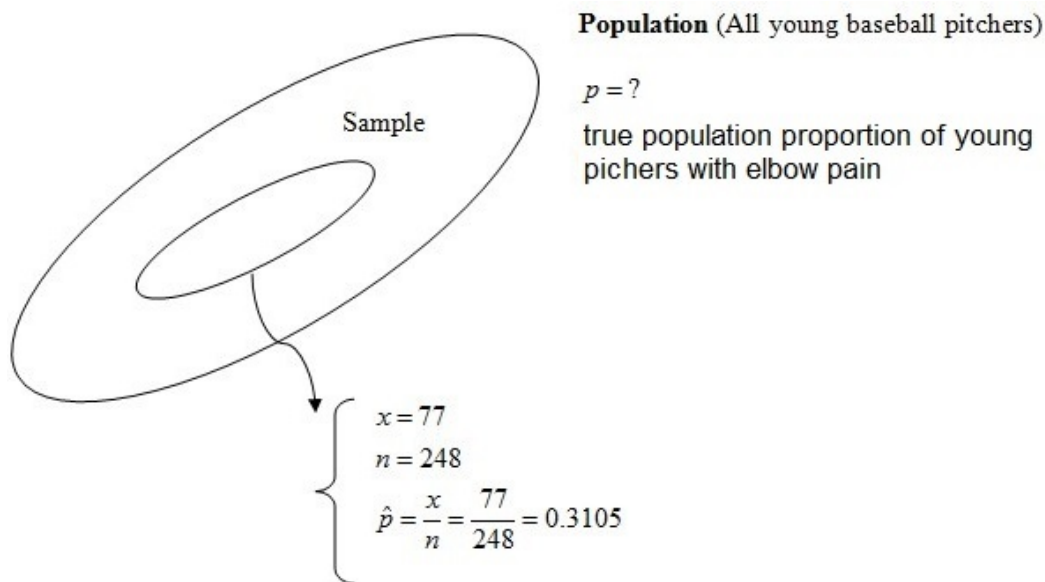


### 1.1 C.I. AROUND ONE POPULATION PROPORTION ( $p$ ) [using sample statistics]:

**Example:** In a 1999-2000 longitudinal study of youth baseball, researchers found that 77 of 248 young pitchers complained of elbow pain after pitching. What is the 90% C.I. for the true population proportion of young pitchers with elbow pain?



#### Solution by hand:

Since the CL is 90% then  $Z_{\alpha/2} = \text{QUANTILE}('Normal', (1 - 0.90)/2) \approx -1.645$  [alternatively,  $Z_{\alpha/2} = \text{probit}((1 - 0.90)/2) \approx -1.645$ ]. Thus,

$$\hat{p} \pm Z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.3105 \pm 1.645 \sqrt{\frac{0.3105(1-0.3105)}{248}} = (0.2621676, 0.3588324) \approx (0.2622, 0.3588).$$

We are 90% confident that the true population proportion of young pitchers with elbow pain is between 26.22% and 35.88%.

#### Solution by SAS:

```

data C1;
input outcome $ count;
cards;
Pain 77
Nopain 171
run;

ods select BinomialCLs;
proc freq data=C1;
table outcome/binomial (level='Pain' CL=WALD ) alpha=0.10;
WEIGHT COUNT;
RUN;

```

Confidence Limits for the Binomial Proportion		
outcome = Pain		
Proportion = 0.3105		
Type	90% Confidence Limits	
Wald	0.2622	0.3588

Alternatively,

```

data C1;
x=77;
n=248;
alpha=0.10;
p=x/n;
z=probit(1-alpha/2);
se=(sqrt(n*p*(1-p)))/n;
L=p-z*se;
U=p+z*se;
RUN;

PROC PRINT DATA=C1;
VAR P L U;
format _numeric_ 9.4;
RUN;

```

Obs	p	L	U
1	0.3105	0.2622	0.3588

## 1.2 C.I. AROUND ONE POPULATION PROPORTION ( $p$ ) [using sample data]:

**Example:** Using the Diabetes dataset from Lab1, what is the 95% confidence interval around the true population proportion of diabetes?

