Sampling Distribution

Aihua Li

Sampling:

- Hard to get the whole population
- Probability allows us to take a sample from a population and make inferences to a population.
- Simple random sampling is one of the core concepts to much of data collection and analysis.
  - ✓ In simple random sampling, each individual or object in a population has an equal probability of being selected into the sample.
  - ✓ The first step in simple random sampling is to define the population of concern (often called a sampling frame).
- Other methods of sampling include systematic random sampling, cluster sampling and stratified random sampling.

Objectives

Upon successful completion of this lesson, you should be able to:

- Understand the meaning of sampling distribution.
- Apply the central limit theorem to calculate approximate probabilities for sample means and sample proportions.
- **Describe the sampling distribution of the sample mean and proportion.**
- Identify situations in which the normal distribution and t-distribution may be used to approximate a sampling distribution.

Sampling Distribution

- The sampling distribution of a statistic is a probability distribution based on a large number of samples of size n from a given population.
- The concept of a sampling distribution is perhaps the most basic concept in inferential statistics. It is also a difficult concept because a sampling distribution is a theoretical distribution rather than an empirical distribution.

Sampling Distribution

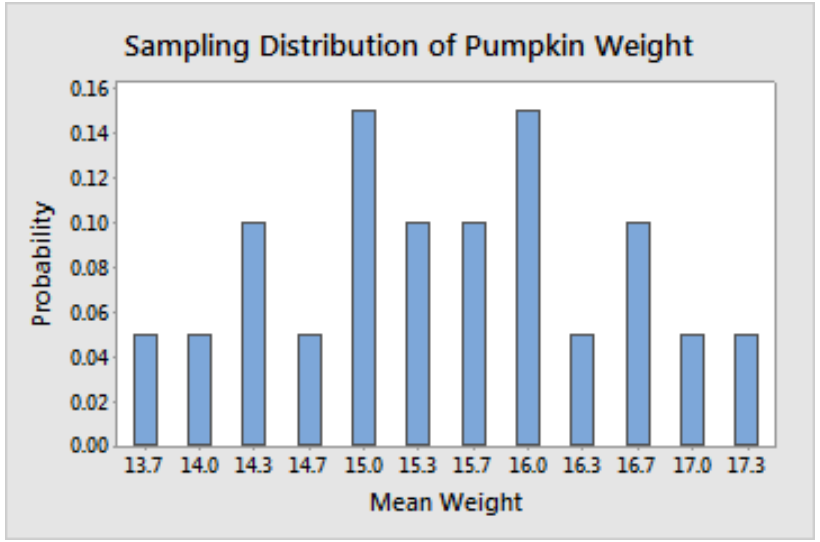
For example, the following table shows the weights of the entire population of 6 pumpkins. The pumpkins can only be one of the weight values listed in the following table.

Pumpkin	1	2	3	4	5	6
Weight	19	14	15	12	16	17

- Take all possible random samples of the population that contain 3 pumpkins (20 random samples).
- Calculate the mean of each sample.

Sample	Weights	Mean Weight	Probability
2, 3, 4	14, 15, 12	13.7	1/20
2, 4, 5	14, 12, 16	14	1/20
2, 4, 6	14, 12, 17	14.3	2/20
3, 4, 5	15, 12, 16		
3, 4, 6	15, 12, 17	14.7	1/20
1, 2, 4	19, 14, 12	15	3/20
2, 3, 5	14, 15, 16		
4, 5, 6	12, 16, 17		
2, 3, 6	14, 15, 17	15.3	2/20
1, 3, 4	19, 15, 12		
1, 4, 5	19, 12, 16	15.7	2/20
2, 5, 6	14, 16, 17		
1, 2, 3	19, 14, 15	16	3/20
3, 5, 6	15, 16, 17		
1, 4, 6	19, 12, 17		
1, 2, 5	19, 14, 16	16.3	1/20
1, 2, 6	19, 14, 17	16.7	2/20
1, 3, 5	19, 15, 16		
1, 3, 6	19, 15, 17	17	1/20
1, 5, 6	19, 16, 17	17.3	1/20

- The sampling distribution of the mean weights is displayed on this graph. The distribution is centered around 15.5, which is also the true value for the population mean.
- The random samples with sample means closer to 15.5 have a greater probability of occurring than the samples with sample means farther away from 15.5.



- Approximate the sampling distribution of the sample statistic: For example, if you sample from the normal population then the sample mean has exactly the normal distribution. But if you sample from a population other than normal population, you may not be able to determine the exact distribution of the sample mean.
- However, because of the central limit theorem, the sample mean is approximately distributed as normal, provided **your samples are large enough**.

An example:

- Question: A large tank of fish from a hatchery is being delivered to the lake. We want to know the average length of the fish in the tank.

- Solution: Randomly sample twenty fish and use the sample mean to estimate the population mean.

  ✓ Denote the sample mean of the twenty fish as $\bar{x}_1$

  ✓ Take a separate sample of size twenty from the same hatchery. Denote that sample mean as $\bar{x}_2$. Would $\bar{x}_1$ equal $\bar{x}_2$?

  ✓ Take 1000 random samples of size twenty and recording all of the sample means and create a histogram.

  ✓The distribution of all of these sample means is the sampling distribution of the sample mean

- We can find the sampling distribution of a any sample statistic that would estimate a certain population parameter of interest.
- We will focus on the sampling distributions for the sample mean $\bar{x}$, and the sample proportion $\hat{p}$

- Sampling Distribution of the Sample Mean
- Sampling Distribution of the Sample Proportion

Sampling Distribution of Means

What is it?

A theoretical probably distribution that represents a statistic (such as a mean) for all possible samples of a given size from a population of interest.

What purpose does it serve?

Helps us determine if our sample statistics are accurate estimates of population parameters.

Imagine drawing all possible samples of a given size and calculating a mean for each of those samples. Then plot all of those means in a frequency distribution, called the sampling distribution of means.

But, only in theory.

The Central Limit Theorem

The **central limit theorem** states that:

- if the size of the sample is large (i.e., $n = 60$), the shape of the sampling distribution of means approximates the shape of the normal distribution.
- the mean of the sampling distribution of means will equal the population mean (μ).
  - This is called the expected value.
- the standard deviation of the sampling distribution is called the standard error of the mean (σ_M):

$$\sigma_M = \frac{\sigma}{\sqrt{n}}$$

  - tells us how much the sample mean (M) is expected to vary from the population mean (μ) due to chance.

As the size of the sample increases, the amount of standard error decreases.

For Example,

- The standard deviations below are identical.
- but the sample sizes are different.

