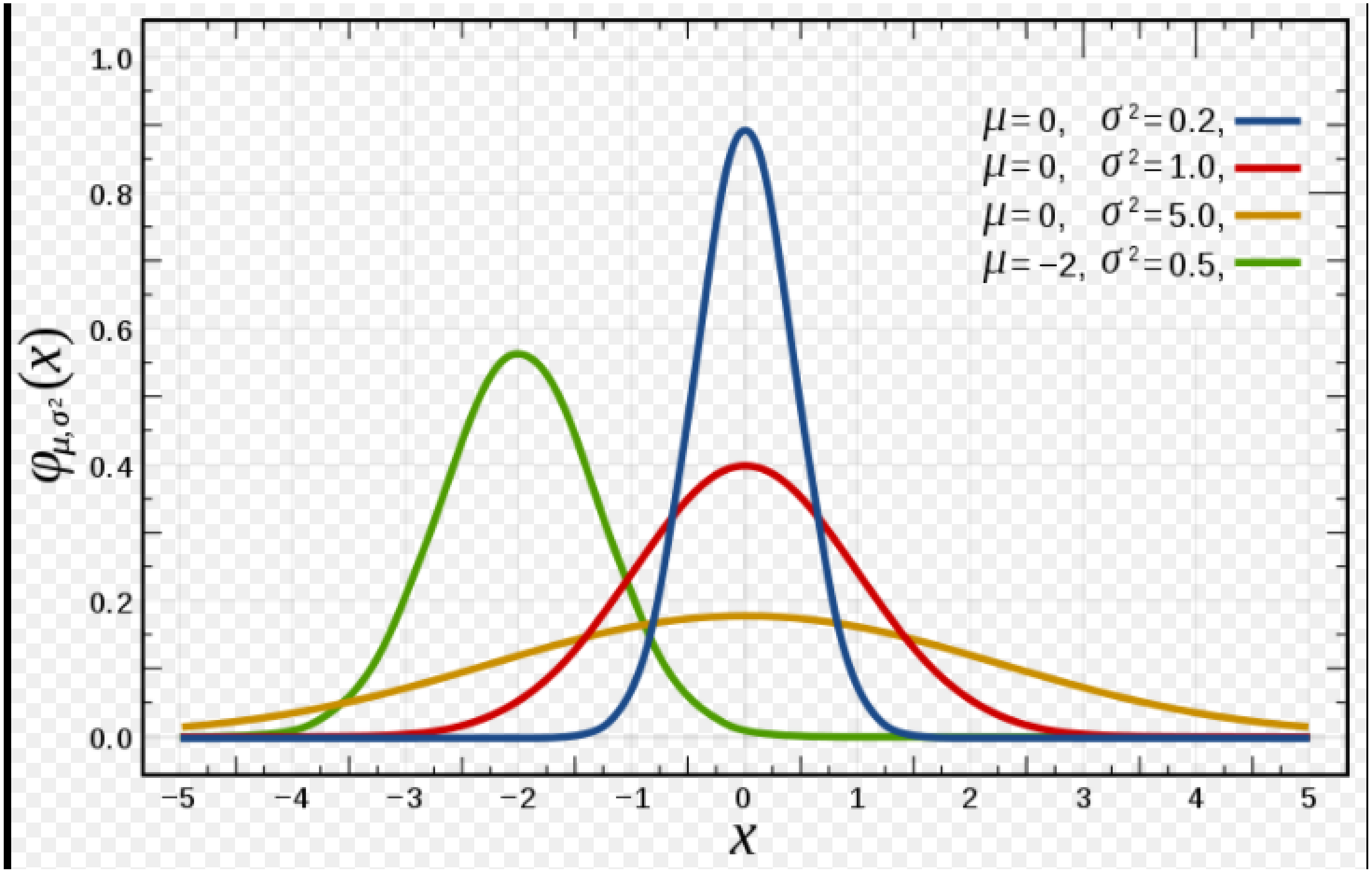
The formula for the standardized normal probability density function is

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

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:

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



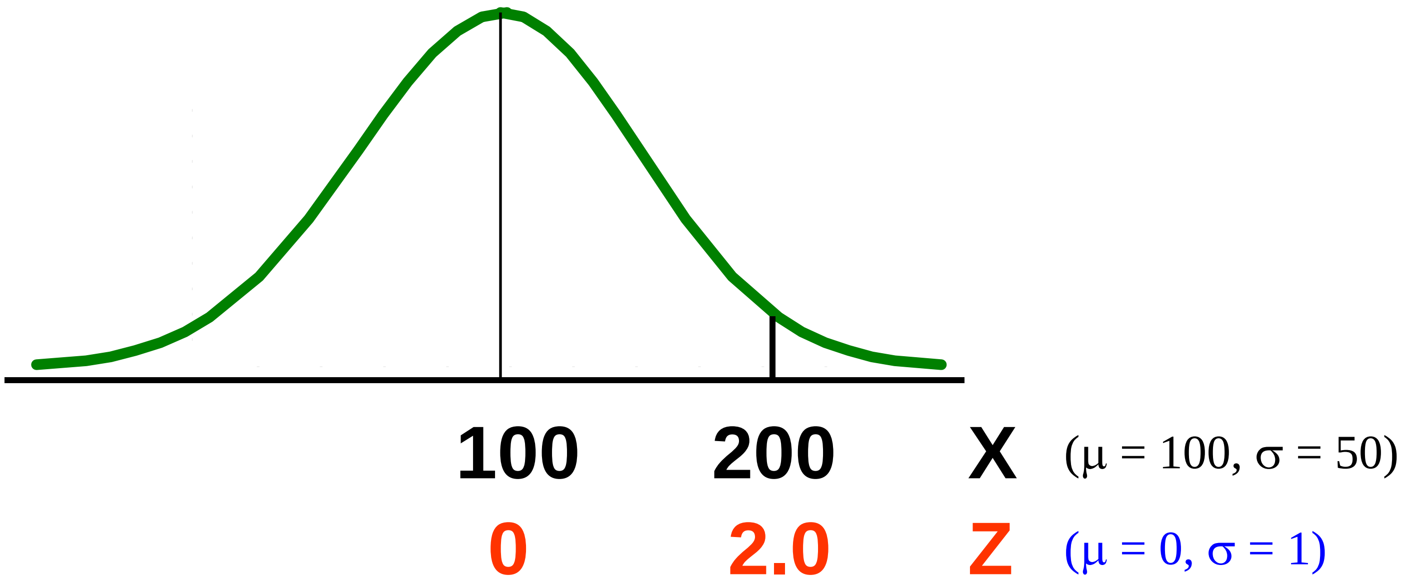
$\mu=0, \quad \sigma^2=0.2,$
$\mu=0, \quad \sigma^2=1.0,$
$\mu=0, \quad \sigma^2=5.0,$
$\mu=-2, \quad \sigma^2=0.5,$
1.0
0.8
0.6
0.4
0.2
0.0
$\varphi_{\mu,\sigma^2}(x)$
−5 −4 −3 −2 −1 0 1 2 3 4 5
x

