#1. Find the proportion of observations from a standard normal distribution that satisfies each of the following statements.

a. $Z < 2.85$

b. $Z > 2.85$

c. $Z > -1.66$

d. $-1.66 < Z < 2.85$

e. $Z < -2.25$

f. $Z > -2.25$

g. $Z > 1.77$

h. $-2.25 < Z < 1.77$

Art of Stat Solution for #1a ONLY (solutions for #1b-#1h are similar)

#1a. $P(Z < 2.85) = .9978$
**R Solution**

```r
# 1a) Calculate Pr[Normal(mean=0, sd=1) <= 2.85]
pnorm(2.85)
## [1] 0.997814

# 1b) Calculate Pr[Normal(mean=0, sd=1) > 2.85]
pnorm(2.85, lower.tail=FALSE)
## [1] 0.002185961

# 1c) Calculate Pr[Normal(mean=0, sd=1) > -1.66]
pnorm(-1.66, lower.tail=FALSE)
## [1] 0.9515428

# 1d) Calculate Pr[-1.66 <= Normal(mean=0, sd=1) <= 2.85]
pnorm(2.85) - pnorm(-1.66)
## [1] 0.9493568

# 1e) Calculate Pr[Normal(mean=0, sd=1) <= -2.25]
pnorm(-2.25)
## [1] 0.01222447

# 1f) Calculate Pr[Normal(mean=0, sd=1) > -2.25]
1 - pnorm(-2.25)
## [1] 0.9877755

# 1g) Calculate Pr[Normal(mean=0, sd=1) > 1.77]
pnorm(1.77, lower.tail=FALSE)
## [1] 0.03836357

# 1h) Calculate Pr[-2.25 <= Normal(mean=0, sd=1) <= 1.77]
pnorm(1.77) - pnorm(-2.25)
## [1] 0.949412
```

**Stata Solution**

```stata
. * 1a) Pr[Z < 2.85]
display "Pr[ Normal(0,1) < 2.85 ] = " normal(2.85)
Pr[ Normal(0,1) < 2.85 ] = .99781404

. * 1b) Pr[Z > 2.85]
display "Pr[ Normal(0,1) > 2.85 ] = " 1 - normal(2.85)
Pr[ Normal(0,1) > 2.85 ] = .00218596

. * 1c Pr[Z > -1.66]
display "Pr[ Normal(0,1) > -1.66 ] = " 1 - normal(-1.66)
Pr[ Normal(0,1) > -1.66 ] = .95154277

. * 1d) Pr[ -1.66 < Z < 2.85]
display "Pr[ -1.66 < Normal(0,1) < 2.85 ] = " normal(2.85) - normal(-1.66)
Pr[ -1.66 < Normal(0,1) < 2.85 ] = .94935681
```

sol_normal.docx
* 1e) $\Pr[ Z < -2.25 ]$
   display "Pr[ Normal(0,1) < -2.25 ] = " normal(-2.25)
   $\Pr[ \text{Normal}(0,1) < -2.25 ] = .01222447$

* 1f) $\Pr[ Z > -2.25 ]$
   display "Pr[ Normal(0,1) > -2.25 ] = " 1 - normal(-2.25)
   $\Pr[ \text{Normal}(0,1) > -2.25 ] = .98777553$

* 1g) $\Pr[ Z > 1.77 ]$
   display "Pr[ Normal(0,1) > 1.77 ] = " 1 - normal(1.77)
   $\Pr[ \text{Normal}(0,1) > 1.77 ] = .03836357$

* 1h) $\Pr[ -2.25 < Z < 1.77 ]$
   display "Pr[ -2.25 < Normal(0,1) < 1.77 ] = " normal(1.77) - normal(-2.25)
   $\Pr[ -2.25 < \text{Normal}(0,1) < 1.77 ] = .94941196$
#2. The height, X, of young American women is distributed normal with mean \( \mu = 65.5 \) and standard deviation \( \sigma = 2.5 \) inches. Find the probability of each of the following events

a. \( X < 67 \)

**Art of Stat Solution**

\[ 2a. .7257 \]

\[ pr(X < 67) = pr\left( \frac{X-\mu}{\sigma} < \frac{67-\mu}{\sigma} \right) \]

\[ = pr\left( \frac{67-65.5}{2.5} \right) \]

\[ = pr[Z < .6] \]

\[ = .7257 \]

**R Solution**

\[ \text{# Calculate Pr[Normal(mean=65.5, sd=2.5) < 67]} \]
\[ \text{pnorm(67, mean=65.5, sd=2.5)} \]
\[ \#\# [1] 0.7257469 \]