Given: $\sigma = 16$
$n = 25$

$$\sigma_M = \frac{16}{\sqrt{25}} = 3.2$$

Given: $\sigma = 16$
$n = 100$

$$\sigma_M = \frac{16}{\sqrt{100}} = 1.6$$

Thus, larger samples generally produce less error (and therefore greater accuracy) in predicting population means.

Probabilities, Proportions, and Percentages of Sample Means

The sampling distribution of means:

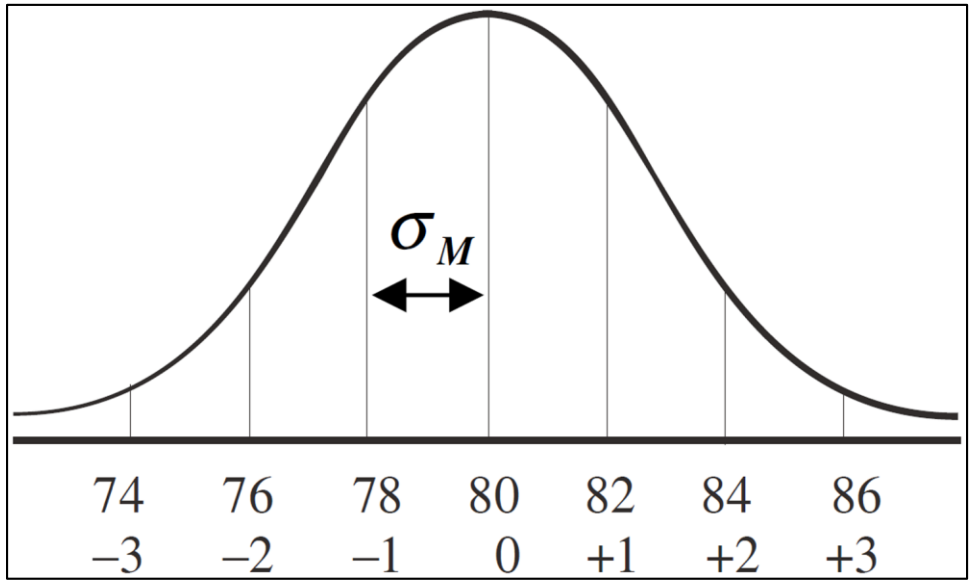
- Is used to determine the probabilities, proportions, and percentages associated with particular sample means.
- Similar to our previous use of the normal distribution curve to determine these figures for particular raw scores.
- But, since we are dealing with sample means rather than raw scores, the formulas involved are modified as follows:

$$z = \frac{M - \mu}{\sigma_M} \qquad \text{and} \qquad M = \mu + (z)(\sigma_M)$$

Population is normal distribution or not distribution

In the sampling distribution below, $\mu = 80$, $\sigma = 14$, and $n = 49$. Thus, the standard error (σ_M) is 2.

$$\sigma_M = \frac{14}{\sqrt{49}} = 2$$



We can now use the sampling distribution of means to answer questions about probabilities, proportions, and percentages.

For Example,

1. In the previous normal distribution with $\mu = 80$ and $\sigma_M = 2$,

   a. What is the probability that the sample mean obtained is below 81?

   - Convert the given sample mean to a z-score:

$$z = \frac{M - \mu}{\sigma_M} = \frac{81 - 80}{2} = +.50$$

   - Then, locate in the *z*-table the probability associated with sample means below a z-score of +.50 (column B):

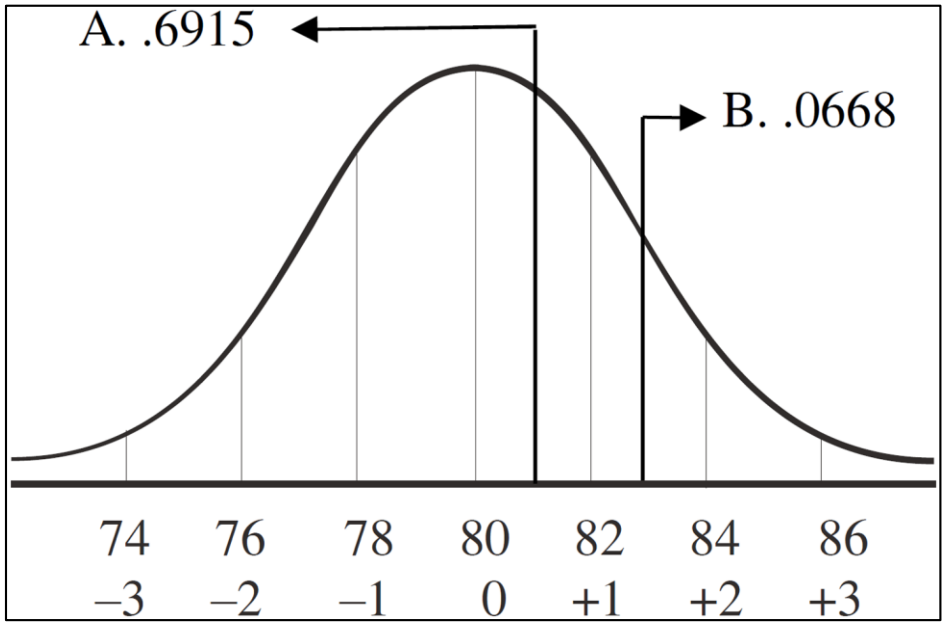
$$p(M < 81) = .6915$$

Notice that the correct notation now specifies a sample mean (*M*) rather than a raw score (*X*).

b. What proportion of the sample means can be expected to have a value greater than 83?

$$z = \frac{M - \mu}{\sigma_M} = \frac{83 - 80}{2} = +1.50$$

$$p(M > 83) = .0668$$



2. Given a normally shaped population distribution with a $\mu = 95$ and $\sigma = 5$, a sample of size $n = 25$ is drawn at random.

a. The probability is .05 that the mean of the sample will be above what value?