- Whenever X assumes a value x, the corresponding value of Z is given by $z = (x - \mu)/\sigma$. Therefore, if X falls between the values $x = x1$ and $x = x2$, the random variable Z will fall between the corresponding values $z1 = (x1 - \mu)/\sigma$ and $z2 = (x2 - \mu)/\sigma$. Consequently, we may write

$$P(x_1 < X < x_2) = \frac{1}{\sqrt{2\pi}\sigma}\int_{x_1}^{x_2} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx = \frac{1}{\sqrt{2\pi}}\int_{z_1}^{z_2} e^{-\frac{1}{2}z^2}\, dz$$

$$= \int_{z_1}^{z_2} n(z;0,1)\, dz = P(z_1 < Z < z_2),$$

Comparing X and Z units



Example

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

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

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

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

Yikes!

But to look up Z= 1.5 in standard normal chart = 0.9332

Practice problem

If birth weights in a population are normally distributed with a mean of 109 oz and a standard deviation of 13 oz,

a. What is the chance of obtaining a birth weight of 141 oz *or heavier* when sampling birth records at random?

b. What is the chance of obtaining a birth weight of 120 *or lighter*?

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	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	0.7852
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	0.8810	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	0.9772	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	0.9864	0.9868	0.9871	0.9875	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	0.9925	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
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Answer

a. What is the chance of obtaining a birth weight of 141 oz *or heavier* when sampling birth records at random?

$$\boxed{Z = \frac{141 - 109}{13} = 2.46}$$

From the chart Z of 2.46 corresponds to a right tail (greater than) area of: $P(Z \geq 2.46) = 1-(0.9931) = 0.0069$ or 0.69 %

Answer

b. What is the chance of obtaining a birth weight of 120 *or lighter*?

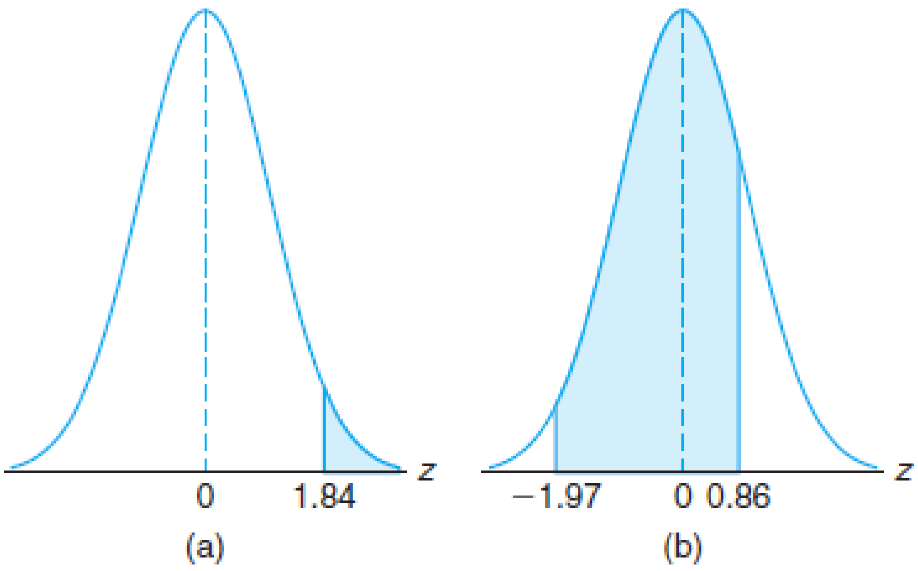
$$\boxed{Z = \frac{120 - 109}{13} = 0.85}$$

From the chart Z of 0.85 corresponds to a left tail area of:
$P(Z \leq 0.85) = 0.8023 = 80.23\%$