**Stata Solution**

\[ . * \text{ Pr[ Normal(mean=65.5, sd=2.5) < 67]} \]
\[ . \text{display "Pr[ Normal(mean=65.5, sd=2.5) < 67 ] = " normal((67-65.5)/2.5)} \]
\[ \text{Pr[ Normal(mean=65.5, sd=2.5) < 67 ] = .72574688} \]
b. $64 < X < 67$

**Art of Stat Solution**

2b. \[.4515\]

\[
pr(64 < X < 67) = pr\left(\frac{64 - 65.5}{2.5} < Z < \frac{67 - 65.5}{2.5}\right)
\]

= pr[-0.6 < Z < +0.6]

= .4515

---

**R Solution**

```r
# Calculate Pr[64 < Normal(mean=65.5, sd=2.5) < 67]
pnorm(67, mean=65.5, sd=2.5) - pnorm(64, mean=65.5, sd=2.5)
## [1] 0.4514938
```

---

**Stata Solution**

```
. * Pr[ 64 < Normal(mean=65.5, sd=2.5) < 67]
. display "Pr[ 64 < Normal(mean=65.5, sd=2.5) < 67 ] = " normal((67-65.5)/2.5) - normal((64-65.5)/2.5)
Pr[ 64 < Normal(mean=65.5, sd=2.5) < 67 ] = .45149376
```

---
#3. Suppose that, in a certain population, the distribution of GRE scores is normal with mean $\mu=600$ and standard deviation $\sigma=80$.

a. What is the probability of a score less than 450 or greater than 750?

**Art of Stat Solution**

**Answer**: .0608

**Solution:**

Define the random variable $X = \text{GRE score}$.

Thus, $X$ is distributed normal with mean $\mu=600$ and standard deviation $\sigma=80$.

We write this more compactly as $X \sim \text{Normal (} \mu=600, \sigma=80 \text{)}$.

$$
\text{Probability \{ score } < 450 \text{ OR score } > 750 \} =
\text{pr}[X < 450] + \text{pr}[X > 750]
= 1 - \text{pr}[450 < X < 750]
= 1 - .9392
= .0608
$$

**R Solution**

```r
# Calculate Pr[ Normal(mean=600, sd=80) < 450 ] + Pr[ Normal(mean=600, sd=80) > 750 ]
pnorm(450, mean=600, sd=80) + pnorm(750, mean=600, sd=80, lower.tail=FALSE)
## [1] 0.06079272
```

**Stata Solution**

```stata
. * Pr[ Normal(mean=600, sd=80) < 450 OR > 750 ] = 1 - Pr[ 450 < Normal(mean=600, sd=80) < 750 ]
. display "Pr[ Normal(mean=600, sd=80) < 450 OR > 750 ] = " 1 - (normal((750-600)/80) - normal((450-600)/80))
Pr[ Normal(mean=600, sd=80) < 450 OR > 750 ] = .06079272
```

sol_normal.docx
b. What proportion of students has scores between 450 and 750?

**Art of Stat Solution**

**Answer: .9392**

**Solution:** “Proportion” of students with scores between 450 and 750 → we want:

\[ \text{Pr}[450 < X < 750] \]

= .9392

---

**R Solution**

```r
# Calculate Pr[ 450 < Normal(mean=600, sd=80) < 750 ]
pnorm(750, mean=600, sd=80) - pnorm(450, mean=600, sd=80)
## [1] 0.9392073
```

**Stata Solution**

```stata
. * Pr[ 450 < Normal(mean=600, sd=80) < 750 ]
. display "Pr[ 450 < Normal(mean=600, sd=80) < 750 ] = " normal((750-600)/80) - normal((450-600)/80)
Pr[ 450 < Normal(mean=600, sd=80) < 750 ] = .93920728
```

---

sol_normal.docx
#3 – Continued. The Chapin Social Insight Test evaluates how accurately the subject appraises other people. In the reference population used to develop the test, Chapin Social Insight Test scores are distributed normal with mean $\mu=25$ and standard deviation $\sigma=5$.

c. What proportion of the population has scores below 20 on the Chapin test?

**Art of Stat Solution**

**Answer:** .1587

**Solution:**

The solution for the “proportion of the population” is a probability calculation.

Define the random variable $X = \text{Chapin Social Insight Test Score}$.

$X$ is distributed Normal ($\mu=25$, $\sigma=5$).