- First, find the standard error:
- Next, find the *z*-score associated with a proportion of .0500, which is 1.65.
- Finally, use the formula to convert the *z*-score to a sample mean:

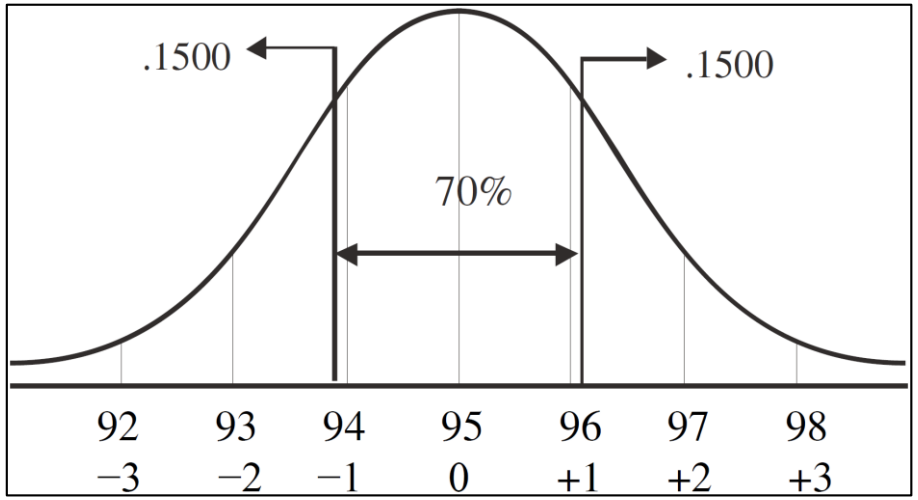
$$\sigma_M = \frac{5}{\sqrt{25}} = 1$$

$$\begin{aligned} M &= \mu + (z)(\sigma_M) \\ &= 95 + (1.65)(1) \\ &= 96.65 \end{aligned}$$

b. What range of sample means would be expected to occur in the middle of the distribution 70% of the time?

- We need to identify the *z*-scores associated with the extreme 30% in the tails (i.e., .1500 at each end) and then use the formula to locate the means at each extreme.
  - The *z*-scores associated with the highest and lowest .1500 extremes of the distribution are ±1.04.

$$M = 95 + (-1.04)(1) = 93.96$$



$$M = 95 + (+1.04)(1) = 96.04$$

Thus, 70% of the time, we can expect to draw sample means that fall within the range of 93.96 and 96.04.

Formula Summary

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

Tells how much a particular raw score deviates from the mean of a population in standard deviation units.

$$z = \frac{M - \mu}{\sigma_M}$$

Tells how much a particular sample mean deviates from the population in standard error units.

$$X = \mu + (z)(\sigma)$$

Use when being asked for a raw score value.

$$M = \mu + (z)(\sigma_M)$$

Use when being asked for the value of a sample mean.

An example: z-Scores

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

- Calculating Z-Scores with the Sampling Distribution of the Sample Means
- An example : let's say we have a production process that produces long-lasting light bulbs. The average lifespan of the bulbs produced is 1,500 hours with a population standard deviation of 300 hours. So…….

  (1) We select 100 light bulbs at random. What is the standard deviation of the sample means?

  (2) What is the probability that one bulb, selected at random, will last longer than 1,800 hours?

  (3) What is the probability that the average of 100 randomly selected bulbs is greater than 1,800 hours?

(1) We select 100 light bulbs at random. What is the standard deviation of the sample means?
(2) What is the probability that one bulb, selected at random, will last longer than 1,800 hours?
(3) What is the probability that the average of 100 randomly selected bulbs is greater than 1,800 hours?

- Solution: μ=1500, $\sigma = 300$

(1) The standard deviation of the sample means equals the known population standard deviation divided by the square root of the sample size (n). Therefore the SD(xbar)= 300/sqrt(100) = 300/10= 30 hours.

(2) Because we are interested in just one bulb (not an average), we use this z-score formula:

Therefore, Z=(1,800 – 1,500)/300 = 300/300 = 1. And we want to know the probability that X>1,800 which translates into Z>1. We look this up in our z table and find that p(Z>1) = 1- 0.8413=0.1587. There is approximately 15.9% chance that one light bulb chosen at random will last longer than 1,800 hours.

(3) Because we are interested in the average, we use the following z-score formula:

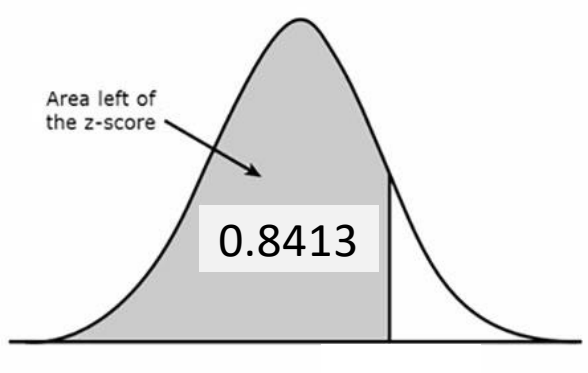
- Therefore, z=(1,800 – 1,500)/(300/sqrt(100))=300/30 = 10. And the probability that z>10 is practically zero.

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

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

A **z**-score **table shows** the percentage of values (usually a decimal figure) to the left of a given **z**-score on a standard normal distribution. The corresponding area is 0.8413 which translates into 84.13% of the standard normal distribution being below (or to the left) of the **z**-score.

$$p(z>1) = 1 - 0.8413 = 0.1587$$



Z=1

http://www.z-table.com/

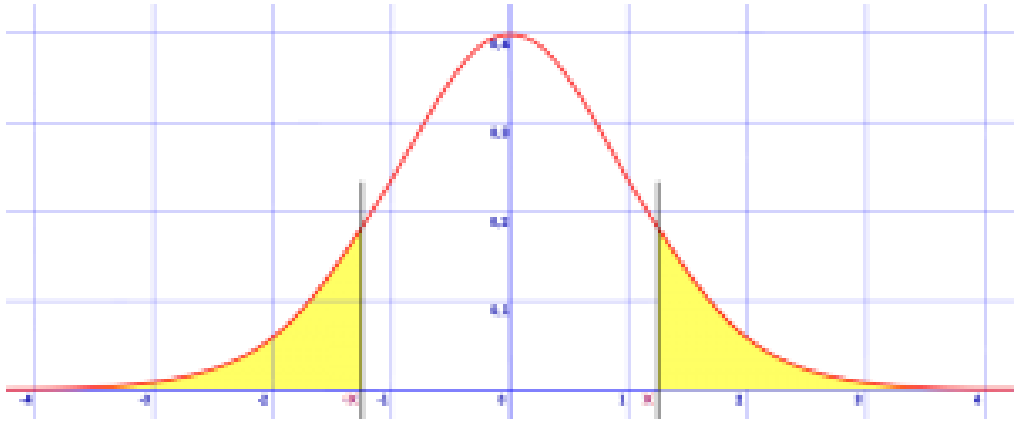
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