- Given a standard normal distribution, find the area under the curve that lies

(a) to the right of $z = 1.84$ and

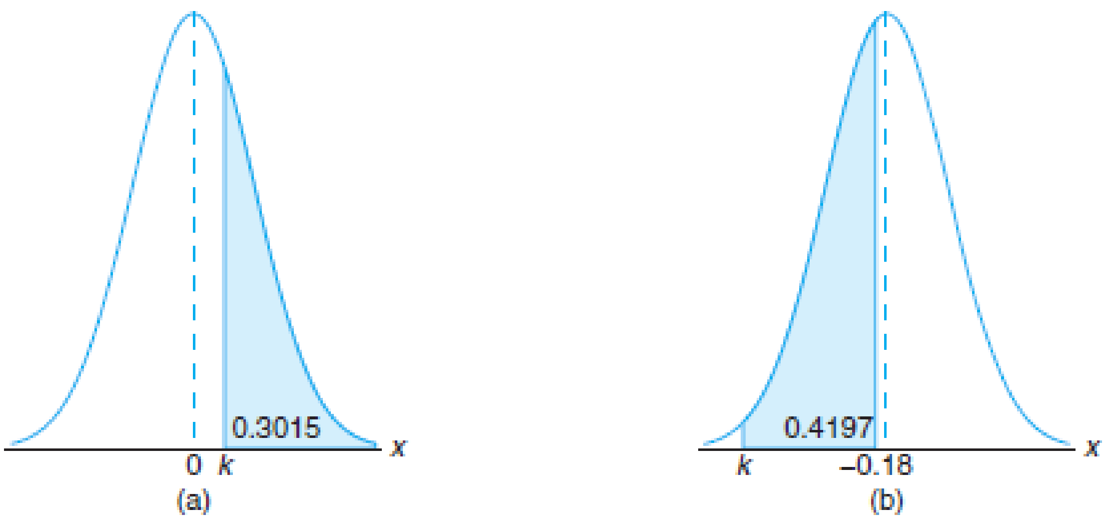
(b) between $z = -1.97$ and $z = 0.86$.



(a) The area in Figure (a) to the right of $z = 1.84$ is equal to 1 minus the area in normal Table to the left of $z = 1.84$, namely, $1 - 0.9671 = 0.0329$.

(b) The area in Figure (b) between $z = -1.97$ and $z = 0.86$ is equal to the area to the left of $z = 0.86$ minus the area to the left of $z = -1.97$. From normal Table we find the desired area to be $0.8051 - 0.0244 = 0.7807$.

- Given a standard normal distribution, find the value of *k* such that
- (a) $P(Z > k) = 0.3015$ and
- (b) $P(k < Z < -0.18) = 0.4197$.

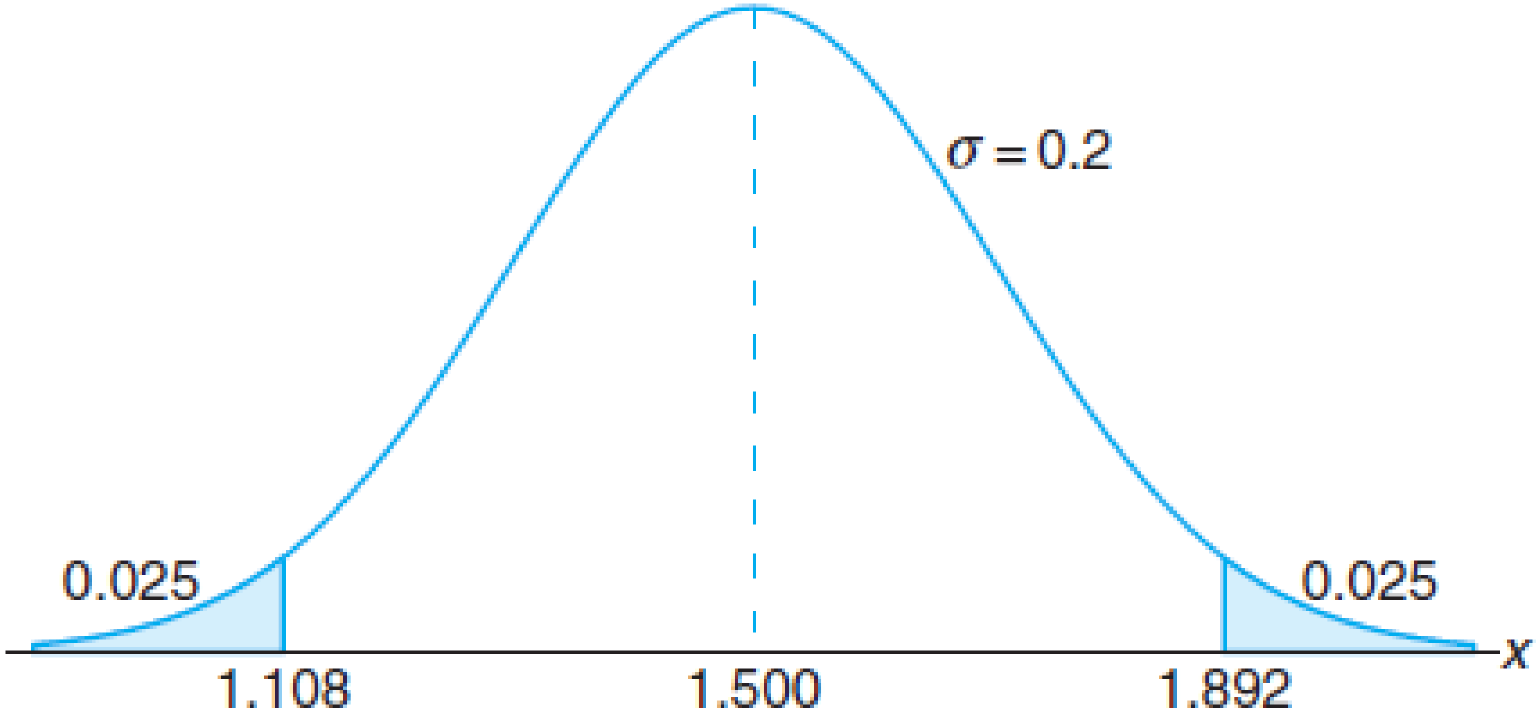


(a) In Figure (a), we see that the k value leaving an area of 0.3015 to the right must then leave an area of 0.6985 to the left. From normal Table it follows that $k = 0.52$.