Want: $\Pr(X < 20) = .1587$

**R Solution**

```r
# Calculate Pr[ Normal(mean=25, sd=5) < 20]  
pnorm(20, mean=25, sd=5)  
## [1] 0.1586553
```

**Stata Solution**

```stata
. * Pr[ Normal(mean=25, sd=5) < 20]  
. display "Pr[ Normal(mean=25, sd=5) < 20] = " normal((20-25)/5)  
Pr[ Normal(mean=25, sd=5) < 20] = .15865525
```

sol_normal.docx
d. What proportion has scores below 10?

Art of Stat Solution

Answer: .0013

Solution:
This is similar to “a”. →
The solution for the “proportion of the population” is a probability calculation.

X is distributed Normal (μ=25, σ=5).
Want: pr(X < 10) = .0013

The Normal Distribution

![Normal Distribution Graph]

R Solution

```
# Calculate Pr[ Normal(mean=25, sd=5) < 10]
pnorm(10, mean=25, sd=5)
## [1] 0.001349898
```

Stata Solution

```
. * Pr[ Normal(mean=25, sd=5) < 10]
. display "Pr[ Normal(mean=25, sd=5) < 10] = " normal((10-25)/5)
Pr[ Normal(mean=25, sd=5) < 10] = .0013499
```
#4. Consider again the setting in questions #3a and #3b: in a certain population, the distribution of GRE scores is normal with mean $\mu = 600$ and standard deviation $\sigma = 80$.

a. What score is equal to the 95th percentile?

**Art of Stat Solution**

**Answer:** 731.6

There are at least two solutions to this question:

**Solution I** – Simple “plug in” variety

**Solution II** – 2 step solution that re-enforces the concepts.

**Step 1:** Obtain the 95$^{th}$ percentile for $Z \sim $ Normal(0,1). Call this $Z_{.95}$

**Step 2:** Use $Z_{.95}$ and the formula on page 26 of the course notes to obtain $X_{.95}$

**Solution I:** Set mean=600 and standard deviation = 80

**Solution II Step I:** Set mean=0 and standard deviation = 1 and then solve for the percentile of X

---

**sol_normal.docx**
Solution II Step 2:
Use the formula on page 26 of the unit 7 notes with the following inputs: 
(1) $Z_{0.95} = 1.645$  
(2) $\mu = 600$ and $\sigma = 80$

$$X_{0.95} = \sigma Z_{0.95} + \mu$$
$$= (80)[1.645] + 600$$
$$= 731.6$$

R Solution

```r
# Calculate 95th percentile of a Normal(mean=600, sd=80)
round(qnorm(.95, mean=600, sd=80), digits=2)
## [1] 731.59
```

Stata Solution

```stata
* 95th percentile of Normal(mean=600, sd=80)
. display " [ 95th percentile of Normal(mean=600, sd=80) ] = " 80*invnormal(.95) + 600
[ 95th percentile of Normal(mean=600, sd=80) ] = 731.58829
```
#4 – Continued. Next, consider again the setting of questions #3c and #3d: The Chapin Social Insight Test evaluates how accurately the subject appraises other people. In the reference population used to develop the test, Chapin Social Insight Test scores are distributed normal with mean $\mu=25$ and standard deviation $\sigma=5$.

b. How high a score must you have in order to be in the top quarter of the population in social insight?

**Art of Stat Solution**

**Answer:** 28.37

**Solution:**
Hone your translation skills here. To be in the “top quarter” your score must be $\geq 75^{th}$ percentile.

Normal Distribution with $\mu = 25$ and $\sigma = 5$

$P(X > 28.372) = 25\%$

Of course you can always just do the brute force right tail probability = .25 to get the same answer!!

Normal Distribution with $\mu = 25$ and $\sigma = 5$

$P(X > 28.372) = 25\%$
**R Solution**

```
# Calculate 75th percentile of a Normal(mean=25, sd=5)
round(qnorm(.75,mean=25,sd=5),digits=2)
## [1] 28.37
```

**Stata Solution**

```
. * 75th percentile of Normal(mean=25, sd=5)
. display " [ 75th percentile of Normal(mean=25, sd=5) ] = " 5*invnormal(.75) + 25
[ 75th percentile of Normal(mean=25, sd=5) ] = 28.372449
```
5. A normal distribution has mean $\mu=100$ and standard deviation $\sigma=15$ (for example, IQ). Give limits, symmetric about the mean, within which 95% of the population would lie:

Solution:
This exercise is asking you to work with the following characteristic of the Normal distribution:

If $X_1, X_2, \ldots, X_n$ are a simple random sample, each distributed $\text{Normal}(\mu, \sigma^2)$
Then the sample mean of $n$ observations is distributed $\text{Normal}(\mu, \sigma^2/n)$