How to find the area under a curve (percentage)

z=0.46 and z=-0.46

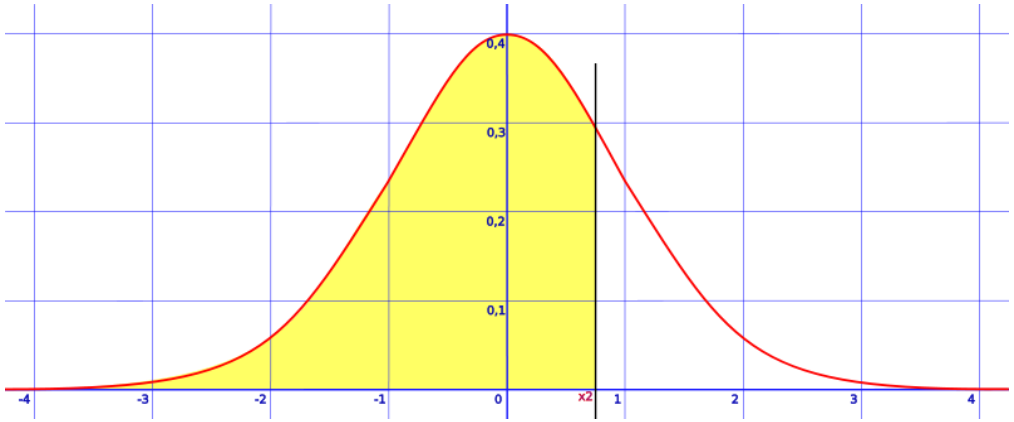
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	[illegible]	
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
–0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
–0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
–0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
–0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
–0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
–0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
–0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641



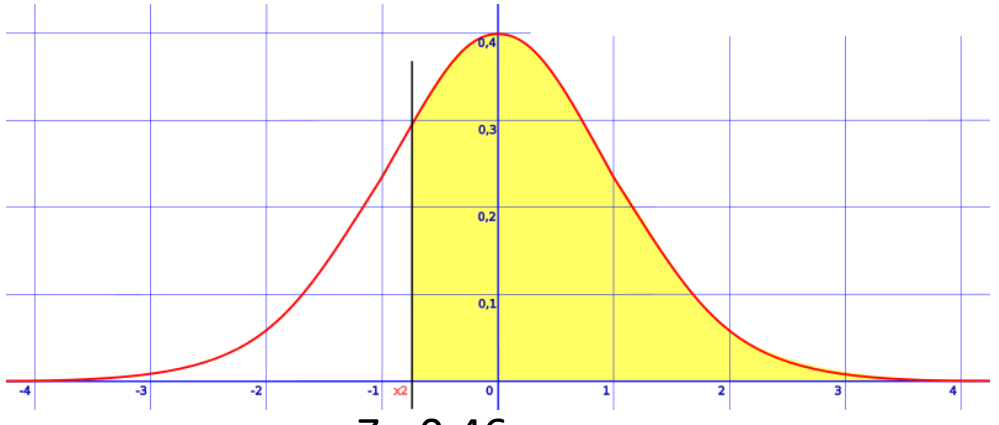
Z=-0.46 Z=0.46

.3228 (left) 1-0.6672=0.3228 (right)



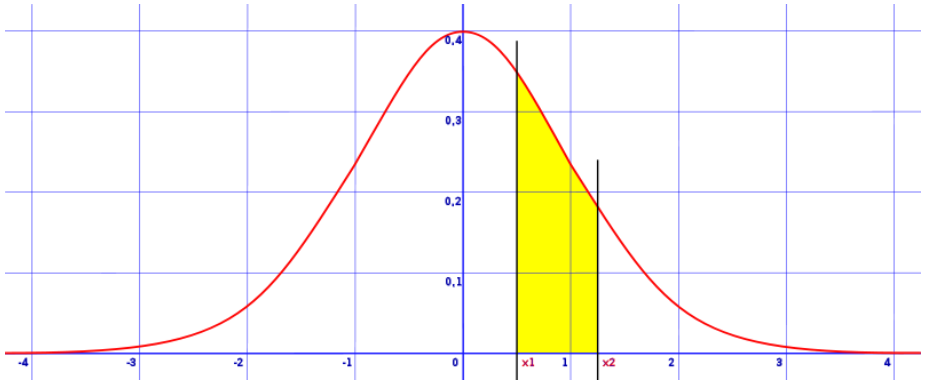
Z=0.46

area to the left of a z score (z is greater than the mean), .6772



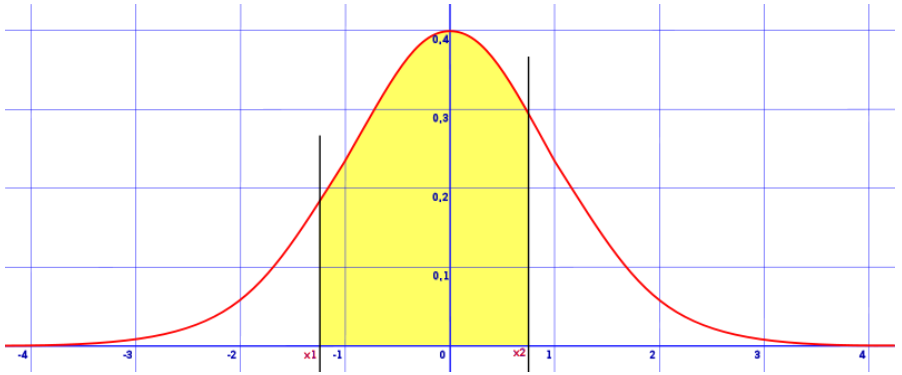
Z=-0.46

area to the right of a z score (z is greater than the mean), 1-0.3228=0.6772



Z=0.40 Z=0.46

find the area from z= 0.46 to z= 0.40,
0.6772- 0.6554=0.0218



Z=-0.46 Z=0.16

area between two z values on opposite sides of mean (find the area between two z values of -0.46 and +1.16):

0.5636-.3228=0.2408

Sampling distribution for the sample mean for a very small population

- [Note: The sampling method is done without replacement.]
- **Sample Means with a Small Population: Pumpkin Weights**

In this example, the population is the weight of six pumpkins (in pounds) displayed in a carnival "guess the weight" game booth. You are asked to guess the average weight of the six pumpkins by taking a random sample without replacement from the population. Since we know the weights from the population, we can find the population mean.