(b) From normal Table we note that the total area to the left of -0.18 is equal to 0.4286. In Figure (b), we see that the area between k and -0.18 is 0.4197, so the area to the left of k must be $0.4286 - 0.4197 = 0.0089$. Hence, from normal Table, we have $k = -2.37$.

- Gauges are used to reject all components for which a certain dimension is not within the specification 1.50 ± *d*. It is known that this measurement is normally distributed with mean 1.50 and standard deviation 0.2. Determine the value *d* such that the specifications “cover” 95% of the measurements.
- $P(-z1<Z<z1) = 0.95$
- $P(Z<z1) = 0.95/2+0.5 = 0.975$
- $z1=1.96$
- $(x1-\mu)/\sigma=z1 \Rightarrow x1=z1^* \sigma + \mu = 1.96^*0.2 + 1.5=0.392+1.5=1.892$

- Negative side = 1.5 – 0.392=1.018


σ = 0.2
0.025
0.025
x
1.108
1.500
1.892

- A certain machine makes electrical resistors having a mean resistance of 40 ohms and a standard deviation of 2 ohms. Assuming that the resistance follows a normal distribution Find the percentage of resistances exceeding 43 ohms resistance is measured to the nearest ohm.
- We assign a measurement of 43 ohms to all resistors whose resistances are greater than 42.5 and less than 43.5. We are actually approximating a discrete distribution by means of a continuous normal distribution.
- $z = (43.5 - 40)/2 = 1.75.$
- $P(X > 43.5) = P(Z > 1.75) = 1 - P(Z < 1.75) = 1 - 0.9599 = 0.0401.$

Are my data normally distributed?

1. Look at the histogram! Does it appear bell shaped?
2. Compute descriptive summary measures—are mean, median, and mode similar?
3. Do 2/3 of observations lie within 1 std dev of the mean? Do 95% of observations lie within 2 std dev of the mean?

Approximation of Binomial distribution using Normal Distribution

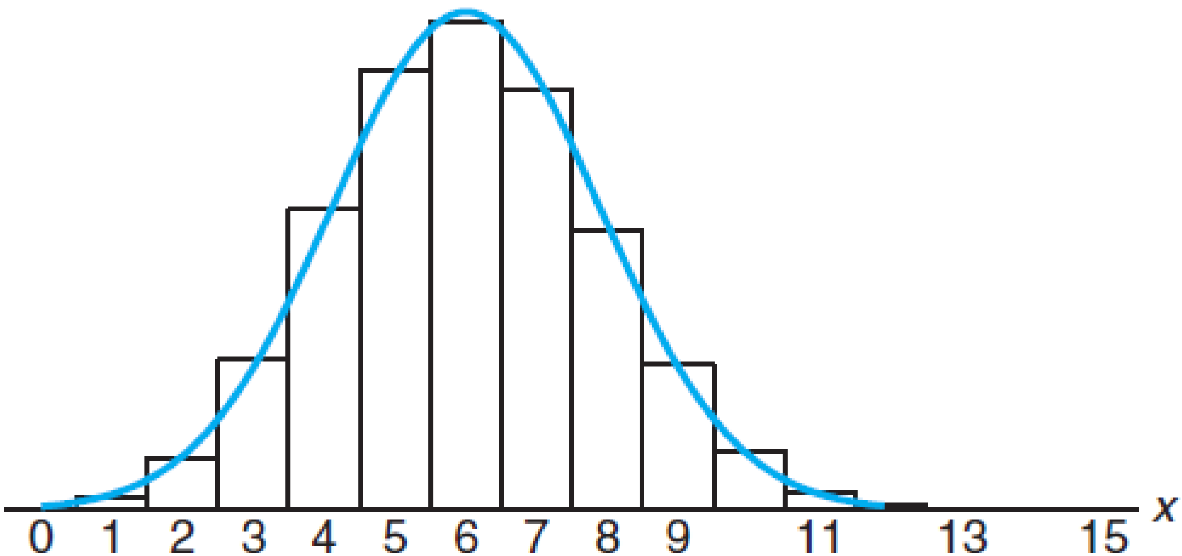
- If X is a binomial random variable with mean $\mu = np$ and variance $\sigma^2 = npq$, then the limiting form of the distribution of

$$Z = \frac{X - np}{\sqrt{npq}},$$

as $n \to \infty$, is the standard normal distribution $n(z; 0, 1)$.

- It turns out that the normal distribution with $\mu = np$ and $\sigma^2 = np(1 - p)$ not only provides a very accurate approximation to the binomial distribution when n is large and p is not extremely close to 0 or 1 but also provides a fairly good approximation even when n is small and p is reasonably close to 1/2.

Example comparison of binomial and normal distribution



Normal approximation of $b(x; 15, 0.4)$

Thumb Rule

If n, p, and q are such that:

np and nq

are both greater than 5.