**Solution by SAS:**

```

proc freq data=biom505.diabetesfall17;
table diab/binomial (level='1');
run;

```

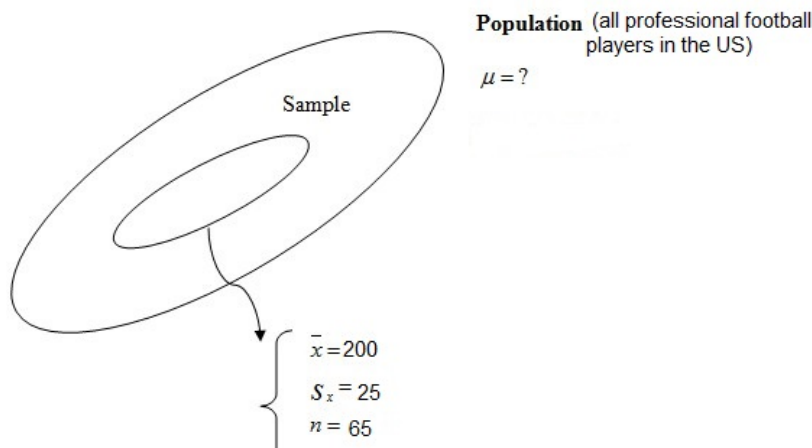
diab				
diab	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	330	84.62	330	84.62
1	60	15.38	390	100.00
Frequency Missing = 13				

Binomial Proportion	
diab = 1	
Proportion	0.1538
ASE	0.0183
95% Lower Conf Limit	0.1180
95% Upper Conf Limit	0.1897
Exact Conf Limits	
95% Lower Conf Limit	0.1195
95% Upper Conf Limit	0.1935

We are 90% confident that the true population proportion of diabetes is between 11.80% and 18.96%.

## 2.1 C.I. AROUND ONE POPULATION MEAN ( $\mu$ ) when $\sigma$ is unknown [using sample statistics]:

**Example:** A study from the US of 65 professional football players found that the weight of football players is roughly normally distributed with a sample mean of 200 pounds and a standard deviation of 25 pounds. What is the 99% C.I. for the true population mean weight of US football players?



**Solution by hand:** Since the CL is 99% then  $t_{n-1, \alpha/2} = \text{tinv}((1 - 0.99)/2, 64) \approx -2.655$ . Thus,

$$\bar{x} \pm t_{n-1, \alpha/2} \left( \frac{s_x}{\sqrt{n}} \right) = 200 \pm 2.655 \left( \frac{25}{\sqrt{65}} \right) = (191.7672, 208.2328) \approx (191.77, 208.23).$$

**Solution by SAS:**

```
data t;
xbar=200;
s=25;
n=65;
alpha=1-0.99;
se=s/sqrt(n);
L=xbar-tinv(1 - alpha/2, 64)*se;
U=xbar+tinv(1 - alpha/2, 64)*se;
run;
```

```
PROC PRINT DATA=t;
VAR xbar L U;
format _numeric_ 9.2;
RUN;
```

Obs	xbar	L	U
1	200.00	191.77	208.23

We are 99% confident that the true population mean weight of US professional football players is between 191.77 and 208.23 pounds.

## 2.2 C.I. AROUND ONE POPULATION MEAN ( $\mu$ ) when $\sigma$ is unknown [using sample data]:

**Example:** Using the Diabetes dataset from Lab1, what is the 95% confidence interval around the true mean age of subjects in the population of African Americans in Virginia?

**Solution by SAS:**

```
proc means data=biom505.diabetesfall17 maxdec=2 alpha=0.05 mean clm;
var age;
run;
```

Analysis Variable : age age		
Mean	Lower 95% CL for Mean	Upper 95% CL for Mean
46.85	45.25	48.45

Alternatively,

```
ods select BasicIntervals;
proc univariate data=biom505.diabetesfall17 cibasic;
var age;
run;
```

Basic Confidence Limits Assuming Normality			
Parameter	Estimate	95% Confidence Limits	
Mean	46.85112	45.25369	48.44854
Std Deviation	16.31233	15.25853	17.52368
Variance	266.09221	232.82288	307.07945

We are 95% confident that the true mean age of subjects in the population of African Americans in Virginia is between 45.25 and 48.45 pounds.

## 2.3 C.I. AROUND ONE POPULATION MEAN ( $\mu$ ) when $\sigma$ is unknown by stratum (group) [using sample data]:

**Example:** Using the Diabetes dataset from Lab1, what is the 95% confidence interval around the true mean age of Female subjects in the population of African Americans in Virginia?