Pumpkin	A	B	C	D	E	F
Weight (in pounds)	19	14	15	9	10	17

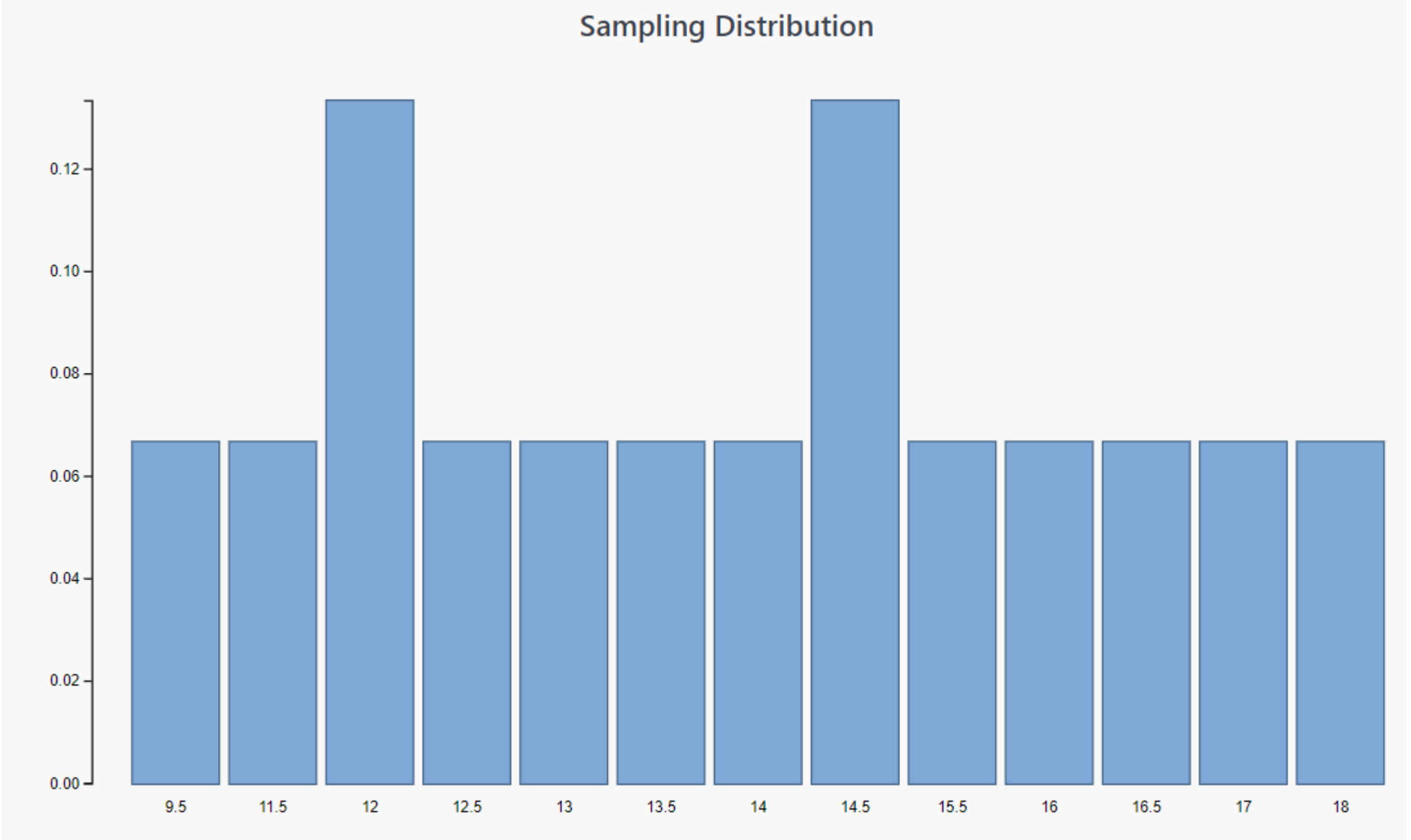
$$\mu = \frac{19 + 14 + 15 + 9 + 10 + 17}{6} = 14 \text{ pounds}$$

- To demonstrate the sampling distribution, let's start with obtaining all of the possible samples of size n=2 from the populations, sampling without replacement. The table below show all the possible samples, the weights for the chosen pumpkins, the sample mean and the probability of obtaining each sample. Since we are drawing at random, each sample will have the same probability of being chosen.

We can combine all of the values and create a table of the possible values and their respective probabilities.

$\bar{x}$	9.5	11.5	12.0	12.5	13.0	13.5	14.0	14.5	15.5	16.0	16.5	17.0	18.0
Probability	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{1}{15}$

The table is the probability table for the sample mean and it is the sampling distribution of the sample mean weights of the pumpkins when the sample size is 2. It is also worth noting that the sum of all the probabilities equals 1. It might be helpful to graph these values.


Sampling Distribution
0.12
0.10
0.08
0.06
0.04
0.02
0.00
9.5
11.5
12
12.5
13
13.5
14
14.5
15.5
16
16.5
17
18

- One can see that the chance that the sample mean is exactly the population mean is only 1 in 15, very small. When using the sample mean to estimate the population mean, some possible error will be involved since the sample mean is random.
- Now that we have the sampling distribution of the sample mean, we can calculate the mean of all the sample means. In other words, we can find the mean (or expected value) of all the possible.
- The mean of the sample means is

$$\mu_{\bar{x}} = \sum \bar{x}_i f(\bar{x}_i) = 9.5\left(\tfrac{1}{15}\right) + 11.5\left(\tfrac{1}{15}\right) + 12\left(\tfrac{2}{15}\right)$$
$$+ 12.5\left(\tfrac{1}{15}\right) + 13\left(\tfrac{1}{15}\right) + 13.5\left(\tfrac{1}{15}\right) + 14\left(\tfrac{1}{15}\right)$$
$$+ 14.5\left(\tfrac{2}{15}\right) + 15.5\left(\tfrac{1}{15}\right) + 16\left(\tfrac{1}{15}\right) + 16.5\left(\tfrac{1}{15}\right)$$
$$+ 17\left(\tfrac{1}{15}\right) + 18\left(\tfrac{1}{15}\right) = 14$$

- Even though each sample may give you an answer involving some error, the expected value is right at the target: exactly the population mean. In other words, if one does the experiment over and over again, the overall average of the sample mean is exactly the population mean.