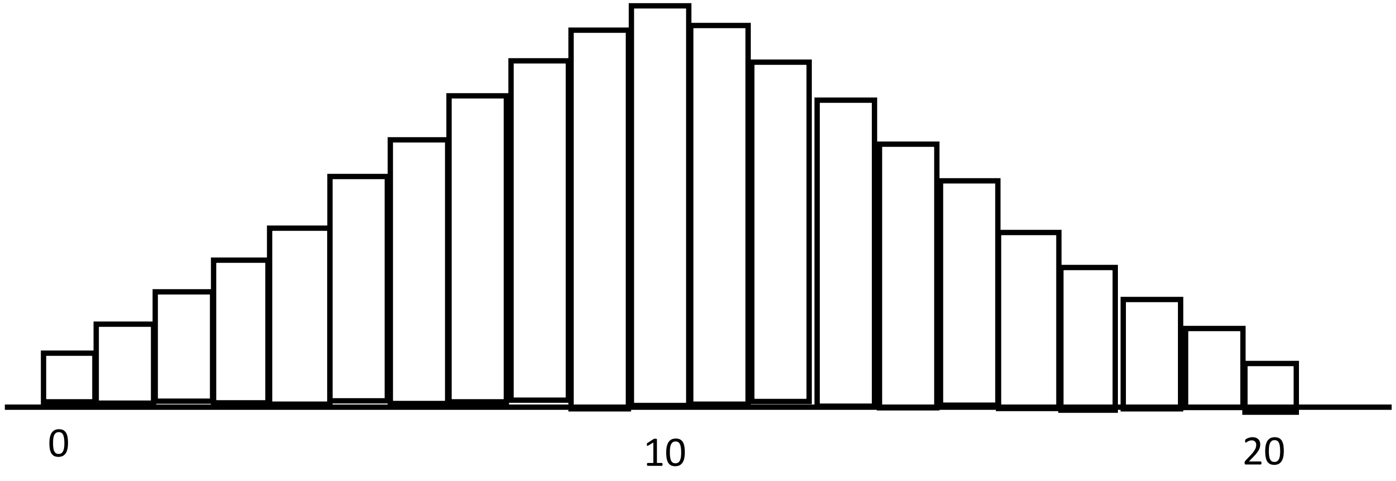
Mean and Standard Deviation: Binomial Distribution

$$\mu = np \text{ and } \sigma = \sqrt{npq}$$

Experiment: tossing a coin 20 times

Problem: Find the probability of getting exactly 10 heads.

Distribution of the number of heads appearing should look like:



Using the Binomial Probability Formula

n = 20

x = 10

p = 0.5

q = 1 – p = 0.5

P(10) = <u>0.176197052</u>

Normal Approximation of the Binomial Distribution

First calculate the mean and standard deviation:

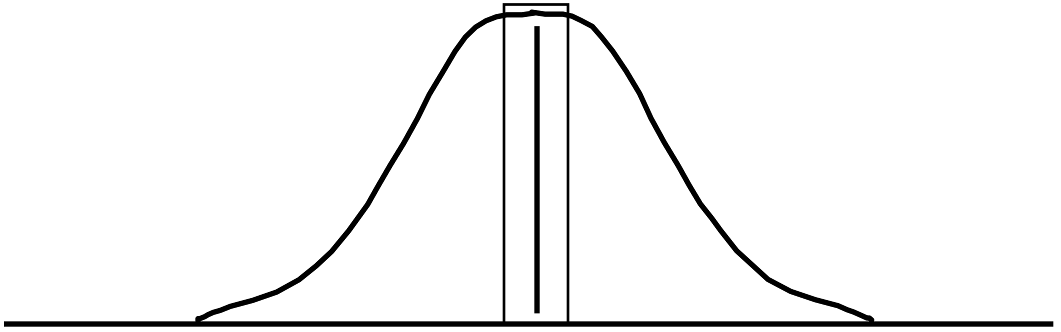
$$\mu = np = 20\,(0.5) = 10$$

$$\sigma = \sqrt{np(1-p)} = \sqrt{20(.5)(.5)} = \sqrt{5} \approx 2.24$$

The Continuity Correction

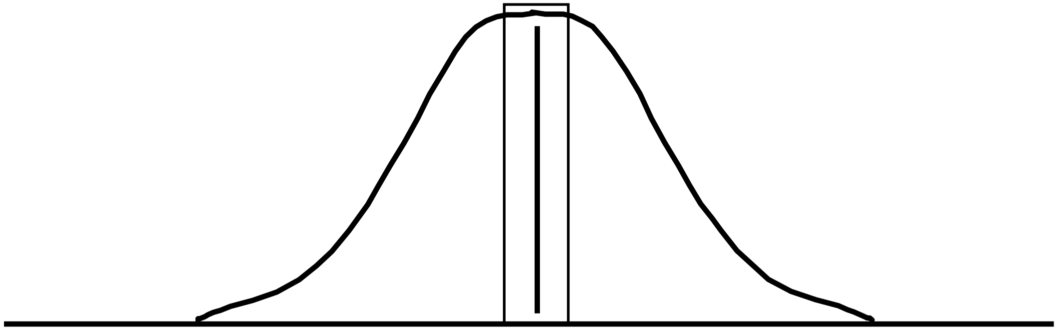
<u>Continuity Correction</u>: to compute the probability of getting <u>exactly</u> 10 heads, find

the probability of getting between 9.5 and 10.5 heads.