**Solution by SAS:**

```
proc means data=biom505.diabetesfall17 maxdec=2 alpha=0.05 mean clm;
CLASS GENDER;
var age;
run;
```

Analysis Variable : age age				
gender	N Obs	Mean	Lower 95% CL for Mean	Upper 95% CL for Mean
female	234	45.83	43.69	47.97
male	169	48.26	45.85	50.67

Alternatively,

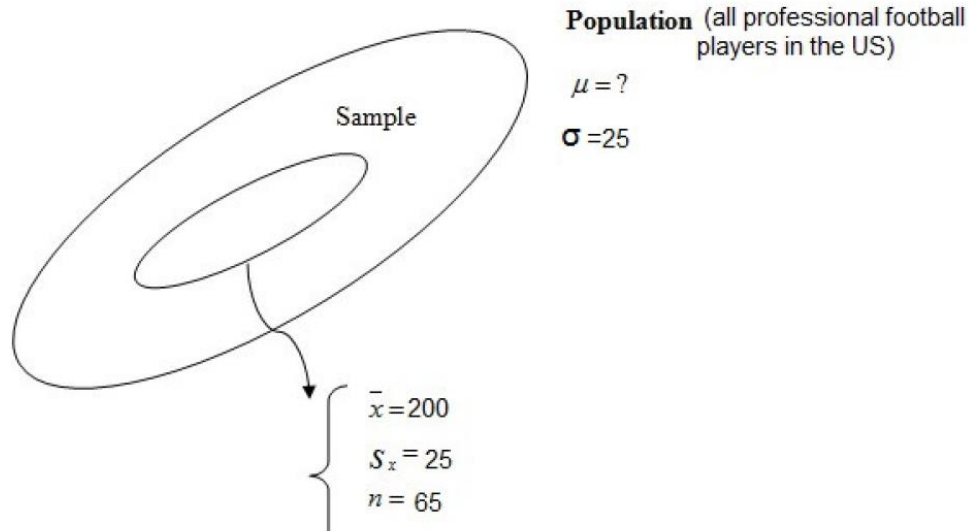
```
ods select BasicIntervals;
proc univariate data=biom505.diabetesfall17 cibasic;
CLASS GENDER;
var age;
run;
```

The UNIVARIATE Procedure Variable: age (age) gender = female				The UNIVARIATE Procedure Variable: age (age) gender = male			
Basic Confidence Limits Assuming Normality				Basic Confidence Limits Assuming Normality			
Parameter	Estimate	95% Confidence Limits		Parameter	Estimate	95% Confidence Limits	
Mean	45.83333	43.69471	47.97196	Mean	48.26036	45.85499	50.66572
Std Deviation	16.60478	15.22430	18.26274	Std Deviation	15.83932	14.31158	17.73509
Variance	275.71888	231.77919	333.52779	Variance	250.88419	204.82136	314.53344

We are 95% confident that the true mean age of Females in the population of African Americans in Virginia is between 43.69 and 47.97 pounds.

### 3.1 C.I. AROUND ONE POPULATION MEAN ( $\mu$ ) when $\sigma$ is known [using sample statistics]:

**Example:** A study from the US of 65 professional football players found that the weight of football players is roughly normally distributed with a sample mean of 200 pounds. Assume that the population standard deviation for the weights is known and equals to 25 pounds. What is the 99% C.I. for the true population mean weight of US football players?



**Solution by hand:** Since the CL is 99% then  $Z_{\alpha/2} = \text{probit}((1 - 0.99)/2) \approx -2.576$ . Thus,

$$\bar{x} \pm Z_{\alpha/2} \left( \frac{\sigma_x}{\sqrt{n}} \right) = 200 \pm 2.576 \left( \frac{25}{\sqrt{65}} \right) = (192.0122, 207.9878) \approx (192.01, 207.99).$$

**Solution by SAS:**

```
data Zci;
xbar=200;
sigma=25;
n=65;
alpha=1-0.99;
se=sigma/sqrt(n);
L=xbar-probit(1 - alpha/2)*se;
U=xbar+probit(1 - alpha/2)*se;
run;
```

```
PROC PRINT DATA=Zci;
VAR xbar L U;
format _numeric_ 9.2;
RUN;
```

Obs	xbar	L	U
1	200.00	192.01	207.99

We are 99% confident that the true population mean weight of US professional football players is between 192.01 and 207.99 pounds.