- Now, let's do the same thing as above but with sample size n=5

Sample	Weights	$\bar{x}$	Probability
A, B, C, D, E	19, 14, 15, 9, 10	13.4	1/6
A, B, C, D, F	19, 14, 15, 9, 17	14.8	1/6
A, B, C, E, F	19, 14, 15, 10, 17	15.0	1/6
A, B, D, E, F	19, 14, 9, 10, 17	13.8	1/6
A, C, D, E, F	19, 15, 9, 10, 17	14.0	1/6
B, C, D, E, F	14, 15, 9, 10, 17	13.0	1/6

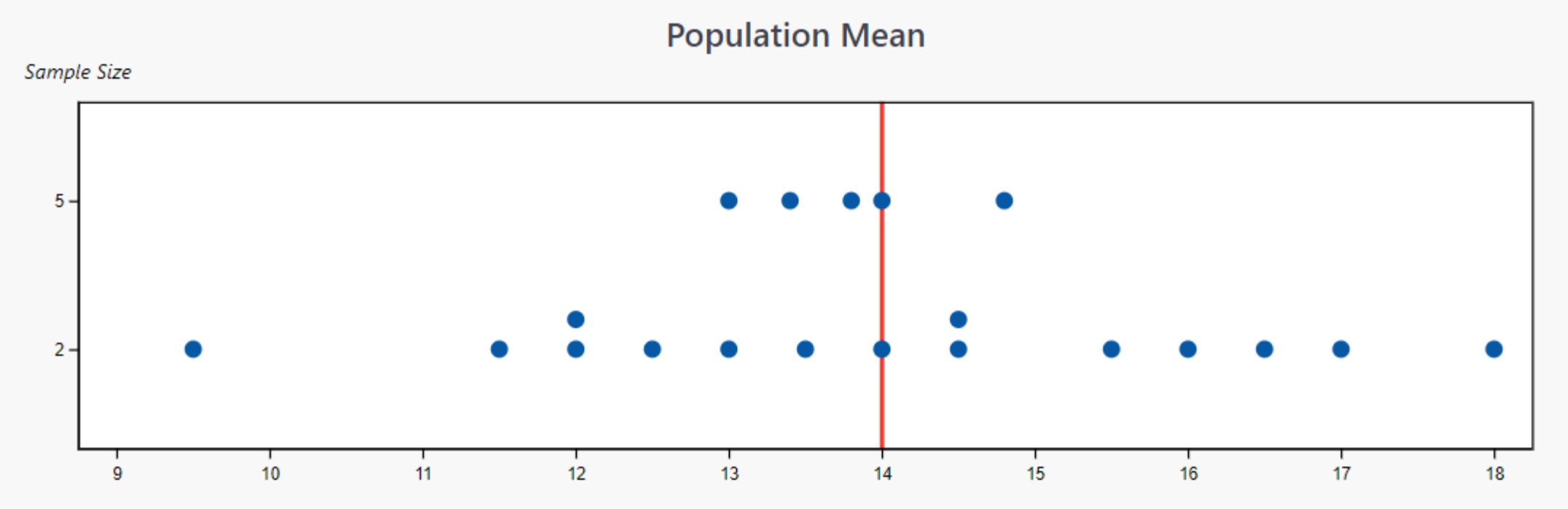
The sampling distribution is:

$\bar{x}$	13.0	13.4	13.8	14.0	14.8	15.0
Probability	1/6	1/6	1/6	1/6	1/6	1/6

The mean of the sample means is...

$$\mu = (\frac{1}{6})(13 + 13.4 + 13.8 + 14.0 + 14.8 + 15.0) = 14 \text{ pounds}$$

- The following dot plots show the distribution of the sample means corresponding to sample sizes of n=2 and of n=5



- Again, we see that using the sample mean to estimate population mean involves sampling error. However, the error with a sample of size is on the average smaller than with a sample of size .
- Sampling Error and Size

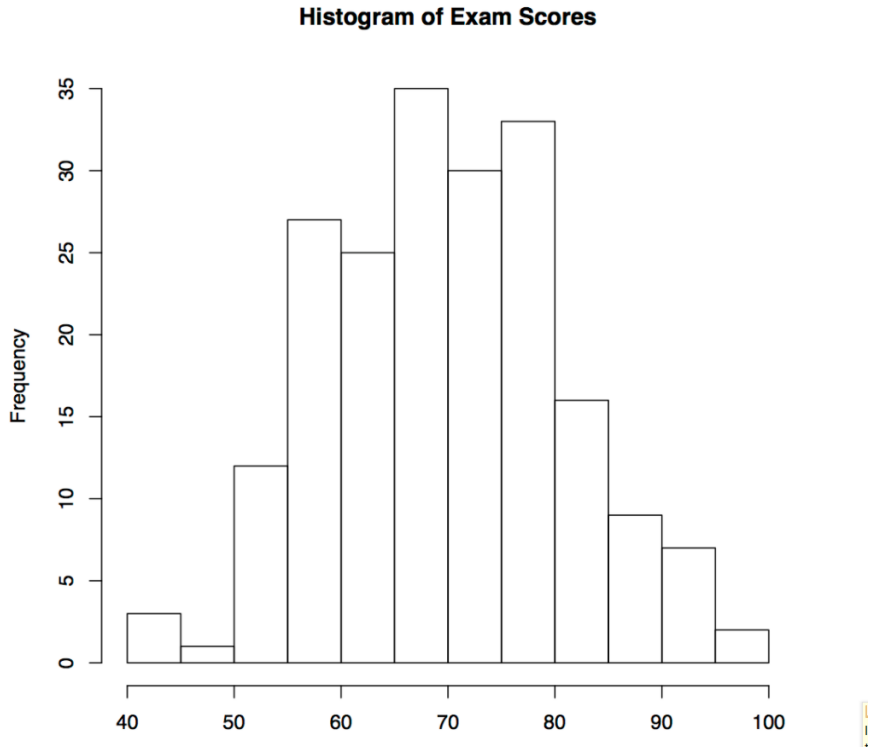
✓The error resulting from using a sample characteristic to estimate a population characteristic.

✓Sample size and sampling error: As the dotplots above show, the possible sample means cluster more closely around the population mean as the sample size increases. Thus, the possible sampling error decreases as sample size increases.

- What happens when the population is not small, as in the pumpkin example?

Sampling distribution for the sample mean for large population

- An instructor of an introduction to statistics course has 200 students. The scores out of 100 points are shown in the histogram.