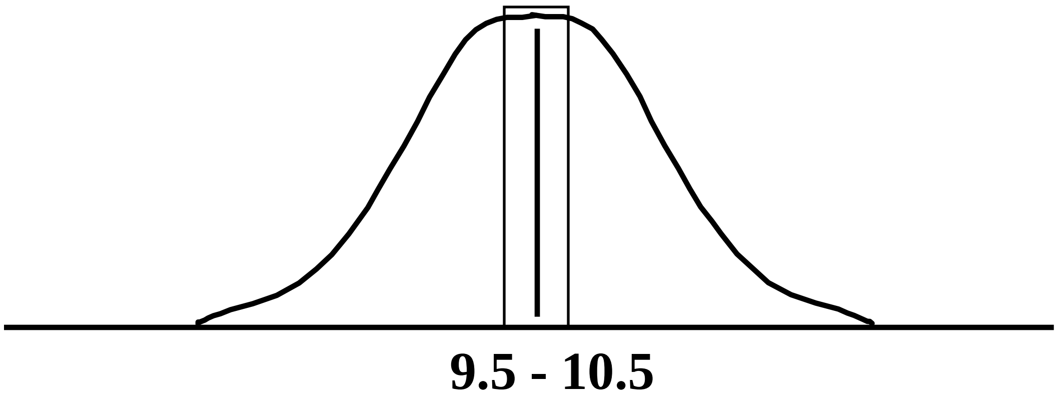
The Continuity Correction

Continuity Correction is needed because we are approximating a discrete probability distribution with a continuous distribution.



The Continuity Correction

We are using the area under the curve to approximate the area of the rectangle.



Using the Normal Distribution

$P(9.5 \leq x \leq 10.5) = ?$

for $x = 9.5$: $z = -0.22$

$P(z < -0.22) = .4129$

Using the Normal Distribution

for $x = 10.5$: $z = = 0.22$

$P(z < .22) = 0.5871$

$P(9.5 \leq x \leq 10.5) = 0.5871 - 0.4129 = 0.1742$

Application of Normal Distribution

If 22% of all patients with high blood pressure have side effects from a certain medication, and 100 patients are treated, find the probability that at least 30 of them will have side effects.

Using the Binomial Probability Formula we would need to compute:

$$P(30) + P(31) + \ldots + P(100) \text{ or } 1 - P(x \leq 29)$$

Using the Normal Approximation to the Binomial Distribution

Is it appropriate to use the normal distribution?

Check: $np =$

$nq =$

Using the Normal Approximation to the Binomial Distribution

$$n\,p = 22$$

$$n\,q = 78$$

Both are greater than five.

Find the mean and standard deviation

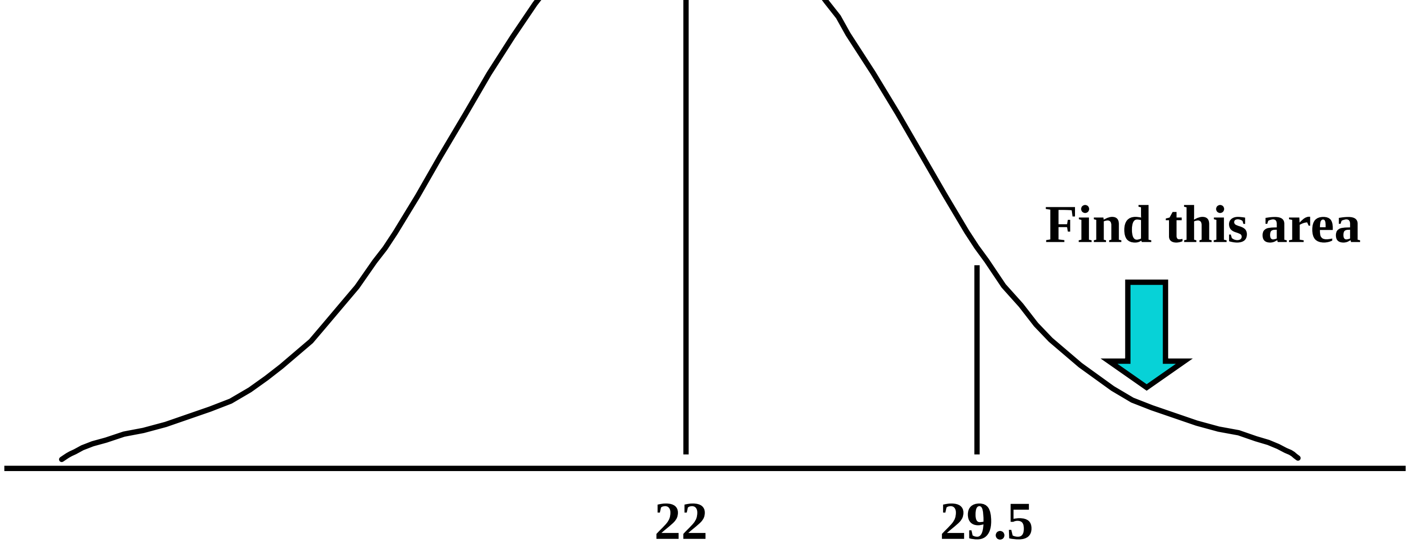
$$\mu = 100(0.22) = 22$$

and $\sigma = \sqrt{100(0.22)(0.78)} =$

$$\sqrt{17.16} = 4.14$$

Applying the Normal Distribution

To find the probability that at least 30 of them will have side effects, find P(x ≥ 29.5)