Tip!
It is necessary to input the value of $\sqrt{\sigma^2/n}$ in the box “standard deviation”

a) Individual observations

Answer: 70.6, 129.4

```
“Individual observations” →
“mean” = $\mu=100$
“standard deviation” = $\sigma = 15$.
```

R Solution

```
# Individual Observations: 2.5th and 97.5th percentiles of Normal(mean=100, sd=15)
paste(round(qnorm(.025,mean=100,sd=15),digits=2)," and ",round(qnorm(.975,mean=100,sd=15),digits=2))
## [1] "70.6 and 129.4"
```
Stata Solution

* 2.5th percentile of Normal(mean=100, sd=15)
display " [ 2.5th percentile of Normal(mean=100, sd=15) ] = " 15*invnormal(.025) + 100
[ 2.5th percentile of Normal(mean=100, sd=15) ] = 70.60054

display " [ 97.5th percentile of Normal(mean=100, sd=15) ] = " 15*invnormal(.975) + 100
[ 97.5th percentile of Normal(mean=100, sd=15) ] = 129.39946

b) Means of 4 observations

Answer: 85.3, 114.7

"Means of 4 observations" \rightarrow
"mean" = \mu = 100
"standard deviation" = SE = \sqrt{\sigma^2/n} = \sigma/\sqrt{4} = 15/2 = 7.5

R Solution

# Means of n=4 Observations: 2.5th and 97.5th percentiles of Normal(mean=100, sd=7.5)
paste(round(qnorm(.025,mean=100,sd=7.5),digits=2)," and ",round(qnorm(.975,mean=100,sd=7.5),digits=2))
## [1] "85.3 and 114.7"

Stata Solution

* Means of 4 Observations: 2.5th and 97.5th percentiles of Normal(mean=100, sd=7.5)
display " [ 2.5th percentile of Normal(mean=100, sd=7.5) ] = " 7.5*invnormal(.025) + 100
[ 2.5th percentile of Normal(mean=100, sd=7.5) ] = 85.30027

display " [ 97.5th percentile of Normal(mean=100, sd=7.5) ] = " 7.5*invnormal(.975) + 100
[ 97.5th percentile of Normal(mean=100, sd=7.5) ] = 114.69973
c) Means of 16 observations

**Answer:** 92.65, 107.35

"Means of 16 observations” →

“mean” = µ=100

“standard deviation” = SE = √(σ²/n) = σ/√16 = 15/4 = 3.75

---

### R Solution

```r
# Means of n=16 Observations: 2.5th and 97.5th percentiles of Normal(mean=100, sd=3.75)
paste(round(qnorm(.025,mean=100,sd=3.75),digits=2)," and ",round(qnorm(.975,mean=100,sd=3.75),digits=2))
## [1] "92.65 and 107.35"
```

---

### Stata Solution

```stata
. * Means of 16 Observations: 2.5th and 97.5th percentiles of Normal(mean=100, sd=3.75)
. display " [ 2.5th percentile of Normal(mean=100, sd=3.75) ] = " 3.75*invnormal(.025) + 100
[ 2.5th percentile of Normal(mean=100, sd=3.75) ] = 92.650135
. display " [ 97.5th percentile of Normal(mean=100, sd=3.75) ] = " 3.75*invnormal(.975) + 100
[ 97.5th percentile of Normal(mean=100, sd=3.75) ] = 107.34986
```

---

Sol_normal.docx
d) Means of 100 observations

Answer: 97.06, 102.94

“Means of 100 observations” →
“mean” = µ = 100
“standard deviation” = SE = \sqrt{\sigma^2/n} = \sigma/\sqrt{100} = 15/10 = 1.5

R Solution

```
# Means of n=100 Observations: 2.5th and 97.5th percentiles of Normal(mean=100, sd=1.5)
paste(round(qnorm(.025,mean=100,sd=1.5),digits=2)," and ",round(qnorm(.975,mean=100,sd=1.5),digits=2))
## [1] "97.06 and 102.94"
```

Stata Solution

```
. * Sd) Means of 100 Observations: 2.5th and 97.5th percentiles of Normal(mean=100, sd=1.5)
. display " [ 2.5th percentile of Normal(mean=100, sd=1.5) ] = " 1.5*invnormal(.025) + 100
   [ 2.5th percentile of Normal(mean=100, sd=1.5) ] = 97.060054
. display " [ 97.5th percentile of Normal(mean=100, sd=1.5) ] = " 1.5*invnormal(.975) + 100
   [ 97.5th percentile of Normal(mean=100, sd=1.5) ] = 102.93995
```