- Let's demonstrate the sampling distribution of the sample means using the **StatKey website**

Inbox - ivale | Mail - Li, Aih | Google Trans | Combination | Z-table (Rig | z table show | Z Score Tabl | Simple rand | 4.1 - Sampli | Sampling Di

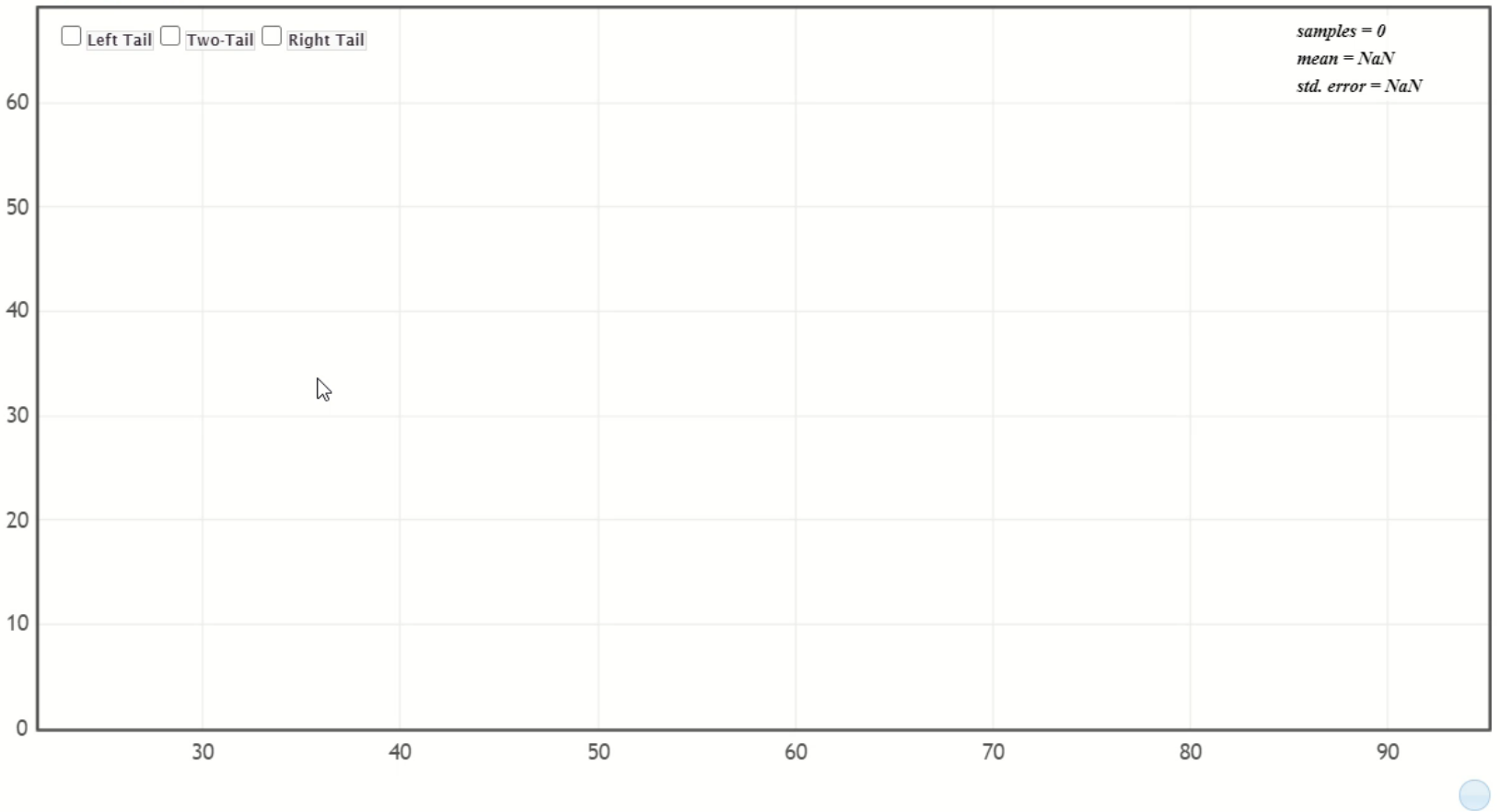
Not secure | lock5stat.com/StatKey/sampling_1_quant/sampling_1_quant.html

StatKey Sampling Distribution for a Mean

Statistics Grad Schools | Show Data Table | Edit Data | Choose samples of size n = 10 | Upload File | Change Column(s)

Generate 1 Sample | Generate 10 Samples | Generate 100 Samples | Generate 1000 Samples | Reset Plot