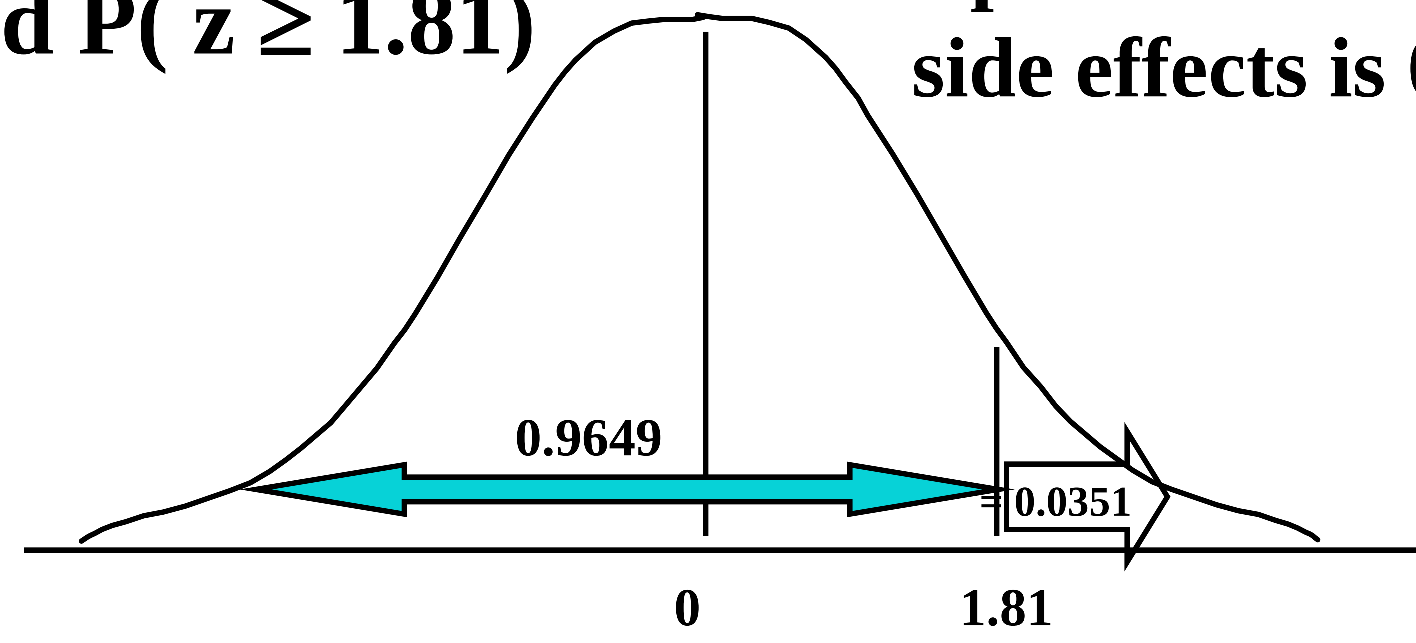


Applying the Normal Distribution

$$z = \frac{29.5 - 22}{4.14} = 1.81$$

Find P(z ≥ 1.81)

The probability that at least 30 of the patients will have side effects is 0.0351.



Reminders:

- Use the normal distribution to approximate the binomial only if both np and nq are greater than 5.

- Always use the continuity correction when approximating the binomial distribution.

The Gamma Function

The Gamma Function

- To define the family of gamma distributions, we first need to introduce a function that plays an important role in many branches of mathematics.

- **Definition** For $\alpha > 0$, the **gamma function** $\Gamma(\alpha)$ is defined by

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} \, dx$$

The Gamma Function

- The most important properties of the gamma function are the following:

- **1.** For any $\alpha > 1$, $\Gamma(\alpha) = (\alpha - 1) \cdot \Gamma(\alpha - 1)$
  [via integration by parts]

- **2.** For any positive integer, *n*, $\Gamma(n) = (n - 1)!$

- **3.** $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

The Gamma Function

- By earlier expression, we can say that
- $$f(x;\alpha) = \begin{cases} \dfrac{x^{\alpha-1}e^{-x}}{\Gamma(\alpha)} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$
- then f(x; α) ≥ 0 and $$\int_0^{\infty} f(x;\alpha)\,dx = \Gamma(\alpha)/\Gamma(\alpha) = 1$$
- so *f*(*x*; a) satisfies the two basic properties of a pdf.

The Gamma Distribution

Exponential Distribution

- Poisson distribution is used to compute the probability of specific numbers of "events" during a particular *period of time or span of space*.

- In many applications, the time period or span of space is the random variable.

- For example, an industrial engineer may be interested in modeling the time T between arrivals of cars at a congested intersection during rush hour in a large city.

- Poisson distribution was developed as a single-parameter distribution with parameter λ, where λ may be interpreted as the mean number of events *per unit time.*
- Consider now the *random variable* described by the time required for the first event to occur.
- Using the Poisson distribution, we find that the probability of no events occurring in the span up to time *t* is given by

$$p(0; \lambda t) = \frac{e^{-\lambda t}(\lambda t)^0}{0!} = e^{-\lambda t}$$

- let X be the random variable representing time to the first Poisson event.
- The probability that the length of time until the first event will exceed x is the same as the probability that no Poisson events will occur in x.

$$P(X > x) = e^{-\lambda x}.$$

- Thus, the cumulative distribution function for X is given by

$$P(0 \leq X \leq x) = 1 - e^{-\lambda x}$$

- We can differentiate the above to achieve the pdf

$$f(x) = \lambda e^{-\lambda x},$$

- In exponential distribution, often we are provided gap between two successive events in terms of parameter β
- Where $\lambda=1/\beta$
- $f(x) = 1/\beta e^{-x/\beta}$

Example

- Suppose that a system contains a certain type of component whose time, in years, to failure is given by *T*. The random variable *T* is modeled nicely by the exponential distribution with mean time to failure *β* = 5. If 5 of these components are installed in different systems, what is the probability that at least 2 are still functioning at the end of 8 years?