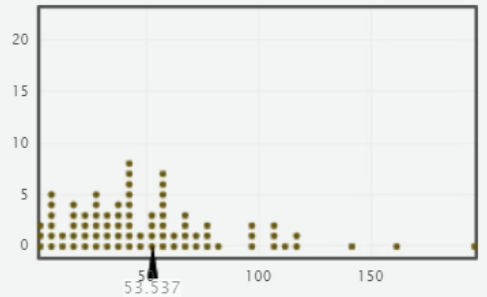
Sampling Dotplot of Mean



Data Plots | Confidence Intervals

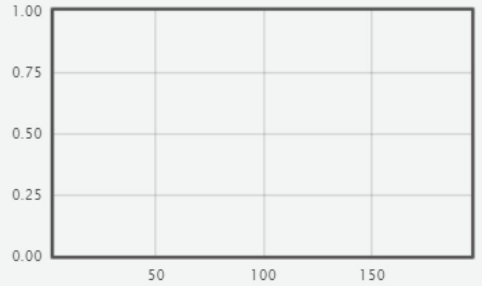
Population

n = 82, mean = 53.537
median = 45.5, stdev = 36.89



Sample

Show Data Table

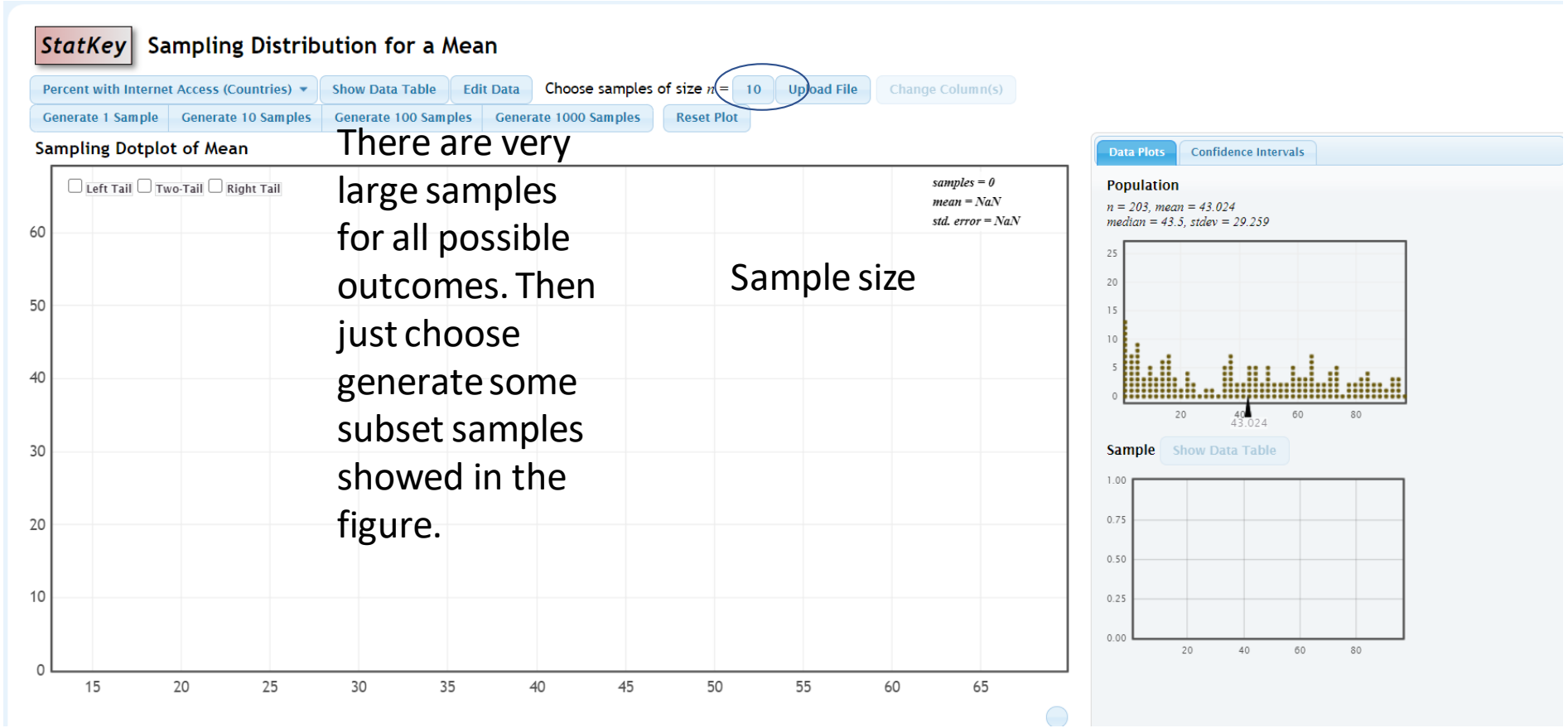


Help

StatKey v. 2.1.1 is written in JavaScript and should work well with any current browser including Chrome, Firefox, Safari, Opera, and IE.
Comments, feedback, and bug reports can be sent to **lock5stat@gmail.com.**

Presentation Mode OFF

- http://www.lock5stat.com/StatKey/sampling_1_quant/sampling_1_quant.html



Sampling Distribution of the Sample Proportion

- Before we begin, let's make sure we review the terms and notation associated with proportions:
- p is the population proportion. It is a fixed value.
- n is the size of the random sample.
- p^ is the sample proportion. It varies based on the sample.

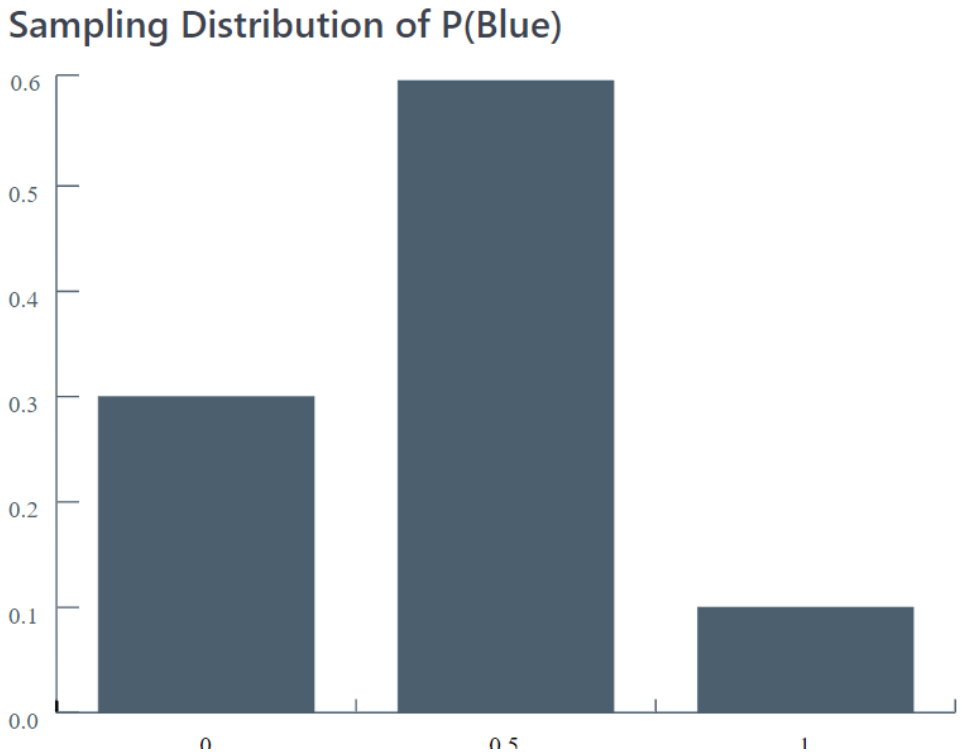
In a particular family, there are five children. Their names are Alex (A), Betina (B), Carly (C), Debbie (D), and Edward (E). The table below shows the child's name and their favorite color.

We are interested in the proportion of children in the family who prefer the color blue, and from the table, we can see that p=.40 of the children prefer blue.
Similar to the pumpkin example earlier in the lesson, let's say we didn't know the proportion of children who like blue as their favorite color. We'll use resampling methods to estimate the proportion. Let's take n=2 repeated samples, taken without replacement. Here are all the possible samples of size n=2 and their respective probabilities of the proportion of children who like blue

Name	Alex (A)	Betina (B)	Carly (C)	Debbie (D)	Edward (E)
Color	Green	Blue	Yellow	Purple	Blue

Sample	P(Blue)	Probability
AB	1/2	1/10
AC	0	1/10
AD	0	1/10
AE	1/2	1/10
BC	1/2	1/10
BD	1/2	1/10
BE	1	1/10
CD	0	1/10
CE	1/2	1/10
DE	1/2	1/10

The probability mass function (PMF) is:
The graph of the PMF:



P(Blue)	0	1/2	1
Probability	3/10	6/10	1/10

- The true proportion is p=P(Blue)=2/5. When the sample size is n=2, you can see from the PMF, it is not possible to get a sampling proportion that is equal to the true proportion.
- Although not presented in detail here, we could find the sampling distribution for a larger sample size, say n=4. The PMF for n=4 is...

P(Blue)	1/4	1/2
Probability	2/5	3/5

As with the sampling distribution of the sample mean, the sampling distribution of the sample proportion will have sampling error. It is also the case that the larger the sample size, the smaller the spread of the distribution.