$$P(T > 8) = \frac{1}{5}\int_8^{\infty} e^{-t/5}\, dt = e^{-8/5} \approx 0.2.$$

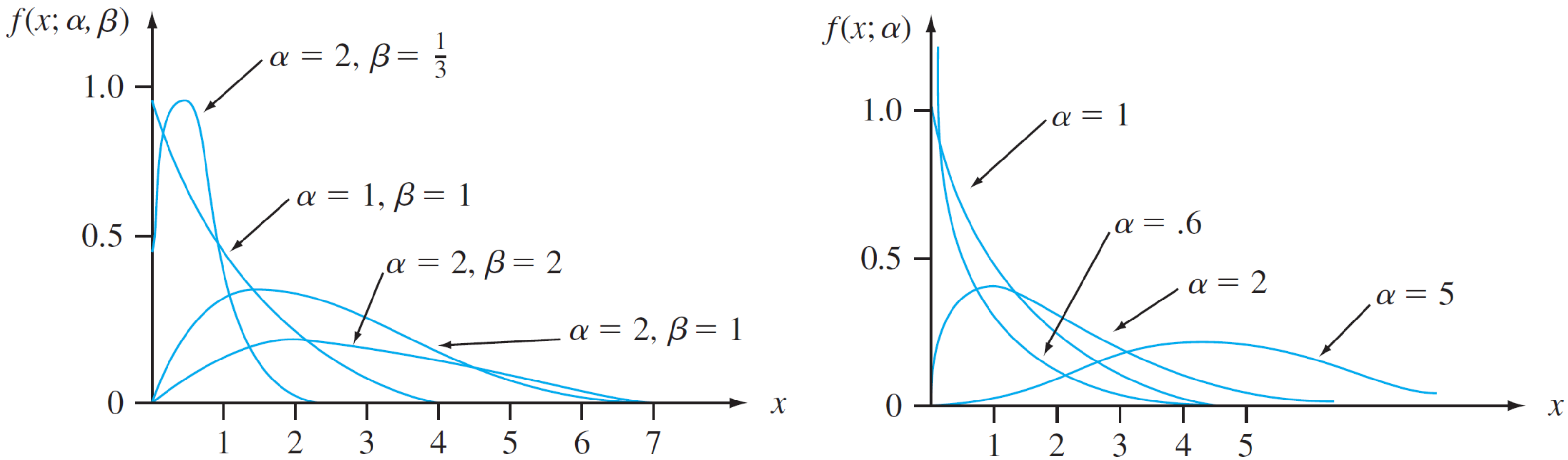
$$P(X \geq 2) = \sum_{x=2}^{5} b(x; 5, 0.2) = 1 - \sum_{x=0}^{1} b(x; 5, 0.2) = 1 - 0.7373 = 0.2627.$$

The Gamma Distribution

- **Definition :** A continuous random variable *X* is said to have a **gamma distribution** if the pdf of *X* is
- $$f(x;\alpha,\beta) = \begin{cases} \dfrac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$
- where the parameters α and β satisfy $\alpha > 0$, $\beta > 0$. The **standard gamma distribution** has $\beta = 1$.

The Gamma Distribution

- The exponential distribution results from taking $\alpha = 1$ and $\beta = 1/\lambda$. Figure (a) illustrates the graphs of the gamma pdf $f(x; \alpha, \beta)$ for several (α, β) pairs, whereas Figure (b) presents graphs of the standard gamma pdf.



Gamma density curves

Figure (a)

standard gamma density curves

Figure (b)

The Gamma Distribution

- For the standard pdf, when $\alpha \leq 1$, $f(x;\ \alpha)$, is strictly decreasing as x increases from 0; when $\alpha > 1$, $f(x;\ \alpha)$ rises from 0 at $x = 0$ to a maximum and then decreases.

- The parameter β is called the *scale parameter* because values other than 1 either stretch or compress the pdf in the x direction.

The Gamma Distribution

- The mean and variance of a random variable X having the gamma distribution $f(x; \alpha, \beta)$ are
- $E(X) = \mu = \alpha\beta$ $\quad V(X) = \sigma^2 = \alpha\beta^2$
- When X is a standard gamma rv, the cdf of X,
-

Application of Gamma Distribution

- Gamma distribution is related the exponential distribution (related to the Poisson process). If the events are occurring according to the Poisson distribution the time till the occurrence of the first event is described by the exponential distribution. The time till the occurrence of the k^{th} event is described by the Gamma distribution.

- Suppose that telephone calls arriving at a particular switchboard follow a Poisson process with an average of 5 calls coming per minute. What is the probability that up to a minute will elapse by the time 2 calls have come in to the switchboard?
- The Poisson process applies, with time until 2 Poisson events following a gamma distribution with *β* = 1/5 and *α* = 2. Denote by *X* the time in minutes that transpires before 2 calls come. The required probability is given by

$$P(X \leq 1) = \int_0^1 \frac{1}{\beta^2} x e^{-x/\beta}\, dx = 25 \int_0^1 x e^{-5x}\, dx = 1 - e^{-5}(1+5) = 0.96.$$

Chi-Squared Distribution

- Another very important special case of the gamma distribution is obtained by letting $\alpha = v/2$ and $\beta = 2$, where v is a positive integer. The result is called **the chi-squared distribution**. The distribution has a single parameter, v, called the **degrees of freedom**.

$$f(x; v) = \begin{cases} \frac{1}{2^{v/2}\Gamma(v/2)} x^{v/2-1} e^{-x/2}, & x > 0, \\ 0, & \text{elsewhere}, \end{cases}$$

$$\mu = v \text{ and } \sigma^2 = 2v.$$